A posteriori error estimation for simulation of diffusion and fluid mechanics problems by finite volume techniques

Anh Ha Le

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS XIII

Spécialité

MATHÉMATIQUES APPLIQUÉES

Présentée par
Anh Ha LE

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ PARIS XIII

Sujet de la thèse:

ESTIMATIONS D'ERREUR A POSTERIORI POUR LA SIMULATION DE PROBLÈMES DE DIFFUSION ET DE MÉCANIQUE DES FLUIDES PAR DES MÉTHODES DE VOLUMES FINIS

Organisme d'accueil:

COMMISSARIAT À L'ÉNERGIE ATOMIQUE, SACLAY

Soutenue le 25 Octobre 2011 devant le jury composé de:

Fayssal BENKHALDOUN	Université Paris XIII	
Claire CHAINAIS-HILLAIRET	Université Lille I	
Emmanuel CREUSÉ	Université Lille I	Rapporteur
Alexandre ERN	Université Paris-Est	
Laurence HALPERN	Université Paris XIII	
Florence HUBERT	Université de Provence	
Pascal OMNES	CEA Saclay	Directeur de thèse
Martin VOHRALÍK	Université Paris VI	Rapporteur

ACKNOWLEDGEMENTS

I am deeply indebted to my thesis advisor Pascal Omnes, for his constant support, invaluable guidance, patience, mathematical inspiration throughout my work at CEA, Saclay. I appreciated his assistance in writing reports (i.e., this thesis, applications to postdoc positions ...) and about my life in France (i.e., documents, studying French …).

I would like to thank Emmanuel Creusé and Martin Vohralík for accepting to be reporters and their careful reading of my thesis. Moreover, I thank again Martin Vohralík for helpful discussions about the nonlinear diffusion equations.

I also would like to acknowledge Fayssal Benkhaldoun, Claire Chainais-Hillairet, Alexandre Ern, Laurence Halpern and Florence Hubert for serving in the thesis committee; special thanks to Fayssal Benkhaldoun for being my thesis advisor during the first and second years, to Laurence Halpern for letting me know about this PhD subject, to Claire Chainais-Hillairet for many useful literature references about the transport equations.

I would like to express my gratitude to Stéphane Dellacherie for inviting me to work on the transport equations and for many useful discussions.

I would also like to thank CEA Saclay for financial assistance and providing an attractive place for me to do and enjoy the mathematics, my colleges at CEA, Jean-Luc, Philippe, Jean-Christophe, Christian G., Christian V., Pierre, Paul-Marie, Samuel, Zac, Bilal, Kun Kun, François, my Vietnamese friends Hai, Hoang, Bin, Binh, Lan, Liem, Huyen, Cuc, Ngoc, Mi for helping me in many ways in my life.

I gratefully thank Nadine and her family for helping me to understand the life in France, explaining many French documents, teaching and talking French with me daily, introducing French food, especially, making me think that I'm living with my family, to my best friend Tam for strong encouraging and helping me during difficult times.

Finally, I would also like to thank my family for encouraging me, believing in me, providing me the necessary things for my studies during difficult periods for the family. The dissertation is dedicated to them.

Contents

Introduction

When approaching the solution of a partial differential equation (PDE), by any numerical method (finite elements, finite volumes, ...), the user may wish, on the one hand, to be able to guarantee the quality of the computation, by imposing for example that a certain norm of the error between the exact solution (called $\hat{\phi}$ in what follows) and the computed solution (called ϕ_h in what follows) is lower than a prescribed tolerance, and, on the other hand, to minimize the computer resources that are needed to find a numerical solution that satisfies this tolerance, by using a mesh well adapted to the solution.

For this purpose, *a priori* error estimation does not help a lot: they are generally written under the form

$$\left\|\hat{\phi} - \phi_h\right\| \leq Ch^{\alpha} \left\|\hat{\phi}\right\|,$$

with norms that need not be specified in this general introduction, and are obtained by supposing some regularity properties to the solution of the considered PDE (a regularity which may not be verified in practice, for example due to the presence of shocks, of geometrical singularities, ...). Moreover, the generic constant C which appears in the above estimation is either unknown or very difficult to estimate. But the main difficulty is that the norm of the exact solution (the term $\left\|\hat{\phi}\right\|$ above) is very often not computable, because the exact solution is unknown, and it does not provide any local information that could show where to refine the mesh in priority.

It is the purpose of *a posteriori* error estimation to give some answers to the wishes discussed above. Some of the first investigators who developed rigorous analysis in that direction were Babuska and Rheinboldt more than thirty years ago [3, 4]. The aim of this kind of estimation is to obtain a general expression of the form

$$\left\|\hat{\phi} - \phi_h\right\| \leq \eta\left(\tau_h, \phi_h, f\right) = \left(\sum_i C_i\left(\omega_i\right) \eta_i^{\beta}(\phi_h, f)\right)^{1/\beta}, \tag{1}$$

where τ_h is the mesh on which the solution ϕ_h was calculated, f represents the data of the problem (source term, boundary conditions, initial conditions in the case of a time-dependant problem, coefficients in the equations, ...). The sum over i is a sum over some elements (named here ω_i) of the mesh (of the domain Ω if the problem is stationary, or of the space-time domain $\Omega \times [0, T]$, where T is the final time of the simulation, if the problem is time-dependent). These elements of the mesh may not always be the cells themselves, but some neighbors of a given cell, for example. In general, the constants $C_i\left(\omega_i\right)$ depend on the shape and/or size of the cells included in ω_i. The local estimators η_i depend only on the numerical solution ϕ_h and of its variations on ω_i, and on the data of the continuous problem,

but they never depend on the exact solution. As far as the real number β is concerned, it may depend on the type of PDE considered, but also on the type of norm in which the error is measured. When the constants C_i are unknown or may not be easily computed, we may not ensure that the actual error is lower than a given tolerance because one is unable to entirely compute the right-hand side in the above inequality. On the contrary, if one may compute them, the bound is said to be fully computable, or guaranteed. But in both cases, one may use the local estimator η_i, which is always computable, to determine the zones of the mesh which have to be refined in priority: these will be the zones where η_i is "large" with respect to other zones. What "large" means and the strategy that may be achieved to refine the mesh may depend on the problem.

In order to ensure that the global estimator η is useful to estimate the error in practice, it is interesting to consider the ratio

$$E := \frac{\eta\left(\tau_h, \phi_h, f\right)}{\left\|\hat{\phi} - \phi_h\right\|},$$

which is named efficiency of the estimator. In particular, it is important to verify that this quantity is bounded by above: the estimator is then said to be efficient. Indeed, if this quantity could be arbitrarily large, it wouldn't mean much to estimate the error by the estimator. One is also interested in the local efficiency

$$E_i := \frac{\eta_i}{\left\|\hat{\phi} - \phi_h\right\|_{\omega_i}},$$

which allows to ensure that the local estimator η_i does not overestimate too much the error on ω_i; if this quantity is bounded independently of the index i, then a refinement strategy based on the local values of the quantities η_i may lead to almost optimal meshes in terms of number of degrees of freedom for a given error tolerance. Moreover, the ideal wish is that the estimator E should be asymptotically exact, meaning that $\lim_{h\to 0} E = 1$; indeed, one would then have an almost exact estimation of the error.

A last property which is desirable for the estimators η_i is that they should be cheap to compute (much cheaper than solving the global problem on the given mesh), so that the overall computational cost is not seriously increased by the error estimation.

The PDEs that we consider in this thesis are all related to fluid dynamics. They may all be considered as submodels of the more general compressible Navier-Stokes equations, where the unknowns ρ, $\mathbf{u}$, p, which are respectively the density, the velocity and the pressure of the fluid, are related by the following equations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{2}$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{3}$$

$$p = p(\rho). \tag{4}$$

Equation (2) is the continuity, or mass conservation equation, while (3) is the momentum equation, in which ν is the viscosity. Finally, (4) is a constitutive law that relates the pressure to the density.

In Chapter 1, we consider a very usual submodel of (2)–(4): the stationary Stokes problem. It is obtained by considering that ρ is a constant with respect to both time and space (the fluid is then said to be incompressible), so that (2) simply becomes $\nabla \cdot \mathbf{u} = 0$. Then neglecting the non-linear term in (3), we get the simpler equation $-\nu\Delta\mathbf{u} + \nabla p = \mathbf{f}$.

In Chapter 2, we consider the non-linear Darcy equation. It is a model for incompressible flows ($\nabla \cdot \mathbf{u} = 0$) in porous media and is widely used to modelize water flow in the underground, with applications in the petroleum industry, or in the field of nuclear waste storage, among others. In the linear setting, this law was obtained by Darcy through experiments and expresses a linear relationship between the fluid velocity and the pressure gradient ($\mathbf{u} = k\nabla p$). It was later justified by the mathematical theory of homogenization (see for example [6] and the references therein) by considering that the viscosity ν of the fluid in the Navier-Stokes equations is related to the size ε of the pores in the media by the (nondimensionalized) relationship $\nu = \varepsilon^\beta$, with some constant $\beta > 0$. We understand from [6] that the critical value $\beta = 3/2$ plays a special role in this theory because the homogenized system is then non-linear. On the other hand, for $\beta < 3/2$, the limit system is the linear Darcy law, but for β close to $3/2$, it is argued that non-linear terms may not completely be neglected if one wishes an accurate modeling. Therefore, it is natural to consider non-linear corrections to this law under the form: $\mathbf{u} = H(p)\nabla p$. Note that non-linear equations of the form $\mathbf{u} = H(\nabla p)\nabla p$ may also be considered, but not in this work.

Finally, in Chapter 3, we consider only the mass conservation equation, where we suppose that the velocity $\mathbf{u}$ is given. Then, this is a linear transport equation for the variable ρ. We shall actually consider the non-conservative form of this equation: $\partial_t\rho + \mathbf{u} \cdot \nabla\rho = 0$, which is equivalent to the conservative form, for a regular solution and a divergence free velocity field. However, for more generality, we shall consider the transport equation with a non divergence free velocity.

Finally, the numerical methods we consider in this work are all of finite volume (FV) type. A general overview of these methods is given in [14]. These methods are very popular in various engineering fields, and in particular in computational fluid dynamics for various reasons, among which we may cite the following: they are based on local balance of very concrete physical quantities, like mass, momentum, energy, so that the conservativity properties of the continuous model are transferred to the discrete one; they may be designed to reproduce certain physical properties, like the maximum principle, which provides them with an acknowledged robustness; they may be used on very general meshes and are thus applicable on complex geometries.

The general idea of these methods is first to split the domain of computation (which may possibly be a space-time domain for time-dependent problems) into small volumes and associate one unknown per volume. Then the PDE is integrated over each control volume and volume integrals of spatial derivatives are transformed into integrals over the edges through the Green-Gauss formula. The difficulty is then to evaluate these so-called fluxes with the help of the unknowns of the scheme in a consistent way. This is particularly difficult when diffusive fluxes have to be evaluated, because one has to approach the gradient of the unknowns, rather than the unknowns themselves, and one thus has to go through some reconstruction step.

In our work, we focus on the traditional upwind scheme for the transport equation, and

on a more recent method, namely the discrete duality finite volume method (DDFV) for the Stokes and non-linear diffusion equations (although the theory developed here for the non-linear Darcy problem also applies to more general conservative methods). The DDFV schemes have been introduced to facilitate the reconstruction step of the gradient on general meshes when one has to evaluate diffusive fluxes, see [12, 18, 19].

Let us now give more details on the subjects treated in the chapters of this work.

In Chapter 1, we consider the Stokes equations in a two dimensional simply connected polygonal domain Ω with boundary Γ:

$$-\Delta\widehat{\mathbf{u}} + \nabla\widehat{p} = \mathbf{f} \text{ in } \Omega, \tag{5}$$

$$\nabla \cdot \widehat{\mathbf{u}} = 0 \text{ in } \Omega, \tag{6}$$

$$\widehat{\mathbf{u}} = \mathbf{g} \text{ on } \Gamma, \tag{7}$$

$$\int_\Omega \widehat{p}(x)dx = 0, \tag{8}$$

where $\widehat{\mathbf{u}}$ is the fluid velocity, $\widehat{p}$ is the pressure, $\mathbf{f}$ is the body forces per unit mass, and the function $\mathbf{g}$ satisfies $\int_\Gamma \mathbf{g}(\sigma)\cdot\mathbf{n}d\sigma = 0$. With $\mathbf{f} \in H^{-1}(\Omega)$ and $\mathbf{g} \in H^{1/2}(\Gamma)$, this problem is well-posed due to the so-called inf-sup condition: there exists $\beta > 0$ such that:

$$\beta = \inf_{q\in L^2_0(\Omega)} \sup_{\mathbf{v}\in(H^1_0(\Omega))^2} \frac{\int_\Omega q\nabla\cdot\mathbf{v}(\mathrm{x})d\mathrm{x}}{\|\mathbf{v}\|_{(H^1_0(\Omega))^2}\|q\|_{L^2(\Omega)}}. \tag{9}$$

Our purpose in this chapter is to compute an a posteriori error estimation between the exact solution $\widehat{\mathbf{u}}$, $\widehat{p}$ of (5)–(8) and its numerical approximation by the penalized discrete duality finite volume scheme (DDFV) as presented in [21].

Like for other equations, the development of a posteriori error estimations for the Stokes problem has followed the a priori investigation of numerical methods. As far as finite element methods are concerned, R. Verfürth [29] made one the very first contributions by getting two a posteriori error estimations for the mini-element discretization: one is based on a suitable evaluation of the residual, the other is based on the solution of local Stokes problems. Later on, R. Verfürth [30] generalized the first estimator developed in [29] to the nonconforming Crouzeix–Raviart finite element method, neglecting however the consistency error in the estimator. It was shown however in E. Dari et al. [10] that this consistency error may not always be neglected, and, in order to take it into account properly, they use a Helmholtz-Hodge like decomposition (adapted to the Stokes problem) of the velocity error. In the resulting error estimator, this gives rise to terms related to the jumps of the tangential velocity components from one cell to another, in addition to the usual jumps of the normal components of the stress tensor. The case of the non-conforming Fortin–Soulie quadratic elements is also treated.

All the above-cited finite element methods satisfy a uniform discrete inf-sup condition. However, it is often found useful in practice to consider discretizations (especially low-order ones) that do not verify a uniform discrete inf-sup condition. In this context, C. Bernardi et al. [5] consider the finite element approximation of the Stokes equations when a penalty

term is added to stabilize the variational formulation. The a posteriori error estimation they obtain includes two contributions: one related to the discretization on a given mesh, the other related to the penalty term. Based on these two contributions, the mesh refinement and the decrease of the penalty term are linked within an adaptive process.

A very recent contribution by A. Hannukainen et al. [17] sets a general framework for obtaining a posteriori error estimations for the discretization of the Stokes equations. The method is based on the reconstruction of postprocessed H_0^1 conforming velocity and $H - div$ conforming stress tensor fields deduced from the numerical approximation, and it may be applied to various conforming and conforming stabilized finite element methods, the discontinuous Galerkin method, the Crouzeix–Raviart nonconforming finite element method, the mixed finite element method, and a general class of finite volume methods.

However, as far as finite volume methods are concerned, the use of arbitrary meshes in [17] requires first to solve local Stokes problems on a conforming subtriangulation of each control volume, and then to apply the above-cited reconstruction on this subtriangulation. Instead, we would like to obtain error estimates for the solution of the DDFV scheme presented in [21] without having to solve any local problem or to compute any reconstruction. To do this, we shall adapt to the Stokes problem the a posteriori error estimation investigated in [27] for the DDFV discretization of the Laplace equation, using the discrete variational formulation verified by this scheme. The non-conformity of the method is dealt with using the the Helmholtz-Hodge like decomposition introduced in [10]. Our estimator also includes a term related to the stabilization term in the incompressibility equation.

In Chapter 2, we consider an approximate solution of the following nonlinear equation:

$$\begin{cases} -\mathrm{div}(H(\hat{u})\nabla\hat{u})(x) &= f(x), \quad x \in \Omega \\ \hat{u}(x) &= 0, \quad x \in \Gamma \end{cases} \tag{10}$$

under the following assumptions

A1. Let $H : \mathbb{R} \mapsto \mathbb{R}$ be continuous such that there exist constants C_1, $C_2 > 0$ satisfying

$$C_1 \leq H(\hat{u}) \leq C_2 \ \text{ for all } \hat{u} \in \mathbb{R}. \tag{11}$$

Moreover, we will assume that there exists a positive constant C such that

$$|H(\hat{u}) - H(\hat{v})| \leq C|\hat{u} - \hat{v}| \quad \text{ for all } \hat{u}, \hat{v} \in \mathbb{R}. \tag{12}$$

A2. Let f belong to $L^2(\Omega)$.

The discrete solution of the nonlinear diffusion equation is investigated by the mixed finite element method in [25], and, more generally, in [28]. More recently, finite volume discretizations have been developed for linear diffusion equations, such as finite volume schemes on admissible meshes [14], DDFV schemes [27] and multipoint flux approximation (MPFA) schemes [1] on arbitrary meshes, etc. Then we are interested in solving the nonlinear diffusion equations by such finite volume discretizations. In this work, we will deal with the above three methods. The discretization process of the problem leads to a system of

nonlinear equations, which is linearized by the fixed point method. These schemes have local conservative property, which is the important ingredient to obtain an a posteriori error estimation.

The theory of a posteriori error estimation is not very developed for the nonlinear diffusion equation. L. El Alaoui et al. [13] have obtained a posteriori error estimate by using a finite element method in the p-Laplace case. D. Kim et al. [20] give an estimate for the mixed finite element discretization. For the finite volume methods, although there are not many important results for the nonlinear diffusion equations, essential development steps on a posteriori error estimation have been achieved for the linear diffusion equations. Nicaise [26] gave an a posteriori error estimation for Morley-type interpolations of the original piecewise constant finite volume approximation. P. Omnes and al. [27] used the equivalence of the DDFV scheme with a finite element like method to derive fully computable a posteriori error estimate for this method applied to the Laplace equation. For local conservative methods, M. Vohralík [32] built an approximate function which depends only on the flux through the segments of the cells and on the values of the unknowns at some control points and provides error estimation with respect to this reconstruction. In addition, the effectivity index is very close to one, which demonstrates a good estimation. M. Vohralík's idea will be applied in this work to estimate the error of finite volume techniques applied to the non-linear diffusion equations.

Given a discrete solution, a stage of iterative process and a given mesh, our a posteriori error estimation is split into 2 terms: the discretization and linearization estimators. This splitting has two main advantages. The first one is that, in practice, when the number of iteration is large enough, the linearization estimator is negligible compared with the discretization one. Thus the balance of these two estimators is an important key to avoid performing an excessive number of nonlinear solver iterations. The other one is that, as a result, the mesh refinement is only based on the discretization estimator, since the linearization estimator is then negligible. This type of analysis is considered in [13] for the p-Laplace case. In this work, we don't deal with the convergence of the linearization and the discretization processes. We only mention them in our numerical experiments.

In Chapter 3, we consider the following transport equation in N space dimensions ($N \geq 1$), with initial condition:

$$\begin{cases} \frac{\partial u}{\partial t} + \mathbf{a}(\mathrm{x}, t) \cdot \nabla u & = 0, \ \forall x \in \mathbb{R}^N, \ t \in \mathbb{R}^+, \\ u(x, 0) & = u_0(x), \ \forall x \in \mathbb{R}^N. \end{cases} \tag{13}$$

We shall consider here, following [23], the unique entropy solution to (13).

The upwind scheme is a standard method to solve in a approximate way this problem, and, more generally, conservation laws (see [14]), but the a priori and a posteriori error estimations have been investigated only recently on scalar equations.

In the case of a conservation law with a divergence free flux function and no source term, C. Chainais-Hillairet [7] derived an a priori error estimate with a convergence order of $1/4$ in the L^1 norm. Based on [7], an a posteriori error estimation is achieved by D. Kröner and M. Ohlberger [22] for the upwind explicit scheme. This result is used to define an algorithm

with an adaptive grid for the finite volume scheme. Recently, in Mamaghani's Ph.D. thesis [24], an a posteriori error estimate is obtained for the implicit upwind finite volume scheme.

Otherwise, in [8], the a priori error estimation was also treated for a conservation law with a non divergence free flux function and with a source term.

In our work, we will deal with the transport equation with a non divergence free velocity field, written as a conservation law with the source term $\nabla \cdot \mathbf{a}(\mathrm{x}, t)u(\mathrm{x}, t)$, and we obtain an a posteriori error estimation for the explicit upwind scheme.

In Chapter 4, we consider the following two Poincaré inequalities: The Friedrichs (also called Poincaré) inequality

$$\int_\Omega u^2(x)dx \leq c_F \int_\Omega |\boldsymbol{\nabla} u(x)|^2 dx \quad , \quad \forall u \in H_0^1(\Omega) \tag{14}$$

and the Poincaré (also called mean Poincaré) inequality

$$\int_\Omega u^2(x)dx \leq c_P \int_\Omega |\boldsymbol{\nabla} u(x)|^2 dx \quad , \quad \forall u \in H^1(\Omega) \text{ such that } \int_\Omega u(x)dx = 0, \tag{15}$$

where c_F and c_P are constants depending only on Ω. These two inequalities play an important role in the theory of partial differential equations.

This chapter considers discrete versions of Poincaré inequalities for the DDFV method of discretization on arbitrary meshes, as presented, e.g., in [12].

The originality of these schemes is that they work well on all kind of meshes, including very distorted, degenerating, or highly nonconforming meshes (see the numerical tests in [12]). The name DDFV comes from the fact that these schemes are based on the definition of discrete gradient and divergence operators which verify a discrete Green formula.

In this introduction, let us only mention that in the DDFV discretization, scalar functions are discretized by their values both at the centers and at the vertices of a given mesh, and their gradients are evaluated on the so-called "diamond-cells" associated to the edges of the mesh. Each internal diamond-cell is a quadrilateral; its vertices are the two nodes of a given internal edge and the centers of the two cells which share this edge. Each boundary diamond cell is a degenerated quadrangle (i.e. a triangle); its vertices are the two nodes of a given boundary edge and the center of the corresponding cell and that of the boundary edge.

Then, the discrete version of the L^2 norm on the left-hand side of (14) and (15) is the half-sum of the L^2 norms of two piecewise constant functions, one defined with the discrete values given at the centers of the original ("primal" in what follows) cells, and the other defined with the discrete values given at the vertices of the primal mesh, to which we associate cells of a dual mesh. Moreover, the discrete version of the gradient L^2 norm on the right-hand side of (14) and (15) is the L^2 norm of the piecewise constant gradient vector field defined with it discrete values on the diamond-cells.

In the finite volume context, discrete Poincaré-Friedrich inequalities have previously been proved in [14, Lemma 9.1, Lemma 10.2] and [16], respectively for so-called "admissible" meshes (roughly speaking, meshes such that each edge is orthogonal to the segment joining the centers of the two cells sharing that edge, see the precise definition in [14, Definition 9.1]) and for Voronoi meshes. Similar results on duals of general simplicial triangulations are

proved in [31]. In the DDFV context, a discrete version of (14) is given for arbitrary meshes in [2]. However, the discrete constant c_F which appears in that paper depends on the mesh regularity in a rather intricate way, see [2, Formula (2.6) and Lemma 3.3].

The main result of our contribution is the proof of discrete versions of both (14) and (15) in the DDFV context, with constants c_F and c_P depending only on the domain and on the minimum angle in the diagonals of the diamond cells of the mesh.

Our proof of the discrete version of (14) is very similar to those given in [14] or [31]. We also prove a discrete version of (14) in a slightly more general situation when the domain is not simply connected and the discrete values of the function vanish only on the exterior boundary of the domain and are constant on each of the internal boundaries (this will have a subsequent application in the last section of the present chapter).

However, the task is more difficult for the mean-Poincaré inequality. Like in [14], it is divided into three steps. The first is the proof of this inequality on a convex subdomain; in the second, our proof differs from that in [14] because we actually do not need to prove a bound on the L^2 norm of the difference of discrete functions and their discrete mean value on the boundary of a convex subset, but rather an easier bound on the L^1 norm of this difference. The final step consists in dividing a general polygonal domain into several convex polygonal subdomains and in combining the first two steps to obtain the result.

As a consequence of these results, we derive a discrete equivalent of the following result (which is a particular case of a result given in [15]): Let us consider open, bounded, simply connected, convex polygonal domains $(\Omega_q)_{q\in[0,Q]}$ of $\mathbb{R}^2$ such that $\Omega_q \subset \Omega_0$ for all $q \in [1,Q]$ and $\bar{\Omega}_{q_1} \cap \bar{\Omega}_{q_2} = \emptyset$ for all $(q_1,q_2) \in [1,Q]^2$ with $q_1 \neq q_2$. Let Ω be defined by $\Omega = \Omega_0 \backslash (\cup_{q=1}^{Q} \Omega_q)$. Let us denote by $\Gamma = \partial\Omega = \cup_{q=0}^{Q} \Gamma_q$, with $\Gamma_q = \partial\Omega_q$ for all $q \in [0,Q]$. Then, there exists a constant C, depending only on Ω, such that for all vector field $\mathbf{u}$ in $H(\mathrm{div},\Omega) \cap H(\mathrm{rot},\Omega)$ with $\mathbf{u}\cdot\mathbf{n} = 0$ on Γ and $(\mathbf{u}\cdot\boldsymbol{\tau}, 1)_{\Gamma_q} = 0$ for all $q \in [1,Q]$, there holds

$$||\mathbf{u}||_{L^2(\Omega)} \leq C(||\nabla\cdot\mathbf{u}||_{L^2(\Omega)} + ||\nabla\times\mathbf{u}||_{L^2(\Omega)}). \tag{16}$$

The discrete equivalent has applications in the derivation of *a priori* error estimates for the DDFV method applied to the Stokes equations ([11]).

References for the Introduction

[1] I. Aavatsmark, *An introduction to multipoint flux approximation for quadrilateral girds*. Computational Geosciences 6, pp. 405–432 (2002).

[2] B. Andreianov, F. Boyer, F. Hubert, *Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes*. Numer. Methods Partial Differ. Equations 23 pp. 145–195 (2007).

[3] I. Babuska and W.C. Rheinboldt, *Error estimates for adaptive finite element computations*. SIAM J. Numer. Anal. 15, pp. 736–754, 1978.

[4] I. Babuska and W.C. Rheinboldt, *A posteriori error estimates for the finite element method*. Int. J. Numer. Meth. Engrg. 12, pp. 1597–1615, 1978.

[5] C. Bernardi, V. Girault, F. Hecht, *A Posteriori Analysis of a Penalty Method and Application to the Stokes Problem*. Mathematical Models and Methods in Applied Sciences, Vol. 13, No. 11, 1599–1628 (2003).

[6] A. Bourgeat, E. Marušić-Paloka, A. Mikelić, *Weak nonlinear corrections for Darcy's law*. Math. Models Methods Appl. Sci. 6, No.8, pp. 1143–1155, 1996.

[7] Chainais–Hillairet C., *Finite volume schemes for a nonlinear equation. Convergence towards the entropy solution and error estimation*. Model. Math. Anal. Numer. 33, 1, 129-156 (1999).

[8] C. Chainais–Hillairet, S. Champier, *Finite Volume for nonhomogeneous Scalar Conservation Laws: Error Estimate*. Numer. Math. vol 88 (4), pp. 607–639 (2001).

[9] S. Delcourte, K. Domelevo, P. Omnes, *A Discrete Duality Finite Volume Approach to Hodge Decomposition and Div–Curl Problems on Almost Arbitrary Two–Dimensional Meshes*. Siam J. Numer. Anal. Vol. 45, No.3, pp. 1142–1174 (2007).

[10] E. Dari, R. Durán and C. Padra, *Error Estimators for Nonconforming Finite Element Approximations of the Stokes Problem*. Math. Comp. 64, 1017-1033 (1995).

[11] S. Delcourte, P. Omnes, *A discrete duality finite volume discretization of the vorticity-velocity-pressure formulation of the 2D Stokes problem on almost arbitrary two-dimensional grids*. 2011.

[12] K. Domelevo, P. Omnes, *A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids.* M2AN Math. Model. Numer. Anal. 39 (6) pp. 1203–1249 (2005).

[13] L. El Alaoui, A. Ern, M. Vohralík, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems.* Comput. Methods Appl. Mech. Engrg., to appear (2011).

[14] R. Eymard, T. Gallouët, and R. Herbin, *Finite Volume methods.* In: Ciarlet, Ph.G. and Lions, J.-L., Eds. Handbook of Numerical Analysis, Vol. VII, Amsterdam: North-holland, pp. 713–1020 (2000).

[15] P. Fernandes, G. Gilardi, *Magnetostatic and electrostatic problems in inhomogeneous anisotropic media with irregular boundary and mixed boundary conditions.* Math. Models Methods Appl. Sci. 7, No.7, pp. 957–991 (1997).

[16] A. Glitzky, J. A. Griepentrog, *Discrete Sobolev-Poincaré inequalities for Voronoi finite volume approximations.* SIAM J. Numer. Anal. 48, pp. 372-391 (2010).

[17] A. Hannukainen, R. Stenberg, M. Vohralík *A Unified Framework for a Posteriori Error Estimation for the Stokes Problem.* Available at http://hal.archives-ouvertes.fr/hal-00470131/fr/.

[18] F. Hermeline, *A finite volume method for the approximation of diffusion operators on distorted meshes.* J. Comput. Phys. 160, pp. 481–499, 2000.

[19] F. Hermeline, *Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes.* Comput. Methods Appl. Mech. Eng. 192, pp. 1939–1959, 2003.

[20] D. Kim, E.J Park, *A priori and a posteriori analysis of mixed finite element methods for nonlinear elliptic equations.* SIAM J. Numer. Anal. Volume 48 (3), pp. 1186–1207 (2010)

[21] S. Krell, *Stabilized DDFV Schemes for Stokes Problem with Variable Viscosity on General 2D Meshes.* Available at http://hal.archives-ouvertes.fr/hal-00385687/fr/.

[22] Kröner D., Ohlberger M., *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in Multi Dimensions.* Math. Comput. 69, pp. 25–39 (1999).

[23] Kruskov S.N., *First order quasilinear equations with several space variable.* Math. USSR. Sb. 10, pp. 217–243, 1970.

[24] M. Mamaghani, *Suivi de Fronts par des Méthodes de Raffinement de Maillage Adaptatif et Application à la Simulation du Procédé de Récupération Steam Assisted Gravity Drainage.* Ph.D thesis (2010).

[25] F. A. Milner, *Mixed finite element methods for quasilinear second-order elliptic problems.* Math. Comp. 44, pp. 303–320 (1985).

[26] S. Nicaise, *A posteriori error estimations of some cell-centered finite volume methods.* SIAM J. Numer. Anal. 43, pp. 1481–1503 (2005).

[27] P. Omnes, Y. Penel, and Y. Rosenbaum, *A Posteriori Error Estimation for the Discrete Duality Finite Volume Discretization of the Laplace Equation.* SIAM J. Numer. Anal. 47, pp. 2782–2807 (2009).

[28] E. J. Park, *Mixed finite element methods for nonlinear second order elliptic problems.* SIAM J. Numer. Anal. v.32, pp. 865–885 (1995).

[29] R. Verfürth, *A Posteriori Error Estimation for the Stokes Equations.* Numer. Math. 55, pp. 309–325 (1989).

[30] R. Verfürth, *A Posteriori Error Estimation for the Stokes Equations II non-conforming discretizations.* Numer. Math. 60, pp. 235–249 (1991).

[31] M. Vohralík, On the Discrete Poincaré-Friedrichs Inequalities for Nonconforming Approximations of the Sobolev Space H^1, Numer. Funct. Anal. Optim. 26 (7-8), pp. 925–952 (2005).

[32] M. Vohralík, *Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods.* Numer. Math. 111, pp. 121–158 (2008).

Chapter 1

Stokes Equations

We derive an a posteriori error estimation for the discrete duality finite volume (DDFV) discretization of the stationary Stokes equations on very general twodimensional meshes, when a penalty term is added in the incompressibility equation to stabilize the variational formulation. Two different estimators are provided: one for the error on the velocity and one for the error on the pressure. They both include a contribution related to the error due to the stabilization of the scheme, and a contribution due to the discretization itself. The estimators are globally upper as well as locally lower bounds for the errors of the DDFV discretization. Numerical experiments illustrate the theoretical results and we especially consider the influence on the error of the penalty parameter for a fixed mesh and also of the mesh size for a fixed value of the penalty parameter.

1.1 Introduction

Let Ω be a two dimensional simply connected polygonal domain with boundary Γ. We consider the Stokes equations

$$-\Delta\widehat{\mathbf{u}} + \nabla\widehat{p} = \mathbf{f} \text{ in } \Omega, \tag{1.1}$$

$$\nabla \cdot \widehat{\mathbf{u}} = 0 \text{ in } \Omega, \tag{1.2}$$

$$\widehat{\mathbf{u}} = \mathbf{g} \text{ on } \Gamma, \tag{1.3}$$

$$\int_\Omega \widehat{p}(x)dx = 0, \tag{1.4}$$

where $\widehat{\mathbf{u}}$ is the fluid velocity, $\widehat{p}$ is the pressure, $\mathbf{f}$ is the body forces per unit mass, and the function $\mathbf{g}$ statifies $\int_\Gamma \mathbf{g}(\sigma)\cdot\mathbf{n}d\sigma = 0$. With $\mathbf{f} \in H^{-1}(\Omega)$ and $\mathbf{g} \in H^{1/2}(\Gamma)$, this problem is well-posed (see [5]) due to the so-called inf-sup condition: there exists $\beta > 0$ such that:

$$\beta = \inf_{q\in L^2_0(\Omega)} \sup_{\mathbf{v}\in(H^1_0(\Omega))^2} \frac{\int_\Omega q\nabla\cdot\mathbf{v}(\mathrm{x})d\mathrm{x}}{\|\mathbf{v}\|_{(H^1_0(\Omega))^2}\|q\|_{L^2(\Omega)}}. \tag{1.5}$$

Our purpose in this work is to compute an a posteriori error estimation between the exact solution $\widehat{\mathbf{u}}$, $\widehat{p}$ of (1.1)–(1.4) and its numerical approximation by the penalized discrete duality finite volume scheme (DDFV) as presented in [7].

Like for other equations, the development of a posteriori error estimations for the Stokes problem has followed the a priori investigation of numerical methods. As far as finite elements methods are concerned, R. Verfürth [10] made one of the very first contributions by getting two a posteriori error estimations for the mini-element discretization: one is based on a suitable evaluation of the residual, the other is based on the solution of local Stokes problems. Later on, R. Verfürth [11] generalized the first estimator developped in [10] to the nonconforming Crouzeix–Raviart finite element method, neglecting however the consitency error in the estimator. It was shown however in E. Dari et al. [4] that this consistency error may not always be neglected, and, in order to take it into account properly, they use a Helmholtz-Hodge like decomposition (adapted to the Stokes problem) of the velocity error. In the resulting error estimator, this gives rise to terms related to the jumps of the tangential velocity components from one cell to another, in addition to the usual jumps of the normal components of the stress tensor. The case of the non-conforming Fortin–Soulie quadratic elements is also treated.

All the above-cited finite element methods satisfy a uniform discrete inf-sup condition. However, it is often found useful in practise to consider discretizations (especially low-order ones) that do not verify a uniform discrete inf-sup condition. In this context, C. Bernardi et al. [1] consider the finite element approximation of the Stokes equations when a penalty term is added to stabilize the variational formulation. The a posteriori error estimation they obtain includes two contributions: one related to the discretization on a given mesh, the other related to the penalty term. Based on these two contributions, the mesh refinement and the decrease of the penalty term are linked within an adaptive process.

A very recent contribution by A. Hannukainen et al. [6] sets a general framework for obtaining a posteriori error estimations for the discretization of the Stokes equations. The method is based on the reconstruction of postprocessed H_0^1 conforming velocity and $H-div$ conforming stress tensor fields deduced from the numerical approximation, and it may be applied to various conforming and conforming stabilized finite element methods, the discontinous Galerkin method, the Crouzeix–Raviart nonconforming finite element method, the mixed finite element method, and general class of finite volume methods.

However, as far as finite volume methods are concerned, the use of arbitrary meshes in [6] requires first to solve local Stokes problems on a conforming subtriangulation of each control volume, and then to apply the above-cited reconstruction on this subtriangulation. Instead, we would like to obtain error estimates for the solution of the DDFV scheme presented in [7] without having to solve any local problem or to compute any reconstruction. To do this, we shall adapt to the Stokes problem the a posteriori error estimation investigated in [8] for the DDFV discretization of the Laplace equation, using the discrete variational formulation verified by this scheme. The non-conformity of the method is dealt with using the the Helmholtz-Hodge like decomposition introduced in [4]. Our estimator also includes a term related to the stabilization term in the incompressibility equation.

This chapter is organized as follows. Section 1.2 sets some notations and definitions related to the meshes, to discrete differential operators, and to discrete functions. In section 1.3, we present the DDVF scheme and its equivalent variational formula is recalled. In section 1.4, representations of the errors are elaborated. This is used in section 1.5 to find a computational upper bound of these errors. In section 1.6, the local efficiency of the error estimators is verified. Numerical experiments are presented in section 1.7.

1.2 Notations and definitions

Let Ω be covered by a primal mesh with polygonal cells denoted by T_i, $i \in [1, I]$, we associate a point G_i located in the interior of T_i. With any S_k, $k \in [1, K]$, we associate a dual cell P_k by joining points G_i associate with the primal cells surrounding S_k to the midpoints of the edges of which S_k is a node. the notations are summarized in Fig. 1.1 and 1.2.

With any primal edge A_j with $j \in [1, J]$, we associate a so-called diamond-cell D_j obtained by joining the vertices $S_{k_1(j)}$ and $S_{k_2(j)}$ of A_j to the points $G_{i_1(j)}$ and $G_{i_2(j)}$ associated with the primal cells that share A_j as a part of their boundaries. When A_j is a boundary edge (there are J^Γ such edges), the associated diamond-cell is a flat quadrilateral (i.e. a triangle) and we denote by $G_{i_2(j)}$ the midpoint of A_j (thus, there are J^Γ such additional points G_i). The unit normal vector to A_j is $\mathbf{n}_j$ and points from $G_{i_1(j)}$ to $G_{i_2(j)}$. We denote by A'_{j1} (resp. A'_{j2}) the segment joining $G_{i_1(j)}$ (resp. $G_{i_2(j)}$) and the midpoint of A_j. Its associated unit normal vector, pointing from $S_{k_1(j)}$ to $S_{k_2(j)}$, is denoted by $\mathbf{n}'_{j1}$ (resp. $\mathbf{n}'_{j2}$). In the case of a boundary diamond-cell, A'_{j2} reduces to $\{G_{i_2(j)}\}$ and does not play any role. Finally, for any diamond-cell D_j, we shall denote by $M_{i_\alpha k_\beta}$ the midpoint of $[G_{i_\alpha(j)} S_{k_\beta(j)}]$, with $(\alpha, \beta) \in \{1; 2\}^2$. With $\mathbf{n}_j$, $\mathbf{n}'_{j1}$ and $\mathbf{n}'_{j2}$, we associate orthogonal unit vectors $\boldsymbol{\tau}_j$, $\boldsymbol{\tau}'_{j1}$ and $\boldsymbol{\tau}'_{j2}$, such that the corresponding orthonormal bases are positively oriented. For any primal T_i such that $A_j \subset \partial T_i$, we shall define $\mathbf{n}_{ji} := \mathbf{n}_j$ if $i = i_1(j)$ and $\mathbf{n}_{ji} := -\mathbf{n}_j$ if $i = i_2(j)$, so that $\mathbf{n}_{ji}$ is always exterior to T_i. With $\mathbf{n}_{ji}$, we associate $\boldsymbol{\tau}_{ji}$ such that $(\mathbf{n}_{ji}, \boldsymbol{\tau}_{ji})$ is positively oriented. Similarly, when A'_{j1} and A'_{j2} belong to ∂P_k, we define $(\mathbf{n}'_{jk1}, \boldsymbol{\tau}'_{jk1})$ and $(\mathbf{n}'_{jk2}, \boldsymbol{\tau}'_{jk2})$ so that $\mathbf{n}'_{jk1}$ and $\mathbf{n}'_{jk2}$ are orthogonal to A'_{j1} and A'_{j2} and exterior to P_k.

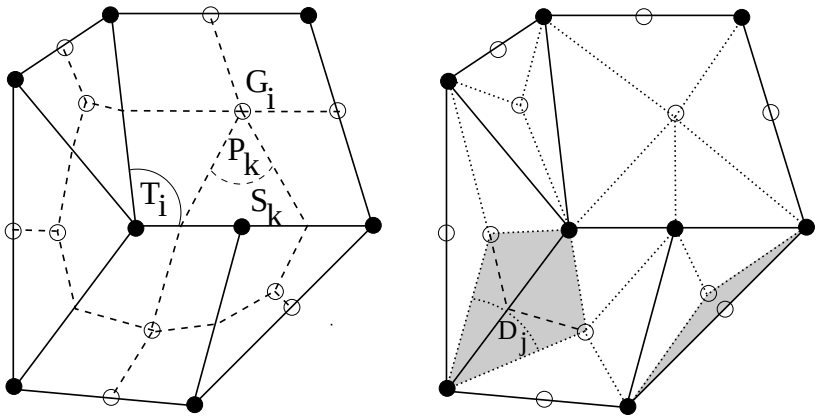

Figure 1.1: A nonconforming primal mesh and its associated dual mesh (left) and diamond mesh (right).

For $\mathbf{v} \in (H^2(\Omega))^2$ with $\mathbf{v} = (v_1, v_2)^t$, we define

$$\nabla \mathbf{v} = \begin{pmatrix} \partial v_1/\partial x & \partial v_1/\partial y \\ \partial v_2/\partial x & \partial v_2/\partial y \end{pmatrix}, \qquad \nabla \times \mathbf{v} = \begin{pmatrix} \partial v_1/\partial y & -\partial v_1/\partial x \\ \partial v_2/\partial y & -\partial v_2/\partial x \end{pmatrix},$$

$$\Delta \mathbf{v} = \begin{pmatrix} \Delta v_1 \\ \Delta v_2 \end{pmatrix}.$$

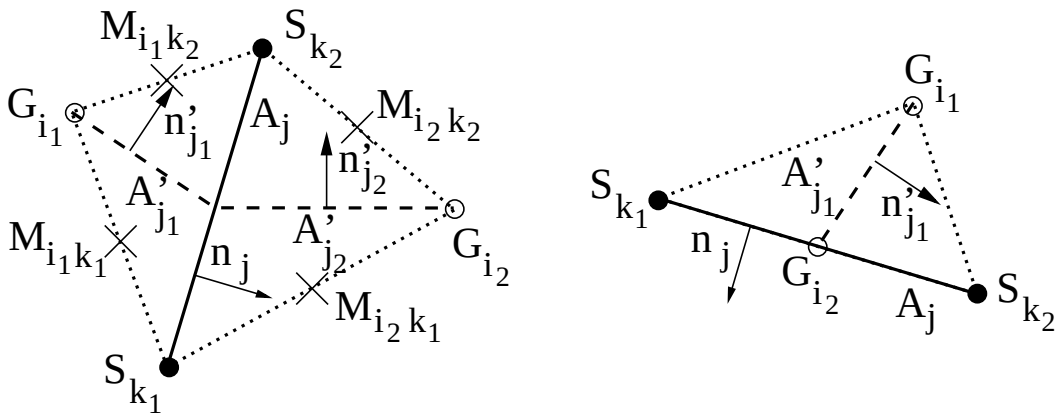

Figure 1.2: Notations for the inner diamond-cell (left) and a boundary diamond mesh (right).

If A and B are two matrices with dimension M, we define the inner product

$$A : B = \sum_{i,j=1}^{M} A_{ij} B_{ij}.$$

For future use, we recall Green's formulae

$$\int_\Omega \Delta \mathbf{v} \cdot \mathbf{w} dx = -\int_\Omega \nabla \mathbf{v} : \nabla \mathbf{w} + \int_{\partial\Omega} (\nabla \mathbf{v}\ \mathbf{n}) \cdot \mathbf{w} ds, \tag{1.6}$$

$$\int_\Omega \nabla \mathbf{v} : \nabla \times \mathbf{w} dx = -\int_{\partial\Omega} (\nabla \mathbf{v} \boldsymbol{\tau}) \cdot \mathbf{w} ds, \tag{1.7}$$

for any $\mathbf{v} \in (H^2(\Omega))^2$ and $\mathbf{w} \in (H^1(\Omega))^2$. Here, $\mathbf{n}$ is the outward normal to $\partial\Omega$ and $\boldsymbol{\tau}$ is the tangent vector to $\partial\Omega$ such that $(\mathbf{n}, \boldsymbol{\tau})$ is positively oriented.

In the definition of the DDFV scheme, we shall associate the velocity unknowns to the points G_i and S_k and the pressure unknowns to the diamond-cells. Moreover the gradient and divergence of the velocity will be defined on the diamond-cells. This leads us to the following definitions.

Definition 1.1. *Let* $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P)$, *and* $\mathbf{v} = (\mathbf{v}_i^T, \mathbf{v}_k^P)$ *be in* $(\mathbb{R}^2)^I \times (\mathbb{R}^2)^K$. *Let* $\Phi = (\Phi_j)$ *and* $\Psi = (\Psi_j)$ *be in* $(\mathbb{R}^{2\times 2})^J$. *And let* $p = (p)_j$ *and* $q = (q)_j$ *be in* $\mathbb{R}^J$. *We define the following scalar products*

$$(\mathbf{u}, \mathbf{v})_{T,P} := \frac{1}{2} \left(\sum_{i\in[1,I]} |T_i| \mathbf{u}_i^T \cdot \mathbf{v}_i^T + \sum_{k\in[1,K]} |P_k| \mathbf{u}_k^P \cdot \mathbf{v}_k^P \right), \tag{1.8}$$

$$(\Phi, \Psi)_D := \sum_{j\in[1,J]} |D_j| \Phi_j : \Psi_j, \qquad (p, q)_D = \sum_{j\in[1,J]} |D_j| p_j q_j. \tag{1.9}$$

Definition 1.2. *Let* $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P)$ *be in* $(\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$. *For any boundary edge* A_j, *with the notation of Fig 2.1, we define* $\tilde{\mathbf{u}}_j$ *as the trace of* $\mathbf{u}$ *over* A_j *by*

$$\tilde{\mathbf{u}}_j = \frac{1}{4}\left(\mathbf{u}^P_{k_1(j)} + 2\mathbf{u}^T_{i_2(j)} + \mathbf{u}^P_{k_2(j)}\right). \tag{1.10}$$

Let $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P)$ *be in* $(\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ *and let* $\mathbf{w} = (\mathbf{w}_j)$ *be defined on the boundary* Γ. *We define the following boundary scalar products*

$$(\mathbf{w}, \tilde{\mathbf{u}})_{\Gamma_h} := \sum_{j\in\Gamma} |A_j| \mathbf{w}_j \cdot \tilde{\mathbf{u}}_j. \tag{1.11}$$

Definition 1.3. *Let* $\Phi = (\Phi_j)$ *be in* $(\mathbb{R}^{2\times 2})^J$. *We define divergence and curl of the tensor field* Φ *on the primal and dual cells by*

$$\begin{aligned}
(\nabla_h^T \cdot \Phi)_i &:= \frac{1}{|T_i|} \sum_{j\in\partial T_i} |A_j| \Phi_j \mathbf{n}_{ji},\\
(\nabla_h^P \cdot \Phi)_k &:= \frac{1}{|P_k|}\left(\sum_{j\in\partial P_k} (|A'_{j_1}| \Phi_j \mathbf{n'}_{j_1} + |A'_{j_2}| \Phi_j \mathbf{n'}_{j_2}) + \sum_{j\in\partial P_k\cap\Gamma} \frac{|A_j|}{2} \Phi_j \mathbf{n}_j\right),\\
(\nabla_h^T \times \Phi)_i &:= \frac{1}{|T_i|} \sum_{j\in\partial T_i} |A_j| \Phi_j \boldsymbol{\tau}_{ji},\\
(\nabla_h^P \times \Phi)_k &:= \frac{1}{|P_k|}\left(\sum_{j\in\partial P_k} (|A'_{j_1}| \Phi_j \boldsymbol{\tau}'_{j_1} + |A'_{j_2}| \Phi_j \boldsymbol{\tau}'_{j_2}) + \sum_{j\in\partial P_k\cap\Gamma} \frac{|A_j|}{2} \Phi_j \boldsymbol{\tau}_j\right),
\end{aligned}$$

where the unit vectors $\boldsymbol{\tau}_j$ *(respectively* $\boldsymbol{\tau}_{ji}$, $\boldsymbol{\tau}'_{j1k}$, $\boldsymbol{\tau}'_{j2k}$*) are such that (*$\mathbf{n}_j$,$\boldsymbol{\tau}_j$*) (resp. (*$\mathbf{n}_{ji}$,$\boldsymbol{\tau}_{ji}$*), (*$\mathbf{n}'_{j1k}$,$\boldsymbol{\tau}'_{j1k}$*), (*$\mathbf{n}'_{j2k}$,$\boldsymbol{\tau}'_{j2k}$*)) are orthonormal positively oriented bases of* $\mathbb{R}^2$.

Definition 1.4. *Let* $\mathbf{u} = (u_1, u_2)$ *be in* $(\mathbb{R}^2)^{I+K+J^\Gamma}$ *with* $u_1 = ((u_1)_i^T, (u_1)_k^P) \in \mathbb{R}^{I+K+J^\Gamma}$ *and* $u_2 = ((u_2)_i^T, (u_2)_k^P) \in \mathbb{R}^{I+K+J^\Gamma}$, *the discrete gradient* $\nabla_h^D \mathbf{u}$ *and the discrete curl* $\nabla_h^D \times \mathbf{u}$ *are defined by their values in the diamond-cells* D_j *by*

$$(\nabla_h^D \mathbf{u})_j = \begin{pmatrix} (\nabla_h^D u_1)^t{}_j \\ (\nabla_h^D u_2)^t{}_j \end{pmatrix}, \qquad (\nabla_h^D \times \mathbf{u})_j = \begin{pmatrix} (\nabla_h^D \times u_1)^t{}_j \\ (\nabla_h^D \times u_2)^t{}_j \end{pmatrix},$$

where, for $\phi \in \mathbb{R}^{I+K+J^\Gamma}$, *we define*

$$\begin{aligned}
(\nabla_h^D \phi)_j &:= \frac{1}{2|D_j|}\left\{[\phi^P_{k_2} - \phi^P_{k_1}](|A'_{j_1}| \mathbf{n'}_{j_1} + |A'_{j_2}| \mathbf{n'}_{j_2}) + [\phi^T_{i_2} - \phi^T_{i_1}]|A_j| \mathbf{n}_j\right\},\\
(\nabla_h^D \times \phi)_j &:= \frac{1}{2|D_j|}\left\{[\phi^P_{k_2} - \phi^P_{k_1}](|A'_{j_1}| \boldsymbol{\tau}'_{j_1} + |A'_{j_2}| \boldsymbol{\tau}'_{j_2}) + [\phi^T_{i_2} - \phi^T_{i_1}]|A_j| \boldsymbol{\tau}_j\right\}.
\end{aligned}$$

From basic geometrical arguments, we obtain some properties of the discrete gradient:

$$\phi^P_{k_2} - \phi^P_{k_1} = \nabla_h^D \phi \cdot \overrightarrow{S_{k_1} S_{k_2}}, \tag{1.12}$$

$$\phi^T_{i_2} - \phi^T_{i_1} = \nabla_h^D \phi \cdot \overrightarrow{G_{i_1} G_{i_2}}. \tag{1.13}$$

We also need a discrete divergence of a vector, which is defined using the discrete gradient

$$(\nabla_h^D \cdot \mathbf{u})_j = \mathrm{Trace}((\nabla_{\mathrm{h}}^{\mathrm{D}} \mathbf{u})_{\mathrm{j}}).$$

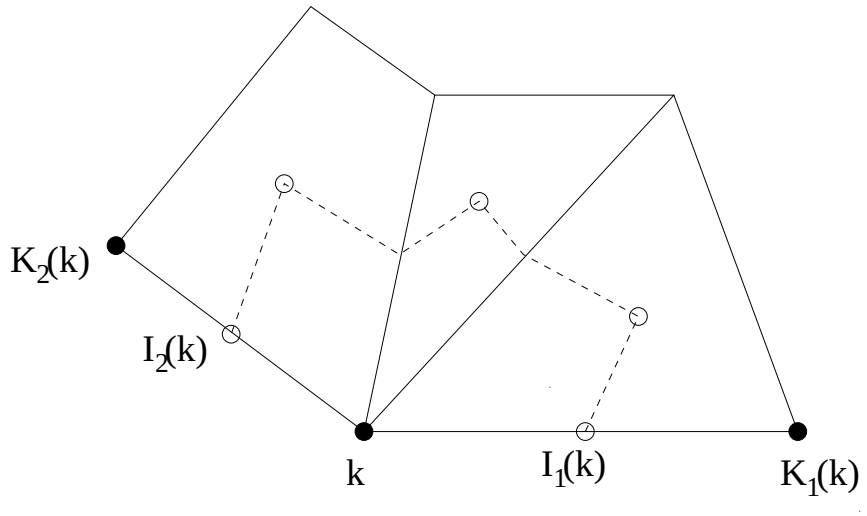

Figure 1.3: Notation for a boundary dual cell in formula (1.22).

For the penalization of the scheme, we need to define the following Laplacian-type operator.

Definition 1.5. *Let $p = (p_j) \in \mathbb{R}^J$, we define:*

$$(\Delta_h^D p)_j = \frac{1}{|D_j|} \sum_{j' \in \partial D_j} \frac{d_j^2 + d_{j'}^2}{d_j^2}(p_{j'} - p_j), \tag{1.14}$$

where ∂D_j is the set of diamond cells which have a common segment with D_j, $d_j = diam(D_j)$ and $d_{j'} = diam(D_{j'})$. In addition, we construct piecewise constant functions corresponding to the approximate pressure and to the penalty term:

$$p_h(\mathrm{x}) = p_j, \quad \forall\ \mathrm{x} \in D_j,\ j \in [1, J], \tag{1.15}$$

$$(\Delta_h^D p)_h(\mathrm{x}) = (\Delta_h^D p)_j, \quad \forall\ \mathrm{x} \in D_j, j \in [1, J]. \tag{1.16}$$

Proposition 1.6. *For $\Phi \in (\mathbb{R}^{2\times 2})^J$ and $= (\mathbf{u}^T, \mathbf{u}^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ and $p \in \mathbb{R}^J$, the following discrete Green formula hold:*

$$(\nabla_h^{T,P} \cdot \Phi, \mathbf{u})_{T,P} = -(\nabla_h^D \mathbf{u}, \Phi)_D + (\Phi \mathbf{n}, \tilde{\mathbf{u}})_{\Gamma,h}, \tag{1.17}$$

$$(\nabla_h^{T,P} \times \Phi, \mathbf{u})_{T,P} = (\nabla_h^D \times \mathbf{u}, \Phi)_D + (\Phi \boldsymbol{\tau}, \tilde{\mathbf{u}})_{\Gamma,h}, \tag{1.18}$$

$$(\nabla_h^{T,P} \cdot pI_2, \mathbf{u})_{T,P} = -(\nabla_h^D \cdot \mathbf{u}, p)_D + (p\mathbf{n}, \tilde{\mathbf{u}})_{\Gamma,h}, \tag{1.19}$$

where I_2 is the 2×2 identity matrix. The formula (1.17) is called discrete Stokes formula and is proved in [7]; it can also be found in [3]. The formulae (1.18) and (1.19) can be demonstrated in the same way.

Next proposition may be found in [3].

Proposition 1.7. *For all* $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$, *there holds*

$$(\nabla_h^T \times (\nabla_h^D \mathbf{u}))_i = 0 \quad \forall i \in [1, I], \tag{1.20}$$

$$(\nabla_h^P \times (\nabla_h^D \mathbf{u}))_k = 0 \quad \forall k \notin \Gamma. \tag{1.21}$$

In addition, for $k \in \Gamma$, *the following equality holds* (*see Fig. 1.3 for the notations*)

$$(\nabla_h^P \times (\nabla_h^D \mathbf{u}))_k = \frac{1}{|P_k|}\left[(\mathbf{u}_{I_2(k)}^T - \mathbf{u}_{I_1(k)}^T) + \frac{1}{2}(\mathbf{u}_{K_1(k)}^P - \mathbf{u}_{K_2(k)}^P)\right]. \tag{1.22}$$

Definition 1.8. *With any* $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$, *we associate the function* $\mathbf{u}_h$ *defined by*

$$(\mathbf{u}_h)|_{D_j} \in (P^1(D_j))^2 \ \forall j \in [1, J],$$
$$\mathbf{u}_h(M_{i_\alpha(j) k_\beta(j)}) = \frac{1}{2}(\mathbf{u}_{i_\alpha(j)}^T + \mathbf{u}_{k_\beta(j)}^P) \ \forall j \in [1, J], \ (\alpha, \beta) \in \{1, 2\}^2.$$

Proposition 1.9. *Let* $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ *and let* $\mathbf{u}_h$ *be defined by Definition 1.8. There holds*

$$(\nabla_h^D \mathbf{u})_j = \nabla \mathbf{u}_h|_{D_j}, \quad \forall\ j \in [1, J], \tag{1.23}$$

$$(\nabla_h^D \cdot \mathbf{u})_j = \nabla \cdot \mathbf{u}_h|_{D_j}, \quad \forall\ j \in [1, J]. \tag{1.24}$$

Definition 1.10. *Let* $\mathbf{u}_h$ *be in* $(P^1(D_j))^2, \ \forall j \in [1, J]$, *and not necessarily continous over the interfaces of neighboring diamond-cells. We define its piecewise gradient and divergence over* Ω *by:*

$$\nabla_h \mathbf{u}_h(\mathrm{x}) = \nabla \mathbf{u}_h|_{D_j}(\mathrm{x}) \text{ and } \nabla_h \cdot \mathbf{u}_h(\mathrm{x}) = \nabla \cdot \mathbf{u}_h|_{D_j}(\mathrm{x}), \ \forall \mathrm{x} \in D_j, \ j \in [1, J]. \tag{1.25}$$

1.3 The finite volume scheme on general meshes

We recall the finite volume scheme used for the numerical approximation of equations (1.1)–(1.4). This scheme is constructed on the basis of the discrete differential operators defined in section 1.2.

$$(\nabla_h^T \cdot (-\nabla_h^D \mathbf{u} + pI_2))_i = \mathbf{f}_i^T \ , \quad \forall i \in [1, I], \tag{1.26}$$

$$(\nabla_h^P \cdot (-\nabla_h^D \mathbf{u} + pI_2))_k = \mathbf{f}_k^P \ , \quad \forall k \in [1, K], \tag{1.27}$$

$$(\nabla_h^D \cdot \mathbf{u})_j + \varepsilon(\Delta_h^D p)_j = 0 \ , \quad \forall j \in [1, J], \tag{1.28}$$

$$\frac{\mathbf{u}_{k_1(j)}^P + 2\mathbf{u}_{i_2(j)}^T + \mathbf{u}_{k_2(j)}^P}{4} = \mathbf{g}_j \ , \quad \forall j \in \Gamma, \tag{1.29}$$

$$\sum_{j=1}^{J} |D_j| p_j = 0, \tag{1.30}$$

with an appropriate choice of $\varepsilon > 0$. We suppose that $\mathbf{g}$ is regular enough, so that we can set $\mathbf{g}_j = \mathbf{g}(G_{i_2(j)})$ in (1.29), while in (1.26) and (1.27), $\mathbf{f}_i^T$ and $\mathbf{f}_k^P$ are the mean values of $\mathbf{f}$ over T_i and P_k, respectively:

$$\mathbf{f}_i^T = \frac{1}{|T_i|}\int_{T_i} \mathbf{f}(\mathrm{x})d\mathrm{x} \quad \text{and} \quad \mathbf{f}_k^P = \frac{1}{|P_k|}\int_{P_k} \mathbf{f}(\mathrm{x})d\mathrm{x}. \tag{1.31}$$

In [7], it is proved that if $\varepsilon > 0$, then the scheme (1.26)–(1.30) has a unique solution.

Proposition 1.11. *Let $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P)$ and $p = (p_j)_{j\in[1,J]}$ be the solution of the scheme (1.26)-(1.30). Let $\mathbf{v} = (\mathbf{v}_i^T, \mathbf{v}_k^P)$ such that $\tilde{\mathbf{v}}_j = 0$ for all $j \in \Gamma$. Let $\mathbf{u}_h$ and $\mathbf{v}_h$ be the solution associated to $\mathbf{u}$ and $\mathbf{v}$ by Definition 1.8. Let us set in addition*

$$\mathbf{v}_h^*(\mathrm{x}) := \frac{1}{2}\left(\sum_{i\in[1,I]} \mathbf{v}_i^T\theta_i^T(\mathrm{x}) + \sum_{k\in[1,K]} \mathbf{v}_k^P\theta_k^P(\mathrm{x})\right), \tag{1.32}$$

where θ_i^T and θ_k^P are respectively the characteristic function of the cells T_i and P_k. Then, there holds

$$\sum_j \int_{D_j} \nabla_h\mathbf{u}_h : \nabla_h\mathbf{v}_h(\mathrm{x})d\mathrm{x} - \sum_j \int_{D_j} \nabla_h\cdot\mathbf{v}_h\, p_h(\mathrm{x})d\mathrm{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}_h^*(\mathrm{x})d\mathrm{x}. \tag{1.33}$$

Proof. We can apply Eq. (1.26) and (1.27), we have

$$-(\nabla_h^T\cdot(\nabla_h^D\mathbf{u})_i\cdot\mathbf{v}_i^T + (\nabla_h^T\cdot(pI_2)_i)\cdot\mathbf{v}_i^T = \mathbf{f}_i^T\cdot\mathbf{v}_i^T \quad \forall i\in[1,I], \tag{1.34}$$

$$-(\nabla_h^P\cdot(\nabla_h^D\mathbf{u}))_i\cdot\mathbf{v}_k^P + (\nabla_h^P\cdot(pI_2))_k\cdot\mathbf{v}_k^P = \mathbf{f}_k^P\cdot\mathbf{v}_k^P \quad \forall k\in[1,K]. \tag{1.35}$$

Multiplying (1.34) by $|T_i|$ and (1.35) by $|P_k|$ and summing over all i and all k, we obtain

$$-(\nabla_h^{T,P}\cdot(\nabla_h^D\mathbf{u}),\mathbf{v})_{T,P} + (\nabla_h^{T,P}\cdot(pI_2),\mathbf{v})_{T,P} = (\mathbf{f},\mathbf{v})_{T,P}.$$

We can apply (1.17), (1.19) and $\tilde{\mathbf{v}}_j = 0$ for all $j\in\Gamma$, we obtain

$$(\nabla_h^D(\mathbf{u}),\nabla_h^D(\mathbf{v}))_D - (\nabla_h^D\cdot\mathbf{v},p)_D = (\mathbf{f},\mathbf{v})_{T,P},$$

or, using the property in (1.23), we have

$$\sum_j \int_{D_j} \nabla_h\mathbf{u}_h : \nabla_h\mathbf{v}_h(\mathrm{x})d\mathrm{x} - \sum_j \int_{D_j} \nabla_h\cdot\mathbf{v}_h\, p_h(\mathrm{x})d\mathrm{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}_h^*(\mathrm{x})d\mathrm{x}.$$

□

1.4 A representation of the error

1.4.1 A representation of the velocity error

The variational formula of (1.1) reads:

$$\int_\Omega \nabla\widehat{\mathbf{u}} : \nabla\mathbf{v}d\mathrm{x} - \int_\Omega \widehat{p}\nabla\cdot\mathbf{v}d\mathrm{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}(\mathrm{x})d\mathrm{x}, \tag{1.36}$$

for all $\mathbf{v} \in (H_0^1(\Omega))^2$. We shall estimate the H^1 semi norm of the error between the exact solution $\widehat{\mathbf{u}}$ and the function $\mathbf{u}_h$ associated to the solution of the DDFV scheme. For this, We shall denote by $\mathbf{e} = \widehat{\mathbf{u}} - \mathbf{u}_h$ and $e_p = \widehat{p} - p_h$ the error in the velocity and pressure, respectively. We have

$$\|\nabla_h \mathbf{e}\|_{L^2(\Omega)} = \left(\sum_j \int_{D_j} |\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h|^2(\mathrm{x}) d\mathrm{x} \right)^{1/2}. \tag{1.37}$$

Since Ω is a simply connected domain and since $\nabla_h \mathbf{e} = \nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h$ belongs to $(L^2(\Omega))^{2\times 2}$, we may decompose it in the following way (see Lemma 3.2 in [4]):

$$\nabla_h \mathbf{e} = \nabla \widehat{\Phi} - q I_2 + \nabla \times \widehat{\Psi}, \tag{1.38}$$

where $q \in L_0^2(\Omega)$, $\widehat{\Phi} \in (H_0^1(\Omega))^2$ with $\nabla \cdot \widehat{\Phi} = 0$ and $\widehat{\Psi} \in (H^1(\Omega))^2$ with integral zero satisfying

$$\begin{aligned} \|\nabla \widehat{\Phi}\|_{L^2(\Omega)} &\leq \|\nabla_h e\|_{L^2(\Omega)}, \\ \|q\|_{L^2(\Omega)} &\leq \frac{2}{\beta} \|\nabla_h e\|_{L^2(\Omega)}, \\ \|\nabla \times \widehat{\Psi}\|_{L^2(\Omega)} &\leq (1 + \frac{2\sqrt{2}}{\beta}) \|\nabla_h e\|_{L^2(\Omega)}, \end{aligned} \tag{1.39}$$

where β is defined by (1.5).

Now, we estimate the velocity error using the decomposition (1.38). First observe that

$$\int_\Omega \nabla_h \mathbf{e} : I_2 q(\mathrm{x}) d\mathrm{x} = \int_\Omega \nabla_h \cdot \mathbf{e}\, q(\mathrm{x}) d\mathrm{x} = \int_\Omega (\nabla \cdot \widehat{\mathbf{u}} - \nabla_h \cdot \mathbf{u}_h) q(\mathrm{x}) d\mathrm{x}.$$

From (1.2) and (1.28), we have

$$\int_\Omega \nabla_h \mathbf{e} : I_2 q(\mathrm{x}) d\mathrm{x} = \varepsilon \int_\Omega (\Delta_h^D p)_h q(\mathrm{x}) d\mathrm{x}. \tag{1.40}$$

Multiplying the term $\nabla_h \mathbf{e}(\mathrm{x})$ with (1.38) side by side and integrating over Ω, there holds

$$\begin{aligned} \|\nabla_h \mathbf{e}\|_{L^2(\Omega)}^2 &= \int_\Omega \nabla_h \mathbf{e} : (\nabla \widehat{\Phi} + \nabla \times \hat{\Psi} - q I_2) d\mathrm{x} \\ &= i_1 + i_2 - \varepsilon \int_\Omega (\Delta_h^D p)_h q(\mathrm{x}) d\mathrm{x}, \end{aligned} \tag{1.41}$$

where

$$i_1 = \sum_j \int_{D_j} (\nabla \widehat{\mathbf{u}} - \nabla \mathbf{u}_h) : \nabla \hat{\Phi}(\mathrm{x}) d\mathrm{x}$$

and

$$i_2 = \sum_j \int_{D_j} (\nabla \widehat{\mathbf{u}} - \nabla \mathbf{u}_h) : \nabla \times \hat{\Psi}(\mathrm{x}) d\mathrm{x}.$$

In order to find a suitable representation of i_1 and i_2, we shall need the following definitions

Definition 1.12. *The boundary ∂D_j of any diamond-cell D_j is composed of the four segments $[G_{i_\alpha(j)}S_{k_\beta(j)}]$ with $(\beta,\alpha)\in\{1,2\}$ (see Fig. 1.2). Let us define by S the set of these edges when j runs over the whole set of diamond-cells and $\overset{\circ}{S}$ those edges that do not lie in the boundary Γ. Each $s\in\overset{\circ}{S}$ is thus a segment that we shall denote by $[G_{i(s)}S_{k(s)}]$. We shall also write $s\in\overset{\circ}{T}_i$ (resp. $s\in\overset{\circ}{P}_k$) if $s\subset T_i$ ($s\subset P_k$) and $s\not\subset\Gamma$. Finally, we shall denote by $\mathbf{n}_s$ one of the two unit normal vectors to s, arbitrarily chosen among the two possible choices but then fixed for the sequel, and $[(\nabla\mathbf{u}_h-p_hI_2)\mathbf{n}_s]_s$ is the jump of the normal component of $\nabla\mathbf{u}_h-p_hI_2$ through segment s.*

Proposition 1.13. *Let $\widehat{\Phi}$ be defined in equation (1.38). Let $\Phi=(\Phi_i^T,\Phi_k^P)\in(\mathbb{R}^2)^{I+J^\Gamma}\times(\mathbb{R}^2)^K$ be such that*

$$\tilde{\Phi}_j=0 \text{ for all } j\in\Gamma. \tag{1.42}$$

Then, it holds that

$$\begin{aligned} i_1=&\frac{1}{2}\sum_{i\in[1,I]}\int_{T_i}\mathbf{f}\cdot(\widehat{\Phi}-\Phi_i^T)(\mathbf{x})d\mathbf{x}+\frac{1}{2}\sum_{k\in[1,K]}\int_{P_k}\mathbf{f}\cdot(\widehat{\Phi}-\Phi_k^P)(\mathbf{x})d\mathbf{x}\\ &-\frac{1}{2}\sum_{i\in[1,I]}\sum_{s\subset\overset{\circ}{T}_i}\int_s[(\nabla_h\mathbf{u}_h-p_hI_2)\mathbf{n}_s]_s\cdot(\widehat{\Phi}-\Phi_i^T)(\sigma)d\sigma\\ &-\frac{1}{2}\sum_{k\in[1,K]}\sum_{s\subset\overset{\circ}{P}_k}\int_s[(\nabla_h\mathbf{u}_h-p_hI_2)\mathbf{n}_s]_s\cdot(\widehat{\Phi}-\Phi_k^P)(\sigma)d\sigma. \end{aligned} \tag{1.43}$$

Proof. First, Since $\widehat{\Phi}\in(H_0^1(\Omega))^2$ and $\nabla\cdot\widehat{\Phi}=0$, we can apply (1.36), we have

$$\begin{aligned} i_1&=\sum_j\int_{D_j}\nabla\widehat{\mathbf{u}}:\nabla\hat{\Phi}(\mathrm{x})d\mathrm{x}-\sum_j\int_{D_j}\nabla_h\mathbf{u}_h:\nabla\hat{\Phi}(\mathrm{x})d\mathrm{x}\\ &=\int_\Omega\nabla\widehat{\mathbf{u}}:\nabla\hat{\Phi}(\mathrm{x})d\mathrm{x}-\sum_j\int_{D_j}\nabla_h\mathbf{u}_h:\nabla\hat{\Phi}(\mathrm{x})d\mathrm{x}\\ &=\int_\Omega\mathbf{f}\cdot\widehat{\Phi}(\mathrm{x})d\mathrm{x}-\sum_j\int_{D_j}\nabla_h\mathbf{u}_h:\nabla\widehat{\Phi}(\mathrm{x})d\mathrm{x}. \end{aligned}$$

For any $\Phi=(\Phi_i^T,\Phi_k^P)$ satisfying (1.42), formula (1.33) leads to

$$\begin{aligned} i_1=&\int_\Omega\mathbf{f}\cdot(\Phi-\Phi_h^*)(\mathrm{x})d\mathrm{x}-\sum_j\int_{D_j}p_h\nabla_h\cdot\Phi_h(\mathrm{x})d\mathrm{x}\\ &-\sum_j\int_{D_j}\nabla_h\mathbf{u}_h:(\nabla\widehat{\Phi}-\nabla_h\Phi_h)(\mathrm{x})d\mathrm{x}. \end{aligned}$$

We know that $p_h\nabla_h\cdot\Phi_h=p_hI_2:\nabla_h\Phi_h$ and $p_hI_2:\nabla\widehat{\Phi}=0$ (since $\nabla\cdot\widehat{\Phi}=0$), then

$$i_1=\int_\Omega\mathbf{f}\cdot(\Phi-\Phi_h^*)(\mathrm{x})d\mathrm{x}-\sum_{j=1}^J H_1(j), \tag{1.44}$$

where

$$H_1(j) = \int_{D_j} (\nabla_h \mathbf{u}_h - p_h I_2) : (\nabla \widehat{\Phi} - \nabla_h \Phi_h)(\mathrm{x}) d\mathrm{x}. \tag{1.45}$$

Let us consider a diamond-cell D_j. Since $\nabla_h \mathbf{u}_h - p_h I_2$ is constant over D_j, we may write, using Green's formula over D_j,

$$H_1(j) = \int_{\partial D_j} (\nabla_h \mathbf{u}_h - p_h I_2) \mathbf{n}_{\partial D_j} \cdot (\widehat{\Phi} - \Phi_h)(\sigma) d\sigma,$$

where $\mathbf{n}_{\partial D_j}$ is the unit normal vector exterior to D_j on its boundary. Moreover, let each s be one of the four boundary edges of D_j, the function Φ_h belongs to P^1 over s and the quantity $\nabla \mathbf{u}_h - p_h I_2$ is a constant; the integral of $(\nabla \mathbf{u}_h - p_h I_2)\mathbf{n}_{\partial D_j} \cdot \Phi_h$ along this edge may thus exactly be computed by the midpoint rule; using the definition of Φ_h, this function equals $\frac{1}{2}(\Phi^T_{i(s)} + \Phi^T_{k(s)})$ at the midpoint of s. There holds:

$$H_1(j) = \sum_{s \in \partial D_j} \int_s (\nabla_h \mathbf{u}_h - p_h I_2) \mathbf{n}_{s,j} \cdot \left[\widehat{\Phi} - \frac{1}{2} \left(\Phi^T_{i(s)} + \Phi^P_{k(s)} \right) \right] (\sigma) d\sigma, \tag{1.46}$$

where $\mathbf{n}_{s,j}$ is the unit normal vector exterior to D_j on s.

In the sum of the $H_1(j)$ in (1.44) over $j \in [1, J]$, there are two types of edges s: those in $\overset{\circ}{S}$ and those included in Γ. First, each $s \in \overset{\circ}{S}$ is the common edge of two diamond-cells; then, in the sum, there are two corresponding integrals over s, in which we can factorize by $\left[\widehat{\Phi} - \frac{1}{2} \left(\Phi^T_{i(s)} + \Phi^P_{k(s)} \right) \right] (\sigma)$. Indeed, the jump of this function through s vanishes because $\widehat{\Phi} \in (H^1_0(\Omega))^2$. Secondly, each diamond-cell D_j whose boundary intersects Γ has two edges of equal length $s = [G_{i_2(j)} S_{k_\beta(j)}]$ with $\beta \in \{1, 2\}$ which are included in Γ, and their union is exactly A_j. Since $(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_j$ is a constant, and $\sum_{\beta \in \{1,2\}} \int_{[G_{i_2(j)} S_{k_\beta(j)}]} (\widehat{\Phi} - \frac{1}{2}(\Phi^T_{i_2(j)} + \Phi^P_{k_\beta(j)}))(\sigma) d\sigma = \int_{A_j} (\widehat{\Phi} - \tilde{\Phi}_h)(\sigma) d\sigma$, we have

$$\begin{aligned}
&\sum_{s \in \partial D_j \cap \Gamma} \int_s (\nabla_h \mathbf{u}_h - p_h I_2) \mathbf{n}_{s,j} \cdot \left[\widehat{\Phi} - \frac{1}{2} \left(\Phi^T_{i(s)} + \Phi^P_{k(s)} \right) \right] (\sigma) d\sigma \\
&= \sum_{\beta \in \{1,2\}} \int_{[G_{i_2(j)} S_{k_\beta(j)}]} (\nabla_h \mathbf{u}_h - p_h I_2) \mathbf{n}_j \cdot \left[\widehat{\Phi} - \frac{1}{2} \left(\Phi^T_{i_2(j)} + \Phi^P_{k_\beta(j)} \right) \right] (\sigma) d\sigma \\
&= \int_{A_j} (\nabla_h \mathbf{u}_h - p_h I_2) \mathbf{n}_j \cdot (\widehat{\Phi} - \tilde{\Phi}_h)(\sigma) d\sigma = 0,
\end{aligned} \tag{1.47}$$

thanks to (1.42) and to the fact that $\widehat{\Phi} \in (H^1_0(\Omega))^2$. With these remarks, we can write

$$\sum_{j \in [1,J]} H_1(j) = \sum_{s \in \overset{\circ}{S}} [(\nabla_h u_h - p_h I_2) \mathbf{n}_s]_s \left[\widehat{\Phi} - \frac{1}{2} (\Phi^T_{i(s)} + \Phi^P_{k(s)}) \right]. \tag{1.48}$$

Then, we may write $\widehat{\Phi} - \frac{1}{2}\left(\Phi^T_{i(s)} + \Phi^P_{k(s)}\right) = \frac{1}{2}\left[\left(\widehat{\Phi} - \Phi^T_{i(s)}\right) + \left(\widehat{\Phi} - \Phi^P_{k(s)}\right)\right]$. Summing in the right-hand side of (1.48) the various contributions of Φ^T_i for a fixed i and the various contributions of Φ^P_k for a fixed k, we obtain the following formula

$$\begin{aligned}\sum_{j=1}^{J} H_1(j) = \frac{1}{2}\sum_{i\in[1,I]}\sum_{s\subset\overset{\circ}{T}_i}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\Phi} - \Phi^T_i)(\sigma)d\sigma \\ + \frac{1}{2}\sum_{k\in[1,K]}\sum_{s\subset\overset{\circ}{P}_k}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\Phi} - \Phi^P_k)(\sigma)d\sigma.\end{aligned} \tag{1.49}$$

Finally, according to (1.44) and definition (1.32) of Φ^*_h, we obtain (1.43). □

Before we turn to a representation formula for i_2 in (1.41), we need some technical lemmas related to the $L^2(\Omega)$ scalar product of discrete gradients and curls.

Lemma 1.14. *Let $\mathbf{u} = (\mathbf{u}^T_i, \mathbf{u}^P_k)$ be the velocity component of the solution of the scheme (1.26)–(1.30) and $\Psi = (\Psi^T_i, \Psi^P_k) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$. There holds*

$$(\nabla^{T,P}_h \times (\nabla^D_h \mathbf{u}), \Psi)_{T,P} = -\sum_{k\in\Gamma}\int_{\partial P_k\cap\Gamma} (\nabla \mathbf{g}(\sigma) - \nabla^D_h \mathbf{u}_h(\sigma))\boldsymbol{\tau}_k \cdot \Psi^P_k d\sigma, \tag{1.50}$$

where $\boldsymbol{\tau}_k$ is the tangent vector to $\partial P_k \cap \Gamma$ which is positively oriented with respect to the unit normal vector exterior to $\partial P_k \cap \Gamma$.

Proof. According to Eq. (1.20) and (1.21), there holds

$$(\nabla^T_h \times (\nabla^D_h \mathbf{u}))_i = 0 \quad \forall i \in [1, I] \text{ and } (\nabla^P_h \times (\nabla^D_h \mathbf{u}))_k = 0 \quad \forall k \notin \Gamma. \tag{1.51}$$

On the other hand, since the solution of the discrete problem verifies (1.29), there holds, for $k \in \Gamma$, with the notations of Fig. 1.3:

$$\mathbf{u}^T_{I_1(k)} = 2\mathbf{g}(G_{I_1(k)}) - \frac{1}{2}(\mathbf{u}^P_k + \mathbf{u}^P_{K_1(k)}) \text{ and } \mathbf{u}^T_{I_2(k)} = 2\mathbf{g}(G_{I_2(k)}) - \frac{1}{2}(\mathbf{u}^P_k + \mathbf{u}^P_{K_2(k)}). \tag{1.52}$$

Following (1.22) and (1.52), we obtain that

$$(\nabla^P_h \times (\nabla^D_h \mathbf{u}))_k = \frac{1}{|P_k|}\left[2(\mathbf{g}(G_{I_2(k)}) - \mathbf{g}(G_{I_1(k)})) + (\mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_{K_2(k)})\right], \ \forall\ k \in \Gamma. \tag{1.53}$$

From (1.51), and using definition of scalar product in (1.8), we obtain

$$(\nabla^{T,P}_h \times (\nabla^D_h \mathbf{u}), \Psi)_{T,P} = \frac{1}{2}\sum_{k\in\Gamma} |P_k|(\nabla^P_h \times (\nabla^D_h \mathbf{u}))_k \cdot \Psi^P_k.$$

Using (1.53) leads to

$$(\nabla^{T,P}_h \times (\nabla^D_h \mathbf{u}), \Psi)_{T,P} = \sum_{k\in\Gamma}(\mathbf{g}(G_{I_2(k)}) - \mathbf{g}(G_{I_1(k)})) \cdot \Psi^P_k + \frac{1}{2}\sum_{k\in\Gamma}(\mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_{K_2(k)}) \cdot \Psi^P_k. \tag{1.54}$$

In addition, we have

$$\mathbf{g}(G_{I_2(k)}) - \mathbf{g}(G_{I_1(k)}) = (\mathbf{g}(G_{I_2(k)}) - \mathbf{g}(S_k)) + (\mathbf{g}(S_k) - \mathbf{g}(G_{I_1(k)}),$$

so that

$$\mathbf{g}(G_{I_2(k)}) - \mathbf{g}(G_{I_1(k)}) = -\int_{\partial P_k \cap \Gamma} \nabla \mathbf{g}(\sigma) \boldsymbol{\tau}_k d\sigma, \tag{1.55}$$

In the same way, we have

$$\mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_{K_2(k)} = (\mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_k) + (\mathbf{u}^P_k - \mathbf{u}^P_{K_2(k)})$$

Applying the property (1.12) of the discrete gradient:

$$|S_k S_{K_1(k)}| \nabla_h \mathbf{u_h}(\sigma) \boldsymbol{\tau}_k = \mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_k, \quad \forall \sigma \in [S_k S_{K_1(k)})] \tag{1.56}$$

and

$$|S_{K_2(k)} S_k| \nabla_h \mathbf{u_h}(\sigma) \boldsymbol{\tau}_k = \mathbf{u}^P_k - \mathbf{u}^P_{K_2(k)}, \quad \forall \sigma \in [S_{K_2(k)} S_k)], \tag{1.57}$$

there holds

$$\mathbf{u}^P_{K_1(k)} - \mathbf{u}^P_{K_2(k)} = 2 \int_{\partial P_k \cap \Gamma} \nabla^D_h \mathbf{u}_h(\sigma) \boldsymbol{\tau}_k d\sigma. \tag{1.58}$$

Combining (1.54) with (1.55) and (1.58) come to

$$(\nabla^{T,P}_h \times (\nabla^D_h \mathbf{u}), \Psi)_{T,P} = -\sum_{k \in \Gamma} \int_{\partial P_k \cap \Gamma} (\nabla \mathbf{g}(\sigma) - \nabla^D_h \mathbf{u}_h(\sigma)) \boldsymbol{\tau}_k \cdot \Psi^P_k d\sigma.$$

This implies our lemma. □

Lemma 1.15. *Let* $\mathbf{u} = (\mathbf{u}^T_i, \mathbf{u}^P_k)$ *be the velocity component of the solution of the scheme (1.26)–(1.30) and* $\Psi = (\Psi^T_i, \Psi^P_k) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$. *Let* $\mathbf{u}_h$ *and* Ψ_h *be their associated functions through Def. 1.8. There holds*

$$\begin{aligned} \sum_j \int_{D_j} \nabla_h \mathbf{u}_h : \nabla_h \times \Psi_h(\mathrm{x}) d\mathrm{x} = &-\sum_{k \in \Gamma} \int_{\partial P_k \cap \Gamma} (\nabla \mathbf{g}(\sigma) - \nabla^D_h \mathbf{u}_h(\sigma)) \boldsymbol{\tau}_k \cdot \Psi^P_k d\sigma \\ &- (\nabla^D_h \mathbf{u} \boldsymbol{\tau}, \tilde{\Psi})_{\Gamma,h}. \end{aligned} \tag{1.59}$$

Proof. Applying the discrete Green formula (1.18), there holds

$$\begin{aligned} \sum_j \int_{D_j} \nabla_h \mathbf{u}_h : \nabla_h \times \Psi_h(\mathrm{x}) d\mathrm{x} &= (\nabla^D_h \mathbf{u}, \nabla^D_h \times \Psi)_D \\ &= (\nabla^{T,P}_h \times (\nabla^D_h \mathbf{u}), \Psi)_{T,P} - (\nabla^D_h \mathbf{u} \boldsymbol{\tau}, \tilde{\Psi})_{\Gamma,h}. \end{aligned}$$

Following the lemma 1.14, we have completed our lemma. □

Proposition 1.16. *Let $\mathbf{u} = (\mathbf{u}_i^T, \mathbf{u}_k^P)$ be the velocity component of the solution of the scheme (1.26)–(1.30) and $\mathbf{u}_h$ the function associated to $\mathbf{u}$ by the definition 1.8. Let $\widehat{\Psi}$ be defined in equation (1.38). Let $\Psi = (\Psi_i^T, \Psi_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ and Ψ_h be its associated function. Then, the following representation holds*

$$\begin{aligned} i_2 = &\frac{1}{2} \sum_{i\in[1,I]} \sum_{s\in\overset{\circ}{T}_i} \int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_i^T)(\sigma) d\sigma \\ &+ \frac{1}{2} \sum_{k\in[1,K]} \sum_{s\in\overset{\circ}{P}_k} \int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_k^P)(\sigma) d\sigma \\ &- \sum_{k\in\Gamma} \int_{\partial P_k \cap \Gamma} (\nabla \mathbf{g}(\sigma) - \nabla_h^D \mathbf{u}_h(\sigma)) \boldsymbol{\tau}_k \cdot (\widehat{\Psi}(\sigma) - \Psi_k^P) d\sigma. \end{aligned} \tag{1.60}$$

Proof. From (1.41), there holds

$$\begin{aligned} i_2 &= \sum_j \int_{D_j} (\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h) : \nabla \times \widehat{\Psi}(\mathrm{x}) d\mathrm{x} \\ &= \int_\Omega \nabla \widehat{\mathbf{u}} : \nabla \times \widehat{\Psi}(\mathrm{x}) d\mathrm{x} - \sum_j \int_{D_j} \nabla_h \mathbf{u}_h : \nabla_h \times \Psi_h(\mathrm{x}) d\mathrm{x} - \sum_j H_2(j), \end{aligned} \tag{1.61}$$

where

$$H_2(j) = \int_{D_j} \nabla_h \mathbf{u}_h : (\nabla \times \widehat{\Psi} - \nabla_h \times \Psi_h)(\mathrm{x}) d\mathrm{x}. \tag{1.62}$$

By application of the continuous Green formula, there holds

$$\int_\Omega \nabla \widehat{\mathbf{u}} : \nabla \times \widehat{\Psi}(\mathrm{x}) d\mathrm{x} = -\int_{\partial\Omega} \nabla \widehat{\mathbf{u}}\, \boldsymbol{\tau} \cdot \widehat{\Psi}(\sigma) d\sigma = -\int_\Gamma \nabla \widehat{\mathbf{g}}\, \boldsymbol{\tau} \cdot \widehat{\Psi}(\sigma) d\sigma. \tag{1.63}$$

We can evaluate the sum of $H_2(j)$ over j just like we evaluated the sum of $H_1(j)$ in Proposition 1.13. There are only two differences. The first is that the gradients of $\widehat{\Phi}$ and Φ_h are replaced by the curls of $\widehat{\Psi}$ and Ψ_h, which implies that normal vectors $\mathbf{n}_s$ are replaced by tangent vectors $-\boldsymbol{\tau}_s$. The second difference is that the boundary integrals do not vanish any more, but can be evaluated like in the discussion that leads to (1.47). Then noting that

$$\sum_{j\in J^\Gamma} \int_{A_j} \nabla_h \mathbf{u}_h \boldsymbol{\tau}_j \cdot \widehat{\Psi}(\sigma) d\sigma = \sum_{k\in\Gamma} \int_{P_k\cap\Gamma} \nabla_h \mathbf{u}_h \boldsymbol{\tau}_k \cdot \widehat{\Psi}(\sigma) d\sigma,$$

we obtain the following formula

$$\begin{aligned} \sum_j H_2(j) = &-\frac{1}{2} \sum_{i\in[1,I]} \sum_{s\in\overset{\circ}{T}_i} \int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_i^T)(\sigma) d\sigma \\ &- \frac{1}{2} \sum_{k\in[1,K]} \sum_{s\in\overset{\circ}{P}_k} \int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_k^P)(\sigma) d\sigma \\ &- \sum_{k\in\Gamma} \int_{P_k\cap\Gamma} \nabla_h \mathbf{u}_h \boldsymbol{\tau}_k \cdot \widehat{\Psi}(\sigma) d\sigma + (\nabla_h^D \mathbf{u}\boldsymbol{\tau}, \tilde{\Psi})_{\Gamma,h}. \end{aligned} \tag{1.64}$$

Combining (1.61), (1.59), (1.63) and (1.64), we have

$$\begin{aligned} i_2 = & \frac{1}{2}\sum_{i\in[1,I]}\sum_{s\in\overset{\circ}{T}_i}\int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_i^T)(\sigma) d\sigma \\ & + \frac{1}{2}\sum_{k\in[1,K]}\sum_{s\in\overset{\circ}{P}_k}\int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_k^P)(\sigma) d\sigma \\ & - \sum_{k\in\Gamma}\int_{P_k\cap\Gamma} (\nabla \mathbf{g}(\sigma) - \nabla_h^D \mathbf{u}_h(\sigma))\boldsymbol{\tau}_k \cdot (\widehat{\Psi}(\sigma) - \Psi_k^P) d\sigma. \end{aligned} \tag{1.65}$$

We have finished this proposition. □

1.4.2 A representation of the pressure error

Proposition 1.17. *Let* $\widehat{\mathbf{v}} \in (H_0^1(\Omega))^2$ *and* $\mathbf{v} = (\mathbf{v}_i^T, \mathbf{v}_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ *be such that* $\tilde{\mathbf{v}}_j = 0$ *for all* $j \in \Gamma$. *We have that*

$$\begin{aligned} \int_\Omega e_p \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} = & \int_\Omega \nabla_h \mathbf{e} : \nabla\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} - \frac{1}{2}\sum_{i\in[1,I]}\int_{T_i} \mathbf{f}\cdot(\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\mathbf{x}) d\mathbf{x} \\ & - \frac{1}{2}\sum_{k\in[1,K]}\int_{P_k} \mathbf{f}\cdot(\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\mathbf{x}) d\mathbf{x} \\ & + \frac{1}{2}\sum_{i\in[1,I]}\sum_{s\subset\overset{\circ}{T}_i}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\sigma) d\sigma \\ & + \frac{1}{2}\sum_{k\in[1,K]}\sum_{s\subset\overset{\circ}{P}_k}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\sigma) d\sigma. \end{aligned} \tag{1.66}$$

Proof. We can use the formula (1.36) to obtain

$$\begin{aligned} \int_\Omega e_p \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} = & \int_\Omega p\nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} - \int_\Omega p_h \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} \\ = & \int_\Omega \nabla_h \mathbf{e} : \nabla\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} - \int_\Omega \mathbf{f}\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} + \int_\Omega \nabla_h \mathbf{u}_h : \nabla\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} \\ & - \int_\Omega p_h \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x}. \end{aligned} \tag{1.67}$$

Using Eq. (1.33), we have

$$\begin{aligned} \int_\Omega e_p \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x}(\mathrm{x}) d\mathrm{x} = & \int_\Omega \nabla_h \mathbf{e} : \nabla\widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} - \int_\Omega \mathbf{f}\cdot(\widehat{\mathbf{v}} - \mathbf{v}_h^*)(\mathrm{x}) d\mathrm{x} \\ & + \int_\Omega (\nabla_h \mathbf{u}_h - p_h I_2) : (\nabla\widehat{\mathbf{v}} - \nabla_h \mathbf{v}_h)(\mathrm{x}) d\mathrm{x}. \end{aligned} \tag{1.68}$$

Just like in Proposition 1.13 (see (1.45) and (1.49)), we have

$$\int_\Omega (\nabla_h \mathbf{u}_h - p_h I_2) : (\nabla \widehat{\mathbf{v}} - \nabla_h \mathbf{v}_h)(\mathrm{x}) d\mathrm{x} = \frac{1}{2} \sum_{i\in[1,I]} \sum_{s\subset \overset{\circ}{T}_i} \int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\sigma) d\sigma + \frac{1}{2} \sum_{k\in[1,K]} \sum_{s\subset \overset{\circ}{P}_k} \int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\sigma) d\sigma. \tag{1.69}$$

From (1.68) and (1.69), we obtain the following formula

$$\begin{aligned}\int_\Omega e_p \nabla \cdot \widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} = &\int_\Omega \nabla_h \mathbf{e} : \nabla \widehat{\mathbf{v}}(\mathrm{x}) d\mathrm{x} - \frac{1}{2} \sum_{i\in[1,I]} \int_{T_i} \mathbf{f} \cdot (\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\mathbf{x}) d\mathbf{x} \\ &- \frac{1}{2} \sum_{k\in[1,K]} \int_{P_k} \mathbf{f} \cdot (\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\mathbf{x}) d\mathbf{x} \\ &+ \frac{1}{2} \sum_{i\in[1,I]} \sum_{s\subset \overset{\circ}{T}_i} \int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\sigma) d\sigma \\ &+ \frac{1}{2} \sum_{k\in[1,K]} \sum_{s\subset \overset{\circ}{P}_k} \int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\sigma) d\sigma.\end{aligned}$$

This implies our proposition. □

1.5 A computable error bound

1.5.1 Preliminaries

In this subsection, we will present some Poincare-type inequalities which are useful to obtain a computable error bound.

Lemma 1.18. *Let ω be an open bounded set which is star-shaped with respect to one of its point. Let $\mathbf{u} \in (H^1(\omega))^2$ and let $\overline{\mathbf{u}}_\omega$ be the mean-value of $\mathbf{u}$ over ω. Then,*

$$\exists C(\omega),\ s.t\ .\ \|\mathbf{u} - \overline{\mathbf{u}}_\omega\|_{L^2(\omega)} \le C(\omega) diam(\omega) \|\nabla \mathbf{u}\|_{L^2(\omega)}. \tag{1.70}$$

Note that when ω is convex, a universal constant $C(\omega)$ is given by $\frac{1}{\pi}$ (see in [9]). When ω is not convex, we may use explicitly computable formulas given, for example, by [2, 13]. The constant $C(\omega)$ only denpends on the shape of ω, not on its diameter.

Finally, we will also need a trace inequality (see [8]).

Lemma 1.19. *Let T be a triangle and let E be one of its edges; let ρ be the distance from E to the vertex of T opposite to E, and let σ be the longest among the two other sides if T. Let $\varepsilon > 0$ be an arbitrary real valued number; then for all $\mathbf{u} \in (H^1(T))^2$, there holds*

$$\|\mathbf{u}\|^2_{L^2(E)} \le \frac{1}{\rho} \left((2 + \varepsilon^{-2}) \|\mathbf{u}\|^2_{L^2(T)} + \varepsilon^2 \sigma^2 \|\nabla \mathbf{u}\|^2_{L^2(T)} \right). \tag{1.71}$$

1.5.2 A computable bound for the velocity error

In the expression (1.43) of i_1, the values of (Φ_i^T, Φ_k^P) are arbitrary, except for the boundary midpoint values chosen so that (1.42) holds. In the expression of i_2 in (1.60), the values of (Ψ_i^T, Ψ_k^P) are arbitrary.

Definition 1.20. *Since $\widehat{\Phi}$, $\widehat{\Psi}$ are not necessarily more regular than $(H^1(\Omega))^2$, we choose as an interpolation their L^2 projection on the primal and dual cells:*

$$\Phi_i^T = \frac{1}{|T_i|}\int_{T_i} \widehat{\Phi}(\mathrm{x})d\mathrm{x}\ \forall i \in [1,I],\ \Phi_k^P = \frac{1}{|P_k|}\int_{P_k} \widehat{\Phi}(\mathrm{x})d\mathrm{x}\ \forall k \in [1,K], \tag{1.72}$$

$$\Psi_i^T = \frac{1}{|T_i|}\int_{T_i} \widehat{\Psi}(\mathrm{x})d\mathrm{x}\ \forall i \in [1,I],\ \Psi_k^P = \frac{1}{|P_k|}\int_{P_k} \widehat{\Psi}(\mathrm{x})d\mathrm{x}\ \forall k \in [1,K]. \tag{1.73}$$

In order to complete the definition of (Φ_i^T, Φ_k^P), for any $i \in \Gamma$, the boundary value of Φ_i^T are given by (1.42). In this problem, it is not necessary to define the value of Ψ_i^T for all $i \in \Gamma$.

Proposition 1.21. *Let $h_i^T := diam(T_i)$, $h_k^P := diam(P_k)$. There exist computable constant $C(T_i)$, and $C(P_k)$ such that*

$$\left|\sum_{i\in[1,I]}\int_{T_i} \mathbf{f}\cdot(\widehat{\Phi} - \Phi_i^T)(\mathrm{x})d\mathrm{x}\right| \leq osc(\mathbf{f},T,\Omega)\|\nabla\widehat{\Phi}\|_{L^2(\Omega)}, \tag{1.74}$$

$$\left|\sum_{k\in[1,K]}\int_{P_k} \mathbf{f}\cdot(\widehat{\Phi} - \Phi_k^P)(\mathrm{x})d\mathrm{x}\right| \leq osc(\mathbf{f},P,\Omega)\|\nabla\widehat{\Phi}\|_{L^2(\Omega)}, \tag{1.75}$$

where

$$osc(\mathbf{f},T,\Omega) = \left(\sum_{i\in[1,I]} (C(T_i)h_i^T)^2\|\mathbf{f} - \mathbf{f}_i^T\|_{L^2(T_i)}\right)^{1/2} \tag{1.76}$$

and

$$osc(\mathbf{f},P,\Omega) = \left(\sum_{k\in[1,K]} (C(P_k)h_k^P)^2\|\mathbf{f} - \mathbf{f}_k^P\|_{L^2(P_k)}\right)^{1/2}. \tag{1.77}$$

Proof. Since Φ_i^T was chosen as the mean-value of $\widehat{\Phi}$ over T_i (see (1.72)), we have

$$\int_{T_i} \mathbf{f}\cdot(\widehat{\Phi} - \Phi_i^T)(\mathrm{x})d\mathrm{x} = \int_{T_i} (\mathbf{f} - \mathbf{f}_i^T)\cdot(\widehat{\Phi} - \Phi_i^T)(\mathrm{x})d\mathrm{x}.$$

Applying the Cauchy-Schwarz inequality, Lemma 1.18 to $\widehat{\Phi}$ over T_i and the discrete Cauchy-Schwarz inequality, we are lead to (1.74).
Similarly, we also obtain (1.75). □

Propositions 1.22 and 1.24 below are proved just like Propositions 5.9 and 5.10 in [8].

Proposition 1.22. *For any primal cell T_i and any dual P_k such that $T_i \cap P_k \neq \emptyset$, Let $s = [G_i S_k]$ and $t_{ik,1}$ and $t_{ik,2}$ be the triangles defined in Fig. 1.4 such that $t_{ik,1} \cup t_{ik,2} = T_i \cap P_k$. Let $\rho_{ik,\alpha}$ be the distance from s to the vertex of $t_{ik,\alpha}$ opposite to s and $\sigma_{ik,\alpha}$ be the length of the longest among the two other edges of $t_{ik,\alpha}$. $C(T_i)$ is the constant that appears in (1.70). For any strictly positive μ, let us define*

$$C_s(\mu) = \frac{\left(1+\sqrt{1+\frac{\sigma_{ik,1}^2}{\mu}}\right)\left(1+\sqrt{1+\frac{\sigma_{ik,2}^2}{\mu}}\right)}{\left(1+\sqrt{1+\frac{\sigma_{ik,1}^2}{\mu}}\right)\rho_{ik,2} + \left(1+\sqrt{1+\frac{\sigma_{ik,2}^2}{\mu}}\right)\rho_{ik,1}}, \tag{1.78}$$

$$\chi_i(\mu) = (C(T_i)h_i^T)^2 + \mu. \tag{1.79}$$

$$\tag{1.80}$$

We define the local and global error estimators related to the primal mesh:

$$(\eta_i^T)^2 = \inf_{\mu>0}\left[\chi_i(\mu)\sum_{s\in\overset{\circ}{T}_i} C_s(\mu)\|[(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s\|^2_{L^2(s)}\right] \text{ and } (\eta^T)^2 = \sum_i (\eta_i^T)^2, \tag{1.81}$$

$$(\eta'^T_i)^2 = \inf_{\mu>0}\left[\chi_i(\mu)\sum_{s\in\overset{\circ}{T}_i} C_s(\mu)\|[\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s\|^2_{L^2(s)}\right] \text{ and } (\eta'^T)^2 = \sum_i (\eta'^T_i)^2. \tag{1.82}$$

With these definitions, there holds:

$$\left|\sum_{i\in[1,I]}\sum_{s\in\overset{\circ}{T}_i}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\hat{\Phi} - \Phi_i^T)(\sigma)d\sigma\right| \leq \eta^T \|\nabla\widehat{\Phi}\|_{L^2(\Omega)}, \tag{1.83}$$

$$\left|\sum_{i\in[1,I]}\sum_{s\in\overset{\circ}{T}_i}\int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\hat{\Psi} - \Psi_i^T)(\sigma)d\sigma\right| \leq \eta'^T \|\nabla\widehat{\Psi}\|_{L^2(\Omega)}. \tag{1.84}$$

Remark 1.23. *The minimization in (1.81) is performed numerically when we effectively compute the estimators. However, we may already get an idea of the behaviour of this quantity by choosing $\mu = h_{T_i}^2$ to evaluate η_i^T. By definition of $\sigma_{ik,\alpha}$, this length is lower than the diameter of T_i, which implies*

$$C_s\left(\left(h_i^T\right)^2\right) \leq \frac{(1+\sqrt{2})^2}{2(\rho_{ik,1}+\rho_{ik,2})}. \tag{1.85}$$

Under the hypothesis that the ratios $\frac{\rho_{ik,\alpha}}{h_i^T}$ are all bounded by below by the same constant which does not depend on the mesh, we obtain the following bound

$$\eta_i^T \leq K h_i^T \sum_{s\in\overset{\circ}{T}_i} \|[(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s\|^2_{L^2(s)},$$

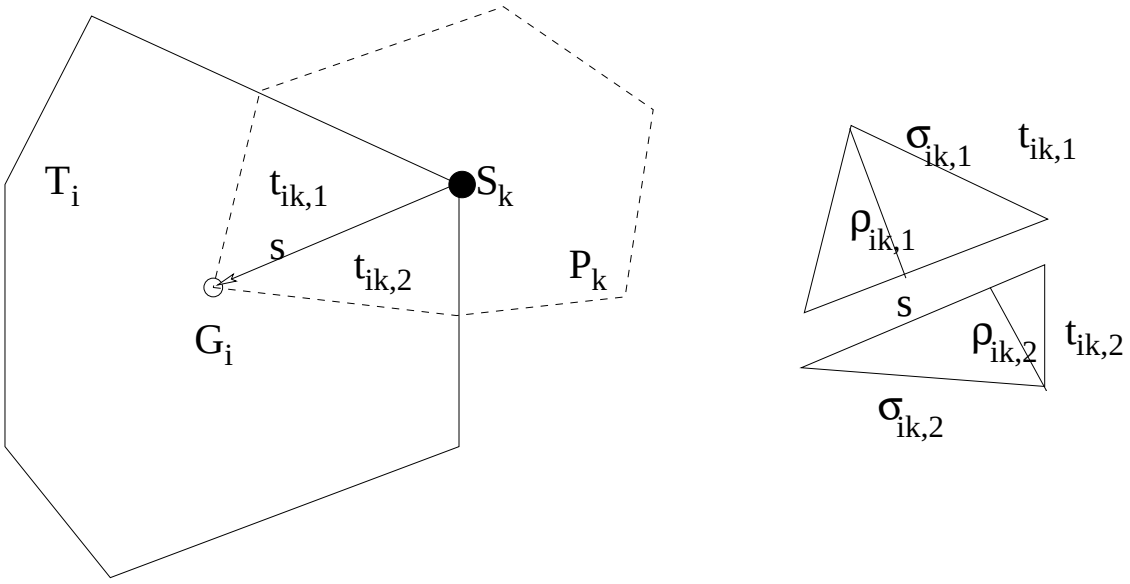

Figure 1.4: For each cell T_i and each vertex S_k of T_i, $T_i \cap P_k$ is split in two triangles $t_{ik,1}$ and $t_{ik,2}$.

where the constant K does not depend on the mesh. The same remark holds for $\eta_i'^T$.

Proposition 1.24. *Let us set the same notations as in Prop. 1.22. Let C_s be defined by (1.78). Let $C(P_k)$ be the constant involved in (1.70). Let us define for any strictly positive μ,*

$$\chi_k(\mu) = (C(P_k)h_k^P)^2 + \mu. \tag{1.86}$$

We define the local and global error estimators related to the dual mesh:

$$(\eta_k^P)^2 = \inf_{\mu>0}\left[\chi_k(\mu)\sum_{s\in\overset{\circ}{P}_k} C_s(\mu)\|[(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s\|^2_{L^2(s)}\right] \text{ and } (\eta^P)^2 = \sum_k (\eta_k^P)^2, \tag{1.87}$$

$$(\eta_k'^P)^2 = \inf_{\mu>0}\left[\chi_k(\mu)\sum_{s\in\overset{\circ}{P}_k} C_s(\mu)\|[\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s\|^2_{L^2(s)}\right] \text{ and } (\eta'^P)^2 = \sum_k (\eta_k'^P)^2. \tag{1.88}$$

With these definitions, there holds:

$$\left|\sum_{k\in[1,K]}\sum_{s\in\overset{\circ}{P}_k}\int_s [(\nabla_h \mathbf{u}_h - p_h I_2)\mathbf{n}_s]_s \cdot (\widehat{\Phi} - \Phi_k^P)(\sigma)d\sigma\right| \leq \eta^P \|\nabla\widehat{\Phi}\|_{L^2(\Omega)}, \tag{1.89}$$

$$\left|\sum_{k\in[1,K]}\sum_{s\in\overset{\circ}{P}_k}\int_s [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s \cdot (\widehat{\Psi} - \Psi_k^P)(\sigma)d\sigma\right| \leq \eta'^P \|\nabla\widehat{\Psi}\|_{L^2(\Omega)}. \tag{1.90}$$

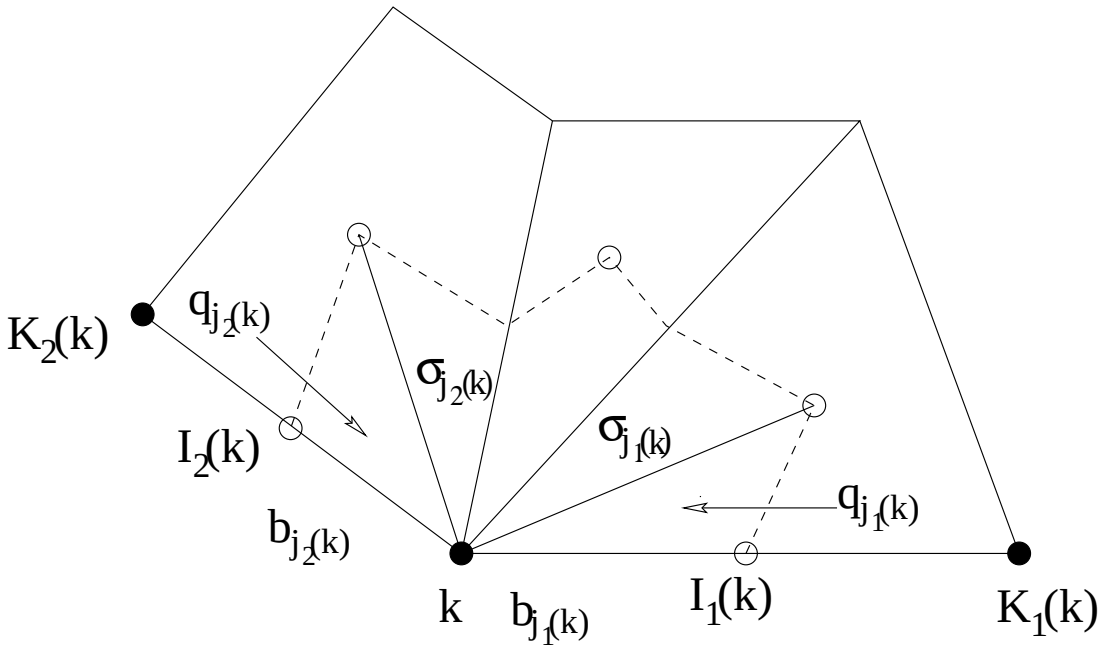

Figure 1.5: For any $k \in \Gamma$, S_k is the common vertex of $q_{j_1(k)}$ and $q_{j_2(k)}$.

Proposition 1.25. *For any $k \in \Gamma$, let us denote by $D_{j_1(k)}$ and $D_{j_2(k)}$ the two diamond cells whose boundary intersect Γ and which have S_k as a vertex. Let $q_{j_1(k)} = P_k \cap D_{j_1(k)}$, $q_{j_2(k)} = P_k \cap D_{j_2(k)}$ and the segment $b_{j_\alpha(k)}$ be the intersection between $\partial q_{j_\alpha(k)}$ and Γ. Let $\rho_{j_\alpha(k)}$ be the distance from $b_{j_\alpha(k)}$ to the vertex of $q_{j_\alpha(k)}$ opposite to $b_{j_\alpha(k)}$ and $\sigma_{j_\alpha(k)}$ be the length of the longest among the two other edges of $q_{j_\alpha(k)}$ (see Fig. 1.5). $C(P_k)$ is the constant that appears in (1.70). For any strictly positive μ, let us define*

$$C_\alpha(\mu) = \frac{2 + \frac{\sigma^2_{j_\alpha(k)}}{\mu + \sqrt{\mu^2 + \mu\sigma^2_{j_\alpha(k)}}}}{\rho_{j_\alpha(k)}}, \tag{1.91}$$

$$\lambda_k(\mu) = (C(P_k) h_k^P)^2 + \mu. \tag{1.92}$$

We define the local and global error estimator related to the boundary:

$$(\zeta_k^P)^2 = \inf_\mu \left[\lambda_k(\mu) \sum_{\alpha=1}^{2} C_\alpha(\mu) \|(\nabla \mathbf{g} - \nabla \mathbf{u}_h)\boldsymbol{\tau}_{j_\alpha(k)}\|^2_{L^2(b_{j_\alpha(k)})} \right] \text{ and } (\zeta^P)^2 = \sum_{k\in\Gamma} (\zeta_k^P)^2. \tag{1.93}$$

With these definitions, there holds:

$$\sum_{k\in\Gamma} \int_{\partial P_k \cap \Gamma} \left| (\nabla \mathbf{g} - \nabla_h \mathbf{u}_h)\boldsymbol{\tau}_k \cdot (\widehat{\Psi} - \Psi_k^P)(\sigma) \right| d\sigma \leq \zeta^P \|\nabla \widehat{\Psi}\|_{L^2(\Omega)}. \tag{1.94}$$

Proof. By application of the Cauchy-Schwarz inequality on each edge $b_{j_\alpha(k)}$ and the weighted discrete Cauchy-Schwarz inequality, we obtain for any set of strictly positive real-valued

numbers C_α^P

$$\int_{\partial P_k\cap\Gamma}\Big|(\nabla\mathbf{g}-\nabla_h\mathbf{u}_h)\boldsymbol{\tau}_k\cdot(\widehat{\Psi}-\Psi_k^P)(\sigma)\Big|\,d\sigma=\sum_{\alpha=1}^{2}\int_{b_{j_\alpha(k)}}\Big|(\nabla\mathbf{g}-\nabla_h\mathbf{u}_h)\boldsymbol{\tau}_{j_\alpha(k)}\cdot(\widehat{\Psi}-\Psi_k^P)(\sigma)\Big|\,d\sigma$$
$$\leq\left[\sum_{\alpha=1}^{2}C_\alpha^P\|(\nabla\mathbf{g}-\nabla_h\mathbf{u}_h)\boldsymbol{\tau}_{j_\alpha(k)}\|^2_{L^2(b_{j_\alpha(k)})}\right]^{1/2}\left[\sum_{\alpha=1}^{2}\frac{1}{C_\alpha^P}\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(b_{j_\alpha(k)})}\right]^{1/2}. \tag{1.95}$$

Now, for each segment $b_{j_\alpha(k)}$, we can apply the trace inequality (1.71) on each triangle $q_{j_\alpha(k)}$, for all $\alpha\in\{1,2\}$ and for all strictly positive $\varepsilon_{j_\alpha(k)}$. With $C_{1,j_\alpha(k)}=\frac{2+\varepsilon^{-2}_{j_\alpha(k)}}{\rho_{j_\alpha(k)}}$ and $C_{2,j_\alpha(k)}=\frac{\varepsilon^2_{j_\alpha(k)}\sigma^2_{j_\alpha(k)}}{\rho_{j_\alpha(k)}}$, we obtain

$$\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(b_{j_\alpha(k)})}\leq C_{1,j_\alpha(k)}\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(q_{j_\alpha(k)})}+C_{2,j_\alpha(k)}\|\nabla\widehat{\Psi}\|^2_{L^2(q_{j_\alpha(k)})}.$$

Let $\mu>0$ be arbitrary. For $b_{j_\alpha(k)}$ for $\alpha\in\{1,2\}$, let us choose $\varepsilon_{j_\alpha(k)}$ so that

$$\varepsilon^2_{j_\alpha(k)}=\frac{\mu+\sqrt{\mu^2+\mu\sigma^2_{j_\alpha(k)}}}{\sigma^2_{j_\alpha(k)}}\Longleftrightarrow C_{2,j_\alpha(k)}=\mu C_{1,j_\alpha(k)} \tag{1.96}$$

and $C_\alpha^P=C_{1,j_\alpha(k)}$. There holds:

$$\begin{aligned}\sum_{\alpha=1}^{2}\frac{1}{C_\alpha^P}\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(b_{j_\alpha(k)})}&\leq\sum_{\alpha=1}^{2}\left(\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(q_{j_\alpha(k)})}+\mu\|\nabla\widehat{\Psi}\|^2_{L^2(q_{j_\alpha(k)})}\right)\\&\leq\|\widehat{\Psi}-\Psi_k^P\|^2_{L^2(P_k)}+\mu\|\nabla\widehat{\Psi}\|^2_{L^2(P_k)}\\&\leq\left[(C(P_k)h_k^P)^2+\mu\right]\|\nabla\widehat{\Psi}\|^2_{L^2(P_k)}.\end{aligned} \tag{1.97}$$

Taking (1.97) into (1.95) and applying the discrete Cauchy-Schwarz inequality leads to (1.94). □

Before estimating the velocity error, we now define the indicator related to the penalization:

$$\zeta_\varepsilon=\varepsilon\|(\Delta_h^D p)_h\|_{L^2(\Omega)}. \tag{1.98}$$

In the term in the right-hand side of (1.41), it is easy to see that

$$\left|\varepsilon\int_\Omega(\Delta_h^D p)_h q(\mathrm{x})d\mathrm{x}\right|\leq\zeta_\varepsilon\|q\|_{L^2(\Omega)}. \tag{1.99}$$

Theorem 1.26. *Let $\|\nabla_h\mathbf{e}\|_{L^2(\Omega)}$ be defined by (1.37), let the definitions (1.76)-(1.77), (1.78)-(1.82), (1.86)-(1.88), (1.91)-(1.93) and (1.98) hold. We have*

$$\|\nabla_h\mathbf{e}\|_{L^2(\Omega)}\leq\eta=\eta_h+\eta_\epsilon. \tag{1.100}$$

where

$$\eta_h = \frac{1}{2}\Big(osc(f,T,\Omega) + osc(f,P,\Omega) + \eta^T + \eta^P\Big) + \left(\frac{1}{2} + \frac{\sqrt{2}}{\beta}\right)\Big(\eta'^T + \eta'^P + 2\zeta^P\Big) \quad (1.101)$$

$$\eta_\epsilon = \frac{2}{\beta}\zeta_\varepsilon. \quad (1.102)$$

η, η_h, and η_ϵ are called the total estimator, the discretization estimator, and the penalization estimator for the velocity, respectively.

Proof. Using (1.41), (1.39), (1.43), (1.60), (1.74)-(1.75), (1.83)-(1.84), (1.89)-(1.90), (1.94) and (1.99), we finish the proof of this theorem. □

1.5.3 A computable bound for the pressure error

Proposition 1.27. *The following estimate holds:*

$$\|e_p\|_{L(\Omega)} \leq \frac{1}{2\beta}\left(2\|\nabla_h \mathbf{e}\|_{L^2(\Omega)} + osc(\mathbf{f},T,\Omega) + osc(\mathbf{f},P,\Omega) + \eta^T + \eta^P\right). \quad (1.103)$$

Proof. Since $e_p \in L^2_0(\Omega)$, there exists $\widehat{\mathbf{v}} \in (H^1_0(\Omega))^2$, such that

$$\|e_p\|_{L^2(\Omega)} \leq \frac{1}{\beta}\frac{\int_\Omega e_p \nabla\cdot\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x}}{\|\nabla\widehat{\mathbf{v}}\|_{L^2(\Omega)}}. \quad (1.104)$$

Using (1.66), we obtain that

$$\begin{aligned}\int_\Omega e_p\nabla\cdot\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x} &= \int_\Omega \nabla_h\mathbf{e} : \nabla\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x} - \frac{1}{2}\sum_{i\in[1,I]}\int_{T_i}\mathbf{f}\cdot(\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\mathbf{x})d\mathbf{x}\\ &- \frac{1}{2}\sum_{k\in[1,K]}\int_{P_k}\mathbf{f}\cdot(\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\mathbf{x})d\mathbf{x}\\ &+ \frac{1}{2}\sum_{i\in[1,I]}\sum_{s\subset\overset{\circ}{T}_i}\int_s[(\nabla_h\mathbf{u}_h - p_hI_2)\mathbf{n}_s]_s\cdot(\widehat{\mathbf{v}} - \mathbf{v}_i^T)(\sigma)d\sigma\\ &+ \frac{1}{2}\sum_{k\in[1,K]}\sum_{s\subset\overset{\circ}{P}_k}\int_s[(\nabla_h\mathbf{u}_h - p_hI_2)\mathbf{n}_s]_s\cdot(\widehat{\mathbf{v}} - \mathbf{v}_k^P)(\sigma)d\sigma.\end{aligned}$$

We choose $\mathbf{v} = (\mathbf{v}_i^T, \mathbf{v}_k^P) \in (\mathbb{R}^2)^{I+J^\Gamma} \times (\mathbb{R}^2)^K$ such that

$$\mathbf{v}_i^T = \frac{1}{|T_i|}\int_{T_i}\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x} \;\; \forall i\in[1,I],\;\; \mathbf{v}_k^P = \frac{1}{|P_k|}\int_{P_k}\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x}\;\forall k\in[1,K] \quad (1.105)$$

and the boundary values of $\mathbf{v}_i,\ i\in\Gamma$ are chosen so that $\tilde{\mathbf{v}}_j = 0$ for all $j\in\Gamma$. Similarly to the calculations involved in propositions 1.21, 1.22 and 1.24, we obtain

$$\left|\int_\Omega e_p\nabla\cdot\widehat{\mathbf{v}}(\mathrm{x})d\mathrm{x}\right| \leq \frac{1}{2}\left(2\|\nabla_h\mathbf{e}\|_{L^2(\Omega)} + osc(\mathbf{f},T,\Omega) + osc(\mathbf{f},P,\Omega) + \eta^T + \eta^P\right)\|\nabla\widehat{\mathbf{v}}\|_{L^2(\Omega)}. \quad (1.106)$$

Taking (1.106) into (1.104), we have finished our proposition. □

1.6 Efficiency of the estimators

Since the estimator η_i^T involves jumps of $\nabla_h \mathbf{u}_h - p_h I_2$ through the common edge $s = [G_i S_k]$ of two neighboring diamond-cells, we shall use functions with a support included in the triangle $t_{ik,\alpha}$ with $\alpha = 1$ or 2, defined in the Fig. 1.4. Since we consider a fixed s in what follows, we simplify the notations to t_1 and t_2. For any triangle t in $\{t_1, t_2\}$, we denote by $\lambda_{t,\beta}$ the barycentric coordinates associated with the tree vertices of t, with $\beta \in \{1, 2, 3\}$. We suppose that the vertices of t_1 and t_2 are locally numbered so that the two nodes of the edge s are the vertices 1 and 2 of each of the triangles t_1 and t_2.

Definition 1.28. *We define the following bubble functions*

$$b_t = 27\lambda_{t,1}\lambda_{t,2}\lambda_{t,3} \quad for\ t = t_1 \quad or \quad t = t_2, \tag{1.107}$$

$$b_s = \begin{cases} 4\lambda_{t_\alpha,1}\lambda_{t_\alpha,2} & on \quad t_\alpha,\ \alpha = \{1,2\} \\ 0 & elsewhere. \end{cases} \tag{1.108}$$

There holds $\omega_t = supp(b_t) \subset t$ and $\omega_s := supp(b_s) = T_i \cap P_k = t_1 \cup t_2$. The following propositions are given for example, in [12].

Proposition 1.29. *There holds*

$$0 \le b_t \le 1,\ 0 \le b_s \le 1, \tag{1.109}$$

$$\int_s b_s(\sigma) d\sigma = \frac{2}{3}|s|. \tag{1.110}$$

Proposition 1.30. *There exists a constant $C > 0$ only depending on the minimal angle in the couple (t_1, t_2) such that, for $t = t_1$ or $t = t_2$ and $h_t = diam(t)$*

$$\frac{1}{C}h_t^2 \le \int_t b_t(\mathrm{x}) d\mathrm{x} = \frac{9}{20}|t| \le Ch_t^2, \tag{1.111}$$

$$\frac{1}{C}s^2 \le \int_t b_s(\mathrm{x}) d\mathrm{x} = \frac{1}{3}|t| \le Cs^2, \tag{1.112}$$

$$\|\nabla b_t\|_{L^2(t)} \le Ch_t^{-1}\|b_t\|_{L^2(t)}, \tag{1.113}$$

$$\|\nabla b_s\|_{L^2(t)} \le Cs^{-1}\|b_s\|_{L^2(t)}. \tag{1.114}$$

In order to prove the local efficiency of the error estimator we shall make the following hypothesis:

Hypothesis 1.31. *We suppose that the triangulation of Ω composed of all the triangles $t_{ik,\alpha}$ is regular in the sense that the minimum angles in those triangles are bounded by below independently of the mesh.*

From this hypothesis, we derive the following propositions.

Proposition 1.32. *For any primal cell T_i and any dual cell P_k such that $T_i \cap P_k \neq \emptyset$, let $s = [G_i S_k]$ and $t_{ik,1}$ and $t_{ik,2}$ be the triangles in Fig. 1.4 such that $t_{ik,1} \cup t_{ik,2} = T_i \cap P_k$. Let $h_i^T = diam(T_i)$, $h_k^P = diam(P_k)$ and $S_{ik} = |T_i \cap P_k|$. Let Hypothesis 1.31 hold. Then, there exists a constant C independent of the mesh such that*

$$(h_i^T)^2 S_{i,k}^{-1} \le C \quad and \quad (h_k^P)^2 S_{i,k}^{-1} \le C.$$

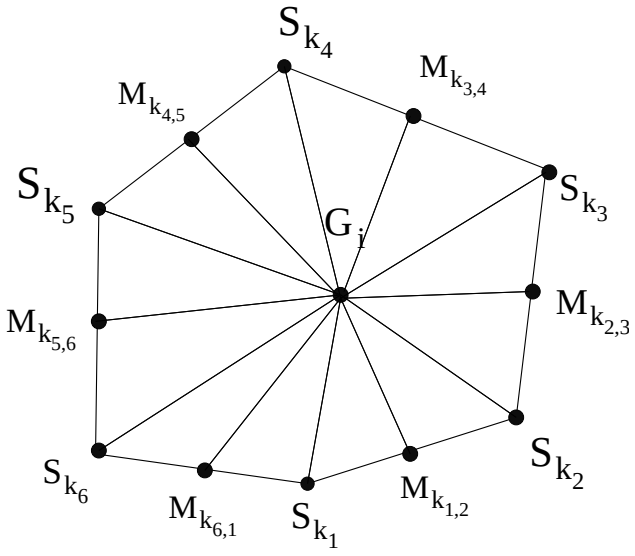

Figure 1.6: Notations of Prop. 1.32.

Proof. We will only prove the first inequality, since the second one can be treated in the same way.

Let $\alpha_0 > 0$ be the lower bound of all the angles of all the triangles $t_{ik,\alpha}$.

For any $i \in [i, I]$, let V_i be the number of vertices of the primal cell T_i. First, we note that

$$V_i \leq V := \frac{2\pi}{2\alpha_0}, \text{ for all } i \in [1, I]. \tag{1.115}$$

Let S_{k_ℓ}, with $\ell \in [1, V_i]$ be the vertices of the primal cell T_i and $M_{k_{\ell,\ell+1}}$ be the midpoint of segment $[S_{k_\ell} S_{k_{\ell+1}}]$, then

$$S_{ik_\ell} = |T_i \cap P_{k_\ell}| = |G_i S_{k_\ell} M_{k_{\ell,\ell+1}}| + |G_i S_{k_\ell} M_{k_{\ell-1,\ell}}|. \tag{1.116}$$

Now, let us estimate the area of triangle $G_i S_{k_\ell} M_{k_{\ell,\ell+1}}$. Following Hypothesis 1.31, all the angles of triangle $G_i S_{k_\ell} M_{k_{\ell,\ell+1}}$ are greater than α_0. Let h_{G_i} be the the maximum distance from point G_i to the boundary of T_i, i.e.,

$$h_{G_i} := \max\{|G_i S_{k_\ell}|, \ \ell \in [1, V_i]\}. \tag{1.117}$$

We have

$$|G_i S_{k_\ell} M_{k_{\ell,\ell+1}}| = \frac{1}{2} \sin(S_{k_\ell} \widehat{G_i M_{k_{\ell,\ell+1}}}) |G_i S_{k_\ell}| |G_i M_{k_{\ell,\ell+1}}|. \tag{1.118}$$

By a calculation on triangles $G_i S_{k_\ell} M_{k_{\ell,\ell+1}}$ and $G_i S_{k_{\ell+1}} M_{k_{\ell,\ell+1}}$, there holds

$$|G_i M_{k_{\ell,\ell+1}}| \geq |G_i S_{k_\ell}| \sin \alpha_0, \qquad |G_i S_{k_{\ell+1}}| \geq |G_i M_{k_{\ell,\ell+1}}| \sin \alpha_0. \tag{1.119}$$

From (1.119), we have the recurrence formula

$$|G_i S_{k_\ell}| \geq (\sin\alpha_0)^2 |G_i S_{k_{\ell+1}}|. \tag{1.120}$$

Starting from the vertex S_k which reaches the max in definition (1.117), the shortest way to go to a given vertex S_{k_ℓ} contains at most $V_i/2$ neighboring vertices for which we may apply (1.120), and we obtain:

$$|G_i S_{k_\ell}| \geq (\sin\alpha_0)^{V_i} h_{G_i}. \tag{1.121}$$

Then, from (1.119), we get that

$$|G_i S_{k_\ell}||G_i M_{k_{\ell,\ell+1}}| \geq (\sin\alpha_0)^{2V_i+1} h_{G_i}^2. \tag{1.122}$$

Combining (1.115), (1.118) and (1.122), and noting that from the definition of h_{G_i}, then $h_{G_i} \geq \frac{h_i^T}{2}$, we obtain

$$|G_i S_{k_\ell} M_{k_{\ell,\ell+1}}| \geq (\sin\alpha_0)^{2V+2} \frac{(h_i^T)^2}{8}. \tag{1.123}$$

In the same way, we have

$$|G_i S_{k_\ell} M_{k_{\ell-1,\ell}}| \geq (\sin\alpha_0)^{2V+2} \frac{(h_i^T)^2}{8}. \tag{1.124}$$

Using (1.116), (1.123)–(1.124), we obtain:

$$S_{ik_\ell} \geq (\sin\alpha_0)^{2V+2} \frac{(h_i^T)^2}{4}. \tag{1.125}$$

Thus, the inequality is proved with $C = 4\,(\sin\alpha_0)^{-\frac{2\pi}{\alpha_0}-2}$. □

Proposition 1.33. *Under Hypothesis 1.31, the positive constants $C(T_i)$ and $C(P_k)$ are bounded independently of the mesh, and the constant C in Prop. 1.30 is bounded by above and by below independently of the mesh.*

Proof. The constants $C(T_i), C(P_k)$ coming from (1.70) were bounded explicitly in [13]. From these expressions, it is easily seen that they are bounded if Hyp. 1.31 holds. Moreover, it is proved in [12] that C in Prop. 1.30 depends only on the regularity of the triangles $t_{ik,\alpha}$. □

Now, we will consider the efficiency of the estimators.

Theorem 1.34. *For any primal cell T_i, let $h_i^T := diam(T_i)$ and $\mathbf{f}_i^T$ be the mean-value of $\mathbf{f}$ over T_i. Let η_i^T (resp. $\eta_i'^T$) be defined in (1.81) (resp. in (1.82)). For any dual cell P_k, let $h_k^P := diam(P_k)$ and $\mathbf{f}_k^P$ be the mean-value of $\mathbf{f}$ over P_k. Let η_k^P (resp. $\eta_k'^P$) be defined in (1.87) (resp. in (1.88)). And for any boundary dual cell P_k, let ζ_k^P be defined in (1.93). Let Hypothesis 1.31 hold. Then, there exists a constant C independent of the mesh such that*

$$(\eta_i^T)^2 \leq C\left(\|\nabla_h \mathbf{u}_h - \nabla\widehat{\mathbf{u}}\|^2_{L^2(T_i)} + \|p_h - \widehat{p}\|^2_{L^2(T_i)}\right) + C(h_i^T)^2 \|\mathbf{f} - \mathbf{f}_i^T\|^2_{L^2(T_i)}, \tag{1.126}$$

$$(\eta_i'^T)^2 \leq C\|\nabla_h \mathbf{u}_h - \nabla\widehat{\mathbf{u}}\|^2_{L^2(T_i)}, \tag{1.127}$$

$$(\eta_k^P)^2 \leq C\left(\|\nabla_h \mathbf{u}_h - \nabla\widehat{\mathbf{u}}\|^2_{L^2(P_k)} + \|p_h - \widehat{p}\|^2_{L^2(P_k)}\right) + C(h_k^P)^2 \|\mathbf{f} - \mathbf{f}_k^P\|^2_{L^2(P_k)}, \tag{1.128}$$

$$(\eta_k'^P)^2 \leq C\|\nabla_h \mathbf{u}_h - \nabla\widehat{\mathbf{u}}\|^2_{L^2(P_k)}, \tag{1.129}$$

$$(\zeta_k^P)^2 \leq C\|\nabla_h \mathbf{u}_h - \nabla\widehat{\mathbf{u}}\|^2_{L^2(P_k)}. \tag{1.130}$$

Proof. Let us consider an element T_i of the primal mesh and a diamond edge s in $\overset{\circ}{T}_i$. Let us recall that by definition, such an edge s does not belong to Γ. Let us consider the function $\mathbf{w}_s = [(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s b_s$, where b_s is defined by (1.108). This function belongs to $(H_0^1(\Omega))^2$ and we may thus apply (1.36), which, taking into account the support of $\mathbf{w}_s$, reduces to

$$\int_{\omega_s} (\nabla \widehat{\mathbf{u}} - I_2 \widehat{p}) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} = \int_{\omega_s} \mathbf{f} \cdot \mathbf{w}_s(\mathrm{x}) d\mathrm{x}. \tag{1.131}$$

Moreover, $\mathbf{u}_h$ belongs to $(P^1(D_j))^2$, p_h is a constant in each D_j and $\mathbf{w}_s$ vanishes on Γ. Thus there holds:

$$\begin{aligned}
\int_\Omega (\nabla_h \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} &= \sum_j \int_{D_j} (\nabla \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} \\
&= \sum_j \int_{\partial D_j} [(\nabla \mathbf{u}_h - I_2 p_h)\mathbf{n}_{\partial D_j}] \cdot \mathbf{w}_s(\sigma) d\sigma \\
&= \sum_i \sum_{s' \subset \overset{\circ}{T}_i} \int_{s'} [(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_{s'}]_{s'} \cdot \mathbf{w}_s(\sigma) d\sigma.
\end{aligned}$$

But since $\mathbf{w}_s$ vanishes on all the other edges $s' \neq s$, taking into account the definition of $\mathbf{w}_s$ and the property of b_s in (1.110), there holds

$$\begin{aligned}
\int_\Omega (\nabla_h \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} &= \int_s [(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s \cdot \mathbf{w}_s(\sigma) d\sigma \\
&= |[(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s|^2 \int_s b_s(\sigma)\sigma \\
&= \frac{2}{3} |s| \, |[(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s|^2 \\
&= \frac{2}{3} \|[(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s\|^2_{L^2(s)}.
\end{aligned} \tag{1.132}$$

And taking into account (1.131), we have

$$\begin{aligned}
M^2 :&= \|[(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s\|^2_{L^2(s)} \\
&= \frac{3}{2} \int_\Omega (\nabla_h \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} = \frac{3}{2} \int_{\omega_s} (\nabla_h \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} \\
&= \frac{3}{2} \left[\int_{\omega_s} (\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}) : \nabla \mathbf{w}_s(\mathrm{x}) d\mathrm{x} - \int_{\omega_s} (p_h - \widehat{p}) \nabla \cdot \mathbf{w}_s(\mathrm{x}) d\mathrm{x} + \int_{\omega_s} \mathbf{f} \cdot \mathbf{w}_s(\mathrm{x}) d\mathrm{x} \right].
\end{aligned} \tag{1.133}$$

Using the Cauchy-Schwarz inequality leads to

$$M^2 \leq \frac{3}{2} \left[\left(\|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|_{L^2(\omega_s)} + \sqrt{2} \|p_h - \widehat{p}\|_{L^2(\omega_s)} \right) \|\nabla \mathbf{w}_s\|_{L^2(\omega_s)} \right] + \frac{3}{2} \|\mathbf{f}\|_{L^2(\omega_s)} \|\mathbf{w}_s\|_{L^2(\omega_s)}. \tag{1.134}$$

Let us now bound $\|\nabla \mathbf{w}_s\|_{L^2(\omega_s)}$ and $\|\mathbf{w}_s\|_{L^2(\omega_s)}$. There holds, thank to (1.114),

$$\|\nabla \mathbf{w}_s\|_{L^2(\omega_s)} = |[(\nabla \mathbf{u}_s - I_2 p_h)\mathbf{n}_s]_s| \|\nabla b_s\|_{L^2(\omega_s)} \leq |[(\nabla \mathbf{u}_s - I_2 p_h)\mathbf{n}_s]_s| C|s|^{-1} \|b_s\|_{L^2(\omega_s)}, \tag{1.135}$$

$$\|\mathbf{w}_s\|_{L^2(\omega_s)} = |[(\nabla \mathbf{u}_s - I_2 p_h)\mathbf{n}_s]_s| \|b_s\|_{L^2(\omega_s)}. \tag{1.136}$$

So there remains to find a bound for $\|b_s\|_{L^2(\omega_s)}$. In order to do this, we first infer from (1.109) that $b_s^2 \leq b_s$. This implies, using (1.112)

$$\|b_s\|_{L^2(\omega_s)} = \left[\|b_s\|^2_{L^2(t_1)} + \|b_s\|^2_{L^2(t_2)}\right]^{1/2} \leq \left[\int_{t_1 \cup t_2} b_s(\mathrm{x}) d\mathrm{x}\right]^{1/2} \leq C|s|. \tag{1.137}$$

Taking into account that $|[(\nabla_h \mathbf{u}_h - I_2 p_h)\mathbf{n}_s]_s| = |s|^{-1/2} M$ and considering (1.134) to (1.137), we obtain

$$M \leq C\left[|s|^{-1/2}(\|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|_{L^2(\omega_s)} + \sqrt{2}\|p_h - \widehat{p}\|_{L^2(\omega_s)}) + |s|^{1/2}\|\mathbf{f}\|_{L^2(\omega_s)}\right]. \tag{1.138}$$

One usually expresses $\|\mathbf{f}\|_{L^2(\omega_s)}$ as a function of $\|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|_{L^2(\omega_s)} + \|p_h - \widehat{p}\|_{L^2(\omega_s)}$ and of higher order terms. Let $t = t_1$ or t_2, and let us denote by $\mathbf{f}_t$ the mean value of $\mathbf{f}$ over t. Then,

$$\|\mathbf{f}\|_{L^2(t)} \leq \|\mathbf{f} - \mathbf{f}_t\|_{L^2(t)} + \|\mathbf{f}_t\|_{L^2(t)}. \tag{1.139}$$

Then, consider $\mathbf{w}_t = \mathbf{f}_t b_t$, where b_t is defined by (1.107). The function $\mathbf{w}_t$ belongs to $(H_0^1)^2$. Thus, taking into account the support of b_t, Eq. (1.36) reduces to

$$\int_t (\nabla \widehat{\mathbf{u}} - I_2 \widehat{p}) : \nabla \mathbf{w}_t(\mathrm{x}) d\mathrm{x} = \int_t \mathbf{f} \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x}. \tag{1.140}$$

Moreover, since $\nabla_h \mathbf{u}_h - I_2 p_h$ is a constant over each t, and since $\mathbf{w}_t$ vanishes on the boundary of t, there holds

$$\int_t (\nabla_h \mathbf{u}_h - I_2 p_h) : \nabla \mathbf{w}_t(\mathrm{x}) d\mathrm{x} = 0. \tag{1.141}$$

Since $\mathbf{f}_t$ is a constant over t, there holds, thanks to (1.111), (1.140) and (1.141),

$$\begin{aligned}
\|\mathbf{f}_t\|^2_{L^2(t)} &= |t| (\mathbf{f}_t)^2 = C (\mathbf{f}_t)^2 \int_t b_t(\mathrm{x}) d\mathrm{x} = C \int_t \mathbf{f}_t \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x} \\
&= C\left[\int_t \mathbf{f} \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x} + \int_t (\mathbf{f}_t - \mathbf{f}) \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x}\right] \\
&= C\left[\int_t (\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h) : \nabla \mathbf{w}_t(\mathrm{x}) d\mathrm{x} - \int_t (\widehat{p} - p_h) \nabla \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x} + \int_t (\mathbf{f}_t - \mathbf{f}) \cdot \mathbf{w}_t(\mathrm{x}) d\mathrm{x}\right] \\
&\leq C\left(\|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(t)} + \sqrt{2}\|\widehat{p} - p_h\|_{L^2(t)}\right) \|\nabla \mathbf{w}_t\|_{L^2(t)} + C\|\mathbf{f}_t - \mathbf{f}\|_{L^2(t)} \|\mathbf{w}_t\|_{L^2(t)},
\end{aligned} \tag{1.142}$$

with $C = 20/9$ in the above expressions. Let us now bound $\|\mathbf{w}_t\|_{L^2(t)}$ and $\|\nabla \mathbf{w}_t\|_{L^2(t)}$. With (1.113), there holds

$$\|\nabla \mathbf{w}_t\|_{L^2(t)} = |\mathbf{f}_t| \|\nabla b_t\|_{L^2(t)} \leq |\mathbf{f}_t| C h_t^{-1} \|b_t\|_{L^2(t)}, \tag{1.143}$$

$$\|\mathbf{w}_t\|_{L^2(t)} = |\mathbf{f}_t| \|b_t\|_{L^2(t)}. \tag{1.144}$$

The remaining term that has to be bounded is $\|b_t\|_{L^2(t)}$. For this, we first infer from (1.109) that $b_t^2(\mathrm{x}) \leq b_t(\mathrm{x})$ and then

$$|\mathbf{f}_t| \, \|b_t\|_{L^2(t)} \leq |\mathbf{f}_t| \left(\int_t b_t(\mathrm{x}) d\mathrm{x} \right)^{1/2} \leq C \, |\mathbf{f}_t| \, |t|^{1/2} = C \, \|\mathbf{f}_t\|_{L^2(t)} \,, \tag{1.145}$$

in which $C = \sqrt{9/20}$. Combining (1.142)–(1.143)–(1.144)–(1.145), we finally get

$$\|\mathbf{f}_t\|_{L^2(t)} \leq C \left(\|\mathbf{f}_t - \mathbf{f}\|_{L^2(t)} + h_t^{-1} \, \|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(t)} + h_t^{-1} \|\widehat{p} - p_h\|_{L^2(t)} \right).$$

Since s is an edge of t, there holds $|s| \leq h_t$; applying (1.139), we obtain

$$\|\mathbf{f}\|_{L^2(t)} \leq C \left(\|\mathbf{f}_t - \mathbf{f}\|_{L^2(t)} + |s|^{-1} \, \|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(t)} + |s|^{-1} \|\widehat{p} - p_h\|_{L^2(t)} \right).$$

Thus, taking into account that $\omega_s = t_1 \cup t_2$, there holds

$$\begin{aligned}
\|\mathbf{f}\|_{L^2(\omega_s)} &\leq \|\mathbf{f}\|_{L^2(t_1)} + \|\mathbf{f}\|_{L^2(t_2)} \\
&\leq C \left(\|\mathbf{f}_{t_1} - \mathbf{f}\|_{L^2(t_1)} + \|\mathbf{f}_{t_2} - \mathbf{f}\|_{L^2(t_2)} \right) + C|s|^{-1} \, \|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(t_1)} \\
&\quad + C|s|^{-1} \left(\|\widehat{p} - p_h\|_{L^2(t_1)} + \|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(t_2)} + \|\widehat{p} - p_h\|_{L^2(t_2)} \right) \\
&\leq C|s|^{-1} \left(\|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(\omega_s)} + \|\widehat{p} - p_h\|_{L^2(\omega_s)} \right) + C \, \|\mathbf{f}_{\omega_s} - \mathbf{f}\|_{L^2(\omega_s)} \,.
\end{aligned} \tag{1.146}$$

In this sequence of inequalities, we have used the fact that $\mathbf{f}_t$ minimizes $\|c - f\|_{L^2(t)}$ when c runs over $\mathbb{R}^2$; in particular, $\|\mathbf{f}_t - \mathbf{f}\|_{L^2(t)} \leq \|\mathbf{f}_{\omega_s} - \mathbf{f}\|_{L^2(t)}$, where $\mathbf{f}_{\omega_s}$ is the mean value of $\mathbf{f}$ over ω_s. Combining (1.138) and (1.146), we obtain

$$\begin{aligned}
M = \|[\nabla_h (\mathbf{u}_h - p_h) \mathbf{n}_s]_s\|_{L^2(s)} &\leq C|s|^{1/2} \, \|\mathbf{f} - \mathbf{f}_{\omega_s}\|_{L^2(\omega_s)} \\
&\quad + C|s|^{-1/2} \left(\|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|_{L^2(\omega_s)} + \|\widehat{p} - p_h\|_{L^2(\omega_s)} \right).
\end{aligned} \tag{1.147}$$

By definition, the local estimator $(\eta_i^T)^2$ is lower than the value taken by the function in (1.81) in $\mu = (h_i^T)^2$. In (1.81), we may bound $C(T_i)$ by $1/\pi$ since the primal cells have been supposed to be convex, and with (1.85) and (1.147), we obtain

$$\begin{aligned}
\left(\eta_i^T\right)^2 &\leq C \left(h_i^T\right)^2 \sum_{s \in \overset{\circ}{T}_i} \frac{|s|^{-1}}{\rho_{ik,1} + \rho_{ik,2}} \left(\|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|^2_{L^2(\omega_s)} + \|p_h - \widehat{p}\|^2_{L^2(\omega_s)} \right) \\
&\quad + C \left(h_i^T\right)^2 \sum_{s \in \overset{\circ}{T}_i} \frac{|s|}{\rho_{ik,1} + \rho_{ik,2}} \, \|\mathbf{f} - \mathbf{f}_{\omega_s}\|^2_{L^2(\omega_s)} \,.
\end{aligned}$$

Using Prop. 1.32, and since by definition $S_{ik} = \frac{1}{2}|s| \, (\rho_{ik,1} + \rho_{ik,2})$ and $|s| \leq h_i^T$, the above inequality leads to (1.126). As far as (1.127) is concerned, let us consider the function $\mathbf{v}_s = [\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s b_s$. There obviously holds

$$\int_\Omega \nabla \widehat{\mathbf{u}} : \nabla \times \mathbf{v}_s(\mathrm{x}) \, d\mathrm{x} = \int_{\omega_s} \nabla \widehat{\mathbf{u}} : \nabla \times \mathbf{v}_s(\mathrm{x}) \, d\mathrm{x} = 0 \,. \tag{1.148}$$

Eq. (1.148) and the calculations that previously led to (1.132) may be used to yield

$$\begin{aligned}\|[\nabla_h \mathbf{u}_h \boldsymbol{\tau}_s]_s\|^2_{L^2(s)} &= \frac{3}{2}\int_{\omega_s} \nabla_h \mathbf{u}_h : \nabla \times \mathbf{v}_s(\mathrm{x}) d\mathrm{x} \\ &= \frac{3}{2}\int_{\omega_s} (\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}) : \nabla \times \mathbf{v}_s(\mathrm{x}) d\mathrm{x} \\ &\leq \frac{3}{2} \|\nabla_h \mathbf{u}_h - \nabla \widehat{\mathbf{u}}\|_{L^2(\omega_s)} \|\nabla \mathbf{v}_s\|_{L^2(\omega_s)} . \end{aligned} \tag{1.149}$$

Just like (1.134) led to (1.138) and then to (1.126), the inequality (1.149) leads to (1.127). The dual inequalities (1.128), (1.129) and (1.130) may be obtained in the same way. We note that in (1.130), we obtained the result so far only in the case $\mathbf{g} = 0$. The proof is very similar to that of (1.127) and (1.129), but some definitions have to be changed because the segment $b_{j_\alpha(k)}$ in the definition (1.93) is a boundary segment, and is thus the edge of only one triangle t; the function b_s is thus defined only in that triangle t. □

1.7 Numerical results

In the part of numerical experiments, first, we study the influence of the parameter ε for a fixed mesh and of the mesh size for a fixed value of the penalty parameter. Secondly, we give an overall process to recursively adapt the value of the penalty parameter and the mesh refinement.

1.7.1 Influence of the penalty parameter

In this subsection, we will work on the domain $\Omega = [0;1]^2$. A triangular mesh with rather uniform triangles is used. The exact solution $(\widehat{\mathbf{u}}, \widehat{p})$ is regular with $\widehat{\mathbf{u}} = (\partial_y \varphi, -\partial_x \varphi)$ given by

$$\varphi(x,y) = 100x^2y^2(1-x)^2(1-y)^2 \text{ and } \widehat{p}(x,y) = 10(x^2+y^2-\frac{2}{3}). \tag{1.150}$$

Fig. 1.7 presents the plots of the errors and the estimators when the penalty parameter ε goes from 10^{-2} to 10^{-8}. They include the actual errors in the $H^1(\Omega)$ and $L^2(\Omega)$ norms for the velocity, i.e. the error in the velocity gradient $\|\nabla \widehat{\mathbf{u}} - \nabla_h \mathbf{u}_h\|_{L^2(\Omega)}$ and in the velocity $\|\widehat{\mathbf{u}} - \mathbf{u}_h\|_{L^2(\Omega)}$, the total estimator, the discretization estimator and the penalization estimator which are given by Theorem 1.26 when we estimate the velocity error. The left (resp. right) figure corresponds to the mesh size $h = 5.69 \times 10^{-2}$ (resp. $h = 3.125 \times 10^{-2}$). We see that for a given mesh, the ratio between the penalization estimator and the penalty parameter ε asymptoticaly tends to a constant, while the discretization estimator is nearly independent of ε. Moreover, the actual errors decrease with ε until a certain level. Then, the discretization error is the dominant error and decreasing ε further does not have any influence on the overall error. As expected, when the mesh size is smaller (right part of the Figure), then the value of the penalty parameter for which the errors saturate is also smaller.

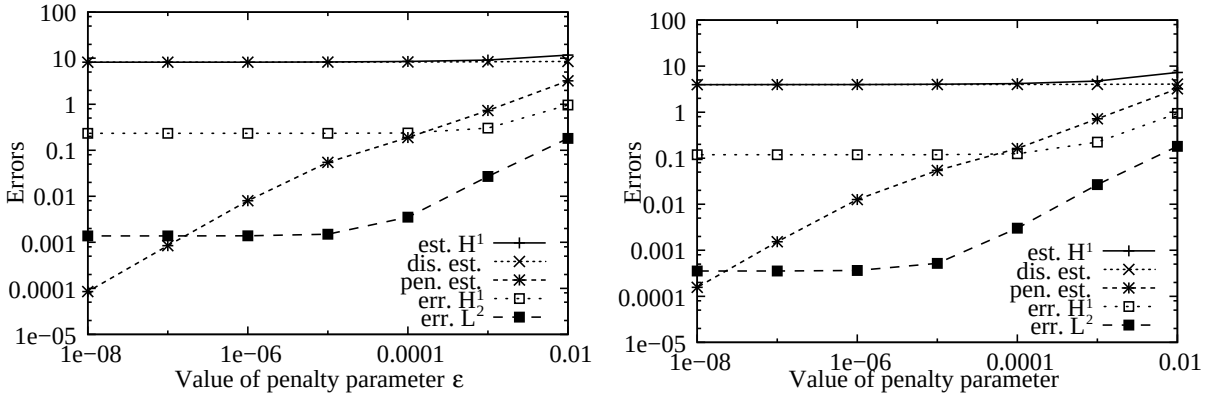

Figure 1.7: Actual errors in $H^1(\Omega)$ and $L^2(\Omega)$ norms, total estimator, discritization estimator and penalization estimator for the velocity. Left: $h = 5.69 \times 10^{-2}$, right: $h = 3.125 \times 10^{-2}$.

1.7.2 Influence of the mesh size

On the same square domain Ω and with the same exact solution as previously, we work with a fixed $\varepsilon = 10^{-3}$ (resp. $\varepsilon = 10^{-7}$) in the left (resp. right) part of Fig. 1.8. Since the solution is regular, only uniformly refined triangular meshes will be considered. Figure 1.8 presents the same curves as in Fig. 1.7, but now as a funtion of h, varying from 0.25 to 1.6×10^{-2}.

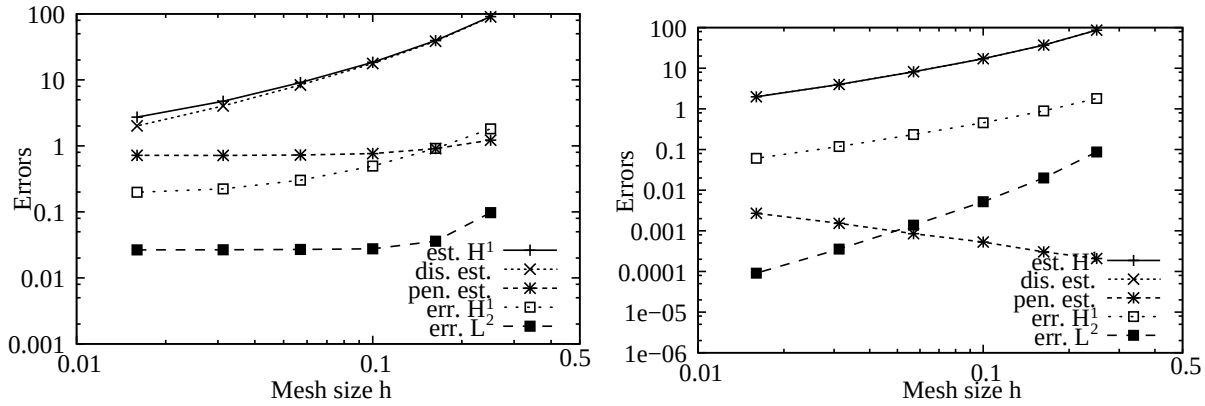

Figure 1.8: Actual errors in $H^1(\Omega)$ and $L^2(\Omega)$ norms, total estimator, discritization estimator and penalization estimator for the velocity. Left: $\varepsilon = 10^{-3}$, right: $\varepsilon = 10^{-7}$.

The actual errors decrease until the mesh size h is so small that the penalization error will dominate the discretization error, and the total error thus stagnates to a certain level. The penalization estimator is nearly independent of h in the left figure but behaves roughly like h^{-1} in the right figure. This behaviour remains unexplained and further investigations have to be conducted about this. The total estimator and the discretization estimator decrease regularly when h decreases, roughly like h, but, then, when h is small enough the total estimator starts to stagnate because the penalization estimator stops being negligeable (better seen in Fig. 1.9, where $\varepsilon = 10^{-2}$).

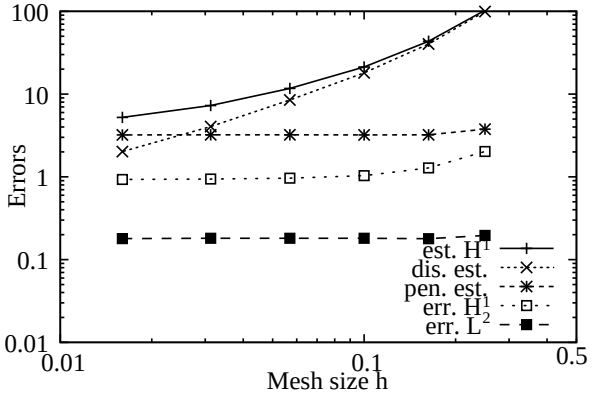

Figure 1.9: Actual errors in $H^1(\Omega)$ and $L^2(\Omega)$ norms, total estimator, discritization estimator and penalization estimator for the velocity for $\varepsilon = 10^{-2}$.

1.7.3 Adaptive penalty parameter and mesh

We propose the following computational process. We start with a given coarse mesh and an initial value of ε, and we fix some ratio $0 < \gamma \leq 1$.

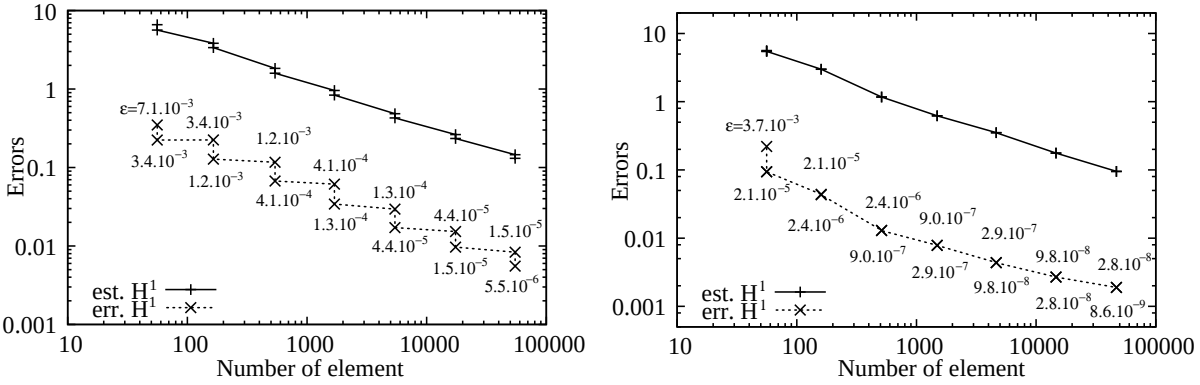

Figure 1.10: Actual errors in $H^1(\Omega)$ and total estimator for the velocity. Left: $\gamma = 1/10$, right: $\gamma = 1/500$.

Then, we compute the numerical solution, and we get η_h and η_ϵ. Then,

- If $\eta_\epsilon \geq \gamma\eta_h$, we adapt a new ε by multiplying the old ε with the ratio $\frac{\gamma\eta_h}{2\eta_\epsilon}$ and keep the same mesh for a new computation. This has the effect of maintaining the error due to the penalization below a certain ratio of the error due to the discretization.

- Otherwise, we adaptively refine the mesh based on the discretization estimator η_h. For this, on the given mesh, we compute the local discretization estimators $\eta_{i,h}$ with $\eta_h^2 = \sum_i \eta_{i,h}^2$ and ask to refine a given primal cell T_i by a factor 4 in terms of area if $\eta_{i,h} \geq (\max_i \eta_{i,h})/2$.

The test we present to illustrate this strategy is also on the domain $\Omega = [0;1]^2$, The exact solution $(\widehat{\mathbf{u}}, \widehat{p})$ is regular with $\widehat{\mathbf{u}} = (\varphi_y, -\varphi_x)$, and φ is given by

$$\varphi(x,y) = x^2(1-x)^2y^2(1-y)^2 \text{ and } \widehat{p}(x,y) = 5(x^2+y^2-\frac{2}{3}). \tag{1.151}$$

For accuracy reasons, the ratio γ may be chosen so that the penalization error is much lower than the discretization error like in the right figure of Fig. 1.10 obtained with $\gamma = 1/500$. We observe that the actual error and the total estimator are not affected by the penalty term. Moreover, we made a test with $\gamma = 1/10$ and we present the result in the left ficture of Fig. 1.10 to show the interplay between the mesh refinement and the decrease of ε.

In the following test, we will compare the exact error and total estimator in $H^1(\Omega)$ for uniform and adaptive refinements. We will combine this work with the adaptive penalty and mesh with $\gamma = 1/500$. Our test is in the domain $\Omega = [0,1[^2$ and the exact couple solution $(\widehat{\mathbf{u}}, \widehat{p})$ is singular with $\widehat{\mathbf{u}} = (\varphi_y, -\varphi_x)$, and φ is

$$\varphi(x,y) = x^{\frac{7}{4}}(1-x)^2y^2(1-y)^2 \text{ and } \widehat{p}(x,y) = \frac{x+y-1}{10}.$$

We observe that the velocity $\widehat{\mathbf{u}}$ is in $[H^{\frac{5}{4}}(\Omega)]^2$ and there is a boundary singularity on the edge $x = 0$.

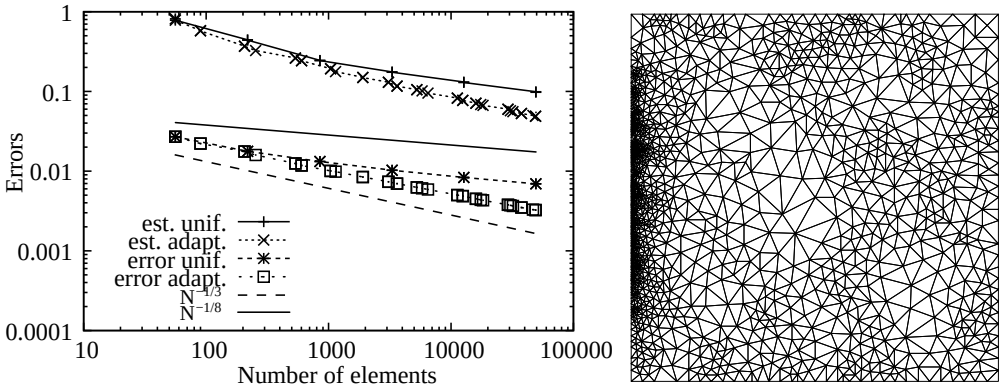

Figure 1.11: Estimated and exact errors in uniform/adaptive refinement (left) and an adaptived mesh (right).

In Fig. 1.11, the penalty parameter ε decreases from 10^{-3} to 7.98×10^{-11} for the adaptive refinement (from 10^{-3} to 1.80×10^{-7} for the uniform refinement). The curve of convergence rate corresponding to the uniform mesh refinement is parallel to $N^{-1/8}$ curve, while the curve of convergence rate corresponding to the adaptive mesh refinement is parallel to $N^{-1/3}$ curve. Moreover, the effectivity of both refinements is almost 15.

Normally, the plot of exact error corresponding to a adaptive mesh refinement is paralell to $N^{-1/2}$, but in our singular case, it is not satisfied. We would like to determine that this problem is caused by our dicretization estimator or not, we include in Fig. 1.12. We compare two exact error corresponding two adaptive refinement process which are driven

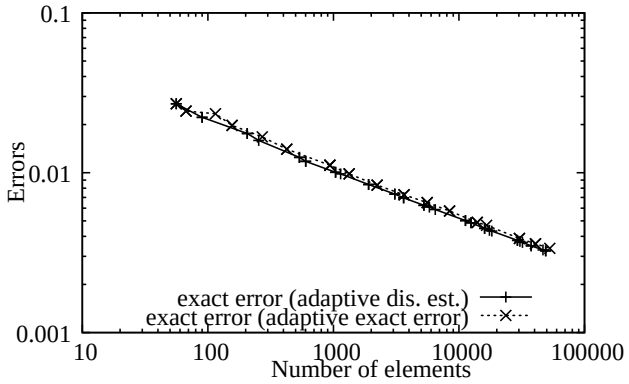

Figure 1.12: Exact errors in the adaptive process using the discretization estimator and exact error.

by our dicretization estimator and the exact error. Clearly, the exact error is not mostly affected by the applied adaptive process.

References for the Chapter 1

[1] C. Bernardi, V. Girault, F. Hecht, *A Posteriori Analysis of a Penalty Method and Application to the Stokes Problem*. Mathematical Models and Methods in Applied Sciences, Vol. 13, No. 11, pp. 1599–1628 (2003).

[2] C. Carstensen and S. Funken, *Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods*, East-West J. Numer. Math. 8, pp. 153–175 (2000).

[3] S. Delcourte, K. Domelevo, P. Omnes, *A Discrte Duality Finite Volume Approach to Hodge Decomposition and Div–Curl Problems on Almost Arbitrary Two–Dimensional Meshes*. Siam J. Numer. Anal. Vol. 45, No.3, pp. 1142–1174 (2007).

[4] E. Dari, R. Durán and C. Padra, *Error Estimators for Nonconforming Finite Element Approximations of the Stokes Problem*. Math. Comp. 64, pp. 1017-1033 (1995).

[5] Y. Girault, P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations*. Springer-Verlag, Berlin, Heidelber, New York, Tokyo (1986).

[6] A. Hannukainen, R. Stenberg, M. Vohralík *A Unified Framework for a Posteriori Error Estimation for the Stokes Problem*. Available at http://hal.archives-ouvertes.fr/hal-00470131/fr/.

[7] S. Krell, *Stabilized DDFV Schemes for Stokes Probblem with Variable Viscosity on General 2D Meshes*. Available at http://hal.archives-ouvertes.fr/hal-00385687/fr/.

[8] P. Omnes, Y. Penel, and Y. Rosenbaum, *A Posteriori Error Estimation for the Discrete Duality Finite Vomume Discretization of the Laplace Equation*. SIAM J. Numer. Anal. 47, pp. 2782–2807 (2009).

[9] L.E Payne, H.F Weiberger, *An optimal Poincaré inequality for convex domain*. Arch. Rational Mech. Anal . 5, pp. 286–292 (1960).

[10] R. Verfürth, *A Posteriori Error Estimation for the Stokes Equations*. Numer. Math. 55, pp. 309–325 (1989).

[11] R. Verfürth, *A Posteriori Error Estimation for the Stokes Equations II noncomforming discretizations*. *Numer. Math. 60, pp. 235–249 (1991).

[12] R. VERFÜRTH, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Teubner–Wiley, Stuttgart (1996).

[13] R. Verfürth, *Error estimates for some quasi-interpolation operators*. ESAIM:M2AN. 33, pp. 695–713 (1999).

Chapter 2

Nonlinear Darcy Equations in Two dimensions

We present in this chapter several schemes for the discrete solution of nonlinear diffusion equations along with related a posteriori error estimation. The estimator includes two terms: discretization and linearization estimators. Hence, the iterative linearization can be stopped whenever the linearization estimator drops below a fraction of the discretization estimator. Thus, this stopping criterion ensures that the actual error and the estimators are only affected by the space discretization. Moreover, this leads to computational savings, because it avoids unnecessary linearization iterations. Numerical tests are performed with several types of nonlinear diffusion equations.

2.1 Introduction

Let Ω be an open bounded polygonal subset of $\mathbb{R}^2$, $\Gamma = \partial\Omega$, f be a given function from Ω to $\mathbb{R}$ and H a given function specified below. We consider an approximate solution of the following nonlinear equation:

$$\begin{cases} -\mathrm{div}(H(\hat{u})\nabla\hat{u})(x) &= f(x), \quad x \in \Omega, \\ \hat{u}(x) &= 0, \quad x \in \Gamma. \end{cases} \tag{2.1}$$

This nonlinear diffusion equation appears in several physical models such as Darcy flows in porous medias.

Let us give some assumptions on this problem.

A1. Let $H : \mathbb{R} \mapsto \mathbb{R}$ be continuous such that there exist constants C_1, $C_2 > 0$ satisfying

$$C_1 \leq H(\hat{u}) \leq C_2, \quad \text{for all } \hat{u} \in \mathbb{R}. \tag{2.2}$$

Moreover, we will assume that there exists a positive constant C such that

$$|H(\hat{u}) - H(\hat{v})| \leq C|\hat{u} - \hat{v}|, \quad \text{for all } (\hat{u}, \hat{v}) \in \mathbb{R}^2. \tag{2.3}$$

A2. Let f belong to $L^2(\Omega)$.

Under assumptions (A1) and (A2), there exist $\hat{u} \in V = H_0^1(\Omega)$ such that $\hat{u}$ is the unique weak solution of problem (2.1), in the sense that

$$\int_\Omega H(\hat{u})\nabla\hat{u} \cdot \nabla\hat{v}(x)dx = \int_\Omega f\hat{v}(x)dx \quad \text{for all } \hat{v} \in V. \tag{2.4}$$

The existence of weak solution of this problem is proved in [9] and results about uniqueness of the solution may be found in [4].

The discrete solution of the nonlinear diffusion equation is investigated by the mixed finite element method in [12], and, more generally, in [15]. More recently, finite volume discretizations have been developed for linear diffusion equations, such as finite volume schemes on admissible meshes [7], DDFV schemes [14] and multipoint flux approximation (MPFA) schemes [1] on arbitrary meshes, etc. Then we are interested in solving the nonlinear diffusion equations by such finite volume discretizations. In this work, we will deal with the above three methods. The discretization process of the problem leads to a system of nonlinear equations, which is linearized by the fixed point method. These schemes have a local conservation property, which is the important ingredient to obtain an a posteriori error estimation.

The theory of a posteriori error estimation is not very developed for the nonlinear diffusion equation. L. El Alaoui et al. [6] have obtained a posteriori error estimate for a finite element method in the p-Laplace case. D. Kim et al. [11] gave an estimate for the mixed finite element discretization. As far as finite volume methods are considered, although there are not many important results for the nonlinear diffusion equations, essential development steps on a posteriori error estimation have been achieved for linear diffusion equations. Nicaise [13] gave a posteriori error estimation for Morley-type interpolations of the original piecewise constant finite volume approximation. P. Omnes et al. [14] used the equivalence of the DDFV scheme with a finite element like method to derive fully computable a posteriori error estimate for this method applied to the Laplace equation. For the local conservative methods, M. Vohralík [19] built an approximate function which depends only on the flux through the edges of the cells and the finite volume unknowns at some control points in the cells and provides error estimation with respect to this reconstruction. In addition, the effectivity index is very close to one, which demonstrates an accurate estimation. M. Vohralík's idea will be applied in this work to estimate the error of finite volume techniques applied to the non-linear diffusion equations.

Given a discrete solution, a stage of iterative process and a mesh, our a posteriori error estimation is split into 2 terms: the discretization and linearization estimators. This splitting has two main advantages. The first one is that, in practice, when the number of iterations is large enough, the linearization estimator is negligible compared with the discretization one. Thus the balance of these two estimators is an important key to avoid performing an excessive number of nonlinear solver iterations. The other one is that the mesh refinement is then only based on the discretization estimator. This type of analysis is considered in [6] for the p-Laplace case. In this work, we do not deal with the convergence of the linearization iterations and of the discretization processes. We only mention them in our numerical experiments.

This chapter is organized as follows. Section 2.2 sets some notations and definitions related to the meshes, the non-linear problem together with its discretization and linearization.

Section 2.3 is devoted to our a posteriori error estimation, and its efficiency is verified in Section 2.4. In Section 2.5, we present results of some numerical tests.

2.2 Construction of the schemes

2.2.1 Notations and definitions

We will follow the definition of general meshes given in [14]. Let Ω be covered by a primal mesh with cells denoted by T_i, $i \in [1, I]$. To each T_i, we associate a point G_i located in the interior of T_i and $V(i)$ is set of the vertices of T_i. With any vertex S_k, with $k \in [1, K]$, we associate a dual cell P_k by joining points G_i associated with the primal cells surrounding S_k to the midpoints of the edges of which S_k is a node. The notations are summarized in Fig. 2.1 and Fig. 2.2.

Let h_i denote the diameter of T_i and ρ_i denote the diameter of the largest ball inscribed in T_i. We make the following shape regularity assumption on the mesh:

Assumption A (shape-regularity of the meshes). There exists a positive constant θ independent of the mesh such that $\max_{i\in[1,I]} h_i/\rho_i \leq \theta$.

With any primal edge A_j with $j \in [1, J]$, we associate a so-called diamond-cell D_j obtained by joining the vertices $S_{k_1(j)}$ and $S_{k_2(j)}$ of A_j to the points $G_{i_1(j)}$ and $G_{i_2(j)}$ associated with the primal cells $T_{i_1(j)}$ and $T_{i_2(j)}$ that share A_j as a part of their boundaries. When A_j is a boundary edge (there are J^Γ such edges), the associated diamond-cell is a flat quadrilateral (i.e. a triangle) and we denote by $G_{i_2(j)}$ the midpoint of A_j (thus, there are J^Γ such additional points G_i). The unit normal vector to A_j is $\mathbf{n}_j$ and points from $G_{i_1(j)}$ to $G_{i_2(j)}$. We denote by A'_{j1} (resp. A'_{j2}) the segment joining $G_{i_1(j)}$ (resp. $G_{i_2(j)}$) and the midpoint of A_j. Its associated unit normal vector, pointing from $S_{k_1(j)}$ to $S_{k_2(j)}$, is denoted by $\mathbf{n}'_{j1}$ (resp. $\mathbf{n}'_{j2}$). In the case of a boundary diamond-cell, A'_{j2} reduces to $\{G_{i_2(j)}\}$ and does not play any role. Finally, for any diamond-cell D_j, we shall denote by $M_{i_\alpha k_\beta}$ the midpoint of $[G_{i_\alpha(j)} S_{k_\beta(j)}]$, with $(\alpha, \beta) \in \{1; 2\}^2$. For any primal T_i, such that $A_j \subset \partial T_i$, we denote $\mathbf{n}_{ji} := \mathbf{n}_j$ if $i = i_1(j)$ and $\mathbf{n}_{ji} := -\mathbf{n}_j$ if $i = i_2(j)$ so that $\mathbf{n}_{ij}$ is always exterior to T_i.

Moreover, for the MPFA O scheme, we will set some more definitions. For any $j \in [1, J]$, and any $k \in [1, K]$ such that $S_k \in \partial A_j$, we associate the segment A_j^k which is obtained by joining the vertex S_k and the midpoint M_j of A_j. For any $i \in [1, I]$, $k \in V(i)$, we denote by $A_{j_1}^{i,k}, A_{j_2}^{i,k}$ the two boundary segments of T_i having S_k as common vertex, and $M_{j_1}^{i,k}, M_{j_2}^{i,k}$ as midpoints. Let $\nu_{j_1}^{i,k}$ (resp. $\nu_{j_2}^{i,k}$) be the interior unit normal vector located on the segment opposite to vertex $M_{j_1}^{i,k}$ (resp. $M_{j_2}^{i,k}$) in the triangle $G_i M_{j_1}^{i,k} M_{j_2}^{i,k}$ (see Fig. 2.3).

Finally, we shall also consider the case in which the primal mesh $(T_i)_{i\in[1,I]}$ is an admissible mesh (see [7]). In that case, for $i \in [1, I]$, such that $A_j \subset \partial T_i$, let $d_{i,j}$ be the distance between the control point G_i and the segment A_j.

By a slight abuse of notations, we shall write $k \in \Gamma$ if the vertex S_k belongs to Γ. Identically, we shall write $i \in \Gamma$ (resp. $j \in \Gamma$) if $G_i \in \Gamma$ (resp. $A_j \subset \Gamma$).

We recall here the discrete differential operators which have been constructed on fairly general two dimensional meshes and some of their properties. For more details, see [5, 1].

Definition 2.1. *Let $\mathbf{u} = (\mathbf{u}_j)$ be in $(\mathbb{R}^2)^J$. We define its discrete divergence on primal and*

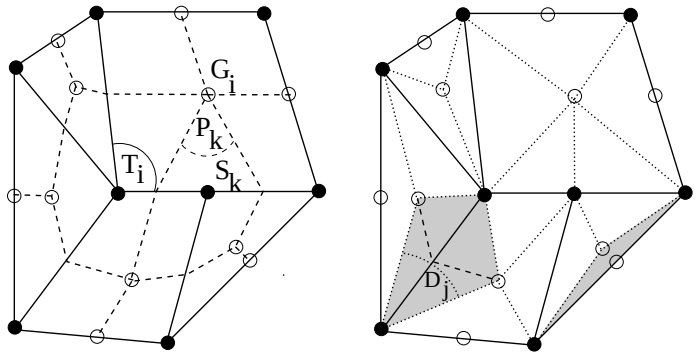

Figure 2.1: A nonconforming primal mesh and its associated dual mesh (left) and diamond-mesh (right).

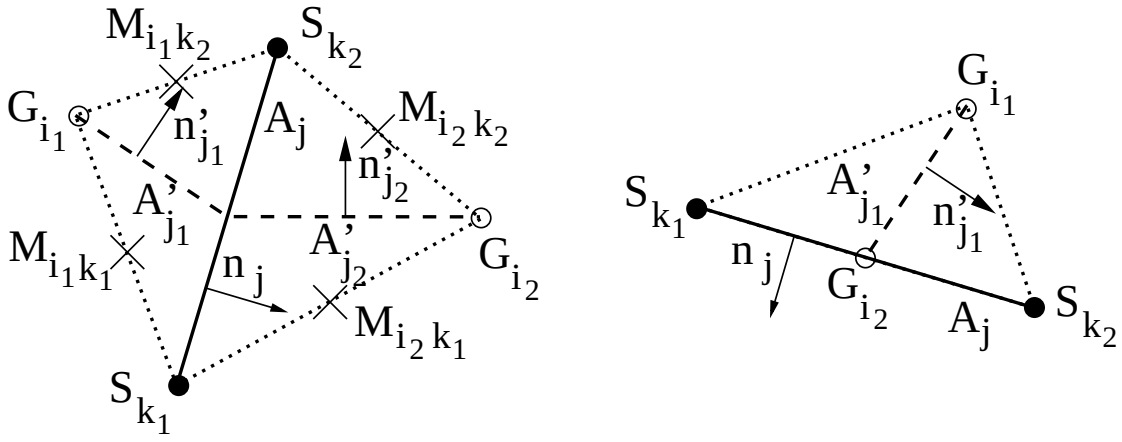

Figure 2.2: Notations for the inner diamond-cell (left) and a boundary diamond mesh (right).

dual cells by

$$(\nabla_h^T \cdot \mathbf{u})_i := \frac{1}{|T_i|} \sum_{j \in \partial T_i} |A_j| \mathbf{u}_j \cdot \mathbf{n}_{ji},$$

$$(\nabla_h^P \cdot \mathbf{u})_k := \frac{1}{|P_k|} \left(\sum_{j \in \partial P_k} (|A'_{j1}| \mathbf{u}_j \cdot \mathbf{n}'_{j1k} + |A'_{j2}| \mathbf{u}_j \cdot \mathbf{n}'_{j2k}) + \sum_{j \in \partial P_k \cap \Gamma} \frac{|A_j|}{2} \mathbf{u}_j \cdot \mathbf{n}_j \right).$$

Definition 2.2. *Let $\phi = (\phi_i^T, \phi_k^P)$ be in $\mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$, its discrete gradient $\nabla_h^D \phi$ is defined by its values on the cells D_j by*

$$(\nabla_h^D \phi)_j := \frac{1}{2|D_j|} \Big\{ [\phi_{k_2}^P - \phi_{k_1}^P](|A'_{j1}| \mathbf{n}'_{j1} + |A'_{j2}| \mathbf{n}'_{j2}) + [\phi_{i_2}^T - \phi_{i_1}^T] |A_j| \mathbf{n}_j \Big\}.$$

Definition 2.3. *Let $\phi = (\phi_i^T, \phi_j^k) \in \mathbb{R}^I \times \mathbb{R}^{2J}$ with $k \in \partial A_j$, its discrete gradient $\nabla_h^{i,k} \phi$ is*

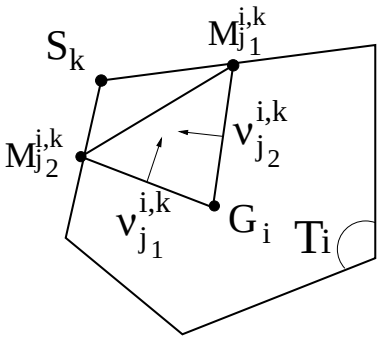

Figure 2.3: Notations for the triangle $G_i M_{j_1}^{i,k} M_{j_2}^{i,k}$.

defined by its values on the triangle $G_i M_{j_1}^{i,k} M_{j_2}^{i,k}$ *by*

$$\nabla_h^{i,k}\phi = \frac{1}{2|G_i M_{j_1}^{i,k} M_{j_2}^{i,k}|}\Big\{(\phi_{j_1}^k - \phi_i^T)\nu_{j_1}^{i,k}|G_i M_{j_2}^{i,k}| + (\phi_{j_2}^k - \phi_i^T)\nu_{j_2}^{i,k}|G_i M_{j_1}^{i,k}|\Big\}.$$

We will define the broken Sobolev space which we use to compute a posteriori error estimation.

Definition 2.4 (Broken Sobolev space)**.**

$$H_h^1(\Omega) := \Big\{\phi \in L^2(\Omega); \phi|_{T_i} \in H^1(T_i)\ \forall i \in [1, I]\Big\}. \tag{2.5}$$

2.2.2 The schemes and their linearizations by a fixed point method

Let us now define three schemes. We set

$$\begin{aligned}
V_h^D = \Big\{ & u_h = ((u_{h,i}^T), (u_{h,k}^P)) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K / \text{ s.t.} \\
& u_{h,i}^T = 0\ ,\ \forall i \in \Gamma \text{ and } u_{h,k}^P = 0\ ,\ \forall k \in \Gamma\Big\}, \\
V_h^O = \Big\{ & u_h = ((u_{h,i}^T), (u_{h,j}^k)) \in \mathbb{R}^I \times \mathbb{R}^{2J} / \text{ s.t. } u_{h,j}^k = 0\ ,\ \forall j \in \Gamma, S_k \in \partial A_j\Big\}, \\
V_h = \Big\{ & u_h = ((u_{h,i}), u_{h,j}) \in \mathbb{R}^I \times \mathbb{R}^J / \text{ s.t. } u_{h,j} = 0\ ,\ \forall j \in \Gamma\Big\}.
\end{aligned}$$

At a given stage m of the iterative scheme, if we know an approximate solution u_h^m, we define some $u_h^{m,*} \in V_h^D$ (resp. V_h^O, V_h) at which the nonlinear function H is linearized. How $u_h^{m,*}$ is chosen depends on the method considered, for example one may choose $u_h^{m,*} = u_h^m$ or $u_h^{m,*} = \alpha^m u_h^{m-1} + (1-\alpha^m)u_h^m$ with $\alpha^m \in [0, 1)$ (relaxation method), etc. Then the discrete linearized scheme is: find $u_h^{m+1} \in V_h^D$ (resp. V_h^O, V_h) such that

$$\text{DDFV}\begin{cases}
-(\nabla_h^T \cdot H(u_h^{m,*})(\nabla_h^D u_h^{m+1}))_i = f_i^T \quad \forall i \in [1, I], \\
-(\nabla_h^P \cdot H(u_h^{m,*})(\nabla_h^D u_h^{m+1}))_k = f_k^P \quad \forall k \notin \Gamma, \\
u_{h,j}^{m,*} = \dfrac{(u^{m,*})_{h,k_2(j)}^P + (u_h^{m,*})_{h,k_1(j)}^P}{2}.
\end{cases} \tag{2.6}$$

$$\text{MPFA O}\begin{cases} \dfrac{-1}{|T_i|}\displaystyle\sum_{j\in\partial T_i}\sum_{k\in\partial A_j}|A_j^k|H((u_h^{m,*})_j^k)\nabla_h^{i,k}u_h^{m+1}\cdot\mathbf{n}_{ji}=f_i^T \quad \forall i\in[1,I], \\ \\ \nabla_h^{i_1(j),k}u_h^{m+1}\cdot\mathbf{n}_{ji_1(j)}+\nabla_h^{i_2(j),k}u_h^{m+1}\cdot\mathbf{n}_{ji_2(j)}=0 \quad \forall j\notin\Gamma, k\in\partial A_j. \end{cases} \tag{2.7}$$

$$\text{FV}\begin{cases} \dfrac{-1}{|T_i|}\displaystyle\sum_{j\in\partial T_i}|A_j|H(u_{h,j}^{m,*})\frac{u_{h,j}^{m+1}-u_{h,i}^{m+1}}{d_{i,j}}=f_i^T \quad \forall i\in[1,I], \\ \\ \dfrac{u_{h,j}^{m+1}-u_{h,i_1(j)}^{m+1}}{d_{i_1(j),j}}+\dfrac{u_{h,j}^{m+1}-u_{h,i_2(j)}^{m+1}}{d_{i_2(j),j}}=0 \quad \forall j\notin\Gamma, \end{cases} \tag{2.8}$$

where f_i^T and f_k^P are the mean-values of f over T_i and P_k, respectively. For a given mesh, following the idea of the authors [6], in practice, starting from an initial guess $u_h^{0,*}\in V_h$, a sequence of discrete solutions $(u_h^m)_{m\geq 1}\in V_h$ is generated through the following iterative algorithm: for $m\geq 0$,

(1) Linearize the nonlinear problem at $u_h^{m,*}$.

(2) Solve the discrete linearized problem (2.6), (2.7) or (2.8) for u_h^{m+1}.

(3) If desired precision is reached, then stop. Else set $m\leftarrow(m+1)$ and go to step (1).

We assume this process will be stopped; it means that there exists $N_h\in\mathbb{N}$ such that the discrete solution $u^{N_h}\in V_h$ satisfies the desired precision. It is convenient to set $u_h=u^{N_h}$ and $w_h=u_h^{N_h-1,*}$; then in all three schemes above, u_h and w_h satisfy that

$$\begin{aligned} &\sum_{j\in\partial T_i}|A_j|F_{ji}(w_h,u_h)=-|T_i|f_i^T \quad \forall i\in[1,I], \\ &F_{ji_1(j)}(w_h,u_h)+F_{ji_2(j)}(w_h,u_h)=0 \quad \forall j\in[1,J-J^\Gamma], \end{aligned} \tag{2.9}$$

where $F_{ji_1(j)}$ and $F_{ji_2(j)}$ are the fluxes through the edge A_j shared by the two cells $T_{i_1(j)}$ and $T_{i_2(j)}$, where for $i\in[1,I]$ and $j\in\partial T_i$, we denote by F_{ji} the flux from control volume T_i through the segment A_j:

$$F_{ji}(w_h,u_h)=\begin{cases} H(w_{h,j})(\nabla_h^D u_h)_j\cdot\mathbf{n}_{ji} & \text{for the DDFV scheme,} \\ \\ \dfrac{1}{2}\displaystyle\sum_{k\in\partial A_j}H(w_{h,j}^k)\nabla_h^{i,k}u_h\cdot\mathbf{n}_{ji} & \text{for the MPFA O scheme,} \\ \\ H(w_{h,j})\dfrac{u_{h,j}-u_{h,i}}{d_{ij}} & \text{for the FV scheme.} \end{cases} \tag{2.10}$$

2.3 An a posteriori error estimate

In this part, we restrict to the case where all (primal) cells T_i are triangular. An extension to non triangular cells may be performed exactly like in [19] by dividing each cell into triangles by joining its center to its vertices, and by solving a local scheme discretizing the original equation with Neumann boundary conditions coming from the numerical scheme on the global mesh.

2.3.1 The construction of an approximate function

In order to define an approximate function whose gradient is a good approximation of $H(\hat{u})\nabla\hat{u}$ and which depends on the discrete solutions u_h and w_h, we will rely on the ideas of R. Eymard et al. [8]. For any $i \in [1, I]$ and $j \in \partial T_i$, let $\phi_{i,j} \in H^1(T_i)/\mathbb{R}$ be the variational solution of the following Neumann problem

$$\begin{aligned}
&\Delta\phi_{i,j}(\mathrm{x}) = \frac{|A_j|}{|T_i|} \quad \text{for a.e } \mathrm{x} \in T_i,\\
&\nabla\phi_{i,j}(\mathrm{y}) \cdot \mathbf{n}_{ji} = 1 \quad \text{for a.e } \mathrm{y} \in A_j,\\
&\nabla\phi_{i,j}(\mathrm{y}) \cdot \mathbf{n}_{j_1 i} = 0 \quad \text{for a.e } \mathrm{y} \in A_{j_1}, j_1 \in \partial T_i,\ j_1 \neq j.
\end{aligned}$$

Since $(T_i)_{i\in[1,I]}$ is a family of triangular cells, following [17] and [8], the functions $\phi_{i,j}$ are the usual polynomial basis functions for the flux in the classical mixed finite element approximation, and their gradient may be written explicitly. Let a be the vertex which is opposite to segment A_j. Then,

$$\nabla\phi_{i,j}(\mathrm{x}) = \frac{\mathrm{x} - \mathrm{a}}{d(\mathrm{a}, A_j)} \text{ for any } \mathrm{x} \in T_i,$$

where $d(\mathrm{x}, A_j)$ denotes the distance between A_j and a.

Next we define the functions $M_i, W_i \in H^1(T_i)$ by: for all $\mathrm{x} \in T_i$

$$\begin{aligned}
M_i(\mathrm{x}) &= \sum_{j\in\partial T_i} F_{ji}(u_h, u_h)\phi_{i,j}(\mathrm{x}) + C_i,\\
W_i(\mathrm{x}) &= \sum_{j\in\partial T_i} F_{ji}(w_h, u_h)\phi_{i,j}(\mathrm{x}).
\end{aligned}$$

For any $i \in [1, I]$, we choose the constant C_i such that

$$M_i(G_i) = \int_0^{u_{h,i}} H(t)dt,$$

where $u_{h,i}$ is the value of u_h on the primal cell T_i and we also define $M \in H^1_h(\Omega)$ by

$$M(\mathrm{x}) = M_i(\mathrm{x}), \text{ for all } \mathrm{x} \in T_i \text{ and } i \in [1, I]. \tag{2.11}$$

It is clear from the definition of W_i, that for any $i \in [1, I]$

$$\nabla W_i \cdot \mathbf{n}_{ji}(\mathrm{x}) = F_{ji}(w_h, u_h) \quad \text{for all } \mathrm{x} \in A_j, j \in \partial T_i. \tag{2.12}$$

and

$$\Delta W_i(\mathrm{x})|_{T_i} = \frac{1}{|T_i|} \sum_{j\in\partial T_i} |A_j| F_{ji}(w_h, u_h) = -f_i^T. \tag{2.13}$$

2.3.2 Poincaré inequality

Lemma 2.5. *Let ω be a polygonal domain. Let $\varphi \in H^1(\omega)$ and let φ_ω be the mean-value of φ over ω. Then, there exists a constant C_ω depending only on the shape of ω such that*

$$\|\varphi - \varphi_\omega\|_{L^2(\omega)} \leq C_\omega diam(\omega) \|\nabla\varphi\|_{L^2(\omega)}. \tag{2.14}$$

Note that when ω is convex, a constant c_ω is given by $\frac{1}{\pi}$; for more details, see [16].

2.3.3 Averaging interpolation operator

Let $P_h^2(\Omega)$ denote the space of polynomials of degree at most 2 on each element. The averaging interpolation operator $\mathcal{I}_{Os} : P_h^2(\Omega) \to P_h^2(\Omega) \cap H_0^1(\Omega)$ has been considered in [10]. Given a function $\varphi_h \in P_h^2(\Omega)$, the value of $I_{Os}(\varphi_h)$ at a Lagrangian node is the average of the values of φ_h at this node.

Let $[\varphi_h]_j$ be the jump of the function φ_h through the edge A_j: $[\varphi_h]_j$ is the difference of the value of φ_h in $T_{i_1(j)}$ and $T_{i_2(j)}$ and if $A_j \subset \Gamma$, then $[\varphi_h]_j = \varphi_h$. We mention the following result which has been proved in [10] and that will be used to prove the efficiency of the estimators.

Lemma 2.6. *Let $\varphi \in P_h^2(\Omega)$, and let $\mathcal{I}_{Os}(\varphi_h) \in P_h^2(\Omega) \cap H_0^1(\Omega)$ be constructed as above. Then*

$$\|\nabla(\varphi_h - \mathcal{I}_{Os}(\varphi_h))\|^2_{L^2(T_i)} \leq C \sum_{A_j \cap T_i \neq \emptyset} |A_j|^{-1} \|[\varphi_h]_j\|^2_{L^2(A_j)}, \tag{2.15}$$

where C depends only on the shape regularity parameter θ defined in subsection 2.2.1.

2.3.4 A posteriori error estimate

We set

$$B(\phi, \varphi) = \sum_{i \in [1,I]} (\nabla\phi, \nabla\varphi)_{T_i} \quad \forall \phi, \varphi \in H_h^1(\Omega) \tag{2.16}$$

and the corresponding energy norm

$$|||\phi|||^2_\Omega = B(\phi, \phi). \tag{2.17}$$

Let $\hat{u}$ be the weak solution of (2.1), we set

$$F(\mathrm{x}) = \int_0^{\hat{u}(\mathrm{x})} H(t)dt. \tag{2.18}$$

It is easy to see that $F \in H_0^1(\Omega)$ and

$$\nabla F(\mathrm{x}) = H(\hat{u}(\mathrm{x}))\nabla\hat{u}(\mathrm{x}). \tag{2.19}$$

For $i \in [1, I]$, let the estimators be defined as

$$\eta_{\mathrm{R},i} := C_{T_i} h_i \|f - f_i\|_{L^2(T_i)},\ \eta_{\mathrm{R}}^2 = \sum_{i\in[1,I]} \eta_{\mathrm{R},i}^2, \tag{2.20}$$

$$\eta_{\mathrm{NC},i} := \|\nabla M - \nabla \mathcal{I}_{Os}(M)\|_{L^2(T_i)},\ \eta_{\mathrm{NC}}^2 = \sum_{i\in[1,I]} \eta_{\mathrm{NC},i}^2, \tag{2.21}$$

$$\eta_{\mathrm{L},i} := \|\nabla M_i - \nabla W_i\|_{L^2(T_i)},\ \eta_{\mathrm{L}}^2 = \sum_{i\in[1,I]} \eta_{\mathrm{L},i}^2. \tag{2.22}$$

Moreover, the (local) discretization estimators is defined following:

$$\eta_{\mathrm{D},i} = \eta_{\mathrm{NC},i} + \eta_{\mathrm{R},i} \text{ and } \eta_D = \eta_{\mathrm{NC}} + \eta_{\mathrm{R}}. \tag{2.23}$$

Now we can state the main result of this section, which is a bound for the L^2 norm of the error between the exact flux $\nabla F = H(\hat{u})\nabla\hat{u}$ and the flux that can be deduced from the solution u_h through the gradient of M defined by (2.11).

Theorem 2.7. *There holds*

$$\|\nabla F - \nabla M\|_{L^2(\Omega)} \leq \eta_{NC} + \eta_R + \eta_L. \tag{2.24}$$

2.3.5 Balancing discretization and linearization estimators

To balance the discretization and linearization estimators, we follow the idea in [6]: we choose a positive parameter γ_{D} and stop the iterative loop in Subsection 2.2.2 whenever

$$\eta_{\mathrm{L}} \leq \gamma_{\mathrm{D}} \eta_{\mathrm{D}}. \tag{2.25}$$

With this condition, unnecessary linearization iterations can be avoided, which can lead to important computational savings.

2.3.6 Proof of the a posteriori error estimate

From the definition of the operator B, there holds (see Lemma 7.1 in [18]):

Theorem 2.8 (Abstract framework). *Let $F, S \in H_0^1(\Omega)$ and let $M \in H_h^1(\Omega)$ be arbitrary. Then*

$$|||F - M|||_\Omega \leq |||M - S|||_\Omega + \left| B(F - M, \frac{F - S}{|||F - S|||_\Omega}) \right|. \tag{2.26}$$

Lemma 2.9. *Let $\Phi \in H_0^1(\Omega)$ be arbitrary such that $\|\nabla\Phi\|_{L^2(\Omega)} = 1$. Then*

$$|B(F - M, \Phi)| \leq \eta_R + \eta_L. \tag{2.27}$$

Proof. Using the bilinearity of $B(.,.)$, combining the property of F in (2.19) with the variational formulation involving the exact solution $\hat{u}$ in (2.3) and the definition of M in (2.11), we obtain

$$\begin{aligned} B\left(F-M,\Phi\right) &= \sum_{i\in[1,I]}\int_{T_i}(\nabla F-\nabla M)\cdot\nabla\Phi(\mathrm{x})d\mathrm{x} \\ &= \int_{\Omega} f\Phi(\mathrm{x})d\mathrm{x} - \sum_{i\in[1,I]}\int_{T_i}\nabla W_i(\mathrm{x})\cdot\nabla\Phi(\mathrm{x})d\mathrm{x} \\ &- \sum_{i\in[1,I]}\int_{T_i}(\nabla M_i-\nabla W_i)\cdot\nabla\Phi(\mathrm{x})d\mathrm{x}. \end{aligned} \tag{2.28}$$

Let Y be defined by

$$Y := \sum_{i\in[1,I]}\int_{T_i}\nabla W_i(x)\cdot\nabla\Phi(\mathrm{x})d\mathrm{x}.$$

Applying the Green formula on each T_i, $i\in[1,I]$, there holds

$$Y = \sum_{i\in[1,I]}\sum_{j\in\partial T_i}\int_{A_j}\nabla W_i\cdot\mathbf{n}_{ji}\Phi(\sigma)d\sigma - \sum_{i\in[1,I]}\int_{T_i}\Delta W_i\Phi(\mathrm{x})d\mathrm{x}.$$

Using the property of W_i in (2.12)-(2.13), for each $i\in[1,I]$ and $j\in\partial T_j$, $F_{ji}(w_h,u_h)$ is constant for all A_j; thus we come to

$$\begin{aligned} Y &= \sum_{i\in[1,I]}\sum_{j\in\partial T_i}F_{ji}(w_h,u_h)\int_{A_j}\Phi(\sigma)d\sigma + \sum_{i\in[1,I]}\int_{T_i}f_i\Phi(\mathrm{x})d\mathrm{x} \\ &= \sum_{j\notin\Gamma}\left[F_{ji_1(j)}(w_h,u_h)+F_{ji_2(j)}(w_h,u_h)\right]\int_{A_j}\Phi(\sigma)d\sigma + \sum_{j\in\Gamma}F_{ji}(w_h,u_h)\int_{A_j}\Phi(\sigma)d\sigma \\ &+ \sum_{i\in[1,I]}\int_{T_i}f_i\Phi(\mathrm{x})d\mathrm{x}. \end{aligned}$$

For any edge $A_j\not\subset\Gamma$, from the second equation in (2.9), the summation of the flux from $T_{i_1(j)}$ to $T_{i_2(j)}$ and from $T_{i_2(j)}$ to $T_{i_1(j)}$ vanishes. In addition, since $\Phi\in H_0^1(\Omega)$, then for any boundary edge A_j, there holds

$$\int_{A_j}\Phi(\sigma)d\sigma = 0.$$

Thus,

$$\sum_{i\in[1,I]}\int_{T_i}\nabla W_i(\mathrm{x})\cdot\nabla\Phi(\mathrm{x})d\mathrm{x} = \sum_{i\in[1,I]}\int_{T_i}f_i\Phi(\mathrm{x})d\mathrm{x}. \tag{2.29}$$

Combining (2.28) and (2.29), there holds

$$B(F-M,\Phi)=\sum_{i\in[1,I]}\int_{T_i}(f-f_i)(\Phi-\Phi_i)(\mathrm{x})d\mathrm{x}-\sum_{i\in[1,I]}\int_{T_i}(\nabla M_i-\nabla W_i)\cdot\nabla\Phi(\mathrm{x})d\mathrm{x},$$

where Φ_i is the mean value of Φ over T_i for all $i\in[1,I]$. Applying the Cauchy-Schwarz inequality, there holds

$$B(F-M,\Phi)\leq\sum_{i\in[1,I]}\|f-f_i\|_{L^2(T_i)}\|\Phi-\Phi_i\|_{L^2(T_i)}+\sum_{i\in[1,I]}\|\nabla M_i-\nabla W_i\|_{L^2(T_i)}\|\nabla\Phi\|_{L^2(T_i)}.$$

Using Lemma 2.5, we obtain that

$$\|\Phi-\Phi_i\|_{L^2(T_i)}\leq C_{T_i}h_i\|\nabla\Phi\|_{L^2(T_i)}. \tag{2.30}$$

Finally, applying the discrete Cauchy-Schwarz inequality and taking into account the fact that $\|\nabla\Phi\|_{L^2(\Omega)}\leq 1$, there holds

$$|B\left(F-M,\Phi\right)|\leq\left(\sum_{i\in[1,I]}\eta_{R,i}^2\right)^{1/2}+\left(\sum_{i\in[1,I]}\eta_{L,i}^2\right)^{1/2}, \tag{2.31}$$

which completes the lemma. Note that in practice, this bound can be improved to

$$|B\left(F-M,\Phi\right)|\leq\left(\sum_{i\in[1,I]}(\eta_{R,i}+\eta_{L,i})^2\right)^{1/2}.$$

□

Now, we come back to the proof of Theorem 2.7. We choose $S=\mathcal{I}_{Os}(M)\in P_h^2(\Omega)\cap H_0^1(\Omega)$ and apply Theorem 2.8 and Lemma 2.9.

2.4 Efficiency of the estimators

Lemma 2.10. *There exists a positive constant C independent of the mesh such that*

$$|A_j|^{-1/2}\|[M]_j\|_{L^2(A_j)}\leq C\sum_{\alpha\in\{1,2\}}\|\nabla M-\nabla F\|_{L^2(T_{i_\alpha(j)})}+|A_j|^{-1/2}\|\langle[M-F]_j,1\rangle_{A_j}|A_j|^{-1}\|_{L^2(A_j)}.$$

Proof. Applying the triangle inequality and the fact that $\langle[F],1\rangle_{A_j}=0$, there holds

$$\|[M]_j\|_{L^2(A_j)}\leq\|[M]_j-\overline{[M]}_j\|_{L^2(A_j)}+\|\langle[M-F]_j,1\rangle_{A_j}|A_j|^{-1}\|_{L^2(A_j)}, \tag{2.32}$$

where $\overline{[M]}_j=\frac{1}{|A_j|}\int_{A_j}[M]_j(\sigma)d\sigma$.

We will next use the inequality

$$\|[\Psi_h]_j\|_{L^2(A_j)} \leq C|A_j|^{1/2} \sum_{\alpha\in\{1,2\}} \|\nabla\Psi_h - \nabla\Phi\|_{L^2(T_{i_\alpha}(j))} \tag{2.33}$$

where C only depends on the parameter θ and $j \in [1, J]$, $\Psi_h \in H_h^1(\Omega)$ such that $\langle[\Psi_h], 1\rangle_{A_j} = 0$ and $\Phi \in H_0^1(\Omega)$ are arbitrary, for more detail, see Theorem 10 in [2]. Applying this inequality to the first term in the right-hand of (2.32) with $\Psi_h|_{T_{i_\alpha(j)}} = M|_{T_{i_\alpha(j)}} - \overline{M}_{A_j,T_{i_\alpha(j)}}$ and $F = \Psi$, where $\overline{M}_{A_j,T_{i_\alpha(j)}} = \frac{1}{|A_j|}\int_{A_j} M|_{T_{i_\alpha(j)}}(\sigma)d\sigma$, we obtain

$$\|[M]_j\|_{L^2(A_j)} \leq C|A_j|^{1/2} \sum_{\alpha\in\{1,2\}} \|\nabla M - \nabla F\|_{L^2(T_{i_\alpha}(j))} + \|\langle[M-F]_j, 1\rangle_{A_j}|A_j|^{-1}\|_{L^2(A_j)}, \tag{2.34}$$

which concludes the proof. □

Using the Lemma 2.6 and Lemma 2.10, we obtain the following theorem.

Theorem 2.11. *There exists a positive constant C depending only on the parameter θ such that*

$$\eta_{NC,i}^2 \leq C \sum_{T_k\cap T_i\neq\emptyset} \|\nabla M - \nabla F\|_{L^2(T_k)}^2 + C \sum_{A_j\cap T_i\neq\emptyset} \|\langle[M-F]_j, 1\rangle_{A_j}|A_j|^{-1}\|_{L^2(A_j)}. \tag{2.35}$$

Remark 2.12. *In this section, we consider the efficiency of η_{NC} only, because η_R will be of higher order as soon as f is more regular that $L^2(\Omega)$ and because η_L is controlled by η_{NC} and η_R through (2.25).*

2.5 Numerical results

We show here some numerical results obtained on a domain $\Omega =]-1;1[\times]-1;1[\backslash[0;1]\times[-1;0]$. The exact solution is $\hat{u}(r,\theta) = r^{2/3}\sin(2\theta/3)$, expressed in cylindrical coordinates (r,θ) centered on $(0,0)$. We use the Triangle mesh generator and the DDFV scheme for all tests. Our tests have two parts. In the first, we compare the estimated and actual errors, the effectivity indices for a uniform and an adaptive refinement. On the mesh given, we compute the discretization estimator $\eta_{\mathrm{D},i} = \eta_{\mathrm{NC},i} + \eta_{\mathrm{R},i}$ and ask to refine a given T_i by a factor 4 in terms of area if $\eta_{\mathrm{D},i} \geq (\max_i \eta_{\mathrm{D},i})/2$. To stop the iteration in the fixed point method in Subsection 2.2.2, we choose the parameter value $\gamma_{\mathrm{D}} = 0.01$. In the second test, we will discuss our stopping criterion with the parameter value $\gamma_{\mathrm{D}} = 0.01$ for the fixed point iteration by comparing it to a classical stopping criterion, namely that the balance equation are satisfied by u_h^{n+1} up to a certain relative tolerance in the discrete L^2 norm:

$$\frac{\left(\sum_{i\in[1,I]} |T_i| \left[\sum_{j\in\partial T_i} \frac{|A_j|}{T_i} F_{i,j}(u_h^{n+1}, u_h^{n+1}) - f_i^T\right]^2\right)^{1/2}}{\left(\sum_{i\in[1,I]} |T_i||f_i^T|^2\right)^{1/2}} \leq 10^{-8}. \tag{2.36}$$

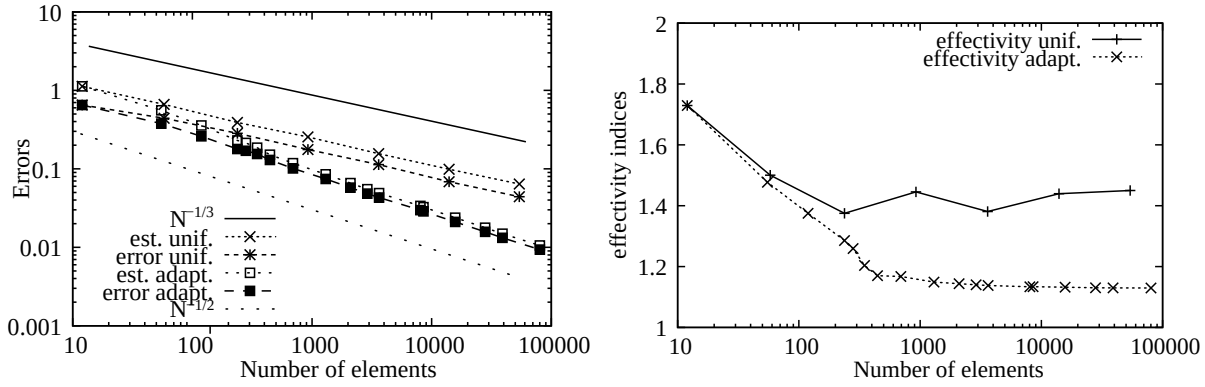

Figure 2.4: Estimated and actual errors for uniform and adaptive refinements (left) and effectivity indices for uniform and adaptive refinements (right) for $H(x) = 1 + 1/(1 + x^2)$.

In figure 2.4(left), we have plotted the curves of the actual errors and the estimators for a uniform and for an adaptive mesh refinement. The curve corresponding to the uniform mesh refinement is parallel to the $N^{-1/3}$ curve, while the curve corresponding to the adaptive mesh refinement is parallel to the $N^{-1/2}$ curve. In Figure 2.4(right), we plot the effectivity indices for the uniform and adaptive refinement. The efficiency for the adaptive refinement is around 1.15 while the efficiency for uniform refinement is around 1.4.

Next, Figure 2.5 presents the true error, the total estimator η, the discretization estimator η_{D}, and the linearization estimator η_{L} on a given mesh as a function of the number of fixed-point iterations. Firstly, we clearly see that the discretization estimator decreases along with the linearization one only during the first few iterations, then while the linearization estimator continues to decrease, the total error and the discretization estimator remain almost constant, and we note that they are almost equal.

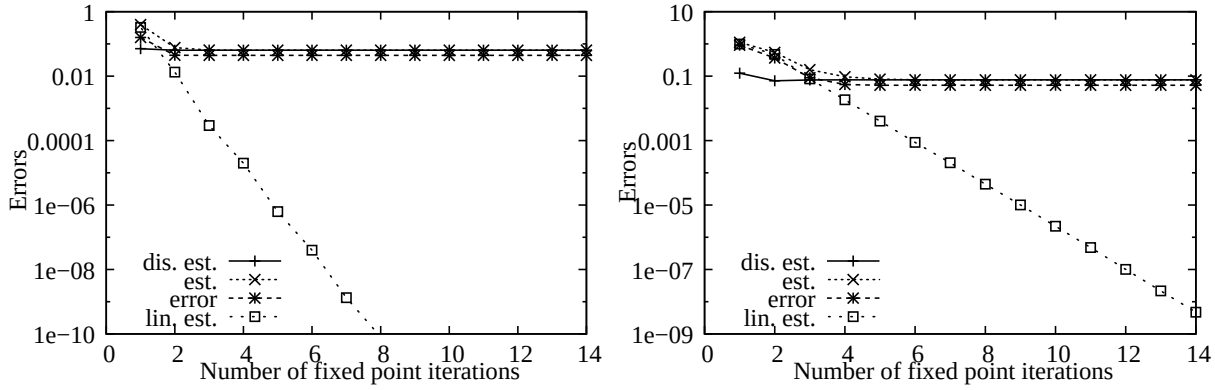

Figure 2.5: Total error, total estimator, discretization and linearization estimators as a function of the fixed point iteration. Left: $H(x) = 1 + 1/(1 + x^2)$ and right: $H(x) = 2 + \sin(10x)$.

Moreover, the number of iterations necessary to reach the stopping criterion increases as the function H becomes more oscillatory. For instance, in the left part of Figure 2.5, the

stopping criterion (2.25) is reached after 3 iterations and the classical one after 8 iterations; in the right part of this figure, these numbers become respectively 7 and 14. This confirms that the fixed point iteration can be stopped earlier than classical iteration.

References for the Chapter 2

[1] I. Aavatsmark, *An introduction to multipoint flux approximation for quadrilateral grids.* Computational Geosciences 6, pp. 405–432 (2002).

[2] Y. Achdou, C. Bernardi, F. Coquel, *A priori and a posteriori analysis of finite volume discretization of Darcy's equations.* Numer. Math. 96, pp. 17–42 (2003).

[3] H. Brézis, *Équations et inéquations non linéaires dans les espaces vectoriels en dualité.* Ann. Inst. Fourier, tome 18, No.1, pp. 115–175 (1968).

[4] M. Chipot, G. Michaille, *Uniqueness results and monotonocity properties for strongly nonlinear elliptic variational inequalities.* Ann. Scuola Norm. Sup. Pisa Cl. Sci (4), 16(1), pp. 137–166 (1989).

[5] K. Domelevo, P. Omnes, *A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids.* M2AN Math. Model. Numer. Anal. 39 (6), pp. 1203–1249 (2005).

[6] L. El Alaoui, A. Ern, M. Vohralík, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems.* Comput. Methods Appl. Mech. Engrg., to appear (2011).

[7] R. Eymard, T. Gallouët, R. Herbin, *Finite Volume methods.* In: Ciarlet, Ph.G. and Lions, J.-L., Eds. Handbook of Numerical Analysis, Vol. VII, Amsterdam: North-holland, pp. 173–1020 (2000).

[8] R. Eymard, T. Gallouët, and R. Herbin, *Finite volume approximation of elliptic problems and convergence of an approximate gradient.* Appl. Numer. Math. 37(1-2), pp. 31–53 (2001).

[9] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires.* Dunod, Gautthier-Villars, Paris (1969).

[10] O.A Karakashian, F. Pascal, *A posteriori error estimate for a discontinuous Galerkin approximation of second-order elliptic problems.* SIAM J. Numer. Anal. 41(6), pp. 2374–2399 (2003).

[11] D. Kim, E.J. Park, *A priori and a posteriori analysis of mixed finite element methods for nonlinear elliptic equations.* SIAM J. Numer. Anal. 48 (3), pp. 1186–1207 (2010).

[12] F. A. Milner, *Mixed finite element methods for quasilinear second-order elliptic problems.* Math. Comp. 44, pp. 303–320 (1985).

[13] S. Nicaise, *A posteriori error estimations of some cell-centered finite volume methods.* SIAM J. Numer. Anal. 43, pp. 1481–1503 (2005).

[14] P. Omnes, Y. Penel, Y. Rosenbaum, *A Posteriori Error Estimation for the Discrete Duality Finite Volume Discretization of the Laplace Equation.* SIAM J. Numer. Anal. 47(4), pp. 2782–2807 (2009).

[15] E. J. Park, *Mixed finite element methods for nonlinear second order elliptic problems.* SIAM J. Numer. Anal. 32, pp. 865–885 (1995).

[16] L.E Payne, H.F Weiberger, *An optimal Poincaré inequality for convex domain.* Arch. Rational Mech. Anal. 5, pp. 286–292 (1960).

[17] J.E Roberts, J.M Thomas, *Mixed and hybrids methods.* in: Handbook of Numerical Analysis II (North-Holland, Amsterdam), pp. 523–640 (1991).

[18] M. Vohralík, *A posteriori error estimates for lowest-order mixed finite element discretization of convection-diffusion-reaction equations.* SIAM J. Numer. Anal. 45(4), pp. 1570–1599 (2007).

[19] M. Vohralík, *Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods.* Numer. Math. 111, pp. 121–158 (2008).

Chapter 3

Transport Equations

In this chapter, we will consider a posteriori error estimation for the transport equation $\partial_t u + \mathbf{a}(x,t)\cdot\nabla u = 0$ *with the initial data* $u_0 \in L^\infty \cap BV_{\rm loc}$ *and the divergence of the velocity field* $\mathbf{a}$ *is not equal to zero. An a posteriori estimate for the error between the exact solution and the solution of an upwind finite volume scheme is derived in the* L^1 *norm.*

3.1 Introduction

We consider here the following transport equation in N space dimensions ($N \geq 1$), with initial condition:

$$\begin{cases} \frac{\partial u}{\partial t} + \mathbf{a}(\mathrm{x},t)\cdot\nabla u &= 0,\ \forall \mathrm{x}\in\mathbb{R}^N,\ t\in\mathbb{R}^+, \\ u(\mathrm{x},0) &= u_0(\mathrm{x}),\ \forall \mathrm{x}\in\mathbb{R}^N. \end{cases} \tag{3.1}$$

The following hypotheses are made on the data:

$$\begin{aligned} &(i)\ u_0 \in L^\infty(\mathbb{R}^N)\cap BV_{\rm loc}(\mathbb{R}^N);\ \text{we call } (U_m,U_M)\in\mathbb{R}^2 \text{ such that } U_m \leq u_0 \leq U_M \ a.e. \\ &(ii)\ \mathbf{a}\in C^1(\mathbb{R}^N\times\mathbb{R}^+,\mathbb{R}^N),\ \text{there exist } V, L_{\mathbf{a}} < \infty \text{ such that} \\ &\qquad |\mathbf{a}(\mathrm{x},t)| \leq V,\ \forall(\mathrm{x},t)\in\mathbb{R}^N\times\mathbb{R}^+, \\ &\qquad |\mathbf{a}(\mathrm{x},t)-\mathbf{a}(\mathrm{y},s)| \leq L_{\mathbf{a}}(|\mathrm{x}-\mathrm{y}|+|t-s|)\ \forall(\mathrm{x},t),(\mathrm{y},s)\in\mathbb{R}^N\times\mathbb{R}^+ \\ &\quad \text{and}\quad \nabla\cdot\mathbf{a}\in L^\infty(\mathbb{R}^N\times\mathbb{R}^+). \end{aligned} \tag{3.2}$$

We shall now define the functional spaces $BV(\Omega)$ and $BV_{\rm loc}(\Omega)$.

Definition 3.1. *Let* $\Omega\subset\mathbb{R}^p$*, with* $p\in\mathbb{N}$*; the functional space* $BV(\Omega)$ *is defined as follows:*

$$BV(\Omega) = \left\{ g :\ \sup\left\{ \int_\Omega g(\mathrm{x})\nabla\cdot\varphi(\mathrm{x})d\mathrm{x},\ \varphi\in C_c^\infty(\Omega,\mathbb{R}^p),\ \|\varphi\|_\infty\leq 1\right\} < \infty\right\}.$$

On $BV(\Omega)$, we define a seminorm:

$$|g|_{BV(\Omega)} = \sup\left\{ \int_\Omega g(\mathrm{x})\nabla\cdot\varphi(\mathrm{x}),\ \varphi\in C_c^\infty(\Omega,\mathbb{R}^p),\ \|\varphi\|_\infty\leq 1\right\}.$$

We also consider $BV_{\text{loc}}(\Omega)$:

$$BV_{\text{loc}}(\Omega) = \{g;\ g \in BV(K) \text{ for all compact } K \subset \Omega\}.$$

Following [5], we say that $u \in L^\infty(\mathbb{R}^N \times \mathbb{R}^+)$ is the unique entropy solution to (3.1), if for all $\kappa \in \mathbb{R}$, $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R}^+, \mathbb{R}^+)$, we have

$$\begin{aligned}\int_{\mathbb{R}^N\times\mathbb{R}^+} \Big[&|u(\mathrm{x},t)-\kappa|\varphi_t(\mathrm{x},t) + (u(\mathrm{x},t)\top\kappa - u(\mathrm{x},t)\bot\kappa)\mathbf{a}(\mathrm{x},t)\cdot\nabla\varphi(\mathrm{x},t) \\ &+(u(\mathrm{x},t)\top\kappa - u(\mathrm{x},t)\bot\kappa)\nabla\cdot\mathbf{a}(\mathrm{x},t)\,\varphi(\mathrm{x},t)\Big] d\mathrm{x}dt + \int_{R^N} |u_0(\mathrm{x})-\kappa|\varphi(\mathrm{x},0)d\mathrm{x} \geq 0,\end{aligned} \tag{3.3}$$

where for $a, b \in \mathbb{R}$, we set $a\top b := \max\{a, b\}$ and $a\bot b := \min\{a, b\}$.

The upwind scheme is a standard method to solve in a approximate way this problem, and, more generally, conservation laws (see [3]), but the a priori and a posteriori error estimations have been investigated only recently on scalar equations.

In the case of a conservation law with a divergence free flux function and no source term, C. Chainais-Hillairet [1] derived an a priori error estimate with a convergence order of $1/4$ in the L^1 norm. Based on [1], an a posteriori error estimation is achieved by D. Kröner and M. Ohlberger [4] for the upwind explicit scheme. This result is used to define an algorithm with an adaptive grid for the finite volume scheme. Recently, in Mamaghani's Ph.D. thesis [6], an a posteriori error estimate is obtained for the implicit upwind finite volume scheme.

Otherwise, in [2], the a priori error estimation was also treated for a conservation law with a non divergence free flux function and with a source term (both "stiff" and "non stiff" source terms are treated). However, the authors did not obtain explicit bounds that could be used in a posteriori error estimates, and they used the property that with the initial data belonging to $BV_{\text{loc}}(\mathbb{R}^N)$, then the entropy solution also belongs to $BV_{\text{loc}}(\mathbb{R}^N \times \mathbb{R}^+)$ but did not actually prove it.

In our work, we will deal with the transport equation with a non divergence free velocity field, written as a conservation law with the source term $\nabla\cdot\mathbf{a}(\mathrm{x},t)u(\mathrm{x},t)$, and we obtain an a posteriori error estimation for the explicit upwind scheme. To obtain this result, we shall prove the property that with the initial data belonging to $BV_{\text{loc}}(\mathbb{R}^N)$, then the entropy solution also belongs to $BV_{\text{loc}}(\mathbb{R}^N \times \mathbb{R}^+)$.

3.2 Definition and stability of the upwind scheme

3.2.1 Notations and definition of the scheme

Firstly, let $\{t_n : n \in \mathbb{N} \text{ s.t. } 0 = t_0 < t_1 < t_2 < \cdots\}$ be a partition of $\mathbb{R}^+$ and $\Delta t_n = t_{n+1} - t_n$ be the steps of this partition. Further, for all $n \in \mathbb{N}$, let $\mathcal{T}$ be a mesh of $\mathbb{R}^N$ such that the common interface of any two elements (which are called control volumes in what follows) of $\mathcal{T}$ is included in a hyperplane of $\mathbb{R}^N$ and the elements are convex subdomains. Let $\mathcal{N}(p)$ denote the set of neighbours of the control volume p; for $q \in \mathcal{N}(p)$, we denote by $\sigma_{p,q}$ the common interface between p and q, and by $\mathbf{n}_{p,q}$ the unit normal vector to $\sigma_{p,q}$ oriented from p to q. For any $(p, q) \in \mathcal{T}^2$, let h_p be the diameter of control volume p and $h = \max\{h_p : p \in \mathcal{T}\}$

and let $h_{p,p}$ be the diameter of common interface $\sigma_{p,q}$. We assume that there exist $\alpha > 0$ such that, for all $p \in \mathcal{T}$:

$$\begin{aligned} \alpha h^N &\leq |p|, \\ \alpha |\partial p| &\leq h^{N-1}. \end{aligned} \tag{3.4}$$

Now, we define the upwind scheme. The discrete unknowns are denoted by u_p^n, $n \in \mathbb{N}$, $p \in \mathcal{T}$. The set $\{u_p^0, p \in \mathcal{T}\}$ is given by the initial condition:

$$u_p^0 = \frac{1}{|p|} \int_p u_0(\mathrm{x}) d\mathrm{x}. \tag{3.5}$$

Assume the CFL condition

$$\Delta t_n \leq (1-\xi) \frac{\alpha^2 h}{V} \ , \quad \forall\ n \in \mathbb{N}, \tag{3.6}$$

where $\xi \in (0,1)$. Let us consider the following explicit numerical scheme:

$$\frac{u_p^{n+1} - u_p^n}{\Delta t_n} + \frac{1}{|p|} \sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_{p,q}^n - \frac{1}{|p|} \sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_p^n = 0, \tag{3.7}$$

where

$$a_{p,q}^n = \frac{1}{\Delta t_n} \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} \mathbf{a}(\gamma, t) \cdot \mathbf{n}_{p,q} d\gamma dt \tag{3.8}$$

and

$$u_{p,q}^n = \begin{cases} u_p^n & \text{if } a_{p,q}^n > 0, \\ \dfrac{u_p^n + u_q^n}{2} & \text{if } a_{p,q}^n = 0, \\ u_q^n & \text{if } a_{p,q}^n < 0. \end{cases} \tag{3.9}$$

We note that the sum $\dfrac{1}{|p|} \displaystyle\sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_{p,q}^n$ is an upwind discretization of $\nabla \cdot (\mathbf{a}u)$ and the sum $-\dfrac{1}{|p|} \displaystyle\sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_p^n$ is a discretization of $-u \nabla \cdot \mathbf{a}$.

The approximate solution, denoted by u_h, is defined from $\mathbb{R}^N \times \mathbb{R}_+$ to $\mathbb{R}$ with the help of the discrete unknowns of the scheme by:

$$u_h(x,t) = u_p^n \text{ if } x \in p,\ t \in [t_n, t_{n+1}[,\ \forall\ p \in \mathcal{T},\ n \in \mathbb{N}, \text{ and } u_{\mathcal{T},0}(\mathrm{x}) = u_h(\mathrm{x}, 0) \tag{3.10}$$

3.2.2 L^∞-stability of the scheme

Lemma 3.2. *Under assumption (3.2), (3.4) and CFL condition (3.6), let u_h be given by (3.10), then*

$$U_m \le u_p^n \le U_M, \quad \forall n \in \mathbb{N},\ p \in \mathcal{T} \tag{3.11}$$

and

$$\|u_h\|_{L^\infty(\mathbb{R}^N \times \mathbb{R}^+)} \le \|u_0\|_{L^\infty(\mathbb{R}^N)}. \tag{3.12}$$

Proof. From the relation (3.7), we express u_p^{n+1} as a function of u_p^n and u_q^n, $q \in \mathcal{N}(p)$,

$$u_p^{n+1} = u_p^n - \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_{p,q}^n + \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} a_{p,q}^n u_p^n.$$

By the definition of $u_{p,q}^n$, there holds

$$u_p^{n+1} = u_p^n - \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} [a_{p,q}^n]^-(u_q^n - u_p^n),$$

where $[a_{p,q}^n]^- = \min(0, a_{p,q}^n)$. Or

$$u_p^{n+1} = u_p^n(1 + \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} [a_{p,q}^n]^-) + \sum_{q \in \mathcal{N}(p)} \frac{-\Delta t_n}{|p|} [a_{p,q}^n]^- u_q^n. \tag{3.13}$$

Following the CFL condition, then

$$1 + \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} [a_{p,q}^n]^- \ge 0, \ \frac{-\Delta t_n}{|p|} [a_{p,q}^n]^- \ge 0$$

$$\text{and } 1 + \frac{\Delta t_n}{|p|} \sum_{q \in \mathcal{N}(p)} [a_{p,q}^n]^- + \sum_{q \in \mathcal{N}(p)} \frac{-\Delta t_n}{|p|} [a_{p,q}^n]^- = 1.$$

Hence, under the CFL condition, u_p^{n+1} is a convex combination of $\{u_q^n : q \in \mathcal{T}\}$ and we obtain

$$\inf_{q \in \mathcal{T}} u_q^n \le u_p^{n+1} \le \sup_{q \in \mathcal{T}} u_q^n\ ,\ \forall p \in \mathcal{T}.$$

This concludes the proof of (3.11), which, in turn, yields (3.12). □

3.2.3 A "weak BV" estimate

We set some notations that will be used in all the sequel. Let $T > 0$ and $R > 0$ be given, and let $B(0, R)$ be the ball of radius R centered on the origin, we define

$$\begin{aligned}
N_T &= \max\{n \in \mathbb{N}, t_n \le T\},\\
\mathcal{T}_R &= \{p \in \mathcal{T}, p \subset B(0,R)\},\\
\mathcal{E}_*^n &= \{(p,q) \in (\mathcal{T})^2,\ q \in \mathcal{N}(p),\ u_p^n > u_q^n\},\\
\mathcal{E}_R^n &= \{(p,q) \in (\mathcal{T})^2, p \text{ or } q \in B(0,R),\ q \in \mathcal{N}(p),\ \sigma_{p,q} \subset B(0,R) \text{ and } u_p^n > u_q^n\},\\
\mathcal{E}_{R,h} &= \{(p,q) \in (\mathcal{T})^2, p \subset B(0,R) \text{ and } q \not\subset B(0,R), q \in \mathcal{N}(p)\}\\
\mathcal{E} &= \big\{(p,q) \in (\mathcal{T})^2,\ q \in \mathcal{N}(p)\big\}.
\end{aligned}$$

Lemma 3.3. *Under assumption (3.2), (3.4) and CFL condition (3.6), then there exists a constant C_{bv} such that*

$$\sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n} |a_{p,q}^n||u_p^n - u_q^n| \leq \frac{C_{bv}}{\sqrt{h}},\ \forall h \in R \tag{3.14}$$

$$\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} |p||u_p^{n+1} - u_n^p| \leq \frac{C_{bv}}{\sqrt{h}},\ \forall h \in R. \tag{3.15}$$

Proof. In this proof, we shall denote by C_i ($i \in \mathbb{N}$) various quantities only depending on $\mathbf{a}$, u_0, ξ, R, T. The mesh size is chosen small enough ($h < R$) so that $\mathcal{T}_R$ is not empty. We first prove (3.14). Multiplying 3.7 by $\delta t_n |p| u_p^n$ and summing the result over $p \in T_R$, $n \in 0, \cdots, N_T$ yields

$$B_1 + B_2 = 0, \tag{3.16}$$

with

$$B_1 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} |p| u_p^n (u_p^{n+1} - u_n^n)$$

and

$$B_2 = \sum_{n=0}^{N_T} \Delta t_n \sum_{p\in\mathcal{T}_R} \sum_{q\in\mathcal{N}(p)} a_{p,q}^n (u_{p,q} - u_p) u_p^n.$$

Gathering the last two summations by edges in B_2, since $a_{p,q}^n = -a_{q,p}^n$ and $u_{p,q}^n = u_{q,p}^n$, we have

$$B_2 = B_3 - \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in E_{R,h}} a_{q,p}^n (u_{p,q} - u_q) u_q^n, \tag{3.17}$$

where

$$B_3 = \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n} \left[a_{p,q}^n (u_{p,q} - u_p) u_p^n - a_{p,q}^n (u_{p,q} - u_q) u_q^n \right].$$

Then

$$|B_2 - B_3| \leq \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_{R,h}} |a_{q,p}^n (u_{p,q} - u_q) u_q^n|$$

We consider that elements p in $\mathcal{E}_{R,h}$ are included in $B(0,R) \setminus B(0,R-h)$, since the measure of $B(0,R) \setminus B(0,R-h)$ is less than $C_2 h$, then number of the elements in $\mathcal{E}_{R,h}$, for fixed n, is lower than $C_2 h/(\alpha h^N) = C_3 h^{1-N}$. Thanks to (3.4), using the fact that $|\partial p| \leq (1/\alpha) h^{N-1}$, that $\mathbf{a}(\mathrm{x}, t) \leq V$ and $(u_{p,q} - u_q) u_q^n \leq 2 \max\{U_m^2, U_M^2\}$. Then

$$\sum_{(p,q)\in\mathcal{E}_{R,h}} |a_{q,p}^n (u_{p,q} - u_q) u_q^n| \leq 2V \max\{U_m^2, U_M^2\} \frac{1}{\alpha} C_3 h^{1-N} h^{N-1} = C_4.$$

Moreover, since $\sum_{n=0}^{N_T} \Delta t_n = t_{N_T+1} \leq 2T$, we deduce that

$$|B_2 - B_3| \leq 2TC_4 = C_5.$$

We now express the term B_3, we have

$$\begin{aligned} a_{p,q}^n(u_{p,q} - u_p)u_p^n - a_{p,q}^n(u_{p,q} - u_q)u_q^n =& a_{p,q}^n\big[(u_p^n - u_q^n)(u_{p,q}^n - \frac{u_p^n + u_q^n}{2})\big] \\ & - \frac{a_{p,q}^n}{2}\big[(u_p^n)^2 - (u_q^n)^2\big], \end{aligned}$$

then

$$B_3 = B_4 + B_5, \tag{3.18}$$

where

$$B_4 = -\sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n} \frac{a_{p,q}^n}{2}\big[(u_p^n)^2 - (u_q^n)^2\big]$$

and

$$B_5 = \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n} a_{p,q}^n\big[(u_p^n - u_q^n)(u_{p,q}^n - \frac{u_p^n + u_q^n}{2})\big].$$

We have

$$B_4 = -\frac{1}{2}\sum_{n=0}^{N_T} \Delta t_n \sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_p^n)^2 + \frac{1}{2}\sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_{R,h}} a_{q,p}(u_q^n)^2. \tag{3.19}$$

The second term of the expression of B_2 can be bounded in same way when we estimated the term $|B_2 - B3|$, then

$$|\sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_{R,h}} a_{q,p}(u_q^n)^2| \leq C_6. \tag{3.20}$$

Therefore,

$$\sum_{n=0}^{N_T} \Delta t_n \sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_p^n)^2 = \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} (u_p^n)^2 \int_{t_n}^{t_{n+1}} \int_p \nabla\cdot\mathbf{a}(\mathrm{x},t)d\mathrm{x}dt$$

By the property of $\nabla\cdot\mathbf{a}$ in (3.2), there holds

$$|\sum_{n=0}^{N_T} \Delta t_n \sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_p^n)^2| \leq 2T|B(0,R)|\max\{U_M^2, U_m^2\}\|\nabla\cdot\mathbf{a}\|_{L^\infty(\mathbb{R}^N\times\mathbb{R}^+)} = C_7. \tag{3.21}$$

Combining (3.19)-(3.21), we obtain

$$|B_4| \leq \frac{C_6 + C_7}{2} = C_8. \tag{3.22}$$

Using the definition of $u^n_{p,q}$, we have

$$a^n_{p,q}\big[(u^n_p - u^n_q)(u^n_{p,q} - \frac{u^n_p + u^n_q}{2})\big] \geq \frac{|a^n_{p,q}|}{2}(u^n_p - u^n_q)^2.$$

Then, there holds

$$B_5 \geq \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}^n_R} \frac{|a^n_{p,q}|}{2}(u^n_p - u^n_q)^2.$$

Using (3.2.3), (3.18), (3.22) and (3.2.3), there holds

$$B_2 \geq \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}^n_R} \frac{|a^n_{p,q}|}{2}(u^n_p - u^n_q)^2 - C_9, \tag{3.23}$$

where $C_9 = C_5 + C_8$.
Let us now turn to B_1. We have

$$\begin{aligned} B_1 = &\frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |p|u^n_p(u^{n+1}_p - u^n_p) - \frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |p|u^{n+1}_p(u^{n+1}_p - u^n_p) \\ &+ \frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |p|u^n_p(u^{n+1}_p - u^n_p) + \frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |p|u^{n+1}_p(u^{n+1}_p - u^n_p) \end{aligned}$$

Combining the first term and the second term, the third term and fourth term in the previous expression, we obtain

$$B_1 = -\frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |p|(u^{n+1}_p - u^n_p)^2 + \frac{1}{2}\sum_{p\in\mathcal{T}_R} |p|(u^{N_T+1}_p)^2 - \frac{1}{2}\sum_{p\in\mathcal{T}_R} |p|(u^0_p)^2. \tag{3.24}$$

Using (3.7) and the Cauchy-Schwarz inequality yields the following inquality:

$$(u^{n+1}_p - u^n_p)^2 \leq \frac{(\Delta t_n)^2}{|p|^2}\sum_{q\in\mathcal{N}(p)} |a^n_{p,q}| \sum_{q\in\mathcal{N}(p)} |[a^n_{p,q}]^-|(u^n_p - u^n_q)^2. \tag{3.25}$$

By the definition of $a^n_{p,q}$, we have

$$\sum_{q\in\mathcal{N}(p)} |a^n_{p,q}| \leq V|\partial p|. \tag{3.26}$$

Using the CFL condition (3.6) and the inequality (3.26) gives

$$(u^{n+1}_p - u^n_p)^2 \leq (1-\xi)\Delta t_n \sum_{q\in\mathcal{N}(p)} |[a^n_{p,q}]^-|(u^n_p - u^n_q)^2. \tag{3.27}$$

Summing inequality (3.27) over $p \in \mathcal{T}_R$ and over $n = 0, \cdots, N_T$ and reordering the summation leads to

$$\begin{aligned} -\frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}|p|(u_p^{n+1}-u_p^n)^2 \le -\frac{1-\xi}{2}\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n|(u_p^n-u_q^n)^2 \\ +\frac{1-\xi}{2}\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_{R,h}}|[a_{q,p}^n]^-|(u_p^n-u_q^n)^2 \end{aligned} \tag{3.28}$$

Using (3.24) and (3.28), we have

$$B_1 \ge -\frac{1-\xi}{2}\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n|(u_p^n-u_q^n)^2 - C_{10}, \tag{3.29}$$

where $C_{10} \ge \frac{1}{2}\sum_{p\in\mathcal{T}_R}|p|(u_p^0)^2$.
Adding (3.23) and (3.29) side by side, and using the equality (3.16), there holds:

$$\frac{\xi}{2}\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n|(u_p-u_q)^2 \le C_{11}, \tag{3.30}$$

where $C_{11} = C_9 + C_{10}$.
Applying the Cauchy-Schwarz inequality to the left-hand side of (3.14), and using (3.30) yields

$$\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n||u_p^n-u_q^n| \le C_{12}\left(\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n|\right)^{1/2}, \tag{3.31}$$

where $C_{12} = \sqrt{\frac{2C_{11}}{\xi}}$.
Noting that

$$\sum_{(p,q)\in\mathcal{E}_R^n}|a_{p,q}^n| \le \sum_{p\in\mathcal{T}_R}V|\partial p| \le V\frac{h^{N-1}}{\alpha}\frac{|B(0,R)|}{\alpha h^N} = \frac{C_{13}}{h}$$

and $\sum_{n=0}^{N_T}\Delta t_n = t_{N_T+1} \le 2T$, one obtians (3.14) from (3.31) with $C_{bv} = C_{12}\sqrt{C_{13}}$.

The estimate on the derivative time (3.15) is a straightforword consequence of (3.7) and (3.15). □

3.3 Entropy inequalities for the approximate solution

3.3.1 Discrete entropy inequality

Lemma 3.4. *Under assumption (3.2), (3.4) and CFL condition (3.6), for all $\kappa \in \mathbb{R}$, $p \in \mathcal{T}$, $n \in \mathbb{N}$, the following inequality holds:*

$$\begin{aligned}\frac{|u_p^{n+1}-\kappa|-|u_p^n-\kappa|}{\Delta t_n}+\frac{1}{|p|}\sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_{p,q}^n\top\kappa-u_{p,q}^n\bot\kappa)\\ -\frac{1}{|p|}\sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_p^n\top\kappa-u_p^n\bot\kappa)\le 0.\end{aligned} \tag{3.32}$$

Proof. We recall the expression of u_p^{n+1} in (3.13),

$$u_p^{n+1}=u_p^n(1+\frac{\Delta t_n}{|p|}\sum_{q\in\mathcal{N}(p)}[a_{p,q}^n]^-)+\sum_{q\in\mathcal{N}(p)}\frac{-\Delta t_n}{|p|}[a_{p,q}^n]^-u_q^n. \tag{3.33}$$

The coefficients of (u_p^n, u_q^n) are positive and their sum is equal to one. Hence

$$u_p^{n+1}\top\kappa\le u_p^n\top\kappa(1+\frac{\Delta t_n}{|p|}\sum_{q\in\mathcal{N}(p)}[a_{p,q}^n]^-)+\sum_{q\in\mathcal{N}(p)}\frac{-\Delta t_n}{|p|}[a_{p,q}^n]^-u_q^n\top\kappa \tag{3.34}$$

and

$$u_p^{n+1}\bot\kappa\ge u_p^n\bot\kappa(1+\frac{\Delta t_n}{|p|}\sum_{q\in\mathcal{N}(p)}[a_{p,q}^n]^-)+\sum_{q\in\mathcal{N}(p)}\frac{-\Delta t_n}{|p|}[a_{p,q}^n]^-u_q^n\bot\kappa. \tag{3.35}$$

Subtracting (3.35) from (3.34) side by side, one has

$$\begin{aligned}u_p^{n+1}\top\kappa-u_p^{n+1}\bot\kappa\le(u_p^n\top\kappa-u_p^n\bot\kappa)(1+\frac{\Delta t_n}{|p|}\sum_{q\in\mathcal{N}(p)}[a_{p,q}^n]^-)\\ +\sum_{q\in\mathcal{N}(p)}\frac{-\Delta t_n}{|p|}[a_{p,q}^n]^-\ (u_q^n\top\kappa-u_q^n\bot\kappa),\end{aligned} \tag{3.36}$$

or

$$\begin{aligned}(u_p^{n+1}\top\kappa-u_p^{n+1}\bot\kappa)-(u_p^n\top\kappa-u_p^n\bot\kappa)+\sum_{q\in\mathcal{N}(p)}\frac{\Delta t_n}{|p|}[a_{p,q}^n]^-(u_q^n\top\kappa-u_q^n\bot\kappa)\\ -\frac{\Delta t_n}{|p|}\sum_{q\in\mathcal{N}(p)}[a_{p,q}^n]^-\ (u_p^n\top\kappa-u_p^n\bot\kappa)\le 0.\end{aligned}$$

Noting that $u_p^{n+1}\top\kappa-u_p^{n+1}\bot\kappa=|u_p^{n+1}-\kappa|$ and $u_p^n\top\kappa-u_p^n\bot\kappa=|u_p^n-\kappa|$, we obtain

$$\begin{aligned}\frac{|u_p^{n+1}-\kappa|-|u_p^n-\kappa|}{\Delta t_n}+\frac{1}{|p|}\sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_{p,q}^n\top\kappa-u_{p,q}^n\bot\kappa)\\ -\frac{1}{|p|}\sum_{q\in\mathcal{N}(p)} a_{p,q}^n(u_p^n\top\kappa-u_p^n\bot\kappa)\le 0.\end{aligned}$$

We have the statement of the lemma. □

3.3.2 Continuous entropy estimates for the approximate solution

For $\Omega = \mathbb{R}^N$ or $\mathbb{R}^N \times \mathbb{R}^+$, we denote by $\mathcal{M}(\Omega)$ the set of positive measures on Ω, that is of σ-additive mappings from the Borel σ-algebra of Ω in $\overline{\mathbb{R}}^+$. If $\mu \in \mathcal{M}(\Omega)$ and $g \in C_c(\Omega)$, one sets $\langle \mu, g \rangle = \int g d\mu$.

Theorem 3.5. *Under assumption (3.2), (3.4) and CFL condition (3.6), let u_h be given by (3.10), there exist $\mu_h \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}^+)$ and $\mu_\mathcal{T} \in \mathcal{M}(\mathbb{R}^N)$ such that:*

$$\begin{aligned}\int_{\mathbb{R}^N\times\mathbb{R}^+} \Big[|u_h - \kappa|\varphi_t(\mathrm{x},t) + (u_h(\mathrm{x},t)\top\kappa - u_h(\mathrm{x},t)\bot\kappa)\mathbf{a}(\mathrm{x},t)\cdot\nabla\varphi(\mathrm{x},t) \\ +(u_h(\mathrm{x},t)\top\kappa - u_h(\mathrm{x},t)\bot\kappa)\nabla\cdot\mathbf{a}(\mathrm{x},t)\varphi(\mathrm{x},t)\Big] d\mathrm{x}dt + \int_{\mathbb{R}^N} |u_0(\mathrm{x}) - \kappa|\varphi(\mathrm{x},0)d\mathrm{x} \\ \geq -\int_{\mathbb{R}^N\times\mathbb{R}^+} (|\varphi_t(\mathrm{x},t)| + |\nabla\varphi(\mathrm{x},t)|) d\mu_h(\mathrm{x},t) - \int_{\mathbb{R}^N} \varphi(\mathrm{x},0) d\mu_\mathcal{T}(\mathrm{x}),\end{aligned} \tag{3.37}$$

for all $\kappa \in \mathbb{R}$ and $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R}^+, \mathbb{R}^+)$. The measures μ_h and $\mu_\mathcal{T}$ satisfy the following properties:

1. *For all $R > 0$ and $T > 0$, there exists C_m depending only on $\mathbf{a}$, u_0, α, ξ, R, T such that, for $h \leq R$:*

$$\mu_h(B(0,R)\times[0,T]) \leq C_m(h + \sqrt{h}). \tag{3.38}$$

2. *The measure $\mu_\mathcal{T}$ is the measure of density $|u_0(\cdot) - u_{\mathcal{T},0}(\cdot)|$ w.r.t the Lebesgue measure. Moreover, for all $R > 0$, there exists D_m depending only on u_0, α and R such that:*

$$\mu_\mathcal{T}(B(0,R)) \leq D_m h |u_0|_{BV(B(0,R+4h))}. \tag{3.39}$$

Proof. Let $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R}^+, \mathbb{R}^+)$ and $\kappa \in \mathbb{R}$. Let T and R be such that $\varphi \neq 0$ implies $|x| \leq R - h$ and $t \leq T$. Let us multiply (3.32) by $\int_{t_n}^{t_{n+1}} \int_p \varphi(\mathrm{x},t) d\mathrm{x}dt$, and sum the result for all $p \in \mathcal{T}$ and $n \in \mathbb{N}$. One obtains:

$$T_1 + T_2 \leq 0, \tag{3.40}$$

with

$$T_1 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \frac{|u_p^{n+1} - \kappa| - |u_p^n - \kappa|}{\Delta t_n} \int_{t_n}^{t_{n+1}} \int_p \varphi(\mathrm{x},t) d\mathrm{x}dt, \tag{3.41}$$

$$T_2 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \frac{1}{|p|} \int_{t_n}^{t_{n+1}} \int_p \varphi(\mathrm{x},t) d\mathrm{x}dt \sum_{q\in\mathcal{N}(p)} a_{p,q}^n (u_{p,q}^n\top\kappa - u_p^n\top\kappa - u_{p,q}^n\bot\kappa + u_p\bot\kappa). \tag{3.42}$$

We set

$$T_1^* = -\int_{\mathbb{R}^N\times\mathbb{R}^+} |u_h(\mathrm{x},t) - \kappa|\varphi_t(\mathrm{x},t) d\mathrm{x}dt - \int_{\mathbb{R}^N} |u_0(\mathrm{x}) - \kappa|\varphi(\mathrm{x},0) d\mathrm{x} \tag{3.43}$$

$$\begin{aligned} T_2^* = &- \int_{\mathbb{R}^N\times\mathbb{R}^+} (u_h(\mathrm{x},t)\top\kappa - u_h(\mathrm{x},t)\bot\kappa)\mathbf{a}(\mathrm{x},t)\cdot\nabla\varphi(\mathrm{x},t)d\mathrm{x}dt \\ &- \int_{\mathbb{R}^N\times\mathbb{R}^+} (u_h(\mathrm{x},t)\top\kappa - u_h(\mathrm{x},t)\bot\kappa)\nabla\cdot\mathbf{a}(\mathrm{x},t)\varphi(\mathrm{x},t)d\mathrm{x}dt. \end{aligned} \tag{3.44}$$

Comparison between T_1 and T_1^*

We note that u_h is constant in $p\times[t_n,t_{n+1}[$ for all $p\in\mathcal{T}$ and $n\in\mathbb{N}$, then T_1^* may be rewritten

$$\begin{aligned} T_1^* &= -\sum_{p\in\mathcal{T}_R}\sum_{n=0}^{N_T}\int_p\int_{t_n}^{t_{n+1}} |u_p^n-\kappa|\varphi_t(\mathrm{x},t)dtd\mathrm{x} - \int_{\mathbb{R}^N}|u_0(\mathrm{x})-\kappa|\varphi(\mathrm{x},0)d\mathrm{x} \\ &= \sum_{p\in\mathcal{T}_R}\sum_{n=0}^{N_T}\int_p |u_p^n-\kappa|\big(\varphi(\mathrm{x},t_n)-\varphi(\mathrm{x},t_{n+1})\big)dtd\mathrm{x} - \int_{\mathbb{R}^N}|u_0(\mathrm{x})-\kappa|\varphi(\mathrm{x},0)d\mathrm{x} \end{aligned}$$

In the first term in the right-hand side of this expression, for all $n\in\overline{0,N_T}$, in each $p\in\mathcal{T}$, we factorize by $\varphi(\cdot,t_n)$. Recalling that $u_{\mathcal{T},0}(\mathrm{x}) = u_p^0\ \forall\mathrm{x}\in p$, and noting that $\varphi(\mathrm{x},t_{N_T+1})=0$ for all $\mathrm{x}\in\mathbb{R}^N$, we get:

$$\begin{aligned} T_1^* = \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}\frac{|u_p^{n+1}-\kappa|-|u_p^n-\kappa|}{\Delta t_n}\int_{t_n}^{t_{n+1}}\int_p\varphi(\mathrm{x},t_{n+1})d\mathrm{x}dt \\ + \int_{\mathbb{R}^N}\big(|u_{\mathcal{T},0}(\mathrm{x})-\kappa|-|u_0(\mathrm{x})-\kappa|\big)\varphi(\mathrm{x},0)d\mathrm{x}dt. \end{aligned} \tag{3.45}$$

Subtracting (3.41) and (3.45) side by side and applying the triangle inequality, there holds

$$\begin{aligned} |T_1-T_1^*| \le \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}\frac{|u_p^{n+1}-u_p^n|}{\Delta t_n}\int_{t_n}^{t_{n+1}}\int_p|\varphi(\mathrm{x},t_{n+1})-\varphi(\mathrm{x},t)|d\mathrm{x}dt \\ + \int_{\mathbb{R}^N}|u_{\mathcal{T},0}(\mathrm{x})-u_0(\mathrm{x})|\varphi(\mathrm{x},0)d\mathrm{x}. \end{aligned}$$

We know that, for all $t\in[t_n,t_{n+1}]$, there holds:

$$|\varphi(\mathrm{x},t_{n+1})-\varphi(\mathrm{x},t)| = |\int_t^{t_{n+1}}\varphi_s(\mathrm{x},s)ds| \le \int_{t_n}^{t_{n+1}}|\varphi_s(\mathrm{x},s)|ds.$$

Hence,

$$\begin{aligned} |T_1-T_1^*| \le \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}|u_p^{n+1}-u_p^n|\int_{t_n}^{t_{n+1}}\int_p|\varphi_t(\mathrm{x},t)|d\mathrm{x}dt \\ + \int_{\mathbb{R}^N}|u_{\mathcal{T},0}(\mathrm{x})-u_0(\mathrm{x})|\varphi(\mathrm{x},0)d\mathrm{x}. \end{aligned} \tag{3.46}$$

We define two measures $\mu_{\mathcal{T}}\in\mathcal{M}(\mathbb{R}^N)$ and $\lambda_h\in\mathcal{M}(\mathbb{R}^N\times\mathbb{R}^+)$ by their action on $C_c(\mathbb{R}^N)$ and $C_c(\mathbb{R}^N\times\mathbb{R}^+)$:

$$\langle \mu_{\mathcal{T}}, g\rangle = \int_{\mathbb{R}^N} |u_0(\mathrm{x}) - u_{\mathcal{T},0}(\mathrm{x})| g(\mathrm{x}) d\mathrm{x},\ \forall g \in C_c(\mathbb{R}^N), \tag{3.47}$$

$$\langle \lambda_h, g\rangle = \sum_{n\in\mathbb{N}} \sum_{p\in\mathcal{T}} |u_p^{n+1} - u_p^n| \int_{t_n}^{t_{n+1}} \int_p g(\mathrm{x},t) d\mathrm{x} dt,\ \forall g \in C_c(\mathbb{R}^N \times \mathbb{R}^+). \tag{3.48}$$

Inequality (3.46) gives

$$|T_1 - T_1^*| \le \int_{\mathbb{R}^N\times\mathbb{R}^+} |\varphi_t(\mathrm{x},t)| d\lambda_h(\mathrm{x},t) + \int_{\mathbb{R}^N} \varphi(\mathrm{x},0) d\mu_{\mathcal{T}}(\mathrm{x}). \tag{3.49}$$

Since $u_0 \in BV_{loc}(\mathbb{R}^N)$, we follow Lemma 6.8 in [3], then, there exists a constant D_m such that

$$\mu_{\mathcal{T}}(B(0,R)) \le D_m h |u_0|_{BV(B(0,R+4h))}.$$

Furthermore, we apply inequality (3.15), we have

$$\lambda_h(B(0,R)\times[0,T]) \le (1-\xi)\frac{\alpha^2 C_{bv}\sqrt{h}}{V},\quad \forall h < R,\ \forall R > 0.$$

Comparison between T_2 and T_2^*

In T_2 we gather the terms by edges. Since $u_{p,q}^n = u_{q,p}^n$ and $a_{p,q}^n = -a_{q,p}^n$, we have $T_2 = T_{21} - T_{22}$, with:

$$T_{21} = \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} \frac{a_{p,q}^n}{|p|} [u_{p,q}^n \top\kappa - u_p^n \top\kappa - u_{p,q}^n \bot\kappa + u_p \bot\kappa] \int_{t_n}^{t_{n+1}} \int_p \varphi(\mathrm{x},t) d\mathrm{x} dt,$$

$$T_{22} = \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} \frac{a_{p,q}^n}{|q|} [u_{p,q}^n \top\kappa - u_q^n \top\kappa - u_{p,q}^n \bot\kappa + u_q \bot\kappa] \int_{t_n}^{t_{n+1}} \int_q \varphi(\mathrm{x},t) d\mathrm{x} dt.$$

Applying Green's formula on each p in T_2^*, there holds

$$T_2^* = -\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} (u_p^n \top\kappa - u_p^n \bot\kappa) \sum_{q\in\mathcal{N}(p)} \int_{\sigma_{p,q}} \int_{t_n}^{t_{n+1}} \mathbf{a}(\gamma,t)\cdot\mathbf{n}_{p,q} \varphi(\gamma,t) dt d\gamma.$$

We can also gather the terms of T_2^* by edges and afterwards decompose T_2^* as $T_{21}^* - T_{22}^*$ with:

$$T_{21}^* = \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} \left(u_{p,q}^n \top\kappa - u_p^n \top\kappa - u_{p,q}^n \bot\kappa + u_p^n \bot\kappa \right) \int_{\sigma_{p,q}} \int_{t_n}^{t_{n+1}} \mathbf{a}(\gamma,t)\cdot\mathbf{n}_{p,q} \varphi(\gamma,t) dt d\gamma,$$

$$T_{22}^* = \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} \left(u_{p,q}^n \top\kappa - u_q^n \top\kappa - u_{p,q}^n \bot\kappa + u_q^n \bot\kappa \right) \int_{\sigma_{p,q}} \int_{t_n}^{t_{n+1}} \mathbf{a}(\gamma,t)\cdot\mathbf{n}_{p,q} \varphi(\gamma,t) dt d\gamma.$$

We introduce the differences of the averages of φ on p and on $\sigma_{p,q}$

$$r_{p,q}^n = \frac{a_{p,q}^n}{\Delta t_n |p|} \int_{t_n}^{t_{n+1}} \int_p \varphi(\mathrm{x},t) d\mathrm{x} dt - \frac{1}{\Delta t_n} \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} \mathbf{a}(\gamma,t)\cdot\mathbf{n}_{p,q} \varphi(\gamma,t) d\gamma dt. \tag{3.50}$$

Hence

$$T_{21} - T_{21}^* = \sum_{n=0}^{N_T} \Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n} \Big(u_{p,q}^n \top \kappa - u_p^n \top \kappa - u_{p,q}^n \bot \kappa + u_p^n \bot \kappa \Big) r_{p,q}^n. \tag{3.51}$$

From the definition of $u_{p,q}^n$ in (3.9), for all $\kappa \in \mathbb{R}$, we have the following inequalities:

$$|u_{p,q}^n \top \kappa - u_p^n \top \kappa| \le \max\{-sgn(a_{p,q}^n), 0\} |u_p^n - u_q^n|, \tag{3.52}$$
$$|u_{p,q}^n \bot \kappa - u_p^n \bot \kappa| \le \max\{-sgn(a_{p,q}^n), 0\} |u_p^n - u_q^n|, \tag{3.53}$$

where for all $a \in \mathbb{R}$, we set the unusual definition of $sgn(a)$:

$$sgn(a) = \begin{cases} 1 & \text{if } a > 0, \\ -1 & \text{if } a < 0, \\ -\frac{1}{2} & \text{if } a = 0. \end{cases}$$

In (3.50), we add the term $-\frac{a_{p,q}^n}{\Delta t_n |\sigma_{p,q}|} \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} \varphi(\gamma, s) d\gamma ds$ in the first term and the second term and change the integration variable t into s in the second term. There holds:

$$\begin{aligned} r_{p,q}^n = {} & \frac{a_{p,q}^n}{(\Delta t_n)^2 |\sigma_{p,q}||p|} \int_{t_n}^{t_{n+1}} \int_{t_n}^{t_{n+1}} \int_p \int_{\sigma_{p,q}} \big[\varphi(\mathrm{x}, t) - \varphi(\gamma, s)\big] d\gamma d\mathrm{x} dt ds \\ & - \frac{1}{(\Delta t_n)^2 |\sigma_{p,q}|} \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} \big[\mathbf{a}(\gamma, s) \cdot \mathbf{n}_{p,q} - \frac{a_{p,q}^n}{|\sigma_{p,q}|}\big] \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} \big[\varphi(\gamma, s) - \varphi(\zeta, \tau)\big] d\zeta d\tau \; d\gamma ds. \end{aligned} \tag{3.54}$$

From the definition of $a_{p,q}^n$, and the conditon in (3.2), there holds: for all $(\gamma, s) \in \sigma_{p,q} \times [t_n, t_{n+1}[$,

$$\begin{aligned} \Big| \mathbf{a}(\gamma, s) \cdot \mathbf{n}_{p,q} - \frac{a_{p,q}^n}{|\sigma_{p,q}|} \Big| &\le \frac{1}{\Delta t_n |\sigma_{p,q}|} \int_{t_n}^{t_{n+1}} \int_{\sigma_{p,q}} |\mathbf{a}(\gamma, s) - \mathbf{a}(\zeta, \tau)| d\zeta d\tau \\ &\le L_{\mathbf{a}} (\Delta t_n + h_{p,q}). \end{aligned} \tag{3.55}$$

Moreover, for all $(\gamma, s, \zeta, \tau) \in (\sigma_{p,q})^n \times [t_n, t_{n+1}[^2$, we have:

$$|\varphi(\gamma, s) - \varphi(\zeta, \tau)| \le \int_0^1 (\Delta t_n + h_{p,q}) (|\nabla_{\mathrm{x}} \varphi| + |\varphi_t|)(\zeta + \theta(\gamma - \zeta), \tau + \theta(s - \tau)) d\theta \tag{3.56}$$

and for all $(\mathrm{x}, \gamma, t, s) \in p \times \sigma_{p,q} \times [t_n, t_{n+1}[^2$:

$$|\varphi(\mathrm{x}, t) - \varphi(\gamma, s)| \le \int_0^1 (h_p + \Delta t_n) (|\nabla_{\mathrm{x}} \varphi| + |\varphi_t|)(\mathrm{x} + \theta(\gamma - \mathrm{x}), t + \theta(s - t)) d\theta \tag{3.57}$$

For all $p \in \mathcal{T}$, $q \in \mathcal{N}(p)$, $n \in \mathbb{N}$, we define some measures $\mu_{p,q}^n \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}^+)$ and $\nu_{p,q}^n \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}^+)$ by their action on $C_c(\mathbb{R}^N \times \mathbb{R}^+)$:

$$\langle \mu_{p,q}^n, g\rangle = \frac{1}{(\Delta t_n)^2|\sigma_{p,q}||p|}\int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}}\int_p\int_{\sigma_{p,q}}\int_0^1 (h_p+\Delta t_n)g(\mathrm{x}+\theta(\gamma-\mathrm{x}), \\ t+\theta(s-t))d\theta d\gamma d\mathrm{x}dtds, \tag{3.58}$$

$$\langle \nu_{p,q}^n, g\rangle = \frac{L_{\mathbf{a}}}{(\Delta t_n)^2|\sigma_{p,q}|}\int_{t_n}^{t_{n+1}}\int_{\sigma_{p,q}}\int_{t_n}^{t_{n+1}}\int_{\sigma_{p,q}}\int_0^1 (\sigma_{p,q}+\Delta t_n)^2 g(\zeta+\theta(\gamma-\zeta), \\ \tau+\theta(s-\tau))d\theta d\zeta d\tau d\gamma ds. \tag{3.59}$$

We have

$$\mu_{p,q}^n(B(0,R)\times[0,T]) \le h_p+\Delta t_n,\ \forall h<R \tag{3.60}$$

and

$$\nu_{p,q}^n(B(0,R)\times[0,T]) \le (h_{p,q}+\Delta t_n)^2,\ \forall h<R. \tag{3.61}$$

Combing (3.54), (3.55), (3.56), (3.57) and definition of $\mu_{p,q}^n$ and $\nu_{p,q}^n$, we have:

$$|r_{p,q}^n| \le |a_{p,q}^n|\langle \mu_{p,q}^n, |\nabla_{\mathrm{x}}\varphi|+|\varphi_t|\rangle + \langle \nu_{p,q}^n, |\nabla_{\mathrm{x}}\varphi|+|\varphi_t|\rangle \tag{3.62}$$

Then, (3.51)–(3.53) and (3.62) lead to

$$|T_{21}-T_{21}^*| \le 2\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}\max\{-sgn(a_{p,q}^n),0\}|a_{p,q}^n||u_p^n-u_q^n|\langle \mu_{p,q}^n, |\nabla\varphi|+|\varphi_t|\rangle \\ +2\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}\max\{-sgn(a_{p,q}^n),0\}|u_p^n-u_q^n|\langle \nu_{p,q}^n, |\nabla\varphi|+|\varphi_t|\rangle. \tag{3.63}$$

We obtain as well

$$|T_{22}-T_{22}^*| \le 2\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}\max\{-sgn(a_{q,p}^n),0\}|a_{p,q}^n||u_p^n-u_q^n|\langle \mu_{p,q}^n, |\nabla\varphi|+|\varphi_t|\rangle \\ +2\sum_{n=0}^{N_T}\Delta t_n \sum_{(p,q)\in\mathcal{E}_R^n}\max\{-sgn(a_{q,p}^n),0\}|u_p^n-u_q^n|\langle \nu_{q,p}^n, |\nabla\varphi|+|\varphi_t|\rangle. \tag{3.64}$$

Finally, we can define $\mu_h \in \mathcal{M}(\mathbb{R}^N\times\mathbb{R}^+)$ by its action on $C_c(\mathbb{R}^N\times\mathbb{R}^+)$:

$$\langle \mu_h, g\rangle = \langle \lambda_h, g\rangle + 2\sum_{n\in\mathbb{N}}\Delta t_n \sum_{(p,q)\in\mathcal{E}_*^n}|a_{p,q}^n||u_p^n-u_q^n| \\ \times\Big[\max\{-sgn(a_{p,q}^n),0\}\langle \mu_{p,q}^n, g\rangle + \max\{-sgn(a_{q,p}^n),0\}\langle \mu_{q,p}^n, g\rangle\Big] \\ +2\sum_{n\in\mathbb{N}}\Delta t_n \sum_{(p,q)\in\mathcal{E}_*^n}|u_p^n-u_q^n|\Big[\max\{-sgn(a_{p,q}^n),0\}\langle \nu_{p,q}^n, g\rangle + \max\{-sgn(a_{q,p}^n),0\}\langle \nu_{q,p}^n, g\rangle\Big]. \tag{3.65}$$

From inequality (3.14), (3.3.2), (3.60)–(3.61), and property of operator sgn, there exists constant C_m such that

$$\mu_h(B(0,R)) \leq C_m(h+\sqrt{h}),\ \forall h < R,\ R > 0. \tag{3.66}$$

Let us call A the left-hand side of (3.37), then

$$A = -T_1^* - T_2^* = -T_1 - T_2 + (T_1 - T_1^*) + (T_2 - T_2^*). \tag{3.67}$$

From (3.40), (3.49), (3.63) and (3.64), we obtain

$$A \geq -\int_{\mathbb{R}^N\times\mathbb{R}^+} (|\varphi_t(\mathrm{x},t)| + |\nabla\varphi(\mathrm{x},t)|)d\mu_h(\mathrm{x},t) - \int_{\mathbb{R}^N} \varphi(\mathrm{x},0)d\mu_{\mathcal{T}}(\mathrm{x}).$$

□

3.4 An a posteriori error estimation

3.4.1 An entropy estimate between approximation and entropy solution

We start with the following lemma which is useful for this subsection.

Lemma 3.6. *For all $\varepsilon > 0$, there exist $\bar{\rho}_1 \in C_c^\infty(\mathbb{R},\mathbb{R})$ and $\rho_N \in C_c^\infty(\mathbb{R}^N,\mathbb{R})$ such that*

$$\begin{aligned} &supp(\bar{\rho}_1) \subset [-1,0], && \bar{\rho}_1 \geq 0, \quad \int_{\mathbb{R}} \bar{\rho}_1(\mathrm{x})d\mathrm{x} = 1, \\ &supp(\rho_N) \subset \{\mathrm{x}\in\mathbb{R}^N :\ |\mathrm{x}| \leq 1\}, && \rho_N \geq 0, \quad \int_{\mathbb{R}^N} \rho_N(\mathrm{x})d\mathrm{x} = 1, \end{aligned} \tag{3.68}$$

with

$$\bar{K}_1 := \int_{\mathbb{R}^+} |\partial_t\bar{\rho}_1(t)|dt \leq 4+\varepsilon, \qquad K_N := \int_{\mathbb{R}^N} |\nabla\rho_N(\mathrm{x})|d\mathrm{x} \leq N+1+\varepsilon. \tag{3.69}$$

Theorem 3.7. *Assume (3.2) and $u_0 \in BV(\mathbb{R}^N)$. Let $\tilde{u} \in L^\infty(\mathbb{R}^N\times\mathbb{R})$ be such that $U_m \leq \tilde{u} \leq U_M$ a.e. Assume there exist $\mu \in \mathcal{M}(\mathbb{R}^N\times\mathbb{R}^+)$ and $\mu_0 \in \mathcal{M}(\mathbb{R}^N)$ such that*

$$\begin{aligned} &\int_{\mathbb{R}^N\mathbb{R}^+} \Big[|\tilde{u}(\mathrm{x},t)-\kappa|\varphi_t(\mathrm{x},t) + \big[\tilde{u}(\mathrm{x},t)\top\kappa - \tilde{u}(\mathrm{x},t)\bot\kappa\big]\mathbf{a}(\mathrm{x},t)\cdot\nabla\varphi(\mathrm{x},t) \\ &+ \big[\tilde{u}(\mathrm{x},t)\top\kappa - \tilde{u}(\mathrm{x},t)\bot\kappa\big]\nabla\cdot\mathbf{a}(\mathrm{x},t)\,\varphi(\mathrm{x},t)\Big]d\mathrm{x}dt + \int_{\mathbb{R}^N} |u_0(\mathrm{x})-\kappa|\varphi(\mathrm{x},0)d\mathrm{x} \\ &\geq -\int_{\mathbb{R}^N\times\mathbb{R}^+} \Big[|\varphi_t(\mathrm{x},t)| + |\nabla\varphi(\mathrm{x},t)|\Big]d\mu(\mathrm{x},t) - \int_{\mathbb{R}^N} |\varphi(\mathrm{x},0)|d\mu_0(\mathrm{x}) \end{aligned} \tag{3.70}$$

for all $\kappa \in \mathbb{R}$ and $\varphi \in C_c^1(\mathbb{R}^N\times\mathbb{R}^+,\mathbb{R}^+)$. Let u be the unique entropy weak solution of (3.1), that is:

$$\begin{aligned} &\int_{\mathbb{R}^N\times\mathbb{R}^+} \Big[|u(\mathrm{y},s)-\kappa|\varphi_s(\mathrm{y},s) + \big[u(\mathrm{y},s)\top\kappa - u(\mathrm{y},s)\bot\kappa\big]\mathbf{a}(\mathrm{y},s)\cdot\nabla\varphi(\mathrm{y},s) + \\ &\big[u(\mathrm{y},s)\top\kappa - u(\mathrm{y},s)\bot\kappa\big]\nabla\cdot\mathbf{a}(\mathrm{y},s)\varphi(\mathrm{y},s)\Big]dyds + \int_{\mathbb{R}^N} |u_0(\mathrm{y})-\kappa|\varphi(\mathrm{y},0)d\mathrm{y} \geq 0 \end{aligned} \tag{3.71}$$

for all $\kappa \in \mathbb{R}$, $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R}^+, \mathbb{R}^+)$. *Let* ω, R, $T \in \mathbb{R}^+$, $\mathrm{x}_0 \in \mathbb{R}^N$ *be given, and* $\rho \in C_0^1(\mathbb{R}^+, [0,1])$ *be such that* $\rho' \le 0$ *and*

$$\begin{aligned} &\rho = 1 \ \textit{on}\ [0,R], \\ &\rho = 0 \ \textit{on}\ [R+1,\infty], \\ &\|\rho'\|_\infty \le 2. \end{aligned} \tag{3.72}$$

Let

$$\begin{aligned} &\psi(\mathrm{x},t) := \frac{T-t}{T} e^{-Ht} \rho(|\mathrm{x}-\mathrm{x}_0| + \omega t) \ \textit{on}\ \mathbb{R}^N \times [0,T], \\ &\psi(\mathrm{x},t) := 0 \ \textit{on}\ \mathbb{R}^N \times [T,\infty), \end{aligned} \tag{3.73}$$

with $H = \max\big\{[\nabla \cdot \mathbf{a}(\mathrm{x},t)]^+ : \ \mathrm{x} \in \mathbb{R}^N, t \in \mathbb{R}^+\big\}$. *Then we have*

$$\begin{aligned} \int_{\mathbb{R}^N \times \mathbb{R}^+} &\Big[|\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)| \psi_t(\mathrm{x},t) + \big[\tilde{u}(\mathrm{x},t) \top u(\mathrm{x},t) - \tilde{u}(\mathrm{x},t) \bot u(\mathrm{x},t)\big] \\ &\times \mathbf{a}(\mathrm{x},t) \cdot \nabla \psi(\mathrm{x},t) + |\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)| \nabla \cdot \mathbf{a}(\mathrm{x},t) \psi(\mathrm{x},t) \Big] d\mathrm{x} dt \\ &\ge -a_0 \mu_0(\{\psi(\cdot,0) \ne 0\}) - (b+c)(\mu(\{\psi \ne 0\}))^{\frac{1}{2}} - a\mu(\{\psi \ne 0\}), \end{aligned} \tag{3.74}$$

where

$$\begin{aligned} a :=& 2\omega + \frac{1}{T} + H + 2, \\ a_0 :=& 1, \\ b :=& N + 5 \\ c :=& \Big[2\omega + \frac{1}{T} + H + 4\|\mathbf{a}\|_\infty + 3\|\nabla \cdot \mathbf{a}\|_\infty + L_{\mathbf{a}}(N+1)\Big] |u|_{BV(K_r)} \\ &+ \Big[(N+1)\|\mathbf{a}\|_\infty + \|\nabla \cdot \mathbf{a}\|_\infty + 1\Big] |u_0|_{BV(K_{0,r})}, \end{aligned} \tag{3.75}$$

with

$$\begin{aligned} K &= \{(\mathrm{x},t) \in \mathbb{R}^N \times \mathbb{R}^+ : \ \psi(\mathrm{x},t) \ne 0\}, \quad K_0 = \{\mathrm{x} \in \mathbb{R}^N : \ \psi(\mathrm{x},0) \ne 0\}, \\ K_{0,r} &= \{\mathrm{x} \in R^N : \exists \mathrm{y} \in K_0, \ d(\mathrm{x},\mathrm{y}) \le \frac{1}{r}\}, \\ K_r &= \{(\mathrm{x},t) \in \mathbb{R}^N \times \mathbb{R}^+ : \exists (\mathrm{y},s) \in K, \ d(\mathrm{x},\mathrm{y}) \le \frac{1}{r}, \ d(t,s) \le \frac{1}{r}\}, \\ r &= \sqrt{\frac{1}{\mu(K)}}. \end{aligned}$$

We remark that $|u|_{BV(K_r)}$ *can be estimated by an explicit constant in Theorem 3.13.*

Proof. For all $\varepsilon > 0$, let $\bar{\rho}_1 \in C_0^\infty(\mathbb{R}, \mathbb{R})$ and $\rho_p \in C_0^\infty(\mathbb{R}^p, \mathbb{R})$ satisfy Lemma 3.6. For $r \in \mathbb{R}, \ r > 0$, one defines $\bar{\rho}_{1,r}$, $\rho_{p,r}$ as

$$\bar{\rho}_{1,r}(t) = r\bar{\rho}_1(rt), \qquad \rho_{p,r}(\mathrm{x}) = r^p \rho_p(r\mathrm{x}). \tag{3.76}$$

One sets

$$\varphi(\mathrm{x},t,\mathrm{y},s)=\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s). \tag{3.77}$$

First of all, we state the following lemma.

Lemma 3.8. *From definitions (3.68), (3.72), (3.73) and (3.76), we have the following properties:*

$$\begin{aligned}&\|\psi\|_\infty\le 1, \qquad \|\psi(\cdot,0)\|_\infty\le 1,\\ &\|\psi_t\|_\infty\le 2\omega+\frac{1}{T}+H, \qquad \|\nabla\psi\|_\infty\le 2,\end{aligned} \tag{3.78}$$

$$\int_{\mathbb{R}^N\times\mathbb{R}^+}\rho_{p,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)d\mathrm{y}ds=1,\quad\forall\ (\mathrm{x},t)\in\mathbb{R}^N\times\mathbb{R}^+,\ \forall r>0, \tag{3.79}$$

$$\int_{\mathbb{R}^N\times\mathbb{R}^+}|\nabla\rho_{N,r}(\mathrm{x}-\mathrm{y})|\bar{\rho}_{1,r}(t-s)d\mathrm{y}ds\le rK_N,\quad\forall\ (\mathrm{x},t)\in\mathbb{R}^N\times\mathbb{R}^+,\ \forall r>0, \tag{3.80}$$

$$\int_{\mathbb{R}^N\times\mathbb{R}^+}\rho_{N,r}(\mathrm{x}-\mathrm{y})|\bar{\rho}_{t,1,r}(t-s)|d\mathrm{y}ds\le r\bar{K}_1,\quad\forall\ (\mathrm{x},t)\in\mathbb{R}^N\times\mathbb{R}^+,\ \forall r>0. \tag{3.81}$$

Now, let us take $\varphi(\cdot,\cdot,\mathrm{y},s)$ for the test function φ and $\kappa=u(\mathrm{y},s)$ in (3.70), then we integrate (3.70) for $(\mathrm{y},s)\in\mathbb{R}^N\times\mathbb{R}^+$; there holds:

$$\begin{aligned}&\int_{(\mathbb{R}^N\times\mathbb{R}^+)^2}\Big[|\tilde{u}(\mathrm{x},t)-u(\mathrm{y},s)|\psi_t(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)+\\ &\qquad|\tilde{u}(\mathrm{x},t)-u(\mathrm{y},s)|\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}'_{1,r}(t-s)+\\ &\qquad[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s)-\tilde{u}(\mathrm{x},t)\bot u(\mathrm{y},s)]\,\mathbf{a}(\mathrm{x},t)\cdot\nabla\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)+\\ &\qquad[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s)-\tilde{u}(\mathrm{x},t)\bot u(\mathrm{y},s)]\,\psi(\mathrm{x},t)\mathbf{a}(\mathrm{x},t)\cdot\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)+\\ &\qquad[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s)-\tilde{u}(\mathrm{x},t)\bot u(\mathrm{y},s)]\,\nabla\cdot\mathbf{a}(\mathrm{x},t)\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big]d\mathrm{x}dtd\mathrm{y}ds\\ &\qquad+\int_{\mathbb{R}^N\times\mathbb{R}^+\times\mathbb{R}^N}|u_0(\mathrm{x})-u(\mathrm{y},s)|\psi(\mathrm{x},0)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-s)d\mathrm{x}d\mathrm{y}ds\ge\\ &-\int_{\mathbb{R}^N\times\mathbb{R}^+}\int_{\mathbb{R}^N\times\mathbb{R}^+}\Big[\big|\rho_{N,r}(\mathrm{x}-\mathrm{y})\left[\psi_t(\mathrm{x},t)\bar{\rho}_{1,r}(t-s)+\psi(\mathrm{x},t)\bar{\rho}'_{1,r}(t-s)\right]\big|+\\ &\qquad\big|\bar{\rho}_{1,r}(t-s)\left[\nabla\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})+\psi(\mathrm{x},t)\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\right]\big|\Big]d\mu(\mathrm{x},t)d\mathrm{y}ds\\ &\qquad-\int_{\mathbb{R}^N\times\mathbb{R}^+}\int_{\mathbb{R}^N}\big|\psi(\mathrm{x},0)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-s)\big|d\mu_0(\mathrm{x})d\mathrm{y}ds.\end{aligned} \tag{3.82}$$

Let us take $\varphi(\mathrm{x},t,\cdot,\cdot)$ for the test function φ and $\kappa=\tilde{u}(\mathrm{x},t)$ in (3.71), then we integrate (3.71) for $(\mathrm{x},t)\in\mathbb{R}^N\times\mathbb{R}^+$; there holds:

$$\begin{aligned}&\int_{(\mathbb{R}^N\times\mathbb{R}^+)^2}\Big[-|\tilde{u}(\mathrm{x},t)-u(\mathrm{y},s)|\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}'_{1,r}(t-s)+\\ &\qquad[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s)-\tilde{u}(\mathrm{x},t)\bot u(\mathrm{y},s)]\,\psi(\mathrm{x},t)\mathbf{a}(\mathrm{y},s)\cdot\nabla_{\mathrm{y}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)+\\ &\qquad[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s)-\tilde{u}(\mathrm{x},t)\bot u(\mathrm{y},s)]\,\psi(\mathrm{x},t)\nabla\cdot\mathbf{a}(\mathrm{y},s)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big]d\mathrm{y}ds\mathrm{x}dt+\\ &\qquad\int_{\mathbb{R}^N\times\mathbb{R}^+\times\mathbb{R}^N}|u_0(\mathrm{y})-\tilde{u}(\mathrm{x},t)|\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t)d\mathrm{y}d\mathrm{x}dt\ge 0.\end{aligned} \tag{3.83}$$

The sum of the second term in the left-hand side of (3.82) and the first term in the left-hand side of (3.83) vanishes. Moreover $\bar{\rho}_{1,r}(t) = 0$ for all $t \in \mathbb{R}^+$, so that the final term in the left-hand side of (3.83) is zero. In addition, $\nabla_{\mathrm{y}}\rho_{N,r}(\mathrm{x}-\mathrm{y}) = -\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})$. Adding the inequalities (3.82) and (3.83) thus yields:

$$E_{11} + E_{12} + E_{13} + E_{14} + E_{15} \geq -E_2, \tag{3.84}$$

where:

$$E_{11} = \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[|\tilde{u}(\mathrm{x},t) - u(\mathrm{y},s)|\psi_t(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big] d\mathrm{x}dtd\mathrm{y}ds, \tag{3.85}$$

$$\begin{aligned} E_{12} = \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[&[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{y},s)] \\ &\mathbf{a}(\mathrm{x},t)\cdot\nabla\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big] d\mathrm{x}dtd\mathrm{y}ds, \end{aligned} \tag{3.86}$$

$$\begin{aligned} E_{13} = -\int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[&[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{y},s)]\,\psi(\mathrm{x},t) \\ &[\mathbf{a}(\mathrm{y},s) - \mathbf{a}(\mathrm{x},t)]\cdot\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big] d\mathrm{x}dtd\mathrm{y}ds + \\ \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[&[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{y},s)]\,\psi(\mathrm{x},t) \\ &\nabla\cdot\mathbf{a}(\mathrm{y},s)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big] d\mathrm{x}dtd\mathrm{y}ds, \end{aligned} \tag{3.87}$$

$$\begin{aligned} E_{14} = \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[&[\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{y},s)]\,\psi(\mathrm{x},t) \\ &\nabla\cdot\mathbf{a}(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big] d\mathrm{x}dtd\mathrm{y}ds, \end{aligned} \tag{3.88}$$

$$E_{15} = \int_{\mathbb{R}^N\times\mathbb{R}^+\times\mathbb{R}^N} |u_0(\mathrm{x}) - u(\mathrm{y},s)|\psi(\mathrm{x},0)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-s)d\mathrm{y}ds d\mathrm{x}, \tag{3.89}$$

$$\begin{aligned} E_2 = \int_{\mathbb{R}^N\times\mathbb{R}^+}\int_{\mathbb{R}^N\times\mathbb{R}^+} \Big[&\big|\rho_{N,r}(\mathrm{x}-\mathrm{y})\left[\psi_t(\mathrm{x},t)\bar{\rho}_{1,r}(t-s) + \psi(\mathrm{x},t)\bar{\rho}'_{1,r}(t-s)\right]\big| + \\ &\big|\bar{\rho}_{1,r}(t-s)\left[\nabla\psi(\mathrm{x},t)\rho_{N,r}(\mathrm{x}-\mathrm{y}) + \psi(\mathrm{x},t)\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\right]\big|\Big] d\mu(\mathrm{x},t)d\mathrm{y}ds \\ &+ \int_{\mathbb{R}^N\times\mathbb{R}^+}\int_{\mathbb{R}^N} \big|\psi(\mathrm{x},0)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-s)\big| d\mu_0(\mathrm{x})d\mathrm{y}ds. \end{aligned} \tag{3.90}$$

One has to study, now, the five terms of (3.84).
Using Lemma 3.8, equality (3.90) leads to:

$$\begin{aligned} E_2 \leq \int_{R^N\times\mathbb{R}^+} \Big[\big|\psi_t(\mathrm{x},t)\big| + \big|\nabla\psi(\mathrm{x},t)\big|\Big] d\mu(\mathrm{x},t) + r(K_N + \bar{K}_1)\int_{\mathbb{R}^N\times\mathbb{R}^+} \psi(\mathrm{x},t)d\mu(\mathrm{x},t) \\ + \int_{\mathbb{R}^N} \psi(\mathrm{x},0)d\mu_0(\mathrm{x}) \leq \Big\{\Big[2\omega + \frac{1}{T} + H + 2\Big] + r\Big[K_N + \bar{K}_1\Big]\Big\}\mu(K) + \mu_0(K_0). \end{aligned} \tag{3.91}$$

Let us now handle the term E_{11}; from (3.79), we obtain

$$\begin{aligned}
&\left|E_{11} - \int_{\mathbb{R}^N\times\mathbb{R}^+} |\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)|\psi_t(\mathrm{x},t)d\mathrm{x}dt\right| \\
&\leq \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[|u(\mathrm{x},t) - u(\mathrm{y},s)||\psi_t(\mathrm{x},t)|\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)d\mathrm{x}dtd\mathrm{y}ds\Big] \\
&= \int_{\mathbb{R}^N\times\mathbb{R}^+} \left[|\psi_t(\mathrm{x},t)| \left[\int_{\mathbb{R}^N\times\mathbb{R}^+} |u(\mathrm{x},t) - u(\mathrm{y},s)|\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)d\mathrm{y}ds\right]\right] d\mathrm{x}dt.
\end{aligned}$$

We change the integration variables (y,s) into (η,τ) such that $\mathrm{y} = \mathrm{x}+\eta$ and $s = t+\tau$, we rewrite the previous inequality as

$$\begin{aligned}
&\left|E_{11} - \int_{\mathbb{R}^N\times\mathbb{R}^+} |\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)|\psi_t(\mathrm{x},t)d\mathrm{x}dt\right| \\
&\leq \int_{\mathbb{R}^N\times\mathbb{R}^+} \left[|\psi_t(\mathrm{x},t)| \left[\int_{\mathbb{R}^N\times\mathbb{R}^+} |u(\mathrm{x},t) - u(\mathrm{x}+\eta,t+\tau)|\rho_{N,r}(-\eta)\bar{\rho}_{1,r}(-\tau)d\eta d\tau\right]\right] d\mathrm{x}dt \\
&= \int_{\mathbb{R}^N\times\mathbb{R}^+} \left[\rho_{N,r}(-\eta)\bar{\rho}_{1,r}(-\tau) \left[\int_{\mathbb{R}^N\times\mathbb{R}^+} |\psi_t(\mathrm{x},t)||u(\mathrm{x},t) - u(\mathrm{x}+\eta,t+\tau)|d\mathrm{x}dt\right]\right] d\eta d\tau.
\end{aligned}$$

From the definition of $\rho_{N,r}$ and $\bar{\rho}_{1,r}$, there holds:

$$\rho_{N,r}(-\eta) = 0 \text{ if } |\eta| \geq \frac{1}{r} \text{ and } \bar{\rho}_{1,r}(-\tau) = 0 \text{ if } \tau \geq \frac{1}{r}.$$

Then

$$\left|E_{11} - \int_{\mathbb{R}^N\times\mathbb{R}^+} |\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)|\psi_t(\mathrm{x},t)d\mathrm{x}dt\right| \leq \|\psi_t\|_\infty \varepsilon(r,K), \tag{3.92}$$

with

$$\varepsilon(r,K) = \sup_{(\eta,\tau)} \left\{\int_K |u(\mathrm{x},t) - u(\mathrm{x}+\eta,t+\tau)|d\mathrm{x}dt : \ |\eta| \leq \frac{1}{r},\ 0 \leq \tau \leq \frac{1}{r}\right\}. \tag{3.93}$$

Combining (3.92) and (3.78) yields:

$$\left|E_{11} - \int_{\mathbb{R}^N\times\mathbb{R}^+} \left[|\tilde{u}(\mathrm{x},t) - u(\mathrm{x},t)|\psi_t(\mathrm{x},t)\right] d\mathrm{x}dt\right| \leq (2\omega + \frac{1}{T} + H)\varepsilon(r,K). \tag{3.94}$$

To treat E_{12}, we may proceed in the same way. We note that $\tilde{u}(\mathrm{x},t)\top u(\mathrm{y},s) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{y},s)$ $= |\tilde{u}(\mathrm{x},t) - u(\mathrm{y},s)|$ and we can treat the term $\mathbf{a}(\mathrm{x},t)\cdot\nabla\psi(\mathrm{x},t)$ in E_{12} as we treated the term $\psi_t(\mathrm{x},t)$ in E_{11}; using that $\|\mathbf{a}\cdot\nabla\psi\|_\infty \leq 2\|\mathbf{a}\|_\infty$, one obtains:

$$\left|E_{12} - \int_{\mathbb{R}^N\times\mathbb{R}^+} \big[\tilde{u}(\mathrm{x},t)\top u(\mathrm{x},t) - \tilde{u}(\mathrm{x},t)\perp u(\mathrm{x},t)\big]\mathbf{a}(\mathrm{x},t)\cdot\nabla\psi(\mathrm{x},t)d\mathrm{x}dt\right| \leq 2\|\mathbf{a}\|_\infty \varepsilon(r,K). \tag{3.95}$$

Let us now turn to E_{13}. We compare this term with

$$\begin{aligned} E_{13b} = -\int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[\big[\tilde{u}(\mathrm{x},t)\top u(\mathrm{x},t) - \tilde{u}(\mathrm{x},t)\bot u(\mathrm{x},t)\big]\psi(\mathrm{x},t) \\ \times(\mathbf{a}(\mathrm{y},s)-\mathbf{a}(\mathrm{x},t))\cdot\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big]d\mathrm{x}dtd\mathrm{y}ds \\ + \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} \Big[\big[\tilde{u}(\mathrm{x},t)\top u(\mathrm{x},t) - \tilde{u}(\mathrm{x},t)\bot u(\mathrm{x},t)\big]\psi(\mathrm{x},t) \\ \times\nabla\cdot\mathbf{a}(\mathrm{y},s)\ \rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)\Big]d\mathrm{x}dtd\mathrm{y}ds. \end{aligned} \tag{3.96}$$

In the first integral of (3.96), we use $\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y}) + \nabla_{\mathrm{y}}\rho_{N,r}(\mathrm{x}-\mathrm{y}) = 0$, then we apply Green's formula with respect to y. Noting that $\nabla_{\mathrm{y}}\cdot(\mathbf{a}(\mathrm{y},s)-\mathbf{a}(\mathrm{x},t)) = \nabla_{\mathrm{y}}\cdot\mathbf{a}(\mathrm{y},s)$, one has $E_{13b}=0$. Subtracting E_{13b} from E_{13}, there holds

$$\begin{aligned} E_{13} \le \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} |u(\mathrm{x},t)-u(\mathrm{y},s)|\psi(\mathrm{x},t)|(\mathbf{a}(\mathrm{y},s)-\mathbf{a}(\mathrm{x},t))\cdot\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})|\bar{\rho}_{1,r}(t-s)d\mathrm{x}dtd\mathrm{y}ds \\ + \int_{(\mathbb{R}^N\times\mathbb{R}^+)^2} |u(\mathrm{x},t)-u(\mathrm{y},s)|\psi(\mathrm{x},t)|\nabla_{\mathrm{y}}\cdot\mathbf{a}(\mathrm{y},s)|\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(t-s)d\mathrm{x}dtd\mathrm{y}ds. \end{aligned} \tag{3.97}$$

The first term in the right-hand side of (3.97) is then smaller than $2K_NL_{\mathbf{a}}\|\psi\|_\infty\varepsilon(r,K)$ since $|(\mathbf{a}(\mathrm{y},s)-\mathbf{a}(\mathrm{x},t)|$ is bounded by $2L_{\mathbf{a}}/r$ and (3.80) holds. The second term can be treated in the same way as E_{11}. One has:

$$E_{13} \le (2L_{\mathbf{a}}K_N\|\psi\|_\infty + \|\nabla\cdot\mathbf{a}\|_\infty\|\psi\|_\infty)\varepsilon(r,K) \le (2L_{\mathbf{a}}K_N + \|\nabla\cdot\mathbf{a}\|_\infty)\varepsilon(r,K). \tag{3.98}$$

We can estimate E_{14} in the same way as we estimated E_{11}, replacing the expression $\varphi_t(\mathrm{x},t)$ by $\varphi(\mathrm{x},t)\nabla\cdot\mathbf{a}(\mathrm{x},t)$. There holds:

$$\left|E_{14} - \int_{\mathbb{R}^N\times\mathbb{R}^+} |\tilde{u}(\mathrm{x},t)-u(\mathrm{x},t)|\psi(\mathrm{x},t)\nabla\cdot\mathbf{a}(\mathrm{x},t)d\mathrm{x}dt\right| \le \|\nabla\cdot\mathbf{a}\|_\infty\varepsilon(r,K). \tag{3.99}$$

In order to study E_{15}, for any given $\mathrm{x}\in\mathbb{R}^N$, let us define $\varphi(\mathrm{x},\mathrm{y},s) = \psi(\mathrm{x},0)\rho_{N,r}(\mathrm{x}-\mathrm{y})\int_s^\infty\bar{\rho}_{1,r}(-\tau)d\tau$ and $\kappa = u_0(\mathrm{x})$. Let us take in (3.71) the function $\varphi(\mathrm{x},\cdot,\cdot)$ which is in $C^\infty(\mathbb{R}^N\times\mathbb{R}^+,\mathbb{R}^+)$ with a compact support in $\mathbb{R}^N\times\mathbb{R}^+$. We then integrate the resulting inequality with respect to $\mathrm{x}\in\mathbb{R}^N$. We get:

$$-E_{15} + E_{16} + E_{17} + E_{18} \ge 0, \tag{3.100}$$

with:

$$\begin{aligned} E_{16} = -\int_{\mathbb{R}^N\times\mathbb{R}^+\times\mathbb{R}^N}\int_s^\infty (u(\mathrm{y},s)\top u_0(\mathrm{x}) - u(\mathrm{y},s)\bot u_0(\mathrm{x}))\psi(\mathrm{x},0) \\ \times\mathbf{a}(\mathrm{y},s)\cdot\nabla_{\mathrm{x}}\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-\tau)d\tau d\mathrm{y}d\mathrm{x}ds, \end{aligned} \tag{3.101}$$

$$\begin{aligned} E_{17} = \int_{\mathbb{R}^N\times\mathbb{R}^+\times\mathbb{R}^N}\int_s^\infty (u(\mathrm{y},s)\top u_0(\mathrm{x}) - u(\mathrm{y},s)\bot u_0(\mathrm{x}))\psi(\mathrm{x},0) \\ \nabla\cdot\mathbf{a}(\mathrm{y},s)\rho_{N,r}(\mathrm{x}-\mathrm{y})\bar{\rho}_{1,r}(-\tau)d\tau d\mathrm{y}d\mathrm{x}ds, \end{aligned} \tag{3.102}$$

$$E_{18} = \int_{\mathbb{R}^N \times \mathbb{R}^+} \int_0^\infty |u_0(\mathrm{x}) - u_0(\mathrm{y})| \psi(\mathrm{x}, 0) \rho_{N,r}(\mathrm{x} - \mathrm{y}) \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x}. \tag{3.103}$$

In order to obtain a bound on E_{16}, one introduces E_{16b} defined as:

$$\begin{aligned} E_{16b} = \int_{\mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^N} \int_s^\infty (u(\mathrm{y}, s) \top u_0(\mathrm{y}) - u(\mathrm{y}, s) \perp u_0(\mathrm{y})) \psi(\mathrm{x}, 0) \\ \times \mathbf{a}(\mathrm{y}, s) \cdot \nabla_{\mathrm{x}} \rho_{N,r}(\mathrm{x} - \mathrm{y}) \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x} ds. \end{aligned} \tag{3.104}$$

Integrating by parts with respect to the x variable yields:

$$\begin{aligned} E_{16b} = - \int_{\mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^N} \int_s^\infty (u(\mathrm{y}, s) \top u_0(\mathrm{y}) - u(\mathrm{y}, s) \perp u_0(\mathrm{y})) \\ (\mathbf{a}(\mathrm{y}, s) \cdot \nabla_{\mathrm{x}} \psi(\mathrm{x}, 0)) \rho_{N,r}(\mathrm{x} - \mathrm{y}) \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x} ds. \end{aligned} \tag{3.105}$$

Then, noting that the time support of this integration is reduced to $[0, \frac{1}{r}]$, one has:

$$E_{16b} \leq \|\mathbf{a}\|_\infty \|\nabla \psi\|_\infty \varepsilon(r, K) \leq 2 \|\mathbf{a}\|_\infty \varepsilon(r, K). \tag{3.106}$$

Furthermore, one has:

$$\begin{aligned} |E_{16} + E_{16b}| \leq \int_{\mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^N} \int_s^\infty |u_0(\mathrm{x}) - u_0(\mathrm{y})| \psi(\mathrm{x}, 0) \\ |\mathbf{a}(\mathrm{y}, s) \cdot \nabla_{\mathrm{x}} \rho_{N,r}(\mathrm{x} - \mathrm{y})| \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x} ds, \end{aligned} \tag{3.107}$$

which is bounded by $K_N \|\mathbf{a}\|_\infty \varepsilon_0(r, K_0)$ (since the time support of the integration is reduced to $[0, \frac{1}{r}]$), where $\varepsilon_0(r, K_0)$ is defined by:

$$\varepsilon_0(r, K_0) = \sup_\eta \left\{ \int_{K_0} |u_0(\mathrm{x}) - u_0(\mathrm{x} + \eta)| d\mathrm{x} : \ |\eta| \leq \frac{1}{r} \right\}. \tag{3.108}$$

Combining (3.106) and (3.107), we obtain

$$E_{16} \leq \|\mathbf{a}\|_\infty \big[2\varepsilon(r, K) + K_N \varepsilon_0(r, K_0) \big]. \tag{3.109}$$

In the same way, one introduces E_{17b} defined as:

$$\begin{aligned} E_{17b} = \int_{\mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^N} \int_s^\infty (u(\mathrm{y}, s) \top u_0(\mathrm{y}) - u(\mathrm{y}, s) \perp u_0(\mathrm{y})) \psi(\mathrm{x}, 0) \\ \nabla \cdot \mathbf{a}(\mathrm{y}, s) \rho_{N,r}(\mathrm{x} - \mathrm{y}) \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x} ds. \end{aligned} \tag{3.110}$$

There holds

$$E_{17b} \leq \|\nabla \cdot \mathbf{a}\|_\infty \|\psi(\cdot, 0)\|_\infty \varepsilon(r, K) \leq \|\nabla \cdot \mathbf{a}\|_\infty \varepsilon(r, K) \tag{3.111}$$

and

$$\begin{aligned} |E_{17} - E_{17b}| \leq \int_{\mathbb{R}^N \times \mathbb{R}^+ \times \mathbb{R}^N} \int_s^\infty |u_0(\mathrm{x}) - u_0(\mathrm{y})| \psi(\mathrm{x}, 0) \\ |\nabla \cdot \mathbf{a}(\mathrm{y}, s)| \rho_{N,r}(\mathrm{x} - \mathrm{y}) \bar{\rho}_{1,r}(-\tau) d\tau d\mathrm{y} d\mathrm{x} ds, \end{aligned}$$

which is bounded by $\|\psi(\cdot,0)\|_\infty\|\nabla\cdot\mathbf{a}\|_\infty\varepsilon_0(r,K_0)$. With (3.111), one has

$$E_{17}\le\|\nabla\cdot\mathbf{a}\|_\infty\big[\varepsilon(r,K)+\varepsilon_0(r,K_0)\big]. \tag{3.112}$$

It is easily seen that

$$E_{18}\le\varepsilon_0(r,K_0). \tag{3.113}$$

From (3.100), (3.109), (3.112) and (3.113), there holds

$$E_{15}\le\big[2\|\mathbf{a}\|_\infty+\|\nabla\cdot\mathbf{a}\|_\infty\big]\varepsilon(r,K)+\big[K_N\|\mathbf{a}\|_\infty+\|\nabla\cdot\mathbf{a}\|_\infty+1\big]\varepsilon_0(r,K_0). \tag{3.114}$$

Using (3.84), (3.91), (3.94), (3.95), (3.98), (3.99), (3.114), there holds

$$\begin{aligned}\int_{\mathbb{R}^N\times\mathbb{R}^+}\Big[&|\tilde u(\mathrm{x},t)-u(\mathrm{x},t)|\psi_t(\mathrm{x},t)+[\tilde u(\mathrm{x},t)\top u(\mathrm{x},t)-\tilde u(\mathrm{x},t)\perp u(\mathrm{x},t)]\\ &\times\mathbf{a}(\mathrm{x},t)\cdot\nabla\psi(\mathrm{x},t)+|\tilde u(\mathrm{x},t)-u(\mathrm{x},t)|\nabla\cdot\mathbf{a}(\mathrm{x},t)\psi(\mathrm{x},t)\Big]d\mathrm{x}dt\\ &\ge-(2\omega+\frac{1}{T}+H+2+r(K_N+\bar K_1))\mu(K)-\mu_0(K_0)-E,\end{aligned} \tag{3.115}$$

with
$E=\Big[2\omega+\frac{1}{T}+H+4\|\mathbf{a}\|_\infty+3\|\nabla\cdot\mathbf{a}\|_\infty+2L_{\mathbf{a}}K_N\Big]\varepsilon(r,K)+\Big[K_N\|\mathbf{a}\|_\infty+\|\nabla\cdot\mathbf{a}\|_\infty+1\Big]\varepsilon_0(r,K_0)$.
Since $u_0\in BV_{\mathrm{loc}}(\mathbb{R}^N)$, From Theorem 3.13 and Lemma 3.12, there holds:

$$\varepsilon_0(r,K_0)\le\frac{|u_0|_{BV(K_0,r)}}{r},\qquad\varepsilon(r,K)\le\frac{|u|_{BV(K_r)}}{r}. \tag{3.116}$$

And From (3.69), we have:

$$\bar K_1+K_N\le N+5+2\varepsilon. \tag{3.117}$$

Choosing $r=\sqrt{\frac{1}{\mu(K)}}$, and realizing that the left-hand side in (3.115) does not depend on ε, then letting ε tend to zero, we obtain (3.74).
This concludes the proof of Theorem 3.7. □

3.4.2 An a posteriori error estimation

Now we can present the a posteriori error estimate, the main result of this document. Let R, ω, T and x_0 be defined in Theorem 3.7, it is convenient to set

$$\begin{aligned}I_0&:=\left\{n:\ 0\le t_n\le\min\left\{\frac{R+1}{\omega},T\right\}\right\},\\ D_{R+1}&:=\{(\mathrm{x},t):\ |\mathrm{x}-\mathrm{x}_0|+\omega t<R+1\},\\ M(t)&:=\{p\in\mathcal{T}:\ \text{ there exists }\mathrm{x}\in p\text{ such that }(\mathrm{x},t)\in D_{R+1}\}.\end{aligned} \tag{3.118}$$

Theorem 3.9. *Assume (3.2) and $u_0 \in BV(\mathbb{R}^N)$. Assume there exist $\mu \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}^+)$ and $\mu_0 \in \mathcal{M}(\mathbb{R}^N)$ satisfying (3.70) and let $u \in BV(\mathbb{R}^N \times \mathbb{R}^+)$ be the unique entropy weak solution of (3.1) and let u_h be given by (3.10). Let $K \subset\subset \mathbb{R}^N \times \mathbb{R}^+$, $\omega = V$ and choose T, R and x_0 such that $T \in]0, \frac{R}{\omega}[$ and*

$$K \subset \bigcup_{0 \le t \le T} B_{R-\omega t}(\mathrm{x}_0) \times \{t\}. \tag{3.119}$$

Then we have

$$\int_K |u(\mathrm{x},t) - u_h(\mathrm{x},t)| d\mathrm{x}dt \le T e^{TH} \left[a_0 \int_{|\mathrm{x}-\mathrm{x}_0| \le R+1} |u_0(\mathrm{x}) - u_{\mathcal{T},0}(\mathrm{x})| d\mathrm{x} + (b+c)\sqrt{Q} + aQ \right] \tag{3.120}$$

where a_0, a, b, c are defined in (3.75),

$$\begin{aligned} Q = & \sum_{n \in I_0} \sum_{p \in M(t_n)} |u_p^{n+1} - u_p^n| \Delta t_n |p| \\ & + 2 \sum_{n \in I_0} \sum_{(p,q) \in \mathcal{E}} \Big[|a_{p,q}^n| |u_p^n - u_q^n| \Delta t_n (\Delta t_n + \max\{h_p, h_q\}) |\sigma_{p,q}| \delta_{p,q}^n \Big] \\ & + 2 L_{\mathbf{a}} \sum_{n \in I_0} \sum_{(p,q) \in \mathcal{E}} \Big[|u_p^n - u_q^n| \Delta t_n (\Delta t_n + h_{p,q})^2 |\sigma_{p,q}| \delta_{p,q}^n \Big], \end{aligned} \tag{3.121}$$

and

$$\begin{aligned} & \delta_{p,q}^n = 0 \quad \textit{if} \quad p \cup q \times [t_n, t_{n+1}] \cap D_{R+1} = \emptyset, \\ & \delta_{p,q}^n = 1 \quad \textit{otherwise.} \end{aligned} \tag{3.122}$$

Proof. Let ψ be defined by (3.73). Theorem 3.7 holds for $\tilde{u}(\mathrm{x},t) := u_h(\mathrm{x},t)$, $\mu_0 = \mu_{\mathcal{T}}$ and $\mu = \mu_h$, where $\mu_{\mathcal{T}}$, μ_h are defined by (3.47) and (3.65). This means that

$$\begin{aligned} & \int_{\mathbb{R}^N \times \mathbb{R}^+} \Big[|u_h(\mathrm{x},t) - u(\mathrm{x},t)| \Big[\frac{T-t}{T} e^{-tH} \rho' \omega - \frac{1}{T} e^{-tH} \rho - H \frac{t-T}{T} e^{-tH} \rho \Big] \\ & + \big[u_h(\mathrm{x},t) \top u(\mathrm{x},t) - u_h(\mathrm{x},t) \perp u(\mathrm{x},t) \big] \frac{T-t}{T} \rho \, \mathbf{a}(\mathrm{x},t) \cdot \frac{\mathrm{x}-\mathrm{x}_0}{|\mathrm{x}-\mathrm{x}_0|} e^{-tH} \\ & + |\tilde{u}_h(\mathrm{x},t) - u(\mathrm{x},t)| \nabla \cdot \mathbf{a}(\mathrm{x},t) \frac{T-t}{T} e^{-Ht} \rho \Big] d\mathrm{x}dt \\ & \ge -a_0 \mu_{\mathcal{T}}(\{\psi \ne 0\}) - (b+c)(\mu_h(\{\psi \ne 0\}))^{1/2} - a \mu_h(\{\psi \ne 0\}), \end{aligned} \tag{3.123}$$

with $\rho = \rho(|\mathrm{x}-\mathrm{x}_0| + \omega t)$ and $\rho' = \rho'(|\mathrm{x}-\mathrm{x}_0| + \omega t)$. Let us estimate the second term in the left-hand side of the previous inequality. One has:

$$\begin{aligned} & \big[u_h(\mathrm{x},t) \top u(\mathrm{x},t) - u_h(\mathrm{x},t) \perp u(\mathrm{x},t) \big] \frac{T-t}{T} \rho' \mathbf{a}(\mathrm{x},t) \cdot \frac{\mathrm{x}-\mathrm{x}_0}{|\mathrm{x}-\mathrm{x}_0|} e^{-tH} \\ & \qquad \le V |u_h(\mathrm{x},t) - u(\mathrm{x},t)| \frac{T-t}{T} |\rho'| e^{-tH} \\ & \qquad \le -\omega |u_h(\mathrm{x},t) - u(\mathrm{x},t)| \frac{T-t}{T} \rho' e^{-tH}. \end{aligned} \tag{3.124}$$

So that

$$\begin{aligned}\big|u_h(\mathrm{x},t)-u(\mathrm{x},t)\big|\frac{T-t}{T}e^{-tH}\rho'\omega+\big[u_h(\mathrm{x},t)\top u(\mathrm{x},t)-u_h(\mathrm{x},t)\perp u(\mathrm{x},t)\big]\\ \times\frac{T-t}{T}\rho\,\mathbf{a}(\mathrm{x},t)\cdot\frac{\mathrm{x}-\mathrm{x}_0}{|\mathrm{x}-\mathrm{x}_0|}e^{-tH}\le 0,\quad\forall\ (\mathrm{x},t)\in\mathbb{R}^N\times[0,T].\end{aligned}\tag{3.125}$$

In addition, from the definition of H, we have

$$|u_h(\mathrm{x},t)-u(\mathrm{x},t)|(-H+\nabla\cdot\mathbf{a}(\mathrm{x},t))\frac{T-t}{T}e^{-tH}\rho\le 0,\quad\forall\ 0\le t\le T,\ \mathrm{x}\in\mathbb{R}^N.\tag{3.126}$$

Combining (3.123), (3.125) and (3.126), there holds

$$\begin{aligned}&\int_{\mathbb{R}^N\times\mathbb{R}^+}|u_h(\mathrm{x},t)-u(\mathrm{x},t)|e^{-tH}\rho d\mathrm{x}dt\\ &\le T\big[a_0\mu_{\mathcal{T}}(\{\psi\neq 0\})+(b+c)(\mu_h(\{\psi\neq 0\}))^{1/2}+a\mu_h(\{\psi\neq 0\})\big].\end{aligned}\tag{3.127}$$

If $(\mathrm{x},t)\in K$, there holds $(\mathrm{x},t)\in B_{R-\omega t}(\mathrm{x}_0)\times\{t\}$, $\rho(|\mathrm{x}-\mathrm{x}_0|+\omega t)=1$, and $e^{-tH}\ge e^{-HT}$, then

$$\int_K|u_h(\mathrm{x},t)-u(\mathrm{x},t)|d\mathrm{x}dt\le e^{TH}\int_{\mathbb{R}^N\times\mathbb{R}^+}|u_h(\mathrm{x},t)-u(\mathrm{x},t)|e^{-tH}\rho d\mathrm{x}dt.\tag{3.128}$$

Hence

$$\begin{aligned}&\int_K|u_h(\mathrm{x},t)-u(\mathrm{x},t)|d\mathrm{x}dt\\ &\le Te^{TH}\big[a_0\mu_{\mathcal{T}}(\{\psi(\cdot,0)\neq 0\})+(b+c)(\mu_h(\{\psi\neq 0\}))^{1/2}+a\mu_h(\{\psi\neq 0\})\big].\end{aligned}\tag{3.129}$$

Now the theorem follows from the following two lemmas.

Lemma 3.10. *Let ψ be defined as in (3.73). Then*

$$\mu_{\mathcal{T}}(\{\psi(\cdot,0)\neq 0\})\le\int_{|\mathrm{x}-\mathrm{x}_0|\le R+1}|u_{\mathcal{T},0}(\mathrm{x})-u_0(\mathrm{x})|d\mathrm{x}.\tag{3.130}$$

Proof. Using the definition of ρ in (3.72), we obtain

$$\mu_{\mathcal{T}}(\{\psi(\cdot,0)\neq 0\})\le\mu_{\mathcal{T}}(\{|\mathrm{x}-\mathrm{x}_0|+\omega t<R+1|_{t=0}\})$$

and from the definition of $\mu_{\mathcal{T}}$, we have

$$\mu_{\mathcal{T}}(\{|\mathrm{x}-\mathrm{x}_0|<R+1\})=\int_{|\mathrm{x}-\mathrm{x}_0|<R+1}|u_{\mathcal{T},0}(\mathrm{x})-u_0(\mathrm{x})|d\mathrm{x}.$$

□

Lemma 3.11. *Let ψ be defined as in (3.73). Then*

$$\begin{aligned}\mu_h(\{\psi\neq 0\}) \leq & \sum_{n\in I_0}\sum_{p\in M(t_n)} \Delta t_n |u_p^{n+1}-u_p^n||p| \\ & + 2\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}}\Big[|a_{p,q}^n||u_p^n-u_q^n|\Delta t_n(\Delta t_n+\max\{h_p,h_q\})|\sigma_{p,q}|\delta_{p,q}^n\Big] \\ & 2L_{\mathbf{a}}\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}}\Big[|u_p^n-u_q^n|\Delta t_n(\Delta t_n+h_{p,q})^2|\sigma_{p,q}|\delta_{p,q}^n\Big].\end{aligned} \tag{3.131}$$

Proof. The definition of μ_h is given by (3.65). Let us consider λ_h, which is part of μ_h (see definition in (3.48)). Since $D_{R+1}=\{(\mathrm{x},t):\ |\mathrm{x}-\mathrm{x}_0|+\omega t<R+1\}$, we obtain

$$\begin{aligned}\lambda_h(\{\psi\neq 0\}) \leq \lambda_h(D_{R+1}) &= \sum_{n\in I_0}\sum_{p\in M(t_n)} |u_p^{n+1}-u_p^n|\int_{t_n}^{t_{n+1}}\int_p \chi_{D_{R+1}}d\mathrm{x}dt \\ &\leq \sum_{n\in I_0}\sum_{p\in M(t_n)} \Delta t_n|u_p^{n+1}-u_p^n||p|.\end{aligned} \tag{3.132}$$

This proves the estimate concerning λ_h. Now we have to estimate $\mu_{p,q}^n$ which is also part of μ_h. This measure is defined in (3.58) and we have

$$\begin{aligned}\langle\mu_{p,q}^n,\psi\rangle = \frac{1}{(\Delta t_n)^2|\sigma_{p,q}||p|}\int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}}\int_p\int_{\sigma_{p,q}}\int_0^1 (\Delta t_n+h_p)\psi(\mathrm{x}+\theta(\gamma-x),\\ t+\theta(s-t))d\theta d\gamma d\mathrm{x}dtds.\end{aligned}$$

First we consider the case when $p\times[t_n,t_{n+1}]\cap D_{R+1}=\emptyset$ and $q\times[t_n,t_{n+1}]\cap D_{R+1}=\emptyset$. Since $(\mathrm{x},t)\in p\times[t_n,t_{n+1}[$, $(\gamma,s)\in\sigma_{p,q}\times[t_n,t_{n+1}[$, p is convex and $0\leq\theta\leq 1$, we get

$$(\mathrm{x}+\theta(\gamma-\mathrm{x}),t+\theta(s-t))\in p\times[t_n,t_{n+1}]$$

and therefore

$$\psi(\mathrm{x}+\theta(\gamma-\mathrm{x}),t+\theta(s-t))=0.$$

Otherwise, there holds

$$|\psi(\mathrm{x}+\theta(\gamma-\mathrm{x}),t+\theta(s-t))|\leq 1.$$

Hence, using the definition of $\delta_{p,q}$ in (3.122), we obtain

$$\langle\mu_{p,q}^n,\psi\rangle \leq (\Delta t_n+h_p)\delta_{p,q}^n. \tag{3.133}$$

In the same way, one has

$$\langle\mu_{q,p}^n,\psi\rangle \leq (\Delta t_n+h_q)\delta_{p,q}^n, \tag{3.134}$$

$$\langle\nu_{p,q}^n,\psi\rangle \leq L_{\mathbf{a}}(\Delta t_n+h_{p,q})^2|\sigma_{p,q}|\delta_{p,q}^n, \tag{3.135}$$

$$\langle\nu_{q,p}^n,\psi\rangle \leq L_{\mathbf{a}}(\Delta t_n+h_{p,q})^2|\sigma_{p,q}|\delta_{p,q}^n. \tag{3.136}$$

Using the estimates (3.132), (3.133), (3.134) and the definition of μ_h, one has

$$\begin{aligned}\mu_h(\{\psi \neq 0\}) \leq & \sum_{n\in I_0}\sum_{p\in M(t_n)} \Delta t_n |u_p^{n+1} - u_p^n||p| \\ & +2\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}_*^n}\Big[\max\{-sgn(a_{p,q}^n),0\}|a_{p,q}^n||u_p^n - u_q^n|\Delta t_n(\Delta t_n + h_p)|\sigma_{p,q}|\delta_{p,q}^n\Big] \\ & +2\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}_*^n}\Big[\max\{-sgn(a_{q,p}^n),0\}|a_{p,q}^n||u_p^n - u_q^n|\Delta t_n(\Delta t_n + h_q)|\sigma_{p,q}|\delta_{p,q}^n\Big] \\ & +2L_{\mathbf{a}}\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}_*^n}\Big[(\max\{-sgn(a_{p,q}^n),0\} + \max\{-sgn(a_{q,p}^n),0\})\Delta t_n(\Delta t_n + h_{p,q})^2|\sigma_{p,q}|\delta_{p,q}^n\Big].\end{aligned}$$

From the definitions of $\mathcal{E}_*^n$ and noting that $\max\{-sgn(a_{p,q}^n),0\} + \max\{-sgn(a_{q,p}^n),0\} = 1$, there holds

$$\begin{aligned}\mu_h(\{\psi \neq 0\}) \leq & \sum_{n\in I_0}\sum_{p\in M(t_n)} \Delta t_n |u_p^{n+1} - u_p^n||p| \\ & +2\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}}\Big[|a_{p,q}^n||u_p^n - u_q^n|\Delta t_n(\Delta t_n + \max\{h_p,h_q\})\delta_{p,q}^n\Big] \\ & +2L_{\mathbf{a}}\sum_{n\in I_0}\sum_{(p,q)\in\mathcal{E}}\Big[|u_p^n - u_q^n|\Delta t_n(\Delta t_n + h_{p,q})^2|\sigma_{p,q}|\delta_{p,q}^n\Big].\end{aligned}$$

This completes this lemma. □

Now the proof of Theorem 3.9 follows if we use (3.129), (3.130) and (3.131). □

3.5 Properties of the entropy solution

Our problem is a special case of what the authors of [2] call the non stiff case, which they discretized by the explicit upwind scheme. But in that paper, the authors didn't give the reason why the term $\varepsilon(r,K)$ defined by (3.93) can bounded by an explicit constant that depends on $\mathbf{a}$, r, K and u_0. In this section, we will determine this constant. We note that we will use one property obtained in [2] for the non stiff case: the convergence of the approximation solution to the entropy solution in weak L^∞.

We follow the proof presented by Chainais–Hillairet in [1], but we do not suppose that the velocity is divergence free and we were able to improve the bounds given in [1].

We start with the following lemma:

Lemma 3.12. *Let $v \in BV_{\mathrm{loc}}(\mathbb{R}^p)$, $p \in \mathbb{N}^*$. Then, for every bounded subset Ω of $\mathbb{R}^p$ and for all $\eta > 0$,*

$$\int_\Omega |v(\mathrm{x}+\eta) - v(\mathrm{x})|d\mathrm{x} \leq |v|_{BV(\Omega_\eta)}|\eta|, \tag{3.137}$$

where $\Omega_\eta = \{\mathrm{x} \in \mathbb{R}^p;\ d(\mathrm{x},\Omega) \leq \eta\}$ and $d(\mathrm{x},\Omega) = \inf\{|\mathrm{x}-\mathrm{y}|,\ y \in \Omega\}$ is the distance from x to Ω. This Lemma was proved in [3, Lemma 6.9].

Let us now consider the main theorem of this section.

Theorem 3.13. *Under assumption (3.2), then the entropy solution u to (3.1) belongs to $BV_{\mathrm{loc}}(\mathbb{R}^N \times [0,T])$ for all $T > 0$. Further, for all compact set $K \subset \mathbb{R}^N \times \mathbb{R}^+$, there exists an explicit constant $C_{K,u_0,\mathbf{a}}$ such that*

$$|u|_{BV(K)} \leq C_{K,u_0,\mathbf{a}}. \tag{3.138}$$

Proof. Theorem 3.13 is proved through the following steps: first, we build a discrete solution of the transport equation by a particular scheme and on a particular family of meshes. This solution satisfies a strong BV estimate and verifies a property like (3.37). Secondly, using the uniqueness of the entropy solution and the properties of the dicrete solution, then, when the size of the particular family of meshes tends to 0, the discrete solution will converge to the unique entropy solution, and moreover, the discrete solution will transfer its properties to the entropy solution.

For the sake of simplicity, we will prove the theorem with $N = 2$. Let us set $z = (x, y)$ a point in $\mathbb{R}^2$. We consider Cartesian meshes made up of squares of size h. We denote p_{ij} to be the cell of center $(x_i = ih,\ y_j = jh)$ with $(i,j) \in \mathbb{Z}^2$. The vertices of this cell are the points with coordinates $(x_{i-\frac{1}{2}}, y_{j-\frac{1}{2}})$, $(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}})$, $(x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}})$, $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$. We set k to be the time step and $t_n = kn$ for all $n \in \mathbb{N}$.

Let the transport velocity be $\mathbf{a}(z,t) = (a_1(z,t), a_2(z,t))$; then the problem (3.1) rewrites:

$$\frac{\partial u}{\partial t} + (a_1(z,t)u(z,t))_x + (a_2(z,t)u(z,t))_y = [(a_1(z,t))_x + (a_2(z,t))_y]u(z,t), \forall (z,t) \in \mathbb{R}^2 \times \mathbb{R}^+,$$
$$u(z,0) = u_0(z),\ \forall z \in \mathbb{R}^2.$$

We split each component of velocity $\mathbf{a}$ into two parts: the first part must be non negative and the second part non positive:

$$a_1(z,t) = b(z,t) + c(z,t),$$
$$a_2(z,t) = d(z,t) + f(z,t).$$

For instance, we can take:

$$\begin{aligned} b(z,t) &= \frac{1}{2}(a_1(z,t) + V), \qquad & c(z,t) &= \frac{1}{2}(a_1(z,t) - V), \\ d(z,t) &= \frac{1}{2}(a_2(z,t) + V), \qquad & f(z,t) &= \frac{1}{2}(a_2(z,t) - V), \end{aligned} \tag{3.139}$$

where V is defined by (3.2).

The discrete unknowns are the $u^n_{i,j}$, where $(i,j) \in \mathbb{Z}^2$ and $n \in \mathbb{N}$. From the positive and negative property of the functions $b,\ c,\ d$ and f, we construct the explicit upwind scheme:

$$\begin{aligned} u^{n+1}_{i,j} = u^n_{i,j} &- \frac{k}{h}\left[b^n_{i+\frac{1}{2},j}u^n_{i,j} - b^n_{i-\frac{1}{2},j}u^n_{i-1,j} + c^n_{i+\frac{1}{2},j}u^n_{i+1,j} - c^n_{i-\frac{1}{2},j}u^n_{i,j}\right] \\ &- \frac{k}{h}\left[d^n_{i,j+\frac{1}{2}}u^n_{i,j} - d^n_{i,j-\frac{1}{2}}u^n_{i,j-1} + f^n_{i,j+\frac{1}{2}}u^n_{i,j+1} - f^n_{i,j-\frac{1}{2}}u^n_{i,j}\right] \\ &+ \frac{k}{h}\left[b^n_{i+\frac{1}{2},j} - b^n_{i-\frac{1}{2},j} + c^n_{i+\frac{1}{2},j} - c^n_{i-\frac{1}{2},j}\right]u^n_{i,j} \\ &+ \frac{k}{h}\left[d^n_{i,j+\frac{1}{2}} - d^n_{i,j-\frac{1}{2}} + f^n_{i,j+\frac{1}{2}} - f^n_{i,j-\frac{1}{2}}\right]u^n_{i,j}. \end{aligned} \tag{3.140}$$

with the initial condition:

$$u_{i,j}^0 = \frac{1}{h^2}\int_{p_{i,j}} u_0(z)dz, \tag{3.141}$$

and the following definitions of the edge velocities:

$$b_{i+\frac{1}{2},j}^n = \frac{1}{h}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} b(x_{i+\frac{1}{2}}, y, t^n)dy, \tag{3.142}$$

$$c_{i+\frac{1}{2},j}^n = \frac{1}{h}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} c(x_{i+\frac{1}{2}}, y, t^n)dy, \tag{3.143}$$

$$d_{i,j+\frac{1}{2}}^n = \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} d(x, y_{j+\frac{1}{2}}, t^n)dx, \tag{3.144}$$

$$f_{i,j+\frac{1}{2}}^n = \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x, y_{j+\frac{1}{2}}, t^n)dx. \tag{3.145}$$

The approximation solution $u_{\mathcal{T},k}$ is defined by:

$$u_{\mathcal{T},k}(x,t) = u_{i,j}^n \ , \quad \forall x \in p_{ij},\ t \in [t_n, t_{n+1}[. \tag{3.146}$$

The BV-norm on $\mathbb{R}^2 \times [0,T[$ can be writen as follows:

$$\begin{aligned} |u_{\mathcal{T},k}|_{BV(\mathbb{R}^2\times[0,T[)} = &\sum_{t_n,t_{n+1}\in[0,T[}\ \sum_{i,j\in\mathbb{Z}} h^2|u_{i,j}^{n+1} - u_{i,j}^n| \\ &+ \sum_{t_n\in[0,T[} k \sum_{i,j\in\mathbb{Z}} \big(h|u_{i+1,j}^n - u_{i,j}^n| + h|u_{i,j+1}^n - u_{i,j}^n|\big). \end{aligned} \tag{3.147}$$

Let Ω be a compact set in $\mathbb{R}^2$ and $T > 0$. The BV-norm of $u_{\mathcal{T},k}$ on $\Omega \times [0,T[$ is given by:

$$\begin{aligned} |u_{\mathcal{T},k}|_{BV(\Omega\times[0,T[)} = &\sum_{t^n,t^{n+1}\in[0,T[}\ \sum_{p_{i,j}\cap\Omega\neq\emptyset} h^2|u_{i,j}^{n+1} - u_{i,j}^n| \\ + \sum_{t_n\in[0,T[} k \sum_{p_{i,j}\cap\Omega\neq\emptyset, p_{i+1,j}\cap\Omega\neq\emptyset} &h|u_{i+1,j}^n - u_{i,j}^n| + \sum_{t_n\in[0,T[} k \sum_{p_{i,j}\cap\Omega\neq\emptyset, p_{i,j+1}\cap\Omega\neq\emptyset} h|u_{i,j+1}^n - u_{i,j}^n|. \end{aligned} \tag{3.148}$$

Shortening the expression of $u_{i,j}^{n+1}$ in (3.140), there holds:

$$\begin{aligned} u_{i,j}^{n+1} = u_{i,j}^n &- \frac{k}{h}\left[b_{i-\frac{1}{2},j}^n u_{i,j}^n - b_{i-\frac{1}{2},j}^n u_{i-1,j}^n + c_{i+\frac{1}{2},j}^n u_{i+1,j}^n - c_{i+\frac{1}{2},j}^n u_{i,j}^n\right] \\ &- \frac{k}{h}\left[d_{i,j-\frac{1}{2}}^n u_{i,j}^n - d_{i,j-\frac{1}{2}}^n u_{i,j-1}^n + f_{i,j+\frac{1}{2}}^n u_{i,j+1}^n - f_{i,j+\frac{1}{2}}^n u_{i,j}^n\right]. \end{aligned} \tag{3.149}$$

Thanks to their definitions in (3.139), all functions $b,\ c,\ d$ and f are bounded by V, and we obtain that the CFL condition

$$k \leq \frac{h}{4V} \tag{3.150}$$

ensures the L^∞ stability of the scheme.

Lemma 3.14. *Assume (3.2), $u_0 \in BV(\mathbb{R}^n)$. Then, there exists $C_{\mathbf{a}}$ depending only on $\mathbf{a}$ such that*

$$\sum_{i,j\in\mathbb{Z}} \big(h|u^n_{i+1,j} - u^n_{i,j}| + h|u^n_{i,j+1} - u^n_{i,j}|\big) \le (1 + C_{\mathbf{a}}k)^n |u_0|_{BV(\mathbb{R}^2)}, \tag{3.151}$$

$$\sum_{i,j\in\mathbb{Z}} h^2|u^{n+1}_{i,j} - u^n_{i,j}| \le 2kV(1 + C_{\mathbf{a}}k)^n |u_0|_{BV(\mathbb{R}^2)}. \tag{3.152}$$

Moreover, the function $u_{\mathcal{T},k}$, defined by (3.146) satisfies:

$$|u_{\mathcal{T},k}|_{BV(\mathbb{R}^2\times[0,T[)} \le (1 + 2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}|u_0|_{BV(\mathbb{R}^2)}. \tag{3.153}$$

Proof. From the BV semi norm of $u_{\mathcal{T},k}$ (3.147), the inequality (3.153) is a straightforward consequence of (3.151) and (3.152).

Next, we shall prove (3.151) and (3.152). We set $BV(u,n) = BV(u,n)_x + BV(u,n)_y$ with:

$$BV(u,n)_x := \sum_{i,j\in\mathbb{Z}} h|u^n_{i+1,j} - u^n_{i,j}| \text{ and } BV(u,n)_y := \sum_{i,j\in\mathbb{Z}} h|u^n_{i,j+1} - u^n_{i,j}|.$$

Then, we have

$$\begin{aligned} BV(u,0)_x &= \sum_{i,j\in\mathbb{Z}} h|u^0_{i+1,j} - u^0_{i,j}| \\ &= \sum_{i,j\in\mathbb{Z}} \frac{1}{h}\left|\int_{p_{i+1,j}} u_0(x,y)dxdy - \int_{p_{i,j}} u_0(x,y)dxdy\right|. \end{aligned}$$

With a change of variable respect to x, there holds

$$\begin{aligned} BV(u,0)_x &\le \sum_{i,j\in\mathbb{Z}} \frac{1}{h}\int_{p_{i,j}} \big|u_0(x+h,y) - u_0(x,y)\big|dxdy \\ &\le \frac{1}{h}\int_{\mathbb{R}^2} \big|u_0(x+h,y) - u_0(x,y)\big|dxdy. \end{aligned} \tag{3.154}$$

By using the same technique with respect to y, we also obtain

$$BV(u,0)_y \le \frac{1}{h}\int_{\mathbb{R}^2} \big|u_0(x,y+h) - u_0(x,y)\big|dxdy. \tag{3.155}$$

Thanks to (3.154) and (3.155), we estimate $BV(u,0)$:

$$\begin{aligned} BV(u,0) \le \frac{1}{h}\Big[&\int_{\mathbb{R}^2} \big|u_0(x+h,y) - u_0(x,y)\big|dxdy \\ &+ \int_{\mathbb{R}^2} \big|u_0(x,y+h) - u_0(x,y)\big|dxdy\Big]. \end{aligned} \tag{3.156}$$

From (3.156) and Lemma 6.9 in [3] (see Lemma 3.12 in this work), there holds:

$$BV(u,0) \leq |u_0|_{BV(\mathbb{R}^2)}. \tag{3.157}$$

In order to prove inequality (3.151), we use the induction method. First, the case when $n = 0$ is given by (3.157). Let us now assume that (3.151) holds for a fixed n, it means that

$$BV(u,n) \leq (1 + C_{\mathbf{a}}k)^n |u_0|_{BV(\mathbb{R}^2)}. \tag{3.158}$$

We shall now estimate $BV(u, n+1)$. Thanks to the defintions of b, c and the CFL condition (3.150), for all $(i,j) \in \mathbb{Z}$ we have

$$\alpha^n_{i-\frac{1}{2},j} := \frac{k}{h} b^n_{i-\frac{1}{2},j} \in [0, \frac{1}{4}],$$

$$\beta^n_{i+\frac{1}{2},j} := \frac{k}{h} c^n_{i+\frac{1}{2},j} \in [-\frac{1}{4}, 0].$$

Then we can rewrite (3.149) as:

$$\begin{aligned} u^{n+1}_{i,j} = u^n_{i,j} &- \alpha^n_{i-\frac{1}{2},j}(u^n_{i,j} - u^n_{i-1,j}) - \beta^n_{i+\frac{1}{2},j}(u^n_{i+1,j} - u^n_{i,j}) \\ &- \frac{k}{h} d^n_{i,j-\frac{1}{2}}(u^n_{i,j} - u^n_{i,j-1}) - \frac{k}{h} f^n_{i,j+\frac{1}{2}}(u^n_{i,j+1} - u^n_{i,j}) \end{aligned} \tag{3.159}$$

and we also have

$$\begin{aligned} u^{n+1}_{i+1,j} = u^n_{i+1,j} &- \alpha^n_{i+\frac{1}{2},j}(u^n_{i+1,j} - u^n_{i,j}) - \beta^n_{i+\frac{3}{2},j}(u^n_{i+2,j} - u^n_{i+1,j}) \\ &- \frac{k}{h} d^n_{i+1,j-\frac{1}{2}}(u^n_{i+1,j} - u^n_{i+1,j-1}) - \frac{k}{h} f^n_{i+1,j+\frac{1}{2}}(u^n_{i+1,j+1} - u^n_{i+1,j}). \end{aligned} \tag{3.160}$$

Let us subtract (3.160) and (3.159) side by side; there holds:

$$\begin{aligned} u^{n+1}_{i+1,j} - u^{n+1}_{i,j} = (1 - \alpha^n_{i+\frac{1}{2},j} + \beta^n_{i+\frac{1}{2},j})(u^n_{i+1,j} - u^n_{i,j}) + \alpha^n_{i-\frac{1}{2},j}(u^n_{i,j} - u^n_{i-1,j}) \\ -\beta^n_{i+\frac{3}{2},j}(u^n_{i+2,j} - u^n_{i+1,j}) + N_d + N_f, \end{aligned} \tag{3.161}$$

where

$$N_d = \frac{k}{h}\Big[d^n_{i,j-\frac{1}{2}}(u^n_{i,j} - u^n_{i,j-1}) - d^n_{i+1,j-\frac{1}{2}}(u^n_{i+1,j} - u^n_{i+1,j-1})\Big], \tag{3.162}$$

$$N_f = \frac{k}{h}\Big[f^n_{i,j+\frac{1}{2}}(u^n_{i,j+1} - u^n_{i,j}) - f^n_{i+1,j+\frac{1}{2}}(u^n_{i+1,j+1} - u^n_{i+1,j})\Big]. \tag{3.163}$$

We set

$$N^*_d = \frac{k}{h}\Big[d^n_{i,j-\frac{1}{2}}(u^n_{i,j} - u^n_{i+1,j}) - d^n_{i,j-\frac{3}{2}}(u^n_{i,j-1} - u^n_{i+1,j-1})\Big], \tag{3.164}$$

$$N^*_f = \frac{k}{h}\Big[f^n_{i,j+\frac{1}{2}}(u^n_{i+1,j} - u^n_{i,j}) - f^n_{i,j+\frac{3}{2}}(u^n_{i+1,j+1} - u^n_{i,j+1})\Big]. \tag{3.165}$$

Let us denote $\gamma^n_{i,j-\frac{1}{2}} := \frac{k}{h} d^n_{i,j-\frac{1}{2}} \in [0, \frac{1}{4}]$ and $\delta^n_{i,j+\frac{1}{2}} := \frac{k}{h} f^n_{i,j+\frac{1}{2}} \in [-\frac{1}{4}, 0]$; then we can write

$$\begin{aligned} N_d^* &= \gamma^n_{i,j-\frac{1}{2}}(u^n_{i,j} - u^n_{i+1,j}) - \gamma^n_{i,j-\frac{3}{2}}(u^n_{i,j-1} - u^n_{i+1,j-1}), \\ N_f^* &= \delta^n_{i,j+\frac{1}{2}}(u^n_{i+1,j} - u^n_{i,j}) - \delta^n_{i,j+\frac{3}{2}}(u^n_{i+1,j+1} - u^n_{i,j+1}). \end{aligned} \tag{3.166}$$

Thanks to (3.161) and (3.166), we have

$$\begin{aligned} u^{n+1}_{i+1,j} - u^{n+1}_{i,j} = (1 &- \alpha^n_{i+\frac{1}{2},j} + \beta^n_{i+\frac{1}{2},j} - \gamma^n_{i,j-\frac{1}{2}} + \delta^n_{i,j+\frac{1}{2}})(u^n_{i+1,j} - u^n_{i,j}) \\ &+ \alpha^n_{i-\frac{1}{2},j}(u^n_{i,j} - u^n_{i-1,j}) - \beta^n_{i+\frac{3}{2},j}(u^n_{i+2,j} - u^n_{i+1,j}) + \gamma^n_{i,j-\frac{3}{2}}(u^n_{i+1,j-1} - u^n_{i,j-1}) \\ &- \delta^n_{i,j+\frac{3}{2}}(u^n_{i+1,j+1} - u^n_{i,j+1}) + N_d - N_d^* + N_f - N_f^*. \end{aligned} \tag{3.167}$$

Now, we will focus on bounding $|N_d - N_d^*|$ and $|N_f - N_f^*|$. We have

$$N_d - N_d^* = \frac{k}{h}\Big[(d^n_{i,j-\frac{1}{2}} - d^n_{i+1,j-\frac{1}{2}})u^n_{i+1,j} + (d^n_{i,j-\frac{3}{2}} - d^n_{i,j-\frac{1}{2}})u^n_{i,j-1} + (d^n_{i+1,j-\frac{1}{2}} - d^n_{i,j-\frac{3}{2}})u^n_{i+1,j-1}\Big].$$

We may express the last term in the right hand-side of the previous equality as:

$$(d^n_{i+1,j-\frac{1}{2}} - d^n_{i,j-\frac{3}{2}})u^n_{i+1,j-1} = (d^n_{i+1,j-\frac{1}{2}} - d^n_{i,j-\frac{1}{2}})u^n_{i+1,j-1} + (d^n_{i,j-\frac{1}{2}} - d^n_{i,j-\frac{3}{2}})u^n_{i+1,j-1}.$$

Then,

$$N_d - N_d^* = \frac{k}{h}\Big[(d^n_{i,j-\frac{1}{2}} - d^n_{i+1,j-\frac{1}{2}})(u^n_{i+1,j} - u^n_{i+1,j-1}) + (d^n_{i,j-\frac{3}{2}} - d^n_{i,j-\frac{1}{2}})(u^n_{i,j-1} - u^n_{i+1,j-1})\Big]. \tag{3.168}$$

Moreover, from the definitions of d and $d^n_{i,j+\frac{1}{2}}$, with a change of variable in x, there holds

$$|d^n_{i,j-\frac{1}{2}} - d^n_{i+1,j-\frac{1}{2}}| \leq \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \big|a_2(x, y_{j-\frac{1}{2}}, t^n) - a_2(x+h, y_{j-\frac{1}{2}}, t^n)\big| dx.$$

Let us now denote $L^x_{a_1}$ and $L^x_{a_2}$ (resp. $L^y_{a_1}$ and $L^y_{a_2}$) the Lipschitz constants of a_1 and a_2 with respect to x (resp. y); we get

$$|d^n_{i,j-\frac{1}{2}} - d^n_{i+1,j-\frac{1}{2}}| \leq L^x_{a_2} h. \tag{3.169}$$

In the same way, we have

$$|d^n_{i,j-\frac{3}{2}} - d^n_{i,j-\frac{1}{2}}| \leq L^y_{a_2} h. \tag{3.170}$$

Combining (3.168), (3.169) and (3.170), there holds:

$$|N_d - N_d^*| \leq k\big[L^x_{a_2}\,|u^n_{i+1,j} - u^n_{i+1,j-1}| + L^y_{a_2}\,|u^n_{i,j-1} - u^n_{i+1,j-1}|\big]. \tag{3.171}$$

To bound $N_f - N_f^*$, we use the same trick as for $N_d - N_d^*$ and obtain

$$|N_f - N_f^*| \le k\big[L_{a_2}^x\,|u_{i+1,j+1}^n - u_{i+1,j}^n| + L_{a_2}^y\,|u_{i+1,j+1}^n - u_{i,j+1}^n|\big]. \tag{3.172}$$

Then, taking the absolute value of (3.167), setting $\mathcal{L} = \max_{w\in\{a_1,a_2\}}\{L_\omega^x, L_\omega^y\}$ and using the fact that $(1 - \alpha_{i+\frac12,j}^n + \beta_{i+\frac12,j}^n - \gamma_{i,j-\frac12}^n + \delta_{i,j+\frac12}^n)$, $\alpha_{i-\frac12,j}^n$, $-\beta_{i+\frac32,j}^n$, $\gamma_{i,j-\frac32}^n$ and $-\delta_{i,j+\frac32}^n$ are all positive, we get:

$$\begin{aligned}
|u_{i+1,j}^{n+1} - u_{i,j}^{n+1}| \le{}& (1 - \alpha_{i+\frac12,j}^n + \beta_{i+\frac12,j}^n - \gamma_{i,j-\frac12}^n + \delta_{i,j+\frac12}^n)|u_{i+1,j}^n - u_{i,j}^n| \\
&+ \alpha_{i-\frac12,j}^n|u_{i,j}^n - u_{i-1,j}^n| - \beta_{i+\frac32,j}^n|u_{i+2,j}^n - u_{i+1,j}^n| \\
&+ \gamma_{i,j-\frac32}^n|u_{i+1,j-1}^n - u_{i,j-1}^n| - \delta_{i,j+\frac32}^n|u_{i+1,j+1}^n - u_{i,j+1}^n| \\
&+ k\mathcal{L}\Big[|u_{i+1,j}^n - u_{i+1,j-1}^n| + |u_{i,j-1}^n - u_{i+1,j-1}^n| \\
&\qquad + |u_{i+1,j+1}^n - u_{i+1,j}^n| + |u_{i+1,j+1}^n - u_{i,j+1}^n|\Big].
\end{aligned} \tag{3.173}$$

Summing (3.173) over i and j and noting that some sums cancel because $\sum \alpha_{i+\frac12,j}^n|u_{i+1,j}^n - u_{i,j}^n| = \sum \alpha_{i-\frac12,j}^n|u_{i,j}^n - u_{i-1,j}^n|$, etc, we obtain

$$BV(u,n+1)_x \le BV(u,n)_x + 2k\mathcal{L}\big[BV(u,n)_x + BV(u,n)_y\big]. \tag{3.174}$$

We obtain a similar bound for $BV(u,n+1)_y$ and combining with (3.158) this leads to

$$BV(u,n+1) \le (1 + 4k\mathcal{L})(1 + kC_{\mathbf{a}})^n|u_0|_{BV(\mathbb{R}^2)},$$

which completes (3.151) if we choose $C_{\mathbf{a}} = 4\mathcal{L}$.

In the next step, we prove (3.152); thanks to (3.149), applying the bound of b, c, d and f, we have

$$\begin{aligned}
|u_{i,j}^{n+1} - u_{i,j}^n| \le \frac{kV}{h}\Big[&|u_{i,j}^n - u_{i-1,j}^n| + |u_{i+1,j}^n - u_{i,j}^n| \\
&+|u_{i,j}^n - u_{i,j-1}| + |u_{i,j+1}^n - u_{i,j}^n|\Big].
\end{aligned}$$

By multiplying this inequality side by side with h^2 and after summing over i and j, we have

$$\sum_{i,j\in\mathbb{Z}} h^2|u_{i,j}^{n+1} - u_{i,j}^n| \le 2kV[BV(u,n)]. \tag{3.175}$$

Using (3.151), we have finished this lemma. □

Lemma 3.15. *Assume (3.2). Then, for all compact set $\Omega \subset \mathbb{R}^2$, for all $T > 0$, there exists $C_{\mathbf{a}}$ depending only on $\mathbf{a}$ and a compact set $\Omega_0^h \subset \mathbb{R}^N$ depending only on Ω, T, $\frac{k}{h}$ and h such that*

$$|u_{\mathcal{T},k}|_{BV(\Omega\times[0,T[)} \le (1 + 2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}|u_0|_{BV(\Omega_0^h)}. \tag{3.176}$$

Proof. We set

$$\Omega_0 = \bigcup_{\mathrm{x}\in\Omega} B_{\max}(\mathrm{x}, \frac{h}{k}T),$$
$$\Omega_0^h = \bigcup_{p_{ij}\cap\Omega_0\neq\emptyset} p_{ij},$$

where $B_{\max}(\mathrm{x}, \frac{h}{k}T) = \{|\mathrm{y}-\mathrm{x}|_{\max} < \frac{h}{k}T\}$ and $|\mathrm{y}-\mathrm{x}|_{\max} = \max\{|y_1-x_1|, |y_2-x_2|\}$, with $\mathrm{x} = (x_1, x_2)$ and $\mathrm{y} = (y_1, y_2)$. Following the proof of Lemma 3.14 without using inequality (3.157), we get

$$|u_{\mathcal{T},k}|_{BV(\Omega\times[0,T[)} \leq (1+2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}|u_{\mathcal{T},0}|_{BV(\Omega_0)}.$$

Noting that

$$|u_{\mathcal{T},0}|_{BV(\Omega_0)} \leq |u_0|_{BV(\Omega_0^h)},$$

then, we obtain

$$|u_{\mathcal{T},k}|_{BV(\Omega\times[0,T[)} \leq (1+2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}|u_0|_{BV(\Omega_0^h)}.$$

We have finished the proof of the lemma. □

Now, we return to Theorem 3.13. Lemma 3.15 generalized to R^N shows that $u_{\mathcal{T},k}$ belongs to $BV(\Omega\times[0,T[)$ for all compact set $\Omega\subset\mathbb{R}^N$ and for all $T>0$. Therefore for all $\varphi\in C_c^\infty(\Omega\times[0,T[,\mathbb{R})$ such that $\|\varphi\|_\infty\leq 1$, from the norm BV of $u_{\mathcal{T},k}$ on $\Omega\times[0,T[$ in (3.148), there holds, thanks to (3.176):

$$\begin{aligned}&\sum_{j=1}^N \int_{\Omega\times[0,T[} u_{\mathcal{T},k}(\mathrm{x},t)\frac{\partial\varphi}{\partial x_j}(\mathrm{x},t)d\mathrm{x}dt + \int_{\Omega\times[0,T[} u_{\mathcal{T},k}(\mathrm{x},t)\frac{\partial\varphi}{\partial t}(\mathrm{x},t)d\mathrm{x}dt\\ &\leq (1+2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}|u_0|_{BV(\Omega_0^h)}.\end{aligned} \tag{3.177}$$

Since $u_{\mathcal{T},k}$ converges toward u in weak $L^\infty(\mathbb{R}^N\times[0,T[)$ (Theorem 3 in [2]) and we can choose the ratio $\frac{k}{h}$ to be a constant, we can pass to the limit in (3.177); then we get the same inequalities with u (resp. $\inf_{h>0}|u_0|_{BV(\Omega_0^h)}$) instead of $u_{\mathcal{T},k}$ (resp. $|u_0|_{BV(\Omega_0^h)}$). This proves that u belongs to $BV_{\mathrm{loc}}(\mathbb{R}^N\times[0,T[)$, for all $T>0$.

Moreover, since $K\subset\mathbb{R}^N\times\mathbb{R}^+$ is a compact set, then, there exist a compact set $\Omega\subset\mathbb{R}^N$ and $T>0$ such that $K\subset\Omega\times[0,T[$ and

$$|u|_{BV(K)} \leq |u|_{BV(\Omega\times[0,T[)} \leq (1+2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}\inf_{h>0}|u_0|_{BV(\Omega_0^h)}.$$

Then we complete this theorem with $C_{K,u_0,\mathbf{a}} = (1+2V)\min\{\frac{1}{C_{\mathbf{a}}}, T\}e^{C_{\mathbf{a}}T}\inf_{h>0}|u_0|_{BV(\Omega_0^h)}$. □

References for the Chapter 3

[1] Chainais–Hillairet C., *Finite volume schemes for a nonlinear equation. Convergence towards the entropy solution and error estimation.* Model. Math. Anal. Numer. 33, 1, 129-156 (1999).

[2] C. Chainais–Hillairet, S. Champier, *Finite Volume for nonhomogeneous Scalar Conservation Laws: Error Estimate.* Numer. Math. vol 88 (4), pp. 607–639 (2001).

[3] Eymard R., Gallouët T., and Herbin R, *Finite volume methods.* In: Ciarlet, Ph.G. and Lions, J.-L., Eds. Handbook of Numerical Analysis, Vol. VII, Amsterdam: North-holland, pp. 173-1020 (2000).

[4] Kröner D., Ohlberger M., *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in Multi Dimensions.* Math. Comput. 69. pp. 25-39 (1999).

[5] Kruskov S.N., *First order quasilinear equations with several space variable.* Math. USSR. Sb. 10. 217-243, 1970.

[6] M. Mamaghani, *Suivi de Fronts par des Méthodes de Raffinement de Maillage Adaptatif et Application à la Simulation du Procédé de Récupération Steam Assisted Gravity Drainage.* Ph.D thesis (2010).

Chapter 4

Discrete Poincaré Inequalities

We establish discrete Poincaré type inequalities on a twodimensional polygonal domain covered by arbitrary, possibly nonconforming meshes. On such meshes, discrete scalar fields are defined by their values both at the cell centers and vertices, while discrete gradients are associated with the edges of the mesh, like in the discrete duality finite volume scheme. We prove that the constants that appear in these inequalities depend only on the domain and on the angles in the diagonals of the diamond cells constructed by joining the two vertices of each mesh edge and the centers of the cells that share that edge.

4.1 Introduction

Let Ω be a two dimensional polygonal domain. Let us introduce the following two Poincaré inequalities which will be mentioned throughout this chapter: The Friedrichs (also called Poincaré) inequality

$$\int_\Omega u^2(x)dx \leq c_F \int_\Omega |\boldsymbol{\nabla} u(x)|^2 dx \quad , \quad \forall u \in H_0^1(\Omega) \tag{4.1}$$

and the Poincaré (also called mean Poincaré) inequality

$$\int_\Omega u^2(x)dx \leq c_P \int_\Omega |\boldsymbol{\nabla} u(x)|^2 dx \quad , \quad \forall u \in H^1(\Omega) \text{ such that } \int_\Omega u(x)dx = 0, \tag{4.2}$$

where c_F and c_P are constants depending only on Ω. These two inequalities play an important role in the theory of partial differential equations. Here, $H^1(\Omega)$ is the Sobolev space of $L^2(\Omega)$ functions with generalized derivatives in $(L^2(\Omega))^2$, and $H_0^1(\Omega)$ is the subspace of $H^1(\Omega)$ with zero boundary values in the sense of traces on $\partial\Omega$. More details on the Sobolev spaces $H^1(\Omega)$, $H_0^1(\Omega)$ may be found, e.g., in [1].

This chapter considers discrete versions of Poincaré inequalities for the so-called discrete duality finite volume (DDFV) method of discretization on arbitrary meshes, as presented, e.g., in [9]. Originally developed for the discretization of (possibly heterogeneous, anisotropic, nonlinear) diffusion equations on arbitrary meshes [2, 5, 9, 13, 14, 18], this technique has found applications in other fields, like electromagnetics [15], div-curl problems [7] and Stokes flows [6, 16, 17], drift diffusion and energy transport models [3].

The originality of these schemes is that they work well on all kind of meshes, including very distorted, degenerating, or highly nonconforming meshes (see the numerical tests in [9]). The name DDFV comes from the fact that these schemes are based on the definition of discrete gradient and divergence operators which verify a discrete Green formula.

Details about this method are recalled in section 4.2. In this introduction, let us only mention that in the DDFV discretization, scalar functions are discretized by their values both at the centers and at the vertices of a given mesh, and their gradients are evaluated on the so-called "diamond-cells" associated to the edges of the mesh. Each internal diamond-cell is a quadrilateral; its vertices are the two nodes of a given internal edge and the centers of the two cells which share this edge. Each boundary diamond cell is a degenerated quadrangle (i.e. a triangle); its vertices are the two nodes of a given boundary edge and the center of the corresponding cell and that of the boundary edge.

Then, the discrete version of the L^2 norm on the left-hand side of (4.1) and (4.2) is the half-sum of the L^2 norms of two piecewise constant functions, one defined with the discrete values given at the centers of the original ("primal" in what follows) cells, and the other defined with the discrete values given at the vertices of the primal mesh, to which we associate cells of a dual mesh. Moreover, the discrete version of the gradient L^2 norm on the right-hand side of (4.1) and (4.2) is the L^2 norm of the piecewise constant gradient vector field defined with it discrete values on the diamond-cells.

In the finite volume context, discrete Poincaré-Friedrich inequalities have previously been proved in [10, Lemma 9.1, Lemma 10.2] and [12], respectively for so-called "admissible" meshes (roughly speaking, meshes such that each edge is orthogonal to the segment joining the centers of the two cells sharing that edge, see the precise definition in [10, Definition 9.1]) and for Voronoi meshes. Similar results on duals of general simplicial triangulations are proved in [19]. In the DDFV context, a discrete version of (4.1) is given for arbitrary meshes in [2]. However, the discrete constant c_F which appears in that paper depends on the mesh regularity in a rather intricate way, see [2, Formula (2.6) and Lemma 3.3].

The main result of our contribution is the proof of discrete versions of both (4.1) and (4.2) in the DDFV context, with constants c_F and c_P depending only on the domain and on the minimum angle in the diagonals of the diamond cells of the mesh.

Our proof of the discrete version of (4.1) is very similar to those given in [10] or [19]. We also prove a discrete version of (4.1) in a slightly more general situation when the domain is not simply connected and the discrete values of the function vanish only on the exterior boundary of the domain and are constant on each of the internal boundaries (this will have a subsequent application in the last section of the present work).

However, the task is more difficult for the mean-Poincaré inequality. Like in [10], it is divided into three steps. The first is the proof of this inequality on a convex subdomain; in the second, our proof differs from that in [10] because we actually do not need to prove a bound on the L^2 norm of the difference of discrete functions and their discrete mean value on the boundary of a convex subset, but rather an easier bound on the L^1 norm of this difference. The final step consists in dividing a general polygonal domain into several convex polygonal subdomains and in combining the first two steps to obtain the result.

As a consequence of these results, we derive a discrete equivalent of the following result (which is a particular case of a result given in [11]): Let us consider open, bounded, simply connected, convex polygonal domains $(\Omega_q)_{q\in[0,Q]}$ of $\mathbb{R}^2$ such that $\Omega_q \subset \Omega_0$ for all $q \in [1, Q]$

and $\bar{\Omega}_{q_1} \cap \bar{\Omega}_{q_2} = \emptyset$ for all $(q_1, q_2) \in [1, Q]^2$ with $q_1 \neq q_2$. Let Ω be defined by $\Omega = \Omega_0 \backslash (\cup_{q=1}^{Q} \Omega_q)$. Let us denote by $\Gamma = \partial\Omega = \cup_{q=0}^{Q} \Gamma_q$, with $\Gamma_q = \partial\Omega_q$ for all $q \in [0, Q]$. Then, there exists a constant C, depending only on Ω, such that for all vector field $\mathbf{u}$ in $H(\mathrm{div}, \Omega) \cap H(\mathrm{rot}, \Omega)$ with $\mathbf{u} \cdot \mathbf{n} = 0$ on Γ and $(\mathbf{u} \cdot \boldsymbol{\tau}, 1)_{\Gamma_q} = 0$ for all $q \in [1, Q]$, there holds

$$||\mathbf{u}||_{L^2(\Omega)} \leq C(||\nabla \cdot \mathbf{u}||_{L^2(\Omega)} + ||\nabla \times \mathbf{u}||_{L^2(\Omega)}). \tag{4.3}$$

The discrete equivalent has applications in the derivation of *a priori* error estimates for the DDFV method applied to the Stokes equations ([8]).

This chapter is organized as follows. Section 4.2 sets some notations and definitions related to the meshes, to discrete differential operators and to discrete functions. In section 4.3, discrete Poincaré inequalities are presented. First, we prove a discrete Poincaré inequality for discrete functions vanishing on the boundary of the polygonal domain and then extend this result to the slightly more general case mentioned above. Then, we prove the discrete mean Poincaré inequality with the 3 steps described above. Finally, we present in section 4.4 an application of the previous results to the derivation of another discrete inequality, relating the norm of discrete vector fields defined on the diamond cells and verifying special boundary conditions, to that of their divergence and curls defined on the primal and dual meshes. In section 4.5, we present the details of the proof of a Lemma involved in our main results.

4.2 Notations and definitions

The following notations are summarized in Fig. 4.1 and Fig. 4.2. Let Ω be defined as above and be covered by a primal mesh with polygonal cells denoted by T_i, $i \in [1, I]$. With each T_i, we associate a point G_i located in the interior of T_i. let us denote by S_k, with $k \in [1, K]$ the nodes of the cells. With any S_k, we associate a dual cell P_k by joining the points G_i associated with the primal cells surrounding S_k to the midpoints of the edges of which S_k is a node.

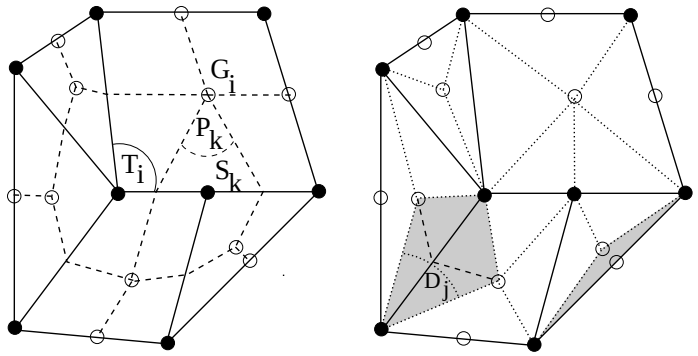

Figure 4.1: A nonconforming primal mesh and its associated dual mesh (left) and diamond-mesh (right).

With any primal edge A_j with $j \in [1, J]$, we associate a so-called diamond-cell D_j obtained by joining the vertices $S_{k_1(j)}$ and $S_{k_2(j)}$ of A_j to the points $G_{i_1(j)}$ and $G_{i_2(j)}$ associated

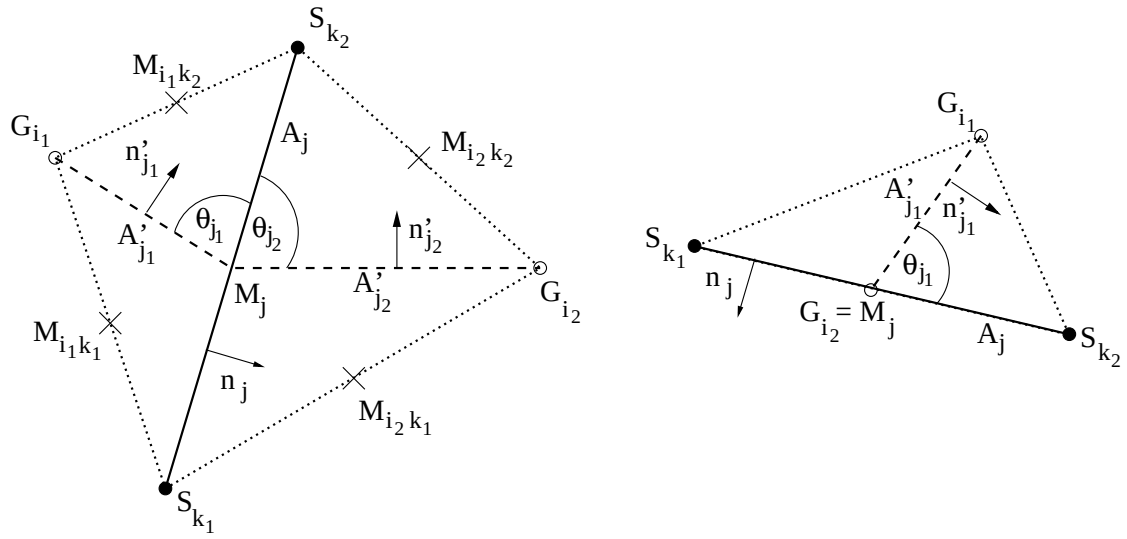

Figure 4.2: Notations for the inner diamond-cell (left) and a boundary diamond mesh (right).

with the primal cells that share A_j as a part of their boundaries. When A_j is a boundary edge (there are J^Γ such edges), the associated diamond-cell is a flat quadrilateral (i.e. a triangle) and we denote by $G_{i_2(j)}$ the midpoint of A_j (thus, there are J^Γ such additional points G_i). The unit normal vector to A_j is $\mathbf{n}_j$ and points from $G_{i_1(j)}$ to $G_{i_2(j)}$. We denote by A'_{j1} (resp. A'_{j2}) the segment joining $G_{i_1(j)}$ (resp. $G_{i_2(j)}$) and the midpoint of A_j. Its associated unit normal vector, pointing from $S_{k_1(j)}$ to $S_{k_2(j)}$, is denoted by $\mathbf{n}'_{j1}$ (resp. $\mathbf{n}'_{j2}$). We also define vectors $\boldsymbol{\tau}_j$, $\boldsymbol{\tau}'_{j1}$ and $\boldsymbol{\tau}'_{j2}$ such that $(\mathbf{n}_j, \boldsymbol{\tau}_j)$, $(\mathbf{n}'_{j1}, \boldsymbol{\tau}'_{j1})$ and $(\mathbf{n}'_{j2}, \boldsymbol{\tau}'_{j2})$ are orthonormal, positively oriented basis of $\mathbb{R}^2$. In the case of a boundary diamond-cell, A'_{j2} reduces to $\{G_{i_2(j)}\}$ and does not play any role. Finally, for any diamond-cell D_j, we shall denote by $M_{i_\alpha k_\beta}$ the midpoint of $[G_{i_\alpha(j)} S_{k_\beta(j)}]$, with $(\alpha, \beta) \in \{1;2\}^2$, M_j the midpoint of $S_{k_1(j)} S_{k_2(j)}$ and θ_{j_1} (resp θ_{j_2}) is defined by the angle, lower than $\pi/2$, between segment $S_{k_1(j)} S_{k_2(j)}$ and segment $G_{i_1(j)} M_j$ (resp $G_{i_2(j)} M_j$).

We shall use the following definition

Definition 4.1. *We denote by $\theta^* > 0$ the greatest angle in the mesh such that*

$$\theta_{j_1} \geq \theta^* \text{ and } \theta_{j_2} \geq \theta^* \qquad \text{for all } j \in [1, J].$$

Now we shall associate discrete scalar values to the points G_i and S_k and discrete twodimensional vector fields to the diamond-cells. This leads us to the following definitions.

Definition 4.2. *Let $\phi = (\phi_i^T, \phi_k^P)$, and $\psi = (\psi_i^T, \psi_k^P)$ be in $\mathbb{R}^I \times \mathbb{R}^K$. Let $\mathbf{u} = (\mathbf{u}_j)$ and*

$\mathbf{w} = (\mathbf{w}_j)$ *be in* $\left(\mathbb{R}^J\right)^2$. *We define the following scalar products and associated norms*

$$(\phi, \psi)_{T,P} := \frac{1}{2}\left(\sum_{i\in[1,I]} |T_i|\phi_i^T\psi_i^T + \sum_{k\in[1,K]} |P_k|\phi_k^P\psi_k^P\right), \tag{4.4}$$

$$\|\phi\|_{T,P}^2 := (\phi,\phi)_{T,P},$$

$$(\mathbf{w},\mathbf{u})_D := \sum_{j\in[1,J]} |D_j|\mathbf{w}_j\cdot\mathbf{u}_j \ \ , \ \ \|\mathbf{u}\|_D^2 := (\mathbf{u},\mathbf{u})_D. \tag{4.5}$$

Definition 4.3. *Let* $\phi = (\phi_i^T, \phi_k^P)$ *be in* $\mathbb{R}^{I+J^\Gamma}\times\mathbb{R}^K$. *We define the trace* $\tilde{\phi}$ *of* ϕ *on the boundary edges* $A_j \subset \Gamma$ *with* $\tilde{\phi}_j := \frac{1}{4}\left(\phi_{k_1(j)}^P + 2\phi_{i_2(j)}^T + \phi_{k_2(j)}^P\right)$

We also define a discrete scalar product for the traces of $\mathbf{u}\cdot\mathbf{n}$ and $\tilde{\phi}$ on the boundaries Γ_q

$$(\mathbf{u}\cdot\mathbf{n},\tilde{\phi})_{\Gamma_q,h} := \sum_{j\in\Gamma_q} |A_j|\,\mathbf{u}_j\cdot\mathbf{n}_j\times\tilde{\phi}_j$$

and on Γ

$$(\mathbf{u}\cdot\mathbf{n},\tilde{\phi})_{\Gamma,h} := \sum_{q\in[0,Q]} (\mathbf{u}\cdot\mathbf{n},\tilde{\phi})_{\Gamma_q,h}\ . \tag{4.6}$$

In the proof of discrete Poincaré inequalities, we often use the piecewise constant functions based on the discrete functions defined at the centers of each mesh; we set the following definitions

Definition 4.4. *Let* $\phi\in\mathbb{R}^{I+J^\Gamma}\times\mathbb{R}^K$. *The piecewise constant functions* $\phi^T(x)$ *and* $\phi^P(x)$ *are defined following, respectively,*

$$\begin{aligned} &\phi^T(x) = \phi_i^T \ \ , \ \ \forall x\in T_i \text{ and } i\in[1,I];\\ &\phi^P(x) = \phi_k^P \ \ , \ \ \forall x\in P_k \text{ and } k\in[1,K]. \end{aligned} \tag{4.7}$$

We recall here the discrete gradient [4, 9] and (vector) curl operators [7] which have been constructed on the diamond cells.

Definition 4.5. *Let* $\phi = (\phi_i^T,\phi_k^P)$ *be in* $\mathbb{R}^{I+J^\Gamma}$, *its discrete gradient* $\boldsymbol{\nabla}_h^D\phi$ *and discrete curl* $\boldsymbol{\nabla}_h^D\times\phi$ *are defined by their values in the cells* D_j *through*

$$(\boldsymbol{\nabla}_h^D\phi)_j := \frac{1}{2|D_j|}\left\{[\phi_{k_2}^P - \phi_{k_1}^P](|A'_{j1}|\mathbf{n}'_{j1} + |A'_{j2}|\mathbf{n}'_{j2}) + [\phi_{i_2}^T - \phi_{i_1}^T]|A_j|\mathbf{n}_j\right\},$$

$$(\boldsymbol{\nabla}_h^D\times\phi)_j := -\frac{1}{2|D_j|}\left\{[\phi_{k_2}^P - \phi_{k_1}^P](|A'_{j1}|\boldsymbol{\tau}'_{j1} + |A'_{j2}|\boldsymbol{\tau}'_{j2}) + [\phi_{i_2}^T - \phi_{i_1}^T]|A_j|\boldsymbol{\tau}_j\right\}.$$

In the proof of our results, we shall use the following theorem which is exactly [7, Theorem 4.7]

Theorem 4.6. *Let $(\mathbf{u}_j)_{j\in[1,J]}$ be a discrete vector field defined by its values on the diamond-cells D_j. There exist unique $\phi = (\phi_i^T, \phi_k^P)_{i\in[1,I+J^\Gamma],k\in[1,K]}$, $\psi = (\psi_i^T, \psi_k^P)_{i\in[1,I+J^\Gamma],k\in[1,K]}$ and $(c_q^T, c_q^P)_{q\in[1,Q]}$ such that:*

$$\mathbf{u}_j = (\boldsymbol{\nabla}_h^D \phi)_j + (\boldsymbol{\nabla}_h^D \times \psi)_j, \quad \forall j \in [1, J] \,, \tag{4.8}$$

$$\sum_{i\in[1,I]} |T_i| \phi_i^T = \sum_{k\in[1,K]} |P_k| \phi_k^P = 0 \,, \tag{4.9}$$

$$\psi_i^T = 0 \,,\ \forall i \in \Gamma_0 \quad , \quad \psi_k^P = 0 \,,\ \forall k \in \Gamma_0 \,, \tag{4.10}$$

and

$$\forall q \in [1, Q] \,, \qquad \psi_i^T = c_q^T \,,\ \forall i \in \Gamma_q \quad , \quad \psi_k^P = c_q^P \,,\ \forall k \in \Gamma_q \,. \tag{4.11}$$

Moreover, the decomposition (4.8) is orthogonal.

We shall also need the following construction of discrete divergence and (scalar) curl operators on both primal and dual cells:

Definition 4.7. *Let $\mathbf{u} = (\mathbf{u}_j)$ be defined in $(\mathbb{R}^2)^J$ by its values on the diamond-cells. We define*

$$\left(\nabla_h^T \cdot \mathbf{u}\right)_i := \frac{1}{|T_i|} \sum_{j\in\partial T_i} |A_j| \mathbf{u}_j \cdot \mathbf{n}_{ji}, \tag{4.12}$$

$$\begin{aligned} \left(\nabla_h^P \cdot \mathbf{u}\right)_k := \frac{1}{|P_k|} \Bigg(& \sum_{j\in\partial P_k} \left(|A'_{j1}| \mathbf{u}_j \cdot \mathbf{n}'_{j1k} + |A'_{j2}| \mathbf{u}_j \cdot \mathbf{n}'_{j2k} \right) \\ & + \sum_{j\in\partial P_k\cap\Gamma} \frac{|A_j|}{2} \mathbf{u}_j \cdot \mathbf{n}_j \Bigg), \end{aligned} \tag{4.13}$$

$$\left(\nabla_h^T \times \mathbf{u}\right)_i := \frac{1}{|T_i|} \sum_{j\in\partial T_i} |A_j| \mathbf{u}_j \cdot \boldsymbol{\tau}_{ji}, \tag{4.14}$$

$$\begin{aligned} \left(\nabla_h^P \times \mathbf{u}\right)_k := \frac{1}{|P_k|} \Bigg(& \sum_{j\in\partial P_k} \left(|A'_{j1}| \mathbf{u}_j \cdot \boldsymbol{\tau}'_{j1k} + |A'_{j2}| \mathbf{u}_j \cdot \boldsymbol{\tau}'_{j2k} \right) \\ & + \sum_{j\in\partial P_k\cap\Gamma} \frac{|A_j|}{2} \mathbf{u}_j \cdot \boldsymbol{\tau}_j \Bigg). \end{aligned} \tag{4.15}$$

The following result [7, Proposition 4.1], which consists in discrete Green formulas, has motivated the name "discrete duality":

Theorem 4.8. *For $\mathbf{u} \in (\mathbb{R}^2)^J$ and $\phi = (\phi^T, \phi^P) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$, it holds that*

$$(\mathbf{u}, \boldsymbol{\nabla}_h^D \phi)_D = -(\nabla_h^{T,P} \cdot \mathbf{u}, \phi)_{T,P} + (\mathbf{u} \cdot \mathbf{n}, \tilde{\phi})_{\Gamma,h}, \tag{4.16}$$

$$(\mathbf{u}, \boldsymbol{\nabla}_h^D \times \phi)_D = (\nabla_h^{T,P} \times \mathbf{u}, \phi)_{T,P} - (\mathbf{u} \cdot \boldsymbol{\tau}, \tilde{\phi})_{\Gamma,h}. \tag{4.17}$$

4.3 Discrete Poincaré inequalities

We first start with a discrete version of (4.1). Our result is a special case of that proved in [2, Lemma 3.3], but our expression of the discrete constant c_F is more precise and simple, in that its dependence on the geometry of the cells occurs only through the angles in the diagonals of the diamond-cells. This is an important result in the DDFV context, since *a priori* error estimations of the discrete solution of the Laplace equation obtained with this method also only depend on the cell geometries through angles in the diamond-cells (see [9]).

Theorem 4.9 (Discrete Poincaré-Friedrichs Inequality). *Let Ω be an open bounded polygonal domain; let us consider $u = (u_i^T, u_k^P) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$ such that*

$$u_k^P = 0,\ \forall k \in \Gamma \quad \text{and} \quad u_i^T = 0,\ \forall i \in \Gamma. \tag{4.18}$$

Let θ^* be defined by Definition 4.1. Then, there exists a constant C only depending on Ω and θ^* such that

$$\|u\|_{T,P} \leq C \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.19}$$

Proof. Let $u^T(\cdot)$ and $u^P(\cdot)$ be the piecewise constant functions defined in Definition 4.4. Then obviously $\|u\|_{T,P}^2 = (\|u^T\|_{L^2(\Omega)}^2 + \|u^P\|_{L^2(\Omega)}^2)/2$, so that, in order to prove (4.19), it suffices to prove

$$\|u^T\|_{L^2(\Omega)} \leq C \|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.20}$$

$$\|u^P\|_{L^2(\Omega)} \leq C \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.21}$$

We shall first prove (4.20). Let $\mathbf{d}_1 = (0,1)^t$ and $\mathbf{d}_2 = (1,0)^t$; for $x \in \Omega$, let $\mathcal{D}_x^1$ and $\mathcal{D}_x^2$ be the straight lines going through x and parallel to the vectors $\mathbf{d}_1$ and $\mathbf{d}_2$. For any edge $j \in [1, J]$ and any $x \in \Omega$, let us define $\chi_j^{T,1}(x)$ and $\chi_j^{T,2}(x)$ by

$$\chi_j^{T,\ell}(x) = \begin{cases} 1 & \text{if } A_j \cap \mathcal{D}_x^\ell \neq \emptyset \\ 0 & \text{if } A_j \cap \mathcal{D}_x^\ell = \emptyset \end{cases} \quad \text{for } \ell = 1, 2. \tag{4.22}$$

Remark 4.10. *For any $x = (x_1, x_2) \in \Omega$, we note that $\chi_j^{T,1}(x)$ only depends on x_1 and $\chi_j^{T,2}(x)$ only depends on x_2.*

From the first formula of definition 4.5 and simple geometry, it is easy to see that

$$(\boldsymbol{\nabla}_h^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} = u_{i_2(j)}^T - u_{i_1(j)}^T, \quad \forall j \in [1, J]. \tag{4.23}$$

Then, for any $i \in [1, I]$ and a.e. $x \in T_i$, let us follow the straight line $\mathcal{D}_x^\ell$ until it intersects the boundary Γ, and let us denote by $v_1(i) := i, v_2(i), \cdots, v_{n-1}(i)$ the indices of the primal cells that it intersects (in the order they are intersected) and by $v_n(i)$ the index in $[I+1, I+J^\Gamma]$ corresponding to the first boundary segment intersected by $\mathcal{D}_x^\ell$ (see Fig. 4.3). Then, since $u_{v_n(i)}^T = 0$ because of the boundary conditions, we may write

$$u_i^T = u_{v_1(i)}^T = (u_{v_1(i)}^T - u_{v_2(i)}^T) + (u_{v_2(i)}^T - u_{v_3(i)}^T) + \cdots + (u_{v_{n-1}(i)}^T - u_{v_n(i)}^T) = \sum_{m=1}^{n-1} (u_{v_m(i)}^T - u_{v_{m+1}(i)}^T),$$

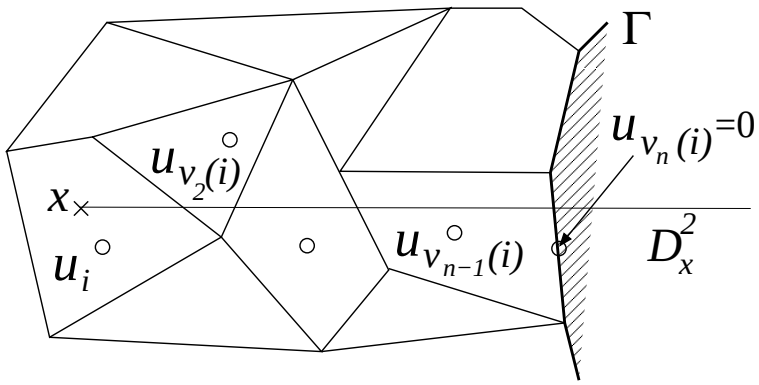

Figure 4.3: Straight line $\mathcal{D}_x^2$ intersecting primal cells from point x to the boundary.

so that, since any couple $(u^T_{v_m(i)}, u^T_{v_{m+1}(i)})$ is a pair of neighboring values through an edge A_j intersected by $\mathcal{D}_x^\ell$, there holds, thanks to (4.23)

$$|u^T(x)| = |u_i^T| \le \sum_{j=1}^{J} \left| (\boldsymbol{\nabla}_j^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right| \chi_j^{T,\ell}(x)$$

for $\ell = 1, 2$. Then, setting $v_j := \left| (\boldsymbol{\nabla}_j^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right|$, one has

$$(u^T(x))^2 \le \left(\sum_{j=1}^{J} v_j \, \chi_j^{T,1}(x) \right) \left(\sum_{j=1}^{J} v_j \, \chi_j^{T,2}(x) \right). \tag{4.24}$$

Integrating the above inequality over T_i and summing over $i \in [1, I]$ yields

$$\|u^T\|^2_{L^2(\Omega)} \le \int_\Omega \left[\left(\sum_{j=1}^{J} v_j \, \chi_j^{T,1}(x) \right) \left(\sum_{j=1}^{J} v_j \, \chi_j^{T,2}(x) \right) \right] dx. \tag{4.25}$$

Let $\alpha = \inf\{x_1;\ (x_1, x_2) \in \Omega\}$ and $\beta = \sup\{x_1;\ (x_1, x_2) \in \Omega\}$. For each $x_1 \in (\alpha, \beta)$, we denote by $H(x_1)$ the set of x_2 such that $x = (x_1, x_2) \in \Omega$. From Remark 4.10 and the fact that $\int_{H(x_1)} \chi_j^{T,2}(x_2) dx_2 \le |A_j|$ and $\int_\alpha^\beta \chi_j^{T,1}(x_1) dx_1 \le |A_j|$, we infer that (4.25) may be written

in the following way:

$$\begin{aligned}
\|u^T\|^2_{L^2(\Omega)} &\leq \int_\alpha^\beta dx_1 \int_{H(x_1)} dx_2 \left[\sum_{j=1}^J v_j\ \chi_j^{T,1}(x_1) \sum_{j=1}^J v_j\ \chi_j^{T,2}(x_2)\right] \\
&\leq \int_\alpha^\beta \sum_{j=1}^J v_j\ \chi_j^{T,1}(x_1) \left(\int_{H(x_1)} \sum_{j=1}^J v_j\ \chi_j^{T,2}(x_2) dx_2\right) dx_1 \\
&\leq \int_\alpha^\beta \sum_{j=1}^J v_j\ \chi_j^{T,1}(x_1) \left(\sum_{j=1}^J v_j \int_{H(x_1)} \chi_j^{T,2}(x_2) dx_2\right) dx_1 \\
&\leq \int_\alpha^\beta \sum_{j=1}^J v_j\ \chi_j^{T,1}(x_1) \left(\sum_{j=1}^J v_j |A_j|\right) dx_1 \\
&\leq \left(\sum_{j=1}^J v_j |A_j|\right) \sum_{j=1}^J v_j \int_\alpha^\beta \chi_j^{T,1}(x_1) dx_1 \leq \left(\sum_{j=1}^J v_j |A_j|\right) \left(\sum_{j=1}^J v_j |A_j|\right).
\end{aligned}$$

We thus obtain

$$\|u^T\|^2_{L^2(\Omega)} \leq \left(\sum_{j=1}^J |(\boldsymbol{\nabla}_h^D u)_j . \overrightarrow{G_{i_1(j)} G_{i_2(j)}}| |A_j|\right)^2. \tag{4.26}$$

Finally, Using the Cauchy-Schwarz inequality, we have

$$\|u^T\|^2_{L^2(\Omega)} \leq \left(\sum_{j=1}^J |(\boldsymbol{\nabla}_h^D u)_j|^2 |G_{i_1(j)} G_{i_2(j)}| |A_j|\right) \left(\sum_{j=1}^J |G_{i_1(j)} G_{i_2(j)}| |A_j|\right). \tag{4.27}$$

Since $|D_j| = \frac{1}{2}\left(|A_j||G_{i_1} M_j| \sin\theta_{j_1} + |A_j||G_{i_2} M_j| \sin\theta_{j_2}\right)$, we have that $|A_j||G_{i_1} G_{i_2}| \leq \frac{2|D_j|}{\sin\theta^*}$ by Definition 4.1 and the triangle inequality. Moreover, since $\sum_{j=1}^J |D_j| = |\Omega|$, there holds

$$\|u^T\|^2_{L^2(\Omega)} \leq \frac{4}{\sin^2\theta^*} |\Omega| \sum_{j=1}^J |(\boldsymbol{\nabla}_h^D u)_j|^2 |D_j|.$$

We have completed inequality (4.20) with $C = \frac{2}{\sin\theta^*} |\Omega|^{1/2}$. We now turn to inequality (4.21). We shall use a very similar process to that employed in the proof of (4.20). A slight difference comes from the fact that dual cells may be non-convex, and that the straight lines $\mathcal{D}_x^\ell$ may thus intersect twice the boundary $A'_{j1} \cup A'_{j2}$ between two adjacent dual cells (see Fig. 4.4), in which case it is not useful to introduce the difference $u^P_{k_2(j)} - u^P_{k_1(j)}$ in the calculation. We thus define $\chi_j^{P,1}(x)$ and $\chi_j^{P,2}(x)$ by

$$\chi_j^{P,\ell}(x) = \begin{cases} 1 & \text{if only } A'_{j1} \cap \mathcal{D}_x^\ell \neq \emptyset \text{ or } A'_{j2} \cap \mathcal{D}_x^\ell \neq \emptyset \\ 0 & \text{if } \left(A'_{j1} \cup A'_{j2}\right) \cap \mathcal{D}_x^\ell = \emptyset \end{cases} \qquad \text{for } \ell = 1, 2.$$

In the above definition, it is meant that if $\mathcal{D}_x^\ell$ intersects both A'_{j1} and A'_{j2}, then $\chi_j^{P,\ell}(x) = 0$. From the first formula of definition 4.5, it is easy to see that

$$(\boldsymbol{\nabla}_j^D u)_j \cdot \overrightarrow{S_{k_1(j)} S_{k_2(j)}} = u^P_{k_2(j)} - u^P_{k_1(j)}, \quad \forall j \in [1, J].$$

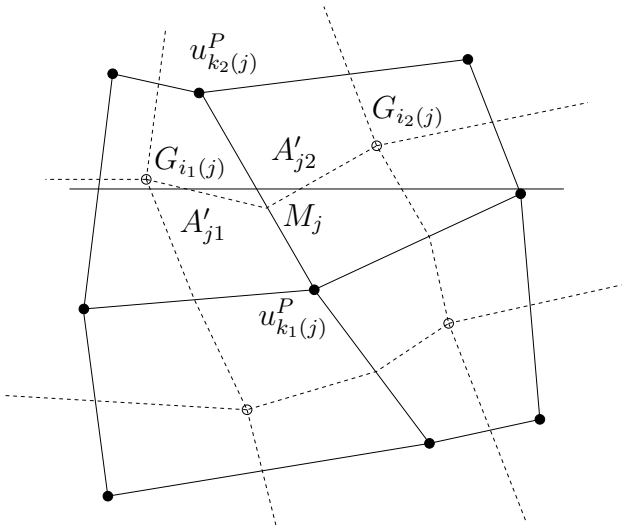

Figure 4.4: The straight line $\mathcal{D}^2_x$ intersects twice the boundary $A'_{j1} \cup A'_{j2}$ of a non convex dual.

Thus, for any $k \in [1, K]$ and a.e. $x \in P_k$, one has

$$|u^P_k| \leq \sum_{j=1}^{J} |(\boldsymbol{\nabla}^D_h u)_j \cdot \overrightarrow{S_{k_1(j)} S_{k_2(j)}}| \; \chi^{P,\ell}_j(x) \;\; , \;\; \ell = 1, 2.$$

Using a similar process as in the proof of (4.20) and taking into account that

$$\int_\alpha^\beta \chi^{P,1}_j(x_1) dx_1 \leq |A'_{j_1}| + |A'_{j_2}| \text{ and } \int_{H(x_1)} \chi^{P,2}_j(x_2) dx_2 \leq |A'_{j_1}| + |A'_{j_2}|,$$

we obtain

$$\|u^P\|^2_{L^2(\Omega)} \leq \left(\sum_{j=1}^{J} |(\boldsymbol{\nabla}^D_h u)_j| |A_j| (|A'_{j_1}| + |A'_{j_2}|) \right)^2$$

which allows to obtain, similarly as above

$$\|u^P\|^2_{L^2(\Omega)} \leq \frac{4}{\sin^2 \theta^*} |\Omega| \sum_{j=1}^{J} |(\boldsymbol{\nabla}^D_h u)_j|^2 |D_j|,$$

which concludes the proof of inequality (4.21) with $C = \frac{2}{\sin \theta^*} |\Omega|^{1/2}$.

□

We now turn to a generalization of Theorem 4.9 which will be useful in the last section of this work.

Theorem 4.11 (Discrete Poincaré-Friedrichs Inequality). *Let us consider open, bounded, simply connected, convex polygonal domains* $(\Omega_q)_{q \in [0,Q]}$ *of* $\mathbb{R}^2$ *such that* $\Omega_q \subset \Omega_0$ *for all*

$q \in [1,Q]$ and $\bar{\Omega}_{q_1} \cap \bar{\Omega}_{q_2} = \emptyset$ for all $(q_1,q_2) \in [1,Q]^2$ with $q_1 \neq q_2$. Let Ω be defined by $\Omega = \Omega_0 \backslash (\cup_{q=1}^{Q} \Omega_q)$. Let us denote by $\Gamma = \partial\Omega = \cup_{q=0}^{Q} \Gamma_q$, with $\Gamma_q = \partial\Omega_q$ for all $q \in [0,Q]$.

Let $u = (u^T, u^P) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$ be such that

$$\begin{aligned} u_k^P = 0,\ \forall k \in \Gamma_0 \quad \text{and} \quad u_i^T = 0,\ \forall i \in \Gamma_0, \\ u_k^P = c_q^P,\ \forall k \in \Gamma_q, \quad \text{and} \quad u_i^T = c_q^T,\ \forall i \in \Gamma_q, \forall q \in [1,Q]. \end{aligned} \tag{4.28}$$

With θ^* given by Definition 4.1, there exists a constant C depending only on Ω and θ^* such that (4.19) holds.

Proof. Like in Theorem 4.9, it suffices to prove both (4.20) and (4.21). We shall only prove (4.20), since the proof of (4.21) follows exactly the same lines.

The only difference in the proof of (4.20) in Theorem 4.11 with respect to Theorem 4.9 is that the straight line $\mathcal{D}_x^\ell$ may now intersect one or several internal boundary(ies) Γ_q, with $q \in [1,Q]$, before intersecting the external boundary Γ_0 (see Fig. 4.5). For the sake of simplicity, we shall consider only one intersection with an internal boundary Γ_q (since the alternative may be treated exactly in the same way), and we denote by $v_{n_q}(i)$ and $v_{n_q+1}(i)$ the indices in $[I+1, I+J^\Gamma]$ corresponding to those intersected boundary edges of Γ_q. We may still write

$$u_i^T = \sum_{m=1}^{n-1} (u_{v_m(i)}^T - u_{v_{m+1}(i)}^T),$$

but, now, the couple $(u_{v_{n_q}(i)}^T, u_{v_{n_q+1}(i)}^T)$ *is not* a pair of neighboring values through an edge A_j intersected by $\mathcal{D}_x^\ell$. However, these two values are equal because of (4.28), so that

$$u_i^T = \sum_{\substack{m \in [1,n-1] \\ m \neq n_q}} (u_{v_m(i)}^T - u_{v_{m+1}(i)}^T).$$

Now, any couple $(u_{v_m(i)}^T, u_{v_{m+1}(i)}^T)$ in the above sum *is* a pair of neighboring values through an edge A_j of the mesh, intersected by $\mathcal{D}_x^\ell$, so that there holds, thanks to (4.23)

$$|u_i^T| \leq \sum_{j=1}^{J} \left| (\boldsymbol{\nabla}_j^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right| \chi_j^{T,\ell}(x)$$

for $\ell = 1,2$ and we finish the proof just like in the proof of (4.20). □

Let us now turn to a discrete version of (4.2). As announced in the Introduction, the proof will be divided in three steps. The first step is to prove it in the case of a convex polygonal domain (Theorem 4.12), then we shall prove an inequality related to the mean value on the boundary of a convex polygonal domain (Theorem 4.15) and we shall conclude by the general case of a possibly non-convex polygonal domain (Theorem 4.17).

Theorem 4.12 (Discrete mean Poincaré Inequality for a convex polygonal domain). *Let Ω be an open bounded polygonal connected domain, and ω be an open convex*

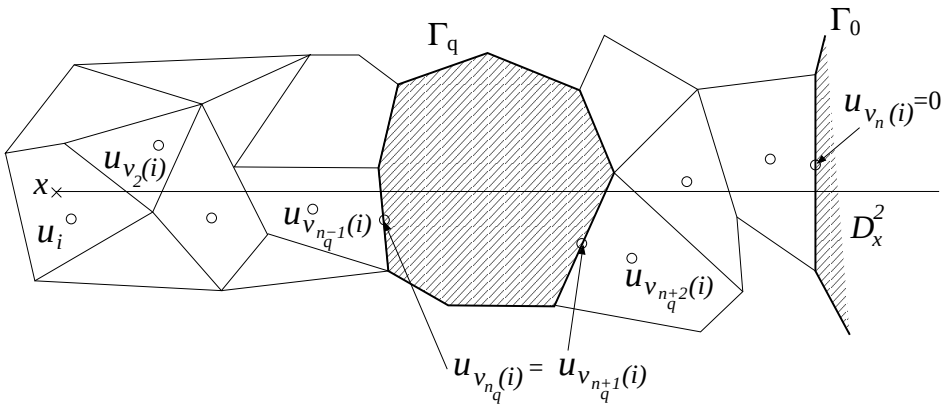

Figure 4.5: Straight line $\mathcal{D}_x^2$ intersecting primal cells from point x to the boundary through internal boundary Γ_q.

polygonal subset of Ω, with $\omega \neq \emptyset$. Let $u = (u_i^T, u_k^P) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$; the associated piecewise constant functions u^T, u^P are defined through Definition 4.4. Let θ^* be defined through Definition 4.1. Let us define the following mean-values:

$$m_\omega^T(u) := \frac{1}{|\omega|}\int_\omega u^T(x)\,dx \quad , \quad m_\omega^P(u) := \frac{1}{|\omega|}\int_\omega u^P(x)\,dx.$$

Then, there exists a constant C only depending on Ω and θ^* such that

$$\|u^T - m_\omega^T(u)\|_{L^2(\omega)} \leq C\|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.29}$$

and

$$\|u^P - m_\omega^P(u)\|_{L^2(\omega)} \leq C\|\boldsymbol{\nabla}_h^D u\|_D, . \tag{4.30}$$

(Choosing $\omega = \Omega$ proves the discrete equivalent of (4.2) if Ω is convex.)

Proof. We only prove inequality (4.29). The proof of (4.30) may be adapted just like in the proof of Theorem 4.9. We first note that

$$\begin{aligned}\int_\omega |u^T(x) - m_\omega^T(u)|^2 dx &= \int_\omega \left|u^T(x) - \frac{1}{|\omega|}\int_\omega u^T(y)dy\right|^2 dx \\ &\leq \frac{1}{|\omega|}\int_\omega\int_\omega |u^T(x) - u^T(y)|^2 dydx.\end{aligned} \tag{4.31}$$

We define points A, B, C, D belonging to $\overline{\omega}$ in the following way

$$\begin{aligned}x_A &= \inf\{x_1; (x_1,x_2) \in \omega\}, \quad x_C = \sup\{x_1; (x_1,x_2) \in \omega\}, \\ y_B &= \inf\{y_2; (y_1,y_2) \in \omega\}, \quad y_D = \sup\{y_2; (y_1,y_2) \in \omega\}.\end{aligned}$$

Remark 4.13. *Up to a rotation of ω, we may always suppose that those four points are different one from the other, except if ω is triangular; in that case, up to a rotation of ω, we may set $A = B$ and the proof is exactly the same as that below.*

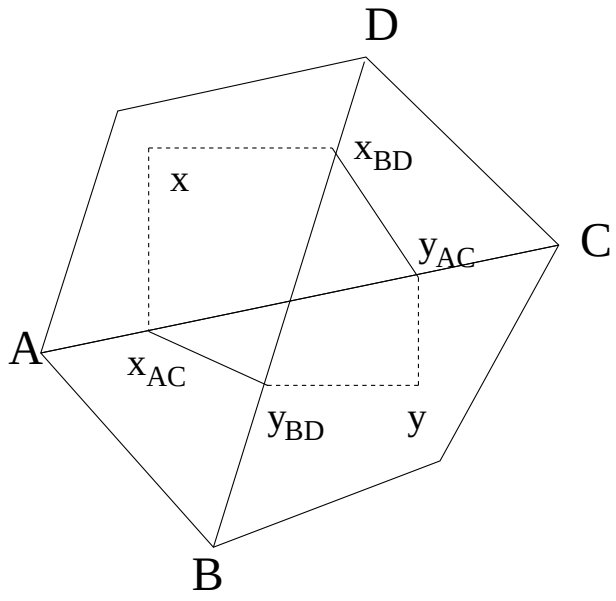

Figure 4.6: Notation for points A, B, C, D and points x_{AC}, x_{BD}, y_{AC}, y_{BD}.

For any $x = (x_1, x_2) \in \omega$, we define $x_{AC} \in [AC]$ such that $(x_{AC})_1 = x_1$ and $x_{BD} \in [BD]$ such that $(x_{BD})_2 = x_2$. The notations are summarized in Fig 4.6.
Applying the triangle inequality leads to

$$\begin{aligned} |u^T(x) - u^T(y)| \leq |u^T(x) - u^T(x_{BD})| + |u^T(x_{BD}) - u^T(y_{AC})| \\ + |u^T(y_{AC}) - u^T(y)|, \end{aligned} \tag{4.32}$$

and also to

$$\begin{aligned} |u^T(x) - u^T(y)| \leq |u^T(x) - u^T(x_{AC})| + |u^T(x_{AC}) - u^T(y_{BD})| \\ + |u^T(y_{BD}) - u^T(y)|. \end{aligned} \tag{4.33}$$

From (4.32) and (4.33), we have

$$\int_\omega \int_\omega |u^T(x) - u^T(y)|^2 dxdy \leq \sum_{i=1}^{9} I_i \tag{4.34}$$

where I_1–I_9 are defined and estimated in what follows:

1.

$$I_1 = \int_\omega \int_\omega |u^T(x) - u^T(x_{BD})| \, |u^T(x) - u^T(x_{AC})| \, dxdy. \tag{4.35}$$

Using again (4.22) and (4.23), we may write

$$|u^T(x) - u^T(x_{AC})| \leq \sum_{j=1}^{J} \chi_j^{T,1}(x) \left| (\boldsymbol{\nabla}_h^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right|. \tag{4.36}$$

and

$$|u^T(x) - u^T(x_{BD})| \leq \sum_{j=1}^{J} \chi_j^{T,2}(x) \left| (\boldsymbol{\nabla}_h^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right|. \tag{4.37}$$

Henceforth, we set for convenience $v_j = \left| (\boldsymbol{\nabla}_h^D u)_j \cdot \overrightarrow{G_{i_1(j)} G_{i_2(j)}} \right|$. Recalling that $\chi_j^{T,1}(x)$ only depends on x_1 and $\chi_j^{T,2}(x)$ only depends on x_2, and noting that the integrand in (4.35) does not depend on y, there holds

$$\begin{aligned} I_1 &\leq |\omega| \left(\int_{x_C}^{x_A} \sum_{j=1}^{J} \chi_j^{T,1}(x) v_j dx_1 \right) \left(\int_{y_B}^{y_D} \sum_{j=1}^{J} \chi_j^{T,2}(x) v_j dx_2 \right) \\ &\leq |\omega| \left(\sum_{j=1}^{J} v_j \int_{x_C}^{x_A} \chi_j^{T,1}(x) dx_1 \right) \left(\sum_{j=1}^{J} v_j \int_{y_B}^{y_D} \chi_j^{T,2}(x) dx_2 \right). \end{aligned}$$

We use that

$$\int_{x_A}^{x_C} \chi_j^{T,1}(x) dx_1 \leq |A_j| \tag{4.38}$$

and

$$\int_{y_B}^{y_D} \chi_j^{T,2}(x) dx_2 \leq |A_j| \tag{4.39}$$

and obtain

$$I_1 \leq |\omega| \left(\sum_{j=1}^{J} |A_j| v_j \right)^2. \tag{4.40}$$

2.

$$I_2 = \int_\omega \int_\omega |u^T(x) - u^T(x_{BD})| \, |u^T(x_{AC}) - u^T(y_{BD})| \, dxdy.$$

Using inequality (4.37), we have

$$I_2 \leq \int_\omega \int_\omega \left(\sum_{j=1}^{J} \chi_j^2(x) \, v_j \right) |u^T(x_{AC}) - u^T(y_{BD})| \, dxdy.$$

By definition, $\chi_j^2(x)$ only depends on x_2 (which is in $[y_B, y_D]$), while x_{AC} only depends on x_1 (which is in $[x_A, x_C]$); of course, y_{BD} does not depend on x, so that

$$I_2 \leq \left(\sum_{j=1}^{J} v_j \int_{y_B}^{y_D} \chi_j^{T,2}(x) dx_2 \right) \int_\omega \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| \, dx_1 dy.$$

Thanks to (4.39), we thus have

$$I_2 \leq \left(\sum_{j=1}^{J} |A_j| v_j \right) \int_\omega \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| \, dx_1 dy.$$

Since y_{BD} only depends on y_2 and x_{AC} does not depend on y, the integration with respect to y_1 (which is in $[x_A, x_C]$) is straightforward and yields

$$I_2 \leq (x_C - x_A) \left(\sum_{j=1}^{J} |A_j| v_j \right) \int_{y_B}^{y_D} \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| \, dx_1 dy_2. \tag{4.41}$$

3.

$$I_3 = \int_\omega \int_\omega |u^T(x) - u^T(x_{BD})| \, |u^T(y_{BD}) - u^T(y)| \, dxdy.$$

This integral clearly decouples into two independent integrals

$$I_3 = \int_\omega |u^T(x) - u^T(x_{BD})| \, dx \int_\omega |u^T(y_{BD}) - u^T(y)| \, dy$$

which may be treated like in the estimation of I_1 thanks to (4.37), (4.39) and the fact that $\chi^{T,2}$ depends only on x_2. We obtain

$$I_3 = (x_C - x_A)^2 \left(\sum_{j=1}^{J} |A_j| v_j \right)^2. \tag{4.42}$$

4.

$$I_4 = \int_\omega \int_\omega |u^T(x_{BD}) - u^T(y_{AC})| \, |u^T(x) - u^T(x_{AC})| \, dxdy.$$

We may proceed very similarly to the estimation of I_2 and we obtain that

$$I_4 \leq (y_D - y_B) \left(\sum_{j=1}^{J} |A_j| v_j \right) \int_{x_A}^{x_C} \int_{y_B}^{y_D} |u^T(x_{BD}) - u^T(y_{AC})| \, dx_2 dy_1. \tag{4.43}$$

5.

$$I_5 = \int_\omega \int_\omega |u^T(x_{BD}) - u^T(y_{AC})| \, |u^T(x_{AC}) - u^T(y_{BD})| \, dxdy.$$

On the one hand, x_{BD} and y_{AC} do not depend on x_1; on the other hand, x_{AC} and y_{BD} do not depend on x_2, so that the integration with respect to x decouples into

$$I_5 \leq \int_\omega \left(\int_{y_B}^{y_D} |u^T(x_{BD}) - u^T(y_{AC})| \, dx_2 \right) \left(\int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| \, dx_1 \right) dy.$$

We also note that y_{BD} and x_{AC} do not depend on y_1 and that y_{AC} and x_{BD} do not depend on y_2, so that the integration with respect to y decouples into

$$I_5 \leq \int_{x_A}^{x_C} \int_{y_B}^{y_D} |u^T(x_{BD}) - u^T(y_{AC})| \, dx_2 dy_1 \int_{y_B}^{y_D} \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| \, dx_1 dy_2. \tag{4.44}$$

6.

$$I_6 = \int_\omega \int_\omega |u^T(x_{BD}) - u^T(y_{AC})|\, |u^T(y_{BD}) - u^T(y)|\, dxdy.$$

We may proceed very similarly to the estimations of I_2 and I_4 and we obtain that

$$I_6 \leq (x_C - x_A) \left(\sum_{j=1}^{J} |A_j| v_j \right) \int_{x_A}^{x_C} \int_{y_B}^{y_D} |u^T(x_{BD}) - u^T(y_{AC})|\, dx_2 dy_1. \tag{4.45}$$

7.

$$I_7 = \int_\omega \int_\omega |u^T(y_{AC}) - u^T(y)|\, |u^T(x) - u^T(x_{AC})|\, dxdy.$$

We may proceed very similarly to the estimation of I_3 and we obtain that

$$I_7 \leq (y_D - y_B)^2 \left(\sum_{j=1}^{J} |A_j| v_j \right)^2. \tag{4.46}$$

8.

$$I_8 = \int_\omega \int_\omega |u^T(y_{AC}) - u^T(y)|\, |u^T(x_{AC}) - u^T(y_{BD})|\, dxdy.$$

We may proceed very similarly to the estimations of I_2, I_4 and I_6 and we obtain that

$$I_8 \leq (y_D - y_B) \left(\sum_{j=1}^{J} |A_j| v_j \right) \int_{y_B}^{y_D} \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})| dx_1 dy_2. \tag{4.47}$$

9.

$$I_9 = \int_\omega \int_\omega |u^T(y_{AC}) - u^T(y)|\, |u^T(y_{BD}) - u^T(y)|\, dxdy.$$

We may proceed very similarly to the estimations of I_1 and we obtain that

$$I_9 \leq |\omega| \left(\sum_{j=1}^{J} |A_j| v_j \right)^2. \tag{4.48}$$

In order to conclude the proof of Theorem 4.12, we need the following lemma, a proof of which is postponed to 4.5.

Lemma 4.14. *There exists a constant C_1 depending only on Ω such that*

$$\int_{y_B}^{y_D} \int_{x_A}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2 \leq C_1 diam(\omega) \left(\sum_{j=1}^{J} |A_j| v_j \right),$$

$$\int_{x_A}^{x_C} \int_{y_B}^{y_D} |u^T(x_{BD}) - u^T(y_{AC})|\, dx_2 dy_1 \leq C_1 diam(\omega) \left(\sum_{j=1}^{J} |A_j| v_j \right).$$

Applying Lemma 4.14 and combining estimations (4.40) to (4.48) with the bound (4.34) results in

$$\int_\Omega \int_\Omega |u^T(x) - u^T(y)|^2 dxdy \leq C_2^2 \left(\sum_{j=1}^{J} |A_j| v_j \right)^2,$$

where $C_2^2 = (4 + 4C_1 + C_1^2) diam^2(\omega)$. Now this inequality may be treated exactly like (4.26), and there holds

$$\int_\omega \int_\omega |u^T(x) - u^T(y)|^2 dxdy \leq \frac{4C_2^2}{\sin^2 \theta^*} |\omega| \sum_{j=1}^{J} |(\boldsymbol{\nabla}_h^D u)_j|^2 |D_j|$$

From (4.31), we have

$$\int_\omega (u^T(x) - m_\omega^T(u))^2 dx \leq \frac{4C_2^2}{\sin^2 \theta^*} \sum_{j=1}^{J} |(\boldsymbol{\nabla}_h^D u)_j|^2 |D_j|,$$

which implies the desired result with $C = \frac{2C_2}{\sin \theta^*}$.

□

The second step in the proof of a discrete version of (4.2) is to establish an inequality related to the mean value on the boundary of a convex polygonal domain

Theorem 4.15 (Mean boundary Inequality). *Let Ω be an open bounded polygonal connected subset of $\mathbb{R}^2$ and let ω be an open polygonal convex subset of Ω and $\mathcal{I} \subset \partial\omega$, with $|\mathcal{I}| > 0$ ($|\mathcal{I}|$ is the one-dimensional Lebesgue measure of $\mathcal{I}$). Assume that $\mathcal{I}$ is included in a hyperplane of $\mathbb{R}^2$. Let $u = (u^T, u^P) \in \mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^J$ be given and the associated piecewise constant functions u^T and u^P be defined through Definition 4.4. Let $\gamma^T(u)(\sigma) = u_i^T$ for all $\sigma \in \overline{T}_i \cap \partial\omega$. (If $\sigma \in \overline{T}_i \cap \overline{T}_{i'}$, then the choice of u_i^T or $u_{i'}^T$ in the definition of γ^T does not matter). Let $\gamma^P(u)(\sigma) = u_k^P$ for all $\sigma \in \overline{P}_k \cap \partial\omega$. (If $\sigma \in \overline{P}_k \cap \overline{P}_{k'}$, then the choice of u_k^P or $u_{k'}^P$ in the definition of γ^P does not matter). Let $m_\mathcal{I}^T(u)$ (resp $m_\mathcal{I}^P(u)$) be the mean value of $\gamma^T(u)$ (resp $\gamma^P(u)$) on I. Let θ^* be defined through Definition 4.1. Then, there exists a constant C, only depending on Ω, ω, $\mathcal{I}$ and θ^* such that*

$$\|u^T - m_\mathcal{I}^T(u)\|_{L^1(\omega)} \leq C \|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.49}$$

$$\|u^P - m_\mathcal{I}^P(u)\|_{L^1(\omega)} \leq C \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.50}$$

Proof. Since $\mathcal{I}$ is included in a hyperplane, it may be assumed, without loss of generality, that $\mathcal{I} = \{0\} \times [a, b]$ and $\omega \subset \mathbb{R}_+ \times \mathbb{R}$ (the convexity of ω is used here). We choose points A, B, C and D, belonging to $\overline{\omega}$, such that

$$x_A = \inf\{x_1; (x_1, x_2) \in \omega\}, \quad x_C = \sup\{x_1; (x_1, x_2) \in \omega\},$$
$$y_B = \inf\{x_2; (x_1, x_2) \in \omega\}, \quad y_D = \sup\{x_2; (x_1, x_2) \in \omega\}.$$

Remark 4.16. *It may happen in particular cases that those four points are not different one from the other, but this does not change the general idea of the proof. If $A = B$ and $\mathcal{I} = [BD]$, then it even simplifies the proof since in that case, we do not have to introduce the point σ_{BD} defined below.*

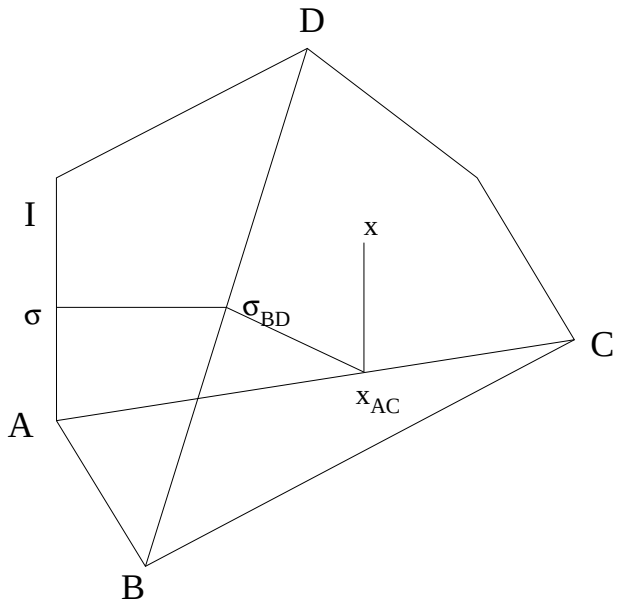

Figure 4.7: Notation for points A, B, C, D and points x_{AC}, σ_{BD}.

For any $x = (x_1, x_2) \in \omega$ and $\sigma = (\sigma_1, \sigma_2) \in \mathcal{I}$, we define $x_{AC} \in AC$ such that $(x_{AC})_1 = x_1$ and $\sigma_{BD} \in BD$ such that $(\sigma_{BD})_2 = \sigma_2$. The notations are summarized in Fig 4.7. The following triangle inequality holds:

$$\begin{aligned} |u^T(x) - \gamma u^T(\sigma)| \leq & \quad |u^T(x) - u^T(x_{AC})| \\ & + \quad |u^T(x_{AC}) - u^T(\sigma_{BD})| + |\gamma u^T(\sigma) - u^T(\sigma_{BD})|. \end{aligned}$$

Moreover, there holds

$$\begin{aligned} \|u^T - m_{\mathcal{I}}^T(u)\|_{L^1(\omega)} &= \int_\omega \left| u^T(x) - \frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} \gamma u^T(\sigma) d\sigma \right| dx \\ &= \int_\omega \left| \frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} [u^T(x) - \gamma u^T(\sigma)] d\sigma \right| dx \\ &\leq \frac{1}{|\mathcal{I}|} \int_\omega \int_{\mathcal{I}} \left| u^T(x) - \gamma u^T(\sigma) \right| d\sigma dx, \end{aligned}$$

so that, taking into account the above triangle inequality, we obtain:

$$\begin{aligned} &\|u^T - m_{\mathcal{I}}^T(u)\|_{L^1(\omega)} \leq \frac{1}{|\mathcal{I}|} \int_\omega \int_{\mathcal{I}} |u^T(x) - u^T(x_{AC})| \, d\sigma dx \\ &+ \frac{1}{|\mathcal{I}|} \int_\omega \int_{\mathcal{I}} |u^T(x_{AC}) - u^T(\sigma_{BD})| \, d\sigma dx + \frac{1}{|\mathcal{I}|} \int_\omega \int_{\mathcal{I}} |\gamma u^T(\sigma) - u^T(\sigma_{BD})| \, d\sigma dx. \end{aligned}$$

We first observe that the function $|u^T(x) - u^T(x_{AC})|$ doesn't depend on the variable σ; then, using similar techniques to those which led to (4.36), and the fact that $\int_{x_A}^{x_C} \chi_j^{T,1}(x) dx_1 \leq |A_j|$, there holds

$$\frac{1}{|\mathcal{I}|} \int_\omega \int_{\mathcal{I}} |u^T(x) - u^T(x_{AC})| \, d\sigma dx \leq diam(\omega) \left(\sum_{j=1}^{J} |A_j| v_j \right), \tag{4.51}$$

where we recall the notation $v_j = |(\boldsymbol{\nabla}_h^D u)_j \cdot \overrightarrow{G_{i_1(j)}G_{i_2(j)}}|$.

Then, we know that the function $|\gamma u^T(\sigma) - u^T(\sigma_{BD})|$ only depends on the variable σ; then, using similar techniques to those which led to (4.37), and the fact that $\int_{\mathcal{I}} \chi_j^{T,2}(\sigma)d\sigma \leq |A_j|$, we have

$$\frac{1}{|\mathcal{I}|}\int_\omega \int_{\mathcal{I}} |\gamma u^T(\sigma) - u^T(\sigma_{BD})|\, d\sigma dx \leq \frac{|\omega|}{|\mathcal{I}|}\left(\sum_{j=1}^J |A_j| v_j\right). \tag{4.52}$$

Now, x_{AC} doesn't depend on the variable x_2, so that

$$\frac{1}{|\mathcal{I}|}\int_\omega \int_{\mathcal{I}} |u^T(x_{AC}) - u^T(\sigma_{BD})|\, d\sigma dx \leq \frac{diam(\omega)}{|\mathcal{I}|}\int_{x_A}^{x_C}\int_{\mathcal{I}} |u^T(x_{AC}) - u^T(\sigma_{BD})|\, d\sigma dx_1.$$

Applying an inequality like in Lemma 4.14 leads to

$$\frac{1}{|\mathcal{I}|}\int_\omega \int_{\mathcal{I}} |u(x_{AC}) - u(\sigma_{BD})|\, d\sigma dx \leq \frac{C_1 diam^2(\omega)}{|\mathcal{I}|}\left(\sum_{j=1}^J |A_j| v_j\right). \tag{4.53}$$

Using (4.51), (4.52) and (4.53), we conclude that

$$\|u^T - m_{\mathcal{I}}^T(u)\|_{L^1(\omega)} \leq \left[diam(\omega) + \frac{|\omega|}{|\mathcal{I}|} + \frac{C^* diam^2(\omega)}{|\mathcal{I}|}\right]\left(\sum_{j=1}^J |A_j| v_j\right).$$

Then, the Cauchy-Schwarz inequality yields (4.49). Similarly, we also obtain (4.50). □

Now, we come to the final step of our result.

Theorem 4.17 (Mean Poincaré Inequality). *Let Ω be an open bounded polygonal connected subset of $\mathbb{R}^2$; let $u = (u^T, u^P)$ be in $\mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$, and $u^T(x)$, $u^P(x)$ be defined through Definition 4.4. Let θ^* be defined through definition 4.1. Then, there exists a constant C only depending on Ω and θ^* such that*

$$\|u^T - m_\Omega^T(u)\|_{L^2(\Omega)} \leq C\|\boldsymbol{\nabla}_h^D u\|_D \tag{4.54}$$

and

$$\|u^P - m_\Omega^P(u)\|_{L^2(\Omega)} \leq C\|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.55}$$

where $m_\Omega^T(u)$ (resp. $m_\Omega^P(u)$) is the mean-value of u^T (resp. u^P) on Ω.

Proof. Since Ω is polygonal, there exists a finite number of disjoint convex polygonal sets, denoted by $\{\Omega_1, ..., \Omega_n\}$, such that $\overline{\Omega} = \cup_{i=1}^n \overline{\Omega}_i$. Let $\mathcal{I}_{i,j} = \overline{\Omega}_i \cap \overline{\Omega}_j$ and B be the set of couples $(i,j) \in \{1, ..., n\}^2$ such that $i \neq j$ and the one-dimensional Lebesgue measure of $\mathcal{I}_{i,j}$, denoted by $|\mathcal{I}_{i,j}|$ is strictly positive.
Let m_i denote the mean value of u^T on Ω_i, $i \in \{1, ..., n\}$, and $m_{i,j}$ denote the mean value of u^T on $\mathcal{I}_{i,j}$, $(i,j) \in B$. Note that $m_{i,j} = m_{j,i}$ for all $(i,j) \in B$.
Theorem 4.12 gives the existence of C_i, $i \in \{1, ..., n\}$ only depending on Ω (since the Ω_i only depend on Ω) and θ^*, such that

$$\|u^T - m_i\|_{L^2(\Omega_i)} \leq C_i \|\boldsymbol{\nabla}_h^D u\|_D, \quad \forall i \in \{1, ..., n\}. \tag{4.56}$$

Applying the Cauchy-Schwarz inequality, we have

$$\|u^T - m_i\|_{L^1(\Omega_i)} \leq |\Omega_i|^{1/2} C_i \, \|\boldsymbol{\nabla}_h^D u\|_D, \quad \forall i \in \{1, ..., n\}. \tag{4.57}$$

Moreover, Theorem 4.15 gives the existence of $C_{i,j}$, $(i,j) \in B$, only depending on Ω and θ^*, such that

$$\|u^T - m_{i,j}\|_{L^1(\Omega_i)} \leq C_{i,j} \, \|\boldsymbol{\nabla}_h^D u\|_D, \quad \forall (i,j) \in B. \tag{4.58}$$

Then, one has, by a triangle inequality

$$|\Omega_i| \, |m_i - m_{i,j}| = \|m_i - m_{i,j}\|_{L^1(\Omega_i)} \leq \left(|\Omega_i|^{1/2} C_i + C_{i,j}\right) \|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.59}$$

for all $(i,j) \in B$. Applying a triangular inequality and using the fact that $m_{i,j} = m_{j,i}$, we get from (4.59) that there exists a constant $C'_{i,j}$ only depending on Ω and θ^* such that

$$|m_i - m_j| \leq C'_{i,j} \, \|\boldsymbol{\nabla}_h^D u\|_D, \tag{4.60}$$

for all $(i,j) \in B$.

Since Ω in connected, we can always connect any $(i,j) \in \{1, ..., n\}^2$ by a finite set of couples belonging to B. Applying triangular inequalities and the related inequalities (4.60), we obtain the existence of $K_{i,j}$, only depending on Ω and θ^*, such that $|m_i - m_j| \leq K_{i,j} \|\boldsymbol{\nabla}_h^D u\|_D$ for all $(i,j) \in \{1, ..., n\}^2$, and therefore, the existence of a constant M_i, only depending on Ω and θ^*, such that

$$\left|m_\Omega^T(u) - m_i\right| = \left| \frac{1}{|\Omega|} \sum_{j \in [1,n]} |\Omega_j| (m_j - m_i) \right| \leq M_i \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.61}$$

Then, (4.56), (4.61) and a triangle inequality yield

$$\|u^T - m_\Omega^T(u)\|_{L^2(\Omega_i)} \leq \|u^T - m_i\|_{L^2(\Omega_i)} + |\Omega_i|^{1/2} \left|m_\Omega^T(u) - m_i\right| \leq \left(C_i + M_i |\Omega_i|^{1/2}\right) \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.62}$$

Summing up the squares of inequalities (4.62) over $i \in \{1, ..., n\}$ yields (4.54). We obtain (4.55) in a similar way. This completes the proof of Theorem 4.17. □

Corollary 4.18. *Let Ω be an open bounded polygonal connected subset of $\mathbb{R}^2$; let $u = (u^T, u^P)$ be in $\mathbb{R}^{I+J^\Gamma} \times \mathbb{R}^K$, and such that*

$$\sum_{i=1}^{I} |T_i| u_i^T = \sum_{k=1}^{K} |P_k| u_k^P = 0.$$

Let θ^* be defined through definition 4.1. Then, there exists a constant C only depending on Ω and θ^* such that

$$\|u\|_{T,P} \leq C \|\boldsymbol{\nabla}_h^D u\|_D. \tag{4.63}$$

4.4 Application

Theorem 4.19. *Let Ω be a two-dimensional polygonal domain with exterior boundary denoted by Γ_0 and internal connected components denoted by Γ_q, with $q \in [1,Q]$. There exists a constant C depending only on Ω and θ^* defined by Definition 4.1, such that for any discrete vector field $(\mathbf{u}_j)_{j\in[1,J]}$ with $\mathbf{u}\cdot\mathbf{n}=0$ on Γ and $(\mathbf{u}\cdot\boldsymbol{\tau},1)_{\Gamma_q,h}=0$ for all $q\in[1,Q]$, there holds*

$$||\mathbf{u}||_D \le C\left(||\nabla^{T,P}\cdot\mathbf{u}||_{T,P}+||\nabla^{T,P}\times\mathbf{u}||_{T,P}\right). \tag{4.64}$$

Proof. Let $(\mathbf{u}_j)_{j\in[1,J]}$ be given with $\mathbf{u}\cdot\mathbf{n}=0$ on Γ and $(\mathbf{u}\cdot\boldsymbol{\tau},1)_{\Gamma_q,h}=0$ for all $q\in[1,Q]$. According to Theorem 4.6, there exists $\phi=(\phi_i^T,\phi_k^P)_{i\in[1,I+J^\Gamma],k\in[1,K]}$, $\psi=(\psi_i^T,\psi_k^P)_{i\in[1,I+J^\Gamma],k\in[1,K]}$ and $(c_q^T,c_q^P)_{q\in[1,Q]}$ such that (4.8) holds, the decomposition being orthogonal. Then there holds

$$||\mathbf{u}||_D^2=(\mathbf{u},\boldsymbol{\nabla}_h^D\phi)_D+(\mathbf{u},\boldsymbol{\nabla}_h^D\times\psi)_D \tag{4.65}$$

and

$$||\boldsymbol{\nabla}_h^D\phi||_D\le||\mathbf{u}||_D\quad\text{and}\quad||\boldsymbol{\nabla}_h^D\times\psi||_D=||\boldsymbol{\nabla}_h^D\psi||_D\le||\mathbf{u}||_D. \tag{4.66}$$

Using the discrete integration by part properties (4.16) and (4.17) in (4.65), we obtain

$$||\mathbf{u}||_D^2=-(\nabla_h^{T,P}\cdot\mathbf{u},\phi^{T,P})_{T,P}+(\mathbf{u}\cdot\mathbf{n},\tilde{\phi})_{\Gamma,h}+(\nabla_h^{T,P}\times\mathbf{u},\psi^{T,P})_{T,P}-(\mathbf{u}\cdot\boldsymbol{\tau},\tilde{\psi})_{\Gamma,h}. \tag{4.67}$$

In (4.67), both boundary terms vanish. The first because $\mathbf{u}\cdot\mathbf{n}=0$ on Γ. As far as the second is concerned, from (4.11) and the definition of the boundary scalar product (4.6) we have

$$(\mathbf{u}\cdot\boldsymbol{\tau},\tilde{\psi})_{\Gamma,h}=(\mathbf{u}\cdot\boldsymbol{\tau},\tilde{\psi})_{\Gamma_0,h}+\sum_{q\in[1,Q]}\left(\frac{c_q^T+c_q^P}{2}\right)(\mathbf{u}\cdot\boldsymbol{\tau},1)_{\Gamma_q,h},$$

so that (4.10) and the fact that $(\mathbf{u}\cdot\boldsymbol{\tau},1)_{\Gamma_q,h}=0$ for all $q\in[1,Q]$ allow us to conclude that $(\mathbf{u}\cdot\boldsymbol{\tau},\tilde{\psi})_{\Gamma,h}=0$. Thus, we have

$$||\mathbf{u}||_D^2=-(\nabla_h^{T,P}\cdot\mathbf{u},\phi^{T,P})_{T,P}+(\nabla_h^{T,P}\times\mathbf{u},\psi^{T,P})_{T,P} \tag{4.68}$$

Using the Cauchy-Schwarz inequality in (4.68), and then applying Theorem 4.11 for ψ and Corollary 4.18 for ϕ, we get (4.64) from (4.66). □

4.5 Proof of Lemma 4.14

We shall only give the proof of the first inequality in Lemma 4.14, since the proof of the other inequality follows exactly the same lines. If the four points (A,B,C,D) are all different, then we may denote by I the intersection of AC and BD, and the angle α between the diagonals AC and BD is different from 0. This is also the case of the angles β_i and γ_i displayed on Fig. 4.8. If ω is a triangle, up to a rotation, we have that $A=B$ and we set $I=A=B$. Then, the angles α, β_1 and γ_1 are all different from 0 and evaluating the term G in (4.69) reduces to the evaluation of H_1, which simplifies the proof. Let us go back to the general case. We set

$$G=\int_{y_B}^{y_D}\int_{x_A}^{x_C}|u^T(x_{AC})-u^T(y_{BD})|\,dx_1dy_2=H_1+H_2+H_3+H_4, \tag{4.69}$$

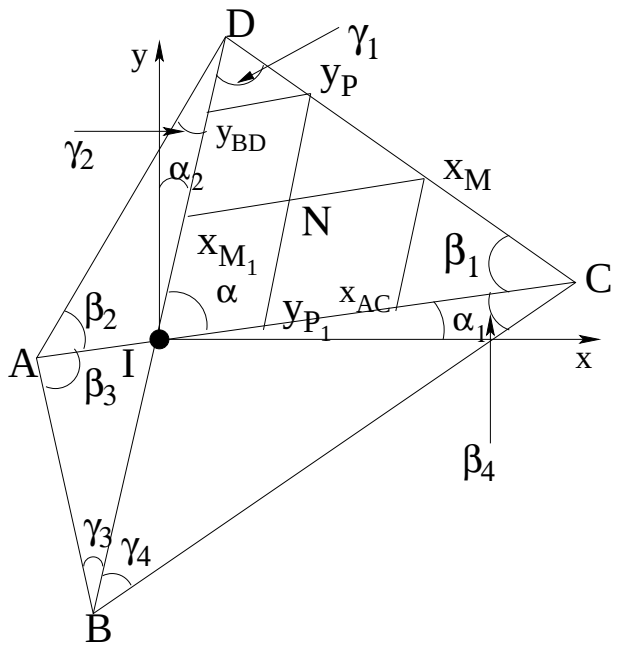

Figure 4.8: $x_{AC}x_M x_{M_1}$ intersects DC before it intersects $y_{BD}y_P y_{P_1}$.

where

$$H_1 = \int_{y_I}^{y_D}\int_{x_I}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2,$$
$$H_2 = \int_{y_I}^{y_D}\int_{x_A}^{x_I} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2,$$
$$H_3 = \int_{y_B}^{y_I}\int_{x_A}^{x_I} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2,$$
$$H_4 = \int_{y_B}^{y_I}\int_{x_I}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2.$$

We only estimate the first term in the right-hand side of inequality (4.69), since the other may be treated similarly. For any $x_{AC} \in IC$ and $y_{BD} \in ID$, let x_M (resp. y_P) be the intersection of DC with the straight line going though x_{AC} (resp. y_{BD}) and parallel to the segment $[ID]$ (resp. $[IC]$), and let x_{M_1} (resp. y_{P_1}) be the intersection of ID (resp. IC) with the straight line going through x_M (resp. x_P) and parallel to the segment IC (resp. ID). Then, we shall examine two cases, according to where the broken line $x_{AC}x_M x_{M_1}$ intersects with the broken line $y_{BD}y_P y_{P_1}$ at point N.

Case 1: The broken line $x_{AC}x_M x_{M_1}$ intersects DC at x_M before it intersects the broken line $y_{BD}y_P y_{P_1}$, (see Fig. 4.8). Then, using the triangle inequality leads to

$$\begin{aligned}|u^T(x_{AC}) - u^T(y_{BD})| \le\ & |u^T(x_{AC}) - u^T(x_M)| + |u^T(x_M) - u^T(N)| \\ & + |u^T(N) - u^T(y_P)| + |u^T(y_P) - u(y_{BD})|.\end{aligned}$$

Let the function χ_j from $\mathbb{R}^2 \times \mathbb{R}^2$ to $\{0,1\}$ be defined by

$$\chi_j(x,y) = \begin{cases} 1 & \text{if } [x,y] \cap A_j \neq \emptyset \\ 0 & \text{if } [x,y] \cap A_j = \emptyset \end{cases} \tag{4.70}$$

Recalling once again the notation $v_j = |(\boldsymbol{\nabla}_h^D u^T)_j \cdot \overrightarrow{G_{i_1(j)}G_{i_2(j)}}|$, we have that

$$|u^T(x_M) - u^T(N)| \leq \sum_{j=1}^{J} \chi_j(x_M, N)\, v_j \leq \sum_{j=1}^{J} \chi_j(x_M, x_{M_1})\, v_j, \tag{4.71}$$

due to the fact that since $N \in [x_M x_{M_1}]$ then $\chi_j(x_M, N) \leq \chi_j(x_M, x_{M_1})$.

Similarly, we obtain that

$$|u^T(N) - u^T(y_P)| \leq \sum_{j=1}^{J} \chi_j(y_P, y_{P_1})\, v_j. \tag{4.72}$$

We also have

$$|u^T(x_{AC}) - u^T(x_M)| \leq \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j, \tag{4.73}$$

and

$$|u^T(y_P) - u^T(y_{BD})| \leq \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j. \tag{4.74}$$

From (4.71)-(4.74), we have

$$\begin{aligned} |u^T(x_{AC}) - u^T(y_{BD})| \leq & \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j + \sum_{j=1}^{J} \chi_j(x_M, x_{M_1})\, v_j \\ & + \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j + \sum_{j=1}^{J} \chi_j(y_P, y_{P_1})\, v_j. \end{aligned}$$

Case 2: The broken line $x_{AC}x_M x_{M_1}$ intersects the broken line $y_{BD}y_P y_{P_1}$ at N before it intersects DC (*see*Fig 4.9). We use the triangle inequality to obtain

$$|u^T(x_{AC}) - u^T(y_{BD})| \leq |u^T(x_{AC}) - u^T(N)| + |u^T(N) - u^T(y_{BD})|. \tag{4.75}$$

Similarly to Case 1, since $N \in [x_{AC}x_M]$ and $N \in [y_{BD}y_P]$, there holds

$$|u^T(x_{AC}) - u^T(N)| \leq \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j \tag{4.76}$$

$$|u^T(N) - u^T(y_{BD})| \leq \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j. \tag{4.77}$$

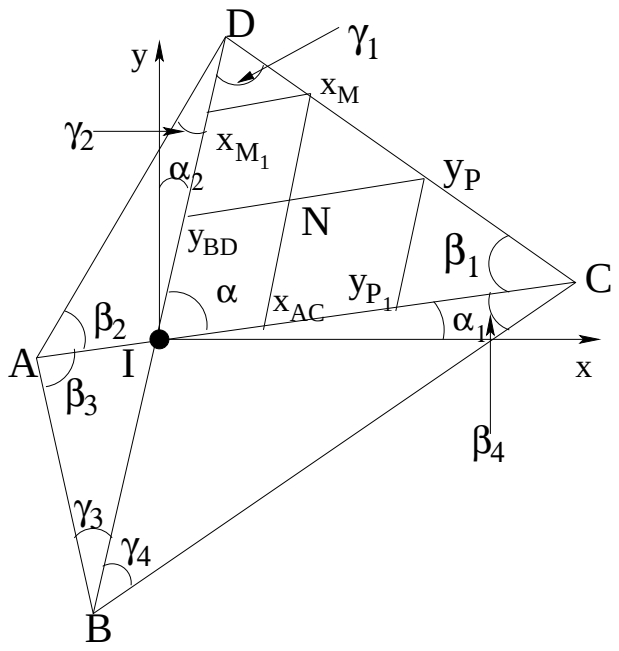

Figure 4.9: $x_{AC}x_M x_{M_1}$ intersects $y_{BD}y_P y_{P_1}$ before it intersects DC.

Adding (4.76) to (4.77), and combining with (4.75) we have

$$|u^T(x_{AC}) - u^T(y_{BD})| \leq \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j + \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j.$$

So that in both cases, we always obtain

$$\begin{aligned} |u^T(x_{AC}) - u^T(y_{BD})| \leq & \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j + \sum_{j=1}^{J} \chi_j(x_M, x_{M_1})\, v_j \\ & + \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j + \sum_{j=1}^{J} \chi_j(y_P, y_{P_1})\, v_j. \end{aligned}$$

We thus always have

$$H_1 = \int_{y_I}^{y_D} \int_{x_I}^{x_C} |u^T(x_{AC}) - u^T(y_{BD})|\, dx_1 dy_2 \leq L_1 + L_2 + L_3 + L_4, \tag{4.78}$$

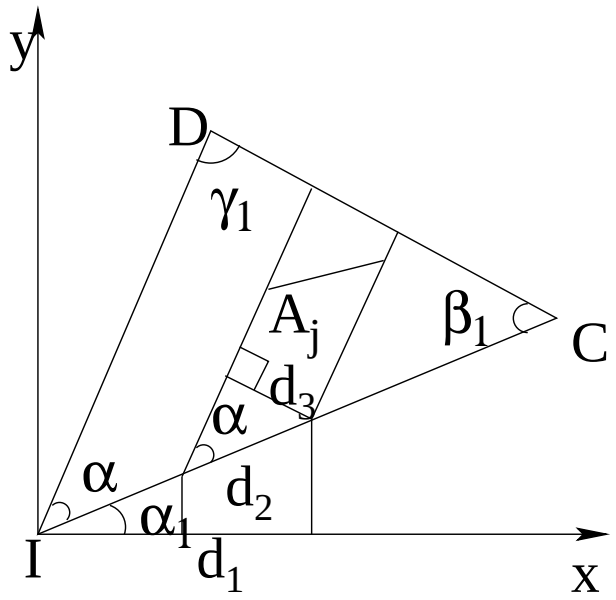

Figure 4.10: How to estimate the term $\int_{x_I}^{x_C} \chi_j(x_{AC}, x_M) dx_1$.

where L_1, L_2, L_3, and L_4 are defined as follows:

$$L_1 = \int_{y_I}^{y_D} \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j dx_1 dy_2,$$

$$L_2 = \int_{y_I}^{y_D} \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(x_M, x_{M_1})\, v_j dx_1 dy_2,$$

$$L_3 = \int_{y_I}^{y_D} \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(y_{BD}, y_P)\, v_j dx_1 dy_2,$$

$$L_4 = \int_{y_I}^{y_D} \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(y_P, y_{P_1})\, v_j dx_1 dy_2.$$

Observing that $\chi_j(x_{AC}, x_M)$ only depends on variable x_1, we find

$$\begin{aligned} L_1 &\leq (y_D - y_I) \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(x_{AC}, x_M)\, v_j dx_1 \\ &= (y_D - y_I) \sum_{j=1}^{J} \int_{x_I}^{x_C} \chi_j(x_{AC}, x_M) dx_1 v_j. \end{aligned}$$

Let us take a look at Fig. 4.10 and its associated notations. Simple geometrical arguments show that

$$\int_{x_I}^{x_C} \chi_j(x_{AC}, x_M) dx_1 =: d_1 = d_2 \cos\alpha_1 = d_3 \frac{\cos\alpha_1}{\sin\alpha} \leq \frac{\cos\alpha_1 |A_j|}{\sin\alpha}.$$

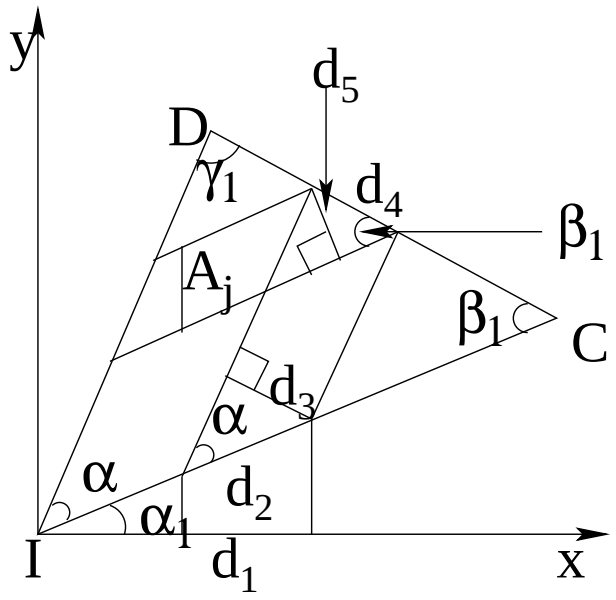

Figure 4.11: How to estimate the term $\int_{x_I}^{x_C} \chi_j(x_M, x_{M_1}) dx_1$.

This results in

$$L_1 \leq (y_D - y_I) \frac{\cos \alpha_1}{\sin \alpha} \left(\sum_{j=1}^{J} |A_j| v_j \right). \tag{4.79}$$

Moreover, there holds

$$\begin{aligned} L_2 &\leq (y_D - y_I) \int_{x_I}^{x_C} \sum_{j=1}^{J} \chi_j(x_M, x_{M_1})\, v_j dx_1 \\ &= (y_D - y_I) \sum_{j=1}^{J} \int_{x_I}^{x_C} \chi_j(x_M, x_{M_1}) dx_1 v_j. \end{aligned}$$

Let us take a look at Fig. 4.11 and its associated notations. Simple geometrical arguments show that

$$\begin{aligned} \int_{x_I}^{x_C} \chi_j(x_M, x_{M_1}) dx_1 \quad &=: \quad d_1 = d_2 \cos \alpha_1 = d_3 \frac{\cos \alpha_1}{\sin \alpha} \\ &= \quad d_4 \frac{\cos \alpha_1 \sin \gamma_1}{\sin \alpha} = d_5 \frac{\cos \alpha_1 \sin \gamma_1}{\sin \alpha \, \sin \beta_1} \leq \frac{\cos \alpha_1 \sin \gamma_1 |A_j|}{\sin \alpha \, \sin \beta_1}. \end{aligned}$$

So that there holds

$$L_2 \leq \frac{\cos \alpha_1 \sin \gamma_1}{\sin \alpha \, \sin \beta_1} (y_D - y_I) \left(\sum_{j=1}^{J} |A_j| v_j \right). \tag{4.80}$$

Similarly,

$$L_3 \leq \frac{\cos\alpha_2}{\sin\alpha}(x_C - x_I)\left(\sum_{j=1}^{J}|A_j|v_j\right). \tag{4.81}$$

$$L_4 \leq \frac{\cos\alpha_2 \sin\beta_1}{\sin\alpha\ \sin\gamma_1}(x_C - x_I)\left(\sum_{j=1}^{J}|A_j|v_j\right). \tag{4.82}$$

From (4.78)-(4.82), we conclude that there exists a constant C depending only on the geometry of ω (since the angles depend only on the geometry of ω) such that

$$H_1 \leq C diam(\omega)\left(\sum_{j=1}^{J}|A_j|v_j\right). \tag{4.83}$$

Using similar techniques, we also obtain that

$$H_2 \leq C diam(\Omega)\left(\sum_{j=1}^{J}|A_j|v_j\right), \tag{4.84}$$

$$H_3 \leq C diam(\Omega)\left(\sum_{j=1}^{J}|A_j|v_j\right), \tag{4.85}$$

$$H_4 \leq C diam(\Omega)\left(\sum_{j=1}^{J}|A_j|v_j\right). \tag{4.86}$$

Combining (4.83)-(4.86) with (4.69), we have

$$\int_{y_B}^{y_D}\int_{x_A}^{x_C}|u^T(x_{AC}) - u^T(y_{BD})|\,dx_1dy_2 \leq C_1 diam(\Omega)\left(\sum_{j=1}^{J}|A_j|v_j\right),$$

where $C_1 = 4C$, which concludes the proof of Lemma 4.14.

References for the Chapter 4

[1] R. Adams, Sobolev Spaces, Academic Press, New York, 1975.

[2] B. Andreianov, F. Boyer, F. Hubert, *Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes*. Numer. Methods Partial Differ. Equations 23, pp. 145–195 (2007).

[3] C. Chainais-Hillairet, *Discrete duality finite volume schemes for two-dimensional drift-diffusion and energy-transport models*. Internat. J. Numer. Methods Fluids 59, pp. 239–257 (2009).

[4] Y. Coudière, J.-P. Vila, P. Villedieu, *Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem*. M2AN Math. Model. Numer. Anal. 33 (3), pp. 493–516 (1999).

[5] Y. Coudière, C. Pierre, O. Rousseau, R. Turpault, A 2D/3D Discrete Duality Finite Volume Scheme. Application to ECG simulation, International Journal On Finite Volumes 6 (1) (2009), electronic only.

[6] S. Delcourte, *Développement de méthodes de volumes finis pour la mécanique des fluides*.PhD thesis of the Université Paul Sabatier, Toulouse, France, 2007. Available at: http://tel.archives-ouvertes.fr/tel-00200833/fr/

[7] S. Delcourte, K. Domelevo, P. Omnes, *A discrete duality finite volume approach to Hodge decomposition and div-curl problems on almost arbitrary two-dimensional meshes*. SIAM J. Numer. Anal. 45, pp. 1142–1174 (2007).

[8] S. Delcourte, P. Omnes, *A discrete duality finite volume discretization of the vorticity-velocity-pressure formulation of the 2D Stokes problem on almost arbitrary two-dimensional grids*. 2011.

[9] K. Domelevo, P. Omnes, *A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids*. M2AN Math. Model. Numer. Anal. 39 (6) pp. 1203–1249 (2005).

[10] R. Eymard, T. Gallouët, *R. Herbin, Finite Volume methods*. In: Ciarlet, P.G. and Lions, J.-L., Eds. Handbook of Numerical Analysis, Vol. VII, Amsterdam: North-holland, pp. 713–1020 (2000).

[11] P. Fernandes, G. Gilardi, *Magnetostatic and electrostatic problems in inhomogeneous anisotropic media with irregular boundary and mixed boundary conditions.* Math. Models Methods Appl. Sci. 7, pp. 957–991 (1997).

[12] A. Glitzky, J. A. Griepentrog, *Discrete Sobolev-Poincaré inequalities for Voronoi finite volume approximations.* SIAM J. Numer. Anal. 48, pp. 372–391 (2010).

[13] F. Hermeline, *A finite volume method for the approximation of diffusion operators on distorted meshes.* J. Comput. Phys. 160, pp. 481–499 (2000).

[14] F. Hermeline, *Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes.* Comput. Methods Appl. Mech. Eng. 192, pp. 1939–1959 (2003).

[15] F. Hermeline, S. Layouni, P. Omnes, *A finite volume method for the approximation of Maxwell's equations in two space dimensions on arbitrary meshes.* Journal of Computational Physics 227, pp. 9365–9388 (2008).

[16] S. Krell, *Stabilized DDFV schemes for Stokes problem with variable viscosity on general 2D meshes.* To be published in Num. Meth. for PDEs, http://dx.doi.org/10.1002/num.20603 (2010).

[17] S. Krell, *Schémas Volumes Finis en mécanique des fluides complexes.* PhD Thesis of the Univeristé de Provence Aix-Marseille I, 2010. Available at: http://tel.archives-ouvertes.fr/tel-00524509/fr/.

[18] A. Njifenjou, A.J. Kinfack, *Convergence analysis of an MPFA method for flow problems in anisotropic heterogeneous porous media.* International Journal On Finite Volumes 5 (1) (2008), electronic only.

[19] M.Vohralík, *On the Discrete Poincaré-Friedrichs Inequalities for Nonconforming Approximations of the Sobolev Space* H^1. Numer. Funct. Anal. Optim. 26 (7-8), pp. 925–952 (2005).