



## Doctoral Thesis

### **Computational Magnetohydrodynamics with Discrete Differential Forms**

**Author(s):**

Pagliantini, Cecilia

**Publication Date:**

2016

**Permanent Link:**

<https://doi.org/10.3929/ethz-a-010722079> →

**Rights / License:**

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS. ETH NO. 23781

# Computational Magnetohydrodynamics with Discrete Differential Forms

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by  
CECILIA PAGLIANTINI

MSc. Mathematical Modelling in Engineering,  
Polytechnic University of Turin  
born on 20.02.1988  
citizen of Italy

accepted on the recommendation of  
Prof. Dr. Ralf Hiptmair, ETH Zurich, examiner  
Prof. Dr. Blanca Ayuso de Dios, Hamburg University of Technology, co-examiner  
Prof. Dr. Siddhartha Mishra, ETH Zurich, co-examiner

Pagliantini, Cecilia  
Computational Magnetohydrodynamics with Discrete Differential Forms  
Dissertation, ETH Zurich, 2016, No. 23781  
DOI: 10.3929/ethz-a-010722079  
ISBN: 978-3-906327-55-6

# Abstract

The equations of magnetohydrodynamics (MHD) model the interaction of conducting fluids with electromagnetic fields, and provide the mathematical description of problems arising in areas as diverse as plasma physics, astrophysics, and thermonuclear fusion. They comprise balance equations for mass, momentum and energy, the magneto-quasistatic Maxwell's equations for the electromagnetic fields, and material laws. This thesis is devoted to the development and analysis of novel numerical methods for the MHD problem based on Galerkin schemes for the electromagnetic fields via finite element exterior calculus (FEEC), coupled with finite volume (FV) schemes for the conservation laws of fluid mechanics.

Finite element exterior calculus relies on discrete differential forms which provide structure-preserving discretizations by supporting a discrete de Rham cohomology. The magneto-quasistatic model underlying resistive MHD yields a magnetic advection-diffusion problem for the magnetic potential. We consider the singular perturbation limit and devise robust numerical discretizations of generalized transient advection problems for differential forms, through an Eulerian method of lines with explicit time-stepping and stabilized Galerkin schemes. The spatial approximations include both conforming discrete differential forms and genuinely discontinuous finite elements, and are designed to accommodate discontinuous advection velocities which inevitably occur in MHD flows. A priori convergence estimates are established for Lipschitz continuous velocities, conforming meshes and polynomial finite element spaces of discrete differential forms. Additionally, we explore an alternative class of numerical schemes for the discretization of the Lie derivative, built on the duality between the contraction operator and the extrusion of manifolds. These methods incorporate an upwinding element and deliver discrete advection operators that commute with the exterior derivative, hence ensuring that closed forms are Lie advected into closed forms. Nonlinear flux limiters, in the form of residual-based artificial viscosity, are designed to curb the spurious oscillations resulting from higher order polynomial discretizations.

In the resistive MHD system, the eddy current model with non-vanishing diffusion gives rise to a parabolic problem in  $H(\mathbf{curl}, \Omega)$ . Discretizations with Galerkin schemes, and implicit-explicit (IMEX) Runge–Kutta time-stepping, entail solving, in each time step, discrete boundary value problems for the double **curl** operator. Spatial discretizations with discrete differential forms pave the way for applying fast iterative solvers, viz. multigrid. For discontinuous Galerkin discretizations, we develop a family of preconditioners based on an auxiliary space of  $H(\mathbf{curl}, \Omega)$ -conforming finite elements, together with a smoother. The resulting iterative solvers are shown to be asymptotically optimal in terms of independence from the mesh width. With particular regard to the case of locally dominant transport, robustness with respect to jumps in the zeroth- and second-order parts of the operator is shown to hold in almost all configurations, except when the problem changes from being **curl**-dominated to reaction-dominated.

The balance laws for the fluid variables can be considered as a system of conservation laws with the magnetic induction field as a space variable coefficient, supplied at every time step by one of the foregoing structure-preserving discretizations. The design of finite volume schemes for this extended Euler problem relies on approximate Riemann solvers, adapted to accommodate the electromagnetic contributions to the momentum and energy directly entering the fluxes. High order spatial accuracy is achieved via non-oscillatory reconstruction techniques, such as TVD limiters and (W)ENO-type reconstructions.

A full discretization of the MHD system results from coupling the FEEC-based numerical schemes for the magnetic advection-diffusion problem with finite volume approximations of the conservation laws for the fluid. The lowest order fully coupled scheme is tested on a set of benchmark tests for the two-dimensional planar ideal MHD equations. The method based on extrusion contraction upwind schemes for the magnetic advection preserves the divergence constraint exactly, and proves first order accurate for smooth solutions, conservative, and stable.

This research was partly supported by the Swiss NSF Grant No. 146355.



# Prefazione

Le equazioni della magnetoidrodinamica (MHD) costituiscono un modello per l'interazione di fluidi elettricamente conduttori e campi elettromagnetici, e forniscono la descrizione matematica di problemi di interesse in varie aree quali la fisica del plasma, l'astrofisica e la fusione termonucleare. Comprendono equazioni di conservazione della massa, della quantità di moto e dell'energia, le equazioni di Maxwell nel modello delle correnti parassite per i campi elettromagnetici, e le relazioni costitutive. Questa tesi verte sullo sviluppo e sull'analisi di nuovi metodi numerici per il problema MHD, basati sull'accoppiamento di schemi di tipo Galerkin, attraverso il *finite element exterior calculus* (FEEC), per i campi elettromagnetici e di metodi ai volumi finiti (FV) per le leggi di conservazione della fluidodinamica.

Il *finite element exterior calculus* si basa su forme differenziali discrete, le quali generano schemi numerici *structure-preserving* in quanto compatibili con una coomologia di de Rham discreta. Il modello magnetostatico alla base del problema MHD porta a un problema di avvezione-diffusione magnetica per il potenziale magnetico. Nel limite di perturbazione singolare, deriviamo discretizzazioni numeriche robuste per problemi di avvezione generalizzati a forme differenziali, tramite schemi in tempo esplicativi e schemi di tipo Galerkin con termini di stabilizzazione. Le approssimazioni in spazio includono forme differenziali discrete conformi e al contempo elementi finiti discontinui, e sono pensate per consentire velocità discontinue che inevitabilmente si presentano in flussi MHD. Dimostriamo stime di convergenza a priori per velocità lipschitziane, griglie conformi e spazi polinomiali di forme differenziali discrete. Inoltre, analizziamo una classe alternativa di schemi numerici per la discretizzazione della derivata di Lie fondati sulla dualità tra la contrazione di forme e l'*extrusion* di varietà. Questi metodi incorporano una componente di *upwinding* e danno luogo a operatori discreti di avvezione che commutano con la derivata esterna, assicurando quindi che forme differenziali chiuse vengono trasportate, nel senso della derivata di Lie, in forme chiuse. Al fine di mitigare le oscillazioni spurie associate a discretizzazioni di grado polinomiale alto, sviluppiamo *flux limiters* non lineari sotto forma di viscosità artificiale basata sul residuo dell'equazione.

Nel problema della magnetoidrodinamica resistiva, il modello magnetostatico con diffusione non nulla dà origine ad un problema parabolico in  $H(\mathbf{curl}, \Omega)$ . Approssimazioni di tipo Galerkin con stabilizzazione e metodi in tempo di tipo Runge–Kutta implicito-explicito (IMEX), richiedono la risoluzione, ad ogni passo di tempo, di problemi ai valori al bordo per l'operatore di doppio  $\mathbf{curl}$ . Discretizzazioni in spazio con forme differenziali discrete consentono l'applicazione di solutori iterativi efficienti, come per esempio metodi multigriglia. Per schemi di tipo Galerkin discontinuo, sviluppiamo una famiglia di precondizionatori basati su uno spazio ausiliario di elementi finiti in  $H(\mathbf{curl}, \Omega)$  ed uno *smoother*. Dimostriamo che i metodi iterativi derivanti sono asintoticamente ottimi in termini di indipendenza dal passo di griglia. Con particolare riguardo al caso di trasporto localmente dominante, mostriamo che i predetti solutori sono robusti rispetto a possibili discontinuità dei coefficienti dell'operatore di ordine zero e dell'operatore di secondo ordine. Questo vale in quasi tutte le configurazioni, fatta eccezione nel caso in cui il problema cambia da  $\mathbf{curl}$ -dominante a reazione-dominante.

Le leggi di conservazione della fluidodinamica possono essere considerate come un sistema di equazioni in cui il campo di induzione magnetica si comporta da coefficiente variabile in spazio ed è fornito, ad ogni passo di tempo, da una delle precedenti discretizzazioni *structure-preserving*. Costruiamo schemi ai volumi finiti per il problema di Eulero esteso fondati su solutori di Riemann approssimati, i quali tengono conto della presenza del contributo elettromagnetico nei flussi delle leggi di conservazione della quantità di moto e dell'energia. L'impiego di tecniche di ricostruzione non oscillatoria, come TVD *limiters* e riconstruzioni di tipo (W)ENO, consente accuratezze di ordine alto.

Deriviamo metodi numerici per l'intero sistema MHD dall'accoppiamento di schemi basati su FEEC per il problema di avvezione-diffusione magnetica, con approssimazioni ai volumi finiti per le equazioni di conservazione della fluidodinamica. Verifichiamo le prestazioni dello schema di ordine polinomiale minimo

così ottenuto tramite una serie di esperimenti standard per le equazioni della magnetoidrodinamica ideale, in dimensione due e in configurazione planare. Il metodo basato sugli schemi di *extrusion contraction upwind* per l'avvezione dei campi magnetici soddisfa il vincolo di divergenza in modo esatto, e si rivela conservativo, stabile e accurato al primo ordine nel caso di soluzioni lisce.

Questo lavoro di ricerca è stato parzialmente finanziato dal Swiss NSF Grant No. 146355.

# Contents

|   |     |
|---|-----|
| <b>Abstract</b>   | iii |
| <b>Introduction</b>   | ix  |
| <b>1. The Magnetohydrodynamic Model</b>   | 1   |
| 1.1. A Macroscopic Model of Plasma . . . . .  | 1   |
| 1.1.1. The Eigenvalue Structure of the Ideal MHD System . . . . .                         | 5   |
| 1.2. Literature on Numerical Methods for MHD . . . . .                                    | 8   |
| 1.2.1. Survey on Divergence Cleaning Techniques . . . . .                                 | 9   |
| 1.2.2. The Semi-Godunov Approach . . . . .  | 10  |
| <b>I. Maxwell's Equations in MHD</b>  | 13  |
| <b>2. Differential Forms in Electromagnetism</b>  | 15  |
| 2.1. Exterior Calculus . . . . .  | 15  |
| 2.2. Generalized Advection-Diffusion Evolution Problem . . . . .                          | 22  |
| 2.2.1. Structure-Preserving Discretizations . . . . .                                     | 23  |
| 2.3. Discrete Differential Forms . . . . .  | 23  |
| 2.4. Polynomial Differential Forms . . . . .  | 27  |
| 2.4.1. Finite Element Discrete Differential Forms . . . . .                               | 29  |
| 2.4.2. Finite Element Spaces for Vector Proxies of Discrete Differential Forms . . . . .  | 33  |
| 2.4.3. Interpolation Estimates . . . . .  | 33  |
| <b>3. Stabilized Galerkin for Transient Advection of Differential Forms</b>               | 37  |
| 3.1. Generalized Pure Advection Evolution Problem . . . . .                               | 37  |
| 3.1.1. Well-posedness . . . . .   | 38  |
| 3.2. Stationary Generalized Advection Problem . . . . .                                   | 39  |
| 3.2.1. Stabilized Galerkin Variational Formulation . . . . .                              | 40  |
| 3.3. Stationary Transport: Estimates for Continuous Velocity Fields . . . . .             | 44  |
| 3.3.1. Averaging Operator . . . . .   | 45  |
| 3.3.2. A Priori Convergence Estimates . . . . .   | 47  |
| 3.4. Fully Discrete Problem . . . . .   | 52  |
| 3.5. Numerical Experiments in 2D: Continuous Velocity . . . . .                           | 59  |
| 3.6. Numerical Experiments in 2D: Discontinuous Velocity . . . . .                        | 59  |
| 3.6.1. Stationary Advection of 2-Forms . . . . .  | 61  |
| 3.6.2. Stationary Advection of 1-Forms . . . . .  | 62  |
| 3.6.3. Transient Advection of 1-Forms . . . . .   | 64  |
| <b>4. Extrusion Contraction Upwind Schemes</b>  | 69  |
| 4.1. Contraction and Extrusion . . . . .  | 69  |
| 4.2. Upwinding the Contraction via Extrusion . . . . .                                    | 72  |
| 4.2.1. Commuting Property of the Discrete Lie Derivative . . . . .                        | 75  |
| 4.3. Numerical Experiments in 2D . . . . .  | 76  |
| 4.3.1. Transient Advection of 0-Forms . . . . .   | 76  |
| 4.3.2. Transient Advection of 1-Forms . . . . .   | 79  |
| 4.4. High Order Schemes . . . . .   | 81  |
| 4.4.1. Nonlinear Residual-Based Viscosity for Extrusion Contraction . . . . .             | 84  |
| 4.4.2. Nonlinear Residual-Based Viscosity for Stabilized Discontinuous Galerkin . . . . . | 87  |

|   |            |
|---|------------|
| <b>5. Preconditioning the Magnetic Advection-Diffusion Problem</b>  | <b>91</b>  |
| 5.1. Symmetric Interior Penalty Discontinuous Galerkin Discretization . . . . .                               | 94         |
| 5.2. Auxiliary Space Preconditioning . . . . .  | 98         |
| 5.2.1. Fictitious Space and Auxiliary Space Method . . . . .  | 98         |
| 5.2.2. Auxiliary Space Preconditioners for the IP-DG Discretization in $H_0(\mathbf{curl}, \Omega)$ . . . . . | 100        |
| 5.2.3. Smoothers for the Auxiliary Space Preconditioner . . . . .   | 101        |
| 5.3. Asymptotic Optimality of the Preconditioner . . . . .  | 105        |
| 5.3.1. $\nu$ -stable Decomposition . . . . .  | 106        |
| 5.3.2. $\beta$ -aware Averaging Operator . . . . .  | 110        |
| 5.3.3. Coarser Auxiliary Space . . . . .  | 117        |
| 5.4. Numerical Experiments in 2D . . . . .  | 120        |
| 5.4.1. Triangular Meshes. Constant Coefficients . . . . .   | 121        |
| 5.4.2. Structured Triangular Mesh. One Discontinuous Coefficient . . . . .                                    | 123        |
| 5.4.3. Structured Triangular Mesh. Two Discontinuous Coefficients . . . . .                                   | 125        |
| 5.4.4. Tensor Product Meshes . . . . .  | 127        |
| <b>II. The Hydrodynamic Equations and the Full Ideal MHD</b>  | <b>131</b> |
| <b>6. The Extended Euler Equations</b>  | <b>133</b> |
| 6.1. Derivation of the Extended Euler Equations . . . . .   | 133        |
| 6.2. Finite Volume Discretization . . . . .   | 135        |
| 6.2.1. Approximate Riemann Solvers . . . . .  | 138        |
| 6.2.2. Numerical Experiments in 2D . . . . .  | 142        |
| 6.3. High Order Schemes . . . . .   | 145        |
| 6.3.1. Piecewise Linear Reconstruction . . . . .  | 145        |
| 6.3.2. Numerical Experiments in 2D . . . . .  | 146        |
| <b>7. Numerical Experiments for Ideal MHD</b>   | <b>149</b> |
| 7.1. Planar Two-Dimensional Ideal MHD . . . . .   | 149        |
| 7.1.1. The FV-FEEC Schemes in 2D . . . . .  | 149        |
| 7.1.2. Accuracy Test: Smooth Vortex . . . . .   | 151        |
| 7.1.3. Super-Fast Expansion: Shock Tube Test . . . . .  | 153        |
| 7.1.4. Orszag–Tang Benchmark . . . . .  | 153        |
| 7.1.5. Rotor Problem . . . . .  | 157        |
| 7.1.6. Blast Wave Problem . . . . .   | 159        |
| 7.2. Assessment of FV-DG Schemes for MHD Coupling . . . . .   | 159        |
| 7.3. Towards Second Order Schemes . . . . .   | 163        |
| <b>8. Concluding Remarks</b>  | <b>165</b> |
| <b>Bibliography</b>   | <b>167</b> |

# Introduction

Plasma phenomena are ubiquitous in several physical applications ranging from controlled thermonuclear fusion [Che16] to astrophysics [Par91]. A plasma is an ionized<sup>1</sup> gas, consisting of freely moving charged particles. The majority of plasma phenomena are accurately described under the so-called plasma approximation, a macroscopic interpretation of plasmas as single species, non-relativistic conducting fluids subject to the influence of magnetic fields. A conducting fluid subject to a magnetic field experiences electric currents and the magnetic field associated with these currents affects the original field and the flow motion. The complex dynamics involved finds a theoretical description in the magnetohydrodynamic (MHD) model.

The MHD equations provide a mathematical interpretation of the interplay of conducting non-magnetic fluids with electromagnetic fields. For compressible fluids, the model comprises conservation laws for mass, momentum and energy together with material laws and Maxwell's equations, in the magneto-quasistatic reduction, to describe the evolution of the electromagnetic fields

$$\left\{ \begin{array}{l} \partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0, \\ \partial_t (\rho \mathbf{u}) + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u} + (p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbb{I} - \mathbf{B} \otimes \mathbf{B}) = \mathbf{0}, \\ \partial_t E + \operatorname{div}((E + p - \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbf{u} + \mathbf{E} \times \mathbf{B}) = 0, \\ \partial_t \mathbf{B} + \operatorname{curl} \mathbf{E} = \mathbf{0}, \\ \operatorname{div} \mathbf{B} = 0. \end{array} \right.$$

Here  $\rho$ ,  $p$  and  $\mathbf{u}$  are the fluid density, pressure and velocity, respectively; the total energy  $E$  is expressed through the equation of state; and  $\mathbf{E}$  and  $\mathbf{B}$  are the electric and magnetic induction fields, respectively. The electric field is described via the current induced by the Lorentz force,  $\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B})$ , and Ampère's law,  $\mathbf{J} = \operatorname{curl}(\mu^{-1} \mathbf{B})$ , where  $\mu$  is the magnetic permeability and  $\sigma$  the electric conductivity.

Although the analysis of the MHD problem at the continuous level is not yet mature, the relevance of the model in several applications has fostered research oriented towards the design of numerical approximation schemes.

This dissertation is concerned with the development, analysis and implementation of stable numerical approximations of the MHD equations, by combining finite volume methods for the fluid conservation laws with numerical schemes based on discrete differential forms for the evolution of the electromagnetic fields.

As described in Chapter 1, the MHD equations form a non-strictly hyperbolic system of conservation laws with non-convex flux functions. Since the characteristic fields are not either genuinely nonlinear or linearly degenerate, possible non-regular waves, like compound waves or overcompressive shocks, can develop, in addition to shocks and discontinuities. Furthermore, violating the divergence constraint at the discrete level might lead to the onset of unphysical plasma transport orthogonal to the magnetic field lines: this causes the loss of conservation of momentum and energy and might trigger numerical instabilities. In view of the complexity of the MHD system, devising numerical schemes which are simultaneously stable, accurate, efficient and physically reliable is rather challenging.

The most popular numerical schemes for the MHD problem, and generally for systems of conservation laws, are finite volume methods [LeV02], in which an integral version of the conservation law is solved inside each control volume of the computational domain. The literature on numerical methods for system of conservation laws has seen also an extensive development of discontinuous Galerkin schemes

---

<sup>1</sup>Collective plasma behaviors are encountered even when the ionization is only partial (*cf.* hot and cold plasmas).

[CS98]. The numerical treatment of the divergence constraint in the context of finite volume and discontinuous Galerkin methods, usually provides control only of some specific approximation of the divergence operator, and not of other approximations. Moreover, the resistive MHD model demands for a suitable discretization of the resistive double **curl** operator. Finite element methods are more amenable to handling these issues. However, the fluid dynamics community generally eschews conforming finite element schemes, owing to the difficulties related to the discretization of solutions with low regularity and displaying physical discontinuities. On the other hand, discrete differential forms in the guise of conforming finite elements, the so-called edge elements and generalizations, represent a well-established tool in the numerical discretization of problems in computational electromagnetism, as described in Chapter 2. The so-called Finite Element Exterior Calculus [AFW06] aims at finite element approximations with good approximation properties, which provide structure-preserving discretizations in the sense that they support a discrete de Rham cohomology. Hence, discrete potential becomes available and the divergence constraint holds exactly, even for the discrete induction field.

The MHD model can be phrased as a combination of the quasi-magnetostatic Maxwell equations with the conservation laws of gas dynamics, together with a mathematical description of the coupling interaction. This “splitting and coupling approach” suggests that a numerical discretization mimicking such strategy is not only reasonable, but could reap the benefits of numerical schemes tailored for different phenomena. In this dissertation we propose a splitting approach, inspired by [FMR09], through the combination of numerical discretizations of the electromagnetic part, founded on finite element exterior calculus, as derived in Part I, and finite volume methods for the fluid conservation laws, as derived in Part II.

The magneto-quasistatic model underlying resistive MHD yields a magnetic advection-diffusion problem for the unknown magnetic induction field. Introducing the magnetic potential, the divergence constraint on the induction field is automatically absorbed into a new advection-diffusion problem, satisfied by the potential. A generalized advection-diffusion problem for differential forms was derived in [Heu11], via the correspondences between differential forms and scalar/vector fields: it includes the aforementioned equations and the well known scalar case as particular instances. Part I is devoted to the analysis of efficient and stable numerical schemes for the generalized transient advection-diffusion problem. In Chapter 3, we derive and analyze robust numerical methods based on an Eulerian method of lines with explicit time-stepping and stabilized Galerkin spatial discretizations which accommodate both  $H\Lambda^k(\Omega)$ -conforming discrete differential forms and (truly) discontinuous finite elements. In the singular perturbation limit, we focus on the generalized pure advection problem, with particular emphasis on the ability of such schemes to cope with discontinuous transport velocity fields which inevitably result from the MHD equations. Rigorous a priori convergence theory is established for Lipschitz continuous velocities, conforming meshes and standard finite element spaces of discrete differential forms, in Chapter 3.

Concurrently with the FEEC-based stabilized Galerkin methods, we investigate, in Chapter 4, another family of discretizations founded on discrete differential forms, the so-called extrusion contraction upwind schemes, introduced in [HH08]. The extrusion contraction upwind methods deliver discrete advection operators commuting with the exterior derivative, hence, implying that closed forms are Lie advected into closed forms. Despite the numerical evidence of stability and robustness of the aforementioned schemes, even in the presence of shocks and discontinuities, rigorous convergence and stability results are still an open problem. The only exception hitherto, is the lowest order extrusion contraction upwind scheme with nodal quadrature which reduces to the scheme of Tabata [Tab77]: the resulting Galerkin matrix enjoys the M-matrix property, ensuring that the discrete evolution operator admits a monotone inverse. Moreover, for the purpose of designing high order numerical schemes satisfying some discrete monotonicity property, the entropy viscosity methods of [GP08] is adapted to high order extrusion contraction upwind approximations. Specifically, the commutativity of the exterior derivative and the discrete Lie derivative allows the construction of nonlinear residual-based viscosity schemes for the advection of the magnetic potential, based, however, on the residual of the magnetic induction equation.

The numerical discretizations, proposed in Chapters 3 and 4, deliver robust approximations also when augmented with a discretization of the diffusion operator. The eddy current model with non-vanishing magnetic diffusion, characterizing resistive MHD flows, gives rise to a parabolic problem in  $H(\mathbf{curl}, \Omega)$ . An implicit-explicit (IMEX) strategy for the temporal discretization seems a natural choice. Efficient and robust iterative solvers are needed to efficiently solve the large sparse linear systems of equations resulting from the discretization, in each time step, of boundary value problems for the double **curl** operator.

In the context of  $H(\mathbf{curl}, \Omega)$ -conforming Galerkin finite element discretizations, subspace correction preconditioners are well-established in the form of multigrid [Hip99a], and domain decomposition methods [HT00]. These fast solvers can also be exploited in the construction of efficient preconditioners for genuinely discontinuous Galerkin discretizations of the magnetic diffusion problem. Specifically, for the symmetric interior penalty discontinuous Galerkin discretization of  $H(\mathbf{curl}, \Omega)$ -elliptic problems on conforming meshes, we develop, in Chapter 5, a family of preconditioners based on the Auxiliary Space Method of [Nep91; Xu96], with an auxiliary space of  $H(\mathbf{curl}, \Omega)$ -conforming finite element functions, together with a relaxation technique (local smoothing). In Theorem 5.3.1, we state that the proposed preconditioners enjoy asymptotic optimality with respect to mesh refinement. Since resistive effects in MHD are local phenomena, the investigation of the asymptotic behavior of the preconditioners is particularly relevant for (locally) dominant transport, namely in the presence of discontinuous coefficients in the second- and zeroth-order parts of the operator. The proposed solvers prove robust with respect to jumps in the coefficients, except when the problem changes from a “ $\mathbf{curl}$ -dominated” regime to a “reaction-dominated” regime, and vice versa.

Assuming that the magnetic field is approximated by one of the numerical methods developed in Part I, the balance laws for mass, momentum and energy can be written as a system of conservation laws, named extended Euler system [FMR09], with the magnetic induction field as a (discontinuous) varying coefficient. On account of the tight coupling of fluid and electromagnetic fields, the electromagnetic contribution to the conservation of momentum and energy directly enters the fluxes. The numerical treatment of the fluid balance laws rests upon finite volume schemes, with numerical fluxes suitably designed, in Chapter 6, to accommodate the discontinuous magnetic induction field. High order approximations rely on piecewise polynomial reconstructions and slope limiters.

Coupling the FEEC-based numerical schemes, developed in Part I, with the finite volume methods devised for the conservation laws for the fluid, in Chapter 6, yields a full discretization of the ideal MHD system. The resulting new algorithm is described and tested in Chapter 7. We implement the two-dimensional planar transient ideal MHD problem on structured meshes with lowest order spatial discretizations, and numerically assess the scheme on a set of benchmark problems. The lowest order method for the full ideal MHD problem, based on extrusion contraction upwind schemes, proves numerically first order accurate (for smooth solutions), stable and robust. The corresponding method associated with fully discontinuous Galerkin discretizations of the advection problem for the magnetic potential, yields unphysical solutions on some challenging benchmark tests. Insights regarding the construction of second order schemes for the full ideal MHD hinging on the splitting approach are given at the end of Chapter 7.



# 1. The Magnetohydrodynamic Model

Plasma interacting with magnetic fields characterizes a range of physical phenomena from controlled thermonuclear fusion experiments to astrophysics. The magnetohydrodynamic approximation provides the mathematical model to describe the macroscopic interplay between plasmas and electromagnetic fields. We first attempt to identify, in Section 1.1, the conditions on the plasma and plasma dynamics under which the magnetohydrodynamic model is valid, largely gleaned from the monographs [GP04], [Che16] and [Par91, Chapter 5]. In Section 1.1.1, we derive the wave structure of the ideal MHD system, which is of crucial importance in designing Riemann solver-based numerical schemes for the fluid equations (*cf.* Chapter 6). A survey on numerical schemes developed to handle the MHD problem in an accurate, robust and efficient way is presented in Section 1.2. Particular emphasis is placed on the numerical strategies to enforce some discrete version of the divergence constraint on the magnetic induction field.

## 1.1. A Macroscopic Model of Plasma

“A plasma is a quasi-neutral gas of charged and neutral particles which exhibits collective behavior” [Che16, p. 2]. In a plasma, the charged particles are subject to long-range Coulomb interactions with many distant encounters, due to the movement of the charged particles in the average electrostatic field created by all other particles, which results in an increasing interacting volume. Collective plasma behavior is encountered under three main conditions: the long-range Coulomb interactions between charged particles dominate over the short-range binary collisions with neutrals; the *quasi-neutrality* assumption, namely the length scale of plasma dynamics is much larger than the *Debye length*, the characteristic size of regions where charge imbalance associated with thermal fluctuations might occur (and electric fields generated); sufficiently many particles are contained in a sphere having the radius equal to the Debye length, so that the Debye shielding is a statistically valid concept. These assumptions can be rephrased in terms of local conditions on the density and temperature. However, the presence of magnetic fields entail a global macroscopic perspective. A macroscopic description of plasmas is valid under the hypothesis of frequent enough collisions among charged particles to establish fluid behavior (the *collisionality assumption*), and large enough length scale (much larger than the ion gyro-radius) and time scale (much larger than the inverse of the ion gyro-frequency) involving the magnetic field. Under the latter assumption, averaging over the gyration of the particles allows to consider only the motion of the guiding center of the cyclotron gyration (the so-called guiding center approximation). In this low-frequency motion, a plasma manifests an apparent charge neutrality, so that the electric field dynamics is dictated by the interaction of electrons and ions towards preservation of neutrality. This behavior justifies the so-called *plasma approximation*, in which the plasma is considered as a single species particle ensemble subject to the influence of magnetic fields. The plasma approximation is a mathematical ploy to enforce the quasi-neutrality condition.

In light of these two characterizations, microscopic and macroscopic, of the plasma state, the description of plasma processes can be seized by different theoretical models. In the presence of low-density plasma, collective effects can be neglected and single-particle motions in given magnetic and electric fields provide an accurate description of the plasma dynamics. Otherwise, two main models are used: the statistical approach of the *kinetic plasma theory*, describing plasma microscopically as a collection of particles and where the dynamics is analyzed in terms of time-dependent distribution functions; and the fluid theory (magnetohydrodynamics), describing plasma in terms of averaged macroscopic functions.

### Derivation of the MHD Equations

Magnetohydrodynamics can be thought of as the branch of fluid dynamics dealing with the motion of electrically conducting materials in the presence of electromagnetic fields. In this perspective, despite plasma phenomena involve charged particles, some of the physical behaviors of ordinary fluids are applicable to the plasma state, when approximated as a conducting fluid. Furthermore, electric and magnetic forces are associated with different effects taking place at different length scales, with the magnetic forces dominant on the longer length scales. Consequently, the bulk of the plasma dynamics involving the electromagnetic fields is well-described by the so-called *pre-Maxwell equations*, namely Maxwell's equations with neglected displacement. As mentioned above, this reduced model, also known as quasi-magnetostatic approximation or eddy-current model, corresponds to the fact that MHD theory deals with low-frequency phenomena.

In light of the above considerations, it comes about natural to formulate the MHD models by combining the pre-Maxwell equations with conservation laws of gas dynamics, concurrently providing a mathematical description of the coupling interaction. Maxwell's equations are the fundamental laws of electromagnetism and describe the evolution and interaction of temporal and spatially varying electromagnetic fields. The quasi-magnetostatic approximation model reads

$$\mathbf{curl} \mathbf{E} = -\partial_t \mathbf{B}, \quad (\text{Faraday's law}) \quad (1.1)$$

$$\mathbf{curl} \mathbf{H} = \mathbf{J}, \quad (\text{Ampère's law}) \quad (1.2)$$

$$\operatorname{div} \mathbf{B} = 0, \quad (1.3)$$

where  $\mathbf{J}$  denotes the current density,  $\mathbf{E}$  the electric field, and  $\mathbf{B}$  and  $\mathbf{H}$  are the induction and magnetic fields, respectively. The model has to be supplemented with initial conditions, decay conditions for the fields at infinity, and with the material law coupling induction and magnetic fields through the magnetic permeability tensor  $\mu$ ,

$$\mathbf{B} = \mu \mathbf{H}. \quad (1.4)$$

Roughly speaking, Faraday's law (1.1) describes the evolution of the magnetic induction field with respect to the spatial variations of the electric field. Ampère's law (neglecting the Maxwell correction associated with the electric displacement field) relates a flowing electric current to the magnetic field winding around the current. Lastly, (1.3) dictates the sole admissibility of solenoidal magnetic induction fields and the absence of magnetic monopoles. Magnetic monopoles are hypothetical elementary particles carrying a "magnetic charge" (isolated magnets) which would provide the magnetic analog of charged particles, as magnetic dipoles for electric dipoles. They were hypothesized by Dirac [Dir31] in order to explain the quantization of electric charge. However, they have never been experimentally observed nor found in nature. This fact is expressed by the Gauss law for magnetism:  $\operatorname{div} \mathbf{B} = 0$ . Faraday's law guarantees, in addition, that initially solenoidal magnetic induction fields satisfy the divergence constraint at all time.

Concerning the description of the moving plasma treated as a fluid, the density  $\rho = \rho(\mathbf{x}, t)$ , the pressure  $p = p(\mathbf{x}, t)$ , and the velocity  $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$  characterize, as in gas dynamics, the fluid at every point  $(\mathbf{x}, t)$  in space and time. Evolution equations for the above-stated variables describe the dynamical behavior of the fluid. In greater detail, in MHD flows as in the case of neutral fluid, the mass of a fluid element is conserved: this is expressed through the continuity equation

$$\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0. \quad (1.5)$$

The interaction between fluid and electromagnetic variables comes into play in the conservation law for the momentum  $\rho \mathbf{u}$ . Indeed, in non-relativistic plasmas with isotropic plasma pressure and neglecting gravitational effects, Newton's equation of motion describes the acceleration of a fluid element associated with the force exerted by the combination of pressure gradient and electromagnetic terms as

$$\rho(\partial_t \mathbf{u} + \mathbf{u} \cdot \mathbf{grad} \mathbf{u}) = -\mathbf{grad} p + \mathbf{J} \times \mathbf{B}. \quad (1.6)$$

Here the operator  $\mathbf{grad} : \mathbb{R}^3 \rightarrow \mathbb{R}^{3,3}$  denotes the gradient of a vector and  $\mathbf{grad} \mathbf{u}$  is nothing but the Jacobian matrix of  $\mathbf{u}$ . By Ampère's law (1.2) and the material law (1.4),  $\mathbf{J} = \mathbf{curl}(\mu^{-1} \mathbf{B})$ . Hence, setting

$\mu = 1$  and recasting (1.6) in divergence form (see also Section 6.1) yields the conservation law for the momentum,

$$\partial_t(\rho\mathbf{u}) + \mathbf{div} \left( \rho\mathbf{u} \otimes \mathbf{u} + \left( p + \frac{1}{2}\|\mathbf{B}\|_{\ell^2}^2 \right) \mathbb{I} - \mathbf{B} \otimes \mathbf{B} \right) = \mathbf{0}, \quad (1.7)$$

where  $\mathbb{I} \in \mathbb{R}^{3,3}$  is the identity matrix,  $\mathbf{div} : \mathbb{R}^{3,3} \rightarrow \mathbb{R}^3$  is the row-wise divergence and  $\mathbf{a} \otimes \mathbf{b}$  denotes the Kronecker (or dyadic) product of the column vector  $\mathbf{a} \in \mathbb{R}^3$  with the row vector  $\mathbf{b} \in \mathbb{R}^3$ . Finally, the fluid density and pressure are related via a thermodynamic equation, called *equation of state*. Let  $\gamma$  denote the ratio of the specific heat at constant pressure  $C_p$  and the specific heat at constant volume  $C_v$ . The fluid is usually assumed to be either adiabatic,  $\partial_t(p\rho^{-\gamma}) + \mathbf{u} \cdot \mathbf{grad}(p\rho^{-\gamma}) = 0$ , or isothermal,  $\partial_t(p\rho^{-1}) + \mathbf{u} \cdot \mathbf{grad}(p\rho^{-1}) = 0$ . The evolution of the fluid pressure  $p$  is described through the conservation law for the specific entropy  $S$ , with  $S = p\rho^{-\gamma}$  in adiabatic fluids, as

$$\partial_t p + \mathbf{u} \cdot \mathbf{grad} p + \gamma p \mathbf{div} \mathbf{u} = 0.$$

Coupling the above-stated entropy conservation equation with the balance law for the momentum (1.7) and Faraday's law (1.1), yields a conservation equation for the total energy density (see e.g. [GP04, Section 4.3.1] and Section 6.1]),

$$\partial_t E + \mathbf{div} \left( \left( E + p - \frac{1}{2}\|\mathbf{B}\|_{\ell^2}^2 \right) \mathbf{u} + \mathbf{E} \times \mathbf{B} \right) = 0. \quad (1.8)$$

The total energy  $E$  is given by the sum of the thermal energy, the kinetic energy density of the fluid motion and the total energy density of the magnetic field,

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho\|\mathbf{u}\|_{\ell^2}^2 + \frac{1}{2}\|\mathbf{B}\|_{\ell^2}^2.$$

The term  $\mathbf{E} \times \mathbf{B}$ , in the conservation of energy (1.8), is the *Poynting vector* and it represents the flow of electromagnetic energy.

Collecting the equations governing the evolution of the fluid variables subject to the influence of electromagnetic fields (1.5), (1.7), (1.8) and the pre-Maxwell equations (1.1), (1.3), results in the MHD system

$$\begin{cases} \partial_t \rho + \mathbf{div}(\rho\mathbf{u}) = 0, \\ \partial_t(\rho\mathbf{u}) + \mathbf{div} \left( \rho\mathbf{u} \otimes \mathbf{u} + \left( p + \frac{1}{2}\|\mathbf{B}\|_{\ell^2}^2 \right) \mathbb{I} - \mathbf{B} \otimes \mathbf{B} \right) = \mathbf{0}, \\ \partial_t E + \mathbf{div} \left( \left( E + p - \frac{1}{2}\|\mathbf{B}\|_{\ell^2}^2 \right) \mathbf{u} + \mathbf{E} \times \mathbf{B} \right) = 0, \\ \partial_t \mathbf{B} + \mathbf{curl} \mathbf{E} = \mathbf{0}, \\ \mathbf{div} \mathbf{B} = 0. \end{cases} \quad (1.9)$$

The system has to be closed by specifying one independent electromagnetic variable, as follows. Electric currents can flow in an MHD fluid owing to its conducting properties. The current density  $\mathbf{J}$  is related to the electric field  $\mathbf{E}$  via Ohm's law

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}),$$

which includes the contribution of the electric field induced by the motion of the fluid across magnetic fields, and where  $\sigma$  is the symmetric positive semi-definite tensor, representing the fluid electric conductivity. The electric field can then be expressed in the terms of the magnetic induction field  $\mathbf{B}$  via Ohm's and Ampère's laws. Considering the electric field as a secondary quantity in MHD models has a physical interpretation: large electrostatic fields are associated with charge imbalances which occur on length scales of the Debye radius and are therefore averaged out in the plasma approximation.

## Ideal and Resistive MHD Models

In a perfectly conducting plasma, the electric field  $\mathbf{E}'$  in the frame moving with the plasma vanishes, namely  $\mathbf{E}' := \mathbf{E} + \mathbf{u} \times \mathbf{B} = \mathbf{0}$ , and the magnetic field lines are *frozen* into the fluid (*cf.* Alfvén's theorem,

e.g. [Bis93, Section 2.2]). This property can be shown by considering the magnetic flux through a surface element  $\mathbf{B} \cdot dS$ , and combining Faraday's and Ohm's equations to obtain that the material derivative of the magnetic flux vanishes. Hence, the magnetic flux through any surface moving with the fluid is constant. Considering the plasma as a perfect conductor results in the so-called *ideal* MHD model, and corresponds to the assumption that the plasma dynamics takes place on time scales faster than the magnetic field decay by Joule or resistive heating (proportional to  $\mathbf{J} \cdot \mathbf{E}$ ). On the other hand, the relative importance of resistive effects over length scales of order  $L$  can be gauged by the *magnetic Reynolds number*  $R_m := \mu_0 \sigma L U$  where  $U$  is a characteristic velocity scale of the MHD flow and  $\mu_0$  is the permeability of free space. Resistive MHD models capture the changes of the magnetic field topology during the flow, the reconnection effects. Large values of  $R_m$  correspond to transport times much shorter than the diffusion time for the magnetic fields. Despite the ideal MHD model is a purely mathematical abstraction, the good electric conductivity properties of hot and strongly magnetized plasmas justify the ideal MHD equations as a good approximation of the dynamical phenomena of such plasmas. Additionally, robust numerical discretizations of ideal MHD models pave the way to reliable approximations of the resistive MHD problem, even in the presence of locally small conductivity. In view of the discussion about resistive reconnection phenomena, the topological configuration of magnetic field lines is particularly relevant. In ideal MHD, an integral invariant is related to the magnetic topology, the magnetic helicity, defined on a material volume  $V$  as

$$h(V) := \int_V \mathbf{A} \cdot \mathbf{B} dV.$$

In view of Alfvén's theorem and of the divergence constraint, the magnetic helicity of any flux tube (cylindrical region containing magnetic field lines parallel to tube surfaces) is conserved. This entails that flux tubes in an ideal conductor preserve their relative topology and their individual fluxes.

### Advection of the Magnetic Potential

In the framework of the eddy-current model, the magnetic advection-diffusion problem for the magnetic induction field  $\mathbf{B}$  can be equivalently reformulated in terms of the advection of the magnetic vector potential, in the so-called  $\mathbf{A}$ -based formulation. More precisely, on domains with trivial topology, the solenoidal constraint  $\operatorname{div} \mathbf{B} = 0$  implies that the magnetic induction field  $\mathbf{B}$  can be written as  $\mathbf{B} = \operatorname{curl} \mathbf{A}$ . This well known result is rigorously justified in Section 2.2. Clearly, the magnetic potential is uniquely defined up to a  $\operatorname{curl}$ -free function of the form  $\operatorname{grad} \varphi$ , the so-called *gauge*. Gauge fixing entails eliminating this further degree of freedom by imposing an additional constraint on the magnetic potential field.

Using Faraday's law (1.1) and rewriting the magnetic induction in terms of the magnetic vector potential results in

$$\operatorname{curl} \mathbf{E}' = -\partial_t \mathbf{B} - \operatorname{curl}(\mathbf{B} \times \mathbf{u}) = \operatorname{curl}(-\partial_t \mathbf{A} - \operatorname{curl} \mathbf{A} \times \mathbf{u}) - \operatorname{curl}(\operatorname{grad}(\mathbf{u} \cdot \mathbf{A})),$$

where the last term has vanishing contribution. Therefore,  $\mathbf{E}' = -\partial_t \mathbf{A} - \operatorname{curl} \mathbf{A} \times \mathbf{u} - \operatorname{grad}(\mathbf{u} \cdot \mathbf{A}) - \operatorname{grad} \varphi$ , where  $\operatorname{grad} \varphi$  represents the gauge freedom. We may envision the *Weyl gauge* (or temporal gauge) and set  $\varphi = 0$ . Furthermore, applying the material law (1.4) to Ampère's law (1.2) yields

$$\operatorname{curl} \mathbf{H} = \operatorname{curl}(\mu^{-1} \mathbf{B}) = \operatorname{curl}(\mu^{-1} \operatorname{curl} \mathbf{A}) = \sigma \mathbf{E}'.$$

Hence, the magnetic potential satisfies the transient advection-diffusion equation

$$\operatorname{curl}(\mu^{-1} \operatorname{curl} \mathbf{A}) + \sigma(\partial_t \mathbf{A} + \operatorname{curl} \mathbf{A} \times \mathbf{u} + \operatorname{grad}(\mathbf{u} \cdot \mathbf{A})) = \mathbf{0}.$$

Notice that under the assumptions made throughout the derivation, the magnetic vector potential is unique provided that  $\sigma > 0$  and suitable boundary and initial conditions are imposed. Further details on the magnetic advection equation and its numerical discretization are articulated in Chapter 3.

**Remark 1.1.1.** In Chapter 2 we shall present an interpretation of the electromagnetic fields in terms of differential forms able to capture the different nature of primal fields  $\mathbf{B}$ ,  $\mathbf{E}$  and dual field  $\mathbf{H}$ . In light of these results the derivation and algebraic manipulations carried out in the present section might appear rather crude. However, when looking at the realizations of differential forms in the Euclidean space, the isomorphism between the “space of fields” and the “space of fluxes” legitimates the derivation.

### 1.1.1. The Eigenvalue Structure of the Ideal MHD System

In  $n = 3$  dimensions, the ideal MHD system is a set of  $m = 8$  equations

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0, \\ \partial_t (\rho \mathbf{u}) + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u} + (p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbb{I} - \mathbf{B} \otimes \mathbf{B}) = \mathbf{0}, \\ \partial_t E + \operatorname{div}((E + p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B}) \mathbf{B}) = 0, \\ \partial_t \mathbf{B} + \operatorname{curl}(\mathbf{B} \times \mathbf{u}) = \mathbf{0}, \end{cases} \quad (1.10)$$

with  $\mathbf{B}$  satisfying the divergence constraint  $\operatorname{div} \mathbf{B} = 0$ .

Let  $G$  be an open subset of  $\mathbb{R}^m$ , called the set of states, and let  $\{\mathbf{f}^\ell\}_{\ell=1}^n : G \rightarrow \mathbb{R}^m$  be smooth functions. The Cauchy problem for systems of conservation laws, such as the ideal MHD equations (1.10), reads: Find  $\mathbf{U}(\mathbf{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow G$  such that

$$\begin{aligned} \partial_t \mathbf{U} + \operatorname{div}(\mathbf{F}(\mathbf{U})) &= \mathbf{0} && \text{in } \mathbb{R}^n \times \mathbb{R}_+, \\ \mathbf{U}(\mathbf{x}, 0) &= \mathbf{U}_0(\mathbf{x}) && \text{in } \mathbb{R}^n, \end{aligned} \quad (1.11)$$

where  $\mathbf{U} = (u_1, \dots, u_m)$  are called *conserved variables* and  $\mathbf{F} = (\mathbf{f}^1, \dots, \mathbf{f}^n) : G \rightarrow \mathbb{R}^{m \times n}$  is the *flux function*. The Jacobian matrix of  $\mathbf{f}^\ell = (f_1^\ell, \dots, f_m^\ell)$ ,  $1 \leq \ell \leq n$ , is defined as  $A_\ell(\mathbf{U}) = (\partial_{u_i} f_i^\ell)_{1 \leq i, j \leq m}$ . The system is called *hyperbolic* if, for any  $\mathbf{U} \in G$  and any direction  $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}^n$  with  $\|\mathbf{w}\|_{\ell^2} = 1$ , the matrix obtained as the convex combination  $\mathbf{F}'(\mathbf{U}) \cdot \mathbf{w} := \sum_{\ell=1}^n w_\ell A_\ell(\mathbf{U})$  has  $m$  real eigenvalues  $\lambda^1(\mathbf{U}, \mathbf{w}) \leq \dots \leq \lambda^m(\mathbf{U}, \mathbf{w})$  and a complete set of linearly independent eigenvectors  $r^1(\mathbf{U}, \mathbf{w}), \dots, r^m(\mathbf{U}, \mathbf{w})$ . If the eigenvalues are all distinct  $\lambda^1(\mathbf{U}, \mathbf{w}) < \dots < \lambda^m(\mathbf{U}, \mathbf{w})$ , the system is *strictly hyperbolic*. If  $\mathbf{F}'(\mathbf{U}) \cdot \mathbf{w}$  has real eigenvalues but is not diagonalizable, then the system is *weakly hyperbolic*. Let  $\partial_{\mathbf{U}} \lambda^p(\mathbf{U}) := (\partial_{u_i} \lambda^p(\mathbf{U}))_{i=1}^m$ , for  $1 \leq p \leq m$ , where the dependence on  $\mathbf{w}$  has been omitted. The  $p$ -th characteristic field  $(\lambda^p, r^p)$  is said to be

- *genuinely nonlinear*, if  $\partial_{\mathbf{U}} \lambda^p(\mathbf{U}) \cdot r^p(\mathbf{U}) \neq 0$  for all  $\mathbf{U} \in G$ ;
- *linearly degenerate*, if  $\partial_{\mathbf{U}} \lambda^p(\mathbf{U}) \cdot r^p(\mathbf{U}) = 0$  for all  $\mathbf{U} \in G$ .

If  $\partial_{\mathbf{U}} \lambda^p(\mathbf{U}) \cdot r^p(\mathbf{U}) = 0$  only for some  $\mathbf{U} \in G$ , then the  $p$ -th characteristic field is said to be *non-convex nonlinear*. The corresponding waves are termed *compound shocks* and are characterized by shocks with attached rarefaction.

Waves generated in a fluid medium, in which the MHD description is valid, are called MHD waves. They are low-frequency waves (in the sense that the frequency is smaller than the ion gyro-frequency) and they can possess both longitudinal and transverse mode. The hyperbolic nature of the MHD system allows the diagonalization of the Jacobian matrix  $A_\ell(\mathbf{U})$  in each space direction  $1 \leq \ell \leq n$ . Since the derivation of the Jacobian matrix for the MHD fluxes requires lengthy algebraic manipulations, we omit it here and refer the interested reader to [BW88, Section IV]. In three dimensions, there are only seven non-trivial equations (1.10) in each direction, and the flux Jacobian has eigenvalues

$$\begin{aligned} \lambda^1 &= \mathbf{u} \cdot \mathbf{n} - c_f, & \lambda^2 &= \mathbf{u} \cdot \mathbf{n} - c_a, & \lambda^3 &= \mathbf{u} \cdot \mathbf{n} - c_s, \\ \lambda^4 &= \mathbf{u} \cdot \mathbf{n}, \\ \lambda^5 &= \mathbf{u} \cdot \mathbf{n} + c_s, & \lambda^6 &= \mathbf{u} \cdot \mathbf{n} + c_a, & \lambda^7 &= \mathbf{u} \cdot \mathbf{n} + c_f, \end{aligned}$$

where  $\mathbf{n}$  is the unit normal vector,  $a := \sqrt{\gamma p / \rho}$  is the *sound speed*,  $c_a := \sqrt{(\mathbf{B} \cdot \mathbf{n})^2 / \rho}$  is the *Alfvén speed* and the slow and fast magnetoacoustic wave speeds are given by

$$\begin{aligned} c_s &= \frac{1}{\sqrt{2}} \sqrt{a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho} - \sqrt{\left(a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho}\right)^2 - 4a^2 \frac{(\mathbf{B} \cdot \mathbf{n})^2}{\rho}}}, \\ c_f &= \frac{1}{\sqrt{2}} \sqrt{a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho} + \sqrt{\left(a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho}\right)^2 - 4a^2 \frac{(\mathbf{B} \cdot \mathbf{n})^2}{\rho}}}. \end{aligned} \quad (1.12)$$

Note that, since  $c_s \leq c_a \leq c_f$ , the system is only non-strictly hyperbolic, for the eigenvalues may coincide at some points in the flow, named *umbilic* points. We delve deeper into the possible umbilic degeneracies in the forthcoming sections. The characteristic fields dictate the (finite) speed of propagation of the seven waves constituting the Riemann solution of the MHD system (in the so-called *regular case*). Specifically,

- $W^{1,7}, \mathbf{u} \cdot \mathbf{n} \mp c_f$  : fast magnetosonic waves, rarefaction/shock to the left/right,
- $W^{2,6}, \mathbf{u} \cdot \mathbf{n} \mp c_a$  : shear Alfvén waves, rotational discontinuity to the left/right,
- $W^{3,5}, \mathbf{u} \cdot \mathbf{n} \mp c_s$  : slow magnetosonic waves, rarefaction/shock to the left/right,
- $W^4, \mathbf{u} \cdot \mathbf{n}$  : entropy wave, contact discontinuity.

Alfvén waves are shear waves peculiar of electromagnetic phenomena. In a homogeneous background plasma, they are related to oscillations caused by the tension of the magnetic field lines trying to restore the background state: the restoring force for the perturbation is the Lorentz force exerting a magnetic tension along the field lines. More interestingly, they propagate in the direction of the magnetic induction field and, since the frequency of the MHD waves depends on the wavelength (see e.g. [GP04, Section 5.2]), the Alfvén waves encode information on the magnetic field topology with no fluctuations of density and pressure. Propagation perpendicular to the magnetic induction field is realized by longitudinal oscillations associated with pressure changes (the magnetic pressure is the restoring force), and represented by magnetosonic (compressional) waves. Lastly, the entropy wave is a longitudinal wave propagating at the fluid velocity, and characterized by variations of the entropy across the wave front, see [JT64, Section 4.3] for further details.

### Well-posedness

The presence of discontinuities and the consequent low regularity of the solutions of (systems of) conservation laws hint to an interpretation of solutions in the sense of distributions. A vector-valued function  $\mathbf{U} \in L^1_{\text{loc}}(\mathbb{R}^n \times \mathbb{R}^+)^m$  is a weak solution of (1.11), if for  $\mathbf{U}(\mathbf{x}, 0) \in L^1_{\text{loc}}(\mathbb{R}^n)^m$  and for all  $\varphi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}_0^+)^m$ , there holds

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}^n} \left( \mathbf{U}(\mathbf{x}, t) \cdot \partial_t \varphi(\mathbf{x}, t) + \sum_{\ell=1}^n \mathbf{f}^\ell(\mathbf{U}) \cdot \partial_{x_\ell} \varphi \right) d\mathbf{x} dt + \int_{\mathbb{R}^n} \mathbf{U}(\mathbf{x}, 0) \cdot \varphi(\mathbf{x}, 0) d\mathbf{x} = 0. \quad (1.13)$$

For piecewise differentiable initial data  $\mathbf{U}_0 \in C_{\text{pw}}^1(\mathbb{R}^n)^m$ , let  $\mathbf{U} \in C_{\text{pw}}^1(\mathbb{R}^n \times \mathbb{R}_+)^m$  with discontinuities across smooth  $n$ -dimensional submanifolds  $\mathcal{M}$  of  $\mathbb{R}^m \times \mathbb{R}_+$  with normal  $(\mathbf{n}, n_t)$ . Then  $\mathbf{U}$  is a weak solution of (1.11) if and only if:  $\mathbf{U}$  is a classical solution at its point of differentiability; and  $\mathbf{U}$  satisfies the *Rankine–Hugoniot conditions* across the curves of discontinuity, namely

$$[\mathbf{U}] n_t + \sum_{\ell=1}^n [\mathbf{f}^\ell(\mathbf{U})] \cdot \mathbf{n} = \mathbf{0},$$

where  $[\cdot]$  denotes the jump across the hypersurface  $\mathcal{M}$ .

Global well-posedness results are not available for nonlinear systems of conservation laws. For genuinely nonlinear one-dimensional systems with Riemann initial data, Lax [Lax57] proved existence and stability of entropy solutions (weak solutions satisfying an entropy condition). The generalization to one-dimensional Cauchy problems for initial data with “small” total variation, was developed by Glimm [Gli65]. No such results are available for non-convex and non-strictly hyperbolic systems, such as the MHD problem, nor have they been extended to systems of hyperbolic conservation laws in multi-dimensions. Non-unique solutions of the MHD Riemann problem are known [BKP96], see also [Tor02] and references therein. For a fixed state  $\mathbf{U}^*$ , the Rankine–Hugoniot conditions determine the so-called *Hugoniot loci*,  $m$  one-parameter families consisting of all states  $\mathbf{U}$  that can be connected to  $\mathbf{U}^*$  via a discontinuous shock moving at speed  $s$ . In presence of convex fluxes, the Lax entropy condition selects the parts of the Hugoniot locus representing physically admissible shocks: jumps in the  $p$ -th wave family, traveling at speed  $s$  and obeying Rankine–Hugoniot conditions, must satisfy  $\lambda^p(\mathbf{U}_L) > s > \lambda^p(\mathbf{U}_R)$ , if the  $p$ -th field is

genuinely nonlinear,  $\lambda^p(\mathbf{U}_L) = s = \lambda^p(\mathbf{U}_R)$ , if the  $p$ -th field is linearly degenerate. Here the subscripts hint at the upstream and downstream sides of the shock. In this situation, the solution of the Riemann problem admits only one type of shocks, sometimes referred to as *regular* or Lax shocks. The latter are characterized by  $m + 1$  characteristics converging into the shock and  $m - 1$  characteristics leaving the shock (*cf.* [DP01], [Daf05, Section 8.3]). On the contrary, the non-convexity of the flux function in MHD entails that several types of admissible shocks can be associated with the same characteristic field, regular shocks and the so-called *intermediate* shocks. If less than  $m + 1$  characteristics converge, the shock is undercompressive, whilst in overcompressive shocks more than  $m + 1$  characteristics impinge on the shock. In overcompressive shocks, the Lax entropy condition is fulfilled simultaneously by distinct families of characteristics, e.g.,  $\lambda^p(\mathbf{U}_L) > \lambda^{p'}(\mathbf{U}_L) > s > \lambda^p(\mathbf{U}_R) > \lambda^{p'}(\mathbf{U}_R)$ . The physical significance of intermediate waves is not yet clear despite the fact that it has been long debated. We refer to [De 99, Section 9] for a comprehensive discussion on the topic.

## MHD Shocks and Discontinuities

In the MHD model, with the exception of the entropy wave, adjacent states may be separated by either a shock or a rarefaction wave generated by abrupt changes in pressure. The entropy and Alfvén waves are linearly degenerate [JT64, Chapter 6]. The discontinuities in linearly degenerate characteristic fields are *contact discontinuities* namely transition layers with no plasma flow and characterized by changes in density but continuous thermal pressure, magnetic induction field and velocity. The magnetosonic waves are genuinely nonlinear and, since the quantities  $\partial_{\mathbf{U}} \lambda^p(\mathbf{U}) \cdot r^p(\mathbf{U})$ ,  $p \in \{1, 3, 5, 7\}$ , may change sign, the characteristic speeds of the magnetosonic waves are non-convex.

For a shock perpendicular to one of the coordinate directions, we can establish the Rankine–Hugoniot conditions in a transformed frame where the shock remains stationary. Let  $s$  denote the normal speed of the shock front and let  $\mathbf{u}' := \mathbf{u} - s\mathbf{n}$  be the fluid velocity in the shock frame of reference. Integration across the shock yields the Rankine–Hugoniot jump conditions for ideal MHD [KBC89],

$$\begin{aligned} [\rho \mathbf{u}' \cdot \mathbf{n}] &= 0, \\ \left[ \rho |\mathbf{u}' \cdot \mathbf{n}|^2 + p + \frac{1}{2} \|\mathbf{B} \times \mathbf{n}\|_{\ell^2}^2 \right] &= 0, \\ \rho \mathbf{u}' \cdot \mathbf{n} [\mathbf{u} \times \mathbf{n}] - \mathbf{B} \cdot \mathbf{n} [\mathbf{B} \times \mathbf{n}] &= \mathbf{0}, \\ \rho \mathbf{u}' \cdot \mathbf{n} \left[ \frac{1}{2} (|\mathbf{u}' \cdot \mathbf{n}|^2 + \|\mathbf{u}' \times \mathbf{n}\|_{\ell^2}^2) + \frac{1}{\rho} \left( \frac{\gamma}{\gamma - 1} p + \|\mathbf{B} \times \mathbf{n}\|_{\ell^2}^2 \right) \right] - \mathbf{B} \cdot \mathbf{n} [(\mathbf{u}' \times \mathbf{n}) \cdot (\mathbf{B} \times \mathbf{n})] &= 0, \\ \rho \mathbf{u}' \cdot \mathbf{n} \left[ \frac{\mathbf{B} \times \mathbf{n}}{\rho} \right] - \mathbf{B} \cdot \mathbf{n} [\mathbf{u}' \times \mathbf{n}] &= \mathbf{0}, \\ [\mathbf{B} \cdot \mathbf{n}] &= 0. \end{aligned}$$

The Rankine–Hugoniot conditions possess six shock solutions: the slow and fast shocks, and the intermediate shocks. The latter are characterized by fluid velocities greater than the Alfvén speed ahead of the shock, and smaller than the Alfvén speed upstream. The slow shocks are transition layers associated with discontinuous density, an entropy increase and a decrease of the tangential component of the magnetic induction field, whilst the fast shocks differ from the slow shocks in that the tangential component of the magnetic induction field increases. The entropy discontinuity does not propagate with respect to the fluid. It separates states with equal total pressure but different densities, temperatures and specific entropies. If, in addition, the normal component of the induction field vanishes, the entropy discontinuity is called MHD tangential discontinuity (by analogy with hydrodynamics) and it is characterized by

$$\begin{aligned} \mathbf{u}' \cdot \mathbf{n} &= 0, \quad \mathbf{B} \cdot \mathbf{n} = 0, \quad \left[ p + \frac{1}{2} \|\mathbf{B} \times \mathbf{n}\|_{\ell^2}^2 \right] = 0, \\ [\rho] &\neq 0, \quad [p] \neq 0, \quad [\mathbf{u}' \times \mathbf{n}] \neq 0, \quad [\mathbf{B} \times \mathbf{n}] \neq 0. \end{aligned}$$

The rotational or intermediate shock is an isentropic non-compressive transition layer characterized by the rotation of the tangential component of the  $\mathbf{B}$  field. Further details on intermediate shocks and admissibility criteria can be found in [KBC89].

In addition to the non-convexity, the umbilic degeneracy corresponding to non-distinct wave speeds, involves a degree of interaction between different modes yielding high singularities. The possible wave speeds degeneracies are as follows:

- If  $\mathbf{B} \cdot \mathbf{n} = 0$ , then  $c_s = c_a = 0$  and  $c_f = \sqrt{a^2 + \|\mathbf{B}\|_{\ell^2}^2/\rho}$ : this is called quintuple umbilic point since the entropy, Alfvén and slow magnetosonic waves all travel at the same speed.
- If the tangential magnetic induction field component vanishes but  $\|\mathbf{B}\|_{\ell^2} \neq 0$  and  $c_a \neq a$ , then  $c_s = c_a = \sqrt{(\mathbf{B} \cdot \mathbf{n})^2/\rho}$  and the fast magnetosonic speed coincides with the sound speed.
- If the tangential magnetic induction field component vanishes and  $\|\mathbf{B}\|_{\ell^2} \neq 0$  but the Alfvén speed coincides with the sound speed  $c_a = a$ , then  $c_s = c_f = c_a = a$ .

Despite the lack of well-posedness results for the MHD system and the complexity of its wave structure, the importance of such model in plasma dynamics has boosted the development of several numerical methods. On the one hand, numerical simulations of MHD flows can confirm the physical validity of the model and provide insight into the role of the plasma approximation. On the other hand, they can improve the current understanding of plasma behavior in a range of instances, from fusion confinement tests to astrophysical plasma dynamics.

## 1.2. Literature on Numerical Methods for MHD

In designing numerical schemes for the MHD system, it would be desirable to achieve physical correct and accurate reproductions of the system wave structure with minimal smearing of the solutions and without introducing spurious oscillations near shocks. Moreover, a robust numerical scheme should generate physically admissible solutions, being stable in the limit of high magnetic Reynolds number and it should satisfy the divergence constraint on the magnetic induction field at the discrete level. The latter is by no means a straightforward task and has been argument of active research in the last years, see Section 1.2.1.

Numerical methods for the MHD equations, and in general for systems of conservation laws, are well-developed in the form of finite volume schemes, see e.g. [GR96; LeV02] and references therein. The finite volume approach is based on conservative local discretizations of the balance laws in integral form, as explained in Chapter 6. Numerical approximations of the interface fluxes are designed via exact or approximate solutions of the Riemann problem at each cell interface. Non-oscillatory piecewise polynomial reconstructions enable high order spatial accuracy, whilst temporal high order accuracy can be achieved with stability preserving Runge–Kutta methods. Finite volume schemes are rather simple to implement, computationally efficient and well-suited to reproduce the physical manifestations of nonlinear terms. However, a range of complications are related to stability and convergence analyses for fully discrete schemes and in handling limiters, to the extension to arbitrarily high order and to the treatment of unstructured meshes. Moreover, finite volume schemes cannot typically satisfy the divergence constraint exactly, as explained in the forthcoming Section 1.2.1.

Other classes of numerical methods have been investigated for the numerical solution of systems of hyperbolic conservation laws, among others, finite element and discontinuous Galerkin methods [CS98], and spectral methods [GH01].

At the best of our knowledge, discrete differential forms in the guise of edge and face elements have only been employed in the development of mixed finite element discretizations of incompressible resistive MHD with partial Lie derivative by [Sch04] and more recently in [HMX16]. However, robustness of the aforementioned schemes with respect to large magnetic Reynolds number is still an issue as they rely on standard discretizations of the transport operator.

### 1.2.1. Survey on Divergence Cleaning Techniques

The divergence-free evolution of the magnetic induction is of core importance in the design of numerical schemes for the MHD problem. Violating the divergence constraint results in the addition of extra source terms in the momentum and energy equations. The effect is the onset of unphysical plasma transport orthogonal to the magnetic induction lines, and the loss of momentum and energy conservation [BB80]. Indeed, a spurious Lorentz force  $\mathbf{F}$  along the magnetic induction  $\mathbf{F} \cdot \mathbf{B} = (\mathbf{J} \times \mathbf{B}) \cdot \mathbf{B} = \|\mathbf{B}\|_{\ell^2}^2 \operatorname{div} \mathbf{B}$  is generated if the magnetic induction is not solenoidal. This can trigger numerical instabilities and possible slump in pressure and density to negative values, see e.g. [Tót00]. Typically, finite volume discretizations only provide control of the error generated by some particular discrete approximation (generally not by others) of the divergence operator of the order of the truncation spatial and temporal error. Nevertheless, near shock discontinuities such error may become very large. Moreover, since the divergence wave is stationary, divergence errors in localized regions can grow in time with mesh refinement unables to counteract the numerical pollution of the solution (see [Bal01]).

In the remainder of this section, we summarize the most commonly used divergence cleaning techniques. An in-depth survey on the many numerical strategies devised to evolve solenoidal discrete magnetic induction fields, in the context of finite volume discretizations, is supplied by [Tót00].

**The Projection Method.** The projection method is a predictor-corrector strategy introduced in [BB80] and based on correcting a non-solenoidal magnetic induction by removing its “unphysical part” represented by the gradient of a scalar potential. In a method of line approach, let  $\bar{\mathbf{B}}_h^{n+1}$  denote the magnetic induction field derived at the  $n$ -th time step of some arbitrary numerical scheme on a topologically trivial computational domain. The Helmholtz–Hodge decomposition of  $\bar{\mathbf{B}}_h^{n+1}$  gives  $\bar{\mathbf{B}}_h^{n+1} = \operatorname{curl} \mathbf{A} + \operatorname{grad} \phi$  for  $\mathbf{A} \in H(\operatorname{curl}, \Omega)$  and  $\phi \in H^1(\Omega)$ . The *corrected* field  $\mathbf{B}_h^{n+1}$  is obtained as  $\mathbf{B}_h^{n+1} = \bar{\mathbf{B}}_h^{n+1} - \operatorname{grad} \phi$  where the scalar potential satisfies the Poisson problem  $\operatorname{div}_h(\operatorname{grad} \phi) = \operatorname{div}_h \bar{\mathbf{B}}_h^{n+1}$ . Note that, if the underlying numerical scheme is conforming, that is  $\bar{\mathbf{B}}_h^{n+1} \in H(\operatorname{div}, \Omega)$ , then the second order diffusion operator is globally well-defined and the divergence constraint is exactly satisfied by  $\mathbf{B}_h^{n+1}$ . On the other hand, in a finite volume approach, the Poisson problem has to be solved on each control volumes and only a local discrete divergence of the corrected field  $\mathbf{B}_h^{n+1}$  can be preserved zero. As advocated in [Tót00, Section 5], a careful implementation of the scheme can preserve the conservative properties of the underlying finite volume discretization and flexibility. The computational cost of adding an elliptic equation to solve at each time step is objectively a drawback of the projection method. Furthermore, some oscillations near strong shocks can arise as the result of the approximation needed to solve the additional elliptic equation.

**The Godunov–Powell Method or Eight-Wave Formulation.** In [Pow94], Powell proposed a semi-conservative form of the MHD equations, which differs from the conservative one by the addition of source terms, proportional to  $\operatorname{div} \mathbf{B}$ , on the system right hand side. Powell’s formulation is symmetrizable and Galilean invariant, and it is known as eight-wave form since it is characterized by a wave fan with one extra wave, a stationary plane wave associated with the propagation of  $\operatorname{div} \mathbf{B}$ . Finite volume schemes, resulting from the discretization of Powell’s formulation, have proven to be numerically robust in many test cases, with the caveat that the formulation is no longer conservative. Moreover, the addition of the non-physical divergence wave might introduce incorrect jump conditions associated with strong shocks, wrecking the reliability of the scheme [Tót00, Section 6.3.2]. Note that the so-called Godunov–Powell source terms require an ad hoc discretization to prevent instabilities of the overall scheme (see [Fuc+11, Section 2.2]). This approach is computationally inexpensive compared to the projection method. Moreover, in the presence of outflow boundary conditions, the non-zero divergence, associated with numerical effects, should be advected away from the computational domain by the flow, since the divergence wave propagates with the fluid velocity:  $\partial_t(\operatorname{div} \mathbf{B}) + \operatorname{div}(\mathbf{u}(\operatorname{div} \mathbf{B})) = 0$ .

**The Generalized Lagrange Multiplier MHD Formulation.** A further approach to handle solenoidal magnetic induction fields is based on enforcing the divergence constraint through a Lagrange multiplier. The generalized Lagrangian multiplier method consists in an operator splitting procedure where, once the hyperbolic balance equations are numerically solved, the divergence constraint is enforced by solving an equation for the Lagrange multiplier. This additional equation can be chosen to be either hyperbolic, parabolic, or elliptic. In the latter case, the scheme reduces to the projection method previously introduced.

In the parabolic case, the local divergence errors are dissipated and diffused (under suitable boundary conditions): the Lagrange multiplier  $\phi$  satisfies the heat equation  $\partial_t \phi - c_p^2 \Delta \phi = 0$  for some  $c_p \in \mathbb{R}$ . Last, the hyperbolic correction boils down to advecting the divergence error with finite speed  $c_h > 0$ , by additionally solving the wave equation  $\partial_{tt} \phi - c_h^2 \Delta \phi = 0$ . The hyperbolic approach was introduced in [Mun+00] for the Maxwell equations and extended to the MHD problem in [Ded+02]. It turned out to be particularly attractive in the context of finite volume discretizations since it provides an add-on to existing finite volume codes. The main disadvantages of the generalized Lagrange multiplier approach are related to the fact that the system has been modified and, as already mentioned, this might result in incorrect jump conditions across strong shocks. Moreover, both the hyperbolic and parabolic corrections require the definition of parameters: the speed of propagation  $c_h$  in the first case, the damping factor  $c_p$  in the second case.

**The Constrained Transport Method.** Since fully discontinuous polynomial spaces (and, hence, finite dimensional spaces associated with finite volume discretizations) do not accommodate a discrete de Rham complex structure (*cf.* Chapter 2), there has been a current of schemes based on interpolations of the discrete magnetic induction into suitable  $H(\text{div}, \Omega)$ -conforming spaces. On the cellular partition of an  $n$ -dimensional domain, these methods consist in staggering the discretizations of the velocity and magnetic induction so that the normal component of the  $\mathbf{B}$  field at each mesh  $(n-1)$ -face is well-defined and the electric field components are collocated at the  $(n-2)$ -cells. Stokes' law applied to Faraday's equation yields a discrete time-update strategy which preserves the magnetic induction field solenoidal. Such an algorithm was first proposed by Yee [Yee66] for the Maxwell equations and its evolutions are referred to as constrained transport methods [HE89]. We refer the interested reader to [Tót00, p. 607 and Section 4] for a detailed literature and comparison of different constrained transport approaches. Staggering of the variables leads to complications on parallel implementation, with adaptive mesh refinement and arbitrary meshes.

**Schemes Based on the Magnetic Potential.** The use of the magnetic potential to numerically solve the MHD equations in the context of finite volume discretizations has been suggested in [HE89] and developed in [LD04] within a constrained transport strategy. In [Ros06], Rossmanith proposed a constrained transport method without staggering: the magnetic induction field is obtained from the discretization of the ideal MHD system through the Riemann solver-based wave propagation method of [LeV97], and further corrected using a discrete magnetic potential. The latter ensues, at each time step, from the discretization of the magnetic advection equation solved with given velocity averaged between two intermediate time steps. Mishra and Tadmor proposed in [MT12] a unified formulation of constrained transport methods where edge centered numerical fluxes are replaced by vertex centered potentials ensuring the preservation of the divergence constraint. All the aforementioned works are contextualized in the category of constrained transport methods, and they do not fit the finite element exterior calculus framework. In this respect the approach we introduce in this work represents a novelty in the context of numerical approximations of the MHD problem.

### 1.2.2. The Semi-Godunov Approach

The splitting approach we pursue in the present work takes the cue from a family of reduction techniques, called *semi-Godunov schemes*, devised to handle systems of hyperbolic conservation laws in a particular form. Semi-Godunov schemes were introduced in [KMR09] to solve one-dimensional non-strictly hyperbolic systems of conservation laws with a triangular structure, i.e., where the evolution of the  $k$ -th system variable is independent of the subsequent unknowns  $(u_{k+1}, \dots, u_m)$ . It consists of decoupling the system into a set of scalar conservation laws for each variable, and to treat the other variables as (discontinuous) coefficients. A Godunov-type approach is then implemented to discretize the new set of equations with discontinuous fluxes. Scalar conservation laws with discontinuous coefficients have been studied by several authors, see [KMR09, p. 561] and references therein.

As suggested in Section 1.1, the physical nature of the MHD models hints at a “splitting and coupling approach” to deal with the entanglement of rather different phenomena, the fluid motion and the electromagnetic effects. The pith of the numerical schemes investigated in the present work is a discretization of the MHD model mimicking such strategy. Analogously to a semi-Godunov scheme and

inspired by [FMR09], we perform a local (in time) reduction of the MHD system into two systems with discontinuous coefficients: the magnetic induction/potential is advected with a known (discontinuous) velocity field, *cf.* Part I, and the **B** field is treated as a discontinuous coefficient in the system of conservation laws for the fluid variables, *cf.* Part II.



# Part I.

## Maxwell's Equations in MHD



## 2. Differential Forms in Electromagnetism

Mathematical and physical research on electromagnetic phenomena in the last forty years disclosed that the peculiar features of different physical quantities (e.g. the distinction between fields and fluxes) can be canonically captured by the wealth of structure inherent in exterior calculus. Differential forms originated in the works of Grassmann and of Cartan, and their connection with field theory dates back to the early 80's (see [Hip02, Bibliographical notes p. 243] for references). To date, the relationship between electromagnetism and algebraic topology, and the interpretation of electromagnetic fields in terms of differential forms has been widely investigated in mathematics, mathematical physics and computational sciences.

As seen in Chapter 1, the standard model for resistive MHD (1.9) comprises balance equations for mass, momentum and energy together with material laws and Maxwell's equations, in their magneto-quasistatic reduction, for the electromagnetic fields. The formulation of the linear eddy current model in the presence of a conducting fluid moving with velocity  $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$  boils down to the evolution PDE

$$\partial_t \mathbf{A} + \operatorname{curl}(\varepsilon \operatorname{curl} \mathbf{A}) + \operatorname{curl} \mathbf{A} \times \mathbf{u} + \operatorname{grad}(\mathbf{u} \cdot \mathbf{A}) = \mathbf{f}, \quad (2.1)$$

governing the evolution of the unknown magnetic vector potential  $\mathbf{A}$  and with  $\varepsilon$  being the magnetic diffusion coefficient. The vector calculus notation (2.1) conceals common features with the familiar and widely studied second order advection-diffusion equation for the unknown scalar function  $w$

$$\partial_t w - \operatorname{div} \varepsilon \operatorname{grad} w + \mathbf{u} \cdot \operatorname{grad} w = f. \quad (2.2)$$

Here the formalism of exterior algebra comes into play as a powerful tool in providing a mathematically rigorous interpretation of the electromagnetic fields, and in allowing for a unified trim treatment of the physical laws of electromagnetism. As we shall see in Section 2.2, problems (2.2) and (2.1) belong to a single family of second order evolution problems for differential forms, which we have dubbed *generalized advection-diffusion*.

In the present chapter, we introduce in Section 2.1 the theory of exterior calculus and differential forms, largely based on the treatises [Hip02], [AFW06] and [Sch95]. The algebraic and topological structure of cellular complexes provide the natural setting to establish discrete analogs of differential forms which are the subject of Section 2.3. With the intent of designing finite element schemes for the generalized advection-diffusion problem, Section 2.4 is devoted to finite element approximations based on polynomial subspaces of some Sobolev spaces with good approximation properties, and which provide structure-preserving discretizations in the sense that they form a subcomplex of the de Rham complex.

### 2.1. Exterior Calculus

Having established that the language of vector calculus is not exhaustive in capturing the physical nature of electromagnetic fields, the exterior calculus structure is able to grasp two manifestations of such fields and fluxes: one as objects integrated over oriented surfaces and a pointwise perspective peculiar to measurement procedures. As "quantities with reference to a  $k$ -dimensional manifold", electromagnetic fields are integral forms according to the following definition.

**Definition 2.1.1** (Integral form, [Hip02, Definition 1]). Let  $\mathcal{M}$  be a piecewise smooth  $n$ -dimensional manifold. Let  $\mathcal{S}_k(\mathcal{M})$ ,  $k \in [0, n] \cap \mathbb{N}$  be the set of compact, oriented, piecewise smooth  $k$ -dimensional submanifold of  $\mathcal{M}$ . An *integral form* of degree  $k$  is a continuous (with respect to the deformation topology) additive alternating mapping  $\omega : \mathcal{S}_k(\mathcal{M}) \rightarrow \mathbb{C}$ . Integral  $k$ -forms on  $\mathcal{M}$  form a vector space denoted by  $\mathcal{F}^k(\mathcal{M})$ .

In a local perspective, as “quantities at a point in space”, we need to introduce the so-called differential forms to represent electromagnetic fields. To this aim, let  $V$  be a real vector space of dimension  $n$  and let  $V^k$ ,  $k \in [1, n] \cap \mathbb{N}$ , denote the Cartesian power of  $V$ ,  $V^k := V \times \dots \times V = \{(v_1, \dots, v_k) : v_i \in V, \forall 1 \leq i \leq k\}$ . We define the space  $\text{Alt}^k V$  of *alternating algebraic  $k$ -forms* on  $V$  as the space of  $k$ -linear maps  $V^k \rightarrow \mathbb{R}$  that assign to a  $k$ -tuple of elements  $(v_1, \dots, v_k) \in V^k$  a real number  $\omega(v_1, \dots, v_k)$ , with reversed sign whenever two arguments are interchanged. For  $k = 0$ ,  $\text{Alt}^0 V := \mathbb{R}$ .

**Definition 2.1.2** (Differential form, [Hip02, Definition 2]). Let  $\mathcal{M}$  be a smooth  $n$ -dimensional manifold. A *differential form* of degree  $k \in [0, n] \cap \mathbb{N}$  is a mapping assigning to each  $x \in \mathcal{M}$  an alternating  $k$ -multilinear form on the tangent space of  $\mathcal{M}$  at  $x$ :  $x \in \mathcal{M} \mapsto \omega_x \in \text{Alt}^k T_x \mathcal{M}$ . A differential  $k$ -form is of class  $C^m$ ,  $m \in \mathbb{N}_0$ , if the map  $x \in \mathcal{M} \mapsto \omega_x(v_1, \dots, v_k) \in \mathbb{R}$  is  $m$ -times continuously differentiable whenever  $(v_1, \dots, v_k) \in (T_x \mathcal{M})^k$  are smooth vector fields.

The space of smooth differential  $k$ -forms on  $\mathcal{M}$  is denoted by  $\Lambda^k(\mathcal{M})$ . It can be seen as the space of smooth sections of the exterior  $k$ -bundle  $\bigcup_{x \in \mathcal{M}} \text{Alt}^k T_x \mathcal{M}$ , as in [Sch95, Definition 1.2.1]. Smooth differential forms possess an algebraic structure when endowed with a multiplication operator.

**Definition 2.1.3** (Wedge product). Let  $j, p \in \mathbb{N}$ ,  $j, p \geq 1$  with  $j < p$ , and let  $G(j, p)$  be the set of permutations  $\sigma$  of  $\{1, \dots, p\}$  such that  $\sigma(1) < \dots < \sigma(j)$  and  $\sigma(j+1) < \dots < \sigma(p)$ . Let  $(v_1, \dots, v_{j+k}) \in V^{j+k}$  be arbitrary smooth vector fields on  $V$ . The *exterior or wedge product* of alternating forms is the operator  $\wedge : \text{Alt}^j V \times \text{Alt}^k V \rightarrow \text{Alt}^{j+k} V$ , defined as

$$(\omega \wedge \eta)(v_1, \dots, v_{j+k}) := \sum_{\sigma \in G(j, j+k)} \text{sign}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(j)}) \eta(v_{\sigma(j+1)}, \dots, v_{\sigma(j+k)}).$$

The wedge product is bilinear, associative and graded anti-commutative, that is,  $\eta \wedge \omega = (-1)^{jk} \omega \wedge \eta$  for all  $\omega \in \text{Alt}^j V$  and  $\eta \in \text{Alt}^k V$ . It can be extended to differential forms on a smooth  $n$ -dimensional manifold  $\mathcal{M}$  by applying the exterior product of algebraic forms pointwise on each  $x \in \mathcal{M}$ , i.e.,  $(\omega \wedge \eta)_x = \omega_x \wedge \eta_x$ , for all  $\omega \in \Lambda^j(\mathcal{M})$  and  $\eta \in \Lambda^k(\mathcal{M})$ .

In order to define transport operators acting on differential forms, we need to introduce the differentiation of forms and a further algebraic operation on  $\Lambda^k(\mathcal{M})$ , the interior product. For the sake of simplicity, we restrict to the case of  $\Omega$  being a domain in  $\mathbb{R}^n$ . The reader interested into a more general case is addressed to [Sch95, Definition 1.2.2 (e)] or [Lan99, Proposition 3.2]. For a domain  $\Omega \subset \mathbb{R}^n$ , each tangent space  $T_x \Omega$  can be identified with  $\mathbb{R}^n$ , and hence, for  $\omega \in \Lambda^k(\Omega)$  and  $(v_1, \dots, v_k) \in (T_x \Omega)^k$ , the mapping  $x \in \Omega \mapsto \omega_x(v_1, \dots, v_k) \in \mathbb{R}$  is smooth. In this setting, we can define the exterior derivative and the contraction operator as follows.

**Definition 2.1.4** (Exterior derivative). Let  $\omega \in \Lambda^k(\Omega)$ ,  $x \in \mathcal{M} \mapsto \omega_x \in \text{Alt}^k T_x \mathcal{M}$  and  $(v_1, \dots, v_{k+1}) \in (T_x \Omega)^{k+1}$ , the *exterior derivative* is the operator  $d^k : \Lambda^k(\Omega) \rightarrow \Lambda^{k+1}(\Omega)$  defined as

$$(d^k \omega_x)(v_1, \dots, v_{k+1}) = \sum_{j=1}^{k+1} (-1)^{j+1} \partial_{v_j} \omega_x(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_{k+1}),$$

where  $\partial_{v_j}$  is the partial derivative in the  $v_j$  direction. The exterior derivative is a graded linear operator of degree +1 and it is a differential, i.e.,  $d^{k+1} \circ d^k = 0$ .

**Definition 2.1.5** (Contraction operator). On an  $n$ -dimensional real vector space  $V$ , let  $(v_1, \dots, v_k) \in V^k$ . The *interior product* or *contraction*  $i_u : \text{Alt}^k V \rightarrow \text{Alt}^{k-1} V$  of an alternating  $k$ -form  $\omega \in \text{Alt}^k V$  by a vector  $u \in V$  is defined as:  $(i_u \omega)(v_1, \dots, v_{k-1}) := \omega(u, v_1, \dots, v_{k-1})$ . Applying pointwise the contraction of algebraic forms defines a contraction operator on the space of differential forms  $\Lambda^k(\Omega)$  on a domain  $\Omega \subset \mathbb{R}^n$ . Since the forms are alternating, it holds  $i_u(i_u \omega) = 0$  for all  $\omega \in \Lambda^k(\Omega)$ .

Local bases for the space of smooth differential forms on a smooth  $n$ -manifold can be described in terms of local coordinate charts on the manifold. The exterior derivative and the contraction operator are independent of the choice of local coordinates. More generally, if  $\phi : \mathcal{M} \rightarrow \widehat{\mathcal{M}}$  is a diffeomorphism of (sufficiently) smooth  $n$ -dimensional manifolds  $\mathcal{M}$  and  $\widehat{\mathcal{M}}$ , then the tangent map  $D\phi_x : T_x \mathcal{M} \rightarrow T_{\phi(x)} \widehat{\mathcal{M}}$

is linear for every  $x \in \mathcal{M}$ . Moreover,  $\phi$  induces a *pullback of differential forms*  $\phi^* : \Lambda^k(\widehat{\mathcal{M}}) \rightarrow \Lambda^k(\mathcal{M})$  defined for all  $\omega \in \Lambda^k(\widehat{\mathcal{M}})$  as

$$\begin{aligned} (\phi^*\omega)_x(v_1, \dots, v_k) &= (\mathrm{D}\phi_x)^*\omega_{\phi(x)}(v_1, \dots, v_k) \\ &= \omega_{\phi(x)}(\mathrm{D}\phi_x(v_1), \dots, \mathrm{D}\phi_x(v_k)), \end{aligned}$$

with  $(v_1, \dots, v_k) \in (T_x\mathcal{M})^k$ . The pullback commutes with the exterior derivative  $\phi^*(\mathrm{d}^k\omega) = \mathrm{d}^k(\phi^*\omega)$ , and with the contraction operator  $\phi^*(\mathbf{i}_{\hat{u}}\omega) = \mathbf{i}_u(\phi^*\omega)$ ,  $\hat{u}$  being the image of  $u$  under the tangent map. Moreover, the pullback preserves the wedge product  $\phi^*(\omega \wedge \eta) = \phi^*\omega \wedge \phi^*\eta$ , for all  $\omega, \eta \in \Lambda^k(\widehat{\mathcal{M}})$ .

Differential forms can be integrated and differentiated without recourse to a metric structure. Specifically, integration of differential forms and exterior derivative are related via Stokes' theorem. If  $\Omega$  is a domain in  $\mathbb{R}^n$  with boundary  $\partial\Omega$  endowed with the induced orientation (the extension to oriented  $n$ -manifolds with boundary holds), then

$$\int_{\Omega} \mathrm{d}^{n-1}\omega = \int_{\partial\Omega} \mathrm{tr}\omega \quad \omega \in \Lambda^{n-1}(\Omega), \quad (2.3)$$

where the trace map  $\mathrm{tr} : \Lambda^{n-1}(\Omega) \rightarrow \Lambda^{n-1}(\partial\Omega)$  is defined as the pullback of the inclusion  $\iota : \partial\Omega \hookrightarrow \Omega$ .

## Vector Proxies

When as manifold we consider the affine space  $A(\mathbb{R}^3)$ , the tangent space  $T_x(A(\mathbb{R}^3))$  at  $x$  can be identified with  $\mathbb{R}^3$  endowed with the Euclidean inner product (denoted by  $\cdot$  henceforth). Therefore, natural identifications of  $\Lambda^0(A(\mathbb{R}^3))$  and  $\Lambda^3(A(\mathbb{R}^3))$  with  $\mathbb{R}$  and of  $\Lambda^1(A(\mathbb{R}^3))$  and  $\Lambda^2(A(\mathbb{R}^3))$  with  $\mathbb{R}^3$  can be established as follows [Hip02, Table 2.1]: Let  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$ ,

$$\begin{aligned} \mathbf{x} &\mapsto \omega(\mathbf{x}) =: u(\mathbf{x}), \\ \mathbf{x} &\mapsto \{\mathbf{v}_1 \mapsto (\omega(\mathbf{x}))(\mathbf{v}_1)\}, & (\omega(\mathbf{x}))(\mathbf{v}_1) &=: \mathbf{u}(\mathbf{x}) \cdot \mathbf{v}_1, \\ \mathbf{x} &\mapsto \{(\mathbf{v}_1, \mathbf{v}_2) \mapsto (\omega(\mathbf{x}))(\mathbf{v}_1, \mathbf{v}_2)\}, & (\omega(\mathbf{x}))(\mathbf{v}_1, \mathbf{v}_2) &=: \mathbf{u}(\mathbf{x}) \cdot (\mathbf{v}_1 \times \mathbf{v}_2), \\ \mathbf{x} &\mapsto \{(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \mapsto (\omega(\mathbf{x}))(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)\}, & (\omega(\mathbf{x}))(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) &=: u(\mathbf{x}) \det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3). \end{aligned}$$

Based on these isomorphisms  $g_k : \Lambda^k(A(\mathbb{R}^3)) \rightarrow \mathbb{R}^3$ , 0- and 3-forms can be associated with continuous scalar functions  $u$ , whilst 1- and 2-forms are naturally identified with continuous vector-valued functions  $\mathbf{u}$ . The associated fields are called *proxy fields* for the forms (the term was coined in [Bos98b]) and all the operations on exterior forms correspond to operations on scalars and vectors. Vector proxies are by definition coordinate-dependent, in that endowing  $\mathbb{R}^3$  with a different inner product yields different vector proxies.

In the forthcoming analysis, we rely on transformations of scalar and vector-valued functions under diffeomorphisms of manifolds. More specifically, let  $\Omega$  and  $\widehat{\Omega}$  be domain in  $\mathbb{R}^3$  and let  $\Phi : \widehat{\mathbf{x}} \in \widehat{\Omega} \mapsto \mathbf{x} \in \Omega$  be a diffeomorphism with Jacobian  $D\Phi$ . The pullbacks on vector proxies of continuous differential forms are [Hip02, Equations (2.16)-(2.19)]:

$$\begin{aligned} h_0^\Phi &= g_0 \circ \Phi^* \circ g_0^{-1}, & (h_0^\Phi u)(\widehat{\mathbf{x}}) &= u(\Phi(\widehat{\mathbf{x}})), \\ h_1^\Phi &= g_1 \circ \Phi^* \circ g_1^{-1}, & (h_1^\Phi u)(\widehat{\mathbf{x}}) &= D\Phi(\widehat{\mathbf{x}})^\top \mathbf{u}(\Phi(\widehat{\mathbf{x}})), \\ h_2^\Phi &= g_2 \circ \Phi^* \circ g_2^{-1}, & (h_2^\Phi u)(\widehat{\mathbf{x}}) &= \det D\Phi(\widehat{\mathbf{x}}) D\Phi(\widehat{\mathbf{x}})^{-1} \mathbf{u}(\Phi(\widehat{\mathbf{x}})), \\ h_3^\Phi &= g_3 \circ \Phi^* \circ g_3^{-1}, & (h_3^\Phi u)(\widehat{\mathbf{x}}) &= \det D\Phi(\widehat{\mathbf{x}}) u(\Phi(\widehat{\mathbf{x}})), \end{aligned} \quad (2.4)$$

where  $u$  is a continuous scalar function and  $\mathbf{u}$  a continuous vector-valued function on  $\Omega$ . The inverse of the maps  $h_1^\Phi$  and  $h_2^\Phi$  are known as the covariant and contravariant Piola transformations, respectively (see [BBF13, Section 2.1.3] for further details).

### Inner Product, Volume Form and Hodge Operator

Let  $V$  be an  $n$ -dimensional vector space. If  $\{b_i\}_{i=1}^n$  and  $\{\tilde{b}_i\}_{i=1}^n$  are two bases of  $V$ , then  $\omega(b_1, \dots, b_n) = \det(g)\omega(\tilde{b}_1, \dots, \tilde{b}_n)$  for any  $\omega \in \text{Alt}^n V$ , where  $g$  is the linear map associated with the change of basis  $g(\tilde{b}_i) = b_i$ ,  $1 \leq i \leq n$ . An equivalence relation on bases of  $V$  can be defined by saying that two bases  $\{b_i\}_{i=1}^n$  and  $\{\tilde{b}_i\}_{i=1}^n$  have the same orientation if and only if  $\omega(b_1, \dots, b_n)$  and  $\omega(\tilde{b}_1, \dots, \tilde{b}_n)$  have the same sign (or equivalently  $\det(g) > 0$ ). Setting an *orientation* of  $V$  corresponds to choosing one of the two equivalence classes. Then we can determine a unique alternating  $n$ -linear form,  $\text{vol} \in \text{Alt}^n V$ , called *volume form*, such that  $\text{vol}(e_1, \dots, e_n) = 1$  for all orthonormal, positively oriented bases  $\{e_i\}_{i=1}^n$ . If an  $n$ -dimensional vector space  $V$  is oriented and equipped with an inner product, then  $\text{Alt}^k V$  is naturally endowed with the inner product

$$\langle \omega, \eta \rangle := \sum_{\sigma} \omega(e_{\sigma(1)}, \dots, e_{\sigma(k)}) \eta(e_{\sigma(1)}, \dots, e_{\sigma(k)}), \quad \forall \omega, \eta \in \text{Alt}^k V, \quad (2.5)$$

where the sum is over increasing sequences  $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$  and  $\{e_i\}_{i=1}^n$  is any orthonormal basis of  $V$ .

For smooth differential forms on an oriented Riemannian manifold  $\mathcal{M}$ , the tangent space  $T_x \mathcal{M}$  at  $x \in \mathcal{M}$  is endowed with an inner product and so are the spaces  $\text{Alt}^k T_x \mathcal{M}$ . The volume form  $\text{vol} \in \Lambda^n(\mathcal{M})$  is the unique differential  $n$ -form such that  $\text{vol}_x$  is the volume form associated with the oriented inner product space  $T_x \mathcal{M}$ , for each  $x \in \mathcal{M}$ . For a fixed  $\omega \in \text{Alt}^k T_x \mathcal{M}$ , the wedge product induces a linear map  $\text{Alt}^{n-k} T_x \mathcal{M} \rightarrow \mathbb{R}$  by composing the map  $\eta \in \text{Alt}^{n-k} T_x \mathcal{M} \mapsto \omega \wedge \eta$  with the canonical isomorphism  $c \text{vol} \in \text{Alt}^n T_x \mathcal{M} \mapsto c \in \mathbb{R}$ . The Riesz representation theorem ensures that there exists a form  $\star \omega \in \text{Alt}^{n-k} T_x \mathcal{M}$  such that

$$\omega \wedge \eta = \langle \star \omega, \eta \rangle \text{vol}, \quad \forall \eta \in \text{Alt}^{n-k} T_x \mathcal{M}.$$

The linear map  $\omega \in \text{Alt}^k T_x \mathcal{M} \mapsto \star \omega \in \text{Alt}^{n-k} T_x \mathcal{M}$  is called the *Hodge operator*. Note that the  $\star$  operator depends on the metric on  $\mathcal{M}$ , is an isometry, and is (graded) idempotent, i.e.,  $\star(\star \omega) = (-1)^{k(n-k)} \omega$ , for  $\omega \in \Lambda^k(\Omega)$ . If not otherwise specified, we denote with  $\star$  the Hodge operator on an  $n$ -manifold  $\mathcal{M}$  with boundary, and with  $\star_{n-1}$  the operator on a  $(n-1)$ -dimensional submanifold of  $\mathcal{M}$ .

Note that in Maxwell's equations the Hodge operator enters the material law (1.4),

$$b = \mu \star h, \quad (2.6)$$

where  $b \in \Lambda^2(\mathbb{R}^n)$  represents the magnetic induction and  $h \in \Lambda^1(\mathbb{R}^n)$  is the magnetic field. The magnetic permeability  $\mu$  relates to the properties of the medium, whilst the  $\star$  operator carries the information about the metric properties of the space. The correspondence between wedge product and Hodge operator for differential forms and vector proxies can be found in Table 2.1.

|                        | Exterior calculus   | Proxy calculus   |
|------------------------|---|--|
| $\wedge$ product       | $\wedge : \Lambda^1(\mathbb{R}^3) \times \Lambda^1(\mathbb{R}^3) \rightarrow \Lambda^2(\mathbb{R}^3)$ | $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (cross product) |
|                        | $\wedge : \Lambda^1(\mathbb{R}^3) \times \Lambda^2(\mathbb{R}^3) \rightarrow \Lambda^3(\mathbb{R}^3)$ | $\cdot : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ (dot product)      |
| Hodge operator $\star$ | $\star : \Lambda^0(\mathbb{R}^3) \rightarrow \Lambda^3(\mathbb{R}^3)$                                 | $\text{id} : \mathbb{R} \rightarrow \mathbb{R}$                                      |
|                        | $\star : \Lambda^1(\mathbb{R}^3) \rightarrow \Lambda^2(\mathbb{R}^3)$                                 | $\text{id} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$                                  |

Table 2.1.: Correspondence between wedge product and Hodge operator for differential forms and vector proxies. More details in [Hip02, Section 2.2], [AFW06, Table 2.1].

### Lie Derivative

With the definitions established formerly, we can introduce the generalization to smooth differential forms of the directional derivative of a function. Let  $u$  be a smooth vector field (i.e. a section of the tangent

bundle) on the  $n$ -dimensional smooth manifold  $\mathcal{M}$ . The exterior derivative and the contraction by  $u$  can be combined into an operator on smooth differential  $k$ -forms, named *Lie derivative*,  $L_u : \Lambda^k(\mathcal{M}) \rightarrow \Lambda^k(\mathcal{M})$  and defined as

$$L_u \omega := i_u d^k \omega + d^{k-1} i_u \omega, \quad \forall \omega \in \Lambda^k(\mathcal{M}). \quad (2.7)$$

This formulation is often referred to as *Cartan's homotopy* or *magic formula* (see e.g. [Sch95, Equation 2.3] or [Lee13, Theorem 14.35]). Equivalently, the Lie derivative can be defined through the homotopy formula of differential geometry (see [Lan02, Chapter V, Section 2]), by introducing the *flow* of the vector field  $u$  on the manifold  $\mathcal{M}$ , namely  $\Phi : \mathbb{R} \times \mathcal{M} \rightarrow \mathcal{M}$  such that  $\partial_t \Phi(t, x) = u(\Phi(t, x), t)$  with  $\Phi(0, x) = x$ . The Lie derivative of a  $k$ -form  $\omega \in \Lambda^k(\Omega)$  is

$$L_u \omega = \frac{d}{dt} \Big|_{t=0} \Phi_t^* \omega = \lim_{t \searrow 0} \frac{\Phi_t^* \omega(t) - \omega(0)}{t}. \quad (2.8)$$

Hence,  $L_u$  is a coordinate-independent operator which measures the rate of change of a differential form along the flow of a vector field. In particular, by looking at the material derivative of a  $k$ -form

$$\begin{aligned} \frac{D}{Dt} \omega(t) &= \frac{d}{d\tau} \Phi_\tau^* \omega(t) \Big|_{\tau=t} = \lim_{\tau \searrow 0} \frac{\omega(t) - \Phi_{-\tau}^* \omega(t - \tau)}{\tau} \\ &= \lim_{\tau \searrow 0} \frac{\omega(t) - \omega(t - \tau) + \omega(t - \tau) - \Phi_{-\tau}^* \omega(t - \tau)}{\tau} = \partial_t \omega + L_u \omega, \end{aligned}$$

the Lie derivative emerges as the generalization to  $k$ -forms of the spatial part of the material derivative. The latter represents differentiation along the characteristic curves, and hence measures the rate of change observed by a material particle moving with the fluid. This “dynamical” formulation of the Lie derivative will be leveraged in Chapter 4 to formulate numerical schemes for the advection of differential forms.

Another characterization of the Lie derivative, as Friedrichs’ operator, is particularly amenable to discretization (see [EG06] for discontinuous Galerkin discretizations of Friedrichs’ systems). Friedrichs’ symmetric systems were introduced in [Fri58] and provide a unified framework that goes beyond the traditional classification of linear stationary PDEs in hyperbolic or elliptic type. More precisely, Friedrichs’ systems are coupled first order PDEs with symmetry and positivity properties (see e.g. [EG04, Section 5.2.2]). Examples are, among others, advection-reaction equations and advection-diffusion-reaction equations, and Maxwell’s equations. Let  $\mathcal{K} : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  and let  $\{\mathcal{A}^m\}_{m=1}^n$  be a family of matrix-valued fields  $\mathcal{A}^m : \Omega \rightarrow \mathbb{R}^{n \times n}$  such that  $\mathcal{A}^m = (\mathcal{A}^m)^\top$ . Assume that  $\mathcal{K} \in L^\infty(\Omega)^{n \times n}$ ,  $\mathcal{A}^m \in L^\infty(\Omega)^{n \times n}$  for  $1 \leq m \leq n$ , and  $\sum_{m=1}^n \partial_{x_m} \mathcal{A}^m \in L^\infty(\Omega)^{n \times n}$ . A *Friedrichs’ symmetric operator* is an  $\mathbb{R}^n$ -valued linear differential operator  $\mathcal{F} = (\mathcal{F}^1, \dots, \mathcal{F}^n)$  such that,

$$\mathcal{F}^i : \Omega \rightarrow \mathbb{R}, \quad \mathcal{F}^i \mathbf{w} := \sum_{j=1}^n \sum_{m=1}^n \mathcal{A}_{i,j}^m \partial_{x_m} w_j + \sum_{j=1}^n \mathcal{K}_{i,j} w_j, \quad \forall 1 \leq i \leq n, \forall \mathbf{w} = (w_1, \dots, w_n) \in \Omega.$$

In order to express the Lie derivative as a Friedrichs’ operator, we employ a Leibniz rule with respect to the wedge product, namely

$$L_u(\omega \wedge \star \eta) = L_u \omega \wedge \star \eta - \omega \wedge \star L_u \eta, \quad \forall \omega \in \Lambda^k(\Omega), \eta \in \Lambda^j(\Omega), \quad (2.9)$$

where  $\mathcal{L}_u$  denotes the Lie co-derivative:  $\mathcal{L}_u = -(-1)^{k(n-k)} \star L_u \star$ , see also Definition 2.1.7.

**Lemma 2.1.6** (Lie derivative as Friedrichs’ operator, [HHP16, Appendix A]). *Let  $\Omega$  be a domain in  $\mathbb{R}^n$  and let  $\mathbf{u} = (u_1, \dots, u_n)$  be a smooth vector field. Let  $\{\mathbf{e}_I^k\}_{I=1}^{\binom{n}{k}} \subset \text{Alt}^k \mathbb{R}^n$  be an orthonormal basis of alternating  $k$ -forms. If  $\mathbf{w}$  is proxy field of the differential  $k$ -form  $\omega \in \Lambda^k(\Omega)$ , denoted by  $v.p.(\omega) = \mathbf{w}$ , then the Lie derivative  $L_{\mathbf{u}} \omega \in \Lambda^k(\Omega)$  of  $\omega$  with respect to the vector field  $\mathbf{u}$  has vector proxy*

$$v.p.(L_{\mathbf{u}} \omega) = \sum_{m=1}^n u_m \partial_{x_m} \mathbf{w} + \mathbf{K} \mathbf{w}, \quad \mathbf{K}_{I,J} := \delta_{I,J} \sum_{m=1}^n \partial_{x_m} u_m + v.p.(\mathbf{e}_J^k \wedge \star \mathcal{L}_{\mathbf{u}} \mathbf{e}_I^k),$$

for  $1 \leq I, J \leq \binom{n}{k}$  and where  $\delta_{I,J}$  denotes the Kronecker delta.

*Proof.* Let  $\omega \in \Lambda^k(\Omega)$ ,  $\omega = \sum_{J=1}^{\binom{n}{k}} \omega_J \mathbf{e}_J^k$  with  $n$ -forms  $\omega_J$  and  $\mathbf{e}_I^k \wedge \star \mathbf{e}_J^k = \delta_{I,J}$ . Then, using the Leibniz rule (2.9), the projection of  $L_u \omega$  onto  $\mathbf{e}_I^k$  yields

$$\begin{aligned} L_u \omega \wedge \star \mathbf{e}_I^k &= L_u(\omega \wedge \star \mathbf{e}_I^k) + \omega \wedge \star L_u \mathbf{e}_I^k \\ &= d_i_u(\omega \wedge \star \mathbf{e}_I^k) + \omega \wedge \star L_u \mathbf{e}_I^k \\ &= d_i_u \omega_I + \sum_{J=1}^{\binom{n}{k}} \omega_J \mathbf{e}_J^k \wedge \star L_u \mathbf{e}_I^k. \end{aligned} \quad (2.10)$$

The conclusion follows using the correspondences from Table 2.2 for  $n$ -forms and  $v.p.(\omega) = \mathbf{w} = (w_1, \dots, w_{\binom{n}{k}})$ , so that the first term in (2.10) has vector proxy  $v.p.(d_i_u \omega_I) = \text{div}(\mathbf{u} w_I)$ .  $\square$

### Sobolev Spaces of Differential Forms

Having endowed the alternating forms with an inner product (2.5), it is possible to introduce an  $L^2$ -inner product of differential  $k$ -forms and define Sobolev spaces as completion of  $\Lambda^k(\Omega)$  in proper norms. More precisely, the  $L^2$ -inner product of differential  $k$ -forms on a domain  $\Omega \subset \mathbb{R}^n$  is defined as the integral of their pointwise inner product (2.5), namely,

$$(\omega, \eta)_{L^2 \Lambda^k(\Omega)} := \int_{\Omega} \langle \omega_x, \eta_x \rangle \text{vol} = \int_{\Omega} \omega \wedge \star \eta. \quad (2.11)$$

The completion of  $\Lambda^k(\Omega)$  in the norm  $\|\cdot\|_{L^2 \Lambda^k(\Omega)}^2 = (\cdot, \cdot)_{L^2 \Lambda^k(\Omega)}$  defines the space  $L^2 \Lambda^k(\Omega)$ . The Hilbert spaces  $H \Lambda^k(\Omega)$  are defined as  $H \Lambda^k(\Omega) = \{\omega \in L^2 \Lambda^k(\Omega), d^k \omega \in L^2 \Lambda^{k+1}(\Omega)\}$  endowed with the norm  $\|\cdot\|_{H \Lambda^k(\Omega)}^2 := \|\cdot\|_{L^2 \Lambda^k(\Omega)}^2 + \|d^k \cdot\|_{L^2 \Lambda^{k+1}(\Omega)}^2$ . Moreover, the spaces  $H^s \Lambda^k(\Omega)$  for  $s \geq 0$  consist of differential  $k$ -forms such that the maps  $x \in \Omega \mapsto \omega_x(v_1, \dots, v_k) \in \mathbb{R}$ ,  $(v_1, \dots, v_k) \in (T_x \Omega)^k$ , belong to  $H^s(\Omega)$ . These Sobolev spaces are Hilbert spaces.

Formal adjoint operators in the  $L^2$ -inner product for the contractions and the Lie derivative can be introduced as in the following definition.

**Definition 2.1.7.** Let  $\Omega \subset \mathbb{R}^n$  be a smooth, oriented  $n$ -dimensional manifold. Let  $u$  be a smooth vector field and let  $\omega \in \Lambda^k(\Omega)$ ,

- (i) the exterior co-derivative operator,  $\delta^k : \Lambda^k(\Omega) \rightarrow \Lambda^{k-1}(\Omega)$ , is defined as  $\star \delta^k \omega := (-1)^k d^{n-k} \star \omega$ ;
- (ii) the co-contraction,  $j_u : \Lambda^k(\Omega) \rightarrow \Lambda^{k+1}(\Omega)$ , is defined as  $\star j_u \omega := (-1)^k i_u \star \omega$ ;
- (iii) the Lie co-derivative,  $\mathcal{L}_u : \Lambda^k(\Omega) \rightarrow \Lambda^k(\Omega)$ , is defined as  $\mathcal{L}_u = -(-1)^{k(n-k)} \star L_u \star$ .

The exterior derivative and co-derivative define the Hodge Laplacian,  $\delta^{k+1} d^k + d^{k-1} \delta^k : \Lambda^k(\Omega) \rightarrow \Lambda^k(\Omega)$ .

A Leibniz rule with respect to the wedge product applies to the exterior derivative and to the contraction operator, namely, for all  $\omega \in \Lambda^k(\Omega)$  and  $\eta \in \Lambda^j(\Omega)$ ,

$$d^{k+n-j}(\omega \wedge \star \eta) = d^k \omega \wedge \star \eta + (-1)^{j+k} \omega \wedge \star \delta^j \eta, \quad (2.12)$$

$$i_u(\omega \wedge \star \eta) = i_u \omega \wedge \star \eta + (-1)^{j+k} \omega \wedge \star j_u \eta, \quad (2.13)$$

Note that (2.13) together with the  $L^2$ -inner product (2.11) yields  $(i_u \omega, \eta)_{L^2 \Lambda^k(\Omega)} = (\omega, j_u \eta)_{L^2 \Lambda^k(\Omega)}$  for all  $\omega \in \Lambda^{k+1}(\Omega)$  and  $\eta \in \Lambda^k(\Omega)$ . On the other hand, the Leibniz rule (2.12) combined with Stokes' law (2.3) gives the following integration by parts rules in the  $L^2$ -inner product (2.11),

$$(d^{k-1} \omega, \eta)_{L^2 \Lambda^k(\Omega)} = (\omega, \delta^k \eta)_{L^2 \Lambda^k(\Omega)} + \int_{\partial \Omega} \text{tr}(\omega \wedge \star \eta), \quad \forall \omega \in \Lambda^{k-1}(\Omega), \eta \in \Lambda^k(\Omega).$$

Thus  $\delta^k : \Lambda^{k+1}(\Omega) \rightarrow \Lambda^k(\Omega)$  is the formal adjoint of  $d^k : \Lambda^k(\Omega) \rightarrow \Lambda^{k+1}(\Omega)$  with respect to the  $L^2$ -inner product, if  $\omega$  or  $\eta$  vanishes near the boundary. As immediate consequence, analogous manipulations on (2.9) result in the integration by part rule for the Lie derivative

$$(L_u \omega, \eta)_{L^2 \Lambda^k(\Omega)} = (\omega, \mathcal{L}_u \eta)_{L^2 \Lambda^k(\Omega)} + \int_{\partial\Omega} \text{tr } i_u(\omega \wedge \star \eta), \quad \forall \omega, \eta \in \Lambda^k(\Omega). \quad (2.14)$$

By the theory of traces in Sobolev spaces, the trace operator  $\text{tr} : \Lambda^k(\Omega) \rightarrow \Lambda^k(\partial\Omega)$  introduced in (2.3) can be extended by continuity to a mapping of  $H^1 \Lambda^k(\Omega)$  onto the Sobolev space  $H^{1/2} \Lambda^k(\partial\Omega)$ : for  $\rho \in H^{1/2} \Lambda^k(\partial\Omega)$ , there exists  $\eta \in H^1 \Lambda^{n-1-k}(\Omega)$  whose trace  $\text{tr } \eta = \star_{n-1} \rho \in H^{1/2} \Lambda^{n-1-k}(\partial\Omega)$ , and  $\|\eta\|_{H^1 \Lambda^k(\Omega)} \leq C \|\star_{n-1} \rho\|_{H^{1/2} \Lambda^{n-1-k}(\partial\Omega)} \leq C \|\rho\|_{H^{1/2} \Lambda^k(\partial\Omega)}$ . The trace operator can further be extended to a bounded operator  $H \Lambda^k(\Omega) \rightarrow H^{-1/2} \Lambda^k(\partial\Omega)$ . Indeed, for  $\omega \in \Lambda^k(\Omega)$ , the integration by parts formula from (2.12) yields

$$\begin{aligned} (\text{tr } \omega, \star_{n-1} \rho)_{L^2 \Lambda^k(\partial\Omega)} &= \int_{\partial\Omega} \text{tr } \omega \wedge \text{tr } \eta = \int_{\Omega} d^k \omega \wedge \eta + (-1)^k \int_{\Omega} \omega \wedge d^{n-1-k} \eta \\ &\leq C \|\omega\|_{H \Lambda^k(\Omega)} \|\eta\|_{H^1 \Lambda^k(\Omega)} \leq C \|\omega\|_{H \Lambda^k(\Omega)} \|\rho\|_{H^{1/2} \Lambda^{n-1-k}(\partial\Omega)}. \end{aligned}$$

**Remark 2.1.8.** On subdomains of the Euclidean space, the Lie derivative associated with a smooth vector field  $u$  and its adjoint possess a rather extraordinary property: the differential operator  $L_u + \mathcal{L}_u$  can be recast as an operator involving only differentiation of the vector field  $u$ . Realizations of this attribute for vector proxies of differential forms in  $\mathbb{R}^3$  are summarized in Table 2.2. It will turn out that this property is of great importance for the well-posedness of transport problems involving the Lie derivative operator and for the derivation of numerical approximations. Indeed, in the context of Galerkin discretizations of boundary value problems in weak form, it allows to weaken the regularity requirements on the advected differential forms and deduce stability estimates by imposing constraints only on the velocity field. This comes at a price: sufficient differentiability of the velocity. We will dwell on the topic in Chapter 3.

| $\omega \in \Lambda^k(\Omega)$      | $k = 0$                           | $k = 1$  | $k = 2$  | $k = 3$                            |
|-------------------------------------|-----------------------------------|--|--|------------------------------------|
| $d^k \omega$                        | $\text{grad } w$                  | $\text{curl } \mathbf{w}$  | $\text{div } \mathbf{w}$   | —                                  |
| $i_u \omega$                        | —                                 | $\mathbf{u} \cdot \mathbf{w}$  | $\mathbf{w} \times \mathbf{u}$   | $w \mathbf{u}$                     |
| $\delta^k \omega$                   | —                                 | $-\text{div } \mathbf{w}$  | $\text{curl } \mathbf{w}$  | $-\text{grad } w$                  |
| $j_u \omega$                        | $w \mathbf{u}$                    | $-\mathbf{w} \times \mathbf{u}$  | $\mathbf{u} \cdot \mathbf{w}$  | —                                  |
| $L_u \omega$                        | $\mathbf{u} \cdot \text{grad } w$ | $\text{grad}(\mathbf{u} \cdot \mathbf{w}) + \text{curl } \mathbf{w} \times \mathbf{u}$ | $\text{curl}(\mathbf{w} \times \mathbf{u}) + \mathbf{u} \text{div } \mathbf{w}$        | $\text{div}(w \mathbf{u})$         |
| $\mathcal{L}_u \omega$              | $-\text{div}(w \mathbf{u})$       | $\text{curl}(\mathbf{u} \times \mathbf{w}) - \mathbf{u} \text{div } \mathbf{w}$        | $\mathbf{u} \times \text{curl } \mathbf{w} - \text{grad}(\mathbf{u} \cdot \mathbf{w})$ | $-\mathbf{u} \cdot \text{grad } w$ |
| $L_u \omega + \mathcal{L}_u \omega$ | $-w \text{div } \mathbf{u}$       | $(Du + (Du)^\top) \mathbf{w} - w \text{div } \mathbf{u}$                               | $w \text{div } \mathbf{u} - (Du + (Du)^\top) \mathbf{w}$                               | $w \text{div } \mathbf{u}$         |
| $\text{tr}$                         | $w$                               | $\mathbf{n} \times \mathbf{w}$   | $\mathbf{w} \cdot \mathbf{n}$  | —                                  |
| $H \Lambda^k(\Omega)$               | $H^1(\Omega)$                     | $H(\text{curl}, \Omega)$   | $H(\text{div}, \Omega)$  | $L^2(\Omega)$                      |

Table 2.2.: Exterior calculus notations and corresponding expressions for vector proxies. For details see [HH13a, Table 2], [Hip02, Section 2.2], [AFW06, Table 2.2].

### De Rham Complex

A cochain complex is a sequence  $\{V_k, d_k\}_k$  of Abelian groups  $V_k$  and homomorphisms  $d_k : V_k \rightarrow V_{k+1}$  such that  $d_{k+1} \circ d_k = 0$  (*cf.* Section 2.3). The fact that the exterior derivative from Definition 2.1.4 is a differential means that the range of the exterior derivative of differential  $(k-1)$ -forms is contained in the kernel of the exterior derivative of differential  $k$ -forms, namely  $\mathcal{R}(d^{k-1}) \subset \mathcal{N}(d^k)$ . Hence, the spaces  $\Lambda^k(\Omega)$  on the domain  $\Omega \subset \mathbb{R}^n$  together with the graded operator  $d^k$  form the sequence

$$0 \longrightarrow \Lambda^0(\Omega) \xrightarrow{d^0} \Lambda^1(\Omega) \xrightarrow{d^1} \dots \xrightarrow{d^{n-1}} \Lambda^n(\Omega) \longrightarrow 0, \quad (2.15)$$

called *de Rham complex*. The de Rham sequence extends to the so-called  $L^2$  de Rham complex,

$$0 \longrightarrow H\Lambda^0(\Omega) \xrightarrow{d^0} H\Lambda^1(\Omega) \xrightarrow{d^1} \dots \xrightarrow{d^{n-1}} H\Lambda^n(\Omega) \longrightarrow 0. \quad (2.16)$$

A de Rham complex can also be constructed for the  $H\Lambda^k(\Omega)$  spaces with boundary conditions  $\overset{\circ}{H}\Lambda^k(\Omega) := \{\omega \in H\Lambda^k(\Omega) : \text{tr } \omega = 0\}$ . The quotient spaces  $\mathcal{CH}^k(\Omega) = \mathcal{N}(d^k)/\mathcal{R}(d^{k-1})$ ,  $1 \leq k \leq n$  and  $\mathcal{CH}^0(\Omega) = \mathcal{N}(d^0)$  form the de Rham cohomology. They are finite dimensional vector spaces with dimensions equal to the Betti numbers of the manifold  $\Omega$ . The sequence is *exact* if the cohomology is trivial, namely if  $\mathcal{N}(d^k) = \mathcal{R}(d^{k-1})$  for all  $1 \leq k \leq n$ . The inclusion  $\mathcal{R}(d^{k-1}) \subset \mathcal{N}(d^k)$  corresponds to the cochain complex property, whilst the reverse inclusion  $\mathcal{N}(d^k) \subset \mathcal{R}(d^{k-1})$  entails that the cohomology groups of the cochain complex are trivial. The cohomology groups measure ‘‘how far’’ the complex is from being exact. Indeed, a differential  $k$ -forms  $\omega \in \Lambda^k(\Omega)$  is closed, i.e.,  $d^k\omega = 0$  if and only if  $\omega = d^{k-1}\eta + \tau$  with  $\eta \in \Lambda^{k-1}(\Omega)$  and  $\tau$  is a representative of the  $k$ -cohomology group  $\mathcal{CH}^k(\Omega)$ .

### Tensor Product of Differential Forms

For the purpose of realizing discrete differential forms to approximate quantities on Cartesian meshes, we need to introduce the tensor product of differential forms. The gist of this construction is to apply the exterior calculus structure to the Cartesian product of domains as in [ABB15, Section 3]. More in details, the projections  $\pi_j : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^j$ , with  $j \in \{m, n\}$  and  $n, m \geq 1$ , induce by pullback the injective maps  $\pi_m^* : \text{Alt}^k \mathbb{R}^m \rightarrow \text{Alt}^k \mathbb{R}^{m+n}$  and  $\pi_n^* : \text{Alt}^\ell \mathbb{R}^n \rightarrow \text{Alt}^\ell \mathbb{R}^{m+n}$ . The linear map

$$\begin{aligned} \text{Alt}^k \mathbb{R}^m \otimes \text{Alt}^\ell \mathbb{R}^n &\longrightarrow \text{Alt}^{k+\ell} \mathbb{R}^{m+n} \\ \omega \otimes \eta &\longmapsto \pi_m^* \omega \wedge \pi_n^* \eta, \end{aligned}$$

is injective. Hence, the tensor product of algebraic  $k$ -forms on  $\mathbb{R}^m$  and  $\ell$ -forms on  $\mathbb{R}^n$  is a  $(k + \ell)$ -form on  $\mathbb{R}^{m+n}$ . On differential forms one can proceed analogously. Let  $\omega \in \Lambda^k(\Omega_m)$  on the domain  $\Omega_m \subset \mathbb{R}^m$  and let  $\eta \in \Lambda^\ell(\Omega_n)$  on  $\Omega_n \subset \mathbb{R}^n$ . Then, the tensor product of  $\omega$  and  $\eta$  is defined as  $\omega \otimes \eta := \pi_m^* \omega \wedge \pi_n^* \eta \in \Lambda^{k+\ell}(\Omega_m \times \Omega_n)$ , where  $\pi_m : \Omega_m \times \Omega_n \rightarrow \Omega_m$  and  $\pi_n : \Omega_m \times \Omega_n \rightarrow \Omega_n$  are the canonical projections. Note that the exterior derivative on the tensor product is  $d^{k+\ell}(\omega \otimes \eta) = \pi_m^*(d^k \omega) \wedge \pi_n^* \eta + (-1)^k \pi_m^* \omega \wedge \pi_n^*(d^\ell \eta)$ .

Similarly, it is possible to construct tensor products of subcomplexes of a given complex as subcomplex of the complex on the Cartesian product of the domains. Let  $\{M^k, d^k\}_{k=0}^m$  and  $\{N^\ell, d^\ell\}_{\ell=0}^n$  be subcomplexes of the de Rham complex on  $\Omega_m$  and  $\Omega_n$ , respectively. Let us introduce the space  $(M \otimes N)^j := \bigoplus_{k+\ell=j} (M^k \otimes N^\ell)$ , for  $0 \leq j \leq m+n$ , and the differential  $d^j(\omega \otimes \eta) = d^k \omega \wedge \eta + (-1)^k \omega \wedge d^\ell \eta$  with  $\omega \in M^k$  and  $\eta \in N^\ell$ . The space  $(M \otimes N)^j$  consists of differential  $j$ -forms on  $\Omega_m \times \Omega_n$  and the differential  $d$  is the restriction of the exterior derivative to  $H\Lambda^k(\Omega_m \times \Omega_n)$ . Hence,  $\{(M \otimes N)^j, d^j\}_{j=0}^{m+n}$  is a subcomplex of the de Rham complex (2.16) on  $\Omega_m \times \Omega_n$ .

## 2.2. Generalized Advection-Diffusion Evolution Problem

In the language of exterior calculus, Maxwell’s equations in the eddy current model in moving conductors can be cast as,

$$\begin{array}{lll} d^1 e + \partial_t b = 0 & (\text{Faraday's law}) & d^2 b = 0 & (\text{Gauss' law}) \\ d^1 h = j & (\text{Ampère's law}) & j = \sigma \star (e - i_u b) & (\text{Ohm's law}) \end{array}$$

where  $e \in \Lambda^1(\Omega)$  is the electric field,  $h \in \Lambda^1(\Omega)$  the magnetic field and  $b \in \Lambda^2(\Omega)$  the magnetic induction over a domain  $\Omega \subset \mathbb{R}^3$  with trivial topology;  $\sigma > 0$  is the conductivity parameter and  $\mathbf{u}$  the velocity. The generalized Ohm law leads to  $d^1(\sigma^{-1} \star d^1 h) = d^1(e - i_u b) = -d^1(\partial_t a) - d^1 i_u d^1 a$  where Faraday’s law has been recast in terms of the magnetic vector potential  $a \in \Lambda^1(\Omega)$  with  $b = d^1 a$ . Introducing a scalar *gauge potential*  $\phi \in \Lambda^0(\Omega)$  results in  $\sigma^{-1} \star d^1 h = -\partial_t a - d^0 \phi - i_u d^1 a$ . Manipulating the material

law  $b = \mu \star h$ , using the definition of exterior co-derivative and the idempotence of the Hodge operator, yields  $\delta^2 \mu^{-1} d^1 a = \star d^1 h$ . A gauge of the form  $\phi = i_{\mathbf{u}} a$ , results in

$$\partial_t a + \sigma^{-1} \delta^2 \mu^{-1} d^1 a + i_{\mathbf{u}} d^1 a + d^0 i_{\mathbf{u}} a = 0,$$

where suitable initial and boundary conditions have to be supplied (refer to [Heu11, Section 3.2] for further details on the derivation).

More generally, for a time-dependent differential  $k$ -form  $\omega(t)$  on the bounded domain  $\Omega \subset \mathbb{R}^n$ , the strong form of the generalized advection-diffusion equation in the space-time domain  $\Omega \times I$ ,  $I := [0, T]$ , reads

$$\begin{aligned} \star \partial_t \omega(t) + (-1)^{k+1} d^{n-k-1} \varepsilon \star d^k \omega(t) + \star \alpha \omega(t) + \star L_{\mathbf{u}} \omega(t) &= f(t), & \text{in } \Omega \times I, \\ \text{tr } \omega(t) &= \text{tr } g(t), & \text{on } (\Gamma_{\text{in}} \cup \Gamma_0) \times I, \\ \text{tr}(i_{\mathbf{n}} \omega(t)) &= \text{tr } s(t), & \text{on } \Gamma_{\text{in}} \times I, \\ \omega(0) &= \omega_0, & \text{in } \Omega, \end{aligned} \quad (2.17)$$

where  $\mathbf{u} : \bar{\Omega} \times I \rightarrow \mathbb{R}^n$  is a *given* velocity field,  $f(t) \in \Lambda^{n-k}(\Omega)$  a source term, and a possibly vanishing reaction term  $\star \alpha \omega$  has been added. The scalar diffusivity parameter  $\varepsilon$  and the reaction coefficient  $\alpha$  are non-negative and bounded functions  $\Omega \rightarrow \mathbb{R}$ , and the boundary conditions are imposed at the inflow boundary  $\Gamma_{\text{in}} := \{\mathbf{x} \in \partial\Omega : \mathbf{u} \cdot \mathbf{n}(\mathbf{x}) < 0\}$  and at the “elliptic boundary” (where the diffusion parameter  $\varepsilon > 0$ ).

We can now preach the connection between (2.1), (2.2) and (2.17) through the vector proxies of differential 1- and 0-forms, respectively (*cf.* Tables 2.1 and 2.2). In the generalized advection-diffusion problem (2.17), the diffusion operator is  $d^{n-k-1} \star d^k$ , the zeroth-order term amounts to a reaction term and the advection operator is the Lie derivative  $L_{\mathbf{u}}$  associated with the velocity field  $\mathbf{u}$ .

### 2.2.1. Structure-Preserving Discretizations

The popularity of numerical schemes based on discrete analogs of differential forms (called *discrete differential forms*, see Section 2.3) started on the premise that spurious solutions resulting from standard numerical methods often originate from inconsistent discretizations of the physical fields and of the differential operators involved. Formulating numerical approximations *compatible* with the geometric and topological structures underlying the continuum model problem and not just approximating them has paved the way for numerical schemes with superior stability properties and yielding physically consistent solutions.

Precursors of discrete differential forms appeared in the geometry literature in the work of Whitney [Whi57] and were brought to the finite element community first by Baker [Bak83] and then in [Bos88] and [Kot84]. We refer to Section 2.4.2 for a thorough historical overview on finite element spaces of differential forms.

The potential of discrete differential forms has been harnessed, among others, by mimetic finite difference methods [HS99], discrete exterior calculus [Des+05] and Finite Element Exterior Calculus [AFW06]. The latter will be our preferred approach in dealing with numerical approximations of (2.17). Finite element exterior calculus aims at constructing finite element spaces incarnating finite dimensional subcomplexes of certain differential complexes valid at the continuous level (in our case the de Rham complex (2.16)). By means of commuting projections the discrete fields inherit the topological and algebraic structure of the continuum setting.

## 2.3. Discrete Differential Forms

To develop a discrete theory, the exterior calculus construction along with operators involving differential forms has to be established on discrete manifolds. Since the numerical discretizations we are interested in are based on tessellations of a physical domain, the discrete manifolds we look at are cellular complexes.

In order to keep the presentation focused, we restrict our attention to simplicial complexes, however the whole discourse extends to the more general structure provided by CW-complexes. In particular, we introduce the concept of cell complex and the construction of topological objects over such complexes. More details on chain and cochain structures can be found in [GK01], [Des+05] and in compendia on algebraic topology, e.g. [Hat02].

### Chains

The *standard  $k$ -simplex* is the non-degenerate convex hull of the unit vectors  $\{e_i\}_{i=1}^{k+1} \subset \mathbb{R}^n$ ,  $n > k$ , along the coordinates axes, namely

$$\Delta^k = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{k+1} t_i e_i \text{ with } t_i \geq 0, \sum_{i=1}^{k+1} t_i = 1 \right\}.$$

Given  $k + 1$  geometrically distinct points  $\{v_i\}_{i=1}^{k+1} \subset \mathbb{R}^n$ ,  $n \geq k$ , a (singular)  $k$ -simplex  $\sigma^k$  is the map

$$\begin{aligned} \sigma^k : \Delta^k &\longrightarrow \mathbb{R}^n \\ \sum_{i=1}^{k+1} t_i e_i &\longmapsto \sum_{i=1}^{k+1} t_i v_i. \end{aligned}$$

The map is affine and its image is the convex hull of  $\{v_i\}_{i=1}^{k+1}$ . The simplex  $\sigma^k$  is denoted by  $[v_1, \dots, v_{k+1}]$  and the points  $\{v_i\}_{i=1}^{k+1}$  are called the vertices of the simplex. The convex hull of a subset of  $\{v_i\}_{i=1}^{k+1}$  is a simplex itself and it is called a face of  $\sigma^k$ . The  $i$ -th face map  $F_k^{(i)} : \Delta^{k-1} \rightarrow \Delta^k$  is defined as  $F_k^{(i)}([t_1, \dots, t_k]) = [t_1, \dots, t_{i-1}, 0, t_i, \dots, t_k]$ .

A *simplicial complex*  $K$  in  $\mathbb{R}^n$  is a collection of simplices in  $\mathbb{R}^n$  such that: every face of a simplex of  $K$  is in  $K$ , and the intersection of any two simplices of  $K$  is a face of each of them. A *simplicial triangulation* of a polytope is a simplicial complex  $K$  such that the union of the simplices of  $K$  recovers the polytope. A *subcomplex* of  $K$  is the simplicial complex formed by a subcollection of simplices of  $K$  containing all faces of its elements. The subcomplex of  $K$  containing all simplices of  $K$  of dimension at most  $k$  is called the  $k$ -*skeleton* of  $K$ .

Any  $n$ -simplex is endowed with an ordering of its vertices  $[v_1, \dots, v_{n+1}]$ . A by-product of ordering the vertices of simplices is that this determines an (internal) orientation of any  $\ell$ -face, described by an ordered  $(\ell + 1)$ -tuple of vertices, up to even permutations of the vertices. Specifically, the vertices of any  $\ell$ -simplex spanned by a (sub)set of vertices, inherit an ordering according to their order in the supporting subspace (i.e., the larger simplex containing the  $\ell$ -simplex). In this way, an orientation of an  $\ell$ -simplex induces an orientation of its boundary. Inner orientation is intrinsic and does not depend on the  $\ell$ -simplex being embedded in a larger space. An external orientation can instead be supplied by giving a “crossing direction” to a surface, namely by providing the surface with a continuous field of normals. An outer orientation of the tangent space at a point is, by definition, an inner orientation of its complement. Hence, if the encompassing space is oriented, outer and inner orientations are dual to each other. Cells which enclose a volume can be outer oriented, through the “inside out” crossing direction. Then, crossing the cell boundary from the interior to the exterior of the cell provides an induced external orientation of the cell boundary. We refer to [Hip02, Section 3.1] and [Bos98a, Section 5.2.1] for further details on simplices orientation.

**Definition 2.3.1** (Chain). Let  $K$  be a simplicial complex. A  $k$ -chain  $c_k$  is a formal (finite) linear combination of  $k$ -simplices in  $K$  with coefficients  $n_{\sigma^k} \in \mathbb{Z}$ ,  $c_k = \sum_{\sigma^k \in K} n_{\sigma^k} \sigma^k$ . The  $k$ -chains form a free abelian group  $\mathcal{C}_k(K; \mathbb{Z})$  generated by a basis consisting of oriented  $k$ -simplices.

The (algebraic) boundary of a  $k$ -simplex  $\sigma^k$  is the  $(k - 1)$ -chain obtained by omitting the vertex  $v_i$ , i.e.,  $\partial_k \sigma^k = \sum_{i=1}^k (-1)^{i+1} \sigma^k \circ F_k^{(i)}$ , where  $F_k^{(i)}$  is the  $i$ -th face map. The boundary operator can be extended linearly, as in the following definition.

**Definition 2.3.2** (Boundary operator). The *boundary operator*  $\partial_k : \mathcal{C}_k(K; \mathbb{Z}) \rightarrow \mathcal{C}_{k-1}(K; \mathbb{Z})$  is a homomorphism defined through its action on a simplex  $\sigma^k$  as

$$\partial_k \left( \sum_{\sigma^k \in K} n_{\sigma^k} \sigma^k \right) = \sum_{\sigma^k \in K} n_{\sigma^k} \partial_k \sigma^k.$$

The boundary operator and the chain group form a complex: On an  $n$ -dimensional simplicial complex, a *chain complex* is a collection of chain groups and homomorphisms  $\partial_k$  such that,

$$0 \longrightarrow \mathcal{C}_n(K; \mathbb{Z}) \xrightarrow{\partial_n} \dots \xrightarrow{\partial_{k+1}} \mathcal{C}_k(K; \mathbb{Z}) \xrightarrow{\partial_k} \dots \xrightarrow{\partial_1} \mathcal{C}_0(K; \mathbb{Z}) \longrightarrow 0$$

and  $\partial_k \circ \partial_{k+1} = 0$ . This implies that  $\mathcal{R}(\partial_{n+1}) \subset \mathcal{N}(\partial_n)$ . The range of  $\partial_{n+1}$ ,  $\mathcal{R}(\partial_{n+1})$ , is a subgroup of  $\mathcal{C}_k(K; \mathbb{Z})$  and contains the so-called  $(n+1)$ -*boundaries*. The null-space  $\mathcal{N}(\partial_n)$  is the subgroup of  $n$ -*cycles*. Boundaries are cycles. On the other hand, functions which are  $n$ -cycles but not  $(n+1)$ -boundaries are characterized by the quotient group  $\mathcal{H}_n(K; \mathbb{R}) := \mathcal{N}(\partial_n)/\mathcal{R}(\partial_{n+1})$ , the  $n$ -th *homology group* of the chain complex. The rank of  $\mathcal{H}_n(K; \mathbb{R})$  is the number of its independent equivalence classes ( $n$ -cycles which differ by a  $(n+1)$ -boundary) and it is called  $n$ -th Betti number of  $K$  (see e.g. [Hat02, p. 130]).

## Cochains

Analogously to integral  $k$ -forms defined with reference to  $k$ -manifolds, we introduce homomorphisms over chains, the so-called cochains. Cochains provide the scaffolding to build discrete differential forms.

**Definition 2.3.3** (Cochain). Let  $K$  be a simplicial complex. A  $k$ -*cochain* is a homomorphism from the chain group  $\mathcal{C}_k(K; \mathbb{Z})$  to the additive group  $\mathbb{R}$ . The space of cochains  $\mathcal{C}^k(K; \mathbb{R})$  has an abelian group structure with the addition of two homomorphisms corresponding to the addition of their values in  $\mathbb{R}$ .

Since a chain is a linear combination of simplices, the value assigned by a cochain to a given chain is the sum of the values on each simplex. Each homomorphism can be regarded as a functional on chains and the action of a cochain  $c^k \in \mathcal{C}^k(K; \mathbb{R})$  on a chain  $c_k \in \mathcal{C}_k(K; \mathbb{Z})$  can be expressed as *chain-cochain pairing*

$$\langle c^k, c_k \rangle := c^k(c_k). \quad (2.18)$$

Note that this is tantamount to integration of integral forms on chains in the smooth manifolds framework. An integral  $k$ -form determines a  $k$ -cochain through the so-called *de Rham maps*:

$$\begin{aligned} \mathbf{l}_k : \mathcal{F}^k(K) &\longrightarrow \mathcal{C}^k(K; \mathbb{R}) \\ \omega &\longmapsto (\mathbf{l}_k(\omega))(F) = \int_F \omega \quad \forall F \in \mathcal{S}_k(K). \end{aligned}$$

It is also possible to devise cochain analogs of the differential operators acting on integral forms. In greater detail, a discrete exterior derivative can be characterized through the *coboundary operator*  $\mathfrak{d}^k : \mathcal{C}^k(K; \mathbb{R}) \rightarrow \mathcal{C}^{k+1}(K; \mathbb{R})$  defined by duality to the boundary operator  $\partial_k$  with respect to the chain-cochain pairing. For a  $k$ -cochain  $c^k \in \mathcal{C}^k(K; \mathbb{R})$ , and a chain  $c_{k+1} \in \mathcal{C}_{k+1}(K; \mathbb{Z})$ ,

$$\langle \mathfrak{d}^k c^k, c_{k+1} \rangle := \langle c^k, \partial_{k+1} c_{k+1} \rangle. \quad (2.19)$$

The coboundary operator induces the cochain complex

$$0 \longleftarrow \mathcal{C}^n(K; \mathbb{R}) \xleftarrow{\mathfrak{d}^{n-1}} \dots \xleftarrow{\mathfrak{d}^k} \mathcal{C}^k(K; \mathbb{R}) \xleftarrow{\mathfrak{d}^{k-1}} \dots \xleftarrow{\mathfrak{d}^0} \mathcal{C}^0(K; \mathbb{R}) \longleftarrow 0,$$

where  $\mathfrak{d}^{k+1} \circ \mathfrak{d}^k = 0$ . Analogously to the chain complex, we can introduce the  $n$ -th cohomology group  $\mathcal{CH}^n(K; \mathbb{R}) := \mathcal{N}(\mathfrak{d}^{n+1})/\mathcal{R}(\mathfrak{d}^n)$ . Note that the coboundary operator on cochains is linear and the associated matrix, the so-called *incidence matrix* of  $\ell$ -faces and  $(\ell+1)$ -faces, belongs to  $\mathbb{C}^{N_{\ell+1}, N_\ell}$  and is sparse with values in  $\{-1, 0, 1\}$  determined by the adjacency relation and relative orientations. The incidence matrix plays a pure combinatorial algebraic role containing however topological information about the discretized domain.

**Theorem 2.3.4** (De Rham theorem for cochains). Let  $\mathcal{T}_h$  be a cellular complex. For every  $c^k \in \mathcal{C}^k(\mathcal{T}_h; \mathbb{R})$ ,  $\mathfrak{d}^k c^k = 0$  if and only if there exist  $c^{k-1} \in \mathcal{C}^{k-1}(\mathcal{T}_h; \mathbb{R})$  and  $a^k \in \mathcal{CH}^k(\mathcal{T}_h; \mathbb{R})$  such that  $c^k = \mathfrak{d}^{k-1} c^{k-1} + a^k$ .

In order to define discrete differential  $k$ -forms we introduce the so-called *Whitney maps*: bijective linear operators  $W_k$  mapping  $k$ -cochain to differential  $k$ -forms defined almost everywhere (see [Hip02, Section 3.2] for further details).

**Definition 2.3.5.** The space of *discrete differential  $k$ -forms*  $\Lambda_h^k(\mathcal{T}_h)$  on a cellular complex  $\mathcal{T}_h$  is the range of  $k$ -cochains  $\mathcal{C}^k(\mathcal{T}_h; \mathbb{R})$  under the bijective linear map  $W_k$  satisfying the following properties:

1. Discrete differential forms are extensions of cochains, in the sense that  $I_k \circ W_k = \text{Id}$ ;
2. The exterior derivatives of cochains and of discrete differential forms commute, namely

$$\begin{array}{ccc} \mathcal{C}^k(\mathcal{T}_h; \mathbb{R}) & \xrightarrow{W_k} & \Lambda_h^k(\mathcal{T}_h) \\ \downarrow \mathfrak{d}^k & & \downarrow \mathfrak{d}^k \\ \mathcal{C}^{k+1}(\mathcal{T}_h; \mathbb{R}) & \xrightarrow{W_{k+1}} & \Lambda_h^{k+1}(\mathcal{T}_h). \end{array}$$

With the notion of exterior derivative, and the relationship between the natural pairing and integration, one can regard (2.19) as a discrete generalized Stokes' theorem. Thus, given a  $(k+1)$ -chain  $c_{k+1}$ , and a discrete  $k$ -form  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$ , the discrete Stokes' theorem states that

$$\langle \mathfrak{d}^k \omega_h, c_{k+1} \rangle = \langle \omega_h, \partial_{k+1} c_{k+1} \rangle. \quad (2.20)$$

It immediately follows that  $\mathfrak{d}^{k+1} \circ \mathfrak{d}^k = 0$ .

### Discrete Hodge Operator and Discrete Contraction

Discrete exterior calculus emerged as a natural tool in the development of numerical methods for electromagnetic problems. Indeed, it boasts the ability of maintaining the separation between topological and geometrical quantities yielding a discrete picture which partakes of the properties of the continuum models. However, not all the operators involved in the electromagnetic field laws and borrowed from differential geometry have straightforward discrete combinatorial analogs as the exterior derivative.

The numerical treatment of the material laws, e.g. (2.6), demands for discrete Hodge operators. Resting on the observation that the constitutive laws do not constrain discrete 2-form fields and 1-form fields to be constructed on the same cellular complex, discrete versions of the Hodge operators can be designed via linear mappings between suitable finite dimensional spaces of discrete differential forms on possibly different and unrelated meshes (under the requirements derived in [Hip01, Section 4]). Galerkin schemes tailored to the weak formulation of the material laws can then be implemented. On the other hand, choosing pairs of *dual meshes* fostered alternative discretizations of the Hodge operator closely related to staggered finite volume approaches à la Yee [Yee66]. Different specializations of discrete Hodge operators turned out to translate in various discrete schemes, as pointed out in [Hip02, p. 268]. The link between various finite element and finite volume numerical schemes and a unified abstract error analysis of the discretization error have been established in [Hip01].

On a related note, the missing dowel in the discretization of the Lie derivative is a suitable description of the discrete contraction operator. The lack of smoothness of discrete differential forms does not allow to define a discrete contraction operator as the restriction to discrete spaces. Indeed, let  $c^k$  be a given  $k$ -chain on the cellular complex  $\mathcal{T}_h$  and let  $\omega_h \in \Lambda_h^{k+1}(\mathcal{T}_h)$ , then  $\langle i_u \omega_h, c^k \rangle \neq \langle -i_{-u} \omega_h, c^k \rangle$ . As suggested by Bossavit in [Bos03], the notion of extrusion of a manifold, see Definition 4.1.1, provides a dual object to the contraction operator which is amenable to a discrete interpretation. We will delve into this type of approach later in Chapter 4. An alternative technique was introduced in [Hir03, Section 8.2], where discrete definitions of the Hodge operator and of the wedge product are combined into a discrete contraction through the formula  $i_u \omega = (-1)^{k(n-k)} \star (\star \omega \wedge u^\flat)$ ,  $\omega \in \Lambda^k(\Omega)$ , where the flat operator  $\flat$  maps vector fields to 1-forms [Hir03, Definition 5.2.2].

## 2.4. Polynomial Differential Forms

Finite element methods for the discretization of initial boundary value problems are grounded in the definition of finite dimensional subspaces of the continuous function spaces. Polynomial finite element spaces are particularly attractive for analysis and implementation. In the present section we introduce spaces of polynomial differential forms and use them to build finite element spaces over simplicial and tensor product cellular complexes.

Let  $\mathcal{P}_r(\mathbb{R}^n)$  denote the space of polynomials of degree at most  $r \geq 0$  in  $n$  variables and let  $\mathcal{H}_r(\mathbb{R}^n)$  be the space of homogeneous polynomials of degree  $r$ . The space of polynomial differential  $k$ -forms  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n)$  of degree at most  $r \geq 0$  contains  $\omega \in \Lambda^k(\mathbb{R}^n)$  such that the map  $x \in \mathbb{R}^n \mapsto \omega_x \in \text{Alt}^k \mathbb{R}^n$  has polynomial coefficients of degree at most  $r$ , and has dimension  $\dim \mathcal{P}_r\Lambda^k(\mathbb{R}^n) = \dim \mathcal{P}_r(\mathbb{R}^n) \dim \text{Alt}^k \mathbb{R}^n = \binom{r+k}{r} \binom{n+r}{n-k}$ . Another family  $\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$  of polynomial differential forms intermediate between polynomial spaces of subsequent degrees generalizes Whitney forms. A unified analysis of these spaces was carried out in [Hip99b] and later proposed in [AFW06, Section 3] in a different perspective through the Koszul complex. The Koszul differential  $\kappa : \Lambda^k(\mathbb{R}^n) \rightarrow \Lambda^{k-1}(\mathbb{R}^n)$  is an instance of the contraction operator (Definition 2.1.5), based on the identification of  $\mathbb{R}^n$  with the tangent space  $T_0\mathbb{R}^n$  at the origin. More in details, we can associate to each  $x \in \mathbb{R}^n$  a vector in  $T_0\mathbb{R}^n$  and an element  $X(x) \in T_x\mathbb{R}^n$  through the isomorphism  $T_0\mathbb{R}^n \rightarrow T_x\mathbb{R}^n$  induced by the translation map  $y \mapsto y + x$ . The Koszul differential is the contraction with the vector  $X(x)$ , namely

$$(\kappa\omega)_x(v_1, \dots, v_{k-1}) := \omega_x(X(x), v_1, \dots, v_{k-1}) \quad \forall \omega \in \Lambda^k(\mathbb{R}^n), \text{ and with } \{v_i\}_{i=1}^{k-1} \subset \mathbb{R}^n.$$

The operator  $\kappa$  is a graded differential and inherits the properties of the contraction operator:  $\kappa \circ \kappa = 0$ , the Leibniz rule  $\kappa(\omega \wedge \eta) = (\kappa\omega) \wedge \eta + (-1)^k \omega \wedge (\kappa\eta)$  for  $\omega \in \Lambda^k(\mathbb{R}^n)$ ,  $\eta \in \Lambda^j(\mathbb{R}^n)$ , and  $\phi^*(\kappa\omega) = \kappa(\phi^*\omega)$  for all  $\omega \in \Lambda^k(\mathbb{R}^n)$  and  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear or affine. Note that, in defining the Koszul differential, the choice of the origin as a base point in  $\mathbb{R}^n$  is arbitrary: if  $z \in \mathbb{R}^n$ , the vector  $X(x) = X_z(x)$  can be defined by assigning to each point  $x$  the translation  $y \mapsto y + x$  of the vector pointing from  $z$  to  $x$ .

A relation between the Koszul differential on homogeneous polynomial differential forms  $\kappa : \mathcal{H}_r\Lambda^k(\mathbb{R}^n) \rightarrow \mathcal{H}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$  and the exterior derivative was derived in [AFW06, Theorem 3.1] and reads  $(d^{k-1}\kappa + \kappa d^k)\omega = (r+k)\omega$  for all  $\omega \in \mathcal{H}_r\Lambda^k(\mathbb{R}^n)$ . It entails the following important properties:

- (i) (*Injectivity.*) If  $\omega \in \mathcal{H}_r\Lambda^k(\mathbb{R}^n)$ ,  $d^{k-1}\kappa\omega = 0$  implies  $\kappa\omega = 0$  and  $\kappa d^k\omega = 0$  implies  $d^k\omega = 0$ .
- (ii) (*Exact sequence property.*) The homogeneous subcomplex of the de Rham complex

$$\mathbb{R} \hookrightarrow \mathcal{H}_r\Lambda^0(\mathbb{R}^n) \xrightarrow{d^0} \mathcal{H}_{r-1}\Lambda^1(\mathbb{R}^n) \xrightarrow{d^1} \dots \xrightarrow{d^{n-1}} \mathcal{H}_{r-n}\Lambda^n(\mathbb{R}^n) \longrightarrow 0 \quad (2.21)$$

is exact for any  $r \geq 0$ . Indeed, if  $\omega \in \mathcal{H}_r\Lambda^k(\mathbb{R}^n)$  is a closed form, then  $(d^{k-1}\kappa + \kappa d^k)\omega = d^{k-1}\kappa\omega = (r+k)\omega$ . Hence,  $\omega = d^{k-1}\kappa\omega/(r+k) =: d^{k-1}\eta$  with  $\eta \in \mathcal{H}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$ . Moreover, since  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n) = \bigoplus_{j=0}^r \mathcal{H}_j\Lambda^k(\mathbb{R}^n)$ , we immediately get a polynomial subcomplex for  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n)$  analogous to (2.21).

- (iii) (*Direct sum decomposition.*) Let  $r+k > 0$ , then  $\mathcal{H}_r\Lambda^k(\mathbb{R}^n) = \kappa\mathcal{H}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) \oplus d^{k-1}\mathcal{H}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$ . Indeed, if  $\omega \in \mathcal{H}_r\Lambda^k(\mathbb{R}^n)$ , then  $\omega = (d^{k-1}\kappa\omega + \kappa d^k\omega)/(r+k) =: d^{k-1}\mu + \kappa\eta$  with  $\mu \in \mathcal{H}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$  and  $\eta \in \mathcal{H}_{r-1}\Lambda^{k+1}(\mathbb{R}^n)$ . The sum is direct since, for  $\omega \in \kappa\mathcal{H}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) \cap d^{k-1}\mathcal{H}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$ ,  $d^k\omega = \kappa\omega = 0$ , owing to  $d^{k+1} \circ d^k = \kappa \circ \kappa = 0$ . Thus,  $\omega = 0$ . Similarly, it holds  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n) = \kappa\mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) \oplus d^{k-1}\mathcal{P}_{r+1}\Lambda^{k-1}(\mathbb{R}^n)$ .

The decomposition of  $\mathcal{H}_r\Lambda^k(\mathbb{R}^n)$  prompts the definition of a polynomial space intermediate between  $\mathcal{P}_{r-1}\Lambda^k(\mathbb{R}^n)$  and  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n)$ , namely

$$\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) := \mathcal{P}_{r-1}\Lambda^k(\mathbb{R}^n) \oplus \kappa\mathcal{H}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) = \mathcal{P}_{r-1}\Lambda^k(\mathbb{R}^n) + \kappa\mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n).$$

Observe that  $\mathcal{P}_r^-\Lambda^0(\mathbb{R}^n) = \mathcal{P}_r\Lambda^0(\mathbb{R}^n)$ , and  $\mathcal{P}_r^-\Lambda^n(\mathbb{R}^n) = \mathcal{P}_{r-1}\Lambda^n(\mathbb{R}^n)$ , whilst  $\mathcal{P}_{r-1}\Lambda^k(\mathbb{R}^n) \subsetneq \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) \subsetneq \mathcal{P}_r\Lambda^k(\mathbb{R}^n)$  for  $0 < k < n$ . It holds,  $\dim \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) = \frac{r}{r+k} \dim \mathcal{P}_r\Lambda^k(\mathbb{R}^n)$ . Moreover, any closed  $k$ -form  $\omega \in \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$  belongs to  $\mathcal{P}_{r-1}\Lambda^k(\mathbb{R}^n)$  (see [AFW06, Theorem 3.4]). Since the base point  $z \in \mathbb{R}^n$  used to define the Koszul differential was arbitrary, the definition of  $\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$  is independent of  $z$ . Hence,

it is possible to define the space  $\mathcal{P}_r^-\Lambda^k(V)$  on any affine subspace  $V \subset \mathbb{R}^n$ . We will sometimes refer to  $\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$  and  $\mathcal{P}_r\Lambda^k(\mathbb{R}^n)$  as the “first family” and “second family” of polynomial differential forms, respectively.

It would be desirable to construct long exact sequences for the intermediate space  $\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$ . In fact, it can be shown [AFW06, Section 3.5] that the complex

$$\mathbb{R} \hookrightarrow \mathcal{P}_r^-\Lambda^0(\mathbb{R}^n) \xrightarrow{d^0} \mathcal{P}_r^-\Lambda^1(\mathbb{R}^n) \xrightarrow{d^1} \dots \xrightarrow{d^{n-1}} \mathcal{P}_r^-\Lambda^n(\mathbb{R}^n) \longrightarrow 0$$

is a subcomplex of the de Rham complex with vanishing cohomology, for every  $r \geq 1$ . This follows by the exact sequence property of (2.21) together with the following result.

**Lemma 2.4.1** ([AFW06, Lemma 3.8]). *The following statements hold:*

(i)  $d^k \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) \subset d^k \mathcal{P}_r\Lambda^k(\mathbb{R}^n) \subset \mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) \subset \mathcal{P}_r^-\Lambda^{k+1}(\mathbb{R}^n)$  for all  $r \geq 1$ ;

(ii) *The restrictions*

$$\begin{aligned} d^k : \mathcal{P}_r\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n), & d^k : \mathcal{P}_{r+1}^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_{r+1}^-\Lambda^{k+1}(\mathbb{R}^n), \\ d^k : \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_r^-\Lambda^{k+1}(\mathbb{R}^n), & d^k : \mathcal{P}_{r+1}^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_r\Lambda^{k+1}(\mathbb{R}^n), \end{aligned}$$

*have the same kernel.*

(iii) *The restrictions*

$$\begin{aligned} d^k : \mathcal{P}_r\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n), & d^k : \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_r^-\Lambda^{k+1}(\mathbb{R}^n), \\ d^k : \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_r^-\Lambda^{k+1}(\mathbb{R}^n), & d^k : \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n) &\rightarrow \mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n), \end{aligned}$$

*have the same image.*

Eventually, polynomial discrete differential forms are used to approximate physical quantities satisfying field equations. In order to apply the finite element machinery and derive interpolation estimates, the pullback of discrete differential forms is of paramount importance, and this requires affine-invariant polynomial spaces. While the space of homogeneous differential  $k$ -forms  $\mathcal{H}_r\Lambda^k(\mathbb{R}^n)$  is linear but not affine invariant, it can be shown that both  $\phi^*(\mathcal{P}_r\Lambda^k(\mathbb{R}^n)) \subset \mathcal{P}_r\Lambda^k(\mathbb{R}^n)$  and  $\phi^*(\mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)) \subset \mathcal{P}_r^-\Lambda^k(\mathbb{R}^n)$  hold for any affine map  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . A complete list and discussion on finite dimensional spaces of polynomial differential forms invariant under affine transformations can be found in [AFW06, Section 3.4].

To the purpose of constructing finite element spaces of discrete differential forms on Cartesian meshes in Section 2.4.1, two other families of polynomial differential forms are considered: the first family  $\mathcal{Q}_r^-\Lambda^k$  of polynomial differential forms on tensor product domains [ABB15], and the class of polynomial spaces  $\mathcal{S}_r\Lambda^k$  obtained in [AA14] as generalization to differential forms of the serendipity finite elements. Let  $\Sigma(k, n)$  denote the set of subsets of  $[1, n] \cap \mathbb{N}$  containing  $k$  elements,  $0 \leq k \leq n$ . Let  $\sigma = \{\sigma_1, \dots, \sigma_k\} \in \Sigma(k, n)$  with  $\sigma_1 < \dots < \sigma_k$ . The algebraic  $k$ -forms  $(d\lambda)_\sigma := d\lambda_{\sigma_1} \wedge \dots \wedge d\lambda_{\sigma_k}$  form a basis for  $\text{Alt}^k \mathbb{R}^n$ , where  $\{\lambda_i\}_i$  are the barycentric coordinate functions. A form monomial in  $n$  variables is defined as  $m = x^\alpha d\lambda_\sigma$  for some multi-index  $\alpha$  and  $\sigma \in \Sigma(k, n)$ . The polynomial degree of  $m$  is  $\deg m = |\alpha| := \|\alpha\|_{\ell^1}$ , and the linear degree of  $m$  is defined as  $\deg m := \#\{i \in \sigma^* : \alpha_i = 1\}$ , where  $\sigma^* \in \Sigma(n - k, n)$  is the complement of  $\sigma$  in  $[1, n] \cap \mathbb{N}$ . The linear degree is the polynomial degree counting only the variables which enter linearly, in contrast to the superlinear degree which corresponds to the polynomial degree ignoring the variables contributing linearly. The concept of linear degree of polynomial differential forms follows by extension, namely  $\deg \omega$  is the minimum of the linear degrees of the monomials in  $\omega \in \mathcal{H}_r\Lambda^k(\mathbb{R}^n)$ . The space of forms in  $\mathcal{H}_r\Lambda^k(\mathbb{R}^n)$  with polynomial degree  $r$  and linear degree at least  $\ell \leq \min\{r, n - k\}$  is denoted by  $\mathcal{H}_{r,\ell}\Lambda^k(\mathbb{R}^n)$ . For  $r \geq 1$ , let  $\mathcal{J}_r\Lambda^k(\mathbb{R}^n) := \bigoplus_{\ell \geq 1} \kappa \mathcal{H}_{r+\ell-1,\ell}\Lambda^{k+1}(\mathbb{R}^n)$ . Then we define [AA14, Equation (17)],

$$\mathcal{S}_r\Lambda^k(\mathbb{R}^n) := \mathcal{P}_r\Lambda^k(\mathbb{R}^n) + \mathcal{J}_r\Lambda^k(\mathbb{R}^n) + d^{k-1} \mathcal{J}_{r+1}\Lambda^{k-1}(\mathbb{R}^n). \quad (2.22)$$

Note that  $\mathcal{J}_r\Lambda^n(\mathbb{R}^n) = \mathcal{J}_r\Lambda^{n-1}(\mathbb{R}^n) = 0$  and  $\mathcal{J}_r\Lambda^{n-2}(\mathbb{R}^n) = \kappa \mathcal{H}_{r,1}\Lambda^{n-1}(\mathbb{R}^n)$ . It immediately follows from the definition that the polynomials spaces  $\mathcal{S}_r\Lambda^k(\mathbb{R}^n)$  form a subcomplex of the de Rham complex. Indeed,  $d^k \mathcal{S}_r\Lambda^k(\mathbb{R}^n) = d^k \mathcal{P}_r\Lambda^k(\mathbb{R}^n) + d^k \mathcal{J}_r\Lambda^k(\mathbb{R}^n) \subset \mathcal{P}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) + \mathcal{S}_{r-1}\Lambda^{k+1}(\mathbb{R}^n) \subset \mathcal{S}_{r-1}\Lambda^{k+1}(\mathbb{R}^n)$ .

As we shall see in the forthcoming Section 2.4.1, the trace of polynomial  $k$ -forms  $\omega \in \mathcal{S}_r \Lambda^k(\mathbb{R}^n)$  on hyperplanes  $f$  is a polynomial  $k$ -form in  $\mathcal{S}_r \Lambda^k(f)$ : it holds  $\text{tr } \mathcal{S}_r \Lambda^k(\mathbb{R}^n) = \mathcal{S}_r \Lambda^k(f)$ .

The spaces  $\mathcal{Q}_r^- \Lambda^k$  rests on a tensor product construction, which allows to build a subcomplex of the de Rham complex, as shown in Section 2.1. Note that the de Rham subcomplexes associated with the spaces  $\mathcal{P}_r \Lambda^k$  and the  $\mathcal{P}_r^- \Lambda^k$  coincide in dimension one. Hence, for  $n = 2$ , taking the tensor product of the aforementioned subcomplex with itself, the spaces  $\mathcal{Q}_r^- \Lambda^k(I \times I)$  can be constructed as  $\mathcal{Q}_r^- \Lambda^0(I \times I) := \mathcal{P}_r(I) \otimes \mathcal{P}_r(I)$ ,  $\mathcal{Q}_r^- \Lambda^1(I \times I) := (\mathcal{P}_{r-1}(I) \otimes \mathcal{P}_r(I))dx^1 \oplus (\mathcal{P}_r(I) \otimes \mathcal{P}_{r-1}(I))dx^2$ , and  $\mathcal{Q}_r^- \Lambda^2(I \times I) := (\mathcal{P}_{r-1}(I) \otimes \mathcal{P}_{r-1}(I))dx^1 \wedge dx^2$ . Iterating this process  $n$ -times, for arbitrary  $n \in \mathbb{N}$ , a new subcomplex  $\{\mathcal{Q}_r^- \Lambda^k(I^n), d^k\}_{k=0}^n$  of the de Rham complex on the unit  $n$ -hypercube  $I^n$  is derived; the shape function spaces are [ABB15, Section 5]

$$\mathcal{Q}_r^- \Lambda^k(I^n) := \bigoplus_{\sigma \in \Sigma(k,n)} \left( \bigotimes_{i=1}^n \mathcal{P}_{r-\delta_{i,\sigma}}(I) \right) dx^{\sigma_1} \wedge \dots \wedge dx^{\sigma_k}, \quad \dim \mathcal{Q}_r^- \Lambda^k(I^n) = \binom{n}{k} (r+1)^{n-k} r^k,$$

where  $\sigma = \{\sigma_1, \dots, \sigma_k\} \in \Sigma(k,n)$  is an increasing map  $\{1, \dots, k\} \rightarrow \{1, \dots, n\}$  and  $\delta_{i,\sigma} = 1$  if  $i \in \{\sigma_1, \dots, \sigma_k\}$ ,  $\delta_{i,\sigma} = 0$  otherwise.

#### 2.4.1. Finite Element Discrete Differential Forms

To define finite element spaces of differential forms we need to specify the triple  $\{T, \zeta, W\}$  where:  $T$  is a closed subset of  $\mathbb{R}^n$  with non-empty interior and Lipschitz continuous boundary;  $\zeta(T)$  is an  $n$ -dimensional space of functions defined on  $T$ ;  $W(T)$  is a unisolvant set of  $n$  linear functionals in the dual space  $\zeta(T)^*$  (*the degrees of freedom*). Moreover, we aim at finite element spaces which, other than carrying good approximation properties, provide a subcomplex of the de Rham complex by admitting projection operators that commute with the exterior derivative.

**Assumption 2.4.2 (Mesh).** Let  $\{\mathcal{T}_h\}_{h>0}$  be a family of cellular decompositions of the domain  $\Omega \subset \mathbb{R}^n$ , in the sense that for each  $h$ ,  $\mathcal{T}_h$  is a finite set of oriented  $n$ -cells such that the union of all the  $n$ -cells of  $\mathcal{T}_h$  is the closure of  $\Omega$  and  $\mathcal{T}_h$  is a cellular complex. That is, any face of a  $n$ -cell in  $\mathcal{T}_h$  is also in  $\mathcal{T}_h$ , and the intersection of any two  $n$ -cells in  $\mathcal{T}_h$  is either empty or a common face. Let  $h_f := \sup\{\mathbf{x} \cdot \mathbf{y}, \mathbf{x}, \mathbf{y} \in f\}$  for  $f \in \Delta_j(\Omega)$ ,  $0 \leq j \leq n$ . The family of meshes is indexed by the mesh width  $h = \max\{h_T; T \in \mathcal{T}_h\}$ . We assume that  $\mathcal{T}_h$ , for every  $h$ , is one of the following:

- (i) A finite element simplicial decomposition of  $\Omega$  in the sense of Ciarlet [Cia78, Section 3.1], namely shape regular so that there exists  $\rho_0$  independent of  $h$  such that  $0 < \rho_0 < \min\{h_T/\rho_T : T \in \mathcal{T}_h\} < \max\{h_T/\rho_T : T \in \mathcal{T}_h\} < \rho_0^{-1}$ , where  $\rho_T$  is the radius of the largest ball inscribed in  $T \in \mathcal{T}_h$ ; and quasi-uniform, i.e., there exists  $\kappa$  independent of  $h$  such that  $0 < \kappa < \max\{h_T : T \in \mathcal{T}_h\}/\min\{h_T : T \in \mathcal{T}_h\} < \kappa^{-1}$ .
- (ii) A tensor product mesh, namely a compatible, conforming, locally quasi-uniform, affine mesh partition of  $\Omega$  into non-degenerate axiparallel parallelotopes.

Furthermore, let  $\Delta_j(\mathcal{T}_h)$  denote the set of all  $j$ -faces of  $\mathcal{T}_h$  with the convention  $\mathcal{T}_h = \Delta_n(\mathcal{T}_h)$ . Let  $\Delta_j(\mathcal{T}_h) = \Delta_j^\circ(\mathcal{T}_h) \cup \Delta_j^\partial(\mathcal{T}_h)$  where  $\Delta_j^\circ(\mathcal{T}_h)$  denotes the set of interior  $j$ -faces and  $\Delta_j^\partial(\mathcal{T}_h)$  the set of  $j$ -faces of  $\mathcal{T}_h$  belonging to the domain boundary  $\partial\Omega$ . The set of all cells of  $\mathcal{T}_h$  is denoted by  $\Delta(\mathcal{T}_h) := \bigcup_{j=0}^n \Delta_j(\mathcal{T}_h)$ . With a small abuse of notation, given an  $m$ -face  $f \in \Delta_m(\mathcal{T}_h)$ , let  $\Delta_j(f)$  denote the set of all  $j$ -dimensional subsfaces of  $f$  if  $j < m$ , the set of all  $j$ -faces containing  $f$  if  $j > m$ , and  $\Delta_m(f) := \{f\}$ . Each interior  $j$ -face is equipped with an intrinsic orientation and the boundary  $(n-1)$ -faces are by convention assumed to be oriented such that the normal vector points outwards.

On a bounded domain  $\Omega \subset \mathbb{R}^n$  with trivial topology we consider a partition  $\mathcal{T}_h$  satisfying Assumption 2.4.2. The spaces  $\zeta(T)$  of *shape functions* on each  $n$ -cell  $T \in \mathcal{T}_h$  are then defined as the restriction to  $T$  of the spaces of polynomial differential  $k$ -forms on  $\mathbb{R}^n$ , previously introduced. Moreover, we need to supply a set of degrees of freedom, namely a basis for the dual space  $\zeta(T)^*$ .

### Finite Element Differential Forms on Simplicial Meshes

In order to characterize the (global) finite element spaces of polynomial differential  $k$ -forms on simplicial meshes, we give geometrical decompositions for the local spaces  $\mathcal{P}_r\Lambda^k(T)$ ,  $\mathcal{P}_r^-\Lambda^k(T)$  and their duals.

Let  $0 \leq k \leq n$  and  $r \geq 1$ . On a  $n$ -simplex  $T \in \mathcal{T}_h$ , the degrees of freedom of  $\mathcal{P}_r\Lambda^k(T)$  are defined on every  $j$ -simplex  $f_j \in \Delta_j(T)$ , with  $k \leq j \leq \min\{n, r+k-1\}$ , as

$$\omega \in \mathcal{P}_r\Lambda^k(T) \longmapsto W_{f_j,T}^\ell(\omega) := \int_{f_j} \operatorname{tr} \omega \wedge \eta_j^\ell \quad \forall \ell = 1, \dots, N_j, \quad (2.23)$$

where  $\{\eta_j^\ell\}_{\ell=1}^{N_j}$  is a basis of  $\mathcal{P}_{r-j+k}^-\Lambda^{j-k}(f_j)$  (see e.g. [Hip02, Sections 3.2 and 3.4]). The decomposition  $\mathcal{P}_r\Lambda^k(T) = \bigoplus_{f_j \in \Delta_j(T)} V(f_j)$  holds with  $V(f_j)$  isomorphic to  $\mathcal{P}_{r-j+k}^-\Lambda^{j-k}(f_j)$ . Moreover, if the moments of  $\omega \in \mathcal{P}_r\Lambda^k(T)$  vanish, that is,  $W_{f_j,T}^\ell(\omega) = 0$  for all  $\ell = 1, \dots, N_j$ , on every  $f_j \in \Delta_j(T)$ ,  $k \leq j \leq \min\{n, r+k-1\}$ , then  $\omega = 0$ . The mapping  $\mathcal{P}_{r-j+k}^-\Lambda^{j-k}(f_j) \rightarrow W(f_j) := \{\phi \in \mathcal{P}_r\Lambda^k(T)^*: \phi(\omega) = \int_{f_j} \operatorname{tr} \omega \wedge \eta \text{ with } \eta \in \mathcal{P}_{r-j+k}^-\Lambda^{j-k}(f_j)\}$  is surjective and a counting argument shows that it is an isomorphism. A set of unisolvant degrees of freedom for  $\mathcal{P}_r\Lambda^k(T)$  is then provided by the decomposition [AFW06, Theorem 4.10],

$$\mathcal{P}_r\Lambda^k(T)^* = \bigoplus_{\substack{f_j \in \Delta_j(T) \\ k \leq j \leq \min\{n, r+k-1\}}} W(f_j). \quad (2.24)$$

Let  $1 \leq k \leq n$  and  $r \geq 1$ . Similarly, the degrees of freedom of  $\mathcal{P}_r^-\Lambda^k(T)$  are defined on every  $j$ -simplex  $f_j \in \Delta_j(T)$ , with  $k \leq j \leq \min\{n, r+k-1\}$ , as

$$\omega \in \mathcal{P}_r^-\Lambda^k(T) \longmapsto W_{f_j,T}^\ell(\omega) := \int_{f_j} \operatorname{tr} \omega \wedge \eta_j^\ell \quad \forall \ell = 1, \dots, N_j, \quad (2.25)$$

where  $\{\eta_j^\ell\}_{\ell=1}^{N_j}$  is a basis of  $\mathcal{P}_{r-j+k-1}^-\Lambda^{j-k}(f_j)$ . Moreover, the local polynomial space  $\mathcal{P}_r^-\Lambda^k(T)$  can be decomposed as in [AFW06, Theorem 4.15], namely  $\mathcal{P}_r^-\Lambda^k(T) = \bigoplus_{f_j \in \Delta_j(T)} V(f_j)$ , with  $V(f_j)$  isomorphic to  $\mathcal{P}_{r-j+k-1}^-\Lambda^{j-k}(f_j)$ . The unisolvence of the degrees of freedom for  $\mathcal{P}_r^-\Lambda^k(T)$  can be established via a geometric decomposition of the dual space  $\mathcal{P}_r^-\Lambda^k(T)^*$  analogous to (2.24) (see [AFW06, Theorem 4.14]).

### Finite Element Differential Forms on Cartesian Meshes

On an  $n$ -parallelopiped  $T$ , the degrees of freedom associated with the shape functions space  $\mathcal{S}_r\Lambda^k(T)$ ,  $0 \leq k \leq n$  and  $r \geq 1$ , are defined on every  $j$ -face  $f_j \in \Delta_j(T)$ , with  $k \leq j \leq \min\{n, \lfloor r/2 \rfloor + k\}$ , as

$$\omega \in \mathcal{S}_r\Lambda^k(T) \longmapsto W_{f_j,T}^\ell(\omega) := \int_{f_j} \operatorname{tr} \omega \wedge \eta_j^\ell \quad \forall \ell = 1, \dots, N_j, \quad (2.26)$$

where  $\{\eta_j^\ell\}_{\ell=1}^{N_j}$  is a basis of  $\mathcal{P}_{r-2(j-k)}\Lambda^{j-k}(f_j)$ . The set of degrees of freedom is unisolvant by [AA14, Theorem 3.6]. If  $f_{n-1} \in \Delta_{n-1}(T)$ , the degrees of freedom (2.26) associated with  $f_{n-1}$  and any  $j$ -subface  $f_j \subset f_{n-1}$  determine the trace of the finite element differential form on  $f_{n-1}$ . Hence,  $\mathcal{S}_r\Lambda^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$ . Moreover, the spaces  $\mathcal{S}_r\Lambda^k(T)$  form a subcomplex of the de Rham complex: the long sequence

$$\mathbb{R} \hookrightarrow \mathcal{S}_r\Lambda^0(T) \xrightarrow{d^0} \mathcal{S}_{r-1}\Lambda^1(T) \xrightarrow{d^1} \dots \xrightarrow{d^{n-1}} \mathcal{S}_{r-n}\Lambda^n(T) \longrightarrow 0$$

is exact.

Analogously, on an  $n$ -parallelopiped  $T \in \mathcal{T}_h$  a set of unisolvant degrees of freedom of  $\mathcal{Q}_r^-\Lambda^k(T)$ ,  $r \geq 1$ , is defined as in [ABB15, p. 12],

$$\omega \in \mathcal{Q}_r^-\Lambda^k(T) \longmapsto W_{f_j,T}^\ell(\omega) := \int_{f_j} \operatorname{tr} \omega \wedge \eta_j^\ell \quad \forall \ell = 1, \dots, N_j, \quad (2.27)$$

on every  $j$ -face  $f_j \in \Delta_j(T)$ , with  $k \leq j \leq \min\{n, r+k-1\}$  and where  $\{\eta_j^\ell\}_{\ell=1}^{N_j}$  is a basis of  $\mathcal{Q}_{r-1}^-\Lambda^{j-k}(f_j)$ . The commutativity of the pullback with the exterior derivative yields  $d^k \mathcal{Q}_r^- \Lambda^k(\mathcal{T}_h) \subset \mathcal{Q}_r^- \Lambda^{k+1}(\mathcal{T}_h)$  (subcomplex property). The family of finite element spaces  $\mathcal{Q}_r^- \Lambda^k(\mathcal{T}_h)$  together with the exterior derivative  $d^k$  form a long exact sequence.

The construction of the polynomial spaces  $\mathcal{Q}_r^- \Lambda^k(T)$  can be generalized to curvilinear cubes (convex quadrilaterals in two dimensions and cubes with straight edges but possibly non-planar faces in three dimensions) as in [ABB15], where approximation properties of the corresponding finite element spaces are also derived.

### The Canonical Projections

Let  $0 \leq k \leq n$  and  $r \geq 1$ , and let  $T$  be an  $n$ -simplex. Local projection operators onto the first and second family of polynomial spaces on  $T$ ,  $\Pi_{r,T}^k : \Lambda^k(T) \rightarrow \mathcal{P}_r \Lambda^k(T)$  and  $\Pi_{r,T}^{k,-} : \Lambda^k(T) \rightarrow \mathcal{P}_r^- \Lambda^k(T)$ , are defined by

$$\begin{aligned} \int_{f_j} \text{tr}(\Pi_{r,T}^k \omega) \wedge \eta &= \int_{f_j} \text{tr} \omega \wedge \eta, \quad \forall \eta \in \mathcal{P}_{r+k-j}^- \Lambda^{j-k}(f_j), f_j \in \Delta_j(T), k \leq j \leq \min\{n, r+k-1\}, \\ \int_{f_j} \text{tr}(\Pi_{r,T}^{k,-} \omega) \wedge \eta &= \int_{f_j} \text{tr} \omega \wedge \eta, \quad \forall \eta \in \mathcal{P}_{r-j+k-1} \Lambda^{j-k}(f_j), f_j \in \Delta_j(T), k \leq j \leq \min\{n, r+k-1\}. \end{aligned}$$

Similarly on an  $n$ -hypercube  $T$ , projections  $\Pi_{r,T}^{k,s} : \Lambda^k(T) \rightarrow \mathcal{S}_r \Lambda^k(T)$  and  $\Pi_{r,T}^{k,q} : \Lambda^k(T) \rightarrow \mathcal{Q}_r^- \Lambda^k(T)$ , with  $0 \leq k \leq n$  and  $r \geq 1$ , are defined by

$$\begin{aligned} \int_{f_j} \text{tr}(\Pi_{r,T}^{k,s} \omega) \wedge \eta &= \int_{f_j} \text{tr} \omega \wedge \eta, \quad \forall \eta \in \mathcal{P}_{r-2(j-k)} \Lambda^{j-k}(f_j), f_j \in \Delta_j(T), k \leq j \leq \min\{n, \lfloor r/2 \rfloor + k\}, \\ \int_{f_j} \text{tr}(\Pi_{r,T}^{k,q} \omega) \wedge \eta &= \int_{f_j} \text{tr} \omega \wedge \eta, \quad \forall \eta \in \mathcal{Q}_{r-1}^- \Lambda^{j-k}(f_j), f_j \in \Delta_j(T), k \leq j \leq \min\{n, r+k-1\}. \end{aligned}$$

The operators above introduced commute with the exterior derivative as shown in the following result.

**Lemma 2.4.3.** *Let  $0 \leq k \leq n$  and  $r \geq 1$ . Let  $T$  be a  $n$ -simplex and let  $Q$  denote a  $n$ -dimensional hypercube. The following four diagrams commute:*

$$\begin{array}{ccc} \begin{array}{ccc} \Lambda^k(T) & \xrightarrow{d^k} & \Lambda^{k+1}(T) \\ \downarrow \Pi_{r,T}^k & & \downarrow \Pi_{r-1,T}^{k+1} \\ \mathcal{P}_r \Lambda^k(T) & \xrightarrow{d^k} & \mathcal{P}_{r-1} \Lambda^{k+1}(T) \end{array} & \quad & \begin{array}{ccc} \Lambda^k(T) & \xrightarrow{d^k} & \Lambda^{k+1}(T) \\ \downarrow \Pi_{r,T}^{k,-} & & \downarrow \Pi_{r,T}^{k+1,-} \\ \mathcal{P}_r^- \Lambda^k(T) & \xrightarrow{d^k} & \mathcal{P}_r^- \Lambda^{k+1}(T) \end{array} \\ \begin{array}{ccc} \Lambda^k(Q) & \xrightarrow{d^k} & \Lambda^{k+1}(Q) \\ \downarrow \Pi_{r,T}^{k,s} & & \downarrow \Pi_{r-1,T}^{k+1,s} \\ \mathcal{S}_r \Lambda^k(Q) & \xrightarrow{d^k} & \mathcal{S}_{r-1} \Lambda^{k+1}(Q) \end{array} & \quad & \begin{array}{ccc} \Lambda^k(Q) & \xrightarrow{d^k} & \Lambda^{k+1}(Q) \\ \downarrow \Pi_{r,T}^{k,q} & & \downarrow \Pi_{r,T}^{k+1,q} \\ \mathcal{Q}_r^- \Lambda^k(Q) & \xrightarrow{d^k} & \mathcal{Q}_r^- \Lambda^{k+1}(Q) \end{array} \end{array}$$

The proof of this result for the local spaces  $\mathcal{P}_r \Lambda^k(T)$  and  $\mathcal{P}_r^- \Lambda^k(T)$  can be found in [AFW06, Lemma 4.24]. Concerning the families of local polynomial differential forms on hypercubes, a similar argument applies based on the subcomplex property  $d^k \mathcal{Q}_r^- \Lambda^k(\mathcal{T}_h) \subset \mathcal{Q}_r^- \Lambda^{k+1}(\mathcal{T}_h)$  and  $d^k \mathcal{S}_r \Lambda^k(\mathcal{T}_h) \subset \mathcal{S}_{r-1} \Lambda^{k+1}(\mathcal{T}_h)$ , the unisolvence of  $\zeta(Q)^*$ , the Leibniz rule for the exterior derivative and the commuting of the exterior derivative with the trace operator.

### $H\Lambda^k(\Omega)$ -conforming Finite Element Differential Forms

The Hilbert space  $H\Lambda^k(\Omega)$ ,  $0 \leq k \leq n$  is characterized by function  $\omega \in L^2\Lambda^k(\Omega)$ , piecewise smooth with respect to the triangulation  $\mathcal{T}_h$ , whose trace  $\text{tr } \omega$  is single-valued on all  $(n-1)$ -cells  $f_{n-1} \in \Delta_{n-1}(\mathcal{T}_h)$ , or equivalently  $\text{tr } \omega$  is single-valued on all  $f_j \in \Delta_j(\mathcal{T}_h)$ , with  $k \leq j \leq n-1$ . Global polynomial subspaces of  $H\Lambda^k(\Omega)$  are defined on a simplicial mesh  $\mathcal{T}_h$  as

$$\mathcal{P}_r\Lambda^k(\mathcal{T}_h) := H\Lambda^k(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_r\Lambda^k(T), \quad \mathcal{P}_r^-\Lambda^k(\mathcal{T}_h) := H\Lambda^k(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_r^-\Lambda^k(T), \quad (2.28)$$

and on a Cartesian mesh  $\mathcal{T}_h$  as,

$$\mathcal{S}_r\Lambda^k(\mathcal{T}_h) := H\Lambda^k(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{S}_r\Lambda^k(T), \quad \mathcal{Q}_r^-\Lambda^k(\mathcal{T}_h) := H\Lambda^k(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{Q}_r^-\Lambda^k(T). \quad (2.29)$$

Furthermore, if  $f_{n-1} \in \Delta_{n-1}(\mathcal{T}_h)$  one can rely on the “trace property” of the shape function spaces to infer  $\text{tr } \omega \in \mathcal{P}_r\Lambda^k(f_{n-1})$ , which is hence determined by the degrees of freedom corresponding to the  $j$ -faces  $f_j \subset f_{n-1}$ . Note that the interelement continuity required for the finite element space  $\Lambda_h^k(\mathcal{T}_h)$  to belong to  $H\Lambda^k(\Omega)$  is directly imposed through the choice of degrees of freedom associated with the  $j$ -dimensional faces in  $\mathcal{T}_h$ , with  $k \leq j \leq n-1$ .

Let  $\Lambda_h^k(T)$  denote one of the local polynomial spaces of differential forms:  $\mathcal{P}_r\Lambda^k(T)$  or  $\mathcal{P}_r^-\Lambda^k(T)$  on simplices, and  $\mathcal{Q}_r^-\Lambda^k(T)$  or  $\mathcal{S}_r\Lambda^k(T)$  on hypercubes. Then, global projection operators, the *canonical projections*, can be obtained by extension of the local projections  $\Pi_T^k : \Lambda^k(T) \rightarrow \Lambda_h^k(T)$  as  $\Pi_h^k : \Lambda^k(\Omega) \rightarrow \Lambda_h^k(\mathcal{T}_h)$ ,  $(\Pi_h^k \omega)|_T = \Pi_T^k(\omega|_T)$  for  $\omega \in \Lambda^k(\Omega)$ . The  $H\Lambda^k(\Omega)$ -conforming finite element spaces formerly presented over a cell complex  $\mathcal{T}_h$  form a discrete de Rham sequence as a cochain projection from the de Rham complex (2.15) through the projection operators  $\Pi_h^k$ , namely

$$\begin{array}{ccccccc} H\Lambda^0(\Omega) & \xrightarrow{d^0} & H\Lambda^1(\Omega) & \xrightarrow{d^1} & \dots & \xrightarrow{d^{n-1}} & H\Lambda^n(\Omega) \\ \downarrow \Pi_h^0 & & \downarrow \Pi_h^1 & & & & \downarrow \Pi_h^n \\ \Lambda_h^0(\mathcal{T}_h) & \xrightarrow{d^0} & \Lambda_h^1(\mathcal{T}_h) & \xrightarrow{d^1} & \dots & \xrightarrow{d^{n-1}} & \Lambda_h^n(\mathcal{T}_h) \end{array} \quad (2.30)$$

where each  $\Lambda_h^k(\mathcal{T}_h) \rightarrow \Lambda_h^{k+1}(\mathcal{T}_h)$  can be substituted by,

$$\begin{array}{ll} \mathcal{P}_r\Lambda^k(\mathcal{T}_h) \rightarrow \mathcal{P}_{r-1}\Lambda^{k+1}(\mathcal{T}_h), & \mathcal{P}_r^-\Lambda^k(\mathcal{T}_h) \rightarrow \mathcal{P}_r^-\Lambda^{k+1}(\mathcal{T}_h), \\ \mathcal{S}_r\Lambda^k(\mathcal{T}_h) \rightarrow \mathcal{S}_{r-1}\Lambda^{k+1}(\mathcal{T}_h), & \mathcal{Q}_r^-\Lambda^k(\mathcal{T}_h) \rightarrow \mathcal{Q}_r^-\Lambda^{k+1}(\mathcal{T}_h), \end{array}$$

for all  $r \geq 1$ , (see [AFW06, Section 5.5] and [Arn13]).

Remark that the second family of conforming elements on an hexahedral partition  $\mathcal{T}_h$  does not satisfy a discrete de Rham diagram (see [Mon03, Section 8.2.3]).

**Theorem 2.4.4** (De Rham theorem for discrete differential forms). *For every  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$  with  $d^k \omega_h = 0$ , there exist  $\mu_h \in \Lambda_h^{k-1}(\mathcal{T}_h)$  and  $\gamma_h \in \mathcal{H}_k(\mathcal{T}_h; \mathbb{R})$  such that  $\omega_h = d^{k-1} \mu_h + \gamma_h$ .*

### Discontinuous Finite Element Differential Forms

Another class of finite element methods, the so-called *discontinuous Galerkin* (DG) schemes, has proven very competitive in terms of flexibility, efficiency and accuracy in approximating boundary value problems, including models arising in electromagnetism: Maxwell’s equations [CS16], eddy current models [Hou+05; BP06; GSS07] and magnetic induction [HH13b].

For a simplicial decomposition  $\mathcal{T}_h$ , the spaces of polynomial totally discontinuous discrete differential  $k$ -forms on  $\mathcal{T}_h$  are defined for  $r \geq 0$  as

$$\mathcal{P}_r^d\Lambda^k(\mathcal{T}_h) := \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_r\Lambda^k(T), \quad \mathcal{P}_r^{d,-}\Lambda^k(\mathcal{T}_h) := \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_r^-\Lambda^k(T), \quad (2.31)$$

and on a Cartesian mesh  $\mathcal{T}_h$  as,

$$\mathcal{S}_r^d \Lambda^k(\mathcal{T}_h) := \bigoplus_{T \in \mathcal{T}_h} \mathcal{S}_r \Lambda^k(T), \quad \mathcal{Q}_r^{d,-} \Lambda^k(\mathcal{T}_h) := \bigoplus_{T \in \mathcal{T}_h} \mathcal{Q}_r^- \Lambda^k(T). \quad (2.32)$$

The discontinuous finite element spaces do not fit the FEEC framework: the exterior derivative operator is well-defined only locally, on  $n$ -cells of  $\mathcal{T}_h$ , and the de Rham complex does not support discrete subcomplexes of discontinuous type. However, their construction justifies their name.

### 2.4.2. Finite Element Spaces for Vector Proxies of Discrete Differential Forms

The correspondence between classical spaces of polynomial scalar/vector-valued functions and polynomial spaces of finite element differential forms is given in [AFW06, Tables 5.1 and 5.2].

In terms of their vector proxy representations in  $\mathbb{R}^3$ , the spaces  $\mathcal{P}_r^- \Lambda^0(\mathcal{T}_h) = \mathcal{P}_r \Lambda^0(\mathcal{T}_h)$  are the generalization to polynomial scalar function of higher degree of the Lagrange finite elements, while  $\mathcal{Q}_r^- \Lambda^0(\mathcal{T}_h)$  is the standard polynomial finite element space  $\mathcal{Q}_r(\mathcal{T}_h) \subset H^1(\Omega)$ . Analogously,  $\mathcal{P}_r \Lambda^3(\mathcal{T}_h)$  is the space of globally discontinuous piecewise polynomials of degree at most  $r$ , and  $\mathcal{P}_r^- \Lambda^3(\mathcal{T}_h) = \mathcal{P}_{r-1} \Lambda^3(\mathcal{T}_h)$ . Similarly,  $\mathcal{Q}_r^- \Lambda^3(\mathcal{T}_h)$  is the piecewise polynomial space  $\mathcal{Q}_{r-1}(\mathcal{T}_h) \subset L^2(\Omega)$ .

Concerning discrete differential 1-forms and 2-forms, higher order  $H\Lambda^k(\Omega)$ -conforming finite elements  $\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$ ,  $\mathcal{Q}_r^- \Lambda^1(\mathcal{T}_h)$  and  $\mathcal{P}_r^- \Lambda^2(\mathcal{T}_h)$ ,  $\mathcal{Q}_r^- \Lambda^2(\mathcal{T}_h)$  can be identified, in two dimensions, with the Raviart–Thomas spaces [RT77], while they were introduced in three dimensions as Nédélec edge and face elements of the first kind [Néd80]. In two dimensions,  $\mathcal{P}_r \Lambda^1(\mathcal{T}_h)$  and  $\mathcal{P}_r \Lambda^2(\mathcal{T}_h)$  are also known as Brezzi–Douglas–Marini (BDM) elements [BDM85], and in three dimensions they were introduced in [Néd86] as  $H(\mathbf{curl}, \Omega)$ -conforming edge element and  $H(\operatorname{div}, \Omega)$ -conforming face elements of the second kind.

The more “exotic” spaces  $\mathcal{S}_r \Lambda^k(\mathcal{T}_h)$  on tensor product meshes  $\mathcal{T}_h$  developed in [AA11] correspond, in two dimensions, to the serendipity space and generalizations, if  $k = 0$ ; they coincide with the rectangular BDM elements [BDM85] for  $k = 1$ . The space  $\mathcal{S}_r \Lambda^3(\mathcal{T}_h)$  is the globally discontinuous space  $\mathcal{P}_r(\mathcal{T}_h)$ .

### 2.4.3. Interpolation Estimates

Interpolation estimates largely resort to techniques based on the pullback of discrete differential forms under affine mappings. In particular, since the finite element spaces previously introduced are affine-equivalent, one can proceed by local arguments on the *reference element* and affine pullbacks.

Let  $\Lambda_h^k(\mathcal{T}_h)$  be a space of polynomial (possibly discontinuous) discrete differential forms as introduced in the previous sections. A discrete differential form  $\omega \in \Lambda_h^k(T)$  admits the local representation

$$\omega(\mathbf{x}) = \sum_{j=k}^{M_{\min}} \sum_{f_j \in \Delta_j(\mathcal{T})} \sum_{\ell=1}^{N_j} W_{f_j, T}^\ell \psi_{j, T}^\ell \quad \forall \mathbf{x} \in T, \quad (2.33)$$

with the expansion coefficients  $W_{f_j, T}^\ell \in \mathbb{R}$ , and the convention that  $W_T^\ell := W_{f_n, T}^\ell$  for all  $\ell = 1, \dots, N_n$ . Moreover,  $\{\psi_{j, T}^\ell\}_{j, \ell}$  is a basis of the local polynomial space on  $T$ , and  $N_j$  is the cardinality of the set  $\Delta_j(\mathcal{T}_h)$ . If  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{S}_r \Lambda^k(\mathcal{T}_h)$ , then  $M_{\min} := \min\{n, \lfloor r/2 \rfloor + k\}$ , otherwise  $M_{\min} := \min\{n, r + k - 1\}$ .

**Lemma 2.4.5.** *Let  $\widehat{T}$  be the reference  $n$ -cell. Let  $\phi : \widehat{T} \mapsto T$  be an affine isomorphism. Then there exist  $C_1, C_2, C_3 > 0$  such that, for all  $\widehat{\omega} \in \Lambda_h^k(\widehat{T})$ , it holds*

$$C_1 \|\widehat{\mathbf{d}}^k \widehat{\omega}\|_{L^2 \Lambda^{k+1}(\widehat{T})}^2 \leq C_1 \|\widehat{\omega}\|_{L^2 \Lambda^k(\widehat{T})}^2 \leq \sum_{j=k}^{M_{\min}} \sum_{f_j \in \Delta_j^\circ(T)} \sum_{\ell=1}^{N_j} (W_{f_j, T}^\ell)^2 \leq C_2 \|\widehat{\omega}\|_{L^2 \Lambda^k(\widehat{T})}^2. \quad (2.34)$$

Analogously, let  $\hat{f}_{n-1}$  be the reference  $(n-1)$ -face. There exist  $C_4, C_5 > 0$  such that

$$C_4 \|\widehat{\operatorname{tr}} \widehat{\omega}\|_{L^2 \Lambda^k(\hat{f}_{n-1})}^2 \leq \sum_{j=k}^{M_{\min}} \sum_{f_j \in \Delta_j(f_{n-1})} \sum_{\ell=1}^{N_j} (W_{f_j, T}^\ell)^2 \leq C_5 \|\widehat{\operatorname{tr}} \widehat{\omega}\|_{L^2 \Lambda^k(\hat{f}_{n-1})}^2.$$

The constants  $C_i$ ,  $i \in \{1, 2, 3, 4, 5\}$ , depend on the polynomial degree and the differential form degree  $k$ .

For vector proxies of discrete differential forms transformations of norms under pullbacks apply, see [Hip02, Section 3.6] for further details. Specifically, let  $H^m(\Omega)^n$ ,  $m \in \mathbb{N}_0$ , denote the Sobolev space associated with the semi-norm  $|\mathbf{v}|_{H^m(\Omega)^n}^2 := \sum_{j=1}^n \int_{\Omega} |\partial^m v_j|^2 d\mathbf{x}$ , for  $\mathbf{v} = (v_1, \dots, v_n)$ . Similarly for Sobolev spaces  $H^m(\Omega)$  of scalar functions.

**Lemma 2.4.6.** *Let  $T = \phi(\hat{T})$  be the image of the reference element under a bijective affine map  $\phi(\hat{\mathbf{x}}) = B_T \hat{\mathbf{x}} + c_T$ . Then, there exists  $C > 0$  depending only on the shape of  $T$ , such that for  $m \in \mathbb{N}_0$  and  $\mathbf{v} \in H^m(T)^n$ ,  $v \in H^m(T)$ ,*

$$\begin{aligned} |h_\Phi^0 v|_{H^m(\hat{T})}^2 &\leq C |\det(B_T)|^{-1} \|B_T\|^{2m} |v|_{H^m(T)}^2 \leq Ch_T^{2m-n} |v|_{H^m(T)}^2, \\ |h_\Phi^1 \mathbf{v}|_{H^m(\hat{T})}^2 &\leq C |\det(B_T)|^{-1} \|B_T\|^{2+2m} |\mathbf{v}|_{H^m(T)}^2 \leq Ch_T^{2+2m-n} |\mathbf{v}|_{H^m(T)}^2, \\ |h_\Phi^2 \mathbf{v}|_{H^m(\hat{T})}^2 &\leq C |\det(B_T)| \|B_T\|^{2m} \|B_T^{-1}\|^2 |\mathbf{v}|_{H^m(T)}^2 \leq Ch_T^{2m-2+n} |\mathbf{v}|_{H^m(T)}^2, \end{aligned} \quad (2.35)$$

where  $\|B_T\|$  denotes the Euclidean matrix norm of  $B_T$ . Moreover, if  $\mathbf{v} \in H(\operatorname{curl}, T)$  and  $\mathbf{z} \in H(\operatorname{div}, T)$ ,

$$\begin{aligned} |h_\Phi^1 \mathbf{v}|_{H(\operatorname{curl}, \hat{T})}^2 &\leq C |\det(B_T)| \|B_T^{-1}\|^2 |\mathbf{v}|_{H(\operatorname{curl}, T)}^2 \leq Ch_T^{n-2} |\mathbf{v}|_{H(\operatorname{curl}, T)}^2, & \text{for } n = 3, \\ |h_\Phi^1 \mathbf{v}|_{H(\operatorname{curl}, \hat{T})}^2 &\leq C |\det(B_T)| |\mathbf{v}|_{H(\operatorname{curl}, T)}^2 \leq Ch_T^n |\mathbf{v}|_{H(\operatorname{curl}, T)}^2, & \text{for } n = 2, \\ |h_\Phi^2 \mathbf{z}|_{H(\operatorname{div}, \hat{T})}^2 &\leq C |\det(B_T)| |\mathbf{z}|_{H(\operatorname{div}, T)}^2 \leq Ch_T^n |\mathbf{z}|_{H(\operatorname{div}, T)}^2. \end{aligned} \quad (2.36)$$

*Proof.* The proof is based on scaling arguments, the pullback formulas (2.4), and the properties of affine mappings:  $|\det(B_T)| \lesssim h_T^n$ ,  $|\det(B_T)|^{-1} \lesssim h_T^{-n}$  and  $\|B_T\| \lesssim h_T$ ,  $\|B_T^{-1}\| \lesssim h_T^{-1}$ , where we have omitted the dependence on the shape regularity constant. In particular, we refer to e.g. [Hip02, Lemma 3.12] for the estimates (2.35), and to e.g. [Mon03, Lemma 5.43], [BBF13, Lemma 2.1.8] for the bounds (2.36). The derivation of (2.36) relies on the Piola covariant and contravariant transformations yielding

$$\operatorname{curl} \mathbf{v}(\mathbf{x}) = \frac{1}{|\det B_T|} B_T \widehat{\operatorname{curl}} \widehat{\mathbf{v}}(\phi^{-1}(\mathbf{x})), \quad \operatorname{div} \mathbf{z}(\mathbf{x}) = \frac{1}{|\det B_T|} \widehat{\operatorname{div}} \widehat{\mathbf{z}}(\phi^{-1}(\mathbf{x})),$$

and, in two dimensions, on the isomorphism between the spaces  $H(\operatorname{curl}, T)$  and  $H(\operatorname{div}, T)$ .  $\square$

As shown in [Hip02, Equation (3.37)], the affine equivalence techniques deployed in the previous Lemma 2.4.6 implies uniform  $L^2$ -stability of the local basis functions of  $\Lambda_h^k(\mathcal{T}_h)$ . As immediate consequence of Lemma 2.4.5 and (2.35), we can obtain bounds on the trace of the jump across  $(n-1)$ -faces of vector proxies of discontinuous discrete differential 1-forms ubiquitous in discontinuous Galerkin approximations of boundary value problems in electromagnetism. Their derivation rests upon affine transformations and in particular the fact that the unit normal  $\hat{\mathbf{n}}$  vector to  $\hat{f}$  is mapped as  $\mathbf{n} \circ \phi = B_T^{-T} \hat{\mathbf{n}} / |B_T^{-T} \hat{\mathbf{n}}|$ .

**Lemma 2.4.7.** *Let  $\mathcal{T}_h$  be a cellular complex as in Assumption 2.4.2. Let  $f \in \Delta_{n-1}(T^-) \cap \Delta_{n-1}(T^+)$  be a  $(n-1)$ -cell;  $T^-, T^+ \in \mathcal{T}_h$ . Let  $\Lambda_h^{d,1}(\mathcal{T}_h)$  be a discontinuous polynomial space of vector proxies of discrete differential 1-forms. Then, there exist  $C_1, C_2 > 0$  depending only on the polynomial degree  $r$  and the shape regularity of  $\mathcal{T}_h$  such that, for all  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$ , it holds*

$$C_1 h^{3-n} \|\operatorname{tr} [\mathbf{v}]_f\|_{L^2(f)}^2 \leq \sum_{j=1}^{M_{\min}} \sum_{f_j \in \Delta_j(f)} \sum_{\ell=1}^{N_j} (V_{f_j, T^+}^\ell - V_{f_j, T^-}^\ell)^2 \leq C_2 h^{3-n} \|\operatorname{tr} [\mathbf{v}]_f\|_{L^2(f)}^2, \quad (2.37)$$

where the jump of  $\mathbf{v}$  across  $f$  is defined as  $[\mathbf{v}]_f := \mathbf{v}|_{T^+} - \mathbf{v}|_{T^-}$  and  $T^+$  has outward normal;  $\{V_{f_j, T^\pm}^\ell\}_{j,\ell}$  are the degrees of freedom associated with the representation of  $\mathbf{v}$ , as in (2.33). If  $\Lambda_h^{d,1}(\mathcal{T}_h) = \mathcal{S}_r^d \Lambda^1(\mathcal{T}_h)$ , then  $M_{\min} := \min\{n, \lfloor r/2 \rfloor + 1\}$ , otherwise  $M_{\min} := \min\{n, r\}$ .

The approximation properties of polynomial finite element spaces of differential forms are established in the following results.

**Theorem 2.4.8** ([AFW06, Theorem 5.3]). *Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain and let  $\mathcal{T}_h$  be a triangulation of  $\Omega$  as in Assumption 2.4.2 (i). Let  $\Pi_h^k$  denote the canonical projection of  $\Lambda^k(\Omega)$  onto  $\mathcal{P}_r\Lambda^k(\mathcal{T}_h)$ , for  $0 \leq k \leq n$  and  $r \geq 1$ . Let  $1 \leq p \leq \infty$  and  $(n-k)/p \leq s \leq r+1$ . Then  $\Pi_h^k$  admits a bounded extension to  $W_p^s \Lambda^k(\Omega)$  and*

$$\|\omega - \Pi_h^k \omega\|_{L^p \Lambda^k(\Omega)} \leq Ch^s |\omega|_{W_p^s \Lambda^k(\Omega)}, \quad \forall \omega \in W_p^s \Lambda^k(\Omega),$$

with constant  $C > 0$  independent of  $h$ .

The definition of the canonical projections from Section 2.4.1 relies on traces on simplices requiring therefore spaces with sufficient smoothness properties. For scalar  $H^1$ -conforming finite element, the Clément quasi-interpolation operator [Clé75] provides an operator bounded in  $L^2$  and which delivers optimal error bounds. However, the Clément operator is not a projection and does not preserve homogeneous boundary conditions. An alternative quasi-interpolation operator fulfilling these properties was proposed by Scott and Zhang [SZ90]. Extensions to vector-valued elements and  $H\Lambda^k(\Omega)$ ,  $k > 0$ , spaces have been introduced first in [Sch01] and further developed in [AFW06, Section 5.4] and [CW08], where a set of operators  $\tilde{\Pi}_h^k$  is constructed by combining the canonical projections with a smoothing operator and multiplying by the inverse operator restricted to the finite element space. Finally in [FW14], Falk and Winther introduced  $L^2$ -bounded, local, commuting projectors. More recently, in [EG15] projections of  $H\Lambda^k(\Omega)$  spaces are constructed by combining a projection onto the discontinuous finite element space with an averaging operator. The smoothed operators  $\tilde{\Pi}_h^k : L^2 \Lambda^k(\Omega) \rightarrow \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$  provide  $L^2$ -bounded projections and commute with the exterior derivative.

**Theorem 2.4.9.** *Let  $\Omega \subset \mathbb{R}^n$  be a bounded Lipschitz domain, and let  $\mathcal{T}_h$  be a triangulation of  $\Omega$  as in Assumption 2.4.2 (i). Let  $\tilde{\Pi}_h^k$  denote the smoothed projection onto  $\mathcal{P}_r \Lambda^k(\mathcal{T}_h)$ . For  $0 \leq s \leq r+1$ , there holds*

$$\|\omega - \tilde{\Pi}_h^k \omega\|_{L^2 \Lambda^k(\Omega)} \leq Ch^s |\omega|_{H^s \Lambda^k(\Omega)}, \quad \forall \omega \in H^s \Lambda^k(\Omega),$$

with constant  $C > 0$  independent of  $h$ .

Optimal error estimates are achieved by applying the following result.

**Theorem 2.4.10.** *Let  $K \subset \mathbb{R}^n$  be a bounded Lipschitz domain. Then, for all  $\omega \in H^{r+1} \Lambda^k(K)$ ,  $r \geq 0$  and  $0 \leq k \leq n$ , it holds*

$$\inf_{p_h \in \mathcal{P}_r \Lambda^k(K)} \|\omega - p_h\|_{L^2 \Lambda^k(K)} \leq Ch_K^{r+1} |\omega|_{H^{r+1} \Lambda^k(K)},$$

where the constant  $C > 0$  depends on the polynomial degree  $r$  and the shape regularity constant.

In particular, for  $r \geq 1$ , on each element  $T \in \mathcal{T}_h$  of a simplicial mesh, analogous estimates with a factor  $h_T^r$  can be obtained for the polynomial space  $\mathcal{P}_r^- \Lambda^k(T)$  since  $\mathcal{P}_{r-1} \Lambda^k(T) \subset \mathcal{P}_r^- \Lambda^k(T) \subset \mathcal{P}_r \Lambda^k(T)$  (convergence of full order is re-established if  $k = 0$ ). On tensor product meshes, one can proceed likewise:  $\mathcal{S}_r \Lambda^k(T)$  yields error estimates of order  $r+1$  since  $\mathcal{P}_r \Lambda^k(T) \subset \mathcal{S}_r \Lambda^k(T)$  with  $r \geq 1$  (and  $\mathcal{S}_r \Lambda^k(T)$  does not contain  $\mathcal{P}_{r+1} \Lambda^k(T)$ ). The case of  $\mathcal{Q}_r \Lambda^k(T)$  is analogous to the first family on simplicial meshes. Indeed  $\mathcal{P}_{r-1} \Lambda^k(T) \subset \mathcal{Q}_r \Lambda^k(T)$  and  $\mathcal{P}_r \Lambda^0(T) \subset \mathcal{Q}_r \Lambda^0(T)$ . Similar approximation estimates hold for spaces of globally discontinuous finite elements with  $r \geq 0$ .



# 3. Stabilized Galerkin for Transient Advection of Differential Forms

This chapter pertains to the numerical discretization of the generalized transient advection problem for differential forms, which include the magnetic advection equation (2.1), relevant for MHD models. We pursue an Eulerian method of lines with explicit time-stepping.

In Section 3.2, we devise a stabilized Galerkin spatial semi-discretization for the generalized stationary advection problem for merely piecewise smooth velocity  $\mathbf{u}$  resulting from numerical discretizations of the extended Euler system. Next, Section 3.3 establishes a priori convergence estimates for the stabilized Galerkin discretization in the stationary setting. For want of well-posedness results for the generalized advection problem in case of discontinuous  $\mathbf{u}$ , these investigations are confined to Lipschitz continuous velocities  $\mathbf{u} \in W^{1,\infty}(\Omega)$ . The stability and consistency results obtained in the stationary case are instrumental for the convergence analysis of the fully discrete scheme in Section 3.4. Finally, in Sections 3.5 and 3.6 the performance of the new method is tested in various numerical experiments for both the stationary and the transient generalized advection problem in 2D. The tests cover both continuous and discontinuous velocities and employ tensor product grids and simplicial meshes.

The main results contained in this chapter have been published in [HHP16].

## 3.1. Generalized Pure Advection Evolution Problem

It is well known that for scalar advection-diffusion equation (2.2) standard Galerkin discretizations with Lagrangian finite elements will break down in the singular perturbation limit of vanishing diffusion. The onset of spurious oscillations reflects the weakly coercive nature of the problem in the energy norm namely the loss of control of the ratio between the continuity constant, which depends on the velocity field, and the coercivity constant, proportional to the diffusion coefficient. A plethora of stabilization mechanisms have been devised to provide additional control in some strong norm (see [RST08] for literature on stabilization techniques for singularly perturbed problems). The addition to the discrete variational formulation of consistent terms weighted by the strong residual of the equation is at the root of the so-called residual-based methods, such as SUPG (streamline upwind Petrov Galerkin) [HB79] or the Galerkin least squares [HFH89]. Another class of (linear) stabilization techniques for Galerkin methods based on discontinuous approximation spaces rests upon *upwind* numerical fluxes at element interfaces, see [Arn+02] and references therein. The gist of both approaches is to systematically add a consistent artificial diffusivity to counteract the spreading of unphysical oscillations while avoiding an excessive smearing of the solution.

Robustness for  $\varepsilon \searrow 0$  is also a key issue for the spatial discretization of (2.1). In this chapter we tackle the challenge of robust Eulerian spatial finite element discretizations for the generalized advection-diffusion problem (2.17). In fact, we focus on the pure advection problem obtained from (2.17) for  $\varepsilon = 0$ ; if a scheme performs well in this case, it will also be suitable for (2.17) when augmented with a standard  $H\Lambda^k(\Omega)$ -conforming Galerkin discretization of the diffusion term. In the case of genuine discontinuous Galerkin finite elements we can rely on interior penalty DG methods (see Chapter 5 for the derivation of interior penalty DG schemes for the double **curl** operator).

The generalized pure advection initial boundary value problem in the space-time domain  $\Omega \times I$ ,  $I := [0, T]$ , reads: For  $f \in C^0(I; L^2\Lambda^k(\Omega))$  and  $\omega_0 \in W|_{t=0}$  find  $\omega \in C^1(I; L^2\Lambda^k(\Omega))$ ,  $\omega(t) \in W|_t$  such that

$$\begin{aligned} \partial_t \omega(t) + \alpha \omega(t) + \mathbf{L}_{\mathbf{u}} \omega(t) &= f(t), & \text{in } \Omega \times I, \\ \omega(0) &= \omega_0, & \text{in } \Omega, \end{aligned} \tag{3.1}$$

where  $\mathbf{u} : \overline{\Omega} \times I \rightarrow \mathbb{R}^n$  is a *given* velocity field, the reaction coefficient  $\alpha : \Omega \rightarrow \mathbb{R}$  is a non-negative and bounded function. The boundary conditions are imposed at the inflow boundary  $\Gamma_{\text{in}} := \{\mathbf{x} \in \partial\Omega : \mathbf{u} \cdot \mathbf{n}(\mathbf{x}) < 0\}$ , and the variational spaces  $V$  and  $W$  are defined as

$$V := \{\omega \in L^2 \Lambda^k(\Omega) : \mathbf{L}_{\mathbf{u}} \omega \in L^2 \Lambda^k(\Omega), \int_{\Gamma_{\text{in}}} \text{tr } i_{-\mathbf{u}}(\omega \wedge \star \omega) < \infty\},$$

$$W|_t := \{\omega \in V : \text{tr } \omega = g, \text{tr } i_{\mathbf{n}} \omega = s \text{ on } \Gamma_{\text{in}}, g(t) \in L^2 \Lambda^k(\Gamma_{\text{in}}), s(t) \in L^2 \Lambda^{k-1}(\Gamma_{\text{in}})\}.$$

All other notation can be found in Chapter 2.

A semi-Lagrangian approach for the discretization of (3.1) was proposed in [HH13a]. Conversely, in the present work we pursue a mesh based Eulerian method of lines approach to (3.1), employing a (jump) stabilized Galerkin discretization and piecewise polynomial discrete differential forms for spatial discretization. Our new methods will be constructed to accommodate discontinuous velocities aligned with the mesh.

A jump-stabilized discontinuous Galerkin method for the stationary advection problem for 0-forms in  $\mathbb{R}^3$  and Lipschitz continuous velocities  $\mathbf{u} \in W^{1,\infty}(\Omega)$ , was introduced and theoretically analyzed in, among others, [HSS00; BMS04]. An extension of these results to the stationary magnetic advection problem (1-forms in  $\mathbb{R}^3$ , cf. (2.1)) was proposed in [HH13b], where a priori convergence rates were derived for both fully discontinuous piecewise polynomial functions and  $H(\mathbf{curl}, \Omega)$ -conforming finite elements. Discontinuous velocity fields were not taken into account. We remark that for discontinuous velocities, even the spatial discretization of the scalar transport problem (2.2), for which existence and uniqueness of weak solutions are known, is discussed only rarely [Boy12; Wal05].

### 3.1.1. Well-posedness

For velocity fields uniformly continuous in time and Lipschitz continuous in space, that is,  $\mathbf{u} \in C^0(I; W^{1,\infty}(\Omega))$ , we assume the following ‘‘coercivity’’ condition on the velocity (cf. Table 3.1 for vector proxy correspondences): there exists a constant  $\alpha_0 > 0$  such that

$$\int_{\Omega} \left( \alpha + \frac{1}{2} (\mathbf{L}_{\mathbf{u}(\cdot,t)} + \mathcal{L}_{\mathbf{u}(\cdot,t)}) \right) \omega \wedge \star \omega \geq \alpha_0 \int_{\Omega} \omega \wedge \star \omega, \quad \forall \omega \in L^2 \Lambda^k(\Omega), \forall t \in I. \quad (3.2)$$

If  $\mathbf{u}$  does not depend on  $t$  and  $\mathbf{u} \in W^{1,\infty}(\Omega)$ , the Hille-Yosida theorem in [Bre11, Theorem 7.4] can be directly applied to show that, under assumption (3.2), the variational problem associated with (3.1): Find  $\omega \in C^1(I; L^2 \Lambda^k(\Omega)) \cap C^0(I; W)$  such that, for all  $\eta \in L^2 \Lambda^k(\Omega)$

$$(\partial_t \omega, \eta)_{\Omega} + (\alpha \omega, \eta)_{\Omega} + (\mathbf{L}_{\mathbf{u}} \omega, \eta)_{\Omega} = (f, \eta)_{\Omega},$$

$$(\omega(0), \eta)_{\Omega} = (\omega_0, \eta)_{\Omega},$$

is well-posed. Here  $(\cdot, \cdot)_{\Omega}$  denotes the  $L^2 \Lambda^k(\Omega)$  inner product  $(\omega, \eta)_{\Omega} := \int_{\Omega} \omega \wedge \star \eta$  defined in (2.11).

Further, for time dependent velocities  $\mathbf{u} \in C^0(I; W^{1,\infty}(\Omega))$ , it can be shown (analogously to [HH13a, Lemma 3.4]) that the ‘‘coercivity’’ condition on the velocity (3.2) ensures that the operator  $\mathcal{A}(t) := \alpha \text{id} + \mathbf{L}_{\mathbf{u}(\cdot,t)} : W \rightarrow L^2 \Lambda^k(\Omega)$  is uniformly maximal and monotone. Hence,  $\mathcal{A}(t)$  is the infinitesimal generator of a strongly continuous semigroup of contractions [Bre11, Proposition 7.1]. Then by [Paz83, Theorems 2.2 and 2.3 pp. 131-133], the family of generators  $\{\mathcal{A}(t)\}_{t \in I}$  is *stable* in the sense of Kato [Paz83, Definition 2.1, p. 130]. We can therefore revert to known results from semi-group theory for hyperbolic evolution systems [Paz83, Chapters 5.2-5.4] for well-posedness statements of (3.1). In particular, a coordinate-based representation of Lie derivatives (see Lemma 2.1.6) highlights that (3.1) falls into the class of evolution problems for the so-called Friedrichs’ symmetric operators [Fri58] and then [Kat76, pp. 143-145] gives well-posedness of (3.1) if  $\Omega = \mathbb{R}^n$  (see also [Paz83, Theorem 3.1 p. 135]).

These results require  $\mathbf{u}$  to be Lipschitz continuous in space. However, MHD solutions feature shocks that give rise to *discontinuous velocities*; discontinuous transport velocities are relevant in the context of magneto-quasistatic Maxwell’s equations, also in the limit of small magnetic diffusion.

A well-posedness theory for velocity fields with less regularity is available only for scalar advection. In [DL89] DiPerna and Lions showed well-posedness of the scalar advection problem for velocity fields  $\mathbf{u} \in L^1_{\text{loc}}(0, T; W^{1,1}(\mathbb{R}^n))$  with  $\text{div } \mathbf{u} \in L^1(0, T; L^\infty(\mathbb{R}^n))$  through the concept of renormalized solutions. Since the transport equation is linear, the idea is to regularize the velocity field and the initial data via mollification and then to consider a priori estimates for the sequence of smooth solutions of the regularized problem. Existence of weak solutions is deduced via a compactness argument. More recently, Ambrosio in [Amb04] provided an extension of this breakthrough to transport velocity fields in  $L^1_{\text{loc}}(0, T; \text{BV}_{\text{loc}}(\mathbb{R}^n))$  and  $\text{div } \mathbf{u} \in L^1(0, T; L^1_{\text{loc}}(\mathbb{R}^n))$ . Moreover, a notion of generalized flow associated with low regular velocity fields (the regular Lagrangian flow) and an extension of the characteristics theory to beyond the smooth context have been subject of investigation of several authors, see [CD08; Amb11; BC13] and references therein. To the best of our knowledge, beside the case of scalar transport, a well-posedness theory for the generalized transport problem (3.1) with low regular advection velocities has not been developed. Even though the aforementioned results have been established for nearly incompressible velocity fields (see [De 06] for a detailed overview), the assumption on the boundedness of the divergence of the velocity (absolute continuity with respect to the Lebesgue measure in the BV case) is of crucial importance for the well-posedness of the scalar advection problem. In the context of the generalized transport problem for a differential  $k$ -form, this corresponds requiring the operator  $L_{\mathbf{u}} + \mathcal{L}_{\mathbf{u}}$  to be bounded in space, which conceals a rather strong assumption on the regularity of the velocity itself, when  $k = 1, 2$ .

| $\Lambda^k(\Omega)$ | “Coercivity” condition (3.2)   |
|---------------------|--|
| $k = 0$             | $\alpha - \frac{1}{2}\text{div } \mathbf{u} \geq \alpha_0 \text{ a.e. in } \Omega$   |
| $k = 1$             | $\lambda_{\min} \left( \alpha \mathbb{I} + \frac{1}{2}(D\mathbf{u} + (D\mathbf{u})^\top) - \frac{1}{2}\text{div } \mathbf{u} \mathbb{I} \right) \geq \alpha_0 \text{ a.e. in } \Omega$ |
| $k = 2$             | $\lambda_{\min} \left( \alpha \mathbb{I} + \frac{1}{2}\text{div } \mathbf{u} \mathbb{I} - \frac{1}{2}(D\mathbf{u} + (D\mathbf{u})^\top) \right) \geq \alpha_0 \text{ a.e. in } \Omega$ |
| $k = 3$             | $\alpha + \frac{1}{2}\text{div } \mathbf{u} \geq \alpha_0 \text{ a.e. in } \Omega$   |

Table 3.1.: Expressions in vector proxies of the “coercivity” condition (3.2) on the velocity field.  $\mathbb{I}$  stands for the identity matrix in  $\mathbb{R}^{3 \times 3}$  and  $\lambda_{\min}$  denotes the minimum eigenvalue on the domain  $\Omega$ .

## 3.2. Stationary Generalized Advection Problem

The Eulerian method of lines policy applies time-stepping after discretization in space. Therefore, we first address the spatial discretization of (3.1) and we start from the *stationary* generalized advection boundary value problem for a  $k$ -form  $\omega$  on the bounded computational domain  $\Omega \subset \mathbb{R}^n$ :

$$\alpha\omega + L_{\mathbf{u}}\omega = f, \quad \text{in } \Omega, \tag{3.3a}$$

$$\text{tr } \omega = g, \quad \text{on } \Gamma_{\text{in}}, \tag{3.3b}$$

$$\text{tr } i_{\mathbf{n}}\omega = s, \quad \text{on } \Gamma_{\text{in}}, \tag{3.3c}$$

with  $f \in L^2\Lambda^k(\Omega)$ ,  $g \in L^2\Lambda^k(\Gamma_{\text{in}})$ ,  $s \in L^2\Lambda^{k-1}(\Gamma_{\text{in}})$ , and piecewise Lipschitz continuous velocity field  $\mathbf{u}$ . As stated in Section 3.1.1, if  $\mathbf{u} \in W^{1,\infty}(\Omega)$  problem (3.3) is well-posed in  $V$  under the assumption (3.2). The transport problems for the vector proxies associated with  $k$ -forms,  $0 \leq k \leq n$  can be derived using the correspondences from Table 2.2.

### Transmission Conditions

We aim for stabilized Galerkin methods that, crudely speaking, involve a penalization of suitable jumps across interfaces inside  $\Omega$ . In order to select the right jump terms, we have to understand the natural transmission conditions across an internal interface  $f \subset \Omega$  satisfied by a solution  $\omega$  of (3.3).

For smooth velocity  $\mathbf{u} \in W^{1,\infty}(\Omega)$ , the requirement  $L_{\mathbf{u}}\omega \in L^2\Lambda^k(\Omega)$  read in distributional sense, involves the transmission condition

$$\text{tr}[\mathbf{i}_{\mathbf{u}}(\omega \wedge \star\eta)]_f = 0, \quad \forall \eta \in C_0^\infty\Lambda^k(\Omega), \quad (3.4)$$

for any oriented (piecewise) smooth  $(n-1)$ -dimensional surface  $f \subset \Omega$ . This formula is a consequence of the integration by parts formula for the Lie derivative. The transmission conditions (3.4) carry over to Lipschitz continuous velocity  $\mathbf{u} \in W^{1,\infty}(\Omega)$ . Clearly, no transmission conditions are imposed across surfaces tangential to  $\mathbf{u}$  (characteristic surfaces).

In case of discontinuous velocity  $\mathbf{u}$ , an interpretation of  $L_{\mathbf{u}}\omega$  in the sense of distributions is no longer available. Therefore, at jumps of  $\mathbf{u}$  resort to a strong interpretation of  $L_{\mathbf{u}}\omega$ . Appealing to Cartan's homotopy formula (2.7),  $L_{\mathbf{u}} = d^{k-1}\mathbf{i}_{\mathbf{u}} + \mathbf{i}_{\mathbf{u}}d^k$ , we find the strong transmission conditions

$$\text{tr}[\omega]_f = \text{tr}[\mathbf{i}_{\mathbf{u}}\omega]_f = 0, \quad \forall \text{ oriented surfaces } f \subset \Omega, [\mathbf{u}]_f \neq 0, \quad (3.5)$$

from demanding  $\omega \in L^2\Lambda^k(\Omega)$ ,  $\mathbf{i}_{\mathbf{u}}\omega \in H^1\Lambda^{k-1}(\Omega)$  and  $d^k\omega \in L^2\Lambda^{k+1}(\Omega)$ .

### 3.2.1. Stabilized Galerkin Variational Formulation

Let  $\mathcal{T}_h = \{T\}$  be a cellular partition of  $\Omega \subset \mathbb{R}^n$  as in Assumption 2.4.2. Note that the forthcoming derivation of the numerical scheme can be extended to more general meshes, with the restrictions introduced in Assumption 2.4.2 needed in the analysis of Section 3.3.

The set of  $(n-1)$ -faces at the inflow boundary is defined as  $\Delta_{n-1}^{\partial,-}(\mathcal{T}_h) := \{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) : f \subset \Gamma_{\text{in}}\}$  and  $\bar{\Gamma}_{\text{in}} = \bigcup\{f : f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)\}$ , whereas  $\Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)$  is the set of faces at the outflow boundary. An oriented  $(n-1)$ -face  $f$  has a distinguished normal  $\mathbf{n}_f$ . Any face  $f$ , as part of the boundary of some element  $T \in \mathcal{T}_h$ , has either  $\mathbf{n}_f = \mathbf{n}_{T|_f}$  or  $\mathbf{n}_f = -\mathbf{n}_{T|_f}$ . Given  $\omega \in \Lambda^k(\Omega)$ , its two restrictions to  $f$  are denoted by  $\omega^+$  and  $\omega^-$ , e.g.  $\omega^+ := \omega|_{T^+}$  where element  $T^+$  has outward normal  $\mathbf{n}_f$ . Hence, we can introduce the notion of jump and average across a facet  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$  as

$$[\omega]_f := \omega^+ - \omega^-, \quad \{\omega\}_f := \frac{1}{2}(\omega^+ + \omega^-). \quad (3.6)$$

For  $f \in \Delta_{n-1}^\partial(\Omega)$  we assume  $f$  to be oriented so that  $\mathbf{n}_f$  points outwards and  $[\omega]_f = \{\omega\}_f := \omega$ .

Further, let  $\Lambda_h^k(\mathcal{T}_h)$  denote some piecewise polynomial approximation space for differential  $k$ -forms as in Section 2.4.1. Here  $\Lambda_h^k(\mathcal{T}_h)$  could be either a  $H\Lambda^k(\Omega)$ -conforming space  $\Lambda_h^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$  or a nonconforming space  $\Lambda_h^k(\mathcal{T}_h) \subset L^2\Lambda^k(\Omega)$  for which  $\Lambda_h^k(\mathcal{T}_h) \not\subset H\Lambda^k(\Omega)$ .

The method is formulated in the general framework of time-dependent velocity fields  $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$  and relies on the assumption that the possible (spatial) discontinuities of the velocity are resolved by the mesh.

**Assumption 3.2.1.** *For every  $t \in I$  we have  $\mathbf{u}(\cdot, t) \in W^{1,\infty}(T)$  for each  $T \in \mathcal{T}_h$ , that is the velocity field is assumed to be  $\mathcal{T}_h$ -piecewise Lipschitz continuous.*

This may seem to be a severe limitation but for our purposes it represents a reasonable condition in view of the fact that the velocity field is obtained from numerically solving the hydrodynamic equations in the MHD system (1.9).

Next, multiplying equation (3.3a) by a test form  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$  and applying the integration by parts rule (2.14), results in

$$(\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\omega_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T + \sum_{T \in \mathcal{T}_h} \int_{\partial T} \text{tr} \mathbf{i}_{\mathbf{u}}(\omega_h \wedge \star\eta_h) = (f, \eta_h)_\Omega, \quad \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h).$$

Applying the product rule (2.13) to the boundary terms, results in

$$(\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\omega_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} [\mathbf{i}_{\mathbf{u}}\omega_h \wedge \star\eta_h]_f + \int_f \text{tr} [\omega_h \wedge \star\mathbf{j}_{\mathbf{u}}\eta_h]_f = (f, \eta_h)_\Omega,$$

for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ . Moreover, it can be easily verified that, for all  $\mu_h \in \Lambda_h^{k-1}(\mathcal{T}_h)$ ,  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , the so-called ‘‘DG magic formula’’ holds,

$$\sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \int_f \text{tr} [\mu_h \wedge \star \eta_h]_f = \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \int_f \text{tr} (\{\mu_h\}_f \wedge \star [\eta_h]_f) + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} ([\mu_h]_f \wedge \star \{\eta_h\}_f).$$

For  $\omega \in W$  solution of problem (3.3), the transmission conditions (3.5) at the mesh facets  $\text{tr} [\omega]_f = \text{tr} [\mathbf{i}_{\mathbf{u}} \omega]_f = 0$ , for all  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ , yield the variational formulation: Find  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$  such that  $\mathbf{a}_h(\omega_h, \eta_h) = l(\eta_h)$  for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , where

$$l(\eta_h) := (f, \eta_h)_\Omega - \sum_{f \in \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f \text{tr} \mathbf{i}_{\mathbf{u}} (g \wedge \star \eta_h), \quad (3.7)$$

$$\begin{aligned} \mathbf{a}_h(\omega_h, \eta_h) := & (\alpha \omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\omega_h, \mathcal{L}_{\mathbf{u}} \eta_h)_T + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f \text{tr} \mathbf{i}_{\mathbf{u}} (\omega_h \wedge \star \eta_h) \\ & + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} (\{\mathbf{i}_{\mathbf{u}} \omega_h\}_f \wedge \star [\eta_h]_f) + \int_f \text{tr} (\{\omega_h\}_f \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f). \end{aligned} \quad (3.8)$$

As mentioned in Section 3.1, classical Galerkin finite element discretizations of advection problems suffer from instabilities. Therefore, devising stabilization techniques to counteract this limitation has been investigated widely. We consider the following stabilization operator

$$\mathbf{s}_h(\omega_h, \eta_h) := \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f c_f \text{tr} ([\mathbf{i}_{\mathbf{u}} \omega_h]_f \wedge \star [\eta_h]_f) + \int_f \bar{c}_f \text{tr} ([\omega_h]_f \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f), \quad \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h), \quad (3.9)$$

where the stabilization coefficients  $c_f$  and  $\bar{c}_f$  are constant on each  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$  but may depend on the velocity field and on the facets diameter  $h_f$ . Throughout this dissertation, the stabilization parameters are assumed to satisfy the following:

**Assumption 3.2.2.** *We assume that  $c_f$  and  $\bar{c}_f$  satisfy:  $c_f \mathbf{u} \cdot \mathbf{n}_f \geq c_0 > 0$  and  $\bar{c}_f \mathbf{u} \cdot \mathbf{n}_f \geq \bar{c}_0 > 0$  uniformly for all facets  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ .*

In particular, by considering the direction of the numerical fluxes as given by the average of the velocity field, the choice

$$c_f = \bar{c}_f = \frac{1}{2} \frac{\{\mathbf{u}\}_f \cdot \mathbf{n}_f}{|\{\mathbf{u}\}_f \cdot \mathbf{n}_f|}, \quad f \in \Delta_{n-1}^\circ(\mathcal{T}_h), \quad (3.10)$$

gives a scheme with *upwind fluxes* (see [Heu11, Remark 4.1.2] in the case  $\mathbf{u} \in W^{1,\infty}(\Omega)$ ). Indeed, from (3.8) together with (3.9) the contribution of the  $(n-1)$ -faces, for  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , reads

$$\begin{aligned} & \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} (\{\mathbf{i}_{\mathbf{u}} \omega_h\}_f \wedge \star [\eta_h]_f) + \int_f c_f \text{tr} ([\mathbf{i}_{\mathbf{u}} \omega_h]_f \wedge \star [\eta_h]_f) \\ & + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} (\{\omega_h\}_f \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f) + \int_f c_f \text{tr} ([\omega_h]_f \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f) \\ & = \frac{1}{2} \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} \left( (1 + 2c_f)(\mathbf{i}_{\mathbf{u}} \omega_h)^+ \wedge \star [\eta_h]_f \right) + \text{tr} \left( (1 - 2c_f)(\mathbf{i}_{\mathbf{u}} \omega_h)^- \wedge \star [\eta_h]_f \right) \\ & + \frac{1}{2} \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr} \left( (1 + 2c_f)\omega_h^+ \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f \right) + \text{tr} \left( (1 - 2c_f)\omega_h^- \wedge \star [\mathbf{j}_{\mathbf{u}} \eta_h]_f \right). \end{aligned}$$

Note that, since the velocity field is discontinuous, the upwind direction on the mesh skeleton may not be well-defined. Here we consider the direction of the stream as the one given by the average of the

velocity. However, other possibilities are feasible: an upwind direction given locally by the velocity field can be used, even if this choice will lead to non-unique numerical fluxes at mesh interfaces.

The evaluation of the terms in (3.8) involving the Lie derivative  $\mathcal{L}_{\mathbf{u}}\eta_h$  requires the knowledge of the first order derivatives of the velocity field  $\mathbf{u}$ . Note that since the velocity is assumed to be a smooth function in all elements  $T \in \mathcal{T}_h$ , the quantity  $(\omega_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T$  is well-defined for all  $T \in \mathcal{T}_h$ . However, as suggested in [HH13b], a different equivalent formulation of the bilinear form  $a_h(\cdot, \cdot)$  is convenient for implementation purposes.

**Proposition 3.2.3.** *The following equality holds for all  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$ ,*

$$\begin{aligned} a_h(\omega_h, \eta_h) &= (\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (i_{\mathbf{u}} d^k \omega_h, \eta_h)_T + (\omega_h, j_{\mathbf{u}} \delta^k \eta_h)_T \\ &\quad + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr}(i_{\mathbf{u}} \omega_h \wedge \star \eta_h) - \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr}(\omega_h \wedge \star j_{\mathbf{u}} \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(\{i_{\mathbf{u}} \omega_h\}_f \wedge \star [\eta_h]_f) - \int_f \text{tr}([\omega_h]_f \wedge \star \{j_{\mathbf{u}} \eta_h\}_f). \end{aligned} \quad (3.11)$$

*Proof.* By using the Leibniz rule for the exterior derivative with respect to the wedge product (2.12) and Stokes' theorem (2.3), it easily follows that

$$\int_{\partial\Omega} \text{tr}(\omega \wedge \star \mu) = (d^k \omega, \mu)_\Omega - (\omega, \delta^{k+1} \mu)_\Omega, \quad \forall \omega \in \Lambda^k(\Omega), \mu \in \Lambda^{k+1}(\Omega). \quad (3.12)$$

Hence, using (3.12) together with Cartan's homotopy formula for the adjoint of the Lie derivative  $\mathcal{L}_{\mathbf{u}}$  results in

$$(\omega_h, \delta^{k+1} j_{\mathbf{u}} \eta_h)_\Omega = (i_{\mathbf{u}} d^k \omega_h, \eta_h)_\Omega - \int_{\partial\Omega} \text{tr}(\omega_h \wedge \star j_{\mathbf{u}} \eta_h).$$

Furthermore, exploiting (2.13), yields

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} (\omega_h, \delta^{k+1} j_{\mathbf{u}} \eta_h)_T &= \sum_{T \in \mathcal{T}_h} (i_{\mathbf{u}} d^k \omega_h, \eta_h)_T - \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h)} \int_f \text{tr}(\omega_h \wedge \star j_{\mathbf{u}} \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} - \int_f \text{tr}(\{\omega_h\}_f \wedge \star [j_{\mathbf{u}} \eta_h]_f) - \int_f \text{tr}([\omega_h]_f \wedge \star \{j_{\mathbf{u}} \eta_h\}_f), \end{aligned} \quad (3.13)$$

where the outflow boundary terms can be recast as

$$\int_f \text{tr}(i_{\mathbf{u}} \omega_h \wedge \star \eta_h) = \int_f \text{tr} i_{\mathbf{u}} (\omega_h \wedge \star \eta_h) - \int_f \text{tr} (\omega_h \wedge \star j_{\mathbf{u}} \eta_h), \quad \forall f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h). \quad (3.14)$$

Finally, substituting (3.13) and (3.14) into the bilinear form (3.8) yields the conclusion.  $\square$

Note that, for spaces of  $H\Lambda^k(\Omega)$ -conforming discrete differential forms, the terms  $\text{tr}([\omega_h]_f \wedge \star \{j_{\mathbf{u}} \eta_h\}_f)$  in (3.11) and  $\bar{c}_f \text{tr}([\omega_h]_f \wedge \star [j_{\mathbf{u}} \eta_h]_f)$  in (3.9) vanish for all  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$  and all  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$ .

**Remark 3.2.4** (Lipschitz continuous velocity field  $\mathbf{u} \in W^{1,\infty}(\Omega)$ ). Let us consider the particular case of velocity fields that feature Lipschitz continuity in space, that is  $\mathbf{u} \in W^{1,\infty}(\Omega)$ . An easy computation allows to write, for all  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$

$$\sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(\{i_{\mathbf{u}} \omega_h\}_f \wedge \star [\eta_h]_f) = \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(i_{\{u\}_f} \{\omega_h\}_f \wedge \star [\eta_h]_f) + \frac{1}{4} \int_f \text{tr}(i_{[u]_f} [\omega_h]_f \wedge \star [\eta_h]_f),$$

$$\sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(\{\omega_h\}_f \wedge \star [j_{\mathbf{u}} \eta_h]_f) = \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(\{\omega_h\}_f \wedge \star j_{\{u\}_f} [\eta_h]_f) + \int_f \text{tr}(\{\omega_h\}_f \wedge \star j_{[u]_f} \{\eta_h\}_f),$$

and similarly for the stabilization terms in (3.9). Since trivially  $[\mathbf{u}]_f \equiv 0$  for all  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ , the terms involving the jump of the velocity can be dropped and the variational problem reduces to: Find  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$  such that  $\mathbf{a}_h(\omega_h, \eta_h) + \mathbf{s}_h(\omega_h, \eta_h) = l(\eta_h)$  for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , where  $l(\eta_h)$  is as in (3.7) while the stabilized bilinear form reads

$$\begin{aligned} \mathbf{a}_h(\omega_h, \eta_h) + \mathbf{s}_h(\omega_h, \eta_h) &= (\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\omega_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr } i_{\mathbf{u}}(\omega_h \wedge \star \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(i_{\mathbf{u}}\{\omega_h\}_f \wedge \star [\eta_h]_f) + \int_f \text{tr}(\{\omega_h\}_f \wedge \star j_{\mathbf{u}}[\eta_h]_f) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f c_f \text{tr}(i_{\mathbf{u}}[\omega_h]_f \wedge \star [\eta_h]_f) + \int_f \bar{c}_f \text{tr}([\omega_h]_f \wedge \star j_{\mathbf{u}}[\eta_h]_f). \end{aligned}$$

If  $c_f = \bar{c}_f$ , using (2.13) the bilinear form can be recast as

$$\begin{aligned} \mathbf{a}_h(\omega_h, \eta_h) + \mathbf{s}_h(\omega_h, \eta_h) &= (\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\omega_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr } i_{\mathbf{u}}(\omega_h \wedge \star \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr } i_{\mathbf{u}}(\{\omega_h\}_f \wedge \star [\eta_h]_f) + \int_f c_f \text{tr } i_{\mathbf{u}}([\omega_h]_f \wedge \star [\eta_h]_f), \end{aligned} \tag{3.15}$$

and the formulation in [Heu11, Equation 4.8, p. 61] is recovered. Note that, owing to the fact that  $\{\mathbf{u}\}_f = \mathbf{u}|_f$ , the choice of stabilization given in (3.10) yields a scheme with genuine upwind fluxes.

Analogously, the bilinear form corresponding to the reformulated variational problem (3.11) for  $\mathbf{u} \in W^{1,\infty}(\Omega)$  and  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$  reads:

$$\begin{aligned} \mathbf{a}_h(\omega_h, \eta_h) &= (\alpha\omega_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (i_{\mathbf{u}} d^k \omega_h, \eta_h)_T + (\omega_h, j_{\mathbf{u}} \delta^k \eta_h)_T \\ &\quad + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr}(i_{\mathbf{u}} \omega_h \wedge \star \eta_h) - \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \text{tr}(\omega_h \wedge \star j_{\mathbf{u}} \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \text{tr}(i_{\mathbf{u}} \{\omega_h\}_f \wedge \star [\eta_h]_f) - \int_f \text{tr}([\omega_h]_f \wedge \star j_{\mathbf{u}} \{\eta_h\}_f). \end{aligned} \tag{3.16}$$

Observe that the stabilized Galerkin formulation (3.7), (3.15) for Lipschitz continuous velocities  $\mathbf{u} \in W^{1,\infty}(\Omega)$  can be also derived by imposing on the mesh facets the transmission conditions (3.4).

Since the solution  $\omega \in W$  of (3.3) can only exhibit jump discontinuities across characteristics surfaces (*cf.* (3.4)),  $\mathbf{u} \cdot \mathbf{n}_f i_{\mathbf{n}_f}([\omega]_f \wedge \star v) = 0$  for all  $v \in \Lambda^k(\Omega)$  and for all oriented smooth, open,  $(n-1)$ -dimensional faces  $f \subset \Omega$ . Hence, the discrete stabilized approximation (3.7), (3.8), (3.9) is fully consistent with problem (3.3), namely

$$\mathbf{a}_h(\omega, \eta_h) + \mathbf{s}_h(\omega, \eta_h) = l(\eta_h), \quad \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h). \tag{3.17}$$

Therefore, if  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$  denotes the numerical solution of the discretized problem, the Galerkin orthogonality property  $\mathbf{a}_h(\omega - \omega_h, \eta_h) + \mathbf{s}_h(\omega - \omega_h, \eta_h) = 0$  for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$  holds.

### Stabilized Galerkin Formulation in Terms of Vector Proxies

For the sake of completeness, we present the vector proxy representation of the stabilized reformulated bilinear form (3.11), (3.9) corresponding to the variational formulation associated with the transport problem of the corresponding  $k$ -form. Tables 2.1 and 2.2 are used to establish the correspondences. Let  $\Omega \subset \mathbb{R}^3$  and let  $V_h$  be finite element spaces of vector proxies associated with the spaces  $\Lambda_h^k(\mathcal{T}_h)$

of polynomial differential  $k$ -forms on the mesh  $\mathcal{T}_h$ . Let  $w, v \in V_h$  or  $\mathbf{w}, \mathbf{v} \in V_h$  be the vector proxy representations of the  $k$ -forms  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$ :

$$\begin{aligned}
 k=0: \quad \mathbf{a}_h(w, v) + \mathbf{s}_h(w, v) &= \int_{\Omega} \alpha w v \, d\mathbf{x} + \sum_{T \in \mathcal{T}_h} \int_T \mathbf{u} \cdot \operatorname{grad} w v \, d\mathbf{x} - \sum_{f \in \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f \mathbf{u} \cdot \mathbf{n}_f w v \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} - \int_f [w]_f \{\mathbf{u} v\}_f \cdot \mathbf{n}_f \, dS + \int_f \bar{c}_f [w]_f [\mathbf{u} v]_f \cdot \mathbf{n}_f \, dS. \\
 k=1: \quad \mathbf{a}_h(\mathbf{w}, \mathbf{v}) + \mathbf{s}_h(\mathbf{w}, \mathbf{v}) &= \int_{\Omega} \alpha \mathbf{w} \cdot \mathbf{v} \, d\mathbf{x} + \sum_{T \in \mathcal{T}_h} \int_T (\operatorname{curl} \mathbf{w} \times \mathbf{u}) \cdot \mathbf{v} \, d\mathbf{x} - \int_T \mathbf{w} \cdot \mathbf{u} \operatorname{div} \mathbf{v} \, d\mathbf{x} \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\partial}(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f (\mathbf{w} \cdot \mathbf{u}) (\mathbf{v} \cdot \mathbf{n}_f) \, dS + \sum_{f \in \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f (\mathbf{w} \times \mathbf{n}_f) \cdot (\mathbf{u} \times \mathbf{v}) \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f \{\mathbf{w} \cdot \mathbf{u}\}_f [\mathbf{v}]_f \cdot \mathbf{n}_f \, dS + \int_f ([\mathbf{w}]_f \times \mathbf{n}_f) \cdot \{\mathbf{u} \times \mathbf{v}\}_f \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f c_f [\mathbf{w} \cdot \mathbf{u}]_f [\mathbf{v}]_f \cdot \mathbf{n}_f \, dS - \int_f \bar{c}_f ([\mathbf{w}]_f \times \mathbf{n}_f) \cdot [\mathbf{u} \times \mathbf{v}]_f \, dS. \\
 k=2: \quad \mathbf{a}_h(\mathbf{w}, \mathbf{v}) + \mathbf{s}_h(\mathbf{w}, \mathbf{v}) &= \int_{\Omega} \alpha \mathbf{w} \cdot \mathbf{v} \, d\mathbf{x} + \sum_{T \in \mathcal{T}_h} \int_T \mathbf{u} \operatorname{div} \mathbf{w} \cdot \mathbf{v} \, d\mathbf{x} + \int_T \mathbf{w} \cdot (\mathbf{u} \times \operatorname{curl} \mathbf{v}) \, d\mathbf{x} \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\partial}(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f (\mathbf{w} \times \mathbf{u}) \cdot (\mathbf{v} \times \mathbf{n}_f) \, dS - \sum_{f \in \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f (\mathbf{w} \cdot \mathbf{n}_f) (\mathbf{v} \cdot \mathbf{u}) \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f \{\mathbf{w} \times \mathbf{u}\}_f \cdot ([\mathbf{v}]_f \times \mathbf{n}_f) \, dS - \int_f [\mathbf{w}]_f \cdot \mathbf{n}_f \{\mathbf{u} \cdot \mathbf{v}\}_f \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f c_f [\mathbf{w} \times \mathbf{u}]_f \cdot ([\mathbf{v}]_f \times \mathbf{n}_f) \, dS + \int_f \bar{c}_f [\mathbf{w}]_f \cdot \mathbf{n}_f [\mathbf{u} \cdot \mathbf{v}]_f \, dS. \\
 k=3: \quad \mathbf{a}_h(w, v) + \mathbf{s}_h(w, v) &= \int_{\Omega} \alpha w v \, d\mathbf{x} - \sum_{T \in \mathcal{T}_h} \int_T w \mathbf{u} \cdot \operatorname{grad} v \, d\mathbf{x} - \sum_{f \in \Delta_{n-1}^{\partial}(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial, -}(\mathcal{T}_h)} \int_f \mathbf{u} \cdot \mathbf{n}_f w v \, dS \\
 &\quad + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f \{\mathbf{u} w\}_f \cdot \mathbf{n}_f [v]_f \, dS + \int_f c_f [\mathbf{u} w]_f \cdot \mathbf{n}_f [v]_f \, dS.
 \end{aligned}$$

### 3.3. Stationary Transport: Estimates for Continuous Velocity Fields

As explained in Section 3.1, a rigorous convergence analysis is only possible in the case  $\mathbf{u} \in W^{1,\infty}(\Omega)$ , for want of a well-posedness result for (3.3) with discontinuous velocity fields. Hence, all theoretical results will rely on the assumption  $\mathbf{u} \in W^{1,\infty}(\Omega)$ . Moreover, without loss of generality, we can assume the scaling  $\|\mathbf{u}\|_{L^\infty(\Omega)} = 1$ .

In the subsequent analysis, we make use of pairs of  $H\Lambda^k(\Omega)$ -conforming spaces  $\Lambda_h^k(\mathcal{T}_h)$  (2.28)-(2.29), and nonconforming spaces  $\Lambda_h^{d,k}(\mathcal{T}_h)$  (2.31)-(2.32), as in the following:

- (I)  $\Lambda_h^{d,k}(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^k(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$  with  $\mathcal{T}_h$  simplicial mesh,  $r \geq 1$ ;
- (II)  $\Lambda_h^{d,k}(\mathcal{T}_h) = \mathcal{P}_{r+1}^d \Lambda^k(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_{r+1}^- \Lambda^k(\mathcal{T}_h)$  with  $\mathcal{T}_h$  simplicial mesh,  $r \geq 0$ ;

(III)  $\Lambda_h^{d,k}(\mathcal{T}_h) = \mathcal{Q}_{r+1}^{\text{d},-}\Lambda^k(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_{r+1}^-\Lambda^k(\mathcal{T}_h)$  with  $\mathcal{T}_h$  tensor product mesh,  $r \geq 0$ .

**Remark 3.3.1.** We pay particular attention to  $H\Lambda^k(\Omega)$ -conforming trial/test spaces because they allow the straightforward Galerkin discretization of the diffusion form  $\mathbf{d}^{n-k-1} \star \mathbf{d}^k$  present in (2.17).

However, if not otherwise specified,  $\Lambda_h^k(\mathcal{T}_h)$  could be either a  $H\Lambda^k(\Omega)$ -conforming approximation space  $\mathcal{P}_r\Lambda^k(\mathcal{T}_h)$  or  $\mathcal{P}_{r+1}^-\Lambda^k(\mathcal{T}_h)$ , but also the totally discontinuous space  $\mathcal{P}_r^{\text{d}}\Lambda^k(\mathcal{T}_h)$  on a simplicial mesh and  $\mathcal{Q}_{r+1}^-\Lambda^k(\mathcal{T}_h)$  on a tensor product mesh. Note that, throughout the forthcoming manipulations, we assume that  $\Lambda_h^n(\mathcal{T}_h) = \mathcal{P}_r^{\text{d}}\Lambda^n(\mathcal{T}_h)$ .

Let  $V(h) := \Lambda_h^k(\mathcal{T}_h) + V$ . We introduce the discrete operators  $A_h, S_h : V(h) \rightarrow \Lambda_h^k(\mathcal{T}_h)$  such that  $(A_h\omega, \eta_h)_\Omega := \mathbf{a}_h(\omega, \eta_h)$  and  $(S_h\omega, \eta_h)_\Omega := \mathbf{s}_h(\omega, \eta_h)$ , for all  $\omega \in V(h)$  and  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , with  $\mathbf{a}_h(\cdot, \cdot)$  and  $\mathbf{s}_h(\cdot, \cdot)$  as in (3.15). Note that the bilinear form  $\mathbf{s}_h(\cdot, \cdot)$  associated with the stabilization operator is symmetric and nonnegative on  $V(h) \times V(h)$ . Moreover, for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ , applying (2.14) to  $\mathbf{a}_h(\eta_h, \eta_h)$ , results in

$$\begin{aligned} \mathbf{a}_h(\eta_h, \eta_h) &= (\alpha\eta_h, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\eta_h, \mathcal{L}_{\mathbf{u}}\eta_h)_T + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}(\eta_h \wedge \star \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}([\eta_h]_f \wedge \star [\eta_h]_f) \\ &= \sum_{T \in \mathcal{T}_h} \left( \eta_h, \alpha + \frac{1}{2}(\mathbf{L}_{\mathbf{u}} + \mathcal{L}_{\mathbf{u}})\eta_h \right)_T + \frac{1}{2} \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}(\eta_h \wedge \star \eta_h) \\ &\quad - \frac{1}{2} \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}(\eta_h \wedge \star \eta_h) \\ &= \frac{1}{2} \int_{\partial\Omega} |\mathbf{u} \cdot \mathbf{n}_{\partial\Omega}| \operatorname{tr} \mathbf{i}_{\mathbf{n}_{\partial\Omega}}(\eta_h \wedge \star \eta_h) - \frac{1}{2} (\Lambda\eta_h, \eta_h)_\Omega, \end{aligned} \quad (3.18)$$

where  $\Lambda := -(2\alpha \operatorname{id} + \mathbf{L}_{\mathbf{u}} + \mathcal{L}_{\mathbf{u}})$ . Let us introduce the following norms on  $V(h)$ ,

$$\|\omega\|_h^2 := \|\omega\|_{L^2\Lambda^k(\Omega)}^2 + |\omega|_h^2,$$

with

$$\begin{aligned} |\omega|_h^2 &:= \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega\|_{f,\mathbf{u}}^2 + \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega\|_{f,-\mathbf{u}}^2 + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \|[\omega]_f\|_{f,c_f\mathbf{u}}^2, \\ \|\omega\|_{f,\mathbf{u}}^2 &:= \int_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}(\omega \wedge \star \omega), \quad \text{and} \quad \|\omega\|_{f,c_f\mathbf{u}}^2 := \int_f c_f \operatorname{tr} \mathbf{i}_{\mathbf{u}}(\omega \wedge \star \omega). \end{aligned} \quad (3.19)$$

Note that these norms are well-defined in view of the definition of inflow and outflow boundary and the fact that  $\operatorname{tr} \mathbf{i}_{\mathbf{u}}(\omega \wedge \star \omega) = (\mathbf{u} \cdot \mathbf{n}_f) \operatorname{tr} \mathbf{i}_{\mathbf{n}_f}(\omega \wedge \star \omega)$  together with Assumption 3.2.2 on the stabilization coefficients  $c_f$ ,  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ . The above norms are combined into

$$\|\omega\|_* := h^{-1/2} \|\omega\|_{L^2\Lambda^k(\Omega)} + h^{1/2} |\omega|_{H\Lambda^k(\mathcal{T}_h)} + \left( \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \|\omega\|_{L^2\Lambda^k(f)}^2 \right)^{1/2} + |\omega|_h, \quad (3.20)$$

where  $|\cdot|_{H\Lambda^k(\mathcal{T}_h)}$  stands for the broken  $H\Lambda^k$ -seminorm on  $\mathcal{T}_h$ .

### 3.3.1. Averaging Operator

Convergence estimates for the spatial discretization of the stationary boundary value problem are key to analyzing the convergence of the fully discrete scheme. They hinge on stability results for the differential operator  $L_h := A_h + S_h$ . In order for these results to hold for both nonconforming and  $H\Lambda^k(\Omega)$ -conforming space discretizations, we approximate discontinuous differential forms by differential

forms in  $H\Lambda^k(\mathcal{T}_h)$ . In particular, if  $\Lambda_h^{d,k}(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h)$  are defined as in either (I), (II) or (III), for every  $\omega \in \Lambda_h^{d,k}(\mathcal{T}_h)$ , there exists  $\omega^c \in \Lambda_h^k(\mathcal{T}_h)$  such that

$$\|\omega - \omega^c\|_{L^2\Lambda^k(\Omega)}^2 \leq Ch \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \|\text{tr}[\omega]_f\|_{L^2\Lambda^k(f)}^2, \quad (3.21)$$

with constant  $C > 0$  depending only on the polynomial degree and the shape regularity of the mesh. The construction of the  $H\Lambda^k(\Omega)$ -conforming approximation  $\omega^c \in \Lambda_h^k(\mathcal{T}_h)$  is based on an averaging interpolation operator which is the extension to discrete differential  $k$ -forms of the operator introduced and studied for scalar functions in  $\mathbb{R}^3$  in [KP03] and [Hou+05, Appendix]. Since the proof of (3.21) is constructive, for the sake of conciseness, we restrict to the case  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$ . The extension to the cases  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r^-\Lambda^k(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_r^-\Lambda^k(\mathcal{T}_h)$  follows alike.

Without loss of generality, in what follows we assume that  $n < r + k - 1$ . If not, the degrees of freedom associated with some subcells of  $\mathcal{T}_h$  become superfluous and can simply be dropped in the analysis.

**Definition 3.3.2.** The averaging interpolation operator  $\mathcal{P}_h : \mathcal{P}_r^d\Lambda^k(\mathcal{T}_h) \rightarrow \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$  is defined through the degrees of freedom of  $\psi := \mathcal{P}_h(\omega) \in \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$  with respect to the (local) degrees of freedom  $\{W_{f_j,T}^\ell\}_{j,\ell}$  associated with  $\omega \in \mathcal{P}_r^d\Lambda^k(\mathcal{T}_h)$  and defined in (2.23). In particular,

- Degrees of freedom on  $n$ -simplices:  $\Psi_T^\ell := W_T^\ell$ , for all  $T \in \Delta_n(\mathcal{T}_h)$ , with  $\ell = 1, \dots, N_n$ .
- Degrees of freedom on  $j$ -simplices for  $k \leq j \leq n-1$ : for  $\ell = 1, \dots, N_j$

$$\Psi_{f_j}^\ell := \begin{cases} \frac{1}{M_j} \sum_{T \in \Delta_n(f_j)} W_{f_j,T}^\ell & \forall f_j \in \Delta_j^\circ(\mathcal{T}_h), \\ W_{f_j,T}^\ell & \forall f_j \in \Delta_j^\partial(\mathcal{T}_h) \cap \Delta_j(T), \end{cases} \quad (3.22)$$

where  $M_j := \sum_{T \in \Delta_n(f_j)} 1$ .

Note that if  $f_{n-1} \in \Delta_{n-1}(T_1) \cap \Delta_{n-1}(T_2)$ , then  $\Psi_{f_{n-1}}^\ell|_{T_1} = \Psi_{f_{n-1}}^\ell|_{T_2}$  for all  $\ell = 1, \dots, N_{n-1}$ . This means that  $\text{tr } \psi|_{T_1} = \text{tr } \psi|_{T_2}$  on every  $f_{n-1} \in \Delta_{n-1}(\mathcal{T}_h)$ .

The averaging operator introduced in Definition 3.3.2 enables to approximate discontinuous nonconforming differential forms by  $H\Lambda^k(\Omega)$ -conforming ones through the following approximation result anticipated in (3.21).

**Proposition 3.3.3.** Let  $\omega \in \mathcal{P}_r^d\Lambda^k(\mathcal{T}_h)$  and  $\mathcal{P}_h$  be the averaging operator from Definition 3.3.2. Then there exists a constant  $C > 0$  depending only on the polynomial degree  $r$  and the shape regularity of the mesh, such that

$$\|\omega - \mathcal{P}_h(\omega)\|_{L^2\Lambda^k(\Omega)}^2 \leq Ch \sum_{f_{n-1} \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \|\text{tr}[\omega]_{f_{n-1}}\|_{L^2\Lambda^k(f_{n-1})}^2.$$

*Proof.* Let  $T^+ \in \mathcal{T}_h$  be the image of the reference  $n$ -simplex  $\widehat{T}$  under the affine isomorphism  $\phi$ ,  $T^+ = \phi(\widehat{T})$ . In view of Lemma 2.4.5 and the Definition 3.3.2 of averaging operator, it holds

$$\|\widehat{\omega} - \widehat{\mathcal{P}_h(\omega)}\|_{L^2\Lambda^k(\widehat{T})}^2 \simeq \sum_{j=k}^{n-1} \sum_{f_j \in \Delta_j(T^+)} \sum_{\ell=1}^{N_j} (W_{f_j,T^+}^\ell - \Psi_{f_j}^\ell)^2.$$

Let us estimate separately each set of degrees of freedom on the  $j$ -subsimplices of  $\mathcal{T}_h$ . First, we consider the  $(n-1)$ -face  $f_{n-1} = \Delta_{n-1}^\circ(T^+) \cap \Delta_{n-1}^\circ(T^-)$ , then owing to (3.22),

$$\sum_{\ell=1}^{N_{n-1}} (W_{f_{n-1},T^+}^\ell - \Psi_{f_{n-1}}^\ell)^2 = \frac{1}{4} \sum_{\ell=1}^{N_{n-1}} (W_{f_{n-1},T^+}^\ell - W_{f_{n-1},T^-}^\ell)^2.$$

Consider now the degrees of freedom on the  $j$ -subsimplices of  $\mathcal{T}_h$ ,  $f_j \in \Delta_j^\circ(T^+)$  with  $k \leq j \leq n-2$ . Note that for every  $k \leq j \leq n-2$ ,  $M_j$  is the cardinality of the set of  $n$ -simplices of  $\mathcal{T}_h$  sharing  $f_j$  and, owing

to the shape regularity, this number is bounded uniformly in the mesh width  $h$ . By the definition in (3.22), it follows

$$\begin{aligned} \sum_{\ell=1}^{N_j} (W_{f_j, T^+}^\ell - \Psi_{f_j}^\ell)^2 &= \sum_{\ell=1}^{N_j} \left( W_{f_j, T^+}^\ell - M_j^{-1} \sum_{T \in \Delta_n^\circ(f_j)} W_{f_j, T}^\ell \right)^2 \\ &\leq C M_j^{-2} \sum_{T \in \Delta_n^\circ(f_j)} \sum_{\ell=1}^{N_j} (W_{f_j, T^+}^\ell - W_{f_j, T}^\ell)^2 \\ &\leq C \sum_{\substack{f_{n-1} \in \Delta_{n-1}^\circ(f_j) \\ f_{n-1} \subset \partial T^+ \cap \partial T^-}} \sum_{\ell=1}^{N_j} (W_{f_j, T^+}^\ell - W_{f_j, T^-}^\ell)^2. \end{aligned}$$

Moreover, by definition of trace and jump across a face  $f_{n-1} = \Delta_{n-1}^\circ(T^+) \cap \Delta_{n-1}^\circ(T^-)$ , the degrees of freedom of  $\text{tr}[\omega]_{f_{n-1}}$  are given by the set  $\{W_{f_j, T^+}^\ell - W_{f_j, T^-}^\ell\}_{j, \ell}$  for  $j = k, \dots, n-1$  and  $\ell = 1, \dots, j$ . Hence, applying Lemma 2.4.5 yields

$$\sum_{j=k}^{n-2} \sum_{f_j \in \Delta_j^\circ(f_{n-1})} \sum_{\ell=1}^{N_j} (W_{f_j, T^+}^\ell - W_{f_j, T^-}^\ell)^2 + \sum_{\ell=1}^{N_{n-1}} (W_{f_{n-1}, T^+}^\ell - W_{f_{n-1}, T^-}^\ell)^2 \leq C \|\text{tr}[\widehat{\omega}]_{\widehat{f}_{n-1}}\|_{L^2 \Lambda^k(\widehat{f}_{n-1})}^2.$$

Using a scaling argument yields the conclusion.  $\square$

### 3.3.2. A Priori Convergence Estimates

**Lemma 3.3.4.** *There exists a constant  $C_S$  depending on the stabilization coefficients  $|c_f|^{1/2}$ , the polynomial degree  $r$  and the shape regularity of the mesh, such that*

$$|\omega_h|_h \leq C_S h^{-1/2} \|\omega_h\|_{L^2 \Lambda^k(\Omega)}, \quad \forall \omega_h \in \Lambda_h^k(\mathcal{T}_h). \quad (3.23)$$

Moreover, for every  $\omega \in V(h)$

$$\|L_h \omega\|_{L^2 \Lambda^k(\Omega)} \leq C_L \|\omega\|_{L^2 \Lambda^k(\Omega)} + |\omega|_{H \Lambda^k(\mathcal{T}_h)} + C'_L h^{-1/2} |\omega|_h, \quad (3.24)$$

where the constant  $C_L$  depends on  $\alpha$  and  $|\mathbf{u}|_{W^{1,\infty}(\Omega)}$ , and the constant  $C'_L$  depends on the stabilization coefficients  $|c_f|^{1/2}$ ,  $|c_f|^{-1/2}$ , the polynomial degree  $r$  and the shape regularity of the mesh.

*Proof.* The first inequality (3.23) immediately follows from the definition of  $h$ -seminorm in (3.19) and inverse trace inequalities [Cia78, p. 146].

In order to show (3.24), let  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ . Integration by parts yields

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} (\omega, \mathbf{j}_\mathbf{u} \delta^k \eta_h)_T &= \sum_{T \in \mathcal{T}_h} (\mathbf{d}^{k-1} \mathbf{i}_\mathbf{u} \omega, \eta_h)_T - \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h)} \int_f \text{tr}(\mathbf{i}_\mathbf{u} \omega \wedge \star \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} - \int_f \text{tr}(\mathbf{i}_\mathbf{u} [\omega]_f \wedge \star \{\eta_h\}_f) - \int_f \text{tr}(\mathbf{i}_\mathbf{u} \{\omega\}_f \wedge \star [\eta_h]_f). \end{aligned}$$

Inserting into the expression of  $\mathbf{a}_h(\omega, \eta_h)$  in (3.16) (given in Proposition 3.2.3) results in

$$\begin{aligned} (L_h \omega, \eta_h)_\Omega &= (\alpha \omega, \eta_h)_\Omega + \sum_{T \in \mathcal{T}_h} (\mathbf{L}_\mathbf{u} \omega, \eta_h)_T - \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h)} \int_f \text{tr} \mathbf{i}_\mathbf{u} (\omega \wedge \star \eta_h) \\ &\quad + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} - \int_f \text{tr} \mathbf{i}_\mathbf{u} ([\omega]_f \wedge \star \{\eta_h\}_f) + \int_f c_f \text{tr} \mathbf{i}_\mathbf{u} ([\omega]_f \wedge \star [\eta_h]_f). \end{aligned} \quad (3.25)$$

Therefore, using Cauchy-Schwarz inequality gives

$$\begin{aligned}
 |(L_h\omega, \eta_h)_\Omega| &\leq (\|\alpha\|_{L^\infty(\Omega)} + |\mathbf{u}|_{W^{1,\infty}(\Omega)}) \|\omega\|_{L^2\Lambda^k(\Omega)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \\
 &+ \|\mathbf{u}\|_{L^\infty(\Omega)} |\omega|_{H\Lambda^k(\mathcal{T}_h)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} + \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega\|_{f,-\mathbf{u}} \|\eta_h\|_{f,-\mathbf{u}} \\
 &+ \sum_{f \in \Delta_{n-1}^o(\mathcal{T}_h)} \|[\omega]_f\|_{f,c_f \mathbf{u}} \|[\eta_h]_f - c_f^{-1} \{\eta_h\}_f\|_{f,c_f \mathbf{u}}.
 \end{aligned} \tag{3.26}$$

The last term can be bounded using inverse trace inequalities as follows

$$\begin{aligned}
 \sum_{f \in \Delta_{n-1}^o(\mathcal{T}_h)} \|[\eta_h]_f - c_f^{-1} \{\eta_h\}_f\|_{f,c_f \mathbf{u}}^2 &\leq \sum_{\substack{f \in \Delta_{n-1}^o(\mathcal{T}_h) \\ f = \partial T^+ \cap \partial T^-}} |c_f| \|\mathbf{u}\|_{L^\infty(\Omega)} \left( \|\eta_h\|_{L^2\Lambda^k(\partial T^+)}^2 + \|\eta_h\|_{L^2\Lambda^k(\partial T^-)}^2 \right. \\
 &\quad \left. + \|c_f^{-1} \eta_h\|_{L^2\Lambda^k(\partial T^+)}^2 + \|c_f^{-1} \eta_h\|_{L^2\Lambda^k(\partial T^-)}^2 \right) \\
 &\leq \sum_{\substack{f \in \Delta_{n-1}^o(\mathcal{T}_h) \\ f = \partial T^+ \cap \partial T^-}} |c_f| \|\mathbf{u}\|_{L^\infty(\Omega)} \max\{1, |c_f|^{-2}\} \|\eta_h\|_{L^2\Lambda^k(\partial T^+ \cup \partial T^-)}^2 \\
 &\leq \|\mathbf{u}\|_{L^\infty(\Omega)} \max_{f \in \Delta_{n-1}^o(\mathcal{T}_h)} \max\{|c_f|, |c_f|^{-1}\} \sum_{T \in \mathcal{T}_h} h^{-1} \|\eta_h\|_{L^2\Lambda^k(T)}^2.
 \end{aligned} \tag{3.27}$$

Combining (3.26) with (3.27) and using inverse trace inequalities on the boundary terms yields

$$\begin{aligned}
 |(L_h\omega, \eta_h)_\Omega| &\leq C_L \|\omega\|_{L^2\Lambda^k(\Omega)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} + \|\mathbf{u}\|_{L^\infty(\Omega)} |\omega|_{H\Lambda^k(\mathcal{T}_h)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \\
 &+ Ch^{-1/2} \|\mathbf{u}\|_{L^\infty(\Omega)}^{1/2} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega\|_{f,-\mathbf{u}} \\
 &+ Ch^{-1/2} \|\mathbf{u}\|_{L^\infty(\Omega)}^{1/2} \max_{f \in \Delta_{n-1}^o(\mathcal{T}_h)} \max\{|c_f|^{1/2}, |c_f|^{-1/2}\} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \sum_{f \in \Delta_{n-1}^o(\mathcal{T}_h)} \|[\omega]_f\|_{f,c_f \mathbf{u}} \\
 &\leq \left( C_L \|\omega\|_{L^2\Lambda^k(\Omega)} + |\omega|_{H\Lambda^k(\mathcal{T}_h)} + C'_L h^{-1/2} |\omega|_h \right) \|\eta_h\|_{L^2\Lambda^k(\Omega)},
 \end{aligned}$$

with  $C'_L > 0$  depending on  $|c_f|^{1/2}$  and  $|c_f|^{-1/2}$ , the polynomial degree and the shape regularity of the mesh. Note that the above-stated result applies to both  $H\Lambda^k(\Omega)$ -conforming and nonconforming space discretization.  $\square$

For the sake of conciseness, the following bounds on orthogonal subscales are presented in the case of full polynomial spaces of discrete differential  $k$ -forms on simplices, case (I), namely for the spaces  $\Lambda_h^{d,k}(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^k(\mathcal{T}_h)$  and  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$ . The proof for the cases (II) and on tensor product meshes (III), follows mutatis mutandis.

**Lemma 3.3.5.** *The following statements hold true:*

- (i) *If  $\Pi_h$  denotes the  $L^2$ -orthogonal projection onto  $\Lambda_h^{d,k}(\mathcal{T}_h)$ , then there exists a constant  $C_\pi$  such that for all  $\omega \in \Lambda_h^{d,k}(\mathcal{T}_h) + V$ ,  $\eta_h \in \Lambda_h^{d,k}(\mathcal{T}_h)$  with  $r \geq 1$ , it holds*

$$|(L_h(\omega - \Pi_h\omega), \eta_h)_\Omega| \leq C_\pi \|\omega - \Pi_h\omega\|_* \|\eta_h\|_h;$$

- (ii) *If  $\Pi_h$  denotes the global  $L^2$ -orthogonal projection onto  $\Lambda_h^k(\mathcal{T}_h)$ , then there exists a constant  $C'_\pi$  such that for all  $\omega \in V(h)$ ,  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$*

$$|(L_h(\omega - \Pi_h\omega), \eta_h)_\Omega| \leq C'_\pi \|\omega - \Pi_h\omega\|_* \|\eta_h\|_h;$$

where the constants  $C_\pi$  and  $C'_\pi$  depend on  $\alpha$ ,  $|\mathbf{u}|_{W^{1,\infty}(\Omega)}$ , the stabilization coefficients  $|c_f|^{1/2}$ ,  $|c_f|^{-1/2}$ , the polynomial degree  $r$  and the shape regularity of the mesh.

*Proof.* In order to show (i), we proceed as in [BMS04, Equation (45)], namely one can add to the bilinear form  $(L_h(\omega - \Pi_h\omega), \eta_h)_\Omega$  given in (3.15) the zero term  $\sum_{T \in \mathcal{T}_h} (\omega - \Pi_h\omega, \mathcal{L}_{\mathbf{u}_h}\eta_h)_T$  where  $\mathbf{u}_h$  is the  $L^2$ -projection of  $\mathbf{u} \in W^{1,\infty}(\Omega)$  onto piecewise constant vector fields. Hence, using estimates on the projection error for  $\mathbf{u}$  in the  $L^\infty$ -norm, Cauchy-Schwarz inequality and inverse inequalities, results in

$$\begin{aligned} |(L_h(\omega - \Pi_h\omega), \eta_h)_\Omega| &\leq \|\alpha\|_{L^\infty(\Omega)} \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \\ &+ (|\mathbf{u} - \mathbf{u}_h|_{W^{1,\infty}(\Omega)} + C|\mathbf{u}|_{W^{1,\infty}(\Omega)}) \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \|\eta_h\|_{L^2\Lambda^k(\Omega)} \\ &+ \|\mathbf{u}\|_{L^\infty(\Omega)}^{1/2} \max_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \max\{|c_f|^{1/2}, |c_f|^{-1/2}\} |\eta_h|_h \sum_{T \in \mathcal{T}_h} \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\partial T)} \\ &+ |\eta_h|_h \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega - \Pi_h\omega\|_{f,\mathbf{u}} \leq C_\pi \|\omega - \Pi_h\omega\|_* \|\eta_h\|_h \quad \forall \eta_h \in \Lambda_h^{d,k}(\mathcal{T}_h), \end{aligned} \quad (3.28)$$

where the interior facet terms have been bounded has in (3.27).

In the case (ii) of  $H\Lambda^k(\Omega)$ -conforming discretization, we can proceed analogously by using estimate (3.28), but the non-zero term  $\sum_{T \in \mathcal{T}_h} (\omega - \Pi_h\omega, -\mathcal{L}_{\mathbf{u}_h}\eta_h)_T$  has to be bounded. We show that for all  $\omega \in V(h)$ ,  $\eta_h \in \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$ ,

$$\sum_{T \in \mathcal{T}_h} |(\omega - \Pi_h\omega, \mathcal{L}_{\mathbf{u}_h}\eta_h)_T| \leq Ch^{-1/2} \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \|\eta_h\|_h, \quad (3.29)$$

with the constant  $C > 0$  depending only on the polynomial degree and the shape regularity of the mesh. In order to do that, we harness the fact that since  $\mathbf{u}_h$  is piecewise constant, for all  $T \in \mathcal{T}_h$ ,  $\mathbf{L}_{\mathbf{u}_h} = -\mathcal{L}_{\mathbf{u}_h}$  and we build  $H\Lambda^k(\Omega)$ -conforming approximations for each of the two terms appearing in Cartan's formula  $\mathbf{L}_{\mathbf{u}_h} = \mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k + \mathbf{d}^{k-1} \mathbf{i}_{\mathbf{u}_h}$ . In particular, in view of Proposition 3.3.3, let  $\gamma_h^{c,k} \in \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$  be the  $H\Lambda^k(\Omega)$ -conforming approximation of  $\mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k \eta_h \in \mathcal{P}_r^d\Lambda^k(\mathcal{T}_h)$  and let  $\gamma_h^{c,k-1} \in \mathcal{P}_{r+1}\Lambda^{k-1}(\mathcal{T}_h)$  be the  $H\Lambda^k(\Omega)$ -conforming approximation of  $\mathbf{i}_{\mathbf{u}_h} \eta_h \in \mathcal{P}_{r+1}^d\Lambda^{k-1}(\mathcal{T}_h)$ . Since  $\gamma_h^{c,k}, \mathbf{d}^{k-1} \gamma_h^{c,k-1} \in \mathcal{P}_r\Lambda^k(\mathcal{T}_h)$ ,

$$(\omega - \Pi_h\omega, \mathbf{L}_{\mathbf{u}_h}\eta_h)_T = \left( \omega - \Pi_h\omega, \mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k \eta_h - \gamma_h^{c,k} \right)_T + \left( \omega - \Pi_h\omega, \mathbf{d}^{k-1} (\mathbf{i}_{\mathbf{u}_h} \eta_h - \gamma_h^{c,k-1}) \right)_T,$$

and by Cauchy-Schwarz inequality and inverse inequality, one has

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} |(\omega - \Pi_h\omega, \mathbf{L}_{\mathbf{u}_h}\eta_h)_T| &\leq \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \|\mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k \eta_h - \gamma_h^{c,k}\|_{L^2\Lambda^k(\Omega)} \\ &+ Ch^{-1} \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \|\mathbf{i}_{\mathbf{u}_h} \eta_h - \gamma_h^{c,k-1}\|_{L^2\Lambda^{k-1}(\Omega)}. \end{aligned}$$

By the approximation results in Proposition 3.3.3, the projection errors can be bounded as

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} |(\omega - \Pi_h\omega, \mathbf{L}_{\mathbf{u}_h}\eta_h)_T| &\leq Ch \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \left( \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \|\text{tr} [\mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k \eta_h]_f\|_{L^2\Lambda^k(f)}^2 \right)^{1/2} \\ &+ \|\omega - \Pi_h\omega\|_{L^2\Lambda^k(\Omega)} \left( \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \|\text{tr} [\mathbf{i}_{\mathbf{u}_h} \eta_h]_f\|_{L^2\Lambda^{k-1}(f)}^2 \right)^{1/2}. \end{aligned}$$

Upper bounds for the facet terms can be derived as follows: Let us decompose the velocity field  $\mathbf{u}$  into its normal component  $\mathbf{u}_n := (\mathbf{u} \cdot \mathbf{n})\mathbf{n}$  and its tangential component  $\mathbf{u}_t := (\mathbf{n} \times \mathbf{u}) \times \mathbf{n}$ . Then we can write

$$\text{tr}(\mathbf{i}_{\mathbf{u}} \eta_h) = \text{tr}(\mathbf{i}_{\mathbf{u}_n} \eta_h + \mathbf{i}_{\mathbf{u}_t} \eta_h) = (\mathbf{u} \cdot \mathbf{n}) \text{tr}(\mathbf{i}_{\mathbf{n}} \eta_h) + \mathbf{i}_{\mathbf{u}_t} \text{tr} \eta_h, \quad \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h), \forall k. \quad (3.30)$$

If  $f = \partial T^+ \cap \partial T^-$ , using estimates on the projection error for  $\mathbf{u}$ , trace and inverse inequalities together with (3.30) and the fact that  $\mathbf{d}^k \eta_h \in H\Lambda^{k+1}(\Omega)$  owing to (2.30), results in

$$\begin{aligned} \|\text{tr} [\mathbf{i}_{\mathbf{u}_h} \mathbf{d}^k \eta_h]_f\|_{L^2\Lambda^k(f)} &\leq \|\text{tr} [\mathbf{i}_{(\mathbf{u}-\mathbf{u}_h)} \mathbf{d}^k \eta_h]_f\|_{L^2\Lambda^k(f)} + \|\text{tr} [\mathbf{i}_{\mathbf{u}} \mathbf{d}^k \eta_h]_f\|_{L^2\Lambda^k(f)} \\ &\leq \|\mathbf{u} - \mathbf{u}_h\|_{L^\infty(\Omega)} \|\mathbf{d}^k \eta_h\|_{L^2\Lambda^{k+1}(f)} + \|(\mathbf{u} \cdot \mathbf{n}_f) \text{tr} \mathbf{i}_{\mathbf{n}_f} [\mathbf{d}^k \eta_h]_f\|_{L^2\Lambda^k(f)} \\ &\leq Ch |\mathbf{u}|_{W^{1,\infty}(\Omega)} h^{-1/2} h^{-1} \|\eta_h\|_{L^2\Lambda^k(T^+ \cup T^-)} + Ch^{-1} \|[\eta_h]_f\|_{f,\mathbf{u}}. \end{aligned}$$

Similarly,  $\eta_h \in \mathcal{P}_r \Lambda^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$  implies  $[\mathbf{i}_{\mathbf{u}_t} \operatorname{tr} \eta_h]_f = 0$  for all  $f \in \Delta_{n-1}(\mathcal{T}_h)$ , hence

$$\begin{aligned} \|\operatorname{tr} [\mathbf{i}_{\mathbf{u}_h} \eta_h]_f\|_{L^2 \Lambda^{k-1}(f)} &\leq \|\operatorname{tr} [\mathbf{i}_{(\mathbf{u}-\mathbf{u}_h)} \eta_h]_f\|_{L^2 \Lambda^{k-1}(f)} + \|\operatorname{tr} [\mathbf{i}_{\mathbf{u}} \eta_h]_f\|_{L^2 \Lambda^{k-1}(f)} \\ &\leq Ch |\mathbf{u}|_{W^{1,\infty}(\Omega)} h^{-1/2} \|\eta_h\|_{L^2 \Lambda^k(T^+ \cup T^-)} + \|[\eta_h]_f\|_{f,\mathbf{u}}, \end{aligned}$$

which leads to the desired estimate (3.29). Finally, combining the estimates (3.28) and (3.29) yields

$$\begin{aligned} |(L_h(\omega - \Pi_h \omega), \eta_h)_\Omega| &\leq C_\pi \|\omega - \Pi_h \omega\|_* \|\eta_h\|_h + Ch^{-1/2} \|\omega - \Pi_h \omega\|_{L^2 \Lambda^k(\Omega)} \|\eta_h\|_h \\ &\leq C'_\pi \|\omega - \Pi_h \omega\|_* \|\eta_h\|_h. \end{aligned}$$

□

Note that (3.24) together with the definition of  $\|\cdot\|_*$  in (3.20), inverse and trace inequalities gives

$$\|L_h \omega\|_{L^2 \Lambda^k(\Omega)} \leq Ch^{-1/2} \|\omega\|_* \leq Ch^{-1} \|\omega\|_{L^2 \Lambda^k(\Omega)}, \quad \forall \omega \in V(h), \quad (3.31)$$

and  $C$  depends only on the polynomial degree and the shape regularity of the mesh.

As shown in the following (*cf.* Theorem 3.4.3), stability of second order Runge–Kutta schemes can be achieved with the standard CFL condition if the space discretization is performed with piecewise linear finite elements. Therefore, this case is tackled separately. In particular, we can establish the following estimate.

**Lemma 3.3.6.** *Let  $\Pi_h^0$  denote the  $L^2$ -orthogonal projection onto  $\mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$ . In the case of space discretization with piecewise affine elements  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1^d \Lambda^k(\mathcal{T}_h)$ , or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1 \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1^- \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^k(\mathcal{T}_h)$ , there exists a constant  $C_\pi$  which depends on  $\alpha$ ,  $|\mathbf{u}|_{W^{1,\infty}(\Omega)}$ , the stabilization coefficients  $|c_f|^{1/2}$ ,  $|c_f|^{-1/2}$  and the shape regularity of the mesh, such that for all  $\omega_h, \eta_h \in \Lambda_h^k(\mathcal{T}_h)$*

$$|(L_h \omega_h, \eta_h - \Pi_h^0 \eta_h)_\Omega| \leq C_\pi h^{-1/2} \|\omega_h\|_h \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)}.$$

*Proof.* The proof we propose is very similar to the proof in [BEF10, Lemma 2.1]. Let us first consider the case of nonconforming truly discontinuous discretization, namely  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1^d \Lambda^k(\mathcal{T}_h)$ . We use the formulation in (3.25) and proceed similarly to the proof of (i) from Lemma 3.3.5; let  $K_f := \max_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \max\{|c_f|^{1/2}, |c_f|^{-1/2}\} \in \mathbb{R}$ , it holds

$$\begin{aligned} |(L_h \omega_h, \eta_h - \Pi_h^0 \eta_h)_\Omega| &\leq \|\alpha\|_{L^\infty(\Omega)} \|\omega_h\|_{L^2 \Lambda^k(\Omega)} \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)} \\ &\quad + (|\mathbf{u} - \mathbf{u}_h|_{W^{1,\infty}(\Omega)} + C |\mathbf{u}|_{W^{1,\infty}(\Omega)}) \|\omega_h\|_{L^2 \Lambda^k(\Omega)} \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)} \\ &\quad + CK_f \|\mathbf{u}\|_{L^\infty(\Omega)}^{1/2} h^{-1/2} \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)} \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \|[\omega_h]_f\|_{f,c_f \mathbf{u}} \\ &\quad + C \|\mathbf{u}\|_{L^\infty(\Omega)}^{1/2} h^{-1/2} \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)} \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \|\omega_h\|_{f,-\mathbf{u}} \\ &\leq C_\pi h^{-1/2} \|\omega_h\|_h \|\eta_h - \Pi_h^0 \eta_h\|_{L^2 \Lambda^k(\Omega)}, \end{aligned} \quad (3.32)$$

where we have used the fact that, for  $\eta_h \in \mathcal{P}_1^d \Lambda^k(\mathcal{T}_h)$ ,  $\Pi_h^0 \eta_h \in \mathcal{P}_1^d \Lambda^k(\mathcal{T}_h)$  and  $\mathbf{L}_{\mathbf{u}_h} \omega_h \in \mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$  thus  $(\mathbf{L}_{\mathbf{u}_h} \omega_h, \eta_h - \Pi_h^0 \eta_h)_T = 0$  for all  $T \in \mathcal{T}_h$ .

In the case of  $H\Lambda^k(\Omega)$ -conforming discretization, we denote  $y_h := \eta_h - \Pi_h^0 \eta_h$ . Then, by the estimate derived in (3.32) and the fact that  $(\mathbf{L}_{\mathbf{u}_h} \omega_h, y_h)_T = 0$  for all  $T \in \mathcal{T}_h$ , it holds

$$\begin{aligned} |(L_h \omega_h, \Pi_h y_h)_\Omega| &= |(\alpha \omega_h, \Pi_h y_h)_\Omega + \sum_{T \in \mathcal{T}_h} ((\mathbf{L}_{\mathbf{u}} - \mathbf{L}_{\mathbf{u}_h}) \omega_h, \Pi_h y_h)_T \\ &\quad + \sum_{T \in \mathcal{T}_h} (\mathbf{L}_{\mathbf{u}_h} \omega_h, \Pi_h y_h)_T + \sum_{f \in \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \operatorname{tr} \mathbf{i}_{-\mathbf{u}} (\omega_h \wedge \star \Pi_h y_h) \end{aligned}$$

$$\begin{aligned}
 & + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f c_f \operatorname{tr} i_{\mathbf{u}}([\omega_h]_f \wedge \star [\Pi_h y_h]_f) - \int_f \operatorname{tr} i_{\mathbf{u}}([\omega_h]_f \wedge \star \{\Pi_h y_h\}_f) \\
 & \leq C_\pi h^{-1/2} \|\omega_h\|_h \|\Pi_h y_h\|_{L^2 \Lambda^k(\Omega)} + \sum_{T \in \mathcal{T}_h} |(\mathbf{L}_{\mathbf{u}_h} \omega_h, \Pi_h y_h)_T| \\
 & \leq C_\pi h^{-1/2} \|\omega_h\|_h \|y_h\|_{L^2 \Lambda^k(\Omega)} + \sum_{T \in \mathcal{T}_h} |(\mathbf{L}_{\mathbf{u}_h} \omega_h, \Pi_h y_h - y_h)_T|.
 \end{aligned}$$

Moreover, using the bound (3.29), there exists  $C > 0$  depending on the mesh shape regularity such that

$$\sum_{T \in \mathcal{T}_h} (\mathbf{L}_{\mathbf{u}_h} \omega_h, \Pi_h y_h - y_h)_T \leq Ch^{-1/2} \|\omega_h\|_h \|y_h - \Pi_h y_h\|_{L^2 \Lambda^k(\Omega)} \leq Ch^{-1/2} \|\omega_h\|_h \|y_h\|_{L^2 \Lambda^k(\Omega)},$$

which concludes the proof.  $\square$

As a consequence of the estimates shown in Lemmas 3.3.4 to 3.3.6, we present a convergence result for the stationary advection problem with Lipschitz continuous velocity fields. Analogous estimates were proposed in [Heu11, Theorem 4.1.8] for nonconforming differential forms in  $\mathbb{R}^n$ . The present result extends to  $H\Lambda^k(\Omega)$ -conforming discrete differential forms in  $\mathbb{R}^n$  the a priori convergence estimates derived in [BMS04, Section 5] for 0-forms in  $\mathbb{R}^3$  and in [Heu11, Theorem 4.1.13 and 4.1.14] for 1- and 2-forms in  $\mathbb{R}^3$ . Note that the numerical experiments presented in Section 3.6 for non-Lipschitz velocities indicate that the following convergence result might hold in a more general setting.

**Theorem 3.3.7.** *Let  $\alpha \in L^\infty(\Omega)$  and  $\mathbf{u} \in W^{1,\infty}(\Omega)$  in (3.1) satisfy the ‘‘coercivity’’ condition (3.2). Furthermore, let the stabilization parameters  $c_f$  fulfill the non-negativity Assumption 3.2.2. Then*

$$\mathbf{a}_h(\omega, \omega) + \mathbf{s}_h(\omega, \omega) \geq \min \left\{ \alpha_0, \frac{1}{2} \right\} \|\omega\|_h^2, \quad \forall \omega \in \Lambda_h^k(\mathcal{T}_h), 0 \leq k \leq n. \quad (3.33)$$

Moreover, if  $\omega \in H^{r+1} \Lambda^k(\Omega)$  is solution of the advection problem (3.1) and  $\omega_h \in \Lambda_h^k(\mathcal{T}_h)$  is solution of the discrete variational formulation with bilinear form given in (3.15), then

$$\|\omega - \omega_h\|_h \leq Ch^{r+1/2} \|\omega\|_{H^{r+1} \Lambda^k(\Omega)},$$

with the constant  $C > 0$  depending on  $|c_f|$ ,  $|c_f|^{-1}$ ,  $\alpha$ ,  $\mathbf{u}$ , the polynomial degree  $r$  and the shape regularity of the mesh.

*Proof.* The proof of stability (3.33) immediately follows from (3.18), the positivity condition (3.2) and the definition of the  $h$ -norm.

Let  $\Pi_h$  denote the  $L^2$ -projection into  $\Lambda_h^k(\mathcal{T}_h)$ . By stability and consistency (3.17), one has

$$\min \left\{ \alpha_0, \frac{1}{2} \right\} \|\omega_h - \Pi_h \omega\|_h^2 \leq |(\mathbf{L}_h(\omega - \Pi_h \omega), \eta_h)_\Omega|,$$

where  $\eta_h := \omega_h - \Pi_h \omega$ . We proceed as in the proof of Lemma 3.3.5 (i) to get (3.28), namely

$$\min \left\{ \alpha_0, \frac{1}{2} \right\} \|\omega_h - \Pi_h \omega\|_h^2 \leq C \left( \|\omega - \Pi_h \omega\|_{L^2 \Lambda^k(\Omega)} + \sum_{T \in \mathcal{T}_h} \|\omega - \Pi_h \omega\|_{L^2 \Lambda^k(\partial T)} \right) \|\eta_h\|_h,$$

and use a multiplicative trace inequality (see [BS08, Theorem 1.6.6]) for the interior facets terms, i.e.

$$\|\omega - \Pi_h \omega\|_{L^2 \Lambda^k(\partial T)}^2 \leq C \left( h_T^{-1} \|\omega - \Pi_h \omega\|_{L^2 \Lambda^k(T)}^2 + h_T |\omega - \Pi_h \omega|_{H^1 \Lambda^k(T)}^2 \right),$$

with  $C$  depending only on the shape of  $T$ .

Moreover, in the case of  $H\Lambda^k(\Omega)$ -conforming discrete differential forms the extra non-zero terms are bounded as in (3.29). The approximation estimates from Theorem 2.4.10,

$$\inf_{\mu_h \in \mathcal{P}_r \Lambda^k(T)} \|\omega - \mu_h\|_{L^2 \Lambda^k(T)} \leq Ch^{r+1} \|\omega\|_{H^{r+1} \Lambda^k(T)},$$

$$\inf_{\mu_h \in \mathcal{P}_r \Lambda^k(T)} \|\omega - \mu_h\|_{H^1 \Lambda^k(T)} \leq Ch^r \|\omega\|_{H^{r+1} \Lambda^k(T)},$$

for  $C > 0$  independent of  $h$ , yield the conclusion.  $\square$

**Remark 3.3.8.** For the scalar advection problem, namely when  $k = 0$ , the  $\|\cdot\|_h$ -norm reduces to the  $L^2 \Lambda^k$ -norm plus the contributions of the weakly enforced boundary conditions. The resulting weaker coercivity of the discrete problem (3.33) does not guarantee control over spurious oscillations which might occur in the presence of internal or boundary layers. Instabilities might also arise in the generalized advection problem for  $k$ -forms without stabilization ( $c_f = 0$ ). Moreover, in this latter case, only sub-optimal convergence is attained, namely  $\|\omega - \omega_h\|_{L^2 \Lambda^k(\Omega)} \leq Ch^r \|\omega\|_{H^{r+1} \Lambda^k(\Omega)}$  holds with  $C > 0$  independent of the mesh width  $h$  [BMS04, Section 6].

### 3.4. Fully Discrete Problem

In the present section, we formulate the fully discrete advection problem for a differential  $k$ -form by coupling the stabilized Galerkin spatial discretization introduced in Section 3.2 with explicit time-stepping schemes. In particular, the forward Euler method and explicit second-order and third-order Runge–Kutta (RK) schemes are investigated. We study a priori error estimates in Sobolev norms for the fully discretized problem following the approach of [LT98] and [BEF10].

On the time interval  $I = [0, T]$ , we consider a uniform partition  $\bigcup_{n=0}^{N-1} [t^n, t^{n+1}]$  for a given positive integer  $N$  and  $t^n = n\Delta t$  with uniform time step  $\Delta t$  such that  $T = N\Delta t$ . The semi-discrete problem reads: Find  $\omega_h(t) \in \Lambda_h^k(\mathcal{T}_h)$  such that

$$\begin{aligned} (\partial_t \omega_h(t), \eta_h)_\Omega + (L_h(t) \omega_h(t), \eta_h)_\Omega &= l(t)(\eta_h) & \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h), \\ (\omega_h(0), \eta_h)_\Omega &= (\omega^0, \eta_h)_\Omega & \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h), \end{aligned} \quad (3.34)$$

where the bilinear forms  $a_h(\cdot, \cdot)$ ,  $s_h(\cdot, \cdot)$  and the linear functional  $l(\cdot)$  are obtained at each time step through spatial discretization as in (3.8), (3.9) and (3.7) with forcing term  $f(\mathbf{x}, t)$  and velocity field  $\mathbf{u}(\mathbf{x}, t)$ . The semi-discrete problem (3.34) can be equivalently recast as the finite dimensional operator evolution equation

$$\partial_t \omega_h(t) + L_h(t) \omega_h(t) = F_h(t), \quad \forall t \in [0, T], \quad (3.35)$$

where  $F_h(t) \in \Lambda_h^k(\mathcal{T}_h)$  is such that  $(F_h(t), \eta_h)_\Omega = l(t)(\eta_h)$  for all  $\eta_h \in \Lambda_h^k(\mathcal{T}_h)$ .

In light of the results established in Lemmas 3.3.4 to 3.3.6, quasi-optimal convergence rates for the  $L^2 \Lambda^k$ -error in space  $L^\infty$ -error in time can be proven for smooth solutions of the problem (3.34), along the lines of the analysis proposed by Burman et al. in [BEF10]. In particular, under CFL-type conditions (see Table 3.2), the efficacy of the proposed space-time discretization lies in the fact that the anti-diffusive nature of explicit RK schemes is compensated by the artificial dissipation introduced through the stabilized spatial discretization.

|            | explicit Euler      | RK2                     | RK3               |
|------------|---------------------|-------------------------|-------------------|
| $p = 0$    | $\Delta t = O(h)$   | $\Delta t = O(h)$       | $\Delta t = O(h)$ |
| $p = 1$    | $\Delta t = O(h^2)$ | $\Delta t = O(h)$       | $\Delta t = O(h)$ |
| $p \geq 2$ | $\Delta t = O(h^2)$ | $\Delta t = O(h^{4/3})$ | $\Delta t = O(h)$ |

Table 3.2.: CFL conditions for the discretization of (3.35) with explicit Runge–Kutta time-stepping:  $p = r$  for  $\mathcal{P}_r^d \Lambda^k$  with  $r \geq 0$  and for  $\mathcal{P}_r \Lambda^k$  with  $r \geq 1$ ;  $p = r + 1$  for  $\mathcal{P}_{r+1}^- \Lambda^k$  and  $\mathcal{Q}_{r+1}^- \Lambda^k$  with  $r \geq 0$ .

In the following paragraphs, we introduce the fully discrete problem for (3.35) by explicitly stating the stages corresponding to the time-stepping. Moreover, we present the convergence results corresponding to each fully discrete scheme.

### Explicit Euler Scheme

The first order explicit Euler scheme for the problem (3.35) reads

$$\omega_h^{n+1} = \omega_h^n - \Delta t L_h^n \omega_h^n + \Delta t F_h^n, \quad (3.36)$$

where  $\omega_h^n = \omega_h(t^n)$ ,  $L_h^n := L_h(t^n)$  and  $F_h^n := F_h(t^n)$ .

**Theorem 3.4.1.** Let  $\omega \in C^0(0, T; H^{r+1} \Lambda^k(\Omega)) \cap C^2(0, T; L^2 \Lambda^k(\Omega))$  be the exact solution of (3.1) and let  $\{\omega_h^n\}_{n=1}^N \subset \Lambda_h^k(\mathcal{T}_h)$  be the discrete solution of problem (3.36). Let Assumption 3.2.2 and the monotonicity condition (3.2) for  $\mathbf{u} \in C^0(0, T; W^{1,\infty}(\Omega))$  hold true. Consider the trial spaces  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_{r+1}^- \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_{r+1}^- \Lambda^k(\mathcal{T}_h)$ . Then there exist constants  $C, C_{\text{CFL}} > 0$  depending only on the constants in Lemmas 3.3.4 to 3.3.6 and the trial/test spaces  $\Lambda_h^k(\mathcal{T}_h)$  such that, if

- (i)  $\Delta t \leq C_{\text{CFL}} h$ , for  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$ ;
- (ii)  $\Delta t \leq C_{\text{CFL}} h^2$ , for all other choices of  $\Lambda_h^k(\mathcal{T}_h)$ ;

then

$$\max_{0 \leq n \leq N} \|\omega(t^n) - \omega_h^n\|_{L^2 \Lambda^k(\Omega)} \leq C(\Delta t + h^{r+1/2}).$$

*Proof.* Let  $\Pi_h$  denote the  $L^2$ -projection onto  $\Lambda_h^k(\mathcal{T}_h)$ . Consistency (3.17) of the stationary problem (3.3) entails that the exact solution  $\omega$  of (3.1) satisfies

$$\Pi_h(\partial_t \omega)(t) + L_h(t)\omega(t) = F_h(t), \quad \forall t \in [0, T]. \quad (3.37)$$

If  $\omega^n := \omega(t^n)$ , the error generated at each stage (here one stage) of the scheme (3.36) can be written as  $\omega^n - \omega_h^n = (\omega^n - \Pi_h \omega^n) - (\omega_h^n - \Pi_h \omega^n) =: e_\pi^n - e_h^n$ . Using (3.36), the error  $e_h^n$  satisfies the equation

$$\begin{aligned} e_h^{n+1} &= e_h^n - \Pi_h(\omega^{n+1} - \omega^n) - \Delta t L_h^n e_h^n + \Delta t L_h^n e_\pi^n - \Delta t L_h^n \omega^n + \Delta t F_h^n \\ &\stackrel{(3.37)}{=} e_h^n - \Delta t L_h^n e_h^n + \Delta t L_h^n e_\pi^n - \Delta t \Pi_h \theta^n, \end{aligned} \quad (3.38)$$

where  $\theta^n := \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (t^{n+1} - t) \partial_{tt} \omega dt$ . Testing (3.38) with  $e_h^n$  yields

$$(e_h^{n+1}, e_h^n)_\Omega = \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 - \Delta t (L_h^n e_h^n, e_h^n)_\Omega + \Delta t (L_h^n e_\pi^n, e_h^n)_\Omega - \Delta t (\Pi_h \theta^n, e_h^n)_\Omega.$$

Note that,

$$\begin{aligned} 2(e_h^{n+1}, e_h^n)_\Omega &= (e_h^{n+1}, e_h^{n+1} - e_h^{n+1} + e_h^n)_\Omega + (e_h^{n+1} - e_h^n + e_h^n, e_h^n)_\Omega \\ &= \|e_h^{n+1}\|_{L^2 \Lambda^k(\Omega)}^2 - \|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2. \end{aligned} \quad (3.39)$$

Hence, applying the stability result (3.18) namely

$$(L_h \eta_h, \eta_h)_\Omega = |\eta_h|_h^2 - \frac{1}{2} (\Lambda \eta_h, \eta_h)_\Omega, \quad \forall \eta_h \in \Lambda_h^k(\mathcal{T}_h), \forall t \in (0, T),$$

yields

$$\begin{aligned} \frac{1}{2} \|e_h^{n+1}\|_{L^2 \Lambda^k(\Omega)}^2 - \frac{1}{2} \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 &= \frac{1}{2} \|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 - \Delta t |e_h^n|_h^2 + \frac{\Delta t}{2} (\Lambda e_h^n, e_h^n)_\Omega \\ &\quad + \Delta t (L_h^n e_\pi^n, e_h^n)_\Omega - \Delta t (\Pi_h \theta^n, e_h^n)_\Omega. \end{aligned} \quad (3.40)$$

The energy identity (3.40) can be bounded using Cauchy-Schwarz inequality and the ‘‘coercivity’’ hypothesis (3.2) as

$$\begin{aligned} \frac{1}{2} \|e_h^{n+1}\|_{L^2 \Lambda^k(\Omega)}^2 - \frac{1}{2} \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 &\lesssim \frac{1}{2} \|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 - \Delta t |e_h^n|_h^2 - \frac{\alpha_0}{2} \Delta t \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 \\ &\quad + \Delta t (L_h^n e_\pi^n, e_h^n)_\Omega + \Delta t \|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)} \|e_h^n\|_{L^2 \Lambda^k(\Omega)}. \end{aligned} \quad (3.41)$$

In view of Lemma 3.3.5,

$$\Delta t (L_h^n(\omega^n - \Pi_h \omega^n), e_h^n)_\Omega \leq C \Delta t \|e_\pi^n\|_* \|e_h^n\|_h = C \Delta t \|e_\pi^n\|_* (\|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + |e_h^n|_h^2)^{1/2}.$$

In order to estimate the term  $\|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2$ , we use the error equation (3.38) and the bound (3.31), to obtain

$$\begin{aligned} \|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 &\leq \Delta t^2 (\|L_h^n e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|L_h^n e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2) + \Delta t^2 \|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)}^2 \\ &\lesssim \Delta t^2 h^{-2} (\|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2) + \Delta t^2 \|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)}^2. \end{aligned}$$

The term  $\|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)}$  can be bounded as

$$\begin{aligned} |\theta^n|^2 &= \frac{1}{\Delta t^2} \left| \int_{t^n}^{t^{n+1}} (t^{n+1} - t) \partial_{tt} \omega dt \right|^2 \lesssim \Delta t \left| \int_{t^n}^{t^{n+1}} |\partial_{tt} \omega|^2 dt \right|, \\ \|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)}^2 &\leq \|\theta^n\|_{L^2 \Lambda^k(\Omega)}^2 \lesssim \Delta t^2 \max_{t \in [t^n, t^{n+1}]} \|\partial_{tt} \omega\|_{L^2 \Lambda^k(\Omega)}^2 \leq C(\omega) \Delta t^2. \end{aligned} \quad (3.42)$$

Hence, under the 2-CFL condition (Case (ii)),

$$\|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 \lesssim \Delta t (\|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2) + C(\omega) \Delta t^4. \quad (3.43)$$

In the Case (i), the bound (3.24) in Lemma 3.3.4 reduces to  $\|L_h \omega\|_{L^2 \Lambda^k(\Omega)} \leq C_L \|\omega\|_{L^2 \Lambda^k(\Omega)} + C'_L h^{-1/2} |\omega|_h$  for all  $\omega \in V(h)$ . Thus,

$$\begin{aligned} \|e_h^{n+1} - e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 &\leq \Delta t^2 (\|L_h^n e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|L_h^n e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2) + \Delta t^2 \|\Pi_h \theta^n\|_{L^2 \Lambda^k(\Omega)}^2 \\ &\lesssim \Delta t^2 (\|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \|e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2) + \Delta t^2 h^{-1} (|e_h^n|_h^2 + |e_\pi^n|_h^2) + C(\omega) \Delta t^4 \\ &\lesssim \Delta t \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \Delta t |e_h^n|_h^2 + \Delta t \|e_\pi^n\|_*^2 + C(\omega) \Delta t^4, \end{aligned} \quad (3.44)$$

under the milder CFL condition  $\Delta t \leq C_{\text{CFL}} h$ .

Therefore, using Young's inequality,  $\Delta t \leq 1$ , and substituting in (3.41) the bounds (5.19) and (3.44) (respectively (3.43)), results in

$$\|e_h^{n+1}\|_{L^2 \Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \Delta t |e_h^n|_h^2 \lesssim \Delta t \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \Delta t (\Delta t^2 + \|e_\pi^n\|_*^2 + \|e_\pi^n\|_{L^2 \Lambda^k(\Omega)}^2).$$

The conclusion follows by a discrete Gronwall type argument [Gro19] and standard estimates on the projection error  $e_\pi^n$ . □

**Remark 3.4.2.** Note that, as shown in the proof of Theorem 3.4.1, the mild time step constraint in Theorem 3.4.1 (i) valid for spatial approximations based on piecewise constants discontinuous elements (finite volume)  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$  hinges on the trivial observation that the bound (3.24) in Lemma 3.3.4 reduces to  $\|L_h \omega\|_{L^2 \Lambda^k(\Omega)} \leq C_L \|\omega\|_{L^2 \Lambda^k(\Omega)} + C'_L h^{-1/2} |\omega|_h$  for all  $\omega \in V(h)$ . This is the standard CFL condition for upwind finite volume or finite difference schemes for scalar advection problems.

### Explicit RK2 Schemes

We consider, as in [BEF10, Section 3.1], explicit Runge–Kutta scheme of order two (RK2) for the problem (3.35) of the form

$$\mu_h^n = \omega_h^n - \Delta t L_h^n \omega_h^n + \Delta t F_h^n, \quad (3.45)$$

$$\omega_h^{n+1} = \frac{1}{2} (\mu_h^n + \omega_h^n) - \frac{1}{2} \Delta t L_h^n \mu_h^n + \frac{1}{2} \Delta t \psi_h^n, \quad (3.46)$$

where  $\psi_h^n := F_h^n + \Delta t (\partial_t F_h)(t^n) + \delta_h^n$ , for  $f$  in (3.1) sufficiently smooth in time and  $\delta_h^n$  such that  $\|\delta_h^n\|_{L^2 \Lambda^k(\Omega)} \leq C \Delta t^2$ .

Similarly to the explicit Euler scheme, convergence of the fully discrete problem with second order two-stage Runge–Kutta schemes of the form (3.45), (3.46) can be established.

**Theorem 3.4.3.** Let  $\omega \in C^0(0, T; H^{r+1} \Lambda^k(\Omega)) \cap C^3(0, T; L^2 \Lambda^k(\Omega))$  be the exact solution of (3.1) and let  $\{\omega_h^n\}_{n=1}^N \subset \Lambda_h^k(\mathcal{T}_h)$  be the discrete solution of problem (3.45)-(3.46). Let Assumption 3.2.2 and the monotonicity condition (3.2) for  $\mathbf{u} \in C^0(0, T; W^{1,\infty}(\Omega))$  hold true. Consider the trial spaces  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_{r+1}^- \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_{r+1}^- \Lambda^k(\mathcal{T}_h)$ . Then there exist constants  $C, C_{\text{CFL}} > 0$  depending only on the constants in Lemmas 3.3.4 to 3.3.6 and the trial/test spaces  $\Lambda_h^k(\mathcal{T}_h)$  such that, if

- (i)  $\Delta t \leq C_{\text{CFL}} h$ , for a nonconforming spatial discretization with  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1^d \Lambda^k(\mathcal{T}_h)$  or for a H $\Lambda^k(\Omega)$ -conforming approximation with spaces  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1 \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_1^- \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^k(\mathcal{T}_h)$ ;
- (ii)  $\Delta t \leq C_{\text{CFL}} h^{4/3}$ , for all other choices of  $\Lambda_h^k(\mathcal{T}_h)$ ;

then

$$\max_{0 \leq n \leq N} \|\omega(t^n) - \omega_h^n\|_{L^2 \Lambda^k(\Omega)} \leq C(\Delta t^2 + h^{r+1/2}).$$

*Proof.* The present proof is based on energy arguments and follows the same idea as in Theorem 3.4.1. It is similar to the proofs of [BEF10, Theorem 3.1 and Theorem 3.2]. We proceed by steps.

Step 1. (Error equations) Let  $\mu(t) := \omega(t) + \Delta t(\partial_t \omega)(t)$ , the errors

$$\begin{aligned} e_\pi^n &:= \omega^n - \Pi_h \omega^n, & z_\pi^n &:= \mu^n - \Pi_h \mu^n, \\ e_h^n &:= \omega_h^n - \Pi_h \omega^n, & z_h^n &:= \mu_h^n - \Pi_h \mu^n, \end{aligned}$$

satisfy the following equations:

$$z_h^n = e_h^n - \Delta t L_h^n e_h^n + \Delta t a_h^n, \quad (3.47)$$

$$e_h^{n+1} = \frac{1}{2}(e_h^n + z_h^n) - \frac{1}{2}\Delta t L_h^n z_h^n + \frac{1}{2}\Delta t b_h^n, \quad (3.48)$$

where  $a_h^n := L_h^n e_\pi^n$  and  $b_h^n := L_h^n z_\pi^n - \Pi_h \theta^n + \delta_h^n$  with  $\theta^n := \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (t^{n+1} - t)^2 \partial_{ttt} \omega dt$ .

Indeed, we can use the consistency property (3.37) of the strong solution  $\omega$  of problem (3.1) together with the definition of  $\mu$ ; thereby,

$$\Pi_h \mu^n = \Pi_h(\omega^n + \Delta t(\partial_t \omega)(t^n)) = \Pi_h \omega^n - \Delta t L_h^n \omega^n + \Delta t F_h^n. \quad (3.49)$$

Subtracting (3.49) from (3.45) and using the linearity of  $L_h^n$ , results in  $z_h^n = e_h^n - \Delta t L_h^n (\omega_h^n - \omega^n) = e_h^n - \Delta t L_h^n e_h^n + \Delta t L_h^n e_\pi^n$ , which gives (3.47). Moreover, integrating (twice) by parts  $\theta^n$  gives

$$\theta^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (t^{n+1} - t)^2 \partial_{ttt} \omega dt = \frac{2}{\Delta t} (\omega^{n+1} - \omega^n) - 2(\partial_t \omega)(t^n) - \Delta t(\partial_{tt} \omega)(t^n),$$

which rearranged and projected onto  $\Lambda_h^k(\mathcal{T}_h)$  via  $\Pi_h$ , together with (3.49), yields

$$\begin{aligned} \Pi_h \omega^{n+1} &= \Pi_h(\omega^n + \Delta t(\partial_t \omega)(t^n)) + \frac{\Delta t^2}{2} \Pi_h(\partial_{tt} \omega)(t^n) + \frac{\Delta t}{2} \Pi_h \theta^n \\ &= \frac{\Delta t}{2} \Pi_h \mu^n + \frac{1}{2} \Pi_h \omega^n - \frac{\Delta t}{2} L_h^n \omega^n + \frac{\Delta t}{2} F_h^n + \frac{\Delta t^2}{2} \Pi_h(\partial_{tt} \omega)(t^n) + \frac{\Delta t}{2} \Pi_h \theta^n. \end{aligned}$$

Moreover, using (3.49) and the definition of  $\mu$  results in

$$\begin{aligned} \Delta t \Pi_h(\partial_{tt} \omega)(t^n) &= (\partial_t(\Pi_h \mu - \Pi_h \omega))(t^n) \\ &= (\partial_t(-\Delta t L_h \omega + \Delta t F_h))(t^n) \\ &= -L_h^n(\mu^n - \omega^n) + \Delta t(\partial_t F_h)(t^n). \end{aligned}$$

Therefore,  $2\Pi_h \omega^{n+1} = (\Pi_h \mu^n + \Pi_h \omega^n) - \Delta t L_h^n \mu^n + \Delta t(F_h^n + \Delta t(\partial_t F_h)(t^n) + \Pi_h \theta^n)$ . Subtracting from (3.46) gives (3.48).

Step 2. (Energy identity) Testing the error equations (3.47) and (3.48) in a weak form with  $e_h^n$  and  $2z_h^n$ , respectively, yields

$$(z_h^n, e_h^n)_\Omega = \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 - \Delta t (L_h^n e_h^n, e_h^n)_\Omega + \Delta t (a_h^n, e_h^n)_\Omega, \quad (3.50)$$

$$2(e_h^{n+1}, z_h^n)_\Omega = (e_h^n, z_h^n)_\Omega + \|z_h^n\|_{L^2\Lambda^k(\Omega)}^2 - \Delta t (L_h^n z_h^n, z_h^n)_\Omega + \Delta t (b_h^n, z_h^n)_\Omega. \quad (3.51)$$

Summing (3.50) and (3.51), and using (3.39), results in

$$\begin{aligned} \|e_h^{n+1}\|_{L^2\Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 &= \|e_h^{n+1} - z_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \Delta t (a_h^n, e_h^n)_\Omega + \Delta t (b_h^n, z_h^n)_\Omega \\ &\quad - \Delta t (L_h^n e_h^n, e_h^n)_\Omega - \Delta t (L_h^n z_h^n, z_h^n)_\Omega. \end{aligned}$$

Similarly to (3.40), equation (3.18) gives the following energy identity

$$\begin{aligned} \|e_h^{n+1}\|_{L^2\Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \Delta t |e_h^n|_h^2 + \Delta t |z_h^n|_h^2 &= \|e_h^{n+1} - z_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \Delta t (a_h^n, e_h^n)_\Omega \\ &\quad + \Delta t (b_h^n, z_h^n)_\Omega + \frac{\Delta t}{2} (\Lambda e_h^n, e_h^n)_\Omega + \frac{\Delta t}{2} (\Lambda z_h^n, z_h^n)_\Omega. \end{aligned} \quad (3.52)$$

Step 3. (Stability) Let us assume the 1-CFL condition namely  $\Delta t \leq C_{\text{CFL}} h$  and  $\Delta t \leq 1$ . We bound the last four terms in the energy equation (3.52). By definition of  $a_h^n$  and the bound (3.31), it holds

$$\Delta t \|a_h^n\|_{L^2\Lambda^k(\Omega)} = \Delta t \|L_h^n e_\pi^n\|_{L^2\Lambda^k(\Omega)} \leq C \Delta t h^{-1/2} \|e_\pi^n\|_* \stackrel{1\text{-CFL}}{\leq} C \Delta t^{1/2} \|e_\pi^n\|_*. \quad (3.53)$$

In view of Lemma 3.3.5,

$$\Delta t (a_h^n, e_h^n)_\Omega = \Delta t (L_h^n (\omega^n - \Pi_h \omega^n), e_h^n)_\Omega \leq C \Delta t \|e_\pi^n\|_* \|e_h^n\|_h = C \Delta t \|e_\pi^n\|_* (\|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + |e_h^n|_h^2)^{1/2}.$$

Similarly, using the definition of  $b_h^n$ , the assumptions on  $\delta_h^n$  and Lemma 3.3.5

$$\begin{aligned} \Delta t (b_h^n, z_h^n)_\Omega &\leq \Delta t (L_h^n z_\pi^n, z_h^n)_\Omega + \Delta t \|\Pi_h \theta^n\|_{L^2\Lambda^k(\Omega)} \|z_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t \|\delta_h^n\|_{L^2\Lambda^k(\Omega)} \|z_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\lesssim \Delta t \|z_\pi^n\|_* \|z_h^n\|_h + \Delta t \|\Pi_h \theta^n\|_{L^2\Lambda^k(\Omega)} \|z_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t^3 \|z_h^n\|_{L^2\Lambda^k(\Omega)}. \end{aligned}$$

Analogously to (3.42), one can show that  $\|\Pi_h \theta^n\|_{L^2\Lambda^k(\Omega)}^2 \lesssim \Delta t^4 \max_{t \in [t^n, t^{n+1}]} \|\partial_{ttt} \omega\|_{L^2\Lambda^k(\Omega)}^2 \lesssim C(\omega) \Delta t^4$ . Hence,

$$\begin{aligned} \Delta t (b_h^n, z_h^n)_\Omega &\lesssim \Delta t \|z_\pi^n\|_* (\|z_h^n\|_{L^2\Lambda^k(\Omega)}^2 + |z_h^n|_h^2)^{1/2} + C(\omega) \Delta t^3 \|z_h^n\|_{L^2\Lambda^k(\Omega)}, \\ \Delta t \|b_h^n\|_{L^2\Lambda^k(\Omega)} &\leq C \Delta t^{1/2} \|z_\pi^n\|_* + C(\omega) \Delta t^3 + \Delta t^3. \end{aligned} \quad (3.54)$$

Moreover, using the error equation (3.47) for  $z_h^n$  leads to

$$\begin{aligned} \|z_h^n\|_{L^2\Lambda^k(\Omega)} &\leq \|e_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t \|L_h^n e_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t \|a_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\stackrel{(3.31)}{\leq} \|e_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t h^{-1} \|e_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t \|a_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\stackrel{1\text{-CFL}}{\leq} (C_{\text{CFL}} + 1) \|e_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t \|a_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\stackrel{(3.53)}{\leq} C \|e_h^n\|_{L^2\Lambda^k(\Omega)} + C \Delta t^{1/2} \|e_\pi^n\|_* \\ &\stackrel{\Delta t \leq 1}{\leq} C (\|e_h^n\|_{L^2\Lambda^k(\Omega)} + \|e_\pi^n\|_*). \end{aligned} \quad (3.55)$$

Collecting the foregoing bounds and using Young's inequality

$$\begin{aligned} (a_h^n, e_h^n)_\Omega + (b_h^n, z_h^n)_\Omega &\lesssim \|e_\pi^n\|_*^2 + \frac{1}{2} \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \frac{1}{2} |e_h^n|_h^2 \\ &\quad + \|z_\pi^n\|_*^2 + \|z_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \frac{1}{2} |z_h^n|_h^2 + \frac{1}{2} \Delta t^4 \\ &\stackrel{(3.55)}{\lesssim} \frac{1}{2} |e_h^n|_h^2 + \frac{1}{2} |z_h^n|_h^2 + \left( \Delta t^4 + \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \|e_\pi^n\|_*^2 + \|z_\pi^n\|_*^2 \right). \end{aligned}$$

Furthermore, applying the ‘‘coercivity’’ condition (3.2) to bound  $(\Lambda \cdot, \cdot)_\Omega$  in (3.52) yields the estimate

$$\begin{aligned} \|e_h^{n+1}\|_{L^2\Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \frac{\Delta t}{2}|e_h^n|_h^2 + \frac{\Delta t}{2}|z_h^n|_h^2 &\leq \|e_h^{n+1} - z_h^n\|_{L^2\Lambda^k(\Omega)}^2 \\ &+ C\Delta t \left( \Delta t^4 + \|e_\pi^n\|_*^2 + \|z_\pi^n\|_*^2 + \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 \right). \end{aligned} \quad (3.56)$$

Step 4. (Convergence estimates) We first consider case (ii). Summing the error equations (3.47) and (3.48) results in

$$\begin{aligned} e_h^{n+1} - z_h^n &= \frac{1}{2}z_h^n - \frac{\Delta t}{2}L_h^n z_h^n + \frac{1}{2}\Delta t L_h^n e_h^n + \frac{1}{2}(\Delta t L_h^n e_h^n - \Delta t a_h^n - e_h^n) + \frac{\Delta t}{2}(b_h^n - a_h^n) \\ &\stackrel{(3.47)}{=} -\frac{\Delta t}{2}L_h^n z_h^n + \frac{\Delta t}{2}L_h^n e_h^n + \frac{\Delta t}{2}(b_h^n - a_h^n - \Delta t L_h^n a_h^n) + \frac{\Delta t}{2}(L_h^n z_h^n - L_h^n e_h^n + \Delta t(L_h^n)^2 e_h^n) \\ &= \frac{\Delta t^2}{2}(L_h^n)^2 e_h^n + \frac{\Delta t}{2}(b_h^n - a_h^n - \Delta t L_h^n a_h^n) =: \frac{\Delta t^2}{2}(L_h^n)^2 e_h^n + R_h^n. \end{aligned}$$

In order to bound  $R_h^n$ , we exploit inequality (3.31) together with (3.53), (3.54) and the 1-CFL condition, namely

$$\begin{aligned} \|R_h^n\|_{L^2\Lambda^k(\Omega)} &\leq \frac{\Delta t}{2}\|b_h^n\|_{L^2\Lambda^k(\Omega)} + \frac{\Delta t}{2}\|a_h^n\|_{L^2\Lambda^k(\Omega)} + \frac{\Delta t^2}{2}\|L_h^n a_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\lesssim \Delta t^{1/2}\|z_\pi^n\|_* + \Delta t^3 + \Delta t^{1/2}\|e_\pi^n\|_* + \frac{\Delta t^2}{2}h^{-1}\|a_h^n\|_{L^2\Lambda^k(\Omega)} \\ &\lesssim \Delta t^{1/2}\|z_\pi^n\|_* + \Delta t^3 + \Delta t^{1/2}\|e_\pi^n\|_* + \Delta t^{3/2}\|e_\pi^n\|_*, \end{aligned}$$

hence  $\|e_h^{n+1} - z_h^n\|_{L^2\Lambda^k(\Omega)}^2 \lesssim \Delta t^4\|(L_h^n)^2 e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \Delta t(\Delta t^5 + \|e_\pi^n\|_*^2 + \|z_\pi^n\|_*^2)$ . Moreover, owing to (3.31) and the 4/3-CFL condition

$$\Delta t^4\|(L_h^n)^2 e_h^n\|_{L^2\Lambda^k(\Omega)}^2 \leq C\Delta t^4 h^{-2}\|L_h^n e_h^n\|_{L^2\Lambda^k(\Omega)}^2 \leq C\Delta t^4 h^{-4}\|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 \leq C\Delta t\|e_h^n\|_{L^2\Lambda^k(\Omega)}^2.$$

Substituting into the stability equation (3.56) results in

$$\|e_h^{n+1}\|_{L^2\Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \frac{\Delta t}{2}|e_h^n|_h^2 + \frac{\Delta t}{2}|z_h^n|_h^2 \lesssim \Delta t\|e_h^n\|_{L^2\Lambda^k(\Omega)}^2 + \Delta t(\Delta t^4 + \|e_\pi^n\|_*^2 + \|z_\pi^n\|_*^2).$$

We now consider case (i) in Theorem 3.4.3 and proceed as follows. Let  $x_h^n := e_h^n - z_h^n$ . The error equations (3.47) and (3.48) yield

$$\begin{aligned} e_h^{n+1} - \frac{1}{2}z_h^n &= \frac{1}{2}z_h^n - \frac{\Delta t}{2}L_h^n z_h^n + \frac{\Delta t}{2}L_h^n e_h^n + \frac{\Delta t}{2}(b_h^n - a_h^n) \\ &= \frac{1}{2}z_h^n + \frac{\Delta t}{2}L_h^n x_h^n + \frac{\Delta t}{2}(b_h^n - a_h^n). \end{aligned}$$

By (3.24) in Lemma 3.3.4,

$$\|e_h^{n+1} - z_h^n\|_{L^2\Lambda^k(\Omega)} \lesssim \Delta t\|x_h^n\|_{L^2\Lambda^k(\Omega)} + \Delta t|x_h^n|_{H\Lambda^k(\mathcal{T}_h)} + \Delta t h^{-1/2}|x_h^n|_h + \Delta t\|b_h^n - a_h^n\|_{L^2\Lambda^k(\Omega)}. \quad (3.57)$$

By inverse inequality, for  $y_h^n := x_h^n - \Pi_h^0 x_h^n$  where  $\Pi_h^0$  is the  $L^2$ -orthogonal projection onto  $\mathcal{P}_0^d \Lambda^k(\mathcal{T}_h)$ ,

$$|x_h^n|_{H\Lambda^k(\mathcal{T}_h)} = |x_h^n - \Pi_h^0 x_h^n|_{H\Lambda^k(\mathcal{T}_h)} =: |y_h^n|_{H\Lambda^k(\mathcal{T}_h)} \lesssim h^{-1}\|y_h^n\|_{L^2\Lambda^k(\Omega)}.$$

Moreover, by the error equation (3.47),  $x_h^n = \Delta t L_h^n e_h^n - \Delta t a_h^n$ . Then,

$$\|y_h^n\|_{L^2\Lambda^k(\Omega)}^2 = (x_h^n, y_h^n)_\Omega - (\Pi_h^0 x_h^n, y_h^n)_\Omega + (\Pi_h^0 x_h^n, x_h^n)_\Omega = \Delta t (L_h^n e_h^n, y_h^n)_\Omega - \Delta t (a_h^n, y_h^n)_\Omega.$$

Applying Lemma 3.3.6 gives

$$\|y_h^n\|_{L^2\Lambda^k(\Omega)}^2 \leq \Delta t |(L_h^n e_h^n, y_h^n)_\Omega| + \Delta t \|a_h^n\|_{L^2\Lambda^k(\Omega)} \|y_h^n\|_{L^2\Lambda^k(\Omega)}$$

$$\lesssim \Delta t h^{-1/2} \|e_h^n\|_h \|y_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t \|a_h^n\|_{L^2 \Lambda^k(\Omega)} \|y_h^n\|_{L^2 \Lambda^k(\Omega)}.$$

Hence, using a 1-CFL condition results in

$$\begin{aligned} |x_h^n|_{H\Lambda^k(\mathcal{T}_h)} &\lesssim h^{-1} \|y_h^n\|_{L^2 \Lambda^k(\Omega)} \lesssim \Delta t h^{-3/2} (\|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + |e_h^n|_h^2)^{1/2} + \Delta t h^{-1} \|a_h^n\|_{L^2 \Lambda^k(\Omega)} \\ &\lesssim \Delta t h^{-3/2} |e_h^n|_h + h^{-1} (\Delta t \|a_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^{1/2} \|e_h^n\|_{L^2 \Lambda^k(\Omega)}). \end{aligned}$$

Furthermore, using (3.31) together with (3.53), gives

$$\Delta t \|x_h^n\|_{L^2 \Lambda^k(\Omega)} \leq \Delta t^2 \|L_h^n e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^2 \|a_h^n\|_{L^2 \Lambda^k(\Omega)} \lesssim \Delta t^2 h^{-1} \|e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^2 h^{-1/2} \|e_\pi^n\|_*.$$

Substituting the above-stated estimates into (3.57) yields

$$\begin{aligned} \|e_h^{n+1} - z_h^n\|_{L^2 \Lambda^k(\Omega)} &\lesssim \Delta t^2 h^{-1} \|e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^2 h^{-1/2} \|e_\pi^n\|_* \\ &\quad + \Delta t h^{-1} (\Delta t \|a_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^{1/2} \|e_h^n\|_{L^2 \Lambda^k(\Omega)}) + \Delta t^2 h^{-3/2} |e_h^n|_h \\ &\quad + \Delta t h^{-1/2} |e_h^n - z_h^n|_h + \Delta t \|b_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t \|a_h^n\|_{L^2 \Lambda^k(\Omega)}. \end{aligned}$$

The definition of  $\|\cdot\|_*$  and the bounds (3.53), (3.54) give

$$\begin{aligned} \|e_h^{n+1} - z_h^n\|_{L^2 \Lambda^k(\Omega)} &\lesssim (\Delta t^2 h^{-1} + \Delta t^{3/2} h^{-1}) \|e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^2 h^{-3/2} |e_h^n|_h + \Delta t h^{-1/2} |e_h^n - z_h^n|_h \\ &\quad + (\Delta t^2 h^{-1/2} + \Delta t^{3/2} h^{-1} + \Delta t^{1/2}) \|e_\pi^n\|_* + \Delta t^{1/2} \|z_\pi^n\|_* + \Delta t^3 \\ &\stackrel{\text{1-CFL}}{\lesssim} (\Delta t + \Delta t^{1/2}) \|e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^{1/2} |e_h^n|_h + \Delta t^{1/2} |e_h^n - z_h^n|_h \\ &\quad + (\Delta t^{3/2} + \Delta t^{1/2}) \|e_\pi^n\|_* + \Delta t^{1/2} \|z_\pi^n\|_* + \Delta t^3 \\ &\lesssim \Delta t^{1/2} (|e_h^n|_h + |z_h^n|_h) + \Delta t^{1/2} (\|e_\pi^n\|_* + \|z_\pi^n\|_* + \|e_h^n\|_{L^2 \Lambda^k(\Omega)} + \Delta t^{5/2}). \end{aligned}$$

Substituting into the stability estimate (3.56) results in

$$\|e_h^{n+1}\|_{L^2 \Lambda^k(\Omega)}^2 - \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \frac{\Delta t}{2} |e_h^n|_h^2 + \frac{\Delta t}{2} |z_h^n|_h^2 \lesssim \Delta t \|e_h^n\|_{L^2 \Lambda^k(\Omega)}^2 + \Delta t (\Delta t^4 + \|e_\pi^n\|_*^2 + \|z_\pi^n\|_*^2).$$

A Gronwall type argument and the approximation properties for  $e_\pi^n$  and  $z_\pi^n$  in the  $\|\cdot\|_*$ -norm yields the conclusion.  $\square$

### Explicit RK3 Schemes

The explicit Runge–Kutta scheme of order three (RK3) for the problem (3.35) as in [BEF10, Section 4.1] reads

$$\mu_h^n = \omega_h^n - \Delta t L_h^n \omega_h^n + \Delta t F_h^n, \quad (3.58)$$

$$\gamma_h^n = \frac{1}{2} (\mu_h^n + \omega_h^n) - \frac{1}{2} \Delta t L_h^n \mu_h^n + \frac{1}{2} \Delta t (F_h^n + \Delta t (\partial_t F_h)(t^n)), \quad (3.59)$$

$$\omega_h^{n+1} = \frac{1}{3} (\mu_h^n + \gamma_h^n + \omega_h^n) - \frac{1}{3} \Delta t L_h^n \gamma_h^n + \frac{1}{3} \Delta t \psi_h^n, \quad (3.60)$$

where  $\psi_h^n := F_h^n + \Delta t (\partial_t F_h)(t^n) + \frac{1}{2} \Delta t^2 (\partial_{tt} F_h)(t^n) + \delta_h^n$  with  $f$  in (3.1) sufficiently smooth in time and  $\|\delta_h^n\|_{L^2 \Lambda^k(\Omega)} \leq C \Delta t^3$ .

A detailed proof of the following theorem can be found in [BEF10, Theorem 4.1].

**Theorem 3.4.4.** *Let  $\omega \in C^0(0, T; H^{r+1} \Lambda^k(\Omega)) \cap C^4(0, T; L^2 \Lambda^k(\Omega))$  be the exact solution of (3.1) and let  $\{\omega_h^n\}_{n=1}^N \subset \Lambda_h^k(\mathcal{T}_h)$  be the discrete solution of problem (3.58), (3.59) and (3.60). Let Assumption 3.2.2 and the monotonicity condition (3.2) for  $\mathbf{u} \in C^0(0, T; W^{1,\infty}(\Omega))$  hold true. Consider the trial spaces  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_r \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{P}_{r+1}^- \Lambda^k(\mathcal{T}_h)$  or  $\Lambda_h^k(\mathcal{T}_h) = \mathcal{Q}_{r+1}^- \Lambda^k(\mathcal{T}_h)$ . Then there exist constants  $C, C_{\text{CFL}} > 0$  depending only on the constants in Lemmas 3.3.4 to 3.3.6 and the trial/test spaces  $\Lambda_h^k(\mathcal{T}_h)$  such that, under the 1-CFL condition  $\Delta t \leq C_{\text{CFL}} h$  it holds*

$$\max_{0 \leq n \leq N} \|\omega(t^n) - \omega_h^n\|_{L^2 \Lambda^k(\Omega)} \leq C (\Delta t^3 + h^{r+1/2}).$$

### 3.5. Numerical Experiments in 2D: Continuous Velocity

All numerical experiments on unstructured meshes presented in this chapter were carried out in Python using the FEniCS 1.0 (2011-12-07) library<sup>1</sup> [LMW12]. The numerical schemes on tensor product meshes were implemented in MATLAB<sup>2</sup>.

The two-dimensional problem describing the transport of differential 1-forms written in terms of vector proxies (see Table 2.2) amounts of solving

$$\begin{aligned} \partial_t \mathbf{w} + \alpha \mathbf{w} + \operatorname{grad}(\mathbf{u} \cdot \mathbf{w}) - \mathbf{u}^\perp \operatorname{div}(\mathbf{w}^\perp) &= \mathbf{f} && \text{in } \Omega \times [0, T], \\ \mathbf{w} &= \mathbf{g} && \text{on } \Gamma_{\text{in}} \times [0, T], \\ \mathbf{w}(0) &= \mathbf{w}_0 && \text{in } \Omega, \end{aligned} \quad (3.61)$$

where the symbol  $\perp$  denotes a clockwise  $\pi/2$ -rotation of a vector field in the Euclidean space.

#### Rotating Hump Problem

We consider the pure transport ( $\alpha = 0$ ) problem in the domain  $\Omega = [-1, 1]^2$  and in the time interval  $[0, 1]$ . The time-dependent Lipschitz continuous velocity field is

$$\mathbf{u}(\mathbf{x}, t) := \sin(2\pi t)(1 - x^2)(1 - y^2) \begin{pmatrix} y \\ -x \end{pmatrix},$$

so that there is no inflow boundary. The initial condition is defined as the “bump”

$$\mathbf{w}_0 := \begin{cases} (\varphi, \varphi)^\top & \text{if } x^2 + (y - 0.25)^2 < 0.25, \\ (0, 0)^\top & \text{otherwise,} \end{cases} \quad (3.62)$$

with  $\varphi(x, y) := \cos(\pi \sqrt{x^2 + (y - 0.25)^2})^4$  and the forcing term is set to zero,  $\mathbf{f} = (0, 0)^\top$ .

Since at time  $t = 1$  the exact solution coincides with the initial condition  $\mathbf{w}_0$  in (3.62), we compare the numerical solution obtained at the final time  $T = 1$  with  $\mathbf{w}_0$ . Figure 3.1 shows the  $L^2$ -error at the final time, for a numerical spatial discretization based on nonconforming fully discontinuous piecewise polynomial vector-valued functions (a) and  $H(\operatorname{curl}, \Omega)$ -conforming rotated Raviart–Thomas elements (b) of polynomial degree  $r$  and explicit Euler time-stepping. The time step is chosen in order to fulfill the CFL condition in Theorem 3.4.1. Note that the error of the fully discrete scheme attains the convergence rates predicted by the theory. Analogously, the convergence rates reported for the Heun method (Figure 3.2) and for the RK3 scheme (Figure 3.3) comply with the error behavior derived in Theorem 3.4.3 and Theorem 3.4.4, respectively.

**Remark 3.5.1.** Although we observed an increased convergence rate of the spatial discretization (also in the stationary case) compared to the predictions of the theory, our results are probably sharp. In the case of scalar transport it is well known [Pet91] that on very special meshes, sometimes called Peterson meshes, the  $L^2$ -norm estimates are sharp.

### 3.6. Numerical Experiments in 2D: Discontinuous Velocity

Lacking a sound convergence theory for the generalized advection problem in the presence of discontinuous velocity fields, the first set of experiments aims at testing the numerical performances of the stabilized Galerkin spatial discretization, proposed in Section 3.2, of the stationary advection problem corresponding to (3.3).

<sup>1</sup><http://fenicsproject.org/> (Accessed June 2016)

<sup>2</sup><http://mathworks.com/products/matlab/> (Accessed June 2016)

### 3. Stabilized Galerkin for Transient Advection of Differential Forms

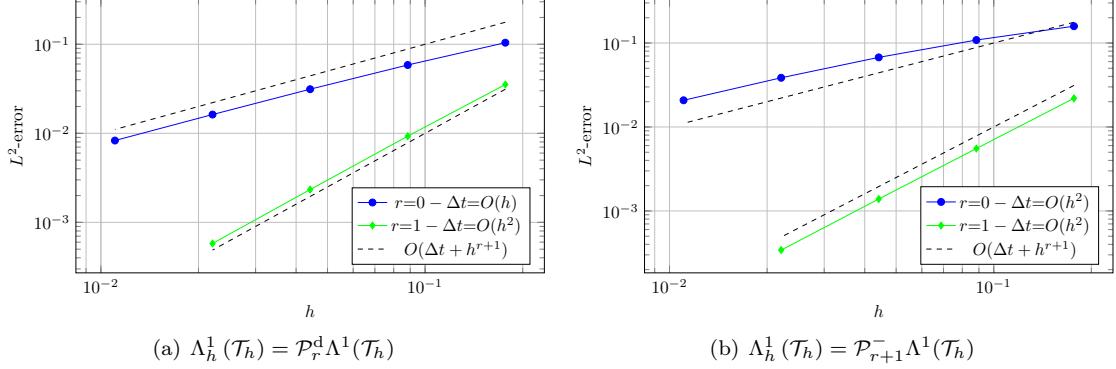


Figure 3.1.: Stabilized Galerkin discretization with upwind stabilization, polynomial differential 1-forms  $\Lambda_h^1(\mathcal{T}_h)$  and explicit Euler time-stepping.

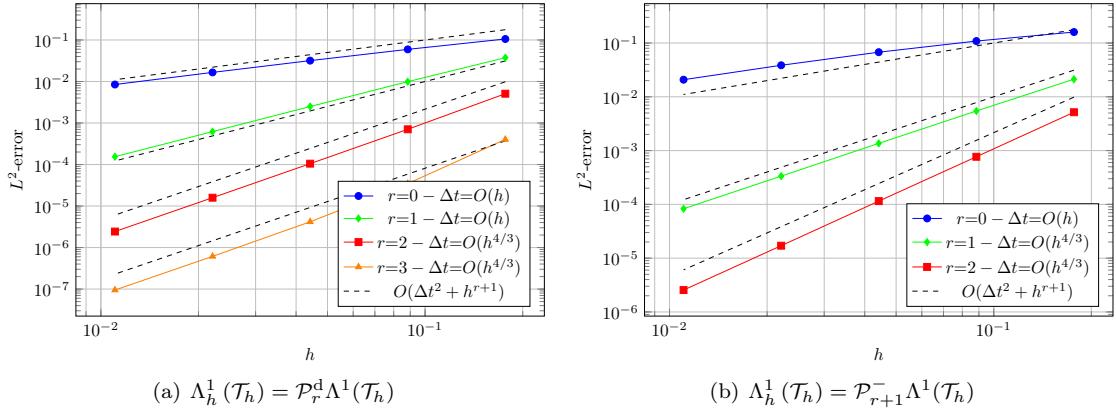


Figure 3.2.: Stabilized Galerkin discretization with upwind stabilization, polynomial differential 1-forms  $\Lambda_h^1(\mathcal{T}_h)$  and second order Heun time-stepping.

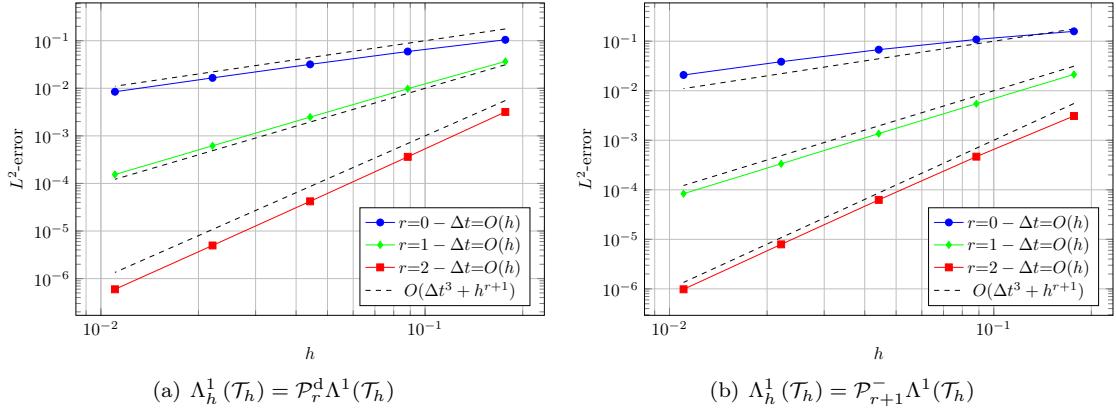


Figure 3.3.: Stabilized Galerkin discretization with upwind stabilization, polynomial differential 1-forms  $\Lambda_h^1(\mathcal{T}_h)$  and third order explicit Runge–Kutta time-stepping.

### 3.6.1. Stationary Advection of 2-Forms

The stationary advection problem for a differential 2-form in  $\Omega \subset \mathbb{R}^2$  in its vector proxy representation is nothing but the linear pure transport problem for the scalar density unknown  $\rho$ , namely

$$\begin{aligned} \alpha\rho + \operatorname{div}(\mathbf{u}\rho) &= f && \text{in } \Omega, \\ \rho &= g && \text{on } \Gamma_{\text{in}}. \end{aligned} \quad (3.63)$$

Note that, for Lipschitz continuous velocities, the stabilized Galerkin scheme (3.15),(3.7) written for (3.63) coincides with the discontinuous Galerkin discretization introduced in [HSS00; BMS04]. We consider an unstructured mesh over the domain  $\Omega = [0, 1]^2$  and the following partition into open subdomains:  $\Omega_1 = (0, 0.25) \times (0, 1)$  and  $\Omega_2 = (0.25, 1) \times (0, 1)$ . For the transport problem with  $\alpha = 2$ , we assume the velocity field  $\mathbf{u} = (u^1, u^2)$  to be defined componentwise as

$$u^1(\mathbf{x}) = \begin{cases} \sin(\pi x) & \mathbf{x} \in \Omega_1, \\ 3\sin(\pi x) & \mathbf{x} \in \Omega_2, \end{cases} \quad u^2(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega_1, \\ 3 & \mathbf{x} \in \Omega_2. \end{cases}$$

The data  $f$  and  $g$  are chosen such that the solution of the scalar transport problem (3.63) is given by the discontinuous scalar function (density)

$$\rho(x, y) = \begin{cases} -3\sin(\pi x) & \mathbf{x} \in \Omega_1, \\ -\sin(\pi x) & \mathbf{x} \in \Omega_2. \end{cases}$$

Notice that the exact solution and the given velocity field satisfy  $\mathbf{w}\rho \in H(\operatorname{div}, \Omega)$ , but  $\mathbf{w} \notin H(\operatorname{div}, \Omega)$ . We consider a conforming discretization based on DG finite element spaces of piecewise polynomial scalar functions of degree at most  $r$ , i.e.  $\Lambda_h^2(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^2(\mathcal{T}_h)$ .

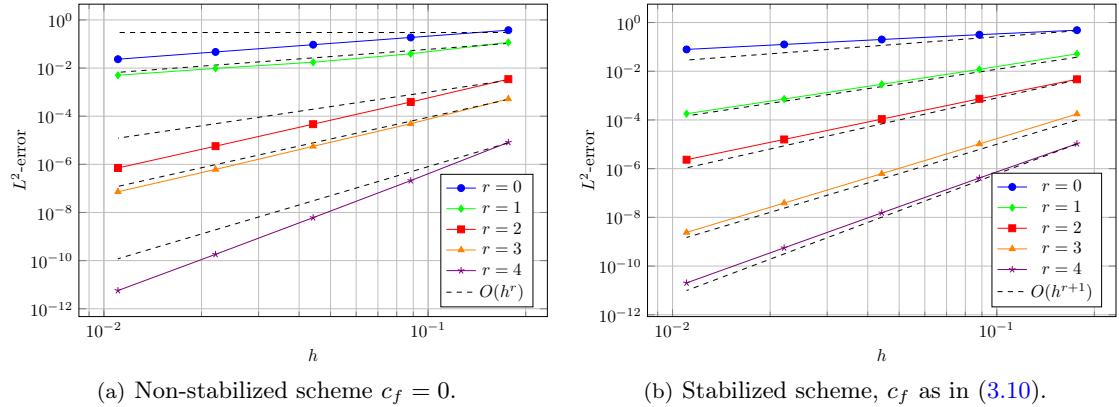


Figure 3.4.: (Conforming) DG finite elements  $\Lambda_h^2(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^2(\mathcal{T}_h)$ .

Fig. 3.4 shows the behavior of the  $L^2$ -error as the mesh is uniformly refined for different polynomial degrees. As in the experiments with Lipschitz continuous velocity field,  $r+1$  convergence rate of the  $L^2$ -error is attained by the stabilized scheme (see Fig. 3.4 (b)). On the other hand, the numerical flux obtained when the non-stabilized scheme is applied leads to sub-optimal order of accuracy  $r$  for odd polynomial degree and  $r+1$  when the polynomial degree is even (optimal rate) as shown in Fig. 3.4 (a). This behavior can be ascribed to the choice of the numerical flux and has been observed in the case of Lipschitz continuous velocity fields in [Shu01].

### 3.6.2. Stationary Advection of 1-Forms

#### Test of Convergence

Let us consider the stationary transport problems for 1-forms associated with (3.61), namely,

$$\begin{aligned} \alpha \mathbf{w} + \operatorname{grad}(\mathbf{u} \cdot \mathbf{w}) - \mathbf{u}^\perp \operatorname{div}(\mathbf{w}^\perp) &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{w} &= \mathbf{g} \quad \text{on } \Gamma_{\text{in}}. \end{aligned} \quad (3.64)$$

Here we focus on the stationary pure transport ( $\alpha = 0$ ) problem in the unit square  $\Omega = [0, 1]^2$ . We perform a set of numerical simulations on a family of unstructured meshes  $\{\mathcal{T}_h\}_h$  obtained by uniform refinement of an initial partition  $\mathcal{T}_0$  which resolves the jump discontinuity in the velocity field  $\mathbf{u} = (u^1, u^2)^\top$ . The velocity is assumed to be piecewise polynomial with respect to the open subdomain partition  $\Omega_1 = (0, 0.5) \times (0, 1)$  and  $\Omega_2 = (0.5, 1) \times (0, 1)$ . Namely,

$$u^1(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega_1, \\ 3 & \mathbf{x} \in \Omega_2, \end{cases} \quad u^2(x, y) = 2y + 1 \quad \text{in } \Omega.$$

The data  $\mathbf{f}$  and  $\mathbf{g}$  in (3.61) are chosen such that the strong solution of the BVP is given by the discontinuous vector field  $\mathbf{w}$  with components

$$w_1(\mathbf{x}) = \begin{cases} 3 & \mathbf{x} \in \Omega_1, \\ 1 & \mathbf{x} \in \Omega_2, \end{cases} \quad w_2(x, y) = (1 - x^2)(1 - y^2) \quad \text{in } \Omega.$$

Note that the exact solution is tangentially continuous and such that the forcing term  $\mathbf{f}$  belongs to  $L^2(\Omega)^2$ . We perform a numerical discretization based on:

- (i)  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{P}_{r+1}^- \Lambda^1(\mathcal{T}_h)$ , rotated Raviart–Thomas elements (Figure 3.5);
- (ii)  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{P}_r \Lambda^1(\mathcal{T}_h)$ , rotated Brezzi–Douglas–Marini (BDM) elements (Figure 3.6);
- (iii)  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ , piecewise polynomial discontinuous 1-forms (Figure 3.7).

Figure 3.5, Figure 3.6 and Figure 3.7 show the behavior of the  $L^2$ -error as the mesh is refined for the non-stabilized and stabilized Galerkin spatial scheme introduced in Section 3.2. As with Lipschitz continuous velocity fields, convergence rate  $r + 1$  of the  $L^2$ -error is attained by the stabilized scheme with edge elements of polynomial degree  $r$ . For lowest order conforming elements, the rate deteriorates by a factor of 1 when the non-stabilized scheme is applied, whereas higher order polynomial discretization yields numerical solutions which suffer of large oscillations.

Note that, for a discretization based on the full polynomial space (Case (ii), Figure 3.6), the error behaves as in the case of Lipschitz continuous velocity fields when the polynomial degree  $r$  is odd. Even polynomial degrees lead to a deteriorated convergence rate of the error in the  $L^2$ -norm.

#### Velocity with Non-Resolved Discontinuities

The derivation of the method, as of Section 3.2, relies on the assumption that the mesh resolves the possible discontinuities of the velocity field (*cf.* Assumption 3.2.1). In the following experiment we investigate an example of normal jump discontinuity in the velocity field not resolved by the mesh or any of its refinements and observe that the instabilities arising downstream of the discontinuity irremediably compromise the accuracy of the numerical solution and wreck the performance of the method. The failure of the numerical scheme in this test case may be ascribed to the fact that, since across the mesh facets the jump of the velocity vanishes, the scheme itself does not capture the discontinuity of the velocity and hence of the solution. Jump discontinuities are only taken into account through numerical quadrature.

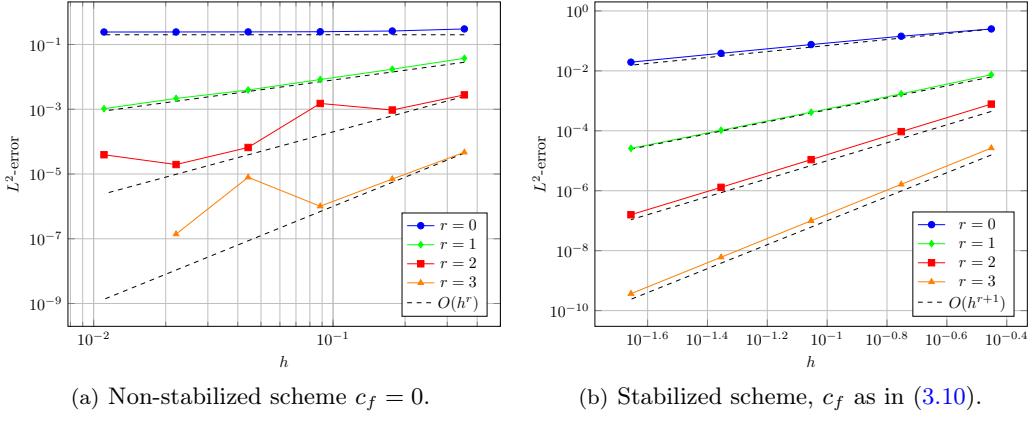
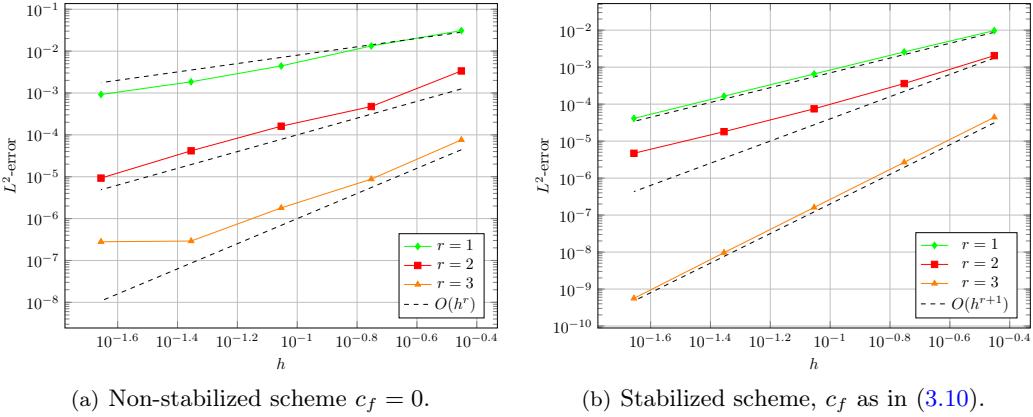
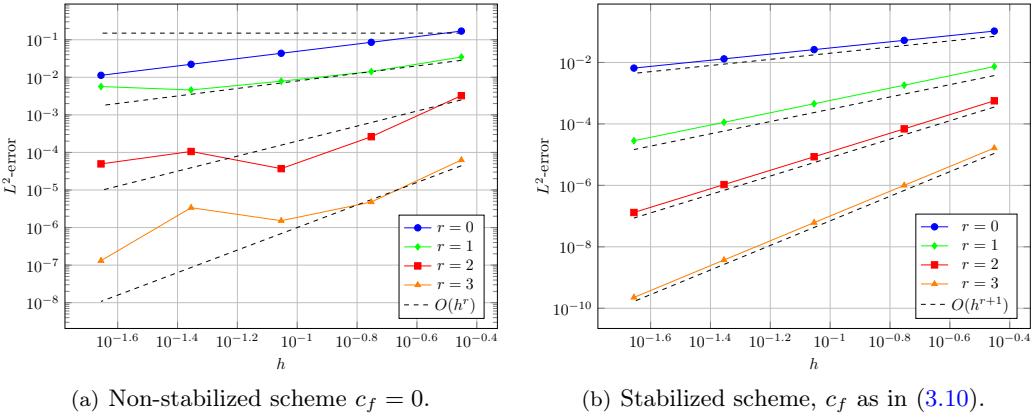

 Figure 3.5.: H(**curl**, \$\Omega\$)-conforming finite elements of the first family, \$\Lambda\_h^1(\mathcal{T}\_h) = \mathcal{P}\_{r+1}^- \Lambda^1(\mathcal{T}\_h)\$.

 Figure 3.6.: H(**curl**, \$\Omega\$)-conforming finite elements of the second family, \$\Lambda\_h^1(\mathcal{T}\_h) = \mathcal{P}\_r \Lambda^1(\mathcal{T}\_h)\$.


Figure 3.7.: Nonconforming DG finite elements \$\Lambda\_h^1(\mathcal{T}\_h) = \mathcal{P}\_r^d \Lambda^1(\mathcal{T}\_h)\$.

In more details, the stationary pure magnetic transport problem (3.64) is solved in the domain \$\Omega = [0, 1]^2\$ with a tensor product mesh and velocity field \$\mathbf{u} = (u^1, u^2)^\top\$ defined componentwise as

$$u^1(x, y) = \begin{cases} 1 & x < y, \\ 3 & x > y, \end{cases} \quad u^2 \equiv 1 \text{ in } \Omega.$$

The data  $\mathbf{f}$  and  $\mathbf{g}$  are such that the strong solution of the stationary problem corresponding to (3.61) is given by  $\mathbf{w} = (w^1, w^2)^\top$  with

$$w^1(x, y) = \begin{cases} 3 \sin(\pi x) & x < y, \\ \sin(\pi x) & x > y, \end{cases} \quad w^2 = (1 - x^2)(1 - y^2) \text{ in } \Omega,$$

as shown in Figure 3.8 (first column). The numerical discretization is performed with lowest order edge elements  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^1(\mathcal{T}_h)$  and upwind stabilization.

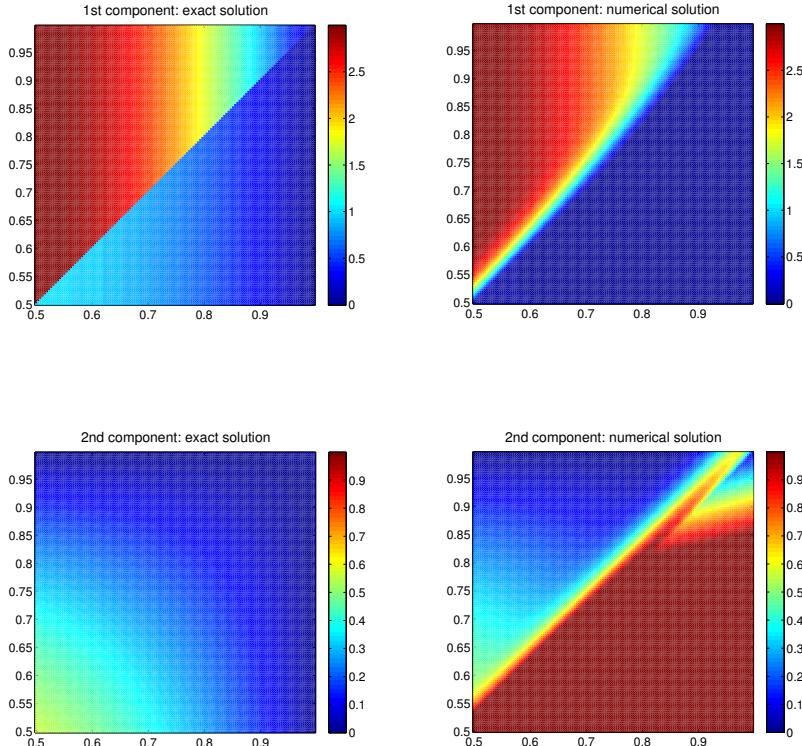


Figure 3.8.: Vector components of the exact and numerical solutions for jumps of the velocity  $\mathbf{u}$  not resolved by the mesh.

On the basis of Figure 3.8, it can be inferred that the numerical solution obtained with the upwind stabilized scheme is not affected by spurious oscillations but fails to reproduce the exact solution. An error analysis provides evidence of large errors along the discontinuity and no convergence is achieved.

### 3.6.3. Transient Advection of 1-Forms

#### Test Cases: Shear and Collisional Velocities

Encouraged by the promising performances of the scheme presented in Section 3.6.2 for the stationary advection problem in the presence of resolved discontinuous velocities, we tackle the full discretization of the transient problem.

Let us consider the two-dimensional transient magnetic advection problem (3.61) in the space domain  $\Omega = [0, 2]^2$  with periodic boundary conditions at the boundary  $\partial\Omega$  and time interval  $I = [0, 2]$ . The numerical discretization is performed on a tensor product mesh with lowest order rotated Raviart–Thomas elements  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^1(\mathcal{T}_h)$  and the second order Heun method as time integrator (in order to avoid restrictive time steps). Let  $\mathbb{1}_S$  be the characteristic function on the subset  $S \subset \Omega$ . We consider different velocity fields whose discontinuities are resolved by the domain partition  $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$  with

$\Omega_1 = (0, 1) \times (0, 2)$  and  $\Omega_2 = (1, 2) \times (0, 2)$ . Numerical simulations have been conducted with the following velocity fields and initial conditions:

- (i)  $\mathbf{u} = (0, \mathbb{1}_{\Omega_1} - \mathbb{1}_{\Omega_2})^\top$ , and  $\mathbf{w}_0(x, y) = (x(2-x)y(2-y), \sin(2\pi x))^\top$ . The initial condition belongs to  $H(\mathbf{curl}, \Omega)$ , and its component in the direction of the velocity field vanishes along the discontinuity;
- (ii)  $\mathbf{u} = (0, \mathbb{1}_{\Omega_1} - \mathbb{1}_{\Omega_2})^\top$ , and  $\mathbf{w}_0(x, y) = (\sin(2\pi x), 4)^\top$ . The initial condition is **curl-free**, while its contraction with the velocity field, namely the vector component in the velocity direction is not in  $H^1(\Omega)$ ;
- (iii)  $\mathbf{u} = (\mathbb{1}_{\Omega_1} - \mathbb{1}_{\Omega_2}, 0)^\top$ , and  $\mathbf{w}_0(x, y) = (\sin(2\pi x), 4)^\top$ . The initial condition is **curl-free**, and its component in the direction of the velocity field vanishes along the discontinuity;
- (iv)  $\mathbf{u} = (\mathbb{1}_{\Omega_1} - \mathbb{1}_{\Omega_2}, 0)^\top$ , and  $\mathbf{w}_0(x, y) = (x(2-x)y(2-y), \sin(2\pi x))^\top$ . The initial condition is in  $H(\mathbf{curl}, \Omega)$ , and its component in the direction of the velocity field is not in  $H^1(\Omega)$ .

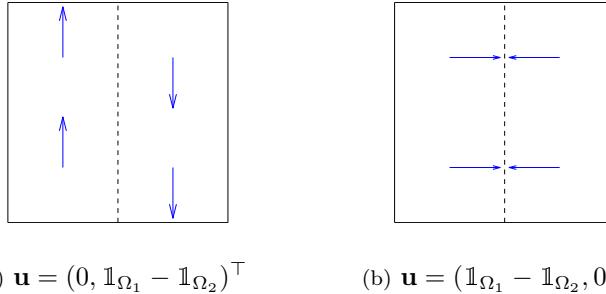


Figure 3.9.: Sketch of shear velocity (a) and collisional velocity (b).

Note that, even in the presence of simple discontinuities of the velocity field (as in the present test cases), a naive approximation of such discontinuities, say by averaging, would yield incorrect treatment of the problem close to the interface. The numerical scheme we proposed in (3.15) bears instead the full information on the velocity.

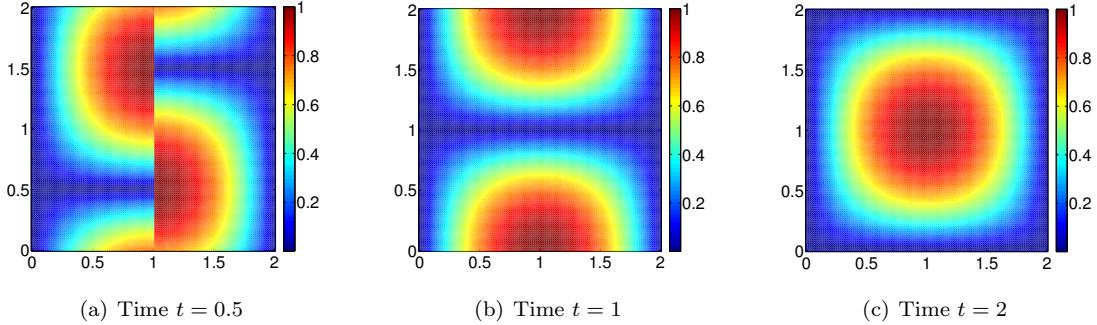


Figure 3.10.: First component of the numerical solution obtained from the stabilized scheme ( $c_f$  as in (3.10)) with  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^1(\mathcal{T}_h)$  and Heun time-stepping ( $\Delta t = 0.1h$ ), for shear velocity and initial condition as in (i).

For the Case (i), we run the simulation for an entire period, namely until  $T = 2$  and compare the solution with the initial condition. Figure 3.10 shows that the initial datum is transported in the two different subdomains  $\Omega_1$  and  $\Omega_2$ , and the  $L^2$ -error computed at the final time reaches the expected first order convergence (see Figure 3.11).

In Case (iv), even if the initial condition is smooth, the magnetic advection with normally discontinuous collisional velocity yields the formation of a shock along the discontinuity (Figure 3.12 (b)) until complete blow-up. An analogous behavior of the numerical solution obtained from the stabilized scheme can be reported in the Case (ii), where instantaneous blow-up of the solution along the discontinuity is observed. However, we expect a blow-up of the solution in these situations: the observed behavior of the

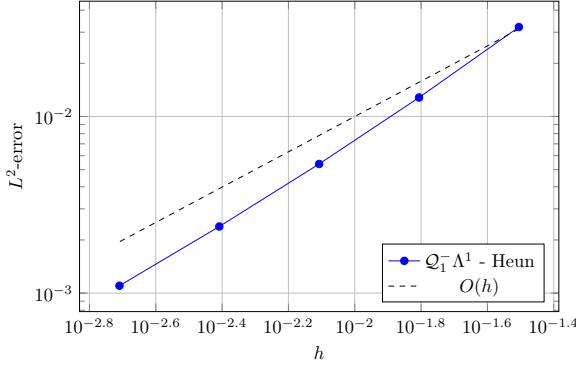
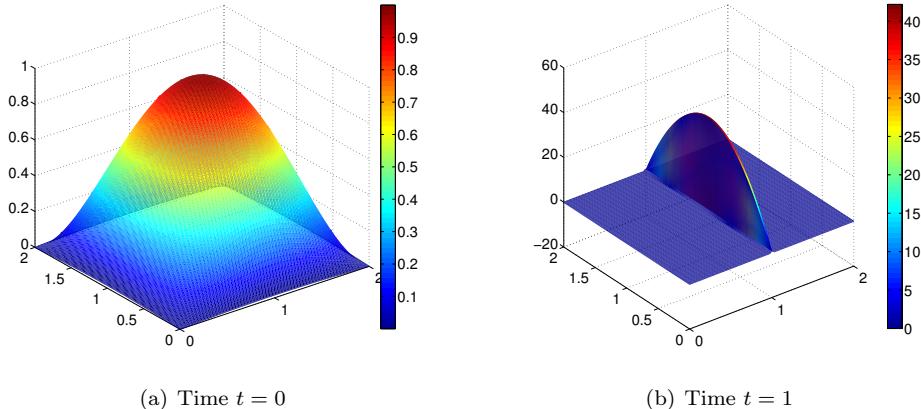


Figure 3.11.:  $L^2$ -error at time  $T=2$  of the fully discrete problem:  $\Lambda_h^1(\mathcal{T}_h) = Q_1^- \Lambda^1(\mathcal{T}_h)$  in space and Heun time-stepping. Shear velocity and initial condition as in (i).

numerical solution is not engendered by instabilities caused by the numerical scheme, but accurately reflects “physical reality”.



(a) Time  $t = 0$

(b) Time  $t = 1$

Figure 3.12.: First component of the numerical solution obtained from the stabilized scheme ( $c_f$  as in (3.10)) with  $\Lambda_h^1(\mathcal{T}_h) = Q_1^- \Lambda^1(\mathcal{T}_h)$  and Heun method ( $\Delta t = 0.1h$ ), for collisional velocity and smooth initial condition as in (iv).

A similar conclusion can be drawn for the Case (iii) in Figure 3.13.

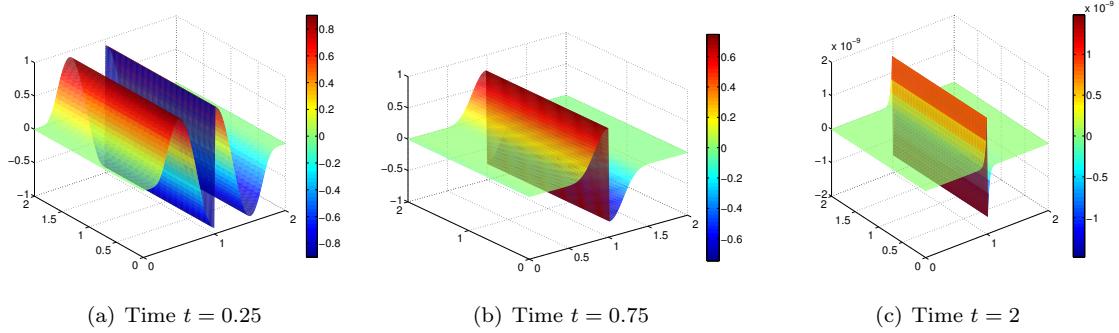


Figure 3.13.: First component of the magnetic induction field obtained from the stabilized scheme ( $c_f$  as in (3.10)) with  $\Lambda_h^1(\mathcal{T}_h) = Q_1^- \Lambda^1(\mathcal{T}_h)$  and Heun method ( $\Delta t = 0.1h$ ), for velocity and initial condition as in (iii).

### Orszag–Tang Benchmark with Given Velocity Field

As last numerical experiment, we test the discretization of the magnetic advection problem (3.61) on a widely used two-dimensional benchmark problem for MHD flows, the so-called Orszag–Tang vortex system, see Section 7.1.4 for further details.

The problem is set in the two-dimensional domain  $\Omega = [0, 2]^2$  with periodic boundary conditions at the boundary  $\partial\Omega$ . On the time interval  $I = [0, 1]$  we consider the uniform time step  $\Delta t = 5 \cdot 10^{-3}$ . The initial condition is the smooth vector field  $\mathbf{w}_0(x, y) = (\sin(2\pi x), \sin(\pi y))^\top$  and the velocity field is piecewise constant with respect to the mesh. In particular, it is given at each time step as the outcome of a high order finite volume simulation of the full MHD system obtained with the ALSVID-UQ 3.0 (2014-03-20) code<sup>3</sup> (using a three-wave HLL solver and modified WENO reconstructions in order to keep the pressure and density positive). Note that even if the initial velocity is smooth, complex structures such as shocks and shock interactions develop in time. Concerning the discretization, the numerical scheme has been implemented on a tensor product mesh with  $200 \times 200$  elements and  $H(\mathbf{curl}, \Omega)$ -conforming lowest order rotated Raviart–Thomas elements  $\Lambda_h^1(\mathcal{T}_h) = Q_1^- \Lambda^1(\mathcal{T}_h)$  have been used for the spatial discretization, while a second order two-stage Runge–Kutta time-stepping is deployed in order to exploit the mild CFL condition of the scheme. The stabilization parameter is as in (3.10), i.e., the upwind direction is assumed to be given by the average of the velocity field.

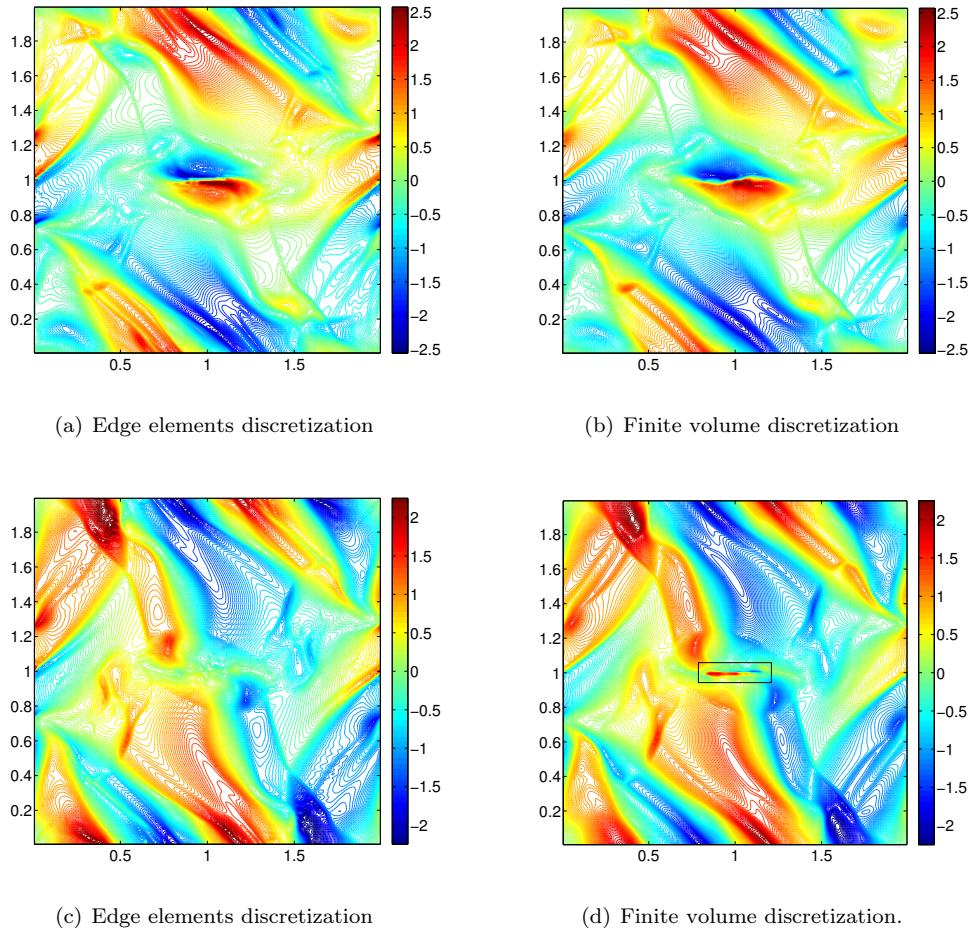


Figure 3.14.: Orszag–Tang benchmark: Comparison at time  $T = 1$  between the magnetic induction field obtained with the stabilized scheme (3.8)–(3.9), lowest order edge elements and second order Heun time-stepping (first column (a)–(c)) and the corresponding components of the reference solution from the MHD simulation with a finite volume scheme (second column (b)–(d)). The current sheet in Figure (d) is framed in the black box.

<sup>3</sup><http://www.sam.math.ethz.ch/alsvid-uq> (Accessed March 2016)

The rotated magnetic induction field obtained using the scheme (3.34) has been compared with the magnetic induction derived from the full MHD system. The numerical method we have proposed well resolves shocks, is stable and no spurious oscillations occur. As can be inferred from Figure 3.14, the current sheet characterizing the second component of the magnetic induction (black box in Figure 3.14 (d)) is not captured by the  $H(\mathbf{curl}, \Omega)$ -conforming scheme owing to the low order polynomial space discretization which is highly diffusive.

# 4. Extrusion Contraction Upwind Schemes

This chapter deals with a class of numerical schemes, devised in [HH08] to provide stable conforming discretizations of the stationary advection-diffusion problems, and based on finite dimensional spaces of discrete differential forms. We recall the strong form of the generalized advection-diffusion problem (2.17) for differential  $k$ -forms in  $\Omega \times I$ , with  $\Omega$  bounded Lipschitz domain in  $\mathbb{R}^n$  and  $I := [0, T]$ , namely

$$\begin{aligned} \star \partial_t \omega(t) + (-1)^{k+1} d^{n-k-1} \varepsilon \star d^k \omega(t) + \star \alpha \omega(t) + \star L_{\mathbf{u}} \omega(t) &= f(t), & \text{in } \Omega \times I, \\ \text{tr } \omega(t) &= \text{tr } g(t), & \text{on } (\Gamma_{\text{in}} \cup \Gamma_0) \times I, \\ \text{tr}(i_{\mathbf{n}} \omega(t)) &= \text{tr } s(t), & \text{on } \Gamma_{\text{in}} \times I, \\ \omega(0) &= \omega_0, & \text{in } \Omega, \end{aligned}$$

with the notations from Chapter 3. We focus on the transient generalized pure advection problem, with the rationale that numerical schemes proving robust in this case are also suitable for the generalized problem with non-vanishing diffusion.

Owing to Cartan's formula,  $L_{\mathbf{u}} = d^{k-1} i_{\mathbf{u}} + i_{\mathbf{u}} d^k$ , the exterior derivative and the Lie derivative commute, namely

$$d^k L_{\mathbf{u}} \omega = L_{\mathbf{u}} d^k \omega, \quad \forall \omega \in \Lambda^k(\Omega). \quad (4.1)$$

The commuting property has the fundamental consequence that closed differential forms are Lie advected into closed forms. In the MHD perspective, this translates into the exact preservation of the divergence constraint at every time. Moreover, it entails that, if  $\omega \in \Lambda^k(\Omega)$  is solution of the advection problem for  $k$ -forms, then  $d^k \omega$  is solution of the advection problem for  $(k+1)$ -forms, under suitable forcing terms, initial and boundary conditions. As an example, in three-dimensional ideal MHD flows, the magnetic potential  $\mathbf{A}$  and the magnetic induction field,  $\mathbf{B} = \text{curl } \mathbf{A}$ , satisfy the advection problem for 1-forms and 2-forms, respectively.

Most of the numerical methods for advection problems, as the stabilized Galerkin scheme in Chapter 3, do not seem to provide finite element solutions satisfying a strong or weak version of (4.1). Defining a discrete Lie derivative through Cartan's formula boils down to finding suitable discretizations of the exterior derivative and of the contraction operator. For  $H\Lambda^k(\Omega)$ -conforming finite element approximation spaces, a well-defined discrete exterior derivative is obtained as the restriction of the exterior derivative to the finite dimensional spaces of discrete differential forms. This chapter focuses on the design of structure-preserving and stable discretizations of the contraction operator based, as in [HH08], on the duality between contraction and extrusion advocated in [Bos03]. Furthermore, in Section 4.4, we adapt the entropy viscosity method, proposed in [GP08], to achieve high order numerical schemes giving non-oscillatory solutions and non-oscillatory derivatives of the solution near discontinuities. This limiting strategy is numerically investigated for both the extrusion contraction and the genuinely discontinuous stabilized Galerkin schemes from Chapter 3.

## 4.1. Contraction and Extrusion

Let  $\mathcal{M}_n$  be a  $n$ -dimensional smooth manifold and let  $\mathbf{u}$  be a smooth vector field on  $\mathcal{M}_n$ . In the definition of the Lie derivative in Section 2.1, we introduced the flow induced by  $\mathbf{u}$ , namely the mapping  $\Phi : \mathbb{R} \times \mathcal{M}_n \rightarrow \mathcal{M}_n$  such that  $\partial_t \Phi(t, x) = \mathbf{u}(\Phi(t, x))$  with  $\Phi(0, x) = x$ . The orbits of smooth manifolds under the flow define the so-called extrusion (see Figure 4.1).

**Definition 4.1.1** (Extrusion). Let  $\mathcal{M}_n$  be an  $n$ -dimensional smooth oriented manifold. Let  $\mathcal{S}_j$  be a  $j$ -dimensional submanifold of  $\mathcal{M}_n$ , the extrusion  $\text{Ext}_{\mathbf{u}}(\mathcal{S}_j, t)$  of  $\mathcal{S}_j$  by the smooth vector field  $\mathbf{u}$ , at time  $t$ , is the  $(j+1)$ -dimensional manifold formed by the union of the submanifolds obtained by sweeping  $\mathcal{S}_j = \Phi(0, \mathcal{S}_j)$  along the flow of  $\mathbf{u}$  to the submanifold  $\Phi(t, \mathcal{S}_j)$ . Specifically,  $\text{Ext}_{\mathbf{u}}(\mathcal{S}_j, t) = \bigcup_{s \in [0, t]} \Phi(s, \mathcal{S}_j)$  with orientation given by  $\partial \text{Ext}_{\mathbf{u}}(\mathcal{S}_j, t) = \Phi(t, \mathcal{S}_j) - \Phi(0, \mathcal{S}_j) - \text{Ext}_{\mathbf{u}}(\partial \mathcal{S}_j, t)$ .

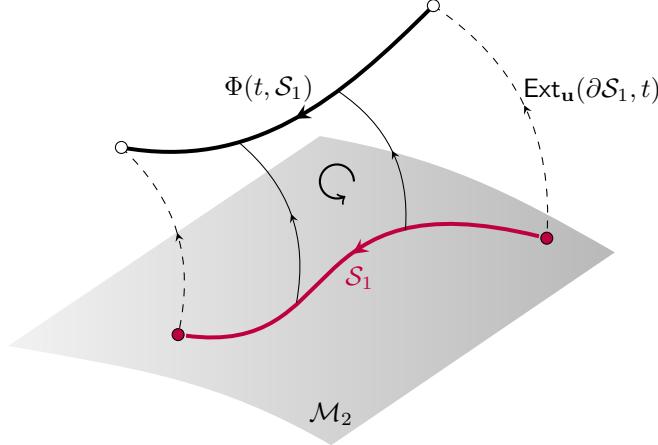


Figure 4.1.: Sketch of the extrusion of an oriented path  $\mathcal{S}_1$ .

In Definition 2.1.5, the contraction of alternating  $(k+1)$ -forms by a smooth vector field  $\mathbf{u}$  was defined as the  $k$ -form such that  $(i_{\mathbf{u}}\omega)(v_1, \dots, v_k) = \omega(\mathbf{u}, v_1, \dots, v_k)$ , for  $\omega \in \text{Alt}^{k+1} V$  and  $(v_1, \dots, v_k) \in V^k$ ,  $V$  being a real vector space. Pointwise application of the foregoing construction yields a definition of the contraction operator on smooth differential forms. The correspondences between the contraction of differential forms and proxy fields is recalled in Table 4.1. Alternatively, with the concept of extrusion, the contraction  $i_{\mathbf{u}}\omega \in \Lambda^k(\mathcal{M}_n)$  of a smooth  $(k+1)$ -form  $\omega \in \Lambda^{k+1}(\mathcal{M}_n)$  on a  $k$ -dimensional smooth oriented submanifold  $\mathcal{M}_k$  can be defined as the instantaneous change of  $\omega$  evaluated on the extrusion of  $\mathcal{M}_k$  [Bos03, Equation (14)], namely

$$\langle i_{\mathbf{u}}\omega, \mathcal{M}_k \rangle = \lim_{t \searrow 0} \frac{1}{t} \langle \omega, \text{Ext}_{\mathbf{u}}(\mathcal{M}_k, t) \rangle, \quad (4.2)$$

where  $\langle \cdot, \cdot \rangle$  denotes the chain-cochain duality pairing (2.18).

| $\omega \in \Lambda^k(\Omega)$ | $k = 0$ | $k = 1$                       | $k = 2$                        | $k = 3$       |
|--------------------------------|---------|-------------------------------|--------------------------------|---------------|
| $i_{\mathbf{u}}\omega$         | –       | $\mathbf{u} \cdot \mathbf{w}$ | $\mathbf{w} \times \mathbf{u}$ | $w\mathbf{u}$ |

Table 4.1.: Contraction of differential forms: exterior calculus notations and corresponding expressions for vector proxies  $\mathbf{w}$ ,  $w$ . See also Table 2.2.

**Remark 4.1.2** (Cartan's homotopy formula). The equivalence between the definition of Lie derivative (2.8) and Cartan's formula (2.7) can be shown using the duality between extrusion and contraction and Stokes' theorem. Indeed, integrating the Lie derivative of a smooth  $k$ -form  $\omega$  over a  $k$ -dimensional manifold  $\mathcal{M}_k$ , results in

$$\begin{aligned} \int_{\mathcal{M}_k} L_{\mathbf{u}}\omega &= \lim_{t \searrow 0} \frac{1}{t} \left( \int_{\mathcal{M}_k} \Phi_t^* \omega - \int_{\mathcal{M}_k} \omega \right) = \lim_{t \searrow 0} \frac{1}{t} \left( \int_{\Phi(t, \mathcal{M}_k)} \omega - \int_{\mathcal{M}_k} \omega \right) \\ &= \lim_{t \searrow 0} \frac{1}{t} \left( \int_{\partial \text{Ext}_{\mathbf{u}}(\mathcal{M}_k, t) + \Phi(0, \mathcal{M}_k) + \text{Ext}_{\mathbf{u}}(\partial \mathcal{M}_k, t)} \omega - \int_{\mathcal{M}_k} \omega \right) \\ &= \lim_{t \searrow 0} \frac{1}{t} \left( \int_{\partial \text{Ext}_{\mathbf{u}}(\mathcal{M}_k, t)} \omega + \int_{\text{Ext}_{\mathbf{u}}(\partial \mathcal{M}_k, t)} \omega \right) \end{aligned}$$

$$= \lim_{t \searrow 0} \frac{1}{t} \left( \int_{\text{Ext}_{\mathbf{u}}(\mathcal{M}_k, t)} \mathbf{d}^k \omega + \int_{\partial \mathcal{M}_k} \mathbf{i}_{\mathbf{u}} \omega \right) = \int_{\mathcal{M}_k} (\mathbf{i}_{\mathbf{u}} \mathbf{d}^k \omega + \mathbf{d}^{k-1} \mathbf{i}_{\mathbf{u}} \omega).$$

An algebraic proof of the equivalence of the two definitions can be found e.g. in [Lan99, Proposition 5.3 **LIE 1.**].

### Discrete Contraction via Extrusion

In a finite element discretization approach, the aim is to find suitable approximations of integrals of forms. The concept of extrusion of a manifold by the flow of a vector field offers a natural way to define the integral of a contracted discrete differential form over such manifold (4.2). A discrete contraction and its combination with the coboundary operator pave the way to discretizations of the Lie derivative.

Let  $\mathcal{T}_h$  be a cellular complex satisfying Assumption 2.4.2, and let  $\Lambda_h^k(\mathcal{T}_h)$ ,  $0 \leq k \leq n$ , be an  $H\Lambda^k(\Omega)$ -conforming space of polynomial discrete differential forms, as in (2.28) and (2.29). The duality formula (4.2) expresses the fact that the contraction of a smooth  $(k+1)$ -form over a  $k$ -dimensional manifold is equal to the instantaneous change of the form over the extrusion of the manifold. Equivalently, one might introduce an “upwinding” concept, by rewriting (4.2) as

$$\langle \mathbf{i}_{\mathbf{u}} \omega, \mathcal{M}_k \rangle = - \lim_{t \searrow 0} \frac{1}{t} \langle \omega, \text{Ext}_{\mathbf{u}}(\mathcal{M}_k, -t) \rangle, \quad \forall \omega \in \Lambda^{k+1}(\mathcal{M}_n), \quad (4.3)$$

where the  $k$ -dimensional submanifold  $\mathcal{M}_k$  is extruded *backward in time*.

As pointed out in Chapter 2, the lack of smoothness of discrete differential forms implies that

$$\langle \mathbf{i}_{\mathbf{u}} \omega_h, c_k \rangle = \lim_{t \searrow 0} \frac{1}{t} \int_{\text{Ext}_{\mathbf{u}}(c_k, t)} \omega_h \neq - \lim_{t \searrow 0} \frac{1}{t} \int_{\text{Ext}_{-\mathbf{u}}(c_k, t)} \omega_h = \langle -\mathbf{i}_{-\mathbf{u}} \omega_h, c_k \rangle, \quad (4.4)$$

with  $\omega_h \in \Lambda_h^{k+1}(\mathcal{T}_h)$  and  $c_k \in \Delta_k(\mathcal{T}_h)$ . The discrete contraction of a discrete  $(k+1)$ -form  $\omega_h \in \Lambda_h^{k+1}(\mathcal{T}_h)$  can be defined by somehow enforcing the equality in (4.4), namely extruding backward in time every  $k$ -cell of  $\mathcal{T}_h$ , and then evaluating the discrete form over the extruded  $(k+1)$ -manifold. Moreover, instead of approximating the “values” of the contraction  $\mathbf{i}_{\mathbf{u}} \omega_h$  at the  $k$ -cells, one can look for an approximation of their extrusion, rephrasing the problem into how to represent the extrusion of a  $k$ -dimensional manifold as a combination of  $(k+1)$ -chains. As an example, the case of 1-manifolds (extrusion of 0-chains) on a two-dimensional mesh  $\mathcal{T}_h$  can be analyzed as in [Bos03, Section III], namely an oriented line  $\overline{\mathbf{x} - \mathbf{y}}$  in a fixed  $n$ -cell  $T \in \mathcal{T}_h$  can be described by a weighted sum of edges (1-cells) as,

$$\overline{\mathbf{x} - \mathbf{y}} = \sum_{e \in \Delta_1(\mathcal{T}_h) \cap \partial T} \langle \psi_e, \overline{\mathbf{x} - \mathbf{y}} \rangle \bar{e},$$

where  $\psi_e$  is the Whitney form associated with the oriented edge  $e = \{\mathbf{x}_1, \mathbf{x}_2\}$ , with vector proxy  $\psi_e = \lambda_{\mathbf{x}_1} \text{grad} \lambda_{\mathbf{x}_2} - \lambda_{\mathbf{x}_2} \text{grad} \lambda_{\mathbf{x}_1}$ , and  $\lambda_{\mathbf{x}_i}$  being the barycentric coordinate at the node  $\mathbf{x}_i$ ,  $i \in \{1, 2\}$ . Let  $\mathbf{x} \in \Delta_0(\mathcal{T}_h)$  be a 0-cell of  $\mathcal{T}_h$  and let us assume that the path  $\overline{\mathbf{x} - \mathbf{y}} := -\lim_{t \searrow 0} \text{Ext}_{-\mathbf{u}}(\mathbf{x}, t)/t$  is contained in an  $n$ -cell  $T_{\mathbf{x}}^{\text{upw}} \in \mathcal{T}_h$ . Given a discrete 1-form  $\omega_h \in \Lambda_h^1(\mathcal{T}_h)$  we can define the degree of freedom of its contraction at  $\mathbf{x} \in \Delta_0(T)$  by enforcing  $\langle \mathbf{i}_{\mathbf{u}} \omega_h, \mathbf{x} \rangle = -\lim_{t \searrow 0} \langle \omega_h, \text{Ext}_{-\mathbf{u}}(\mathbf{x}, t) \rangle / t$ . In this way,

$$\begin{aligned} \langle \mathbf{i}_{\mathbf{u}} \omega_h, \mathbf{x} \rangle &= \langle \omega_h, \overline{\mathbf{x} - \mathbf{y}} \rangle = \sum_{e \in \Delta_1(\mathcal{T}_h) \cap \partial T_{\mathbf{x}}^{\text{upw}}} \langle \psi_e, \overline{\mathbf{x} - \mathbf{y}} \rangle \langle \omega_h, \bar{e} \rangle \\ &= \langle \sum_{e \in \Delta_1(\mathcal{T}_h) \cap \partial T_{\mathbf{x}}^{\text{upw}}} W_{e, T_{\mathbf{x}}^{\text{upw}}}^1(\omega_h) \psi_e, \overline{\mathbf{x} - \mathbf{y}} \rangle = \langle \omega_h|_{T_{\mathbf{x}}^{\text{upw}}}, \overline{\mathbf{x} - \mathbf{y}} \rangle = \langle \mathbf{i}_{\mathbf{u}} \omega_h|_{T_{\mathbf{x}}^{\text{upw}}}, \mathbf{x} \rangle, \end{aligned}$$

where  $W_{e, T_{\mathbf{x}}^{\text{upw}}}^1(\omega_h) \in \mathbb{R}$  is the degree of freedom of  $\omega_h \in \Lambda_h^1(\mathcal{T}_h)$  associated with  $e$  as in (2.33).

## 4.2. Upwinding the Contraction via Extrusion

The terminology *extrusion contraction upwind schemes* was introduced in [HH08] where, taking the cue from the monotone upwind quadrature scheme of Tabata [Tab77], the authors proposed an upwind discretization of the Lie derivative in the advection-diffusion problem exploiting the duality between contraction and extrusion suggested in [Bos03]. While the focus of the work [HH08] is on the stationary scalar advection-diffusion problem, emphasis on the transient pure advection problem for 1-forms is placed in the discretization scheme suggested in [Mul+11]. The approach of the latter work resembles more a finite volume technique: the contraction of a discrete  $(k+1)$ -form is approximated by projecting the extrusion of every  $k$ -cell of  $\mathcal{T}_h$  onto the mesh aligned  $(k+1)$ -dimensional subspaces, and then summing the contributions of  $(k+1)$ -dimensional finite volume discretizations applied to each of the projections, (*cf.* [Mul+11, Section 4.2] for further details). In a FEEC perspective, we adopt the strategy of [HH08].

Recall the generalized pure advection initial boundary value problem (3.1). Following a method of lines strategy, the extrusion contraction upwind discretization of the time-independent advection operator can be coupled with an explicit Runge–Kutta scheme. Therefore, we first consider the stationary generalized advection boundary value problem for a  $k$ -form  $\omega \in \Lambda^k(\Omega)$  on the bounded computational domain  $\Omega \subset \mathbb{R}^n$ , as in (3.3) and [Heu11, Equation (4.1)]. More precisely, the problem reads: Find  $\omega \in V$  such that

$$\begin{aligned}\alpha\omega + L_{\mathbf{u}}\omega &= f, \quad \text{in } \Omega, \\ \operatorname{tr} \omega &= g, \quad \text{on } \Gamma_{\text{in}}, \\ \operatorname{tr} i_{\mathbf{n}}\omega &= s, \quad \text{on } \Gamma_{\text{in}},\end{aligned}$$

with  $f \in L^2\Lambda^k(\Omega)$ ,  $g \in L^2\Lambda^k(\Gamma_{\text{in}})$ ,  $s \in L^2\Lambda^{k-1}(\Gamma_{\text{in}})$  and where

$$V := \{\omega \in L^2\Lambda^k(\Omega) : L_{\mathbf{u}}\omega \in L^2\Lambda^k(\Omega), \int_{\Gamma_{\text{in}}} \operatorname{tr} i_{-\mathbf{u}}(\omega \wedge \star\omega) < \infty\}.$$

Let us assume, for the time being, that  $\mathbf{u} \in W^{1,\infty}(\Omega)$  and satisfies the “coercivity” condition (3.2). The well-posedness of the generalized advection problem has been discussed in Section 3.1.1. We aim at finding a stable numerical discretization of the Lie operator  $L_{\mathbf{u}}$  based on polynomial  $H\Lambda^k(\Omega)$ -conforming discrete differential forms. With the Cartan’s formula in mind, this boils down to finding discrete counterparts of the exterior derivative and of the contraction operator.

Let  $\Lambda_{h,r}^k(\mathcal{T}_h)$ ,  $0 \leq k \leq n$ , be the  $H\Lambda^k(\Omega)$ -conforming space of piecewise polynomial discrete differential  $k$ -forms of degree at most  $r \geq 1$ , as introduced in Section 2.4.1, namely

$$\Lambda_{h,r}^k(\mathcal{T}_h) := \{\omega_h \in H\Lambda^k(\Omega) : \omega_{h|_T} \in \zeta_r(T), T \in \mathcal{T}_h\}, \quad (4.5)$$

where the local shape functions are  $\zeta_r(T) = \mathcal{P}_r\Lambda^k(T)$  or  $\zeta_r(T) = \mathcal{P}_r^-\Lambda^k(T)$  on simplicial meshes, and  $\zeta_r(T) = \mathcal{Q}_r^-\Lambda^k(T)$  or  $\zeta_r(T) = \mathcal{S}_r\Lambda^k(T)$  on Cartesian meshes. Recall that, in view of Stokes’s theorem, a discrete definition of exterior derivative is directly given through the coboundary operator (*cf.* (2.20)), i.e.,  $\langle d^k\omega_h, c_{k+1} \rangle = \langle \omega_h, \partial c_{k+1} \rangle$  where  $\partial c_{k+1}$  is a  $k$ -chain, boundary of  $c_{k+1}$ . Hence, for every  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$ ,  $d^k\omega_h$  can be easily computed through the mesh incidence matrix and  $d^k\omega_h \in \Lambda_{h,r+}^{k+1}(\mathcal{T}_h)$  for  $r^+ = r^+(r) \in \mathbb{N}$ ,  $r^+ \geq 1$  and  $r^+ \in \{r-1, r\}$  according to the choice of  $\zeta_r(T)$ , as in (2.30). The first step in devising a discrete definition of the contraction operator consists in a global reconstruction, obtained by interpolating the contraction  $i_{\mathbf{u}}\omega_h$  of  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$  into some space of  $H\Lambda^{k-1}(\Omega)$ -conforming discrete differential forms. Since we are interested in piecewise polynomial discretizations, we consider interpolation spaces of  $H\Lambda^k(\Omega)$ -conforming polynomial discrete differential forms  $\Lambda_{h,p}^k(\mathcal{T}_h)$  of type (4.5) for some polynomial degree  $p \geq 1$  which might differ from the polynomial degree  $r$  of the approximation spaces. Finding a discretization of the contraction operator is therefore rephrased as associating suitable degrees of freedom to the discrete  $(k-1)$ -form  $i_{\mathbf{u}}\omega_h$ , [HH08, Section 1.2]. This is achieved by exploiting the duality between the contraction of a differential form by a vector field and the extrusion of a manifold, as presented in Section 4.1, and imposing in the discrete setting the equality (4.4) valid at the continuous level.

**Definition 4.2.1** (Upwind interpolation of contraction). Let  $\mathcal{T}_h$  be a cellular complex as in Assumption 2.4.2. Let  $\omega_h \in \Lambda_h^{k+1}(\mathcal{T}_h)$ ,  $0 \leq k \leq n-1$ , be a discrete differential  $(k+1)$ -form and let  $\mathbf{u}$  be a smooth vector field. Given a polynomial space of differential forms  $\Lambda_{h,p}^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$ ,  $p \geq 1$ , as in (4.5), the upwind interpolation operator  $\mathcal{I}_{\mathbf{u},p}^k : L^2\Lambda^k(\Omega) \rightarrow \Lambda_{h,p}^k(\mathcal{T}_h)$  is defined as,

$$\mathcal{I}_{\mathbf{u},p}^k(i_{\mathbf{u}}\omega_h) = \sum_{j=k}^{M_{\min}^{p,k}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell (i_{\mathbf{u}}\omega_h|_{T_{f_j}^{\text{upw}}}) \psi_j^\ell,$$

where  $T_{f_j}^{\text{upw}} \in \mathcal{T}_h$  is the  $n$ -cell laying in the upwind direction of  $f_j$ , determined by the vector field  $\mathbf{u}$ . Here,  $\{W_{f_j}^\ell\}_{j,\ell} \subset \mathbb{R}$  are the degrees of freedom associated with the  $j$ -dimensional mesh cells,  $\{\psi_j^\ell\}_{j,\ell}$  is a basis of  $\Lambda_{h,p}^k(\mathcal{T}_h)$ , and

$$M_{\min}^{p,k} := \begin{cases} \min\{n, \lfloor p/2 \rfloor + k\} & \text{if } \Lambda_{h,p}^k(\mathcal{T}_h) = \mathcal{S}_p \Lambda^k(\mathcal{T}_h), \\ \min\{n, p+k-1\} & \text{otherwise.} \end{cases} \quad (4.6)$$

Note that,  $\mathcal{I}_{\mathbf{u},p}^k(i_{\mathbf{u}}\omega_h) = i_{\mathbf{u}}\omega_h$  for every  $i_{\mathbf{u}}\omega_h \in \Lambda_{h,p}^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$ . Moreover, the degrees of freedom interior to the mesh elements are not affected by upwinding. This is analogous to the Galerkin methods derived in Chapter 3, where flux upwinding is confined to the mesh interface terms.

On account of the discretization of the contraction operator ensuing from Definition 4.2.1, the *discrete* Lie derivative is defined as,

$$\begin{aligned} L_{\mathbf{u}}^h : \Lambda_{h,r}^k(\mathcal{T}_h) &\longrightarrow \Lambda_{h,p}^k(\mathcal{T}_h) \\ \omega_h &\longmapsto \mathcal{I}_{\mathbf{u},p}^k(i_{\mathbf{u}}d^k\omega_h) + d^{k-1}\mathcal{I}_{\mathbf{u},p}^{k-1}(i_{\mathbf{u}}\omega_h), \end{aligned} \quad (4.7)$$

where  $p^- \geq 1$  is such that  $d^{k-1}\Lambda_{h,p}^{k-1}(\mathcal{T}_h) \subset \Lambda_{h,p}^k(\mathcal{T}_h)$ , see Section 2.4. A few observations can be made. First, the polynomial interpolation order  $p$  has to be chosen such that the consistency error does not debase the accuracy order related to the finite element approximation. Moreover, in view of (4.3), the characterization (4.7) automatically incorporates an upwinding of the Lie derivative. Lastly, since  $\mathbf{u}$  is Lipschitz continuous and the discrete differential forms are piecewise polynomials, the moments of the contracted forms on each  $j$ -cell of  $\mathcal{T}_h$  are well-defined from within the  $n$ -cell in the upwind direction of the flow. As a result of the discretization of the Lie derivative, the discrete advection problem, recast in weak form, reads: Find  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$  such that

$$a_h(\omega_h, \eta_h) = (f, \eta_h)_\Omega, \quad \forall \eta_h \in \Lambda_{h,r}^k(\mathcal{T}_h), \quad (4.8)$$

where  $(\cdot, \cdot)_\Omega$  denotes the  $L^2\Lambda^k(\Omega)$  inner product  $(\omega, \eta)_\Omega := \int_\Omega \omega \wedge \star \eta$  in (2.11). The bilinear form  $a_h(\cdot, \cdot)$  is defined, for all  $\omega_h, \eta_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$ , as

$$a_h(\omega_h, \eta_h) := (\alpha\omega_h, \eta_h)_\Omega + \int_\Omega (\mathcal{I}_{\mathbf{u},p}^k(i_{\mathbf{u}}d^k\omega_h) \wedge \star \eta_h + d^{k-1}\mathcal{I}_{\mathbf{u},p}^{k-1}(i_{\mathbf{u}}\omega_h) \wedge \star \eta_h).$$

**Remark 4.2.2** (Boundary conditions at the inflow boundary). Incorporating the boundary conditions in the discrete definition of the Lie derivative for  $k$ -forms, requires the computation of the moment-based degrees of freedom at the boundary  $j$ -cells,  $k \leq j \leq M_{\min}^{p,k}$ , using the boundary data. Note that the trace of the form and the trace of its contraction,  $\text{tr}(i_{\mathbf{n}}\omega_h)$ , at the inflow boundary of the domain  $\Omega$  are supplied by the boundary conditions. One can decompose the velocity field  $\mathbf{u}$  in its normal component  $\mathbf{u}_n := (\mathbf{u} \cdot \mathbf{n})\mathbf{n}$  and its tangential component  $\mathbf{u}_t := (\mathbf{n} \times \mathbf{u}) \times \mathbf{n}$ , such that

$$i_{\mathbf{u}}\omega_h = i_{\mathbf{u}_n}\omega_h + i_{\mathbf{u}_t}\omega_h = (\mathbf{u} \cdot \mathbf{n})i_{\mathbf{n}}\omega_h + i_{\mathbf{u}_t}\omega_h, \quad \forall \omega_h \in \Lambda_h^k(\mathcal{T}_h), \forall k.$$

However, in the case of non-smooth boundaries, the evaluation of the contraction might not be well-defined at a given boundary  $j$ -cell,  $j \leq n-1$ . One could average the contributions from the  $n$ -cells sharing the  $j$ -cell. For example, in two dimensions, at a boundary node  $\mathbf{x} \in \Delta_0(\mathcal{T}_h) \cap \partial\Omega$  such that  $x \in \Delta_0(e_1) \cap \Delta_0(e_2)$  for some edges  $e_1, e_2 \in \Delta_1(\mathcal{T}_h) \cap \partial\Omega$ , one can approximate the contraction of a 1-form  $\omega_h$  as

$$i_{\mathbf{u}}\omega_h \approx \frac{1}{2}((\mathbf{u} \cdot \mathbf{n}_1)i_{\mathbf{n}_1}\omega_h + i_{\mathbf{u}_t}\omega_h) + \frac{1}{2}((\mathbf{u} \cdot \mathbf{n}_2)i_{\mathbf{n}_2}\omega_h + i_{\mathbf{u}_t}\omega_h).$$

### Discrete Lie Derivative in Terms of Vector Proxies

Let  $V_h$  be finite element spaces of vector proxies associated with the spaces  $\Lambda_{h,r}^k(\mathcal{T}_h)$  of polynomial differential  $k$ -forms of degree at most  $r \geq 1$  on the three-dimensional cellular complex  $\mathcal{T}_h$ . Let  $\mathcal{I}_{\mathbf{u},p}^k$ ,  $0 \leq k \leq 2$  and  $\mathcal{I}_{\mathbf{u},p^-}^k$ ,  $1 \leq k \leq 3$ , be upwind interpolation operators onto piecewise polynomial spaces of degree at most  $p \geq 1$  and  $p^- \geq 1$ , respectively. Let  $w_h \in V_h$  or  $\mathbf{w}_h \in V_h$  be the vector proxy representation of the  $k$ -form  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$  and let  $M_{\min}^{p,k}$  be defined as in (4.6) for  $n = 3$ . Tables 2.1, 2.2 and 4.1 suggest the correspondences between differential forms and their vector proxies. The discrete Lie derivative (4.7) reads

$$\begin{aligned} k = 0 : \quad \mathbf{L}_{\mathbf{u}}^h w_h &= \mathcal{I}_{\mathbf{u},p}^0(\mathbf{i}_{\mathbf{u}} \mathbf{d}^0 w_h) = \mathcal{I}_{\mathbf{u},p}^0(\mathbf{u} \cdot \operatorname{grad} w_h) \\ &= \sum_{j=0}^{M_{\min}^{p,0}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\mathbf{u} \cdot \operatorname{grad} w_h|_{T_{f_j}^{\text{upw}}}) \lambda_j^\ell, \end{aligned} \quad (4.9)$$

where  $\{\lambda_j^\ell\}_{j,\ell}$  is a basis of  $H^1$ -conforming polynomials of degree at most  $p$ . As anticipated,  $T_{f_j}^{\text{upw}}$  denotes the  $n$ -cell of  $\mathcal{T}_h$  laying in the upwind direction of  $f_j \in \Delta_j(\mathcal{T}_h)$  in the sense that the extrusion  $\operatorname{Ext}_{\mathbf{u}}(f_j, t)$  is contained in  $T_{f_j}^{\text{upw}}$ , for  $t$  small enough, and with the assumption that  $T_{f_n}^{\text{upw}} = T$  and that  $T_{f_0}^{\text{upw}} = T$  for all  $f_0 \in \Delta_0(T) \setminus \partial T$ .

Let  $\{\phi_j^\ell\}_{j,\ell}$  be a basis of  $H(\operatorname{curl}, \Omega)$ -conforming polynomials of degree at most  $p$  and  $\{\lambda_j^\ell\}_{j,\ell}$  a basis of  $H^1$ -conforming polynomials of degree at most  $p^-$ . There holds,

$$\begin{aligned} k = 1 : \quad \mathbf{L}_{\mathbf{u}}^h \mathbf{w}_h &= \mathcal{I}_{\mathbf{u},p}^1(\mathbf{i}_{\mathbf{u}} \mathbf{d}^1 \mathbf{w}_h) + \mathbf{d}^0 \mathcal{I}_{\mathbf{u},p^-}^0(\mathbf{i}_{\mathbf{u}} \mathbf{w}_h) = \mathcal{I}_{\mathbf{u},p}^1(\operatorname{curl} \mathbf{w}_h \times \mathbf{u}) + \operatorname{grad}(\mathcal{I}_{\mathbf{u},p^-}^0(\mathbf{u} \cdot \mathbf{w}_h)) \\ &= \sum_{j=1}^{M_{\min}^{p,1}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\operatorname{curl} \mathbf{w}_h \times \mathbf{u}|_{T_{f_j}^{\text{upw}}}) \phi_j^\ell \\ &\quad + \sum_{j=0}^{M_{\min}^{p^-,0}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\mathbf{u} \cdot \mathbf{w}_h|_{T_{f_j}^{\text{upw}}}) \operatorname{grad} \lambda_j^\ell. \end{aligned}$$

Let  $\{\varphi_j^\ell\}_{j,\ell}$  be a basis of  $H(\operatorname{div}, \Omega)$ -conforming polynomials of degree at most  $p$  and  $\{\phi_j^\ell\}_{j,\ell}$  a basis of  $H(\operatorname{curl}, \Omega)$ -conforming polynomials of degree at most  $p^-$ , then

$$\begin{aligned} k = 2 : \quad \mathbf{L}_{\mathbf{u}}^h \mathbf{w}_h &= \mathcal{I}_{\mathbf{u},p}^2(\mathbf{i}_{\mathbf{u}} \mathbf{d}^2 \mathbf{w}_h) + \mathbf{d}^1 \mathcal{I}_{\mathbf{u},p^-}^1(\mathbf{i}_{\mathbf{u}} \mathbf{w}_h) = \mathcal{I}_{\mathbf{u},p}^2(\mathbf{u} \operatorname{div} \mathbf{w}_h) + \operatorname{curl}(\mathcal{I}_{\mathbf{u},p^-}^1(\mathbf{w}_h \times \mathbf{u})) \\ &= \sum_{j=2}^{M_{\min}^{p,2}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\mathbf{u} \operatorname{div} \mathbf{w}_h|_{T_{f_j}^{\text{upw}}}) \varphi_j^\ell \\ &\quad + \sum_{j=1}^{M_{\min}^{p^-,1}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\mathbf{w}_h \times \mathbf{u}|_{T_{f_j}^{\text{upw}}}) \operatorname{curl} \phi_j^\ell. \end{aligned}$$

For  $\{\varphi_j^\ell\}_{j,\ell}$  basis of  $H(\operatorname{div}, \Omega)$ -conforming polynomials of degree at most  $p^-$ ,

$$k = 3 : \quad \mathbf{L}_{\mathbf{u}}^h w_h = \mathbf{d}^2 \mathcal{I}_{\mathbf{u},p^-}^2(\mathbf{i}_{\mathbf{u}} w_h) = \operatorname{div}(\mathcal{I}_{\mathbf{u},p^-}^2(\mathbf{u} w_h)) = \sum_{j=2}^{M_{\min}^{p^-,2}} \sum_{f_j \in \Delta_j(\mathcal{T}_h)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell(\mathbf{u} w_h|_{T_{f_j}^{\text{upw}}}) \operatorname{div} \varphi_j^\ell.$$

**Remark 4.2.3** (Tabata's scheme). Let  $\mathcal{T}_h$  denote a simplicial triangulation of  $\Omega \subset \mathbb{R}^n$  of weakly acute type. We consider the extrusion contraction upwind discretization of the scalar advection problem with linear Lagrangian finite element spaces  $V_h$ . Using the discrete Lie derivative (4.9) in the weak formulation (4.8), yields the bilinear form

$$a_h(w_h, v_h) = \sum_{\ell=1}^{N_0} \mathbf{u}(x_\ell) \cdot (\operatorname{grad} w_h)|_{T_{x_\ell}^{\text{upw}}}(x_\ell) \int_{\Omega} \lambda^\ell v_h, \quad \forall w_h, v_h \in V_h.$$

where  $\{\lambda^\ell\}_\ell$  are the barycentric coordinates and  $N_0 := \dim V_h = \#\Delta_0(\mathcal{T}_h)$ . Approximating the integration on  $\Omega$  using local quadrature rules  $Q(T) = \{a_{i,T}, q_{i,T}\}_{i=0}^n$  with weights  $\{q_{i,T}\}_i$  and nodes  $\{a_{i,T}\}_i$  at the mesh 0-cells (vertices of the  $n$ -simplices), results in

$$\begin{aligned} a_h(w_h, v_h) &= \sum_{\ell=1}^{N_0} \mathbf{u}(x_\ell) \cdot (\text{grad} w_h)|_{T_{x_\ell}^{\text{upw}}}(x_\ell) \sum_{T \in \mathcal{T}_h} \sum_{a_{i,T} \in \Delta_0(T)} q_{i,T} \lambda^\ell(a_{i,T}) v_h(a_{i,T}) \\ &= \sum_{T \in \mathcal{T}_h} \sum_{a_{i,T} \in \Delta_0(T)} q_{i,T} \mathbf{u}(a_{i,T}) \cdot (\text{grad} w_h)|_{T_{a_{i,T}}^{\text{upw}}}(a_{i,T}) v_h(a_{i,T}), \end{aligned}$$

and the so-called *upwind quadrature* or Tabata's scheme [Tab77] is recovered. The method proposed by Tabata to solve the transient scalar advection-diffusion problem with homogeneous Dirichlet boundary conditions at the domain boundaries, is first order accurate, and it delivers an algebraic system M-matrix (i.e., a non-singular matrix whose entries  $a_{i,j}$  satisfy  $a_{i,j} \leq 0$  for  $i \neq j$  and the entries  $b_{i,j}$  of the inverse matrix are non-negative,  $b_{i,j} \geq 0$ ). This entails that the discrete solution operator is inverse monotone. Therefore, when augmented with a standard linear finite element discretization of the diffusion operator, the resulting scheme is able to preserve the inverse-monotonicity property and hence the maximum principle characterizing the problem at the continuous level [Tab77, Theorem 1].

#### 4.2.1. Commuting Property of the Discrete Lie Derivative

The discretization of the Lie derivative proposed in (4.7) yields an advection operator satisfying the commuting property (4.1) in the discrete setting.

**Proposition 4.2.4.** *Let  $\Omega \subset \mathbb{R}^n$  be a bounded Lipschitz domain and let  $\mathcal{T}_h$  be a cellular complex on  $\Omega$  as in Assumption 2.4.2. Let  $\mathbf{u} \in W^{1,\infty}(\Omega)$  and let  $L_u^h$  be the extrusion contraction upwind discretization of the Lie derivative  $L_u$  in (4.7). For all  $0 \leq k \leq n-1$ , the following diagram*

$$\begin{array}{ccc} \Lambda_{h,r}^k(\mathcal{T}_h) & \xrightarrow{d^k} & \Lambda_{h,r+}^{k+1}(\mathcal{T}_h) \\ \downarrow L_u^h & & \downarrow L_u^h \\ \Lambda_{h,p}^k(\mathcal{T}_h) & \xrightarrow{d^k} & \Lambda_{h,p+}^{k+1}(\mathcal{T}_h), \end{array}$$

commutes, namely,  $L_u^h d^k \omega_h = d^k L_u^h \omega_h$ , for all  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h) \subset H\Lambda^k(\Omega)$ ,  $r \geq 1$ .

*Proof.* The commutativity of the discrete exterior and Lie derivatives follows immediately from the topological properties of the (discrete) exterior derivative. Indeed, if  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$ , the definition of discrete Lie derivative in (4.7) results in

$$\begin{aligned} L_u d^k \omega_h &= \mathcal{I}_{\mathbf{u},p+}^{k+1}(\mathbf{i}_{\mathbf{u}} d^{k+1} d^k \omega_h) + d^k \mathcal{I}_{\mathbf{u},p}^k(\mathbf{i}_{\mathbf{u}} d^k \omega_h) = d^k (\mathcal{I}_{\mathbf{u},p}^k(\mathbf{i}_{\mathbf{u}} d^k \omega_h)) \\ &= d^k (d^{k-1} \mathcal{I}_{\mathbf{u},p-}^{k-1}(\mathbf{i}_{\mathbf{u}} \omega_h) + \mathcal{I}_{\mathbf{u},p}^k(\mathbf{i}_{\mathbf{u}} d^k \omega_h)) = d^k L_u \omega_h, \end{aligned}$$

owing to the fact that the discrete exterior derivative satisfies  $d^{k+1} \circ d^k = 0$ .  $\square$

**Remark 4.2.5.** The result of Proposition 4.2.4 has two major consequences:

- (i) Under the assumption of unique solvability of the discrete time-dependent problem corresponding to (4.8), discrete closed  $k$ -forms are Lie advected into closed  $k$ -forms. There are different ways to proceed, we show it from an algebraic perspective. Consider the semi-discrete problem  $\partial_t \omega_h + \alpha \omega_h + L_u^h \omega_h = f$ , and suppose the initial datum is a closed form. Let  $\mathbb{L}$  be the matrix associated with the discretization of the Lie derivative (4.7). Using linear explicit temporal discretizations (we refer to (6.10) for explicit formulas), the discrete solution  $\omega_h^n$  at time  $t^n$  is obtained as a linear combination of the solution at the previous time steps. Let  $W^n$  be the vector of degrees of freedom for  $\omega_h$  at time  $t^n$ . In the simplest case of explicit Euler time-stepping,  $\mathbb{M}W^{n+1} = \mathbb{M}W^n - \Delta t^n \alpha \mathbb{M}W^n - \Delta t^n \mathbb{L}W^n + \Delta t^n F^n$ , where  $\mathbb{M}$  is the mass matrix and  $F^n$  is the load vector

associated with the source term at time  $t^n$ . By Proposition 4.2.4, the incidence matrix  $\mathbb{D}^k$ , representing the exterior derivative operator, commutes with  $\mathbb{L}$ . Hence, under suitable boundary conditions,  $\mathbb{D}^k W^n = 0$  for all  $n \geq 1$ , provided the right hand side is closed at all times.

(ii) Recall that the long sequences

$$\begin{aligned}\mathbb{R} \hookrightarrow \mathcal{P}_r \Lambda^0(\mathcal{T}_h) &\xrightarrow{\mathbb{d}^0} \mathcal{P}_{r-1} \Lambda^1(\mathcal{T}_h) \xrightarrow{\mathbb{d}^1} \dots \xrightarrow{\mathbb{d}^{n-1}} \mathcal{P}_{r-n} \Lambda^n(\mathcal{T}_h) \longrightarrow 0, \\ \mathbb{R} \hookrightarrow \mathcal{P}_r^- \Lambda^0(\mathcal{T}_h) &\xrightarrow{\mathbb{d}^0} \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h) \xrightarrow{\mathbb{d}^1} \dots \xrightarrow{\mathbb{d}^{n-1}} \mathcal{P}_r^- \Lambda^n(\mathcal{T}_h) \longrightarrow 0,\end{aligned}$$

are exact for all  $r \geq 1$ , and similarly for polynomial spaces of discrete differential forms over Cartesian complexes (2.29). At least for polynomial discrete differential forms of the first family, if a priori convergence results (in a certain norm) independent of the form degree  $k$  can be established, then, upon suitably tuning the polynomial approximation and interpolation orders, no accuracy is lost in solving the advection problem for the magnetic potential, rather than the magnetic induction advection.

### 4.3. Numerical Experiments in 2D

Since stability and a priori convergence estimates are still open questions, we test the method on a set of numerical experiments with the aim of deriving heuristic convergence properties in the  $L^2$ -norm and in some energy norm, and to verify the robustness of the numerical scheme in tackling shocks and complex MHD flows. Finally, we look for numerical confirmation of the discrete commuting property stated in Proposition 4.2.4.

As will be discussed in Chapter 7, in this dissertation, we focus on numerical simulations for the planar two-dimensional MHD model. In this setting, the magnetic potential is a scalar function associated with the transverse out-plane component of the magnetic induction field, and  $\mathbf{B} = \mathbf{curl}_{2D} A = \text{grad}^\perp A := (\partial_y A, -\partial_x A)^\top$ . For this reason we restrict, in the present section, to numerical simulations for the advection of 0-forms and 1-forms in two dimensions.

The velocity field has been assumed to be a Lipschitz continuous vector-valued function hitherto. However, in MHD flow simulations we are dealing with  $\mathcal{T}_h$ -piecewise polynomial velocities. We consider the averaged values at every interpolation  $j$ -cells whenever the velocity is not well-defined. The average will also determine the upwind direction of the flow. A more sophisticated approach to avoid the possible shortcoming of averaging the velocity values will be pointed out in Chapter 7.

#### 4.3.1. Transient Advection of 0-Forms

On the two-dimensional domain  $\Omega \subset \mathbb{R}^2$  and on the time interval  $I \subset \mathbb{R}$ , we consider the pure advection problem for the scalar magnetic potential  $A$ , namely

$$\begin{aligned}\partial_t A(t) + \mathbf{u} \cdot \text{grad} A(t) &= f(t), \quad \text{in } \Omega \times I, \\ A(0) &= A_0, \quad \text{in } \Omega,\end{aligned}\tag{4.10}$$

with suitable boundary conditions.

##### Test of Convergence: Constant Velocity

On the domain  $\Omega = [0, 2]^2$  with periodic boundary conditions and the time interval  $I = [0, 0.5]$ , the initial condition  $A_0 = \frac{1}{\pi} \cos(\pi y) + \frac{1}{2\pi} \cos(2\pi x)$  is advected at constant velocity  $\mathbf{u} = (4, 4)^\top$ . The forcing term is set to zero  $f = 0$ . On a family of Cartesian meshes  $\{\mathcal{T}_h\}_h$ , we discretize the Lie derivative as in (4.9) using the approximation spaces  $\Lambda_{h,r}^k(\mathcal{T}_h)$  of bilinear ( $r = 1$ ) and biquadratic ( $r = 2$ ) Lagrangian finite elements. The polynomial degree of the upwind interpolation is chosen to coincide with the polynomial degree of the finite element trial and test spaces, that is  $p = r$  in (4.9). In order to gauge the spatial accuracy

of the extrusion contraction scheme, we use Heun time-stepping with uniform time step  $\Delta t = 0.1h$  for bilinear Lagrangian finite element approximations and  $\Delta t = 0.01h$  for biquadratic Lagrangian finite elements. Owing to the periodicity of the domain, we can compare the numerical solution at the final time  $t = 0.5$  with  $A_0$ .

The projection of the numerical and exact solution at final time onto the one-dimensional line  $\{\mathbf{x} \in \Omega, y = 1\}$  is shown in Figure 4.2, bottom row. The piecewise linear discretization yields a rather diffusive solution and first order accuracy in both the  $L^2$ - and  $H^1$ -norms, as reported in Figure 4.2 where the  $L^2$ -error of the potential  $A$  and of its two-dimensional **curl**, the magnetic induction field  $\mathbf{B}$ , is reported. Second order convergence is attained with the piecewise biquadratic discretization (and interpolation).

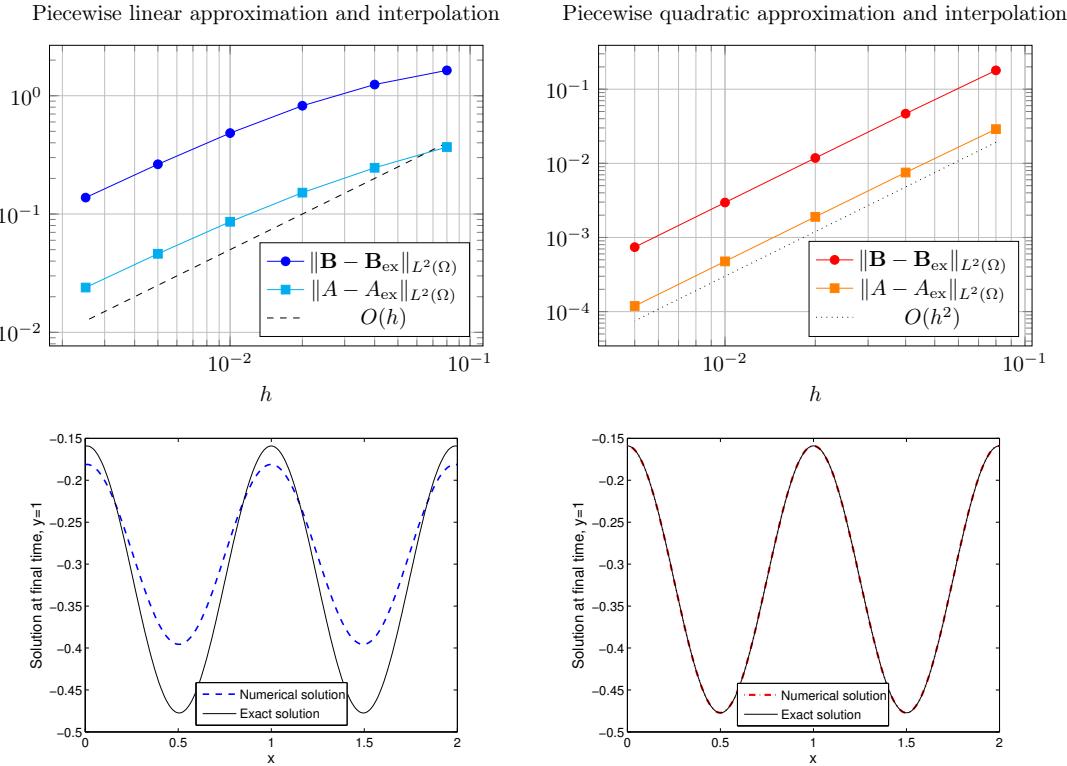


Figure 4.2.: Test for advection of 0-forms with constant velocity. Convergence plot of the error in the  $L^2$ -norm at the final time  $t = 0.5$  for the magnetic potential  $A$  and the induction field  $\mathbf{B}$  (top row). The exact solution is  $A_{\text{ex}} \equiv A_0$  and analogous for the  $\mathbf{B}$  field. Projection of the exact and numerical solution for  $y = 1$  on a  $200 \times 200$  Cartesian mesh (bottom row). Left: first order interpolation, bilinear Lagrangian finite elements. Right: second order interpolation, biquadratic Lagrangian finite elements.

### Accuracy Test for MHD: Smooth Vortex

As a second numerical test we consider the electromagnetic phenomena involved in the evolution of an MHD flow. The present test will be used for the numerical convergence analysis of the numerical schemes for the full ideal MHD system in Chapter 7. Restricting to the advection equation for the magnetic potential, the solution is known at every point in space and time on the domain  $\Omega \times I := [-5, 5]^2 \times [0, 0.5]$  and it is given analytically by

$$\begin{aligned}\mathbf{u}(x, y, t) &= \mathbf{u}_0 + \frac{\kappa}{2\pi} e^{1/2(1-r^2)} (t - y, x - t)^\top, \\ \mathbf{B}(x, y, t) &= \frac{\mu}{2\pi} e^{1/2(1-r^2)} (t - y, x - t)^\top, \\ A(x, y, t) &= \frac{\mu}{2\pi} e^{1/2(1-r^2)},\end{aligned}$$

with  $r(x, y, t) := \sqrt{(x - u_0^1 t)^2 + (y - u_0^2 t)^2}$ ,  $\mathbf{u}_0 = (u_0^1, u_0^2)^\top = (1, 1)^\top$  and  $\kappa = \mu = 1$ . On a Cartesian mesh, piecewise linear and piecewise quadratic extrusion contraction upwind discretizations are coupled to Heun time-stepping. First order convergence in the  $L^2$ -norm and  $H(\mathbf{curl}_{2D}, \Omega)$ -seminorm are recorded in the former case, Figure 4.3 (plot on the right). Second order convergence in the same norms is observed for the piecewise quadratic discretization in Figure 4.3 (left plot).

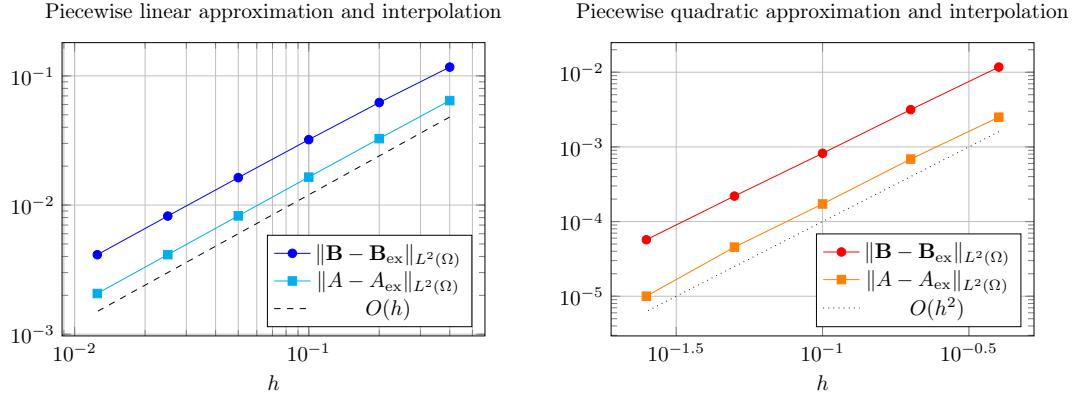


Figure 4.3.: Numerical convergence study of the piecewise linear (left) and piecewise quadratic (right) extrusion contraction upwind schemes for the MHD smooth vortex accuracy test with given velocity. Heun time-stepping with uniform time step  $\Delta t = 0.4h$  for bilinear Lagrangian finite element approximations and  $\Delta t = 0.01h$  for biquadratic approximations.

### Orszag–Tang Benchmark with Given Velocity Field

As for the stabilized Galerkin schemes of Chapter 3, we assess the performance of the extrusion contraction scheme in solving the more challenging MHD problem given by the Orszag–Tang vortex system (see Section 7.1.4 for further details). The velocity field is supplied at each time step. On the spatial domain  $\Omega = [0, 2]^2$  with periodic boundary conditions we consider the initial magnetic potential  $A_0 = \frac{1}{\pi} \cos(\pi y) + \frac{1}{2\pi} \cos(2\pi x)$ . The velocity field at each time step is the output of a high order finite volume discretization of the full ideal MHD system (obtained with the ALSVID-UQ code). The scalar advection problem is discretized in space on a Cartesian mesh using extrusion contraction piecewise linear and piecewise quadratic upwind schemes with velocity field averaged at the interpolation nodes and upwind direction at each node given by the averaged velocity. The polynomial order of the upwind interpolation operator coincides with the polynomial approximation degree. Heun time-stepping is used for the temporal approximation on the time interval  $I = [0, 1]$  with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ .

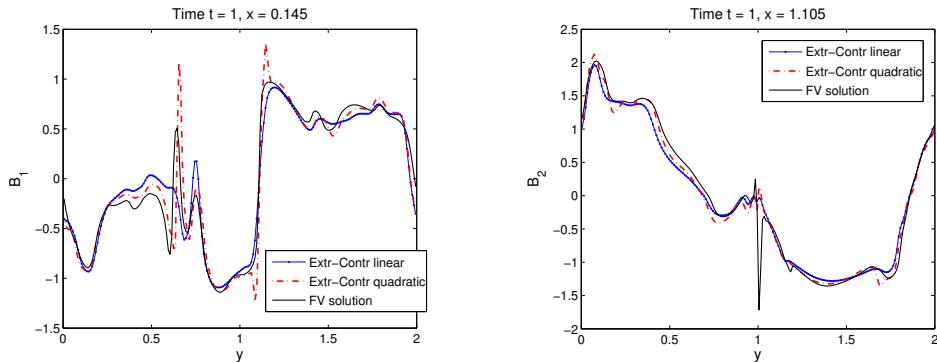


Figure 4.4.: Orszag–Tang benchmark with given velocity field. Comparison plots of the projection of the  $\mathbf{B}$  field for constant values of the  $x$ -coordinate on a Cartesian mesh with  $200 \times 200$  elements. Discretization of the advection problem for 0-forms using piecewise linear and piecewise quadratic extrusion contraction upwind schemes. As reference solution, a high order finite volume solution of the full ideal MHD system is considered.

The projection of the magnetic induction field, obtained with the foregoing discretizations, on lines at constant  $x$  is compared with the reference finite volume discrete solution in Figure 4.4. As expected, the piecewise quadratic approximation produces more accurate solutions than the piecewise linear discretization, although not as accurate as the finite volume solution. However, it can be noticed that, near shocks and discontinuities, the piecewise quadratic solution exhibits “overshoots” and “undershoots”. This is a typical by-product of numerical discretizations higher than first order accurate, as symptom of lack of monotonicity. We will reexamine this issue later in Section 4.4 and try to mend this phenomenon.

### 4.3.2. Transient Advection of 1-Forms

On the simply connected bounded domain  $\Omega \subset \mathbb{R}^2$  with Lipschitz boundary, we address the discretization of the initial boundary value problem describing the advection of the magnetic induction

$$\begin{aligned}\partial_t \mathbf{B} + \mathbf{u} \operatorname{div} \mathbf{B} + \operatorname{grad}^\perp (\mathbf{B} \cdot \mathbf{u}) &= \mathbf{f}, \quad \text{in } \Omega \times I, \\ \mathbf{B}(0) &= \mathbf{B}_0, \quad \text{in } \Omega,\end{aligned}\tag{4.11}$$

with periodic boundary conditions. In order to simplify the forthcoming presentation, we consider lowest order finite element approximations with finite element spaces of polynomial discrete differential forms of the first family, (2.25) and (2.27); extension to higher polynomial orders and second families is straightforward. Let  $\Lambda_{h,1}^1(\mathcal{T}_h) = \mathcal{P}_1^- \Lambda^1(\mathcal{T}_h)$  on a simplicial decomposition  $\mathcal{T}_h$  of  $\Omega$ ,  $\Lambda_{h,1}^1(\mathcal{T}_h) = \mathcal{Q}_1^- \Lambda^1(\mathcal{T}_h)$  on a Cartesian mesh  $\mathcal{T}_h$ . Note that the aforementioned spaces correspond to the lowest order *rotated* Raviart–Thomas elements  $\mathcal{RT}_0$ . The semi-discrete magnetic advection variational problem reads: Find  $\mathbf{B}_h(t) \in \Lambda_{h,1}^1(\mathcal{T}_h)$  such that

$$\int_{\Omega} \partial_t \mathbf{B}_h \cdot \mathbf{v} + \sum_{T \in \mathcal{T}_h} \int_T \mathcal{I}_{\mathbf{u},1}^1(\mathbf{u} \operatorname{div} \mathbf{B}_h) \cdot \mathbf{v} + \sum_{T \in \mathcal{T}_h} \int_T \operatorname{grad}^\perp (\mathcal{I}_{\mathbf{u},1}^0(\mathbf{B}_h \cdot \mathbf{u}^\perp)) \cdot \mathbf{v} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \Lambda_{h,1}^1(\mathcal{T}_h),$$

where  $\mathcal{I}_{\mathbf{u},1}^1$  is the upwind interpolation operator of polynomial degree 1 onto the polynomial space of discrete differential 1-forms  $\Lambda_{h,1}^1(\mathcal{T}_h)$ , and  $\mathcal{I}_{\mathbf{u},1}^0$  is the upwind interpolation onto  $H^1(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_1(T) = \mathcal{P}_1 \Lambda^0(\mathcal{T}_h)$ . Let  $\{\psi_{e_j}\}_{j=1}^{\#\Delta_1(\mathcal{T}_h)}$  be a basis of  $\Lambda_{h,1}^1(\mathcal{T}_h)$  and let  $\{\lambda_i\}_{i=1}^{\#\Delta_0(\mathcal{T}_h)}$  be a basis of the linear Lagrangian finite element space  $\mathcal{P}_1 \Lambda^0(\mathcal{T}_h)$ . Consider the expansion  $\mathbf{B}_h = \sum_{e_j \in \Delta_1(\mathcal{T}_h)} W_{e_j} \psi_{e_j}$ , for all  $\mathbf{B}_h \in \Lambda_{h,1}^1(\mathcal{T}_h)$ , with  $W_{e_j} \in \mathbb{R}$ . The bilinear form (4.8) associated with the discrete Lie derivative can be written, for all  $\mathbf{v} \in \Lambda_{h,1}^1(\mathcal{T}_h)$  as,

$$\begin{aligned}a_h(\mathbf{B}_h, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \int_T \mathcal{I}_{\mathbf{u},1}^1(\mathbf{u} \operatorname{div} \mathbf{B}_h) \cdot \mathbf{v} + \sum_{T \in \mathcal{T}_h} \int_T \operatorname{grad}^\perp (\mathcal{I}_{\mathbf{u},1}^0(\mathbf{B}_h \cdot \mathbf{u}^\perp)) \cdot \mathbf{v} \\ &= \sum_{T \in \mathcal{T}_h} \sum_{e_i \in \Delta_1(\mathcal{T}_h)} \langle (\mathbf{u} \operatorname{div} \mathbf{B}_h)|_{T_{e_i}^{\text{upw}}}, e_i \rangle \int_T \psi_{e_i} \cdot \mathbf{v} \\ &\quad + \sum_{T \in \mathcal{T}_h} \int_T \operatorname{grad}^\perp \left( \sum_{x_i \in \Delta_0(\mathcal{T}_h)} (\mathbf{B}_h \cdot \mathbf{u}^\perp)|_{T_{x_i}^{\text{upw}}}(x_i) \lambda_i \right) \cdot \mathbf{v} \\ &= \sum_{e_i \in \Delta_1(\mathcal{T}_h)} \sum_{e_j \in \Delta_1(\mathcal{T}_h)} W_{e_j} \langle (\mathbf{u} \operatorname{div} \psi_{e_j})|_{T_{e_i}^{\text{upw}}}, e_i \rangle \sum_{T \in \mathcal{T}_h} \int_T \psi_{e_i} \cdot \mathbf{v} \\ &\quad + \sum_{x_i \in \Delta_0(\mathcal{T}_h)} \sum_{e_j \in \Delta_1(\mathcal{T}_h)} W_{e_j} \langle \psi_{e_j} \cdot \mathbf{u}^\perp|_{T_{x_i}^{\text{upw}}}(x_i) \sum_{T \in \mathcal{T}_h} \int_T \operatorname{grad}^\perp \lambda_i \cdot \mathbf{v}.\end{aligned}\tag{4.12}$$

In this setting, employing the lowest order extrusion contraction upwind scheme with bilinear form as in (4.12), we perform a set of numerical simulations.

#### Test of Convergence: Constant Velocity

The goal of this experiment is twofold: infer the possible accuracy of the scheme and verify that solenoidal vector fields are indeed advected into solenoidal vector fields, as asserted in Proposition 4.2.4. The

magnetic advection problem (4.11) is considered on the domain  $\Omega = [0, 2]^2$  with periodic boundary conditions and in the time interval  $I = [0, 0.5]$ . The initial condition is set to  $\mathbf{B}_0 = (-\sin(\pi y), \sin(2\pi x))^\top$ , the advection velocity is constant  $\mathbf{u} = (4, 4)^\top$  and the forcing term is assumed to vanish,  $\mathbf{f} = (0, 0)^\top$ . In view of the periodic boundary conditions, we compare the solution at the final time with the initial condition.

Figure 4.5 shows that the  $L^2$ -error converges at first order rate with respect to the mesh width  $h$  and the divergence of the magnetic induction field is maintained zero up to machine precision.

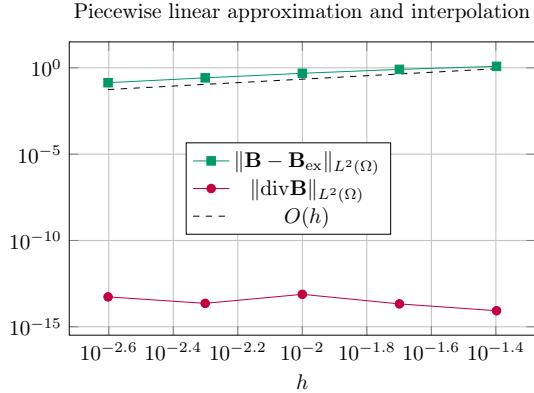


Figure 4.5.: Test for advection of 1-forms with constant velocity. Convergence plot of the error in the  $L^2$ -norm at the final time  $t = 0.5$ . Piecewise linear extrusion contraction upwind schemes with Heun time-stepping with uniform time step  $\Delta t = 0.1h$ .

### Advection of Non-Solenoidal Magnetic Induction

As a second test case, we consider the advection of the magnetic induction field with non-zero divergence with the aim of monitoring the convergence rate of the solution in the  $L^2$ -norm and in the energy norm. On the unit square  $\Omega = [0, 1]^2$  with periodic boundary conditions and in the time interval  $I = [0, 0.5]$ , we consider the magnetic advection problem (4.11) with initial condition given by

$$\mathbf{B}_0 := \begin{cases} (\varphi, \varphi)^\top & \text{if } x^2 + (y - 0.25)^2 < 0.25, \\ (0, 0)^\top & \text{otherwise,} \end{cases}$$

with  $\varphi(x, y) := \cos(\pi\sqrt{x^2 + (y - 0.25)^2})^4$ . The ‘‘hump’’ is Lie advected on the diagonal of the domain with velocity field  $\mathbf{u} = (2, 2)^\top$ . The forcing term in this experiment is set to zero,  $\mathbf{f} = (0, 0)^\top$ . We compute the numerical errors associated with the spatial discretization, at the final time  $t = 0.5$ .

Figure 4.6 shows that first order convergence is attained both in the  $L^2$ -norm and in the  $H(\operatorname{div}, \Omega)$ -seminorm.

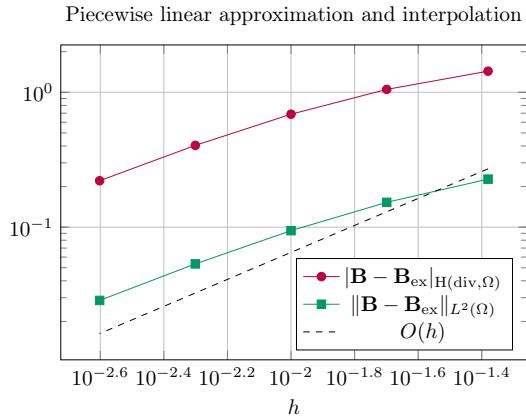


Figure 4.6.: Advection of 1-forms. Non-solenoidal magnetic induction. Numerical convergence study of the piecewise linear extrusion contraction upwind schemes. Heun time-stepping with uniform time step  $\Delta t = 0.1h$ .

### MHD Rotor Problem

As a final test case for the extrusion contraction upwind discretization of the magnetic advection problem we consider a full MHD simulation. The extrusion contraction scheme for the advection is coupled with a lowest order finite volume discretization for the fluid part of the MHD problem (1.10). We are aware that, at this point of the dissertation, we have not yet introduced all the tools needed for the discretization of the full MHD model. However, this single MHD test is presented to highlight the robustness of the extrusion contraction scheme in dealing with complex flows characterized by steep layers and shock discontinuities.

The rotor problem was introduced in [BS99, Section 3.1] and will be described in Section 7.1.5. It consists of a dense cylinder spinning on the computational domain  $\Omega = [0, 1]^2$ . The initial magnetic induction field is  $\mathbf{B}_0 = (2.5/\sqrt{4\pi}, 0)^\top$ . The velocity field is piecewise constant and is the output of the finite volume discretization of the fluid equations. The initial data concerning the fluid variables are not relevant at the moment but we refer to Section 7.1.5 for the fully detailed simulation. The numerical experiment runs until time  $T = 0.295$ . The  $\mathbf{B}$  field and its divergence are shown in Figure 4.7. The fact that the  $\mathbf{B}$  field is maintained solenoidal over the numerical computation is verified also in this case.

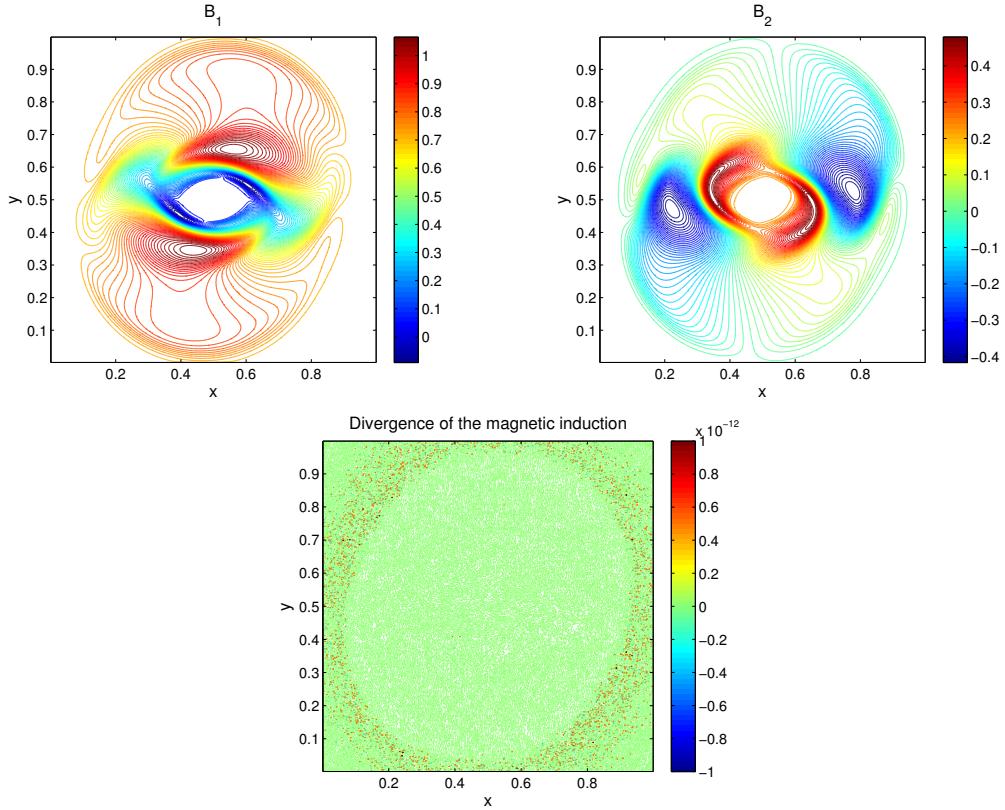


Figure 4.7.: MHD rotor problem. Lowest order extrusion contraction upwind scheme for the magnetic advection equation within the ideal MHD problem. Contour plots on a Cartesian mesh with  $400 \times 400$  elements of the magnetic induction field (top row), of the pressure (bottom left) and of the divergence of the  $\mathbf{B}$  field (bottom right). Heun time-stepping with CFL constant  $C_{\text{CFL}} = 0.4$ .

### 4.4. High Order Schemes

The stabilization terms introduced in the numerical schemes of Chapter 3 and the artificial viscosity implicitly included (through upwinding) in the discretization of the Lie derivative (4.7) do not cure the onset of spurious localized oscillations near discontinuities when high order discretizations are employed. The so-called Gibbs phenomenon characterizing high order accurate numerical solutions is manifestation

of the fact that high order numerical methods can introduce artificial extrema from initially monotone data, and are therefore *non-monotonicity preserving*. Monotone schemes mimic a property of the exact solution of conservation laws: if the initial data  $\mathbf{w}_0$  and  $\mathbf{v}_0$  satisfy  $\mathbf{w}_0(\mathbf{x}) \geq \mathbf{v}_0(\mathbf{x})$  for all  $\mathbf{x}$ , then the solutions of the two problems associated with  $\mathbf{w}_0$  and  $\mathbf{v}_0$  satisfy  $\mathbf{w}(\mathbf{x}, t) \geq \mathbf{v}(\mathbf{x}, t)$  for all  $\mathbf{x}$  and for all  $t > 0$ . By Godunov's theorem [God59], no linear high order method can yield monotone solutions. Hence, only nonlinear schemes, even in the study of linear equations, can face the conundrum of achieving high order accuracy while avoiding spurious oscillations near steep gradients.

Typically, numerical mechanisms to avoid the Gibbs effect in high order schemes rely on limiting procedures applied in the whole computational domain with the aim of suppressing spurious oscillations by reducing the accuracy of the scheme to first order in the vicinity of steep gradients or discontinuities of the solution. Monotone methods are often based upon ad hoc constructions and frequently involve trade-offs, such as violation of global conservation. The monotone scheme of Tabata [Tab77] is an example of a monotone method which is not conservative, *cf.* [Ike83, Section 4.1].

Total variation diminishing (TVD) schemes are a prominent class of nonlinear methods: high order accuracy is obtained through data reconstruction further constrained via slope limiters in order to avoid spurious oscillations. Goodman and LeVeque [GL85] showed that, in several space dimensions, TVD schemes are at most first order accurate. Since any generalized slope limiter that enforces the TVD property would unavoidably reduce to first order accuracy, high order accuracy at extrema can be restored with schemes satisfying a relaxed monotonicity-preserving requirement. Popular categories are the (weighted) essentially non-oscillatory (W)ENO reconstruction schemes, resumed in Section 6.3.1, and Runge–Kutta discontinuous Galerkin methods with total variation bounded (TVB) limiters pioneered in [CHS90; CS98]. The difficulties related to the extension of limiters to multi-dimensional problem, finite element schemes, general meshes and higher than second order polynomial degree, together with the limited theoretical understanding of the stability and convergence properties of the resulting numerical schemes have lead to alternative nonlinear stabilization approaches.

The class of schemes, called *shock-capturing* or spurious oscillations at layers diminishing (SOLD) methods [HMM86], relies on, other than a stabilization in the streamline direction, the addition of a further stabilization term acting orthogonally to the streamlines, in the crosswind direction. This idea was first implemented in [MH85] on a Petrov–Galerkin method where the test functions are modified with the added discontinuity-capturing term in the direction of the solution gradient. Since the shock-capturing term is a function of the discrete solution gradient, the resulting numerical scheme is nonlinear. Alternatively, the *entropy viscosity* methods are a class of high order schemes introduced in [GP08; GPP11] and based on the addition, to numerical schemes for hyperbolic conservation laws, of a degenerate nonlinear diffusion term. The latter is tuned locally by a numerical viscosity proportional to the local entropy production, with the rationale of introducing a “penalization” wherever the entropy conservation is violated. The  $L^2$ -stability of the entropy viscosity methods for scalar nonlinear conservation equations has recently been shown in [BGP14]. As for linear stabilization techniques (see Section 3.1), all the aforementioned nonlinear stabilization mechanisms entail the addition, although via different strategies, of some nonlinear second order viscosity. The attempt is to circumvent Godunov's theorem allowing the construction of high order schemes with controlled oscillations near large gradients of the solution. Of a slightly different flavor are the so-called *algebraic flux correction* methods [KM05], which are rooted in the flux-corrected transport idea [BB97]. After a (possibly unstable) high order discretization of a conservation law, the resulting advection matrices are modified a posteriori so to enjoy the M-matrix property. This modification of the discrete advection operator yields a low order scheme which is non-oscillatory but overly diffusive. The redundant amount of artificial diffusion is removed by adding a suitable compensating local antidiiffusion.

In the present dissertation, when using a splitting technique to solve the ideal MHD problem, the advection equation for the magnetic potential is solved in order to derive a well-defined divergence-free magnetic induction field. High order numerical schemes designed to that purpose, should ideally minimize the development of spurious oscillations in both the magnetic potential and the magnetic induction field. Let us consider the following simple example, which will be studied in detail later on. A piecewise linear scalar potential  $A_0 \in C^0([0, 1])$  (4.17) is advected in a flow at positive constant velocity. Owing to the domain periodicity, we compare the numerical solution obtained with the piecewise linear and piecewise quadratic extrusion contraction upwind schemes against the initial condition. It is evident from

Figure 4.8 that the traits of the numerical solution concern also, possibly ‘‘amplified’’, its gradient: the piecewise linear scheme is rather diffusive, the piecewise quadratic approximation yields a more accurate solution affected by mild oscillations close to the function singularities which translate into pronounced wiggles downstream of the discontinuities of the magnetic induction field.

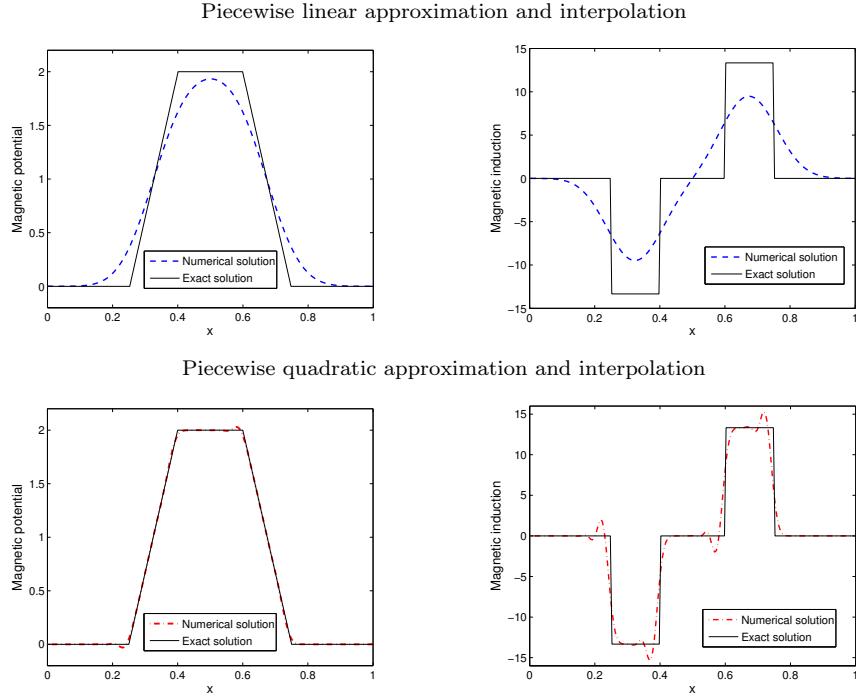


Figure 4.8.: Exact and numerical solutions on a mesh with 200 elements. Discretization with piecewise linear (top row) and piecewise quadratic (bottom row) extrusion contraction upwind schemes and Heun time-stepping with uniform time step  $\Delta t = 0.05h$ . Magnetic potential (left) and magnetic induction field (right).

In the present section we investigate how to design high order (almost) non-oscillatory numerical schemes for the generalized advection problem, with particular emphasis on the scalar transport of the magnetic potential entering the splitting scheme for the two-dimensional planar MHD problem, see Section 7.1. As much as the literature on limiters and TVD schemes to deal with discontinuities in the solution of hyperbolic problems is rich, numerical strategies to tackle Gibbs phenomena in some derivative of the solution have been rarely considered. At the best of our knowledge, the only attempt in this direction, and in a similar context, was made by Rossmanith in [Ros06]. The author develops a constrained transport discretization of the advection problem for the magnetic potential based on a modification of the wave propagation method [LeV97]. The scheme consists in a tailored-made finite volume discretization of the Taylor expansion of the unknown function at the new time instant. Second order accuracy (for smooth solutions) is achieved through high resolution corrections of the numerical fluxes associated with second order and cross derivatives. An oscillations-free magnetic induction is then obtained via the wave limiting strategy of LeVeque [LeV97] with TVD slope limiters based instead on ‘‘wave differences’’, central finite difference approximations of the second derivative of the discrete solution in each mesh element.

**Remark 4.4.1.** We must acknowledge that the difficulties related to the design of limiting strategies to counteract spurious oscillations in the gradient of the numerical solutions provide one of the reason why divergence cleaning techniques for MHD flows based on the evolution of the magnetic potential have seldom been designed. Since the discrete Lie derivative, derived with the extrusion contraction approach, enjoys a commuting diagram property (*cf.* Proposition 4.2.4), the splitting algorithm for the discretization of the ideal MHD system can be based directly on the advection problem for the  $\mathbf{B}$  field without extra actions to enforce the divergence constraint.

#### 4.4.1. Nonlinear Residual-Based Viscosity for Extrusion Contraction

The entropy viscosity methods introduced in [GP08] are particularly attractive for the aforementioned task of rendering high order extrusion contraction upwind schemes devoid of oscillations in the exterior derivative of the solution of the generalized advection problem. We rephrase the entropy viscosity method for nonlinear conservation laws proposed in [GPP11], to accommodate the generalized advection problem (3.1), recalled in Section 4.2.

Let us consider the generalized advection problem for  $k$ -forms,  $0 \leq k \leq n - 1$ , with, for simplicity, no source term  $f = 0$ . The semi-discrete variational formulation (4.8) is augmented with a *degenerate* viscosity term so that the problem reads: Find  $\omega_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$  such that  $(\omega_h(0), \eta_h)_\Omega = (\omega_0, \eta_h)_\Omega$  for all  $\eta_h \in \Lambda_{h,r}^k(\mathcal{T}_h)$  and

$$(\partial_t \omega_h, \eta_h)_\Omega + a_h(\omega_h, \eta_h) + b_h(\omega_h, \eta_h) = 0, \quad \forall \eta_h \in \Lambda_{h,r}^k(\mathcal{T}_h), \quad (4.13)$$

where

$$b_h(\omega_h, \eta_h) := \sum_{T \in \mathcal{T}_h} \nu_T(t, \omega_h(t)) (\mathbf{d}^k \omega_h, \mathbf{d}^k \eta_h)_T, \quad \forall \omega_h, \eta_h \in \Lambda_{h,r}^k(\mathcal{T}_h). \quad (4.14)$$

The function  $\nu(\mathbf{x}, t, \omega_h(t)) = \nu_T(t, \omega_h(t)) \mathbb{1}_{T \in \mathcal{T}_h} \in \mathcal{P}_0(\mathcal{T}_h \times (t^n, t^n + \Delta t^n))$  for all  $n$ , is the *viscosity coefficient*. The discretization of the bilinear form  $b_h(\cdot, \cdot)$  is standard. To the purpose of defining the viscosity coefficient, we conjecture that the oscillations in the discrete magnetic induction field can be controlled by tuning the artificial viscosity with an indicator function based on the  $\mathbf{B}$  field and not on the magnetic potential. This reasoning is justified only in view of Proposition 4.2.4: the exterior derivative of the discrete differential form obtained via extrusion contraction upwinding of the advection problem for  $k$ -forms, satisfies in its turn an extrusion contraction upwind discretization of the advection problem for  $(k+1)$ -form. Hereafter, we conduct the derivation of the viscosity coefficient based on the equation residual rather than on some entropy residual. At a fixed time  $t^n$ , let  $\omega_h^n$  be the numerical solution of (4.13) and let  $\gamma_h^n := \mathbf{d}^k \omega_h^n$ . The idea of how to define the viscosity coefficient is presented along the lines of [GPP11, Section 2] with the details left to the forthcoming section. As a first step, the residual operator associated with the generalized advection problem for a  $(k+1)$ -form  $\gamma_h$  is given by

$$(R(t))(\gamma_h) := (\partial_t \gamma_h)(t) + \mathbf{i}_{\mathbf{u}} \mathbf{d}^{k+1} \gamma_h(t) + \mathbf{d}^k \mathbf{i}_{\mathbf{u}} \gamma_h(t). \quad (4.15)$$

The discretization  $R_h$  of  $R$  is based on extrusion contraction upwind schemes and on a suitable time-stepping. Concerning the advection operator, since  $\omega_h^n \in H\Lambda^k(\Omega)$ , then  $\mathbf{d}^{k+1} \gamma_h^n \equiv 0$  by construction, whilst

$$\mathbf{d}^k \mathbf{i}_{\mathbf{u}} \gamma_h^n \approx \mathbf{d}^k \mathcal{I}_{\mathbf{u}, p^-}^k(\mathbf{i}_{\mathbf{u}} \gamma_h(t)) = \mathbf{d}^k \sum_{j=k}^{M_{\min}^{p-,k}} \sum_{f_j \in \Delta_j(T)} \sum_{\ell=1}^{N_j} W_{f_j}^\ell (\mathbf{i}_{\mathbf{u}} \gamma_h|_{T_{f_j}^{\text{upw}}}) \psi_{j,T}^\ell.$$

The artificial viscosity term (4.14) is clearly nonlinear in the sense that  $\nu$  depends on the solution. However, it can be rendered explicit in time by choosing  $\nu = \nu(\mathbf{x}, t^n, \omega_h(s))$  for  $s \leq t^n$ , namely approximating the temporal derivative in (4.15) with an explicit time-stepping (of appropriate order),  $(\partial_t \gamma_h)(t^n) \approx (\partial_t \gamma_h)_h^n$ . The viscosity coefficient is then computed based on the discrete residual  $R_h^n(\gamma_h^n) := (\partial_t \gamma_h)_h^n + \mathbf{d}^k \mathcal{I}_{\mathbf{u}, p^-}^k(\mathbf{i}_{\mathbf{u}} \gamma_h(t^n))$  under suitable normalizations and controlled by a threshold value having the dimension of a first order viscosity. In greater detail,  $\nu_T(t^n, \omega_h(t^n)) =: \nu_T^n = \min\{\nu_T^E(t^n), \nu_T^{\max}(t^n)\}$  with  $\nu_T^E$  derived as in [GPP11, Equation (2.4)], namely,

$$\nu_T^E(t^n) := c_E h_T^2 \frac{\|R_h^n(\gamma_h^n)\|_{L^\infty(T)}}{\left\| \gamma_h^n - \int_{\Omega} \gamma_h^n \right\|_{L^\infty(\Omega)}},$$

where the  $L^\infty$ -norm on each element  $T \in \mathcal{T}_h$  is taken over quadrature points on  $T$  and  $c_E \in \mathbb{R}$ ,  $c_E > 0$  is a positive constant. The first order viscosity reads as in [GPP11, Equation (2.5)],

$$\nu_T^{\max}(t^n) := c_{\max} h_T \| \mathbf{F}'(\gamma_h^n) \|_{L^\infty(T)}, \quad (4.16)$$

where  $\mathbf{F}'$  is the Jacobian matrix of the flux  $\mathbf{F}$  associated with the advection problem for  $(k+1)$ -forms in conservation form and  $c_{\max} \in \mathbb{R}$ ,  $c_{\max} > 0$ . In two dimensions,  $k \in \{0, 1\}$  and the flux function  $\mathbf{F}_{k+1}$  associated to the advection of  $(k+1)$ -forms can be expressed in vector proxies as:  $\mathbf{F}_1(\mathbf{w}) = \mathbf{w} \otimes \mathbf{u} - \mathbf{u} \otimes \mathbf{w}$  and  $\mathbf{F}_2(w) = \mathbf{u} w$ .

### Scalar Advection Problem in 2D

In the present section, spurred on by the two-dimensional planar MHD problem, we consider the residual-based scheme for the scalar advection equation (4.10) with no inflow boundary or with periodic boundary conditions. Let  $\Lambda_{h,r}^0(\mathcal{T}_h) \subset H^1(\Omega)$  be the space of piecewise polynomial scalar functions of degree at most  $r \geq 1$  and let  $\mathcal{D}_h^0 = \Lambda_{h,p}^0(\mathcal{T}_h)$  in Definition 4.2.1 with  $p \geq 1$  and  $N_D := \dim \mathcal{D}_h^0$ . The semi-discrete approximation (4.13) of (4.10) via extrusion contraction upwind scheme has bilinear form, for every  $A_h, v_h \in \Lambda_{h,r}^0(\mathcal{T}_h)$ ,

$$a_h(A_h, v_h) = \int_{\Omega} (\mathbf{u} \cdot \operatorname{grad} A_h) v_h = \int_{\Omega} \mathcal{I}_{\mathbf{u},p}^0(\mathbf{u} \cdot \operatorname{grad} A_h) v_h = \sum_{k=1}^{N_D} \mathbf{u}(x_k) \cdot (\operatorname{grad} A_h)|_{T_{x_k}^{\text{upw}}}(x_k) \int_{\Omega} \lambda_k v_h,$$

where  $\{\lambda_k\}_{k=1}^{N_D}$  is the barycentric coordinates basis of  $\mathcal{D}_h^0$  and  $\{x_k\}_{k=1}^{N_D}$  are the global interpolation nodes for the Lagrangian finite element. The viscosity term (4.14) is

$$b(A_h, v_h) := \sum_{T \in \mathcal{T}_h} \nu_T(t, \mathbf{B}_h(t)) (\operatorname{grad} A_h, \operatorname{grad} v_h)_T, \quad \forall A_h, v_h \in \Lambda_{h,r}^0(\mathcal{T}_h).$$

The entropy viscosity coefficient in each mesh cell  $T \in \mathcal{T}_h$  at a fixed time  $t^n$  is derived as follows. Let  $A_h^n$  be the numerical solution of (4.13) with bilinear forms defined above and suitable time-stepping. Let the induction field at time  $t^n$  be  $\mathbf{B}_h^n := \operatorname{grad}^\perp A_h^n$ .

1. Define the residual of the advection problem (4.11) for the magnetic induction  $\mathbf{B}_h$ , as

$$\mathbf{R}(\mathbf{B}_h) := \partial_t \mathbf{B}_h + \mathbf{u} \operatorname{div} \mathbf{B}_h + \operatorname{grad}^\perp (\mathbf{B}_h \cdot \mathbf{u}^\perp).$$

The discretization  $\mathbf{R}_h$  of  $\mathbf{R}$  is based on the extrusion contraction upwind approximation of the advection operator,

$$\operatorname{grad}_h^\perp (\mathbf{B}_h^n \cdot \mathbf{u}^\perp(t^n)) \approx \operatorname{grad}^\perp (\mathcal{I}_{\mathbf{u},p^-}^0(\mathbf{B}_h^n \cdot \mathbf{u}^\perp(t^n))) = \sum_{k=1}^{N_D} (\mathbf{B}_h^n)|_{T_{x_k}^{\text{upw}}}(x_k) \cdot \mathbf{u}^\perp(x_k, t^n) \operatorname{grad}^\perp \lambda_k,$$

where  $N_{D^-} := \dim \Lambda_{h,0}^{p^-}(\mathcal{T}_h)$  for  $p^- \geq 1$ . Note that  $\operatorname{div} \mathbf{B}_h^n \equiv 0$  since  $A_h^n \in H^1(\Omega)$ . Furthermore, in order to make the viscosity explicit without affecting the accuracy, a second order explicit time-stepping is applied. In the numerical simulations we used, as suggested in [GPP11, Remark 2.3], a BDF2 temporal discretization, namely

$$(\partial_t \mathbf{B}_h)_h^n = \frac{\Delta t^{n-1}}{\Delta t^{n-2}(\Delta t^{n-1} + \Delta t^{n-2})} \mathbf{B}_h^{n-2} - \frac{\Delta t^{n-2} + \Delta t^{n-1}}{\Delta t^{n-2} \Delta t^{n-1}} \mathbf{B}_h^{n-1} + \frac{\Delta t^{n-2} + 2\Delta t^{n-1}}{\Delta t^{n-1}(\Delta t^{n-1} + \Delta t^{n-2})} \mathbf{B}_h^n.$$

Hence,  $\mathbf{R}_h^n(\mathbf{B}_h^n) = (\partial_t \mathbf{B}_h)_h^n + \sum_{k=1}^{N_{D^-}} (\mathbf{B}_h^n)|_{T_{x_k}^{\text{upw}}}(x_k) \cdot \mathbf{u}^\perp(x_k, t^n) \operatorname{grad}^\perp \lambda_k$ .

2. Define the viscosity coefficient as

$$\nu_T^E(t^n) := c_E h_T^2 \frac{\left( \sum_{i=1}^n \sup_{\mathbf{x} \in T} |(\mathbf{R}_h^n(\mathbf{B}_h^n))_i(\mathbf{x})|^2 \right)^{1/2}}{\left( \sum_{i=1}^n \sup_{\mathbf{x} \in \Omega} \left| (\mathbf{B}_h^n)_i(\mathbf{x}) - \int_{\Omega} (\mathbf{B}_h^n)_i \right|^2 \right)^{1/2}},$$

where the supremum on each element  $T \in \mathcal{T}_h$  is taken over quadrature points on  $T$ . Specifically, let  $Q(T) = \{\mathbf{x}_j, q_j\}_j$  be a local quadrature rule on  $T$ , then, for all  $1 \leq i \leq n$ ,

$$\begin{aligned} \sup_{\mathbf{x} \in T} |(\mathbf{R}_h^n(\mathbf{B}_h^n))_i(\mathbf{x})|^2 &\approx \sup_{\mathbf{x}_j \in Q(T)} |(\mathbf{R}_h^n(\mathbf{B}_h^n))_i(\mathbf{x}_j)|^2, \\ \sup_{\mathbf{x} \in \Omega} \left| (\mathbf{B}_h^n)_i(\mathbf{x}) - \int_{\Omega} (\mathbf{B}_h^n)_i \right| &\approx \sup_{T \in \mathcal{T}_h} \sup_{\mathbf{x}_j \in Q(T)} \left| (\mathbf{B}_h^n)_i(\mathbf{x}_j) - \frac{1}{|\Omega|} \sum_{T \in \mathcal{T}_h} \sum_{\ell} q_{\ell} (\mathbf{B}_h^n)_i(\mathbf{x}_{\ell}) \right|. \end{aligned}$$

3. Introduce a first order viscosity as upper bound, namely

$$\nu_T^{\max} := c_{\max} h_T \|\mathbf{S}\|, \quad (\mathbf{S})_{i,j} := \sup_{\mathbf{x} \in T} |\mathbf{F}'_{i,j}(\mathbf{B}_h^n)(\mathbf{x})|, \quad i, j \in \{1, 2\},$$

where  $\|\cdot\|$  is some matrix norm and  $\mathbf{F}'$  is the Jacobian matrix of the flux  $\mathbf{F}$  appearing in the magnetic advection equation (4.11) recast as a system of conservation laws. Here, using the Euclidean vector norm, since  $\mathbf{u} \in \mathcal{P}_0(\mathcal{T}_h)^2$ ,  $\|\mathbf{S}\| = \|\mathbf{u}|_T(t^n)\|_{\ell^2}$ .

4. Define the entropy viscosity coefficient in (4.14) as  $\nu_T(t^n, \mathbf{B}_h(t^n)) = \min\{\nu_T^E, \nu_T^{\max}\}$ .

In the alternative situation occurring by directly discretizing the magnetic induction problem, as pointed out in Remark 4.4.1, one should rely on the addition of artificial magnetic diffusion, based again on the induction residual as above derived. The augmented discrete operator obtained from the Lie derivative plus the second order artificial diffusion will still satisfy a commuting diagram property, owing to the fact that the exterior derivative is a differential, i.e.,  $d^{k+1} \circ d^k = 0$ . On the other hand, it is not straightforward to gauge the effectiveness of the artificial viscosity since the nonlinear second order stabilization will have no impact on the kernel of the exterior derivative. The question arises an interesting topic of further investigations.

### Advection of the Magnetic Potential with Discontinuous Magnetic Induction

In order to appraise the ability of the extrusion contraction upwind scheme with residual-based viscosity in providing stable and accurate solutions of the scalar advection and simultaneously of the solution derivatives, we consider a two-dimensional variant of the one-dimensional advection problem, proposed in [Ros06, Fig. 5.1]. The piecewise linear initial datum

$$A_0(x) = \begin{cases} 0 & \text{if } x < 0.25, \\ \frac{x - 0.25}{0.075} & \text{if } 0.25 < x < 0.4, \\ 2 & \text{if } 0.4 < x < 0.6, \\ \frac{0.75 - x}{0.075} & \text{if } 0.6 < x < 0.75, \\ 0 & \text{if } x > 0.75, \end{cases} \quad (4.17)$$

is advected with constant and positive velocity field  $\mathbf{u} = (1, 0)^\top$  on the domain  $\Omega = [0, 1]^2$  until time  $T = 0.25$ .

In Figure 4.9, the numerical solutions of the extrusion contraction upwind scheme with piecewise linear, piecewise quadratic and piecewise quadratic residual-based viscosity approximations are compared. The residual-based viscosity has parameters  $c_E = 0.5$  and  $c_{\max} = 0.05$ . The magnetic induction field obtained with the addition of the residual-based viscosity terms delivers solutions which are more accurate than the piecewise linear ones and non-oscillatory if compared to the piecewise quadratic discretization.

The choice of the tunable parameters  $c_E$  and  $c_{\max}$  has to be adjusted to the problem data at hand. Poor choices of the viscosity parameters might generate unstable derivatives of the solutions, as in Figure 4.10, or yield stable but overdissusive solutions. It appears to us that there is no consistent way to tune the positive constants  $c_E$  and  $c_{\max}$ . This constitutes a drawback of the entropy and residual-based viscosity schemes.

In the numerical simulations we performed, the amount of added nonlinear viscosity did not allow to achieve second order accurate (for smooth solutions) schemes for the advection of the potential alongside non-oscillatory discrete induction fields. Usually, oscillations-free inductions were associated with almost first order convergence (in  $L^2$ -norm), whilst weighting the residual-based viscosity so to have second order convergence resulted in polluted or unstable induction fields. The same considerations held when using the first order viscosity in (4.16) only as indicator threshold function not entering the viscosity coefficient, namely for  $\nu_T(t^n, \mathbf{B}_h(t^n)) = \nu_T^E$ , if  $\nu_T^E < \nu_T^{\max}$ , and  $\nu_T(t^n, \mathbf{B}_h(t^n)) = 0$  otherwise.

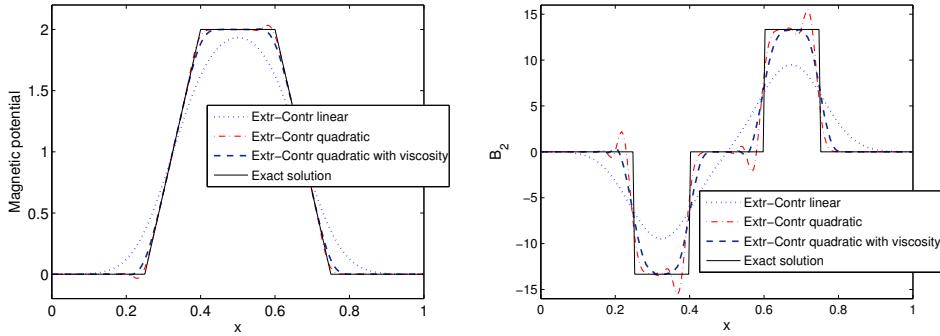


Figure 4.9.: Advection of piecewise linear continuous magnetic potential. Comparison of the projection of the exact and numerical solution (left) and second component of the magnetic induction field (right) for different numerical discretizations on a  $200 \times 200$  Cartesian mesh. Heun time-stepping with uniform time step  $\Delta t = 0.01h$ . The residual-based viscosity has parameters  $c_E = 0.5$  and  $c_{\max} = 0.05$ .

Piecewise quadratic approximation and interpolation. Residual-based nonlinear viscosity.

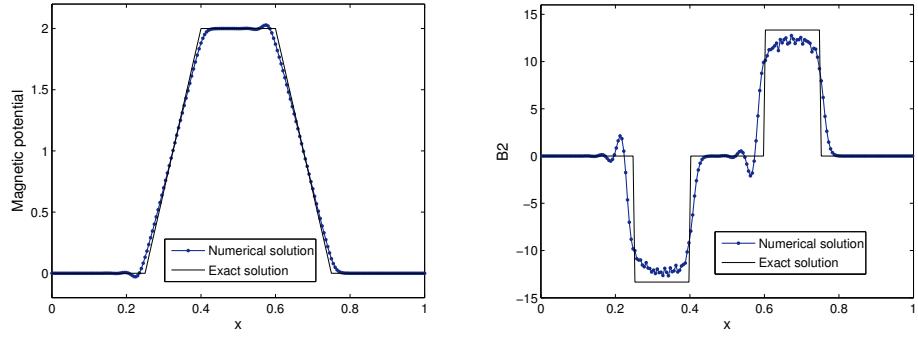


Figure 4.10.: Test case piecewise linear continuous magnetic potential. Projection of the exact and numerical solution (left) and second component of the magnetic induction field (right). Discretization with piecewise biquadratic extrusion contraction upwind scheme on a  $200 \times 200$  Cartesian mesh and Heun time-stepping with uniform time step  $\Delta t = 0.01h$ . The residual-based viscosity has parameters  $c_E = 0.1$  and  $c_{\max} = 0.5$ .

### Orszag–Tang Benchmark with Given Velocity

We compare the performances, in terms of development of oscillations, of piecewise quadratic extrusion contraction upwind schemes on the Orszag–Tang benchmark with given velocity field presented in Section 4.3.1. The advection problem for the magnetic potential is discretized on the spatial domain  $\Omega = [0, 2]^2$  with periodic boundary conditions and on the time interval  $I = [0, 1]$  with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ . The initial magnetic potential is  $A_0 = \frac{1}{\pi} \cos(\pi y) + \frac{1}{2\pi} \cos(2\pi x)$ . The residual-based viscosity is designed as in Section 4.4.1 with parameters  $c_E = 0.5$  and  $c_{\max} = 0.05$ .

Figure 4.11 shows the projection of the magnetic induction field on lines at constant  $x$ . The residual-based viscosity extrusion contraction upwind scheme exhibits accuracy comparable with the one of the piecewise quadratic scheme (with no extra viscosity) and, in addition, the “overshoots” and “undershoots” in the magnetic induction field are mitigated when not completely eliminated.

#### 4.4.2. Nonlinear Residual-Based Viscosity for Stabilized Discontinuous Galerkin

Taking the cue from the entropy viscosity methods for DG discretizations introduced in [Zin+13], we adapt the scheme discussed in Section 4.4.1 to the stabilized DG scheme (from Chapter 3) for the scalar

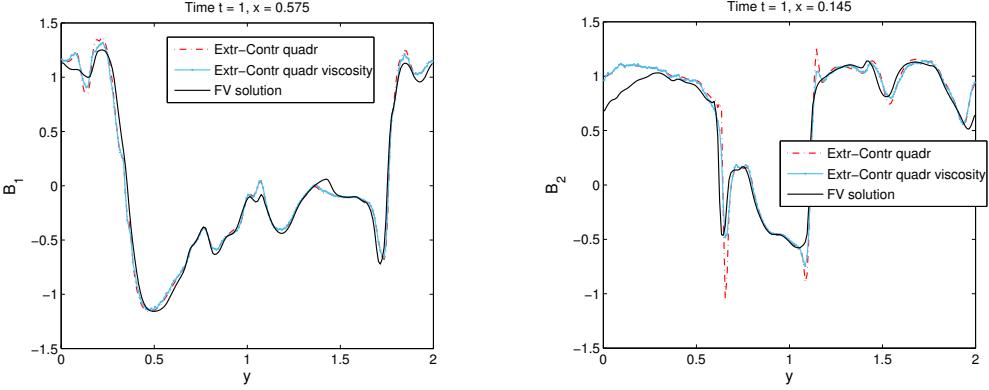


Figure 4.11.: Orszag–Tang benchmark with given velocity field. Comparison plots of the projection of the  $\mathbf{B}$  field for constant values of the  $x$ -coordinate. Discretization of the advection problem for the magnetic potential using piecewise biquadratic extrusion contraction upwind scheme with and without residual-based viscosity terms. The coefficient for the viscosity term are  $c_E = 0.5$  and  $c_{\max} = 0.05$ . Heun time-stepping with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ . Cartesian mesh with  $200 \times 200$  elements. As reference solution we consider the output of a high order finite volume discretization of the full ideal MHD system.

advection problem, where we assume, for simplicity, constant velocities<sup>1</sup>. Let  $\Lambda_h^{d,k}(\mathcal{T}_h)$  be the space of piecewise polynomials of degree at most  $r \geq 0$  on the mesh  $\mathcal{T}_h$  satisfying Assumption 2.4.2. With the notations introduced in Chapter 3, the semi-discrete stabilized discontinuous Galerkin scheme (3.7)-(3.9) for (4.10) with Dirichlet inflow boundary conditions  $A = g \in L^2(\Gamma_{\text{in}})$  on  $\Gamma_{\text{in}}$  reads: Find  $A_h(t) \in \Lambda_h^{d,k}(\mathcal{T}_h)$  such that

$$(\partial_t A_h, v_h)_\Omega + a_h(A_h, v_h) = l(v_h), \quad \forall v_h \in \Lambda_h^{d,k}(\mathcal{T}_h). \quad (4.18)$$

The bilinear form  $a_h(\cdot, \cdot)$  and the linear function  $l(\cdot)$  are defined, for all  $A_h, v_h \in \Lambda_h^{d,k}(\mathcal{T}_h)$  as

$$\begin{aligned} a_h(A_h, v_h) := & \sum_{T \in \mathcal{T}_h} \int_T -A_h \operatorname{div}(\mathbf{u} v_h) d\mathbf{x} + \sum_{f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \setminus \Delta_{n-1}^{\partial,-}(\mathcal{T}_h)} \int_f \mathbf{u} \cdot \mathbf{n}_f A_h v_h dS \\ & + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \{A_h\}_f [\mathbf{u} v_h]_f \cdot \mathbf{n}_f dS + \sum_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)} \int_f \bar{c}_f [A_h]_f [\mathbf{u} v_h]_f \cdot \mathbf{n}_f dS, \\ l(v_h) := & \int_\Omega f v_h d\mathbf{x} - \sum_{f \in \Delta_{n-1}^{\partial,-}} \int_f \mathbf{u} \cdot \mathbf{n}_f g v_h dS. \end{aligned}$$

“Upwind” fluxes are obtained with the choice of the stabilization parameter  $\bar{c}_f = \operatorname{sign}(\{\mathbf{u}\}_f \cdot \mathbf{n}_f)/2$ , for all  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ . Upon semi-discretization of the advection problem, we add a second order consistent diffusion term as in (4.13); thereby,

$$(\partial_t A_h, v_h)_\Omega + a_h(A_h, v_h) + \sum_{T \in \mathcal{T}_h} \nu_T(t, A_h(t)) (\operatorname{grad} A_h, \operatorname{grad} v_h)_T = l(v_h), \quad \forall v_h \in \Lambda_h^{d,k}(\mathcal{T}_h).$$

The residual-based coefficient  $\nu_T$  at time  $t^n$  and on a fixed element  $T \in \mathcal{T}_h$  is set to  $\nu_T(t^n, A_h^n) = \min\{\nu_T^E(t^n), \nu_T^{\max}(t^n)\}$ , where the first order viscosity  $\nu_T^{\max}$  is as in (4.16), whereas the coefficient  $\nu_T^E$  is defined as,

$$\nu_T^E(t^n) := c_E h_T^2 \frac{\max\{\|R_h^n(A_h^n)\|_{L^\infty(T)}, \max_{f \in \Delta_{n-1}(T)} \max_{x \in f} |(J_f^n(A_h^n))(x)|\}}{\left\| A_h^n - \int_\Omega A_h^n \right\|_{L^\infty(\Omega)}}.$$

Here  $J_f^n(A_h^n) := h_f^{-1} \{ \mathbf{u} \}_f [A_h^n]_f$  is the residual associated with the jump of the solution across the mesh interfaces, all other quantities and the approximation of the  $L^\infty$ -norm are as in Section 4.4.1.

<sup>1</sup>The viscosity method is designed for conservation laws. For non-constant velocity the advection problem has to be considered as a conservation law in quasilinear form with variable coefficients.

The derivation of entropy viscosity methods for discontinuous Galerkin discretizations arises an ambiguity, namely whether the artificial diffusion is a local contribution active only in the interior of each cell or if it entails an interface flux contribution as well. In [Zin+13] a further degenerate penalization is added on the mesh  $(n - 1)$ -skeleton. Specifically, the numerical fluxes associated with the second order diffusion are defined in the spirit of an interior penalty method, so that the consistent artificial viscosity term reads

$$\sum_{T \in \mathcal{T}_h} \nu_T (\operatorname{grad} A_h, \operatorname{grad} v_h)_T + \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h)} \int_f -\{\nu \operatorname{grad} A_h\}_f \cdot \mathbf{n}_f v_h dS + \int_f \delta h_f^{-1} \{\nu\}_f [A_h]_f [v_h]_f dS,$$

with  $\delta \in \{0, 1\}$  and  $\nu$  as above-stated. In all the numerical tests we tried, no significant difference could be recorded between the two variants of the scheme.

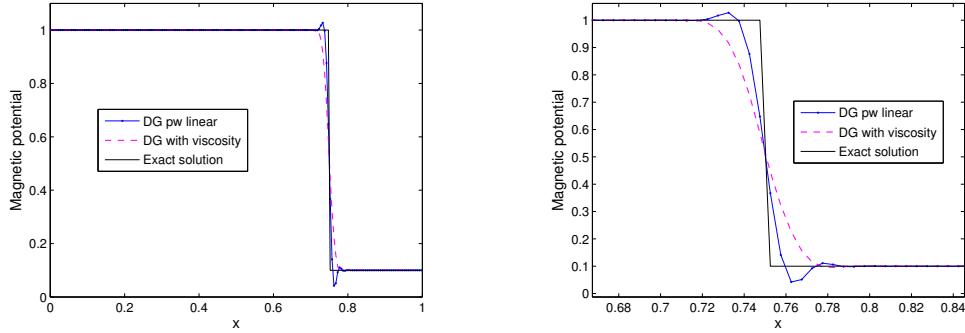
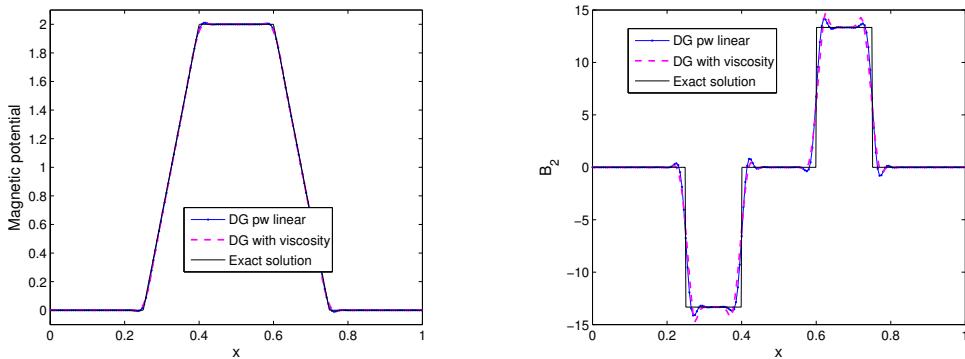


Figure 4.12.: Advection of step function at constant velocity field. Stabilized discontinuous Galerkin scheme on a  $200 \times 200$  mesh. Heun time-stepping with uniform time step  $\Delta t = 0.1h$ . Left: numerical solution obtained with and without residual-based viscosity terms. Right: closeup of the numerical solution close to the discontinuity.

As shown in Figure 4.12, the residual-based viscosity DG scheme is well-suited to deal with discontinuities in the magnetic potential. However, if not properly adjusted to heed the behavior of the magnetic induction, the scheme does not supply an effective limiting of the oscillations arising in the derivative of the solution, see Figure 4.13. Moreover, as it can be observed in Figure 4.13 (top left plot), not even the oscillations of the magnetic potential are cured: this might hint to the fact that the indicator function regulating the amount of artificial viscosity is too weak to deal with rather small oscillations and non-discontinuous solutions. Weighting the entropy viscosity term in the stabilized Galerkin numerical scheme for the magnetic potential advection, in order to control the oscillations of the  $\mathbf{B}$  field, seems not a trivial task, owing to the fact that a discrete commuting property is not satisfied by the discretization (4.18).



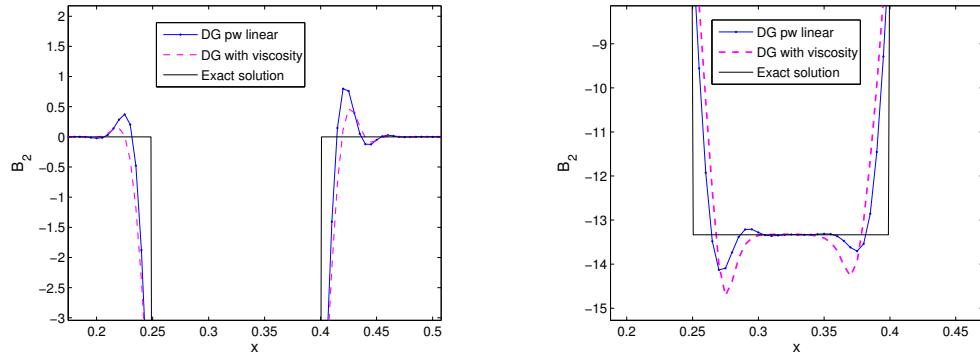


Figure 4.13.: Advection of piecewise linear magnetic potential at constant velocity field. Stabilized discontinuous Galerkin scheme on a  $200 \times 200$  mesh. Heun time-stepping with uniform time step  $\Delta t = 0.1h$ . Numerical solution obtained with and without residual-based viscosity terms: magnetic potential (top left) and magnetic induction field (top right). In the bottom row, closeup of the magnetic induction field in the vicinity of the discontinuity.

## 5. Preconditioning the Magnetic Advection-Diffusion Problem

As result of the onset of small-scale structures like current sheets or magnetic turbulence, resistive effects act on the magnetic configurations of plasmas, *cf.* Section 1.1. In particular, plasma motions can “push together” magnetic fields lines: while in ideal MHD flows this dynamics is (magnetic) flux conserving and hence keeps the magnetic topology unchanged, in resistive MHD the magnetic field configuration can be restructured in a new topology with a lower energy state. The time scale at which reconnection takes place is dictated by the electric resistivity  $\eta = \sigma^{-1}$  and by the magnetic permeability  $\mu$ .

The  $\mathbf{A}$ -based eddy current model derived in Section 2.2 with non-vanishing resistive diffusion reads

$$\sigma \partial_t \mathbf{A} + \mathbf{curl}(\mu^{-1} \mathbf{curl} \mathbf{A}) + \sigma \mathbf{curl} \mathbf{A} \times \mathbf{u} + \sigma \operatorname{grad}(\mathbf{u} \cdot \mathbf{A}) = \mathbf{0},$$

with the double  $\mathbf{curl}$  representing the diffusion operator. This is a parabolic problem in  $H(\mathbf{curl}, \Omega)$ : L-stable implicit Runge–Kutta schemes are a natural choice for the time integration of the electromagnetic fields, winnowing out too stringent time-step restrictions imposed by explicit methods. On the other hand, the hyperbolic (non-symmetric) character of the advective operator and the potential presence of nonlinear stabilizations or slope limiters (see Chapter 4), suggests that explicit time-stepping would be efficient. Treating the advection explicitly and the second order operator implicitly can be realized through the so-called *implicit-explicit* (IMEX) time-stepping schemes.

The use of (discontinuous) Galerkin spatial discretizations, as in Chapter 3, together with an implicit-explicit time-stepping entails solving discrete boundary value problems for the double  $\mathbf{curl}$  operator in each time step. This chapter is devoted to the construction and analysis of fast iterative solvers to efficiently tackle the resulting large sparse linear systems of equations.

Specifically, in each time step of a partly implicit time-stepping scheme we have to solve an  $H(\mathbf{curl}, \Omega)$ -elliptic boundary value problem, which, in abstract form, reads

$$\begin{aligned} \mathbf{curl}(\nu \mathbf{curl} \mathbf{v}) + \beta \mathbf{v} &= \mathbf{f}, & \text{in } \Omega, \\ \operatorname{tr} \mathbf{v} &= \mathbf{0}, & \text{on } \partial\Omega. \end{aligned} \tag{5.1}$$

Here  $\Omega \subset \mathbb{R}^n$ ,  $n \in \{2, 3\}$ , is a bounded domain with *trivial topology* and Lipschitz boundary  $\partial\Omega$ ,  $\mathbf{f} \in L^2(\Omega)^n$ , and  $\nu(\mathbf{x})$  and  $\beta(\mathbf{x})$  are possibly discontinuous coefficients, which are assumed to be positive and bounded functions in  $\Omega$ . They represent properties of the medium or material: in resistive MHD flows,  $\nu$  is typically the inverse of the magnetic permeability and  $\beta$  is proportional to the ratio of electrical conductivity and the time step. Resistive terms become sizable in association with specific resistive modes like tearing modes, turbulence, and high current concentration, all small-scale local phenomena. Therefore, it is desirable to develop robust approximations (and hence robust solvers) to properly handle the switch from relative large  $\nu$  to relative large  $\beta$ .

The boundary value problem (5.1) cast in weak form delivers the  $H(\mathbf{curl}, \Omega)$ -elliptic variational formulation: Find  $\mathbf{v} \in H_0(\mathbf{curl}, \Omega)$  such that

$$a(\mathbf{v}, \mathbf{z}) := (\nu \mathbf{curl} \mathbf{v}, \mathbf{curl} \mathbf{z})_{L^2(\Omega)} + (\beta \mathbf{v}, \mathbf{z})_{L^2(\Omega)} = (\mathbf{f}, \mathbf{z})_{L^2(\Omega)}, \quad \forall \mathbf{z} \in H_0(\mathbf{curl}, \Omega), \tag{5.2}$$

where the Hilbert space  $H_0(\mathbf{curl}, \Omega)$  is defined as in Section 2.1 (see Table 2.2),

$$H_0(\mathbf{curl}, \Omega) := \overset{\circ}{H}\Lambda^1(\Omega) = \{\mathbf{v} \in H(\mathbf{curl}, \Omega) : \mathbf{n} \times \mathbf{v} = \mathbf{0} \text{ on } \partial\Omega\},$$

and endowed with the graph norm  $\|\mathbf{v}\|_{H\Lambda^1(\Omega)}^2 := \|\mathbf{v}\|_{L^2(\Omega)}^2 + \|\mathbf{curl} \mathbf{v}\|_{L^2(\Omega)}^2$ . The assumptions  $\nu > 0$ ,  $\beta > 0$  a.e. in  $\Omega$  ensure existence and uniqueness of solutions of (5.2) (see e.g. [Eva98, Chapter 6]).

## Preconditioners for Conforming Finite Element Methods

Subspace correction preconditioners in the context of conforming Galerkin finite element discretizations of  $H(\mathbf{curl}, \Omega)$ -elliptic variational problems are well-established both in the form of multigrid [Hip99a; Hip03; AFW00; HX07; HZ09] and domain decomposition methods [HT00; Tos00b]. While the authors prove uniform performance with respect to mesh refinement, the presence of discontinuities in the coefficients is not included in their analyses. However, in [Hip99a] numerical evidence hints that multigrid methods in  $H(\mathbf{curl}, \Omega)$  are affected by jumping coefficients in a similar way as their scalar counterparts.

For conforming and nonconforming discretizations of scalar second order elliptic problems, the design preconditioning strategies, aiming at robustness with respect to the jumps in the diffusion coefficient, has received considerable attention. However, the asymptotic convergence of multilevel solvers for conforming discretizations in the presence of two different coefficients (i.e., a reaction-diffusion problem or one resulting upon time discretization of a parabolic model), has only been recently addressed in [KXZ15]. The solvers designed in [KXZ15] prove robust only when one of the two coefficients is constant, or if both coefficients have the same pattern distribution.

Much less progress has been made in the context of  $H(\mathbf{curl}, \Omega)$ -elliptic problems with two variable coefficients. In spite of the relevance of the problem, the bulk of contributions is largely restricted to conforming finite element approximations in the two-dimensional case. Non-overlapping domain decomposition methods of substructuring type are studied in [TWW01; DW12], Neumann-Neumann methods in [Tos00a], and FETI and FETI-DP in [RT01] and [TV05], respectively. Besides the work in [DW12], where the authors use non-standard coarse spaces based on energy minimization, all the other aforementioned works reflect a dependence on the coefficients in the asymptotic convergence that predicts deterioration of the preconditioner when the coefficient in the zero-th order part of the operator is dominant.

The three-dimensional  $H(\mathbf{curl}, \Omega)$ -elliptic problem with two variable coefficients, has rarely been investigated. This is related to the significant challenges that emerge in the three-dimensional continuous problem, and to the much intricate construction of the finite element discretizations. FETI-DP algorithms for conforming approximations were introduced in [Tos06] and the analysis reveals explicit dependence on the ratio between the reaction and the  $\mathbf{curl}$  coefficients, as appeared in all first works for the two-dimensional case. Other significant contributions are contained in [HSZ08; HSZ13], where the authors further extended the research from [HZ04; HZ03]. In [HSZ08] a novel mortar method for  $H(\mathbf{curl}, \Omega)$ -conforming finite element discretizations is introduced and analyzed. A weighted (with respect to the reaction coefficient) Helmholtz decomposition is derived in [HSZ13] and used in the analysis of the substructuring preconditioner of [HZ04], applied to the problem with variable coefficients. The assumptions on the distribution of the coefficients seems however quite restrictive, ruling out many cases of interest. More recently in [DW16], the authors have devised a non-overlapping BDDC algorithm able to improve the dependence on the quotient  $H/h$  between the coarse and fine meshes by saving two logarithmic factors. Nevertheless, their analysis still reflects the same dependence on the coefficients as in [Tos06].

## Preconditioners for Discontinuous Galerkin Methods

As shown in Chapter 3, truly discontinuous finite element discretizations provide an attractive alternative to standard and stabilized conforming Galerkin approximations. In MHD simulations the rationale for using DG discretizations is to cope with (locally) dominating transport.

Over the last fifteen years a considerable effort has been devoted to the development of efficient and robust preconditioning techniques for DG discretizations. Most analysis and especially convergence results have dealt with DG approximations of simple (mostly second order) elliptic problems. The first efforts were focused on the development and analysis of classical domain decomposition methods: overlapping Schwarz methods were studied in [FK01; FK05; BW05] for Interior Penalty (IP) DG approximations of second and fourth order problems, whilst simple Schwarz methods with no overlap were introduced and proved to be convergent (unlike to the conforming case) in [FK01; AA07; AA08] for all the DG methods considered in [Arn+02]. The analysis in the aforementioned works uses an augmented version of classical

Schwarz theory in order to deal with the nonconformity of the finite element spaces. Simultaneously, first attempts to design and analyze efficient multigrid solvers in [GK03; BZ05] followed the classical multigrid theories of [BPX91] and [BH83; Bre02], respectively. Nowadays, there is still active research in these directions trying to harness classical theories. In particular, Schwarz preconditioners [AH11; Bar+11; BS06; Ant+12; CKW13] and multigrid methods [BO07; Bre+11; KT08; Coc+14; ASV15] have been investigated for newly introduced DG discretizations and for  $hp$ -DG approximations of elliptic problems. Furthermore, more sophisticated non-overlapping domain decomposition preconditioners of substructuring type have been recently studied for DG discretizations of second order elliptic problems in two dimensions. In [DGS07; DGS12; DGS13; Ant+15], non-overlapping BDDC, N-N, FETI-DP and substructuring methods have been introduced and analyzed for a Nitsche-type approximation. BDDC preconditioners are also studied in [BPS13] for a weakly penalized IP method; in [SL13] for the  $p$ -version of a hybrid DG method; and in [CPP14] for  $hp$ -IP-DG spectral methods. While different approaches have been considered in the analysis, all the foregoing works provide quasi-optimality results (with respect to the mesh size and in [CPP14; Ant+15] also with respect to the polynomial degree) and robustness of the preconditioners with respect to possible high variations or jumps in the diffusion coefficient.

The evolution of domain decomposition and multigrid preconditioners has been paralleled by the design and analysis of other subspace correction methods (for two and three-dimensional problems), erected on the construction of suitable splittings of DG finite element spaces. At least two main approaches, based on different principles, have been pursued: the use of a suitable subspace, and the construction of orthogonal splittings. Optimal multilevel preconditioners, based on an orthogonal space decomposition of the DG space, were introduced in [AZ09] for symmetric and non-symmetric piecewise linear IP approximations of elliptic problems. This technique has been adapted and extended to deal with a larger family of problems including elliptic problems with jumping coefficients [Ayu+14b], linear elasticity [Ayu+13a] and convection dominated problems corresponding to drift-diffusion models for the transport of species [Ayu+13b].

A different direction was followed in [Dob+06] and [BCD08; Bri+09], where the authors introduced two-level and multilevel preconditioners, respectively, for the Interior Penalty (IP) DG methods. The aforementioned works are related to the construction of the solvers presented in this chapter. Indeed, the conceptual foundation of the two approaches (although not explicitly mentioned in the referred works on multilevel preconditioners) is the fictitious space lemma and the Auxiliary Space Method (ASM). The gist of the ASM is to introduce an auxiliary finite element space in which preconditioning techniques are available: spaces pf piecewise constant functions in the first case, and conforming linear finite elements in the present derivation. The ASM has been further exploited recently in [Ayu+14a] to construct optimal preconditioners for a family of  $H(\text{div}, \Omega)$ -DG discretizations of the Stokes problem, and in [Bri+15] to develop optimal multilevel preconditioners for spectral DG discretizations (see also [CPP14] where these results are used for designing a BDDC preconditioner). In particular, the analysis in [Ayu+14a] requires a suitable extension of the Fictitious Space Lemma.

There is also a relatively big body of work on DG discretizations for boundary values problems like (5.1). Different varieties of DG for different extended and regularized versions of (5.1) have been presented in, among others, [BHP07; BP06; Hou+05; PS03]. By and large it seems that numerical analysis has entirely focused on a priori and a posteriori error estimate and no attention has been paid to the design and analysis of preconditioners.

For the linear systems of equations arising from a piecewise polynomial symmetric Interior Penalty Discontinuous Galerkin (IP-DG) discretization of  $H(\mathbf{curl}, \Omega)$ -elliptic boundary value problems (5.1), presented in Section 5.1, we design a family of preconditioners based on the *auxiliary space method* (ASM), explained in Section 5.2. Since the construction and analysis of the proposed solvers relies on auxiliary  $H(\mathbf{curl}, \Omega)$ -conforming finite element spaces,

we take the availability of a (direct) solver for any standard  $H(\mathbf{curl}, \Omega)$ -conforming Galerkin discretization of (5.2) for granted.

In Section 5.3, we explore the asymptotic optimality of the proposed preconditioners with respect to the mesh width and the problem coefficients. Extensive numerical experiments in two dimensions are included in Section 5.4 to verify the theory and assess the performance of the preconditioners.

The results contained in this chapter have been published in [AHP16].

## 5.1. Symmetric Interior Penalty Discontinuous Galerkin Discretization

Let  $\mathcal{T}_h$  be a cellular partition of  $\Omega$  as in Assumption 2.4.2. We are not interested in the capability of DG to accommodate rather general meshes: we acknowledge an extension of our approach to more general meshes is an open problem. As described in Section 3.2.1, if  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$  is an interior  $(n-1)$ -face shared by two elements  $T^+, T^- \in \mathcal{T}_h$ , we denote by  $\mathbf{v}^\pm$  the traces on  $f$  of a piecewise smooth vector-valued function  $\mathbf{v}$  taken from within  $T^\pm$ . Let  $\mathbf{n}_f^+$  and  $\mathbf{n}_f^-$  denote the unit normal vectors on  $f$  pointing outwards from  $T^+$  and  $T^-$ , respectively. The average across  $f \in \Delta_{n-1}(\mathcal{T}_h)$  is defined as in (3.6) while, throughout this chapter, the jump  $[\mathbf{v}]_{f,\tau}$  represents the jump of the trace (the subscript  $\tau$  hints at the tangential component of  $\mathbf{v}$  on  $f$ ):

$$\{\mathbf{v}\}_f := \frac{\mathbf{v}^+ + \mathbf{v}^-}{2}, \quad [\mathbf{v}]_{f,\tau} := [\operatorname{tr} \mathbf{v}]_f = \mathbf{n}_f^+ \times \mathbf{v}^+ + \mathbf{n}_f^- \times \mathbf{v}^-.$$

On a boundary face  $f \in \Delta_{n-1}^\partial(\mathcal{T}_h)$ , we set  $\{\mathbf{v}\}_f := \mathbf{v}$  and  $[\mathbf{v}]_{f,\tau} := \mathbf{n}_f \times \mathbf{v}$ .

Let the finite element space  $\Lambda_h^{d,1}(\mathcal{T}_h)$  be one of the discontinuous polynomial spaces  $\mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ ,  $\mathcal{P}_r^{d,-} \Lambda^1(\mathcal{T}_h)$  or  $\mathcal{Q}_r^{d,-} \Lambda^1(\mathcal{T}_h)$  with  $r \geq 1$  as in (2.31) and (2.32), namely

$$\Lambda_h^{d,1}(\mathcal{T}_h) := \{\mathbf{v} \in L^2(\Omega)^n, \mathbf{v}|_T \in \zeta_r(T), T \in \mathcal{T}_h\}, \quad (5.3)$$

with local shape functions  $\zeta_r(T) = \mathcal{P}_r \Lambda^1(T)$ ,  $\mathcal{P}_r^- \Lambda^1(T)$  or  $\mathcal{Q}_r^- \Lambda^1(T)$ , respectively. The corresponding  $H_0(\operatorname{curl}, \Omega)$ -conforming finite element spaces are as in (2.28) and (2.29),

$$\Lambda_h^1(\mathcal{T}_h) := \Lambda_h^{d,1}(\mathcal{T}_h) \cap H_0(\operatorname{curl}, \Omega). \quad (5.4)$$

Henceforth, if not otherwise specified, we assume that functions in the aforementioned finite element spaces comply with the boundary conditions in (5.1).

Every  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  has a local representation (2.33) in the basis of  $\zeta_r(T)$ :

$$\mathbf{v}|_T(\mathbf{x}) = \sum_{j=1}^{\min\{n,r\}} \sum_{f_j \in \Delta_j^\circ(T)} \sum_{\ell=1}^{N_j} V_{f_j,T}^\ell \varphi_{j,T}^\ell(\mathbf{x}), \quad \forall \mathbf{x} \in T, \forall T \in \mathcal{T}_h,$$

where  $\{\varphi_{j,T}^\ell, 1 \leq j \leq \min\{n,r\}, 1 \leq \ell \leq N_j\}$  is a basis of  $\zeta_r(T)$ . We assume that the polynomial degree  $r$  satisfies  $r \geq n$ , with  $r \geq 1$ . If  $r < n$ , the forthcoming discussion proceeds verbatim by dropping the superfluous terms. We adopt the notation:  $e \in \Delta_{n-2}(\mathcal{T}_h)$  and  $\mathbf{v}_{e,T} = \{V_{e,T}^i\}_{i=1}^{N_e}$ ,  $f \in \Delta_{n-1}(\mathcal{T}_h)$  and  $\mathbf{v}_{f,T} = \{V_{f,T}^i\}_{i=1}^{N_f}$ ,  $\mathbf{v}_T = \{V_T^i\}_{i=1}^{N_b}$  so that,

$$\mathbf{v}|_T(\mathbf{x}) = \sum_{e \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} V_{e,T}^i \varphi_{e,T}^i(\mathbf{x}) + \sum_{f \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} V_{f,T}^i \varphi_{f,T}^i(\mathbf{x}) + \sum_{i=1}^{N_b} V_T^i \varphi_T^i(\mathbf{x}), \quad \forall \mathbf{x} \in T, \forall T \in \mathcal{T}_h. \quad (5.5)$$

### Symmetric Interior Penalty Method

We propose a discretization of the problem (5.1) along the lines of [Hou+05], based on the symmetric Interior Penalty discontinuous Galerkin (SIP-DG) method, as pioneered in [Arn82; Whe78; Bak77] for scalar second order problems (see also [Arn+02, Section 3.4]). To deal with the discontinuous coefficients of the problem and provide a robust approximation method, we modify the classical SIP-DG method, similarly to [Dry03, Section 4], where weighted averages of the discontinuous coefficients on the mesh

skeleton are introduced. Moreover, the choice of penalizing the tangential jump of the solution across mesh interfaces, weighted by stabilization coefficients which depend on the distribution of the coefficient  $\nu$  over  $\mathcal{T}_h$ , is aimed at ensuring the robustness of the approximation and of the preconditioners with respect to variations of the problem coefficients.

More precisely, we consider the discrete variational formulation: Find  $\mathbf{v}_h \in \Lambda_h^{d,1}(\mathcal{T}_h)$  such that

$$a_{\text{DG}}(\mathbf{v}_h, \mathbf{z}_h) = (\mathbf{f}, \mathbf{z}_h)_{\mathcal{T}_h} \quad \forall \mathbf{z}_h \in \Lambda_h^{d,1}(\mathcal{T}_h), \quad (5.6)$$

where

$$(\mathbf{f}, \mathbf{z}_h)_{\mathcal{T}_h} := \sum_{T \in \mathcal{T}_h} (\mathbf{f}, \mathbf{z}_h)_{L^2(T)}, \quad \forall \mathbf{z}_h \in \Lambda_h^{d,1}(\mathcal{T}_h),$$

and the bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  is defined, for all  $\mathbf{v}, \mathbf{z} \in \Lambda_h^{d,1}(\mathcal{T}_h)$ , as

$$\begin{aligned} a_{\text{DG}}(\mathbf{v}, \mathbf{z}) := & \sum_{T \in \mathcal{T}_h} (\nu_T \mathbf{curl}_{\mathbf{h}} \mathbf{v}, \mathbf{curl}_{\mathbf{h}} \mathbf{z})_{L^2(T)} + \sum_{T \in \mathcal{T}_h} (\beta_T \mathbf{v}, \mathbf{z})_{L^2(T)} - \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} (\{\nu \mathbf{curl}_{\mathbf{h}} \mathbf{v}\}_{f,\gamma}, [\mathbf{z}]_{f,\tau})_{L^2(f)} \\ & - \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} ([\mathbf{v}]_{f,\tau}, \{\nu \mathbf{curl}_{\mathbf{h}} \mathbf{z}\}_{f,\gamma})_{L^2(f)} + \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} (s_f [\mathbf{v}]_{f,\tau}, [\mathbf{z}]_{f,\tau})_{L^2(f)}. \end{aligned} \quad (5.7)$$

Here  $\mathbf{curl}_{\mathbf{h}}$  denotes the broken operator;  $\nu_T$  and  $\beta_T$  are the restriction of the coefficients  $\nu$  and  $\beta$  to the element  $T \in \mathcal{T}_h$ , respectively. We assume that the partition  $\mathcal{T}_h$  of  $\Omega$  resolves the coefficients, i.e.,  $\nu, \beta \in \mathcal{P}_0(\mathcal{T}_h)$ . Moreover, in (5.7), the function  $s_f$  penalizes the tangential jumps over the  $(n-1)$ -cells of the skeleton of the partition. In particular,  $s_f$  is defined as in [Arn82; GSS08; PS03],

$$s_f := c_0 h_f^{-1}, \quad \forall f \in \Delta_{n-1}(\mathcal{T}_h), \quad (5.8)$$

where  $c_0$  is a strictly positive constant independent of the mesh size and the coefficients of the problem, and depending only on the shape regularity constant of  $\mathcal{T}_h$ . Every interface jump is weighted with a coefficient function  $\{\alpha_T^\nu\}_{T \in \mathcal{T}_h} \in \mathcal{P}_0(\mathcal{T}_h)$  defined elementwise as

$$\alpha_T^\nu := \begin{cases} \max_{f \in \Delta_{n-1}(T)} \max_{e \in \Delta_n(e)} \nu_T & \forall T \in \Delta_n^\circ(\mathcal{T}_h), \\ \nu_T & \forall T \in \Delta_n^\partial(\mathcal{T}_h). \end{cases} \quad (5.9)$$

Observe that, in view of the above definition, the coefficient  $\alpha_T^\nu$  is taking the maximum coefficient  $\nu$  over a patch of elements surrounding  $T$ . Moreover, in (5.7), the weighted average  $\{\cdot\}_{f,\gamma}$  is defined as the plain trace for a boundary face, whereas for  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$ ,

$$\{\mathbf{v}\}_{f,\gamma} := \gamma_f^+ \mathbf{v}^+ + \gamma_f^- \mathbf{v}^-, \quad \text{with} \quad \gamma_f^- = 1 - \gamma_f^+,$$

for weights  $\gamma_f^\pm$  that depend on the coefficient  $\nu$  and might vary over all interior faces. More precisely, for any  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h) \cap \Delta_{n-1}(T^+) \cap \Delta_{n-1}(T^-)$ , we take  $\gamma_f^\pm$  as follows:

$$\gamma_f^\pm = \frac{\nu^\mp}{\nu^+ + \nu^-}, \quad \text{where} \quad \nu^\pm := \nu|_{T^\pm}.$$

With this choice of  $\{\gamma_f\}_{f \in \Delta_{n-1}^\circ(\mathcal{T}_h)}$ , the weighted average  $\{\cdot\}_{f,\gamma}$  can be expressed in terms of the standard average as

$$\{\nu \mathbf{curl}_{\mathbf{h}} \mathbf{v}\}_{f,\gamma} = \{\nu\}_{f,H} \{\mathbf{curl}_{\mathbf{h}} \mathbf{v}\}_f, \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h), \forall f \in \Delta_{n-1}(\mathcal{T}_h), \quad (5.10)$$

and  $\{\nu\}_{f,H}$  denotes the harmonic average of the coefficient  $\nu$  across a face  $f \in \Delta_{n-1}(\mathcal{T}_h)$ , namely

$$\{\nu\}_{f,H} := \begin{cases} \frac{2\nu^+ \nu^-}{\nu^+ + \nu^-} & \forall f \in \Delta_{n-1}^\circ(\mathcal{T}_h) \cap \Delta_{n-1}(T^+) \cap \Delta_{n-1}(T^-), \\ \nu_T & \forall f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \cap \Delta_{n-1}(T). \end{cases} \quad (5.11)$$

Notice that  $\min\{\nu^+, \nu^-\} \leq \{\nu\}_{f,H} \leq 2 \min\{\nu^+, \nu^-\}$  and hence, in particular,  $\{\nu\}_{f,H} \leq 2\nu^\pm$ . Thus, the harmonic average of  $\nu$  can also be bounded by the stabilization coefficients (5.9),

$$\{\nu\}_{f,H} \leq \alpha_T^\nu, \quad \forall f \in \Delta_{n-1}(T), \forall T \in \mathcal{T}_h. \quad (5.12)$$

When the variational formulation (5.6) is restricted to  $\Lambda_h^1(\mathcal{T}_h)$ , the corresponding  $H_0(\mathbf{curl}, \Omega)$ -conforming discretization of (5.1) is obtained. On  $\Lambda_h^{d,1}(\mathcal{T}_h)$  we introduce the seminorms

$$\begin{aligned} \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 &:= \sum_{T \in \mathcal{T}_h} \nu_T \|\mathbf{curl}_h \mathbf{v}\|_{L^2(T)}^2 & \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h), \\ |\mathbf{v}|_{*,\nu}^2 &:= \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 & \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h), \end{aligned}$$

and norms

$$\|\mathbf{v}\|_{0,\beta,\mathcal{T}_h}^2 := \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{v}\|_{L^2(T)}^2 \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h), \quad (5.13)$$

$$\|\mathbf{v}\|_{\text{DG}}^2 := \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 + \|\mathbf{v}\|_{0,\beta,\mathcal{T}_h}^2 + |\mathbf{v}|_{*,\nu}^2 \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h). \quad (5.14)$$

Since the coefficients  $\nu$  and  $\beta$  are assumed to be  $\mathcal{T}_h$ -piecewise constant,  $\nu, \beta \in \mathcal{P}_0(\mathcal{T}_h)$ , we can distinguish two regimes:

$$\Upsilon_h := \{T \in \mathcal{T}_h : h_T^2 \beta_T < \alpha_T^\nu\}, \quad \Upsilon_h^\Gamma := \{T \in \mathcal{T}_h : h_T^2 \beta_T \geq \alpha_T^\nu\}. \quad (5.15)$$

Elements in  $\Upsilon_h$  are in the so-called *curl-dominated regime*, whereas  $\Upsilon_h^\Gamma$  is the region with relative large  $\beta$  and will be referred to as *reaction-dominated regime* thereof.

### Convergence of the Approximation

The bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  defined in (5.7) is continuous and coercive in  $\Lambda_h^{d,1}(\mathcal{T}_h)$  with respect to the  $\|\cdot\|_{\text{DG}}$  norm (5.14), with constants independent of the mesh size  $h$  and the coefficients of the problem.

**Proposition 5.1.1.** *Let the bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  be defined as in (5.7). Then, there exists a constant  $C_{\text{cont}} > 0$  depending on the shape regularity of the mesh and on the polynomial degree of  $\Lambda_h^{d,1}(\mathcal{T}_h)$  such that*

$$|a_{\text{DG}}(\mathbf{v}, \mathbf{z})| \leq C_{\text{cont}} \|\mathbf{v}\|_{\text{DG}} \|\mathbf{z}\|_{\text{DG}}, \quad \forall \mathbf{v}, \mathbf{z} \in \Lambda_h^{d,1}(\mathcal{T}_h). \quad (5.16)$$

Moreover, there exists  $0 < c \in \mathbb{R}$  independent of the mesh width and the problem coefficients such that for  $c_0 > c$  in (5.8),

$$a_{\text{DG}}(\mathbf{v}, \mathbf{v}) \geq C_{\text{stab}} \|\mathbf{v}\|_{\text{DG}}^2, \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h). \quad (5.17)$$

The constant  $C_{\text{stab}} > 0$  depends only on the shape regularity of the mesh and on the polynomial degree.

*Proof.* We first show the coercivity (5.17). Taking  $\mathbf{z} = \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  in (5.7), results in

$$a_{\text{DG}}(\mathbf{v}, \mathbf{v}) \geq \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 + \|\mathbf{v}\|_{0,\beta,\mathcal{T}_h}^2 + c_0 |\mathbf{v}|_{*,\nu}^2 - 2 \left| \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \int_f \{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma} \cdot [\mathbf{v}]_{f,\tau} \, ds \right|. \quad (5.18)$$

The Cauchy-Schwarz inequality and the arithmetic-geometric inequality together with the relation (5.10) and the bound (5.12) on the harmonic average, give, for  $\epsilon > 0$ ,

$$\begin{aligned} \left| \int_f \{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma} \cdot [\mathbf{v}]_{f,\tau} \, ds \right| &\leq C \left( \frac{h_f}{\{\nu\}_{f,H}} \|\{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma}\|_{L^2(f)}^2 \right)^{1/2} \left( \{\nu\}_{f,H} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \right)^{1/2} \\ &\stackrel{(5.10)}{\leq} C\epsilon \{\nu\}_{f,H} h_f \|\{\mathbf{curl}_h \mathbf{v}\}_f\|_{L^2(f)}^2 + \frac{C}{4\epsilon} \{\nu\}_{f,H} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned} \quad (5.19)$$

Now, let  $f \in \Delta_{n-1}(T^+) \cap \Delta_{n-1}(T^-)$ . The trace inequality [Agr65, Theorem 3.10] and the inverse inequality [Cia91, Theorem 3.2.6] together with the fact that  $\{\nu\}_{f,H} \leq 2\nu_{T^\pm}$  yield

$$\{\nu\}_{f,H} h_f \|\{\mathbf{curl}_h \mathbf{v}\}_f\|_{L^2(f)}^2 \leq C(\nu_{T^+} \|\mathbf{curl}_h \mathbf{v}\|_{0,T^+}^2 + \nu_{T^-} \|\mathbf{curl}_h \mathbf{v}\|_{0,T^-}^2),$$

where  $C$  depends on the shape regularity and the local polynomial space. Moreover, by (5.12) and the definition (5.9), it holds

$$\begin{aligned} \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \{\nu\}_{f,H} \|\cdot\|_{L^2(f)} &\leq C \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{f \in \Delta_{n-1}(T)} \|\cdot\|_{L^2(f)} \\ &\leq C \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{f \in \Delta_{n-1}(T)} \|\cdot\|_{L^2(f)} + \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \Delta_{n-1}(T)} \|\cdot\|_{L^2(f)}. \end{aligned}$$

Hence, summing (5.19) over all  $(n-1)$ -faces yields

$$\begin{aligned} \left| \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \int_f \{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma} \cdot [\mathbf{v}]_{f,\tau} \, ds \right| &\leq C\epsilon \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 \\ &\quad + \frac{C}{4\epsilon} \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned}$$

Plugging the estimate into (5.18), results in

$$a_{\text{DG}}(\mathbf{v}, \mathbf{v}) \geq (1 - 2\epsilon C) \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 + \|\mathbf{v}\|_{0,\beta,\mathcal{T}_h}^2 + \left( c_0 - \frac{C}{2\epsilon} \right) |\mathbf{v}|_{*,\nu}^2.$$

By taking  $\epsilon$  sufficiently small and the constant penalty parameter  $c_0$  sufficiently large, coercivity (5.17) is achieved with a constant  $C_{\text{stab}}$  depending only on the shape regularity of  $\mathcal{T}_h$  and the polynomial degree of  $\Lambda_h^{\text{d},1}(\mathcal{T}_h)$ .

Concerning the continuity (5.16), the Cauchy-Schwarz inequality gives

$$\begin{aligned} \left| \sum_{T \in \mathcal{T}_h} \int_T \nu_T \mathbf{curl}_h \mathbf{v} \cdot \mathbf{curl}_h \mathbf{z} + \beta_T \mathbf{v} \cdot \mathbf{z} \right| &\leq \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h} \|\mathbf{curl}_h \mathbf{z}\|_{0,\nu,\mathcal{T}_h} + \|\mathbf{v}\|_{0,\beta,\mathcal{T}_h} \|\mathbf{z}\|_{0,\beta,\mathcal{T}_h}, \\ \left| \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \int_f c_0 h_f^{-1} [\mathbf{v}]_{f,\tau} \cdot [\mathbf{z}]_{f,\tau} \right| &\leq C |\mathbf{v}|_{*,\nu} |\mathbf{z}|_{*,\nu} \leq C \|\mathbf{v}\|_{\text{DG}} \|\mathbf{z}\|_{\text{DG}}. \end{aligned}$$

A similar reasoning as in (5.19) (but without using the arithmetic-geometric inequality) gives

$$\left| \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} \int_f \{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma} \cdot [\mathbf{z}]_{f,\tau} \, ds \right| \leq C \|\mathbf{curl}_h \mathbf{v}\|_{0,\nu,\mathcal{T}_h} |\mathbf{z}|_{*,\nu},$$

and continuity (5.16) follows from the definition (5.14) of the  $\|\cdot\|_{\text{DG}}$  norm.  $\square$

The a priori error analysis of the discretization introduced in (5.6) can be conducted similarly to [Hou+05; GSS07], [GSS08, Section 5] where the time-dependent Maxwell equations and the time-harmonic Maxwell problem in the high frequency regime are considered. In particular, in  $\mathbf{V}(h) := \Lambda_h^{\text{d},1}(\mathcal{T}_h) + \mathbf{H}_0(\mathbf{curl}, \Omega)$  the discrete bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  is extended to an auxiliary *augmented* bilinear form well-defined on  $\mathbf{V}(h)$ ,

$$\begin{aligned} \tilde{a}_{\text{DG}}(\mathbf{v}, \mathbf{z}) &= \sum_{T \in \mathcal{T}_h} (\nu_T \mathbf{curl}_h \mathbf{v}, \mathbf{curl}_h \mathbf{z})_{L^2(T)} + \sum_{T \in \mathcal{T}_h} (\beta_T \mathbf{v}, \mathbf{z})_{L^2(T)} - \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} (\{\nu \Pi_h(\mathbf{curl}_h \mathbf{v})\}_{f,\gamma}, [\mathbf{z}]_{f,\tau})_{L^2(f)} \\ &\quad - \sum_{f \in \Delta_{n-1}(\mathcal{T}_h)} ([\mathbf{v}]_{f,\tau}, \{\nu \Pi_h(\mathbf{curl}_h \mathbf{z})\}_{f,\gamma})_{L^2(f)} + \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} (s_f [\mathbf{v}]_{f,\tau}, [\mathbf{z}]_{f,\tau})_{L^2(f)}, \end{aligned}$$

where  $\Pi_h$  is the  $L^2$ -orthogonal projection onto  $\Lambda_h^{\text{d},1}(\mathcal{T}_h)$ . Observing that  $\tilde{a}_{\text{DG}}(\mathbf{v}, \mathbf{z}) = a_{\text{DG}}(\mathbf{v}, \mathbf{z})$  for all  $\mathbf{z} \in \Lambda_h^{\text{d},1}(\mathcal{T}_h)$ , coercivity of the augmented form in  $\Lambda_h^{\text{d},1}(\mathcal{T}_h)$  follows from the coercivity of  $a_{\text{DG}}(\cdot, \cdot)$ .

Continuity in  $\mathbf{V}(h)$  is shown arguing as for  $a_{\text{DG}}(\cdot, \cdot)$ , using in addition the stability of the  $L^2$ -projection. By estimating the consistency error, quasi-optimal error estimates in the DG-norm can be shown, provided the solution of (5.2) is regular enough. More precisely, let the problem parameters  $\nu$  and  $\beta$  be smooth functions, positive and uniformly bounded in  $\Omega$ . Let  $\mathbf{v}_h \in \Lambda_h^{\text{d},1}(\mathcal{T}_h) := \mathcal{P}_r^{\text{d}}\Lambda^1(\mathcal{T}_h)$  be the discrete solution of (5.6) and let  $\mathbf{v} \in H^s(\Omega)^n$  be the exact solution of (5.2), with  $\mathbf{curl}\mathbf{v} \in H^s(\Omega)^n$  for  $s > 1/2$ . Then,

$$\|\mathbf{v} - \mathbf{v}_h\|_{\text{DG}} \leq Ch^{\min(r,s)} (\|\mathbf{v}\|_{H^s(\Omega)} + \|\mathbf{curl}\mathbf{v}\|_{H^s(\Omega)}),$$

where the constant  $C > 0$  is independent of the mesh width but depends on the shape regularity of the mesh, the polynomial degree  $r$  and the problem coefficients  $\nu$  and  $\beta$ . We will not delve into further details and the reason is twofold: the a priori error analysis of the DG approximation (5.6) is out of the scope of the present analysis, and, on the other hand, since the proof of the a priori error estimates in [Hou+05] hinges on duality techniques, the extension to non-smooth coefficients does not carry over straightforwardly.

A naive application of an iterative solver (e.g. the conjugate gradient method) for the solution of the linear system ensuing from the discretization (5.6) would be undermined by the dimension of the system and by the problem coefficients. Indeed the spectral condition number of the Galerkin matrix associated with the discrete bilinear form (5.7) is proportional to a factor  $h^{-2}\mathcal{J}(\nu, \beta, \alpha_T^\nu)$ , where

$$\mathcal{J}(\nu, \beta, \alpha_T^\nu) := \frac{\max_T \alpha_T^\nu}{\min_T \nu_T} + h^2 \frac{\max_T \beta_T}{\min_T \beta_T}.$$

Hence, designing a preconditioner able to tame the combined effect of the mesh width and of highly varying coefficients is of crucial importance.

## 5.2. Auxiliary Space Preconditioning

In this section we present the key ideas of preconditioning approaches based on fictitious or auxiliary spaces in a general abstract framework. The design and analysis of the proposed preconditioning technique relies on the theory of the *fictitious space method*. For this reason, following the guidelines given in [HX07], we provide the main steps required to apply this theory to our particular problem. Then we introduce a family of preconditioners for the IP-DG discretization of problem (5.1) presented in Section 5.1.

### 5.2.1. Fictitious Space and Auxiliary Space Method

The auxiliary space method was introduced as a technique to develop and analyze optimal (in terms of independence on the dimension of the system) multilevel preconditioners for elliptic discretizations on general unstructured meshes in [Xu96] and for nonconforming methods in [Osw96]. It can be interpreted as a further generalization of the fictitious space approach, based on the so-called *fictitious space lemma* originally introduced by Nepomnyaschikh in [Nep91].

Let us use the symbols ' and \* to label dual spaces and adjoint operators, respectively. Let  $V$  be a real Hilbert space. Two main ingredients are required in the construction of a fictitious space preconditioner for the operator  $A : V \rightarrow V'$  associated with the inner product  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  and norm  $\|\cdot\|_A := a(\cdot, \cdot)^{1/2}$ :

- (1) Another real Hilbert space  $U$ , the *fictitious space*, endowed with the inner product  $b(\cdot, \cdot) : U \times U \rightarrow \mathbb{R}$ , associated operator  $B : U \rightarrow U'$  and induced norm  $\|\cdot\|_B := b(\cdot, \cdot)^{1/2}$ .
- (2) A continuous, linear and surjective transfer operator  $\Pi : U \rightarrow V$  (the so-called prolongation operator in domain decomposition and multigrid methods).

Then, the fictitious space preconditioner  $M : V' \rightarrow V$  is defined as

$$M := \Pi \circ B^{-1} \circ \Pi^*. \quad (5.20)$$

The adjoint operator  $\Pi^* : V' \rightarrow U'$  is defined through the duality pairings  $\langle \varphi, \Pi u \rangle_{V' \times V} = \langle \Pi^* \varphi, u \rangle_{U' \times U}$  for all  $\varphi \in V'$ ,  $u \in U$ . Naturally, the convergence properties of  $M$  depend on the auxiliary space  $U$  and on  $B$ . The fact that  $\Pi$  is surjective, ensures that  $M$  is an isomorphism, and in particular a valid preconditioner, [HX07, Lemma 2.1]. The analysis of the fictitious space preconditioner is grounded on the fictitious space lemma, which we recall next without proof (see e.g. [HX07, Theorem 2.2]).

**Theorem 5.2.1.** [Nep91, Lemma 2.2]. *Let  $V$  and  $U$  be Hilbert spaces endowed with inner product  $a(\cdot, \cdot)$  and  $b(\cdot, \cdot)$  and associated isomorphisms  $A : V \rightarrow V'$  and  $B : U \rightarrow U'$ , respectively. Let  $\Pi : U \rightarrow V$  be a linear operator and  $M := \Pi \circ B^{-1} \circ \Pi^* : V' \rightarrow V$ . Assume that*

- (i)  $\Pi$  is bounded, i.e.  $\exists C_1 > 0 : \|\Pi u\|_A^2 \leq C_1 \|u\|_B^2 \quad \forall u \in U$ ;
- (ii)  $\Pi$  is surjective and  $\exists C_0 > 0 : \forall v \in V \quad \exists u \in U : v = \Pi u \quad \text{and} \quad \|u\|_B^2 \leq C_0 \|v\|_A^2$ .

Then,

$$C_0^{-1} \|v\|_A^2 \leq a(M \circ Av, v) \leq C_1 \|v\|_A^2 \quad \forall v \in V.$$

In the present work, the space  $V$  is the finite dimensional space  $\Lambda_h^{d,1}(\mathcal{T}_h)$  in (5.3) and it is endowed with the inner product  $a_{DG}(\cdot, \cdot)$  given in (5.7). The coercivity and symmetry of the bilinear form  $a_{DG}(\cdot, \cdot)$  shown in Proposition 5.1.1, ensure that the associated operator  $A_{DG} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow (\Lambda_h^{d,1}(\mathcal{T}_h))'$  is self-adjoint and positive definite. Its algebraic representation yields a symmetric positive definite matrix  $\mathbb{A}$ . The fictitious space Theorem 5.2.1 provides an estimate of the spectral condition number of the preconditioned system matrix, namely  $\kappa(\mathbb{M}\mathbb{A}) \leq C_0 C_1$ , where  $\mathbb{M}$  denotes the matrix representation of the preconditioner (5.20). In addition, the bilinear forms  $b(\cdot, \cdot)$  on the fictitious space  $U$  can be replaced with spectrally equivalent bilinear forms: under the assumptions of Theorem 5.2.1, if a preconditioner  $N : U' \rightarrow U$  for  $B$  in the space  $U$  is available, the fictitious space preconditioner can be constructed as  $M = \Pi \circ N \circ \Pi^*$ . Furthermore, the spectral condition number of the preconditioned system in  $V$  can be gauged in terms of the spectral condition number of the preconditioned system in  $U$ , namely  $\kappa(\mathbb{M}\mathbb{A}) \leq \kappa(\mathbb{N}\mathbb{B}) C_0 C_1$ .

In the auxiliary space method developed in [Xu96], the fictitious space is chosen as a product space having  $V$  as one of its components. Such choice eases the construction of a *surjective* map  $\Pi$ . The fictitious space is then given by the Cartesian product  $U = V \times W_1 \times \dots \times W_m$ ,  $m \in \mathbb{N}$ , where  $\{W_i\}_{i=1}^m$  are Hilbert spaces representing the “true” auxiliary spaces. We focus on the simplest case of nonnested two level preconditioning based on the fictitious space  $U = V \times W$ , where  $W$  is endowed with a symmetric positive definite bilinear form  $a_w(\cdot, \cdot)$  and associated operator  $A_w : W \rightarrow W'$ . As component of the fictitious space, the space  $V$  is equipped with another (in the sense of different from  $a(\cdot, \cdot)$ ) inner product  $s(\cdot, \cdot)$  with associated mapping  $S : V \rightarrow V'$ . The operator  $S$  is typically referred to as *smoother*. This approach can be thought of as the fictitious space technique with the inner product

$$b(u, u) = s(v, v) + a_w(w, w), \quad \forall u = (v, w) \in V \times W, \quad (5.21)$$

and the auxiliary space preconditioner operator is given by

$$B = S^{-1} + \Pi_w \circ A_w^{-1} \circ \Pi_w^*, \quad (5.22)$$

where the linear transfer operator  $\Pi_w : W \rightarrow V$  yields the surjective map

$$\Pi := \begin{pmatrix} \text{Id} \\ \Pi_w \end{pmatrix} : U \rightarrow V.$$

Here, the adjoint operator  $\Pi_w^* : V' \rightarrow W'$  is defined by  $\langle \varphi, \Pi_w w \rangle_{V' \times V} = \langle \Pi_w^* \varphi, w \rangle_{W' \times W}$  for all  $\varphi \in V'$ ,  $w \in W$ . If  $\mathbb{S} \in \mathbb{R}^{N_v \times N_v}$  with  $N_v := \dim V$  and  $\mathbb{A}_w \in \mathbb{R}^{N_w \times N_w}$ ,  $N_w := \dim W$ , then the preconditioner (5.22) in algebraic form reads

$$\mathbb{B} = \mathbb{S}^{-1} + \mathbb{P} \mathbb{A}_w^{-1} \mathbb{P}^\top, \quad (5.23)$$

where  $\mathbb{P} \in \mathbb{R}^{N_v \times N_w}$  is the matrix representation of the transfer operator  $\Pi_w$ .

The analysis of the auxiliary space preconditioner hinges on the fictitious space Theorem 5.2.1. Its assumptions boil down to fulfilling the conditions of the following theorem (see [HX07, Section 2] and [HWZ06, Lemma 2.1]).

**Theorem 5.2.2.** *Let  $V$  and  $W$  be real Hilbert spaces endowed with symmetric positive definite bilinear forms  $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  and  $a_w(\cdot, \cdot) : W \times W \rightarrow \mathbb{R}$ , respectively. Let  $s(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$  be an inner product on  $V$  with associated operator  $S : V \rightarrow V'$ . Assume that the following conditions are satisfied.*

**Property (F0):** *The transfer operator  $\Pi_w : W \rightarrow V$  is uniformly bounded, i.e.  $\exists c_w > 0$  independent of the mesh width and the parameters of the problem, and depending on the mesh only through its shape regularity constant such that*

$$a(\Pi_w w, \Pi_w w) \leq c_w a_w(w, w), \quad \forall w \in W. \quad (5.24)$$

**Property (F1):** *The operator  $S^{-1} : V' \rightarrow V$  is continuous, namely there exists  $c_s > 0$ , independent of the mesh width and the parameters of the problem, and depending only on the shape regularity of the mesh such that*

$$a(v, v) \leq c_s s(v, v), \quad \forall v \in V.$$

**Property (F2):** *(Stable decomposition) For every  $u \in V$  there exist  $w \in W$  and  $v \in V$  such that  $u = v + \Pi_w w$ , and there exists  $C_0 > 0$  independent of  $v$ , such that*

$$\inf_{\substack{v \in V, w \in W \\ u=v+\Pi_w w}} \{s(v, v) + a_w(w, w)\} \leq C_0 a(u, u), \quad \forall u = v + \Pi_w w.$$

Then, a direct application of the fictitious space Theorem 5.2.1 yields  $\kappa(\mathbb{B}\mathbb{A}) \leq C_0(c_s + c_w)$ .

### 5.2.2. Auxiliary Space Preconditioners for the IP-DG Discretization in $H_0(\mathbf{curl}, \Omega)$

In constructing an auxiliary space preconditioner for the operator  $A_{\text{DG}} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow (\Lambda_h^{d,1}(\mathcal{T}_h))'$  associated with the bilinear form  $a_{\text{DG}} : \Lambda_h^{d,1}(\mathcal{T}_h) \times \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \mathbb{R}$  in (5.7), the following choices of the auxiliary space  $W$  are adopted:

- (a)  $W$  is the finite element space  $\Lambda_h^1(\mathcal{T}_h) := \Lambda_h^{d,1}(\mathcal{T}_h) \cap H_0(\mathbf{curl}, \Omega)$  endowed with the bilinear form  $a_w(\cdot, \cdot)$  deriving from the  $H_0(\mathbf{curl}, \Omega)$ -conforming finite element approximation of the model problem (5.1):

$$a_w(\mathbf{v}, \mathbf{z}) := \sum_{T \in \mathcal{T}_h} (\nu_T \mathbf{curl} \mathbf{v}, \mathbf{curl} \mathbf{z})_{L^2(T)} + \sum_{T \in \mathcal{T}_h} (\beta_T \mathbf{v}, \mathbf{z})_{L^2(T)}, \quad \forall \mathbf{v}, \mathbf{z} \in \Lambda_h^1(\mathcal{T}_h).$$

Note that  $a_w(\cdot, \cdot)$  is the restriction of the bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  to the  $H_0(\mathbf{curl}, \Omega)$ -conforming finite element space  $\Lambda_h^1(\mathcal{T}_h)$ , namely  $a_w(\mathbf{v}, \mathbf{z}) = a_{\text{DG}}(\mathbf{v}, \mathbf{z})$  for all  $\mathbf{v}, \mathbf{z} \in \Lambda_h^1(\mathcal{T}_h)$ . The associated operator  $A_w : \Lambda_h^1(\mathcal{T}_h) \rightarrow (\Lambda_h^1(\mathcal{T}_h))'$  is self-adjoint and positive definite. The transfer operator  $\Pi_w : \Lambda_h^1(\mathcal{T}_h) \rightarrow \Lambda_h^{d,1}(\mathcal{T}_h)$  is trivially the standard inclusion.

- (b) On a simplicial mesh, if  $\Lambda_h^{d,1}(\mathcal{T}_h) = \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ ,  $r \geq 1$ , we consider a second possible choice for the auxiliary space, namely  $W = \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$ . Note that  $\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h) \subset \mathcal{P}_r \Lambda^1(\mathcal{T}_h) \subset \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ . The space  $\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$  is endowed with the inner product  $a_w(\cdot, \cdot)$  corresponding to the  $H_0(\mathbf{curl}, \Omega)$ -conforming approximation based on  $\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$ . Hence, the corresponding associated operator  $A_w : \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h) \rightarrow (\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h))'$  is self-adjoint and positive definite, and the transfer operator  $\Pi_w : \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h) \rightarrow \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$  is defined as the standard inclusion.

The condition (F0) in Theorem 5.2.2 is trivially satisfied for both the previous choices since the transfer operator  $\Pi_w$  is the standard inclusion. Furthermore, in both cases  $W = \Lambda_h^1(\mathcal{T}_h)$  or  $W = \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$ , inequality (5.24) becomes an identity with constant  $c_w = 1$ , owing to the fact that  $a_w(\cdot, \cdot)$  is the restriction of  $a_{\text{DG}}(\cdot, \cdot)$  to the corresponding  $H(\mathbf{curl}, \Omega)$ -conforming spaces.

### 5.2.3. Smoothers for the Auxiliary Space Preconditioner

To assess the performance of the proposed family of preconditioners, the choices of the auxiliary space described in Section 5.2.2 are combined with different possible smoothing operators  $S : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow (\Lambda_h^{d,1}(\mathcal{T}_h))'$ . We consider relaxation techniques as subspace correction methods [Xu92]: pointwise relaxation (non-overlapping) and patch smoothers (overlapping).

**Pointwise relaxation or non-overlapping subspace correction methods.** Since the Galerkin matrix associated with the DG discretization (5.6) is symmetric positive definite, we focus on Jacobi-type smoothers. Indeed, by virtue of [Zik08, Lemma 3.3], the pointwise symmetric Gauss-Seidel smoother is spectrally equivalent to the corresponding Jacobi smoother, with constants independent of the problem coefficients and mesh size. We introduce non-overlapping additive Schwarz smoothers based on the following splittings of  $\Lambda_h^{d,1}(\mathcal{T}_h)$ :

- (i) Pointwise Jacobi smoother:

$$\Lambda_h^{d,1}(\mathcal{T}_h) = \bigoplus_{T \in \mathcal{T}_h} \left( \sum_{e \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} \text{span}\{\varphi_{e,T}^i\} + \sum_{f \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} \text{span}\{\varphi_{f,T}^i\} + \sum_{i=1}^{N_b} \text{span}\{\varphi_T^i\} \right). \quad (5.25)$$

- (ii) Block Jacobi smoother:

$$\Lambda_h^{d,1}(\mathcal{T}_h) = \bigoplus_{T \in \mathcal{T}_h} \text{span}\{\varphi_{e,T}^1, \dots, \varphi_{e,T}^{N_e}, \varphi_{f,T}^1, \dots, \varphi_{f,T}^{N_f}, \varphi_T^1, \dots, \varphi_T^{N_b}\}. \quad (5.26)$$

In Lemma 5.2.4, we show that the (non-overlapping) block Jacobi smoother associated with (5.26) is spectrally equivalent to the pointwise Jacobi relaxation relative to the splitting (5.25).

**Patch smoothers or overlapping subspace correction methods.** The description and analysis of patch smoothers can rely on overlapping additive Schwarz methods [DW87] or on subspace correction methods [Xu92]. Let  $\{\Omega_j\}_{j=1}^J$  with  $J \in \mathbb{N}$ ,  $J > 2$ , be an overlapping subdomain decomposition of the domain  $\Omega$ , namely a collection of open subsets of  $\Omega$  whose boundaries are aligned with the mesh  $\mathcal{T}_h$ . The subspaces  $\{\Lambda_h^{d,1}(\Omega_j)\}_{j=1}^J$  associated with the domain partition  $\{\Omega_j\}_{j=1}^J$  are defined as

$$\Lambda_h^{d,1}(\Omega_j) := \{\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h) : \text{supp}(\mathbf{v}) \subseteq \Omega_j\}, \quad j \in [1, J] \cap \mathbb{N}.$$

In particular, we consider the following (overlapping) space decompositions:

$$\Lambda_h^{d,1}(\mathcal{T}_h) = \sum_{x_j \in \Delta_0(\mathcal{T}_h)} \Lambda_h^{d,1}(\Omega_j), \quad \Omega_j = \{T \in \mathcal{T}_h : x_j \in \partial T\} = \Delta_n(x_j), \quad (5.27a)$$

$$\Lambda_h^{d,1}(\mathcal{T}_h) = \sum_{e_j \in \Delta_1(\mathcal{T}_h)} \Lambda_h^{d,1}(\Omega_j), \quad \Omega_j = \{T \in \mathcal{T}_h : e_j \in \partial T\} = \Delta_n(e_j), \quad (5.27b)$$

$$\Lambda_h^{d,1}(\mathcal{T}_h) = \sum_{T_j \in \mathcal{T}_h} \Lambda_h^{d,1}(\Omega_j), \quad \Omega_j = \{T \in \mathcal{T}_h : \partial T_j \cap \partial T \in \Delta_{n-2}(\mathcal{T}_h)\}. \quad (5.27c)$$

Observe that

$$\Omega \subseteq \bigcup_{x_j \in \Delta_0(\mathcal{T}_h)} \Omega_j, \quad \Omega \subseteq \bigcup_{e_j \in \Delta_1(\mathcal{T}_h)} \Omega_j, \quad \Omega \subseteq \bigcup_{T_j \in \mathcal{T}_h} \Omega_j.$$

From the shape regularity assumption and by construction, for each of the domain decompositions, the patches (or subdomains)  $\Omega_j$  are of comparable size. Moreover, all the domain decompositions (5.27) have the *finite covering property*: for every  $\mathbf{x} \in \Omega$ , there is a finite number  $N(\mathbf{x})$  of patches containing  $\mathbf{x}$ . We define  $N_c = \max_{\mathbf{x} \in \Omega} N(\mathbf{x})$  which is a finite (and moderate) constant depending on the connectivity of the mesh. Introducing a partition of unity associated with each of the domain decompositions in (5.27), it can be shown that the subspaces  $\{\Lambda_h^{d,1}(\Omega_j)\}_{j=1}^J$  provide a space decomposition of  $\Lambda_h^{d,1}(\mathcal{T}_h)$  (see e.g. [BS08, Lemma 7.4.12]). Moreover, since  $\Lambda_h^{d,1}(\mathcal{T}_h)$  is a space of discontinuous piecewise polynomials, there are no continuity constraints imposed in the preceding space splittings.

In order to define the additive overlapping Schwarz (or additive subspace correction method) associated with a given domain decomposition of  $\Omega$  as in (5.27), we first consider the restriction of the IP-DG method (5.6) to the subspace  $\Lambda_h^{d,1}(\Omega_j)$ . The bilinear form in (5.7) becomes: for all  $\mathbf{v} \in \Lambda_h^{d,1}(\Omega_j)$ ,

$$\begin{aligned} a_{\text{DG}}^{\Omega_j}(\mathbf{v}, \mathbf{v}) &= \sum_{T \in \Omega_j} \left( \nu_T \|\mathbf{curl}_h \mathbf{v}\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}\|_{L^2(T)}^2 \right) - 2 \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h) \cap \Omega_j \setminus \partial \Omega_j} (\{\nu \mathbf{curl}_h \mathbf{v}\}_{f,\gamma}, [\mathbf{v}]_{f,\tau})_{L^2(f)} \\ &\quad - 2 \sum_{f \in \Delta_{n-1}^{\circ}(\mathcal{T}_h) \cap \partial \Omega_j} \{\nu\}_{f,H} (\mathbf{n}_f \times \mathbf{v}, \mathbf{curl}_h \mathbf{v})_{L^2(f)} \\ &\quad + \sum_{T \in \Omega_j} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \partial \Omega_j} (s_f [\mathbf{v}]_{f,\tau}, [\mathbf{v}]_{f,\tau})_{L^2(f)}, \\ &\quad + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \cap \partial \Omega_j} (s_f \mathbf{n}_f \times \mathbf{v}, \mathbf{n}_f \times \mathbf{v})_{L^2(f)}. \end{aligned} \tag{5.28}$$

The foregoing bilinear form defines the local solvers for  $\Omega_j \cap \partial \Omega = \emptyset$ . For the patches touching the boundary of  $\Omega$ , the third term in the sum is modified having  $\nu_T$  instead of  $\{\nu\}_{f,H}$ . The additive Schwarz smoother is defined as

$$s_O(\mathbf{v}, \mathbf{v}) := \sum_{j=1}^J a_{\text{DG}}^{\Omega_j}(\mathbf{v}_j, \mathbf{v}_j), \quad \text{with } \mathbf{v} = \sum_{j=1}^J \mathbf{v}_j, \quad \mathbf{v}_j \in \Lambda_h^{d,1}(\Omega_j). \tag{5.29}$$

If  $A_j : \Lambda_h^{d,1}(\Omega_j) \rightarrow (\Lambda_h^{d,1}(\Omega_j))'$  is the local operator associated with  $a_{\text{DG}}^{\Omega_j}(\cdot, \cdot)$ , and  $\iota_j : \Lambda_h^{d,1}(\Omega_j) \hookrightarrow \Lambda_h^{d,1}(\mathcal{T}_h)$  is the natural embedding, then the additive Schwarz smoothing operator reads  $S = \sum_{j=1}^J \iota_j \circ A_j^{-1} \circ \iota_j^*$ .

**Remark 5.2.3.** As we will show in Lemma 5.3.15 and observe numerically, the use of a block oriented smoother in the preconditioner  $B$  (5.22) is only essential when the auxiliary space  $W$  is “coarser” than  $\Lambda_h^1(\mathcal{T}_h)$ , i.e., is of type (b) (see Section 5.2.2). The requirement of using overlapping block-type smoothers in two-level and multi-level preconditioners for the finite element approximation of (5.1) has been observed (and illustrated numerically) in the literature by many authors and for different approaches [HT00; Hip03; AFW00]. The mathematical justification that a pointwise smoother in a two level preconditioner will not provide a convergent method is given in [Zik08], where a lower bound of order  $1 - ch^2$  on the convergence rate of such two level preconditioner is derived. More precisely, in [Zik08, Section 4], the author provides an explicit construction of a function that “cannot be seen” by the coarse solver and cannot be damped by a pointwise relaxation. Here, as we will show in Lemma 5.3.15, the use of a pointwise smoother would break the edge bubbles (not seen by the auxiliary space) leading to a component with arbitrary high energy which cannot be damped.

### Properties of the Smoothing Operators

For each of the smoothers introduced in (5.25), (5.26) and (5.29), we prove that the associated smoothing operator  $S$  has continuous and uniformly bounded inverse (property (F1)). Furthermore, we determine how the operator  $S$  scales in relation with the identity operator.

**Lemma 5.2.4.** *Let  $s(\cdot, \cdot)$  denote the bilinear form associated with the pointwise Jacobi smoother or with the block Jacobi smoother, as defined in the splittings of  $\Lambda_h^{d,1}(\mathcal{T}_h)$  in (5.25) and (5.26), respectively. Then, there exists  $c_s > 0$  independent of the mesh size and the coefficients of the problems but depending on the local polynomial space and on the shape regularity of the mesh such that*

$$a_{\text{DG}}(\mathbf{v}, \mathbf{v}) \leq c_s s(\mathbf{v}, \mathbf{v}), \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h).$$

Moreover,  $s(\cdot, \cdot)$  satisfies

$$s(\mathbf{v}, \mathbf{v}) \leq C \sum_{T \in \mathcal{T}_h} \left( \nu_T h_T^{-2} \|\mathbf{v}\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}\|_{L^2(T)}^2 + \alpha_T^\nu h_T^{-2} \|\mathbf{v}\|_{L^2(T)}^2 \right), \quad \forall \mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h), \tag{5.30}$$

with  $C > 0$  depending on the local polynomial space and on the shape regularity of the mesh.

*Proof.* To define the bilinear form  $s(\cdot, \cdot)$  relative to Jacobi smoother associated with the space splitting (5.25) we make use of the representation (5.5) of functions in  $\Lambda_h^{d,1}(\mathcal{T}_h)$ , where each basis function is considered as a *global* basis function on  $\Omega$  extended by zero outside of its support. The pointwise Jacobi relaxation reads,

$$\begin{aligned} s_J(\mathbf{v}, \mathbf{v}) := & \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e,T}^i)(V_{e,T}^i)^2 \\ & + \sum_{T \in \mathcal{T}_h} \sum_{f \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} a_{\text{DG}}(\varphi_{f,T}^i, \varphi_{f,T}^i)(V_{f,T}^i)^2 + \sum_{T \in \mathcal{T}_h} \sum_{i=1}^{N_b} a_{\text{DG}}(\varphi_T^i, \varphi_T^i)(V_T^i)^2, \end{aligned} \quad (5.31)$$

whilst the block Jacobi operator, using as blocks the elements  $T \in \mathcal{T}_h$  of the mesh, has bilinear form

$$\begin{aligned} s_{Jb}(\mathbf{v}, \mathbf{v}) := & \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_{n-2}(T)} \sum_{e' \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e',T}^i)V_{e,T}^i V_{e',T}^i \\ & + \sum_{T \in \mathcal{T}_h} \sum_{f \in \Delta_{n-1}(T)} \sum_{f' \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} a_{\text{DG}}(\varphi_{f,T}^i, \varphi_{f',T}^i)V_{f,T}^i V_{f',T}^i \\ & + \sum_{T \in \mathcal{T}_h} \sum_{i=1}^{N_b} a_{\text{DG}}(\varphi_T^i, \varphi_T^i)(V_T^i)^2. \end{aligned} \quad (5.32)$$

It is enough to focus on the lowest order case, namely for local degrees of freedom given by  $\{V_{e,T}^i\}_{i=1}^{N_e}$  for  $T \in \mathcal{T}_h$ ,  $e \in \Delta_1(T)$  with  $N_e \leq 2$ . The general case (as given in (5.31) and (5.32)) can be shown by arguing in the same way for the terms involving faces and elements degrees of freedom.

First we prove continuity of the pointwise smoother. Using the representation of  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  in (5.5), The Cauchy-Schwarz inequality and the arithmetic-geometric inequality yield

$$\begin{aligned} a_{\text{DG}}(\mathbf{v}, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{T' \in \mathcal{T}_h} \sum_{e' \in \Delta_1(T')} \sum_{i=1}^{N_e} a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e',T'}^i)V_{e,T}^i V_{e',T'}^i \\ &\leq \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{T' \in \mathcal{T}_h} \sum_{\substack{e' \in \Delta_1(T') \\ \partial T' \cap \partial T \neq \emptyset}} \sum_{i=1}^{N_e} \sqrt{a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e,T}^i)} \sqrt{a_{\text{DG}}(\varphi_{e',T'}^i, \varphi_{e',T'}^i)} V_{e,T}^i V_{e',T'}^i \\ &\leq C \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{i=1}^{N_e} a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e,T}^i)(V_{e,T}^i)^2 = C s_J(\mathbf{v}, \mathbf{v}), \end{aligned}$$

where the constant  $C$  depends on the shape regularity of  $\mathcal{T}_h$ . In order to prove (5.30), note that continuity and stability of the bilinear form  $a_{\text{DG}}(\cdot, \cdot)$  hold for each of the basis functions (considered as global functions). In particular, for all  $i = 1, \dots, N_e$ , we have  $C_{\text{stab}} \|\varphi_{e,T}^i\|_{\text{DG}}^2 \leq a_{\text{DG}}(\varphi_{e,T}^i, \varphi_{e,T}^i) \leq C_{\text{cont}} \|\varphi_{e,T}^i\|_{\text{DG}}^2$ . Moreover, under the assumption of shape regularity of  $\mathcal{T}_h$ , the standard inverse inequalities  $\|\mathbf{curl}_h \varphi_{e,T}^i\|_{L^2(T)}^2 \leq Ch_T^{-2} \|\varphi_{e,T}^i\|_{L^2(T)}^2$  and  $\|\mathbf{n}_f \times \varphi_{e,T}^i\|_{L^2(f)}^2 \leq Ch_T^{-1} \|\varphi_{e,T}^i\|_{L^2(T)}^2$  hold (see Lemma 2.4.6). Hence, the definition of the  $\|\cdot\|_{\text{DG}}$  norm (5.14) gives

$$\begin{aligned} s_J(\mathbf{v}, \mathbf{v}) &\simeq \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{i=1}^{N_e} (V_{e,T}^i)^2 \left( \nu_T \|\mathbf{curl}_h \varphi_{e,T}^i\|_{L^2(T)}^2 + \beta_T \|\varphi_{e,T}^i\|_{L^2(T)}^2 \right) \\ &\quad + \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_1(T)} \sum_{f \in \Delta_{n-1}(e) \cap \Delta_{n-1}(T)} \sum_{i=1}^{N_e} h_f^{-1} (V_{e,T}^i)^2 \|\mathbf{n}_f \times \varphi_{e,T}^i\|_{L^2(f)}^2 \\ &\leq C \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{i=1}^{N_e} (V_{e,T}^i)^2 \left( \nu_T h_T^{-2} \|\varphi_{e,T}^i\|_{L^2(T)}^2 + \beta_T \|\varphi_{e,T}^i\|_{L^2(T)}^2 \right) \end{aligned}$$

$$\begin{aligned}
 & + C \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_1(T)} \sum_{i=1}^{N_e} (V_{e,T}^i)^2 \sum_{f \in \Delta_{n-1}(e) \cap \Delta_{n-1}(T)} h_f^{-1} h_T^{-1} \|\varphi_{e,T}^i\|_{L^2(T)}^2 \\
 & \leq C \sum_{T \in \mathcal{T}_h} \left( \nu_T h_T^{-2} \|\mathbf{v}\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}\|_{L^2(T)}^2 + \alpha_T^\nu h_T^{-2} \|\mathbf{v}\|_{L^2(T)}^2 \right),
 \end{aligned}$$

with  $C > 0$  depending on the mesh  $\mathcal{T}_h$  through its shape regularity. For the block Jacobi smoother  $s_{Jb}$  in (5.32), using the short-hand notation  $\mathbf{v}_i := \mathbf{v}|_{T_i}$ , the Cauchy-Schwarz inequality yields

$$\begin{aligned}
 a_{DG}(\mathbf{v}, \mathbf{v}) & = a_{DG} \left( \sum_{T_i \in \mathcal{T}_h} \mathbf{v}_i, \sum_{T_j \in \mathcal{T}_h} \mathbf{v}_j \right) \leq \sum_{T_i \in \mathcal{T}_h} \sum_{\substack{T_j \in \mathcal{T}_h \\ \partial T_i \cap \partial T_j \neq \emptyset}} \sqrt{a_{DG}(\mathbf{v}_i, \mathbf{v}_i)} \sqrt{a_{DG}(\mathbf{v}_j, \mathbf{v}_j)} \\
 & \leq \frac{1}{2} \sum_{T_i \in \mathcal{T}_h} \sum_{\substack{T_j \in \mathcal{T}_h \\ \partial T_i \cap \partial T_j \neq \emptyset}} (a_{DG}(\mathbf{v}_i, \mathbf{v}_i) + a_{DG}(\mathbf{v}_j, \mathbf{v}_j)) \leq C \sum_{T_i \in \mathcal{T}_h} a_{DG}(\mathbf{v}_i, \mathbf{v}_i) = Cs_{Jb}(\mathbf{v}, \mathbf{v}).
 \end{aligned}$$

Moreover, the non-overlapping block Jacobi smoother enjoys the same spectral scaling of the pointwise Jacobi relaxation. Indeed, by (5.32),

$$\begin{aligned}
 s_{Jb}(\mathbf{v}, \mathbf{v}) & = \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{e' \in \Delta_1(T)} \sum_{i=1}^{N_e} a_{DG}(\varphi_{e,T}^i, \varphi_{e',T}^i) V_{e,T}^i V_{e',T}^i \\
 & \leq \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{e' \in \Delta_1(T)} \sum_{i=1}^{N_e} \sqrt{a_{DG}(\varphi_{e,T}^i, \varphi_{e,T}^i)} \sqrt{a_{DG}(\varphi_{e',T}^i, \varphi_{e',T}^i)} V_{e,T}^i V_{e',T}^i \\
 & \leq C \sum_{T \in \mathcal{T}_h} \sum_{e \in \Delta_1(T)} \sum_{i=1}^{N_e} a_{DG}(\varphi_{e,T}^i, \varphi_{e,T}^i) (V_{e,T}^i)^2 = Cs_J(\mathbf{v}, \mathbf{v}),
 \end{aligned}$$

and therefore, the estimate (5.30) holds also in the case of block Jacobi relaxation.  $\square$

The next Lemma pertains to the continuity and spectral properties of the overlapping smoothers. In particular, it establishes the property (F1) in Theorem 5.2.2, and provides the scaling of the bilinear form  $s_O(\cdot, \cdot)$  in (5.29).

**Lemma 5.2.5.** *Let  $\Lambda_h^{d,1}(\mathcal{T}_h) = \sum_{j=1}^J \Lambda_h^{d,1}(\Omega_j)$  with subspaces  $\Lambda_h^{d,1}(\Omega_j)$  defined as in (5.27), and let  $s_O(\cdot, \cdot)$  be the corresponding (overlapping) additive Schwarz method given in (5.29). Then, for every  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and any choice of  $\{\mathbf{v}_j \in \Lambda_h^{d,1}(\Omega_j)\}_{j=1}^J$  for which  $\mathbf{v} = \sum_{j=1}^J \mathbf{v}_j$ , it holds*

$$a_{DG}(\mathbf{v}, \mathbf{v}) \leq c_s s_O(\mathbf{v}, \mathbf{v}), \quad (5.33)$$

where the constant  $c_s > 0$  depends on the local polynomial space, the shape regularity and connectivity of the mesh and on the amount of overlapping  $N_c$  in the subdomain partition  $\{\Omega_j\}_{j=1}^J$ . Moreover,  $s_O(\cdot, \cdot)$  satisfies

$$\begin{aligned}
 s_O(\mathbf{v}, \mathbf{v}) & \simeq \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \left( \nu_T \|\mathbf{curl}_h \mathbf{v}_j\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}_j\|_{L^2(T)}^2 \right) \right. \\
 & \quad + \sum_{T \in \Omega_j} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \partial \Omega_j} h_f^{-1} \|[\mathbf{v}_j]_{f,\tau}\|_{L^2(f)}^2 \\
 & \quad \left. + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \cap \partial \Omega_j} h_f^{-1} \|\mathbf{n}_f \times \mathbf{v}_j\|_{L^2(f)}^2 \right).
 \end{aligned} \quad (5.34)$$

*Proof.* Note that (5.33) states that in the decomposition  $\mathbf{v} = \sum_{j=1}^J \mathbf{v}_j$ , the energy of the parts bounds the energy of the function in  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$ . To measure the overlap of the domain splitting we introduce

the constants

$$c_{jk} = \begin{cases} 1 & \text{if } \Omega_j \cap \Omega_k \neq \emptyset \\ 0 & \text{if } \Omega_j \cap \Omega_k = \emptyset \end{cases} \quad j, k = 1, \dots, J.$$

The Cauchy-Schwarz inequality, the arithmetic-geometric inequality and the fact that  $c_{jk} = c_{kj}$  for all  $j, k = 1, \dots, J$ , gives

$$\begin{aligned} a_{\text{DG}}(\mathbf{v}, \mathbf{v}) &= a_{\text{DG}}\left(\sum_{j=1}^J \mathbf{v}_j, \sum_{k=1}^J \mathbf{v}_k\right) = \sum_{j,k=1}^J c_{jk} a_{\text{DG}}(\mathbf{v}_j, \mathbf{v}_k) \leq \sum_{j,k=1}^J \sqrt{c_{jk} a_{\text{DG}}(\mathbf{v}_j, \mathbf{v}_j)} \sqrt{c_{kj} a_{\text{DG}}(\mathbf{v}_k, \mathbf{v}_k)} \\ &\leq \frac{1}{2} \sum_{j,k=1}^J (c_{jk} a_{\text{DG}}(\mathbf{v}_j, \mathbf{v}_j) + c_{kj} a_{\text{DG}}(\mathbf{v}_k, \mathbf{v}_k)) \leq \sum_{j,k=1}^J c_{jk} a_{\text{DG}}(\mathbf{v}_j, \mathbf{v}_j) \\ &= \sum_{j=1}^J a_{\text{DG}}(\mathbf{v}_j, \mathbf{v}_j) \left( \sum_{k=1}^J c_{jk} \right) \leq N_c \sum_{j=1}^J a_{\text{DG}}^{\Omega_j}(\mathbf{v}_j, \mathbf{v}_j) = c_s s_O(\mathbf{v}, \mathbf{v}), \end{aligned}$$

where  $c_s = \max_{1 \leq j \leq J} \#\{k : \Omega_j \cap \Omega_k \neq \emptyset\}$ . The proof of (5.34) reduces to use the continuity and coercivity of  $a_{\text{DG}}^{\Omega_j}(\cdot, \cdot)$  on the subspace  $\Lambda_h^{\text{d},1}(\Omega_j)$  for every  $j = 1, \dots, J$ .  $\square$

### 5.3. Asymptotic Optimality of the Preconditioner

This section is devoted to the stability analysis of the preconditioners introduced in Section 5.2.2 and Section 5.2.3. We first state and discuss the main results.

**Theorem 5.3.1.** *Let  $\mathbb{B}$  be the auxiliary space preconditioner as defined in (5.23) and associated with one of the following:*

- (i) Auxiliary space  $\Lambda_h^1(\mathcal{T}_h)$  as in (5.4) and Jacobi smoother relative to the splitting of  $\Lambda_h^{\text{d},1}(\mathcal{T}_h)$  in (5.25) or (5.26);
- (ii) Auxiliary space  $\mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$  and overlapping patch smoother as in (5.29) and (5.28).

Then,

$$\kappa(\mathbb{B}\mathbb{A}) \leq c_a (1 + c_s) \max\{1, \delta(\nu, \beta)\},$$

where

$$\delta(\nu, \beta) := \min \left\{ \max_{T \in \mathcal{T}_h} \frac{h_T^2 \beta_T}{\nu_T}, \max_{\substack{T, T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\beta_T}{\beta_{T'}}, \max_{\substack{T \in \Upsilon_h, T' \in \Upsilon'_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\alpha_T^\nu}{\alpha_{T'}^\nu} \right\}, \quad (5.35)$$

with  $\Upsilon_h$  and  $\Upsilon'_h$  as in (5.15). The constants  $c_s, c_a$  depend only on the polynomial degree and on the shape regularity of the mesh.

The proof of the above-stated theorem relies on the application of the fictitious space Theorem 5.2.1 and therefore boils down to the verification of conditions **(F0)**, **(F1)** and **(F2)** in Theorem 5.2.2. As pointed out in the description of the preconditioners in Section 5.2.2, condition **(F0)** is trivially satisfied in all cases, as can be seen from the definition of  $a_w(\cdot, \cdot)$  and owing to the fact that  $\Pi_w$  is the standard inclusion. Moreover, Property **(F1)** has been verified in Section 5.2.3 for the overlapping and non-overlapping smoothing operators. Therefore, we need to establish Property **(F2)** regarding the stability of the decomposition (5.21).

### 5.3.1. $\nu$ -stable Decomposition

Typically, in the auxiliary space framework, the first step towards a stable decomposition relies on the construction of an *averaging* interpolation operator from the nonconforming polynomial space  $\Lambda_h^{d,1}(\mathcal{T}_h)$  to its  $H(\mathbf{curl}, \Omega)$ -conforming counterpart. This coincides with the operator introduced in Definition 3.3.2, when considering the vector proxies of differential 1-forms in  $\mathbb{R}^n$ ,  $n \in \{2, 3\}$ , and with the averaging in [Hou+05, Appendix]. However, aiming at a decomposition uniformly stable in the problems coefficients, we introduce an averaging operator  $\mathcal{P}_h^\nu : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^{d,1}(\mathcal{T}_h) \cap H(\mathbf{curl}, \Omega)$  which takes into account the presence of the coefficients in the model problem (5.1). It is worth stressing that the approximation and stability properties of the interpolation operator introduced in [Hou+05, Appendix] has been already used in the different contexts related to the DG approximation of the time harmonic Maxwell problem; in the analysis of spectral approximation [BP06] (see also [BHP07] for the numerical verification of the results), and in different works related to a posteriori error estimates [HPS05; HPS07].

As a side note, the asymptotic analysis and estimates derived in the remainder of this chapter, require to introduce constants. With a small abuse of notation, by  $C$  we will denote a generic positive constant whose value may vary among different occurrences, but, unless otherwise specified, will be independent of the mesh width and the coefficients of the problem and may only depend on the polynomial degree, the shape regularity and the connectivity of the mesh partition.

#### $\nu$ -weighted Averaging Operator

To ease the notation, we set  $\mathbf{z} := \mathcal{P}_h^\nu(\mathbf{v})$  and, since  $\mathbf{z} \in \Lambda_h^{d,1}(\mathcal{T}_h) \cap H(\mathbf{curl}, \Omega)$ , we can write  $\mathbf{z}|_T$  in the basis of  $\zeta_r(T)$  for all  $T \in \mathcal{T}_h$ . More precisely, we have the following representation:

$$\mathbf{z}(\mathbf{x}) = \sum_{e \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} Z_e^i \varphi_{e,T}^i(\mathbf{x}) + \sum_{f \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} Z_f^i \varphi_{f,T}^i(\mathbf{x}) + \sum_{T}^{N_b} Z_T^i \varphi_T^i(\mathbf{x}), \quad \forall \mathbf{x} \in T, \forall T \in \mathcal{T}_h. \quad (5.36)$$

Hence, in order to define  $\mathbf{z}$ , it is enough to specify the coefficients  $\mathbf{z}_e = \{Z_e^i\}_{i=1}^{N_e}$ ,  $\mathbf{z}_f = \{Z_f^i\}_{i=1}^{N_f}$  and  $\mathbf{z}_T = \{Z_T^i\}_{i=1}^{N_b}$  in terms of those of  $\mathbf{v}$ , and simultaneously ensure the  $H(\mathbf{curl}, \Omega)$ -conformity of the global function. First, we define two sets of weights associated with  $(n-1)$ - and  $(n-2)$ -cells of the mesh  $\mathcal{T}_h$ . More precisely, let  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h)$  be an interior  $(n-1)$ -face such that  $f = \partial T^+ \cap \partial T^-$ . We define

$$\omega_{f,T^+} := \frac{\sqrt{\nu_{T^+}}}{\sqrt{\nu_{T^+}} + \sqrt{\nu_{T^-}}}, \quad \text{and} \quad \omega_{f,T^-} := 1 - \omega_{f,T^+} = \frac{\sqrt{\nu_{T^-}}}{\sqrt{\nu_{T^+}} + \sqrt{\nu_{T^-}}}. \quad (5.37)$$

Note that  $2\nu_{T^\pm}\omega_{f,T^\mp} \leq \{\nu\}_{f,H} \leq \{\nu\}_{*,f} \leq \alpha_{T^+}(\nu), \alpha_{T^-}(\nu)$ , with  $\{\cdot\}_{*,f}$  and  $\alpha_f^\nu$  defined as in (5.9) and  $\{\cdot\}_{f,H}$  as in (5.11). In order to define the weights on the  $(n-2)$ -cells of the mesh, let  $e \in \Delta_{n-2}^\circ(\mathcal{T}_h)$ . The cardinality of the set  $\Delta_n(e)$  is bounded by a finite constant depending on the shape regularity and the connectivity of the mesh  $\mathcal{T}_h$ , uniformly with respect to  $h$ . Let  $\nu_j := \nu|_{T_j}$ , we define

$$\omega_{e,T_\ell} := \frac{\sqrt{\nu_\ell}}{\sum_{T_j \in \Delta_n(e)} \sqrt{\nu_j}} \quad \forall T_\ell \in \Delta_n(e). \quad (5.38)$$

The weighted averaging projection operator  $\mathcal{P}_h^\nu$  is defined as:

**Definition 5.3.2.** Let  $\mathcal{P}_h^\nu : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$  be such that, for any  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$ ,  $\mathbf{z} = \mathcal{P}_h^\nu(\mathbf{v})$  is given by (5.36), with coefficients  $\mathbf{z}_e = \{Z_e^i\}_{i=1}^{N_e}$ ,  $\mathbf{z}_f = \{Z_f^i\}_{i=1}^{N_f}$  and  $\mathbf{z}_T = \{Z_T^i\}_{i=1}^{N_b}$  defined as follows.

- For every  $T \in \mathcal{T}_h$  the coefficients  $\mathbf{z}_T$  associated with volume degrees of freedom are set equal to those of  $\mathbf{v}_T$ :  $Z_T^i = V_T^i$ , for all  $i = 1, \dots, N_b$ .
- The coefficients  $\mathbf{z}_f$  associated with  $(n-1)$ -cell moments are defined, for all  $i = 1, \dots, N_f$ , as

$$Z_f^i = \begin{cases} \omega_{f,T^+} V_{f,T^+}^i + \omega_{f,T^-} V_{f,T^-}^i & \forall f \in \Delta_{n-1}^\circ(\mathcal{T}_h) \cap \Delta_{n-1}(T^+) \cap \Delta_{n-1}(T^-), \\ V_{f,T}^i & \forall f \in \Delta_{n-1}^\partial(\mathcal{T}_h) \cap \Delta_{n-1}(T), \end{cases} \quad (5.39)$$

where the weights  $\omega_{f,T^+}, \omega_{f,T^-}$  are as in (5.37).

- The coefficients  $\mathbf{z}_e$  associated with  $(n-2)$ -cell moments are defined, for all  $i = 1, \dots, N_e$ , as the convex combination

$$Z_e^i = \begin{cases} \sum_{T_\ell \in \Delta_n(e)} \omega_{e,T_\ell} V_{e,T_\ell}^i & \forall e \in \Delta_{n-2}^\circ(\mathcal{T}_h), \\ V_{e,T}^i & \forall e \in \Delta_{n-2}^\partial(\mathcal{T}_h) \cap \Delta_{n-2}(T), \end{cases} \quad (5.40)$$

with weights  $\{\omega_{e,T_\ell}\}_\ell$  defined in (5.38).

The definition of  $\mathcal{P}_h^\nu$  is completely general with respect to the distribution of the coefficient  $\nu$  which is only required to be piecewise constant on  $\mathcal{T}_h$  for every  $h$ . In particular, for  $\nu \equiv 1$  in  $\Omega$ , the averaging operator  $\mathcal{P}_h^\nu$  reduces to the projection operator proposed in [Hou+05, Appendix]. The following result provides the approximation properties of  $\mathcal{P}_h^\nu$  in the local  $\nu$ -weighted  $L^2$ -norm.

**Lemma 5.3.3.** *Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^\nu : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$  be the averaging operator introduced in Definition 5.3.2. Then, there exists a constant  $C > 0$  depending only on the polynomial degree and the shape regularity of the mesh such that, for all  $T \in \mathcal{T}_h$ , there holds*

$$\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \leq C \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f \|\llbracket \mathbf{v} \rrbracket_{f,\tau}\|_{L^2(f)}^2. \quad (5.41)$$

The estimate (5.41) holds also for the local  $\nu$ -weighted norm of the approximation error, since  $\nu_T < \alpha_T^\nu$ , for all  $T \in \mathcal{T}_h$ .

Before giving the proof of Lemma 5.3.3, we present a stability result of the averaging operator  $\mathcal{P}_h^\nu$  in the  $\beta$ -weighted  $L^2$ -norm defined in (5.13).

**Corollary 5.3.4.** *Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^\nu : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$  be the averaging operator as in Definition 5.3.2. Then, there exists  $C > 0$  depending only on the polynomial degree and the shape regularity of the mesh such that*

$$\|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \leq C \max\{1, \theta(\nu, \beta)\} \|\mathbf{v}\|_{DG}^2, \quad (5.42)$$

where  $\theta(\nu, \beta)$  is defined as

$$\theta(\nu, \beta) := \min \left\{ \max_{T \in \mathcal{T}_h} \frac{h_T^2 \beta_T}{\nu_T}, \max_{\substack{T, T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\beta_T}{\beta_{T'}} \right\}. \quad (5.43)$$

*Proof.* The proof boils down to showing the following estimates,

$$\|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \leq C \sum_{T \in \mathcal{T}_h} h_T^2 \frac{\beta_T}{\nu_T} \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_T^\nu h_f^{-1} \|\llbracket \mathbf{v} \rrbracket_{f,\tau}\|_{L^2(f)}^2, \quad (5.44)$$

$$\|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \leq C \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{v}\|_{L^2(T)} + \sum_{\substack{T, T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\beta_T}{\beta_{T'}} \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2. \quad (5.45)$$

To prove (5.44), we multiply and divide by  $\nu_T$  the elementwise  $\beta$ -weighted norms of the approximation error and exploit the local  $\nu$ -weighted  $L^2$ -estimate (5.41) (since  $\nu_T < \alpha_T^\nu$  for all  $T \in \mathcal{T}_h$ ), namely

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 &= \sum_{T \in \mathcal{T}_h} \frac{\beta_T}{\nu_T} \nu_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \\ &\stackrel{(5.41)}{\leq} C \sum_{T \in \mathcal{T}_h} \frac{\beta_T}{\nu_T} \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_T^\nu h_f \|\llbracket \mathbf{v} \rrbracket_{f,\tau}\|_{L^2(f)}^2 \\ &\leq C \sum_{T \in \mathcal{T}_h} h_T^2 \frac{\beta_T}{\nu_T} \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_T^\nu h_f^{-1} \|\llbracket \mathbf{v} \rrbracket_{f,\tau}\|_{L^2(f)}^2. \end{aligned}$$

In order to show (5.45), we use the estimate (5.41) with  $\alpha_T^\nu = 1$  for all  $T \in \mathcal{T}_h$ , trace inequalities and inverse inequalities to get

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 &\leq C \sum_{T \in \mathcal{T}_h} \beta_T \sum_{f \in \Delta_{n-1}(T)} h_f \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\quad + \sum_{T \in \mathcal{T}_h} \beta_T \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \Delta_{n-1}(T)} h_f \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\leq C \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{v}\|_{L^2(T)}^2 + C \sum_{T \in \mathcal{T}_h} \sum_{\substack{T' \in \mathcal{T}_h \setminus \{T\} \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\beta_T}{\beta_{T'}} \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2. \end{aligned}$$

Owing to the definition of the DG-norm (5.14), the estimates (5.44) and (5.45) are combined to obtain (5.42).  $\square$

*Proof of Lemma 5.3.3.* First, let  $T \in \mathcal{T}_h$  be an arbitrary element that does not intersect the boundary  $\partial\Omega$ . To estimate the difference  $\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})$  we use the representations (5.5) and (5.36). The norm equivalence (2.34), and the transformation of the  $L^2$ -norm under pullback (2.35), result in

$$\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \leq Ch_T^{n-2} \left( \sum_{f \in \Delta_{n-1}(T)} \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 \right). \quad (5.46)$$

We estimate each of the contributions on the right hand side above separately. Let  $f \in \Delta_{n-1}^\circ(T)$  be an  $(n-1)$ -interior face such that  $f = \partial T' \cap \partial T$  with  $T, T' \in \mathcal{T}_h$ . Combining the definition of the averaging operator and the corresponding degrees of freedom (5.39), with (2.37) gives

$$\begin{aligned} \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 &\stackrel{(5.39)}{=} \alpha_T^\nu \|(1 - \omega_{f,T})\mathbf{v}_{f,T} - \omega_{f,T'} \mathbf{v}_{f,T'}\|_{\ell^2}^2 \stackrel{(5.37)}{=} \alpha_T^\nu (\omega_{f,T'})^2 \|\mathbf{v}_{f,T} - \mathbf{v}_{f,T'}\|_{\ell^2}^2 \\ &\leq \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{v}_{f,T'}\|_{\ell^2}^2 \stackrel{(2.37)}{\leq} C \alpha_T^\nu h_f^{3-n} \int_f |[\mathbf{v}]_{f,\tau}|^2 ds, \end{aligned} \quad (5.47)$$

since the weights satisfy  $\omega_{f,T} < 1$  for all  $T \in \mathcal{T}_h$  and  $f \in \Delta_{n-1}(\mathcal{T}_h)$ . Concerning the degrees of freedom on the  $(n-2)$ -faces, using the definition in (5.40) results in

$$\sum_{e \in \Delta_{n-2}(T)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 \leq C \sum_{e \in \Delta_{n-2}(T)} \sum_{\substack{T_\ell \in \Delta_n(e) \\ T_\ell \neq T}} \alpha_{T_\ell}^\nu (\omega_{e,T_\ell})^2 \|\mathbf{v}_{e,T} - \mathbf{v}_{e,T_\ell}\|_{\ell^2}^2. \quad (5.48)$$

To estimate such terms, we introduce a numbering (ordering) of the elements in the set  $\Delta_n(e)$  for a fixed  $e \in \Delta_{n-2}(T)$ , such that  $\Delta_n(e) = \bigcup_{\ell=0}^{M_e} T_\ell$  with  $M_e := |\Delta_n(e)| - 1$ . The ordering is such that  $T_0 := T$  and

$$\begin{aligned} \partial T_0 \cap \partial T_1 &\in \Delta_{n-1}(T_0), & \text{and} && \partial T_0 \cap \partial T_{M_e} &\in \Delta_{n-1}(T_0), \\ \partial T_0 \cap \partial T_\ell &\in \Delta_{n-2}(T_0), & \text{but} && \partial T_0 \cap \partial T_\ell &\notin \Delta_{n-1}(T_0), & \forall \ell = 2, \dots, M_e - 1, \\ \partial T_\ell \cap \partial T_{\ell+1} &\in \Delta_{n-1}(T_\ell) \cap \Delta_{n-1}(T_{\ell+1}) \neq \emptyset, & \forall \ell = 2, \dots, M_e - 1. \end{aligned} \quad (5.49)$$

With the foregoing numbering, summing and subtracting suitable degrees of freedom (see [KP03, Lemma 2.2] for a similar trick), results in

$$\begin{aligned} \sum_{\substack{T_\ell \in \Delta_n(e) \\ T_\ell \neq T_0}} \alpha_{T_0}^\nu (\omega_{e,T_\ell})^2 \|\mathbf{v}_{e,T_0} - \mathbf{v}_{e,T_\ell}\|_{\ell^2}^2 &\leq \alpha_{T_0}^\nu (\omega_{e,T_1})^2 \|\mathbf{v}_{e,T_0} - \mathbf{v}_{e,T_1}\|_{\ell^2}^2 \\ &\quad + \sum_{\ell=2}^{M_e-1} \alpha_{T_0}^\nu (\omega_{e,T_\ell})^2 \ell \sum_{j=0}^{\ell-1} \|\mathbf{v}_{e,T_j} - \mathbf{v}_{e,T_{j+1}}\|_{\ell^2}^2 \\ &\quad + \alpha_{T_0}^\nu (\omega_{e,T_{M_e}})^2 \|\mathbf{v}_{e,T_0} - \mathbf{v}_{e,T_{M_e}}\|_{\ell^2}^2. \end{aligned}$$

In view of the fact that  $\omega_{e,T} < 1$ , for all  $T \in \mathcal{T}_h$  and  $e \in \Delta_{n-2}(\mathcal{T}_h)$ , we can use the bounds on the trace of the interfaces jump (2.37), and the fact that  $\ell$  is uniformly bounded ( $\ell \leq M_e$ ); thereby

$$\begin{aligned} \sum_{\substack{T_\ell \in \Delta_n(e) \\ T_\ell \neq T_0}} \alpha_{T_0}^\nu (\omega_{e,T_\ell})^2 h_f^{3-n} \|\mathbf{v}_{e,T_0} - \mathbf{v}_{e,T_\ell}\|_{\ell^2}^2 &\leq C \alpha_{T_0}^\nu \sum_{\substack{T_\ell \in \Delta_n(e) \\ T_\ell \neq T_0}} \sum_{f \in \Delta_{n-1}(e) \cap \Delta_{n-1}(T_\ell)} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\leq C \alpha_{T_0}^\nu \sum_{f \in \Delta_{n-1}(e)} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned} \quad (5.50)$$

Hence, substituting (5.50) into (5.48) yields

$$\sum_{e \in \Delta_{n-2}(T)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 \leq C \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2, \quad (5.51)$$

where the constant  $C > 0$  only depends on the shape regularity and connectivity of the mesh.

For elements  $T \in \mathcal{T}_h$  at the domain boundary  $\partial T \cap \partial \Omega \neq \emptyset$ , the same type of estimates can be obtained by exploiting the fact that, in view of the boundary conditions, the degrees of freedom  $\mathbf{z}_{f,T}$  and  $\mathbf{z}_{e,T}$  are set to zero on the boundary faces (see (5.39) and (5.40)). In greater detail, analogously to (5.46), it holds

$$\begin{aligned} \nu_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 &\leq C h_T^{n-2} \left( \sum_{f \in \Delta_{n-1}(T) \cap \Delta_{n-1}^\partial} \nu_T \|\mathbf{v}_{f,T}\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T) \cap \Delta_{n-2}^\partial} \nu_T \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right. \\ &\quad \left. + \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\partial} \nu_T \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^\partial} \nu_T \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 \right). \end{aligned} \quad (5.52)$$

The last two contributions are estimated as in (5.47) and (5.51), respectively. The first two terms can be bounded by arguing similarly, but using equivalence (2.37) on the faces  $f \in \Delta_{n-1}(T)$ , namely

$$\begin{aligned} \nu_T \left( \sum_{f \in \Delta_{n-1}(T) \cap \Delta_{n-1}^\partial} \|\mathbf{v}_{f,T}\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T) \cap \Delta_{n-2}^\partial} \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right) &\leq 2\nu_T \sum_{f \in \Delta_{n-1}(T) \cap \Delta_{n-1}^\partial} \left( \|\mathbf{v}_{f,T}\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(f)} \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right) \\ &\leq C \sum_{f \in \Delta_{n-1}(T) \cap \Delta_{n-1}^\partial} \alpha_T^\nu h_f^{3-n} \int_f |\mathbf{n}_f \times \mathbf{v}|^2 ds. \end{aligned} \quad (5.53)$$

To conclude, the fact that the “jump” on the boundary  $\partial \Omega$  reduces to the plain tangential trace is used. Therefore, substituting into (5.46) and (5.52) the local contributions from the interior faces (5.47), from the interior  $(n-2)$ -faces (5.51) and from the boundary (5.53), yields (5.41).  $\square$

### Stable Decomposition

By means of the approximation properties of the averaging operator in Definition 5.3.2, we establish a first stability result for the splitting associated with the Case (a) in Section 5.2.2, paying particular attention to the distribution of the coefficients  $\nu$  and  $\beta$ .

**Proposition 5.3.5.** *Let the mesh  $\mathcal{T}_h$  be shape regular and local quasi-uniform as in Assumption 2.4.2. Let  $\Lambda_h^{d,1}(\mathcal{T}_h)$  be defined as in (5.3) and let  $\Lambda_h^1(\mathcal{T}_h) = \Lambda_h^{d,1}(\mathcal{T}_h) \cap H_0(\mathbf{curl}, \Omega)$  be the corresponding  $H(\mathbf{curl}, \Omega)$ -conforming finite element space. Let  $s(\cdot, \cdot)$  be a pointwise smoother as defined in (5.31) or in (5.32). Then, for any  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  there exist  $\mathbf{v}_0 \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and  $\mathbf{z} \in \Lambda_h^1(\mathcal{T}_h)$  such that  $\mathbf{v} = \mathbf{v}_0 + \mathbf{z}$  and*

$$s(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{z}, \mathbf{z}) \leq c_0^2 \max\{1, \theta(\nu, \beta)\} a_{DG}(\mathbf{v}, \mathbf{v}), \quad (5.54)$$

where  $\theta(\nu, \beta)$  is defined as in (5.43) and the constant  $c_0^2 > 0$  depends only the polynomial degree and on the shape regularity of the mesh.

*Proof.* Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^\nu$  be the averaging operator introduced in Definition 5.3.2. Since by construction,  $\mathcal{P}_h^\nu(\mathbf{v}) \in \Lambda_h^1(\mathcal{T}_h)$ , for all  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  we take  $\mathbf{v}_0 = \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}) \in \Lambda_h^{d,1}(\mathcal{T}_h)$ . Then, the scaling (5.30) of the pointwise smoothers, together with the approximation estimates (5.41) and (5.42) and the fact that  $\nu_T \leq \alpha_T^\nu$  for any  $T \in \mathcal{T}_h$  results in

$$\begin{aligned} s(\mathbf{v}_0, \mathbf{v}_0) &\lesssim \sum_{T \in \mathcal{T}_h} \nu_T h_T^{-2} \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 + \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 + \sum_{T \in \mathcal{T}_h} h_T^{-2} \alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \\ &\lesssim \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + \max\{1, \theta(\nu, \beta)\} \|\mathbf{v}\|_{DG}^2. \end{aligned}$$

To conclude we need to consider  $a_w(\mathcal{P}_h^\nu(\mathbf{v}), \mathcal{P}_h^\nu(\mathbf{v}))$ . Since

$$a_w(\mathcal{P}_h^\nu(\mathbf{v}), \mathcal{P}_h^\nu(\mathbf{v})) \leq a_w(\mathbf{v}, \mathbf{v}) + a_w(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})),$$

the stability proof reduces to bound the weighted  $H(\mathbf{curl}, \Omega)$ -norm of the difference  $\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})$  (using the continuity of  $a_w(\cdot, \cdot)$  in that norm). Hence, a standard application of inverse inequality together with (5.41) and (5.42), yields

$$\begin{aligned} a_w(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})) &\leq \|\mathbf{curl}_{\mathbf{h}}(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}))\|_{0,\nu,\mathcal{T}_h}^2 + \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \\ &\leq C \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \\ &\leq C \max\{1, \theta(\nu, \beta)\} \|\mathbf{v}\|_{DG}^2. \end{aligned}$$

Collecting the foregoing two estimates results in (5.54). Since the choice of  $\mathbf{v}$  was arbitrary this concludes the proof.  $\square$

**Remark 5.3.6.** The results derived in Proposition 5.3.5 entail that:

- If the problem is **curl**-dominated in the whole domain (i.e.  $\nu_T \geq \beta_T h_T^2$  for all  $T \in \mathcal{T}_h$ ), the estimate (5.54) guarantees that the auxiliary space preconditioner is uniformly convergent and robust with respect to jumps in the coefficients  $\beta$  and  $\nu$ . In this case one could replace the solution operator in the auxiliary space  $A_w$  by the domain decomposition preconditioner proposed in [DW16], getting an optimal solver.
- If the reaction coefficient  $\beta$  is such that there exists  $B > 1$  for which  $B^{-1} \leq \beta_T/\beta_{T'} \leq B$  for all  $T, T' \in \mathcal{T}_h$  with  $\partial T \cap \partial T' \neq \emptyset$ , (5.54) ensures the uniform convergence of the auxiliary space preconditioner and the robustness with respect to possible jumps in the coefficient  $\nu$ . This could be considered in agreement with the results available in the literature concerning auxiliary space type preconditioners for discretizations of second order problems with only one jumping coefficient [CPP14; Bri+15].
- If the problem is reaction-dominated in the whole domain and  $\beta$  is allowed to have high variations in different regions, the application of Proposition 5.3.5 would predict a convergence affected significantly by the size of the largest ratio  $h_T^2 \beta_T / \nu_T$  and the largest jump on the reaction coefficient  $\beta$ . However, such prediction might be pessimistic and would not endorse the results obtained in the actual computations, as reported in the numerical experiments of Section 5.4. In addition, if the problem is reaction-dominated in the whole domain, one might expect that the auxiliary space solver is not required in the preconditioner (5.23). In fact, as we will show in the subsequent Proposition 5.3.10 and in the numerical experiments, by suitably *turning off* the auxiliary space solver in the preconditioner, its convergence will not be jeopardized by the largest ratio  $h_T^2 \beta_T / \nu_T$  nor by the largest jump on the reaction coefficient.

### 5.3.2. $\beta$ -aware Averaging Operator

In order to address the most general case, namely when the problem is **curl**-dominated in some regions and reaction-dominated in others, we introduce an additional averaging operator  $\mathcal{P}_h^{\nu,\beta} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$ . It is defined in terms of the operator  $\mathcal{P}_h^\nu$  from Definition 5.3.2, but, additionally, it takes into account the

distribution of the coefficient  $\beta$ , by *turning off* the auxiliary space correction in the reaction-dominated regime.

**Definition 5.3.7.** Let  $\mathcal{P}_h^{\nu,\beta} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^{d,1}(\mathcal{T}_h) \cap H_0(\mathbf{curl}, \Omega)$  and let  $\bar{\mathbf{z}} := \mathcal{P}_h^{\nu,\beta}(\mathbf{v})$  be defined through the local representation (analogue of (5.36)),

$$\bar{\mathbf{z}}|_T(\mathbf{x}) = \sum_{e \in \Delta_{n-2}(T)} \sum_{i=1}^{N_e} \bar{Z}_e^i \varphi_{e,T}^i(\mathbf{x}) + \sum_{f \in \Delta_{n-1}(T)} \sum_{i=1}^{N_f} \bar{Z}_f^i \varphi_{f,T}^i(\mathbf{x}) + \sum_{i=1}^{N_b} \bar{Z}_T^i \varphi_T^i(\mathbf{x}), \quad \forall \mathbf{x} \in T, \forall T \in \mathcal{T}_h.$$

The degrees of freedom of  $\bar{\mathbf{z}}$  are imposed equal to the degrees of freedom of  $\mathbf{z} := \mathcal{P}_h^\nu(\mathbf{v})$  (introduced in Definition 5.3.2) or are set to zero according to the following criteria:

- The coefficients associated with volume degrees of freedom on  $T \in \mathcal{T}_h$ , for all  $i = 1, \dots, N_b$  are:

$$\bar{Z}_T^i = \begin{cases} 0 & \text{if } h_T^2 \beta_T \geq \alpha_T^\nu, \\ Z_T^i & \text{otherwise.} \end{cases}$$

- If  $f \in \Delta_{n-1}^\circ(\mathcal{T}_h) \cap \Delta_{n-1}(T^+) \cap \Delta_{n-1}(T^-)$  the coefficients associated with  $(n-1)$ -face moments, for all  $i = 1, \dots, N_f$  are:

$$\bar{Z}_f^i = \begin{cases} 0 & \text{if } h_{T^+}^2 \beta_{T^+} \geq \alpha_{T^+}^\nu \text{ or } h_{T^-}^2 \beta_{T^-} \geq \alpha_{T^-}^\nu, \\ Z_f^i & \text{otherwise.} \end{cases}$$

On boundary faces  $f \in \Delta_{n-1}^\partial(\mathcal{T}_h)$  we set  $\bar{\mathbf{z}}_f = \mathbf{z}_f$ .

- If  $e \in \Delta_{n-2}^\circ(\mathcal{T}_h)$ , the coefficients associated with  $(n-2)$ -face moments, for all  $i = 1, \dots, N_e$  are:

$$\bar{Z}_e^i = \begin{cases} 0 & \text{if } \exists T' \in \Delta_n(e) : h_{T'}^2 \beta_{T'} \geq \alpha_{T'}^\nu, \\ Z_e^i & \text{otherwise.} \end{cases}$$

On boundary faces  $e \in \Delta_{n-2}^\partial(\mathcal{T}_h)$  we set  $\bar{\mathbf{z}}_e = \mathbf{z}_e$ .

Analogously to Proposition 5.3.5, uniform stability of the auxiliary space decomposition relies on the approximation properties of the operator  $\mathcal{P}_h^{\nu,\beta}$ . Let  $T \in \mathcal{T}_h$  be a fixed element in the **curl**-dominated regime. We define the following sets (see Figure 5.1 for an example in a 2D schematic representation where an element is depicted as a triangle):

$$\begin{aligned} \Delta_{n-1}^\Gamma(T) &:= \{f \in \Delta_{n-1}(T) \cap \Delta_{n-1}(T') \text{ with } T' \in \Upsilon_h^\Gamma\}, \\ \Delta_{n-2}^\Gamma(T) &:= \{e \in \Delta_{n-2}(T) : \Delta_n(e) \cap \Upsilon_h^\Gamma \neq \emptyset\}, \\ \Delta_n^\Gamma(T) &:= \{T' \in \Upsilon_h^\Gamma : \Delta_{n-2}(T) \cap \Delta_{n-2}(T') \neq \emptyset, \Delta_{n-1}(T) \cap \Delta_{n-1}(T') = \emptyset\}. \end{aligned} \tag{5.55}$$

With this notation in mind, local approximation estimates for the operator  $\mathcal{P}_h^{\nu,\beta}$  can be derived. We refer to Figure 5.2 for a 2D sketch of the mesh configurations covered in Lemmas 5.3.8 and 5.3.9.

**Lemma 5.3.8.** Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^{\nu,\beta} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$  be the operator introduced in Definition 5.3.7. Let  $T \in \mathcal{T}_h$  be fixed. Assume that  $\Delta_{n-2}^\Gamma(T) \neq \emptyset$  but  $\Delta_{n-1}^\Gamma(T) = \emptyset$ . Then,

$$\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq C \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{T' \in \Delta_n^\Gamma(T)} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} h_{T'}^2 \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2, \tag{5.56}$$

where the constants  $C > 0$  depend only on the polynomial degree and the shape regularity of the mesh.

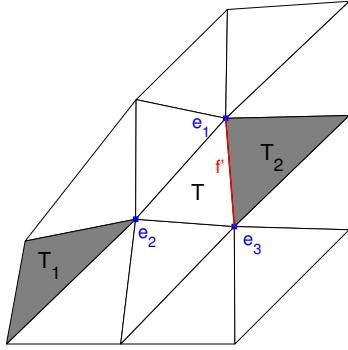


Figure 5.1.: 2D sketch of the sets defined in (5.55). Assume that the elements in the white region belong to  $\Upsilon_h$ , while the elements in gray ( $T_1$  and  $T_2$ ) are contained in  $\Upsilon_h^\Gamma$ . For the element  $T$  (in the white region):  $\Delta_{n-1}^\Gamma(T) = \{f'\}$ ,  $\Delta_{n-2}^\Gamma(T) = \{e_1, e_2, e_3\}$ ,  $\Delta_n^\Gamma(T) = \{T_1\}$ .

*Proof.* Without loss of generality we can assume that  $T \in \Upsilon_h$  is an interior element, for the degrees of freedom on the boundaries of the domain can be bounded exactly as in Lemma 5.3.3. Using the Definition 5.3.7 of the operator  $\mathcal{P}_h^{\nu, \beta}$ , and the representation in terms of degrees of freedom (5.5) together with (2.34), (2.35), results in

$$\begin{aligned} h_T^{-2} \alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu, \beta}(\mathbf{v})\|_{L^2(T)}^2 &\leq Ch_T^{n-2} h_T^{-2} \alpha_T^\nu \left( \sum_{f \in \Delta_{n-1}(T)} \|\mathbf{v}_{f,T} - \bar{\mathbf{z}}_f\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T)} \|\mathbf{v}_{e,T} - \bar{\mathbf{z}}_e\|_{\ell^2}^2 \right) \\ &\leq Ch_T^{n-4} \left( \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 + \sum_{f \in \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T}\|_{\ell^2}^2 \right. \\ &\quad \left. + \sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right). \end{aligned}$$

The degrees of freedom associated with  $(n-1)$ -faces not belonging to the interface can be bounded as in (5.47) in Lemma 5.3.3,

$$\sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 \leq C \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \quad (5.57)$$

The  $(n-2)$ -faces not lying at the interfaces can be treated as in (5.51) from Lemma 5.3.3, so that

$$\sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 \leq C \sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^\Gamma(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_T^\nu h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \quad (5.58)$$

Concerning the interface faces, the degrees of freedom associated with the  $(n-2)$ -faces in  $\Delta_{n-2}^\Gamma(T)$  can be bounded using triangle inequality as follows. Let  $e \in \Delta_{n-2}^\Gamma(T)$  be fixed and let  $\Delta_n(e) = \bigcup_{\ell=0}^{M_e} T_\ell$  be as in (5.49), with  $T_0 := T$ . Let  $T_j \in \mathcal{T}_h$ , with  $2 \leq j \leq M_e - 1$ , be any element in the non-empty set  $\Delta_n^\Gamma(T) \cap \Delta_n(e)$ . Then,

$$\alpha_{T_0}(\nu) \|\mathbf{v}_{e,T_0}\|_{\ell^2}^2 \leq C(j+1) \alpha_{T_0}(\nu) \left( \|\mathbf{v}_{e,T_0} - \mathbf{v}_{e,T_1}\|_{\ell^2}^2 + \sum_{1 \leq \ell < j} \|\mathbf{v}_{e,T_\ell} - \mathbf{v}_{e,T_{\ell+1}}\|_{\ell^2}^2 + \|\mathbf{v}_{e,T_j}\|_{\ell^2}^2 \right),$$

where  $C(j+1) > 0$  can be bounded by a constant which depends on the connectivity  $M_e$  of each  $(n-2)$ -face of  $T$ . Hence, using the jump characterization (2.37) together with (2.34), (2.35) yields

$$\begin{aligned} \sum_{e \in \Delta_{n-2}^\Gamma(T)} \alpha_T^\nu h_f^{3-n} \|\mathbf{v}_{e,T}\|_{\ell^2}^2 &\leq C \sum_{f \in \Delta_{n-1}(T)} \alpha_T^\nu h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + \sum_{T' \in \Delta_n^\Gamma(T)} \alpha_T^\nu h_f^{3-n} h_{T'}^{-1} \|\mathbf{v}\|_{L^2(T')}^2 \\ &\quad + \sum_{e \in \Delta_{n-2}^\Gamma(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \Delta_{n-1}(T)} \alpha_T^\nu h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned} \quad (5.59)$$

The  $L^2$ -norm on the elements  $T' \in \Upsilon_h^\Gamma$  can be gauged using the fact that  $h_{T'}^{-2} \alpha_{T'}^\nu < \beta_{T'}$ , namely

$$\sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}(T)}} h_T^{-2} \alpha_T^\nu \|\mathbf{v}\|_{L^2(T')}^2 \leq \sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}(T)}} \frac{\alpha_T^\nu}{\alpha_{T'}^\nu} h_T^{-2} \alpha_{T'}^\nu \|\mathbf{v}\|_{L^2(T')}^2 \leq \sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}(T)}} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2.$$

Finally, the degrees of freedom corresponding to  $(n-1)$ -faces, where the degrees of freedom of the conforming approximation have been set to zero, are estimated by means of the jump characterization (2.37) and the norm equivalence (2.34), (2.35) as

$$\begin{aligned} \sum_{f \in \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T}\|_{\ell^2}^2 &\leq \sum_{\substack{f \in \Delta_{n-1}^\Gamma(T) \\ f = \partial T \cap \partial T'}} 2\alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{v}_{f,T'}\|_{\ell^2}^2 + 2\alpha_T^\nu \|\mathbf{v}_{f,T'}\|_{\ell^2}^2 \\ &\leq C \sum_{f \in \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}^\Gamma(T)}} \alpha_{T'}^\nu h_f^{3-n} h_{T'}^{-1} \|\mathbf{v}\|_{L^2(T')}^2. \end{aligned} \quad (5.60)$$

Collecting (5.57), (5.58) and (5.59) yields

$$h_T^{-2} \alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq C \sum_{T' \in \Delta_n^\Gamma(T)} \frac{\alpha_T^\nu}{\alpha_{T'}^\nu} \alpha_{T'}^\nu h_{T'}^{-2} \|\mathbf{v}\|_{L^2(T')}^2 + \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_{T'}^\nu h_T^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2.$$

Using the fact that  $h_{T'}^{-2} \alpha_{T'}^\nu < \beta_{T'}$ , for all  $T' \in \Delta_n^\Gamma(T)$ , yields (5.56) and concludes the proof.  $\square$

If  $\Delta_n^\Gamma(T) = \emptyset$ , one can achieve slightly stronger bounds in the approximation error estimates for the operator  $\mathcal{P}_h^{\nu,\beta}$ , as in the following result.

**Lemma 5.3.9.** *Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^{\nu,\beta} : \Lambda_h^{d,1}(\mathcal{T}_h) \rightarrow \Lambda_h^1(\mathcal{T}_h)$  be the averaging operator in Definition 5.3.7. Let  $T \in \Upsilon_h$  be fixed. Assume that  $\Delta_{n-1}^\Gamma(T) \neq \emptyset$ . If  $\Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T) \neq \emptyset$  and  $\Delta_n^\Gamma(T) = \emptyset$ , then*

$$\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq C \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}^\Gamma(T)}} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} h_{T'}^2 \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2. \quad (5.61)$$

If  $\Delta_{n-1}^\Gamma(T) \equiv \Delta_{n-1}(T)$ , then

$$\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq C \alpha_T^\nu \sum_{f \in \Delta_{n-1}(T)} h_f \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{\substack{T' \in \Upsilon_h \\ \partial T \cap \partial T' \in \Delta_{n-1}(T)}} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} h_{T'}^2 \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2, \quad (5.62)$$

where the constants  $C > 0$  depend only on the polynomial degree and the shape regularity of the mesh.

*Proof.* As in the proof of Lemma 5.3.8, we consider a fixed interior element  $T \in \Upsilon_h$ ; the degrees of freedom on the boundaries of the domain can be bounded exactly as in Lemma 5.3.3. Two cases have to be considered: first assume that  $\Delta_{n-1}^\Gamma(T) \neq \emptyset$  and  $\Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T) \neq \emptyset$ ,  $\Delta_n^\Gamma(T) = \emptyset$ . In view of the Definition 5.3.7 of the operator  $\mathcal{P}_h^{\nu,\beta}$ , by the norm equivalence (2.34), (2.35),

$$\begin{aligned} h_T^{-2} \alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 &\leq Ch_T^{n-2} h_T^{-2} \alpha_T^\nu \left( \sum_{f \in \Delta_{n-1}(T)} \|\mathbf{v}_{f,T} - \bar{\mathbf{z}}_f\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T)} \|\mathbf{v}_{e,T} - \bar{\mathbf{z}}_e\|_{\ell^2}^2 \right) \\ &\leq Ch_T^{n-4} \left( \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 + \sum_{f \in \Delta_{n-1}^\Gamma(T)} \alpha_T^\nu \|\mathbf{v}_{f,T}\|_{\ell^2}^2 \right. \\ &\quad \left. + \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^\Gamma(T)} \sum_{e \in \Delta_{n-2}(f)} \alpha_T^\nu \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 + \sum_{f \in \Delta_{n-1}^\Gamma(T)} \sum_{e \in \Delta_{n-2}(f)} \alpha_T^\nu \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right). \end{aligned}$$

We estimate each of the contributions on the right hand side separately. The degrees of freedom associated with  $(n-1)$ -faces of  $T$  not at the interface, can be bounded as in (5.47) from Lemma 5.3.3:

$$\sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^{\Gamma}(T)} \alpha_T^{\nu} \|\mathbf{v}_{f,T} - \mathbf{z}_f\|_{\ell^2}^2 \leq C \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^{\Gamma}(T)} \alpha_T^{\nu} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \quad (5.63)$$

Moreover, the terms associated with the  $(n-1)$ -faces of  $T$  at the interface can be gauged as in (5.60). Analogously, the degrees of freedom corresponding to  $(n-2)$ -faces belonging to  $(n-1)$ -faces at the interface can be bounded as

$$\begin{aligned} \sum_{f \in \Delta_{n-1}^{\Gamma}(T)} \sum_{e \in \Delta_{n-2}(f)} \alpha_e^{\nu} \|\mathbf{v}_{e,T}\|_{\ell^2}^2 &\leq \sum_{\substack{f \in \Delta_{n-1}^{\Gamma}(T) \\ f = \partial T \cap \partial T'}} \sum_{e \in \Delta_{n-2}(f)} 2\alpha_e^{\nu} \|\mathbf{v}_{e,T} - \mathbf{v}_{e,T'}\|_{\ell^2}^2 + 2\alpha_e^{\nu} \|\mathbf{v}_{e,T'}\|_{\ell^2}^2 \\ &\leq C \sum_{f \in \Delta_{n-1}^{\Gamma}(T)} \alpha_f^{\nu} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{\substack{T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \in \Delta_{n-1}^{\Gamma}(T)}} \alpha_{T'}^{\nu} h_{T'}^{-1} \|\mathbf{v}\|_{L^2(T')}^2. \end{aligned} \quad (5.64)$$

Finally, the bound on the remaining  $(n-2)$ -faces can be derived as in (5.51) from Lemma 5.3.3,

$$\begin{aligned} \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^{\Gamma}(T)} \sum_{e \in \Delta_{n-2}(f)} \alpha_e^{\nu} \|\mathbf{v}_{e,T} - \mathbf{z}_e\|_{\ell^2}^2 &\leq C \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^{\Gamma}(T)} \sum_{e \in \Delta_{n-2}(f)} \sum_{f \in \Delta_{n-1}(e)} \alpha_f^{\nu} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\leq C \sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^{\Gamma}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_f^{\nu} h_f^{3-n} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned} \quad (5.65)$$

Combining the estimates (5.63), (5.60), (5.64) and (5.65) results in

$$\begin{aligned} h_T^{-2} \alpha_T^{\nu} \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 &\leq C \sum_{f \in \Delta_{n-1}(T) \setminus \Delta_{n-1}^{\Gamma}(T)} \alpha_f^{\nu} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + \sum_{f \in \Delta_{n-1}^{\Gamma}(T)} \alpha_f^{\nu} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\quad + \sum_{\substack{T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \in \Delta_{n-1}^{\Gamma}(T)}} \alpha_{T'}^{\nu} h_{T'}^{-2} \|\mathbf{v}\|_{L^2(T')}^2 + \sum_{e \in \Delta_{n-2}(T) \setminus \Delta_{n-2}^{\Gamma}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_f^{\nu} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \\ &\leq C \sum_{\substack{T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \in \Delta_{n-1}^{\Gamma}(T)}} \frac{\alpha_T^{\nu}}{\alpha_{T'}^{\nu}} \alpha_{T'}^{\nu} h_{T'}^{-2} \|\mathbf{v}\|_{L^2(T')}^2 + \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} \alpha_f^{\nu} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned}$$

Using the shape regularity of the mesh and the fact that  $h_{T'}^{-2} \alpha_{T'}^{\nu} < \beta_{T'}$  for all  $T' \in \mathcal{T}_h$  with  $\partial T \cap \partial T' \in \Delta_{n-1}^{\Gamma}(T)$ , yields (5.61).

For the case  $\Delta_{n-1}^{\Gamma}(T) \equiv \Delta_{n-1}(T)$ , by Definition 5.3.7 of  $\mathcal{P}_h^{\nu,\beta}$ , one has  $\bar{\mathbf{z}}_e = 0$  for every  $e \in \Delta_{n-2}(T)$ ,  $\bar{\mathbf{z}}_f = 0$  for every  $f \in \Delta_{n-1}(T)$  and  $\bar{\mathbf{z}}_T = \mathbf{z}_T$ . Therefore, the approximation error can be estimated as

$$h_T^{-2} \alpha_T^{\nu} \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq Ch_T^{4-n} \left( \sum_{f \in \Delta_{n-1}(T)} \alpha_f^{\nu} \|\mathbf{v}_{f,T}\|_{\ell^2}^2 + \sum_{e \in \Delta_{n-2}(T)} \alpha_e^{\nu} \|\mathbf{v}_{e,T}\|_{\ell^2}^2 \right).$$

For the  $(n-1)$ -faces terms we use (5.60) whereas the terms on the  $(n-2)$ -faces can be estimated through (5.64) with  $\Delta_{n-1}^{\Gamma}(T) \equiv \Delta_{n-1}(T)$ . This results in

$$h_T^{-2} \alpha_T^{\nu} \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \leq C \sum_{f \in \Delta_{n-1}(T)} \alpha_f^{\nu} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + \sum_{\substack{T' \in \mathcal{T}_h \\ \partial T \cap \partial T' \in \Delta_{n-1}(T)}} \frac{\alpha_T^{\nu}}{\alpha_{T'}^{\nu}} \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2,$$

where we have used again the fact that for all  $T' \in \mathcal{T}_h^{\Gamma}$ , it holds  $h_{T'}^{-2} \alpha_{T'}^{\nu} < \beta_{T'}$ . □

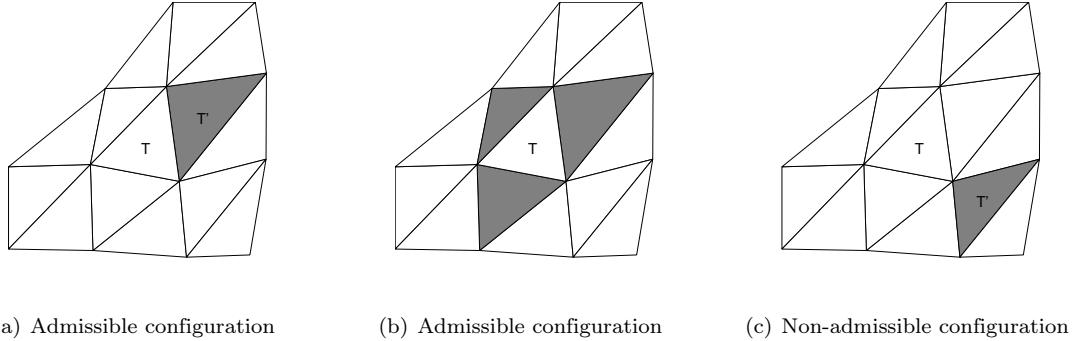


Figure 5.2.: 2D sketch of the required assumptions in Lemma 5.3.9. The elements  $T$  in the white region belong to  $\Upsilon_h$ , while the elements in gray (as  $T'$ ) are contained in  $\Upsilon_h^\Gamma$ . The element  $T \in \Upsilon_h$  in Case (a) satisfies the hypotheses of Lemma 5.3.9 for the bound (5.61), whereas  $T \in \Upsilon_h$  in (b) fulfills the assumptions of Lemma 5.3.9 for the bound (5.62). The configuration in Case (c) does not satisfy the hypotheses of Lemma 5.3.9.

### Stable Decomposition

The stability of the decomposition (5.21) can be established as in the following result.

**Proposition 5.3.10.** *Let  $\mathcal{T}_h$  be a shape regular and locally quasi-uniform partition of  $\Omega$  as in Assumption 2.4.2. Let  $\Lambda_h^1(\mathcal{T}_h) = \Lambda_h^{d,1}(\mathcal{T}_h) \cap H_0(\mathbf{curl}; \Omega)$  and  $s(\cdot, \cdot)$  be the pointwise smoother defined in (5.25). Then, for any  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  there exist  $\mathbf{v}_0 \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and  $\mathbf{z} \in \Lambda_h^1(\mathcal{T}_h)$  such that  $\mathbf{v} = \mathbf{v}_0 + \mathbf{z}$  and*

$$s(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{z}, \mathbf{z}) \leq c_0 \max\{1, \vartheta(\nu)\} a_{DG}(\mathbf{v}, \mathbf{v}), \quad (5.66)$$

with

$$\vartheta(\nu) := \max_{\substack{T \in \Upsilon_h, T' \in \Upsilon_h^\Gamma \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\alpha_T^\nu}{\alpha_{T'}^\nu}, \quad (5.67)$$

where the sets  $\Upsilon_h$  and  $\Upsilon_h^\Gamma$  are defined in (5.15) and the constant  $c_0$  depends only on the polynomial degree and the shape regularity of the mesh.

*Proof.* Let  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$  and let  $\mathcal{P}_h^{\nu, \beta}$  be the averaging operator from Definition 5.3.7. By construction,  $\mathcal{P}_h^{\nu, \beta}(\mathbf{v}) \in \Lambda_h^1(\mathcal{T}_h)$ . Henceforth we set  $\mathbf{v}_0 = \mathbf{v} - \mathcal{P}_h^{\nu, \beta}(\mathbf{v})$ . We distinguish several cases depending on the regions of the domain, and proceed by local estimates. In particular, for a fixed element  $T \in \mathcal{T}_h$ , we distinguish three possible cases:

- (i)  $T \in \Upsilon_h^\Gamma$  i.e.  $h_T^2 \beta_T \geq \alpha_T^\nu \geq \nu_T$  i.e. the element is in the reaction-dominated regime;
- (ii)  $T \in \Upsilon_h$  i.e.  $h_T^2 \beta_T < \alpha_T^\nu$  and  $\Delta_{n-2}^\Gamma(T) = \emptyset$  (“isolated”  $\mathbf{curl}$ -dominated region);
- (iii)  $T \in \Upsilon_h$  i.e.  $h_T^2 \beta_T < \alpha_T^\nu$  and  $\Delta_{n-2}^\Gamma(T) \neq \emptyset$ ;

The last case includes the elements on  $\Upsilon_h$  having a  $(n-2)$ - or a  $(n-1)$ -face at the interface between the reaction-dominated and  $\mathbf{curl}$ -dominated regime. Let us consider one by one the previous cases.

We will typically use

$$a_w(\mathcal{P}_h^{\nu, \beta}(\mathbf{v}), \mathcal{P}_h^{\nu, \beta}(\mathbf{v}))_{|T} \leq a_w(\mathbf{v}, \mathbf{v})_{|T} + a_w(\mathbf{v} - \mathcal{P}_h^{\nu, \beta}(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^{\nu, \beta}(\mathbf{v}))_{|T}, \quad (5.68)$$

and therefore to bound  $a_w(\mathcal{P}_h^{\nu, \beta}(\mathbf{v}), \mathcal{P}_h^{\nu, \beta}(\mathbf{v}))_{|T}$  it will be enough to bound the last term in (5.68).

*Case (i).* In this case,  $T$  is an element in the reaction-dominated region and therefore the auxiliary space solver is “turned off”:  $\mathcal{P}_h^{\nu, \beta}(\mathbf{v})_{|T} = 0$  and hence  $\mathbf{v}_0|_T = \mathbf{v}|_T$ . Since, the correction in the auxiliary space is (locally) left out, in order to prove (5.66) one only needs to consider the smoother. For the pointwise

or block Jacobi smoother, using the scaling in (5.30) together with the fact that  $h_T^{-2}\nu_T \leq h_T^{-2}\alpha_T^\nu \leq \beta_T$ , results in

$$s(\mathbf{v}_0, \mathbf{v}_0)|_T \leq C \left( h_T^{-2}\nu_T \|\mathbf{v}\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}\|_{L^2(T)}^2 + h_T^{-2}\alpha_T^\nu \|\mathbf{v}\|_{L^2(T)}^2 \right) \leq C\beta_T \|\mathbf{v}\|_{L^2(T)}^2.$$

*Case (ii).* On account of the construction of the operator  $\mathcal{P}_h^{\nu,\beta}$  in Definition 5.3.7, it holds  $\mathcal{P}_h^{\nu,\beta}(\mathbf{v})|_T = \mathcal{P}_h^\nu(\mathbf{v})|_T$  and hence  $\mathbf{v}_0|_T = \mathbf{v}|_T - \mathcal{P}_h^\nu(\mathbf{v})|_T$ . Therefore, we can directly argue as in the proof of Proposition 5.3.5. Taking into account (5.68) and using inverse inequalities, the local approximation properties of  $\mathcal{P}_h^\nu$  in (5.41) and the assumption  $h_T^2\beta_T < \alpha_T^\nu$  yields

$$\begin{aligned} a_w(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}))|_T &\leq \nu_T \|\mathbf{curl}_{\mathbf{h}}(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}))\|_{L^2(T)}^2 + \beta_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \\ &\leq 2h_T^{-2}\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \\ &\leq C\alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2. \end{aligned}$$

An analogous reasoning applies to the smoothing operator. For the pointwise or block Jacobi smoother, the scaling in (5.30) together with (5.41), gives

$$s(\mathbf{v}_0, \mathbf{v}_0)|_T \leq C \left( 2h_T^{-2}\nu_T \|\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})\|_{L^2(T)}^2 \right) \leq C\alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2.$$

*Case (iii).* The last case covers the configuration where the current element is in the **curl**-dominated regime but shares at least one  $(n-2)$ -face with an element in the reaction-dominated region. Under these assumptions, the inverse inequality together with the estimate (5.56) from Lemma 5.3.8 gives

$$\begin{aligned} a_w(\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v}))|_T &\leq \nu_T \|\mathbf{curl}_{\mathbf{h}}(\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v}))\|_{L^2(T)}^2 + \beta_T \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \\ &\leq 2h_T^{-2}\nu_T \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \\ &\leq C\alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 + C \sum_{T' \in \Delta_n^\Gamma(T)} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} \beta_{T'} \|\mathbf{v}\|_{L^2(T')}^2. \end{aligned}$$

Moreover, the same bound can be derived for the pointwise Jacobi operator introduced in (5.31) (or (5.32)) by using the scaling (5.30) and applying estimate (5.56),

$$\begin{aligned} s(\mathbf{v}_0, \mathbf{v}_0)|_T &\leq C \left( h_T^{-2}\nu_T \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 + \beta_T \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 + h_T^{-2}\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2 \right) \\ &\leq Ch_T^{-2}\alpha_T^\nu \|\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{L^2(T)}^2, \end{aligned} \quad (5.69)$$

since by assumption the current element  $T$  satisfies  $h_T^2\beta_T < \alpha_T^\nu$ . Collecting the local contributions from all the discussed cases, results in

$$s(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{z}, \mathbf{z}) \leq \|\mathbf{curl}_{\mathbf{h}}\mathbf{v}\|_{0,\nu,\mathcal{T}_h}^2 + C|\mathbf{v}|_{*,\nu}^2 + C \max \left\{ 1, \max_{T \in \mathcal{T}_h, T' \in \mathcal{T}'_h} \frac{\alpha_{T'}^\nu}{\alpha_{T'}^\nu} \right\} \|\mathbf{v}\|_{0,\beta,\mathcal{T}_h}^2,$$

for all  $\mathbf{v} \in \Lambda_h^{d,1}(\mathcal{T}_h)$ ,  $\mathbf{v} = \mathbf{v}_0 + \mathbf{z}$  and  $\mathbf{v}_0 \in \Lambda_h^{d,1}(\mathcal{T}_h)$ ,  $\mathbf{z} \in \Lambda_h^1(\mathcal{T}_h)$ . Coercivity of  $a_{DG}(\cdot, \cdot)$  yields the conclusion.  $\square$

**Remark 5.3.11.** The result presented in Theorem 5.3.1 encompasses the following coefficients distributions:

- There exist  $\beta_{\max}$  and  $\nu_{\min}$  such that

$$\nu_T \geq \nu_{\min} > 0, \quad \text{and} \quad 0 < \beta_T \leq \beta_{\max}, \quad \forall T \in \mathcal{T}_h.$$

- $\nu > 0$  is arbitrary and there exists  $B > 1$  such that

$$B^{-1} \leq \frac{\beta_T}{\beta_{T'}} \leq B, \quad \forall T, T' \in \mathcal{T}_h, \quad \text{with} \quad \partial T \cap \partial T' \neq \emptyset.$$

- $\beta > 0$  is arbitrary and there exists  $\Lambda > 1$  such that

$$\Lambda^{-1} \leq \frac{\alpha_T^\nu}{\alpha_{T'}^\nu} \leq \Lambda, \quad \forall T \in \Upsilon_h, T' \in \Upsilon_h^\Gamma, \quad \text{with } \partial T \cap \partial T' \neq \emptyset.$$

The only left out case is when there exist  $T \in \Upsilon_h$  and  $T' \in \Upsilon_h$  such that simultaneously it holds

$$\begin{aligned} \Delta_{n-1}(T) \cap \Delta_{n-1}(T') &= \emptyset, & \Delta_{n-2}(T) \cap \Delta_{n-2}(T') &\neq \emptyset, \\ \nu_T &\nearrow \infty & \text{and} & \beta_T \searrow 0, \\ \nu_{T'} &\searrow 0 & \text{and} & \beta_{T'} \nearrow \infty. \end{aligned}$$

In this case, since all elements sharing a  $(n-2)$ -face contribute to the construction of a conforming approximation of a given function in the DG space  $\Lambda_h^{d,1}(\Upsilon_h)$ , it is not possible, with the techniques presented here, to bound simultaneously the  $L^2$ -norm of the approximation error weighted by the coefficients  $\nu$  and  $\beta$ .

**Remark 5.3.12.** The analysis presented in Section 5.3.2 could be carried out analogously by considering the local ratio  $h_T^2 \beta_T / \nu_T$  instead of  $h_T^2 \beta_T / \alpha_T^\nu$ . The decomposition of  $\mathbf{v} \in \Lambda_h^{d,1}(\Upsilon_h)$ , corresponding to (5.66) would yield a bound of the form

$$s(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{z}, \mathbf{z}) \leq c_0^2 \max \left\{ 1, \max_{\substack{T \in \Upsilon_h, T' \in \Upsilon_h' \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\alpha_T^\nu}{\nu_{T'}}, \max_{T \in \Upsilon_h^\Gamma} \frac{\alpha_T^\nu}{\nu_T} \right\} a_{\text{DG}}(\mathbf{v}, \mathbf{v}),$$

with  $\mathbf{v} = \mathbf{v}_0 + \mathbf{z}$ ,  $\mathbf{v}_0 \in \Lambda_h^{d,1}(\Upsilon_h)$ , and  $\mathbf{z} \in \Lambda_h^1(\Upsilon_h)$ . Here the uniform bound required on the ratio  $\alpha_T^\nu / \nu_{T'}$  implies that the coefficients  $\nu_{T_i}$  for elements  $T_i$  in a **curl**-dominated region and belonging to the neighborhood of  $T$  cannot be arbitrarily small within the patch. The choice of dealing with the less natural ratio  $h_T^2 \beta_T / \alpha_T^\nu$  as in (5.66) is aimed at avoiding this shortcoming.

### 5.3.3. Coarser Auxiliary Space

On a simplicial mesh  $\Upsilon_h$ , a DG discretization based on the local space  $\zeta_r(T) = \mathcal{P}_r \Lambda^1(T)$ ,  $r \geq 1$ , can be combined to the  $H_0(\mathbf{curl}, \Omega)$ -conforming finite element space associated with the local space  $\mathcal{P}_r^- \Lambda^1(T)$  (reproducing Case (b) in Section 5.2.2). Using an overlapping additive smoother of the type (5.29), the stable decomposition property (F2) in Theorem 5.2.2 is fulfilled and the bounds in Propositions 5.3.5 and 5.3.10 apply verbatim.

**Proposition 5.3.13.** Let  $\Upsilon_h$  be a shape regular and quasi-uniform simplicial mesh as in Assumption 2.4.2 (i). Let  $s_O(\cdot, \cdot)$  be an overlapping smoother as in (5.29). Then, for every  $\mathbf{v} \in \mathcal{P}_r^d \Lambda^1(\Upsilon_h)$  there exist  $\mathbf{v}_0 \in \mathcal{P}_r^d \Lambda^1(\Upsilon_h)$  and  $\mathbf{w} \in \mathcal{P}_r^- \Lambda^1(\Upsilon_h)$  such that  $\mathbf{v} = \mathbf{v}_0 + \mathbf{w}$  and

$$s_O(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{w}, \mathbf{w}) \leq \tilde{c}_a \max\{1, \delta(\nu, \beta)\} a_{\text{DG}}(\mathbf{v}, \mathbf{v}), \quad (5.70)$$

where  $\delta(\nu, \beta)$  is defined as in (5.35), namely

$$\delta(\nu, \beta) := \min \left\{ \max_{\substack{T \in \Upsilon_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{h_T^2 \beta_T}{\nu_T}, \max_{\substack{T, T' \in \Upsilon_h \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\beta_T}{\beta_{T'}}, \max_{\substack{T \in \Upsilon_h, T' \in \Upsilon_h' \\ \partial T \cap \partial T' \neq \emptyset}} \frac{\alpha_T^\nu}{\alpha_{T'}^\nu} \right\},$$

and the constant  $\tilde{c}_a > 0$  depends only on the polynomial degree and the shape regularity of the mesh.

*Proof.* Let  $\mathbf{z} := \mathcal{P}_h^\nu(\mathbf{v}) \in \mathcal{P}_r \Lambda^1(\Upsilon_h)$  and  $\bar{\mathbf{z}} := \mathcal{P}_h^{\nu, \beta}(\mathbf{v}) \in \mathcal{P}_r^- \Lambda^1(\Upsilon_h)$  where  $\mathcal{P}_h^\nu$  and  $\mathcal{P}_h^{\nu, \beta}$  are the averaging operators from Definitions 5.3.2 and 5.3.7, respectively. We consider simultaneously two splittings of  $\mathbf{v} \in \mathcal{P}_r^d \Lambda^1(\Upsilon_h)$ . Let  $\Pi_h^{1,-} : \mathcal{P}_r \Lambda^1(\Upsilon_h) \rightarrow \mathcal{P}_r^- \Lambda^1(\Upsilon_h)$  be the projection operator associated with the second family of  $H(\mathbf{curl}, \Omega)$ -conforming finite element (*cf.* Section 2.4.1 and [Néd86]). We define  $\mathbf{w} := \Pi_h^{1,-}(\mathbf{z})$  and  $\bar{\mathbf{w}} := \Pi_h^{1,-}(\bar{\mathbf{z}})$ . To bound the  $H(\mathbf{curl}, \Omega)$ -conforming part of the decompositions,

local approximation estimates for the interpolation operator  $\Pi_h^{1,-}$  together with the estimates from Propositions 5.3.5 and 5.3.10, give

$$\begin{aligned} a_w(\mathbf{w}, \mathbf{w}) &\leq \|\mathbf{curl}(\Pi_h^{1,-}(\mathbf{z}))\|_{0,\nu,\mathcal{T}_h}^2 + \|\Pi_h^{1,-}(\mathbf{z})\|_{0,\beta,\mathcal{T}_h}^2 \leq C (\|\mathbf{curl}(\mathcal{P}_h^\nu(\mathbf{v}))\|_{0,\nu,\mathcal{T}_h}^2 + \|\mathcal{P}_h^\nu(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2) \\ &\leq C \max\{1, \theta(\nu, \beta)\} \|\mathbf{v}\|_{\text{DG}}^2, \\ a_w(\overline{\mathbf{w}}, \overline{\mathbf{w}}) &\leq C \|\mathbf{curl}(\mathcal{P}_h^{\nu,\beta}(\mathbf{v}))\|_{0,\nu,\mathcal{T}_h}^2 + C \|\mathcal{P}_h^{\nu,\beta}(\mathbf{v})\|_{0,\beta,\mathcal{T}_h}^2 \leq C \max\{1, \vartheta(\nu)\} \|\mathbf{v}\|_{\text{DG}}^2, \end{aligned}$$

with  $\theta$  and  $\vartheta$  as in (5.43) and (5.67), respectively. Concerning the patch smoother, we set  $\mathbf{v}_0 = (\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})) + (\mathbf{z} - \Pi_h^{1,-}(\mathbf{z}))$  and similarly  $\overline{\mathbf{v}}_0 = (\mathbf{v} - \mathcal{P}_h^{\nu,\beta}(\mathbf{v})) + (\overline{\mathbf{z}} - \Pi_h^{1,-}(\overline{\mathbf{z}}))$ . On account of Lemma 2.4.1,  $\mathbf{curl}(\mathcal{P}_r^- \Lambda^1(T)) = \mathbf{curl}(\mathcal{P}_r \Lambda^1(T))$ , and therefore the difference  $\mathbf{z} - \Pi_h^{1,-}(\mathbf{z})$  (resp.  $\overline{\mathbf{z}} - \Pi_h^{1,-}(\overline{\mathbf{z}})$ ) is  $\mathbf{curl}$ -free. Hence, we can rely on the following  $L^2$ -orthogonal discrete Helmholtz-Hodge decompositions (see [AFW06, Section 5.6]),

$$\mathbf{z} = \Pi_h^{1,-}(\mathbf{z}) + \nabla q, \quad \overline{\mathbf{z}} = \Pi_h^{1,-}(\overline{\mathbf{z}}) + \nabla \bar{q}, \quad q, \bar{q} \in H^1(\Omega) \cap \bigoplus_{T \in \mathcal{T}_h} \mathcal{P}_{r+1}(T). \quad (5.71)$$

In view of the decompositions (5.71), it holds

$$s_O(\mathbf{v}_0, \mathbf{v}_0) \leq s_O(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})) + s_O(\nabla q, \nabla q). \quad (5.72)$$

To bound the first term, the scaling of the smoothing operator (5.34) in Lemma 5.2.5 together with the estimates on the approximation error for  $\mathcal{P}_h^\nu$  in Lemma 5.3.3 and Corollary 5.3.4, gives

$$\begin{aligned} s_O(\mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v}), \mathbf{v} - \mathcal{P}_h^\nu(\mathbf{v})) &\simeq \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \nu_T \|\mathbf{curl} \mathbf{h}(\mathbf{v}_j - \mathcal{P}_h^\nu(\mathbf{v}_j))\|_{L^2(T)}^2 + \beta_T \|\mathbf{v}_j - \mathcal{P}_h^\nu(\mathbf{v}_j)\|_{L^2(T)}^2 \right. \\ &\quad + \sum_{T \in \Omega_j} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \setminus \partial \Omega_j} h_f^{-1} \|[\mathbf{v}_j]_{f,\tau}\|_{L^2(f)}^2 \\ &\quad \left. + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \cap \partial \Omega_j} h_f^{-1} \|\mathbf{n}_f \times (\mathbf{v}_j - \mathcal{P}_h^\nu(\mathbf{v}_j))\|_{L^2(f)}^2 \right) \\ &\leq C \max\{1, \theta(\nu, \beta)\} \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \beta_T \|\mathbf{v}_j\|_{L^2(T)}^2 + \sum_{T \in \Omega_j} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}_j]_{f,\tau}\|_{L^2(f)}^2 \right), \end{aligned}$$

where  $\mathbf{v}_j = \theta_j \mathbf{v}$  and  $\{\theta_j\}_{j=1}^J$  is a partition of unity relative to the decomposition  $\{\Omega_j\}_{j=1}^J$ .

As regards the second term in (5.72), we write  $\nabla q_j = \mathbf{z}_j - \Pi_h^{1,-}(\mathbf{z}_j) = \mathcal{P}_h^\nu(\mathbf{v}_j) - \Pi_h^{1,-}(\mathcal{P}_h^\nu(\mathbf{v}_j))$  for  $j = 1, \dots, J$ . Then, the scaling (5.34) together with  $\nabla q \in \mathcal{N}(\mathbf{curl})$ , inverse and trace inequalities and the local error estimates for  $\mathbf{z}_j - \Pi_h^{1,-}(\mathbf{z}_j)$  (see [HT00], [TW05, Lemma 10.4 and Lemma 10.8]) yield

$$\begin{aligned} s_O(\nabla q, \nabla q) &\simeq \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \beta_T \|\nabla q_j\|_{L^2(T)}^2 + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e) \cap \partial \Omega_j} h_f^{-1} \|\mathbf{n}_f \times \nabla q_j\|_{L^2(f)}^2 \right) \\ &\leq C \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \beta_T \|\mathbf{z}_j - \Pi_h^{1,-}(\mathbf{z}_j)\|_{L^2(T)}^2 + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu h_T^{-2} \|\mathbf{z}_j - \Pi_h^{1,-}(\mathbf{z}_j)\|_{L^2(T)}^2 \right) \\ &\leq C \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \beta_T \|\mathcal{P}_h^\nu(\mathbf{v}_j)\|_{L^2(T)}^2 + \sum_{\substack{T \in \Omega_j \\ \partial T \cap \partial \Omega_j \neq \emptyset}} \alpha_T^\nu h_T^{-2} \|\mathcal{P}_h^\nu(\mathbf{v}_j)\|_{L^2(T)}^2 \right) \\ &\leq C \max\{1, \theta(\nu, \beta)\} \sum_{j=1}^J \left( \sum_{T \in \Omega_j} \beta_T \|\mathbf{v}_j\|_{L^2(T)}^2 + \sum_{T \in \Omega_j} \alpha_T^\nu \sum_{e \in \Delta_{n-2}(T)} \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|[\mathbf{v}_j]_{f,\tau}\|_{L^2(f)}^2 \right), \end{aligned}$$

where the last inequality follows by Lemma 5.3.3 and Corollary 5.3.4. To “glue” together the contributions from the different patches, we use the properties of  $\{\theta_j\}_{j=1}^J$ , namely (see [TW05, Chapter 10])

$$\begin{aligned}\|\mathbf{v}_j\|_{L^2(T)}^2 &= \|\theta_j \mathbf{v}\|_{L^2(T)}^2 \leq \|\theta_j\|_{L^\infty(\Omega)}^2 \|\mathbf{v}\|_{L^2(T)}^2 \leq C \|\mathbf{v}\|_{L^2(T)}^2, & \forall T \in \Omega_j, \\ \|[\mathbf{v}_j]_{f,\tau}\|_{L^2(f)}^2 &= \|[\theta_j \mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \leq \|\theta_j\|_{L^\infty(\Omega)}^2 \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2 \leq C \|[\mathbf{v}]_{f,\tau}\|_{L^2(f)}^2, & \forall f \in \mathcal{F}_h \cap \Omega_j.\end{aligned}$$

To bound  $s_O(\bar{v}_0, \bar{v}_0)$  (second splitting), the proof proceeds mutatis mutandis but using the estimates from Proposition 5.3.10, so that  $s_O(\bar{v}_0, \bar{v}_0) \leq C \max\{1, \vartheta(\nu)\} \|\mathbf{v}\|_{\text{DG}}^2$ .

Collecting all estimates and taking the minimum over the two splittings the conclusion (5.70) follows.  $\square$

**Remark 5.3.14.** On hexahedral meshes, the stability results of Proposition 5.3.13 do not hold true. Indeed, as pointed out in [BP06, Remark 7.14], and observed numerically in [BHP07, Section 5] and [GSS08, Section 6.2], the local full polynomial space  $\mathcal{Q}_r \Lambda^1(T)$  yields a discretization which triggers spurious modes. This is confirmed by the numerical experiments in Section 5.4.4 (see in particular Figures 5.12 and 5.13). Therefore, using a spectrally correct auxiliary space in the preconditioner for the DG discretization based on  $\mathcal{Q}_r \Lambda^1(T)$  does not seem to result in a convergent solver, independently of the type of the smoother. However, if the auxiliary space consists of full polynomials, Theorem 5.3.1 applies and the resulting preconditioner is uniform (see Table 5.6 “Q1-Q1 Jacobi”). Similarly, for the choice  $\Lambda_h^{d,1}(\mathcal{T}_h) = \mathcal{Q}_r^{d,-} \Lambda^1(\mathcal{T}_h)$ ,  $r \geq 1$ , together with the auxiliary space  $\Lambda_h^1(\mathcal{T}_h) = \mathcal{Q}_r^- \Lambda^1(\mathcal{T}_h)$ , the results of Theorem 5.3.1 carry over.

On the other hand, one might consider as local space in  $\Lambda_h^{d,1}(\mathcal{T}_h)$  the spectrally correct element  $\mathcal{S}_r \Lambda^k(\mathcal{T}_h)$  in (2.22), and use as auxiliary space  $\mathcal{Q}_r^- \Lambda^1(\mathcal{T}_h)$ ,  $r \geq 1$ . This combination together with an overlapping smoother provides, in the lowest order case and in dimension two (see Section 5.4.4), a uniform preconditioner (for continuous coefficients). However, the proof of Proposition 5.3.13 does not apply to this element (not even when  $r = 1$ ), since a connection between the two mentioned families through a discrete Helmholtz decomposition (in the spirit of (5.71)) does not seem to extend straightforwardly. For polynomial degree  $r > 1$ , the lack of a clear relation or inclusion between the two families hinders the construction of an auxiliary space preconditioner.

Similarly to [Zik08, Theorem 4.5], we can construct an example to show that the use of an overlapping smoother in Proposition 5.3.13 is indeed necessary.

**Lemma 5.3.15.** *Let  $\mathcal{T}_h$  be a mesh of simplices, shape regular and quasi-uniform as in Assumption 2.4.2 (i). Let  $s(\cdot, \cdot)$  be a pointwise smoother as in (5.31) or in (5.32). Then, there exist  $\mathbf{v} \in \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ ,  $\mathbf{w} \in \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$  and  $\mathbf{v}_0 \in \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$  such that  $\mathbf{v} = \mathbf{v}_0 + \mathbf{w}$  and*

$$s(\mathbf{v}_0, \mathbf{v}_0) + a_w(\mathbf{w}, \mathbf{w}) \leq C \max \left\{ 1, \max_{T \in \mathcal{T}_h} \frac{\alpha_T^\nu h_T^{-2}}{\beta_T} \right\} a_{\text{DG}}(\mathbf{v}, \mathbf{v}),$$

where  $C > 0$  depends on the shape regularity of the mesh and on the polynomial degree.

**Remark 5.3.16.** As a consequence, under the assumptions of Lemma 5.3.15, the spectral condition number of the preconditioned system would depend on the mesh size and the problem coefficients, except in the reaction-dominated regime.

*proof of Lemma 5.3.15.* The proof is constructive. Let  $T \in \mathcal{T}_h$  be a fixed  $n$ -simplex with barycentric coordinates  $\{\lambda_i\}_{i=0}^n$ . Let  $e = e_{i,j} \in \Delta_1(T)$ ,  $i, j \in \{0, n\} \cap \mathbb{N}$ , be a fixed 1-simplex (edge) of  $T$  such that  $\{a_i, a_j\} \in \Delta_0(e)$  are its endpoints. Let  $\mathbf{b}_e$  be a basis function of  $\mathcal{P}_r \Lambda^1(T)$ , associated with  $e$ , of the form  $\nabla(3\lambda_i \lambda_j)$ , i.e., the gradient of a quadratic edge bubble. Note that  $\mathbf{b}_e \in \mathcal{P}_r \Lambda^1(T) \setminus \mathcal{P}_r^- \Lambda^1(T)$  and can be extended to a global function equal to zero outside of its support (the “macroelement” consisting of the union of the tetrahedra sharing the edge  $e$ ). Then  $\mathbf{b}_e \in \mathcal{P}_r \Lambda^1(\mathcal{T}_h)$  but  $\mathbf{b}_e \notin \mathcal{P}_r^- \Lambda^1(\mathcal{T}_h)$ . Arguing as in the proof of Proposition 5.3.13, let  $\mathbf{z} := \mathcal{P}_h^\nu(\mathbf{b}_e) \equiv \mathbf{b}_e \in \mathcal{P}_r \Lambda^1(\mathcal{T}_h)$  and  $\mathbf{w} := \Pi_h^{1,-}(\mathbf{b}_e) \equiv \mathbf{0}$ . Therefore,

$a_w(\Pi_h^{1,-}(\mathbf{b}_e), \Pi_h^{1,-}(\mathbf{b}_e)) = 0$ , and taking into account the definition of the pointwise smoother (see the proof of the scaling (5.30) given in Lemma 5.2.4), it holds

$$\begin{aligned} s(\mathbf{b}_e, \mathbf{b}_e) &\lesssim \sum_{T \in \mathcal{T}_h} \beta_T \|\mathbf{b}_e\|_{L^2(T)}^2 + \sum_{T \in \mathcal{T}_h} \alpha_T^\nu \sum_{f \in \Delta_{n-1}(e)} h_f^{-1} \|\mathbf{n} \times \mathbf{b}_e\|_{L^2(f)}^2 \\ &\lesssim \sum_{T \in \Delta_n(e)} \beta_T \|\mathbf{b}_e\|_{L^2(T)}^2 + \sum_{T \in \Delta_n(e)} \alpha_T^\nu h_T^{-2} \|\mathbf{b}_e\|_{L^2(T)}^2. \end{aligned}$$

On the other hand, since  $\mathbf{b}_e \in \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h) \cap H_0(\mathbf{curl}, \Omega)$ , its tangential jump  $[\mathbf{b}_e]_{f,\tau}$  vanishes across the mesh faces. Moreover,  $\mathbf{b}_e \in \mathcal{N}(\mathbf{curl})$  by construction, hence

$$a_{DG}(\mathbf{b}_e, \mathbf{b}_e) \simeq \sum_{T \in \Delta_n(e)} \beta_T \|\mathbf{b}_e\|_{L^2(T)}^2.$$

Therefore, for  $\mathbf{v}_0 := \mathbf{b}_e \in \mathcal{P}_r^d \Lambda^1(\mathcal{T}_h)$ , the splitting  $\mathbf{v} = \mathbf{b}_e + \Pi_h^{1,-}(\mathbf{b}_e)$  gives

$$s(\mathbf{b}_e, \mathbf{b}_e) + a_w(\Pi_h^{1,-}(\mathbf{b}_e), \Pi_h^{1,-}(\mathbf{b}_e)) = s(\mathbf{b}_e, \mathbf{b}_e) \leq C \max \left\{ 1, \max_{T \in \Delta_n(e)} \frac{\alpha_T^\nu h_T^{-2}}{\beta_T} \right\} a_{DG}(\mathbf{b}_e, \mathbf{b}_e),$$

where  $C$  depends only on the shape regularity of the mesh and on the polynomial degree  $r \geq 1$ . The last inequality shows that, unless the problem is reaction-dominated (i.e.  $\beta_T \geq h_T^{-2} \alpha_T^\nu$  for all  $T \in \Delta_n(e)$ ), the smoother would not be effective in damping the function  $\mathbf{b}_e$  (not seen by the auxiliary space) and the spectral condition number of the preconditioned system would show dependence on the mesh size, deteriorating as the mesh is refined.  $\square$

## 5.4. Numerical Experiments in 2D

In the following numerical simulations, we restrict to the two-dimensional problem (5.1) in the unit square  $\Omega = [0, 1]^2$ . The numerical effort required for the validation of the theoretical results on very fine meshes, where the asymptotical behavior of the proposed preconditioner emerges, deterred us from dealing with the three-dimensional case. The two-dimensional operators are defined as the scalar curl  $\operatorname{curl} \mathbf{v} = \partial v_2 / \partial x_1 - \partial v_1 / \partial x_2$  for all  $\mathbf{v} = (v_1, v_2) \in H(\mathbf{curl}, \Omega)$  and the vectorial curl,  $\operatorname{curl}_{2D} \phi = \nabla^\perp \phi := (\partial \phi / \partial x_2, -\partial \phi / \partial x_1)^\top$  for all  $\phi \in H(\mathbf{curl}, \Omega)$ . Throughout this section, the constant penalty parameter  $c_0$  in (5.8) entering the IP-DG discretization is set to  $c_0 = 10$ .

Concerning the solver, we will compare the performances of the unpreconditioned conjugate gradient (CG) and preconditioned conjugate gradient (PCG) algorithms [GV96, Section 10.2 and Section 10.3] with zero vector as initial guess. For each numerical experiment, we report the number of iterations required by the (P)CG algorithm to achieve convergence with a tolerance of  $10^{-7}$ . The spectral condition number of the unpreconditioned matrix is derived from the eigenvalues computed with the MATLAB built-in routine `eig`. This is also used for the preconditioned system whenever the matrix size allows it, namely in the numerical experiments on non-uniform and locally refined triangular meshes in Section 5.4.1 and for (5.73). On fine meshes, a Lanczos procedure within the PCG routine (see [GV96, Chapter 9 and Section 10.2.5]) is used to compute the extremal eigenvalues of the preconditioned system with a control on the quality of the approximation of the eigenvalues up to a tolerance of  $10^{-10}$ . To validate the theoretical results, we consider numerical experiments on both structured, locally refined and quasi-uniform triangular meshes, with continuous (the first three test cases) and strongly varying discontinuous coefficients  $\nu$  and  $\beta$  (the last three set of experiments). Finally, we present some numerical results and considerations for the case of tensor product meshes along the lines of Remark 5.3.14.

Since, in two dimensions, the space  $H(\mathbf{curl}, \Omega)$  is isomorphic to  $H(\operatorname{div}, \Omega)$  through a  $\pi/2$  rotation, we derive  $H(\mathbf{curl}, \Omega)$ -conforming finite element spaces from  $H(\operatorname{div}, \Omega)$ -conforming spaces by means of a vector rotation: *rotated* Raviart–Thomas (RT) finite element space  $\mathcal{RT}_k$  and *rotated* Brezzi–Douglas–Marini (BDM) element space  $\mathcal{BDM}_k$ . With a small abuse of notation and because of space constraints, in some of the plots and tables the term “rotated” is omitted.

### 5.4.1. Triangular Meshes. Constant Coefficients.

#### Structured Triangular Mesh. Constant Coefficients.

As first test case, the problem with constant coefficients  $\beta = \nu = 1$  is considered on a uniform structured triangular mesh. The IP-DG discretization is based on the full polynomial space  $\mathcal{P}_1^d \Lambda^1(\mathcal{T}_h)$  and is preconditioned with lowest order rotated BDM elements (5.4) for the auxiliary space and Jacobi pointwise smoother (5.31). As shown in Figure 5.3, the spectral condition number of the preconditioned matrix is independent of the mesh width (see also Table 5.1 for the number of iterations required for convergence).

In the same graphic and table are reported the results obtained with the preconditioner of type (b) (see Section 5.2) based on the lowest order rotated Raviart–Thomas elements in the discretization of the auxiliary space combined with different pointwise and patch smoothers. Note that only in the case of a block relaxation (overlapping additive Schwarz), the condition number is independent of the mesh width. The non-efficacy of pointwise (Jacobi or block Jacobi) smoothers is in agreement with Remark 5.2.3 and Lemma 5.3.15. The performances of the different smoothers can be also compared in Table 5.1.

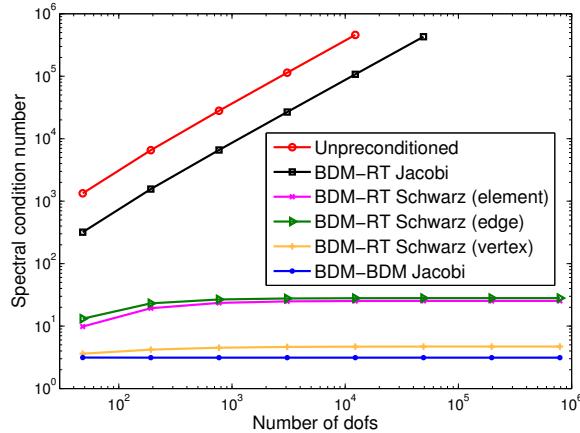


Figure 5.3.: Spectral condition number vs. number of dofs. The DG discretization is based on the full polynomial space (lowest order rotated BDM elements) and lowest order rotated RT elements for the auxiliary space (BDM-RT) or lowest order rotated BDM elements (BDM-BDM). Different choices of the smoother are considered.

| $\#\mathcal{T}_h$        | $2^3$ | $2^5$ | $2^7$ | $2^9$ | $2^{11}$ | $2^{13}$ | $2^{15}$ | $2^{17}$ |
|--------------------------|-------|-------|-------|-------|----------|----------|----------|----------|
| BDM Unpreconditioned     | 45    | 190   | 434   | 889   | 1784     | —        | —        | —        |
| BDM-RT Jacobi            | 36    | 73    | 90    | 90    | 93       | 94       | 93       | 90       |
| BDM-RT block Jacobi      | 32    | 67    | 78    | 82    | 85       | 86       | 77       | 74       |
| BDM-RT Schwarz (element) | 15    | 26    | 29    | 28    | 27       | 26       | 26       | 24       |
| BDM-RT Schwarz (edge)    | 20    | 32    | 34    | 34    | 32       | 31       | 29       | 28       |
| BDM-RT Schwarz (vertex)  | 12    | 15    | 16    | 17    | 17       | 16       | 16       | 16       |
| BDM-BDM Jacobi           | 11    | 12    | 12    | 11    | 11       | 11       | 10       | 10       |

Table 5.1.: Number of iterations for decreasing mesh width. Cases as in Figure 5.3.

#### Quasi-uniform Triangular Mesh. Constant Coefficients.

For coefficients  $\beta = \nu = 1$ , the spectral condition number of the preconditioned system, with an auxiliary space based on lowest order rotated BDM elements and pointwise relaxation, prove independent of the mesh width even on quasi-uniform triangular meshes (see Figure 5.4).

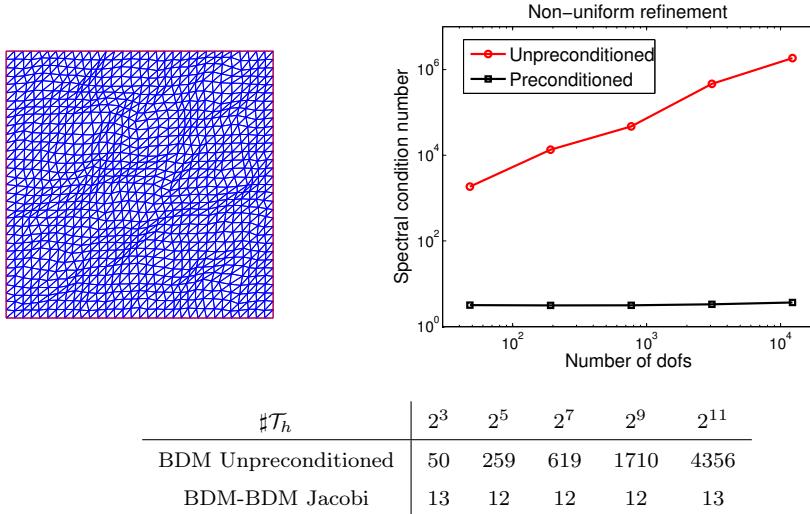


Figure 5.4.: Condition number vs. number of dofs for quasi-uniform meshes. In the table, number of iterations for decreasing mesh width. Numerical discretization based on full polynomial DG spaces and second family edge elements (rotated  $\mathcal{BDM}_1$ ) for the auxiliary space. Pointwise Jacobi smoother.

#### Triangular Mesh, Local Refinement. Constant Coefficients.

Under the same discretization as in the previous test case, with  $\beta = \nu = 1$ , the preconditioner proposed in Section 5.2.2 Case (a), prove competitive for locally refined meshes, see Figure 5.5 and Table 5.2. In particular, we consider three different test cases which often occur in applications, namely a local refinement towards a corner of the domain, towards a boundary side and towards a point/region inside the domain  $\Omega$ . The refinement strategy is not driven by any error estimator. As can be easily observed, in all cases, the convergence of the preconditioner (measured from the spectral condition number or number of iterates) is uniform with respect to the mesh size.

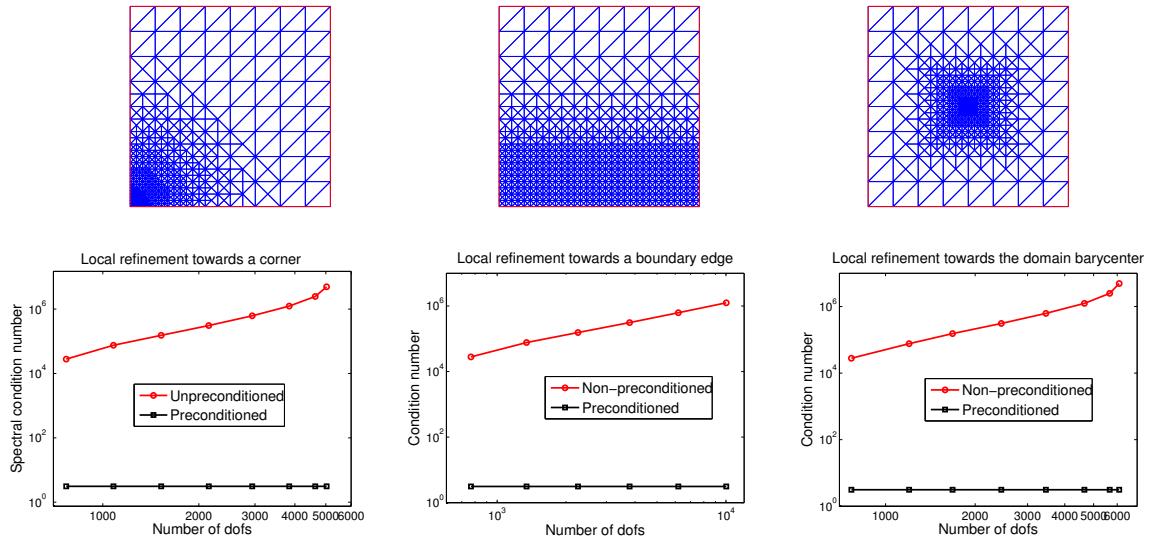


Figure 5.5.: Three different local refinement strategies. Condition number vs. number of dofs for a numerical discretization based on full polynomial DG spaces and second family edge elements (rotated  $\mathcal{BDM}_0$ ) for the auxiliary space. Pointwise Jacobi smoother.

| Case 1, in Figure 5.5 |     |     |      |      |      |      |      |      |
|-----------------------|-----|-----|------|------|------|------|------|------|
| $\#\mathcal{T}_h$     | 128 | 180 | 254  | 358  | 490  | 640  | 772  | 838  |
| BDM Unpreconditioned  | 434 | 705 | 1065 | 1488 | 2171 | 3073 | 4439 | 6024 |
| BDM-BDM Jacobi        | 12  | 12  | 12   | 12   | 12   | 12   | 12   | 12   |
| Case 2, in Figure 5.5 |     |     |      |      |      |      |      |      |
| $\#\mathcal{T}_h$     | 128 | 224 | 376  | 632  | 1032 | 1672 |      |      |
| BDM Unpreconditioned  | 434 | 678 | 1042 | 1445 | 2123 | 3090 |      |      |
| BDM-BDM Jacobi        | 12  | 12  | 12   | 12   | 12   | 12   |      |      |
| Case 3, in Figure 5.5 |     |     |      |      |      |      |      |      |
| $\#\mathcal{T}_h$     | 128 | 200 | 280  | 408  | 576  | 776  | 944  | 1016 |
| BDM Unpreconditioned  | 434 | 682 | 1046 | 1399 | 2094 | 3049 | 4324 | 6156 |
| BDM-BDM Jacobi        | 12  | 12  | 12   | 12   | 12   | 12   | 12   | 12   |

Table 5.2.: Number of iterations for decreasing mesh width. Numerical discretization based on full polynomial DG spaces and second family edge elements (rotated  $\mathcal{BDM}_1$ ) for the auxiliary space. Pointwise Jacobi smoother.

#### 5.4.2. Structured Triangular Mesh. One Discontinuous Coefficient.

**Structured Triangular Mesh. Coefficients:**  $\beta = 1$ ,  $\nu$  Discontinuous.

Let us assume that  $\beta = 1$ , whilst the magnetic diffusivity  $\nu$  is discontinuous and piecewise constant, namely

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 & \text{if } \mathbf{x} \in \Omega_1 := [0, 0.5]^2 \cup [0.5, 1]^2, \\ \nu_2 & \text{otherwise.} \end{cases}$$

The coefficient  $\nu_2 = 1$  is fixed. Note that the initial uniform triangulation resolves the jump discontinuities of the coefficient  $\nu$ . We compare the performances of the preconditioner as the mesh is uniformly refined and for different values of the coefficient  $\nu_1$ . The discretization is based on lowest order rotated BDM elements  $\mathcal{P}_1^d \Lambda^1(\mathcal{T}_h)$  and preconditioned with auxiliary space as in (5.4) with pointwise Jacobi smoother (5.31).

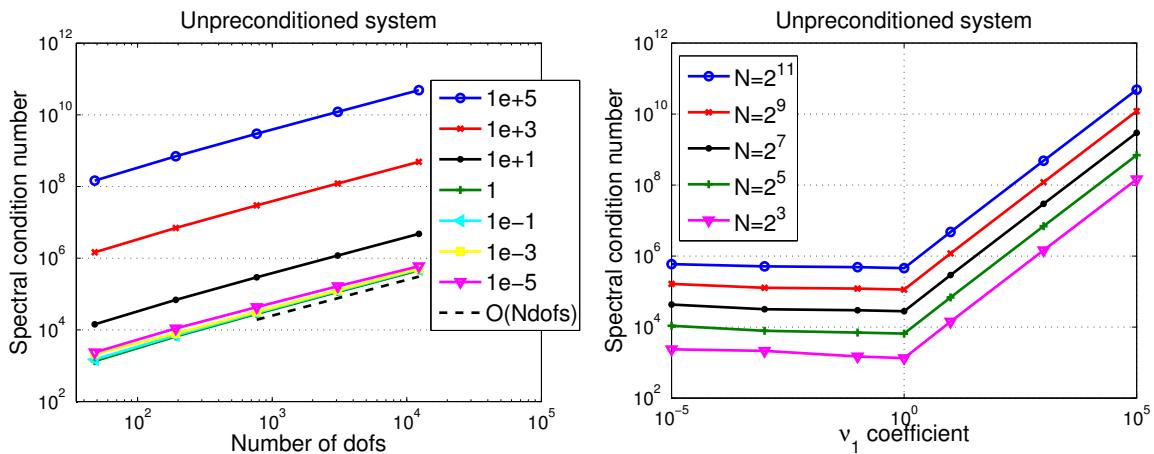


Figure 5.6.: Condition number vs. number of dofs for different values of the coefficient  $\nu$  (left). Condition number vs. values of  $\nu$  for different mesh widths (right). Discretization based on full polynomial DG space. Unpreconditioned system.

When the problem is solved without appealing to a preconditioner, the spectral condition number depends, as expected, on the mesh width and on the magnitude of the jump (Figure 5.6 and Table 5.3). Concerning

the preconditioned system, the condition number is independent of the mesh width (Figure 5.7, left) and it is asymptotically independent on the magnitude of the jump of the coefficient  $\nu$  (see Figure 5.7, right). This can be readily checked also from Table 5.3, which report the condition number and number of iterations for different values of the coefficient  $\nu_1$ . This is in agreement with Theorem 5.3.1 (see also Remark 5.3.6) which predicts uniform convergence when only one of the two coefficients is allowed to vary.

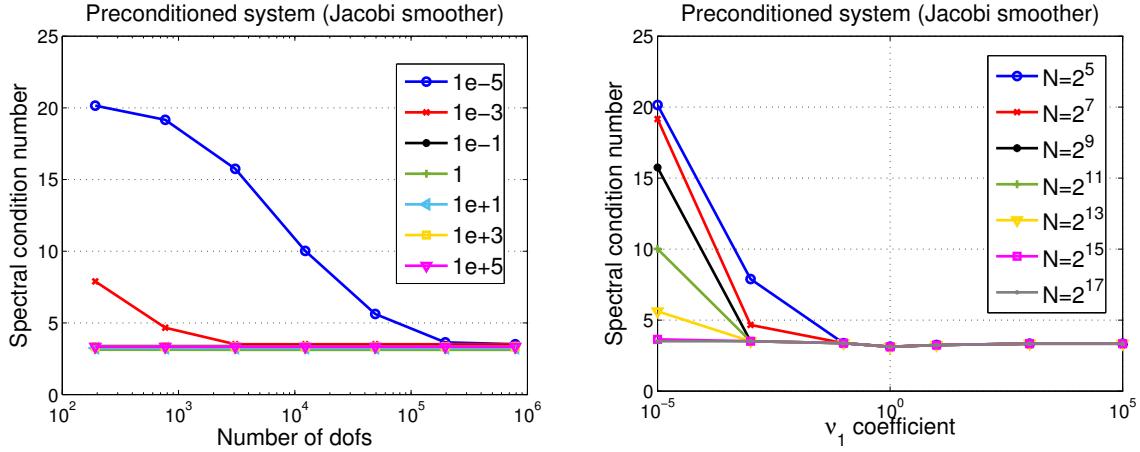


Figure 5.7.: Condition number vs. number of dofs for different values of the coefficient  $\nu$  (left). Condition number vs. values of  $\nu$  for different mesh widths (right). Discretization based on full polynomial DG space and second family edge elements for the auxiliary space. Pointwise Jacobi smoother.

|                  |                    | Unpreconditioned system |                |                |                | Preconditioned system |                   |                |                |                |                 |                 |                 |                 |
|------------------|--------------------|-------------------------|----------------|----------------|----------------|-----------------------|-------------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
|                  |                    | # $\mathcal{T}_h$       | 2 <sup>5</sup> | 2 <sup>7</sup> | 2 <sup>9</sup> | 2 <sup>11</sup>       | # $\mathcal{T}_h$ | 2 <sup>5</sup> | 2 <sup>7</sup> | 2 <sup>9</sup> | 2 <sup>11</sup> | 2 <sup>13</sup> | 2 <sup>15</sup> | 2 <sup>17</sup> |
| $\nu_1$          | $\# \mathcal{T}_h$ |                         |                |                |                |                       |                   |                |                |                |                 |                 |                 |                 |
| 10 <sup>-5</sup> |                    |                         | 172            | 424            | 810            | 1598                  |                   | 27             | 33             | 31             | 22              | 16              | 11              | 9               |
| 10 <sup>-3</sup> |                    |                         | 158            | 465            | 1281           | 3333                  |                   | 20             | 16             | 13             | 12              | 12              | 11              | 11              |
| 10 <sup>-1</sup> |                    |                         | 199            | 619            | 1405           | 2777                  |                   | 13             | 13             | 12             | 12              | 12              | 11              | 11              |
| 1                |                    |                         | 144            | 374            | 760            | 1490                  |                   | 12             | 12             | 12             | 11              | 11              | 11              | 10              |
| 10 <sup>1</sup>  |                    |                         | 280            | 905            | 2073           | 3807                  |                   | 12             | 12             | 12             | 12              | 11              | 11              | 11              |
| 10 <sup>3</sup>  |                    |                         | 640            | 3732           | 12324          | 31630                 |                   | 13             | 15             | 13             | 12              | 12              | 12              | 12              |
| 10 <sup>5</sup>  |                    |                         | 1155           | 7091           | 27563          | >50000                |                   | 14             | 15             | 14             | 14              | 14              | 13              | 13              |

Table 5.3.: Number of iterations. Coefficients:  $\beta = 1$ ,  $\nu$  discontinuous. Discretization based on full polynomial DG space and second family edge elements for the auxiliary space. Pointwise Jacobi smoother.

### Structured Triangular Mesh. Coefficients: $\beta$ Discontinuous, $\nu = 1$ .

In this experiment, we assume the magnetic diffusivity to be  $\nu = 1$ , while the reaction coefficient  $\beta$  is discontinuous and piecewise constant, namely

$$\beta(\mathbf{x}) = \begin{cases} \beta_1 & \text{if } \mathbf{x} \in \Omega_1 := [0, 0.5]^2 \cup [0.5, 1]^2, \\ \beta_2 & \text{otherwise.} \end{cases}$$

The coefficient  $\beta_1 = 1$  is fixed. Lowest order rotated BDM elements  $\mathcal{P}_1^d \Lambda^1(\mathcal{T}_h)$  are used for both the DG discretization and the auxiliary space (5.4). The auxiliary space preconditioner uses pointwise Jacobi smoother (5.31). Table 5.4 reports the number of iterations as the coefficient  $\beta_1$  varies. As it can be observed in Figure 5.8, the condition number of the preconditioned system is asymptotically independent

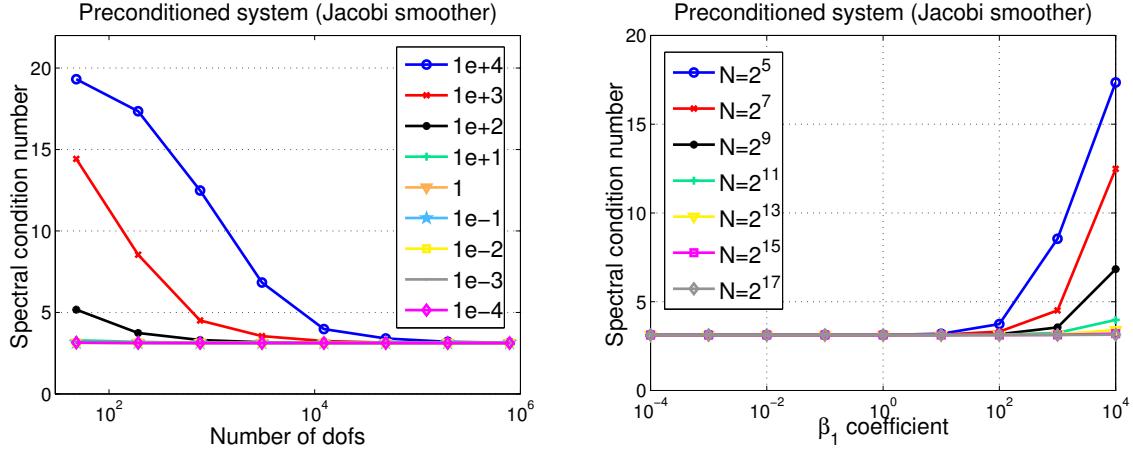


Figure 5.8.: Condition number vs. number of dofs for different values of the coefficient  $\beta$  (left). Condition number vs. values of  $\beta$  for different mesh widths (right). Discretization based on full polynomial DG space and second family edge elements for the auxiliary space. Pointwise Jacobi smoother.

both on the mesh width and on the magnitude of the jump of the coefficient  $\beta$ . This is in agreement with Theorem 5.3.1 (and Remark 5.3.6) which predicts uniform convergence when only one of the two coefficients is allowed to vary.

|           |                  | Unpreconditioned system |                |                | Preconditioned system |                   |                |                |                |                 |                 |                 |                 |
|-----------|------------------|-------------------------|----------------|----------------|-----------------------|-------------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
|           |                  | $\#\mathcal{T}_h$       | 2 <sup>5</sup> | 2 <sup>7</sup> | 2 <sup>9</sup>        | $\#\mathcal{T}_h$ | 2 <sup>5</sup> | 2 <sup>7</sup> | 2 <sup>9</sup> | 2 <sup>11</sup> | 2 <sup>13</sup> | 2 <sup>15</sup> | 2 <sup>17</sup> |
| $\beta_2$ | $\beta_1$        |                         |                |                |                       |                   |                |                |                |                 |                 |                 |                 |
|           | 10 <sup>-4</sup> | 284                     | 2250           | 3685           |                       | 10 <sup>-4</sup>  | 12             | 12             | 12             | 11              | 11              | 10              | 10              |
|           | 10 <sup>-3</sup> | 252                     | 1980           | 4823           |                       | 10 <sup>-3</sup>  | 12             | 12             | 12             | 11              | 11              | 10              | 10              |
|           | 10 <sup>-2</sup> | 305                     | 1559           | 4048           |                       | 10 <sup>-2</sup>  | 12             | 12             | 12             | 11              | 11              | 10              | 10              |
|           | 10 <sup>-1</sup> | 248                     | 877            | 2003           |                       | 10 <sup>-1</sup>  | 12             | 12             | 12             | 11              | 11              | 10              | 10              |
|           | 1                | 144                     | 374            | 760            |                       | 1                 | 12             | 12             | 12             | 11              | 11              | 11              | 10              |
|           | 10 <sup>1</sup>  | 173                     | 570            | 1302           |                       | 10 <sup>1</sup>   | 12             | 12             | 12             | 11              | 11              | 10              | 10              |
|           | 10 <sup>2</sup>  | 124                     | 538            | 1254           |                       | 10 <sup>2</sup>   | 15             | 13             | 12             | 11              | 11              | 10              | 10              |
|           | 10 <sup>3</sup>  | 116                     | 356            | 751            |                       | 10 <sup>3</sup>   | 21             | 16             | 13             | 11              | 10              | 9               | 9               |
|           | 10 <sup>4</sup>  | 121                     | 340            | 555            |                       | 10 <sup>4</sup>   | 28             | 24             | 18             | 13              | 11              | 9               | 9               |

Table 5.4.: Number of iterations. Coefficients:  $\nu = 1$ ,  $\beta$  discontinuous. Discretization based on full polynomial DG space and second family edge elements for the auxiliary space. Pointwise Jacobi smoother.

### 5.4.3. Structured Triangular Mesh. Two Discontinuous Coefficients.

We now turn to the more interesting and challenging case of both  $\beta$  and  $\nu$  discontinuous. Let  $\Omega_1 := [0, 0.5]^2 \cup [0.5, 1]^2$ , we define

$$\nu(\mathbf{x}) = \begin{cases} 10^{-4} & \text{if } \mathbf{x} \in \Omega_1, \\ 10^{-2} & \text{otherwise,} \end{cases} \quad \text{and} \quad \beta(\mathbf{x}) = \begin{cases} 2^\delta \cdot 10^{-4} & \text{if } \mathbf{x} \in \Omega_1, \\ 2^\delta \cdot 10^{-2} & \text{otherwise.} \end{cases} \quad (5.73)$$

where  $\delta \in [-20, 45] \cap \mathbb{Z}$ . For a given mesh with  $\#\mathcal{T}_h = 512$  elements, we analyze the spectral condition number of the unpreconditioned and preconditioned system as the ratio

$$L(h, \nu, \beta) := h_T^2 \frac{\beta_T}{\nu_T},$$

varies. On account of the quasi-uniformity of the mesh and the choice of the coefficients in (5.73), the ratio  $L(h, \nu, \beta)$  is constant on every  $T \in \mathcal{T}_h$  and only depends on the parameter  $\delta$ . For a discretization based on lowest order piecewise polynomials  $\mathcal{P}_1^d \Lambda^1(\mathcal{T}_h)$ , we consider three different preconditioners: a single pointwise Jacobi as in (5.31), the ASM preconditioner based on the lowest order second kind edge elements for the auxiliary space (5.4) with pointwise Jacobi as smoother (5.31) and, as third case, the ASM preconditioner based on lowest order first kind edge elements for the auxiliary space (5.4) and overlapping additive Schwarz smoother (edge based) (5.29). As predicted by Proposition 5.3.5, when  $L(h, \nu, \beta) > 1$  the auxiliary space is not needed to ensure uniform convergence with respect to both the problem coefficients and the mesh width, see Figure 5.9.

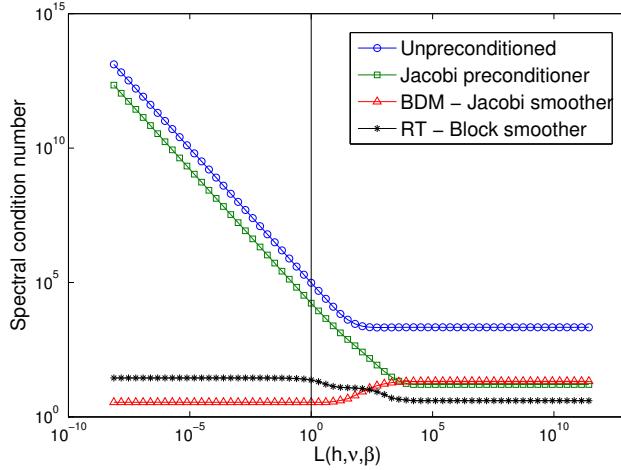


Figure 5.9.: Condition number vs. ratio  $L$  for a fixed mesh size  $h$ . The condition number refers to the unpreconditioned system (—○—), the pointwise Jacobi preconditioner (—□—), the auxiliary space preconditioner based on rotated  $\mathcal{BDM}_1$  elements with pointwise Jacobi smoother (—△—) and the auxiliary space preconditioner based on rotated  $\mathcal{RT}_0$  elements with overlapping Schwarz smoother (—\*—). Discretization: discontinuous lowest order rotated  $\mathcal{BDM}_1$  elements.

### Checkerboard Experiment

We consider an experiment where the distribution of the coefficients follows a checkerboard pattern according to the partition:

$$\Omega_1 := \bigcup_{i=0}^3 \left[ \frac{1}{4}i, \frac{1}{4}(i+1) \right]^2 \cup \bigcup_{i=0}^3 \left[ \frac{1}{4}i, \frac{1}{4}(i+1) \right] \times \left[ \frac{1}{4}(i+2)\text{mod}4, \frac{1}{4}(i+2)\text{mod}4 + \frac{1}{4} \right],$$

as depicted in Figure 5.10 (white patches correspond to  $\Omega_1$ ). We define

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 & \text{if } \mathbf{x} \in \Omega_1, \\ \nu_2 & \text{otherwise,} \end{cases} \quad \text{and} \quad \beta(\mathbf{x}) = \begin{cases} \beta_1 & \text{if } \mathbf{x} \in \Omega_1, \\ \beta_2 & \text{otherwise.} \end{cases}$$

where  $\nu_1, \nu_2, \beta_1$  and  $\beta_2$  are set to different values for the three experiments carried out, see Table 5.5.

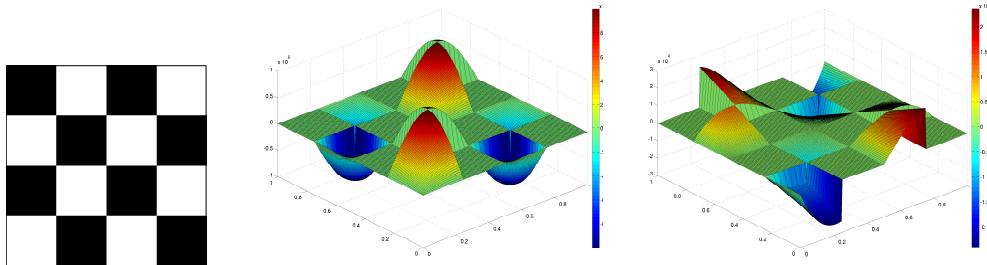


Figure 5.10.: Components of the analytical vector-valued solution of the checkerboard experiment with  $\nu_1 = 10^{-2}, \nu_2 = 10^4, \beta_1 = 10^3, \beta_2 = 1$ .

Lowest order rotated BDM elements  $\mathcal{P}_1^d \Lambda^1(\mathcal{T}_h)$  are used both in the DG discretization and for the auxiliary space (5.4) in the ASM preconditioner together with a pointwise Jacobi smoother (5.31). In Figure 5.10 (center and rightmost) are represented the two components of the approximate solution. We note that the weak regularity of the solution leads to a significantly reduced convergence rate of the DG scheme. In Table 5.5, the estimated spectral condition numbers and the number of iterations required for convergence for the three different configurations of the coefficients are given. As can be observed in Table 5.5, the preconditioner significantly outperforms the unpreconditioned system even if a slight dependence of the jump of the coefficients might be recorded. Such dependence however seems to hinge on the possible transition from reaction-dominated to **curl**-dominated. Notice that the first and third cases reported in Table 5.5, correspond to cases where **curl**-dominated and reaction-dominated regimes alternate in the checkerboard pattern (for the first case the problem becomes **curl**-dominated in the whole domain for the two finest meshes), while the second case reported in Table 5.5 corresponds to the **curl**-dominated regime.

| $\#\mathcal{T}_h$ | $2^7$                       | $2^9$   | $2^{11}$                     | $2^{13}$           | $2^{15}$           |
|-------------------|-----------------------------|---|------------------------------|--------------------|--------------------|
|                   |                             | $\nu_1 = 10^{-2}, \quad \nu_2 = 10^4 ; \quad \beta_1 = 10^3, \quad \beta_2 = 1$     |                              |                    |                    |
| CG                | $2.45\text{e}+8 - (1881)$   | $1.17\text{e}+9 - (7840)$   | $5.00\text{e}+9 - (18899)$   | —                  | —                  |
| PCG               | $19.035376 - (28)$          | $15.740534 - (30)$  | $9.965288 - (26)$            | $5.791281 - (19)$  | $3.643158 - (14)$  |
|                   |                             | $\nu_1 = 10^4, \quad \nu_2 = 10 ; \quad \beta_1 = 10^{-2}, \quad \beta_2 = 10^{-4}$ |                              |                    |                    |
| CG                | $2.45\text{e}+12 - (15275)$ | $1.17\text{e}+13 - (>50000)$  | $5.03\text{e}+13 - (>50000)$ | —                  | —                  |
| PCG               | $3.509092 - (13)$           | $3.509857 - (13)$   | $3.509857 - (13)$            | $3.509857 - (13)$  | $3.509857 - (16)$  |
|                   |                             | $\nu_1 = 10^{-3}, \quad \nu_2 = 1 ; \quad \beta_1 = 10^4, \quad \beta_2 = 10^2$     |                              |                    |                    |
| CG                | $1.35\text{e}+3 - (148)$    | $2.23\text{e}+3 - (228)$  | $5.56\text{e}+3 - (363)$     | —                  | —                  |
| PCG               | $20.525256 - (28)$          | $20.714987 - (30)$  | $20.524956 - (30)$           | $19.729734 - (29)$ | $17.183196 - (28)$ |

Table 5.5.: Condition number and number of iterations (in brackets) for the checkerboard experiment. Discretization based on lowest order rotated BDM elements. Auxiliary space preconditioner based on lowest order edge element of the second family and pointwise Jacobi smoother.

#### 5.4.4. Tensor Product Meshes

On tensor product meshes, we only deal with constant coefficients  $\nu$  and  $\beta$ . As first test case, we consider the model problem (5.1) on  $\Omega = [0, \pi]^2$  with  $\nu = 1$  and  $\beta = 0$  as in [BHP07, Section 5.1, Example 1]. The exact eigenvalues are given by  $n^2 + m^2$  for  $n$  and  $m$  positive integers. We compute and show in Figure 5.11 the lower part of the spectrum using a DG discretization based on the full polynomials space  $\mathcal{Q}_1^d \Lambda^1(\mathcal{T}_h)^2$  (right), the rotated version of the space  $\mathcal{S}_1^d \Lambda^1(\mathcal{T}_h)$  (center),  $\mathcal{S}_1 \Lambda^1(T) := \mathcal{RT}_0(T) + \{\mathbf{curl}(x^2y), \mathbf{curl}(xy^2), \mathbf{curl}(x^2), \mathbf{curl}(y^2)\} = \mathcal{BDM}_1(T)$ , and the rotated Nédélec elements of the first family  $\mathcal{RT}_0$  (left). As pointed out in Remark 5.3.14 and numerically observed in Figure 5.11, on quadrilateral meshes a DG discretization based on the full polynomial space of degree  $r$  in each variable, is not spectrally correct. Therefore, a preconditioner built on an auxiliary space where the  $H_0(\mathbf{curl}, \Omega)$ -conforming discretization is spectrally correct (e.g. Nédélec elements of the first family) is not effective, independently of the choice of the smoother and the amount of domain overlaps involved in its construction as it can be inferred from Figure 5.12 and Table 5.6.

However, as stated in Remark 5.3.14, an effective ‘‘coarser’’ auxiliary space (with overlapping smoothers) can be constructed for the special case of a discontinuous approximation based on the finite element space  $\mathcal{S}_1^d \Lambda^1(\mathcal{T}_h)$ . On the other hand, in view of Theorem 5.3.1, using the same local spaces for the discretization and for the  $H(\mathbf{curl}, \Omega)$ -conforming auxiliary space provides an effective solver. This can be easily checked in Figure 5.12, Figure 5.13 and Table 5.6 where the results obtained with DG discretizations based on local spaces of first family edge elements and full polynomials space  $\mathcal{Q}_r^d \Lambda^1(T)^2$ , preconditioned with auxiliary space built on the  $H(\mathbf{curl}, \Omega)$ -conforming global elements of the same family in each case are (also) reported.

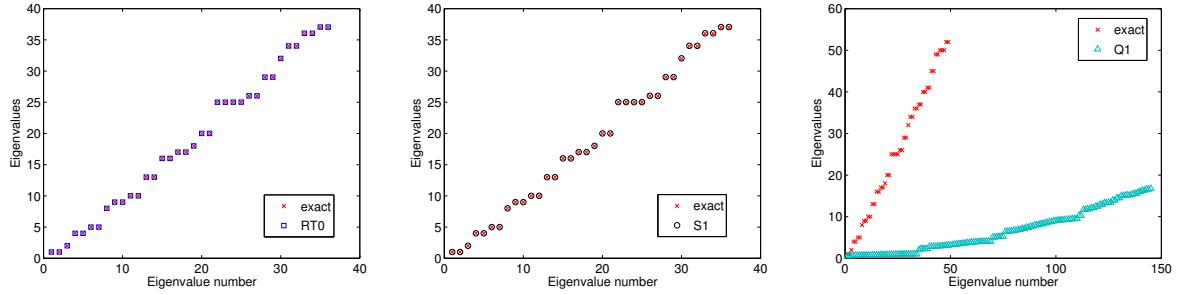


Figure 5.11.: Lower part of the spectrum obtained with different DG discretizations: rotated Nédélec elements of the first family  $\mathcal{RT}_0$  (left), rotated  $\mathcal{S}_1$  (center), and the full polynomial space ( $Q_1$ )<sup>2</sup> (right).

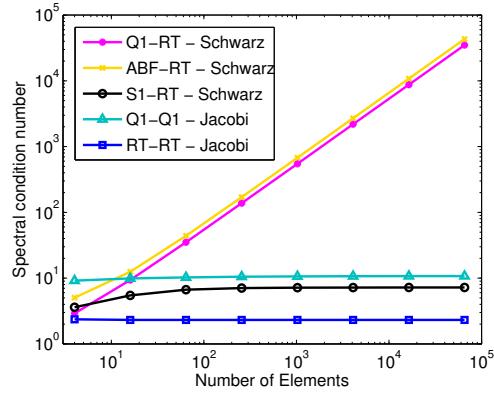


Figure 5.12.: Condition number vs. number of dofs (left): piecewise bilinear Lagrangian elements discretization with ASM based on rotated  $\mathcal{RT}_0$  elements with overlapping additive Schwarz smoother (Q1-RT); same auxiliary space and smoother coupled with DG discretizations associated with  $\mathcal{ABF}_0$  (ABF-RT) and  $\mathcal{S}_1^d \Lambda^1(\mathcal{T}_h)$  (S1-RT); DG discretization with rotated  $\mathcal{RT}_0$  discontinuous elements and rotated  $\mathcal{RT}_0$  as auxiliary space with pointwise Jacobi smoother (RT-RT); discontinuous bilinear Lagrangian elements with  $H(\mathbf{curl}, \Omega)$ -conforming full polynomial auxiliary space and Jacobi smoother (Q1-Q1).

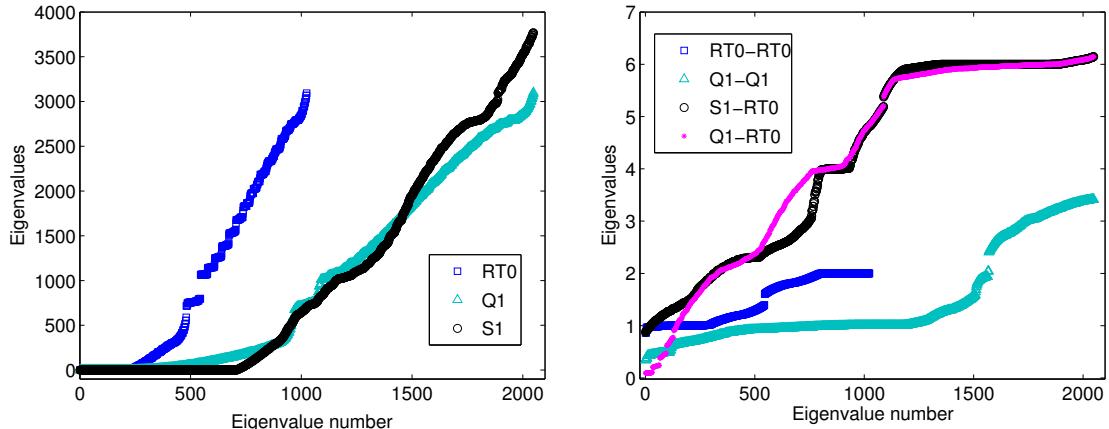


Figure 5.13.: Spectrum of the Galerkin matrix associated with different DG discretizations (left) and spectrum of the corresponding preconditioned matrices (right), for different combinations as in Figure 5.12.

Piecewise polynomial approximations based on the  $H(\mathbf{curl}, \Omega)$ -conforming version of the  $H(\mathbf{div}, \Omega)$ -conforming element  $\mathcal{ABF}_k = \mathcal{P}_{k+2,k} \times \mathcal{P}_{k,k+2}$  [ABF05, Section 5] have also been implemented for  $k = 0$ . Even in this case, a preconditioner based on rotated  $\mathcal{RT}_0$  elements performs poorly (see Table 5.6).

| $\#\mathcal{T}_h$            | $16 \times 16$ | $32 \times 32$ | $64 \times 64$ | $128 \times 128$ | $256 \times 256$ |
|------------------------------|----------------|----------------|----------------|------------------|------------------|
| RT0 Unpreconditioned         | 128            | 204            | 376            | 753              | 1504             |
| Q1 Unpreconditioned          | 410            | 815            | 1454           | 2796             | 4554             |
| S1 Unpreconditioned          | 543            | 1083           | 2031           | 4056             | 7316             |
| RT0-RT0 Jacobi               | 9              | 9              | 9              | 9                | 9                |
| Q1-Q1 Jacobi                 | 22             | 21             | 20             | 19               | 19               |
| Q1-RT0: Jacobi   overlapping | 259   61       | 471   113      | 844   202      | 1622   337       | 2936   618       |
| ABF0-RT0 overlapping         | 59             | 116            | 230            | 458              | 874              |
| S1-RT0: Jacobi   overlapping | 88   18        | 72   19        | 49   20        | 34   20          | 36   19          |

Table 5.6.: Number of iterations for different discretizations and preconditioners on a tensor product mesh.



## Part II.

# The Hydrodynamic Equations and the Full Ideal MHD



# 6. The Extended Euler Equations

The second block of the MHD system (1.10) consists of the conservation equations for mass, momentum and (hydrodynamic) energy with the magnetic induction field entering through the Lorentz force. As outlined in Chapter 1, and in the spirit of semi-Godunov schemes (Section 1.2.2), the approach we pursue here is based on the local (in time) reduction of the MHD system into two systems with discontinuous coefficients: the advection of magnetic induction/potential with a known discontinuous velocity field (see Chapters 3 and 4) and the conservation laws for the fluid variables with the  $\mathbf{B}$  field treated as a discontinuous known function.

The numerical discretization of the so-called extended Euler equations is performed using standard finite volume schemes, Section 6.2. Approximate Riemann solvers for the underlying reduced MHD system are introduced in Section 6.2.1: in an attempt to foil the loss of physically relevant information, the wave structure of the extended Euler Riemann problem presented in Section 6.1 is considered as a subset of the MHD fan with wave speeds always taken from the latter. Higher order finite volume schemes are considered in Section 6.3.

## 6.1. Derivation of the Extended Euler Equations

As introduced in Chapter 1, MHD theory describes the combined collective interaction of moving electrically conducting fluid and electromagnetic fields. A conductor moving across a magnetic field experiences an electromotive force which in turn drives electric currents flowing in the conductor. The magnetic field associated with these currents will modify the original magnetic field altering the flow motion accordingly.

With the name of extended Euler equations<sup>1</sup> we denote the Euler equations of gas dynamics for a compressible fluid, subject to a combination of electric and magnetic forces related to the presence of electromagnetic fields, namely

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0 \\ \partial_t(\rho \mathbf{u}) + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u} + p \mathbb{I}) = \mathbf{J} \times \mathbf{B} \\ \partial_t E^{\text{hd}} + \operatorname{div}((E^{\text{hd}} + p) \mathbf{u}) = \mathbf{J} \cdot \mathbf{E} \end{cases} \quad (6.1)$$

where  $E^{\text{hd}}$  denotes the hydrodynamic energy, sum of the thermal and kinetic energies. We can combine Ampère's law  $\mathbf{J} = \operatorname{curl} \mathbf{H}$ , the material law (1.4) with  $\mu = 1$ , and Ohm's law for a perfectly conducting fluid  $\mathbf{E} = \mathbf{B} \times \mathbf{u}$ ; thereby:

$$\mathbf{J} \cdot \mathbf{E} = \operatorname{curl} \mathbf{B} \cdot (\mathbf{B} \times \mathbf{u}) = \operatorname{curl}(\mathbf{B} \times \mathbf{u}) \cdot \mathbf{B} - \operatorname{div}((\mathbf{B} \times \mathbf{u}) \times \mathbf{B}).$$

Using the magnetic advection equation for the  $\mathbf{B}$  field together with the vector identity  $(\mathbf{B} \times \mathbf{u}) \times \mathbf{B} = (\mathbf{B} \cdot \mathbf{B})\mathbf{u} - (\mathbf{u} \cdot \mathbf{B})\mathbf{B}$  results in  $\mathbf{J} \cdot \mathbf{E} = -\partial_t \mathbf{B} \cdot \mathbf{B} - \operatorname{div}((\mathbf{B} \cdot \mathbf{B})\mathbf{u}) + \operatorname{div}((\mathbf{u} \cdot \mathbf{B})\mathbf{B})$ . Substituting into the conservation of energy, yields the conservation law

$$\partial_t E + \operatorname{div} \left( \left( E + p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B})\mathbf{B} \right) = 0$$

for the total energy  $E = E^{\text{hd}} + \|\mathbf{B}\|_{\ell^2}^2/2$  which includes the magnetic energy.

---

<sup>1</sup>the term appeared in [FMR09].

Moreover, the contribution of the Lorentz force gives

$$\mathbf{J} \times \mathbf{B} = \mathbf{curl} \mathbf{B} \times \mathbf{B} = \mathbf{div} \left( \mathbf{B} \otimes \mathbf{B} - \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \mathbb{I} \right) - \mathbf{B} \mathbf{div} \mathbf{B},$$

coupled with the divergence constraint on the magnetic induction field yields,

$$\partial_t(\rho \mathbf{u}) + \mathbf{div} \left( \rho \mathbf{u} \otimes \mathbf{u} + \left( p + \frac{\|\mathbf{B}\|_{\ell^2}^2}{2} \right) \mathbb{I} - \mathbf{B} \otimes \mathbf{B} \right) = \mathbf{0},$$

where  $\rho \mathbf{u} \otimes \mathbf{u}$  is the Reynolds stress tensor and  $(\|\mathbf{B}\|_{\ell^2}^2/2) \mathbb{I} - \mathbf{B} \otimes \mathbf{B}$  is the magnetic part of the Maxwell stress tensor. It is thus manifest that the extended Euler system constitutes the fluid part of the ideal MHD equations, namely (6.1) can be recast as

$$\begin{cases} \partial_t \rho + \mathbf{div}(\rho \mathbf{u}) = 0 \\ \partial_t(\rho \mathbf{u}) + \mathbf{div} \left( \rho \mathbf{u} \otimes \mathbf{u} + \left( p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \mathbb{I} - \mathbf{B} \otimes \mathbf{B} \right) = \mathbf{0} \\ \partial_t E + \mathbf{div} \left( \left( E + p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B}) \mathbf{B} \right) = 0 \end{cases} \quad (6.2)$$

and in the absence of magnetic phenomena  $\mathbf{B} \equiv \mathbf{0}$ , the system (6.2) reduces to the compressible Euler equations.

### Characteristic Waves of the Extended Euler System

The extended Euler system is a parametric hyperbolic system of conservation laws. In three dimensions, if  $\mathbf{U} \in \mathbb{R}^m$ ,  $m \in \mathbb{N}$ , denotes the vector of the conserved fluid variables  $\mathbf{U} := (\rho, \rho u^1, \rho u^2, \rho u^3, E)$ , and  $\mathbf{x} = (x^1, x^2, x^3) \in \mathbb{R}^3$ , then (6.2) can be written as

$$\partial_t \mathbf{U} + \sum_{\ell=1}^3 \partial_{x^\ell} f^\ell(\mathbf{U}, \mathbf{B}) = \mathbf{0}. \quad (6.3)$$

If  $\delta_{i,\ell}$  denotes the Kronecker delta, the directional fluxes  $\{f^\ell\}_{\ell=1}^3$  are defined as

$$f^\ell(\mathbf{U}, \mathbf{B}) = \begin{pmatrix} \rho u^\ell \\ \rho u^1 u^\ell - B_1 B_\ell + \left( p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \delta_{1,\ell} \\ \rho u^2 u^\ell - B_2 B_\ell + \left( p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \delta_{2,\ell} \\ \rho u^3 u^\ell - B_3 B_\ell + \left( p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) \delta_{3,\ell} \\ \left( E + p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2 \right) u^\ell - (\mathbf{u} \cdot \mathbf{B}) B_\ell \end{pmatrix}.$$

In order to draw information on the spectral structure on the system, one can derive a quasi-linear form of (6.3) in the primitive variables unknowns  $\mathbf{V} = (\rho, u^1, u^2, u^3, p)$ , namely

$$\partial_t \mathbf{V} + \sum_{\ell=1}^3 A^\ell(\mathbf{V}, \mathbf{B}) \partial_{x^\ell} \mathbf{V} = S(\mathbf{V}, \mathbf{B}, D\mathbf{B}),$$

where the right hand side  $S$  might depend on the induction field  $\mathbf{B}$  and its Jacobian  $D\mathbf{B}$ . The directional Jacobian  $\mathbf{A}(\mathbf{V}, \mathbf{B}) \cdot \mathbf{n} := \sum_{\ell=1}^3 A^\ell(\mathbf{V}, \mathbf{B}) n_\ell$  in the direction of the unit vector  $\mathbf{n} = (n_1, n_2, n_3)$  is given by

$$\mathbf{A}(\mathbf{V}, \mathbf{B}) \cdot \mathbf{n} = \begin{pmatrix} \mathbf{u} \cdot \mathbf{n} & \rho n_1 & \rho n_2 & \rho n_3 & 0 \\ 0 & \mathbf{u} \cdot \mathbf{n} & 0 & 0 & 1/\rho n_1 \\ 0 & 0 & \mathbf{u} \cdot \mathbf{n} & 0 & 1/\rho n_2 \\ 0 & 0 & 0 & \mathbf{u} \cdot \mathbf{n} & 1/\rho n_3 \\ 0 & \rho(a^{E,1}(\mathbf{V}, \mathbf{B}))^2 & \rho(a^{E,2}(\mathbf{V}, \mathbf{B}))^2 & \rho(a^{E,3}(\mathbf{V}, \mathbf{B}))^2 & \mathbf{u} \cdot \mathbf{n} \end{pmatrix},$$

where the sound speeds of the acoustic waves of the extended Euler system in each direction are

$$a^{E,\ell}(\mathbf{V}, \mathbf{B}) = \frac{1}{\sqrt{\rho}} \sqrt{\gamma p n_\ell + (\gamma - 1) [n_\ell (\|\mathbf{B}\|_{\ell^2}^2 - B_\ell^2) - B_\ell (\mathbf{B} \cdot \mathbf{n} - B_\ell n_\ell)]}, \quad \ell \in \{1, 2, 3\}, \quad (6.4)$$

and the superscript E serves to distinguish the wave speeds of the extended Euler system to the waves associated with the MHD equations (1.12). If  $\mathbf{u}^E := (a^{E,1}(\mathbf{V}, \mathbf{B}), a^{E,2}(\mathbf{V}, \mathbf{B}), a^{E,3}(\mathbf{V}, \mathbf{B}))$ , the eigenvalues of the directional Jacobian are

$$\lambda^1 = \mathbf{u} \cdot \mathbf{n} - \mathbf{u}^E \cdot \mathbf{n}, \quad \lambda^{2,3,4} = \mathbf{u} \cdot \mathbf{n}, \quad \lambda^5 = \mathbf{u} \cdot \mathbf{n} + \mathbf{u}^E \cdot \mathbf{n}. \quad (6.5)$$

Hence, the extended Euler system is only weakly hyperbolic and the Riemann solution is characterized by five waves:

$$\begin{aligned} W^{1,5}, \mathbf{u} \cdot \mathbf{n} \mp \mathbf{u}^E \cdot \mathbf{n} : & \text{ acoustic waves, rarefaction/shock to the left/right,} \\ W^{2,3,4}, \mathbf{u} \cdot \mathbf{n} : & \text{ shear wave, contact discontinuity.} \end{aligned} \quad (6.6)$$

The characteristic wave structure of the system is particularly relevant for the design of finite volume numerical schemes in the forthcoming section. Ideally, we aim at Riemann solvers (see Section 6.2) whose exact solution does not contain a vacuum state and which generate *admissible* solutions, in the sense of physically meaningful states. The set of admissible states for the MHD equations consists of vectors of conserved variables (including the  $\mathbf{B}$  field) such that the density and the pressure are positive, namely

$$G := \{\hat{\mathbf{U}} = (\rho, \rho \mathbf{u}, E, \mathbf{B}) : \rho > 0, p(\hat{\mathbf{U}}) > 0\}. \quad (6.7)$$

Finite volume schemes preserving admissible states are called *positively conservative*.

## 6.2. Finite Volume Discretization

In the last decades, finite volume methods have been widely used as numerical schemes to approximate conservation laws. In these methods, the computational domain is divided into cells, and the integral formulation of the balance equations is numerically solved by evolving in time the cell averages of the conserved variables.

For the sake of simplicity, in the present and the following chapter, we restrict to Cartesian domains  $\Omega = J_1 \times \dots \times J_d \subset \mathbb{R}^d$ , with  $J_\ell \subset \mathbb{R}$ ,  $\ell = 1, \dots, d$ , bounded and connected. We proceed via a *dimensional splitting* by designing numerical discretizations of one-dimensional problems in each space direction and combining the solutions thus obtained. Let  $\{\mathcal{T}_h\}_h$  be a family of partitions of  $\Omega$  obtained, for all  $h > 0$ , as  $\mathcal{T}_h = \mathcal{T}_{h_1}^1 \times \dots \times \mathcal{T}_{h_d}^d$ , where  $\mathcal{T}_{h_\ell}^\ell$  is a uniform mesh on  $J_\ell$  with  $M_h^\ell := \#\mathcal{T}_{h_\ell}^\ell$  elements. Hence, the mesh width in the  $\ell$ -direction is given by  $h_\ell = |J_\ell|/M_h^\ell$  and  $h = \max_\ell h_\ell$ . Every element  $T_j \in \mathcal{T}_h$  (also called *control volume* in the FV jargon) is identified by its barycenter  $\mathbf{x}_j = (x_{j_1}^1, \dots, x_{j_d}^d)$ , where  $j = (j_1, \dots, j_d)$  is a multi-index in  $\mathfrak{J} := \mathbb{N}^d \cap ([1, M_h^1] \times \dots \times [1, M_h^d])$ , which “selects” the spatial direction. The  $(d-1)$ -skeleton of the mesh is fundamental in determining the fluxes: the interfaces of the element  $T_j \in \mathcal{T}_h$  are denoted by  $\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_t}$  where  $\mathbf{e}_t$  is the  $t$ -th unit vector in  $\mathbb{R}^d$ .

The weak solution (*cf.* (1.13))  $\mathbf{U}(\mathbf{x}, t)$  of the extended Euler system is approximated by means of *cell averages*

$$\mathbf{U}(\mathbf{x}, t)|_{T_j} \approx \mathbf{U}_j(t) := \frac{1}{|T_j|} \int_{T_j} \mathbf{U}(\mathbf{x}, t) d\mathbf{x}, \quad \forall T_j \in \mathcal{T}_h, j \in \mathfrak{J}. \quad (6.8)$$

Integrating the conservation law (6.3) on a fixed element  $T_j \in \mathcal{T}_h$  results into a semi-discrete finite volume scheme of the form

$$\partial_t \mathbf{U}_j(t) = - \sum_{\ell=1}^d \frac{f_{j+\frac{1}{2}\mathbf{e}_\ell}^\ell(t) - f_{j-\frac{1}{2}\mathbf{e}_\ell}^\ell(t)}{h_\ell}, \quad f_{j \pm \frac{1}{2}\mathbf{e}_\ell}^\ell(t) := f^\ell(\mathbf{U}(\mathbf{x}_{j \pm \frac{1}{2}\mathbf{e}_\ell}, t), \mathbf{B}(\mathbf{x}_{j \pm \frac{1}{2}\mathbf{e}_\ell}, t)), \quad \forall j \in \mathfrak{J}.$$

The conservation property is maintained at the discrete level: on each element, the rate of change of the cell average is given by the difference of the numerical fluxes across the element boundaries,

$$\int_{t^n}^{t^{n+1}} \partial_t \mathbf{U}_j(t) dt = - \sum_{\ell=1}^d \frac{1}{h_\ell} \left( \int_{t^n}^{t^{n+1}} f_{j+\frac{1}{2}\mathbf{e}_\ell}^\ell(t) dt - \int_{t^n}^{t^{n+1}} f_{j-\frac{1}{2}\mathbf{e}_\ell}^\ell(t) dt \right), \quad \forall j \in \mathfrak{J}.$$

Here  $\bigcup_{n=0}^{N-1} [t^n, t^{n+1}]$ ,  $N \in \mathbb{N}$ , is a (not necessarily uniform) partition of the temporal interval  $I = [0, T]$  such that  $t^{n+1} = t^n + \Delta t^n$ , with time step  $\Delta t^n$  satisfying a CFL-type condition which will be discussed in the forthcoming section.

At each time  $t^n$ , the semi-discrete formulation requires an approximation of the fluxes which otherwise would need a priori knowledge of the solution at each interface. With the aim of finding approximations of the interface fluxes, Godunov observed in [God59] that the  $\mathcal{T}_h$ -piecewise constant cell averages define a Riemann problem at each  $(d-1)$ -cell of the mesh skeleton, namely a conservation law with initial data consisting of two constant states separated by a simple discontinuity. In the  $x^\ell$ -direction for all  $\ell = 1, \dots, d$ :

$$\begin{aligned} \partial_t \mathbf{U} + \partial_{x^\ell} f^\ell(\mathbf{U}, \mathbf{B}) &= \mathbf{0}, \\ \mathbf{U}(\mathbf{x}, t^n) &= \begin{cases} \mathbf{U}_j(t^n) & \text{if } \mathbf{x} < \mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}, \forall j \in \mathfrak{J}, \\ \mathbf{U}_{j+\mathbf{e}_\ell}(t^n) & \text{if } \mathbf{x} > \mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}, \forall j \in \mathfrak{J}. \end{cases} \end{aligned} \quad (6.9)$$

The solution of (6.3) at each time can be approximated by superposition of solutions of local Riemann problems representing waves emanating from each mesh interface in the following way. Since the solution  $\bar{\mathbf{U}}(\mathbf{x}, t)$  of the Riemann problem (6.9) is self-similar  $\bar{\mathbf{U}}(\mathbf{x}, t) = \bar{\mathbf{U}}((\mathbf{x} - \mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell})/(t - t^n))$ , the interface flux is constant i.e.

$$f^\ell(\bar{\mathbf{U}}(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}, t)) = f^\ell(\bar{\mathbf{U}}(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}, t^n)) =: f^\ell(\bar{\mathbf{U}}_{j+\frac{1}{2}\mathbf{e}_\ell}), \quad \forall j \in \mathfrak{J}, \forall \ell = 1, \dots, d,$$

(the dependence on  $\mathbf{B}$  has been omitted). Moreover, if  $\lambda_{\max}^\ell$  is the maximum wave speed of the system of conservation laws in the  $\ell$ -th direction (as will be defined in (6.12)), under the CFL condition

$$\sum_{\ell=1}^d \frac{\lambda_{\max}^\ell}{h_\ell} \Delta t \leq \frac{1}{2},$$

the waves from neighboring elements do not intersect. In this way, the approximate numerical fluxes are well-defined at the interfaces: if  $\bar{\mathbf{U}}$  is discontinuous at  $\mathbf{x} = \mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}$  for all  $t > t^n$ , a stationary shock is located at the interface and Rankine–Hugoniot conditions ensure that

$$\lim_{\epsilon \rightarrow 0} \bar{\mathbf{U}}(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell} - \epsilon) = \lim_{\epsilon \rightarrow 0} \bar{\mathbf{U}}(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell} + \epsilon).$$

In conclusion, the “exact” fluxes at time  $t^{n+1}$  can then be approximated as

$$f_{j \pm \frac{1}{2}\mathbf{e}_\ell}^\ell(t^{n+1}) \approx f^\ell(\bar{\mathbf{U}}_{j \pm \frac{1}{2}\mathbf{e}_\ell}) =: F_{j \pm \frac{1}{2}\mathbf{e}_\ell}^{\ell, n+1}(\mathbf{B}(\mathbf{x}_{j \pm \frac{1}{2}\mathbf{e}_\ell}, t^{n+1})), \quad \forall j \in \mathfrak{J}, \forall \ell = 1, \dots, d.$$

### Time-stepping

With the temporal partition above-stated, let  $\mathbf{U}_j^n$  denote the cell average (6.8) on  $T_j$  at time  $t^n$ , i.e.,  $\mathbf{U}_j^n := \mathbf{U}_j(t^n)$ . A piecewise constant finite volume discretization in space can be combined with the first order explicit Euler time-stepping to deliver the approximation

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \sum_{\ell=1}^d \frac{\Delta t^n}{h_\ell} (F_{j+\frac{1}{2}\mathbf{e}_\ell}^{\ell, n} - F_{j-\frac{1}{2}\mathbf{e}_\ell}^{\ell, n}), \quad \forall j \in \mathfrak{J},$$

where  $F_{j+\frac{1}{2}\mathbf{e}_\ell}^{\ell, n}$  is a shorthand notation for  $F_{j+\frac{1}{2}\mathbf{e}_\ell}^{\ell, n}(\mathbf{B}(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}, t^n)) =: F^\ell(\mathbf{U}_j^n, \mathbf{U}_{j+\mathbf{e}_\ell}^n, \mathbf{B}_{j+\frac{1}{2}\mathbf{e}_\ell}^n)$  and

$$F^\ell(\mathbf{U}, \mathbf{U}, \mathbf{B}_{j \pm \frac{1}{2}\mathbf{e}_\ell}) = f^\ell(\mathbf{U}, \mathbf{B}_{j \pm \frac{1}{2}\mathbf{e}_\ell}), \quad \forall \mathbf{U} \in \mathbb{R}^m, \quad \forall j \in \mathfrak{J}, \ell = 1, \dots, d,$$

that is  $F^\ell$  is a numerical flux consistent with the directional flux  $f^\ell$  for all  $1 \leq \ell \leq d$ . The global approximate solution  $\mathbf{U}_h$  is obtained as  $\mathbf{U}_h(\mathbf{x}, t) = \sum_{n,j} \mathbb{1}_{T_j}(\mathbf{x}, t) \mathbb{1}_{[t^n, t^{n+1})}(\mathbf{x}, t) \mathbf{U}_j^n$ .

Higher order temporal discretizations can be achieved using *strong-stability preserving* (SSP) Runge–Kutta schemes [GST01] (first called total variation diminishing time discretizations). The family of explicit SSP-RK temporal schemes achieves strong stability in a certain norm,  $\|\mathbf{U}^{n+1}\| \leq \|\mathbf{U}^n\|$ , under a suitable CFL condition, provided stability holds for the forward Euler time-stepping (possibly under a different CFL-type restriction on the time step).

We first restate the system of conservation laws (6.3) in the semi-discrete form

$$\partial_t \mathbf{U}_j(t) = - \sum_{\ell=1}^d \frac{1}{h_\ell} (F_{j+\frac{1}{2}\mathbf{e}_\ell}^\ell(t) - F_{j-\frac{1}{2}\mathbf{e}_\ell}^\ell(t)), \quad \forall j \in \mathfrak{J},$$

as the finite dimensional operator evolution equation analogous to (3.35),

$$\partial_t \mathbf{U}(t) + L_h(\mathbf{U}(t)) = \mathbf{0}, \quad \text{with } \mathbf{U}(t) = \sum_{j \in \mathfrak{J}} \mathbb{1}_{T_j} \mathbf{U}_j(t).$$

A general  $s$ -stage SSP Runge–Kutta method can be written in the Shu–Osher form [SO88] for each component of the vector-valued solution  $\mathbf{U}(t) = \{U_b(t)\}_{b=1}^m$  as,

$$\begin{aligned} U_b^{n;0} &= U_b^n, \\ U_b^{n;i} &= \sum_{j=0}^{i-1} (\alpha_{i,j} U_b^{n;j} + \Delta t^n \delta_{i,j} L_h(U_b^{n;j})), \quad i = 1, \dots, s, \\ U_b^{n+1} &= U_b^{n;s}, \end{aligned} \tag{6.10}$$

with coefficients  $\alpha_{i,j}, \delta_{i,j} \in \mathbb{R}$  for  $1 \leq i \leq s$  and  $0 \leq j \leq i-1$ . If  $\alpha_{i,j}, \delta_{i,j} \geq 0$  with  $\alpha_{i,j} \neq 0$  whenever  $\delta_{i,j} \neq 0$  and  $\sum_{j=0}^{i-1} \alpha_{i,j} = 1$  for all  $1 \leq i \leq s$ , then the intermediate states  $U_b^{n;i}$  are convex combinations of explicit Euler steps with time step  $\delta_{i,j}/\alpha_{i,j} \Delta t^n$ . Hence, if the explicit Euler step is stable i.e.  $\|\mathbf{U}^n + \Delta t^n L_h(\mathbf{U}^n)\| \leq \|\mathbf{U}^n\|$  in a given norm and for  $\Delta t^n \leq Ch$ , then the  $s$ -stage RK scheme (6.10) is stable under the CFL condition  $\Delta t^n \leq Cc(\alpha, \delta)h$ , where  $c(\alpha, \delta) := \min_{i,j} (\delta_{i,j}/\alpha_{i,j})$ . An optimal (in terms of CFL condition,  $c(\alpha, \delta) = 1$ ) two-stage second order SSP-RK scheme is as in (3.45)-(3.46) (with zero forcing term), namely

$$\begin{aligned} U_b^{n;1} &= U_b^n - \Delta t^n L_h(U_b^n), \\ U_b^{n+1} &= \frac{1}{2}(U_b^n + U_b^{n;1}) - \frac{1}{2}\Delta t^n L_h(U_b^{n;1}), \end{aligned}$$

and an optimal three-stage third order SSP-RK scheme reads

$$\begin{aligned} U_b^{n;1} &= U_b^n - \Delta t^n L_h(U_b^n), \\ U_b^{n;2} &= \frac{1}{4}(3U_b^n + U_b^{n;1}) - \frac{1}{4}\Delta t^n L_h(U_b^{n;1}), \\ U_b^{n+1} &= \frac{1}{3}(U_b^n + 2U_b^{n;2}) - \frac{2}{3}\Delta t^n L_h(U_b^{n;2}). \end{aligned}$$

For sufficiently smooth solutions of scalar conservation laws and symmetrizable systems, convergence analyses in Sobolev norms of the fully discrete scheme with discontinuous Galerkin spatial discretizations can be carried out similarly to Section 3.4 and with analogous restrictions on the time step (see [ZS04] and [ZS06]).

Concerning the choice of the local time step, we adopt a perspective where the extended Euler equations are considered as embedded in the physics of the full ideal MHD system. Since in MHD the fast magnetosonic waves propagate on a time scale much faster than the fluid velocity, these waves dictate

the time step restrictions necessary for a stable numerical update. For the temporal discretization of the extended Euler system, the  $n$ -th time step is taken to be

$$\Delta t^n = C_{\text{CFL}} \left( \sum_{\ell=1}^d \frac{\lambda_{\max}^\ell}{h_\ell} \right)^{-1}, \quad (6.11)$$

where  $C_{\text{CFL}} > 0$  is a constant and  $\{\lambda_{\max}^\ell\}_{\ell=1}^d$  are the maximum eigenvalues associated with the fast magnetosonic waves (1.12), namely

$$\lambda_{\max}^\ell = \max_{j \in J} \lambda_f^{M,\ell}(\mathbf{U}_j^n, \mathbf{B}_j^n) := \max_{j \in J} (|u_j^{\ell,n}| + c_f^{M,\ell}(\mathbf{U}_j^n, \mathbf{B}_j^n)), \quad \ell = 1, \dots, d, \quad (6.12)$$

$$c_f^{M,\ell} = \frac{1}{\sqrt{2}} \sqrt{a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho} + \sqrt{\left(a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho}\right)^2 - 4a^2 \frac{B_\ell^2}{\rho}}}, \quad \ell = 1, \dots, d.$$

### 6.2.1. Approximate Riemann Solvers

In finite volume discretizations of conservation laws, exact solutions of the Riemann problems yield the so-called Godunov scheme. A major drawback of Godunov-type methods is related to exactly solving the nonlinear Riemann problem, especially in the case of systems of conservation laws in multi-dimensions: exact solutions are computationally expensive and largely unavailable (under uniqueness conditions an exact Riemann solver for ideal MHD was derived in [Tor02]). Moreover, Godunov fluxes are already an approximation of the exact fluxes. Hence, linear (Roe-type [Roe81]) and nonlinear approximate Riemann solvers have become standard in finite volume schemes.

Brio and Wu [BW88] pioneered the use of Roe-type linearized Riemann solver for MHD with gas constant  $\gamma = 2$ , and extension of the Roe-type matrix and corresponding Roe-average for general  $\gamma$  was developed in [CG97]. Practically, linearized solvers are often based on simply averaging the Jacobians of the two states across each interface. However, the cumbersome computational cost and the failure of providing positivity preserving schemes [Ein+91] has thwarted the extensive application of linearized Roe-type solver in MHD simulations.

An attractive alternative is provided by reduced-wave nonlinear solvers (the so-called *HLL solvers*) which approximate the wave structure of the full Riemann problem by a simplified set of known waves. The HLL solvers are prominent in marrying limited computational cost, accurate reproduction of the physical features of the flow, and robustness of the resulting numerical scheme. On the other hand, no HLL solver has the resolution of the Roe solver. We present HLL-type approximate Riemann solvers for the extended Euler equations based on solvers developed for the ideal MHD system and inspired by [FMR09]. Each Riemann solver is specified only in the  $x := x^1$ -direction; modifications in other directions are derived likewise. For the sake of better readability, we switch from the multi-index  $j$  to the index  $i$  treating the problem as one-dimensional and we omit the superscript  $\ell = 1$ .

#### Two-wave HLL Solver

The Harten–Lax–van Leer (HLL) approximate Riemann solver, introduced in [HLL83] for the inviscid gas dynamic equations, assumes a wave configuration for the Riemann problem solution consisting of three constant states separated by two shock waves moving to the left and to the right of the interface. Possible intermediate states are lumped into the middle constant one, so that the Riemann solution has the form

$$\mathbf{U}_{i+1/2}(\mathbf{x}, t) = \begin{cases} \mathbf{U}_i & \text{if } x < s_{i+1/2}^L t, \\ \mathbf{U}_{i+1/2}^* & \text{if } s_{i+1/2}^L t < x < s_{i+1/2}^R t, \\ \mathbf{U}_{i+1} & \text{if } x > s_{i+1/2}^R t. \end{cases}$$

The selection of the left  $s_{i+1/2}^L$  and right  $s_{i+1/2}^R$  acoustic wave speeds determines different variants of the approximate flux, which will be discussed thereafter. The resulting HLL flux solver at time  $t^n$  is

$$F_{i+1/2}^{n,\text{HLL}} = F(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) = \begin{cases} f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) & \text{if } s_{i+1/2}^L > 0, \\ f_{i+1/2}^{*,\text{HLL}} & \text{if } s_{i+1/2}^L < 0 < s_{i+1/2}^R, \\ f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) & \text{if } s_{i+1/2}^R < 0. \end{cases}$$

Note that the middle flux  $f_{i+1/2}^{*,\text{HLL}}$  is generally different from  $f(\mathbf{U}_{i+1/2}^*, \mathbf{B}_{i+1/2}^*)$ . It is determined, together with the intermediate subsonic state  $\mathbf{U}_{i+1/2}^*$ , by applying local conservation through Rankine–Hugoniot conditions, namely

$$\begin{aligned} f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) - f_{i+1/2}^{*,\text{HLL}} &= s_{i+1/2}^R (\mathbf{U}_{i+1}^n - \mathbf{U}_{i+1/2}^*), \\ f_{i+1/2}^{*,\text{HLL}} - f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) &= s_{i+1/2}^L (\mathbf{U}_{i+1/2}^* - \mathbf{U}_i^n), \end{aligned}$$

which results in

$$\begin{aligned} f_{i+1/2}^{*,\text{HLL}} &= \frac{s_{i+1/2}^R f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) - s_{i+1/2}^L f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) + s_{i+1/2}^R s_{i+1/2}^L (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)}{s_{i+1/2}^R - s_{i+1/2}^L}, \\ \mathbf{U}_{i+1/2}^* &= \frac{s_{i+1/2}^R \mathbf{U}_{i+1}^n - s_{i+1/2}^L \mathbf{U}_i^n + f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) - f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n)}{s_{i+1/2}^R - s_{i+1/2}^L}. \end{aligned} \quad (6.13)$$

Note that the computation of the flux  $f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n)$  at the interface requires the knowledge of both the tangential and normal components of the  $\mathbf{B}$  field at the interface. Whenever one of these quantities is not well-defined, we will take some average  $\{B^1\}_{i+1/2}$  and/or  $\{B^2\}_{i+1/2}$  across the interface.

Choosing opposite wave speeds  $s_{i+1/2}^L = -s_{i+1/2}^R$ , the expression for the intermediate flux simplifies to

$$f_{i+1/2}^{*,\text{HLL}} = \frac{f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) + f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n)}{2} - \frac{s_{i+1/2}^R}{2} (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n).$$

### 1. Lax–Friedrichs flux

Using the maximum wave speeds compatible with the CFL condition, namely  $s_{i+1/2}^R = h/\Delta t^n$  and  $s_{i+1/2}^L = -s_{i+1/2}^R$ , yields the Lax–Friedrichs flux

$$F_{i+1/2}^{n,\text{LxF}} := \frac{f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) + f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n)}{2} - \frac{h}{2\Delta t^n} (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n).$$

### 2. Rusanov flux

Locally adapted opposite wave speeds  $s_{i+1/2}^R = \max\{|f'(\mathbf{U}_i^n, \mathbf{B}_i^n)|, |f'(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1}^n)|\}$  and  $s_{i+1/2}^L = -s_{i+1/2}^R$  give the Rusanov (or local Lax–Friedrichs) flux

$$\begin{aligned} F_{i+1/2}^{n,\text{Rus}} &:= \frac{f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) + f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n)}{2} \\ &\quad - \frac{\max\{|f'(\mathbf{U}_i^n, \mathbf{B}_i^n)|, |f'(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1}^n)|\}}{2} (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n). \end{aligned}$$

For the extended Euler system within the MHD equations, the artificial diffusion term in the Rusanov flux is computed using the eigenvalue associated with the fast wave of the full MHD equations, i.e.,

$$\max\{|f'(\mathbf{U}_i^n, \mathbf{B}_i^n)|, |f'(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1}^n)|\} = \max\{\lambda_f^{\text{M},1}(\mathbf{U}_i^n, \mathbf{B}_i^n), \lambda_f^{\text{M},1}(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1}^n)\},$$

with  $\lambda_f^{\text{M},1}(\mathbf{U}_i^n, \mathbf{B}_i^n) := |u_i^{1,n}| + c_f^{\text{M},1}(\mathbf{U}_i^n, \mathbf{B}_i^n)$  and  $c_f^{\text{M},1}$  as in (1.12).

### 3. Einfeldt wave speeds

Similarly to the HLL solver for the gas dynamic problem proposed in [HLL83], right and left speeds can be obtained as Lipschitz continuous approximations of the smallest and largest wave speeds, respectively. However, Einfeldt suggested in [Ein88] to use the minimum and maximum eigenvalues of a Roe-average to restrain the spreading of contact discontinuities and increase the resolution at isolated shocks:

$$\begin{aligned} s_{i+1/2}^L &:= \min\{u_i^{1,n} - a^{\text{E},1}(\mathbf{V}_i^n, \mathbf{B}_i^n), \{u^{1,n}\}_{i+1/2} - \hat{a}^{\text{E},1}(\{\mathbf{V}^n\}_{i+1/2}, \{\mathbf{B}^n\}_{i+1/2})\}, \\ s_{i+1/2}^R &:= \max\{u_{i+1}^{1,n} + a^{\text{E},1}(\mathbf{V}_{i+1}^n, \mathbf{B}_{i+1}^n), \{u^{1,n}\}_{i+1/2} + \hat{a}^{\text{E},1}(\{\mathbf{V}^n\}_{i+1/2}, \{\mathbf{B}^n\}_{i+1/2})\}. \end{aligned} \quad (6.14)$$

The speed  $a^{\text{E},1}$  is defined as in (6.4) whereas  $\hat{a}^{\text{E},1}$  is the “extended” sound speed

$$\begin{aligned} \{\rho^n(a^{\text{E},1})^2\}_{i+1/2} &= \hat{a}^{\text{E},1}(\{\mathbf{V}^n\}_{i+1/2}, \{\mathbf{B}^n\}_{i+1/2}) \\ &= \sqrt{\frac{1}{\rho_i^n} + \frac{1}{\rho_{i+1}^n}} \sqrt{\gamma(p_i^n + p_{i+1}^n) + (\gamma - 1)[(B_2^n)_i^2 + (B_2^n)_{i+1}^2 + (B_3^n)_i^2 + (B_3^n)_{i+1}^2]}, \end{aligned}$$

associated with the averaged directional Jacobian

$$\widehat{A}_{i+1/2}^1 = \begin{pmatrix} \{u^{1,n}\}_{i+1/2} & \{\rho^n\}_{i+1/2} & 0 & 0 & 0 \\ 0 & \{u^{1,n}\}_{i+1/2} & 0 & 0 & \{1/\rho^n\}_{i+1/2} \\ 0 & 0 & \{u^{1,n}\}_{i+1/2} & 0 & 0 \\ 0 & 0 & 0 & \{u^{1,n}\}_{i+1/2} & 0 \\ 0 & \{\rho^n(a^{\text{E},1})^2\}_{i+1/2} & \{\rho^n(a^{\text{E},2})^2\}_{i+1/2} & \{\rho^n(a^{\text{E},3})^2\}_{i+1/2} & \{u^{1,n}\}_{i+1/2} \end{pmatrix}.$$

Despite the efficiency and robustness of HLL-type Riemann solvers, the two-wave configuration hinders the resolution of physical features, in particular Alfvén and slow waves and contact discontinuities, yielding overdiffusive solutions. Contact discontinuities are “restored” in the structure of the Riemann solution with the modified HLL solver introduced by Toro, Spruce and Speares [TSS94] and dubbed HLLC.

### Three-wave HLLC Solver

The HLLC solver approximates the Riemann solution by three waves allowing for two intermediate states. The fast magneto-sonic waves are modeled as in the HLL solver, whilst the intermediate states are separated by a wave moving with speed  $s_{i+1/2}^M$  and modeling a contact discontinuity (associated with the multiple eigenvalue  $\lambda^{2,3,4}$  (6.5)). The HLLC numerical flux is defined as

$$F_{i+1/2}^{\text{HLLC}} = F(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) = \begin{cases} f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) & \text{if } s_{i+1/2}^L > 0, \\ f_L^* & \text{if } s_{i+1/2}^L < 0 < s_{i+1/2}^M, \\ f_R^* & \text{if } s_{i+1/2}^M < 0 < s_{i+1/2}^R, \\ f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) & \text{if } s_{i+1/2}^R < 0, \end{cases}$$

and the Riemann problem solution reads

$$\mathbf{U}_{i+1/2}(\mathbf{x}, t) = \begin{cases} \mathbf{U}_i & \text{if } x < s_{i+1/2}^L t, \\ \mathbf{U}_L^* & \text{if } s_{i+1/2}^L t < x < s_{i+1/2}^M t, \\ \mathbf{U}_R^* & \text{if } s_{i+1/2}^M t < x < s_{i+1/2}^R t, \\ \mathbf{U}_{i+1} & \text{if } x > s_{i+1/2}^R t. \end{cases}$$

The left and right speeds model the fast magneto-sonic waves and are as in (6.14), the middle wave speed  $s_{i+1/2}^M$  is the velocity of the averaged Jacobian  $s_{i+1/2}^M = \{u^{1,n}\}_{i+1/2}$  since it models the contact discontinuity. The intermediate fluxes are determined by applying local conservation through the Rankine–Hugoniot conditions

$$\begin{aligned} s_{i+1/2}^L \mathbf{U}_L^* - f_L^* &= s_{i+1/2}^L \mathbf{U}_i^n - f(\mathbf{U}_i, \mathbf{B}_{i+1/2}^n), \\ s_{i+1/2}^M \mathbf{U}_L^* - f_L^* &= s_{i+1/2}^M \mathbf{U}_R^* - f_R^*, \\ s_{i+1/2}^R \mathbf{U}_{i+1}^n - f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) &= s_{i+1/2}^R \mathbf{U}_R^* - f_R^*, \end{aligned} \quad (6.15)$$

where  $\mathbf{U}_R^*$  and  $\mathbf{U}_L^*$  denote the right and left intermediate states, respectively. Simple algebraic manipulations yield the intermediate fluxes

$$\begin{aligned} f_R^* &= f_{i+1/2}^{*,\text{HLL}} + \frac{s_{i+1/2}^R(s_{i+1/2}^M - s_{i+1/2}^L)}{s_{i+1/2}^R - s_{i+1/2}^L} (\mathbf{U}_R^* - \mathbf{U}_L^*), \\ f_L^* &= f_{i+1/2}^{*,\text{HLL}} - \frac{s_{i+1/2}^L(s_{i+1/2}^R - s_{i+1/2}^M)}{s_{i+1/2}^R - s_{i+1/2}^L} (\mathbf{U}_R^* - \mathbf{U}_L^*). \end{aligned} \quad (6.16)$$

However, since the system (6.15) is underdetermined, a further constraint on the intermediate states needs to be imposed.

### 1. Linde’s “adequate” solver

The constraint introduced in [Lin98, Section 4.3.3] characterizes the jump of the intermediate states as a non-negative fraction of the initial jump across the middle wave, namely  $\mathbf{U}_R^* - \mathbf{U}_L^* = \alpha(\mathbf{U}_i^n - \mathbf{U}_{i+1}^n)$  for  $\alpha \in [0, 1]$ . Indeed, the jump across an isolated shock is  $\mathbf{U}_i^n - \mathbf{U}_{i+1}^n$ , while it vanishes for any other type of wave. This choice aims at robustness of the Riemann solver together with good resolution of isolated shocks. For the purpose of detecting the isolated discontinuities and hence computing  $\alpha$ , a certain norm of the quantity  $f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) - f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) - s_{i+1/2}^M(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)$  would give information about the initial Riemann data and allows to distinguish the two limit cases  $\alpha = 0$  and  $\alpha = 1$ . Indeed, by the Rankine–Hugoniot conditions and the definition of the intermediates speed,  $f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) - f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) = s_{i+1/2}^M(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)$  holds in the presence of an initial contact wave. In this case  $\alpha$  has to be 1. On the other hand, in the presence of a fast shock  $|f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) - f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) - s_{i+1/2}^M(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)| = c_* |\mathbf{U}_{i+1}^n - \mathbf{U}_i^n|$ , with  $c_* := |a^{E,1}(V_i^n, \mathbf{B}_i^n) - s_{i+1/2}^M|$ . Hence,

$$\alpha := \max \left\{ 0, 1 - \frac{s}{c_*} \right\}, \quad s := \frac{\|f(\mathbf{U}_{i+1}^n, \mathbf{B}_{i+1/2}^n) - f(\mathbf{U}_i^n, \mathbf{B}_{i+1/2}^n) - s_{i+1/2}^M(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)\|_{\ell^1}}{\|\mathbf{U}_{i+1}^n - \mathbf{U}_i^n\|_{\ell^1}}.$$

As pointed out in [Lin98],  $s$  gives an indication of the speed of the dominant wave in the frame of reference of the middle wave. Note that  $s$  can be in principle computed using a different norm. Finally, with an analytic expression for the weight  $\alpha$ , the intermediate fluxes can be determined as in (6.16).

### 2. Gurski’s solver

Let  $\mathbf{U}_R^* = (\rho_R^*, (\rho\mathbf{u})_R^*, E_R^*)$  and analogously for  $\mathbf{U}_L^*$ . In order to determine the intermediate states, the Rankine–Hugoniot conditions across a contact discontinuity entail constant normal velocity  $\mathbf{u}_L^* \cdot \mathbf{n} = \mathbf{u}_R^* \cdot \mathbf{n} = s_{i+1/2}^M$  and total pressure (sum of fluid and magnetic pressure). Gurski [Gur04] extended to ideal MHD the idea in [Bat+97] of deriving the middle wave speed  $s_{i+1/2}^M$  from (6.15) through the HLL middle state formula (6.13). Specifically, using the local conservation of mass and momentum (in the  $x$ -direction) results in

$$\begin{aligned} s_{i+1/2}^M &= \frac{(\rho^n u^{1,n})^*}{(\rho^n)^*} = \frac{\rho_i^n u_i^{1,n} (u_i^{1,n} - s_{i+1/2}^L) - \rho_{i+1}^n u_{i+1}^{1,n} (u_{i+1}^{1,n} - s_{i+1/2}^R)}{\rho_i^n (u_i^{1,n} - s_{i+1/2}^L) - \rho_{i+1}^n (u_{i+1}^{1,n} - s_{i+1/2}^R)} \\ &\quad + \frac{p_i^n - p_{i+1}^n + ((\|\mathbf{B}_i^n\|_{\ell^2} - \|\mathbf{B}_{i+1}^n\|_{\ell^2})/2 + (\mathbf{B}_{i+1}^n \cdot \mathbf{n})^2 - (\mathbf{B}_i^n \cdot \mathbf{n})^2)}{\rho_i^n (u_i^{1,n} - s_{i+1/2}^L) - \rho_{i+1}^n (u_{i+1}^{1,n} - s_{i+1/2}^R)}. \end{aligned}$$

The intermediate conserved variables can be determined through (6.15) using the definition of  $s^M$ . However, since  $\mathbf{U}_{R/L}^*$  depends on  $f(\mathbf{U}_{R/L}^*)$ , it is not straightforward how the magnetic induction field  $\mathbf{B}$  enters the intermediate fluxes. One possibility would be to simply use the information from the neighboring elements  $\mathbf{B}_i^n$ ,  $\mathbf{B}_{i+1}^n$  and from the interface  $\mathbf{B}_{i+1/2}^n$ . Alternatively, in the spirit of Riemann solvers for the MHD system, one could consider intermediates states also for the  $\mathbf{B}$  field. The computation of the intermediate states  $\mathbf{B}_R^*$  and  $\mathbf{B}_L^*$  require the introduction of “artificial” fluxes which could be provided by the conservation law for  $\mathbf{B}$ :  $\partial_t \mathbf{B} + \operatorname{div}(\mathbf{B} \otimes \mathbf{u} - \mathbf{u} \otimes \mathbf{B}) = 0$ . This ambiguity deterred us from using Gurski’s HLLC solver in the numerical simulations for the full ideal MHD equations in Chapter 7.

The two-wave solvers for the equations of gas dynamics and based on a suitable choice of acoustic and contact wave speeds yield positivity preserving and entropy stable schemes [Ein88; Bat+97]. Furthermore, they provide exact resolution of isolated shock but cannot exactly resolve isolated contact discontinuities and consequently are too dissipative. The HLLC solvers for the ideal MHD Riemann problem resolve isolated shock and contact waves exactly, albeit their numerical resolution is not comparable with that of linearized solvers except for fast and entropy waves. Linde’s scheme is easy to implement but lacks positivity preserving properties and a sound definition of the weight  $\alpha$ . On the other hand, Gurski’s solver has been proven to be positivity preserving (under some mild conditions on the wave speeds, see [Gur04, Section 4.1] for details), although computational results [FMR09, Remark p. 648] show that it is slightly more dissipative than Linde’s solver. It seems that there is no agreement on an optimal HLL solver: Several other versions of three-wave HLL schemes for ideal MHD have been developed, e.g. in [BKW07] (the so-called relaxed HLLC solver), in [MK05] (the HLLD solver) and in [Li05].

### Limitations of Finite Volume Schemes for the Extended Euler System

Lacking well-posedness and stability results for multi-dimensional systems of conservation laws, convergence estimates for finite volume approximations of the MHD problems and general nonlinear systems are not available, not even in one dimension. Nevertheless, it would be desirable to construct numerical discretizations able to capture the structure of the underlying continuum MHD model, in particular in terms of waves structure, conservation of admissible states and entropy stability.

Concerning the first aspect, the formulation of approximate Riemann solvers for the extended Euler equations based on HLL-type solvers for the MHD Riemann problem is indeed an attempt to capture the wave structure (6.6) viewed as a subset of the ideal MHD wave fan. However, as pointed out in [FMR09], the lack of control on the  $\mathbf{B}$  field and the unavailability of a Roe-average, which consequently affects the choice of the wave speeds, does not guarantee that the resulting scheme is able to exactly capture fast magnetosonic shocks or isolated contact discontinuities. Analogously, none of the foregoing solvers for the extended Euler system is provably positively conservative. We were not able to show such property for  $\mathbf{B}$  fields resulting from the stabilized Galerkin discretization (3.7), (3.8), (3.9) nor from the extrusion contraction approximation (4.7) of the magnetic advection problem for the magnetic field/potential. Furthermore, the presence of the  $\mathbf{B}$  field as a discontinuous parameter entering the fluxes impinges on the design of numerical fluxes for the extended Euler equations tailored to satisfy a discrete version of the entropy inequality  $\partial_t(\rho s) + \operatorname{div}(\rho \mathbf{u} s) \leq 0$  with the thermodynamic entropy defined as  $s := \log(p) - \gamma \log(\rho)$ .

These limitations hamper the construction of arbitrarily high order robust finite volume schemes for the extended Euler system, see also Section 6.3.

#### 6.2.2. Numerical Experiments in 2D

##### Accuracy Test: Smooth Vortex

To experimentally gauge the performances of lowest order finite volume schemes for the extended Euler system, we propose a two-dimensional MHD test with given magnetic induction field [Bal04, Section 6]. More in details, the solution is known analytically at any point in space and time in the domain

$\Omega \times I = [-5, 5]^2 \times [0, 0.5]$ , see Section 7.1.2 for further details. Let  $r(x, y, t) := \sqrt{(x - u_0^1 t)^2 + (y - u_0^2 t)^2}$ , with  $\mathbf{u}_0 = (u_0^1, u_0^2)^\top = (1, 1)^\top$ , the flow is characterized by the following set of data,

$$\begin{aligned}\rho(x, y, t) &= 1, \\ p(x, y, t) &= 1 + \frac{1}{8\pi}(\mu^2(1 - r^2) - \kappa^2)e^{1-r^2}, \\ \mathbf{u}(x, y, t) &= \mathbf{u}_0 + \frac{\kappa}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top, \\ \mathbf{B}(x, y, t) &= \frac{\mu}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top.\end{aligned}$$

As initial datum we take  $\mathbf{U}(\mathbf{x}, 0)$  and parameters  $\kappa = \mu = 1$ . The ratio of specific heats is  $\gamma = 5/3$ . The magnetic induction field  $\mathbf{B}$  is given at each time step in analytic form. We aim at assessing the convergence properties of the scheme. Explicit Euler time-stepping ( $C_{\text{CFL}} = 0.4$ ) is coupled with a piecewise constant finite volume discretization in space and tested with different approximate Riemann solvers. In Figure 6.1, the  $L^2$ -error,  $L^1$ -error and  $L^\infty$ -error of the primitive variables at the final time  $T = 0.5$  are reported. As expected, first order convergence is observed.

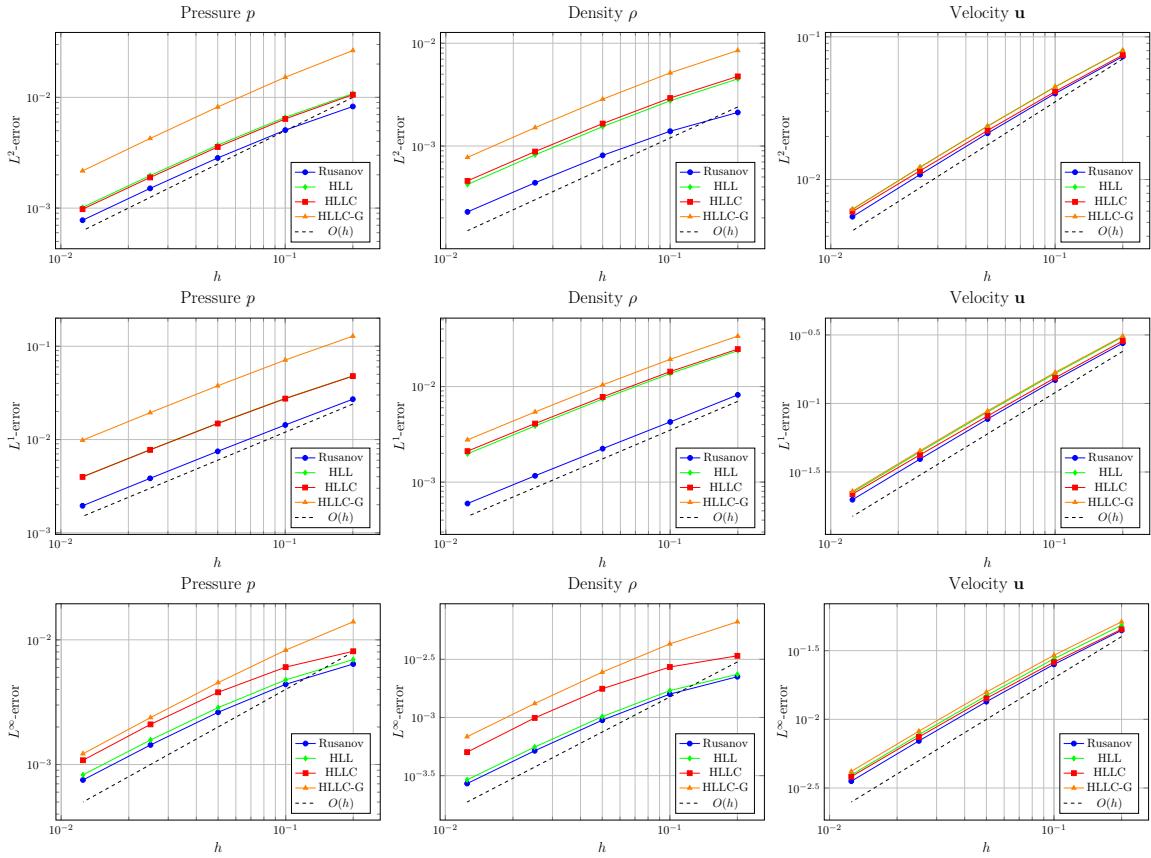


Figure 6.1.: Smooth vortex. Accuracy test for finite volume discretizations of the extended Euler system with given analytic  $\mathbf{B}$  field.  $L^2$ -error (top row),  $L^1$ -error (middle row),  $L^\infty$ -error (bottom row). Different approximate Riemann solvers are considered. In the legend, HLL refers to the two-wave Riemann solver with Einfeldt wave speeds, HLLC to the three-wave Linde solver, and HLLC-G denotes the three-wave Gurski solver.

### Orszag–Tang Benchmark with Given B Field

As a second test case we consider the Orszag–Tang benchmark for ideal MHD [OT79]. We refer to Section 7.1.4 for more details on the test features. On the domain  $\Omega \times I = [0, 2]^2 \times [0, 1]$  with periodic boundary conditions, we run a simulation of the extended Euler system discretized with a piecewise

constant finite volume scheme and SSP-RK2 as time-stepping with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ . The magnetic induction field  $\mathbf{B}$  is given at each time step as the output of a high order finite volume discretization of the full MHD equations implemented in the ALSVID-UQ code (see Section 3.6.3 for further details). The primitive variables at time  $t = 0$  are  $\rho = \gamma^2$ ,  $p = \gamma$ ,  $u^1(x, y, 0) = -\sin(\pi y)$ ,  $u^2(x, y, 0) = \sin(\pi x)$ . The ratio of specific heats is  $\gamma = 5/3$  and  $\mathbf{B}_0 = (-\sin(\pi y), \sin(2\pi x))^\top$ .

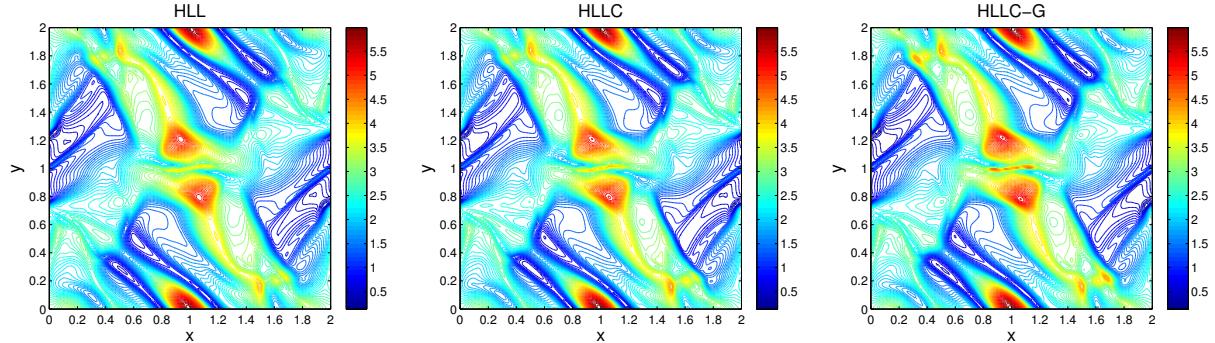


Figure 6.2.: Orszag–Tang benchmark with given  $\mathbf{B}$  field. Numerical pressure obtained on a  $200 \times 200$  Cartesian mesh. Heun time-stepping with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ . Piecewise constant finite volume discretization in space with HLL (left) fluxes, HLLC (center) solver with Linde’s scheme, Gurski’s HLLC solver (right).

In Figure 6.2 we report the isolines plots of the numerical pressure obtained with lowest order finite volume spatial discretization and the different approximate Riemann solvers introduced in Section 6.2.1. Figure 6.3 shows two projections of the discrete density onto the one-dimensional interval  $[0, 2]$  for constant values of  $y = 1.425$  and  $x = 0.465$ , respectively. Two main features are conspicuous when comparing the outcome of the finite volume discretization of the extended Euler system with the reference solution (same resolution but better accuracy): the lowest order methods for the extended Euler system are rather diffusive independently of the Riemann solver employed in the fluxes approximation; moreover the three-wave solvers deliver better results than the two-wave HLL solvers (Rusanov and HLL with Einfeldt speeds). However, a precise comparison of the ability of each Riemann solver to provide approximations able to reproduce the characteristic structures of the MHD solution is not a straightforward task. Indeed we conjecture that the nuances in the full MHD wave structure cannot be genuinely captured by the different Riemann solvers for the extended Euler equations, owing to the fact that magnetic induction field  $\mathbf{B}$  is imposed at each time step.

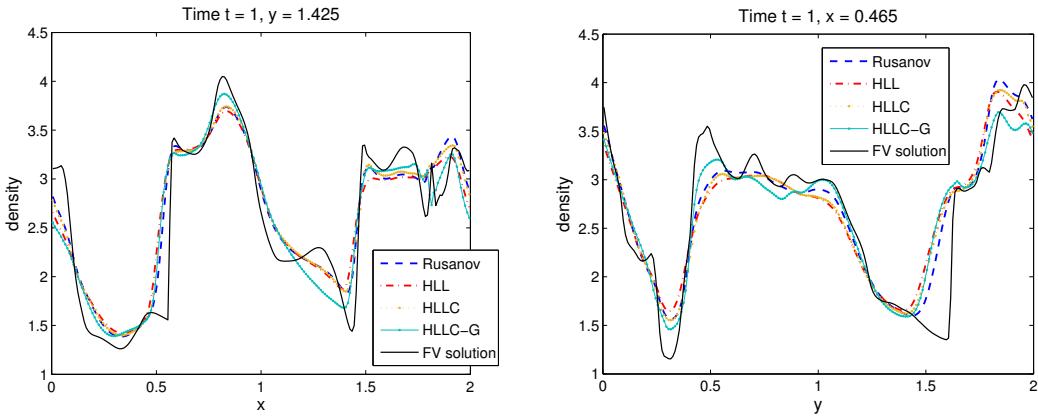


Figure 6.3.: Orszag–Tang benchmark with given  $\mathbf{B}$  field. Projections of the numerical density obtained on a  $200 \times 200$  Cartesian mesh. Heun time-stepping with uniform time step  $\Delta t = 5 \cdot 10^{-4}$ . Piecewise constant finite volume discretization in space with Rusanov, HLL solvers with Einfeldt speeds, HLLC fluxes with Linde’s scheme, and Gurski’s HLLC solver.

### 6.3. High Order Schemes

Second and higher order accuracy can be achieved by combining high order SSP-RK time-stepping with spatial finite volume discretizations based on suitable reconstructions, i.e. piecewise polynomial approximations of the solution from its cell averages.

#### The REA Algorithm

Three major steps can be singled out in Godunov-type schemes:

1. **Reconstruction.** At time  $t^n$ , the cell averages  $\{\mathbf{U}_j^n\}_{j \in \mathfrak{J}}$  are *reconstructed* in each direction into a  $\mathcal{T}_h$ -piecewise  $\mathbb{R}^m$ -valued polynomial function  $\mathbf{p}^n \in \mathcal{P}_{\tau_\ell}(T_j)^m$  for all  $j \in \mathfrak{J}$ , where  $\tau_\ell \in \mathbb{N}^d$  specifies the polynomial degree in each direction and whose  $l$ th-component is  $(\tau_\ell)_l := r\delta_{\ell,l}$  for  $r \geq 0$  and  $\delta_{\ell,l}$  denoting the Kronecker delta. The reconstruction must be of order  $O(h^{r+1})$  and conservative, i.e., the cell averages of  $\mathbf{p}^n$  on each  $T_j \in \mathcal{T}_h$  equal  $\mathbf{U}_j^n$ . In the lowest order approximation ( $r = 0$ ),  $\mathbf{p}^n$  is the  $\mathcal{T}_h$ -piecewise constant function obtained as  $\mathbf{p}^n(\mathbf{x}) = \mathbf{U}_j^n$  for all  $\mathbf{x} \in T_j$  and all  $j \in \mathfrak{J}$ .
2. **Evolution.** The system of conservation laws (6.3) is solved, exactly or approximately, with initial datum  $\mathbf{p}^n(\mathbf{x})$  yielding the update  $\mathbf{p}^{n+1}(\mathbf{x})$ . In the approximate Riemann solver policy, the traces of the reconstructed function  $\mathbf{p}^n(\mathbf{x})$  at the mesh interfaces enter the approximate fluxes  $F_{j+\frac{1}{2}\mathbf{e}_\ell}^{\ell,n} = F^\ell(\widehat{\mathbf{U}}_j^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}), \widehat{\mathbf{U}}_{j+\mathbf{e}_\ell}^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}), \mathbf{B}_{j+\frac{1}{2}\mathbf{e}_\ell}^n)$ , for  $1 \leq \ell \leq d$ , where,
$$\widehat{\mathbf{U}}_j^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}) := \lim_{\varepsilon \searrow 0} \mathbf{p}^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell} - \varepsilon), \quad \text{and} \quad \widehat{\mathbf{U}}_{j+\mathbf{e}_\ell}^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell}) = \lim_{\varepsilon \searrow 0} \mathbf{p}^n(\mathbf{x}_{j+\frac{1}{2}\mathbf{e}_\ell} + \varepsilon), \quad \forall j \in \mathfrak{J}.$$
3. **Averaging.** The solution  $\mathbf{p}^{n+1}(\mathbf{x})$  at time  $t^{n+1}$  is averaged to obtain the updated cell-averages  $\{\mathbf{U}_j^{n+1}\}_{j \in \mathfrak{J}}$ .

Roughly speaking, the accuracy of the reconstruction dictates the order of the finite volume scheme. Other than the classical piecewise linear reconstruction with minmod limiting, we consider second-order ENO and WENO reconstruction techniques tailored to finite volume discretizations of the extended Euler equations. Essentially non-oscillatory (ENO) [Har+87] and weighted essentially non-oscillatory (WENO) [SO89] reconstructions are widely used to design high-order accurate schemes for the MHD equations, see [Fuc+11] and references therein, also in view of their simple extensions to high polynomial orders.

The reconstruction is performed on the primitive variables  $\mathbf{V} = (\rho, \mathbf{u}, p)$ , while the conserved variables  $\mathbf{U} = (\rho, \rho\mathbf{u}, E)$  are evolved in time.

#### 6.3.1. Piecewise Linear Reconstruction

The piecewise linear reconstructed polynomial  $\mathbf{p}^n(\mathbf{x})$  has the form  $\mathbf{p}^n(\mathbf{x}) = \mathbf{V}_j^n + \boldsymbol{\sigma}_j^n(\mathbf{x} - \mathbf{x}_j)$  for all  $\mathbf{x} \in T_j$  and all  $T_j \in \mathcal{T}_h$ , where  $\boldsymbol{\sigma}_j^n = (\boldsymbol{\sigma}_j^{1,n} | \dots | \boldsymbol{\sigma}_j^{d,n}) \in \mathbb{R}^{m,d}$  is the matrix of the slopes of the components of  $\mathbf{V} \in \mathbb{R}^m$  in each direction  $1 \leq \ell \leq d$ . The slopes  $\{\boldsymbol{\sigma}_j^n\}_{j \in \mathfrak{J}}$  have to be carefully chosen in order to avoid oscillatory and unstable solutions near shock discontinuities.

##### Minmod Limiter

The so-called *minmod limiter* yields a TVD reconstruction which is “limited” based on local approximation of the solution gradients. In particular,

$$\boldsymbol{\sigma}_j^{\ell,n} := \text{minmod} \left( \frac{\mathbf{V}_{j+\mathbf{e}_\ell}^n - \mathbf{V}_j^n}{h_\ell}, \frac{\mathbf{V}_j^n - \mathbf{V}_{j-\mathbf{e}_\ell}^n}{h_\ell} \right), \quad \forall j \in \mathfrak{J}, \ell = 1, \dots, d,$$

where

$$\text{minmod}(a, b) := \begin{cases} \text{sign}(a) \min\{|a|, |b|\} & \text{if } \text{sign}(a) = \text{sign}(b), \\ 0 & \text{otherwise.} \end{cases}$$

The minmod limiter selects the smallest between the slopes in the upwind and downwind direction if they have the same sign, otherwise the reconstructed solution will coincide locally with the cell average.

### ENO Reconstruction

In the ENO procedure, the reconstructed primitive variables  $\widehat{\mathbf{V}}_j^n(\mathbf{x})$  are derived from the piecewise constant cell averages  $\mathbf{V}_j^n$  as

$$\widehat{\mathbf{V}}_j^n(\mathbf{x}) = \mathbf{V}_j^n + \sum_{\ell=1}^d \frac{1}{h_\ell} D^\ell \mathbf{V}_j^n(\mathbf{x} - \mathbf{x}_j), \quad \forall \mathbf{x} \in T_j, \quad j \in \mathfrak{J}, \quad (6.17)$$

where the ENO gradients are defined as in [Fuc+11, Section 2.5.1], for all  $1 \leq \ell \leq d$  and  $j \in \mathfrak{J}$ ,

$$D^\ell \mathbf{V}_j^n := \begin{cases} \mathbf{V}_{j+\mathbf{e}_\ell}^n - \mathbf{V}_j^n & \text{if } \Gamma_j^\ell \leq 1, \\ \mathbf{V}_j^n - \mathbf{V}_{j-\mathbf{e}_\ell}^n & \text{otherwise,} \end{cases} \quad \Gamma_j^\ell := \frac{|\psi_{j+\mathbf{e}_\ell}^n - \psi_j^n|}{|\psi_j^n - \psi_{j-\mathbf{e}_\ell}^n|}.$$

The function  $\psi$ , called *smoothness indicator*, depends on the conserved variables and should carry information on the wave structure and on the discontinuities of the solution. As observed in [Fuc+11], the total energy provides a potentially good indicator function since it is discontinuous across all discontinuities of the one-dimensional MHD problem. Straightforward computations show that the requirements in the reconstruction step of the REA algorithm are met by the ENO reconstruction procedure.

### WENO Reconstruction

The WENO algorithm is a modified version of the ENO reconstruction. Specifically, it considers the linear reconstruction (6.17) with cell gradients

$$D^\ell \mathbf{V}_j^n := \omega_j^\ell (\mathbf{V}_{j+\mathbf{e}_\ell}^n - \mathbf{V}_j^n) + (1 - \omega_j^\ell) (\mathbf{V}_j^n - \mathbf{V}_{j-\mathbf{e}_\ell}^n), \quad \forall j \in \mathfrak{J}, \quad \ell = 1, \dots, d.$$

The weights are

$$\omega_j^\ell := \frac{a_j^0}{a_j^0 + a_j^1}, \quad a_j^0 = \frac{1}{3(\epsilon + b_j^{\ell,0})}, \quad a_j^1 = \frac{1}{3(\epsilon + b_j^{\ell,1})},$$

with coefficients  $b_j^{\ell,0} = (\psi_{j+\mathbf{e}_\ell}^n - \psi_j^n)^2$  and  $b_j^{\ell,1} = (\psi_j^n - \psi_{j-\mathbf{e}_\ell}^n)^2$ . The function  $\psi$  is the smoothness indicator and the parameter  $\epsilon$  is a small positive number. In the forthcoming numerical experiments, we use as smoothness indicator the total energy  $E$  and set  $\epsilon = 10^{-8}$ .

ENO and WENO reconstruction techniques are robust and provide non-oscillatory solutions, but they do not ensure that the resulting methods are positively conservative. To overcome the lack of positivity, a further modification consists in limiting the slope in each direction, such that the reconstructed variables are in the admissible set of states (6.7). The updated variables obtained from the reconstructed ones are however not guaranteed to have positive density and pressure. A further positive fix is in order: modifications of the ENO and WENO reconstructions for the ideal MHD system are derived in [Waa09] and [Fuc+11, Section 2.5]. Note that a similar reasoning for the extended Euler equations does not apply since the lowest order scheme is not provably positively conservative in the first place.

#### 6.3.2. Numerical Experiments in 2D

To the purpose of appraising the impact of piecewise polynomial reconstructions within finite volume discretizations of the extended Euler problem, we conduct the same numerical experiments as in Section 6.2.2.

### Accuracy Test: Smooth Vortex

As detailed in Section 6.2.2, the solution of the smooth vortex test is known analytically at any point in space and time in the domain  $\Omega \times I = [-5, 5]^2 \times [0, 0.5]$ . Specifically, let  $r(x, y, t) := \sqrt{(x - u_0^1 t)^2 + (y - u_0^2 t)^2}$ , with  $\mathbf{u}_0 = (u_0^1, u_0^2)^\top = (1, 1)^\top$ , the flow is characterized by the following set of data,

$$\begin{aligned}\rho(x, y, t) &= 1, \\ p(x, y, t) &= 1 + \frac{1}{8\pi}(\mu^2(1 - r^2) - \kappa^2)e^{1-r^2}, \\ \mathbf{u}(x, y, t) &= \mathbf{u}_0 + \frac{\kappa}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top, \\ \mathbf{B}(x, y, t) &= \frac{\mu}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top,\end{aligned}$$

with parameters  $\kappa = \mu = 1$ . The ratio of specific heats is  $\gamma = 5/3$ . The magnetic induction field  $\mathbf{B}$  is provided in analytic form and the initial condition is given by  $\mathbf{U}(\mathbf{x}, 0)$ . We test the two- and three-wave Riemann solvers from Section 6.2.1 with different piecewise linear reconstructions and second order SSP-RK as time-stepping. In order to minimize the error pollution, we consider a relatively small time step by setting  $C_{\text{cfl}} = 0.01$  and feed the finite volume scheme with a magnetic induction field  $\mathbf{B}$  computed at each stage of the RK2 time-stepping.

The  $L^2$ -error and  $L^1$ -error of the numerical pressure at the final time  $T = 0.5$  are reported in Figure 6.4. The second order accuracy is attained with all the piecewise linear reconstructions proposed in Section 6.3.1, with the minmod limiter proving slightly overdiffusive.

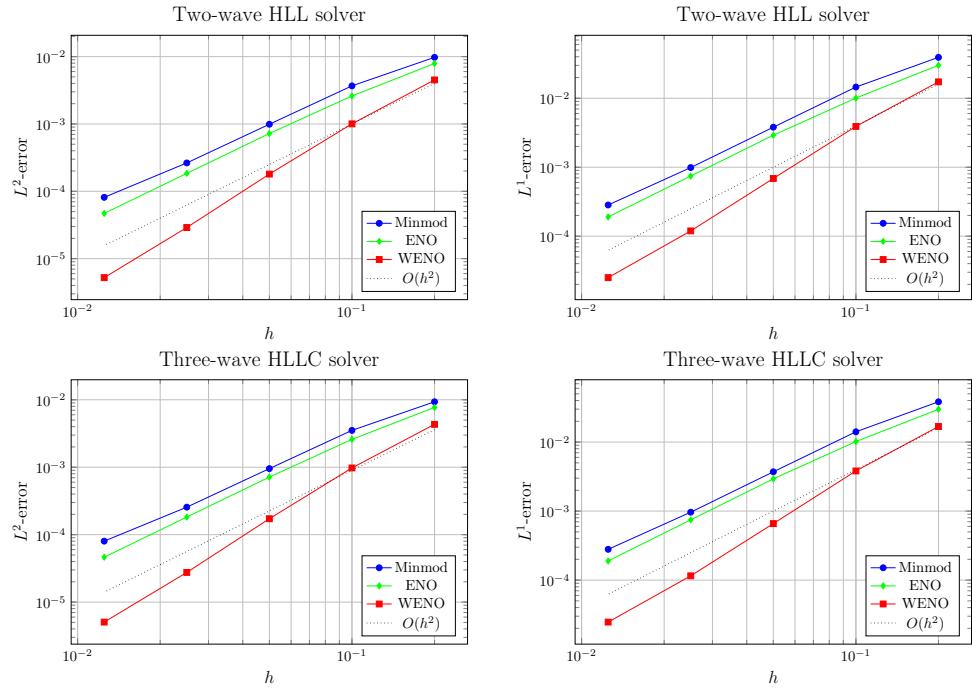


Figure 6.4.: Smooth vortex. Accuracy test for finite volume discretizations of the extended Euler system with given analytic  $\mathbf{B}$  field. HLL approximate Riemann solvers are tested in combination with different piecewise linear reconstructions. The  $L^2$ -error of the total pressure at the final time is reported in the first column, and the  $L^1$ -error in the second column.

### Orszag–Tang Benchmark with Given $\mathbf{B}$ Field

As a second test case, we consider the Orszag–Tang test for ideal MHD [OT79]. We refer to Section 7.1.4 for details on the benchmark. The magnetic induction field  $\mathbf{B}$  is given at each time step as the output of a high order FV discretization of the full MHD system, as in Section 6.2.2. In Figure 6.5, the projections of the discrete velocity field onto the one-dimensional interval  $[0, 2]$  for constant values of  $y = 0.56719$  and  $y = 1.2637$ , respectively, are presented. The finite volume scheme for the extended Euler system uses the two-wave HLL solver of Linde (see Section 6.2.1). On a  $200 \times 200$  Cartesian mesh, we compare the output of the lowest order FV scheme based on the HLL solver and explicit Euler time-stepping with the FV scheme using the HLL solver, piecewise linear ENO reconstruction and SSP-RK2 time-stepping. From Figure 6.5 we can infer that, compared with the high order reference solution, the FV scheme with reconstruction yields non-diffusive accurate results without oscillations at the shocks.

Note that for the present Orszag–Tang test with given  $\mathbf{B}$  field on a relatively coarse mesh ( $200 \times 200$  elements), no positive fix has been implemented on the reconstruction step: notwithstanding, negative pressures or densities are not encountered for any of the Riemann solvers from Section 6.2.1.

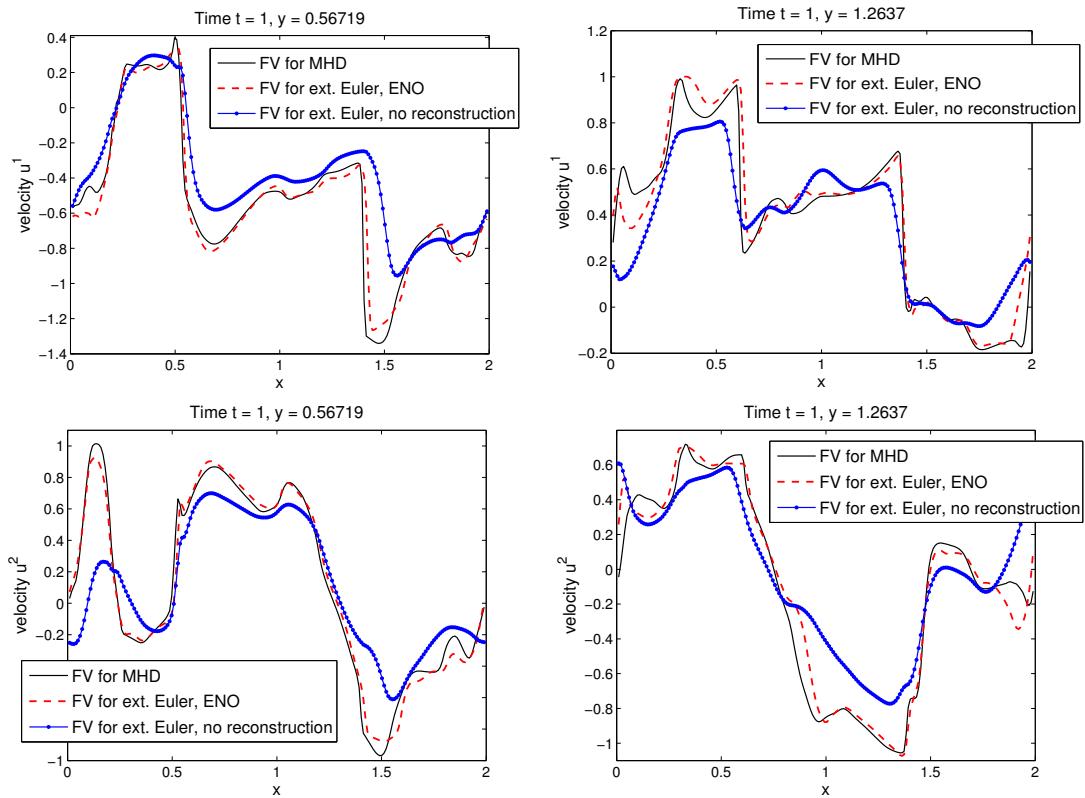


Figure 6.5.: Orszag–Tang benchmark with given  $\mathbf{B}$  field. Projections of the numerical velocity on a  $200 \times 200$  Cartesian mesh for constant values of  $y$ . Comparison of HLL solver without reconstruction and explicit Euler time-stepping, and of HLL solver, piecewise linear ENO reconstruction, and RK2 time-stepping with the reference solution from a high order FV discretization of the full MHD system. Uniform time step  $\Delta t = 5 \cdot 10^{-4}$ .

# 7. Numerical Experiments for Ideal MHD

The numerical methods derived and analyzed in Chapters 3 and 4 for the magnetic induction equation and in Chapter 6 for the extended Euler system can be blended into discretization schemes for the full ideal MHD problem, which we coin FV-FEEC (Finite Volume-Finite Element Exterior Calculus).

The present chapter is devoted to testing the new family of FV-FEEC schemes, for the two-dimensional planar MHD problem, on a set of numerical benchmark experiments. The scope is to provide numerical evidence of their accuracy, good stability properties and ability in preserving the physical distinctive features of the model at the continuous level.

## 7.1. Planar Two-Dimensional Ideal MHD

When numerically tackling the full MHD problem we focus, in the present chapter, on the planar two-dimensional problem. The two-dimensional setting provides a first step to assess the robustness of numerical schemes maintaining many of the physical features of the three-dimensional model and paves the way to three-dimensional simulations. In the transverse magnetic setting, the translational symmetry allows to reduce the three-dimensional magnetic advection-diffusion problem to a two-dimensional initial boundary value problem on the space-time cylinder with spatial domain obtained from a cross section of the three-dimensional domain. The case of systems of reduced dimensionality because of axial symmetries is common to many applications and the theory and numerical analysis extend mutatis mutandis to the lower dimensional case. In particular, in a planar two-dimensional MHD model, the magnetic potential is a scalar function representing the transverse out-plane component of the three-dimensional vector magnetic potential and  $\mathbf{B} = \text{curl}_{2D} A = \text{grad}^\perp A := (\partial_y A, -\partial_x A)^\top$ . The formalism of exterior calculus is flexible in accommodating this change of dimensionality without undermining the physical meaning of the quantities at hand. In vector calculus, by identifying flux fields and field intensities as vectors, the distinction between the two associated (isomorphic) vector spaces is not explicitly conveyed. In two dimensions the two vector spaces are no longer isomorphic: while 1-forms are still represented by vector fields, 2-forms are associated with scalar quantities.

### 7.1.1. The FV-FEEC Schemes in 2D

Let  $\Omega \subset \mathbb{R}^d$  be a Cartesian domain as from Section 6.2 and let  $I = [0, T]$  be a temporal interval. Let  $\mathcal{T}_h$  denote a tensor product decomposition of the domain  $\Omega$  under Assumption 2.4.2 and let the temporal interval  $I$  be divided into subintervals  $I = \bigcup_{n=0}^{N-1} (t^n, t^{n+1}]$ ,  $N \in \mathbb{N}$ , with  $t^{n+1} = t^n + \Delta t^n$  and time step  $\Delta t^n$  satisfying the CFL condition dictated by the fast MHD wave, as in (6.11). In the present section, we consider the piecewise constant finite volume schemes, derived in Section 6.2, for the extended Euler equations with suitable boundary and initial conditions,

$$\begin{cases} \partial_t \rho + \text{div}(\rho \mathbf{u}) = 0, \\ \partial_t(\rho \mathbf{u}) + \text{div}(\rho \mathbf{u} \otimes \mathbf{u} + (p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbb{I} - \mathbf{B} \otimes \mathbf{B}) = \mathbf{0}, \\ \partial_t E + \text{div}((E + p + \frac{1}{2} \|\mathbf{B}\|_{\ell^2}^2) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B}) \mathbf{B}) = 0, \end{cases}$$

in  $\Omega \times I$ , coupled with compatible finite element discretizations of the pure advection problem for the scalar magnetic potential  $A$ ,

$$\partial_t A + \mathbf{u} \cdot \text{grad} A = 0, \quad \text{in } \Omega \times I,$$

via lowest order extrusion contraction upwind schemes, see Section 4.2.

Two possible operator splitting algorithms can be implemented. A *sequential* one, in the spirit of the temporal Godunov splitting, where the extended Euler system is solved first, and the output velocity enters the discrete magnetic advection equation as in Algorithm 1 (the choice of which problem has to be addressed first is totally arbitrary since the operators do not necessarily commute).

---

**Algorithm 1** Sequential splitting algorithm

```

1: Set  $n = 0, t = 0$ . Given initial conditions  $(\rho_0, \mathbf{u}_0, E_0, \mathbf{B}_0)$ .
2: while time  $t < T$  do
3:    $\mathbf{u}^{n+1} \leftarrow$  Solve the extended Euler equations given  $\mathbf{B}^n$ .
4:    $A^{n+1} \leftarrow$  Solve the advection of 0-forms given  $\mathbf{u}^{n+1}$ .
5:    $\mathbf{B}^{n+1} \leftarrow$  Compute the (discrete) curl of  $A^{n+1}$ .
6:   Set  $t = t + \Delta t^n, n = n + 1$ .
7: end while
```

---

Alternatively, a *synchronous* splitting, Algorithm 2, envisages that the two systems, extended Euler and magnetic advection, are concurrently advanced in time, in the sense that after spatial discretization, the two problems are re-coupled to form a unique system of ODEs. The latter is solved via explicit SSP Runge–Kutta time-stepping, as described in (6.10). In this way, the coupling fields, the velocity and the magnetic induction, are updated within each subsystems at every intermediate stage of the temporal scheme.

---

**Algorithm 2** Synchronous splitting algorithm

```

1: Set  $n = 0, t = 0$ . Given initial conditions  $(\rho_0, \mathbf{u}_0, E_0, \mathbf{B}_0)$ .
2: while time  $t < T$  do
3:    $\mathbf{u}^{n;0} = \mathbf{u}^n; \mathbf{B}^{n;0} = \mathbf{B}^n$ .
4:   for each step  $1 \leq i \leq s$  of an  $s$ -stage SSP-RK time-stepping do
5:      $\mathbf{u}^{n;i} \leftarrow$  Solve the extended Euler equations given  $\{\mathbf{B}^{n;j}\}_{j=0}^{i-1}$ .
6:      $A^{n;i} \leftarrow$  Solve the advection of 0-forms given  $\{\mathbf{u}^{n;j}\}_{j=0}^{i-1}$ .
7:      $\mathbf{B}^{n;i} \leftarrow$  Compute the (discrete) curl of  $A^{n;i}$ .
8:   end for
9:    $\mathbf{u}^{n+1} = \mathbf{u}^{n;s}; A^{n+1} = A^{n;s}; \mathbf{B}^{n+1} = \mathbf{B}^{n;s}$ .
10:  Set  $t = t + \Delta t^n, n = n + 1$ .
11: end while
```

---

The sequential approach yields a *splitting error*, associated with the time lag, which might affect the accuracy order of the overall FV-FEEC schemes. For this reason, even if we have not recorded a significant difference in the two splitting algorithms in the lowest order discretizations, we will prefer hereafter the synchronous over the sequential approach.

**Remark 7.1.1** (Collisional fluid velocities). On a tensor product partition  $\mathcal{T}_h$  of the computational domain  $\Omega$ , the fluid velocity resulting from lowest order finite volume discretizations of the extended Euler equations is a  $\mathcal{T}_h$ -piecewise constant function  $\mathbf{u} \in \mathcal{P}_0(\mathcal{T}_h)$  (collecting the cell averages of the fluid velocity in each mesh element). However, the extrusion contraction upwind discretization of the advection problem entails an upwind interpolation of the Lie derivative which requires the knowledge of the velocity field at the mesh nodes. One can approximate, at time  $t$ , the value at a given node by averaging the values from the elements sharing the node. Using the notation from Chapter 6,

$$\mathbf{u}(\mathbf{x}_{j+\frac{1}{2}(\mathbf{e}_1+\mathbf{e}_2)}, t) = \frac{1}{4} \left( \mathbf{u}(\mathbf{x}_j, t) + \mathbf{u}(\mathbf{x}_{j+\mathbf{e}_1}, t) + \mathbf{u}(\mathbf{x}_{j+\mathbf{e}_1+\mathbf{e}_2}, t) + \mathbf{u}(\mathbf{x}_{j+\mathbf{e}_2}, t) \right), \quad \forall j \in \mathfrak{J}, t \in I.$$

This pointwise interpolation might appear rather crude. It is especially the case in the presence of colliding or diverging velocity at a node or an edge, as depicted in Figure 7.1, where one might lose information on the local dynamics.

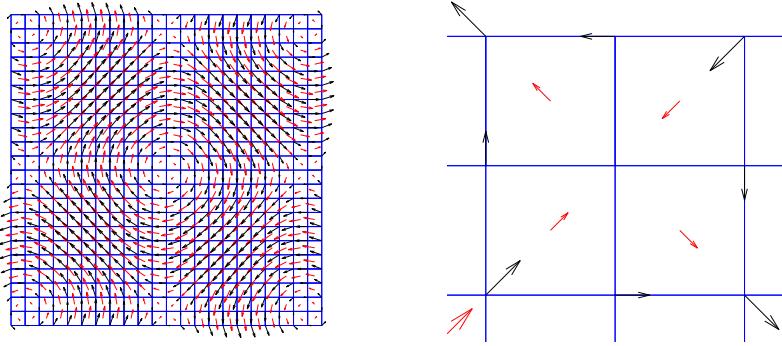


Figure 7.1.: Projection of the initial velocity field for the Orszag–Tang test (Section 7.1.4) onto the space of  $\mathcal{T}_h$ -piecewise constant functions. Values of the velocity at the mesh cells (red) and its averaged values at the mesh nodes (black). On the right a close-up of a mesh node where velocities from the neighboring elements collide or diverge.

In MHD, the fast magnetosonic wave speed provides a good indicator of the flow dynamics. Hence, in order to overcome the shortcomings related to averaging, the upwind direction at the mesh cells where the velocity field is colliding or diverging is determined by the fast MHD wave speed (1.12), namely

$$c_f^M := (c_f^{M,1}, c_f^{M,2}), \quad \text{with} \quad c_f^{M,\ell} = \frac{1}{\sqrt{2}} \sqrt{a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho}} + \sqrt{\left(a^2 + \frac{\|\mathbf{B}\|_{\ell^2}^2}{\rho}\right)^2 - 4a^2 \frac{B_\ell^2}{\rho}}, \quad \ell \in \{1, 2\}.$$

To recapitulate, in the FV-FEEC algorithms implemented in the forthcoming numerical experiments, the upwind direction entering the discrete Lie derivative at the nodes where the fluid velocity is collisional or diverging is determined by the fast MHD wave speed at the nodes, whilst the pointwise advection velocity at the nodes is taken from within the upwind element.

### 7.1.2. Accuracy Test: Smooth Vortex

In order to experimentally assess the accuracy of the lowest order FV-FEEC scheme, we present a genuinely two-dimensional (non-trivial) MHD test where the solution is known analytically at every point in space and time. The smooth vortex test was proposed in [Bal04, Section 6] (a scaling factor  $\sqrt{4\pi}$  has been absorbed in the definition of  $\mathbf{B}$ ). The problem is associated with a smoothly varying fluid vortex which propagates at a  $\pi/4$  angle to the computational mesh on the domain  $\Omega = [-5, 5]^2$  with periodic boundaries. The initial condition is given by a vortex characterized by fluctuations of the velocity and of the magnetic field, superimposed to an unperturbed MHD flow  $\mathbf{U}_0 = (\rho_0, p_0, u_0^1, u_0^2) = (1, 1, 1, 1)$ ,  $\mathbf{B}_0 = \mathbf{0}$ . Let  $r(x, y, t) := \sqrt{(x-t)^2 + (y-t)^2}$ , the flow is described by the following set of data,

$$\begin{aligned} \rho(x, y, t) &= 1, \\ p(x, y, t) &= 1 + \frac{1}{8\pi}(\mu^2(1 - r^2) - \kappa^2)e^{1-r^2}, \\ \mathbf{u}(x, y, t) &= \mathbf{u}_0 + \frac{\kappa}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top, \\ \mathbf{B}(x, y, t) &= \frac{\mu}{2\pi}e^{1/2(1-r^2)}(t - y, x - t)^\top, \\ A(x, y, t) &= \frac{\mu}{2\pi}e^{1/2(1-r^2)}. \end{aligned}$$

As initial datum we take  $\mathbf{U}(\mathbf{x}, 0)$  with  $\mathbf{u}_0 = (u_0^1, u_0^2)^\top = (1, 1)^\top$ ,  $\kappa = \mu = 1$  and  $A(\mathbf{x}, 0)$  for the magnetic advection subproblem. The ratio of specific heats is  $\gamma = 5/3$ . The time interval is  $I = [0, 0.5]$ . Explicit Euler is used as time-stepping with CFL constant  $C_{\text{CFL}} = 0.4$ .

The numerical convergence study on smooth solutions displays first order convergence in the  $L^2$ - and  $H^1$ -norms of the scalar magnetic potential at the final time, see Figure 7.2 (left). On the right plot in Figure 7.2, the errors in the norms associated with the Bochner spaces  $L^\infty(I, L^2(\Omega))$  and  $L^\infty(I, H^1(\Omega))$ , defined as  $\|A\|_{L^\infty(I, H)} := \text{ess sup}_{t \in I} \|A(t)\|_H$  on the Sobolev space  $(H, \|\cdot\|_H)$ , are reported. Analogous conclusion can be drawn from Figure 7.3, where the errors of the  $\mathcal{T}_h$ -piecewise constant MHD primitive variables, in different norms, are illustrated.

The results obtained with sequential and parallel coupling were comparable.

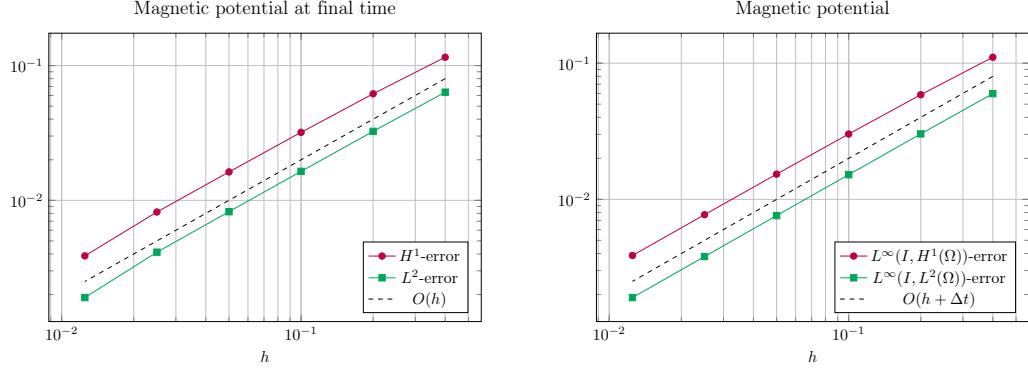


Figure 7.2.: Accuracy test for full ideal MHD system.  $L^2$ - and  $H^1$ -error of the magnetic potential at the final time  $T = 0.5$  (left);  $L^\infty(I, L^2(\Omega))$ - and  $L^\infty(I, H^1(\Omega))$ -error of the magnetic potential (right). Finite volume discretization of the extended Euler system with HLL Riemann solver and piecewise linear extrusion contraction scheme for the advection of the scalar magnetic potential. Explicit Euler time-stepping and  $C_{\text{CFL}} = 0.4$ .

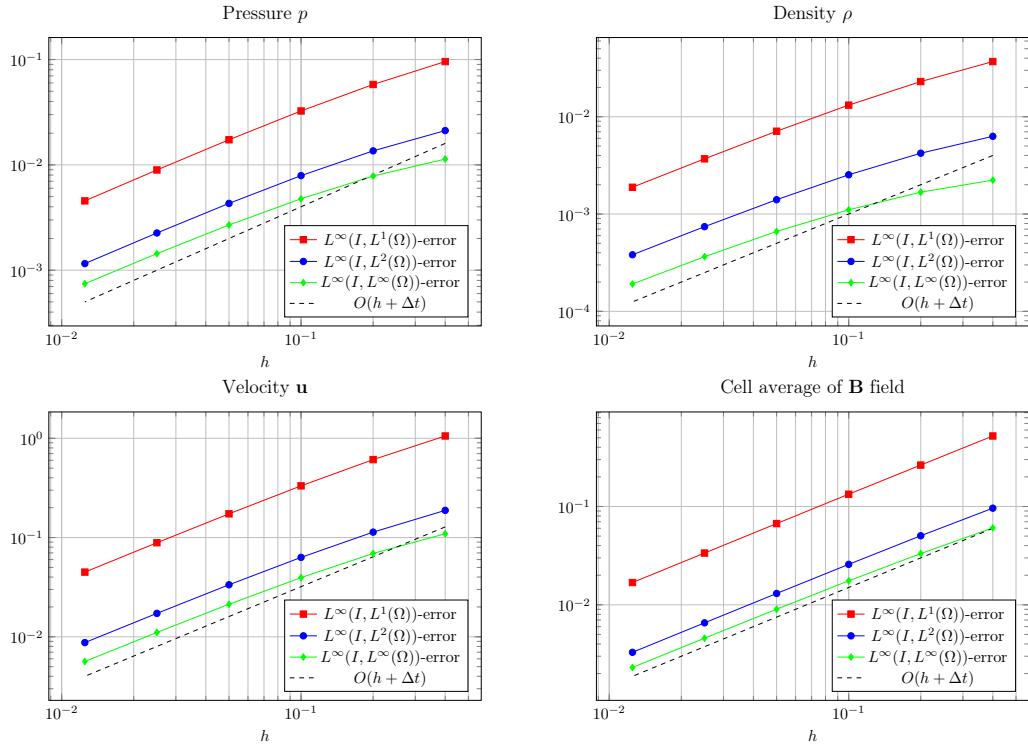


Figure 7.3.: Accuracy test for full ideal MHD system. Space-time error of the MHD primitive variables in different norms. Results obtained with a finite volume discretization of the extended Euler system with HLL Riemann solver and piecewise linear extrusion contraction scheme for the advection of the scalar magnetic potential, explicit Euler time-stepping and  $C_{\text{CFL}} = 0.4$ .

### 7.1.3. Super-Fast Expansion: Shock Tube Test

To the aim of testing the robustness of numerical schemes in delivering physically admissible solutions of the one-dimensional MHD equations, a super-fast expansion simulation has been used in e.g. [MK05, Section 6.1 p. 338] and [FMR09, Section 3.2]. With the same goal, we study a variant of the foregoing test case in a formally two-dimensional setting. Let us consider the domain  $\Omega = [0, 1]^2$  with periodic upper  $\{\mathbf{x} = (x, y) \in \Omega : y = 1\}$  and lower  $\{\mathbf{x} = (x, y) \in \Omega : y = 0\}$  boundaries. In the remaining part of  $\partial\Omega$ , non-reflecting Neumann type boundary conditions are applied to the conserved variables of the extended Euler system, and outflow boundary conditions for the advection problem since the evolution of the velocity field does not induce any inflow boundary on the considered time interval  $I = [0, 0.2]$ . We perform planar two-dimensional simulations of the one-dimensional (in the  $x$ -direction) shock tube test with initial data

$$\begin{aligned}\rho_0(x, y) &= 1, \\ p_0(x, y) &= 0.45, \\ u_0^1(x, y) &= \begin{cases} -3.1 & \text{if } x < 0.5, \\ 3.1 & \text{if } x > 0.5, \end{cases} \quad u_0^2(x, y) = 0, \\ \mathbf{B}_0(x, y) &= (0, 0.5)^\top, \\ A_0(x, y) &= -0.5x,\end{aligned}$$

and  $\gamma = 5/3$  as ratio of specific heats. As the problem involves a left-moving and a right-moving rarefaction waves, the central region is subject to a super-fast expansion yielding very low density and pressure.

In [FMR09, Section 3.2], it has been observed that linearized Roe solvers for finite volume discretizations of the full MHD system usually run into negative pressure and density in such test case. The FV-FEEC scheme proves positively conservative at all tested resolutions, see Figure 7.4.

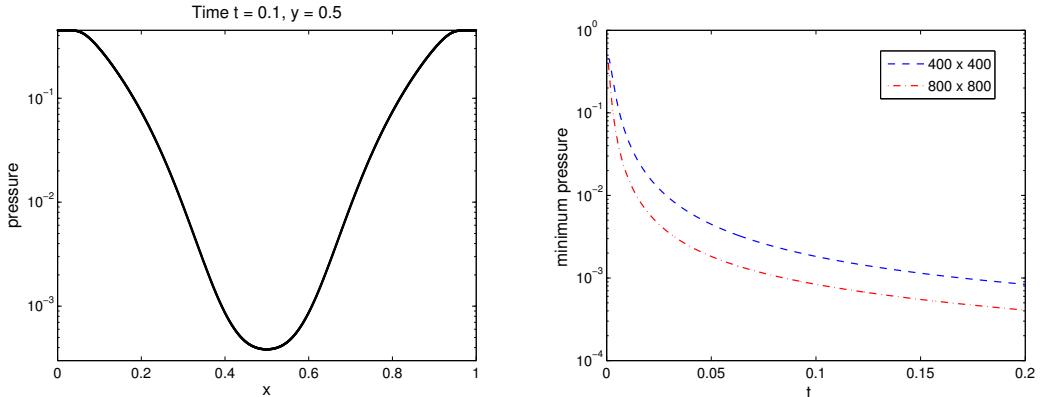


Figure 7.4.: Super-fast expansion. Semi-logarithmic plot of the projection onto  $\{\mathbf{x} \in \Omega : y = 0.5\}$  of the numerical pressure (left) at the intermediate time  $t = 0.1$ . Numerical discretization on a  $1600 \times 1600$  Cartesian mesh with Heun time-stepping and  $C_{\text{CFL}} = 0.4$ . Lowest order extrusion contraction for the advection of the magnetic potential and finite volume scheme for the extended Euler equations using the HLL approximate Riemann solver. On the right, semi-logarithmic evolution plot of the minimum of the discrete pressure until time  $t = 0.2$  and at  $400 \times 400$  and  $800 \times 800$  mesh resolutions.

### 7.1.4. Orszag–Tang Benchmark

The so-called Orszag–Tang vortex system was introduced in [OT79, Section 3], and describes the transition to supersonic turbulence in the MHD equations. The development of shock waves and the complex interaction between various shocks with different speed, which characterized the solution, makes the Orszag–Tang benchmark a challenging test for numerical methods.

Let us consider the domain  $\Omega = [0, 2]^2$  with periodic boundary conditions at the boundary  $\partial\Omega$ . The time interval is  $I = [0, 1]$ . The initial conditions for the primitive fluid variables, the magnetic induction field  $\mathbf{B}$  and the magnetic potential  $A$  are

$$\begin{aligned}\rho_0(x, y) &= \gamma^2, \\ p_0(x, y) &= \gamma, \\ \mathbf{u}_0(x, y) &= (-\sin(\pi y), \sin(\pi x))^\top, \\ \mathbf{B}_0(x, y) &= (-\sin(\pi y), \sin(2\pi x))^\top, \\ A_0(x, y) &= \frac{1}{\pi} \cos(\pi y) + \frac{1}{2\pi} \cos(2\pi x),\end{aligned}$$

and  $\gamma = 5/3$  is the gas constant. Since the Orszag–Tang test is a widely used benchmark, we can compare the performances of the FV-FEEC scheme with the values presented in [FMR09, Table 2.2] for different numerical methods. The comparison parameter is the maximum value of the discrete pressure at the final time  $t = 1$ . As reported in [FMR09, Section 3.3], finite volume schemes for the full MHD system based on Roe solvers, two-wave HLL and three-wave HLLC solvers, with and without divergence cleaning (via projection methods), experienced negative pressures at fine mesh resolution.

In Table 7.1, we compare the maximum pressure of the lowest order FV-FEEC discretization with the first order finite volume schemes HLL/SUS and HLLC/SUS from [FMR09] and [Fuc+09]. The reported values are comparable at all mesh resolutions: the HLLC solver for FV-FEEC gives slightly more “accurate” results, as experienced in Section 6.2.2 when testing approximate Riemann solvers on the extended Euler system, and shown in Figure 7.5 in comparison with a second order accurate reference solution<sup>1</sup>.

| $\#\mathcal{T}_h$  | HLL/SUS<br>[FMR09, Table 2.2] | FV-FEEC<br>HLL | HLLC/SUS<br>[FMR09, Table 2.2] | FV-FEEC<br>HLLC |
|--------------------|-------------------------------|----------------|--------------------------------|-----------------|
| $100 \times 100$   | 4.94                          | 4.38           | 5.04                           | 5.05            |
| $200 \times 200$   | 5.39                          | 5.29           | 5.41                           | 5.65            |
| $400 \times 400$   | 5.79                          | 5.84           | 5.81                           | 5.91            |
| $800 \times 800$   | 6.05                          | 6.06           | 6.07                           | 6.10            |
| $1600 \times 1600$ | 6.21                          | 6.20           | 6.22                           | 6.23            |

Table 7.1.: Orszag–Tang benchmark. Maximum value of the discrete pressure at the final time obtained with the FV-FEEC scheme with two-wave HLL and three-wave HLLC Riemann solvers, and compared with the values from the finite volume discretizations HLL(C)/SUS in [FMR09].

Finite volume schemes are, by construction, conservative methods. However, the conservation property is not naturally inherited by the coupled FV-FEEC discretizations. In order to numerically assess the conservation properties of the FV-FEEC scheme, we monitor the evolution of the mean, on the computational domain  $\Omega$ , of the conserved variables over time. Given the scalar function  $f(\mathbf{x}, t)$  and the initial datum  $f_0(\mathbf{x})$ , we compute, on the partition of the temporal interval, the following error,

$$\mathcal{E}(f) := \max_{1 \leq n \leq N} |\bar{f}^n - \bar{f}_0|, \quad \bar{f}^n := \oint_{\Omega} f(\mathbf{x}, t^n) d\mathbf{x} \quad \text{and} \quad \bar{f}_0 := \oint_{\Omega} f_0(\mathbf{x}) d\mathbf{x}. \quad (7.1)$$

Whenever needed we consider the relative error  $\mathcal{E}_r(f) := \mathcal{E}(f)/|\bar{f}_0|$ . Table 7.2 reports the values of the aforementioned error on the conserved MHD variables and for different mesh refinements.

<sup>1</sup>The reference solutions used throughout the present chapter were provided by R. Käppeli, SAM, ETH Zürich, and based on the FISH code [Käp+11].

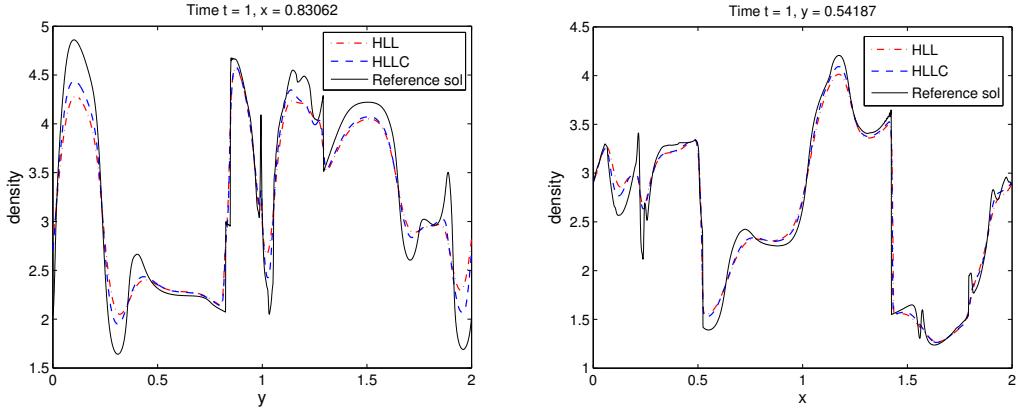


Figure 7.5.: Orszag–Tang benchmark. Projections of the discrete density for constant values of the  $x$ -coordinate (left) and  $y$ -coordinate (right). Numerical discretization on a  $1600 \times 1600$  Cartesian mesh. The finite volume scheme for the extended Euler equations is based on two-wave HLL and three-wave HLLC approximate Riemann solver. Heun time-stepping with  $C_{\text{CFL}} = 0.4$ . A second order accurate finite volume solution on  $3200 \times 3200$  mesh elements is used as reference solution.

| $\#\mathcal{T}_h$  | $\mathcal{E}_r(\rho)$ | $\mathcal{E}(\rho u^1)$ | $\mathcal{E}(\rho u^2)$ | $\mathcal{E}(B_1)$  | $\mathcal{E}(B_2)$  | $\mathcal{E}_r(E)$  |
|--------------------|-----------------------|-------------------------|-------------------------|---------------------|---------------------|---------------------|
| $200 \times 200$   | $1.3097\text{e-}15$   | $2.1723\text{e-}11$     | $4.0363\text{e-}12$     | $1.8402\text{e-}11$ | $4.5034\text{e-}12$ | $2.1552\text{e-}15$ |
| $400 \times 400$   | $1.8335\text{e-}15$   | $1.3512\text{e-}10$     | $3.3280\text{e-}11$     | $2.0040\text{e-}10$ | $3.4728\text{e-}10$ | $1.8236\text{e-}15$ |
| $800 \times 800$   | $2.2264\text{e-}15$   | $4.1668\text{e-}10$     | $1.0388\text{e-}10$     | $9.8822\text{e-}10$ | $7.6756\text{e-}10$ | $2.8183\text{e-}15$ |
| $1600 \times 1600$ | $3.6671\text{e-}15$   | $5.6204\text{e-}09$     | $5.9614\text{e-}10$     | $2.0064\text{e-}09$ | $4.8494\text{e-}09$ | $3.1498\text{e-}15$ |

Table 7.2.: Orszag–Tang benchmark. Conservation properties of the FV-FEEC scheme. Error (7.1) of the MHD conserved variables at different mesh resolutions.

In Figure 7.6, we report the  $L^1$ -error of the primitive MHD variables at the final time  $t = 1$  computed with respect to the second order reference solution. The observed convergence rate is around 0.6.

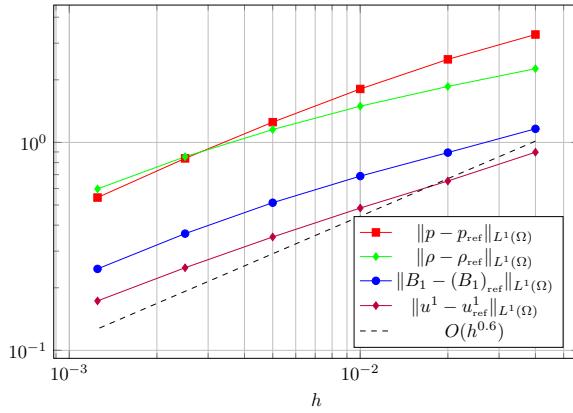


Figure 7.6.: Orszag–Tang benchmark. Plot of the  $L^1$ -error vs. the mesh width  $h$ . The error of the MHD primitive variables is computed at the final time  $t = 1$  and with respect to a reference solution on a  $3200 \times 3200$  mesh.

Finally, the ability of the FV-FEEC scheme to reproduce physically reliable solutions with rather sharp resolution of the shock fronts is gauged in Figure 7.7, see also the results available in literature e.g. [Fuc+11, Section 3.4] or [Tót00, Section 6.4]. The lowest order FV-FEEC is admittedly diffusive and does not capture all the complex shock interaction features visible in the second order accurate solution in Figure 7.7 (right column). The FV-FEEC nor the second order reference solution manage to reproduce

the current sheet appearing in the center of the domain in the second component of the  $\mathbf{B}$  field, refer to Figure 3.14 in Chapter 3.

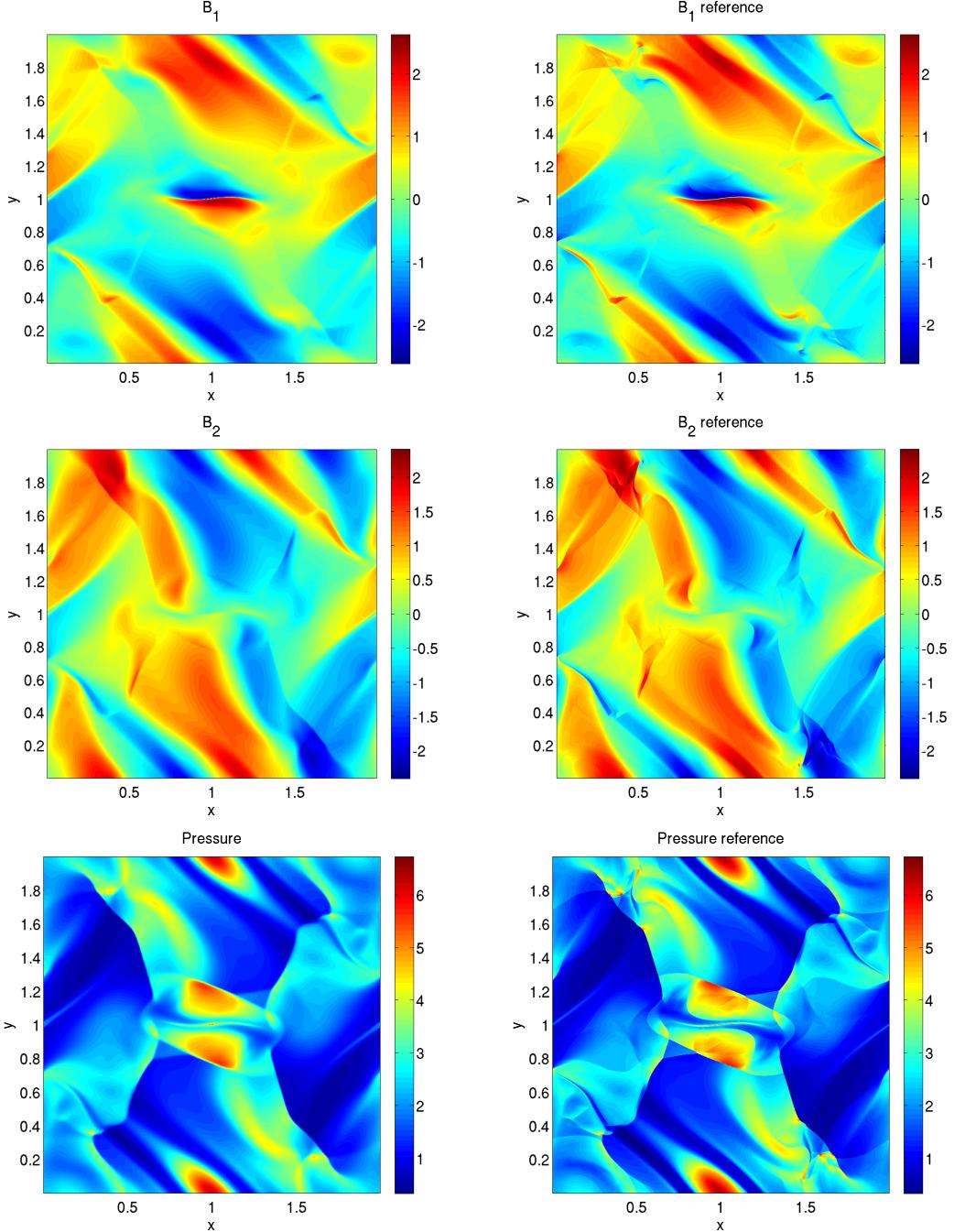


Figure 7.7.: Orszag–Tang benchmark. On the left, numerical solution on a  $1600 \times 1600$  Cartesian mesh obtained with the FV-FEEC scheme and Heun time-stepping with  $C_{\text{CFL}} = 0.4$ . Finite volume scheme for the extended Euler equations based on the HLL approximate Riemann solver. The color map is scaled to the extrema of the reference solution on a  $3200 \times 3200$  mesh (right column).

### 7.1.5. Rotor Problem

A key dimensionless parameter for ideal MHD models is the so-called *plasma beta* defined as the ratio of thermal to magnetic pressure,

$$\beta = 2\mu_0 \frac{p}{\|\mathbf{B}\|_{\ell^2}^2}.$$

Delivering physically admissible solutions in presence of low-beta plasmas ( $\beta \ll 1$ ) is particularly challenging for numerical schemes. The rotor problem provides a numerical test for the low-beta plasma setting. The rotor benchmark was introduced in [BS99, Section 3.1] to test the emergence and propagation of torsional Alfvén waves. The interior of the rotor is characterized by low values of the pressure, so that the test is also well-suited to attest the robustness of a numerical method in preserving positivity. The initial set up consists of a dense spinning cylinder (the rotor) of radius 0.05, surrounded by the ambient fluid at rest which occupies the remaining part of the computational domain  $\Omega = [0, 1]^2$ . The initial magnetic field is uniform but, as the rotor spins with the initial rotating velocity, the magnetic field in the  $x$ -direction starts wrapping around the rotor causing torsional Alfvén waves to propagate into the ambient fluid. As a result, the angular momentum will eventually decrease while the rotor will experience compression under the effect of the increased magnetic pressure assuming an oblong shape.

The physical problem is set up on an unbounded domain. This translates into artificial non-reflecting Neumann-type boundary conditions for the conserved variables entering the extended Euler system. Concerning the magnetic advection problem, the evolution of the velocity field guarantees that no inflow boundaries will occur at any time. The initial data are as explained above and characterized by the following data,

$$\begin{aligned} \rho_0(x, y) &= \begin{cases} 10 & \text{if } r < 0.1, \\ 1 + 9f & \text{if } 0.1 < r < 0.115, \\ 1 & \text{if } r > 0.115, \end{cases} \\ p_0(x, y) &= 0.5, \\ u_0^1(x, y) &= \begin{cases} 5 - 10y & \text{if } r < 0.1, \\ (5 - 10y)f & \text{if } 0.1 < r < 0.115, \\ 0 & \text{if } r > 0.115, \end{cases} \quad u_0^2(x, y) = \begin{cases} 10x - 5 & \text{if } r < 0.1, \\ (10x - 5)f & \text{if } 0.1 < r < 0.115, \\ 0 & \text{if } r > 0.115, \end{cases} \\ \mathbf{B}_0(x, y) &= \left( \frac{2.5}{\sqrt{4\pi}}, 0 \right)^\top, \\ A_0(x, y) &= \frac{2.5}{\sqrt{4\pi}}y, \end{aligned}$$

where  $r := \sqrt{(x - 0.5)^2 + (y - 0.5)^2}$ ,  $f := (23 - 200r)/3$  and the gas constant is  $\gamma = 5/3$ . The simulation runs until time  $T = 0.295$ .

In order to numerically analyze the conservative properties of the FV-FEEC scheme tested on the rotor problem, Table 7.3 reports the error (7.1) on the conserved variables.

| $\#\mathcal{T}_h$ | $\mathcal{E}_r(\rho)$ | $\mathcal{E}_r(\rho u^1)$ | $\mathcal{E}_r(\rho u^2)$ | $\mathcal{E}_r(B_1)$ | $\mathcal{E}_r(B_2)$ | $\mathcal{E}_r(E)$ |
|-------------------|-----------------------|---------------------------|---------------------------|----------------------|----------------------|--------------------|
| 200 × 200         | 4.3462e-04            | 4.1243e-11                | 2.3873e-12                | 4.8734e-04           | 3.0872e-04           | 9.3236e-04         |
| 400 × 400         | 8.2870e-05            | 9.7000e-12                | 4.3911e-11                | 9.9092e-05           | 6.5086e-08           | 1.7748e-04         |
| 800 × 800         | 5.5013e-06            | 3.5593e-11                | 4.5438e-10                | 6.7642e-06           | 1.1632e-10           | 1.1712e-05         |
| 1600 × 1600       | 1.0410e-07            | 1.6785e-08                | 5.7213e-07                | 1.3463e-07           | 1.7050e-09           | 2.2139e-07         |

Table 7.3.: Rotor problem. Conservation properties of the FV-FEEC scheme. Error (7.1) of the MHD conserved variables at different mesh resolutions.

As in the Orszag–Tang benchmark, we compare the FV-FEEC solution with a second order accurate finite volume reference solution. Figure 7.8, on the left, shows the  $L^1$ -error of the primitive MHD variables as the mesh is refined. Since the error stagnates rather quickly, we also check the convergence of the scheme under mesh refinement by evaluating the  $L^1$ -error with respect to the numerical solution obtained with the same FV-FEEC scheme but on a finer mesh ( $1600 \times 1600$  elements).

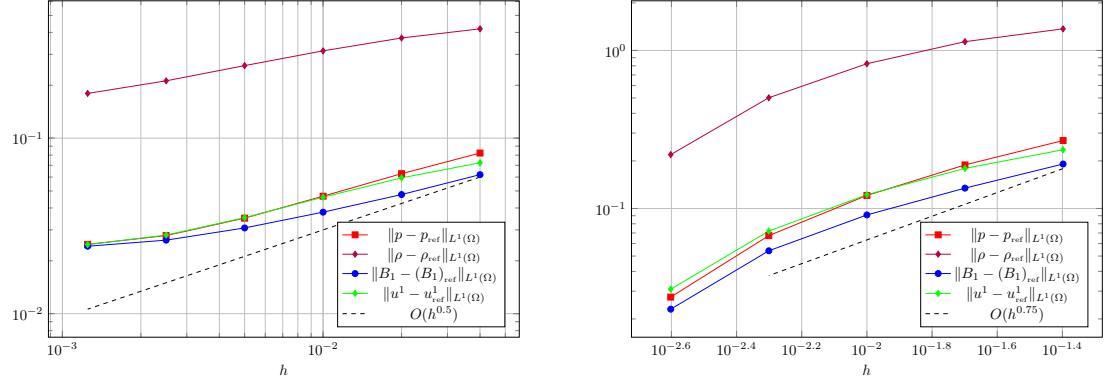


Figure 7.8.: Rotor problem. Plot of the  $L^1$ -error vs. the mesh width  $h$ . The error of the MHD primitive variables is computed at the final time  $t = 1$  and with respect to a second order accurate reference solution on a  $3200 \times 3200$  mesh (left). The plot on the right shows the  $L^1$ -error of the MHD primitive variables at the final time  $t = 1$ , computed with respect to the FV-FEEC solution on the finest mesh, namely at resolution  $1600 \times 1600$ .

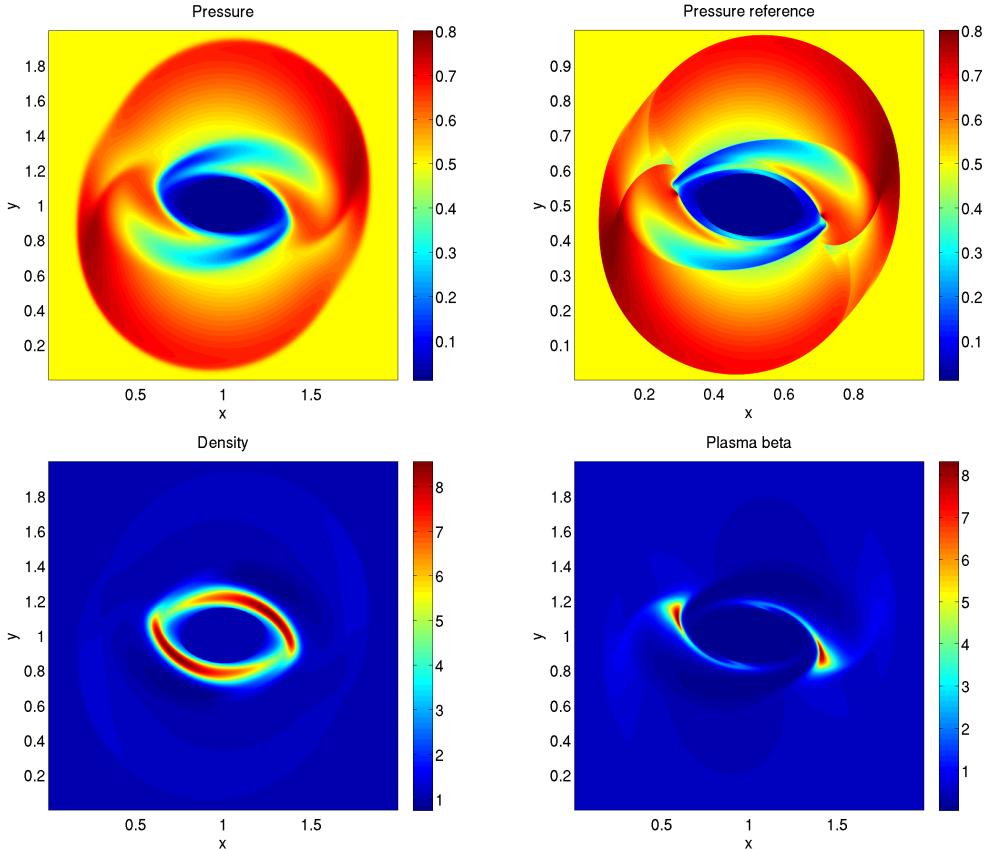


Figure 7.9.: Rotor problem. Numerical solution on a  $1600 \times 1600$  Cartesian mesh obtained with the lowest order FV-FEEC scheme and Heun time-stepping with  $C_{\text{CFL}} = 0.4$ . Finite volume scheme for the extended Euler equations based on the HLL approximate Riemann solver. The color map of the pressure plot is scaled to the extrema of the reference solution on a  $3200 \times 3200$  mesh, shown on the top right plot.

The FV-FEEC performs robustly also in the rotor test, as attested by Figure 7.9. The scheme captures many of the features of the MHD rotor flow being however rather diffusive when compared with the second order reference solution.

### 7.1.6. Blast Wave Problem

As a last test case, we consider the isothermal blast wave problem proposed in [Bal98, Section 6.2.2]. Note that we did not develop a numerical scheme tailored to the isothermal MHD problem, we rather “emulated” the isothermal behavior by setting the ratio of specific heats close to unitary ( $\gamma = 1.001$  in the forthcoming simulations). The blast wave benchmark is numerically challenging because it is characterized by a highly anisotropic explosion spreading out from a high density cloud initialized in a circular region of the domain. As pointed out in [Bal98], failing to provide a control of the divergence of the induction field can engender detrimental small-scale fluctuations. In the computational domain  $\Omega = [0, 1]^2$  the initial data are,

$$\begin{aligned}\rho_0(x, y) &= p_0(x, y) = \begin{cases} 100 & \text{if } \sqrt{(x - 0.5)^2 + (y - 0.5)^2} < 0.05, \\ 1 & \text{otherwise,} \end{cases} \\ \mathbf{u}_0(x, y) &= \mathbf{0}, \\ \mathbf{B}_0(x, y) &= \left( \frac{5}{\sqrt{\pi}}, 0 \right)^\top, \\ A_0(x, y) &= \frac{5}{\sqrt{\pi}}y.\end{aligned}$$

The simulation spans the time interval  $I = [0, 0.09]$ . Boundary conditions are of non-reflecting Neumann-type for the extended Euler variables, and the velocity field gives no inflow boundary.

The conservative properties of the scheme are numerically gauged in Table 7.4 where the error defined in (7.1) on the conserved variables is reported.

| $\#\mathcal{T}_h$  | $\mathcal{E}_r(\rho)$ | $\mathcal{E}(\rho u^1)$ | $\mathcal{E}(\rho u^2)$ | $\mathcal{E}_r(B_1)$ | $\mathcal{E}(B_2)$  | $\mathcal{E}_r(E)$  |
|--------------------|-----------------------|-------------------------|-------------------------|----------------------|---------------------|---------------------|
| $200 \times 200$   | $8.5943\text{e-}14$   | $1.3769\text{e-}12$     | $5.8536\text{e-}11$     | $7.8382\text{e-}15$  | $4.3338\text{e-}11$ | $8.9687\text{e-}14$ |
| $400 \times 400$   | $2.6538\text{e-}15$   | $8.6998\text{e-}12$     | $7.5751\text{e-}10$     | $1.7876\text{e-}15$  | $1.7969\text{e-}10$ | $3.9629\text{e-}15$ |
| $800 \times 800$   | $2.0471\text{e-}15$   | $1.8790\text{e-}11$     | $1.4140\text{e-}09$     | $2.8231\text{e-}15$  | $1.1313\text{e-}10$ | $9.6211\text{e-}15$ |
| $1600 \times 1600$ | $2.1693\text{e-}14$   | $7.6349\text{e-}08$     | $6.4150\text{e-}07$     | $9.9568\text{e-}14$  | $5.3582\text{e-}09$ | $1.8401\text{e-}14$ |

Table 7.4.: Blast wave test. Conservation properties of the FV-FEEC scheme. Error (7.1) of the MHD conserved variables at different mesh resolutions.

Numerical instabilities are a typical outcome of the blast wave test even for lowest order finite volume approximations of the full MHD system, see [Fuc+11, pp. 356–357]. The blast wave MHD flow is characterized by outward- and inward-going fast magnetosonic shock and the magnetic induction field experiences a strong compression on account of the explosion. As it can be observed in Figure 7.10, the FV-FEEC scheme is robust, oscillations-free and it approximates the shocks rather sharply, being however diffusive on account of the first order accuracy. The second row of Figure 7.10 shows the magnetic induction field lines: no fluctuations are observed, not even in the middle of the computational domain (compare with [Bal98, Figure 9] and [Fuc+11, Figure 8]).

## 7.2. Assessment of FV-DG Schemes for MHD Coupling

As extensively discussed in Chapter 3, genuinely discontinuous stabilized Galerkin discretizations of the magnetic advection problem are a viable alternative to FEEC-based schemes. This section pertains to

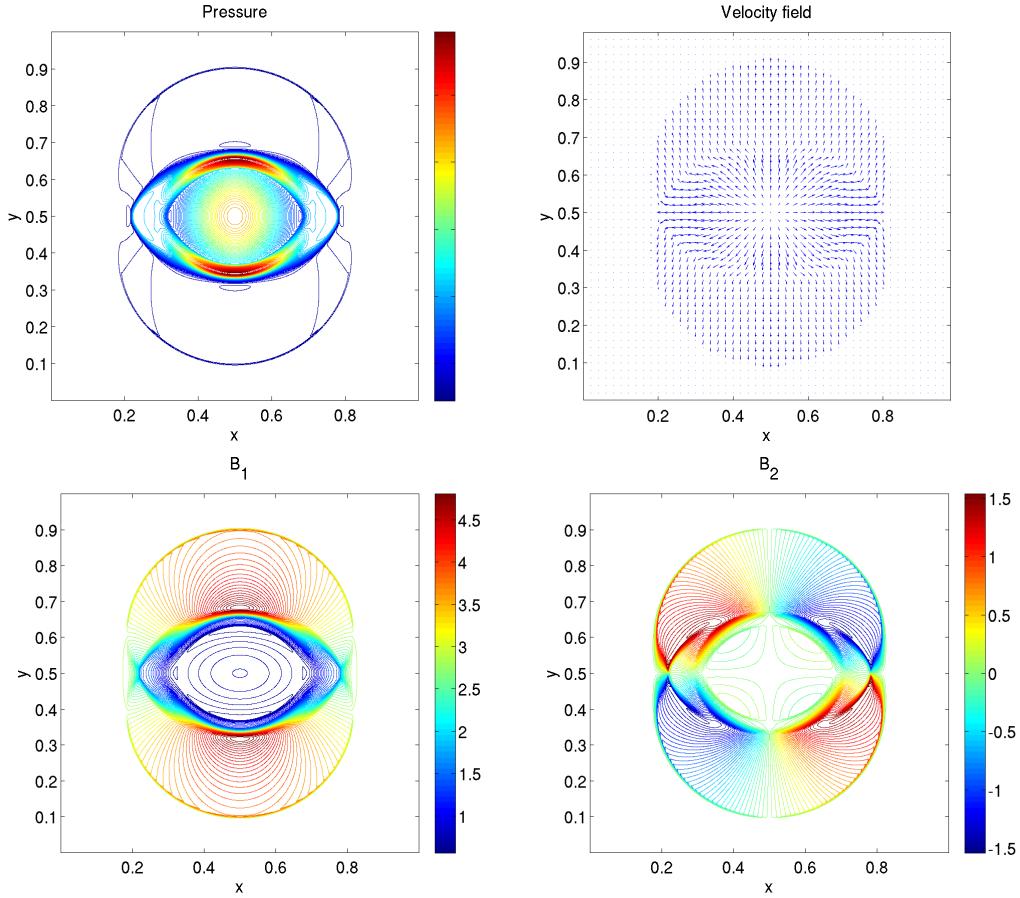


Figure 7.10.: Blast wave test. Numerical discretization on a  $1600 \times 1600$  Cartesian mesh with lowest order FV-FEEC scheme and Heun time-stepping,  $C_{\text{CFL}} = 0.4$ . Finite volume scheme for the extended Euler equations using HLL approximate Riemann solver.

stabilized discontinuous Galerkin-based schemes for the ideal MHD problem, which we name FV-DG, and to their comparison with the FV-FEEC discretization tested in Section 7.1. A hasty application of the DG schemes devised in Chapter 3 to the scalar advection problem within the framework of the two-dimensional planar MHD system runs into issues: a piecewise constant magnetic potential on a given Cartesian mesh does not allow to determine a magnetic induction field with globally well-defined traces at the mesh interfaces, if a standard discretization of the exterior derivative (the two-dimensional **curl** here) is used; in addition, since the two-dimensional **curl** of the magnetic potential will not be (globally) well-defined, we need to introduce a suitable definition of a local discrete divergence operator. While little can be done to prevent the latter shortcoming, owing to the intrinsic nature of DG schemes, the first issue can be addressed by staggering the discretization of the magnetic potential with respect to the fluid variables in a way that resembles a Yee-type scheme and constrained transport methods. This would yield a meaningful definition of the incidence matrix (*cf.* Section 2.3) connecting the potential to the induction field. The entwinement of electromagnetic fields and fluid variables in this staggered approach requires, in addition, an interpolation of the velocity field on the dual mesh where the magnetic advection problem is discretized.

### Piecewise Constant Upwind Discontinuous Galerkin on Dual Mesh

Let  $\Omega \subset \mathbb{R}^2$  be a Cartesian domain. We consider a tensor product partition  $\mathcal{T}_h$  under Assumption 2.4.2 and the corresponding dual mesh  $\mathcal{T}_h^d$ . Throughout the present section, we use integer indices  $i, j \in \mathbb{N}$  to refer to the elements on the dual mesh; fractional indices are associated with the primal mesh as in Figure 7.11. In order to keep the presentation focused and avoid the proliferation of many terms, we

restrict to periodic boundary conditions at the boundary  $\partial\Omega$ . The derivation of the scheme when boundary conditions at the inflow boundary are supplied, can be derived straightforwardly as in Chapter 3.

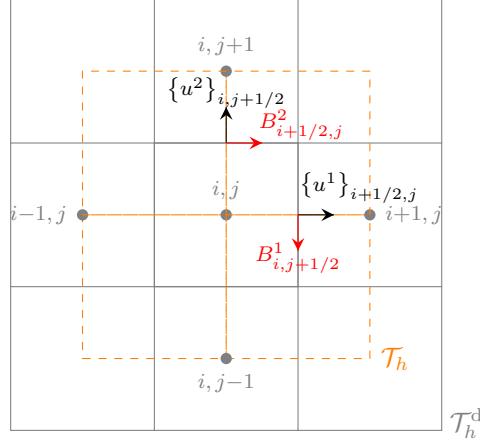


Figure 7.11.: Sketch of the primal-dual stencil for piecewise constant stabilized discontinuous Galerkin schemes.

Let us consider the initial boundary value problem for the scalar advection of the magnetic potential:

$$\begin{aligned}\partial_t A + \mathbf{u} \cdot \operatorname{grad} A &= f, && \text{in } \Omega \times I, \\ A(\mathbf{x}, 0) &= A_0(\mathbf{x}), && \text{in } \Omega.\end{aligned}$$

The lowest order stabilized discontinuous Galerkin semi-discrete scheme can be derived from (3.8) and (3.9) and reads: Find  $A(t) \in \mathcal{P}_0(\mathcal{T}_h^d)$  such that, for all  $\varphi \in \mathcal{P}_0(\mathcal{T}_h^d)$ ,

$$\sum_{T_{i,j} \in \mathcal{T}_h^d} \int_{T_{i,j}} \partial_t A \varphi - A \varphi \operatorname{div} \mathbf{u} + \sum_{f \in \Delta_1(\mathcal{T}_h^d)} \int_f \mathbf{u} \cdot \mathbf{n}_f (\{A\}_f [\varphi]_f + c_f [A]_f [\varphi]_f) = 0. \quad (7.2)$$

Since  $\mathbf{u} \in \mathcal{P}_0(\mathcal{T}_h)^2$ , the trace of the velocity field at the mesh interfaces of the dual mesh is not uniquely defined. A standard averaging approach yields an approximation of the normal velocity  $\mathbf{u} \cdot \mathbf{n}_f$  on each  $f \in \Delta_1(\mathcal{T}_h^d)$  as follows,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{n}_{f_{i \pm 1/2,j}} &\approx \{u^1\}_{i \pm 1/2,j} = \frac{u_{i \pm 1/2,j+1/2}^1 + u_{i \pm 1/2,j-1/2}^1}{2}, && \forall i, j, \\ \mathbf{u} \cdot \mathbf{n}_{f_{i,j \pm 1/2}} &\approx \{u^2\}_{i,j \pm 1/2} = \frac{u_{i+1/2,j \pm 1/2}^2 + u_{i-1/2,j \pm 1/2}^2}{2}, && \forall i, j.\end{aligned}$$

Let  $f = f_{i+1/2,j} \in \Delta_1(\mathcal{T}_h^d)$ , with the convention that the normal  $\mathbf{n}_f := \mathbf{n}_{f_{i+1/2,j}}$  points from the element  $T_{i,j}$  to the element  $T_{i+1,j}$ . The choice of  $c_f = 1/2 \operatorname{sign}(\mathbf{u} \cdot \mathbf{n}_f)$  entails upwind fluxes. Indeed,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{n}_f (\{A\}_f + c_f [A]_f) &\approx \{u^1\}_{i+1/2,j} \begin{cases} A_{i,j} & \text{if } \{u^1\}_{i+1/2,j} > 0, \\ A_{i+1,j} & \text{if } \{u^1\}_{i+1/2,j} < 0, \end{cases} \\ &= \max\{0, \{u^1\}_{i+1/2,j}\} A_{i,j} + \min\{0, \{u^1\}_{i+1/2,j}\} A_{i+1,j},\end{aligned}$$

and analogous equivalences hold for all other edges in  $\Delta_1(\mathcal{T}_h^d)$ . Moreover, a discrete divergence operator on the dual mesh  $\mathcal{T}_h^d$  needs to be defined to evaluate the cell terms in (7.2). The central differences operator is a natural choice, i.e.,

$$(\operatorname{div} \mathbf{w})_{i,j} \approx \frac{w_{i+1/2,j}^1 - w_{i-1/2,j}^1}{h} + \frac{w_{i,j+1/2}^2 - w_{i,j-1/2}^2}{h}, \quad (7.3)$$

for  $\mathbf{w} \in \mathcal{P}_0(\mathcal{T}_h^d)$ . Therefore, straightforward computations allow to reformulate the piecewise constant upwind stabilized Galerkin scheme (7.2) as a finite difference schemes with upwind fluxes. On each

element  $T_{i,j} \in \mathcal{T}_h^d$  of the dual mesh, the local spatial piecewise constant scheme with explicit Euler time-stepping reads

$$A_{i,j}^{n+1} = A_{i,j}^n - \Delta t^n \max\{0, \{u^1\}_{i-1/2,j}\} \frac{A_{i,j}^n - A_{i-1,j}^n}{h} - \Delta t^n \min\{0, \{u^1\}_{i+1/2,j}\} \frac{A_{i+1,j}^n - A_{i,j}^n}{h} \\ - \Delta t^n \max\{0, \{u^2\}_{i,j-1/2}\} \frac{A_{i,j}^n - A_{i,j-1}^n}{h} - \Delta t^n \min\{0, \{u^2\}_{i,j+1/2}\} \frac{A_{i,j+1}^n - A_{i,j}^n}{h}.$$

Following the steps of the splitting Algorithms 1 and 2, once the magnetic advection problem for the potential is numerically solved, the magnetic induction field  $\mathbf{B}$  has to be computed from the potential  $A$ . Here we project the magnetic induction field into the space of lowest order Raviart–Thomas finite element. The traces of the  $\mathbf{B}$  field, required in the numerical fluxes of the finite volume scheme for the extended Euler equations, are defined at the interfaces of a fixed element  $T_{i+1/2,j+1/2} \in \mathcal{T}_h$  of the primal mesh as,

$$B_{i+1,j+1/2}^1 = \frac{A_{i+1,j+1} - A_{i+1,j}}{h}, \quad B_{i,j+1/2}^1 = \frac{A_{i,j+1} - A_{i,j}}{h}, \\ B_{i+1/2,j}^2 = \frac{A_{i,j} - A_{i+1,j}}{h}, \quad B_{i+1/2,j+1}^2 = \frac{A_{i,j+1} - A_{i+1,j+1}}{h},$$

and analogously for all other elements. The tangential trace is averaged at each mesh edge. It easily follows that the numerical scheme ensures, by construction, that the divergence constraint is satisfied up to machine precision.

The FV-DG scheme performed robustly on the Orszag–Tang benchmark at all the tested resolutions (until  $400 \times 400$  mesh elements), with pressure values comparable with those in Table 7.1. However, the first order piecewise constant discretization of the advection problem in the full ideal MHD coupling proved meager in tackling low-beta plasma flows: on the rotor problem described in Section 7.1.5, negative values of the pressure are encountered (on a  $200 \times 200$  Cartesian mesh for CFL constants  $C_{\text{CFL}} \in \{0.05, 0.4\}$ ).

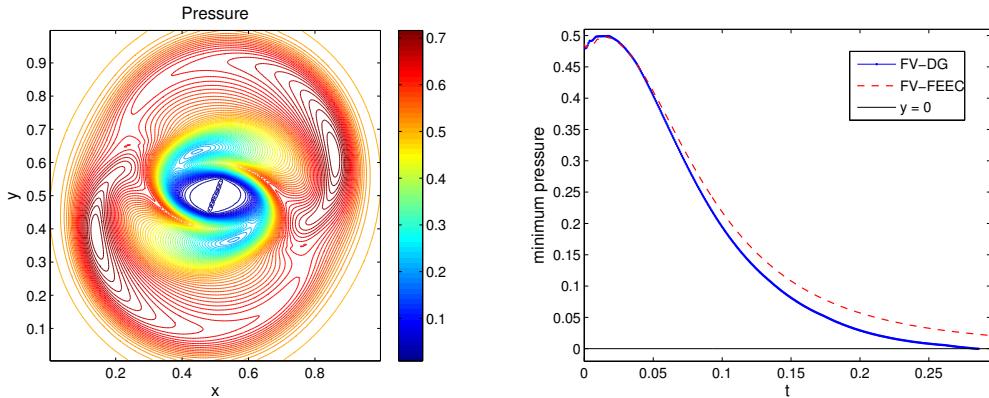


Figure 7.12.: Rotor problem. Pressure at time  $t = 0.2761$  on a  $200 \times 200$  Cartesian mesh (left) and minimum of the pressure vs. time (right). Piecewise constant DG discretization of the magnetic advection problem coupled with lowest order finite volume schemes with HLL solver for the extended Euler equations. Explicit Euler time-stepping.

Figure 7.12 (right) shows the evolution of the minimum of the pressure over time compared with the FEEC-FV scheme under the same parameters (same approximate Riemann solver, CFL condition, mesh resolution and problem data). The numerical pressure obtained via the FV-DG scheme has dropped to negative values in the interior of the rotor (Figure 7.12 on the left). The onset of spurious solutions and negative pressures cannot be ascribed to the failure of the scheme in preserving a solenoidal magnetic field (it has been also numerically verified the divergence of the  $\mathbf{B}$  field vanishes up to machine precision). This lack of correlation has been observed before in [T6t00, Section 6.6] for a eight-wave finite volume scheme based on constrained transport methods.

### 7.3. Towards Second Order Schemes

Ascertained in Section 7.1 that the lowest order FV-FEEC schemes are first order accurate for smooth solutions, possess built-in structure-preserving properties, and perform robustly in many challenges MHD benchmark tests, the natural next step would be to design formally second and higher order accurate FV-FEEC methods.

As a first attempt, we simply couple the second order extrusion contraction upwind scheme for the advection of the magnetic potential to finite volume discretizations of the extended Euler system, based on the two-wave HLL Riemann solver, piecewise linear reconstruction and the minmod limiter. A second order two-stage SSP-RK time-stepping, with CFL constant  $C_{\text{CFL}} = 0.1$ , provides the temporal discretization. The algorithm is tested on the Orszag–Tang benchmark from Section 7.1.4. On a  $200 \times 200$  Cartesian mesh, the code runs into a (local) point with negative pressure, encountered at the level of the updated variables and not on the reconstructed ones. We suspect that spurious build-up of the magnetic field values close to the bottom boundary, as shown in Figure 7.13, causes an anomalous magnetic energy with associated decrease of the pressure until the latter drops to negative values.

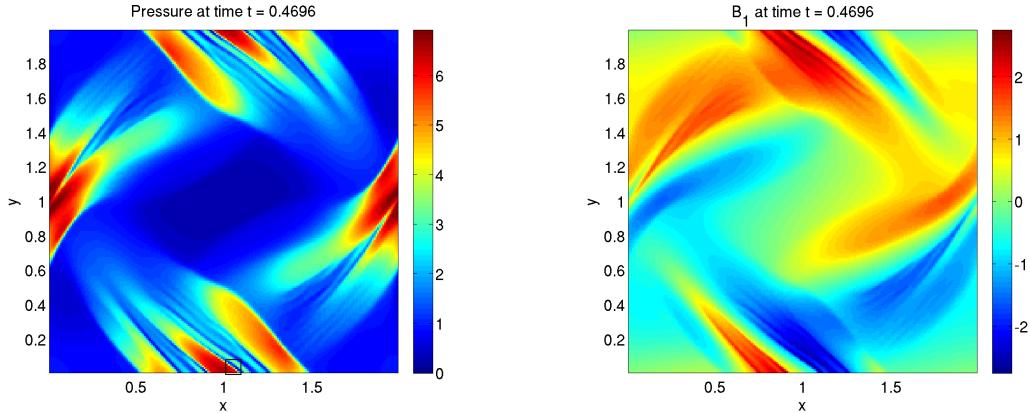


Figure 7.13.: Orszag–Tang benchmark. Plots of the discrete pressure and first component of the magnetic induction field on a  $200 \times 200$  Cartesian mesh, at the time of the algorithm crash,  $t = 0.4696$ . The black box, in the plot on the right, frames the mesh element where the pressure is negative.

Modifying the aforementioned FV-FEEC algorithm with the addition, to the extrusion contraction upwind scheme, of a residual-based nonlinear stabilization, as described in Section 4.4.1, yields similar results. Localized instabilities, in the vicinity of shocks in the solution, provoke a drop of the pressure to negative values. On the other hand, coupling the FEEC-based second order scheme with a lowest order finite volume discretization does not lead to non-admissible states.

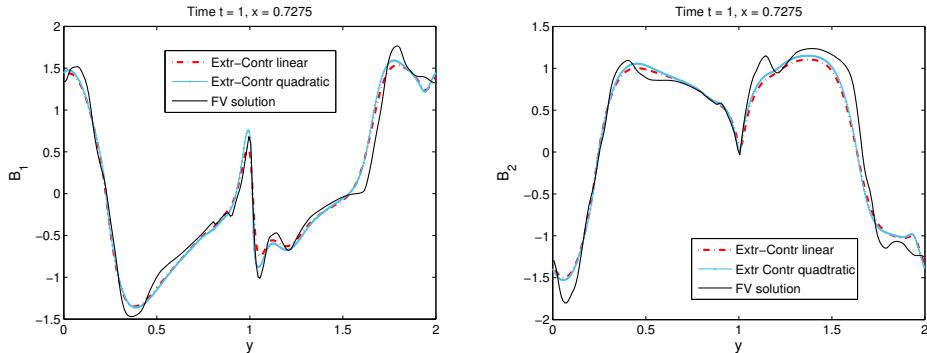


Figure 7.14.: Orszag–Tang benchmark. Projections of the numerical induction field on a  $200 \times 200$  Cartesian mesh for constant values of  $x$ . Comparison of FV-FEEC discretizations based on piecewise bilinear and piecewise biquadratic extrusion contraction upwind schemes. Second order two-stage SSP-RK time-stepping with  $C_{\text{CFL}} = 0.1$ , and lowest order finite volume discretizations of the extended Euler system are employed in both cases.

We are well aware that such FV-FEEC scheme cannot be more than first order accurate, but we would like to monitor the impact of the increased accuracy of the magnetic induction field on the fluid variables approximation. Simulations on the Orszag–Tang benchmark, Figure 7.14, show that the piecewise quadratic extrusion contraction upwind scheme yields slightly more accurate solutions than the piecewise linear scheme, when compared to the reference finite volume solution of the full MHD system. Although the nonlinear degenerate viscosity terms have not been added to the scheme in the present experiment, the numerical induction field is non-oscillatory near steep gradients. We conjecture that the diffusive behavior of the second order extrusion contraction scheme can be ascribed to the tight interplay between the discrete electromagnetic fields and the fluid variables, which are approximated at first order. On the rotor benchmark and on the fast wave expansion test, the FV-FEEC with piecewise quadratic extrusion contraction performed similarly at all the mesh resolution we tried (up until  $800 \times 800$  mesh elements).

With the aim of conceiving ways to counteract the development of non-admissible numerical states, we identify three main ingredients for the design of second order FV-FEEC schemes:

1. A second order discretization of the transient advection problem,
  - a) able to supply a magnetic induction field accurate to second order;
  - b) endowed with a nonlinear mechanism of oscillations damping, capable of ensuring some TVD-like property without affecting the accuracy of the scheme.
2. A second order extension of the finite volume scheme for the extended Euler equations via reconstruction and limiting, with controls on the preservation and evolution of admissible states.
3. A suitable tuning of the splitting algorithm in the coupling step, so that no additional error associated with the splitting is introduced.

In light of the above considerations, we can single out two possible sources of issues in discretely evolving admissible solutions: the nonlinear artificial viscosity entering the extrusion contraction upwind might not be “strong enough”; second, a clipping of the reconstructed fluid variables and a further positive fix to guarantee admissible updated variables might be indispensable. The difficulties in properly tuning the residual-based viscosity for extrusion contraction schemes have been highlighted in Section 4.4. Devising a positive fix for the finite volume discretizations of the extended Euler system seems an even more challenging task since, as pointed out in Section 6.3.1, the lowest order scheme is not provably positively conservative. Both questions yield interesting topics for further investigations.

## 8. Concluding Remarks

We conclude this thesis with a short summary of its content and an outlook on possible directions of further investigation.

We have developed a family of numerical methods to solve the single-fluid standard MHD problem by coupling two different spatial discretizations of fluid and electromagnetic variables. The evolution of the electromagnetic fields relies on FEEC-based finite element approximations, designed to accommodate discontinuous advection velocities which inevitably occur in MHD flows. FEEC-based methods applied to the magnetic advection problem for the potential ensure that the divergence constraint is satisfied exactly, and no mesh-staggering of fluid and electromagnetic variables, typical of constrained transport and “central schemes” for hyperbolic problems, is required. Stability and a priori convergence estimates have been established for explicit Runge–Kutta time-stepping and stabilized Galerkin discretizations of the generalized advection problem on structured and unstructured meshes. Concerning extrusion contraction upwind schemes, the observed numerical results suggest that the methods possess good stability and approximation properties. Two features are particularly attractive: the intrinsic upwinding, which acts as a linear stabilization in the presence of boundary and internal layers, and the commuting property of the discrete Lie derivative with the exterior derivative, which guarantees that closed forms are Lie advected into closed forms. The latter property allows to directly discretize the transport problem for the induction field without compromising the divergence constraint. It would naturally be desirable to supply the extrusion contraction upwind schemes with theoretical foundation by establishing stability and convergence estimates.

High order polynomial discretizations are subject to the development of spurious oscillations in the vicinity of steep gradients and discontinuities. We have considered the construction of flux limiting through the addition of residual-based artificial viscosity to the discrete advection problem for the magnetic potential. Owing to the discrete commuting diagram property, the extrusion contraction upwind scheme can accommodate nonlinear viscosity terms weighted on the basis of the residual of the magnetic induction equation. This strategy cannot be applied straightforwardly to (discontinuous) Galerkin discretizations since a discrete commuting property does not seem to hold. In this case it might be advisable to design ad hoc total variation bounded limiters calibrated by functionals of the magnetic induction field. Open to further investigations is also the combination of extrusion contraction upwind approximations with residual- and entropy-based viscosity schemes for the magnetic induction equation itself.

Concerning the fluid dynamics part of the MHD model, we treated the balance equations for the fluid as a system of conservation laws with a varying coefficient, the magnetic induction field. Finite volume schemes have been used for the numerical discretizations of the extended Euler equations and they hinge on approximate Riemann solvers tailored to accommodate the presence of the magnetic induction. The further adaptation of this construction to design numerical fluxes yielding a discrete version of the entropy inequality would pave the way to entropy stable schemes of arbitrarily high order.

Coupling the FEEC-based methods for the magnetic advection with the finite volume schemes for the fluid results in numerical methods for the ideal MHD system, which proved stable and competitive in several two-dimensional tests. The synchronous-coupling algorithm allows to parallelize, at each time step, the solvers for the two subsystems. The extension of the fully coupled scheme to unstructured meshes and efficient three-dimensional simulations are essential for realistic applications. At this stage, we did not aim at computationally efficient realistic simulations but rather at evidence of the increased robustness of the novel numerical schemes on model problems available in literature. The promising numerical results obtained when using discretizations of the electromagnetic fields based on discrete differential forms, even in the presence of complex flows, suggest that structure-preserving conforming discretizations can be competitive also in computational fluid dynamics, typically preserve of finite volume and discontinuous Galerkin methods. To design formally high order schemes for the full MHD

## 8. Concluding Remarks

---

model some additional challenges have to be addressed. Specifically, how to couple the residual-based viscosity schemes for extrusion contraction with non-oscillatory reconstructions for the equations of fluid mechanics, how should the latter cope with the presence of the magnetic induction field, and how to implement a positive fix to ensure that high order finite volume schemes evolve physically admissible solutions.

In the resistive MHD problem, using implicit-explicit time-stepping together with FEEC-based spatial discretizations requires fast iterative solvers to handle the large algebraic systems which ensue from the discretization, at each time step, of boundary value problems for the double **curl** operator. Once the discretization relies on discrete differential forms, efficient solvers are available. For discontinuous Galerkin schemes, we have designed a family of preconditioners based on the auxiliary space method and proved their robustness with respect to the mesh width and to locally dominant transport (under some assumptions on the coefficients distribution). Hence, the FEEC-based conforming and non-conforming schemes provide a natural way to discretize the double **curl** operator and might enjoy a decisive advantage, also in terms of efficiency, over existing finite volume methods.

# Bibliography

- [Agm65] Shmuel Agmon. *Lectures on elliptic boundary value problems*. Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Van Nostrand Mathematical Studies, No. 2. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965, pp. v+291 (cit. on p. 97).
- [Amb04] Luigi Ambrosio. “Transport equation and Cauchy problem for  $BV$  vector fields.” In: *Invent. Math.* 158.2 (2004), pp. 227–260 (cit. on p. 39).
- [Amb11] Luigi Ambrosio. “The flow associated to weakly differentiable vector fields: recent results and open problems.” In: *Nonlinear conservation laws and applications*. Vol. 153. IMA Vol. Math. Appl. Springer, New York, 2011, pp. 181–193 (cit. on p. 39).
- [Ant+15] Paola F. Antonietti, Blanca Ayuso de Dios, Silvia Bertoluzza, and Micol Pennacchio. “Substructuring preconditioners for an  $h\text{-}p$  domain decomposition method with interior penalty mortaring.” In: *Calcolo* 52.3 (2015), pp. 289–316 (cit. on p. 93).
- [Ant+12] Paola F. Antonietti, Blanca Ayuso de Dios, Susanne C. Brenner, and Li-Yeng Sung. “Schwarz methods for a preconditioned WOPSIP method for elliptic problems.” In: *Comput. Methods Appl. Math.* 12.3 (2012), pp. 241–272 (cit. on p. 93).
- [AA07] Paola F. Antonietti and Blanca Ayuso. “Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case.” In: *M2AN Math. Model. Numer. Anal.* 41.1 (2007), pp. 21–54 (cit. on p. 92).
- [AA08] Paola F. Antonietti and Blanca Ayuso. “Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems.” In: *M2AN Math. Model. Numer. Anal.* 42.3 (2008), pp. 443–469 (cit. on p. 92).
- [AH11] Paola F. Antonietti and Paul Houston. “A class of domain decomposition preconditioners for  $hp$ -discontinuous Galerkin finite element methods.” In: *J. Sci. Comput.* 46.1 (2011), pp. 124–149 (cit. on p. 93).
- [ASV15] Paola F. Antonietti, Marco Sarti, and Marco Verani. “Multigrid algorithms for  $hp$ -discontinuous Galerkin discretizations of elliptic problems.” In: *SIAM J. Numer. Anal.* 53.1 (2015), pp. 598–618 (cit. on p. 93).
- [Arn82] Douglas N. Arnold. “An interior penalty finite element method with discontinuous elements.” In: *SIAM J. Numer. Anal.* 19.4 (1982), pp. 742–760 (cit. on pp. 94, 95).
- [Arn13] Douglas N. Arnold. “Spaces of finite element differential forms.” In: *Analysis and numerics of partial differential equations*. Vol. 4. Springer INdAM Ser. Springer, Milan, 2013, pp. 117–140 (cit. on p. 32).
- [AA11] Douglas N. Arnold and Gerard Awanou. “The serendipity family of finite elements.” In: *Found. Comput. Math.* 11.3 (2011), pp. 337–344 (cit. on p. 33).
- [AA14] Douglas N. Arnold and Gerard Awanou. “Finite element differential forms on cubical meshes.” In: *Math. Comp.* 83.288 (2014), pp. 1551–1570 (cit. on pp. 28, 30).
- [ABB15] Douglas N. Arnold, Daniele Boffi, and Francesca Bonizzoni. “Finite element differential forms on curvilinear cubic meshes and their approximation properties.” In: *Numer. Math.* 129.1 (2015), pp. 1–20 (cit. on pp. 22, 28–31).
- [ABF05] Douglas N. Arnold, Daniele Boffi, and Richard S. Falk. “Quadrilateral  $\mathbf{H}(\text{div})$  finite elements.” In: *SIAM J. Numer. Anal.* 42.6 (2005), pp. 2429–2451 (cit. on p. 128).
- [Arn+02] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. “Unified analysis of discontinuous Galerkin methods for elliptic problems.” In: *SIAM J. Numer. Anal.* 39.5 (2001/02), pp. 1749–1779 (cit. on pp. 37, 92, 94).
- [AFW00] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. “Multigrid in  $\mathbf{H}(\text{div})$  and  $\mathbf{H}(\text{curl})$ .” In: *Numer. Math.* 85.2 (2000), pp. 197–217 (cit. on pp. 92, 102).
- [AFW06] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. “Finite element exterior calculus, homological techniques, and applications.” In: *Acta Numer.* 15 (2006), pp. 1–155 (cit. on pp. x, 15, 18, 21, 23, 27, 28, 30–33, 35, 118).
- [Ayu+14a] Blanca Ayuso de Dios, Franco Brezzi, L. Donatella Marini, Jinchao Xu, and Ludmil Zikatanov. “A simple preconditioner for a discontinuous Galerkin method for the Stokes problem.” In: *J. Sci. Comput.* 58.3 (2014), pp. 517–547 (cit. on p. 93).
- [Ayu+13a] Blanca Ayuso de Dios, Ivan Georgiev, Johannes Kraus, and Ludmil Zikatanov. “A subspace correction method for discontinuous Galerkin discretizations of linear elasticity equations.” In: *ESAIM Math. Model. Numer. Anal.* 47.5 (2013), pp. 1315–1333 (cit. on p. 93).

- [AHP16] Blanca Ayuso de Dios, Ralf Hiptmair, and Cecilia Pagliantini. “Auxiliary space preconditioners for SIP-DG discretizations of  $H(\text{curl})$ -elliptic problems with discontinuous coefficients.” In: *IMA Journal of Numerical Analysis* 10.1093/imanum/drw018 (2016) (cit. on p. 94).
- [Ayu+14b] Blanca Ayuso de Dios, Michael Holst, Yunrong Zhu, and Ludmil Zikatanov. “Multilevel preconditioners for discontinuous Galerkin approximations of elliptic problems with jump coefficients.” In: *Math. Comp.* 83.287 (2014), pp. 1083–1120 (cit. on p. 93).
- [Ayu+13b] Blanca Ayuso de Dios, Ariel Lombardi, Paola Pietra, and Ludmil Zikatanov. “A block solver for the exponentially fitted IIPG-0 method.” In: *Domain decomposition methods in science and engineering XX*. Vol. 91. Lect. Notes Comput. Sci. Eng. Springer, Heidelberg, 2013, pp. 239–246 (cit. on p. 93).
- [AZ09] Blanca Ayuso de Dios and Ludmil Zikatanov. “Uniformly convergent iterative methods for discontinuous Galerkin discretizations.” In: *J. Sci. Comput.* 40.1-3 (2009), pp. 4–36 (cit. on p. 93).
- [Bak77] Garth A. Baker. “Finite element methods for elliptic equations using nonconforming elements.” In: *Math. Comp.* 31.137 (1977), pp. 45–59 (cit. on p. 94).
- [Bak83] Garth A. Baker. “Combinatorial Laplacians and Sullivan-Whitney forms.” In: *Differential geometry (College Park, Md., 1981/1982)*. Vol. 32. Progr. Math. Birkhäuser, Boston, Mass., 1983, pp. 1–33 (cit. on p. 23).
- [Bal98] Dinshaw S. Balsara. “Total variation diminishing scheme for adiabatic and isothermal magnetohydrodynamics.” In: *The Astrophysical Journal Supplement Series* 116.1 (1998), p. 133 (cit. on p. 159).
- [Bal01] Dinshaw S. Balsara. “Divergence-free adaptive mesh refinement for magnetohydrodynamics.” In: *J. Comput. Phys.* 174.2 (Dec. 2001), pp. 614–648 (cit. on p. 9).
- [Bal04] Dinshaw S. Balsara. “Second-order-accurate schemes for magnetohydrodynamics with divergence-free reconstruction.” In: *The Astrophysical Journal Supplement Series* 151.1 (2004), p. 149 (cit. on pp. 142, 151).
- [BS99] Dinshaw S. Balsara and Daniel S. Spicer. “A staggered mesh algorithm using high order Godunov fluxes to ensure solenoidal magnetic fields in magnetohydrodynamic simulations.” In: *J. Comput. Phys.* 149.2 (1999), pp. 270–292 (cit. on pp. 81, 157).
- [Bar+11] Andrew T. Barker, Susanne C. Brenner, Eun-Hee Park, and Li-Yeng Sung. “Two-level additive Schwarz preconditioners for a weakly over-penalized symmetric interior penalty method.” In: *J. Sci. Comput.* 47.1 (2011), pp. 27–49 (cit. on p. 93).
- [BKP96] Aleksei A. Barmin, Andrei G. Kulikovskiy, and Nikolai V. Pogorelov. “Shock-capturing approach and nonevolutionary solutions in magnetohydrodynamics.” In: *J. Comput. Phys.* 126.1 (1996), pp. 77–90 (cit. on p. 6).
- [Bat+97] Paul Batten, Nicholas Clarke, Claire Lambert, and Derek M. Causon. “On the choice of wavespeeds for the HLLC Riemann solver.” In: *SIAM J. Sci. Comput.* 18.6 (1997), pp. 1553–1570 (cit. on pp. 141, 142).
- [Bis93] Dieter Biskamp. *Nonlinear magnetohydrodynamics*. Cambridge Monographs on Plasma Physics 1. Cambridge University Press, Cambridge, U.K., 1993 (cit. on p. 4).
- [BBF13] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed finite element methods and applications*. Vol. 44. Springer Series in Computational Mathematics. Springer, Heidelberg, 2013, pp. xiv+685 (cit. on pp. 17, 34).
- [BGP14] Andrea Bonito, Jean-Luc Guermond, and Bojan Popov. “Stability analysis of explicit entropy viscosity methods for non-linear scalar conservation equations.” In: *Math. Comp.* 83.287 (2014), pp. 1039–1062 (cit. on p. 82).
- [BB97] Jay P. Boris and David L. Book. “Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works.” In: *J. Comput. Phys.* 135.2 (1997), pp. 170–186 (cit. on p. 82).
- [Bos88] Alain Bossavit. “Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism.” In: *Physical Science, Measurement and Instrumentation, Management and Education - Reviews, IEE Proceedings A* 135.8 (Nov. 1988), pp. 493–500 (cit. on p. 23).
- [Bos98a] Alain Bossavit. *Computational electromagnetism*. Electromagnetism. Academic Press, Inc., San Diego, CA, 1998, pp. xx+352 (cit. on p. 24).
- [Bos98b] Alain Bossavit. “On the geometry of electromagnetism. (4): ‘Maxwell’s house’.” In: *J. Japan Soc. Appl. Electromagn. and Mech.* 6.4 (1998), pp. 318–326 (cit. on p. 17).
- [Bos03] Alain Bossavit. “Extrusion, contraction: their discretization via Whitney forms.” In: *COMPEL* 22.3 (2003), pp. 470–480 (cit. on pp. 26, 69–72).
- [BC13] François Bouchut and Gianluca Crippa. “Lagrangian flows for vector fields with gradient given by a singular integral.” In: *J. Hyperbolic Differ. Equ.* 10.2 (2013), pp. 235–282 (cit. on p. 39).
- [BKW07] François Bouchut, Christian Klingenberg, and Knut Waagan. “A multiwave approximate Riemann solver for ideal MHD based on relaxation. I. Theoretical framework.” In: *Numer. Math.* 108.1 (2007), pp. 7–42 (cit. on p. 142).
- [Boy12] Franck Boyer. “Analysis of the upwind finite volume method for general initial- and boundary-value transport problems.” In: *IMA J. Numer. Anal.* 32.4 (2012), pp. 1404–1439 (cit. on p. 38).
- [BB80] Jeremiah U. Brackbill and Daniel C. Barnes. “The effect of nonzero  $\nabla \cdot \mathbf{B}$  on the numerical solution of the magnetohydrodynamic equations.” In: *J. Comput. Phys.* 35.3 (1980), pp. 426–430 (cit. on p. 9).
- [BH83] Dietrich Braess and Wolfgang Hackbusch. “A new convergence proof for the multigrid method including the  $V$ -cycle.” In: *SIAM J. Numer. Anal.* 20.5 (1983), pp. 967–975 (cit. on p. 93).

- [BPX91] James H. Bramble, Joseph E. Pasciak, and Jinchao Xu. “The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms.” In: *Math. Comp.* 56.193 (1991), pp. 1–34 (cit. on p. 93).
- [Bre02] Susanne C. Brenner. “Convergence of the multigrid  $V$ -cycle algorithm for second-order boundary value problems without full elliptic regularity.” In: *Math. Comp.* 71.238 (2002), pp. 507–525 (cit. on p. 93).
- [Bre+11] Susanne C. Brenner, Jintao Cui, Thirupathi Gudi, and Li-Yeng Sung. “Multigrid algorithms for symmetric discontinuous Galerkin methods on graded meshes.” In: *Numer. Math.* 119.1 (2011), pp. 21–47 (cit. on p. 93).
- [BO07] Susanne C. Brenner and Luke Owens. “A  $W$ -cycle algorithm for a weakly over-penalized interior penalty method.” In: *Comput. Methods Appl. Mech. Engrg.* 196.37–40 (2007), pp. 3823–3832 (cit. on p. 93).
- [BPS13] Susanne C. Brenner, Eun-Hee Park, and Li-Yeng Sung. “A balancing domain decomposition by constraints preconditioner for a weakly over-penalized symmetric interior penalty method.” In: *Numer. Linear Algebra Appl.* 20.3 (2013), pp. 472–491 (cit. on p. 93).
- [BS08] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*. Third. Vol. 15. Texts in Applied Mathematics. Springer, New York, 2008, pp. xviii+397 (cit. on pp. 51, 101).
- [BS06] Susanne C. Brenner and Li-Yeng Sung. “Multigrid algorithms for  $C^0$  interior penalty methods.” In: *SIAM J. Numer. Anal.* 44.1 (2006), pp. 199–223 (cit. on p. 93).
- [BW05] Susanne C. Brenner and Kening Wang. “Two-level additive Schwarz preconditioners for  $C^0$  interior penalty methods.” In: *Numer. Math.* 102.2 (2005), pp. 231–255 (cit. on p. 92).
- [BZ05] Susanne C. Brenner and Jie Zhao. “Convergence of multigrid algorithms for interior penalty methods.” In: *Appl. Numer. Anal. Comput. Math.* 2.1 (2005), pp. 3–18 (cit. on p. 93).
- [Bre11] Haïm Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011, pp. xiv+599 (cit. on p. 38).
- [BDM85] Franco Brezzi, Jim Douglas Jr., and L. Donatella Marini. “Two families of mixed finite elements for second order elliptic problems.” In: *Numer. Math.* 47.2 (1985), pp. 217–235 (cit. on p. 33).
- [BMS04] Franco Brezzi, L. Donatella Marini, and Endre Süli. “Discontinuous Galerkin methods for first-order hyperbolic problems.” In: *Math. Models Methods Appl. Sci.* 14.12 (2004), pp. 1893–1903 (cit. on pp. 38, 49, 51, 52, 61).
- [BW88] Moysey Brio and Cheng-Chin Wu. “An upwind differencing scheme for the equations of ideal magnetohydrodynamics.” In: *J. Comput. Phys.* 75.2 (1988), pp. 400–422 (cit. on pp. 5, 138).
- [Bri+15] Kolja Brix, Martin Campos Pinto, Claudio Canuto, and Wolfgang Dahmen. “Multilevel preconditioning of discontinuous Galerkin spectral element methods. Part I: geometrically conforming meshes.” In: *IMA J. Numer. Anal.* 35.4 (2015), pp. 1487–1532 (cit. on pp. 93, 110).
- [BCD08] Kolja Brix, Martin Campos Pinto, and Wolfgang Dahmen. “A multilevel preconditioner for the interior penalty discontinuous Galerkin method.” In: *SIAM J. Numer. Anal.* 46.5 (2008), pp. 2742–2768 (cit. on p. 93).
- [Bri+09] Kolja Brix, Martin Campos Pinto, Wolfgang Dahmen, and Ralf Massjung. “Multilevel preconditioners for the interior penalty discontinuous Galerkin method. II. Quantitative studies.” In: *Commun. Comput. Phys.* 5.2–4 (2009), pp. 296–325 (cit. on p. 93).
- [BHP07] Annalisa Buffa, Paul Houston, and Ilaria Perugia. “Discontinuous Galerkin computation of the Maxwell eigenvalues on simplicial meshes.” In: *J. Comput. Appl. Math.* 204.2 (2007), pp. 317–333 (cit. on pp. 93, 106, 119, 127).
- [BP06] Annalisa Buffa and Ilaria Perugia. “Discontinuous Galerkin approximation of the Maxwell eigenproblem.” In: *SIAM J. Numer. Anal.* 44.5 (2006), 2198–2226 (electronic) (cit. on pp. 32, 93, 106, 119).
- [BEF10] Erik Burman, Alexandre Ern, and Miguel A. Fernández. “Explicit Runge-Kutta schemes and finite elements with symmetric stabilization for first-order linear PDE systems.” In: *SIAM J. Numer. Anal.* 48.6 (2010), pp. 2019–2042 (cit. on pp. 50, 52, 54, 55, 58).
- [CS16] Martin Campos Pinto and Eric Sonnendrücker. “Gauss-compatible Galerkin schemes for time-dependent Maxwell equations.” In: *Math. Comp.* 85.302 (2016), pp. 2651–2685 (cit. on p. 32).
- [CPP14] Claudio Canuto, Luca F. Pavarino, and Alexandre B. Pieri. “BDDC preconditioners for continuous and discontinuous Galerkin methods using spectral/ $hp$  elements with variable local polynomial degree.” In: *IMA J. Numer. Anal.* 34.3 (2014), pp. 879–903 (cit. on pp. 93, 110).
- [CG97] Patricia Cargó and Gérard Gallice. “Roe matrices for ideal MHD and systematic construction of Roe matrices for systems of conservation laws.” In: *J. Comput. Phys.* 136.2 (1997), pp. 446–466 (cit. on p. 138).
- [Che16] Francis F. Chen. *Introduction to plasma physics and controlled fusion*. Third Edition. Switzerland: Springer, 2016, pp. xii+490 (cit. on pp. ix, 1).
- [CW08] Snorre H. Christiansen and Ragnar Winther. “Smoothed projections in finite element exterior calculus.” In: *Math. Comp.* 77.262 (2008), pp. 813–829 (cit. on p. 35).
- [CKW13] Eric T. Chung, Hyea Hyun Kim, and Olof B. Widlund. “Two-level overlapping Schwarz algorithms for a staggered discontinuous Galerkin method.” In: *SIAM J. Numer. Anal.* 51.1 (2013), pp. 47–67 (cit. on p. 93).
- [Cia78] Philippe G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978, pp. xix+530 (cit. on pp. 29, 47).
- [Cia91] Philippe G. Ciarlet. “Basic error estimates for elliptic problems.” In: *Handbook of numerical analysis, Vol. II*. Handb. Numer. Anal., II. Amsterdam: North-Holland, 1991, pp. 17–351 (cit. on p. 97).

- [Clé75] Philippe Clément. “Approximation by finite element functions using local regularization.” In: *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique* 9.R-2 (1975), pp. 77–84 (cit. on p. 35).
- [Coc+14] Bernardo Cockburn, Olivier Dubois, Jay Gopalakrishnan, and Shuguang Tan. “Multigrid for an HDG method.” In: *IMA J. Numer. Anal.* 34.4 (2014), pp. 1386–1425 (cit. on p. 93).
- [CHS90] Bernardo Cockburn, Suchung Hou, and Chi-Wang Shu. “The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case.” In: *Math. Comp.* 54.190 (1990), pp. 545–581 (cit. on p. 82).
- [CS98] Bernardo Cockburn and Chi-Wang Shu. “The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems.” In: *J. Comput. Phys.* 141.2 (1998), pp. 199–224 (cit. on pp. x, 8, 82).
- [CD08] Gianluca Crippa and Camillo De Lellis. “Regularity and compactness for the DiPerna-Lions flow.” In: *Hyperbolic problems: theory, numerics, applications*. Springer, Berlin, 2008, pp. 423–430 (cit. on p. 39).
- [Daf05] Constantine M. Dafermos. *Hyperbolic conservation laws in continuum physics*. Second. Vol. 325. Fundamental Principles of Mathematical Sciences. Springer-Verlag, Berlin, 2005, pp. xx+626 (cit. on p. 7).
- [De 06] Camillo De Lellis. “Notes on hyperbolic systems of conservation laws and transport equations.” In: *Handbook of Differential Equations: Evolutionary Equations* 3 (2006), pp. 277–383 (cit. on p. 39).
- [De 99] Hans De Sterck. “Numerical simulation and analysis of magnetically dominated MHD bow shock flows with applications in space physics.” PhD thesis. Katholieke Universiteit Leuven, Belgium, 1999, p. 325 (cit. on p. 7).
- [DP01] Hans De Sterck and Stefaan Poedts. “Overcompressive shocks and compound shocks in 2D and 3D magnetohydrodynamic flows.” In: *Proc. 8 Int. Conf. Hyperbolic Problems, Intl. Series of Numerical Mathematics* 141 (2001), pp. 791–801 (cit. on p. 7).
- [Ded+02] Andreas Dedner, Friedemann Kemm, Dietmar Kröner, Claus-Dieter Munz, Thomas Schnitzer, and Matthias Wesenberg. “Hyperbolic divergence cleaning for the MHD equations.” In: *J. Comput. Phys.* 175.2 (2002), pp. 645–673 (cit. on p. 10).
- [Des+05] Mathieu Desbrun, Anil N. Hirani, Melvin Leok, and Jerrold E. Marsden. “Discrete exterior calculus.” In: arXiv:math/0508341v2 (2005) (cit. on pp. 23, 24).
- [DL89] Ronald J. DiPerna and Pierre-Louis Lions. “Ordinary differential equations, transport theory and Sobolev spaces.” In: *Invent. Math.* 98.3 (1989), pp. 511–547 (cit. on p. 39).
- [Dir31] Paul A. M. Dirac. “Quantised singularities in the electromagnetic field.” In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 133.821 (1931), pp. 60–72 (cit. on p. 2).
- [Dob+06] Veselin A. Dobrev, Raytcho D. Lazarov, Panayot S. Vassilevski, and Ludmil T. Zikatanov. “Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations.” In: *Numer. Linear Algebra Appl.* 13.9 (2006), pp. 753–770 (cit. on p. 93).
- [DW12] Clark R. Dohrmann and Olof B. Widlund. “An iterative substructuring algorithm for two-dimensional problems in  $H(\text{curl})$ .” In: *SIAM J. Numer. Anal.* 50.3 (2012), pp. 1004–1028 (cit. on p. 92).
- [DW16] Clark R. Dohrmann and Olof B. Widlund. “A BDDC algorithm with deluxe scaling for three-dimensional  $H(\text{curl})$  problems.” In: *Comm. Pure Appl. Math.* 69.4 (2016), pp. 745–770 (cit. on pp. 92, 110).
- [Dry03] Maksymilian Dryja. “On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients.” In: *Comput. Methods Appl. Math.* 3.1 (2003), 76–85 (electronic) (cit. on p. 94).
- [DGS07] Maksymilian Dryja, Juan Galvis, and Marcus Sarkis. “BDDC methods for discontinuous Galerkin discretization of elliptic problems.” In: *J. Complexity* 23.4-6 (2007), pp. 715–739 (cit. on p. 93).
- [DGS12] Maksymilian Dryja, Juan Galvis, and Marcus Sarkis. “Neumann-Neumann methods for a DG discretization of elliptic problems with discontinuous coefficients on geometrically nonconforming substructures.” In: *Numer. Methods Partial Differential Equations* 28.4 (2012), pp. 1194–1226 (cit. on p. 93).
- [DGS13] Maksymilian Dryja, Juan Galvis, and Marcus Sarkis. “A FETI-DP preconditioner for a composite finite element and discontinuous Galerkin method.” In: *SIAM J. Numer. Anal.* 51.1 (2013), pp. 400–422 (cit. on p. 93).
- [DW87] Maksymilian Dryja and Olof B. Widlund. *An additive variant of the Schwarz alternating method in the case of many subregions*. Tech. rep. 339. New York: Department of Computer Science, Courant Institute, 1987 (cit. on p. 101).
- [Ein88] Bernd Einfeldt. “On Godunov-type methods for gas dynamics.” In: *SIAM J. Numer. Anal.* 25.2 (1988), pp. 294–318 (cit. on pp. 140, 142).
- [Ein+91] Bernd Einfeldt, Claus-Dieter Munz, Philip L. Roe, and Björn Sjögren. “On Godunov-type methods near low densities.” In: *J. Comput. Phys.* 92.2 (1991), pp. 273–295 (cit. on p. 138).
- [EG04] Alexandre Ern and Jean-Luc Guermond. *Theory and practice of finite elements*. Vol. 159. Applied Mathematical Sciences. Springer-Verlag, New York, 2004, pp. xiv+524 (cit. on p. 19).
- [EG06] Alexandre Ern and Jean-Luc Guermond. “Discontinuous Galerkin methods for Friedrichs’ systems. I. General theory.” In: *SIAM J. Numer. Anal.* 44.2 (2006), pp. 753–778 (cit. on p. 19).
- [EG15] Alexandre Ern and Jean-Luc Guermond. “Finite element quasi-interpolation and best approximation.” 2015 (cit. on p. 35).

- [Eva98] Lawrence C. Evans. *Partial Differential Equations*. Vol. 19. Graduate Studies in Mathematics. Providence, Rhode Island: American Mathematical Society, 1998 (cit. on p. 91).
- [FW14] Richard S. Falk and Ragnar Winther. “Local bounded cochain projections.” In: *Math. Comp.* 83.290 (2014), pp. 2631–2656 (cit. on p. 35).
- [FK01] Xiaobing Feng and Ohannes A. Karakashian. “Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems.” In: *SIAM J. Numer. Anal.* 39.4 (2001), 1343–1365 (electronic) (cit. on p. 92).
- [FK05] Xiaobing Feng and Ohannes A. Karakashian. “Two-level non-overlapping Schwarz preconditioners for a discontinuous Galerkin approximation of the biharmonic equation.” In: *J. Sci. Comput.* 22/23 (2005), pp. 289–314 (cit. on p. 92).
- [Fri58] Kurt O. Friedrichs. “Symmetric positive linear differential equations.” In: *Comm. Pure Appl. Math.* 11 (1958), pp. 333–418 (cit. on pp. 19, 38).
- [Fuc+09] Franz G. Fuchs, Kenneth H. Karlsen, Siddharta Mishra, and Nils H. Risebro. “Stable upwind schemes for the magnetic induction equation.” In: *M2AN Math. Model. Numer. Anal.* 43.5 (2009), pp. 825–852 (cit. on p. 154).
- [Fuc+11] Franz G. Fuchs, Andrew D. McMurry, Siddhartha Mishra, Nils Henrik Risebro, and Knut Waagan. “Approximate Riemann solvers and robust high-order finite volume schemes for multi-dimensional ideal MHD equations.” In: *Commun. Comput. Phys.* 9.2 (2011), pp. 324–362 (cit. on pp. 9, 145, 146, 155, 159).
- [FMR09] Franz G. Fuchs, Siddhartha Mishra, and Nils Henrik Risebro. “Splitting based finite volume schemes for ideal MHD equations.” In: *J. Comput. Phys.* 228.3 (2009), pp. 641–660 (cit. on pp. x, xi, 11, 133, 138, 142, 153, 154).
- [Gli65] James Glimm. “Solutions in the large for nonlinear hyperbolic systems of equations.” In: *Comm. Pure Appl. Math.* 18 (1965), pp. 697–715 (cit. on p. 6).
- [GR96] Edwige Godlewski and Pierre-Arnaud Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Vol. 118. Applied Mathematical Sciences. Springer-Verlag, New York, 1996, pp. viii+509 (cit. on p. 8).
- [God59] Sergei K. Godunov. “A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics.” In: *Mat. Sb. (N.S.)* 47(89).3 (1959), 271–306, (in Russian) (cit. on pp. 82, 136).
- [GP04] Hans Goedbloed and Stefaan Poedts. *Principles of Magnetohydrodynamics*. Cambridge University Press, 2004 (cit. on pp. 1, 3, 6).
- [GV96] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Third. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 1996, pp. xxx+698 (cit. on p. 120).
- [GL85] Jonathan B. Goodman and Randall J. LeVeque. “On the accuracy of stable schemes for 2D scalar conservation laws.” In: *Math. Comp.* 45.171 (1985), pp. 15–21 (cit. on p. 82).
- [GK03] Jay Gopalakrishnan and Guido Kanschat. “A multilevel discontinuous Galerkin method.” In: *Numer. Math.* 95.3 (2003), pp. 527–550 (cit. on p. 93).
- [GH01] David Gottlieb and Jan S. Hesthaven. “Spectral methods for hyperbolic problems.” In: *J. Comput. Appl. Math.* 128.1-2 (2001), pp. 83–131 (cit. on p. 8).
- [GST01] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. “Strong stability-preserving high-order time discretization methods.” In: *SIAM Rev.* 43.1 (2001), 89–112 (electronic) (cit. on p. 137).
- [Gro19] Thomas H. Gronwall. “Note on the derivatives with respect to a parameter of the solutions of a system of differential equations.” In: *Ann. of Math. (2)* 20.4 (1919), pp. 292–296 (cit. on p. 54).
- [GK01] Paul W. Gross and Robert P. Kotiuga. “Data structures for geometric and topological aspects of finite element algorithms.” In: *Geometric Methods for Computational Electromagnetics (F. Teixeira. ed.)* 32 (2001), pp. 151–169 (cit. on p. 24).
- [GSS07] Marcus J. Grote, Anna Schneebeli, and Dominik Schötzau. “Interior penalty discontinuous Galerkin method for Maxwell’s equations: energy norm error estimates.” In: *J. Comput. Appl. Math.* 204.2 (2007), pp. 375–386 (cit. on pp. 32, 97).
- [GSS08] Marcus J. Grote, Anna Schneebeli, and Dominik Schötzau. “Interior penalty discontinuous Galerkin method for Maxwell’s equations: optimal  $L^2$ -norm error estimates.” In: *IMA J. Numer. Anal.* 28.3 (2008), pp. 440–468 (cit. on pp. 95, 97, 119).
- [GP08] Jean-Luc Guermond and Richard Pasquetti. “Entropy-based nonlinear viscosity for Fourier approximations of conservation laws.” In: *C. R. Math. Acad. Sci. Paris* 346.13-14 (2008), pp. 801–806 (cit. on pp. x, 69, 82, 84).
- [GPP11] Jean-Luc Guermond, Richard Pasquetti, and Bojan Popov. “Entropy viscosity method for nonlinear conservation laws.” In: *J. Comput. Phys.* 230.11 (2011), pp. 4248–4267 (cit. on pp. 82, 84, 85).
- [Gur04] Katharine F. Gurski. “An HLLC-type approximate Riemann solver for ideal magnetohydrodynamics.” In: *SIAM J. Sci. Comput.* 25.6 (2004), 2165–2187 (electronic) (cit. on pp. 141, 142).
- [Har+87] Ami Harten, Björn Engquist, Stanley Osher, and Sukumar R. Chakravarthy. “Uniformly high-order accurate essentially nonoscillatory schemes. III.” In: *J. Comput. Phys.* 71.2 (1987), pp. 231–303 (cit. on p. 145).
- [HLL83] Amiram Harten, Peter D. Lax, and Bram van Leer. “On upstream differencing and Godunov-type schemes for hyperbolic conservation laws.” In: *SIAM Rev.* 25.1 (1983), pp. 35–61 (cit. on pp. 138, 140).

## Bibliography

---

- [Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002, pp. xii+544 (cit. on pp. 24, 25).
- [HE89] John F. Hawley and Charles R. Evans. “Constraint preserving transport for magnetohydrodynamics.” In: *Frontiers in numerical relativity (Urbana-Champaign, IL, 1988)*. Cambridge Univ. Press, Cambridge, 1989, pp. 179–193 (cit. on p. 10).
- [Heu11] Holger Heumann. “Eulerian and semi-Lagrangian methods for advection-diffusion of differential forms.” PhD thesis. ETH Zürich, 2011, pp. xiv+144 (cit. on pp. x, 23, 41, 43, 51, 72).
- [HH08] Holger Heumann and Ralf Hiptmair. *Extrusion contraction upwind schemes for convection-diffusion problems*. Tech. rep. 2008-30. Switzerland: Seminar for Applied Mathematics, ETH Zürich, 2008 (cit. on pp. x, 69, 72).
- [HH13a] Holger Heumann and Ralf Hiptmair. “Convergence of lowest order semi-Lagrangian schemes.” In: *Found. Comput. Math.* 13.2 (2013), pp. 187–220 (cit. on pp. 21, 38).
- [HH13b] Holger Heumann and Ralf Hiptmair. “Stabilized Galerkin methods for magnetic advection.” In: *ESAIM Math. Model. Numer. Anal.* 47.6 (2013), pp. 1713–1732 (cit. on pp. 32, 38, 42).
- [HHP16] Holger Heumann, Ralf Hiptmair, and Cecilia Pagliantini. “Stabilized Galerkin for transient advection of differential forms.” In: *Discrete Contin. Dyn. Syst. Ser. S* 9.1 (2016), pp. 185–214 (cit. on pp. 19, 37).
- [Hip99a] Ralf Hiptmair. “Multigrid method for Maxwell’s equations.” In: *SIAM J. Numer. Anal.* 36.1 (1999), pp. 204–225 (cit. on pp. xi, 92).
- [Hip99b] Ralf Hiptmair. “Canonical construction of finite elements.” In: *Math. Comp.* 68.228 (1999), pp. 1325–1346 (cit. on p. 27).
- [Hip01] Ralf Hiptmair. “Discrete Hodge operators.” In: *Numer. Math.* 90.2 (2001), pp. 265–289 (cit. on p. 26).
- [Hip02] Ralf Hiptmair. “Finite elements in computational electromagnetism.” In: *Acta Numer.* 11 (2002), pp. 237–339 (cit. on pp. 15–18, 21, 24, 26, 30, 34).
- [Hip03] Ralf Hiptmair. “Analysis of multilevel methods for eddy current problems.” In: *Math. Comp.* 72.243 (2003), pp. 1281–1303 (cit. on pp. 92, 102).
- [HT00] Ralf Hiptmair and Andrea Toselli. “Overlapping and multilevel Schwarz methods for vector valued elliptic problems in three dimensions.” In: *Parallel solution of partial differential equations (Minneapolis, MN, 1997)*. Vol. 120. IMA Vol. Math. Appl. Springer, New York, 2000, pp. 181–208 (cit. on pp. xi, 92, 102, 118).
- [HWZ06] Ralf Hiptmair, Gisela Widmer, and Jun Zou. “Auxiliary space preconditioning in  $\mathbf{H}_0(\mathbf{curl}; \Omega)$ .” In: *Numer. Math.* 103.3 (2006), pp. 435–459 (cit. on p. 100).
- [HX07] Ralf Hiptmair and Jinchao Xu. “Nodal auxiliary space preconditioning in  $\mathbf{H}(\mathbf{curl})$  and  $\mathbf{H}(\mathbf{div})$  spaces.” In: *SIAM J. Numer. Anal.* 45.6 (2007), 2483–2509 (electronic) (cit. on pp. 92, 98–100).
- [HZ09] Ralf Hiptmair and Wei-Ying Zheng. “Local multigrid in  $\mathbf{H}(\mathbf{curl})$ .” In: *J. Comput. Math.* 27.5 (2009), pp. 573–603 (cit. on p. 92).
- [Hir03] Anil N. Hirani. “Discrete exterior calculus.” PhD thesis. California Institute of Technology, 2003, p. 103 (cit. on p. 26).
- [Hou+05] Paul Houston, Ilaria Perugia, Anna Schneebeli, and Dominik Schötzau. “Interior penalty method for the indefinite time-harmonic Maxwell equations.” In: *Numer. Math.* 100.3 (2005), pp. 485–518 (cit. on pp. 32, 46, 93, 94, 97, 98, 106, 107).
- [HPS05] Paul Houston, Ilaria Perugia, and Dominik Schötzau. “Energy norm a posteriori error estimation for mixed discontinuous Galerkin approximations of the Maxwell operator.” In: *Comput. Methods Appl. Mech. Engrg.* 194.2–5 (2005), pp. 499–510 (cit. on p. 106).
- [HPS07] Paul Houston, Ilaria Perugia, and Dominik Schötzau. “An a posteriori error indicator for discontinuous Galerkin discretizations of  $\mathbf{H}(\mathbf{curl})$ -elliptic partial differential equations.” In: *IMA J. Numer. Anal.* 27.1 (2007), pp. 122–150 (cit. on p. 106).
- [HSS00] Paul Houston, Christoph Schwab, and Endre Süli. “Stabilized  $hp$ -finite element methods for first-order hyperbolic problems.” In: *SIAM J. Numer. Anal.* 37.5 (2000), pp. 1618–1643 (cit. on pp. 38, 61).
- [HMX16] Kaibo Hu, Yicong Ma, and Jinchao Xu. “Stable finite element methods preserving  $\nabla \cdot \mathbf{B} = 0$  exactly for MHD models.” In: *Numerische Mathematik* (2016), pp. 1–26 (cit. on p. 8).
- [HSZ08] Qiya Hu, Shi Shu, and Jun Zou. “A mortar edge element method with nearly optimal convergence for three-dimensional Maxwell’s equations.” In: *Math. Comp.* 77.263 (2008), pp. 1333–1353 (cit. on p. 92).
- [HSZ13] Qiya Hu, Shi Shu, and Jun Zou. “A discrete weighted Helmholtz decomposition and its application.” In: *Numer. Math.* 125.1 (2013), pp. 153–189 (cit. on p. 92).
- [HZ03] Qiya Hu and Jun Zou. “A nonoverlapping domain decomposition method for Maxwell’s equations in three dimensions.” In: *SIAM J. Numer. Anal.* 41.5 (2003), pp. 1682–1708 (cit. on p. 92).
- [HZ04] Qiya Hu and Jun Zou. “Substructuring preconditioners for saddle-point problems arising from Maxwell’s equations in three dimensions.” In: *Math. Comp.* 73.245 (2004), pp. 35–61 (cit. on p. 92).
- [HB79] Thomas J. R. Hughes and Alexander Brooks. “A multidimensional upwind scheme with no crosswind diffusion.” In: *Finite element methods for convection dominated flows (Papers, Winter Ann. Meeting Amer. Soc. Mech. Engrs., New York, 1979)*. Vol. 34. AMD. Amer. Soc. Mech. Engrs. (ASME), New York, 1979, pp. 19–35 (cit. on p. 37).

- [HFH89] Thomas J. R. Hughes, Leopoldo P. Franca, and Gregory M. Hulbert. “A new finite element formulation for computational fluid dynamics. VIII. The Galerkin/least-squares method for advective-diffusive equations.” In: *Comput. Methods Appl. Mech. Engrg.* 73.2 (1989), pp. 173–189 (cit. on p. 37).
- [HMM86] Thomas J. R. Hughes, Michel Mallet, and Akira Mizukami. “A new finite element formulation for computational fluid dynamics. II. Beyond SUPG.” In: *Comput. Methods Appl. Mech. Engrg.* 54.3 (1986), pp. 341–355 (cit. on p. 82).
- [HS99] James M. Hyman and Mikhail Shashkov. “Mimetic discretizations for Maxwell’s equations.” In: *J. Comput. Phys.* 151.2 (1999), pp. 881–909 (cit. on p. 23).
- [Ike83] Tsutomu Ikeda. *Maximum principle in finite element models for convection-diffusion phenomena*. Vol. 4. Lecture Notes in Numerical and Applied Analysis. Kinokuniya Book Store Co., Ltd., Tokyo, 1983, pp. ix+159 (cit. on p. 82).
- [JT64] Alan Jeffrey and Tosiya Taniuti. *Non-linear wave propagation. With applications to physics and magnetohydrodynamics*. New York: Academic Press, 1964, pp. x + 369 (cit. on pp. 6, 7).
- [Käp+11] Roger Käppeli, Stuart C. Whitehouse, Simon Scheidegger, Ue-Li Pen, and Matthias Liebendörfer. “FISH: a three-dimensional parallel magnetohydrodynamics code for astrophysical applications.” In: *The Astrophysical Journal Supplement Series* 195.2 (2011), p. 20 (cit. on p. 154).
- [KP03] Ohannes A. Karakashian and Frédéric Pascal. “A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems.” In: *SIAM J. Numer. Anal.* 41.6 (2003), 2374–2399 (electronic) (cit. on pp. 46, 108).
- [KMR09] Kenneth Hvistendahl Karlsen, Siddhartha Mishra, and Nils Henrik Risebro. “Convergence of finite volume schemes for triangular systems of conservation laws.” In: *Numer. Math.* 111.4 (2009), pp. 559–589 (cit. on p. 10).
- [Kat76] Tosio Kato. “Linear and quasi-linear equations of evolution of hyperbolic type.” In: *Hyperbolicity*. Ed. by Giuseppe Da Prato and Giuseppe Geymonat. Vol. 72. C.I.M.E. Summer Schools. Springer Berlin Heidelberg, 1976, pp. 125–191 (cit. on p. 38).
- [KBC89] Charles F. Kennel, Roger D. Blandford, and Paolo Coppi. “MHD intermediate shock discontinuities. Part 1. Rankine-Hugoniot conditions.” In: *Journal of Plasma Physics* 42 (1989), pp. 299–319 (cit. on p. 7).
- [KXZ15] Tzanio V. Kolev, Jinchao Xu, and Yunrong Zhu. “Multilevel preconditioners for reaction-diffusion problems with discontinuous coefficients.” In: *Journal of Scientific Computing* (2015), pp. 1–27 (cit. on p. 92).
- [Kot84] P. Robert Kotiuga. *Hodge decompositions and computational electromagnetics*. PhD in Electrical Engineering. 1984 (cit. on p. 23).
- [KT08] Johannes K. Kraus and Satyendra K. Tomar. “Multilevel preconditioning of two-dimensional elliptic problems discretized by a class of discontinuous Galerkin methods.” In: *SIAM J. Sci. Comput.* 30.2 (2008), pp. 684–706 (cit. on p. 93).
- [KM05] Dmitri Kuzmin and Matthias Möller. “Algebraic flux correction. I. Scalar conservation laws.” In: *Flux-corrected transport*. Sci. Comput. Springer, Berlin, 2005, pp. 155–206 (cit. on p. 82).
- [Lan99] Serge Lang. *Fundamentals of differential geometry*. Vol. 191. Graduate Texts in Mathematics. New York: Springer-Verlag, 1999, pp. xviii+535 (cit. on pp. 16, 71).
- [Lan02] Serge Lang. *Introduction to differentiable manifolds*. Second. Universitext. Springer-Verlag, New York, 2002, pp. xii+250 (cit. on p. 19).
- [Lax57] Peter D. Lax. “Hyperbolic systems of conservation laws. II.” In: *Comm. Pure Appl. Math.* 10 (1957), pp. 537–566 (cit. on p. 6).
- [Lee13] John M. Lee. *Introduction to smooth manifolds*. Second. Vol. 218. Graduate Texts in Mathematics. New York: Springer, 2013, pp. xvi+708 (cit. on p. 19).
- [LeV97] Randall J. LeVeque. “Wave propagation algorithms for multidimensional hyperbolic systems.” In: *J. Comput. Phys.* 131 (1997), pp. 327–353 (cit. on pp. 10, 83).
- [LeV02] Randall J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002, pp. xx+558 (cit. on pp. ix, 8).
- [LT98] Doron Levy and Eitan Tadmor. “From semidiscrete to fully discrete: stability of Runge-Kutta schemes by the energy method.” In: *SIAM Rev.* 40.1 (1998), 40–73 (electronic) (cit. on p. 52).
- [Li05] Shengtai Li. “An HLLC Riemann solver for magneto-hydrodynamics.” In: *J. Comput. Phys.* 203.1 (2005), pp. 344–357 (cit. on p. 142).
- [Lin98] Timur J. Linde. “A three-dimensional adaptive multifluid MHD model of the heliosphere.” PhD thesis. The university of Michigan, Ann Arbor, MI, 1998 (cit. on p. 141).
- [LMW12] Anders Logg, Kent-Andre Mardal, and Garth N. Wells. *Automated solution of differential equations by the finite element method. The FEniCS book*. Springer, 2012 (cit. on p. 59).
- [LD04] Pasquale Londrillo and Luca Del Zanna. “On the divergence-free condition in Godunov-type schemes for ideal magnetohydrodynamics: the upwind constrained transport method.” In: *J. Comput. Phys.* 195.1 (2004), pp. 17–48 (cit. on p. 10).

- [MT12] Siddhartha Mishra and Eitan Tadmor. “Constraint preserving schemes using potential-based fluxes. III. Genuinely multi-dimensional schemes for MHD equations.” In: *ESAIM Math. Model. Numer. Anal.* 46.3 (2012), pp. 661–680 (cit. on p. 10).
- [MK05] Takahiro Miyoshi and Kanya Kusano. “A multi-state HLL approximate Riemann solver for ideal magnetohydrodynamics.” In: *J. Comput. Phys.* 208.1 (2005), pp. 315–344 (cit. on pp. 142, 153).
- [MH85] Akira Mizukami and Thomas J. R. Hughes. “A Petrov-Galerkin finite element method for convection-dominated flows: an accurate upwinding technique for satisfying the maximum principle.” In: *Comput. Methods Appl. Mech. Engrg.* 50.2 (1985), pp. 181–193 (cit. on p. 82).
- [Mon03] Peter Monk. *Finite element methods for Maxwell’s equations*. Numerical Mathematics and Scientific Computation. New York: Oxford University Press, 2003, pp. xiv+450 (cit. on pp. 32, 34).
- [Mul+11] Patrick Mullen, Alexander McKenzie, Dmitry Pavlov, Luke Durant, Yiying Tong, Eva Kanso, Jerrold E. Marsden, and Mathieu Desbrun. “Discrete Lie advection of differential forms.” In: *Found. Comput. Math.* 11.2 (2011), pp. 131–149 (cit. on p. 72).
- [Mun+00] Claus-Dieter Munz, Pascal Omnes, Rudolf Schneider, Eric Sonnendrücker, and Ursula Voß. “Divergence correction techniques for Maxwell solvers based on a hyperbolic model.” In: *J. Comput. Phys.* 161.2 (2000), pp. 484–511 (cit. on p. 10).
- [Néd80] Jean-Claude Nédélec. “Mixed finite elements in  $\mathbf{R}^3$ .” In: *Numer. Math.* 35.3 (1980), pp. 315–341 (cit. on p. 33).
- [Néd86] Jean-Claude Nédélec. “A new family of mixed finite elements in  $\mathbf{R}^3$ .” In: *Numer. Math.* 50.1 (1986), pp. 57–81 (cit. on pp. 33, 117).
- [Nep91] Sergey V. Nepomnyaschikh. “Mesh theorems on traces, normalizations of function traces and their inversion.” In: *Soviet J. Numer. Anal. Math. Modelling* 6.3 (1991), pp. 223–242 (cit. on pp. xi, 98, 99).
- [OT79] Steven A. Orszag and Cha-Mei Tang. “Small-scale structure of two-dimensional magnetohydrodynamic turbulence.” In: *J. Fluid Mech.* 90.1 (1979), pp. 129–143 (cit. on pp. 143, 148, 153).
- [Osw96] Peter Oswald. “Preconditioners for nonconforming discretizations.” In: *Math. Comp.* 65.215 (1996), pp. 923–941 (cit. on p. 98).
- [Par91] George K. Parks. *Physics of space plasmas: an introduction*. Redwood City, Calif.: Addison-Wesley, 1991 (cit. on pp. ix, 1).
- [Paz83] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*. Vol. 44. Applied Mathematical Sciences. Springer-Verlag, New York, 1983, pp. viii+279 (cit. on p. 38).
- [PS03] Ilaria Perugia and Dominik Schötzau. “The  $hp$ -local discontinuous Galerkin method for low-frequency time-harmonic Maxwell equations.” In: *Math. Comp.* 72.243 (2003), pp. 1179–1214 (cit. on pp. 93, 95).
- [Pet91] Todd E. Peterson. “A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation.” In: *SIAM J. Numer. Anal.* 28.1 (1991), pp. 133–140 (cit. on p. 59).
- [Pow94] Kenneth G. Powell. *An approximate Riemann solver for magnetohydrodynamics (that works in more than one dimension)*. Tech. rep. 94-24. Langley, VA, 1994 (cit. on p. 9).
- [RT01] Francesca Rapetti and Andrea Toselli. “A FETI preconditioner for two-dimensional edge element approximations of Maxwell’s equations on nonmatching grids.” In: *SIAM J. Sci. Comput.* 23.1 (2001), pp. 92–108 (cit. on p. 92).
- [RT77] Pierre-Arnaud Raviart and Jean-Marie Thomas. “A mixed finite element method for 2nd order elliptic problems.” In: *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*. Springer, Berlin, 1977, 292–315. Lecture Notes in Math., Vol. 606 (cit. on p. 33).
- [Roe81] Philip L. Roe. “Approximate Riemann solvers, parameter vectors, and difference schemes.” In: *J. Comput. Phys.* 43.2 (1981), pp. 357–372 (cit. on p. 138).
- [RST08] Hans-Görg Roos, Martin Stynes, and Lutz Tobiska. *Robust numerical methods for singularly perturbed differential equations*. Second. Vol. 24. Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2008, pp. xiv+604 (cit. on p. 37).
- [Ros06] James A. Rossmanith. “An unstaggered, high-resolution constrained transport method for magnetohydrodynamic flows.” In: *SIAM J. Sci. Comput.* 28.5 (2006), pp. 1766–1797 (cit. on pp. 10, 83, 86).
- [Sch01] Joachim Schöberl. *Commuting quasi-interpolation operators for mixed finite elements*. Tech. rep. ISC-01-10-MATH. Texas A&M University, College Station, 2001 (cit. on p. 35).
- [SL13] Joachim Schöberl and Christoph Lehrenfeld. “Domain decomposition preconditioning for high order hybrid discontinuous Galerkin methods on tetrahedral meshes.” In: *Advanced finite element methods and applications*. Vol. 66. Lect. Notes Appl. Comput. Mech. Springer, Heidelberg, 2013, pp. 27–56 (cit. on p. 93).
- [Sch04] Dominik Schötzau. “Mixed finite element methods for stationary incompressible magneto-hydrodynamics.” In: *Numer. Math.* 96.4 (2004), pp. 771–800 (cit. on p. 8).
- [Sch95] Günter Schwarz. *Hodge decomposition – a method for solving boundary value problems*. Vol. 1607. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1995, pp. viii+155 (cit. on pp. 15, 16, 19).
- [SZ90] L. Ridgway Scott and Shangyou Zhang. “Finite element interpolation of nonsmooth functions satisfying boundary conditions.” In: *Math. Comp.* 54.190 (1990), pp. 483–493 (cit. on p. 35).

- [Shu01] Chi-Wang Shu. "Different formulations of the discontinuous Galerkin method for the viscous terms." In: *Advances in Scientific Computing*. Shi, Z.-C., Mu, M., Xue, W., and Zou, J. (eds.) 47 (06 2001), pp. 144–155 (cit. on p. 61).
- [SO88] Chi-Wang Shu and Stanley Osher. "Efficient implementation of essentially nonoscillatory shock-capturing schemes." In: *J. Comput. Phys.* 77.2 (1988), pp. 439–471 (cit. on p. 137).
- [SO89] Chi-Wang Shu and Stanley Osher. "Efficient implementation of essentially nonoscillatory shock-capturing schemes. II." In: *J. Comput. Phys.* 83.1 (1989), pp. 32–78 (cit. on p. 145).
- [Tab77] Masahisa Tabata. "A finite element approximation corresponding to the upwind finite differencing." In: *Mem. Numer. Math.* 4 (1977), pp. 47–63 (cit. on pp. x, 72, 75, 82).
- [TSS94] Eleuterio F. Toro, Michael Spruce, and William Speares. "Restoration of the contact surface in the HLL-Riemann solver." In: *Shock Waves* 4.1 (1994), pp. 25–34 (cit. on p. 140).
- [Tor02] Manuel Torrilhon. *Exact solver and uniqueness conditions for Riemann problems of ideal magnetohydrodynamics*. Tech. rep. 2002-06. Switzerland: Seminar for Applied Mathematics, ETH Zürich, 2002 (cit. on pp. 6, 138).
- [Tos00a] Andrea Toselli. "Neumann-Neumann methods for vector field problems." In: *Electron. Trans. Numer. Anal.* 11 (2000), pp. 1–24 (cit. on p. 92).
- [Tos00b] Andrea Toselli. "Overlapping Schwarz methods for Maxwell's equations in three dimensions." In: *Numer. Math.* 86.4 (2000), pp. 733–752 (cit. on p. 92).
- [Tos06] Andrea Toselli. "Dual-primal FETI algorithms for edge finite-element approximations in 3D." In: *IMA J. Numer. Anal.* 26.1 (2006), pp. 96–130 (cit. on p. 92).
- [TV05] Andrea Toselli and Xavier Vasseur. "Dual-primal FETI algorithms for edge element approximations: two-dimensional  $H$  and  $P$  finite elements on shape-regular meshes." In: *SIAM J. Numer. Anal.* 42.6 (2005), pp. 2590–2611 (cit. on p. 92).
- [TW05] Andrea Toselli and Olof B. Widlund. *Domain decomposition methods—algorithms and theory*. Vol. 34. Springer Series in Computational Mathematics. Berlin: Springer-Verlag, 2005, pp. xvi+450 (cit. on pp. 118, 119).
- [TWW01] Andrea Toselli, Olof B. Widlund, and Barbara I. Wohlmuth. "An iterative substructuring method for Maxwell's equations in two dimensions." In: *Math. Comp.* 70.235 (2001), pp. 935–949 (cit. on p. 92).
- [Tót00] Gábor Tóth. "The  $\nabla \cdot B = 0$  constraint in shock-capturing magnetohydrodynamics codes." In: *J. Comput. Phys.* 161.2 (2000), pp. 605–652 (cit. on pp. 9, 10, 155, 162).
- [Waa09] Knut Waagan. "A positive MUSCL-Hancock scheme for ideal magnetohydrodynamics." In: *J. Comput. Phys.* 228.23 (2009), pp. 8609–8626 (cit. on p. 146).
- [Wal05] Noel J. Walkington. "Convergence of the discontinuous Galerkin method for discontinuous solutions." In: *SIAM J. Numer. Anal.* 42.5 (2005), 1801–1817 (electronic) (cit. on p. 38).
- [Whe78] Mary F. Wheeler. "An elliptic collocation-finite element method with interior penalties." In: *SIAM J. Numer. Anal.* 15.1 (1978), pp. 152–161 (cit. on p. 94).
- [Whi57] Hassler Whitney. *Geometric integration theory*. Princeton University Press, Princeton, N. J., 1957, pp. xv+387 (cit. on p. 23).
- [Xu92] Jinchao Xu. "Iterative methods by space decomposition and subspace correction." In: *SIAM Rev.* 34.4 (1992), pp. 581–613 (cit. on p. 101).
- [Xu96] Jinchao Xu. "The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids." In: *Computing* 56.3 (1996), pp. 215–235 (cit. on pp. xi, 98, 99).
- [Yee66] Kane Yee. "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media." In: *IEEE Transactions on Antennas and Propagation* 14.3 (1966), pp. 302–307 (cit. on pp. 10, 26).
- [ZS04] Qiang Zhang and Chi-Wang Shu. "Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws." In: *SIAM J. Numer. Anal.* 42.2 (2004), 641–666 (electronic) (cit. on p. 137).
- [ZS06] Qiang Zhang and Chi-Wang Shu. "Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin method for symmetrizable systems of conservation laws." In: *SIAM J. Numer. Anal.* 44.4 (2006), 1703–1720 (electronic) (cit. on p. 137).
- [Zik08] Ludmil T. Zikatanov. "Two-sided bounds on the convergence rate of two-level methods." In: *Numer. Linear Algebra Appl.* 15.5 (2008), pp. 439–454 (cit. on pp. 101, 102, 119).
- [Zin+13] Valentin Zingan, Jean-Luc Guermond, Jim Morel, and Bojan Popov. "Implementation of the entropy viscosity method with the discontinuous Galerkin method." In: *Comput. Methods Appl. Mech. Engrg.* 253 (2013), pp. 479–490 (cit. on pp. 87, 89).