Space-time Hybridizable Discontinuous Galerkin Method for the Advection-Diffusion Problem

by

Yuan Wang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Yves Bourgault
Professor, Dept. of Mathematics and Statistics
University of Ottawa

Supervisor: Sander Rhebergen
Associate Professor, Dept. of Applied Mathematics
University of Waterloo

Internal Member: David Del Rey Fernández
Assistant Professor, Dept. of Applied Mathematics
University of Waterloo

Giang Tran
Assistant Professor, Dept. of Applied Mathematics
University of Waterloo

Internal-External Member: Nasser Mohieddin Abukhdeir
Associate Professor, Dept. of Chemical Engineering
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Yuan Wang was the sole author of this thesis, which was written under the supervision of Prof. Sander Rhebergen.

Chapter 4: Research presented in Chapter 4 and part of Chapter 3 (some results in sections 3.1 to 3.3) has been submitted for publication [104] (`https://arxiv.org/abs/2308.12130`).

Chapter 5: Research presented in Chapter 5 and part of Chapter 3 (some results in sections 3.1 to 3.3 and the entirety of sections 3.4 to 3.6) has been submitted for publication [105] (`https://arxiv.org/abs/2404.04130`).

Abstract

In this thesis, we analyze a space-time hybridizable discontinuous Galerkin (HDG) method for the time-dependent advection-dominated advection-diffusion problem. It is well-known that solutions to these problems may admit sharp boundary and interior layers and that many numerical methods are prone to non-physical oscillations when resolving these solutions. This challenge has prompted the design of many new numerical methods and stabilization mechanisms. Among others, HDG methods prove to be capable of resolving the sharp layers in a robust manner. The design principles of HDG methods consist of discontinuous Galerkin (DG) methods and their strong stability properties, as well as hybridization to reduce the computational cost of the numerical method.

The analysis in this work focuses on a space-time formulation of the time-dependent advection-diffusion problem and an HDG discretization in both space and time. This provides a straightforward approach to discretize the problem on a time-dependent domain, with arbitrary higher-order spatial and temporal accuracy. We present an a priori error analysis that provides Péclet-robust error estimates that are also valid on moving meshes. A key intermediate step towards our error estimates is a Péclet-robust inf-sup stability condition.

The second contribution of this thesis is an a posteriori error analysis of the space-time HDG method for the time-dependent advection-dominated advection-diffusion problem on fixed domains. This is motivated by the efficiency of combining a posteriori error estimators with adaptive mesh refinement (AMR) to locally refine or coarsen a mesh in the presence of sharp layers. When the solution admits sharp layers, AMR may still lead to optimal rates of convergence in terms of the number of degrees-of-freedom, unlike uniform mesh refinement.

In this thesis, we present an a posteriori error estimator for the space-time HDG method with respect to a locally computable norm. We prove its reliability and local efficiency. The proof of reliability is based on a combination of a Péclet-robust coercivity type result and a saturation assumption. In addition, efficiency, which is local both in space and time, is shown using bubble function techniques. The error estimator in this thesis is fully local, hence it is an estimator for local space and time adaptivity in the AMR procedure.

Finally, numerical simulations are presented to demonstrate and verify the theory. Both uniform and adaptive refinement strategies are performed on problems which admit boundary and interior layers.

Acknowledgements

As much as the common wisdom would suggest that PhD is a 4-year personal and inward adventure - and there is a lot of truth in that - this trip would not have come nearly as far as it has without the influence, inspiration, friendship and mentorship I received for which I owe a great amount of gratitude to many.

My first and foremost thank is to my supervisor, Prof. Sander Rhebergen. From Sander I learned, among other things, work ethic and professionalism. Much of my growth as an academic is thanks to his invaluable influence and his commitment to my studies and my work.

Many thanks to Keegan Kirk, for his mentorship. To Thanin Quartz (Juju), for the time we had enduring the Pandemic together that made it much less of an endurance. To Paulo Zúñiga, for his Stoicism that had and still is having a great impact on me. To Giselle Sosa Jones, for the shared journey in our faith. To Somayeh Fouladi, for her generous friendship. To Hetian Liu, for being the brother us one-child generation never had. To Weiqi Wang, for many movie nights and for the camaraderie all these years.

I wish to extend my heartfelt appreciation to my defense committee: Prof. Yves Bourgault, Prof. Nasser Mohieddin Abukhdeir, Dr. David Del Rey Fernández and Dr. Giang Tran. Special thanks to David for his constant and generous support. I'm also truly appreciative of the great company within the Applied Math Department: Abdullah Sivas, Yunhui He, Esteban Henríquez, Elizabeth Yackoboski, Amirhossein Dehghanizadeh, Ala' Alalabi and Milad Moshayedi.

Special acknowledgement goes to Luke Smith, from whom I learned Linux; and to Prof. Wolfgang Bangerth and the deal.II community, from whom I learned a great deal of C++, finite element implementation and parallel computing.

And to my expat family: Bo Na, Yipin Lu, Xuanrui Li, Yun Su. To my family, from whom I've been regrettably physically separated during the entirety of my PhD.

Lastly, to Christ The Saviour Antiochian Orthodox Church fellowship and Fr. Christopher Rigden-Briscall for the guidance.

Dedication

To my parents Jierong Wang and Zhongxia Pan.

Table of Contents

List of Figures

List of Symbols

T The final time 1

$\overline{\nabla}$ Spatial gradient: $\overline{\nabla} = (\partial_{x_1}, \partial_{x_2}, \dots, \partial_{x_d})$ 1

$\bar{\beta}$ A given divergence-free advective field 1

ε A positive constant diffusion coefficient 1

$\Omega(t)$ A time-dependent polygonal ($d = 2$) or polyhedral ($d = 3$) domain in $\mathbb{R}^d$. When the domain is fixed, we denote $\Omega(t) \equiv \Omega$. 1

$\mathcal{E}$ $(d+1)$-dimensional polyhedral space-time domain 1

n The outward space-time normal vector to $\partial\mathcal{E}$: $n := (n_t, \bar{n})$, where n_t and $\bar{n}$ are temporal and spatial components of the space-time normal vector, respectively 17

β Space-time advective field: $(1, \bar{\beta})$ 17

∇ Space-time gradient: $\nabla := (\partial_t, \overline{\nabla})$ 17

$\partial\mathcal{E}$ The boundary of the space-time domain $\mathcal{E}$ which has a nonoverlapping partition $\partial\mathcal{E} = \partial\mathcal{E}_D \cup \partial\mathcal{E}_N$ where $\partial\mathcal{E}_D$ and $\partial\mathcal{E}_N$ denote Dirichlet and Neumann boundaries respectively 17

ζ^- Indicator function for the inflow part of the boundary of the space-time domain $\mathcal{E}$ 17

$\mathcal{K}$ A space-time element 17

K A spatial element 17

Υ The transformation that describes the deformation of the domain 17

$\mathcal{T}_h$ The space-time mesh 18

$\mathcal{R}_{\mathcal{K}}$ Part of the boundary of the element $\mathcal{K}$ with $t \in (t_*, t^*)$ on which $\bar{n}$ is the zero vector. Alternatively, $\mathcal{R}_{\mathcal{K}} := K_* \cup K^*$. 18

K_* Given a space-time element $\mathcal{K}$ with $t \in (t_*, t^*)$, the facet at time $t = t_*$ 18

K^* Given a space-time element $\mathcal{K}$ with $t \in (t_*, t^*)$, the facet at time $t = t^*$ 18

$\mathcal{Q}_{\mathcal{K}}$ Part of the boundary of the element $\mathcal{K}$ on which $\bar{n} \neq 0$. Furthermore, $\mathcal{Q}_{\mathcal{K}}$ and $\mathcal{R}_{\mathcal{K}}$ form a nonoverlapping partition of $\partial\mathcal{K}$. 18

$E_{\mathcal{K}}$ A $(d-1)$-dimensional edge of K^* and K_* 18

$\mathcal{F}_h$ The set of all facets of $\mathcal{T}_h$ 18

$\mathcal{F}_h^i$ The set of all interior facets in $\mathcal{F}_h$ 18

$\mathcal{F}_h^b$ The set of all boundary facets in $\mathcal{F}_h$ 18

$\mathcal{F}_{\mathcal{Q},h}$ The set of all $\mathcal{Q}$-facets in $\mathcal{F}_h$ 18

$\mathcal{F}_{\mathcal{R},h}$ The set of all $\mathcal{R}$-facets in $\mathcal{F}_h$ 18

Γ The union of all facets in $\mathcal{F}_h$ 18

$\partial\mathcal{T}_h$ The set of element boundaries 18

$\mathcal{Q}_h$ The set that consists of parts of an element boundary on which $\bar{n} \neq 0$ 18

$\mathcal{R}_h$ The set that consists of parts of an element boundary on which $\bar{n} = 0$ 18

$\partial\mathcal{T}_h^i$ The set of element boundaries excluding the part of the element boundary that lies on $\partial\mathcal{E}$ 18

$\omega_{\mathcal{K}}$ The union of elements $\mathcal{K}'$ such that $\partial\mathcal{K} \cap \partial\mathcal{K}' \neq \emptyset$ 18

$\sigma_{\mathcal{K}}$ The union of elements that share at least one vertex with $\mathcal{K}$ 18

ω_F The set of elements that contain a facet F' such that $F \cap F'$ is itself a facet 18

$\Phi_{\mathcal{K}}$ The mapping between the fixed reference element $\widehat{\mathcal{K}} = (-1,1)^{d+1}$ and space-time element $\mathcal{K} \in \mathcal{T}_h$ 18

$G_{\mathcal{K}}$ The affine part of $\Phi_{\mathcal{K}}$ that sets the size of $\mathcal{K}$ 18

h_K The spatial size of the element $\mathcal{K}$ 18

$\delta t_{\mathcal{K}}$ The time-step of the element $\mathcal{K}$ 18

$\phi_{\mathcal{K}}$ The diffeomorphism part of $\Phi_{\mathcal{K}}$ that sets the shape of $\mathcal{K}$ 18

$J_{\phi_{\mathcal{K}}}$ The Jacobian of the diffeomorphism $\phi_{\mathcal{K}}$ 20

$F_{\mathcal{Q}}^{j}$ A $\mathcal{Q}$-face where $\widetilde{x}_j$ is fixed in its affine domain 20

$J_{\phi_{\mathcal{K}}}^{j}$ A matrix obtained by removing the j^{th} column vector from $J_{\phi_{\mathcal{K}}}$ 20

$\Delta t_{\mathcal{K}}$ The time-step of the space-time slab that $\mathcal{K}$ is in 20

$V_h^{(p_t,p_s)}$ The finite element space defined on the elements. The superscript (p_t, p_s) will be omitted. 21

$M_h^{(p_t,p_s)}$ The finite element space defined on the facets. The superscript (p_t, p_s) will be omitted. 21

$\boldsymbol{V}_h$ The finite element space for the HDG method $\boldsymbol{V}_h = V_h \times M_h$. 21

$[\boldsymbol{v}_h]$ HDG jump: for $\boldsymbol{v}_h = (v_h, \mu_h) \in \boldsymbol{V}_h$, $(v_h - \mu_h)$ 21

$[\![v_h]\!]$ DG jump: on a facet $F \in \mathcal{F}_h^i$, where $F \subset \partial\mathcal{K}_1 \cap \partial\mathcal{K}_2$, $(v_{h1}n_1 + v_{h2}n_2)$ 21

$\langle\!\langle \mu_h \rangle\!\rangle$ Edge jump: consider two elements $\mathcal{K}_1$ and $\mathcal{K}_2$ such that $K_1^* = K_{2,*}$. Denote the restriction of μ_h to $\mathcal{Q}_{\mathcal{K}_1}$ and $\mathcal{Q}_{\mathcal{K}_2}$ by μ_{h1} and μ_{h2}, respectively. The jump of μ_h across edges of K_1^* is defined by $\mu_{h1} - \mu_{h2}$ 21

$\boldsymbol{V}$ $:= V \times M$ where $V := \{v \in H^1(\mathcal{E}) \,|\, v|_{\partial\mathcal{E}_D} = 0\} \cap H^2(\mathcal{E})$ and M its trace space 22

$\boldsymbol{V}(h)$ $:= V(h) \times M(h)$ where $V(h) := V_h + V$ and $M(h) := M_h + M$ 22

Π_h The L^2-projection onto V_h 22

τ_ε $:= \Delta t_{\mathcal{K}} \tilde{\varepsilon}$ as the coefficient of the norm of the time derivative 22

$\tilde{\varepsilon}$ A parameter that depends on the size of the space-time element compared to the diffusion parameter ε 22

$\mathcal{T}_h^d$ Space-time elements in $\mathcal{T}_h$ such that $\delta t_{\mathcal{K}} \leq h_K \leq \varepsilon$ 22

$\mathcal{T}_h^x$ Space-time elements in $\mathcal{T}_h$ such that $\delta t_{\mathcal{K}} \leq \varepsilon < h_K$ 22

$\mathcal{T}_h^c$ Space-time elements in $\mathcal{T}_h$ such that $\varepsilon < \delta t_{\mathcal{K}} \leq h_K$ 22

β_s The stabilization parameter in the advective part of the HDG bilinear form 22

β_s The penalty parameter in the diffusive part of the HDG bilinear form 23

ζ^+ Indicator function for the outflow boundary 23

$\Pi_h^{\mathcal{F}}$ The L^2-projection onto M_h 29

$\lambda_{\mathcal{K}}$ Coefficient of the element residual $\|R_h^{\mathcal{K}}\|_{\mathcal{K}}$ 31

$\mathcal{I}_h^c$ An averaging operator (also known as the Oswald approximation operator) 32

$\check{\mathcal{Q}}_{\mathcal{K}}^i$ The union of $\mathcal{Q}$-facets in $\mathcal{F}_h^i$ that have a non-empty intersection with $\partial\mathcal{K}$ 32

$\check{\mathcal{R}}_{\mathcal{K}}^i$ The union of $\mathcal{R}$-facets in $\mathcal{F}_h^i$ that have a non-empty intersection with $\partial\mathcal{K}$ 32

$\mathcal{T}_{\mathfrak{h}}$ The subgrid obtained by halving the time-step of each element in $\mathcal{T}_h$ 35

$\mathcal{T}_{\mathcal{K}}$ $:= \{\mathring{\mathcal{K}}^*, \mathring{\mathcal{K}}_*\}$ where $\mathring{\mathcal{K}}^*$ and $\mathring{\mathcal{K}}_*$ denote the two resulting space-time elements from $\mathcal{K} \in \mathcal{T}_h$. Furthermore, every $\mathcal{Q}$-facet $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$ is divided into $F_{\mathcal{Q}}^*$ and $F_{\mathcal{Q},*}$. 35

$F_{\mathring{\mathcal{R}}}$ $:= \partial\mathring{\mathcal{K}}^* \cap \partial\mathring{\mathcal{K}}_*$ 35

$E_{\mathring{\mathcal{K}}}$ An edge of $F_{\mathring{\mathcal{R}}}$ 35

$\gamma_{\mathfrak{h}}$ The restriction operator in the subgrid projection estimates. As part of its definition, another restriction operator $\gamma_{\mathcal{F},\mathfrak{h}}$ is defined. 35

i_h The projection operator in the subgrid projection estimates. As part of its definition, projection operators $i_h^{\mathcal{K}}$ and $i_h^{\mathcal{F}}$ are defined. 35

c_T A constant independent of h_K, $\delta t_{\mathcal{K}}$, and ε, but linear in T 45

φ The weighting function that is essential in proving the Péclet-robust inf-sup conditions 46

$R_h^{\mathcal{K}}$ Element residual 74

R_h^N Facet residual for the Neumann condition 74

$\eta^{\mathcal{K}}$	Error estimator on an element $\mathcal{K}$	75
$\psi_{\mathcal{K}}$	Element bubble function	101
ψ_F	Facet bubble function	104

Chapter 1

Introduction

This thesis presents an a priori and an a posteriori error analysis of a space-time hybridizable discontinuous Galerkin (HDG) method for the time-dependent advection-diffusion problem. The a priori error analysis considers the problem on a time-dependent polygonal ($d = 2$) or polyhedral ($d = 3$) domain $\Omega(t) \subset \mathbb{R}^d$, that evolves continuously in the time interval $t \in [0, T]$. The problem is given by

$$\partial_t u + \overline{\nabla} \cdot (\bar{\beta} u) - \varepsilon \overline{\nabla}^2 u = f \quad \text{in } \Omega(t),\ 0 < t \leq T, \tag{1.1}$$

in which $\overline{\nabla} = (\partial_{x_1}, \partial_{x_2}, \dots, \partial_{x_d})$ denotes the spatial gradient, $\bar{\beta}$ is a given divergence-free advective field, $\varepsilon > 0$ is a constant diffusion coefficient, and f is a forcing term. The a posteriori error analysis considers eq. (1.1) on a fixed spatial domain $\Omega(t) \equiv \Omega$. We also introduce the $(d+1)$-dimensional polyhedral space-time domain as $\mathcal{E} := \{(t, x) : x \in \Omega(t), 0 < t < T\} \subset \mathbb{R}^{d+1}$. In section 2.1, we will recast eq. (1.1) to its space-time formulation.

We will assume that $\bar{\beta} \in [W^{1,\infty}(\mathcal{E})]^d$, $\|\bar{\beta}\|_{L^\infty(\mathcal{E})} \leq 1$ and, following [12], that $\|\bar{\beta}\|_{W^{1,\infty}(\mathcal{E})} \leq c \|\bar{\beta}\|_{L^\infty(\mathcal{E})} \leq c$. We further assume that the size of Ω is order 1, following [31, 95], and hence ε^{-1} is the Péclet number of eq. (1.1). The focus in this thesis is the advection-dominated regime ($\varepsilon \ll 1$).

1.1 Stabilization of the advection-dominated advection-diffusion problem

The time-dependent advection-dominated advection-diffusion equation eq. (1.1) arises in various application areas [92]. Examples include, but are not limited to, the linearized Navier–Stokes equations of fluid dynamics with large Reynolds number [60, 80], the simulation of oil extraction from underground reservoirs [49], convective heat transport problems with large Péclet numbers [65], and multiphase flows [58].

When advection dominates in the advection-diffusion equation, its solution is well-known to potentially admit sharp boundary and interior layers [43, 57]. Accurately capturing these solutions proves to be nontrivial, and it is well understood that when applied to such problems, standard finite element methods are prone to global nonphysical oscillations. In response to this difficulty, various stabilization strategies have been proposed over the recent decades [11, 68, 73, 91, 92]. A prominent example is the streamline upwind Petrov–Galerkin method (SUPG) [21, 24, 39, 72] which achieves robust solutions by introducing artificial diffusion in the streamline direction of the advective field. However, spurious oscillations in the narrow boundary/interior layer region have been observed in SUPG solutions [2]. To smear out these oscillations, nonlinear artificial crosswind diffusion terms are added and this is the design principle of the spurious oscillations at layers diminishing method (SOLD) [70, 69, 71]. An alternative approach is the continuous interior penalty method (CIP) which enhances stability by penalizing the jump of the streamline derivative on interior faces of the mesh [22, 26, 27, 42]. The stabilization term of CIP methods is symmetric, unlike SUPG methods. Another member of symmetric stabilization techniques is the local projection stabilization method (LPS). It was introduced under the framework of projection-based stabilizations and is capable of attaining SUPG-type stability [77, 78, 79]. Finally, we remark that nonlinear artificial crosswind diffusion terms have also been combined with CIP and LPS methods, see [17, 25].

The numerical methods discussed above can be considered as stabilized variants of continuous finite element methods (CG) which use piecewise polynomial approximations that are continuous across interior facets of the mesh. An alternative is to use a discontinuous Galerkin (DG) finite element method which uses a discontinuous piecewise polynomial approximation.

With the DG method the PDE is discretized locally on each element and adjacent local discrete systems are coupled through a numerical flux defined on the element boundary. See the left panel of fig. 1.1 for an illustration. Under the framework of numerical fluxes, many existing DG methods can be unified and categorized by the specific choice of the

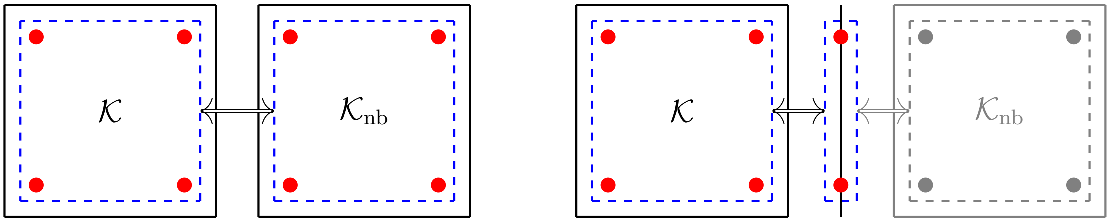


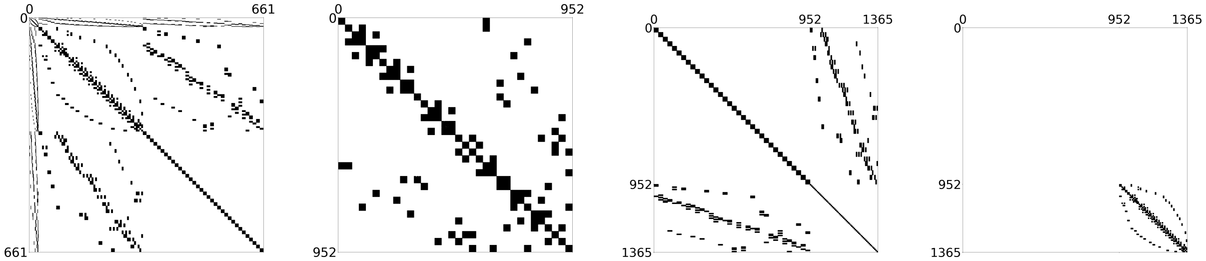
Figure 1.1: Red dots depict dofs. Left: DG dofs on the element $\mathcal{K}$ directly interact with dofs on a neighboring element $\mathcal{K}_{\mathrm{nb}}$. Right: HDG dofs on the element $\mathcal{K}$ only interact with dofs on facets.

numerical flux [10]. Suitably devised numerical fluxes render DG methods locally conservative, which is an important feature when local conservation of certain physical quantities is desired. Furthermore, the localized nature of DG methods opens up opportunities to highly parallelizable implementations, general meshes (meshes with hanging nodes, elements with nonstandard shapes, etc.), and *hp*-adaptivity (*h*-adaptivity: refining and coarsening local elements; *p*-adaptivity: the polynomial degree may vary between elements).

It is because of the aforementioned reasons that we consider a class of DG methods in this thesis. In the context of the advection-dominated advection-diffusion problem, DG methods have been extensively studied in [12, 37, 38, 46, 87]. Comparison studies of different stabilization techniques, including DG methods, on advection-dominated advection-diffusion problems can be found in [11, 20]. We also mention that published monographs on DG methods include [30, 41, 74, 86, 90].

1.2 Space-time hybridizable discontinuous Galerkin methods

DG methods are known to be expensive; on the same mesh, and when using polynomials of the same degree, DG methods have a larger number of degrees-of-freedom (dofs) compared to, for example, CG methods. Hybridizable DG (HDG) methods have been designed specifically to reduce the number of globally coupled degrees-of-freedom by using hybridization [36, 35]. This is achieved by introducing new dofs on the facets and designing the numerical flux such that element dofs communicate only with facet dofs, see the right panel of fig. 1.1 for an illustration. As such, element dofs are local dofs and can be cheaply eliminated through static condensation. This results in a reduced system of equations for only the globally coupled facet dofs. For higher-order approximations this reduced system



(a) CG (b) DG (c) HDG (d) HDG after static condensation

Figure 1.2: Using CG, DG, HDG, we solve a Poisson problem on a unit square mesh with 34 elements. We use polynomials of order 6 and plot the sparsity pattern. We observe that DG leads to a larger linear system (952) than CG (661). HDG further adds and additional (413) facet dofs. However, after static condensation, the facet dofs form the global system to be solved, which is smaller than CG's system (due to using high-order finite elements). Implementations for this example are done in the finite element library NGSolve [94].

of equations may be smaller than that of a CG discretization on the same mesh. See fig. 1.2 for the sparsity patterns of CG, DG, HDG before static condensation, and HDG after static condensation when applied to a Poisson problem.

Space-time discontinuous Galerkin methods: To discretize the time-dependent advection-diffusion equation, this thesis considers a space-time HDG method. In the space-time framework, a time-dependent PDE on a d-dimensional spatial domain is first converted to a 'stationary' PDE on a $(d+1)$-dimensional space-time domain. This space-time problem is then discretized simultaneously in space and time by a finite element method on the $(d+1)$-dimensional space-time mesh. At the expense of increased memory requirement, space-time methods excel at the automatic treatment of time-dependent domains, are arbitrarily higher-order accurate in both space and time, are straightforward to parallelize, and easily allows for local space and time adaptivity.

The space-time HDG method traces back to using DG time-stepping in space-time discretizations [66, 67]. On fixed domains, for example, DG time-stepping combined with SUPG was analyzed for the advection-diffusion equation in [64], while space-time DG, in which DG is applied both in space and time, was analyzed for a nonlinear advection-diffusion problem in [51]. The space-time DG method for the (linear) advection-diffusion problem on a time-dependent domain was analyzed in [98] by considering the space-time discretization on a space-time mesh consisting of anisotropic (in space and time) elements. This enabled them to obtain error estimates in terms of the spatial mesh size and the time-step. Their work was an extension of the analysis of DG methods for the stationary advection-diffusion problem on anisotropic spatial meshes [53] to space-time.

The extension of HDG to space-time, in which HDG is used to discretize a PDE in both space and time, was presented in [88, 89]. Application and analysis of HDG methods for the stationary advection-dominated advection-diffusion equation can be found in [34, 44, 85, 106]. The first a priori error analysis of a space-time HDG method for the advection-diffusion problem on time-dependent domains appeared in [76], which extended the space-time anisotropic framework used in [98] to HDG. However, despite the space-time HDG method in [76] performing well in practice for $\varepsilon \ll 1$, the well-posedness result proven in [76] does not hold in the advection dominated regime.

The a priori error analysis of this thesis. In this thesis we revisit the analysis in [76], however, with focus on the advection-dominated regime. We start by identifying the standard coercivity argument as the main source of the error estimate being nonrobust with respect to the Péclet number. Specifically, coercivity is a special stability bound derived by choosing the test function as exactly the trial function in the weak formulation. This typically leads to a bound in terms of an energy-type norm that involves the H^1-seminorm scaled by $\varepsilon^{1/2}$ and the L^2-norm. The coercivity bound typically has a constant factor that requires a positive reaction coefficient when the advective field is divergence-free (see, for example, [106, Lemma 4.2], [86, Lemma 4.59]). This means that a simple advection-diffusion problem with a constant advective field, which is necessarily divergence-free, would lose coercivity in the energy-type norm.

In [76, Lemma 4.3], a similar coercivity argument was able to circumvent the need of a strictly positive reaction term while retaining a bound for the energy-type norm. However, the resulting coercivity constant depends on ε, entailing a weakened stability when $\varepsilon \ll 1$. Moreover, this ε-dependence of the stability constant eventually manifests in the error analysis, resulting in a nonrobust a priori error estimate with respect to the Péclet number.

The pivotal development in deriving a stronger stability bound of finite element methods for advection-diffusion-reaction equations appeared in [12] which is inspired by an analysis on the PDE itself from decades earlier [40]. The latter provides a well-posedness analysis of the PDE in its pure hyperbolic limit ($\varepsilon = 0$) by imposing the following regularity conditions on the advective field $\bar{\beta}$: (1) $\bar{\beta} \in W^{1,\infty}(\Omega)$; (2) $\bar{\beta}$ has no stationary point in the domain, i.e., $|\bar{\beta}(x)| \neq 0$ for any $x \in \Omega$; (3) $\bar{\beta}$ has no closed curves. The last condition means that any subcharacteristic $\xi_x(\tau)$, defined as the solution of the ordinary differential equation $\frac{\mathrm{d}\xi}{\mathrm{d}\tau} = \bar{\beta}(\xi(\tau))$ with $\xi(0) = x \in \Omega$, leaves the domain Ω in a finite time. See also [92, Part III Chapter 1] for related discussions.

One theoretical implication of the aforementioned set of assumptions on $\bar{\beta}$ is the existence of a smooth function ψ such that $\bar{\beta} \cdot \overline{\nabla}\psi(x) \geq b_0$ for some constant $b_0 > 0$, which depends on the inverse of the diameter of the domain Ω. This function ψ turns out to be the key theoretical devise which [12] employs to obtain a stronger stability estimate for the DG method therein. In particular, they define a weighting function $\varphi := \exp(-\psi) + \chi$, with χ a free to choose positive constant. Then, instead of choosing the test function as the trial function itself, they use the product of the trial function and the weighting function φ as the new test function. This results in a coercivity-type bound with respect to an energy-type norm and simultaneously, a stability constant independent of the diffusion parameter ε. By projecting this weighted test function to the DG finite element space and by taking into account the corresponding projection estimate, they are able to prove a discrete inf-sup stability in the advection-dominated regime. A Péclet-robust a priori error analysis follows in a standard fashion. Additionally, they demonstrate that the inf-sup condition can be further enhanced to bound a norm that also provides control of the streamline derivative. The same idea is used to analyze an HDG method for the stationary advection-diffusion problem in the advection-dominated regime in [52].

Inspired by the weighted test function approach, we will construct a weighted test function to show stability of the space-time HDG method. However, we make the important observation that in the space-time formulation of the time-dependent advection-diffusion problem, the space-time advective field combines the time derivative and the spatial advective field, $\partial_t u + \overline{\nabla} \cdot (\bar{\beta} u) = \nabla \cdot (\beta u)$ with $\beta := (1, \bar{\beta})$ and $\nabla := (\partial_t, \overline{\nabla})$, and hence has a constant component in the time direction (see section 2.1 for the detailed setup). With the assumption that $\bar{\beta} \in W^{1,\infty}(\Omega)$, this constant component ensures that (1) $\beta \in W^{1,\infty}(\mathcal{E})$; (2) β has no stationary point in the space-time domain $\mathcal{E}$; and (3) β has no closed curves. The last statement is because any subcharacteristic leaves the space-time domain $\mathcal{E}$ in a finite time bounded by T. Therefore, differing from [12, 52], we are guaranteed for free a smooth function ψ such that $\beta \cdot \nabla\psi(x) \geq b_0$ for some constant $b_0 > 0$, which depends on the inverse of the diameter of the space-time domain $\mathcal{E}$.

A further development shows that we are able to simplify the analysis by explicitly constructing the smooth function as $\psi = t/T$ and the weighting function as $\varphi = eT \exp(-t/T) + \chi$ (see eq. (4.3)), i.e., the weighted test function depending only on the time variable and b_0 depending only on $1/T$ (the inverse of the diameter of the space-time domain in the time direction). Based on this choice of the weighted test function, we prove an inf-sup stability with its constant independent of the diffusion parameter ε (see eq. (4.2a)) in place of the coercivity result [76, Lemma 4.3]. The proof is similar to its counterparts in [12, 52] where a projection estimate of the weighted test function (see lemma 4.4) is combined with a coercivity-type Péclet robust stability bound (see lemma 4.1).

Based on this new Péclet-robust inf-sup stability, we prove the second Péclet-robust inf-sup stability in an enhanced norm which provides control on the time derivative. This result finds its counterpart in [76, Theorem 4.4]. However, the choice of our test function proves more convoluted (see eq. (4.42)). Finally, analogous to [12], we further enhance the second inf-sup stability to a norm that also provides control on the streamline derivative. This results in the main Péclet-robust inf-sup stability of our a priori error analysis, see theorem 4.1. The Péclet-robust a priori error estimate can be shown based on the inf-sup stability in a standard fashion. To the best of the author's knowledge, this is the first a priori error analysis of an HDG method for the time-dependent advection-dominated advection-diffusion problem on moving domains.

1.3 Adaptivity and a posteriori error analysis

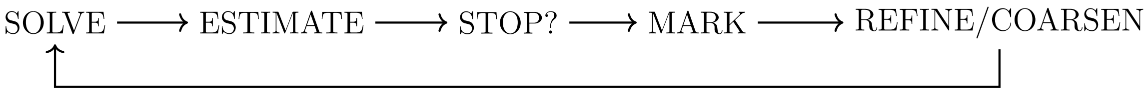
A segue: why adaptivity? When solving advection-dominated advection-diffusion problems with uniform mesh refinement, which is the assumed refinement strategy in the a priori error analysis, and when sharp boundary and/or interior layers are present in the solution, the local approximation error in the narrow boundary/interior layer region tends to dominate the global error. This imbalance of error distribution can manifest itself in a dramatic fashion where only a small portion of the elements contribute to, for example, more than 99% of the error.

The objective of adaptivity is therefore to allocate more elements/dofs to areas of the domain where the local numerical approximation has the largest errors. This alternative mesh refinement strategy is known as adaptive mesh refinement (AMR). In the case of advection-dominated advection-diffusion problems, resolving sharp layers with sufficiently small approximation errors requires mesh elements at a similar scale as the boundary/interior layer width. Meanwhile, outside this narrow layer region, a similar level of local approximation error can be achieved with much coarser elements. A successfully executed AMR procedure, therefore, balances the local errors throughout the mesh and has the potential to yield an "optimal mesh" associated with a specific global error tolerance. Figure 1.3 shows a test case of eq. (1.1) on the space-time domain $[0,1]^3$ with the exact solution being

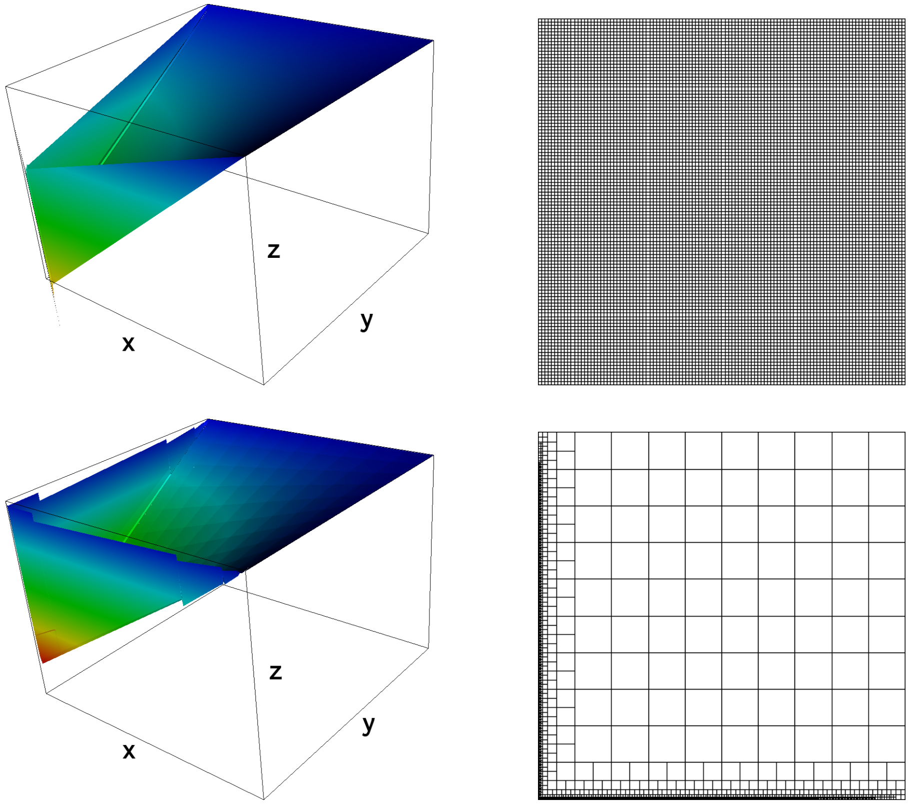
$$u(t,x,y) = (1-\exp(-t))\left(\tfrac{\exp((x-1)/\varepsilon)-1}{\exp(-1/\varepsilon)-1} + x - 1\right)\left(\tfrac{\exp((y-1)/\varepsilon)-1}{\exp(-1/\varepsilon)-1} + y - 1\right).$$

The solution exhibits boundary layers of width $\mathcal{O}(\varepsilon)$ near the boundary of the domain where $x = 1$ or $y = 1$. Two solutions are shown in fig. 1.3 implemented with uniform mesh refinement and AMR respectively. We observe that with less dofs, the solution on the adaptively refined mesh successfully resolves the boundary layer whereas its counterpart on the uniformly refined mesh does not.

We remark that, under the same principle, special layer-adapted meshes have been devised for layer problems [92]. This, however, requires the location of the layer to be known a priori which is typically not possible especially for time-dependent problems. AMR, on the other hand, does not require a priori information of the solution and hence is generally a more suitable approach. The lack of a priori information of the solution also rules out a priori error estimates as a viable guide for the AMR procedure. A standard form of an a priori error estimate is $\|u - u_h\|_{1,\Omega}^2 \leq \sum_{K\in\Omega} ch_K^2 \, |u|_{2,K}^2$, which contains unknown local quantity $|u|_{2,K}$ of the exact solution. Thus it does not provide a computable local error estimate and it is usually used to show the asymptotic convergence rate under uniform mesh refinement. This motivates the a posteriori error estimation, which purports to provide local error estimates that can be computed at a low cost using known and computable quantities only, such as the finite element solution u_h, the problem data, the boundary conditions, and the geometric data of the mesh. With an a posteriori error estimator, the standard AMR procedure is enabled and proceeds as follows:



Inside this loop, in the SOLVE step we obtain the finite element solution u_h on the current mesh $\mathcal{T}_h$, and in the ESTIMATE step we compute the local error estimate on each element. This estimate is denoted by $\eta_{\mathcal{K}}$. The summation of $\eta_{\mathcal{K}}$ over all $\mathcal{K} \in \mathcal{T}_h$ gives an estimate for the global error $\|u - u_h\|_{\mathcal{T}_h}$. In the STOP step we check whether the global error estimate is smaller than a prescribed error tolerance. This serves as the stopping criterion of the procedure. If the error tolerance has not been reached, we proceed with the MARK step in which we mark all elements with $\eta_{\mathcal{K}}$ bigger than a prescribed threshold. Common marking strategies include: (1) a certain percentage of the elements with the biggest local error estimates are marked; (2) elements with error estimates bigger than a certain percentage of the biggest local estimate are marked; and (3) elements whose local error estimates together constitute a certain percentage of the global error estimate are marked. Similarly, a portion of elements with relatively smaller error estimates are marked for coarsening. Finally the REFINE/COARSEN step applies a prescribed refinement strategy to the elements marked for refinement and coarsens the elements marked for coarsening. We then proceed again to the SOLVE step, now on this new mesh. In this thesis, we only consider regular refinement



Refinement	Number of elements	Number of facet dofs	L^2-error	Spatial H^1-error
Uniform	1,404,928	17,009,664	4.1e-3	1.6e-1
Adaptive	1,173,990	16,801,500	7.9e-4	9.7e-2

Figure 1.3: This is a boundary layer example implemented using the space-time HDG method in this thesis. Uniform refinement (upper row) and adaptive refinement (lower row) are employed and their spatial solutions at the final time are plotted. With slightly less dofs, adaptive refinement resolves the boundary layer whereas the uniform refinement still has not. Furthermore, numbers of elements and dofs on the global space-time mesh as well as global L^2- and H^1-errors are tabulated. The degree to which the boundary layer has been resolved is reflected in the errors.

whereby a hexahedral element is divided into eight smaller hexahedral elements by joining the midpoints of edges.

For the AMR loop to properly function, we expect two crucial properties of the a posteriori error estimator $\eta_{\mathcal{K}}$. Firstly, the stopping criterion requires the global error estimate to bound the exact global error as follows

$$\|u - u_h\|_{\Omega} \leq c^* \Big(\sum_{K \in \mathcal{T}} \eta_{\mathcal{K}}^2 \Big)^{1/2} . \tag{1.2}$$

With eq. (1.2), if $\left(\sum_{\mathcal{K} \in \mathcal{T}} \eta_{\mathcal{K}}^2\right)^{1/2} < \tau$, the prescribed error tolerance, we can conclude that $\|u - u_h\|_{\Omega} < \tau$, up to a multiplicative constant c^*. The constant c^* should be independent of the mesh-size parameters and ideally be $\mathcal{O}(1)$ at all levels of refinement. This is known as the *reliability* property of an a posteriori error estimator. Secondly, for a well-informed selection of elements during the MARK step, we want the local error estimate to be a lower bound for the exact local error

$$\eta_{\mathcal{K}} \leq c_* \|u - u_h\|_{\mathcal{K}} , \quad \forall \mathcal{K} \in \mathcal{T}_h . \tag{1.3}$$

With eq. (1.3), if $\eta_{\mathcal{K}}$ surpasses the marking threshold, we deduce that $\|u - u_h\|_{\mathcal{K}}$ is also greater than the marking threshold, up to a multiplicative constant c_*. Similarly, the constant c_* should be independent of the mesh-size parameters and ideally be $\mathcal{O}(1)$ at all levels of refinement. This property of the error estimator is known as *local efficiency*. Furthermore, the ratio between the estimated error, $\left(\sum_{\mathcal{K} \in \mathcal{T}_h} \eta_{\mathcal{K}}^2\right)^{1/2}$, and the exact error, $\|u - u_h\|_{\Omega}$ measures the quality of the error estimator and should ideally be $\mathcal{O}(1)$ at all levels of refinement. This ratio is called the *efficiency index*. Combining reliability and (local) efficiency leads to a bound for the efficiency index, $[1/c^*, c_*]$.

For solutions of the advection-dominated advection-diffusion equation, a reliable and locally efficient error estimator might still prove insufficient in driving the AMR procedure. This usually manifests when the efficiency index is dependent on the problem parameter, which, in this case, is the Péclet number ε^{-1}. Particularly, when the Péclet number grows, the sharpness of the reliability and local efficiency bounds eqs. (1.2) and (1.3) deteriorates. In other words, as the layers become sharper, which corresponds to a larger Péclet number and which signals a greater demand for AMR, the quality of the error estimator worsens. Therefore, it is of great interest to aim for the independence of the reliability and local efficiency constants with respect to the problem parameter that behaves singularly, a property aptly termed *robustness* of the error estimator.

The a posteriori error analysis of this thesis. We present an a posteriori error analysis for a space-time HDG method for the time-dependent advection-diffusion problem on fixed domains.

For the stationary advection-diffusion problem, in the advection-dominated limit, a posteriori error analysis has been done for various finite element methods. Examples include a posteriori error analysis for conforming finite element methods [81, 93, 100, 103], DG methods [47, 48, 56, 95, 109], and HDG methods [6, 33, 96]. The focus of these studies has been the robustness of the error estimator with respect to the Péclet number.

Nonrobustness of the error estimator for the standard energy norm was first observed in [100]. The analysis therein was based on the standard coercivity (which is used to show reliability) and boundedness (which is used to show local efficiency) with respect to the standard energy norm of the weak formulation of the PDE (see [100, Section 4]). It showed that the upper and lower bound constants (as in eqs. (1.2) and (1.3) respectively) differ by a factor $\varepsilon^{-1/2}$ particularly when narrow layer regions have not been sufficiently resolved. This behaviour may be attributed to the lack of measurement of the streamline derivative in the standard energy norm. Furthermore, a space-time version of the error estimator in [100] is derived and analyzed in [7] for the time-dependent advection-diffusion problem, inheriting the nonrobustness.

To attain robustness, one possible approach is to augment the energy norm to measure the streamline derivative. Then, based on the enhanced norm, one may derive an inf-sup condition in place of the coercivity and a new boundedness result. Ideally, these two results lead to upper and lower bound constants that are independent of ε. This idea forms the basis of the newly introduced dual norm in [103] which measures the error in the streamline derivative. A Péclet robust continuous inf-sup condition was proved with respect to the augmented energy norm by the dual norm. Meanwhile, a boundedness result with respect to the augmented norm is shown with no dependence on the Péclet number. Combining the corresponding upper and lower bounds results in robustness (see [103, Lemma 3.1 and Theorem 4.1]). This approach was also used in the a posteriori error analysis of DG methods for the stationary advection-diffusion problem (see [48, 95]). An alternative dual norm, argued to be more suitable for advection-dominated problems, was presented in [93]. Their residual-based estimator was shown to be almost robust in one spatial dimension.

Based on [103], the dual norm technique is extended to analyzing an a posteriori error estimator for the time-dependent advection-diffusion problem in [102], which is shown to be robust. Computing the space-time error estimator therein is not trivial and requires solving an auxiliary stationary reaction-diffusion problem. Similarly, the error estimator in [95] for the DG method, which is also robust with respect to an energy norm augmented by a dual norm, is extended to the time-dependent problem in [31, 32]. The latter extension follows the elliptic reconstruction technique [55, 82, 83], which provides a general framework to extend error estimators for the stationary problem to the time-dependent problem.

However, the dual norm approach is not without its downside. Particularly, the nature of it being dual with respect to a global (energy) norm renders its local evaluation or estimation impossible. Therefore, the efficiency bound (as in eq. (1.3)) for the error estimator can only be global. An a posteriori error analysis, not involving dual norms, was presented in [33] and was later extended to the Oseen problem in [6]. The reliability analysis is based on a Péclet-robust coercivity-type result. This result, as an alternative of the Péclet robust continuous inf-sup condition proved for the dual norm, was inspired by the a priori error analysis in [12]. As we discussed previously in this chapter on the a priori error analysis, the analysis in [12] is based on the weighted test function and the assumptions that $\bar{\beta}$ lives in $W^{1,\infty}(\Omega)$ and has no closed curves nor stationary points. Using the weighted test function, a coercivity-type bound is derived (see [12, Lemma 4.4]) and it is closely related to the bound used to show reliability in [33, 6] (see, respectively, [33, Lemma 4.1], [6, Lemma 3.6]). Robustness of the a posteriori error estimator was shown in [33] for the stationary advection-diffusion problem. Furthermore, without any dual norm, the norm in [33] is locally-computable and a local efficiency result is provided.

This has naturally led us to exploit the Péclet-robust a priori error analysis in this thesis in order to obtain an a posteriori error estimator for the space-time HDG discretization of the time-dependent advection-diffusion problem. Analogous to [33], the basis for the a posteriori error analysis in this thesis is the intermediate Péclet-robust coercivity result (see lemma 4.1) we proved for the a priori error analysis. This results in a reliability bound for the L^2- and spatial H^1-norms of the error, but not for the error of the time derivative. For the latter, we use a saturation assumption, inspired by [23].

Let $\mathcal{T}_h$ be a given mesh and let $\mathcal{T}_{\mathfrak{h}}$ be a mesh obtained by applying a level of refinement on $\mathcal{T}_h$. Let u_h be the finite element solution on mesh $\mathcal{T}_h$ and $u_{\mathfrak{h}}$ the finite element solution on mesh $\mathcal{T}_{\mathfrak{h}}$. A saturation assumption supposes that $u_{\mathfrak{h}}$ has a strictly smaller error than u_h. In other words, we have

$$\|u - u_{\mathfrak{h}}\|_{\Omega} \leq \rho \|u - u_h\|_{\Omega} \quad \text{for} \quad \rho < 1.$$

By a triangle inequality, we then have

$$\|u - u_h\|_{\Omega} \leq \frac{1}{1-\rho} \|u_{\mathfrak{h}} - u_h\|_{\Omega}.$$

The saturation assumption provides an approach to estimate $\|u - u_h\|_{\Omega}$ by estimating $\|u_{\mathfrak{h}} - u_h\|_{\Omega}$ instead. For the latter, one can typically rely on a combination of discrete inf-sup stability and Galerkin orthogonality. This was done in [23] for an a posteriori error analysis of the advection-reaction equation. Certain restrictions had to be placed on the subgrid refinement to construct theoretically viable $\mathcal{T}_{\mathfrak{h}}$ in [23] which prevents the analysis to be applicable in three-dimensions. In this thesis, since desired error estimation is for the error in the time derivative, we rely on a subgrid constructed by halving the time-step of every space-time element. A time derivative error estimate is then obtained by combining a Galerkin orthogonality and a discrete inf-sup stability. The latter comes from the inf-sup condition we proved for the a priori error analysis with respect to a norm that involves a term that measures the time derivative. See eq. (4.2b) and theorem 5.3. Finally, we remark that the saturation assumption may fail in general. See [1, Section 5.2] and [18].

Due to that the saturation assumption of our interest does not hold for constant polynomial approximation in time, we will use linear polynomial approximation in time in the space-time HDG discretization. The resulting a posteriori error analysis is thus for a second order accurate in time and arbitrary order accurate in space space-time HDG discretization of the time-dependent advection-diffusion problem. We remark that despite a nonrobust a posteriori error bound, as shown in theorems 5.1 and 5.2, the norm we use is locally computable and also measures the error in the time derivative. Furthermore, the error estimator in this work is fully local hence it is an estimator for local space and time adaptivity in the AMR procedure.

1.4 Implementations in this thesis

Numerical experiments in this thesis are implemented in the finite element library deal.II [8, 9] with distributed memory parallelization [16]. In contrast with shared memory parallelization, the mesh is decomposed by the p4est library [29] and each processor only stores a subset of elements with a distributed data structure. The communication between machines is then handled by an implementation of Message Passing Interface (MPI). This allows our implementation to run test cases with up to 1000 processors and 50 million dofs (after static condensation). The linear system arising from the space-time HDG discretization is solved all-at-once using the Multifrontal Massively Parallel Solver (MUMPS) [3, 4]. We remark that on uniformly refined meshes, the solution process can alternatively be carried out using a slab-by-slab approach. By partitioning the global time interval into subintervals, the initial space-time domain is divided into space-time slabs. Each space-time slab is then tessellated and the PDE is discretized and solved on the space-time slab mesh from one time subinterval to the next using the solution on the current space-time slab as an initial condition for the next. See fig. 1.4 for an illustration. Besides MUMPS, we also used preconditioned GMRES in PETSc [14, 13, 15] to solve the linear system arised from the slab-by-slab approach. The GMRES is preconditioned by classical algebraic multigrid

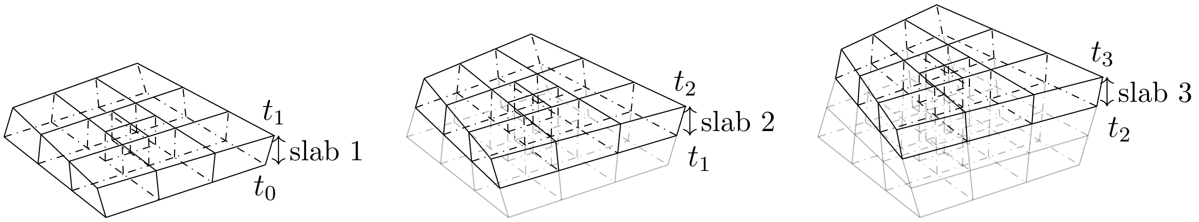


Figure 1.4: The slab-by-slab approach: The discretization on the global space-time mesh is solved one slab at a time, exemplified by the three slabs above. The solution of slab 1 is used as initial condition for the problem on slab 2, etc. When solving on a time-dependent domain, the slab moves forward in time according to the domain deformation mapping.

from BoomerAMG [59] with an absolute solver tolerance of 10^{-12}. Finally, we remark that the memory requirement is always lower for the slab-by-slab approach compared to the all-at-once approach. However, the all-at-once approach allows for a straightforward implementation of space-time adaptivity and so we adopt the all-at-once approach for the AMR procedure in this thesis.

We acknowledge that the research in this thesis is enabled by support provided by

- Math Faculty Computing Facility at the University of Waterloo (`https://uwaterloo.ca/math-faculty-computing-facility/`);
- Simon Fraser University (`https://www.sfu.ca/research/supercomputer-cedar`);
- Compute Ontario (`https://www.computeontario.ca/`);
- Calcul Québec (`https://www.calculquebec.ca/`);
- Digital Research Alliance of Canada (`https://alliancecan.ca`).

1.5 Thesis outline

The subsequent chapters of this thesis are organized as follows:

Chapter 2: The time-dependent advection-diffusion problem on a moving domain is recast into its space-time formulation on the $(d+1)$-dimensional space-time domain. This formulation automatically accounts for the domain deformation. To tessellate the space-time domain, we introduce geometric objects such as space-time slabs, elements, facets

(d-dimensional faces) and edges (($d-1$)-dimensional faces) and describe regularity conditions thereof. Finite element spaces are then defined on the space-time mesh and a space-time HDG method is introduced.

Chapter 3: Before we lay out the a priori and a posteriori error analyses of the space-time HDG method, we present in this chapter the useful theoretical tools such that scaling arguments, inverse and trace inequalities, local projection estimates, approximation estimates of an averaging operator (also known as the Oswald approximation operator), and a subgrid projection estimate. All these bounds, inequalities, and estimates are formulated and proved with space-time anisotropy.

Chapter 4: As the first contribution of this thesis, this chapter presents a Péclet-robust a priori error analysis of the space-time HDG method when applied to the advection-dominated advection-diffusion problem on moving domains. Based on a weighted test function technique, a novel inf-sup condition is proved as the key result to attain Péclet-robustness. This inf-sup stability is then extended to a norm that also measures the error of the streamline derivative. The error analysis and numerical examples conclude this chapter.

Chapter 5: The second contribution of this thesis is an a posteriori error analysis of the space-time HDG method when applied to the advection-dominated advection-diffusion problem on fixed domains. We present the a posteriori error estimator and prove its reliability and local efficiency. A novel saturation assumption along with a subgrid projection estimation are employed to estimate the error in the time derivative. Finally, we illustrate the theory with numerical examples that involve boundary and interior layers.

Chapter 6: The thesis concludes with discussions on potential future work based on the research in chapters 3 to 5.

Chapter 2

Space-time HDG for advection-diffusion problems

In this chapter, we introduce the space-time formulation of the time-dependent advection-diffusion equation and a space-time HDG method. The analysis of the latter will be the focus of this thesis. In section 2.1, the time-dependent advection-diffusion equation eq. (1.1) is reformulated into eq. (2.1), which is more convenient for the analysis. In section 2.2, the space-time HDG method is subsequently introduced in detail with two parts: in section 2.2.1 we describe geometries of space-time slabs, elements and facets as well as regularity conditions imposed on these geometric objects; in section 2.2.2, we present the finite element spaces, norms, conditions on the problem data, and finally, the space-time HDG discretization eq. (2.9).

2.1 The advection-diffusion problem

The space-time formulation of the advection-diffusion equation consists in recasting eq. (1.1) as a problem in $(d+1)$-dimensional space-time. For this, we define the $(d+1)$-dimensional polyhedral space-time domain as $\mathcal{E} := \{(t,x) : x \in \Omega(t), 0 < t < T\} \subset \mathbb{R}^{d+1}$. Its boundary, $\partial\mathcal{E}$, consists of three disjoint parts

$$\begin{aligned} \Omega(0) &:= \{(t,x) \in \partial\mathcal{E} \,:\, t = 0\}, \\ \Omega(T) &:= \{(t,x) \in \partial\mathcal{E} \,:\, t = T\}, \\ \mathcal{Q}_{\mathcal{E}} &:= \{(t,x) \in \partial\mathcal{E} \,:\, 0 < t < T\}. \end{aligned}$$

The outward space-time normal vector to $\partial\mathcal{E}$ is denoted by $n := (n_t, \bar{n})$, where n_t and $\bar{n}$ are temporal and spatial components of the space-time normal vector, respectively. Introducing the space-time advective field $\beta := (1, \bar{\beta})$ and the space-time gradient operator $\nabla := (\partial_t, \overline{\nabla})$, the space-time formulation of eq. (1.1) is given by

$$\nabla \cdot (\beta u) - \varepsilon \overline{\nabla}^2 u = f \text{ in } \mathcal{E}. \tag{2.1a}$$

We consider a nonoverlapping partition of the domain boundary, $\partial\mathcal{E} = \partial\mathcal{E}_D \cup \partial\mathcal{E}_N$, and impose the boundary conditions

$$-\zeta^- u\beta \cdot n + \varepsilon \overline{\nabla} u \cdot \bar{n} = g \text{ on } \partial\mathcal{E}_N, \tag{2.1b}$$

$$u = 0 \text{ on } \partial\mathcal{E}_D. \tag{2.1c}$$

The Dirichlet $\partial\mathcal{E}_D$ and Neumann $\partial\mathcal{E}_N$ boundaries are defined by:

$$\begin{aligned} \partial\mathcal{E}_D &:= \{(t, x) : x \in \Gamma_D(t), 0 < t \leq T\}, \\ \partial\mathcal{E}_N &:= \{(t, x) : x \in \Gamma_N(t) \cup \Omega(0) \cup \Omega(T), 0 < t \leq T\}, \end{aligned}$$

where we also prescribe a nonoverlapping partition of the boundary of $\Omega(t)$, i.e., $\partial\Omega(t) = \Gamma_D(t) \cup \Gamma_N(t)$. Furthermore, ζ^- is an indicator function for the inflow (where $\beta \cdot n < 0$) part of the boundary of $\mathcal{E}$. Therefore, the boundary condition on $\partial\mathcal{E}_N$ also imposes the initial condition $u(x, 0) = g(x)$ on $\Omega(0)$. Finally, we assume that the forcing term f lies in $L^2(\mathcal{E})$ and that the Neumann boundary data g lies in $L^2(\partial\mathcal{E}_N)$.

2.2 The space-time HDG method

2.2.1 Description of space-time slabs, elements, facets and edges

An initial partition of the space-time domain $\mathcal{E}$ consists of dividing the time interval $[0, T]$ into time levels $0 = t_0 < t_1 < \cdots < t_N = T$ and defining the nth time interval as $I_n = (t_n, t_{n+1})$. The space-time domain is divided into space-time slabs $\mathcal{E}^n := \mathcal{E} \cap (I_n \times \mathbb{R}^d)$, which are then divided into space-time elements, $\mathcal{E}^n = \cup_j \mathcal{K}_j^n$. To construct the space-time element $\mathcal{K}_j^n$, we divide the domain $\Omega(t_n)$ into nonoverlapping spatial elements K_j^n so that $\Omega(t_n) = \cup_j K_j^n$. Let Υ be the transformation describing the deformation of the domain. The spatial elements K_j^{n+1} at t_{n+1} are obtained by mapping the nodes of the elements K_j^n into their new position via the transformation Υ. Each space-time element $\mathcal{K}_j^n$ is obtained by connecting the elements K_j^n and K_j^{n+1} via linear interpolation in time following [99].

When domain is fixed, we remark that Υ becomes the identity mapping. We denote the set of all space-time elements tessellating the space-time domain by $\mathcal{T}_h$.

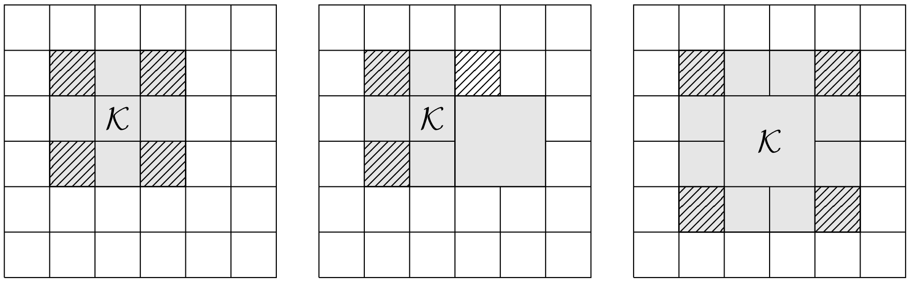
The boundary of a space-time element $\mathcal{K}$ with $t \in (t_*, t^*)$ is partitioned as $\partial\mathcal{K} = \mathcal{Q}_\mathcal{K} \cup \mathcal{R}_\mathcal{K}$ where $\mathcal{R}_\mathcal{K} := K_* \cup K^*$, $\mathcal{Q}_\mathcal{K} \cap \mathcal{R}_\mathcal{K} = \emptyset$, and where K_* denotes the facet of $\mathcal{K}$ at time $t = t_*$ and K^* denotes the facet of $\mathcal{K}$ at time t^*. On $\partial\mathcal{K}$, the outward unit space-time normal vector is denoted by $n^\mathcal{K} = (n_t^\mathcal{K}, \bar{n}^\mathcal{K})$, where $n_t^\mathcal{K}$ and $\bar{n}^\mathcal{K}$ are the temporal and spatial components of the space-time normal vector, respectively. For ease of notation, we omit the superscript $\mathcal{K}$ from now on. Note that $\bar{n}$ is the zero vector on an $\mathcal{R}$-facet, i.e., that $\bar{n} = 0$ on K^* and K_*, and that $\bar{n} \neq 0$ on a $\mathcal{Q}$-facet. The a posteriori error analysis in this thesis also requires the $(d-1)$-dimensional edges of K^* and K_*. We denote such an edge by $E_\mathcal{K}$.

We will allow at most 1-irregularly refined space-time elements in the space-time mesh $\mathcal{T}_h$. The facets in the mesh can be divided into three cases: (1) boundary facets; (2) interior facets shared by two elements at the same refinement level; (3) interior facets shared between more than two elements. We denote the set of all facets by $\mathcal{F}_h$. Within this set, the sets of all interior facets, boundary facets, $\mathcal{Q}$-facets (facets on which $\bar{n} \neq 0$), and $\mathcal{R}$-facets (facets on which $\bar{n} = 0$) are denoted by $\mathcal{F}_h^i$, $\mathcal{F}_h^b$, $\mathcal{F}_{\mathcal{Q},h}$, and $\mathcal{F}_{\mathcal{R},h}$, respectively. The union of all facets in $\mathcal{F}_h$ is denoted by Γ. Furthermore, we denote by $\partial\mathcal{T}_h$ the set of element boundaries, by $\mathcal{Q}_h$ the set that consists of parts of an element boundary on which $\bar{n} \neq 0$, by $\mathcal{R}_h$ the set that consists of parts of an element boundary on which $\bar{n} = 0$, and by $\partial\mathcal{T}_h^i$ the set of element boundaries excluding the part of the element boundary that lies on $\partial\mathcal{E}$.

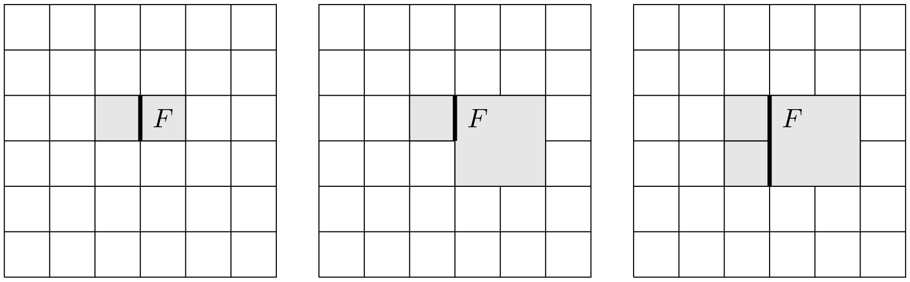
We denote by $\omega_\mathcal{K}$ the union of elements $\mathcal{K}'$ such that $\partial\mathcal{K} \cap \partial\mathcal{K}' \neq \emptyset$, and denote by $\sigma_\mathcal{K}$ the union of elements that share at least one vertex with $\mathcal{K}$. Consider now a facet F. Any elements containing facets F' such that $F \cap F'$ is itself a facet belong to the set ω_F. See fig. 2.1 for a depiction of $\omega_\mathcal{K}$, $\sigma_\mathcal{K}$, and ω_F.

To define the finite element spaces, we require the mapping $\Phi_\mathcal{K}$ between a fixed reference element $\widehat{\mathcal{K}} = (-1,1)^{d+1}$ and space-time element $\mathcal{K} \in \mathcal{T}_h$. Following [53] and [98], this mapping $\Phi_\mathcal{K}(\widehat{\mathcal{K}}) = \mathcal{K}$ is decomposed into two parts. First, $G_\mathcal{K}(\widehat{\mathcal{K}}) = \widetilde{\mathcal{K}}$ denotes the affine mapping defined by $G_\mathcal{K}(\widehat{x}) = A_\mathcal{K}\widehat{x} + b$, where $A_\mathcal{K} = \text{diag}\ (\delta t_\mathcal{K}/2, h_K/2, \ldots, h_K/2)$ and $b \in \mathbb{R}^{d+1}$ a constant translation vector such that the brick $\widetilde{\mathcal{K}} := (0, \delta t_\mathcal{K}) \times (0, h_K)^d$, see fig. 2.2. In the following, h_K is used to denote the spatial size of the element $\mathcal{K}$ and $\delta t_\mathcal{K}$ the time-step. We then define $\Phi_\mathcal{K} := \phi_\mathcal{K} \circ G_\mathcal{K}$, where $\phi_\mathcal{K}$ is a diffeomorphism such that $\phi_\mathcal{K}(\widetilde{\mathcal{K}}) = \mathcal{K}$ (see fig. 2.2). Note that $G_\mathcal{K}$ sets the size of the element $\mathcal{K}$ while $\phi_\mathcal{K}$ sets its shape. Following [53] and [98], we assume that $\phi_\mathcal{K}$ is close to the identity, i.e., we will assume that $\phi_\mathcal{K}$ satisfies:

$$c^{-1} \leq |\det J_{\phi_\mathcal{K}}| \leq c, \quad \|(J_{\phi_\mathcal{K}})_{ij}\|_{L^\infty(\widetilde{\mathcal{K}})} \leq c \quad 0 \leq i,j \leq d, \quad \forall \mathcal{K} \in \mathcal{T}_h, \tag{2.2}$$



(a) $\omega_{\mathcal{K}}$ and $\sigma_{\mathcal{K}}$



(b) ω_F

Figure 2.1: Depiction of sets $\omega_{\mathcal{K}}$, $\sigma_{\mathcal{K}}$, and ω_F on conforming and 1-irregularly refined meshes. Figure (a): elements in the set $\omega_{\mathcal{K}}$ are the grey colored elements excluding the hatched elements; elements in the set $\sigma_{\mathcal{K}}$ are colored grey and include the hatched elements. Figure (b): elements in the set ω_F are colored grey.

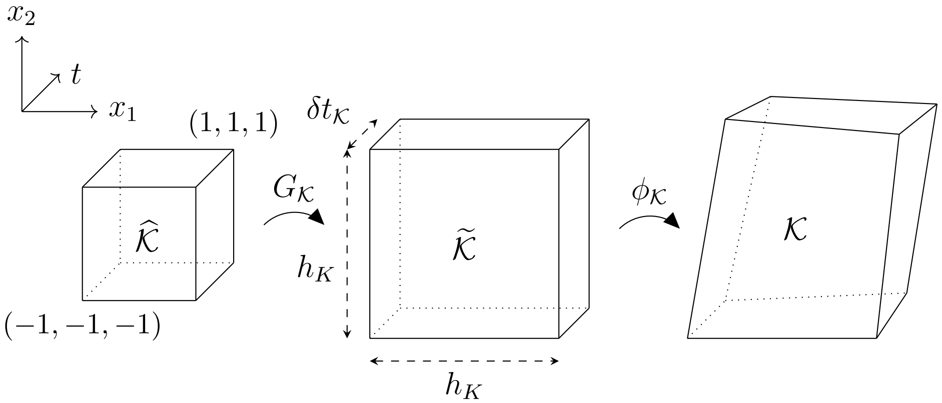


Figure 2.2: Construction of the space-time element $\mathcal{K}$ through an affine mapping $G_{\mathcal{K}} : \widehat{\mathcal{K}} \to \widetilde{\mathcal{K}}$ and a diffeomorphism $\phi_{\mathcal{K}} : \widetilde{\mathcal{K}} \to \mathcal{K}$ [98]. Note that the front and back faces of $\mathcal{K}$ have constant t-coordinate and hence are parallel to each other.

where c is a generic constant independent of h_K, $\delta t_{\mathcal{K}}$, ε, and T, where $J_{\phi_{\mathcal{K}}} \in \mathbb{R}^{(n+1)\times(n+1)}$ is the Jacobian of the diffeomorphism $\phi_{\mathcal{K}}$, and where the index 0 denotes the coordinate of the time direction. Since t only depends on $\widetilde{t}$,

$$\left(J_{\phi_{\mathcal{K}}}\right)_{0k} = 0, \quad 1 \le k \le d, \quad \forall \mathcal{K} \in \mathcal{T}_h. \tag{2.3a}$$

We remark that when domain is fixed, x_k is independent of $\tilde{t}$ for $1 \le k \le d$, thus, we have

$$\left(J_{\phi_{\mathcal{K}}}\right)_{k0} = 0, \quad 1 \le k \le d, \quad \forall \mathcal{K} \in \mathcal{T}_h. \tag{2.3b}$$

For the inverse of $J_{\phi_{\mathcal{K}}}$, let $\det J_{\phi_{\mathcal{K}} \backslash mn}$ denote the (m,n) minor of $J_{\phi_{\mathcal{K}}}$. We will assume that:

$$c^{-1} \le |\det J_{\phi_{\mathcal{K}}}^{-1}| \le c, \quad \|\det J_{\phi_{\mathcal{K}} \backslash mn}\|_{L^\infty(\widetilde{\mathcal{K}})} \le c, \quad \forall \mathcal{K} \in \mathcal{T}_h. \tag{2.4}$$

Let $F_{\mathcal{Q}}^j$ be a $\mathcal{Q}$-face where $\widetilde{x}_j$ is fixed in its affine domain. The parametrization of $F_{\mathcal{Q}}^j$, obtained from the restriction of $\phi_{\mathcal{K}}$ to the boundary of $\widetilde{\mathcal{K}}$ where $\widetilde{x}_j$ is fixed, is denoted by $\phi_{F_{\mathcal{Q}}}$. Then, (see [84, Theorem 21.3 and Definition on page 189] and appendix A),

$$\int_{F_{\mathcal{Q}}^j} f(x)\, \mathrm{d}s = \int_{\widetilde{F}_{\mathcal{Q}}^j} f\left(\phi_{F_{\mathcal{Q}}}(\widetilde{x})\right) \left(\det\left((J_{\phi_{\mathcal{K}}}^j)^\intercal J_{\phi_{\mathcal{K}}}^j\right)\right)^{1/2} \mathrm{d}\widetilde{s}, \tag{2.5}$$

where $J_{\phi_{\mathcal{K}}}^j \in \mathbb{R}^{(n+1)\times n}$ is obtained by removing the j^{th} column vector from $J_{\phi_{\mathcal{K}}}$. We will assume that

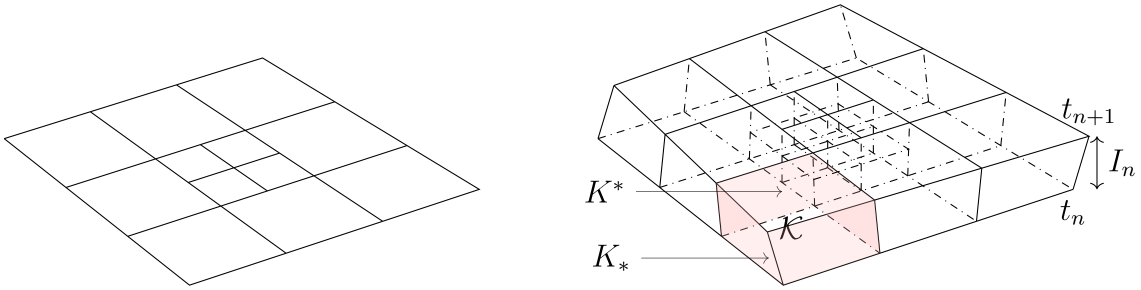
$$c^{-1} \le \left(\det\left((J_{\phi_{\mathcal{K}}}^i)^\intercal J_{\phi_{\mathcal{K}}}^i\right)\right)^{1/2} \le c, \quad 0 \le i \le d. \tag{2.6}$$

To account for local time-stepping, consider a space-time element $\mathcal{K}$ in space-time slab $\mathcal{E}^n$. Then we introduce, in addition to the local time-step $\delta t_{\mathcal{K}}$ set by $\Phi_{\mathcal{K}}$, the slab time-step $\Delta t_{\mathcal{K}} := t_{n+1} - t_n$, i.e., the length of I_n. Note that $\delta t_{\mathcal{K}} \le \Delta t_{\mathcal{K}}$ with $\delta t_{\mathcal{K}} < \Delta t_{\mathcal{K}}$ when using local time-stepping. We will assume that $\Delta t_{\mathcal{K}}/\delta t_{\mathcal{K}} \le c$ for all $\mathcal{K} \in \mathcal{T}_h$. An illustration of a $(d+1)$-dimensional space-time mesh in slab $\mathcal{E}^n$, with $d = 2$, is shown in fig. 2.3.

2.2.2 Finite element spaces and the discretization

Let $\partial_x^\alpha v$, with α a multi-index, be the weak derivative of v and let $H^s(U) := \{v \in L^2(U) : \partial_x^\alpha v \in L^2(U) \text{ for } |\alpha| \le s\}$, where s is a nonnegative integer and $U \subset \mathbb{R}^r$ is an open domain with $x := (x_1, \dots, x_r)$ denoting the coordinates of $\mathbb{R}^r$. The norm of $H^s(U)$ is defined by $\|v\|_{H^s(U)}^2 := \sum_{|\alpha|\le s} \|\partial_x^\alpha v\|_U^2$, where $\|\cdot\|_U$ is the usual L^2-norm on U.

We also require anisotropic Sobolev spaces. Following [98] we only consider anisotropy between spatial and temporal variables with no anisotropy between the spatial variables.



(a) A spatial mesh $\Omega(t_n) = \cup_j K_j^n$

(b) A space-time slab $\mathcal{E}_n$. The facets dyed in red form $\mathcal{Q}_{\mathcal{K}}$ while $\mathcal{R}_{\mathcal{K}}$ is the union of K^* and K_*.

Figure 2.3: Illustration of a moving spatial domain $\Omega(t) \subset \mathbb{R}^d$ for $t \in I_n$ resulting in the space-time slab $\mathcal{E}^n \subset \mathbb{R}^{(d+1)}$ (with $d = 2$). Local time-stepping within a space-time slab is featured in fig. 2.3b. Here $K^* = K^{n+1}$ and $K_* = K^n$.

As such, let s_s and s_t denote the spatial and temporal Sobolev indices, respectively. For $\alpha_t, \alpha_{s_i} \geq 0$, $1 \leq i \leq d$, the anisotropic Sobolev space of order (s_t, s_s) is defined on an open domain $U \subset \mathbb{R}^{d+1}$ by (see [53]):

$$H^{(s_t,s_s)}(U) := \{v \in L^2(U) : \partial_t^{\alpha_t}\partial_x^{\alpha_s} v \in L^2(U) \text{ for } \alpha_t \leq s_t, |\alpha_s| \leq s_s\},$$

where $\alpha_s = (\alpha_{s_1}, \ldots, \alpha_{s_d})$ and $x := (x_1, \ldots, x_d)$ denotes the spatial coordinates. The anisotropic Sobolev norm reads $\|v\|^2_{H^{(s_t,s_s)}(U)} := \sum_{\alpha_t \leq s_t, |\alpha_s| \leq s_s} \|\partial_t^{\alpha_t}\partial_x^{\alpha_s} v\|_U^2$.

For the HDG method, we require the following finite element spaces

$$\begin{aligned} V_h^{(p_t,p_s)} &:= \{v_h \in L^2(\mathcal{E}) : v_h|_{\mathcal{K}} \circ \phi_{\mathcal{K}} \circ G_{\mathcal{K}} \in Q^{(p_t,p_s)}(\widehat{\mathcal{K}}) \quad \forall \mathcal{K} \in \mathcal{T}_h\}, \\ M_h^{(p_t,p_s)} &:= \{\mu_h \in L^2(\Gamma) : \mu_h|_F \circ \phi_{\mathcal{K}} \circ G_{\mathcal{K}} \in Q^{(p_t,p_s)}(\widehat{F}) \quad \forall F \in \mathcal{F}_h, \mu_h = 0 \text{ on } \partial\mathcal{E}_D\}, \end{aligned}$$

where $Q^{(p_t,p_s)}(U)$ denotes the set of all tensor product polynomials of degree p_t in the temporal direction and p_s in each spatial direction on a domain U. For simplicity of notation, we omit the superscript (p_t, p_s) from now on and define $\boldsymbol{V}_h = V_h \times M_h$ and denote the pairs $(v, \mu) \in \boldsymbol{V}_h$ and $(u, \lambda) \in \boldsymbol{V}_h$ as $\boldsymbol{v} = (v, \mu)$ and $\boldsymbol{u} = (u, \lambda)$.

On an element boundary we denote the HDG jump by $[\boldsymbol{v}_h] := (v_h - \mu_h)$ and on a facet $F \in \mathcal{F}_h^i$, where $F \subset \partial\mathcal{K}_1 \cap \partial\mathcal{K}_2$, we denote the usual DG jump by $[\![v_h]\!] := (v_{h1}n_1 + v_{h2}n_2)$. Next, consider two elements $\mathcal{K}_1$ and $\mathcal{K}_2$ such that $K_1^* = K_{2,*}$. Denote the restriction of μ_h to $\mathcal{Q}_{\mathcal{K}_1}$ and $\mathcal{Q}_{\mathcal{K}_2}$ by μ_{h1} and μ_{h2}, respectively. The jump of μ_h across edges of K_1^* is defined by $\langle\!\langle \mu_h \rangle\!\rangle := \mu_{h1} - \mu_{h2}$. Note that for pairs of K_1 and K_2 such that $K_1^* \subsetneq K_{2,*}$ or $K_{2,*} \subsetneq K_1^*$, we do not define any edge jump.

To end this section we introduce $\boldsymbol{V} := V \times M$, where $V := \{v \in H^1(\mathcal{E}) \,|\, v|_{\partial\mathcal{E}_D} = 0\} \cap H^2(\mathcal{E})$ and M its trace space, and define the extended function space $\boldsymbol{V}(h) := V(h) \times M(h)$ where $V(h) := V_h + V$ and $M(h) := M_h + M$. We will require the following three norms on $\boldsymbol{V}(h)$:

$$\begin{aligned}
|||\boldsymbol{v}|||_{v,h}^2 :=& \sum_{\mathcal{K}\in\mathcal{T}_h} \|v\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2} [\boldsymbol{v}] \right\|_{\partial\mathcal{K}}^2 + \sum_{F\in\partial\mathcal{E}_N} \left\| |\tfrac{1}{2}\beta\cdot n|^{1/2} \mu \right\|_F^2 && (2.7a)\\
&+ \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} v\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \|[\boldsymbol{v}]\|_{\mathcal{Q}_{\mathcal{K}}}^2, &&\\
|||\boldsymbol{v}|||_{s,h}^2 :=& |||\boldsymbol{v}|||_{v,h}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t v\|_{\mathcal{K}}^2, && (2.7b)\\
|||\boldsymbol{v}|||_{ss,h}^2 :=& |||\boldsymbol{v}|||_{s,h}^2 + \|v\|_{sd,h}^2 := |||\boldsymbol{v}|||_{s,h}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \|\Pi_h (\beta\cdot\nabla v)\|_{\mathcal{K}}^2, && (2.7c)
\end{aligned}$$

where Π_h denotes the L^2-projection onto V_h and the parameter τ_ε in the definition of $|||\boldsymbol{v}|||_{s,h}$ depends on the size of the space-time element compared to the diffusion parameter ε:

$$\tau_\varepsilon := \Delta t_{\mathcal{K}} \tilde{\varepsilon},$$

where

$$\tilde{\varepsilon} := \begin{cases} 1 & \text{if } \mathcal{K} \in \mathcal{T}_h^d := \left\{\mathcal{K} \in \mathcal{T}_h | \delta t_{\mathcal{K}} \le h_K \le \varepsilon\right\}, \\ \varepsilon^{1/2} & \text{if } \mathcal{K} \in \mathcal{T}_h^x := \left\{\mathcal{K} \in \mathcal{T}_h | \delta t_{\mathcal{K}} \le \varepsilon < h_K\right\}, \\ \varepsilon & \text{if } \mathcal{K} \in \mathcal{T}_h^c := \left\{\mathcal{K} \in \mathcal{T}_h | \varepsilon < \delta t_{\mathcal{K}} \le h_K\right\}. \end{cases}$$

Finally, $\beta_s := \sup_{(x,t)\in F} |\beta\cdot n|$, for $F \subset \partial\mathcal{K}$. It is useful to remark that

$$\inf_{(x,t)\in F} (\beta_s - \tfrac{1}{2}\beta\cdot n) \ge \tfrac{1}{2} \max_{(x,t)\in F} |\beta\cdot n| \qquad \forall F \in \partial\mathcal{K},\ \forall\mathcal{K}\in\mathcal{T}_h. \tag{2.8}$$

Let $u, v \in [L^2(U)]^r$ for $1 \le r \le d+1$. We will write $(u,v)_U = \int_U u\cdot v \,\mathrm{d}x$ if $U \subset \mathbb{R}^{d+1}$ and $\langle u,v\rangle_U = \int_U u\cdot v\,\mathrm{d}s$ if $U \subset \mathbb{R}^d$. Furthermore, we define $(u,v)_{\mathcal{T}_h} := \sum_{\mathcal{K}\in\mathcal{T}_h} (u,v)_{\mathcal{K}}$, $\langle u,v\rangle_{\partial\mathcal{T}_h} := \sum_{\mathcal{K}\in\mathcal{T}_h}\langle u,v\rangle_{\partial\mathcal{K}}$, $\langle u,v\rangle_{\mathcal{Q}_h} := \sum_{\mathcal{K}\in\mathcal{T}_h}\langle u,v\rangle_{\mathcal{Q}_{\mathcal{K}}}$, and $\langle u,v\rangle_{\partial\mathcal{E}_N} := \sum_{F\in\mathcal{F}_h^b\cap\partial\mathcal{E}_N}\langle u,v\rangle_F$.

The space-time HDG method for eq. (2.1) is given by: Find $\boldsymbol{u}_h \in \boldsymbol{V}_h$ such that

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = (f, v_h)_{\mathcal{T}_h} + \langle g, \mu_h\rangle_{\partial\mathcal{E}_N} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.9}$$

with $a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) := a_{h,d}(\boldsymbol{u}_h, \boldsymbol{v}_h) + a_{h,c}(\boldsymbol{u}_h, \boldsymbol{v}_h)$ and where

$$\begin{aligned}
a_{h,d}\left(\boldsymbol{u}, \boldsymbol{v}\right) &:= \left(\varepsilon \overline{\nabla} u, \overline{\nabla} v\right)_{\mathcal{T}_h} + \langle \varepsilon \alpha h_K^{-1} \left[\boldsymbol{u}\right], \left[\boldsymbol{v}\right] \rangle_{\mathcal{Q}_h} - \langle \varepsilon \left[\boldsymbol{u}\right], \overline{\nabla}_{\bar{n}} v \rangle_{\mathcal{Q}_h} - \langle \varepsilon \overline{\nabla}_{\bar{n}} u, \left[\boldsymbol{v}\right] \rangle_{\mathcal{Q}_h}, \\
a_{h,c}\left(\boldsymbol{u}, \boldsymbol{v}\right) &:= -\left(\beta u, \nabla v\right)_{\mathcal{T}_h} + \langle \zeta^+ \beta \cdot n \lambda, \mu \rangle_{\partial \mathcal{E}_N} + \langle \left(\beta \cdot n\right) \lambda + \beta_s \left[\boldsymbol{u}\right], \left[\boldsymbol{v}\right] \rangle_{\partial \mathcal{T}_h}.
\end{aligned}$$

Here, $\overline{\nabla}_{\bar{n}} v := \overline{\nabla} v \cdot \bar{n}$ denotes the directional derivative, $\beta_s > 0$ is a penalty parameter and ζ^+ denotes the outflow boundary indicator on a facet.

Chapter 3

Inequalities, approximations and projections

This chapter presents theoretical results that are useful for the a priori error analysis in chapter 4 and the a posteriori error analysis in chapter 5. Sections 3.1 to 3.3 list our scaling arguments, inverse inequalities, trace inequalities and local projection estimates in the space-time setting. For this we take into account the anisotropy between spatial and temporal variables. When the domain is fixed, which we assume for the a posteriori error analysis, special cases of these inequalities and additional useful inequalities are included at the end of each section.

Furthermore, sections 3.4 to 3.6 cover results that are only relevant to the a posteriori error analysis and hence a fixed domain is assumed. Specifically, section 3.4 presents useful local projection estimates; section 3.5 presents approximation estimates of an averaging operator (also known as the Oswald approximation operator); finally, section 3.6 presents several results related to a subgrid projection.

3.1 Scaling arguments

Following [53, Definition 2.9], we define

$$H^1(\mathcal{K}) := \{v \in L^2(\mathcal{K}) : \ (v \circ \phi_{\mathcal{K}}) \in H^1(\widetilde{\mathcal{K}})\} \, .$$

Consider an element $\widetilde{\mathcal{K}}$ and let $\widetilde{F}_{\mathcal{Q}} \subset \mathcal{Q}_{\widetilde{\mathcal{K}}}$ and $\widetilde{F}_{\mathcal{R}} \subset \mathcal{R}_{\widetilde{\mathcal{K}}}$. For $\widetilde{v} \in H^1(\widetilde{\mathcal{K}})$, the following scaling arguments between the reference domain and the affine domain can be shown based on [97, Lemma B.7]

$$\|\widehat{v}\|^2_{\widehat{\mathcal{K}}} = \left(\tfrac{1}{2}\right)^{-d-1} \delta t_{\mathcal{K}}^{-1} h_K^{-d} \, \|\widetilde{v}\|^2_{\widetilde{\mathcal{K}}} \,, \tag{3.1a}$$

$$\|\widehat{v}\|^2_{\widehat{F}_{\mathcal{Q}}} = \left(\tfrac{1}{2}\right)^{-d} \delta t_{\mathcal{K}}^{-1} h_K^{-d+1} \, \|\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{Q}}} \,, \tag{3.1b}$$

$$\|\widehat{v}\|^2_{\widehat{F}_{\mathcal{R}}} = \left(\tfrac{1}{2}\right)^{-d} h_K^{-d} \, \|\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{R}}} \,, \tag{3.1c}$$

$$\|\partial_{\widehat{x}_i} \widehat{v}\|^2_{\widehat{\mathcal{K}}} = \left(\tfrac{1}{2}\right)^{-d+1} \delta t_{\mathcal{K}}^{-1} h_K^{-d+2} \, \|\partial_{\widetilde{x}_i} \widetilde{v}\|^2_{\widetilde{\mathcal{K}}} \,, \quad \forall 1 \leq i \leq d, \tag{3.1d}$$

$$\|\partial_{\widehat{t}} \widehat{v}\|^2_{\widehat{\mathcal{K}}} = \left(\tfrac{1}{2}\right)^{-d+1} \delta t_{\mathcal{K}} h_K^{-d} \, \|\partial_{\widetilde{t}} \widetilde{v}\|^2_{\widetilde{\mathcal{K}}} \,. \tag{3.1e}$$

Furthermore, we have the following scaling arguments between the affine domain and the physical domain

$$c^{-1} \|v\|_{\mathcal{K}} \leq \|\widetilde{v}\|_{\widetilde{\mathcal{K}}} \leq c \, \|v\|_{\mathcal{K}} \,, \tag{3.2a}$$

$$c^{-1} \|v\|_{F_{\mathcal{Q}}} \leq \|\widetilde{v}\|_{\widetilde{F}_{\mathcal{Q}}} \leq c \, \|v\|_{F_{\mathcal{Q}}} \,, \tag{3.2b}$$

$$c^{-1} \|v\|_{F_{\mathcal{R}}} \leq \|\widetilde{v}\|_{\widetilde{F}_{\mathcal{R}}} \leq c \, \|v\|_{F_{\mathcal{R}}} \,, \tag{3.2c}$$

$$\|\partial_{\widetilde{x}_i} \widetilde{v}\|_{\widetilde{\mathcal{K}}} \leq c \, \|\overline{\nabla} v\|_{\mathcal{K}} \,, \quad \forall 1 \leq i \leq d, \tag{3.2d}$$

$$\|\partial_{\widetilde{t}} \widetilde{v}\|_{\widetilde{\mathcal{K}}} \leq c \left(\|\partial_t v\|_{\mathcal{K}} + \|\overline{\nabla} v\|_{\mathcal{K}} \right) . \tag{3.2e}$$

Here, eq. (3.2a) follows from a change of variables, eqs. (2.2) and (2.4), while eqs. (3.2b) and (3.2c) follow from a change of variables and eqs. (2.5) and (2.6).

Equation (3.2e) follows from the chain rule, eq. (2.2), change of variables and eq. (2.4):

$$\begin{aligned}
\|\partial_{\widetilde{t}} \widetilde{v}\|^2_{\widetilde{\mathcal{K}}} &= \int_{\widetilde{\mathcal{K}}} \left(\sum_{j=0}^{d} \partial_{x_j} v \circ \phi_{\mathcal{K}} \left(\tfrac{\partial x_j}{\partial \widetilde{t}} \right) \right)^2 \mathrm{d}\widetilde{t} \, \mathrm{d}\widetilde{x} \\
&\leq c \left(\int_{\widetilde{\mathcal{K}}} \left(\partial_t v \circ \phi_{\mathcal{K}} \right)^2 \mathrm{d}\widetilde{t} \, \mathrm{d}\widetilde{x} + \sum_{j=1}^{d} \int_{\widetilde{\mathcal{K}}} \left(\partial_{x_j} v \circ \phi_{\mathcal{K}} \right)^2 \mathrm{d}\widetilde{t} \, \mathrm{d}\widetilde{x} \right) \\
&\leq c \left(\int_{\mathcal{K}} \left(\partial_t v \right)^2 |\det J_{\phi_{\mathcal{K}}}^{-1}| \, \mathrm{d}t \, \mathrm{d}x + \sum_{j=1}^{d} \int_{\mathcal{K}} \left(\partial_{x_j} v \right)^2 |\det J_{\phi_{\mathcal{K}}}^{-1}| \, \mathrm{d}t \, \mathrm{d}x \right) \\
&\leq c \left(\|\partial_t v\|_{\mathcal{K}} + \|\overline{\nabla} v\|_{\mathcal{K}} \right)^2 .
\end{aligned}$$

Similar steps, combined with eq. (2.3a), are used to show eq. (3.2d).

Special cases and additional inequalities on a fixed domain: Following similar steps in showing eq. (3.2e), combined with eq. (2.3b), we have the following reduced version of eq. (3.2e)

$$\|\partial_{\widetilde{t}} \widetilde{v}\|_{\widetilde{\mathcal{K}}} \leq c \, \|\partial_t v\|_{\mathcal{K}} \,. \tag{3.3}$$

Furthermore, when $\widetilde{v}_h \in Q(\widetilde{\mathcal{K}})$ we have

$$\|\partial_{\hat{t}} \widehat{v}_h\|_{\widehat{F}_{\mathcal{R}}} = \left(\tfrac{1}{2}\right)^{1-d/2} \delta t_{\mathcal{K}} h_K^{-d/2} \, \|\partial_{\tilde{t}} \widetilde{v}_h\|_{\widetilde{F}_{\mathcal{R}}} \,, \tag{3.4a}$$

$$\|\partial_{\tilde{t}} \widetilde{v}_h\|_{\widetilde{F}_{\mathcal{R}}} \leq c \, \|\partial_t v_h\|_{F_{\mathcal{R}}} \,. \tag{3.4b}$$

Equation (3.4a) follows by extending [53, Lemma A.3] to $(d+1)$-dimensions in the space-time setting, while eq. (3.4b) follows from the chain rule, eq. (2.3b), a change of variables, and eqs. (2.5) and (2.6):

$$\begin{aligned}
\|\partial_{\tilde{t}} \widetilde{v}_h\|^2_{\widetilde{F}_{\mathcal{R}}} &= \int_{\widetilde{F}_{\mathcal{R}}} (\partial_{\tilde{t}} \widetilde{v}_h)^2 \, \mathrm{d}\widetilde{x} = \int_{\widetilde{F}_{\mathcal{R}}} \left(\sum_{j=0}^{d} \partial_{x_j} v_h \circ \phi_{\mathcal{K}} \left(\tfrac{\partial x_j}{\partial \tilde{t}}\right) \right)^2 \mathrm{d}\widetilde{x} \\
&= \int_{\widetilde{F}_{\mathcal{R}}} (\partial_t v_h \circ \phi_{\mathcal{K}})^2 \left(\tfrac{\partial t}{\partial \tilde{t}}\right)^2 \mathrm{d}\widetilde{x} = \int_{\widetilde{F}_{\mathcal{R}}} (\partial_t v_h \circ \phi_{\mathcal{K}})^2 \, \mathrm{d}\widetilde{x} \\
&= \int_{\widetilde{F}_{\mathcal{R}}} (\partial_t v_h \circ \phi_{F_{\mathcal{R}}})^2 \, \mathrm{d}\widetilde{x} \leq c \int_{\widetilde{F}_{\mathcal{R}}} (\partial_t v_h \circ \phi_{F_{\mathcal{R}}})^2 \left(\det\left((J^0_{\phi_{\mathcal{K}}})^{\intercal} J^0_{\phi_{\mathcal{K}}}\right)\right)^{1/2} \mathrm{d}\widetilde{x} \leq c \, \|\partial_t v_h\|^2_{F_{\mathcal{R}}} \,.
\end{aligned}$$

When we consider an edge $E_{\mathcal{K}}$ such that $\phi_{\mathcal{K}}(\widetilde{E}_{\mathcal{K}}) = E_{\mathcal{K}}$, we have

$$c^{-1} \, \|\widetilde{v}_h\|_{\widetilde{E}_{\mathcal{K}}} \leq \|v_h\|_{E_{\mathcal{K}}} \leq c \, \|\widetilde{v}_h\|_{\widetilde{E}_{\mathcal{K}}} \,. \tag{3.5}$$

Equation (3.5) is analogous to eq. (3.2b) in an integral domain with one lower dimension and can be shown with similar steps.

3.2 Anisotropic inverse and trace inequalities

Consider again an element $\widetilde{\mathcal{K}}$ and let $\widetilde{F}_{\mathcal{Q}} \subset \mathcal{Q}_{\widetilde{\mathcal{K}}}$ and $\widetilde{F}_{\mathcal{R}} \subset \mathcal{R}_{\widetilde{\mathcal{K}}}$. For any $\mathcal{K} \in \mathcal{T}_h$ and $v \in H^1(\mathcal{K})$, we have the following trace inequalities from [97, Lemma B.6]

$$\|\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{Q}}} \leq c \, (h_K^{-1} \, \|\widetilde{v}\|^2_{\widetilde{\mathcal{K}}} + \|\widetilde{v}\|_{\widetilde{\mathcal{K}}} \, \|\widetilde{\overline{\nabla}} \widetilde{v}\|_{\widetilde{\mathcal{K}}}) \,, \tag{3.6a}$$

$$\|\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{R}}} \leq c \, (\delta t_{\mathcal{K}}^{-1} \, \|\widetilde{v}\|^2_{\widetilde{\mathcal{K}}} + \|\widetilde{v}\|_{\widetilde{\mathcal{K}}} \, \|\partial_{\tilde{t}} \widetilde{v}\|_{\widetilde{\mathcal{K}}}) \,. \tag{3.6b}$$

Adapting [53, Corollaries 3.49, 3.54] to the space-time context, specifically taking into account the spatial mesh size h_K and time-step $\delta t_{\mathcal{K}}$ of a space-time element $\mathcal{K} \in \mathcal{T}_h$, we have the following anisotropic inverse and trace inequalities, which hold for all $v_h \in V_h$:

$$\|\partial_t v_h\|_{\mathcal{K}} \le c\,(\delta t_{\mathcal{K}}{}^{-1} + h_K^{-1})\,\|v_h\|_{\mathcal{K}}\,, \tag{3.7a}$$

$$\|\overline{\nabla} v_h\|_{\mathcal{K}} \le c h_K^{-1}\,\|v_h\|_{\mathcal{K}}\,, \tag{3.7b}$$

$$\|v_h\|_{\mathcal{Q}_{\mathcal{K}}} \le c_\star h_K^{-1/2}\,\|v_h\|_{\mathcal{K}}\,, \tag{3.7c}$$

$$\|v_h\|_{\partial\mathcal{K}} \le c\,(\delta t_{\mathcal{K}}{}^{-1/2} + h_K^{-1/2})\,\|v_h\|_{\mathcal{K}}\,, \tag{3.7d}$$

where $c_\star$ is a constant independent of h_K, $\delta t_{\mathcal{K}}$, ε, and T. (We distinguish $c_\star$ from c to prove lemma 4.1.) The following lemma introduces an additional inequality.

Lemma 3.1. *Let $\mathcal{K} \in \mathcal{T}_h$ be a space-time element. For all $\mu_h \in M_h$,*

$$\|\partial_t \mu_h\|_{F_{\mathcal{Q}}} \le c\,(\delta t_{\mathcal{K}}{}^{-1} + h_K^{-1})\,\|\mu_h\|_{F_{\mathcal{Q}}}\,. \tag{3.8}$$

Proof. The d-dimensional hypersurface $F_{\mathcal{Q}}$ is embedded in $\mathbb{R}^{d+1}$ and in general it may be curved. Therefore, we cannot use eq. (3.7a) directly to conclude eq. (3.8). Instead, we first map $F_{\mathcal{Q}}$ to the affine domain. For this, let $\phi_{F_{\mathcal{Q}}}(\widetilde{F}_{\mathcal{Q}}) = F_{\mathcal{Q}}$, i.e., the transformation of a face from the affine domain to the physical domain. We then observe that one of the spatial coordinates, which is denoted by $\widetilde{x}_j$ without loss of generality, of $\widetilde{F}_{\mathcal{Q}}$ is fixed. This means we can view $\widetilde{F}_{\mathcal{Q}}$ in the $\mathbb{R}^d$ domain with coordinates $(\widetilde{t}, \widetilde{x}_1, \dots, \widetilde{x}_{j-1}, \widetilde{x}_{j+1}, \dots, \widetilde{x}_d)$ and apply the d-dimensional versions of eq. (3.7a) and eq. (3.7b) to $\widetilde{F}_{\mathcal{Q}}$:

$$\|\partial_{\widetilde{t}} \widetilde{\mu}_h\|_{\widetilde{F}_{\mathcal{Q}}} \le c\,(\delta t_{\mathcal{K}}{}^{-1} + h_K^{-1})\,\|\widetilde{\mu}_h\|_{\widetilde{F}_{\mathcal{Q}}}\,, \quad \|\widetilde{\overline{\nabla}} \widetilde{\mu}_h\|_{\widetilde{F}_{\mathcal{Q}}} \le c h_K^{-1}\,\|\widetilde{\mu}_h\|_{\widetilde{F}_{\mathcal{Q}}}\,. \tag{3.9}$$

Using the mapping $\phi_{F_{\mathcal{Q}}}$, eqs. (2.5) and (2.6),

$$\begin{aligned}
\|\partial_t \mu_h\|_{F_{\mathcal{Q}}}^2 &= \int_{\widetilde{F}_{\mathcal{Q}}} \left[\left(\partial_t\left(\widetilde{\mu}_h \circ \phi_{F_{\mathcal{Q}}}^{-1}\right)\right) \circ \phi_{F_{\mathcal{Q}}}\right]^2 \left[\det\left((J_{\phi_{\mathcal{K}}}^j)^\intercal J_{\phi_{\mathcal{K}}}^j\right)\right]^{1/2} \mathrm{d}\widetilde{s} \\
&\le c \int_{\widetilde{F}_{\mathcal{Q}}} \left[\left(\partial_t\left(\widetilde{\mu}_h \circ \phi_{F_{\mathcal{Q}}}^{-1}\right)\right) \circ \phi_{F_{\mathcal{Q}}}\right]^2 \mathrm{d}\widetilde{s}.
\end{aligned}$$

By the chain rule,

$$\partial_t\left(\widetilde{\mu}_h \circ \phi_{F_{\mathcal{Q}}}^{-1}\right) = \left((\partial_{\widetilde{t}}\widetilde{\mu}_h) \circ \phi_{F_{\mathcal{Q}}}^{-1}\right) \frac{\partial \widetilde{t}}{\partial t} + \sum_{1 \le i \le d, i \ne j} \left((\partial_{\widetilde{x}_i}\widetilde{\mu}_h) \circ \phi_{F_{\mathcal{Q}}}^{-1}\right) \frac{\partial \widetilde{x}_i}{\partial t}.$$

We note that $\frac{\partial \widetilde{x}_i}{\partial t}$ is the $(i,0)$-element of $J_{\phi_{\mathcal{K}}}^{-1}$ which equals $(-1)^i \det J_{\phi_{\mathcal{K}\setminus i0}} / \det J_{\phi_{\mathcal{K}}}$. Similarly, $\frac{\partial \widetilde{t}}{\partial t}$ corresponds to the $(0,0)$-element of $J_{\phi_{\mathcal{K}}}^{-1}$ which equals $\det J_{\phi_{\mathcal{K}\setminus 00}} / \det J_{\phi_{\mathcal{K}}}$. Now using eq. (2.2), eq. (2.4), eq. (3.9), definition eq. (2.5), and eq. (2.6) we find that:

$$\begin{aligned} \|\partial_t \mu_h\|^2_{F_{\mathcal{Q}}} &\leq c \left(\int_{\widetilde{F}_{\mathcal{Q}}} (\partial_{\widetilde{t}} \widetilde{\mu}_h)^2 \, \mathrm{d}\widetilde{s} + \int_{\widetilde{F}_{\mathcal{Q}}} \left(\widetilde{\overline{\nabla}} \widetilde{\mu}_h\right)^2 \mathrm{d}\widetilde{s} \right) \\ &\leq c \left(\delta t_{\mathcal{K}}{}^{-2} + h_K^{-2}\right) \int_{\widetilde{F}_{\mathcal{Q}}} \widetilde{\mu}_h^2 \left[\det\left(\left(J^j_{\phi_{\mathcal{K}}}\right)^{\intercal} J^j_{\phi_{\mathcal{K}}}\right)\right]^{1/2} \mathrm{d}\widetilde{s} \leq c \left(\delta t_{\mathcal{K}}{}^{-2} + h_K^{-2}\right) \|\mu_h\|^2_{F_{\mathcal{Q}}}, \end{aligned}$$

which is eq. (3.8). □

Special cases and additional inequalities on a fixed domain: The following versions of eqs. (3.7a), (3.7d) and (3.8) adapted to fixed domains can be shown by considering eq. (2.3b)

$$\|\partial_t v_h\|_{\mathcal{K}} \leq c\delta t_{\mathcal{K}}^{-1} \|v_h\|_{\mathcal{K}} \quad \forall v_h \in V_h, \tag{3.10a}$$

$$\|v_h\|_{\mathcal{R}_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-1/2} \|v_h\|_{\mathcal{K}} \quad \forall v_h \in V_h, \tag{3.10b}$$

$$\|\partial_t \mu_h\|_{\mathcal{Q}_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-1} \|\mu_h\|_{\mathcal{Q}_{\mathcal{K}}} \quad \forall \mu_h \in M_h. \tag{3.10c}$$

The following lemma introduces additional inequalities for fixed domains:

Lemma 3.2. *Let $\mathcal{K} \in \mathcal{T}_h$ be a space-time element and $\mu_h \in M_h$. For any $F_{\mathcal{R}} \subset \mathcal{R}_{\mathcal{K}}$ and $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$, we have*

$$\|\overline{\nabla} \mu_h\|_{F_{\mathcal{R}}} \leq c h_K^{-1} \|\mu_h\|_{F_{\mathcal{R}}}, \tag{3.11a}$$

$$\|\mu_h\|_{E_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-1/2} \|\mu_h\|_{F_{\mathcal{Q}}} \quad \forall E_{\mathcal{K}} \subset F_{\mathcal{Q}}, \tag{3.11b}$$

$$\|\mu_h\|_{E_{\mathcal{K}}} \leq c h_K^{-1/2} \|\mu_h\|_{F_{\mathcal{R}}} \quad \forall E_{\mathcal{K}} \subset F_{\mathcal{R}}. \tag{3.11c}$$

Proof. Equation (3.11a) is a result of applying eq. (3.7b) on $F_{\mathcal{R}}$ while eq. (3.11c) is a direct application of a standard isotropic trace inequality on $E_{\mathcal{K}}$ (see, for example, [45, Lemma 12.8]). As for eq. (3.11b), consider $\phi_{F_{\mathcal{Q}}}(\widetilde{E}_{\mathcal{K}}) = E_{\mathcal{K}}$. Applying eq. (3.7d) on the affine domain gives us

$$\|\widetilde{\mu}_h\|_{\widetilde{E}_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-1/2} \|\widetilde{\mu}_h\|_{\widetilde{F}_{\mathcal{Q}}}.$$

We conclude eq. (3.11b) via scaling arguments eqs. (3.2b) and (3.5). □

3.3 Anisotropic projection estimates

Let $\Pi_h^{\mathcal{F}}$ be the L^2-projection onto M_h. It can be shown that for $v \in H^1(\mathcal{K})$, we have

$$\|v - \Pi_h v\|_{\mathcal{K}} \le c\left(\delta t_{\mathcal{K}} \|\partial_t v\|_{\mathcal{K}} + h_K \|\overline{\nabla} v\|_{\mathcal{K}}\right), \tag{3.12a}$$

$$\|\overline{\nabla}\left(v - \Pi_h v\right)\|_{\mathcal{K}} \le c\|\overline{\nabla} v\|_{\mathcal{K}}, \tag{3.12b}$$

$$\|\partial_t\left(v - \Pi_h v\right)\|_{\mathcal{K}} \le c\left(\|\partial_t v\|_{\mathcal{K}} + \|\overline{\nabla} v\|_{\mathcal{K}}\right), \tag{3.12c}$$

$$\|\Pi_h v - \Pi_h^{\mathcal{F}} v\|_{\mathcal{Q}_{\mathcal{K}}} \le c h_K^{1/2}\|\overline{\nabla} v\|_{\mathcal{K}}. \tag{3.12d}$$

Proof. To begin, we note that the projection operator on $\widehat{\mathcal{K}}$ and $\widetilde{\mathcal{K}}$ are related to the projection operator on $\mathcal{K}$ as follows (see [53, Definition 3.12]):

$$\widetilde{\Pi}_h \widetilde{v} = \left(\Pi_h\left(\widetilde{v}\circ\phi_{\mathcal{K}}^{-1}\right)\right)\circ\phi_{\mathcal{K}} \quad \forall \widetilde{v}\in L^2(\widetilde{\mathcal{K}}), \qquad \widehat{\Pi}\widehat{v} = \left(\widetilde{\Pi}_h\left(\widehat{v}\circ G_{\mathcal{K}}^{-1}\right)\right)\circ G_{\mathcal{K}} \quad \forall\widehat{v}\in L^2(\widehat{\mathcal{K}}).$$

Similarly, on any $F \in \partial\mathcal{K}$, we have the following relations:

$$\widetilde{\Pi}_h^{\mathcal{F}}\widetilde{v} = \left(\Pi_h^{\mathcal{F}}\left(\widetilde{v}\circ\phi_F^{-1}\right)\right)\circ\phi_F \ \forall\widetilde{v}\in L^2(\widetilde{F}), \quad \widehat{\Pi}^{\mathcal{F}}\widehat{v} = \left(\widetilde{\Pi}_h^{\mathcal{F}}\left(\widehat{v}\circ G_F^{-1}\right)\right)\circ G_F \ \forall \widehat{v}\in L^2(\widehat{F}),$$

where G_F and ϕ_F are the restrictions of $G_{\mathcal{K}}$ and $\phi_{\mathcal{K}}$ on F, respectively.

Equation (3.12a) is the standard anisotropic projection estimate (see, for example, [53, Lemma 3.13]) and hence we omit its proof here. To show eq. (3.12b) we first note that the following projection estimate holds on $\widehat{\mathcal{K}}$ for any $1 \le i \le d$ (see [53, Lemma 3.7, eq. (3.12)]):

$$\|\partial_{\widehat{x}_i}\left(\widehat{v}-\widehat{\Pi}\widehat{v}\right)\|_{\widehat{\mathcal{K}}} \le c\|\partial_{\widehat{x}_i}\widehat{v}\|_{\widehat{\mathcal{K}}}. \tag{3.13}$$

By the chain rule $\partial_{\widetilde{x}_i}\left(\widetilde{v}-\widetilde{\Pi}_h\widetilde{v}\right) = 2h_K^{-1}\partial_{\widehat{x}_i}\left(\left(\widetilde{v}-\widetilde{\Pi}_h\widetilde{v}\right)\circ G_{\mathcal{K}}\right)\circ G_{\mathcal{K}}^{-1}$, using that $|\det G_{\mathcal{K}}| = \delta t_{\mathcal{K}} h_K^d 2^{-d-1}$, and using eqs. (3.1d) and (3.13), we find

$$\|\partial_{\widetilde{x}_i}\left(\widetilde{v}-\widetilde{\Pi}_h\widetilde{v}\right)\|_{\widetilde{\mathcal{K}}}^2 \le c\delta t_{\mathcal{K}} h_K^{d-2}\|\partial_{\widehat{x}_i}\widehat{v}\|^2 \le c\|\partial_{\widetilde{x}_i}\widetilde{v}\|_{\widetilde{\mathcal{K}}}^2. \tag{3.14}$$

To obtain the result on the physical element, consider first that by the chain rule,

$$\partial_{x_i}(v-\Pi_h v) = \sum_{1\le j\le d}\left(\partial_{\widetilde{x}_j}\left(\left(v-\Pi_h v\right)\circ\phi_{\mathcal{K}}\right)\circ\phi_{\mathcal{K}}^{-1}\right)\left((-1)^{i+j}(\det J_{\phi_{\mathcal{K}}})^{-1}\det J_{\phi_{\mathcal{K}}\setminus ij}\right),$$

where we used that $\widetilde{t}$ only depends on t in $\phi_{\mathcal{K}}^{-1}$ and that $\frac{\partial\widetilde{x}_j}{\partial x_i} = (-1)^{i+j}(\det J_{\phi_{\mathcal{K}}})^{-1}\det J_{\phi_{\mathcal{K}}\setminus ij}$. By assumptions eq. (2.2) and eq. (2.4), and using eq. (3.14), we therefore find that:

$$\|\partial_{x_i}\left(v-\Pi_h v\right)\|_{\mathcal{K}}^2 \le c\sum_{1\le j\le d}\|\partial_{\widetilde{x}_i}\left(\widetilde{v}-\widetilde{\Pi}_h\widetilde{v}\right)\|_{\widetilde{\mathcal{K}}}^2 \le c\sum_{1\le j\le d}\|\partial_{\widetilde{x}_j}\widetilde{v}\|_{\widetilde{\mathcal{K}}}^2 \le c\|\overline{\nabla}v\|_{\mathcal{K}}^2,$$

where the last step uses eq. (3.2d). The proof for eq. (3.12c) is similar and therefore omitted.

For eq. (3.12d), we consider a d-dimensional hypersurface $F_{\mathcal{Q}} \in \mathcal{Q}_{\mathcal{K}}$. We first map $F_{\mathcal{Q}}$ to the reference domain. For this, let $\phi_{F_{\mathcal{Q}}} \circ G_{F_{\mathcal{Q}}}(\widehat{F}_{\mathcal{Q}}) = F_{\mathcal{Q}}$, i.e., the transformation of a face from the reference domain to the physical domain. We then observe that one of the spatial coordinates, which is denoted by $\widehat{x}_j$ without loss of generality, of $\widehat{F}_{\mathcal{Q}}$ is fixed. We further consider a decomposition of $\widehat{\Pi} = \widehat{\pi}_{\widehat{t}} \Pi_{1 \leq i \leq d} \widehat{\pi}_{\widehat{x}_i}$ where $\widehat{\pi}_{\widehat{t}}$ and $\widehat{\pi}_{\widehat{x}_i}$ are the one-dimensional L^2-projection operators applied in the time direction and in the spatial direction $\widehat{x}_i$, respectively. Similarly, $(\widehat{\Pi}^{\mathcal{F}})\,|_{\widehat{F}_{\mathcal{Q}}} = \widehat{\pi}_{\widehat{t}} \Pi_{1 \leq i \leq d, i \neq j} \widehat{\pi}_{\widehat{x}_i}$. By [53, Definitions 3.1, 3.6], we have:

$$\|\widehat{\Pi}\widehat{v} - \widehat{\Pi}^{\mathcal{F}}\widehat{v}\|_{\widehat{F}_{\mathcal{Q}}} = \|\widehat{\pi}_{\widehat{t}} \Pi_{1 \leq i \leq d, i \neq j} \widehat{\pi}_{\widehat{x}_i} \left(\widehat{v} - \widehat{\pi}_{\widehat{x}_j} \widehat{v}\right)\|_{\widehat{F}_{\mathcal{Q}}} \leq c \,\|\widehat{v} - \widehat{\pi}_{\widehat{x}_j}\widehat{v}\|_{\widehat{F}_{\mathcal{Q}}} \leq c \,\|\partial_{\widehat{x}_j}\widehat{v}\|_{\widehat{\mathcal{K}}}\,, \tag{3.15}$$

where the equality is by commutativity of $\widehat{\pi}_{\widehat{x}_i}$ and $\widehat{\pi}_{\widehat{x}_j}$ $(i \neq j)$ and the last two inequalities are due to the boundedness of any composition of projections $\widehat{\pi}$ and [53, Lemma 3.3]. Next, using eq. (3.1b), eq. (3.15), and eq. (3.1d),

$$\|\widetilde{\Pi}_h\widetilde{v} - \widetilde{\Pi}_h^{\mathcal{F}}\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{Q}_i}} \leq c h_K \,\|\partial_{\widetilde{x}_j}\widetilde{v}\|^2_{\widetilde{\mathcal{K}}}\,.$$

Therefore, also using eq. (2.5) and eq. (2.6),

$$\|\Pi_h v - \Pi_h^{\mathcal{F}} v\|^2_{F_{\mathcal{Q}}} \leq c \,\|\widetilde{\Pi}_h\widetilde{v} - \widetilde{\Pi}_h^{\mathcal{F}}\widetilde{v}\|^2_{\widetilde{F}_{\mathcal{Q}}} \leq c h_K \,\|\partial_{\widetilde{x}_j}\widetilde{v}\|^2_{\widetilde{\mathcal{K}}} \leq c h_K \,\|\overline{\nabla} v\|^2_{\mathcal{K}}\,,$$

where we reverse the scaling arguments in the final inequality, proving eq. (3.12d). ☐

Special cases and additional inequalities on a fixed domain: It can be shown that by using eq. (2.3b) in the proof of eq. (3.12c), eq. (3.12c) reduces to the following

$$\|\partial_t(v - \Pi_h v)\|_{\mathcal{K}} \leq c \,\|\partial_t v\|_{\mathcal{K}}\,. \tag{3.16}$$

Additionally, the following projection estimate can be shown similarly as eq. (3.12d):

$$\|(\Pi_h - \Pi_h^{\mathcal{F}})\, v\|_{F_{\mathcal{R}}} \leq c \delta t_{\mathcal{K}}^{1/2} \,\|\partial_t v\|_{\mathcal{K}}\,. \tag{3.17}$$

3.4 Projection estimates for the a posteriori error analysis

The following lemma presents local projection estimates that will be useful in showing reliability of the error estimator eq. (5.1).

Lemma 3.3. *Let $v \in H^1(\mathcal{K})$ and consider Π_h, the L^2-projection onto V_h. For any $\mathcal{K} \in \mathcal{T}_h$ and any $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$, $F_{\mathcal{R}} \subset \mathcal{R}_{\mathcal{K}}$, assuming that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, the following projection estimates can be shown*

$$\left\| v - \Pi_h v \right\|_{\mathcal{K}} \leq c\lambda_{\mathcal{K}} \left(h_K \varepsilon^{1/2} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} + \left\| v \right\|_{\mathcal{K}} \right), \tag{3.18a}$$

$$\left\| v - \Pi_h v \right\|_{F_{\mathcal{Q}}} \leq c h_K^{1/2} \varepsilon^{-1/2} \left(h_K \varepsilon^{1/2} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right), \tag{3.18b}$$

$$\left\| v - \Pi_h v \right\|_{F_{\mathcal{R}}} \leq c \varepsilon^{-1/2} \left(h_K \varepsilon^{1/2} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right), \tag{3.18c}$$

where $\lambda_{\mathcal{K}} := \min \{1, h_K \varepsilon^{-1/2}\}$.

Proof. Consider the following local trace inequality, which holds for all $\mathcal{K} \in \mathcal{T}_h$, $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$ and $v \in H^1(\mathcal{K})$ and it can be shown by combining eq. (3.6) and scaling arguments eqs. (3.2b) to (3.2d) and (3.3)

$$\left\| v \right\|_{F_{\mathcal{Q}}}^2 \leq c \left(h_K^{-1} \left\| v \right\|_{\mathcal{K}}^2 + \left\| v \right\|_{\mathcal{K}} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right), \tag{3.19a}$$

$$\left\| v \right\|_{F_{\mathcal{R}}}^2 \leq c \left(\delta t_{\mathcal{K}}^{-1} \left\| v \right\|_{\mathcal{K}}^2 + \left\| v \right\|_{\mathcal{K}} \left\| \partial_t v \right\|_{\mathcal{K}} \right). \tag{3.19b}$$

Using eq. (3.12a), we have

$$\begin{aligned} \left\| v - \Pi_h v \right\|_{\mathcal{K}} &\leq c h_K \varepsilon^{-1/2} \left(\varepsilon^{1/2} h_K^{-1} \delta t_{\mathcal{K}} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right) \\ &\leq c h_K \varepsilon^{-1/2} \left(\varepsilon^{1/2} h_K^{-1} \delta t_{\mathcal{K}} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} + \left\| v \right\|_{\mathcal{K}} \right), \end{aligned} \tag{3.20}$$

while boundedness of the projection operator Π_h gives

$$\left\| v - \Pi_h v \right\|_{\mathcal{K}} \leq c \left\| v \right\|_{\mathcal{K}} \leq c \left(\varepsilon^{1/2} h_K^{-1} \delta t_{\mathcal{K}} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} + \left\| v \right\|_{\mathcal{K}} \right). \tag{3.21}$$

Combining eqs. (3.20) and (3.21) yields

$$\left\| v - \Pi_h v \right\|_{\mathcal{K}} \leq c\lambda_{\mathcal{K}} \left(\varepsilon^{1/2} h_K^{-1} \delta t_{\mathcal{K}} \left\| \partial_t v \right\|_{\mathcal{K}} + \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} + \left\| v \right\|_{\mathcal{K}} \right). \tag{3.22}$$

Furthermore, combining the trace inequalities eq. (3.19) with the projection bounds eqs. (3.12b) and (3.16), we obtain:

$$\left\| v - \Pi_h v \right\|_{F_{\mathcal{Q}}} \leq c \varepsilon^{-1/2} \left(\delta t_{\mathcal{K}} h_K^{-1/2} \varepsilon^{1/2} \left\| \partial_t v \right\|_{\mathcal{K}} + h_K^{1/2} \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right), \tag{3.23a}$$

$$\left\| v - \Pi_h v \right\|_{F_{\mathcal{R}}} \leq c \varepsilon^{-1/2} \left(\delta t_{\mathcal{K}}^{1/2} \varepsilon^{1/2} \left\| \partial_t v \right\|_{\mathcal{K}} + h_K \delta t_{\mathcal{K}}^{-1/2} \varepsilon^{1/2} \left\| \overline{\nabla} v \right\|_{\mathcal{K}} \right). \tag{3.23b}$$

Lemma 3.3 is now an immediate consequence of eq. (3.23), eq. (3.22), and using $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. □

3.5 Approximation estimates of an averaging operator

We define an averaging operator $\mathcal{I}_h^c : V_h \to V_h \cap C^0(\mathcal{E})$. For any $v_h \in V_h$, we first construct the coarsest conforming refinement $\mathcal{T}_h^c$ of $\mathcal{T}_h$; the operator $\mathcal{I}_h^c v_h$ is prescribed at vertices of $\mathcal{T}_h^c$ by the average of the values of v_h at the vertex (see [33, Lemma 3.4] and [45, Section 22.2]). For the Dirichlet boundary nodes, i.e. nodes on $\partial\mathcal{E}_D$, $\mathcal{I}_h^c v_h$ is prescribed by zero. Furthermore, given a space-time element $\mathcal{K} \in \mathcal{T}_h$, we introduce $\check{\mathcal{Q}}_{\mathcal{K}}^i$ to denote the union of $\mathcal{Q}$-facets in $\mathcal{F}_h^i$ that have a non-empty intersection with $\partial\mathcal{K}$. Similarly we introduce $\check{\mathcal{R}}_{\mathcal{K}}^i$.

Lemma 3.4. *For a space-time element $\mathcal{K} \in \mathcal{T}_h$, the averaging operator $\mathcal{I}_h^c : V_h \to V_h \cap C^0(\mathcal{E})$ satisfies the following*

$$\|v_h - \mathcal{I}_h^c v_h\|_{\mathcal{K}} \leq c \Big(\sum_{F \in \check{\mathcal{Q}}_{\mathcal{K}}^i} h_K^{1/2} \big\| [\![v_h]\!] \big\|_F + \sum_{F \in \check{\mathcal{R}}_{\mathcal{K}}^i} \delta t_{\mathcal{K}}^{1/2} \big\| [\![v_h]\!] \big\|_F \Big) . \tag{3.24}$$

Proof. The proof below combines an estimate for the averaging operator on conforming meshes (extended from [45, Lemma 22.3] to space-time meshes), and an auxiliary mesh technique (see, for example, [63, 75, 108, 109]).

We start by proving eq. (3.24) on a conforming $(d+1)$-dimensional space-time mesh. Within this conforming mesh, consider $\sigma_{\mathcal{K}}$ (see fig. 2.1) which consists of a space-time element $\mathcal{K}$ and $\mathcal{K}_i$, $i = 1, \ldots, 3^{(d+1)} - 1$. We map $\sigma_{\mathcal{K}}$ to the reference domain while preserving connectivity relations between the elements. This is achieved by combining $\Phi_{\mathcal{K}}$ with $\Phi_{\mathcal{K}_i}$ $(i = 1, \ldots, 3^{(d+1)} - 1)$, where $\Phi_{\mathcal{K}_i}$ are $\Phi_{\mathcal{K}}$ with suitable linear translations.

Applying [45, Lemma 22.3] to $\sigma_{\mathcal{K}}$ in the reference domain,

$$\|\widehat{v}_h - \mathcal{I}_h^c \widehat{v}_h\|_{\widehat{\mathcal{K}}} \leq c \big(\sum_{\widehat{F} \subset \check{\mathcal{Q}}_{\widehat{\mathcal{K}}}^i} \|[\![\widehat{v}_h]\!]\|_{\widehat{F}} + \sum_{\widehat{F} \subset \check{\mathcal{R}}_{\widehat{\mathcal{K}}}^i} \|[\![\widehat{v}_h]\!]\|_{\widehat{F}} \big) . \tag{3.25}$$

We remark that in the proof of [45, Lemma 22.3], the only intermediate result that restricts the domain dimension to be lower than or equal to three is [45, Lemma 21.4]. We argue that [45, Lemma 21.4] can be extended to the space-time domain $\mathcal{E} \subset \mathbb{R}^{d+1}$ due to it being Lipschitz. With scaling arguments eqs. (3.1a) to (3.1c) and (3.2), eq. (3.25) is transformed back to the physical domain:

$$\|v_h - \mathcal{I}_h^c v_h\|_{\mathcal{K}} \leq c \big(\sum_{F \subset \check{\mathcal{Q}}_{\mathcal{K}}^i} h_K^{1/2} \|[\![v_h]\!]\|_F + \sum_{F \subset \check{\mathcal{R}}_{\mathcal{K}}^i} \delta t_{\mathcal{K}}^{1/2} \|[\![v_h]\!]\|_F \big) . \tag{3.26}$$

We now consider the case of a 1-irregular mesh. Let $\mathcal{K} \in \mathcal{T}_h$ and let $\mathcal{T}_h^c$ be the coarsest refinement of $\mathcal{T}_h$. We consider two cases: (1) $\mathcal{K}$ is not refined on $\mathcal{T}_h^c$; and (2) $\mathcal{K}$ is refined on $\mathcal{T}_h^c$. In fig. 3.1, we provide examples of both cases to illustrate the geometric objects involved in the following proof.

Case 1. If $\mathcal{K}$ is not refined on $\mathcal{T}_h^c$, we denote by $\sigma_{\mathcal{K}}^c$ the local patch of elements associated with $\mathcal{K}$ on $\mathcal{T}_h^c$. Applying eq. (3.26) on $\sigma_{\mathcal{K}}^c$ gives

$$\|v_h - \mathcal{I}_h^c v_h\|_{\mathcal{K}} \le c\,\Big(\sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}}^{i,c}\setminus\check{\mathfrak{Q}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F + \sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}}^{i,c}\setminus\check{\mathfrak{R}}_{\mathcal{K}}^{i}} \delta t_{\mathcal{K}}^{1/2}\,\|[\![v_h]\!]\|_F\Big)\,, \tag{3.27}$$

where $\check{\mathcal{Q}}_{\mathcal{K}}^{i,c}$ and $\check{\mathcal{R}}_{\mathcal{K}}^{i,c}$ are defined similarly as $\check{\mathcal{Q}}_{\mathcal{K}}^{i}$ and $\check{\mathcal{R}}_{\mathcal{K}}^{i}$, but for $\mathcal{K}$ on $\mathcal{T}_h^c$, and where $\check{\mathfrak{Q}}_{\mathcal{K}}^{i}$ and $\check{\mathfrak{R}}_{\mathcal{K}}^{i}$ are unions of newly generated $\mathcal{Q}$-faces and $\mathcal{R}$-faces that divide an element in $\mathcal{T}_h$ to create $\mathcal{T}_h^c$. Note that $[\![v_h]\!]$ vanishes on $\check{\mathfrak{Q}}_{\mathcal{K}}^{i}$ and $\check{\mathfrak{R}}_{\mathcal{K}}^{i}$, explaining why they are excluded from the summation in eq. (3.27). Equation (3.24) then follows from eq. (3.27) by noting that

$$\begin{aligned}
\sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}}^{i,c}\setminus\check{\mathfrak{Q}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F &\le c \sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F\,,\\
\sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}}^{i,c}\setminus\check{\mathfrak{R}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F &\le c \sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F\,.
\end{aligned}$$

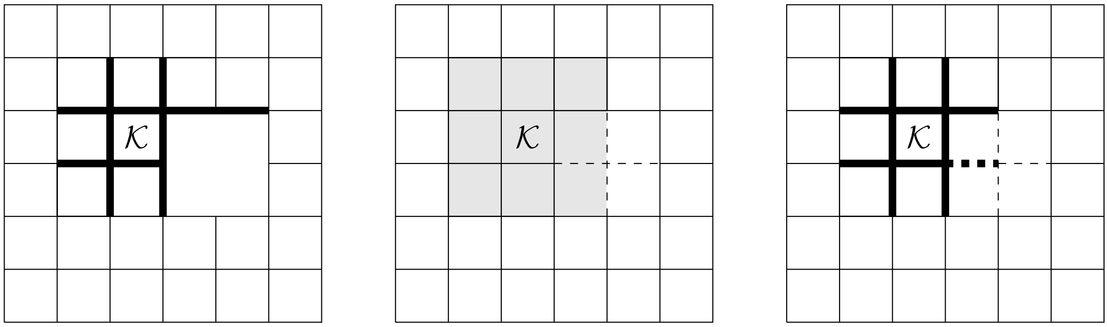
Case 2. When $\mathcal{K}$ is refined on $\mathcal{T}_h^c$ into $M_{\mathcal{K}}$ elements, $\mathcal{K} = \cup_{j=1}^{M_{\mathcal{K}}} \mathcal{K}_j$, where $M_{\mathcal{K}} \le 2^{d+1}$. We apply eq. (3.26) on each $\sigma_{\mathcal{K}_j}^c$ resulting in

$$\|v_h - \mathcal{I}_h^c v_h\|_{\mathcal{K}_j} \le c\,\Big(\sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}_j}^{i,c}\setminus\check{\mathfrak{Q}}_{\mathcal{K}_j}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F + \sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}_j}^{i,c}\setminus\check{\mathfrak{R}}_{\mathcal{K}_j}^{i}} \delta t_{\mathcal{K}}^{1/2}\,\|[\![v_h]\!]\|_F\Big)\,. \tag{3.28}$$

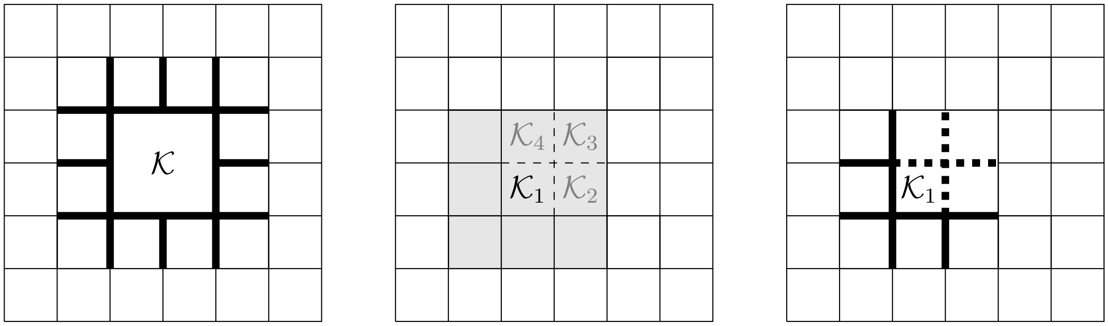
Combining eq. (3.28) for all $j = 1, \ldots, M_{\mathcal{K}}$ gives eq. (3.24) by noting that

$$\begin{aligned}
\sum_{j=1}^{M_{\mathcal{K}}} \sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}_j}^{i,c}\setminus\check{\mathfrak{Q}}_{\mathcal{K}_j}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F &\le c \sum_{F\subset \check{\mathcal{Q}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F\,,\\
\sum_{j=1}^{M_{\mathcal{K}}} \sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}_j}^{i,c}\setminus\check{\mathfrak{R}}_{\mathcal{K}_j}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F &\le c \sum_{F\subset \check{\mathcal{R}}_{\mathcal{K}}^{i}} h_K^{1/2}\,\|[\![v_h]\!]\|_F\,.
\end{aligned}$$

□



(a) An example of Case 1 in the proof of lemma 3.4



(b) An example of Case 2 in the proof of lemma 3.4

Figure 3.1: Examples of Case 1 and 2 in the proof of lemma 3.4.
Left column: the space-time element $\mathcal{K}$ on the 1-irregular mesh $\mathcal{T}_h$. Thick solid lines are the union of facets in $\check{\mathcal{R}}^i_{\mathcal{K}}$ and $\check{\mathcal{Q}}^i_{\mathcal{K}}$.
Centre column: Coarsest refinement of $\mathcal{T}_h$ (in dashed lines) is applied to construct $\mathcal{T}^c_h$. In fig. 3.1a, $\mathcal{K}$ is not refined in $\mathcal{T}^c_h$. Elements in $\sigma^c_{\mathcal{K}}$ are colored in grey. In fig. 3.1b, $\mathcal{K}$ is refined to $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \mathcal{K}_4$ (i.e., $M_{\mathcal{K}} = 4$). We only highlight $\mathcal{K}_1$ for this illustration and elements in $\sigma^c_{\mathcal{K}_1}$ are colored in grey.
Right column: thick lines (solid and dashed) are the union of facets in $\check{\mathcal{R}}^{i,c}_{\mathcal{K}}$ and $\check{\mathcal{Q}}^{i,c}_{\mathcal{K}}$. In fig. 3.1a, thick dashed lines are the union in facets of $\check{\mathfrak{R}}^i_{\mathcal{K}}$ and $\check{\mathfrak{Q}}^i_{\mathcal{K}}$. In fig. 3.1b, thick dashed lines are the union of facets of $\check{\mathfrak{R}}^i_{\mathcal{K}_1}$ and $\check{\mathfrak{Q}}^i_{\mathcal{K}_1}$.

3.6 Subgrid projection estimates

Let $\mathcal{T}_{\mathfrak{h}}$ be the subgrid obtained by halving the time-step of each element in $\mathcal{T}_h$. We now define a set of objects within the subgrid that will be useful in the ensuing analysis. As a rule of thumb, objects that are associated with the upper half time-step are denoted with an asterisk superscript while objects that are associated with the lower half time-step are denoted with an asterisk subscript.

For each $\mathcal{K} \in \mathcal{T}_h$, we introduce $\mathring{\mathcal{K}}^*$ and $\mathring{\mathcal{K}}_*$ to denote the two resulting space-time elements in $\mathcal{T}_{\mathfrak{h}}$, i.e., $\mathcal{K} = \mathring{\mathcal{K}}^* \cup \mathring{\mathcal{K}}_*$, and write $\mathcal{T}_{\mathcal{K}} := \{\mathring{\mathcal{K}}^*, \mathring{\mathcal{K}}_*\}$. Furthermore, every $\mathcal{Q}$-facet $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$ is divided into $F_{\mathcal{Q}}^*$ and $F_{\mathcal{Q},*}$. We define $F_{\mathring{\mathcal{R}}} := \partial\mathring{\mathcal{K}}^* \cap \partial\mathring{\mathcal{K}}_*$ and introduce $E_{\mathring{\mathcal{K}}}$ to denote any edge of $F_{\mathring{\mathcal{R}}}$. Finally, for any $v_{\mathfrak{h}} \in V_{\mathfrak{h}}$, when considering a $\mathcal{K} \in \mathcal{T}_h$ with $\mathcal{K} = \mathring{\mathcal{K}}^* \cup \mathring{\mathcal{K}}_*$, we let $v_{\mathfrak{h}}^*$ and $v_{\mathfrak{h},*}$ denote $v_{\mathfrak{h}}|_{\mathring{\mathcal{K}}^*}$ and $v_{\mathfrak{h}}|_{\mathring{\mathcal{K}}_*}$, respectively. See figs. 3.2 and 3.3 for illustrations in $(1+1)$, $(2+1)$ and $(3+1)$-dimensional space-time domains respectively.

The following trace inequalities are obtained by applying eq. (3.11b) and eq. (3.11c).

Lemma 3.5. *On the subgrid $\mathcal{T}_{\mathfrak{h}}$, the following trace inequalities hold (where, for each inequality, it is implicitly assumed that $E_{\mathring{\mathcal{K}}}$ is an edge of the facet on the right-hand side):*

$$\begin{aligned} \|v_{\mathfrak{h}}^*\|_{E_{\mathring{\mathcal{K}}}} &\le c\delta t_{\mathcal{K}}^{-1/2} \|v_{\mathfrak{h}}^*\|_{F_{\mathcal{Q}}^*}, \quad \|v_{\mathfrak{h},*}\|_{E_{\mathring{\mathcal{K}}}} \le c\delta t_{\mathcal{K}}^{-1/2} \|v_{\mathfrak{h},*}\|_{F_{\mathcal{Q},*}}, \\ \|v_{\mathfrak{h}}^*\|_{E_{\mathring{\mathcal{K}}}} &\le ch_K^{-1/2} \|v_{\mathfrak{h}}^*\|_{F_{\mathring{\mathcal{R}}}}, \quad \|v_{\mathfrak{h},*}\|_{E_{\mathring{\mathcal{K}}}} \le ch_K^{-1/2} \|v_{\mathfrak{h},*}\|_{F_{\mathring{\mathcal{R}}}}. \end{aligned} \tag{3.29}$$

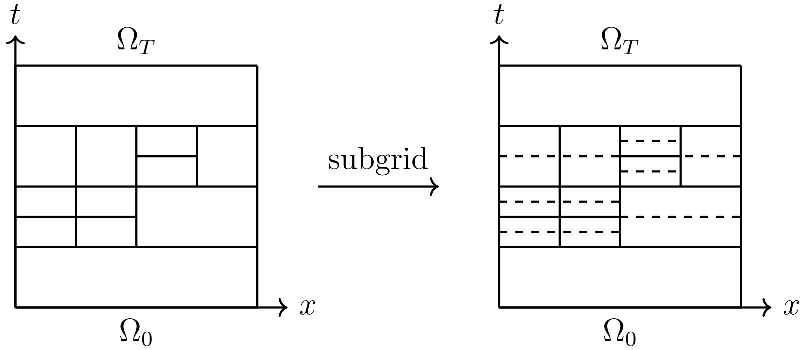
Definition 3.1. *We define the following restriction operator:*

$$\begin{aligned} &\gamma_{\mathfrak{h}} : \boldsymbol{V}_h \to \boldsymbol{V}_{\mathfrak{h}} \ : \ (v_h, \mu_h) \mapsto \left(v_h, \gamma_{\mathcal{F},\mathfrak{h}}(\boldsymbol{v}_h)\right), \\ &\gamma_{\mathcal{F},\mathfrak{h}}(\boldsymbol{v}_h) := \begin{cases} \mu_h, & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ v_h, & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}. \end{cases} \end{aligned} \tag{3.30}$$

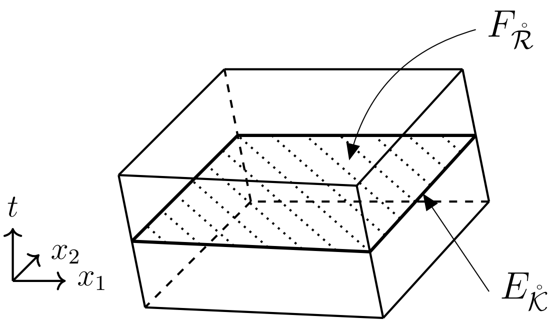
Furthermore, let $i_h^{\mathcal{K}}(v_{\mathfrak{h}})$ denote the L^2-projection of $v_{\mathfrak{h}}$ onto V_h, and let $i_h^{\mathcal{F}}(\mu_{\mathfrak{h}})$ be defined as follows. For any facet $F \in \mathcal{F}_h$, if $F \in \mathcal{F}_{\mathfrak{h}}$, $(i_h^{\mathcal{F}}(\mu_{\mathfrak{h}}))|_F := (\mu_{\mathfrak{h}})|_F$; else, $(i_h^{\mathcal{F}}(\mu_{\mathfrak{h}}))|_F$ is the L^2-projection of $\mu_{\mathfrak{h}}$ onto M_h. If $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$ we define $(i_h^{\mathcal{F}}(\mu_{\mathfrak{h}}))|_F := (\mu_{\mathfrak{h}})|_F$. See fig. 3.4 for an illustration of how $i_h^{\mathcal{F}}$ projects onto interior $\mathcal{Q}$-facets in $\mathcal{F}_{\mathcal{Q},h}$. We then define the projection operator:

$$i_h : \boldsymbol{V}_{\mathfrak{h}} \to \boldsymbol{V}_{\mathfrak{h}} \ : \ (v_{\mathfrak{h}}, \mu_{\mathfrak{h}}) \mapsto \gamma_{\mathfrak{h}}\left(i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{F}}(\mu_{\mathfrak{h}})\right).$$

We have the following projection estimates.



(a) $(1 + 1)$-dimensional example



(b) $(2 + 1)$-dimensional example

Figure 3.2: In the left-hand side of the figure we show a $(1 + 1)$-dimensional example of constructing the subgrid while the right-hand side of the figure gives a $(2+1)$-dimensional illustration of the new facets and edges resulting from the subgrid refinement. We point readers to fig. 3.3 for a $(3 + 1)$-dimensional illustration.

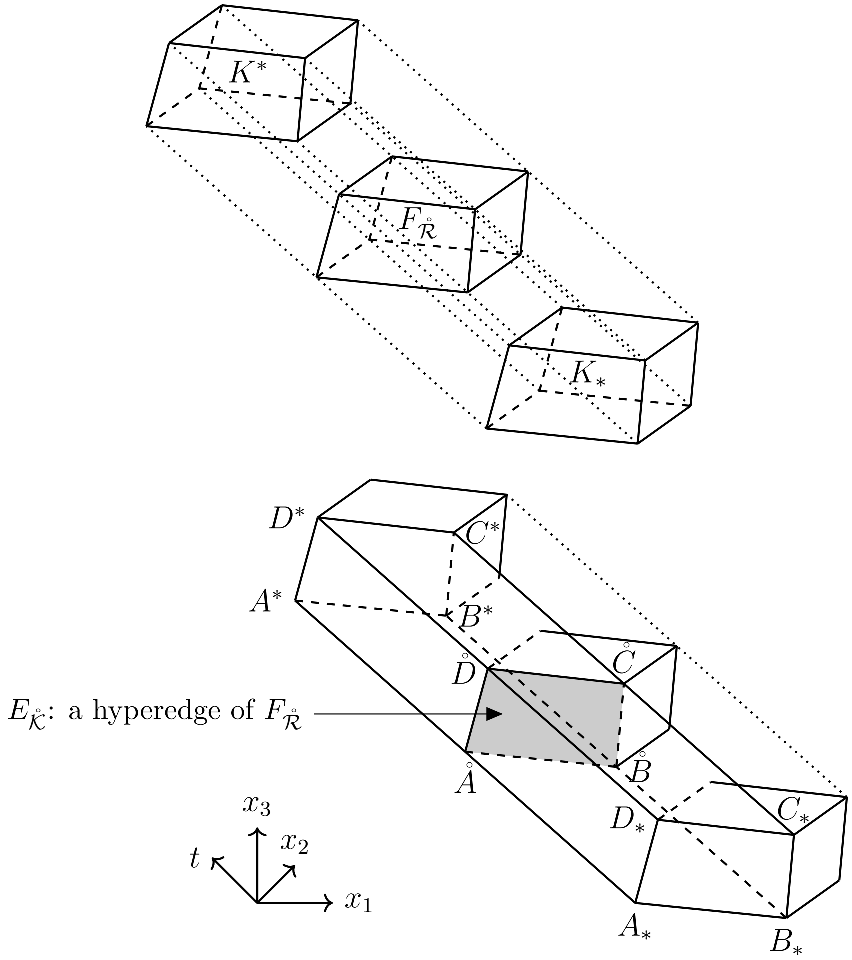


Figure 3.3: A (3 + 1)-dimensional illustration of new hyperfaces of dimension 2 and 3 resulting from the subgrid refinement. In the left-hand side panel, a 4-dimensional space-time element $\mathcal{K}$ is shown. Hexahedra K^* and K_* are boundary facets of $\mathcal{K}$ in $\mathcal{R}_{\mathcal{K}}$. The subgrid facet $F_{\mathring{\mathcal{R}}}$ is obtained by halving the time-step of $\mathcal{K}$. In the right-hand side panel, one of the six $\mathcal{Q}$-facets of $\mathcal{K}$ is shown by connecting A^* with A_*, B^* with B_*, C^* with C_* and D^* with D_*. The quadrilateral formed by $\mathring{A}$, $\mathring{B}$, $\mathring{C}$ and $\mathring{D}$ is highlighted as one of the six hyperedges of $F_{\mathring{\mathcal{R}}}$.

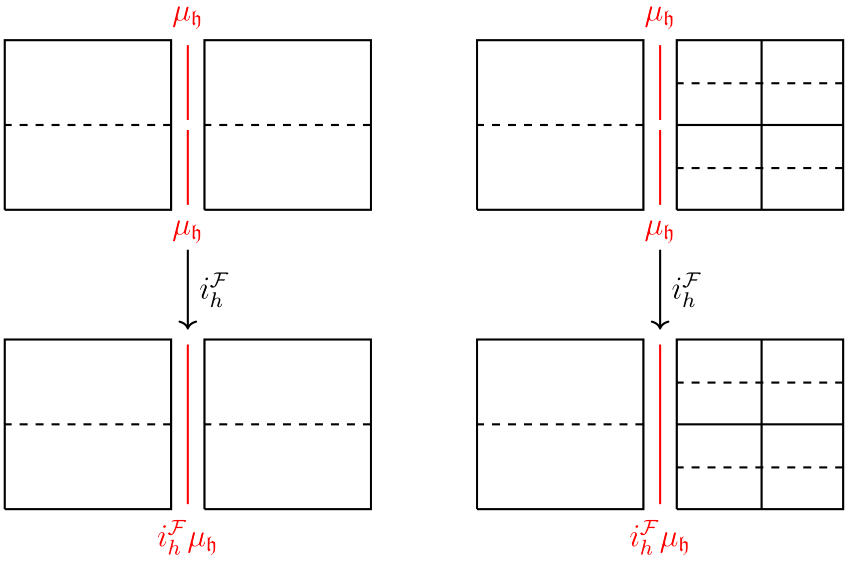


Figure 3.4: Illustration of subgrid projection $i_h^\mathcal{F}$ onto an interior $\mathcal{Q}$-facet in $\mathcal{F}_{\mathcal{Q},h}$. The neighboring elements of the $\mathcal{Q}$-facet are on the same refinement level in the left column and are on different refinement levels in the right column.

Lemma 3.6. *Let $\boldsymbol{v}_\mathfrak{h} \in \boldsymbol{V}_\mathfrak{h}$, let the projection operator i_h be defined as in definition 3.1. Then,*

$$\|(I - i_h^\mathcal{K})\, v_\mathfrak{h}\|_\mathcal{K} \le c\left(\delta t_\mathcal{K}^{1/2}\big\|[\![v_\mathfrak{h}]\!]\big\|_{F_{\mathring{\mathcal{R}}}} + \delta t_\mathcal{K}^{3/2}\big\|[\![\partial_t v_\mathfrak{h}]\!]\big\|_{F_{\mathring{\mathcal{R}}}}\right) \qquad \text{for } \mathcal{K} \in \mathcal{T}_h, \tag{3.31a}$$

$$\|(I - i_h^\mathcal{F})\, \mu_\mathfrak{h}\|_{F_\mathcal{Q}} \le c\left(\delta t_\mathcal{K}^{1/2}\big\|\langle\!\langle \mu_\mathfrak{h}\rangle\!\rangle\big\|_{E_{\mathring{\mathcal{K}}}} + \delta t_\mathcal{K}^{3/2}\big\|\langle\!\langle \partial_t \mu_\mathfrak{h}\rangle\!\rangle\big\|_{E_{\mathring{\mathcal{K}}}}\right) \qquad \text{for } F_\mathcal{Q} \in \mathcal{F}_{\mathcal{Q},h}, \tag{3.31b}$$

where $\mathcal{K}$ on the right-hand side of eq. (3.31b) is chosen such that $F_\mathcal{Q} \subset \mathcal{Q}_\mathcal{K}$.

Proof. We show eq. (3.31) based on an idea in the proof of [19, Lemma 3.1]. On the reference element $\widehat{\mathcal{K}}$, let $\widehat{v}_\mathfrak{h}$ be defined as follows

$$\widehat{v}_\mathfrak{h} := \begin{cases} \sum_{0\le p_i\le p_s, 1\le i\le d} k^*_{p_1\dots p_d}\widehat{t}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} \\ \qquad + \sum_{0\le p_i\le p_s, 1\le i\le d} b^*_{p_1\dots p_d}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} & \text{on } \widehat{\mathring{\mathcal{K}}}^*, \\ \sum_{0\le p_i\le p_s, 1\le i\le d} k_{p_1\dots p_d,*}\widehat{t}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} \\ \qquad + \sum_{0\le p_i\le p_s, 1\le i\le d} b_{p_1\dots p_d,*}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} & \text{on } \widehat{\mathring{\mathcal{K}}}_*, \end{cases}$$

and let

$$\widehat{w}_h^\circ := \sum_{0\le p_i\le p_s, 1\le i\le d} \tfrac{1}{2}\left(k^*_{p_1\dots p_d} + k_{p_1\dots p_d,*}\right)\widehat{t}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} \\ + \sum_{0\le p_i\le p_s, 1\le i\le d} \tfrac{1}{2}\left(b^*_{p_1\dots p_d} + b_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}.$$

Then, by Hölder's inequality for sums and Fubini's theorem, we have

$$\begin{aligned}\|\widehat{v}_{\mathfrak{h}} - \widehat{w}_h^\circ\|^2_{\widehat{\mathcal{K}}} =& \int_{\widehat{\mathcal{K}}} \Big(\sum_{0\le p_i\le p_s} \tfrac{1}{2}\left(k^*_{p_1\dots p_d} - k_{p_1\dots p_d,*}\right)\widehat{t}\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d} \\ &+ \sum_{0\le p_i\le p_s} \tfrac{1}{2}\left(b^*_{p_1\dots p_d} - b_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\Big)^2 \,\mathrm{d}\widehat{x}\mathrm{d}\widehat{t} \\ \le& c\int_{\widehat{\mathcal{K}}} \Big(\sum_{0\le p_i\le p_s} \left(k^*_{p_1\dots p_d} - k_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\Big)^2 \widehat{t}^2 \,\mathrm{d}\widehat{x}\mathrm{d}\widehat{t} \\ &+ c\int_{\widehat{\mathcal{K}}} \Big(\sum_{0\le p_i\le p_s} \left(b^*_{p_1\dots p_d} - b_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\Big)^2 \,\mathrm{d}\widehat{x}\mathrm{d}\widehat{t} \\ \le& c\int_{\widehat{x}\in[-1,1]^d} \Big(\sum_{0\le p_i\le p_s} \left(k^*_{p_1\dots p_d} - k_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\Big)^2 \,\mathrm{d}\widehat{x} \\ &+ c\int_{\widehat{x}\in[-1,1]^d} \Big(\sum_{0\le p_i\le p_s} \left(b^*_{p_1\dots p_d} - b_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\Big)^2 \,\mathrm{d}\widehat{x},\end{aligned} \tag{3.32}$$

We further observe that

$$\begin{aligned}\|[\![\widehat{v}_{\mathfrak{h}}]\!]\|^2_{\widehat{F}_{\mathring{\mathcal{R}}}} + \|[\![\partial_{\widehat{t}}\widehat{v}_{\mathfrak{h}}]\!]\|^2_{\widehat{F}_{\mathring{\mathcal{R}}}} =& \int_{\widehat{x}\in[-1,1]^d} \big(\sum_{0\le p_i\le p_s} \left(b^*_{p_1\dots p_d} - b_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\big)^2 \,\mathrm{d}\widehat{x} \\ &+ \int_{\widehat{x}\in[-1,1]^d} \big(\sum_{0\le p_i\le p_s} \left(k^*_{p_1\dots p_d} - k_{p_1\dots p_d,*}\right)\widehat{x}_1^{p_1}\widehat{x}_2^{p_2}\cdots\widehat{x}_d^{p_d}\big)^2 \,\mathrm{d}\widehat{x}.\end{aligned} \tag{3.33}$$

Combining eqs. (3.32) and (3.33) we conclude that

$$\inf_{\widehat{w}_h\in Q^{(1,p_s)}(\widehat{\mathcal{K}})} \|\widehat{v}_{\mathfrak{h}} - \widehat{w}_h\|_{\widehat{\mathcal{K}}} \le \|\widehat{v}_{\mathfrak{h}} - \widehat{w}_h^\circ\|_{\widehat{\mathcal{K}}} \le c\left(\|[\![\widehat{v}_{\mathfrak{h}}]\!]\|_{\widehat{F}_{\mathring{\mathcal{R}}}} + \|[\![\partial_{\widehat{t}}\widehat{v}_{\mathfrak{h}}]\!]\|_{\widehat{F}_{\mathring{\mathcal{R}}}}\right).$$

Since the L^2-projection is optimal [45, eq.(18.32)], we have shown eq. (3.31a) on the reference element, i.e.,

$$\|(I - \widehat{i^{\mathcal{K}}})\,\widehat{v}_{\mathfrak{h}}\|_{\widehat{\mathcal{K}}} \le c\left(\|[\![\widehat{v}_{\mathfrak{h}}]\!]\|_{\widehat{F}_{\mathring{\mathcal{R}}}} + \|[\![\partial_{\widehat{t}}\widehat{v}_{\mathfrak{h}}]\!]\|_{\widehat{F}_{\mathring{\mathcal{R}}}}\right).$$

Scaling arguments eqs. (3.1a), (3.1c) and (3.4a) now give us

$$\|(I - \widetilde{i}_h^{\mathcal{K}})\, \widetilde{v}_{\mathfrak{h}}\|_{\widetilde{\mathcal{K}}} \leq c\,(\delta t_{\mathcal{K}}^{1/2}\, \|\llbracket \widetilde{v}_{\mathfrak{h}} \rrbracket\|_{\widetilde{F}_{\mathring{\mathcal{R}}}} + \delta t_{\mathcal{K}}^{3/2}\, \|\llbracket \partial_{\tilde{t}} \widetilde{v}_{\mathfrak{h}} \rrbracket\|_{\widetilde{F}_{\mathring{\mathcal{R}}}})\,. \tag{3.34}$$

Combining eq. (3.34), eqs. (3.2a), (3.2c) and (3.4b), we conclude eq. (3.31a). The proof of eq. (3.31b) is similar, but in one lower spatial dimension. □

Lemma 3.7. *Let $\boldsymbol{v}_{\mathfrak{h}} \in \boldsymbol{V}_{\mathfrak{h}}$, there holds:*

$$\|\llbracket v_{\mathfrak{h}} \rrbracket\|_{F_{\mathring{\mathcal{R}}}} \leq c \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \||\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2}\, [\boldsymbol{v}_{\mathfrak{h}}]\|_{\partial\mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} \leq c \||\boldsymbol{v}_{\mathfrak{h}}\||_{s,\mathfrak{h}}, \tag{3.35a}$$

$$\|\llbracket \partial_t v_{\mathfrak{h}} \rrbracket\|_{F_{\mathring{\mathcal{R}}}} \leq c \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \|\partial_t v_{\mathfrak{h}}\|_{\partial\mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} \leq c \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \delta t_{\mathcal{K}}^{-1/2}\, \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}\,, \tag{3.35b}$$

$$\begin{aligned}\|\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} \leq c h_K^{-1/2} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} & \||\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2}\, [\boldsymbol{v}_{\mathfrak{h}}]\|_{\partial\mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} \\ & + c\delta t_{\mathcal{K}}^{-1/2} \left(\|[\boldsymbol{v}_{\mathfrak{h}}^*]\|_{F_{\mathcal{Q}}^*} + \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \right),\end{aligned} \tag{3.35c}$$

$$\begin{aligned}\|\langle\!\langle \partial_t \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} \leq c h_K^{-1/2} \delta t_{\mathcal{K}}^{-1/2} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} & \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} \\ & + c\delta t_{\mathcal{K}}^{-3/2} \left(\|[\boldsymbol{v}_{\mathfrak{h}}^*]\|_{F_{\mathcal{Q}}^*} + \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \right).\end{aligned} \tag{3.35d}$$

Proof. For eq. (3.35a), we write the DG jump in terms of HDG jumps by inserting the facet variable:

$$\begin{aligned}(\llbracket v_{\mathfrak{h}} \rrbracket|_{F_{\mathring{\mathcal{R}}}})^2 &= ([\boldsymbol{v}_{\mathfrak{h}}]\,|_{\partial\mathring{\mathcal{K}}^* \cap F_{\mathring{\mathcal{R}}}} - [\boldsymbol{v}_{\mathfrak{h}}]\,|_{\partial\mathring{\mathcal{K}}_* \cap F_{\mathring{\mathcal{R}}}})^2 \\ &= (\sqrt{2/3}\,|\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2}\, [\boldsymbol{v}_{\mathfrak{h}}]\,|_{\partial\mathring{\mathcal{K}}^* \cap F_{\mathring{\mathcal{R}}}} - \sqrt{2}\,|\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2}\, [\boldsymbol{v}_{\mathfrak{h}}]\,|_{\partial\mathring{\mathcal{K}}_* \cap F_{\mathring{\mathcal{R}}}})^2\,,\end{aligned}$$

where we factor in $|\beta_s - \frac{1}{2}\beta \cdot n|^{1/2}$ due to that $|\beta_s - \frac{1}{2}\beta \cdot n|^{1/2} = \sqrt{3/2}$ on an $\mathcal{R}$-face if $n_t = -1$ and $|\beta_s - \frac{1}{2}\beta \cdot n|^{1/2} = \sqrt{1/2}$ on an $\mathcal{R}$-face if $n_t = 1$; eq. (3.35a) then follows by the triangle inequality and the definition of $\||\cdot\||_{s,\mathfrak{h}}$.

For eq. (3.35b), we expand the DG jump and apply the triangle inequality; the trace inequality eq. (3.7d) then concludes the bound.

To show eq. (3.35c) we require a more involved splitting. For the edge $E_{\mathring{\mathcal{K}}}$ on a $\mathcal{Q}$-face $F_{\mathcal{Q}}$, we observe the following:

$$\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle|_{E_{\mathring{\mathcal{K}}}} = -\,[\boldsymbol{v}_{\mathfrak{h}}^*]\,|_{\mathcal{Q}_{\mathring{\mathcal{K}}^*} \cap E_{\mathring{\mathcal{K}}}} + [\boldsymbol{v}_{\mathfrak{h},*}]\,|_{\mathcal{Q}_{\mathring{\mathcal{K}}_*} \cap E_{\mathring{\mathcal{K}}}} + \langle\!\langle v_{\mathfrak{h}} \rangle\!\rangle|_{E_{\mathring{\mathcal{K}}}}, \tag{3.36}$$

where $\boldsymbol{v}_{\mathfrak{h}}^{*}$ and $\boldsymbol{v}_{\mathfrak{h},*}$ denote the HDG solution pairs on $\mathring{\mathcal{K}}^{*}$ and $\mathring{\mathcal{K}}_{*}$, respectively. We apply the triangle inequality on $\|\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}}$ and use trace inequalities eq. (3.29) to obtain:

$$
\begin{aligned}
\|\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} &\leq \|\langle\!\langle v_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} + \|[\boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{E_{\mathring{\mathcal{K}}}} + \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{E_{\mathring{\mathcal{K}}}} \\
&\leq c \left(h_K^{-1/2} \|[\![v_{\mathfrak{h}}]\!]\|_{F_{\mathring{\mathcal{R}}}} + \delta t_{\mathcal{K}}^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{F_{\mathring{\mathcal{Q}}}^{*}} + \delta t_{\mathcal{K}}^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \right) \\
&\leq c \left(h_K^{-1/2} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \| |\beta_s - \tfrac{1}{2} \beta \cdot n|^{1/2} [\boldsymbol{v}_{\mathfrak{h}}] \|_{\partial \mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} + \delta t_{\mathcal{K}}^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{F_{\mathring{\mathcal{Q}}}^{*}} + \delta t_{\mathcal{K}}^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \right).
\end{aligned}
$$

To show eq. (3.35d) we again use the splitting eq. (3.36), followed by the triangle inequality, trace inequalities eq. (3.29) and eq. (3.7d), and inverse inequality eq. (3.8) to obtain:

$$
\begin{aligned}
\|\langle\!\langle \partial_t \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} &\leq \|\langle\!\langle \partial_t v_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} + \|[\partial_t \boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{E_{\mathring{\mathcal{K}}}} + \|[\partial_t \boldsymbol{v}_{\mathfrak{h},*}]\|_{E_{\mathring{\mathcal{K}}}} \\
&\leq c \left(h_K^{-1/2} \|[\![\partial_t v_{\mathfrak{h}}]\!]\|_{F_{\mathring{\mathcal{R}}}} + \delta t_{\mathcal{K}}^{-1/2} \|[\partial_t \boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{F_{\mathring{\mathcal{Q}}}^{*}} + \delta t_{\mathcal{K}}^{-1/2} \|[\partial_t \boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \right) \\
&\leq c \Big(\sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} h_K^{-1/2} \delta t_{\mathcal{K}}^{-1/2} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} + \delta t_{\mathcal{K}}^{-3/2} \|[\boldsymbol{v}_{\mathfrak{h}}^{*}]\|_{F_{\mathring{\mathcal{Q}}}^{*}} + \delta t_{\mathcal{K}}^{-3/2} \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \Big).
\end{aligned}
$$

□

Lemma 3.8. *Let $\boldsymbol{v}_{\mathfrak{h}} \in \boldsymbol{V}_{\mathfrak{h}}$ and let the projection operator i_h be defined as in definition 3.1. Consider an element $\mathcal{K} \in \mathcal{T}_h$ such that it has a $\mathcal{Q}$-facet $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$ such that $F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}$. There holds that $(i_h^{\mathcal{K}} v_{\mathfrak{h}})\,|_{F_{\mathcal{Q}}} \equiv i_h^{\mathcal{F}} \left(v_{\mathfrak{h}}|_{F_{\mathcal{Q}}} \right)$ on $\mathcal{F}_{\mathcal{Q},h}$.*

Proof. We verify the equivalence by showing that $\widehat{i_h^{\mathcal{K}}} \widehat{v}_{\mathfrak{h}} \equiv \widehat{i_h^{\mathcal{F}}} \widehat{v}_{\mathfrak{h}}$ on the reference domain. On $\widehat{\mathcal{K}}$, let $\widehat{v}_{\mathfrak{h}}$ be defined as follows

$$
\widehat{v}_{\mathfrak{h}} := \begin{cases} \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d} k_{p_0 p_1 \dots p_d}^{*} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_d^{p_d} & \text{on } \widehat{\mathring{\mathcal{K}}}^{*}, \\ \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d} k_{p_0 p_1 \dots p_d, *} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_d^{p_d} & \text{on } \widehat{\mathring{\mathcal{K}}}_{*}. \end{cases}
$$

Suppose that

$$
\widehat{i_h^{\mathcal{K}}} \widehat{v}_{\mathfrak{h}} = \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d} \widetilde{k}_{p_0 p_1 \dots p_d} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_d^{p_d}.
$$

By definition of the projection, for any $0 \leq q_0 \leq p_t$ and $0 \leq q_i \leq p_s$, $1 \leq i \leq d$

$$
\int_{\widehat{\mathcal{K}}} \left(\widehat{v}_{\mathfrak{h}} - \widehat{i_h^{\mathcal{K}}} \widehat{v}_{\mathfrak{h}} \right) \left(\widehat{t}^{q_0} \widehat{x}_1^{q_1} \widehat{x}_2^{q_2} \cdots \widehat{x}_d^{q_d} \right) \mathrm{d}\widehat{x}\, \mathrm{d}\widehat{t} = 0.
$$

Let us denote the $\mathcal{Q}$-face on which $\widehat{x}_d = 1$ by $\widehat{F}_d$. Then, without loss of generality, and using Fubini's theorem,

$$\begin{aligned}
&\int_{\widehat{\mathring{\mathcal{K}}}^*} \sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d} (k^*_{p_0\ldots p_d} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}\, \mathrm{d}\widehat{t} \\
&\quad + \int_{\widehat{\mathring{\mathcal{K}}}_*} \sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d} (k_{p_0\ldots p_d,*} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}\, \mathrm{d}\widehat{t} \\
&= \sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d} \Bigg(\int_{\widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d} (k^*_{p_0\ldots p_d} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \int_{-1}^{1} \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}_d \\
&\quad + \int_{\widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d} (k_{p_0\ldots p_d,*} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \int_{-1}^{1} \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}_d \Bigg) \\
&= \sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d} \int_{-1}^{1} \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}_d \Bigg(\int_{\widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d} (k^*_{p_0\ldots p_d} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \\
&\quad + \int_{\widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d} (k_{p_0\ldots p_d,*} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \Bigg).
\end{aligned}$$

Note that $\int_{-1}^{1} \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}_d = 0$ for $p_d + q_d$ odd. Then, for each $0 \le q_d \le p_s$ leaving out p_d's such that $p_d + q_d$ is odd, we have:

$$\begin{aligned}
&\sum_{p_d \text{ s.t. } p_d+q_d \text{ is even}} \left(\int_{-1}^{1} \widehat{x}_d^{p_d+q_d}\, \mathrm{d}\widehat{x}_d \right) \cdot \\
&\Bigg(\sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d-1} \Big(\int_{\widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d} (k^*_{p_0\ldots p_d} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \\
&\quad + \int_{\widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d} (k_{p_0\ldots p_d,*} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \Big) \Bigg) = 0.
\end{aligned}$$

Using that $\int_{-1}^{1} \widehat{x}_d^{2k}\, \mathrm{d}\widehat{x}_d = 2/(2k+1)$ and introducing

$$\begin{aligned}
z_{p_d} := \sum_{0\le p_0\le p_t, 0\le p_i\le p_s, 1\le i\le d-1} \Big(&\int_{\widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d} (k^*_{p_0\ldots p_d} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \\
&+ \int_{\widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d} (k_{p_0\ldots p_d,*} - \widetilde{k}_{p_0\ldots p_d})\, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}}\, \mathrm{d}\widehat{s} \Big),
\end{aligned}$$

we have for each $0 \le q_d \le p_s$

$$\sum_{p_d \text{ s.t. } p_d+q_d \text{ is even}} \frac{2z_{p_d}}{p_d+q_d+1} = 0.$$

Writing this as a linear system we find, for p_s even,

$$\begin{bmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & \cdots & \frac{1}{p_s+1} \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \cdots & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \cdots & \frac{1}{p_s+3} \\ 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \cdots & 0 \\ \frac{1}{5} & 0 & \frac{1}{7} & 0 & \frac{1}{9} & \cdots & \frac{1}{p_s+5} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{p_s+1} & 0 & \frac{1}{p_s+3} & 0 & \frac{1}{p_s+5} & \cdots & \frac{1}{2p_s+1} \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_{p_s-2} \\ z_{p_s-1} \\ z_{p_s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

When p_s is odd we find:

$$\begin{bmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & \cdots & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \cdots & \frac{1}{p_s+2} \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \cdots & 0 \\ 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \cdots & \frac{1}{p_s+4} \\ \frac{1}{5} & 0 & \frac{1}{7} & 0 & \frac{1}{9} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{p_s+2} & 0 & \frac{1}{p_s+4} & 0 & \cdots & \frac{1}{2p_s+1} \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_{p_s-2} \\ z_{p_s-1} \\ z_{p_s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

After suitable transformations, the above matrices can be written as, when p_s is even,

$$\begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{5} & \cdots & \frac{1}{p_s+1} & 0 & \cdots & \cdots & 0 \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & \cdots & \frac{1}{p_s+3} & 0 & \cdots & \cdots & 0 \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{9} & \cdots & \frac{1}{p_s+5} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \cdots & 0 \\ \frac{1}{p_s+1} & \frac{1}{p_s+3} & \frac{1}{p_s+5} & \cdots & \frac{1}{2p_s+1} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{3} & \frac{1}{5} & \cdots & \frac{1}{p_s+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \frac{1}{5} & \frac{1}{7} & \cdots & \frac{1}{p_s+3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{p_s+1} & \frac{1}{p_s+3} & \cdots & \frac{1}{2p_s-1} \end{bmatrix}, \tag{3.37}$$

and, when p_s is odd,

$$\begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{5} & \cdots & \frac{1}{p_s} & 0 & \cdots & \cdots & 0 \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & \cdots & \frac{1}{p_s+2} & 0 & \cdots & \cdots & 0 \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{9} & \cdots & \frac{1}{p_s+4} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \cdots & 0 \\ \frac{1}{p_s} & \frac{1}{p_s+2} & \frac{1}{p_s+4} & \cdots & \frac{1}{2p_s-1} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{3} & \frac{1}{5} & \cdots & \frac{1}{p_s+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \frac{1}{5} & \frac{1}{7} & \cdots & \frac{1}{p_s+4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{p_s+2} & \frac{1}{p_s+4} & \cdots & \frac{1}{2p_s+1} \end{bmatrix}. \tag{3.38}$$

We observe that two diagonal block matrices in eqs. (3.37) and (3.38) are examples of the Hankel matrix (a square matrix in which elements on each skew-diagonal are constant). They can further be shown to be totally positive, see [50, Example 0.1.8]. Therefore, we conclude that both block matrices eqs. (3.37) and (3.38) are nonsingular and $z_i = 0$ for $0 \leq i \leq p_s$, i.e.,

$$\begin{aligned} \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d-1} \Big(\int_{\widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d} & (k^*_{p_0 \ldots p_d} - \widetilde{k}_{p_0 \ldots p_d}) \, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}} \, \mathrm{d}\widehat{s} \\ & + \int_{\widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d} (k_{p_0 \ldots p_d, *} - \widetilde{k}_{p_0 \ldots p_d}) \, \widehat{t}^{p_0+q_0} \widehat{x}_1^{p_1+q_1} \cdots \widehat{x}_{d-1}^{p_{d-1}+q_{d-1}} \, \mathrm{d}\widehat{s} \Big) = 0, \end{aligned} \tag{3.39}$$

for any $0 \leq q_0 \leq p_t$ and $0 \leq q_i \leq p_s$, $1 \leq i \leq d-1$. Therefore, considering

$$\begin{aligned} \widehat{\mu}_{\mathfrak{h}, p_d} &:= \begin{cases} \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d-1} k^*_{p_0 \ldots p_d} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_{d-1}^{p_{d-1}} & \text{on } \widehat{\mathring{\mathcal{K}}}^* \cap \widehat{F}_d, \\ \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d-1} k_{p_0 \ldots p_d, *} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_{d-1}^{p_{d-1}} & \text{on } \widehat{\mathring{\mathcal{K}}}_* \cap \widehat{F}_d, \end{cases} \\ \widehat{\lambda}_{\mathfrak{h}, p_d} &:= \sum_{0 \leq p_0 \leq p_t, 0 \leq p_i \leq p_s, 1 \leq i \leq d-1} \widetilde{k}_{p_0 \ldots p_d} \widehat{t}^{p_0} \widehat{x}_1^{p_1} \widehat{x}_2^{p_2} \cdots \widehat{x}_{d-1}^{p_{d-1}} \text{ on } \widehat{F}_d, \end{aligned}$$

we conclude from eq. (3.39) that $\widehat{i}_h^{\mathcal{F}} \widehat{\mu}_{\mathfrak{h}, p_d} = \widehat{\lambda}_{\mathfrak{h}, p_d}$. Further, observing that $\widehat{v}_{\mathfrak{h}}|_{\widehat{F}_d} = \sum_{p_d} \widehat{\mu}_{\mathfrak{h}, p_d}$, that $(\widehat{i}_h^{\mathcal{K}} \widehat{v}_{\mathfrak{h}})|_{\widehat{F}_d} = \sum_{p_d} \widehat{\lambda}_{\mathfrak{h}, p_d}$ and that projection is linear, we conclude that $(\widehat{i}_h^{\mathcal{K}} \widehat{v}_{\mathfrak{h}})|_{\widehat{F}_d} = \widehat{i}_h^{\mathcal{F}} (\widehat{v}_{\mathfrak{h}}|_{\widehat{F}_d})$. □

Chapter 4

Péclet-robust a priori error analysis

An a priori error analysis of a space-time HDG discretization of the time-dependent advection-diffusion equation is presented in this chapter. We prove the inf-sup stability of the space-time HDG method in section 4.1. Based on the inf-sup stability estimate, a Galerkin orthogonality and projection estimates, we present the error analysis in section 4.2. Finally, in section 4.3, we illustrate the theoretical results with numerical examples.

4.1 Stability

The main goal of this section is to prove theorem 4.1 which states stability of the space-time HDG method for the advection-diffusion equation with respect to a norm that includes measurement of the streamline derivative, i.e., $|||\cdot|||_{ss,h}$ defined in eq. (2.7c). We will prove that this result is robust with respect to the Péclet number. A similar result for the stationary problem is shown in [12, Theorem 4.6]. For this and following sections, c_T denotes a constant independent of h_K, $\delta t_{\mathcal{K}}$, and ε, but linear in T.

Theorem 4.1. *There exists δt_0, independent of ε and T, such that when $\delta t_{\mathcal{K}} \le \min(h_K, \delta t_0)$ on all $\mathcal{K} \in \mathcal{T}_h$, and for all $\boldsymbol{w}_h \in \boldsymbol{V}_h$,*

$$c_T^{-1} |||\boldsymbol{w}_h|||_{ss,h} \le \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{a_h(\boldsymbol{w}_h, \boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{s,h}}. \tag{4.1}$$

The following two inf-sup conditions with respect to, respectively, $|||\cdot|||_{v,h}$ and $|||\cdot|||_{s,h}$, and which hold under the same conditions as theorem 4.1, are used to prove theorem 4.1:

$$c_T^{-1}|||\boldsymbol{w}_h|||_{v,h} \leq \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{a_h(\boldsymbol{w}_h, \boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{v,h}} \qquad \forall \boldsymbol{w}_h \in \boldsymbol{V}_h, \tag{4.2a}$$

$$c_T^{-1}|||\boldsymbol{w}_h|||_{s,h} \leq \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{a_h(\boldsymbol{w}_h, \boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{s,h}} \qquad \forall \boldsymbol{w}_h \in \boldsymbol{V}_h. \tag{4.2b}$$

We prove eqs. (4.2a) and (4.2b) in sections 4.1.1 and 4.1.2, respectively. We then prove theorem 4.1 in section 4.1.3. To prove these results we introduce, for $T \geq 1$, the weighting function

$$\varphi = eT\exp(-t/T) + \chi, \tag{4.3}$$

where the positive constant χ will be determined later. For $0 < T < 1$ we propose $\varphi(t) = e\exp(-t) + \chi$. In our analysis, however, we will only consider $T \geq 1$; the analysis for $T < 1$ follows identical steps as the $T \geq 1$ case, resulting in inf-sup conditions theorem 4.1 and eqs. (4.2a) and (4.2b) independent of T. Denoting the cell mean of $\bar{\beta}$ by $\bar{\beta}_0$, we will also use that (see [28]),

$$\|\bar{\beta} - \bar{\beta}_0\|_{L^\infty(\mathcal{K})} \leq ch_K|\bar{\beta}|_{W^{1,\infty}(\mathcal{K})} \quad \forall \mathcal{K} \in \mathcal{T}_h. \tag{4.4}$$

4.1.1 The inf-sup condition with respect to $|||\cdot|||_{v,h}$

To prove eq. (4.2a) we first require the following lemmas.

Lemma 4.1. *Let φ be defined as in eq. (4.3) with χ chosen such that $\chi > (e-\sqrt{2})T/(\sqrt{2}-1)$. Furthermore, choose the penalty parameter α in eq. (2.9) such that $\alpha > 1 + 4c_\star^2$, with $c_\star$ the constant in eq. (3.7c). Then for all $\boldsymbol{w}_h := (w_h, \varkappa_h) \in \boldsymbol{V}_h$:*

$$\begin{aligned} a_h(\boldsymbol{w}_h, \varphi\boldsymbol{w}_h) \geq &\tfrac{1}{2}(T+\chi)\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla} w_h\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1}\|[\boldsymbol{w}_h]\|_{\mathcal{Q}_\mathcal{K}}^2\Big) + \tfrac{1}{2}\sum_{\mathcal{K}\in\mathcal{T}_h}\|w_h\|_{\mathcal{K}}^2 \\ &+ (T+\chi)\Big(\big\||\tfrac{1}{2}\beta\cdot n|^{1/2}\varkappa_h\big\|_{\partial\mathcal{E}_N}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h}\big\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}[\boldsymbol{w}_h]\big\|_{\partial\mathcal{K}}^2\Big). \end{aligned}$$

Proof. On an element $\mathcal{K} \in \mathcal{T}_h$ we have $-w_h\beta\cdot\nabla(\varphi w_h) = -\frac{1}{2}\nabla\cdot(\varphi\beta w_h^2) - \frac{1}{2}w_h^2\beta\cdot\nabla\varphi$. Using Gauss's theorem, $[\varphi\boldsymbol{w}_h] = \varphi[\boldsymbol{w}_h]$, and that $\zeta^+\beta\cdot n = (\beta\cdot n + |\beta\cdot n|)/2$, we note that

$$\begin{aligned} a_{h,c}(\boldsymbol{w}_h, \varphi\boldsymbol{w}_h) = &-(\tfrac{1}{2}w_h^2, \beta\cdot\nabla\varphi)_{\mathcal{T}_h} + \langle\tfrac{1}{2}\varphi\varkappa_h^2, \beta\cdot n + |\beta\cdot n|\rangle_{\partial\mathcal{E}_N} \\ &-\langle\tfrac{1}{2}\varphi w_h^2, \beta\cdot n\rangle_{\partial\mathcal{T}_h} + \langle\varphi[\boldsymbol{w}_h]^2, \sup|\beta\cdot n|\rangle_{\partial\mathcal{T}_h} + \langle\varphi\varkappa_h[\boldsymbol{w}_h], \beta\cdot n\rangle_{\partial\mathcal{T}_h}. \end{aligned}$$

Since $-\frac{1}{2}w_h^2+\varkappa_h\left[\boldsymbol{w}_h\right] = -\frac{1}{2}\left[\boldsymbol{w}_h\right]^2-\frac{1}{2}\varkappa_h^2$, $\varkappa_h$ is single-valued on element boundaries, $\varkappa_h=0$ on $\partial\mathcal{E}_D$, we have by definition of φ and using that $-\beta\cdot\nabla\varphi\geq 1$:

$$a_{h,c}(\boldsymbol{w}_h,\varphi\boldsymbol{w}_h) \geq (T+\chi)\left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2}\varkappa_h\right\|^2_{\partial\mathcal{E}_N} + \tfrac{1}{2}\sum_{\mathcal{K}\in\mathcal{T}_h}\|w_h\|^2_{\mathcal{K}} \\ + (T+\chi)\sum_{\mathcal{K}\in\mathcal{T}_h}\left\| \left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{w}_h\right]\right\|^2_{\partial\mathcal{K}}. \tag{4.5}$$

Next, noting that $\overline{\nabla}\varphi=0$, and using the Cauchy–Schwarz inequality and eq. (3.7c),

$$a_{h,d}(\boldsymbol{w}_h,\varphi\boldsymbol{w}_h) \geq (T+\chi)\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon\left\|\overline{\nabla}w_h\right\|^2_{\mathcal{K}} + (T+\chi)\,\alpha\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K^{-1}\left\|\left[\boldsymbol{w}_h\right]\right\|^2_{\mathcal{Q}_{\mathcal{K}}} \\ - \sum_{\mathcal{K}\in\mathcal{T}_h}2\varepsilon^{1/2}c_\star\,(eT+\chi)\left\|\overline{\nabla}w_h\right\|_{\mathcal{K}}\varepsilon^{1/2}h_K^{-1/2}\left\|\left[\boldsymbol{w}_h\right]\right\|_{\mathcal{Q}_{\mathcal{K}}}.$$

Using Hölder's inequality for sums and the inequality $ax^2-2bxy+dy^2\geq(ad-b^2)(x^2+y^2)/(a+d)$, which holds for positive real numbers a,b,d and $ad>b^2$ (see [86]) allows us to obtain

$$a_{h,d}(\boldsymbol{w}_h,\varphi\boldsymbol{w}_h) \geq (T+\chi)\,\frac{\alpha-\left(\frac{c_\star(eT+\chi)}{T+\chi}\right)^2}{1+\alpha}\,\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon\left\|\overline{\nabla}w_h\right\|^2_{\mathcal{K}}+\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K^{-1}\left\|\left[\boldsymbol{w}_h\right]\right\|^2_{\mathcal{Q}_{\mathcal{K}}}\Big).$$

Since χ and α are chosen such that $\chi>(e-\sqrt{2})T/(\sqrt{2}-1)$, so that $T+\chi>(eT+\chi)/\sqrt{2}$, and $\alpha>1+4c_\star^2$, it follows that

$$a_{h,d}(\boldsymbol{w}_h,\varphi\boldsymbol{w}_h) \geq \tfrac{1}{2}(T+\chi)\,\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon\left\|\overline{\nabla}w_h\right\|^2_{\mathcal{K}}+\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K^{-1}\left\|\left[\boldsymbol{w}_h\right]\right\|^2_{\mathcal{Q}_{\mathcal{K}}}\Big). \tag{4.6}$$

The result follows after combining eqs. (4.5) and (4.6). □

The following lemma extends the L^2-projection estimates of [12, Lemma 4.2] to space-time elements, taking into account the spatial mesh size h_K and time-step $\delta t_{\mathcal{K}}$.

Lemma 4.2. *Let φ be the function defined in eq. (4.3). For any $w_h\in V_h$ the following estimates hold:*

$$\left\|(I-\Pi_h)\,(\varphi w_h)\right\|_{\mathcal{K}} \leq c\delta t_{\mathcal{K}}\left\|w_h\right\|_{\mathcal{K}}, \tag{4.7a}$$

$$\left\|\overline{\nabla}\left((I-\Pi_h)\,(\varphi w_h)\right)\right\|_{\mathcal{K}} \leq c\delta t_{\mathcal{K}}h_K^{-1}\left\|w_h\right\|_{\mathcal{K}}, \tag{4.7b}$$

$$\left\|\overline{\nabla}\left((I-\Pi_h)\,(\varphi w_h)\right)\right\|_{\mathcal{Q}_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}h_K^{-3/2}\left\|w_h\right\|_{\mathcal{K}}, \tag{4.7c}$$

$$\left\|(I-\Pi_h)\,(\varphi w_h)\right\|_{\mathcal{Q}_{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}h_K^{-1/2}\left\|w_h\right\|_{\mathcal{K}}, \tag{4.7d}$$

$$\left\|(I-\Pi_h)\,(\varphi w_h)\right\|_{\mathcal{R}_{\mathcal{K}}} \leq c{\delta t_{\mathcal{K}}}^{1/2}\left\|w_h\right\|_{\mathcal{K}}. \tag{4.7e}$$

Proof. We first observe that when φ is a function of the time variable only on $\mathcal{K}$, so are $\widetilde{\varphi}$ and $\widehat{\varphi}$ on $\widetilde{\mathcal{K}}$ and $\widehat{\mathcal{K}}$ respectively. Therefore, for $w_h \in V_h$, and for $1 \leq i, j \leq d$, $j \neq i$,

$$\partial_{\widetilde{x}_i}^{p_s+1}\left(\widetilde{\varphi w_h}\right) = \partial_{\widetilde{x}_i}^{p_s+1}\partial_{\widetilde{x}_j}\left(\widetilde{\varphi w_h}\right) = \partial_{\widetilde{x}_i}^{p_s+1}\partial_{\widetilde{t}}\left(\widetilde{\varphi w_h}\right) = 0. \tag{4.8}$$

The equivalent derivatives above are also zero on the reference domain $\widehat{\mathcal{K}}$. Furthermore,

$$\|\partial_{\widetilde{t}}^{p_t+1}\left(\widetilde{\varphi w_h}\right)\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-p_t}\left\|\widetilde{w}_h\right\|_{\widetilde{\mathcal{K}}}, \tag{4.9a}$$

$$\|\partial_{\widetilde{t}}^{p_t+1}\partial_{\widetilde{x}_i}\left(\widetilde{\varphi w_h}\right)\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-p_t}h_K^{-1}\left\|\widetilde{w}_h\right\|_{\widetilde{\mathcal{K}}}. \tag{4.9b}$$

Equation (4.9a) can be shown using the general Leibniz rule, that $\partial_{\widetilde{t}}^{p_t+1}\widetilde{w_h} = 0$, that $\|\partial_{\widetilde{t}}^{j_t}\widetilde{\varphi}\|_{\widetilde{\mathcal{K}}} \leq e$ for all $1 \leq j_t \leq p_t + 1$, that eq. (3.7a) reduces to $\|\partial_{\widetilde{t}}\widetilde{v}_h\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{-1}\left\|\widetilde{v}_h\right\|_{\widetilde{\mathcal{K}}}$ on the axiparallel element $\widetilde{\mathcal{K}}$ (see [53, Corollary 3.54]) and using that $\delta t_{\mathcal{K}} < 1$. Similar arguments can be used to show eq. (4.9b).

To prove eq. (4.7a) we follow the proof of [12, Lemma 4.2], and apply the projection estimates in [62, Lemma 3.4] when considered on the affine domain $\widetilde{\mathcal{K}}$, eq. (4.8) and eq. (4.9a),

$$\|\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}^{p_t+1}\left\|\partial_{\widetilde{t}}^{p_t+1}\left(\widetilde{\varphi w_h}\right)\right\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}\left\|\widetilde{w_h}\right\|_{\widetilde{\mathcal{K}}}. \tag{4.10}$$

A scaling argument applied to eq. (4.10) from $\widetilde{\mathcal{K}}$ to $\mathcal{K}$ yields eq. (4.7a).

We next prove eq. (4.7b). First note,

$$\left\|\widetilde{\overline{\nabla}}\left(\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\right)\right\|_{\widetilde{\mathcal{K}}}^2 \leq \sum_{1\leq i\leq d} ch_K^{d-2}\delta t_{\mathcal{K}}\left\|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\right\|_{\widehat{\mathcal{K}}}^2. \tag{4.11}$$

Following similar steps as in the proof of [54, Lemma 7.5], the right-hand side of eq. (4.11) can be bound further using the triangle inequality, commutativity of $\partial_{\widehat{x}_i}$ with $\widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i}\widehat{\pi}_{x_j}$, boundedness of $\widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i}\widehat{\pi}_{x_j}$, the projection estimates in [62, Lemma 3.4] and [54, Lemma 7.3], eq. (4.8), eq. (3.1a), and eq. (4.9b),

$$\begin{aligned}\left\|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\right\|_{\widehat{\mathcal{K}}} &\leq \left\|\partial_{\widehat{x}_i}\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\partial_{\widehat{x}_i}\widehat{\varphi w_h}\right)\right\|_{\widehat{\mathcal{K}}} + c\left\|\widehat{\pi}_{x_i}\left(\partial_{\widehat{x}_i}\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\left(\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\right\|_{\widehat{\mathcal{K}}} \\ &\leq c\left\|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_i}\widehat{\varphi w_h}\right\|_{\widehat{\mathcal{K}}} \leq ch_K^{-d/2}\delta t_{\mathcal{K}}^{1/2}\left\|\widetilde{w_h}\right\|_{\widetilde{\mathcal{K}}}.\end{aligned} \tag{4.12}$$

Combining the right-hand side of eq. (4.12) with eq. (4.11), we find:

$$\|\widetilde{\overline{\nabla}}\left(\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\right)\|_{\widetilde{\mathcal{K}}} \leq c\delta t_{\mathcal{K}}h_K^{-1}\left\|\widetilde{w_h}\right\|_{\widetilde{\mathcal{K}}}. \tag{4.13}$$

A scaling argument applied to eq. (4.13) from $\widetilde{\mathcal{K}}$ to $\mathcal{K}$ yields eq. (4.7b). With similar steps it can be shown that

$$\|\partial_{\tilde{t}}\left(\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\right)\|_{\widetilde{\mathcal{K}}} \leq c \, \|\tilde{w}_h\|_{\widetilde{\mathcal{K}}} \, . \tag{4.14}$$

We next prove eq. (4.7c). We start with a scaling argument to transform the integral on a $\mathcal{Q}$-face $\widetilde{F}_{\mathcal{Q},m}$ from the affine domain to the reference domain. Note that, without loss of generality, subscript m denotes the index of the spatial coordinate for which $\widehat{x}_m \equiv 1$. Using eq. (3.1b) we find

$$\|\widetilde{\overline{\nabla}}\left(\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\right)\|^2_{\widetilde{F}_{\mathcal{Q},m}} \leq \sum_{1 \leq i \leq d} c \delta t_{\mathcal{K}} h_K^{d-3} \, \|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\|^2_{\widehat{F}_{\mathcal{Q},m}} \, . \tag{4.15}$$

Consider now the right-hand side term. Following [54, Lemma 7.9] we consider the cases $i = m$ and $i \neq m$ separately, starting with $i = m$. Using the commutativity of $\partial_{\widehat{x}_m}$ with $\widehat{\pi}_t \Pi_{1 \leq j \leq d, j \neq m} \widehat{\pi}_{x_j}$, the triangle inequality, and [53, Lemma 3.47] (see also [54, Lemma 7.8]),

$$\begin{aligned} \|\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \leq & \|\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right) - \widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ & + c \, \|\widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right) - \widehat{\pi}_t \Pi_{1 \leq j \leq d, j \neq m} \widehat{\pi}_{x_j}\left(\widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}} \\ & + c \, \|\widehat{\pi}_t \Pi_{1 \leq j \leq d, j \neq m} \widehat{\pi}_{x_j}\left(\widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_m}\widehat{\pi}_{x_m}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}} \, . \end{aligned} \tag{4.16}$$

The first and third terms on the right-hand side of eq. (4.16) vanish by [54, Lemmas 7.2 and 7.3] and eq. (4.8). The second term on the right-hand side of eq. (4.16) is bounded using the same argument as in the proof of [62, Lemma 3.4] by noting that $\widehat{\pi}_{x_m}$ and $\widehat{\pi}_t$ are one-dimensional L^2-projections applied in the spatial direction $\widehat{x}_m$ and time direction, respectively, the commutativity of $\widehat{\pi}_{x_m}$ with $\partial_{\widehat{t}}^{p_t+1}$ and $\partial_{\widehat{x}_j}^{p_s+1}$ $(j \neq m)$, the boundedness of $\widehat{\pi}_{x_m}$, and eq. (4.8),

$$\|\widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right) - \widehat{\pi}_t \Pi_{1 \leq j \leq d, j \neq m} \widehat{\pi}_{x_j}\left(\widehat{\pi}_{x_m}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}} \leq c \, \|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}} \, , \tag{4.17}$$

so that eq. (4.16) becomes:

$$\|\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \leq c \, \|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_m}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}} \, . \tag{4.18}$$

We now consider the right-hand side of eq. (4.15) with $i \neq m$. We have by a triangle inequality

$$\begin{aligned}\|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \leq & \|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \widehat{\pi}_{x_m}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ & + \|\left(I - \widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i, j\neq m}\widehat{\pi}_{x_j}\right)\widehat{\pi}_{x_m}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ & + \|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ & + \|\left(I - \widehat{\pi}_{x_m}\right)\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ & + \|\left(I - \widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i, j\neq m}\widehat{\pi}_{x_j}\right)\widehat{\pi}_{x_m}\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}}.\end{aligned} \tag{4.19}$$

For the second and the fifth terms on the right-hand side, we observe that the functions inside the norms are polynomials in the $\widehat{x}_m$-direction. Therefore, [54, Lemma 7.8] gives us, using similar steps used to find eq. (4.17),

$$\begin{aligned}&\|\left(I - \widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i, j\neq m}\widehat{\pi}_{x_j}\right)\widehat{\pi}_{x_m}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ &\quad + \|\left(I - \widehat{\pi}_t\Pi_{1\leq j\leq d, j\neq i, j\neq m}\widehat{\pi}_{x_j}\right)\widehat{\pi}_{x_m}\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \\ \leq & c\left(\|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}} + \|\partial_{\widehat{t}}\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}}\right) \\ &\quad + c\left(\sum_{1\leq j\leq d, j\neq m, j\neq i}\|\partial_{\widehat{x}_j}\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}}\right).\end{aligned}$$

Next, using that $\widehat{\pi}_{x_i}$ and $\partial_{\widehat{x}_j}$ commute, using [54, eq.(7.5) in Lemma 7.3] and eq. (4.8), we find that for any $j \neq i$ where $1 \leq i \leq d$ and $0 \leq j \leq d$ with $\widehat{x}_0 = \widehat{t}$,

$$\|\partial_{\widehat{x}_j}\left(\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right) - \partial_{\widehat{x}_i}\widehat{\pi}_{x_i}\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{\mathcal{K}}}^2 \leq c\,\|\partial_{\widehat{x}_i}^{p_s+1}\partial_{\widehat{x}_j}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}}^2 = 0.$$

Therefore, the second and fifth terms on the right-hand side of eq. (4.19) are bounded by $c\,\|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}}$. All remaining terms on the right-hand side of eq. (4.19) vanish by combining [54, Lemmas 7.2 and 7.3] eqs. (3.6) and (4.8) and so, for $i \neq m$,

$$\|\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h} - \widehat{\Pi}_h\left(\widehat{\varphi w_h}\right)\right)\|_{\widehat{F}_{\mathcal{Q},m}} \leq c\,\|\partial_{\widehat{t}}^{p_t+1}\partial_{\widehat{x}_i}\left(\widehat{\varphi w_h}\right)\|_{\widehat{\mathcal{K}}}. \tag{4.20}$$

Combining eqs. (4.15), (4.18) and (4.20) and applying eqs. (3.1a) and (4.9b) we obtain:

$$\|\widetilde{\overline{\nabla}}\left(\widetilde{\varphi w_h} - \widetilde{\Pi}_h\left(\widetilde{\varphi w_h}\right)\right)\|_{\widetilde{F}_{\mathcal{Q},m}} \leq c\delta t_{\mathcal{K}} h_K^{-3/2}\,\|\widetilde{w}_h\|_{\widetilde{\mathcal{K}}}. \tag{4.21}$$

A scaling argument applied to eq. (4.21) from $\widetilde{\mathcal{K}}$ to $\mathcal{K}$ yields eq. (4.7c).

Equation (4.7d) follows directly by combining the local trace inequality eq. (3.6a) with eqs. (4.10) and (4.13):

$$\|\widetilde{\varphi w_h} - \widetilde{\Pi}_h \left(\widetilde{\varphi w_h}\right)\|_{\widetilde{\mathcal{Q}}_{\mathcal{K}}} \le c h_K^{-1/2} \delta t_{\mathcal{K}} \, \|\widetilde{w_h}\|_{\widetilde{\mathcal{K}}} \,, \tag{4.22}$$

and a scaling argument applied to eq. (4.22) from $\widetilde{\mathcal{K}}$ to $\mathcal{K}$. Lastly, eq. (4.7e) follows by combining the local trace inequality eq. (3.6b) with eqs. (4.10) and (4.14):

$$\|\widetilde{\varphi w_h} - \widetilde{\Pi}_h \left(\widetilde{\varphi w_h}\right)\|_{\widetilde{\mathcal{R}}_{\mathcal{K}}} \le c \delta t_{\mathcal{K}}{}^{1/2} \, \|\widetilde{w_h}\|_{\widetilde{\mathcal{K}}} \,, \tag{4.23}$$

and a scaling argument applied to eq. (4.23) from $\widetilde{\mathcal{K}}$ to $\mathcal{K}$. $\square$

Lemma 4.3. *Let* $\boldsymbol{\Pi}_h \left(\varphi \boldsymbol{w}_h\right) := \left(\Pi_h \left(\varphi w_h\right), \Pi_h^{\mathcal{F}} \left(\varphi \varkappa_h\right)\right)$ *for all* $\boldsymbol{w}_h := \left(w_h, \varkappa_h\right) \in \boldsymbol{V}_h$. *The following holds:*

$$|||\boldsymbol{\Pi}_h \left(\varphi \boldsymbol{w}_h\right)|||_{v,h} \le c_T |||\boldsymbol{w}_h|||_{v,h}.$$

Proof. We start with volume terms in the definition of $|||\cdot|||_{v,h}$ eq. (2.7a). Due to boundedness of Π_h, and using eq. (3.12b),

$$\sum_{\mathcal{K} \in \mathcal{T}_h} \|\Pi_h \left(\varphi w_h\right)\|_{\mathcal{K}}^2 + \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \, \|\overline{\nabla} \left(\Pi_h \left(\varphi w_h\right)\right)\|_{\mathcal{K}}^2 \le c \left(eT + \chi\right)^2 |||\boldsymbol{w}_h|||_{v,h}^2. \tag{4.24}$$

Next, the diffusive facet terms are bounded using a triangle inequality, eq. (3.12d), and boundedness of $\Pi_h^{\mathcal{F}}$:

$$\begin{aligned} &\sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1} \, \|\Pi_h \left(\varphi w_h\right) - \Pi_h^{\mathcal{F}} \left(\varphi \varkappa_h\right)\|_{\mathcal{Q}_{\mathcal{K}}}^2 \\ &\quad \le c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1} h_K \, \|\overline{\nabla} \left(\varphi w_h\right)\|_{\mathcal{K}}^2 + c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1} \, \|\varphi \left[\boldsymbol{w}_h\right]\|_{\mathcal{Q}_{\mathcal{K}}}^2 \le c \left(eT + \chi\right)^2 |||\boldsymbol{w}_h|||_{v,h}^2. \end{aligned} \tag{4.25}$$

For the Neumann boundary term in the definition of $|||\cdot|||_{v,h}$, consider first a single facet $F \in \partial \mathcal{E}_N$. Then,

$$\begin{aligned} &\| |\tfrac{1}{2} \beta \cdot n|^{1/2} \, \Pi_h^{\mathcal{F}} \left(\varphi \varkappa_h\right)\|_F \\ &\qquad \le \| |\tfrac{1}{2} \beta \cdot n|^{1/2} \left(\Pi_h^{\mathcal{F}} \left(\varphi \varkappa_h\right) - \Pi_h \left(\varphi w_h\right)\right)\|_F + \| |\tfrac{1}{2} \beta \cdot n|^{1/2} \, \Pi_h \left(\varphi w_h\right)\|_F := I + II. \end{aligned}$$

For term I, using that $|\beta\cdot n| \le (\max_{(t,x)\in F}|\beta\cdot n|)$, that $\Pi_h^{\mathcal{F}}\Pi_h(\varphi w_h) = \Pi_h(\varphi w_h)$ on F, boundedness of $\Pi_h^{\mathcal{F}}$ and a triangle inequality we have:

$$
\begin{aligned}
\big\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\big(\Pi_h^{\mathcal{F}}(\varphi\varkappa_h) - \Pi_h(\varphi w_h)\big)\big\|_F &\le (\tfrac{1}{2}\max_{(t,x)\in F}|\beta\cdot n|)^{1/2}\left\|\varphi\varkappa_h - \Pi_h(\varphi w_h)\right\|_F \\
&\le (\tfrac{1}{2}\max_{(t,x)\in F}|\beta\cdot n|)^{1/2}\left\|\varphi\varkappa_h - \varphi w_h\right\|_F \\
&\quad + (\tfrac{1}{2}\max_{(t,x)\in F}|\beta\cdot n|)^{1/2}\left\|(I-\Pi_h)(\varphi w_h)\right\|_F .
\end{aligned}
$$

Using that $|\varphi| \le eT + \chi$ and eq. (2.8) for the first term on the right-hand side, and using eq. (4.7d) and that $\delta t_{\mathcal{K}} \le h_K$ for the second term, we obtain

$$
I \le (eT+\chi)\left\|(\beta_s - \tfrac{1}{2}\beta\cdot n)^{1/2}(\varkappa_h - w_h)\right\|_F + c\left\|w_h\right\|_{\mathcal{K}_F}, \tag{4.26}
$$

where $\mathcal{K}_F$ is the space-time element of which F is a facet. Next, for term II, by a triangle inequality, using eq. (4.7d), that $\delta t_{\mathcal{K}} \le h_K$, $|\varphi| \le eT+\chi$, and eq. (2.8),

$$
\begin{aligned}
II &\le \left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}(I-\Pi_h)(\varphi w_h)\right\|_F + \left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\varphi(w_h - \varkappa_h)\right\|_F + (eT+\chi)\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\varkappa_h\right\|_F \\
&\le c\left\|w_h\right\|_{\mathcal{K}} + (eT+\chi)\left\| |\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}(w_h - \varkappa_h)\right\|_F + (eT+\chi)\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\varkappa_h\right\|_F .
\end{aligned}
$$

For a facet $F \in \partial\mathcal{E}_N$ we therefore conclude that

$$
\begin{aligned}
\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\Pi_h^{\mathcal{F}}(\varphi\varkappa_h)\right\|_F \le{}& c\,(eT+\chi)\left\|(\beta_s - \tfrac{1}{2}\beta\cdot n)^{1/2}(\varkappa_h - w_h)\right\|_F \\
&+ c\left\|w_h\right\|_{\mathcal{K}} + (eT+\chi)\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\varkappa_h\right\|_F .
\end{aligned} \tag{4.27}
$$

We find for the Neumann term in the definition of $|||\cdot|||_{v,h}$:

$$
\begin{aligned}
\sum_{F\in\partial\mathcal{E}_N}\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\Pi_h^{\mathcal{F}}(\varphi\varkappa_h)\right\|_F^2 \le{}& c\left\|\beta\right\|_{L^\infty(\mathcal{E})}\sum_{\mathcal{K}\in\mathcal{T}_h}\left\|w_h\right\|_{\mathcal{K}}^2 \\
&+ (eT+\chi)^2\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\left\|(\beta_s - \tfrac{1}{2}\beta\cdot n)^{1/2}(\varkappa_h - w_h)\right\|_{\partial\mathcal{K}}^2 + \sum_{F\in\partial\mathcal{E}_N}\left\| |\tfrac{1}{2}\beta\cdot n|^{1/2}\varkappa_h\right\|_F^2\Big) .
\end{aligned} \tag{4.28}
$$

Finally, we consider the advective facet terms in the definition of $|||\cdot|||_{v,h}$. On a single facet we have:

$$
\left\| |\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}(\Pi_h(\varphi w_h) - \Pi_h^{\mathcal{F}}(\varphi\varkappa_h))\right\|_F \le c(\max_{(t,x)\in F}|\beta\cdot n|)^{1/2}\left\|(\Pi_h(\varphi w_h) - \Pi_h^{\mathcal{F}}(\varphi\varkappa_h))\right\|_F .
$$

Using identical steps as used to find the bound for I in eq. (4.26), we find:

$$\begin{aligned}\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left(\Pi_h\left(\varphi w_h\right)-\Pi_h^{\mathcal{F}}\left(\varphi\varkappa_h\right)\right)\|_F&\\ \le c\left(eT+\chi\right)\|\left(\beta_s-\tfrac{1}{2}\beta\cdot n\right)^{1/2}\left(\varkappa_h-w_h\right)\|_F+c\,\|w_h\|_{\mathcal{K}}\,,&\end{aligned}$$

so that

$$\begin{aligned}\sum_{\mathcal{K}\in\mathcal{T}_h}\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left(\Pi_h\left(\varphi w_h\right)-\Pi_h^{\mathcal{F}}\left(\varphi\varkappa_h\right)\right)\|_{\partial\mathcal{K}}^2&\\ \le c\left(eT+\chi\right)^2\sum_{\mathcal{K}\in\mathcal{T}_h}\|\left(\beta_s-\tfrac{1}{2}\beta\cdot n\right)^{1/2}\left(\varkappa_h-w_h\right)\|_{\partial\mathcal{K}}^2+c\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\|w_h\|_{\mathcal{K}}^2\Big)\,.&\qquad(4.29)\end{aligned}$$

The result follows after collecting the bounds in eqs. (4.24), (4.25), (4.28) and (4.29). ☐

Lemma 4.4. *For any* $\boldsymbol{w}:=(w,\varkappa)\in L^2(\mathcal{E})\times L^2(\Gamma)$, *let* $\boldsymbol{\delta w}:=(w-\Pi_h w,\varkappa-\Pi_h^{\mathcal{F}}\varkappa)$. *The following holds for all* $\boldsymbol{w}_h\in\boldsymbol{V}_h$*:*

$$\begin{aligned}a_h(\boldsymbol{w}_h,\boldsymbol{\delta}\left(\varphi\boldsymbol{w}_h\right))\le & c_T\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon\,\|\overline{\nabla}w_h\|_{\mathcal{K}}^2+\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K^{-1}\,\|[\boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}}^2\Big)+\sum_{\mathcal{K}\in\mathcal{T}_h}\left(1/8+\delta t_{\mathcal{K}}\right)\|w_h\|_{\mathcal{K}}^2\\ &+c_T\Big(\sum_{F\in\partial\mathcal{E}_N}\|\left|\tfrac{1}{2}\beta\cdot n\right|^{1/2}\varkappa_h\|_F^2+\sum_{\mathcal{K}\in\mathcal{T}_h}\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}[\boldsymbol{w}_h]\|_{\partial\mathcal{K}}^2\Big)\,.\end{aligned}$$

Proof. Let $z\in H^1(\mathcal{T}_h)$ and $\varpi\in L^2(\mathcal{F}_h)$ such that $\varpi|_{\partial\mathcal{E}_D}=0$. Let $\boldsymbol{z}:=(z,\varpi)$. Integrating $(\beta w_h,\nabla z)_{\mathcal{T}_h}$ by parts and using that $\langle(\beta\cdot n)\,\varkappa_h,\varpi\rangle_{\partial\mathcal{T}_h}=\langle(\beta\cdot n)\,\varkappa_h,\varpi\rangle_{\partial\mathcal{E}_N}$, because $\varkappa_h$ and ϖ are single-valued on Γ and zero on $\partial\mathcal{E}_D$, we have:

$$\begin{aligned}a_{h,c}(\boldsymbol{w}_h,\boldsymbol{z})=&\left(\beta\cdot\nabla w_h,z\right)_{\mathcal{T}_h}-\langle\tfrac{1}{2}\left(\beta\cdot n\right)[\boldsymbol{w}_h]\,,z\rangle_{\partial\mathcal{T}_h}\\ &+\langle\big(\beta_s-\tfrac{1}{2}\left(\beta\cdot n\right)\big)[\boldsymbol{w}_h]\,,z\rangle_{\partial\mathcal{T}_h}-\langle\beta_s\,[\boldsymbol{w}_h]\,,\varpi\rangle_{\partial\mathcal{T}_h}+\langle\tfrac{1}{2}\big(|\beta\cdot n|-\beta\cdot n\big)\,\varkappa_h,\varpi\rangle_{\partial\mathcal{E}_N}.\end{aligned}\qquad(4.30)$$

At this point, note that $\boldsymbol{\delta}\left(\varphi\boldsymbol{w}_h\right)=\boldsymbol{\delta}\left(eT\exp(-t/T)\boldsymbol{w}_h\right)$ because $\boldsymbol{\delta}\left(\chi\boldsymbol{w}_h\right)=0$. Furthermore, let $\beta_0=(1,\bar{\beta}_0)$. By definition of Π_h, the following vanishes

$$\left(\beta_0\cdot\nabla w_h,\left(I-\Pi_h\right)\left(eT\exp(-t/T)\right)w_h\right)_{\mathcal{T}_h}=0.$$

From eq. (4.30), with $\boldsymbol{z}=\boldsymbol{\delta}\left(eT\exp(-t/T)\boldsymbol{w}_h\right)$, we now find that:

$$\begin{aligned}a_{h,c}\big(\boldsymbol{w}_h,\boldsymbol{\delta}\left(eT\exp(-t/T)\right)\boldsymbol{w}_h\big)=&\left(\left(\beta-\beta_0\right)\cdot\nabla w_h,\left(I-\Pi_h\right)\left(eT\exp(-t/T)w_h\right)\right)_{\mathcal{T}_h}\\ &-\langle\tfrac{1}{2}\left(\beta\cdot n\right)[\boldsymbol{w}_h]\,,\left(I-\Pi_h\right)\left(eT\exp(-t/T)w_h\right)\rangle_{\partial\mathcal{T}_h}\\ &+\langle\left(\beta_s-\tfrac{1}{2}\left(\beta\cdot n\right)\right)[\boldsymbol{w}_h]\,,\left(I-\Pi_h\right)\left(eT\exp(-t/T)w_h\right)\rangle_{\partial\mathcal{T}_h}\\ &+\langle\tfrac{1}{2}\left(|\beta\cdot n|-\beta\cdot n\right)\varkappa_h,\left(I-\Pi_h^{\mathcal{F}}\right)\left(eT\exp(-t/T)\varkappa_h\right)\rangle_{\partial\mathcal{E}_N}\\ =:&M_1+M_2+M_3+M_4,\end{aligned}$$

where, by definitions of $\Pi_h^{\mathcal{F}}$ and β_s, $\langle \beta_s \left[\boldsymbol{w}_h\right], (I - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)\varkappa_h\right)\rangle_{\partial\mathcal{T}_h} = 0$. We will now bound each of the terms M_i, $i = 1, \dots, 4$ separately.

We observe that $(\beta - \beta_0)\cdot\nabla w_h = (\bar{\beta} - \bar{\beta}_0)\cdot\overline{\nabla} w_h$ because the first components of β and β_0 are 1. We then bound M_1 using the Cauchy–Schwarz inequality, eq. (4.4), eq. (4.7a), and eq. (3.7b):

$$M_1 \le \sum_{\mathcal{K}\in\mathcal{T}_h} c h_K h_K^{-1} \left\| w_h \right\|_{\mathcal{K}} \delta t_{\mathcal{K}} \left\| w_h \right\|_{\mathcal{K}} = c \sum_{\mathcal{K}\in\mathcal{T}_h} \delta t_{\mathcal{K}} \left\| w_h \right\|_{\mathcal{K}}^2 . \tag{4.31}$$

We proceed with bounding M_2 and M_3. Using the Cauchy–Schwarz inequality, eq. (2.8), eq. (4.7d), and eq. (4.7e), we find that

$$\begin{aligned} M_2 + M_3 \le c \sum_{\mathcal{K}\in\mathcal{T}_h} \delta t_{\mathcal{K}} h_K^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} \left[\boldsymbol{w}_h\right] \right\|_{\mathcal{Q}_{\mathcal{K}}} \left\| w_h \right\|_{\mathcal{K}} \\ + c \sum_{\mathcal{K}\in\mathcal{T}_h} \delta t_{\mathcal{K}}{}^{1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} \left[\boldsymbol{w}_h\right] \right\|_{\mathcal{R}_{\mathcal{K}}} \left\| w_h \right\|_{\mathcal{K}} . \end{aligned}$$

Since $\delta t_{\mathcal{K}} \le h_K$, $\delta t_{\mathcal{K}} h_K^{-1/2}$ can be bounded by 1. Therefore, applying Young's inequality,

$$\begin{aligned} M_2 + M_3 \le& c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\left(\beta\cdot n\right)\right|^{1/2} \left[\boldsymbol{w}_h\right] \right\|_{\partial\mathcal{K}} \left\| w_h \right\|_{\mathcal{K}} \\ \le& \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| w_h \right\|_{\mathcal{K}}^2 + c\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\left(\beta\cdot n\right)\right|^{1/2} \left[\boldsymbol{w}_h\right] \right\|_{\partial\mathcal{K}}^2 . \end{aligned} \tag{4.32}$$

For M_4, we first apply the Cauchy–Schwarz inequality and the triangle inequality:

$$M_4 \le c \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F \left(\left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F + \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \Pi_h^{\mathcal{F}}\left(eT\exp(-t/T)\varkappa_h\right) \right\|_F \right) . \tag{4.33}$$

The second term in parentheses on the right-hand side of eq. (4.33) is bounded following identical steps in showing eq. (4.27). Applying also Young's inequality, and denoting by

$\mathcal{K}_F$ the space-time element of which F is a facet,

$$\begin{aligned}
M_4 \leq & c \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F^2 \\
& + c \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F \left(c_T \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F + c_T \left\| \left(\beta_s - \tfrac{1}{2}\beta\cdot n\right)^{1/2} [\boldsymbol{w}_h] \right\|_{\partial\mathcal{K}} + \|w_h\|_{\mathcal{K}_F} \right) \\
= & (c + c_T) \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F^2 \\
& + c_T \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F \left\| \left(\beta_s - \tfrac{1}{2}\beta\cdot n\right)^{1/2} [\boldsymbol{w}_h] \right\|_{\partial\mathcal{K}} + c \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F \|w_h\|_{\mathcal{K}_F} \\
\leq & \tfrac{\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \|w_h\|_{\mathcal{K}}^2 + (c_T + c\delta^{-1}) \sum_{F\in\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h \right\|_F^2 + c_T \sum_{F\in\partial\mathcal{E}_N} \left\| \left(\beta_s - \tfrac{1}{2}\beta\cdot n\right)^{1/2} [\boldsymbol{w}_h] \right\|_{\partial\mathcal{K}}^2 .
\end{aligned} \tag{4.34}$$

We proceed with the diffusive term $a_{h,d}$. With test function $\boldsymbol{z} = \boldsymbol{\delta}\left(eT\exp(-t/T)\boldsymbol{w}_h\right)$,

$$\begin{aligned}
a_{h,d}\left(\boldsymbol{w}_h, \boldsymbol{\delta}\left(eT\exp(-t/T)\right)\boldsymbol{w}_h\right) = & \left(\varepsilon\overline{\nabla}w_h, \overline{\nabla}\left((I-\Pi_h)\left(eT\exp(-t/T)\right)w_h\right)\right)_{\mathcal{T}_h} \\
& - \langle \varepsilon\alpha h_K^{-1}[\boldsymbol{w}_h], (\Pi_h - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)w_h\right)\rangle_{\mathcal{Q}_h} \\
& + \langle \varepsilon\alpha h_K^{-1}[\boldsymbol{w}_h], (I - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)[\boldsymbol{w}_h]\right)\rangle_{\mathcal{Q}_h} \\
& - \langle \varepsilon[\boldsymbol{w}_h], \overline{\nabla}_{\bar{n}}\left((I-\Pi_h)\left(eT\exp(-t/T)w_h\right)\right)\rangle_{\mathcal{Q}_h} \\
& + \langle \varepsilon\overline{\nabla}_{\bar{n}}w_h, (\Pi_h - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)w_h\right)\rangle_{\mathcal{Q}_h} \\
& - \langle \varepsilon\overline{\nabla}_{\bar{n}}w_h, (I - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)[\boldsymbol{w}_h]\right)\rangle_{\mathcal{Q}_h} \\
=: & M_5 + M_6 + M_7 + M_8 + M_9 + M_{10}.
\end{aligned}$$

To bound M_5 we use the Cauchy–Schwarz inequality, eq. (4.7b), the assumption that $\delta t_{\mathcal{K}} \leq h_K$, and Young's inequality:

$$M_5 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla}w_h\|_{\mathcal{K}} \delta t_{\mathcal{K}} h_K^{-1} \|w_h\|_{\mathcal{K}} \leq \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \|w_h\|_{\mathcal{K}}^2 + c\varepsilon\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla}w_h\|_{\mathcal{K}}^2 . \tag{4.35}$$

To bound M_6 we apply the Cauchy–Schwarz inequality, eq. (3.12d), and Young's inequality:

$$\begin{aligned}
M_6 \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \|[\boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}} h_K^{1/2} \|\overline{\nabla}\left(eT\exp(-t/T)w_h\right)\|_{\mathcal{K}} \\
\leq & c_T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \|[\boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}}^2 + c_T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla}w_h\|_{\mathcal{K}}^2 .
\end{aligned} \tag{4.36}$$

M_7 can be bounded using the boundedness of $\Pi_h^{\mathcal{F}}$:

$$M_7 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}} \left\|(I - \Pi_h^{\mathcal{F}})\left(eT\exp(-t/T)\left[\boldsymbol{w}_h\right]\right)\right\|_{\mathcal{Q}_\mathcal{K}} \leq c_T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}}^2 . \tag{4.37}$$

Terms M_9 and M_{10} are bounded in a similar way as M_6 and M_7, and using eq. (3.7c):

$$M_9 + M_{10} \leq c_T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}}^2 + c_T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} w_h\right\|_{\mathcal{K}}^2 . \tag{4.38}$$

Finally, we bound M_8 using the Cauchy–Schwarz inequality, eq. (4.7c), the assumption that $\delta t_\mathcal{K} \leq h_K$, and Young's inequality:

$$\begin{aligned} M_8 &\leq \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}} \left\|\overline{\nabla}\left((I - \Pi_h)\left(eT\exp(-t/T)w_h\right)\right)\right\|_{\mathcal{Q}_\mathcal{K}} \\ &\leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}} \delta t_\mathcal{K} h_K^{-3/2} \left\|w_h\right\|_\mathcal{K} \\ &\leq \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|w_h\right\|_\mathcal{K}^2 + c\varepsilon\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}}^2 . \end{aligned} \tag{4.39}$$

Collecting eqs. (4.31), (4.32) and (4.34) to (4.39) we find that

$$\begin{aligned} a_h(\boldsymbol{w}_h, \boldsymbol{\delta}\left(\varphi\boldsymbol{w}_h\right)) \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \delta t_\mathcal{K} \left\|w_h\right\|_\mathcal{K}^2 + 2\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|w_h\right\|_\mathcal{K}^2 \\ & + (c_T + c\delta^{-1}) \Bigg(\sum_{F\in\partial\mathcal{E}_N} \left\|\left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h\right\|_F^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|\left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{w}_h\right]\right\|_{\partial\mathcal{K}}^2 \Bigg) \\ & + (c_T + c\varepsilon\delta^{-1}) \Bigg(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} w_h\right\|_\mathcal{K}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}}^2 \Bigg) . \end{aligned}$$

The result follows by choosing $\delta = 1/16$. □

We are now ready to prove eq. (4.2a).

Proof of eq. (4.2a). Choose $\delta t_0 = 1/8$. When $\delta t_\mathcal{K} \leq \delta t_0$ for all $\mathcal{K} \in \mathcal{T}_h$ we find, by combining lemmas 4.1 and 4.4,

$$\begin{aligned} a_h(\boldsymbol{w}_h, \boldsymbol{\Pi}_h\left(\varphi\boldsymbol{w}_h\right)) \geq & \left(\tfrac{1}{4}(T + \chi) - c_T\right)\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} w_h\right\|_\mathcal{K}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|_{\mathcal{Q}_\mathcal{K}}^2 \Big) + \tfrac{1}{4} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|w_h\right\|_\mathcal{K}^2 \\ & + (T + \chi - c_T)\Big(\sum_{F\in\partial\mathcal{E}_N} \left\|\left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} \varkappa_h\right\|_F^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|\left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{w}_h\right]\right\|_{\partial\mathcal{K}}^2 \Big) . \end{aligned}$$

Choosing χ to satisfy $\chi \geq 4c_T$ in addition to the conditions of lemma 4.1, we obtain

$$a_h(\boldsymbol{w}_h, \boldsymbol{\Pi}_h\,(\varphi \boldsymbol{w}_h)) \geq \tfrac{1}{4}|\!|\!|\boldsymbol{w}_h|\!|\!|_{v,h}^2 \geq c_T^{-1}|\!|\!|\boldsymbol{w}_h|\!|\!|_{v,h}|\!|\!|\boldsymbol{\Pi}_h(\varphi \boldsymbol{w}_h)|\!|\!|_{v,h}, \tag{4.40}$$

where the second inequality is due to lemma 4.3. We therefore conclude eq. (4.2a). □

4.1.2 The inf-sup condition with respect to $|\!|\!| \cdot |\!|\!|_{s,h}$

The following boundedness result will be useful in the proof of stability eq. (4.2b)

$$|a_{h,d}(\boldsymbol{u}_h, \boldsymbol{v}_h)| \leq c|\!|\!|\boldsymbol{u}_h|\!|\!|_{v,h}|\!|\!|\boldsymbol{v}_h|\!|\!|_{v,h} \qquad \forall \boldsymbol{u}_h, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{4.41}$$

To prove this, we have by the Cauchy–Schwarz inequality and eq. (3.7c),

$$\begin{aligned}
&|a_{h,d}(\boldsymbol{u}_h, \boldsymbol{v}_h)| \\
&\leq \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2}\,\|\overline{\nabla} u_h\|_{\mathcal{K}}\,\varepsilon^{1/2}\,\|\overline{\nabla} v_h\|_{\mathcal{K}} + \alpha \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2}\,\|[\boldsymbol{u}_h]\|_{\mathcal{Q}_{\mathcal{K}}}\,\varepsilon^{1/2} h_K^{-1/2}\,\|[\boldsymbol{v}_h]\|_{\mathcal{Q}_{\mathcal{K}}} \\
&+ c\,\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2}\,\|[\boldsymbol{u}_h]\|_{\mathcal{Q}_{\mathcal{K}}}\,\varepsilon^{1/2}\,\|\overline{\nabla} v_h\|_{\mathcal{K}} + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2}\,\|\overline{\nabla} u_h\|_{\mathcal{K}}\,\varepsilon^{1/2} h_K^{-1/2}\,\|[\boldsymbol{v}_h]\|_{\mathcal{Q}_{\mathcal{K}}}\Big)\,,
\end{aligned}$$

after which eq. (4.41) follows immediately.

To prove eq. (4.2b), we first construct the test function $\boldsymbol{y}_h := (y_h, \vartheta_h)$ as a function of $\boldsymbol{w}_h = (w_h, \varkappa_h) \in \boldsymbol{V}_h$. The elemental test function y_h is defined as:

$$y_h := \tau_\varepsilon \partial_t w_h. \tag{4.42a}$$

To define the facet test function ϑ_h we consider four different sets of facets. First we consider facets F in $\partial\mathcal{K}_1 \cap \partial\mathcal{K}_2 \cap \mathcal{Q}_h^i$ and such that there is no difference in the refinement level in the time direction between $\mathcal{K}_1$ and $\mathcal{K}_2$. This means that $\delta t_{\mathcal{K}_1} = \delta t_{\mathcal{K}_2} =: \delta t_{\mathcal{K}}$ and, since $\mathcal{K}_1$ and $\mathcal{K}_2$ must come from the same space-time slab, $\Delta t_{\mathcal{K}_1} = \Delta t_{\mathcal{K}_2} =: \Delta t_{\mathcal{K}}$. We then define:

$$\vartheta_h := \begin{cases} \Delta t_{\mathcal{K}} \partial_t \varkappa_h, & \delta t_{\mathcal{K}} \leq h_{K_1} \leq \varepsilon, \delta t_{\mathcal{K}} \leq h_{K_2} \leq \varepsilon, \\ \Delta t_{\mathcal{K}} \varepsilon^{1/2} \partial_t \varkappa_h, & \delta t_{\mathcal{K}} \leq \varepsilon < h_{K_1}, \delta t_{\mathcal{K}} \leq \varepsilon < h_{K_2}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.42b}$$

We next consider facets F in $\partial\mathcal{K}_1 \cap \partial\mathcal{K}_2 \cap \mathcal{Q}_h^i$ and such that there is one level of refinement difference between $\mathcal{K}_1$ and $\mathcal{K}_2$ in the time direction. Without loss of generality, we assume that $2\delta t_{\mathcal{K}_1} = \delta t_{\mathcal{K}_2}$. Furthermore, since $\mathcal{K}_1$ and $\mathcal{K}_2$ must come from the same space-time slab, $\Delta t_{\mathcal{K}_1} = \Delta t_{\mathcal{K}_2} := \Delta t_{\mathcal{K}}$. We then define:

$$\vartheta_h := \begin{cases} \Delta t_{\mathcal{K}} \partial_t \varkappa_h, & \delta t_{\mathcal{K}_1} \le h_{K_1} \le \varepsilon, \delta t_{\mathcal{K}_2} \le h_{K_2} \le \varepsilon, \\ \Delta t_{\mathcal{K}} \varepsilon^{1/2} \partial_t \varkappa_h, & \delta t_{\mathcal{K}_1} \le \varepsilon < h_{K_1}, \delta t_{\mathcal{K}_2} \le \varepsilon < h_{K_2}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.42c}$$

For facets F in $\partial\mathcal{K} \cap \mathcal{Q}_h^b$, we define:

$$\vartheta_h := \begin{cases} \Delta t_{\mathcal{K}} \partial_t \varkappa_h, & \delta t_{\mathcal{K}} \le h_K \le \varepsilon, \\ \Delta t_{\mathcal{K}} \varepsilon^{1/2} \partial_t \varkappa_h, & \delta t_{\mathcal{K}} \le \varepsilon < h_K, \\ 0, & \text{otherwise.} \end{cases} \tag{4.42d}$$

Finally, for facets F in $\mathcal{R}_h$, we define:

$$\vartheta_h := 0. \tag{4.42e}$$

We observe from definition eq. (4.42) that $\vartheta_h \equiv 0$ on $\partial\mathcal{T}_h^c$, which denotes the set of element boundaries of space-time elements in $\mathcal{T}_h^c$. Furthermore, for any space-time element $\mathcal{K} \in \mathcal{T}_h^{dx} := \mathcal{T}_h^d \cup \mathcal{T}_h^x$, we introduce $\mathcal{Q}_{\mathcal{K}}^0$ to denote those $\mathcal{Q}$-faces on which ϑ_h is prescribed in eqs. (4.42b) and (4.42c) to be zero. We will define $\mathcal{Q}_h^0 := \cup_{\mathcal{K}\in\mathcal{T}_h} \mathcal{Q}_{\mathcal{K}}^0$. Consider now $\mathcal{K} \in \mathcal{T}_h^{dx,0}$, which denotes the set of space-time elements in $\mathcal{T}_h^{dx}$ for which $\mathcal{Q}_{\mathcal{K}}^0 \ne \emptyset$. Then, there exists a $\mathcal{K}'$ such that $\partial\mathcal{K}' \cap \partial\mathcal{K} \ne \emptyset$ and that either $h_K \le \varepsilon \le h_{K'}$ (or $h_{K'} \le \varepsilon \le h_K$), or $\delta t_{\mathcal{K}} \le \varepsilon \le \delta t_{\mathcal{K}'}$ (or $\delta t_{\mathcal{K}'} \le \varepsilon \le \delta t_{\mathcal{K}}$). For the former case, since spatial elements are shape-regular and the difference of refinement levels in the spatial direction between two adjacent space-time elements is at most one, we have $c^{-1}h_{K'} \le h_K \le c h_{K'}$. If the latter case holds, since $\delta t_{\mathcal{K}} = \frac{1}{2}\delta t_{\mathcal{K}'}$, it holds that $\delta t_{\mathcal{K}} \sim \varepsilon$. Therefore,

$$c^{-1}h_K \le \varepsilon \le c h_K \quad \text{or} \quad c^{-1}\delta t_{\mathcal{K}} \le \varepsilon \le c\,\delta t_{\mathcal{K}} \qquad \forall \mathcal{K} \in \mathcal{T}_h^{dx,0}. \tag{4.43}$$

Lemmas 4.5 and 4.6 will be used to prove eq. (4.2b). The proofs of these lemmas will repeatedly use the following set of inequalities: For all $\mathcal{K} \in \mathcal{T}_h$,

$$h_K^{-1} \le \delta t_{\mathcal{K}}^{-1}, \quad \Delta t_{\mathcal{K}} \le c\,\delta t_{\mathcal{K}}, \quad \tau_\varepsilon \le \Delta t_{\mathcal{K}}, \quad \varepsilon \le 1. \tag{4.44}$$

Lemma 4.5. *Assume that* $\delta t_{\mathcal{K}} \leq h_K$ *for all space-time elements* $\mathcal{K} \in \mathcal{T}_h$*. Let* $\boldsymbol{w}_h = (w_h, \varkappa_h) \in \boldsymbol{V}_h$ *and let* $\boldsymbol{y}_h$ *be defined by eq. (4.42). The following holds:*

$$|||\boldsymbol{y}_h|||_{s,h} \leq c|||\boldsymbol{w}_h|||_{s,h}. \tag{4.45}$$

Proof. We start with the volume terms of $|||\cdot|||_{s,h}$. Using eq. (3.7a) and eq. (4.44), we have:

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \|y_h\|_{\mathcal{K}}^2 \leq \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t w_h\|_{\mathcal{K}}^2 . \tag{4.46}$$

For the diffusive volume term, using commutativity of $\overline{\nabla}$ and ∂_t, eq. (3.7a) and eq. (4.44):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla} y_h\|_{\mathcal{K}}^2 = \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon^2 \varepsilon \|\partial_t (\overline{\nabla} w_h)\|_{\mathcal{K}}^2 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \|\overline{\nabla} w_h\|_{\mathcal{K}}^2 . \tag{4.47}$$

The time-derivative volume term is treated similarly, using eq. (4.44):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t y_h\|_{\mathcal{K}}^2 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t w_h\|_{\mathcal{K}}^2 . \tag{4.48}$$

For the diffusive facet term in the definition of $|||\cdot|||_{s,h}$, we use lemma 3.1, eq. (3.7c), that ϑ_h vanishes on $\partial\mathcal{T}_h^c$, that ϑ_h vanishes on $\mathcal{Q}_{\mathcal{K}}^0$ when $\mathcal{K} \in \mathcal{T}_h^{dx}$, that $\varepsilon \leq \delta t_{\mathcal{K}}$ and $\varepsilon h_K^{-2}\tau_\varepsilon \leq 1$ on $\mathcal{T}_h^c$, and that $\varepsilon h_K^{-2}\tau_\varepsilon \leq c$ on $\mathcal{T}_h^{dx,0}$ due to eq. (4.43):

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \|[\boldsymbol{y}_h]\|_{\mathcal{Q}_{\mathcal{K}}}^2 \\
=&\sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \varepsilon h_K^{-1}\tau_\varepsilon^2 \|[\partial_t \boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}\setminus\mathcal{Q}_{\mathcal{K}}^0}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \varepsilon h_K^{-1}\tau_\varepsilon^2 \|\partial_t w_h\|_{\mathcal{Q}_{\mathcal{K}}^0}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h^{c}} \varepsilon h_K^{-1}\tau_\varepsilon^2 \|\partial_t w_h\|_{\mathcal{Q}_{\mathcal{K}}}^2 \\
\leq& c\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \|[\boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}}^2 + c\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t w_h\|_{\mathcal{K}}^2 .
\end{aligned} \tag{4.49}$$

For the advective facet term, using lemma 3.1, eq. (3.7d), eq. (4.44), that $\tilde{\varepsilon}^2 \leq \varepsilon h_K^{-1}$ on $\mathcal{K} \in \mathcal{T}_h^{dx}$ since $h_K \leq \varepsilon$ on $\mathcal{T}_h^d$ and $\tilde{\varepsilon} = \varepsilon^{1/2}$ on $\mathcal{T}_h^x$:

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2} [\boldsymbol{y}_h] \right\|_{\partial\mathcal{K}}^2 \\
\leq& c\Big(\sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon^2 \|[\partial_t \boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}\setminus\mathcal{Q}_{\mathcal{K}}^0}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon^2 \|\partial_t w_h\|_{\mathcal{R}_{\mathcal{K}}\cup\mathcal{Q}_{\mathcal{K}}^0}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h^{c}} \tau_\varepsilon^2 \|\partial_t w_h\|_{\partial\mathcal{K}}^2 \Big) \\
\leq& c\Big(\sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \varepsilon h_K^{-1} \|[\boldsymbol{w}_h]\|_{\mathcal{Q}_{\mathcal{K}}\setminus\mathcal{Q}_{\mathcal{K}}^0}^2 \Big) + c\sum_{\mathcal{K}\in\mathcal{T}_h^{c}} \tau_\varepsilon \|\partial_t w_h\|_{\mathcal{K}}^2 .
\end{aligned} \tag{4.50}$$

Finally, the Neumann boundary term is bounded using the triangle inequality, Young's inequality, lemma 3.1, eq. (3.7c), eq. (4.44), and that $h_K \le \varepsilon$ for $\mathcal{K} \in \mathcal{T}_h^d$:

$$\begin{aligned} \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \left\| \left|\tfrac{1}{2}\beta \cdot n\right|^{1/2} \vartheta_h \right\|^2_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} \le& c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon^2 \left\|[\partial_t \boldsymbol{w}_h]\right\|^2_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} + c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon^2 \left\|\partial_t w_h\right\|^2_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} \\ \le& c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \varepsilon h_K^{-1} \left\|[\boldsymbol{w}_h]\right\|^2_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} + c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon \left\|\partial_t w_h\right\|^2_{\mathcal{K}}. \end{aligned} \tag{4.51}$$

Combining eqs. (4.46) to (4.51) yields eq. (4.45). ☐

Lemma 4.6. *Assume that $\delta t_{\mathcal{K}} \le h_K$ for all space-time elements $\mathcal{K} \in \mathcal{T}_h$. Let $\boldsymbol{w}_h = (w_h, \varkappa_h) \in \boldsymbol{V}_h$, let $\boldsymbol{y}_h$ be defined as in eq. (4.42), and let $\boldsymbol{\Pi}_h(\varphi \boldsymbol{w}_h)$ be defined as in lemma 4.3. There exists a positive constant c such that*

$$|||\boldsymbol{w}_h|||^2_{s,h} \le a_h(\boldsymbol{w}_h, 2(\boldsymbol{y}_h + c\boldsymbol{\Pi}_h(\varphi\boldsymbol{w}_h))).$$

Proof. Let us first note that ϑ_h vanishes on $\mathcal{R}_h$ and $\partial\mathcal{T}_h^c$. Therefore, defining $\mathcal{Q}_h^{dx} := \partial\mathcal{T}_h^{dx} \cap \mathcal{Q}_h$, we find after some algebraic manipulation that:

$$\begin{aligned} a_{h,c}(\boldsymbol{w}_h, \boldsymbol{y}_h) = & \left(\nabla \cdot (\beta w_h), y_h\right)_{\mathcal{T}_h} + \langle (\beta_s - \tfrac{1}{2}\beta \cdot n)[\boldsymbol{w}_h], y_h\rangle_{\partial\mathcal{T}_h} - \langle \tfrac{1}{2}\beta \cdot n [\boldsymbol{w}_h], y_h\rangle_{\partial\mathcal{T}_h} \\ & - \langle \tfrac{1}{2}(|\beta\cdot n| - \beta\cdot n)\varkappa_h, [\boldsymbol{y}_h]\rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h^{dx}} + \langle \tfrac{1}{2}(|\beta\cdot n| - \beta\cdot n)\varkappa_h, y_h\rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h^{dx}} \\ & + \langle \beta_s [\boldsymbol{w}_h], [\boldsymbol{y}_h]\rangle_{\mathcal{Q}_h^{dx}\setminus\mathcal{Q}_h^0} - \langle \beta_s [\boldsymbol{w}_h], y_h\rangle_{\mathcal{Q}_h^{dx}\setminus\mathcal{Q}_h^0}. \end{aligned}$$

Furthermore, since $\left(\nabla \cdot (\beta w_h), y_h\right)_{\mathcal{T}_h} = (\partial_t w_h, \tau_\varepsilon \partial_t w_h)_{\mathcal{T}_h} + (\overline{\nabla} \cdot (\bar{\beta} w_h), \tau_\varepsilon \partial_t w_h)_{\mathcal{T}_h}$, we find that

$$\begin{aligned} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t w_h\right\|^2_{\mathcal{K}} =& a_h(\boldsymbol{w}_h, \boldsymbol{y}_h) - a_{h,d}(\boldsymbol{w}_h, \boldsymbol{y}_h) - (\overline{\nabla} \cdot (\bar{\beta} w_h), \tau_\varepsilon \partial_t w_h)_{\mathcal{T}_h} \\ & - \langle (\beta_s - \tfrac{1}{2}\beta \cdot n)[\boldsymbol{w}_h], \tau_\varepsilon \partial_t w_h\rangle_{\partial\mathcal{T}_h} + \langle \tfrac{1}{2}\beta\cdot n [\boldsymbol{w}_h], \tau_\varepsilon \partial_t w_h\rangle_{\partial\mathcal{T}_h} \\ & - \langle \beta_s [\boldsymbol{w}_h], \tau_\varepsilon [\partial_t \boldsymbol{w}_h]\rangle_{\mathcal{Q}_h^{dx}\setminus\mathcal{Q}_h^0} + \langle \beta_s [\boldsymbol{w}_h], \tau_\varepsilon \partial_t w_h\rangle_{\mathcal{Q}_h^{dx}\setminus\mathcal{Q}_h^0} \\ & + \langle \tfrac{1}{2}(|\beta\cdot n| - \beta\cdot n)\varkappa_h, \tau_\varepsilon [\partial_t \boldsymbol{w}_h]\rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h^{dx}} \\ & - \langle \tfrac{1}{2}(|\beta\cdot n| - \beta\cdot n)\varkappa_h, \tau_\varepsilon \partial_t w_h\rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h^{dx}} \\ =& a_h(\boldsymbol{w}_h, \boldsymbol{y}_h) - a_{h,d}(\boldsymbol{w}_h, \boldsymbol{y}_h) - (\overline{\nabla} \cdot (\bar{\beta} w_h), \tau_\varepsilon \partial_t w_h)_{\mathcal{T}_h} \\ & + T_1 + T_2 + T_3 + T_4 + T_5 + T_6. \end{aligned} \tag{4.52}$$

We will bound the last eight terms on the right-hand side of the above equation. First, by eq. (4.41), lemma 4.5, and Young's inequality, we have

$$
\begin{aligned}
a_{h,d}(\boldsymbol{w}_h, \boldsymbol{y}_h) \leq & c |\!|\!| \boldsymbol{w}_h |\!|\!|_{v,h} |\!|\!| \boldsymbol{w}_h |\!|\!|_{s,h} \leq c \, ((\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2)^{1/2} + |\!|\!| \boldsymbol{w}_h |\!|\!|_{v,h}) \, |\!|\!| \boldsymbol{w}_h |\!|\!|_{v,h} \\
\leq & c \, (1 + \delta^{-1}) \, |\!|\!| \boldsymbol{w}_h |\!|\!|_{v,h}^2 + \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2 .
\end{aligned}
\tag{4.53}
$$

Next, using the Cauchy–Schwarz and Young's inequalities, $\delta t_{\mathcal{K}} \leq \varepsilon$ for $\mathcal{K} \in \mathcal{T}_h^{dx}$, eq. (3.7b) and eq. (4.44):

$$
\begin{aligned}
(\overline{\nabla} \cdot (\bar{\beta} w_h) \, , \tau_\varepsilon \partial_t w_h)_{\mathcal{T}_h} \leq & \frac{\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2 + c\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \overline{\nabla} w_h \right\|_{\mathcal{K}}^2 \\
\leq & \frac{\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2 + c\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} w_h \right\|_{\mathcal{K}}^2 + c\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| w_h \right\|_{\mathcal{K}}^2 .
\end{aligned}
\tag{4.54}
$$

T_1 and T_2 can be bounded using the Cauchy–Schwarz inequality, $\frac{1}{2} \left| \beta \cdot n \right| \leq \left| \beta_s - \frac{1}{2} \beta \cdot n \right|$ for all $F \in \partial\mathcal{T}_h$, eq. (3.7d), eq. (4.44), and Young's inequality:

$$
T_1 + T_2 \leq \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2 + c\delta^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{w}_h] \right\|_{\partial\mathcal{K}}^2 . \tag{4.55}
$$

Similarly T_3 and T_4 are bounded using the Cauchy–Schwarz inequality, lemma 3.1, eq. (3.7c), eq. (2.8), $\delta t_{\mathcal{K}} \leq h_K \leq \varepsilon$ for $\mathcal{K} \in \mathcal{T}_h^d$, eq. (4.44), and Young's inequality. Note that we also make use of $\tilde{\varepsilon} \leq \varepsilon^{1/2} h_K^{-1/2}$ on $\mathcal{T}_h^{dx}$ since on $\mathcal{T}_h^d$, $\tilde{\varepsilon} = 1$ and $h_K \leq \varepsilon$ while on $\mathcal{T}_h^x$, $\tilde{\varepsilon} = \varepsilon^{1/2}$ and $h_K \leq 1$:

$$
\begin{aligned}
T_3 + T_4 \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \tau_\varepsilon \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} ((\delta t_{\mathcal{K}}^{-1} + h_K^{-1}) \left\| [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} + h_K^{-1/2} \left\| \partial_t w_h \right\|_{\mathcal{K}}) \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} (\tilde{\varepsilon} \left\| [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} + \Delta t_{\mathcal{K}}^{1/2} \tilde{\varepsilon} \left\| \partial_t w_h \right\|_{\mathcal{K}}) \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h^{dx}} \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} (\varepsilon^{1/2} h_K^{-1/2} \left\| [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K} \backslash \mathcal{Q}_\mathcal{K}^0} + \tau_\varepsilon^{1/2} \left\| \partial_t w_h \right\|_{\mathcal{K}}) \\
\leq & \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t w_h \right\|_{\mathcal{K}}^2 + c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\| [\boldsymbol{w}_h] \right\|_{\mathcal{Q}_\mathcal{K}}^2 + c \, (1 + \delta^{-1}) \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{w}_h] \right\|_{\partial\mathcal{K}}^2 .
\end{aligned}
\tag{4.56}
$$

Similarly, to bound T_5 and T_6, we use the Cauchy–Schwarz inequality, lemma 3.1, eq. (3.7c), that $\delta t_{\mathcal{K}} \le h_K \le \varepsilon$ for $\mathcal{K} \in \mathcal{T}_h^d$, eq. (4.44), $\tilde{\varepsilon} \le \varepsilon^{1/2} h_K^{-1/2}$ on $\mathcal{T}_h^{dx}$ and Young's inequality:

$$T_5+T_6 \le \tfrac{1}{2}\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t w_h\right\|_{\mathcal{K}}^2 + c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|[\![\boldsymbol{w}_h]\!]\right\|_{\mathcal{Q}_{\mathcal{K}}}^2 + c\left(1+\delta^{-1}\right) \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\tfrac{1}{2}\beta \cdot n\right|^{1/2} \varkappa_h \right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2 . \tag{4.57}$$

Combining eqs. (4.52) to (4.57) and choosing $\delta = 1/5$ we obtain

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t w_h\right\|_{\mathcal{K}}^2 \le a_h(\boldsymbol{w}_h, 2\boldsymbol{y}_h) + c|\!|\!|\boldsymbol{w}_h|\!|\!|_{v,h}^2. \tag{4.58}$$

Adding $|\!|\!|\boldsymbol{w}_h|\!|\!|_{v,h}^2$ to both sides of eq. (4.58), the first bound in eq. (4.40) yields the result. □

We end this section by proving eq. (4.2b).

Proof of eq. (4.2b). By eq. (3.12c) and using that $\tau_\varepsilon \le c\varepsilon$, because on $\mathcal{T}_h^{dx}$, $\tau_\varepsilon \le \Delta t \le c\delta t \le c\varepsilon$ and on $\mathcal{T}_h^c$, $\tau_\varepsilon = \Delta t_{\mathcal{K}} \varepsilon \le \varepsilon (\le c\varepsilon)$, we find

$$\begin{aligned}
\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t \left(\Pi_h\left(\varphi w_h\right)\right)\right\|_{\mathcal{K}}^2 \le& c\left(\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t\left(\varphi w_h\right)\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\overline{\nabla}\left(\varphi w_h\right)\right\|_{\mathcal{K}}^2\right) \\
\le& c\left(c_T^2\left(\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t w_h\right\|_{\mathcal{K}}^2 + c\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} w_h\right\|_{\mathcal{K}}^2\right) + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|w_h\right\|_{\mathcal{K}}^2\right) \\
\le& c_T^2 |\!|\!|\boldsymbol{w}_h|\!|\!|_{s,h}^2.
\end{aligned}$$

Therefore, using lemma 4.3, we conclude that

$$|\!|\!|\boldsymbol{\Pi}_h\left(\varphi \boldsymbol{w}_h\right)|\!|\!|_{s,h} \le c_T |\!|\!|\boldsymbol{w}_h|\!|\!|_{s,h}. \tag{4.59}$$

Equation (4.2b) can now be shown to hold after combining eq. (4.59) with lemmas 4.5 and 4.6. □

4.1.3 The inf-sup condition with respect to $|\!|\!|\cdot|\!|\!|_{ss,h}$

Proof of theorem 4.1. We construct the test function $\boldsymbol{\varkappa}_h := (\varkappa_h, \varsigma_h)$ such that for $\mathcal{K} \in \mathcal{T}_h$, $\varkappa_h|_{\mathcal{K}} := \frac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \Pi_h\left(\beta \cdot \nabla w_h\right)$ while ς_h vanishes on all faces of $\mathcal{F}_h$. We first show that there exists a positive constant c_1, independent of h_K, $\delta t_{\mathcal{K}}$, ε, and T such that the following holds:

$$|||\boldsymbol{\kappa}_h|||_{s,h} \leq c_1 \|w_h\|_{sd,h} \,. \tag{4.60}$$

We bound each term of $|||\cdot|||_{s,h}$, starting with the volume terms. Noting that $\frac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \leq 1$ and using the definition of $\|\cdot\|_{sd,h}$ in eq. (2.7c), we have:

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \|\kappa_h\|_{\mathcal{K}}^2 \leq \|w_h\|_{sd,h}^2 \,. \tag{4.61}$$

The diffusive volume term is bounded using $\varepsilon \leq 1$, $\frac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} h_K^{-2} \leq 1$ and eq. (3.7b):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \, \|\overline{\nabla}\kappa_h\|_{\mathcal{K}}^2 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K}\right)^2 h_K^{-2} \left\|\Pi_h \left(\beta \cdot \nabla w_h\right)\right\|_{\mathcal{K}}^2 \leq c\|w_h\|_{sd,h}^2 \,. \tag{4.62}$$

For the time derivative volume term, we need eq. (4.44) and eq. (3.7a):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \|\partial_t \kappa_h\|_{\mathcal{K}}^2 \leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \delta t_{\mathcal{K}} \left(\delta t_{\mathcal{K}}^{-1} + h_K^{-1}\right)^2 \left(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K}\right)^2 \left\|\Pi_h \left(\beta \cdot \nabla w_h\right)\right\|_{\mathcal{K}}^2 \leq c\|w_h\|_{sd,h}^2 \,. \tag{4.63}$$

Next we turn to the facet terms. To bound the diffusive facet term, we apply eq. (3.7c):

$$\begin{aligned} \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \, \|[\![\boldsymbol{\kappa}_h]\!]\|_{\mathcal{Q}_{\mathcal{K}}}^2 &= \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K}\right)^2 \|\Pi_h \left(\beta \cdot \nabla w_h\right)\|_{\mathcal{Q}_{\mathcal{K}}}^2 \\ &\leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \|\Pi_h \left(\beta \cdot \nabla w_h\right)\|_{\mathcal{K}}^2 \leq c \, \|w_h\|_{sd,h}^2 \,. \end{aligned} \tag{4.64}$$

We use eq. (3.7d) and that $\left(\delta t_{\mathcal{K}}^{-1/2} + h_K^{-1/2}\right)^2 \frac{\delta t_{\mathcal{K}} h_K}{\delta t_{\mathcal{K}} + h_K} \leq 2$ to bound the advective facet term:

$$\begin{aligned} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \, [\boldsymbol{\kappa}_h] \right\|_{\partial\mathcal{K}}^2 &\leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K}\right)^2 \|\Pi_h \left(\beta \cdot \nabla w_h\right)\|_{\partial\mathcal{K}}^2 \\ &\leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \|\Pi_h \left(\beta \cdot \nabla w_h\right)\|_{\mathcal{K}}^2 \leq c\|w_h\|_{sd,h}^2 \,. \end{aligned} \tag{4.65}$$

The Neumann boundary term vanishes since $\varsigma_h \equiv 0$. We can therefore conclude eq. (4.60) from eqs. (4.61) to (4.65).

We next show that there exists a positive constant c_2, independent of h_K, $\delta t_{\mathcal{K}}$, ε, and T such that

$$\|w_h\|_{sd,h}^2 - c_2 |||\boldsymbol{w}_h|||_{s,h} \|w_h\|_{sd,h} \leq a_h(\boldsymbol{w}_h, \boldsymbol{\kappa}_h). \tag{4.66}$$

We first write the advective part of the bilinear form as:

$$a_{h,c}(\boldsymbol{w}_h, \boldsymbol{\kappa}_h) = \big(\nabla \cdot (\beta w_h), \kappa_h\big)_{\mathcal{T}_h} + \langle (\beta_s - \beta \cdot n) [\boldsymbol{w}_h], \kappa_h \rangle_{\partial \mathcal{T}_h} =: T_1 + T_2. \tag{4.67}$$

We bound T_1 using the definition of the projection operator Π_h:

$$\begin{aligned} T_1 &= \Big((I - \Pi_h)(\beta \cdot \nabla w_h), \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \Pi_h (\beta \cdot \nabla w_h) \Big)_{\mathcal{T}_h} + \Big(\Pi_h (\beta \cdot \nabla w_h), \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \Pi_h (\beta \cdot \nabla w_h) \Big)_{\mathcal{T}_h} \\ &= \| w_h \|_{sd,h}^2 . \end{aligned} \tag{4.68}$$

For T_2, we note that $|\beta_s - \beta \cdot n| \leq 2 |\beta_s - \frac{1}{2} \beta \cdot n|$ for any $F \in \mathcal{T}_h$. Then, also using eq. (3.7d) and Hölder's inequality for sums,

$$T_2 \leq c \sum_{\mathcal{K} \in \mathcal{T}_h} \Big\| \big|\beta_s - \tfrac{1}{2} \beta \cdot n\big|^{1/2} [\boldsymbol{w}_h] \Big\|_{\partial \mathcal{K}} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \big\| \Pi_h (\beta \cdot \nabla w_h) \big\|_{\partial \mathcal{K}} \leq c |\!|\!| \boldsymbol{w}_h |\!|\!|_{s,h} \| w_h \|_{sd,h} . \tag{4.69}$$

For the diffusive part of the bilinear form, we write:

$$\begin{aligned} a_{h,d}(\boldsymbol{w}_h, \boldsymbol{\kappa}_h) &= (\varepsilon \overline{\nabla} w_h, \overline{\nabla} \kappa_h)_{\mathcal{T}_h} - \langle \varepsilon [\boldsymbol{w}_h], \overline{\nabla}_{\bar{n}} \kappa_h \rangle_{\mathcal{Q}_h} - \langle \varepsilon \kappa_h, \overline{\nabla}_{\bar{n}} w_h \rangle_{\mathcal{Q}_h} + \langle \alpha \varepsilon h_K^{-1} [\boldsymbol{w}_h], \kappa_h \rangle_{\mathcal{Q}_h} \\ &=: I_1 + I_2 + I_3 + I_4. \end{aligned} \tag{4.70}$$

For I_1, we apply the Cauchy–Schwarz inequality, eq. (3.7b), and Hölder's inequality for sums:

$$\begin{aligned} I_1 &\leq \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \| \overline{\nabla} w_h \|_{\mathcal{K}} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \| \overline{\nabla} (\Pi_h (\beta \cdot \nabla w_h)) \|_{\mathcal{K}} \\ &\leq c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon^{1/2} \| \overline{\nabla} w_h \|_{\mathcal{K}} \Big(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \Big)^{1/2} \| \Pi_h (\beta \cdot \nabla w_h) \|_{\mathcal{K}} \leq c |\!|\!| \boldsymbol{w}_h |\!|\!|_{s,h} \| w_h \|_{sd,h} . \end{aligned} \tag{4.71}$$

For I_2, we apply the Cauchy–Schwarz inequality, eqs. (3.7b) and (3.7c), and Hölder's inequality for sums:

$$\begin{aligned} I_2 &\leq \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \| [\boldsymbol{w}_h] \|_{\mathcal{Q}_{\mathcal{K}}} \tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \| \overline{\nabla} (\Pi_h (\beta \cdot \nabla w_h)) \|_{\mathcal{Q}_{\mathcal{K}}} \\ &\leq c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2} \| [\boldsymbol{w}_h] \|_{\mathcal{Q}_{\mathcal{K}}} \Big(\tfrac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \Big)^{1/2} \| \Pi_h (\beta \cdot \nabla w_h) \|_{\mathcal{K}} \leq c |\!|\!| \boldsymbol{w}_h |\!|\!|_{s,h} \| w_h \|_{sd,h} . \end{aligned} \tag{4.72}$$

Similarly for I_3 and I_4, we apply the Cauchy–Schwarz inequality, eq. (3.7c), and Hölder's inequality for sums:

$$I_3 + I_4 \leq c|||\boldsymbol{w}_h|||_{s,h} \|w_h\|_{sd,h} \,. \tag{4.73}$$

Combining eqs. (4.67) to (4.73), we conclude eq. (4.66).

Combining eq. (4.66) and eq. (4.60) then yields:

$$c_1^{-1}\left(\|w_h\|_{sd,h} - c_2|||\boldsymbol{w}_h|||_{s,h}\right) \leq \frac{a_h(\boldsymbol{w}_h, \boldsymbol{\kappa}_h)}{c_1\|w_h\|_{sd,h}} \leq \frac{a_h(\boldsymbol{w}_h, \boldsymbol{\kappa}_h)}{|||\boldsymbol{\kappa}_h|||_{s,h}} \leq \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{a_h(\boldsymbol{w}_h, \boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{s,h}}.$$

By combining the above with eq. (4.2b),

$$\left(1 + (c_1^{-1}c_2 + 1)c_T\right) \sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{a_h(\boldsymbol{w}_h, \boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{s,h}} \geq c_1^{-1}\|w_h\|_{sd,h} + |||\boldsymbol{w}_h|||_{s,h} \geq c|||\boldsymbol{w}_h|||_{ss,h},$$

proving eq. (4.1). □

4.2 Error analysis

The following projection estimates for Π_h and $\Pi_h^{\mathcal{F}}$ were shown to hold for any $u|_{\mathcal{K}} \in H^{(p_t+1,p_s+1)}(\mathcal{K})$, $\mathcal{K} \in \mathcal{T}_h$, see [76, Lemma 5.2], [98, Lemma 6.1 and Remark 6.2], and [53, Lemmas 3.13 and 3.17]:

$$\|u - \Pi_h u\|_{\mathcal{K}}^2 \leq c\left(h_K^{2p_s+2} + \delta t_{\mathcal{K}}{}^{2p_t+2}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,, \tag{4.74a}$$

$$\|\overline{\nabla}\left(u - \Pi_h u\right)\|_{\mathcal{K}}^2 \leq c\left(h_K^{2p_s} + \delta t_{\mathcal{K}}{}^{2p_t+2}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,, \tag{4.74b}$$

$$\|\partial_t\left(u - \Pi_h u\right)\|_{\mathcal{K}}^2 \leq c\left(h_K^{2p_s} + \delta t_{\mathcal{K}}{}^{2p_t}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,, \tag{4.74c}$$

$$\|\overline{\nabla}_{\bar{n}}\left(u - \Pi_h u\right)\|_{\mathcal{Q}_{\mathcal{K}}}^2 \leq c\left(h_K^{2p_s-1} + h_K^{-1}\delta t_{\mathcal{K}}{}^{2p_t+2}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,, \tag{4.74d}$$

$$\|u - \Pi_h u\|_{\partial\mathcal{K}}^2 \leq c\left(h_K^{2p_s+1} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,, \tag{4.74e}$$

$$\|u - \Pi_h^{\mathcal{F}} u\|_{\partial\mathcal{K}}^2 \leq c\left(h_K^{2p_s+1} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right)\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 \,. \tag{4.74f}$$

Let us define $h := \max_{\mathcal{K}\in\mathcal{T}_h} h_K$ and $\delta t := \max_{\mathcal{K}\in\mathcal{T}_h} \delta t_{\mathcal{K}}$. An immediate consequence of eq. (4.74) is the following estimate.

Lemma 4.7. *Let u, with $u|_{\mathcal{K}} \in H^{(p_t+1,p_s+1)}(\mathcal{K})$ for all $\mathcal{K} \in \mathcal{T}_h$, and define $\boldsymbol{u} := (u, u|_\Gamma)$. Let $\boldsymbol{\Pi}_h\boldsymbol{u} = (\Pi_h u, \Pi_h^{\mathcal{F}} u)$. Then,*

$$|||\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}|||_{ss,h}^2 \leq c\left[h^{2p_s}(h + \varepsilon + \tilde{\varepsilon}\delta t) + \delta t^{2p_t}(\delta t + \varepsilon h^{-1}\delta t)\right],$$

where the constant c depends on $\sum_{\mathcal{K}\in\mathcal{T}_h} \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}$.

Proof. By eq. (4.74a),

$$\|u - \Pi_h u\|_{\mathcal{K}}^2 \leq c \left(h_K^{2p_s+2} + \delta t_{\mathcal{K}}{}^{2p_t+2}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 . \tag{4.75}$$

Next, eq. (4.74b) gives us:

$$\varepsilon \left\|\overline{\nabla}\left(u - \Pi_h u\right)\right\|_{\mathcal{K}}^2 \leq c\varepsilon \left(h_K^{2p_s} + \delta t_{\mathcal{K}}{}^{2p_t+2}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 . \tag{4.76}$$

For the advective facet terms, we use eqs. (4.74e) and (4.74f) and the triangle inequality:

$$\begin{aligned}\left\|\left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2}\left((u - \Pi_h u) - (\gamma(u) - \Pi_h^{\mathcal{F}}\gamma(u))\right)\right\|_{\partial\mathcal{K}}^2 + \left\|\left|\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left(\gamma(u) - \Pi_h^{\mathcal{F}}\gamma(u)\right)\right\|_{\partial\mathcal{K}\cap\partial\mathcal{E}_N}^2 \\ \leq c\left(h_K^{2p_s+1} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 .\end{aligned} \tag{4.77}$$

Similarly, for the diffusive facet term, we again apply the triangle inequality and eqs. (4.74e) and (4.74f):

$$\varepsilon h_K^{-1} \left\|(u - \Pi_h u) - (\gamma(u) - \Pi_h^{\mathcal{F}}\gamma(u))\right\|_{\mathcal{Q}_{\mathcal{K}}}^2 \leq c\varepsilon\left(h_K^{2p_s} + h_K^{-1}\delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 . \tag{4.78}$$

For the streamline derivative term, we use eqs. (4.74b) and (4.74c) and that $\frac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K} \leq \delta t_{\mathcal{K}} h_K$:

$$\begin{aligned}\frac{\delta t_{\mathcal{K}} h_K^2}{\delta t_{\mathcal{K}} + h_K}\left\|\Pi_h\left(\beta\cdot\nabla\left(u - \Pi_h u\right)\right)\right\|_{\mathcal{K}}^2 &\leq c\delta t_{\mathcal{K}} h_K \left(\left\|\overline{\nabla}\left(u - \Pi_h u\right)\right\|_{\mathcal{K}}^2 + \left\|\partial_t\left(u - \Pi_h u\right)\right\|_{\mathcal{K}}^2\right) \\ &\leq c\delta t_{\mathcal{K}} h_K \left(h_K^{2p_s} + \delta t_{\mathcal{K}}{}^{2p_t}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 .\end{aligned} \tag{4.79}$$

Finally, for the time-derivative term, using eq. (4.74c),

$$\begin{aligned}\mathcal{T}_\varepsilon \left\|\partial_t\left(u - \Pi_h u\right)\right\|_{\mathcal{K}}^2 &\leq \begin{cases} c\left(h_K^{2p_s}\delta t_{\mathcal{K}} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 & \text{if } \mathcal{K} \in \mathcal{T}_h^d, \\ c\varepsilon^{1/2}\left(h_K^{2p_s}\delta t_{\mathcal{K}} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 & \text{if } \mathcal{K} \in \mathcal{T}_h^x, \\ c\varepsilon\left(h_K^{2p_s}\delta t_{\mathcal{K}} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 & \text{if } \mathcal{K} \in \mathcal{T}_h^c. \end{cases} \\ &\leq c\tilde{\varepsilon}\left(h_K^{2p_s}\delta t_{\mathcal{K}} + \delta t_{\mathcal{K}}{}^{2p_t+1}\right) \|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}^2 .\end{aligned} \tag{4.80}$$

The result follows after combining eqs. (4.75) to (4.80) and summing over all $\mathcal{K} \in \mathcal{T}_h$. □

The following lemma will be used to prove the global error estimate of theorem 4.2.

Lemma 4.8. *Let u, with $u|_{\mathcal{K}} \in H^{(p_t+1,p_s+1)}(\mathcal{K})$ for all $\mathcal{K} \in \mathcal{T}_h$, solve eq. (2.1) and define $\boldsymbol{u} := (u, \lambda)$ with $\lambda = u|_{\Gamma}$. Let $\boldsymbol{\Pi}_h\boldsymbol{u} = (\Pi_h u, \Pi_h^{\mathcal{F}} u)$ and let $\boldsymbol{u}_h = (u_h, \lambda_h) \in \boldsymbol{V}_h$ be the solution to eq. (2.9). The following holds:*

$$
\begin{aligned}
&|a_h(\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}, \boldsymbol{v}_h)| \\
&\leq \Big[c|\!|\!|\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}|\!|\!|_{ss,h} + c\, \big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} (u - \Pi_h u) \big\|_{\partial\mathcal{T}_h} + \big(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K \left\|\overline{\nabla}_{\bar{n}} \left(u - \Pi_h u\right)\right\|_{\mathcal{Q}_{\mathcal{K}}}^2\big)^{1/2}\Big] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdot |\!|\!|\boldsymbol{v}_h|\!|\!|_{s,h}.
\end{aligned}
$$

Proof. We start with the advective part of $a_h(\cdot,\cdot)$. Writing $\zeta^+\beta \cdot n = (\beta \cdot n + |\beta \cdot n|)/2$ and using the triangle inequality,

$$
\begin{aligned}
|a_{h,c}(\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}, \boldsymbol{v}_h)| \leq{}& |(\beta\left(u - \Pi_h u\right), \nabla v_h)_{\mathcal{T}_h}| + |\langle \tfrac{1}{2}\left(\beta \cdot n + |\beta \cdot n|\right)\left(\lambda - \Pi_h^{\mathcal{F}}\lambda\right), \mu_h\rangle_{\partial\mathcal{E}_N}| \\
&+ |\langle (\beta \cdot n)\left(\lambda - \Pi_h^{\mathcal{F}}\lambda\right) + \beta_s \left[\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}\right], \left[\boldsymbol{v}_h\right]\rangle_{\partial\mathcal{T}_h}| =: I_1 + I_2 + I_3.
\end{aligned}
$$

To bound I_1, we follow the proof of [12, Theorem 5.1] by noting that if $\beta_0 = (1, \bar{\beta}_0)$ then $(\beta_0\left(u - \Pi_h u\right), \nabla v_h)_{\mathcal{T}_h} = 0$ and $((\beta - \beta_0)\left(u - \Pi_h u\right), \nabla v_h)_{\mathcal{T}_h} = ((\bar{\beta} - \bar{\beta}_0)\left(u - \Pi_h u\right), \overline{\nabla} v_h)_{\mathcal{T}_h}$. Then, using the Cauchy–Schwarz inequality, eq. (4.4), eq. (3.7b), and Hölder's inequality for sums, we obtain

$$
I_1 \leq \sum_{\mathcal{K}\in\mathcal{T}_h} c \left\|u - \Pi_h u\right\|_{\mathcal{K}} \left\|v_h\right\|_{\mathcal{K}}.
$$

Using the Cauchy–Schwarz inequality, we bound I_2 as:

$$
I_2 \leq c \,\big\| |\tfrac{1}{2}\beta \cdot n|^{1/2} (\lambda - \Pi_h^{\mathcal{F}}\lambda) \big\|_{\partial\mathcal{E}_N} \big\| |\tfrac{1}{2}\beta \cdot n|^{1/2} \mu_h \big\|_{\partial\mathcal{E}_N}.
$$

With $\beta \cdot n \leq \sup|\beta \cdot n| \leq 2\left(\sup|\beta \cdot n| - \frac{1}{2}\beta \cdot n\right)$, for all $F \in \partial\mathcal{T}_h$, and the Cauchy–Schwarz inequality, we bound I_3 as:

$$
\begin{aligned}
I_3 &\leq c \left|\langle (\sup|\beta \cdot n| - \tfrac{1}{2}\beta \cdot n)\left(\lambda - \Pi_h^{\mathcal{F}}\lambda + \left[\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}\right]\right), \left[\boldsymbol{v}_h\right]\rangle_{\partial\mathcal{T}_h}\right| \\
&\leq c \sum_{\mathcal{K}\in\mathcal{T}_h} \big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} (u - \Pi_h u) \big\|_{\partial\mathcal{K}} \big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} \left[\boldsymbol{v}_h\right] \big\|_{\partial\mathcal{K}}.
\end{aligned}
$$

Collecting the bounds for I_1, I_2, and I_3, and using Hölder's inequality for sums,

$$
|a_{h,c}(\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}, \boldsymbol{v}_h)| \leq \left(c|\!|\!|\boldsymbol{u} - \boldsymbol{\Pi}_h\boldsymbol{u}|\!|\!|_{ss,h} + c\,\big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} (u - \Pi_h u) \big\|_{\partial\mathcal{T}_h}\right) |\!|\!|\boldsymbol{v}_h|\!|\!|_{s,h}. \tag{4.81}
$$

We now proceed with the diffusive part of $a_h(\cdot,\cdot)$. By the triangle inequality,

$$\begin{aligned}|a_{h,d}(\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u},\boldsymbol{v}_h)| \leq{}& |(\varepsilon\overline{\nabla}\left(u-\Pi_h u\right),\overline{\nabla}v_h)_{\mathcal{T}_h}| + |\langle\varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}\right],\left[\boldsymbol{v}_h\right]\rangle_{\mathcal{Q}_h}| \\ &+ |\langle\varepsilon\left[\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}\right],\overline{\nabla}_{\bar{n}}v_h\rangle_{\mathcal{Q}_h}| + |\langle\varepsilon\overline{\nabla}_{\bar{n}}\left(u-\Pi_h u\right),\left[\boldsymbol{v}_h\right]\rangle_{\mathcal{Q}_h}|\,.\end{aligned}$$

By applying the Cauchy–Schwarz inequality, the first two terms on the right-hand side can be bounded by $(1+\alpha)\,|||\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}|||_{s,h}|||\boldsymbol{v}_h|||_{s,h}$. For the last two terms, by the Cauchy–Schwarz inequality and eq. (3.7c),

$$\begin{aligned}|\langle\varepsilon\left[\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}\right],\overline{\nabla}_{\bar{n}}v_h\rangle_{\mathcal{Q}_h}|&+|\langle\varepsilon\overline{\nabla}_{\bar{n}}\left(u-\Pi_h u\right),\left[\boldsymbol{v}_h\right]\rangle_{\mathcal{Q}_h}| \\ &\leq c\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon^{1/2}h_K^{-1/2}\left\|\left[\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}\right]\right\|_{\mathcal{Q}_\mathcal{K}}\varepsilon^{1/2}\left\|\overline{\nabla}v_h\right\|_{\mathcal{K}} \\ &\quad+\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon^{1/2}h_K^{1/2}\left\|\overline{\nabla}_{\bar{n}}\left(u-\Pi_h u\right)\right\|_{\mathcal{Q}_\mathcal{K}}\varepsilon^{1/2}h_K^{-1/2}\left\|\left[\boldsymbol{v}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}}\,.\end{aligned}$$

Therefore, using Hölder's inequality for sums,

$$|a_{h,d}(\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u},\boldsymbol{v}_h)| \leq \big(c|||\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}|||_{ss,h} + \big(\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K\left\|\overline{\nabla}_{\bar{n}}\left(u-\Pi_h u\right)\right\|_{\mathcal{Q}_\mathcal{K}}^2\big)^{1/2}\big)\,|||\boldsymbol{v}_h|||_{s,h}. \tag{4.82}$$

The result follows by combining eq. (4.81) and eq. (4.82). □

Theorem 4.2 (Global error estimate)**.** *Let* $\boldsymbol{u}$ *and* $\boldsymbol{u}_h$ *be as in lemma 4.8. Then*

$$|||\boldsymbol{u}-\boldsymbol{u}_h|||_{ss,h}^2 \leq c_T\left[h^{2p_s}(h+\varepsilon+\tilde{\varepsilon}\delta t)+\delta t^{2p_t+1}(1+\varepsilon h^{-1})\right].$$

where c_T *depends on* $\sum_{\mathcal{K}\in\mathcal{T}_h}\|u\|_{H^{(p_t+1,p_s+1)}(\mathcal{K})}$.

Proof. We start by noting that Galerkin orthogonality was shown in [76]:

$$a_h(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)=0\quad\forall\boldsymbol{v}_h:=(v_h,\mu_h)\in\boldsymbol{V}_h. \tag{4.83}$$

By a triangle inequality, theorem 4.1, and eq. (4.83) we find:

$$|||\boldsymbol{u}_h-\boldsymbol{u}|||_{ss,h}\leq|||\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}|||_{ss,h}+c_T\sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h}\frac{a_h(\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u},\boldsymbol{v}_h)}{|||\boldsymbol{v}_h|||_{s,h}}.$$

Using lemma 4.8,

$$\begin{aligned}|||\boldsymbol{u}_h-\boldsymbol{u}|||_{ss,h}\leq c_T|||\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}|||_{ss,h}&+c_T\left\||\beta_s-\tfrac{1}{2}\beta\cdot n|^{1/2}\left(u-\Pi_h u\right)\right\|_{\partial\mathcal{T}_h} \\ &+c_T\big(\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon h_K\left\|\overline{\nabla}_{\bar{n}}\left(u-\Pi_h u\right)\right\|_{\mathcal{Q}_\mathcal{K}}^2\big)^{1/2}.\end{aligned} \tag{4.84}$$

The second term on the right-hand side of eq. (4.84) is bounded using eq. (4.74e):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}(u - \Pi_h u)\right\|^2_{\partial\mathcal{K}_h} \le c\,(h^{2p_s+1} + \delta t^{2p_t+1}) \sum_{\mathcal{K}\in\mathcal{T}_h} \|u\|^2_{H^{(p_t+1,p_s+1)}(\mathcal{K})}\,. \tag{4.85}$$

The last term on the right-hand side of eq. (4.84) is bounded using eq. (4.74d):

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K \left\|\overline{\nabla}_{\bar{n}}(u - \Pi_h u)\right\|^2_{\mathcal{Q}_\mathcal{K}} \le c\varepsilon\,(h^{2p_s} + \delta t^{2p_t+2}) \sum_{\mathcal{K}\in\mathcal{T}_h} \|u\|^2_{H^{(p_t+1,p_s+1)}(\mathcal{K})}\,. \tag{4.86}$$

The result follows after combining eqs. (4.84) to (4.86) and lemma 4.7. □

Remark 1. *The error estimate of theorem 4.2 shows that if* $\varepsilon < \delta t = h$ *then* $|||\boldsymbol{u}-\boldsymbol{u}_h|||_{ss,h} = \mathcal{O}(h^{p_s+1/2} + \delta t^{p_t+1/2})$, *while if* $\delta t = h < \varepsilon$ *then* $|||\boldsymbol{u}-\boldsymbol{u}_h|||_{ss,h} = \mathcal{O}(h^{p_s} + \delta t^{p_t})$.

4.3 Numerical examples

The space-time HDG method eq. (2.9) is implemented in this section using the finite element library deal.II [8] on unstructured hexahedral space-time meshes with p4est [16] to obtain distributed mesh information. We use PETSc [14, 13, 15] to solve the linear systems arising at each time-step (GMRES preconditioned by classical algebraic multigrid from BoomerAMG [59] with an absolute solver tolerance of 10^{-12}).

In our implementation we furthermore choose the penalty parameter $\alpha = 8p_s^2$ (see, for example, [90]). For both numerical examples, we show the rates of convergence for different polynomial degrees when the error is measured in $|||\cdot|||_{ss,h}$ for $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-8}$, respectively.

4.3.1 A rotating Gaussian pulse test on moving domain

We consider the solution of a two-dimensional rotating Gaussian pulse on a deforming domain [89] to demonstrate the convergence properties of the space-time HDG method predicted by theorem 4.2. In eq. (2.1) we set $\beta = (1, -4x_2, 4x_1)^T$ and $f = 0$. Defining $\widetilde{x}_1 := x_1\cos(4t) + x_2\sin(4t)$ and $\widetilde{x}_2 := -x_1\sin(4t) + x_2\cos(4t)$, the exact solution to this problem is given by

$$u(t, x_1, x_2) = \tfrac{\sigma^2}{\sigma^2+2\varepsilon t}\exp\left(-\tfrac{(\widetilde{x}_1-x_{1c})^2+(\widetilde{x}_2-x_{2c})^2}{2\sigma^2+4\varepsilon t}\right),$$

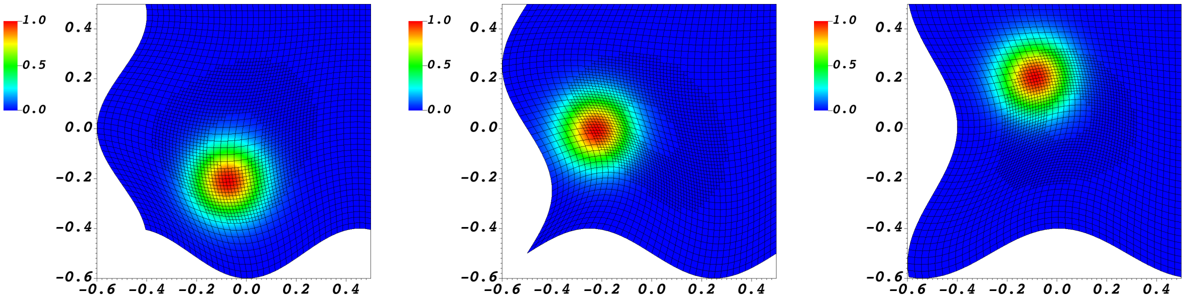


Figure 4.1: The spatial mesh and the ring of elements with an extra level of refinement deform over time. The solution shown is for $\varepsilon = 10^{-8}$. Plots correspond to time levels $t = 0.2, 0.5, 0.8$ from left to right.

with initial and boundary conditions set appropriately. We choose $\sigma = 0.1$ and $(x_{1c}, x_{2c}) = (-0.2, 0.1)$. The deforming domain $\Omega(t)$ is obtained by transforming a uniform mesh, with coordinates $(x_1^u, x_2^u) \in (-0.5, 0.5)^2$, to

$$x_i = x_i^u + A(\tfrac{1}{2} - x_i^u)\sin(2\pi(\tfrac{1}{2} - x_i^* + t)), \quad i = 1, 2,$$

where $(x_1^*, x_2^*) = (x_2, x_1)$ and $A = 0.1$. We consider this problem for $t \in [0, 1]$.

To create our coarsest mesh, we start with an initial mesh with elements of size $h \approx \delta t = 10^{-1}$. Space-time elements in a ring prescribed by $|((x_1^c)^2 + (x_2^c)^2)^{1/2} - 0.2| < 0.1$, where (x_1^c, x_2^c) is the spatial coordinate of the centre of a space-time element, are then uniformly refined once and are of size $h \approx \delta t = 0.05$. See fig. 4.1 for plots of the solution and spatial mesh at different time levels. The reason to consider two sets of elements is to verify that the analysis of previous sections hold on 1-irregular space-time meshes. Finer meshes are obtained by uniformly refining our coarsest mesh.

In the third row of table 4.1 we have that $h \approx \delta t = 1.25 \times 10^{-2}$ inside the refined ring while elsewhere $h \approx \delta t = 2.5 \times 10^{-2}$. Therefore, for the first three rows in table 4.1, $h \approx \delta t \geq \varepsilon = 10^{-2}$ and we observe a rate of convergence of approximately $p + 1/2$. In the following three rows we observe a drop in the rate of convergence to approximately p. This happens in two stages since there are two sets of elements in our mesh, see fig. 4.1. In the first stage (the fourth row of table 4.1), elements in the refined ring are such that $h \approx \delta t = 6.25 \times 10^{-3} < \varepsilon$, but elsewhere $h \approx \delta t = 1.25 \times 10^{-2} > \varepsilon$. In the next stage (fifth row of table 4.1), all elements satisfy $h \approx \delta t < \varepsilon$, resulting in a rate of convergence of p after the fifth row. In table 4.2 we observe that $h \approx \delta t > \varepsilon = 10^{-8}$ for all cycles and the error converges at a rate of approximately $p + 1/2$. These observations from tables 4.1 and 4.2 are in agreement with remark 1.

Table 4.1: The solution errors measured in $|||\cdot|||_{ss,h}$ and corresponding rates of convergence when using polynomial approximation $p = 1, 2, 3$ for the case $\varepsilon = 10^{-2}$.

Cells per slab	Number of slabs	$p = 1$	Rates	$p = 2$	Rates	$p = 3$	Rates
296	10	4.7e-2	-	7.8e-3	-	1.3e-3	-
1100	20	1.8e-2	1.4	1.6e-3	2.4	1.2e-4	3.6
4372	40	7.7e-3	1.3	3.2e-4	2.3	1.7e-5	3.4
17572	80	3.7e-3	1.1	7.3e-5	2.1	1.4e-6	3.2
70540	160	2.0e-3	0.9	2.3e-5	1.7	2.4e-7	2.4
282580	320	9.0e-4	1.1	4.9e-6	2.2	2.5e-8	3.3

Table 4.2: The solution errors measured in $|||\cdot|||_{ss,h}$ and corresponding rates of convergence when using polynomial approximation $p = 1, 2, 3$ for the case $\varepsilon = 10^{-8}$.

Cells per slab	Number of slabs	$p = 1$	Rates	$p = 2$	Rates	$p = 3$	Rates
289	10	1.1e-1	-	1.6e-2	-	2.8e-3	-
1086	20	3.9e-2	1.5	2.8e-3	2.7	2.3e-4	3.8
4372	40	1.1e-2	1.8	4.4e-4	2.7	1.8e-5	3.7
17572	80	3.4e-3	1.7	7.1e-5	2.6	1.4e-6	3.7
70540	160	1.1e-3	1.6	1.2e-5	2.6	1.1e-7	3.6
282580	320	4.0e-4	1.5	2.1e-6	2.5	9.7e-9	3.6

4.3.2 A boundary layer test case on a fixed domain

In this example, we consider the solution of a two-dimensional boundary layer test case on a fixed domain $(x_1, x_2) \in (0,1)^2$ and $t \in [0,1]$ [52]. In eq. (2.1) we set $\beta = (1,1,1)^T$ and with suitably chosen source term f, the exact solution is given by

$$u(t, x_1, x_2) = \big(1 - \exp(-t)\big)$$
$$\cdot\Big(\sin(\frac{\pi x_1}{2}) + \sin(\frac{\pi x_2}{2}) - \sin(\frac{\pi x_1}{2})\sin(\frac{\pi x_2}{2}) + \frac{\exp(-1/\varepsilon) - \exp(-(1-x_1)(1-x_2)/\varepsilon)}{1 - \exp(-1/\varepsilon)}\Big).$$

The solution develops boundary layers of width $\mathcal{O}(\varepsilon)$ near the domain boundary where $x = 1$ and $y = 1$. The trigonometry terms in the solution are added such that the exact solution does not behave like quadratic polynomials away from the boundary layer (see [12, Section 6, example 4]). This helps to verify the rates of convergence when using $p > 1$.

In table 4.3, the convergence history is presented for $\varepsilon = 10^{-2}$. In the fourth row of table 4.1, all elements satisfy $h = \delta t = 0.0125 > \varepsilon$, while in the fifth row of table 4.1, all elements satisfy $h = \delta t = 6.25 \times 10^{-3} < \varepsilon$. Hence, the boundary layer has been resolved in the last two rows, which show a rate of convergence p. The rate of convergence $p+1/2$ in the pre-asymptotic regime, i.e., prior to the layer being resolved, is not observed in this case. We remark that this may be due to that, in the pre-asymptotic regime, $\|u\|_{H^{(p_t+1, p_s+1)}(\mathcal{K})}$ for a space-time element $\mathcal{K}$ within the layer region may be as large as $\mathcal{O}(\varepsilon^{-1})$. The errors in these elements dominate $|||\boldsymbol{u} - \boldsymbol{u}_h|||_{ss,h}$, but they are not accounted for in the error estimate in lemma 4.7. See the left panel of fig. 4.2 for the solution solved with 320 space-time slabs and plotted at the final time $t = 1.0$.

In table 4.4, the convergence history is presented for $\varepsilon = 10^{-8}$. For this case, the bounday layer is unresolved throughout all refinement levels due to the most refined level corresponding to $h = \delta t = 3.125 \times 10^{-3} \gg 10^{-8}$. As a result, the sharp gradient in the boundary layer region is not resolved. This becomes the dominating source of the global error $|||\boldsymbol{u} - \boldsymbol{u}_h|||_{ss,h}$. Therefore, we choose to compute the solution error measured in $|||\cdot|||_{ss,h}$ only in that part of the domain that excludes the boundary layer, i.e., in $[0,1] \times [0,0.9] \times [0,0.9] \subsetneq \mathcal{E}$. Outside of the boundary layer, the solution is "smooth" and hence $\|u\|_{H^{(p_t+1, p_s+1)}(\mathcal{K})}$ can be bounded by $\mathcal{O}(1)$. We observe that the error converges at a rate of approximately $p+1/2$ in agreement with remark 1 in regard to the pre-asymptotic regime. See the right panel of fig. 4.2 for the solution solved with 320 space-time slabs and plotted at the final time $t = 1.0$.

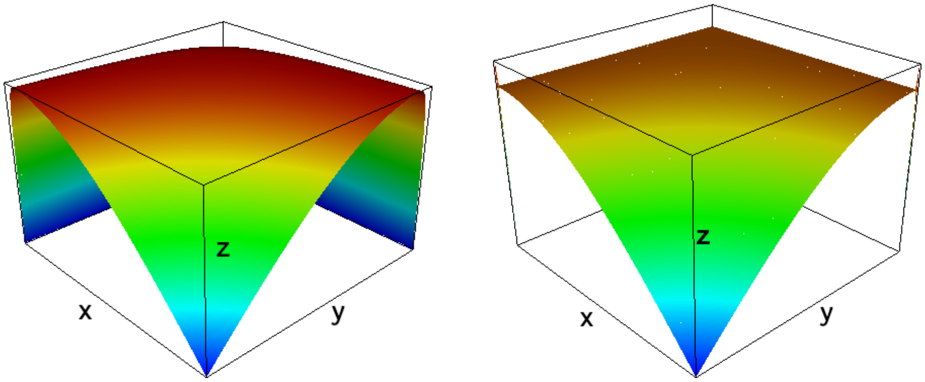


Figure 4.2: The solutions shown are for the boundary layer test case with $\varepsilon = 10^{-2}$ (left panel) and $\varepsilon = 10^{-8}$ (right panel) respectively. Both are solved with 320 space-time slabs (corresponding to the last rows of table 4.3 and table 4.4) and plotted at the final time $t = 1.0$. We observe that the boundary layer has been resolved for $\varepsilon = 10^{-2}$ whereas for $\varepsilon = 10^{-8}$, the layer has not yet been resolved.

Table 4.3: The solution errors measured in $|||\cdot|||_{ss,h}$ and corresponding rates of convergence when using polynomial approximation $p = 1, 2, 3$ for the case $\varepsilon = 10^{-2}$.

Cells per slab	Number of slabs	$p = 1$	Rates	$p = 2$	Rates	$p = 3$	Rates
100	10	4.0e-1	-	2.5e-1	-	1.3e-1	-
400	20	2.6e-1	0.6	1.1e-1	1.2	3.3e-2	2.0
1600	40	1.5e-1	0.8	3.5e-2	1.6	6.0e-3	2.5
6400	80	8.0e-2	0.9	9.7e-3	1.9	8.7e-4	2.8
25600	160	4.0e-2	1.0	2.5e-3	2.0	1.1e-4	3.0
102400	320	2.0e-2	1.0	6.2e-4	2.0	1.4e-5	3.0

Table 4.4: The solution errors measured in $|||\cdot|||_{ss,h}$ for space-time elements that lie in $[0, 1] \times [0, 0.9] \times [0, 0.9]$ and corresponding rates of convergence when using polynomial approximation $p = 1, 2, 3$ for the case $\varepsilon = 10^{-8}$.

Cells per slab	Number of slabs	$p = 1$	Rates	$p = 2$	Rates	$p = 3$	Rates
100	10	1.4e-2	-	7.0e-3	-	4.8e-3	-
400	20	1.3e-3	3.3	1.0e-5	9.4	5.3e-8	16.5
1600	40	4.4e-4	1.6	1.6e-6	2.7	3.6e-9	3.9
6400	80	1.5e-4	1.6	2.7e-7	2.6	2.8e-10	3.7
25600	160	5.1e-5	1.5	4.6e-8	2.6	2.3e-11	3.6
102400	320	1.8e-5	1.5	7.9e-9	2.5	2.1e-12	3.5

Chapter 5

A posteriori error analysis

This chapter is dedicated to the a posteriori error analysis of the space-time HDG method. In section 5.1, the error estimators and the main reliability and local efficiency results are presented. Section 5.2 introduces a saturation assumption and a subgrid projection. These two theoretical devices are combined to derive an a posteriori error estimation of the error of the time derivative. The reliability of the a posteriori error estimator is proven in section 5.3 and the local efficiency in section 5.4. Finally, we illustrate the theoretical results with numerical examples in section 5.5.

5.1 The error estimator and the main results

We present the residual-based a posteriori error estimator for the space-time HDG method, eq. (2.9), in this section. Firstly, we need the following element and facet residuals

$$\begin{aligned} R_h^{\mathcal{K}} &:= f + \varepsilon \overline{\nabla}^2 u_h - \nabla \cdot (\beta u_h) && \forall \mathcal{K} \in \mathcal{T}_h, \\ R_h^{N} &:= g - \varepsilon \overline{\nabla} u_h \cdot \overline{n} + \zeta^- u_h \beta \cdot n && \forall F \in \partial\mathcal{E}_N. \end{aligned}$$

Furthermore, we define

$$\begin{aligned} \eta_R^{\mathcal{K}} &:= \lambda_{\mathcal{K}} \, \| R_h^{\mathcal{K}} \|_{\mathcal{K}} \,, & \eta_{J,1}^{\mathcal{K}} &:= h_K^{1/2} \varepsilon^{1/2} \big\| [\![\overline{\nabla}_{\overline{n}} u_h]\!] \big\|_{\mathcal{Q}_{\mathcal{K}} \setminus \partial\mathcal{E}} \,, \\ \eta_{J,2}^{\mathcal{K}} &:= \Big(\big(\eta_{J,2,1}^{\mathcal{K}}\big)^2 + \big(\eta_{J,2,2}^{\mathcal{K}}\big)^2 \Big)^{1/2} , & \eta_{J,3}^{\mathcal{K}} &:= \Big(\big(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\big)^2 + \big(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\big)^2 \Big)^{1/2} , \\ \eta_{BC,1}^{\mathcal{K}} &:= h_K^{1/2} \varepsilon^{-1/2} \, \| R_h^{N} \|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} \,, & \eta_{BC,2}^{\mathcal{K}} &:= \| R_h^{N} \|_{\partial\mathcal{K} \cap \Omega_0} = \| g - u_h \|_{\partial\mathcal{K} \cap \Omega_0} \,, \end{aligned}$$

where

$$\eta^{\mathcal{K}}_{J,2,1} := h_K^{-1/2}\varepsilon^{1/2}\big\|[\boldsymbol{u}_h]\big\|_{\mathcal{Q}_\mathcal{K}}\,, \qquad \eta^{\mathcal{K}}_{J,2,2} := h_K^{1/4}\varepsilon^{-1/2}\big\|[\boldsymbol{u}_h]\big\|_{\mathcal{Q}_\mathcal{K}}\,,$$

$$\eta^{\mathcal{K}}_{J,3,\mathcal{Q}} := \big\| \big|\beta_s - \tfrac{1}{2}\beta\cdot n\big|^{1/2}[\boldsymbol{u}_h]\big\|_{\mathcal{Q}_\mathcal{K}}\,, \qquad \eta^{\mathcal{K}}_{J,3,\mathcal{R}} := \big\| \big|\beta_s - \tfrac{1}{2}\beta\cdot n\big|^{1/2}[\boldsymbol{u}_h]\big\|_{\mathcal{R}_\mathcal{K}}\,.$$

We then introduce the following a posteriori error estimator for the solution $\boldsymbol{u}_h \in \boldsymbol{V}_h$ to eq. (2.9):

$$\eta^2 := \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}\right)^2, \tag{5.1}$$

where $\left(\eta^{\mathcal{K}}\right)^2 := \left(\eta^{\mathcal{K}}_R\right)^2 + \sum_{i=1}^3 \left(\eta^{\mathcal{K}}_{J,i}\right)^2 + \sum_{j=1}^2 \left(\eta^{\mathcal{K}}_{BC,j}\right)^2$. Finally, a modified version of the norm $|||\cdot|||_{s,h}$ is needed. For $\boldsymbol{v} \in \boldsymbol{V}(h)$, we define

$$\begin{aligned}|||\boldsymbol{v}|||^2_{sT,h} := &\sum_{\mathcal{K}\in\mathcal{T}_h} \|v\|^2_{\mathcal{K}} + \sum_{\mathcal{K}\in\mathcal{T}_h} \big\| \big|\beta_s - \tfrac{1}{2}\beta\cdot n\big|^{1/2}[\boldsymbol{v}]\big\|^2_{\partial\mathcal{K}} + T\sum_{F\in\partial\mathcal{E}_N} \big\| \big|\tfrac{1}{2}\beta\cdot n\big|^{1/2}\mu\big\|^2_F \\ &+ T\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon\,\|\overline{\nabla} v\|^2_{\mathcal{K}} + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1}\,\|[\boldsymbol{v}]\|^2_{\mathcal{Q}_\mathcal{K}} + \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon\,\|\partial_t v\|^2_{\mathcal{K}}\,.\end{aligned}$$

The following theorems establish reliability and efficiency of the a posteriori error estimator eq. (5.1). Their proofs are given in sections 5.3 and 5.4, respectively.

Theorem 5.1 (Reliability)**.** *Let u solve eq. (2.1), $\boldsymbol{u} = (u, u|_\Gamma)$, and let $\boldsymbol{u}_h$ solve eq. (2.9). Assuming that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we have the following reliability estimate*

$$|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h} \le cT\varepsilon^{-1/2}\eta. \tag{5.2}$$

Theorem 5.2 (Efficiency)**.** *Let $\boldsymbol{u}$ and $\boldsymbol{u}_h$ be as in theorem 5.1 and assume that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. Furthermore, let $\operatorname{osc}_h^{\mathcal{K}} := \lambda_{\mathcal{K}}\,\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}$ and $\operatorname{osc}_h^N := h_K^{1/2}\varepsilon^{-1/2}\,\|(I-\Pi_h^{\mathcal{F}})R_h^N\|_F$, where $\Pi_h^{\mathcal{F}}$ denotes the L^2-projection onto M_h. Then, for all $\mathcal{K}\in\mathcal{T}_h$,*

$$\eta^{\mathcal{K}} \le c\sum_{\mathcal{K}\subset\omega_{\mathcal{K}}} \varepsilon^{-1/2}\tilde{\varepsilon}^{-1/2}|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + c\operatorname{osc}_h^{\mathcal{K}} + c\operatorname{osc}_h^N, \tag{5.3}$$

where

$$\begin{aligned}|||\boldsymbol{v}|||^2_{sT,h,\mathcal{K}} := &\|v\|^2_{\mathcal{K}} + \big\| \big|\beta_s - \tfrac{1}{2}\beta\cdot n\big|^{1/2}[\boldsymbol{v}]\big\|^2_{\partial\mathcal{K}} + T\sum_{F\in\partial\mathcal{E}_N\cap\partial\mathcal{K}} \big\| \big|\tfrac{1}{2}\beta\cdot n\big|^{1/2}\mu\big\|^2_F \\ &+ T\varepsilon\,\|\overline{\nabla} v\|^2_{\mathcal{K}} + \varepsilon h_K^{-1}\,\|[\boldsymbol{v}]\|^2_{\mathcal{Q}_\mathcal{K}} + \tau_\varepsilon\,\|\partial_t v\|^2_{\mathcal{K}}\,.\end{aligned}$$

Remark 2. *From theorem 5.2, and by definition of $\tilde{\varepsilon}$, we have that on sufficiently refined elements $\mathcal{K}\in\mathcal{T}_h^d$ the following estimate holds:*

$$\eta^{\mathcal{K}} \le c\sum_{\mathcal{K}\subset\omega_{\mathcal{K}}} \varepsilon^{-1/2}|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + c\operatorname{osc}_h^{\mathcal{K}} + c\operatorname{osc}_h^N\,.$$

5.2 Saturation assumption and time derivative error estimation

We pose problem eq. (2.9) on the subgrid mesh $\mathcal{T}_{\mathfrak{h}}$, i.e., find $\boldsymbol{u}_{\mathfrak{h}} \in \boldsymbol{V}_{\mathfrak{h}}$ such that

$$a_{\mathfrak{h}}\left(\boldsymbol{u}_{\mathfrak{h}}, \boldsymbol{v}_{\mathfrak{h}}\right) = \left(f, v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} + \langle g, \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} \quad \forall \boldsymbol{v}_{\mathfrak{h}} \in \boldsymbol{V}_{\mathfrak{h}}. \tag{5.4}$$

Lemma 5.1 (Galerkin orthogonality)**.** *Let* $\boldsymbol{u}_h$ *and* $\boldsymbol{u}_{\mathfrak{h}}$ *be the solutions of eq.* (2.9) *and eq.* (5.4)*, respectively. With the restriction operator defined in eq.* (3.30)*, we have the following Galerkin orthogonality result:*

$$a_{\mathfrak{h}}\left(\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right), \gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right) = 0 \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{5.5}$$

Proof. For any $\boldsymbol{v}_h \in \boldsymbol{V}_h$

$$\begin{aligned} &a_h\left(\boldsymbol{u}_h, \boldsymbol{v}_h\right) = \left(f, v_h\right)_{\mathcal{T}_h} + \langle g, \mu_h\rangle_{\partial\mathcal{E}_N}, \\ &a_{\mathfrak{h}}\left(\boldsymbol{u}_{\mathfrak{h}}, \gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right) = \left(f, v_h\right)_{\mathcal{T}_{\mathfrak{h}}} + \langle g, \gamma_{\mathcal{F},\mathfrak{h}}\left(\boldsymbol{v}_h\right)\rangle_{\partial\mathcal{E}_N} = \left(f, v_h\right)_{\mathcal{T}_{\mathfrak{h}}} + \langle g, \mu_h\rangle_{\partial\mathcal{E}_N}, \end{aligned}$$

and thus $a_h\left(\boldsymbol{u}_h, \boldsymbol{v}_h\right) = a_{\mathfrak{h}}\left(\boldsymbol{u}_{\mathfrak{h}}, \gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right)$. To show eq. (5.5), it remains to show that $a_h\left(\boldsymbol{u}_h, \boldsymbol{v}_h\right) = a_{\mathfrak{h}}\left(\gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right), \gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right)$. First, note for the element integrals we have,

$$\left(\varepsilon\overline{\nabla}u_h, \overline{\nabla}v_h\right)_{\mathcal{T}_h} - \left(\beta u_h, \nabla v_h\right)_{\mathcal{T}_h} = \left(\varepsilon\overline{\nabla}u_h, \overline{\nabla}v_h\right)_{\mathcal{T}_{\mathfrak{h}}} - \left(\beta u_h, \nabla v_h\right)_{\mathcal{T}_{\mathfrak{h}}},$$

and for the diffusion facet terms,

$$\langle\varepsilon\overline{\nabla}_{\overline{n}}u_h, \left[\boldsymbol{v}_h\right]\rangle_{\mathcal{Q}_h} + \langle\varepsilon\left[\boldsymbol{u}_h\right], \overline{\nabla}_{\overline{n}}v_h\rangle_{\mathcal{Q}_h} = \langle\varepsilon\overline{\nabla}_{\overline{n}}u_h, \left[\gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}} + \langle\varepsilon\left[\gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right)\right], \overline{\nabla}_{\overline{n}}v_h\rangle_{\mathcal{Q}_{\mathfrak{h}}}.$$

Next, since $\left[\gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right] = 0$ on $\mathcal{R}_{\mathfrak{h}} \setminus \mathcal{R}_h$, we have

$$\langle\left(\beta\cdot n\right)\gamma_{\mathcal{F},\mathfrak{h}}\left(\boldsymbol{u}_h\right) + \beta_s\left[\gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right)\right], \left[\gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} = \langle\left(\beta\cdot n\right)\lambda_h + \beta_s\left[\boldsymbol{u}_h\right], \left[\boldsymbol{v}_h\right]\rangle_{\partial\mathcal{T}_h},$$

and similarly, on the Neumann boundary, we have for the advective facet terms

$$\langle\zeta^+\beta\cdot n\gamma_{\mathcal{F},\mathfrak{h}}\left(\boldsymbol{u}_h\right), \gamma_{\mathcal{F},\mathfrak{h}}\left(\boldsymbol{v}_h\right)\rangle_{\partial\mathcal{E}_N} = \langle\zeta^+\beta\cdot n\lambda_h, \mu_h\rangle_{\partial\mathcal{E}_N}.$$

Finally, for the penalty term,

$$\langle\varepsilon\alpha h_K^{-1}\left[\gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right)\right], \left[\gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}} = \langle\varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_h\right], \left[\boldsymbol{v}_h\right]\rangle_{\mathcal{Q}_h},$$

because the spatial element size parameter h_K does not change from $\mathcal{K}$ to $\mathring{\mathcal{K}}$. Therefore, $a_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = a_{\mathfrak{h}}\left(\gamma_{\mathfrak{h}}\left(\boldsymbol{u}_h\right), \gamma_{\mathfrak{h}}\left(\boldsymbol{v}_h\right)\right)$ for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and hence eq. (5.5). ☐

Following [23, Section 4, especially Remark 2], we assume that the following *saturation assumption* holds uniformly on the family of meshes $\{\mathcal{T}_h\}_h$: There exists $\rho < 1$, independent of h_K, $\delta t_{\mathcal{K}}$, and ε, such that:

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \big\|\partial_t(u - u_{\mathfrak{h}})\big\|_{\mathcal{K}}^2 \le \rho^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \big\|\partial_t(u - u_h)\big\|_{\mathcal{K}}^2 . \tag{5.6}$$

With the saturation assumption we prove the following useful theorem.

Theorem 5.3 (Time derivative estimation)**.** *Let u be the solution to eq.* (2.1) *and let $\boldsymbol{u}_h = (u_h, \lambda_h)$ be the solution to eq.* (2.9)*. If the saturation assumption eq.* (5.6) *holds, and if $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, then*

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \big\|\partial_t\left(u - u_h\right)\big\|_{\mathcal{K}}^2 \le cT^2\varepsilon^{-1}\eta^2. \tag{5.7}$$

Proof. By the triangle inequality and eq. (5.6) we find

$$\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t\left(u - u_h\right)\right\|_{\mathcal{K}}^2\Big)^{1/2} \le \frac{1}{1-\rho}\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t\left(u_{\mathfrak{h}} - u_h\right)\right\|_{\mathcal{K}}^2\Big)^{1/2} .$$

By the inf-sup condition eq. (4.2b), we have

$$\Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t\left(u_{\mathfrak{h}} - u_h\right)\right\|_{\mathcal{K}}^2\Big)^{1/2} \le c|\!|\!|\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(\boldsymbol{u}_h)|\!|\!|_{s,\mathfrak{h}} \le c_T \sup_{\boldsymbol{v}_{\mathfrak{h}}\in\boldsymbol{V}_{\mathfrak{h}}} \frac{a_{\mathfrak{h}}\left(\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(\boldsymbol{u}_h), \boldsymbol{v}_{\mathfrak{h}}\right)}{|\!|\!|\boldsymbol{v}_{\mathfrak{h}}|\!|\!|_{s,\mathfrak{h}}}. \tag{5.8}$$

Using Galerkin orthogonality eq. (5.5), that Ω_T consists of $\mathcal{R}$-facets only and that $\left(\mathrm{I} - i_h^{\mathcal{F}}\right)\mu_{\mathfrak{h}}$ vanishes on $\mathcal{R}$-facets, we have

$$\begin{aligned} a_{\mathfrak{h}}(\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(\boldsymbol{u}_h), \boldsymbol{v}_{\mathfrak{h}}) &= a_{\mathfrak{h}}(\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(\boldsymbol{u}_h), \left(\mathrm{I} - i_h\right)\boldsymbol{v}_{\mathfrak{h}}) \\ &= (f, (\mathrm{I} - i_h^{\mathcal{K}})\, v_{\mathfrak{h}})_{\mathcal{T}_{\mathfrak{h}}} + \langle g, (\mathrm{I} - i_h^{\mathcal{F}})\,\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_{\mathfrak{h}}} - a_{\mathfrak{h}}(\gamma_{\mathfrak{h}}(\boldsymbol{u}_h), \left(\mathrm{I} - i_h\right)\boldsymbol{v}_{\mathfrak{h}}). \end{aligned}$$

Using integration by parts on $(\varepsilon\overline{\nabla}u_{\mathfrak{h}}, \overline{\nabla}v_{\mathfrak{h}})_{\mathcal{T}_{\mathfrak{h}}}$ and $\left(\beta u_{\mathfrak{h}}, \nabla v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}}$, using the definition of the residual $R_{\mathfrak{h}}^{\mathcal{K}}$, and applying the Dirichlet and the Neumann boundary conditions, we have

$$\begin{aligned} a_{\mathfrak{h}}(\boldsymbol{u}_{\mathfrak{h}}, \boldsymbol{v}_{\mathfrak{h}}) = {} & \left(f - R_{\mathfrak{h}}^{\mathcal{K}}, v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} + \langle\varepsilon\overline{\nabla}_{\overline{n}}u_{\mathfrak{h}}, \mu_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}\setminus\partial\mathcal{E}} - \langle\beta\cdot n u_{\mathfrak{h}}, \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{T}_{\mathfrak{h}}\setminus\partial\mathcal{E}} \\ & + \langle\varepsilon\overline{\nabla}_{\overline{n}}u_{\mathfrak{h}} - \zeta^-\beta\cdot n u_{\mathfrak{h}}, \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} - \langle\zeta^+\beta\cdot n\left[\boldsymbol{u}_{\mathfrak{h}}\right], \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} \\ & - \langle\varepsilon\left[\boldsymbol{u}_{\mathfrak{h}}\right], \overline{\nabla}_{\overline{n}}v_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}} + \langle\varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_{\mathfrak{h}}\right], \left[\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}} + \langle\left(\beta_s - \beta\cdot n\right)\left[\boldsymbol{u}_{\mathfrak{h}}\right], \left[\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}}. \end{aligned} \tag{5.9}$$

We will replace the trial and test functions $(\boldsymbol{u}_{\mathfrak{h}}, \boldsymbol{v}_{\mathfrak{h}})$ in the above by the trial and test functions $(\gamma_{\mathfrak{h}}(\boldsymbol{u}_h), (I - i_h)\boldsymbol{v}_{\mathfrak{h}})$. To simplify what follows, we write out the definitions of these trial and test functions:

$$\gamma_{\mathfrak{h}}(\boldsymbol{u}_h) = (u_h, \gamma_{\mathcal{F},\mathfrak{h}}(\boldsymbol{u}_h)) = \begin{cases} (u_h, \lambda_h), & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ (u_h, u_h), & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \backslash \mathcal{F}_{\mathcal{R},h}, \end{cases}$$

and

$$\begin{aligned} (I - i_h)\boldsymbol{v}_{\mathfrak{h}} &= \boldsymbol{v}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{F}}(\mu_{\mathfrak{h}})) = \boldsymbol{v}_{\mathfrak{h}} - (i_h^{\mathcal{K}}(v_{\mathfrak{h}}), \gamma_{\mathcal{F},\mathfrak{h}}(i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{F}}(\mu_{\mathfrak{h}}))) \\ &= \boldsymbol{v}_{\mathfrak{h}} - \begin{cases} (i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{F}}(\mu_{\mathfrak{h}})), & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ (i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{K}}(v_{\mathfrak{h}})), & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \backslash \mathcal{F}_{\mathcal{R},h}, \end{cases} \\ &= \begin{cases} (v_{\mathfrak{h}}, \mu_{\mathfrak{h}}) - (i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{F}}(\mu_{\mathfrak{h}})), & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ (v_{\mathfrak{h}}, \mu_{\mathfrak{h}}) - (i_h^{\mathcal{K}}(v_{\mathfrak{h}}), i_h^{\mathcal{K}}(v_{\mathfrak{h}})), & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \backslash \mathcal{F}_{\mathcal{R},h}, \end{cases} \\ &= \begin{cases} ((I - i_h^{\mathcal{K}})v_{\mathfrak{h}}, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}), & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ ((I - i_h^{\mathcal{K}})v_{\mathfrak{h}}, \mu_{\mathfrak{h}} - i_h^{\mathcal{K}}(v_{\mathfrak{h}})), & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \backslash \mathcal{F}_{\mathcal{R},h}. \end{cases} \end{aligned}$$

We now consider each term of $a_{\mathfrak{h}}(\gamma_{\mathfrak{h}}(\boldsymbol{u}_h), (I - i_h)\boldsymbol{v}_{\mathfrak{h}})$ separately. First, let us note that

$$\begin{aligned} \left(f - R_{\mathfrak{h}}^{\mathcal{K}}, v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} &\to \left(f - R_{\mathfrak{h}}^{\mathcal{K}}, (I - i_h^{\mathcal{K}})v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} \\ \langle \varepsilon \overline{\nabla}_{\overline{n}} u_{\mathfrak{h}}, \mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_{\mathfrak{h}} \backslash \partial\mathcal{E}} &\to \langle \varepsilon \overline{\nabla}_{\overline{n}} u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_{\mathfrak{h}} \backslash \partial\mathcal{E}} \end{aligned}$$

Next,

$$\langle \beta \cdot n u_{\mathfrak{h}}, \mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} \to \begin{cases} \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h} \\ \langle \beta \cdot n u_h, \mu_{\mathfrak{h}} - i_h^{\mathcal{K}}(v_{\mathfrak{h}}) \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h} \end{cases}$$

Note, however, that $\langle \beta \cdot n u_h, \mu_{\mathfrak{h}} - i_h^{\mathcal{K}}(v_{\mathfrak{h}}) \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} = 0$ on facets $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$. This is because $\mu_{\mathfrak{h}} - i_h^{\mathcal{K}}(v_{\mathfrak{h}})$ and u_h are single-valued on $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$ and $\beta \cdot n^- = -\beta \cdot n^+$ on $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$. Let $\partial\mathcal{T}_{\mathfrak{h}}^{\mathcal{R}}$ denote that set $\partial\mathcal{T}_{\mathfrak{h}}$ excluding all facets $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$. We may therefore write:

$$\begin{aligned} \langle \beta \cdot n u_{\mathfrak{h}}, \mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} &\to \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}}^{\mathcal{R}} \backslash \partial\mathcal{E}} \\ &= \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}}^{\mathcal{R}} \backslash \partial\mathcal{E}} + \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\mathcal{F}_{\mathcal{R},\mathfrak{h}} \backslash \mathcal{F}_{\mathcal{R},h}} \\ &= \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \backslash \partial\mathcal{E}} \end{aligned}$$

where we added the zero term $\langle \beta\cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\mathcal{F}_{\mathcal{R},\mathfrak{h}}\setminus\mathcal{F}_{\mathcal{R},h}}$ (indeed, by definition of $i_h^{\mathcal{F}}$ on $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$, $(I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}} = 0$). Next, we have:

$$\begin{aligned}
\langle \varepsilon \left[\boldsymbol{u}_{\mathfrak{h}}\right], \overline{\nabla}_{\overline{n}} v_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}} &\to \langle \varepsilon \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \overline{\nabla}_{\overline{n}}((I - i_h^{\mathcal{K}})v_{\mathfrak{h}})\rangle_{\mathcal{Q}_{\mathfrak{h}}}, \\
\langle \varepsilon\alpha h_K^{-1} \left[\boldsymbol{u}_{\mathfrak{h}}\right], \left[\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}} &\to \langle \varepsilon\alpha h_K^{-1} \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I - i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}}, \\
\langle (\beta_s - \beta\cdot n) \left[\boldsymbol{u}_{\mathfrak{h}}\right], \left[\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} &\to \langle (\beta_s - \beta\cdot n) \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I - i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}}, \\
\langle \zeta^+\beta\cdot n \left[\boldsymbol{u}_{\mathfrak{h}}\right], \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} &\to \langle \zeta^+\beta\cdot n \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N}, \\
\langle \varepsilon\overline{\nabla}_{\overline{n}} u_{\mathfrak{h}} - \zeta^-\beta\cdot n u_{\mathfrak{h}}, \mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} &\to \langle \varepsilon\overline{\nabla}_{\overline{n}} u_h - \zeta^-\beta\cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}\cap\partial\mathcal{E}_N},
\end{aligned}$$

where the third term can be divided into two cases

$$\begin{aligned}
&\langle (\beta_s - \beta\cdot n) \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I - i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} \\
&\qquad = \begin{cases} \langle (\beta_s - \beta\cdot n) \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I - i_h^{\mathcal{K}})v_{\mathfrak{h}} - (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} & \forall F \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}} \cup \mathcal{F}_{\mathcal{R},h}, \\ \langle (\beta_s - \beta\cdot n) \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I - i_h^{\mathcal{K}})v_{\mathfrak{h}} - (\mu_{\mathfrak{h}} - i_h^{\mathcal{K}}(v_{\mathfrak{h}}))\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} = 0 & \forall F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}. \end{cases}
\end{aligned}$$

Returning to eq. (5.9), we find

$$\begin{aligned}
&a_{\mathfrak{h}}(\boldsymbol{u}_{\mathfrak{h}} - \gamma_{\mathfrak{h}}(\boldsymbol{u}_h), \boldsymbol{v}_{\mathfrak{h}}) \\
&\qquad = \left(R_{\mathfrak{h}}^{\mathcal{K}}, (I - i_h^{\mathcal{K}})v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} \\
&\qquad\quad + \left[-\langle \varepsilon\overline{\nabla}_{\overline{n}} u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}\setminus\partial\mathcal{E}} + \langle \beta\cdot n u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{T}_{\mathfrak{h}}\setminus\partial\mathcal{E}}\right] \\
&\qquad\quad + \langle \varepsilon \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \overline{\nabla}_{\overline{n}}((I - i_h^{\mathcal{K}})v_{\mathfrak{h}})\rangle_{\mathcal{Q}_{\mathfrak{h}}} \\
&\qquad\quad - \langle \varepsilon\alpha h_K^{-1} \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I - i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\mathcal{Q}_{\mathfrak{h}}} \\
&\qquad\quad + \left[-\langle (\beta_s - \beta\cdot n) \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I - i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}} + \langle \zeta^+\beta\cdot n \left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N}\right] \\
&\qquad\quad + \langle g - \varepsilon\overline{\nabla}_{\overline{n}} u_h + \zeta^-(\beta\cdot n)u_h, (I - i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\mathcal{Q}_{\mathfrak{h}}\cap\partial\mathcal{E}_N} \\
&\qquad =: M_1 + M_2 + M_3 + M_4 + M_5 + M_6.
\end{aligned} \tag{5.10}$$

We will bound the M_i's separately.

Bound for M_1. M_1 is bounded using the Cauchy–Schwarz inequality, eq. (3.31a), eqs. (3.35a) and (3.35b), and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$:

$$\begin{aligned}
&\left(R_h^{\mathcal{K}}, (I - i_h^{\mathcal{K}})\, v_{\mathfrak{h}}\right)_{\mathcal{T}_{\mathfrak{h}}} = \left(R_h^{\mathcal{K}}, (I - i_h^{\mathcal{K}})\, v_{\mathfrak{h}}\right)_{\mathcal{T}_h} \\
\leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \lambda_{\mathcal{K}} \|R_h^{\mathcal{K}}\|_{\mathcal{K}} \max\{h_K^{-1}\varepsilon^{1/2}, 1\} \|(I - i_h^{\mathcal{K}})\, v_{\mathfrak{h}}\|_{\mathcal{K}} \\
\leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \lambda_{\mathcal{K}} \|R_h^{\mathcal{K}}\|_{\mathcal{K}} \max\{\varepsilon^{1/2}, h_K\} \left(\sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} \left[\boldsymbol{v}_{\mathfrak{h}}\right] \right\|_{\partial\mathring{\mathcal{K}}\cap F_{\mathring{\mathcal{R}}}} + \sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}} \delta t_{\mathcal{K}} h_K^{-1} \left\|\partial_t v_{\mathfrak{h}}\right\|_{\mathring{\mathcal{K}}} \right).
\end{aligned}$$

On elements where $\max\{\varepsilon^{1/2}, h_K\} = \varepsilon^{1/2}$, using that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we find $\varepsilon^{1/2}\delta t_{\mathcal{K}} h_K^{-1} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} \leq c\tau_\varepsilon^{1/2} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}$. On elements where $\max\{\varepsilon^{1/2}, h_K\} = h_K$ we have, by eq. (3.10a), $h_K \delta t_{\mathcal{K}} h_K^{-1} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} \leq c \|v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}$. Therefore,

$$\max\{\varepsilon^{1/2}, h_K\} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \delta t_{\mathcal{K}} h_K^{-1} \big\|\partial_t v_{\mathfrak{h}}\big\|_{\mathring{\mathcal{K}}} \leq \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \tau_\varepsilon^{1/2} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} + \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \|v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}.$$

Using furthermore that $\max\{\varepsilon^{1/2}, h_K\} \leq 1$ and Hölder's inequality for sums, we find

$$\begin{aligned} M_1 \leq & c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_R^{\mathcal{K}})^2 \Big)^{1/2} \Bigg[\sum_{\mathcal{K} \in \mathcal{T}_h} \Bigg(\sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \Big(\big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} [\boldsymbol{v}_{\mathfrak{h}}] \big\|_{\mathcal{R}_{\mathring{\mathcal{K}}}} + \tau_\varepsilon^{1/2} \|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} + \|v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}} \Big) \Bigg)^2 \Bigg]^{1/2} \\ \leq & c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_R^{\mathcal{K}})^2 \Big)^{1/2} |||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}. \end{aligned}$$

Bound for M_2. We first write

$$M_2 = \underbrace{-\langle \varepsilon \overline{\nabla}_{\overline{n}} u_h, (I - i_h^{\mathcal{F}}) \mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_{\mathfrak{h}} \setminus \partial\mathcal{E}}}_{M_{21}} + \underbrace{\langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}}) \mu_{\mathfrak{h}} \rangle_{\partial\mathcal{T}_{\mathfrak{h}} \setminus \partial\mathcal{E}}}_{M_{22}}.$$

For M_{21}, using that $\langle \cdot, \cdot \rangle_{\mathcal{Q}_{\mathfrak{h}} \setminus \partial\mathcal{E}} = \langle \cdot, \cdot \rangle_{\mathcal{Q}_h \setminus \partial\mathcal{E}}$, writing element-wise integrals as facet integrals on interior facets, using the Cauchy–Schwarz inequality and the projection estimate eq. (3.31b), we find

$$\begin{aligned} M_{21} \leq & c \sum_{F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}^i} \big| \langle \llbracket \varepsilon \overline{\nabla}_{\overline{n}} u_h \rrbracket, (I - i_h^{\mathcal{F}}) \mu_{\mathfrak{h}} \rangle_{F_{\mathcal{Q}}} \big| \\ \leq & c \sum_{F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}^i} \delta t_{\mathcal{K}}^{1/2} \|\llbracket \varepsilon \overline{\nabla}_{\overline{n}} u_h \rrbracket\|_{F_{\mathcal{Q}}} \big(\|\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} + \delta t_{\mathcal{K}} \|\langle\!\langle \partial_t \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} \big), \end{aligned} \tag{5.11}$$

where $\mathcal{K}$ in the last step is chosen such that $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathcal{K}}$. Consider the two terms on the right-hand side of eq. (5.11) separately. First, using eq. (3.35c) and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we have

$$\begin{aligned} & \sum_{F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}^i} \|\llbracket \varepsilon \overline{\nabla}_{\overline{n}} u_h \rrbracket\|_{F_{\mathcal{Q}}} \delta t_{\mathcal{K}}^{1/2} \|\langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle\|_{E_{\mathring{\mathcal{K}}}} \\ \leq & c \sum_{F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}^i} h_K^{1/2} \varepsilon^{-1/2} \|\llbracket \varepsilon \overline{\nabla}_{\overline{n}} u_h \rrbracket\|_{F_{\mathcal{Q}}} \\ & \quad \cdot \Big(\delta t_{\mathcal{K}}^{1/2} h_K^{-1} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \big\| |\beta_s - \tfrac{1}{2}\beta \cdot n|^{1/2} [\boldsymbol{v}_{\mathfrak{h}}] \big\|_{\partial\mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} + \varepsilon^{1/2} h_K^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h}}^*]\|_{F_{\mathcal{Q}}^*} + \varepsilon^{1/2} h_K^{-1/2} \|[\boldsymbol{v}_{\mathfrak{h},*}]\|_{F_{\mathcal{Q},*}} \Big) \\ \leq & c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,1}^{\mathcal{K}})^2 \Big)^{1/2} |||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}. \end{aligned}$$

The second term on the right-hand side of eq. (5.11) can be bounded similarly by using eq. (3.35d) and $\delta t_{\mathcal{K}} h_K^{-1} \varepsilon^{1/2} \le c \tau_\varepsilon^{1/2}$:

$$\sum_{F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},h}^i} \| [\![\varepsilon \overline{\nabla}_{\overline{n}} u_h]\!] \|_{F_{\mathcal{Q}}} \, \delta t_{\mathcal{K}}^{3/2} \, \| \langle\!\langle \partial_t \mu_{\mathfrak{h}} \rangle\!\rangle \|_{E_{\mathring{\mathcal{K}}}} \le c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,1}^{\mathcal{K}})^2 \Big)^{1/2} |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.$$

Therefore, we have that

$$M_{21} \le c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,1}^{\mathcal{K}})^2 \Big)^{1/2} |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.$$

For M_{22}, we first note that since $(I - i_h^{\mathcal{F}})\, \mu_{\mathfrak{h}}$ vanishes on $\mathcal{R}_{\mathfrak{h}}$ we have that

$$M_{22} = \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}}) \mu_{\mathfrak{h}} \rangle_{\partial \mathcal{T}_{\mathfrak{h}} \backslash \partial \mathcal{E}} = \langle \beta \cdot n u_h, (I - i_h^{\mathcal{F}}) \mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_{\mathfrak{h}} \backslash \partial \mathcal{E}}.$$

Then, similar to eq. (5.11), we have using eqs. (3.31b), (3.35c) and (3.35d), that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, $\varepsilon \le 1$, $h_K \le 1$, that $\|\bar{\beta}\|_{L^\infty(\mathcal{E})} \le 1$, and noting that $\langle \beta \cdot n \lambda_h, (I - i_h^{\mathcal{F}})\, \mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_h \backslash \partial \mathcal{E}} = 0$ by single-valuedness of λ_h, $\beta \cdot n$, and $(I - i_h^{\mathcal{F}})\, \mu_{\mathfrak{h}}$ on element boundaries,

$$\begin{aligned}
M_{22} =& \langle \beta \cdot n \left[\boldsymbol{u}_h \right], (I - i_h^{\mathcal{F}})\, \mu_{\mathfrak{h}} \rangle_{\mathcal{Q}_h \backslash \partial \mathcal{E}} \\
\le& c \sum_{\mathcal{K} \in \mathcal{T}_h} \delta t_{\mathcal{K}}^{1/2} \, \| [\boldsymbol{u}_h] \|_{\mathcal{Q}_{\mathcal{K}}} \big(\| \langle\!\langle \mu_{\mathfrak{h}} \rangle\!\rangle \|_{E_{\mathring{\mathcal{K}}}} + \delta t_{\mathcal{K}} \, \| \langle\!\langle \partial_t \mu_{\mathfrak{h}} \rangle\!\rangle \|_{E_{\mathring{\mathcal{K}}}} \big) \\
\le& c \sum_{\mathcal{K} \in \mathcal{T}_h} \| [\boldsymbol{u}_h] \|_{\mathcal{Q}_{\mathcal{K}}} \cdot \Bigg(\sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \delta t_{\mathcal{K}}^{1/2} h_K^{-1/2} \, \| |\beta_s - \tfrac{1}{2} \beta \cdot n|^{1/2} \left[\boldsymbol{v}_{\mathfrak{h}} \right] \|_{\partial \mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} \\
& \qquad + \delta t_{\mathcal{K}} h_K^{-1/2} \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \| \partial_t v_{\mathfrak{h}} \|_{\mathring{\mathcal{K}}} + \| [\boldsymbol{v}_{\mathfrak{h}}^*] \|_{F_{\mathcal{Q}}^*} + \| [\boldsymbol{v}_{\mathfrak{h},*}] \|_{F_{\mathcal{Q},*}} \Bigg) \\
\le& c \sum_{\mathcal{K} \in \mathcal{T}_h} h_K^{1/2} \varepsilon^{-1/2} \, \| [\boldsymbol{u}_h] \|_{\mathcal{Q}_{\mathcal{K}}} \cdot \Bigg(\sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \| |\beta_s - \tfrac{1}{2} \beta \cdot n|^{1/2} \left[\boldsymbol{v}_{\mathfrak{h}} \right] \|_{\partial \mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} \\
& \qquad + \sum_{\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}} \tau_\varepsilon^{1/2} \, \| \partial_t v_{\mathfrak{h}} \|_{\mathring{\mathcal{K}}} + \varepsilon^{1/2} h_K^{-1/2} \, \| [\boldsymbol{v}_{\mathfrak{h}}^*] \|_{F_{\mathcal{Q}}^*} + \varepsilon^{1/2} h_K^{-1/2} \, \| [\boldsymbol{v}_{\mathfrak{h},*}] \|_{F_{\mathcal{Q},*}} \Bigg) \\
\le& c \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,2,2}^{\mathcal{K}})^2 \Big)^{1/2} |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.
\end{aligned}$$

Combining the bounds for M_{21} and M_{22} we obtain:

$$M_2 \le c \Big[\Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,1}^{\mathcal{K}})^2 \Big)^{1/2} + \Big(\sum_{\mathcal{K} \in \mathcal{T}_h} (\eta_{J,2,2}^{\mathcal{K}})^2 \Big)^{1/2} \Big] |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.$$

Bound for M_3. For M_3, using the Cauchy–Schwarz inequality, the trace inequality eq. (3.7c), the inverse inequality eq. (3.7b), the subgrid projection estimate eq. (3.31a), eqs. (3.35a) and (3.35b), that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, and Hölder's inequality for sums,

$$\begin{aligned}
M_3 \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}}} h_K^{1/2} \left\| \overline{\nabla} \left(\left(I - i_h^{\mathcal{K}} \right) v_{\mathfrak{h}} \right) \right\|_{\mathcal{Q}_{\mathcal{K}}} \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}}} h_K^{-1} \left(\delta t_{\mathcal{K}}^{1/2} \left\| [\![v_{\mathfrak{h}}]\!] \right\|_{F_{\mathring{\mathcal{R}}}} + \delta t_{\mathcal{K}}^{3/2} \left\| [\![\partial_t v_{\mathfrak{h}}]\!] \right\|_{F_{\mathring{\mathcal{R}}}} \right) \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}}} h_K^{-1} \delta t_{\mathcal{K}}^{1/2} \\
& \quad \cdot \Big(\sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}} \left\| \left| \beta_s - \tfrac{1}{2} \beta \cdot n \right|^{1/2} [\boldsymbol{v}_{\mathfrak{h}}] \right\|_{\partial\mathring{\mathcal{K}} \cap F_{\mathring{\mathcal{R}}}} + \sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}} \tau_{\varepsilon}^{1/2} \left\| \partial_t v_{\mathfrak{h}} \right\|_{\mathring{\mathcal{K}}} \Big) \\
\leq & c \Big(\sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}} \right)^2 \Big)^{1/2} |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.
\end{aligned}$$

Bound for M_4. Let $\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}$ and $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathring{\mathcal{K}}}$. We write $M_4 := M_{41} + M_{42}$ where M_{41} is the sum of integrals over $F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$ and M_{42} the sum of integrals over $F_{\mathcal{Q}} \notin \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$. The latter case occurs when the neighboring element of $\mathring{\mathcal{K}}$ over $F_{\mathcal{Q}}$ is coarser than $\mathring{\mathcal{K}}$. To bound M_{41}, we first note that for $F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$, we have

$$[(I - i_h) \boldsymbol{v}_{\mathfrak{h}}] = \left(I - i_h^{\mathcal{K}} \right) v_{\mathfrak{h}} - \left(I - i_h^{\mathcal{F}} \right) \mu_{\mathfrak{h}} = \left(I - i_h^{\mathcal{F}} \right) \left(v_{\mathfrak{h}} - \mu_{\mathfrak{h}} \right), \tag{5.12}$$

where the last step is by lemma 3.8. Then, note that by the Cauchy–Schwarz inequality and boundedness of the projection $i_h^{\mathcal{F}}$, we have

$$\langle \varepsilon \alpha h_K^{-1} [\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)], [(I - i_h) \boldsymbol{v}_{\mathfrak{h}}] \rangle_{F_{\mathcal{Q}}} \leq c \left(\varepsilon^{1/2} h_K^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}}} \right) \left(\varepsilon^{1/2} h_K^{-1/2} \left\| [\boldsymbol{v}_{\mathfrak{h}}] \right\|_{\mathcal{Q}_{\mathcal{K}}} \right), \tag{5.13}$$

so that

$$M_{41} \leq c \left(\sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}} \right)^2 \right)^{1/2} |\!|\!| \boldsymbol{v}_{\mathfrak{h}} |\!|\!|_{s,\mathfrak{h}}.$$

We now consider M_{42}. Consider an $F_{\mathcal{Q}} \notin \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$. Denote the coarser neighboring element of $\mathring{\mathcal{K}}$ over $F_{\mathcal{Q}}$ by $\mathring{\mathcal{K}}_{nb}$ and denote the restriction of $v_{\mathfrak{h}}$ to $\mathring{\mathcal{K}}_{nb}$ by $v_{nb,\mathfrak{h}}$. We have

$$\begin{aligned}
[(I - i_h) \boldsymbol{v}_{\mathfrak{h}}] &= \left(I - i_h^{\mathcal{K}} \right) v_{\mathfrak{h}} + v_{nb,\mathfrak{h}} - \mu_{\mathfrak{h}} + i_h^{\mathcal{F}} \mu_{\mathfrak{h}} - v_{nb,\mathfrak{h}} \\
&= \left(I - i_h^{\mathcal{K}} \right) v_{\mathfrak{h}} + \left(I - i_h^{\mathcal{F}} \right) \left(v_{nb,\mathfrak{h}} - \mu_{\mathfrak{h}} \right) - \left(I - i_h^{\mathcal{K}} \right) v_{nb,\mathfrak{h}},
\end{aligned} \tag{5.14}$$

where the last step is by lemma 3.8. We have:

$$\begin{aligned}\langle \varepsilon\alpha h_K^{-1}\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I-i_h)\,\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{F_{\mathcal{Q}}} =& \langle \varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_h\right], (I-i_h^{\mathcal{K}})\,v_{\mathfrak{h}}\rangle_{F_{\mathcal{Q}}} \\ &+ \langle \varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_h\right], (I-i_h^{\mathcal{F}})\left[\boldsymbol{v}_{nb,\mathfrak{h}}\right]\rangle_{F_{\mathcal{Q}}} \\ &- \langle \varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_h\right], (I-i_h^{\mathcal{K}})\,v_{nb,\mathfrak{h}}\rangle_{F_{\mathcal{Q}}}.\end{aligned} \tag{5.15}$$

The second term on the right-hand side of eq. (5.15) is bounded in the same way as eq. (5.13). For the first term on the right-hand side of eq. (5.15), using the Cauchy–Schwarz inequality, the trace inequality eq. (3.7c), the subgrid projection bound eq. (3.31a), that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, and eqs. (3.35a) and (3.35b), we find

$$\begin{aligned}&\langle \varepsilon\alpha h_K^{-1}\left[\boldsymbol{u}_h\right], (I-i_h^{\mathcal{K}})\,v_{\mathfrak{h}}\rangle_{F_{\mathcal{Q}}} \\ &\qquad\le c\left(\varepsilon^{1/2}h_K^{-1/2}\left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_{\mathcal{K}}}\right)\left(\sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}}\left(\left\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{v}_{\mathfrak{h}}\right]\right\|_{\partial\mathring{\mathcal{K}}\cap F_{\mathring{\mathcal{R}}}} + \tau_\varepsilon^{1/2}\left\|\partial_t v_{\mathfrak{h}}\right\|_{\mathring{\mathcal{K}}}\right)\right).\end{aligned}$$

The third term on the right-hand side of eq. (5.15) is bound in the same way. For M_{42} we therefore find that

$$M_{42} \le c\left(\sum_{\mathcal{K}\in\mathcal{T}_h}\left(\eta_{J,2,1}^{K}\right)^2\right)^{1/2}|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

Combining the bounds for M_{41} and M_{42},

$$M_4 \le c\left(\sum_{\mathcal{K}\in\mathcal{T}_h}\left(\eta_{J,2,1}^{\mathcal{K}}\right)^2\right)^{1/2}|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

Bound for M_5. For M_5 we first write

$$M_5 = -\underbrace{\langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I-i_h)\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{\partial\mathcal{T}_{\mathfrak{h}}}}_{M_{51}} + \underbrace{\langle\zeta^+\beta\cdot n\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I-i_h^{\mathcal{F}})\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N}}_{M_{52}}.$$

To bound M_{51} we consider the $\mathcal{Q}$-facets and $\mathcal{R}$-facets separately. For the $\mathcal{Q}$-facets we follow the same steps as used in bounding M_4. Let $\mathring{\mathcal{K}} \in \mathcal{T}_{\mathcal{K}}$ and $F_{\mathcal{Q}} \subset \mathcal{Q}_{\mathring{\mathcal{K}}}$. If $F_{\mathcal{Q}} \in \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$, we use eq. (5.12), the Cauchy–Schwarz inequality, boundedness of the projection $i_h^{\mathcal{F}}$, and eq. (2.8):

$$\langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I-i_h)\,\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{F_{\mathcal{Q}}} \le c\left\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_{\mathcal{K}}}\left\|\left|\beta_s-\tfrac{1}{2}\beta\cdot n\right|^{1/2}\left[\boldsymbol{v}_{\mathfrak{h}}\right]\right\|_{\mathcal{Q}_{\mathcal{K}}}. \tag{5.16}$$

If $F_{\mathcal{Q}} \notin \mathcal{F}_{\mathcal{Q},\mathfrak{h}}$, we have, using eq. (5.14),

$$\begin{aligned}\langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], \left[(I-i_h)\,\boldsymbol{v}_{\mathfrak{h}}\right]\rangle_{F_{\mathcal{Q}}} =& \langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I-i_h^{\mathcal{K}})\,v_{\mathfrak{h}}\rangle_{F_{\mathcal{Q}}} \\ &\langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I-i_h^{\mathcal{F}})\left[\boldsymbol{v}_{nb,\mathfrak{h}}\right]\rangle_{F_{\mathcal{Q}}} \\ &\langle(\beta_s-\beta\cdot n)\left[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)\right], (I-i_h^{\mathcal{K}})\,v_{nb,\mathfrak{h}}\rangle_{F_{\mathcal{Q}}}.\end{aligned} \tag{5.17}$$

The second term on the right-hand side of eq. (5.17) is bounded in the same way as eq. (5.16). For the first term on the right-hand side of eq. (5.17) we use the Cauchy–Schwarz inequality, the trace inequality eq. (3.7c), the subgrid projection estimate eq. (3.31a), the estimates eqs. (3.35a) and (3.35b), and eq. (2.8) to find:

$$\begin{aligned}&\langle(\beta_s - \beta\cdot n)\,[\boldsymbol{u}_h]\,,(I - i_h^{\mathcal{K}})\,v_{\mathfrak{h}}\rangle_{F_{\mathcal{Q}}}\\ &\quad\le c\varepsilon^{-1/2}\,\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\,[\boldsymbol{u}_h]\|_{\mathcal{Q}_{\mathcal{K}}}\Big(\sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}}\big(\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\,[\boldsymbol{v}_{\mathfrak{h}}]\|_{\partial\mathring{\mathcal{K}}\cap F_{\mathring{\mathcal{R}}}} + \tau_\varepsilon^{1/2}\,\|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}\big)\Big)\,.\end{aligned} \tag{5.18}$$

The third term on the right-hand side of eq. (5.17) is bounded in the same way. Combining eqs. (5.16) and (5.18), we bound the contributions from the $\mathcal{Q}$-facets to M_{51} as follows:

$$\langle(\beta_s - \beta\cdot n)\,[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)]\,,[(I - i_h)\,\boldsymbol{v}_{\mathfrak{h}}]\rangle_{\mathcal{Q}_{\mathfrak{h}}} \le c\varepsilon^{-1/2}\,\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\big)^2\Big)^{1/2}\,|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}. \tag{5.19}$$

Next we consider the contributions of the $\mathcal{R}$-facets to M_{51}. Using that $(I - i_h^{\mathcal{F}})\,\mu_{\mathfrak{h}} = 0$ on $F \in \mathcal{F}_{\mathcal{R},h}$, and that $[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)] = 0$ on $F \in \mathcal{F}_{\mathcal{R},\mathfrak{h}} \setminus \mathcal{F}_{\mathcal{R},h}$, using the Cauchy–Schwarz inequality, the trace inequality eq. (3.10b), the subgrid projection estimate eq. (3.31a), the estimates eqs. (3.35a) and (3.35b) the inverse estimate eq. (3.11a), we find

$$\begin{aligned}\langle(\beta_s - \beta\cdot n)\,[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)]\,,[(I - i_h)\,\boldsymbol{v}_{\mathfrak{h}}]\rangle_{\mathcal{R}_{\mathfrak{h}}} &= \langle(\beta_s - \beta\cdot n)\,[\boldsymbol{u}_h]\,,(I - i_h^{\mathcal{K}})\,v_{\mathfrak{h}}\rangle_{\mathcal{R}_h}\\ &\le c\sum_{\mathcal{K}\in\mathcal{T}_h}\varepsilon^{-1/2}\,\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\,[\boldsymbol{u}_h]\|_{\mathcal{R}_{\mathcal{K}}}\\ &\qquad\cdot\Big(\sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}}\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\,[\boldsymbol{v}_{\mathfrak{h}}]\|_{\partial\mathring{\mathcal{K}}} + \sum_{\mathring{\mathcal{K}}\in\mathcal{T}_{\mathcal{K}}}\tau_\varepsilon^{1/2}\|\partial_t v_{\mathfrak{h}}\|_{\mathring{\mathcal{K}}}\Big)\\ &\le c\varepsilon^{-1/2}\,\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\big)^2\Big)^{1/2}\,|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.\end{aligned} \tag{5.20}$$

We can now bound M_{51} by combining eqs. (5.19) and (5.20):

$$M_{51} \le c\varepsilon^{-1/2}\,\Big[\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\big)^2\Big)^{1/2} + \Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\big)^2\Big)^{1/2}\Big]\,|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

For M_{52} we use that $(I - i_h^{\mathcal{F}})\,\mu_{\mathfrak{h}} = 0$ on $F \in \mathcal{F}_{\mathcal{R},h}$, the Cauchy–Schwarz inequality, the boundedness of the projection $i_h^{\mathcal{F}}$, and eq. (2.8) to find that

$$\langle\zeta^+\beta\cdot n\,[\gamma_{\mathfrak{h}}(\boldsymbol{u}_h)]\,,(I - i_h^{\mathcal{F}})\,\mu_{\mathfrak{h}}\rangle_{\partial\mathcal{E}_N} \le c\sum_{\mathcal{K}\in\mathcal{T}_h}\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\,[\boldsymbol{u}_h]\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}\,\beta_s^{1/2}\,\|\mu_{\mathfrak{h}}\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}\,. \tag{5.21}$$

To bound $\beta_s^{1/2}\left\|\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}$, consider a facet $F_{\mathcal{Q}}\subset\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N$. By the mean value theorem for definite integrals (see, for example, [5, Theorem 14.16]), there exists $(t_m,x_m)\in F_{\mathcal{Q}}$ such that

$$\left\||\beta\cdot n|^{1/2}\mu_{\mathfrak{h}}\right\|_{F_{\mathcal{Q}}}^2=\int_{F_{\mathcal{Q}}}|\beta\cdot n|\,\mu_{\mathfrak{h}}^2\,\mathrm{d}s=|\beta\cdot n|\,|_{(t_m,x_m)}\int_{F_{\mathcal{Q}}}\mu_{\mathfrak{h}}^2\,\mathrm{d}s=|\beta\cdot n|\,|_{(t_m,x_m)}\left\|\mu_{\mathfrak{h}}\right\|_{F_{\mathcal{Q}}}^2.\tag{5.22}$$

Let $(t_M,x_M)\in F_{\mathcal{Q}}$ be the point on which $|\beta\cdot n|$ attains its maximum β_s on $F_{\mathcal{Q}}$. Since β is Lipschitz continuous and n is constant on $F_{\mathcal{Q}}$ (since $\mathcal{Q}$-facets are flat), we deduce that $\beta\cdot n$ is Lipschitz continuous on $F_{\mathcal{Q}}$. Thus, using that $\delta t_{\mathcal{K}}\le h_K$, we have

$$\left|\beta_s-|\beta\cdot n|\,|_{(t_m,x_m)}\right|\le c\left|(t_m,x_m)-(t_M,x_M)\right|\le ch_K.\tag{5.23}$$

A consequence of eq. (5.22), eq. (5.23), and eq. (3.7c) is the following bound:

$$\begin{aligned}\beta_s\left\|\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2&\le|\beta\cdot n|\,|_{(t_m,x_m)}\left\|\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2+ch_K\left\|\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2\\&\le\left\||\beta\cdot n|^{1/2}\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2+ch_K\left\|v_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2+ch_K\left\|[\boldsymbol{v}_{\mathfrak{h}}]\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2\\&\le c\varepsilon^{-1}\left[\left\||\beta\cdot n|^{1/2}\mu_{\mathfrak{h}}\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2+\left\|v_{\mathfrak{h}}\right\|_{\mathcal{K}}^2+\varepsilon h_K^{-1}\left\|[\boldsymbol{v}_{\mathfrak{h}}]\right\|_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N}^2\right].\end{aligned}\tag{5.24}$$

Combining eqs. (5.21) and (5.24), we find the following bound for M_{52}:

$$M_{52}\le c\varepsilon^{-1/2}\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\big)^2\Big)^{1/2}|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

Combining the bounds for M_{51} and M_{52} we find that

$$M_5\le c\varepsilon^{-1/2}\left[\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\big)^2\Big)^{1/2}+\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\big)^2\Big)^{1/2}\right]|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

Bound for M_6. The derivation of a bound for M_6 is similar to that of the bound for M_{22}:

$$M_6\le c\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{BC,1}^{\mathcal{K}}\big)^2\Big)^{1/2}|||\boldsymbol{v}_{\mathfrak{h}}|||_{s,\mathfrak{h}}.$$

Combining eqs. (5.8) and (5.10) with the bounds for M_1 to M_6 we find:

$$\begin{aligned}\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\tau_\varepsilon\left\|\partial_t\left(u_{\mathfrak{h}}-u_h\right)\right\|_{\mathcal{K}}^2\Big)^{1/2}&\le c_T\Big(\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_R^{\mathcal{K}}\big)^2\Big)^{1/2}+\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,1}^{\mathcal{K}}\big)^2\Big)^{1/2}\\&+\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,2,1}^{\mathcal{K}}\big)^2\Big)^{1/2}+\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,2,2}^{\mathcal{K}}\big)^2\Big)^{1/2}+\varepsilon^{-1/2}\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{J,3}^{\mathcal{K}}\big)^2\Big)^{1/2}+\Big(\sum_{\mathcal{K}\in\mathcal{T}_h}\big(\eta_{BC,1}^{\mathcal{K}}\big)^2\Big)^{1/2}\Big).\end{aligned}$$

Equation (5.7) follows by using Hölder's inequality for sums. □

5.3 Reliability of the error estimator

In this section we prove theorem 5.1. Let $e_u := u - u_h$ denote the true error. To derive an upper bound for e_u we follow [56, 61, 95, 107] and consider the following decomposition of $u_h = \mathcal{I}_h^c u_h + u_h^r$. Here $\mathcal{I}_h^c$ is the averaging operator defined in section 3.5 and $u_h^r := u_h - \mathcal{I}_h^c u_h$. We further introduce $e_u^c := u - \mathcal{I}_h^c u_h$. Applying the triangle inequality,

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u\right\|_{\mathcal{K}}^2 \leq c \left(\sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u^c\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} u_h^r\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|u_h^r\right\|_{\mathcal{K}}^2 \right).$$

In this section we will use the same weighting function as in eq. (4.3)

$$\varphi := eT\exp(-t/T) + \chi,$$

where the positive constant χ will be determined later. We further introduce the following forms (see [56, 95, 109]):

$$\begin{aligned}
k_h(\boldsymbol{u}, \boldsymbol{v}) &= -\langle \varepsilon \left[\boldsymbol{u}\right], \overline{\nabla}_{\overline{n}} v\rangle_{\mathcal{Q}_h} - \langle \varepsilon \overline{\nabla}_{\overline{n}} u, \left[\boldsymbol{v}\right]\rangle_{\mathcal{Q}_h}, \\
b_h(\lambda, \mu) &= \langle \zeta^+ \beta \cdot n \lambda, \mu\rangle_{\partial\mathcal{E}_N}, \\
\widetilde{a}_h(\boldsymbol{u}, \boldsymbol{v}) &= a_h(\boldsymbol{u}, \boldsymbol{v}) - k_h(\boldsymbol{u}, \boldsymbol{v}) - b_h(\lambda, \mu).
\end{aligned}$$

Lemma 5.2. *Let φ be as in eq. (4.3). Then,*

$$\chi \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_{\mathcal{K}}^2 + \tfrac{1}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u^c\right\|_{\mathcal{K}}^2 + \tfrac{1}{2}\chi \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| |\beta \cdot n|^{1/2} e_u^c \right\|_{\partial\mathcal{E}_N}^2 \leq \sum_{i=1}^{6} T_i, \tag{5.25}$$

where

$$\begin{aligned}
T_1 &= \left(R_h^{\mathcal{K}}, (I - \Pi_h)(\varphi e_u^c)\right)_{\mathcal{T}_h}, \\
T_2 &= -\langle \varepsilon \overline{\nabla}_{\overline{n}} u_h, \left(I - \Pi_h^{\mathcal{F}}\right)\left(\varphi e_u^c\right)\rangle_{\mathcal{Q}_h^i} + \langle R_h^N, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N \setminus \Omega_T}, \\
T_3 &= \langle \varepsilon \alpha h_K^{-1} \left[\boldsymbol{u}_h\right], \left(\Pi_h - \Pi_h^{\mathcal{F}}\right)\left(\varphi e_u^c\right)\rangle_{\mathcal{Q}_h} - \langle \varepsilon \left[\boldsymbol{u}_h\right], \overline{\nabla}_{\overline{n}} \left(\Pi_h \left(\varphi e_u^c\right)\right)\rangle_{\mathcal{Q}_h}, \\
T_4 &= \langle \beta \cdot n u_h, \left(I - \Pi_h^{\mathcal{F}}\right)\left(\varphi e_u^c\right)\rangle_{\partial\mathcal{T}_h^i} + \langle \left(\beta_s - \beta \cdot n\right)\left[\boldsymbol{u}_h\right], \left(\Pi_h - \Pi_h^{\mathcal{F}}\right)\left(\varphi e_u^c\right)\rangle_{\partial\mathcal{T}_h^i}, \\
T_5 &= \left(\varepsilon \overline{\nabla} \left(I - \mathcal{I}_h^c\right) u_h, \overline{\nabla}\left(\varphi e_u^c\right)\right)_{\mathcal{T}_h} - \left(\beta \left(I - \mathcal{I}_h^c\right) u_h, \nabla \left(\varphi e_u^c\right)\right)_{\mathcal{T}_h}, \\
T_6 &= \langle \zeta^+ \beta \cdot n \left(u_h - \mathcal{I}_h^c u_h\right), \varphi e_u^c\rangle_{\partial\mathcal{E}_N} - \langle \zeta^+ \beta \cdot n \left[\boldsymbol{u}_h\right], \Pi_h^{\mathcal{F}}\left(\varphi e_u^c\right)\rangle_{\partial\mathcal{E}_N} \\
&\quad + \langle \left(\beta_s - \beta \cdot n\right)\left[\boldsymbol{u}_h\right], \left(\Pi_h - \Pi_h^{\mathcal{F}}\right)\left(\varphi e_u^c\right)\rangle_{\partial\mathcal{E} \setminus \Omega_T}.
\end{aligned}$$

Proof. Using the definition of the weighting function eq. (4.3), that $\varphi \geq \chi$, that $\beta \cdot \nabla \varphi = \partial_t \varphi = -e \exp(-t/T)$ and that $\overline{\nabla} \varphi = 0$, we have

$$(\varepsilon \overline{\nabla} e_u^c, \overline{\nabla} (\varphi e_u^c))_{\mathcal{T}_h} \geq \chi \varepsilon (\overline{\nabla} e_u^c, \overline{\nabla} e_u^c)_{\mathcal{T}_h} \quad \text{and} \quad -\tfrac{1}{2} ((\beta \cdot \nabla \varphi) e_u^c, e_u^c)_{\mathcal{T}_h} \geq \tfrac{1}{2} (e_u^c, e_u^c)_{\mathcal{T}_h},$$

so that

$$\chi \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \|\overline{\nabla} e_u^c\|_{\mathcal{K}}^2 + \tfrac{1}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \|e_u^c\|_{\mathcal{K}}^2 \leq (\varepsilon \overline{\nabla} e_u^c, \overline{\nabla} (\varphi e_u^c))_{\mathcal{T}_h} - \tfrac{1}{2} ((\beta \cdot \nabla \varphi) e_u^c, e_u^c)_{\mathcal{T}_h}. \tag{5.26}$$

For the right-hand side of eq. (5.26), using that $-\frac{1}{2}(\beta \cdot \nabla \varphi)(e_u^c)^2 = \varphi e_u^c \nabla \cdot (\beta e_u^c) - \frac{1}{2} \nabla \cdot (\beta \varphi (e_u^c)^2)$ because $\nabla \cdot \beta = 0$, integration by parts, that $\beta \cdot n$, e_u^c, and φ are single-valued on element boundaries, that e_u^c vanishes on $\partial \mathcal{E}_D$, the divergence theorem, and eq. (2.1a), we find:

$$\begin{aligned}
&(\varepsilon \overline{\nabla} e_u^c, \overline{\nabla} (\varphi e_u^c))_{\mathcal{T}_h} - \tfrac{1}{2} ((\beta \cdot \nabla \varphi) e_u^c, e_u^c)_{\mathcal{T}_h} \\
&= - (\varepsilon \overline{\nabla}^2 u, \varphi e_u^c)_{\mathcal{T}_h} + \langle \varepsilon \overline{\nabla}_{\overline{n}} u, \varphi e_u^c \rangle_{\mathcal{Q}_h \cap \partial \mathcal{E}_N} - (\varepsilon \overline{\nabla} \mathcal{I}_h^c u_h, \overline{\nabla} (\varphi e_u^c))_{\mathcal{T}_h} \\
&\qquad + (\nabla \cdot (\beta u), \varphi e_u^c)_{\mathcal{T}_h} - (\nabla \cdot (\beta \mathcal{I}_h^c u_h), \varphi e_u^c)_{\mathcal{T}_h} - \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\partial \mathcal{E}_N} \\
&= (f, \varphi e_u^c)_{\mathcal{T}_h} + \langle \varepsilon \overline{\nabla}_{\overline{n}} u, \varphi e_u^c \rangle_{\mathcal{Q}_h \cap \partial \mathcal{E}_N} - (\varepsilon \overline{\nabla} \mathcal{I}_h^c u_h, \overline{\nabla} (\varphi e_u^c))_{\mathcal{T}_h} \\
&\qquad + (\beta \mathcal{I}_h^c u_h, \nabla (\varphi e_u^c))_{\mathcal{T}_h} - \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}} - \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\partial \mathcal{E}_N} \\
&= (f, \varphi e_u^c)_{\mathcal{T}_h} - \widetilde{a}_h((\mathcal{I}_h^c u_h, \mathcal{I}_h^c u_h), (\varphi e_h^c, \varphi e_h^c)) + \langle \varepsilon \overline{\nabla}_{\overline{n}} u, \varphi e_u^c \rangle_{\mathcal{Q}_h \cap \partial \mathcal{E}_N} \\
&\qquad - \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}_N} - \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\partial \mathcal{E}_N}.
\end{aligned} \tag{5.27}$$

Using $\zeta^- \beta \cdot n = \frac{1}{2} (\beta \cdot n - |\beta \cdot n|)$, the last term above, excluding $\Omega_T \subset \partial \mathcal{E}_N$, is rewritten as follows

$$\begin{aligned}
&- \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} \\
&\qquad = - \langle \zeta^- \beta \cdot n u, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} - \tfrac{1}{2} \langle |\beta \cdot n| \, u, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} + \tfrac{1}{2} \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T}.
\end{aligned}$$

Therefore, using the Neumann boundary condition eq. (2.1b), the right-hand side of eq. (5.27) becomes

$$\begin{aligned}
&(f, \varphi e_u^c)_{\mathcal{T}_h} + \langle g, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} - \widetilde{a}_h((\mathcal{I}_h^c u_h, \mathcal{I}_h^c u_h), (\varphi e_h^c, \varphi e_h^c)) - \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}_N} \\
&\qquad - \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\Omega_T} - \tfrac{1}{2} \langle |\beta \cdot n| \, u, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} + \tfrac{1}{2} \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} \\
&= (f, \varphi e_u^c)_{\mathcal{T}_h} + \langle g, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} - \widetilde{a}_h((\mathcal{I}_h^c u_h, \mathcal{I}_h^c u_h), (\varphi e_h^c, \varphi e_h^c)) - \tfrac{1}{2} \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} \\
&\qquad - \langle \beta \cdot n \mathcal{I}_h^c u_h, \varphi e_u^c \rangle_{\Omega_T} - \tfrac{1}{2} \langle \beta \cdot n e_u^c, \varphi e_u^c \rangle_{\Omega_T} - \tfrac{1}{2} \langle |\beta \cdot n| \, u, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} \\
&= (f, \varphi e_u^c)_{\mathcal{T}_h} + \langle g, \varphi e_u^c \rangle_{\partial \mathcal{E}_N \setminus \Omega_T} - \widetilde{a}_h((\mathcal{I}_h^c u_h, \mathcal{I}_h^c u_h), (\varphi e_h^c, \varphi e_h^c)) - \mathfrak{B}_h,
\end{aligned} \tag{5.28}$$

where in the last step we collect remaining boundary terms in $\mathfrak{B}_h$:

$$\begin{aligned}\mathfrak{B}_h = \tfrac{1}{2}\langle \beta \cdot n\mathcal{I}_h^c u_h, \varphi e_u^c\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} &+ \tfrac{1}{2}\langle \beta \cdot n\mathcal{I}_h^c u_h, \varphi e_u^c\rangle_{\Omega_T} \\ &+ \tfrac{1}{2}\langle \beta \cdot nu, \varphi e_u^c\rangle_{\Omega_T} + \tfrac{1}{2}\langle |\beta \cdot n|\, u, \varphi e_u^c\rangle_{\partial\mathcal{E}_N\backslash\Omega_T}.\end{aligned}$$

Next, the HDG method eq. (2.9) with test functions $\boldsymbol{\Pi}_h(\varphi e_u^c, \varphi e_u^c) := (\Pi_h(\varphi e_u^c), \Pi_h^{\mathcal{F}}(\varphi e_u^c))$, and noting that $\Pi_h^{\mathcal{F}}(\varphi e_u^c) = 0$ on $\partial\mathcal{E}_D$, becomes:

$$\begin{aligned}0 = &-(f, \Pi_h(\varphi e_u^c))_{\mathcal{T}_h} - \langle g, \Pi_h^{\mathcal{F}}(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} \\ &+ \widetilde{a}_h(\boldsymbol{u}_h, \boldsymbol{\Pi}_h(\varphi e_u^c, \varphi e_u^c)) + k_h(\boldsymbol{u}_h, \boldsymbol{\Pi}_h(\varphi e_u^c, \varphi e_u^c)) + b_h(\lambda_h, \Pi_h^{\mathcal{F}}(\varphi e_u^c)).\end{aligned} \tag{5.29}$$

Adding eq. (5.29) to eq. (5.28), the right-hand side of eq. (5.26) becomes

$$\begin{aligned}(\varepsilon\overline{\nabla} e_u^c, \overline{\nabla}\,(\varphi e_u^c))_{\mathcal{T}_h} &- \tfrac{1}{2}\left((\beta\cdot\nabla\varphi)\, e_u^c, e_u^c\right)_{\mathcal{T}_h} \\ &= (f, (I - \Pi_h)(\varphi e_u^c))_{\mathcal{T}_h} + \langle g, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} \\ &\quad - \widetilde{a}_h(\boldsymbol{u}_h, (\boldsymbol{I} - \boldsymbol{\Pi}_h)(\varphi e_u^c, \varphi e_u^c)) - \mathfrak{R}_h \\ &\quad + k_h(\boldsymbol{u}_h, \boldsymbol{\Pi}_h(\varphi e_u^c, \varphi e_u^c)) + b_h(\lambda_h, \Pi_h^{\mathcal{F}}(\varphi e_u^c)) - \mathfrak{B}_h,\end{aligned} \tag{5.30}$$

where $\mathfrak{R}_h := \widetilde{a}_h((\mathcal{I}_h^c u_h, \mathcal{I}_h^c u_h), (\varphi e_h^c, \varphi e_h^c)) - \widetilde{a}_h(\boldsymbol{u}_h, (\varphi e_u^c, \varphi e_u^c))$. By definition of $\widetilde{a}_h$, the first three terms on the right-hand side of eq. (5.30) become

$$\begin{aligned}&(f, (I - \Pi_h)(\varphi e_u^c))_{\mathcal{T}_h} + \langle g, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} \\ &\quad - (\varepsilon\overline{\nabla} u_h, \overline{\nabla}\,(I - \Pi_h)\,(\varphi e_u^c))_{\mathcal{T}_h} - \langle \varepsilon\alpha h_K^{-1}\,[\boldsymbol{u}_h]\,, [(\boldsymbol{I} - \boldsymbol{\Pi}_h)\,(\varphi e_u^c, \varphi e_u^c)]\rangle_{\mathcal{Q}_h} \\ &\quad + (\beta u_h, \nabla(I - \Pi_h)\,(\varphi e_u^c))_{\mathcal{T}_h} - \langle \beta\cdot n\lambda_h + \beta_s\,[\boldsymbol{u}_h]\,, [(\boldsymbol{I} - \boldsymbol{\Pi}_h)\,(\varphi e_u^c, \varphi e_u^c)]\rangle_{\partial\mathcal{T}_h} \\ &= (R_h^{\mathcal{K}}, (I - \Pi_h)(\varphi e_u^c))_{\mathcal{T}_h} + \langle R_h^N, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} + \langle \varepsilon\alpha h_K^{-1}\,[\boldsymbol{u}_h]\,, (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\mathcal{Q}_h} \\ &\quad - \langle \varepsilon\overline{\nabla}_{\overline{n}} u_h, (I - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\mathcal{Q}_h^i} + \langle \varepsilon\overline{\nabla}_{\overline{n}} u_h, (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\mathcal{Q}_h} \\ &\quad + \langle (1 - \zeta^-)\,\beta\cdot nu_h, (I - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\backslash\Omega_T} + \langle \beta\cdot nu_h, (I - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h^i\cup\Omega_T} \\ &\quad + \langle \beta\cdot n\lambda_h + \beta_s\,[\boldsymbol{u}_h]\,, (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h} - \langle \beta\cdot nu_h, (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h}.\end{aligned} \tag{5.31}$$

For the k_h, b_h and $\mathfrak{R}_h$ terms on the right-hand side of eq. (5.30), we have

$$\begin{aligned}&k_h(\boldsymbol{u}_h, \boldsymbol{\Pi}_h\,(\varphi e_u^c, \varphi e_u^c)) = -\langle \varepsilon\,[\boldsymbol{u}_h]\,, \overline{\nabla}_{\overline{n}}\,(\Pi_h\,(\varphi e_u^c))\rangle_{\mathcal{Q}_h} - \langle \varepsilon\overline{\nabla}_{\overline{n}} u_h, (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\mathcal{Q}_h}, \\ &b_h(\lambda_h, \Pi_h^{\mathcal{F}}\,(\varphi e_u^c)) = \langle \zeta^+\beta\cdot n\lambda_h, \Pi_h^{\mathcal{F}}\,(\varphi e_u^c)\rangle_{\partial\mathcal{E}_N}, \\ &\mathfrak{R}_h = -\left(\varepsilon\overline{\nabla}\,(I - \mathcal{I}_h^c)\,u_h, \overline{\nabla}\,(\varphi e_u^c)\right)_{\mathcal{T}_h} + \left(\beta\,(I - \mathcal{I}_h^c)\,u_h, \nabla\,(\varphi e_u^c)\right)_{\mathcal{T}_h}.\end{aligned} \tag{5.32}$$

Using eqs. (5.30) to (5.32), and the definition of $-\mathfrak{B}_h$ we obtain eq. (5.25). □

Lemma 5.3. *Let φ be as in eq. (4.3) and assume that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. The following estimate holds:*

$$
\begin{aligned}
T \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 + \tfrac{1}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \| e_u^c \|_{\mathcal{K}}^2 + T \sum_{F \in \partial \mathcal{E}_N} \left\| \left| \beta \cdot n \right|^{1/2} e_u^c \right\|_F^2 & \\
\leq c \sum_{\mathcal{K} \in \mathcal{T}_h} \Big(T^2 \left(\eta_R^{\mathcal{K}} \right)^2 + T^2 \left(\eta_{J,1}^{\mathcal{K}} \right)^2 + T^2 \varepsilon^{-1} \left(\eta_{J,2,1}^{\mathcal{K}} \right)^2 + T^2 \varepsilon^{-1} \left(\eta_{J,2,2}^{\mathcal{K}} \right)^2 & \\
+ T^2 \varepsilon^{-1} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}} \right)^2 + T^2 \left(\eta_{J,3}^{\mathcal{K}} \right)^2 + T^2 \left(\eta_{BC,1}^{\mathcal{K}} \right)^2 + T \left(\eta_{BC,2}^{\mathcal{K}} \right)^2 \Big). &
\end{aligned}
\tag{5.33}
$$

Proof. We start by bounding the T_i, $i = 1, \dots, 6$, terms in lemma 5.2.

Bound for T_1. Using the Cauchy–Schwarz inequality, the local projection estimate eq. (3.18a), that $T + \chi \leq |\varphi| \leq eT + \chi$ and that $1 \leq |\partial_t \varphi| \leq e$, and Young's inequality, and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ and $T \geq 1$, we find that

$$
\begin{aligned}
T_1 \leq & \sum_{\mathcal{K} \in \mathcal{T}_h} \| R_h^{\mathcal{K}} \|_{\mathcal{K}} \, \| (I - \Pi_h) \left(\varphi e_u^c \right) \|_{\mathcal{K}} \\
\leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \lambda_{\mathcal{K}} \| R_h^{\mathcal{K}} \|_{\mathcal{K}} \left(eT + \chi \right) \left(h_K \varepsilon^{1/2} \| \partial_t e_u^c \|_{\mathcal{K}} + \varepsilon^{1/2} \| \overline{\nabla} e_u^c \|_{\mathcal{K}} + \| e_u^c \|_{\mathcal{K}} \right) \\
& \qquad + \sum_{\mathcal{K} \in \mathcal{T}_h} \lambda_{\mathcal{K}} \| R_h^{\mathcal{K}} \|_{\mathcal{K}} \left(e h_K \varepsilon^{1/2} + eT + \chi \right) \| e_u^c \|_{\mathcal{K}} \\
\leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \eta_R^{\mathcal{K}} \left(T + \chi \right) \left(h_K \varepsilon^{1/2} \| \partial_t e_u^c \|_{\mathcal{K}} + \varepsilon^{1/2} \| \overline{\nabla} e_u^c \|_{\mathcal{K}} + \| e_u^c \|_{\mathcal{K}} \right) \\
\leq & \tfrac{c}{\delta} \left(T + \chi \right)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_R^{\mathcal{K}} \right)^2 + \tfrac{c\delta}{2} \left(T + \chi \right) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \| \overline{\nabla} e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \tau_\varepsilon \| \partial_t e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \| e_u^c \|_{\mathcal{K}}^2 .
\end{aligned}
$$

Bound for T_2. We write T_2 as

$$
T_2 = \underbrace{-\langle \varepsilon \overline{\nabla}_{\overline{n}} u_h, \left(I - \Pi_h^{\mathcal{F}} \right) \left(\varphi e_u^c \right) \rangle_{\mathcal{Q}_h^i}}_{T_{21}} + \underbrace{\langle R_h^N, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c) \rangle_{\partial \mathcal{E}_N \setminus (\Omega_T \cup \Omega_0)}}_{T_{22}} + \underbrace{\langle R_h^N, (I - \Pi_h^{\mathcal{F}})(\varphi e_u^c) \rangle_{\Omega_0}}_{T_{23}} .
$$

For T_{21} we write element boundary integrals as facet integrals in which we use that $\left(I - \Pi_h^{\mathcal{F}} \right) \left(\varphi e_u^c \right)$ is continuous across a facet, use the Cauchy–Schwarz inequality, the triangle inequality, the local projection estimate eq. (3.18b), the projection bound eq. (3.12d), and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, $T \geq 1$, and Young's inequality to find

$$
\begin{aligned}
T_{21} \leq & \sum_{F \in \mathcal{F}^i_{\mathcal{Q},h}} \| \llbracket \varepsilon \overline{\nabla}_{\bar{n}} u_h \rrbracket \|_F \, \| (I - \Pi_h^{\mathcal{F}}) \, (\varphi e_u^c) \|_F \\
\leq & \sum_{F \in \mathcal{F}^i_{\mathcal{Q},h}} \| \llbracket \varepsilon \overline{\nabla}_{\bar{n}} u_h \rrbracket \|_F \left(\| (I - \Pi_h) \, (\varphi e_u^c) \|_F + \| (\Pi_h - \Pi_h^{\mathcal{F}}) \, (\varphi e_u^c) \|_F \right) \\
\leq & c \sum_{F \in \mathcal{F}^i_{\mathcal{Q},h}} h_K^{1/2} \varepsilon^{-1/2} \| \llbracket \varepsilon \overline{\nabla}_{\bar{n}} u_h \rrbracket \|_F \, (T + \chi) \left(h_K \varepsilon^{1/2} \| \partial_t e_u^c \|_{\mathcal{K}} + \varepsilon^{1/2} \| \overline{\nabla} e_u^c \|_{\mathcal{K}} \right) \\
& + c \sum_{F \in \mathcal{F}^i_{\mathcal{Q},h}} h_K^{1/2} \varepsilon^{-1/2} \| \llbracket \varepsilon \overline{\nabla}_{\bar{n}} u_h \rrbracket \|_F \, \| e_u^c \|_{\mathcal{K}} \\
\leq & \tfrac{c}{\delta} \, (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,1}^{\mathcal{K}} \right)^2 + \tfrac{c\delta}{2} \, (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \, \| \overline{\nabla} e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \tau_\varepsilon \, \| \partial_t e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \| e_u^c \|_{\mathcal{K}}^2 \, .
\end{aligned}
$$

Term T_{22} can be bounded similarly:

$$
T_{22} \leq \tfrac{c}{\delta} \, (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{BC,1}^{\mathcal{K}} \right)^2 + \tfrac{c\delta}{2} \, (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \, \| \overline{\nabla} e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \tau_\varepsilon \, \| \partial_t e_u^c \|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \| e_u^c \|_{\mathcal{K}}^2 \, .
$$

For T_{23} we have, by the Cauchy–Schwarz inequality, boundedness of $\Pi_h^{\mathcal{F}}$, that $|\beta \cdot n| = 1$ on Ω_0, and Young's inequality:

$$
\begin{aligned}
T_{23} \leq & \sum_{F_{\mathcal{R}} \subset \Omega_0} \| R_h^N \|_{F_{\mathcal{R}}} \, \| (I - \Pi_h^{\mathcal{F}}) \, (\varphi e_u^c) \|_{F_{\mathcal{R}}} \\
\leq & c \, (T + \chi) \sum_{F_{\mathcal{R}} \subset \Omega_0} \| R_h^N \|_{F_{\mathcal{R}}} \, \| |\tfrac{1}{2} \beta \cdot n|^{1/2} \, e_u^c \|_{F_{\mathcal{R}}} \\
\leq & \tfrac{c}{2\delta} \, (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{BC,2}^{\mathcal{K}} \right)^2 + \tfrac{c\delta}{2} \, (T + \chi) \sum_{F_{\mathcal{R}} \subset \Omega_0} \| |\tfrac{1}{2} \beta \cdot n|^{1/2} \, e_u^c \|_{F_{\mathcal{R}}}^2 \, .
\end{aligned}
$$

Combining the bounds for T_{21}, T_{22}, and T_{23}, we obtain:

$$
\begin{aligned}
T_2 \leq & \tfrac{c}{\delta} \, (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,1}^{\mathcal{K}} \right)^2 + \tfrac{c}{\delta} \, (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{BC,1}^{\mathcal{K}} \right)^2 + \tfrac{c}{2\delta} \, (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{BC,2}^{\mathcal{K}} \right)^2 \\
& + c\delta \, (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \, \| \overline{\nabla} e_u^c \|_{\mathcal{K}}^2 + c\delta \sum_{\mathcal{K} \in \mathcal{T}_h} \tau_\varepsilon \, \| \partial_t e_u^c \|_{\mathcal{K}}^2 + c\delta \sum_{\mathcal{K} \in \mathcal{T}_h} \| e_u^c \|_{\mathcal{K}}^2 \\
& + \tfrac{c\delta}{2} \, (T + \chi) \sum_{F_{\mathcal{R}} \subset \Omega_0} \| |\tfrac{1}{2} \beta \cdot n|^{1/2} \, e_u^c \|_{F_{\mathcal{R}}}^2 \, .
\end{aligned}
$$

Bound for T_3. We write T_3 as

$$T_3 = \underbrace{\langle \varepsilon \alpha h_K^{-1} \left[\boldsymbol{u}_h\right], (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\mathcal{Q}_h}}_{T_{31}} \underbrace{-\langle \varepsilon \left[\boldsymbol{u}_h\right], \overline{\nabla}_{\overline{n}}\,(\Pi_h\,(\varphi e_u^c))\rangle_{\mathcal{Q}_h}}_{T_{32}}.$$

Term T_{31} is bounded using the Cauchy–Schwarz inequality, the projection bound eq. (3.12d), and Young's inequality:

$$\begin{aligned}
T_{31} \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon h_K^{-1} \left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}} \left\|(\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_h^c)\right\|_{\mathcal{Q}_\mathcal{K}} \\
\leq & c\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2} \left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}} \varepsilon^{1/2} \left\|\overline{\nabla} e_u^c\right\|_\mathcal{K} \\
\leq & \tfrac{c}{2\delta}\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} (\eta_{J,2,1}^{\mathcal{K}})^2 + \tfrac{c\delta}{2}\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_\mathcal{K}^2.
\end{aligned}$$

For T_{32} we use the Cauchy–Schwarz inequality, the trace inequality eq. (3.7c), the first bound in eq. (3.16), and Young's inequality to find:

$$\begin{aligned}
T_{32} \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}} \left\|\overline{\nabla}\,(\Pi_h\,(\varphi e_u^c))\right\|_{\mathcal{Q}_\mathcal{K}} \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}} h_K^{-1/2} \left\|\overline{\nabla}\,(\Pi_h\,(\varphi e_u^c))\right\|_\mathcal{K} \\
\leq & c\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{1/2} h_K^{-1/2} \left\|\left[\boldsymbol{u}_h\right]\right\|_{\mathcal{Q}_\mathcal{K}} \varepsilon^{1/2} \left\|\overline{\nabla} e_u^c\right\|_\mathcal{K} \\
\leq & \tfrac{c}{2\delta}\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} (\eta_{J,2,1}^{\mathcal{K}})^2 + \tfrac{c\delta}{2}\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_\mathcal{K}^2.
\end{aligned}$$

Combining the bounds for T_{31} and T_{32} we obtain:

$$T_3 \leq \tfrac{c}{\delta}\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} (\eta_{J,2,1}^{\mathcal{K}})^2 + c\delta\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_\mathcal{K}^2.$$

Bound for T_4. Using that $\langle \beta \cdot n \lambda_h, (I - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h^i} = 0$ we start by writing T_4 as

$$T_4 = \langle \beta \cdot n \left[\boldsymbol{u}_h\right], (I - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h^i} + \langle (\beta_s - \beta \cdot n) \left[\boldsymbol{u}_h\right], (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c)\rangle_{\partial\mathcal{T}_h^i}.$$

Next, by a triangle inequality, using eq. (2.8), that $|\beta_s - \frac{1}{2}\beta \cdot n|^{1/2} \leq c$, and the Cauchy–

Schwarz inequality,

$$\begin{aligned}
T_4 \leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_\mathcal{K}} \left(\left\| (I-\Pi_h)\,(\varphi e_u^c) \right\|_{\mathcal{Q}_\mathcal{K}} + \left\| (\Pi_h - \Pi_h^\mathcal{F})\,(\varphi e_u^c) \right\|_{\mathcal{Q}_\mathcal{K}} \right) \\
&+ c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_\mathcal{K}} \left(\left\| (I-\Pi_h)\,(\varphi e_u^c) \right\|_{\mathcal{R}_\mathcal{K}} + \left\| (\Pi_h - \Pi_h^\mathcal{F})\,(\varphi e_u^c) \right\|_{\mathcal{R}_\mathcal{K}} \right) \\
=:& T_{41} + T_{42}.
\end{aligned}$$

For T_{41} we use the local projection estimate eq. (3.18b), the projection estimate eq. (3.12d), and Young's inequality to find:

$$\begin{aligned}
T_{41} \leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} h_K^{1/2}\varepsilon^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_\mathcal{K}} (T+\chi) \left(\tau_\varepsilon^{1/2} \left\| \partial_t e_u^c \right\|_\mathcal{K} + \varepsilon^{1/2} \left\| \overline{\nabla} e_u^c \right\|_\mathcal{K} \right) \\
&+ c \sum_{\mathcal{K}\in\mathcal{T}_h} h_K^{1/2}\varepsilon^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_\mathcal{K}} \left\| e_u^c \right\|_\mathcal{K} \\
\leq& \tfrac{c}{\delta} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^\mathcal{K}\right)^2 + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_\mathcal{K}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_\mathcal{K}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_\mathcal{K}^2 .
\end{aligned}$$

For T_{42} we use the local projection estimate eq. (3.18c), the projection estimate eq. (3.17) using that $\delta t_\mathcal{K} = \mathcal{O}(h_K^2)$, and Young's inequality,

$$\begin{aligned}
T_{42} \leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_\mathcal{K}} (T+\chi) \left(\tau_\varepsilon^{1/2} \left\| \partial_t e_u^c \right\|_\mathcal{K} + \varepsilon^{1/2} \left\| \overline{\nabla} e_u^c \right\| \right) \\
&+ c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_\mathcal{K}} \left\| e_u^c \right\|_\mathcal{K} \\
\leq& \tfrac{c}{\delta}\varepsilon^{-1} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^\mathcal{K}\right)^2 + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_\mathcal{K}^2 \\
&+ \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_\mathcal{K}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_\mathcal{K}^2 .
\end{aligned}$$

Combining the bounds for T_{41} and T_{42} we obtain:

$$\begin{aligned}
T_4 \leq& \tfrac{c}{\delta} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^\mathcal{K}\right)^2 + \tfrac{c}{\delta}\varepsilon^{-1} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^\mathcal{K}\right)^2 \\
&+ c\delta (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_\mathcal{K}^2 + c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_\mathcal{K}^2 + c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_\mathcal{K}^2 .
\end{aligned}$$

Bound for T_5. We write T_5 as

$$T_5 = \underbrace{\left(\varepsilon \overline{\nabla} \left(I - \mathcal{I}_h^c\right) u_h, \overline{\nabla} \left(\varphi e_u^c\right)\right)_{\mathcal{T}_h}}_{T_{51}} \underbrace{- \left(\beta \left(I - \mathcal{I}_h^c\right) u_h, \nabla \left(\varphi e_u^c\right)\right)_{\mathcal{T}_h}}_{T_{52}}.$$

For T_{51} we use the Cauchy–Schwarz inequality, the inverse inequality eq. (3.7b), the approximation estimate of the averaging operator eq. (3.24), that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, and Young's inequality to find

$$\begin{aligned}
T_{51} \leq & \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} \left(I - \mathcal{I}_h^c\right) u_h \right\|_{\mathcal{K}} \left\| \overline{\nabla} \left(\varphi e_u^c\right) \right\|_{\mathcal{K}} \\
\leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1} \left\| \left(I - \mathcal{I}_h^c\right) u_h \right\|_{\mathcal{K}} \left\| \overline{\nabla} \left(\varphi e_u^c\right) \right\|_{\mathcal{K}} \\
\leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1} \Big(\sum_{F \subset \check{\mathcal{Q}}_{\mathcal{K}}^i} h_K^{1/2} \left\| [\![u_h]\!] \right\|_F + \sum_{F \subset \check{\mathcal{R}}_{\mathcal{K}}^i} \delta t_{\mathcal{K}}^{1/2} \left\| [\![u_h]\!] \right\|_F \Big) \left\| \overline{\nabla} \left(\varphi e_u^c\right) \right\|_{\mathcal{K}} \\
\leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon h_K^{-1/2} \\
& \cdot \Big(\sum_{F \subset \check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}' \subset \omega_F} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}'}} + h_K^{1/2} \sum_{F \subset \check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}' \subset \omega_F} \left\| \left|\beta_s - \tfrac{1}{2}\beta \cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}'}} \Big) \left(T + \chi\right) \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}} \\
\leq & \tfrac{c}{2\delta} \left(T + \chi\right) \sum_{\mathcal{K} \in \mathcal{T}_h} \Big(\left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 \Big) + \tfrac{c\delta}{2} \left(T + \chi\right) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 .
\end{aligned}$$

For T_{52} we first write

$$T_{52} = \underbrace{- \left(u_h - \mathcal{I}_h^c u_h, \partial_t \left(\varphi e_u^c\right)\right)_{\mathcal{T}_h}}_{T_{521}} \underbrace{- \left(\overline{\beta}(u_h - \mathcal{I}_h^c u_h), \overline{\nabla} \left(\varphi e_u^c\right)\right)_{\mathcal{T}_h}}_{T_{522}}.$$

We bound T_{521} using the Cauchy–Schwarz inequality, the approximation estimate of the averaging operator eq. (3.24), and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. We further note that on $\mathcal{T}_h^x$ we have that $\tilde{\varepsilon}^{-1/2} h_K^{-1/2} < h_K^{1/4} \varepsilon^{-1}$ and on $\mathcal{T}_h^c$ we have that $\tilde{\varepsilon}^{-1/2} h_K^{-1/2} \leq \varepsilon^{-1} h_K^{1/2}$. Therefore,

$$\begin{aligned}
T_{521} \leq & c \sum_{\mathcal{K} \in \mathcal{T}_h} \Big(\sum_{F \subset \check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}' \subset \omega_F} \widetilde{\varepsilon}^{-1/2} h_K^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}'}} \Big) \left(\left(T + \chi\right) \tau_{\varepsilon}^{1/2} \left\| \partial_t e_u^c \right\|_{\mathcal{K}} + \left\| e_u^c \right\|_{\mathcal{K}} \right) \\
& + c \sum_{\mathcal{K} \in \mathcal{T}_h} \Big(\sum_{F \subset \check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}' \subset \omega_F} \widetilde{\varepsilon}^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta \cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}'}} \Big) \left(\left(T + \chi\right) \tau_{\varepsilon}^{1/2} \left\| \partial_t e_u^c \right\|_{\mathcal{K}} + \left\| e_u^c \right\|_{\mathcal{K}} \right) \\
\leq & \tfrac{c}{\delta} \varepsilon^{-1} \left(T + \chi\right)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta} \varepsilon^{-1} \left(T + \chi\right)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 \\
& + \tfrac{c}{\delta} \varepsilon^{-1} \left(T + \chi\right)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + 2c\delta \sum_{\mathcal{K} \in \mathcal{T}_h} \tau_{\varepsilon} \left\| \partial_t e_u^c \right\|_{\mathcal{K}}^2 + 2c\delta \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 .
\end{aligned}$$

For T_{522}, using the Cauchy–Schwarz inequality, the approximation estimate of the averaging operator eq. (3.24)

$$\begin{aligned} T_{522} \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\sum_{F\subset\check{\mathcal{Q}}^i_{\mathcal{K}}} \sum_{\mathcal{K}'\subset\omega_F} h_K^{1/2} \varepsilon^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}'}} + \sum_{F\subset\check{\mathcal{R}}^i_{\mathcal{K}}} \sum_{\mathcal{K}'\subset\omega_F} \varepsilon^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}'}} \Big) \\ & \cdot (T+\chi)\, \varepsilon^{1/2} \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}} \\ \leq & \tfrac{c}{2\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\left(\eta^{\mathcal{K}}_{J,2,2}\right)^2 + \varepsilon^{-1} \left(\eta^{\mathcal{K}}_{J,3,\mathcal{R}}\right)^2 \right) + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|^2_{\mathcal{K}}. \end{aligned}$$

$$T_{522} \leq \tfrac{c}{2\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\left(\eta^{\mathcal{K}}_{J,2,2}\right)^2 + \varepsilon^{-1} \left(\eta^{\mathcal{K}}_{J,3,\mathcal{R}}\right)^2 \right) + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|^2_{\mathcal{K}}.$$

Combining the bounds for T_{521} and T_{522} we find that

$$\begin{aligned} T_{52} \leq & \tfrac{c}{\delta} \varepsilon^{-1} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,2,1}\right)^2 \\ & + \tfrac{c}{\delta} (T+\chi) \left[\varepsilon^{-1} (T+\chi) + \tfrac{1}{2} \right] \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,2,2}\right)^2 \\ & + \tfrac{c}{\delta} \varepsilon^{-1} (T+\chi) \left[(T+\chi) + \tfrac{1}{2} \right] \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,3,\mathcal{R}}\right)^2 \\ & + 2c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|^2_{\mathcal{K}} + 2c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|^2_{\mathcal{K}} + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|^2_{\mathcal{K}}. \end{aligned}$$

Combining the bounds for T_{51} and T_{52}, we obtain:

$$\begin{aligned} T_5 \leq & \tfrac{c}{\delta} (T+\chi) \left[\varepsilon^{-1} (T+\chi) + \tfrac{1}{2} \right] \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,2,1}\right)^2 \\ & + \tfrac{c}{\delta} (T+\chi) \left[\varepsilon^{-1} (T+\chi) + \tfrac{1}{2} \right] \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,2,2}\right)^2 \\ & + \tfrac{c}{\delta} (T+\chi) \left\{ \varepsilon^{-1} \left[(T+\chi) + \tfrac{1}{2} \right] + \tfrac{1}{2} \right\} \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta^{\mathcal{K}}_{J,3,\mathcal{R}}\right)^2 \\ & + 2c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|^2_{\mathcal{K}} + 2c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|^2_{\mathcal{K}} + c\delta (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|^2_{\mathcal{K}}. \end{aligned}$$

Bound for T_6. We write T_6 as follows:

$$\begin{aligned}
T_6 =& \langle (\beta_s - \beta\cdot n)\, [\boldsymbol{u}_h]\,, (\Pi_h - \Pi_h^{\mathcal{F}})\, (\varphi e_u^c)\rangle_{\partial\mathcal{E}_D} \\
&+ \Big[\langle \zeta^+ \beta\cdot n\, (u_h - \mathcal{I}_h^c u_h)\,, \varphi e_u^c\rangle_{\Omega_T} - \langle \zeta^+ \beta\cdot n\, [\boldsymbol{u}_h]\,, \Pi_h^{\mathcal{F}}\, (\varphi e_u^c)\rangle_{\Omega_T} \Big] \\
&+ \big[\langle \zeta^+ \beta\cdot n\, (u_h - \mathcal{I}_h^c u_h)\,, \varphi e_u^c\rangle_{\partial\mathcal{E}_N\setminus\Omega_T} - \langle \zeta^+ \beta\cdot n\, [\boldsymbol{u}_h]\,, \Pi_h^{\mathcal{F}}\, (\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\setminus\Omega_T} \\
&\qquad + \langle (\beta_s - \beta\cdot n)\, [\boldsymbol{u}_h]\,, (\Pi_h - \Pi_h^{\mathcal{F}})\, (\varphi e_u^c)\rangle_{\partial\mathcal{E}_N\setminus(\Omega_T\cup\Omega_0)} \big] \\
&+ \langle (\beta_s - \beta\cdot n)\, [\boldsymbol{u}_h]\,, (\Pi_h - \Pi_h^{\mathcal{F}})\, (\varphi e_u^c)\rangle_{\Omega_0} \\
=&:\, T_{61} + T_{62} + T_{63} + T_{64}.
\end{aligned}$$

For T_{61}, we use the Cauchy–Schwarz inequality and the projection bound eq. (3.12d)

$$\begin{aligned}
T_{61} \leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \|[\boldsymbol{u}_h]\|_{\partial\mathcal{K}\cap\partial\mathcal{E}_D}\, h_K^{1/2}\, \|\overline{\nabla}\, (\varphi e_u^c)\|_{\mathcal{K}} \\
\leq& \tfrac{c}{2\delta}\, (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2}\, (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon\, \|\overline{\nabla} e_u^c\|_{\mathcal{K}}^2 .
\end{aligned}$$

For T_{62} we first write

$$T_{62} = \underbrace{\langle \zeta^+ \beta\cdot n\, (u_h - \mathcal{I}_h^c u_h)\,, \varphi e_u^c\rangle_{\Omega_T}}_{T_{621}} \underbrace{- \langle \zeta^+ \beta\cdot n\, [\boldsymbol{u}_h]\,, \Pi_h^{\mathcal{F}}\, (\varphi e_u^c)\rangle_{\Omega_T}.}_{T_{622}}$$

We bound T_{621}, using the Cauchy–Schwarz inequality, the trace inequality eq. (3.10b), the approximation estimate of the averaging operator eq. (3.24), that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, and Young's inequality:

$$\begin{aligned}
T_{621} \leq& \sum_{\mathcal{K}\in\mathcal{T}_h} \|u_h - \mathcal{I}_h^c u_h\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T}\, \|\varphi e_u^c\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T} \\
\leq& c \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\varepsilon^{-1/2} \sum_{F\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} \varepsilon^{1/2} h_K^{-1/2}\, \|[\boldsymbol{u}_h]\|_{\mathcal{Q}_{\mathcal{K}'}} + \sum_{F\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} \||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\, [\boldsymbol{u}_h]\|_{\partial\mathcal{K}'} \Big) \\
&\quad \cdot (T+\chi)\, \||\beta\cdot n|^{1/2}\, e_u^c\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T} \\
\leq& \tfrac{c}{2\delta} \varepsilon^{-1}\, (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{2\delta}\, (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 \\
&\quad + c\delta\, (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \||\beta\cdot n|^{1/2}\, e_u^c\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T}^2 .
\end{aligned}$$

Next, we bound T_{622} using the Cauchy–Schwarz inequality, the boundedness of the projection operator $\Pi_h^{\mathcal{F}}$, and Young's inequality:

$$\begin{aligned}
T_{622} \leq & \sum_{\mathcal{K}\in\mathcal{T}_h} \left\lVert \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\rVert_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T} (T+\chi) \left\lVert \left|\beta\cdot n\right|^{1/2} e_u^c \right\rVert_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T} \\
\leq & \tfrac{c}{2\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left\lVert \left|\beta\cdot n\right|^{1/2} e_u^c \right\rVert_{\mathcal{R}_{\mathcal{K}}\cap\Omega_T}^2 .
\end{aligned}$$

Combining the bounds for T_{621} and T_{622} we find that

$$T_{62} \leq \tfrac{c}{2\delta}\varepsilon^{-1} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + \tfrac{3c\delta}{2} (T+\chi) \sum_{F\subset\Omega_T} \left\lVert \left|\beta\cdot n\right|^{1/2} e_u^c \right\rVert_F^2 .$$

For T_{63} we write $T_{63} = T_{631} + T_{632} + T_{633}$ where

$$\begin{aligned}
T_{631} &:= \langle \zeta^+ \beta\cdot n \left(u_h - \mathcal{I}_h^c u_h\right), \varphi e_u^c \rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h}, \\
T_{632} &:= -\langle \zeta^+ \beta\cdot n \, [\boldsymbol{u}_h], \Pi_h^{\mathcal{F}} \left(\varphi e_u^c\right) \rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h}, \\
T_{633} &:= \langle \left(\beta_s - \beta\cdot n\right) [\boldsymbol{u}_h], \left(\Pi_h - \Pi_h^{\mathcal{F}}\right) \left(\varphi e_u^c\right) \rangle_{\partial\mathcal{E}_N\cap\mathcal{Q}_h} .
\end{aligned}$$

To bound T_{631}, we use the Cauchy–Schwarz inequality, the trace inequality eq. (3.7c), the approximation estimate of the averaging operator eq. (3.24), and Young's inequality:

$$\begin{aligned}
T_{631} \leq & \sum_{\mathcal{K}\in\mathcal{T}_h} c h_K^{-1/2} \left\lVert u_h - \mathcal{I}_h^c u_h \right\rVert_{\mathcal{K}} \left\lVert \left|\beta\cdot n\right|^{1/2} \varphi e_u^c \right\rVert_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} \\
\leq & \sum_{\mathcal{K}\in\mathcal{T}_h} c \Big(\sum_{F'\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_{F'}} \left\lVert [\boldsymbol{u}_h] \right\rVert_{\mathcal{Q}_{\mathcal{K}'}} + \sum_{F'\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_{F'}} \left\lVert [\boldsymbol{u}_h] \right\rVert_{\mathcal{R}_{\mathcal{K}'}} \Big) \left\lVert \left|\beta\cdot n\right|^{1/2} \varphi e_u^c \right\rVert_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} \\
\leq & \sum_{\mathcal{K}\in\mathcal{T}_h} c \Big(\varepsilon^{-1/2} \sum_{F'\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_{F'}} h_K^{-1/2} \varepsilon^{1/2} \left\lVert [\boldsymbol{u}_h] \right\rVert_{\mathcal{Q}_{\mathcal{K}'}} + \sum_{F'\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_{F'}} \left\lVert \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\rVert_{\mathcal{R}_{\mathcal{K}'}} \Big) \\
& \cdot \left\lVert \left|\beta\cdot n\right|^{1/2} \varphi e_u^c \right\rVert_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} \\
\leq & \tfrac{c}{2\delta}\varepsilon^{-1} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{2\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + c\delta (T+\chi) \sum_{F\in\partial\mathcal{E}_N} \left\lVert \left|\beta\cdot n\right|^{1/2} e_u^c \right\rVert_F^2 .
\end{aligned}$$

For T_{632}, using the Cauchy–Schwarz inequality, eq. (2.8), and boundedness of the projection operator $\Pi_h^{\mathcal{F}}$, we find

$$T_{632} \leq \sum_{\mathcal{K}\in\mathcal{T}_h} \left\lVert \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\rVert_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} (T+\chi) \beta_s^{1/2} \left\lVert e_u^c \right\rVert_{\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N} .$$

Using similar steps as used to bound M_{52} in the proof of theorem 5.3, we note that $\beta_s \left\|e_u^c\right\|_F^2 \le \left\||\beta \cdot n|^{1/2} e_u^c\right\|_F^2 + ch_K \left\|e_u^c\right\|_F^2$. Furthermore, using eq. (3.19) and Young's inequality we then find

$$
\begin{aligned}
T_{632} \le & c \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta \cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} (T + \chi) \left(\left\| |\beta \cdot n|^{1/2} e_u^c \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} + h_K^{1/2} \left\| e_u^c \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} \right) \\
\le & c \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta \cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} (T + \chi) \\
& \quad \cdot \left(\left\| |\beta \cdot n|^{1/2} e_u^c \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} + \left\| e_u^c \right\|_{\mathcal{K}} + h_K^{1/2} \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^{1/2} \left\| e_u^c \right\|_{\mathcal{K}}^{1/2} \right) \\
\le & c \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta \cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} (T + \chi) \\
& \quad \cdot \left(\left\| |\beta \cdot n|^{1/2} e_u^c \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} + \left\| e_u^c \right\|_{\mathcal{K}} + h_K \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}} \right) \\
\le & \tfrac{c}{\delta} (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,3,\mathcal{Q}}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} (T + \chi) \sum_{F \subset \partial\mathcal{E}_N} \left\| |\beta \cdot n|^{1/2} e_u^c \right\|_F^2 \\
& + \tfrac{c}{2\delta} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 .
\end{aligned}
$$

Next, we consider T_{633}. Using the Cauchy–Schwarz inequality, the projection estimate eq. (3.12d), and Young's inequality we find

$$
\begin{aligned}
T_{633} \le & c \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} \left\| \left(\Pi_h - \Pi_h^{\mathcal{F}}\right) (\varphi e_u^c) \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} \\
\le & c \sum_{\mathcal{K} \in \mathcal{T}_h} h_K^{1/2} \varepsilon^{-1/2} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N} \varepsilon^{1/2} (T + \chi) \left\| \overline{\nabla} (e_u^c) \right\|_{\mathcal{K}} \\
\le & \tfrac{c}{2\delta} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 .
\end{aligned}
$$

Combining the bounds for T_{631}, T_{632}, and T_{633} we find that

$$
\begin{aligned}
T_{63} \le & \tfrac{c}{2\delta} \varepsilon^{-1} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta} (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta} (T + \chi)^2 \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\eta_{J,3}^{\mathcal{K}}\right)^2 \\
& + \tfrac{c\delta}{2} \sum_{\mathcal{K} \in \mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 + c\delta (T + \chi) \sum_{\mathcal{K} \in \mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 + \tfrac{3c\delta}{2} (T + \chi) \sum_{F \subset \partial\mathcal{E}_N} \left\| |\beta \cdot n|^{1/2} e_u^c \right\|_F^2 .
\end{aligned}
$$

For T_{64}, we use the Cauchy–Schwarz inequality, the projection estimate eq. (3.17), and

Young's inequality to find

$$\begin{aligned}
T_{64} \leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_0} \left\| (\Pi_h - \Pi_h^{\mathcal{F}})\,(\varphi e_u^c) \right\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_0} \\
\leq & c \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon^{-1/2} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_0} (T+\chi)\, \tau_\varepsilon^{1/2} \left\| \partial_t e_u^c \right\|_{\mathcal{K}} \\
& + c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}}\cap\Omega_0} \left\| e_u^c \right\|_{\mathcal{K}} \\
\leq & \tfrac{c}{\delta}\varepsilon^{-1} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_{\mathcal{K}}^2 + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 .
\end{aligned}$$

Combining the bounds for T_{61}, T_{62}, T_{63} and T_{64}, we obtain

$$\begin{aligned}
T_6 \leq & \tfrac{c}{\delta}\varepsilon^{-1} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{3c}{2\delta} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 \\
& + \tfrac{2c}{\delta}\varepsilon^{-1} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,R}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta} (T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3}^{\mathcal{K}}\right)^2 \\
& + \tfrac{c\delta}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_{\mathcal{K}}^2 + c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 \\
& + \tfrac{3c\delta}{2} (T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 + 3c\delta\,(T+\chi) \sum_{F\subset\partial\mathcal{E}_N} \left\| \left|\beta\cdot n\right|^{1/2} e_u^c \right\|_F^2 .
\end{aligned}$$

With each of the terms T_i, $i = 1, \dots, 6$ bounded, we now bound $\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_{\mathcal{K}}^2$. By the triangle inequality and eq. (5.7),

$$\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u^c \right\|_{\mathcal{K}}^2 \leq c T^2 \varepsilon^{-1} \eta^2 + c \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t (u_h - \mathcal{I}_h^c u_h) \right\|_{\mathcal{K}}^2 .$$

For the second term on the right-hand side, using the inverse inequality eq. (3.10a), the approximation estimate of the averaging operator eq. (3.24), Hölder's inequality for sums, and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$,

$$\begin{aligned}
&\sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t(u_h - \mathcal{I}_h^c u_h)\right\|_{\mathcal{K}}^2 \\
&\qquad \le c \sum_{\mathcal{K}\in\mathcal{T}_h} \tilde{\varepsilon} \big(\sum_{F\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} h_K \delta t_{\mathcal{K}}^{-1} \left\|[\![u_h]\!]\right\|_F^2 + \sum_{F\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \left\|[\![u_h]\!]\right\|_F^2 \big) \\
&\qquad \le c \sum_{\mathcal{K}\in\mathcal{T}_h} \tilde{\varepsilon} \big(\sum_{F\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} h_K^{-1} \left\|[\boldsymbol{u}_h]\right\|_{\mathcal{Q}_{\mathcal{K}'}}^2 + \sum_{F\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\partial\mathcal{K}'} \big) \\
&\qquad \le c\varepsilon^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + c \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3}^{\mathcal{K}}\right)^2 .
\end{aligned} \tag{5.34}$$

Combining eqs. (5.25) and (5.34) with the bounds for T_1 to T_6, we have

$$\begin{aligned}
&\chi \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_{\mathcal{K}}^2 + \tfrac{1}{2} \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u^c\right\|_{\mathcal{K}}^2 + \tfrac{1}{2}\chi \sum_{F\subset\partial\mathcal{E}_N} \left\| \left|\beta\cdot n\right|^{1/2} e_u^c \right\|_F^2 \\
&\quad \le c\delta\,(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_{\mathcal{K}}^2 + c\delta \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u^c\right\|_{\mathcal{K}}^2 + c\delta\,(T+\chi) \sum_{F\subset\mathcal{E}_N} \left\| \left|\beta\cdot n\right|^{1/2} e_u^c \right\|_F^2 \\
&\qquad + \tfrac{c}{\delta}(T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_R^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta}(T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,1}^{\mathcal{K}}\right)^2 \\
&\qquad + \tfrac{c}{\delta}(T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{BC,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta}(T+\chi) \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{BC,2}^{\mathcal{K}}\right)^2 \\
&\qquad + c\left(\tfrac{1}{\delta}+\delta\right)(T+\chi)^2 \varepsilon^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \tfrac{c}{\delta}(T+\chi)^2 \varepsilon^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 \\
&\qquad + \tfrac{c}{\delta}(T+\chi)^2 \varepsilon^{-1} \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 + c\left(\tfrac{1}{\delta}+\delta\right)(T+\chi)^2 \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,3}^{\mathcal{K}}\right)^2 .
\end{aligned}$$

The result eq. (5.33) follows by choosing $\chi = T$ and $\delta = 1/(8c)$. □

We end this section by proving theorem 5.1.

Proof of theorem 5.1. Using the triangle inequality, Young's inequality, and eq. (2.8), we have

$$\begin{aligned}
\left|\!\left|\!\left| \boldsymbol{u} - \boldsymbol{u}_h \right|\!\right|\!\right|_{sT,h}^2 \le{}& c\,\big(T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\|\overline{\nabla} e_u^c\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\|e_u^c\right\|_{\mathcal{K}}^2 + T \sum_{F\subset\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} e_u^c \right\|_F^2 \big) \\
&+ \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\|\partial_t e_u\right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + T\left(\eta_{J,3}^{\mathcal{K}}\right)^2 \right) + I_1 + I_2 + I_3,
\end{aligned} \tag{5.35}$$

where

$$I_1 = cT \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla}(I - \mathcal{I}_h^c)u_h \right\|_{\mathcal{K}}^2, \qquad I_2 = c \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| (I - \mathcal{I}_h^c)u_h \right\|_{\mathcal{K}}^2,$$

$$I_3 = cT \sum_{F\subset\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} (I - \mathcal{I}_h^c)u_h \right\|_F^2 .$$

Using the inverse inequality eq. (3.7b), the approximation estimate of the averaging operator eq. (3.24), and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we bound I_1 as follows:

$$\begin{aligned} I_1 &\le cT \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \Big(\sum_{F\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} h_K^{-1} \left\| [\![u_h]\!] \right\|_F^2 + \sum_{F\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \left\| [\![u_h]\!] \right\|_F^2 \Big) \\ &\le cT \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\sum_{F\subset\check{\mathcal{Q}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} \varepsilon h_K^{-1} \left\| [\boldsymbol{u}_h] \right\|_{\mathcal{Q}_{\mathcal{K}'}}^2 + \sum_{F\subset\check{\mathcal{R}}_{\mathcal{K}}^i} \sum_{\mathcal{K}'\subset\omega_F} \left\| \left|\beta_s - \tfrac{1}{2}\beta\cdot n\right|^{1/2} [\boldsymbol{u}_h] \right\|_{\mathcal{R}_{\mathcal{K}'}}^2 \Big) \\ &\le cT \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 \Big) . \end{aligned}$$

Using the approximation estimate of the averaging operator eq. (3.24), then similar to the bound of I_1 we have:

$$I_2 \le c \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 \Big) .$$

Finally, using the trace inequalities eqs. (3.7c) and (3.10b), the approximation estimate of the averaging operator eq. (3.24), and that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we can bound I_3 as follows:

$$I_3 \le cT \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\varepsilon^{-1} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2 \Big) .$$

Combining the bounds for I_1, I_2, and I_3 with eq. (5.35) we find that

$$\begin{aligned} \left|\!\left|\!\left| \boldsymbol{u} - \boldsymbol{u}_h \right|\!\right|\!\right|_{sT,h}^2 \le c \Big(& T \sum_{\mathcal{K}\in\mathcal{T}_h} \varepsilon \left\| \overline{\nabla} e_u^c \right\|_{\mathcal{K}}^2 + \sum_{\mathcal{K}\in\mathcal{T}_h} \left\| e_u^c \right\|_{\mathcal{K}}^2 + T \sum_{F\subset\partial\mathcal{E}_N} \left\| \left|\tfrac{1}{2}\beta\cdot n\right|^{1/2} e_u^c \right\|_F^2 \Big) \\ & + \sum_{\mathcal{K}\in\mathcal{T}_h} \tau_\varepsilon \left\| \partial_t e_u \right\|_{\mathcal{K}}^2 + c \sum_{\mathcal{K}\in\mathcal{T}_h} \left(\eta_{J,2,2}^{\mathcal{K}}\right)^2 + cT \sum_{\mathcal{K}\in\mathcal{T}_h} \Big(\varepsilon^{-1} \left(\eta_{J,2,1}^{\mathcal{K}}\right)^2 + \left(\eta_{J,3}^{\mathcal{K}}\right)^2 \Big) . \end{aligned}$$

By lemma 5.3 this is further bound as:

$$\begin{aligned}\|\|\boldsymbol{u}-\boldsymbol{u}_h\|\|^2_{sT,h} \leq & c\sum_{\mathcal{K}\in\mathcal{T}_h}\Big(T^2\left(\eta_R^{\mathcal{K}}\right)^2+T^2\left(\eta_{J,1}^{\mathcal{K}}\right)^2+T^2\varepsilon^{-1}\left(\eta_{J,2,1}^{\mathcal{K}}\right)^2+T^2\varepsilon^{-1}\left(\eta_{J,2,2}^{\mathcal{K}}\right)^2\\ & \qquad +T^2\varepsilon^{-1}\left(\eta_{J,3,\mathcal{R}}^{\mathcal{K}}\right)^2+T^2\left(\eta_{J,3}^{\mathcal{K}}\right)^2+T^2\left(\eta_{BC,1}^{\mathcal{K}}\right)^2+T\left(\eta_{BC,2}^{\mathcal{K}}\right)^2\Big)\\ & +\sum_{\mathcal{K}\in\mathcal{T}_h}\tau_\varepsilon\left\|\partial_t e_u\right\|^2_{\mathcal{K}}\\ \leq & cT^2\varepsilon^{-1}\eta^2+\sum_{\mathcal{K}\in\mathcal{T}_h}\tau_\varepsilon\left\|\partial_t e_u\right\|^2_{\mathcal{K}}.\end{aligned}$$

We conclude eq. (5.2) using theorem 5.3. □

5.4 Local efficiency of the error estimator

In this section we prove theorem 5.2. Given any space-time element $\mathcal{K}$, we introduce element bubble function $\psi_{\mathcal{K}}=c_\theta\Pi_{i=1}^{2^{d+1}}\theta_{\mathcal{K},i}$, where $\theta_{\mathcal{K},i}$ denotes the linear Lagrangian basis polynomial associated with the i-th vertex of $\mathcal{K}$, and the constant factor c_θ is such that $\|\psi_{\mathcal{K}}\|_{L^\infty(\mathcal{K})}=1$. We observe that $(\psi_{\mathcal{K}})\,|_{\partial\mathcal{K}}=0$. Given any $v\in V_h$, the element bubble function satisfies the following estimates (see [100, Lemma 3.3] and [103, Lemma 3.6]):

$$\left\|\psi_{\mathcal{K}}v\right\|_{\mathcal{K}}\leq c\left\|v\right\|_{\mathcal{K}}, \qquad c\left\|v\right\|^2_{\mathcal{K}}\leq\left(v,\psi_{\mathcal{K}}v\right)_{\mathcal{K}}, \tag{5.36}$$

We also need facet bubble functions. For an element $\mathcal{K}$ and one of its $\mathcal{Q}$-facets $F\in\mathcal{Q}_{\mathcal{K}}$, we first transform to the reference domain and consider $\widehat{\mathcal{K}}=\Phi_{\mathcal{K}}^{-1}(\mathcal{K})$ and $\widehat{F}=\Phi_{\mathcal{K}}^{-1}(F)$. Without loss of generality, we let $\hat{x}_i$ denote the spatial coordinate such that $\hat{x}_i\equiv-1$ on $\widehat{F}$. Given any number $\kappa\in(0,1]$, we denote by Ψ_κ the mapping from $(\hat{t},\hat{x}_1,\dots,\hat{x}_i,\dots,\hat{x}_d)$ to $(\hat{t},\hat{x}_1,\dots,\kappa(\hat{x}_i+1)-1,\dots,\hat{x}_d)$ and we let $\widehat{\mathcal{K}}_\kappa:=\Psi_\kappa(\widehat{\mathcal{K}})$. We introduce the following facet bubble function

$$\hat{\psi}_{\mathcal{K},F,\kappa}=\begin{cases}c_{\theta,F}\Pi_{i=1}^{2^d}\hat{\theta}_{\mathcal{K},F,i,\kappa} & \text{on } \widehat{\mathcal{K}}_\kappa,\\ 0 & \text{on } \widehat{\mathcal{K}}\setminus\widehat{\mathcal{K}}_\kappa,\end{cases}$$

where $\hat{\theta}_{\mathcal{K},F,i,\kappa}$ denotes the linear Lagrangian basis polynomial associated with the i-th vertex of $\widehat{\mathcal{K}}_\kappa$ that is also on $\widehat{F}$. Similarly, the constant factor $c_{\theta,F}$ is such that $\|\hat{\psi}_{\mathcal{K},F,\kappa}\|_{L^\infty(\widehat{F})}=1$.

Furthermore, given any $\mu \in M_h$ and considering $\hat{\mu} = \mu \circ \Phi_{\mathcal{K}}$, we have the following estimates:

$$\begin{aligned} &\|\hat{\psi}_{\mathcal{K},F,\kappa}\hat{\mu}\|_{\widehat{F}} \leq c\,\|\hat{\mu}\|_{\widehat{F}}\,, && c\,\|\hat{\mu}\|^2_{\widehat{F}} \leq \langle \hat{\mu}, \hat{\psi}_{\mathcal{K},F,\kappa}\hat{\mu}\rangle_{\widehat{F}}, \\ &\|\hat{\psi}_{\mathcal{K},F,\kappa}\hat{\mu}\|_{\widehat{\mathcal{K}}} \leq c\kappa^{1/2}\,\|\hat{\mu}\|_{\widehat{F}}\,, && \|\widehat{\overline{\nabla}}\hat{\psi}_{\mathcal{K},F,\kappa}\hat{\mu}\|_{\widehat{\mathcal{K}}} \leq c\kappa^{-1/2}\,\|\hat{\mu}\|_{\widehat{F}}\,, \end{aligned} \tag{5.37}$$

where the first estimate is a result of $\|\hat{\psi}_{\mathcal{K},F,\kappa}\|_{L^\infty(\widehat{F})} = 1$ and the remaining estimates are shown in [101, Lemma 3.4].

We remark that the facet function μ in eq. (5.37) is continued to functions on elements using the continuation operator defined in [100]. We furthermore remark that eqs. (5.36) and (5.37) are proven in [100, 101, 103] on n-simplices and parallelepipeds, with $n \geq 2$. However, these inequalities also hold for our mesh due to the assumptions on $\phi_{\mathcal{K}}$ eqs. (2.2) and (2.4) resulting in a Jacobian bounded independent of h_K and $\delta t_{\mathcal{K}}$.

To define the facet bubble function on ω_F, we consider three cases in fig. 5.1:

- Case 1 The neighboring element of $\mathcal{K}$ across F, denoted by $\mathcal{K}_{nb}$, is at the same refinement element as $\mathcal{K}$.
- Case 2 The 2^d neighboring elements of $\mathcal{K}$ across F, denoted by $\mathcal{K}_{nb,i}$ with $i = 1, \dots, 2^d$, are finer.
- Case 3 The neighboring element of $\mathcal{K}$ with respect to facet F is coarser, which is denoted by $\mathcal{K}_{nb,0}$.

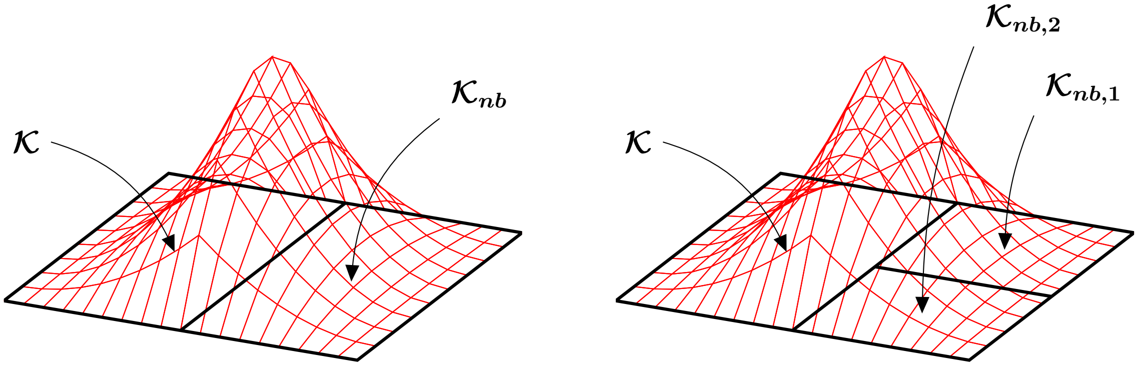
For Case 1, we let

$$\psi_{F,\kappa} := \begin{cases} \hat{\psi}_{\mathcal{K},F,\kappa} \circ \Phi_{\mathcal{K}}^{-1} & \text{on } \mathcal{K}, \\ \hat{\psi}_{\mathcal{K}_{nb},F,\kappa} \circ \Phi_{\mathcal{K}_{nb}}^{-1} & \text{on } \mathcal{K}_{nb}. \end{cases} \tag{5.38}$$

For Case 2, we consider the refinement of $\mathcal{K} := \cup_{i=1}^{2^d}\mathcal{K}_i$ such that F is refined to the set of $\{F_i\}_{i=1}^{2^d}$ where $F_i = \mathcal{Q}_{\mathcal{K}_i} \cap \mathcal{Q}_{\mathcal{K}_{nb,i}}$. We further denote by ω_{F_i} the union of $\mathcal{K}_i$ and $\mathcal{K}_{nb,i}$ and define a $\psi_{F,\kappa,i}$ on ω_{F_i} as in eq. (5.38) on each F_i.

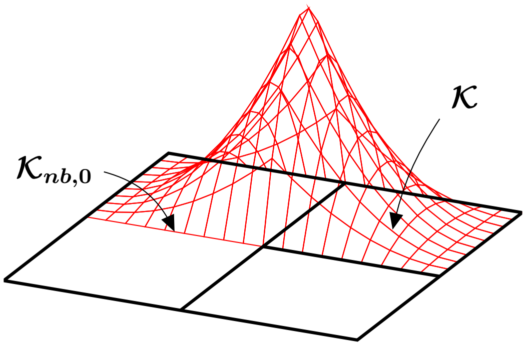
For Case 3, we consider the coarsest refinement of $\mathcal{K}_{nb,0}$ such that one of the refined elements $\mathcal{K}_{nb}$ has the property that $F = \mathcal{Q}_{\mathcal{K}} \cap \mathcal{Q}_{\mathcal{K}_{nb}}$. We denote the union of $\mathcal{K}$ and $\mathcal{K}_{nb}$ by $\omega_{F,*}$. Then, $\psi_{F,\kappa}$ is defined on $\omega_{F,*}$ as in eq. (5.38).

Applying the scaling arguments eqs. (3.1) and (3.2) to eq. (5.37), using the definition of $\psi_{F,\kappa}$ described above, choosing $\kappa = \tilde{\varepsilon}^{1/2}\varepsilon^{1/2}$, and dropping the subscript κ from $\psi_{F,\kappa}$, we



(a) Facet bubble when the neighboring element, $\mathcal{K}_{nb}$, of $\mathcal{K}$ is at the same refinement level

(b) Facet bubble when the neighboring element, $\mathcal{K}_{nb}$, of $\mathcal{K}$ is finer



(c) Facet bubble when the neighboring element, $\mathcal{K}_{nb}$, of $\mathcal{K}$ is coarser

Figure 5.1: Given an element $\mathcal{K}$ and a $\mathcal{Q}$-facet $F \in \mathcal{Q}_{\mathcal{K}}$, depending on the refinement level of $\mathcal{K}$'s neighboring element(s), we consider three different cases of the facet bubble function ψ_F for $F \in \mathcal{Q}_{\mathcal{K}}$.

obtain the following estimates:

$$\begin{aligned} &\left\|\psi_F\mu\right\|_F \leq c\left\|\mu\right\|_F, && c\left\|\mu\right\|_F^2 \leq \langle \mu, \psi_F\mu\rangle_F, \\ &\left\|\psi_F\mu\right\|_{\omega_F} \leq ch_K^{1/2}\tilde{\varepsilon}^{1/4}\varepsilon^{1/4}\left\|\mu\right\|_F, && \left\|\overline{\nabla}\psi_F\mu\right\|_{\omega_F} \leq ch_K^{-1/2}\tilde{\varepsilon}^{-1/4}\varepsilon^{-1/4}\left\|\mu\right\|_F. \end{aligned} \tag{5.39}$$

With the above bubble functions defined, we proceed with proving theorem 5.2.

Proof of theorem 5.2. Each term of $\eta^{\mathcal{K}}$ will be bound separately. However, let us first note that since $\eta^{\mathcal{K}}_{J,2}$ and $\eta^{\mathcal{K}}_{J,3}$ are part of $\left|\!\left|\!\left| \boldsymbol{u} - \boldsymbol{u}_h \right|\!\right|\!\right|_{sT,h,\mathcal{K}}$, these terms are trivially bounded.

Bound for $\eta^{\mathcal{K}}_R$. By the triangle inequality and Young's inequality,

$$\|R_h^{\mathcal{K}}\|_{\mathcal{K}}^2 \leq 2\,\|\Pi_h R_h^{\mathcal{K}}\|_{\mathcal{K}}^2 + 2\,\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}^2. \tag{5.40}$$

We bound the first term on the right-hand side. Using estimate eq. (5.36), with c_1 and c_2 the constants in the first and second inequalities of eq. (5.36), respectively, the Cauchy–Schwarz inequality, and Young's inequality with constant c_1, we note that

$$\tfrac{c_1}{2}\,\|\Pi_h R_h^{\mathcal{K}}\|_{\mathcal{K}}^2 \leq (R_h^{\mathcal{K}}, \psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} + \tfrac{c_2^2}{2c_1}\,\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}^2. \tag{5.41}$$

Combining eqs. (5.40) and (5.41), and using the boundedness of the projection Π_h so that $\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}^2 \leq c\,\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}$, we obtain

$$\lambda_{\mathcal{K}}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}^2 \leq c\lambda_{\mathcal{K}}\,(R_h^{\mathcal{K}}, \psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} + c\lambda_{\mathcal{K}}\,\|(I-\Pi_h)R_h^{\mathcal{K}}\|_{\mathcal{K}}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}. \tag{5.42}$$

To bound the first term on the right-hand side of eq. (5.42), we use the definition of $R_h^{\mathcal{K}}$, integrate by parts, and use that $\nabla\cdot\beta = 0$, to find for any $z \in H_0^1(\mathcal{K})$,

$$(R_h^{\mathcal{K}}, z)_{\mathcal{K}} = (\varepsilon^{1/2}\overline{\nabla}(u-u_h), \varepsilon^{1/2}\overline{\nabla}z)_{\mathcal{K}} + (\overline{\beta}\cdot\overline{\nabla}(u-u_h), z)_{\mathcal{K}} + (\partial_t(u-u_h), z)_{\mathcal{K}}. \tag{5.43}$$

Choosing $z = \psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}}$, we bound each term on the right-hand side of eq. (5.43) separately. Using the Cauchy–Schwarz inequality, the inequality eq. (3.7b), estimate eq. (5.36), and boundedness of the projection Π_h, we obtain:

$$(\varepsilon^{1/2}\overline{\nabla}(u-u_h), \varepsilon^{1/2}\overline{\nabla}(\psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}}))_{\mathcal{K}} \leq c\varepsilon^{1/2}\,\|\overline{\nabla}(u-u_h)\|_{\mathcal{K}}\,\varepsilon^{1/2}h_K^{-1}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}, \tag{5.44a}$$

$$(\overline{\beta}\cdot\overline{\nabla}(u-u_h), \psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} \leq c\varepsilon^{1/2}\,\|\overline{\nabla}(u-u_h)\|_{\mathcal{K}}\,\varepsilon^{-1/2}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}, \tag{5.44b}$$

$$(\partial_t(u-u_h), \psi_{\mathcal{K}}\Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} \leq c\tau_{\varepsilon}^{1/2}\,\|\partial_t(u-u_h)\|_{\mathcal{K}}\,\tau_{\varepsilon}^{-1/2}\,\|R_h^{\mathcal{K}}\|_{\mathcal{K}}. \tag{5.44c}$$

From eq. (5.43) with $z = \psi_{\mathcal{K}} \Pi_h R_h^{\mathcal{K}}$ and eq. (5.44) we therefore obtain:

$$(R_h^{\mathcal{K}}, \psi_{\mathcal{K}} \Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} \leq c \Bigg((\varepsilon^{1/2} h_K^{-1} + \varepsilon^{-1/2}) \varepsilon^{1/2} \left\| \overline{\nabla} (u - u_h) \right\|_{\mathcal{K}} \\ + \tau_{\varepsilon}^{-1/2} \tau_{\varepsilon}^{1/2} \left\| \partial_t (u - u_h) \right\|_{\mathcal{K}} \Bigg) \left\| R_h^{\mathcal{K}} \right\|_{\mathcal{K}}. \tag{5.45}$$

Using that $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ we note that $\lambda_{\mathcal{K}} (\varepsilon^{1/2} h_K^{-1} + \varepsilon^{-1/2}) < c \tilde{\varepsilon}^{-1/2} \varepsilon^{-1/2}$. Therefore, multiplying both sides of eq. (5.45) by $\lambda_{\mathcal{K}}$, we find

$$\lambda_{\mathcal{K}} \, (R_h^{\mathcal{K}}, \psi_{\mathcal{K}} \Pi_h R_h^{\mathcal{K}})_{\mathcal{K}} \leq c \varepsilon^{-1/2} \tilde{\varepsilon}^{-1/2} \left(\varepsilon^{1/2} \left\| \overline{\nabla} (u - u_h) \right\|_{\mathcal{K}} + \tau_{\varepsilon}^{1/2} \left\| \partial_t (u - u_h) \right\|_{\mathcal{K}} \right) \left\| R_h^{\mathcal{K}} \right\|_{\mathcal{K}}. \tag{5.46}$$

Combining eqs. (5.42) and (5.46), and using the definitions of $|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}}$, $\eta_R^{\mathcal{K}}$ and $\operatorname{osc}_h^{\mathcal{K}}$,

$$\eta_R^{\mathcal{K}} \leq c \varepsilon^{-1/2} \tilde{\varepsilon}^{-1/2} |||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + c \operatorname{osc}_h^{\mathcal{K}}.$$

Bound for $\eta_{J,1}^{\mathcal{K}}$. Let F be a facet such that $F \subset \mathcal{Q}_{\mathcal{K}} \setminus \partial \mathcal{E}$. To bound $\eta_{J,1}^{\mathcal{K}}$ we consider separately Case 1, Case 2, and Case 3.

Case 1. For any $F \subset \mathcal{Q}_{\mathcal{K}}$ and $z \in H_0^1(\omega_F)$, we have

$$\begin{aligned} \langle \varepsilon [\![\overline{\nabla}_{\overline{n}} u_h]\!], z \rangle_F = & - \left(\varepsilon^{1/2} \overline{\nabla} (u - u_h), \varepsilon^{1/2} \overline{\nabla} z \right)_{\omega_F} - \left(\partial_t (u - u_h), z \right)_{\omega_F} \\ & - \left(\overline{\beta} \cdot \overline{\nabla} (u - u_h), z \right)_{\omega_F} + \left(R_h^{\mathcal{K}}, z \right)_{\omega_F}. \end{aligned} \tag{5.47}$$

Choosing $z = \psi_F \varepsilon [\![\overline{\nabla}_{\overline{n}} u_h]\!]$, using eq. (5.39), and the Cauchy–Schwarz inequality, we obtain

$$h_K^{1/2} \varepsilon^{1/2} \left\| [\![\overline{\nabla}_{\overline{n}} u_h]\!] \right\|_F \leq c \varepsilon^{-1/4} \tilde{\varepsilon}^{-1/4} \varepsilon^{1/2} \left\| \overline{\nabla} (u - u_h) \right\|_{\omega_F} \\ + h_K \varepsilon^{-1/2} \varepsilon^{1/4} \tilde{\varepsilon}^{1/4} \left(\tau_{\varepsilon}^{-1/2} \tau_{\varepsilon}^{1/2} \left\| \partial_t (u - u_h) \right\|_{\omega_F} + \varepsilon^{-1/2} \varepsilon^{1/2} \left\| \overline{\nabla} (u - u_h) \right\|_{\omega_F} + \lambda_{\mathcal{K}}^{-1} \lambda_{\mathcal{K}} \left\| R_h^{\mathcal{K}} \right\|_{\omega_F} \right). \tag{5.48}$$

Using $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, $h_K \varepsilon^{-1/2} \tilde{\varepsilon}^{1/2} \leq 1$ and $\varepsilon^{-1/4} \tilde{\varepsilon}^{1/4} \max \{ h_K, \varepsilon^{1/2} \} \leq 1$, we find

$$h_K^{1/2} \varepsilon^{1/2} \left\| [\![\overline{\nabla}_{\overline{n}} u_h]\!] \right\|_F \leq c \sum_{\mathcal{K} \subset \omega_F} \varepsilon^{-1/4} \tilde{\varepsilon}^{-1/4} |||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + \lambda_{\mathcal{K}} \left\| R_h^{\mathcal{K}} \right\|_{\omega_F}.$$

Case 2. Identical steps as in Case 1 gives

$$h_K^{1/2} \varepsilon^{1/2} \left\| [\![\overline{\nabla}_{\overline{n}} u_h]\!] \right\|_{F_i} \leq c \varepsilon^{-1/4} \tilde{\varepsilon}^{-1/4} \varepsilon^{1/2} \left\| \overline{\nabla} (u - u_h) \right\|_{\omega_{F_i}} + c \varepsilon^{1/4} \tilde{\varepsilon}^{1/4} \tau_{\varepsilon}^{1/2} \left\| \partial_t (u - u_h) \right\|_{\omega_{F_i}} + \lambda_{\mathcal{K}} \left\| R_h^{\mathcal{K}} \right\|_{\omega_{F_i}}.$$

Summing over all F_i's,

$$h_K^{1/2}\varepsilon^{1/2}\left\|[\![\overline{\nabla}_{\overline{n}}u_h]\!]\right\|_F \leq c\sum_{\mathcal{K}\subset\omega_F}\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}|\!|\!|\boldsymbol{u}-\boldsymbol{u}_h|\!|\!|_{sT,h,\mathcal{K}} + \lambda_{\mathcal{K}}\left\|R_h^{\mathcal{K}}\right\|_{\omega_F}.$$

Case 3. Identical steps as in Case 1 gives

$$h_K^{1/2}\varepsilon^{1/2}\left\|[\![\overline{\nabla}_{\overline{n}}u_h]\!]\right\|_F \leq c\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}\varepsilon^{1/2}\left\|\overline{\nabla}(u-u_h)\right\|_{\omega_{F,*}}+c\varepsilon^{1/4}\tilde{\varepsilon}^{1/4}\tau_\varepsilon^{1/2}\left\|\partial_t(u-u_h)\right\|_{\omega_{F,*}}+\lambda_{\mathcal{K}}\left\|R_h^{\mathcal{K}}\right\|_{\omega_{F,*}}.$$

Since $\omega_{F,*}\subset\omega_F$, we then find

$$h_K^{1/2}\varepsilon^{1/2}\left\|[\![\overline{\nabla}_{\overline{n}}u_h]\!]\right\|_F \leq c\sum_{\mathcal{K}\subset\omega_F}\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}|\!|\!|\boldsymbol{u}-\boldsymbol{u}_h|\!|\!|_{sT,h,\mathcal{K}} + \lambda_{\mathcal{K}}\left\|R_h^{\mathcal{K}}\right\|_{\omega_F}.$$

For each of the three cases, summing over all facets $F\subset\mathcal{Q}_{\mathcal{K}}\setminus\partial\mathcal{E}$, and using the definitions of $\eta_{J,1}^{\mathcal{K}}$ and $\eta_R^{\mathcal{K}}$, we find

$$\eta_{J,1}^{\mathcal{K}} \leq \sum_{F\in\mathcal{Q}_{\mathcal{K}}\setminus\partial\mathcal{E}}\sum_{\mathcal{K}\subset\omega_F}\left[c\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}|\!|\!|\boldsymbol{u}-\boldsymbol{u}_h|\!|\!|_{sT,h,\mathcal{K}}+\eta_R^{\mathcal{K}}\right].$$

Bound for $\eta_{BC,1}^{\mathcal{K}}$. To bound $\eta_{BC,1}^{\mathcal{K}}$, let F be a facet such that $F\subset\mathcal{Q}_{\mathcal{K}}\cap\partial\mathcal{E}_N$. By the triangle inequality and Young's inequality,

$$\left\|R_h^N\right\|_F^2 \leq 2\left\|\Pi_h^{\mathcal{F}}R_h^N\right\|_F^2+2\left\|(I-\Pi_h^{\mathcal{F}})R_h^N\right\|_F^2. \tag{5.49}$$

We bound the first term on the right-hand side. Using estimate eq. (5.39), with c_1 and c_2 the constants in the first and second inequalities of eq. (5.39), respectively, the Cauchy–Schwarz inequality, and Young's inequality with constant c_1, we note that

$$\tfrac{c_1}{2}\left\|\Pi_h^{\mathcal{F}}R_h^N\right\|_F^2 \leq \langle R_h^N,\psi_F\Pi_h^{\mathcal{F}}R_h^N\rangle_F+\tfrac{c_2^2}{2c_1}\left\|(I-\Pi_h^{\mathcal{F}})R_h^N\right\|_F^2. \tag{5.50}$$

Combining eqs. (5.49) and (5.50), and using the boundedness of the projection $\Pi_h^{\mathcal{F}}$ so that $\left\|(I-\Pi_h^{\mathcal{F}})R_h^N\right\|_F^2 \leq c\left\|(I-\Pi_h^{\mathcal{F}})R_h^N\right\|_F\left\|R_h^N\right\|_F$, we obtain

$$\left\|R_h^N\right\|_F^2 \leq c\langle R_h^N,\psi_F\Pi_h^{\mathcal{F}}R_h^N\rangle_F+c\left\|(I-\Pi_h^{\mathcal{F}})R_h^N\right\|_F\left\|R_h^N\right\|_F. \tag{5.51}$$

Let $z\in H^1(\omega_F)$ be such that $z|_{\partial\omega_F\setminus F}=0$. Note that $\omega_F=\mathcal{K}$. Similar to eq. (5.47), we have:

$$(R_h^{\mathcal{K}},z)_{\mathcal{K}}=(\varepsilon^{1/2}\overline{\nabla}(u-u_h),\varepsilon^{1/2}\overline{\nabla}z)_{\mathcal{K}}+(\partial_t(u-u_h),z)_{\mathcal{K}}+(\overline{\beta}\cdot\overline{\nabla}(u-u_h),z)_{\mathcal{K}}-\langle\varepsilon\overline{\nabla}_{\bar{n}}(u-u_h),z\rangle_F.$$

The last term on the right-hand side can be rewritten using eq. (2.1b) resulting in

$$\begin{aligned}\langle R_h^N, z\rangle_F =& (\varepsilon^{1/2}\overline{\nabla}(u-u_h), \varepsilon^{1/2}\overline{\nabla}z)_{\mathcal{K}} + (\partial_t(u-u_h), z)_{\mathcal{K}} + (\overline{\beta}\cdot\overline{\nabla}(u-u_h), z)_{\mathcal{K}} \\ &- (R_h^{\mathcal{K}}, z)_{\mathcal{K}} - \langle \zeta^-(u-\mu_h)\beta\cdot n, z\rangle_F + \langle \zeta^-\left[\boldsymbol{u}_h\right]\beta\cdot n, z\rangle_F.\end{aligned} \tag{5.52}$$

Choosing $z = \psi_F \Pi_h^{\mathcal{F}} R_h^N$ in eq. (5.52) and using eqs. (2.8) and (5.39) and boundedness of $\Pi_h^{\mathcal{F}}$

$$\begin{aligned} c h_K^{1/2}\varepsilon^{-1/2}\langle R_h^N, \psi_F \Pi_h^{\mathcal{F}} R_h^N\rangle_F \le& \Big(\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}\varepsilon^{1/2}\left\|\overline{\nabla}(u-u_h)\right\|_{\mathcal{K}} \\ &+ c h_K \varepsilon^{-1/2}\varepsilon^{1/4}\tilde{\varepsilon}^{1/4}\left(\|\partial_t(u-u_h)\|_{\mathcal{K}} + \|\overline{\nabla}(u-u_h)\|_{\mathcal{K}} + \|R_h^{\mathcal{K}}\|_{\mathcal{K}}\right) \\ &+ c h_K^{1/2}\varepsilon^{-1/2}\left(\left\||\tfrac{1}{2}\beta\cdot n|^{1/2}(u-\mu_h)\right\|_F + \left\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\left[\boldsymbol{u}_h\right]\right\|_F\right)\Big)\|R_h^{\mathcal{F}}\|_F\,. \end{aligned}$$

The first two terms on the right-hand side are identical to the right-hand side in eq. (5.48) and so can be bounded similarly:

$$\begin{aligned} c h_K^{1/2}\varepsilon^{-1/2}\langle R_h^N, \psi_F \Pi_h^{\mathcal{F}} R_h^N\rangle_F \le& \Big(\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + \lambda_{\mathcal{K}}\left\|R_h^{\mathcal{K}}\right\|_{\mathcal{K}} \\ &+ c h_K^{1/2}\varepsilon^{-1/2}\left(\left\||\tfrac{1}{2}\beta\cdot n|^{1/2}(u-\mu_h)\right\|_F + \left\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\left[\boldsymbol{u}_h\right]\right\|_F\right)\Big)\|R_h^{\mathcal{F}}\|_F\,. \end{aligned} \tag{5.53}$$

At this point, let us note that $h_K^{1/2}\varepsilon^{-1/2} \le \tilde{\varepsilon}^{-1/2}$ for $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. Therefore, for the last term on the right-hand side of eq. (5.53) we have

$$c h_K^{1/2}\varepsilon^{-1/2}\left(\left\||\tfrac{1}{2}\beta\cdot n|^{1/2}(u-\mu_h)\right\|_F + \left\||\beta_s - \tfrac{1}{2}\beta\cdot n|^{1/2}\left[\boldsymbol{u}_h\right]\right\|_F\right) \le c\tilde{\varepsilon}^{-1/2}|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h,\mathcal{K}}. \tag{5.54}$$

Combining eqs. (5.51), (5.53) and (5.54), summing over all $F \in \mathcal{Q}_{\mathcal{K}} \cap \partial\mathcal{E}_N$, using that $\tilde{\varepsilon}^{-1/2} \le \varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}$, and the definitions of $\eta_{BC,1}^{\mathcal{K}}$ and $\eta_R^{\mathcal{K}}$, we find that

$$\eta_{BC,1}^{\mathcal{K}} \le c\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4}|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h,\mathcal{K}} + c\eta_R^{\mathcal{K}} + c\,\mathrm{osc}_h^N.$$

Bound for $\eta_{BC,2}^{\mathcal{K}}$. Let F be a facet such that $F \subset \mathcal{R}_{\mathcal{K}} \cap \Omega_0$. By eq. (2.1b) we have that $g = -u\beta\cdot n = u$. Therefore,

$$\eta_{BC,2}^{\mathcal{K}} = \|u-u_h\|_F \le \|u-\mu_h\|_F + \|\left[\boldsymbol{u}_h\right]\|_F \le c|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h,\mathcal{K}}.$$

Bound for $\eta_{J,2,2}^{\mathcal{K}}$. Let F be a facet such that $F \subset \mathcal{Q}_{\mathcal{K}} \backslash \partial \mathcal{E}$. Using again that $h_K^{1/2} \varepsilon^{-1/2} \leq \tilde{\varepsilon}^{-1/2}$ for $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we have

$$\eta_{J,2,2}^{\mathcal{K}} \leq \varepsilon^{-1/4} \tilde{\varepsilon}^{-3/4} |||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h,\mathcal{K}}.$$

Combining the bounds for $\eta_R^{\mathcal{K}}$, $\eta_{J,1}^{\mathcal{K}}$, $\eta_{BC,1}^{\mathcal{K}}$, $\eta_{BC,2}^{\mathcal{K}}$ and $\eta_{J,2,2}^{\mathcal{K}}$, and since $\varepsilon^{-1/4}\tilde{\varepsilon}^{-1/4} \leq \varepsilon^{-1/2}\tilde{\varepsilon}^{-1/2}$, we conclude eq. (5.3). $\square$

5.5 Numerical examples

In this section, we solve the space-time HDG method eq. (2.9) with AMR using the a posteriori error estimator $\eta^{\mathcal{K}}$ introduced in eq. (5.1). The implementation uses the finite element library deal.II [8, 9] on unstructured hexahedral space-time meshes with p4est [16] to obtain distributed mesh information. Furthermore, in our implementation we choose the penalty parameter $\alpha = 8p_s^2$ (see, for example, [90]). The linear system is solved all-at-once using the Multifrontal Massively Parallel Solver (MUMPS) [3, 4]. In each refinement cycle, the local error estimate $\eta^{\mathcal{K}}$ is computed for all $\mathcal{K} \in \mathcal{T}_h$ and then ordered according to the magnitude of $\eta^{\mathcal{K}}$. The top 25% of elements are marked for refinement and the bottom 10% of elements are marked for coarsening. The test cases in this section are implemented for both $\delta t_{\mathcal{K}} = h_K$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. In each example we will also investigate the efficiency index, which is defined as $\eta / |||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h}$.

Remark 3. *By theorem 5.1, theorem 5.2 and remark 2 we expect the efficiency index to be bounded below by $\mathcal{O}(\varepsilon^{1/2})$ and above by $\mathcal{O}(\varepsilon^{-1})$ in the pre-asymptotic regime and above by $\mathcal{O}(\varepsilon^{-1/2})$ in the asymptotic regime.*

5.5.1 A rotating Gaussian pulse test

This test case involves a Gaussian pulse on the spatial domain $\Omega = (-0.5, 0.5)^2$ and we simulate its rotation in the time interval $I = (0, 1]$. We set $\beta = (1, -4x_2, 4x_1)^{\intercal}$ and $f = 0$. Initial and boundary conditions are then chosen such that the exact solution to the problem is given by

$$u(t, x_1, x_2) = \tfrac{\sigma^2}{\sigma^2 + 2\varepsilon t} \exp\left(-\tfrac{(\widetilde{x}_1 - x_{1c})^2 + (\widetilde{x}_2 - x_{2c})^2}{2\sigma^2 + 4\varepsilon t}\right),$$

where $\widetilde{x}_1 := x_1 \cos(4t) + x_2 \sin(4t)$ and $\widetilde{x}_2 := -x_1 \sin(4t) + x_2 \cos(4t)$. We choose $\sigma = 0.1$ and $(x_{1c}, x_{2c}) = (-0.2, 0.1)$. To demonstrate the motion of the pulse and the adaptive mesh

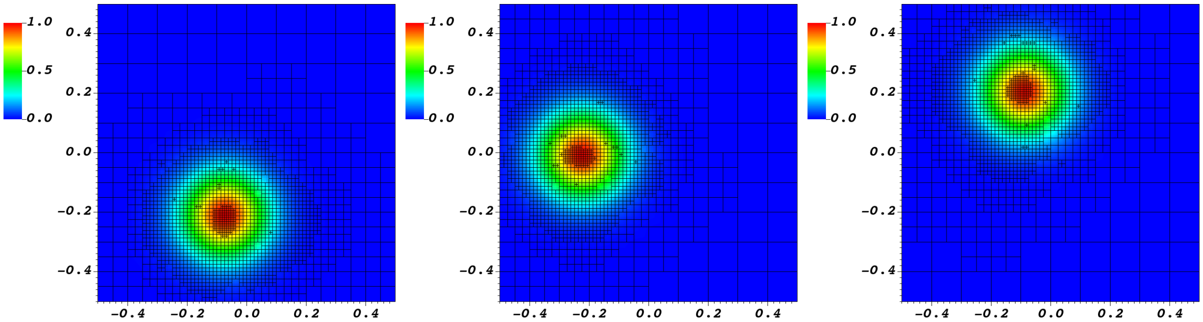


Figure 5.2: The spatial mesh and the rotating pulse. The solution is shown for $\varepsilon = 10^{-4}$. Plots correspond to time levels $t = 0.2, 0.5, 0.8$ from left to right.

refinement, we plot the spatial meshes and the solutions at $t = 0.2, 0.5, 0.8$ for $\varepsilon = 10^{-4}$ in fig. 5.2.

We perform three convergence tests with $\varepsilon = 10^{-3}$, 10^{-4}, and 10^{-5}. In fig. 5.3, for $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we present the convergence histories of the error estimator η, the true error $|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h}$ when using AMR, and the true error $|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h}$ when using uniform refinement. Additionally, we compute the efficiency index after each refinement cycle and plot its history. All tests are implemented with $p_t = p_s = 1$.

For both $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ we observe on fig. 5.3 that solutions on adaptively refined meshes are slightly more accurate than their counterparts on uniformly refined meshes although there is not too much advantage of using AMR for this smooth test case. Both solutions exhibit convergence rate $\mathcal{O}(N^{-1/2})$ which is optimal in the pre-asymptotic regime (see remark 1). These results correspond to what we expect from reliability and efficiency of the estimator proven in theorem 5.1 and theorem 5.2. Nonrobustness of the error estimator η is observed with the efficiency index being of order $\varepsilon^{-1/2}$. This lies within the interval commented on in remark 3.

5.5.2 A boundary layer test

We now consider problem eq. (2.1) in which the solution exhibits boundary layers. The problem is set up on the spatial domain $\Omega = (0,1)^2$ and the time interval $I = (0,1]$ with $\beta = (1,1,1)^\intercal$. The initial and boundary conditions and the source term are chosen such that the exact solution is given by

$$u(t,x_1,x_2) = (1-\exp(-t))\left(\tfrac{\exp((x_1-1)/\varepsilon)-1}{\exp(-1/\varepsilon)-1} + x_1 - 1\right)\left(\tfrac{\exp((x_2-1)/\varepsilon)-1}{\exp(-1/\varepsilon)-1} + x_2 - 1\right).$$

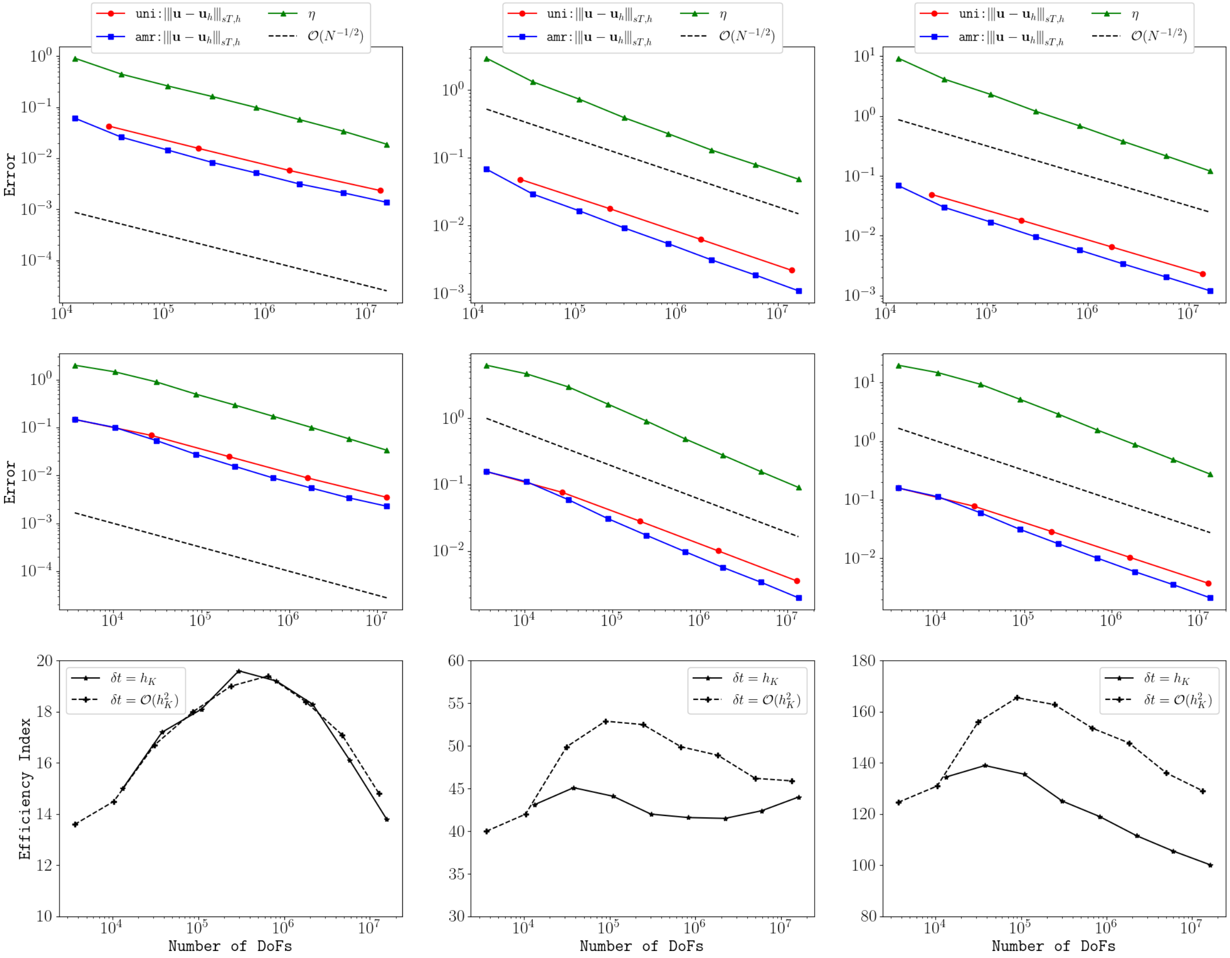
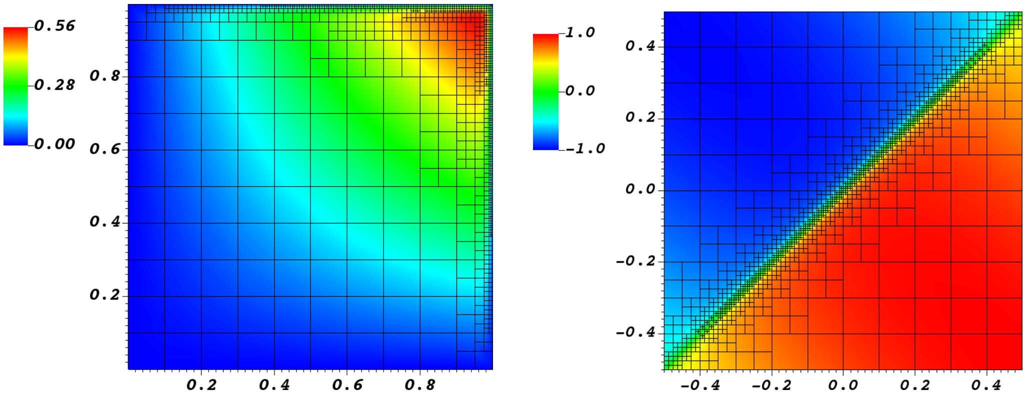


Figure 5.3: Convergence histories of the rotating pulse test case. From left to right: $\varepsilon = 10^{-3}$, $\varepsilon = 10^{-4}$ and $\varepsilon = 10^{-5}$. Top row: $\delta t_{\mathcal{K}} = h_K$; middle row: $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$; bottom row: efficiency index for both $\delta t_{\mathcal{K}} = h_K$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$.



(a) Boundary layer test case.

(b) Interior layer test case.

Figure 5.4: The boundary and interior layer solutions at time $t = 1.0$ for sections 5.5.2 and 5.5.3 respectively. Both solutions are for $\varepsilon = 10^{-3}$.

It is known that for small ε, the solution features boundary layers of width $\mathcal{O}(\varepsilon)$ at the outflow boundary of the spatial domain. See fig. 5.4a for an example when $\varepsilon = 10^{-3}$ and $\mathcal{T}_h$ has 20663 elements.

Set $p_t = p_s = 1$. We perform three convergence tests with $\varepsilon = 10^{-2}$, 10^{-3}, and 10^{-4}. For $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ we present in fig. 5.5 the convergence histories of $|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h}$, for both uniform and adaptive mesh refinements, and of η for adaptive mesh refinement.

For both $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ we observe that for $\varepsilon = 10^{-2}, 10^{-3}$ and with AMR, the error $|||\boldsymbol{u} - \boldsymbol{u}_h|||_{sT,h}$ converges with optimal rate $\mathcal{O}(N^{-1/3})$ in the asymptotic regime where the layer has been sufficiently resolved. This is not the case for $\varepsilon = 10^{-4}$ where more refinement cycles are needed to resolve the layer. However, solutions on adaptively refined meshes show better accuracy than those on uniformly refined meshes. These results verify reliability and efficiency of the estimator proven in theorem 5.1 and theorem 5.2. Furthermore, the efficiency indices depicted in fig. 5.5 show nonrobustness of order $\varepsilon^{-1/2}$ in the pre-asymptotic regime and robustness in the asymptotic regime. These results once again lie within the interval commented on in remark 3.

5.5.3 An interior layer test

In this test case, problem eq. (2.1) is set up on the spatial domain $\Omega = (-0.5, 0.5)^2$ and the time interval $I = (0, 1]$. We set $\beta = (1, 1, 1)^\intercal$ and set the initial condition, boundary

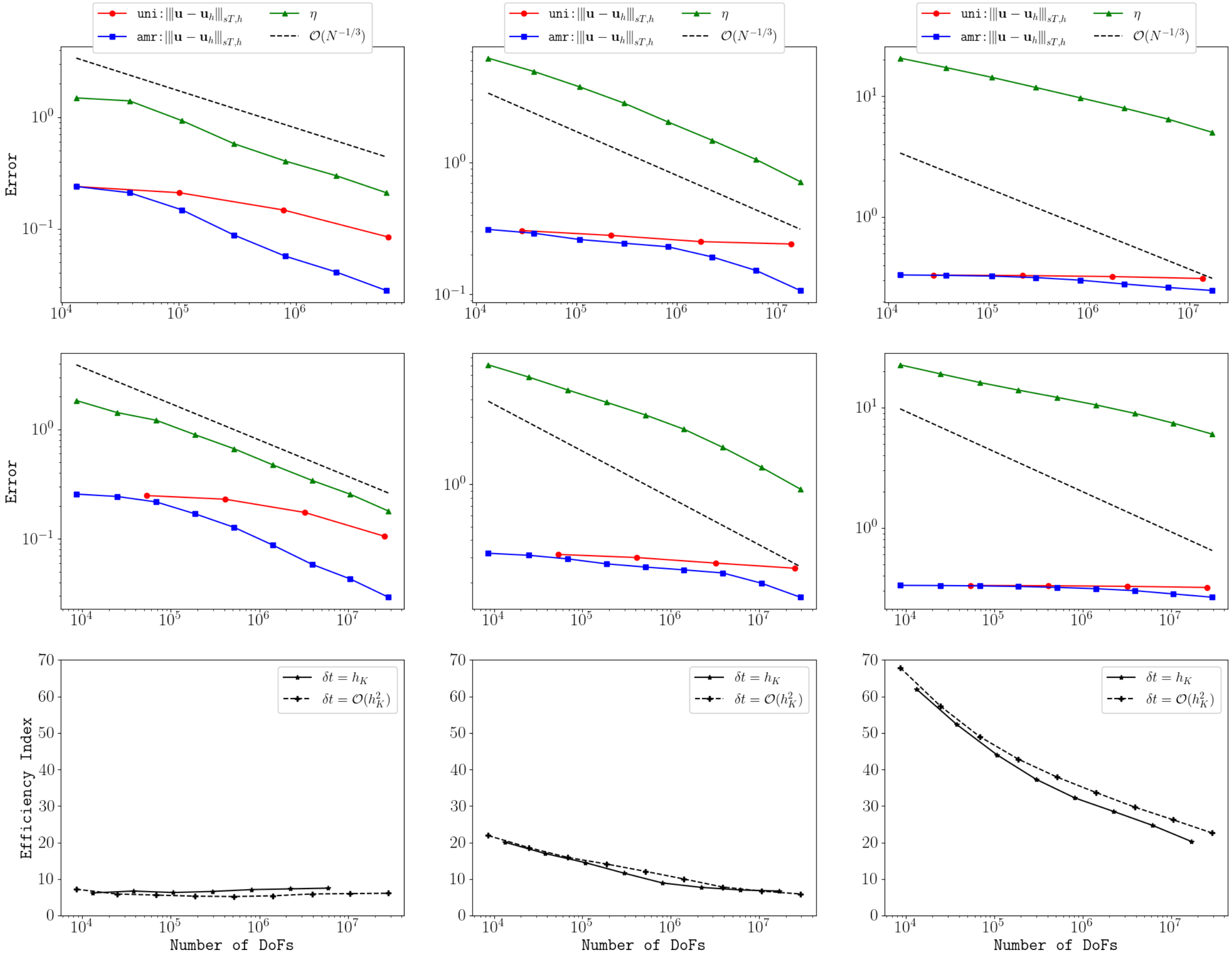


Figure 5.5: Convergence histories of the boundary layer test case. From left to right: $\varepsilon = 10^{-2}$, $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$. Top row: $\delta t_{\mathcal{K}} = h_K$; middle row: $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$; bottom row: efficiency index for both $\delta t_{\mathcal{K}} = h_K$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$.

condition and the source term such that the exact solution is given by

$$u(t,x_1,x_2) = (1-\exp(-t))\left(\arctan\left(\tfrac{y-x}{\sqrt{2}\varepsilon}\right)\right)\left(1-\tfrac{(x+y)^2}{2}\right).$$

This solution has a diagonal interior layer on the spatial domain. See fig. 5.4b for an example when $\varepsilon = 10^{-3}$ and when $\mathcal{T}_h$ has 23169 elements.

As in section 5.5.2, we perform three convergence tests with $\varepsilon = 10^{-2}$, 10^{-3}, and 10^{-4}. For $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, we present in fig. 5.6 the convergence histories of $|||\boldsymbol{u}-\boldsymbol{u}_h|||_{sT,h}$, for both uniform and adaptive mesh refinements, and of η for the adaptive mesh refinement.

Both for $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$, when $\varepsilon = 10^{-2}$, solutions obtained on adaptively refined meshes converge with the optimal rate $\mathcal{O}(N^{-1/3})$. On uniformly refined meshes, solutions converge slightly slower than the optimal rate. For $\varepsilon = 10^{-3}$, adaptive meshes yield better solutions which converge slightly faster than the optimal rate in the asymptotic regime. Solutions on uniformly refined meshes converge with a sub-optimal rate. For $\varepsilon = 10^{-4}$, both solutions on adaptively refined meshes and uniformly refined meshes converge sub-optimally. However, the former still performs better than the latter. Efficiency indices for all three cases are bounded above by 10, demonstrating robustness of the error estimator η for this test case.

The results from fig. 5.6 verify reliability and efficiency of the estimator proven in theorem 5.1 and theorem 5.2. The robustness result of the error estimator η again lies within the interval commented on in remark 3.

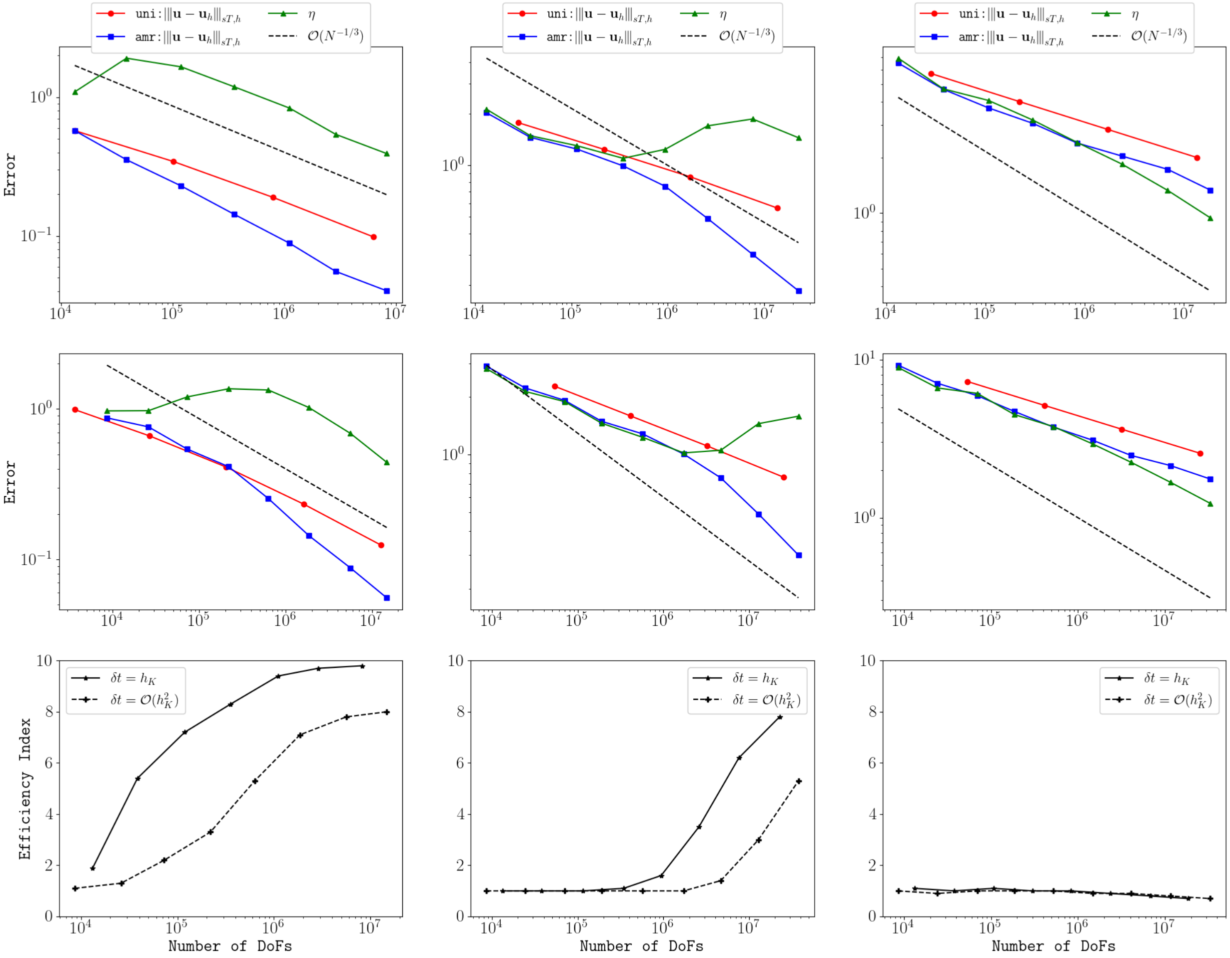


Figure 5.6: Convergence histories of the interior layer test case. From left to right: $\varepsilon = 10^{-2}$, $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$. Top row: $\delta t_{\mathcal{K}} = h_K$; middle row: $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$; bottom row: efficiency index for both $\delta t_{\mathcal{K}} = h_K$ and $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$.

Chapter 6

Conclusion

In this thesis we provided an a priori error analysis of a space-time HDG method for the time-dependent advection-dominated advection-diffusion equation on deforming domains. We proved a novel inf-sup stability result in theorem 4.1 for the space-time HDG discretization (eq. (2.9)) in the advection-dominated regime with respect to a norm ($|||\cdot|||_{ss,h}$ in eq. (2.7c)) that measures the error in its usual energy-type norm, its time derivative and its streamline derivative. Based on this inf-sup stability result, we derived in theorem 4.2 an a priori error estimate that shows a drop from $p + 1/2$ to p in the rate of convergence when transitioning from a mesh size larger than the diffusion parameter ε to a mesh size smaller than ε. A numerical example with a smooth Gaussian rotating pulse supports our error estimate. When the exact solution exhibits sharp layers, and when the mesh size is sufficiently small to resolve the layer, the error estimate predicts a rate of convergence p. This prediction is supported by a boundary layer example. We also demonstrated that in the pre-asymptotic regime, and when measuring the error only in that part of the domain that excludes the layer, we obtain a rate of convergence $p + 1/2$, again in agreement with the error estimate.

We then presented and analyzed an a posteriori error estimator for the space-time HDG method of the time-dependent advection-diffusion problem with adaptive mesh refinement on fixed domains. We proved, and verified numerically, reliability and local efficiency of the error estimator with respect to a locally computable norm. Numerical simulations showed, through an AMR procedure, that the error estimator is able to produce meshes on which solutions converge optimally. In particular, when sharp layers are present, optimal convergence occurs in the asymptotic regime. Furthermore, both the reliability and the local efficiency results derived in theorems 5.1 and 5.2 are nonrobust and together they lead to a bound for the efficiency index in the interval $[\varepsilon^{1/2}, \varepsilon^{-1}]$. In the numerical simulations, we observed the efficiency index to fall within this range. Finally, we remark that the proofs of theorems 5.1 and 5.2 assume $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$. The numerical examples, however, have shown that this assumption may be relaxed since similar numerical results are obtained with $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$.

We conclude this thesis by discussing potential directions for future work. As shown in section 5.5, the mesh size ratio constraint $\delta t_{\mathcal{K}} = \mathcal{O}(h_K^2)$ appears to be an assumption that can be relaxed in practice. It is therefore of significant interest to pursue an a posteriori error analysis based on $\delta t_{\mathcal{K}} = \mathcal{O}(h_K)$. An equally interesting direction lies in removing the saturation assumption in the reliability proof. Besides the obvious theoretical improvement, this would potentially lead to an a posteriori error analysis for arbitrary order accurate in time space-time HDG discretizations. Two additional extensions of the theory should be mentioned and, in the author's opinion, pose less of a challenge. The first is to incorporate *hp*-adaptivity in the reliability and local efficiency bounds and into the AMR procedure; the second is to extend the analysis so that it applies to problems that evolve on moving domains.

Finally, a possible next venue of the Péclet-robust a priori error analysis and the novel inf-sup stability therein is the time-dependent Oseen equation (which can be viewed as advection-diffusion of the linear momentum in fluid dynamics) on moving domains. With such an a priori error analysis available, it would be a natural next step to derive and analyze an a posteriori error estimator and implement the AMR procedure for the time-dependent Oseen equation.

References

[1] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* John Wiley & Sons, Inc., 2000.

[2] R. Almeida and R. Silva. A stable Petrov-Galerkin method for convection-dominated problems. *Comput. Methods Appl. Mech. Engrg.*, 140(3–4):291–304, 1997.

[3] P. Amestoy, I. Duff, J. L'Excellent, and J. Koster. MUMPS: a general purpose distributed memory sparse solver. *International Workshop on Applied Parallel Computing*, pages 121–130, 2000.

[4] P. Amestoy, A. Guermouche, J. L'Excellent, and S. Parlet. Hybrid scheduling for the parallel solution of linear systems. *Parallel Comput.*, 32:136–156, 2006.

[5] T. M. Apostol. *Mathematical Analysis.* Addison-Wesley, 1974.

[6] R. Araya, M. Solano, and P. Vega. A posteriori error analysis of an HDG method for the Oseen problem. *Appl. Numer. Math.*, 146:291–308, 2019.

[7] R. Araya and P. Venegas. An a posteriori error estimator for an unsteady advection–diffusion–reaction problem. *Comput. Math. Appl.*, 66(12):2456–2476, 2014.

[8] D. Arndt, W. Bangerth, M. Bergbauer, M. Feder, M. Fehling, J. Heinz, T. Heister, L. Heltai, M. Kronbichler, M. Maier, P. Munch, J.-P. Pelteret, B. Turcksin, D. Wells, and S. Zampini. The `deal.II` library, version 9.5. *J. Numer. Math.*, 31(3):231–246, 2023.

[9] D. Arndt, W. Bangerth, D. Davydov, T. Heister, L. Heltai, M. Kronbichler, M. Maier, J.-P. Pelteret, B. Turcksin, and D. Wells. The deal.II finite element library: Design, features, and insights. *Comput. Math. Appl.*, 81:407–422, 2021.

[10] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2002.

[11] M. Augustin, A. Caiazzo, A. Fiebach, J. Fuhrmann, V. John, A. Linke, and R. Umla. An assessment of discretizations for convection-dominated convection–diffusion equations. *Comput. Methods Appl. Mech. Engrg.*, 200(47–48):3395–3409, 2011.

[12] B. Ayuso and L. D. Marini. Discontinuous Galerkin methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 47(2):1391–1420, 2009.

[13] S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. C. McInnes, R. T. Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, and J. Zhang. PETSc/TAO Users Manual. Technical Report ANL-21/39 - Revision 3.19, Argonne National Laboratory, 2023.

[14] S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. M. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. C. McInnes, R. T. Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, and J. Zhang. PETSc Web page. `https://petsc.org/`, 2023.

[15] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith. Efficient Management of Parallelism in Object Oriented Numerical Software Libraries. In E. Arge, A. M. Bruaset, and H. P. Langtangen, editors, *Modern Software Tools in Scientific Computing*, pages 163–202. Birkhäuser Press, 1997.

[16] W. Bangerth, C. Burstedde, T. Heister, and M. Kronbichler. Algorithms and data structures for massively parallel generic adaptive finite element codes. *ACM Trans. Math. Software*, 38(2):14:1–14:28, 2011.

[17] G. Barrenechea, V. John, and P. Knobloch. A local projection stabilization finite element method with nonlinear crosswind diffusion for convection-diffusion-reaction equations. *ESAIM Math. Model. Numer. Anal.*, 47(5):1335–1366, 2013.

[18] F. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM J. Numer. Anal.*, 33(3):2431–2444, 1996.

[19] D. Braess and R. Verfürth. A posteriori error estimators for the Raviart–Thomas element. *SIAM J. Numer. Anal.*, 33(6):2431–2444, 1996.

[20] F. Brezzi, B. Cockburn, L. D. Marini, and E. Süli. Stabilization mechanisms in discontinuous Galerkin finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 195(25):3293–3310, 2006.

[21] A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 32(1–3):199–259, 1982.

[22] E. Burman. A unified analysis for conforming and nonconforming stabilized finite element methods using interior penalty. *SIAM J. Numer. Anal.*, 43(5):2012–2033, 2005.

[23] E. Burman. A posteriori error estimation for interior penalty finite element approximations of the advection-reaction equation. *SIAM J. Numer. Anal.*, 47(5):3584–3607, 2009.

[24] E. Burman. Consistent SUPG-method for transient transport problems: Stability and convergence. *Comput. Methods Appl. Mech. Engrg.*, 199(17–20):1114–1123, 2010.

[25] E. Burman and A. Ern. Stabilized Galerkin approximation of convection-diffusion-reaction equations: discrete maximum principle and convergence. *Math. Comp.*, 74(252):1637–1652, 2005.

[26] E. Burman and M. A. Fernández. Finite element methods with symmetric stabilization for the transient convection-diffusion-reaction equation. *Comput. Methods Appl. Mech. Engrg.*, 198(33–36):2508–2519, 2009.

[27] E. Burman and P. Hansbo. Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems. *Comput. Methods Appl. Mech. Engrg.*, 193(15–16):1437–1453, 2004.

[28] E. Burman and B. Stamm. Minimal stabilization for discontinuous Galerkin finite element methods for hyperbolic problems. *J. Sci. Comput.*, 33:183–208, 2007.

[29] C. Burstedde, L. C. Wilcox, and O. Ghattas. `p4est`: Scalable Algorithms for Parallel Adaptive Mesh Refinement on Forests of Octrees. *SIAM J. Sci. Comput.*, 33(3):1103–1133, 2011.

[30] A. Cangiani, Z. Dong, E. H. Georgoulis, and P. Houston. *hp-Version Discontinuous Galerkin Methods on Polygonal and Polyhedral Meshes.* Springer Briefs in Mathematics. Springer Cham, 2017.

[31] A. Cangiani, E. H. Georgoulis, S. Giani, and S. Metcalfe. *hp*-adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems. *Comput. Math. Appl.*, 78:3090–3104, 2019.

[32] A. Cangiani, E. H. Georgoulis, and S. Metcalfe. Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems. *IMA J. Numer. Anal.*, 34:1578–1597, 2014.

[33] H. Chen, J. Li, and W. Qiu. Robust a posteriori error estimates for HDG method for convection-diffusion equations. *IMA J. Numer. Anal.*, 36:437–462, 2016.

[34] Y. Chen and B. Cockburn. Analysis of variable-degree HDG methods for convection-diffusion equations. Part I: general nonconforming meshes. *IMA J. Numer. Anal.*, 32(4):1267–1293, 2012.

[35] B. Cockburn. Static Condensation, Hybridization, and the Devising of the HDG Methods. In G. Barrenechea, F. Brezzi, A. Cangiani, and E. Georgoulis, editors, *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, pages 129—177. Springer, Cham, 2016.

[36] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[37] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463, 1998.

[38] B. Cockburn and C.-W. Shu. Runge–Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16:173–261, 2001.

[39] J. de Frutos, B. García-Archilla, and J. Novo. Local error estimates for the SUPG method applied to evolutionary convection-reaction-diffusion equations. *J. Sci. Comput.*, 66:528—554, 2016.

[40] A. Devinatz, R. Ellis, and A. Friedman. The asymptotic behavior of the first real eigenvalue of second order elliptic operators with a small parameter in the highest derivatives. II. *Indiana Univ. Math. J.*, 23(11):991–1011, 1975.

[41] V. Dolejší and M. Feistauer. *Discontinuous Galerkin Method*, volume 48 of *Springer Series in Computational Mathematics.* Springer Cham, 2015.

[42] J. Douglas and T. Dupont. Interior penalty procedures for elliptic and parabolic Galerkin methods. In R. Glowinski and J. L. Lions, editors, *Computing Methods in Applied Sciences*, pages 207–216. Springer Berlin Heidelberg, 1976.

[43] W. Eckhaus. Boundary layers in linear elliptic singular perturbation problems. *SIAM Review*, 14(2):225–270, 1972.

[44] H. Egger and J. Schöberl. A hybrid mixed discontinuous Galerkin finite-element method for convection-diffusion problems. *IMA J. Numer. Anal.*, 30(4):1206–1234, 2010.

[45] A. Ern and J. Guermond. *Finite Elements I*, volume 72 of *Texts in Applied Mathematics.* Springer Cham, 1 edition, 2021.

[46] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. I. General theory. *SIAM J. Numer. Anal.*, 44(2):753–778, 2006.

[47] A. Ern and A. Stephansen. A posteriori energy-norm error estimates for advection-diffusion equations approximated by weighted interior penalty methods. *J. Comput. Math.*, 26(4):488–510, 2008.

[48] A. Ern, A. Stephansen, and M. Vohralík. Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems. *J. Comput. Appl. Math.*, 234:114–130, 2010.

[49] R. E. Ewing, editor. *The Mathematics of Reservoir Simulation.* Society for Industrial and Applied Mathematics, 1983.

[50] S. M. Fallat and C. R. Johnson. *Totally nonnegative matrices.* Princeton University Press, 2022.

[51] M. Feistauer, V. Kučera, K. Najzar, and J. Prokopová. Analysis of space-time discontinuous Galerkin method for nonlinear convection-diffusion problems. *Numer. Math.*, 117:251–288, 2011.

[52] G. Fu, W. Qiu, and W. Zhang. An analysis of HDG methods for convection-dominated diffusion problems. *ESAIM Math. Model. Numer. Anal.*, 49(1):225–256, 2015.

[53] E. H. Georgoulis. *Discontinuous Galerkin Methods on Shape-Regular and Anisotropic Meshes.* PhD thesis, University of Oxford, 2003.

[54] E. H. Georgoulis. *hp*-version interior penalty discontinuous Galerkin finite element methods on anisotropic meshes. *Int. J. Numer. Anal. Model.*, 3(1):52–79, 2006.

[55] E. H. Georgoulis, O. Lakkis, and J. M. Virtanen. A posteriori error control for discontinuous Galerkin methods for parabolic problems. *SIAM J. Numer. Anal.*, 49(2):427–458, 2011.

[56] S. Giani, D. Schötzau, and L. Zhu. An a-posteriori error estimate for *hp*-adaptive DG methods for convection–diffusion problems on anisotropically refined meshes. *Comput. Math. Appl.*, 67(4):869–887, 2014.

[57] H. Goering, A. Felgenhauer, G. Lube, H. Roos, and L. Tobiska. *Singularly perturbed differential equations.* Akademie-Verlag, Berlin, 1983.

[58] S.-Y. Hahn, J. Bigeon, and J.-C. Sabonnadiere. An 'upwind' finite element method for electromagnetic field problems in moving media. *Int. J. Numer. Methods Eng.*, 24:2071–2086, 1987.

[59] V. E. Henson and U. M. Yang. BoomerAMG: A parallel algebraic multigrid solver and preconditioner. *Appl. Numer. Math.*, 41(1):155–177, 2002.

[60] C. Hirsch. *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics.* Elsevier, 2007.

[61] P. Houston, D. Schötzau, and T. P. Wihler. Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Models Methods Appl. Sci.*, 17(1):33–62, 2007.

[62] P. Houston, C. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163, 2002.

[63] P. Houston, E. Süli, and T. P. Wihler. A posteriori error analysis of *hp*-version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs. *IMA J. Numer. Anal.*, 28:245–273, 2008.

[64] T. J. R. Hughes, L. P. Franca, and M. Mallet. A new finite element formulation for computational fluid dynamics: VI. Convergence analysis of the generalized SUPG formulation for linear time-dependent multidimensional advective-diffusive systems. *Comput. Methods Appl. Mech. Engrg.*, 63(1):97–112, 1987.

[65] M. Jakob. *Heat transfer.* Wiley, New York, 1959.

[66] P. Jamet. Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM J. Numer. Anal.*, 15(5):912–928, 1978.

[67] P. Jamet and B. Bonnerot. Numerical solution of the Eulerian equations on compressible flow by a finite element method which follows the free boundary and the interfaces. *J. Comput. Phys.*, 18(1):21–45, 1975.

[68] V. John, , P. Knobloch, and J. Novo. Finite elements for scalar convection-dominated equations and incompressible flow problems: a never ending story? *Comput. Visualization Sci.*, 19:47–63, 2018.

[69] V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I – A review. *Comput. Methods Appl. Mech. Engrg.*, 196(17–20):2197–2215, 2007.

[70] V. John and P. Knobloch. On the performance of SOLD methods for convection-diffusion problems with interior layers. *Int. J. Comput. Sci. Math.*, 1(2–4):245–258, 2007.

[71] V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part II – Analysis for P_1 and Q_1 finite elements. *Comput. Methods Appl. Mech. Engrg.*, 197(21–24):1997–2014, 2008.

[72] V. John and J. Novo. Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations. *SIAM J. Numer. Anal.*, 49(3):1149–1176, 2011.

[73] M. Kadalbajoo and V. Gupta. A brief survey on numerical methods for solving singularly perturbed problems. *Appl. Math. Comput.*, 217(8):3641–3716, 2010.

[74] G. Kanschat. *Discontinuous Galerkin Methods for Viscous Incompressible Flow.* Teubner Research, 2008.

[75] O. A. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(6):2374–2399, 2003.

[76] K. L. A. Kirk, T. L. Horvath, A. Cesmelioglu, and S. Rhebergen. Analysis of a space-time hybridizable discontinuous Galerkin method for the advection-diffusion problem on time-dependent domains. *SIAM J. Numer. Anal.*, 57(4):1677–1696, 2019.

[77] P. Knobloch. A generalization of the local projection stabilization for convection-diffusion-reaction equations. *SIAM J. Numer. Anal.*, 48(2):659–680, 2010.

[78] P. Knobloch and G. Lube. Local projection stabilization for advection–diffusion–reaction problems: One-level vs. two-level approach. *Appl. Numer. Math.*, 59(12):2891–2907, 2009.

[79] P. Knobloch and L. Tobiska. On the stability of finite-element discretizations of convection–diffusion–reaction equations. *IMA J. Numer. Anal.*, 31(1):147–164, 2011.

[80] H. O. Kreiss and J. Lorenz. *Initial-boundary value problems and the Navier–Stokes equations.* Society for Industrial and Applied Mathematics, 2004.

[81] G. Kunert. A posteriori error estimation for convection dominated problems on anisotropic meshes. *Math. Methods Appl. Sci.*, 26(7):589–617, 2003.

[82] O. Lakkis and C. Makridakis. Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, 75(256):1627–1658, 2006.

[83] C. Makridakis and R. H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594, 2003.

[84] J. R. Munkres. *Analysis on Manifolds.* Addison-Wesley, Redwood City, CA, 1991.

[85] N. C. Nguyen, J. Peraire, and B. Cockburn. An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations. *J. Comput. Phys.*, 228(9):3232–3254, 2009.

[86] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69 of *Mathématiques et Applications.* Springer–Verlag Berlin Heidelberg, 2012.

[87] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical report, Los Alamos Scientific Laboratory, Tech. Report LA-UR-73-479, 1973.

[88] S. Rhebergen and B. Cockburn. A space-time hybridizable discontinuous Galerkin method for incompressible flows on deforming domains. *J. Comput. Phys.*, 231(11):4185–4204, 2012.

[89] S. Rhebergen and B. Cockburn. Space-time hybridizable discontinuous Galerkin method for the advection-diffusion equation on moving and deforming meshes. In C. A. de Moura and C. S. Kubrusly, editors, *The Courant–Friedrichs–Lewy (CFL) condition, 80 years after its discovery*, pages 45–63. Birkhäuser Science, 2013.

[90] B. Rivière. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*, volume 35 of *Frontiers in Applied Mathematics.* Society for Industrial and Applied Mathematics, Philadelphia, 2008.

[91] H. Roos. Robust numerical methods for singularly perturbed differential equations: a survey covering 2008–2012. 2012, 2012.

[92] H. Roos, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations.* Springer Series in Computational Mathematics. Springer Berlin, Heidelberg, 2008.

[93] G. Sangalli. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Math. Comp.*, 77:41–70, 2008.

[94] J. Schöberl. C++11 implementation of finite elements in NGSolve. Technical Report Technical Report ASC-2014-30, Institute for Analysis and Scientific Computing, 2014.

[95] D. Schötzau and L. Zhu. A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations. *Appl. Numer. Math.*, 59:2236–2255, 2009.

[96] N. Sharma. Robust a-posteriori error estimates for weak Galerkin method for the convection-diffusion problem. *Appl. Numer. Math.*, 170:384–397, 2021.

[97] J. J. Sudirham. *Space-time discontinuous Galerkin methods for convection-diffusion problems: Application to wet-chemical etching.* PhD thesis, University of Twente, 2005.

[98] J. J. Sudirham, J. J. W. van der Vegt, and R. M. J. van Damme. Space-time discontinuous Galerkin method for advection-diffusion problems on time-dependent domains. *Appl. Numer. Math.*, 56(12):1491–1518, 2006.

[99] J. J. W. van der Vegt and H. van der Ven. Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flows: I. General formulation. *J. Comput. Phys.*, 182(2):546–585, 2002.

[100] R. Verfürth. A posteriori error estimators for convection-diffusion equations. *Numer. Math.*, 80:641–663, 1998.

[101] R. Verfürth. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation. *Numer. Math.*, 78:479–493, 1998.

[102] R. Verfürth. Robust a posteriori error estimates for nonstationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1783–1802, 2005.

[103] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1766–1782, 2005.

[104] Y. Wang and S. Rhebergen. Space-time hybridizable discontinuous Galerkin method for advection-diffusion on deforming domains: The advection-dominated regime. Submitted. arXiv:2308.12130, 2023.

[105] Y. Wang and S. Rhebergen. A posteriori error analysis of a space-time hybridizable discontinuous Galerkin method for the advection-diffusion problem. Submitted. arXiv:2404.04130, 2024.

[106] G. N. Wells. Analysis of an interface stabilized finite element method: the advection-diffusion-reaction equation. *SIAM J. Numer. Anal.*, 49(1):87–109, 2011.

[107] L. Zhu. *Robust A Posteriori Error Estimation for Discontinuous Galerkin Methods for Convection Diffusion Problems.* PhD thesis, University of British Columbia, 2010.

[108] L. Zhu, S. Giani, P. Houston, and D. Schötzau. Energy norm a posteriori error estimation for *hp*-adaptive discontinuous Galerkin methods for elliptic problems in three dimensions. *Math. Models Methods Appl. Sci.*, 21:267–306, 2011.

[109] L. Zhu and D. Schötzau. A robust a posteriori error estimate for *hp*-adaptive DG methods for convection-diffusion equations. *IMA J. Numer. Anal.*, 31(3):971–1005, 2011.

APPENDICES

Appendix A

Some facts from differential geometry

Given a parameterization $\varrho : (-1,1)^k \to S$, where S is a k-dimensional hypersurface in $\mathbb{R}^n$ $(n \geq k)$, integration over the surface S is defined as (see [84, Theorem 21.3 and Definition on page 189]):

$$\int_S f(x)\,\mathrm{d}x = \int_{(-1,1)^k} f\left(\varrho(\xi)\right) \mathrm{vol}_k \left(\frac{\partial \varrho}{\partial \xi_1}, \frac{\partial \varrho}{\partial \xi_2}, \dots, \frac{\partial \varrho}{\partial \xi_k}\right) \mathrm{d}\xi, \tag{A.1}$$

where

$$\mathrm{vol}_k \left(\frac{\partial \varrho}{\partial \xi_1}, \frac{\partial \varrho}{\partial \xi_2}, \dots, \frac{\partial \varrho}{\partial \xi_k}\right) := \left(\det\left(\varrho_k^T \varrho_k\right)\right)^{1/2},$$

in which ϱ_k denotes the $n \times k$ matrix with column vectors $\{\frac{\partial \varrho}{\partial \xi_i}\}_{1 \leq i \leq k}$.

In the context of the space-time element $\mathcal{K}$, the diffeomorphism $\phi_{\mathcal{K}}$ maps $\mathcal{K}$ from its affine domain to its physical domain. Given a $\mathcal{Q}$-facet $F_{\mathcal{Q}}$ of $\mathcal{K}$, which is in general curved in its physical domain, we denote the restriction of $\phi_{\mathcal{K}}$ on F by $\phi_{F_{\mathcal{Q}}}$. Furthermore, without loss of generality, we assume that $\widetilde{x}_j$ is fixed on $\widetilde{F}$ for some $1 \leq j \leq d$. Based on eq. (A.1), we can view ϕ_F as a parameterization of the facet F from its affine domain to its physical domain and define the integration over F as follows

$$\begin{aligned}
\int_{F_{\mathcal{Q}}} f(x)\,\mathrm{d}x &= \int_{\widetilde{F}_{\mathcal{Q}}} f\left(\phi_{F_{\mathcal{Q}}}(\widetilde{x})\right) \mathrm{vol}_d \left(\frac{\partial \phi_{F_{\mathcal{Q}}}}{\partial \widetilde{t}}, \frac{\partial \phi_{F_{\mathcal{Q}}}}{\partial \widetilde{x}_1}, \dots, \frac{\partial \phi_{F_{\mathcal{Q}}}}{\partial \widetilde{x}_{j-1}}, \frac{\partial \phi_{F_{\mathcal{Q}}}}{\partial \widetilde{x}_{j+1}}, \dots, \frac{\partial \phi_{F_{\mathcal{Q}}}}{\partial \widetilde{x}_d}\right) \mathrm{d}\widetilde{x} \\
&= \int_{\widetilde{F}_{\mathcal{Q}}} f\left(\phi_{\mathcal{K}}(\widetilde{x})\right) \mathrm{vol}_d \left(\frac{\partial \phi_{\mathcal{K}}}{\partial \widetilde{t}}, \frac{\partial \phi_{\mathcal{K}}}{\partial \widetilde{x}_1}, \dots, \frac{\partial \phi_{\mathcal{K}}}{\partial \widetilde{x}_{j-1}}, \frac{\partial \phi_{\mathcal{K}}}{\partial \widetilde{x}_{j+1}}, \dots, \frac{\partial \phi_{\mathcal{K}}}{\partial \widetilde{x}_d}\right) \mathrm{d}\widetilde{x} \\
&= \int_{\widetilde{F}_{\mathcal{Q}}} f\left(\phi_{\mathcal{K}}(\widetilde{x})\right) \left(\det\left(\left(J^j_{\phi_{\mathcal{K}}}\right)^T J^j_{\phi_{\mathcal{K}}}\right)\right)^{\frac{1}{2}} \mathrm{d}\widetilde{x},
\end{aligned} \tag{A.2}$$

where $J_{\phi_\mathcal{K}}^j$ denotes the $(n+1) \times n$ submatrix of $J_{\phi_\mathcal{K}}$ by selecting all but its j^{th} column vectors.

In [99, Appendix B], an alternative definition of eq. (A.2) is given:

$$\int_{F_\mathcal{Q}} f(x)\,\mathrm{d}x = \int_{\widetilde{F}_\mathcal{Q}} f\left(\phi_\mathcal{K}(\widetilde{x})\right) \left| \frac{\partial \phi_\mathcal{K}}{\partial \widetilde{t}} \wedge \frac{\partial \phi_\mathcal{K}}{\partial \widetilde{x}_1} \wedge \cdots \wedge \frac{\partial \phi_\mathcal{K}}{\partial \widetilde{x}_{j-1}} \wedge \frac{\partial \phi_\mathcal{K}}{\partial \widetilde{x}_{j+1}} \wedge \cdots \wedge \frac{\partial \phi_\mathcal{K}}{\partial \widetilde{x}_d} \right| \mathrm{d}\widetilde{x}, \quad \text{(A.3)}$$

where the outer product "∧" is used. In general, $v = w_1 \wedge \cdots \wedge w_{n-1}$, for $n-1$ vectors w_i in $\mathbb{R}^n$, is defined component-wise by the rule:

$$v^j = \det(w_1, \cdots, w_{n-1}, e_j),$$

with v^j denoting the j^{th} component of the vector v and e_j denoting the j^{th} basis vector in $\mathbb{R}^n$.

The two definitions, eq. (A.2) and eq. (A.3), are indeed equivalent. Below, we show the equivalence using generic notations as in eq. (A.1) where we consider a parameterization $\varrho : (-1,1)^{n-1} \to S$ with S being a $(n-1)$-dimensional hypersurface in $\mathbb{R}^n$:

$$\begin{aligned}
\left| \frac{\partial \varrho}{\partial \xi_1} \wedge \frac{\partial \varrho}{\partial \xi_2} \wedge \cdots \wedge \frac{\partial \varrho}{\partial \xi_{n-1}} \right|^2 &= \sum_{i=1}^{n} \left(\det\left(\frac{\partial \varrho}{\partial \xi_1}, \ldots, \frac{\partial \varrho}{\partial \xi_{n-1}}, e_i \right) \right)^2 \\
&= \sum_{i=1}^{n} \left(\det\left(\left[\frac{\partial \varrho}{\partial \xi_1}, \ldots, \frac{\partial \varrho}{\partial \xi_{n-1}}, e_i \right]^T \left[\frac{\partial \varrho}{\partial \xi_1}, \ldots, \frac{\partial \varrho}{\partial \xi_{n-1}}, e_i \right] \right) \right) \\
&= \sum_{i=1}^{n} \left(\det\left(\begin{bmatrix} \varrho_{n-1}^T \varrho_{n-1} & \varrho_{n-1}^T e_i \\ e_i^T \varrho_{n-1} & 1 \end{bmatrix} \right) \right) \\
&= \sum_{i=1}^{n} \left(\det\left(\varrho_{n-1}^T \varrho_{n-1} \right) \left(1 - e_i^T \varrho_{n-1} \left(\varrho_{n-1}^T \varrho_{n-1} \right)^{-1} \varrho_{n-1}^T e_i \right) \right) \\
&= \det\left(\varrho_{n-1}^T \varrho_{n-1} \right) \left(n - \operatorname{tr}\left(\varrho_{n-1} \left(\varrho_{n-1}^T \varrho_{n-1} \right)^{-1} \varrho_{n-1}^T \right) \right) \\
&= \det\left(\varrho_{n-1}^T \varrho_{n-1} \right),
\end{aligned}$$

where we used a few facts from linear algebra: $\det(AB) = \det(A)\det(B)$, $\det\left(A^T\right) = \det(A)$, $\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \right) = \det(A)\det\left(D - CA^{-1}B\right)$ and finally, $\operatorname{tr}\left(A\left(A^T A\right)^{-1} A^T\right) = \operatorname{rank}(A)$.