



**HAL**  
open science

# Mesh requirements for the finite element approximation of problems with sign-changing coefficients

Anne-Sophie Bonnet-Ben Dhia, Camille Carvalho, Patrick Ciarlet

► **To cite this version:**

Anne-Sophie Bonnet-Ben Dhia, Camille Carvalho, Patrick Ciarlet. Mesh requirements for the finite element approximation of problems with sign-changing coefficients. 2016. hal-01335153v1

**HAL Id: hal-01335153**

**<https://hal.science/hal-01335153v1>**

Preprint submitted on 21 Jun 2016 (v1), last revised 9 Oct 2017 (v3)

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

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

# Mesh requirements for the finite element approximation of problems with sign-changing coefficients

Anne-Sophie Bonnet-Ben Dhia      Camille Carvalho      Patrick Ciarlet Jr.

June 21, 2016

## Abstract

Transmission problems with sign-changing coefficients occur in electromagnetic theory in the presence of negative materials surrounded by classical materials. For general geometries, establishing Fredholmness of these transmission problems is well-understood thanks to the T-coercivity approach. Moreover, for a plane interface, there exist meshing rules that guarantee an optimal convergence rate for the finite element approximation. We propose here a new treatment at the corners of the interface which allows to design meshing rules for an arbitrary polygonal interface and then recover standard error estimates. This treatment relies on the use of simple geometrical transforms to define the meshes. Numerical results illustrate the importance of this new design.

**Keywords:** T-coercivity, transmission problem, sign-changing coefficient, conforming finite element method, T-conforming mesh.

## 1 Introduction and setting of the transmission problem with sign-changing coefficients

Our aim is to solve the problem

$$\begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that:} \\ -\operatorname{div}(\sigma \nabla u) = f \text{ in } \Omega, \end{cases} \quad (1)$$

where  $\Omega \subset \mathbb{R}^2$  is a bounded domain partitioned into two regions,  $\sigma$  is a scalar, sign-changing coefficient, and  $f$  is some given data. In electromagnetic theory, this problem can be interpreted as a transmission problem, in a domain composed of a classical dielectric material ( $\sigma > 0$ ), and a negative material ( $\sigma < 0$ ). A negative material can be for example a metal at optical frequencies, or a metamaterial (e.g. [22, 1]), for which some physical parameters become negative (the permittivity for metals, both the permittivity and the permeability for metamaterials). Due to the sign-changing coefficient  $\sigma$ , well-posedness of problem (1) is not guaranteed. In particular, classical tools such as Lax-Milgram theorem do not apply since the coercivity on  $H_0^1(\Omega) \times H_0^1(\Omega)$  of the corresponding sesquilinear form

$$a : (v, w) \mapsto \int_{\Omega} \sigma \nabla v \cdot \overline{\nabla w} \, d\Omega, \quad (2)$$

is lost. However, over the past decade, techniques have been developed to establish well-posedness, under appropriate conditions, via the T-coercivity theory: introduced in [5], it consists in building isomorphisms  $T$  such that the form  $(v, w) \mapsto a(v, Tw)$  is coercive on  $H_0^1(\Omega) \times H_0^1(\Omega)$ . For short, we say that  $a(\cdot, \cdot)$  is T-coercive. What is less clear is the discrete counterpart of this approach.

In this paper, we consider problem (1) with the following hypothesis on  $\sigma$ :

$$\begin{cases} \sigma_1 := \sigma|_{\Omega_1} \text{ is a constant such that } \sigma_1 > 0, \\ \sigma_2 := \sigma|_{\Omega_2} \text{ is a constant such that } \sigma_2 < 0. \end{cases}$$

The ratio  $\kappa_{\sigma} := \sigma_2/\sigma_1$  is called the *contrast*. Let  $\Omega$  be a *domain* of  $\mathbb{R}^2$ , that is a connected bounded open subset of  $\mathbb{R}^2$  with a Lipschitz boundary. It is split as  $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$ , where  $\Omega_1$  and  $\Omega_2$  are two disjoint domains.

The interface separating  $\Omega_1$  and  $\Omega_2$  is called  $\Sigma$ : we assume that it is a polygonal line made of straight edges and corners. Given  $v$  defined over  $\Omega$ , we use the notation  $v_i := v|_{\Omega_i}$ ,  $i = 1, 2$ .

The equivalent variational formulation of (1) reads:

$$\text{Find } u \in V \text{ such that } \forall w \in V, a(u, w) = \langle f, w \rangle, \quad (3)$$

where  $V = H_0^1(\Omega)$ ,  $a$  is the form defined in (2), and  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $V$  and its dual  $V'$ .

Let us first describe a simple configuration for which everything is well understood. It is the symmetric geometry:  $\Sigma$  is a part of a straight line and  $\Omega_2$  is the symmetric of  $\Omega_1$  with respect to  $\Sigma$ . Then one can easily prove that the problem is well-posed if and only if  $\kappa_\sigma \neq -1$ , by considering the following operators  $\mathsf{T}$ :

$$\mathsf{T}v = \begin{cases} v_1 & \text{in } \Omega_1 \\ -v_2 + 2\mathcal{S}v_1 & \text{in } \Omega_2 \end{cases}, \text{ respectively } \mathsf{T}v = \begin{cases} v_1 - 2\mathcal{S}v_2 & \text{in } \Omega_1 \\ -v_2 & \text{in } \Omega_2 \end{cases},$$

where  $\mathcal{S}v_1(x) = v_1(\mathcal{S}x)$ , with  $\mathcal{S}$  denoting the symmetry with respect to  $\Sigma$ . On one hand, one can prove that  $a(\cdot, \cdot)$  is  $\mathsf{T}$ -coercive if  $\kappa_\sigma \neq -1$ . On the other hand, if  $\kappa_\sigma = -1$ , the problem (1) is ill-posed since it has a kernel (set of solutions with zero right-hand side  $f$ ) which is infinite dimensional. Thus we say that the condition  $\kappa_\sigma \neq -1$  is *optimal* for the well-posedness of the continuous problem.

Suppose now that  $\kappa_\sigma \neq -1$  and that we want to approximate the solution with a conforming finite element method (for short, a cFE method). This leads to the discrete problems

$$\text{Find } u^h \in V^h \text{ such that } \forall w^h \in V^h, a(u^h, w^h) = \langle f, w^h \rangle, \quad (4)$$

where  $(V^h)_h$  denotes a sequence of finite-dimensional subspaces of  $V$ , with  $h$  a positive parameter that goes to 0. If  $\mathsf{T}(V^h) \subset V^h$  for small  $h$ , then  $\mathsf{T}$ -coercivity can be exploited at the discrete level. In particular, the discrete problem is well-posed and, by Céa's lemma, the error is controlled by the best approximation error. Hence, we just have to ensure the condition  $\mathsf{T}(V^h) \subset V^h$ , which is achieved in this symmetric geometry by using a symmetric mesh (for a uniform degree of approximation). When  $\mathsf{T}(V^h) \subset V^h$ , we say that the mesh is  *$\mathsf{T}$ -conforming*. Let us emphasize that using non-symmetric meshes can deteriorate drastically the convergence of the cFE method when  $\kappa_\sigma$  is close to  $-1$  (cf. [10]).

The general objective of our paper is to generalize this type of result. To do so for an arbitrary geometry, one has to build an operator  $\mathsf{T}$  such that: 1) the form  $a$  is  $\mathsf{T}$ -coercive ; 2) it is possible to choose  $V^h$  such that  $\mathsf{T}(V^h) \subset V^h$ . To our knowledge, it has been provided operators  $\mathsf{T}$  satisfying 1) (see [2]) for any 2D configuration, however condition 2) is not always ensured. The main difficulty lies in the treatment of the corners' interface  $\Sigma$ . Indeed, proceeding similarly as presented for the symmetric case, one can build *geometry-based operators* realizing 1) near the corners [2]. Unfortunately these operators involve angular dilations, therefore 2) is never satisfied since the polynomial nature of basis FE functions is not preserved by such transform. The goal is to propose another construction of operator  $\mathsf{T}$ , based on isometries [18], such that for appropriate meshes condition 2) holds. In other words, as 2) boils down to some condition on the mesh, our aim is to provide meshing rules for an arbitrary geometry, ensuring that the standard convergence rate is recovered, as soon as  $\kappa_\sigma$  is such that the continuous problem (1) with sign-changing coefficients is well-posed.

The outline of the paper is as follows. In the next section, we provide a review of the techniques proposed so far to approximate problem (1) with sign-changing coefficients and motivate why and how to satisfy condition 2). Section 3 is dedicated to the construction of new explicit  $\mathsf{T}$ -coercivity operators for elementary geometries whose interface has only one corner. Then in section 4 we develop the theory that allows one to study the well-posedness of problem (1) for an arbitrary polygonal interface, and to derive the optimality condition on the contrast  $\kappa_\sigma$ . Our aim is to provide tools that can be extended to the discrete problems, namely the approximation of problem (1) by cFE methods. This is the main topic of section 5, where convergence is derived as soon as the optimality condition on the contrast is fulfilled. Numerical results are presented in section 6. Finally some concluding remarks are given.

## 2 Historical background: $\mathsf{T}$ -coercivity and $\mathsf{T}$ -conforming meshes

Let us review known results concerning the  $\mathsf{T}$ -coercivity approach. Introduced in [5], the  $\mathsf{T}$ -coercivity technique has then been developed in [2]. Similarly to the symmetric case, the idea is to build linear continuous operators

T on V:

$$\mathbf{T}v = \begin{cases} v_1 & \text{in } \Omega_1 \\ -v_2 + 2\mathbf{R}v_1 & \text{in } \Omega_2 \end{cases}, \text{ respectively } \mathbf{T}'v = \begin{cases} v_1 - 2\mathbf{R}'v_2 & \text{in } \Omega_1 \\ -v_2 & \text{in } \Omega_2 \end{cases}, \quad (5)$$

where  $\mathbf{R}$ ,  $\mathbf{R}'$  are operators (to be precised) such that  $\mathbf{R}v_1|_\Sigma = v_1|_\Sigma$ ,  $\mathbf{R}'v_2|_\Sigma = v_2|_\Sigma$ . This construction is attractive because once the operators  $\mathbf{R}$ ,  $\mathbf{R}'$  are settled, the operators  $\mathbf{T}$ ,  $\mathbf{T}'$  define isomorphisms of  $V$ . Let us present the available constructions of  $\mathbf{R}$  (note that  $\mathbf{R}'$  is obtained by inverting the roles of  $\Omega_1$  and  $\Omega_2$ ):

- the first idea (see [5]) was to consider an operator  $\mathbf{R}^{(1)}$  that acts from  $H_{00}^{1/2}(\Sigma)$  (the space of traces on  $\Sigma$  of functions of  $V$ ) to  $V_2$ , where  $V_i := \{v_i, v \in V\}$ ,  $i = 1, 2$ . Then it was proven that T-coercivity is realized with  $\mathbf{T}^{(1)}$  (defined as (5) with  $\mathbf{R}^{(1)}$ ) under the condition that  $|\kappa_\sigma|$  is "small enough" (with no explicit bound) ;
- an alternative was next proposed in [18] with an operator  $\mathbf{R}^{(2)}$  that acts from  $V_1$  to  $V_2$ . Let us point out the difference between the two approaches: in  $\mathbf{T}^{(1)}$ ,  $\mathbf{R}^{(1)}$  acts on the trace of  $v_1$  on  $\Sigma$ , while in  $\mathbf{T}^{(2)}$  (defined as (5) with  $\mathbf{R}^{(2)}$ ),  $\mathbf{R}^{(2)}$  acts on the whole function  $v_1$ , defined on  $\Omega_1$ . The case of the symmetry-based operator  $\mathbf{S}$  fits into this second category but not in the first one. In that case, T-coercivity is realized under the condition

$$\kappa_\sigma \notin [-\|\mathbf{R}'\|^2; -1/\|\mathbf{R}\|^2], \quad (6)$$

with

$$\|\mathbf{R}\| = \sup_{w_1 \in V_1 \setminus \{0\}} \frac{\|\nabla \mathbf{R}w_1\|_{\Omega_2}}{\|\nabla w_1\|_{\Omega_1}}, \quad \text{and} \quad \|\mathbf{R}'\| = \sup_{w_2 \in V_2 \setminus \{0\}} \frac{\|\nabla \mathbf{R}'w_2\|_{\Omega_1}}{\|\nabla w_2\|_{\Omega_2}}; \quad (7)$$

- the case of an arbitrary geometry was completely clarified in [2]. The main conclusions are twofold. First, it appears that proving the well-posedness of (3) is too restrictive, and possible only for very simple cases. A more relevant objective is to ensure that the problem is well-posed in the Fredholm sense (see section 4). And it has been proved by a localization process that this property depends only on the value of the contrast  $\kappa_\sigma$  and on the geometry of the interface  $\Sigma$ , but not on the global geometry of  $\Omega$ . To do so, the authors introduce

$$\mathbf{T}^{(3)}v = \begin{cases} v_1 & \text{on } \Omega_1 \\ -v_2 + 2\chi \mathbf{R}^{(3)}v_1 & \text{on } \Omega_2 \end{cases}, \quad (8)$$

where, thanks to the introduction of a cutoff function  $\chi$  (whose support coincides with a neighborhood of  $\Sigma$ ), the operator  $\mathbf{R}^{(3)}$  needs only to be defined with the help of local geometrical transforms, which map a part of  $\Omega_1$  to a part of  $\Omega_2$ . For instance,  $\mathbf{R}^{(3)}$  is based on the composition of a central symmetry and an angular dilation near a corner. The main result is that, for a given closed polygonal line  $\Sigma$  (possibly with curved edges), problem (3) is well-posed in the Fredholm sense if and only if  $\kappa_\sigma \notin [-I_\alpha; -1/I_\alpha]$  where  $\alpha$  is the smallest corner angle of the polygonal line and

$$I_\alpha := \max\left(\frac{2\pi - \alpha}{\alpha}, \frac{\alpha}{2\pi - \alpha}\right). \quad (9)$$

What are the discrete counterparts of each approach ?

- At the discrete step, the condition  $\mathbf{T}^{(1)}(V^h) \subset V^h$  was not sought. Instead, the idea was to introduce a discrete counterpart of the operator  $\mathbf{T}^{(1)}$ : for  $v^h \in V^h$ , one defines

$$\mathbf{T}_h^{(1)}v^h = \begin{cases} v_1^h & \text{in } \Omega_1 \\ -v_2^h + 2\mathbf{R}_h^{(1)}(v_1^h|_\Sigma) & \text{in } \Omega_2 \end{cases},$$

where  $\mathbf{R}_h^{(1)}(v_1^h|_\Sigma)$  is defined as the discrete-harmonic element of  $V_2^h$  ( $V_i^h := \{v_i^h, v^h \in V^h\}$ ,  $i = 1, 2$ ), with a trace on  $\Sigma$  equal to  $v_1^h|_\Sigma$ . It is furthermore proved in [5] that, for a quasi-uniformly meshed interface, under a condition on the contrast which is more restrictive than the one for the continuous problem, convergence of the cFE method is achieved. Obviously, in the symmetric geometry, this result is less satisfactory than the one obtained for the symmetry-based operator  $\mathbf{T}$ .

- To avoid the use of quasi-uniform meshes, in [18] is proposed the discrete counterpart, for  $v^h \in V^h$ ,

$$\mathbf{T}_h^{(2)}v^h = \begin{cases} v_1^h & \text{in } \Omega_1 \\ -v_2^h + 2I_h(\mathbf{R}^{(2)}v_1^h) & \text{in } \Omega_2 \end{cases}$$

---

<sup>1</sup>From now on, we denote by  $\|\cdot\|_{\mathcal{O}}$  the  $L^2$ -norm over the open set  $\mathcal{O}$ .

where  $I_h$  is a projection operator, with values in  $V_2^h$ . Note that in the case of a symmetric domain, one does not need to introduce an operator  $I_h$  if the meshes are symmetric. However, despite the analogy to the symmetric geometry, this second approach has the same drawback than the first one, at least for a general geometry: due to the presence of the projection operator  $I_h$ , numerical convergence is not guaranteed even if the continuous problem is T-coercive.

Interestingly, the authors of [18] exhibit a special geometry, different from the symmetric one, where the operator  $\mathbf{R}$  can be again built using symmetries. This is the case where  $\Omega$  is a square, and  $\Omega_1$  is one quadrant of this square (right-angle geometry). The definition of  $\mathbf{R}$  now only involves axial symmetries with respect to axes  $Ox_1$  and  $Ox_2$  (if the origin  $O$  is located at the center of  $\Omega$ ). For the continuous problem, this allows to prove the well-posedness if  $\kappa_\sigma \notin [-3; -1/3]$  (condition that has been proved to be optimal in [2]). Then at the discrete level, one doesn't have to introduce an operator  $I_h$ . Indeed, if the mesh is symmetric with respect to  $Ox_1$  and  $Ox_2$ , then the condition  $\mathbf{T}^{(2)}(V^h) \subset V^h$  is satisfied (T-conforming mesh) and the *optimality of the approximation* is obtained: in others words, the convergence is ensured as soon as the continuous problem is well-posed.

Summing up, imposing  $\mathbf{T}(V^h) \subset V^h$  is essential to guarantee convergence of the cFE method under the same conditions as the ones to prove well-posedness of the continuous problem, and providing T-conforming meshes is an attractive method to satisfy this condition.

With the results of [2], since the global geometry of  $\Omega$  does not matter for the well-posedness (in the Fredholm sense) of the continuous problem, one can expect that the properties of the mesh "far from the interface"  $\Sigma$  should not matter either. In other words, can one prove that a local T-conformity of the mesh around  $\Sigma$  is sufficient to ensure convergence of the cFE method as soon as the continuous problem is well-posed in the Fredholm sense?

- A partial answer to this question has been given in [10, p. 17], considering  $\mathbf{T}^{(3)}$  with the discrete version, for  $v^h \in V^h$ ,

$$\mathbf{T}_h^{(3)} v^h = \begin{cases} v_1^h & \text{in } \Omega_1 \\ -v_2^h + 2\chi_h \mathbf{R}^{(3)} v_1^h & \text{in } \Omega_2 \end{cases}, \quad (10)$$

with  $\chi_h$  the interpolation of  $\chi$  on  $V^h$ . In this framework, the constraint on the meshes boils down to a simple requirement: namely, that the mesh on  $\Omega_2$  is locally deduced from the mesh on  $\Omega_1$  by some geometrical transforms. This is the so-called *local T-conformity* of the mesh [10, Definition 10]. In principle, optimality of the cFE method follows. However, they are two shortcomings. The first one is technical, namely there is a mistake in the definition (10) of the discrete operator  $\mathbf{T}_h^{(3)}$ : as a matter of fact, even if  $\mathbf{R}^{(3)} v_1^h$  is locally in  $V_2^h$ ,  $\chi_h \mathbf{R}^{(3)} v_1^h$  does not belong to  $V_2^h$  in general. The second shortcoming is that, except for special geometries such as the right-angle geometry described above, no interface with corners was tackled. Again, let us emphasize that, in the case of an interface with corners, for the operator  $\mathbf{T}^{(3)}$  defined in [2], the condition  $\mathbf{T}^{(3)}(V^h) \subset V^h$  is never satisfied because the angular dilation of a polynomial function is not a polynomial function.

The questions are then: is it possible to find another optimal construction of the operator  $\mathbf{T}$  such that for appropriate meshes there holds  $\mathbf{T}(V^h) \subset V^h$ , as it was the case with the axial symmetries for the right-angle geometry? This topic is addressed in §3. More generally, for a polygonal interface, can we recover optimality of the cFE method under a generalized *local T-conformity* condition of the meshes? This is the subject of §4.

### 3 New T-coercivity operators and associated T-conforming meshes around corners

In this section, we consider a domain  $\Omega$  whose interface  $\Sigma$  separating  $\Omega_1$  and  $\Omega_2$  is a polygonal line with only one interior corner  $c$  of angle  $\alpha$  measured in  $\Omega_1$ , and two endpoints on the boundary  $\partial\Omega$ . Moreover, if  $(\rho, \theta)$  denote the local polar coordinates centered at  $c$ , then  $\Omega_1$  locally coincides with the cone  $\{(\rho \cos \theta, \rho \sin \theta) \mid 0 < \rho, 0 < \theta < \alpha\}$ . In the particular case of a disk  $\Omega$ , one considers the operators (5) realizing T-coercivity [2], with  $\mathbf{R}, \mathbf{R}'$  such that, for all  $(v_1, v_2) \in V_1 \times V_2$ ,

$$\mathbf{R}v_1(\rho, \theta) = v_1(\rho, \frac{\alpha}{\alpha - 2\pi}(\theta - 2\pi)), \quad \forall(\rho, \theta) \in \Omega_2, \quad \text{and} \quad \mathbf{R}'v_2(\rho, \theta) = v_2(\rho, \frac{\alpha - 2\pi}{\alpha}\theta + 2\pi), \quad \forall(\rho, \theta) \in \Omega_1.$$

It is proved in [2] that  $a$  (defined in (2)) is T-coercive if and only if  $\kappa_\sigma \notin [-I_\alpha; -1/I_\alpha]$  where  $I_\alpha$  is defined in (9)<sup>(2)</sup>. The goal of this section is to propose an alternative operator  $\mathbf{R}_{\text{new}}$  (resp.  $\mathbf{R}'_{\text{new}}$ ) such that:

- T-coercivity holds for  $\mathbf{T}^{(2)}$  with  $\mathbf{R}_{\text{new}}$  (resp.  $\mathbf{T}'^{(2)}$  with  $\mathbf{R}'_{\text{new}}$ ) under the same conditions as above;
- one can provide meshing rules such that the condition  $\mathbf{T}^{(2)}(V^h) \subset V^h$  (resp.  $\mathbf{T}'^{(2)}(V^h) \subset V^h$ ) holds (such as using a symmetric mesh when one considers a symmetry-based operator).

This alternative method consists in building *rotation-* and *symmetry-based operators*, and can be applied to other geometries than a disk. In fact the key idea is to consider a *pattern-based domain*  $\Omega$  (see figure 1).

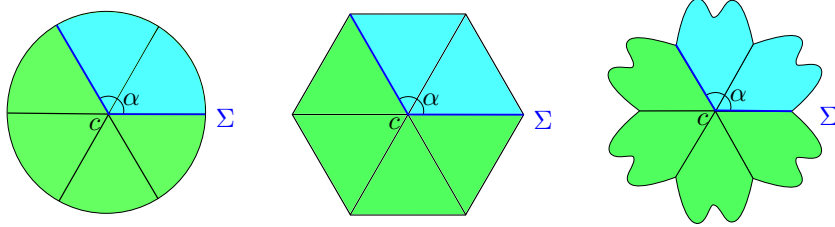


Figure 1: Examples of pattern-based geometries with a corner  $c$  of angle  $\alpha = 2\pi/3$  measured in  $\Omega_1$ : (left) sector-based, (middle) triangle-based, and (right) leaf-based geometry.

Let us introduce some notations. Suppose one can identify a pattern  $\mathcal{P}$  such that  $\Omega$  is a pattern-based domain of pattern  $\mathcal{P}$  (sector, triangle, etc.). Introduce  $p, q > 0$ ,  $p \neq q$ , with  $p + q$  even such that  $\Omega_1$  is composed of  $p$  patterns, and  $\Omega_2$  is composed of  $q$  patterns (in figure 1, one can choose for instance  $p = 2$  and  $q = 4$  with respectively a sector, a triangle, a leaf as pattern  $\mathcal{P}$ ). Then one rewrites  $\alpha := 2\pi \frac{p}{p+q}$  and due to the symmetry of the geometry it follows that

$$I_\alpha := \max\left(\frac{p}{q}, \frac{q}{p}\right). \quad (11)$$

Note that the case  $p = q$  corresponds to a straight angle (this case is already treated in [2, 10]). Also, the condition on the parity of  $p + q$  is by no means restrictive. Indeed if  $p + q$  is odd, one simply doubles  $p$  and  $q$  without changing the value of  $\alpha$ . Introduce also :

- $\Omega_1^k$ ,  $k = 1, p$ , the open set corresponding to the  $k$ -th pattern of  $\Omega_1$ :  $\overline{\Omega_1} = \bigcup_{k=1}^p \overline{\Omega_1^k}$ . Respectively define  $\Omega_2^l$ ,  $l = 1, q$ , the open set corresponding to the  $l$ -th pattern of  $\Omega_2$ :  $\overline{\Omega_2} = \bigcup_{l=1}^q \overline{\Omega_2^l}$ . The numbering is chosen counterclockwise.
- $v_1^k := v|_{\Omega_1^k}$ ,  $k = 1, p$  the restriction to  $\Omega_1^k$ . Similarly,  $v_2^l := v|_{\Omega_2^l}$ ,  $l = 1, q$ .
- $e_n$ ,  $n = 1, 2$ , the two edges of  $\Sigma$  such that  $e_1$  coincides locally with  $\{(\rho, 0), 0 < \rho\}$ , while  $e_2$  coincides locally with the line  $\{(\rho \cos \alpha, \rho \sin \alpha), 0 < \rho\}$ .

To define  $\mathbf{R}_{\text{new}}$  as a rotation- and symmetry-based operator from  $V_1$  to  $V_2$ , one introduces  $\mathcal{S}^n$  the axial symmetry of axis  $e_n$ ,  $n = 1, 2$ , and  $\mathcal{R}_m$  the rotation of angle  $m\alpha/p$ ,  $m = 1, p + q - 1$ . Define also their inverse,  $\mathcal{S}^n$  (being its own inverse) and  $\mathcal{R}_{-m}$  the rotation of angle  $-m\alpha/p$ . These transforms are such that

- for  $(\rho, \theta) \in \Omega_1^k$ , for all  $k \in I_1 := \llbracket 1, \min(p, q) \rrbracket$ ,  $\mathcal{S}^1(\rho, \theta) \in \Omega_2^{q+1-k}$ ,
- for  $(\rho, \theta) \in \Omega_1^k$ , for all  $k \in I_2 := \llbracket p + 1 - \min(p, q), p \rrbracket$ ,  $\mathcal{S}^2(\rho, \theta) \in \Omega_2^{p+1-k}$ ,
- for  $(\rho, \theta) \in \Omega_1^k$ , for all  $k \in \llbracket 1, p \rrbracket$ ,  $\mathcal{R}_{p+l-k}(\rho, \theta) \in \Omega_2^l$ , for all  $l \in \llbracket 1, q \rrbracket$ .

<sup>2</sup>In general one has to be careful in the presence of the two exterior corners  $c_1^{\text{ext}}$ ,  $c_2^{\text{ext}}$ , that is the endpoints of  $\Sigma$  on the boundary  $\partial\Omega$ . Nonetheless, in the case where the angles at the corner  $c_1^{\text{ext}}$  (resp.  $c_2^{\text{ext}}$ ) are the same in  $\Omega_1$  and  $\Omega_2$  (as it is always the case in this paper), (9) holds. If not, then the expression of  $I_\alpha$  is modified and it depends on the corners' angles at  $c_1^{\text{ext}}$  and  $c_2^{\text{ext}}$  (see [2]).

Then one defines symmetry-based operators  $\mathcal{S}^n$  ( $n = 1, 2$ ) and rotation-based operators  $\mathcal{R}_m$  ( $m = 1, p+q-1$ ) from  $V_1^k := \{v_1^k, v \in V\}$  to  $V_2^l := \{v_2^l, v \in V\}$  by:

- $\mathcal{S}^1 v_1^k(\rho, \theta) = v_1^k(\mathcal{S}^1(\rho, \theta))$  for  $(\rho, \theta) \in \Omega_2^{q+1-k}$ , for all  $k \in I_1$ ,
- $\mathcal{S}^2 v_1^k(\rho, \theta) = v_1^k(\mathcal{S}^2(\rho, \theta))$  for  $(\rho, \theta) \in \Omega_2^{p+1-k}$ , for all  $k \in I_2$ ,
- $\mathcal{R}_{k-(p+l)} v_1^k(\rho, \theta) = v_1^k(\mathcal{R}_{k-(p+l)}(\rho, \theta))$  for  $(\rho, \theta) \in \Omega_2^l$ , for all  $l = 1, q$ , for all  $k = 1, p$ .

Finally one can define global operators  $\mathbf{R}$  from  $V_1$  to  $V_2$  composing these rotation-based and symmetry-based operators. To be an *admissible operator*,  $\mathbf{R}$  has to ensure continuity at all interfaces between the patterns. Roughly speaking, the definition of an admissible operator is based on three steps (the general construction is given in the Appendix A.1):

- 1) apply  $\mathcal{S}^1$  to  $v_1^k$  for some ad hoc  $k$ , for  $(\rho, \theta) \in \Omega_2^{q+1-k}$ ;
- 2) apply  $\mathcal{S}^2$  to  $v_1^k$  for some ad hoc  $k$ , for  $(\rho, \theta) \in \Omega_2^{p+1-k}$ ;
- 3) ensure continuity by applying geometry-based operators to the remaining  $(v_1^{k'})_{k' \in I}$ ,  $I$  being an appropriate subset of indices (see Appendix A.1).

Let us give some examples. In the following we consider a triangle-based domain  $\Omega$ .

**First example** ( $\alpha < \pi$ ): consider for instance  $\alpha = 2\pi/3$ ,  $p = 2$  and  $q = 4$ . Then one can define two admissible operators (illustrated in figure 2):

$$\mathbf{R}_1^{\text{adm}} v_1 = \begin{cases} \mathcal{S}^2 v_1^2 & \text{in } \Omega_2^1 \\ \mathcal{R}_{-2} v_1^2 & \text{in } \Omega_2^2 \\ \mathcal{S}^1 v_1^2 & \text{in } \Omega_2^3 \\ \mathcal{S}^1 v_1^1 & \text{in } \Omega_2^4 \end{cases}, \quad \text{and} \quad \mathbf{R}_2^{\text{adm}} v_1 = \begin{cases} \mathcal{S}^2 v_1^2 & \text{in } \Omega_2^1 \\ \mathcal{S}^2 v_1^1 & \text{in } \Omega_2^2 \\ \mathcal{R}_{-4} v_1^1 & \text{in } \Omega_2^3 \\ \mathcal{S}^1 v_1^1 & \text{in } \Omega_2^4 \end{cases}.$$

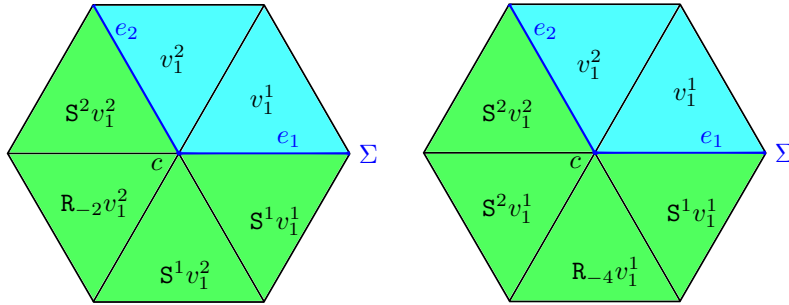


Figure 2: Representation of the two admissible geometry-based operators for the first example.

**Second example** ( $\alpha > \pi$ ): consider for instance  $\alpha = 4\pi/3$ ,  $p = 4$  and  $q = 2$  (one simply inverts the roles of  $p$  and  $q$  from the first example). In that case one "folds  $v_1$  up" such that  $\mathcal{R}v_1$  is defined over  $\Omega_2$ . Then one obtains the two admissible operators described in figure 3:

$$\mathbf{R}_1^{\text{adm}} v_1 = \begin{cases} \mathcal{S}^2 v_1^4 & \text{in } \Omega_2^1 \\ \mathcal{S}^1 v_1^1 - \mathcal{R}_{-4} v_1^2 + \mathcal{S}^2 v_1^3 & \text{in } \Omega_2^2 \end{cases}, \quad \text{and} \quad \mathbf{R}_2^{\text{adm}} v_1 = \begin{cases} \mathcal{S}^2 v_1^4 - \mathcal{R}_{-2} v_1^3 + \mathcal{S}^1 v_1^2 & \text{in } \Omega_2^1 \\ \mathcal{S}^1 v_1^1 & \text{in } \Omega_2^2 \end{cases}.$$

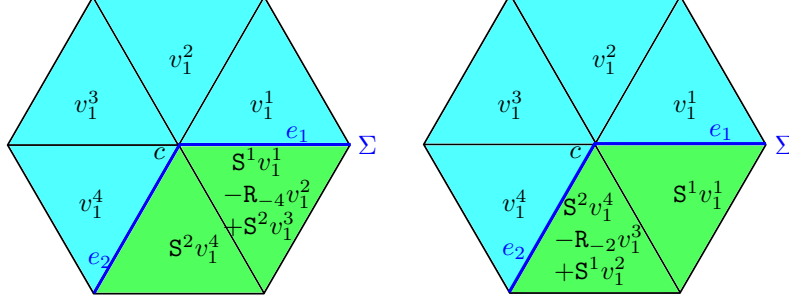


Figure 3: Representation of the two admissible geometry-based operators for the second example: (left) one applies an axial symmetry in  $\Omega_1^4$  to reach  $\Omega_2^1$  and folds the rest of  $v_1$  into  $\Omega_2^2$  to ensure continuity, (right) one applies an axial symmetry in  $\Omega_1^1$  to reach  $\Omega_2^2$  and folds the rest of  $v_1$  into  $\Omega_2^1$  to ensure continuity.

Remark that we can build admissible operators  $\mathbf{R}'$  with the same operators (using the inverse of the rotation-based and symmetry-based operators), by simply inverting the roles of  $p$  and  $q$ . One can easily generalize this method to any  $p, q > 0$  (see the Appendix A.1 for details).

To realize T-coercivity under the same conditions as in [2], one finally has to check that these operators  $\mathbf{R}, \mathbf{R}'$  are optimal in the sense that  $\|\mathbf{R}\|^2 = \|\mathbf{R}'\|^2 = I_\alpha$  defined in (9) or (11), the operator norms being defined in (7). For the two previous examples,  $I_\alpha = 2$ . Let us check whether  $\|\mathbf{R}_n^{\text{adm}}\|^2 = \|\mathbf{R}_n^{\text{adm}'}\|^2 = 2$ ,  $n = 1, 2$ . Let us go back to our first example. Since  $\mathbf{S}^n$  and  $\mathbf{R}_{-m}$  ( $n = 1, 2, m = 1, 5$ ) are isometry-based operators, one finds easily that

$$\|\nabla \mathbf{R}_1^{\text{adm}} v_1\|_{\Omega_2}^2 = 3\|\nabla v_1^2\|_{\Omega_1^2}^2 + \|\nabla v_1^1\|_{\Omega_1^1}^2 \leq 3\|\nabla v_1\|_{\Omega_1}^2, \quad \text{and} \quad \|\nabla \mathbf{R}_2^{\text{adm}} v_1\|_{\Omega_2}^2 = \|\nabla v_1^2\|_{\Omega_1^2}^2 + 3\|\nabla v_1^1\|_{\Omega_1^1}^2 \leq 3\|\nabla v_1\|_{\Omega_1}^2,$$

which is less satisfactory than the expected optimal norm (that is 2). However, considering the average of these two admissible operators gives us the result. More precisely, define

$$\mathbf{R}_{\text{new}} v_1 := \frac{1}{2}(\mathbf{R}_1^{\text{adm}} + \mathbf{R}_2^{\text{adm}})v_1 = \begin{cases} \mathbf{S}^2 v_1^2 & \text{in } \Omega_2^1 \\ \frac{1}{2}(\mathbf{R}_{-2} v_1^2 + \mathbf{S}^2 v_1^1) & \text{in } \Omega_2^2 \\ \frac{1}{2}(\mathbf{S}^1 v_1^2 + \mathbf{R}_{-4} v_1^1) & \text{in } \Omega_2^3 \\ \mathbf{S}^1 v_1^1 & \text{in } \Omega_2^4 \end{cases},$$

then one can prove that  $\|\mathbf{R}_{\text{new}}\|^2 = 2$ . Indeed, for  $v_1 \in V_1$

$$\|\nabla(\mathbf{R}_{\text{new}} v_1)\|_{\Omega_2}^2 = \|\nabla(\mathbf{S}^2 v_1^2)\|_{\Omega_2^1}^2 + \|\frac{1}{2}\nabla(\mathbf{R}_{-2} v_1^2 + \mathbf{S}^2 v_1^1)\|_{\Omega_2^2}^2 + \|\frac{1}{2}\nabla(\mathbf{S}^1 v_1^2 + \mathbf{R}_{-4} v_1^1)\|_{\Omega_2^3}^2 + \|\nabla(\mathbf{S}^1 v_1^1)\|_{\Omega_2^4}^2$$

and since  $\mathbf{S}^n$  and  $\mathbf{R}_{-m}$  ( $n = 1, 2, m = 1, 5$ ) are isometry-based operators, using the triangle inequality one finds that

$$\|\nabla(\mathbf{R}_{\text{new}} v_1)\|_{\Omega_2}^2 \leq \|\nabla v_1^2\|_{\Omega_1^2}^2 + (\frac{1}{2}\|\nabla v_1^2\|_{\Omega_1^2} + \frac{1}{2}\|\nabla v_1^1\|_{\Omega_1^1})^2 + (\frac{1}{2}\|\nabla v_1^2\|_{\Omega_1^2} + \frac{1}{2}\|\nabla v_1^1\|_{\Omega_1^1})^2 + \|\nabla v_1^1\|_{\Omega_1^1}^2.$$

Define the matrix  $M = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$ , and  $\vec{W} \in \mathbb{R}^2$  such that  $\vec{W} := (\|\nabla v_1^1\|_{\Omega_1^1}, \|\nabla v_1^2\|_{\Omega_1^2})^\top$ . Then one has

$M\vec{W} = (\|\nabla v_1^2\|_{\Omega_1^2}, \frac{1}{2}\|\nabla v_1^1\|_{\Omega_1^1} + \frac{1}{2}\|\nabla v_1^2\|_{\Omega_1^2}, \frac{1}{2}\|\nabla v_1^1\|_{\Omega_1^1} + \frac{1}{2}\|\nabla v_1^2\|_{\Omega_1^2}, \|\nabla v_1^1\|_{\Omega_1^1})^\top$  and one remarks that

$$\|\nabla \mathbf{R}_{\text{new}} v_1\|_{\Omega_2}^2 \leq \|M\vec{W}\|_2^2 \leq (M^\top M \vec{W}, \vec{W}) \leq \|M^\top M\|_2 \|\nabla v_1\|_{\Omega_1}^2,$$

where  $\|\cdot\|_2$  denotes the euclidean norm. Hence  $\|\mathbf{R}_{\text{new}}\|^2 \leq \|M^\top M\|_2$ . Finally a straightforward computation shows that  $\|M^\top M\|_2 = 2$ . One can operate similarly for the second example by considering  $\mathbf{R}'_{\text{new}} v_1$  over  $\Omega_2$



as the average of the two operators described in figure 3:

$$\mathbf{R}'_{\text{new}} v_1 = \begin{cases} \frac{1}{2}(2\mathbf{S}^2 v_1^4 - \mathbf{R}_{-2} v_1^3 + \mathbf{S}^1 v_1^2) & \text{in } \Omega_2^1 \\ \frac{1}{2}(2\mathbf{S}^1 v_1^4 - \mathbf{R}_{-4} v_1^3 + \mathbf{S}^2 v_1^2) & \text{in } \Omega_2^2 \end{cases}.$$

Then one finds  $\|\mathbf{R}'_{\text{new}}\|^2 \leq \|M'^\top M'\|_2 = 2$ , with the matrix  $M' = \begin{pmatrix} 0 & 1/2 & 1/2 & 1 \\ 1 & 1/2 & 1/2 & 0 \end{pmatrix} = M^\top$ . Remark that it is expected to find that  $M' = M^\top$  as the role of  $p$  and  $q$  are exchanged. Thus, one obtains the same results for  $\mathbf{R}'_{\text{new}}$ . Finally, we note that equality for the norms, that is  $\|\mathbf{R}_{\text{new}}\|^2 = \|M^\top M\|_2$ , resp.  $\|\mathbf{R}'_{\text{new}}\|^2 = \|M'^\top M'\|_2$ , can be recovered easily<sup>(3)</sup>.

Summing up, for any  $p, q$ , one simply defines  $\mathbf{R}_{\text{new}}, \mathbf{R}'_{\text{new}}$  as the average of all admissible geometry-based operators. The general expression of these operators is given in the Appendix A.1 and propositions 1 and 2 there give us that for any  $p, q$ ,  $\max(\|\mathbf{R}_{\text{new}}\|^2, \|\mathbf{R}'_{\text{new}}\|^2) = I_\alpha$ <sup>(3)</sup>.

Let us make some comments regarding this approach:

- this approach is optimal when the corner angle  $\alpha$  can be expressed as a rational fraction times  $2\pi$ , that is  $\alpha \in 2\pi\mathbb{Q}$ . Since  $\mathbb{Q}$  is dense in  $\mathbb{R}$ , one can theoretically come as close to  $\alpha$  as desired.
- in the case of a general polygonal interface  $\Sigma$ ,  $\Omega$  is *locally pattern-based* in a neighborhood of any interior corner. Consequently this approach can be adapted using a localization process (see next section).

At this point we have provided new operators that satisfy the first requirement (namely optimality). Is it possible also to satisfy the second requirement  $\mathbf{T}(V^h) \subset V^h$ ?

Suppose that  $\Omega$  is a pattern-based domain, and that the pattern is polygonal (for instance a triangle). Once this pattern is meshed, the rule to get  $\mathbf{T}$ -conforming meshes is to duplicate by symmetry this mesh in each pattern. When one considers a uniform degree of approximation on the whole domain, one has automatically  $\mathbf{T}(V^h) \subset V^h$  (resp.  $\mathbf{T}'(V^h) \subset V^h$ ) with the new operators described above. Then, as already mentioned, classical errors estimates directly follows from Céa's lemma [10, Corollary 1]. Note that there is no need for additional symmetry requirements for the meshing of the pattern.

Let us give some numerical results. We consider the geometry described previously in the second example. In that case  $\Omega^{\text{hex}}$  is a hexagon where  $\Omega_1^{\text{hex}}$  locally coincides with a cone of angle  $\alpha = 4\pi/3$ . In this case  $I_c = [-2; -1/2]$ . Let us construct an exact solution (of a problem of type (1) with  $\kappa_\sigma \notin I_c$ , see (12) below): consider  $u_r \in H^1(\Omega^{\text{hex}})$  such that in polar coordinates

$$u_r(\rho, \theta) = \begin{cases} \sigma_1^{-1} \rho^2 \sin(\frac{p+q}{2}(\theta-\alpha)) & \text{in } \Omega_1^{\text{hex}}, \\ \sigma_2^{-1} \rho^2 \sin(\frac{p+q}{2}(\theta-\alpha)) & \text{in } \Omega_2^{\text{hex}}, \end{cases}$$

where  $p$  (resp.  $q$ ) still denotes the number of patterns in  $\Omega_1^{\text{hex}}$  (resp.  $\Omega_2^{\text{hex}}$ ) and  $f := -\text{div}(\sigma \nabla u_r) = \frac{1}{4}((p+q)^2 - 16) \sin(\frac{p+q}{2}(\theta-\alpha)) \in L^2(\Omega^{\text{hex}})$ <sup>(4)</sup>. By construction,  $u_r$  is piecewise smooth [16]:  $u_r|_{\Omega_i^{\text{hex}}} \in H^{3-\varepsilon}(\Omega_i^{\text{hex}})$ ,  $\forall \varepsilon > 0$ ,  $i = 1, 2$ . To illustrate the importance of  $\mathbf{T}$ -conforming meshes around corners, let us add to  $u_r$  a singular part, that is some  $u_s(\rho, \theta) = \rho^\lambda \Phi(\theta)$ , with  $\lambda := \lambda(\sigma) \in \mathbb{R}$ , such that  $\text{div}(\sigma \nabla u_s) = 0$  in  $\Omega^{\text{hex}}$ . For example we consider [6, 18]

$$\Phi(\theta) = \begin{cases} \frac{\cosh(\lambda(\theta - \alpha/2))}{\cosh(\lambda\alpha/2)} & 0 \leq \theta \leq \alpha/2 \\ \frac{\cosh(\lambda(\theta + (2\pi - \alpha)/2))}{\cosh(\lambda(2\pi - \alpha)/2)} & -(2\pi - \alpha)/2 \leq \theta \leq 0 \end{cases},$$

$$\Phi(\theta) = \Phi(2\pi - \theta - \alpha/2) \quad -\pi \leq \theta \leq -(2\pi - \alpha) \text{ or } \alpha/2 \leq \theta \leq \pi.$$

<sup>3</sup>For instance by choosing  $v_1$  such that  $v_1^k \in H_0^1(\Omega_1^k)$ , for  $k = 1, p$ , with  $v_1^1$  a symmetric function w.r.t. the line  $\theta = \frac{\pi}{p+q}$ , and  $v_1^k(\rho, \theta) = v_1^1(\mathcal{R}_{k-1}(\rho, \theta))$ , for  $(\rho, \theta) \in \Omega_1^k$ , for  $k = 2, p$ .

<sup>4</sup>With this choice  $u_r(\cdot, \theta) = 0$  for  $\theta \in \{n \frac{2\pi}{p+q}, n \in \mathbb{Z}\}$ , then continuity is easily ensured at the crossing of the interfaces.

One chooses for  $\lambda$  the smallest possible positive real number. Its value depends on  $\kappa_\sigma$ , and one can prove that since  $\kappa_\sigma \notin I_c$ ,  $\lambda \neq 0$  [3], so  $u_s \in H^1(\Omega^{\text{hex}})$ , with  $u_s|_{\Omega_i^{\text{hex}}} \in H^{1+\lambda-\varepsilon}(\Omega_i^{\text{hex}})$ ,  $\forall \varepsilon > 0$ ,  $i = 1, 2$ . Defining  $g := (u_r + u_s)|_{\partial\Omega^{\text{hex}}}$ , one checks that  $u = u_r + u_s$  is the unique solution of the problem

$$\begin{cases} \text{Find } u \in H^1(\Omega^{\text{hex}}) \text{ such that:} \\ -\operatorname{div}(\sigma \nabla u) = f \quad \text{in } \Omega^{\text{hex}} \\ u = g \quad \text{on } \partial\Omega^{\text{hex}} \end{cases} . \quad (12)$$

With the help of a lifting of the non-zero boundary condition, one easily checks that problem (12) set in this hexagonal domain is well-posed when  $\kappa_\sigma \notin I_c = [-2; -1/2]$ . We consider two kinds of meshes (see figure 4), namely standard meshes (without rotation- or symmetry-based invariance) and T-conforming meshes, and discretize the problem for several FE orders, for two chosen contrasts:  $\kappa_\sigma = -3$  and  $\kappa_\sigma = -2.1$  (see figure 5). For this geometry, computations give  $\lambda = 0.5$  when  $\kappa_\sigma = -3$ , and  $\lambda = 0.2051664777$  when  $\kappa_\sigma = -2.1$ . Classically the regularity of the solution  $u = u_r + u_s$  is driven by the singular behavior  $u_s$ . Namely one finds that  $u_i$ ,  $i = 1, 2$ , belongs to  $H^{1+\lambda-\varepsilon}(\Omega_i^{\text{hex}})$ ,  $\forall \varepsilon > 0$ . Consequently one expects an order of convergence equal to  $\lambda$  for the relative errors in  $H_0^1$ -norm. With figure 5 one concludes that T-conforming meshes ensure optimal convergence speed while standard meshes are not stable with respect to the mesh size. Note that in figure 5 one does not improve the order of convergence using higher FE orders, due to the low regularity of the solution.

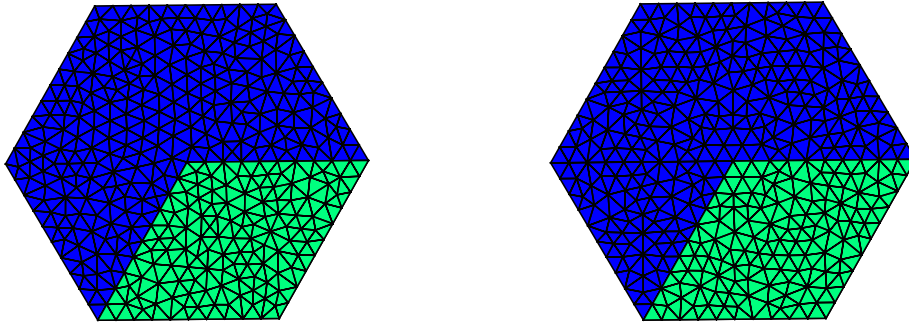


Figure 4: Left: standard mesh ( $h = 0.2$ , 385 nodes),  $\Omega_1^{\text{hex}}$  corresponds to blue region while  $\Omega_2^{\text{hex}}$  corresponds to the green one. Right: T-conforming mesh ( $h = 0.2$ , 379 nodes).

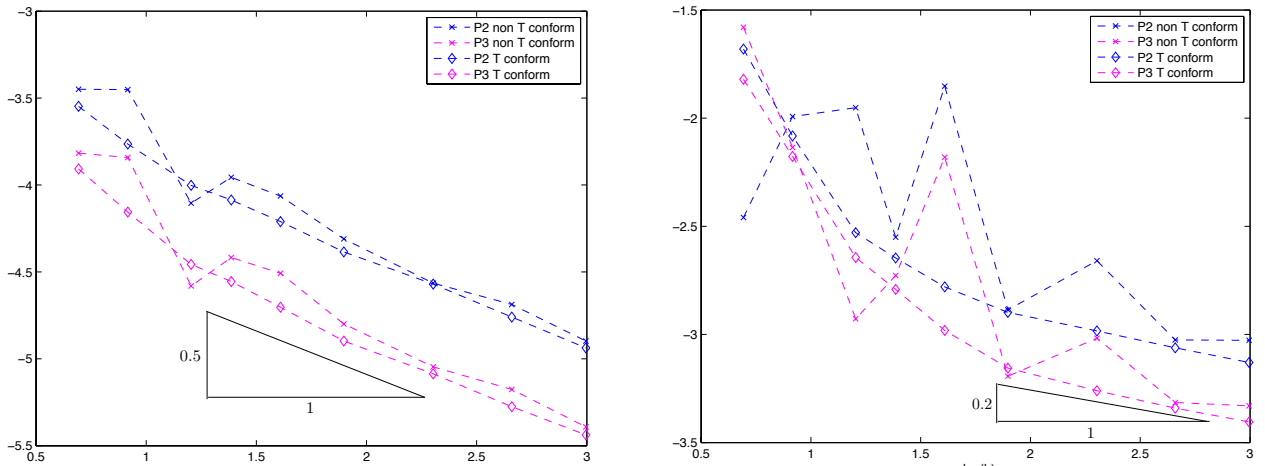


Figure 5: Left: relative error in  $H^1$  norm for different mesh sizes  $h$  (log-log scale) for  $\kappa_\sigma = -3$ . Right: relative error in  $H^1$  norm for different mesh sizes  $h$  (log-log scale) for  $\kappa_\sigma = -2.1$ .

In a more general case of a polygonal interface with several interior corners, one has to apply locally this *tilings* method in the neighborhood of the corners. This strategy is explained in the next two sections.

## 4 Weak T-coercivity for a general polygonal interface

In this section we recall some theoretical results regarding the T-coercivity approach, and prove well-posedness of problem (1) for an arbitrary geometry using the new geometry-based operators introduced in section 3. Consider a Hilbert space  $E$  with its dual  $E'$ , a bilinear form  $b$  defined over  $E \times E$  and  $B$  a (linear continuous operator) from  $E$  to  $E'$  such that  $\langle Bv, w \rangle = b(v, w)$ , for all  $v, w \in E$ . Then, for some data  $f \in E'$ , solving

$$\text{Find } u \in E \text{ such that } b(u, w) = \langle f, w \rangle, \quad \forall w \in E, \quad (13)$$

is equivalent to solving

$$\text{Find } u \in E \text{ such that } Bu = f \text{ in } E'. \quad (14)$$

Classically [17, 7], we recall that  $B$  is said to be a Fredholm operator when  $\dim(\ker(B)) < \infty$ , its range  $R(B)$  is closed and  $\text{codim}(R(B)) < \infty$ ; in this case its index is equal to  $\dim(\ker(B)) - \text{codim}(R(B))$ . If in addition the associated form  $b$  is hermitian, the index is automatically equal to 0. When  $B$  is a Fredholm operator of index 0, we say that (13)-(14) is *well-posed in the Fredholm sense*. On the other hand, (13)-(14) is *well-posed* if, and only if, for all  $f \in E'$ , it has one and only one solution  $u$ , with continuous dependence: there exists  $C > 0$  such that, for all  $f \in E'$ , the solution  $u$  verifies  $\|u\|_E \leq C\|f\|_{E'}$ . In terms of operators, it means that  $B^{-1}$  is well-defined as a continuous operator from  $E'$  to  $E$ .

To prove well-posedness (in the Fredholm sense) of problem (1), we will apply the theory of T-coercivity [5, 2, 10]. Let us recall some results. Within the Banach-Necas-Babuska framework, one can define a *weak stability condition*, also called an inf-sup condition.

**Definition 1.** Let  $b(\cdot, \cdot)$  be a continuous sesquilinear form on  $E \times E$ . It verifies a weak stability condition if

$$\exists \mathbb{C} \text{ a compact operator, } \exists \alpha' > 0, \beta' \in \mathbb{R}, \forall v \in E, \sup_{w \in E \setminus \{0\}} \frac{|b(v, w)|}{\|w\|_E} \geq \alpha' \|v\|_E - \beta' \|\mathbb{C}v\|_E. \quad (15)$$

Let us now introduce an *a priori* intermediate condition (cf. [5]).

**Definition 2.** Let  $b(\cdot, \cdot)$  be a continuous sesquilinear form on  $E \times E$ . It is weakly T-coercive if

$$\exists \mathbb{C} \text{ a compact operator, } \exists \mathbb{T} \text{ a bijective operator, } \exists \underline{\alpha} > 0, \underline{\beta} \in \mathbb{R}, \forall v \in E, |b(v, \mathbb{T}v)| \geq \underline{\alpha} \|v\|_E^2 - \underline{\beta} \|\mathbb{C}v\|_E^2. \quad (16)$$

In other words, the form  $b(\cdot, \mathbb{T}\cdot)$  fulfills a Gårding's inequality [21].

**Remark 1.** When  $\beta' \leq 0$  in (15), one recovers the classical stability condition. Respectively, when  $\underline{\beta} \leq 0$  in (16), one obtains T-coercivity for  $b(\cdot, \cdot)$ .

The operator  $\mathbb{T}$  introduced in (16) realizes the inf-sup condition (15): it is sometimes called an inf-sup operator. Let us state a simple result.

**Lemma 1.** Let  $b(\cdot, \cdot)$  be a continuous, sesquilinear hermitian form on  $E \times E$ . Then the three assertions below are equivalent:

- (i) (13)-(14) is well-posed in the Fredholm sense;
- (ii) the form  $b$  satisfies a weak stability condition;
- (iii) the form  $b$  is weakly T-coercive.

As seen in the introduction, problem (1) set in the Hilbert space  $V = H_0^1(\Omega)$  can be expressed in variational form as (3). Let us introduce the (continuous linear) operator  $A$  from  $V$  to  $V'$  such that  $\langle Av, w \rangle = a(v, w)$  for all  $v, w \in V$ , with the form  $a$  defined in (2): by construction,  $A : v \mapsto -\text{div}(\sigma \nabla v)$ . In operator form, (1)-(3) writes equivalently

$$\text{Find } u \in V \text{ such that } Au = f \text{ in } V'. \quad (17)$$

We propose below some realizations of the inf-sup operator that can be used for problem (1) for a polygonal interface  $\Sigma$ . From now on, we suppose for simplicity that  $\Sigma$  is a polygonal line without endpoints, ie. it is a loop. Let  $N$  denote the number of its corners,  $(c_n)_{n=1,N}$  its corners,  $(\alpha_n)_{n=1,N}$  the corner angles measured in  $\Omega_1$ , and  $(e_n)_{n=1,N}$  its edges. We introduce the polar coordinates  $(\rho_n, \theta_n)$  centered at  $c_n$  such that  $\Omega_1$  coincides locally with the cone  $\{(\rho_n \cos \theta_n, \rho_n \sin \theta_n) \mid 0 < \rho_n, 0 < \theta_n < \alpha_n\}$ .

First, we build a partition of unity on  $\bar{\Omega}$ . So, let  $(\chi_n)_{n=1,2N} \in (C^\infty(\bar{\Omega}; [0, 1]))^N$  with supports localized in a neighborhood of the interface, such that, for  $n = 1, N$ ,  $\chi_n = 1$  near the corner  $c_n$  and  $\chi_n = 0$  "far" from  $c_n$ ; respectively  $\chi_{n+N} = 0$  on  $\Sigma \setminus e_n$ . Plus, one defines  $(\chi_n)_n$  so that it holds  $\sum_{n=1,2N} \chi_n = 1$  on  $\Sigma$ . Then we define  $\chi_0 = 1 - \sum_{n=1,2N} \chi_n$ , which vanishes in a neighborhood of the interface  $\Sigma$  and, setting  $P = 2N$ , we obtain that  $(\chi_p)_{p=0,P} \in (C^\infty(\bar{\Omega}; [0, 1]))^P$  is a partition of unity on  $\bar{\Omega}$ . Finally we denote by  $S_p := \text{supp}(\chi_p)$  the support of the cutoff function  $\chi_p$ .

Now, let  $(\mathbf{R}_p)_{p=1,P}$  be (linear continuous) operators that act from  $V_1$  to  $V_2$  (resp.  $(\mathbf{R}'_p)_{p=1,P}$  from  $V_2$  to  $V_1$ ). We suppose *in addition* that these operators are such that, for all  $p = 1, P$ ,

$$\begin{aligned} \exists C, C' > 0, \forall w_1 \in V_1, \quad \|\mathbf{R}_p w_1\|_{\Omega_2} &\leq C \|w_1\|_{\Omega_1} \quad \text{and} \quad \|\chi_p^{1/2} \nabla(\mathbf{R}_p w_1)\|_{\Omega_2 \cap S_p} \leq C' \|\chi_p^{1/2} \nabla w_1\|_{\Omega_1 \cap S_p}, \\ \exists C, C' > 0, \forall w_2 \in V_2, \quad \|\mathbf{R}'_p w_2\|_{\Omega_1} &\leq C \|w_2\|_{\Omega_2} \quad \text{and} \quad \|\chi_p^{1/2} \nabla(\mathbf{R}'_p w_2)\|_{\Omega_1 \cap S_p} \leq C' \|\chi_p^{1/2} \nabla w_2\|_{\Omega_2 \cap S_p}. \end{aligned} \quad (18)$$

We introduce

$$\|\mathbf{R}\| := \max_{p=1,P} \|\mathbf{R}_p\|, \quad \|\mathbf{R}'\| := \max_{p=1,P} \|\mathbf{R}'_p\|, \quad \text{resp.} \quad |\mathbf{R}| := \max_{p=1,P} |\mathbf{R}_p|, \quad |\mathbf{R}'| := \max_{p=1,P} |\mathbf{R}'_p|, \quad (19)$$

where for all  $p = 1, P$ ,

$$\begin{aligned} \|\mathbf{R}_p\| &:= \sup_{w_1 \in V_1, \|\chi_p^{1/2} \nabla w_1\|_{\Omega_1 \cap S_p} = 1} \|\chi_p^{1/2} \nabla \mathbf{R}_p w_1\|_{\Omega_2 \cap S_p}, \quad |\mathbf{R}_p| := \sup_{w_1 \in L^2(\Omega_1), \|w_1\|_{\Omega_1} = 1} \|\mathbf{R}_p w_1\|_{\Omega_2}, \\ \|\mathbf{R}'_p\| &:= \sup_{w_2 \in V_2, \|\chi_p^{1/2} \nabla w_2\|_{\Omega_2 \cap S_p} = 1} \|\chi_p^{1/2} \nabla \mathbf{R}'_p w_2\|_{\Omega_1 \cap S_p}, \quad |\mathbf{R}'_p| := \sup_{w_2 \in L^2(\Omega_2), \|w_2\|_{\Omega_2} = 1} \|\mathbf{R}'_p w_2\|_{\Omega_1}. \end{aligned} \quad (20)$$

Finally, we assume matching conditions on the traces:

$$\forall p, \forall v_1 \in V_1, \quad \mathbf{R}_p v_1|_{\Sigma \cap S_p} = v_1|_{\Sigma \cap S_p}, \quad \forall p, \forall v_2 \in V_2, \quad \mathbf{R}'_p v_2|_{\Sigma \cap S_p} = v_2|_{\Sigma \cap S_p}. \quad (21)$$

**Remark 2.** *Later on at the end of the section, we will show that in practice we can provide  $\chi_p, \mathbf{R}_p, \mathbf{R}'_p$  which fulfill all the conditions above.*

Finally we define the two operators

$$\mathbf{T}v = \begin{cases} v_1 & \text{on } \Omega_1 \\ -v_2 + 2 \sum_{p=1,P} \chi_p \mathbf{R}_p v_1 & \text{on } \Omega_2 \end{cases}, \quad \mathbf{T}'v = \begin{cases} v_1 - 2 \sum_{p=1,P} \chi_p \mathbf{R}'_p v_2 & \text{on } \Omega_1 \\ -v_2 & \text{on } \Omega_2 \end{cases}. \quad (22)$$

Now one can prove the following

**Lemma 2.** *Suppose that  $\chi_p, \mathbf{R}_p, \mathbf{R}'_p$  satisfy (18), (19), (20) for all  $p = 1, P$ . If the contrast  $\kappa_\sigma$  does not belong to  $[-\|\mathbf{R}'\|^2; -1/\|\mathbf{R}\|^2]$ , then the form  $a$  is weakly T-coercive for  $\mathbf{T}$  defined in (22), and problem (1) is well-posed in the Fredholm sense.*

**Remark 3.** *As for any geometry the case  $\kappa_\sigma = -1$  is always ill-posed [14, 2], necessarily  $\|\mathbf{R}\| \geq 1, \|\mathbf{R}'\| \geq 1$ .*

*Proof.* We assume for instance that  $\kappa_\sigma \in (-1/\|\mathbf{R}\|^2; 0)$ . By lemma 1, we just have to show that the form  $a$  is weakly T-coercive, namely

$$\exists \mathbf{C} \text{ a compact operator, } \exists \mathbf{T} \text{ a bijective operator, } \exists \underline{\alpha} > 0, \underline{\beta} \in \mathbb{R}, \forall v \in V, \quad a(v, \mathbf{T}v) \geq \underline{\alpha} \|v\|_V^2 - \underline{\beta} \|\mathbf{C}v\|_V^2.$$

We consider operator  $\mathbf{T}$  defined in (22)-left to prove the above condition. Due to the matching conditions (21) satisfied by  $(\mathbf{R}_p)_p$ ,  $\mathbf{T}v \in V$  for all  $v \in V$  and, in addition one checks easily that  $\mathbf{T} \circ \mathbf{T} = \mathbf{I}_V$  so  $\mathbf{T}$  is bijective.

Then:

$$\begin{aligned}
a(v, \mathbf{T}w) &= |\sigma_1|(\nabla v_1, \nabla w_1)_{\Omega_1} + |\sigma_2|(\nabla v_2, \nabla w_2)_{\Omega_2} - 2|\sigma_2| \sum_{p=1, P} (\nabla v_2, \nabla(\chi_p \mathbf{R}_p w_1))_{\Omega_2 \cap S_p} \\
&= b(v, w) + c(v, w), \text{ where the forms } b \text{ and } c \text{ are respectively defined by:} \\
b(v, w) &= |\sigma_1|(\nabla v_1, \nabla w_1)_{\Omega_1} + |\sigma_2|(\nabla v_2, \nabla w_2)_{\Omega_2} - 2|\sigma_2| \sum_{p=1, P} (\nabla v_2, \chi_p \nabla(\mathbf{R}_p w_1))_{\Omega_2 \cap S_p}, \\
c(v, w) &= -2|\sigma_2| \sum_{p=1, P} (\nabla v_2, \mathbf{R}_p w_1 \nabla \chi_p)_{\Omega_2 \cap S_p}.
\end{aligned}$$

First we prove that  $b$  is coercive. Using Young's inequality with  $\eta > 0$  on  $b(v, v)$ , we get

$$b(v, v) \geq |\sigma_1| \|\nabla v_1\|_{\Omega_1}^2 + |\sigma_2| \|\nabla v_2\|_{\Omega_2}^2 - |\sigma_2| \sum_{p=1, P} \left( \eta \|\chi_p^{1/2} \nabla v_2\|_{\Omega_2 \cap S_p}^2 + \eta^{-1} \|\chi_p^{1/2} \nabla(\mathbf{R}_p v_1)\|_{\Omega_2 \cap S_p}^2 \right).$$

Then we find a first lower bound:

$$\begin{aligned}
-|\sigma_2| \eta \sum_{p=1, P} \|\chi_p^{1/2} \nabla v_2\|_{\Omega_2 \cap S_p}^2 &\geq -|\sigma_2| \eta \sum_{p=1, P} \|\chi_p^{1/2} \nabla v_2\|_{\Omega_2}^2 \\
&= -|\sigma_2| \eta \sum_{p=1, P} (\chi_p \nabla v_2, \nabla v_2)_{\Omega_2} \\
&= -|\sigma_2| \eta ((1 - \chi_0) \nabla v_2, \nabla v_2)_{\Omega_2} \\
&\geq -|\sigma_2| \eta \|\nabla v_2\|_{\Omega_2}^2, \text{ as } \|1 - \chi_0\|_{L^\infty(\Omega_2)} = 1.
\end{aligned}$$

Next, using the definitions of  $(\|\mathbf{R}_p\|)_{p=1, P}$  and  $\|\mathbf{R}\|$ , we find a second lower bound:

$$\begin{aligned}
-|\sigma_2| \eta^{-1} \sum_{p=1, P} \|\chi_p^{1/2} \nabla(\mathbf{R}_p v_1)\|_{\Omega_2 \cap S_p}^2 &\geq -|\sigma_2| \eta^{-1} \sum_{p=1, P} \|\mathbf{R}_p\|^2 \|\chi_p^{1/2} \nabla v_1\|_{\Omega_1 \cap S_p}^2 \\
&\geq -|\sigma_2| \eta^{-1} \|\mathbf{R}\|^2 \sum_{p=1, P} \|\chi_p^{1/2} \nabla v_1\|_{\Omega_1}^2 \\
&= -|\sigma_2| \eta^{-1} \|\mathbf{R}\|^2 ((1 - \chi_0) \nabla v_1, \nabla v_1)_{\Omega_1} \\
&\geq -|\sigma_2| \eta^{-1} \|\mathbf{R}\|^2 \|\nabla v_1\|_{\Omega_1}^2, \text{ as } \|1 - \chi_0\|_{L^\infty(\Omega_1)} = 1.
\end{aligned}$$

Hence,

$$b(v, v) \geq (|\sigma_1| - |\sigma_2| \eta^{-1} \|\mathbf{R}\|^2) \|\nabla v_1\|_{\Omega_1}^2 + |\sigma_2| (1 - \eta) \|\nabla v_2\|_{\Omega_2}^2.$$

We remark now that we can choose  $\eta > 0$  such that

$$(|\sigma_1| - |\sigma_2| \eta^{-1} \|\mathbf{R}\|^2) > 0 \quad \text{and} \quad |\sigma_2| (1 - \eta) > 0.$$

Indeed, the former condition is equivalent to  $\|\mathbf{R}\|^2 |\kappa_\sigma| < \eta$ . Because  $\kappa_\sigma \in (-1/\|\mathbf{R}\|^2; 0)$ , we can choose  $\eta < 1$  such that this condition is fulfilled.

Regarding  $c(v, v)$ , if we let  $\mathbf{C} : V \rightarrow V$  such that  $(\mathbf{C}v, w)_V = c(v, w)$  for all  $v, w \in V$ , then by Cauchy-Schwarz inequality one gets

$$\|\mathbf{C}v\|_V^2 = c(v, \mathbf{C}v) \leq \|\nabla v\|_\Omega G P \|\mathbf{R}\| \|\mathbf{C}v\|_\Omega,$$

where  $G := 2|\sigma_2| \max_{p=1, P} (\|\chi_p\|_{W^{1, \infty}(\Omega_2)})$ . By Rellich's theorem, one concludes that  $\mathbf{C}$  is compact. Then by using Young's inequality with  $\eta' > 0$  we find

$$c(v, v) \geq -\frac{1}{2} ((\eta')^{-1} \|\mathbf{C}v\|_V^2 + \eta' \|v\|_V^2),$$

which ends the proof for  $\eta'$  small enough.

On the other hand, if  $\kappa_\sigma \in (-\infty; -\|\mathbf{R}'\|^2)$ , one can reverse the roles of  $\Omega_1$  and  $\Omega_2$  by using this time the operator  $\mathbf{T}'$  defined in (22)-right. The proof then proceeds as above to prove that there exist  $\underline{\alpha}' > 0$ ,  $\underline{\beta}'$ , and a compact operator  $\mathbf{C}'$  such that

$$\forall v \in V, \quad a(v, \mathbf{T}'v) \geq \underline{\alpha}' \|v\|_V^2 - \underline{\beta}' \|\mathbf{C}'v\|_V^2,$$

which is the condition (16). □

To conclude the study of the well-posedness of problem (1), we let us explicit  $\chi_p$ ,  $(\mathbb{R}_p)_{p=1,P}$ , (resp.  $(\mathbb{R}'_p)_{p=1,P}$ ), and compute the bounds. In the case where  $\Sigma$  is a polygonal line with all angles  $\alpha_n \in 2\pi\mathbb{Q}$ ,  $n = 1, N$ , one can explicit these bounds using the results of §3. To do so, let us define  $(B_p)_{p=1,P}$  a sequence of connected open sets so that  $\bigcup_{p=1,P} B_p$  is a neighborhood of the interface. For all  $n = 1, N$ , define  $B_n$  as a triangle-based neighborhood of  $c_n$ , and  $B_{N+n}$  as a neighborhood of a part of  $e_n$  (excluding its endpoints) which is symmetric with respect to  $e_n$  (see figure 6). More precisely, for all  $n = 1, N$ ,  $B_n$  is a cyclic polygon centered at  $c_n$  composed of  $p_n > 0$  triangles in  $\Omega_1$  and  $q_n > 0$  triangles in  $\Omega_2$ : define  $\rho_{c_n}$  the radius and  $s_n$  the side length of this polygon. Then for all  $n = 1, N$  one defines  $B_{N+n}$  as a trapezoid-based open set with one trapezoid in  $\Omega_1$  and one in  $\Omega_2$ , each trapezoid being of side lengths  $s_n$  and  $s_{n+1}$ <sup>(5)</sup>. For technical purposes one chooses  $p_n \geq 2$  so that all part edge neighborhoods  $B_{N+n}$ ,  $n = 1, N$  are disjoint open sets. Note that a part edge neighborhood intersects with two corners neighborhoods. Namely for  $n = 1, N - 1$ , one has  $\overline{B_{N+n}} \cap \overline{B_n} \neq \emptyset$ ,  $\overline{B_{N+n}} \cap \overline{B_{n+1}} \neq \emptyset$ , and for  $n = N$  one has  $\overline{B_P} \cap \overline{B_N} \neq \emptyset$ ,  $\overline{B_P} \cap \overline{B_1} \neq \emptyset$ . Then one defines  $(\mathbb{R}_p)_{p=1,P}$  from  $V_1$  to  $V_2$  (and also from  $L^2(\Omega_1)$  to  $L^2(\Omega_2)$ ) such that for all  $w_1 \in V_1$  and for all  $n = 1, N$

$$\mathbb{R}_n w_1(x) = \begin{cases} 0 & \text{if } x \notin B_n \\ \mathbb{R}_{\text{new}}^n w_1(x) & \text{if } x \in B_n \end{cases}, \quad (23)$$

with  $\mathbb{R}_{\text{new}}^n$  defined as in §3 (the general expression is given in Appendix A.1), and

$$\mathbb{R}_{N+n} w_1(x) = \begin{cases} 0 & \text{if } x \notin B_{N+n} \\ w_1(x_\Sigma, -y_\Sigma) & \text{if } x \in B_{N+n} \end{cases}, \quad (24)$$

where  $(x_\Sigma, y_\Sigma)$  denotes the local cartesian coordinates, and the first axis coincides with  $\Sigma$ .

Finally, let us precise some properties the cutoff functions  $(\chi_p)_p$ . For all  $p = 1, N$ ,  $\chi_p(x) = \chi_p(\rho_p)$ <sup>(6)</sup>, and  $\chi_p = 1$  for  $0 \leq \rho_p \leq \rho_{\min,p}$  for some  $\rho_{\min,p} > 0$ ; respectively  $\chi_{N+p}$  is a symmetric function with respect to the interface. This can be realized by introducing  $\psi \in C^\infty(\overline{\Omega}; [0, 1])$  equal to 1 in a neighborhood of the interface, for which we remark that  $\psi - \sum_{p=1,N} \chi_p$  is a (smooth) function that vanishes in a neighborhood of the corners, and also that "close" to the interface, it is a function which is symmetric with respect to the interface<sup>(7)</sup>. Then one chooses  $S_p$  the support of  $\chi_p$  such that  $S_p \subset B_p$  for the corner case ( $p = 1, N$ )<sup>(8)</sup> and in the edge case  $S_p \subset \overline{B_p} \cup \overline{B_{p-N}} \cup \overline{B_{p+1-N}}$  ( $p = N + 1, P - 1$ ), resp.  $S_P \subset \overline{B_P} \cup \overline{B_N} \cup \overline{B_1}$ . One can check that (21) is satisfied for all  $\mathbb{R}_p$ ,  $p = 1, P$ . We can finally state the result on the well-posedness of problem (1).

<sup>5</sup> $B_{2N}$  is composed of two trapezoids of side lengths  $s_{2N}$  and  $s_1$ .

<sup>6</sup>Recall that  $(\rho_p, \theta_p)$  denote the polar coordinates centered at  $c_p$ ,  $p = 1, N$ .

<sup>7</sup>The strictly positive, upper limit distance up to which this property applies is a function of  $\min_p \rho_{\min,p}$  and of  $(\alpha_p)_{p=1,N}$ .

<sup>8</sup>In accordance with the previous notations, one chooses  $\rho_{\min,p}$  such that  $\rho_{\min,p} < \rho_{c_p}$ .

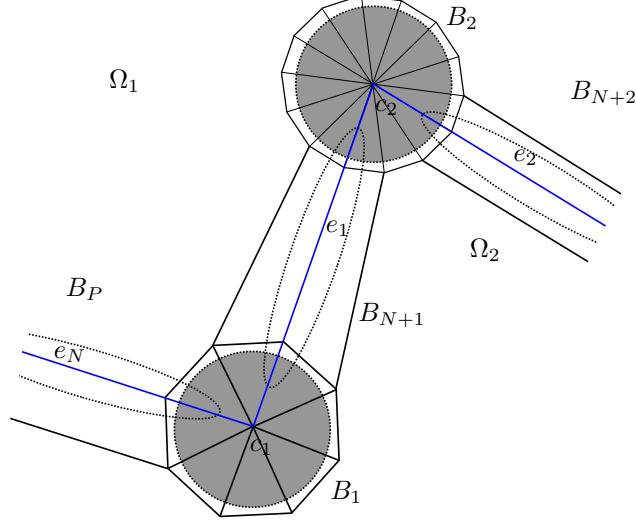


Figure 6: Notations around two corners  $c_1$  and  $c_2$ : here  $p_1 = 2$ ,  $q_1 = 6$ ,  $p_2 = 11$  and  $q_2 = 3$ . The gray zones represent the support of  $\chi_p$  while the empty-dashed domains represent the support of  $\chi_{N+p}$ ,  $p = 1, N$ .

**Theorem 1.** *Assume that all the corners' angles of the interface belong to  $2\pi\mathbb{Q}$ . If  $\kappa_\sigma \notin [-I_\alpha; -1/I_\alpha]$ , with  $I_\alpha := \max_{n=1, N} \max(\frac{2\pi-\alpha_n}{\alpha_n}, \frac{\alpha_n}{2\pi-\alpha_n})$ , then the form  $a$  is weakly T-coercive and problem (1) is well-posed in the Fredholm sense.*

*Proof.* With the operators (23)-(24) and using propositions 3 and 4 (cf. Appendix A.1), one finds that  $\min(\|\mathbf{R}\|^2, \|\mathbf{R}'\|^2) \geq 1$  and  $\max(\|\mathbf{R}\|^2, \|\mathbf{R}'\|^2) \leq I_\alpha$ . Consequently  $[-\|\mathbf{R}'\|^2; -1/\|\mathbf{R}\|^2] \subseteq [-I_\alpha; -1/I_\alpha]$ . Then if  $\kappa_\sigma \notin [-I_\alpha; -1/I_\alpha]$  the result follows from lemma 2.  $\square$

Note that if  $\kappa_\sigma \in [-I_\alpha; -1/I_\alpha]$ , then the problem (1) is not well-posed in the Fredholm sense, according to section 6 of [2]. Hence, we conclude that the condition on the contrast  $\kappa_\sigma \notin [-I_\alpha; -1/I_\alpha]$  is optimal for a 2D geometry with a polygonal interface separating the two subdomains.

## 5 T-conforming meshes for a general polygonal interface and error estimates

We study the approximation of problem (1) when the contrast  $\kappa_\sigma$  does not belong to the critical interval  $[-I_\alpha; -1/I_\alpha]$ . According to the previous section (see theorem 1), it follows that problem (1) is well-posed in the Fredholm sense (namely there might be a finite dimensional kernel). From now on, we impose further that it is well-posed, to ensure the uniqueness of the solution  $u$  to be approximated. We study the family of discrete problems (4), focusing on conforming<sup>(9)</sup> finite element approximations. Our aim is to prove that they can be solved in geometries with interfaces that include corners, and to prove optimality of the approximation, that is the unique solution  $u^h$  tends towards the exact solution  $u$  when  $h$  tends to 0 with optimal convergence rate.

We assume throughout §5 that  $\kappa_\sigma \in (-1/I_\alpha; 0)$  so that we can use the operator  $\mathbf{T}$  as defined in (22)-left. The conforming approximations in  $V$  are defined on meshes that match with the interface (all elements are either subsets of  $\overline{\Omega_1}$  or  $\overline{\Omega_2}$ ), with piecewise polynomial approximations. Let us consider  $(\mathcal{T}_h)_h$  a regular family of meshes of  $\overline{\Omega}$ , made of triangles; for all triangles  $\tau$ , one has either  $\tau \subset \overline{\Omega_1}$  or  $\tau \subset \overline{\Omega_2}$ . Define the family of Lagrange FE spaces (for a degree of approximation  $d \geq 1$ )

$$V_{(d)}^h := \{v \in V : v|_\tau \in \mathbb{P}_d(\tau), \forall \tau \in \mathcal{T}_h\},$$

<sup>9</sup>A Discontinuous Galerkin approach has recently been studied in [11].

where  $\mathbb{P}_d(\tau)$  is the space of polynomials of degree at most  $d$  on the triangle  $\tau$ . Unless otherwise specified, the index  $(d)$  is omitted. The parameter, or meshsize,  $h$  is defined as  $h := \max_{\tau \in \mathcal{T}_h} h_\tau$ , where  $h_\tau$  is the diameter of  $\tau$ . We recall that  $V_1^h$  (resp.  $V_2^h$ ) denotes the FE subspace of  $V_1$  (resp.  $V_2$ ) built on the triangulation  $\mathcal{T}_h$ . Throughout this section,  $C$  is used to denote a generic positive constant which is independent of the meshsize, the triangulation, and the data/unknown of interest.

**Definition 3.** Given  $h$ , let  $\mathcal{T}_{h,i}^p := \{\tau \in \mathcal{T}_h : \tau \cap \text{int}(S_p) \cap \Omega_i \neq \emptyset\}$ , for  $i = 1, 2$  and  $p = 1, P$ .

The meshes  $(\mathcal{T}_h)_h$  are locally T-conform if, for all  $h \lesssim 1$ , for all  $p = 1, P$ , for all  $\tau \in \mathcal{T}_{h,1}^p$ , the image of  $\tau$  by  $\mathbb{R}_p$  belongs to  $\mathcal{T}_{h,2}^p$ .

Note that if the conditions of the above definition are met, then for all  $v^h \in V^h$ , it holds that  $\mathbb{R}_p v_1^h$  is equal to the restriction of some element of  $V_2^h$  on  $\Omega_2 \cap S_p$ . We set  $\Omega_2^h := \text{int}(\bigcup_{p=1,P} \bigcup_{\tau \in \mathcal{T}_{h,2}^p} \tau)$ .

In general, for  $v^h \in V^h$ ,  $\mathbb{T}v^h \notin V^h$  due to the cutoff functions  $(\chi_p)_{p=1,P}$  that need to be interpolated. Then our goal is to prove optimality of the approximation, using some discrete operators  $(\mathbb{T}_h)_h$  (precised below). In [10] it was proposed to replace  $\chi_p$  by its interpolation of degree 1  $\chi_{p,h}$ . In our case, it leads to consider (10) with an additional summation  $\sum_{p=1,P}$  to fit with the definition of T in (22)-left:

$$\mathbb{T}_h^{(3)} v^h = \begin{cases} v_1^h & \text{in } \Omega_1 \\ -v_2^h + 2 \sum_{p=1,P} \chi_{p,h} \mathbb{R}_p v_1^h & \text{in } \Omega_2 \end{cases} \quad (25)$$

Unfortunately, for all  $p = 1, P$ , the degree of  $\chi_{p,h} \mathbb{R}_p v_1^h$  (restricted to a triangle) is too high. For instance, if one is dealing with Lagrange FE of degree 1 ( $d = 1$ ), its degree is 2: hence,  $\mathbb{T}_h^{(3)} v^h$  does not belong to  $V^h$  in general. This is why we introduce instead the discrete operator:

$$\mathbb{T}_h v^h = \begin{cases} v_1^h & \text{in } \Omega_1 \\ -v_2^h + 2I_h(\sum_{p=1,P} \chi_{p,h} \mathbb{R}_p v_1^h) & \text{in } \Omega_2 \end{cases} \quad (26)$$

with  $I_h$  the interpolation operator on  $V_{(d)}^h$ . Moreover from [15, Corollaries 1.109-1.110] we have for  $p = 1, P$  and for  $h$  small enough, the stability and approximation estimates

$$\exists C > 0, \|\chi_{p,h}\|_{W^{1,\infty}(\Omega_2)} \leq C \|\chi_p\|_{W^{1,\infty}(\Omega_2)}, \quad (27)$$

$$\exists C > 0, \|\chi_p - \chi_{p,h}\|_{W^{1,\infty}(\Omega_2)} \leq C h |\chi_p|_{W^{2,\infty}(\Omega_2)}. \quad (28)$$

We emphasize that the analysis presented in [10] remains valid if one replaces (the generalization of)  $\mathbb{T}_h^{(3)}$  by  $\mathbb{T}_h$ , as we explain below. Our aim is to prove that the form  $a$  is *uniformly weakly T-coercive*, namely

$$\exists \underline{\alpha}' > 0, \underline{\beta} \in \mathbb{R}, \forall h \text{ small enough}, \forall v^h \in V^h, |a(v^h, \mathbb{T}_h v^h)| \geq \underline{\alpha}' \|v^h\|_V^2 - \underline{\beta} \|Cv^h\|_V^2, \quad (29)$$

and from there one can follow the last part of the proof of Proposition 3 in [10] to derive a uniform discrete inf-sup condition (by contradiction). Classically, this yields an error estimate.

In order to obtain (29), we shall evaluate  $|a(v^h, (\mathbb{T} - \mathbb{T}_h)v^h)|$  for all  $v^h \in V^h$ . More precisely we want to prove

$$\exists C_0 > 0, \forall h \text{ small enough}, \forall v^h \in V^h, |a(v^h, (\mathbb{T} - \mathbb{T}_h)v^h)| \leq C_0 h \|v^h\|_V^2, \quad (30)$$

which is a consequence of proving that there exists  $C > 0$  such that, for all  $h$  small enough, for all  $v^h \in V^h$ ,  $\|(\mathbb{T} - \mathbb{T}_h)v^h\|_V \leq C h \|v^h\|_V$ . This result is the object of lemma 3.

**Lemma 3.** Assume that the meshes are locally T-conform then there holds

$$\exists C > 0, \forall h \text{ small enough}, \forall v^h \in V^h, \|(\mathbb{T} - \mathbb{T}_h)v^h\|_V \leq C h \|v^h\|_V, \quad (31)$$

and the form  $a$  is *uniformly weakly T-coercive*.



*Proof.* Given  $v^h \in V^h$ , let us apply the triangle inequality:

$$\|(\mathbf{T} - \mathbf{T}_h)v^h\|_V \leq \|(\mathbf{T} - \mathbf{T}_h^{(3)})v^h\|_V + \|(\mathbf{T}_h^{(3)} - \mathbf{T}_h)v^h\|_V.$$

Using the continuity of  $(\mathbf{R}_p)_{p=1,P}$  and (28), one has for the first term of the right-hand side:

$$\|(\mathbf{T} - \mathbf{T}_h^{(3)})v^h\|_V \leq C \sum_{p=1,P} \|\chi_p - \chi_{p,h}\|_{W^{1,\infty}(\Omega_2)} \|v^h\|_V \leq C h \|v^h\|_V.$$

On the other hand, for the last term of the right-hand side setting  $w_{p,2}^h = \mathbf{R}_p v_1^h$ ,  $p = 1, P$ , one finds

$$\|(\mathbf{T}_h^{(3)} - \mathbf{T}_h)v^h\|_V \leq 2 \sum_{p=1,P} \|\nabla(\chi_{p,h} w_{p,2}^h - I_h(\chi_{p,h} w_{p,2}^h))\|_{\Omega_2^h}.$$

The last term is technical to bound due to the presence of  $\chi_{p,h}$  (see details in the Appendix A.3). Since the meshes are locally  $\mathbf{T}$  conform, one has  $w_2^h \in V^h$ , then using lemma 5 in §A.3 and finally the continuity of the operators  $(\mathbf{R}_p)_{p=1,P}$ , one finds

$$\|(\mathbf{T}_h^{(3)} - \mathbf{T}_h)v^h\|_V \leq C h \|v^h\|_V,$$

which leads to (31), and thus to (30). It now follows that

$$\exists \underline{\alpha} > 0, \underline{\beta} \in \mathbb{R}, \forall h \lesssim 1, \forall v^h \in V^h, |a(v^h, \mathbf{T}_h v^h)| \geq (\underline{\alpha} - C_0 h) \|v^h\|_V^2 - \underline{\beta} \|C v^h\|_V^2,$$

with  $\underline{\alpha}$  and  $\underline{\beta}$  constants that express the weak  $\mathbf{T}$ -coercivity of  $a(\cdot, \cdot)$  with the operator  $\mathbf{T}$  as in (22), see lemma 2 and theorem 1. Hence, one finds the desired result (29), which in turns yields the error estimate (cf. [10]).  $\square$

**Theorem 2.** *Assume that problem (1) is well-posed and that the meshes are locally  $\mathbf{T}$ -conform. Then for  $h$  small enough there exists one and only one solution  $u^h$  to (4), with the estimate*

$$\|u - u^h\|_V \leq C \inf_{v^h \in V^h} \|u - v^h\|_V. \quad (32)$$

We conclude that we have optimality of the approximation for problem (1), thanks to the choice of the discrete  $\mathbf{T}$ -coercivity operator as in (25), assuming only that the meshes are locally  $\mathbf{T}$ -conform.

## 6 Numerical experiments

We are now in position to present some numerical illustrations. Consider the case of a squared cavity with a triangular inclusion. Define the square  $\Omega := (-4; 4) \times (-4; 4)$ :  $\Omega_2$  is an equilateral triangle of height 4 (the center of gravity  $G$  of  $\Omega_2$  is located at  $(x_1, x_2) = (0, -1/3)$ ) and  $\Omega_1 = \Omega \setminus \overline{\Omega_2}$ . We want to approximate problem (1) set in  $\Omega$  with a constant right-hand side  $f$ :

$$\begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that:} \\ -\operatorname{div}(\sigma \nabla u) = 1 \quad \text{in } \Omega \end{cases}.$$

Denoting  $(c_n)_{n=1,3}$  the corners of  $\Sigma$ , then  $\Omega_1$  locally coincides near  $c_n$ ,  $n = 1, 3$  with the cone of aperture  $\alpha := 5\pi/3$ . For this configuration one finds  $I_\alpha := 5$ , and with theorem 1 one can prove that if  $\kappa_\sigma \notin [-5; -1/5]$  then problem (1) (with the data  $f = 1$ ) is well-posed in the Fredholm sense. Then in order to define locally  $\mathbf{T}$ -conforming meshes, proceeding as in sections 3 and 5, one defines first a neighborhood of the interface  $\Sigma$  made of three polygonal neighborhoods (one for each corner) where one applies the ad hoc symmetry- and rotation-based operator (23), and three trapezoid-based neighborhoods (one for each edge) where one applies the symmetry-based operator (24). The associated locally  $\mathbf{T}$ -conforming mesh is deduced by meshing one pattern of each neighborhood, and duplicating it by symmetry (see figure 7 (right)): in this configuration we choose  $p_1 = p_2 = p_3 = 10$  and  $q_1 = q_2 = q_3 = 2^{(10)}$ .

<sup>10</sup>As explained in section 4, we choose  $q, p \geq 2$  to avoid the neighborhoods of the straight parts of  $\Sigma$  to overlap.

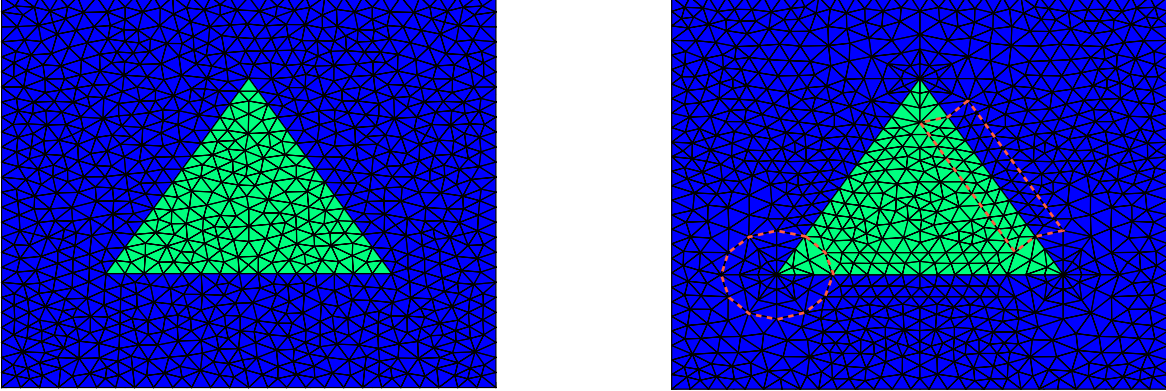


Figure 7: Left: example a standard triangulation of  $\Omega_h$  (1430 nodes). Right: example of locally T-conforming mesh (1424 nodes). The dashed lines enclose regions where we built T-conforming meshes near a corner, respectively a straight line.

In figure 8 are represented the relative errors in  $L^2$ -norm for several standard meshes (figure 7 (left)) and several T-conforming meshes (figure 7 (right)), for a contrast  $\kappa_\sigma = -5.2$ , and for several orders of approximation ( $d = 1, 2, 3$ ). Note that we do not compare the discrete solution with the exact solution. As the optimality of the approximation (see theorem 2) with T-conforming meshes is satisfied, one can compare with the discrete solution obtained on the finest mesh and one can check in figure 8 that convergence is ensured with a monotonic behavior of the error. On the other hand, with standard meshes the simulation gives unsatisfying results with a slightly erratic behavior of the error. Numerical results illustrate the importance of using T-conforming meshes to guarantee convergence of the numerical method.

Due to the symmetry of the geometry (and the right-hand side), one is expecting a symmetric solution with

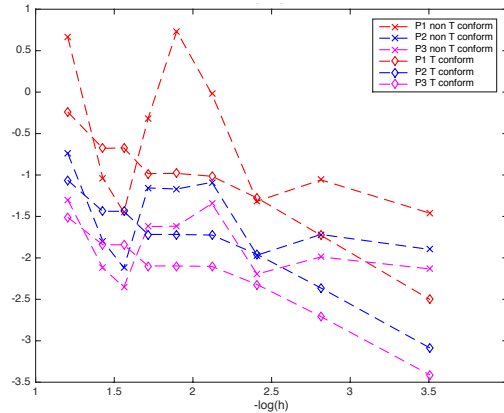


Figure 8: Relative error in  $L^2$ -norm for different mesh sizes  $h$  (log-log scale) for  $\kappa_\sigma = -5.2$ .

respect to  $Ox_2$ . One can check that, even for a refined, but standard mesh, the discrete solution does not satisfy this symmetry principle (see figure 9). This also explains why the approximation considering standard meshing does not converge so well.

Let us make some remarks regarding the convergence's orders. First let us mention that problem (1) with  $f = 1$  might have a non trivial kernel as we only proved Fredholmness. Nonetheless numerical computations did not present non trivial solutions of the homogeneous problem. With [10, Proposition 2] the regularity of the solution  $u$  of problem (1) is such that, away from the corners,  $u_i$ ,  $i = 1, 2$  is piecewise- $H^2$  whereas near the corners there exists a singularity exponent  $s \in (0, 1]$  such that  $u_i$  is only guaranteed to be piecewise- $H^{1+s}$

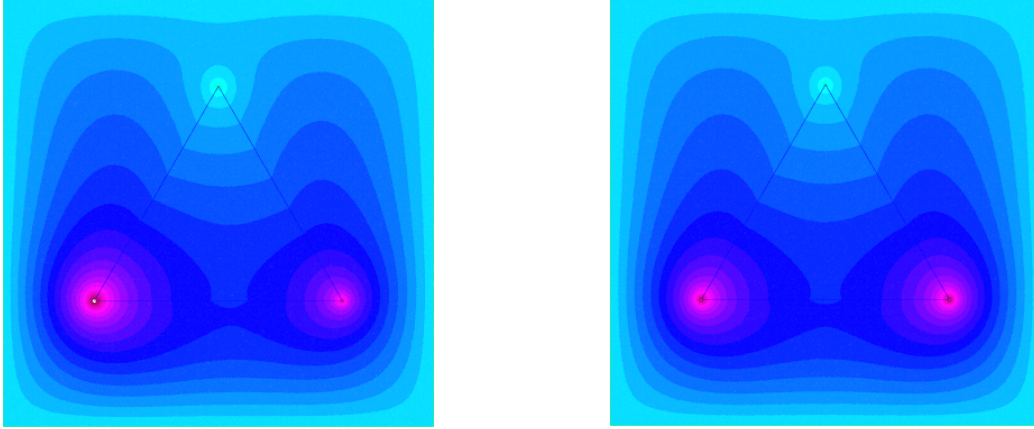


Figure 9: Left: solution  $u_h$  obtained with P3 finite elements on the finer standard mesh (101576 nodes). Right: solution  $u_h$  obtained with P3 finite elements on the finer T-conforming mesh (101716 nodes).

due to the presence of a singular part here (cf. §3). Then one has the estimates

$$\|u - u_h\|_V \leq C h^s, \quad \text{and} \quad \|u - u_h\|_\Omega \leq C h^{2s}.$$

The last estimate is obtained with the so-called Aubin-Nitsche lemma [15]. Note that in absence of the singular part, one expects that  $s = 1$ .

For  $\kappa_\sigma = -5.2$  computations yield a singularity exponent  $s = 0.13$ , whereas the results of figure 8 show an average convergence's order of 0.9 (that is  $s = 0.45$ ). This higher convergence order is due to the fact that we did not compare the discrete solution with the exact solution.

## 7 Concluding remarks

In this paper we extended the approximation theory for transmission problems with sign-changing coefficients with the T-coercivity approach. The novelty lies in the treatment of polygonal interfaces. This construction can be directly applied to the discrete problem by designing T-conforming meshes, which allows one to guarantee convergence of conforming finite element methods as soon as the continuous problem is well-posed in the Fredholm sense. Below, we list some possible applications of our work.

- One can study the transmission problem in the time-harmonic regime:

$$\begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that:} \\ \text{div}(\sigma \nabla u) + \omega^2 \zeta u = f \text{ in } \Omega \end{cases}, \quad (33)$$

at the frequency  $\omega \neq 0$ , with  $\zeta \in \mathbb{R}$ . Indeed, the added term  $\omega^2 \zeta u$  is a compact perturbation, so one concludes that when  $\kappa_\sigma \notin I_c$ , problem (33) is well-posed in the Fredholm sense and, when it is well-posed, one obtains convergence of the cFE method using locally T-conforming meshes.

- One can extend the results on problems (1) and (33) to the transmission problem with (one or two) sign-changing coefficients and boundary conditions other than Dirichlet boundary condition (see for instance [9, Chap.1 §1.7] for changes).
- The previous proofs can be easily adapted to 2D geometries with a curved boundary or a curved interface. In particular, the case of the curved interface can be covered with the help of Theorem 3.1-1 of [12]. For an interface without corners, well-posedness in the Fredholm sense is established as soon as  $\kappa_\sigma \neq -1$  (see [8]). We refer to [14] for the first relation – and proof with the help of integral equations – of this result. Results for the optimality of the approximation holds using for instance isoparametric quadrilateral FE (see [8]).

- One may apply similar techniques to a transmission problem with sign-changing coefficients in 3D geometries with straight or curved boundary and interface. However the optimality of the T-coercivity approach has yet to be established in the general case [2].
- Once the well-posedness of problem (33) is established, one can easily solve the associated eigenvalue problem

$$\begin{cases} \text{Find } (u, \lambda) \in H_0^1(\Omega) \setminus \{0\} \times \mathbb{C} \text{ such that:} \\ -\text{div}(\sigma \nabla u) = \lambda u \text{ in } \Omega \end{cases}, \quad (34)$$

and derive classical error estimates for the eigenvalues' approximation by following [19] (see [8]).

- Let us mention that a posteriori error estimates for diffusion problems with sign-changing coefficients have been carried out in [13] and lead to similar meshing requirements near the interfaces.
- On the other hand, the study of problem (1) when  $\kappa_\sigma$  belongs to the critical interval cannot be handled as previously, due to the appearance of singularities [20, 6] that do not belong to  $H^1(\Omega)$ . This problem has been investigated in [4] for particular geometries: it requires a new functional framework explicitly taking into account the singularities. For a general polygonal interface, it is proposed in [3] a new numerical approximation method which relies on the use of Perfectly Matched Layers at the corners.

## References

- [1] S. Anantha Ramakrishna, *Physics of negative refractive index materials*, Rep. Prog. Phys., **68** (2005), pp. 449–521.
- [2] A.-S. Bonnet-Ben Dhia, L. Chesnel, P. Ciarlet Jr., *T-coercivity for scalar interface problems between dielectrics and metamaterials*, Math. Mod. Num. Anal., **46** (2012), pp. 1363–1387.
- [3] A.-S. Bonnet-Ben Dhia, C. Carvalho, L. Chesnel, P. Ciarlet Jr., *On the use of Perfectly Matched Layers at corners for scattering problems with sign-changing coefficients*, in revision.
- [4] A.-S. Bonnet-Ben Dhia, L. Chesnel, X. Claeys, *Radiation condition for a non-smooth interface between a dielectric and a metamaterial*, Math. Model. Meth. App. Sci., **23** (2013), pp. 1629–1662.
- [5] A.-S. Bonnet-Ben Dhia, P. Ciarlet Jr., C.M. Zwölf, *Time harmonic wave diffraction problems in materials with sign-shifting coefficients*, J. Comput. Appl. Math., **234** (2010), pp. 1912–1919, Corrigendum p. 2616.
- [6] A.-S. Bonnet-Ben Dhia, M. Dauge, K. Ramdani, *Analyse spectrale et singularités d'un problème de transmission non coercif*, C. R. Acad. Sci. Paris, Ser. I, **328** (1999), pp. 717–720.
- [7] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.
- [8] C. Carvalho, *Étude mathématique et numérique de structures plasmoniques avec des coins*, PhD. thesis, École Polytechnique, 2015.
- [9] L. Chesnel, *Étude de quelques problèmes de transmission avec changement de signe. Application aux métamatériaux*, PhD. thesis, École Polytechnique, 2012.
- [10] L. Chesnel, P. Ciarlet Jr., *T-coercivity and continuous Galerkin methods: application to transmission problems with sign changing coefficients*, Numer. Math., **124** (2013), pp. 1–29.
- [11] E.T. Chung, P. Ciarlet Jr., *A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials*, J. Comput. Appl. Math., **239** (2013), pp. 189–207.
- [12] P.G. Ciarlet, *Mathematical Elasticity, Vol. III : Theory of Shells*, North-Holland, Amsterdam, 2000.
- [13] P. Ciarlet Jr., M. Vohralik, *Robust a posteriori error control for transmission problems with sign-changing coefficients using localization of dual norms*, hal-01148476, 2015.

- [14] M. Costabel, E. Stephan, *A direct boundary integral method for transmission problems*, J. of Math. Anal. and Appl., **106** (1985), pp. 367–413.
- [15] A. Ern, J.-L. Guermond, *Theory and practice of finite elements*, Springer-Verlag, Berlin, 2004.
- [16] P. Grisvard, *Singularities in boundary value problems*, Masson, RMA **22**, Paris, 1992.
- [17] W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.
- [18] S. Nicaise, J. Venel, *A posteriori error estimates for a finite element approximation of transmission problems with sign changing coefficients*, J. Comput. Appl. Math., **235** (2011), pp. 4272–4282.
- [19] J.E. Osborn, *Spectral approximation for compact operators*, Math. Comp., **29** (1975), pp. 712–725.
- [20] K. Ramdani, *Lignes supraconductrices : analyse mathématique et numérique*, Ph.D. thesis, Université Paris 6, 1999.
- [21] A.H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., **28** (1974), pp. 959–962.
- [22] D.R. Smith, J.B. Pendry, M.C.K. Wiltshire, *Metamaterials and negative refractive index*, Science, **305** (2004), pp. 788–792.

## A Appendix

### A.1 General construction of $\mathbf{R}$ around corners

In this section we generalize the tilings method presented in section 3 to any corner of angle  $\alpha \in 2\pi\mathbb{Q}$ . Recall that we define  $\alpha = 2\pi\frac{p}{p+q}$  in  $\Omega_1$ , with  $p, q > 0$ ,  $p \neq q$  and  $p + q$  even. Proceeding as in section 3, one builds several admissible rotation- and symmetry-based operators  $\mathbf{R}^{\text{adm}}$ , and then take the average of all admissible operators to obtain the desired result (that is operators  $\mathbf{R}$  and  $\mathbf{R}'$  with the same minimal norm as in [2]). We propose  $\min(p, q)$  admissible operators below based on a simple algorithm.

Consider for instance that  $p < q$ . One constructs  $p$  admissible operators  $(\mathbf{R}_i^{\text{adm}})_{i=1,p}$  from  $V_1$  to  $V_2$ , the  $i$ -th operator being obtained by (see figure 10 for an illustration):

- 1) in  $\Omega_2^{q+1-k}$ ,  $k = 1, i$ : apply  $\mathbf{S}^1$  to  $v_1^k$ ;
- 2) in  $\Omega_2^{p+1-k}$ ,  $k = i, p$ : apply  $\mathbf{S}^2$  to  $v_1^k$ ;
- 3) in  $\Omega_2^l$ ,  $l \in I := \llbracket p+2-i, q-i \rrbracket$ : from  $l = p+2-i$  to  $l = q-i$ , apply  $\mathbf{R}_{i-(p+l)}$  to  $v_1^i$ , then  $\mathbf{S}^2 \circ \mathbf{R}_{p+1-i-(l+1)}$  to  $v_1^i$ , update  $l \rightarrow l+2$  and so on. In other words, alternatively apply a rotation-based operator and a rotation+symmetry-based operator to  $v_1^i$ .

At step 3), since  $q + p$  is even ( $|I| = q - p - 1$  is odd), one always finishes by  $\mathbf{R}_{2i-(p+q)}$  (which ensures continuity of  $\mathbf{R}_i^{\text{adm}}v_1$  on  $\partial\Omega_2^{q+1-i} \cap \partial\Omega_2^{q-i}$ ).

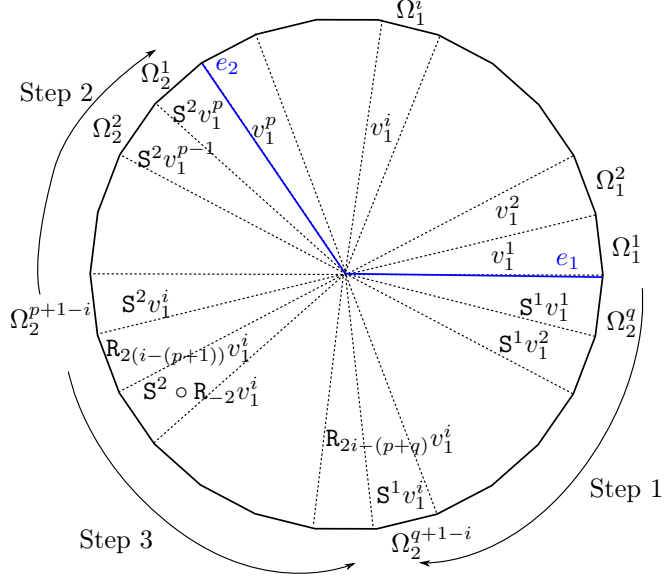


Figure 10: Scheme representing steps 1)-2)-3) for building  $R_i^{\text{adm}}$ .

One constructs  $p$  admissible operators  $(R_i^{\text{adm}'})_{i=1,p}$  from  $V_2$  to  $V_1$  similarly, the  $i$ -th operator being obtained by:

- 4) in  $\Omega_1^k$ ,  $k = 1, \min(i, p-1)$ : apply  $S^1$  to  $v_2^{q+1-k}$ ;
- 5) in  $\Omega_1^k$ ,  $k = \min(i+1, p), p$ : apply  $S^2$  to  $v_2^{p+1-k}$ ;
- 6) in  $\Omega_1^i$ , if  $i < p$ , add up the remaining contributions  $(v_2^l)_{l \in I'}$ , with  $I' := \llbracket p+1-i, q-i \rrbracket$ : for  $l = p+1-i$  to  $l = q-i$ , start by  $S^2 \circ R_{l-(p+1-i)}$  then apply  $-R_{(l+1)+p-i}$ , update  $l \rightarrow l+2$  and so on. If  $i = p$  one adds up  $(v_2^l)_{l \in I' = \llbracket 2, q-p+1 \rrbracket}$ : for  $l = 2$  to  $l = q-p+1$  start with  $-R_l$  then apply the rotation+symmetry-based operator  $S^1 \circ R_{(q-p+1)-(l+1)}$ , update  $l \rightarrow l+2$  and so on.

At step 6), in both cases,  $q+p$  is even ( $|I'| = q-p$  is even) so one applies successively pairs of operators. Note also that at step 6), one adds contributions in the  $i$ -th pattern of  $\Omega_1$ , already considered at step 4) or 5), and remark that when the index is zero,  $R_0 = I$  so that the rotation+symmetry-based operator simply becomes a symmetry-based operator.

Since  $p+q$  is even<sup>11</sup>, one can check that continuity is ensured at all interfaces of the patterns. The algorithm produces  $p$  operators from  $V_1$  to  $V_2$ , respectively from  $V_2$  to  $V_1$ . Let us give an example: for  $p = 4$ ,  $q = 6$  (that is  $\alpha = 4\pi/5$ ). Then one finds 4 admissible operators from  $V_1$  to  $V_2$  (each one corresponds to a column):

$R_1^{\text{adm}} v_1$	$R_2^{\text{adm}} v_1$	$R_3^{\text{adm}} v_1$	$R_4^{\text{adm}} v_1$	
$S^2 v_1^4$ ,	$S^2 v_1^4$ ,	$S^2 v_1^4$ ,	$S^2 v_1^4$ ,	in $\Omega_2^1$
$S^2 v_1^3$ ,	$S^2 v_1^3$ ,	$S^2 v_1^3$ ,	$R_{-2} v_1^4$ ,	in $\Omega_2^2$
$S^2 v_1^2$ ,	$S^2 v_1^2$ ,	$R_{-4} v_1^3$ ,	$S^1 v_1^4$ ,	in $\Omega_2^3$
$S^2 v_1^1$ ,	$R_{-6} v_1^2$ ,	$S^1 v_1^3$ ,	$S^1 v_1^3$ ,	in $\Omega_2^4$
$R_{-8} v_1^1$ ,	$S^1 v_1^2$ ,	$S^1 v_1^2$ ,	$S^1 v_1^2$ ,	in $\Omega_2^5$
$S^1 v_1^1$ ,	$S^1 v_1^1$ ,	$S^1 v_1^1$ ,	$S^1 v_1^1$ ,	in $\Omega_2^6$

<sup>11</sup>Whereas if  $p+q$  were odd, continuity would not be not ensured at all interfaces: e.g. for 1)-2)-3), at the interface between  $\Omega_2^{p-i}$  and  $\Omega_2^{p-i+1}$ .

In this example,  $|I| = 1$  so one applies only one rotation-based operator (following step 3)). Conversely, one finds 4 admissible operators from  $V_2$  to  $V_1$ :

$$\begin{array}{ccccccc}
\mathbf{R}_1^{\text{adm}'} v_1 & & \mathbf{R}_2^{\text{adm}'} v_1 & & \mathbf{R}_3^{\text{adm}'} v_1 & & \mathbf{R}_4^{\text{adm}'} v_1 \\
\parallel & & \parallel & & \parallel & & \parallel \\
\mathbf{S}^1 v_2^6 - \mathbf{R}_8 v_2^5 + \mathbf{S}^2 v_2^4, & & \mathbf{S}^1 v_2^6, & & \mathbf{S}^1 v_2^6, & & \mathbf{S}^1 v_2^6, & \text{in } \Omega_1^1 \\
\mathbf{S}^2 v_2^3, & & \mathbf{S}^1 v_2^5 - \mathbf{R}_6 v_2^4 + \mathbf{S}^2 v_2^3, & & \mathbf{S}^1 v_2^5, & & \mathbf{S}^1 v_2^5, & \text{in } \Omega_1^2 \\
\mathbf{S}^2 v_2^2, & & \mathbf{S}^2 v_2^2, & & \mathbf{S}^1 v_2^4 - \mathbf{R}_4 v_2^3 + \mathbf{S}^2 v_2^2, & & \mathbf{S}^1 v_2^4, & \text{in } \Omega_1^3 \\
\mathbf{S}^2 v_2^1, & & \mathbf{S}^2 v_2^1, & & \mathbf{S}^2 v_2^1, & & \mathbf{S}^2 v_2^1 - \mathbf{R}_2 v_2^2 + \mathbf{S}^1 v_2^3, & \text{in } \Omega_1^4
\end{array}$$

Here  $|I'| = 2$  so one applies two additional operators to  $v_2^l$ ,  $l \in I'$  (following step 6)).

With these guidelines one can write all operators  $(\mathbf{R}_i^{\text{adm}})_{i=1,p}$ ,  $(\mathbf{R}_i^{\text{adm}'})_{i=1,p}$  for any  $p, q$  such that  $0 < p < q$  and  $p + q$  even. Note that by exchanging  $p$  with  $q$  (and  $\Omega_1$  with  $\Omega_2$ ), one addresses similarly the case  $q < p$ . In the following, we set  $p < q$ . As mentioned in section 3, the problem is that, taken individually, each admissible operator  $\mathbf{R}_i^{\text{adm}}$  (resp.  $\mathbf{R}_i^{\text{adm}'}$ ),  $i = 1, p$ , does not satisfy  $\|\mathbf{R}_i^{\text{adm}}\|^2 = I_\alpha$  (resp.  $\|\mathbf{R}_i^{\text{adm}'}\|^2 = I_\alpha$ ), with  $I_\alpha$  defined in (11). Indeed, for all  $v_1 \in V_1$ ,  $v_2 \in V_2$ , for  $i = 1, p$ ,  $\mathbf{R}_i^{\text{adm}}$  built from 1)-2)-3) and  $\mathbf{R}_i^{\text{adm}'}$  built from 4)-5)-6), one gets the bounds

$$\|\nabla(\mathbf{R}_i^{\text{adm}} v_1)\|_{\Omega_2}^2 \leq \sum_{k=1,p} \|\nabla v_1^k\|_{\Omega_1^k}^2 + (|I| + 1) \|\nabla v_1^i\|_{\Omega_1^i}^2 \leq (q - p + 1) \|\nabla v_1\|_{\Omega_1}^2,$$

$$\|\nabla(\mathbf{R}_i^{\text{adm}'} v_2)\|_{\Omega_1}^2 \leq \sum_{l \in \llbracket 1, q \rrbracket \setminus I'} \|\nabla v_2^l\|_{\Omega_2^l}^2 + (|I'| + 1) \sum_{l \in I'} \|\nabla v_2^l\|_{\Omega_2^l}^2 \leq (q - p + 1) \|\nabla v_2\|_{\Omega_2}^2.$$

The bounds are sharp since:

- in the first case, one may choose  $v_1$  such that  $v_1^k = 0$  for  $k \neq i$ : then  $\|\nabla(\mathbf{R}_i^{\text{adm}} v_1)\|_{\Omega_2}^2 = (q - p + 1) \|\nabla v_1\|_{\Omega_1}^2$ ;
- in the second case, given  $l \in I' (\neq \emptyset)$ , one may choose  $v_2$  such that  $v_2^k = 0$  for  $k \neq l$ : then  $\|\nabla(\mathbf{R}_i^{\text{adm}'} v_2)\|_{\Omega_1}^2 = (q - p + 1) \|\nabla v_2\|_{\Omega_2}^2$ .

One can check that  $I_\alpha < q - p + 1$  for all  $1 < p < q$ :  $I_\alpha - (q - p + 1) = (p - q)(p - 1)/p < 0$  (12).

To get optimal operators (that is of norm equal to  $I_\alpha$ ), one defines  $\mathbf{R}$ ,  $\mathbf{R}'$  as the average of all admissible operators:

$$\mathbf{R} = \frac{1}{p} \sum_{i=1,p} \mathbf{R}_i^{\text{adm}}, \quad \text{and} \quad \mathbf{R}' = \frac{1}{p} \sum_{i=1,p} \mathbf{R}_i^{\text{adm}'}. \quad (35)$$

Then one obtains for instance for  $v_1 \in V_1$

$$\|\nabla(\mathbf{R} v_1)\|_{\Omega_2}^2 = \sum_{l=1,q} \|\nabla(\mathbf{R} v_1)\|_{\Omega_2^l}^2 \leq \sum_{l=1,q} \left( \sum_{k=1,p} M_{lk} \|\nabla v_1^k\|_{\Omega_1^k} \right)^2, \quad \text{with} \quad M_{lk} = \sup_{w_1 \in V_1} \frac{\|\nabla(\mathbf{R} w_1)\|_{\Omega_2^l}}{\|\nabla w_1^k\|_{\Omega_1^k}}. \quad (36)$$

Above, a sum over  $k$  appears due to the fact that for  $l = 1, q$ , the  $(\mathbf{R}_i^{\text{adm}} v_1)|_{\Omega_2^l}$  is a linear combination of isometry-based operators applied to some  $v_1^k$ . To evaluate  $\|\mathbf{R}\|$ ,  $\|\mathbf{R}'\|$ , it is convenient to introduce the matrix  $M \in \mathcal{M}_{q,p}(\mathbb{R})$  of coefficients  $(M_{lk})_{l=1,q,k=1,p}$ . Then from above one has<sup>(3)</sup>

$$\|\mathbf{R}\|^2 = \|M^\top M\|_2.$$

<sup>12</sup>Equality is obtained for  $p = q$  (corresponding to a plane interface) for which  $\|\mathbf{R}_i^{\text{adm}}\|^2 = \|\mathbf{R}_i^{\text{adm}'}\|^2 = 1$ , and in addition  $p = 1$ : in that case there is only one admissible operator.

Let us give the general expression of the matrix  $M$  for any  $p, q$ , and evaluate  $\|M^\top M\|_2$  to conclude. For  $n \in \llbracket 1, p \rrbracket$ , define  $M_{p-n+1}, \widetilde{M}_{p-n+1} \in \mathcal{M}_{p-n+1, p}(\mathbb{R})$  such that

$$\begin{array}{c} M_{p-n+1} \\ \parallel \\ \left( \begin{array}{cccccccc} 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \frac{p}{p} \\ \vdots & & & & & & & \ddots & \frac{p-1}{p} & \frac{1}{p} \\ \vdots & & \{0\} & & & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \frac{j}{p} & \ddots & \left\{ \frac{1}{p} \right\} & \vdots & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \frac{n}{p} & \frac{1}{p} & \dots & \dots & \dots & \dots & \frac{1}{p} \end{array} \right), \end{array} \quad \begin{array}{c} \widetilde{M}_{p-n+1} \\ \parallel \\ \left( \begin{array}{cccccccc} \frac{1}{p} & \dots & \dots & \dots & \dots & \frac{1}{p} & \frac{n}{p} & 0 & \dots & 0 \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \left\{ \frac{1}{p} \right\} & \ddots & \frac{p-j+1}{p} & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \{0\} & \vdots & \vdots \\ \frac{1}{p} & \frac{p-1}{p} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \end{array} \right), \end{array}$$

where  $j$  denotes the column index. All entries are equal to 0 or  $\frac{1}{p}$ , except on one diagonal where they range from  $\frac{n}{p}$  to  $\frac{p}{p} = 1$ . Then the shape of  $M$  depends on whether  $p \leq \frac{q}{2}$  or not:

- if  $p \leq q/2$ : let  $m \in \mathbb{N}$  such that  $q = 2p + m$ . Then the matrix  $M$  is written  $M = \begin{pmatrix} M_p \\ M'_m \\ \widetilde{M}_p \end{pmatrix}$  with  $M_p, \widetilde{M}_p \in \mathcal{M}_p(\mathbb{R})$  (defined above for  $n = 1$ ) and  $M'_m \in \mathcal{M}_{m,p}(\mathbb{R})$  whose entries are all equal to  $1/p$ , with the convention that if  $m = 0$ ,  $M'_m$  is empty.

- if  $p > q/2$ : let  $m \in \mathbb{N}^*$  such that  $q = 2p - m$ . If  $m > 2$  then the matrix  $M$  is written  $M = \begin{pmatrix} M_{p-m+1} \\ M''_{m-2} \\ \widetilde{M}_{p-m+1} \end{pmatrix}$ , with  $M_{p-m+1}, \widetilde{M}_{p-m+1} \in \mathcal{M}_{p-m+1, p}(\mathbb{R})$  (defined above for  $n = m$ ) and  $M''_{m-2} \in \mathcal{M}_{m-2, p}(\mathbb{R})$  such that (below  $i$  denotes the row index)

$$M''_{m-2} = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 & \frac{m-1}{p} & \frac{1}{p} & \dots & \frac{1}{p} & \frac{2}{p} & 0 \\ \vdots & & & & \ddots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \{0\} & \ddots & \frac{m-i}{p} & \ddots & \left\{ \frac{1}{p} \right\} & \ddots & \frac{1+i}{p} & \ddots & \{0\} & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \frac{2}{p} & \frac{1}{p} & \dots & \frac{1}{p} & \frac{m-1}{p} & 0 & \dots & \dots & \dots & 0 \end{pmatrix}.$$

$\underbrace{\hspace{10em}}_{p-m-1}$

One simply finds the matrix  $M = \begin{pmatrix} M_{p-1} \\ \frac{1}{p} \dots \frac{1}{p} \\ \widetilde{M}_{p-1} \end{pmatrix}$  when  $q = 2p - 1$ , and the matrix  $M = \begin{pmatrix} M_{p-1} \\ \widetilde{M}_{p-1} \end{pmatrix}$  when  $q = 2p - 2$ .

Note that the cases  $p > q$  are obtained by taking the transposed matrix presented above. Moreover, one checks easily by direct inspection the next result.

**Proposition 1.** *The matrix  $M \in \mathcal{M}_{q,p}(\mathbb{R})$  satisfies the following properties:*

$$[1] \begin{cases} \forall l = 1, q, \forall k = 1, p, M_{lk} \geq 0 \\ \exists l_0, k_0 \text{ s.t. } M_{l_0 k_0} = 1 \end{cases}; [2] \forall l, \begin{cases} \sum_{k=1}^p M_{lk} = 1 & \text{if } p \leq q \\ \sum_{k=1}^p M_{lk} \leq \frac{p}{q} & \text{if } p \geq q \end{cases}; [3] \forall k, \begin{cases} \sum_{l=1}^q M_{lk} \leq \frac{q}{p} & \text{if } p \leq q \\ \sum_{l=1}^q M_{lk} = 1 & \text{if } p \geq q \end{cases}.$$

It follows that



**Proposition 2.** For all  $p, q > 0$  let  $M \in \mathcal{M}_{q,p}(\mathbb{R})$  be a matrix which satisfies the properties of proposition 1. Then

$$\|M^\top M\|_2 \leq \max\left(\frac{q}{p}, \frac{p}{q}\right).$$

*Proof.* Define  $A := M^\top M$ , it holds  $\|A\|_2 = \lambda_{\max} = \max_{\lambda \in \sigma(A)} \lambda(A) > 0$ , where  $\sigma(A)$  is the set of eigenvalues of  $A$ . Using Gershgorin circle theorem, one bounds the spectrum of  $A$  as follows:

$$\forall \lambda \in \sigma(A), \forall k = 1, p, |\lambda - A_{kk}| \leq \sum_{l \neq k} A_{kl},$$

then

$$\begin{aligned} \lambda_{\max} &\leq \max_{k=1,p} \left( A_{kk} + \sum_{l \neq k} A_{kl} \right) = \max_{k=1,p} \left( \sum_{l=1,p} A_{kl} \right) \leq \max_{k=1,p} \left( \sum_{l=1}^p \sum_{h=1}^q (M^\top)_{kh} M_{hl} \right) = \max_{k=1,p} \left( \sum_{l=1}^p \sum_{h=1}^q M_{hk} M_{hl} \right) \\ &\leq \max_{k=1,p} \left( \sum_{h=1}^q \left( \sum_{l=1}^p M_{hl} \right) M_{hk} \right). \end{aligned}$$

Due to the second and the third properties in proposition 1 satisfied by  $M$ , one finds

$$\lambda_{\max} \leq \max_{k=1,p} \left( \max\left(1, \frac{p}{q}\right) \sum_{h=1}^q M_{hk} \right) \leq \max\left(1, \frac{p}{q}\right) \max\left(1, \frac{q}{p}\right) = \max\left(\frac{p}{q}, \frac{q}{p}\right).$$

□

Consequently, the proposed operator  $\mathbf{R}$  in (35) is of optimal norm. One proceeds similarly for  $\mathbf{R}'$  by considering  $M^\top$  instead of  $M$  as the roles of  $p$  and  $q$  are exchanged:

$$\|\mathbf{R}\|^2 = \|M^\top M\|_2 \leq \max\left(\frac{p}{q}, \frac{q}{p}\right), \quad \text{and} \quad \|\mathbf{R}'\|^2 = \|M M^\top\|_2 \leq \max\left(\frac{p}{q}, \frac{q}{p}\right).$$

**Remark 4.** There is always a unit entry in  $M$  according to proposition 1 (denoted by  $M_{k_0 l_0}$ ), one readily checks that if  $w_1^0 \in V_1$  with  $\text{supp}(w_1^0) \subset \Omega_1^{k_0}$ , then it follows  $\|\nabla(\mathbf{R}w_1^0)\|_{\Omega_2}^2 \geq \|\nabla w_1^0\|_{\Omega_1}^2$ . Hence  $\|\mathbf{R}\|^2 \geq 1$ . Similarly,  $\|\mathbf{R}'\|^2 \geq 1$ .

## A.2 Weighted estimates for operators $\mathbf{R}$

In §A.1 we provided bounds for the norms of the geometry-based operators  $\mathbf{R}$ . Here we provide bounds for the operator norm when we use a localization process (see section 4), that is when the operator  $\mathbf{R}$  is locally applied in the neighborhood of the interface thanks to a cut-off function  $\xi$  (defined as in section 4) whose support is localized either near a corner (proposition 3) or a straight line (proposition 4) of the interface  $\Sigma$ . We use the same notations as in §A.1.

**Proposition 3.** Let  $\xi$  be a smooth positive function with support  $S$ , that depends only on the distance to the corner of angle  $\alpha = 2\pi \frac{p}{p+q}$ . Then

$$\forall w_1 \in V_1, \quad \|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 \leq I_\alpha \|\xi^{1/2} \nabla w_1\|_{\Omega_1 \cap S}^2,$$

where  $I_\alpha = \max\left(\frac{p}{q}, \frac{q}{p}\right)$ , and  $\mathbf{R}$  is a rotation- and symmetry-based operator from  $V_1$  to  $V_2$  defined as in (23).

*Proof.* Let  $w_1 \in V_1$ .

$$\begin{aligned} \|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 &= \int_{\Omega_2 \cap S} \xi(\rho) \left( \left| \frac{\partial(\mathbf{R}w_1)}{\partial \rho} \right|^2 + \frac{1}{\rho^2} \left| \frac{\partial(\mathbf{R}w_1)}{\partial \theta} \right|^2 \right) \rho d\rho d\theta \\ &= \sum_{l=1}^q \int_{\Omega_2^l \cap S} \xi(\rho) \left( \left| \frac{\partial(\mathbf{R}w_1)}{\partial \rho} \right|^2 + \frac{1}{\rho^2} \left| \frac{\partial(\mathbf{R}w_1)}{\partial \theta} \right|^2 \right) \rho d\rho d\theta. \end{aligned}$$

Then similar to (36) with the change of variables  $(\rho, \theta) \mapsto (r_k, \varphi_k)$  for  $k = 1, p$  that maps  $\Omega_2^l$  to  $\Omega_1^k$  (note that  $\xi(r_k) = \xi(\rho)$ ) one finds

$$\begin{aligned} \|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 &\leq \sum_{l=1}^q \left( \sum_{k=1}^p M_{lk} \left( \int_{\Omega_1^k \cap S} \xi(r_k) \left( \left| \frac{\partial w_1^k(r_k, \varphi_k)}{\partial r_k} \right|^2 + \frac{1}{r_k^2} \left| \frac{\partial w_1^k(r_k, \varphi_k)}{\partial \varphi_k} \right|^2 \right) r_k dr_k d\varphi_k \right)^{1/2} \right)^2 \\ &\leq \sum_{l=1}^q \left( \sum_{k=1}^p M_{lk} \|\xi^{1/2} \nabla w_1^k\|_{\Omega_1^k \cap S} \right)^2 \end{aligned}$$

Introducing  $\vec{W}_\xi = (\|\xi^{1/2} \nabla w_1^1\|_{\Omega_1^1 \cap S}, \dots, \|\xi^{1/2} \nabla w_1^p\|_{\Omega_1^p \cap S})^\top$ , then one has  $M \vec{W}_\xi = \sum_{k=1}^p M_{lk} \|\xi^{1/2} \nabla w_1^k\|_{\Omega_1^k \cap S}$ , using proposition 2 one finally gets

$$\|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 \leq (M^\top M \vec{W}_\xi, \vec{W}_\xi) \leq I_\alpha \|\xi^{1/2} \nabla w_1\|_{\Omega_1 \cap S}^2.$$

□

**Remark 5.** Following remark 4, one can find  $w_1^0 \in V_1$  such that  $\|\xi^{1/2} \nabla(\mathbf{R}w_1^0)\|_{\Omega_2 \cap S}^2 \geq \|\xi^{1/2} \nabla w_1^0\|_{\Omega_1 \cap S}^2$ .

**Proposition 4.** Let  $\xi$  be a smooth positive function with support  $S$  that is symmetric with respect to the interface. Then for all  $w_1 \in V_1$

$$\|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 = \|\xi^{1/2} \nabla w_1\|_{\Omega_1 \cap S}^2,$$

where  $\mathbf{R}$  is the symmetry-based operator (24).

*Proof.* This inequality is obtained using the change of variables  $(x^\Sigma, y^\Sigma) \rightarrow (x^\Sigma, -y^\Sigma)$  in  $\Omega_2 \cap S$ . For all  $w_1 \in V_1$

$$\begin{aligned} \|\xi^{1/2} \nabla(\mathbf{R}w_1)\|_{\Omega_2 \cap S}^2 &= \int_{\Omega_2 \cap S} \xi(x^\Sigma, y^\Sigma) \left( \left| \frac{\partial(\mathbf{R}w_1)}{\partial x^\Sigma} \right|^2 + \left| \frac{\partial(\mathbf{R}w_1)}{\partial y^\Sigma} \right|^2 \right) dx^\Sigma dy^\Sigma \\ &= \int_{\Omega_1 \cap S} \xi(x^\Sigma, y^\Sigma) \left( \left| \frac{\partial w_1}{\partial x^\Sigma} \right|^2 + \left| \frac{\partial w_1}{\partial y^\Sigma} \right|^2 \right) dx^\Sigma dy^\Sigma = \|\xi^{1/2} \nabla w_1\|_{\Omega_1 \cap S}^2. \end{aligned}$$

□

### A.3 Local and global interpolation estimates

Let  $\hat{\tau}$  be the reference triangle, with corners  $(1, 0)$ ,  $(0, 1)$  and  $(0, 0)$ , and let  $\hat{I}_d$  be the Lagrange interpolation operator over  $\mathbb{P}_d(\hat{\tau})$ . Given  $(\mathcal{T}_h)_h$  a regular family of triangulations of a domain  $\Omega$  we call  $I_\tau^d$  the Lagrange interpolation operator over  $\mathbb{P}_d(\tau)$ , for all  $h$  and all  $\tau \in \mathcal{T}_h$ .

**Lemma 4.** On the reference triangle  $\hat{\tau}$ , one has the estimate

$$\exists \hat{C} > 0, \forall \hat{\chi} \in \mathbb{P}_1(\hat{\tau}), \forall \hat{v} \in \mathbb{P}_d(\hat{\tau}), \|\hat{\nabla}(\hat{\chi}\hat{v} - \hat{I}_d(\hat{\chi}\hat{v}))\|_{\hat{\tau}} \leq \hat{C} |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} \|\hat{\nabla}\hat{v}\|_{\hat{\tau}}. \quad (37)$$

As a consequence, one has the uniform local estimate

$$\exists C > 0, \forall h, \forall \tau \in \mathcal{T}_h, \forall \chi \in \mathbb{P}_1(\tau), \forall v \in \mathbb{P}_d(\tau), \|\nabla(\chi v - I_\tau^d(\chi v))\|_{\tau} \leq C h_\tau |\chi|_{W^{1,\infty}(\tau)} \|\nabla v\|_{\tau}. \quad (38)$$

*Proof.* Write  $\hat{\chi}(\hat{x}, \hat{y}) = \alpha + \hat{\chi}_1(\hat{x}, \hat{y})$ , resp.  $\hat{v}(\hat{x}, \hat{y}) = \hat{v}_-(\hat{x}, \hat{y}) + \hat{v}_d(\hat{x}, \hat{y})$  where  $\hat{\chi}_1(\hat{x}, \hat{y}) = \beta\hat{x} + \gamma\hat{y}$ , resp.  $\hat{v}_d(\hat{x}, \hat{y}) = \sum_{i=0,d} a_i \hat{x}^i \hat{y}^{d-i}$  and  $\deg(\hat{v}_-) < d$ . For  $i = 0, d$ , define  $\hat{e}_i(\hat{x}, \hat{y}) = \hat{x}^{i+1} \hat{y}^{d-i} - \hat{I}_d(\hat{x}^{i+1} \hat{y}^{d-i})$ , resp.

$\hat{f}_i(\hat{x}, \hat{y}) = \hat{x}^i \hat{y}^{d+1-i} - \hat{I}_d(\hat{x}^i \hat{y}^{d+1-i})$ . Note that  $|\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} = \max(|\beta|, |\gamma|)$ .

Since there holds  $\hat{\chi}\hat{v} - \hat{I}_d(\hat{\chi}\hat{v}) = \hat{\chi}_1\hat{v}_d - \hat{I}_d(\hat{\chi}_1\hat{v}_d) = \beta \sum_{i=0,d} a_i \hat{e}_i(\hat{x}, \hat{y}) + \gamma \sum_{i=0,d} a_i \hat{f}_i(\hat{x}, \hat{y})$ , one finds

$$\begin{aligned} \left| \hat{\nabla} \left( \hat{\chi}\hat{v} - \hat{I}_d(\hat{\chi}\hat{v}) \right) (\hat{x}, \hat{y}) \right| &\leq |\beta| \sum_{i=0,d} |a_i| |\hat{\nabla} \hat{e}_i(\hat{x}, \hat{y})| + |\gamma| \sum_{i=0,d} |a_i| |\hat{\nabla} \hat{f}_i(\hat{x}, \hat{y})| \\ &\leq |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} \sum_{i=0,d} |a_i| \left( |\hat{\nabla} \hat{e}_i(\hat{x}, \hat{y})| + |\hat{\nabla} \hat{f}_i(\hat{x}, \hat{y})| \right) \\ &\leq |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} \left( \sum_{i=0,d} |a_i|^2 \right)^{1/2} \left( \sum_{i=0,d} \left( |\hat{\nabla} \hat{e}_i(\hat{x}, \hat{y})| + |\hat{\nabla} \hat{f}_i(\hat{x}, \hat{y})| \right)^2 \right)^{1/2} \\ &\leq |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} \left( \sum_{i=0,d} |a_i|^2 \right)^{1/2} \left( 2 \sum_{i=0,d} \left( |\hat{\nabla} \hat{e}_i(\hat{x}, \hat{y})|^2 + |\hat{\nabla} \hat{f}_i(\hat{x}, \hat{y})|^2 \right) \right)^{1/2}. \end{aligned}$$

It follows that  $\|\hat{\nabla}(\hat{\chi}\hat{v} - \hat{I}_d(\hat{\chi}\hat{v}))\|_{\hat{\tau}} \leq \hat{C}_d |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} (\sum_{i=0,d} |a_i|^2)^{1/2}$ , with  $\hat{C}_d = \sqrt{2} [\sum_{i=0,d} (\|\hat{\nabla} \hat{e}_i\|_{\hat{\tau}}^2 + \|\hat{\nabla} \hat{f}_i\|_{\hat{\tau}}^2)]^{1/2}$ . Then, we remark that the  $\ell^2$ -norm of the coefficients is a norm over  $\mathbb{P}_d(\hat{\tau})$ , hence it is also a norm over its vector subspace  $\mathbb{P}_d^{zmv}(\hat{\tau})$  made of zero-mean value polynomials on  $\hat{\tau}$ . Now, over  $\mathbb{P}_d^{zmv}(\hat{\tau})$ , the semi-norm  $\|\hat{\nabla} \cdot\|_{\hat{\tau}}$  is also a norm and because  $\mathbb{P}_d^{zmv}(\hat{\tau})$  is a finite dimensional vector space, both norms are equivalent. Noting finally that, starting from  $\hat{v}$ , one has only to modify the degree-0 coefficient to obtain a zero-mean value polynomial, we finally get that there exists  $\hat{C}'$  independent of  $\hat{v}$  such that  $(\sum_{i=0,d} |a_i|^2)^{1/2} \leq \hat{C}' \|\hat{\nabla} \hat{v}\|_{\hat{\tau}}$ . Taking  $\hat{C} = \hat{C}_d \hat{C}'$  leads to (37).

We recall that provided the family of triangulations  $(\mathcal{T}_h)_h$  is regular, there exists  $\mathbf{s} > 0$  such that, for all  $h$  and for all  $\tau \in \mathcal{T}_h$ , there holds  $h_\tau \leq \mathbf{s} \rho_\tau$ , where  $\rho_\tau$  is the diameter of the largest ball that can be inscribed in  $\tau$ . One can then derive (38) from (37) by using the affine mapping from the reference triangle  $\hat{\tau}$  to the triangle  $\tau \in \mathcal{T}_h$ . We refer for instance to [15]. We report here the computations for the sake of completeness ( $C_0, C_1, \dots$  are constants that are independent of  $h$  and  $\tau$ ):

$$\begin{aligned} \|\nabla(\chi v - I_\tau^d(\chi v))\|_\tau &\leq C_0 \mathbf{s} \|\hat{\nabla}(\hat{\chi}\hat{v} - \hat{I}_d(\hat{\chi}\hat{v}))\|_{\hat{\tau}} && \text{cf. Lemmas 1.100-1.101 of [15]} \\ &\leq C_1 |\hat{\chi}|_{W^{1,\infty}(\hat{\tau})} \|\hat{\nabla} \hat{v}\|_{\hat{\tau}} && \text{cf. (37)} \\ &\leq C_2 h_\tau |\chi|_{W^{1,\infty}(\tau)} \times C_3 \mathbf{s} \|\nabla v\|_\tau && \text{cf. Lemmas 1.100-1.101 of [15]}, \end{aligned}$$

that is (38) with  $C = C_2 C_3 \mathbf{s}$ . □

Let us prove the estimate over  $\Omega^h$ . Recall that  $V := H_0^1(\Omega)$ ,  $V_{(d)}^h := \{v \in V : v|_\tau \in \mathbb{P}_d(\tau), \forall \tau \in \mathcal{T}_h\}$ , and  $I_h$  is the interpolation operator on  $V_{(d)}^h$ . Consider a cutoff function  $\chi$  whose support is included in  $\Omega$ , and denote by  $\chi_h$  its interpolation of degree 1.

**Lemma 5.** *For all  $v^h \in V_{(d)}^h$ ,*

$$\exists C > 0, \quad \|\nabla(\chi_h v^h - I_h(\chi_h v^h))\|_{\Omega^h} \leq C h \|\chi\|_{W^{1,\infty}(\Omega)} \|\nabla v^h\|_\Omega. \quad (39)$$

*Proof.* To obtain (39), we evaluate the  $L^2$ -norm of  $\nabla(\chi_h v^h - I_h(\chi_h v^h))$  on  $\Omega^h$  by splitting  $\Omega^h$  into triangles, and then going back to the reference triangle to use the uniform estimate (38):

$$\|\nabla(\chi_h v^h - I_h(\chi_h v^h))\|_{\Omega^h}^2 = \sum_{\tau \subset \Omega^h} \int_\tau |\nabla(\chi_h|_\tau v^h|_\tau - I_\tau^d(\chi_h|_\tau v^h|_\tau))|^2 \leq C \sum_{\tau \subset \Omega^h} h_\tau^2 |\chi_h|_\tau|_{W^{1,\infty}(\tau)}^2 \|\nabla v^h|_\tau\|_\tau^2.$$

Using the definition of the meshsize  $h$  yields

$$\|\nabla(\chi_h v^h - I_h(\chi_h v^h))\|_{\Omega^h} \leq C h \left( \sup_{\tau \subset \Omega^h} |\chi_h|_\tau|_{W^{1,\infty}(\tau)} \right) \left( \sum_{\tau \subset \Omega^h} \|\nabla v^h|_\tau\|_\tau^2 \right)^{1/2} \leq C h \|\chi_h\|_{W^{1,\infty}(\Omega)} \|\nabla v^h\|_\Omega.$$

One concludes using the stability estimate (27). □