N° d'ordre : 07/32

THÈSE

présentée à

L'UNIVERSITÉ DE VALENCIENNES ET DU HAINAUT-CAMBRÉSIS

en vue d'obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ

Spécialité : MATHÉMATIQUES APPLIQUÉES

par

Sarah COCHEZ-DHONDT

Méthodes d'éléments finis et estimations d'erreur a posteriori.

Soutenue le 30 Novembre 2007 devant le jury composé de

Directeur de thèse:	S. Nicaise, Université de Valenciennes
Rapporteurs:	A. Ern, Ecole Nationale des Ponts, Marne la Vallée
	J. Schöberl, RWTH Aachen University, Aix-la-Chapelle
Examinateurs:	F. Ben Belgacem, Université de Technologie de Compiègne
	E. Creusé, Université de Valenciennes
	L. Paquet, Université de Valenciennes

Remerciements

Je voudrais exprimer ma profonde reconnaissance à Monsieur Serge Nicaise pour m'avoir proposé ce sujet de recherche et l'avoir dirigé. Il a su être disponible et m'apporter l'aide et les conseils dont j'avais besoin pour mener à bien mon travail.

Je tiens à exprimer ma gratitude à Messieurs Alexandre Ern et Joachim Schöberl qui ont accepté de rapporter sur cette thèse.

Je souhaiterais aussi adresser mes vifs remerciements à Messieurs Faker Ben Belgacem, Emmanuel Creusé et Luc Paquet pour leur participation à l'examen et au jury de cette thèse.

Je remercie également l'équipe de Simula+ qui m'a accueillie. L'utilisation des codes déjà développés dans ce projet m'a permis de gagner un temps précieux au cours de mon travail de thèse.

Mes remerciements vont aussi à mes collègues du LAMAV pour leur accueil et, en particulier, à Nabila pour son aide précieuse dans les démarches administratives.

Enfin, je voudrais remercier ma famille et plus particulièrement Matthieu qui m'a écouté parler de l'avancement de mon travail patiemment au cours de ces trois années.

Table des matières

Introduction **3**

1 Residual based a posteriori error estimators for the heterogeneous Maxwell equations **9**
1.1 Setting of the problem 9
1.1.1 Quasistatic electromagnetic fields in conductors 9
1.1.2 Electromagnetic fields in dielectrics 11
1.1.3 A common variational formulation 12
1.2 The heterogeneous Maxwell equations 14
1.2.1 The discrete problem 15
1.2.2 Basic tools 16
1.2.3 Bubble functions 18
1.3 Robust a posteriori error estimation 19
1.3.1 Helmholtz Decomposition 19
1.3.2 Interpolation error estimates 20
1.3.3 Error estimates 25
1.3.4 Extension to three-dimensional polyhedral domains 33
1.3.5 Numerical experiments 34
1.3.6 Conclusion 35
1.4 Uniform a posteriori error estimation for the Maxwell equations with discontinuous coefficients 36
1.4.1 Helmholtz Decomposition 40
1.4.2 Interpolation error estimates 42
1.4.3 Error estimates 54
1.4.4 Extension to three-dimensional polyhedral domains 64
1.4.5 Numerical experiments 65
1.4.6 Conclusion 69

2 A posteriori error estimators based on equilibrated fluxes **71**
2.1 Introduction 71
2.2 The two-dimensional reaction-diffusion equation 72
2.3 Upper and lower bounds for the error estimator 74
2.3.1 Upper bound for the discretization error 75

2.3.2 Numerical results 79
2.4 The time-Harmonic Maxwell equations in 3D 84
2.4.1 The approximated problem 84
2.4.2 Conforming approximated problems 86
2.4.3 The a posteriori error estimator 88
2.4.4 Numerical tests 92

3 Comparison of the three a posteriori error estimators 97
3.1 A solution with a boundary layer 97
3.2 The checkerboard : a test with discontinuous coefficients 105
3.3 A singular solution on the L-shape domain 105
3.4 Conclusion 105

4 Equilibrated error estimators for discontinuous Galerkin methods 115
4.1 Introduction 115
4.2 The boundary value problem and its discretization 117
4.2.1 Discontinuous Galerkin approximated problem 117
4.3 The a posteriori error analysis based on Raviart-Thomas finite elements . . 120
4.3.1 Upper bound 124
4.3.2 Lower bound 125
4.4 Numerical tests 127

Conclusion 133

Bibliographie 135

INTRODUCTION

Parmi les méthodes communément utilisées pour approcher numériquement des problèmes apparaissant en ingénierie, comme, par exemple, l'équation de Laplace, le système de Maxwell ([13–15,17,39]), la méthode des éléments finis est l'une des plus populaires. Dans nombre de ces applications, les techniques adaptatives utilisant les estimateurs d'erreur a posteriori sont devenues un outil indispensable. Ces estimateurs permettent de mesurer la qualité de la solution calculée et fournissent une information pour contrôler l'algorithme d'adaptation de maillage.

Estimateurs d'erreur a posteriori

Dans la méthode des éléments finis, on s'intéresse à l'erreur commise entre la solution exacte et la solution approchée. En effet, on se donne une forme bilinéaire coercive B sur un espace de Hilbert V et on s'intéresse à un problème variationnel du type : étant donné f, trouver u dans V solution de

$$B(u,v) = (f,v), \forall v \in V.$$

On peut alors construire une solution discrète $u_h \in V_h$, espace d'approximation de V, satisfaisant

$$B(u_h,v_h) = (f,v_h), \forall v_h \in V_h,$$

et établir des estimations d'erreur *a priori* qui se présentent sous la forme

$$\|u - u_h\| \leq F(h,u),$$

où la fonction F dépend de la solution exacte u et du pas h de la triangulation. Pour que la méthode converge, il faut alors que la solution u soit suffisamment régulière, et, en général, cette solution exacte u n'est pas connue.

Les estimations d'erreur *a posteriori*, introduites en 1978 par Babuška et Rheinboldt, permettent, elles aussi, de contrôler l'erreur exacte en en donnant une approximation, mais, contrairement aux estimations a priori, sans nécessairement connaître la solution exacte ni sa régularité. Elles se présentent sous la forme

$$\|u - u_h\| \leq F(h,u_h,f)$$

où la fonction F peut se calculer explicitement et ne dépend que de la triangulation, de l'approximation éléments finis u_h et de la donnée du problème f. On note η le second membre $F(h,u_h,f)$, appelé estimateur d'erreur a posteriori. Il peut alors s'exprimer en fonction de quantités locales, relatives aux éléments T de la triangulation $\mathcal{T}_h$ que l'on se donne, en s'écrivant sous la forme :

$$\eta = \left(\sum_{T \in \mathcal{T}_h} \eta_T^2 \right)^{1/2}.$$

En pratique, chaque indicateur local η_T est calculé à partir de la solution discrète et des données du problème. Ces indicateurs peuvent alors donner un bon aperçu de la répartition locale de l'erreur et sont donc un outil intéressant pour l'adaptation de maillage.

On attribue à un estimateur certaines propriétés qui attestent de sa qualité. Ainsi, il doit satisfaire trois conditions :

- *fiabilité :* $\|u - u_h\| \lesssim C\eta + \xi, C > 0$,
- *efficacité :* $\eta \lesssim C\|u - u_h\| + \xi, C > 0$,
- *localité :* l'estimateur doit donner des informations sur la distribution locale de l'erreur,

où ξ est une quantité ne dépendant que du second membre f et des conditions de bord du problème et qui est négligeable devant l'estimateur. Ces propriétés indiquent que l'erreur est globalement équivalente à l'estimateur d'erreur. L'efficacité représente l'optimalité de l'estimateur c'est-à-dire la garantie que l'erreur obtenue est petite sans que le coût de calcul ne soit trop élevé.

La qualité d'un estimateur a posteriori est mesurée par son indice d'efficacité correspondant à $\dfrac{\eta}{\|u - u_h\|}$. Si cet indice et son inverse restent bornés, quel que soit le maillage considéré, l'estimateur est dit *efficace*. L'optimalité d'un estimateur est d'avoir un indice d'efficacité égal à 1 et si cet indice tend vers 1 quand la taille du maillage h tend vers 0, l'estimateur est dit asymptotiquement exact.

En général, les bornes supérieures (fiabilité) et inférieures (efficacité) font intervenir des constantes qui dépendent de la régularité des élements et/ou du saut des coefficients intervenant dans les équations, et donc dans la forme bilinéaire, mais cette dépendance est rarement donnée. Le produit de ces constantes mesure la qualité de l'estimateur et si l'équation contient des paramètres critiques, comme par exemple, si elle est perturbée singulièrement, cette quantité doit rester bornée même si le paramètre prend des valeurs extrêmes. L'estimateur est alors dit robuste si les constantes apparaissant dans les bornes inférieures et supérieures sont uniformes par rapport aux coefficients intervenant dans les équations.

Il existe différents types d'estimateurs d'erreur a posteriori et on peut notamment citer les estimateurs de type résiduel [9, 19, 39, 45, 49], les estimateurs d'erreur hiérarchiques [3, 6–8] ou encore les estimateurs basés sur la résolution de problèmes locaux [3, 29]. Les estimateurs résiduels se calculent à partir des sauts du flux discret à travers les interfaces de la triangulation tandis que les estimateurs basés sur des flux équilibrés font intervenir la résolution sur des patches d'éléments de problèmes de type Neumann qui s'expriment en fonction de la solution approchée et la minimisation d'une fonctionnelle.

L'intérêt pour de telles estimations est principalement dû à la volonté des ingénieurs d'obtenir des résultats numériques précis sans que le coût de calcul soit trop élevé. Afin d'optimiser les calculs, les estimations a posteriori permettent de raffiner certaines parties de la triangulation en fonction de la solution approchée. L'adaptation de maillage est donc devenu une outil important dans l'analyse numérique des équations aux dérivés partielles car la performance d'une méthode de résolution numérique est étroitement liée à la qualité du maillage.

Maillages adaptatifs

Lorsqu'on calcule numériquement la solution d'une équation, on est amené à construire successivement des maillages et à résoudre les systèmes linéaires associés. On se heurte ainsi au coût des calculs issus de ces résolutions car les matrices des systèmes contiennent de plus en plus de degrés de liberté. On cherche alors à réduire ce nombre de degrés de liberté en imposant un raffinement uniquement en certaines régions du maillage. En effet, grâce aux estimateurs d'erreur a posteriori et notamment aux indicateurs locaux, on connaît la répartition de l'erreur et l'on sait atteindre uniquement les éléments où elle est la plus élévée.

On introduit alors une procédure de raffinement, basée sur ces indicateurs, pour raffiner localement le maillage et la procédure itérative fonctionne comme suit : on parcourt tous les éléments du maillage et lorsqu'un indicateur local η_T est jugé trop grand sur un élément de la triangulation, cet élément est marqué pour être raffiné. On peut alors imposer un angle minimal, interdisant aux triangles de s'aplatir, c'est-à-dire que les éléments de la triangulation doivent toujours avoir des angles plus grands que cet angle minimal, et on raffine, dans le cas bidimensionnel, un élément suivant trois possibilités :

- soit cet élément sera coupé en deux, si l'angle minimal à respecter le permet,
- soit il sera coupé en trois, si l'angle minimal à respecter le permet,
- soit il sera coupé en quatre triangles.

Pour chaque triangle divisé, il faut alors marquer ses voisins afin qu'ils soient eux-mêmes coupés, suivant les mêmes critères, afin que le maillage reste conforme.

Méthode de type Galerkin discontinue

Dans la méthode des éléments finis, on parle de méthode de Galerkin, et on la dit conforme, lorsque que l'on choisit de calculer la solution élément fini u_h dans un sous-espace V_h, de l'espace V, contenant la solution exacte u. La solution u_h, construite éléments par éléments, vérifie alors des propriétés de continuité aux interfaces entre les éléments.

Lorsqu'on décide de prendre la solution u_h dans un espace V_h qui n'est plus inclus dans l'espace V, on parle alors d'approximation non conforme. On ne s'assure plus une continuité complète entre les éléments mais la solution approchée peut garder une certaine continuité en quelques points des interfaces.

Introduite en 1973 par Reed et Hill, la méthode de type Galerkin discontinue correspond à une approximation non conforme et repose sur le choix d'une base de fonctions discontinues d'un élément à l'autre. La convergence de la méthode est assurée par des contraintes imposées aux interfaces entre les éléments. La solution approchée n'est alors plus continue et l'ordre d'approximation peut être choisi arbitrairement dans chaque élement.
La discontinuité de la représentation permet de n'imposer aucune contrainte sur le maillage. En particulier les maillages non-conformes sont autorisés.

Plan de la thèse

Dans le cas de l'équation de Maxwell en régime harmonique, relativement peu de résultats existent sur les estimations d'erreur a posteriori ; quelques approches ayant cependant été récemment développées pour ce cas ([9,40,45,49]). Ainsi, on peut citer Monk qui soulignait dans son livre [40] que, pour l'équation de Maxwell, les constantes intervenant dans les estimations d'erreur a posteriori et leur dépendance en fonction des coefficients n'avaient jamais été explicitées. C'est notamment à ce problème que nous nous sommes attaqués.

Dans le chapitre 1, nous parlerons d'estimateurs de type résiduel dont nous étudierons la dépendance en fonction des coefficients de l'équation. Nous présenterons d'abord les équations de Maxwell, le problème continu et le problème approché par des sous-espaces conformes, puis nous traiterons séparément, dans un premier temps, le cas de coefficients constants puis, dans un second temps, le cas de coefficients constants par morceaux. Le but est d'y exprimer les bornes supérieures et inférieures en fonction de normes appropriées. Nous serons alors amenés à prouver des estimations d'erreur d'interpolation et pour cela à introduire un nouvel opérateur d'interpolation du type Clément/Nédélec. Nous préciserons la dépendance des constantes intervenant dans les bornes en fonction de la variation des coefficients.

Dans le chapitre 2, notre approche consiste à utiliser celle des flux équilibrés présentée dans [3,13] et nous y proposerons des estimateurs pour des équations de réaction-diffusion et pour le système de Maxwell. Ainsi, pour l'équation de Laplace, l'idée principale est de construire un champ de vecteur j_h, approximation du champ des contraintes,et d'utiliser le terme $j_h - \nabla u_h$ comme estimateur, u_h étant l'approximation élément fini de la solution exacte. Les termes d'ordre zéro, importants en pratique, présentent alors une difficulté supplémentaire, surtout dans le cas de Maxwell, et qui est ici traitée. En effet, dans ce

cas, il faudra introduire une seconde approximation q_h qui prend en compte le fait que l'approximation élément fini (basée sur les éléments finis de Nédélec de plus bas degré) ne respecte pas la contrainte relative à la divergence ; cette deuxième approximation n'ayant pas besoin d'être introduite s'il n'y a pas de terme d'ordre zéro.

Dans le chapitre 3, nous présenterons un bilan des chapitres précédents en établissant une comparaison, au travers de tests numériques sur des solutions particulières présentant des singularités typiques (couche limite, singularité de coin), des estimateurs construits pour l'équation de Maxwell. Nous établirons notamment, sur des algorithmes d'adaptation de maillages, les différences entre les maillages obtenus successivement pour les différents estimateurs lors d'une même procédure de raffinement.

Dans le chapitre 4, nous proposerons l'extension, pour l'équation de diffusion, des estimateurs équilibrés, pour des méthodes éléments finis de type Galerkin discontinues, la difficulté majeure restant actuellement, pour cette méthode, la gestion du terme d'ordre zéro.

Pour tous nos estimateurs et dans chaque chapitre, nous présenterons des tests numériques qui valident les résultats théoriques.

Notons que les chapitres 1, 2 et 4 correspondent à quatre articles acceptés ou soumis. Nous avons donc gardé la structure générale de ces articles ; seules les références ont été regroupées dans une bibliographie commune.

Chapitre 1

Residual based a posteriori error estimators for the heterogeneous Maxwell equations

1.1 Setting of the problem

Let $O = \Omega \times I \subset \mathbb{R}^2 \times \mathbb{R}$ be a bounded domain of $\mathbb{R}^3$ with a polygonal boundary ∂O. The classical Maxwell equations are given by

$$\begin{cases} \partial_t \mathcal{B} + \mathbf{curl}\, \mathcal{E} = 0 & \text{in } O, \\ \operatorname{div} \mathcal{D} = \rho & \text{in } O, \\ \partial_t \mathcal{D} - \mathbf{curl}\, \mathcal{H} = -\mathcal{J} & \text{in } O, \\ \operatorname{div} \mathcal{B} = 0 & \text{in } O, \end{cases} \tag{1.1}$$

where $\mathcal{E}$, $\mathcal{D}$, $\mathcal{B}$, $\mathcal{H}$ and $\mathcal{J}$ are vector functions of position x in $\mathbb{R}^3$ and time t in $\mathbb{R}$.
$\mathcal{E}$ and $\mathcal{H}$ are the electric and magnetic field intensities, $\mathcal{D}$ and $\mathcal{B}$, are respectively the electric displacement and the magnetic induction. $\mathcal{J}(\cdot, t)$ is the source current density which is supposed to satisfy

$$\operatorname{div} \mathcal{J}(\cdot, t) = 0 \text{ in } \Omega, \forall t \geq 0. \tag{1.2}$$

By setting $\mathcal{D} = \epsilon \mathcal{E}$ and $\mathcal{B} = \mu \mathcal{H}$ where ϵ and μ are positive, bounded, scalar functions, respectively called the *electric permittivity* and the *magnetic permeability*, we can find relationships between $\mathcal{E}$ and $\mathcal{H}$ and obtain second-order Maxwell's equations depending either on the magnetic field $\mathcal{H}$ or on the electric field $\mathcal{E}$. In this paper, we arbitrary choose to eliminate $\mathcal{H}$ rather than $\mathcal{E}$.

1.1.1 Quasistatic electromagnetic fields in conductors

The computation of quasistatic electromagnetic fields in conductors usually employs the eddy current model [9]. In this case, $\mathcal{J}$ is given by $\sigma \mathcal{E} + \mathcal{J}_a$ where σ is the conductivity

of the body occupying O and $\mathcal{J}_a(\cdot,t)$ is the source current density which is supposed to satisfy

$$\operatorname{div}\mathcal{J}_a(\cdot,t)=0 \text{ in } \Omega, \forall t\geq 0.$$

This identity allows to transform (1.1) into

$$\left\{\begin{array}{l}\partial_t\mathcal{B}+\mathbf{curl}\,\mathcal{E}=0\\ \partial_t^2\mathcal{D}-\mathbf{curl}\,\partial_t\mathcal{H}=-\sigma\partial_t\mathcal{E}-\partial_t\mathcal{J}_a\end{array}\right.$$

$$\Leftrightarrow\quad\left\{\begin{array}{l}\mu\partial_t\mathcal{H}+\mathbf{curl}\,\mathcal{E}=0\\ \epsilon\partial_t^2\mathcal{E}-\mathbf{curl}\,\partial_t\mathcal{H}=-\sigma\partial_t\mathcal{E}-\partial_t\mathcal{J}_a\end{array}\right.$$

$$\Leftrightarrow\quad\left\{\begin{array}{l}\partial_t\mathcal{H}=\mu^{-1}\,\mathbf{curl}\,\mathcal{E}\\ \epsilon\partial_t^2\mathcal{E}+\sigma\partial_t\mathcal{E}+\mathbf{curl}(\mu^{-1}\,\mathbf{curl}\,\mathcal{E})=-\partial_t\mathcal{J}_a.\end{array}\right.$$

For good conductors, we can assume that $\epsilon\partial_t^2\mathcal{E}=0$ and obtain the parabolic initial boundary value problem [9, 12]

$$\begin{array}{rcll}\partial_t(\sigma\mathcal{E})+\mathbf{curl}(\chi\,\mathbf{curl}\,\mathcal{E})= & -\partial_t\mathcal{J}_a & \text{in } O,\\ \mathcal{E}\times\boldsymbol{n}= & 0 & \text{on } \partial O,\\ \mathcal{E}(\cdot,t=0)= & \mathcal{E}_0 & \text{in } O,\end{array}$$

where $\mathcal{E}$ is the unknown electric field, χ denotes the inverse of the magnetic permeability, and $\boldsymbol{n}$ denotes the unit outward normal vector along ∂O.

Using a time discretization of the above problem we have to solve at each time step Maxwell's equations

$$\left\{\begin{array}{ll}\mathbf{curl}(\chi\,\mathbf{curl}\,\boldsymbol{u})+\beta\boldsymbol{u}=\boldsymbol{f} & \text{in } O,\\ \boldsymbol{u}\times\boldsymbol{n}=0 & \text{on } \partial O,\end{array}\right. \tag{1.3}$$

where $\boldsymbol{u}$ is the time approximation of the electric field $\mathcal{E}$, the coefficient β is equal to $\sigma/\Delta t$ (where Δt is the time step discretization) and $\boldsymbol{f}$ depends on $\mathcal{J}_a$ and the approximation of $\mathcal{E}$ in the previous step. Therefore we may assume that $\boldsymbol{f}$ satisfies

$$\operatorname{div}\boldsymbol{f}=0 \text{ in } O. \tag{1.4}$$

1.1.2 Electromagnetic fields in dielectrics

We now return to the time-dependent problem (1.1) and reduce it to the time-harmonic Maxwell system setting

$$\begin{aligned}
\mathcal{E}(x,t) &= \Re(\exp(-i\omega t)\widehat{\boldsymbol{E}}(x)) \\
\mathcal{D}(x,t) &= \Re(\exp(-i\omega t)\widehat{\boldsymbol{D}}(x)) \\
\mathcal{H}(x,t) &= \Re(\exp(-i\omega t)\widehat{\boldsymbol{H}}(x)) \\
\mathcal{B}(x,t) &= \Re(\exp(-i\omega t)\widehat{\boldsymbol{B}}(x)) \\
\mathcal{J}(x,t) &= \Re(\exp(-i\omega t)\widehat{\boldsymbol{J}}(x)) \\
\rho(x,t) &= \Re(\exp(-i\omega t)\widehat{\rho}(x))
\end{aligned}$$

where $\widehat{\boldsymbol{E}}$ (and similarly other hat variables) are now complex-valued functions depending on the space variables but not on the time variable ([40]). We introduce the linear, inhomogeneous constitutive equations

$$\widehat{\boldsymbol{D}} = \epsilon\widehat{\boldsymbol{E}} \text{ and } \widehat{\boldsymbol{B}} = \mu\widehat{\boldsymbol{H}}.$$

As dielectric materials are characterized by a small conductivity, we take $\sigma = 0$ and the constitutive relation for the currents reduces to $\widehat{\boldsymbol{J}} = \widehat{\boldsymbol{J}}_a$, where the vector function $\widehat{\boldsymbol{J}}_a$ describes the applied current density. As $i\omega\widehat{\rho} = \operatorname{div}\widehat{\boldsymbol{J}}$, we arrive at the following time-harmonic system :

$$\left\{\begin{array}{rcl}
-i\omega\widehat{\boldsymbol{H}} + \mathbf{curl}\,\widehat{\boldsymbol{E}} &=& 0 \\
\operatorname{div}(\epsilon\widehat{\boldsymbol{E}}) &=& \dfrac{1}{i\omega}\operatorname{div}\widehat{\boldsymbol{J}}_a \\
-i\omega\widehat{\boldsymbol{E}} + \sigma\widehat{\boldsymbol{E}} - \mathbf{curl}\,\widehat{\boldsymbol{H}} &=& -\widehat{\boldsymbol{J}}_a \\
\operatorname{div}(\mu\widehat{\boldsymbol{H}}) &=& 0
\end{array}\right.$$

Defining $\boldsymbol{E}$ as $\epsilon_0^{\frac{1}{2}}\widehat{\boldsymbol{E}}$ and $\boldsymbol{H}$ as $\mu_0^{\frac{1}{2}}\widehat{\boldsymbol{H}}$ where ϵ_0 and μ_0 respectively represents the electric permittivity and the magnetic permeability in vacuum, we obtain the second-order Maxwell system for the electric field $\boldsymbol{E} \in \mathbb{R}^3$ [40] :

$$\mathbf{curl}\,\mu_r^{-1}\,\mathbf{curl}\,\boldsymbol{E} - \kappa^2\epsilon_r\boldsymbol{E} = \boldsymbol{f} \text{ in } O, \tag{1.5}$$

$$\boldsymbol{E}\times\boldsymbol{n} = 0 \text{ on } \partial O, \tag{1.6}$$

where $\boldsymbol{f}$ depends on $\widehat{\boldsymbol{J}}_a$, $\kappa = \omega\sqrt{\epsilon_0\mu_0} = \omega c^{-1}$ is called the *wavenumber*, $\omega \geq 0$ is the frequency of the electromagnetic wave and c is the speed of light in vacuum. Moreover, μ_r and ϵ_r are the relative permeability and permittivity of the medium occupying O defined by :

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} \text{ and } \mu_r = \frac{\mu}{\mu_0}.$$

We assume that ϵ_r and μ_r are uniformly bounded from below and above. To get the same system than before, we now set $\beta = \omega^2 c^{-2}\epsilon_r$ and $\chi = \mu_r^{-1}$. With these notations, our equations become

$$\begin{cases} \mathbf{curl}(\chi\,\mathbf{curl}\,\boldsymbol{u}) - \beta\boldsymbol{u} = \boldsymbol{f} & \text{in } O, \\ \boldsymbol{u}\times\boldsymbol{n} = 0 & \text{on } \partial O, \end{cases} \tag{1.7}$$

where $\boldsymbol{u}$ corresponds to $\boldsymbol{E}$, the datum $\boldsymbol{f}$ is once more a multiple of $\boldsymbol{J}$ and so is divergence free.

1.1.3 A common variational formulation

From now on, we reduce the problem (1.3) (or (1.7)) to a problem in the two-dimensional domain Ω, namely assuming that u depends only on the x_1, x_2 variables, then the equations are reduced to :

$$\begin{cases} \mathbf{curl}(\chi\,\mathrm{curl}\,\boldsymbol{u}) + s\beta\boldsymbol{u} = \boldsymbol{f} & \text{in } \Omega, \\ \boldsymbol{u}\cdot\boldsymbol{t} = 0 & \text{on } \Gamma, \end{cases} \tag{1.8}$$

where $\boldsymbol{t}$ is the unit tangential vector along Γ, $s = 1$ in the quasistatic case and $s = -1$ in the dielectric case. For the sake of simplicity, we assume that Ω is simply connected and that its boundary Γ is connected.
We suppose that χ and β are piecewise constant, namely we assume that there exists a partition $\mathcal{P}$ of Ω into a finite set of Lipschitz polygonal domains $\Omega_1, \cdots, \Omega_J$ such that, on each Ω_j, $\chi = \chi_j$ and $\beta = \beta_j$, where χ_j and β_j are positive constants (see Fig. 1.1).

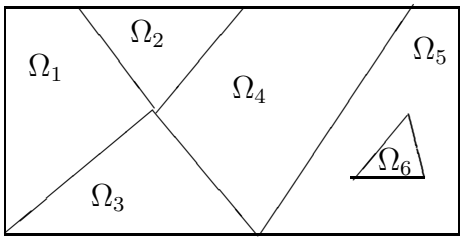


FIG. 1.1 – Partition of the domain Ω.

The variational formulation of (1.8) is well known and involves the space

$$H_0(\mathrm{curl},\Omega) = \{\boldsymbol{u}\in[L^2(\Omega)]^2 : \mathrm{curl}\,\boldsymbol{u}\in L^2(\Omega);\ \boldsymbol{u}\cdot\boldsymbol{t} = 0 \text{ on } \Gamma\}$$

and the bilinear form

$$a(\boldsymbol{u},\boldsymbol{v}) = \int_\Omega (\chi\,\mathrm{curl}\,\boldsymbol{u}\,\mathrm{curl}\,\boldsymbol{v} + s\beta\boldsymbol{u}\cdot\boldsymbol{v})\,dx.$$

For $\boldsymbol{f}\in[L^2(\Omega)]^2$ satisfying (1.4), the weak formulation of (1.8) consists in finding $\boldsymbol{u}\in H_0(\mathrm{curl},\Omega)$ such that

$$a(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{f},\boldsymbol{v}), \forall\boldsymbol{v}\in H_0(\mathrm{curl},\Omega), \tag{1.9}$$

where $(\cdot,\cdot)$ is the $[L^2(\Omega)]^2$-inner product.

In the sequel, we assume that a is coercive on $H_0(\mathrm{curl},\Omega)$, namely we assume that there exists $\alpha > 0$ such that

$$a(\boldsymbol{u},\boldsymbol{u}) \geq \alpha \|\boldsymbol{u}\|^2_{\beta,\chi},\ \forall \boldsymbol{u}\ \in H_0(\mathrm{curl},\Omega) : \mathrm{div}(\beta\boldsymbol{u}) = 0, \tag{1.10}$$

where $\|u\|_{\beta,\chi} = \left(\int_\Omega \chi|\,\mathrm{curl}\,u|^2 + \beta|u|^2\right)^{1/2}$. This coerciveness assumption guarantees that problem (1.8) has a unique solution by the Lax-Milgram lemma.

In the quasistatic case, thanks to the positivity of β and χ, a clearly satisfies coerciveness with $\alpha = 1$.

In the dielectric case, the variational formulation is given by

$$\begin{cases} \text{Find } \boldsymbol{u} \in H_0(\mathrm{curl},\Omega) \text{ such that} \\ (\mu_r^{-1}\,\mathrm{curl}\,\boldsymbol{u}, \mathrm{curl}\,\boldsymbol{v}) - \omega^2 c^{-2}(\epsilon_r \boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f},\boldsymbol{v}),\ \forall \boldsymbol{v} \in H_0(\mathrm{curl},\Omega), \end{cases} \tag{1.11}$$

with $\boldsymbol{u}$ satisfying the divergence constraint $\mathrm{div}(\epsilon_r\boldsymbol{u}) = 0$. If $\omega = 0$, (1.11) has a unique solution. Otherwise, problem (1.11) enters within the framework of the Fredholm alternative and has a unique solution provided ω^2 does not belong to the spectrum of the involved operator. In this paper, we assume that ω is small enough in order to guarantee the coerciveness of a, given here by :

$$a(\boldsymbol{u},\boldsymbol{u}) = \int_\Omega (\mu_r^{-1}|\,\mathrm{curl}\,\boldsymbol{u}|^2 - \omega^2 c^{-2}\epsilon_r|\boldsymbol{u}|^2)\,dx.$$

It means that, if we denote by λ_M^2 the smallest positive eigenvalue of the Maxwell system [40], we assume that $\omega c^{-1} < \lambda_M$. Under this condition, we can estimate the optimal constant of coerciveness and find that :

$$\alpha = \frac{\lambda_M^2 - \omega^2 c^{-2}}{\lambda_M^2 + \omega^2 c^{-2}}.$$

Let us finish this introduction with some notation used in the whole paper : For shortness the $L^2(D)$-norm will be denoted by $\|\cdot\|_D$. In the case $D = \Omega$, we will drop the index Ω. The weighted norm $\|\cdot\|_{D,\beta}$ is defined by

$$\|u\|^2_{D,\beta} := \sum_{j=1}^{J} \beta_j \|u\|^2_{D\cap\Omega_j}.$$

Obviously this norm is equivalent to the $L^2(D)$-norm. As previously if D is equal to Ω, we will drop the index Ω. The standard $H(\mathrm{curl},D)$-norm is denoted by $\|\cdot\|_{H(\mathrm{curl},D)} = \|\cdot\|_D + \|\,\mathrm{curl}\,\|_D$. The usual norm and seminorm of $H^1(D)$ are denoted by $\|\cdot\|_{1,D}$ and $|\cdot|_{1,D}$. For later uses we further need to introduce the space of functions which are piecewise H^k, for $k \in \mathbb{N}$, with respect to the partition of Ω, namely

$$PH^k(\Omega) = \{v \in L^2(\Omega) : v_{|\Omega_j} \in H^k(\Omega_j), \forall j = 1,\cdots,J\},$$

equipped with the norm and semi-norm

$$\|v\|_{PHk,\beta} := \left(\sum_{j=1}^{J} \beta_j \|v_{|\Omega_j}\|^2_{k,\Omega_j}\right)^{1/2},$$

$$|v|_{PHk,\beta} := \left(\sum_{j=1}^{J} \beta_j |v_{|\Omega_j}|^2_{k,\Omega_j}\right)^{1/2},$$

and define $\nabla_P v$ by

$$\nabla_P v_{|\Omega_j} = \nabla(v_{|\Omega_j}),\ \forall j = 1, \ldots, J.$$

The notation $\boldsymbol{u}$ means that the quantity u is a vector and $\nabla \boldsymbol{u}$ means the matrix $(\partial_j u_i)_{1\le i,j\le d}$ (i being the index of row and j the index of column). Finally, the notation $a \lesssim b$ and $a \sim b$ means the existence of positive constants C_1 and C_2, which are independent of $\mathcal{T}$, of the quantities a and b under consideration and of the coefficients β and χ such that $a \lesssim C_2 b$ and $C_1 b \lesssim a \lesssim C_2 b$, respectively.

1.2 The heterogeneous Maxwell equations

Problem (1.9) is approximated in a conforming finite element space V_h of $H_0(\mathrm{curl}, \Omega)$ based on a triangulation $\mathcal{T}$ of the domain made of isotropic triangles, the space V_h is assumed to contain the lowest order Nédélec edge element space (cf. [41]). If $\boldsymbol{u}_h$ is the solution of the discretization of (1.9) we consider an efficient and reliable robust residual a posteriori error estimator for the error $\boldsymbol{e} = \boldsymbol{u} - \boldsymbol{u}_h$ in the $H_0(\mathrm{curl}, \Omega)$-norm.

A posteriori error estimators for standard elliptic boundary value problems are in our days well understood [52]. The analysis of isotropic a posteriori error estimators for the edge elements were successfully initiated in [9, 39] in the context of a "smooth" Helmholtz decomposition. The methods from [9] and [31] were combined in [45] to the case of anisotropic meshes and for a "nonsmooth" Helmholtz decomposition. Alternatively, using a local $H(\mathrm{curl})$ decomposition of the interpolation error (of Clément type) and its local Helmholtz decomposition from [47], J. Schöberl proves in [49] an a posteriori estimate in the case χ and β constant and Ω not necessarily convex. In these papers, the dependence of the constants in the lower and upper bounds with respect to the variation of the constants is not explicitely given. Therefore, the goal of this chapter is to give this dependence in the case of piecewise constant coefficients β and χ, extending to Maxwell's equations what have already been shown for second order elliptic operators with piecewise constant coefficients [11]. Note that the question of the dependence of these constants with respect to the coefficients was raised in [40] page 362.

In this chapter, we will first consider the case when χ and β are constant in the whole Ω, introducing an interpolant of Clément/Nédélec type, and analyze an a posteriori error estimator. Then we extend some ideas to the case of piecewise constant coefficients, but contrary to the previous section, we introduce two different estimators. For the sake of

simplicity, we have restricted ourselves to 2D problems, the extension to the 3D setting is mainly direct (see section 1.4.4).

The schedule of this chapter is the following one : Section 1.2.1 recalls the discretization of our problem. In section 1.2.2, we recall some basic tools for the error estimation analysis. We further study the case of constant coefficients. In this part, we state the adapted Helmholtz decomposition of the error. In section 1.4.2 we give some interpolation error estimates for Clément and Nédélec interpolants, build a new interpolation operator based on the first two ones and prove appropriate interpolation error estimates. The efficiency and reliability of the estimator are established in section 1.4.3. The extension of our results to three-dimensional problems is shortly described in section 1.4.4. Section 1.4.5 is devoted to numerical tests which confirm our theoretical analysis.

1.2.1 The discrete problem

We consider a triangulation $\mathcal{T}_h$ made of triangles denoted by T, T_i or T' whose edges are denoted by e and nodes by x.
We assume that this triangulation is regular i.e. for any element T, the ratio $\frac{h_T}{\rho_T}$ is bounded by a constant $\sigma > 0$ independent of T and of $h = max_{T\in\mathcal{T}_h} h_T$, where h_T is the diameter of T and ρ_T the diameter of its largest inscribed ball.
We will denote by h_e the length of an edge e. The set of edges will be denoted by $\mathcal{E}_h$. Let $\mathcal{N}_\Omega$ be the set of internal nodes of the triangulation and $\mathcal{E}_{h\Omega}$ the set of its internal edges. For an edge e of an element T, we introduce the *outer normal vector* by $\boldsymbol{n}_e$. We define the jump of a function v across an edge as :

$$\left[\left[v(\boldsymbol{y})\right]\right]_e = \lim_{\alpha\to+0} v(\boldsymbol{y}+\alpha\boldsymbol{n}_e) - v(\boldsymbol{y}-\alpha\boldsymbol{n}_e),\ \boldsymbol{y}\ \in e.$$

Note that the sign of $\left[\left[v\right]\right]_e$ depends on the orientation of $\boldsymbol{n}_e$.However, terms such as a gradient jump $\left[\left[\nabla v\cdot\boldsymbol{n}_e\right]\right]_e$ are independent of this orientation.
At least, one uses so called patches :

- ω_T is the union of all elements having a common edge with T,
- ω_e the union of both elements having e as edge,
- ω_x the union of all elements having x as node.

Problem (1.9) is approximated in a curl-conforming finite element subspace V_h of $H_0(\mathrm{curl},\Omega)$ containing the lowest order Nédélec finite element space :

$$V_{h,1} = \{\boldsymbol{v}_h\ \in H_0(\mathrm{curl},\Omega) : \boldsymbol{v}_{h|T} \in \mathcal{ND}_1, \forall T\ \in \mathcal{T}_h\},$$

where $\mathcal{ND}_1$ is given by :

$$\mathcal{ND}_1 = \left\{ p\in\mathcal{P}_1(T)^2 | \exists a\in\mathbb{R}^2, b\in\mathbb{R}, \forall x\in T, p(x) = a + b\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}\right\}.$$

For instance, we may take for V_h the subspace of $H_0(\mathrm{curl},\Omega)$ consisting of functions which are piecewise in $\mathcal{ND}_k$ with $k\geq 1$, as considered in [9, 41] (see [45]).

The discretized problem of problem (1.8) is to find $u_h \in V_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h), \ \forall \boldsymbol{v}_h \in V_h. \tag{1.12}$$

We define the error by :

$$\boldsymbol{e} = \boldsymbol{u} - \boldsymbol{u}_h,$$

and from (1.9), we obtain the defect equation

$$a(\boldsymbol{e}, \boldsymbol{v}) = r(\boldsymbol{v}), \ \forall \boldsymbol{v} \in H_0(\mathrm{curl}, \Omega), \tag{1.13}$$

where the residual is given by

$$r(\boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) - a(\boldsymbol{u}_h, \boldsymbol{v}), \ \forall \boldsymbol{v} \in H_0(\mathrm{curl}, \Omega), \tag{1.14}$$

Then, we deduce from (1.12) the "**Galerkin orthogonality**" relation

$$a(\boldsymbol{e}, \boldsymbol{v}_h) = r(\boldsymbol{v}_h) = 0, \ \forall \boldsymbol{v}_h \in V_h. \tag{1.15}$$

1.2.2 Basic tools

In this section, we introduce some notations and important tools. Some basic relations and lemmas are given as well.

Let us first introduce auxiliary subdomains also called patches : For any triangle T of $\mathcal{T}_h$ and any edge e of $\mathcal{E}_h$, we denote by

- ω_T : the union of all elements having a common edge with T,
- ω_e : the union of both elements having e as edge.

If the parameter χ is small, our Maxwell equations are singularly perturbed. Therefore as in [53], for a real number $\delta \in (0, 1]$ we employ a squeezed element $T_{e,\delta} \subset T$ associated with T and an edge e of T (see Fig. 1.2) introduced in [31, 33, 34, 53] ; the main properties of $T_{e,\delta}$ is to be included into T, to have e as edge and to be of height $\sim \delta h_e$. More precisely, if T is the triangle OQ_1Q_2 and the edge $e = Q_1Q_2$, denote by S_e the midpoint of the edge e, then $T_{e,\delta}$ is the triangle PQ_1Q_2, where the point P lies on the line S_eO such that $|S_eP| = \delta|S_eO|$ (see Fig. 1.2 and [33]).

Now, recall that for any T of $\mathcal{T}_h$, we can define a continuous affine linear mapping transforming the *reference triangle* $\widehat{T}$, whose vertices are given by $\left\{(0,0)^T, \ (1,0)^T, \ (0,1)^T\right\}$, onto T.

Then, in order to use efficiently $T_{e,\delta}$, we require an affine linear transformation $F_{T,e,\delta}$ that maps the reference triangle onto $T_{e,\delta}$. This affine linear mapping is unique.

In the same way, we can now introduce the following transformation [31] :

from a patch ω_T, $T \in \mathcal{T}_h$, to a reference patch :

Denote by $\widehat{\omega}_T$ the reference patch corresponding to ω_T (see Figure 1.3 for the case when

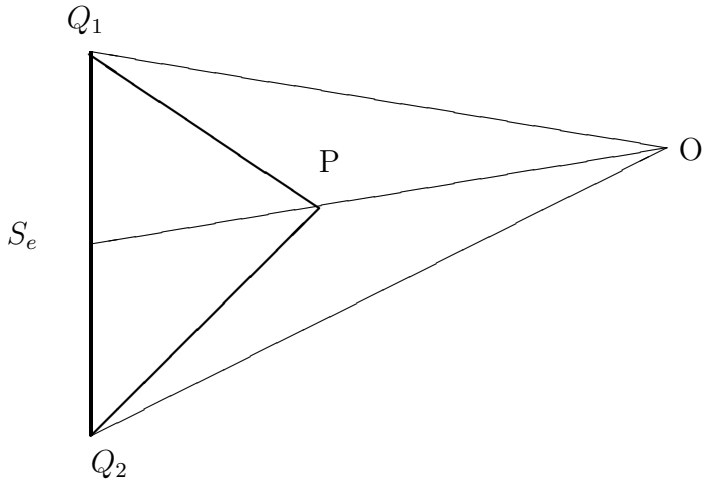


FIG. 1.2 – Triangles $T = OQ_1Q_2$ and $T_{e,\delta} = PQ_1Q_2$.

$T \cap \Gamma$ is empty or reduced to a vertex). Then there exists a continuous, piecewise linear mapping F_T that satisfies :

$$\begin{array}{rcccl} F_T & : & \widehat{\omega}_T & \to & \omega_T \\ F_{T|T_i} = F_{T_i} & : & \widehat{T}_i & \to & T_i,\ \forall i = 1, \ldots, l_T \\ & & \widehat{x} & \to & B_{T_i}\widehat{x} + b_{T_i} \end{array}$$

where $l_T = 2, 3$ or 4, $B_{T_i} \in \mathbb{R}^{2\times 2}$ and $b_{T_i} \in \mathbb{R}^2$. On each $T_i \subset \omega_T$, we set

$$\widehat{\boldsymbol{u}}(\widehat{x}) = B_{T_i}(\boldsymbol{u} \circ F_{T_i}(\widehat{x})). \tag{1.16}$$

from a patch ω_x, $x \in \mathcal{N}_\Omega$, to a reference patch :
Assume that ω_x consists of N triangles arbitrary numbered, and denote by $\widehat{\omega}_x$ the regular N-polygon with the midpoint in the coordinate origin. Then there exists a continuous, piecewise linear mapping F_x that satisfies :

$$\begin{array}{rcccl} F_T & : & \widehat{\omega}_x & \to & \omega_x \\ F_{x|\widehat{T}_i} = F_i & : & \widehat{T}_i & \to & T_i,\ i = 1, \cdots, N \\ & & \widehat{x} & \to & B_i\widehat{x} + b_i \end{array}$$

where $B_i \in \mathbb{R}^{2\times 2}$ and $b_i \in \mathbb{R}^2$. On each $T_i \subset \omega_x,\ i = 1, \cdots, N$, we set

$$\widehat{u}(\widehat{x}) = B_i(u \circ F_i(\widehat{x})). \tag{1.17}$$

Remark 1.2.1. $|\widehat{T}_i| = |\widehat{T}_j|,\ \forall i, j = 1, \cdots, N$.

At least, the lemma below has been proved in [53]. It will play an important role in interpolation estimates.

Lemma 1.2.2. *Let T be an arbitrary triangle and e an edge of it. For $\boldsymbol{v} \in H^1(T)^2$, the trace inequality holds :*

$$\|\boldsymbol{v}\|_e^2 \lesssim \|\boldsymbol{v}\|_T \cdot \left(h_T^{-1}\|\boldsymbol{v}\|_T + \|\nabla \boldsymbol{v}\|_T\right). \tag{1.18}$$

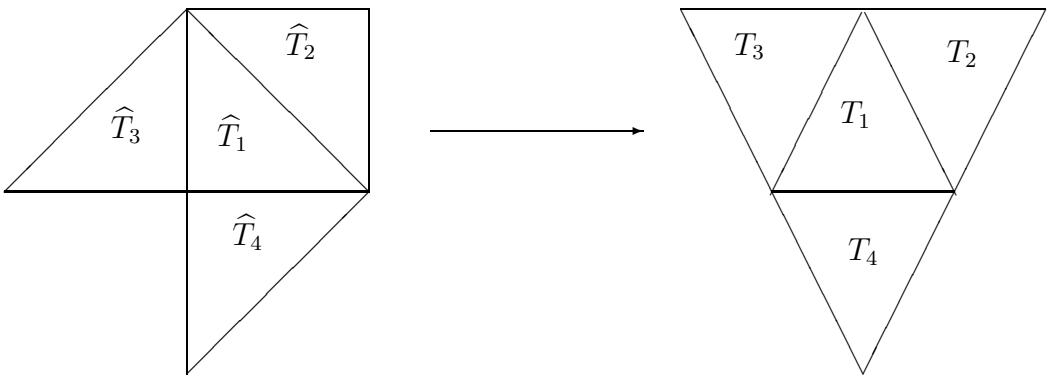


FIG. 1.3 – Transformation F_T that associate the patches $\widehat{\omega}_T = \cup_{i=1,\dots,4}\widehat{T}_i$ and $\omega_T = \cup_{i=1,\dots,4}T_i$.

1.2.3 Bubble functions

For the analysis, we need to introduce bubble functions satisfying some properties. We first define the *element bubble function* $b_{\widehat{T}} \in C(\widehat{T})$ given by $b_{\widehat{T}}(\widehat{x},\widehat{y}) = 3^3\widehat{x}\widehat{y}(1-\widehat{x}-\widehat{y})$, where $\widehat{T}$ is the reference element, and then an *edge bubble function* $b_{\widehat{e},\widehat{T}} \in C(\widehat{T})$ for the edge $\widehat{e} \subset \partial\widehat{T} \cap \{\widehat{y}=0\}$ defined by $b_{\widehat{e},\widehat{T}} = 2^2\widehat{x}(1-\widehat{x}-\widehat{y})$. Furthermore, we require an extension operator $F_{ext} : C(\widehat{e}) \to C(\widehat{T})$, $F_{ext}(v_{\widehat{e}})(\widehat{x},\widehat{y}) := v_{\widehat{e}}(\widehat{x})$.

For a given element T of the triangulation, we obtain the bubble function b_T by the affine linear transformation F_T and the edge bubble function $b_{e,T}$ is similarly defined. We also introduce an edge bubble function b_e on the domain $\omega_e = T_1 \cup T_2$ with an elementwise definition : $b_{e|T_i} := b_{e,T_i}$, $i = 1,2$. Analoguously the extension operator is defined for functions $v_e \in C(e)$ and a same elementwise definition implies that $F_{ext}(v_e) \in C(\omega_e)$.
We recall that $b_T = 0$ on ∂T, $b_e = 0$ on $\partial\omega_e$ and $\|b_T\|_{\infty,T} = \|b_e\|_{\infty,\omega_e} = 1$.
Now, we can state *inverse inequalities* (proved in [52] for instance) :

Lemma 1.2.3. *Let $v_T \in \mathbb{P}^{k_0}(T)$ and $v_e \in \mathbb{P}^{k_1}(e)$. Then the following equivalences and inequalities hold. The implicit constants depend on the polynomial degree k_0 and k_1 but not on T, e or v_T, v_e.*

$$\|v_T b_T^{\frac{1}{2}}\|_T \sim \|v_T\|_T \tag{1.19}$$
$$\|\nabla(v_T b_T)\|_T \lesssim h_T^{-1}\|v_T\|_T \tag{1.20}$$
$$\|v_e b_e^{\frac{1}{2}}\|_e \sim \|v_e\|_e \tag{1.21}$$
$$\|F_{ext}(v_e)b_e\|_T \lesssim h_T^{\frac{1}{2}}\|v_e\|_e \tag{1.22}$$
$$\|\nabla(F_{ext}(v_e)b_e)\|_T \lesssim h_T^{-\frac{1}{2}}\|v_e\|_e \tag{1.23}$$

These bubble functions do not suffice to analyse our residual error estimators. We further need to introduce modified edge bubble functions, cf. also [33]. For some triangle

T and an edge e thereof consider the subtriangle $T_{e,\delta}$ (cf. Figure 1.2). Define the *squeezed edge bubble function* $b_{e,T,\delta}$ by

$$b_{e,T,\delta} := \begin{cases} b_{\widehat{e}} \circ F_{e,T,\delta}^{-1} & \text{on } T_{e,\delta} \\ 0 & \text{on } T \backslash T_{e,\delta} \end{cases} \tag{1.24}$$

where $b_{\widehat{e}}$ is the standard edge bubble function for the edge $\widehat{e} = F_{e,T,\delta}^{-1}(e)$ of the triangle $\widehat{T} = F_{e,T,\delta}^{-1}(T_{e,\delta})$. In other words, $b_{e,T,\delta}$ is the usual bubble function for the edge e in the triangle $T_{e,\delta}$, and it is extended by zero in $T \backslash T_{e,\delta}$.
Standard scaling arguments using the transformation $F_{e,T,\delta} : \widehat{T} \to T_{e,\delta}$ yield the next inverse inequalities for the squeezed edge bubble function, see [33, 34, 53].

Lemma 1.2.4. *Let e be an arbitrary edge of T and assume that $v_e \in \mathbb{P}^{k_1}(e)$. Then the following equivalences and inequalities hold. The implicit constants depend on the polynomial degree k_1 but not on T, e or v_e.*

$$\|F_{ext}(v_e) b_{e,T,\delta}\|_T \lesssim \delta^{\frac{1}{2}} h_T^{\frac{1}{2}} \|v_e\|_e, \tag{1.25}$$

$$\|\nabla(F_{ext}(v_e) b_{e,T,\delta})\|_T \lesssim \delta^{\frac{1}{2}} h_T^{-\frac{1}{2}} \min\{\delta, 1\}^{-1} \|v_e\|_e. \tag{1.26}$$

1.3 Robust a posteriori error estimation

To our knowledge, a robust estimation was not yet considered for the Maxwell system. Our method relies on the introduction of an interpolant of Clément/Nédélec type satisfying appropriate interpolation error estimates.

We consider a robust a posteriori error estimator of residual type for the Maxwell equations in a bounded two (and three) dimensional domain. The continuous problem is approximated using conforming approximated spaces. The main goal is to express the lower and upper bounds with respect to appropriate norms. For that purpose, a new interpolant of Clément/Nédélec type is introduced and some interpolation error estimates are proved. Numerical tests are presented which confirm our theoretical results.

We consider first that the coefficients β and χ are positive constants and we take $s = 1$ in the bilinear form.

1.3.1 Helmholtz Decomposition

Here we mainly recall the standard Helmholtz decomposition of the space $H_0(\mathrm{curl}, \Omega)$. Recall that $H_0(\mathrm{curl}, \Omega)$ was equipped with the inner product

$$(\boldsymbol{v}, \boldsymbol{w})_{\beta,\chi} = (\beta \boldsymbol{v}, \boldsymbol{w}) + (\chi \,\mathrm{curl}\, \boldsymbol{v}, \mathrm{curl}\, \boldsymbol{w}),$$

its associated norm $\|\boldsymbol{v}\|_{\beta,\chi}$ being equivalent to the usual norm $(\|\boldsymbol{v}\|^2 + \|\,\mathrm{curl}\, \boldsymbol{v}\|^2)^{1/2}$.

Lemma 1.3.1. *If Ω is simply connected and its boundary Γ is connected then*

$$H_0(\mathrm{curl},\Omega) = H_0^0(\mathrm{curl},\Omega) \overset{\perp}{\oplus} \mathcal{W}, \tag{1.27}$$

where $H_0^0(\mathrm{curl},\Omega)$ and $\mathcal{W}$ are closed subspaces of $H_0(\mathrm{curl},\Omega)$ defined by

$$\begin{aligned} H_0^0(\mathrm{curl},\Omega) &= \{\boldsymbol{v} \in H_0(\mathrm{curl},\Omega) : \mathrm{curl}\,\boldsymbol{v} = 0 \text{ in } \Omega\}, & (1.28)\\ \mathcal{W} &= \{\boldsymbol{v} \in H_0(\mathrm{curl},\Omega) : \mathrm{div}\,\boldsymbol{v} = 0 \text{ in } \Omega\}, & (1.29) \end{aligned}$$

and the symbol $\overset{\perp}{\oplus}$ means that the decomposition is direct and orthogonal with respect to the inner product $(\cdot,\cdot)_{1,1}$. Furthermore one has

$$H_0^0(\mathrm{curl},\Omega) = \nabla H_0^1(\Omega). \tag{1.30}$$

Then the error $\boldsymbol{e}$ admits the splitting

$$\boldsymbol{e} = \boldsymbol{e}_0 + \boldsymbol{e}_\perp, \tag{1.31}$$

with $\boldsymbol{e}_0 = \nabla\phi$ where $\phi \in H_0^1(\Omega)$ and $\boldsymbol{e}_\perp \in \mathcal{W}$. Moreover $\boldsymbol{e}_\perp$ admits the splitting

$$\boldsymbol{e}_\perp = \nabla\psi + \mathbf{w}, \tag{1.32}$$

where $\psi \in H_0^1(\Omega)$ and $\mathbf{w} \in H_0^1(\Omega)^2$ and satisfies

$$\begin{aligned} \|\nabla\mathbf{w}\|_\beta &\lesssim \beta^{1/2}\chi^{-1/2}\|\boldsymbol{e}_\perp\|_{\beta,\chi}, & (1.33)\\ \|\mathbf{w}\|_\beta &\lesssim \|\boldsymbol{e}_\perp\|_{\beta,\chi}. & (1.34) \end{aligned}$$

Proof: All the results have been proved in Lemma 3.10 and Corollary 3.11 of [45], except the decomposition (1.32) and the estimates (1.33) and (1.34). The decomposition (1.32) of $\boldsymbol{e}_\perp$ was proved in Lemma 2.2 of [47] (for three-dimensional polyhedral domains, but their proof is also valid for two-dimensional polygonal domains) with the estimate

$$\begin{aligned} &\|\mathbf{w}\| + \|\nabla\psi\| \lesssim \|\boldsymbol{e}_\perp\|,\\ &\|\nabla\mathbf{w}\| \lesssim \|\,\mathrm{curl}\,\boldsymbol{e}_\perp\|. \end{aligned}$$

These estimates directly lead to (1.33) and (1.34), because β and χ are constant. ∎

1.3.2 Interpolation error estimates

Clément interpolation

We first recall that the Clément interpolation operator [33, 45] of some function $\phi \in H_0^1(\Omega)$ is defined by :

$$\begin{aligned} \mathrm{I}_{\mathrm{Cl}} : \; H_0^1(\Omega) &\to S(\Omega,\mathcal{T}_h)\\ \phi &\to \sum_{x\,\in\mathcal{N}_\Omega} \frac{1}{|\omega_x|}\left(\int_{\omega_x}\phi\right)\varphi_x = \sum_{x\,\in\mathcal{N}_\Omega} \mathrm{I}_{\mathrm{Cl},x}(\phi)\varphi_x \end{aligned}$$

where $S(\Omega, \mathcal{T}_h)$ is the space of continuous piecewise linear functions on the triangulation which are zero on the boundary and φ_x is the nodal basis function associated with the node x, uniquely determined by the condition :

$$\varphi_x(y) = \delta_{x,y}, \ \forall y \ \in \mathcal{N}_\Omega.$$

We now can state the following interpolation error estimates (see [18]) :

Lemma 1.3.2. *For every function $\phi \ \in H_0^1(\Omega)$, we have*

$$\sum_{T\in\mathcal{T}_h} h_T^{-2}\|\phi - \mathrm{I}_{\mathrm{Cl}}\phi\|_T^2 \lesssim \|\nabla\phi\|^2, \tag{1.35}$$

$$\sum_{T\in\mathcal{T}_h} \|\nabla\left(\phi - \mathrm{I}_{\mathrm{Cl}}\phi\right)\|_T^2 \lesssim \|\nabla\phi\|^2. \tag{1.36}$$

Since our problem also involves functions in $\mathcal{W}$, we need a Nédélec-type interpolant in order to approximate such functions by an $H(\mathrm{curl})$-conforming interpolant. We start by recalling the definition of the Nédélec operator given in [37] and then, as we need a L^2-stability of our operator, we introduce a new interpolant based on the definitions of the previous ones.

Nédélec interpolation

Let $T \in \mathcal{T}_h$ be a triangle and $\mathcal{E}_T$ the set of its edges. For $e \in \mathcal{E}_T$, we fix a unit tangential vector $\boldsymbol{t}_e$ along the edge e. We define (see [37]) the set of linear forms $\{l_e, e \in \mathcal{E}_T\}$ by

$$\begin{aligned} l_e : L^1(e) &\rightarrow \mathbb{R} \\ u &\rightarrow \int_e u\cdot\boldsymbol{t}_e\, ds, \end{aligned}$$

and consider the (basis) functions $\lambda_e \in \mathcal{ND}_1$ satisfying the condition

$$\forall e \in \mathcal{E}_T, \int_{e'} \lambda_e\cdot\boldsymbol{t}_{e'} = \delta_{e,e'}.$$

We further introduce the local interpolation operator $\mathrm{I}_{\mathrm{Ned}|T}(u) \in \mathcal{ND}_1$ defined for u satisfying $u_{|e} \in \left(L^1(e)\right)^2$ by the conditions

$$l_e\left(\mathrm{I}_{\mathrm{Ned}|T}(u)\right) = l_e(u), \forall e \in \mathcal{E}_T.$$

This means that

$$\mathrm{I}_{\mathrm{Ned}|T}(u) = \sum_{e\in\mathcal{E}_T}\left(\int_e u\cdot\boldsymbol{t}_e\, ds\right)\lambda_e.$$

The global interpolation operator $\mathrm{I}_{\mathrm{Ned}}$ is then given by $(\mathrm{I}_{\mathrm{Ned}}u)_{|T} = \mathrm{I}_{\mathrm{Ned}\,|T}(u_{|T}) \in \mathcal{ND}_1, \forall T \in \mathcal{T}_h$ as

$$\begin{array}{rcl} \mathrm{I}_{\mathrm{Ned}} : H_0(\mathrm{curl},\Omega) & \to & V_h \\ u & \to & \displaystyle\sum_{e\in\mathcal{E}_\Omega} \left(\int_e u\cdot \boldsymbol{t}_e\, ds\right)\lambda_e. \end{array}$$

A Clément-Nédélec interpolant

Let us define a Clément-Nédélec interpolant by :

$$\begin{array}{rcl} \mathrm{I}_{\mathrm{CN}} : L^2(\Omega)^2 & \to & V_h \\ \boldsymbol{u} & \to & \displaystyle\sum_{e\in\mathcal{E}_{h\,\Omega}} \alpha_e(\boldsymbol{u})\tilde{\lambda}_e \end{array}$$

where $\alpha_e(\boldsymbol{u}) = \dfrac{1}{|\omega_e|}\displaystyle\int_{\omega_e} \boldsymbol{u}\cdot\boldsymbol{t}_e$ and $\tilde{\lambda}_e = \lambda_e|e|$.

This new interpolant is well-defined and stable relatively to the L^2-norm and the H^1-semi-norm and satisfies standard interpolant error estimates, i.e. we have the following estimates :

Theorem 1.3.3. *For every function $\boldsymbol{u} \in H_0(\mathrm{curl},\Omega)\cap H^1(\Omega)^2$, we have*

$$\|\mathrm{I}_{\mathrm{CN}}\boldsymbol{u}\|_T \lesssim \|\boldsymbol{u}\|_{\omega_T} \tag{1.37}$$

$$\|\boldsymbol{u}-\mathrm{I}_{\mathrm{CN}}\boldsymbol{u}\|_T \lesssim h_T\|\nabla\boldsymbol{u}\|_{\omega_T} \tag{1.38}$$

$$\|\nabla(\boldsymbol{u}-\mathrm{I}_{\mathrm{CN}}\boldsymbol{u})\|_T \lesssim \|\nabla\boldsymbol{u}\|_{\omega_T}. \tag{1.39}$$

Proof: We first define

$$\mathcal{R}_0(\omega_T) = \left\{\mathbf{c}\in H(\mathrm{curl},\omega_T)\ :\ \mathbf{c}_{|T'}\in\mathbb{R}^2, \forall T'\subset\omega_T \text{ and } \mathbf{c}\cdot\mathbf{t}=0 \text{ on } \partial\omega_T\cap\Gamma\right\}$$

and prove that $\mathrm{I}_{\mathrm{CN}}\mathbf{c} = \mathbf{c}$ on T if $\mathbf{c}\in\mathcal{R}_0(\omega_T)$. Indeed, for $e\subset T$, we have

$$\alpha_e(\mathbf{c}) = \frac{1}{|\omega_e|}\int_{\omega_e}\mathbf{c}\cdot\boldsymbol{t}_e = \mathbf{c}\cdot\boldsymbol{t}_e = \frac{1}{|e|}\int_e \mathbf{c}\cdot\boldsymbol{t}_e.$$

Then, the definition of $\mathrm{I}_{\mathrm{Ned}}$ implies that $\mathrm{I}_{\mathrm{CN}\,|T}\mathbf{c} = \mathrm{I}_{\mathrm{Ned}\,|T}\mathbf{c} = \mathbf{c}$.

Let us now show (1.37) : By Cauchy-Schwarz's inequality, we may write

$$|\alpha_e(\boldsymbol{u})| \leq \frac{1}{|\omega_e|}\|\boldsymbol{u}\|_{\omega_e}\|\boldsymbol{t}_e\|_{\omega_e}$$

Since

$$\|\boldsymbol{t}_e\|_{\omega_e} \leq \|\boldsymbol{t}_e\|_\infty|\omega_e|^{\frac{1}{2}}$$

and $\|\boldsymbol{t}_e\|_\infty = 1$, we get

$$|\alpha_e(\boldsymbol{u})| \leq \frac{1}{|\omega_e|^{\frac{1}{2}}}\|\boldsymbol{u}\|_{\omega_e}.$$

By the definition of I_{CN}, we obtain

$$\|\mathrm{I}_{\mathrm{CN}}\,\boldsymbol{u}\|_{T'} \leq \sum_{e\subset\partial T'} \frac{1}{|\omega_e|^{\frac{1}{2}}}\|\boldsymbol{u}\|_{\omega_e}\|\tilde{\lambda}_e\|_{T'}.$$

Moreover, if $\widehat{\lambda}$ denote a basis function on the reference element, we have

$$\|\tilde{\lambda}_e\|_{T'} = |e|\|\lambda_e\|_{T'} = |e|\|B_{T'}^{-t}\widehat{\lambda}\|_{T'} \lesssim |e|h_{T'}^{-1}\|\widehat{\lambda}\|_{\widehat{T'}}h_{T'} \lesssim |e|$$

where $B_{T'}$ is the matrix refering to the affine transformation $F_{T'}$ that maps $\widehat{T}' \subset \widehat{\omega}_T$ onto $T' \subset \omega_T$.

As $|e| = h_e$ and $|\omega_e| \sim h_e^2$, we conclude that

$$\|\mathrm{I}_{\mathrm{CN}}\,\boldsymbol{u}\|_{T'} \lesssim \sum_{e\subset\partial T'} \|\boldsymbol{u}\|_{\omega_e}$$

which implies (1.37).

Now, for any $\mathbf{p} \in \mathcal{R}_0(\omega_T)$, $\boldsymbol{u} - \mathrm{I}_{\mathrm{CN}}\,\boldsymbol{u} = (\mathrm{I} - \mathrm{I}_{\mathrm{CN}})(\boldsymbol{u} - \mathbf{p})$ on T and therefore by (1.37) :

$$\begin{aligned}\|\boldsymbol{u} - \mathrm{I}_{\mathrm{CN}}\,\boldsymbol{u}\|_T &= \|(\mathrm{I} - \mathrm{I}_{\mathrm{CN}})(\boldsymbol{u} - \mathbf{p})\|_T \\ &\lesssim \|\boldsymbol{u} - \mathbf{p}\|_{\omega_T}.\end{aligned}$$

Now we define

$$PH^1(\omega_T) = \{v \in L^2(\omega_T) : v_{|T'} \in H^1(T')\forall T' \subset \omega_T\}.$$

From the above estimate, we see that (1.38) holds if we can bound from below the ratio

$$\frac{h_T^2 \sum\limits_{T'\subset\omega_T} \|\nabla\boldsymbol{u}\|_{T'}^2}{\sum\limits_{T'\subset\omega_T} \|\boldsymbol{u} - \mathbf{p}\|_{T'}^2}$$

for $\boldsymbol{u} \in H(\mathrm{curl}, \omega_T) \cap PH^1(\omega_T)^2$ such that $\boldsymbol{u}\cdot\boldsymbol{t} = 0$ on $\partial\omega_T \cap \Gamma$ and $\mathbf{p} \in \mathcal{R}_0(\omega_T)$, which is equivalent, by applying the affine transformation F_T mapping the patch $\widehat{\omega}_T$ to ω_T (see section 1.2.2), to bound the ratio

$$\frac{\sum\limits_{\widehat{T}'\subset\widehat{\omega}_T} \|\widehat{\nabla}\widehat{\boldsymbol{u}}\|_{\widehat{T}'}^2}{\sum\limits_{\widehat{T}'\subset\widehat{\omega}_T} \|\widehat{\boldsymbol{u}} - \widehat{\mathbf{p}}\|_{\widehat{T}'}^2} \tag{1.40}$$

for $\widehat{\boldsymbol{u}} \in H(\mathrm{curl}, \widehat{\omega}_T) \cap PH^1(\widehat{\omega}_T)^2$ such that $\widehat{\boldsymbol{u}}\cdot\widehat{\boldsymbol{t}} = 0$ on $\widehat{\Gamma}_T$ and $\widehat{\mathbf{p}} \in \mathcal{R}_0(\widehat{\omega}_T)$, where $\widehat{\Gamma}_T$ is made of some edges of the boundary of $\widehat{\omega}_T$. This last ratio will be estimated from below

using the min-max principle.

Indeed, let us set $V = \{\widehat{\boldsymbol{v}} \in H(\mathrm{curl}, \widehat{\omega}_T) \cap PH^1(\widehat{\omega}_T)^2 : \widehat{\boldsymbol{v}} \cdot \widehat{\mathbf{t}} = 0 \text{ on } \widehat{\Gamma}_T\}$ and $H = L^2(\widehat{\omega}_T)^2$. Define the bilinear form

$$l(u, v) = \sum_{\widehat{T}' \subset \widehat{\omega}_T} \int_{\widehat{T}'} \widehat{\nabla} \boldsymbol{u} : \widehat{\nabla} \boldsymbol{v}, \ \forall (\boldsymbol{u}, \boldsymbol{v}) \in V \times V$$

and the inner product $(\boldsymbol{u}, \boldsymbol{v}) = \displaystyle\sum_{\widehat{T}' \subset \widehat{\omega}_T} \int_{\widehat{T}'} \boldsymbol{u} \cdot \boldsymbol{v}$, for $\boldsymbol{u}$ and $\boldsymbol{v}$ in H.

The corresponding spectral problem consists in finding $\lambda \in \mathbb{R}$ and $\boldsymbol{u} \in V$, $\boldsymbol{u} \neq 0$ solution of

$$l(\boldsymbol{u}, \boldsymbol{v}) = \lambda(\boldsymbol{u}, \boldsymbol{v}), \ \forall \boldsymbol{v} \in V. \tag{1.41}$$

We now define the self-adjoint operator A associated with this problem (1.88) by

$$\begin{array}{rccl} A \ : \ D(A) \subset H & \rightarrow & H \\ \boldsymbol{u} & \rightarrow & A\boldsymbol{u} \end{array}$$

such that

$$\forall \boldsymbol{u} \in D(A), \ \exists \mathbf{f} \in H \ : \ l(\boldsymbol{u}, \boldsymbol{v}) = (\mathbf{f}, \boldsymbol{v}), \ \forall \boldsymbol{v} \in V \text{ and } A\boldsymbol{u} = \boldsymbol{f}.$$

Since V is compactly embedded into H, A has a compact inverse. Therefore this operator admits a discrete spectrum and, by the min-max principle, its first positive eigenvalue satisfies :

$$\lambda_1 = \min_{\substack{\mathbf{v} \in V, \mathbf{v} \neq 0 \\ \mathbf{v} \perp \ker A}} \frac{l(\boldsymbol{v}, \boldsymbol{v})}{\|\boldsymbol{v}\|_H^2}.$$

Since $\ker A = \mathcal{R}_0(\widehat{\omega}_T)$, we deduce that

$$\lambda_1 = \min_{\substack{\widehat{\mathbf{u}} \in V, \mathbf{u} \neq 0 \\ \widehat{\mathbf{u}} \perp \mathcal{R}_0(\widehat{\omega}_T)}} \frac{\displaystyle\sum_{\widehat{T}' \subset \widehat{\omega}_T} \|\widehat{\nabla} \widehat{\boldsymbol{u}}\|_{\widehat{T}'}^2}{\displaystyle\sum_{\widehat{T}' \subset \widehat{\omega}_T} \|\widehat{\boldsymbol{u}} - \widehat{\mathbf{p}}\|_{\widehat{T}'}^2}.$$

This gives, by choosing in (1.87) $\widehat{\mathbf{p}}$ as the projection of $\widehat{\boldsymbol{u}}$ on $\mathcal{R}_0(\widehat{\omega}_T)$ with respect to the inner product $(\cdot, \cdot)$, the following estimate :

$$\|\widehat{\boldsymbol{u}} - \widehat{\mathbf{p}}\|_{\widehat{\omega}_T} \ \lesssim \ \lambda_1^{-1/2} \left(\sum_{\widehat{T}' \subset \widehat{\omega}_T} \|\widehat{\nabla} \widehat{\boldsymbol{u}}\|_{\widehat{T}'}^2 \right)^{1/2}.$$

This implies (1.38) by a scaling argument.

We now prove the third estimate. First as $\mathrm{I}_{\mathrm{CN}}\boldsymbol{u} \in [\mathbb{P}^1(T)]^2$, a standard inverse inequality [17] and the estimate (1.37) yield

$$\|\nabla(\mathrm{I}_{\mathrm{CN}}\boldsymbol{u})\|_T \lesssim h_T^{-1}\|\mathrm{I}_{\mathrm{CN}}\boldsymbol{u}\|_T \lesssim h_T^{-1}\|\boldsymbol{u}\|_{\omega_T}.$$

By the triangular inequality we get

$$\begin{aligned}\|\nabla(\boldsymbol{u}-\mathrm{I}_{\mathrm{CN}}\boldsymbol{u})\|_T &\lesssim \|\nabla\boldsymbol{u}\|_T + \|\nabla(\mathrm{I}_{\mathrm{CN}}\boldsymbol{u})\|_T\\ &\lesssim \|\nabla\boldsymbol{u}\|_T + h_T^{-1}\|\boldsymbol{u}\|_{\omega_T}.\end{aligned}$$

Moreover, as for any $\mathbf{p} \in \mathcal{R}_0(\omega_T)$, $\boldsymbol{u}-\mathrm{I}_{\mathrm{CN}}\boldsymbol{u} = (\mathrm{I}-\mathrm{I}_{\mathrm{CN}})(\boldsymbol{u}-\mathbf{p})$ on T, we find

$$\begin{aligned}\|\nabla(\boldsymbol{u}-\mathrm{I}_{\mathrm{CN}}\boldsymbol{u})\|_T &\lesssim \|\nabla\left[(\mathrm{I}-\mathrm{I}_{\mathrm{CN}})(\boldsymbol{u}-\mathbf{p})\right]\|_T\\ &\lesssim \|\nabla(\boldsymbol{u}-\mathbf{p})\|_T + h_T^{-1}\|\boldsymbol{u}-\mathbf{p}\|_{\omega_T}\\ &\lesssim \|\nabla\boldsymbol{u}\|_T + h_T^{-1}\|\boldsymbol{u}-\mathbf{p}\|_{\omega_T}.\end{aligned}$$

Since we have shown by the min-max principle that

$$\|\boldsymbol{u}-\mathbf{p}\|_{\omega_T} \lesssim h_T\|\nabla\boldsymbol{u}\|_{\omega_T}$$

the conclusion follows. ∎

Remark 1.3.4. Another interpolation operator of Clément-Nédélec type satisfying the commuting diagram property and satisfying the estimates (1.37) to (1.84) was introduced in [48]. Our construction is simpler than in [48], since we do not require the commuting diagram property.

1.3.3 Error estimates

Residual error estimators

On a element T, we define by $\boldsymbol{R}_T := \boldsymbol{f} - (\mathbf{curl}(\chi\,\mathrm{curl}\,\boldsymbol{u}_h) + \beta\boldsymbol{u}_h)$ the exact residual and denote by $\boldsymbol{r}_T$ its approximated residual.
Introduce the jump of $\boldsymbol{u}_h$ in the normal direction and the jump of $\mathrm{curl}\,\boldsymbol{u}_h$ in the tangential direction by

$$\begin{aligned}\mathbf{J}_{e,n} &:= \begin{cases}\left[\!\left[\beta\boldsymbol{u}_h\cdot\boldsymbol{n}_e\right]\!\right]_e & \text{for interior edges}\\ 0 & \text{for boundary edges,}\end{cases}\\ \mathbf{J}_{e,t} &:= \begin{cases}\left[\!\left[\chi\,\mathrm{curl}\,\boldsymbol{u}_h\right]\!\right]_e & \text{for interior edges}\\ 0 & \text{for boundary edges.}\end{cases}\end{aligned}$$

In this section, we build a local error estimator of the solenoidal part of the error inspired from [33], where convection-reaction-diffusion problems are considered.

Definition 1.3.5. *The local and global residual error estimators are defined by*

$$\begin{aligned}
\eta_0^2 &:= \sum_{T\in\mathcal{T}_h} \eta_{T,0}^2,\\
\eta_\perp^2 &:= \sum_{T\in\mathcal{T}_h} \eta_{T,\perp}^2,\\
\eta^2 &:= \eta_0^2+\eta_\perp^2,\\
\zeta^2 &:= \sum_{T\in\mathcal{T}_h} \zeta_T^2,
\end{aligned}$$

$$\begin{aligned}
\eta_{T,0}^2 &:= h_T^2\beta\|\operatorname{div}\boldsymbol{u}_h\|_T^2+\sum_{e\subset\partial T} h_e\beta^{-1}\|\mathbf{J}_{e,n}\|_e^2,\\
\eta_{T,\perp}^2 &:= \alpha_T^2\|\boldsymbol{r}_T\|_T^2+\sum_{e\subset\partial T}\chi^{-\frac12}\alpha_T\|\mathbf{J}_{e,t}\|_e^2,\\
\zeta_T^2 &:= \sum_{T'\subset\omega_T}\alpha_{T'}^2\|\boldsymbol{r}_{T'}-\boldsymbol{R}_{T'}\|_{T'}^2,
\end{aligned}$$

where $\alpha_T := \min\{\beta^{-\frac12},\chi^{-\frac12}h_T\}$.

Proof of the lower error bound : the irrotational part

Theorem 1.3.6. *For all elements T, we have the following local error bound :*

$$\eta_{T,0}\lesssim \|e\|_{\beta,\omega_T}. \tag{1.42}$$

Proof:
$\diamond$ *Divergence*
By the inverse inequality (1.19) and Green's formula,

$$\begin{aligned}
\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T^2 &\sim \int_T b_T(\operatorname{div}(\beta\boldsymbol{u}_h))^2\\
&\sim -\int_T \nabla(b_T\operatorname{div}(\beta\boldsymbol{u}_h))\beta\boldsymbol{u}_h\\
&\sim r(\nabla(b_T\operatorname{div}(\beta\boldsymbol{u}_h))) \quad \text{by (1.14) and (1.4)}\\
&\sim a(\boldsymbol{e},\nabla(b_T\operatorname{div}(\beta\boldsymbol{u}_h))) \quad \text{by (1.13)}\\
&\sim \int_T \beta\boldsymbol{e}\nabla(b_T\operatorname{div}(\beta\boldsymbol{u}_h))\\
&\lesssim \beta^{1/2}\|\nabla(b_T\operatorname{div}(\beta\boldsymbol{u}_h))\|_T\|\beta^{1/2}\boldsymbol{e}\|_T\\
&\lesssim \beta^{1/2}h_T^{-1}\|\operatorname{div}(\beta\boldsymbol{u}_h))\|_T\|\beta^{1/2}\boldsymbol{e}\|_T \quad \text{by (1.20)}
\end{aligned}$$

This shows that

$$\|\operatorname{div}(\beta\boldsymbol{u}_h))\|_T\lesssim\beta^{1/2}h_T^{-1}\|\boldsymbol{e}\|_{\beta,T} \tag{1.43}$$

$\diamond$ *Normal jump*
Let e be an interior edge; we recall that $\mathbf{J}_{e,n} \in \mathbb{P}^k(e)$ with $k \in \mathbb{N}$ depending on the chosen finite element space. Set

$$w_e := F_{ext}(\mathbf{J}_{e,n})\, b_e \ \in \ [H_0^1(\omega_e)]^2.$$

An elementwise partial integration gives

$$\begin{aligned}
\int_e \mathbf{J}_{e,n} w_e &= -\int_e \left[\!\left[\beta(\boldsymbol{u}-\boldsymbol{u}_h)\cdot \boldsymbol{n}_e\right]\!\right]_e w_e \\
&= \pm \sum_{T\subset\omega_e} \left[\int_T \beta \boldsymbol{e}\nabla w_e - \int_T \operatorname{div}(\beta \boldsymbol{e}) w_e\right] \\
&= \pm \sum_{T\subset\omega_e} \left[\int_T \beta \boldsymbol{e}\nabla w_e + \int_T \operatorname{div}(\beta \boldsymbol{u}_h) w_e\right] \\
&\lesssim \sum_{T\subset\omega_e} \left(\|\beta^{1/2}\boldsymbol{e}\|_T \beta^{1/2}\|\nabla w_e\|_T + \|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T \|w_e\|_T\right) \\
&\lesssim \sum_{T\subset\omega_e} \left(\|\boldsymbol{e}\|_{\beta,T}\beta^{1/2}h_T^{-1/2}\|\mathbf{J}_{e,n}\|_e + \|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T h_T^{1/2}\|\mathbf{J}_{e,n}\|_e\right) \qquad \text{by (1.22) and (1.23).}
\end{aligned}$$

Since (1.21) yields $\displaystyle\int_e \mathbf{J}_{e,n} w_e \sim \|\mathbf{J}_{e,n}\|_e^2$, we obtain

$$\|\boldsymbol{J}_{e,n}\|_e \lesssim \sum_{T\subset\omega_e} \left(\beta^{1/2}h_T^{-1/2}\|\boldsymbol{e}\|_{\beta,T} + h_T^{1/2}\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T\right).$$

This estimate coupled with (1.91) implies :

$$\|\mathbf{J}_{e,n}\|_e \lesssim \sum_{T\subset\omega_e} \left(\beta^{1/2}h_T^{-1/2}\|\boldsymbol{e}\|_{\beta,T}\right)$$

As $\mathcal{T}_h$ is regular, $h_T \sim h_e$, we obtain :

$$\|\mathbf{J}_{e,n}\|_e \lesssim \beta^{1/2}h_e^{-1/2}\|\boldsymbol{e}\|_{\beta,\omega_e}. \tag{1.44}$$

The estimates (1.91) and (1.92) lead to the conclusion. ∎

Proof of the lower error bound : the solenoidal part

Theorem 1.3.7. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,\perp} \lesssim \|e\|_{\beta,\chi,\omega_T} + \zeta_T \tag{1.45}$$

Proof:
$\diamond$ *Element residual*
Let T be an element of the triangulation. Set $w_T := \boldsymbol{r}_T b_T \ \in \ [H_0^1(T)]^2$.

By the inverse inequality (1.19), Green's formula and the fact that b_T is zero on the boundary of T, we write

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\sim \int_T \boldsymbol{r}_T w_T \\
&\sim \int_T \boldsymbol{R}_T w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T (\boldsymbol{f} - \mathbf{curl}(\chi \operatorname{curl} \boldsymbol{u}_h) - \beta \boldsymbol{u}_h) w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T (\boldsymbol{f} - \beta \boldsymbol{u}_h) w_T - \int_T \chi \operatorname{curl} \boldsymbol{u}_h \operatorname{curl} w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim r(w_T) + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T.
\end{aligned}$$

The relation (1.13) implies

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\sim a(\boldsymbol{e}, w_T) + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T \chi \operatorname{curl} \boldsymbol{e} \operatorname{curl} w_T + \int_T \beta \boldsymbol{e} w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\lesssim \|\chi^{1/2} \operatorname{curl} \boldsymbol{e}\|_T \chi^{1/2}\| \operatorname{curl} w_T\|_T + \|\beta^{1/2} \boldsymbol{e}\|_T \beta^{1/2} \|w_T\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|w_T\|_T.
\end{aligned}$$

The inverse inequalities (1.19) and (1.20) give

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\lesssim \|\chi^{1/2} \operatorname{curl} \boldsymbol{e}\|_T \chi^{1/2} h_T^{-1} \|\boldsymbol{r}_T\|_T + \|\beta^{1/2} \boldsymbol{e}\|_T \beta^{1/2} \|\boldsymbol{r}_T\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|\boldsymbol{r}_T\|_T \\
&\lesssim \left[\left(\|\chi^{1/2} \operatorname{curl} \boldsymbol{e}\|_T + \|\beta^{1/2} \boldsymbol{e}\|_T \right)^{1/2} \left(\chi^{1/2} h_T^{-1} + \beta^{1/2} \right) + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \right] \|\boldsymbol{r}_T\|_T.
\end{aligned}$$

By the definition of α_T, we obtain

$$\|\boldsymbol{r}_T\|_T \lesssim \alpha_T^{-1} \|\boldsymbol{e}\|_{\beta,\chi,T} + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T. \tag{1.46}$$

$\diamond$ *Tangential jump*
Set $w_e := F_{ext}(\mathbf{J}_{e,t}) b_{e,\gamma_e} \in [H_0^1(\omega_e)]^2$ with $\gamma_e \in (0,1]$. For $\omega_e = T_1 \cup T_2$, b_{e,γ_e} is defined as follow

$$b_{e,\gamma_e} := \begin{cases} b_{e,T_1,\gamma_1} & \text{on } T_1 \\ b_{e,T_2,\gamma_2} & \text{on } T_2 \end{cases}$$

and

$$\gamma_e := \begin{cases} \gamma_1 & \text{on } T_1 \\ \gamma_2 & \text{on } T_2 \end{cases}$$

where we choose (see [33])

$$\gamma_i := \frac{1}{2} \chi^{1/2} h_{T_i}^{-1} \alpha_{T_i} = \frac{1}{2} \min \left\{ 1, \beta^{-1/2} \chi^{1/2} h_{T_i}^{-1} \right\}. \tag{1.47}$$

Note that, $b_{e,T_1,\gamma_{1|e}} = b_{e,T_2,\gamma_{2|e}} = b_{e|e}$. It comes from an elementwise partial integration that

$$\begin{aligned}
\|\mathbf{J}_{e,t}\|_e^2 &\sim \int_e \mathbf{J}_{e,t} w_e \cdot \boldsymbol{t}_e \\
&\sim -\int_e \left[\!\left[\chi \operatorname{curl} \boldsymbol{u}_h\right]\!\right]_e w_e \cdot \boldsymbol{t}_e \\
&\sim \pm \sum_{T_i \subset \omega_e} \left[\int_{T_i} \chi \operatorname{curl} \boldsymbol{u}_h \operatorname{curl} w_e - \int_{T_i} \mathbf{curl}(\chi \operatorname{curl} \boldsymbol{u}_h) w_e \right] \\
&\lesssim r(w_e) + \sum_{T_i \subset \omega_e} \int_{T_i} \boldsymbol{R}_{T_i} w_e \\
&\lesssim a(\boldsymbol{e}, w_e) + \sum_{T_i \subset \omega_e} \int_{T_i} \boldsymbol{r}_{T_i} w_e + \sum_{T_i \subset \omega_e} \int_{T_i} (\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}) w_e \\
&\lesssim \sum_{T_i \subset \omega_e} \left[\int_{T_i} \chi \operatorname{curl} \boldsymbol{e} \operatorname{curl} w_e + \int_{T_i} \beta \boldsymbol{e} w_e + \int_{T_i} \boldsymbol{r}_{T_i} w_e + \int_{T_i} (\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}) w_e \right] \\
&\lesssim \sum_{T_i \subset \omega_e} \left[\|\chi^{1/2} \operatorname{curl} \boldsymbol{e}\|_{T_i} \chi^{1/2} \| \operatorname{curl} w_e\|_{T_i} + \|\beta^{1/2} \boldsymbol{e}\|_{T_i} \beta^{1/2} \| w_e\|_{T_i} \right. \\
&+ \left. \|\boldsymbol{r}_{T_i}\|_{T_i} \|w_e\|_{T_i} + \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \|w_e\|_{T_i} \right]. \\
&\lesssim \sum_{T_i \subset \omega_e} \left[\|\boldsymbol{e}\|_{\beta,\chi,T_i} \left(\chi^{1/2} \| \operatorname{curl} w_e\|_{T_i} + \beta^{1/2} \|w_e\|_{T_i} \right) \right. \\
&+ \left. \left(\|\boldsymbol{r}_{T_i}\|_{T_i} + \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \right) \|w_e\|_{T_i} \right].
\end{aligned}$$

By the discrete Cauchy-Schwarz inequality and the inverse estimates (1.25),(1.26), we find

$$\|w_e\|_{\beta,\chi,T_i} \lesssim \gamma_i^{1/2} h_{T_i}^{1/2} \left(\beta^{1/2} + \chi^{1/2} \gamma_i^{-1} h_{T_i}^{-1} \right) \|\mathbf{J}_{e,t}\|_e, \tag{1.48}$$

and (1.94) and (1.100) lead to

$$\begin{aligned}
\|\mathbf{J}_{e,t}\|_e &\lesssim \sum_{T_i \subset \omega_e} \left[\left(\chi^{1/2} h_{T_i}^{-1/2} \gamma_i^{-1/2} + \beta^{1/2} h_{T_i}^{1/2} \gamma_i^{1/2} + \alpha_{T_i}^{-1} h_{T_i}^{1/2} \gamma_i^{1/2} \right)^{1/2} \|\boldsymbol{e}\|_{\beta,\chi,T_i} \right. \\
&+ \left. h_{T_i}^{1/2} \gamma_i^{1/2} \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \right].
\end{aligned}$$

Then, by (1.99), we obtain

$$\|\mathbf{J}_{e,t}\|_e \lesssim \chi^{1/4} \sum_{T_i \subset \omega_e} \left[\alpha_{T_i}^{-1/2} \|\boldsymbol{e}\|_{\beta,\chi,T_i} + \alpha_{T_i}^{1/2} \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \right]. \tag{1.49}$$

Using (1.94), (1.101) and the definition of $\eta_{T,\perp}$, we get :

$$\begin{aligned}
\eta_{T,\perp} &\lesssim \|\boldsymbol{e}\|_{\beta,\chi,T} + \alpha_T \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \\
&+ \sum_{e \subset \partial T} \chi^{-1/4} \alpha_T^{1/2} \sum_{T_i \subset \omega_e} \chi^{1/4} \alpha_{T_i}^{-1/2} \left(\|\boldsymbol{e}\|_{\beta,\chi,T_i} + \alpha_{T_i} \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \right) \\
&\lesssim \|\boldsymbol{e}\|_{\beta,\chi,T} + \sum_{e \subset \partial T} \sum_{T_i \subset \omega_e} \alpha_T^{1/2} \alpha_{T_i}^{-1/2} \|\boldsymbol{e}\|_{\beta,\chi,T_i} \\
&+ \alpha_T \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T + \sum_{e \subset \partial T} \sum_{T_i \subset \omega_e} \alpha_T^{1/2} \alpha_{T_i}^{-1/2} \alpha_{T_i} \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i}.
\end{aligned}$$

As the triangulation is regular, $h_T \sim h_{T_i}$, we get $\alpha_T \alpha_{T_i}^{-1} \sim 1$. That leads to the conclusion. ■

Corollary 1.3.8. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,0} + \eta_{T,\perp} \lesssim \|\mathbf{e}\|_{\beta,\chi,\omega_T} + \zeta_T \tag{1.50}$$

Proof of the upper error bound : the irrotational part

Theorem 1.3.9. *The β-norm of the irrotational part of the error is globally bounded from above by :*

$$\|\boldsymbol{e}_0\|_\beta \lesssim \eta_0. \tag{1.51}$$

Proof:
Let $\phi \in H_0^1(\Omega)$ be the function introduced in the Helmholtz decomposition such that the irrotational part of the error $\boldsymbol{e}_0 = \nabla\phi$. We are interested in $\|e_0\|_\beta = \|e_0\|_{\beta,\chi} = \|\nabla\phi\|_\beta$. By (1.13), we know that

$$a(\boldsymbol{e}_0, \nabla\psi) = (\beta\nabla\phi, \nabla\psi) = r(\nabla\psi),\ \forall\psi \in H_0^1(\Omega).$$

Let $\psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h)$. Then, $\nabla\psi_h \in V_{h,1} \subset V_h$ and the Galerkin orthogonality relation (1.15) gives

$$(\beta\nabla\phi, \nabla\psi) = r(\nabla(\psi - \psi_h)),\ \forall\psi \in H_0^1(\Omega), \psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h).$$

As $\boldsymbol{f}$ is divergence free and $\psi - \psi_h$ belongs to $H_0^1(\Omega)$, we obtain, by Green's formula and an elementwise integration by parts : $\forall\psi \in H_0^1(\Omega), \psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h)$,

$$(\beta\nabla\phi, \nabla\psi) = \sum_{T\in\mathcal{T}_h} \left(\int_T \operatorname{div}(\beta\boldsymbol{u}_h)(\psi - \psi_h) - \sum_{e\subset\partial T} \int_e \mathbf{J}_{e,n}(\psi - \psi_h) \right).$$

Setting $\phi = \psi$ and using Cauchy-Schwarz's inequality give

$$\begin{aligned}
(\beta\boldsymbol{e}_0, \boldsymbol{e}_0) &= (\beta\nabla\phi, \nabla\phi) \\
&\leq \sum_{T\in\mathcal{T}_h} \left(\int_T \operatorname{div}(\beta\boldsymbol{u}_h)(\phi - \psi_h) - \sum_{e\subset\partial T} \int_e \mathbf{J}_{e,n}(\phi - \psi_h) \right) \\
&\leq \sum_{T\in\mathcal{T}_h} \left[\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T \|\phi - \psi_h\|_T + \sum_{e\subset\partial T} \|\mathbf{J}_{e,n}\|_e \|\phi - \psi_h\|_e \right].
\end{aligned}$$

We now introduce the notations $\mu_T = h_T\beta^{-1/2}$ and $\mu_e = h_e\beta^{-1}$.

By the discrete Cauchy-Schwarz inequality we obtain :

$$\begin{aligned}\|\boldsymbol{e}_0\|_\beta^2 &\lesssim \left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^2\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T^2+\sum_{e\subset\partial T}\mu_e\|\mathbf{J}_{e,n}\|_e^2\right)\right\}^{1/2}\\ &\qquad\cdot\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2+\sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{1/2}\\ &\lesssim \eta_0\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2+\sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{1/2}.\end{aligned}$$

To achieve our estimate (1.103), we choose $\psi_h = \mathrm{I}_{\mathrm{Cl}}\phi \in \mathcal{S}(\Omega,\mathcal{T}_h)$ and apply (1.35),(1.36) to obtain

$$\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2+\sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{1/2}\lesssim\|\nabla\phi\|_\beta, \tag{1.52}$$

noting that, by the trace inequality (1.18), we have

$$\begin{aligned}\sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T}h_e^{-1}\|\phi-\psi_h\|_e^2 &\lesssim \sum_{T\in\mathcal{T}_h}h_T^{-1}\|\phi-\psi_h\|_T\left(h_T^{-1}\|\phi-\psi_h\|_T+\|\nabla(\phi-\psi_h)\|_T\right)\\ &\lesssim \left(\sum_{T\in\mathcal{T}_h}h_T^{-2}\|\phi-\psi_h\|_T^2\right)^{1/2}\\ &\qquad\cdot\left(\sum_{T\in\mathcal{T}_h}\left(h_T^{-2}\|\phi-\psi_h\|_T^2+\|\nabla(\phi-\psi_h)\|_T^2\right)\right)^{1/2}.\end{aligned}$$

■

Proof of the upper error bound : the solenoidal part

Theorem 1.3.10. *The $\beta-\chi$-norm of the solenoidal part of the error is globally bounded from above by*

$$\|\boldsymbol{e}_\perp\|_{\beta,\chi}\lesssim\eta+\zeta.$$

Proof: As $\boldsymbol{e}_\perp \in \mathcal{W}\subset H_0(\operatorname{curl},\Omega)$,

$$\|\boldsymbol{e}_\perp\|_{\beta,\chi}^2\le a(\boldsymbol{e}_\perp,\boldsymbol{e}_\perp)=r(\boldsymbol{e}_\perp)=r(\boldsymbol{w})+r(\nabla\psi), \tag{1.53}$$

according to the decomposition (1.32) of $\boldsymbol{e}_\perp$. Using the Galerkin orthogonality relation

(1.15) for any $\boldsymbol{v}_h \in V_h$ and an elementwise integration by parts, we get

$$\begin{aligned} r(\boldsymbol{w}) &= r(\boldsymbol{w}-\boldsymbol{v}_h) \\ &= (\boldsymbol{f}, \boldsymbol{w}-\boldsymbol{v}_h) - a(\boldsymbol{u}_h, \boldsymbol{w}-\boldsymbol{v}_h) \\ &= (\boldsymbol{f}-\beta\boldsymbol{u}_h, \boldsymbol{w}-\boldsymbol{v}_h) \\ &- \sum_{T\in\mathcal{T}_h}\left(\int_T \mathbf{curl}(\chi \operatorname{curl}\boldsymbol{u}_h)(\boldsymbol{w}-\boldsymbol{v}_h) - \sum_{e\subset\partial T}\int_e \chi \operatorname{curl}\boldsymbol{u}_h(\boldsymbol{w}-\boldsymbol{v}_h)\cdot\boldsymbol{t}_e\right) \\ &= \sum_{T\in\mathcal{T}_h}(\boldsymbol{R}_T, \boldsymbol{w}-\boldsymbol{v}_h) - \sum_{e\in\mathcal{E}_h}\int_e \mathbf{J}_{e,t}(\boldsymbol{w}-\boldsymbol{v}_h)\cdot\boldsymbol{t}_e. \end{aligned}$$

Cauchy-Schwarz's inequality leads to

$$\begin{aligned} r(\boldsymbol{w}) &\lesssim \sum_{T\in\mathcal{T}_h}\left[\alpha_T\|\boldsymbol{R}_T\|_T\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T + \sum_{e\subset\partial T}\chi^{-1/4}\alpha_T^{1/2}\|\mathbf{J}_{e,t}\|_e\chi^{1/4}\alpha_T^{-1/2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e\right] \\ &\lesssim \left\{\sum_{T\in\mathcal{T}_h}\left[\alpha_T^2\|\boldsymbol{r}_T\|_T^2 + \alpha_T^2\|\boldsymbol{r}_T-\boldsymbol{R}_T\|_T^2 + \sum_{e\subset\partial T}\chi^{-1/2}\alpha_T\|\mathbf{J}_{e,t}\|_e^2\right]\right\}^{1/2} \\ &\quad\cdot\left\{\sum_{T\in\mathcal{T}_h}\left[\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T}\chi^{1/2}\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2\right]\right\}^{1/2}. \end{aligned}$$

Then, by taking $\boldsymbol{v}_h = \mathrm{I}_{\mathrm{CN}}\boldsymbol{w} \in V_h$, we can prove that :

$$\sum_{T\in\mathcal{T}_h}\left[\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T}\chi^{1/2}\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2\right] \lesssim \|\boldsymbol{w}\|_\beta^2 + \chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2. \tag{1.54}$$

Indeed, the definition of α_T implies $\alpha_T^{-1} = \max\{\beta^{1/2}, \chi^{1/2}h_T^{-1}\}$. It follows, by the estimates (1.37)-(1.38) and the triangular inequality, that

$$\begin{aligned} \sum_{T\in\mathcal{T}_h}\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 &= \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\geq\chi^{1/2}h_T^{-1}}}\beta\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 + \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\leq\chi^{1/2}h_T^{-1}}}\chi h_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 \\ &\lesssim \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\geq\chi^{1/2}h_T^{-1}}}\left(\|\boldsymbol{w}\|_{\beta,T}^2 + \beta\|\boldsymbol{v}_h\|_T^2\right) + \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\leq\chi^{1/2}h_T^{-1}}}\chi\|\nabla\boldsymbol{w}\|_{\omega_T}^2 \\ &\lesssim \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\geq\chi^{1/2}h_T^{-1}}}\left(\|\boldsymbol{w}\|_{\beta,T}^2 + \|\boldsymbol{w}\|_{\beta,\omega_T}^2\right) + \sum_{\substack{T\in\mathcal{T}_h \\ \beta^{1/2}\leq\chi^{1/2}h_T^{-1}}}\chi\beta^{-1}\|\nabla\boldsymbol{w}\|_{\beta,\omega_T}^2 \\ &\lesssim \|\boldsymbol{w}\|_\beta^2 + \chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2. \end{aligned} \tag{1.55}$$

On the other hand, by the trace inequality (1.18) and by the estimates (1.113) and (1.84),

we find

$$\begin{aligned}
\sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T}\chi^{1/2}\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2 &\lesssim \chi^{1/2}\sum_{T\in\mathcal{T}_h}\left[\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T\right.\\
&\quad\left.\cdot\left(h_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T+\|\nabla(\boldsymbol{w}-\boldsymbol{v}_h)\|_T\right)\right]\\
&\lesssim \chi^{1/2}\left(\sum_{T\in\mathcal{T}_h}\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2\right)^{1/2}\\
&\quad\cdot\left(\sum_{T\in\mathcal{T}_h}\left(h_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2+\|\nabla(\boldsymbol{w}-\boldsymbol{v}_h)\|_T^2\right)\right)\\
&\lesssim \chi^{1/2}\left(\|\boldsymbol{w}\|_\beta^2+\chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2\right)^{1/2}\left(\sum_{T\in\mathcal{T}_h}\|\nabla\boldsymbol{w}\|_{\omega_T}^2\right)^{1/2}\\
&\lesssim \chi^{1/2}\left(\|\boldsymbol{w}\|_\beta^2+\chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2\right)^{1/2}\|\nabla\boldsymbol{w}\|\\
&\lesssim \|\boldsymbol{w}\|_\beta^2+\chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2. \qquad (1.56)
\end{aligned}$$

The estimates (1.113) and (1.114) show (1.112). Therefore, from the definitions of $\eta_{T,\perp}$ and ζ_T and the estimate (1.112), we deduce

$$r(\mathbf{w})\lesssim(\eta_\perp+\zeta)\left(\|\boldsymbol{w}\|_\beta^2+\chi\beta^{-1}\|\nabla\boldsymbol{w}\|_\beta^2\right)^{1/2}. \qquad (1.57)$$

Using the bounds (1.33) and (1.34) from the Helmholtz decomposition, we get

$$r(\mathbf{w})\lesssim(\eta_\perp+\zeta)\|\boldsymbol{e}_\perp\|_{\beta,\chi}. \qquad (1.58)$$

On the other hand the arguments of Theorem 1.3.9 yield

$$r(\nabla\psi)\lesssim\eta_0\|\nabla\psi\|_\beta.$$

Using the decomposition (1.32) and the estimate (1.34), we deduce that

$$r(\nabla\psi)\lesssim\eta_0\|\boldsymbol{e}_\perp\|_{\beta,\chi}. \qquad (1.59)$$

The conclusion follows from the estimates (1.53), (1.58) and (1.59). ∎

Corollary 1.3.11. *The error is globally bounded from above by*

$$\|\boldsymbol{e}\|_{\beta,\chi}\lesssim\eta+\zeta. \qquad (1.60)$$

1.3.4 Extension to three-dimensional polyhedral domains

All the results of this paper extend to a three-dimensional polyhedral domain Ω which is bounded and simply connected with a connected boundary Γ. In that domain we consider the Maxwell system

$$\begin{cases}\mathbf{curl}(\chi\,\mathbf{curl}\,\boldsymbol{u})+\beta\boldsymbol{u}=\boldsymbol{f} & \text{in }\Omega,\\ \boldsymbol{u}\times\boldsymbol{n}=0 & \text{on }\Gamma,\end{cases} \qquad (1.61)$$

where $\boldsymbol{f}$ satisfies (1.4) and β and χ are as before.

This problem is then approximated using regular meshes made of tetrahedra and the finite element space V_h is simply assumed to contain lowest order Nédélec elements.

In this setting all the results from section 1.2.2 remain valid, especially Lemma 2.4.1 (the Helmholtz decomposition) due to the results from [45, 47]. Moreover in 3D the Clément-Nédélec interpolant is defined by

$$\begin{array}{rcl} \mathrm{I}_{\mathrm{CN}} : L^2(\Omega)^3 & \to & V_h \\ \boldsymbol{u} & \to & \displaystyle\sum_{e \in \mathcal{E}_{h\Omega}} \alpha_e(\boldsymbol{u}) |e| \lambda_e \end{array}$$

where, as usual $\mathcal{E}_{h\Omega}$ is the set of interior edges of the mesh, λ_e is the standard basis function of lowest order Nédélec elements and we here set $\alpha_e(\boldsymbol{u}) = \dfrac{1}{|\omega_e|} \displaystyle\int_{\omega_e} \boldsymbol{u} \cdot \boldsymbol{t}_e$, when ω_e is made of all tetrahedra having e as edge. The regularity of the mesh allows then to show that Theorem 1.4.9 holds.

As the basic tools of section 1.2.2, the interpolation error estimates from section 1.4.2 and some integrations by parts are the only ingredients that we used for the proof of the lower and upper error bounds, we can conclude that the estimates (1.102) and (1.110) hold in 3D, with the same definition for the local estimators, except that $\mathbf{J}_{e,n}$ and $\mathbf{J}_{e,t}$ are defined for the faces F of the mesh and for the tangential jump where curl $\boldsymbol{u}_h$ is replaced by $\mathbf{curl}\, \boldsymbol{u}_h \times \boldsymbol{n}_F$, see section 4.1 of [45].

1.3.5 Numerical experiments

The following experiments underline and confirm our theoretical predictions. Our examples consist in solving the Maxwell equation (1.8) on the unit square $\Omega = (0,1)^2$ with different values of χ and β and different solutions. In all examples uniform meshes and the lowest order Nédélec finite elements are used.

As first example we consider the exact solution :

$$\boldsymbol{u} = \begin{pmatrix} e^{-y/\sqrt{\varepsilon}} y(1-y) \\ e^{-x/\sqrt{\varepsilon}} x(1-x) \end{pmatrix},$$

fix $\beta = 1$ and take $\chi = \varepsilon$, for different values of ε. Note that for small ε, the gradient of this solution presents exponential boundary layers of width $O(\sqrt{\varepsilon})$ along the lines $x = 0$ and $y = 0$.

To begin, we check that the numerical solution $\boldsymbol{u}_h$ converges toward the exact solution for differents values of ε. To this end, we plot the curve $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi}$ as a function of DoF in Figure 1.4. There a double logarithmic scale is used such that the slope of the curves corresponds to the approximation order. As we can see the convergence rate is of order 1 as theoretically expected.

Now we analyze the upper and lower error bounds. In order to present them in an appropriate manner, we consider the ratios

$$\begin{aligned} q_{up} &= \frac{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\beta,\chi}}{\eta+\xi}, \\ q_{low} &= \max_{T\in\mathcal{T}_h} \frac{\eta_{T,0}+\eta_{T,\perp}}{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\beta,\chi,\omega_T}+\zeta_T}, \end{aligned}$$

as a function of DoF. The first ratio q_{up}, the so-called effectivity index, is related to the global upper error bound and measures the reliability of the estimator. The second ratio is related to the local lower error bound and measures the efficiency of the estimator.

These ratios are presented in Figure 1.5 and 1.6 for different values of ε. There we see that q_{up} decreases in function of ε and is bounded by 0.12. Similarly we remark that q_{low} increases in function of ε and is bounded by 6.73.

As second example we take the exact solution

$$\boldsymbol{u} = \nabla\left(e^{-x/\sqrt{\varepsilon}}x(1-x)y(1-y)\right)$$

fix once more $\beta = 1$ and take $\chi = \varepsilon$ for different values of ε. Here the solution presents an exponential boundary layer of width $O(\sqrt{\varepsilon})$ along the line $x = 0$.

As before we see from Figure 1.7 that the numerical solution $\boldsymbol{u}_h$ converges toward $\boldsymbol{u}$ with a convergence rate of order 1. For this example, we see in Figure 1.8 that the effectivity index is bounded by 0.22, while Figure 1.9 indicates that the ratio q_{low} is bounded by 4 as soon as a reasonable resolution of the layer is achieved.

For the last test, we consider the exact solution :

$$\boldsymbol{u} = \begin{pmatrix} y(1-y) \\ e^{-x/\sqrt{\varepsilon}}x(1-x) \end{pmatrix}$$

where we fix $\chi = 1$, $\varepsilon = 0.001$ and take different values of β. In this case, we see in Figures 1.10 to 1.12 that the convergence rate is 1, that the effectivity index remains bounded by 0.16, and that q_{low} is bounded by 5.8.

Note finally that other examples are tested and give rise to ratios q_{up} and q_{low} that are uniformly bounded with respect to different parameters β and χ.

1.3.6 Conclusion

We have proposed and rigorously analysed a robust a posteriori error estimator of residual type for the Maxwell equations in a bounded two (and three) dimensional domain using conforming finite element spaces of Nédélec type. A new interpolant of Clément/Nédélec type has been introduced and some interpolation error estimates have been proved. We have shown that this estimator is reliable and efficient. Some numerical experiments confirm our theoretical predictions.

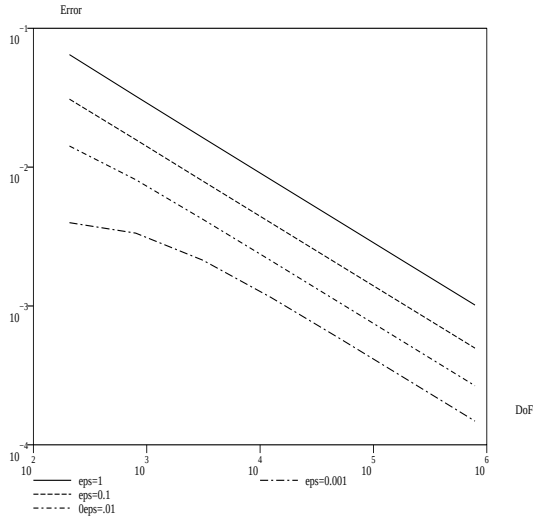


FIG. 1.4 – The error norm $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi}$ as a function of DoF for example 1

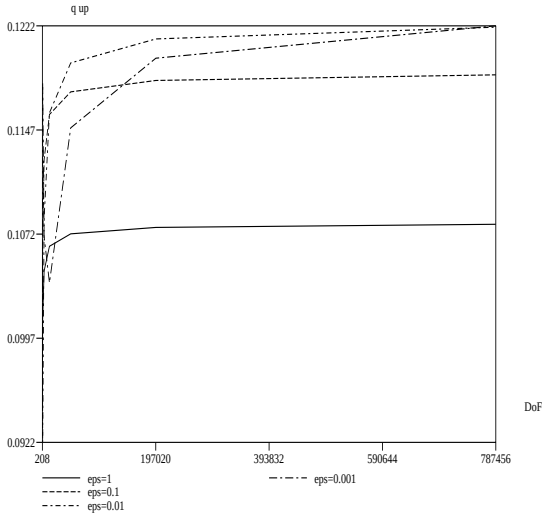


FIG. 1.5 – q_{up} wrt DoF for example 1

1.4 Uniform a posteriori error estimation for the Maxwell equations with discontinuous coefficients

We consider residual based a posteriori error estimators for the heterogeneous Maxwell equations with discontinuous coefficients in bounded two and three dimensional domains.

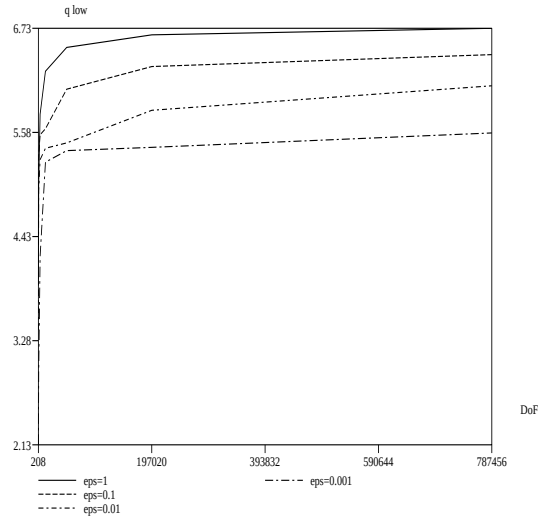


FIG. 1.6 – q_{low} wrt DoF for example 1

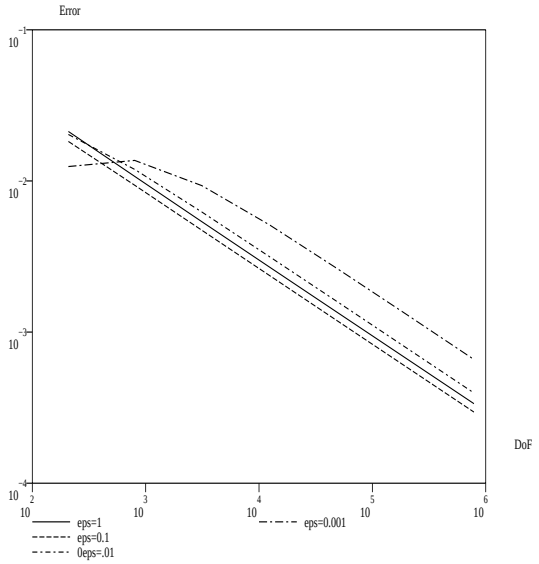


FIG. 1.7 – The error norm $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi}$ as a function of DoF for example 2

The continuous problem is approximated using conforming approximated spaces. The main goal is to express the dependence of the constants in the lower and upper bounds with respect to a chosen norm and to the variation of the coefficients. For that purpose, some new interpolation operators of Clément/Nédélec type are introduced and some interpolation error estimates are proved. Some numerical tests are presented which confirm our theoretical

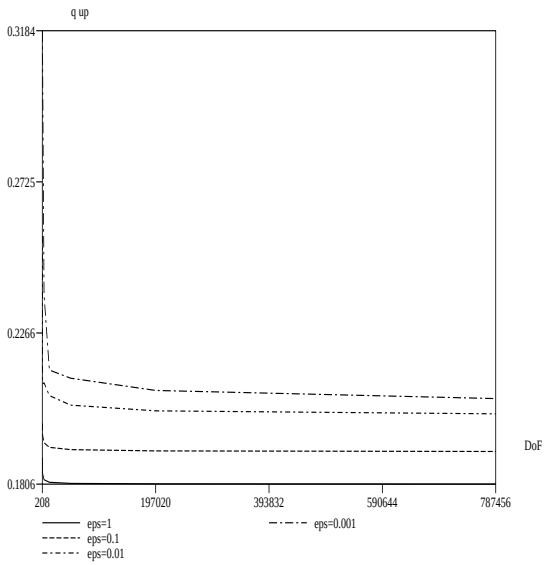


FIG. 1.8 – q_{up} wrt DoF for example 2

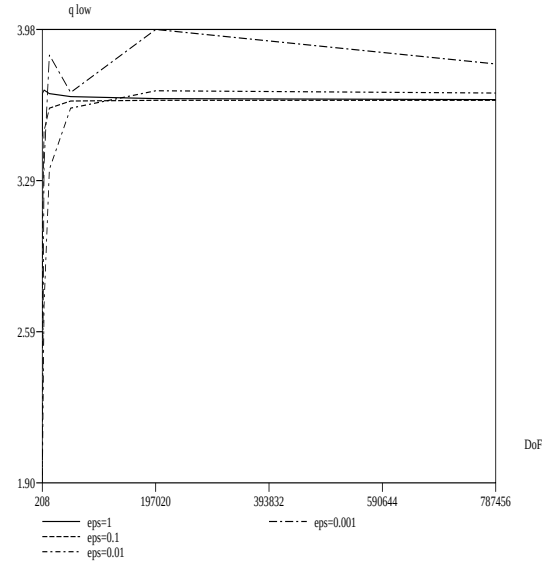


FIG. 1.9 – q_{low} wrt DoF for example 2

results.

The schedule of the section is the following one : Section 1.2.1 recalls the discretization of our problem. In section 1.4.1, we state the adapted Helmholtz decomposition of the error. Some basic tools for the error estimation analysis are recalled in section 1.2.2. In section 1.4.2 we give some interpolation error estimates for Clément and Nédélec interpolants,

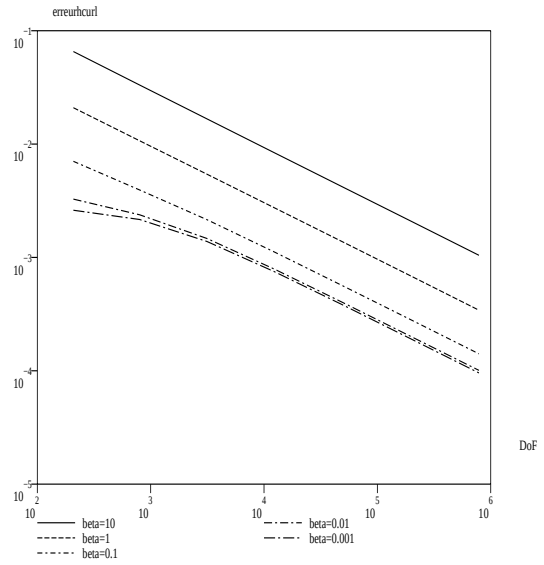


FIG. 1.10 – The error norm $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi}$ as a function of DoF for example 3

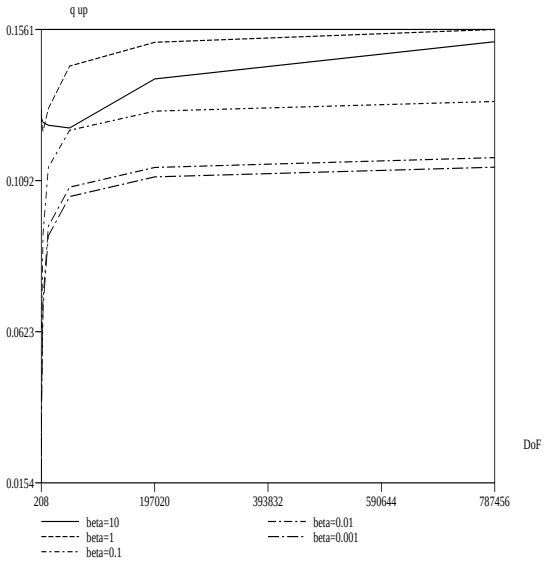


FIG. 1.11 – q_{up} wrt DoF for example 3

introduce a new interpolation operator of Clément-Nédélec type and prove suitable error estimates. The efficiency and reliability of two different estimators are established in section 1.4.3. The extension of our results to three-dimensional problems is shortly described in section 1.4.4. Finally section 1.4.5 is devoted to some numerical tests which confirm our theoretical analysis.

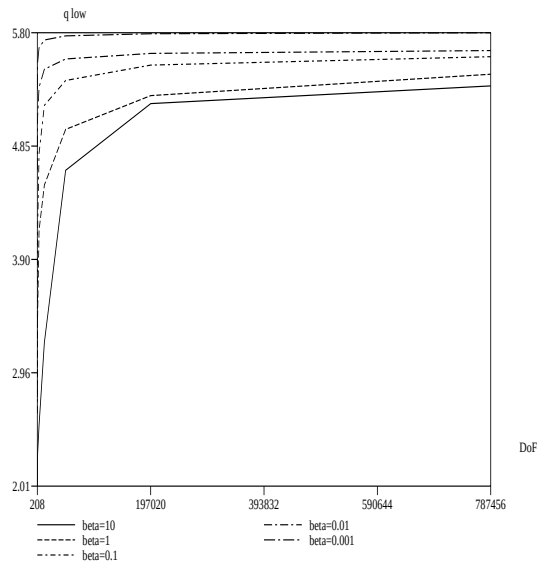


FIG. 1.12 – q_{low} wrt DoF for example 3

1.4.1 Helmholtz Decomposition

We first recall a decomposition of the space $H_0(\mathrm{curl},\Omega)$ of Helmholtz type related to the weight β (namely for $\beta = 1$ the next result is simply the standard Helmholtz decomposition).

Recall that $H_0(\mathrm{curl},\Omega)$ was equipped with the inner product

$$(\boldsymbol{v},\boldsymbol{w})_{\beta,\chi} = (\beta\boldsymbol{v},\boldsymbol{w}) + (\chi\,\mathrm{curl}\,\boldsymbol{v},\mathrm{curl}\,\boldsymbol{w}),$$

its associated norm $\|\boldsymbol{v}\|_{\beta,\chi}$ being equivalent to the usual norm $(\|\boldsymbol{v}\|^2 + \|\,\mathrm{curl}\,\boldsymbol{v}\|^2)^{1/2}$.

Lemma 1.4.1. *If Ω is simply connected and its boundary Γ is connected then*

$$H_0(\mathrm{curl},\Omega) = H_0^0(\mathrm{curl},\Omega) \overset{\perp}{\oplus} W_\beta, \tag{1.62}$$

where $H_0^0(\mathrm{curl},\Omega)$ and W_β are closed subspaces of $H_0(\mathrm{curl},\Omega)$ defined by

$$H_0^0(\mathrm{curl},\Omega) = \{\boldsymbol{v}\in H_0(\mathrm{curl},\Omega) : \mathrm{curl}\,\boldsymbol{v} = 0 \text{ in } \Omega\}, \tag{1.63}$$

$$W_\beta = \{\boldsymbol{v}\in H_0(\mathrm{curl},\Omega) : \mathrm{div}\,(\beta\boldsymbol{v}) = 0 \text{ in } \Omega\}, \tag{1.64}$$

and the symbol $\overset{\perp}{\oplus}$ means that the decomposition is direct and orthogonal with respect to the inner product $(\cdot,\cdot)_{\beta,\chi}$. Furthermore one has

$$H_0^0(\mathrm{curl},\Omega) = \nabla H_0^1(\Omega). \tag{1.65}$$

Proof: Lemma I.2.1 of [26] yield (1.65). It then remains to prove the Helmholtz decomposition (2.23). With the inner product $(\cdot,\cdot)_{\beta,\chi}$, the decomposition (2.23) holds with

$$W_\beta = \{\boldsymbol{v} \in H_0(\mathrm{curl},\Omega) : (\beta\boldsymbol{v},\boldsymbol{w}) + (\chi\,\mathrm{curl}\,\boldsymbol{v},\mathrm{curl}\,\boldsymbol{w}) = 0, \forall \boldsymbol{w} \in H_0^0(\mathrm{curl},\Omega)\}.$$

According to (1.65) this is equivalent to

$$W_\beta = \{\boldsymbol{v} \in H_0(\mathrm{curl},\Omega) : (\beta\boldsymbol{v},\nabla\psi) = 0, \forall \psi \in H_0^1(\Omega)\}.$$

By Green's formula we deduce (2.24). ∎

For our next purposes, we need to decompose any element $\boldsymbol{v}$ from W_β into a singular part $\boldsymbol{v}_S$ and a regular part $\boldsymbol{v}_R$ in the space $H_N(\Omega,\beta)$ defined by

$$H_N(\Omega,\beta) = \{\boldsymbol{w} \in H_0(\mathrm{curl},\Omega) \cap [PH^1(\Omega)]^2 : \mathrm{div}(\beta\boldsymbol{w}) \in L^2(\Omega)\},$$

and equipped with the norm $\|\cdot\|_{PH1,\beta}$. This decomposition is now well known [5, 21, 22], and is obtained by looking at W_β as a (closed) subspace of

$$X_N(\Omega,\beta) = \{\boldsymbol{w} \in H_0(\mathrm{curl},\Omega) : \mathrm{div}(\beta\boldsymbol{w}) \in L^2(\Omega)\},$$

equipped with the norm $\|\cdot\|_{X_N,\beta,\chi}$ defined by

$$\|\boldsymbol{v}\|^2_{X_N,\beta,\chi} = \int_\Omega (\chi|\,\mathrm{curl}\,\boldsymbol{v}|^2 + \beta^{-1}|\,\mathrm{div}(\beta\boldsymbol{v})|^2 + \beta|\boldsymbol{v}|^2).$$

By functional analysis arguments, there exists a positive constant C such that

$$\|\boldsymbol{v}_R\|_{PH1,\beta} + \|\boldsymbol{v}_S\|_{X_N,\beta,\chi} \leq C\|\boldsymbol{v}\|_{X_N,\beta,\chi}.$$

But for our next purposes, we would need to specify the dependence of C with respect to the coefficients β and χ. Unfortunately this dependence is difficult to establish. We therefore use a Helmholtz decomposition from [47] obtained for $\beta = 1$:

Lemma 1.4.2. *Any $\boldsymbol{v} \in W_\beta$ admits the splitting*

$$\boldsymbol{v} = \boldsymbol{w} + \nabla\phi_0, \tag{1.66}$$

with $\boldsymbol{w} \in (H_0^1(\Omega))^2$, $\phi_0 \in H_0^1(\Omega)$ with the estimate

$$\|\boldsymbol{w}\|_\beta + \|\nabla\phi_0\|_{PH1,\beta} \lesssim C_1(\beta,\chi)\|\boldsymbol{v}\|_{\beta,\chi}, \tag{1.67}$$

$$|\boldsymbol{w}|_{PH1,\beta} \lesssim C_2(\beta,\chi)\|\boldsymbol{v}\|_{\beta,\chi}. \tag{1.68}$$

where

$$C_1(\beta,\chi) = \max_{i=1,\cdots,J}\beta_i^{1/2}\max_{j=1,\cdots,J}\beta_j^{-1/2},$$

$$C_2(\beta,\chi) = \max_{i=1,\cdots,J}\beta_i^{1/2}\max_{j=1,\cdots,J}\chi_j^{-1/2}.$$

Proof: According to [47], the splitting (1.66) holds with $\boldsymbol{w} \in (H_0^1(\Omega))^2$, $\phi_0 \in H_0^1(\Omega)$ with the estimate

$$\begin{aligned} \|\boldsymbol{w}\| + \|\nabla\phi_0\| &\lesssim \|\boldsymbol{v}\|, \\ |\boldsymbol{w}|_{1,\Omega} &\lesssim \|\operatorname{curl}\boldsymbol{v}\|. \end{aligned}$$

The requested estimates follow directly from the definition of the norm $\|\cdot\|_{\beta,\chi}$. ∎

Note that the decomposition (1.66) (and therefore the estimates (1.67) and (1.68)) is not unique in general, and, in some particular cases, the estimates could be improved. The two advantages of the presented results are that they do not depend on the singularities of $\boldsymbol{v} \in W_\beta$ and that the involved constants are explicit. We further see that if β and χ are constant on the whole Ω, then the constants reduce to $C_1 = 1$ and $C_2 = \beta\chi^{-1}$, which are the optimal ones.

Corollary 1.4.3. *The error $\boldsymbol{e}$ admits the splitting*

$$\boldsymbol{e} = \boldsymbol{e}_0 + \boldsymbol{e}_\perp,$$

with $\boldsymbol{e}_0 = \nabla\phi$ where $\phi \in H_0^1(\Omega)$ and $\boldsymbol{e}_\perp \in W_\beta$ which admits the decomposition

$$\boldsymbol{e}_\perp = \boldsymbol{w} + \nabla\phi_0, \tag{1.69}$$

with $\boldsymbol{w} \in (H_0^1(\Omega))^2$, $\phi_0 \in H_0^1(\Omega)$ with the estimate

$$\|\boldsymbol{w}\|_\beta + \|\phi_0\|_{PH1,\beta} \lesssim C_1(\beta,\chi)\|\boldsymbol{e}_\perp\|_{\beta,\chi}, \tag{1.70}$$

$$|\boldsymbol{w}|_{PH1,\beta} \lesssim C_2(\beta,\chi)\|\boldsymbol{e}_\perp\|_{\beta,\chi}. \tag{1.71}$$

with $C_i(\beta,\chi), i = 1,2$ as in Lemma 1.4.2. Furthermore the defect equation is equivalent to the two above equations :

$$(\beta\nabla\phi, \nabla\psi) = r(\nabla\psi), \forall\psi \in H_0^1(\Omega), \tag{1.72}$$

$$a(\boldsymbol{e}_\perp, \boldsymbol{w}) = r(\boldsymbol{w}), \forall\boldsymbol{w} \in W_\beta. \tag{1.73}$$

Proof: Direct consequence of the above Lemma recalling that the decomposition (2.23) is orthogonal with respect to the inner product $(\cdot,\cdot)_{\beta,\chi}$. ∎

1.4.2 Interpolation error estimates

Clément interpolation

Let us first modify the standard Clément interpolant as in [11]. With each vertex x, we associate a number $l(x)$ in $\{1,\ldots,J\}$ such that :
$\star$ x is contained in $\bar{\Omega}_{l(x)}$,

$\star$ $\beta_{l(x)} = \max\{\beta_j, 1 \leq j \leq J : x \ \in \bar{\Omega}_j\}$.
Then we define the Clément type interpolant as follow :

$$\begin{aligned} \mathrm{I}_{\mathrm{Cl}} : \ H_0^1(\Omega) &\rightarrow S(\Omega, \mathcal{T}_h) \\ \phi &\rightarrow \sum_{x \ \in \mathcal{N}_\Omega} \frac{1}{|\omega_x \cap \Omega_{l(x)}|} \left(\int_{\omega_x \cap \Omega_{l(x)}} \phi \right) \varphi_x = \sum_{x \ \in \mathcal{N}_\Omega} \mathrm{I}_{\mathrm{Cl,x}}(\phi) \varphi_x \end{aligned}$$

where $S(\Omega, \mathcal{T}_h)$ is the space of continuous piecewise linear functions on the triangulation which are zero on the boundary and φ_x is the nodal basis function associated with the node x, uniquely determined by the condition :

$$\varphi_x(y) = \delta_{x,y}, \ \forall y \ \in \mathcal{N}_\Omega.$$

Under the geometric assumptions that at most 3 subdomains $\bar{\Omega}_j$ share a common point, it has been shown in [11] the following estimates :
For every function $\phi \ \in H_0^1(\Omega)$, every element T and every edge e of T,

$$\begin{aligned} \|\phi - \mathrm{I}_{\mathrm{Cl}}\phi\|_{L^2(T)} &\lesssim h_T \beta_T^{-\frac{1}{2}} \|\nabla\phi\|_{\beta,\Delta_T}, \\ \|\phi - \mathrm{I}_{\mathrm{Cl}}\phi\|_{L^2(e)} &\lesssim h_e^{\frac{1}{2}} \beta_e^{-\frac{1}{2}} \|\nabla\phi\|_{\beta,\Delta_e}, \end{aligned}$$

where $\beta_e = \max_{T \subset \omega_e} \beta_T$. Here, Δ_T (resp. Δ_e) denotes the union of all elements sharing at least one vertex with T (resp. e).

In order to remove the above geometric assumptions, we introduce another interpolant based on a weighted average and defined by :

$$\mathrm{I}_{\mathrm{Cl}}^{new}\phi = \sum_{x \in \mathcal{N}_\Omega} (\mathcal{M}_x\phi)\, \varphi_x, \ \forall \ \phi \in H_0^1(\Omega), \tag{1.74}$$

where $\mathcal{M}_x\phi = \dfrac{\displaystyle\sum_{T \subset \omega_x} \beta_T \frac{1}{|T|} \int_T \phi}{\displaystyle\sum_{T \subset \omega_x} \beta_T}$.

We start with the following lemma :

Lemma 1.4.4. *For every function $\phi \ \in H_0^1(\Omega)$, every node $x \in \mathcal{N}_\Omega$, one has*

$$\|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x} \ \lesssim \ C_{x,Neu}(\beta) h_x \|\nabla\phi\|_{\beta,\omega_x} \tag{1.75}$$

where $h_x = \max_{T \subset \omega_x} h_T$, $C_{x,Neu}(\beta)$ is the Poincaré constant corresponding to the converse of the squareroot of the first positive eigenvalue of the following Neumann problem :

$$\begin{cases} -\operatorname{div}(\beta\nabla\phi) = \lambda\beta\phi & \text{in } \widehat{\omega}_x, \\ \dfrac{\partial\phi}{\partial n} = 0 & \text{on } \partial\widehat{\omega}_x. \end{cases}$$

Proof: Let x be an arbitrary node of the triangulation $\mathcal{T}_h$ and ϕ a function in $H_0^1(\Omega)$. We want to estimate the constant c_1 in

$$\|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x} \leq c_1 h_x \|\nabla\phi\|_{\beta,\omega_x}.$$

This will be done by the min-max principle. This problem is equivalent to bound from below the ratio

$$\frac{\displaystyle\sum_{T\subset\omega_x} \beta_T h_T^2 \|\nabla\phi\|_T^2}{\displaystyle\sum_{T\subset\omega_x} \beta_T \|\phi - \mathcal{M}_x\phi\|_T^2}.$$

We denote by λ_1 the first eigenvalue of the operator Δ_β with Neumann boundary conditions in $\widehat{\omega}_x$. Then, by the min-max principle, one has

$$\lambda_1 = \min_{\widehat{u}\neq 0, \widehat{u}\perp_\beta 1} \frac{\displaystyle\int_{\widehat{\omega}_x} \beta |\nabla\widehat{u}|^2}{\displaystyle\int_{\widehat{\omega}_x} \beta |\widehat{u}|^2},$$

where $\widehat{u} \perp_\beta 1$ means that $\displaystyle\sum_{\widehat{T}\subset\widehat{\omega}_x} \int_{\widehat{T}} \beta_T \widehat{u} = 0$.

If we set $\phi = \widehat{\phi}\circ F_x$, where F_x is the transformation that maps $\widehat{\omega}_x$ onto ω_x we have

$$\begin{aligned}\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\nabla}\widehat{\phi}\|_{\widehat{T}}^2 &= \sum_{T\subset\omega_x} \beta_T \int_T \|B_T^t \nabla\phi\|_T^2 |T|^{-1}\, dx \\ &\lesssim \sum_{T\subset\omega_x} \beta_T \|B_T^t\|^2 |T|^{-1} \|\nabla\phi\|_T^2.\end{aligned}$$

As $\mathcal{T}_h$ is regular, $|T| \sim h_T^2$ and $\|B_T^t\|^2 \lesssim h_T^2$, so :

$$\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\nabla}\widehat{\phi}\|_{\widehat{T}}^2 \lesssim \sum_{T\subset\omega_x} \beta_T \|\nabla\phi\|_T^2.$$

Moreover, we remark that the weighted average is preserved by these transformations :

$$\widehat{\mathcal{M}}\widehat{\phi} = \frac{\displaystyle\sum_{\widehat{T}\subset\widehat{\omega}} \beta_T \int_{\widehat{T}} \widehat{\phi}}{\displaystyle\sum_{\widehat{T}\subset\widehat{\omega}} \beta_T |\widehat{T}|} = \frac{\displaystyle\sum_{T\subset\omega_x} \beta_T \frac{|\widehat{T}|}{|T|} \int_T \phi}{\displaystyle\sum_{T\subset\omega_x} \beta_T |\widehat{T}|} = \mathcal{M}_x\phi. \tag{1.76}$$

In a similar manner, we have :

$$\begin{aligned}\sum_{T\subset\omega_x} \beta_T \|\phi - \mathcal{M}_x\phi\|_T^2 &= \sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} |\widehat{\phi} - \widehat{\mathcal{M}}\widehat{\phi}|_{\widehat{T}}^2 |T| \\ &\lesssim \sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T h_T^2 \|\widehat{\phi} - \widehat{\mathcal{M}}\widehat{\phi}\|_{\widehat{T}}^2.\end{aligned}$$

This means that we are reduced to bound $\min_{\widehat{\phi}\in H^1(\widehat{\omega}_x)} \dfrac{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\nabla}\widehat{\phi}\|^2_{\widehat{T}}}{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\phi}-\widehat{\mathcal{M}}\widehat{\phi}\|^2_{\widehat{T}}}$.

Now, setting $\widehat{u}=\widehat{\phi}-\widehat{\mathcal{M}}\widehat{\phi}$, the condition $\widehat{u}\perp_\beta 1$ means :

$$\begin{aligned}
\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} \left(\widehat{\phi}-\widehat{\mathcal{M}}\widehat{\phi}\right)\cdot 1 = 0 \quad &\Leftrightarrow \quad \sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} \widehat{\phi} = \sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} \widehat{\mathcal{M}}\widehat{\phi} \\
&\Leftrightarrow \quad \sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} \widehat{\phi} = \left(\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T |\widehat{T}|\right) \widehat{\mathcal{M}}\widehat{\phi} \\
&\Leftrightarrow \quad \widehat{\mathcal{M}}\widehat{\phi} = \frac{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \int_{\widehat{T}} \widehat{\phi}}{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T |\widehat{T}|}.
\end{aligned}$$

Therefore,

$$\min_{\widehat{\phi}\in H^1(\widehat{\omega}_x)} \frac{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\nabla}\widehat{\phi}\|^2_{\widehat{T}}}{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\phi}-\widehat{\mathcal{M}}\widehat{\phi}\|^2_{\widehat{T}}} = \min_{\widehat{u}\perp_\beta 1} \frac{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{\nabla}\widehat{u}\|^2_{\widehat{T}}}{\sum_{\widehat{T}\subset\widehat{\omega}_x} \beta_T \|\widehat{u}\|^2_{\widehat{T}}}.$$

By the inverse change of variables, we find :

$$\lambda_1 \lesssim \frac{\sum_{T\subset\omega_x} \beta_T h_T^2 \|\nabla\phi\|^2_T}{\sum_{T\subset\omega_x} \beta_T \|\phi-\mathcal{M}\phi\|^2_T}.$$

■

Lemma 1.4.5. *For every function $\phi \in H^1_0(\Omega)$, any triangle $T\in\mathcal{T}_h$ and any edge $e\in\mathcal{E}_{h\Omega}$, one has*

$$\|\phi - \mathrm{I}^{new}_{\mathrm{Cl}}\phi\|_{\beta,T} \lesssim C_{Neu}(\beta) h_T \|\nabla\phi\|_{\beta,\Delta_T}, \tag{1.77}$$

$$\beta_e^{\frac{1}{2}} \|\phi - \mathrm{I}^{new}_{\mathrm{Cl}}\phi\|_e \lesssim (1+C_{Neu}(\beta))\, h_e^{\frac{1}{2}} \|\nabla\phi\|_{\beta,\Delta_e}, \tag{1.78}$$

where $C_{Neu}(\beta)=\max_{x\in\mathcal{N}_\Omega} C_{x,Neu}(\beta)$.

Proof: Let T be a triangle of $\mathcal{T}_h$. By the definition of $\mathrm{I}^{new}_{\mathrm{Cl}}$ and the estimate (1.75), we have

$$\begin{aligned}
\|\phi - \mathrm{I}^{new}_{\mathrm{Cl}}\phi\|_{\beta,T} &\lesssim \sum_{x\in\mathcal{N}_T} \|\phi-\mathcal{M}_x\phi\|_{\beta,T} \\
&\lesssim \sum_{x\in\mathcal{N}_T} C_{x,Neu}(\beta) h_x \|\nabla\phi\|_{\beta,\omega_x}
\end{aligned}$$

where $\mathcal{N}_T$ denotes the set of vertices of T. The regularity of the triangulation leads to (1.77).
Now, let $e \in \mathcal{E}_{h\Omega}$. We set $\omega_e = T_1 \cup T_2$ and can assume that $\beta_e = \max\{\beta_{T_1}, \beta_{T_2}\} = \beta_{T_1}$. We apply in T_1 the standard trace theorem ([54], Lemma 3.2) :

$$\|\varphi\|_e \lesssim h_e^{-\frac{1}{2}} \|\varphi\|_{T_1} + h_e^{\frac{1}{2}} |\varphi|_{1,T_1}, \tag{1.79}$$

and obtain

$$\begin{aligned} \beta_e^{\frac{1}{2}} \|\phi - \mathrm{I}_{\mathrm{Cl}}^{new}\phi\|_e &= \beta_{T_1}^{\frac{1}{2}} \sum_{x \in \mathcal{N}_\Omega \cap \bar{e}} \|\phi - \mathcal{M}_x\phi\|_e \\ &\lesssim \beta_{T_1}^{\frac{1}{2}} \left(h_e^{-\frac{1}{2}} \sum_{x \in \mathcal{N}_\Omega \cap \bar{e}} \|\phi - \mathcal{M}_x\phi\|_{T_1} + h_e^{\frac{1}{2}} \|\nabla\phi\|_{T_1} \right). \end{aligned}$$

The estimate (1.75) then gives

$$\begin{aligned} \beta_e^{\frac{1}{2}} \|\phi - \mathrm{I}_{\mathrm{Cl}}^{new}\phi\|_e &\lesssim \beta_{T_1}^{\frac{1}{2}} h_e^{-\frac{1}{2}} \sum_{x \in \mathcal{N}_{T_1}} \|\phi - \mathcal{M}_x\phi\|_{T_1} + h_e^{\frac{1}{2}} \|\nabla\phi\|_{\beta,\omega_e} \\ &\lesssim h_e^{-\frac{1}{2}} \sum_{x \in \mathcal{N}_{T_1}} h_x C_{x,Neu}(\beta) \|\nabla\phi\|_{\beta,\omega_x} + h_e^{\frac{1}{2}} \|\nabla\phi\|_{\beta,\omega_e} \\ &\lesssim h_e^{\frac{1}{2}} (1 + C_{Neu}(\beta)) \|\nabla\phi\|_{\beta,\Delta_e}. \end{aligned}$$

■

Lemma 1.4.6. *The interpolants* $\mathrm{I}_{\mathrm{Cl}}^{new}$ *and* I_{Cl} *are equivalent, in the sense that*

$$\|\phi - \mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x} \sim \|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x}, \ \forall x \in \mathcal{N}_\Omega.$$

Proof: On one hand, for every x in $\mathcal{N}_\Omega$, we may write

$$\begin{aligned} \phi - \mathrm{I}_{\mathrm{Cl,x}}\phi &= \phi - \mathrm{I}_{\mathrm{Cl,x}}\phi + \mathcal{M}_x\phi - \mathrm{I}_{\mathrm{Cl,x}}\mathcal{M}_x\phi \\ &= (\mathrm{I} - \mathrm{I}_{\mathrm{Cl,x}})(\mathrm{I} - \mathcal{M}_x)\phi, \end{aligned}$$

and therefore

$$\begin{aligned} \|\phi - \mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x} &= \|(\mathrm{I} - \mathrm{I}_{\mathrm{Cl,x}})(\phi - \mathcal{M}_x\phi)\|_{\beta,\omega_x} \\ &\lesssim \|\mathrm{I} - \mathrm{I}_{\mathrm{Cl,x}}\|_{\beta,\omega_x} \|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x} \\ &\lesssim \left(\|\mathrm{I}\|_{\beta,\omega_x} + \|\mathrm{I}_{\mathrm{Cl,x}}\|_{\beta,\omega_x} \right) \|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x}. \end{aligned}$$

with

$$\|\mathrm{I}\|_{\beta,\omega_x} = 1 = \max_{u \in L^2(\omega_x), \|u\|_{\beta,\omega_x}=1} \|u\|_{\beta,\omega_x}$$

and

$$\|\mathrm{I}_{\mathrm{Cl}}\|_{\beta,\omega_x} = \max_{u \in L^2(\omega_x), \|u\|_{\beta,\omega_x}=1} \|\mathrm{I}_{\mathrm{Cl,x}}u\|_{\beta,\omega_x} = \max_{u \in L^2(\omega_x)\setminus\{0\}} \frac{\|\mathrm{I}_{\mathrm{Cl,x}}u\|_{\beta,\omega_x}}{\|u\|_{\beta,\omega_x}}.$$

By the definition of I_{Cl}, we have :

$$\|\mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x} = \left(\sum_{T\subset\omega_x} \beta_T \|\mathrm{I}_{\mathrm{Cl,x}}\phi\|_T^2\right)^{\frac{1}{2}} = |\mathrm{I}_{\mathrm{Cl,x}}\phi| \left(\sum_{T\subset\omega_x} \beta_T |T|\right)^{\frac{1}{2}}.$$

As, by Cauchy-Schwarz's inequality,

$$|\mathrm{I}_{\mathrm{Cl,x}}\phi| = \left|\frac{1}{|\omega_x \cap \Omega_{l(x)}|} \left(\int_{\omega_x\cap\Omega_{l(x)}} \phi\right)\right| \leq \frac{|\omega_x \cap \Omega_{l(x)}|^{\frac{1}{2}}}{|\omega_x \cap \Omega_{l(x)}|} \|\phi\|_{L^2(\omega_x\cap\Omega_{l(x)})}$$

we obtain that

$$\begin{aligned}
\|\mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x} &\leq |\omega_x \cap \Omega_{l(x)}|^{-\frac{1}{2}} \left(\sum_{T\subset\omega_x} \beta_T |T|\right)^{\frac{1}{2}} \left(\sum_{T\subset\omega_x\cap\Omega_{l(x)}} \beta_T \beta_T^{-1} \|\phi\|_T^2\right)^{\frac{1}{2}} \\
&\lesssim h_x^{-1} \left(\sum_{T\subset\omega_x} \beta_T h_x^2\right)^{\frac{1}{2}} \beta_{l(x)}^{-\frac{1}{2}} \left(\sum_{T\subset\omega_x\cap\Omega_{l(x)}} \beta_T \|\phi\|_T^2\right)^{\frac{1}{2}} \\
&\lesssim \|\phi\|_{\beta,\omega_x}.
\end{aligned}$$

Hence $\|\mathrm{I}_{\mathrm{Cl,x}}\|_{\beta,\omega_x} \lesssim 1$ that leads to $\|\phi - \mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x} \lesssim \|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x}$.
On the other hand, for x in $\mathcal{N}_\Omega$,

$$\begin{aligned}
\phi - \mathcal{M}_x\phi &= \phi - \mathcal{M}_x\phi + \mathrm{I}_{\mathrm{Cl,x}}\phi - \mathcal{M}_x\mathrm{I}_{\mathrm{Cl,x}}\phi \\
&= (\mathrm{I} - \mathcal{M}_x)(\mathrm{I} - \mathrm{I}_{\mathrm{Cl,x}})\phi,
\end{aligned}$$

and then

$$\|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x} \lesssim \|\mathrm{I} - \mathcal{M}_x\|_{\beta,\omega_x} \|\phi - \mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x}.$$

Since

$$\begin{aligned}
\|\mathcal{M}_x\phi\|_{\beta,\omega_x} &= \left(\sum_{T\subset\omega_x} \beta_T \|\mathcal{M}_x\phi\|_T^2\right)^{\frac{1}{2}} \\
&= |\mathcal{M}_x\phi| \left(\sum_{T\subset\omega_x} \beta_T |T|\right)^{\frac{1}{2}},
\end{aligned}$$

and

$$\begin{aligned}
|\mathcal{M}_x\phi| &\leq \frac{\sum_{T\subset\omega_x} \beta_T \frac{1}{|T|} |T|^{\frac{1}{2}} \|\phi\|_T}{\sum_{T\subset\omega_x} \beta_T} \leq \frac{\left(\sum_{T\subset\omega_x} \beta_T \frac{1}{|T|}\right)^{\frac{1}{2}} \left(\sum_{T\subset\omega_x} \beta_T \|\phi\|_T^2\right)^{\frac{1}{2}}}{\sum_{T\subset\omega_x} \beta_T} \\
&\leq \frac{\left(\sum_{T\subset\omega_x} \beta_T \frac{1}{|T|}\right)^{\frac{1}{2}}}{\sum_{T\subset\omega_x} \beta_T} \|\phi\|_{\beta,\omega_x},
\end{aligned}$$

we obtain

$$\|\mathcal{M}_x\phi\|_{\beta,\omega_x} \lesssim \frac{\left(\sum_{T\subset\omega_x} \beta_T \frac{1}{|T|}\right)^{\frac{1}{2}}}{\sum_{T\subset\omega_x} \beta_T} \left(\sum_{T\subset\omega_x} \beta_T |T|\right)^{\frac{1}{2}} \|\phi\|_{\beta,\omega_x}.$$

This implies that

$$\|\mathcal{M}_x\phi\|_{\beta,\omega_x} \lesssim \|\phi\|_{\beta,\omega_x},$$

hence

$$\|\mathcal{M}_x\|_{\beta,\omega_x} \lesssim 1.$$

Therefore we conclude that $\|\phi - \mathcal{M}_x\phi\|_{\beta,\omega_x} \lesssim \|\phi - \mathrm{I}_{\mathrm{Cl,x}}\phi\|_{\beta,\omega_x}$. ∎

Remark 1.4.7. Lemma 1.4.6 and Lemma 2.8 from [11] imply that under the above mentioned geometric assumptions, the constant $C_{Neu}(\beta) \lesssim 1$. Note further that Lemma 1.4.6 shows that the use of I_{Cl} or $\mathrm{I}^{new}_{\mathrm{Cl}}$ is equivalent.

Nédélec interpolation

Let $T \in \mathcal{T}_h$ be a triangle and $\mathcal{E}_{hT}$ the set of its edges. For $e \in \mathcal{E}_{h\Omega}$, we fix $\boldsymbol{t}_e$ one of the unit tangential vectors along the edge e. For $T \in \mathcal{T}_h$, we define the set of linear forms $\{l_e, e \in \mathcal{E}_{hT}\}$ by

$$\begin{aligned} l_e : L^1(e) &\rightarrow \mathbb{R} \\ u &\rightarrow \int_e u \cdot \boldsymbol{t}_e \, ds, \end{aligned}$$

and consider the (basis) functions $\lambda_e \in \mathcal{NED}_1$ satisfying the condition (see [37])

$$\forall e \in \mathcal{E}_{hT}, \int_{e'} \lambda_e \cdot \boldsymbol{t}_{e'} = \delta_{e,e'}.$$

We further introduce the local interpolation operator $\mathrm{I}_{\mathrm{Ned}|T}(u) \in \mathcal{NED}_1$ defined, for u satisfying $u_{|e} \in (L^1(e))^2$, by the conditions

$$l_e\left(\mathrm{I}_{\mathrm{Ned}|T}(u)\right) = l_e(u), \forall e \in \mathcal{E}_{hT}.$$

This means that

$$\mathrm{I}_{\mathrm{Ned}|T}(u) = \sum_{e\in\mathcal{E}_{hT}} \left(\int_e u \cdot \boldsymbol{t}_e \, ds\right) \lambda_e.$$

The global interpolation operator $\mathrm{I}_{\mathrm{Ned}}$ is then given by $(\mathrm{I}_{\mathrm{Ned}}u)_{|T} = \mathrm{I}_{\mathrm{Ned}|T}(u_{|T}) \in \mathcal{NED}_1, \forall T \in \mathcal{T}_h$ as

$$\begin{aligned} \mathrm{I}_{\mathrm{Ned}} : [PH^1(\Omega)]^2 \bigcap H_0(\mathrm{curl},\Omega) &\rightarrow V_h \\ u &\rightarrow \sum_{e\in\mathcal{E}_{h\Omega}} \left(\int_e u \cdot \boldsymbol{t}_e \, ds\right) \lambda_e. \end{aligned}$$

Lemma 1.4.8. *Let T be an element and e an edge of the triangulation. For every function $\boldsymbol{w} \in [PH^1(\Omega)]^2 \bigcap H_0(\mathrm{curl}, \Omega)$, we have :*

$$\|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_T \lesssim h_T \beta_T^{-\frac{1}{2}} \|\nabla_P \boldsymbol{w}\|_{\beta,T} \tag{1.80}$$

$$\|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_e \lesssim h_e^{\frac{1}{2}} \beta_e^{-\frac{1}{2}} \|\nabla_P \boldsymbol{w}\|_{\beta,\omega_e} \tag{1.81}$$

Proof: We consider an element $T \in \mathcal{T}_h$. As $\mathrm{I}_{\mathrm{Ned}\,|T}$ depends only on the triangle T,

$$\begin{aligned} \|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_T &= \|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}\,|T}\boldsymbol{w}_{|T}\|_T \\ &\lesssim diam(T)\|\nabla \boldsymbol{w}_{|T}\|_T, \end{aligned}$$

by a Bramble-Hilbert argument. The estimate (1.80) directly follows as $diam(T) \sim h_T$.

For an edge e, we apply the trace estimate (1.79) with T adjacent to e such that $\beta_T = \beta_e$ and find

$$\|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_e \lesssim h_e^{-\frac{1}{2}} \|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_T + h_e^{\frac{1}{2}} \|\nabla(\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w})\|_T.$$

By (1.80) applied to T, the regularity of the triangulation, the triangular inequality and the trivial inequality

$$\|\nabla \boldsymbol{w}\|_T^2 \leq \beta_T^{-1} \|\nabla_P \boldsymbol{w}\|_{\beta,\omega_e}^2,$$

we deduce that :

$$\|\boldsymbol{w} - \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w}\|_e \lesssim h_e^{\frac{1}{2}} \beta_e^{-\frac{1}{2}} \|\nabla_P \boldsymbol{w}\|_{\beta,\omega_e} + h_e^{\frac{1}{2}} \|\nabla(\mathrm{I}_{\mathrm{Ned}}\boldsymbol{w})\|_T.$$

Moreover as $\mathrm{I}_{\mathrm{Ned}\,|T}\boldsymbol{w} \in \mathcal{ND}_1$, we know (see [37]) that

$$\begin{aligned} \|\nabla(\mathrm{I}_{\mathrm{Ned}}\boldsymbol{w})\|_T &= \frac{\sqrt{2}}{2} \|\,\mathrm{curl}(\mathrm{I}_{\mathrm{Ned}}\boldsymbol{w})\|_T \\ &\lesssim \|\,\mathrm{curl}\,\boldsymbol{w}\|_T \\ &\lesssim \|\nabla \boldsymbol{w}\|_T \end{aligned}$$

that leads to the conclusion. ∎

A Clément-Nédélec interpolant

Let us define a Clément-Nédélec interpolant by :

$$\begin{aligned} \mathrm{I}_{\mathrm{CN}}^{\beta} : L^2(\Omega) &\to V_h \\ u &\to \sum_{e \in \mathcal{E}_{h\Omega}} \Theta_e(u) \tilde{\lambda}_e \end{aligned}$$

where $\Theta_e(u) = \dfrac{1}{|T_e|} \displaystyle\int_{T_e} u \cdot \boldsymbol{t}_e$, for $T_e \in \mathcal{T}_h$ such that $\beta_{T_e} = \beta_e$ and $\tilde{\lambda}_e = \lambda_e |e|$.

This new interpolant is well-defined, is stable relatively to the β-norm and the β-H^1-seminorm and satisfies standard interpolant error estimates, i.e. we have the following estimates :

Theorem 1.4.9. *For every function $u \in [PH^1(\Omega)]^2 \cap H(\mathrm{curl},\Omega)$, any $T \in \mathcal{T}_h$ and any $e \in \mathcal{E}_{h\Omega}$, we have*

$$\begin{aligned}
\|\mathrm{I}^{\beta}_{\mathrm{CN}} u\|_{\beta,T} &\lesssim \|u\|_{\beta,\omega_T}, && (1.82)\\
\|u - \mathrm{I}^{\beta}_{\mathrm{CN}} u\|_{\beta,T} &\lesssim C^{\star}_{Neu}(\beta) h_T \|\nabla u\|_{\beta,\omega_T}, && (1.83)\\
\|\nabla(u - \mathrm{I}^{\beta}_{\mathrm{CN}} u)\|_{\beta,T} &\lesssim \|\nabla u\|_{\beta,\omega_T}, && (1.84)\\
\beta_e^{1/2}\|u - \mathrm{I}^{\beta}_{\mathrm{CN}} u\|_e &\lesssim h_e^{1/2}(C^{\star}_{Neu}(\beta)+1)\|\nabla u\|_{\beta,\Delta_e}, && (1.85)
\end{aligned}$$

where $C^{\star}_{Neu}(\beta)$ is the converse of the squareroot of the first positive eigenvalue of the following Neumann-type problem :

$$\left\{\begin{array}{ll}
-\Delta u = \lambda u & \text{in } \widehat{T},\ \forall \widehat{T} \subset \widehat{\omega}_T,\\
\big[\!\big[u_T\big]\!\big]_e = 0 & \text{on } \widehat{e} \subset \ \text{int } \widehat{\omega}_T,\\
\displaystyle\sum_{\widehat{T}\subset\widehat{\omega}_T:\widehat{e}\subset\partial\widehat{T}} \beta_{\widehat{T}} \frac{\partial u_T}{\partial n} = 0 & \text{on } \widehat{e} \subset \widehat{\omega}_T,\\
\dfrac{\partial u_N}{\partial n} = 0 & \text{on } \widehat{e} \subset \widehat{\omega}_T,
\end{array}\right. \qquad (1.86)$$

where u_T and u_N denote respectively the tangential and normal components of u.

Proof: We first define

$$\mathcal{R}_0(\omega_T) = \big\{ c \in H(\mathrm{curl},\omega_T) \ :\ c_{|T'} \in \mathbb{R}^2, \forall T' \subset \omega_T \big\}$$

and prove that $\mathrm{I}^{\beta}_{\mathrm{CN}} u = u$ on T if $u \in \mathcal{R}_0(\omega_T)$. Indeed, for $u \in \mathcal{R}_0(\omega_T)$ and $e \subset T$, we have

$$\Theta_e(u) = \frac{1}{|T_e|}\int_{T_e} c \cdot \boldsymbol{t}_e = c \cdot \boldsymbol{t}_e = \frac{1}{|e|}\int_e c \cdot \boldsymbol{t}_e.$$

Then, the definition of $\mathrm{I}_{\mathrm{Ned}}$ implies that $\mathrm{I}^{\beta}_{\mathrm{CN}\,|T} u = \mathrm{I}_{\mathrm{Ned}\,|T} u = u$.

Let us now show (1.82) : By Cauchy-Schwarz's inequality, we may write

$$|\Theta_e(u)| \ \leq\ \frac{1}{|T_e|}\|u\|_{T_e}\|\boldsymbol{t}_e\|_{T_e}$$

Since

$$\|\boldsymbol{t}_e\|_{T_e} \leq \|\boldsymbol{t}_e\|_{\infty}|T_e|^{1/2}$$

and $\|\boldsymbol{t}_e\|_{\infty} = 1$, we get

$$|\Theta_e(u)| \leq \frac{1}{|T_e|^{1/2}}\|u\|_{T_e}.$$

By the definition of $\mathrm{I}^{\beta}_{\mathrm{CN}}$, we obtain

$$\beta_{T'}^{1/2}\|\mathrm{I}^{\beta}_{\mathrm{CN}} u\|_{T'} \leq \sum_{e\subset\partial T'} \frac{1}{|T_e|^{1/2}}\beta_{T'}^{1/2}\|u\|_{T_e}\|\tilde{\lambda}_e\|_{T'}.$$

Moreover, if $\widehat{\lambda}$ denotes a basis function on the reference element, we have

$$\|\tilde{\lambda}_e\|_{T'} = |e|\|\lambda_e\|_{T'} = |e|\|B_{T'}^{-t}\widehat{\lambda}\|_{T'} \lesssim |e|h_{T'}^{-1}\|\widehat{\lambda}\|_{\widehat{T'}}h_{T'} \lesssim |e|$$

where $B_{T'}$ is the 2×2 matrix corresponding to the affine transformation $F_{T'}$ that maps $\widehat{T}' \subset \widehat{\omega}_T$ onto T'.

As $|e| = h_e$, $|T_e| \sim h_e^2$ and $\beta_{T'} \leq \beta_e$, we conclude that

$$\beta_{T'}^{1/2}\|\mathrm{I}_{\mathrm{CN}}^{\beta}u\|_{\beta,T'} \lesssim \sum_{e\subset\partial T'} \|u\|_{\beta,T_e}$$

which implies (1.82).

Now, for any $p \in \mathcal{R}_0(\omega_T)$, $u - \mathrm{I}_{\mathrm{CN}}^{\beta}u = (\mathrm{I} - \mathrm{I}_{\mathrm{CN}}^{\beta})(u-p)$ and therefore by (1.82) :

$$\begin{aligned}\beta_T^{1/2}\|u - \mathrm{I}_{\mathrm{CN}}^{\beta}u\|_T &= \beta_T^{1/2}\|(\mathrm{I} - \mathrm{I}_{\mathrm{CN}}^{\beta})(u-p)\|_T \\ &\lesssim \|u-p\|_{\beta,\omega_T}.\end{aligned}$$

For this estimate, we see that (1.83) holds if we can bound from below the ratio

$$\frac{h_T^2 \sum_{T'\subset\omega_T} \beta_{T'}\|\nabla u\|_{T'}^2}{\sum_{T'\subset\omega_T} \beta_{T'}\|u-p\|_{T'}^2}$$

for $u \in H(\mathrm{curl},\omega_T) \cap [PH^1(\omega_T)]^2$ and $p \in \mathcal{R}_0(\omega_T)$, which is equivalent, by applying the affine transformation F_T mapping the patch $\widehat{\omega}_T$ to ω_T (see section 1.2.2) and making the change of unknown $\widehat{u}_{|\widehat{T}'} = B_{T'}^t\, u_{|T'}, \forall T' \subset \omega_T$, to bound the ratio

$$\frac{\sum_{\widehat{T}'\subset\widehat{\omega}_T} \beta_{T'}\|\widehat{\nabla}\widehat{u}\|_{\widehat{T}'}^2}{\sum_{\widehat{T}'\subset\widehat{\omega}_T} \beta_{T'}\|\widehat{u}-\widehat{p}\|_{\widehat{T}'}^2} \tag{1.87}$$

for $\widehat{u} \in H(\mathrm{curl},\widehat{\omega}_T) \cap [PH^1(\widehat{\omega}_T)]^2$ and $\widehat{p} \in \mathcal{R}_0(\widehat{\omega}_T)$. This last ratio will be estimated from below using the min-max principle.

Indeed, we set $V = H(\mathrm{curl},\widehat{\omega}_T) \cap [PH^1(\widehat{\omega}_T)]^2$, $H = L^2(\widehat{\omega}_T)$, define the bilinear form

$$l(u,v) = \sum_{\widehat{T}'\subset\widehat{\omega}_T} \beta_{T'} \int_{\widehat{T}'} \widehat{\nabla}u : \widehat{\nabla}v,\ \forall (u,v) \in V \times V$$

and introduce the inner product $(u,v)_\beta = \sum_{\widehat{T}'\subset\widehat{\omega}_T} \beta_{T'} \int_{\widehat{T}'} u \cdot v$, for u and v in H.

The corresponding spectral problem consists in finding $\lambda \in \mathbb{R}$ and $u \in V$, $u \neq 0$ solution of

$$l(u,v) = \lambda (u,v)_\beta, \ \forall v \in V. \tag{1.88}$$

Taking first v in $\mathcal{D}(\widehat{T}')$ for $\widehat{T}' \subset \widehat{\omega}_T$ and applying Green's formula give that u satifies, in the distribution sense,

$$-\Delta u = \lambda u \text{ on } \widehat{T}'.$$

Moreover, as $u \in H(\mathrm{curl}, \widehat{\omega}_T)$,

$$\left[\left[u_T \right]\right]_e = 0, \forall \widehat{e} \subset \widehat{\omega}_T.$$

Now, for any $v \in V$, if we use Green's formula on each $\widehat{T}'$, we obtain that

$$\sum_{\widehat{T}' \subset \widehat{\omega}_T} \beta_{\widehat{T}'} \int_{\partial \widehat{T}'} \frac{\partial u}{\partial n} \cdot v = \sum_{\widehat{T}' \subset \widehat{\omega}_T} \beta_{\widehat{T}'} \int_{\partial \widehat{T}'} \left(\frac{\partial u_{N|\widehat{T}'}}{\partial n} v_{N|\widehat{T}'} + \frac{\partial u_{T|\widehat{T}'}}{\partial n} v_{T|\widehat{T}'} \right) = 0, \ \forall v \in V. \tag{1.89}$$

This implies the third and fourth conditions of (1.86). Indeed, let $\widehat{e}$ be an arbitrary edge of the patch $\widehat{\omega}_T$ and fix the unit normal and tangential vectors $n_{\widehat{e}}$ and $t_{\widehat{e}}$ along this edge. We consider a function $\varphi \in \mathcal{D}(\omega_{\widehat{e}})$ and first prove the third condition :
Set $v = \varphi \, t_{\widehat{e}}$ on $\widehat{\omega}_T$. Then, $v_N = 0$ on every edge $\widehat{e}' \subset \widehat{\omega}_T$ and as $v \in H(\mathrm{curl}, \widehat{\omega}_T)$, v_T is continuous on the interfaces. That's why (1.89) becomes :

$$\begin{aligned} & \sum_{\widehat{T} \subset \widehat{\omega}_T} \sum_{\widehat{e} \subset \widehat{T}} \beta_{\widehat{T}} \int_{\widehat{e}} \frac{\partial u_T}{\partial n} \varphi = 0 \\ \Leftrightarrow \quad & \left\langle \sum_{\widehat{T} \subset \widehat{\omega}_T \, : \, \widehat{e} \subset \widehat{T}} \beta_{\widehat{T}} \frac{\partial u_T}{\partial n}, \varphi \right\rangle = 0. \end{aligned}$$

That leads to the conclusion.
Now, we set $v = \psi \, n_{\widehat{e}}$ on $\widehat{\omega}_T$ where $\psi = \left\{ \begin{array}{ll} 0 & \text{on } T_1 \subset \omega_{\widehat{e}} \\ \varphi & \text{on } T_2 \subset \omega_{\widehat{e}} \end{array} \right.$. This time, $v_T = 0$ and (1.89) :

$$\begin{aligned} & \sum_{\widehat{e} \subset \widehat{\omega}_T} \int_{\widehat{e}} \frac{\partial u_N}{\partial n} \psi = 0 \\ \Leftrightarrow \quad & \left\langle \sum_{\widehat{e} \subset T_2 \cap \omega_{\widehat{e}}} \frac{\partial u_N}{\partial n}, \varphi \right\rangle = 0 \\ \Leftrightarrow \quad & \frac{\partial u_N}{\partial n} = 0 \text{ on } e' = T_2 \cap \omega_{\widehat{e}}. \end{aligned}$$

Exchanging the roles of the triangles T_1 and T_2 and applying the same proof to all edges $\widehat{e} \subset \widehat{\omega}_T$ give the fourth condition.

We now define the self-adjoint operator A associated with the problem (1.88) by

$$\begin{array}{rccl} A \ : \ D(A) \subset H & \to & H \\ u & \to & Au, \end{array}$$

where $u \in D(A)$ iff $\exists f \in H \ : \ l(u,v) = (f,v), \ \forall v \in V$ and then set $Au = f$.
Since V is compactly embedded into H, A has a compact inverse. Therefore this operator admits a discrete spectrum and, by the min-max principle, its first eigenvalue satisfies :

$$\lambda_1 = \min_{\substack{v \in V, v \neq 0 \\ v \perp_\beta ker A}} \frac{l(v,v)}{\|v\|_H^2}.$$

Since $ker A = \mathcal{R}_0(\widehat{\omega}_T)$, we deduce that

$$\lambda_1 = \min_{\widehat{u} \in V, \ \widehat{u} \perp_\beta \mathcal{R}_0(\widehat{\omega}_T)} \frac{\displaystyle\sum_{\widehat{T}' \subset \widehat{\omega}_T} \beta_{T'} \|\widehat{\nabla} \widehat{u}\|_{\widehat{T}'}^2}{\displaystyle\sum_{\widehat{T}' \subset \widehat{\omega}_T} \beta_{T'} \|\widehat{u} - \widehat{p}\|_{\widehat{T}'}^2}.$$

This gives, by choosing in (1.87) $\widehat{p}$ as the projection of $\widehat{u}$ on $\mathcal{R}_0(\widehat{\omega}_T)$ with respect to the inner product $(\cdot,\cdot)_\beta$, the following estimate :

$$\|\widehat{u} - \widehat{p}\|_{\beta, \widehat{\omega}_T} \ \lesssim \ \lambda^{-1/2} \|\widehat{\nabla} \widehat{u}\|_{\beta, \widehat{\omega}_T}.$$

This implies (1.83) by the above mentioned scaling argument.

We now prove the third estimate. First as $\mathrm{I}_{\mathrm{CN}}^\beta u \in [\mathcal{P}_1(T)]^2$, a standard inverse inequality [17] and the estimate (1.82) yield

$$\beta_T^{1/2} \|\nabla(\mathrm{I}_{\mathrm{CN}}^\beta u)\|_T \ \lesssim \ h_T^{-1} \beta_T^{1/2} \|\mathrm{I}_{\mathrm{CN}}^\beta u\|_T \lesssim h_T^{-1} \|u\|_{\beta, \omega_T}.$$

By the triangular inequality we get

$$\begin{aligned} \beta_T^{1/2} \|\nabla(u - \mathrm{I}_{\mathrm{CN}}^\beta u)\|_T \ &\lesssim \ \beta_T^{1/2} \|\nabla u\|_T + \beta_T^{1/2} \|\nabla(\mathrm{I}_{\mathrm{CN}}^\beta u)\|_T \\ &\lesssim \ \|\nabla u\|_{\beta,T} + h_T^{-1} \|u\|_{\beta, \omega_T}. \end{aligned}$$

Moreover, as for any $p \in \mathcal{R}_0(\omega_T)$, $u - \mathrm{I}_{\mathrm{CN}}^\beta u = (\mathrm{I} - \mathrm{I}_{\mathrm{CN}}^\beta)(u - p)$, we find

$$\begin{aligned} \beta_T^{1/2} \|\nabla(u - \mathrm{I}_{\mathrm{CN}}^\beta u)\|_T \ &\lesssim \ \beta_T^{1/2} \|\nabla \left[(\mathrm{I} - \mathrm{I}_{\mathrm{CN}}^\beta)(u - p) \right] \|_T \\ &\lesssim \ \beta_T^{1/2} \|\nabla(u - p)\|_T + h_T^{-1} \|u - p\|_{\beta, \omega_T} \\ &\lesssim \ \|\nabla u\|_{\beta,T} + h_T^{-1} \|u - p\|_{\beta, \omega_T}. \end{aligned}$$

Since we have shown by the min-max principle that

$$\|u - p\|_{\beta, \omega_T} \lesssim h_T \|\nabla u\|_{\beta, \omega_T}$$

the conclusion follows.
The last estimate (1.85) is a direct consequence of the trace estimate (1.79) (applied in T_e) and the estimates (1.83) and (1.84). ∎

1.4.3 Error estimates

Residual error estimators

On an element T, let $\boldsymbol{R}_T := \boldsymbol{f} - (\operatorname{curl}(\chi \operatorname{curl} \boldsymbol{u}_h) + \beta \boldsymbol{u}_h)$ be the exact residual, and denote by $\boldsymbol{r}_T$ its approximated residual.
Introduce the jump of $\boldsymbol{u}_h$ in the normal direction and the jump of $\operatorname{curl} \boldsymbol{u}_h$ in the tangential direction by

$$\begin{aligned} \mathbf{J}_{e,n} &:= \begin{cases} \left[\!\left[\beta \boldsymbol{u}_h \cdot \boldsymbol{n}_e\right]\!\right]_e & \text{for interior edges} \\ 0 & \text{for boundary edges,} \end{cases} \\ \mathbf{J}_{e,t} &:= \begin{cases} \left[\!\left[\chi \operatorname{curl} \boldsymbol{u}_h\right]\!\right]_e & \text{for interior edges} \\ 0 & \text{for boundary edges.} \end{cases} \end{aligned}$$

In the following study, we will build two different local error estimators of the solenoidal part of the error. The first one is inspired from [11] and the second one has been adapted from [33, 53], where convection-reaction-diffusion problems are considered. Note that the second estimator reduces to the one analyzed in [19] in the case χ and β constant.

Definition 1.4.10. *The local and global residual error estimators are defined by*

$$\begin{aligned} \eta_0^2 &:= \sum_{T \in \mathcal{T}_h} \eta_{T,0}^2, \\ \eta_\perp^2 &:= \sum_{T \in \mathcal{T}_h} \eta_{T,\perp}^2, \\ \eta^2 &:= \eta_0^2 + \eta_\perp^2, \\ \zeta^2 &:= \sum_{T \in \mathcal{T}_h} \zeta_T^2, \end{aligned}$$

$$\begin{aligned} \eta_{T,0}^2 &:= h_T^2 \beta_T^{-1} \| \operatorname{div}(\beta \boldsymbol{u}_h) \|_T^2 + \sum_{e \subset \partial T} h_e \beta_e^{-1} \|\mathbf{J}_{e,n}\|_e^2, \\ \bullet 1^{rst}\ \textit{method} : \eta_{T,\perp}^2 &:= h_T^2 \beta_T^{-1} \|\boldsymbol{r}_T\|_T^2 + \sum_{e \subset \partial T} h_e \beta_e^{-1} \|\mathbf{J}_{e,t}\|_e^2, \\ \zeta_T^2 &:= h_T^2 \beta_T^{-1} \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T^2, \\ \bullet 2^{nd}\ \textit{method} : \eta_{T,\perp}^2 &:= \alpha_T^2 \|\boldsymbol{r}_T\|_T^2 + \sum_{e \subset \partial T} \chi_e^{-1/2} \alpha_e \|\mathbf{J}_{e,t}\|_e^2, \\ \zeta_T^2 &:= \sum_{T' \subset \omega_T} \alpha_{T'}^2 \|\boldsymbol{r}_{T'} - \boldsymbol{R}_{T'}\|_{T'}^2, \end{aligned}$$

where $\beta_e := \max\limits_{\partial T_1 \cap \partial T_2 = \{e\}} \{\beta_{T_1}, \beta_{T_2}\}$, $\chi_e := \max\limits_{\partial T_1 \cap \partial T_2 = \{e\}} \{\chi_{T_1}, \chi_{T_2}\}$ *and, for any* $S \in \mathcal{T}_h \cup \mathcal{E}_h$, $\alpha_S := \min\{\beta_S^{-1/2}, \chi_S^{-1/2} h_S\}$.

Proof of the lower error bound : the irrotational part

Theorem 1.4.11. *For all elements T, we have the following local error bound :*

$$\eta_{T,0} \lesssim \|e\|_{\beta,\omega_T}. \tag{1.90}$$

Proof:
◇ *Divergence*
By the inverse inequality (1.19) and Green's formula,

$$\begin{aligned}
\|\operatorname{div}(\beta \boldsymbol{u}_h)\|_T^2 &\sim \int_T b_T(\operatorname{div}(\beta \boldsymbol{u}_h))^2 \\
&\sim -\int_T \nabla(b_T \operatorname{div}(\beta \boldsymbol{u}_h))\beta \boldsymbol{u}_h \\
&\sim r(\nabla(b_T \operatorname{div}(\beta \boldsymbol{u}_h))) \quad \text{by (1.14) } \textit{and} \text{ (1.4)} \\
&\sim a(\boldsymbol{e}, \nabla(b_T \operatorname{div}(\beta \boldsymbol{u}_h))) \\
&\sim \int_T \beta \boldsymbol{e} \nabla(b_T \operatorname{div}(\beta \boldsymbol{u}_h)) \\
&\lesssim \beta_T^{\frac{1}{2}} \|\nabla(b_T \operatorname{div}(\beta \boldsymbol{u}_h))\|_T \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \\
&\lesssim \beta_T^{\frac{1}{2}} h_T^{-1} \|\operatorname{div}(\beta \boldsymbol{u}_h))\|_T \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \quad \text{by (1.20).}
\end{aligned}$$

This shows that

$$\|\operatorname{div}(\beta \boldsymbol{u}_h))\|_T \lesssim \beta_T^{\frac{1}{2}} h_T^{-1} \|\boldsymbol{e}\|_{\beta,T}. \tag{1.91}$$

◇ *Normal jump*
Let e be an interior edge; we recall that $\mathbf{J}_{e,n} \in \mathbb{P}^k(e)$ with $k \in \mathbb{N}$ depending on the chosen finite element space. Set

$$w_e := F_{ext}(\mathbf{J}_{e,n}) b_e \in [H_0^1(\omega_e)]^2.$$

An elementwise partial integration gives

$$\begin{aligned}
\int_e \mathbf{J}_{e,n} w_e &= -\int_e \left[\!\left[\beta(\boldsymbol{u} - \boldsymbol{u}_h) \cdot \boldsymbol{n}_e\right]\!\right]_e w_e \\
&= \pm \sum_{T \subset \omega_e} \left[\int_T \beta \boldsymbol{e} \nabla w_e - \int_T \operatorname{div}(\beta \boldsymbol{e}) w_e \right] \\
&= \pm \sum_{T \subset \omega_e} \left[\int_T \beta \boldsymbol{e} \nabla w_e + \int_T \operatorname{div}(\beta \boldsymbol{u}_h) w_e \right] \\
&\lesssim \sum_{T \subset \omega_e} \left(\|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \beta_T^{\frac{1}{2}} \|\nabla w_e\|_T + \|\operatorname{div}(\beta \boldsymbol{u}_h)\|_T \|w_e\|_T \right) \\
&\lesssim \sum_{T \subset \omega_e} \left(\|\boldsymbol{e}\|_{\beta,T} \beta_T^{\frac{1}{2}} h_T^{-\frac{1}{2}} \|\mathbf{J}_{e,n}\|_e + \|\operatorname{div}(\beta \boldsymbol{u}_h)\|_T h_T^{\frac{1}{2}} \|\mathbf{J}_{e,n}\|_e \right)
\end{aligned}$$

by (1.22) and (1.23). Since (1.21) yields $\displaystyle\int_e \mathbf{J}_{e,n} w_e \sim \|\mathbf{J}_{e,n}\|_e^2$, we obtain

$$\|\mathbf{J}_{e,n}\|_e \lesssim \sum_{T \subset \omega_e} \left(\beta_T^{\frac{1}{2}} h_T^{-\frac{1}{2}} \|\boldsymbol{e}\|_{\beta,T} + h_T^{\frac{1}{2}} \|\operatorname{div}(\beta \boldsymbol{u}_h)\|_T \right).$$

This estimate coupled with (1.91) implies :

$$\|\mathbf{J}_{e,n}\|_e \lesssim \sum_{T\subset\omega_e} \left(\beta_T^{\frac{1}{2}} h_T^{-\frac{1}{2}} \|\boldsymbol{e}\|_{\beta,T}\right).$$

As $\mathcal{T}_h$ is regular, $h_T \sim h_e$, and $\beta_e = \max\{\beta_{T'} | e \subset \partial T'\} \geq \beta_T$ for $T \subset \omega_e$, we obtain :

$$\|\mathbf{J}_{e,n}\|_e \lesssim \beta_e^{\frac{1}{2}} h_e^{-\frac{1}{2}} \|\boldsymbol{e}\|_{\beta,\omega_e}. \tag{1.92}$$

The estimates (1.91) and (1.92) lead to the conclusion. ■

Proof of the lower error bound : the solenoidal part - first method

Theorem 1.4.12. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,\perp} \lesssim \sum_{T'\subset\omega_T} \left(\chi_{T'}^{\frac{1}{2}} \beta_{T'}^{-\frac{1}{2}} + h_{T'}\right) \|e\|_{\beta,\chi,T'} + \sum_{T'\subset\omega_T} \zeta_{T'}\mu. \tag{1.93}$$

Proof:
$\diamond$ *Element residual*
Let T be an element of the triangulation. Set $w_T := \boldsymbol{r}_T b_T \in [H_0^1(T)]^2$.
By the inverse inequality (1.19), Green's formula and the fact that b_T is zero on the boundary of T, we write

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\sim \int_T \boldsymbol{r}_T w_T \\
&\sim \int_T \boldsymbol{R}_T w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T (\boldsymbol{f} - \operatorname{curl}(\chi \operatorname{curl} \boldsymbol{u}_h) - \beta \boldsymbol{u}_h) w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T (\boldsymbol{f} - \beta \boldsymbol{u}_h) w_T - \int_T \chi \operatorname{curl} \boldsymbol{u}_h \operatorname{curl} w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim r(w_T) + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T.
\end{aligned}$$

The relation (1.13) implies

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\sim a(\boldsymbol{e}, w_T) + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\sim \int_T \chi \operatorname{curl} \boldsymbol{e} \operatorname{curl} w_T + \int_T \beta \boldsymbol{e} w_T + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_T \\
&\lesssim \|\chi^{\frac{1}{2}} \operatorname{curl} \boldsymbol{e}\|_T \chi_T^{\frac{1}{2}} \| \operatorname{curl} w_T\|_T + \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \beta_T^{\frac{1}{2}} \|w_T\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|w_T\|_T.
\end{aligned}$$

The inverse inequalities (1.19) and (1.20) give

$$\begin{aligned}
\|\boldsymbol{r}_T\|_T^2 &\lesssim \|\chi^{\frac{1}{2}} \operatorname{curl} \boldsymbol{e}\|_T \chi_T^{\frac{1}{2}} h_T^{-1} \|\boldsymbol{r}_T\|_T + \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \beta_T^{\frac{1}{2}} \|\boldsymbol{r}_T\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|\boldsymbol{r}_T\|_T \\
&\lesssim \left(\|\chi^{\frac{1}{2}} \operatorname{curl} \boldsymbol{e}\|_T + \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T\right)^{\frac{1}{2}} \left(\chi_T h_T^{-2} + \beta_T\right)^{\frac{1}{2}} \|\boldsymbol{r}_T\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|\boldsymbol{r}_T\|_T.
\end{aligned}$$

Then,

$$\|\boldsymbol{r}_T\|_T \lesssim \left(\chi_T h_T^{-2} + \beta_T\right)^{\frac{1}{2}} \|\boldsymbol{e}\|_{\beta,\chi,T} + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T. \tag{1.94}$$

$\diamond$ *Tangential jump*
Set $w_e := F_{ext}(\mathbf{J}_{e,t}) b_e \in [H_0^1(\omega_e)]^2$. It comes from the inverse inequality (1.21) and an elementwise partial integration that

$$\begin{aligned}
\|\mathbf{J}_{e,t}\|_e^2 &\sim \int_e \mathbf{J}_{e,t} w_e \cdot \boldsymbol{t}_e \\
&\sim -\int_e \left[\!\left[\chi \operatorname{curl} \boldsymbol{u}_h\right]\!\right]_e w_e \cdot \boldsymbol{t}_e \\
&\sim \pm \sum_{T\subset\omega_e} \left[\int_T \chi \operatorname{curl} \boldsymbol{u}_h \operatorname{curl} w_e - \int_T \operatorname{curl}(\chi \operatorname{curl} \boldsymbol{u}_h) w_e\right] \\
&\lesssim r(w_e) + \sum_{T\subset\omega_e} \int_T \boldsymbol{R}_T w_e \\
&\lesssim a(\boldsymbol{e}, w_e) + \sum_{T\subset\omega_e} \int_T \boldsymbol{r}_T w_e + \sum_{T\subset\omega_e} \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_e \\
&\lesssim \sum_{T\subset\omega_e} \left[\int_T \chi \operatorname{curl} \boldsymbol{e} \operatorname{curl} w_e + \int_T \beta \boldsymbol{e} w_e + \int_T \boldsymbol{r}_T w_e + \int_T (\boldsymbol{r}_T - \boldsymbol{R}_T) w_e\right] \\
&\lesssim \sum_{T\subset\omega_e} \Big[\|\chi^{\frac{1}{2}} \operatorname{curl} \boldsymbol{e}\|_T \chi_T^{\frac{1}{2}} \| \operatorname{curl} w_e\|_T + \|\beta^{\frac{1}{2}} \boldsymbol{e}\|_T \beta_T^{\frac{1}{2}} \|w_e\|_T \\
&+ \|\boldsymbol{r}_T\|_T \|w_e\|_T + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \|w_e\|_T\Big].
\end{aligned}$$

By the discrete Cauchy-Schwarz inequality and the inverse estimates (1.22),(1.23), we find

$$\|\mathbf{J}_{e,t}\|_e \lesssim \sum_{T\subset\omega_e} h_T^{\frac{1}{2}} \left[\left(\chi_T h_T^{-2} + \beta_T\right)^{\frac{1}{2}} \|\boldsymbol{e}\|_{\beta,\chi,T} + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T\right]. \tag{1.95}$$

Using (1.94), (1.101) and the definition of $\eta_{T,\perp}$, we get :

$$\begin{aligned}
\eta_{T,\perp} &\lesssim \left(\chi_T^{\frac{1}{2}} \beta_T^{-\frac{1}{2}} + h_T\right) \|\boldsymbol{e}\|_{\beta,\chi,T} + h_T \beta_T^{-\frac{1}{2}} \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \\
&+ \sum_{e\subset\partial T} h_e^{\frac{1}{2}} \beta_e^{-\frac{1}{2}} \sum_{T'\subset\omega_e} \left[h_{T'}^{\frac{1}{2}} \left(\chi_{T'}^{\frac{1}{2}} h_{T'}^{-1} + \beta_{T'}^{\frac{1}{2}}\right) \|\boldsymbol{e}\|_{\beta,\chi,T'} + h_{T'}^{\frac{1}{2}} \|\boldsymbol{r}_{T'} - \boldsymbol{R}_{T'}\|_{T'}\right] \\
&\lesssim \sum_{e\subset\partial T} h_e^{\frac{1}{2}} \beta_e^{-\frac{1}{2}} \sum_{T'\subset\omega_e} \left[h_{T'}^{\frac{1}{2}} \left(\chi_{T'}^{\frac{1}{2}} h_{T'}^{-1} + \beta_{T'}^{\frac{1}{2}}\right) \|\boldsymbol{e}\|_{\beta,\chi,T'} + h_{T'}^{\frac{1}{2}} \|\boldsymbol{r}_{T'} - \boldsymbol{R}_{T'}\|_{T'}\right] \\
&\lesssim \sum_{e\subset\partial T} \sum_{T'\subset\omega_e} \left[\left(\chi_{T'}^{\frac{1}{2}} \beta_{T'}^{-\frac{1}{2}} + h_{T'}\right) \|\boldsymbol{e}\|_{\beta,\chi,T'} + h_{T'} \beta_{T'}^{-\frac{1}{2}} \|\boldsymbol{r}_{T'} - \boldsymbol{R}_{T'}\|_{T'}\right] \\
&\lesssim \sum_{T'\subset\omega_T} \left[\left(\chi_{T'}^{\frac{1}{2}} \beta_{T'}^{-\frac{1}{2}} + h_{T'}\right) \|\boldsymbol{e}\|_{\beta,\chi,T'} + h_{T'} \beta_{T'}^{-\frac{1}{2}} \|\boldsymbol{r}_{T'} - \boldsymbol{R}_{T'}\|_{T'}\right].
\end{aligned}$$

■

Corollary 1.4.13. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,0} + \eta_{T,\perp} \lesssim \sum_{T'\subset\omega_T} \left(\chi_{T'}^{\frac{1}{2}} \beta_{T'}^{-\frac{1}{2}} + 1\right) \|e\|_{\beta,\chi,T'} + \sum_{T'\subset\omega_T} \zeta_{T'}. \tag{1.96}$$

Proof of the lower error bound : the solenoidal part - second method

Theorem 1.4.14. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,\perp} \lesssim \|e\|_{\beta,\chi,T} + \sum_{T' \subset \omega_T} \zeta_{T'}. \tag{1.97}$$

Remark 1.4.15. This new estimator has been built in order to have a coefficient in front of $\|\boldsymbol{e}\|_{\beta,\chi}$ equivalent to 1.

Proof:
$\diamond$ *Element residual*
Let T be an element of the triangulation and set $w_T := \boldsymbol{r}_T b_T \in [H_0^1(T)]^2$.
By the definition of α_T, it immediately follows from (1.94)

$$\|\boldsymbol{r}_T\|_T \lesssim \alpha_T^{-1} \|\boldsymbol{e}\|_{\beta,\chi,T} + \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T. \tag{1.98}$$

$\diamond$ *Tangential jump*
Set $w_e := F_{ext}(\mathbf{J}_{e,t}) b_{e,\gamma_e} \in [H_0^1(\omega_e)]^2$ with $\gamma_e \in (0,1]$. For $\omega_e = T_1 \cup T_2$, b_{e,γ_e} is defined as follow

$$b_{e,\gamma_e} := \begin{cases} b_{e,T_1,\gamma_1} & \text{on } T_1 \\ b_{e,T_2,\gamma_2} & \text{on } T_2 \end{cases}$$

and

$$\gamma_e := \begin{cases} \gamma_1 & \text{on } T_1 \\ \gamma_2 & \text{on } T_2 \end{cases}$$

where we choose (see [33])

$$\gamma_i := \frac{1}{2} \chi_{T_i}^{1/2} h_{T_i}^{-1} \alpha_{T_i} = \frac{1}{2} \min\left\{1, \beta_{T_i}^{-1/2} \chi_{T_i}^{1/2} h_{T_i}^{-1}\right\}. \tag{1.99}$$

Note that, $b_{e,T_1,\gamma_{1|e}} = b_{e,T_2,\gamma_{2|e}} = b_{e|e}$. Then, using the argument of section 1.4.3, we have

$$\begin{aligned} \|\mathbf{J}_{e,t}\|_e^2 &\lesssim \sum_{i=1,2} \Big[\|\boldsymbol{e}\|_{\beta,\chi,T_i} \left(\chi_{T_i}^{1/2} \|\operatorname{curl} w_e\|_{T_i} + \beta_{T_i}^{1/2} \|w_e\|_{T_i} \right) \\ &+ \left(\|\boldsymbol{r}_{T_i}\|_{T_i} + \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \right) \|w_e\|_{T_i} \Big]. \end{aligned}$$

By the discrete Cauchy-Schwarz inequality and the inverse estimates (1.25),(1.26), we find

$$\|w_e\|_{\beta,\chi,T_i} \lesssim \gamma_i^{1/2} h_{T_i}^{1/2} \left(\beta_{T_i}^{1/2} + \chi_{T_i}^{1/2} \gamma_i^{-1} h_{T_i}^{-1} \right) \|\mathbf{J}_{e,t}\|_e, \tag{1.100}$$

and (1.94) and (1.100) lead to

$$\begin{aligned} \|\mathbf{J}_{e,t}\|_e &\lesssim \sum_{i=1,2} \Big[\left(\chi_{T_i}^{1/2} h_{T_i}^{-1/2} \gamma_i^{-1/2} + \beta_{T_i}^{1/2} h_{T_i}^{1/2} \gamma_i^{1/2} + \alpha_{T_i}^{-1} h_{T_i}^{1/2} \gamma_i^{1/2} \right)^{1/2} \|\boldsymbol{e}\|_{\beta,\chi,T_i} \\ &+ h_{T_i}^{1/2} \gamma_i^{1/2} \|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i} \Big]. \end{aligned}$$

Then, by (1.99), we obtain

$$\|\mathbf{J}_{e,t}\|_e \lesssim \sum_{i=1,2} \chi_{T_i}^{1/4} \left[\alpha_{T_i}^{-1/2}\|\boldsymbol{e}\|_{\beta,\chi,T_i} + \alpha_{T_i}^{1/2}\|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i}\right]. \tag{1.101}$$

Using (1.94), (1.101) and the definition of $\eta_{T,\perp}$, we get :

$$\begin{aligned}
\eta_{T,\perp} &\lesssim \|\boldsymbol{e}\|_{\beta,\chi,T} + \alpha_T\|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T \\
&+ \sum_{e\subset\partial T} \chi_e^{-1/4}\alpha_e^{1/2} \sum_{i=1,2} \chi_{T_i}^{1/4}\alpha_{T_i}^{-1/2} \left(\|\boldsymbol{e}\|_{\beta,\chi,T_i} + \alpha_{T_i}\|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i}\right) \\
&\lesssim \|\boldsymbol{e}\|_{\beta,\chi,T} + \sum_{e\subset\partial T}\sum_{i=1,2} \chi_{T_i}^{-1/4}\alpha_{T_i}^{1/2}\chi_{T_i}^{1/4}\alpha_{T_i}^{-1/2}\|\boldsymbol{e}\|_{\beta,\chi,T_i} \\
&+ \alpha_T\|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T + \sum_{e\subset\partial T}\sum_{i=1,2} \chi_{T_i}^{-1/4}\alpha_{T_i}^{1/2}\chi_{T_i}^{1/4}\alpha_{T_i}^{1/2}\|\boldsymbol{r}_{T_i} - \boldsymbol{R}_{T_i}\|_{T_i}.
\end{aligned}$$

as $\chi_e \geq \chi_{T_i}$ and $\alpha_e \leq \alpha_{T_i}$. That leads to the conclusion. ∎

Corollary 1.4.16. *For all elements T, the following local lower error bound holds :*

$$\eta_{T,0} + \eta_{T,\perp} \lesssim \|e\|_{\beta,\chi,\omega_T} + \zeta_T. \tag{1.102}$$

Proof of the upper error bound : the irrotational part

Theorem 1.4.17. *The β-norm of the irrotational part of the error is globally bounded from above by η_0, i.e.,*

$$\|\boldsymbol{e}_0\|_\beta \lesssim (1 + C_{Neu}(\beta))\eta_0. \tag{1.103}$$

Proof:
Let $\phi \in H_0^1(\Omega)$ be the function introduced in the Helmholtz decomposition of the error e such that the irrotational part of the error $\boldsymbol{e}_0 = \nabla\phi$ (cf. Corollary 1.4.3). We are interested in $\|e_0\|_\beta = \|e_0\|_{\beta,\chi} = \|\nabla\phi\|_\beta$. By (1.72), we know that

$$a(\boldsymbol{e}_0, \nabla\psi) = (\beta\nabla\phi, \nabla\psi) = r(\nabla\psi),\ \forall\psi \in H_0^1(\Omega).$$

Let $\psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h)$. Then, $\nabla\psi_h \in V_{h,1} \subset V_h$ and the Galerkin orthogonality relation (1.15) gives

$$(\beta\nabla\phi, \nabla\psi) = r(\nabla(\psi - \psi_h)),\ \forall\psi \in H_0^1(\Omega), \psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h).$$

As $\boldsymbol{f}$ is divergence free and $\psi - \psi_h$ belongs to $H_0^1(\Omega)$, we obtain, by Green's formula and an elementwise integration by parts : $\forall\psi \in H_0^1(\Omega), \psi_h \in \mathcal{S}(\Omega, \mathcal{T}_h)$,

$$(\beta\nabla\phi, \nabla\psi) = \sum_{T\in\mathcal{T}_h} \left(\int_T \mathrm{div}(\beta\boldsymbol{u}_h)(\psi - \psi_h) - \sum_{e\subset\partial T}\int_e \mathbf{J}_{e,n}(\psi - \psi_h)\right).$$

Setting $\phi = \psi$ and using Cauchy-Schwarz's inequality give

$$\begin{aligned}(\beta \boldsymbol{e}_0, \boldsymbol{e}_0) &= (\beta\nabla\phi, \nabla\phi) \\ &\leq \sum_{T\in\mathcal{T}_h}\left(\int_T \operatorname{div}(\beta\boldsymbol{u}_h)(\phi-\psi_h) - \sum_{e\subset\partial T}\int_e \mathbf{J}_{e,n}(\phi-\psi_h)\right) \\ &\leq \sum_{T\in\mathcal{T}_h}\Big[h_T\beta_T^{-\frac{1}{2}}\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T h_T^{-1}\beta_T^{\frac{1}{2}}\|\phi-\psi_h\|_T \\ &+ \sum_{e\subset\partial T} h_T h_e^{-\frac{1}{2}}\beta_e^{-\frac{1}{2}}\|\mathbf{J}_{e,n}\|_e h_T^{-1}h_e^{\frac{1}{2}}\beta_e^{\frac{1}{2}}\|\phi-\psi_h\|_e\Big].\end{aligned}$$

We now introduce the notations $\mu_T = h_T\beta_T^{-\frac{1}{2}}$ and $\mu_e = h_e\beta_e^{-1}$.

By the discrete Cauchy-Schwarz inequality we obtain :

$$\begin{aligned}\|\boldsymbol{e}_0\|_\beta^2 &\lesssim \left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^2\|\operatorname{div}(\beta\boldsymbol{u}_h)\|_T^2 + \sum_{e\subset\partial T}\mu_e\|\mathbf{J}_{e,n}\|_e^2\right)\right\}^{\frac{1}{2}} \\ &\qquad\cdot\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2 + \sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{\frac{1}{2}} \\ &\lesssim \eta_0\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2 + \sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{\frac{1}{2}}.\end{aligned}$$

To achieve our estimate (1.103), we choose $\psi_h = \mathrm{I}_{\mathrm{Cl}}^{new}\phi \in \mathcal{S}(\Omega,\mathcal{T}_h)$ and apply (1.75), (1.78) to obtain

$$\left\{\sum_{T\in\mathcal{T}_h}\left(\mu_T^{-2}\|\phi-\psi_h\|_T^2 + \sum_{e\subset\partial T}\mu_e^{-1}\|\phi-\psi_h\|_e^2\right)\right\}^{\frac{1}{2}} \lesssim (1+C_{Neu}(\beta))\|\nabla\phi\|_\beta. \qquad (1.104)$$

■

Proof of the upper error bound : the solenoidal part - first method

Theorem 1.4.18. *The solenoidal part of the error satisfies*

$$\|\boldsymbol{e}_\perp\|_{\beta,\chi} \lesssim \alpha^{-1}\left[(\eta_\perp+\zeta)C_2(\beta,\chi) + (1+C_{Neu}(\beta))\eta_0 C_1(\beta,\chi)\right].$$

Proof: As $\boldsymbol{e}_\perp \in W_\beta$,

$$\begin{aligned}\alpha\|\boldsymbol{e}_\perp\|_{\beta,\chi}^2 &\leq a(\boldsymbol{e}_\perp,\boldsymbol{e}_\perp) \\ &= r(\boldsymbol{e}_\perp) \\ &= r(\boldsymbol{w}) + r(\nabla\phi_0)\end{aligned}$$

where $\phi_0 \in H^1_0(\Omega)$ and $\boldsymbol{w}$ is the function introduced in the Helmholtz decomposition of the solenoidal part of the error $e_\perp$ (cf. Corollary 1.4.3). Inspired by the proof of the irrotational part, we obtain that

$$r(\nabla\phi_0) \lesssim (1 + C_{Neu}(\beta))\eta_0 \|\nabla\phi_0\|_\beta. \tag{1.105}$$

Now, using the Galerkin orthogonality relation (1.15) for any $\boldsymbol{v}_h \in V_h$ and an elementwise integration by parts, we get

$$\begin{aligned} r(\boldsymbol{w}) &= r(\boldsymbol{w} - \boldsymbol{v}_h) \\ &= (\boldsymbol{f}, \boldsymbol{w} - \boldsymbol{v}_h) - a(\boldsymbol{u}_h, \boldsymbol{w} - \boldsymbol{v}_h) \\ &= (\boldsymbol{f} - \beta\boldsymbol{u}_h, \boldsymbol{w} - \boldsymbol{v}_h) \\ &- \sum_{T\in\mathcal{T}_h} \left(\int_T \mathbf{curl}(\chi \operatorname{curl} \boldsymbol{u}_h)(\boldsymbol{w} - \boldsymbol{v}_h) - \sum_{e\subset\partial T} \int_e \chi \operatorname{curl} \boldsymbol{u}_h (\boldsymbol{w} - \boldsymbol{v}_h) \cdot \boldsymbol{t}_T \right) \\ &= \sum_{T\in\mathcal{T}_h} (\boldsymbol{R}_T, \boldsymbol{w} - \boldsymbol{v}_h) - \sum_{e\in\mathcal{E}_h} \int_e \mathbf{J}_{e,t}(\boldsymbol{w} - \boldsymbol{v}_h) \cdot \boldsymbol{t}_e. \end{aligned} \tag{1.106}$$

Cauchy-Schwarz's inequality leads to

$$\begin{aligned} r(\boldsymbol{w}) &\lesssim \sum_{T\in\mathcal{T}_h} \left[\mu_T \|\boldsymbol{R}_T\|_T \mu_T^{-1} \|\boldsymbol{w} - \boldsymbol{v}_h\|_T + \sum_{e\subset\partial T} \mu_e^{\frac{1}{2}} \|\mathbf{J}_{e,t}\|_e \mu_e^{-\frac{1}{2}} \|\boldsymbol{w} - \boldsymbol{v}_h\|_e \right] \\ &\lesssim \left\{ \sum_{T\in\mathcal{T}_h} \left[\mu_T^2 \|\boldsymbol{R}_T\|_T^2 + \sum_{e\subset\partial T} \mu_e \|\mathbf{J}_{e,t}\|_e^2 \right] \right\}^{\frac{1}{2}} \\ &\quad \cdot \left\{ \sum_{T\in\mathcal{T}_h} \left[\mu_T^{-2} \|\boldsymbol{w} - \boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T} \mu_e^{-1} \|\boldsymbol{w} - \boldsymbol{v}_h\|_e^2 \right] \right\}^{\frac{1}{2}} \\ &\lesssim \left\{ \sum_{T\in\mathcal{T}_h} \left[\mu_T^2 \|\boldsymbol{r}_T\|_T^2 + \mu_T^2 \|\boldsymbol{r}_T - \boldsymbol{R}_T\|_T^2 + \sum_{e\subset\partial T} \mu_e \|\mathbf{J}_{e,t}\|_e^2 \right] \right\}^{\frac{1}{2}} \\ &\quad \cdot \left\{ \sum_{T\in\mathcal{T}_h} \left[\mu_T^{-2} \|\boldsymbol{w} - \boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T} \mu_e^{-1} \|\boldsymbol{w} - \boldsymbol{v}_h\|_e^2 \right] \right\}^{\frac{1}{2}}. \end{aligned}$$

Then, by taking $\boldsymbol{v}_h = \mathrm{I}_{\mathrm{Ned}}\boldsymbol{w} \in V_h$ and using the estimates (1.80)-(1.81), we obtain :

$$\left\{ \sum_{T\in\mathcal{T}_h} \left[\mu_T^{-2} \|\boldsymbol{w} - \boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T} \mu_e^{-1} \|\boldsymbol{w} - \boldsymbol{v}_h\|_e^2 \right] \right\}^{\frac{1}{2}} \lesssim \|\nabla_P \boldsymbol{w}\|_\beta. \tag{1.107}$$

Therefore, from the definitions of $\eta_{T,\perp}$ and ζ_T, we find

$$r(\boldsymbol{w}) \lesssim \left(\sum_{T\in\mathcal{T}_h} (\eta_{T,\perp}^2 + \zeta_T^2) \right)^{\frac{1}{2}} \|\nabla_P \boldsymbol{w}\|_\beta. \tag{1.108}$$

By (1.105) and (1.115), we conclude that

$$\alpha\|\boldsymbol{e}_\perp\|^2_{\beta,\chi} \lesssim (\eta_\perp^2+\zeta^2)^{\frac{1}{2}}\|\nabla_P \boldsymbol{w}\|_\beta + (1+C_{Neu}(\beta))\eta_0\|\nabla\phi_0\|_\beta. \tag{1.109}$$

Using the bounds (1.70) and (1.71) from corollary 1.4.3, we get

$$\alpha\|\boldsymbol{e}_\perp\|^2_{\beta,\chi} \lesssim [(\eta_\perp+\zeta)C_2(\beta,\chi) + (1+C_{Neu}(\beta))\eta_0 C_1(\beta,\chi)]\,\|\boldsymbol{e}_\perp\|_{\beta,\chi}.$$

This leads to the conclusion. ∎

Corollary 1.4.19. *The error is globally bounded from above by*

$$\|\boldsymbol{e}\|_{\beta,\chi} \lesssim (1+\alpha^{-1})(\eta+\zeta)\max\left\{(1+C_{Neu}(\beta))C_1(\beta,\chi), C_2(\beta,\chi)\right\}. \tag{1.110}$$

Proof of the upper error bound : the solenoidal part - second method

Theorem 1.4.20. *The following upper bound holds :*

$$\begin{aligned}\|\boldsymbol{e}_\perp\|_{\beta,\chi} \;&\lesssim\; \alpha^{-1}\{(1+C_{Neu}(\beta))C_1(\beta,\chi)\eta_0 \\ &+\; [C_1(\beta,\chi)+(1+C^\star_{Neu}(\beta))C_2(\beta,\chi)\max_{j=1,\dots,J}\{\chi_j^{1/2}\beta_j^{-1/2}\}]\,(\eta_\perp+\zeta)\}\end{aligned} \tag{1.111}$$

Proof: This time, from (1.106), we obtain :

$$\begin{aligned} r(\boldsymbol{w}) \lesssim& \sum_{T\in\mathcal{T}_h}\left[\alpha_T\|\boldsymbol{R}_T\|_T\alpha_T^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T + \sum_{e\subset\partial T}\chi_e^{-1/4}\alpha_e^{1/2}\|\mathbf{J}_{e,t}\|_e\chi_e^{1/4}\alpha_e^{-1/2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e\right] \\ \lesssim& \left\{\sum_{T\in\mathcal{T}_h}\left[\alpha_T^2\|\boldsymbol{r}_T\|_T^2 + \alpha_T^2\|\boldsymbol{r}_T-\boldsymbol{R}_T\|_T^2 + \sum_{e\subset\partial T}\chi_e^{-1/2}\alpha_e\|\mathbf{J}_{e,t}\|_e^2\right]\right\}^{1/2} \\ &\cdot\left\{\sum_{T\in\mathcal{T}_h}\left[\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 + \sum_{e\subset\partial T}\chi_e^{1/2}\alpha_e^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2\right]\right\}^{1/2}. \end{aligned}$$

Then, by taking $\boldsymbol{v}_h = \mathrm{I}^\beta_{\mathrm{CN}}\boldsymbol{w} \in V_h$, we can prove that :

$$\begin{aligned}\sum_{T\in\mathcal{T}_h}[\alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 \;&+\; \sum_{e\subset\partial T}\chi_e^{1/2}\alpha_e^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2] \\ &\lesssim\; \|\boldsymbol{w}\|_\beta^2 + (1+2\;C^\star_{Neu}(\beta))^2\left(\max_{j=1,\dots,J}\{\chi_j\beta_j^{-1}\}\right)^{1/2}\|\nabla_P\boldsymbol{w}\|_\beta^2.\end{aligned} \tag{1.112}$$

Indeed, the definition of α_T implies $\alpha_T^{-1} = \max\{\beta_T^{1/2}, \chi_T^{1/2}h_T^{-1}\}$. It follows, by the esti-

mates (1.82)-(1.83) and the triangular inequality, that

$$\begin{aligned}
\sum_{T\in\mathcal{T}_h} \alpha_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 &= \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\geq \chi_T^{1/2}h_T^{-1}}} \beta_T\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 + \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\leq \chi_T^{1/2}h_T^{-1}}} \chi_T h_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_T^2 \\
&\lesssim \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\geq \chi_T^{1/2}h_T^{-1}}} \left(\|\boldsymbol{w}\|_{\beta,T}^2 + \beta_T\|\boldsymbol{v}_h\|_T^2\right) + \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\leq \chi_T^{1/2}h_T^{-1}}} \chi_T\beta_T^{-1}h_T^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_{\beta,T}^2 \\
&\lesssim \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\geq \chi_T^{1/2}h_T^{-1}}} \left(\|\boldsymbol{w}\|_{\beta,T}^2 + \|\boldsymbol{w}\|_{\beta,\omega_T}^2\right) + \sum_{\substack{T\in\mathcal{T}_h\\ \beta_T^{1/2}\leq \chi_T^{1/2}h_T^{-1}}} \chi_T\beta_T^{-1}C_{Neu}^{\star}(\beta)^2\|\nabla_P\boldsymbol{w}\|_{\beta,\omega_T}^2 \\
&\lesssim \|\boldsymbol{w}\|_\beta^2 + \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right) C_{Neu}^{\star}(\beta)^2\|\nabla_P\boldsymbol{w}\|_\beta^2. \qquad (1.113)
\end{aligned}$$

On the other hand, with the trace inequality (1.18) applied on T_e such that $\beta_e = \beta_{T_e}$ and the estimates (1.113), (1.84) and (1.85), we find

$$\begin{aligned}
\sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T} \chi_e^{1/2}\alpha_e^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2 &= \sum_{T\in\mathcal{T}_h} \sum_{\substack{e\subset\partial T\\ \beta_e^{1/2}\geq \chi_e^{1/2}h_e^{-1}}} \chi_e^{1/2}\beta_e^{-1/2}\beta_e\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2 \\
&+ \sum_{T\in\mathcal{T}_h} \sum_{\substack{e\subset\partial T\\ \beta_e^{1/2}\leq \chi_e^{1/2}h_e^{-1}}} \chi_e h_e^{-1}\beta_e^{-1}\beta_e\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2 \\
&\lesssim \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right)^{1/2} \sum_{T_e\in\mathcal{T}_h} \beta_{T_e}^{1/2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_{T_e}\cdot \\
&\quad \left(\beta_{T_e}^{1/2}h_{T_e}^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_{T_e} + \beta_{T_e}^{1/2}\|\nabla(\boldsymbol{w}-\boldsymbol{v}_h)\|_{T_e}\right) \\
&+ \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right) \sum_{T_e\in\mathcal{T}_h} (C_{Neu}^{\star}(\beta)+1)^2\|\nabla_P\boldsymbol{w}\|_{\beta,\omega_T}^2 \\
&\lesssim \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right)^{1/2} \left(\sum_{T_e\in\mathcal{T}_h} \beta_{T_e}\|\boldsymbol{w}-\boldsymbol{v}_h\|_{T_e}^2\right)^{1/2} \cdot \\
&\quad \left(\sum_{T_e\in\mathcal{T}_h} \left(\beta_{T_e}h_{T_e}^{-2}\|\boldsymbol{w}-\boldsymbol{v}_h\|_{T_e}^2 + \beta_{T_e}\|\nabla(\boldsymbol{w}-\boldsymbol{v}_h)\|_{T_e}^2\right)\right)^{1/2} \\
&+ \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right) \sum_{T_e\in\mathcal{T}_h} (C_{Neu}^{\star}(\beta)+1)^2\|\nabla_P\boldsymbol{w}\|_{\beta,\omega_T}^2 \\
&\lesssim \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right)^{1/2} \|\boldsymbol{w}\|_\beta(1+C_{Neu}^{\star}(\beta))\|\nabla_P\boldsymbol{w}\|_\beta \\
&+ \left(\max_{j=1,\ldots,J}\{\chi_j\beta_j^{-1}\}\right) (C_{Neu}^{\star}(\beta)+1)^2\|\nabla_P\boldsymbol{w}\|_\beta^2
\end{aligned}$$

$$\sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T}\chi_e^{1/2}\alpha_e^{-1}\|\boldsymbol{w}-\boldsymbol{v}_h\|_e^2 \lesssim \|\boldsymbol{w}\|_\beta^2+\left(\max_{j=1,\dots,J}\{\chi_j\beta_j^{-1}\}\right)(1+C^\star_{Neu}(\beta))^2\|\nabla_P\boldsymbol{w}\|_\beta^2 \quad (1.114)$$

The estimates (1.113) and (1.114) show (1.112). Therefore, from the definitions of $\eta_{T,\perp}$ and ζ_T and the estimate (1.112), we deduce

$$r(\boldsymbol{w}) \lesssim (\eta_\perp+\zeta)\left\{\|\boldsymbol{w}\|_\beta+\|\nabla\boldsymbol{w}\|_\beta(1+C^\star_{Neu}(\beta))\left(\max_{j=1,\dots,J}\{\chi_j\beta_j^{-1}\}\right)^{1/2}\right\}. \quad (1.115)$$

Using the bound (1.68) from the Helmholtz decomposition, we get the conclusion. ■

Corollary 1.4.21. *The error is globally bounded from above by*

$$\begin{aligned}\|\boldsymbol{e}\|_{\beta,\chi} \lesssim\ & (1+\alpha^{-1})\max\{(1+C_{Neu}(\beta))C_1(\beta,\chi),\\ & C_1(\beta,\chi)+(1+C^\star_{Neu}(\beta))\max_{j=1,\dots,J}\{\chi_j^{1/2}\beta_j^{-1/2}\}C_2(\beta,\chi)\}\quad(\eta+\zeta).\end{aligned} \quad (1.116)$$

1.4.4 Extension to three-dimensional polyhedral domains

All the results of this paper extend to a three-dimensional polyhedral domain O which is bounded and simply connected with a connected boundary. In that domain we consider the Maxwell system (1.3), where $\boldsymbol{f}$ satisfies (1.4) and β and χ are as before.

This problem is then approximated using regular meshes made of tetrahedra and the finite element space V_h is simply assumed to contain lowest order Nédélec elements.

In this setting all the results from section 1.2.2 remain valid, especially Lemma 2.4.1 (the Helmholtz decomposition) due to the results from [45, 47]. Moreover in 3D the Clément-Nédélec interpolant is defined by

$$\begin{aligned}\mathrm{I}_{\mathrm{CN}} : L^2(O)^3 &\to V_h\\ \boldsymbol{u} &\to \sum_{e\in\mathcal{E}_{h\Omega}}\alpha_e(\boldsymbol{u})|e|\lambda_e\end{aligned}$$

where, as usual $\mathcal{E}_{h\Omega}$ is the set of interior edges of the mesh, λ_e is the standard basis function of lowest order Nédélec elements and we here set $\alpha_e(\boldsymbol{u})=\dfrac{1}{|T_e|}\displaystyle\int_{T_e}\boldsymbol{u}\cdot\boldsymbol{t}_e$, when T_e is a tetrahedron having e as edge such that $\beta_{T_e}=\max_{e\subset T}\beta_T$. The regularity of the mesh allows then to show that Theorem 1.4.9 holds.

As the basic tools of section 1.2.2, the interpolation error estimates from section 1.4.2 and some integrations by parts are the only ingredients that we used for the proof of the lower and upper error bounds, we can conclude that the estimates (1.97), (1.102), (1.110) and (1.116) hold in 3D, with the same definition for the local estimators, except that $\mathbf{J}_{e,n}$ and $\mathbf{J}_{e,t}$ are defined for the faces F of the mesh and for the tangential jump where curl $\boldsymbol{u}_h$ is replaced by $\mathbf{curl}\,\boldsymbol{u}_h\times\boldsymbol{n}_F$, see section 4.1 of [45].

1.4.5 Numerical experiments

The following experiments underline and confirm our theoretical predictions. Our examples consist in solving the Maxwell equation (1.8) on the unit square $\Omega = (0,1)^2$ with different values of χ and β. In all examples uniform meshes of size $h = \frac{1}{n}, n = 32, 64, 128$ and the lowest order Nédélec finite elements are used. Both estimators are tested and compared. For that purpose, when an exact solution is known we analyze the upper and lower error bounds for each estimator. In order to present them in an appropriate manner, we consider the ratios

$$\begin{aligned} q_{up} &= \frac{\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi}}{\eta + \xi}, \\ q_{low} &= \max_{T \in \mathcal{T}_h} \frac{\eta_{T,0} + \eta_{T,\perp}}{\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\beta,\chi,\omega_T} + \zeta_T}, \end{aligned}$$

as a function of the parameters β and χ. The first ratio q_{up}, the so-called effectivity index, is related to the global upper error bound and measures the reliability of the estimator. The second ratio is related to the local lower error bound and measures the efficiency of the estimator. The theoretical bounds for q_{up} and q_{low} of the previous sections are summarized in Table 1.1.

	$1^{st} method$	$2^{nd} method$
q_{low}	$1 + \max_{T \in \mathcal{T}_h} \left\{ \chi_T^{1/2} \beta_T^{-1/2} \right\}$	1
q_{up}	$\max\{(1 + C_{Neu}(\beta)) C_1(\beta,\chi), C_2(\beta,\chi)\}$	see (1.116)

TAB. 1.1 – Bounds for q_{low} and q_{up} for both methods

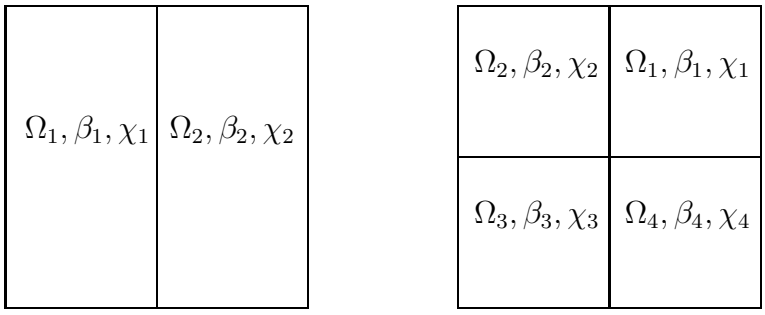


FIG. 1.13 – The decomposition of the domain Ω on 2 subdomains, on the left, and on 4 subdomains, on the right.

In the first example, we suppose that Ω admits a decomposition in two subdomains $\Omega_1 = (0, 1/2)\times(0, 1), \Omega_2 = (1/2, 1)\times(0, 1)$ (see Figure 1.13 left) and take as exact solution :

$$\boldsymbol{u} = \operatorname{curl}\varphi, \text{ where } \varphi = [y(1-y)x(1-x)(2x-1)]^2.$$

In that case, one can show that $C_{Neu}(\beta) \lesssim 1$, $C^{\star}_{Neu}(\beta) \lesssim 1$ and that $W_\beta \subset H^1(\Omega)^2$, so that we can take $C_1(\beta, \chi) = 1$ and $C_2(\beta, \chi)$ as before. Therefore the bounds of q_{up} and q_{low} are simpler and Table 1 is reduced to the next Table 2 :

	$1^{st}method$	$2^{nd}method$
q_{low}	$1 + \max\limits_{T\in\mathcal{T}_h}\left\{\chi_T^{1/2}\beta_T^{-1/2}\right\}$	1
q_{up}	$\max\left\{1, \max\limits_{j=1,\dots,J}\left(\beta_j\chi_j^{-1}\right)\right\}$	$\max\left\{1, \sqrt{\max\limits_{j=1,\dots,J}\left(\beta_j\chi_j^{-1}\right)\max\limits_{j=1,\dots,J}\left(\chi_j\beta_j^{-1}\right)}\right\}$

TAB. 1.2 – Bounds for q_{low} and q_{up} for example 1

In a first case, we fix $\chi_1 = \chi_2 = \beta_2 = 1$ and take different values of β_1. The ratios q_{up} and q_{low} are presented in Figure 1.14 for the first estimator and in Figure 1.15 for the second estimator for different values of β_1. To see more easily the dependence on the involved parameters, all figures are plotted in a double logarithmic scale. For the first estimator, we see that q_{up} behaves like $\sqrt{\beta_1}$ for $\beta_1 \leq 1$ and is mainly constant for $\beta_1 \geq 1$, while q_{low} has a slow variation. For the second estimator, we can say that q_{up} and q_{low} remain constant for any β_1. In both cases, our numerical bounds are better than the theoretical ones (see Table 3).

Now, we fix $\beta_1 = 4$, $\beta_2 = 1$ and take different values of $\chi_1 = \chi_2 = \chi$. The ratios q_{up} and q_{low} are presented for the first (resp. second) estimator in Figure 1.16 (resp. 1.17) for different values of χ. For the first estimator, we see that q_{up} behaves like $\chi^{-1/2}$ for $\chi \geq 1$, while is slightly decreasing as $\chi \leq 1$ decreases. As before we also remark that q_{low} presents slow variations. For the second estimator, again we can say that q_{up} and q_{low} remain quasi constant for any χ. As before, our numerical bounds are better than the theoretical ones (see Table 4).

As second example, we suppose that Ω admits a decomposition into four subdomains $\Omega_i, i = 1, 2, 3, 4$ as shown in Figure 1.13 (right) and introduce the exact solution

$$\boldsymbol{u} = \operatorname{curl}\varphi \text{ where } \varphi = [y(1-y)(2y-1)x(1-x)(2x-1)]^2.$$

We fix $\chi_i = 1$, for all $i = 1, \dots, 4$, $\beta_2 = \beta_4 = 1$ and take $\beta_1 = \beta_3 = \varepsilon$ for different values of ε. For this example, the corner point $S = (1/2, 1/2)$ induces a singularity for any element

Example 1.1	$1^{st}method$	$2^{nd}method$
q_{low}	$1+\max\left\{\beta_1^{-1/2},1\right\}$	1
q_{up}	$\max\{1,\beta_1\}$	$\max\left\{\sqrt{\beta_1},\sqrt{\beta_1^{-1}}\right\}$

TAB. 1.3 – Bounds for q_{low} and q_{up} for example 1.1

Example 1.2	$1^{st}method$	$2^{nd}method$
q_{low}	$1+\sqrt{\chi}$	1
q_{up}	$\max\{1,\chi^{-1}\}$	1

TAB. 1.4 – Bounds for q_{low} and q_{up} for example 1.2

in W_β (see [22]). The constants involved in the bounds of q_{up} can be estimated, namely, $C_{Neu}(\beta)\sim C^\star_{Neu}(\beta)\sim\max\{\sqrt{\epsilon},\frac{1}{\sqrt{\epsilon}}\}$. Therefore the theoretical bounds are easily estimated. The computed ratios q_{up} and q_{low} are presented in Figure 1.18 (resp. in Figure 1.19) for the first estimator (resp. for the second estimator). For the first estimator, we see that q_{up} behaves like $\sqrt{\varepsilon}$ for $\varepsilon\leq 1$ and is mainly constant for $\varepsilon\geq 1$, while q_{low} varies slowly. For the second estimator, we can say that q_{up} and q_{low} remain constant for any ε. In all cases, the numerical bounds are quite better than the theoretical ones.

As a third example, we consider the problem from examples 1.1 and 1.2 with datum $\mathbf{f}=(1,0)^\top$ and for which no exact solution is known. To compare our two estimators we then have computed the ratio η_{Ned}/η_{CN}, where clearly η_{Ned} (resp. η_{CN}) is the estimator of the first (resp. second) method. From Tables 3 and 4, we can obtain theoretical bounds for this ratio that are presented in Table 1.5. The numerical values of this ratio are plotted in Figure 1.20. There we can see that the ratio tends to 1 for β_1 large or χ small, while for β_1 small, the ratio behaves like $\beta_1^{-1/2}$ (better than the upper bound), while for χ large, the ratio behaves like $\sqrt{\chi}$ as the theoretical upper bound. Let us further remark that these results are in accordance with the results presented in Figures 5 to 8.

Note that for the second example with the right-hand side $\mathbf{f}=(1,0)^\top$, the ratio η_{Ned}/η_{CN} behaves like 1 for ϵ large and like $\epsilon^{-1/2}$ for ϵ small. Again these results are in accordance with the ones from Figures 9 and 10.

From these numerical experiments, we can conclude that our second estimator is strongly robust with respect to the variation of the parameters, while the first one is less stable. Surprisingly in all our tests, the first estimator is also stable in the singular perturbation

case (i. e. the case β_1 large or χ small for examples 1 and 3 and the case ε for example 2). This can be justified by the fact that in these cases the contribution of the jumps of $\chi \operatorname{curl} \boldsymbol{u}_h$ in the estimators is too small.

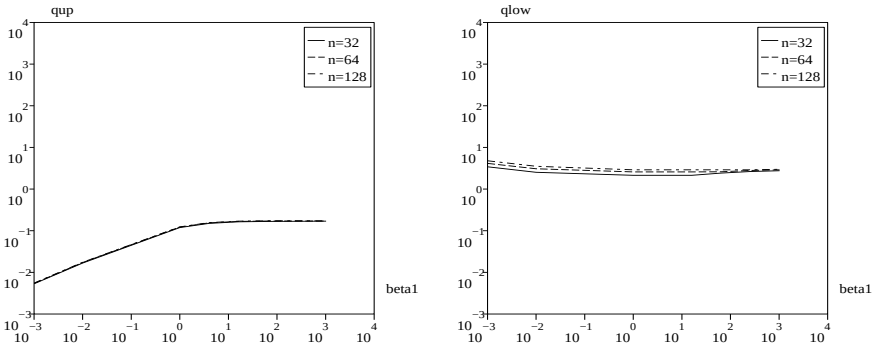


FIG. 1.14 – q_{up} and q_{low} as a function of β_1 for example 1 and for the first estimator.

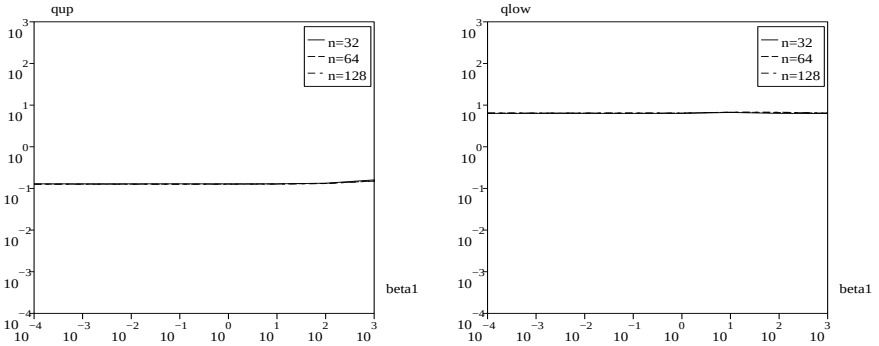


FIG. 1.15 – q_{up} and q_{low} as a function of β_1 for example 1 and for the second estimator.

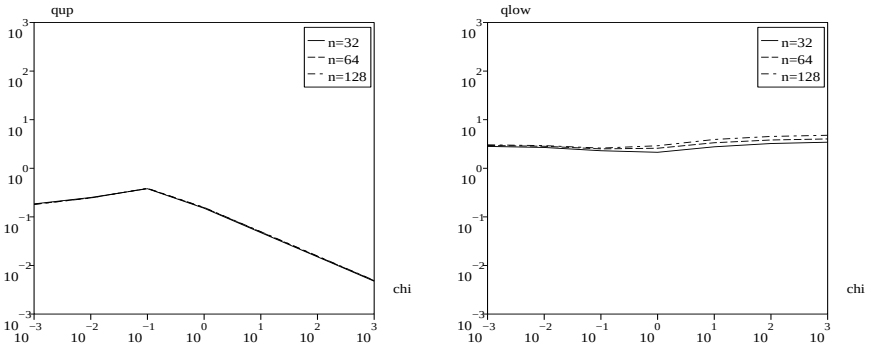


FIG. 1.16 – q_{up} and q_{low} as a function of χ for example 1 and for the first estimator.

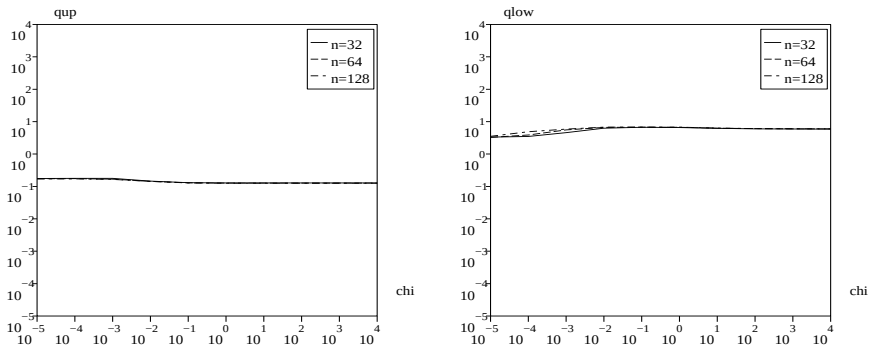


FIG. 1.17 – q_{up} and q_{low} as a function of χ for example 1 and for the second estimator.

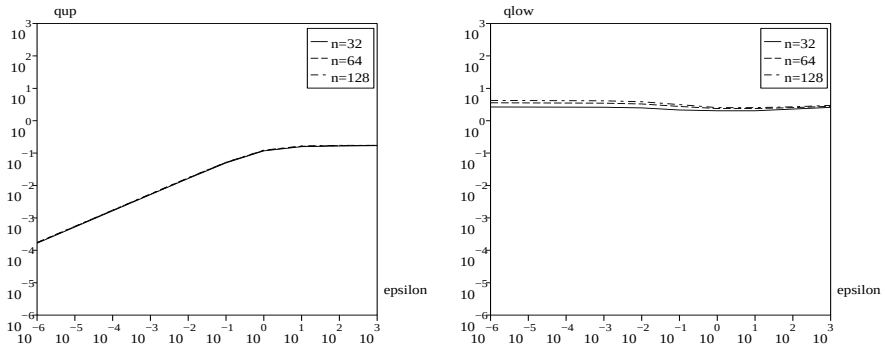


FIG. 1.18 – q_{up} and q_{low} as a function of ε for example 2 and for the first estimator.

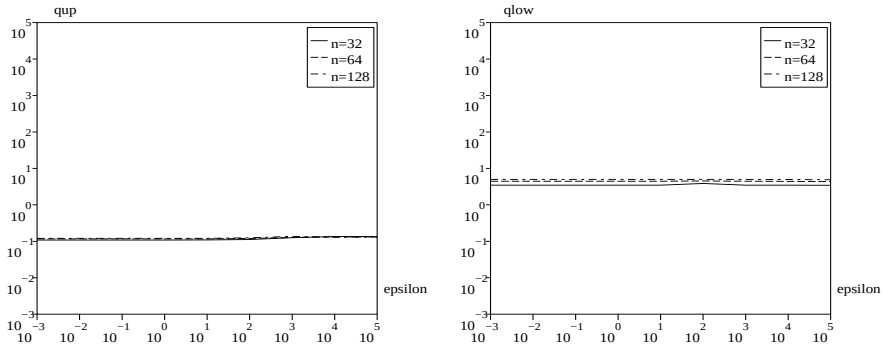


FIG. 1.19 – q_{up} and q_{low} as a function of ε for example 2 and for the second estimator.

1.4.6 Conclusion

We have proposed and rigorously analysed a posteriori error estimators of residual type for the Maxwell equations in a bounded two (and three) dimensional domain using conforming finite element spaces of Nédélec type. A new interpolant of Clément/Nédélec type has been introduced and some interpolation error estimates have been proved. We

η_{Ned}/η_{CN}	Lower Bound	Upper Bound
$\beta_1 \leq 1$ $\chi = 1 = \beta_2$	1	β_1^{-1}
$\beta_1 \geq 1$ $\chi = 1 = \beta_2$	β_1^{-1}	$\sqrt{\beta_1}$
$\chi \leq 1$ $\beta_1 = 4$ $\beta_2 = 1$	χ	$1 + \sqrt{\chi}$
$\chi \geq 1$ $\beta_1 = 4$ $\beta_2 = 1$	1	$\sqrt{\chi}$

TAB. 1.5 – Bounds for η_{Ned}/η_{CN} for examples 1.1 and 1.2

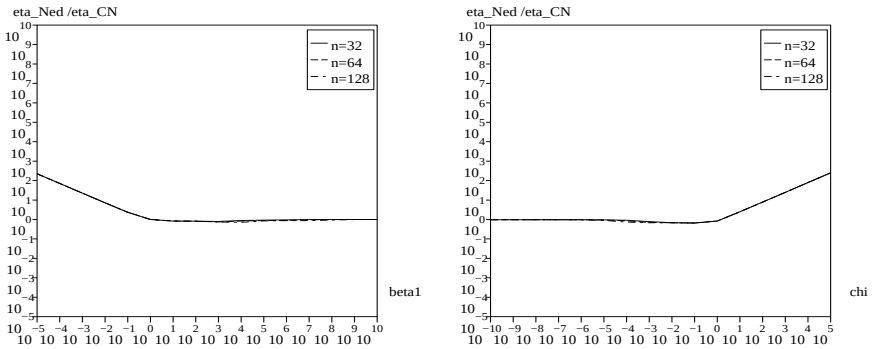


FIG. 1.20 – The ratio η_{Ned}/η_{CN} for 2 subdomains, for different values of β_1 on the left, and for different values of χ, on the right.

have shown that our estimators are reliable and efficient and have explicitly given the dependence of the bounds with respect to the parameters. We further have shown that the second estimator is robust. Some numerical experiments confirm our theoretical predictions.

Chapitre 2

A posteriori error estimators based on equilibrated fluxes

We consider conforming finite element approximations of reaction-diffusion problems and time-harmonic Maxwell equations. We propose new a posteriori error estimators based on $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming finite elements and equilibrated fluxes. It is shown that these estimators give rise to an upper bound where the constant is one up to higher order terms. Lower bounds can also be established with constants depending on the shape regularity of the mesh and the local variation of the coefficients. The reliability and efficiency of the proposed estimator are confirmed by various numerical tests.

2.1 Introduction

Among other methods, the finite element method is widely used for the numerical approximation of partial differential equations, see, e.g., [13–15, 17, 39]. In many engineering applications, adaptive techniques based on a posteriori error estimators have become an indispensable tool to obtain reliable results. Nowadays there exists a vast amount of literature on locally defined a posteriori error estimators for problems in structural mechanics or electromagnetism. We refer to the monographs [3, 6, 40, 52] for a good overview on this topic. In general, local upper and lower bounds are established in order to guarantee the reliability and the efficiency of the proposed estimator. Most of the existing approaches involve constants depending on the shape regularity of the elements and/or of the jumps in the coefficients ; but these dependencies are often not given. Only a few number of approaches gives rise to estimates with explicit constants, see, e.g., [3, 13, 35, 38, 42, 46]. For Maxwell's system, only relatively few results exist. Different well established approaches, for the Laplace operator, have been generalized and adapted to this special situation. Residual type error estimators which measure the jump of the discrete flux have been considered in [9, 19, 39, 45, 49] ; hierarchical error estimators e.g. in [8], and estimators based on the solution of local problems have been introduced in [29]. Here we use an approach based on equilibrated fluxes and $H(\mathrm{div})$- or $H(\mathrm{curl})$-conforming elements. Similar ideas can be

found, e.g., in [13,38,46]. For an overview on equilibration techniques, we refer to [3,35]. For reaction-diffusion problems, in contrast to [13], we first define on the edges an equilibrated flux and then a $H(\mathrm{div})$-conforming element being locally conservative by construction. In [13], the authors directly compute suitable conforming elements by solving local Neumann problems. On the contrary for Maxwell's system the construction of equilibrated fluxes seems to be impossible and therefore we use the construction from [13]. In both cases, the error estimator is locally defined and yields, up to higher order terms, an upper bound with constant one for the discretization error. We note that our error estimators are made for partial differential equations with zero order terms, and the upper bound one is still valid in this more general situation. Special care is required by the lower order terms. In the case of Maxwell's equations, we have to introduce a second approximation that takes into account the non-fulfilment of the divergence constraint of the finite element approximation. This second approximation has not to be introduced if the zero order term is not present. Finally lower bounds are proved, moreover for reaction-diffusion problems, we trace the dependency of the constants with respect to the variation of the coefficients for all proposed estimators. For Maxwell's system this dependency is partially given.

The outline of the chapter is as follows : We recall, in Section 2, the scalar reaction-diffusion problem and its numerical approximation. Section 3 is devoted to the introduction of the locally defined error estimators based on Raviart–Thomas or Brezzi–Douglas–Marini (BDM) elements and the proofs of the upper and lower bounds. The upper bound directly follows from the construction of the estimators, while the proof of the lower bound relies on suitable norm equivalences and some properties of the equilibrated fluxes. Finally in Section 4, we treat the time-harmonic Maxwell equations. For both problem classes, some numerical tests are presented that confirm the reliability and efficiency of our error estimators.

2.2 The two-dimensional reaction-diffusion equation

Let Ω be a bounded domain of $\mathbb{R}^2$ and Γ its polygonal boundary. We consider the following elliptic second order boundary value problem with homogeneous mixed boundary conditions :

$$\begin{aligned} -\mathrm{div}\,(a\,\nabla u) + u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ a\nabla u \cdot n &= 0 \quad \text{on } \Gamma_N, \end{aligned} \tag{2.1}$$

where $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$.

In the sequel, we suppose that a is piecewise constant, namely we assume that there exists a partition $\mathcal{P}$ of Ω into a finite set of Lipschitz polygonal domains $\Omega_1, \cdots, \Omega_J$ such that, on each Ω_j, $a = a_j$ where a_j is a positive constant. For simplicity of notation, we assume that Γ_D has a non-vanishing measure. The variational formulation of (4.1) involves the bilinear form

$$B(u,v) = \int_\Omega (a\nabla u \cdot \nabla v + uv)\,.$$

Given $f \in L^2(\Omega)$, the weak formulation consists in finding $u \in H^1_D(\Omega) := \{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_D\}$ such that

$$B(u, v) = (f, v) = \int_\Omega fv, \ \forall v \in H^1_D(\Omega). \tag{2.2}$$

We consider a triangulation $\mathcal{T}_h$ made of triangles T whose edges are denoted by e and assume that this triangulation is shape-regular, i.e., for any element T, the ratio h_T/ρ_T is bounded by a constant $\sigma > 0$ independent of $T \in \mathcal{T}_h$ and of the mesh-size $h = \max_{T\in\mathcal{T}_h} h_T$, where h_T is the diameter of T and ρ_T the diameter of its largest inscribed ball. We further assume that $\mathcal{T}_h$ is conforming with the partition $\mathcal{P}$ of Ω, i.e., any $T \in \mathcal{T}_h$ is included in one and only one Ω_i. With each edge e of the triangulation, we associate a fixed unit normal vector n_e, and n_T stands for the outer unit normal vector of T. For boundary edges $e \subset \partial\Omega \cap \partial T$, we set $n_e = n_T$. $\mathcal{E}_h$ represents the set of edges of the triangulation, and we assume that the Dirichlet boundary can be written as union of edges. In the sequel, a_T denotes the value of the piecewise constant coefficient a restricted to the element T.

In the following, the L^2-norm on a subdomain D will be denoted by $\|\cdot\|_D$; the index will be dropped if $D = \Omega$. We use $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$ to denote the standard norm and semi-norm on $H^s(D)$ $(s \geq 0)$, respectively. The energy norm is defined by $|||v|||^2 = B(v, v)$, for any $v \in H^1(\Omega)$. Finally, the notation $r \lesssim s$ and $r \sim s$ means the existence of positive constants C_1 and C_2, which are independent of the mesh size, of the coefficients of the partial differential equation and of the quantities r and s such that $r \lesssim C_2 s$ and $C_1 s \lesssim r \lesssim C_2 s$, respectively.

Problem (4.2) is approximated by a conforming finite element subspace of $H^1_D(\Omega)$:

$$X_h = \left\{v_h \in H^1_D(\Omega) | v_{h|T} \in \mathcal{P}_1(T), T \in \mathcal{T}_h\right\}$$

and the finite element solution $u_h \in X_h$ satisfies the discretized problem

$$B(u_h, v_h) = (f, v_h), \ \forall v_h \in X_h. \tag{2.3}$$

For further purposes we introduce a set of fluxes $\{g_e \in \mathcal{P}_1(e) | e \in \mathcal{E}_h\}$ that satisfy the local variational problem

$$B_T(u_h, v_h) = \int_T f\, v_h + \int_{\partial T} g_T v_h, \ \forall v_h \in \mathcal{P}_1(T), \ T \in \mathcal{T}_h, \tag{2.4}$$

where $B_T(\cdot, \cdot)$ represents the local contribution of the bilinear form $B(\cdot, \cdot)$ on the element T and $g_{T|e} = g_e\ n_e \cdot n_T$. The existence of such fluxes is guaranteed and g_e can be locally constructed in terms of its moments and the solution of a local vertex based system, see, e.g., [3, 38]. We note that g_e approximates the flux of the exact solution and thus we set $g_e = 0$, if $e \subset \Gamma_N$.

2.3 Upper and lower bounds for the error estimator

Error estimators can be constructed in many different ways as, for example, using residual type error estimators which measure locally the jump of the discrete flux. A different method, based on equilibrated fluxes, consists in solving local Neumann boundary value problems [3]. Here, introducing the flux as auxiliary variable, we locally define an error estimator based on a $H(\mathrm{div})$-conforming approximation of this variable. This method avoids solving the supplementary above-mentioned local subproblems. Indeed in many applications, the flux $j = a\nabla u$ is an important quantity, and introducing this auxiliary variable, we transform the original problem (4.2) into a first order system. Its weak formulation gives rise to the following saddle point problem : Find $(j, u) \in H_N(\mathrm{div}\,, \Omega) \times L^2(\Omega)$ such that

$$\int_\Omega a^{-1} j\ \tau + \int_\Omega \mathrm{div}\,\tau\ u \ = \ 0,\ \forall \tau \in H_N(\mathrm{div}\,, \Omega), \tag{2.5}$$

$$\int_\Omega \mathrm{div}\, j\ w - \int_\Omega u\ w \ = \ -\int_\Omega f\ w,\ \forall w \in L^2(\Omega), \tag{2.6}$$

the natural space for the flux being

$$H_N(\mathrm{div}\,, \Omega) = \left\{ q \in [L^2(\Omega)]^2 | \mathrm{div}\, q \in L^2(\Omega) \text{ and } q \cdot n = 0 \text{ on } \Gamma_N \right\}.$$

Therefore the discrete flux approximation j_h will be searched in a $H(\mathrm{div})$-conforming space based on standard mixed finite elements. Hence different error estimators can be defined in terms of different mixed finite element spaces such as, e.g., Raviart–Thomas finite elements or BDM elements. Here, for simplicity we only consider low order finite elements but all ideas can be easily generalized to higher order finite elements. We consider three different cases and introduce the inf-sup stable pairs (V_h^i, W_h^i), $i = 1, 2, 3$ by

$$\begin{aligned} V_h^i &= \left\{ v_h \in H_N(\mathrm{div}\,, \Omega) |\, v_{h|T} \in V^i(T), T \in \mathcal{T}_h \right\}, \\ W_h^i &= \left\{ w_h \in L^2(\Omega) |\, w_{h|T} \in W^i(T), T \in \mathcal{T}_h \right\}, \end{aligned}$$

where $V^1(T) = RT_0(T)$, $V^2(T) = BDM_1(T)$, $V^3(T) = RT_1(T)$ and $W^1(T) = W^2(T) = \mathcal{P}_0(T)$, $W^3(T) = \mathcal{P}_1(T)$. Here, we use the definition of the local Raviart–Thomas and BDM elements $RT_l(T) = (\mathcal{P}_l(T))^2 + \mathcal{P}_l(T)\boldsymbol{x}$, $l = 0, 1$ and $BDM_1 = (\mathcal{P}_1(T))^2$. We note that $V^1(T) \subset V^2(T) \subset V^3(T)$. Then it is well known, see, e.g., [15] that $\mathrm{div}\, V_h^i = W_h^i$. We denote by Π_h^i the L^2-projection onto W_h^i. Now we introduce a locally defined flux $j_h^i \in V_h^i$. It is uniquely defined in terms of its degrees of freedom and can be determined with the help of g_e and u_h :

– $i = 1$: for all edges $e \in \mathcal{E}_h$

$$\int_e j_h^1 \cdot n_e = \int_e g_e,$$

– $i = 2$: for all edges $e \in \mathcal{E}_h$

$$\int_e j_h^2 \cdot n_e q = \int_e g_e q, \quad \forall q \in \mathcal{P}_1(e),$$

– $i = 3$ for all edges $e \in \mathcal{E}_h$ and all elements $T \in \mathcal{T}_h$

$$\int_e j_h^3 \cdot n_e q = \int_e g_e q, \quad \forall q \in \mathcal{P}_1(e), \quad \int_T j_h^3 \nabla w = \int_T a \nabla u_h \nabla w, \quad \forall w \in \mathcal{P}_1(T).$$

The global error estimator η_h^i is now given in terms of its elementwise contributions, i.e., $(\eta_h^i)^2 = \sum_{T \in \mathcal{T}_h} (\eta_T^i)^2$, where η_T^i is given by means of j_h^i and Π_h^i :

$$\eta_T^i = \eta_{T;1}^i + \eta_{T;0}^i, \quad \eta_{T;1}^i = \|a^{-\frac{1}{2}}(a\nabla u_h - j_h^i)\|_T, \quad \eta_{T;0}^i = \alpha_T \|u_h - \Pi_h^i u_h\|_T, \tag{2.7}$$

where $\alpha_T = \min\{1, h_T a_T^{-1/2}\}$. We note that if h_T tends to zero, the minimum will be given by $h_T a_T^{-1/2}$. Observing that $\Pi_h^3 u_h = u_h$, $\eta_{T,0}^3 = 0$. To get suitable bounds, we have to consider additionally the data oscillation given by

$$(\mathrm{osc}_i(f))^2 = \sum_{T \in \mathcal{T}_h} \alpha_T^2 \|f - \Pi_h^i f\|_T^2.$$

Remark 2.3.1. *If f is smooth, $osc_i(f)$ is asymptotically a higher order term and thus can be neglected asymptotically. We note that for $a_T \ll 1$ and coarse meshes the case $i = 3$ might be more attractive than the cases $i = 1, 2$.*

2.3.1 Upper bound for the discretization error

The proof of the upper bound is basically based on the observation that all our fluxes j_h^i are $H(\mathrm{div})$-conforming elements and on the following projection lemma.

Lemma 2.3.2. $\mathrm{div}\, j_h^i - \Pi_h^i u_h = -\Pi_h^i f$.

Proof: We start with the observation that $\mathrm{div}\, V_h^i = W_h^i$. Using the definition (2.4) of g_e and of j_h^i, we find for $w \in W_h^i$

$$\begin{aligned} \int_\Omega (\mathrm{div}\, j_h^i - \Pi_h^i u_h) w &= \sum_{T \in \mathcal{T}_h} \left(\int_{\partial T} j_h \cdot n_T w - \int_T j_h^i \nabla w - \int_T u_h w \right) \\ &= \sum_{T \in \mathcal{T}_h} \left(\int_{\partial T} g_T w - \int_T a \nabla u_h \nabla w - \int_T u_h w \right) = -(f, w). \end{aligned}$$

■

Theorem 2.3.3. *The energy norm of the discretization error is bounded by the estimator η_h^i, $i = 1, 2, 3$, and the data oscillation, namely*

$$|||\, u - u_h \,||| \leq \eta_h^i + osc_i(f). \tag{2.8}$$

Proof: Using the definition of the energy norm, inserting the $H(\mathrm{div})$-conforming flux, applying Green's formula and Lemma 2.3.2, we find

$$\begin{aligned}
|||u-u_h|||^2 &= \int_\Omega a\nabla(u-u_h)\nabla(u-u_h) + \int_\Omega (u-u_h)(u-u_h) \\
&= \int_\Omega (j_h^i - a\nabla u_h)\nabla(u-u_h) + \int_\Omega (\Pi_h^i u_h - u_h)(u-u_h) + \int_\Omega (f-\Pi_h^i f)(u-u_h).
\end{aligned}$$

Cauchy–Schwarz's inequality yields

$$\int_\Omega (j_h^i - a\nabla u_h)\nabla(u-u_h) \le \sum_{T\in\mathcal{T}_h} \|a^{-\frac{1}{2}}(j_h^i - a\nabla u_h)\|_T |||u-u_h|||_T = \sum_{T\in\mathcal{T}_h} \eta_{T,1}^i |||u-u_h|||_T,$$

where $|||\cdot|||_T$ stands for the contribution of the energy norm restricted to the element T. We note that $\|w-\Pi_h^i w\|_T \le \|w-\Pi_h^1 w\|_T \le h_T\|\nabla w\|_T$, $w\in H^1(T)$, see, e.g., Lemma 3.5 of [44]. Then it is easy to see that the second and the third term can be bounded by

$$\begin{aligned}
\int_\Omega (\Pi_h^i u_h - u_h)(u-u_h) &\le \sum_{T\in\mathcal{T}_h} \alpha_T \|u_h - \Pi_h^i u_h\|_T |||u-u_h|||_T = \sum_{T\in\mathcal{T}_h} \eta_{T;0}^i |||u-u_h|||_T, \\
\int_\Omega (f-\Pi_h^i f)(u-u_h) &\le \sum_{T\in\mathcal{T}_h} \alpha_T \|f-\Pi_h^i f\|_T |||u-u_h|||_T \le \mathrm{osc}_i(f)\,|||u-u_h|||,
\end{aligned}$$

respectively. Taking into account the definition of η_h^i, we find

$$\begin{aligned}
|||u-u_h|||^2 &\le \sum_{T\in\mathcal{T}_h} (\eta_{T;1}^i + \eta_{T;0}^i)\,|||u-u_h|||_T + \mathrm{osc}_i(f)\,|||u-u_h||| \\
&= \sum_{T\in\mathcal{T}_h} \eta_T^i\,|||u-u_h|||_T + \mathrm{osc}_i(f)\,|||u-u_h||| \le (\eta_h^i + \mathrm{osc}_i(f))\,|||u-u_h|||.
\end{aligned}$$

■

Remark 2.3.4. *Note that our upper bound is independent of the shape regularity of the mesh. More precisely it also holds for so-called anisotropic meshes, i.e., meshes for which σ tends to zero as the mesh size h goes to zero.*

Local upper bound for the discretization error

To show that the error estimator is locally bounded by the discretization error and higher order terms, we apply a suitable norm equivalence for mixed finite elements. Define for each element $T\in\mathcal{T}_h$ the quantities $m_{\partial T}(\cdot)$ and $m_T(\cdot)$ by

$$m_{\partial T}(v) = \|v\cdot n_T\|_{\partial T} \qquad m_T(v) = \|\int_T v\|_2, \tag{2.9}$$

where $\|\cdot\|_2$ denotes the Euclidean norm for vectors or matrices. We note that the two quantities are well defined if, e.g., the components of v are polynomials.

Lemma 2.3.5. *Let $v_h \in V^i(T)$, $T \in \mathcal{T}_h$, then*

$$\|v_h\|_T \sim (h_T^{\frac{1}{2}} m_{\partial T}(v_h) + \frac{\beta_i}{h_T} m_T(v_h)), \tag{2.10}$$

where $\beta_1 = \beta_2 = 0$ and $\beta_3 = 1$.

Proof: For convenience of the reader, we sketch the basic steps of the proof. Using the reference element $\widehat{T}$ with vertices $(0,0)$, $(1,0)$ and $(0,1)$, we find for $\hat{v}_h \in V^i(\widehat{T})$ that

$$\|\hat{v}_h\|_{\widehat{T}} \sim (m_{\partial\widehat{T}}(\hat{v}_h) + \beta_i m_{\widehat{T}}(\hat{v}_h)).$$

This simply follows from the fact that all norms on finite dimensional spaces are equivalent. Now we can use the Piola transformation to define for $v_h \in V^i(T)$ a corresponding $\hat{v}_h \in V^i(\widehat{T})$ by

$$\hat{v}_h(\widehat{\boldsymbol{x}}) = \det B_T B_T^{-1} v_h(\boldsymbol{x}),$$

where $\widehat{T}$ is mapped onto T by the affine mapping $\boldsymbol{x} = B_T\widehat{\boldsymbol{x}} + b_T$ and $B_T \in \mathbb{R}^{2\times 2}$ and $b_T \in \mathbb{R}^2$. We recall that $\|B_T\|_2 \sim |\det B_T|\ \|B_T^{-1}\|_2 \sim h_T$ and $|\det B_T| \sim h_T^2$. Then it is easy to see that $\|v_h\|_T \sim \|\hat{v}_h\|_{\widehat{T}}$. Using the relation $\|B_T^{-\top} n_{\widehat{T}}\|_2 n_T = B_T^{-\top} n_{\widehat{T}}$, we find $\det B_T \|B_T^{-\top} n_{\widehat{T}}\|_2 v_h \cdot n_T = \hat{v}_h \cdot n_{\widehat{T}}$ and thus $\|v_h \cdot n_T\|_{\partial T}^2 \sim h_T^{-1} \|\hat{v}_h \cdot n_{\widehat{T}}\|_{\partial\widehat{T}}^2$. For the volume integral we find $\int_T v_h = B_T \int_{\widehat{T}} \hat{v}_h$ and thus $\| \int_T v_h\|_2 \sim h_T \| \int_{\widehat{T}} \hat{v}_h\|_2$. ∎

We consider the two terms of the error estimators separately, and recall that $\eta_{T;0}^3 = 0$ and $\eta_{T;0}^1 = \eta_{T;0}^2$.

Lemma 2.3.6. *For each $T \in \mathcal{T}_h$ and for $i = 1, 2$, we have*

$$\eta_{T;0}^i = \alpha_T \|\Pi_h^i u_h - u_h\|_T \ \lesssim\ \alpha_T \|f - \Pi_h^i f\|_T + \frac{\sqrt{h_T}}{\sqrt{a_T}} \|g_T - a_T \nabla u_h \cdot n_T\|_{\partial T}. \tag{2.11}$$

Proof: Observing $u_h - \Pi_h^i u_h \in \mathcal{P}_1(T)$ and $a_T \mathrm{div}\, \nabla u_h = 0$ on T, then (2.4) and Green's formula yield

$$\begin{aligned}
\|\Pi_h^i u_h - u_h\|_T^2 &= \int_T u_h(u_h - \Pi_h^i u_h) = \int_T f(u_h - \Pi_h^i u_h) \\
&+ \int_{\partial T} g_T(u_h - \Pi_h^i u_h) - \int_T a \nabla u_h \nabla(u_h - \Pi_h^i u_h) \\
&= \int_T (f - \Pi_h^i f)(u_h - \Pi_h^i u_h) + \int_{\partial T} (g_T - a_T \nabla u_h \cdot n_T)(u_h - \Pi_h^i u_h) \\
&\leq \|f - \Pi_h^i f\|_T \|u_h - \Pi_h^i u_h\|_T + \|g_T - a_T \nabla u_h \cdot n_T\|_{\partial T} \|u_h - \Pi_h^i u_h\|_{\partial T} \\
&\lesssim \left(\|f - \Pi_h^i f\|_T + \tfrac{1}{\sqrt{h_T}} \|g_T - a_T \nabla u_h \cdot n_T\|_{\partial T} \right) \|u_h - \Pi_h^i u_h\|_T.
\end{aligned}$$

From the definition of α_T it follows directly that $\alpha_T/\sqrt{h_T} \leq \sqrt{h_T}/\sqrt{a_T}$. ∎

We recall that the constant only depends on the shape regularity of the element, and can be easily explicitly computed if required. In the following lemma, we provide an upper bound for $\eta_{T;1}^i$.

Lemma 2.3.7. *For each element $T \in \mathcal{T}_h$ and $i = 1, 2, 3$ we have*

$$\eta^i_{T;1} \lesssim \frac{\sqrt{h_T}}{\sqrt{a_T}} \|a_T \nabla u_h \cdot n_T - g_T\|_{\partial T}. \tag{2.12}$$

Proof: The proof is based on the discrete norm equivalence given in Lemma 2.3.5 and the observation that $a_T \nabla u_h \in V^i(T)$ for $i = 1, 2, 3$. Using the definition of the flux j^i_h and of β_i, we find $\beta_i m_T(j^i_h - a_T \nabla u_h) = 0$. Then, the norm equivalence (2.10) yields

$$\eta^i_{T;1} \lesssim \frac{\sqrt{h_T}}{\sqrt{a_T}} m_{\partial T}(a \nabla u_h - j^i_h) = \frac{\sqrt{h_T}}{\sqrt{a_T}} \|a_T \nabla u_h \cdot n_T - j^i_h \cdot n_T\|_{\partial T}.$$

Next, we observe that $(a_T \nabla u_h - j^i_h) \cdot n_e \in S^i(e)$, where $S^1(e) = S^2(e) = \mathcal{P}_0(e)$ and $S^3(e) = \mathcal{P}_1(e)$. Let $\Pi^i_{\partial T}$ be the L^2-projection onto $\prod_{e \subset \partial T} S^i(e) = S^i(\partial T)$, then $\Pi^i_{\partial T}(a_T \nabla u_h \cdot n_T) = a_T \nabla u_h \cdot n_T$ and $j^i_h \cdot n_T = \Pi^i_{\partial T}(j^i_h \cdot n_T) = \Pi^i_{\partial T} g_T$. Here we have used the definition of j^i_h and the fact that $j^i_h \cdot n_T \in S^i(\partial T)$. These preliminary considerations give now the upper bound

$$\eta^i_{T;1} \lesssim \frac{\sqrt{h_T}}{\sqrt{a_T}} \|\Pi^i_{\partial T}(a_T \nabla u_h \cdot n_T - g_T)\|_{\partial T} \leq \frac{\sqrt{h_T}}{\sqrt{a_T}} \|a_T \nabla u_h \cdot n_T - g_T\|_{\partial T}.$$

■

Theorem 2.3.8. *For each element $T \in \mathcal{T}_h$ the following estimate holds*

$$\eta^i_T \lesssim \max\{1, h_T a_T^{-1/2}\} \left(\max_{T' \subset \omega_T} \{\frac{\sqrt{a_{T'}}}{\sqrt{a_T}}\} |||u - u_h|||_{\omega_T} + osc_{|\omega_T}(f) \right), \tag{2.13}$$

where ω_T denotes the patch consisting of all the triangles of $\mathcal{T}_h$ sharing an edge with T.

Proof: Lemmas 2.3.6 and 2.3.7 and the definition (2.7) of the error estimator give

$$\eta^i_T \lesssim \frac{\sqrt{h_T}}{\sqrt{a_T}} \|a_T \nabla u_h \cdot n_T - g_T\|_{\partial T} + \alpha_T \|f - \Pi^i_h f\|_T.$$

The first term on the right side is bounded by the edge contributions $\sqrt{h_e}/\sqrt{a_T} \|a_T \nabla u_h \cdot n_e - g_e\|^2_e$ which is a part of the equilibrated error estimator that can be bounded in terms of the discretization error. Theorem 6.2 of [3] yields

$$\sum_{e \subset \partial T} h_e \|a_T \nabla u_h \cdot n_e - g_e\|^2_e \lesssim \sum_{T' \subset \omega_T} h^2_{T'} \|R_{T'}\|^2_{T'} + \sum_{e \subset \omega_T} h_e \|J_{e,n}\|^2_e,$$

where $R_T = f + \operatorname{div}(a \nabla u_h) - u_h$ is the exact residual on the element T and $J_{e,n}$ stands for the jump of the flux over edges :

$$J_{e,n} = \begin{cases} \left[\!\left[a \nabla u_h \cdot n_e \right]\!\right]_e & \text{for interior edges,} \\ 0 & \text{for Dirichlet boundary edges,} \\ \nabla u_h \cdot n_e & \text{for Neumann boundary edges.} \end{cases}$$

Introducing, for an edge e, $a_e = \max\{a_{T_1}, a_{T_2}\},\ e = \partial T_1 \cap \partial T_2$ we get

$$\begin{aligned} {\eta_T^i}^2 &\lesssim a_T^{-1} \max_{T' \subset \omega_T} \{a_{T'}\} \left(\sum_{T' \subset \omega_T} a_{T'}^{-1} h_{T'}^2 \|R_{T'}\|_{T'}^2 + \sum_{e \subset \omega_T} a_e^{-1} h_e \|J_{e,n}\|_e^2 \right) \\ &+ \alpha_T^2 \|f - \Pi_h^i f\|_T^2. \end{aligned} \tag{2.14}$$

The residual and the jump are terms appearing in the residual based error estimator. It is well known, see, e.g., [52], that these terms can be locally bounded by the error. Introducing element and edge bubble, we can bound, by inverse inequalities, those terms by local contributions of the discretization error. ■

2.3.2 Numerical results

Our first example consists in solving the equation (4.1) on the unit square $\Omega = (0,1)^2$ with $\Gamma_N = \Gamma$. The coefficient a is fixed to be constant and equal to 1. We take isotropic meshes composed of triangles, and we compute j_h^i, $i = 1, 2, 3$. The test is performed with different types of solutions. In the first case, we consider the exact solution

$$u(x,y) = \frac{1}{2} cos(\pi x) cos(\pi y). \tag{2.15}$$

To begin, we check that the numerical solution u_h converges toward the exact solution. To this end, we plot the curve $|||u - u_h|||$ (and the estimators) as a function of DoF (see Fig. 2.1). We see that the approximated solution converges toward the exact one with a convergence rate of one and that the estimators are very close to the error (see Fig. 2.1 and 2.2). In all our test settings, we find that the so-called effectivity indices, i.e., the ratios $|||u-u_h|||/\eta_h^i$, are smaller than one. Indeed we remark in Figure 2.2 that they vary between 0.67 and 0.87, in other words they remain smaller than one.

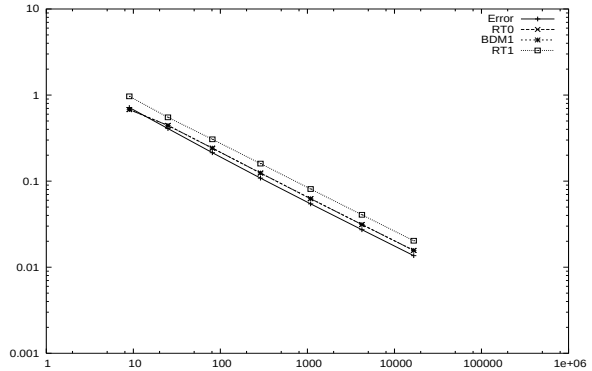


FIG. 2.1 – $|||u - u_h|||$ and η_h^i, $i = 1, 2, 3$ wrt DoF for the first solution.

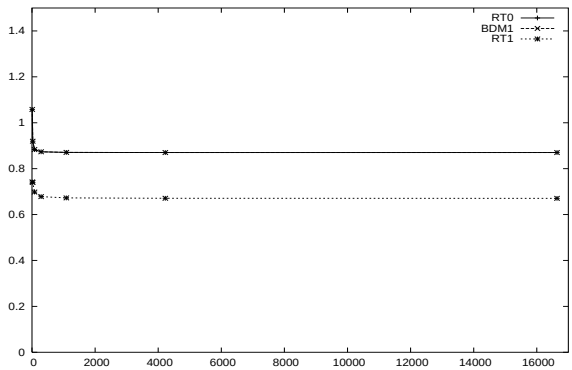


FIG. 2.2 – The ratios $|||u-u_h|||/\eta_h^i$, $i = 1, 2, 3$ wrt DoF for the first solution.

Now we take for the exact solution :

$$u(x,y) = e^{(3x^2-2x^3+3y^2-2y^3)}. \tag{2.16}$$

As before Figure 2.3 shows the error and the estimators wrt the DoF, while Figure 2.4 gives the effectivity indices. Here we can make the same conclusion as before, except that the effectivity indices are even smaller.

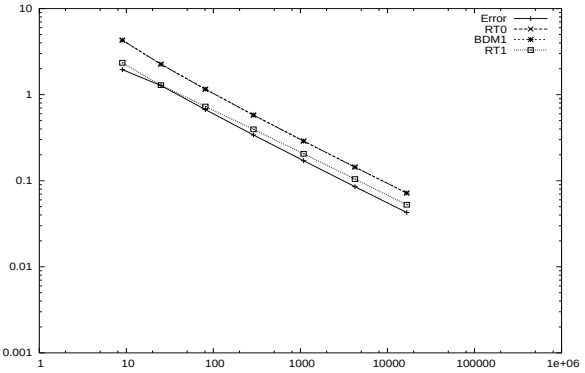


FIG. 2.3 – $|||u-u_h|||$ and η_h^i, $i=1,2,3$ wrt DoF for the second solution.

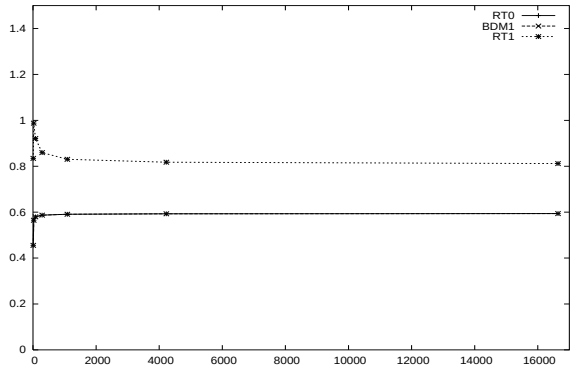


FIG. 2.4 – The ratios $|||u-u_h|||/\eta_h^i$, $i=1,2,3$ wrt DoF for the second solution.

As third example, we consider a solution of problem (4.1) on the unit square $\Omega=(0,1)^2$ with $\Gamma_D=\Gamma$ that exhibits an exponential layer along the y-axis. Namely we take

$$u(x,y) = 4y(1-y)(1-e^{-\alpha x}-(1-e^{-\alpha})x) \tag{2.17}$$

with different values of the parameter α, the coefficient in (4.1) being taken as $a=\frac{1}{\alpha^2}$. Here in order to resolve appropriately the boundary layer of the solution we use anisotropic meshes of Shishkin type as described in [32, 45] for instance (see Remark 2.3.4). First, we compute the estimator η_h^3 and compare it with the exact error. According to Fig. 2.5 we see a good convergence of the approximated solution to the exact one, moreover the estimator remains close to the error as far as the mesh size is small enough, this is confirmed by Fig. 2.6, where the effectivity index is presented for the four values of α with respect to DoF. Secondly, we have computed the global estimator η_h^1 (based on RT_0) and compare it with the exact error and the two contributions η_0^1 and η_1^1, these comparisons are presented in Fig. 2.7 and 2.8 for $\alpha=1$ and 10. In Fig. 2.7, we may see that as far as the mesh size is small enough with respect to the size of α, the term η_0^1 is much smaller than η_1^1, as theoretically expected. On the contrary if the mesh size is relatively rough with respect to the size of α, the term η_1^1 is comparable with η_1^1 (see Fig. 2.7 right). Note further that the use of η_h^1 is more time consuming than η_h^3 since we were unable to achieve the value of $h=1/128$ for $\alpha=100$ and 1000 in a reasonable time.

Now in order to illustrate the performance of our estimator η_h^3, for three examples taken from [38] we show the meshes obtained after some iterations using an iterative algorithm

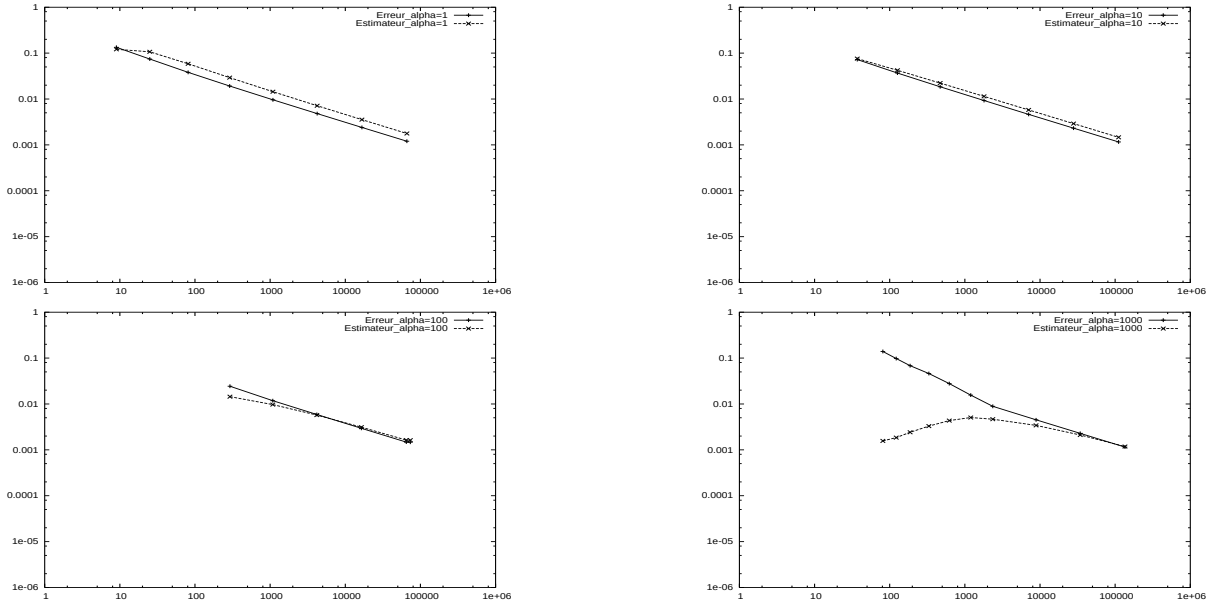


FIG. 2.5 – $|||u - u_h|||$ and η_h^3 wrt DoF for different values of α : top-left : $\alpha = 1$, top-right $\alpha = 10$; bottom-left $\alpha = 100$, bottom-right $\alpha = 1000$.

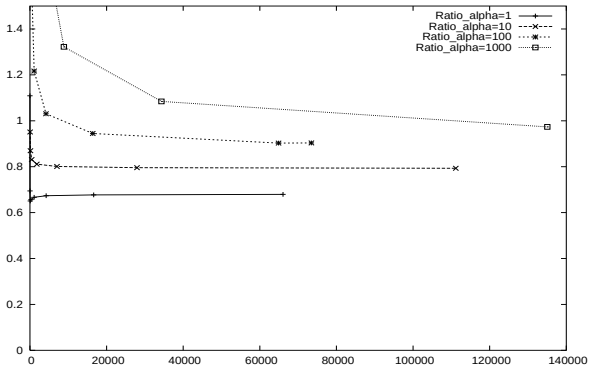


FIG. 2.6 – The ratio $|||u - u_h|||/\eta_h^3$ wrt DoF for different values of α.

based on the marking procedure

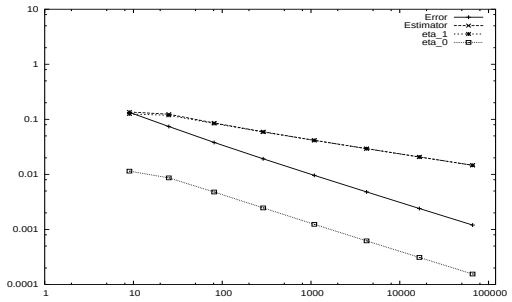
$$\eta_T > 0.5 \max_{T'} \eta_{T'} \text{ or } \eta_T > 0.75 \max_{T'} \eta_{T'},$$

and a standard refinement procedure with a limitation on the minimal angle.

For the first example we take $\Omega = (0,1)^2$, $a = 1$, $\Gamma_D = \Gamma$ and as exact solution :

$$u(x,y) = x(x-1)y(y-1)e^{(-100(x-1/2)^2 - 100(y-117/1000)^2)}.$$

This solution has a large gradient around the point $(\frac{1}{2}, \frac{117}{100})$. Therefore a refinement of the mesh near this point can be expected. This is confirmed by Figure 2.9.



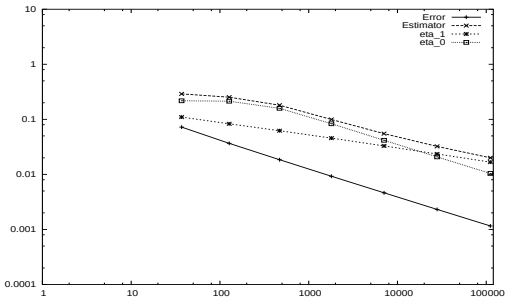
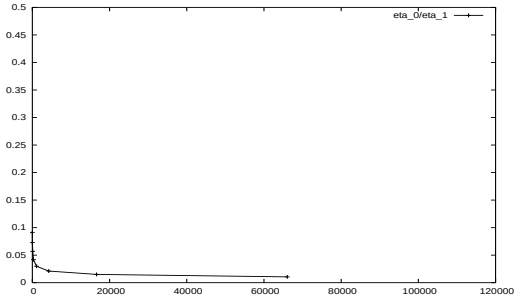


FIG. 2.7 – $|||u - u_h|||$, η_h^1, η_0^1 and η_1^1 wrt DoF for different values of α : on the left $\alpha = 1$, on the right $\alpha = 10$.



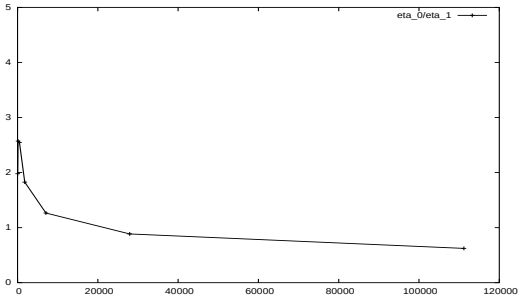


FIG. 2.8 – The ratio η_0^1/η_1^1 wrt DoF for different values of α : on the left $\alpha = 1$, on the right $\alpha = 10$.

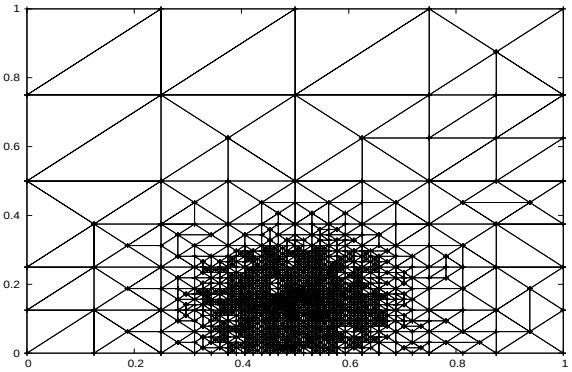


FIG. 2.9 – Adaptive mesh after 10 iterations for the first example and criterion $\eta_T > 0.5 \max_{T'} \eta_{T'}$.

For the second example we take $\Omega = (-1, 1)^2$ and $\Gamma_D = \Gamma$ but a discontinuous coefficient a. Namely we decompose Ω into 4 sub-domains Ω_i, $i = 1, \ldots, 4$ with $\Omega_1 = (0, 1) \times (0, 1)$, $\Omega_2 = (-1, 0) \times (0, 1)$, $\Omega_3 = (-1, 0) \times (-1, 0)$ and $\Omega_4 = (0, 1) \times (-1, 0)$ and take $a = a_i$ on Ω_i, with $a_1 = a_3$ and $a_2 = a_4 = 1$. Using polar coordinates centered at $(0, 0)$, we take as

exact solution,

$$S(x,y) = r^{\alpha}\phi(\theta), \tag{2.18}$$

where $\alpha \in (0,1)$ and ϕ are chosen such that S is harmonic on each sub-domain Ω_i, $i = 1,\ldots,4$ and satisfies the jump conditions :

$$[\![S]\!] = 0 \text{ and } [\![a\nabla S{\cdot}n]\!] = 0$$

on the interfaces (i.e. the segments $\bar{\Omega}_i \cap \bar{\Omega}_{i+1} \pmod 4$, $i = 1,\ldots,4$). We fix non-homogeneous Dirichlet boundary conditions on Γ accordingly.

It is easy to see (see for instance [22]) that α is the root of the transcendental equation

$$\tan\frac{\alpha\pi}{4} = \sqrt{a_1}.$$

This solution has a singular behavior around the point $(0,0)$ (because $\alpha < 1$). Therefore a refinement of the mesh near this point can be expected. This can be checked in Figures 2.10 and 2.11 on the meshes obtained for $a_1 = 5$ and $a_1 = 100$ respectively and for which $\alpha \approx 0.53544094560$ and $\alpha \approx 0.1269020697$.

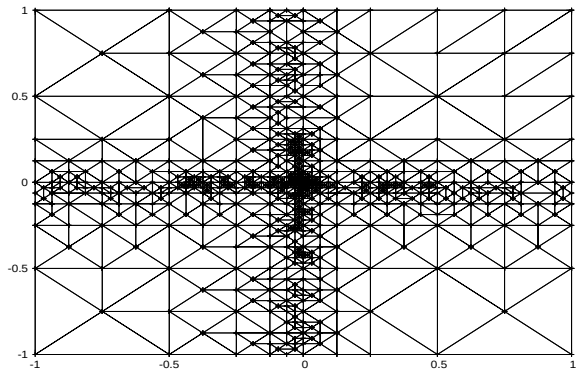


FIG. 2.10 – Adaptive mesh after 20 iterations for the second example ($a_1 = 5$ and criterion $\eta_T > 0.75\max_{T'}\eta_{T'}$).

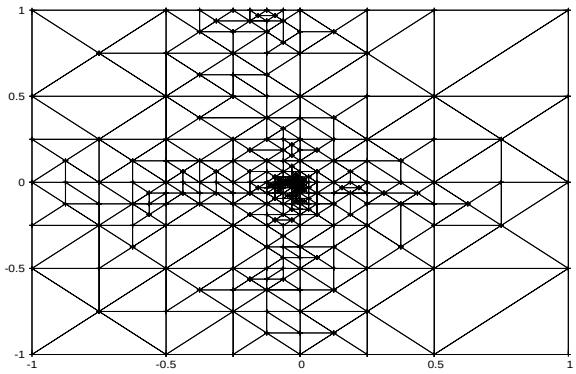


FIG. 2.11 – Adaptive mesh after 20 iterations for the second example ($a_1 = 100$ and criterion $\eta_T > 0.75\max_{T'}\eta_{T'}$).

Finally as last example, we take the L-shape domain $\Omega = (-1,1)^2 \setminus (-1,0)\times(0,1)$, $a = 1$, $\Gamma_D = \Gamma$ and as exact solution

$$S = r^{2/3}\sin(2\theta/3). \tag{2.19}$$

This solution has a singular behavior at $(0,0)$ and the meshes has to be refined near this point. This can be seen in Figure 2.12.

From all these tests we can confirm the reliability and efficiency of our proposed error estimators. Nevertheless for $a \ll 1$ and coarse meshes the estimator η_h^3 based on RT_1 is more attractive and less expensive than the estimators η_h^1 and η_h^2.

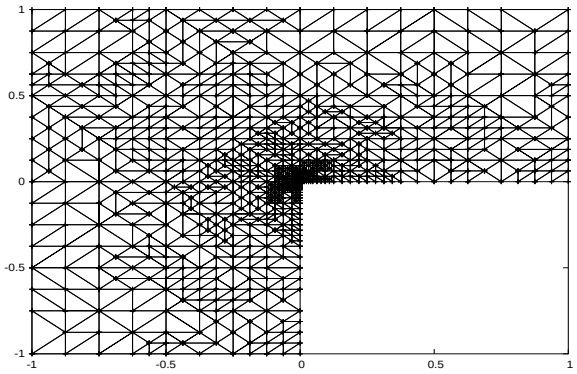


FIG. 2.12 – Adaptive mesh after 10 iterations for the third example and criterion $\eta_T > 0.5 \max_{T'} \eta_{T'}$.

2.4 The time-Harmonic Maxwell equations in 3D

Now, Ω represents a bounded domain of $\mathbb{R}^3$ with a polyhedral boundary Γ. For the sake of simplicity, we further assume that Ω is simply connected and that its boundary is connected. We are interested in the following problem :

$$\begin{aligned} \operatorname{curl}(\chi \operatorname{curl} u) + \beta u &= f \text{ in } \Omega, \\ u \times n &= 0 \text{ on } \Gamma. \end{aligned} \tag{2.20}$$

In the rest of the chapter, we suppose that χ and β are piecewise positive constants. For any $f \in [L^2(\Omega)]^3$ satisfying $\operatorname{div} f = 0$ in Ω, the weak formulation of (2.20) is given by : Find $u \in H_0(\operatorname{curl}, \Omega) = \{v \in [L^2(\Omega)]^3 | \operatorname{curl} v \in [L^2(\Omega)]^3 \text{ and } v \times n = 0 \text{ on } \Gamma\}$ such that

$$B(u,v) = \int_\Omega (\chi \operatorname{curl} u \cdot \operatorname{curl} v + \beta u \cdot v) = \int_\Omega f \cdot v, \ \forall v \in H_0(\operatorname{curl}, \Omega). \tag{2.21}$$

As χ and β are uniformly positive, B is coercive on $H_0(\operatorname{curl}, \Omega)$ with respect to the norm $\|u\|_{\beta,\chi} = (B(u,u))^{1/2}$ and, by the Lax-Milgram lemma, problem (2.20) admits a unique solution.

2.4.1 The approximated problem

The triangulation $\mathcal{T}_h$ is now made of tetrahedra T. Its faces are denoted by F and n_F stands for one of the unit normal vectors of this face. We use the notation $\mathcal{F}$ for the set of faces and $\mathcal{V}_h$ for the set of vertices of the triangulation. All notation introduced before remain valid, except that the elements T are now tetrahedra. In the sequel, χ_T (resp. β_T) denotes the value of the piecewise constant χ (resp. β) restricted to an element T.

Problem (2.21) is approximated in a curl-conforming finite element subspace X_h of $H_0(\operatorname{curl}, \Omega)$ build using the lowest order Nédélec finite elements :

$$X_h = \left\{ v_h \in H_0(\operatorname{curl}, \Omega) | v_{h|T} \in \mathcal{ND}_1(T), T \in \mathcal{T}_h \right\}$$

where $\mathcal{ND}_1(T) = [\mathcal{P}_0(T)]^3 + [\mathcal{P}_0(T)]^3 \times \boldsymbol{x}$ with $\boldsymbol{x} = (x_1, x_2, x_3)^\top$. The discretized problem consists in finding $u_h \in X_h$ such that

$$B(u_h, v_h) = (f, v_h), \ \forall v_h \in X_h. \tag{2.22}$$

We now recall a decomposition of the space $H_0(\mathrm{curl}, \Omega)$ of Helmholtz type related to the weight β.

Lemma 2.4.1. *If Ω is simply connected and its boundary Γ is connected then*

$$H_0(\mathrm{curl}, \Omega) = \nabla H_0^1(\Omega) \overset{\perp}{\oplus} W_\beta, \tag{2.23}$$

where W_β is a closed subspace of $H_0(\mathrm{curl}, \Omega)$ defined by

$$W_\beta = \{v \in H_0(\mathrm{curl}, \Omega) | \mathrm{div}\,(\beta v) = 0 \text{ in } \Omega\}, \tag{2.24}$$

and the symbol $\overset{\perp}{\oplus}$ means that the decomposition is direct and orthogonal with respect to the inner product $B(\cdot, \cdot)$. Then the error $u - u_h$ admits the splitting

$$u - u_h = \nabla\phi + e_\perp, \tag{2.25}$$

with $\phi \in H_0^1(\Omega)$, $e_\perp \in W_\beta$ and

$$\|u - u_h\|_{\beta,\chi}^2 = \|\nabla\phi\|_{\beta,\chi}^2 + \|e_\perp\|_{\beta,\chi}^2. \tag{2.26}$$

Moreover, there exists $\epsilon \in (0, 1]$ (depending on β and on the geometry of Ω) and a constant $C(\beta)$ such that $e_\perp \in (H^\epsilon(\Omega))^3$ with the estimate

$$\|e_\perp\|_{\epsilon,\Omega} \leq C(\beta) \max_{j=1,\cdots,J} \left\{1, \chi_j^{-1/2}\right\} \|u - u_h\|_{\beta,\chi}. \tag{2.27}$$

Proof: As $\nabla H_0^1(\Omega)$ is a closed subspace of $H_0(\mathrm{curl}, \Omega)$ (see Lemma I.2.1 of [26]), the decomposition (2.23) holds with

$$W_\beta = \{v \in H_0(\mathrm{curl}, \Omega) : (\beta v, \nabla\psi) = 0, \forall \psi \in H_0^1(\Omega)\}.$$

By Green's formula we deduce (2.24).

For the requested regularity of $e_\perp$, we apply Theorem 3.5 of [22], which further yields

$$\|e_\perp\|_{\epsilon,\Omega} \leq C(\beta) \left(\|\,\mathrm{curl}\, e_\perp\| + \|\mathrm{div}\,(\beta e_\perp)\|\right).$$

As $e_\perp \in W_\beta$, $\mathrm{div}\,(\beta e_\perp) = 0$. Moreover from the splitting of the error, we see that $\mathrm{curl}(u - u_h) = \mathrm{curl}\, e_\perp$, therefore

$$\|e_\perp\|_{\epsilon,\Omega} \leq C(\beta)\|u - u_h\|_{\beta,1}.$$

■

2.4.2 Conforming approximated problems

Again our idea is based on a saddle point approach. Namely introducing as auxiliary variable $j = \chi \operatorname{curl} u$, then (2.21) becomes : Find $(j,u) \in H(\operatorname{curl},\Omega) \times [L^2(\Omega)]^3$ solution of

$$\begin{array}{rcl} \displaystyle\int_\Omega \chi^{-1} j \cdot v - \int_\Omega \operatorname{curl} v \cdot u &=& 0,\ \forall v \in H(\operatorname{curl},\Omega), \\ \displaystyle\int_\Omega \operatorname{curl} j \cdot w + \int_\Omega \beta u \cdot w &=& \displaystyle\int_\Omega f \cdot w,\ \forall w \in [L^2(\Omega)]^3. \end{array} \tag{2.28}$$

The lowest order approximated mixed finite element pair for this problem is the pair (V_h, W_h) where

$$\begin{array}{rcl} V_h &=& \left\{ v_h \in H(\operatorname{curl},\Omega) | v_{h|T} \in \mathcal{ND}_1(T), T \in \mathcal{T}_h \right\}, \\ W_h &=& \{ w_h \in H(\operatorname{div},\Omega) | w_{h|T} \in RT_0(T), \forall T \in \mathcal{T}_h \text{ and } \operatorname{div} w_h = 0 \text{ in } \Omega \}. \end{array}$$

Therefore a natural choice for our approximated flux is $j_h \in V_h$. But here, contrary to the reaction-diffusion case, βu_h no more belongs to W_h, essentially because βu_h is no more divergence free. Hence we first construct a correction q_h that fulfils this constraint. For that purpose we introduce equilibrated fluxes for the divergence part. Namely let l_F be in $\mathcal{P}_1(F)$ and satisfying the divergence flux equations :

$$\int_T \beta u_h \cdot \nabla w_h = \int_{\partial T} l_T w_h, \forall w_h \in \mathcal{P}_1(T), T \in \mathcal{T}_h, \tag{2.29}$$

where, as usual, $l_T = l_F n_T \cdot n_F$. The existence of such fluxes is proved as for the equilibrated flux equation (2.4) (see [3]) because the discrete problem (2.22) guarantees that (because f is supposed to be divergence free)

$$\int_{\omega_x} \beta u_h \cdot \nabla \lambda_x = \int_{\omega_x} f \cdot \nabla \lambda_x = 0,$$

for all nodes x (when λ_x is the standard hat function). We now fix the discrete divergence flux as the unique $q_h \in V_h^3$ satisfying

$$\int_F q_h \cdot n_F q = \int_F l_F q, \forall q \in \mathcal{P}_1(F), F \subset T, T \in \mathcal{T}_h, \tag{2.30}$$

$$\int_T q_h = \int_T \beta u_h, \forall T \in \mathcal{T}_h, \tag{2.31}$$

and prove the following projection lemma :

Lemma 2.4.2. *If $q_h \in V_h^3$ satisfies (2.30) and (2.31), then* $\operatorname{div} q_h = 0$.

Proof: By Green's formula and (2.30), (2.31), for any $w_h \in W_h^3$, it follows that

$$\int_\Omega \operatorname{div} q_h w_h = \sum_{T \in \mathcal{T}_h} (- \int_T \beta u_h \cdot \nabla w_h + \int_{\partial T} l_T w_h) = 0,$$

due to (2.29). ∎

As f and q_h do not belong to W_h, we shall consider their projection on this space. Namely denoting by Π_h the projection onto W_h with respect to the inner product $(w_h, v_h)_{\beta^{-1}} = \int_\Omega \beta^{-1} w_h \cdot v_h$, we set

$$\tilde{f}_h = \Pi_h f - \Pi_h q_h.$$

Now we consider the alternative problem : Find $\tilde{u} \in X = \{v \in H_0(\mathrm{curl}, \Omega) | \mathrm{div}\,(\beta v) = 0 \text{ in } \Omega\}$ solution of

$$\int_\Omega \chi \,\mathrm{curl}\,\tilde{u} \cdot \mathrm{curl}\, v = \int_\Omega \tilde{f}_h \cdot v, \quad \forall v \in X. \tag{2.32}$$

In order to make an adequate approximation of this problem we use the discrete Helmholtz decomposition of X_h (see [37]) into a subspace of X_h made of discrete β-divergence free functions and curl-free functions, namely we use the splitting

$$X_h = \tilde{X}_h \overset{\perp}{\oplus} \nabla S_h$$

where $\tilde{X}_h = \{w_h \in X_h | (\beta w_h, \nabla \varphi_h) = 0, \forall \varphi_h \in S_h\}$ and $S_h = \{\varphi_h \in H_0^1(\Omega) | \varphi_{h_{|T}} \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h\}$. The decomposition being orthogonal with respect to the inner product $(\beta \cdot, \cdot)$.

Hence the approximated problem of (2.32) is : Find $\tilde{u}_h \in \tilde{X}_h$ satisfying

$$\int_\Omega \chi \,\mathrm{curl}\,\tilde{u}_h \,\mathrm{curl}\,\tilde{v}_h = \int_\Omega \tilde{f}_h \,\tilde{v}_h, \quad \forall \tilde{v}_h \in \tilde{X}_h. \tag{2.33}$$

This problem is well posed since its left-hand side is coercive on $\tilde{X}_h$, due to the discrete Friedrichs inequality.

At this stage we are able to apply Theorem 15 of [13] to the problem (2.32) and its approximation (2.33) that prove the next results :

Lemma 2.4.3. *There exists (an explicitly computable) $j_h \in V_h$ satisfying* $\mathrm{curl}\, j_h = \Pi_h f - \Pi_h q_h$ *with the following estimates*

$$\|\chi^{-1/2}(j_h - \chi \,\mathrm{curl}\,\tilde{u_h})\| \lesssim \|\chi^{1/2} \,\mathrm{curl}(\tilde{u} - \tilde{u_h})\| \leq \|\chi^{-1/2}(j_h - \chi \,\mathrm{curl}\,\tilde{u_h})\|. \tag{2.34}$$

Proof: Following what is done in [13], we first remark that

$$\mathrm{div}\,\tilde{f}_h = \mathrm{div}\,(\Pi_h f) - \mathrm{div}\,(\Pi_h q_h) = I_{RT0}(\mathrm{div}\, f) - I_{RT0}(\mathrm{div}\, q_h) = 0,$$

where I_{RT0} is the Raviart-Thomas interpolation operator mapping $H_N(\mathrm{div}\,, \Omega)$ onto V_h^1. The discrete Helmholtz decomposition of X_h implies that, for any basis function λ_e of X_h, there exist $\tilde{\lambda}_e \in \tilde{X}_h$ and $\varphi \in S_h$ such that $\tilde{\lambda}_e = \lambda_e - \nabla\varphi$ and, by its definition, the approximation $\tilde{u}_h$ satisfies

$$(\chi \,\mathrm{curl}\,\tilde{u}_h, \mathrm{curl}(\lambda_e - \nabla\varphi)) = (\tilde{f}_h, \lambda_e - \nabla\varphi), \forall e \in \mathcal{E}_h.$$

We obtain the orthogonality relation

$$\langle \tilde{f}_h - \operatorname{curl} \chi \operatorname{curl} \tilde{u}_h, \lambda_e \rangle = 0, \forall e \in \mathcal{E}_h,$$

with

$$\operatorname{div} (\tilde{f}_h - \operatorname{curl} \chi \operatorname{curl} \tilde{u}_h) = 0,$$

and from Lemma 14 of [13], the following local decomposition holds :

$$\tilde{f}_h - \operatorname{curl} \chi \operatorname{curl} \tilde{u}_h = \sum_{V \in \mathcal{V}_h} \tilde{f}_{\omega_V} \text{ with } \operatorname{div} \tilde{f}_{\omega_V} = 0, \ \forall V \in \mathcal{V}_h,$$

ω_V denoting the patch consisting of all the elements of $\mathcal{T}_h$ containing the vertex V.
We conclude by Lemma 15 of [13] that there exists $j^\Delta = \sum_{V \in \mathcal{V}_h} j_{\omega_V}$, with $supp(j_{\omega_V}) \subset \omega_V$, satisfying $\operatorname{curl} j_{\omega_V} = \tilde{f}_{\omega_V}$.
If we introduce $j_h = j^\Delta + \chi \operatorname{curl} \tilde{u}_h$, this discrete flux clearly verifies $\operatorname{curl} j_h = \tilde{f} = \Pi_h f - \Pi_h q_h$ and the following estimate can be proved

$$C \|\chi^{-1/2} j^\Delta\| \leq \|\chi^{1/2} \operatorname{curl}(\tilde{u} - \tilde{u_h})\| \leq \|\chi^{-1/2} j^\Delta\|, \text{ with } C > 0. \tag{2.35}$$

■

2.4.3 The a posteriori error estimator

We now introduce local indicators of the error $u - u_h$ on an element T of the triangulation as follows

$$\eta_{\perp,T}^2 = \|\chi^{-1/2} (\chi \operatorname{curl} u_h - j_h)\|_T^2, \quad \eta_{0,T}^2 = \|\beta^{-1/2} (\beta u_h - q_h)\|_T^2.$$

The associated global estimator is then given by $\eta = \left(\sum_{T \in \mathcal{T}_h} (\eta_{0,T}^2 + \eta_{\perp,T}^2) \right)^{1/2}$. The oscillation of a function f is here defined by

$$osc(f)^2 = \sum_{T \in \mathcal{T}_h} h_T^{2\epsilon} \beta_T^{-2} \|f - \Pi_h f\|_T^2,$$

where $\epsilon \in (0, 1]$ is the one from Lemma 2.4.1.

Upper bound

Theorem 2.4.4. *There exists $C(\beta, \chi) > 0$ depending on β and χ such that the following estimate holds*

$$\|u - u_h\|_{\beta,\chi} \leq \eta + C(\beta, \chi) \ (osc(f) + osc(q_h)). \tag{2.36}$$

Proof: From the definition of the norm, introducing the variable j_h and applying Green's formula, we get

$$\begin{aligned}\|u-u_h\|^2_{\beta,\chi} &= \int_\Omega \chi \operatorname{curl}(u-u_h)\operatorname{curl}(u-u_h) + \int_\Omega \beta(u-u_h)(u-u_h)\\ &= \int_\Omega (j_h - \chi \operatorname{curl} u_h)\operatorname{curl}(u-u_h) + \int_\Omega (f-\beta u_h - \Pi_h f + \Pi_h q_h)(u-u_h).\end{aligned}$$

Cauchy-Schwarz's inequality gives

$$\begin{aligned}\|u-u_h\|^2_{\beta,\chi} &\le \int_\Omega (f-\Pi_h f+\Pi_h q_h-\beta u_h)(u-u_h)\\ &+ \sum_{T\in\mathcal{T}_h} \|\chi^{-1/2}(j_h-\chi\operatorname{curl} u_h)\|_T \|\chi^{1/2}\operatorname{curl}(u-u_h)\|_T. \qquad (2.37)\end{aligned}$$

Introducing the Helmholtz decomposition of the error (2.25) and the divergence free flux q_h we get :

$$\begin{aligned}\int_\Omega (f-\Pi_h f+\Pi_h q_h-\beta u_h)\cdot(u-u_h) &= \int_\Omega (f-\operatorname{curl} j_h - q_h)\cdot\nabla\phi\\ &+ \int_\Omega (f-\Pi_h f+\Pi_h q_h - q_h)\cdot e_\perp + \int_\Omega (q_h-\beta u_h)\cdot(u-u_h). \qquad (2.38)\end{aligned}$$

We now estimate each term of this right-hand side. For the first term, applying Green's formula, we get

$$\int_\Omega (f-\operatorname{curl} j_h - q_h)\nabla\phi = \sum_{T\in\mathcal{T}_h}\int_T \operatorname{div}(f-\operatorname{curl} j_h - q_h)\phi = 0, \qquad (2.39)$$

as $\operatorname{div} f = \operatorname{div}\ \operatorname{curl} j_h = \operatorname{div} q_h = 0$.

For the second term, we notice that $(\Pi_h q_h - q_h, I_{RT0}(\beta e_\perp))_{\beta^{-1}} = 0$, as $I_{RT0}(\beta e_\perp)$ (the RT_0 interpolant of $\beta e_\perp$) belongs to W_h. Hence we may write

$$\int_\Omega (\Pi_h q_h - q_h)\cdot e_\perp = \int_\Omega \beta^{-1}(\Pi_h q_h - q_h)(\beta e_\perp - I_{RT0}(\beta e_\perp)).$$

Since $e_\perp$ belongs to $(H^\epsilon(\Omega))^3$, a scaling argument yields $\|\beta e_\perp - I_{RT0}(\beta e_\perp)\| \lesssim h^\epsilon \|\beta e_\perp\|_\epsilon$ (see Theorem 3.4 of [1]) and therefore

$$\int_\Omega (\Pi_h q_h - q_h)\cdot e_\perp \lesssim \left(\sum_{T\in\mathcal{T}_h} \beta_T^{-2} h_T^{2\epsilon}\|q_h-\Pi_h q_h\|^2_T\right)^{1/2} \|\beta e_\perp\|_\epsilon.$$

By the estimate (2.27), we arrive at

$$\int_\Omega (\Pi_h q_h - q_h)\cdot e_\perp \le Cosc(q_h)\|u-u_h\|_{\beta,\chi}, \qquad (2.40)$$

$$\int_\Omega (f-\Pi_h f)\cdot e_\perp \le Cosc(f)\|u-u_h\|_{\beta,\chi}, \qquad (2.41)$$

for some $C > 0$ depending on β and χ.

Finally for the third term, Cauchy-Schwarz's inequality directly yields

$$\int_\Omega (q_h - \beta u_h) \cdot (u - u_h) \leq (\sum_{T \in \mathcal{T}_h} \eta_{0,T}^2)^{1/2} \|\beta^{1/2}(u - u_h)\|. \tag{2.42}$$

The estimate (2.36) directly follows from the estimate (2.37), the identities (2.38) and (2.39) and the bounds (2.40), (2.41) and (2.42). ∎

Before going on let us point out that the terms $osc(f)$ and $osc(q_h)$ are higher order terms. First we remark that even for smooth f, the solution u of problem (2.20) will only have the regularity $u \in (H^\epsilon(\Omega))^3$. Therefore the expected order of convergence for the error will be ϵ, namely $\|u - u_h\|_{\beta,\chi} \leq C(\beta, \chi)h^\epsilon$, for some $C(\beta, \chi) > 0$ depending on β and χ. For the term $osc(f)$, if f belongs to $H^1(\Omega)^3$, then by scaling arguments we have $osc(f) \lesssim h^{1+\epsilon}\|f\|_{1,\Omega}$, and therefore $osc(f)$ tends to zero faster than the error (this will be achieved if β and χ are fixed and if h is small enough).

For the second term $osc(q_h)$, no global regularity results on u are necessary, namely using a scaling argument on each element T, we have $\|q_h - I_{RT0}q_h\|_T \lesssim h_T\|\nabla q_h\|_T$. Therefore we may write (here we do not trace the dependence on β and χ and write for shortness C for a constant depending on these two functions)

$$\begin{aligned} osc(q_h)^2 &\leq Ch^{2\epsilon} \min_{w_h \in W_h} \sum_{T \in \mathcal{T}_h} \beta_T^{-1}\|q_h - w_h\|_T^2 \leq Ch^{2\epsilon} \sum_{T \in \mathcal{T}_h} \beta_T^{-1}\|q_h - I_{RT0}q_h\|_T^2 \\ &\leq Ch^{2\epsilon} \sum_{T \in \mathcal{T}_h} h_T^2\|\nabla q_h\|_T^2 \\ &\leq Ch^{2\epsilon} \sum_{T \in \mathcal{T}_h} h_T^2(\|\nabla(q_h - \beta u_h)\|_T^2 + \|\nabla(\beta u_h)\|_T^2). \end{aligned}$$

As $\|\nabla(\beta u_h)\|_T \lesssim \beta_T\|\operatorname{curl} u_h\|_T$ (see e.g. Lemma 4.1 of [43]) and using a standard inverse inequality we obtain

$$osc(q_h)^2 \leq Ch^{2\epsilon}\left(\|q_h - \beta u_h\|^2 + h^2\|\operatorname{curl} u_h\|^2\right).$$

Since it will be proved below that $\|q_h - \beta u_h\| \leq C\|u - u_h\|_{\beta,\chi}$ (see the estimate (2.43)) and since the variational formulation (2.22) leads to $\|\chi^{1/2} \operatorname{curl} u_h\| \lesssim \|\beta^{-1/2}f\|$, we obtain

$$osc(q_h)^2 \leq Ch^{2\epsilon}(\|u - u_h\|_{\beta,\chi}^2 + h^2\|\beta^{-1/2}f\|^2).$$

This last estimate finally shows that $osc(q_h)$ tends to zero faster than the error.

Lower bound

As for the reaction-diffusion problems our lower bound is based on a suitable norm equivalence :

Theorem 2.4.5. *There exists a positive constant $C(\beta,\chi)$ (depending on β and χ) such that the following local and global lower bounds hold*

$$\eta_{0,T} \lesssim \beta_T^{-1/2} \max_{T'\subset\omega_T} \beta_{T'}^{1/2} \|u-u_h\|_{\beta,\chi,\omega_T}, \tag{2.43}$$

$$(\sum_{T\in\mathcal{T}_h} \eta_{\perp,T}^2)^{1/2} \leq C(\beta,\chi)(\|u-u_h\|_{\beta,\chi} + osc(f) + osc(q_h)). \tag{2.44}$$

Proof: On one hand, as $u_h \in \mathcal{NCD}_1(T) \subset RT_1(T)$, Lemma 2.3.5 and the construction of q_h yield

$$\|q_h - \beta u_h\|_T \lesssim \sum_{F\subset\partial T} h_F^{1/2} \|(q_h-\beta u_h)\cdot n_F\|_F \lesssim \sum_{F\subset\partial T} h_F^{1/2} \|\left[\!\left[\beta u_h\cdot n_F\right]\!\right]_F\|_F. \tag{2.45}$$

This right-hand side is a part of the estimator presented in [45]. By standard inverse inequality we can prove that

$$\|\left[\!\left[\beta u_h\cdot n_F\right]\!\right]_F\|_F \lesssim \sum_{T'\subset\omega_F} \beta_{T'}^{1/2} h_{T'}^{-1/2} \|u-u_h\|_{\beta,\chi,T'}.$$

These two estimates directly prove the local bound (2.43).

Now, using Lemma 2.4.3 we have

$$\begin{aligned}\|\chi^{-1/2}(j_h - \chi\operatorname{curl} u_h)\| &\leq \|\chi^{-1/2}(j_h-\chi\operatorname{curl}\tilde{u}_h)\| + \|\chi^{1/2}\operatorname{curl}(\tilde{u}_h-u_h)\| \\ &\lesssim \|\chi^{1/2}\operatorname{curl}(\tilde{u}-\tilde{u}_h)\| + \|\chi^{1/2}\operatorname{curl}(\tilde{u}_h-u_h)\| \\ &\lesssim \|\chi^{1/2}\operatorname{curl}(\tilde{u}-\tilde{u}_h)\| + \|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\| + \|\chi^{1/2}\operatorname{curl}(u-u_h)\|. \end{aligned}\tag{2.46}$$

The first term of this right-hand side can be bounded using the second Strang lemma :

$$\begin{aligned}\|\chi^{1/2}\operatorname{curl}(\tilde{u}-\tilde{u}_h)\| &\leq \inf_{v_h\in\tilde{X}_h} \|\chi^{1/2}\operatorname{curl}(\tilde{u}-v_h)\| + \sup_{w_h\in\tilde{X}_h} \frac{|(\chi\operatorname{curl}\tilde{u},\operatorname{curl}w_h) - (\tilde{f},w_h)|}{\|\chi^{1/2}\operatorname{curl}w_h\|} \\ &\leq \|\chi^{1/2}\operatorname{curl}(\tilde{u}-u_h+\nabla\varphi)\| \leq \|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\| + \|\chi^{1/2}\operatorname{curl}(u-u_h)\|,\end{aligned}$$

noting that there exists $\varphi\in S_h$ such that $u_h - \nabla\varphi \in \tilde{X}_h$ and that $(\chi\operatorname{curl}\tilde{u},\operatorname{curl}w_h) - (\tilde{f}_h,w_h) = 0$. This estimate and (2.46) yield

$$\|\chi^{-1/2}(j_h-\chi\operatorname{curl}u_h)\| \lesssim \|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\| + \|\chi^{1/2}\operatorname{curl}(u-u_h)\|.$$

It now remains to bound $\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\|$. Applying Green's formula we get

$$\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\|^2 = \int_\Omega \operatorname{curl}(\chi\operatorname{curl}\tilde{u}-\chi\operatorname{curl}u)\cdot(\tilde{u}-u) = \int_\Omega (\Pi_h f - \Pi_h q_h - f + \beta u)\cdot(\tilde{u}-u).$$

By the definition of the projection Π_h, we get

$$\begin{aligned}\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\|^2 =& \int_\Omega \beta^{-1}(\Pi_h f - f)\cdot(\beta(\tilde{u}-u) - I_{RT0}(\beta(\tilde{u}-u))) + \int_\Omega \beta^{1/2}(u-u_h)\cdot\beta^{1/2}(\tilde{u}-u) \\ &+ \int_\Omega \beta^{-1}(q_h-\Pi_h q_h)\cdot(\beta(\tilde{u}-u) - I_{RT0}(\beta(\tilde{u}-u))) + \int_\Omega \beta^{-1/2}(\beta u_h - q_h)\cdot\beta^{1/2}(\tilde{u}-u) \\ \leq& \; C(\beta)(osc(f)+osc(q_h))\|\beta(\tilde{u}-u)\|_\epsilon + (\|\beta^{1/2}(u-u_h)\| + \eta_0)\|\beta^{1/2}(\tilde{u}-u)\|_\epsilon.\end{aligned}$$

By the discrete Cauchy-Schwarz inequality and (2.43), we obtain

$$\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\|^2 \leq C(\beta)(osc(f)+osc(q_h)+\|u-u_h\|_{\beta,\chi})\|u-\tilde{u}\|_\epsilon,$$

and by the estimate $\|u-\tilde{u}\|_\epsilon \leq C\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\|$, for some $C>0$ depending on β and χ, we conclude that

$$\|\chi^{1/2}\operatorname{curl}(\tilde{u}-u)\| \leq C(\beta,\chi)(\|u-u_h\|_{\beta,\chi}+osc(f)+osc(q_h)).$$

■

2.4.4 Numerical tests

We first check the reliability of our estimator. For that purpose, we solve the two-dimensional Maxwell equations on the unit square $\Omega=(0,1)^2$. We take isotropic meshes composed of triangles and use the lowest order Nédélec, the $\mathcal{P}_1$-conforming and the first order Raviart-Thomas finite elements to compute the finite element solution u_h and the fluxes j_h and q_h, respectively.

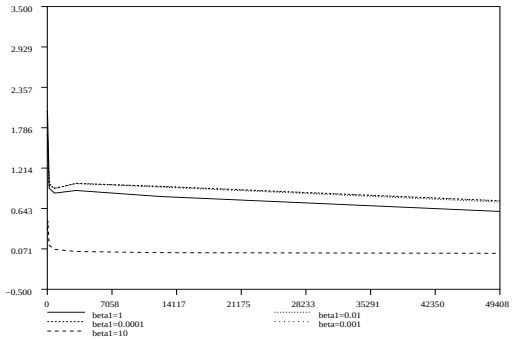


FIG. 2.13 – The ratio $|||u-u_h|||/\eta$ wrt to DoF for the first example

As first example, we suppose that Ω admits a decomposition into four sub-domains $\Omega_1=(0,0.5)^2$, $\Omega_2=(0.5,1)\times(0,0.5)$, $\Omega_3=(0.5,1)^2$ and $\Omega_4=(0,0.5)\times(0.5,1)$ and introduce the exact solution

$$u=\operatorname{curl}\varphi \text{ where } \varphi=[y(1-y)(2y-1)x(1-x)(2x-1)]^2.$$

We fix $\chi_i=1$, for all $i=1,\ldots,4$, $\beta_2=\beta_4=1$ and take different values of $\beta_1=\beta_3$. In Figure 2.14, we have plotted the error and the estimator for two values of β (the other values of β give rise to similar results), there we see that the approximated solution converges toward the exact one with a convergence rate of 1 and that the estimator has a similar behavior. This is confirmed in Figure 2.13, where we present, for some values of β, the effectivity index, i.e., the ratio $|||u-u_h|||/\eta$. From this figure we can say that our estimator is reliable since the effectivity index is bounded by approximatively 0.75.

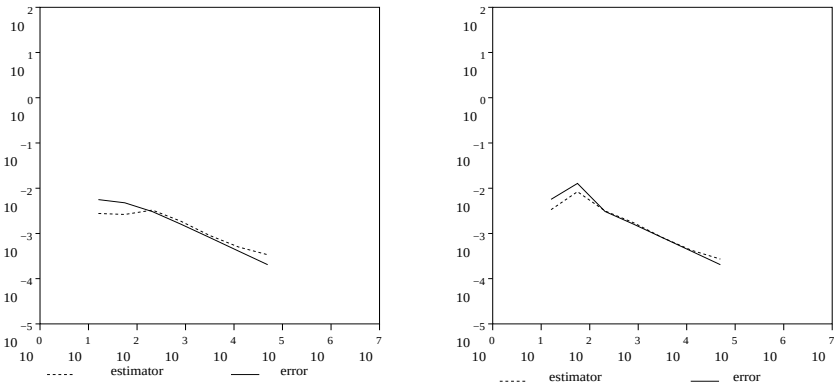


FIG. 2.14 – $|||u - u_h|||$ and η wrt DoF for example 1 with $\beta_1 = 1$ (left) and $\beta_1 = 0.0001$ (right).

As second example we take the exact solution

$$u = \nabla \left(e^{-x/\sqrt{\varepsilon}} x(1-x)y(1-y) \right)$$

on the domain Ω and fix $\beta = 1$ and $\chi = \varepsilon$ for different values of ε. This solution presents an exponential boundary layer of width $O(\sqrt{\varepsilon})$ along the line $x = 0$.

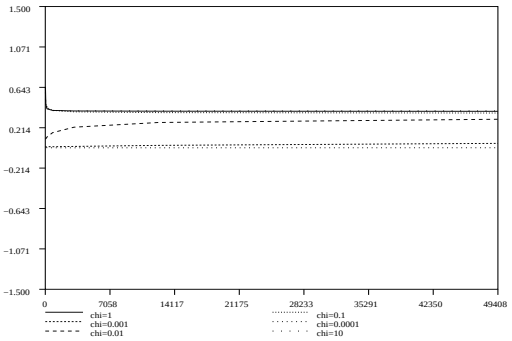


FIG. 2.15 – The ratio $|||u - u_h|||/\eta$ wrt to DoF for example 2.

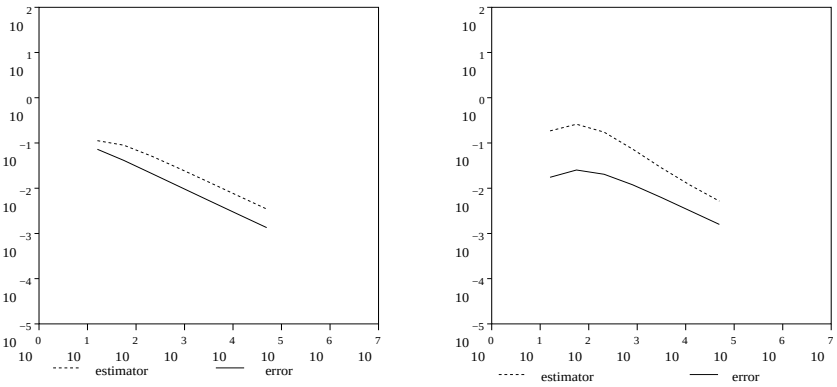


FIG. 2.16 – $|||u - u_h|||$ and η wrt to DoF for example 2 with $\chi = 1$ (left) and $\chi = 0.01$ (right).

As before we show in Figure 2.16 the error and the estimator for two values of ϵ and we see a convergence rate of 1 for the error and the estimator. In Figure 2.15, we present the effectivity index, for some values of ϵ. Again we can assert that our estimator is reliable since the effectivity index is bounded by approximatively 0.38.

As for second order scalar problems, to illustrate the performance of our estimator, we present on two typical examples the meshes obtained after some iterations using an iterative algorithm based on the same marking and refinement procedures.

For the first example we take $\Omega = (-1, 1)^2$, with $\chi = 1$ and a discontinuous coefficient $\beta = a$, corresponding to the decomposition of Ω into 4 sub-domains Ω_i, $i = 1, \ldots, 4$ from the second example of section 2.3.2. As exact solution, we take $u = \nabla S$, where S is given by (4.25). Such a solution is a typical singularity of the Maxwell system at $(0, 0)$ [22] (it belongs to $H(\text{curl}) \cap H(\text{div})$ but not to $(H^1)^2$). Therefore a refinement of the mesh near this point can be expected. This is confirmed by Figures 2.17 on the meshes obtained for $a_1 = 5$ and $a_1 = 100$ respectively.

Remark 2.4.6. *Unlike the domain, that is symetrical with respect to the straight line of equation $y = -x$, the exact solution, i.e. $u = \nabla S$, is not symetrical with respect to this straight line. This implies that the adpative meshes obtained are not necessarily symetrical.*

Finally as second example, we take the L-shape domain $\Omega = (-1, 1)^2 \setminus (-1, 0) \times (0, 1)$, $\chi = \beta = 1$, and as exact solution $u = \nabla S$, where S is given by (2.19). Again this solution is a typical singularity of the Maxwell system at $(0, 0)$ [22]. A refinement of the mesh near this point is once more confirmed numerically in Figure 2.18.

As for the reaction-diffusion problems, all these tests allow to conclude that our proposed estimator is reliable and efficient. Note further that the effectivity index always remains under the value of 1, as theoretically predicted.

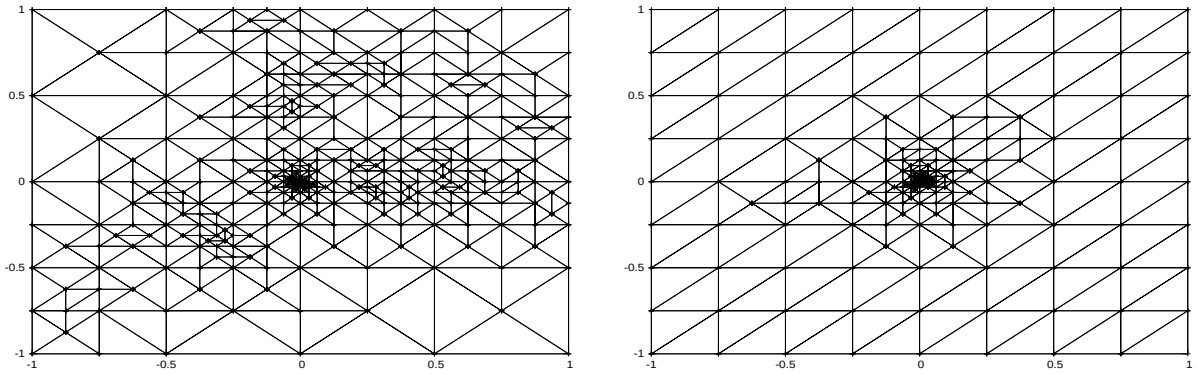


FIG. 2.17 – Adaptive mesh after 15 iterations on the left and 10 iterations on the right for the first example and criterion $\eta_T > 0.75 \max_{T'} \eta_{T'}$, with respectively $a_1 = 5$ and $a_1 = 100$.

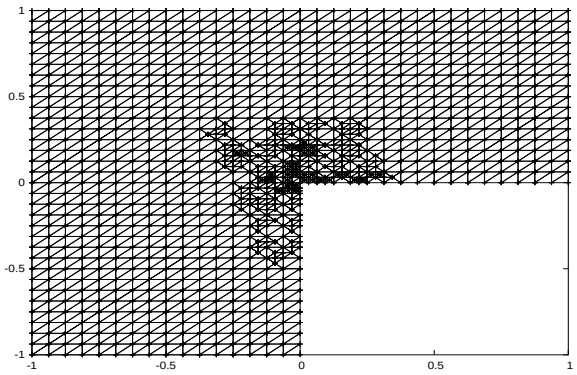


FIG. 2.18 – Adaptive mesh after 10 iterations for the second example and criterion $\eta_T > 0.75 \max_{T'} \eta_{T'}$.

Chapitre 3

Comparison of the three a posteriori error estimators

In this chapter, we compare all the estimators we have constructed for the Maxwell equations and that we will denote as follows

$$\begin{aligned}
\eta_{T,0}^2 &= h_T^2\beta_T^{-1}\|\operatorname{div}(\beta \boldsymbol{u}_h)\|_T^2 + \sum_{e\subset\partial T} h_e\beta_e^{-1}\|\mathbf{J}_{e,n}\|_e^2,\\
\eta_{Ned}^2 &= \sum_{T\in\mathcal{T}_h}\left(\eta_{T,0}^2 + h_T^2\beta_T^{-1}\|\boldsymbol{r}_T\|_T^2 + \sum_{e\subset\partial T} h_e\beta_e^{-1}\|\mathbf{J}_{e,t}\|_e^2\right),\\
\eta_{CN}^2 &= \sum_{T\in\mathcal{T}_h}\left(\eta_{T,0}^2 + \alpha_T^2\|\boldsymbol{r}_T\|_T^2 + \sum_{e\subset\partial T} \chi_e^{-1/2}\alpha_e\|\mathbf{J}_{e,t}\|_e^2\right),\\
\eta_{flux}^2 &= \sum_{T\in\mathcal{T}_h}\left(\|\beta^{-1/2}(\beta u_h - q_h)\|_T^2 + \|\chi^{-1/2}\left(\chi\operatorname{curl} u_h - j_h\right)\|_T^2\right).
\end{aligned}$$

They are tested with the iterative algorithm using the marking procedure presented in the second chapter :

$$\eta_T > 0.75\max_{T'}\eta'_{T'}.$$

For that purpose, we solve the two-dimensional Maxwell equations and take meshes composed of triangles. We use the lowest order Nédélec, the $\mathcal{P}_1$-conforming and the first order Raviart-Thomas finite elements to compute the finite element solution u_h and the fluxes j_h and q_h, respectively. We present here, for three kind of exact solutions, the meshes obtained for the different estimators, with the same number of iterations, where the triangles of the meshes are colored with respect to the value of their local error (See Fig. 3.1)

3.1 A solution with a boundary layer

As first example, we consider, on the unit square $\Omega = (0,1)^2$, the exact solution :

$$u = \nabla\left(e^{-x/\sqrt{\varepsilon}}x(1-x)y(1-y)\right)$$

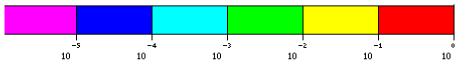


FIG. 3.1 – value of the local error.

presenting a boundary layer along the axis $x = 0$, for two different values of ε. In a first time, we impose a limitation on the minimal angle to refine our mesh and compare our estimators for different steps of the refinement procedure, in Figures 3.2 and 3.3 for $\varepsilon = 10^{-1}$ and Figures 3.4 and 3.5 for $\varepsilon = 10^{-3}$.

As expected, the meshes obtained, in all the tests, are refined on the boundary layer. In Figure 3.2, η_{CN} and η_{Ned} seems to have a similar behaviour and this is confirmed by the test for $\varepsilon = 10^{-3}$ (see Fig. 3.4) where the meshes obtained after 15 iterations are exactly the same for the two estimators.

Now comparing η_{CN} with the flux estimator η_{flux}, we notice, in Figures 3.3 and 3.5, that this last one needs fewer iterations than η_{CN} to get a fine mesh and the local errors decrease fastlier. This is pointed up with the zooms we make on a part of the layer (see Fig. 3.6 - 3.7)

In a second time, we no more impose this minimal angle on the refinement procedure and have a look on the meshes obtained after 15 iterations (see Fig. 3.8 - 3.9) for the different value of ε. We do not represent the mesh obtained for the Nédélec estimator as it is the same as for the Clément-Nédélec estimator. In this case, we know that the mesh in the boundary layer should be composed of thin triangles with large edges parallel to the boundary axis $x = 0$. This phenomenon is quite presented in the case that we use the equilibrated estimator η_{flux} but it is no more the case for η_{CN}. This can be explained by the fact that, for η_{flux}, the theory for the upper bound, presented for isotropic meshes, remains valid for anisotropic meshes, with the constant appearing in the upper bound still equal to 1. On the contrary, for η_{CN}, this constant depends on the mesh and we can conclude that the theory has to be adapted for an anisotropic mesh.

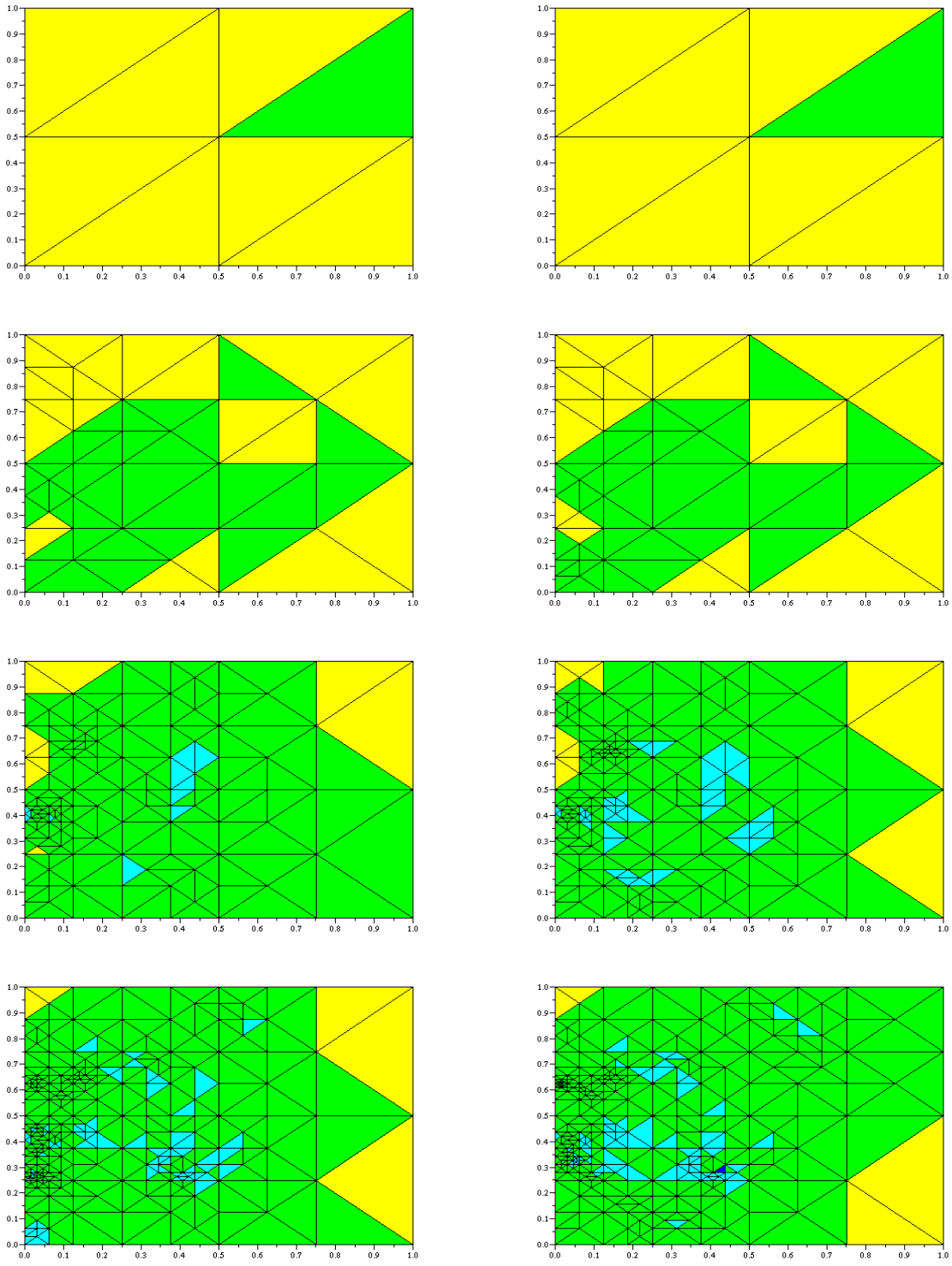


FIG. 3.2 – Iterative meshes obtained for the solution with boundary layer for $\varepsilon = 0.1$: on the left for η_{CN}, on the right for η_{Ned}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

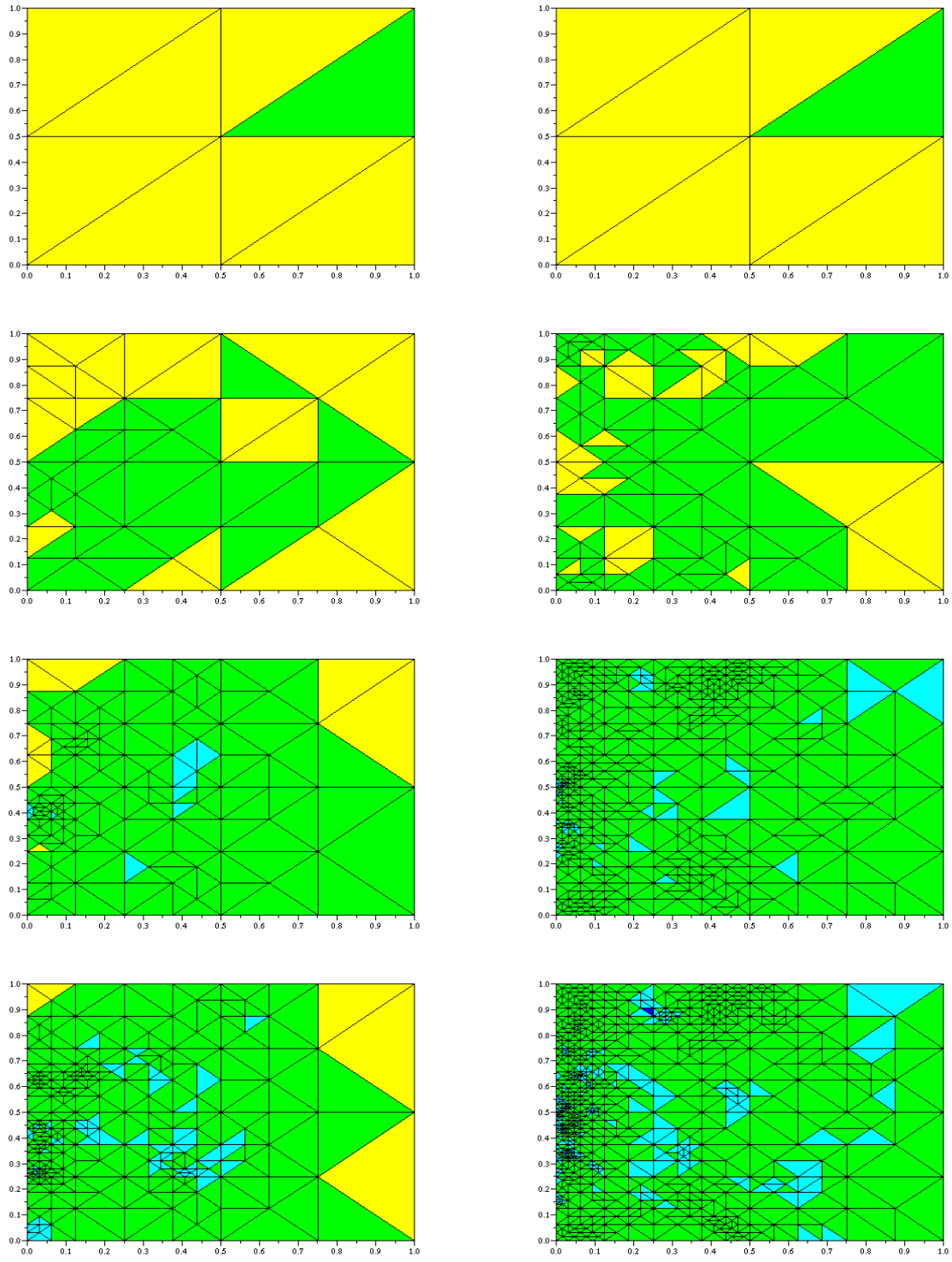


FIG. 3.3 – Iterative meshes obtained for the solution with boundary layer for $\varepsilon = 0.1$: on the left for η_{CN}, on the right for η_{flux}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

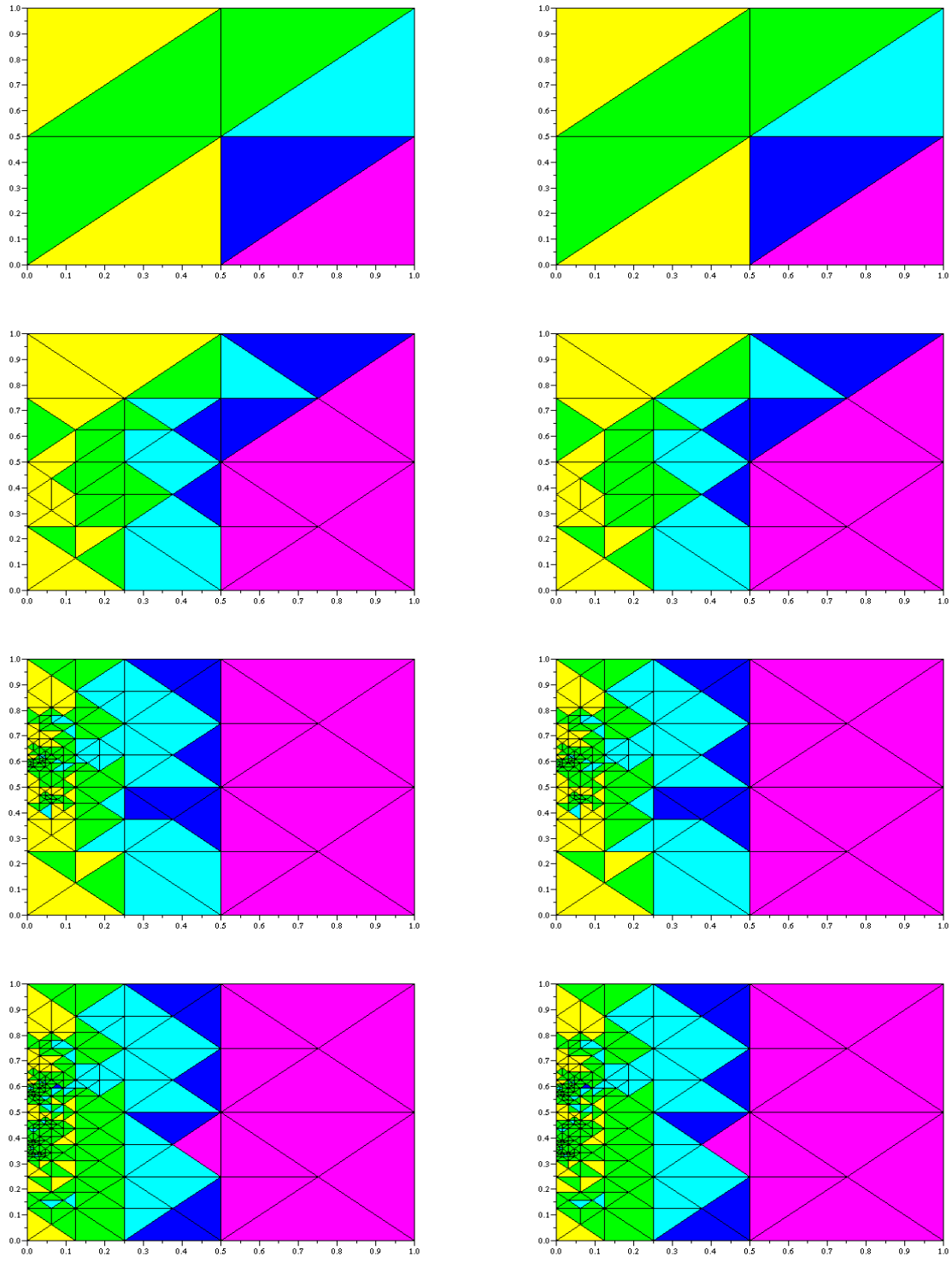


FIG. 3.4 – Iterative meshes obtained for the solution with boundary layer for $\varepsilon = 0.001$: on the left for η_{CN}, on the right for η_{Ned}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

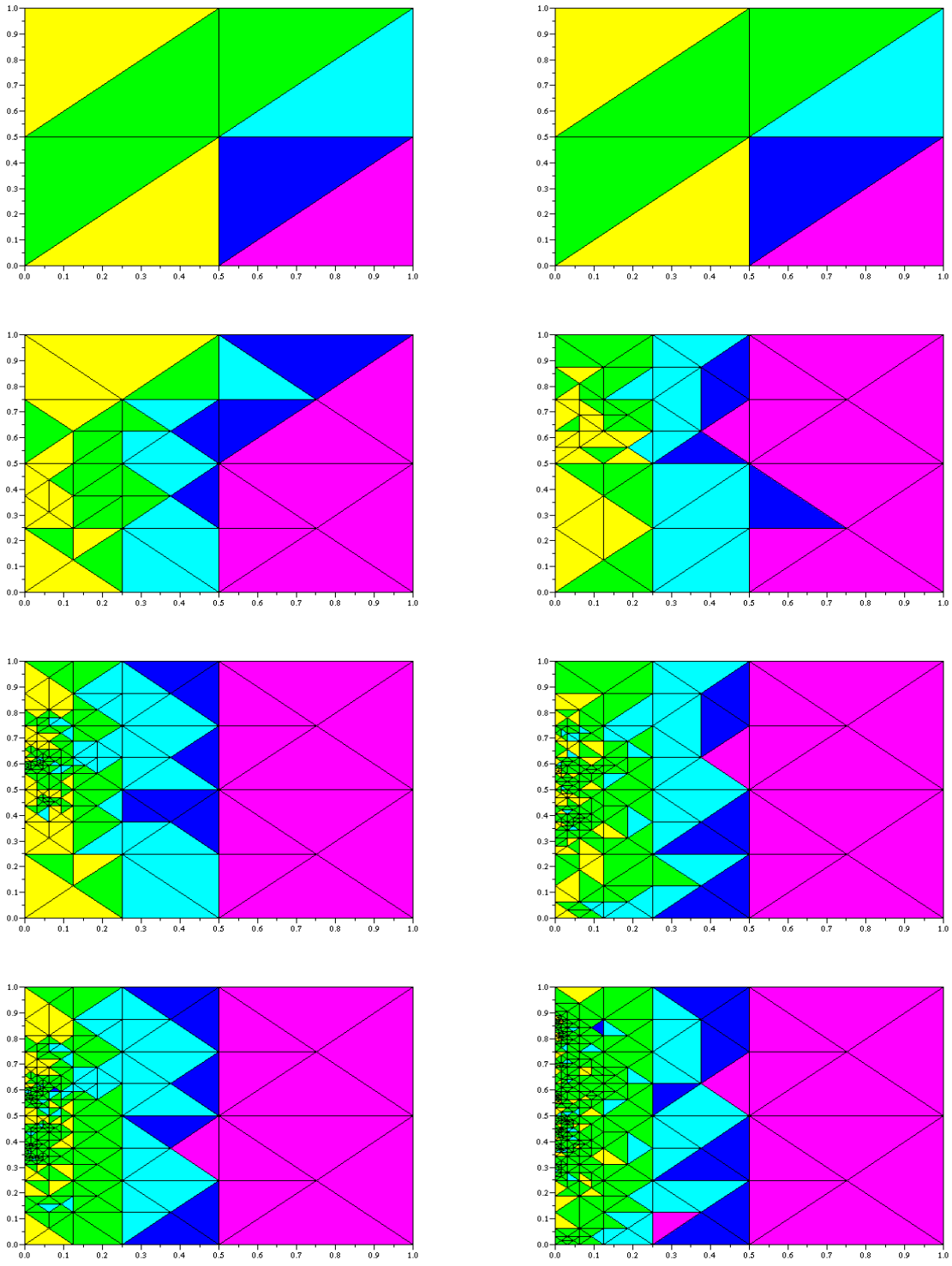


FIG. 3.5 – Iterative meshes obtaineded for the solution with boundary layer for $\varepsilon = 0.001$: on the left for η_{CN}, on the right for η_{flux}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

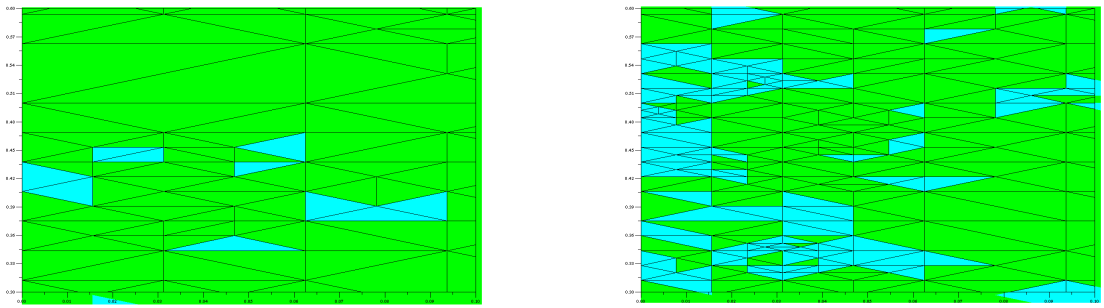


FIG. 3.6 – Iterative meshes obtained for the solution with boundary layer $\varepsilon = 0.1$, zoom on the intervall $(0, 0.1) \times (0.3, 0.6)$ in the layer : on the left for η_{CN}, on the right for η_{flux}, after 15 iterations.

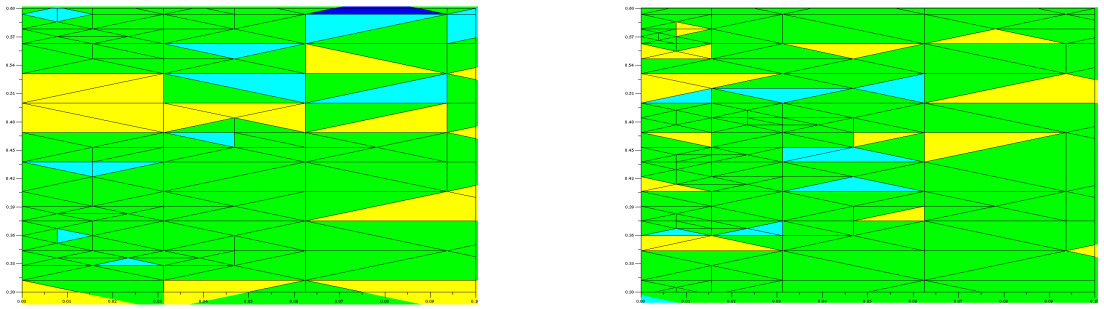


FIG. 3.7 – Iterative meshes obtained for the solution with boundary layer $\varepsilon = 0.001$, zoom on the intervall $(0, 0.3) \times (0.3, 0.6)$ in the layer : on the left for η_{CN}, on the right for η_{flux}, after 15 iterations.

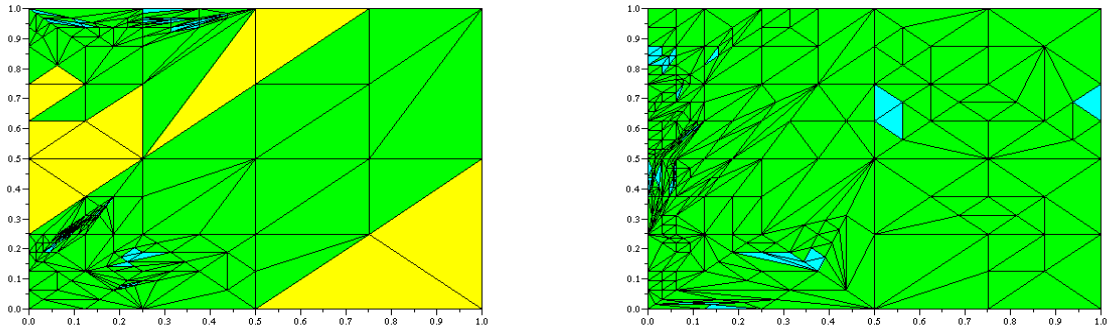


FIG. 3.8 – Iterative meshes obtained for the solution with boundary layer $\varepsilon = 0.1$ without minimal angle : on the left for η_{CN}, on the right for η_{flux}. On the top, the meshes obtained after 15 iterations, on the bottom, we zoom on the intervall $(0, 0.06) \times (0.2, 0.4)$ in the layer.

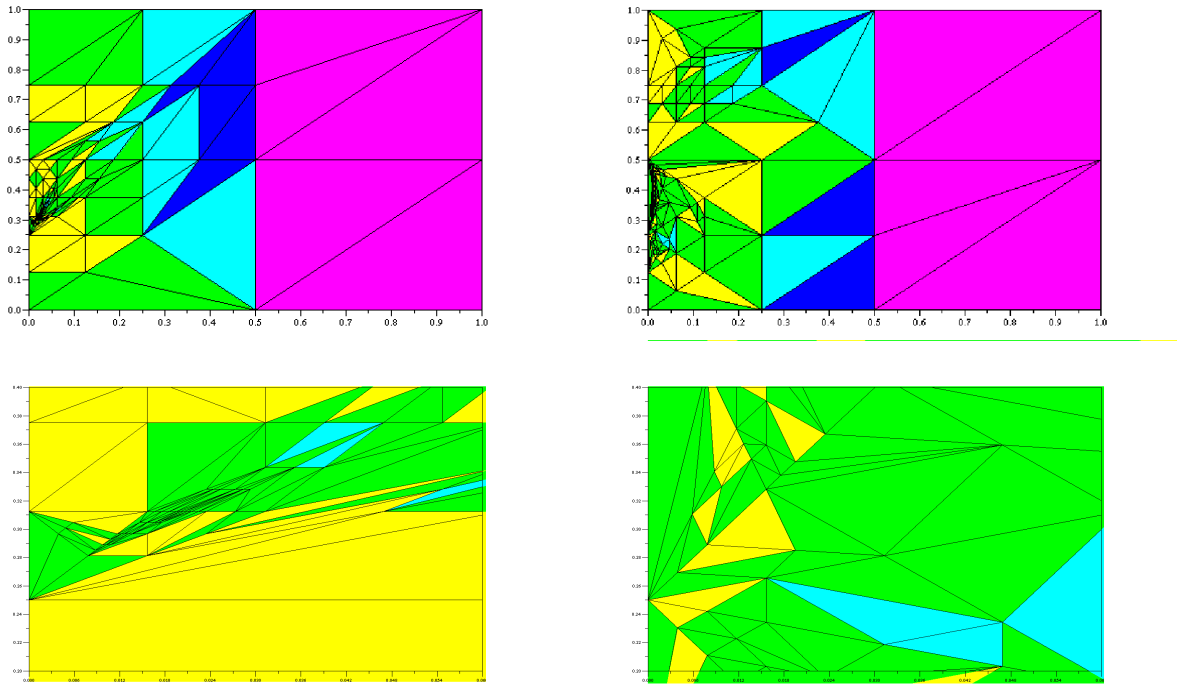


FIG. 3.9 – Iterative meshes obtained for the solution with boundary layer $\varepsilon = 0.001$ without minimal angle : on the left for η_{CN}, on the right for η_{flux}. On the top, the meshes obtained after 15 iterations, on the bottom, we zoom on the intervall $(0, 0.06) \times (0.2, 0.4)$ in the layer.

3.2 The checkerboard : a test with discontinuous coefficients

As second example, we take $\Omega = (-1,1)^2$, with $\chi = 1$ and a discontinuous coefficient β, corresponding to the decomposition of Ω into 4 sub-domains Ω_i, $i = 1, \ldots, 4$ from the second example of section 2.3.2. As exact solution, we take $u = \nabla S$, where S is given by (4.25). We know that a refinement of the mesh near this point is expected. Figures 3.10 - 3.11 and 3.12 - 3.13 represent the meshes obtained for 4 subdomains with $\beta_1 = 5$ and $\beta_1 = 100$ respectively.

The estimators η_{CN} and η_{Ned} have, one more time, quite the same behaviour. In such tests, we use large values of β and we remark that those two estimators have a common part, corresponding to $\sum_{T \in \mathcal{T}_h} \eta_{T,0}^2$, which involves the jump of the component $\beta u_h \cdot n$ over edges in the term denoted $\mathbf{J}_{e,n}$. This part of the estimator is the dominant one, that explain such a similar behaviour.
Comparing now η_{CN} with η_{flux}, we still first remark that η_{flux} is more efficient because it refines fastlier than η_{CN} in few iterations. If we look at Figure 3.11 we notice that, in the beginning, when the mesh is coarse, the residual estimator better localises the singularity and the area to refine. But finer becomes the mesh, better the equilibrated estimator localises and refines efficiently the singularity. This is confirmed by the zooms we make on the singularity, where we notice that the local error near the point $(0,0)$ is smaller for η_{flux} after 15 iterations.

3.3 A singular solution on the L-shape domain

Finally, we take the L-shape domain $\Omega = (-1,1)^2 \setminus (-1,0) \times (0,1)$, $\chi = \beta = 1$, and as exact solution $u = \nabla S$, where S is given by (2.19). A refinement of the mesh near this point is once more obtained numerically in Figures 3.17 and 3.18.

As already mentioned before, we find the same meshes for the residual estimators. We have to remark that, this time, they are better than the ones obtained for η_{flux}. Indeed, they better localise the refinement of the singularity and the local error better decreases. They are less diffusive.

3.4 Conclusion

Unlike in chapter 1, the two kind of residual estimators η_{CN} and η_{Ned} are very similar in all tests we presented. Indeed, in chapter 1, it was proved that, when the coefficients take

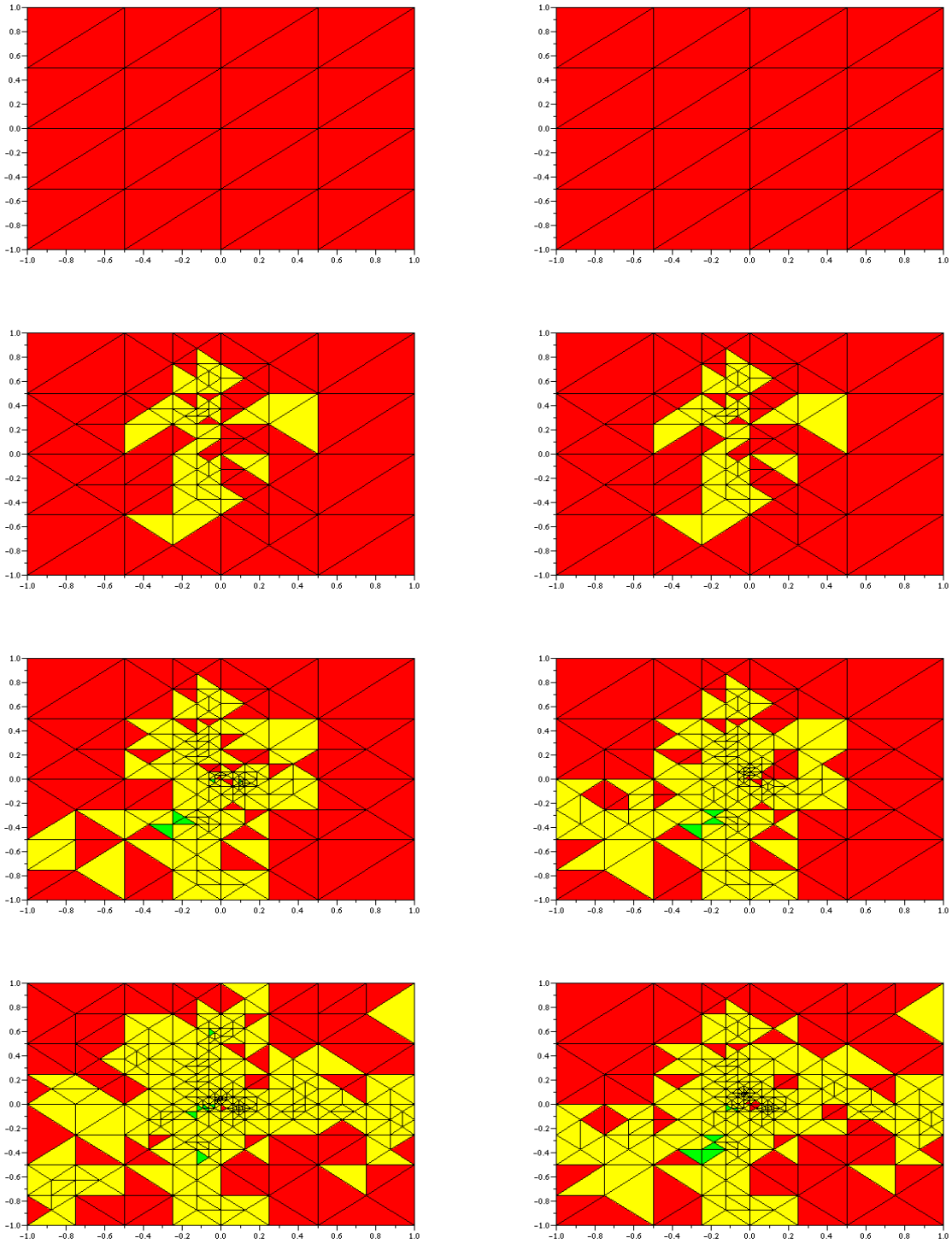


FIG. 3.10 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 5$: on the left for η_{CN}, on the right for η_{Ned}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

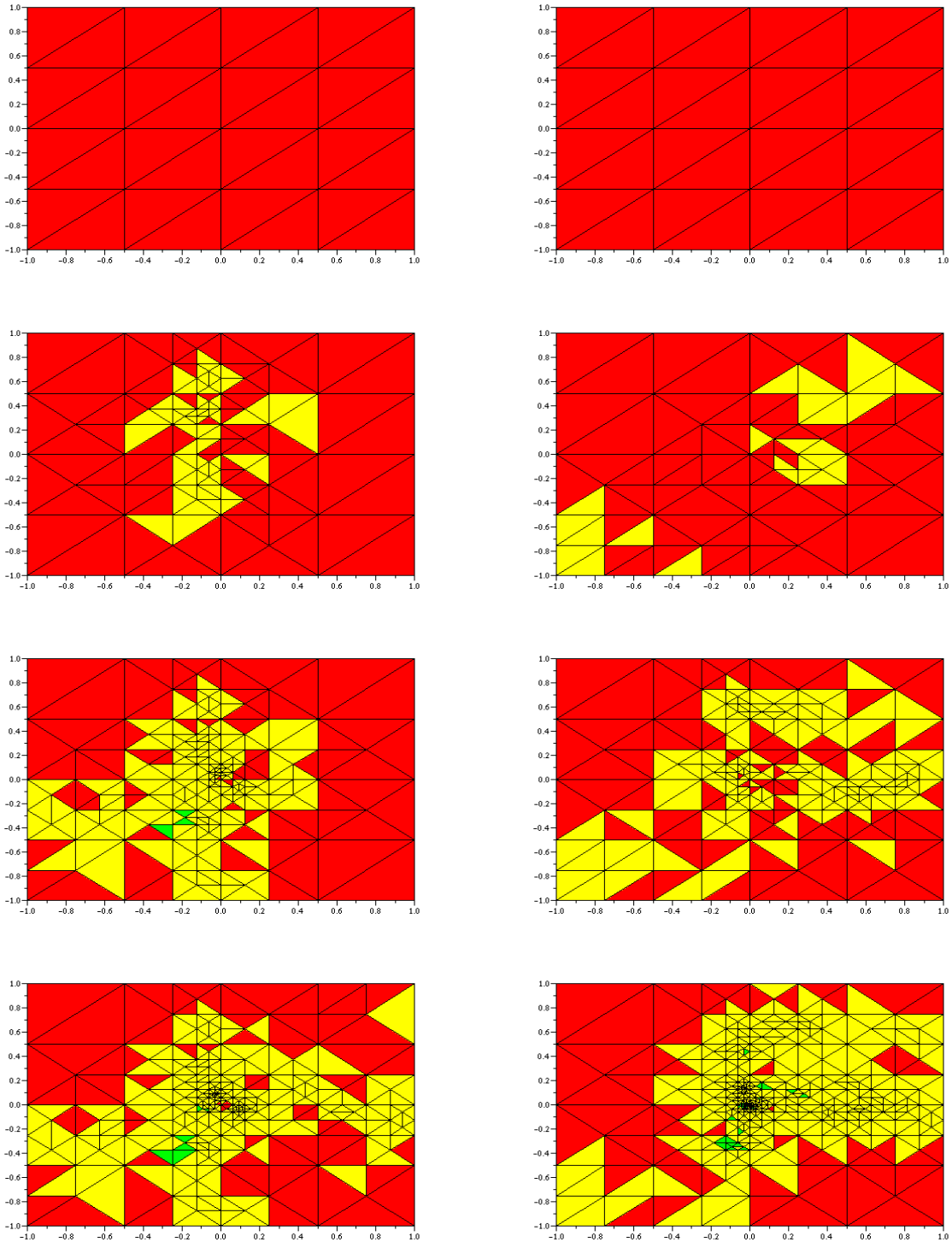


FIG. 3.11 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 5$: on the left for η_{Ned}, on the right for η_{flux}, from the top to the bottom, the initial mesh, after 5, 10 and 15 iterations.

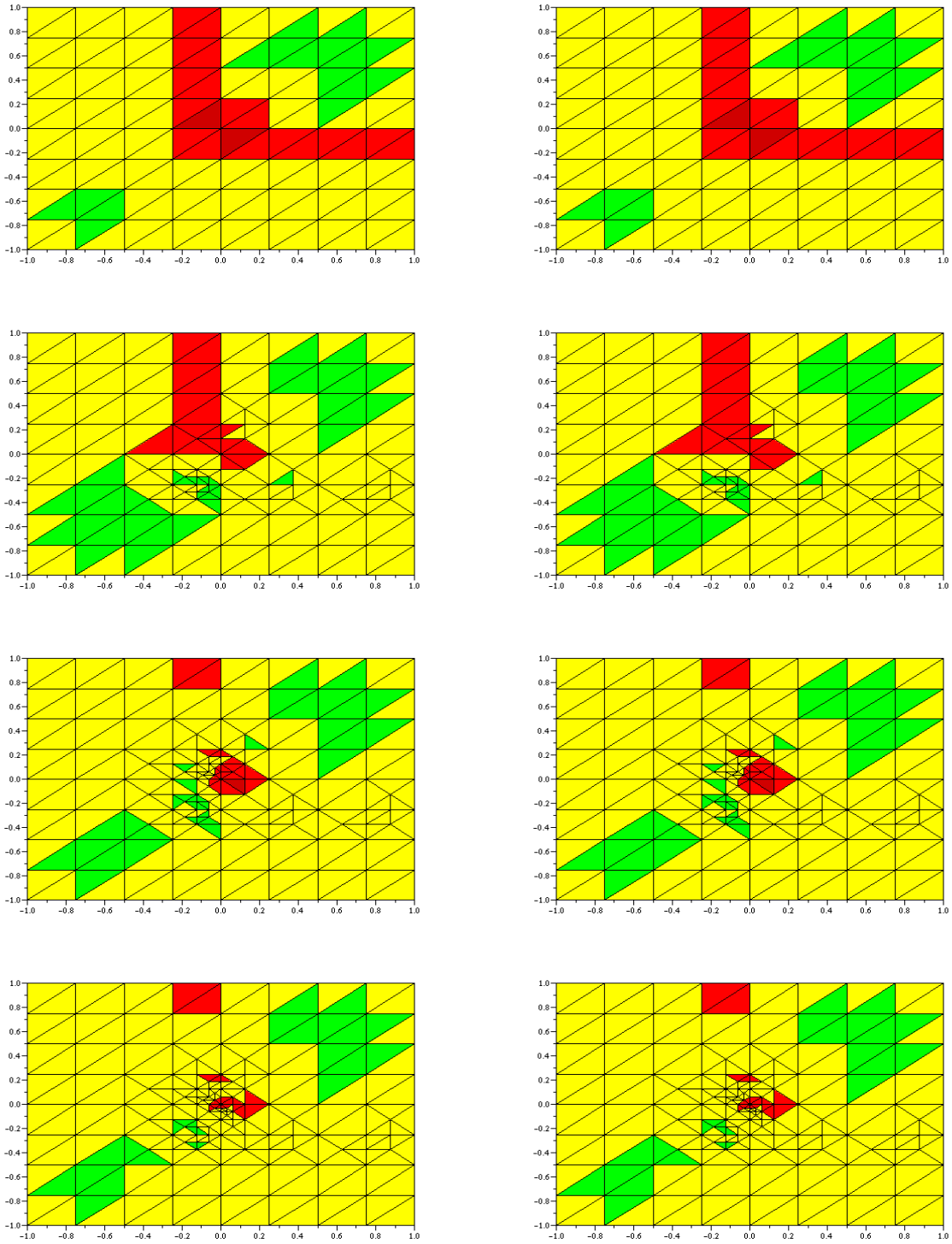


FIG. 3.12 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 100$: on the left for η_{CN}, on the right for η_{Ned}, from the top to the bottom, the initial mesh, after 3, 5 and 7 iterations.

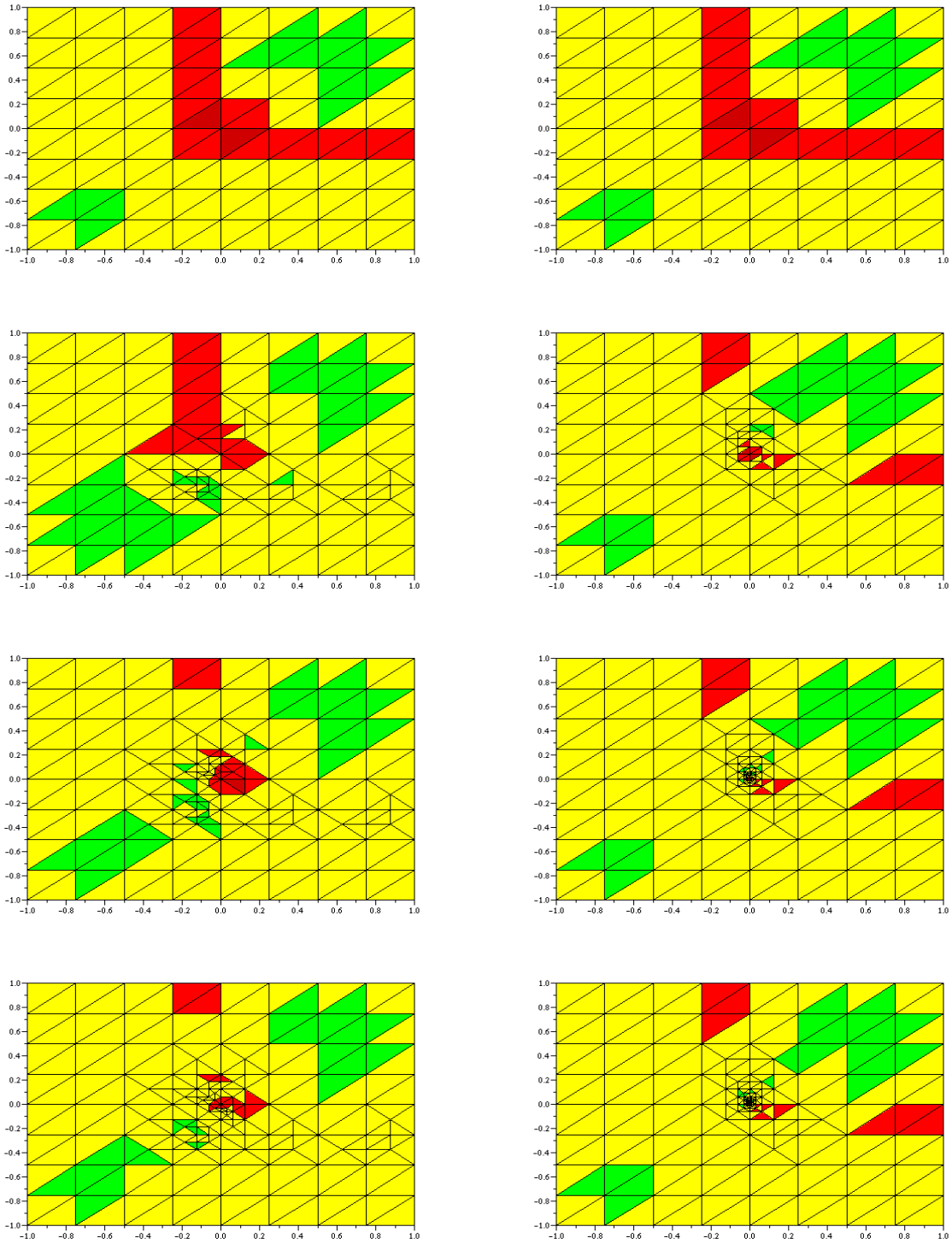


FIG. 3.13 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 100$: on the left for η_{CN}, on the right for η_{flux}, from the top to the bottom, the initial mesh, after 3, 5 and 7 iterations.

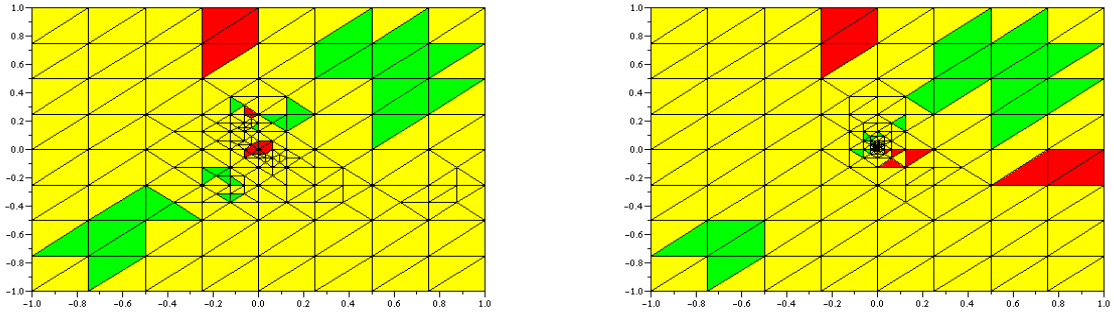


FIG. 3.14 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 100$: on the left for η_{CN}, on the right for η_{flux}, after 9 iterations.

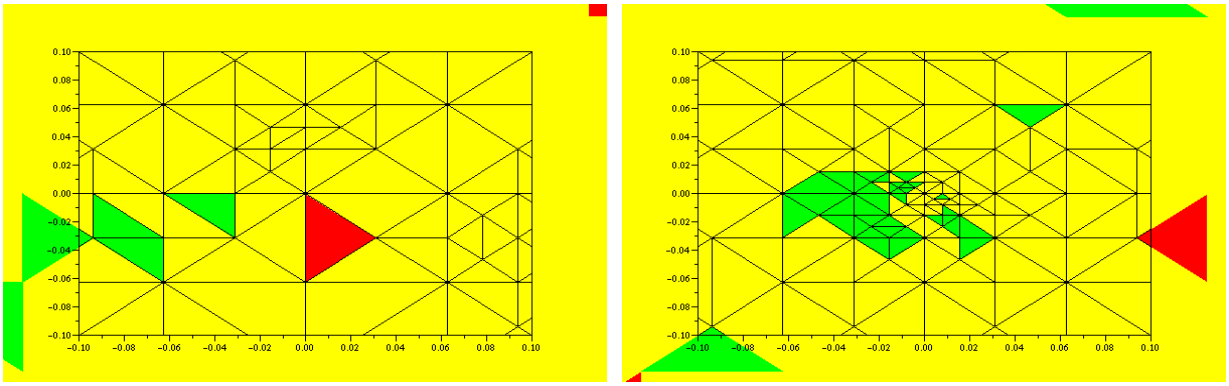


FIG. 3.15 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 5$, zoom on the singularity on $(-0.1, 0.1)^2$: on the left for η_{CN}, on the right for η_{flux}, after 15 iterations.

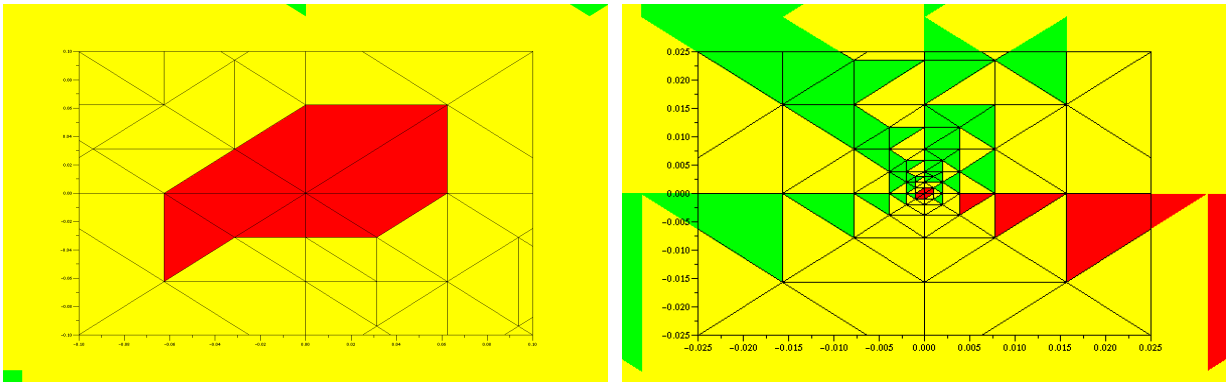


FIG. 3.16 – Iterative meshes obtained for the second solution with 4 subdomains for $\beta_1 = 100$, zoom on the singularity : on the left for η_{CN} on $(-0.025, 0.0.025)^2$, on the right for η_{flux} on $(-0.1, 0.1)^2$, after 9 iterations.

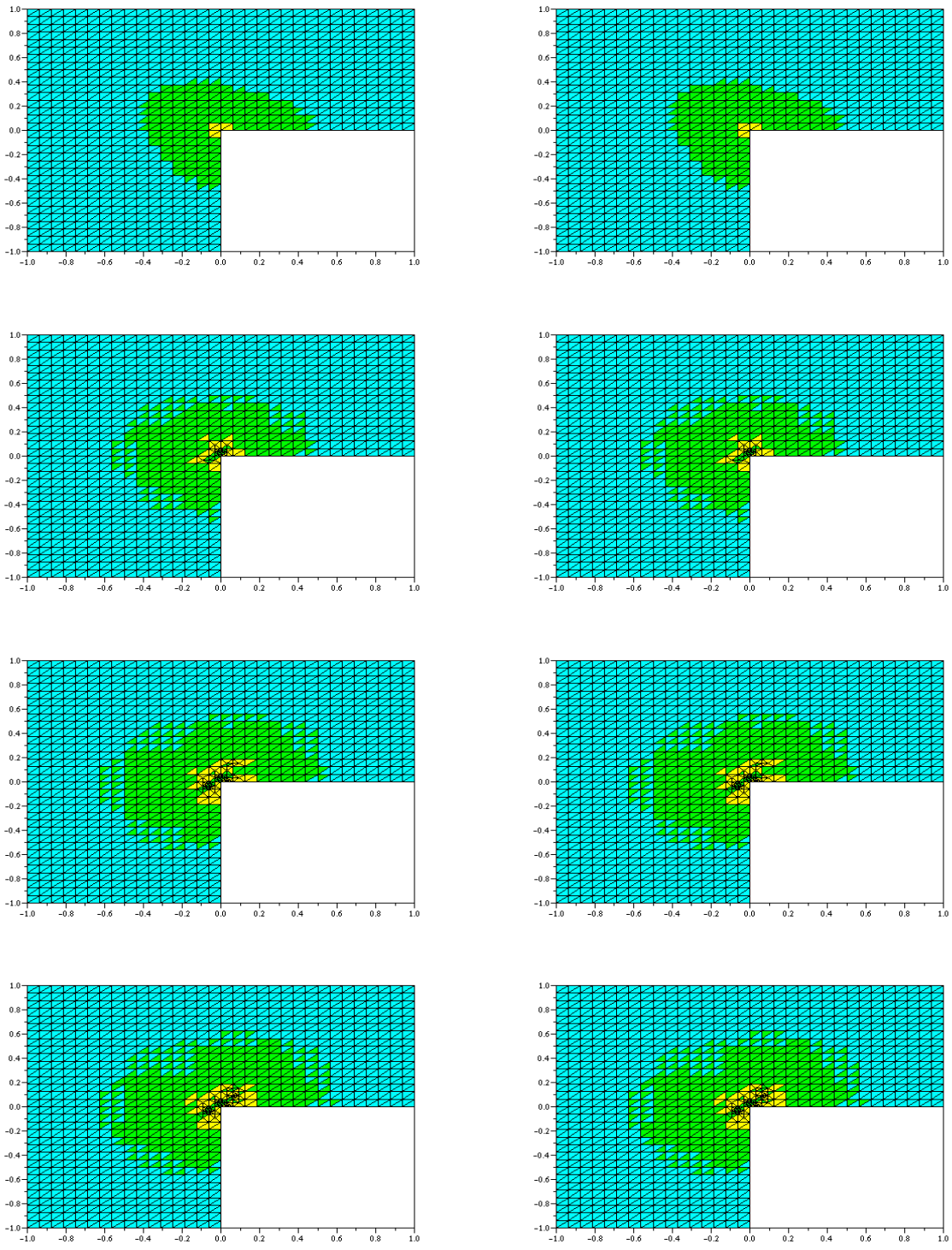


FIG. 3.17 – Iterative meshes obtained for the third example : on the left for η_{CN}, on the right for η_{Ned}, from the top to the bottom, the initial mesh, after 2, 4 and 6 iterations.

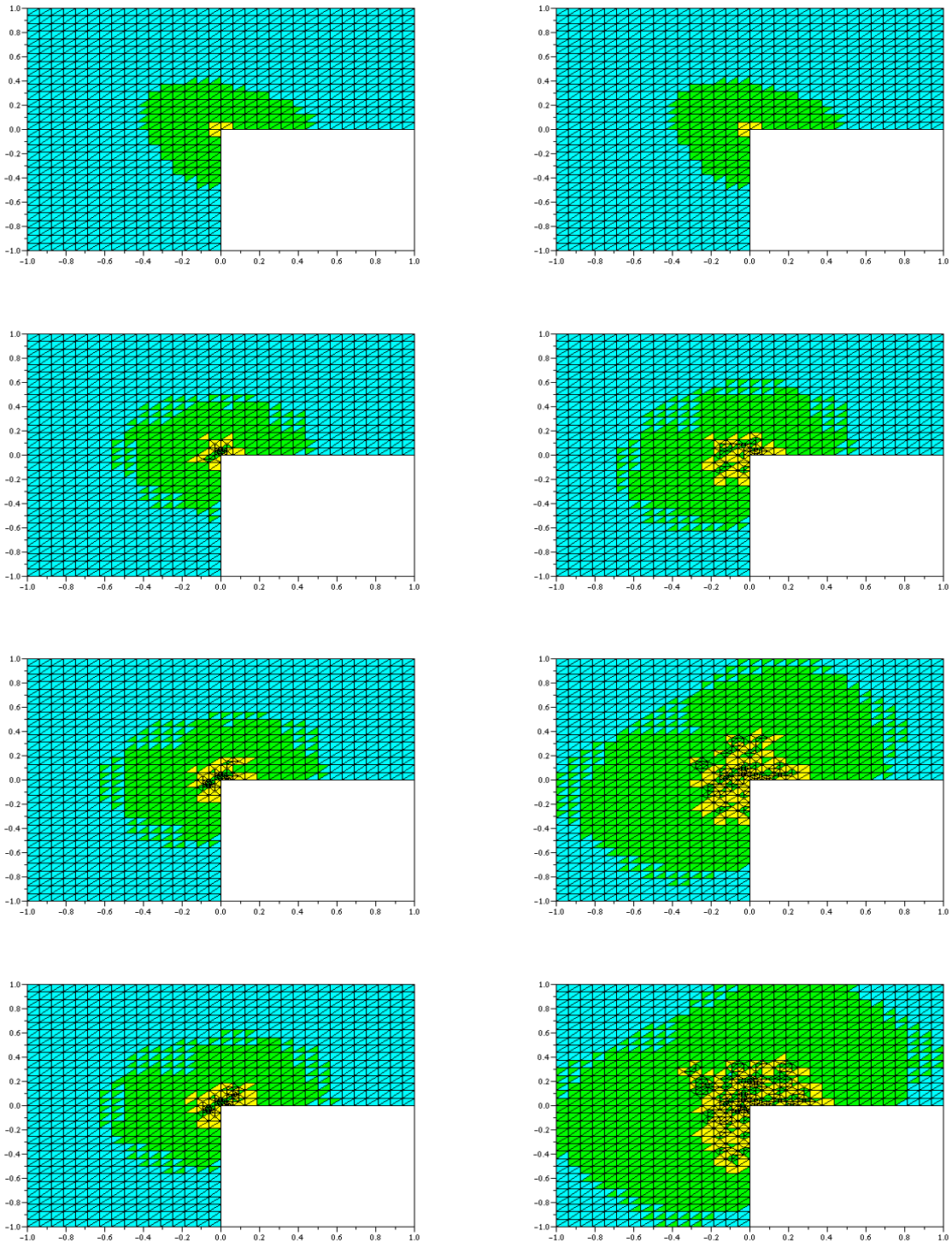


FIG. 3.18 – Iterative meshes obtained for the third example : on the left for η_{CN}, on the right for η_{flux}, from the top to the bottom, the initial mesh, after 2, 4 and 6 iterations.

a large type of values, η_{CN} remains more robust than η_{Ned}. This can may be explained by the fact that the values of the coefficients are constant in two of the three tests and that in the last one, for a solution in the chekerboard, the coefficient β is to large to notice the difference between the two estimators. To be sure, we program this test for small values of the coefficient like 10^{-1} or 10^{-3} and we see one more time that they are equivalent. Indeed, altough the second part of the estimator get a different value, depending on whether it is η_{Ned} or η_{CN}, the irrotational part η_0 remain much more dominant.

All the tests presented before prove that the estimator built with fluxes, η_{flux}, is more performant than the two others. Indeed, unless we can see, in particular on the L-shape domain, that it is more diffusive, it builds, fastlier than the other, an adapted mesh. We need less iterations, compared to the residual estimators, to obtain a mesh well refined near the singularities as expected. this might be explained by the fact that the constant in the upper bound (2.36) is equal to 1 whereas for the others the constant in the bound not only depend on the coefficients but also on the triangulation. These constants may underestimate the error and, altough the value of the local indicators decreases, the error remains large locally.

Chapitre 4

Equilibrated error estimators for discontinuous Galerkin methods

We consider some diffusion problems in domains of $\mathbb{R}^d$, $d = 2$ or 3 approximated by a discontinuous Galerkin method with polynomials of any degree. We propose a new a posteriori error estimator based on $H(\mathrm{div})$- conforming elements. It is shown that this estimator gives rise to an upper bound where the constant is one up to higher order terms. The lower bound is also established with a constant depending on the aspect ratio of the mesh, the dependence with respect to the coefficients being also traced. The reliability and efficiency of the proposed estimator is confirmed by some numerical tests.

4.1 Introduction

Among other methods, the finite element method is one of the more popular that is commonly used in the numerical realization of different problems appearing in engineering applications, like the Laplace equation, the Lamé system, the Stokes system, the Maxwell system, etc.... (see [14, 17, 39]). More recently discontinuous Galerkin finite element methods become very attractive since they present some advantages, like flexibility, adaptivity, etc... In our days a quite large literature exists on the subject, we refer to [4, 20] and the references cited there. Adaptive techniques based on a posteriori error estimators have become indispensable tools for such methods. For continuous Galerkin finite element methods, there now exists a vast amount of literature on a posteriori error estimation for problems in mechanics or electromagnetism and obtaining locally defined a posteriori error estimates. We refer to the monographs [3, 6, 40, 52] for a good overview on this topic. On the other hand a similar theory for discontinuous methods is less developed, let us quote [10, 16, 24, 27, 28, 30, 51].

Usually upper and lower bounds are proved in order to guarantee the reliability and the efficiency of the proposed estimator. Most of the existing approaches involve constants depending on the shape regularity of the elements and/or of the jumps in the coefficients ; but these dependences are oftenly not given. Only a few number of approaches gives rise to

estimates with explicit constants, let us quote [3, 13, 35, 38, 42, 46] for continuous methods. For discontinuous methods, we may cite the recent preprints [2, 36] : in the first one, the author considers second order elliptic operators in two-dimensional domains and a discontinuous method with polynomials of degree 1, while in the second preprint the authors present essentially numerical experiments.

Our goal is therefore to consider second order elliptic operators in two- or three-dimensional domains with mixed boundary conditions and a discontinuous method with polynomials of any degree and further to derive an a posteriori estimator with an explicit constant in the upper bound (more precisely 1) up to oscillating terms. Our approach, called the equilibrated approach [2, 13, 36, 46], is based on the following idea : it consists to build a vector field j_h, which is a $H(\mathrm{div})$-conforming approximation of the stress, i.e., it solves

$$\mathrm{div}\, j_h = -\Pi f,$$

where Πf is the L^2 projection of the right-hand side f on the set of piecewise polynomial functions on the triangulation. Then we use $j_h - a\nabla u_h$ as estimator for the conforming part of the error, when u_h is the finite element approximation of the exact solution. The difference with [2] relies on the determination of j_h that we obtain here by using Raviart-Thomas finite elements instead of $\mathcal{P}_1$ elements. The use of Raviart-Thomas finite elements seems to be more natural, allows to use polynomials of any degree and to consider any space dimension.

Note that the non conforming part of the error is managed using a Helmholtz decomposition of the error and a standard Oswald interpolation operator [2, 30]. Furthermore using standard inverse inequalities, we show that our estimator is locally efficient but in the lower bound, we trace the dependence of the constant with respect to the variation of the coefficients of the differential operator.

The schedule of the chapter is as follows : We recall in section 2 the diffusion problem, its numerical approximation and an appropriate Helmholtz decomposition of the error. Section 3 is devoted to the introduction of the estimator based on Raviart-Thomas elements and the proofs of the upper and lower bounds. The upper bound directly follows from the construction of the estimator, while the lower bound requires the use of some inverse inequalities and some properties of the equilibrated fluxes. Finally in section 4 some numerical tests are presented that confirm the reliability and efficiency of our estimator.

Let us finish this introduction with some notation used in the remainder of the chapter : On D, the $L^2(D)$-norm will be denoted by $\|\cdot\|_D$. In the case $D = \Omega$, we will drop the index Ω. The usual norm and semi-norm of $H^s(D)$ $(s \geq 0)$ are denoted by $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$, respectively. Finally, the notation $a \lesssim b$ and $a \sim b$ means the existence of positive constants C_1 and C_2, which are independent of the mesh size, of the quantities a and b under consideration and of the coefficients of the operators such that $a \lesssim C_2 b$ and $C_1 b \lesssim a \lesssim C_2 b$, respectively. In other words, the constants may depend on the aspect ratio of the mesh as well as the polynomial degree l (see below).

4.2 The boundary value problem and its discretization

Let Ω be a bounded domain of $\mathbb{R}^d, d = 2$ or 3 with a Lipschitz boundary Γ that we suppose to be polygonal ($d = 2$) or polyhedral ($d = 3$). We further assume that Ω is simply connected and that Γ is connected. We consider the following elliptic second order boundary value problem with non homogeneous mixed boundary conditions :

$$\left\{\begin{array}{rcl} -\mathrm{div}\,(a\,\nabla u) & = & f \text{ in } \Omega, \\ u & = & g_D \text{ on } \Gamma_D, \\ a\nabla u \cdot n & = & g_N \text{ on } \Gamma_N, \end{array}\right. \tag{4.1}$$

where $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. If $\Gamma_D = \emptyset$ we further impose that $\int_\Omega f + \int_{\Gamma_N} g_N = 0$ and an unique solution exists if we require $\int_\Omega u = 0$.

In the sequel, we suppose that a is piecewise constant, namely we assume that there exists a partition $\mathcal{P}$ of Ω into a finite set of Lipschitz polygonal/polyhedral domains $\Omega_1, \cdots, \Omega_J$ such that, on each Ω_j, $a = a_j$ where a_j is a real positive constant. The variational formulation of (4.1) involves the bilinear form

$$B(u, v) = \int_\Omega a\nabla u \cdot \nabla v$$

and the Hilbert space

$$H^1_D(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_D \text{ and } \int_\Omega u = 0 \text{ if } \Gamma_D = \emptyset\}.$$

Given $f \in L^2(\Omega)$, $g_D \in H^{\frac{1}{2}}(\Gamma_D)$ and $g_N \in L^2(\Gamma_N)$ (satisfying $\int_\Omega f + \int_{\Gamma_N} g_N = 0$ if $\Gamma_D = \emptyset$), the weak formulation consists in finding $u \in w + H^1_D(\Omega)$ such that

$$B(u, v) = \int_\Omega fv + \int_{\Gamma_N} g_N v, \ \forall v \in H^1_D(\Omega), \tag{4.2}$$

where $w \in H^1(\Omega)$ is a lifting for g_D, i.e., $w = g_D$ on Γ_D. Invoking the positiveness of a, the bilinear form B is coercive on $H^1_D(\Omega)$ with respect to the norm $\left(\int_\Omega a|\nabla u|^2\right)^{1/2}$ and this coerciveness guarantees that problem (4.2) has a unique solution by the Lax-Milgram lemma.

4.2.1 Discontinuous Galerkin approximated problem

Following [4,30], we consider the following discontinuous Galerkin approximation of our continuous problem : We consider a triangulation $\mathcal{T}_h$ made of triangles T if $d = 2$ and of tetrahedra if $d = 3$ whose edges/faces are denoted by e. We assume that this triangulation is

regular, i.e., for any element T, the ratio $\dfrac{h_T}{\rho_T}$ is bounded by a constant $\sigma > 0$ independent of T and of the mesh size $h = \max_{T\in\mathcal{T}_h} h_T$, where h_T is the diameter of T and ρ_T the diameter of its largest inscribed ball. We further assume that $\mathcal{T}_h$ is conforming with the partition $\mathcal{P}$ of Ω, i.e., the function a is constant on each $T \in \mathcal{T}_h$, we then denote by a_T the value of a restricted to an element T. With each edge/face e of an element T, we associate a unit normal vector n_e, and n_T stands for the outer unit normal vector of T. $\mathcal{E}$ represents the set of edges/faces of the triangulation. In the sequel, we need to distinguish between edges/faces included into Ω, Γ_D or Γ_N, in other words, we set

$$\begin{aligned}\mathcal{E}_{int} &= \{e \in \mathcal{E} : e \subset \Omega\},\\ \mathcal{E}_D &= \{e \in \mathcal{E} : e \subset \Gamma_D\},\\ \mathcal{E}_N &= \{e \in \mathcal{E} : e \subset \Gamma_N\}.\end{aligned}$$

For shortness, we also write $\mathcal{E}_{ID} = \mathcal{E}_{int} \cup \mathcal{E}_D$.

Problem (4.2) is approximated by the (discontinuous) finite element space :

$$X_h = \left\{ v_h \in L^2(\Omega) | v_{h|T} \in \mathcal{P}_l(T), T \in \mathcal{T}_h \text{ and } \int_\Omega v_h = 0 \text{ if } \Gamma_D = \emptyset \right\},$$

where l is a fixed positive integer. The space X_h is equipped with the norm

$$\|q\|_{DG,h} := \left(\|a^{1/2}\nabla_h q\|_\Omega^2 + \gamma \sum_{e\in\mathcal{E}_{h\,ID}} h_e^{-1} \| \left[\!\left[q\right]\!\right]_e \|_e^2 \right)^{1/2},$$

where γ is a positive parameter fixed below.

For our further analysis we need to define some jumps and means through any $e \in \mathcal{E}$ of the triangulation. For $e \in \mathcal{E}$ such that $e \subset \Omega$, we denote by T^+ and T^- the two elements of $\mathcal{T}_h$ containing e. Let $q \in X_h$, we denote by $q^\pm$, the traces of q taken from $T^\pm$, respectively. Then we define the mean of q on e by

$$\{\{q\}\} = \frac{q^+ + q^-}{2}.$$

For $v \in [X_h]^d$, we denote similarly

$$\{\{v\}\} = \frac{v^+ + v^-}{2}.$$

The jump of q on e is now defined as follows :

$$\left[\!\left[q\right]\!\right]_e = q^+ n_{T^+} + q^- n_{T^-}.$$

Remark that the jump $\left[\!\left[q\right]\!\right]_e$ of q is vector-valued.

For a boundary edge/face e, i. e., $e \subset \partial\Omega$, there exists a unique element $T^+ \in \mathcal{T}_h$ such that $e \subset \partial T^+$. Therefore the mean and jump of q are defined by $\{\{q\}\} = q^+$ and $[\![q]\!]_e = q^+ n_{T^+}$.

For $q \in X_h$, we define its broken gradient $\nabla_h q$ in Ω by :

$$(\nabla_h q)_{|T} = \nabla q_{|T}, \forall T \in \mathcal{T}_h.$$

With these notations, we define the bilinear form $B_h(.,.)$ as follows :

$$\begin{aligned} B_h(u_h, v_h) &:= \sum_{T\in\mathcal{T}_h} \int_T a\nabla u_h \cdot \nabla v_h - \sum_{e\in\mathcal{E}_{h\,ID}} \int_e (\{\{a\nabla_h v_h\}\} \cdot [\![u_h]\!]_e + \{\{a\nabla_h u_h\}\} \cdot [\![v_h]\!]_e) \\ &+ \gamma \sum_{e\in\mathcal{E}_{h\,ID}} h_e^{-1} \int_e [\![u_h]\!]_e \cdot [\![v_h]\!]_e, \qquad \forall u_h, v_h \in X_h, \end{aligned}$$

where the positive parameter γ is chosen large enough to ensure coerciveness of the bilinear form B_h on X_h (see, e.g., Lemma 2.1 of [30]), namely according to the results from [50], the choice

$$\gamma > \frac{(l+1)(l+d)}{d} \max_{T\in\mathcal{T}_h} \Big(a_T \sum_{e\subset\partial T} h_e \frac{|\partial T|}{|T|} \Big) \tag{4.3}$$

yields the coerciveness of B_h.

The discontinuous Galerkin approximation of problem (4.2) reads now : Find $u_h \in X_h$, such that

$$B_h(u_h, v_h) = F(v_h), \tag{4.4}$$

where

$$F(v_h) = \int_\Omega f v_h + \sum_{e\in\mathcal{E}_D} \int_e g_D(\gamma h_e^{-1} v_h - a\nabla v_h \cdot n_T) + \int_{\Gamma_N} g_N v_h, \ \forall v_h \in X_h.$$

As our approximated scheme is a non conforming one (i.e. the solution does not belong to $H^1_D(\Omega)$), as usual we need to use an appropriate Helmholtz decomposition of the error (see Lemma 3.2 of [25] or Theorem 1 of [2] in 2D and Lemma 6.5 of [23] in 3D) :

Lemma 4.2.1 (Helmholtz decomposition of the error)**.** *We have the following error decomposition*

$$a\nabla_h(u - u_h) = a\nabla\varphi + \operatorname{curl}\chi, \tag{4.5}$$

with $\chi \in H^1(\Omega)$ *if* $d = 2$ *and* $\chi \in H^1(\Omega)^3$ *if* $d = 3$ *is such that*

$$\operatorname{curl}\chi \cdot n = 0 \ \text{on}\ \Gamma_N, \tag{4.6}$$

and $\varphi \in H^1_D(\Omega)$*. Moreover the next identity holds :*

$$\|a^{1/2}\nabla_h(u - u_h)\|^2 = \|a^{1/2}\nabla_h\varphi\|^2 + \|a^{-1/2}\operatorname{curl}\chi\|^2. \tag{4.7}$$

Proof: We consider the following problem : find $\varphi \in H^1_D(\Omega)$ solution of

$$\begin{cases} \operatorname{div} a(\nabla_h(u-u_h) - \nabla\varphi) = 0 & \text{in } \Omega, \\ \varphi = 0 & \text{on } \Gamma_D, \\ a(\nabla_h(u-u_h) - \nabla\varphi)\cdot n = 0 & \text{in } \Gamma_N. \end{cases} \tag{4.8}$$

The weak formulation of that problem (4.8) is :

$$\int_\Omega a\nabla\varphi\cdot\nabla v = \int_\Omega a\nabla_h(u-u_h)\cdot\nabla v, \quad \forall v \in H^1_D(\Omega). \tag{4.9}$$

As the vector field $a(\nabla_h(u-u_h) - \nabla\varphi)$ is divergence free in Ω, i.e.,

$$\operatorname{div} a(\nabla_h(u-u_h) - \nabla\varphi) = 0 \text{ in } \Omega,$$

by Theorem I.3.1 of [26] if $d = 2$ or Theorem I.3.4 of [26] if $d = 3$, there exists $\chi \in H^1(\Omega)$ if $d = 2$ and $\chi \in H^1(\Omega)^3$ if $d = 3$ such that

$$\operatorname{curl}\chi = a(\nabla_h(u-u_h) - \nabla\varphi).$$

This proves the identity (4.5). The boundary condition (4.6) satisfied by χ follows from the boundary condition satisfied by $a(\nabla_h(u-u_h) - \nabla\varphi)$ on Γ_N.

The identity (4.7) directly follows by using Green's formula and the boundary condition (4.6). Indeed using (4.5) we may write

$$\|a^{1/2}\nabla_h(u-u_h)\|^2 = \|a^{1/2}\nabla_h\varphi\|^2 + \|a^{-1/2}\operatorname{curl}\chi\|^2 + 2\int_\Omega \nabla\varphi\cdot\operatorname{curl}\chi.$$

In the last term, using Green's formula we have

$$\int_\Omega \nabla\varphi\cdot\operatorname{curl}\chi = -\int_\Omega \varphi\operatorname{div}\,\operatorname{curl}\chi + \int_\Gamma \varphi\operatorname{curl}\chi\cdot n\,ds = 0,$$

since the boundary term is zero by using the boundary condition $\varphi = 0$ on Γ_D and by using (4.6) on Γ_N. ∎

4.3 The a posteriori error analysis based on Raviart-Thomas finite elements

Error estimators can be constructed in many different ways as, for example, using residual type error estimators which measure locally the jump of the discrete flux [30]. A different method, based on equilibrated fluxes, consists in solving local Neumann boundary value problems [3]. Here, introducing the flux as auxiliary variable, we locally define an error estimator based on a $H(\operatorname{div})$-conforming approximation of this variable. This method avoids solving the supplementary above-mentioned local subproblems. Indeed in many

applications, the flux $j = a\nabla u$ is an important quantity, introducing this auxiliary variable, we transform the original problem (4.1) into a first order system. If $g_N = 0$, its weak formulation gives rise to the following saddle point problem : Find $(j, u) \in H_N(\mathrm{div}\,, \Omega) \times L^2(\Omega)$ such that

$$\int_\Omega a^{-1} j\ \tau + \int_\Omega \mathrm{div}\,\tau\ u = \int_{\Gamma_D} g_D \tau \cdot n,\ \forall \tau \in H_N(\mathrm{div}\,, \Omega), \tag{4.10}$$

$$\int_\Omega \mathrm{div}\, j\ w = -\int_\Omega f\ w,\ \forall w \in L^2(\Omega), \tag{4.11}$$

the natural space for the flux being

$$H_N(\mathrm{div}\,, \Omega) = \left\{ q \in [L^2(\Omega)]^2 | \mathrm{div}\, q \in L^2(\Omega) \text{ and } q \cdot n = 0 \text{ on } \Gamma_N \right\}.$$

Therefore the discrete flux approximation j_h will be searched in a $H(\mathrm{div}\,)$-conforming space V_h based on the Raviart-Thomas finite elements. This means that our error estimate of the conforming part of the error is based on the error between $a\nabla_h u_h$ and an approximating flux j_h of j that we search in the Raviart-Thomas finite element space

$$V_h = \left\{ v_h \in H(\mathrm{div}\,, \Omega) | v_{h|T} \in RT_{l-1}(T), T \in \mathcal{T}_h \right\},$$

where $RT_l(T) = [\mathcal{P}_l(T)]^d + \tilde{\mathcal{P}}_l(T) \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$ and $\tilde{\mathcal{P}}_l(T)$ stands for the space of homogeneous polynomials of degree l.

On a triangle/tetrahedron T, an element p of $RT_{l-1}(T)$ is characterized by the degrees of freedom given by

- $\displaystyle\int_e p \cdot n\ q, \quad \forall q \in \mathcal{P}_{l-1}(e),\ \forall e \subset \partial T,$
- $\displaystyle\int_T p \cdot q, \quad \forall q \in [\mathcal{P}_{l-2}(T)]^d.$

Therefore we fix the discrete flux j_h by setting

$$\int_e j_h \cdot n_T q = \int_e g_{T,e}\ q, \qquad \forall q \in \mathcal{P}_{l-1}(e),\ \forall e \subset \partial T, \tag{4.12}$$

$$\int_T j_h \cdot q = \int_T a\nabla u_h \cdot q - a_T l_{\partial T}(q), \quad \forall q \in [\mathcal{P}_{l-2}(T)]^d, \tag{4.13}$$

where for all $e \subset \partial T$, $g_{T,e}$ is defined by

$$\begin{aligned} g_{T,e} &= \left(\left\{ \left\{ a\nabla_h u_h \right\} \right\} - \gamma h_e^{-1} \left[\left[u_h \right] \right]_e \right) \cdot n_T && \text{if } e \in \mathcal{E}_{int}, \\ g_{T,e} &= a\nabla_h u_h \cdot n_T - \gamma h_e^{-1}(u_h - g_D) && \text{if } e \in \mathcal{E}_D, \\ g_{T,e} &= g_N && \text{if } e \in \mathcal{E}_N, \end{aligned}$$

and the linear form $l_{\partial T}$ is given by

$$l_{\partial T}(q) = \frac{1}{2} \sum_{e \subset \partial T \setminus \Gamma} \int_e \left[\!\left[u_h \right]\!\right]_e \cdot q + \sum_{e \subset \partial T \cap \Gamma_D} \int_e (u_h - g_D)\, q \cdot n_T.$$

Denote by Π_{l-1} the L^2-projection on $W_h = \left\{ w_h \in L^2(\Omega) | w_{h|T} \in \mathcal{P}_{l-1}(T),\ T \in \mathcal{T}_h \right\}$. Then, we have the following projection lemma.

Lemma 4.3.1. *Assume that $j_h \in V_h$ satisfies (4.12)-(4.13) on each element T of $\mathcal{T}_h$. Then, we obtain*

$$\operatorname{div} j_h = -\Pi_{l-1} f. \tag{4.14}$$

Proof: Let T be an element of the triangulation. As $j_h \in V_h$, $\operatorname{div} j_h \in W_h$ and by Green's formula, it follows that, for all $w \in \mathcal{P}_{l-1}(T)$,

$$\int_T \operatorname{div} j_h \; w = -\int_T j_h \; \nabla w + \int_{\partial T} j_h \cdot n_T \; w.$$

Now, from (4.12), (4.13), we get

$$\begin{aligned} \int_T \operatorname{div} j_h \; w &= -B_h(u_h, \tilde{w}) + \int_{\partial T \cap \Gamma_N} g_N \, w \\ &+ \int_{\partial T \cap \Gamma_D} g_D(\gamma h_e^{-1} w - a \nabla w \cdot n_T), \end{aligned}$$

where $\tilde{w}$ mean the extension of w by zero outside T. By the discontinous Galerkin formulation (4.4), we conclude that

$$\int_T \operatorname{div} j_h \; w = -\int_T f w.$$

■

Remark 4.3.2. If $l = 1$ and $d = 2$, alternative constructions of j_h are given in Lemma 6 of [2], our proposed construction has the advantage to hold for any space dimension as well as any degree l.

Remark 4.3.3. The terms $g_{T,e}$ in (4.12) actually play the role of flux functions (in the terminology of [3]). They further fulfil the so-called equilibrated equations (compare with Lemma 5 of [2])

$$\sum_{e \subset \partial T} \int_e g_{T,e} = -\int_T f,$$

due to the above proof with $w = 1$.

We introduce the conforming part of the estimator η_{CF} that only involves the difference between the discrete flux approximation j_h and $a\nabla u_h$:

$$\eta_{CF}^2 = \sum_{T\in\mathcal{T}_h} \eta_{CF,T}^2, \tag{4.15}$$

where the indicator $\eta_{CF,T}$ is defined by

$$\eta_{CF,T} = \|a^{-1/2}\,(a\nabla u_h - j_h)\,\|_T.$$

For the nonconforming part of the error, we associate with u_h, its Oswald interpolation operator, namely the unique element $w_h \in X_h \cap H^1(\Omega)$ defined in the following natural way (see Theorem 2.2 of [30]) : to each node n of the mesh corresponding to Lagrangian-type degree of freedom of $X_h \cap H^1(\Omega)$, the value of w_h is the average of the values of u_h at this node n if it belongs to $\Omega \cup \Gamma_N$ (i.e., $w_h(n) = \frac{\sum_{n\in T} |T| u_{h|T}(n)}{\sum_{n\in T}|T|}$) and the value of g_D at this node if it belongs to $\bar{\Gamma}_D$ (here we assume that $g_D \in C(\bar{\Gamma}_D)$). Then the non conforming indicator $\eta_{NC,T}$ is simply

$$\eta_{NC,T} = \|a^{1/2}\nabla(w_h - u_h)\|_T.$$

The non conforming part of the estimator is then

$$\eta_{NC}^2 = \sum_{T\in\mathcal{T}_h} \eta_{NC,T}^2. \tag{4.16}$$

Similarly we introduce the estimator corresponding to jumps of u_h :

$$\eta_J^2 = \sum_{e\in\mathcal{E}_{h\,ID}} \eta_{J,e}^2,$$

with

$$\eta_{J,e}^2 = \begin{cases} \frac{\gamma}{h_e}\| \left[\!\left[u_h\right]\!\right]_e\|_e^2 & \text{if } e \in \mathcal{E}_{int}, \\ \frac{\gamma}{h_e}\|u_h - g_D\|_e^2 & \text{if } e \in \mathcal{E}_D. \end{cases}$$

The higher order terms depending on the data f and g_N are defined as

$$\begin{aligned} osc(f)^2 &= \sum_{T\in\mathcal{T}_h} h_T^2 a_T^{-1} \|f - \Pi_{l-1} f\|_T^2, \\ osc(g_N)^2 &= \sum_{T\in\mathcal{T}_h} h_T a_T^{-1} \sum_{e\in\mathcal{E}_N : e\subset\partial T} \|g_N - \pi_{l-1} g_N\|_e^2, \end{aligned}$$

where $\pi_{l-1}g$ means the L^2-projection of g on $\{w \in L^2(\Gamma_N) : w_{|e} \in \mathcal{P}_{l-1}(e), \forall e \in \mathcal{E}_N\}$.

4.3.1 Upper bound

Theorem 4.3.4. *Assume that there exists $v_h \in X_h \cap H^1(\Omega)$ such that $g_D = v_{h|\Gamma_D}$. Then the energy norm of the error between the exact solution and its finite element approximation is bounded from above by the estimator and the higher order oscillation terms, this means that there exists $C > 0$ such that*

$$\|a^{1/2}\nabla_h(u-u_h)\| \leq (\eta_{CF}^2 + \eta_{NC}^2)^{1/2} + C\ (osc(f) + osc(g_N)), \tag{4.17}$$

and consequently

$$\|u-u_h\|_{DG,h} \leq (\eta_{CF}^2 + \eta_{NC}^2 + \eta_J^2)^{1/2} + C\ (osc(f) + osc(g_N)). \tag{4.18}$$

Proof: From the Helmholtz decomposition of the error we have

$$\|a^{1/2}\nabla_h(u-u_h)\|^2 = \|a^{1/2}\nabla\varphi\|^2 + \|a^{-1/2}\operatorname{curl}\chi\|^2. \tag{4.19}$$

We are then reduced to estimate each term of this right-hand side.

For the non conforming part, we proceed as in [2], namely by Green's formula we have

$$\begin{aligned}
\|a^{-1/2}\operatorname{curl}\chi\|^2 &= \int_\Omega \nabla_h(u-u_h)\cdot\operatorname{curl}\chi \\
&= -\int_\Omega \nabla_h u_h\cdot\operatorname{curl}\chi + \int_{\Gamma_D} g_D \operatorname{curl}\chi\cdot n \\
&= \int_\Omega \nabla_h(w_h-u_h)\cdot\operatorname{curl}\chi,
\end{aligned}$$

since $\int_\Omega \nabla w_h\cdot\operatorname{curl}\chi = \int_{\Gamma_D} g_D\operatorname{curl}\chi\cdot n$. By Cauchy-Schwarz's inequality we directly obtain

$$\|a^{-1/2}\operatorname{curl}\chi\|^2 \leq \eta_{NC}\|a^{-1/2}\operatorname{curl}\chi\|. \tag{4.20}$$

For the conforming part, we write

$$\begin{aligned}
\|a^{1/2}\nabla\varphi\|^2 &= \int_\Omega a\nabla_h(u-u_h)\cdot\nabla\varphi \\
&= \int_\Omega (a\nabla u - j_h)\cdot\nabla\varphi + \int_\Omega (j_h - a\nabla_h u_h)\cdot\nabla\varphi.
\end{aligned}$$

Applying Green's formula in the first term of this right-hand side, we obtain

$$\begin{aligned}
\|a^{1/2}\nabla\varphi\|^2 &= \int_\Omega (-\operatorname{div}(a\nabla u) + \operatorname{div} j_h)\varphi \\
&\quad + \sum_{T\in\mathcal{T}_h}\int_T a_T^{-1/2}(j_h - a\nabla u_h)a_T^{1/2}\nabla\varphi + \int_{\Gamma_N}(g_N - \pi_{l-1}g_N)\,\varphi \\
&= \int_\Omega (f-\Pi_{l-1}f)\varphi + \sum_{T\in\mathcal{T}_h}\int_T a_T^{-1/2}(j_h - a\nabla u_h)\cdot a_T^{1/2}\nabla\varphi + \int_{\Gamma_N}(g_N-\pi_{l-1}g_N)\,\varphi.
\end{aligned}$$

As $f - \Pi_{l-1} f \perp w_h$, $\forall w_h \in \mathcal{P}_{l-1}(T)$, it follows

$$\begin{aligned}
\|a^{1/2}\nabla\varphi\|^2 &\leq \sum_{T\in\mathcal{T}_h} \|f-\Pi_{l-1}f\|_T \|\varphi - \Pi_{l-1}\varphi\|_T + \sum_{T\in\mathcal{T}_h}\sum_{e\in\mathcal{E}_N : e\subset\partial T} \|g_N - \pi_{l-1}g_N\|_e \|\varphi-\pi_{l-1}\varphi\|_e \\
&+ \sum_{T\in\mathcal{T}_h} \|a^{1/2}\nabla\varphi\|_T \|a^{-1/2}(j_h - a\nabla u_h)\|_T \\
&\leq \sum_{T\in\mathcal{T}_h} \big(C\|f-\Pi_{l-1}f\|_T h_T \|\nabla\varphi\|_T + C \sum_{e\in\mathcal{E}_N : e\subset\partial T} \|g_N - \pi_{l-1}g_N\|_e h_T^{1/2} \|\nabla\varphi\|_T \\
&+ \|a^{1/2}\nabla\varphi\|_T \|a^{-1/2}(j_h - a\nabla u_h)\|_T \big).
\end{aligned}$$

This last estimate follows from standard interpolation error estimates. This finally yields

$$\|a^{1/2}\nabla\varphi\|^2 \leq (\eta_{CF} + Cosc(f) + Cosc(g_N))\|a^{1/2}\nabla\varphi\|. \tag{4.21}$$

Coming back to the identity (4.19), and using the estimates (4.20) and (4.21) we conclude by discrete Cauchy-Schwarz's inequality and again using (4.19) :

$$\begin{aligned}
\|a^{1/2}\nabla_h(u-u_h)\|^2 &\leq \eta_{NC}\|a^{-1/2}\operatorname{curl}\chi\| + (\eta_{CF} + C(osc(f) + osc(g_N)))\|a^{1/2}\nabla\varphi\| \\
&\leq (\eta_{NC}^2 + \eta_{CF}^2)^{1/2}(\|a^{-1/2}\operatorname{curl}\chi\|^2 + \|a^{1/2}\nabla\varphi\|^2)^{1/2} \\
&+ C(osc(f) + osc(g_N))\|a^{1/2}\nabla\varphi\| \\
&\leq [(\eta_{NC}^2 + \eta_{CF}^2)^{1/2} + C(osc(f) + osc(g_N))]\|a^{1/2}\nabla_h(u-u_h)\|.
\end{aligned}$$

■

4.3.2 Lower bound

Our lower bound is based on the equivalence of the L^2-norm of any element in V_h with a discrete mesh dependent norm invoking the degrees of freedom of RT_{l-1}.

Lemma 4.3.5. *Let $v_h \in RT_l(T)$ with $T \in \mathcal{T}_h$, then the following equivalence holds*

$$\begin{aligned}
\|v_h\|_T^2 &\sim h_T^{1/2} \sum_{e\subset\partial T} \sup_{q\in P_l(T):\|q\|_e=1} \left| \int_e v_h \cdot n_T q \right| \\
&+ \sup_{q\in[P_{l-1}(T)]^d:\|q\|_T=1} \left| \int_T v_h \cdot q \right|.
\end{aligned} \tag{4.22}$$

Proof: The proof is standard and simply uses scaling arguments and the so-called Piola transformation. ■

Using the above lemma, we are now able to provide a local lower bound for our estimator $\eta_{CF,T}$.

Theorem 4.3.6. *For each element $T \in \mathcal{T}_h$ the following estimate holds*

$$\eta_{CF,T} \lesssim a_T^{-1/2} \max\{1, a_T\} \max_{T' \subset \omega_T} \{a_{T'}^{1/2}\} \|u - u_h\|_{DG,\omega_T}, \tag{4.23}$$

where ω_T denotes the patch consisting of all the triangles/tetrahedra of $\mathcal{T}_h$ having a non-empty intersection with T and

$$\|v\|_{DG,\omega_T}^2 = \|a^{1/2}\nabla_h v\|_{\omega_T}^2 + \gamma \sum_{e \in \mathcal{E}_{ID}: e \subset \omega_T} h_e^{-1} \| \left[\!\left[v \right]\!\right]_e \|_e^2.$$

Proof: By its definition (4.15), we deduce, from Lemma 4.3.5, that

$$\begin{aligned}
\eta_{CF,T} &= a_T^{-1/2} \|a\nabla u_h - j_h\|_T \\
&\sim a_T^{-1/2} \big[h_T^{1/2} \sum_{e \subset \partial T} \sup_{q \in P_{l-1}(T): \|q\|_e = 1} \left| \int_e (j_h - a\nabla u_h) \cdot n_T q \right| \\
&+ \sup_{q \in [P_{l-2}(T)]^d: \|q\|_T = 1} \left| \int_T (j_h - a\nabla u_h) \cdot q \right| \big]
\end{aligned}$$

By (4.12) and (4.13), we see that

$$\begin{aligned}
\eta_{CF,T} &\lesssim a_T^{-1/2} \big[h_T^{1/2} \sum_{e \subset \partial T \setminus \Gamma} \sup_{q \in P_{l-1}(T): \|q\|_e = 1} \left| \int_e \left[\!\left[a \tfrac{\partial u_h}{\partial n_T} \right]\!\right]_e \cdot q \right| \\
&+ h_T^{1/2} \sum_{e \subset \partial T \cap \Gamma_N} \sup_{q \in P_{l-1}(T): \|q\|_e = 1} \left| \int_e (a \frac{\partial u_h}{\partial n_T} - g_N) \cdot q \right| \\
&+ h_T^{-1/2} \sum_{e \subset \partial T \setminus \Gamma} \sup_{q \in P_{l-1}(T): \|q\|_e = 1} \left| \int_e \left[\!\left[u_h \right]\!\right]_e \cdot q \right| \\
&+ h_T^{-1/2} \sum_{e \subset \partial T \cap \Gamma_D} \sup_{q \in P_{l-1}(T): \|q\|_e = 1} \left| \int_e (u_h - g_D) \cdot q \right| \\
&+ a_T \sup_{q \in [P_{l-2}(T)]^d: \|q\|_T = 1} \sum_{e \subset \partial T \setminus \Gamma} \left| \int_e \left[\!\left[u_h \right]\!\right]_e \cdot q \right| \\
&+ a_T \sup_{q \in [P_{l-2}(T)]^d: \|q\|_T = 1} \sum_{e \subset \partial T \cap \Gamma_D} \left| \int_e (u_h - g_D) \cdot q \right| \big].
\end{aligned}$$

Using Cauchy-Schwarz's inequality and the inverse estimate $\|q\|_e \lesssim h_T^{-1/2} \|q\|_T$, we arrive at the estimate

$$\begin{aligned}
\eta_{CF,T} &\lesssim a_T^{-1/2} h_T^{1/2} \sum_{e \subset T \setminus \Gamma} \| \left[\!\left[a \tfrac{\partial u_h}{\partial n_T} \right]\!\right]_e \|_e + a_T^{-1/2} h_T^{1/2} \sum_{e \subset T \cap \Gamma_N} \| a \frac{\partial u_h}{\partial n_T} - g_N \|_e \\
&+ a_T^{-1/2} h_T^{-1/2} \max\{1, a_T\} \sum_{e \subset T \setminus \Gamma} \| \left[\!\left[u_h \right]\!\right]_e \|_e + a_T^{-1/2} h_T^{-1/2} \max\{1, a_T\} \sum_{e \subset T \cap \Gamma_D} \| u_h - g_D \|_e.
\end{aligned}$$

The two first terms of this right hand side are parts of the standard residual error estimator and it is by now standard that (using appropriate bubble functions and Green's formula)

$$h_T^{1/2} \sum_{e \subset T \setminus \Gamma} \| \left[\left[a \frac{\partial u_h}{\partial n_T} \right]\right]_e \|_e + h_T^{1/2} \sum_{e \subset T \cap \Gamma_N} \| a \frac{\partial u_h}{\partial n_T} - g_N \|_e \lesssim \| a \nabla_h (u - u_h) \|_{\omega_e}.$$

The two other terms are parts of the DG-norm and are here left in the right-hand side. ■

For the non conforming part of the estimator, we make use of Theorem 2.2 of [30] to directly obtain the

Theorem 4.3.7. *Let the assumption of Theorem 4.3.4 be satisfied. For each element $T \in \mathcal{T}_h$ the following estimate holds*

$$\eta_{NC,T} \lesssim a_T^{1/2} \| u - u_h \|_{DG,\omega_T}. \tag{4.24}$$

Remark 4.3.8. From Theorems 4.3.4, 4.3.6 and 4.3.7, we see that the estimator $(\eta_{CF}^2 + \eta_{NC}^2 + \eta_J^2)^{1/2}$ is reliable for the *DG*-norm with an effectivity index (up to higher order terms) equal to 1 and is further locally efficient. Nevertheless the arguments of Theorem 3 of [2] (that are readily extended to the case $d = 3$) show that if γ is large enough, then

$$\eta_J \leq C(\gamma)(\| a^{1/2} \nabla_h (u - u_h) \| + osc(f) + osc(g_N)),$$

where $C(\gamma)$ is a positive constant depending on γ and the aspect ratio of the mesh. This means that the estimator $(\eta_{CF}^2 + \eta_{NC}^2)^{1/2}$ is reliable for the semi-norm $\| a^{1/2} \nabla_h (u - u_h) \|$ with an effectivity index (up to higher order terms) equal to 1, but it is no more locally efficient. The numerical tests of the next section confirm these facts.

4.4 Numerical tests

Our two examples consist in solving the equation (4.1) on the square $\Omega = (-1, 1)^2$ with $\Gamma_D = \Gamma$ and a discontinuous coefficient a. Namely we decompose Ω into 4 sub-domains Ω_i, $i = 1, \ldots, 4$ with $\Omega_1 = (0, 1) \times (0, 1)$, $\Omega_2 = (-1, 0) \times (0, 1)$, $\Omega_3 = (-1, 0) \times (-1, 0)$ and $\Omega_4 = (0, 1) \times (-1, 0)$ and take $a = a_i$ on Ω_i, with $a_1 = a_3$ and $a_2 = a_4 = 1$.

For the first test, for different values of a_1, we take as exact solution the smooth function $u(x, y) = (1+x)^2(1-x)^2 y^2(1-y)^2$, which clearly satisfies (4.1) with $g_D = 0$, the right-hand side f being fixed accordingly. The numerical tests are performed with $l = 1$ and 2 and the penalization parameter $\gamma = 10$ and $\gamma = 20$, respectively. To begin, we check that the numerical solution u_h converges toward the exact solution. To this end, we plot the curves $||u - u_h||_{DG,h}$ and $\| a^{1/2} \nabla_h (u - u_h) \|$ as well as the estimators $\eta_h = (\eta_{CF}^2 + \eta_{NC}^2 + \eta_J^2)^{1/2}$ and $\eta_{h,s} = (\eta_{CF}^2 + \eta_{NC}^2)^{1/2}$ as a function of DoF (see Fig. 4.1 and 4.2). A double logarithmic scale was used so that the slope of the curves yields the order of convergence and that parallel curves correspond to quantities having a constant ratio. For the different values of a_1, we see that the approximated solution converges toward the exact one with a convergence rate of l,

that the estimator η_h (resp. $\eta_{h,s}$) is close to the error $||u-u_h||_{DG,h}$ (resp. $\|a^{1/2}\nabla_h(u-u_h)\|$) and that the contribution of η_J is very small. In all cases, we find that the effectivity indices, i.e., the ratios $||u-u_h||_{DG,h}/\eta_h$ are smaller than one, as theoretically expected. Indeed if we compute these effectivity indices, we remark in Figures 4.1 and 4.2 bottom-right that they are around 0.6 for $l=1$ and 0.05 for $l=2$, in other words they remain smaller than one.

As second test, in order to illustrate the performance of our estimator η_h, we show the meshes obtained after some iterations using an iterative algorithm based on the marking procedure

$$\eta_T > 0.75 \max_{T'} \eta_{T'},$$

and a standard refinement procedure with a limitation on the minimal angle. Using polar coordinates centered at $(0,0)$, we take as exact solution (see Example 3 from [38])

$$S(x,y) = r^\alpha \phi(\theta), \tag{4.25}$$

where $\alpha \in (0,1)$ and ϕ are chosen such that S is harmonic on each sub-domain Ω_i, $i = 1,\ldots,4$ and satisfies the jump conditions :

$$[\![S]\!] = 0 \text{ and } [\![a\nabla S \cdot n]\!] = 0$$

on the interfaces (i.e. the segments $\bar{\Omega}_i \cap \bar{\Omega}_{i+1} \pmod 4$, $i = 1,\ldots,4$). We fix non-homogeneous Dirichlet boundary conditions on Γ accordingly.

It is easy to see (see for instance [22]) that α is the root of the transcendental equation

$$\tan\frac{\alpha\pi}{4} = \sqrt{a_1}.$$

This solution has a singular behavior around the point $(0,0)$ (because $\alpha < 1$). Therefore a refinement of the mesh near this point can be expected. This can be checked in Figures 4.3 and 4.4 on the meshes obtained after 20 iterations for $a_1 = 5$ and $a_1 = 100$ respectively and for which $\alpha \approx 0.53544094560$ and $\alpha \approx 0.1269020697$ (compare with the meshes from Example 3 of [38]). Note that the tests are performed with $l=1$, $\gamma = 25$ and $\gamma = 500$ respectively and with $l=2$, $\gamma = 75$ and $\gamma = 750$ respectively. As expected, we may notice a better final mesh for $l=2$ than for $l=1$.

Let us remark that the choice of the parameter γ has an influence on the performance of our algorithm. Indeed, for example 2, we compare and report in Tables 4.1 to 4.4 for $a_1 = 5$, $a_1 = 100$ and for $l=1$ and $l=2$, the CPU times needed by our algorithm to obtain the same mesh after 20 iterations. These tables show that for $l=1$ the optimal choice of γ is around 25 (resp. 500) for $a_1 = 5$ (resp. $a_1 = 100$), while for $l=2$, the optimal value is around 75 (resp. 750) for $a_1 = 5$ (resp. $a_1 = 100$). From these tables, we may notice that if we go away from the "optimal" value of γ, then the CPU time increases drastically. Note further that the obtained optimal values are mainly in accordance with (4.3).

Finally, we check that adaptive refinements are superior to uniform ones by displaying in Figure 4.5 the decrease of the DG-norm of the error as a function of the total degrees of

freedom for uniform and adaptive strategies for example 2 with $a_1 = 5$ and $a_1 = 100$ and for $l = 1$ and $l = 2$.

From these examples, we can conclude the efficiency and reliability of our proposed estimator.

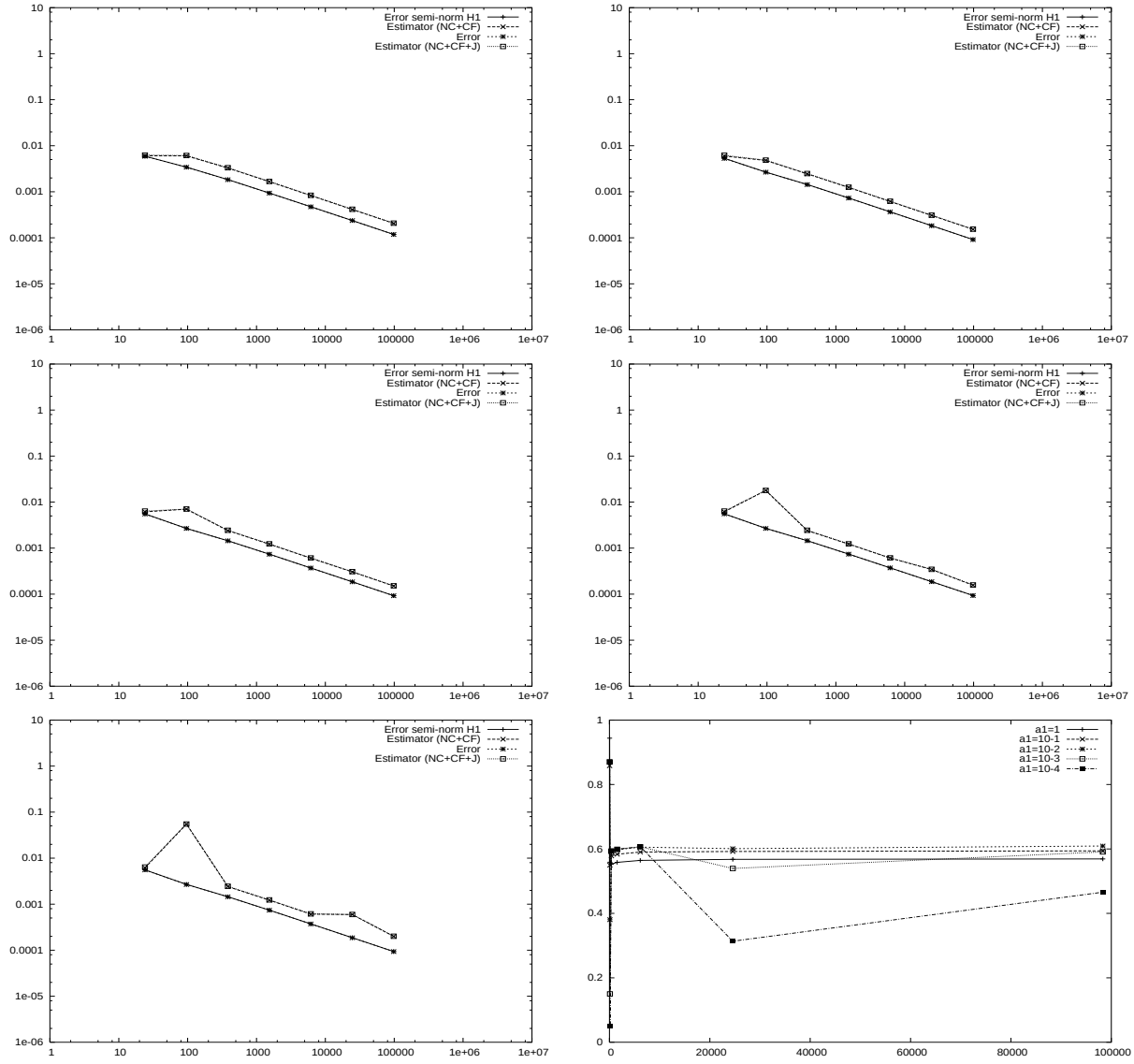


FIG. 4.1 – First example with $l = 1$: : top-left : $a_1 = a_3 = 1$, top-right $a_1 = a_3 = 0.1$; middle-left $a_1 = a_3 = 0.01$, middle-right $a_1 = a_3 = 0.001$; bottom-left $a_1 = a_3 = 0.0001$; bottom-right ratios $\|u - u_h\|_{DG,h}/\eta_h^2$.

γ	10	10.5	11	13	15	20
CPU Time ($\times 10^3$)	bad refinement	305 985	268 656	199 938	60 485	65 453

γ	25	40	50	100
CPU Time ($\times 10^3$)	54 609	61 781	65 469	76 125

TAB. 4.1 – Influence of the parameter γ on the CPU time for the coefficient $a_1 = 5$ and $l = 1$ with 20 refinements.

γ	10	100	250	500	750	1000	10 000	15 000
CPU Time ($\times 10^3$)	146 190	129 510	99 220	27 130	33 040	34 410	57 070	60 960

TAB. 4.2 – Influence of the parameter γ on the CPU time for the coefficient $a_1 = 100$ and $l = 1$ with 20 refinements.

γ	20	50	75	100	200
CPU Time ($\times 10^3$)	272 860	56 520	41 860	51 940	69 990

TAB. 4.3 – Influence of the parameter γ on the CPU time for the coefficient $a_1 = 5$ and $l = 2$ with 20 refinements.

γ	500	750	1000	1500	2000
CPU Time ($\times 10^3$)	317 350	79 080	88 540	118 840	124 830

TAB. 4.4 – Influence of the parameter γ on the CPU time for the coefficient $a_1 = 100$ and $l = 2$ with 20 refinements.

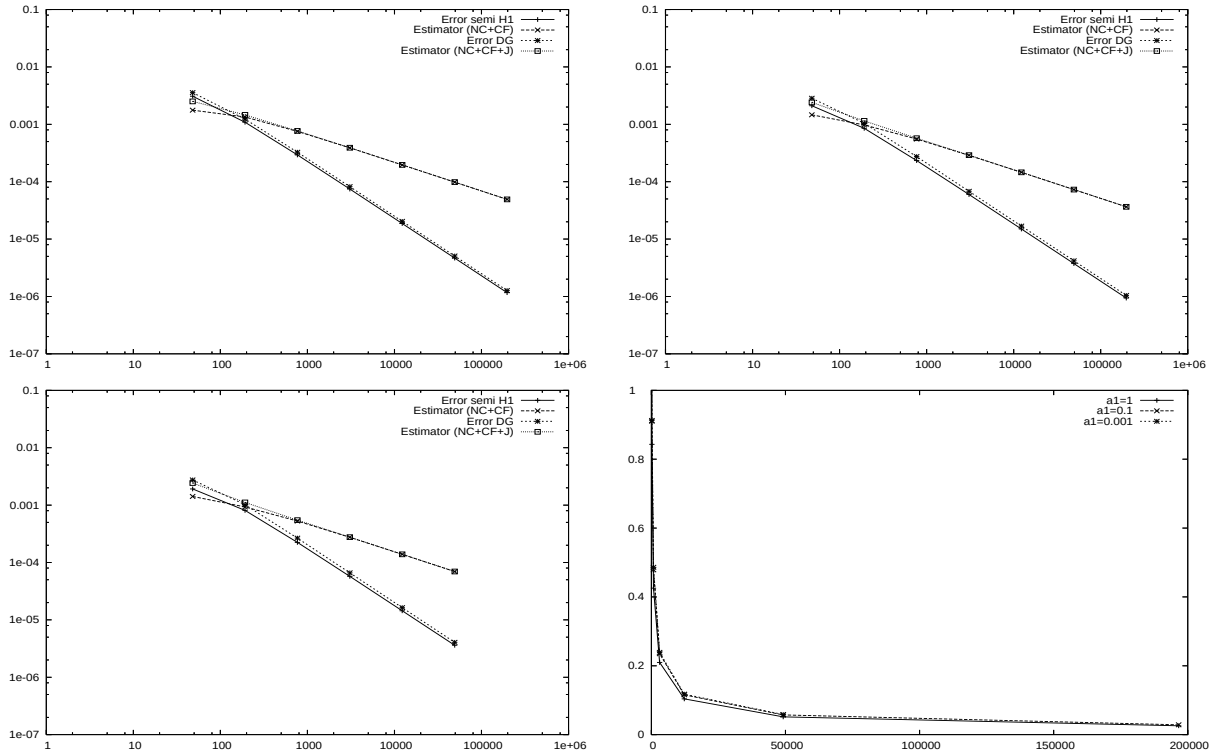


FIG. 4.2 – First example with $l = 2$: top-left : $a_1 = a_3 = 1$, top-right $a_1 = a_3 = 0.1$; bottom-left $a_1 = a_3 = 0.001$; bottom-right ratios $\|u - u_h\|_{DG,h}/\eta_h^2$.

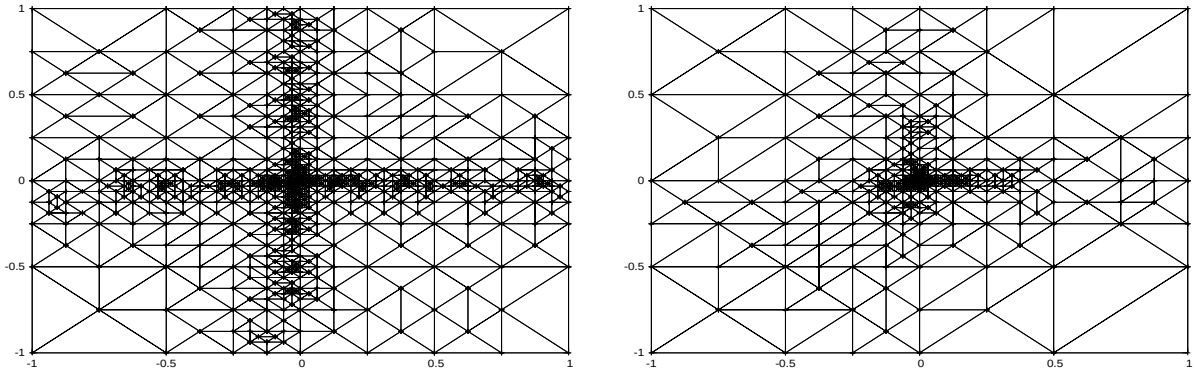


FIG. 4.3 – Adaptive mesh after 20 iterations for the second example with $l = 1$: left $a_1 = a_3 = 5$, $a_2 = a_4 = 1$; right $a_1 = a_3 = 100$, $a_2 = a_4 = 1$.

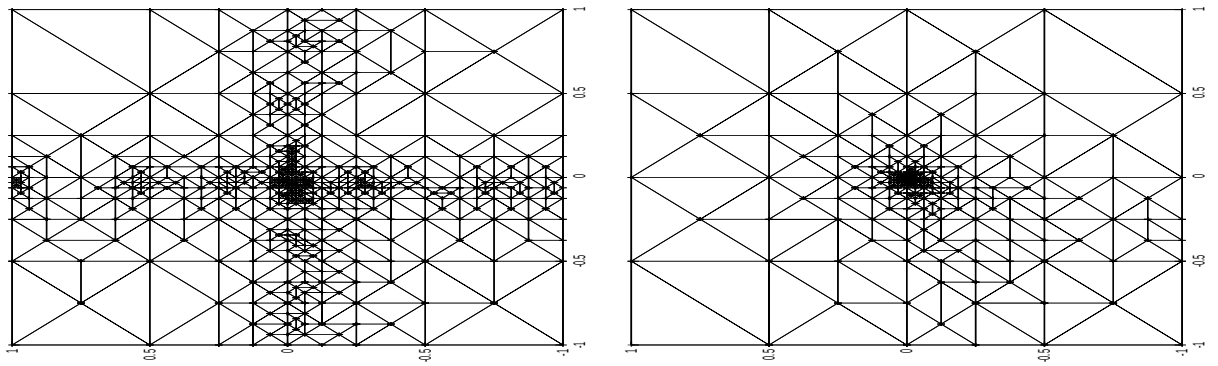


FIG. 4.4 – Adaptive mesh after 20 iterations for the second example with $l = 2$: left $a_1 = a_3 = 5$, $a_2 = a_4 = 1$; right $a_1 = a_3 = 100$, $a_2 = a_4 = 1$.

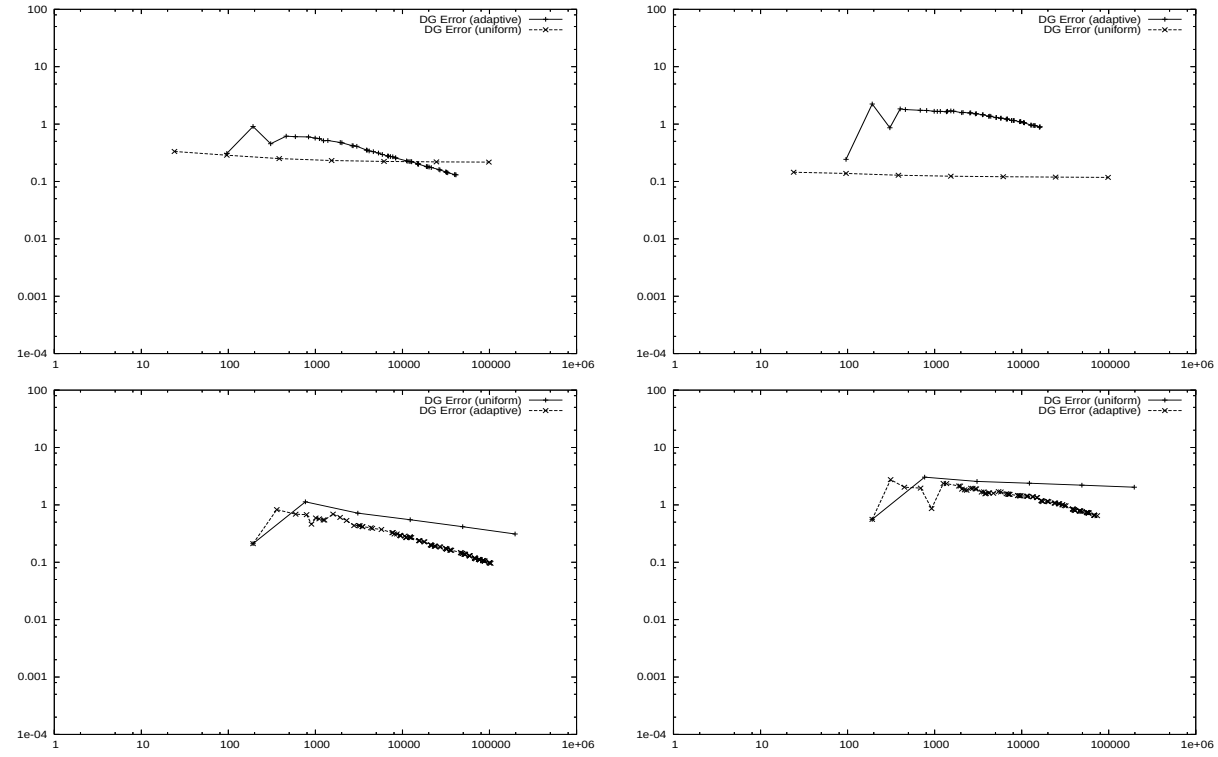


FIG. 4.5 – Comparison between uniform and adaptive refinement procedures for $l = 1, 2$. On the top-left : $a_1 = a_3 = 5, l = 1$, on the top-right : $a_1 = a_3 = 100, l = 1$; on the bottom-left : $a_1 = a_3 = 5, l = 2$, on the bottom-right : $a_1 = a_3 = 100, l = 2$.

Conclusion

Dans ce travail, nous sommes partis du système de Maxwell, et nous avons construit différents types d'estimateurs, à savoir de type résiduel et basés sur des flux équilibrés issus de la résolution de problèmes locaux. Nous avons calculés explicitement, en fonction des coefficients intervenant dans les équations, les constantes apparaissant dans les bornes inférieures et supérieures. Nous avons ainsi montré que ces estimateurs étaient robustes et l'avons validé numériquement. Nous les avons ensuite comparés, au travers de tests numériques présentant des solutions singulières, en confrontant les maillages successivement obtenus par une procédure itérative de raffinement.

Face à l'efficacité notable et la rapidité de calculs des estimateurs basés sur des flux équilibrés, nous avons souhaité étendre cette théorie aux méthodes de type Galerkin discontinues. Ainsi, dans le chapitre 4, nous avons regardé l'équation de diffusion ; la gestion d'un terme d'ordre zéro, comme dans le cas de l'équation de réaction-diffusion, présente encore actuellement quelques difficultés pour cette méthode. En effet, pour démontrer la borne supérieure, nous sommes amenés à exprimer le gradient de l'erreur à l'aide d'une décomposition de type Helmholtz. Il apparaît alors, lorsqu'on majore supérieurement l'erreur par l'estimateur, des termes d'ordre zéro faisant intervenir $u - u_h$ contre des termes issus de la décomposition du gradient de l'erreur et on ne sait pas gérer ces termes. Une difficulté supplémentaire intervient lorsque l'on passe aux équations de Maxwell, puisqu'il faut étendre la théorie de Braess et Schöberl pour construire les flux dans le cas discontinu.

La méthode des volumes finis est proche d'une méthode de type Galerkin discontinue, puisque la solution approchée est construite en tant que constante par morceaux sur les éléments. Elle représente alors un bon moyen d'appréhender la construction des flux dans le cas discontinu. Actuellement en cours de développement, pour l'équation de diffusion, la construction, à partir des volumes finis, d'un estimateur basé sur des flux équilibrés est aussi une perspective de suite à ce travail, dans le cadre des équations de Maxwell.

Bibliographie

[1] A. Agouzal and J.-M. Thomas. An extension theorem for equilibrium finite elements spaces. *Japan J. of Indust. Appl. Math.*, 13 :257–266, 1996.

[2] M. Ainsworth. A posteriori error estimation for discontinuous Galerkin finite element approximation. REPORT 2006-16, Univ. of Strathclyde, 2006.

[3] M. Ainsworth and J. Oden. *A posteriori error estimation in finite element analysis*. John Wiley and Sons, New-York, 2000.

[4] D. G. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39 :1749–1779, 2001.

[5] F. Assous, P. Ciarlet, and E. Sonnendrücker. Resolution of the maxwell equations in a domain with reentrant corners. *RAIRO Math. Mod. Num. Anal.*, 32 :359–389, 1998.

[6] I. Babuška and T. Strouboulis. *The finite element methods and its reliability*. Clarendon Press, Oxford, 2001.

[7] R. Bank and R. Smith. A posteriori error estimates based on hierarchical bases. *SIAM Journal on Numerical Analysis*, 30 :921–935, 1993.

[8] R. Beck, P. Deuflhard, R. Hiptmair, R. Hoppe, and B. Wohlmuth. Adaptive multilevel methods for edge element discretizations of Maxwell's equations. *Surveys of Math. in Industry*, 8 :271–312, 1999.

[9] R. Beck, R. Hiptmair, R. Hoppe, and B. Wohlmuth. Residual based a posteriori error estimators for eddy current computation. *RAIRO Math. Model. Numer. Anal.*, 34 :159–182, 2000.

[10] R. Becker, P. Hansbo, and M. G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Meth. Appl. Mech. Engrg.*, 192 :723–733, 2003.

[11] C. Bernardi and R. Verfürth. Adaptative finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.*, 85 :579–608, 2000.

[12] A. Bossavit. *Computational electromagnetism. Variational formulations, complementarity, edge elements*. San Diego, CA : Academic Press, 1998.

[13] D. Braess and J. Schöberl. Equilibrated residual error estimator for Maxwell's equations. Preprint RICAM 2006-19, Austrian Acad. Sciences, July 2006.

[14] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Springer, New York, 1994.

[15] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, New-York, 1991.

[16] B.Rivière and M. F. Wheeler. A posteriori error estimates for a discontinuous galerkin method applied to elliptic problems. *Comput. Math. Appl.*, 46(1) :141–163, 2003.

[17] P. Ciarlet. The finite element method for elliptic problems. *North Holland*, 1978.

[18] P. Clément. Approximation by finite element functions using local regularization. *R.A.I.R.O.*, 9(2) :77–84, 1975.

[19] S. Cochez-Dhondt and S. Nicaise. Robust a posteriori error estimation for the Maxwell equations. *Comput. Meth. Applied Mechanics Eng.*, 196 :2583–2595, 2007.

[20] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. *The development of discontinuous Galerkin methods*, volume 11 of *Lect. Notes Comput. Sci. Eng.* Springer Verlag, Berlin, 2000.

[21] M. Costabel and M. Dauge. Singularities of electromagnetic fields in polyhedral domains. *Arch. Rational Mech. Anal.*, 151 :221–276, 2000.

[22] M. Costabel, M. Dauge, and S. Nicaise. Singularities of Maxwell interface problems. *RAIRO Modél. Math. Anal. Numér.*, 33 :627–649, 1999.

[23] E. Creusé, G. Kunert, and S. Nicaise. A posteriori error estimation for the Stokes problem : Anisotropic and isotropic discretizations. *Math. Models Methods Appl. Sci.*, 14 :1297–1341, 2004.

[24] E. Creusé and S. Nicaise. Anisotropic a posteriori error estimation for the mixed discontinuous Galerkin approximation of the Stokes problem. *Numer. Meth. PDE*, 22 :449–483, 2006.

[25] E. Dari, R. Duràn, C. Padra, and V. Vampa. A posteriori error estimators for nonconforming finite element methods. *M2AN*, 30 :385–400, 1996.

[26] V. Girault and P.-A. Raviart. *Finite elements methods for Navier-Stokes equations, Theory and Algorithms*. Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986.

[27] P. Houston, I. Perugia, and D. Schötzau. Energy norm a posteriori error estimation for mixed discontinuous Galerkin approximations of the Maxwell operator. *Comput. Meth. Applied Mechanics Eng.*, 194 :499–510, 2005.

[28] P. Houston, I. Perugia, and D. Schötzau. A posteriori error estimation for discontinuous Galerkin discretization of the H(curl)-elliptic partial differential operator. *IMA J. Numer. Analysis*, 27 :122–150, 2007.

[29] F. Izsák, D. Harutyunyan, and J. van der Vegt. A posteriori implicit error estimation for the Maxwell equations. Preprint, University Twente, 2005. submitted to Math. Comp.

[30] O. A. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order problems. *SIAM J. Numer. Anal.*, 41 :2374–2399, 2003.

[31] G. Kunert. A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes. *Logos Verlag, Berlin,*, 1999. Also PhD thesis, TU Chemnitz,http ://archiv.tu-chemnitz.de/pub/1999/0012/index.html.

[32] G. Kunert. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. *Numer. Math.*, 86(3) :471–490, 2000. DOI 10.1007/s002110000170.

[33] G. Kunert. Robust a posteriori error estimation for a singularly perturbed reaction–diffusion equation on anisotropic tetrahedral meshes. *Adv. Comp. Math.*, 15(1–4) :237–259, 2001.

[34] G. Kunert. A posteriori error estimation for convection dominated problems on anisotropic meshes. *Math. Methods Appl. Sci.*, 26 :589–617, 2003.

[35] P. Ladevèze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20 :485–509, 1983.

[36] R. Lazarov, S. Repin, and S. Tomar. Functional a posteriori error estimates for discontinuous Galerkin approximations of elliptic problems. Report, RICAM, Austria, 2006.

[37] S. Lohrengel and S. Nicaise. A discontinuous Galerkin method on refined meshes for the two-dimensional time-harmonic Maxwell equations in composite materials. *J. Comput. Appl. Math.*, 2006.

[38] R. Luce and B. Wohlmuth. A local a posteriori error estimator based on equilibrated fluxes. *SIAM J. Numer. Anal.*, 42 :1394–1414, 2004.

[39] P. Monk. A posteriori error indicators for Maxwell's equations. *J. Comput. Appl. Math.*, 100 :73–190, 1998.

[40] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, 2003.

[41] J.-C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. Numer. Math., 35 :315–341, 1980.

[42] P. Neittaanmaäki and S. Repin. *Reliable methods for computer simulation : error control and a posteriori error estimates*, volume 33 of *Studies in Mathematics and its applications*. Elsevier, Amsterdam, 2004.

[43] S. Nicaise. Edge elements on anisotropic meshes and approximation of the Maxwell equations. *SIAM J. Numer. Anal.*, 39 :784–816, 2001.

[44] S. Nicaise. A posteriori error estimations of some cell centered finite volume methods. *SIAM J. Numer. Anal.*, 43 :1481–1503, 2005.

[45] S. Nicaise and E. Creusé. A posteriori error estimation for the heterogeneous Maxwell equations on isotropic and anisotropic meshes. *Calcolo*, 40 :249–271, 2003.

[46] S. Nicaise, K. Witowski, and B. Wohlmuth. An a posteriori error estimator for the Lamé equation based on equilibrated fluxes. Preprint IANS 2006/005, Universität Stuttgart, 2006.

[47] J. E. Pasciak and J. Zhao. Overlapping Schwarz methods in H(curl) on polyhedral domains. J. Numer. Math., 10 :221–234, 2002.

[48] J. Schöberl. Commuting quasi-interpolation operators for mixed finite elements. Preprint isc-01-10-math, Texas A M University, 2001.

[49] J. Schöberl. A posteriori error estimates for Maxwell equations. Preprint, J. Kepler University Linz, 2005. accepted for publication in Math. Comp.

[50] K. Shahbazi. An explicit expression for the penalty parameter of the interiori penalty method. *J. Comput. Phys.*, 205 :401–407, 2005.

[51] S. Sun and M. F. Wheeler. $L^2(H^1)$ norm a posteriori error estimation for discontinuous Galerkin approximations of reactive transport problems. *J. Sci. Comput.*, 22-23 :501–530, 2005.

[52] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley and Teubner, Chichester and Stuttgart, 1996.

[53] R. Verfürth. Robust a posteriori error estimators for singularly perturbed reaction-diffusion equations. *Numer. Math.*, 78 :479–493, 1998.

[54] R. Verfürth. Error estimates for some quasi-interpolant operators. *RAIRO Math. Model. Numer. Anal.*, 33 :695–713, 1999.

Méthodes d'éléments finis et estimations d'erreur a posteriori

Dans cette thèse, on développe des estimateurs d'erreur a posteriori, pour l'approximation par éléments finis des équations de Maxwell en régime harmonique et des équations de réaction-diffusion. Introduisant d'abord, pour le système de Maxwell, des estimateurs de type résiduel, on étudie la dépendance des constantes intervenant dans les bornes inférieures et supérieures en fonction de la variation des coefficients de l'équation, en les considérant d'abord constants puis constants par morceaux. On construit ensuite un autre type d'estimateur, basé sur des flux équilibrés et la résolution de problèmes locaux, que l'on étudie dans le cadre des équations de réaction-diffusion et du système de Maxwell. Ayant introduit plusieurs estimateurs pour l'équation de Maxwell, on en propose une étude comparative, au travers de tests numériques présentant le comportement de ces estimateurs pour des solutions particulières sur des maillages uniformes ainsi que les maillages obtenus par des procédures de raffinement de maillages adaptatifs. Enfin, dans le cadre des équations de diffusion, on étend la construction des estimateurs équilibrés aux méthodes éléments finis de type Galerkin discontinues.

Mots clefs : Eléments finis, équations de Maxwell, estimations d'erreur, estimations a posteriori, résidu, flux équilibrés, équations de réaction-diffusion, méthodes de Galerkin discontinues.

Finite element methods and a posteriori error estimations

In this thesis, we develop a posteriori error estimators, for the finite element approximation of the time-harmonic Maxwell and reaction-diffusion equations. Introducing first, for Maxwell's system, residual type estimators, we study the dependence of the constants appearing in the lower and upper bounds with respect to the variation of the coefficients of the equation we consider. Then, we construct another type of estimator, based on equilibrated fluxes and the resolution of local problems, that we study for the reaction-diffusion equations and Maxwell's system. With all the estimators built for the Maxwell equation, we propose a comparison through numerical tests involving particular solutions on uniform meshes and refinement procedures with adaptive meshes. Finally, we propose an extension, for diffusion equations, of the equilibrated estimators to the discontinuous Galerkin finite element methods.

Key words : Finite elements, Maxwell's equations, error estimations, a posteriori estimations, residual, equilibrated fluxes, reaction-diffusion equations, discontinuous Galerkin methods.

Spécialité : Mathématiques Appliquées

Laboratoire de Mathématiques et leurs Applications (LAMAV), Université de Valenciennes et du Hainaut-Cambrésis, Le Mont-Houy, 59313 Valenciennes Cedex 9