



**HAL**  
open science

# Un schéma éléments finis non-conformes / volumes finis pour l'approximation en maillages non-structurés des écoulements à faible nombre de Mach

Guillaume Ansanay-Alex

► **To cite this version:**

Guillaume Ansanay-Alex. Un schéma éléments finis non-conformes / volumes finis pour l'approximation en maillages non-structurés des écoulements à faible nombre de Mach. Mathématiques [math]. Université de Provence - Aix-Marseille I, 2009. Français. NNT : 2009AIX11070 . tel-00483598

**HAL Id: tel-00483598**

**<https://theses.hal.science/tel-00483598v1>**

Submitted on 14 May 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE PROVENCE  
U.F.R. M.I.M.  
ÉCOLE DOCTORALE DE MATHÉMATIQUES ET INFORMATIQUE E.D. 184

## THÈSE

présentée pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ DE PROVENCE

*Spécialité : Mathématiques*

par

**Guillaume ANSANAY-ALEX**

sous la direction du Pr. Raphaèle HERBIN

*Titre :*

**UN SCHEMA ELEMENTS FINIS NON-CONFORMES / VOLUMES FINIS  
POUR L'APPROXIMATION EN MAILLAGES NON-STRUCTURES  
DES ECOULEMENTS A FAIBLE NOMBRE DE MACH**

soutenue publiquement le 17 juin 2009

### JURY

M.	Stéphane	CLAIN	MIP, Toulouse	<i>Rapporteur</i>
M.	Cédric	GALUSINSKI	Université de Toulon	<i>Examinateur</i>
M.	Thierry	GOUDON	INRIA, Lille	<i>Rapporteur</i>
Mme	Raphaèle	HERBIN	LATP, Aix-Marseille I	<i>Directrice de thèse</i>
Mme	Florence	HUBERT	LATP, Aix-Marseille I	<i>Examinatrice</i>
M.	Jean-Claude	LATCHE	IRSN, Cadarache	<i>Examinateur</i>
M.	Didier	VOLA	IRSN, Cadarache	<i>Encadrant</i>



# Remerciements

Je remercie d'abord Didier Vola, qui a encadré ma thèse à l'IRSN, ce qui est déjà en soi une épreuve, mais qui m'a aussi aidé à déménager, à perfectionner ma conduite avant que je n'achète une voiture, puis à dégager ma voiture de la neige. Tu m'as surtout soutenu dans les moments difficiles, et transmis une manière de travailler, de la rigueur. Aujourd'hui j'applique ces leçons tous les jours dans mon travail, en me disant "Que ferait Didier?".

Je tiens à remercier Raphaële Herbin, ma directrice de thèse, qui a toujours été là pour me donner de précieux conseils et me sauver dans les situations où je ne savais plus quoi faire. Merci beaucoup, pour tout. Un grand merci également à Thierry Gallouët et Jean-Claude Latché, qui m'ont plusieurs fois sauvé la mise quand j'étais perdu dans divers espaces!

Je souhaite remercier Stéphane Clain et Thierry Goudon, qui ont été rapporteurs de ce mémoire, pour l'attention qu'ils ont porté à ce travail. Je remercie également Florence Hubert d'avoir accepté d'être la présidente du jury, et Cédric Galusinski d'y avoir participé.

Merci à Laurence Rigollet qui m'a accueilli dans son laboratoire pendant ma thèse, et merci aussi à tous les membres du laboratoire, en particulier à Bruno et Fabrice pour leur influence bénéfique sur ma pratique de la programmation, qui sert maintenant dans d'autres domaines de la simulation. Un merci particulier à mes géniales voisines de bureau, Laura et Libuška!

Merci à tous les doctorants et stagiaires que j'ai connus au laboratoire, en particulier à Céline qui a toujours été là pour me rassurer mais aussi pour me gronder un peu, et pour m'encourager à aller à la salle de sport.

Un merci spécial à Aurélien, je ne détaille pas plus, Aurel c'est un peu comme un frère. D'autres mercis spéciaux à Marc et Alexandre, mes amis depuis les jeunes années!

Merci à tous mes amis aixois, les littéraires (bisous à Florence qui est venue à ma soutenance!), ceux qui gravitent autour, ceux du California... merci aussi à l'équipe de La Rotonde! Des mercis particuliers à mes compagnons du samedi, François-Jean avec qui nos discussions sur la vie, le monde, la société, la physique et la métaphysique ne se comptent pas en heures mais en jours. Avec lui et d'autres j'ai partagé les moments de bonheur et les moments plus difficiles : Edena, Serge, Nathaniel, Agnès, Estelle, William, Alain,... Merci aussi aux "générations" successives de 6 mois d'habitants du Hameau, et aux différentes promos de l'INSTN avec qui j'ai partagé mes repas du soir à la cantine.

Encore des mercis pour ceux qui m'ont accueilli à Paris, surtout à Bruno, Nicolas et Paul avec qui nous avons formé une équipe de développeurs de choc! Et un clin d'oeil à tout mon petit milieu parisien!



# Table des matières

<b>1 Synthèse générale</b>	<b>7</b>
1.1 Objectif de la thèse . . . . .	10
1.2 Discrétisation spatiale des équations de la dynamique . . . . .	11
1.2.1 Construction d'une approximation stable de l'opérateur d'advection . . . . .	12
1.2.2 Expérimentations numériques . . . . .	15
1.3 Discrétisation spatiale du transport diffusif de scalaire . . . . .	19
1.3.1 Approximation des flux convectifs . . . . .	19
1.3.2 Approximation des flux diffusifs . . . . .	22
1.3.3 Expérimentations numériques . . . . .	28
1.4 Plan du manuscrit . . . . .	32
<b>2 Approximation du terme de convection non linéaire</b>	<b>33</b>
2.1 Introduction . . . . .	33
2.2 Meshes and discretization spaces . . . . .	35
2.3 A $L^2$ -stable convection operator . . . . .	38
2.3.1 General form of the convection operator and stability analysis . . . . .	38
2.3.2 Interpolating the mass flowrates . . . . .	41
2.3.3 The discrete momentum balance equation . . . . .	46
2.3.4 Concluding remarks . . . . .	47
2.4 A fractional step scheme . . . . .	47
2.5 Numerical experiments . . . . .	50
2.5.1 Incompressible flows . . . . .	50
2.5.2 Variable density flows . . . . .	56
2.6 Conclusion . . . . .	58
<b>3 Advection–diffusion de scalaires par volumes finis</b>	<b>61</b>
3.1 Introduction . . . . .	61
3.2 Notations . . . . .	63
3.3 Discretization of the convection operator . . . . .	64
3.3.1 The intersection multislope MUSCL method . . . . .	65
3.3.2 The center multislope MUSCL method . . . . .	67
3.3.3 Proposition for an extension to non-simplexes and/or non-matching grids . . . . .	68
3.4 Discretization of the diffusion operator . . . . .	69
3.5 The complete discrete scheme . . . . .	70
3.6 Numerical examples . . . . .	71

3.6.1	Multislope MUSCL verification tests . . . . .	72
3.6.2	SUSHI verification tests . . . . .	76
3.6.3	Multislope MUSCL - SUSHI coupling verification . . . . .	78
3.7	Discussion, conclusions . . . . .	79
<b>4</b>	<b>Simulations numériques</b>	<b>87</b>
4.1	Ecoulement dans un tube coudé – Hexaèdres . . . . .	87
4.2	Flamme laminaire 3D . . . . .	93
4.3	Aéraulique anisotherme à faible nombre de Mach . . . . .	97
<b>A</b>	<b>Elements finis non conformes</b>	<b>103</b>
A.1	Element de Rannacher-Turek . . . . .	103
A.2	Element de Crouzeix-Raviart . . . . .	106
<b>B</b>	<b>Contribution FVCA5</b>	<b>107</b>
B.1	Introduction . . . . .	107
B.2	A finite volume result . . . . .	108
B.3	A convection operator for low-order non-conforming FE . . . . .	109
B.3.1	Discretization spaces . . . . .	109
B.3.2	A convection operator . . . . .	109
B.4	A pressure correction scheme . . . . .	110
B.5	Numerical test . . . . .	112
<b>C</b>	<b>Contribution FVCA5 – Benchmark</b>	<b>115</b>
C.1	Presentation of the scheme . . . . .	115
C.2	Implementation . . . . .	116
C.3	Numerical results . . . . .	117
C.4	Comments on the results . . . . .	128

# Chapitre 1

## Synthèse générale

Dans le cadre de ses études de sûreté nucléaire, la Direction de la Prévention des Accidents Majeurs (DPAM) de l’Institut de Radioprotection et de Sûreté Nucléaire (IRSN) s’intéresse à différentes problématiques physiques pour la simulation desquelles sont développés un ensemble de codes de calcul dédiés en mécanique des fluides numérique :

- le code ISIS pour les incendies dans des locaux confinés,
- le code ARTIC pour la dispersion atmosphérique de polluants,
- le code TOPASE pour les écoulements diphasiques.

Ces logiciels sont construits sur un noyau numérique commun qui est, lui-même, basé sur la plate-forme de développement logiciel PELICANS [92]. Ces différents développements sont effectués au Laboratoire d’étude de l’Incendie et du développement de Méthodes pour la Simulation et les Incertitudes (LIMSI). Dans les modèles d’intérêt de chacun de ces codes, on retrouve le transport d’une température, d’une fraction massique, d’une concentration,... couplé à une dynamique d’écoulement modélisable par le système d’équations de Navier–Stokes dit “à faible nombre de Mach”. On peut donc extraire des différents modèles associés à ces codes un modèle simplifié faisant office de dénominateur commun : on s’intéressera dans ce mémoire au système (1.1)–(1.4), qui comporte une dynamique d’écoulement représentée par le système d’équations de Navier–Stokes dit à “faible nombre de Mach”, associée à une grandeur scalaire  $z$  sujette à transport et diffusion.

$$\text{Bilan de masse} \quad \frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \quad (1.1)$$

$$\text{Bilan de quantité de mouvement} \quad \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) = \mathbf{f} + \nabla \cdot \boldsymbol{\tau} - \nabla p \quad (1.2)$$

$$\text{Transport du scalaire } \phi \quad \frac{\partial \rho \phi}{\partial t} + \nabla \cdot \rho \phi \mathbf{u} = \nabla \cdot (\kappa \nabla \phi) \quad (1.3)$$

$$\text{Loi d’état} \quad \rho = \mathcal{G}(\phi) \quad (1.4)$$

où  $t$  est le temps,  $\mathbf{u}$  la vitesse du fluide,  $p$  la pression, et  $\rho$  la masse volumique du fluide. La masse volumique est donnée en fonction de la variable scalaire  $\phi$  par la loi d’état  $\mathcal{G}$ , et le coefficient de diffusion volumique  $\kappa$  est une fonction strictement positive de  $\phi$ . Le tenseur  $\boldsymbol{\tau}$  est la partie visqueuse du tenseur des contraintes :

$$\boldsymbol{\tau} = \mu (\nabla \mathbf{u} + \nabla^t \mathbf{u}) - \frac{2}{3} \mu (\nabla \cdot \mathbf{u}) \mathbf{I}, \quad (1.5)$$



où  $\mu$  est la viscosité du fluide et  $\mathbf{I}$  le tenseur identité. Pour une viscosité constante, cette relation devient :

$$\nabla \cdot \boldsymbol{\tau} = \mu \left[ \Delta \mathbf{u} + \frac{1}{3} \nabla \nabla \cdot \mathbf{u} \right]. \quad (1.6)$$

Notons que  $\rho$  est connue en fonction de la pression par la loi d'état (1.4). L'une des équations (1.1) ou (1.3) peut être vue comme une contrainte qui définit  $p$ , mais on peut également obtenir l'expression de  $\nabla \cdot \mathbf{u}$  à partir de (1.3).

On résout ces équations munies de leurs conditions initiales et aux limites sur un domaine ouvert, borné et connexe  $\Omega \subset \mathbb{R}^d$ ,  $d = 2$  ou  $d = 3$ . On suppose que les conditions aux bord sur la vitesse sont mixtes, la frontière  $\partial\Omega$  de  $\Omega$  étant constituée de  $\partial\Omega_D$  et  $\partial\Omega_N$  telles que :

$$\mathbf{u} = \mathbf{u}_D \quad \text{sur } \partial\Omega_D \quad (1.7)$$

$$(\boldsymbol{\tau}(\mathbf{u}) + p\mathbf{I})\mathbf{n} = \mathbf{g} \quad \text{sur } \partial\Omega_N \quad (1.8)$$

où  $\mathbf{u}_D$  est un champ de vitesses connu sur  $\partial\Omega_D$ ,  $\mathbf{n}$  est la normale sortante à  $\partial\Omega$  et  $\mathbf{g}$  est un terme connu. L'inconnue  $\phi$  sera généralement soumise à des contraintes physiques, comme de rester dans des bornes fixées. On peut par exemple penser à une concentration en espèce chimique qui doit rester entre 0 et 1.

La formulation à faible nombre de Mach des équations de Navier–Stokes est mathématiquement proche des équations de Navier–Stokes en régime incompressible. On applique par conséquent pour l'approximation de ce système les schémas initialement construits pour les écoulements incompressibles. Parmi ceux-ci, on distingue les méthodes de projection, introduites par Chorin et Temam [23, 104], qui sont très populaires. Elles consistent à découpler à chaque pas de temps les équations de bilan de quantité de mouvement et de bilan de masse, ce qui permet de remplacer la résolution difficile du système couplé initial par la résolution d'une succession de problèmes elliptiques plus simples. Le choix retenu dans le noyau numérique pour découpler le système d'équations est celui d'une méthode à pas fractionnaires, qui consiste en un ensemble d'étapes basé sur une méthode de projection incrémentale [54] pour le couple vitesse/pression. Celle-ci se compose d'abord d'une prédiction de la vitesse par la résolution de l'équation de bilan de quantité de mouvement dans laquelle la pression est explicitée, puis d'une étape de correction qui s'apparente à un problème de Darcy réécrit comme un problème elliptique pour l'incrément de pression  $p^{n+1} - p^n$ .

On considère une subdivision  $(t^n)_{n \in [1, N]}$  de  $(0, T)$  où  $N$  est le nombre d'intervalles de la subdivision, et telle que  $t^0 = 0$  et  $t^N = T$ . Si on pose  $\Delta t^n = t^n - t^{n-1}$  et  $\Delta t = \max_{n \in [1, N]} \Delta t^n$ , le schéma permettant de passer des variables connues au temps  $t^n$  aux variables au temps  $t^{n+1}$  se compose des étapes suivantes :

#### 1. Bilan scalaire

Cette étape correspond à la semi-discrétisation en temps de l'équation d'advection–diffusion (1.3), et permet de déterminer la valeur mise à jour du scalaire  $\phi^{n+1}$  :

$$\frac{\rho^n \phi^{n+1} - \rho^{n-1} \phi^n}{\Delta t} + \nabla \cdot (\rho^n \phi^{n+1} \mathbf{u}^n) = \nabla \cdot (\kappa \nabla \phi^{n+1}) \quad (1.9)$$

## 2. Loi d'état

Cette étape correspond à la semi-discrétisation en temps de la loi d'état (1.4), et permet de déterminer la masse volumique mise à jour  $\rho^{n+1}$  :

$$\rho^{n+1} = \mathcal{G}(\phi^{n+1}) \quad (1.10)$$

## 3. Prédiction de vitesse

Cette étape correspond à la semi-discrétisation en temps de l'équation de bilan de quantité de mouvement (1.2), et permet de déterminer une prédiction  $\tilde{\mathbf{u}}^{n+1}$  de la vitesse :

$$\frac{\rho^n \tilde{\mathbf{u}}^{n+1} - \rho^{n-1} \mathbf{u}^n}{\Delta t} + \nabla \cdot (\tilde{\mathbf{u}}^{n+1} \otimes \rho^n \mathbf{u}^n) - \nabla \cdot \tau(\tilde{\mathbf{u}}^{n+1}) + \nabla p^n = \mathbf{f}^{n+1} \quad (1.11)$$

## 4. Correction

Cette étape détermine le couple vitesse/pression mis à jour  $\mathbf{u}^{n+1}, p^{n+1}$  de manière à ce que le bilan de masse (1.1) soit vérifié.

$$\begin{cases} \frac{\rho^n \mathbf{u}^{n+1} - \rho^{n-1} \mathbf{u}^{\tilde{n}+1}}{\Delta t} + \nabla(p^{n+1} - p^n) = 0 \\ \nabla \cdot (\rho^{n+1} \mathbf{u}^{n+1}) = -\frac{\rho^{n+1} - \rho^n}{\Delta t} \end{cases} \quad (1.12)$$

Lors de l'étape de prédiction de vitesse, le bilan de masse discret est vérifié par  $\rho^n$  mais pas par  $\rho^{n+1}$ . Or, pour la construction d'une formulation stable de l'opérateur d'advection, la vérification de ce bilan de masse sera une hypothèse nécessaire. C'est pourquoi on opère à l'étape 3 un décalage d'un pas de temps en arrière de la masse volumique discrète. L'autre alternative serait d'ajouter une étape de détermination *ad hoc* de la masse volumique avant la prédiction de vitesse afin qu'elle vérifie le bilan de masse.

Les tests numériques effectués montrent une convergence d'ordre un en espace et en temps, que l'on opère ou non le décalage en temps de la masse volumique.

Le découplage des équations de bilan de quantité de mouvement et de bilan de masse par une méthode de projection présente des inconvénients. D'une part, sa convergence impose la vérification d'une condition de compatibilité *inf-sup* entre les espaces d'approximation de la vitesse et de la pression. Cette condition est difficilement atteignable en volumes finis non-structurés, mais des paires d'éléments finis permettent de la vérifier et sont donc beaucoup plus indiquées pour l'approximation du couple vitesse/pression. Parmi celles-ci, ce sont des éléments finis non conformes, les éléments de Rannacher–Turek [96], qui ont été choisis pour la discrétisation des champs de vitesse et de pression.

Pour l'approximation des équations de bilan scalaires, il est crucial dans le cadre des applications qui nous concernent que le schéma respecte les bornes physiques des quantités approchées, le dépassement de celles-ci pouvant conduire à des résultats non physiques : concentrations hors de  $[0, 1]$ , températures négatives, ... Cette condition de cohérence avec la physique se traduit mathématiquement par la vérification d'un principe du maximum. Une autre condition requise de cohérence avec la physique est la conservativité des flux discrets. Ces conditions de monotonie sont difficilement atteignables par éléments finis, mais les méthodes de discrétisation par volumes finis sont connues pour permettre de créer des schémas les vérifiant. Ainsi, dans le noyau numérique, les équations de bilan scalaires sont discrétisées par des méthodes de volumes finis.

## 1.1 Objectif de la thèse

Une caractéristique commune des différents codes de simulation cités plus haut est la nécessité de traiter des échelles spatiales très diverses. Par exemple, dans le cadre des études de sûreté incendie, la simulation des feux en milieu confiné concerne des pièces de grande taille (plusieurs centaines de mètres cube) munies de portes et de bouches de ventilation, où peuvent se trouver divers objets usuellement présents dans des installations : boîtes à gants, armoires électriques, canalisations, ... voir figure 1.1. Pour conduire des simulations précises d'incendies, la nécessité d'approcher très précisément les phénomènes de front de flamme impose également l'utilisation de mailles très fines. Ainsi, la nécessité de considérer de si différentes échelles spatiales conduit à raffiner localement le maillage, la discrétisation à l'aide d'un maillage homogène étant inenvisageable en termes de temps de calcul et d'occupation mémoire.

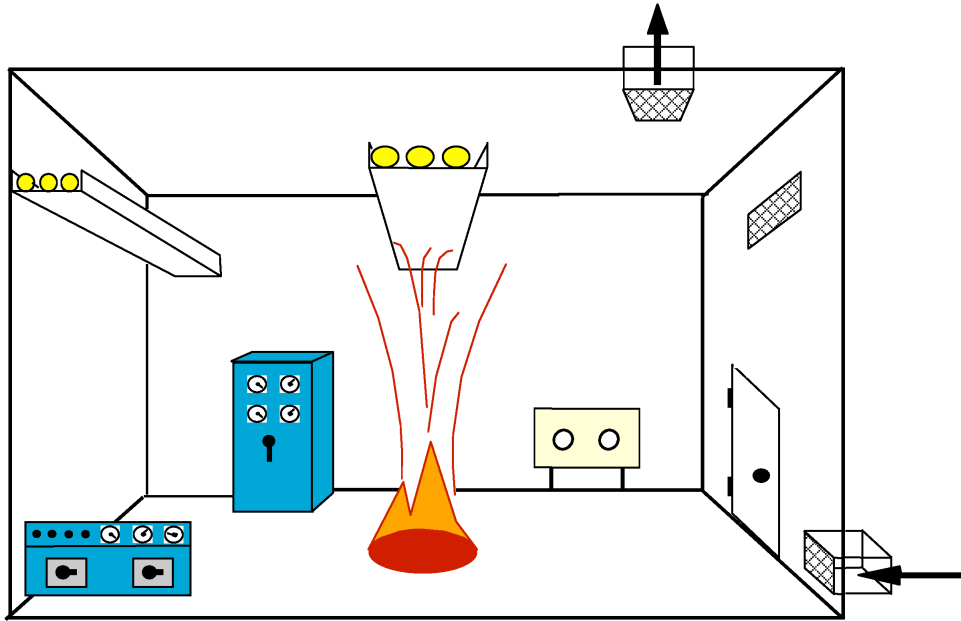


FIG. 1.1 – Schéma d'un local type

Initialement, les schémas implémentés dans le noyau numérique sont construits pour des maillages structurés, c'est-à-dire pour lesquels les faces des mailles sont parallèles aux axes. Ainsi, pour raffiner le maillage près d'un point, il faut propager ce raffinement dans toutes les directions principales du maillage afin de conserver le caractère structuré de celui-ci. La discrétisation à l'aide d'un maillage non-structuré permet de raffiner seulement localement le maillage, en réalisant une économie importante en terme de nombre de mailles, et ainsi une économie en occupation mémoire et en temps de calcul. Afin d'illustrer cette économie, on montre sur la figure 1.2 un domaine raffiné localement dans son coin inférieur gauche, en utilisant un maillage de rectangles ou un maillage de triangles.

Nous développons ici un outil complet pour la simulation en maillages non-structurés d'écoulements couplés à un phénomène de transport d'une quantité scalaire, et nous l'appliquons au modèle simplifié (1.1)–(1.4). Afin d'être conformes aux attentes correspondant à leur utilisation dans un code semi-industriel, les schémas développés pour cet outil doivent remplir des condi-

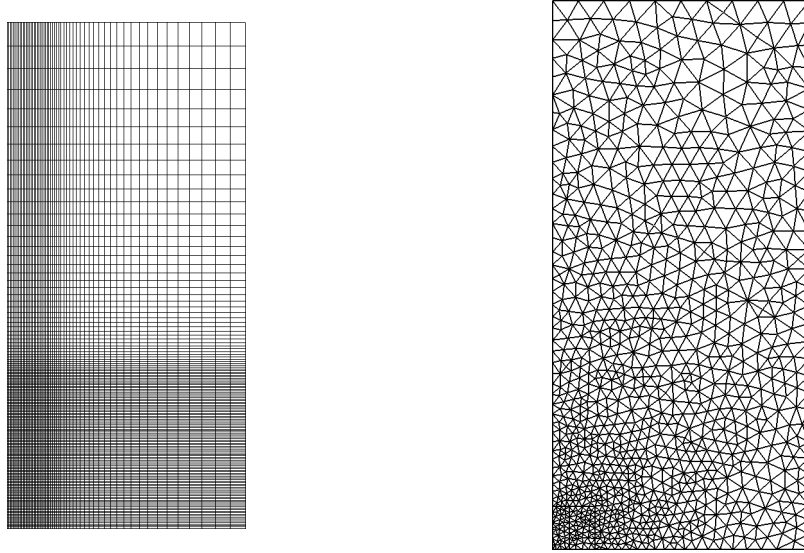


FIG. 1.2 – (a) Maillage structuré raffiné localement ; (b) Maillage non-structuré raffiné localement

tions de robustesse, de précision et d’efficacité en termes de temps de calcul et d’occupation mémoire. Cet outil doit donc, de manière générale, apporter une amélioration en termes de précision sans sacrifier les propriétés de stabilité et de consistance, tout en restant d’une complexité de mise en oeuvre raisonnable à la fois sur le plan de l’algorithmique et sur celui des structures de données. D’autre part, la génération automatisée de maillages de domaines complexes avec un raffinement localisé conduit généralement à des maillages de “mauvaise qualité”, c’est-à-dire ne possédant pas les propriétés nécessaires pour que les schémas classiques puissent se montrer stables et consistants. Afin d’être robustes vis-à-vis de ceux-ci, les schémas développés ne doivent pas être mis en échec par la présence de mailles de mauvaise qualité. Dans le pire des cas, un schéma moins précis et moins exigeant en termes de régularité de maillage doit pouvoir être localement appliqué.

## 1.2 Discrétisation spatiale des équations de la dynamique

Dans la version structurée du noyau numérique, les maillages étant composés de rectangles ou de parallélépipèdes, on emploie l’élément non-conforme de Rannacher–Turek. Dans le cadre de l’extension de ce noyau numérique aux maillages non-structurés, on emploie l’élément de Rannacher–Turek pour les quadrangles et hexaèdres, et l’élément de Crouzeix–Raviart [30] pour les triangles et tétraèdres, voir figure 1.3.

Ces éléments construisent une approximation linéaire par maille des vitesses et une approximation constante par maille des pressions. Les noeuds de vitesse sont situés aux centres des faces et valent la moyenne du champ sur celles-ci, ce qui permet un couplage aisé avec la méthode de volumes finis employée pour l’approximation des équations de transport. Je présente plus en détail les définitions et propriétés de ces éléments dans l’annexe A. L’utilisation de la méthode de Galerkin classique pour l’approximation des équations de Navier–Stokes peut conduire à des solutions présentant des oscillations non-physiques, voire à des pertes de stabilité [57]. De plus,

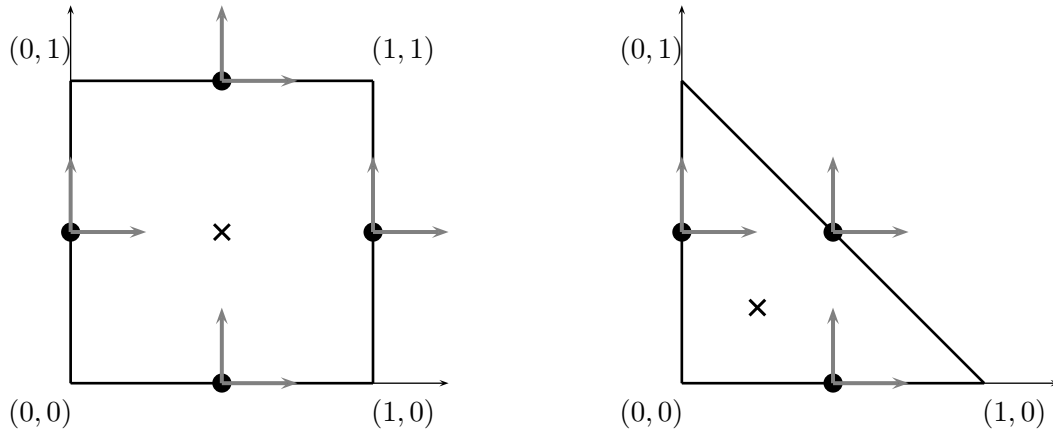


FIG. 1.3 – Noeuds de vitesse (●) et de pression (×) des éléments de Rannacher–Turek (gauche) et de Crouzeix–Raviart (droite)

ces instabilités augmentent si l'écoulement devient sensiblement advectif et/ou en présence de couches limites. Les techniques de stabilisation les plus couramment utilisées sont les techniques d'ajout de résidus décentrés en amont de l'écoulement, dont l'emploi est rendu difficile dans le cadre de discrétisations par des éléments finis non-conformes : du fait de la seule continuité faible entre deux mailles adjacentes de ces approximations, les biais classiques des méthodes d'ajout de résidus décentrés se trouvent exacerbés [3]. Plus récemment, des techniques regroupées sous la terminologie de méthodes variationnelles multi-échelles et des techniques de décentrage algébriques [78] ont été proposées dans la littérature pour pallier aux défauts des méthodes d'ajout de résidus, mais elles présentent divers inconvénients : ajout de degrés de libertés, nécessité de construire des sous-mailles, perte de la symétrie de la formulation, perte de la possibilité de *lumper* la matrice de masse. D'autres méthodes, basées sur l'ajout d'un terme de pénalisation faisant intervenir les sauts de la solution ou de son gradient aux faces, permettent de corriger les instabilités provenant de la formulation Galerkin standard ou de la non-conformité de l'espace d'approximation, sans pour autant présenter les biais des méthodes précédentes. L'idée d'une telle méthode est introduite par Douglas *et al.* dans [32], et a été reprise par Burman *et al.* pour la stabilisation des équations de Navier–Stokes [18]. Toutefois, si elles permettent de retrouver la stabilité de la formulation, elles n'assurent pas naturellement la vérification d'un principe de maximum discret sur la solution approchée. Pour cela, il faut rendre la formulation non-linéaire [16] et utiliser des stabilisations supplémentaires par pénalisation aux faces, ce qui rend l'implémentation peu adaptée dans le contexte d'un code industriel.

### 1.2.1 Construction d'une approximation stable de l'opérateur d'advection

Une approximation stable des termes advectifs basée sur un résultat obtenu en volumes finis a été développée au LIMSI spécifiquement pour la version structurée d'ISIS [49, 9]. Elle n'est pas soumise au défaut majeur des méthodes variationnelles multi-échelles qui couplent par assemblage des éléments finis non-adjacents et augmentent ainsi la largeur de bande des matrices de résolution. Elle n'est pas soumise non plus au biais classique des méthodes de décentrage algébrique qui n'offrent pas de compromis entre la perte d'un ordre de convergence en espace

et l'ajout de termes supplémentaires non-linéaires. L'objet du chapitre 2 est l'extension de cette technique originale d'approximation des termes advectifs aux mailles triangulaires/tétraédriques. L'idée de base de cette technique est l'identité suivante qui est vérifiée si le bilan de masse est conservé :

$$\frac{d}{dt} \int_{\Omega} \frac{1}{2} \rho \mathbf{u}^2 = \int_{\Omega} \left( \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) \right) \mathbf{u} \quad (1.13)$$

Le membre de droite de cette équation contient les termes instationnaires et advectifs du bilan de quantité de mouvement alors que le membre de gauche représente la variation temporelle de l'énergie cinétique. Ainsi l'égalité (1.13) assure que si les bilans de masse et de quantité de mouvements sont vérifiés, alors la variation de l'énergie cinétique est contrôlée par les forces de dissipation interne.

En multipliant (1.2) par  $\mathbf{u}$  et en intégrant sur  $\Omega$ , on obtient ainsi :

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} \rho |\mathbf{u}|^2 dx - \int_{\Omega} \tau(\mathbf{u}) : \nabla \mathbf{u} dx = \int_{\Omega} p \nabla \cdot \mathbf{u} dx \quad (1.14)$$

Cette relation fournit un contrôle de la vitesse  $\mathbf{u}$  si le terme de droite est borné, ce qui est le cas pour des écoulements incompressibles ou barotropiques compressibles.

Pour construire le schéma, il est nécessaire de définir un maillage dual du maillage éléments finis initial, voir figure 1.4. Pour ce faire, on construit autour de chaque face  $\sigma = K|L$  du maillage initial des cônes  $D_{K,\sigma}$  et  $D_{L,\sigma}$ , de base  $\sigma$  et de sommet respectivement les barycentres de  $K$  et  $L$ . On nommera diamant  $D_{\sigma}$  l'union de ces deux cônes. Dans le cas des faces frontières  $\sigma = K|\text{ext}$ , le volume  $D_{\sigma}$  sera restreint au cône  $D_{K,\sigma}$ . On notera  $\epsilon = D_{\sigma} \cap D_{\sigma'}$  la face du maillage dual adjacente aux diamants  $D_{\sigma}$  et  $D_{\sigma'}$ . L'ensemble des faces  $\sigma$  du maillage initial est noté  $\mathcal{E} = \mathcal{E}_{\text{int}} \cup \mathcal{E}_{\text{ext}}$ , où  $\mathcal{E}_{\text{int}}$  et  $\mathcal{E}_{\text{ext}}$  sont respectivement l'ensemble des faces intérieures et extérieures, et  $\mathcal{E}_{\text{ext}} = \mathcal{E}_D \cup \mathcal{E}_N$ , où  $\mathcal{E}_D$  et  $\mathcal{E}_N$  sont respectivement l'ensemble des faces extérieures où des conditions de Dirichlet et Neumann sont imposées. L'ensemble des faces d'une maille duale  $D_{\sigma}$  est noté  $\tilde{\mathcal{E}}(D_{\sigma})$ .

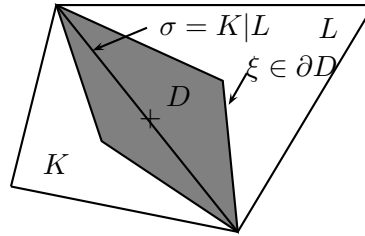


FIG. 1.4 – Construction de la maille duale (grisée)  $D$  autour de la face  $\sigma$  du maillage initial

L'approximation stable de la prédiction de vitesse (1.11) nécessite comme principale hypothèse, indépendamment de l'algorithme de marche en temps, que la masse volumique et la vitesse soient telles que le bilan de masse soit vérifié sur le maillage initial. La construction s'effectue ensuite en deux étapes :

- On suppose d'abord qu'un bilan de masse discret similaire au bilan de masse discret sur le maillage initial est vérifié sur l'ensemble des diamants du maillage dual, et l'on montre le résultat de stabilité en volumes finis suivant, analogue discret de (1.13) :

$$\sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} |D_\sigma| (Cu)_\sigma u_\sigma \geq \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} |D_\sigma| [\rho_\sigma u_\sigma^2 - \rho_\sigma^* (u_\sigma^*)^2] + \frac{1}{2} \sum_{\epsilon = D_\sigma | \text{ext} \in \mathcal{E}_N} F_{\sigma, \epsilon} u_\sigma^2 \quad (1.15)$$

où  $u_\sigma^*$ ,  $u_\sigma$ ,  $\rho_\sigma^*$ ,  $\rho_\sigma$  sont respectivement des approximations d'une composante de la vitesse et de la masse volumique à la face  $\sigma$ , au temps précédent et au temps courant.  $F_{\sigma, \epsilon}$  est une approximation du flux de masse sortant de la face  $\epsilon$  de  $D_\sigma$  :  $F_{\sigma, \epsilon} = \int_\epsilon \rho \mathbf{u} \cdot \mathbf{n}_{\sigma, \epsilon}$ , où  $\mathbf{n}_{\sigma, \epsilon}$  est la normale à  $\epsilon$  qui pointe vers  $\sigma$ .  $C$  est un opérateur sur le maillage dual, défini par :

$$(Cu)_\sigma = \frac{1}{\delta t} (\rho_\sigma u_\sigma - \rho_\sigma^* u_\sigma^*) + \frac{1}{|D_\sigma|} \sum_{\epsilon \in \mathcal{E}(D_\sigma)} F_{\sigma, \epsilon} u_\epsilon \quad (1.16)$$

$u_\epsilon$  est pour les faces intérieures, une approximation centrée des  $u_\sigma$  :  $u_\epsilon = \frac{1}{2}(u_\sigma + u_{\sigma'})$ , et pour les faces frontières, une approximation décentrée amont.

- Afin d'appliquer les techniques développées plus haut, il faut écrire une discrétisation du terme d'advection dont la structure soit identique à celle de l'opérateur volumes finis vérifiant une propriété de stabilité. En pratique, cela consiste à définir des flux aux faces des diamants tels que, si le bilan de masse discret est vérifié sur les mailles primales, alors il l'est également sur les mailles duales en diamants. Le résultat suivant décrit la méthode à suivre pour construire ces flux :

**Lemme 1** *Bilan de masse dans un sous-volume de maille*

Soit  $K \in \mathcal{T}$ ,  $\rho_K$  et  $\rho_K^*$  des réels, et  $(F_{K, \sigma})_{\sigma \in \mathcal{E}(K)}$  des approximations des flux de masse à travers les faces  $\sigma$  tels que le bilan de masse discret sur  $K$  :

$$\frac{|K|}{\delta t} (\rho_K - \rho_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_{K, \sigma} = 0 \quad (1.17)$$

soit vérifié. Soit  $\mathbf{w}^K$  un champ de quantité de mouvement sur  $K$  tel que  $\nabla \cdot \mathbf{w}^K$  est constant sur  $K$  et tel que

$$\int_K \nabla \cdot \mathbf{w}^K d\mathbf{x} = \sum_{\sigma \in \mathcal{E}(K)} \int_\Omega \mathbf{w}^K \cdot \mathbf{n}_{K, \sigma} d\sigma = \sum_{\sigma \in \mathcal{E}(K)} F_{K, \sigma} \quad (1.18)$$

Soit  $D \subset K$  de frontière  $\partial D$ , et  $\mathbf{n}_{\partial D}$  le vecteur normal à  $\partial D$  sortant de  $D$ , alors :

$$\frac{|D|}{\delta t} (\rho_K - \rho_K^*) + \int_{\partial D} \mathbf{w}^K \cdot \mathbf{n}_{\partial D} d\sigma = 0 \quad (1.19)$$

Ayant construit pour tout  $K \in \mathcal{T}$  un champ  $\mathbf{w}$  tel que décrit dans ce résultat, on obtiendra les flux aux faces du demi-diamant  $D_{K, \sigma}$  par intégration de  $\mathbf{w}^K \cdot \mathbf{n}$  sur celles-ci. On décrira ainsi dans le chapitre 2 la définition des flux aux faces des diamants permettant de vérifier le bilan de masse discret sur ceux-ci pour les éléments de Crouzeix–Raviart, et pour les éléments de Rannacher–Turek en coordonnées cartésiennes et axisymétriques.

La discrétisation complète de l'équation de bilan de quantité de mouvement 1.2 est enfin obtenue en utilisant l'opérateur de convection défini ici, et une technique d'éléments finis classique

pour les termes de viscosité et de gradient de pression. Le bilan de quantité de mouvement discret s'écrit alors :

$$\begin{aligned} \forall \sigma \in \mathcal{E} \setminus \mathcal{E}_D, \quad 1 \leq i \leq d, \\ \frac{1}{\delta t} (\rho_\sigma(\mathbf{u}^{(i)})_\sigma - \rho_\sigma^*(\mathbf{u}^{(i)})_\sigma^*) + \frac{1}{|D_\sigma|} \sum_{\epsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\epsilon}(\mathbf{u}^{(i)})_\epsilon + a_d(\mathbf{u}, \varphi_\sigma^{(i)}) \\ - \sum_{K \in \mathcal{T}} \int_K p^n \nabla \cdot \varphi_\sigma^{(i)} d\mathbf{x} = \int_\Omega \mathbf{f} \cdot \varphi_\sigma^{(i)} d\mathbf{x} + \int_{\partial\Omega_N} \mathbf{g} \cdot \varphi_\sigma^{(i)} d\sigma \end{aligned} \quad (1.20)$$

où  $\mathbf{u}^{(i)}$  est la  $i^{\text{ème}}$  composante de la vitesse, et  $\varphi_\sigma^{(i)} = \varphi_\sigma \mathbf{e}^{(i)}$  est la fonction de forme associée à ce degré de liberté.

### 1.2.2 Expérimentations numériques

#### Écoulements incompressibles

Ces écoulements sont régis par le système d'équations (1.21)–(1.22).

$$\rho \left( \frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u} \otimes \mathbf{u}) \right) + \nabla p - \nabla \cdot \tau(\mathbf{u}) = \mathbf{f} \quad (1.21)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (1.22)$$

#### Tourbillons de Green–Taylor

Les tourbillons de Green–Taylor sont une solution manufacturée classique dans  $[0, \frac{7}{4}]^2 \times [0, T]$  des équations de Navier–Stokes pour l'écoulement d'un fluide newtonien incompressible. L'aspect de la solution exacte est représenté sur la figure 1.5. À l'aide de cette solution, on vérifie qu'avec des maillages de quadrangles structurés ou de triangles non structurés, on retrouve les ordres de convergence en espace attendus, à savoir respectivement des ordres 2 et 1 pour les erreurs en norme  $L^2$  et  $H^1$  sur la vitesse, et l'ordre 1 pour l'erreur en norme  $L^2$  sur la pression.

#### Écoulement derrière une marche descendante

L'écoulement derrière une marche descendante est une situation d'écoulement d'un fluide newtonien incompressible souvent étudiée dans la littérature, qui décrit le comportement d'un jet provenant d'un tube droit où le régime permanent est établi et se déversant dans un domaine d'expansion de plus grand diamètre. Ces conditions conduisent à un écoulement présentant des zones de recirculation dont le nombre et la longueur dépendent du nombre de Reynolds de l'écoulement.

La dépendance de la longueur de la première zone de recirculation, qui se forme au pied de la marche, au nombre de Reynolds, permet d'évaluer la précision du schéma numérique utilisé et de comparer ses résultats avec des valeurs de référence. On étudie donc cet écoulement en deux dimensions avec la géométrie définie dans [6] et l'on compare les longueurs de recirculation obtenues avec celles de Chiang *et al.* dans [22]. On observe un très bon accord des résultats, voir figure 1.7.



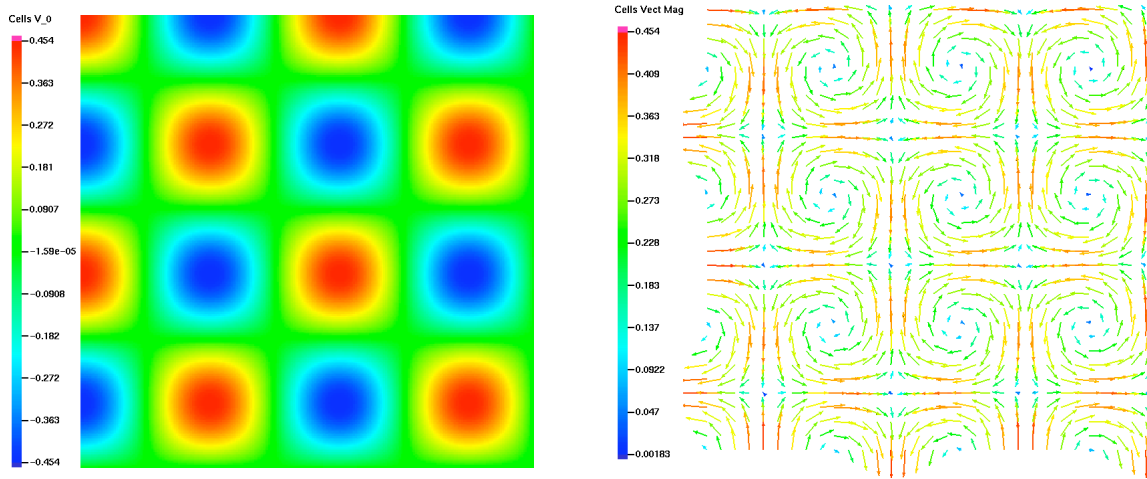


FIG. 1.5 – Amplitude de vitesse (à gauche) et vecteurs vitesse (à droite) pour la solution des tourbillons de Green–Taylor

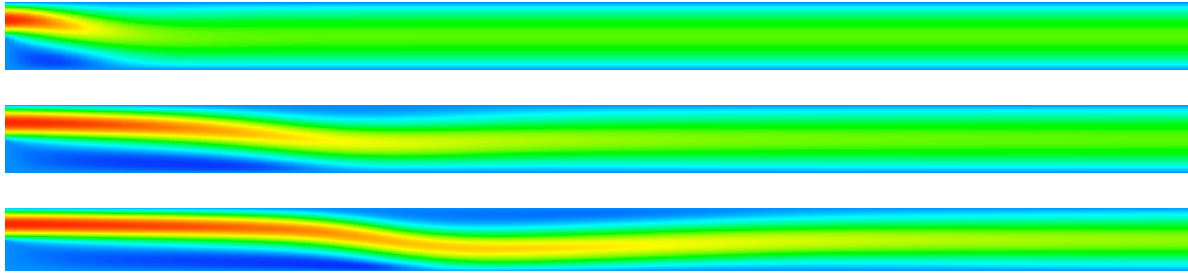


FIG. 1.6 – Écoulement derrière une marche descendante en 2D avec  $\mathbf{Re} = 100$  (en haut),  $\mathbf{Re} = 500$  (au milieu),  $\mathbf{Re} = 1000$  (en bas)

### Écoulement derrière un cylindre

L'écoulement derrière un cylindre est également souvent choisi pour évaluer la précision des méthodes numériques, grâce aux différentes grandeurs physiques caractéristiques qu'il est possible d'associer à une solution donnée : coefficients de portance et de traînée, différence de pression entre les extrémités avant et arrière du cylindre, nombre de Strouhal. Des configurations de référence en 2 et 3 dimensions, les résultats obtenus par différentes équipes, et des fourchettes de valeurs de référence sont présentés dans [100]. On étudie ici l'écoulement à  $\mathbf{Re} = 100$  en 2 dimensions et l'écoulement à  $\mathbf{Re} = 20$  en 3 dimensions et comparons les résultats obtenus avec ceux de Turek *et al.*, voir les tableaux 1.1 et 1.2.

### Écoulements à masse volumique variable

Ces écoulements sont régis par le système d'équations (1.23)–(1.25).

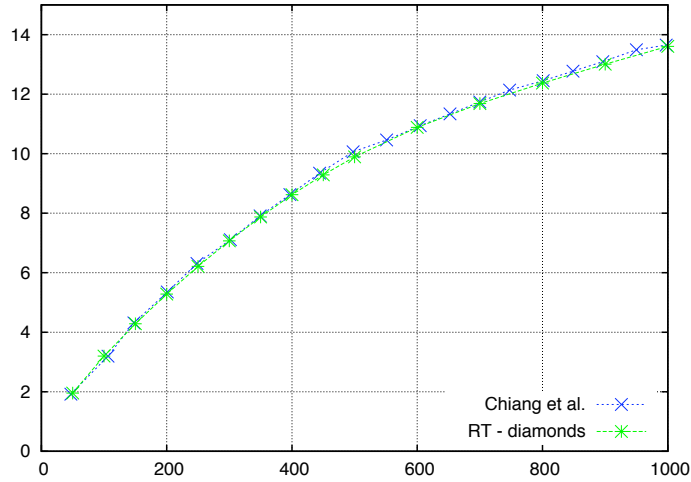


FIG. 1.7 – Ecoulement derrière une marche descendante : longueurs de recirculation pour différentes valeurs du nombre de Reynolds. Comparaison avec les résultats de Chiang *et al.* [22]

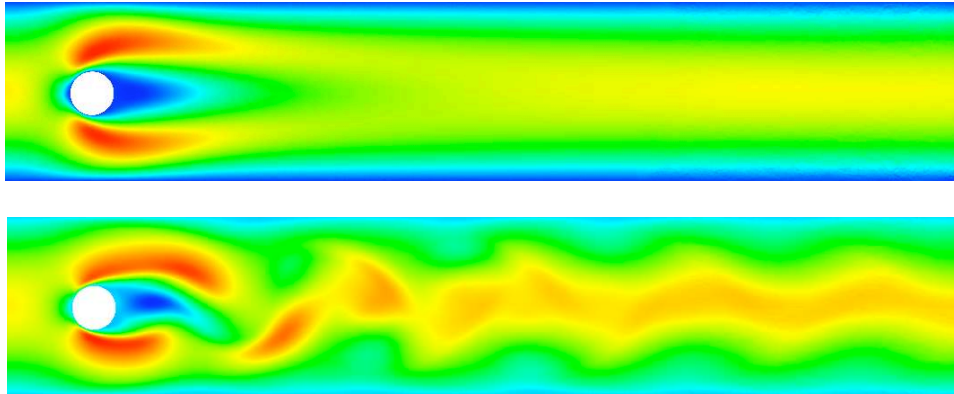


FIG. 1.8 – Ecoulement derrière un cylindre en 2D à  $\mathbf{Re} = 20$  (en haut) et  $\mathbf{Re} = 100$  (en bas)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) \quad (1.23)$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \nabla \cdot \tau(\mathbf{u}) = \mathbf{f} \quad (1.24)$$

$$\frac{\partial \rho y}{\partial t} + \nabla \cdot (\rho y \mathbf{u}) - \lambda \Delta y = g \quad (1.25)$$

### Solution manufacturée

Afin de vérifier que l'utilisation de l'opérateur d'advection construit permet toujours d'obtenir les ordres de convergence attendus en vitesse et en pression, on étudie la convergence du schéma vers une solution analytique de (1.23)–(1.25).

mesh	space unks	time unks	$c_{D_{\max}}$	$c_{L_{\max}}$	St	$\Delta P$
2D#1	294 882	660	<b>3.2395</b>	0.9334	<b>0.3030</b>	2.4533
2D#2	352 722	664	3.2511	<b>0.9983</b>	<b>0.3012</b>	<b>2.4961</b>
Turek 1	167 232	188	3.2498	<b>1.0081</b>	0.2927	2.4410
Turek 2	667 264	612	<b>3.2314</b>	<b>0.9999</b>	<b>0.2973</b>	<b>2.4707</b>
Exact range			3.22 – 3.24	0.99 – 1.01	0.295 – 0.305	2.46 – 2.50

TAB. 1.1 – Écoulement derrière un cylindre 2D,  $\mathbf{Re} = 100$  : coefficients maximaux de traînée  $c_{D_{\max}}$  et de portance  $c_{L_{\max}}$ , nombre de Strouhal et différence de pression entre les extrémités du cylindre. Comparaison avec les résultats de Turek *et al.*

mesh	space unks	$c_D$	$c_L$	$\Delta P$
3D#1	2 271 870	<b>6.175</b>	<b>8.14E-03</b>	<b>0.1673</b>
Turek 1	98 128	5.8431	6.10E-03	0.1482
Turek 2	771 392	5.9731	5.90E-03	0.1605
Turek 3	6 116 608	<b>6.1043</b>	7.90E-03	<b>0.1672</b>
Exact range		6.05 – 6.25	0.008 – 0.01	0.165 – 0.175

TAB. 1.2 – Écoulement derrière un cylindre 3D,  $\mathbf{Re} = 20$  : coefficients de traînée  $c_D$  et de portance  $c_L$ , différence de pression entre les extrémités du cylindre. Comparaison avec les résultats de Turek *et al.*

Une convergence en temps d'ordre un, et des convergences en espace d'ordres respectivement deux et un pour la vitesse et la pression sont retrouvées pour des maillages structurés, et également sur des maillages de triangles, à condition que le terme de diffusion soit écrit sous forme coercive  $\nabla \cdot (\tau(\mathbf{u})) = \Delta \mathbf{u} + \nabla(\nabla \cdot \mathbf{u})$ . Si ce n'est pas le cas, sur des maillages généraux on observe de faibles ordres de convergence en espace, ce qui est probablement dû à l'absence de vérification du lemme de Korn. Cette situation nécessite l'application d'une stabilisation supplémentaire pour respecter le lemme de Korn.

### Instabilité de Rayleigh–Taylor

Cette simulation représente l'instabilité de l'interface entre deux fluides de masses volumiques différentes. Dans le cas de l'instabilité de Rayleigh–Taylor, deux couches de fluides non-miscibles sont placées l'une en dessous de l'autre, le fluide lourd étant au dessus du fluide léger. Les deux fluides étant soumis à la gravité, l'équilibre de l'interface est instable, et l'apparition de perturbations déclenche la descente de fluide lourd, qui s'accompagne de la remontée de fluide léger. On présente ici la simulation de cet écoulement en suivant les caractéristiques de fluide, de domaine de calcul et de maillage spécifiés dans [45]. Les résultats obtenus semblent proches de ceux de cette référence.

### 1.3 Discrétisation spatiale du transport diffusif de scalaire

On cherche au chapitre 3 à approcher numériquement la quantité scalaire  $\phi(\mathbf{x}, t)$  définie sur  $\Omega \times [0, T]$  et satisfaisant l'équation d'advection diffusion suivante :

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \nabla \cdot [\mathbf{u}(\mathbf{x})\phi(\mathbf{x}, t) - \kappa \nabla \phi(\mathbf{x}, t)] = f(\mathbf{x}, t) \quad (1.26)$$

avec conditions aux limites et initiale.

Pour satisfaire les bornes physiques, l'approximation est réalisée par volumes finis. Dans l'objectif d'étendre le noyau numérique du code ISIS à des discrétisations spatiales non-structurées, nous sommes confrontés à deux problèmes :

- L'implémentation de schémas de type MUSCL dans le cas de maillages non-structurés,
- L'implémentation de schémas volumes finis sur maillage quelconque, ne vérifiant pas forcément la condition d'orthogonalité requise pour la consistance des flux dans la méthode des volumes finis classiques.

L'objectif du chapitre 3 est de proposer un schéma volumes finis complet pour l'approximation de l'équation d'advection–diffusion sur des maillages non-structurés généraux, couplant deux techniques nouvelles d'approximation des flux convectifs et diffusifs.

#### 1.3.1 Approximation des flux convectifs

Pour approcher le flux convectif, on cherche à définir une approximation de la quantité  $\phi$  à la face  $\sigma$ . Dans le cas d'un flux convectif linéaire, on peut montrer la convergence du schéma décentré amont (*upwind*) [28], mais il est seulement d'ordre 1. Dans le cas d'un flux convectif non linéaire, les conditions de conservativité et de consistance des flux numériques ne suffisent pas à assurer la convergence des schémas [64], il faut y ajouter une condition de monotonie. On peut citer comme exemples de flux monotones les flux de Godunov [55] et de Lax–Friedrichs. Osher introduit plus tard une classe de flux numériques monotones, les E-flux [90]. En 1973, Lax [80] observe qu'en présence de discontinuités dans la solution, de l'information est perdue, et la variation totale  $\int \left| \frac{\partial \phi}{\partial \mathbf{x}} \right| d\mathbf{x}$  décroît. Cette observation motive la définition par Harten d'une contrainte de diminution de la variation totale (TVD, *Total Variation Diminishing*), [61]. Les schémas de type MUSCL (*Monotone Upwind Scheme for Conservation Laws*) introduits par Van Leer en 1979 [108] entrent dans la famille des discrétisations TVD et étendent le nombre de points utilisés pour la reconstruction : on calcule ainsi des pentes amont et aval, limitées pour préserver la stabilité du schéma. Ces techniques permettent d'atteindre des ordres de convergence en espace plus élevés. Goodman et Leveque démontrent en 1985 dans [56] que l'extension d'un schéma TVD à une dimension supérieure à 1 est au mieux d'ordre 1. Pour pallier à ce défaut inhérent à la contrainte TVD, Harten introduit des contraintes plus faibles sur la variation totale de la solution et propose les schémas ENO (*Essentially Non-Oscillatory*) [65, 66, 62, 63]. En 1986, Jameson et Lax généralisent la condition TVD à des schémas multi-points [72], puis Jameson introduit en 1995 les schémas LED [71]. Un schéma LED assure que les maxima locaux n'augmentent pas, et que les minima locaux ne diminuent pas : c'est une condition plus faible que le principe du maximum local. Un schéma vérifiant le critère LED vérifie le critère TVD, et donc est au plus d'ordre 1 en 2 et 3 dimensions. Le théorème de Goodman et Leveque a encouragé la définition de schémas partant de la considération du principe du maximum et conservant la monotonie et la positivité :

- Schémas à coefficients positifs sur maillages structurés (Spekreijse [102]) ou non structurés (Cournède *et al.*, INRIA [29])
- Schémas à reconstruction linéaire sur maillage non structuré. Ces schémas introduisent la notion de restriction de la reconstruction pour rester dans une zone où le principe du maximum est préservé. (Barth *et al.* [12])
- Opérateurs de reconstruction linéaire sur des simplexes (Barth *et al.* [12], Hubbard [69]) ou des mailles quelconques
- Opérateurs de reconstruction d'ordre  $p$  sur maillage non structuré (Barth *et al.* [11])

On choisit ici d'utiliser une méthode de type MUSCL basée sur la reconstruction linéaire des gradients sur des maillages non-structurés. Par la suite, on supposera que le flux à la face  $\sigma = K|L$ ,  $\mathbf{u} \cdot \mathbf{n}_{KL}$  est positif, où  $\mathbf{n}_{KL}$  est la normale unitaire sortante à  $\sigma$ . Dans les méthodes de type MUSCL, le flux convectif à la face  $\sigma$  d'un volume de contrôle  $K$  est approché à la face  $\sigma = K \cap L$  à l'aide d'une approche amont : la reconstruction s'écrit respectivement depuis le volume de contrôle  $K$  si  $\mathbf{u} \cdot \mathbf{n}_{KL} > 0$  (on dira alors que le flux de masse "vient de"  $K$ ), ou depuis le volume de contrôle  $L$  dans le cas contraire. Si le flux de masse vient de  $K$ , le flux convectif s'écrit  $F_{KL}(\phi_{KL}, \phi_{LK}) = |\sigma| \phi_{KL} \mathbf{u} \cdot \mathbf{n}_{KL}$ . Si la solution est suffisamment régulière,  $\phi_{KL}$  est une approximation linéaire de la solution, de telle sorte que l'on atteint une précision d'ordre 2 en espace. Si la solution présente de forts gradients, cette reconstruction linéaire doit être limitée afin de prévenir l'apparition d'oscillations parasites de la solution approchée. La forme générale de la reconstruction  $\phi_{KL}$  est donc  $\phi_{KL} = \phi_K + s_{KL} |\mathbf{x}_K \mathbf{x}_{KL}|$ , où  $s_{KL}$  est une pente constante, et  $\mathbf{x}_{KL}$  un point de  $\sigma$  à définir. Le choix naturel pour la position de  $\mathbf{x}_{KL}$  est le point d'intersection entre le segment liant les centres  $\mathbf{x}_K$  et  $\mathbf{x}_L$  et la face  $\sigma$  : les points  $\mathbf{x}_K$ ,  $\mathbf{x}_{KL}$  et  $\mathbf{x}_L$  étant alignés, la pente s'écrit naturellement comme la différence entre les valeurs du champ aux centres de  $K$  et  $L$  divisée par la distance qui les sépare :

$$s_{KL} = \frac{\phi_L - \phi_K}{|\mathbf{x}_K \mathbf{x}_L|} \quad (1.27)$$

Détaillons d'abord la construction d'un schéma MUSCL dans la configuration simple d'un maillage structuré de quadrangles. La figure 1.9 représente les mailles et centres de mailles pris en compte pour la reconstruction du flux convectif à la face  $\sigma$ . Afin d'effectuer cette reconstruction, on détermine une pente sur la maille  $i-1$  en amont du flux de masse par rapport à  $\sigma$ . Pour cela, on définit une pente centrée  $p_\sigma^+ = \frac{\phi(\mathbf{x}_i) - \phi(\mathbf{x}_{i-1})}{d(\mathbf{x}_{i-1}, \mathbf{x}_i)}$  et une pente amont  $p_\sigma^- = \frac{\phi(\mathbf{x}_{i-1}) - \phi(\mathbf{x}_{i-2})}{d(\mathbf{x}_{i-2}, \mathbf{x}_{i-1})}$ . La pente  $p_\sigma$  est ensuite définie à l'aide d'un limiteur  $\Phi : p_\sigma = p_\sigma^- = \Phi\left(\frac{p_\sigma^+}{p_\sigma^-}\right)$ , et la reconstruction  $\phi_\sigma = \phi_{KL}$  s'écrit :

$$\phi_{KL} = \phi_K + p_\sigma d(\mathbf{x}_{i-1}, \mathbf{x}_\sigma) \quad (1.28)$$

où  $\mathbf{x}_\sigma$  est le centre de la face  $\sigma$ .

Pour les maillages non-structurés généraux, l'intersection entre une face et le segment liant les centres des mailles voisines n'existe pas nécessairement, comme illustré sur la figure 1.10. Dans cette situation, il est nécessaire de définir une variante de la méthode précédente utilisant un autre point, par exemple le centre de la face.

Les propriétés du maillage contraignent donc dans le choix d'une variante de MUSCL. Dans le chapitre 3, nous proposons un choix de techniques inspirées de la *méthode multipentes* définie

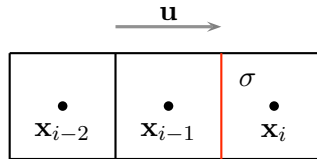
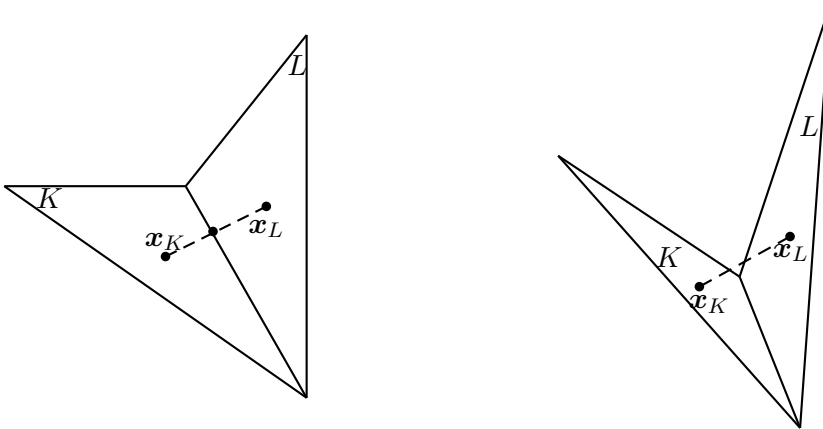


FIG. 1.9 – Motif pour la reconstruction des flux convectifs en MUSCL structuré

FIG. 1.10 – Triangles vérifiant (gauche) ou non (droite) l'existence d'une intersection entre le segment  $[x_K x_L]$  et la face  $\sigma$ 

par Buffard et Clain [15] et Clauzon [27] possédant des conditions sur le maillage et des propriétés de stabilité différentes, et permettant d'utiliser des maillages non-structurés généraux conformes ou non. Ces méthodes peuvent être classées en fonction des maillages avec lesquels elles sont compatibles :

- Pour les maillages conformes de triangles ou tétraèdres vérifiant la condition d'existence de points d'intersections entre  $[x_K x_L]$  et  $\sigma$ , cette intersection peut être choisie comme point de colocation, et une pente naturelle peut être définie. On l'appellera *méthode aux intersections*.
- Pour les maillages de triangles et tétraèdres conformes ne vérifiant pas cette condition, on choisit comme points de colocation les centres des faces, et l'on doit définir la pente comme une extrapolation de la pente naturelle. On appellera cette méthode *méthode aux centres*.
- Dans le cas de maillages possédant des volumes de contrôle avec plus de  $d + 1$  faces, c'est-à-dire pour les maillages non-conformes ou les maillages de quadrangles ou d'hexaèdres, les deux schémas ci-dessus ne peuvent plus s'appliquer. Nous décrirons les modifications nécessaires pour les adapter à de tels maillages.

Il reste une condition nécessaire à l'application de tous ces schémas : pour tout volume de contrôle  $K$ , son centre  $x_K$  (qui sera, pour toutes les méthodes présentées, pris au barycentre de  $K$ ) doit se trouver strictement à l'intérieur du polygone/polyèdre formé par les centres des volumes de contrôle voisins.

**Remarque 1** *Conditions géométriques et maillages 3D*

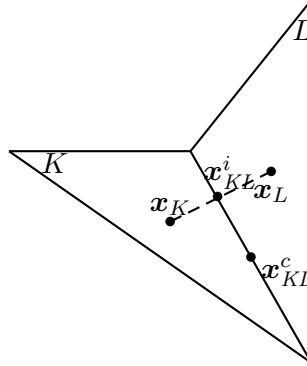


FIG. 1.11 – Notations pour la définition des méthodes MUSCL multipentes

Les conditions d'intersection, voire de stricte inclusion de  $\mathbf{x}_K$  dans le polyèdre des centres des volumes de contrôle voisins s'avèrent très difficile à respecter lors de la génération de maillages 3D non-structurés. A notre connaissance, aucun des logiciels de génération de maillages tridimensionnels que nous avons utilisés (Tetgen, Netgen, encapsulés dans GMSH [52]) ne parvient à créer des maillages vérifiant cette condition à tous coups. Le lecteur souhaitant implémenter ces méthodes est donc amené à utiliser la méthode aux centres lorsque la condition d'intersection n'est pas vérifiée, et une méthode upwind d'ordre 1 lorsque  $\mathbf{x}_K$  n'est pas strictement inclus dans le polyèdre des centres des volumes de contrôle voisins.

### 1.3.2 Approximation des flux diffusifs

L'approximation classique du flux diffusif  $-\nabla\phi \cdot \mathbf{n}_{KL}$  à travers une arête  $\sigma = K|L$  est la suivante :

$$G_{KL}(z_K, z_L) = \frac{z_L - z_K}{d_{K|L}} \quad (1.29)$$

où  $d_{K|L}$  est la distance entre  $\mathbf{x}_K$  et  $\mathbf{x}_L$  signée par  $\mathbf{n}_{KL}$ .

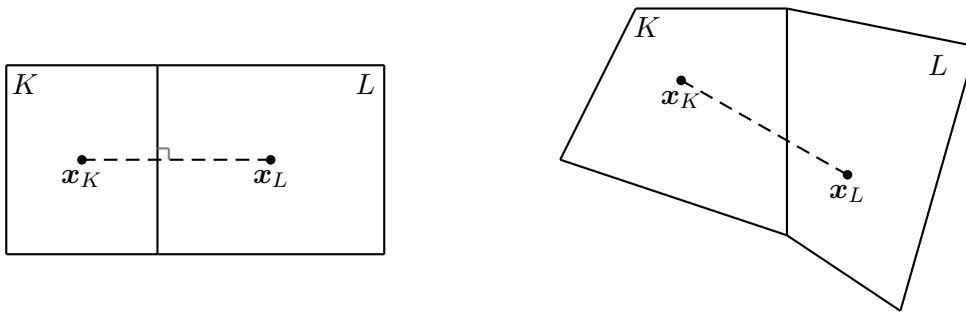


FIG. 1.12 – Maillages admissible (gauche) et non admissible (droite) de quadrangles pour l'approximation des termes diffusifs

Pour que le schéma qui en découle soit stable et d'ordre 2, il faut que, pour chaque face  $\sigma$  du maillage, le segment joignant les centres de mailles  $\mathbf{x}_K$  et  $\mathbf{x}_L$  soit perpendiculaire à  $\sigma$

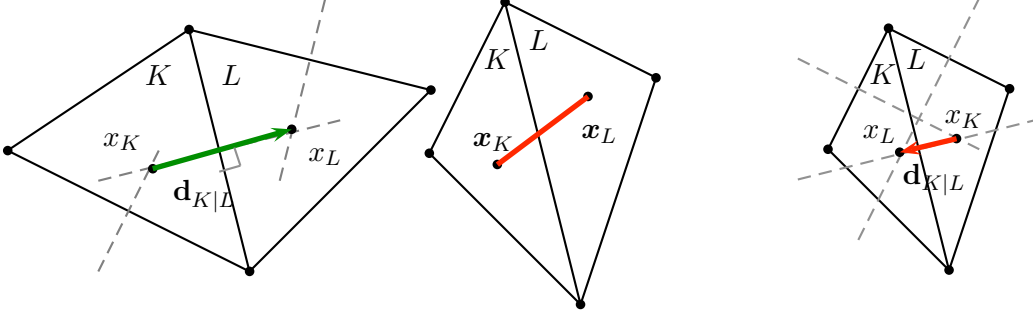


FIG. 1.13 – (a) Orthogonalité et  $d_{K|L} > 0$ ; (b) Non-orthogonalité; (c) Inversion des centres :  $d_{K|L} < 0$

(propriété dite d'orthogonalité du maillage), et que la distance  $d_{K|L}$  entre  $x_K$  et  $x_L$  signée par la normale  $\mathbf{n}_{KL}$  à la face  $\sigma$  sortant de  $K$  soit positive. Une telle configuration est représentée sur les figures 1.12 et 1.13(a). Si l'orthogonalité n'est pas vérifiée (voir figure 1.13(b)), le schéma reste stable mais n'est plus consistant. Cela conduit à utiliser comme centres de mailles non pas les barycentres, mais les centres des cercles circonscrits, qui sont les points de concours des médiatrices des faces. Ainsi, l'orthogonalité reste assurée, mais les centres de mailles peuvent se trouver hors de celles-ci, et conduire à des situations où la positivité des  $d_{K|L}$  n'est plus vérifiée (voir figure 1.13(c)). On ne peut alors plus démontrer la convergence du schéma vers la solution attendue. Ces contraintes suggèrent la définition d'une notion de maillage admissible : on qualifiera d'admissible un maillage tel que l'approximation classique du flux diffusif soit consistante. Eymard *et al.* définissent dans [35] une notion d'admissibilité pour les maillages généraux :

**Définition 1** *Maillage admissible pour l'approximation des termes diffusifs*

*Ces éléments vérifient les propriétés suivantes :*

1.  $\overline{\bigcup_{K \in \mathcal{T}} K} = \bar{\Omega}$
2.  $\forall K \in \mathcal{T}, \exists \mathcal{E}_K \subset \mathcal{E} : \partial K = \bar{K} \setminus K = \bigcup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$ , et  $\mathcal{E} = \bigcup_{K \in \mathcal{T}} \mathcal{E}_K$
3.  $\forall (K, L) \in \mathcal{T}^2, K \neq L$ , soit  $\bar{K} \cap \bar{L}$  est vide, soit égal à  $\bar{\sigma}$  avec  $\sigma \in \mathcal{E}$ , alors noté  $K|L$ .
4. La famille de points  $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$  est telle que  $\forall K \in \mathcal{T}, x_K \in \bar{K}$ . Si  $\bar{K} \cap \bar{L} = \bar{\sigma}$ ,  $x_K \neq x_L$  et la ligne droite de  $x_K$  à  $x_L$  est perpendiculaire à  $\sigma = K|L$ .
5. Pour tout  $\sigma \in \mathcal{E}$  tel que  $\sigma \subset \partial\Omega$ , soit  $K$  le volume de contrôle tel que  $\sigma \in \mathcal{E}_K$ . Si  $x_K \notin \sigma$  et si on note  $\mathcal{D}_{K,\sigma}$  la droite passant par  $x_K$  et perpendiculaire à  $\sigma$ , alors l'intersection  $\mathcal{D}_{K,\sigma} \cap \sigma$  est non vide et notée  $y_\sigma$ .

Si  $\mathcal{T}$  est une famille d'ouverts triangulaires disjoints recouvrant  $\Omega$  et tels que deux volumes de contrôle ayant une face commune ont deux sommets communs, et si les angles de tous les triangles de  $\mathcal{T}$  ont une mesure inférieure à  $\frac{\pi}{2}$ , alors les médiatrices des faces de ces triangles se coupent à l'intérieur de ceux-ci et on peut les définir comme centres de mailles et ainsi obtenir un maillage admissible (Fig. 1.15) au sens de la définition précédente. Cependant, cette contrainte



de triangulation (qu'on nommera *triangulation aigue*) est difficile à obtenir en pratique, si ce n'est en découpant le domaine en carrés eux-mêmes découpés en triangles aigus suivant un motif prédéfini (Fig. 1.14), ce qui présente peu d'intérêt dans la pratique : si le domaine de calcul peut être aisément découpé en carrés, on ne s'intéressera généralement pas à le recouvrir d'un maillage non structuré.

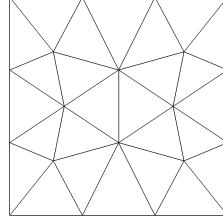


FIG. 1.14 – Découpage d'un carré en 26 triangles aigus

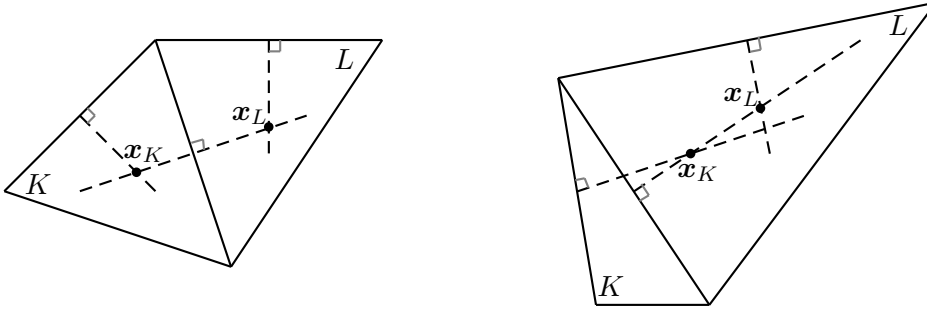


FIG. 1.15 – Maillages admissible (gauche) et non admissible (droite) de triangles pour l'approximation des termes diffusifs

On peut toutefois relaxer cette condition restrictive de triangulation aigue. La condition de Delaunay impose seulement que la fermeture du cercle circonscrit à un triangle n'englobe pas un de ses voisins. Par conséquent, un  $\mathbf{x}_K$  peut se trouver dans la maille  $L$  voisine de  $K$  tant que la distance signée  $d_{K|L}$  reste positive, ce qui permet de préserver la consistance du flux diffusif numérique (Fig. 1.16). On peut donc considérer les triangulations de Delaunay comme des maillages admissibles pour la formulation à deux points du flux diffusif (1.29).

### Diffusion pure sur un maillage non admissible

Géométriquement, la perte de la positivité des transmissibilités est provoquée par l'inversion des centres de deux mailles voisines ( $d_{K,L} < 0$ ), et pour les faces frontières  $\sigma \in \mathcal{E}_{\text{ext}}$ , par des centres de mailles hors de  $\Omega$  ( $d_{K,\sigma} < 0$ ). Les mailleurs offrant de générer des triangulations de Delaunay parviennent, dans le meilleur des cas :

- en 2D, à créer des maillages préservant la positivité de tous les  $d_{K,L}$  et  $d_{K,\sigma}, \sigma \in \mathcal{E}_{\text{ext}}$ ,
- en 3D, à créer des maillages préservant la positivité de tous les  $d_{K,L}$ , mais pas de tous les  $d_{K,\sigma}, \mathcal{E}_{\text{ext}}$ .

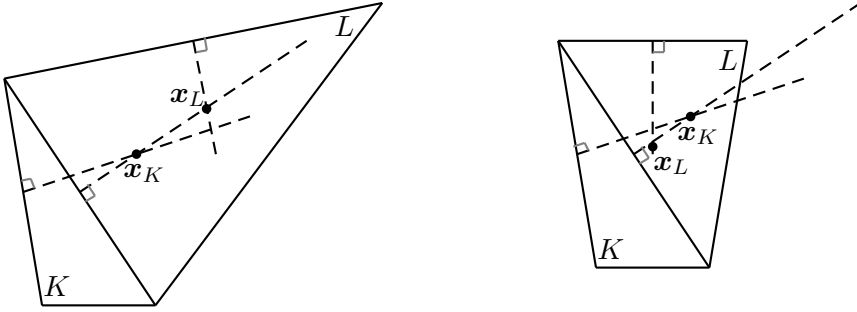


FIG. 1.16 – Maillages de triangles, vérifiant (gauche) ou non (droite) le critère de Delaunay

Lorsque les domaines ont des géométries complexes, ces objectifs sont encore plus difficiles à atteindre. Pour illustrer l'effet d'un maillage ne vérifiant pas la condition d'admissibilité sur l'approximation classique des flux diffusifs, considérons le problème de diffusion pure ( $\mathbf{u} = 0$ ) :

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} - \nabla \cdot (\kappa \nabla \phi(\mathbf{x}, t)) = f(\mathbf{x}, t), \quad \mathbf{x} \in \Omega, \quad t \in (0, T) \quad (1.30)$$

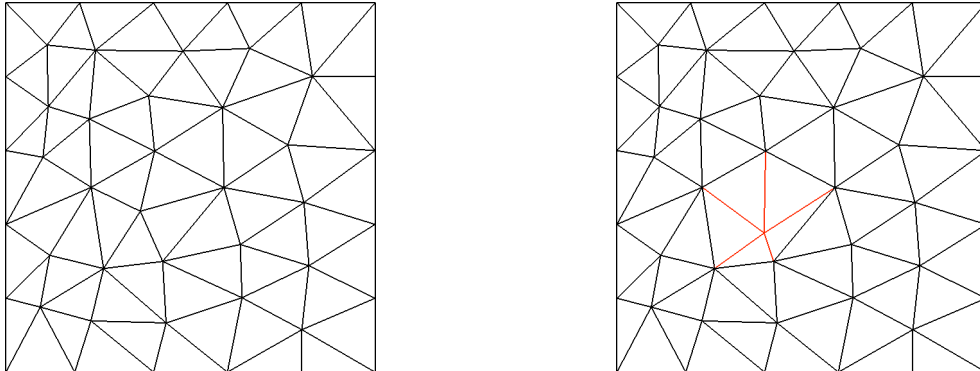
avec le temps final  $T = 1s$ . Munissons le problème d'une condition de Dirichlet homogène sur toute la frontière  $\partial\Omega$ , et de la condition initiale :

$$\phi(\mathbf{x}, t = 0) = \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{\kappa}\right), \quad \mathbf{x} \in \Omega \quad (1.31)$$

La solution du problème est :

$$\phi(\mathbf{x}, t) = \frac{\exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{\kappa(4t + 1)}\right)}{4t + 1}, \quad \mathbf{x} \in \Omega \quad (1.32)$$

avec  $\kappa = 0.01$  et  $(x_0, y_0) = (0.5, 0.5)$ . Pour étudier l'influence d'une distance intercentres  $d_{K|L}$  négative sur le schéma volumes finis classique, on crée à l'aide du mailleur EMC2 une triangulation de  $\Omega$ , qui vérifie initialement le critère de Delaunay (voir figure 1.17(a)). Le champ

FIG. 1.17 – (a) Maillage initial vérifiant le critère de Delaunay; (b) Maillage modifié avec un seul  $d_{K|L} < 0$

obtenu sur ce maillage à  $t = 1s$  avec un pas de temps  $\Delta t = 0.01s$  est représenté sur la figure 1.18.

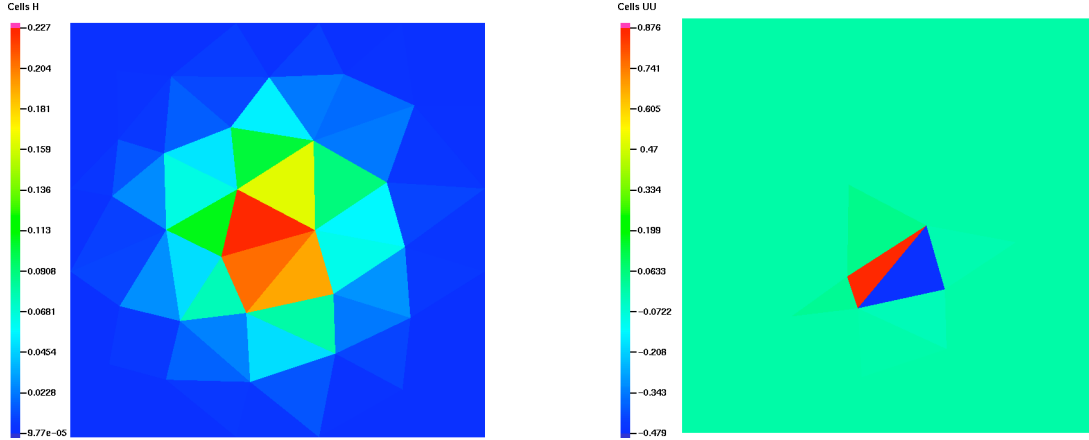


FIG. 1.18 – (a) Champ obtenu avec le maillage Delaunay,  $\Delta t = 0.01s$ ; (b) Champ obtenu à  $\Delta t = 0.01s$  sur le maillage non admissible

Puis on déplace un sommet de ce maillage pour en obtenir un nouveau où une seule face  $\sigma = K|L$  est telle que  $d_{K,L} < 0$  (voir figure 1.17(b)). Sur ce nouveau maillage, non admissible, avec un pas de temps  $\Delta t = 0.01s$ , on converge vers une solution où  $u_K$  et  $u_L$  prennent des valeurs très éloignées de la solution et de signes opposés (Fig. 1.18(b)). Dans le premier cas, on observe une perte de coercivité, le vecteur des approximations de la solution à chaque point de calcul  $\mathbf{z}$  étant tel que  $\mathbf{z}^T \mathbf{A} \mathbf{z} < 0$ , où  $\mathbf{A}$  est la matrice de l'opérateur de diffusion discrétisé.  $\mathbf{A}$  n'est donc pas une matrice symétrique définie positive, et elle possède donc au moins une valeur propre négative. Cet exemple illustre la perte de coercivité due à l'apparition de transmissibilités négatives sur un maillage non admissible, et encourage à définir une méthode d'approximation consistante des flux diffusifs se satisfaisant de conditions peu restrictives sur le maillage.

## Méthode SUSHI

Comme on l'a vu plus haut, sur un maillage vérifiant les conditions d'admissibilité pour l'approximation des flux diffusifs, on peut obtenir un flux numérique consistant avec la formule classique à deux points. Sur des maillages généraux, non conformes, où la condition d'orthogonalité est perdue, ce flux numérique n'est plus consistant. Le schéma SUSHI (*Scheme Using Stabilization Hybrid Interfaces*) suggère une solution pour dépasser cette limitation par la recherche d'un troisième point (triangles), ou d'un quatrième point (tétraèdres). Il a été introduit par Eymard *et al.* dans [39] pour la discrétisation de problèmes de diffusion anisotrope et hétérogène sur des maillages généraux non conformes. Considérons le problème de diffusion stationnaire (1.26) avec  $\mathbf{u}(\mathbf{x}) = 0$ , et conditions aux limites de Dirichlet homogènes. Ce problème s'écrit sous forme variationnelle :

$$\text{Trouver } \phi \in H_0^1(\Omega) \text{ t.q. } : \int_{\Omega} \nabla \phi \nabla \psi = \int_{\Omega} f \psi d\mathbf{x} \text{ pour tout } \psi \in H_0^1(\Omega).$$

L'idée de la méthode SUSHI vient d'une conception "éléments finis", le schéma volumes finis

étant dérivé d'une formulation faible discrète. Les flux numériques sont déterminés à l'aide d'un gradient discret  $\nabla_{\mathcal{D}}\phi$ , défini sur la maille diamant  $D_{K,\sigma}$  par :

$$\nabla_{\mathcal{D}}\phi(\mathbf{x}) = \tilde{\nabla}_{\mathcal{D}}\phi(\mathbf{x}) + R_{\sigma}\phi\mathbf{n}_{\sigma} \quad \forall \mathbf{x} \in D_{K,\sigma} \quad (1.33)$$

avec

$$\tilde{\nabla}_{\mathcal{D}}\phi(\mathbf{x}) = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma|(\phi_{\sigma} - \phi_K)\mathbf{n}_{\sigma} \quad \forall K \in \mathcal{T}, \forall \phi \in \mathcal{X}_{\mathcal{D}} \quad (1.34)$$

où  $\mathcal{X}_{\mathcal{D}}$  est un espace d'approximation construit à l'aide des valeurs  $\phi_{\sigma} = \mathcal{A}_{\sigma}(\phi)$  du champ aux faces, avec

$$\mathcal{A}_{\sigma}(\phi) = \sum_{L \in \mathcal{S}(\sigma)} \beta_L \phi_L \quad \forall \sigma \in \mathcal{E}_{\text{int}} \quad (1.35)$$

et tel que

$$\mathbf{x}_{\sigma} = \sum_{L \in \mathcal{S}(\sigma)} \beta_L \mathbf{x}_L \quad \forall \sigma \in \mathcal{E}_{\text{int}} \quad (1.36)$$

avec

$$\sum_{L \in \mathcal{S}(\sigma)} \beta_L = 1 . \quad (1.37)$$

Les coefficients  $\beta_L$  représentent donc des coefficients provenant des mailles voisines pour la reconstruction du champ à la face  $\sigma$ . Le terme additionnel  $R_{\sigma}\phi$  permet d'obtenir le caractère défini de la formulation :

$$R_{\sigma}\phi = \frac{\sqrt{d}}{|K,\sigma|} \left( \phi_{\sigma} - \phi_K - \tilde{\nabla}_{\mathcal{D}}\phi(\mathbf{x}) \cdot (\mathbf{x}_{\sigma} - \mathbf{x}_K) \right) \quad (1.38)$$

Ce schéma ne vérifie pas la conservativité locale des flux, mais permet d'atteindre un ordre de convergence  $\mathcal{O}(h^2)$  en norme  $L^2$  discrète sur des quadrilatères ou hexaèdres généraux [39]. Dans nos expérimentations numériques, on retrouve le même ordre sur des maillages généraux de triangles et tétraèdres.

### Retour à la diffusion pure sur un maillage non admissible

Le schéma SUSHI opérant une reconstruction de flux consistant même sur des maillages non admissibles, nous pouvons revenir au problème qui se posait au début de cette section. Dans le cas d'un pas de temps  $\Delta t = 0.01s$ , là où avec le schéma de diffusion standard on perdait la coercivité de la formulation et où deux mailles prenaient des valeurs extrêmes opposées, avec le schéma SUSHI on retrouve la solution attendue avec précision. En faisant l'étude de la convergence en espace du schéma SUSHI sur ce cas, on retrouve un ordre  $\mathcal{O}(h^2)$ .

### 1.3.3 Expérimentations numériques

Les problèmes séparés de convection pure et de diffusion pure étant plus difficiles à résoudre séparément que le problème complet, on se concentre d’abord sur la vérification séparée des méthodes d’approximation des flux convectifs et diffusifs, avec des tests en convection pure et diffusion pure, et on utilise ensuite l’approche couplée pour approcher la convection–diffusion d’une impulsion gaussienne. Pour effectuer ces tests et vérifier la robustesse des méthodes développées vis-à-vis de maillages de topologies et de “qualité” différentes, nous avons généré huit séries de maillages 2D, et trois séries de maillages 3D, voir figures 1.19 et 1.20.

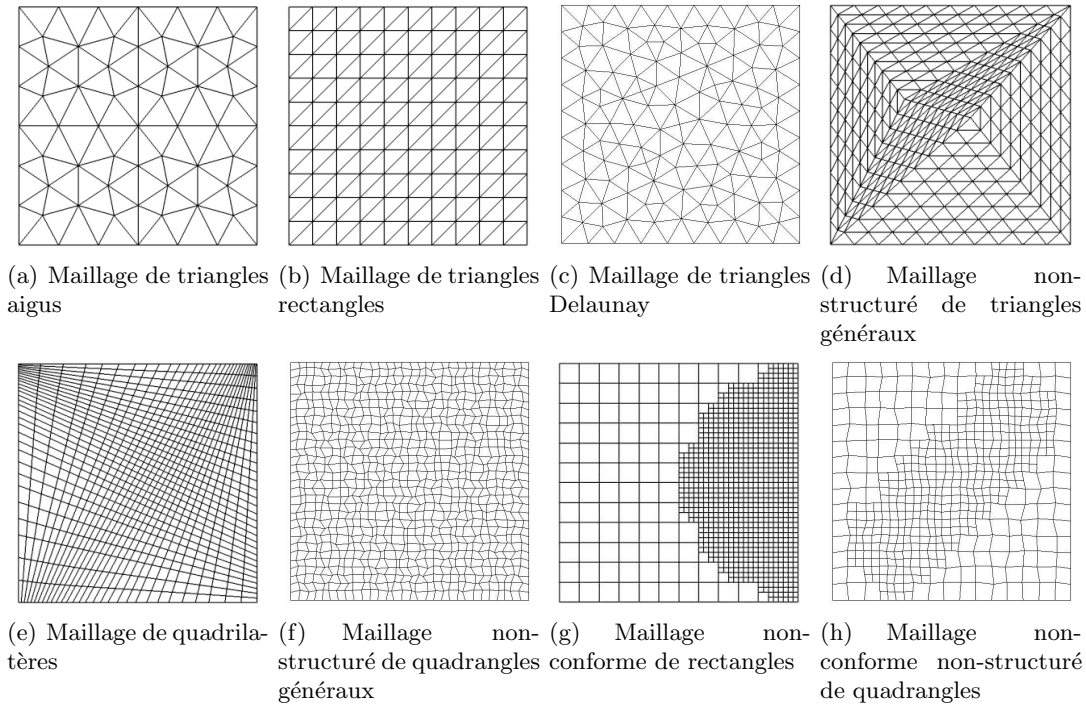


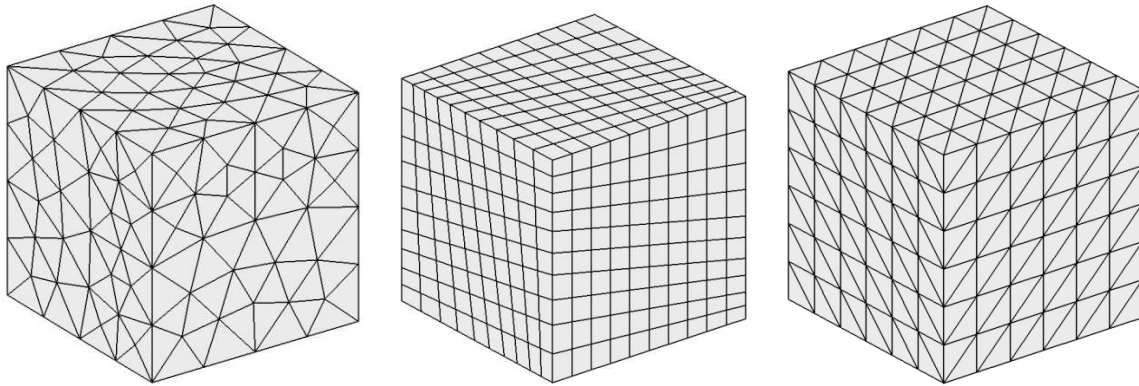
FIG. 1.19 – Huit maillages 2D utilisés pour l’étude de convergence des méthodes MUSCL et SUSHI

#### Convection pure d’une impulsion sinusoïdale

Ce test, défini par Buffard *et al.* [15], décrit l’advection à vitesse  $\mathbf{a}$  constante d’une impulsion sinusoïdale le long de la diagonale du carré unité. L’impulsion est définie par :

$$\phi(\mathbf{x}, t = 0) = \begin{cases} \frac{1}{2}(\cos(5\pi r) + 1) & \text{if } r < \frac{1}{5} \\ 0 & \text{if } r > \frac{1}{5} \end{cases} \quad \text{on } \Omega. \quad (1.39)$$

On complète la définition du problème par des conditions de Dirichlet homogènes sur toute la frontière du domaine. La condition initiale et la solution obtenue à  $t = 1s$  sont représentées sur la figure 1.21. Les courbes de convergence obtenues sont tracées sur les figures 1.22(a) et 1.22(b). Comme le veut la théorie, le principe du maximum discret est vérifié. Comme aucun ordre de convergence théorique n’est connu, nous comparons les résultats obtenus avec ceux de



(a) Maillage de tétraèdres Delaunay (b) Maillage non-structuré d'hexa-  
èdres (c) Maillage de tétraèdres réguliers

FIG. 1.20 – Trois maillages 3D non-structurés utilisés pour l'étude de convergence des méthodes MUSCL multipentes et SUSHI.

la méthode décentrée amont classique, et également avec les simulations numériques conduites par *Buffard & Clain* dans [15].

### Rotation de corps solides

La rotation de corps solides est un test en deux dimensions qui décrit la rotation d'un motif composé d'un carré et d'un cône, voir figure 1.23(a), autour de l'origine, dans le domaine  $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$ . La solution exacte à  $t = \pi$  est donc égale à la solution initiale. Dans le cas précédent de la convection d'une impulsion sinusoïdale, la solution était régulière. Dans le cas présent, la solution est discontinue pour le carré, et non dérivable pour le cône. En approchant la solution de ce problème sur un maillage de triangles généré à l'aide de GMSH et vérifiant le critère de Delaunay en toutes faces, on n'observe pas d'oscillations. Les extrema de la solution initiale sont conservés mais non dépassés : la précision est élevée tout en gardant la vérification du principe du maximum discret. Les solutions obtenues en utilisant les différents schémas sont présentés sur les figures 1.23(b), 1.23(c) et 1.23(d). Les solutions obtenues à l'aide de la méthode MUSCL multipentes, de la méthode MUSCL en structuré et de la méthode décentrée amont sont présentées sur les figures 1.23(b), 1.23(c) et 1.23(d). La seule méthode qui conduise à une conservation exacte des bornes de la solution est la méthode MUSCL multipentes.

### Diffusion pure

Soient  $\kappa = 1$  et  $\phi(\mathbf{x}) = 0$ . On considère le domaine  $\Omega = [0, 1]^2$  et la solution exacte suivante de (1.26) :

$$\phi(\mathbf{x} = (x, y)) = \frac{\sinh(\pi x)}{\sinh(\pi)} \sin(\pi y) \quad \text{on } \Omega. \quad (1.40)$$

On complète la définition du problème par des conditions aux bords de Dirichlet non-homogènes, ou par des conditions mixtes Dirichlet–Neumann. Dans les deux cas, l'ordre 2 de convergence en espace en norme  $L^2$  discrète  $L^2_{\mathcal{D}}$  est obtenu, sur des maillages généraux non-structurés ou non-conformes de quadrangles et de triangles. Ces résultats sont en accord avec ceux obtenus

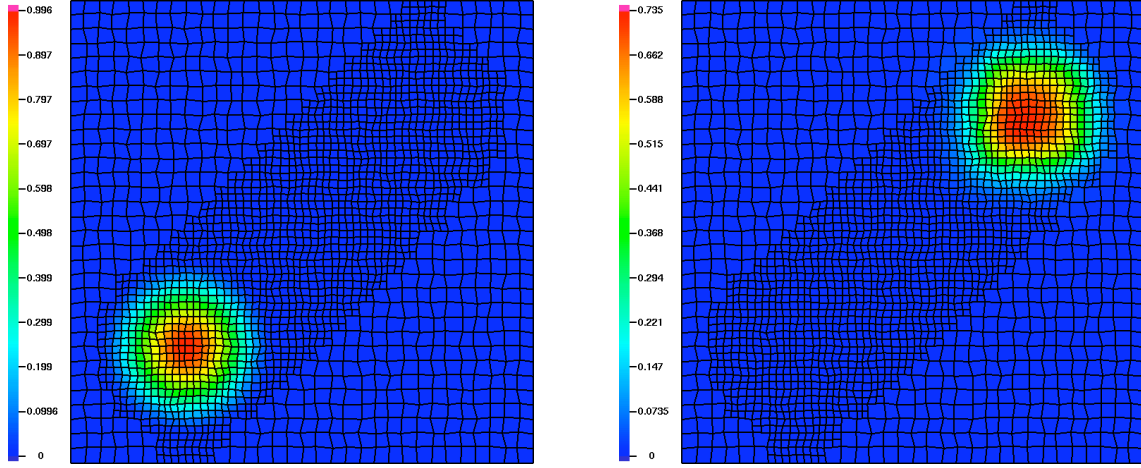
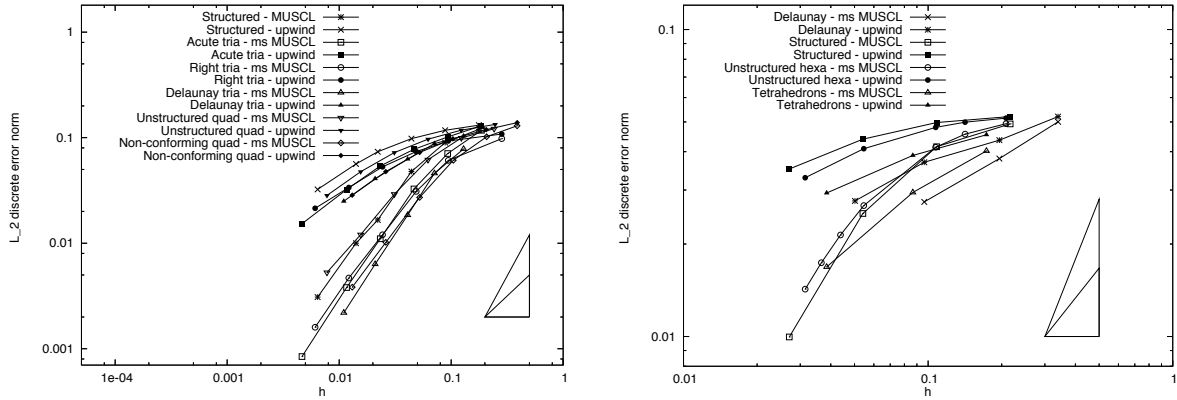


FIG. 1.21 – Condition initiale et solution à  $t = 1s$  pour la convection 2D d’une impulsion sinusoïdale sur le maillage 3.4(h)



(a) Courbes de convergence en espace, problème 2D

(b) Courbes de convergence en espace, problème 3D

FIG. 1.22 – Courbes de convergence pour le problème de convection pure

par Eymard *et al.* dans [40, 36] et Herbin *et al.* dans [68]. Bien que théoriquement, le principe du maximum local ne soit pas attendu pour la méthode SUSHI, celui-ci est respecté dans nos expérimentations numériques. Pour envisager des applications plus complexes, on effectue une étude de convergence en 3D avec des conditions de Dirichlet homogènes sur toute la frontière du domaine  $\Omega = [0, 1] \times [-\frac{1}{2}, \frac{1}{2}] \times [0, 1]$ . L’étude est conduite sur un maillage d’hexaèdres réguliers et deux maillages d’hexaèdres non-structurés représentés sur les figures Fig.3.5(b)–3.5(c). Les ordres de convergence en norme  $L^2_{\mathcal{D}}$  sont respectivement 2.00 et 1.87 pour le maillage d’hexaèdres réguliers et le maillage d’hexaèdres non-structurés. L’ordre de convergence mesuré pour la méthode SUSHI sur un maillage de tétraèdres (Fig. 3.5(c)) est de 1.64.

### Convection–diffusion d’une impulsion gaussienne

On présente dans cette section le couplage de la méthode MUSCL multipentes et du schéma SUSHI. On étudie d’abord la convergence de l’erreur sur l’exemple de l’advection d’une impul-

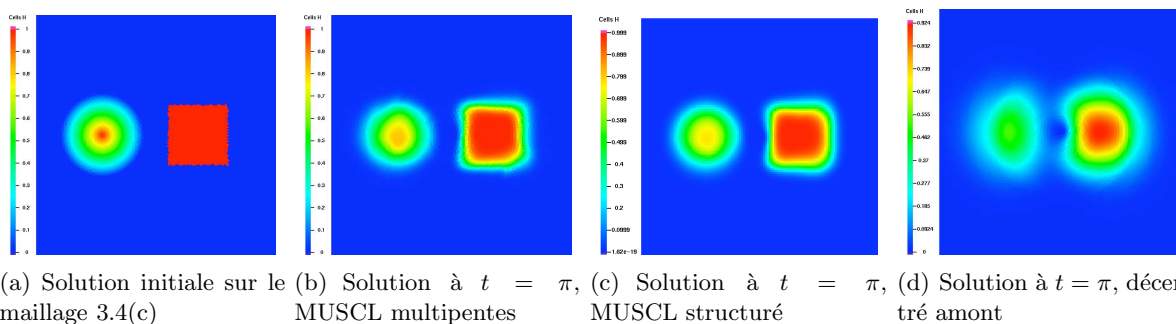


FIG. 1.23 – Rotation de corps solides en 2D : (a) configuration initiale, (b) et (d) solutions obtenues à  $t = \pi$  sur un maillage non-structuré de triangles de Delaunay par les méthodes MUSCL multipentes et décentrée amont, (c) solution obtenue à  $t = \pi$  sur un maillage de carrés par la méthode MUSCL en structuré

sion gaussienne. Considérons le problème instationnaire d’advection–diffusion sans terme source dont la solution analytique en deux dimensions est donnée par Noye *et al.* dans [87], que nous étendons ici en trois dimensions. Ce test décrit la diffusion d’une impulsion gaussienne initiale alors qu’elle est advectée le long de la diagonale principale du domaine  $\Omega = [0, 2]^3 \subset \mathbb{R}^3$ . L’impulsion gaussienne, d’amplitude 1, est initialement centrée sur le point  $(0.5, 0.5, 0.5)$ . On impose

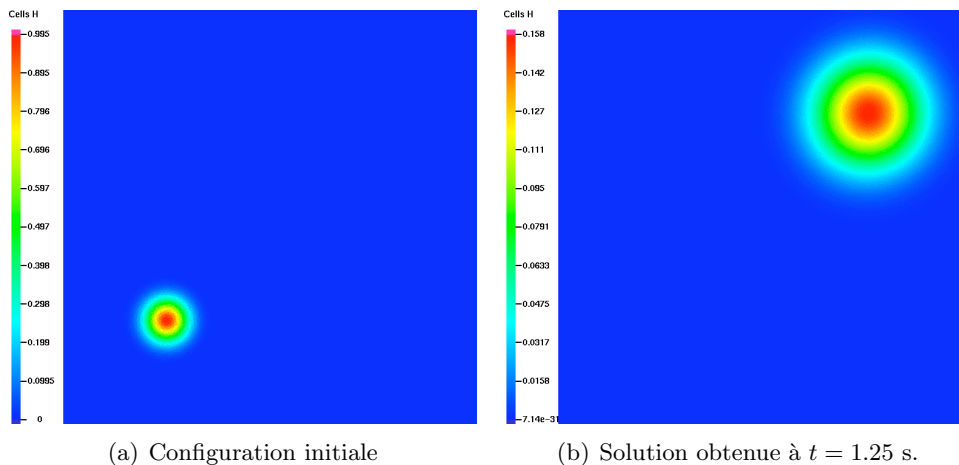


FIG. 1.24 – Solutions à  $t = 0$  (exacte) et  $t = 1.25$  (simulée) de l’advection d’une impulsion gaussienne sur le maillage Delaunay, avec  $h = 0.0219932$ .

des conditions de Dirichlet sur  $\partial\Omega$ . Au temps final  $t = 1.25$  s, l’impulsion est centrée sur le point  $(1.5, 1.5, 1.5)$  et est d’amplitude  $\frac{1}{6}$ . Nous avons conduit l’étude de convergence pour la solution 3D, ainsi que pour son équivalent 2D obtenu en considérant le plan  $z = 0$ . L’ordre de convergence obtenu par l’approche couplée MUSCL–SUSHI est approximativement le double de celui obtenu par le couplage décentré amont–SUSHI. La solution obtenue avec l’approche MUSCL multipentes–SUSHI est en bon accord avec la solution exacte. La solution obtenue avec l’approche décentrée amont–SUSHI présente une erreur de 10 – 20% selon le type de maillage.



## 1.4 Plan du manuscrit

Les deux méthodes présentées ci-dessus font l'objet des deux articles repris en tant que chapitres 2 et 3. Dans le chapitre 4, nous présentons trois cas concrets tridimensionnels d'application du code ISIS utilisant le schéma complet détaillé dans ce manuscrit :

- La simulation d'un écoulement dans un tube coudé, qui est une modélisation simplifiée d'un écoulement sanguin dans une crosse aortique modèle. Cette simulation est basée sur un maillage non-structuré d'hexaèdres.
- La simulation d'une flamme laminaire dans un réacteur cylindrique. Cet écoulement présente la particularité d'être à viscosité variable.
- La simulation d'une aéraulique anisotherme à faible nombre de Mach, qui est une modélisation de la propagation de chaleur dans un local muni d'une porte. Cette simulation est basée sur un maillage non-structuré de tétraèdres.

L'annexe A est consacrée à la présentation détaillée des éléments finis non-conformes (de bas degré, inf-sup stables) utilisés dans ISIS pour l'approximation spatiale : l'élément de Rannacher–Turek pour les mailles quadrangulaires et hexaédriques, et son analogue pour les mailles triangulaires et tétraédriques, l'élément de Crouzeix–Raviart.

L'annexe B est un article publié dans les actes de la conférence *Finite Volumes for Complex Applications 5* qui s'est tenue à Aussois en juin 2008. Cet article présente dans un premier temps la construction citée ci-dessus d'une discrétisation  $L^2$ -stable de l'opérateur de convection des équations de Navier–Stokes dans le cadre de la discrétisation spatiale par des éléments finis de Rannacher–Turek et de Crouzeix–Raviart. Il présente ensuite la définition de conditions aux limites artificielles adaptées aux frontières ouvertes où l'on observe des écoulements réentrants dont la vitesse n'est pas connue. Cette contribution a donné lieu à la présentation d'un poster.

L'annexe C est une contribution au benchmark de méthodes d'approximation des problèmes de diffusion anisotrope organisé dans le cadre de la conférence *Finite Volumes for Complex Applications 5*, dont le but était la comparaison de différentes stratégies volumes finis. Dans ce contexte, les organisateurs ont également demandé que l'approximation de ces problèmes par éléments finis de Galerkin classiques soit présentée pour référence : c'est l'objet de ce document.

## Chapitre 2

# Une approximation $L^2$ -stable de l'opérateur convection de Navier–Stokes pour des éléments finis non conformes de bas degré

### 2.1 Introduction

We address in this paper the numerical solution of the variable density unstationary Navier-Stokes equations :

$$\partial_t \varrho + \nabla \cdot \varrho \mathbf{u} = 0 \quad (2.1a)$$

$$\partial_t(\varrho \mathbf{u}) + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u}) - \nabla \cdot \boldsymbol{\tau} + \nabla p = \mathbf{f} \quad (2.1b)$$

where  $\partial_t$  stands for the derivation operator with respect to time,  $\mathbf{u}$  for the fluid velocity,  $p$  for the (dynamic) pressure (the thermodynamic pressure is assumed to be fixed),  $\mathbf{f}$  for a known volumic forcing term and  $\varrho$  for the fluid density, which is supposed to be a known positive quantity. The tensor  $\boldsymbol{\tau}$  is the viscous part of the stress tensor, given by the following expression :

$$\boldsymbol{\tau}(\mathbf{u}) = \mu \left[ \nabla \mathbf{u} + \nabla^t \mathbf{u} - \frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I} \right] \quad (2.2)$$

where the viscosity  $\mu$  is a positive constant real number. The problem is posed over  $\Omega$ , an open bounded connected subset of  $\mathbb{R}^d$  with  $d = 2$  or  $d = 3$ . An initial condition for  $\mathbf{u}$  must be provided, and we suppose that the boundary conditions are of mixed type : the boundary  $\partial\Omega$  of  $\Omega$  is split into  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ , the velocity is prescribed over the part  $\partial\Omega_D$  of positive  $(d-1)$ -dimensional measure while, on  $\partial\Omega_N$ , we suppose that the external forces are given :

$$\text{on } \partial\Omega_D, \quad \mathbf{u} = \mathbf{u}_D \quad (2.3a)$$

$$\text{on } \partial\Omega_N, \quad (\boldsymbol{\tau}(\mathbf{u}) + p\mathbf{I}) \mathbf{n} = \mathbf{g} \quad (2.3b)$$

where  $\mathbf{u}_D$  is a known velocity field,  $\mathbf{n}$  is the outward normal vector to  $\partial\Omega$  and  $\mathbf{g}$  is a known surfacic forcing term.

System (2.1) is a building block for many physical problems. For instance, supplementing it by an energy balance and supposing  $\varrho$  given by the equation of state of perfect gases evaluated at a fixed pressure yields the asymptotic governing equations of natural convection flows in the low Mach number limit. Considering chemical reactions in the flow, radiative transfer and turbulence phenomena, one obtains a model for the simulation of fires, which is the aim of the ISIS free software, developed in France at the Institut de Radioprotection et de Sûreté Nucléaire (IRSN).

The numerical scheme developed in this paper is based on low-order non-conforming finite elements, namely the Crouzeix-Raviart [30] elements for simplicial meshes or the Rannacher-Turek element [96] for quadrilaterals and hexahedra. These pairs of finite element spaces are the lowest order ones that fulfill the discrete version of the so-called Babuska–Brezzi stability condition. They are thus well suited for a coupling with other balance equations discretized by a finite volume technique for monotonicity reasons, as encountered in the above-mentioned context.

For vanishing external forces  $\mathbf{f}$  and  $\mathbf{g}$  and prescribed velocity  $\mathbf{u}_D$ , if one supposes that the unknown functions  $\mathbf{u}$ ,  $p$  and  $\varrho$  are regular, multiplying Equation (2.1b) by  $\mathbf{u}$  and integrating over  $\Omega$  yields the following *a priori* estimate :

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} \varrho |\mathbf{u}|^2 \, d\mathbf{x} + \int_{\Omega} \boldsymbol{\tau}(\mathbf{u}) : \boldsymbol{\nabla} \mathbf{u} \, d\mathbf{x} = \int_{\Omega} p \boldsymbol{\nabla} \cdot \mathbf{u} \, d\mathbf{x} \quad (2.4)$$

Since the second term at the left hand side is non-negative, this relation provides an estimate for the velocity  $\mathbf{u}$ , provided that an adequate treatment is possible for the right hand side, which is the case, for instance, for incompressible or for compressible barotropic flows. The key ingredient to establish (2.4) is the following identity :

$$\int_{\Omega} [\partial_t(\varrho \mathbf{u}) + \boldsymbol{\nabla} \cdot (\varrho \mathbf{u} \otimes \mathbf{u})] \cdot \mathbf{u} \, d\mathbf{x} = \frac{1}{2} \frac{d}{dt} \int_{\Omega} \varrho |\mathbf{u}|^2 \, d\mathbf{x} \quad (2.5)$$

which holds provided that the mass balance equation (2.1a) is satisfied. Unfortunately, if standard finite element techniques are used to discretize the terms  $\partial_t(\varrho \mathbf{u}) + \boldsymbol{\nabla} \cdot (\varrho \mathbf{u} \otimes \mathbf{u})$ , this identity does not hold at the discrete level with the chosen (non-conforming) finite element spaces, essentially because some terms involving jumps across the interfaces of the elements appear when performing the integration by parts necessary to its derivation. As a consequence, some blow up of the solution was observed in our simulations of moderate-to-high Reynolds number flows, especially with rather coarse meshes, as often used in real life applications.

Some techniques to stabilize the discretization of convection-dominant flows suitable for the Crouzeix-Raviart or Rannacher-Turek elements have been presented in the literature. In [78], the authors introduce some upwinding *a posteriori*, *i.e.* after the assembling of the discrete operators, by directly modifying the obtained algebraic system. An edge stabilization method consisting in adding penalization terms involving the jumps in the gradient of the velocity is proposed in [33] (see also [91] for a review); at the present time, this technique seems to be analysed only for Oseen equations (*i.e.* considering the convection field as a given function), and has the drawback of enlarging the stencil of the discrete convection operator. Another approach, which consists in discretizing the convection term by a finite volume technique based on a dual mesh, has been proposed in the literature.

It has first been applied for the Crouzeix-Raviart element to a scalar linear [89] and nonlinear [2, 31, 42] convection diffusion equation, with the goal to obtain a monotone approximation of

the convection operator and a general discretization of the diffusion (in particular, suitable for anisotropic diffusion). Then this technique has been extended to incompressible stationary flows, for the Crouzeix-Raviart element [98, 99], and further for the Rannacher-Turek element [106, section 3.1.4, pp. 115–134]. The common point of these works is to introduce some upwinding, either for the sake of monotonicity [89, 2, 31, 42] or stability [98, 99, 106].

In this paper, we follow the finite-volume route, and extend it to the case of variable density flows. In addition, essentially because we address the unstationary case, and thus we may rely on the stability induced by the conservation of the kinetic energy (2.4) for the control of the solution, we develop a centered approximation.

This paper is structured as follows. We first describe the space discretization (Section 2.2). Then the proposed approximation for the unstationary and convection terms is given and its stability is proven in Section 2.3. A fractional step algorithm for the Problem (2.1) is then built (Section 2.4) on this basis. We finally compute the solution to different analytical and benchmark tests (Section 2.5) in order to assert the potentialities of the scheme.

The discretization presented here enters as an ingredient in already published (entropy preserving) schemes for compressible flows [48, 50, 51]; compared to these works, we make here a much more involved description, with a detailed proof of the essential stability result, deal with general boundary conditions, extend the discussion to implementation aspects and report an in-depth numerical study.

## 2.2 Meshes and discretization spaces

Let  $\mathcal{T}$  be a decomposition of the domain  $\Omega$  either in simplices (triangles in 2D or tetrahedra in 3D) or, in the case where the shape of  $\Omega$  allows it, in rectangles or rectangular parallelepipeds. This decomposition  $\mathcal{T}$  is assumed to be regular in the usual sense of the finite element literature (*e.g.* [24]). By  $\mathcal{E}(K)$ , we denote the set of the faces  $\sigma$  of the element  $K \in \mathcal{T}$ . The set of all faces of the mesh is denoted by  $\mathcal{E}$ ; the set of faces included in the boundary of  $\Omega$  is denoted by  $\mathcal{E}_{\text{ext}}$  and the set of internal faces (*i.e.*  $\mathcal{E} \setminus \mathcal{E}_{\text{ext}}$ ) is denoted by  $\mathcal{E}_{\text{int}}$ . The set  $\mathcal{E}_{\text{ext}}$  itself decomposes into the set of external faces included in  $\partial\Omega_{\text{D}}$ , denoted by  $\mathcal{E}_{\text{D}}$ , and the set of external faces included in  $\partial\Omega_{\text{N}}$ , denoted by  $\mathcal{E}_{\text{N}}$ . The face  $\sigma \in \mathcal{E}_{\text{int}}$  separating the cells  $K, L \in \mathcal{T}$  is denoted by  $K|L$  and an external face  $\sigma \in \mathcal{E}_{\text{ext}}$  of the cell  $K \in \mathcal{T}$  is denoted by  $K|\text{ext}$ . By  $|K|$  and  $|\sigma|$ , we denote the measure, respectively, of an element  $K$  and of a face  $\sigma$ . For  $\sigma \in \mathcal{E}(K)$ ,  $\mathbf{n}_{K,\sigma}$  stands for the unit normal vector to  $\sigma$  outward to  $K$ .

The finite elements used in this paper are the Crouzeix-Raviart element for simplicial meshes (see [30] for the seminal paper and, for instance, [34, p. 199–201] for a synthetic presentation), and the so-called "rotated bilinear element" introduced by Rannacher and Turek for quadrilateral or hexahedric meshes [96].

For the discretization of the velocity components, the reference element for the Crouzeix-Raviart is the unit  $d$ -simplex and the discrete functional space is the space  $\mathbb{P}_1$  of affine polynomials. The reference element  $\widehat{K}$  for the rotated bilinear element is the unit  $d$ -cube and the discrete functional space on  $\widehat{K}$  is  $\tilde{Q}_1(\widehat{K})$  :

$$\tilde{Q}_1(\widehat{K}) = \text{span} \{1, (\mathbf{x}_i)_{i=1,\dots,d}, (\mathbf{x}_i^2 - \mathbf{x}_{i+1}^2)_{i=1,\dots,d-1}\}$$

For both elements used here, the degrees of freedom are determined by the following set of nodal functionals  $\{m_\sigma, \sigma \in \mathcal{E}(K)\}$  with :

$$m_\sigma(v) = \frac{1}{|\sigma|} \int_\sigma v d\sigma \quad (2.6)$$

The mapping from the reference element to the actual discretization cell is the standard affine mapping for the Crouzeix-Raviart element, and the standard  $Q_1$  mapping for the Rannacher-Turek element. Finally, in both cases, the continuity of the average value  $m_\sigma(v)$  of a discrete function  $v$  across each face of the mesh is required, thus the discrete space  $V_{\mathcal{D}}^{(i)}$  for the  $i^{\text{th}}$  component of the velocity is defined as follows :

$$V_{\mathcal{D}}^{(i)} = \left\{ v \in L^2(\Omega) : v|_K \in W(K), \forall K \in \mathcal{T}; \right. \\ \left. m_\sigma(v|_K) = m_\sigma(v|_L), \forall \sigma = K|L \in \mathcal{E}_{\text{int}}; m_\sigma(v) = m_\sigma(\mathbf{u}_{\mathcal{D}}^{(i)}), \forall \sigma \in \mathcal{E}_{\mathcal{D}} \right\} \quad (2.7)$$

where the space  $W(K)$  is thus the space of affine functions over  $K$  for the Crouzeix-Raviart element and the space of functions obtained from  $\tilde{Q}_1(\hat{K})$  through the  $Q_1$  mapping for the Rannacher-Turek element. Since only the continuity of the integral over the faces of the mesh is imposed, the velocity is discontinuous through each face; the discretization is thus non-conforming in  $H^1(\Omega)^d$ .

From the definition (2.6), each velocity degree of freedom can be associated to a face of an element. Hence, the velocity degrees of freedom may be indexed by the number of the component and the associated face, and the set of velocity degrees of freedom reads :

$$\{\mathbf{v}_{\sigma,i}, \sigma \in \mathcal{E} \setminus \mathcal{E}_{\mathcal{D}}, 1 \leq i \leq d\}$$

We define  $\mathbf{v}_\sigma = \sum_{i=1}^d \mathbf{v}_{\sigma,i} \mathbf{e}^{(i)}$  where  $\mathbf{e}^{(i)}$  is the  $i^{\text{th}}$  vector of the canonical basis of  $\hat{x}^d$ . We denote by  $\varphi_\sigma^{(i)}$  the vector shape function associated to  $\mathbf{v}_{\sigma,i}$ , which, by the definition of the Crouzeix-Raviart and Rannacher-Turek finite elements, reads :

$$\varphi_\sigma^{(i)} = \varphi_\sigma \mathbf{e}^{(i)}$$

where  $\varphi_\sigma$  is the scalar basis function.

For both the Crouzeix-Raviart and the Rannacher-Turek discretizations, the approximation space for the pressure is the space of piecewise constant functions. In addition, since we want to design an algorithm suitable for coupling with possible other balance equations the unknown of which governs the value of the density, we will suppose that  $\varrho$  is approximated by a discrete function, belonging to the same space than the pressure. The degrees of freedom for the pressure thus are  $\{p_K, K \in \mathcal{T}\}$  and the density is defined by  $\{\varrho_K, K \in \mathcal{T}\}$ .

In the definition of the scheme, we will also need a dual mesh, which is defined as follows. For any  $K \in \mathcal{T}$  and any face  $\sigma \in \mathcal{E}(K)$ , let  $D_{K,\sigma}$  be the cone of basis  $\sigma$  and of opposite vertex the mass center of  $K$ . The volume  $D_{K,\sigma}$  is referred to as the half-diamond mesh associated to  $K$  and  $\sigma$ . We now define the diamond mesh  $D_\sigma$  associated to  $\sigma$  as follows : if  $\sigma \in \mathcal{E}_{\text{int}}$ ,  $\sigma = K|L$ ,  $D_\sigma = D_{K,\sigma} \cup D_{L,\sigma}$ ; if  $\sigma \in \mathcal{E}_{\text{ext}}$ ,  $\sigma = K|\text{ext}$ ,  $D_\sigma = D_{K,\sigma}$ . The set of volumes  $(D_\sigma)_{\sigma \in \mathcal{E}}$  provides the dual mesh  $\bar{\mathcal{T}}$  of  $\Omega$ . We denote by  $\bar{\mathcal{E}}(D_\sigma)$  the set of faces of any  $D_\sigma \in \bar{\mathcal{T}}$ , and by  $\varepsilon = D_\sigma|D_{\sigma'}$  the face separating two diamond meshes  $D_\sigma$  and  $D_{\sigma'}$  (see Figure 2.1). Note that, for a diamond cell

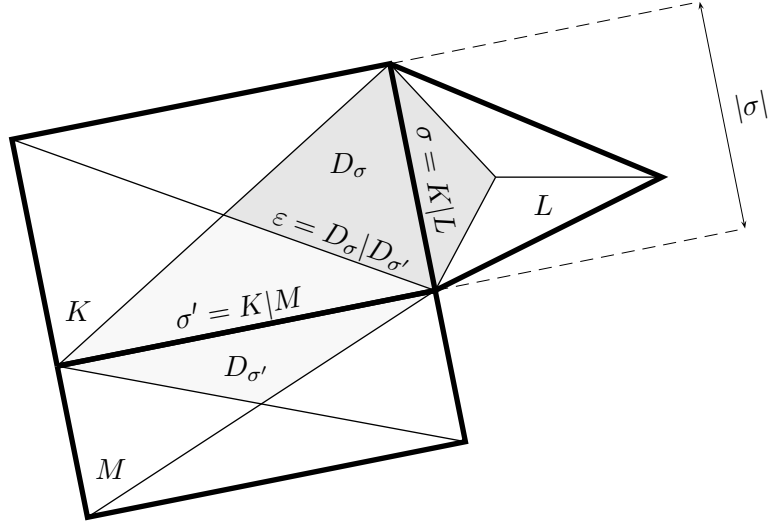


FIG. 2.1 – Notations for control volumes and diamond cells.

$D_\sigma$  adjacent to a boundary of the domain (*i.e.* a diamond cell  $D_\sigma$  associated to a face  $\sigma \in \mathcal{E}_{\text{ext}}$ ) the external face is also a face of the primal mesh; we denote such a face by  $D_\sigma|_{\text{ext}}$ . The unit vector normal to  $\varepsilon \in \bar{\mathcal{E}}(D_\sigma)$  outward to  $D_\sigma$  is denoted by  $\mathbf{n}_{\sigma,\varepsilon}$ .

Since the velocity is prescribed on the faces of  $\mathcal{E}_D$ , the associated diamond cells  $(D_\sigma)_{\sigma \in \mathcal{E}_D}$  do not play any role in the definition of the scheme, whereas their internal faces play a special role. We thus decide to remove these diamond cells from the dual mesh, and to define the set  $\bar{\mathcal{E}}_D$  as their internal faces (see Figure 2.2). On the part  $\partial\Omega_N$  of the boundary, the faces of the primal and dual mesh are the same, and we define  $\bar{\mathcal{E}}_N$  by  $\bar{\mathcal{E}}_N = \mathcal{E}_N$ .

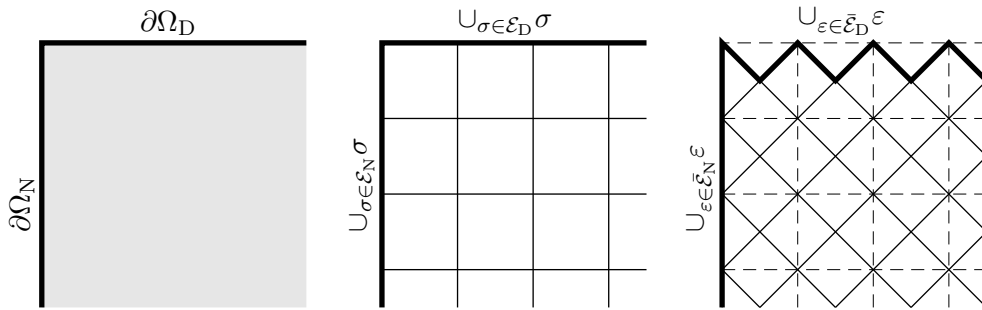


FIG. 2.2 – Notations for the boundaries of the computational domain, of the primal mesh and of the dual mesh.

## 2.3 A $L^2$ -stable convection operator

The goal of this section is to present the proposed discretization for the convection operator. This presentation does not need to be linked to any specific time-marching algorithm, and is thus disconnected from the algorithm used for the solution of the complete system : we just suppose that we know a density field, at the end and beginning of the time step (respectively denoted by  $\varrho$  and  $\varrho^*$ ), the beginning-of-step velocity ( $\mathbf{u}^*$ ) and a convection field such a finite-volume like mass balance on the primal mesh is satisfied (Equation (2.15) below), and build on this basis the desired discrete convection operator. This form of the discrete mass balance is of course the form which can be expected from the discretization of (2.1a) by the considered finite elements.

This construction is made in two steps : first we suppose that a similar mass balance relation holds for each diamond cell (section 2.3.1), then we show how to get it from a discrete velocity field satisfying the mass balance on the primal mesh (section 2.3.2). We mention in passing that in [19], an approximation for the velocity satisfying the mass conservation over a dual vertex-centered mesh is derived from a (possibly high-order) mixed finite element solution of the momentum and mass balance equations.

The general form (*i.e.* without specifying the time marching algorithm) of the discrete momentum balance is then given (section 2.3.3), and we conclude by some general remarks (section 2.3.4).

### 2.3.1 General form of the convection operator and stability analysis

Let us address in this section a sub-problem of the discretization of (2.1) which consists in building a discrete convection operator  $\partial_t(\varrho u) + \nabla \cdot (\mathbf{u} \mathbf{q})$  for a scalar unknown  $u$  satisfying Dirichlet and Neumann boundary conditions on  $\partial\Omega_D$  and  $\partial\Omega_N$  respectively, supposing that the momentum field  $\mathbf{q}$  is known and such that the mass balance holds. More specifically, we assume that  $\mathbf{q}$  and  $\varrho$  are such that :

$$\partial_t \varrho + \nabla \cdot \mathbf{q} = 0 \quad (2.8)$$

and that this yields a discrete mass balance of the form :

$$\forall \sigma \in \mathcal{E} \setminus \mathcal{E}_D, \quad \frac{|D_\sigma|}{\delta t} (\varrho_\sigma - \varrho_\sigma^*) + \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} = 0 \quad (2.9)$$

In this relation,  $\varrho_\sigma$  and  $\varrho_\sigma^*$  stand for an approximation of the density in the diamond cell  $D_\sigma$  at the current and previous time step respectively, and  $F_{\sigma,\varepsilon}$  is an approximation of the outward mass flux through the face  $\varepsilon$  of  $D_\sigma$ , associated to the momentum  $\mathbf{q}$  :

$$F_{\sigma,\varepsilon} = \int_\varepsilon \mathbf{q} \cdot \mathbf{n}_{\sigma,\varepsilon} d\sigma$$

The families of real numbers  $(\varrho_\sigma)_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D}$  and  $(\varrho_\sigma^*)_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D}$  are supposed to be positive. The mass flux is assumed to be outward on  $\partial\Omega_N$ , which means that, for any  $\varepsilon \in \bar{\mathcal{E}}_N$ ,  $F_{\sigma,\varepsilon} \geq 0$ .

Let  $u$  and  $u^*$  be two (scalar) Crouzeix-Raviart or Rannacher-Turek functions, the prescribed value for  $u$  on  $\partial\Omega_D$  being  $u_D$ . For the discretization of the convection operator  $u \mapsto \partial_t(\varrho u) + \nabla \cdot (\mathbf{u} \mathbf{q})$  (which acts in (2.1b) on each component of the velocity), we propose a finite volume

operator  $C$  based on the dual mesh and defined as follows :

$$\forall \sigma \in \mathcal{E} \setminus \mathcal{E}_D, \quad (Cu)_\sigma = \frac{1}{\delta t} (\varrho_\sigma u_\sigma - \varrho_\sigma^* u_\sigma^*) + \frac{1}{|D_\sigma|} \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} u_\varepsilon \quad (2.10)$$

where  $u_\varepsilon$  is defined as follows :

$$\begin{aligned} \text{for } \varepsilon \in \bar{\mathcal{E}}_{\text{int}} \cup \bar{\mathcal{E}}_D, \quad \varepsilon = D_\sigma | D_{\sigma'}, \quad u_\varepsilon &= \frac{1}{2} (u_\sigma + u_{\sigma'}) \\ \text{for } \varepsilon \in \bar{\mathcal{E}}_N, \quad \varepsilon = D_\sigma | \text{ext}, \quad u_\varepsilon &= u_\sigma \end{aligned} \quad (2.11)$$

For internal and Dirichlet faces, the choice thus corresponds to a centered one (remember that the faces of  $\bar{\mathcal{E}}_D$  indeed separate two diamond-cells, one of which is associated to a primal Dirichlet face where the value of the unknown is given by the boundary data (2.7)). An upwind choice would lead to no difference in terms of stability. For faces located on  $\partial\Omega_N$ , since the mass flux is outward, this choice is the upwind one, and also seems to be the only reasonable one, since no natural value is readily available for the external side of the boundary.

At the continuous level, supposing that (2.8) holds and that the functions appearing in the following relations are regular, we have :

$$\int_{\Omega} [\partial_t(\varrho u) + \nabla \cdot (u\mathbf{q})] u \, d\mathbf{x} = \int_{\Omega} [\varrho \partial_t u + \mathbf{q} \cdot \nabla u] u \, d\mathbf{x} = \frac{1}{2} \int_{\Omega} [\varrho \partial_t u^2 + \mathbf{q} \cdot \nabla u^2] \, d\mathbf{x}$$

Integrating by parts and using once again (2.8), we get :

$$\begin{aligned} \int_{\Omega} [\partial_t(\varrho u) + \nabla \cdot (u\mathbf{q})] u \, d\mathbf{x} &= \frac{1}{2} \int_{\Omega} [\varrho \partial_t u^2 - u^2 \nabla \cdot \mathbf{q}] \, d\mathbf{x} + \frac{1}{2} \int_{\partial\Omega} u^2 \mathbf{q} \cdot \mathbf{n} d\sigma \\ &= \frac{1}{2} \int_{\Omega} \partial_t(\varrho u^2) \, d\mathbf{x} + \frac{1}{2} \int_{\partial\Omega} u^2 \mathbf{q} \cdot \mathbf{n} d\sigma \end{aligned}$$

and thus, finally :

$$\int_{\Omega} [\partial_t(\varrho u) + \nabla \cdot (u\mathbf{q})] u \, d\mathbf{x} = \frac{1}{2} \int_{\Omega} \partial_t(\varrho u^2) \, d\mathbf{x} + \frac{1}{2} \int_{\partial\Omega_N} u^2 \mathbf{q} \cdot \mathbf{n} d\sigma + \frac{1}{2} \int_{\partial\Omega_D} u_D^2 \mathbf{q} \cdot \mathbf{n} d\sigma \quad (2.12)$$

which, since  $\mathbf{q} \cdot \mathbf{n}$  is supposed to be non-negative over  $\partial\Omega_N$ , provides a stability estimate for  $u$  by an integration with respect to the time. The following theorem states that the operator  $C$  satisfies a discrete analogue of this relation in the case where  $u_D = 0$ . It adapts to the case of a discretization based on a dual mesh and generalizes to mixed boundary conditions a similar stability result which can be found in [48] (theorem 3.1, p. 317).

**Theorem 2.3.1** *Provided that the discrete mass balance (2.9) holds and that  $u_D = 0$ , the convection operator defined by (2.10)-(2.11) satisfies the following stability result :*

$$\sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} |D_\sigma| (Cu)_\sigma u_\sigma \geq \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} [\varrho_\sigma u_\sigma^2 - \varrho_\sigma^* (u_\sigma^*)^2] + \frac{1}{2} \sum_{\varepsilon = D_\sigma | \text{ext} \in \bar{\mathcal{E}}_N} F_{\sigma,\varepsilon} u_\sigma^2 \quad (2.13)$$



**Proof** We have :

$$\sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} |D_\sigma| (Cu)_\sigma u_\sigma = T_1 + T_2$$

with :

$$T_1 = \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} [\varrho_\sigma u_\sigma - \varrho_\sigma^* u_\sigma^*] u_\sigma, \quad T_2 = \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} u_\sigma \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} u_\varepsilon$$

For the first term, we get  $T_1 = T_{1,1} + T_{1,2} + T_{1,3}$  with :

$$\begin{aligned} T_{1,1} &= \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} (\varrho_\sigma - \varrho_\sigma^*) u_\sigma^2 \\ T_{1,2} &= \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} \varrho_\sigma^* [u_\sigma^2 - (u_\sigma^*)^2] \\ T_{1,3} &= \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} \varrho_\sigma^* [u_\sigma - u_\sigma^*]^2 \end{aligned}$$

The term  $T_{1,3}$  is always positive and can be seen as a dissipation associated to the backward Euler time discretization. Turning now to  $T_2$ , we get  $T_2 = T_{2,1} + T_{2,2}$  with :

$$T_{2,1} = \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} u_\sigma^2 \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon}, \quad T_{2,2} = \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} (u_\varepsilon - u_\sigma) u_\sigma$$

By Equation (2.9), the term  $T_{2,1}$  cancels with  $T_{1,1}$ . Using the identity  $2a(a-b) = a^2 + (a-b)^2 - b^2$  valid for any real number  $a$  and  $b$ , we get :

$$T_{2,2} = -\frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} [u_\sigma^2 + (u_\sigma - u_\varepsilon)^2 - (u_\varepsilon)^2]$$

So  $T_{2,2} = T_{2,2,1} + T_{2,2,2}$  with :

$$T_{2,2,1} = -\frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} u_\sigma^2 \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} = \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \frac{|D_\sigma|}{\delta t} (\varrho_\sigma - \varrho_\sigma^*) u_\sigma^2$$

and, introducing the notation  $(u^2)_{\sigma,\varepsilon} = (u_\varepsilon)^2 - (u_\sigma - u_\varepsilon)^2$  :

$$T_{2,2,2} = \frac{1}{2} \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} (u^2)_{\sigma,\varepsilon}$$

Reordering the summations, we get :

$$\begin{aligned} T_{2,2,2} &= \frac{1}{2} \sum_{\varepsilon = D_\sigma | D_{\sigma'} \in \bar{\mathcal{E}}_{\text{int}}} F_{\sigma,\varepsilon} [(u^2)_{\sigma,\varepsilon} - (u^2)_{\sigma',\varepsilon}] \\ &\quad + \frac{1}{2} \sum_{\varepsilon = D_\sigma | \text{ext} \in \bar{\mathcal{E}}_D} F_{\sigma,\varepsilon} (u^2)_{\sigma,\varepsilon} + \frac{1}{2} \sum_{\varepsilon = D_\sigma | \text{ext} \in \bar{\mathcal{E}}_N} F_{\sigma,\varepsilon} (u^2)_{\sigma,\varepsilon} \end{aligned} \quad (2.14)$$

For  $\varepsilon = D_\sigma | D_{\sigma'} \in \bar{\mathcal{E}}_{\text{int}}$ , we have  $u_\varepsilon = (u_\sigma + u_{\sigma'})/2$  and so  $(u^2)_{\sigma,\varepsilon} = u_\sigma u_{\sigma'}$  and finally  $(u^2)_{\sigma,\varepsilon} - (u^2)_{\sigma',\varepsilon} = 0$ . For  $\varepsilon = D_\sigma | \text{ext} \in \bar{\mathcal{E}}_D$ , supposing that the prescribed value is  $u_D = 0$ , we have

$u_\varepsilon = u_\sigma/2$  and so  $(u^2)_{\sigma,\varepsilon} = 0$ . Finally, for  $\varepsilon = D_\sigma|\text{ext} \in \bar{\mathcal{E}}_N$ ,  $u_\varepsilon = u_\sigma$  and  $(u^2)_{\sigma,\varepsilon} = u_\sigma^2$ . Gathering all the terms, we thus have :

$$\sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_D} |D_\sigma| (Cu)_\sigma u_\sigma \geq T_{1,2} + T_{2,2,1} + \frac{1}{2} \sum_{\varepsilon = D_\sigma|\text{ext} \in \bar{\mathcal{E}}_N} F_{\sigma,\varepsilon} u_\sigma^2$$

which concludes the proof.  $\blacksquare$

**Remark 1 (Upwind scheme)** *An upwind choice for the value of the unknowns at the dual faces is also possible, and its effect is to introduce an artificial dissipation.*

Indeed, the only changes induced by this choice in the preceding proof lie in the evaluation of the terms appearing in  $T_{2,2,2}$  in Relation (2.14). For an internal face  $\varepsilon = D_\sigma|D_{\sigma'}$ , supposing without loss of generality that the chosen orientation for the face is such that  $F_{\sigma,\varepsilon} \geq 0$ , we would have  $u_\varepsilon = u_\sigma$  and thus  $(u^2)_{\sigma,\varepsilon} - (u^2)_{\sigma',\varepsilon} = (u_{\sigma'} - u_\sigma)^2$ . The corresponding energy flux in (2.14), which reads  $F_{\sigma,\varepsilon} [(u^2)_{\sigma,\varepsilon} - (u^2)_{\sigma',\varepsilon}]$  is thus always positive. Let us compare this term with the dissipation which would be induced by a diffusion term. For a two-point flux finite volume scheme, this latter takes the form  $\lambda |\varepsilon| (u_{\sigma'} - u_\sigma)^2/h_\varepsilon$  where  $\lambda$  is the diffusion coefficient and  $h_\varepsilon$  is a geometric quantity associated to the face  $\varepsilon$  and of the same magnitude as its diameter. We thus see that the "numerical diffusion" through the dual face  $\varepsilon$  is given by  $h_\varepsilon \mathbf{q}_\varepsilon \cdot \mathbf{n}$ , where  $\mathbf{q}_\varepsilon$  stands for the mean value of  $\mathbf{q}$  over  $\varepsilon$ .

For faces of  $\bar{\mathcal{E}}_D$ , by a similar computation, we easily check that the energy flux is either zero or positive. For faces of  $\bar{\mathcal{E}}_N$ , the computation is unchanged, since the performed choice was already the upwind one.

**Remark 2 (Non-homogeneous Dirichlet boundary conditions)** *When the prescribed value  $u_D$  is not zero, the "energy flux" over  $\varepsilon = D_\sigma|D_{\sigma'} \in \bar{\mathcal{E}}_D$  (i.e. the quantity appearing in the second summation of Relation (2.14)) reads :*

$$F_{\sigma,\varepsilon} (u^2)_{\sigma,\varepsilon} = F_{\sigma,\varepsilon} u_\sigma u_{\sigma'} = F_{\sigma,\varepsilon} u_\sigma \frac{1}{|\varepsilon'|} \int_{\varepsilon'} u_D$$

which is consistent with the inlet flux term appearing at the continuous level (i.e. the last term of Equation (2.12)), but with the following two differences :

- (i) the persistence of the unknown  $u_\sigma$  makes that the identity (2.13) does not yield, at least in its present form, a control on the solution ;
- (ii) the momentum flux  $F_{\sigma,\varepsilon}$  is not exactly the one entering the domain. This slight difference only has a weak impact if the momentum field  $\mathbf{q}$  is given, but, for Navier-Stokes equations, this field is itself a function of the unknown  $u$  (which, in this context, will be a component of the velocity), and, once again, the solution reappears in an a priori uncontrolled term.

Note however that this difficulty associated to boundary conditions seems to have no impact in practice, since no uncontrolled growth of the kinetic energy was ever observed in our computations with non-homogeneous Dirichlet boundary conditions.

### 2.3.2 Interpolating the mass flowrates

We now turn to the case where the field  $\mathbf{q} = \varrho \mathbf{u}$  is itself obtained from a discretization of the complete problem (2.1). With the chosen finite elements and a backward Euler discretization

with respect to time, the pressure and density being piecewise constant per primal cell, the discrete mass balance takes the form :

$$\forall K \in \mathcal{T}, \quad \frac{|K|}{\delta t} (\varrho_K - \varrho_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma} = 0 \quad (2.15)$$

In this relation,  $\varrho_K$  and  $\varrho_K^*$  stand for an approximation of the density in the primal cell  $K$  at the end and beginning of the time step, and  $F_{K,\sigma}$  is an approximation of the outward mass flux through the face  $\sigma$  of  $K$ . Since the density is discontinuous through  $\sigma$ , the expression of  $F_{K,\sigma}$  must be written as :

$$F_{K,\sigma} = |\sigma| \varrho_\sigma^m \mathbf{u}_\sigma \cdot \mathbf{n}_{K,\sigma} \quad (2.16)$$

where any reasonable approximation for  $\varrho_\sigma^m$  seems to be suitable in the present context, because the density is supposed to be given and positive. Throughout this paper, the centered choice is performed for the internal faces, a value is computed from the data on  $\partial\Omega_D$  and the upwind choice is made on  $\partial\Omega_N$  :

$$\begin{aligned} \text{for } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, & \quad \varrho_\sigma^m = \frac{1}{2} (\varrho_K + \varrho_L) \\ \text{for } \sigma \in \mathcal{E}_D, & \quad \varrho_\sigma^m = \frac{1}{|\sigma|} \int_\sigma \varrho d\sigma \\ \text{for } \sigma \in \mathcal{E}_N, \sigma = K|\text{ext}, & \quad \varrho_\sigma^m = \varrho_K \end{aligned} \quad (2.17)$$

**Remark 3** *In a more general context, choosing for  $\varrho_\sigma^m$  an upwind discretization may be a convenient way to ensure the positivity of the density (see [48, 46, 47, 51] for works exploiting this argument in the context of compressible flows).*

For the Crouzeix-Raviart element and, with the specific meshes considered here, for the Rannacher-Turek element, we have

$$\int_K \varphi_\sigma d\mathbf{x} = |D_{K,\sigma}|$$

where  $|D_{K,\sigma}|$  is the measure of the half diamond cell  $D_{K,\sigma}$ . Thus, still with a backward Euler time discretization, applying a mass lumping technique to the finite element discretization of the unstationary term  $\partial_t(\varrho \mathbf{u})$  in Equation (2.1b) yields, in the discrete equation associated with  $\sigma$  (*i.e.* obtained with the test functions  $\varphi_\sigma^{(i)}$ ,  $1 \leq i \leq d$ ) :

$$\frac{|D_\sigma|}{\delta t} (\varrho_\sigma \mathbf{u}_\sigma - \varrho_\sigma^* \mathbf{u}_\sigma^*) + \text{other terms} = 0 \quad (2.18)$$

where  $\mathbf{u}$  and  $\mathbf{u}^*$  stand for the velocity at the current time step and the previous time step respectively, and  $\varrho_\sigma$  is defined by :

$$\begin{aligned} \text{for } \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, & \quad |D_\sigma| \varrho_\sigma = |D_{K,\sigma}| \varrho_K + |D_{L,\sigma}| \varrho_L \\ \text{for } \sigma \in \mathcal{E}_N, \sigma = K|\text{ext}, & \quad \varrho_\sigma = \varrho_K \end{aligned} \quad (2.19)$$

No value for  $\varrho_\sigma$  needs to be specified for  $\sigma \in \mathcal{E}_D$ , since the velocity is prescribed on this boundary, and no equation is consequently written for the associated degrees of freedom.

To be in position to apply the theory developed in the previous section, the task we have to complete is thus the following : obtain a discretization of the term  $\nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u})$  such that, associated to the time derivative term of Equation (2.18) with the density defined by (2.19), the structure of Equation (2.10) is recovered. In fact, this problem reduces to the definition of momentum fluxes through the faces of the diamond cells such that a discrete mass balance over the diamond cell (*i.e.* Equation (2.9)) holds, starting from the mass balance over the primal cells (2.15). The construction of these fluxes is the goal of the remainder of this section. We first give an argument which provides a general technique for this purpose, then successively address the case of the Crouzeix-Raviart elements, the Rannacher-Turek elements in two and three dimensions and, finally, in axisymetrical coordinates, for a specific type of mesh.

### A general argument

The approach adopted here is based on the following elementary result.

**Lemma 1 (Mass balance in a sub-volume of a mesh)** *Let  $K \in \mathcal{T}$ , let  $\varrho_K$  and  $\varrho_K^*$  be two real numbers, and consider a family of real numbers  $(F_{K,\sigma})_{\sigma \in \mathcal{E}(K)}$  such that (2.15) holds. Let  $\mathbf{w}^K$  be a momentum field on  $K$ , such that  $\nabla \cdot \mathbf{w}^K$  is constant over  $K$  and satisfying :*

$$\int_K \nabla \cdot \mathbf{w}^K \, dx = \sum_{\sigma \in \mathcal{E}(K)} \int_{\sigma} \mathbf{w}^K \cdot \mathbf{n}_{K,\sigma} d\sigma = \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma} \quad (2.20)$$

*Let  $D$  be a subset of  $K$  with boundary  $\partial D$ , and  $\mathbf{n}_{\partial D}$  be the normal vector to  $\partial D$  outward  $D$ . Then the following property holds :*

$$\frac{|D|}{\delta t} (\varrho_K - \varrho_K^*) + \int_{\partial D} \mathbf{w}^K \cdot \mathbf{n}_{\partial D} d\sigma = 0$$

**Proof** Using the fact that the divergence of  $\mathbf{w}$  is constant over  $K$ , then Relation (2.15), we have :

$$\int_{\partial D} \mathbf{w} \cdot \mathbf{n}_{\partial D} d\sigma = \int_D \nabla \cdot \mathbf{w} \, dx = \frac{|D|}{|K|} \int_K \nabla \cdot \mathbf{w} \, dx = -\frac{|D|}{|K|} \frac{|K|}{\delta t} (\varrho_K - \varrho_K^*)$$

■

Suppose now that we are able to build for any  $K \in \mathcal{T}$  a constant divergence field  $\mathbf{w}$  such that (2.20) holds, and that we evaluate the fluxes at each face of a half-diamond cell  $D_{K,\sigma}$  by integration of  $\mathbf{w}^K \cdot \mathbf{n}$  over the face. The set of the faces of  $D_{K,\sigma}$ , denoted by  $\tilde{\mathcal{E}}(D_{K,\sigma})$ , is the union of  $\sigma$  and a set of faces of the dual mesh. By definition of  $\mathbf{w}^K$ , we thus get a flux on  $\sigma$  which is  $F_{K,\sigma}$ , and a family of additional fluxes  $(F_{\sigma,\varepsilon})_{\varepsilon \in \tilde{\mathcal{E}}(D_{K,\sigma}) \setminus \{\sigma\}}$  such that :

$$\frac{|D_{K,\sigma}|}{\delta t} (\varrho_K - \varrho_K^*) + F_{K,\sigma} + \sum_{\varepsilon \in \tilde{\mathcal{E}}(D_{K,\sigma}) \setminus \{\sigma\}} F_{\sigma,\varepsilon} = 0$$

If  $D_{K,\sigma}$  is associated to  $\sigma \in \mathcal{E}_N$ ,  $D_{K,\sigma} = D_\sigma$  and the preceding relation is exactly (2.9), thanks to the definition of  $\varrho_\sigma$  by Equation (2.19). If  $\sigma \in \mathcal{E}_{\text{int}}$ ,  $\sigma = K|L$ , summing the equation for  $D_{K,\sigma}$  and  $D_{L,\sigma}$ , we get, since  $F_{K,\sigma} = -F_{L,\sigma}$  by their definition (2.16) :

$$\frac{1}{\delta t} \left[ (|D_{K,\sigma}| \varrho_K + |D_{L,\sigma}| \varrho_L) - (|D_{K,\sigma}| \varrho_K^* + |D_{L,\sigma}| \varrho_L^*) \right] + \sum_{\varepsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} = 0$$

which is once again (2.9), thanks to (2.19).

### The Crouzeix-Raviart element

Since the shape functions for the Crouzeix-Raviart element are linear over each cell  $K \in \mathcal{T}$ , the field  $\mathbf{w}^K$  may be itself a Crouzeix-Raviart function. A first possible choice is to derive it by direct interpolation of the fluxes at the faces  $(F_{K,\sigma})_{\sigma \in \mathcal{E}(K)}$  :

$$\mathbf{w}^K(\mathbf{x}) = \sum_{\sigma \in \mathcal{E}(K)} \frac{F_{K,\sigma}}{|\sigma|} \mathbf{n}_{K,\sigma} \varphi_\sigma(\mathbf{x})$$

However, if the fluxes  $(F_{K,\sigma})_{\sigma \in \mathcal{E}(K)}$  are computed by relation (2.16), *i.e.* take the form  $F_{K,\sigma} = |\sigma| \varrho_\sigma^m \mathbf{u}_\sigma \cdot \mathbf{n}_{K,\sigma}$ , another possible choice for  $\mathbf{w}^K$  is :

$$\mathbf{w}^K(\mathbf{x}) = \sum_{\sigma \in \mathcal{E}(K)} \varrho_\sigma^m \mathbf{u}_\sigma \varphi_\sigma(\mathbf{x}) \quad (2.21)$$

This is this latter formula which is chosen for the numerical experiments presented hereafter.

### The Rannacher-Turek element

For the Rannacher-Turek element, the divergence of discrete functions is not constant, so  $\mathbf{w}^K$  cannot belong to this discrete space. We thus suppose that each cell  $K \in \mathcal{T}$  is a rectangular parallelepiped, and we use the following interpolation formula :

$$\mathbf{w}^K(\mathbf{x}) = \sum_{\sigma \in \mathcal{E}(K)} \frac{F_{K,\sigma}}{|\sigma|} \mathbf{n}_{K,\sigma} \alpha_\sigma(\mathbf{x} \cdot \mathbf{n}_{K,\sigma})$$

where the functions  $\alpha_\sigma$  are affine interpolation functions which are determined in such a way that (2.20) holds. For instance, let us suppose that  $d = 2$  and that the considered element is  $(\mathbf{x}_{1,O}, \mathbf{x}_{1,E}) \times (\mathbf{x}_{2,S}, \mathbf{x}_{2,N})$ . Using the notations introduced in figure 2.3, we get :

$$\mathbf{w}^K(\mathbf{x}) = \left[ \begin{array}{cc} \frac{\mathbf{x}_{1,E} - \mathbf{x}_1}{\delta \mathbf{x}_1} \frac{-F_O}{\delta \mathbf{x}_2} + \frac{\mathbf{x}_1 - \mathbf{x}_{1,O}}{\delta \mathbf{x}_1} \frac{F_E}{\delta \mathbf{x}_2} \\ \frac{\mathbf{x}_{1,N} - \mathbf{x}_2}{\delta \mathbf{x}_2} \frac{-F_S}{\delta \mathbf{x}_1} + \frac{\mathbf{x}_2 - \mathbf{x}_{1,S}}{\delta \mathbf{x}_2} \frac{F_N}{\delta \mathbf{x}_1} \end{array} \right]$$

Integrating the quantity  $\mathbf{w}^K \cdot \mathbf{n}$  over each dual face of the mesh leads to an expression of the flux at each dual face as a linear combination of the fluxes at the primal ones ; for the present example, this relation takes the form :

$$F_{\sigma,\varepsilon} = \alpha_O F_O + \alpha_E F_E + \alpha_S F_S + \alpha_N F_N \quad (2.22)$$

where the coefficients  $\alpha_O$ ,  $\alpha_E$ ,  $\alpha_S$  and  $\alpha_N$  are given in the following table :

	$\alpha_O$	$\alpha_E$	$\alpha_S$	$\alpha_N$
$F_{O S}$	$-3/8$	$1/8$	$3/8$	$-1/8$
$F_{S E}$	$-1/8$	$3/8$	$-3/8$	$1/8$
$F_{E N}$	$1/8$	$-3/8$	$-1/8$	$3/8$
$F_{N O}$	$3/8$	$-1/8$	$1/8$	$-3/8$

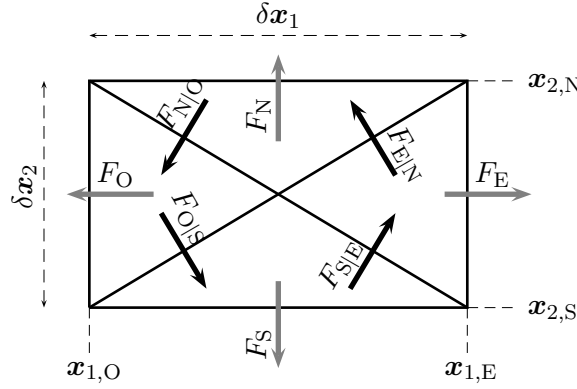


FIG. 2.3 – Rannacher-Turek element in Cartesian coordinates – Local notations for the definition of the interpolation field  $\mathbf{w}^K$ .

### The Rannacher-Turek element in axisymetrical coordinates

Let us now suppose that we are using axisymetrical coordinates, with  $d = 2$ , the first coordinate axis being associated to the distance  $r$  to the symmetry axis and the second coordinate axis being parallel to the symmetry axis (coordinate  $z$ ). We assume that the mesh is a (possibly non-uniform) rectangular grid. Then, with the notations defined on Figure 2.4, we get :

$$\mathbf{w}^K(\mathbf{x}) = \begin{bmatrix} \frac{1 - \psi(r)}{r} (-F_O) + \frac{\psi(r)}{r} F_E \\ z_N - z \frac{-F_S}{\pi(r_E^2 - r_O^2)} + \frac{z - z_S}{\delta z} \frac{F_N}{\pi(r_E^2 - r_O^2)} \end{bmatrix} \quad \text{with } \psi(r) = \frac{r^2 - r_O^2}{r_E^2 - r_O^2}$$

Note that the divergence of this function is indeed constant, since, with this system of coordinates,  $\nabla \cdot \mathbf{w} = \partial_r(r\mathbf{w}_r)/r + \partial_z \mathbf{w}_z$ .

In addition, it may be checked that the integral of each shape function  $\varphi_\sigma$  defines the same dual mesh than in the Cartesian system of coordinates, *i.e.* is equal to the volume of the torus having for section the triangle delimited by the diagonals of the element and  $\sigma$ , as sketched on Figure 2.4. As previously, integrating the quantity  $\mathbf{w}^K \cdot \mathbf{n}$  over each dual face of the mesh leads to a relation of the form :

$$F_{\sigma,\varepsilon} = \alpha_O F_O + \alpha_E F_E + \alpha_S F_S + \alpha_N F_N$$

The coefficients  $\alpha_O$ ,  $\alpha_E$ ,  $\alpha_S$  and  $\alpha_N$  are given as a function of the geometrical features of the element in the following table :

	$\alpha_O$	$\alpha_E$	$\alpha_S$	$\alpha_N$
$F_{O S}$	$-(1 - \beta)/2$	$\beta/2$	$-\frac{r_O}{8\bar{r}} + 2\beta$	$\frac{r_O}{8\bar{r}} - \beta$
$F_{S E}$	$-\gamma/2$	$(1 - \gamma)/2$	$\frac{r_E}{8\bar{r}} - 2\gamma$	$-\frac{r_E}{8\bar{r}} + \gamma$
$F_{E N}$	$\gamma/2$	$-(1 - \gamma)/2$	$\frac{r_E}{8\bar{r}} - \gamma$	$-\frac{r_E}{8\bar{r}} + 2\gamma$
$F_{N O}$	$(1 - \beta)/2$	$-\beta/2$	$-\frac{r_O}{8\bar{r}} + \beta$	$\frac{r_O}{8\bar{r}} - 2\beta$

where  $\bar{r}$ ,  $\beta$  and  $\gamma$  are defined as :

$$\bar{r} = \frac{r_O + r_E}{2}, \quad \beta = \frac{r_O + \delta r/6}{4\bar{r}}, \quad \gamma = \frac{r_E - \delta r/6}{4\bar{r}} \quad (2.23)$$

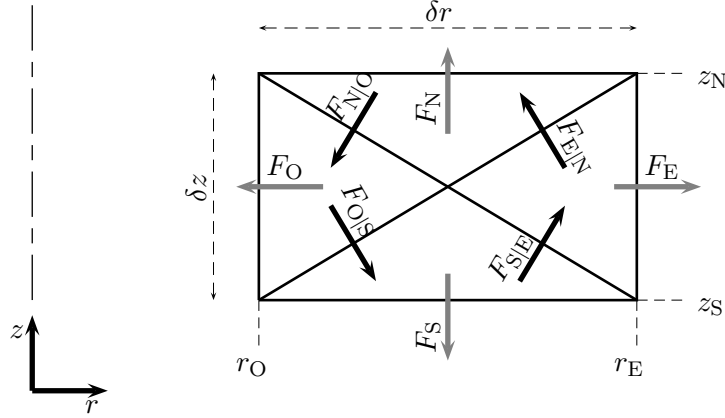


FIG. 2.4 – Rannacher-Turek element in axisymmetrical coordinates – Local notations for the definition of the interpolation field  $\mathbf{w}^K$ .

### 2.3.3 The discrete momentum balance equation

The discretization of (2.1b) is obtained by applying the above-defined convection operator to each component of the velocity and using the standard finite element technique to approximate the pressure gradient and viscosity terms. We thus obtain the following discrete momentum balance equation :

$$\forall \sigma \in \mathcal{E} \setminus \mathcal{E}_D, \text{ for } 1 \leq i \leq d,$$

$$\begin{aligned} \frac{|D_\sigma|}{\delta t} (\varrho_\sigma(\mathbf{u}^{(i)})_\sigma - \varrho_\sigma^*(\mathbf{u}^{(i)})_\sigma^*) + \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma, \varepsilon}(\mathbf{u}^{(i)})_\varepsilon + a_d(\mathbf{u}, \varphi_\sigma^{(i)}) \\ - \sum_{K \in \mathcal{T}} \int_K p^n \nabla \cdot \varphi_\sigma^{(i)} \, d\mathbf{x} = \int_\Omega \mathbf{f} \cdot \varphi_\sigma^{(i)} \, d\mathbf{x} + \int_{\partial\Omega_N} \mathbf{g} \cdot \varphi_\sigma^{(i)} \, d\sigma \end{aligned} \quad (2.24)$$

where, for any discrete velocity field  $\mathbf{v}$  and  $\mathbf{w}$  :

$$a_d(\mathbf{v}, \mathbf{w}) = \mu \sum_{K \in \mathcal{T}} \int_K [\nabla \mathbf{v} : \nabla \mathbf{w} + \frac{1}{3} \nabla \cdot \mathbf{v} \nabla \cdot \mathbf{w}] \, d\mathbf{x}$$

This relation is obtained by remarking that, since the viscosity  $\mu$  is supposed to be constant, the divergence of the stress tensor  $\boldsymbol{\tau}$  given by Equation (2.2) reads  $\nabla \cdot (\boldsymbol{\tau}(\mathbf{u})) = \mu \Delta \mathbf{u} + \mu/3 \nabla \nabla \cdot (\mathbf{u})$ .

### 2.3.4 Concluding remarks

**Extension to the MAC scheme.** The content of this section extends in a straightforward way to the MAC scheme [60, 44], and thus allows to build for this spatial approximation a discrete convection operator satisfying the kinetic energy theorem. A surprising consequence of the computation of the mass fluxes at the faces of the dual mesh (for the MAC scheme as well as for the finite element schemes considered here) is that the flux at a face  $\varepsilon$  included in the element  $K$  generally involves the value of the density in the neighbours of  $K$ .

**Comparison with the convection operator proposed in [98, 99].** In [98, 99], in the two-dimensional incompressible case and for the Crouzeix-Raviart element, the authors introduce a convection operator which is obtained by an upwind finite volume discretization based on the dual cell. This work is adapted for the Rannacher-Turek element in [106]. This operator is written in a non-conservative form, in the following sense. Let us assume that, for any diamond cell of the mesh  $D_\sigma$ , the following discrete divergence-free condition holds :

$$\sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} = 0 \quad (2.25)$$

Then a convection operator which, applied to a discrete function  $u$ , reads :

$$(Cu)_\sigma = \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} u_\varepsilon$$

can equivalently be written :

$$(Cu)_\sigma = \sum_{\varepsilon \in \bar{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon} (u_\varepsilon - u_\sigma)$$

The first form may be considered as a discretization of the convection term written in divergence form (*i.e.*, for a given transport field  $\mathbf{q}$  of fluxes  $F_{\sigma,\varepsilon}$ ,  $\nabla \cdot (u\mathbf{q})$ ), and the second one as a discretization of the same term in gradient form (*i.e.*  $\mathbf{q} \cdot \nabla u$ ). This is this latter form which is used in [98, 99, 106]. As a consequence, even if the divergence-free constraint (2.25) does not hold, the trilinear form associated to the convection term,  $b(\mathbf{q}, u, v) = \sum_\sigma (Cu)_\sigma v_\sigma$ , enjoys the same stability property that in the continuous case : for any discrete field  $u$ ,  $b(\mathbf{q}, u, u) = 0$  if  $u_\varepsilon$  is obtained by a centered approximation and  $b(\mathbf{q}, u, u) \geq 0$  in the upwind case.

In [98, 99, 106], the fluxes  $F_{\sigma,\varepsilon}$  are obtained from a direct interpolation of the transport field (*i.e.*, with the notations of the present paper, up to the multiplication by a constant density,  $\mathbf{w} = \mathbf{u}$ ). For the Crouzeix-Raviart discretization, the resulting convection operator thus coincides (for the incompressible case) with an upwind version of the operator proposed in this work. This is no more the case for the Rannacher-Turek element (the condition (2.25) is not fulfilled by the choice made in [106]).

## 2.4 A fractional step scheme

In this section, we address the solution of the full system (2.1). To this purpose, we build an incremental projection-like algorithm.



Let us consider a partition  $0 = t^0 < t^1 < \dots < t^N = T$  of the time interval  $(0, T)$ , which, for the sake of simplicity, we suppose uniform. Let  $\delta t$  be the constant time step  $\delta t = t^{n+1} - t^n$  for  $n = 0, 1, \dots, N - 1$ . In a semi-discrete time setting, the proposed algorithm consists in the following two step scheme :

With given data  $\varrho^{n-1}$ ,  $\varrho^n$ ,  $\varrho^{n+1}$ , and for known (from the previous time step)  $\mathbf{u}^n$  and  $p^n$ , and supposing that :

$$\frac{1}{\delta t}(\varrho^n - \varrho^{n-1}) + \nabla \cdot (\varrho^n \mathbf{u}^n) = 0 \quad (2.26)$$

1 - Solve for  $\tilde{\mathbf{u}}^{n+1}$  :

$$\frac{1}{\delta t}(\varrho^n \tilde{\mathbf{u}}^{n+1} - \varrho^{n-1} \mathbf{u}^n) + \nabla \cdot (\tilde{\mathbf{u}}^{n+1} \otimes \varrho^n \mathbf{u}^n) + \nabla p^n - \nabla \cdot [\tau(\tilde{\mathbf{u}}^{n+1})] = \mathbf{f}^{n+1} \quad (2.27)$$

2 - Solve for  $p^{n+1}$  and  $\mathbf{u}^{n+1}$  :

$$\frac{1}{\delta t}[\varrho^n (\mathbf{u}^{n+1} - \tilde{\mathbf{u}}^{n+1})] + \nabla(p^{n+1} - p^n) = 0 \quad (2.28a)$$

$$\frac{1}{\delta t}(\varrho^{n+1} - \varrho^n) + \nabla \cdot (\varrho^{n+1} \mathbf{u}^{n+1}) = 0 \quad (2.28b)$$

Step 1 consists in a semi-implicit solution of the momentum balance equation to obtain a predicted velocity. Its discretization is given by Relation (2.24), with  $\mathbf{u} = \tilde{\mathbf{u}}^{n+1}$ ,  $\mathbf{u}^* = \mathbf{u}^n$ ,  $p = p^n$ ,  $\varrho = \varrho^n$  and  $\varrho^* = \varrho^{n-1}$ .

Step 2 is a pressure correction step, which boils down to the usual projection step used in incompressible flow solvers when the density is constant (e.g. [84]). When the density varies, the pressure is determined by just solving one Poisson equation per time step, see [59]. The time-derivative term in (2.28a) is lumped, as in the discretization of the prediction step, and the gradient operator is also the same as in the prediction step. Note that  $\varrho^{n+1}$  and  $\varrho^n$  are known data, so that (2.28b) is actually a constraint for  $p$ . The mass balance (Equation (2.28b)) is discretized according to (2.15), with the mass fluxes given by (2.16), the density at the face being given by (2.17), with, for this equation,  $\varrho = \varrho^{n+1}$ ,  $\varrho^* = \varrho^n$  and  $\mathbf{u} = \mathbf{u}^{n+1}$ . At the discrete level, taking the divergence of (2.28a) and using (2.28b) to eliminate the unknown velocity  $\mathbf{u}^{n+1}$  yields a linear elliptic problem for the pressure, the operator of which looks like a finite volume diffusion operator (see [48, section 3.4] for a description of a similar computation). Once the pressure is computed, the first relation yields the updated velocity.

Finally, Equation (2.26) is just the mass balance equation taken at the previous times step.

It may be checked that the arguments of section 2.3 hold, so that the convection operator satisfies the stability stated in Theorem 2.3.1. In the incompressible case, the scheme may be shown to be unconditionnally stable, in the sense that the velocity satisfies discrete versions of the  $L^\infty(L^2)$  and  $L^2(H^1)$  *a priori* estimates of the continuous problem.

Writing the projection step without  $\rho$  leads to :

$$\frac{1}{\delta t}(\tilde{\mathbf{u}}^{n+1} - \mathbf{u}^n) + \nabla \cdot (\tilde{\mathbf{u}}^{n+1} \otimes \mathbf{u}^n) + \nabla p^n - \nabla \cdot (\tau(\tilde{\mathbf{u}}^{n+1})) = \mathbf{f}^{n+1} \quad (2.29)$$

$$\frac{1}{\delta t}(\mathbf{u}^{n+1} - \tilde{\mathbf{u}}^{n+1}) + \nabla p^{n+1} - \nabla p^n = 0 \quad (2.30)$$

$$\nabla \cdot (\mathbf{u}^{n+1}) = 0 \quad (2.31)$$

(2.30) gives :

$$\left(\frac{\mathbf{u}^{n+1}}{\delta t} + \nabla p^{n+1}\right)^2 = \left(\frac{\tilde{\mathbf{u}}^{n+1}}{\delta t} + \nabla p^n\right)^2 \quad (2.32)$$

that is :

$$\left(\frac{\mathbf{u}^{n+1}}{\delta t}\right)^2 - \frac{2}{\delta t} p^{n+1} \nabla \cdot (\mathbf{u}^{n+1}) + |\nabla p^{n+1}|^2 = \left(\frac{\tilde{\mathbf{u}}^{n+1}}{\delta t}\right)^2 - \frac{2}{\delta t} p^n \nabla \cdot (\tilde{\mathbf{u}}^{n+1}) + |\nabla p^n|^2 \quad (2.33)$$

The kinetic energy theorem states that :

$$\frac{1}{2\delta t} \int_{\Omega} (\tilde{\mathbf{u}}^{n+1})^2 - (\mathbf{u}^n)^2 dx \leq \int_{\Omega} \frac{1}{\delta t} ((\tilde{\mathbf{u}}^{n+1} - \mathbf{u}^n) + \nabla \cdot (\tilde{\mathbf{u}}^{n+1} \otimes \mathbf{u}^n)) \tilde{\mathbf{u}}^{n+1} dx \quad (2.34)$$

Combining (2.29) and (2.34) then leads to :

$$\frac{1}{2\delta t} \int_{\Omega} (\tilde{\mathbf{u}}^{n+1})^2 - (\mathbf{u}^n)^2 + \int_{\Omega} \nabla p^n \tilde{\mathbf{u}}^{n+1} - \nabla \cdot (\tau(\tilde{\mathbf{u}}^{n+1}) \tilde{\mathbf{u}}^{n+1}) \leq \int_{\Omega} f^{n+1} \tilde{\mathbf{u}}^{n+1} \quad (2.35)$$

and by (2.33) :

$$\int_{\Omega} \nabla p^n \tilde{\mathbf{u}}^{n+1} = - \int_{\Omega} p^n \nabla \cdot (\tilde{\mathbf{u}}^{n+1}) = \int_{\Omega} \frac{(\mathbf{u}^{n+1})^2}{2\delta t} + \frac{\delta t}{2} |\nabla p^{n+1}|^2 - \frac{(\tilde{\mathbf{u}}^{n+1})^2}{2\delta t} - \frac{\delta t}{2} |\nabla p^n|^2 \quad (2.36)$$

that is :

$$\int_{\Omega} \left( \frac{(\mathbf{u}^{n+1})^2}{2\delta t} - \frac{(\mathbf{u}^n)^2}{2\delta t} \right) dx + \int_{\Omega} \frac{\delta t}{2} (|\nabla p^{n+1}|^2 - |\nabla p^n|^2) - \nabla \cdot (\tau(\tilde{\mathbf{u}}^{n+1}) \tilde{\mathbf{u}}^{n+1}) \leq \int_{\Omega} f^{n+1} \tilde{\mathbf{u}}^{n+1} \quad (2.37)$$

Multiplicating by  $2\delta t$  and summing over time leads to the previously mentioned discrete estimates :

$$|u^N|^2 + \sum_{0..N} \delta t |\tilde{\mathbf{u}}^{n+1}|_{H^1}^2 \leq |f| |\tilde{\mathbf{u}}^{n+1}| + |\mathbf{u}^0|^2 \quad (2.38)$$

The stability of a similar scheme, even if more complex (in particular, involving an additional renormalization step for the pressure), is also proven in the compressible barotropic case in [48]. Note that these results also hold (with a simpler proof) for the semi-implicit coupled (*i.e.* without the prediction and pressure correction technique) scheme.

If an additional balance equation must be solved for another variable on which the density depends (as in Section 2.5.2), let say  $y$ , this is performed before the first step, with the following time discretization :

$$\frac{1}{\delta t} (\varrho^n y^{n+1} - \varrho^{n-1} y^n) + \nabla \cdot (\varrho^n y^{n+1} \mathbf{u}^n) - \nabla \cdot (\nabla y^{n+1}) = 0 \quad (2.39)$$

and with a standard finite volume scheme. In this case, for a convenient discretization of the convection term (upwind or MUSCL, for instance), the condition (2.26) allows to obtain a monotone scheme [79, 7].

## 2.5 Numerical experiments

Theoretical proofs of convergence for schemes similar to the proposed one, *i.e.* combining a Crouzeix-Raviart finite element discretization for the diffusion and a finite volume discretization for the convection, are available in some cases : incompressible stationary flows [98, 99], convection-diffusion equations [2, 31, 42]. However, the only analysis for Navier-Stokes equations [98, 99] is performed in the incompressible framework and for an upwind approximation for the convection, the resulting scheme being only first order in space. Our goal here is thus to perform numerical experimentations to check whether these convergence results extend in essentially three directions : obtaining, with the centered approximation, second order in space convergence for the velocity, in the unstationary case and for variable density flows. In addition, we also address some classical benchmarks to assert the robustness of the scheme. We begin with incompressible flows, then turn to variable density flows.

### 2.5.1 Incompressible flows

This section is devoted to the solution of unstationary incompressible Navier–Stokes equations :

$$\begin{aligned} \varrho [\partial_t \mathbf{u} + \nabla \cdot (\mathbf{u} \otimes \mathbf{u})] + \nabla p - \mu \Delta \mathbf{u} &= \mathbf{f} \\ \nabla \cdot \mathbf{u} &= 0 \end{aligned} \quad (2.40)$$

We first assess the accuracy of the scheme against a well-known analytical solution, namely the so-called Taylor-Green vortices, then compute various benchmarks of the literature : backward-facing step flows and flows behind obstacles.

#### Green–Taylor vortices

The so-called Green-Taylor vortex flows are a known solutions to System (2.40) posed on  $\hat{x}^2 \times [0, T]$  with a zero forcing term  $\mathbf{f}$ . The solution is given by :

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t) &= \begin{bmatrix} -\cos(2\pi \mathbf{x}_1) \sin(2\pi \mathbf{x}_2) \\ \sin(2\pi \mathbf{x}_1) \cos(2\pi \mathbf{x}_2) \end{bmatrix} \exp\left(-\frac{8\pi^2}{\mathbf{Re}} t\right) \\ p(\mathbf{x}, t) &= -\frac{\cos(4\pi \mathbf{x}_1) + \cos(4\pi \mathbf{x}_2)}{4} \exp\left(-\frac{16\pi^2}{\mathbf{Re}} t\right) \end{aligned} \quad (2.41)$$

The Reynolds number chosen here is  $\mathbf{Re} = 100$  ( $\varrho = 1$  and  $\mu = 0.01$ ), the computational domain is set to  $(0, 1/2)^2$ , and the final time is  $T = 1$ . The velocity is prescribed on the whole boundary, and the initial and boundary conditions are chosen to match the analytical solution.

For the Rannacher-Turek element, the domain is meshed by  $n \times n$  regular grids, with  $n = 16, 32, 64$  and  $128$ . For the Crouzeix-Raviart element, the meshes are obtained as follows : first, we build a regular grid ; then, this latter is perturbed by moving each inner vertex  $\mathbf{a}$  of the mesh to a random point of a circle centered on  $\mathbf{a}$  and of radius equal to the length of the smallest edge issued from  $\mathbf{a}$  multiplied by  $0.1$  ; finally, each cell of this mesh is cut along its diagonals in four triangles. The resulting mesh is quite general, in the sense that it does not enjoy any particular symmetry property. Four meshes are built, referred to by mesh 1, mesh 2, mesh 3 and mesh 4 and obtained from an initial  $16 \times 16, 32 \times 32, 64 \times 64$  and  $128 \times 128$  regular grid, respectively.

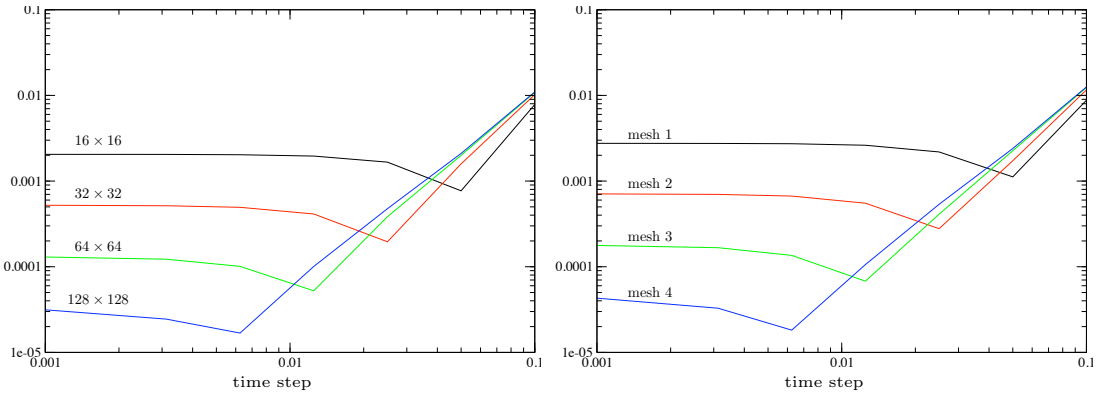


FIG. 2.5 – Green-Taylor vortices problem : time convergence in  $L^2$  norm for the velocity, for various meshes. *left* : quadrangles, *right* : triangles.

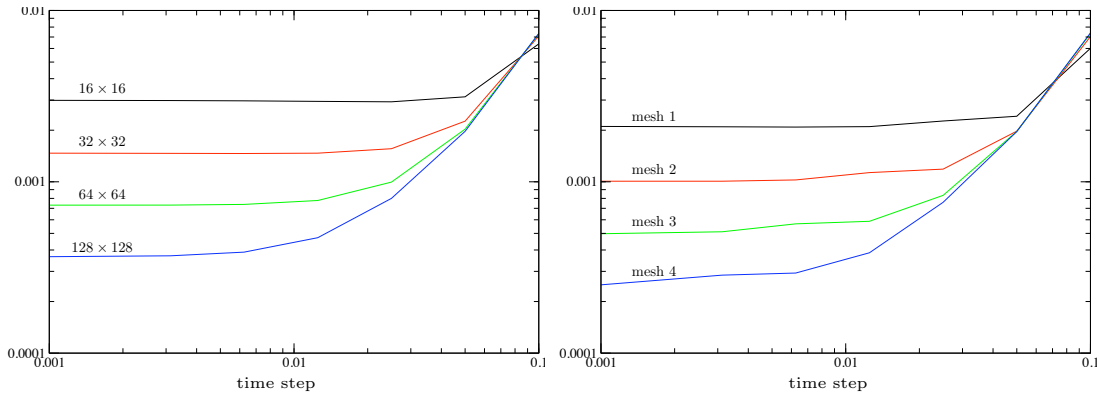


FIG. 2.6 – Green-Taylor vortices problem : time convergence in  $L^2$  norm for the pressure, for various meshes. *left* : quadrangles, *right* : triangles.

The  $L^2$  errors for the velocity and the pressure at  $t = 1$ , as a function of the time step and for the various considered meshes, are reported on Figure 2.5 and 2.6 respectively. From large to small time steps, curves first decrease, according to a surprising second-order time convergence rate. In fact, computations with the semi-implicit coupled scheme show that this latter is much more accurate at large time step; reported errors are thus essentially splitting errors, which indeed are known to behave as  $\delta t^2$  [58]. Then, at small time steps, a plateau is obtained, which corresponds to a second order for the velocity and first order for the pressure spatial error.

### Backward-facing step

We now address flows over a backward-facing step. The computational domain starts at the channel expansion, and the inlet velocity is prescribed, with a parabolic profile, at the upper part of height  $h$  of the left wall. The velocity is set to zero at the top and bottom walls, and a "do-nothing" (or homogeneous Neumann) boundary condition is imposed at the right-hand side of the domain. The expansion rate is  $H/h = 1.9423$ , and the domain length is set to  $L = 60$ .

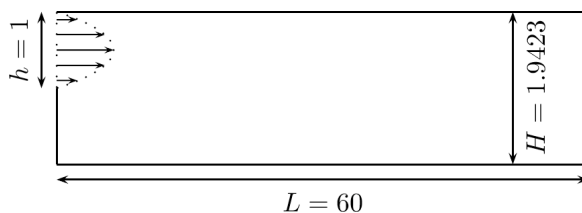
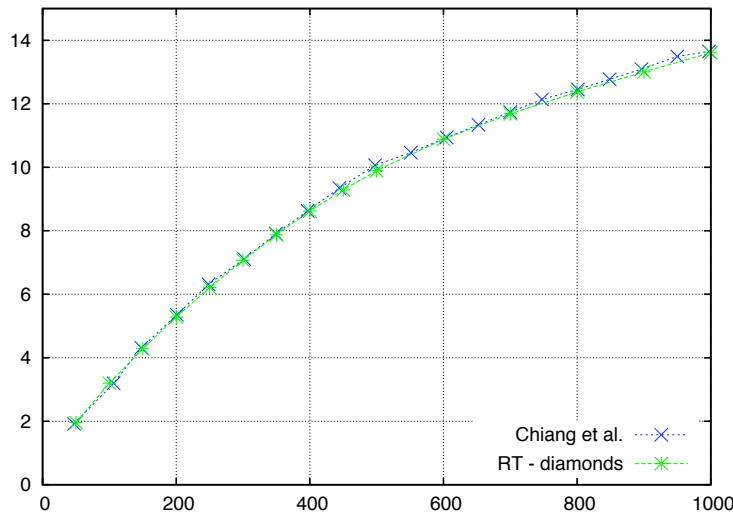


FIG. 2.7 – Computational domain for the backward-facing step flow.

The mesh is a regular uniform  $600 \times 291$  grid, and the computation is performed with the Rannacher-Turek element. The steady state is obtained by a fictitious transient, starting from a zero velocity.

This test, experimentally studied in [6], is commonly used to assert the accuracy of numerical schemes, thanks to the dependency of the reattachment length on Reynolds number. This latter quantity is defined here as  $\mathbf{Re} = \rho u_{\max} h / \mu$ , where  $u_{\max}$  is the maximum value of the velocity in the inlet section. On Figure 2.8, we compare the results of our computations for various Reynolds number (from  $\mathbf{Re} = 50$  to  $\mathbf{Re} = 1000$ ) with the numerical results published in [22]; a very good agreement is observed.

FIG. 2.8 – Computed recirculation length past a backward-facing step, as a function of the Reynolds number. Ordinate :  $x_R / (H - h)$ , where  $x_R$  is the reattachment length.

On Figure 2.9, we present the obtained transient development phase, for  $\mathbf{Re} = 2500$ , at the same times than [22, fig. 5]; once again, results compare satisfactorily.

### Flow past a cylinder

We address in this section some test cases which are part of a benchmark proposed in [100]. The first considered case is two-dimensional, the second one is three-dimensional. We give here

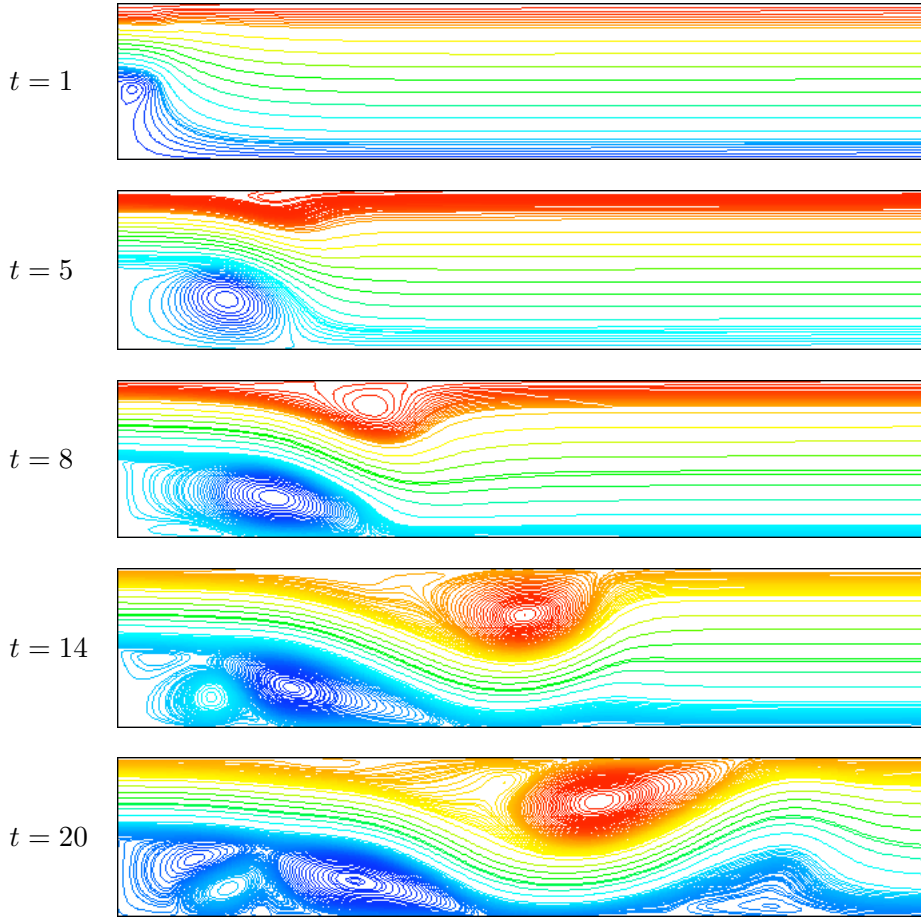


FIG. 2.9 – Streamlines past a backward-facing step for  $\mathbf{Re} = 2500$ , at different times.

a brief description of each case and refer to [100] for a complete presentation.

**2D flow,  $\mathbf{Re} = 100$**  This test corresponds to the 2D-2 case in [100]. The geometry for this test is sketched on figure 2.10. The fluid enters the domain on the left boundary, with an imposed velocity profile :

$$u_x(0, y) = 4u_m y \frac{H - y}{H^2}, \quad u_y(0, y) = 0$$

where  $H = 0.41m$  is the height of the channel and  $u_m = 1.5$  m/s; a zero velocity is prescribed at the other boundaries except for the right-hand side, where we use an inlet/outlet boundary condition which ensures the stability of the problem even in presence of inward velocities [4].

The density is  $\rho = 1$  and the viscosity is  $\mu = 0.01$ , so the Reynolds number, defined as  $\mathbf{Re} = \rho \bar{u} D / \mu$ , where  $D = 0.1$  is the diameter of the cylinder and  $\bar{u} = 2 u_x(0, H/2) / 3$ , is equal to 100.

A "coarse version" of the meshes used for the presented computation is sketched on Figure 2.11; real meshes are considerably refined with respect to this one, by diminishing the discretization step along the characteristic lines (the boundaries and the concentric circles around the cylinder).

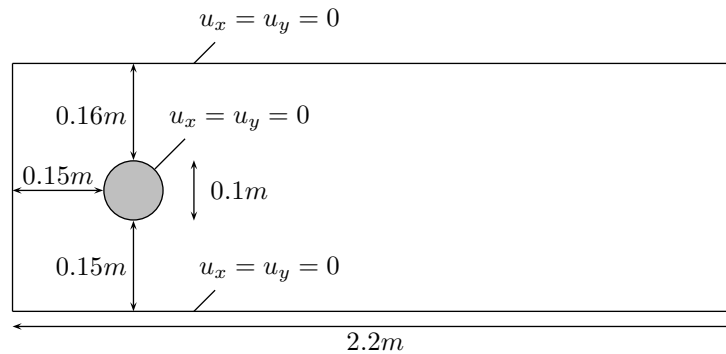


FIG. 2.10 – Geometry for the 2D flow around a cylinder.

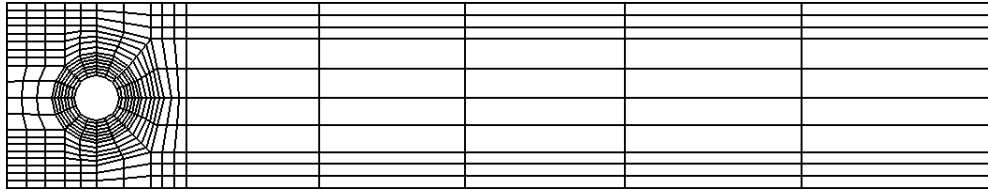
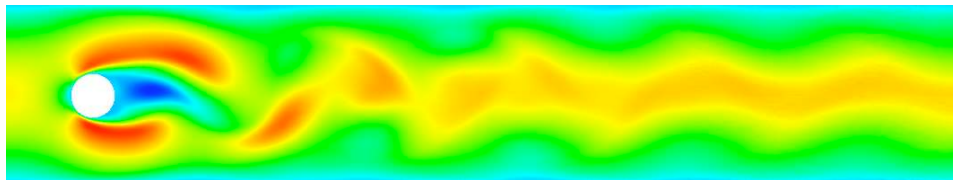


FIG. 2.11 – A "coarse version" of the 2D mesh.

The space discretization is performed with the Rannacher–Turek elements in their parametric variant. The cells are not rectangular, so the definition of the discrete convection operator given in the preceding section needs to be generalized : in fact, we simply keep for the expression of the mass fluxes at the dual faces the same linear combination of the fluxes on the primal faces as in the rectangular case (*i.e.* we suppose that Relation (2.22) applies). Since the integral over a deformed element  $K$  of a shape function associated to a face of  $K$  is no longer  $|K|/4$ , the geometrical interpretation of this term as a finite-volume discretization over a dual cell built from the diagonal lines of the primal mesh does not hold. However, the discrete trilinear form associated to the convection term is still antisymmetrical. Finally, note also that the deviation from the rectangular case attenuates as the mesh is refined.

In all our computations, the time step is  $\delta t = 5.10^{-4}$  s.

FIG. 2.12 – 2D flow with  $\mathbf{Re} = 100$ .

The flow is unsteady, and the main characteristic flow quantities quoted in [100] are the maximum drag coefficient  $c_{D\max}$ , the maximum lift coefficient  $c_{L\max}$ , the Strouhal number  $St$  and an instantaneous pressure difference  $\Delta P$  between the front and end points of the cylinder,

*i.e.* the points  $(0.15m, 0.20m)$  and  $(0.25m, 0.20m)$  (see [100, section 2.2] for a precise definition of these quantities). The obtained values for two space discretizations (computations C#1 and C#2) are gathered in the following table, together with the results of Rannacher and Turek using the same finite element as in the present work (computations RT#1 and RT#2), although in the non-parametric variant, and a plausible range for the results derived from the set of the contributions to the benchmark. Values entering this reference interval are typeset in bold in the table.

mesh	space unks	time steps	$c_{D_{\max}}$	$c_{L_{\max}}$	St	$\Delta P$
C#1	294 882	660	<b>3.2395</b>	0.9334	<b>0.3030</b>	2.4533
C#2	352 722	664	3.2511	<b>0.9983</b>	<b>0.3012</b>	<b>2.4961</b>
RT#1	167 232	188	3.2498	<b>1.0081</b>	0.2927	2.4410
RT#2	667 264	612	<b>3.2314</b>	<b>0.9999</b>	<b>0.2973</b>	<b>2.4707</b>
Reference range			3.22 – 3.24	0.99 – 1.01	0.295 – 0.305	2.46 – 2.50

The obtained results either enter the reference range or are very close to it, without too much refining the mesh (the finest one only leads to a little bit more than half the number of space unknowns used in computation RT#2).

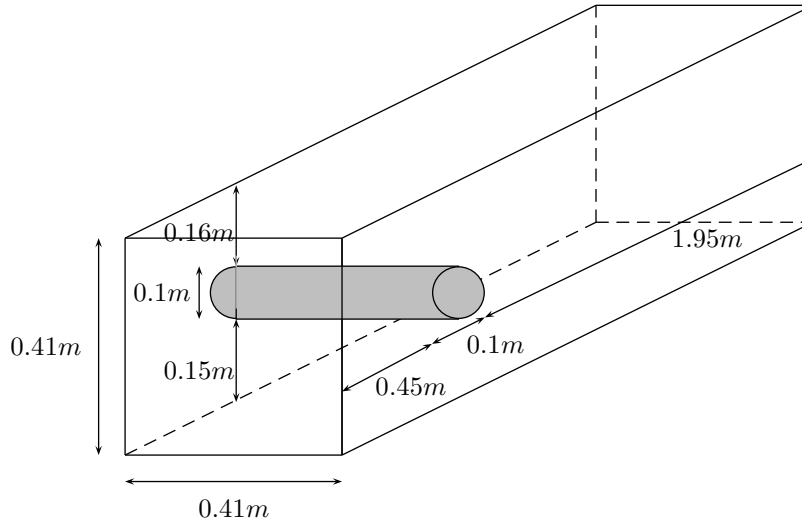


FIG. 2.13 – Geometry for the 3D flow around a cylinder.

**3D flow,  $\mathbf{Re} = 20$**  We now turn to the steady three-dimensional flow referred to as the 3D-1Z case in [100]. The geometry of the computational domain is sketched on Figure 2.13. The flow is governed by the system of equations (2.40). The inlet velocity profile is parabolic :

$$u_x(0, y, z) = 16u_m yz \frac{(H-y)(H-z)}{H^4}, \quad u_y(0, y, z) = u_z(0, y, z) = 0$$

with  $u_m = 0.45$  m/s. The viscosity of the fluid is  $\mu = 10^{-3}$ , and its density is  $\rho = 1$ , so the Reynolds number, defined as  $\mathbf{Re} = \rho \bar{u} D / \mu$  with  $\bar{u} = 4 u_x(0, H/2, H/2) / 9$ , is equal to 20.



The mesh is obtained by first adapting the mesh used for the previous case to triangulate a cut of the domain along a  $x, z$ -plane, and then building on this basis tree-dimensional cells by extrusion along the  $y$ -axis, with a uniform step. The resulting cells are thus hexahedra, and the space discretization is performed with the parametric version of the Rannacher-Turek element. The convection operator is obtained by extending the definition of Section 2.3 as previously.

As required in [100], we compute the drag coefficient  $c_D$ , the lift coefficient  $c_L$ , and the pressure difference  $\Delta P$  between the front  $((0.45, 0.20, 0.205))$  and end  $((0.55, 0.20, 0.205))$  points of the cylinder (see [100, section 2.3] for a precise definition of these quantities). The obtained values (computations C#3) are gathered in the following table, together with the results of some participants using the same finite element as in the present work (computations RT#3, RT#4 and RT#5), although in the non-parametric variant, and a plausible range for the results derived from the set of the contributions to the benchmark. The three computed characteristic quantities lie in these reference intervals.

mesh	space unks	$c_D$	$c_L$	$\Delta P$
C#3	2 271 870	<b>6.175</b>	<b>8.14E-03</b>	<b>0.1673</b>
RT#3	98 128	5.8431	6.10E-03	0.1482
RT#4	771 392	5.9731	5.90E-03	0.1605
RT#5	6 116 608	<b>6.1043</b>	7.90E-03	<b>0.1672</b>
Exact range		6.05 – 6.25	0.008 – 0.01	0.165 – 0.175

## 2.5.2 Variable density flows

In this section, we consider the system of equations governing a two-component flow :

$$\begin{aligned}
 \partial_t \varrho + \nabla \cdot (\varrho \mathbf{u}) &= 0 \\
 \partial_t (\varrho \mathbf{u}) + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \nabla \cdot \boldsymbol{\tau}(\mathbf{u}) &= \mathbf{f} \\
 \partial_t (\varrho y) + \nabla \cdot (\varrho y \mathbf{u}) - \lambda \Delta y &= g
 \end{aligned} \tag{2.42}$$

where the density  $\varrho$  is given as a function of the unknown  $y$  by :

$$\varrho = \varrho(y) = \frac{1}{\frac{y}{\varrho_1} + \frac{1-y}{\varrho_2}} \tag{2.43}$$

The component densities  $\varrho_1$  and  $\varrho_2$  are supposed to be two constant positive real numbers.

Let us suppose that the velocity is prescribed to zero on the whole boundary of the computational domain  $\partial\Omega$ . In this condition, integrating the mass balance in (2.42) over  $\Omega$  yields the total mass balance :

$$\frac{d}{dt} \int_{\Omega} \varrho \, d\mathbf{x} = 0 \tag{2.44}$$

On the other side, the density  $\varrho$  is given as a function of  $y$ , itself solution of the third equation of (2.42). Integrating this relation over  $\Omega$  and supposing that the diffusion flux of the mass of component 1 vanishes at the boundary, *i.e.*  $\lambda \nabla y \cdot \mathbf{n} = 0$  on  $\partial\Omega$ , we obtain :

$$\frac{d}{dt} \int_{\Omega} \varrho y \, d\mathbf{x} = 0 \tag{2.45}$$

By an easy manipulation of the equation of state (2.43), we get :

$$\varrho = \varrho_2 + \left(1 - \frac{\varrho_2}{\varrho_1}\right) \varrho y$$

and thus, since  $\varrho$  appears as an affine function of  $\varrho y$ , the relations (2.44) and (2.45) are fortunately compatible. With the proposed time-stepping procedure, this property does not hold anymore, because of the time shift of the density in the computation of  $y$ , which is done by solving (2.39) at the first step of the algorithm. Thus a renormalisation of the density is necessary to ensure the existence of a solution to the projection step :

$$\varrho^{n+1} = \frac{\int_{\Omega} \varrho(y^{n+1}) \, d\mathbf{x}}{\int_{\Omega} \varrho^n \, d\mathbf{x}} \varrho(y^{n+1})$$

Note that this relation is reminiscent of the scaling of the density obtained through its dependency versus the so-called thermodynamical pressure in the asymptotic model for low Mach number flows [82].

### Convergence to an analytical solution

In this section, we choose the following solution to system (2.42) :

$$\begin{aligned} \varrho(\mathbf{x}, t) &= 1 + \frac{1}{4} \sin(\pi t) [\cos(\pi \mathbf{x}_1) + \cos(\pi \mathbf{x}_2)] \\ \varrho(\mathbf{x}, t) \mathbf{u}(\mathbf{x}, t) &= -\frac{1}{4} \cos(\pi t) \begin{bmatrix} \sin(\pi \mathbf{x}_1) \\ \sin(\pi \mathbf{x}_2) \end{bmatrix} \\ p(\mathbf{x}, t) &= \sin(\pi t) [\sin(\pi \mathbf{x}_1) + \sin(\pi \mathbf{x}_2)] \end{aligned}$$

We suppose that  $\varrho_1 = 0.25$  and  $\varrho_2 = 2$ , so relation (2.43) yields  $y = (2 - \varrho)/(7\varrho)$ . The viscosity is supposed to be  $\mu = 0.01$  and, in the third relation of (2.42), we choose  $\lambda = 0$ . The computational domain is  $\Omega = (0, 1) \times (0, 1)$ , so the normal velocity is always zero at the boundary, and the integral of the density over  $\Omega$  does not vary with time. With this choice for  $\varrho$  and  $\varrho \mathbf{u}$ , the mass balance (first relation of (2.42)) is verified. The right-hand side in the momentum balance and in the  $y$  transport equation, the initial conditions and the boundary conditions (prescribed value for  $\mathbf{u}$ ) are chosen to match the analytical solution.

The domain is meshed by  $n \times n$  structured grids, with  $n = 16, 32, 64$  and  $128$ . The discretization of the convection term for  $y$  is performed by an upwind finite volume scheme. Errors obtained at  $t = 0.5$  are displayed on figures 2.14 and 2.15, in  $L^2$  norm for the velocity and the pressure, and, for  $y$ , in the following (usual in the finite volume) context discrete  $L^2$  norm :

$$|y|_{L_h^2} = \sum_{K \in \mathcal{T}} |K| y(\mathbf{x}_K)^2$$

where  $\mathbf{x}_K$  is the mass center of the cell  $K$ . Curves show a decrease at large time step which correspond to an approximate first order in time convergence, and then a plateau, the value on which corresponds to a second order space convergence for the velocity and a first order space convergence for the pressure.

For simplicial meshes with a Crouzeix-Raviart element discretization, the same behaviour is observed.

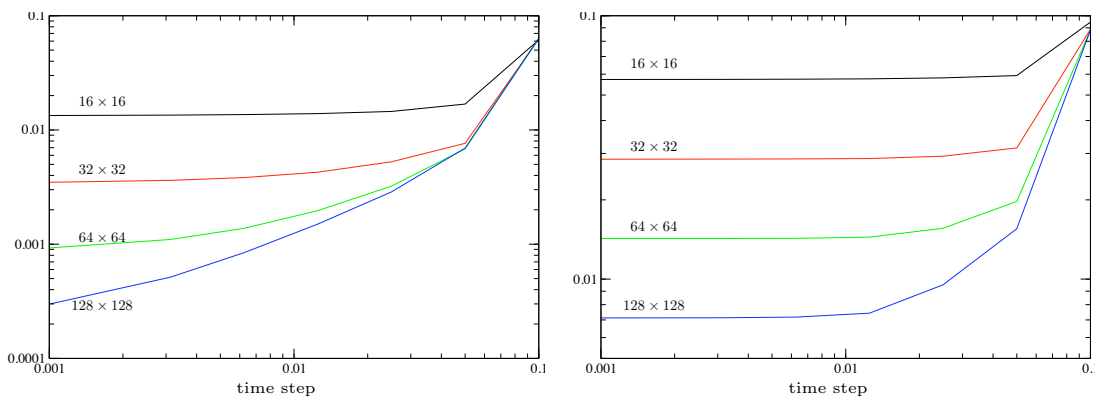


FIG. 2.14 – Variable density analytical solution : time convergence in  $L^2$  norm for the velocity and the pressure, for various meshes.

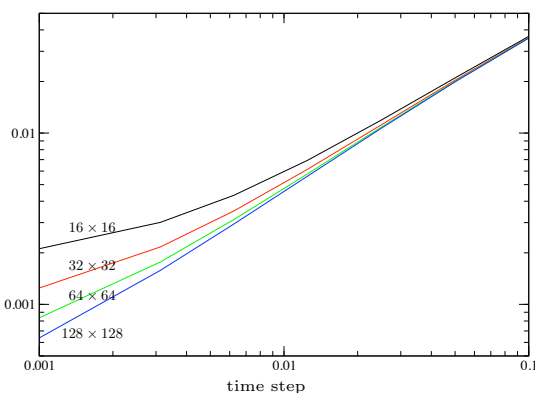


FIG. 2.15 – Variable density analytical solution : time convergence in discrete  $L^2$  norm for the  $y$  variable, for various meshes.

### A Rayleigh-Taylor instability flow

We now address a case studied in [105, 45], consisting in a Rayleigh-Taylor instability flow. Since we take  $\lambda = 0$  and  $g = 0$ , we get from (2.42) that  $\partial_t \rho + \mathbf{u} \nabla \rho = \mathbf{0}$  and  $\nabla \cdot \mathbf{u} = \mathbf{0}$ , so that the model equations are the same as those of [45]. Hence the data are exactly the same as in the case referred to by " $\mathbf{Re} = 1000$ " in [45]. The chosen mesh is the same as in [45] for the finite volume computation, namely a  $256 \times 512$  uniform grid. A MUSCL technique with a Van Leer limiter is implemented for the solution of the balance equation for  $y$ .

Results are plotted on Figure 2.16, and seem to be quite close to those presented in [45, Figure 1, p.898].

## 2.6 Conclusion

We have presented in this paper a discrete operator for the approximation with low-order non-conforming finite element spaces of the convection terms in Navier-Stokes equation in va-

riable density flows. This operator is built by a finite volume technique, based on a dual mesh. We prove that this operator satisfies a discrete counterpart of the the kinetic conservation identity. This stability property has been observed to greatly improve the robustness of computations, specifically with relatively coarse meshes as often encountered in real-life applications. In addition, the assembling cost of this operator is low, and it does not extend the stencil of the scheme beyond the stencil of the diffusion terms.

This discretization is now routinely used in simulations performed with the ISIS [70] code, implemented on the basis of the software component library PELICANS [92], both freewares being developped at the French Institut de Radioprotection et de Sûreté Nucléaire (IRSN). It is one of the ingredient of entropy preserving schemes for the simulation of monophasic [48] or diphasic [51] compressible barotropic flows. The assesment of its efficiency in the context of Large Eddy Simulations is foreseen in a next future.

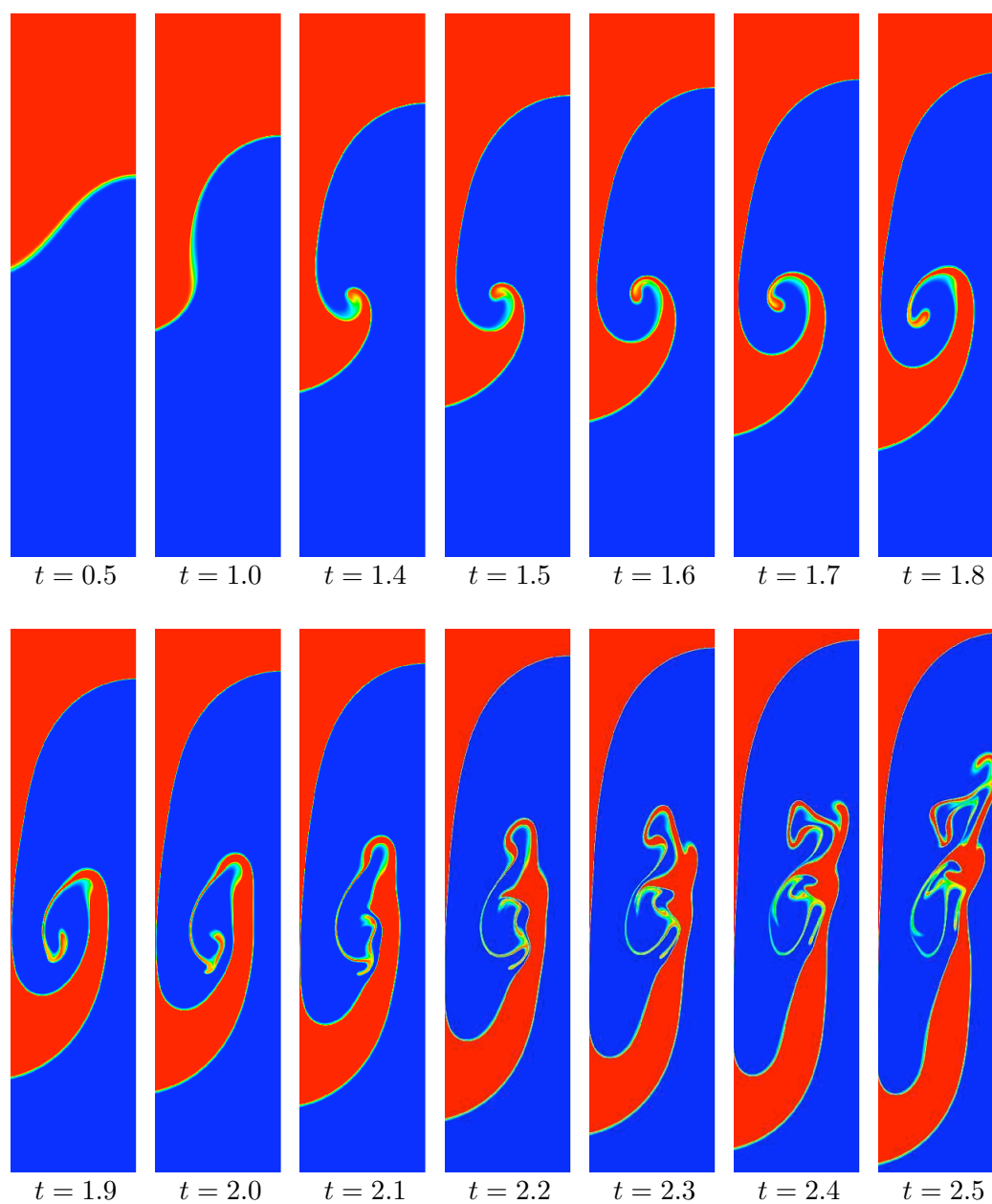


FIG. 2.16 – Density in a Rayleigh–Taylor instability flow.

## Chapitre 3

# Approximation volumes finis de l'advection–diffusion de scalaires sur maillages quelconques

### 3.1 Introduction

In this paper, we address the convection–diffusion problem, find a scalar function  $\phi(\mathbf{x}, t)$  defined over  $\Omega \times [0, T]$  such that :

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \nabla \cdot [\mathbf{u}(\mathbf{x})\phi(\mathbf{x}, t) - \kappa \nabla \phi(\mathbf{x}, t)] = f(\mathbf{x}, t) \quad \text{on } \Omega \times [0, T] \quad (3.1a)$$

$$\phi(\mathbf{x}, t) = \phi_D(\mathbf{x}, t) \quad \text{on } \partial\Omega_D \times [0, T] \quad (3.1b)$$

$$-\kappa \nabla \phi(\mathbf{x}, t) \cdot \mathbf{n} = g(\mathbf{x}, t) \quad \text{on } \partial\Omega_N \times [0, T], \quad (3.1c)$$

$$\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x}) \quad \text{in } \Omega, \quad (3.1d)$$

where :

- $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , is an open connected and bounded domain, with a piecewise continuously differentiable boundary,
- $\mathbf{u}(\mathbf{x})$  is the divergence-free convective velocity, where the diffusion coefficient tensor function  $\kappa$  verifies an uniform ellipticity condition,
- $f(\mathbf{x}, t) \in L^2(\Omega)$  is the source term and  $L^2(\Omega)$  is the space of square integrable functions,
- $\mathbf{n}$  is the unit normal vector to the boundary, outward to  $\Omega$ ,
- $\partial\Omega_D$  and  $\partial\Omega_N$  are the Dirichlet and Neuman boundaries, which form a partition of  $\partial\Omega$ , and are such that the measure of  $\partial\Omega_D$  is positive.
- $\phi_0$ ,  $\phi_D$ ,  $g$  denote the initial value, the Dirichlet boundary value (on the downwind boundary) and the Neumann boundary value (on the upwind boundary), respectively.

Under classical regularity assumptions on the problem data and if the Dirichlet part of the boundary is not empty, the system of equations (3.1a)–(3.1d) admits a unique solution [81] and this solution fulfills a weak maximum principle [53] so that it reaches its extrema on the domain boundary  $\partial\Omega$ .

The discretization of such an equation on usual regular grids has been the subject of numerous works, see e.g. the review [86]. Our aim in the present work is to test some schemes which can

be used on any type of two or three dimensional grids, including non conforming grids such as those generated by local refinement. Furthermore, the scheme is required to :

- (i) be stable for convection dominated regimes without requiring low order approximation of convective fluxes on sufficiently regular solutions,
- (ii) integrate higher order approximation of diffusive and convective fluxes,
- (iii) be robust in regard to the discrete maximum principle (discrete maximum principle).

This last point is especially important when considering coupled problems where some of the equations (e.g. equations of state) are only defined for certain ranges of values of the variables : the approximate solutions should belong to that range. In a very recent work, *Manzini et al.* have proposed a methodology that meets very similar specifications for the 2D space dimension, [83].

The Galerkin method naturally applies to unstructured grids, allows high order approximation of the diffusive fluxes but only fulfills the discrete maximum principle for the Laplace equation on specific grids, [112]. The classical spurious oscillations of the solution in the convective dominant regime or in presence of sharp boundary layers have motivated several variants of the Galerkin method, among which we can cite the GLS, SUPG or VMS methods [86, 97, 73, 103]. But only a few of those also fulfill the discrete maximum principle. However, in order to improve the spatial convergence order without losing the discrete maximum principle, non-linear terms must be added to the Galerkin formulation [77, 17]. Our experience with those methods is that the additional terms they involve must be implicit and are stiff enough to deteriorate the convergence rates of classical linearization algorithms, thus preventing their use in industrial applications.

Finite volume based methods have been widely used to approximate conservation (hyperbolic) laws due to their intrinsic local conservation property. To this extend, monotony-preserving methods have been developed that ensure both the second order approximations of the convective fluxes and the fulfillment of the discrete maximum principle, at least on structured grids. Some of those methods are variants of the Monotone Upstream-centered Schemes for Conservation Laws (MUSCL) of *Van Leer* [107]. They involve an upwind approximation of the convective fluxes at the control volume edges but, contrary to Godunov methods, this approximation uses a piecewise linear reconstruction of the solution at the edge center. To respect the monotony criterion, the gradient of the reconstruction is corrected using slope limiters [95]. Some higher order methods exist based (or not) on extended-stencils [10]. In his reference work, *Hubbard* [69] developed a variant of the MUSCL method for 2D unstructured grids. Several works have been proposed to extend this method to 3D unstructured and non-conforming grids imposing more or less strong assumptions on the mesh topology [93, 43, 109, 110, 25, 27, 26].

Contrary to the variants of the Galerkin method, a key issue for finite volume methods is the approximation of diffusive fluxes on unstructured grids. Indeed, the natural approximation of the gradient at the side intersection of two control volumes, that only involves the average values of the solution in these two control volumes and the signed distance between their circumcenters for simplexes or cuboids, imposes strong assumptions on the mesh topology to get both a coercive form and a consistent gradient approximation, [35]. For instance, a simplicial mesh must fulfill at least a weak Delaunay condition, which ensures that this signed distance remains positive ; however, under this condition, the circumcenter of the boundary cells may be outside the computational domain. This leads to negative transmissibilities for the diffusive fluxes at the corresponding boundary sides and thus to possible coercivity losses. Most of the unstructured meshing software can generate Delaunay grids on enough regular geometrical domains using

Voronoi diagrams but according to our experience with some of them [67, 52, 5], they often generate boundary cells with circumcenters outside the domain. Moreover, to our knowledge, no meshing software can generate admissible grids composed of quadrilaterals or hexahedra. Many methods that tackle this issue, including various mixed finite element/finite volume methods, are available in the literature. The approximation of anisotropic diffusive fluxes on unstructured non-conforming grids has been recently the subject of a benchmark for which the results obtained with various cell centred schemes, control volume finite element schemes, discontinuous Galerkin schemes, discrete duality finite volume schemes, lattice Boltzmann schemes, mixed or hybrid methods and non-linear methods have been compared to the classical Galerkin methods, see [68].

In the present paper, we propose a strategy that meets the above mentioned specifications. It is based on the Scheme Using Stabilization Hybrid Interfaces (SUSHI) proposed by *Eymard et al.* in [38] to solve diffusion problems and on a variant of the MUSCL method using a multislope approximation of the convective fluxes inspired by the work of *Clauzon et al.*, [25, 27]. The outline of the paper is the following. In Section 3.2, we introduce notations adopted throughout this paper and we define a generic form of the finite volume approximation. Sections 3.3 and 3.4 are devoted to the definition of approximations of the convective and the diffusive fluxes, respectively. In Section 3.6, the verification tests for each approximation as well as the benchmark of convection-diffusion equation on unstructured non-conforming 2D/3D grids are presented.

## 3.2 Notations

Let us start by some notations for the discretization mesh. A finite volume discretization of  $\Omega$ , denoted  $\mathcal{D}$ , is defined by a triplet  $\mathcal{D} = (\mathcal{T}, \mathcal{E}, \mathcal{P})$ , where :

- $\mathcal{T}$  is the set of control volumes  $K$  (also called cells).  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} \bar{K}$ . Herewith, we consider that each control volume is a non-empty convex open disjoint subset of  $\Omega$ .
- $\mathcal{E}$  is the set of edges (in 2D) or faces (in 3D), denoted by  $\sigma$  and such that  $\text{meas}(\sigma) > 0$  for any  $\sigma \in \mathcal{E}$ . We denote by  $\mathcal{E}_K \subset \mathcal{E}$  the set of faces of  $K \in \mathcal{T} : \partial K = \bigcup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$ ,  $\mathcal{E}_{\text{ext}}$  (resp.  $\mathcal{E}_{\text{int}}$ ) the set of boundary (resp. interior) edges of  $\Omega$ . Each  $\sigma \in \mathcal{E}_{\text{int}}$  has exactly two neighbouring cells, say  $K$  and  $L \in \mathcal{T}$ , and  $\bar{K} \cap \bar{L} = \bar{\sigma}$  which we also write :  $\sigma = K|L$ . The normal vector to a face  $\sigma$  of  $K$  outward to  $K$  is denoted by  $\mathbf{n}_{K,\sigma}$ .
- $\mathcal{P}$  is a set of points of  $\Omega$   $\mathcal{P} = (\mathbf{x}_K)_{K \in \mathcal{T}}$  and  $\mathbf{x}_K$  is a point of  $K$ , which we shall call the center (usually the circumcenter or the barycenter). In the following, we denote by  $\mathbf{x}_\sigma$  the barycenter of the face  $\sigma$  and by  $\mathbf{y}_\sigma$  the intersection of  $\sigma$  with the line passing by the centers of the neighbouring cells : in general,  $\mathbf{y}_\sigma \neq \mathbf{x}_\sigma$ , , and in fact one may even have  $\mathbf{y}_\sigma \notin \sigma$  : see Figure 3.1.

Moreover, we shall use the following notations :

- $|\mathbf{x}\mathbf{y}|$  denotes the usual Euclidean distance between two points  $\mathbf{x}$  and  $\mathbf{y}$  of  $\mathbb{R}^2$  or  $\mathbb{R}^3$ . Hereafter, we assume that the distance  $|\mathbf{x}_K\mathbf{y}_\sigma|$  between a center point  $\mathbf{x}_K$  of any cell  $K$  and a face  $\sigma$  of  $K$  is positive.
- $D_{K,\sigma}$  denotes the cone (diamond) with basis  $K|L$  and vertex  $\mathbf{x}_K$ ,  $|K|$  for the cell volume and  $|\sigma|$  for the area of the face  $\sigma$ .
- $\delta t$  is the time step, assumed to be constant for simplicity.
- $H_{\mathcal{T}}$  is the space of piecewise constant functions on the cells of the mesh.
- $X_{\mathcal{T}}$  is the set of families of values on the cells and on the Dirichlet boundary edges :



$X_{\mathcal{T}} = \mathbb{R}^T \times \mathbb{R}^{\mathcal{E}_D}$ , where  $\mathcal{E}_D$  denotes the set of edges that are included in  $\partial\Omega_D$  ; .

In order to introduce the scheme, we begin by a semi-implicit Euler semi-discretization in time. We denote by  $\phi^n$  a function of  $H^1(\Omega)$  (that is a square integrable function with square integrable derivative), which is expected to be an approximation of the solution to a weak solution of Problem (3.1) at time  $t_n = n\delta t$  :

Find  $\phi^{n+1} \in H^1(\Omega)$  such that  $\phi^{n+1} = \phi_D$  a.e. on  $\partial\Omega_D$  and

$$\begin{aligned} \int_{\Omega} \frac{\phi^{n+1} - \phi^n}{\delta t} \psi \, d\mathbf{x} + \int_{\Omega} \operatorname{div}(\mathbf{u}\phi^n) \psi \, d\mathbf{x} + \int_{\Omega} \kappa \nabla \phi^{n+1} \cdot \nabla \psi \, d\mathbf{x} \\ = \int_{\Omega} f(\mathbf{x}, t^{n+1}) \psi \, d\mathbf{x} - \int_{\partial\Omega_N} g \psi \, d\gamma(\mathbf{x}), \text{ for any } \psi \in H_D^1(\Omega), \end{aligned} \quad (3.2)$$

where  $H_D^1(\Omega)$  is the space of  $H^1(\Omega)$  functions such that  $\phi = 0$  a.e. on  $\partial\Omega_D$ .

In the next section, we define a discrete divergence operator using a MUSCL type scheme, which will be used to approximate the convective term  $\int_{\Omega} \operatorname{div}(\mathbf{u}\phi^n) \psi \, d\mathbf{x}$  of (3.2); in Section 3.4, we describe a discrete gradient which is used to approximate the diffusive term  $\int_{\Omega} \kappa \nabla \phi^{n+1} \cdot \nabla \psi \, d\mathbf{x}$  of (3.2). In these two sections, for readability reasons, the exponents referring to time will be omitted and the solution must be understood as the solution at time  $t^{n+1}$ . The complete discrete scheme is then presented in Section 3.5.

### 3.3 Discretization of the convection operator

The aim of this section is to construct a discrete divergence operator to approximate the convective term  $\int_{\Omega} \operatorname{div}(\mathbf{u}\phi^{n+1}) \psi \, d\mathbf{x}$  of (3.2). Mimicking the Stokes formula, we write the integral of the discrete divergence, which is constant on each cell, as the sum of discrete fluxes :

$$|K| \operatorname{div}_{\mathcal{T}}(\mathbf{u}\phi_{\mathcal{T}})|_K = \sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma}(\mathbf{u}, \phi_{\mathcal{T}}) \quad (3.3)$$

where  $F_{K,\sigma}(\mathbf{u}, \phi_{\mathcal{T}}^{n+1})$  is an approximation of  $\int_{\sigma} \mathbf{u}(\mathbf{x})\phi(\mathbf{x}, t) \cdot \mathbf{n}_{\sigma} \, d\sigma$ . The construction of  $F_{K,\sigma}(\mathbf{u}, \phi_{\mathcal{T}})$  is based on the MUSCL technique. It is obtained thanks to the reconstruction of two interface values  $\phi_{KL}$  and  $\phi_{LK}$  (one on each cell  $K$  and  $L$  neighboring the interface), and an upwind choice of the value in the flux.

$$F_{K,\sigma}(\mathbf{u}, \phi_{\mathcal{T}}) = \begin{cases} |\sigma| \phi_{KL} \mathbf{u} \cdot \mathbf{n}_{K,\sigma} & \text{if } \mathbf{u} \cdot \mathbf{n}_{K,\sigma} > 0, \\ |\sigma| \phi_{LK} \mathbf{u} \cdot \mathbf{n}_{K,\sigma} & \text{if } \mathbf{u} \cdot \mathbf{n}_{K,\sigma} < 0, \end{cases} \text{ for } \sigma = K|L \in \mathcal{E}_{\text{int}}. \quad (3.4)$$

We seek a linear reconstruction for  $\phi_{KL}$  to enable second order accuracy in space for smooth solutions : A generic form for  $\phi_{KL}$  reads as follows :

$$\phi_{KL} = \phi_K + \pi_{KL} |\mathbf{x}_K \mathbf{z}_{\sigma}| \quad (3.5)$$

where  $\pi_{KL}$  and  $\mathbf{z}_{\sigma}$  are respectively a constant slope and a physical point of  $\sigma$  to be defined. It is wellknown that the slope must be limited to prevent spurious oscillations of the approximate

solution when steep gradients occur. Indeed, it is easily shown that the explicit MUSCL scheme is  $L^\infty$  stable if the linear reconstruction  $\phi_{KL}$  satisfies

$$\phi_{KL} - \phi_K = \alpha_{KL}(u_L - u_K), \text{ and } \phi_{KL} - \phi_K = \sum_{\substack{\sigma' \in \mathcal{E}_K \\ \sigma' = K|M \neq \sigma}} \alpha_{KM}(u_K - u_M) \text{ for any } \sigma = K|L,$$

where  $\alpha_{KL}$  and  $\alpha_{KM}$  are positive numbers (note that these conditions were introduced in [20, Conditions (4.12) et (4.13) p. 73] for the convergence study of the MUSCL schemes).

The MUSCL methods differ by the way the slopes  $\pi_{KL}$  are built, by the choice the physical point  $\mathbf{x}_\sigma$  on  $\sigma$ , and by the limiting method to ensure these latter conditions,. In this paper, we introduce three MUSCL method variants, based on the works of *Buffard & Clain* [15] and *Clauzon* [27]; their general characteristics are :

- *the intersection method* is used on conforming simplicial grids, in the case where  $\mathbf{y}_\sigma \in \sigma$ , which is true if the edge or face  $\sigma = K|L$  intersects the line segment  $\mathbf{x}_K\mathbf{x}_L$  where  $\mathbf{x}_K$  and  $\mathbf{x}_L$  are the centres of the neighbouring cells  $K$  and  $L$ . The point  $\mathbf{z}_\sigma$  is then chosen as the intersection point :  $\mathbf{z}_\sigma = \mathbf{y}_\sigma$ , see Fig. 3.1, left side.
- *the center method* is applicable on conforming simplicial grids, whenever the edge or face  $\sigma$  does not intersect the line segment  $\mathbf{x}_K, \mathbf{x}_L$  such as in Fig. 3.1, right side. In this case, the point  $\mathbf{x}_{KL}$  is chosen as the barycenter of the face  $\sigma$  :  $\mathbf{z}_\sigma = \mathbf{x}_\sigma$ .
- *the modified intersection/center method* is designed for general grids, which may be non simplicial, but also non matching.

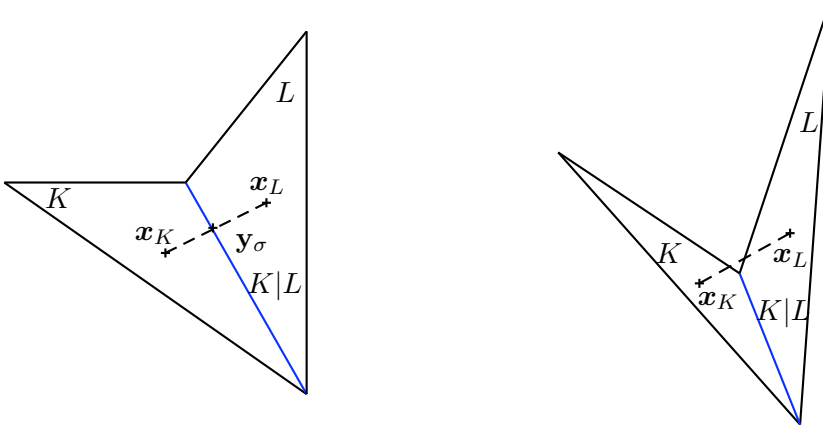


FIG. 3.1 – Triangles meeting (left) or not (right) the condition for the existence of an intersection between the line segment  $x_K x_L$  and the edge  $\sigma$ .

Note that in all these methods, the center  $\mathbf{x}_K$  is always the barycenter of the cell  $K$  and that it has to be strictly inside the volume delimited by centers  $\mathbf{x}_L$  of all adjacent control volumes  $L$  of the cell  $K$ , see Fig. 3.2.

### 3.3.1 The intersection multislope MUSCL method

Using (3.5), one reconstructs  $\phi_{KL}$  at the intersection  $\mathbf{y}_\sigma$  between  $\sigma$  and the line segment  $\mathbf{x}_K\mathbf{x}_L$ , see Fig. 3.3 (Left) with a slope  $\pi_{KL} = p_{KL}$  determined as follows : we first introduce a

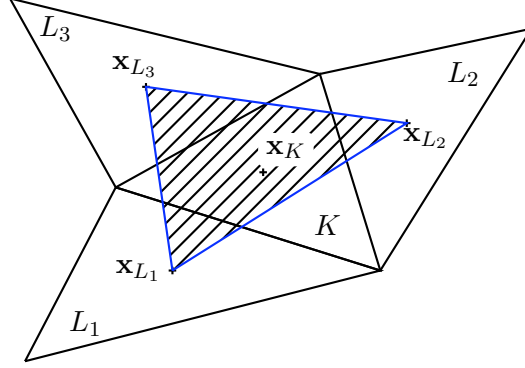


FIG. 3.2 – The center  $\mathbf{x}_K$  of the control volume  $K$  has to be strictly inside the volume delimited by centers  $\mathbf{x}_L$  of all adjacent control volumes  $L$  of  $K$ .

“natural” slope :

$$p_{KL}^+ = \frac{\phi_L - \phi_K}{|\mathbf{x}_K \mathbf{x}_L|}. \quad (3.6)$$

which we shall call “downward” in the following. We also use an “upward” slope which is reconstructed on the cell  $K$  from a linear combination of downward slopes of neighbouring cells in a way we now proceed to explain : we first note that, provided that there is no degenerate element in  $\mathcal{T}$ , the normalized downward slope directions  $\mathbf{t}_{KL} = \mathbf{x}_K \mathbf{x}_L / |\mathbf{x}_K \mathbf{x}_L|$  for all  $\sigma = K|L \in \mathcal{E}_K$ , can be decomposed as follows :

$$\mathbf{t}_{KL} = \sum_{\substack{M:K|M \in \mathcal{E}_K \\ M \neq L}} \beta_{KLM} \mathbf{t}_{KM} \quad (3.7)$$

with constants  $\beta_{KLM} \in \mathbb{R}$ , see Fig. 3.3 (Right). These constants can be inferred from the barycentric coordinates  $\lambda_{KL}$  of  $\mathbf{x}_K$  in the system of the adjacent cells barycenters  $\mathbf{x}_L$ , that is :

$$\mathbf{x}_K = \sum_{\sigma=K|L \in \mathcal{E}_K} \lambda_{KL} \mathbf{x}_L$$

with  $\forall K, \sum_{\sigma=K|L \in \mathcal{E}_K} \lambda_{KL} = 1$ . We then write that

$$\beta_{KLM} = -\frac{\lambda_{KM} |\mathbf{x}_K \mathbf{x}_M|}{\lambda_{KL} |\mathbf{x}_K \mathbf{x}_L|} \text{ for } \sigma = K|L \text{ and } \sigma' = K|M \neq \sigma. \quad (3.8)$$

Upward slopes are then defined as follows :

$$p_{KL}^- = \sum_{\substack{\sigma'=K|M \in \mathcal{E}_K \\ \sigma' \neq \sigma}} \beta_{KLM} p_{KM}^+. \quad (3.9)$$

Finally, the slope used in (3.5) is :

$$p_{KL} = p_{KL}^- \psi_I \left( \frac{p_{KL}^+}{p_{KL}^-} \right) \quad (3.10)$$

where  $\psi_I$  is a limiting operator verifying :

- Extrema degeneracy :  $\psi_I(r) = 0$  if  $r \leq 0$
- Consistency :  $\psi_I(1) = 1$
- Symmetry :  $\psi_I(r) = r\psi_I\left(\frac{1}{r}\right)$

It must also verify the following condition in order to lead to a  $L^\infty$ -stable scheme, [27] :

$$\forall r \in \mathbb{R}, \psi_I(r) \leq \min_{\substack{K \in \mathcal{T} \\ L \in \mathcal{T}: K|L \in \mathcal{E}_K}} \frac{|\mathbf{x}_K \mathbf{x}_L|}{|\mathbf{x}_K \mathbf{y}_\sigma|}. \quad (3.11)$$

Note that the admissibility condition defined in the introduction (Fig. 3.2) means that  $\forall L \in \mathcal{T} : K|L \in \mathcal{E}_K, \lambda_{KL} > 0$ , and implies  $\forall (K, L, M) \in \mathcal{T}^3, \beta_{KLM} < 0$ . The scheme derived from the interface method is  $L^\infty$ -stable, and ensures the verification of a local maximum principle, [27].

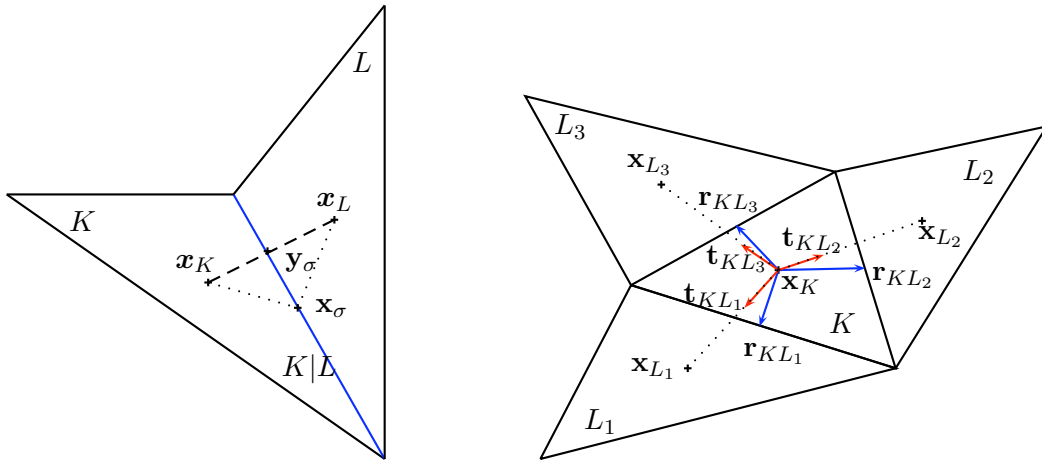


FIG. 3.3 – Geometrical notions for the definition of multislope MUSCL methods. Left : the center method predicts the reconstruction  $\phi_{KL}$  at the face barycenter  $\mathbf{x}_\sigma$ , the intersection method at point  $\mathbf{y}_\sigma$ . Right : geometrical definition of vectors  $\mathbf{r}_{KL}$  and  $\mathbf{t}_{KL}$ .

### 3.3.2 The center multislope MUSCL method

The *center method* uses  $\mathbf{z}_\sigma = \mathbf{x}_\sigma$  in the reconstruction (3.5) of  $\phi_{KL}$ , see Fig. 3.3 (Left), with a slope  $\pi_{KL} = q_{KL}$  determined using a limiter and the downward ( $q_{KL}^+$ ) and upward ( $q_{KL}^-$ ) slopes :

$$q_{KL} = q_{KL}^- \psi_C \left( \frac{q_{KL}^+}{q_{KL}^-} \right). \quad (3.12)$$

The downward slope cannot be defined as the natural slope any more, because  $\mathbf{x}_K, \mathbf{x}_\sigma$  and  $\mathbf{x}_L$  are not aligned, see Fig. 3.3. It is then defined as

$$q_{KL}^+ = a_{KL} p_{KL}^+ + b_{KL} \sum_{\substack{M \in \mathcal{T}: K|M \in \mathcal{E}_K \\ M \neq L}} \gamma_{KLM}^\perp p_{KM}^+ \quad (3.13)$$

where  $a_{KL} = \mathbf{r}_{KL} \cdot \mathbf{t}_{KL}$ ,  $b_{KL} = \mathbf{r}_{KL} \cdot \mathbf{t}_{KL}^\perp$ , and  $\gamma_{KLM}^\perp$  are the coefficients in the decomposition of  $\mathbf{r}_{KL}$  in  $(\mathbf{t}_{KL}, \mathbf{t}_{KL}^\perp)$ :

$$\mathbf{r}_{KL} = a_{KL} \mathbf{t}_{KL} + b_{KL} \mathbf{t}_{KL}^\perp \quad (3.14)$$

$$= a_{KL} \mathbf{t}_{KL} + b_{KL} \sum_{\substack{M \in \mathcal{T}: K|M \in \mathcal{E}_K \\ M \neq L}} \gamma_{KLM}^\perp \mathbf{t}_{KM} \quad (3.15)$$

where  $\mathbf{r}_{KL} = \frac{\mathbf{x}_K \mathbf{x}_\sigma}{|\mathbf{x}_K \mathbf{x}_\sigma|}$  and  $\mathbf{t}_{KL} = \frac{\mathbf{x}_K \mathbf{y}_\sigma}{|\mathbf{x}_K \mathbf{y}_\sigma|}$ ,  $p_{KL}^+$  and  $p_{KM}^+$  are the natural slopes from the intersection method, see Fig. 3.3 (Right).

Upward slopes are defined as in the intersection method :

$$q_{KL}^- = \sum_{\substack{M \in \mathcal{T}: K|M \in \mathcal{E}_K \\ M \neq L}} \beta_{KLM} q_{KM}^+, \quad (3.16)$$

where

$$\beta_{KLM} = -\frac{\lambda_{KM} |\mathbf{x}_K \mathbf{x}_M|}{\lambda_{KL} |\mathbf{x}_K \mathbf{x}_L|} \quad (3.17)$$

and the  $\lambda_{KL}$  are the barycentric coordinates of  $\mathbf{x}_K$  in the system of the adjacent cells barycenters :

$$\mathbf{x}_K = \sum_{L \in \mathcal{T}: K|L \in \mathcal{E}_K} \lambda_{KL} \mathbf{x}_L. \quad (3.18)$$

Without further modifications, the scheme derived from the center method is not LED-stable and does not verify a local maximum principle. A cut-off must be applied to the reconstructed value  $\phi_{KL}$  in order to recover the LED property. The cut-off function is defined as

$$C_{\phi^{min}}^{\phi^{max}}(\phi) = \begin{cases} \phi^{min} & \text{if } \phi^{min} > \phi \\ \phi & \text{if } \phi^{min} < \phi < \phi^{max} \\ \phi^{max} & \text{if } \phi^{max} < \phi \end{cases}. \quad (3.19)$$

Choosing  $\phi^{min} = \min(\phi_K, \phi_K + p_{KL} |\mathbf{x}_K \mathbf{x}_L|)$  and  $\phi^{max} = \max(\phi_K, \phi_K + p_{KL} |\mathbf{x}_K \mathbf{x}_L|)$  leads to a LED scheme, but not to the preservation of a local maximum principle, [27]. Nevertheless, in the numerical experimentations we have conducted, the maximum principle was verified.

### 3.3.3 Proposition for an extension to non-simplexes and/or non-matching grids

In the two above methods, upward slopes are computed as linear combinations of downward slopes (3.9), (3.16), with coefficients deduced from the barycentric coordinates of  $\mathbf{x}_K$  in the system of the adjacent cells barycenters. On a non-matching grid and/or on a grid composed of elements other than simplexes, these barycentric coordinates cannot be uniquely defined any more. The best n-set of coordinates is the one with the smallest variance : the search for this best n-set of barycentric coordinates then expresses itself as a constraint optimization problem. We propose a simpler determination of n-sets of  $\lambda_{KL}$  valid for any 2D/3D meshings. A first simple idea is to use a fixed  $\lambda : \forall L \in \mathcal{T} : K|L \in \mathcal{E}_K, \lambda = \frac{1}{\text{card}(\mathcal{E}_K)}$ . However, this choice does not

discriminate between neighbouring cells. A more efficient set of weights is tested, introducing the distance between cell centers :

$$\lambda_{KL} = \frac{\sum_{\substack{M \in \mathcal{T}: K|M \in \mathcal{E}_K \\ M \neq L}} |\mathbf{x}_K \mathbf{x}_M|}{(\text{card}(\mathcal{E}_K) - 1) \sum_{M \in \mathcal{T}: K|M \in \mathcal{E}_K} |\mathbf{x}_K \mathbf{x}_M|}. \quad (3.20)$$

The implementation of these weights is easy regardless to grid type. We will assert its performance in the numerical examples.

**Remark on the intersection condition for 3D grids.** The condition  $\mathbf{y}_\sigma \in \sigma$  is very difficult to achieve for 3D grids. None of the 3D meshing generators we used (Tetgen, Netgen, embedded in Gmsh [52]) is able to create admissible grids in this respect. We use the intersection method in all cases where it applies, and switch to the center method otherwise.

**Remark on boundary conditions.** For a Dirichlet boundary edge or face, the downward slope is defined as :

$$p_{K\sigma}^+ = \begin{cases} \frac{\phi_K - \phi_D}{|\mathbf{x}_K \mathbf{y}_\sigma|} & \text{for a Dirichlet boundary edge or face,} \\ g(\mathbf{x}_\sigma) & \text{for a Neumann edge or face.} \end{cases}$$

where  $\mathbf{n}_\sigma$  is the boundary edge  $\sigma \in \mathcal{E}_{\text{ext}}$  outward unit normal.

### 3.4 Discretization of the diffusion operator

The diffusion operator is discretized with the a second order cell centred scheme which was first presented in [41] and tested in [1] for oil engineering simulation problems. It is based on the construction of an *ad hoc* discrete gradient.

Its convergence analysis was performed in [38] in the framework of its more general version SUSHI (Scheme Using Stabilization and Hybrid Interfaces) for the discretization of anisotropic and heterogeneous diffusion problems on general non-conforming grids, and more recently in [37] for nonlinear operators of the Leray-Lions type.

The idea, related to that of the Galerkin methods, is to approximate the diffusive term  $\int_\Omega \nabla \phi \cdot \nabla \psi \, d\mathbf{x}$  in (3.2) by  $\int_\Omega \nabla_{\mathcal{T}} \phi \cdot \nabla_{\mathcal{T}} \psi \, d\mathbf{x}$ , where  $\nabla_{\mathcal{T}}$  denotes the discrete gradient operator.

We start by choosing, for any internal  $\sigma \in \mathcal{E}_{\text{int}}$  and any edge of the Neumann boundary  $\sigma \in \mathcal{E}_{\text{N}}$ , some real coefficients  $(\beta_\sigma^L)_{L \in \mathcal{T}}$  such that the barycenter  $\mathbf{x}_\sigma$  of  $\sigma$  is expressed by :

$$\mathbf{x}_\sigma = \sum_{L \in \mathcal{T}} \beta_\sigma^L \mathbf{x}_L, \quad \sum_{L \in \mathcal{T}} \beta_\sigma^L = 1. \quad (3.21)$$

Note that the coefficients  $\beta_\sigma^L$  are not necessarily non-negative. In three space dimensions, it is always possible to restrict the number of nonzero coefficients  $\beta_\sigma^L$  to four. In practice, we try to choose the coefficients  $\beta_\sigma^L$  close to the value  $\frac{1}{2}$  and avoid large variations of their values. In order

to define a discrete gradient, we introduce a second order interpolation operator  $\Pi_\sigma$  at the point  $\mathbf{x}_\sigma$ ; because of the non homogeneous Dirichlet boundary conditions, we need to take into account the discrete values on the Dirichlet edges; more precisely, for a family  $\Phi = ((\phi_K)_{K \in \mathcal{T}}, (\phi_\sigma)_{\sigma \in \mathcal{E}_D})$  of the set of discrete values  $X_{\mathcal{T}}$ , let

$$\Pi_\sigma \Phi = \sum_{L \in \mathcal{T}} \beta_\sigma^L \phi_L, \quad \forall \sigma \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_N, \quad (3.22a)$$

$$\Pi_\sigma \Phi = \phi_\sigma, \quad \forall \sigma \in \mathcal{E}_D, \quad (3.22b)$$

where  $\mathcal{E}_N$  (resp.  $\mathcal{E}_D$ ) denotes the set of faces which are included in the Neuman (resp. Dirichlet) boundary  $\Gamma_N$  (resp.  $\Gamma_D$ ). For  $\Phi \in X_{\mathcal{T}}$ , we introduce a first piecewise constant gradient  $\tilde{\nabla}_{\mathcal{T}} \Phi \in H_{\mathcal{T}}$ , defined by its constant value  $\tilde{\nabla}_K \phi$  on each cell  $K$ :

$$\tilde{\nabla}_K \Phi = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\Pi_\sigma \Phi - \phi_K) \mathbf{n}_\sigma \quad \forall K \in \mathcal{T}. \quad (3.23)$$

The discrete gradient thus defined is consistent, thanks to the following geometrical formula:

$$\sum_{\sigma \in \mathcal{E}_K} |\sigma| \mathbf{n}_\sigma (\mathbf{x}_\sigma - \mathbf{x}_K)^T = |K| \text{Id} \quad \forall K \in \mathcal{T}. \quad (3.24)$$

Indeed, let  $\psi$  be an affine function:  $\mathbf{x} \mapsto \mathbf{w} \cdot \mathbf{x} + z$ , with  $\mathbf{w}, z \in \mathbb{R}^d$  (so that  $\nabla \psi = \mathbf{w}$ ). Let  $\Psi^{\mathcal{T}} = ((\psi_K)_{K \in \mathcal{T}}, (\psi_\sigma)_{\sigma \in \mathcal{E}_D})$  its  $X_{\mathcal{T}}$ -interpolate, defined by  $\psi_K = \psi(\mathbf{x}_K)$  and  $\psi_\sigma = \psi(\mathbf{x}_\sigma)$  for any  $\sigma \in \mathcal{E}_D$ ; let us verify that  $\tilde{\nabla}_K \Psi^{\mathcal{T}} = \mathbf{w}$  for any  $K \in \mathcal{T}$ . First, since  $\Pi_\sigma$  is a second order interpolation operator and  $\psi$  is affine, we have:  $\Pi_\sigma \Psi^{\mathcal{T}} - \psi_K = \psi(\mathbf{x}_\sigma) - \psi(\mathbf{x}_K) = \mathbf{w} \cdot (\mathbf{x}_\sigma - \mathbf{x}_K)$ . Therefore:

$$\tilde{\nabla}_K \Psi^{\mathcal{T}} = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| \mathbf{w} \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) = \mathbf{w},$$

thanks to (3.24). Note that this is not sufficient to ensure the convergence of the scheme. We also need a weak convergence property (see [38]) and a stability property which is not satisfied by the above gradient: indeed, as noted in [38] and illustrated in [21] in particular in the case of Cartesian grids, this discrete gradient may vanish for non zero functions; we thus stabilize it by a consistency error term, which reads:

$$R_{K,\sigma} \Phi = \frac{\sqrt{d}}{|\mathbf{x}_K \mathbf{y}_\sigma|} \left( \Pi_\sigma \Phi - \phi_K - \tilde{\nabla}_K \Phi \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) \right) \quad (3.25)$$

Note that  $R_{K,\sigma} \Psi^{\mathcal{T}}$  vanishes if  $\Psi^{\mathcal{T}}$  is the  $X_{\mathcal{T}}$  interpolate of an affine function  $\psi$ . We then define the discrete gradient of a family  $\Phi$  of  $X_{\mathcal{T}}$  as the piecewise constant functions on the cones  $D_{K,\sigma}$ :

$$\nabla_{\mathcal{T}} \Phi = \tilde{\nabla}_{\mathcal{T}} \Phi + R_{K,\sigma} \Phi \mathbf{n}_{K,\sigma}, \quad \text{on } D_{K,\sigma}. \quad (3.26)$$

The diffusive term  $\int_{\Omega} \nabla \phi \cdot \nabla \psi \, d\mathbf{x}$  in (3.2) is then approximated by  $\int_{\Omega} \nabla_{\mathcal{T}} \Phi \cdot \nabla_{\mathcal{T}} \Psi \, d\mathbf{x}$ .

### 3.5 The complete discrete scheme

We wish to find an approximate value of the solution  $\phi^{n+1}$  of (3.2). The discrete unknowns of the scheme are the values  $\{(\phi_K), K \in \mathcal{T}\}$ , which define the piecewise constant (unknown)

function  $\phi_{\mathcal{T}}^{n+1} \in H_{\mathcal{T}}$ . Assume  $\phi_{\mathcal{T}}^n \in H_{\mathcal{T}}$  to be known. Using the discrete divergence operator  $\text{div}_{\mathcal{T}}$  of a function and discrete gradient operators introduced in the previous sections, we summarize the semi implicit scheme under the weak discrete formulation :

Find  $\phi_{\mathcal{T}}^{n+1} \in H_{\mathcal{T}}$  such that

$$\begin{aligned} & \int_{\Omega} \frac{\phi_{\mathcal{T}}^{n+1} - \phi_{\mathcal{T}}^n}{\delta t} \psi \, d\mathbf{x} + \int_{\Omega} \text{div}(\mathbf{u} \phi_{\mathcal{T}}^n) \psi \, d\mathbf{x} + \int_{\Omega} \kappa \nabla_{\mathcal{T}} \Phi_{\mathcal{T}}^{n+1} \cdot \nabla_{\mathcal{T}} \Psi \, d\mathbf{x} \\ & = \int_{\Omega} f(\mathbf{x}, t^{n+1}) \psi \, d\mathbf{x} + \sum_{\sigma \in \mathcal{E}_N} \Pi_{\sigma}(\psi) \int_{\sigma} \phi_N \, d\gamma(\mathbf{x}), \text{ for any } \psi \in H_{\mathcal{T}}, \end{aligned} \quad (3.27)$$

with

$$\begin{aligned} \Phi_{\mathcal{T}}^{n+1} &= ((\phi_K^{n+1})_{K \in \mathcal{T}}, (\phi_{\sigma}^{n+1})_{\sigma \in \mathcal{E}}) \in X_{\mathcal{T}} \text{ with } \phi_{\sigma}^{n+1} = \phi_D(\mathbf{x}_{\sigma}), \\ \Psi &= ((\psi_K)_{K \in \mathcal{T}}, (\psi_{\sigma})_{\sigma \in \mathcal{E}}) \in X_{\mathcal{T}} \text{ with } \psi_{\sigma} = 0. \end{aligned}$$

Note that in this formulation, the discrete gradient applies once to a set of values  $\Phi_{\mathcal{T}}^{n+1} \in X_{\mathcal{T}}$  which contains the discrete unknowns  $(\phi_K^{n+1})_{K \in \mathcal{T}}$  and the Dirichlet boundary conditions  $(\phi_D(\mathbf{x}_{\sigma}))_{\sigma \in \mathcal{E}}$ , and once to the set of values  $\Psi_{\mathcal{T}} \in X_{\mathcal{T}}$  which contains the test function values  $(\psi_K)_{K \in \mathcal{T}}$  and zero values for  $(\phi_D(\mathbf{x}_{\sigma}))_{\sigma \in \mathcal{E}}$ . This mimicks the continuous weak formulation (3.2).

### 3.6 Numerical examples

In order to illustrate the convergence properties of the above proposed approach, we present in this section a set of examples with analytical or reference solution. First, we focus on separate verifications of both methods : the multislope MUSCL for the approximation of convective fluxes in a pure transient convection problem is verified on the example of the transport of a sinusoidal bump and the solid body rotation problem, the SUSHI method for diffusion problem is tested on two 2D non-homogeneous boundary cases and one 3D homogeneous boundary test with exact solutions. Then, we use the coupled approach to solve the convection-diffusion of a Gaussian pulse problem. Throughout our numerical verification cases, we perform a set of convergence tests according to the following finite volume discrete error norm  $\|\cdot\|_{L_{\mathcal{D}}^2}$  :

$$\|\phi - \phi_{\text{ref}}\|_{L_{\mathcal{D}}^2(\Omega)} = \left[ \sum_{K \in \mathcal{T}} |K| |\phi_K - \phi_{\text{ref}}(\mathbf{x}_K)|^2 \right]^{1/2} \quad (3.28)$$

obtained through a one point integration at the finite volume center. Here,  $\phi_{\text{ref}}$  refers for the given exact solution and  $\phi_K$  its numerical counterpart. The finite volume discrete error norm (3.28) is given as a function of a cell diameter  $h$  which corresponds to the maximum distance between the control volume vertices.

For our verification tests, we generated eight types of successively refined 2D unstructured/non-conforming grids of different topology as well as three types of 3D unstructured grids. For comparison reasons, we completed these by structured rectangular/hexahedral ones. The unstructured grids have been provided by the meshing software Gmsh, [52] (Fig. 3.4(c), Fig. 3.5(a)), EMC2, [67] (Fig. 3.4(d)), or Gambit, [5] (Fig. 3.4(e), Fig. 3.5(b)). The structured and the rest of the unstructured/non-conforming 2D and 3D grids (structured rectangular and hexahedral grid, Fig. 3.4(a)-3.4(b), 3.4(f)-3.4(h), Fig. 3.5(c)) have been generated using PELICANS framework, [92].



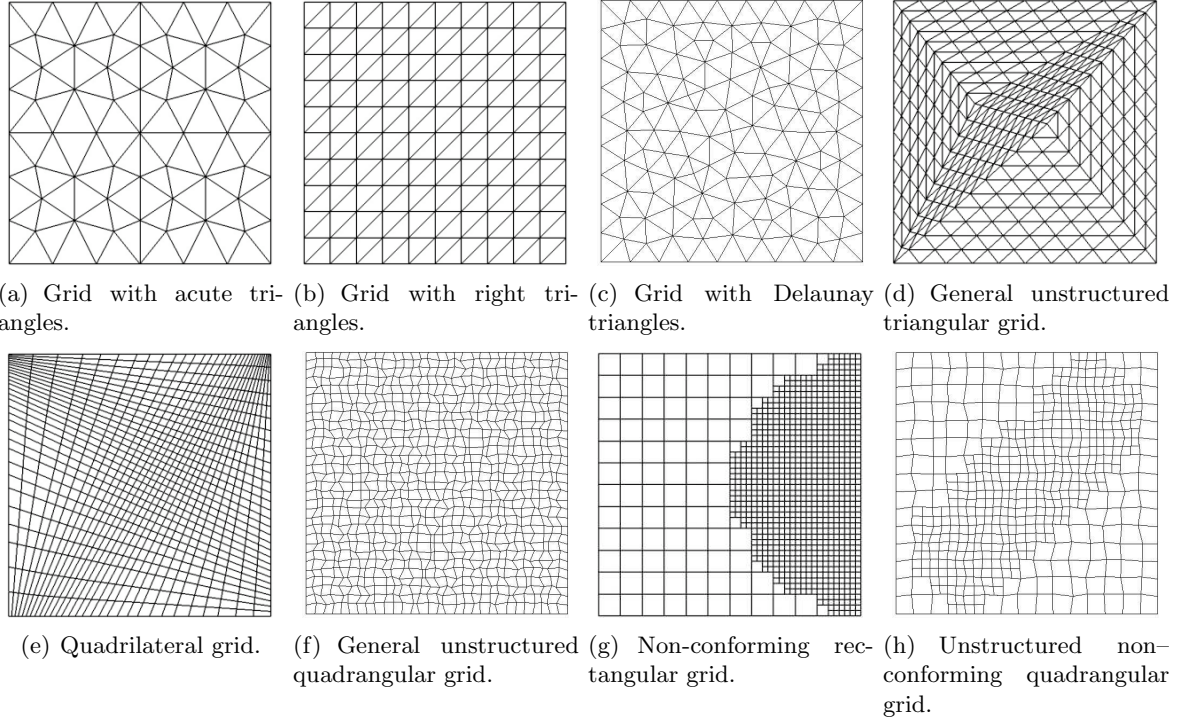


FIG. 3.4 – Eight different 2D grids used for the convergence study of multislope MUSCL and SUSHI methods.

Note that most tests focus on separate convection and diffusion problems, and only one coupled convection–diffusion problem is presented. Actually, the separate problems are harder to solve than the coupled one : considering the approximation of convective fluxes, the presence of a diffusion term tends to stabilize the formulation, and vice versa, considering the approximation of diffusive fluxes, the time-stepping approach used for the approximation of eventual convective terms tends to stabilize the formulation.

### 3.6.1 Multislope MUSCL verification tests

In this section, we use two analytical solutions of the unsteady, pure convection problem with zero source term to verify that the multislope MUSCL method for convective fluxes, and especially its center version, meets the following specifications : (i) non-conforming unstructured grids must be handled, (ii) the discrete maximum principle must be fulfilled and, (iii) even if no theoretical convergence result is yet available, an order higher than 1 must be reached on sufficiently regular solutions. In both tests, we compare the results obtained using the center multislope MUSCL method on unstructured grids with the first order upwind method on the same grids, and also with the classical formulation of the MUSCL method on structured grids (*structured MUSCL*). Note that the multislope MUSCL variants do not degenerate to the classical structured MUSCL method on structured grids.

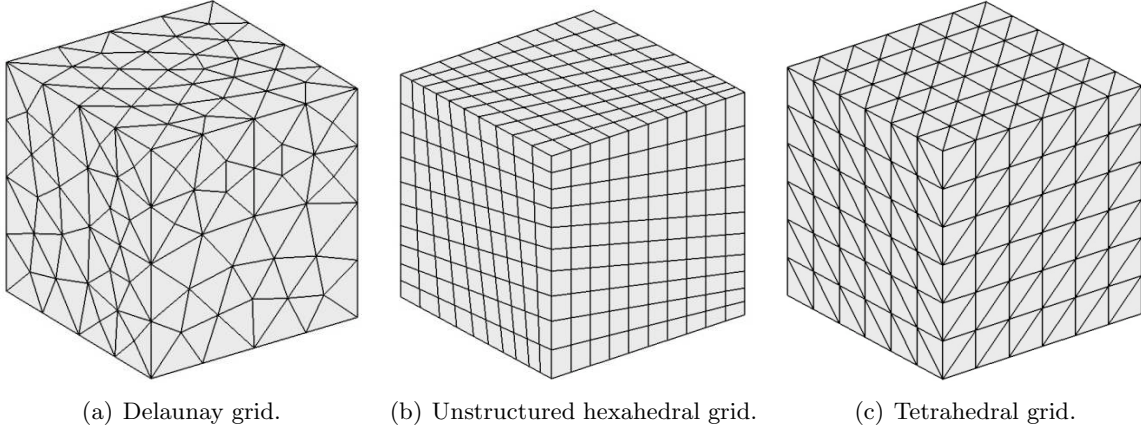


FIG. 3.5 – Three unstructured 3D grids used for the convergence study of multislope MUSCL and SUSHI methods.

### Convection of a sinusoidal bump

This is a 3D version of the 2D test defined by *Buffard & Clain* in [15]; it describes the convection of a sinusoidal bump along the diagonal of a domain  $\Omega = [0, 1]^d$ ,  $d = 2, 3$ , from the point  $\mathbf{x}_0 = (x_0, y_0, z_0) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$  to the point  $\mathbf{x}_1 = (x_1, y_1, z_1) = (\frac{3}{4}, \frac{3}{4}, \frac{3}{4})$ , where the plane  $z = 0$  is supposed for  $d = 2$ . The instantaneous position of the bump is defined as

$$\mathbf{x}_t = (x_t, y_t, z_t) = (x_0 + u_x t, y_0 + u_y t, z_0 + u_z t) \quad (3.29)$$

where  $\mathbf{u}(\mathbf{x}) = (u_x, u_y, u_z) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  is convective velocity and  $t \in [0, 1]$  designates the time. The definition of the bump writes as follows :

$$\phi(\mathbf{x}, t) = \begin{cases} \frac{1}{2}(\cos(5\pi r_t) + 1) & \text{if } r_t < \frac{1}{5} \\ 0 & \text{otherwise} \end{cases} \quad \text{on } \Omega \times t \quad (3.30)$$

with the instantaneous bump center  $r_t = \sqrt{(x - x_t)^2 + (y - y_t)^2 + (z - z_t)^2}$ . The problem is completed with homogeneous Dirichlet conditions on the whole boundary  $\partial\Omega$ .

In order to provide the stability, the time step  $\Delta t$  must be adapted in such a manner to have the CFL parameter

$$CFL = \max_{K \in \mathcal{T}} \frac{\Delta t}{|K|} |F_{K,\sigma}(\mathbf{u}, \phi_{\mathcal{T}})| \quad (3.31)$$

less than 0.6 [15] for which all the computations are experienced to be stable.

We performed a 2D convergence study on six different grids displayed on Fig. 3.4(a), 3.4(b), 3.4(c), 3.4(f), 3.4(h), respectively, and on one structured rectangular grid. The obtained convergence slopes are depicted in Fig. 3.7 and the attained convergence orders are summarized in Table 3.1. As theoretically expected, the solution stays in its prescribed bounds, i.e. the discrete maximum principle is verified. Since no theoretical convergence order is known, we compare our results with the classical upwind method and also with the numerical simulations given by *Buffard & Clain* in [15]. As we can see in Fig. 3.7 and more precisely in the Table 3.1, the convergence order of the center multislope MUSCL method is twice larger than that of the upwind method.

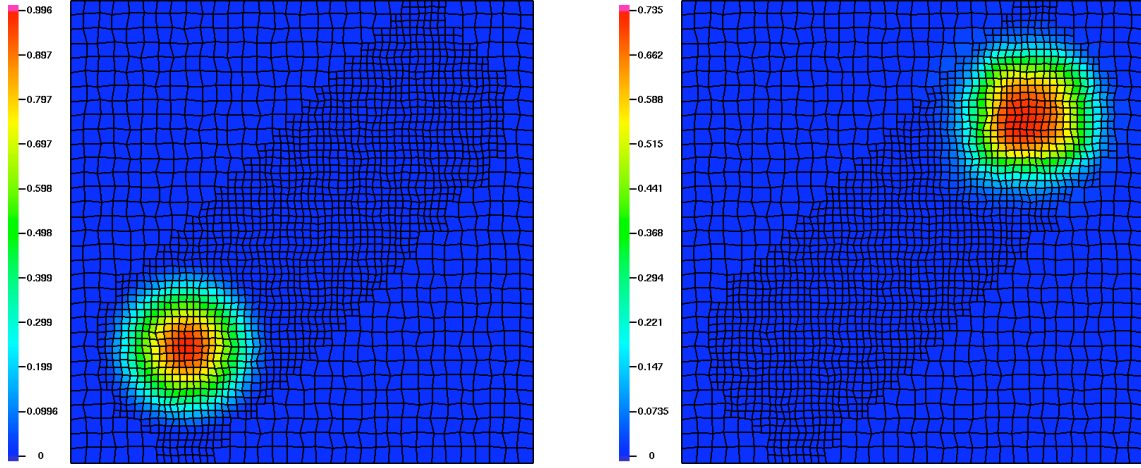


FIG. 3.6 – Initial and final solutions for the 2D convection of a sinusoidal bump on the unstructured non-conforming quadrangular grid, Fig. 3.4(h).

TAB. 3.1 – 2D convection of sinusoidal bump : convergence orders for multislope MUSCL verification test on six types of grids. Comparison with upwind method.

<i>Type of grid</i>	<i>Convergence order</i>	
	<i>Multislope MUSCL or Structured MUSCL (★)</i>	<i>Upwind</i>
Acute triangles, Fig. 3.4(a)	1.64	0.81
Right triangles, Fig. 3.4(b)	1.55	0.65
Delaunay triangles, Fig. 3.4(c)	1.64	0.76
Unstructured quadrangles, Fig. 3.4(f)	1.19	0.74
Unstructured non-conforming quadrangles, Fig. 3.4(h)	1.41	0.74
Structured rectangular (★)	1.49	0.70

Applying a center multislope MUSCL method on a grid composed of Delaunay triangles (see Fig. 3.4(c)), we find a convergence order of 1.64. *Buffard & Clain* obtained a convergence order of 1.17 using the intersection multislope MUSCL method and 1.21 using the center multislope MUSCL method, [15]. This improvement of the order convergence may result from mesh effects : numerical experiments at IRSN have shown that mesh effects can indeed influence the quality of the results.

In order to demonstrate the multislope MUSCL method extension to three dimensions, we performed the test on three unstructured grids (Delaunay, unstructured hexahedral and tetrahedral grids, Fig. 3.5(a)-3.5(c)) and on one structured hexahedral grid. The resulting convergence slopes are shown in Fig. 3.8 and obtained convergence orders are resumed in Table 3.2. Here again, we compared the multislope MUSCL and structured MUSCL methods with the first order upwind method and our results are analogous to these in 2D : the convergence order obtained using the multislope MUSCL method double the upwind one.

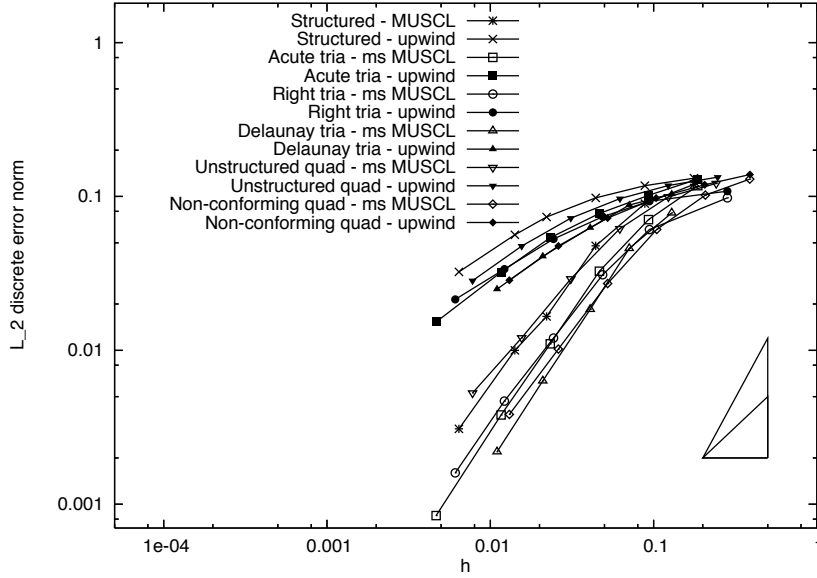


FIG. 3.7 – Logarithmic scale plot of the accuracy of the solution for the 2D homogeneous Dirichlet convection problem (3.41) on six types of grids (Fig. 3.4(a)–3.4(c), 3.4(f), 3.4(h) and regular square grid). Comparison between multislope MUSCL/structured MUSCL method and upwind method. The  $L^2_{\mathcal{D}}$  is traced as a function of the cell diameter  $h$ .

**Solid body rotation**

Solid body rotation is a 2D test describing the rotation of a pattern composed of a square and a cone, see Fig. 3.9(a), around the origin in the square domain  $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$ . The convective velocity is set to

$$\mathbf{u}(\mathbf{x} = (x, y), t) = (2y, -2x) \tag{3.32}$$

and the initial solution is

$$\phi(\mathbf{x} = (x, y), t = 0) = \begin{cases} 1 & \text{if } \frac{1}{10} < x < \frac{3}{5} \wedge -\frac{1}{4} < y < \frac{1}{4} \\ 1 - \frac{\sqrt{(x+0.45)^2 + y^2}}{0.35} & \text{if } \sqrt{(x + 0.45)^2 + y^2} < 0.35 \\ 0 & \text{otherwise} \end{cases} . \tag{3.33}$$

TAB. 3.2 – 3D convection of sinusoidal bump : convergence orders for multislope MUSCL verification test on four types of grids. Comparison with upwind method.

Type of grid	Convergence order	
	Multislope MUSCL	Upwind
	or Structured MUSCL (★)	
Delaunay tetrahedra, Fig. 3.5(a)	0.46	0.23
Unstructured hexahedra, Fig. 3.5(b)	1.30	0.39
Unstructured tetrahedra, Fig. 3.5(c)	0.69	0.37
Structured hexahedra (★)	1.34	0.32

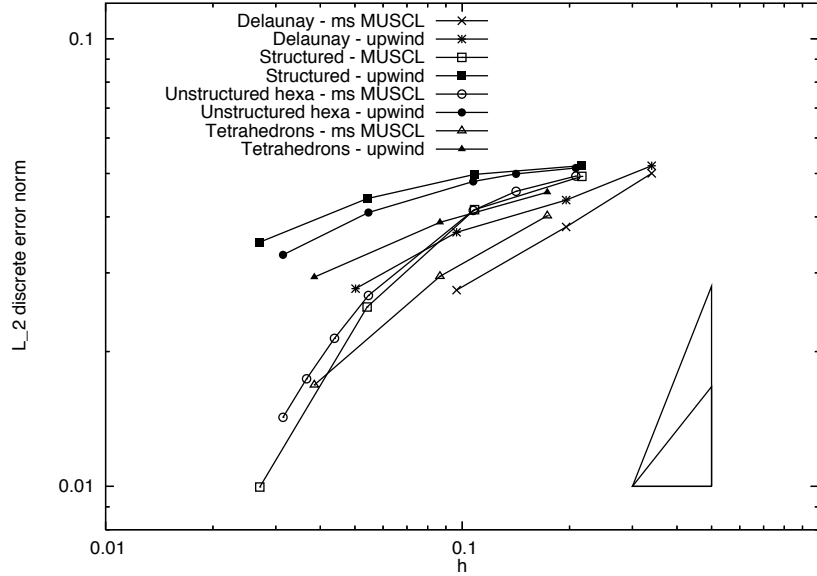


FIG. 3.8 – Logarithmic scale plot of the accuracy of the solution for the 3D homogeneous Dirichlet convection problem (3.41) on three types of grids (regular hexahedral grid, Delaunay tetrahedra, cubes with tetrahedra corresponding to the Fig. 3.5(c)).

At the final time  $t = \pi$  s, the exact solution is equal to the initial solution (3.33).

In the previous case of the convection of a sinusoidal bump, the solution was smooth. In the present case, the solution is discontinuous for the square, and non-differentiable for the cone. Approximating this solution on a grid composed of Delaunay triangles (see Fig. 3.4(c)) with the cell diameter  $h = 0.0353$ , we do not observe any oscillations. The precision is much better than with the upwind scheme (see Fig. 3.9(b) and 3.9(d) for comparison), the discrete maximum principle being verified.

The solutions obtained using the multislope center scheme, the classical MUSCL scheme on a structured mesh, and the classical upwind scheme are shown in Fig. 3.9(b), 3.9(c) and 3.9(d). The only method which has lead to an exact conservation of the initial bounds is the multislope MUSCL method, see Fig. 3.10 for a cutline through the two patterns of the domain at the final time  $t = \pi$  s.

### 3.6.2 SUSHI verification tests

This section is devoted to the numerical results of the SUSHI method for a steady-state diffusion problem in two and three dimensions.

#### 2D tests

Let  $\kappa = 1$  and  $f(\mathbf{x}) = 0$ . We consider a square domain  $\Omega = [0, 1]^2$  and we choose boundary conditions such that the following function  $\xi$  of  $\mathbf{x} = (x, y)$

$$\xi(x, y) = \frac{\sinh(\pi x)}{\sinh(\pi)} \sin(\pi y) \quad \text{on } \Omega.$$

TAB. 3.3 – Ordres of convergence for 2D SUSHI verification tests on five types of grids (Fig. 3.4(a), 3.4(b), 3.4(d), 3.4(e) and 3.4(g)).

<i>Grid type</i>	<i>2D non-homogeneous Dirichlet problem</i>	<i>2D mixed Dirichlet–Neumann problem</i>
Acute triangles, Fig. 3.4(a)	1.99	2.01
Right triangles, Fig. 3.4(b)	1.98	2.12
Unstructured triangles, Fig. 3.4(d)	2.01	1.99
Quadrilaterals, Fig. 3.4(e)	1.99	2.05
Non-conforming grid, Fig. 3.4(g)	1.99	1.99

is solution to the system (3.1a)–(3.1c), which, in the present setting, reads :

$$-\Delta\phi = 0, \quad (3.34)$$

$$\phi = \xi \text{ on } \partial\Omega_D \quad (3.35)$$

$$\nabla\phi \cdot \mathbf{n} = \nabla\xi \cdot \mathbf{n} \text{ on } \partial\Omega_N = \partial\Omega \setminus \partial\Omega_D. \quad (3.36)$$

The set  $\partial\Omega_D$  is defined for each test ; in each case, a convergence study is performed on five different grids that are depicted in Fig. 3.4(a), 3.4(b), 3.4(d), 3.4(e) and 3.4(g).

**Non-homogeneous Dirichlet problem.** Here we take :

$$\partial\Omega_D = \partial\Omega \text{ and } \partial\Omega_N = \emptyset. \quad (3.37)$$

The convergence slopes for this test case are shown in Fig. 3.12 ; they are obtained with respect to the discrete finite volume error norm defined by relation (3.28).

**Mixed Dirichlet–Neumann problem.** Here we take

$$D\text{dom}c_N = \{1\} \times [0, 1] \text{ and } \partial\Omega_D = \partial\Omega \setminus \partial\Omega_N. \quad (3.38)$$

The convergence slopes for this test case, obtained on five different grids (Fig. 3.4(a), 3.4(b), 3.4(d), 3.4(e) and 3.4(g)) are shown in Fig. 3.13.

The non-homogeneous Dirichlet problem as well as the mixed Dirichlet–Neumann problem show the convergence slopes of order 2 in the  $L^2_D$ -norm on general unstructured or non-conforming quadrangular and triangular grids (see Table 3.3 for more accurate results of convergence orders). This is in agreement with results obtained by *Eymard et al.* in [35, 38] and *Herbin & Hubert* in [68]. Although the discrete maximum principle may be violated by the SUSHI method, in our experiments, the solution remains in its prescribed bounds.

### 3D test

Let us consider a pure diffusion, steady-state problem defined on  $\Omega = [0, 1] \times [-\frac{1}{2}, \frac{1}{2}] \times [0, 1]$  with unitary diffusion coefficient  $\kappa = 1$ . T Homogeneous Dirichlet boundary conditions are

applied on the whole domain boundary  $\partial\Omega$ . We consider the source term  $f$  such that the function  $(x, y, z) \mapsto \sin(\pi x) \cos(\pi y) \sin(\pi z)$  is the exact solution to :

$$-\Delta\phi = f \text{ on } \Omega, \quad (3.39)$$

$$\phi = 0 \text{ on } \partial\Omega \quad (3.40)$$

The convergence study is performed on three types of 3D grids, one regular hexahedral and two unstructured that are depicted in Fig.3.5(b)–3.5(c), respectively.

The resulting convergence slopes, depicted in Fig. 3.14, show that for the regular hexahedral and the unstructured hexahedral grids, the convergence order in the  $L^2_{\mathcal{D}}$ -norm is 2.00 and 1.87, respectively. The measured order of the convergence on tetrahedral grid is 1.64 (Fig. 3.5(c)).

### 3.6.3 Multislope MUSCL - SUSHI coupling verification

In this section, we present the coupling of the multislope MUSCL method together with the SUSHI scheme.

Let us consider the coupled, unsteady convection–diffusion problem without source term for which the 2D analytical solution is given by *Noye & Tan* in [88], that we extend to 3D here. The computational domain is  $\Omega = [0, 2]^d \subset \mathbb{R}^d$ ,  $d = 2, 3$ . The test describes the diffusion of an initial Gaussian pulse as it is advected along the main diagonal line of the domain. The exact solution of (3.1a) reads as follows :

$$\phi(x, y, z, t) = \frac{1}{(4t+1)^{d/2}} \exp \left[ -\frac{(x - u_x t - x_0)^2}{\kappa_x(4t+1)} - \frac{(y - u_y t - y_0)^2}{\kappa_y(4t+1)} - \frac{(z - u_z t - z_0)^2}{\kappa_z(4t+1)} \right] \quad (3.41)$$

where the diffusion coefficients are set to  $\kappa_x = \kappa_y = \kappa_z = 0.01$  and the convective velocities  $u_x = u_y = u_z = 0.8$ . The points  $x_0 = y_0 = z_0 = 0.5$  are the coordinates of the initial position of the center of the Gaussian pulse with an unitary pulse height. The Dirichlet boundary conditions are imposed on  $\partial\Omega$  and the final time of calculations is set to  $t = 1.25$  s when the pulse will take a position (1.5, 1.5, 1.5) and its height will be 1/6.

We performed the convergence study for 3D solution (3.41), as well as for its 2D counterpart given by [88] supposing the plane  $z = 0$  in (3.41). The initial 2D configuration and the computed solution on the Delaunay type grid (Fig. 3.4(c)) are depicted in Fig. 3.15. The results of 2D and 3D convergence tests are displayed in Fig. 3.16 and Fig. 3.17, respectively, and the convergence rates are summarized in Table 3.4, where we compared the convergence rates obtained using multislope or structured MUSCL method and the classical upwind method. We observe that the convergence rate of the multislope MUSCL-SUSHI coupling approach approximately double the order achieved with the upwind-SUSHI coupling. In Fig. 3.18, we show the cut curves at  $t = 1.25$  s through the mesh diagonal of the computed solution on four different 2D grids. The solution computed with the multislope MUSCL-SUSHI approach matches well with the exact solution. The solution obtained using the upwind-SUSHI strategy is increased by a factor 10 – 20% depending on the grid type.

TAB. 3.4 – 2D and 3D convection of a Gaussian pulse : convergence orders for multislope MUSCL-SUSHI coupling verification test on six types of 2D grids (Fig. 3.4(a), 3.4(b), 3.4(c), 3.4(f), 3.4(h) and regular square grid) and two types of 3D grids (Delaunay type, Fig. 3.5(a), and regular hexahedral). Comparison with upwind-SUSHI coupling.

<i>Type of grid</i>	<i>Convergence order</i>	
	<i>Multislope MUSCL or Structured MUSCL (★)</i>	<i>Upwind</i>
Acute triangles, Fig. 3.4(a)	1.16	0.71
Right triangles, Fig. 3.4(b)	2.98	0.80
Delaunay triangles, Fig. 3.4(c)	1.37	0.71
Unstructured quadrangles, Fig. 3.4(f)	1.21	0.86
Unstructured non-conforming quadrangles, Fig. 3.4(h)	1.48	0.86
Structured rectangular (★)	2.37	0.86
Delaunay tetrahedra, Fig. 3.5(a)	1.27	0.50
Structurde hexahedra (★)	0.97	0.27

### 3.7 Discussion, conclusions

In this paper, a cell centred scheme based on the multislope MUSCL and SUSHI methods to solve a convection–diffusion equation on general unstructured and non-conforming grids was presented. The control volumes taken into the consideration were triangles and quadrangles in 2D and tetrahedra and hexahedra in 3D. In this sense, we generalized the original work introducing the multislope MUSCL method, [25, 27].

The adopted approach was verified on a set of examples with known analytical solutions. The obtained results shown a very good convergence properties of multislope MUSCL-SUSHI coupled approach that was verified against the exact solution and compared with the upwind-SUSHI strategy. We found that the convergence rate is approximately twice the one obtained using the upwind method in both 2D and 3D computations.

Such an approach is therefore interesting to apply in the numerous engineering applications where computational grids are highly unstructured or where local refinement of the grid is necessary in order to capture more accurately a physical phenomenon, e.g. in the atmospheric dispersion or combustion modelling.

### Acknowledgements

The multislope MUSCL method and the SUSHI implementations as well as all numerical tests presented in this paper were realized using the computational fluid dynamics software ISIS, [70], and the PELICANS application framework, [92].



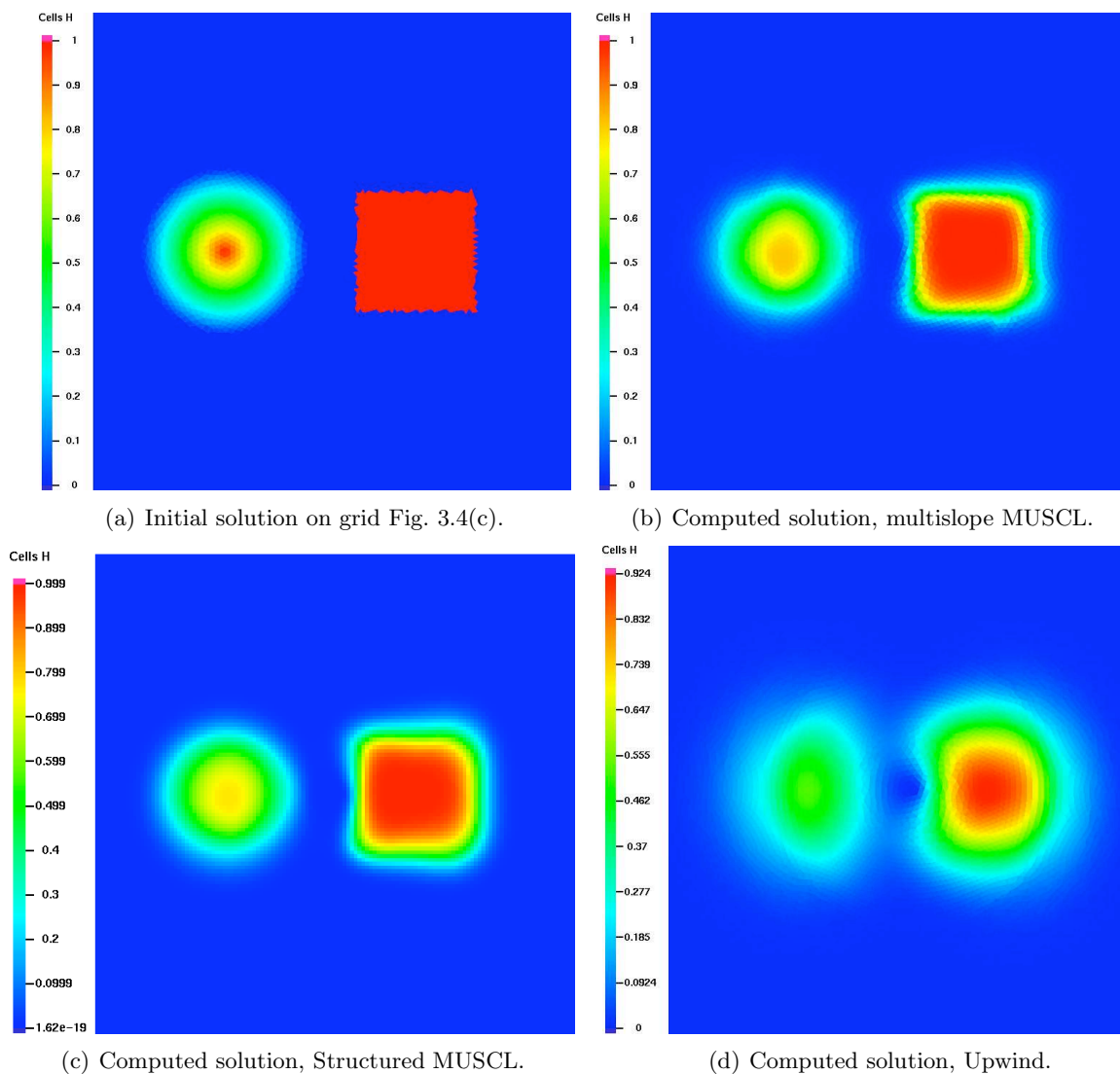


FIG. 3.9 – 2D solid body rotation : 3.9(a) the initial configuration of square and cone in the domain  $\Omega$ , 3.9(b)-3.9(d) the computed solution at  $t = \pi$  on two grid types : the unstructured Delaunay triangles (multislope MUSCL method, 3.9(b), and upwind method, 3.9(d)) and the regular square domain (structured MUSCL method, 3.9(c)).

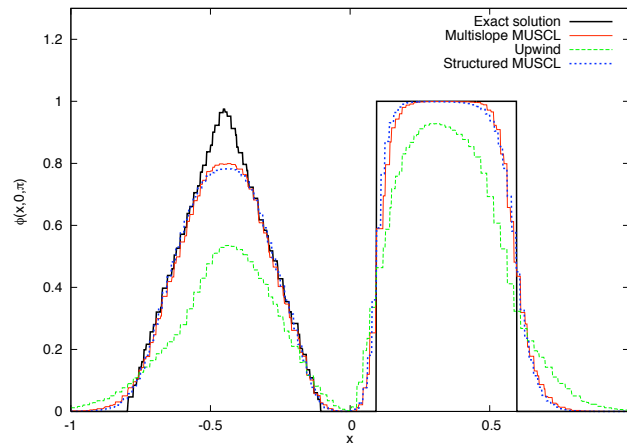


FIG. 3.10 – 2D solid body rotation : cut through the  $x$ -axis at  $y = 0$  of the solution computed via the multislope MUSCL, structured MUSCL and upwind schemes.

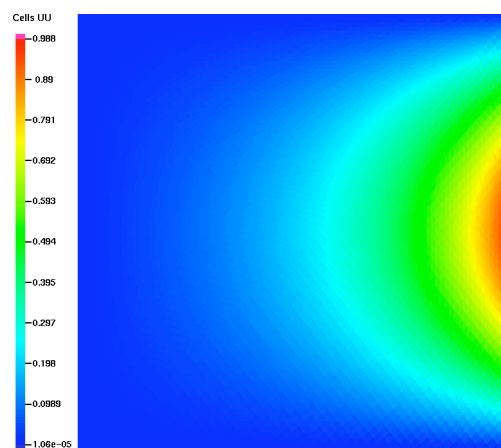


FIG. 3.11 – Solution for the SUSHI verification test with non-homogeneous Dirichlet boundary conditions.

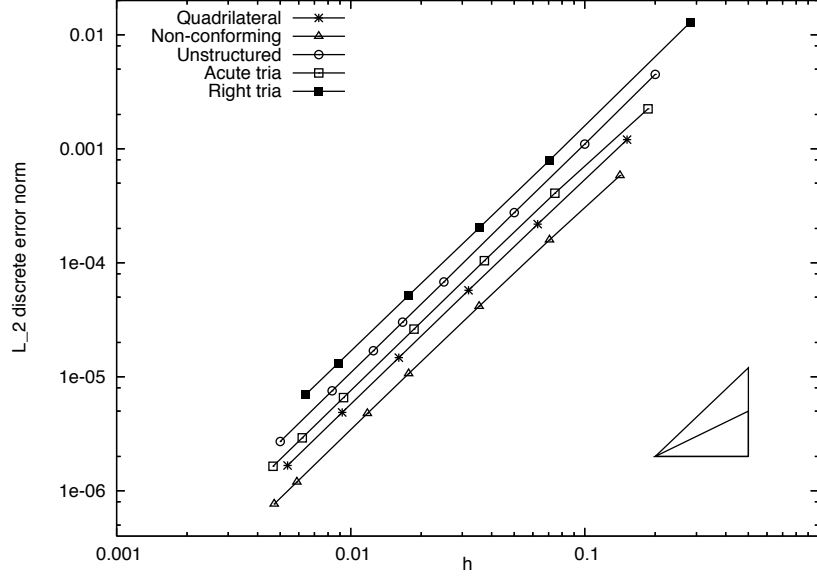


FIG. 3.12 – Logarithmic scale plot of the accuracy of the solution for the non-homogeneous Dirichlet diffusion problem (3.34) with non homogeneous boundary conditions defined by (3.37) on five types of grids depicted in Fig. 3.4(a), 3.4(b), 3.4(d), 3.4(e) and 3.4(g). The  $L^2_{\mathcal{D}}$ -norm is traced as a function of the cell diameter  $h$ .

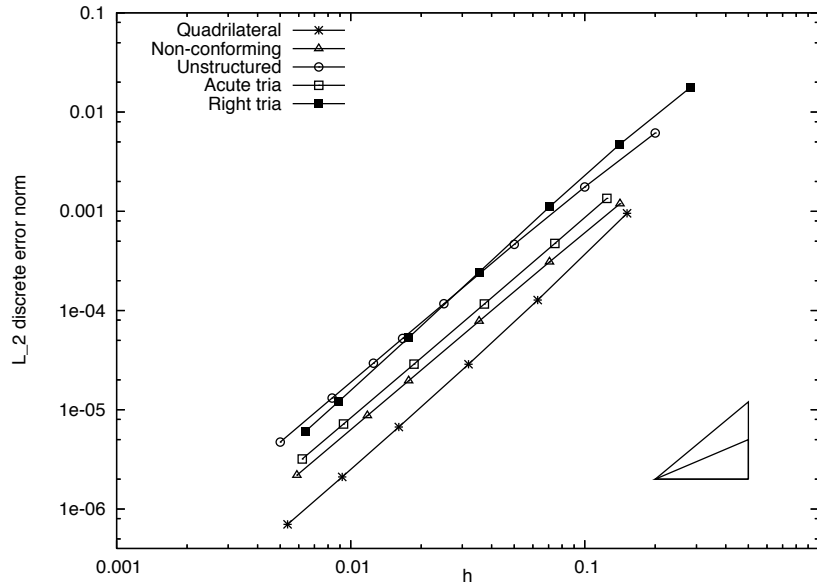


FIG. 3.13 – The accuracy of the solution for the diffusion problem (3.34), with mixed Dirichlet-Neumann boundary conditions defined by (3.38) on five types of grids depicted in Fig. 3.4(a), 3.4(b), 3.4(d), 3.4(e) and 3.4(g). The  $L^2_{\mathcal{D}}$ -norm is traced as a function of the cell diameter  $h$  in logarithmic scale.

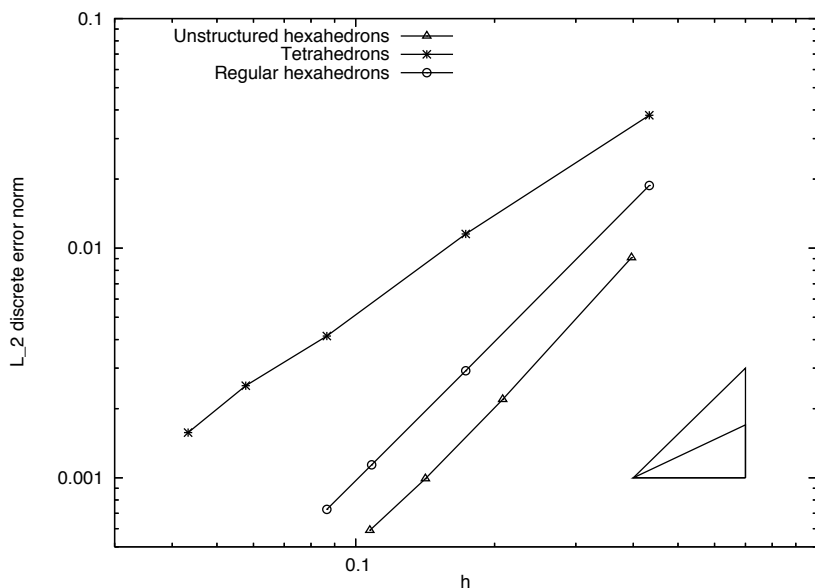


FIG. 3.14 – The  $L^2_{\mathcal{D}}$ -norm as a function of the cell diameter  $h$  in logarithmic scale for three-dimensional verification test of SUSHI method.

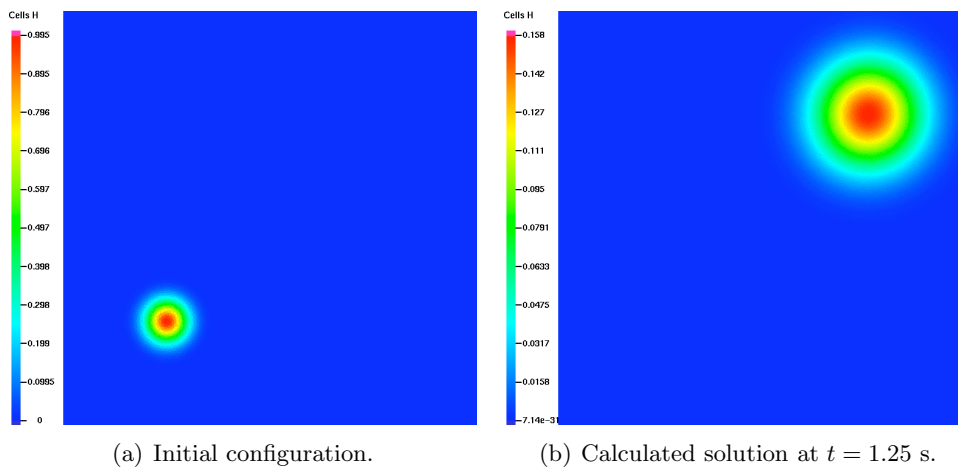


FIG. 3.15 – Initial and final solutions for the 2D convection of a Gaussian pulse on the Delaunay type grid, Fig. 3.4(c), with  $h = 0.0219932$ .

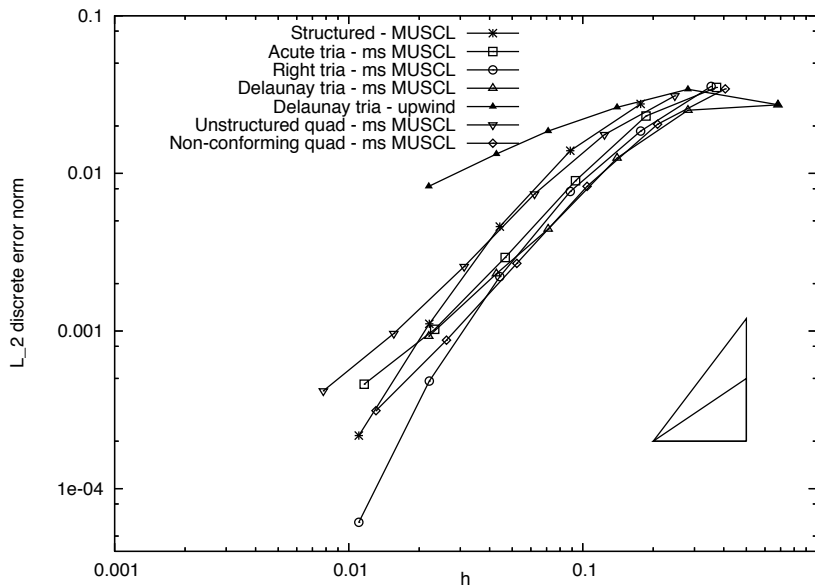


FIG. 3.16 – Multislope MUSCL - SUSHI coupling results : logarithmic scale plot of the accuracy of the solution for the 2D convection of a Gaussian pulse problem on six different grids (Fig. 3.4(a), 3.4(b), 3.4(c), 3.4(f), 3.4(h) and regular square grid).

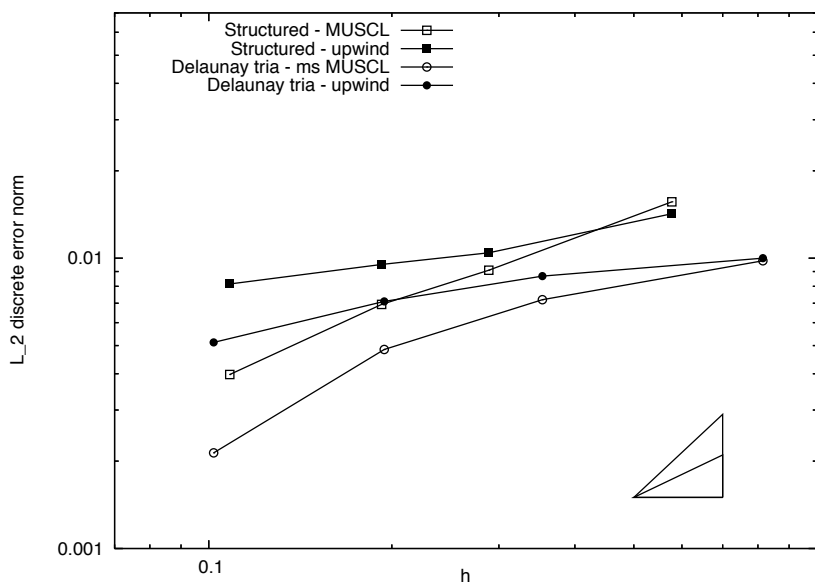
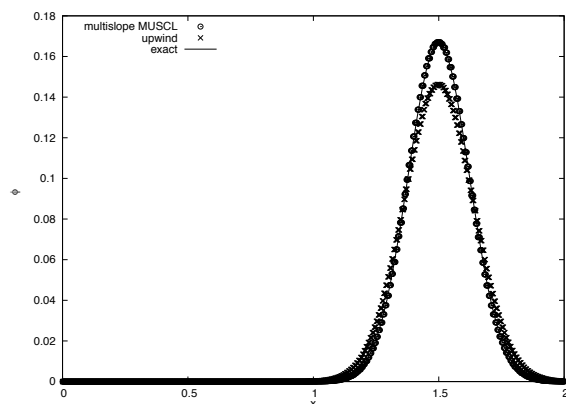
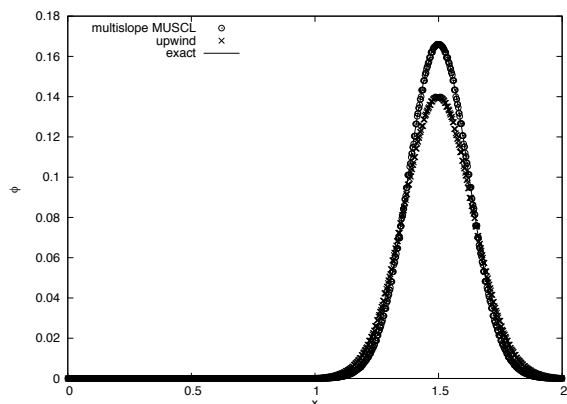
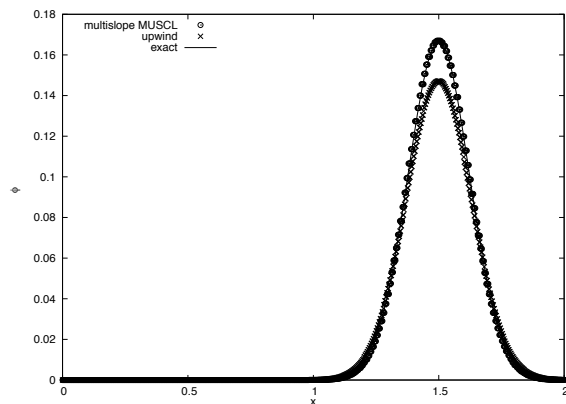
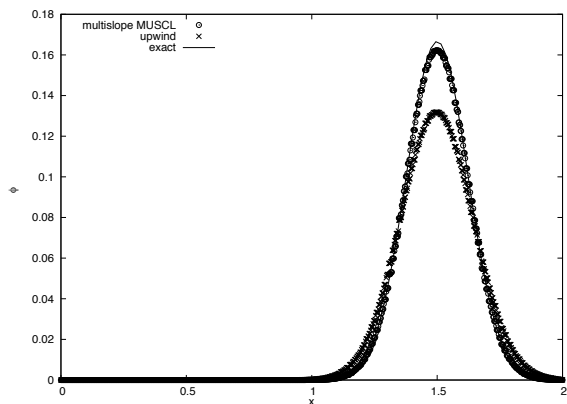


FIG. 3.17 – 3D convection of a Gaussian pulse : logarithmic scale plot of the accuracy of the solution resulting from the multislope MUSCL-SUSHI methods coupling.



(a) Grid with acute triangles ( $h = 0.0116462$ ), Fig. 3.4(a). (b) Grid with right triangles ( $h = 0.0110485$ ), Fig. 3.4(b).



(c) Delaunay grid ( $h = 0.0219932$ ), Fig. 3.4(c). (d) Structured rectangular grid ( $h = 0.00552427$ ).

FIG. 3.18 – 2D convection of a Gaussian pulse : comparison between computed solution with the multislope/structured MUSCL method and the upwind method, respectively, and the exact solution at time  $t = 1.25$  s along the diagonal  $x = y$  of the computational domain for four different grids.



## Chapitre 4

# Simulations numériques

Dans ce dernier chapitre, nous présentons l'étude de trois problèmes concrets d'application des méthodes numériques développées dans ce mémoire. La première application est la simulation d'un écoulement hémodynamique dans un tube coudé représentant un modèle de crosse aortique à paroi fixe. La deuxième application correspond à la simulation d'une flamme laminaire tridimensionnelle dans un cylindre entouré par une paroi. Enfin, la troisième application associe à une dynamique aéraulique un transport de chaleur pour décrire la propagation de la température dans un local comportant une porte ouverte : cette situation est dérivée d'une application concrète du code ISIS dans le cadre de campagnes d'étude des incendies. Dans chacun de ces cas, on construit initialement un maillage adapté, raffiné dans la zone d'intérêt de l'écoulement, et qui est utilisé tout au long de la simulation.

### 4.1 Écoulement dans un tube coudé – Hexaèdres

Cet écoulement est une modélisation simplifiée d'un écoulement sanguin dans une crosse aortique modèle (figure 4.2(a)) lors d'une impulsion de pression : le débit d'entrée, imposé à l'entrée du côté court du tube, provient du modèle d'écoulement sanguin de Womersley [111], voir figure 4.1.

L'écoulement est modélisé par les équations de Navier-Stokes classiques pour un fluide incompressible (bilan de masse et de quantité de mouvement) ; les conditions aux limites sont de type "inlet-outlet", elles sont stables en cas de flot réentrant, voir annexe B section B.4 110. L'écoulement est fortement tri-dimensionnel et se caractérise par des couches limites très fines et des phénomènes de recirculation. Pour en reproduire les caractéristiques principales (couches limites à la paroi), le maillage utilisé compte 672 000 mailles. Le transitoire se compose de 600 pas de temps de  $2,5 \cdot 10^{-3}$  seconde, la simulation étant conduite jusqu'au temps  $t = 1.5$  s. Le maillage (figure 4.2(b)), composé d'hexaèdres, est généré par extrusion d'un disque générateur (figure 4.2(c)) à l'aide du mailleur Mefisto [94].

Sur la figure 4.3, on représente la composante en  $z$  de la vitesse. On voit apparaître à la sortie des vitesses dont la composante en  $z$  est négative, et donc la présence d'écoulements réentrants. Sur la figure 4.4, on représente la norme  $L^2$  de la vitesse dans le plan  $(x, z)$ . On s'intéresse notamment au comportement de l'écoulement au sommet du tube, où apparaissent le plus clairement les phénomènes de recirculation. Le plan de coupe  $IO$  représenté sur la figure 4.2(a) indique son emplacement. Les vecteurs vitesse projetés sur ce plan de coupe à différents



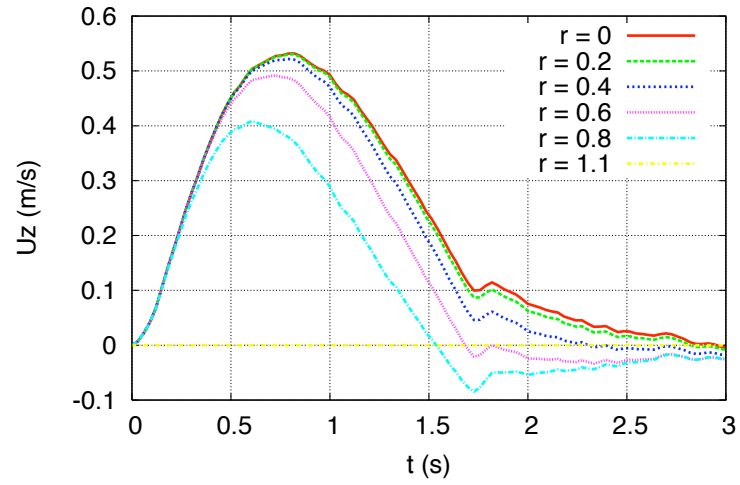


FIG. 4.1 – Profil d'entrée de Womersley

pas de temps, colorés par l'amplitude de la vitesse selon l'axe  $x$ , sont rassemblés sur la figure 4.5. On y voit la formation de deux zones de recirculation symétriques, entre l'écoulement dans le sens du courant (vitesses dont la composante  $x$  est négative), et l'écoulement revenant en sens contraire (vitesses dont la composante  $x$  est positive).

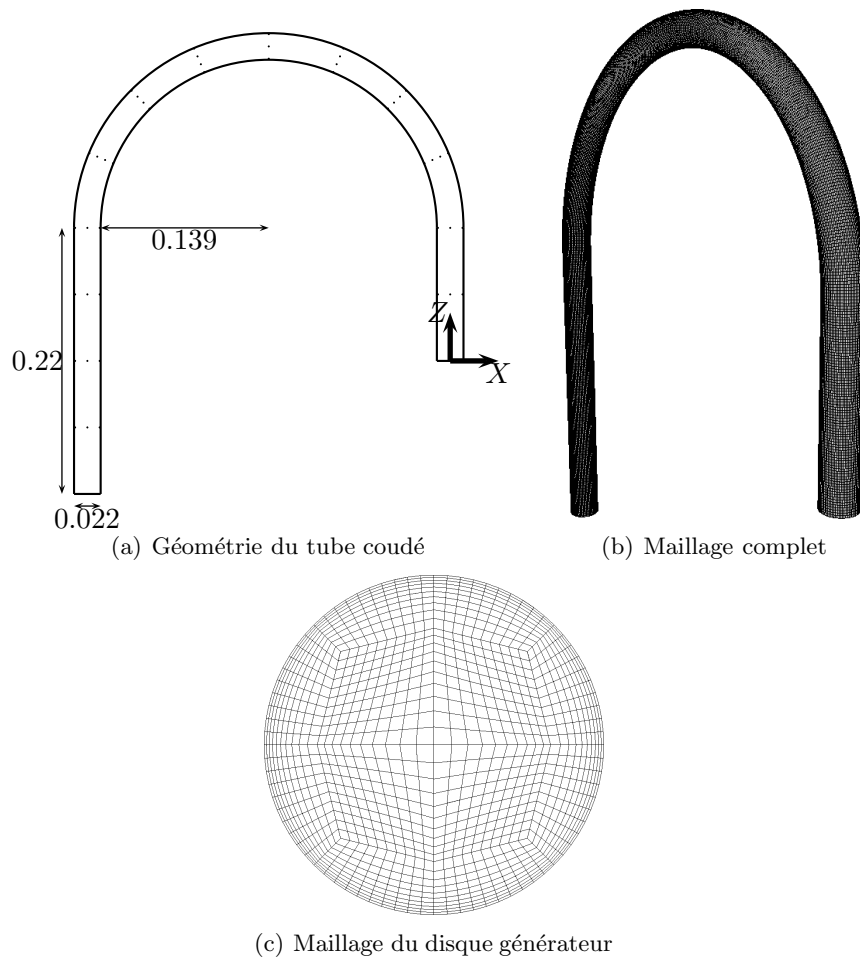
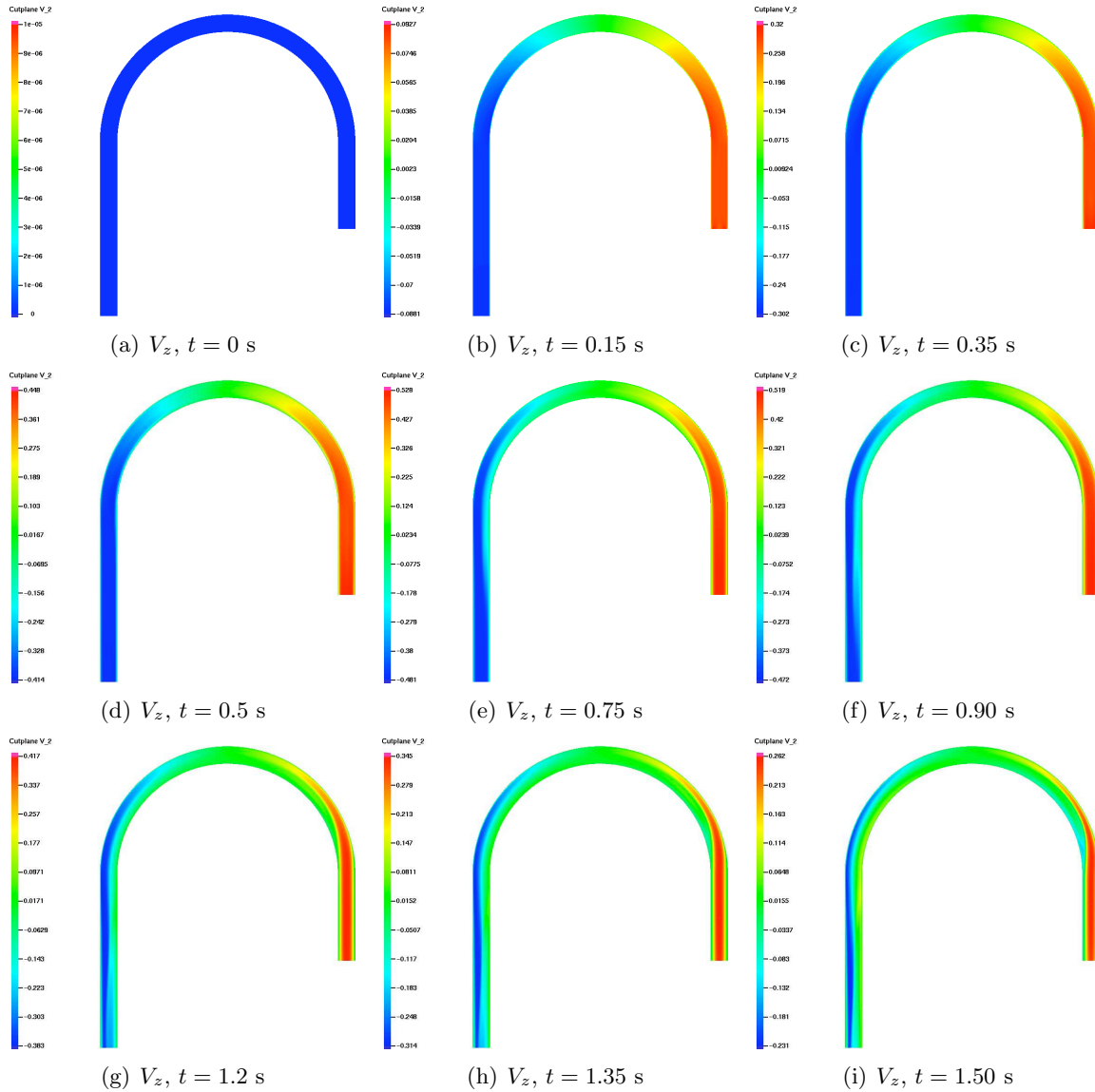


FIG. 4.2 – Géométrie et profil d'entrée pour le modèle de crosse aortique

FIG. 4.3 – Vitesse  $V_z$

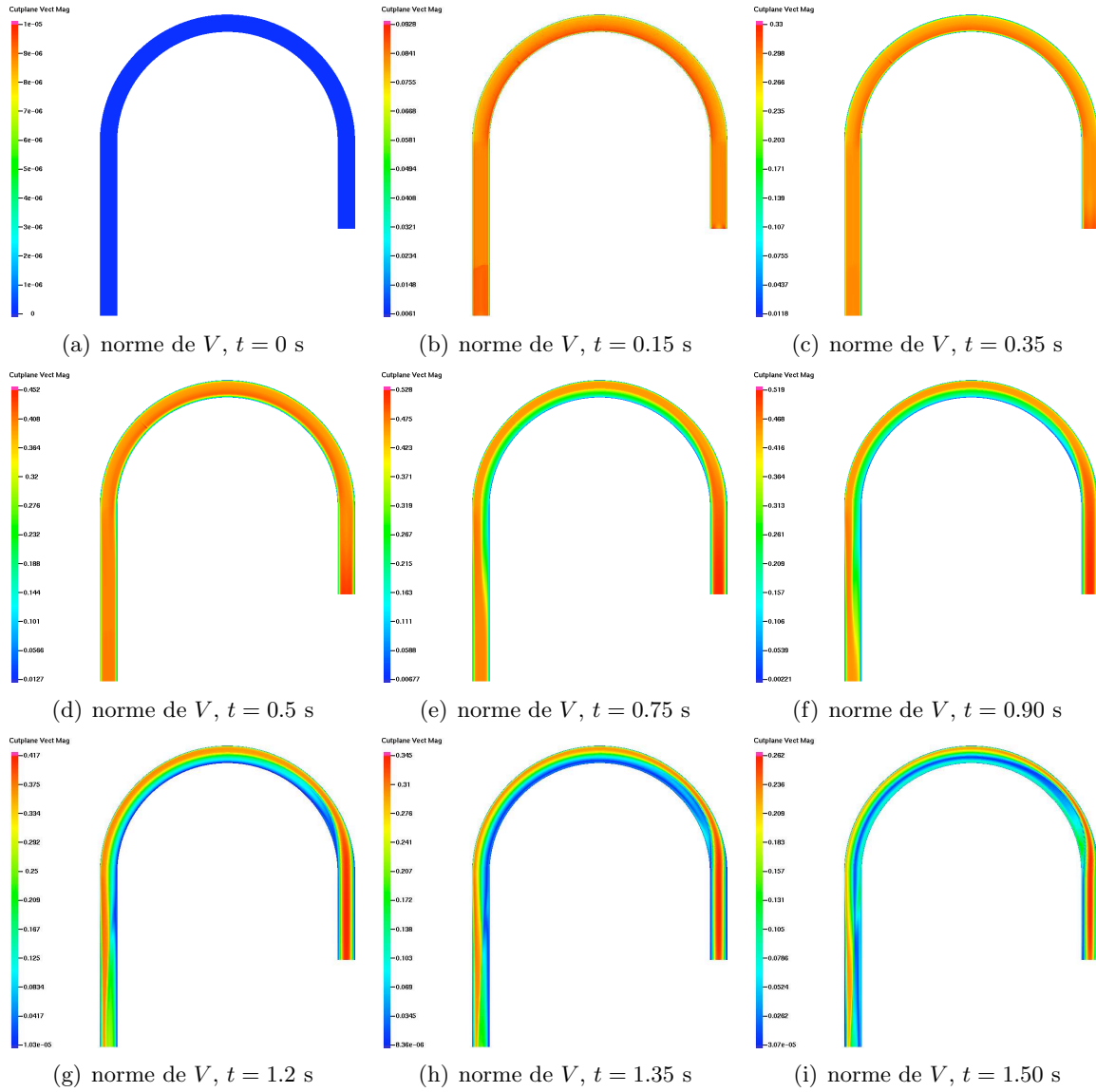


FIG. 4.4 – norme de  $V$

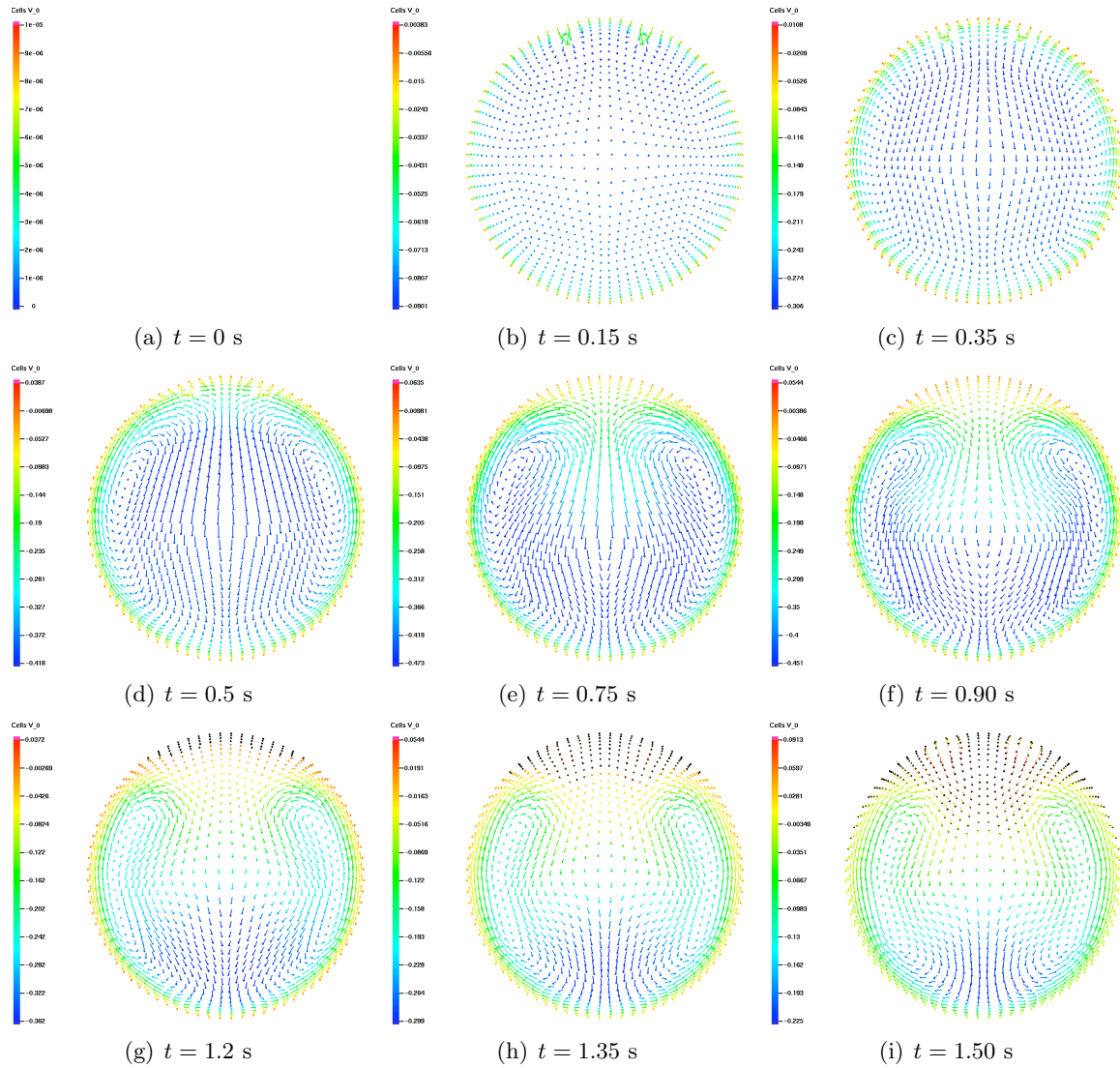


FIG. 4.5 – Vecteurs vitesse sur le plan de coupe IO, colorés par l'amplitude de la vitesse selon l'axe  $x$

## 4.2 Flamme laminaire 3D

On s'intéresse ici à la simulation d'une flamme de diffusion laminaire dans un cylindre. Le concept de flamme laminaire est celui proposé par Burke et Schumann : la réaction chimique est supposée se produire en une seule étape infiniment rapide et irréversible. La flamme est produite par un écoulement en jet provenant de la base d'un cylindre. Au centre du cylindre se trouve une injection de combustible entourée d'une injection d'air. L'écoulement est isolé de l'extérieur par une paroi cylindrique. Le dispositif correspondant à l'expérience simulée est représenté sur la figure 4.6. On se réfère à [85] pour la géométrie du cylindre et des rayons d'injection : le rayon de l'injection de combustible est  $r_I = 0.635$  cm et le rayon de l'injection d'air est celui du cylindre réacteur  $r_O = 2.54$  cm. La longueur de la paroi est  $L = 30$  cm.

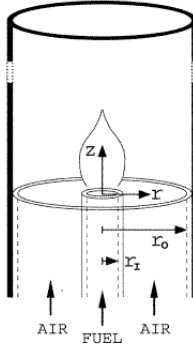


FIG. 4.6 – Schema du dispositif expérimental pour la flamme laminaire 3D

Les équations de bilan de masse, de quantité de mouvement et de fraction de mélange régissant le phénomène simulé sont :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \quad (4.1)$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) = -\nabla p + \nabla \cdot \tau + (\rho - \rho_0) \nabla g \quad (4.2)$$

$$\frac{\partial \rho z}{\partial t} + \nabla \cdot \rho \mathbf{u} z = \nabla \cdot (\rho \mathcal{D} \nabla z) \quad (4.3)$$

où  $\tau = \mu (\nabla \mathbf{u} + \nabla^t \mathbf{u} - \frac{2}{3} (\nabla \cdot \mathbf{u}) \nabla I)$ , et la viscosité s'écrit sous la forme  $\mu = \mu_0 \left( \frac{T}{T_0} \right)^r$  avec  $r = 0.7$ ,  $T_0 = 298$  K et  $\mu_0 = 1.85 \times 10^{-5}$  kg/m-s. Les gaz sont supposés suivre la loi des gaz parfaits :

$$P = \rho \frac{R}{W} T, \quad \frac{1}{W} = \sum_{k=1}^N \frac{Y_k}{W_k} \quad (4.4)$$

où  $W$  est la masse moléculaire moyenne du mélange.  $W_k$  et  $Y_k$  sont respectivement la masse atomique et la fraction massique de l'espèce  $k$ . La diffusivité  $\mathcal{D}$ , que l'on prend ici égale à 1, vérifie

$\rho D = \frac{\mu}{Pr}$ . Dans le modèle de flamme laminaire de Burke–Schumann, les fractions massiques et températures sont données par :

– Dans l’injection de combustible, ( $z > z_{st}$ )

$$\begin{aligned} Y_F(z) &= z Y_F^0 + (z - 1) \frac{Y_O^0}{s} = Y_F^0 \frac{z - z_{st}}{1 - z_{st}} \\ Y_O(z) &= 0 \\ T(z) &= z T_F^0 + (1 - z) T_O^0 + z_{st} \frac{Q Y_F^0}{C_p} \frac{1 - z}{1 - z_{st}} \end{aligned}$$

– Dans l’injection d’air, ( $z < z_{st}$ )

$$\begin{aligned} Y_F(z) &= 0 \\ Y_O(z) &= Y_O^0 \left(1 - \frac{z}{z_{st}}\right) \\ T(z) &= z T_F^0 + (1 - z) T_O^0 + \frac{Q Y_F^0}{C_p} z \end{aligned}$$

La valeur stœchiométrique  $z_{st}$  est donnée par :

$$z_{st} = \left(1 + \frac{s Y_F^0}{Y_O^0}\right)^{-1} = \left(1 + \frac{\nu_O W_O Y_F^0}{\nu_F W_F Y_O^0}\right)^{-1} \quad (4.5)$$

Les propriétés des réactifs et de l’environnement sont :

Constante des gaz	$R = 287.04 \text{ J/kg/K}$
Pression atmosphérique	$P = 101325 \text{ Pa}$
Chaleur spécifique	$c_p = 1383.2 \text{ J/kg/K}$
Taux de dissipation de chaleur	$Q = 44205545.4 \text{ J/kg}$
Masse atomique du combustible ( $\text{CH}_4$ )	$W_F = 16. \times 10^{-3} \text{ kg/mol}$
Masse atomique de l’air ( $\text{O}_2$ )	$W_O = 32. \times 10^{-3} \text{ kg/mol}$
Nombre de Prandtl	$Pr = 0.75$
Gravité	$g_z = -9.81 \text{ m/s}^2$

Le paramètre de dissipation de chaleur  $Q/c_p$  est déterminé à l’aide d’une estimation du pic de température (Xu and Smooke, Mitchell *et al.*). Les conditions initiales et aux limites sont les suivantes :

– Conditions initiales :

$$\mathbf{u}(r, z) = 0, \quad z(r, z) = 0; \quad (4.6)$$

– Conditions aux limites :

–  $r \leq r_I, z = 0$

$$u_r(r, 0) = 0, \quad u_z(r, 0) = 0.045, \quad z(r, 0) = 1; \quad (4.7)$$

–  $r_I < r < r_O, z = 0$

$$u_r(r, 0) = 0, \quad u_z(r, 0) = 0.0988, \quad z(r, 0) = 0; \quad (4.8)$$

–  $r = r_O$

$$u_r(r_O, z) = 0, \quad u_z(r_O, z) = 0, \quad \frac{\partial z}{\partial n}(r_O, z) = 0; \quad (4.9)$$

–  $r = 0$

$$u_r(0, z) = 0, \quad \frac{\partial u_z}{\partial n}(0, z) = 0, \quad \frac{\partial z}{\partial n}(0, z) = 0; \quad (4.10)$$

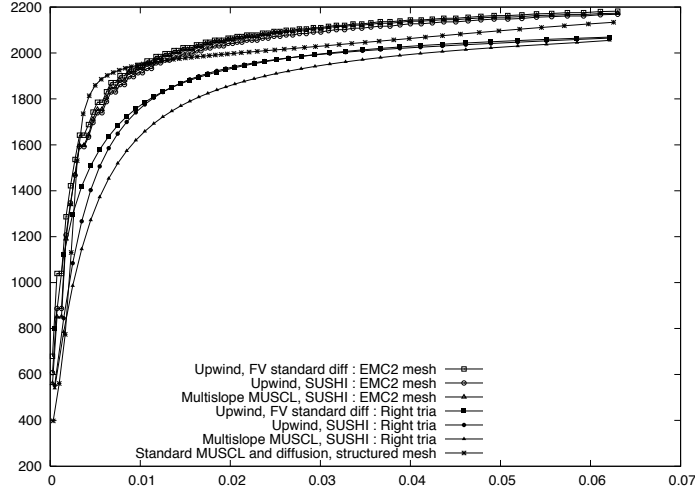


FIG. 4.7 – Flamme laminaire 3D : coupes axiale de la température sur l’axe  $y$ . Comparaison avec les résultats obtenus par Chiang *et al.*

$$-z = L \quad \tau - p \nabla I = 0, \quad \frac{\partial z}{\partial n}(r, L) = 0; \quad (4.11)$$

On génère à l’aide de GMSH [52] le maillage d’un quart du domaine par des tétraèdres, voir figure 4.8(a), raffiné près du cylindre généré par l’interface entre l’injection de combustible et d’air. On impose des conditions de symétrie sur les murs intérieurs (droits). Le maillage ainsi créé est composé de 33398 tétraèdres, soit 70842 noeuds de vitesse. On utilise un pas de temps  $\Delta t = 5.10^{-5}$  s.

Les résultats montrent un très bon accord avec les simulations conduites par Chiang *et al.*, avec une hauteur de flamme (distance entre l’injection et le point où la température commence à décroître) identique, et la même température en haut du cylindre, à la sortie de l’écoulement, voir figure 4.7. On représente sur la figure 4.8(b) le champ de température sur un plan de coupe du domaine, et sur la figure 4.8(c) les surfaces d’isotempérature dans un demi-domaine, au temps  $t = 0.87$  s.

Contrairement à ce que l’on peut observer lors de la simulation sur un maillage 2D axi-symétrique de rectangles, la solution ne s’approche pas au cours du temps d’un écoulement permanent. Cela est très probablement dû au caractère non-structuré du maillage du quart de cylindre qui, n’étant pas à symétrie radiale, fait apparaître des dissymétries dans la solution.



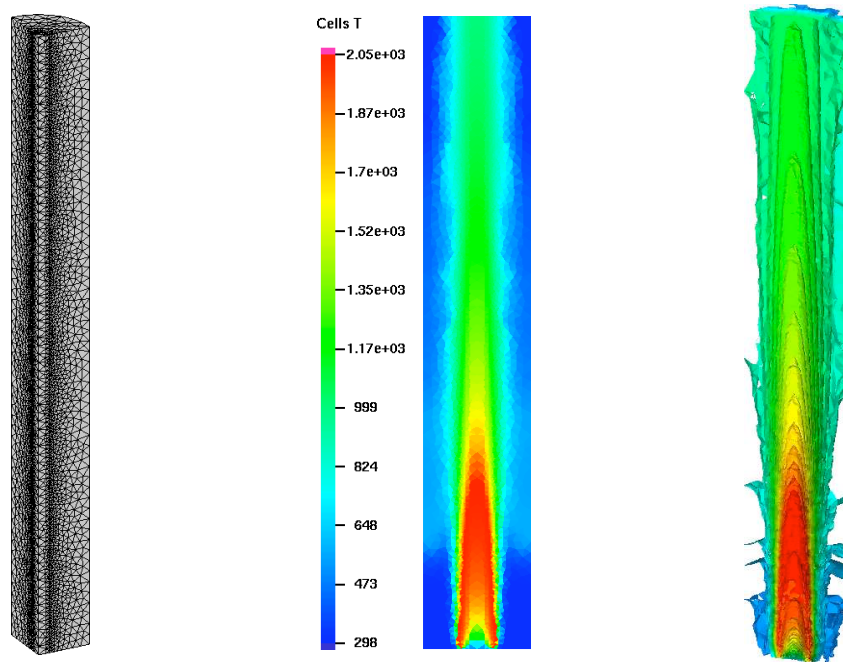


FIG. 4.8 – (a) Maillage utilisé; (b) Plan de coupe (symétrisé); (c) Surfaces d'isotherme

### 4.3 Aéraulique anisotherme à faible nombre de Mach – Tétraèdres

Ce deuxième cas correspond à la simulation d'une aéraulique anisotherme due à une source de chaleur rectangulaire dans un local de  $17,28\text{ m}^3$  muni d'une porte. Il est dérivé d'un cas de validation du code ISIS dans le cadre de l'étude des incendies dans des milieux confinés et ventilés. La modélisation s'effectue par les équations de Navier–Stokes classiques, incluant la conservation de l'énergie sous la forme d'une équation de convection-diffusion pour la température, voir annexe B section B.5.

On raffine le maillage près des deux zones d'intérêt que sont la source de chaleur et la porte, situées respectivement sur la face inférieure et sur une des faces verticales du local, voir figure 4.9. Le maillage est ainsi composé de 43450 tétraèdres (soit 88939 noeuds de vitesse et 43450 noeuds de pression), et a été généré par le logiciel libre GMSH [52] couplé au générateur de maillages tétraédriques de Delaunay TetGen [101]. La simulation est conduite jusqu'au temps  $t = 100\text{ s}$  à l'aide de 200000 pas de temps de  $5 \cdot 10^{-4}\text{ s}$ . Le calcul est réparti sur 16 processeurs entre lesquels les portions de maillage sont équitablement réparties à l'aide de la librairie METIS [75, 74], conduisant à une utilisation d'environ 180 Mo de mémoire par ordinateur.

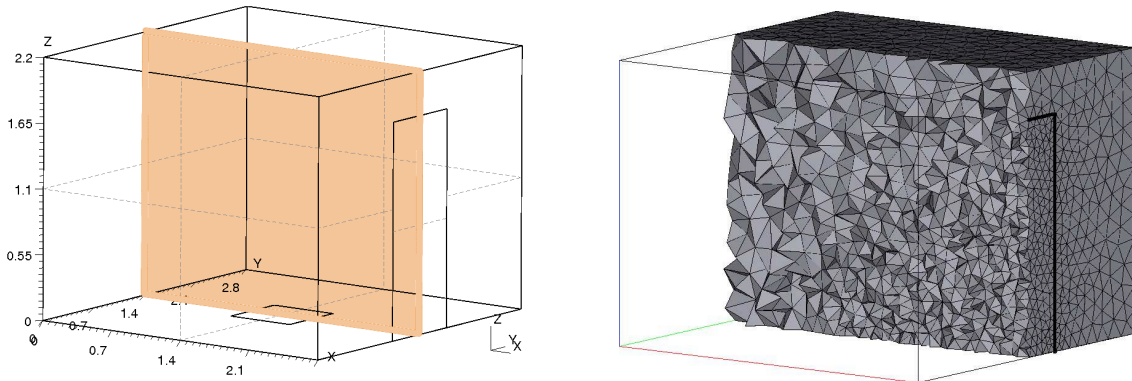


FIG. 4.9 – Schéma du local avec plan de coupe utilisé pour la figure 4.10 ; Coupe du maillage

Dans les premiers instants, un panache se forme, génère des rouleaux de convection et n'interfère que peu avec l'extérieur du local. Par la suite, quand le panache a atteint le plafond, la répartition de température se stratifie peu à peu, ce qui est illustré par l'aplatissement des isosurfaces de température sur les figures 4.11-4.13, à distance du panache et de la porte. De l'air froid est aspiré par le bas de la porte alors que de l'air chaud est rejeté par le haut, ce qui abaisse la température dans le voisinage de la porte. Ce phénomène est illustré par la distortion des isosurfaces correspondant aux températures les plus basses qui suivent le contour de la porte, et par les résultats aux temps avancés, qui présentent un panache incliné par les échanges avec l'extérieur, le reste du domaine étant essentiellement stratifié. On représente sur la figure 4.10 la carte des vecteurs vitesses à l'instant  $t = 100\text{ s}$ , dans le plan de coupe défini sur la figure 4.9, colorés par la valeur de la composante horizontale  $x$  de la vitesse.

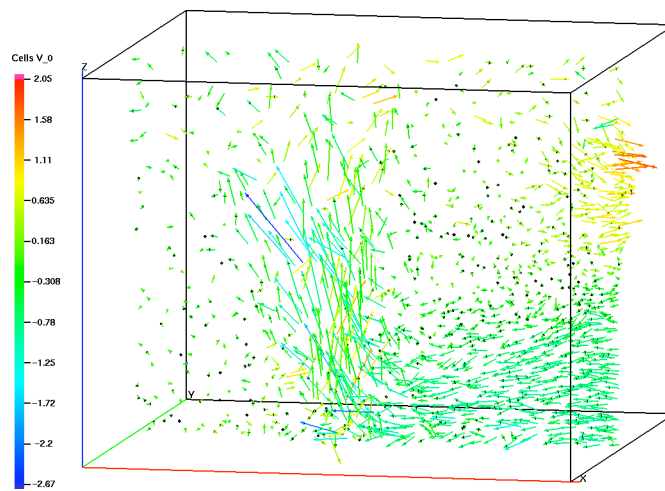


FIG. 4.10 – Vecteurs vitesse colorés par la vitesse suivant l'axe  $x$  à  $t = 100$  s

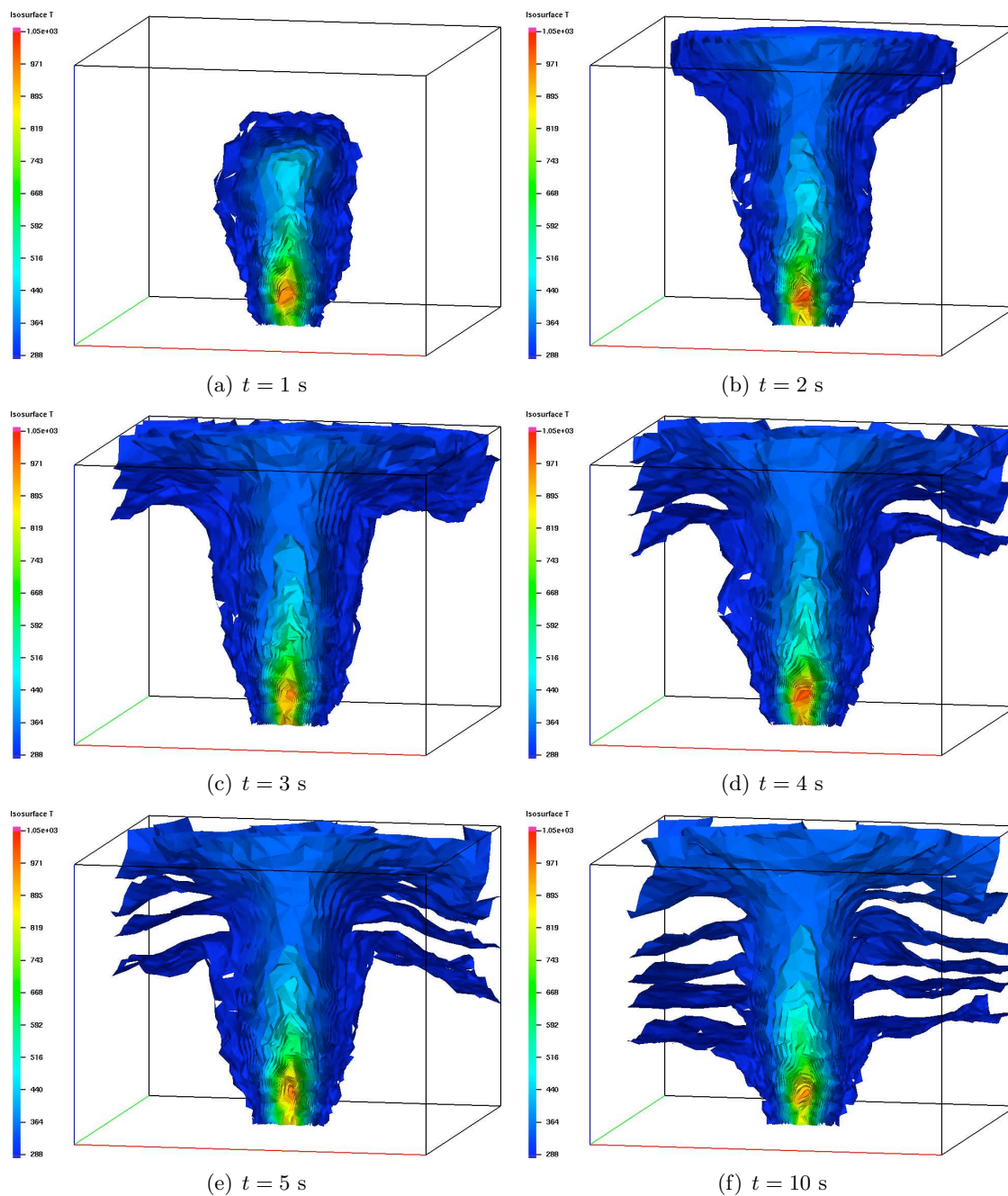


FIG. 4.11 – Isosurfaces de température dans un demi-domaine - 1

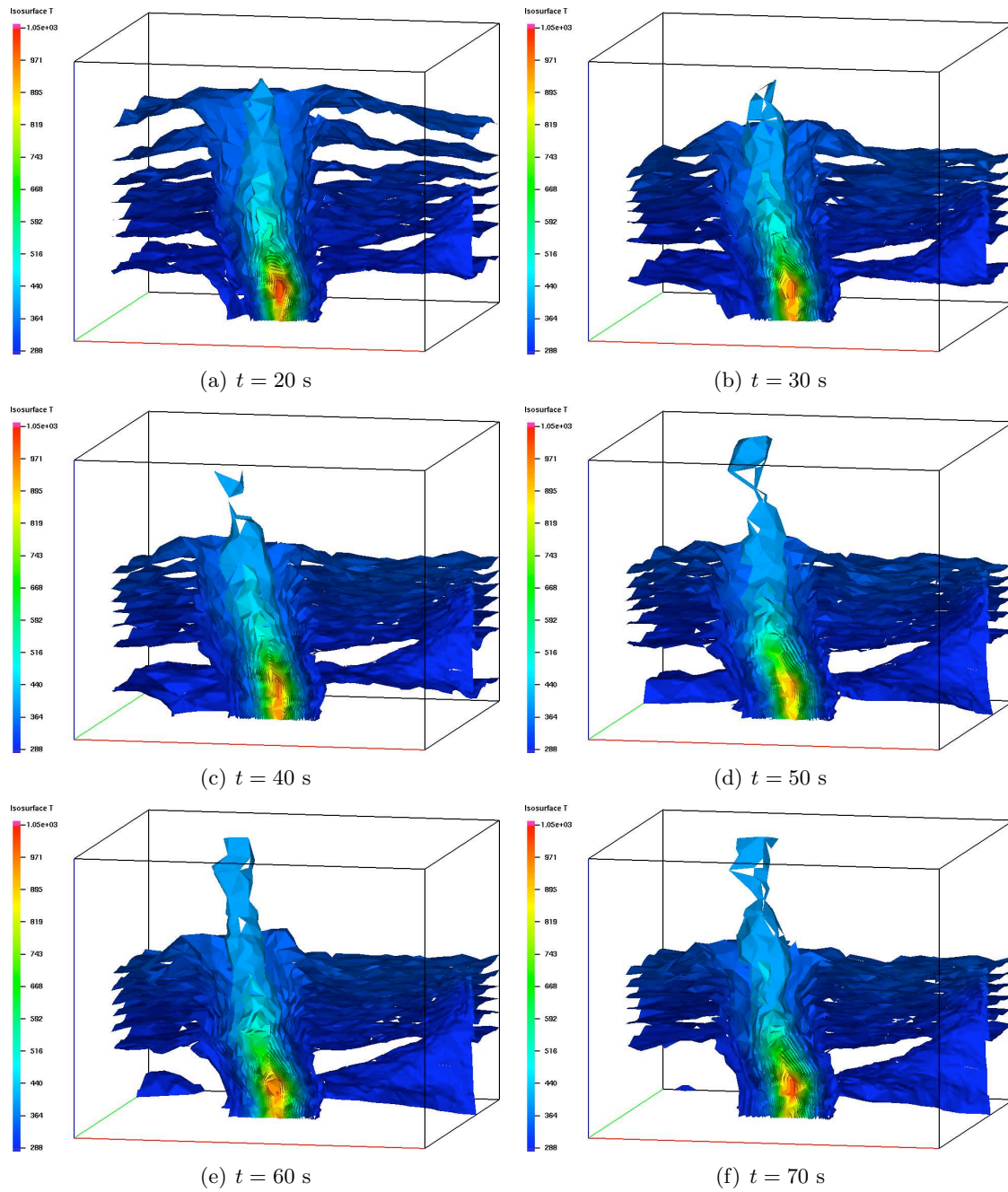


FIG. 4.12 – Isosurfaces de température dans un demi-domaine - 2

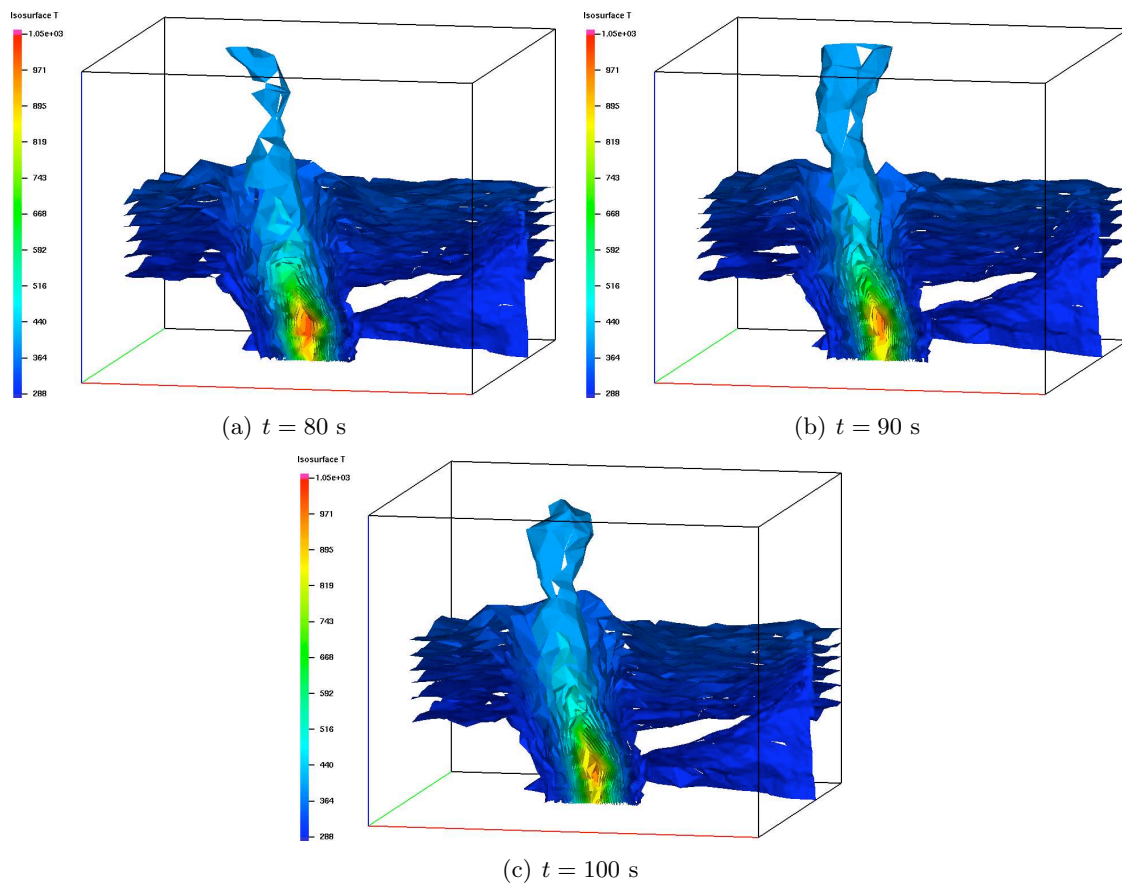


FIG. 4.13 – Isosurfaces de température dans un demi-domaine - 3

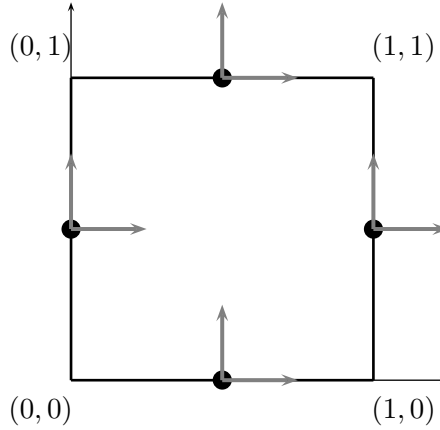


# Annexe A

## Elements finis non conformes

### A.1 Element de Rannacher-Turek

On décrit ici l'élément non-conforme de Rannacher-Turek [96], qui se décline en plusieurs versions aux propriétés spécifiques.



Soit  $\mathcal{T}_h$  un ensemble de mailles  $T$  tel que  $\bigcup_{T \in \mathcal{T}_h} \bar{T} = \bar{\Omega}$ . Les fonctions sont approchées sur chaque élément  $T \in \mathcal{T}_h$  dans l'ensemble

$$\tilde{Q}_1(T) = \{q \circ \Psi_T^{-1} : q \in \text{span}(1, x_i, x_i^2 - x_{i+1}^2), i = 1 \dots n\} \quad (\text{A.1})$$

où  $\Psi_T : \hat{T} \rightarrow T$  est la transformation bilinéaire qui permet d'obtenir  $T$  à partir du carré de référence  $\hat{T}$ .

L'espace d'approximation en vitesse est défini par

$$\mathbf{H}_h^{(a/b)} = \left\{ \begin{array}{l} \mathbf{v}_h \in \mathbf{L}^2(\Omega) : \mathbf{v}_h \in \tilde{Q}_1(T)^n, \forall T \in \mathcal{T}_h, \mathbf{v}_h \text{ continue vis-à-vis} \\ \text{de toutes les fonctionnelles nodales } F_{\Gamma,i}^{(a/b)}(\cdot), \\ \text{et } F_{\Gamma,i}^{(a/b)}(\mathbf{v}_h) = 0 \text{ si } \Gamma \subset \partial\Omega \end{array} \right\} \quad (\text{A.2})$$

où les fonctionnelles nodales sont, pour la variante (a), la moyenne intégrale pondérée sur la face  $\Gamma$  :



$$F_{\Gamma,i}^{(a)}(\mathbf{v}) = \frac{1}{|\Gamma|} \int_{\Gamma} v_i(x) d\sigma \quad (\text{A.3})$$

et pour la variante (b), la valeur du champ au centre de la face :

$$F_{\Gamma,i}^{(b)}(\mathbf{v}) = v_i(x_{\Gamma}) \quad (\text{A.4})$$

où  $x_{\Gamma}$  est le centre de la face  $\Gamma$ .

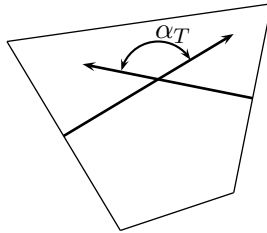
Les espaces  $\mathbf{H}_h^{(a/b)}$  étant non conformes, il faut définir des formes bilinéaires et des normes par morceaux :

$$a_h(v, w) = \sum_{T \in \mathcal{T}_h} a(v, w)_T, \quad |v|_h = a_h(v, v)^{1/2} \quad (\text{A.5})$$

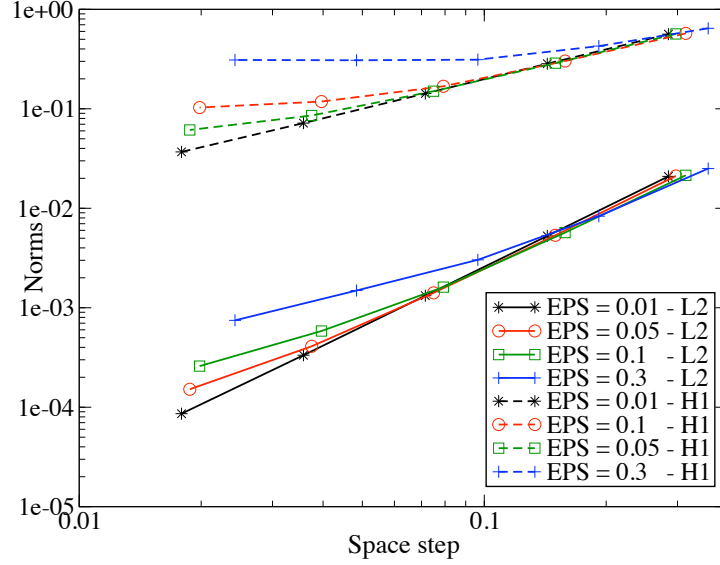
Pour les deux versions, (a) et (b), l'estimation d'erreur d'interpolation s'écrit :

$$|\mathbf{v} - i_h \mathbf{v}| + h |\mathbf{v} - i_h \mathbf{v}|_h \leq ch(h + \sigma_h) |\mathbf{v}|_2, \quad \mathbf{v} \in \mathbf{H}_0^1 \cap \mathbf{H}^2 \quad (\text{A.6})$$

où  $\sigma_h = \max\{|\pi - \alpha_T|, T \in \mathcal{T}_h\}$ ,  $\alpha_T \in (0, \pi)$  étant l'angle maximum entre les normales unitaires de deux faces opposées de  $T$ .  $\sigma_h$  est une mesure de la dégénérescence du maillage  $\mathcal{T}_h$ .



On peut illustrer l'effet de la dégénérescence du maillage sur l'erreur d'interpolation. Sur quatre séries de maillages différents caractérisées par une mesure  $\varepsilon$  de la perturbation depuis un maillage régulier identique et des raffinements successifs, on mesure l'erreur commise en interpolant la fonction  $(x, y) \mapsto \sin(\pi x) \cos(\pi y)$ . On obtient les courbes d'erreur d'interpolation suivantes en norme  $L^2$  et  $H^1$  :



Le paramètre de dégénérescence  $\sigma_h$  est lié au paramètre de perturbation  $\varepsilon$ . Si pour des maillages grossiers, on retrouve l'ordre  $O(h^2)$  en norme  $L^2$  et  $O(h)$  en norme  $H^1$ , en revanche, en raffinant le maillage, le paramètre  $\sigma_h$  devient prépondérant dans le terme  $h + \sigma_h$  et on retrouve bien en augmentant la valeur de  $\varepsilon$  une convergence d'ordre  $O(h)$  en norme  $L^2$  et  $O(1)$  en norme  $H^1$ . Cette perte d'ordre de précision est due au caractère non isoparamétrique des espaces  $\mathbf{H}_h^{(a/b)}$  : les transformations multilinéaires  $\Psi_T : \hat{T} \rightarrow T$  ne sont pas du même type polynomial que les fonctions de forme sur  $\hat{T}$ . Pour l'élément conforme isoparamétrique, on retrouve l'estimation d'erreur optimale. Pour pallier à ce problème tout en gardant le caractère non-conforme, il faut utiliser la version non paramétrique de l'élément de Rannacher-Turek. Pour tout élément  $T \in \mathcal{T}_h$ , soit  $\xi_i$  un système de coordonnées local obtenu en reliant les centres des faces opposées de  $T$ . Pour chaque  $T \in \mathcal{T}_h$ , on re-définit l'ensemble d'approximation des fonctions sur cette maille sans passer par une fonction de transformation :

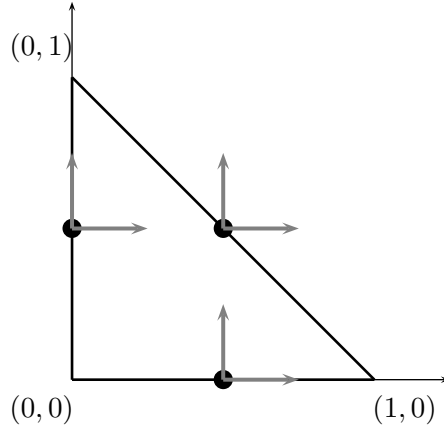
$$\tilde{Q}_1(T) = \text{span}(1, \xi_i, \xi_i^2 - \xi_{i+1}^2), i = 1 \dots n \quad (\text{A.7})$$

Ainsi,  $\tilde{Q}_1(T)$  contient automatiquement les polynômes linéaires, et l'on retrouve l'estimation d'erreur optimale en appliquant le lemme de Bramble-Hilbert à tout élément  $T \in \mathcal{T}_h$  sans se référer à  $\hat{T}$  :

$$|\mathbf{v} - i_h \mathbf{v}| + h|\mathbf{v} - i_h \mathbf{v}|_h \leq ch^2 |\mathbf{v}|_2, \quad \mathbf{v} \in \mathbf{H}_0^1 \cap \mathbf{H}^2 \quad (\text{A.8})$$

Pour l'extension d'ISIS aux maillages non-structurés de triangles ou de tétraèdres, on utilise l'élément de Crouzeix-Raviart [30].

## A.2 Element de Crouzeix-Raviart



Pour cet élément, les fonctions sont approchées sur chaque maille  $T \in \mathcal{T}_h$  dans l'ensemble

$$P_1(T) = \{q \circ \Psi_T^{-1} : q \in \text{span}(1, x_i, x_i x_{i+1}), i = 1 \dots n\} \quad (\text{A.9})$$

où  $\Psi_T : \hat{T} \rightarrow T$  est la transformation bilinéaire qui permet d'obtenir  $T$  à partir du carré de référence  $\hat{T}$ .

Comme l'approximation est bilinéaire sur chaque maille, les fonctionnelles  $F_{\Gamma,i}^{(a)}$  et  $F_{\Gamma,i}^{(b)}$  renvoient la même valeur, et donc l'espace d'approximation est défini de manière analogue à celui de l'élément de Rannacher-Turek :

$$\mathbf{H}_h = \left\{ \mathbf{v}_h \in \mathbf{L}^2(\Omega) : \mathbf{v}_h \in P_1(T)^n, \forall T \in \mathcal{T}_h, \mathbf{v}_h \text{ continue vis-à-vis} \right. \\ \left. \begin{array}{l} \text{de toutes les fonctionnelles nodales } F_{\Gamma,i}(\cdot), \\ \text{et } F_{\Gamma,i}(\mathbf{v}_h) = 0 \text{ si } \Gamma \subset \partial\Omega \end{array} \right\} \quad (\text{A.10})$$

où la fonctionnelle nodale  $F_{\Gamma,i}(\cdot)$  est de manière équivalente, la moyenne intégrale pondérée sur la face  $\Gamma$  et la valeur du champ au centre de la face :

$$F_{\Gamma,i}(\mathbf{v}) = \frac{1}{|\Gamma|} \int_{\Gamma} v_i(x) d\sigma = v_i(x_{\Gamma}) \quad (\text{A.11})$$

où  $x_{\Gamma}$  est le centre de la face  $\Gamma$ .

## Annexe B

# A Finite Volume Stability Result for the Convection Operator in Compressible Flows... and Some Finite Element Applications

### B.1 Introduction

Let  $\rho$  and  $\mathbf{u}$  be a scalar and a vector smooth function respectively, defined over a domain  $\Omega$  of  $\hat{x}^d$ ,  $d = 2$  or  $d = 3$ , and such that the following identity holds in  $\Omega$  :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (\text{B.1})$$

Let  $z$  be a smooth scalar function defined over  $\Omega$ . Then the following stability identity is known to hold :

$$\int_{\Omega} \left[ \frac{\partial \rho z}{\partial t} + \nabla \cdot (\rho z \mathbf{u}) \right] z = \frac{1}{2} \frac{d}{dt} \int_{\Omega} \rho z^2 + \frac{1}{2} \int_{\partial \Omega} \rho z^2 \mathbf{u} \cdot \mathbf{n} \quad (\text{B.2})$$

When  $\rho$  stands for the density and  $\mathbf{u}$  for the velocity, equation (B.1) is the usual mass balance in a variable density flow. Choosing for  $z$  a component of the velocity, equation (B.2) yields the central argument of the kinetic energy conservation theorem.

In this paper, we first derive a finite volume analog of relation (B.2) ; the statement of this stability estimate is the object of section B.2. Then, in section B.3, we show how this result may be used to build a  $L^2$ -stable convection operator for the Rannacher-Turek [96] or Crouzeix-Raviart [30] low order non-conforming Stokes finite elements, switching for this term from the finite element discretization to a finite volume approximation based on a dual mesh ; note that a similar technique is implemented in [2, 42] for the solution of convection-diffusion type equations (so with a known continuous velocity field), however, for a different purpose, namely to satisfy a discrete maximum principle. This discretization is applied to solve the balance equations of the asymptotic model for low Mach number flows, which reads :

$$\begin{cases} \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \nabla \cdot \tau(\mathbf{u}) = \mathbf{f}_v \\ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \end{cases} \quad (\text{B.3})$$

where  $p$  stands for the pressure,  $\tau(\mathbf{u})$  is the shear stress tensor and  $\mathbf{f}_v$  is a forcing term. The density  $\rho$  is supposed to depend on a state variable of the fluid (*e.g.* the temperature, the composition... but not the pressure) which is solution to an additional balance equation. For this problem, we show in particular how the necessity of controlling the terms which arise in the kinetic energy balance under the form of boundary integrals, both from the second term at the right-hand side of (B.2) and from the integration by parts of the divergence of the stress tensor, suggests a boundary condition for artificial boundaries. This has been implemented in the ISIS code developed at IRSN, with the purpose to allow this open-source CFD tool to cope with the simulation of fires in an open atmosphere ; it is assessed in section B.5 against a natural convection model problem.

## B.2 A finite volume result

Let a finite volume admissible mesh  $\mathcal{M}$  (in the sense of [35], Chapter 3) of the calculational domain  $\Omega$  be given. This mesh is composed of a family  $\mathcal{T}$  of control volumes, which are convex disjoint polygons ( $d = 2$ ) or polyhedrons ( $d = 3$ ) included in  $\Omega$  and such that  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} \bar{K}$ . For each neighbouring control volume  $L$  of  $K \in \mathcal{M}$ ,  $\sigma = K|L$  denotes the common edge or face of  $K$  and  $L$ . The sets  $\mathcal{E}_{\text{int}}$ ,  $\mathcal{E}_{\text{ext}}$  and  $\mathcal{E}(K)$  stand respectively for the internal edges or faces (*i.e.* separating two control volumes), the external ones (*i.e.* included in the boundary) and the edges or faces of the control volume  $K$ . By  $|K|$  and  $|\sigma|$ , we denote hereafter the  $d$ - and  $(d - 1)$ -dimensional measures of  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}$ , respectively.

Let  $(\rho_K^*)_{K \in \mathcal{T}}$  and  $(\rho_K)_{K \in \mathcal{T}}$  be two families of positive real numbers satisfying the following set of equations :

$$\forall K \in \mathcal{T}, \quad \frac{|K|}{\delta t} (\rho_K - \rho_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_{\sigma,K} = 0 \quad (\text{B.4})$$

where  $F_{\sigma,K}$  is a quantity associated with the edge  $\sigma$  and with the control volume  $K$  ; we suppose that, for any internal edge  $\sigma = K|L$ ,  $F_{\sigma,K} = -F_{\sigma,L}$ . Equation (B.4) may be seen as the finite-volume counterpart of the continuous mass balance (B.1).

Let  $(z_K^*)_{K \in \mathcal{T}}$  and  $(z_K)_{K \in \mathcal{T}}$  be two families of real numbers. For any internal edge or face  $\sigma = K|L$ , we define  $z_\sigma$  either by  $z_\sigma = \frac{1}{2}(z_K + z_L)$ , or by  $z_\sigma = z_K$  if  $F_{\sigma,K} \geq 0$  and  $z_\sigma = z_L$  otherwise. The first choice is usually referred to as “the centered choice”, and the second one as “the upwind choice” with respect to the quantity  $F_{\sigma,K}$ . For an external edge or face, if  $F_{\sigma,K} \geq 0$ , we suppose that  $z_\sigma = z_K$  (*i.e.* that the upwind choice is made, which seems to be the only natural possibility in this case), and if  $F_{\sigma,K} \leq 0$ , we suppose that  $z_\sigma$  is given by a relation which we do not need to precise for the moment. Then we can state the following stability result.

**Théorème 2 (Stability of the convection operator)** *With the above definitions, the following stability estimate holds, for both the centered or upwind choice for the quantities  $z_\sigma$  :*

$$\sum_{K \in \mathcal{T}} z_K \left[ \frac{|K|}{\delta t} (\rho_K z_K - \rho_K^* z_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_{\sigma,K} z_\sigma \right] \geq \frac{1}{2} \sum_{K \in \mathcal{T}} \frac{|K|}{\delta} t \left[ \rho_K z_K^2 - \rho_K^* z_K^{*2} \right] + \frac{1}{2} \sum_{\substack{\sigma \in \mathcal{E}_{\text{ext}} \\ (\sigma \in \mathcal{E}(K))}} F_{\sigma,K} z_\sigma^2 \quad (\text{B.5})$$

In the case of a velocity vanishing on the boundary of the calculational domain, a proof of this result can be found in [48].

## B.3 A convection operator for low-order non-conforming finite elements

We now turn to the discretization of Navier-Stokes equations (B.3) by a low-order mixed finite element method.

### B.3.1 Discretization spaces

We now suppose that the control volumes are either convex quadrilaterals ( $d = 2$ ), hexahedra ( $d = 3$ ) or simplices. In the first case, the spatial discretization relies on the so-called “rotated bilinear element”/ $P_0$  introduced by Rannacher and Turek [96] (RT in the following); for simplicial meshes, the Crouzeix-Raviart element [30] (CR in the following) is used. The reference element  $\hat{K}$  for the RT element is the unit  $d$ -cube (with edges parallel to the coordinate axes); the discrete functional space on  $\hat{K}$  is  $\tilde{Q}_1(\hat{K})^d$ , where  $\tilde{Q}_1(\hat{K})$  is defined as follows :

$$\tilde{Q}_1(\hat{K}) = \text{span} \{1, (x_i)_{i=1,\dots,d}, (x_i^2 - x_{i+1}^2)_{i=1,\dots,d-1}\}$$

The reference element for the CR element is the unit  $d$ -simplex and the discrete functional space is the space  $P_1$  of affine polynomials. For both velocity elements used here, the degrees of freedom are determined by the following set of nodal functionals on the discrete velocity space :

$$\{\varphi_{\sigma,i}, \sigma \in \mathcal{E}(K), i = 1, \dots, d\}, \quad \varphi_{\sigma,i}(\mathbf{v}) = |\sigma|^{-1} \int_{\sigma} \mathbf{v}_i d\sigma$$

The mapping from the reference element to the actual one is, for the RT element, the standard  $Q_1$  mapping and, for the CR element, the standard affine mapping. Finally, in both cases, the continuity of the average value of discrete velocities (*i.e.*, for a discrete velocity field  $\mathbf{v}$ ,  $\varphi_{\sigma,i}(\mathbf{v})$ ,  $1 \leq i \leq d$ ) across each edge or face of the mesh is required, and, as usual in finite elements methods, Dirichlet conditions are built-in in the approximation space, thus the discrete space  $\mathbf{W}_h$  is defined as follows :

$$\begin{aligned} \mathbf{W}_h = & \{ \mathbf{v}_h \in L^2(\Omega)^d : \mathbf{v}_h|_K \in W(K)^d, \forall K \in \mathcal{T}; \\ & \varphi_{\sigma,i}(\mathbf{v}_h) \text{ continuous across each edge } \sigma \in \mathcal{E}_{\text{int}}, 1 \leq i \leq d; \\ & \varphi_{\sigma,i}(\mathbf{v}_h) = |\sigma|^{-1} \int_{\sigma} \mathbf{u}_{D,i} d\sigma, \forall \sigma \in \mathcal{E}_{\text{ext,D}}, 1 \leq i \leq d \} \end{aligned}$$

where  $W(K)$  is the discrete functions space on  $K$ ,  $\mathcal{E}_{\text{ext,D}}$  is the set of the external edges included in the part of the boundary where the velocity is prescribed and  $\mathbf{u}_D$  is this prescribed velocity. From this definition, each velocity degree of freedom can be associated with an element edge. Hence, the set of velocity degrees of freedom may be written as  $\{\mathbf{v}_{\sigma,i}, \sigma \in \mathcal{E} \setminus \mathcal{E}_{\text{ext,D}}, 1 \leq i \leq d\}$ . We define  $\mathbf{v}_{\sigma} = \sum_{i=1}^d \mathbf{v}_{\sigma,i} e^{(i)}$  where  $e^{(i)}$  is the  $i^{\text{th}}$  vector of the canonical basis of  $\hat{x}^d$ .

For both RT and CR discretizations, the pressure is approximated by piecewise constant functions. The same approximation is used for the density.

### B.3.2 A convection operator

The natural finite element method for RT and CR elements cannot be  $L^2$ -stable; indeed, the derivation of a stability estimate of the form of (B.2) involves integrations by parts, which,

because of the non-conformity of the discretization, makes uncontrolled jumps across the edges ( $d = 2$ ) or faces ( $d = 3$ ) of the elements appear. We thus approximate both the unsteady term (*i.e.*  $\partial \rho \mathbf{u} / \partial t$ ) and the convection term (*i.e.*  $\nabla \cdot \rho \mathbf{u} \otimes \mathbf{u}$ ) by a finite volume discretization, using for control volumes a dual mesh. From the definition of the velocity degrees of freedom  $\mathbf{v}_\sigma$ , a control volume for each of these latter must be associated with an edge or a face  $\sigma$ . For  $\sigma \in \mathcal{E}_{\text{int}}$ ,  $\sigma = K|L$ , this dual cell is defined as the union of the two cones of common basis  $\sigma$  and vertices  $x_K$  and  $x_L$  respectively, where  $x_K$  (resp.  $x_L$ ) is the mass center of  $K$  (resp.  $L$ ), see Figure B.1; for  $\sigma \in \mathcal{E}_{\text{ext}} \setminus \mathcal{E}_{\text{ext,D}}$ , the cell is restricted to the cone included in the adjacent primal control volume. For each considered  $\sigma$ , the corresponding dual cell is denoted by  $D_\sigma$  and called in the following the “diamond cell associated with  $\sigma$ ”.

For simplicial or parallelepipedic meshes, an important property is that  $|D_\sigma|$ , the measure of  $D_\sigma$ , is also the integral over  $\Omega$  of the shape function associated with  $\sigma$ , which shows that the definition of this cell is in some sense consistent with the results of a classical mass lumping of the unsteady term. This is the discretization used here.

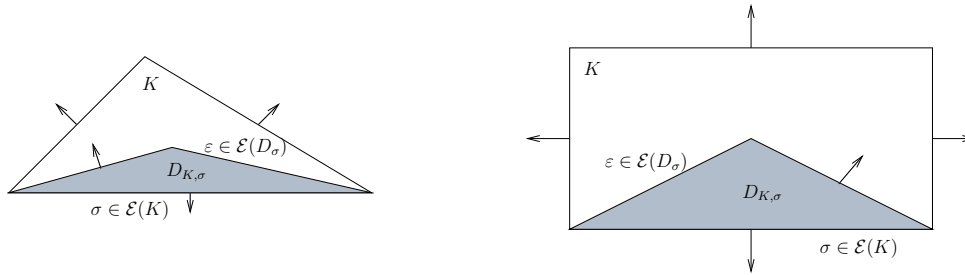


FIG. B.1 – Diamond-cells for the Crouzeix-Raviart and Rannacher-Turek element.

Making use of a finite volume technique for the term  $\nabla \cdot \rho \mathbf{u} \otimes \mathbf{u}$  with the goal of applying theorem 2 raises the problem to approximate the fluxes on the edges of the diamond cells in such a way that discrete mass balance (B.4) holds. Indeed, as pressure discrete functions are piecewise constant over the primal cells, the mixed finite element formulation yields a finite-volume-like discrete mass balance based on the primal mesh, and not on the dual one. For the CR element, it may be seen that evaluating the mass flux on the boundary of the diamond cell  $\partial D_\sigma$  from the mass fluxes at the edges or faces of the primal mesh through the finite element expansion makes the mass balance hold on  $D_\sigma$  also. The proof of this elementary result relies on the fact that the divergence of a discrete velocity is constant over each primal cell. This result is extended to RT elements on parallelepipedic meshes by designing a specific interpolation, such that the divergence of the reconstructed mass flux field is also constant mesh-by-mesh [8].

## B.4 A pressure correction scheme for low Mach number flows with open boundaries

On the basis of the preceding developments, we now derive a pressure correction scheme for the solution of system (B.3). The first equation, usually referred to as the velocity prediction step, consists in solving for a (non-divergence free) tentative velocity the momentum balance equation with the explicit pressure (*i.e.* the pressure at the previous time step); the convection term is linearized, by taking the explicit velocity as advective field. In variational form, this

discrete equation consists in searching  $\tilde{\mathbf{u}}^{n+1} \in \mathbf{W}_h$  such that,  $\forall \mathbf{v} \in \mathbf{W}_h$  :

$$\begin{aligned} \frac{1}{\delta t}(\rho^n \tilde{\mathbf{u}}^{n+1} - \rho^{n-1} \mathbf{u}^n, \mathbf{v})_h + (\nabla \cdot_h \rho^n \tilde{\mathbf{u}}^{n+1} \otimes \mathbf{u}^n, \mathbf{v})_h \\ + a(\tilde{\mathbf{u}}^{n+1}, \mathbf{v}) + b(p^n, \mathbf{v}) = (\mathbf{f}^{n+1}, \mathbf{v})_h \end{aligned} \quad (\text{B.6})$$

In this relation, the bilinear forms  $a(\cdot, \cdot)$  and  $b(\cdot, \cdot)$  stand for the usual finite element discretizations of the viscous dissipation and the pressure gradient term, respectively, and the right hand side  $\mathbf{f}$  gathers the effects of the forcing term and of the non-homogeneous Dirichlet boundary conditions. The notation  $(\cdot, \cdot)_h$  stands for a discrete  $L^2$  inner product defined by  $(\mathbf{v}, \mathbf{w})_h = \sum_{\sigma \in \mathcal{E} \setminus \mathcal{E}_{\text{ext}, D}} |D_\sigma| \mathbf{v}_\sigma \cdot \mathbf{w}_\sigma$ ,  $\forall \mathbf{v} \in \mathbf{W}_h$ ,  $\forall \mathbf{w} \in \mathbf{W}_h$ . The discrete divergence operator is defined as described in the previous section :

$$(\nabla \cdot_h \rho^n \tilde{\mathbf{u}}^{n+1} \otimes \mathbf{u}^n)_\sigma = \frac{1}{|D_\sigma|} \sum_{\substack{\varepsilon \in \mathcal{E}(D_\sigma) \\ \varepsilon \in \mathcal{E} \setminus \mathcal{E}_{\text{ext}, D}}} |\varepsilon| (\rho^n \mathbf{u}^n)_\varepsilon (\tilde{\mathbf{u}}^{n+1})_\varepsilon$$

where  $|\varepsilon|$  stands for the measure of a face or edge  $\varepsilon$  of  $D_\sigma$ , and the centered choice is made for the approximation of  $\tilde{\mathbf{u}}_\varepsilon^{n+1}$  on internal bounds; the external edges or faces of the dual mesh are also the external edges or faces of the primal one, and the approximation  $\tilde{\mathbf{u}}_\varepsilon^{n+1} = \tilde{\mathbf{u}}_\sigma^{n+1}$  is thus natural in this case. The quantities  $(\rho^n \mathbf{u}^n)_\varepsilon$  are obtained, by the interpolation previously described, from the mass fluxes appearing in the discrete mass balance at the previous time-step, because the mass balance at the current one is not solved at this stage of the algorithm : this is the reason of the time-shift of the density in this prediction step (B.6).

The second step of the algorithm is a standard algebraic projection step, and is not detailed here.

We now turn to the derivation of an artificial boundary condition. The question that we address is the following one : what could be a suitable condition for the inflow boundaries where the velocity is not prescribed ? A part of the answer may come from energy estimates : indeed, it seems reasonable to require from this boundary condition not to lead to an unstable problem. A such energy estimate is obtained by taking  $\mathbf{v} = \tilde{\mathbf{u}}^{n+1}$  in equation (B.6) (see [48] for this calculation), which yields, using the stability of the discrete convection operator (theorem 2) :

$$\frac{1}{2}(\rho^n \tilde{\mathbf{u}}^{n+1}, \tilde{\mathbf{u}}^{n+1})_h + \delta t T_{\text{visc}}^{n+1} + \delta t T_{\text{pres}}^{n+1} \leq \frac{1}{2}(\rho^{n-1} \mathbf{u}^n, \mathbf{u}^n)_h + \delta t T_D^{n+1} + \delta t T_{\partial\Omega}^{n+1}$$

The first terms in the left-hand and right-hand sides are the discrete kinetic energy at time  $t = t^{n+1}$  and  $t = t^n$  respectively. Combining this estimate with additional bounds derived from the projection step, the pressure work term  $T_{\text{pres}}^{n+1}$  would more or less cancel if the flow was incompressible and would provide a control of the discrete time derivative of the elastic potential in the compressible case [48]. In the present case, this term is unfortunately not controlled, because the low Mach number model does not appear to be energetically consistent ; we do not develop this point further here. The term  $T_D^{n+1}$  represents the contribution of the forcing term and of the non-homogeneous Dirichlet conditions collected in  $\mathbf{f}^{n+1}$  and is obtained, as usual, by absorbing the contribution of the test function  $\tilde{\mathbf{u}}^{n+1}$  to the inner product  $(\mathbf{f}^{n+1}, \tilde{\mathbf{u}}^{n+1})_h$  in the viscous dissipation term  $T_{\text{visc}}^{n+1}$ . By theorem 2, we obtain for the last term (recall that the



boundary edges or faces are the same for the primal and the dual cells) :

$$T_{\partial\Omega}^{n+1} \leq \sum_{\sigma \in \mathcal{E}_{\text{ext}} \setminus \mathcal{E}_{\text{ext},D}} -\frac{1}{2} |\sigma| (\rho^n \mathbf{u}^n)_\sigma |(\tilde{\mathbf{u}}^{n+1})_\sigma|^2 + \int_{\sigma} (\tau(\tilde{\mathbf{u}}^{n+1}) \mathbf{n}_\sigma - p^n \mathbf{n}_\sigma) \cdot \tilde{\mathbf{u}}^{n+1}$$

The scheme thus will be stable if this term can be controlled by the boundary condition, which may be obtained by replacing in variational formulation (B.6), for the inflow edges or faces where the velocity is not prescribed, the terms multiplying the test function  $\mathbf{v}$  in the following expression :

$$-\frac{1}{2} |\sigma| (\rho^n \mathbf{u}^n)_\sigma (\tilde{\mathbf{u}}^{n+1})_\sigma \cdot \mathbf{v}_\sigma + \int_{\sigma} (\tau(\tilde{\mathbf{u}}^{n+1}) \mathbf{n}_\sigma - p^n \mathbf{n}_\sigma) \cdot \mathbf{v}$$

by a known quantity. This is consistent with the following continuous boundary condition :

$$-\frac{1}{2} \rho \mathbf{u} \cdot \mathbf{n}_\sigma \mathbf{u} + \tau(\mathbf{u}) \mathbf{n}_\sigma - p \mathbf{n}_\sigma = \mathbf{f}_{\partial\Omega} \quad (\text{B.7})$$

where the field  $\mathbf{f}_{\partial\Omega}$ , defined on  $\partial\Omega$ , is a part of the data of the problem. For incompressible flows, a theoretical study of the Navier-Stokes problem complemented with this condition can be found in [13, chapter V]; the conclusion is that the problem is well-posed. Note that this operation may be realised very simply in practice : for a concerned edge or face  $\sigma$ , the integrals involving the stress tensor and the pressure are not computed, the convection term is divided by 2 and the integral of  $\mathbf{f}_{\partial\Omega} \cdot \mathbf{v}$  over  $\sigma$  is added.

## B.5 Numerical test : a natural convection flow with open boundaries

To assess the behaviour of the presented scheme, we address a natural convection flow with artificial boundary conditions. The geometry of the calculational domain is shown in Figure B.2. For this test, system (B.3) is complemented by a standard energy balance, *i.e.* a linear convection-diffusion equation for the temperature, which is solved by an usual finite volume method. The boundary conditions are the following : on  $\partial\Omega_D$ , the velocity is set to zero and the temperature is fixed to  $T = 900^\circ\text{C}$ ; on  $\partial\Omega_{S_1}$  and  $\partial\Omega_{S_2}$ , the normal velocity is set to zero and a free slip is allowed, while the normal gradient of the temperature is set to zero; on the outflow part of  $\partial\Omega_{I/O}$ , a zero traction ( $\tau(\mathbf{u})\mathbf{n} - p\mathbf{n} = 0$ ) and a zero normal temperature gradient are imposed; on the inflow part of  $\partial\Omega_{I/O}$ , the artificial boundary condition (B.7) developed in the previous section is applied with  $\mathbf{f}_{\partial\Omega} = 0$  while the temperature is set to  $T = 300^\circ\text{C}$ . Note that the partition of  $\partial\Omega_{I/O}$  in an inflow and outflow part is determined by the calculation itself. The fluid obeys the ideal gas law, with a constant equal to  $R = 287$  and a constant pressure of  $101325 \text{ Pa}$ . The viscosity is fixed at the value of  $1.68 \cdot 10^{-5} \text{ Pa}\cdot\text{s}$ , the specific heat capacity under constant pressure is given by  $c_p = R\gamma/(\gamma - 1)$  with  $\gamma = 1.4$  and the Prandtl number is equal to 0.7. The width of the domain is  $l = 0.01 \text{ m}$  and the height  $h$  is adjusted in such a way that the Rayleigh number, based on the height of the heated part of the boundary, is equal to  $10^6$  (so  $h = 0.062 \text{ m}$ ).

Results obtained with a  $58 \times 248$  regular grid are shown in Figure B.2. The steady state is obtained through a transient, starting from the initial condition  $\mathbf{u} = 0$  and  $T = 300^\circ\text{C}$ . The

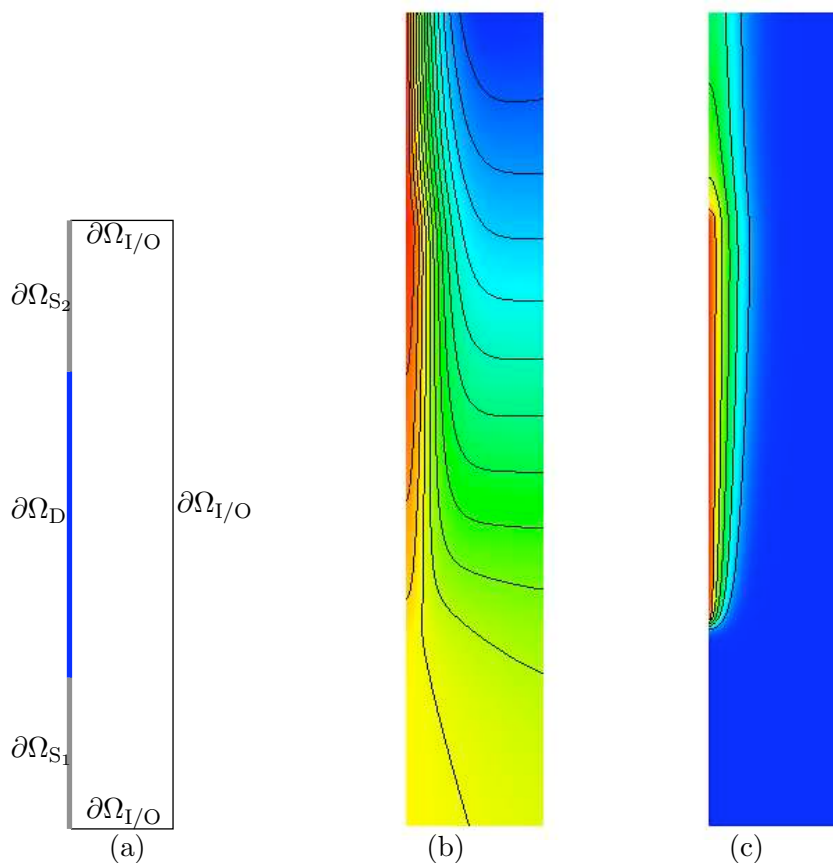


FIG. B.2 – (a) calculational domain, (b) streamlines and (c) isovalues of the temperature ( $T = 400^\circ C$ ,  $T = 500^\circ C$ ,  $T = 600^\circ C$ ,  $T = 700^\circ C$  and  $T = 900^\circ C$ )

flow enters the domain on almost the whole part of  $\partial\Omega_{I/O}$ , except in the left side of the top boundary.

In addition, calculations with a Rayleigh number of  $10^7$  and  $10^8$  were performed. No instability was seen during any of these runs. The essential effect observed when raising the Rayleigh number is a shrinking of the velocity and temperature boundary layer near the left boundary. Finally, we conducted calculations with a larger domain; the obtained velocity and temperature profiles remain in remarkable agreement with the initial results, which shows that this boundary condition does not perturb the flow in the zone of interest.



## Annexe C

# Benchmark on Anisotropic Problems – A Galerkin Finite Element Solution

### C.1 Presentation of the scheme

In this paper, we address some of the benchmark problems posed for the *Finite Volumes for Complex Applications (FVCA 5)* conference. These tests are described in [68], and take all the form of an anisotropic diffusion problem :

$$\begin{cases} -\nabla \cdot (\mathbf{K} \nabla u) & \text{in } \Omega, \\ u = \bar{u} & \text{on } \Gamma_D, \\ \mathbf{K} \nabla u \cdot \mathbf{n} = g & \text{on } \Gamma_N, \end{cases} \quad (\text{C.1})$$

where  $u$  is the unknown,  $\Omega$  is the  $(0, 1) \times (0, 1)$  domain of  $\mathbb{R}^2$ , the boundary of which is split into  $\Gamma_D$  and  $\Gamma_N$ . The problem data are varied, but, in any case, regular :  $f \in \mathcal{L}^2(\Omega)$ ,  $\bar{u} \in \mathcal{C}^{0,1}(\Gamma_D)$ ,  $g \in \mathcal{L}^2(\Gamma_N)$ ,  $\mathbf{K} \in [\mathcal{L}^\infty(\Omega)]^{2 \times 2}$  and there exists  $C > 0$  such that for all  $\mathbf{w} \in \mathbb{R}^2$  :

$$\sum_{1 \leq i, j \leq 2} \mathbf{K}_{ij} \mathbf{w}_i \mathbf{w}_j \geq C \sum_{1 \leq i \leq 2} \mathbf{w}_i^2$$

Problem (C.1) can be formulated in a weak sense as follows :

Find  $u \in \mathcal{V}$  such that ,  $\forall v \in \mathcal{V}$ ,

$$\int_{\Omega} (\mathbf{K} \nabla u) \cdot \nabla v = \int_{\Omega} f v - \int_{\Omega} (\mathbf{K} \nabla \bar{u}) \cdot \nabla v + \int_{\Gamma_N} g v \quad (\text{C.2})$$

where  $\mathcal{V}$  is the subspace of  $\mathcal{H}^1(\Omega)$  of functions whose trace vanishes on  $\Gamma_D$ ,  $\bar{u}$  denotes a lifting of  $\bar{u}$  on  $\Omega$ , and the change of variables from  $u$  to  $u - \bar{u}$  has been performed without renaming the unknown function.

To solve (C.1), we use a Galerkin technique which consists in replacing in variational equation (C.2) the space  $\mathcal{V}$  for the solution and the test functions by a finite dimensional counterpart  $\mathcal{V}_h$ , and the function  $\bar{u} \in \mathcal{V}$  by an approximation  $\bar{u}_h \in \mathcal{V}_h$ .

When the proposed meshes are conforming, we use standard Lagrange finite elements (triangles or quadrilaterals) :

$$\mathcal{V}_h = \left\{ v_h \in \mathcal{V}; \forall K \in \mathcal{T}_h \left| \begin{array}{l} (v_h)|_K \in \mathbb{P}^k(K) \text{ on triangle } K \in \mathcal{T}_h \\ (v_h)|_K \in \mathbb{Q}^k(K) \text{ on quadrilateral } K \in \mathcal{T}_h \end{array} \right. \right\}$$

where  $\mathcal{T}_h$  is a meshing of the domain  $\Omega$ .

The series of meshes 3 proposed for Test 1.2 may be obtained from successive local refinement of an initial conforming mesh. In this case, the associated approximation space is built using the CHARMS (or Conforming, Hierarchical Adaptive Refinement MethodS) adaptation procedure [76]. The basic principle of CHARMS is to refine/unrefine primarily basis functions instead of cells. Consider an initial conforming mesh (possibly unstructured) together with its finite element structure. A hierarchy of nested grids is defined by successive divisions of cells into cells of the same type (uniformly applying the same subdivision pattern). A nested sequence of Lagrange finite element spaces follows, with the property that each basis function at a given level can be written as a linear combination of (few) basis functions belonging to the next finer level ; this generates a natural parent/child relationship between basis functions of two successive levels. Within a set of basis functions, the refinement (resp. unrefinement) of a parent is then defined by the addition (resp. removal) of all its childs whose node is not located at the same place. Cells are accordingly split and coarsened, leading to non-conforming grids, but their role is limited to be integration domains and support of the basis functions that, more importantly, span conforming finite element spaces. Each mesh of the series of meshes 3 corresponds to a grid obtained from three levels and the approximation is built on a finite element space  $\mathcal{V}_h$  that involves basis functions of those three levels.

Finally, we do not propose in this contribution any result for other nonconforming meshes, i.e. for Test 4 with Mesh 5 and for Test 6 with Mesh 7.

## C.2 Implementation

The computations presented in this contribution have been performed with PELICANS. This software development platform is a C++ application framework with a set of integrated reusable components, that has been designed at IRSN to simplify the task of developing applications for numerical mathematics and scientific computing, particularly those concerning PDEs and initial boundary value problems. PELICANS provides functionalities for dealing with multi-dimensional discretizations on structured and unstructured meshes with finite volume and finite element like methods. The following situations or techniques are accessible : moving boundaries, grid motion, characteristic methods, adaptative multilevel mesh refinement. Moreover, facilities are supplied for distributed processing and consistent coupling with external established numerical libraries. Nowadays PELICANS is the numerical kernel of several industrial codes developed at IRSN for nuclear safety studies.

The implementation of a finite element application with PELICANS only supposes that a set of finite element basis functions is given, regardless of its detailed structure or the way it was constructed ; so, as a striking by-product of the object programming methodology, the same code is used whatever the element may be, local refinement should be used or not [14].

PELICANS is distributed under the CeCILL-C license agreement (an adaptation of LGPL to the French law, see <http://www.cecill.info>). While being the property of IRSN, it is fully

free, open-source, and may be reused in softwares distributed under any license. The project page is <https://gforge.irsn.fr/gf/project/pelicans>.

### C.3 Numerical results

As we approximate the problem with a Galerkin finite element method, the computations for post-processing purpose of the fluxes at the domain boundaries, of the  $\mathcal{L}^2$ -norm of the error and of the  $\mathcal{H}^1$ -seminorm of the error have been made by a straightforward application of the definition of this quantities, considering the numerical solution  $u_h$  as a continuous function. For these computations, the numerical quadrature rule is of order 5.

- **Test 1.1 Mild anisotropy**,  $u(x, y) = 16x(1 - x)y(1 - y)$ ,  $\min = 0$ ,  $\max = 1$ , **regular triangular mesh, mesh1 P1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	21	109	-1.81E+00	3.44E-02	5.88E-01		
2	97	593	-1.12E+00	8.60E-03	2.95E-01	1.81	0.902
3	417	2737	-6.19E-01	2.14E-03	1.47E-01	1.91	0.955
4	1729	11729	-3.25E-01	5.34E-04	7.35E-02	1.95	0.975
5	7041	48529	-1.66E-01	1.33E-04	3.67E-02	1.98	0.989
6	28417	197393	-8.41E-02	3.33E-05	1.84E-02	1.99	0.990
7	114177	796177	-4.23E-02	8.32E-06	9.18E-03	1.99	1.00

**ocvl2**= 2.00, **ocvgradl2**= 1.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	1.13E-01	1.13E-01	1.13E-01	1.13E-01	0.00	1.03
2	7.01E-02	7.01E-02	7.01E-02	7.01E-02	0.00	1.01
3	3.87E-02	3.87E-02	3.87E-02	3.87E-02	0.00	1.00
4	2.03E-02	2.03E-02	2.03E-02	2.03E-02	0.00	1.00
5	1.04E-02	1.04E-02	1.04E-02	1.04E-02	0.00	1.00
6	5.26E-03	5.26E-03	5.26E-03	5.26E-03	0.00	1.00
7	2.64E-03	2.64E-03	2.64E-03	2.64E-03	0.00	1.00

**P2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	97	905	1.19E-02	1.54E-03	6.28E-02		
2	417	4337	-6.58E-03	1.80E-04	1.50E-02	2.94	1.96
3	1729	18929	-2.79E-03	2.22E-05	3.69E-03	2.94	1.97
4	7041	79025	-8.37E-04	2.76E-06	9.20E-04	2.97	1.98
5	28417	322865	-2.27E-04	3.45E-07	2.30E-04	2.98	1.99
6	114177	1305137	-5.88E-05	4.31E-08	5.74E-05	2.99	2.00
7	457729	5248049	-1.50E-05	5.38E-09	1.44E-05	3.00	1.99

**ocvl2**= 3.01, **ocvgradl2**= 2.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	7.46E-04	7.46E-04	7.46E-04	7.46E-04	0.00	1.00
2	4.11E-04	4.11E-04	4.11E-04	4.11E-04	0.00	1.00
3	1.74E-04	1.74E-04	1.74E-04	1.74E-04	0.00	1.00
4	5.23E-05	5.23E-05	5.23E-05	5.23E-05	0.00	1.00
5	1.42E-05	1.42E-05	1.42E-05	1.42E-05	0.00	1.00
6	3.67E-06	3.68E-06	3.67E-06	3.68E-06	0.00	1.00
7	9.36E-07	9.36E-07	9.36E-07	9.36E-07	0.00	1.000

- **Test 1.1 Mild anisotropy**,  $u(x, y) = 16x(1 - x)y(1 - y)$ ,  $\min = 0$ ,  $\max = 1$ , **distorted quadrangular meshes**, mesh4\_1\_i,  $i = 1..6$  **Q1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	256	2116	-1.13E+00	6.71E-02	7.60E-01		
2	1089	9409	-4.91E-01	2.40E-02	4.35E-01	1.42	0.770
3	2500	21904	-3.20E-01	1.19E-02	3.01E-01	1.69	0.890
4	4489	39601	-2.38E-01	7.06E-03	2.29E-01	1.78	0.933
5	7056	62500	-1.90E-01	4.64E-03	1.84E-01	1.86	0.967
6	10201	90601	-1.58E-01	3.27E-03	1.54E-01	1.90	0.965

ocvl2= 1.92, ocvgradl2= 0.976.

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
1	1.02E-01	1.28E-01	4.60E-02	7.21E-03	0.00	0.861
2	4.41E-02	5.13E-02	2.52E-02	2.23E-03	0.00	0.949
3	2.67E-02	3.00E-02	1.89E-02	4.46E-03	0.00	0.974
4	1.88E-02	2.07E-02	1.51E-02	4.90E-03	0.00	0.984
5	1.45E-02	1.56E-02	1.26E-02	4.76E-03	0.00	0.990
6	1.17E-02	1.25E-02	1.08E-02	4.47E-03	0.00	0.993

**Q2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	1089	16129	8.85E-03	3.51E-04	2.31E-02		
2	4489	69169	2.32E-03	3.55E-05	6.04E-03	3.24	1.89
3	10201	159201	1.05E-03	9.12E-06	2.71E-03	3.31	1.95
4	18225	286225	5.95E-04	3.61E-06	1.53E-03	3.19	1.97
5	28561	450241	3.83E-04	1.79E-06	9.78E-04	3.12	1.99
6	41209	651249	2.67E-04	1.00E-06	6.80E-04	3.18	1.98

ocvl2= 3.19, ocvgradl2= 1.99.

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
1	6.24E-05	2.31E-03	1.51E-05	5.06E-05	0.00	0.999
2	4.19E-06	5.87E-04	6.05E-07	3.72E-06	0.00	1.00
3	8.39E-07	2.63E-04	1.33E-07	4.55E-07	0.00	1.00
4	2.67E-07	1.49E-04	4.50E-08	1.16E-07	0.00	1.00
5	1.10E-07	9.58E-05	1.91E-08	3.71E-08	0.00	1.00
6	5.29E-08	6.67E-05	9.50E-09	1.62E-08	0.00	1.00

- **Test 1.1 Mild anisotropy**,  $u(x, y) = 16x(1 - x)y(1 - y)$ ,  $\min = 0$ ,  $\max = 1$ , **distorted quadrangular meshes**, mesh4\_2\_i,  $i = 1..6$  **Q1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	1024	8836	-4.95E-01	2.50E-02	4.41E-01		
2	4225	37249	-2.44E-01	7.59E-03	2.34E-01	1.68	0.896
3	9604	85264	-1.62E-01	3.56E-03	1.58E-01	1.84	0.956
4	17161	152881	-1.21E-01	2.05E-03	1.19E-01	1.90	0.976
5	26896	240100	-9.71E-02	1.32E-03	9.57E-02	1.96	0.970
6	38809	346921	-8.09E-02	9.26E-04	7.99E-02	1.93	0.984

ocvl2= 1.94, ocvgradl2= 0.99.

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
1	4.47E-02	5.23E-02	2.04E-02	6.36E-03	0.00	0.937
2	1.93E-02	2.14E-02	1.31E-02	7.16E-03	0.00	0.980
3	1.20E-02	1.30E-02	9.59E-03	5.97E-03	0.00	0.990
4	8.65E-03	9.20E-03	7.53E-03	4.97E-03	0.00	0.995
5	6.75E-03	7.11E-03	6.19E-03	4.22E-03	0.00	0.996
6	5.53E-03	5.78E-03	5.25E-03	3.65E-03	0.00	0.998

**Q2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	4225	65025	5.69E-04	4.05E-05	6.35E-03		
2	17161	269361	1.47E-04	4.06E-06	1.60E-03	3.28	1.97
3	38809	613089	6.57E-05	1.11E-06	7.15E-04	3.18	1.97
4	69169	1096209	3.71E-05	4.55E-07	4.02E-04	3.09	1.99
5	108241	1718721	2.38E-05	2.29E-07	2.58E-04	3.07	1.98
6	156025	2480625	1.66E-05	1.32E-07	1.79E-04	3.01	2.00

ocvl2= 3.02, ocvgradl2= 2.01.

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
1	4.62E-06	1.49E-04	8.26E-07	3.08E-06	0.00	1.00
2	2.96E-07	3.70E-05	4.33E-08	1.37E-07	0.00	1.00
3	5.86E-08	1.65E-05	8.90E-09	2.18E-08	0.00	1.00
4	1.86E-08	9.30E-06	2.91E-09	6.10E-09	0.00	1.00
5	7.62E-09	5.96E-06	1.21E-09	2.31E-09	0.00	1.00
6	3.67E-09	4.14E-06	6.00E-10	1.05E-09	0.00	1.00

- **Test 1.2 Mild anisotropy**,  $u(x, y) = \sin((1 - x)(1 - y)) + (1 - x)^3(1 - y)^2$ ,  $\min = 0$ ,  $\max = 1 + \sin 1$ , **regular triangular mesh**, mesh1 **P1 element**



i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	21	109	1.86E-01	9.35E-03	1.89E-01		
2	97	593	7.56E-02	2.28E-03	9.34E-02	1.84	0.920
3	417	2737	3.30E-02	5.67E-04	4.66E-02	1.91	0.953
4	1729	11729	1.52E-02	1.42E-04	2.33E-02	1.95	0.975
5	7041	48529	7.29E-03	3.54E-05	1.16E-02	1.98	0.993
6	28417	197393	3.56E-03	8.85E-06	5.82E-03	1.99	0.989
7	114177	796177	1.76E-03	2.21E-06	2.91E-03	2.00	0.997

ocvl2= 2.01, ocvgradl2= 1.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	5.51E-02	2.63E-02	2.81E-02	8.64E-02	0.00	1.84
2	2.61E-02	5.64E-03	1.73E-02	4.27E-02	0.00	1.84
3	1.26E-02	8.43E-04	9.38E-03	2.10E-02	0.00	1.84
4	6.16E-03	8.47E-05	4.85E-03	1.03E-02	0.00	1.84
5	3.05E-03	1.71E-04	2.46E-03	5.10E-03	0.00	1.84
6	1.51E-03	1.18E-04	1.24E-03	2.53E-03	0.00	1.84
7	7.55E-04	6.69E-05	6.21E-04	1.26E-03	0.00	1.84

**P2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	97	905	1.14E-03	2.21E-04	9.15E-03		
2	417	4337	-4.90E-04	2.68E-05	2.24E-03	2.89	1.93
3	1729	18929	-2.18E-04	3.32E-06	5.57E-04	2.94	1.96
4	7041	79025	-6.61E-05	4.14E-07	1.39E-04	2.97	1.98
5	28417	322865	-1.80E-05	5.17E-08	3.47E-05	2.98	1.99
6	114177	1305137	-4.68E-06	6.47E-09	8.68E-06	2.99	1.99
7	457729	5248049	-1.19E-06	8.08E-10	2.17E-06	3.00	2.00

ocvl2= 3.01, ocvgradl2= 2.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	1.95E-03	7.66E-03	8.45E-04	3.91E-03	0.00	1.84
2	3.67E-04	1.98E-03	2.82E-04	7.09E-04	0.00	1.84
3	7.60E-05	5.03E-04	7.85E-05	1.44E-04	0.00	1.84
4	1.70E-05	1.26E-04	2.06E-05	3.20E-05	0.00	1.84
5	4.00E-06	3.17E-05	5.26E-06	7.48E-06	0.00	1.84
6	9.68E-07	7.94E-06	1.33E-06	1.81E-06	0.00	1.84
7	2.38E-07	1.99E-06	3.34E-07	4.44E-07	0.00	1.84

- **Test 1.2 Mild anisotropy**,  $u(x, y) = \sin((1-x)(1-y)) + (1-x)^3(1-y)^2$ ,  $\min = 0$ ,  $\max = 1 + \sin 1$ , **locally refined non-conforming rectangular mesh, mesh3 Q1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	25	207	2.72E-01	5.06E-03	6.51E-02		
2	129	1547	1.32E-01	1.26E-03	3.20E-02	1.69	0.867
3	577	8035	6.47E-02	3.14E-04	1.58E-02	1.86	0.941
4	2433	36179	3.20E-02	7.84E-05	7.90E-03	1.93	0.963
5	9985	153139	1.59E-02	1.96E-05	3.94E-03	1.96	0.985

ocvl2= 2.00, ocvgradl2= 1.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	4.26E-02	4.36E-02	4.84E-03	1.35E-01	-3.50E-02	1.84
2	2.24E-02	1.14E-02	3.24E-03	6.54E-02	-1.03E-02	1.84
3	1.15E-02	2.91E-03	1.84E-03	3.20E-02	-2.81E-03	1.84
4	5.83E-03	7.35E-04	9.73E-04	1.58E-02	-7.37E-04	1.84
5	2.93E-03	1.85E-04	5.00E-04	7.87E-03	-1.88E-04	1.84

Q2 element

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	129	2257	-7.24E-03	1.31E-04	4.12E-03		
2	577	12865	-2.01E-03	1.62E-05	1.01E-03	2.79	1.88
3	2433	60449	-5.20E-04	2.02E-06	2.52E-04	2.89	1.93
4	9985	260833	-1.32E-04	2.52E-07	6.28E-05	2.95	1.97
5	40449	1082465	-3.32E-05	3.14E-08	1.57E-05	2.98	1.98

ocvl2= 3.00, ocvgradl2= 2.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	1.31E-03	1.74E-02	9.52E-05	2.84E-03	-6.56E-04	1.84
2	3.04E-04	4.46E-03	2.84E-05	6.55E-04	-8.26E-05	1.84
3	7.36E-05	1.13E-03	7.37E-06	1.62E-04	-1.05E-05	1.84
4	1.81E-05	2.83E-04	1.85E-06	4.05E-05	-1.37E-06	1.84
5	4.49E-06	7.10E-05	4.61E-07	1.01E-05	-1.75E-07	1.84

- **Test 2 Numerical locking**,  $u(x, y) = \sin(2\pi x)e^{-2\pi\sqrt{\frac{1}{3}}y}$ ,  $\delta = 10^5$ ,  $\min = -1$ ,  $\max = 1$ , regular triangular mesh, mesh1 **P1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	37	221	3.22E+00	7.00E-01	4.39E+00		
2	129	833	9.37E-01	1.49E-01	1.91E+00	2.48	1.33
3	481	3233	5.19E-02	3.89E-02	9.87E-01	2.04	1.01
4	1857	12737	-1.73E-03	9.82E-03	4.97E-01	2.04	1.02
5	7297	50561	-1.65E-03	2.46E-03	2.49E-01	2.02	1.01
6	28929	201473	-6.35E-04	6.16E-04	1.25E-01	2.01	1.00
7	115201	804353	-3.24E-04	1.54E-04	6.23E-02	2.01	1.01

ocvl2= 2.00, ocvgradl2= 1.01.

i	erflx0	erflx1	fluy0	fluy1	umin	umax
1	1.00E-00	1.00E-00	1.37E+00	1.85E+00	-9.48E-03	9.75E-03
2	3.64E-01	3.63E-01	4.18E-01	5.25E-01	-1.00	1.00
3	1.00E-01	9.93E-02	2.08E-02	3.54E-02	-1.00	1.00
4	2.57E-02	2.53E-02	-6.20E-04	1.25E-03	-1.00	1.00
5	6.51E-03	6.31E-03	-2.40E-04	-2.02E-04	-1.00	1.00
6	1.65E-03	1.56E-03	-6.38E-05	3.53E-05	-1.00	1.00
7	4.26E-04	3.77E-04	-1.63E-06	-1.73E-05	-1.00	1.00

**P2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	129	1353	-1.69E-01	2.79E-02	5.29E-01		
2	481	5281	-1.14E-01	1.49E-02	3.88E-01	0.95	0.47
3	1857	20865	-3.02E-02	1.92E-03	9.96E-02	3.03	2.01
4	7297	82945	-4.07E-03	2.42E-04	2.51E-02	3.03	2.01
5	28929	330753	-3.00E-04	3.02E-05	6.28E-03	3.02	2.01
6	115201	1320961	-1.86E-05	3.78E-06	1.57E-03	3.01	2.01
7	459777	5279745	-1.17E-06	4.83E-07	3.93E-04	2.97	2.00

ocvl2= 2.97, ocvgradl2= 2.00.

i	erflx0	erflx1	fluy0	fluy1	umin	umax
1	2.15E-01	2.17E-01	-5.59E-02	-1.05E-01	-0.956	0.956
2	1.58E-01	1.58E-01	-4.16E-02	-7.07E-02	-1.00	1.00
3	4.80E-02	4.81E-02	-1.27E-02	-1.72E-02	-1.00	1.00
4	1.26E-02	1.26E-02	-1.72E-03	-2.32E-03	-1.00	1.00
5	3.18E-03	3.18E-03	-1.13E-04	-1.83E-04	-1.00	1.00
6	7.98E-04	7.98E-04	-8.00E-06	-1.00E-05	-1.00	1.00
7	2.00E-04	2.00E-04	-1.70E-09	-1.25E-06	-1.00	1.00

- **Test 2 Numerical locking**,  $u(x, y) = \sin(2\pi x)e^{-2\pi\sqrt{\frac{1}{8}}y}$ ,  $\delta = 10^6$ ,  $\min = -1$ ,  $\max = 1$ , regular triangular mesh, mesh1 **P1 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	37	221	1.03E+01	7.05E-01	4.42E+00		
2	129	833	3.01E+00	1.50E-01	1.92E+00	2.48	1.33
3	481	3233	1.80E-01	3.92E-02	9.94E-01	2.04	1.00
4	1857	12737	1.86E-03	9.89E-03	5.01E-01	2.04	1.01
5	7297	50561	-1.55E-03	2.48E-03	2.51E-01	2.02	1.01
6	28929	201473	-5.17E-04	6.20E-04	1.25E-01	2.01	1.01
7	115201	804353	-1.21E-04	1.55E-04	6.28E-02	2.01	1.00

ocvl2= 2.00, ocvgradl2= 0.995.

i	erflx0	erflx1	fluy0	fluy1	umin	umax
1	1.00E-00	1.00E+00	4.91E+00	5.39E+00	-2.93E-03	3.02E-03
2	3.64E-01	3.63E-01	1.45E+00	1.56E+00	-1.00	1.00
3	9.98E-02	9.96E-02	8.33E-02	9.80E-02	-1.00	1.00
4	2.56E-02	2.54E-02	3.64E-04	2.24E-03	-1.00	1.00
5	6.44E-03	6.38E-03	-3.84E-04	-7.81E-04	-1.00	1.00
6	1.62E-03	1.59E-03	-2.87E-04	-3.66E-05	-1.00	1.00
7	4.09E-04	3.94E-04	1.65E-05	-4.08E-05	-1.00	1.00

**P2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	129	1353	-5.16E-01	2.81E-02	5.33E-01		
2	481	5281	-3.59E-01	1.51E-02	3.93E-01	0.94	0.46
3	1857	20865	-9.59E-02	1.94E-03	1.01E-01	3.04	2.01
4	7297	82945	-1.31E-02	2.45E-04	2.54E-02	3.02	2.02
5	28929	330753	-9.72E-04	3.09E-05	6.36E-03	3.01	2.01
6	115201	1320961	-6.13E-05	3.83E-06	1.59E-03	3.02	2.01
7	459777	5279745	2.80E-06	4.92E-07	3.97E-04	2.96	2.00

**ocvl2**= 2.96, **ocvgradl2**= 2.00.

i	erflx0	erflx1	fluy0	fluy1	umin	umax
1	2.16E-01	2.16E-01	-2.32E-01	-2.81E-01	-0.95	0.95
2	1.58E-01	1.58E-01	-1.64E-01	-1.94E-01	-1.00	1.00
3	4.83E-02	4.83E-02	-4.56E-02	-5.02E-02	-1.00	1.00
4	1.26E-02	1.26E-02	-6.22E-03	-6.83E-03	-1.00	1.00
5	3.20E-03	3.20E-03	-4.57E-04	-5.13E-04	-1.00	1.00
6	8.02E-04	8.02E-04	-3.71E-05	-2.40E-05	-1.00	1.00
7	2.01E-04	2.00E-04	-8.54E-07	2.96E-07	-1.00	1.00

- **Test 3 Oblique flow**,  $\min = 0$ ,  $\max = 1$ , **uniform rectangular mesh**, mesh2 The reference mesh is composed of  $500 \times 500$  squares.

**Q1 element**

i	nunkw	nnmat	sumflux	umin	umax
1	9	49	-7.29E-16	0.00	1.00
2	49	361	-1.25E-15	0.00	1.00
3	225	1849	3.79E-16	0.00	1.00
4	961	8281	2.37E-15	0.00	1.00
5	3969	34969	3.55E-14	0.00	1.00
6	16129	143641	-1.97E-07	0.00	1.00
7	65025	582169	-9.16E-08	0.00	1.00
ref	249001	2235025	-2.18E-13	0.00	1.00

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1	-0.305	0.305	-0.250	0.250	0.221	0.349	0.367
2	-0.196	0.196	-0.125	0.125	0.267	0.255	0.0449
3	-0.198	0.198	-0.111	0.111	0.300	0.310	0.0323
4	-0.197	0.197	-0.104	0.104	0.270	0.300	0.100
5	-0.196	0.196	-0.101	0.101	0.255	0.280	0.0893
6	-0.195	0.195	-0.0998	0.0998	0.247	0.262	0.0573
7	-0.194	0.194	-0.0993	0.0993	0.244	0.252	0.0317
ref	-0.193	0.193	-0.0990	0.0990	0.243	0.247	0.0162

**Q2 element**

i	nunkw	nnmat	sumflux	umin	umax
1	49	529	-4.51E-16	0.00	1.00
2	225	3025	-1.86E-15	0.00	1.00
3	961	14161	-2.52E-15	0.00	1.00
4	3969	61009	-1.90E-14	0.00	1.00
5	16129	253009	-7.52E-14	0.00	1.00
6	65025	1030225	2.84E-08	0.00	1.00
7	261121	4157521	1.59E-08	0.00	1.00
ref	998001	15928081	-4.45E-12	0.00	1.00

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1	-0.127	0.127	-0.0360	0.0360	0.264	0.174	0.341
2	-0.193	0.193	-0.0940	0.0940	0.281	0.319	0.119
3	-0.192	0.192	-0.0961	0.0961	0.248	0.260	0.0462
4	-0.192	0.192	-0.0975	0.0975	0.245	0.248	0.0121
5	-0.193	0.193	-0.0981	0.0981	0.243	0.243	0.00
6	-0.193	0.193	-0.0984	0.0984	0.242	0.242	0.00
7	-0.193	0.193	-0.0986	0.0986	0.242	0.242	0.00
ref	-0.193	0.193	-0.0986	0.0986	0.242	0.242	0.00

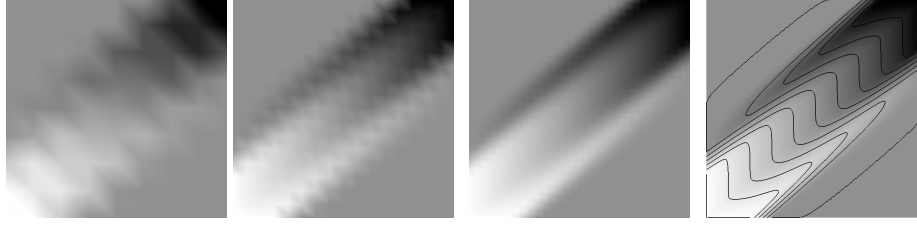


FIG. C.1 – Solutions for the oblique flow on mesh2<sub>i</sub> for  $i=2,3,4$ , and on the reference  $500 \times 500$  mesh obtained with the Q2 element

- **Test 4 Vertical fault**,  $\min = 0$ ,  $\max = 1$ , **non-conforming rectangular mesh**, mesh5  
We choose as a reference mesh the provided one, which is a conforming regular mesh of  $320 \times 320$  cells and has a step size  $h = 4.42 \cdot 10^{-3}$ .

#### Q1 element

i	nunkw	nnmat	sumflux	umin	umax
1					
reg	361	3025	2.52E+00	0.00	1.00
ref	101761	912025	1.56E-01	0.00	1.00

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1							
reg	-43.3	45.4	0.433	8.26E-04	44.5	43.2	2.92E-02
ref	-42.2	44.5	-2.16	7.96E-04	43.3	43.2	2.31E-03

#### Q2 element

i	nunkw	nnmat	sumflux	umin	umax
1					
reg	1521	22801	-7.07E-03	0.00	1.00
ref	408321	6507601	-2.60E-05	0.00	1.00

i	flux0	flux1	fluy0	fluy1	ener1	ener2	eren
1							
reg	-42.4	44.7	-2.28	7.84E-04	43.5	43.4	2.30E-03
ref	-42.1	44.5	-2.32	7.96E-04	43.2	43.2	0.00E+00

- **Test 5 Heterogeneous rotating anisotropy**,  $\min = 0$ ,  $\max = 1$ , **uniform rectangular mesh**, mesh2 **Q1 element**

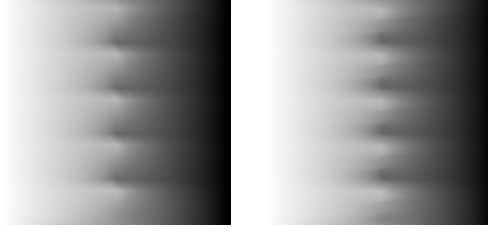


FIG. C.2 – Solution for the vertical fault on conforming rectangular meshes (left : coarse  $20 \times 20$  mesh, right : reference  $320 \times 320$  mesh) obtained with the Q2 element

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	9	49	-5.06E-01	3.07E-02	5.04E-01		
2	49	361	-1.28E-01	7.89E-03	2.52E-01	1.60	0.819
3	225	1849	-3.20E-02	2.00E-03	1.26E-01	1.80	0.909
4	961	8281	-7.97E-03	5.04E-04	6.30E-02	1.90	0.955
5	3969	34969	-1.99E-03	1.26E-04	3.15E-02	1.95	0.977
6	16129	143641	-4.97E-04	3.16E-05	1.57E-02	1.97	0.993
7	65025	582169	-1.24E-04	7.91E-06	7.87E-03	1.99	0.991

**ocvl2**= 2.01, **ocvgradl2**= 1.00.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	1.04E-01	8.46E-02	1.04E-01	8.46E-02	0.00	1.05
2	2.73E-02	1.64E-02	2.73E-02	1.64E-02	0.00	1.01
3	6.97E-03	3.36E-03	6.97E-03	3.36E-03	0.00	1.00
4	1.76E-03	7.46E-04	1.76E-03	7.46E-04	0.00	1.00
5	4.41E-04	1.75E-04	4.41E-04	1.75E-04	0.00	1.00
6	1.10E-04	4.25E-05	1.10E-04	4.25E-05	0.00	1.00
7	2.76E-05	1.05E-05	2.76E-05	1.05E-05	0.00	1.00

**Q2 element**

i	nunkw	nnmat	sumflux	erl2	ergrad	rl2	rgrad
1	49	529	2.98E-01	4.45E-03	1.04E-01		
2	225	3025	7.89E-02	5.58E-04	2.65E-02	2.72	1.79
3	961	14161	2.00E-02	7.00E-05	6.63E-03	2.86	1.91
4	3969	61009	5.03E-03	8.78E-06	1.66E-03	2.93	1.95
5	16129	253009	1.26E-03	1.10E-06	4.15E-04	2.96	1.98
6	65025	1030225	3.15E-04	1.37E-07	1.04E-04	2.99	1.99
7	261121	4157521	7.86E-05	1.72E-08	2.60E-05	2.99	1.99

**ocvl2**= 3.02, **ocvgradl2**= 2.02.

i	erflx0	erflx1	erfly0	erfly1	umin	umax
1	6.70E-02	5.59E-02	6.89E-02	5.51E-02	0.00	1.00
2	1.80E-02	1.35E-02	1.83E-02	1.35E-02	0.00	1.00
3	4.60E-03	3.30E-03	4.68E-03	3.30E-03	0.00	1.00
4	1.16E-03	8.15E-04	1.18E-03	8.15E-04	0.00	1.00
5	2.90E-04	2.02E-04	2.95E-04	2.02E-04	0.00	1.00
6	7.25E-05	5.04E-05	7.38E-05	5.04E-05	0.00	1.00
7	1.81E-05	1.26E-05	1.84E-05	1.26E-05	0.00	1.00

- **Test 6 Oblique drain**,  $\min = -1.2$ ,  $\max = 0$ , **coarse (C) and fine (F) oblique meshes, mesh6 and mesh7** As the solution is linear (and thus belongs to the approximation space), the finite element method is exact. We thus only give the results for the coarse mesh.

**Q1 element**

grid	nunkw	nmat	sumflux	erl2	ergrad
C	180	1450	3.31E-14	3.28E-16	6.52E-15

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
C	1.34E-15	2.99E-15	1.39E-15	1.53E-15	-1.20E+00	-0.00E+00

**Q2 element**

grid	nunkw	nmat	sumflux	erl2	ergrad
C	779	11289	-9.65E-14	4.54E-16	1.12E-14

grid	erflx0	erflx1	erfly0	erfly1	umin	umax
C	1.19E-15	1.73E-14	1.39E-16	7.22E-15	-1.20E+00	-0.00E+00

- **Test 7 Oblique barrier**,  $\min = -5.575$ ,  $\max = 0.575$ , **coarse oblique mesh mesh6** As the solution is linear (and thus belongs to the approximation space), the finite element method is exact.

**Q1 element**

nunkw	nmat	sumflux	erl2	ergrad
180	1450	-6.48E-14	3.79E-15	6.49E-14

erflx0	erflx1	erfly0	erfly1	umin	umax
5.84E-14	7.98E-14	4.44E-16	3.62E-14	-5.575	0.575

**Q2 element**

nunkw	nmat	sumflux	erl2	ergrad
779	11289	5.75E-15	6.12E-15	8.19E-14

erflx0	erflx1	erfly0	erfly1	umin	umax
5.59E-14	7.85E-14	2.00E-15	3.00E-14	-5.575	0.575



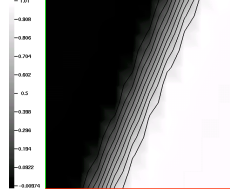


FIG. C.3 – Solution for the anisotropy with wells (Test 9) obtained with the Q2 element

- **Test 8 Perturbed parallelograms, min = 0, perturbed parallelogram mesh mesh8**  
The condition  $\int_{cell(6,6)} f(x) dx = 1$  is satisfied using for right hand side term  $f(x) = 1 / |cell(6,6)|$  if  $x \in cell(6,6)$ ,  $f = 0$  elsewhere.

**Q1 element**

nunkw	nmat	sumflux	umin	umax	
100	784	3.53E-02	-4.17E-03	4.90E-02	
		flux0	flux1	fluy0	fluy1
		5.51E-06	7.15E-05	5.46E-01	4.89E-01

**Q2 element**

nunkw	nmat	sumflux	umin	umax	
441	6241	-8.43E-04	-5.07E-03	8.04E-02	
		flux0	flux1	fluy0	fluy1
		-5.52E-05	1.96E-05	4.98E-01	5.01E-01

- **Test 9 Anisotropy with wells, min = 0, max = 1., square uniform grid mesh9**  
The conditions  $u(x) = 0$ , for  $x \in cell(4,6)$ , and  $u(x) = 1$ , for  $x \in cell(8,6)$ , are satisfied by imposing all degrees of freedom associated with nodes that lie in cells (4,6) and (8,6) to 0 and 1 respectively.

**Q1 element**

nunkw	nmat	sumflux	umin	umax
136	1044	-2.64E-10	-2.36E-02	1.02E+00

**Q2 element**

nunkw	nmat	sumflux	umin	umax
511	7407	-2.72E-11	-5.94E-03	1.01E+00

**C.4 Comments on the results**

On both triangles and quadrilaterals, with a finite element of degree  $k$ , rates of convergence of order  $k$  in  $\mathcal{H}^1$ -norm and of order  $(k + 1)$  in  $\mathcal{L}^2$ -norm are expected, i.e. if  $u \in \mathcal{H}^{k+1}$  :

$$\|u - u_h\|_1 \leq c_1 h^k \|u\|_{k+1}, \quad \|u - u_h\|_0 \leq c_0 h^{k+1} \|u\|_{k+1}, \quad (\text{C.3})$$

where  $\|\cdot\|_1$  and  $\|\cdot\|_0$  denote the  $\mathcal{H}^1$ -norm and  $\mathcal{L}^2$ -norm respectively and  $c_1$  and  $c_0$  two positive constants. For the tests that admit analytical solutions, these theoretical convergence orders are recovered in any case.

The maximum principle is in general not fulfilled, but, since the solutions are smooth enough ( $u \in \mathcal{W}^{k+1,\infty}$ ), a convergence of order  $k + 1$  is likely to be observed also in  $\mathcal{L}^\infty$ -norm, which explains that possible overshoots or undershoots quickly decrease when refining the mesh.



# Bibliographie

- [1] L. Agelas and D. A. Di Pietro. Benchmark on Anisotropic Problems – a symmetric finite volume scheme for anisotropic heterogeneous second-order elliptic problems. In *FVCA5 – Finite Volumes for Complex Applications V*, pages 705–716. Wiley, 2008.
- [2] P. Angot, V. Dolejší, M. Feistauer, and J. Felcman. Analysis of a combined barycentric finite volume–nonconforming finite element method for nonlinear convection-diffusion problems. *Application of Mathematics*, 43 :263–310, 1998.
- [3] G. Ansanay-Alex. Eléments finis non conformes de bas degré pour les écoulements stationnaires de fluides incompressibles. Master’s thesis, EGIM/U. Provence, 2005.
- [4] G. Ansanay-Alex, F. Babik, L. Gastaldo, A. Larcher, C. Lapuerta, J.-C. Latché, and D. Vola. A finite volume stability result for the convection operator in compressible flows ...and some finite element applications. In *Finite Volumes for Complex Applications V - Problems and Perspectives - Aussois, France*, pages 185–192, 2008.
- [5] ANSYS. *GAMBIT*. <http://www.fluent.com/software/gambit/index.htm>.
- [6] B.F. Armaly, F. Durst, J.C.F. Pereira, and B. Schönung. Experimental and theoretical investigation of backward-facing step flow. *Journal of Fluid Mechanics*, 127 :473–496, 1983.
- [7] F. Babik, T. Gallouët, J.-C. Latché, S. Suard, and D. Vola. On two fractional step finite volume and finite element schemes for reactive low Mach number flows. In *The International Symposium on Finite Volumes for Complex Applications IV - Problems and Perspectives - Marrakech, July 4-8*, 2005.
- [8] F. Babik, J.-C. Latché, and D. Vola. An  $L^2$ -stable approximation of the Navier-Stokes advective operator for non conforming finite elements. In *Mini-Workshop on Variational Multiscale Methods and Stabilized Finite Elements, Lausanne*, 2007.
- [9] F. Babik, J.C. Latché, and D. Vola. An  $l_2$ -stable approximation of the navier-stokes advective operator for nonconforming finite elements. Mini-Workshop on Variational Multiscale Methods and Stabilized Finite Elements, feb 2007.
- [10] F. Badrot-Nico, F. Brissaud, and V. Guinot. A finite volume upwind scheme for the solution of the linear advection-diffusion equation with sharp gradients in multiple dimensions. *Advances in Water Resources*, 30(9) :2002–2025, 2007.
- [11] T. J. Barth and P. O. Frederickson. Higher order solution of the Euler equations on unstructured grids using quadratic reconstruction. Technical Report AIAA 90-0013, American Institute for Aeronautics and Astronautics, 1990.
- [12] T. J. Barth and D. C. Jespersen. The design and application of upwind schemes on unstructured meshes. Technical Report AIAA 89-0366, American Institute for Aeronautics and Astronautics, 1989.
- [13] F. Boyer and P. Fabrie. *Eléments d’analyse pour l’étude de quelques modèles d’écoulements*

- de fluides visqueux incompressibles*. Springer-Verlag, 2006.
- [14] F. Boyer, C. Lapuerta, S. Minjeaud, and B. Piar. Hierarchical refinement for the finite element solution of Cahn-Hilliard/Navier-Stokes equations. In *Trends in Numerical and Physical Modeling for Industrial Multiphase Flows*, Cargèse, September 2007.
- [15] T. Buffard and S. Clain. Monoslope and multislope MUSCL methods for unstructured meshes. preprint du Laboratoire de Mathématiques, Université Blaise Pascal, 2008.
- [16] E. Burman and A. Ern. Stabilized galerkin approximation of convection-diffusion-reaction equations : discrete maximum principle and convergence. *Mathematics of Computation*, 74 :1637–1652, 2005.
- [17] E. Burman and A. Ern. Stabilized galerkin approximation of convection-diffusion-reaction equations : discrete maximum principle and convergence. *Mathematics of Computation*, 74 :1637–1652, 2005.
- [18] E. Burman, M. A. Fernandez, and P. Hansbo. Edge stabilization for the incompressible Navier-Stokes equations : a continuous interior penalty finite element method. Technical report, EPFL/IACS, 2004.
- [19] C. Calgaro, E. Creusé, and T. Goudon. An hybrid finite volume-finite element method for variable density incompressible flows. *J. Comput. Phys.*, 227(9) :4671–4696, 2008.
- [20] C. Chainais-Hillairet. *Schémas volumes finis pour des problèmes hyperboliques : convergence et estimations d'erreur*. PhD thesis, Université Paris 6, 1998.
- [21] E. Chénier, R. Eymard, and Herbin R. A collocated finite volume scheme to solve free convection for general non-conforming grids. *J. Comput. Phys.*, to appear, see also <http://hal.archives-ouvertes.fr>.
- [22] T.P. Chiang, T. W.H. Sheu, and C.C. Fang. Numerical investigation of vortical evolution in a backward-facing step expansion flow. *Applied Mathematical Modelling*, 23(12) :915–932, dec 1999.
- [23] A.J. Chorin. Numerical solution of the Navier–Stokes equations. *Mathematics of Computation*, 22 :745–762, 1968.
- [24] P. G. Ciarlet. Handbook of numerical analysis volume II : Finite elements methods – Basic error estimates for elliptic problems. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume II*, pages 17–351. North Holland, 1991.
- [25] S. Clain. The multi-slope MUSCL method. In R. Eymard and J.-M. Hérard, editors, *Finite volumes for complex applications V*, problems & perspectives, pages 297–304. Wiley, 2008.
- [26] S. Clain and V. Clauzon. The  $l^\infty$  stability of the muscl methods. submitted.
- [27] V. Clauzon. *Analyse de schémas d'ordre élevé pour les écoulements compressibles. Application à la simulation numérique d'une torche*. PhD thesis, Université Blaise Pascal, 2008.
- [28] R. Courant, E. Isaacson, and M. Rees. On the solution of nonlinear hyperbolic differential equations by finite differences. *Communications on Pure and Applied Mathematics*, 5(3) :243–255, 1952.
- [29] P. H. Cournède, C. Debiez, and A. Dervieux. A positive MUSCL scheme for triangulations. Technical Report 3465, Institut National de Recherche en Informatique et Automatique, 1998.
- [30] M. Crouzeix and P.A. Raviart. Conforming and non-conforming finite element methods for solving the stationary Stokes equations. *RAIRO Anal. Numer.*, 7 :33–76, 1973.
- [31] V. Dolejší, M. Feistauer, J. Felcman, and A. Kliková. Error estimates for barycentric finite volumes combined with nonconforming finite elements applied to nonlinear convection-

- diffusion problems. *Application of Mathematics*, 47 :301–340, 2002.
- [32] J. Douglas and T. Dupont. Interior penalty procedures for elliptic and parabolic Galerkin methods. *Lecture Notes in Physics*, 58, 1976.
- [33] Burman E. and Hansbo P. A stabilized non-conforming finite element method for incompressible flow. *Computer Methods in Applied Mechanics and Engineering*, 195(23–24) :2881–2899, 2005.
- [34] A. Ern and J.-L. Guermond. Theory and practice of finite elements. Number 159 in Applied Mathematical Sciences. Springer, New York, 2004.
- [35] R. Eymard, T. Gallouët, and R. Herbin. The finite volume method. In *Handbook for Numerical Analysis*. Ciarlet, P. and Lions, J.L., 2000. Version updated in 2003.
- [36] R. Eymard, T. Gallouët, and R. Herbin. Discretization schemes for heterogeneous and anisotropic diffusion problems on general nonconforming meshes. jan 2008.
- [37] R. Eymard, T. Gallouët, and R. Herbin. Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids. to appear in J. Numer. Math., preprint available at <http://arxiv.org/abs/0801.1430>, jan 2009.
- [38] R. Eymard, T. Gallouët, and R. Herbin. Discretization schemes for heterogeneous and anisotropic diffusion problems on general nonconforming meshes. to appear in IMAJNA, preprint available at <http://arxiv.org/abs/0801.1430>, jan 2009.
- [39] R. Eymard, T. Gallouët, and R. Herbin. A cell-centred finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. *IMA Journal of Numerical Analysis*, 26 :326–353, 2005.
- [40] R. Eymard, T. Gallouët, and R. Herbin. Benchmark on Anisotropic Problems – SUSHI : a scheme using stabilization and hybrid interfaces for anisotropic heterogeneous diffusion problems. In *FVCA5 – Finite Volumes for Complex Applications V*, pages 801–814. Wiley, 2008.
- [41] R. Eymard and R. Herbin. A new collocated finite volume scheme for the incompressible Navier-Stokes equations on general non matching grids. *C. R. Math. Acad. Sci. Paris*, 344(10) :659–662, 2007.
- [42] R. Eymard, D. Hilhorst, and M. Vohralík. A combined finite volume–nonconforming/mixed-hybrid finite element scheme for degenerate parabolic problems. *Numerische Mathematik*, 105 :73–131, 2006.
- [43] C. Farre, C.D. Perez-Segarra, M. Soria, and A. Oliva. Analysis of different numerical schemes for the resolution of convection-diffusion equations using finite-volume methods on three-dimensional unstructured grids. Part II : Numerical analysis. *Numerical Heat Transfer : Part B : Fundamentals*, 49(4) :351–375, 2006.
- [44] J.H. Ferziger and M. Perić. *Computational methods for fluid dynamics*. Springer-Verlag, 2002.
- [45] Y. Fraigneau, J.-L. Guermond, and L. Quartapelle. Approximation of variable density incompressible flows by means of finite elements and finite volumes. *Communications in Numerical Methods in Engineering*, 17 :893–902, 2001.
- [46] T. Gallouët, L. Gastaldo, R. Herbin, and J.-C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. Part I : the isothermal case. *to appear in Mathematics of Computation*, 2008.
- [47] T. Gallouët, L. Gastaldo, R. Herbin, and J.-C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. Part II : the isentropic case. *submitted*, 2008.

- [48] T. Gallouët, L. Gastaldo, R. Herbin, and J.-C. Latché. An unconditionally stable pressure correction scheme for compressible barotropic Navier-Stokes equations. *Mathematical Modelling and Numerical Analysis*, 42 :303–331, 2008.
- [49] L. Gastaldo, F. Babik, R. Herbin, and J.-C. Latché. An unconditionally stable pressure correction scheme for barotropic compressible Navier-Stokes equations. In *European Conference on Computational Fluid Dynamics*, 2006.
- [50] L. Gastaldo, R. Herbin, and J.-C. Latché. A discretization of phase mass balance in fractional step algorithms for the drift-flux model. *to appear in IMA Journal of Numerical Analysis*, 2009.
- [51] L. Gastaldo, R. Herbin, and J.-C. Latché. An entropy preserving finite element-finite volume scheme for the drift-flux model. *submitted*, 2009.
- [52] C. Geuzaine and J. C. Remacle. *Gmsh Reference Manual*. <http://www.geuz.org/gmsh>, 2.2 edition, 2008.
- [53] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, 2001.
- [54] K. Goda. A multistep technique with implicit difference schemes for calculating two- or three-dimensional cavity flows. *Journal of Computational Physics*, 30 :76–95, 1979.
- [55] S. K. Godunov. Difference methods for the numerical calculation of the equations of fluid dynamics. *Math. Sbornik*, 47 :271–306, 1959.
- [56] J. D. Goodman and R. J. Leveque. On the accuracy of stable schemes for 2d conservation laws. *Math. Comp.*, 45(171) :15–21, 1985.
- [57] P.M. Gresho and R.L. Lee. Don’t suppress the wiggles – they’re telling you something. In *Finite element methods for convection dominated flows*, New York, 1979.
- [58] J.-L. Guermond. Un résultat de convergence d’ordre deux en temps pour l’approximation des équations de Navier-Stokes par une technique de projection incrémentale. *Mathematical Modelling and Numerical Analysis*, 33(1) :169–189, 1999.
- [59] J.-L. Guermond and A. Salgado. A splitting method for incompressible flows with variable density based on a pressure poisson equation. *Journal of Computational Physics*, 228 :2834–2846, 2009.
- [60] F.H. Harlow and J.E. Welsh. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Physics of Fluids*, 8 :2182–2189, 1965.
- [61] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49 :357–393, 1983.
- [62] A. Harten. ENO schemes with subcell resolution. *J. Comp. Phys.*, 83 :148–184, 1989.
- [63] A. Harten and S. Chakravarthy. Multi-dimensional ENO schemes for general geometries. Technical Report ICASE-91-76, Institute for Computer Applications in Science and Engineering, 2004.
- [64] A. Harten, J. M. Hyman, and P. D. Lax. On finite difference approximations and entropy conditions. *Communication Pure Applied Mathematics*, 29 :297–322, 1976.
- [65] A. Harten, S. Osher, B. Engquist, and S. Chakravarthy. Some results on uniformly high order accurate essentially non-oscillatory schemes. *Appl. Num. Math.*, 2 :347–377, 1986.
- [66] A. Harten, S. Osher, B. Engquist, and S. Chakravarthy. Uniformly high-order accurate essentially nonoscillatory schemes III. *J. Comp. Phys.*, 71(2) :231–303, 1987.
- [67] F. Hecht and E. Saltel. *EMC2*. <http://pauillac.inria.fr/cdrom/www/emc2/fra.htm>.
- [68] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In *FVCA5 – Finite Volumes for Complex Applications V*, pages

- 659–692. Wiley, 2008.
- [69] M.E. Hubbard. Multidimensional slope limiters for muscl-type finite volume schemes on unstructured grids. *Journal of Computational Physics*, 155 :54–74, 1999.
- [70] ISIS. Computational Fluid Dynamics software. <https://gforge.irsn.fr/gf/project/isis>.
- [71] A. Jameson. Positive schemes and shock modelling for compressible flows. *International Journal for Numerical Methods in Fluids*, 20 :743–776, 1995.
- [72] A. Jameson and P. D. Lax. Conditions for the construction of multipoint variation diminishing difference schemes. *Appl. Numer. Math.*, 2(3–5) :335–345, 1986.
- [73] V. John, S. Kaya, and W. Layton. A two-level variational multiscale method for convection-dominated convection-diffusion equations. *Computer Methods in Applied Mechanics and Engineering*, 195(33-36) :4594–4603, 2006.
- [74] G. Karypis and V. Kumar. *METIS*. [http://people.scs.fsu.edu/~burkardt/c\\_src/metis/metis.html](http://people.scs.fsu.edu/~burkardt/c_src/metis/metis.html).
- [75] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1) :359–392, 1998.
- [76] P. Krysl, A. Trivedi, and B. Zhu. Object-oriented hierarchical mesh refinement with CHARMS. *International Journal for Numerical Methods in Engineering*, 60 :1401–1424, 2004.
- [77] D. Kuzmin and M. Möler. *Flux-corrected transport. Principles, Algorithms and Applications.*, chapter Algebraic Flux Correction I. Scalar conservation laws, pages 155–203. Springer-Verlag, 2005.
- [78] D. Kuzmin and S. Turek. *Flux-corrected transport. Principles, Algorithms and Applications.*, chapter Algebraic Flux Correction III. Incompressible Flow Problems. Springer-Verlag, 2005.
- [79] B. Larrouturou. How to preserve the mass fractions positivity when computing compressible multi-component flows. *Journal of Computational Physics*, 95 :59–84, 1991.
- [80] P.D. Lax. *Hyperbolic systems of conservation laws*. SIAM Pub., Philadelphia, 1973.
- [81] J.-L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications*. Springer-Verlag, 1968.
- [82] A. Majda and J. Sethian. The derivation and numerical solution of the equations for zero Mach number solution. *Combustion Science and Techniques*, 42 :185–205, 1985.
- [83] G. Manzini and A. Russo. A finite volume method for advection-diffusion problems in convection-dominated regimes. *Computer Methods in Applied Mechanics and Engineering*, 197(13-16) :1242–1261, 2008.
- [84] M. Marion and R. Temam. Navier-Stokes equations : Theory and approximation. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume VI*. North Holland, 1998.
- [85] R. E. Mitchell, A. F. Sarofim, and L. A. Clomburg. Experimental and numerical investigation of confined laminar diffusion flames. *Combustion and Flame*, 37 :227–244, 1980.
- [86] K.W. Morton. *Numerical Solution of Convection-Diffusion Problems*. Chapman & Hall/CRC, 1996.
- [87] B. J. Noye and H. H. Tan. Finite difference methods for solving the two-dimensional convection-diffusion equation. *International Journal for Numerical Methods in Fluids*, 9 :75–98, 1989.
- [88] B.J. Noye and H.H. Tan. Finite difference methods for solving the two-dimensional convection-diffusion equation. *International Journal for Numerical Methods in Fluids*, 9 :75–98, 1989.



- [89] K. Ohmori and T. Ushijima. A technique of upstream type applied to a linear nonconforming finite element approximation of convective diffusion equations. *RAIRO analyse numérique*, 18 :309–322, 1984.
- [90] S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Numer. Anal.*, 21(2) :217–235, 1984.
- [91] A. Ouazzi and S. Turek. Unified edge-oriented stabilization of nonconforming finite element methods for incompressible flows problems. *Journal of Numerical Mathematics*, 15(4) :299–322, 2007.
- [92] PELICANS. Collaborative development environment. <https://gforge.irsn.fr/gf/project/pelicans/>.
- [93] C.D. Perez-Segarra, C. Farre, J. Cadafalch, and A. Oliva. Analysis of different numerical schemes for the resolution of convection-diffusion equations using finite-volume methods on three-dimensional unstructured grids. Part I : Discretization schemes. *Numerical Heat Transfer : Part B : Fundamentals*, 49(4) :333–350, 2006.
- [94] A. Perronnet. *Méfisto*. <http://www.ann.jussieu.fr/~perronnet/mefisto.gene.html>.
- [95] S. Piperno and S. Depeyre. Criteria for the design of limiters yielding efficient high resolution TVD schemes. *Computer & Fluids*, 1998.
- [96] R. Rannacher and S. Turek. Simple nonconforming quadrilateral Stokes element. *Numerical Methods for Partial Differential Equations*, 8 :97–111, 1992.
- [97] H.-G. Roos, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations – Convection-Diffusion-Reaction and Flow Problems*, volume 24 of *Series in Computational Mathematics*. Springer, second edition, 2008.
- [98] F. Schieweck and L. Tobiska. A nonconforming finite element method of upstream type applied to the Navier-Stokes equation. *Mathematical Modelling and Numerical Analysis*, 23 :627–647, 1989.
- [99] F. Schieweck and L. Tobiska. An optimal order error estimate for an upwind discretization of the navier-stokes equations. *Numerical Methods for Partial Differential Equations*, 12 :407–421, 1996.
- [100] M. Schäfer and S. Turek. Benchmark computations of laminar flow around a cylinder. In E.H. Hirschel, editor, *Flow Simulation with High-Performance Computers II*, volume 52 of *Notes on Numerical Fluid Mechanics*, pages 547–566, 1996. co. F. Durst, E. Krause, R. Rannacher.
- [101] H. Si. *TetGen*. <http://tetgen.berlios.de>.
- [102] S. Spekreijse. Multigrid solution of monotone second-order discretizations of hyperbolic conservation laws. *Mathematics of Computation*, 49(179) :135–155, 1987.
- [103] M. Stynes. Convection-diffusion problems, sdfem/supg and a priori meshes. In G. Lube and G. Rapin, editors, *International Conference on Boundary and Interior Layers*, 2006.
- [104] R. Temam. Sur l’approximation de la solution des équations de navier-stokes par la méthode des pas fractionnaires ii. *Arch. Rat. Mech. Anal.*, 33 :377–385, 1969.
- [105] G. Tryggvason. Numerical simulations of the Rayleigh-Taylor instability. *Journal of Computational Physics*, 75 :253–282, 1988.
- [106] S. Turek. *Efficient solvers for incompressible flow problems : an algorithmic approach in view of computational aspects*. Springer-Verlag, 1999.
- [107] B. Van Leer. Towards the ultimate conservative difference scheme. ii. monotonicity and conservation combined in a second order scheme. *Journal of Computational Physics*, 14 :361–370, 1974.

- [108] B. Van Leer. Towards the ultimate conservative difference scheme. V. a second-order sequel to Godunov's methods. *Journal of Computational Physics*, 32 :101–136, 1979.
- [109] V. Venkatakrishnan. A perspective on unstructured grid flow solvers. Technical Report 95-3, NASA/ICASE, 1995.
- [110] M. Wierse. A new theoretically motivated higher order upwind scheme on unstructured grids of simplices. *Advances in Computational Mathematics*, 7 :303–335, 1997.
- [111] J.R. Womersley. Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. *The Journal of Physiology*, 127 :553–563, 1955.
- [112] J. Xu and L. Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation*, 69(229) :1429–1446, 1999.

## Un schéma éléments finis non-conformes/volumes finis pour l'approximation en maillages non-structurés des écoulements à faible nombre de Mach

Nous développons dans cette thèse un schéma numérique pour la résolution sur des maillages non-structurés d'un système d'équations couplant les équations de Navier–Stokes dites "à faible nombre de Mach" à un ensemble d'équations de bilan pour des quantités scalaires. Un schéma de marche en temps du type méthode à pas fractionnaire est appliqué au système global. La contribution principale de la thèse est le développement d'une approximation stable de la prédiction de vitesse discrétisée par éléments finis non-conformes et la mise au point d'un schéma par volumes finis qui soit à la fois stable et robuste vis-à-vis du principe du maximum pour les équations de bilan scalaires.

L'approximation des équations de Navier–Stokes par les éléments finis non-conformes de Crouzeix–Raviart ou de Rannacher–Turek permet de satisfaire la condition de compatibilité inf-sup discrète d'une part et, d'autre part, ces éléments placent les degrés de liberté de vitesse aux centres des faces, ce qui facilite le couplage avec les autres équations de bilan qui sont approximées par volumes finis. La non-conformité inhérente à ces éléments conduit à deux difficultés. L'approximation de Galerkin de la prédiction de vitesse est plus particulièrement sensible aux régimes à convection dominante et aux couches limites. Nous avons ainsi développé, pour des maillages quelconques en hexahédres ou tétraèdres et en maillage structuré axisymétrique, une approximation des termes d'inertie, construite sur un maillage dual et sur la base d'un résultat de volumes finis, qui respecte une inégalité d'énergie et permet le contrôle au niveau discret de la variation d'énergie cinétique par la dissipation visqueuse. La stabilité au sens  $L^2$  de la vitesse est obtenue et on vérifie expérimentalement que l'ordre optimal de convergence en espace est conservé. Une deuxième difficulté est le contrôle de certaines composantes de vitesse en présence de termes de flottabilité. Nous avons proposé une discrétisation "bien équilibrée" du second membre permettant de contrôler les vitesses tout en gardant l'ordre optimal de convergence spatial.

Dans la définition d'un schéma volumes finis pour l'approximation des équations de convection–diffusion, nous sommes confrontés à la nécessité de s'adapter à des maillages provenant du raffinement localisé du maillage dans des géométries complexes ou près des fortes variations de la solution, maillages potentiellement non-structurés voire non-conformes, auxquels les méthodes classiques d'approximation par volumes finis ne sont pas adaptées. Une autre contrainte omniprésente dans la simulation de ces phénomènes est le respect d'un principe de maximum discret, préservant les grandeurs simulées dans leurs bornes physiques. Nous avons donc proposé un couplage nouveau de schémas volumes finis pour l'équation de convection–diffusion : une méthode d'approximation d'ordre élevé des flux convectifs sur des maillages non-structurés et/ou non-conformes inspirée des méthodes MUSCL multipentes, et une méthode de type volumes finis d'approximation consistante des flux diffusifs. On vérifie numériquement, séparément puis de manière couplée, la précision et le respect du principe du maximum pour ces deux méthodes.

Tous les développements effectués sont enfin validés sur des cas concrets d'intérêt pour la simulation des écoulements turbulents réactifs.