A new ϕ-FEM approach for problems with natural boundary conditions

Michel Duprez, Vanessa Lleras, Alexei Lozinski

A new ϕ-FEM approach for problems with natural boundary conditions

Michel Duprez[*†] and Vanessa Lleras[‡] and Alexei Lozinski[§]

January 7, 2022

Abstract

We present a new finite element method, called ϕ-FEM, to solve numerically elliptic partial differential equations with natural (Neumann or Robin) boundary conditions using simple computational grids, not fitted to the boundary of the physical domain. The boundary data are taken into account using a level-set function, which is a popular tool to deal with complicated or evolving domains. Our approach belongs to the family of fictitious domain methods (or immersed boundary methods) and is close to recent methods of cutFEM/XFEM type. Contrary to the latter, ϕ-FEM does not need any non-standard numerical integration on cut mesh elements or on the actual boundary, while assuring the optimal convergence orders with finite elements of any degree and providing reasonably well conditioned discrete problems. In the first version of ϕ-FEM, only essential (Dirichlet) boundary conditions was considered. Here, to deal with natural boundary conditions, we introduce the gradient of the primary solution as an auxiliary variable. This is done only on the mesh cells cut by the boundary, so that the size of the numerical system is only slightly increased . We prove theoretically the optimal convergence of our scheme and a bound on the discrete problem conditioning, independent of the mesh cuts. The numerical experiments confirm these results.

1 Introduction

We consider a second order elliptic partial differential equation with Neumann boundary conditions

$$-\Delta u + u = f \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \Gamma \tag{1}$$

in a bounded domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ with smooth boundary Γ assuming that Ω and Γ are given by a level-set function ϕ:

$$\Omega := \{\phi < 0\} \text{ and } \Gamma := \{\phi = 0\}. \tag{2}$$

Such a representation is a popular and useful tool to deal with problems with evolving surfaces or interfaces [20]. In the present article, the level-set function is supposed known on $\mathbb{R}^d$, smooth, and to behave near Γ similar to the signed distance to Γ.

Our goal is to develop a finite element method for (1) using a mesh which is not fitted to Γ, i.e. we allow the boundary Γ to cut the mesh cells in an arbitrary manner. The existing finite elements methods on non-matching meshes, such as the fictitious domain/penalty method [11], XFEM [19, 18, 21, 12], CutFEM [8, 7], CutIGA [10] (see also [17] for a review on immersed boundary methods) contain the integrals over the physical domain Ω and thus necessitate non-standard numerical integration on the parts of mesh cells cut

[*]Corresponding author : Michel Duprez, 1 place de l'hôpital, 67000 Strasbourg, France. michel.duprez@inria.fr
[†]Mimesis team, IGG team, Inria, Université de Strasbourg,France
[‡]IMAG, Université de Montpellier, CNRS, France
[§]Laboratoire de mathématiques de Besançon, Université de Bourgogne Franche-Comté, CNRS, France

by Γ. In this article, we propose a finite element method, based on an alternative variational formulation on an extended domain matching the computational mesh, thus avoiding any non-standard quadrature while maintaining the optimal accuracy and controlling the conditioning uniformly with respect to the position of Ω over the mesh.[1]

In the recent article [9], we have proposed such a method for the Poisson problem with homogeneous Dirichlet boundary conditions $u = 0$ on Γ. The idea behind this method, baptised ϕ-FEM, is to put $u = \phi w$ so that $u = 0$ on Γ for whatever w since $\phi = 0$ there. We then replace ϕ and w by the finite element approximations ϕ_h and w_h, substitute $u \approx \phi_h w_h$ into an appropriate variational formulation and get an easily implementable discretization in terms of the new unknown w_h. Such a simple idea cannot be used directly to discretize the Neumann boundary conditions in (1). Indeed, multiplication by ϕ works well to strongly impose the **essential** Dirichlet boundary conditions whereas Neumann conditions are **natural**, i.e. they come out of the usual variational formulation without imposing them into the functional spaces. We want thus to reformulate Problem (1) so that Neumann conditions become essential. The way to go is the dualization of this problem, in the terminology of [4], consisting in introducing an auxiliary (vector-valued) variable for the gradient ∇u. In the present article, we want to use the usual conforming scalar finite elements as much as possible. Accordingly, we do not pursue the classical route of mixed methods, as in Chapter 7 of [4]. We shall rather introduce the additional unknowns only where they are needed, i.e. in the vicinity of boundary Γ.

More specifically, let us assume that Ω lies inside a simply shaped domain $\mathcal{O}$ (typically a box in $\mathbb{R}^d$) and introduce a quasi-uniform simplicial mesh $\mathcal{T}_h^{\mathcal{O}}$ on $\mathcal{O}$ (the background mesh). Let $\mathcal{T}_h$ be a submesh of $\mathcal{T}_h^{\mathcal{O}}$ obtained by getting rid of mesh elements lying entirely outside Ω (the definition of $\mathcal{T}_h$ will be slightly changed afterwords). Denote by Ω_h the domain covered by mesh $\mathcal{T}_h$ (Ω_h only slightly larger than Ω) and by Ω_h^Γ the domain covered by mesh elements of $\mathcal{T}_h$ cut by Γ (a narrow strip of width $\sim h$ around Γ). Assume that the right-hand side f is actually well defined on Ω_h and imagine for the moment that the solution u of eq. (1) can be extended to a function on Ω_h, still denoted by u, which solves the same equation, now on Ω_h:

$$-\Delta u + u = f, \quad \text{in } \Omega_h\,. \tag{3}$$

As announced above, we now introduce an auxiliary vector-valued unknown y on Ω_h^Γ, setting $y = -\nabla u$ there, so that u, y satisfy the dual form of the original equation

$$y + \nabla u = 0\,, \quad \operatorname{div} y + u = f, \quad \text{in } \Omega_h^\Gamma\,. \tag{4}$$

This allows us to rewrite the natural boundary condition $\frac{\partial u}{\partial n} = 0$ on Γ as the essential condition on y: $y \cdot n = 0$ on Γ. The latter can now be imposed using the idea of multiplication by the level-set ϕ. To this end, we note that the outward-looking unit normal n is given on Γ by $n = \frac{1}{|\nabla\phi|}\nabla\phi$. Hence, we have $y \cdot n = 0$ on Γ if we put

$$y \cdot \nabla\phi + p\phi = 0, \quad \text{in } \Omega_h^\Gamma, \tag{5}$$

where p is yet another (scalar-valued) auxiliary unknown on Ω_h^Γ. The variable p is artificial and has no physical meaning (roughly speaking, it represents the second derivative in the normal direction on Γ, extended somehow to Ω_h^Γ). The only purpose of its introduction is to avoid the terms with integral on Γ in our variational formulation.

Our finite element method, cf. (6) below, will be based on a variational formulation of system (3)–(5) treating eqs. (4)–(5) in a least squares manner and adding a stabilization in the vein of the Ghost penalty [6, 10]. As in [9], we coin our method ϕ-FEM in accordance with the tradition of denoting the level-sets by ϕ. Contrary to [9], we need here additional finite element unknowns discretizing y and p on Ω_h^Γ. Since, the latter represents only a small portion of the whole computational domain Ω_h, the extra cost induced by these unknowns is negligible as $h \to 0$. We want to emphasize that the reformulation (3)–(5) is very

[1] Another approach that can bypass non-standard quadrature is the Fat Boundary Method [16, 3] where a non-matching global mesh is combined with a local matching mesh around the boundary. Our method, on the other hand, avoids completely the construction of boundary fitted meshes.

formal and will serve only as a motivation for our discrete scheme (6). The system (3)–(5) itself is clearly over-determined and may well be ill-posed (the "boundary" conditions hidden in (5) are actually not on the boundary of domain Ω_h where the problem is now posed). We shall assume neither the existence of a continuous solution to (3)–(5), nor any properties of such a solution in the theoretical analysis of our scheme, cf. Theorem 2.1.

The article is organized as follows: our ϕ-FEM method is presented in the next section. We also give there the assumptions on the level-set ϕ and on the mesh, and announce our main result: the *a priori* error estimate for ϕ-FEM in the Neumann case. We work with standard continuous $\mathbb{P}_k$ finite elements ($k \geq 1$) on a simplicial mesh and prove the optimal order h^k for the error in the H^1 norm and the (slightly) suboptimal order $h^{k+1/2}$ for the error in the L^2 norm. We note in passing that employing finite elements of any order is quite straightforward in our approach contrary to more traditional schemes of CutFEM type, cf. [5, 14] for a special treatment of the case $k > 1$. The proofs of the error estimates are the subject of Section 3. Moreover, we show in Section 4 that the associated finite element matrix has the condition number of order $1/h^2$, i.e. of the same order as that of a standard finite element method on a matching grid of comparable size. In particular, the conditioning of our method does not suffer from arbitrarily bad intersections of Γ with the mesh. Numerical illustrations are given in Section 5.

2 Definitions, assumptions, description of ϕ-FEM, and the main result

Assume $\Omega \subset \mathcal{O}$ and let $\mathcal{T}_h^{\mathcal{O}}$ be a quasi-uniform simplicial mesh on $\mathcal{O}$ with $h = \max_{T\in\mathcal{T}_h} \operatorname{diam} T$ and $\rho(T) \geq \beta h$ for all $T \in \mathcal{T}_h^{\mathcal{O}}$ with the mesh regularity parameter $\beta > 0$ fixed once for all (here $\rho(T)$ is the radius of the largest ball inscribed in T). Fix integers $k, l \geq 1$ and let ϕ_h be the FE interpolation of ϕ on $\mathcal{T}_h^{\mathcal{O}}$ by the usual continuous finite elements of degree l.[2] Let $\Gamma_h := \{\phi_h = 0\}$ and introduce the computational mesh $\mathcal{T}_h$ (approximately) covering Ω and the auxiliary mesh $\mathcal{T}_h^\Gamma$ covering Γ_h:

$$\begin{aligned} \mathcal{T}_h &= \{T \in \mathcal{T}_h^{\mathcal{O}} : T \cap \{\phi_h < 0\} \neq \varnothing\} && \text{and } \Omega_h = (\cup_{T\in\mathcal{T}_h} T)^\circ, \\ \mathcal{T}_h^\Gamma &= \{T \in \mathcal{T}_h : T \cap \Gamma_h \neq \varnothing\} && \text{and } \Omega_h^\Gamma = (\cup_{T\in\mathcal{T}_h^\Gamma} T)^\circ. \end{aligned}$$

We shall also denote by $\Omega_h^i = \Omega_h \setminus \Omega_h^\Gamma$ the domain of mesh elements completely inside Ω and set $\Gamma_h^i = \partial\Omega_h^i$.

We now introduce the finite element spaces

$$\begin{aligned} V_h^{(k)} &= \{v_h \in H^1(\Omega_h) : v_h|_T \in \mathbb{P}_k(T) \quad \forall T \in \mathcal{T}_h\}, \\ Z_h^{(k)} &= \{z_h \in H^1(\Omega_h^\Gamma)^d : z_h|_T \in \mathbb{P}_k(T)^d \quad \forall T \in \mathcal{T}_h^\Gamma\}, \\ Q_h^{(k)} &= \{q_h \in L^2(\Omega_h^\Gamma) : q_h|_T \in \mathbb{P}_{k-1}(T) \quad \forall T \in \mathcal{T}_h^\Gamma\}, \\ W_h^{(k)} &= V_h^{(k)} \times Z_h^{(k)} \times Q_h^{(k)} \end{aligned}$$

and the finite element problem: Find $(u_h, y_h, p_h) \in W_h^{(k)}$ such that

$$a_h(u_h, y_h, p_h; v_h, z_h, q_h) = \int_{\Omega_h} f v_h + \gamma_{div} \int_{\Omega_h^\Gamma} f(\operatorname{div} z_h + v_h), \tag{6}$$

[2] The integer k is the degree of finite elements which will be used to approximate the principal unknown u while ϕ is approximated by finite elements of degree l. We shall require $l \geq k+1$ in our convergence Theorem 2.1. Note, that we cannot set $l = k$ unlike the Dirichlet case in [9]. This is essentially due to the fact that ϕ_h is used here to approximate the normal on Γ in addition to approximating Γ itself.

for all $(v_h, z_h, q_h) \in W_h^{(k)}$, where

$$a_h(u,y,p;v,z,q) = \int_{\Omega_h} \nabla u \cdot \nabla v + \int_{\Omega_h} uv + \int_{\partial\Omega_h} y \cdot nv + \gamma_{div} \int_{\Omega_h^\Gamma} (\operatorname{div} y + u)(\operatorname{div} z + v) + \gamma_u \int_{\Omega_h^\Gamma} (y + \nabla u) \cdot (z + \nabla v) + \frac{\gamma_p}{h^2} \int_{\Omega_h^\Gamma} (y \cdot \nabla\phi_h + \frac{1}{h} p\phi_h)(z \cdot \nabla\phi_h + \frac{1}{h} q\phi_h) + \sigma h \int_{\Gamma_h^i} \left[\frac{\partial u}{\partial n}\right] \left[\frac{\partial v}{\partial n}\right]$$

with some positive numbers γ_{div}, γ_u, γ_p, and σ properly chosen in a manner independent of h. We have assumed here that f is well defined on Ω_h, rather than on Ω only.

The finite element problem (6) is inspired by (3)–(5). The first line in the definition of a_h comes from multiplying (3) by a test function v, integrating by parts

$$\int_{\Omega_h} \nabla u \cdot \nabla v + \int_{\Omega_h} uv - \int_{\partial\Omega_h} \nabla u \cdot nv = \int_{\Omega_h} fv$$

and noting that $-\nabla u \cdot n = y \cdot n$ on $\partial\Omega_h$ by (4). Equations (4)–(5) are then added in least squares manner, introducing the test functions z and q corresponding to y and p respectively (a similar idea is also used in the construction of CutIGA method [10]). Note that we replace p by $\frac{1}{h}p$ in the term stemming from (5). This rescaling does not affect the discretization of u (which is the only quantity that interests us) and will be crucial to control the conditioning of the method. Finally, the terms multiplied by σh is the Ghost penalty from [6] (we need to penalize the jumps only on Γ_h^i because some continuity of ∇u_h on the facets inside Ω_h^Γ is already enforced by assimilating ∇u_h to y_h which is continuous).

We now recall some technical assumptions on the domain and the mesh, the same as in [15, 9]. These assumptions hold true for smooth domains and sufficiently refined meshes.

Assumption 1. *There exists a neighborhood of Γ, a domain Ω^Γ, which can be covered by open sets $\mathcal{O}_i$, $i = 1, \dots, I$ and one can introduce on every $\mathcal{O}_i$ local coordinates $\xi_1, \dots, \xi_d$ with $\xi_d = \phi$ such that all the partial derivatives $\partial^\alpha \xi / \partial x^\alpha$ and $\partial^\alpha x / \partial \xi^\alpha$ up to order $k+1$ are bounded by some $C_0 > 0$. Thus, ϕ is of class C^{k+2} on Ω^Γ. Moreover, $|\nabla\phi| \geq m$ on Ω^Γ with some $m > 0$.*

Assumption 2. *$\Omega_h^\Gamma \subset \Omega^\Gamma$ and $|\nabla\phi_h| \geq \frac{m}{2}$ on all the mesh elements of Ω_h^Γ.*

Assumption 3. *The approximate boundary Γ_h can be covered by element patches $\{\Pi_k\}_{k=1,\dots,N_\Pi}$ having the following properties:*

- *Each Π_k is composed of a mesh element T_k lying inside Ω and some elements cut by Γ, more precisely $\Pi_k = T_k \cup \Pi_k^\Gamma$ where $T_k \in \mathcal{T}_h$, $T_k \subset \bar{\Omega}$, $\Pi_k^\Gamma \subset \mathcal{T}_h^\Gamma$, and Π_k^Γ contains at most M mesh elements;*
- *Each mesh element in a patch Π_k shares a facet with at least one other mesh element in the same patch (so that Π_k is a connected set). In particular, T_k shares a facet F_k with an element in Π_k^Γ;*
- *$\mathcal{T}_h^\Gamma = \cup_{k=1}^{N_\Pi} \Pi_k^\Gamma$ and $\Gamma_h^i = \cup_{k=1}^{N_\Pi} F_k$;*
- *Π_k and Π_l are disjoint if $k \neq l$.*

Assumption 3 prevents strong oscillations of Γ on the length scale h. It can be reformulated by saying that each cut element $T \in \mathcal{T}_h^\Gamma$ can be connected to an uncut element $T' \subset \Omega_h^i$ by a path consisting of a small number of mesh elements adjacent to one another; see [15] for a more detailed discussion and an illustration (Fig. 2).

Theorem 2.1. *Suppose that Assumptions 1–3 hold true, $l \geq k+1$, $\Omega \subset \Omega_h$ and $f \in H^k(\Omega_h)$. Let $u \in H^{k+2}(\Omega)$ be the solution to (1) and $(u_h, y_h, p_h) \in W_h^{(k)}$ be the solution to (6). Provided γ_{div}, γ_u, γ_p, σ are sufficiently big, it holds*

$$|u - u_h|_{1,\Omega} \leq Ch^k \|f\|_{k,\Omega_h} \quad \text{and} \quad \|u - u_h\|_{0,\Omega} \leq Ch^{k+1/2} \|f\|_{k,\Omega_h} \tag{7}$$

with $C > 0$ depending on the constants in Assumptions 1, 3 (and thus on the norm of ϕ in C^{k+2}), on the mesh regularity, on the polynomial degrees k and l, and on Ω, but independent of h, f, and u.

Remark 1 ((Condition $\Omega \subset \Omega_h$))**.** *The assumptions of Theorem 2.1 include $\Omega \subset \Omega_h$. Note that one would automatically have $\Omega \subset \Omega_h$, were Ω_h defined as the set of mesh cells having a non empty intersection with $\Omega = \{\phi < 0\}$. However, Ω_h is based on the intersections with $\{\phi_h < 0\}$ which can result in some rare situation where tiny portions of Ω lie outside Ω_h. In such a case, the a priori estimates (7)will control the error only on $\Omega \cap \Omega_h$.*

Remark 2 ((non-homogeneous Neumann and Robin conditions))**.** *We can also treat the case of more general boundary conditions:*

(i) non-homogeneous Neumann boundary conditions $\frac{\partial u}{\partial n} = g$ on Γ by adding the term

$$-\frac{\gamma_p}{h^2} \int_{\Omega_h^\Gamma} \tilde{g} |\nabla \phi_h| (z_h \cdot \nabla \phi_h + \frac{1}{h} q_h \phi_h)$$

in the right-hand side of (6) where $\tilde{g} \in H^{k+1}(\Omega_h^\Gamma)$ is lifting of g from Γ to a vicinity of Γ.

(ii) Robin boundary condition $\frac{\partial u}{\partial n} + \alpha u = g$ on Γ ($\alpha \in \mathbb{R}$) by replacing the penultimate term in a_h by

$$\frac{\gamma_p}{h^2} \int_{\Omega_h^\Gamma} (y \cdot \nabla \phi_h - |\nabla \phi_h| \alpha u + \frac{1}{h} p \phi_h)(z \cdot \nabla \phi_h - |\nabla \phi_h| \alpha v + \frac{1}{h} q \phi_h)$$

and by adding the term

$$-\frac{\gamma_p}{h^2} \int_{\Omega_h^\Gamma} \tilde{g} |\nabla \phi_h| (z_h \cdot \nabla \phi_h - |\nabla \phi_h| \alpha v + \frac{1}{h} q_h \phi_h)$$

in the right-hand side of (6) where $\tilde{g} \in H^{k+1}(\Omega_h^\Gamma)$ is defined as before.

Theorem 2.1 remains valid, adding $\|\tilde{g}\|_{k+1,\Omega_h^\Gamma}$ to $\|f\|_{k,\Omega_h}$ in (7). This framework will be used in first test case of the numerical simulations performed in Section 5: Fig. 2-8 for (i) and Fig. 9 for (ii).

3 Proof of the *a priori* error estimates

From now on, we shall use the letter C for positive constants (which can vary from one line to another) that depend only on the regularity of the mesh and on the constants in Assumptions 1–3.

We shall begin with some technical results, mostly adapted from [15] and [9] to be used later in the proofs of the coercivity of a_h (Section 3.2) and the *a priori* error estimates (Sections 3.3 and 3.4).

3.1 Technical lemmas

We recall first a lemma from [9]:

Lemma 3.1. *Let T be a triangle/tetrahedron, E one of its sides and p a polynomial on T such that $p = a$ on E for some $a \in \mathbb{R}$, $\frac{\partial p}{\partial n} = 0$ on E, and $\Delta p = 0$ on T. Then $p = a$ on T.*

We now adapt a lemma from [15]:

Lemma 3.2. *Let B_h be the strip between $\partial\Omega_h$ and Γ_h. For any $\beta > 0$, there exist $0 < \alpha < 1$ and $\delta > 0$ depending only on the mesh regularity and geometrical assumptions such that, for all $v_h \in V_h^{(k)}, z_h \in Z_h^{(k)}$*

$$\left|\int_{B_h} z_h \cdot \nabla v_h\right| \le \alpha |v_h|^2_{1,\Omega_h} + \delta \|z_h + \nabla v_h\|^2_{0,\Omega_h^\Gamma} + \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,\Gamma_h^i} + \beta h^2 \|\operatorname{div} z_h + v_h\|^2_{0,\Omega_h^\Gamma} + \beta h^2 \|v_h\|^2_{0,\Omega_h^\Gamma}. \quad (8)$$

Proof. The boundary Γ can be covered by element patches $\{\Pi_k\}_{k=1,\dots,N_\Pi}$ as in Assumption 3. Choose any $\beta > 0$ and consider

$$\alpha := \max_{\Pi_k, (z_h, v_h) \neq (0,0)} F(\Pi_k, z_h, v_h) \quad (9)$$

with

$$F(\Pi_k, z_h, v_h) = \frac{\|z_h\|_{0,\Pi_k^\Gamma} |v_h|_{1,\Pi_k^\Gamma} - \beta \|z_h + \nabla v_h\|^2_{0,\Pi_k^\Gamma} - \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,F_k} - \frac{\beta}{2} h^2 \|\operatorname{div} z_h\|^2_{0,\Pi_k^\Gamma}}{\frac{1}{2}\|z_h\|^2_{0,\Pi_k^\Gamma} + \frac{1}{2}|v_h|^2_{1,\Pi_k}},$$

where the maximum is taken over all the possible configurations of a patch Π_k allowed by the mesh regularity and over all $v_h \in V_h^{(k)}$ and $z_h \in Z_h^{(k)}$ restricted to Π_k. Note that $F(\Pi_k, z_h, v_h)$ is invariant under the scaling transformation $x \mapsto \frac{1}{h}x$, $v_h \mapsto \frac{1}{h} v_h$, $z_h \mapsto z_h$. We can thus assume $h = 1$ when computing the maximum in (9). Moreover, $F(\Pi_k, z_h, v_h)$ is homogeneous with respect to v_h, z_h, i.e. $F(\Pi_k, z_h, v_h) = F(\Pi_k, \mu z_h, \mu v_h)$ for any $\mu \neq 0$. Thus, the maximum in (9) is indeed attained since it can be taken over a closed bounded set in a finite dimensional space (all the admissible patches on a mesh with $h = 1$ and all v_h, z_h such that $|v_h|^2_{1,\Pi_k} + \|z_h\|^2_{0,\Pi_k^\Gamma} = 1$).

Clearly, $\alpha \le 1$. Supposing $\alpha = 1$ leads to a contradiction. Indeed, if $\alpha = 1$, we can then take Π_k, v_h, z_h yielding this maximum (in particular, $|v_h|^2_{1,\Pi_k} + \|z_h\|^2_{0,\Pi_k^\Gamma} > 0$). We observe then

$$\frac{1}{2}|v_h|^2_{1,\Pi_k} - \|z_h\|_{0,\Pi_k^\Gamma}|v_h|_{1,\Pi_k^\Gamma} + \frac{1}{2}\|z_h\|^2_{0,\Pi_k^\Gamma} + \beta\|z_h + \nabla v_h\|^2_{0,\Pi_k^\Gamma} + \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,F_k} + \frac{\beta}{2} h^2 \|\operatorname{div} z_h\|^2_{0,\Pi_k^\Gamma} = 0$$

and consequently (recall $|v_h|^2_{1,\Pi_k} = |v_h|^2_{1,T_k} + |v_h|^2_{1,\Pi_k^\Gamma}$)

$$\frac{1}{2}|v_h|^2_{1,T_k} + \beta\|z_h + \nabla v_h\|^2_{0,\Pi_k^\Gamma} + \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,F_k} + \frac{\beta}{2} h^2 \|\operatorname{div} z_h\|^2_{0,\Pi_k^\Gamma} = 0. \quad (10)$$

This implies $|v_h|_{1,T_k} = 0$ so that $v_h = \text{const}$ on T_k. Moreover, $\|z_h + \nabla v_h\|_{0,\Pi_k^\Gamma} = 0$ so that $\nabla v_h = -z_h$ on Π_k^Γ, hence ∇v_h is continuous on Π_k^Γ and $\Delta v_h = 0$ on Π_k^Γ since $\operatorname{div} z_h = 0$ there. The jump $\left[\frac{\partial v_h}{\partial n}\right]$ vanishes also on the facet F_k separating T_k from Π_k^Γ, as implied directly by (10). Combining these observations with Lemma 3.1, starting from T_k and its neighbor in Π_k^Γ and then propagating to other elements of Π_k^Γ, we see that $v_h = \text{const}$ on the whole Π_k. We have thus $\nabla v_h = 0$ on Π_k and $z_h = 0$ on Π_k^Γ, which is in contradiction with $|v_h|^2_{1,\Pi_k} + \|z_h\|^2_{0,\Pi_k^\Gamma} > 0$.

Thus $\alpha < 1$ and

$$\|z_h\|_{0,\Pi_k^\Gamma}|v_h|_{1,\Pi_k^\Gamma} \le \frac{\alpha}{2}\|z_h\|^2_{0,\Pi_k^\Gamma} + \frac{\alpha}{2}|v_h|^2_{1,\Pi_k} + \beta\|z_h + \nabla v_h\|^2_{0,\Pi_k^\Gamma} + \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,\partial T_k \cap \partial\Pi_k^\Gamma} + \frac{\beta}{2} h^2 \|\operatorname{div} z_h\|^2_{0,\Pi_k^\Gamma}$$

for all v_h, z_h and all admissible patches Π_k. We now observe

$$\begin{aligned}\left|\int_{B_h} z_h \cdot \nabla v_h\right| &\le \sum_k \left|\int_{B_h \cap \Pi_k^\Gamma} z_h \cdot \nabla v_h\right| \le \sum_k \|z_h\|_{0,\Pi_k^\Gamma}|v_h|_{1,\Pi_k^\Gamma} \\ &\le \frac{\alpha}{2}\|z_h\|^2_{0,\Omega_h^\Gamma} + \frac{\alpha}{2}|v_h|^2_{1,\Omega_h} + \beta\|z_h + \nabla v_h\|^2_{0,\Omega_h^\Gamma} + \beta h \left\| \left[\frac{\partial v_h}{\partial n}\right] \right\|^2_{0,\Gamma_h^i} + \frac{\beta}{2} h^2 \|\operatorname{div} z_h\|^2_{0,\Omega_h^\Gamma}.\end{aligned}$$

We now use the Young inequality with any $\varepsilon > 0$ to obtain

$$\|z_h\|^2_{0,\Omega_h^\Gamma} = \|z_h+\nabla v_h\|^2_{0,\Omega_h^\Gamma}+\|\nabla v_h\|^2_{0,\Omega_h^\Gamma}-2(z_h+\nabla v_h,\nabla v_h)_{0,\Omega_h^\Gamma} \le \left(1+\frac{1}{\varepsilon}\right)\|z_h+\nabla v_h\|^2_{0,\Omega_h^\Gamma}+(1+\varepsilon)|v_h|^2_{1,\Omega_h},$$

which leads to

$$\left|\int_{B_h} z_h\cdot\nabla v_h\right| \le \alpha\left(1+\frac{\varepsilon}{2}\right)|v_h|^2_{1,\Omega_h}+\left(\beta+\frac{\alpha}{2}+\frac{\alpha}{2\varepsilon}\right)\|z_h+\nabla v_h\|^2_{0,\Omega_h^\Gamma}+\beta h\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|^2_{0,\Gamma_h^i}+\beta h^2\|\operatorname{div} z_h\|^2_{0,\Omega_h^\Gamma}.$$

Taking ε sufficiently small, redefining α as $\alpha\left(1+\frac{\varepsilon}{2}\right)$ and putting $\delta = \left(\beta+\frac{\alpha}{2}+\frac{\alpha}{2\varepsilon}\right)$ we obtain

$$\left|\int_{B_h} z_h\cdot\nabla v_h\right| \le \alpha|v_h|^2_{1,\Omega_h}+\delta\|z_h+\nabla v_h\|^2_{0,\Omega_h^\Gamma}+\beta h\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|^2_{0,\Gamma_h^i}+\beta h^2\|\operatorname{div} z_h\|^2_{0,\Omega_h^\Gamma}.$$

This leads to (8) by the triangle inequality $\|\operatorname{div} z_h\|_{0,\Omega_h^\Gamma} \le \|\operatorname{div} z_h + v_h\|_{0,\Omega_h^\Gamma}+\|v_h\|_{0,\Omega_h^\Gamma}$. □

Lemma 3.3. *For all $v\in H^1(\Omega_h^\Gamma)$, $\|v\|_{0,\Omega_h^\Gamma} \le C\left(\sqrt{h}\|v\|_{0,\Gamma_h^i}+h|v|_{1,\Omega_h^\Gamma}\right)$*
and for all $v\in H^1(\Omega_h\backslash\Omega)$, $\|v\|_{0,\Omega_h\backslash\Omega} \le C\left(\sqrt{h}\|v\|_{0,\Gamma}+h|v|_{1,\Omega_h\backslash\Omega}\right)$.

We refer to [15] for the first inequality. The second one can be treated similarly.

The following lemma is borrowed from [9]. It's a partial generalization of Lemma 3.3 to derivatives of higher order.

Lemma 3.4. *Under Assumption 1, it holds for all $v\in H^s(\Omega_h)$ with integer $1\le s\le k+1$, v vanishing on Ω, $\|v\|_{0,\Omega_h\backslash\Omega} \le Ch^s\|v\|_{s,\Omega_h\backslash\Omega}$.*

Lemma 3.5. *For all piecewise polynomial (possibly discontinuous) functions v_h on $\mathcal{T}_h^\Gamma$, $\|v_h\|_{0,\Gamma_h} \le \frac{C}{\sqrt{h}}\|v_h\|_{0,\Omega_h^\Gamma}$ with a constant $C>0$ depending on the maximal degree of polynomials in v_h and on the constants in Assumptions 1–3.*

Proof. A scaling argument on all $T\in\mathcal{T}_h^\Gamma$. □

Finally, we recall a Hardy-type lemma, cf. [9].

Lemma 3.6. *Assume that the domain Ω^Γ is a neighborhood of Γ, given by (2), and satisfies Assumption 1. Then, for any $u\in H^{s+1}(\Omega^\Gamma)$ vanishing on Γ and an integer $s\in[0,k]$, it holds $\left\|\frac{u}{\phi}\right\|_{s,\Omega^\Gamma} \le C\|u\|_{s+1,\Omega^\Gamma}$ with $C>0$ depending only on the constants in Assumption 1 and on s.*

3.2 Coercivity of the bilinear form a

It will be convenient to rewrite the bilinear form a_h in a manner avoiding the integral on $\partial\Omega_h$. To this end, we recall that B_h is the strip between $\partial\Omega_h$ and Γ_h and observe for any $y\in H^1(B_h)^d$, $v\in H^1(B_h)$, $q\in L^2(\Gamma_h)$:

$$\begin{aligned}\int_{\partial\Omega_h} y\cdot nv &= \int_{\partial\Omega_h} y\cdot nv-\int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}(y\cdot\nabla\phi_h)v+\int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}(y\cdot\nabla\phi_h+\frac{1}{h}q\phi_h)v\\ &= \int_{B_h}(v\operatorname{div} y+y\cdot\nabla v)+\int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}(y\cdot\nabla\phi_h+\frac{1}{h}q\phi_h)v.\end{aligned}$$

Indeed, $\phi_h = 0$ on Γ_h and the unit normal to Γ_h, looking outward from B_h, is equal to $-\nabla\phi_h/|\nabla\phi_h|$. Thus,

$$a_h(u,y,p;v,z,q) = \int_{\Omega_h} \nabla u\cdot\nabla v + \int_{\Omega_h} uv + \int_{B_h} (v\,\mathrm{div}\, y + y\cdot\nabla v)$$
$$+ \int_{\Gamma_h} \frac{1}{|\nabla\phi_h|}(y\cdot\nabla\phi_h + \frac{1}{h}q\phi_h)v + \gamma_{div}\int_{\Omega_h^\Gamma}(\mathrm{div}\, y + u)(\mathrm{div}\, z + v) + \gamma_u\int_{\Omega_h^\Gamma}(y+\nabla u)\cdot(z+\nabla v)$$
$$+ \sigma h\int_{\Gamma_h^i}\left[\frac{\partial u}{\partial n}\right]\left[\frac{\partial v}{\partial n}\right] + \frac{\gamma_p}{h^2}\int_{\Omega_h^\Gamma}(y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h)(z\cdot\nabla\phi_h + \frac{1}{h}q\phi_h). \quad (11)$$

Proposition 1. *Provided $\gamma_{div},\gamma_u,\gamma_p,\sigma$ are sufficiently big, there exists an h-independent constant $c>0$ such that*

$$a_h(v_h,z_h,q_h;v_h,z_h,q_h) \geq c|||v_h,z_h,q_h|||_h^2, \quad \forall (v_h,z_h,q_h)\in W_h^{(k)}$$

with

$$|||v,z,q|||_h^2 = \|v\|_{1,\Omega_h}^2 + \|\,\mathrm{div}\, z + v\|_{0,\Omega_h^\Gamma}^2 + \|z+\nabla v\|_{0,\Omega_h^\Gamma}^2 + h\left\|\left[\frac{\partial v}{\partial n}\right]\right\|_{0,\Gamma_h^i}^2 + \frac{1}{h^2}\left\|z\cdot\nabla\phi_h + \frac{1}{h}q\phi_h\right\|_{0,\Omega_h^\Gamma}^2.$$

Proof. Using the reformulation of the bilinear form a_h given by (11), we have for all $(v_h,z_h,q_h)\in W_h^{(k)}$,

$$a_h(v_h,z_h,q_h;v_h,z_h,q_h) = |v_h|_{1,\Omega_h}^2 + \|v_h\|_{0,\Omega_h}^2 + \int_{B_h}(v_h\,\mathrm{div}\, z_h + z_h\cdot\nabla v_h)$$
$$+ \int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}(z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h)v_h + \gamma_{div}\|\,\mathrm{div}\, z_h + v_h\|_{0,\Omega_h^\Gamma}^2 + \gamma_u\|z_h+\nabla v_h\|_{0,\Omega_h^\Gamma}^2$$
$$+ \sigma h\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|_{0,\Gamma_h^i}^2 + \frac{\gamma_p}{h^2}\|z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h\|_{0,\Omega_h^\Gamma}^2.$$

Since $B_h\subset\Omega_h^\Gamma$, we remark that the integral of $v_h\,\mathrm{div}\, z_h$ can be combined with that of v_h on Ω_h^Γ to give

$$\|v_h\|_{0,\Omega_h^\Gamma}^2 + \int_{B_h} v_h\,\mathrm{div}\, z_h \geq \int_{B_h} v_h\,(\mathrm{div}\, z_h + v_h) \geq -\|v_h\|_{0,\Omega_h^\Gamma}\|\,\mathrm{div}\, z_h + v_h\|_{0,\Omega_h^\Gamma}.$$

We also use an inverse inequality from Lemma 3.5 and the fact that $1/|\nabla\phi_h|$ is uniformly bounded by Assumption 2, to estimate

$$\left|\int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}(z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h)v_h\right| \leq \frac{C}{h}\|z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h\|_{0,\Omega_h^\Gamma}\|v_h\|_{0,\Omega_h^\Gamma}.$$

Applying the Young inequality (for any $\varepsilon>0$) to the last two bounds and combining this with (8) yields

$$a_h(v_h,z_h,q_h;v_h,z_h,q_h) \geq (1-\alpha)|v_h|_{1,\Omega_h}^2 + \|v_h\|_{0,\Omega_h^i}^2 - (\varepsilon+\beta h^2)\|v_h\|_{0,\Omega_h^\Gamma}^2 + \left(\gamma_{div} - \frac{1}{2\varepsilon} - \beta h^2\right)\|\,\mathrm{div}\, z_h + v_h\|_{0,\Omega_h^\Gamma}^2$$
$$+ (\gamma_u-\delta)\|z_h+\nabla v_h\|_{0,\Omega_h^\Gamma}^2 + (\sigma-\beta)h\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|_{0,\Gamma_h^i}^2 + \left(\frac{\gamma_p}{h^2} - \frac{C^2}{2\varepsilon h^2}\right)\|z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h\|_{0,\Omega_h^\Gamma}^2.$$

To bound further from below the first 3 terms we note, using Lemma 3.3 and the trace inverse inequality,

$$\|v_h\|_{0,\Omega_h^\Gamma}^2 \leq C(h\|v_h\|_{0,\Gamma_h^i}^2 + h^2|v_h|_{1,\Omega_h^\Gamma}^2) \leq C(\|v_h\|_{0,\Omega_h^i}^2 + h^2|v_h|_{1,\Omega_h}^2)$$

so that, introducing any $\kappa \geq 0$ and observing $h \leq h_0 := \mathrm{diam}(\Omega)$,

$$\begin{aligned}(1-\alpha)|v_h|^2_{1,\Omega_h} + \|v_h\|^2_{0,\Omega_h^i} - (\varepsilon+\beta h^2)\|v_h\|^2_{0,\Omega_h^\Gamma} \\ \geq (1-\alpha)|v_h|^2_{1,\Omega_h} + \|v_h\|^2_{0,\Omega_h^i} + \kappa\|v_h\|^2_{0,\Omega_h^\Gamma} - (\varepsilon+\beta h_0^2+\kappa)\|v_h\|^2_{0,\Omega_h^\Gamma} \\ \geq (1-\alpha-C(\varepsilon+\beta h_0^2+\kappa)h_0^2)|v_h|^2_{1,\Omega_h} + (1-C(\varepsilon+\beta h_0^2+\kappa))\|v_h\|^2_{0,\Omega_h^i} + \kappa\|v_h\|^2_{0,\Omega_h^\Gamma}.\end{aligned}$$

Taking $\varepsilon, \kappa, \beta$ sufficiently small and $\gamma_u, \gamma_p, \gamma_{div}$ sufficiently big, gives the announced lower bound for $a_h(v_h, z_h, q_h; v_h, z_h, q_h)$. □

3.3 Proof of the H^1 error estimate in Theorem 2.1

Under the Theorem's assumptions, the solution to (1) is indeed in $H^{k+2}(\Omega)$ and it can be extended to a function $\tilde{u} \in H^{k+2}(\Omega_h)$ such that $\tilde{u} = u$ on Ω and

$$\|\tilde{u}\|_{k+2,\Omega_h} \leq C(\|f\|_{k,\Omega} + \|g\|_{k+1/2,\Gamma}) \leq \|f\|_{k,\Omega}. \tag{12}$$

Introduce $y = -\nabla\tilde{u}$ and $p = -\frac{h}{\phi} y \cdot \nabla\phi$ on Ω_h^Γ. Then, $y \in H^{k+1}(\Omega_h^\Gamma)$ and $p \in H^k(\Omega_h^\Gamma)$ by Lemma 3.6. Moreover,

$$\|y\|_{k+1,\Omega_h^\Gamma} \leq C\|\tilde{u}\|_{k+2,\Omega_h} \leq C\|f\|_{k,\Omega} \quad \text{and} \quad \|p\|_{k,\Omega_h^\Gamma} \leq Ch\|y\|_{k+1,\Omega_h^\Gamma} \leq Ch\|f\|_{k,\Omega}. \tag{13}$$

Clearly, $\tilde{u}$, y, p satisfy

$$a_h(\tilde{u}, y, p; v_h, z_h, q_h) = \int_{\Omega_h} \tilde{f} v_h + \gamma_{div}\int_{\Omega_h^\Gamma} \tilde{f}(\mathrm{div}\, z_h + v_h) + \frac{\gamma_p}{h^2}\int_{\Omega_h^\Gamma} (y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h)(z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h),$$
$$\forall (v_h, z_h, q_h) \in W_h^{(k)}$$

with $\tilde{f} := -\Delta\tilde{u} + \tilde{u}$. It entails a Galerkin orthogonality relation

$$\begin{aligned}a_h(\tilde{u}-u_h, y-y_h, p-p_h; v_h, z_h, q_h) = \int_{\Omega_h} (\tilde{f}-f)v_h + \gamma_{div}\int_{\Omega_h^\Gamma} (\tilde{f}-f)(\mathrm{div}\, z_h + v_h) \\ + \frac{\gamma_p}{h^2}\int_{\Omega_h^\Gamma} (y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h)(z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h), \quad \forall (v_h, z_h, q_h) \in W_h^{(k)}. \end{aligned} \tag{14}$$

Introducing the standard nodal interpolation I_h or, if necessary, a Clément interpolation (recall that p is only in $H^1(\Omega_h^\Gamma)$ if $k = 1$), we then have by Proposition 1,

$$\begin{aligned}c\,|||\, u_h - I_h\tilde{u}, y_h - I_h y, p_h - I_h p|||_h &\leq \sup_{(v_h,z_h,q_h)\in W_h^{(k)}} \frac{a_h(u_h - I_h\tilde{u}, y_h - I_h y, p_h - I_h p; v_h, z_h, q_h)}{|||v_h, z_h, q_h|||_h} \\ &\leq \sup_{(v_h,z_h,q_h)\in W_h^{(k)}} \frac{I - II - III}{|||v_h, z_h, q_h|||_h},\end{aligned}$$

where

$$I = a_h(e_u, e_y, e_p; v_h, z_h, q_h), \quad II = \int_{\Omega_h} (\tilde{f}-f)v_h + \gamma_{div}\int_{\Omega_h^\Gamma} (\tilde{f}-f)(\mathrm{div}\, z_h + v_h),$$
$$III = \frac{\gamma_p}{h^2}\int_{\Omega_h^\Gamma} (y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h)(z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h),$$

with $e_u = \tilde{u} - I_h\tilde{u},\ e_y = y - I_h\tilde{y}$ and $e_p = p - I_h\tilde{p}$.

We now estimate each term separately. Recalling (11), we have

$$\begin{aligned} I \le \|e_u\|_{1,\Omega_h}\|v_h\|_{1,\Omega_h} + \|\mathrm{div} e_y\|_{0,B_h}\|v_h\|_{0,B_h} + \|e_y\|_{0,B_h}|v_h|_{1,B_h} \\ + \|\frac{1}{|\nabla\phi_h|}(e_y\cdot\nabla\phi_h + \frac{1}{h}e_p\phi_h)\|_{0,\Gamma_h}\|v_h\|_{0,\Gamma_h} + \gamma_{div}\|\,\mathrm{div}\, e_y + e_u\|_{0,\Omega_h^\Gamma}\|\,\mathrm{div}\, z_h + u_h\|_{0,\Omega_h^\Gamma} \\ + \gamma_u\|e_y + \nabla e_u\|_{0,\Omega_h^\Gamma}\|z_h + \nabla v_h\|_{0,\Omega_h^\Gamma} + \sigma h\left\|\left[\frac{\partial e_u}{\partial n}\right]\right\|_{0,\Gamma_h^i}\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|_{0,\Gamma_h^i} \\ + \frac{\gamma_p}{h^2}\|e_y\cdot\nabla\phi_h + \frac{1}{h}e_p\phi_h\|_{0,\Omega_h^\Gamma}\|z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h\|_{0,\Omega_h^\Gamma}. \end{aligned}$$

Applying Lemma 3.5 to the L^2 norms on Γ_h, recalling that $1/|\nabla\phi_h|$ is uniformly bounded on Ω_h^Γ (cf. Assumption 2), and recombining the terms, we get

$$I \le C\left(\|e_u\|^2_{1,\Omega_h} + \|e_y\|^2_{1,\Omega_h^\Gamma} + h\left\|\left[\frac{\partial e_u}{\partial n}\right]\right\|^2_{0,\Gamma_h^i} + \frac{1}{h^2}\|e_y\cdot\nabla\phi_h + \frac{1}{h}e_p\phi_h\|^2_{0,\Omega_h^\Gamma}\right)^{1/2} |||\, v_h, z_h, q_h\,|||_h\,.$$

The usual interpolation estimates give

$$\|e_u\|^2_{1,\Omega_h} + \|e_y\|^2_{1,\Omega_h^\Gamma} + h\left\|\left[\tfrac{\partial e_u}{\partial n}\right]\right\|^2_{0,\Gamma_h^i} \le Ch^{2k}(\|\tilde{u}\|^2_{k+1,\Omega_h} + \|y\|^2_{k+1,\Omega_h^\Gamma})\,.$$

Moreover, recalling that $|\nabla\phi_h|$ and $\frac{1}{h}|\phi_h|$ are uniformly bounded on Ω_h^Γ, we get

$$\frac{1}{h^2}\|e_y\cdot\nabla\phi_h + \frac{1}{h}e_p\phi_h\|^2_{0,\Omega_h^\Gamma} \le \frac{C}{h^2}\left(\|e_y\|^2_{0,\Omega_h^\Gamma} + \|e_p\|^2_{0,\Omega_h^\Gamma}\right) \le Ch^{2k}(|y|^2_{k+1,\Omega_h^\Gamma} + \frac{1}{h^2}|p|^2_{k,\Omega_h^\Gamma}).$$

Thus, by regularity estimates (12), $I \le Ch^k\|f\|_{k,\Omega}\,|||\, v_h, z_h, q_h\,|||_h\,.$

We now estimate the second term

$$\begin{aligned} |II| &\le C(\|\tilde{f} - f\|_{0,\Omega_h}\|v_h\|_{0,\Omega_h} + \|\tilde{f} - f\|_{0,\Omega_h^\Gamma}\|\,\mathrm{div}\, z_h + v_h\|_{0,\Omega_h^\Gamma}) \\ &\le C\|\tilde{f} - f\|_{0,\Omega_h}\,|||\, v_h, z_h, q_h|||_h \le Ch^k\|f\|_{k,\Omega\cup\Omega_h}\,|||\, v_h, z_h, q_h\,|||_h\,. \end{aligned}$$

Indeed, thanks to Lemma 3.4 and $f = \tilde{f}$ on Ω,

$$\|\tilde{f} - f\|_{0,\Omega_h} = \|\tilde{f} - f\|_{0,\Omega_h\setminus\Omega} \le Ch^k\|\tilde{f} - f\|_{k,\Omega_h\setminus\Omega} \le Ch^k\|f\|_{k,\Omega\cup\Omega_h}. \tag{15}$$

Finally,

$$|III| \le \frac{C}{h}\|y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h\|_{0,\Omega_h^\Gamma}\,|||\, v_h, z_h, q_h|||_h$$

and, recalling $y\cdot\nabla\phi + \frac{1}{h}p\phi = 0$ on Ω_h^Γ,

$$\begin{aligned} \frac{1}{h}\|y\cdot\nabla\phi_h + \frac{1}{h}p\phi_h\|_{0,\Omega_h^\Gamma} &= \frac{1}{h}\|y\cdot\nabla(\phi_h - \phi) + \frac{1}{h}p(\phi_h - \phi)\|_{0,\Omega_h^\Gamma} \\ &\le \frac{1}{h}\|y\|_{0,\Omega_h^\Gamma}\|\nabla(\phi_h - \phi)\|_\infty + \frac{1}{h^2}\|p\|_{0,\Omega_h^\Gamma}\|\phi_h - \phi\|_\infty \\ &\le Ch^k(\|y\|_{0,\Omega_h^\Gamma} + \|p\|_{0,\Omega_h^\Gamma}) \le Ch^k\|f\|_{k,\Omega} \end{aligned}$$

by regularity estimates (13). Note that the optimal order is achieved here since ϕ is assumed of regularity C^{k+2} and it is approximated by finite elements of degree at least $k+1$.

Combining the estimate for the terms I–III leads to

$$|||u_h - I_h\tilde{u}, y_h - I_h y, p_h - I_h p|||_h \le Ch^k\|f\|_{k,\Omega\cup\Omega_h},$$

so that, by the triangle inequality together with interpolation estimate, we get

$$|||u_h - \tilde{u}, y_h - y, p_h - p|||_h \le Ch^k\|f\|_{k,\Omega\cup\Omega_h}. \tag{16}$$

This implies the announced H^1 error estimate for $u - u_h$.

3.4 Proof of the L^2 error estimate in Theorem 2.1

Since $\Omega \subset \Omega_h$, we can introduce $w : \Omega \to \mathbb{R}$ such that

$$-\Delta w + w = u - u_h \text{ in } \Omega, \quad \frac{\partial w}{\partial n} = 0 \text{ on } \Gamma.$$

By elliptic regularity, $\|w\|_{2,\Omega} \leq C\|u - u_h\|_{0,\Omega}$. Let $\tilde{w}$ be an extension of w from Ω to Ω_h preserving the H^2 norm estimate and set $w_h = I_h\tilde{w}$. We observe

$$\|u-u_h\|^2_{0,\Omega} = \int_\Omega \nabla(u-u_h)\cdot\nabla(w-w_h) + \int_\Omega (u-u_h)(w-w_h) + \int_\Omega \nabla(u-u_h)\cdot\nabla w_h + \int_\Omega (u-u_h)w_h \leq Ch^{k+1}\|f\|_{k,\Omega_h}|\tilde{w}|_{2,\Omega_h} + \left|\int_\Omega \nabla(u-u_h)\cdot\nabla w_h + \int_\Omega (u-u_h)w_h\right|$$

by the already proven H^1 error estimate and interpolation estimates for $I_h\tilde{w}$ (recall also $\Omega \subset \Omega_h$). Taking $v_h = w_h$, $z_h = 0$ and $q_h = 0$ in the Galerkin orthogonality relation (14), we obtain, thanks to (11),

$$\int_{\Omega_h} \nabla(\tilde{u}-u_h)\cdot\nabla w_h + \int_{\Omega_h}(\tilde{u}-u_h)w_h + \int_{B_h}(w_h \operatorname{div}(y-y_h) + (y-y_h)\cdot\nabla w_h) + \int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}((y-y_h)\cdot\nabla\phi_h + \frac{1}{h}(p-p_h)\phi_h)w_h + \gamma_{div}\int_{\Omega_h^\Gamma}(\operatorname{div}(y-y_h)+\tilde{u}-u_h)w_h + \gamma_u\int_{\Omega_h^\Gamma}((y-y_h)+\nabla(\tilde{u}-u_h))\cdot\nabla w_h + \sigma h\int_{\Gamma_h^i}\left[\frac{\partial(\tilde{u}-u_h)}{\partial n}\right]\left[\frac{\partial w_h}{\partial n}\right] = (1+\gamma_{div})\int_{\Omega_h}(\tilde{f}-f)w_h.$$

Using the last relation in the bound for $\|u-u_h\|^2_{0,\Omega}$, we can further bound it as

$$\begin{aligned}\|u-u_h\|^2_{0,\Omega} \leqslant\ & Ch^{k+1}\|f\|_{k,\Omega_h}|\tilde{w}|_{2,\Omega_h} + \left|\int_{\Omega_h\setminus\Omega}\nabla(\tilde{u}-u_h)\cdot\nabla w_h + \int_{\Omega_h\setminus\Omega}(\tilde{u}-u_h)w_h\right| \\ &+ \left|\int_{B_h}(w_h\operatorname{div}(y-y_h)+(y-y_h)\cdot\nabla w_h)\right| + \left|\int_{\Gamma_h}\frac{1}{|\nabla\phi_h|}((y-y_h)\cdot\nabla\phi_h + \frac{1}{h}(p-p_h)\phi_h)w_h\right| \\ &+ \left|\gamma_{div}\int_{\Omega_h^\Gamma}(\operatorname{div}(y-y_h)+\tilde{u}-u_h)w_h\right| + \left|\gamma_u\int_{\Omega_h^\Gamma}((y-y_h)+\nabla(\tilde{u}-u_h))\cdot\nabla w_h\right| \\ &+ \left|\sigma h\int_{\Gamma_h^i}\left[\frac{\partial(\tilde{u}-u_h)}{\partial n}\right]\left[\frac{\partial w_h}{\partial n}\right]\right| + (1+\gamma_{div})\left|\int_{\Omega_h}(\tilde{f}-f)w_h\right| \\ \leqslant\ & Ch^{k+1}\|f\|_{k,\Omega_h}|\tilde{w}|_{2,\Omega_h} + C\,|||\,\tilde{u}-u_h, y-y_h, p-p_h\,|||_h \times \big(\|w_h\|_{1,\Omega_h\setminus\Omega} \\ &+\|w_h\|_{1,\Omega_h^\Gamma} + h\|w_h\|_{0,\Gamma_h} + \sqrt{h}\|[\nabla w_h]\|_{0,\Gamma_h^i}\big) + C\|\tilde{f}-f\|_{0,\Omega_h\setminus\Omega}\|w_h\|_{1,\Omega_h\setminus\Omega}.\end{aligned}$$

It remains to bound different norms of w_h featuring in the estimate above. By Lemma 3.3 and interpolation estimates

$$\|w_h\|_{0,\Omega_h\setminus\Omega} \leq \|\tilde{w}-I_h\tilde{w}\|_{0,\Omega_h\setminus\Omega} + \|\tilde{w}\|_{0,\Omega_h\setminus\Omega} \leq Ch^2|\tilde{w}|_{2,\Omega_h\setminus\Omega} + C\left(\sqrt{h}\|\tilde{w}\|_{0,\Gamma} + h|\tilde{w}|_{1,\Omega_h\setminus\Omega}\right) \leq C\sqrt{h}\|\tilde{w}\|_{2,\Omega_h}.$$

Similarly,

$$\|\nabla w_h\|_{0,\Omega_h\setminus\Omega} \leq \|\nabla(\tilde{w}-I_h\tilde{w})\|_{0,\Omega_h\setminus\Omega} + \|\nabla\tilde{w}\|_{0,\Omega_h\setminus\Omega} \leq Ch|\tilde{w}|_{2,\Omega_h\setminus\Omega} + C\left(\sqrt{h}\|\nabla\tilde{w}\|_{0,\Gamma} + h|\nabla\tilde{w}|_{1,\Omega_h\setminus\Omega}\right) \leq C\sqrt{h}\|\tilde{w}\|_{2,\Omega_h}.$$

Analogous estimates also hold for $\|w_h\|_{1,\Omega_h^\Gamma}$. Moreover, by interpolation estimates,

$$\|[\nabla w_h]\|_{0,\Gamma_h^i} = \|[\nabla(\tilde{w} - I_h\tilde{w})]\|_{0,\Gamma_h^i} \leqslant C\sqrt{h}|\tilde{w}|_{2,\Omega_h}$$

and, by Lemma 3.5,

$$h\|w_h\|_{0,\Gamma_h} \leq C\sqrt{h}\|w_h\|_{0,\Omega_h^\Gamma} \leq C\sqrt{h}\|\tilde{w}\|_{2,\Omega_h}.$$

Hence,

$$\|u - u_h\|_{0,\Omega}^2 \leq Ch^{k+1}\|f\|_{k,\Omega_h}|\tilde{w}|_{2,\Omega_h} + C\sqrt{h}(|||\tilde{u} - u_h, y - y_h, p - p_h|||_h + \|\tilde{f} - f\|_{0,\Omega_h\setminus\Omega})\|\tilde{w}\|_{2,\Omega_h}.$$

This implies, by (15) and (16), $\|u - u_h\|_{0,\Omega}^2 \leq Ch^{k+\frac{1}{2}}\|f\|_{k,\Omega_h}\|\tilde{w}\|_{2,\Omega_h}$, which entails the announced error estimate in $L^2(\Omega)$ since $\|\tilde{w}\|_{2,\Omega_h} \leq C\|u - u_h\|_{0,\Omega}$.

4 Conditioning

We are now going to prove that the condition number of the finite element matrix associated to the bilinear form a_h is of order $1/h^2$.

Theorem 4.1. *Under Assumptions 1–3 and recalling that the mesh $\mathcal{T}_h$ is quasi-uniform, the condition number defined by $\kappa(\mathbf{A}) := \|\mathbf{A}\|_2\|\mathbf{A}^{-1}\|_2$ of the matrix $\mathbf{A}$ associated to the bilinear form a_h on $W_h^{(k)}$ satisfies $\kappa(\mathbf{A}) \leq Ch^{-2}$. Here, $\|\cdot\|_2$ stands for the matrix norm associated to the vector 2-norm $|\cdot|_2$.*

Proof. The proof is divided into 4 steps:

Step 1. We shall prove for all $q_h \in Q_h^{(k)}$

$$\|q_h\phi_h\|_{0,\Omega_h^\Gamma} \geq Ch\|q_h\|_{0,\Omega_h^\Gamma}. \tag{17}$$

We have

$$\min_{T, q_h\neq 0, \phi_h \neq 0} \frac{\|q_h\phi_h\|_{0,T}}{h_T\|q_h\|_{0,T}\|\nabla\phi_h\|_{\infty,T}} \geqslant C, \tag{18}$$

where the minimum is taken over all simplexes T with $h_T = \operatorname{diam}(T)$ satisfying the regularity assumptions and all polynomials q_h of degree $\leqslant k$ and ϕ_h of degree $\leqslant l$, with ϕ_h vanishing at at least one point on T. Note that this excludes $\|\nabla\phi_h\|_{\infty,T} = 0$ because ϕ_h would then vanish identically on T. The minimum in (18) is indeed attained since, by homogeneity, it can be taken over the compact set $\|q_h\|_{0,T} = \|\nabla\phi_h\|_{\infty,T} = 1$ and simplexes with $h_T = 1$. Hence, (18) is valid with some $C > 0$. Applying (18) on any mesh element $T \in \mathcal{T}_h^\Gamma$ to any $q_h \in Q_h^{(k)}$ and ϕ_h approximation to ϕ satisfying Assumption 2 leads to $\|q_h\phi_h\|_{0,T} \geqslant Ch_T^{\frac{m}{2}}\|q_h\|_{0,T}$. Taking the square on both sides and summing over all $T \in \mathcal{T}_h^\Gamma$ yields (17).

Step 2. We shall prove for all $(v_h, z_h, q_h) \in W_h^{(k)}$

$$a_h(v_h, z_h, q_h; v_h, z_h, q_h) \geq c\|v_h, z_h, q_h\|_0^2 \tag{19}$$

with $\|v_h, z_h, q_h\|_0^2 = \|v_h\|_{0,\Omega_h}^2 + \|z_h\|_{0,\Omega_h^\Gamma}^2 + \|q_h\|_{0,\Omega_h^\Gamma}^2$. Indeed, by Lemma 1,

$$a_h(v_h, z_h, q_h; v_h, z_h, q_h) \geq c|||v_h, z_h, q_h|||_h^2 \geq c(\|v_h\|_{1,\Omega_h}^2 + \|z_h + \nabla v_h\|_{0,\Omega_h^\Gamma}^2 + \|z_h\cdot\nabla\phi_h + \frac{1}{h}q_h\phi_h\|_{0,\Omega_h^\Gamma}^2).$$

We have assumed here (without loss of generality) $h \leq 1$. By Young's inequality with any $\epsilon_1 \in (0,1)$,

$$\|z_h + \nabla v_h\|_{0,\Omega_h^\Gamma}^2 = \|z_h\|_{0,\Omega_h^\Gamma}^2 + \|\nabla v_h\|_{0,\Omega_h^\Gamma}^2 + 2(z_h, \nabla v_h)_{0,\Omega_h^\Gamma} \geq (1-\epsilon_1)\|z_h\|_{0,\Omega_h^\Gamma}^2 - \frac{1-\epsilon_1}{\epsilon_1}\|\nabla v_h\|_{0,\Omega_h^\Gamma}^2. \tag{20}$$

Similarly, for any $\epsilon_2 \in (0,1)$, using that $\nabla\phi_h$ is uniformly bounded,

$$\|z_h \cdot \nabla\phi_h + \frac{1}{h} q_h\phi_h\|^2_{0,\Omega_h^\Gamma} \geq \frac{1-\epsilon_2}{h^2}\|\phi_h q_h\|^2_{0,\Omega_h^\Gamma} - C\frac{1-\epsilon_2}{\epsilon_2}\|z_h\|^2_{0,\Omega_h^\Gamma}. \tag{21}$$

Thus, combining (20), (21) and (17),

$$\begin{aligned} &a_h(v_h, z_h, q_h; v_h, z_h, q_h) \\ &\qquad \geq c\left(\left(1 - \frac{1-\epsilon_1}{\epsilon_1}\right)\|v_h\|^2_{1,\Omega_h} + \left(1 - \epsilon_1 - C\frac{1-\epsilon_2}{\epsilon_2}\right)\|z_h\|^2_{0,\Omega_h^\Gamma} + C(1-\epsilon_2)\|q_h\|^2_{0,\Omega_h^\Gamma}\right). \end{aligned}$$

Taking ϵ_1, ϵ_2 close to 1, we get (19).

Step 3. We shall prove for all $(v_h, z_h, q_h) \in W_h^{(k)}$

$$a_h(v_h, z_h, q_h; v_h, z_h, q_h) \leq \frac{C}{h^2}\|v_h, z_h, q_h\|_0^2. \tag{22}$$

By definition of a_h (11) and Cauchy-Schwarz inequality,

$$\begin{aligned} &a_h(v_h, z_h, q_h; v_h, z_h, q_h) \\ &\qquad \leq C\left(\|v_h\|^2_{1,\Omega_h} + \|z_h\|^2_{1,\Omega_h^\Gamma} + h\left\|\left[\frac{\partial v_h}{\partial n}\right]\right\|^2_{0,\Gamma_h^i} + \frac{1}{h^2}\|z_h\cdot\nabla\phi_h\|^2_{0,\Omega_h^\Gamma} + \frac{1}{h^4}\|q_h\phi_h\|^2_{0,\Omega_h^\Gamma}\right). \end{aligned}$$

This leads to (22) thanks to inverse inequalities and to the fact that both $\nabla\phi_h$ and $\frac{1}{h}\phi_h$ are uniformly bounded on Ω_h^Γ.

Step 4. We combine (19) and (22), and observe that the norm $\|v_h, z_h, q_h\|_0$ is equivalent to the 2-norm of the vector representing (v_h, z_h, q_h). This leads to the desired result as at the end of the proof of Theorem 4.1 from [9]. □

5 Numerical simulations

In this section, we illustrate ϕ-FEM on three different test cases, cf. Fig. 1, exploring the errors with respect to exact "manufactured" solutions. The numerical results for the 1st test case (in 2D) confirm the predicted theoretical estimates (in fact, better than theoretically predicted convergence rate is observed for the L^2 error). In the 2nd test case (also in 2D), we show that the optimal convergence is recovered even when the level-set function ϕ is less regular than assumed by the theory. Our method is also compared with CutFEM [8] in the last case. Finally, a 3D example is given in the 3rd test case.

In the first test case, we will treat some examples with non-homogeneous Neumann condition (Fig. 2-8) and Robin condition (Fig. 9) thanks to the modification of the scheme given in Remark 2. In the two last test cases, we will consider homogeneous Neumann conditions.

The surrounding domains $\mathcal{O}$ are always chosen as boxes aligned with the Cartesian coordinates and the background meshes $\mathcal{T}_h^{\mathcal{O}}$ are obtained from uniform Cartesian grids, dividing the cells into the simplexes (semi-cross meshes in 2D) . We always use the numerical quadrature of a high enough order so that all the integrals in (6) are computed exactly.

We have implemented ϕ-FEM both in `FreeFEM` [13] and in `multiphenics` [2]. Both implementations give the same results in our test cases and we present here only those obtained with `FreeFEM`. The implementation scripts can be consulted on GitHub.[3]

[3] `https://github.com/michelduprez/PhiFEM-Neumann`

Remark 3. *Our method (6) features "mixed" terms, such as* $\gamma_1 \int_{\Omega_h^\Gamma} (y_h + \nabla u_h) \cdot (z_h + \nabla v_h)$ *involving* u_h, v_h *defined on mesh* $\mathcal{T}_h$ *and* y_h, z_h *defined on mesh* $\mathcal{T}_h^\Gamma$*, a submesh of* $\mathcal{T}_h$*. Such integrals cannot be implemented in the current version of* `FEniCS` *since it requires all the finite elements involved in a problem to be defined on the same mesh. This is why we have turned to the* `multiphenics` *library, a spin-off of* `FEniCS`*, that does not have such a restriction. On the other hand,* `FreeFEM` *features interpolations between meshes in a user-friendly manner. However, we have discovered that a straightforward implementation of (6) in* `FreeFEM` *involving an implicit interpolation from* $\mathcal{T}_h$ *to* $\mathcal{T}_h^\Gamma$ *can lead to some spurious oscillations in the error curves. Much better results (reported below) are obtained if we introduce explicitly the interpolation matrix from* $V_h^{(k)}$ *to its restriction on* $\mathcal{T}_h^\Gamma$*, using the* `FreeFEM` *function* `interpolate`*, or, equivalently but more efficiently, the renumbering between the degrees of freedom of these two finite element spaces, , using the* `FreeFEM` *function* `restrict`*.*

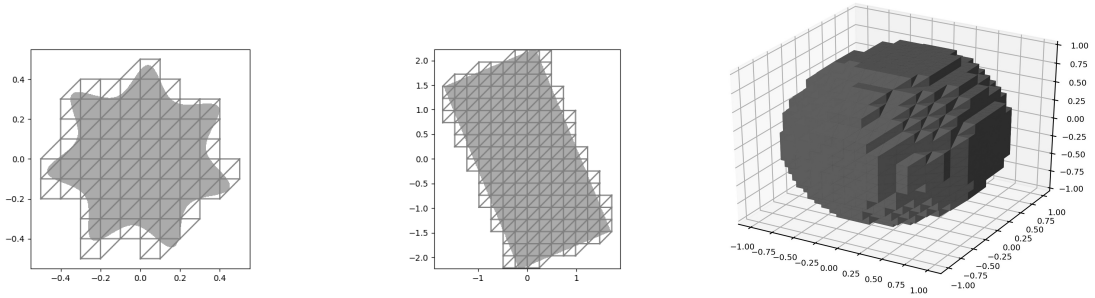

Figure 1: Domains and meshes considered in ϕ-FEM for the test case 1 (left), test case 2 (center) and test case 3 (right).

5.1 1st test case

Domain Ω (see Fig. 1 left) is defined by the level-set function ϕ given in the polar coordinates (r, θ) by

$$\phi(r, \theta) = r^4(5 + 3\sin(7(\theta - \theta_0) + 7\pi/36))/2 - R^4,$$

where $R = 0.47$ and $\theta_0 \in [0, 2\pi)$. The surrounding domain $\mathcal{O}$ is fixed to $(-0.5, 0.5)^2$. Varying the angle θ_0 results in a rotation of Ω, so that the boundary Γ cuts the triangles of the background mesh in a different manner, creating sometimes the "dangerous" situations when certain mesh triangles of $\mathcal{T}_h$ have only a tiny portion inside the physical domain Ω.

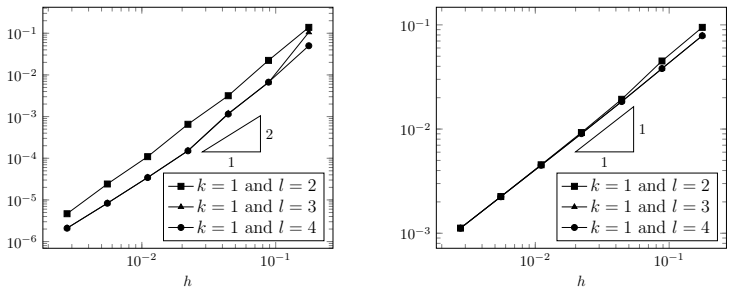

Figure 2: ϕ-FEM for the test case 1, $\theta_0 = 0$, $\sigma = 0.01$ and $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $k = 1$ and different values of l. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

We use ϕ-FEM to solve numerically Poisson-Neumann problem (1) with non-homogeneous boundary conditions $\frac{\partial u}{\partial n} = g$ adjusting f and g so that the exact solution is given by $u(x, y) = \sin(x)\exp(y)$. The Neumann boundary condition is extrapolated to a vicinity of Γ by $\tilde{g} = \frac{\nabla u \cdot \nabla \phi}{|\nabla \phi|} + u\phi$, cf. Remark 2. The addition of $u\phi$ here does not perturb $\tilde{g}$ on Γ. Its purpose is to mimick the real life situation where g is known on Γ only and $\tilde{g}$ is some extension of g, not necessarily the natural one $\nabla u \cdot \nabla \phi / |\nabla \phi|$.

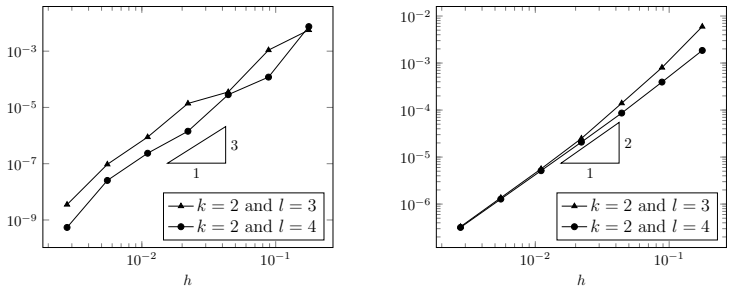

Figure 3: ϕ-FEM for the test case 1, $\theta_0 = 0$, $\sigma = 0.01$ $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $k = 2$ and different values of l. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

We report at Figs. 2 and 3 the evolution of the relative error under the mesh refinement for a fixed position of Ω ($\theta_0 = 0$), using finite element spaces $W_h^{(k)}$ with $k = 1$ ($\mathbb{P}_1$ FE for u_h) and $k = 2$ ($\mathbb{P}_2$ FE for u_h). We also try there different values of l, the degree of finite element used to approximate the level-set ϕ, recalling that it should be chosen as $k+1$ or greater. The experiments reported in these figures confirms the optimal convergence order of the method in both H^1 and L^2 norms (orders k and $k+1$ respectively). The convergence order in the L^2 norm is thus better than in theory. An interesting experimental observation comes from exploring the degree l: while the lowest possible value $l = k + 1$ ensures indeed the optimal convergence orders, it seems advantageous to increase the degree to $l = k + 2$, leading to more accurate results, especially in the L^2 norm. Another series of experiments is reported at Figs. 4 and 5. We explore there the errors with respect to the rotation of Ω over the background mesh (varying θ_0). We restrict ourselves here with finite elements degree $k = 1$ but compare two different values of l: $l = k + 1 = 2$ at Fig. 4 vs. $l = k + 2 = 3$ at Fig. 5. We observe again an advantage of the choice $l = k + 2$: the oscillations on any given background mesh become less important when increasing l and fade away under the mesh refinement in the case $l = k+2$ (this concerns mostly the L^2 errors; the H^1 errors are pretty much the same in both cases). The influence of the parameters σ, γ_{div}, γ_u, γ_p on the accuracy of the method is explored by the numerical experiments reported at Figs. 6 and 7. Although a full assessment of the role of all the 4 parameters is difficult (we have chosen, somewhat arbitrarily, two scenarios of parameter variations out of endless other possibilities), the conclusion of our numerical experiments seems clear: the method is not sensible to variation of the parameters in the wide range from 10^{-6} to 10, and there is no need to take these parameters greater than 10. Finally, we report at Fig. 8 evolution of the condition number of the ϕ-FEM matrix under the mesh refinement and also its sensitivity with respect to the rotations of Ω. The theoretically predicted behavior of $\sim 1/h^2$ is confirmed. The conditioning of the method is also found to be rather insensitive to the position of Ω over the mesh.

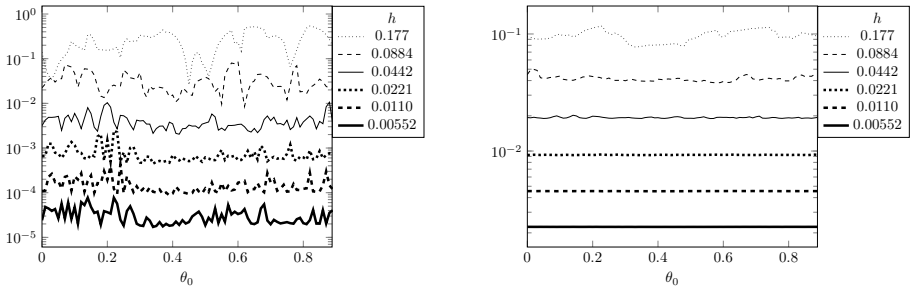

Figure 4: Sensitivity of the relative error with respect to θ_0 in ϕ-FEM for the test case 1, $\sigma = 0.01$ and $\gamma_u = \gamma_p = \gamma_{\text{div}} = 20$, $k = 1$ and $l = 2$. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

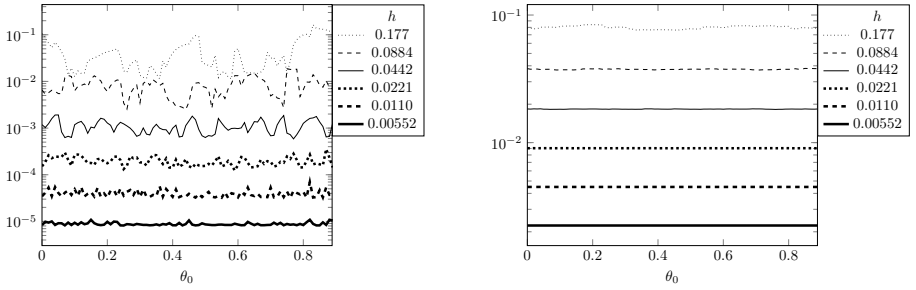

Figure 5: Sensitivity of the relative error with respect to θ_0 in ϕ-FEM for the test case 1, $\sigma = 0.01$ and $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $k = 1$ and $l = 3$. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

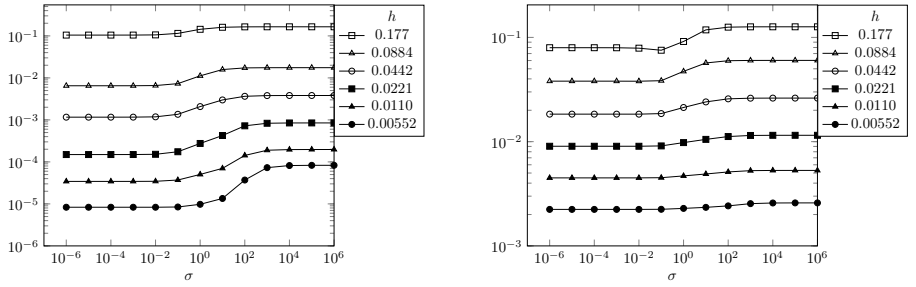

Figure 6: Sensitivity of the relative error in ϕ-FEM with respect to σ with $\gamma_u = \gamma_p = \gamma_{div} = 10$ being fixed for the test case 1, $\theta_0 = 0$, $k = 1$ and $l = 3$. Left: L^2 relative error; Right: H^1 relative error.

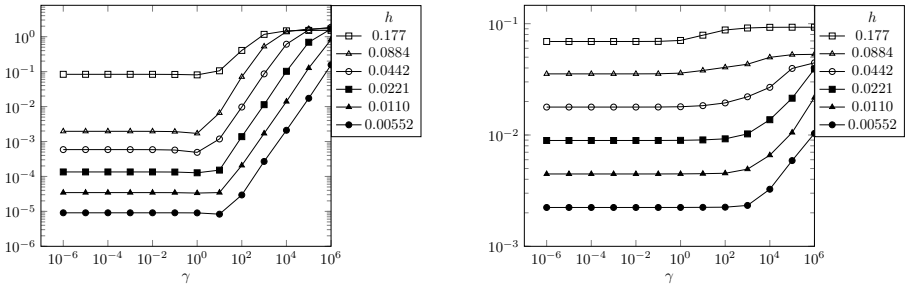

Figure 7: Sensitivity of the relative error in ϕ-FEM with respect to $\gamma_u = \gamma_p = \gamma_{div} = \gamma$ with $\sigma = 0.01$ fixed for the test case 1, $\theta_0 = 0$, $k = 1$ and $l = 3$. Left: L^2 relative error; Right: H^1 relative error.

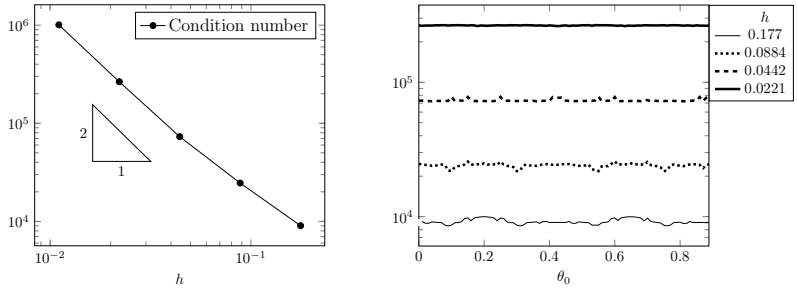

Figure 8: Condition number in ϕ-FEM for the test case 1, $\sigma = 0.01$, $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $\theta_0 = 0$, $k = 1$ and $l = k + 2$. Left: $\theta_0 = 0$; Right: different values of θ_0.

We end this section by given an example with Robin boundary condition with $\alpha = 1$ thanks to modification of main scheme presented in Remark 2. We consider the same domain Ω, level-set function ϕ and solution u as before. The Robin condition is extrapolated by $\tilde{g} = \frac{\nabla u \cdot \nabla \phi}{|\nabla \phi|} + \alpha u + u\phi$. In Fig. 9, we report the L^2 errors and the H^1 error (left) and the condition number (right) for $k = 1$ and $l = 3$. We observe that optimal convergence order and standard condition number remain valid for our Robin formulation.

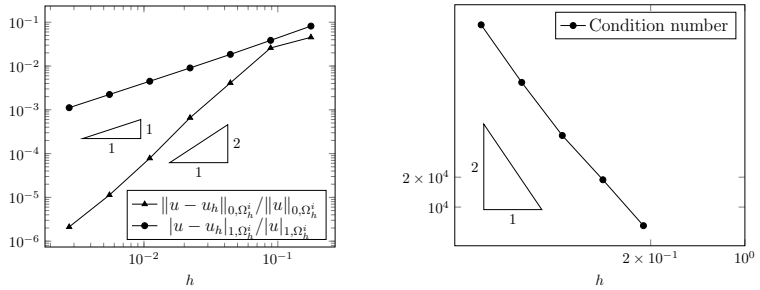

Figure 9: ϕ-FEM for the test case 1 and Robin boundary conditions, $\sigma = 0.01$, $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $\alpha = 1$, $k = 1$ and $l = 3$. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i}/\|u\|_{0,\Omega_h^i}$ and H^1 relative error $|u - u_h|_{1,\Omega_h^i}/|u|_{1,\Omega_h^i}$. ; Right: Condition number.

5.2 2nd test case

In this test case, the domain Ω is the rectangle $(-1,1)\times(-2,2)$ rotated by an angle θ_0 counter-clockwise around the origin. It is defined by the level-set function ϕ given by $\phi(x,y)=\Phi\circ\Pi(x,y)$, with $\Phi(x,y)=\max(|x|,|y|/2)-1$ and $\Pi\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\cos(\theta_0) & -\sin(\theta_0)\\ \sin(\theta_0) & \cos(\theta_0)\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$. The surrounding domain is taken as $\mathcal{O}=(-R,R)^2$, with $R=1.1\sqrt{5}$, cf. Fig. 1 middle.

We use ϕ-FEM to solve numerically Poisson-Neumann problem (1) with the exact solution given by $u(x,y)=U\circ\Pi(x,y)$, where $U(x,y)=\cos(\pi x)\cos(\pi y/2)$.

The results are presented at Figs. 10 (left) and 11, first choosing a fixed inclination angle $\theta_0=\pi/8$, and then varying θ_0 from 0 to $2\pi/7$. The numerical tests show again the optimal convergence of ϕ-FEM with $\mathbb{P}_1$ finite elements in the L^2 and H^1 norms, notwithstanding the fact that the level-set function ϕ is less regular than assumed in our theoretical results. Note that we have used here the FE of degree $l=3$ to represent the level-set, which is higher than the minimal degree $k+1=2$ suggested by the theory. The situation is here similar to that of the tests case 1-2: the implementation using the lower degree $l=2$ elements (not reported here) is also optimally convergent but turns out to be less robust than $l=3$ wih respect to the placement of Ω over the mesh (higher oscillations, especially in the L^2 error, when varying θ_0).

We have also compared our method with CutFEM [8]: Find $u_h\in V_h^{(k)}$ s.t.

$$\int_\Omega \nabla u_h\cdot\nabla v_h+\int_\Omega u_h v_h+\sigma h\sum_{E\in\mathcal{F}^\Gamma}\int_E\left[\frac{\partial u_h}{\partial n}\right]\left[\frac{\partial v_h}{\partial n}\right]=\int_\Omega f v_h+\int_\Gamma g v_h\ \ \forall\ v_h\in V_h^{(k)},$$

where $\mathcal{F}^\Gamma=\{E(\text{internal facet of }\mathcal{T}_h)\text{ such that }\exists T\in\mathcal{T}_h: T\cap\Gamma\neq\varnothing\text{ and }E\in\partial T\}$.

The results are reported at Figs. 10 (right, the simulation at fixed inclination angle θ_0) and 12 (simulations with the rotating domain Ω). Comparing two parts of Fig. 10, we conclude that ϕ-FEM and CutFEM are both optimally convergent and produce very similar results. However, looking closer at Figs. 11 and 12, we can point out an advantage of the ϕ-FEM over the CutFEM: the former seems more robust with respect to the position of Ω over the background mesh, the oscillations of the L^2 errors with rotating the domain are more pronounced for the latter method (the H^1 errors are almost the same in both cases).

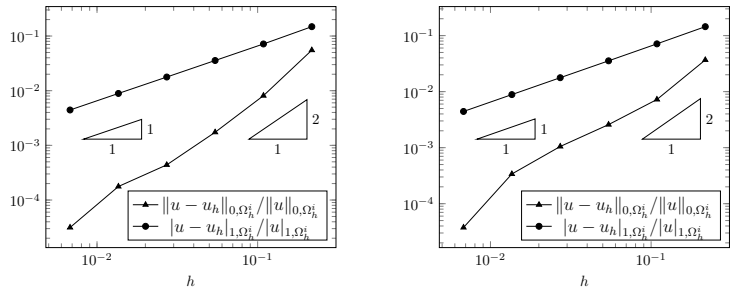

Figure 10: L^2 and H^1 relative error for the test case 2. Left: ϕ-FEM with $\sigma=0.01$, $\gamma_u=\gamma_p=\gamma_{\text{div}}=10$, $k=1$ and $l=3$; Right: CutFEM , $\theta_0=\pi/8$, $\sigma=0.01$.

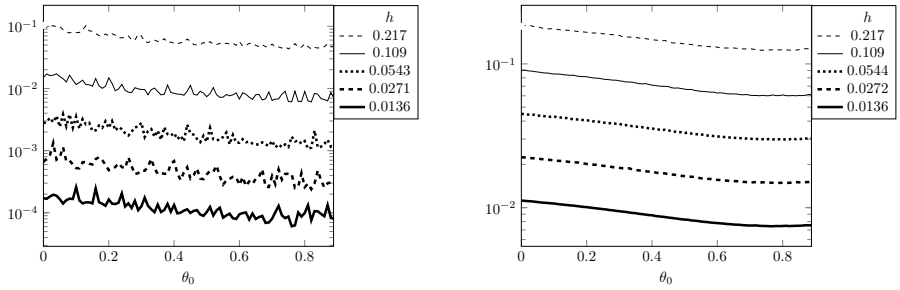

Figure 11: Sensitivity of the relative error with respect to θ_0 in ϕ-FEM for the test case 2, $\sigma = 0.01$, $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $k = 1$ and $l = 3$. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

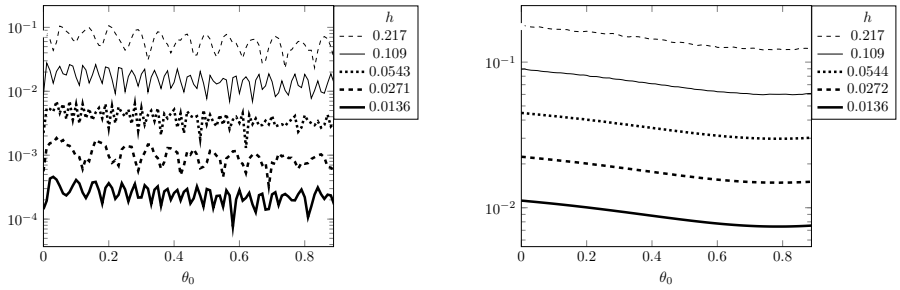

Figure 12: Sensitivity of the relative error with respect to θ_0 in CutFEM for the test case 2, $\sigma = 0.01$. Left: L^2 relative error $\|u - u_h\|_{0,\Omega_h^i} / \|u\|_{0,\Omega_h^i}$; Right: H^1 relative error $|u - u_h|_{1,\Omega_h^i} / |u|_{1,\Omega_h^i}$.

5.3 3rd test case

We here take $\Omega \subset \mathbb{R}^3$ as the ball of radius $R = 0.75$ centered at the origin encapsulated into the box $\mathcal{O} = (-1, 1)^3$. Ω is defined by the level-set function $\phi(x, y, z) = x^2 + y^2 + z^2 - R^2$. Fig. 1 (right) gives an example of mesh $\mathcal{T}_h$ for this test case. We choose the exact solution as $u(x, y, z) = \cos\left(\sqrt{x^2 + y^2 + z^2}\right)$. The normal derivative $\frac{\partial u}{\partial n}$ turns out to be constant on Γ in this case, equal to $-\sin(R)$. Accordingly, we choose to extrapolate the Neumann boundary condition to a vicinity of Γ by $\tilde{g} = -\sin(R)$, cf. Remark 2. Again, we observe in Fig. 13 (left) the optimal orders of convergence for the L^2 and H^1 errors. We have done the experiments on moderate meshes only, using a direct solver (MUMPS) for the linear systems. The condition numbers are reported at Fig. 13 (right). Estimating numerically the condition numbers is tricky even on these relatively coarse meshes in 3D since their direct evaluation, by eg. the standard function of NumPy used in our 2D tests, is too expensive on some of our meshes in 3D. We have computed the condition numbers reported here by tracking the largest and smallest singular values on iterations of the GMRES method without any preconditioner, as provided by PETSc [1]. The condition numbers obtained like this can be thus under-estimated. This can explain why they seem to behave more like $\sim 1/h$ rather than the theoretically expected $\sim 1/h^2$. Another possible explanation is that the meshes explored here are yet too coarse to see the asymptotic regime.

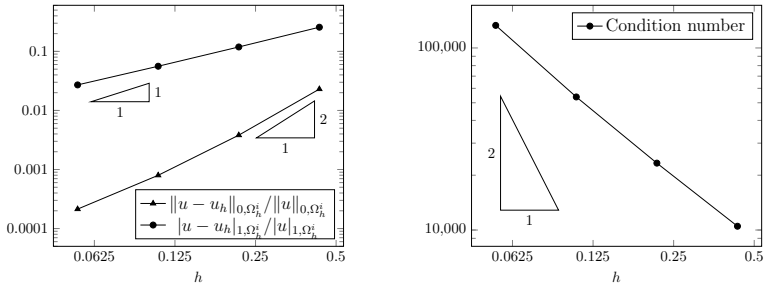

Figure 13: ϕ-FEM for the test case 3, $\sigma = 0.01$, $\gamma_u = \gamma_p = \gamma_{\text{div}} = 10$, $k = 1$ and $l = 3$. Left: L^2 and H^1 relative error; Right: condition number.

References

[1] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. D. Gropp, D. Karpeyev, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, R. T. Mills, T. Munson, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, H. Zhang, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.15, Argonne National Laboratory, 2021.

[2] F. Ballarin and G. Rozza. multiphenics. `https://mathlab.sissa.it/multiphenics`, 2020.

[3] S. Bertoluzza, M. Ismail, and B. Maury. Analysis of the fully discrete fat boundary method. *Numerische Mathematik*, 118(1):49–77, 2011.

[4] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[5] T. Boiveau, E. Burman, S. Claus, and M. Larson. Fictitious domain method with boundary value correction using penalty-free Nitsche method. *J. Numer. Math.*, 26(2):77–95, 2018.

[6] E. Burman. Ghost penalty. *Comptes Rendus Mathematique*, 348(21):1217–1220, 2010.

[7] E. Burman, S. Claus, P. Hansbo, M. Larson, and A. Massing. Cutfem: discretizing geometry and partial differential equations. *International Journal for Numerical Methods in Engineering*, 104(7):472–501, 2015.

[8] E. Burman and P. Hansbo. Fictitious domain finite element methods using cut elements: Ii. A stabilized Nitsche method. *Applied Numerical Mathematics*, 62(4):328–341, 2012.

[9] M. Duprez and A. Lozinski. ϕ-FEM: a finite element method on domains defined by level-sets. *SIAM J. Num. Anal., to appear; preprint on arXiv:1903.03703 [math.NA]*, 2020.

[10] D. Elfverson, M. G. Larson, and K. Larsson. A new least squares stabilized nitsche method for cut isogeometric analysis. *Computer Methods in Applied Mechanics and Engineering*, 349:1–16, 2019.

[11] R. Glowinski and T. Pan. Error estimates for fictitious domain/penalty/finite element methods. *Calcolo*, 29(1):125–141, 1992.

[12] J. Haslinger and Y. Renard. A new fictitious domain approach inspired by the extended finite element method. *SIAM Journal on Numerical Analysis*, 47(2):1474–1499, 2009.

[13] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[14] C. Lehrenfeld. High order unfitted finite element methods on level set domains using isoparametric mappings. *Comput. Methods Appl. Mech. Engrg.*, 300:716–733, 2016.

[15] A. Lozinski. CutFEM without cutting the mesh cells: a new way to impose Dirichlet and Neumann boundary conditions on unfitted meshes. *Comput. Methods Appl. Mech. Engrg.*, 356:75–100, 2019.

[16] B. Maury. A fat boundary method for the poisson problem in a domain with holes. *Journal of scientific computing*, 16(3):319–339, 2001.

[17] R. Mittal and G. Iaccarino. Immersed boundary methods. *Annu. Rev. Fluid Mech.*, 37:239–261, 2005.

[18] N. Moës, E. Béchet, and M. Tourbier. Imposing dirichlet boundary conditions in the extended finite element method. *International Journal for Numerical Methods in Engineering*, 67(12):1641–1669, 2006.

[19] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International journal for numerical methods in engineering*, 46(1):131–150, 1999.

[20] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2003.

[21] N. Sukumar, D. Chopp, N. Moës, and T. Belytschko. Modeling holes and inclusions by level sets in the extended finite-element method. *Computer methods in applied mechanics and engineering*, 190(46-47):6183–6200, 2001.