A posteriori error estimates for mixed finite element discretizations of the Neutron Diffusion equations

Patrick Ciarlet, Minh Hieu Do, François Madiot

A posteriori error estimates for mixed finite element discretizations of the Neutron Diffusion equations

Patrick Ciarlet [*1], Minh Hieu Do [†2], and François Madiot [‡2]

[1] *POEMS, ENSTA Paris, Institut Polytechnique de Paris, 828 Bd des Maréchaux, 91762 Palaiseau Cedex, France.*

[2] *CEA Saclay , DES-Service d'études des réacteurs et de mathématiques appliquées (SERMA) CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France.*

May 20, 2020

Abstract

We analyse *a posteriori* error estimates for the discretization of the neutron diffusion equations with mixed finite elements. We provide guaranteed and locally efficient estimators on a base block equation, the one-group neutron diffusion equation. We pay particular attention to AMR strategies on Cartesian meshes, since such structures are common for nuclear reactor core applications. We exhibit a robust marker strategy for this specific constraint, the *direction marker* strategy. The approach is further extended to a Domain Decomposition Method, the so-called DD+L^2 jumps method, as well as to the multigroup neutron diffusion equation.

Keywords—
Neutronics, diffusion equation, eigenvalue problem, mixed formulation, low regularity solution, *a posteriori* error estimates, mesh refinement.

Introduction

The diffusion equation can model different physical phenomena, for instance Darcy's law, Fick's law or the neutron diffusion. Among models that are used in the nuclear industry, the multigroup neutron diffusion equation plays a central role [15]. The base block is the one-group neutron diffusion equation. In [11, 10], the first author and co-authors carried out the numerical analysis of this one-group neutron diffusion equation with a source term, discretized with mixed finite elements. The numerical analysis is also performed when a domain decomposition method, called the DD+L^2-jumps method, is applied. The analysis included in particular the case of low-regularity solutions. *A priori estimates* were derived in the process. A natural question is then the *a posteriori* analysis of the method, to further optimize the cost of the numerical method. This the main topic we address in this paper.

A posteriori analysis for mixed finite elements has been extensively studied, see [7, 22, 23, 29] and references therein for the Poisson equation, [31] for the diffusion-reaction equation (one-group neutron diffusion equation), and [28] for the convection-diffusion-reaction equation. We also mention [1, 30] where *a posteriori* estimates are derived for a domain decomposition method.

[*]patrick.ciarlet@ensta-paris.fr
[†]minh-hieu.do@cea.fr
[‡]francois.madiot@cea.fr

Nuclear reactor cores often have a Cartesian geometry. Indeed, in the models, the base brick, which is called a cell, is a rectangular cuboid of $\mathbb{R}^3$. The global layout is a set of cells, that are distributed on a 3D grid, so that the global domain of the reactor core is represented by a rectangular cuboid of $\mathbb{R}^3$. Each cell can be made of fuel, absorbing or reflector material. To account for the different materials, the coefficients in the models are piecewise-polynomials (possibly piecewise-constant) with respect to the position, ie. their restriction to each cell is a polynomial [15, 19, 20]. In practice the coefficients characterizing the materials may differ from one cell to another by a factor of order 10 or more.

The outline is as follows.
In Sections 1 and 2, we introduce some notations and our model problem. Then in Section 3, we recall how it can be solved in a mixed setting. To that aim we build the well-known equivalent variational formulation, and provide the existing *a priori* numerical analysis results that allow one to compare the discrete solution to the exact one. For the discretization, we choose the well-known Raviart-Thomas-Nédélec finite element RTN_k, where $k \geq 0$ denotes the order.
In Section 4, we propose the *a posteriori* analysis of the model. We begin by the reconstruction of the solution (via post-processing), which can be devised in at least two ways: one is specific to the lowest-order, and the second one can be applied to any order. We also mention an averaging approach for the reconstruction. In Section 5, we propose some numerical experiments to compare the resulting strategies. For that, we focus on a specific discretization, based on Cartesian meshes. This kind of discretization is of particular importance for nuclear core simulations.
In Section 6, we consider the same problem, now solved with the DD+L^2-jumps method. Finally, we carry out the same analysis on a more general model for the neutron diffusion, the multigroup neutron diffusion equation. This is the subject of Section 7.

1 Notations

We choose the same notations as in [11, 10]. Throughout the paper, C is used to denote a generic positive constant which is independent of the mesh size, the mesh and the quantities/fields of interest. We also use the shorthand notation $A \lesssim B$ for the inequality $A \leq CB$, where A and B are two scalar quantities, and C is a generic constant.
Vector-valued (resp. tensor-valued) function spaces are written in boldface character (resp. blackboard characters) ; for the latter, the index *sym* indicates symmetric fields. Given an open set $\mathcal{O} \subset \mathbb{R}^{\mathtt{d}}$, $\mathtt{d} = 1, 2, 3$, we use the notation $(\cdot,\cdot)_{0,\mathcal{O}}$ (respectively $\|\cdot\|_{0,\mathcal{O}}$) for the $L^2(\mathcal{O})$ and $\mathbf{L}^2(\mathcal{O}) := (L^2(\mathcal{O}))^{\mathtt{d}}$ scalar products (resp. norms). More generally, $(\cdot,\cdot)_{s,\mathcal{O}}$ and $\|\cdot\|_{s,\mathcal{O}}$ (respectively $|\cdot|_{s,\mathcal{O}}$) denote the scalar product and norm (resp. semi-norm) of the Sobolev spaces $H^s(\mathcal{O})$ and $\mathbf{H}^s(\mathcal{O}) := (H^s(\mathcal{O}))^{\mathtt{d}}$ for $s \in \mathbb{R}$ (resp. for $s > 0$).
If moreover the boundary $\partial\mathcal{O}$ is Lipschitz, $\mathbf{n}$ denotes the unit outward normal vector field to $\partial\mathcal{O}$. Finally, it is assumed that the reader is familiar with vector-valued function spaces related to the diffusion equation, such as $\mathbf{H}(\mathrm{div}\,;\mathcal{O})$, $\mathbf{H}_0(\mathrm{div}\,;\mathcal{O})$ etc.
Specifically, we let Ω be a bounded, connected and open subset of $\mathbb{R}^d$ for $d = 2, 3$, having a Lipschitz boundary which is piecewise smooth. We split Ω into N open disjoint parts $\{\Omega_i\}_{1\leq i\leq N}$ with Lipschitz, piecewise smooth boundaries: $\overline{\Omega} = \cup_{1\leq i\leq N}\overline{\Omega_i}$ and the set $\{\Omega_i\}_{1\leq i\leq N}$ is called a partition of Ω. For a field v defined over Ω, we shall use the notations $v_i = v_{|\Omega_i}$, for $1 \leq i \leq N$.
Given a partition $\{\Omega_i\}_{1\leq i\leq N}$ of Ω, we introduce a function space with piecewise regular elements:

$$\mathcal{P}W^{1,\infty}(\Omega) \;=\; \left\{D \in L^\infty(\Omega) \,|\, D_i \in W^{1,\infty}(\Omega_i),\, 1 \leq i \leq N\right\}.$$

To measure $\psi \in \mathcal{P}W^{1,\infty}(\Omega)$, we use the natural norm $\|\psi\|_{\mathcal{P}W^{1,\infty}(\Omega)} = \max_{i=1,N} \|\psi_i\|_{W^{1,\infty}(\Omega_i)}$.

2 The model

Given a source term $S_f \in L^2(\Omega)$, we consider the following neutron diffusion equation, with vanishing Dirichlet boundary condition. In its primal form, it is written:

$$\left\{ \begin{array}{l} \text{Find } \phi \in H^1_0(\Omega) \text{ such that} \\ -\mathrm{div}\, \mathbb{D}\, \mathbf{grad}\, \phi + \Sigma_a\, \phi = S_f \text{ in } \Omega, \end{array} \right. \tag{1}$$

where ϕ, $\mathbb{D}$, and Σ_a denote respectively the neutron flux, the diffusion coefficient and the macroscopic absorption cross section. Finally, S_f denotes the fission source. When solving the neutron diffusion equation, $\mathbb{D}$ is scalar-valued. We choose to consider more generally that $\mathbb{D}$ is a (symmetric) tensor-valued coefficient. The coefficients defining Problem (1) satisfy the assumptions:

$$\left\{ \begin{array}{l} (\mathbb{D}, \Sigma_a) \in \mathbb{L}^\infty_{sym}(\Omega) \times L^\infty(\Omega)\,, \\ \exists D_*, D^* > 0,\ \forall \boldsymbol{z} \in \mathbb{R}^d,\ D_* \|\boldsymbol{z}\|^2 \leq (\mathbb{D}\boldsymbol{z}, \boldsymbol{z}) \leq D^* \|\boldsymbol{z}\|^2 \text{ a.e. in } \Omega\,, \\ \exists (\Sigma_a)_*, (\Sigma_a)^* > 0,\ 0 < (\Sigma_a)_* \leq \Sigma_a \leq (\Sigma_a)^* \text{ a.e. in } \Omega. \end{array} \right. \tag{2}$$

Classically, Problem (1) is equivalent to the following variational formulation:

$$\left\{ \begin{array}{l} \text{Find } \phi \in H^1_0(\Omega) \text{ such that} \\ \forall \psi \in H^1_0(\Omega), \quad (\mathbb{D}\, \mathbf{grad}\, \phi, \mathbf{grad}\, \psi)_{0,\Omega} + (\Sigma_a \phi, \psi)_{0,\Omega} = (S_f, \psi)_{0,\Omega}. \end{array} \right. \tag{3}$$

Under the assumptions (2) on the coefficients, the primal problem (1) is well-posed, in the sense that for all $S_f \in L^2(\Omega)$, there exists one and only one solution $\phi \in H^1_0(\Omega)$ that solves (1), with the bound $\|\phi\|_{1,\Omega} \lesssim \|S_f\|_{0,\Omega}$. Provided that the coefficient $\mathbb{D}$ is piecewise smooth, the solution is smoother (see eg. Proposition 1 in [11]). Instead of imposing a Dirichlet boundary condition on $\partial\Omega$, one can consider a Neumann or Fourier boundary condition $\mu_F \phi + (\mathbb{D}\, \mathbf{grad}\, \phi) \cdot \mathbf{n} = 0$, with $\mu_F \geq 0$. Results are similar.

3 Single domain variational formulation and discretization

Let us introduce the function space:

$$\mathbf{X} \;=\; \left\{ \xi := (\mathbf{q}, \psi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega) \right\}\,,\ \|\xi\|_{\mathbf{X}} := \left(\|\mathbf{q}\|^2_{\mathbf{H}(\mathrm{div}\,,\Omega)} \,+\, \|\psi\|^2_{0,\Omega} \right)^{1/2}\,.$$

From now on, we use the notations: $\zeta = (\mathbf{p}, \phi)$ and $\xi = (\mathbf{q}, \psi)$.

3.1 Mixed variational formulation

The solution ϕ to (1) belongs to $H^1(\Omega)$, so if one lets $\mathbf{p} := -\mathbb{D}\, \mathbf{grad}\, \phi \in \mathbf{L}^2(\Omega)$, the neutron diffusion problem may be written as:

$$\left\{ \begin{array}{l} \text{Find } (\mathbf{p}, \phi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times H^1_0(\Omega) \text{ such that} \\ -\mathbb{D}^{-1}\, \mathbf{p} \,-\, \mathbf{grad}\, \phi = 0 \text{ in } \Omega, \\ \mathrm{div}\, \mathbf{p} \,+\, \Sigma_a \phi = S_f \text{ in } \Omega. \end{array} \right. \tag{4}$$

Solving the mixed problem (4) is equivalent to solving (1).

Proposition 1. *Let $\mathbb{D}, \Sigma_a$ satisfy* (2)*. The solution $(\mathbf{p}, \phi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times H^1_0(\Omega)$ to* (4) *is such that ϕ is a solution to* (1) *with the same data. Conversely, the solution $\phi \in H^1_0(\Omega)$ to* (1) *is such that $(-\mathbb{D}\, \mathbf{grad}\, \phi, \phi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times H^1_0(\Omega)$ is a solution to* (4) *with the same data.*

To obtain the variational formulation for the mixed problem (4), let $\mathbf{q} \in \mathbf{H}(\mathrm{div}\,, \Omega)$ and $\psi \in L^2(\Omega)$, multiply the first equation of (4) by $\mathbf{q}$, the second equation of (4) by $\psi \in L^2(\Omega)$, and integrate over Ω. Adding up the contributions, one finds that:

$$-(\mathbb{D}^{-1}\,\mathbf{p}, \mathbf{q})_{0,\Omega} - (\mathbf{grad}\,\phi, \mathbf{q})_{0,\Omega} + (\psi, \mathrm{div}\,\mathbf{p})_{0,\Omega} + (\Sigma_a \phi, \psi)_{0,\Omega} = (S_f, \psi)_{0,\Omega}. \tag{5}$$

One may integrate by parts the second term in the left-hand side, which yields: $-(\mathbf{grad}\,\phi, \mathbf{q})_{0,\Omega} = (\phi, \mathrm{div}\,\mathbf{q})_{0,\Omega}$. We conclude that the solution to (4) also solves:

$$\left\{ \begin{array}{l} \text{Find } (\mathbf{p}, \phi) \in \mathbf{X} \text{ such that} \\ \forall (\mathbf{q}, \psi) \in \mathbf{X}, \quad -(\mathbb{D}^{-1}\,\mathbf{p}, \mathbf{q})_{0,\Omega} + (\phi, \mathrm{div}\,\mathbf{q})_{0,\Omega} + (\psi, \mathrm{div}\,\mathbf{p})_{0,\Omega} + (\Sigma_a\,\phi, \psi)_{0,\Omega} = (S_f, \psi)_{0,\Omega}. \end{array} \right. \tag{6}$$

Because $\mathbb{D}$ is a symmetric tensor field, the form

$$c\ :\ ((\mathbf{p}, \phi), (\mathbf{q}, \psi)) \mapsto -(\mathbb{D}^{-1}\,\mathbf{p}, \mathbf{q})_{0,\Omega} + (\phi, \mathrm{div}\,\mathbf{q})_{0,\Omega} + (\psi, \mathrm{div}\,\mathbf{p})_{0,\Omega} + (\Sigma_a\,\phi, \psi)_{0,\Omega} \tag{7}$$

is continuous, bilinear and symmetric on $\mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega)$. For further use, we introduce:

$$a : \left\{ \begin{array}{rcl} \mathbf{H}(\mathrm{div}\,, \Omega) \times \mathbf{H}(\mathrm{div}\,, \Omega) & \to & \mathbb{R} \\ (\mathbf{p}, \mathbf{q}) & \mapsto & -(\mathbb{D}^{-1}\,\mathbf{p}, \mathbf{q})_{0,\Omega} \end{array} \right. ; \tag{8}$$

$$b : \left\{ \begin{array}{rcl} \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega) & \to & \mathbb{R} \\ (\mathbf{q}, \psi) & \mapsto & (\psi, \mathrm{div}\,\mathbf{q})_{0,\Omega} \end{array} \right. ; \tag{9}$$

$$t : \left\{ \begin{array}{rcl} L^2(\Omega) \times L^2(\Omega) & \to & \mathbb{R} \\ (\phi, \psi) & \mapsto & (\Sigma_a \phi, \psi)_{0,\Omega} \end{array} \right. ; \tag{10}$$

and the continuous linear form

$$f : \left\{ \begin{array}{rcl} \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega) & \to & \mathbb{R} \\ (\mathbf{q}, \psi) & \mapsto & (S_f, \psi)_{0,\Omega} \end{array} \right. . \tag{11}$$

We may rewrite the variational formulation (6) as:

$$\left\{ \begin{array}{l} \text{Find } (\mathbf{p}, \phi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega) \text{ such that} \\ \forall (\mathbf{q}, \psi) \in \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega), \quad c((\mathbf{p}, \phi), (\mathbf{q}, \psi)) = f((\mathbf{q}, \psi)). \end{array} \right. \tag{12}$$

Proposition 2. *The solution $\zeta = (\mathbf{p}, \phi)$ to* (12) *satisfies* (4)*. Hence, problems* (12) *and* (4) *are equivalent.*

One may prove that the mixed formulation (12) is well-posed, see Theorem 4.4 in [10]. As a matter of fact, the result is obtained by proving an inf-sup condition in $\mathbf{X} = \mathbf{H}(\mathrm{div}\,, \Omega) \times L^2(\Omega)$, which we recall here.

Theorem 1. *Let $\mathbb{D}$ and Σ_a satisfy (2). Then, the bilinear symmetric form c fulfills an inf-sup condition:*

$$\exists \eta > 0, \quad \inf_{\zeta \in \mathbf{X}} \sup_{\xi \in \mathbf{X}} \frac{c(\zeta, \xi)}{\|\zeta\|_{\mathbf{X}}\, \|\xi\|_{\mathbf{X}}} \geq \eta. \tag{13}$$

3.2 Discretization and *a priori* error analysis

We study conforming discretizations of (12). Let $(\mathcal{T}_h)_h$ be a family of meshes, made for instance of simplices, or of rectangles ($d = 2$), resp. cuboids ($d = 3$), indexed by a parameter h equal to the largest diameter of elements of a given mesh. We introduce discrete, finite-dimensional, spaces indexed by h as follows:

$$\mathbf{Q}_h \subset \mathbf{H}(\mathrm{div}\,, \Omega), \text{ and } L_h \subset L^2(\Omega).$$

The conforming discretization of the variational formulation (12) is then:

$$\begin{cases} \text{Find } (\mathbf{p}_h, \phi_h) \in \mathbf{Q}_h \times L_h \text{ such that} \\ \forall (\mathbf{q}_h, \psi_h) \in \mathbf{Q}_h \times L_h, \quad c((\mathbf{p}_h, \phi_h), (\mathbf{q}_h, \psi_h)) = f((\mathbf{q}_h, \psi_h)). \end{cases} \tag{14}$$

Following Definition 2.14 in [16], we assume that $(\mathbf{Q}_h)_h$, resp. $(L_h)_h$ have the *approximability property* in the sense that

$$\begin{aligned} &\forall \mathbf{q} \in \mathbf{H}(\mathrm{div}\,, \Omega),\ \lim_{h\to 0} \left(\inf_{\mathbf{q}_h \in \mathbf{Q}_h} \|\mathbf{q} - \mathbf{q}_h\|_{\mathbf{H}(\mathrm{div}\,,\Omega)} \right) = 0, \\ &\forall \psi \in L^2(\Omega),\ \lim_{h\to 0} \left(\inf_{\psi_h \in L_h} \|\psi - \psi_h\|_{0,\Omega} \right) = 0, \end{aligned} \tag{15}$$

We also impose that the space L_h^0 of piecewise constant fields on the mesh is included in L_h, and that $\mathrm{div}\, \mathbf{Q}_h \subset L_h$. We finally define:

$$\mathbf{X}_h = \{\, \xi_h := (\mathbf{q}_h, \psi_h) \in \mathbf{Q}_h \times L_h \} \text{ , endowed with } \|\cdot\|_{\mathbf{X}}\,.$$

Remark 1. *At some point, the discrete spaces are considered locally, i.e.. restricted to one element of the mesh. So, one introduces the local spaces $\mathbf{Q}_h(K)$, $L_h(K)$, $\mathbf{X}_h(K)$ for every $K \in \mathcal{T}_h$.*

Provided the above conditions are fulfilled, one may derive a uniform discrete inf-sup condition under the same assumptions as in theorem 1 (cf. Theorem 4.5 in [10]).

Theorem 2. *Let $\mathbb{D} \in \mathcal{P}\mathbb{W}^{1,\infty}(\Omega)$, resp. $\Sigma_a \in \mathcal{P}W^{1,\infty}(\Omega)$, satisfy (2). Assume that $(\mathbf{Q}_h)_h$, $(L_h)_h$ fulfill (15), $L_h^0 \subset L_h$ and $\mathrm{div}\, \mathbf{Q}_h \subset L_h$ for all h. Then the bilinear form c fulfills a uniform discrete inf-sup condition in $\mathbf{X}_h$.*

$$\exists\, \eta' > 0, \quad \forall h, \quad \inf_{\zeta_h \in \mathbf{X}_h} \sup_{\xi_h \in \mathbf{X}_h} \frac{c(\zeta_h, \xi_h)}{\|\zeta_h\|_{\mathbf{X}}\, \|\xi_h\|_{\mathbf{X}}} \geq \eta'. \tag{16}$$

The classical *a priori* error analysis follows. Let $\zeta_h = (\mathbf{p}_h, \phi_h)$ be the solution to (14).

Corollary 1. *Let $\mathbb{D} \in \mathcal{P}\mathbb{W}^{1,\infty}(\Omega)$, resp. $\Sigma_a \in \mathcal{P}W^{1,\infty}(\Omega)$, satisfy (2). Assume that $(\mathbf{Q}_h)_h$, $(L_h)_h$ fulfill (15), $L_h^0 \subset L_h$ and $\mathrm{div}\, \mathbf{Q}_h \subset L_h$ for all h. Then there holds:*

$$\exists\, C > 0, \quad \forall h, \quad \|\zeta - \zeta_h\|_{\mathbf{X}_h} \leq C \inf_{\xi_h \in \mathbf{X}_h} \|\zeta - \xi_h\|_{\mathbf{X}_h}. \tag{17}$$

In this paper, we focus on the Raviart-Thomas-Nédélec (RTN) Finite Element [26, 24]. For *simplicial meshes*, that is meshes made of simplices, the finite element spaces RTN_k can be described as follows, where $k \geq 0$ is the order of the discretization for the scalar fields of L_h, see eg. [5].
The boundary of a simplex $K \in \mathcal{T}_h$ is made of the union of $(d-1)$-simplices, called *facets* from now on, and denoted by $(F_e^K)_{1\leq e\leq d+1}$. We let $\mathbb{P}_k(K)$ be the space of polynomials of maximal degree k on K, resp. $\mathbb{P}_k(F_e^K)$ the space of polynomials of maximal degree k on F_e^K. The definition is

$$\begin{aligned} \mathrm{RTN}_k(K) = \{\mathbf{q} \in \mathbf{L}^2(K) \,|\, &\exists \mathbf{a} \in (\mathbb{P}_k(T))^d,\ \exists b \in \mathbb{P}_k(T),\ \forall \mathbf{x} \in K,\ \mathbf{q}(\mathbf{x}) = \mathbf{a} + b\mathbf{x} \\ &\text{and } \forall e \in \{1, \cdots, d+1\},\ (\mathbf{q}\cdot\mathbf{n})_{|F_e^K} \in \mathbb{P}_k(F_e^K)\}. \end{aligned}$$

The definitions of the finite element spaces RTN_k are then

$$\mathbf{Q}_h = \{\mathbf{q}_h \in \mathbf{H}(\mathrm{div}\,, \Omega) \,|\, \forall K \in \mathcal{T}_h,\ \mathbf{q}_{h|K} \in \mathrm{RTN}_k(K)\},\ L_h = \{\psi_h \in L^2(\Omega) \,|\, \forall K \in \mathcal{T}_h,\ \psi_{h|K} \in \mathbb{P}_k(K)\}.$$

Remark 2. *For rectangular or Cartesian meshes, a description of the Raviart-Thomas-Nédélec (RTN) finite element spaces can be found for instance in Section 4.2 of [20]. We consider those meshes explicitly for the numerical examples, see Section 5.*

4 *A posteriori* studies for a mixed Finite element discretization

To develop the study of *a posteriori* estimates, we use the so-called reconstruction of the discrete solution. To that aim, we consider that

$$V := H_0^1(\Omega), \text{the original space of solutions, see (1),}$$

is the *default* space of scalar reconstructed fields. We also introduce the broken space $\mathbf{H}(\mathrm{div}\,;\mathcal{T}_h) = \{\psi \in L^2(\Omega) \mid \psi \in \mathbf{H}(\mathrm{div}\,;K), \forall K \in \mathcal{T}_h\}$.

A first approach has been suggested in [18, Chapter 8]. The reconstruction $\tilde{\zeta}_h = (\tilde{\mathbf{p}}_h, \tilde{\phi}_h)$ is defined as

$$\begin{aligned}&\tilde{\mathbf{p}}_h = \mathbf{p}_h \in \mathbf{Q}_h \subset \mathbf{H}(\mathrm{div}\,;\Omega),\\ &\tilde{\phi}_h \in V.\end{aligned}$$

Remark 3. *For other boundary conditions, i.e.. for a Neumann of Fourier boundary condition, the* default *space V of scalar reconstructed fields would be equal to $H^1(\Omega)$.*

In Section 4.1, we recall some reconstruction approaches for RTN finite element spaces. Section 4.2 is devoted to the derivation of *a posteriori* estimates.

4.1 Reconstruction of the discrete solution

In this section, we investigate some approaches to devise a reconstruction of the discrete solution $(\mathbf{p}_h, \phi_h) \in \mathbf{X}_h$, here obtained with the RTN$_k$ finite element discretization, for $k \geq 0$.
For illustrative purposes, we consider simplicial meshes. Let us introduce some further notations, given such a mesh $\mathcal{T}_h$. The set of facets of $\mathcal{T}_h$ is denoted $\mathcal{F}_h$, and it is split as $\mathcal{F}_h = \mathcal{F}_h^i \cup \mathcal{F}_h^e$, with $\mathcal{F}_h^e$ (resp. $\mathcal{F}_h^i$) being the set of boundary facets (resp. interior facets). We denote by $\mathbb{P}_k(\mathcal{T}_h)$ the space of piecewise polynomials of maximal degree k on each simplex $K \in \mathcal{T}_h$. We let $\mathcal{V}_h^k$ be the set of interpolation points (or nodes) where the degrees of freedom of the V-conforming Lagrange Finite Element space of order k are defined. And, for a node $a \in \mathcal{V}_h^k$, we denote by $\mathcal{T}_a$ the set of simplices K such that $a \in K$.
We recall the definition of the (original) Oswald interpolation operator [25] $\mathcal{I}_{Os} : \mathbb{P}_k(\mathcal{T}_h) \to \mathbb{P}_k(\mathcal{T}_h) \cap V$ such that

$$\forall \phi_h \in \mathbb{P}_k(\mathcal{T}_h),\ \forall a \in \mathcal{V}_h^k, \quad \mathcal{I}_{Os}(\phi_h)(a) = \frac{1}{|\mathcal{T}_a|} \sum_{K \in \mathcal{T}_a} \phi_{h|K}(a).$$

A second, modified Oswald operator is defined in [28] as follows. Let

$$W_0(\mathcal{T}_h) = \left\{ \psi_h \in L^2(\mathcal{T}_h) \mid \forall K \in \mathcal{T}_h, \psi_{h|K} \in H^1(K); \forall F \in \mathcal{F}_h^i, \int_F [\psi_h] = 0; \forall F \in \mathcal{F}_h^e, \int_F \psi_h = 0 \right\},$$

where $[\psi_h]|_F = \psi_h|_{K_1}\mathbf{n}_{K_1} + \psi_h|_{K_2}\mathbf{n}_{K_2}$ denotes the jump of ψ_h on the facet $F \in \mathcal{F}_h^i$ shared by elements K_1 and K_2 and $\mathbf{n}_{K_{1,2}}$ is the unit outer normal of the mesh element $K_{1,2} \in \mathcal{T}_h$. Then, the modified Oswald operator[1] $\mathcal{I}_{MO} : \mathbb{P}_2(\mathcal{T}_h) \cap W_0(\mathcal{T}_h) \to \mathbb{P}_d(\mathcal{T}_h) \cap V$ is such that

$$\begin{aligned}&\forall \phi_h \in \mathbb{P}_2(\mathcal{T}_h) \cap W_0(\mathcal{T}_h),\ \forall a \in \mathcal{V}_h^d,\\ &\qquad \mathcal{I}_{MO}(\phi_h)(a) = \begin{cases} \mathcal{I}_{Os}(\phi_h)(a) & \text{if } a \text{ is not located at a barycenter of a facet} \\ m(\phi_h, a) & \text{else} \end{cases}.\end{aligned}$$

Above, the values $(m(\phi_h, a_F))_{F \in \mathcal{F}_h}$ at the barycenters of the facets are then defined so that the means of $\mathcal{I}_{MO}(\phi_h)$ on every facet is equal to the means of ϕ_h.

Remark 4. *Observe that the results presented in this section can be extended to the case of rectangular or cuboid meshes [29].*

[1] Recall that $d = 2$ or $d = 3$.

4.1.1 Averaging operator

We introduce the averaging operator of the neutron flux $\mathcal{I}_{av} : \mathbb{P}_k(\mathcal{T}_h) \to \mathbb{P}_{k+1}(\mathcal{T}_h) \cap V$ such that

$$\forall \phi_h \in \mathbb{P}_k(\mathcal{T}_h),\ \forall a \in \mathcal{V}_h^{k+1}, \quad \mathcal{I}_{av}(\phi_h)(a) = \frac{1}{|\mathcal{T}_a|} \sum_{K \in \mathcal{T}_a} \phi_{h|K}(a).$$

The *average reconstruction* is $\tilde{\zeta}_h = (\mathbf{p}_h, \mathcal{I}_{av}(\phi_h))$.

4.1.2 Post-processing approaches

In order to recover the relation $\mathbf{p} = -\mathbb{D}\mathbf{grad}\,\phi$ at the discrete level, some post-processing techniques have been introduced for mixed finite element method [28, 29]. The first one is specific to a discretization with the RTN_0 finite element, whereas the second one can be applied to any discretization with a RTN_k finite element, i.e.. k can be any integer, possibly equal to 0.

RTN_0 post-processing A post-processing is introduced in [28] for the RTN_0 finite element, that is with $k = 0$. Briefly, the author proposes one post-processed scalar variable $\mathcal{I}_{pp}(\mathbf{p}_h, \phi_h) = \widehat{\phi}_h \in \mathbb{P}_2(\mathcal{T}_h)$, which is such that

$$\forall K \in \mathcal{T}_h, \quad -\mathbb{D}_K(\mathbf{grad}\,\widehat{\phi}_h)_{|K} = \mathbf{p}_{h|K},\ \frac{(\widehat{\phi}_h, 1)_{0,K}}{|K|} = \phi_{h|K}. \tag{18}$$

Problems (18) are local and independent on each element $K \in \mathcal{T}_h$. We define the RTN_0 post-processing by $\mathcal{I}_{Os} \circ \mathcal{I}_{pp}$. The *reconstruction* associated to the RTN_0 post-processing is $\tilde{\zeta}_h = (\mathbf{p}_h, \mathcal{I}_{Os} \circ \mathcal{I}_{pp}(\mathbf{p}_h, \phi_h))$.

RTN post-processing In the general case of the RTN_k finite element, for $k \geq 0$, there exists no solution to Problem (18). We present here the approach proposed in [2] for the general case. It is shown there that the solution to (14), $\zeta_h = (\mathbf{p}_h, \phi_h) \in \mathbf{X}_h$, is also equal to the first argument of the solution of a hybrid formulation, where the constraint of continuity of the normal trace of the current $\mathbf{p}_h \in \mathbf{Q}_h$ is relaxed. Let

$$\Lambda_h = \left\{ \lambda_h \in L^2(\mathcal{F}_h^i) \mid \exists \mathbf{q}_h \in \mathbf{Q}_h,\ \lambda_{h|F} = \mathbf{q}_h \cdot \mathbf{n}_{|F},\ \forall F \in \mathcal{F}_h^i \right\},$$

be the space of the Lagrange multipliers and let $\tilde{\mathbf{X}}_h := \Pi_{K \in \mathcal{T}_h} \mathbf{X}_h(K)$ be the unconstrained approximation space with the RTN_k local finite element spaces. By definition, $\mathbf{X}_h$ is a strict subset of $\tilde{\mathbf{X}}_h$.
The hybrid formulation is:

$$\left\{ \begin{array}{l} \text{Find } (\zeta_h, \lambda_h) \in \tilde{\mathbf{X}}_h \times \Lambda_h \text{ such that} \\ \forall (\xi_h, \mu_h) \in \tilde{\mathbf{X}}_h \times \Lambda_h, \quad c(\zeta_h, \xi_h) - \displaystyle\sum_{F \in \mathcal{F}_h^i} \int_F \lambda_h [\mathbf{q}_h \cdot \mathbf{n}] + \sum_{F \in \mathcal{F}_h^i} \int_F \mu_h [\mathbf{p}_h \cdot \mathbf{n}] = f(\xi_h). \end{array} \right. \tag{19}$$

Let $\Pi_{M_h} : \tilde{\mathbf{X}}_h \times \Lambda_h \to M_h$ be the projection onto an appropriate space[2] such that, given $(\zeta_h, \lambda_h) \in \tilde{\mathbf{X}}_h \times \Lambda_h$, its projection $\widehat{\phi}_h = \Pi_{M_h}(\zeta_h, \lambda_h)$ is governed by

$$\begin{aligned} &\forall \psi_h \in L_h, \quad (\widehat{\phi}_h, \psi_h)_{0,\Omega} = (\phi_h, \psi_h)_{0,\Omega}, \\ &\forall \mu_h \in \Lambda_h, \quad \sum_{F \in \mathcal{F}_h^i} \int_F \widehat{\phi}_h \mu_h = \sum_{F \in \mathcal{F}_h^i} \int_F \lambda_h \mu_h. \end{aligned}$$

Then, the *reconstruction* associated to the RTN post-processing is $\tilde{\zeta}_h = (\mathbf{p}_h, \Pi_{M_h}(\zeta_h, \lambda_h))$.

Finally, we refer to [17] for an application of this technique in the field of neutronics. The RTN post-processing is defined here by $\mathcal{I}^2_{\mathrm{RTN}} : \tilde{\mathbf{X}}_h \times \Lambda_h \to \mathbb{P}_{k+2}(\mathcal{T}_h) \cap V$ such that

$$\forall (\zeta_h, \lambda_h) \in \tilde{\mathbf{X}}_h \times \Lambda_h,\ \forall a \in \mathcal{V}_h^{k+2}, \quad \mathcal{I}^2_{\mathrm{RTN}}(\zeta_h, \lambda_h)(a) = \frac{1}{|\mathcal{T}_a|} \sum_{K \in \mathcal{T}_a} (\Pi_{M_h}(\zeta_h, \lambda_h))_{|K}(a),$$

The *reconstruction* associated to the RTN post-processing is $\tilde{\zeta}_h = (\mathbf{p}_h, \mathcal{I}^2_{\mathrm{RTN}}(\zeta_h, \lambda_h))$.

4.2 *A posteriori* error estimates

In this section, we detail the derivation of *a posteriori* estimates.

We define

$$\begin{aligned} d_S(\zeta, \xi) &= -a(\mathbf{p}, \mathbf{q}) + t(\phi, \psi) \\ d_A(\zeta, \xi) &= b(\mathbf{p}, \psi) - b(\mathbf{q}, \phi) \\ d(\zeta, \xi) &= d_S(\zeta, \xi) + d_A(\zeta, \xi) = c(\zeta, (-\mathbf{q}, \psi)). \end{aligned}$$

The definition is extended to piecewise smooth fields on $\mathcal{T}_h$ by replacing $\displaystyle\int_\Omega$ by $\displaystyle\sum_{K \in \mathcal{T}_h} \int_K$.

We define the following norm on $\mathbf{X}$, for all $\zeta \in \mathbf{X}$,

$$\begin{aligned} \|\zeta\|_S^2 &= d_S(\zeta, \zeta) + \sum_{K \in \mathcal{T}_h} \|\Sigma_a^{-1/2} \mathrm{div}\, \mathbf{p}\|_{0,K}^2 \\ &= (\mathbb{D}^{-1}\, \mathbf{p}, \mathbf{p})_{0,\Omega} + (\Sigma_a\, \phi, \phi)_{0,\Omega} + \sum_{K \in \mathcal{T}_h} \|\Sigma_a^{-1/2} \mathrm{div}\, \mathbf{p}\|_{0,K}^2. \end{aligned}$$

For $K \in \mathcal{T}_h$, we introduce $N(K) = \{K' \in \mathcal{T}_h, \dim_H(\partial K' \cap \partial K) = d - 1\}$, where $\dim_H$ is the Hausdorff dimension, and $\mathbf{X}_K = \{\zeta = (\mathbf{p}, \phi) \in \mathbf{X} \mid \mathrm{Supp}(\phi) \subset K, \mathrm{Supp}(\mathbf{p}) \subset N(K)\}$. Then one can define the following $\mathbf{X}_K$-local norm, for all $\zeta \in \mathbf{X}$,

$$|\zeta|_{+,K} = \sup_{\xi \in \mathbf{X}_K, \|\xi\|_S \le 1} d(\zeta, \xi). \tag{20}$$

[2]The space M_h is defined here as $M_h = \Pi_{K \in \mathcal{T}_h} M_h(K)$, with for all $K \in \mathcal{T}_h$,

$$M_h(K) = \begin{cases} \{\psi_h \in \mathbb{P}_{k+3}(K) : \psi_h|_{F_e^K} \in \mathbb{P}_{k+1}(F_e^K) \text{ for } 1 \le e \le d+1\} & \text{if } k \text{ is even,} \\ \{\psi_h \in \mathbb{P}_{k+3}(K) : \psi_h|_{F_e^K} \in \mathbb{P}_k(F_e^K) \oplus \tilde{\mathbb{P}}_{k+2}(F) \text{ for } 1 \le e \le d+1\} & \text{if } k \text{ is odd,} \end{cases}$$

where $\tilde{\mathbb{P}}_{k+2}(F)$ denotes the $L^2(F)$-orthogonal complement of $\mathbb{P}_{k+1}(F)$ in $\mathbb{P}_{k+2}(F)$ for any facet $F \in \mathcal{F}_h$. We refer to [2] for the definition of *ad hoc* finite-dimensional spaces M_h for various families and types of elements.

Observe that the norm $\|\cdot\|_S$ measures elements of $\mathbf{X}$ in $\mathbf{H}(\mathrm{div}\,, \mathcal{T}_h) \times L^2(\Omega)$ norm. This corresponds precisely to the energy norm (cf. [18, Chapter 8]).
We propose two alternatives, the first one is to measure the error with respect to the $|\cdot|_{+,K}$ norm (20), the second one is to measure the elements of $\mathbf{X}$ in the weaker $\mathbf{L}^2(\Omega) \times L^2(\Omega)$ norm. Both approaches are respectively developed in sections 4.2.1 and 4.2.2.

4.2.1 Estimates in $|\cdot|_{+,K}$ norm (20)

Lemma 1. *Let ζ and ζ_h be respectively the solution to* (12) *and* (14)*. Let $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h)$ be a reconstruction of ζ_h. We have for all $\xi \in \mathbf{X}$,*

$$d(\zeta - \tilde{\zeta}_h, \xi) = (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\,\tilde{\phi}_h, \mathbf{q})_{0,\Omega}. \tag{21}$$

Proof. Let ξ be in $\mathbf{X}$. According to (12), we have

$$d(\zeta - \tilde{\zeta}_h, \xi) = (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h, \mathbf{q})_{0,\Omega} + (\tilde{\phi}_h, \mathrm{div}\,\mathbf{q})_{0,\Omega}.$$

Owing to the fact that $\tilde{\phi}_h$ is in V, we can integrate by part the last integral:

$$d(\zeta - \tilde{\zeta}_h, \xi) = (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h, \mathbf{q})_{0,\Omega} - (\mathbf{grad}\,\tilde{\phi}_h, \mathbf{q})_{0,\Omega}.$$

This concludes the proof. □

We now derive a similar lemma where the reconstruction $\tilde{\phi}_h^{CR}$ belongs to the Crouzeix-Raviart approximation space

$$V_h^{CR} = \{\psi_h \in L^2(\mathcal{T}_h) \mid \forall K \in \mathcal{T}_h, \psi_{h|K} \in H^1(K); \forall F \in \mathcal{F}_h^i, \int_F [\psi_h] = 0; \forall F \in \mathcal{F}_h^e, \psi_{h|F} = 0\}.$$

Lemma 2. *Let ζ and ζ_h be respectively the solution to* (12) *and* (14)*. Let $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h^{CR})$ be such that $\mathbf{p}_h \in \mathbf{Q}_h$ and $\tilde{\phi}_h^{CR} \in V_h^{CR}$. We have for all $\xi = (\mathbf{q}, \psi) \in \mathbf{X}$, such that $\mathbf{q} \in \mathbf{H}^s(\Omega)$ for some $s \in (0, 1/2)$,*

$$\begin{aligned} d(\zeta - \tilde{\zeta}_h, \xi) = &\sum_{K \in \mathcal{T}_h} \int_K (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h^{CR})\psi - \sum_{K \in \mathcal{T}_h} \int_K (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\,\tilde{\phi}_h^{CR}) \cdot \mathbf{q} \\ &+ \sum_{F \in \mathcal{F}_h^i} \int_F \left(\mathbf{q} - \frac{1}{|F|}\int_F \mathbf{q}\right) \cdot \left[\tilde{\phi}_h^{CR}\right]. \end{aligned} \tag{22}$$

Remark 5. *Let $s \in (0, 1/2)$. For a vector field $\mathbf{q} \in \mathbf{H}^s(K)$ with $\mathrm{div}\,\mathbf{q} \in L^2(K)$, it holds that $\mathbf{q} \cdot \mathbf{n}_{|\partial K} \in H^{-1/2+s}(\partial K)$. Hence one may split duality brackets over ∂K into the sum of duality brackets over the facets. To emphasize this point, we use integrals instead of duality brackets.*

Proof. Let ξ be in $\mathbf{X}$. According to (12), we have

$$d(\zeta - \tilde{\zeta}_h, \xi) = \sum_{K \in \mathcal{T}_h} \int_K (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h^{CR})\psi - \sum_{K \in \mathcal{T}_h} \int_K \mathbb{D}^{-1}\mathbf{p}_h \cdot \mathbf{q} + \sum_{K \in \mathcal{T}_h} \int_K \tilde{\phi}_h^{CR} \mathrm{div}\,\mathbf{q}.$$

Owing to the fact that $\tilde{\phi}_h^{CR}$ is in V_h^{CR}, we can integrate by part the last term:

$$\begin{aligned} d(\zeta - \tilde{\zeta}_h, \xi) = &\sum_{K \in \mathcal{T}_h} \int_K (S_f - \mathrm{div}\,\mathbf{p}_h - \Sigma_a \tilde{\phi}_h^{CR})\psi - \sum_{K \in \mathcal{T}_h} \int_K \mathbb{D}^{-1}\mathbf{p}_h \cdot \mathbf{q} - \sum_{K \in \mathcal{T}_h} \int_K \mathbf{grad}\,\tilde{\phi}_h^{CR} \cdot \mathbf{q} \\ &+ \sum_{F \in \mathcal{F}_h^i} \int_F \mathbf{q} \cdot \left[\tilde{\phi}_h^{CR}\right]. \end{aligned}$$

Above, we use integrals over facets, see remark 5. Using the definition of V_h^{CR}, we obtain that

$$d(\zeta - \tilde{\zeta}_h, \xi) = \sum_{K\in\mathcal{T}_h} \int_K (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h^{CR})\psi - \sum_{K\in\mathcal{T}_h} \int_K \mathbb{D}^{-1}\mathbf{p}_h \cdot \mathbf{q} - \sum_{K\in\mathcal{T}_h} \int_K \mathbf{grad}\ \tilde{\phi}_h^{CR} \cdot \mathbf{q}$$
$$+ \sum_{F\in\mathcal{F}_h^i} \int_F \left(\mathbf{q} - \frac{1}{|F|}\int_F \mathbf{q}\right) \cdot \left[\tilde{\phi}_h^{CR}\right].$$

This concludes the proof. □

Theorem 3. *Let ζ and ζ_h be respectively the solution to* (12) *and* (14)*. Let $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h)$ be a reconstruction of $\zeta_h = (\mathbf{p}_h, \phi_h)$ in $\mathbf{Q}_h \times V$. For any $K \in \mathcal{T}_h$, we define the residual estimator*

$$\eta_{r,K} = \|\Sigma_a^{-1/2}(S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a\tilde{\phi}_h)\|_{0,K}, \tag{23}$$

the flux estimator

$$\eta_{f,K} = \|\mathbb{D}^{1/2}(\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\tilde{\phi}_h)\|_{0,K}, \tag{24}$$

and the non-conforming estimator

$$\eta_{nc,K} = \|\Sigma_a^{1/2}(\tilde{\phi}_h - \phi_h)\|_{0,K}. \tag{25}$$

Then it stands for all $K \in \mathcal{T}_h$

$$|\zeta - \tilde{\zeta}_h|_{+,K} \leq \left(\eta_{r,K}^2 + \sum_{K'\in N(K)} \eta_{f,K'}^2\right)^{1/2}, \tag{26}$$

$$|\zeta - \zeta_h|_{+,K} \leq \left(\eta_{r,K}^2 + \sum_{K'\in N(K)} \eta_{f,K'}^2\right)^{1/2} + \left(\eta_{nc,K}^2 + \sum_{K'\in N(K)} \eta_{nc,K'}^2\right)^{1/2}. \tag{27}$$

Proof. Using the triangle inequality, we obtain for all $K \in \mathcal{T}_h$

$$|\zeta - \zeta_h|_{+,K} \leq |\zeta - \tilde{\zeta}_h|_{+,K} + |\tilde{\zeta}_h - \zeta_h|_{+,K}. \tag{28}$$

According to lemma 1, we have

$$d(\zeta - \tilde{\zeta}_h, \xi) = (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a\tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\ \tilde{\phi}_h, \mathbf{q})_{0,\Omega}. \tag{29}$$

Let $K \in \mathcal{T}_h$ and $\xi = (\mathbf{q}, \psi) \in \mathbf{X}$ be such that $\text{Supp}(\psi) \subset K,\ \text{Supp}(\mathbf{q}) \subset N(K)$. Applying Cauchy Schwarz inequalities successively in $L^2(K)$, $L^2(K')$ for $K' \in N(K)$, and then in $\mathbb{R}^{1+N(K)}$, we get

$$d(\zeta - \tilde{\zeta}_h, \xi) \leq \eta_{r,K}\|\Sigma_a^{1/2}\psi\|_{0,K} + \sum_{K'\in N(K)} \eta_{f,K'}\|\mathbb{D}^{-1/2}\mathbf{q}\|_{0,K'}$$
$$\leq \left(\eta_{r,K}^2 + \sum_{K'\in N(K)} \eta_{f,K'}^2\right)^{1/2} \left(\|\Sigma_a^{1/2}\psi\|_{0,K}^2 + \sum_{K'\in N(K)} \|\mathbb{D}^{-1/2}\mathbf{q}\|_{0,K'}^2\right)^{1/2}$$

We infer (26) from the definition of the $|\cdot|_{+,K}$ norm (20). Now, we want to bound the second term of the right-hand side of (28). We look for an upper bound to

$$\begin{aligned} d(\tilde{\zeta}_h - \zeta_h, \xi) &= d_S(\tilde{\zeta}_h - \zeta_h, \xi) + d_A(\tilde{\zeta}_h - \zeta_h, \xi) \\ &\leq (\Sigma_a(\tilde{\phi}_h - \phi_h), \psi)_{0,K} - (\operatorname{div} \mathbf{q}\,, \tilde{\phi}_h - \phi_h)_{0,\Omega} \\ &\leq \eta_{nc,K} \|\Sigma_a^{1/2}\psi\|_{0,K} + \sum_{K' \in N(K)} \eta_{nc,K'} \|\Sigma_a^{-1/2} \operatorname{div} \mathbf{q}\|_{0,K'} \\ &\leq \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,K'}^2\right)^{1/2} \left(\|\Sigma_a^{1/2}\psi\|_{0,K}^2 + \sum_{K' \in N(K)} \|\Sigma_a^{-1/2} \operatorname{div} \mathbf{q}\|_{0,K'}^2\right)^{1/2}, \end{aligned} \tag{30}$$

where we used Cauchy-Schwarz inequalities in the last two lines. Collecting (26), (28) and (30), we get the desired estimate. □

Remark 6. *Using the same arguments as in the proof of theorem 3, there holds for all $K \in \mathcal{T}_h$,*

$$|\zeta - \tilde{\zeta}_h|_{+,K} \leq \left((\widehat{\eta}_{r,K} + \eta_{nc,K})^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2},$$

$$|\zeta - \zeta_h|_{+,K} \leq \left((\widehat{\eta}_{r,K} + \eta_{nc,K})^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2} + \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,K'}^2\right)^{1/2},$$

where $\widehat{\eta}_{r,K} = \|\Sigma_a^{-1/2}(S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \phi_h)\|_{0,K}$. By definition, it holds that $\eta_{r,K} \leq \widehat{\eta}_{r,K} + \eta_{nc,K}$. Actually, we observe numerically that the residual estimator $\eta_{r,K}$ is "close" to the non-conforming estimator $\eta_{nc,K}$, see Section 5.5. This property is closely related to the weak variational formulation (14)*. In other words, the second equation in* (4) *being imposed weakly, it is expected that $\widehat{\eta}_{r,K} << \eta_{nc,K}$.*

In order to state the next theorem, we will use the following assumption.

Assumption 1. *The parameters $\mathbb{D}$, Σ_a are piecewise constant on $\mathcal{T}_h$ and $S_f \in L_h$.*

Under Assumption 1, one may define

$$\mathbb{D}_K^{max} = \sup_{\mathbf{q} \in \mathbf{L}^2(K) \setminus \{0\}} \frac{(\mathbb{D}\mathbf{q}, \mathbf{q})_{0,K}}{\|\mathbf{q}\|_{0,K}^2}, \quad \mathbb{D}_K^{min} = \inf_{\mathbf{q} \in \mathbf{L}^2(K) \setminus \{0\}} \frac{(\mathbb{D}\mathbf{q}, \mathbf{q})_{0,K}}{\|\mathbf{q}\|_{0,K}^2}.$$

Theorem 4 (local efficiency of the *a posteriori* error estimators)**.** *Let $K \in \mathcal{T}_h$ and let $\eta_{r,K}$ and $\eta_{f,K}$ be the residual estimators respectively given by* (23) *and* (24)*. Under Assumption 1, the following estimates hold true*

$$\eta_{r,K} \leq |\zeta - \tilde{\zeta}_h|_{+,K} \mathtt{c}, \tag{31}$$

$$\eta_{f,K} \leq |\zeta - \tilde{\zeta}_h|_{+,K} \left\{ \mathtt{c}^2 \frac{\mathbb{D}_K^{max}}{\mathbb{D}_K^{min}} + \mathtt{C}^2 \frac{\mathbb{D}_K^{max}}{h_K^2 \Sigma_{a,K}} \right\}^{1/2}, \tag{32}$$

where $\mathtt{c}$ and $\mathtt{C}$ are constants which depend only on the polynomial degree k of S_f, d, and the shape-regularity parameter $\kappa_K := |K|/h_K^d$.

Proof. The proof follows that given in [28, Lemma 7.6]. Let ψ_K be the bubble function on K, given as the product of the $d+1$ linear functions that take the value 1 at one vertex of K and vanish at the other vertices, and let us denote $\psi_r = (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \widetilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that ψ_r is a polynomial in K by Assumption 1. Then the equivalence of norms on finite-dimensional spaces gives

$$c\|\psi_r\|_{0,K}^2 \leq (\psi_r, \psi_K \psi_r)_{0,K}, \tag{33}$$

$$\|\psi_K \psi_r\|_{0,K} \leq \|\psi_r\|_{0,K}, \tag{34}$$

with the constant c depending only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{r,K} = (0, \psi_K \psi_r) \in \mathbf{X}$, we immediately have (cf. the proof of lemma 1)

$$d(\zeta - \tilde{\zeta}_h, \xi_{r,K}) = (\psi_r, \psi_K \psi_r)_{0,K},$$

and by definition (20) of the $|\cdot|_{+,K}$ norm,

$$\begin{aligned} d(\zeta - \tilde{\zeta}_h, \xi_{r,K}) &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\xi_{r,K}\|_S \\ &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\Sigma_a^{1/2} \psi_K \psi_r\|_{0,K}. \end{aligned} \tag{35}$$

Combining (33), (34) and (35), one comes to

$$c\|\psi_r\|_{0,K}^2 \leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\psi_r\|_{0,K} (\Sigma_{a,K})^{1/2}.$$

Using the definition of $\eta_{r,K}$ by (23) concludes the proof of (31):

$$\eta_{r,K} = (\Sigma_{a,K})^{-1/2} \|\psi_r\|_{0,K} \leq \frac{1}{c} |\zeta - \tilde{\zeta}_h|_{+,K}.$$

We now proceed similarly for the second estimate. Let us denote $\mathbf{q}_f = (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\ \tilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that $\mathbf{q}_f$ is a polynomial in K by Assumption 1. Then the equivalence of norms on finite-dimensional spaces and the inverse inequality (cf., e.g., [8, Theorem 3.2.6]) give

$$c\|\mathbf{q}_f\|_{0,K}^2 \leq (\mathbf{q}_f, \psi_K \mathbf{q}_f)_{0,K}, \tag{36}$$

$$\|\psi_K \mathbf{q}_f\|_{0,K} \leq \|\mathbf{q}_f\|_{0,K}, \tag{37}$$

$$\|\operatorname{div}(\psi_K \mathbf{q}_f)\|_{0,K} \leq \frac{C}{h_K} \|\psi_K \mathbf{q}_f\|_{0,K}, \tag{38}$$

with the constants c and C depending only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{f,K} = (\psi_K \mathbf{q}_f, 0) \in \mathbf{X}$, we immediately have (cf. the proof of lemma 1)

$$-d(\zeta - \tilde{\zeta}_h, \xi_{f,K}) = (\mathbf{q}_f, \psi_K \mathbf{q}_f)_{0,K},$$

and again by definition (20) of the $|\cdot|_{+,K}$ norm,

$$\begin{aligned} -d(\zeta - \tilde{\zeta}_h, \xi_{f,K}) &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\xi_{f,K}\|_S \\ &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \{\|\mathbb{D}^{-1/2}(\psi_K \mathbf{q}_f)\|_{0,K}^2 + \|\Sigma_a^{-1/2} \operatorname{div}(\psi_K \mathbf{q}_f)\|_{0,K}^2\}^{1/2} \\ &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \{(\mathbb{D}_K^{min})^{-1} \|(\psi_K \mathbf{q}_f)\|_{0,K}^2 + (\Sigma_{a,K})^{-1} \|\operatorname{div}(\psi_K \mathbf{q}_f)\|_{0,K}^2\}^{1/2} \\ &\leq |\zeta - \tilde{\zeta}_h|_{+,K} \{(\mathbb{D}_K^{min})^{-1} + C^2 (h_K^2 \Sigma_{a,K})^{-1}\}^{1/2} \|(\psi_K \mathbf{q}_f)\|_{0,K}, \end{aligned} \tag{39}$$

where we used the inverse inequality (38) to reach the last line. Combining (36), (37) and (39), one comes to

$$c\|\mathbf{q}_f\|_{0,K}^2 \leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\mathbf{q}_f\|_{0,K} \{(\mathbb{D}_K^{min})^{-1} + C^2 (h_K^2 \Sigma_{a,K})^{-1}\}^{1/2}.$$

Considering the definition of $\eta_{f,K}$ by (24) concludes the proof. □

Remark 7. *Assume in addition in Theorem 4 that there exists a constant $\kappa > 0$, such that $\min_{K \in \mathcal{T}_h} \kappa_K \geq \kappa$, for all $h > 0$. Then, the constants $\mathtt{c}$ and $\mathtt{C}$ do not depend on κ_K (but on κ).*

4.2.2 Estimates in $\mathbf{L}^2(\Omega) \times L^2(\Omega)$ norm

In this section, we define

$$|||\xi|||_{\mathcal{T}_h}^2 := d_S(\xi, \xi) = \sum_{K \in \mathcal{T}_h} |||\xi|||_K^2,$$

where

$$|||\xi|||_K^2 := \|\mathbb{D}^{-1/2}\mathbf{q}\|_{0,K}^2 + \|\Sigma_a^{1/2}\psi\|_{0,K}^2. \tag{40}$$

Theorem 5. *Let ζ be the weak solution of Problem* (12)*. If $\zeta_h = (\mathbf{p}_h, \phi_h)$ is the discrete solution to Problem* (14) *and $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h) \in \mathbf{X}$ is a reconstruction of ζ_h, then*

$$|||\zeta - \tilde{\zeta}_h|||_{\mathcal{T}_h} \leq \left(\sum_{K \in \mathcal{T}_h} \eta_{r,K}^2 + \eta_{f,K}^2 \right)^{1/2}, \tag{41}$$

$$|||\zeta - \zeta_h|||_{\mathcal{T}_h} \leq \left(\sum_{K \in \mathcal{T}_h} \eta_{r,K}^2 + \eta_{f,K}^2 \right)^{1/2} + \left(\sum_{K \in \mathcal{T}_h} \eta_{nc,K}^2 \right)^{1/2}. \tag{42}$$

Proof. We have

$$\begin{aligned} |||\zeta - \tilde{\zeta}_h|||_{\mathcal{T}_h}^2 &= d_S(\zeta - \tilde{\zeta}_h, \zeta - \tilde{\zeta}_h) = d(\zeta - \tilde{\zeta}_h, \zeta - \tilde{\zeta}_h) \\ &\leq |||\zeta - \tilde{\zeta}_h|||_{\mathcal{T}_h} \sup_{\xi \in \mathbf{X}, |||\xi|||_{\mathcal{T}_h} = 1} d(\zeta - \tilde{\zeta}_h, \xi). \end{aligned} \tag{43}$$

We observe that Equation (21) in lemma 1 may be written as

$$d(\zeta - \tilde{\zeta}_h, \xi) = \sum_{K \in \mathcal{T}_h} (S_f + \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,K} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h, \mathbf{q})_{0,K}. \tag{44}$$

Using Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned} d(\zeta - \tilde{\zeta}_h, \xi) &\leq \sum_{K \in \mathcal{T}_h} (\eta_{r,K} \|\Sigma_a^{1/2}\psi\|_{0,K} + \eta_{f,K} \|\mathbb{D}^{-1/2}\mathbf{q}\|_{0,K}) \\ &\leq \sum_{K \in \mathcal{T}_h} (\eta_{r,K}^2 + \eta_{f,K}^2)^{1/2} |||\xi|||_K \\ &\leq \left(\sum_{K \in \mathcal{T}_h} \eta_{r,K}^2 + \eta_{f,K}^2 \right)^{1/2} |||\xi|||_{\mathcal{T}_h}. \end{aligned} \tag{45}$$

Collecting (43), (44) and (45), we infer (41). Using the triangle inequality, we obtain

$$|||\zeta - \zeta_h|||_{\mathcal{T}_h} \leq |||\zeta - \tilde{\zeta}_h|||_{\mathcal{T}_h} + |||\tilde{\zeta}_h - \zeta_h|||_{\mathcal{T}_h}.$$

From the definition of the reconstruction, we get

$$|||\tilde{\zeta}_h - \zeta_h|||_{\mathcal{T}_h} = \left(\sum_{K \in \mathcal{T}_h} \eta_{nc,K}^2 \right)^{1/2}.$$

This concludes the proof. □

Theorem 6 (local efficiency of the *a posteriori* error estimators)**.** *Let $K \in \mathcal{T}_h$ and let $\eta_{r,K}$ and $\eta_{f,K}$ be the residual estimators respectively given by* (23) *and* (24)*. Under Assumption 1, the following estimates hold true*

$$\eta_{r,K} \leq |||\zeta - \tilde{\zeta}_h|||_K \left(\mathtt{c}^2 + \mathtt{C}^2 \frac{\mathbb{D}_K^{max}}{h_K^2 \Sigma_{a,K}} \right)^{1/2}, \tag{46}$$

$$\eta_{f,K} \leq |||\zeta - \tilde{\zeta}_h|||_K \left(\mathtt{c}^2 \frac{\mathbb{D}_K^{max}}{\mathbb{D}_K^{min}} + \mathtt{C}^2 \frac{\mathbb{D}_K^{max}}{h_K^2 \Sigma_{a,K}} \right)^{1/2} \tag{47}$$

where $\mathtt{c}$ *and* $\mathtt{C}$ *are constants which depend only on the polynomial degree k of S_f, d, and the shape-regularity parameter κ_K.*

Proof. The proof follows that given in [28, Lemma 7.6]. Let ψ_K be the bubble function on K defined as in the proof of Theorem 4, and let us denote $\psi_r = (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \widetilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that ψ_r is a polynomial in K by Assumption 1. Then the equivalence of norms on finite-dimensional spaces and the inverse inequality (cf., e.g., [8, Theorem 3.2.6]) give (33), (34) and

$$\|\mathbb{D}^{1/2} \mathbf{grad}\,(\psi_K \psi_r)\|_{0,K} \leq C (\mathbb{D}_K^{max})^{1/2} h_K^{-1} \|\psi_K \psi_r\|_{0,K}, \tag{48}$$

with the constants c and C depending only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{r,K} = (0, \psi_K \psi_r) \in \mathbf{X}$, we immediately have (cf. the proof of lemma 1)

$$d(\zeta - \tilde{\zeta}_h, \xi_{r,K}) = (\psi_r, \psi_K \psi_r)_{0,K},$$

and ,

$$\begin{aligned} d(\zeta - \tilde{\zeta}_h, \xi_{r,K}) &= t(\phi - \tilde{\phi}_h, \psi_K \psi_r) + b(\mathbf{p} - \mathbf{p}_h, \psi_K \psi_r) \\ &= t(\phi - \tilde{\phi}_h, \psi_K \psi_r) - \int_K (\mathbf{p} - \mathbf{p}_h) \cdot \mathbf{grad}\,(\psi_K \psi_r) \\ &\leq |||\zeta - \tilde{\zeta}_h|||_K \left(\|\mathbb{D}^{1/2} \mathbf{grad}\,(\psi_K \psi_r)\|_{0,K}^2 + \|\Sigma_a^{1/2} \psi_K \psi_r\|_{0,K}^2 \right)^{1/2}, \end{aligned} \tag{49}$$

where we integrated by parts the second term in the second line. Combining (33), (34), (48) and (49), one comes to

$$c \|\psi_r\|_{0,K}^2 \leq |||\zeta - \tilde{\zeta}_h|||_K \|\psi_r\|_{0,K} \left(C^2 \mathbb{D}_K^{max} h_K^{-2} + \Sigma_{a,K} \right)^{1/2}.$$

Considering the definition (23) of $\eta_{r,K}$ concludes the proof of (46).
We now proceed similarly for the second estimate. Let us denote $\mathbf{q}_f = (\mathbb{D}^{-1} \mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that $\mathbf{q}_f$ is a polynomial in K by Assumption 1. Then the equivalence of norms on finite-dimensional spaces, and the inverse inequality (cf., e.g., [8, Theorem 3.2.6]) give (36)-(37)-(38) with the constants c and C depending only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{f,K} = (\psi_K \mathbf{q}_f, 0) \in \mathbf{X}$, we immediately have (cf. the proof of lemma 1)

$$-d(\zeta - \tilde{\zeta}_h, \xi_{f,K}) = (\mathbf{q}_f, \psi_K \mathbf{q}_f)_{0,K},$$

and

$$\begin{aligned} &- d(\zeta - \tilde{\zeta}_h, \xi_{f,K}) \\ &= a(\mathbf{p} - \mathbf{p}_h, \psi_K \mathbf{q}_f) + b(\psi_K \mathbf{q}_f, \phi - \tilde{\phi}_h) \\ &\leq |||\zeta - \tilde{\zeta}_h|||_K \left(\|\mathbb{D}^{-1/2} (\psi_K \mathbf{q}_f)\|_{0,K}^2 + \|\Sigma_a^{-1/2} \operatorname{div}(\psi_K \mathbf{q}_f)\|_{0,K}^2 \right)^{1/2}. \end{aligned} \tag{50}$$

Combining (36), (37), (38) and (50), one comes to

$$c \|\mathbf{q}_f\|_{0,K}^2 \leq |||\zeta - \tilde{\zeta}_h|||_K \|\mathbf{q}_f\|_{0,K} \left(C^2 (h_K^2 \Sigma_{a,K})^{-1} + (\mathbb{D}_K^{min})^{-1} \right)^{1/2}.$$

Considering the definition of $\eta_{f,K}$ by (24) concludes the proof. □

4.2.3 Estimates in the primal energy norm

In this section, we aim to briefly recall the *a posteriori* error framework introduced in [28]. The energy norm associated to the primal form is

$$|||\phi|||_p^2 = \|\mathbb{D}^{1/2}\mathbf{grad}\,\phi\|_{0,\Omega}^2 + \|\Sigma_a^{1/2}\phi\|_{0,\Omega}^2.$$

Therefore, we define

$$|||\psi|||_{p,\mathcal{T}_h}^2 = \sum_{K\in\mathcal{T}_h} |||\psi|||_{p,K}^2,$$

where

$$|||\psi|||_{p,K}^2 = \|\mathbb{D}^{1/2}\mathbf{grad}\,\psi\|_{0,K}^2 + \|\Sigma_a^{1/2}\psi\|_{0,K}^2. \tag{51}$$

We have the following *a posteriori* error estimate [28, Theorem 4.2, p.1578].

Theorem 7. *Let ζ and ζ_h be respectively the solution to* (12) *and* (14) *with RTN$_0$ finite elements, and let $\widehat{\phi}_h = \mathcal{I}_{pp}(\phi_h)$. For all $K \in \mathcal{T}_h$, we define the residual estimator*

$$\tilde{\eta}_{r,K} = m_K\|S_f + \operatorname{div}(\mathbb{D}\mathbf{grad}\,\widehat{\phi}_h) - \Sigma_a\widehat{\phi}_h\|_{0,K}, \tag{52}$$

with

$$m_K^2 := \min\left\{C_{P,d}\frac{h_K^2}{\mathbb{D}_K^{min}}, \frac{1}{\Sigma_{a,K}}\right\},$$

where $C_{P,d}$ is the Poincaré constant defined in [28, Definition (2.1)], and the nonconformity estimator

$$\tilde{\eta}_{nc,K} := |||\widehat{\phi}_h - \mathcal{I}_{MO}(\widehat{\phi}_h)|||_{p,K}.$$

Then, under Assumption 1, it holds that

$$|||\phi - \widehat{\phi}_h|||_{p,\mathcal{T}_h} \le \left\{\sum_{K\in\mathcal{T}_h}\tilde{\eta}_{nc,K}^2\right\}^{1/2} + \left\{\sum_{K\in\mathcal{T}_h}\tilde{\eta}_{r,K}^2\right\}^{1/2}. \tag{53}$$

The following theorem states the local efficiency of the residual estimator [28, Theorem 4.4, p.1578-1579].

Theorem 8. *Let ζ and ζ_h be respectively the solution to* (12) *and* (14) *with RTN$_0$ finite elements, and let $\widehat{\phi}_h = \mathcal{I}_{pp}(\phi_h)$. Under Assumption 1, there holds on every $K \in \mathcal{T}_h$*

$$\tilde{\eta}_{r,K} \le \mathtt{C}|||\phi - \widehat{\phi}_h|||_{p,K}\left(\frac{\mathbb{D}_K^{max}}{\mathbb{D}_K^{min}}\right)^{1/2}, \tag{54}$$

with the constant $\mathtt{C}$ depending only on the polynomial degree k of S_f, the space dimension d, and the shape-regularity parameter κ_K.

5 Numerical results

This section is devoted to the numerical experiments we performed on adaptive mesh refinement strategies (AMR). In fact, the AMR strategy can be classified into some categories: the h-refinement (mesh subdivision), which amounts to refining the mesh where large errors occur [27]; the p-refinement (local high order approximation), which increases the order of the polynomial functions [4], or the r-refinement (moving mesh) that moves the nodes of the mesh to increase the mesh density [6], in the regions of interest where large variations of the solution occur. The above strategies can be mixed, such as hp-refinement [3, 12] and hr-refinement [21].

We are interested in the case of heterogeneous coefficients which may induce some singularities in the solution of Problem (1), that is a loss of regularity of the solution due to the interaction among the materials. Therefore, we focus on mesh subdivision strategy in this section.

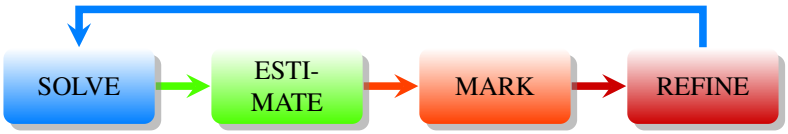
The performance of adaptive mesh refinement is assessed with respect to various criteria such as the error estimator, the marker strategy and the threshold parameter. We recall that the context of our applications is modelling nuclear reactor cores, in particular geometries composed of rectangular cuboids of $\mathbb{R}^3$. This is the reason why the discretization in this section is applied on Cartesian meshes.

Mesh subdivision strategies are introduced in Section 5.1. Section 5.2 presents the set of test cases considered throughout the whole Section 5. Section 5.3 focuses on the marker strategies. The sensitivity with respect to the threshold parameter is investigated in Section 5.4. Section 5.5 compares various reconstruction approaches. Section 5.6 examines different error estimators.

5.1 An adaptive mesh refinement strategy

We recall in this section the h-refinement approach.

From the initial mesh $\mathcal{T}_{h_0}$, we generate a sequence of meshes $\mathcal{T}_{h_k}$ by using the AMR strategy which is in general an iterative loop where at each iteration, we consider the following four modules:



Assuming that the mesh $\mathcal{T}_{h_k}$ is computed, module **Solve** indeed corresponds to solving Problem (14). The output of the module **Estimate** is $(\eta_K)_{K\in\mathcal{T}_{h_k}}$ where η_K is an *a posteriori* error estimator. The stopping criterion of the algorithm is given by

$$\max_{K\in\mathcal{T}_h} \eta_K \leq \epsilon_{AMR},$$

where $\epsilon_{AMR} > 0$. Module **Mark** returns the set of the marked cells based on the error estimators $(\eta_K)_{K\in\mathcal{T}_{h_k}}$. In other words, this module selects a set of elements to be refined. To be convenient, for $S \subset \mathcal{T}_{h_k}$, let us denote $\eta(S) = (\sum_{K\in S} \eta_K^2)^{1/2}$. For a fixed threshold parameter $0 < \theta \leq 1$, the classical bulk chasing (Doffler's marking strategy [14]) is to select the (smallest) set of elements such that

$$\eta(S) \geq \theta\eta(\mathcal{T}_{h_k}). \tag{55}$$

Lastly, module **Refine** performs the mesh refinement according to the selected mesh elements.

This strategy is generic and can be applied to any kind of mesh.

In order to have Cartesian mesh preserving, it is essential to refine the mesh according to the *whole lines* in each direction $(\mathbf{e}_x)_{x=1,d}$ which contain at least one of the selected cells. As a consequence, it is obvious to see that this *cell marker strategy* is really costly since we use the error indicator of just some selected cells to refine the other cells located in the same line for a given direction. Due to this drawback, it is extremely important to point out some other marker cell strategies based on more information. Therefore, instead of using the classical bulk chasing (Doffler's marking strategy) defined on a single cell, we modify it to propose some other error indicators according to the lines of each direction and also on the "cross" value (the total error indicators of all the lines) respectively denoted the *direction marker* and *cross marker* method.
The *direction marker* method consists in selecting for each direction $\mathbf{e}_x, x = 1, \ldots, d$, the small set of lines $L_x \subset \mathcal{T}_{h_k}$ parallel to $\mathbf{e}_x$ such that

$$\eta(L_x) \geq \theta_l \eta(\mathcal{T}_{h_k}), \tag{56}$$

where $0 < \theta_l \leq 1$ is a fixed threshold parameter. The resulting selected set is $\cup_{x=1,d}(L_x)$.
The *cross marker* method corresponds to selecting for each $K \in \mathcal{T}_h$, the small set of elements $S \subset \mathcal{T}_{h_k}$ such that

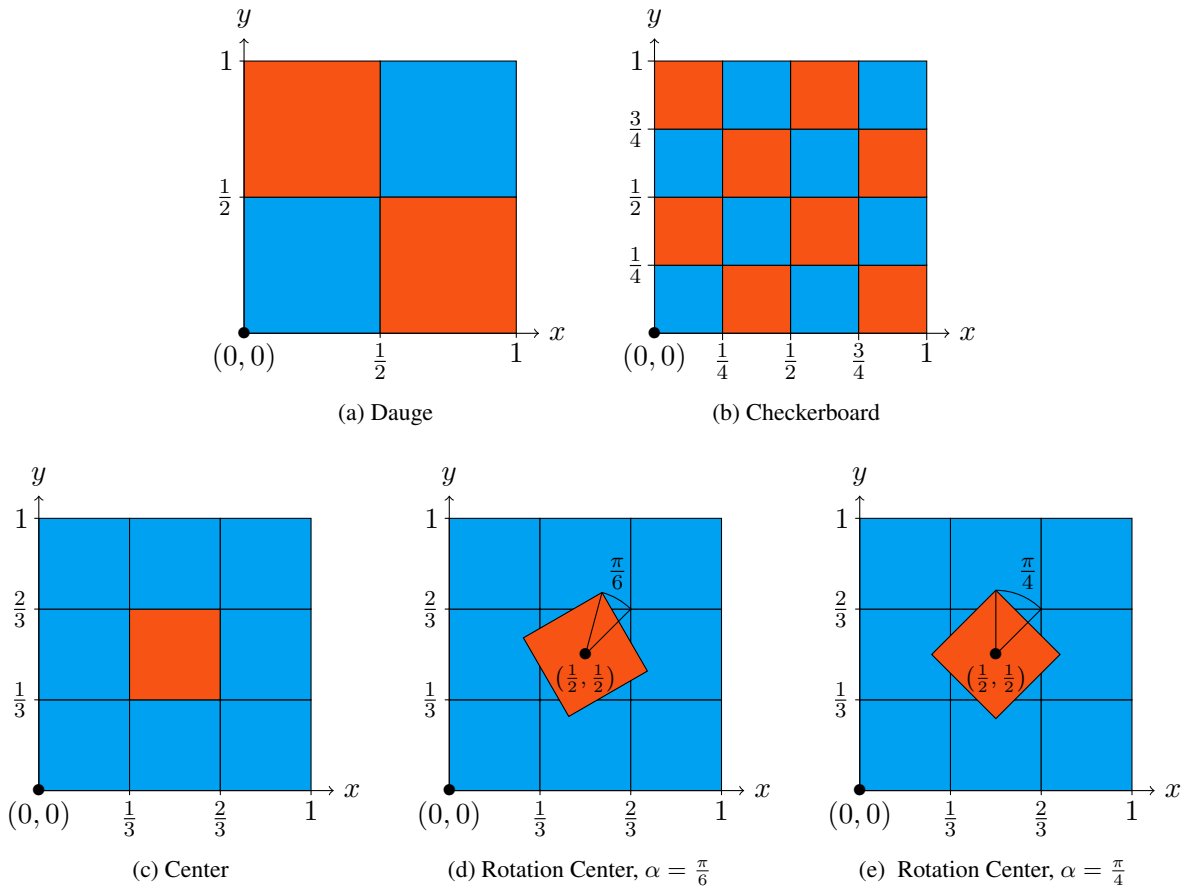
$$\sum_{K \in S} \eta(C_K) \geq \theta_c \eta(\mathcal{T}_{h_k}), \tag{57}$$

where $0 < \theta_c \leq 1$ is a fixed threshold parameter and C_K is the union for all direction $(\mathbf{e}_x)_{x=1,d}$ of the lines containing K. The resulting selected set is $\cup_{K \in S}(C_K)$. Interestingly, performing the mesh refinement is straightforward with both the *direction marker* and *cross marker* methods. In addition, they both preserve the Cartesian structure of the mesh.

5.2 Setting of the test cases

This section is devoted to the definition of the test cases considered, namely the Dauge test case, the Checkerboard test case, the Center test case and the Rotation Center test cases. In the following test cases, we perform the numerical simulations on the domain $\Omega = (0, 1)^2$. We consider here a simple source term given by $S_f = 1$. Moreover, we assume that the diffusion coefficient $\mathbb{D}$ is a scalar, piecewise constant given by Figure 1. We also set $\Sigma_a = 1$.

In the following, the initial mesh of the AMR strategy is chosen to be uniform in all directions. The mesh size of the initial mesh of the Dauge test case, the Checkerboard test case, the Center test case and the Rotation Center test cases are respectively equal to 1/4, 1/8, 1/6 and 1/8.



(a) Dauge (b) Checkerboard

(c) Center (d) Rotation Center, $\alpha = \frac{\pi}{6}$ (e) Rotation Center, $\alpha = \frac{\pi}{4}$

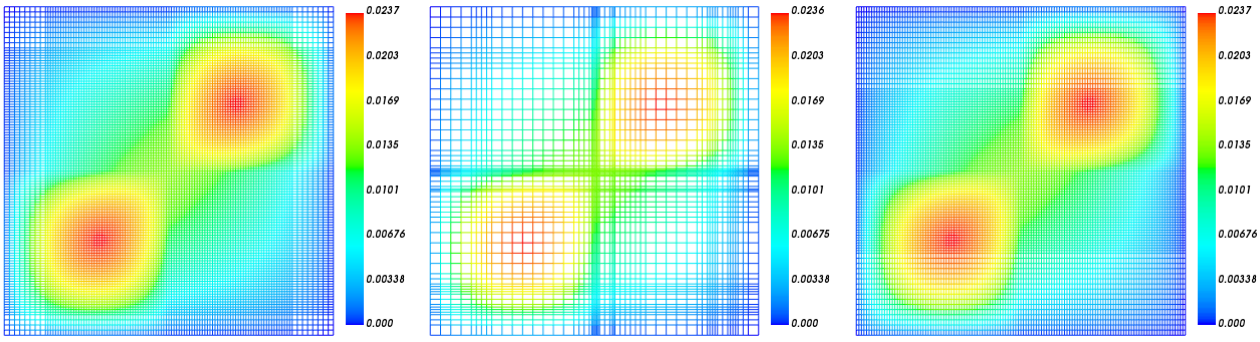
Figure 1: The diffusion coefficient $\mathbb{D}$: the region ■ corresponds to $\mathbb{D} = 10$ and the other region ■ stands for $\mathbb{D} = 1$.

5.3 Influence of the marker cell strategy

We now study the influence of the marker cell strategy on the AMR approach for our set of test cases. In this section, the Dauge test case, the Center test case and the Checkerboard test case are performed with RTN_0 finite elements, while the Rotation Center test cases are performed with RTN_1 finite elements. The AMR strategies are based on the error estimator introduced in (27) and the average reconstruction defined in Section 4.1.1.

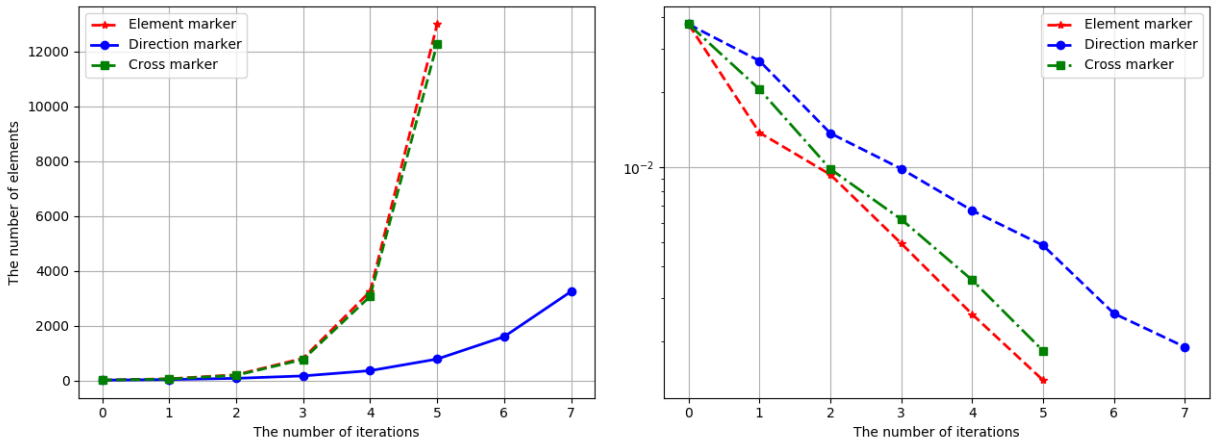
The Dauge test case is a singular toy problem (see also in [13], [11], [18] and references therein for more details). In this test case, the singularity is located at $(0.5, 0.5)$ and we expect refinement in this region. Adaptive mesh refinement is performed with a stopping criterion equal to $\epsilon_{AMR} = 2 \times 10^{-3}$. Figure 2 shows that mesh refinement is more located near the singularity for the *direction marker* strategy than the other strategies. Moreover, Figure 3 shows that the *direction marker* needs at least three times less mesh elements than the other strategies. All the other test cases yield the same conclusions.

So, from now on, the adaptive mesh refinement is always performed with the *direction marker* method.



(a) Element marker (b) Direction marker (c) Cross marker

Figure 2: Dauge test case: the numerical flux on refined meshes for different marker strategies with RTN_0.



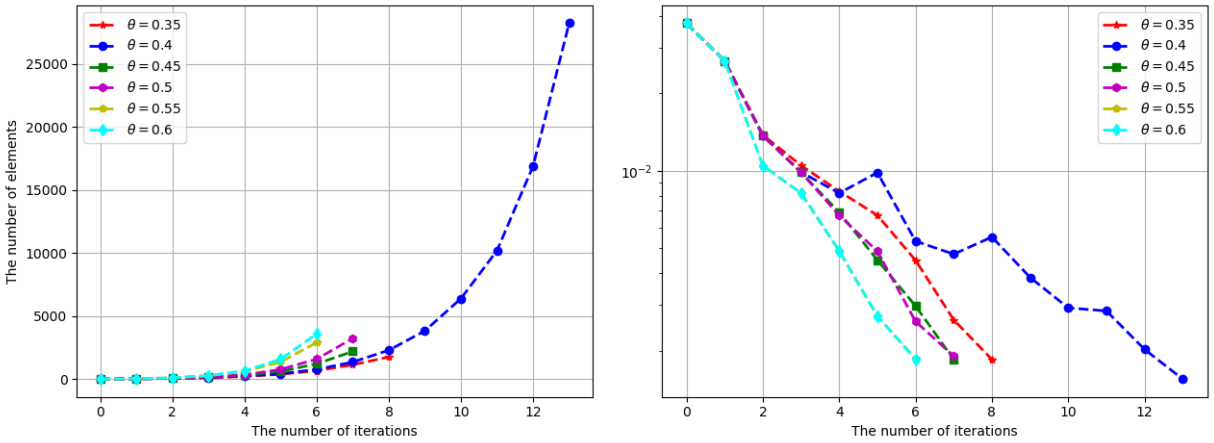
(a) Evolution of the number of elements (b) Evolution of the maximum of total error estimator

Figure 3: Dauge test case: Evolution of the number of elements and the maximum of the total error estimator for different marker cell strategies.

5.4 Sensitivity with respect to the threshold parameter

In this section, we evaluate the sensitivity with respect to the threshold parameter θ_l defined in (56) on the set of test cases presented in Section 5.2. For unstructured meshes like triangular meshes, the typical value for the threshold parameter θ_l is 0.5. However, the choice of an optimal value for the threshold parameter θ_l remains a difficult question. Therefore, we numerically investigate the optimal value of the threshold parameter.

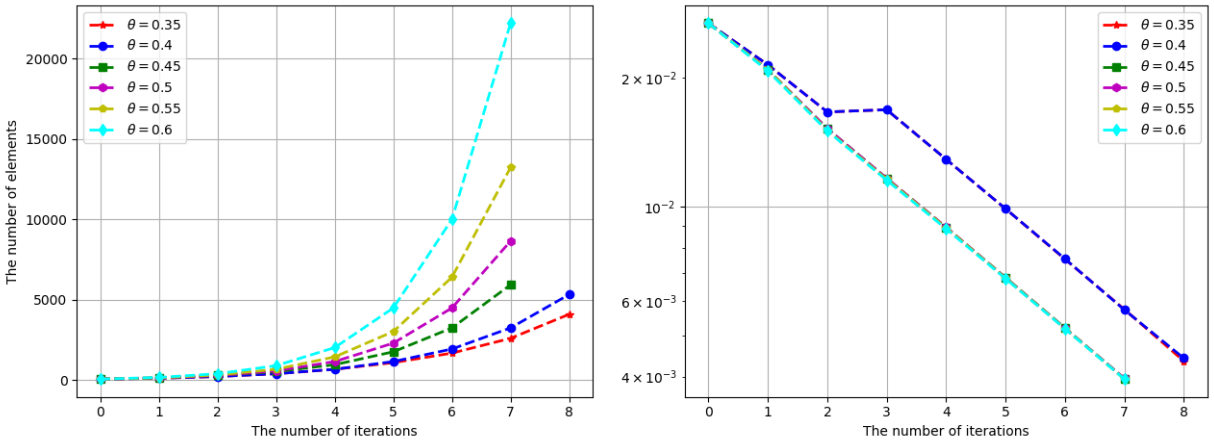
The stopping criterion of the Checkerboard test case is $\epsilon_{AMR} = 5 \times 10^{-3}$. For the other test cases, the stopping criterion is set to $\epsilon_{AMR} = 2 \times 10^{-3}$.
Figures 4 and 5 indicate that the optimal value of θ_l for the Dauge and Checkerboard test case is around 0.35 while Figure 6 shows that the optimal value of this parameter for the Center test case is around 0.6. Moreover, the optimal value of parameter θ_l for the Rotation Center test case with $\alpha = \frac{\pi}{6}$ and $\alpha = \frac{\pi}{4}$ are around 0.45 and 0.4 respectively, see Figures 7 and 8. Figure 9 shows the numerical flux on the refined mesh with an optimal value of θ_l for the different test cases.



(a) Evolution of the number of elements

(b) Evolution of the maximum of total error estimator

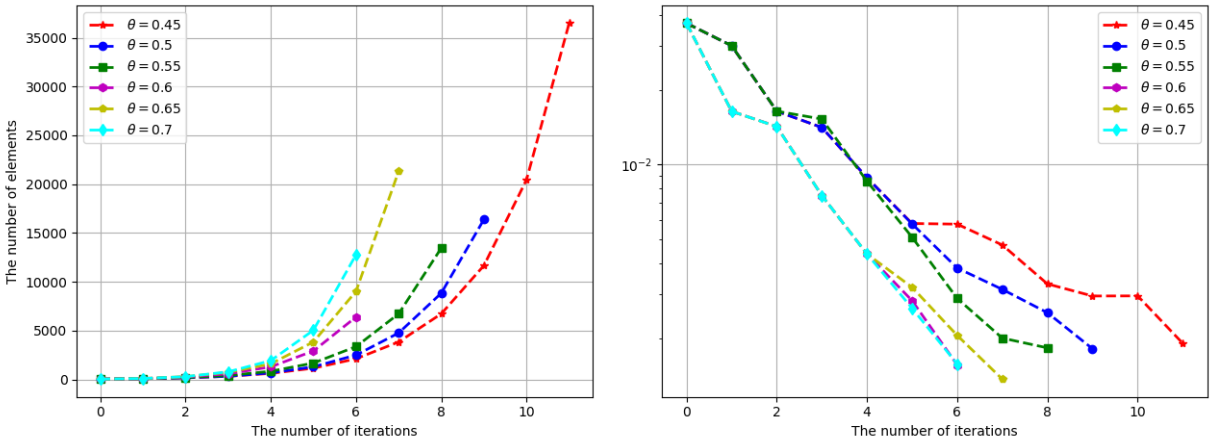
Figure 4: Dauge test case.



(a) Evolution of the number of elements

(b) Evolution of the maximum of total error estimator

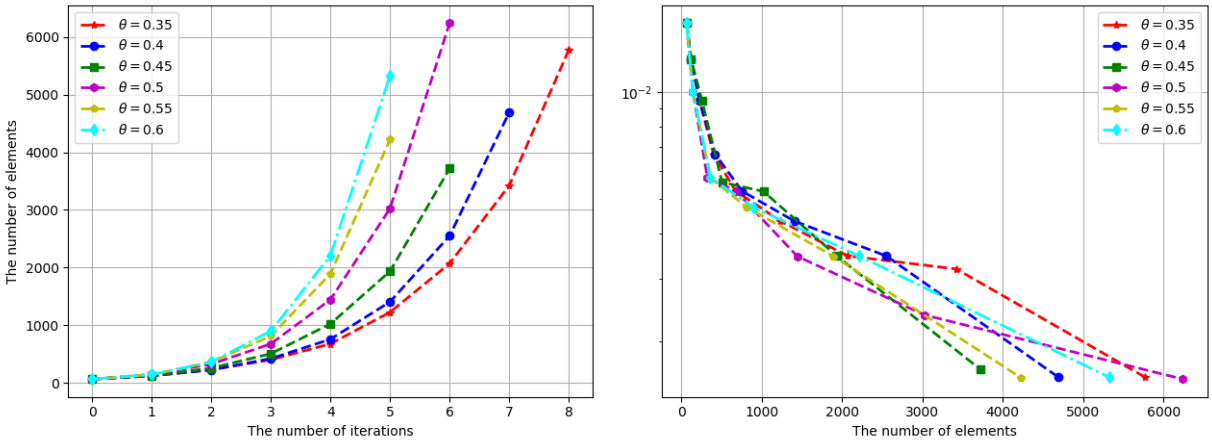
Figure 5: Checkerboard test case.



(a) Evolution of the number of elements

(b) Evolution of the maximum of total error estimator

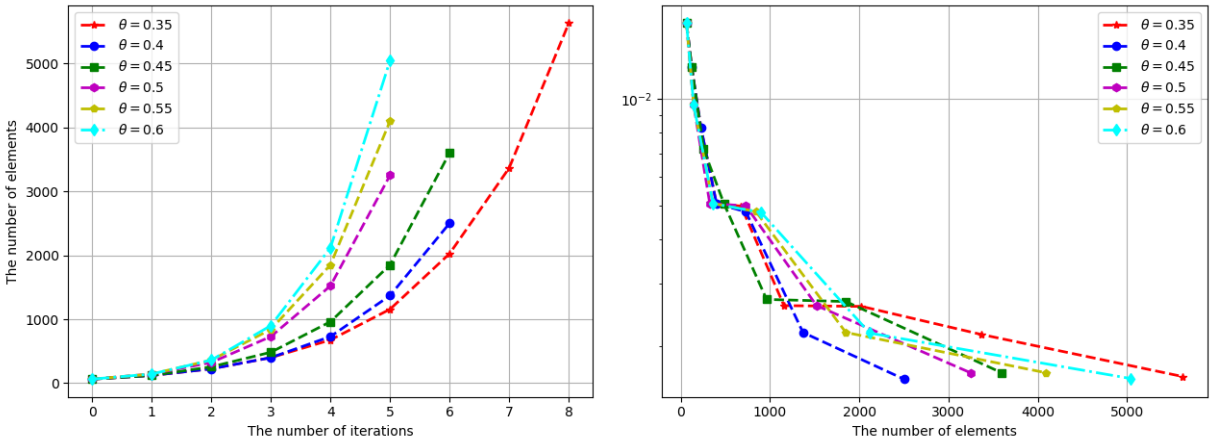
Figure 6: Center test case.



(a) Evolution of the number of elements

(b) Evolution of the maximum of total error estimator

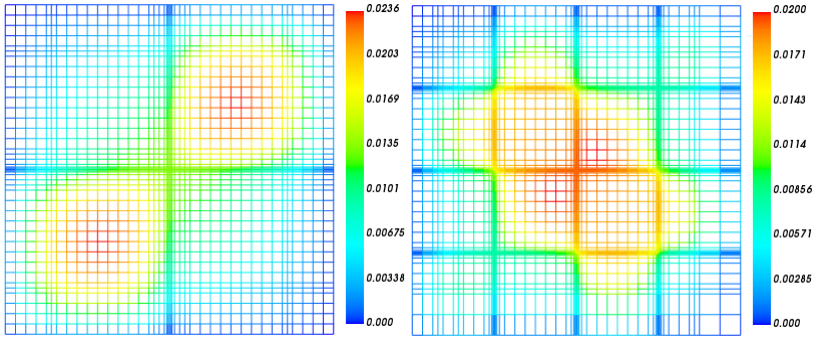
Figure 7: Rotation Center, $\alpha = \pi/6$.



(a) Evolution of the number of elements

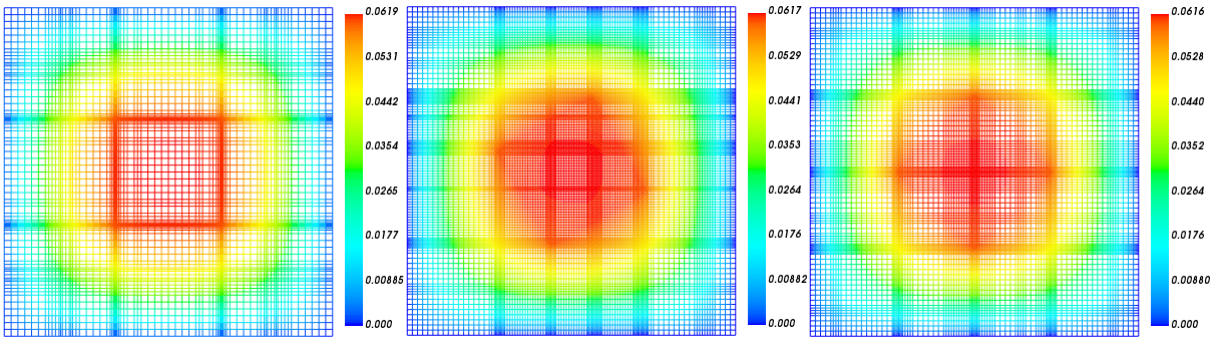
(b) Evolution of the maximum of total error estimator

Figure 8: Rotation Center, $\alpha = \pi/4$.



(a) Dauge, $\theta_l = 0.35$

(b) Checkerboard, $\theta_l = 0.35$



(c) Center, $\theta_l = 0.6$

(d) Rotation Center, $\alpha = \pi/6$, $\theta_l = 0.45$

(e) Rotation Center, $\alpha = \pi/4$, $\theta_l = 0.4$

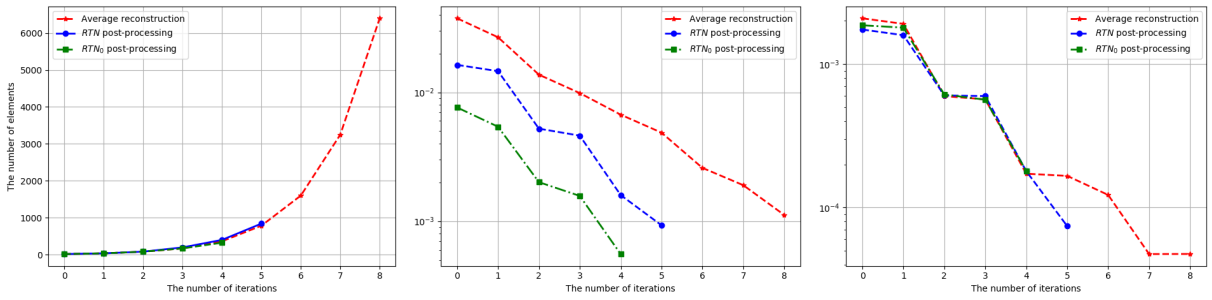
Figure 9: The numerical flux ϕ_h of the different test cases

5.5 Influence of the reconstruction

In this section, we investigate the influence of the reconstruction on the error estimator defined in (27). To this aim, we compare the reconstruction approaches defined in Section 4.1 on the Dauge test case.

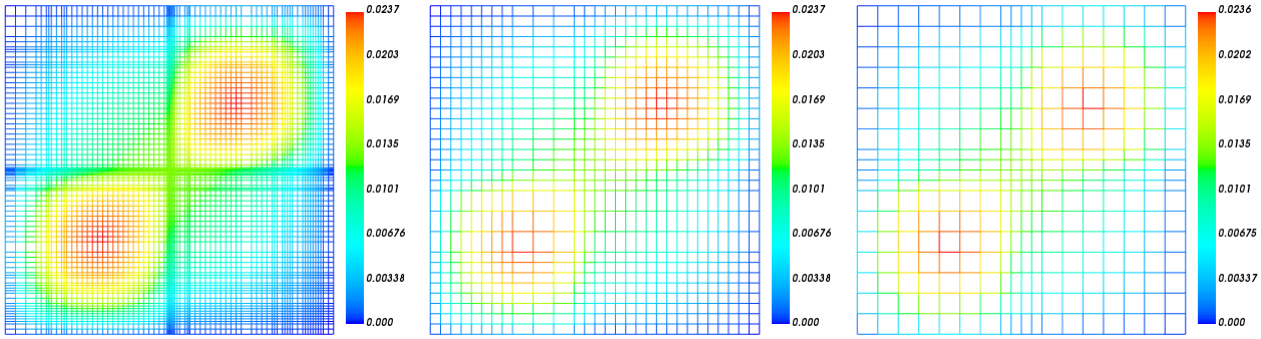
First, the stopping criterion is fixed at $\epsilon_{AMR} = 1.5 \times 10^{-3}$. As can be seen in Figure 10, the average re-

construction and RTN post-processing need more elements to reach the stopping criterion than the RTN_0 post-processing. It is related to the fact that the flux estimator is the dominant contribution of the total estimator. We do not show here the residual estimator which is similar to the non-conforming estimator. Figure 11 shows the numerical flux on the refined mesh for the different reconstructions.



(a) Number of elements (b) Maximum of the total estimator (c) Maximum of the non-conforming estimator

Figure 10: Evolution of the number of elements and the maximum of the total estimator by using different reconstruction methods.



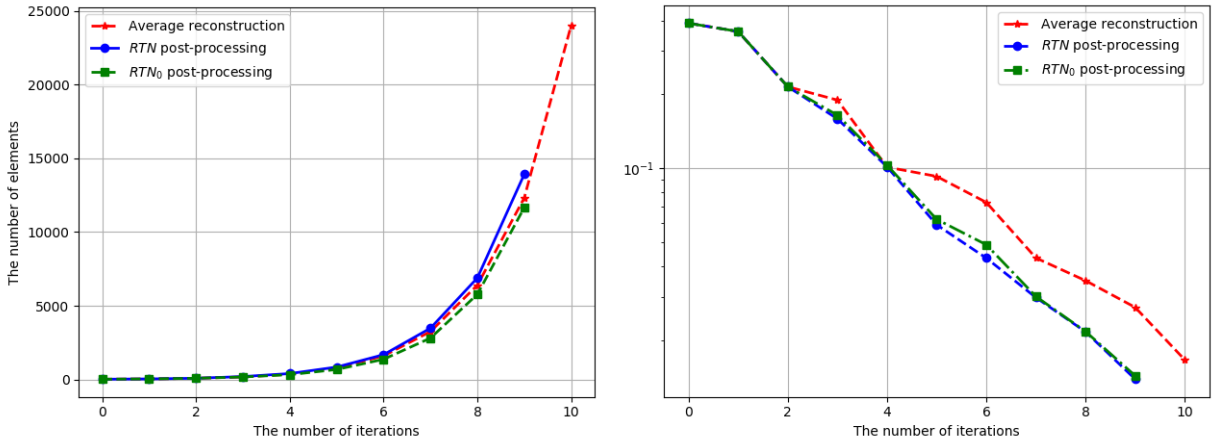
(a) Average reconstruction (b) RTN post-processing (c) RTN_0 post-processing

Figure 11: The numerical flux ϕ_h for different reconstruction methods.

Second, we modify the stopping criterion. Now, the stopping criterion is based on the L_2 error with respect to a reference solution. That is to say, the algorithm is stopped when

$$\frac{\|\phi_{ref} - \phi_h\|_{0,\Omega}}{\|\phi_{ref}\|_{0,\Omega}} \leq \epsilon_{Acr}, \tag{58}$$

where ϕ_{ref} is a reference solution computed on a fine mesh. We fix $\epsilon_{Acr} = 2 \times 10^{-2}$. Figure 12 shows that the RTN_0 post-processing and RTN post-processing give similar AMR strategies and that the resulting mesh have fewer elements than with the average reconstruction.



(a) Number of elements

(b) Relative L^2 error (58)

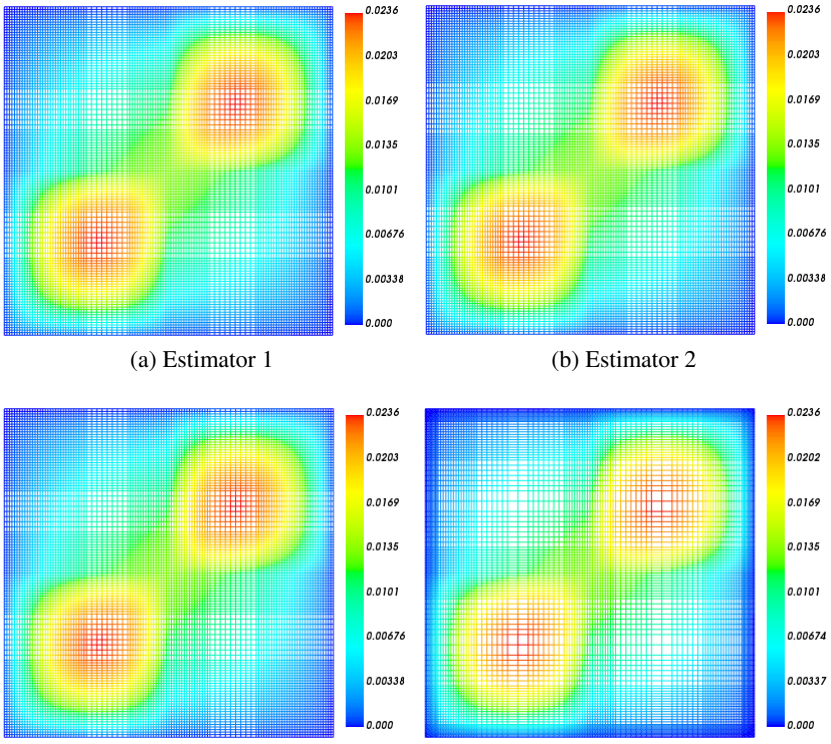
Figure 12: Evolution of the number of elements and the relative L^2 error (58) for different reconstruction methods.

5.6 Comparison of the error estimators

In this section, we perform the Dauge test case with a stopping criterion on the relative L^2 error (58) with respect to a reference solution at $\epsilon_{Acr} = 2 \times 10^{-2}$. To be convenient, let Estimator 1, Estimator 2, Estimator 3 and Estimator 4 respectively stand for the error estimator defined in [Theorem 8.4, p.117][18], (27), (42) and (53). For the sake of completeness, we recall here the different estimators for all $K \in \mathcal{T}_h$,

$$
\begin{aligned}
\eta_K^1 &= \left(\widehat{\eta}_{r,K}^2 + \eta_{f,K}^2 + 9\eta_{nc,K}^2\right)^{1/2} \quad \text{where } \widehat{\eta}_{r,K} \text{ is defined in Remark 6,} \\
\eta_K^2 &= \left(\eta_{r,K}^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2} + \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,K'}^2\right)^{1/2}, \\
\eta_K^3 &= \left(\eta_{r,K}^2 + \eta_{f,K}^2\right)^{1/2} + \eta_{nc,K}, \\
\eta_K^4 &= \tilde{\eta}_{nc,K} + \tilde{\eta}_{r,K}.
\end{aligned}
$$

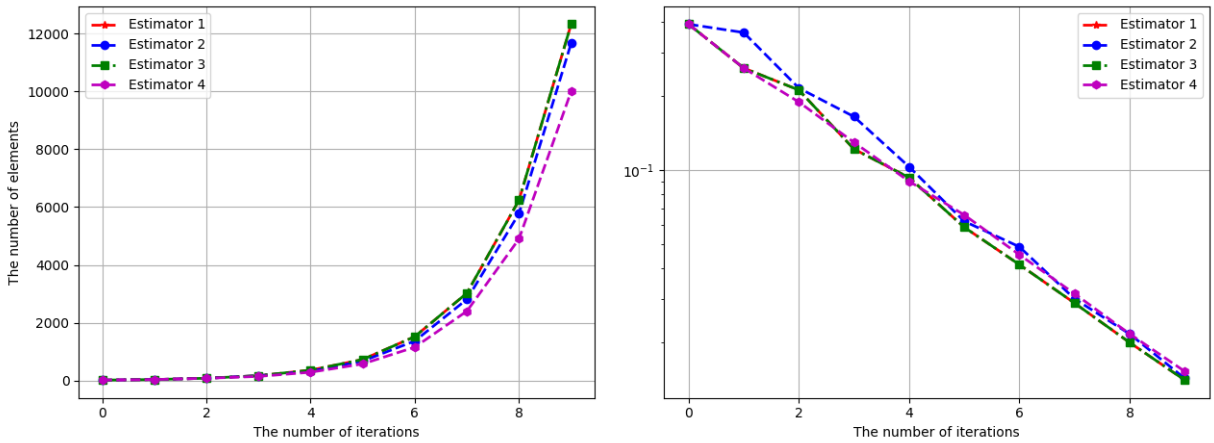
We apply the reconstruction associated to the RTN post-processing to each error estimators. As can be seen in Figure 13, we obtain similar meshes for the AMR strategies using Estimators 1, 2 and 3. On the other hand, there is more refinement near the boundary for Estimator 4. The relative L^2 error on the neutron flux are similar for the different AMR strategies according to Figure 14.



(a) Estimator 1 (b) Estimator 2

(c) Estimator 3 (d) Estimator 4

Figure 13: The numerical flux ϕ_h for different error estimators.



(a) Number of elements (b) Relative L^2 error (58)

Figure 14: Evolution of the number of elements and the relative L^2 error (58) for different error estimators

6 Extension to the DD+L^2-jumps method

In this section, we extend the derivation of the *a posteriori* estimators to a domain decomposition method introduced in [11], namely the DD+L^2-jumps methods.

To this aim, let us consider a partition $\{\tilde{\Omega}_{\tilde{\imath}}\}_{1\leq\tilde{\imath}\leq\tilde{N}}$ of Ω which is independent from the physical partition $\{\Omega_i\}_{1\leq i\leq N}$ introduced in section 1. For a field v defined over Ω, we shall use the notation $v_{\tilde{\imath}} = v|_{\tilde{\Omega}_{\tilde{\imath}}}$, for $1 \leq \tilde{\imath} \leq \tilde{N}$. We denote by $\Gamma_{\tilde{\imath}\tilde{\jmath}}$ the interface between two subdomains $\tilde{\Omega}_{\tilde{\imath}}$ and $\tilde{\Omega}_{\tilde{\jmath}}$ for $\tilde{\imath} \neq \tilde{\jmath}$: if $\dim_H \left(\partial\tilde{\Omega}_{\tilde{\imath}} \cap \partial\tilde{\Omega}_{\tilde{\jmath}}\right) = d-1$, then $\Gamma_{\tilde{\imath}\tilde{\jmath}} = \text{int}(\partial\tilde{\Omega}_{\tilde{\imath}} \cap \partial\tilde{\Omega}_{\tilde{\jmath}})$; otherwise, $\Gamma_{\tilde{\imath}\tilde{\jmath}} = \emptyset$. We define the interface Γ_S by

$$\Gamma_S = \cup_{\tilde{\imath}=1}^{\tilde{N}} \cup_{\tilde{\jmath}=\tilde{\imath}+1}^{\tilde{N}} \overline{\Gamma_{\tilde{\imath}\tilde{\jmath}}}.$$

We also introduce the spaces

$$\begin{aligned}
\tilde{P}H_0^1(\Omega) &= \{\psi \in L^2(\Omega) | \psi_{\tilde{\imath}} \in H^1(\tilde{\Omega}_{\tilde{\imath}}), \psi|_{\partial\tilde{\Omega}_{\tilde{\imath}}\setminus\overline{\Gamma_S}} = 0, 1 \leq \tilde{\imath} \leq \tilde{N}\},\\
\tilde{\mathbf{P}}\mathbf{H}(\text{div}\,,\Omega) &= \{\mathbf{q} \in L^2(\Omega) \mid \mathbf{q}_{\tilde{\imath}} \in \mathbf{H}(\text{div}\,,\tilde{\Omega}_{\tilde{\imath}}), 1 \leq \tilde{\imath} \leq \tilde{N}\},\\
M &= \{\psi_S \in \prod_{\tilde{\imath}<\tilde{\jmath}} L^2(\Gamma_{\tilde{\imath}\tilde{\jmath}})\},\\
\tilde{\mathbf{Q}} &= \{\mathbf{q} \in \tilde{\mathbf{P}}\mathbf{H}(\text{div}\,,\Omega) \mid [\mathbf{q}\cdot\mathbf{n}] \in M\},\\
\mathtt{W} &= \tilde{\mathbf{Q}} \times L^2(\Omega) \times M,
\end{aligned}$$

where $[\mathbf{q}\cdot\mathbf{n}]$ stands for the global jump of the normal component on the interface and is defined by

$$[\mathbf{q}\cdot\mathbf{n}]|_{\Gamma_{\tilde{\imath}\tilde{\jmath}}} = \mathbf{q}_{\tilde{\imath}}\cdot\mathbf{n}_{\tilde{\imath}} + \mathbf{q}_{\tilde{\jmath}}\cdot\mathbf{n}_{\tilde{\jmath}}, \text{ for } 1 \leq \tilde{\imath} < \tilde{\jmath} \leq \tilde{N}.$$

We consider the following problem:

$$\left\{\begin{array}{l}
\text{Find } (\mathbf{p},\phi,\phi_s) \in \tilde{\mathbf{Q}} \times \tilde{P}H_0^1(\Omega) \times M \text{ such that}\\
-\mathbb{D}_{\tilde{\imath}}^{-1}\mathbf{p}_{\tilde{\imath}} - \mathbf{grad}\phi_{\tilde{\imath}} = 0 \quad \text{in } \tilde{\Omega}_{\tilde{\imath}}, \quad \text{for } 1 \leq \tilde{\imath} \leq \tilde{N},\\
\text{div}\,\mathbf{p}_{\tilde{\imath}} + \Sigma_{a,\tilde{\imath}}\phi_{\tilde{\imath}} = S_{f,\tilde{\imath}} \quad \text{in } \tilde{\Omega}_{\tilde{\imath}}, \quad \text{for } 1 \leq \tilde{\imath} \leq \tilde{N},\\
\phi_{\tilde{\imath}} = \phi_s \quad \text{on } \tilde{\Omega}_{\tilde{\imath}} \cap \Gamma_s, \quad \text{for } 1 \leq \tilde{\imath} \leq \tilde{N},\\
[\mathbf{p}\cdot\mathbf{n}] = 0 \quad \text{on } \Gamma_s.
\end{array}\right.$$

The variational formulation writes

$$\text{Find } \mathtt{u} = (\mathbf{p},\phi,\phi_S) \in \mathtt{W} \text{ such that for all } \mathtt{w} = (\mathbf{q},\psi,\psi_S) \in \mathtt{W}, \quad c_S(\mathtt{u},\mathtt{w}) = f(\mathtt{w}), \tag{59}$$

where

$$c_S(\mathtt{u},\mathtt{w}) = c((\mathbf{p},\phi),(\mathbf{q},\psi)) + \int_{\Gamma_S}[\mathbf{p}\cdot\mathbf{n}]\psi_S - \int_{\Gamma_S}[\mathbf{q}\cdot\mathbf{n}]\phi_S, \quad \text{and } f(\mathtt{w}) = (S_f,\psi)_{0,\Omega}.$$

Above, we extended the definition (7) of the bilinear form c to piecewise smooth fields. We do the same for the forms a, b and t. We introduce discrete, finite-dimensional, spaces indexed by h as follows: $Q_{\tilde{\imath},h} \subset \mathbf{H}(\text{div}\,,\tilde{\Omega}_{\tilde{\imath}})$ and $L_{\tilde{\imath},h} \subset L^2(\tilde{\Omega}_{\tilde{\imath}})$, for $1 \leq \tilde{\imath} \leq \tilde{N}$. We impose the following requirements for all $1 \leq \tilde{\imath} \leq \tilde{N}$:

- $\mathbf{q}_{\tilde{\imath},h}\cdot\mathbf{n} \in L^2(\partial\tilde{\Omega}_{\tilde{\imath}})$ for all $h > 0$, for all $\mathbf{q}_{\tilde{\imath},h} \in Q_{\tilde{\imath},h}$;
- $\text{div}\,\mathbf{Q}_{\tilde{\imath},h} \subset L_{\tilde{\imath},h}$ for all $h > 0$;
- $(\mathbf{Q}_{\tilde{\imath},h})_h$ and $(L_{\tilde{\imath},h})_h$ satisfy the approximability property (15) in $\tilde{\Omega}_{\tilde{\imath}}$.

We observe that, to build conforming discretizations, one uses meshes that are *conforming* with respect to the partition $\{\tilde{\Omega}_{\tilde{\imath}}\}_{1\leq\tilde{\imath}\leq\tilde{N}}$: for all h, for all $K \in \mathcal{T}_h$, there exists $\tilde{\imath}$ such that $int(K) \subset \tilde{\Omega}_{\tilde{\imath}}$: we denote $\tilde{\Omega}_{\tilde{\imath}}$ by $\tilde{\Omega}_K$. We then set

$$\tilde{\mathbf{Q}}_h = \prod_{\tilde{\imath}=1}^{\tilde{N}} \mathbf{Q}_{\tilde{\imath},h}, \quad \tilde{L}_h = \prod_{\tilde{\imath}=1}^{\tilde{N}} L_{\tilde{\imath},h}, \quad \mathtt{W}_h = \tilde{\mathbf{Q}}_h \times \tilde{L}_h \times M_h,$$

where $M_h \subset M$ is the discrete space of Lagrange multipliers. We assume that the space of piecewise constant fields is included in M_h. Following [11, Section 5], we introduce the discrete projection operators $(\Pi_{\tilde{\imath}})_{1\le\tilde{\imath}\le\tilde{N}}$ from the spaces of normal traces

$$T_{\tilde{\imath},h} = \{q_{\tilde{\imath},h} \in L^2(\partial\tilde{\Omega}_{\tilde{\imath}} \cap \Gamma_S) \mid \exists \mathbf{q}_{\tilde{\imath},h} \in \mathbf{Q}_{\tilde{\imath},h},\ q_{\tilde{\imath},h} = \mathbf{q}_{\tilde{\imath},h} \cdot \mathbf{n}_{\tilde{\imath}_{|\partial\tilde{\Omega}_{\tilde{\imath}} \cap \Gamma_S}}\},\ \text{for } 1 \le \tilde{\imath} \le \tilde{N},$$

to M_h, resp. the discrete projection operators $(\pi_{\tilde{\imath}})_{1\le\tilde{\imath}\le\tilde{N}}$ from M_h to $(T_{\tilde{\imath},h})_{1\le\tilde{\imath}\le\tilde{N}}$, which are defined by

$$\forall q_{\tilde{\imath},h} \in T_{\tilde{\imath},h}, \forall \psi_{S,h} \in M_h,$$
$$\begin{cases} \displaystyle\int_{\partial\tilde{\Omega}_{\tilde{\imath}} \cap \Gamma_S} (\Pi_{\tilde{\imath}}(q_{\tilde{\imath},h}) - q_{\tilde{\imath},h})\psi_{S,h} = 0 \\ \displaystyle\int_{\partial\tilde{\Omega}_{\tilde{\imath}} \cap \Gamma_S} (\pi_{\tilde{\imath}}(\psi_{S,h}) - \psi_{S,h}) q_{\tilde{\imath},h} = 0. \end{cases}$$

Next, let $\mathbf{p}_h \in \tilde{\mathbf{Q}}_h$. We define the *discrete jump* of the normal component of $\mathbf{p}_h$ on the interface $\Gamma_{\tilde{\imath}\tilde{\jmath}}$ as $[\mathbf{p}_h \cdot \mathbf{n}]_{h,\tilde{\imath}\tilde{\jmath}} := \Pi_{\tilde{\imath}}(\mathbf{p}_{\tilde{\imath},h} \cdot \mathbf{n}_{\tilde{\imath}|\Gamma_{\tilde{\imath}\tilde{\jmath}}}) + \Pi_{\tilde{\jmath}}(\mathbf{p}_{\tilde{\jmath},h} \cdot \mathbf{n}_{\tilde{\jmath}|\Gamma_{\tilde{\imath}\tilde{\jmath}}})$.

The discrete variational formulation associated to (59) writes

$$\text{Find } \mathtt{u}_h = (\mathbf{p}_h, \phi_h, \phi_{S,h}) \in \mathtt{W}_h \text{ such that for all } \mathtt{w}_h = (\mathbf{q}_h, \psi_h, \psi_{S,h}) \in \mathtt{W}_h, \quad c_S(\mathtt{u}_h, \mathtt{w}_h) = f(\mathtt{w}_h). \tag{60}$$

We define

$$\begin{aligned} d_S(\mathtt{u},\mathtt{w}) &= -a(\mathbf{p},\mathbf{q}) + t(\phi,\psi) \\ d_A(\mathtt{u},\mathtt{w}) &= b(\mathbf{p},\psi) - b(\mathbf{q},\phi) - \int_{\Gamma_S} [\mathbf{p}\cdot\mathbf{n}]\psi_S + \int_{\Gamma_S} [\mathbf{q}\cdot\mathbf{n}]\phi_S \\ d(\mathtt{u},\mathtt{w}) &= d_S(\mathtt{u},\mathtt{w}) + d_A(\mathtt{u},\mathtt{w}) = c_S(\mathtt{u},(-\mathbf{q},\psi,-\psi_S)). \end{aligned}$$

We define the following norm on $\mathtt{W}$, for all $\mathtt{u} \in \mathtt{W}$,

$$\begin{aligned} \|\mathtt{u}\|_S^2 &= d_S(\mathtt{u},\mathtt{u}) + \sum_{K\in\mathcal{T}_h} \|\Sigma_a^{-1/2}\mathrm{div}\,\mathbf{p}\|_{0,K}^2 + \sum_{F\in\Gamma_S} \|[\mathbf{p}\cdot\mathbf{n}]\|_{0,F}^2 + \|\phi_S\|_M^2 \\ &= (\mathbb{D}^{-1}\,\mathbf{p},\mathbf{p})_{0,\Omega} + (\Sigma_a\,\phi,\phi)_{0,\Omega} + \sum_{K\in\mathcal{T}_h} \|\Sigma_a^{-1/2}\mathrm{div}\,\mathbf{p}\|_{0,K}^2 + \sum_{F\in\Gamma_S} \|[\mathbf{p}\cdot\mathbf{n}]\|_{0,F}^2 + \|\phi_S\|_M^2, \end{aligned}$$

and the following $\mathtt{W}_K$-local norm, for all $\mathtt{u} \in \mathtt{W}$,

$$|\mathtt{u}|_{+,K} = \sup_{\mathtt{w}\in\mathtt{W}_K, \|\mathtt{w}\|_S\le 1} d(\mathtt{u},\mathtt{w}), \tag{61}$$

where

$$\mathtt{W}_K = \left\{\mathtt{u} = (\mathbf{p},\phi,\phi_S) \in \mathtt{W},\ \mathrm{Supp}(\phi) \subset K,\ \mathrm{Supp}(\phi_S) \subset \partial K \cap \Gamma_S,\ \mathrm{Supp}(\mathbf{p}) \subset \tilde{N}(K)\right\},$$

with $\tilde{N}_K := N(K) \cap \overline{\tilde{\Omega}_K}$. Since $\mathbf{p}$ is in $\tilde{\mathbf{P}}\mathbf{H}(\mathrm{div}\,,\Omega)$, only the elements K' of $N(K)$ that belong to $\overline{\tilde{\Omega}_K}$ have to be considered above.

Assumption 2. *We assume that there exists $\beta_h > 0$ such that for all $\mathbf{q}_h \in \tilde{\mathbf{Q}}_h$,*

$$\int_{\Gamma_S} [\mathbf{q}_h\cdot\mathbf{n}]_h[\mathbf{q}_h\cdot\mathbf{n}] \ge \beta_h \int_{\Gamma_S} [\mathbf{q}_h\cdot\mathbf{n}]^2, \tag{62}$$

and that there exists $\gamma_h > 0$ such that for all $\psi_{S,h} \in M_h$,

$$\sum_{\tilde{\imath}=1}^{\tilde{N}} \sum_{\tilde{\jmath}=\tilde{\imath}+1}^{\tilde{N}} \int_{\Gamma_{\tilde{\imath}\tilde{\jmath}}} (\pi_{\tilde{\imath}}(\psi_{S,h})^2 + \pi_{\tilde{\jmath}}(\psi_{S,h})^2) \ge \gamma_h \|\psi_{S,h}\|_M^2. \tag{63}$$

It is proven in [11, Section 5.1] that, under Assumption 2, the discrete problem (60) is well-posed, and also that the discrete solution fulfills $[\mathbf{p}_h \cdot \mathbf{n}] = 0$.

Lemma 3. *Let* $\mathtt{u}$ *and* $\mathtt{u}_h$ *be respectively the solution to* (59) *and* (60). *Let* $\tilde{\mathtt{u}}_h = (\mathbf{p}_h, \tilde{\phi}_h, \tilde{\phi}_{S,h})$ *be a reconstruction of* $\mathtt{u}_h$ *in* $\tilde{\mathbf{Q}}_h \times V \times M$ *such that* $\tilde{\phi}_{S,h} = \tilde{\phi}_h$ *on* Γ_S. *We have for all* $\mathtt{w} \in \mathtt{W}$,

$$d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}) = (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h, \mathbf{q})_{0,\Omega}. \tag{64}$$

Proof. Let $\tilde{\mathtt{u}}_h$ be a reconstruction of $\mathtt{u}_h$ in $\tilde{\mathbf{Q}} \times V \times M$. We have for all $\mathtt{w} \in \mathtt{W}$,

$$\begin{aligned}
&d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}) \\
&= (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h, \mathbf{q})_{0,\Omega} + (\tilde{\phi}_h, \operatorname{div} \mathbf{q})_{0,\Omega} - \int_{\Gamma_S} [\mathbf{q} \cdot \mathbf{n}] \tilde{\phi}_{S,h} \\
&= (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h, \mathbf{q})_{0,\Omega} + \int_{\Gamma_S} [\mathbf{q} \cdot \mathbf{n}] (\tilde{\phi}_h - \tilde{\phi}_{S,h}),
\end{aligned}$$

where we integrate by parts using [11, Theorem 5].

Noticing that

$$\tilde{\phi}_{S,h} = \tilde{\phi}_h \text{ on } \Gamma_S,$$

we obtain the desired result. □

Theorem 9. *We suppose that Assumption 2 holds. Let* $\mathtt{u}$ *and* $\mathtt{u}_h$ *be respectively the solution to* (59) *and* (60). *Let* $\tilde{\mathtt{u}}_h = (\mathbf{p}_h, \tilde{\phi}_h, \tilde{\phi}_{S,h})$ *be a reconstruction of* $\mathtt{u}_h$ *in* $\tilde{\mathbf{Q}}_h \times V \times M$ *such that* $\tilde{\phi}_{S,h} = \tilde{\phi}_h$ *on* Γ_S. *For any* $K \in \mathcal{T}_h$, *we define the residual estimator* $\eta_{r,K}$ *as in (23), the flux estimator* $\eta_{f,K}$ *as in (24) and the non-conforming estimator* $\eta_{nc,K}$ *as in (25) ; finally, for any* $F \in \Gamma_S$, *we define the interface continuity estimator by*

$$\eta_{ic,F} = \|\tilde{\phi}_{S,h} - \phi_{S,h}\|_{0,F}. \tag{65}$$

Then, it stands for all $K \in \mathcal{T}_h$

$$|\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \leq \left(\eta_{r,K}^2 + \sum_{K' \in \tilde{N}(K)} \eta_{f,K'}^2 \right)^{1/2}, \tag{66}$$

$$\begin{aligned}
|\mathtt{u} - \mathtt{u}_h|_{+,K} \leq &\left(\eta_{r,K}^2 + \sum_{K' \in \tilde{N}(K)} \eta_{f,K'}^2 \right)^{1/2} \\
&+ \left(\eta_{nc,K}^2 + \sum_{K' \in \tilde{N}(K)} \eta_{nc,K'}^2 + \sum_{F \in \Gamma_{S,K}} \eta_{ic,F}^2 \right)^{1/2},
\end{aligned} \tag{67}$$

where $\Gamma_{S,K}$ *is the set of facets associated to* $\tilde{N}(K)$ *belonging to* Γ_S.

Proof. Using the triangle inequality, we obtain for all $K \in \mathcal{T}_h$

$$|\mathtt{u} - \mathtt{u}_h|_{+,K} \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} + |\tilde{\mathtt{u}}_h - \mathtt{u}_h|_{+,K}. \tag{68}$$

We observe that Equation (64) in lemma 3 may be written as

$$d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}) = \sum_{K \in \mathcal{T}_h} (S_f + \operatorname{div} \mathbf{p}_h - \Sigma_a \tilde{\phi}_h, \psi)_{0,K} - (\mathbb{D}^{-1}\mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h, \mathbf{q})_{0,K}.$$

Let $K \in \mathcal{T}_h$ and $\mathtt{w} = (\mathbf{q}, \psi, \psi_S) \in \mathtt{W}$ be such that $\mathrm{Supp}(\psi) \subset K$ and $\mathrm{Supp}(\mathbf{q}) \subset \tilde{N}(K)$. Using Cauchy-Schwarz inequalities, we obtain successively

$$d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}) \leq \eta_{r,K} \|\Sigma_a^{1/2}\psi\|_{0,K} + \sum_{K' \in \tilde{N}(K)} \eta_{f,K'} \|\mathbb{D}^{-1/2}\mathbf{q}\|_{0,K}$$
$$\leq \left(\eta_{r,K}^2 + \sum_{K' \in \tilde{N}(K)} \eta_{f,K'}^2 \right)^{1/2} \|\mathtt{w}\|_S.$$

We conclude from (61) that (66) holds.
Next, we bound the second term of the right-hand side of (68). Let $\mathtt{w} \in \mathtt{W}$, we look for an upper bound to

$$\begin{aligned} d(\tilde{\mathtt{u}}_h - \mathtt{u}_h, \mathtt{w}) &= d_S(\tilde{\mathtt{u}}_h - \mathtt{u}_h, \mathtt{w}) + d_A(\tilde{\mathtt{u}}_h - \mathtt{u}_h, \mathtt{w}) \\ &\leq (\Sigma_a(\tilde{\phi}_h - \phi_h), \psi)_{0,\Omega} - (\tilde{\phi}_h - \phi_h, \mathrm{div}\,\mathbf{q})_{0,\Omega} + \int_{\Gamma_S} [\mathbf{q}\cdot\mathbf{n}](\tilde{\phi}_{S,h} - \phi_{S,h}) \\ &\leq \eta_{nc,K}\|\Sigma_a^{1/2}\psi\|_{0,K} + \sum_{K' \in \tilde{N}(K)} \eta_{nc,K'}\|\Sigma_a^{-1/2}\mathrm{div}\,\mathbf{q}\|_{0,K'} + \sum_{F \in \Gamma_{S,K}} \eta_{ic,F}\|[\mathbf{q}\cdot\mathbf{n}]\|_{0,F} \\ &\leq \left(\eta_{nc,K}^2 + \sum_{K' \in \tilde{N}(K)} \eta_{nc,K'}^2 + \sum_{F \in \Gamma_{S,K}} \eta_{ic,F}^2 \right)^{1/2} \|\mathtt{w}\|_S, \end{aligned} \tag{69}$$

where we used Cauchy-Schwarz in the last two lines. Collecting (67), (68) and (69), we get the estimate. □

Remark 8. *We give here an example of reconstruction based on the averaging operator defined in Section 4.1.1 in the case where $\tilde{L}_h = \mathbb{P}_n(\mathcal{T}_h)$. Let $\tilde{\mathcal{T}}_h$ be a V-conforming mesh, that is without hanging nodes, such that $\tilde{\mathcal{T}}_h$ is a refinement of $\mathcal{T}_h$. We define here $\tilde{\mathcal{V}}_h^{n+1}$ as the set of nodes of $V \cap \mathbb{P}_{n+1}(\tilde{\mathcal{T}}_h)$, and $\tilde{\mathcal{T}}_a$ is the set of simplices sharing a node $a \in \tilde{\mathcal{V}}_h^{n+1}$. For a node a on $int(\Gamma_S) \cap \tilde{\mathcal{V}}_h^{n+1}$, we define $\tilde{\mathcal{E}}_a$ has the set of interface facets sharing a. Given $a \in \tilde{\mathcal{V}}_h^{n+1}$, we distinguish three cases:*

1. *If $a \in \tilde{\Omega}_j$, $\tilde{\phi}_h(a) = \dfrac{1}{|\tilde{\mathcal{T}}_a|} \displaystyle\sum_{K \in \tilde{\mathcal{T}}_a} \phi_{h|K}(a)$;*
2. *If $a \in int(\Gamma_S)$, $\tilde{\phi}_h(a) = \dfrac{1}{|\tilde{\mathcal{E}}_a|} \displaystyle\sum_{E \in \tilde{\mathcal{E}}_a} \phi_{S,h|E}(a)$;*
3. *If $a \in \partial\Omega \cap \partial\tilde{\Omega}_j$, $\tilde{\phi}_h(a) = 0$.*

Interestingly, this approach share some similarities with the construction of the discrete space of the Lagrange multipliers M_h detailed in [11, Section 5.2].

Theorem 10 (local efficiency of the *a posteriori* error estimators)**.** *Let $K \in \mathcal{T}_h$ and let $\eta_{r,K}$ and $\eta_{f,K}$ be the residual estimators respectively given by* (23) *and* (24)*. Under Assumptions 1 and 2, the following estimates holds true*

$$\eta_{r,K} \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K}\mathtt{c}, \tag{70}$$
$$\eta_{f,K} \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K}\left\{ \mathtt{c}^2\frac{\mathbb{D}_K^{max}}{\mathbb{D}_K^{min}} + \mathtt{C}^2\frac{\mathbb{D}_K^{max}}{h_K^2\Sigma_{a,K}} \right\}^{1/2}, \tag{71}$$

where $\mathtt{c}$ and $\mathtt{C}$ are constants which depend only on the polynomial degree k of S_f, d, and the shape-regularity parameter κ_K, and with the $|\cdot|_{+,K}$ norm defined as in (61).

Proof. The proof follows that given in [28, Lemma 7.6]. Let ψ_K be the bubble function on K, and let us denote $\psi_r = (S_f - \operatorname{div} \mathbf{p}_h - \Sigma_a \widetilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. We recall that, since ψ_r is a polynomial in K by Assumption 1, the equivalence of norms on finite-dimensional spaces yields the bounds (33) and (34), with the constant c there depending only on the polynomial degree k of S_f, d, and κ_K. Let $\mathtt{w}_{r,K} = (0, \psi_K \psi_r, 0) \in \mathtt{W}$, we immediately have (cf. the proof of lemma 3)

$$d(\mathtt{u} - \tilde{\mathtt{u}_h}, \mathtt{w}_{r,K}) = (\psi_r, \psi_K \psi_r)_{0,K}.$$

Then, by definition (61) of the $|\cdot|_{+,K}$ norm,

$$d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}_{r,K}) \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \|\mathtt{w}_{r,K}\|_S \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \|\Sigma_a^{1/2} \psi_K \psi_r\|_{0,K}. \tag{72}$$

Combining (33), (34) and (72), one comes to

$$c \|\psi_r\|_{0,K}^2 \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \|\psi_r\|_{0,K} (\Sigma_{a,K})^{1/2}.$$

Using the definition of $\eta_{r,K}$ by (23) concludes the proof of (70).
We now proceed similarly for the second estimate. Let us denote $\mathbf{q}_f = (\mathbb{D}^{-1} \mathbf{p}_h + \mathbf{grad}\, \tilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that $\mathbf{q}_f$ is a polynomial in K by Assumption 1, hence the equivalence of norms on finite-dimensional spaces gives the bounds (36) and (37), while the inverse inequality (38) holds. The constants c and C there depend only on the polynomial degree k of S_f, d, and κ_K. Let $\mathtt{w}_{f,K} = (\psi_K \mathbf{q}_f, 0, 0) \in \mathtt{W}$, we immediately have (cf. the proof of lemma 1)

$$-d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}_{f,K}) = (\mathbf{q}_f, \psi_K \mathbf{q}_f)_{0,K},$$

and, moreover,

$$\begin{aligned} -d(\mathtt{u} - \tilde{\mathtt{u}}_h, \mathtt{w}_{f,K}) &\leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \|\mathtt{w}_{f,K}\|_S \\ &\leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \Big(\|\mathbb{D}^{-1/2} \psi_K \mathbf{q}_f\|_{0,K}^2 + \|\Sigma_a^{-1/2} \operatorname{div}(\psi_K \mathbf{q}_f)\|_{0,K}^2 \Big)^{1/2} \\ &\leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \big((\mathbb{D}_K^{min})^{-1} + C(h_K^2 \Sigma_{a,K})^{-1} \big)^{1/2} \|\psi_K \mathbf{q}_f\|_{0,K}, \end{aligned} \tag{73}$$

where we used (38) in the last line. Combining (36), (37) and (73), one comes to

$$c \|\mathbf{q}_f\|_{0,K}^2 \leq |\mathtt{u} - \tilde{\mathtt{u}}_h|_{+,K} \|\mathbf{q}_f\|_{0,K} \big((\mathbb{D}_K^{min})^{-1} + C(h_K^2 \Sigma_{a,K})^{-1} \big)^{1/2}.$$

Considering the definition of $\eta_{f,K}$ by (24) concludes the proof. □

7 Extension to the multigroup diffusion equations

The neutron flux density in the reactor core is determined by solving the transport equation which depends on seven variables: the space (3), the direction (2), the energy or the speed (1), and the time (1). It physically states the balance between the emission of neutrons by fission and the absorption, scattering, and leakage at the boundaries of neutrons. The most common discretization of the energy variable is the multigroup approximation where the energy domain is divided into subintervals called energy groups. In practice, the neutron flux density is usually modeled by the multigroup diffusion equations [15, Chapter 7] at the reactor core scale.

In many practical situations, only a steady-state solution is relevant and it requires to solve a generalized eigenvalue problem. In the companion paper [9], we perform adaptive mesh refinement for the multigroup diffusion

case on this so-called criticality problem. We present in this section some theoretical results underlying this approach on a source problem.

For $G \geq 2$, we let $\mathcal{I}_G := \{1, ..., G\}$ be the set of energy groups. Given a function space W, we denote by $\underline{W}$ the product space W^G. We extend the notation $(\cdot,\cdot)_{0,\mathcal{O}}$ (respectively $\|\cdot\|_{0,\mathcal{O}}$) for the $\underline{L^2}(\mathcal{O})$ and $\underline{\mathbf{L}^2}(\mathcal{O})$ scalar products (resp. norms). Let $\mathbf{q} := (q_x^g)_{x=1,d}^{g=1,G}$, $\mathbf{q}^g = (q_x^g)_{x=1,d} \in \mathbb{R}^d$ for $1 \leq g \leq G$, and $\operatorname{div} \mathbf{q} = (\operatorname{div}_{\mathbf{x}} \mathbf{q}^g)^{g=1,G} \in \mathbb{R}^G$.

Let $\mathbb{T}_e \in (\mathbb{R})^{G\times G}$ be the even removal matrix. It is a full matrix such that

$$\forall (g,g') \in \mathcal{I}_G \times \mathcal{I}_G, \quad (\mathbb{T}_e)_{g,g'} = \begin{cases} \Sigma_t^g - \Sigma_{s,0}^{g\to g} & \text{if } g = g', \\ -\Sigma_{s,0}^{g'\to g} & \text{if } g \neq g', \end{cases}$$

where $\Sigma_{s,0}^{g'\to g}$ are the Legendre moments of order 0 of the macroscopic scattering cross sections from energy group g' to energy group g and the coefficient Σ_t^g is the macroscopic total cross section of energy group g [15, Part 2, Chapter IV, Section A, p.124-127].
We denote $\mathbb{T}_o \in (\mathbb{R})^{G\times G}$ the odd removal matrix. It is a diagonal matrix such that

$$\forall g \in \mathcal{I}_G, \quad (\mathbb{T}_o)_{g,g} = 1/D^g.$$

where D^g is the diffusion coefficient of energy group g. The coefficients of the matrices $\mathbb{T}_{e,o}$ are supposed to be such that:

$$\begin{cases} (0) & \forall g, g' \in \mathcal{I}_G,\ (D^g, \Sigma_{r,0}^g, \Sigma_{s,0}^{g'\to g}, \underline{\nu\Sigma_f^g}) \in \mathcal{P}W^{1,\infty}(\Omega) \times \mathcal{P}W^{1,\infty}(\Omega) \times L^\infty(\Omega) \times L^\infty(\Omega)\,, \\ (i) & \exists\, (D)_*, (D)^* > 0,\ \forall\, g \in \mathcal{I}_G,\ (D)_* \leq D^g \leq (D)^* \text{ a.e. in } \Omega\,, \\ (ii) & \exists\, (\Sigma_{r,0})_*, (\Sigma_{r,0})^* > 0,\ \forall\, g \in \mathcal{I}_G,\ (\Sigma_{r,0})_* \leq \Sigma_{r,0}^g \leq (\Sigma_{r,0})^* \text{ a.e. in } \Omega\,, \\ (iii) & \exists \epsilon \in (0, (G-1)^{-1}),\ \forall\, g, g' \in \mathcal{I}_G, g' \neq g,\ |\Sigma_{s,0}^{g\to g'}| \leq \epsilon \Sigma_{r,0}^g \text{ a.e. in } \Omega. \end{cases} \tag{74}$$

As a consequence of (74), the matrix $\mathbb{T}_e$ is strictly diagonally dominant, so it is invertible and so is the diagonal matrix $\mathbb{T}_o$.

Remark 9. *Hypothesis* (74)$-(iii)$ *models accurately the core of a pressurized water reactor and, in this case,* ϵ *is a small fraction of* $(G-1)^{-1}$*. So on every row of* $\mathbb{T}_e$*, the off-diagonal entries are much smaller than the diagonal entries. Hence, the inverse of* $\mathbb{T}_e$ *is well-approximated by the inverse of its diagonal.*

In the multigroup case, the bilinear forms read:

$$\underline{a} : \begin{cases} \underline{Q} \times \underline{Q} & \to & \mathbb{R} \\ (\mathbf{p}, \mathbf{q}) & \mapsto & (-\mathbb{T}_o\, \mathbf{p}, \mathbf{q})_{0,\Omega} \end{cases}; \tag{75}$$

$$\underline{b} : \begin{cases} \underline{Q} \times \underline{L} & \to & \mathbb{R} \\ (\mathbf{q}, \psi) & \mapsto & (\psi, \operatorname{div} \mathbf{q})_{0,\Omega} \end{cases}; \tag{76}$$

$$\underline{t} : \begin{cases} \underline{L} \times \underline{L} & \to & \mathbb{R} \\ (\phi, \psi) & \mapsto & (\mathbb{T}_e\, \phi, \psi)_{0,\Omega} \end{cases}; \tag{77}$$

and:

$$\underline{c} : \begin{cases} \underline{\mathbf{X}} \times \underline{\mathbf{X}} & \to & \mathbb{R} \\ (\zeta, \xi) & \mapsto & \underline{a}(\mathbf{p}, \mathbf{q}) + \underline{b}(\mathbf{q}, \phi) + \underline{b}(\mathbf{p}, \psi) + \underline{t}(\phi, \psi) \end{cases}, \tag{78}$$

$$\underline{f} : \begin{cases} \underline{Q} \times \underline{L} & \to & \mathbb{R} \\ (\mathbf{q}, \psi) & \mapsto & (S_f, \psi)_{0,\Omega} \end{cases}, \tag{79}$$

where $S_f \in \underline{L}$.

We may write the variational formulation as:

$$\text{Find } \zeta \in \underline{\mathbf{X}} \text{ such that for all } \xi \in \underline{\mathbf{X}},\ \underline{c}(\zeta, \xi) = f(\xi). \tag{80}$$

We discretize this variational formulation using the same approach as before, see section 3.2, here applied to each group. The associated discrete problem reads,

$$\text{Find } \zeta_h \in \underline{\mathbf{X}}_h \text{ such that for all } \xi_h \in \underline{\mathbf{X}}_h,\ \underline{c}(\zeta_h, \xi_h) = \underline{f}(\xi_h). \tag{81}$$

We define

$$\begin{aligned}
&\underline{d}_S(\zeta, \xi) = -\underline{a}(\mathbf{p}, \mathbf{q}) + \underline{t}(\phi, \psi) \\
&\underline{d}_A(\zeta, \xi) = \underline{b}(\mathbf{p}, \psi) - \underline{b}(\mathbf{q}, \phi) \\
&\underline{d}(\zeta, \xi) = \underline{d}_S(\zeta, \xi) + \underline{d}_A(\zeta, \xi) = \underline{c}(\zeta, (-\mathbf{q}, \psi)).
\end{aligned}$$

We define the following norm on $\underline{\mathbf{X}}$, for all $\zeta \in \underline{\mathbf{X}}$,

$$\|\zeta\|_{S,MG}^2 = \sum_{K \in \mathcal{T}_h} \|d\mathbb{T}_o^{1/2}\mathbf{p}\|_{0,K}^2 + \|d\mathbb{T}_e^{1/2}\phi\|_{0,K}^2 + \sum_{K \in \mathcal{T}_h} \|d\mathbb{T}_e^{-1/2}(\operatorname{div} \mathbf{p})\|_{0,K}^2,$$

where $d\mathbb{T}_{e,o}^{1/2}$, resp. $d\mathbb{T}_{e,o}^{-1/2}$, is the square root of the diagonal part of $\mathbb{T}_{e,o}$, resp. the inverse of the diagonal part of $\mathbb{T}_{e,o}$. For the multigroup diffusion model, we introduce the following $\underline{\mathbf{X}}_K$-local norm, for all $\zeta \in \underline{\mathbf{X}}$,

$$|\zeta|_{+,K} = \sup_{\xi \in \underline{\mathbf{X}}_K, \|\xi\|_{S,MG} \le 1} \underline{d}(\zeta, \xi), \tag{82}$$

with

$$\underline{\mathbf{X}}_K = \{\zeta = (\mathbf{p}, \phi) \in \underline{\mathbf{X}},\ \operatorname{Supp}(\phi) \subset K,\ \operatorname{Supp}(\mathbf{p}) \subset N(K)\}.$$

Observe that the norm $\|\cdot\|_{S,MG}$ measures elements of $\underline{\mathbf{X}}$ in $\underline{\mathbf{H}}(\operatorname{div}, \mathcal{T}_h) \times \underline{L}$ norm. This corresponds precisely to the energy norm (cf. [18, Chapter 8]). In this section, we derive *a posteriori* estimates following the approach of section 4.2.1.

Remark 10. *It is also possible to extend the approach developed in section 4.2.2 to the multigroup diffusion model.*

Lemma 4. *Let ζ and ζ_h be respectively the solution to* (80) *and* (81). *Let $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h) \in \underline{\mathbf{Q}}_h \times \underline{V}$ be a reconstruction of ζ_h. We have,*

$$\forall \xi \in \underline{\mathbf{X}}, \quad \underline{d}(\zeta - \tilde{\zeta}_h, \xi) = (S_f - \operatorname{div} \mathbf{p}_h - \mathbb{T}_e\tilde{\phi}_h, \psi)_{0,\Omega} - (\mathbb{T}_o\mathbf{p}_h + \mathbf{grad}\ \tilde{\phi}_h, \mathbf{q})_{0,\Omega}. \tag{83}$$

We skip the proof which is identical to the proof of lemma 1. We are now in position to state the following theorem.

Theorem 11. *Let ζ and ζ_h be respectively the solution to* (80) *and* (81). *Let $\tilde{\zeta}_h = (\mathbf{p}_h, \tilde{\phi}_h)$ be a reconstruction of ζ_h in $\underline{\mathbf{Q}}_h \times \underline{V}$. For any $K \in \mathcal{T}_h$, we define the residual estimators*

$$\eta_{r,K} = \|d\mathbb{T}_e^{-1/2}(S_f - \operatorname{div} \mathbf{p}_h - \mathbb{T}_e\tilde{\phi}_h)\|_{0,K}, \tag{84}$$

the flux estimator

$$\eta_{f,K} = \|d\mathbb{T}_o^{-1/2}(\mathbb{T}_o\mathbf{p}_h + \mathbf{grad}\tilde{\phi}_h)\|_{0,K}, \tag{85}$$

and the two non-conforming estimators

$$\eta_{nc,K} = \|d\mathbb{T}_e^{-1/2}\mathbb{T}_e(\tilde{\phi}_h - \phi_h)\|_{0,K}, \quad \eta_{nc,\star,K} = \|d\mathbb{T}_e^{1/2}(\tilde{\phi}_h - \phi_h)\|_{0,K}. \tag{86}$$

Then it stands for all $K \in \mathcal{T}_h$,

$$|\zeta - \tilde{\zeta}_h|_{+,K} \leq \left(\eta_{r,K}^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2}, \tag{87}$$

$$|\zeta - \zeta_h|_{+,K} \leq \left(\eta_{r,K}^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2} + \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,\star,K'}^2\right)^{1/2}. \tag{88}$$

Proof. Using the triangle inequality, we obtain for any $K \in \mathcal{T}_h$

$$|\zeta - \zeta_h|_+ \leq |\zeta - \tilde{\zeta}_h|_+ + |\tilde{\zeta}_h - \zeta_h|_+. \tag{89}$$

According to Lemma 4, (83) holds for all $\xi \in \underline{\mathbf{X}}$. Let $K \in \mathcal{T}_h$ and $\xi = (\mathbf{q}, \psi)$ be such that $\text{Supp}(\psi) \subset K$, and $\text{Supp}(\mathbf{q}) \subset N(K)$. Applying Cauchy-Schwarz inequalities, we get successively

$$\begin{aligned}
\underline{d}(\zeta - \tilde{\zeta}_h, \xi) &\leq \eta_{r,K}\|d\mathbb{T}_e^{1/2}\psi\|_{0,K} + \sum_{K' \in N(K)} \eta_{f,K'}\|\mathbb{T}_o^{1/2}\mathbf{q}\|_{0,K'} \\
&\leq \left(\eta_{r,K}^2 + \sum_{K' \in N(K)} \eta_{f,K'}^2\right)^{1/2} \|\xi\|_{S,MG}.
\end{aligned}$$

We infer from the definition (82) of the $|\cdot|_{+,K}$ norm that (87) holds.
Now, we want to bound the second term of the right-hand side of (89). We bound

$$\begin{aligned}
\underline{d}(\tilde{\zeta}_h - \zeta_h, \xi) &= \underline{d}_S(\tilde{\zeta}_h - \zeta_h, \xi) + \underline{d}_A(\tilde{\zeta}_h - \zeta_h, \xi) \\
&\leq (\mathbb{T}_e(\phi_h - \tilde{\phi}_h), \psi)_{0,\Omega} - (\operatorname{div}\mathbf{q}, \tilde{\phi}_h - \phi_h)_{0,\Omega} \\
&\leq \eta_{nc,K}\|d\mathbb{T}_e^{1/2}\psi\|_{0,K} + \sum_{K' \in N(K)} \eta_{nc,\star,K'}\|d\mathbb{T}_e^{-1/2}(\operatorname{div}\mathbf{q})\|_{0,K'} \\
&\leq \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,\star,K'}^2\right)^{1/2} \|\xi\|_{S,MG},
\end{aligned}$$

where we used Cauchy-Schwarz in the last two lines. Hence, $|\tilde{\zeta}_h - \zeta_h|_{+,K} \leq \left(\eta_{nc,K}^2 + \sum_{K' \in N(K)} \eta_{nc,\star,K'}^2\right)^{1/2}$.
Using the triangle inequality (89), we get the desired estimate (88). □

Remark 11. *Using the same arguments as in the proof of theorem 11, we can show under the same assumptions that*

$$|\zeta - \tilde{\zeta}_h|_+ \leq \left(\sum_{K \in \mathcal{T}_h} \eta_{r,K}^2 + \eta_{f,K}^2\right)^{1/2},$$

$$|\zeta - \zeta_h|_+ \leq \left(\sum_{K \in \mathcal{T}_h} \eta_{r,K}^2 + \eta_{f,K}^2\right)^{1/2} + \left(\sum_{K \in \mathcal{T}_h} \eta_{nc,K}^2 + \eta_{nc,\star,K}^2\right)^{1/2},$$

where the global $|\cdot|_+$ norm is defined for all $\zeta \in \underline{\mathbf{X}}$ by,

$$|\zeta|_+ = \sup_{\xi \in \underline{\mathbf{X}}, \|\xi\|_{S,MG} \leq 1} \underline{d}(\zeta, \xi).$$

These estimates are similar to the one stated in the companion paper [9, Estimate (6)] where we use a slightly different definition of the flux estimator.

In order to state the next theorem, we will use the following assumption.

Assumption 3. *The matrices $\mathbb{T}_{e,o}$ are piecewise constant on $\mathcal{T}_h$ and $S_f \in L_h$.*

Under Assumption 3, one may define

$$d\mathbb{T}_{e,K}^{max} = [\max_{g \in \mathcal{I}_G} (\mathbb{T}_e)_{g,g}]_{|K},\ d\mathbb{T}_{e,K}^{min} = [\min_{g \in \mathcal{I}_G} (\mathbb{T}_e)_{g,g}]_{|K};\ d\mathbb{T}_{o,K}^{max} = [\max_{g \in \mathcal{I}_G} (\mathbb{T}_o)_{g,g}]_{|K},\ d\mathbb{T}_{o,K}^{min} = [\min_{g \in \mathcal{I}_G} (\mathbb{T}_o)_{g,g}]_{|K}.$$

Theorem 12 (local efficiency of the *a posteriori* error estimators)**.** *Let $K \in \mathcal{T}_h$ and let $\eta_{r,K}$ and $\eta_{f,K}$ be the residual estimators respectively given by* (84) *and* (85)*. Under Assumption 3, the following estimates holds true*

$$\eta_{r,K} \leq |\zeta - \tilde{\zeta}_h|_{+,K} \mathtt{c} \left(\frac{d\mathbb{T}_{e,K}^{max}}{d\mathbb{T}_{e,K}^{min}} \right)^{1/2}, \tag{90}$$

$$\eta_{f,K} \leq |\zeta - \tilde{\zeta}_h|_{+,K} \left\{ \mathtt{c}^2 \frac{d\mathbb{T}_{o,K}^{max}}{d\mathbb{T}_{o,K}^{min}} + \mathtt{C}^2 \frac{(d\mathbb{T}_{o,K}^{min})^{-1}}{h_K^2 d\mathbb{T}_{e,K}^{min}} \right\}^{1/2}, \tag{91}$$

where $\mathtt{c}$ and $\mathtt{C}$ are constants which depend only on the polynomial degree k of S_f, d, and the shape-regularity parameter κ_K, with the $|\cdot|_{+,K}$ norm defined as in (82).

Proof. The proof follows that given in [28, Lemma 7.6]. Let ψ_K be the bubble function on K, and let us denote $\psi_r = d\mathbb{T}_e^{-1/2}(S_f - \operatorname{div} \mathbf{p}_h - \mathbb{T}_e \tilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Again, because ψ_r is a polynomial in K by Assumption 3, the equivalence of norms on finite-dimensional spaces gives the bounds equivalent to (33) and (34), with the constant c there depending only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{r,K} = (0, \psi_K \psi_r) \in \underline{\mathbf{X}}$, we immediately have (cf. lemma 4)

$$\underline{d}(\zeta - \tilde{\zeta}_h, \xi_{r,K}) = (d\mathbb{T}_e^{1/2} \psi_r, \psi_K \psi_r)_{0,K},$$

so that

$$\underline{d}(\zeta - \tilde{\zeta}_h, \xi_{r,K}) \leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\xi_{r,K}\|_{S,MG} \leq |\zeta - \tilde{\zeta}_h|_{+,K} \|d\mathbb{T}_e^{1/2}{}_{|K} \psi_K \psi_r\|_{0,K}. \tag{92}$$

Combining (33), (34) and (92), one comes to

$$c(d\mathbb{T}_{e,K}^{min})^{1/2} \|\psi_r\|_{0,K}^2 \leq |\zeta - \tilde{\zeta}_h|_{+,K} \|\psi_r\|_{0,K} (d\mathbb{T}_{e,K}^{max})^{1/2}.$$

Using the definition of $\eta_{r,K}$ by (84) concludes the proof of (90).

We now proceed similarly for the second estimate. Let us denote $\mathbf{q}_f = d\mathbb{T}_o^{-1/2}(\mathbb{T}_e \mathbf{p}_h + \mathbf{grad}\ \tilde{\phi}_h)$ on a given $K \in \mathcal{T}_h$. Note that $\mathbf{q}_f$ is a polynomial in K by Assumption 3. Then the equivalence of norms on finite-dimensional spaces gives the bounds analogous to (36) and (37), while the corresponding inverse inequality (38) holds. The constants c and C there depend only on the polynomial degree k of S_f, d, and κ_K. Let $\xi_{f,K} = (\psi_K \mathbf{q}_f, 0) \in \underline{\mathbf{X}}$, we immediately have, according to Lemma 4, that

$$-\underline{d}(\zeta - \tilde{\zeta}_h, \xi_{f,K}) = (d\mathbb{T}_o^{1/2} \mathbf{q}_f, \psi_K \mathbf{q}_f)_{0,K}.$$

Then,

$$
\begin{aligned}
-\underline{d}(\zeta-\tilde{\zeta}_h,\xi_{f,K}) &\leq |\zeta-\tilde{\zeta}_h|_{+,K}\|\xi_{f,K}\|_{S,MG} \\
&\leq |\zeta-\tilde{\zeta}_h|_{+,K}\Big(\|d\mathbb{T}_o^{1/2}(\psi_K\mathbf{q}_f)\|_{0,K}^2+\|d\mathbb{T}_e^{-1/2}\mathrm{div}\,(\psi_K\mathbf{q}_f)\|_{0,K}^2\Big)^{1/2} \\
&\leq |\zeta-\tilde{\zeta}_h|_{+,K}\big(d\mathbb{T}_{o,K}^{max}+C(h_K^2(d\mathbb{T}_{e,K}^{min}))^{-1}\big)^{1/2}\|\psi_K\mathbf{q}_f\|_{0,K} \qquad (93)
\end{aligned}
$$

Combining (36), (37) and (93), one comes to

$$
c(d\mathbb{T}_{o,K}^{min})^{1/2}\|\mathbf{q}_f\|_{0,K}^2 \leq |\zeta-\tilde{\zeta}_h|_{+,K}\big(d\mathbb{T}_{o,K}^{max}+C(h_K^2(d\mathbb{T}_{e,K}^{min}))^{-1}\big)^{1/2}\|\mathbf{q}_f\|_{0,K}
$$

Considering the definition of $\eta_{f,K}$ by (85) concludes the proof. □

8 Conclusion

In this manuscript, we derive *a posteriori* estimates associated to different norms for the numerical solution of the neutron diffusion equation.
As a starting point, we consider the classical diffusion equation and observe that, although the approach presented in [18, Chapter 8] is guaranteed, it remains difficult to prove the local efficiency of the estimator. We address this issue by proposing *a posteriori* estimators that are guaranteed and locally efficient.

We focus on Cartesian meshes since such structures are relevant in nuclear core applications, and outline a robust marker strategy for this specific constraint, the *direction marker* strategy. We observe numerically that the AMR strategy is sensitive to the choice of the threshold parameter. We compare various *a posteriori* estimators under different criteria. We show that the choice of the reconstruction has a strong influence on the AMR strategy. The post-processing approaches are shown to be more efficient than the average reconstruction. In the case of the lowest-order Raviart-Thomas-Nédélec finite element, the RTN_0 post-processing gives a more accurate reconstruction compared to the RTN post-processing. Also, we compare the different estimators with the same choice of reconstruction. And we note that, if the stopping criterion is based on the L^2 error with respect to a reference solution, the various refinement strategies yield similar results.

Finally, we consider more general models or settings. First, we extend our *a posteriori* estimators to a Domain Decomposition Method, the so-called DD+L^2 jumps method. Then, we choose a more general model, widely used for nuclear core simulations, the multigroup diffusion problem, for which we also provide *a posteriori* estimators. We refer to [9] for an example of application.

References

[1] S. Ali Hassan, C. Japhet, M. Kern, and M. Vohralík. A posteriori stopping criteria for optimized Schwarz Domain Decomposition algorithms in mixed formulations. *Comput. Methods Appl. Math.*, 18:2369–2384, 2018.

[2] T. Arbogast and Z. Chen. On the implementation of mixed methods as nonconforming methods for second-order elliptic problems. *Math. Comp.*, 64(211):943–972, 1995.

[3] I. Babuška and M. Suri. The p and h-p versions of the finite element method, basic principles and properties. *SIAM Rev.*, 36(4):578–632, 1994.

[4] I. Babuška, B.A. Szabo, and I.N. Katz. The p-version of the finite element method. *SIAM J. Numer. Anal.*, 18(3):515–545, 1981.

[5] D. Boffi, F. Brezzi, and M. Fortin. *Mixed and hybrid finite element methods and applications*. Springer-Verlag, 2013.

[6] W. Cao, W. Huang, and R.D. Russell. An r-adaptive finite element method based upon moving mesh PDEs. *J. Comput. Phys.*, 149(2):221–244, 1999.

[7] C. Carstensen. A posteriori error estimate for the mixed finite element method. *Math. Comp.*, 66(218):465–476, 1997.

[8] P.G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathmetics*. SIAM, 2002.

[9] P. Ciarlet, Jr., M.-H. Do, and F. Madiot. Adaptive solution of the neutron diffusion equation with heterogeneous coefficients using the mixed finite element method on structured meshes. In *PHYSOR 2020, Cambridge, United Kingdom, March 29 - April 2, 2020*, 2020.

[10] P. Ciarlet, Jr., L. Giret, E. Jamelot, and F. D. Kpadonou. Numerical analysis of the mixed finite element method for the neutron diffusion eigenproblem with heterogeneous coefficients. *ESAIM: Math. Modell. Numer. Anal.*, 52:2003–2035, 2018.

[11] P. Ciarlet, Jr., E. Jamelot, and F. D. Kpadonou. Domain decomposition methods for the diffusion equation with low-regularity solution. *Comput. Math. Applic.*, 74:2369–2384, 2017.

[12] P. Daniel, A. Ern, I. Smears, and M. Vohralík. An adaptive hp-refinement strategy with computable guaranteed bound on the error reduction factor. *Comput. Math. Applic.*, 76(5):967–983, 2018.

[13] M. Dauge. Benchmark computations for Maxwell equations for the approximation of highly singular solutions. *Available at: https://perso.univ-rennes1.fr/monique.dauge/core/index.html*, 2004.

[14] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[15] J. J. Duderstadt and L. J. Hamilton. *Nuclear reactor analysis*. John Wiley & Sons, Inc., 1976.

[16] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*. Springer-Verlag, 2004.

[17] F. Févotte. Une méthode de post-traitement des éléments finis de Raviart-Thomas appliquée à la neutronique. CMAP Seminar, Palaiseau, May 21, 2019.

[18] L. Giret. *Non-conforming domain decomposition for the multigroup neutron SP_N equations*. PhD thesis, Université Paris Saclay, 2018.

[19] E. Jamelot, A.-M. Baudron, and J.-J. Lautard. Domain decomposition for the SP_N solver MINOS. *Transport Theory and Statistical Physics*, 41(7):495–512, 2012.

[20] E. Jamelot and P. Ciarlet, Jr. Fast non-overlapping Schwarz domain decomposition methods for solving the neutron diffusion equation. *J. Comput. Phys.*, 241:445–463, 2013.

[21] J. Lang, W. Cao, W. Huang, and R.D. Russell. A two-dimensional moving finite element method with local refinement based on a posteriori error estimates. *Appl. Numer. Math.*, 46(1):75–94, 2003.

[22] M. G. Larson and A. Målqvist. A posteriori error estimates for mixed finite element approximations of elliptic problems. *Numer. Math.*, 108(3):487–500, 2008.

[23] C. Lovadina and R. Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Math. Comp.*, 75(256):1659–1674, 2006.

[24] J.-C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35(3):315–341, 1980.

[25] P. Oswald. On a BPX-preconditioner for P1 elements. *Computing*, 51(2):125–133, 1993.

[26] P.-A. Raviart and J.-M. Thomas. A mixed finite element method for second order elliptic problems. In *Mathematical aspects of finite element methods*, volume 606 of *Lecture Notes in Mathematics*, pages 292–315. Springer, 1977.

[27] R. Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.*, 50(1-3):67–83, 1994.

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

[29] M. Vohralík. Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods. *Math. Comp.*, 79(272):2001–2032, 2010.

[30] M. F. Wheeler and I. Yotov. A posteriori error estimates for the mortar mixed finite element method. *SIAM J. Numer. Anal.*, 43(3):1021–1042, 2005.

[31] B. Wohlmuth and R. Hoppe. A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements. *Math. Comp.*, 68(228):1347–1378, 1999.