MOX–Report No. 10/2009

Numerical approximation of incompressible flows with net flux defective boundary conditions by means of penalty techniques

Paolo Zunino

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Numerical approximation of incompressible flows with net flux defective boundary conditions by means of penalty techniques

Paolo Zunino[♯]

March 20, 2009

[♯] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
paolo.zunino@polimi.it

Keywords: incompressible fluid dynamics, artificial boundary conditions, finite element approximation, penalty methods, computational hemodynamics.

Abstract

We consider incompressible flow problems with defective boundary conditions prescribing only the net flux on some inflow and outflow sections of the boundary. As a pradigm for such problems, we simply refer to Stokes flow. After a brief review of the problem and of its well posedness, we discretize the corresponding variational formulation by means of finite elements and looking at the boundary conditions as constraints, we exploit a penalty method to account for them. We perform the analysis of the method in terms of consistency, boundedness and stability of the discrete bilinear form and we show that the application of the penalty method does not affect the optimal convergence properties of the finite element discretization. Since the additional terms introduced to account for the defective boundary conditions are non local, we also analyze the spectral properties of the equivalent algebraic formulation and we exploit them to set up an efficient solution strategy. In contrast to alternative discretization methods based for instance on Lagrange multipliers accounting for the constraints on the boundary, the present scheme is particularly effective because it only mildly affects the computational cost of the numerical approximation. Indeed, it does not require neither multipliers nor sub-iterations or additional adjoint problems with respect to the reference problem at hand.

1 Introduction

Many fluid dynamics problems arise in unbounded domains, but the application of numerical approximation techniques often require to restrict to a bounded region. For this reason, artificial boundaries and artificial boundary conditions must be introduced, as illustrated in the seminal work by Heywood, Rannacher and Turek [1]. This difficulty arises for a wide spectrum of engineering applications, in particular we mention computational hemodynamics [2, 3, 4, 5, 6]. In this context, one of the effective techniques to provide artificial boundary conditions consists to prescribe the net flux on the inflow and outflow sections of the truncated domain. Such conditions are *defective* because the full velocity profile may not be available on these sections. As a result of that, to obtain a well posed problem, additional conditions on the stresses are mandatory. In alternative, see [1, 2], one could prescribe the mean pressure over the artificial sections. Although the variational formulation of the defective net flux problem can be easily casted into a standard framework [1], such non standard boundary conditions require a special numerical treatment. Since this difficulty does not arise for the mean pressure problem, see [2], we restrict here to the net flux conditions.

Following the seminal work of Babuska for the alternative treatment of essential conditions for second order elliptic problems, see [7], an effective possibility proposed and analyzed in [2] and later in [3, 4] consists in accounting for the constraints at the artificial boundaries by means of Lagrange multipliers. It has been proved that this technique is very flexible for the purposes of computational hemodynamics, but it considerably increases the computational cost of the numerical approximation, see in particular [3]. More recently, a new strategy based on the satisfaction of the boundary constraints by means of a control problem has been proposed in [6]. This approach seems to be extremely general and flexible, but it involves the iterative solution of the reference problem with an adjoint problem. As a consequence of that, the overall cost of the computation might increase remarkably.

Getting inspiration from the case of elliptic equations with essential boundary conditions, we address the application of the classical method by Nitsche, i.e. the *penalty method*, see [8, 9, 10], to Stokes problem with net flux defective boundary conditions. We will show that this technique, applied in the framework of the finite element method, leads to a well posed discrete problem that is consistent, stable and convergent with the optimal convergence properties of the selected finite elements. We will also analyze the spectral properties of the corresponding algebraic problem, propose a convenient solution strategy and finally show that at the computational level the penalty method turns out to be more effective than the application of the Lagrange multipliers, because the computational cost of the former is almost equivalent to the solution of a Stokes problem with full essential (Dirichlet) boundary data.

Although not addressed here, the present technique can be straightforwardly

applied to Oseen or Navier-Stokes equations in the steady or time dependent cases. From a wider perspective, it could also be useful to set up multiscale methods, designed for instance to account of the arterial peripheral resistances by means of lumped or one-dimensional reduced models, see [2, 5].

2 Problem setting

Let $\Omega \subset \mathbb{R}^d$ be a possibly convex polygon/polyhedron with $d = 2, 3$. We aim to prescribe only the average of the velocity field $\mathbf{u}$ on a finite number of subsets of the boundary that is denoted by $\Gamma_k \subseteq \partial\Omega$ with measure $|\Gamma_k|$, where $k = 1, \ldots, N$ and $\Gamma = \cup_{k=1}^{N} \Gamma_k$. For simplicity, we address the case of homogeneous Dirichlet boundary conditions on $\partial\Omega \setminus \Gamma$, but the present treatment can be easily generalized to include Neumann conditions. To avoid confusion, we denote with $\mathcal{L}^2(\Omega) := [L^2(\Omega)]^d$ and $\mathcal{H}^1(\Omega) := [H^1(\Omega)]^d$ the usual Sobolev spaces for vector valued functions.

Following [1], the Stokes problem with *vector defective* boundary conditions is defined as follows: given $\mathbf{f} \in \mathcal{L}^2(\Omega)$ and N *vectors* $\mathbf{U}_k \in \mathbb{R}^d$ such that $\sum_{k=1}^{N} (\mathbf{U}_k \cdot \mathbf{n})|\Gamma_k| = 0$ to satisfy mass balance, we aim to find a couple of functions $(\mathbf{u}, p)$ and N *vectors* $\mathbf{c}_k \in \mathbb{R}^d$ such that

$$\begin{cases} -\nabla^2 \mathbf{u} + \nabla p = \mathbf{f}, \ \nabla \cdot \mathbf{u} = 0, & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0}, & \text{on } \partial\Omega \setminus \Gamma, \\ \frac{1}{|\Gamma_k|} \int_{\Gamma_k} \mathbf{u} = \mathbf{U}_k, \ p\mathbf{n} - \nabla \mathbf{u} \cdot \mathbf{n} = \mathbf{c}_k & \text{on } \Gamma_k, \ k = 1, \ldots, N, \end{cases} \tag{1}$$

where $(\cdot)$ denotes both the standard scalar product and the matrix-vector multiplication, while $\mathbf{n}$ is the outer unit normal vector with respect to $\partial\Omega$. The more usual formulation with *scalar* defective boundary conditions will be briefly discussed later on. In the case of non homogeneous boundary data on Γ_k, we introduce suitable lifting vector functions $\mathbf{w}_k \in \mathcal{H}^1(\Omega)$ with $\nabla \cdot \mathbf{w}_k = 0$ such that $\frac{1}{|\Gamma_k|} \int_{\Gamma_k} \mathbf{w}_k = \mathbf{1}$, $\int_{\Gamma_k} \mathbf{w}_j = \mathbf{0}$ for $k \neq j$ and $\mathbf{w}_k = \mathbf{0}$ on $\partial\Omega \setminus \Gamma$, being $\mathbf{1} \in \mathbb{R}^d$ the vector with unit components and $\mathbf{0}$ the null vector. To address the vatiational formulation of (1) we apply divergence-free spaces. More precisely, we define

$$\begin{aligned} \mathcal{V} &:= \{\mathbf{v} \in \mathcal{H}^1(\Omega) : \ \mathbf{v}|_{\partial\Omega\setminus\Gamma} = 0\}, \\ \mathcal{V}^{\mathbf{0}} &:= \{\mathbf{v} \in \mathcal{V} : \ \int_{\Gamma_k} \mathbf{v} = \mathbf{0}, \ \forall k\}, \\ \mathcal{V}^{\mathbf{0}}_{div} &:= \{\mathbf{v} \in \mathcal{V}^{\mathbf{0}} : \ \nabla \cdot \mathbf{v} = 0\}. \end{aligned}$$

Let us multiply (1) by $\mathbf{v} \in \mathcal{V}^{\mathbf{0}}_{div}$ and integrate over Ω. By means of Green's formula we obtain,

$$\begin{aligned} &\int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v} + \textstyle\sum_{k=1}^{N} \displaystyle\int_{\Gamma_k} (p\mathbf{n} - \nabla \mathbf{u} \cdot \mathbf{n}) \cdot \mathbf{v} = \int_\Omega \mathbf{f} \cdot \mathbf{v}, \ \forall \mathbf{v} \in \mathcal{V}^{\mathbf{0}}_{div}, \\ &\int_{\Gamma_k} (p\mathbf{n} - \nabla \mathbf{u} \cdot \mathbf{n}) \cdot \mathbf{v} = \int_{\Gamma_k} \mathbf{c}_k \cdot \mathbf{v} = \mathbf{c}_k \cdot \int_{\Gamma_k} \mathbf{v} = 0, \ \forall \mathbf{v} \in \mathcal{V}^{\mathbf{0}}_{div}. \end{aligned}$$

Then, applying the change of variable $\mathbf{u^0} = \mathbf{u} - \mathbf{w}$, being $\mathbf{w}^{(i)} = \sum_{k=1}^{N} \mathbf{U}_k^{(i)} \mathbf{w}_k^{(i)}$ with $i = 1, \ldots, d$, the variational formulation of problem (1) requires to find $\mathbf{u^0} \in \mathcal{V}_{div}^{\mathbf{0}}$ such that

$$a(\mathbf{u^0}, \mathbf{v}) := \int_\Omega \nabla \mathbf{u^0} : \nabla \mathbf{v} = F(\mathbf{v}) := \int_\Omega \mathbf{f} \cdot \mathbf{v} - \int_\Omega \nabla \mathbf{w} : \nabla \mathbf{v}, \ \forall \mathbf{v} \in \mathcal{V}_{div}^{\mathbf{0}}. \quad (2)$$

2.1 Existence and uniqueness of a solution

Let $\|\mathbf{v}\|_{0,\Sigma}$, $\|\mathbf{v}\|_{1,\Sigma}$ and $|\mathbf{v}|_{1,\Sigma}$ be the $\mathcal{L}^2$ norm, $\mathcal{H}^1$ norm and seminorm on Σ, respectively, where Σ is a d or $(d-1)$-dimensional subset of Ω. We also denote with $(\mathbf{u}, \mathbf{v})_\Sigma$ the $\mathcal{L}^2$ inner product over Σ. Furthermore, for the ease of notation we introduce a symmetric bilinear operator and its corresponding seminorm,

$$\langle \mathbf{u}, \mathbf{v} \rangle_\Sigma := \frac{1}{|\Sigma|} \Big(\int_\Sigma \mathbf{u} \Big) \cdot \Big(\int_\Sigma \mathbf{v} \Big), \quad \langle \mathbf{v} \rangle_\Sigma^2 := \langle \mathbf{v}, \mathbf{v} \rangle_\Sigma \, .$$

For the analysis of problem (2) we proceed by means of the Lax-Milgram Lemma. In this context, we will make use of the following Poincaré-Friedrichs inequality that we address here for vector valued functions, for notational convenience.

Property 2.1 *For all* $\mathbf{v} \in \mathcal{H}^1(\Omega)$, $\Gamma \subseteq \partial\Omega$ *with* $|\Gamma| > 0$ *there exists a positive constant* C_{PF} *such that*

$$\|\mathbf{v}\|_{0,\Omega}^2 \leq C_{PF} \big(|\mathbf{v}|_{1,\Omega}^2 + \langle \mathbf{v} \rangle_\Gamma^2 \big). \quad (3)$$

Owing to (3), we prove that (2) is well posed independently of the conditions enforced on the complementary part of Γ. We will first prove the existence of a velocity field in $\mathbf{u} \in \mathcal{V}_{div}^{\mathbf{0}}$. The existence of a corresponding pressure $p \in Q := \mathcal{L}_0^2(\Omega) := \{q \in \mathcal{L}^2(\Omega) : \int_\Omega q = 0\}$ follows from standard arguments, for which we refer to [13].

Corollary 2.2 *Problem* (2) *admits a unique solution.*

Proof We verify that $a(\cdot, \cdot)$ is coercive in $\mathcal{V}_{div}^{\mathbf{0}}$, endowed with the $\mathcal{H}^1$ norm,

$$a(\mathbf{v}, \mathbf{v}) = |\mathbf{v}|_{1,\Omega}^2 + \sum_{k=1}^{N} \langle \mathbf{v} \rangle_{\Gamma_k}^2 \geq \tfrac{1}{2} \min \Big[1, \tfrac{1}{C_{PF}} \Big] \|\mathbf{v}\|_{1,\Omega}^2, \ \forall \mathbf{v} \in \mathcal{V}_{div}^{\mathbf{0}}.$$

Owing to the Cauchy-Schwarz inequality, it is straightforward to show that $a(\cdot, \cdot)$ is continuous and that $F(\cdot)$ is a linear and bounded functional in $\mathcal{H}^1(\Omega)$. The result follows from the application of the Lax-Milgram Lemma in $\mathcal{V}_{div}^{\mathbf{0}}$. □

2.2 Additional properties

We notice that problem (2) can be equivalently reformulated for velocity-pressure couples. Let $\mathcal{W} := \mathcal{V} \times Q$ and $\mathcal{W}^{\mathbf{0}} := \mathcal{V}^{\mathbf{0}} \times Q$ be the natural spaces for the weak solution of (1), where $Q := \mathcal{L}^2_0(\Omega)$ denotes the subspace of functions $q \in \mathcal{L}^2(\Omega)$ with zero mean value over Ω. Let us introduce $c^{\pm}\big((\mathbf{u}, p), (\mathbf{v}, q) \big)$ such that

$$c^{\pm}\big((\mathbf{u}, p), (\mathbf{v}, q) \big) := a(\mathbf{u}, \mathbf{v}) + b(p, \mathbf{v}) \pm b(q, \mathbf{u}),$$
$$b(p, \mathbf{v}) := -\left(p, \nabla \cdot \mathbf{v}\right)_{\Omega}, \ \forall (\mathbf{u}, p), (\mathbf{v}, q) \in \mathcal{W}.$$

Then, setting $K(\mathbf{v}, q) := F(\mathbf{v})$ for any $(\mathbf{v}, q) \in \mathcal{W}^{\mathbf{0}}$, problem (2) is equivalent to find $(\mathbf{u^0}, p) \in \mathcal{W}^{\mathbf{0}}$ such that,

$$c^{\pm}\big(\left(\mathbf{u^0}, p\right), (\mathbf{v}, q) \big) = K(\mathbf{v}, q), \ \forall (\mathbf{v}, q) \in \mathcal{W}^{\mathbf{0}}, \tag{4}$$

where the sign $(+)$ or $(-)$ in $c^{\pm}(\cdot, \cdot)$ is indifferent with respect to the well posedness of (4), but it determines the following properties that will be useful in the sequel.

Lemma 2.3 *The bilinear form $c^{+}(\cdot, \cdot)$ is symmetric and $c^{-}(\cdot, \cdot)$ is semi-definite in $\mathcal{W}$.*

Proof Since $a(\cdot, \cdot)$, defined in (2) is symmetric, it is straightforward to verify that $c^{+}\big((\mathbf{u}, p), (\mathbf{v}, q) \big) = c^{+}\big((\mathbf{v}, q), (\mathbf{u}, p) \big)$. Finally, owing to Corollary 2.2, we obtain $c^{-}\big((\mathbf{v}, q), (\mathbf{v}, q) \big) \geq C \|\mathbf{v}\|^2_{1,\Omega}$, i.e. $c^{-}(\cdot, \cdot)$ is semi-definite in $\mathcal{W}$. □

Lemma 2.3 confirms that problem (1) is self-adjoint. To conclude, we assume that problem (4) is *regularizing*.

Hypothesis 2.4 *Given $\mathbf{g} \in \mathcal{L}^2(\Omega)$, let $(\mathbf{u}^{\mathbf{0}}_g, p_g)$ be the solution of (4) with $K(\mathbf{v}, q) := (\mathbf{g}, \mathbf{v})_{\Omega}$. Then $(\mathbf{u}^{\mathbf{0}}_g, p_g) \in \big(\mathcal{H}^2(\Omega) \cap \mathcal{V}^{\mathbf{0}}\big) \times \big(H^1(\Omega) \cap Q\big)$ and the following a-priori estimate holds,*

$$|||(\mathbf{u}^{\mathbf{0}}_g, p_g)|||_* \lesssim \|\mathbf{g}\|_{0,\Omega}, \ \forall \mathbf{g} \in \mathcal{L}^2(\Omega), \quad |||(\mathbf{u}^{\mathbf{0}}_g, p_g)|||^2_* := \|\mathbf{u}^{\mathbf{0}}_g\|^2_{2,\Omega} + \|p_g\|^2_{1,\Omega}. \tag{5}$$

Given $\mathbf{g} \in \mathcal{L}^2(\Omega)$, let $(\mathbf{z}^{\mathbf{0}}_g, r_g) \in \mathcal{W}^{\mathbf{0}}$ be the weak solution of the adjoint problem with respect to (4), that is the solution of

$$c^{+}\big((\mathbf{v}, q), \left(\mathbf{z}^{\mathbf{0}}_g, r_g\right) \big) = (\mathbf{g}, \mathbf{v})_{\Omega}, \ \forall (\mathbf{v}, q) \in \mathcal{W}^{\mathbf{0}}. \tag{6}$$

The following result will be useful in the sequel.

Lemma 2.5 *If assumption 2.4 holds true, then problem (6) is also regularizing.*

Proof The result is straightforward since problem (4) with $c^{+}(\cdot, \cdot)$ is self-adjoint and regularizing. □

Remark We notice that there exists a similar (but not equivalent) formulation of problem (1). It consists to prescribe only the flow rates on Γ_k, instead of the mean value of the velocity vector. By consequence, it leads to *scalar defective* boundary conditions. The corresponding Stokes problem is defined as follows: given $\mathbf{f} \in \mathcal{L}^2(\Omega)$ and N *constants* U_k such that $\sum_{k=1}^{N} U_k |\Gamma_k| = 0$, we aim to find a function $\mathbf{u}$ and N *constants* c_k,

$$\begin{cases} -\nabla^2 \mathbf{u} + \nabla p = \mathbf{f},\ \nabla \cdot \mathbf{u} = 0, & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0}, & \text{on } \partial\Omega \setminus \Gamma, \\ \frac{1}{|\Gamma_k|} \int_{\Gamma_k} \mathbf{u} \cdot \mathbf{n} = U_k,\ p\mathbf{n} - \nabla \mathbf{u} \cdot \mathbf{n} = c_k \mathbf{n} & \text{on } \Gamma_k,\ k = 1, \ldots, N. \end{cases} \tag{7}$$

Proceeding as in the case of vector defective conditions, the variational formulation of problem (7) requires to define $\widehat{\mathcal{V}}^0_{div} := \{\mathbf{v} \in \mathcal{V} \ :\ \nabla \cdot \mathbf{v} = 0,\ \int_{\Gamma_k} \mathbf{v} \cdot \mathbf{n} = 0,\ \forall k\}$ and to find $\widehat{\mathbf{u}}^0 \in \widehat{\mathcal{V}}^0_{div}$ such that

$$a(\widehat{\mathbf{u}}^0, \mathbf{v}) = F(\mathbf{v}),\ \forall \mathbf{v} \in \widehat{\mathcal{V}}^0_{div}. \tag{8}$$

However, when $\partial\Omega \setminus \Gamma = \emptyset$, problem (8) might be ill-posed. Indeed, we observe that (3) would not be true if only the normal component of $\mathbf{v}$ were accounted on Γ. This is easily seen by considering a constant field $\mathbf{v}_x = 0,\ \mathbf{v}_y = 1$ for $(x, y) \in (0, 1) \times (0, 1)$ with $\Gamma_1 := \{x = 0\} \times (0, 1)$, $\Gamma_2 := \{x = 1\} \times (0, 1)$ such that

$$|\mathbf{v}|^2_{1,\Omega} + \langle \mathbf{v} \cdot \mathbf{n} \rangle^2_{\Gamma} = 0,\ \|\mathbf{v}\|^2_{0,\Omega} > 0.$$

This shows that the scalar defective conditions are not sufficient to ensure the positivity of problem (8). To this purpose, it is mandatory to set up a Dirichlet condition on $\partial\Omega \setminus \Gamma$, such that the positivity of the bilinear form follows from the application of the standard Poincaré inequality. Owing to these observations, we proceed to investigate the case of vector defective conditions, keeping in mind that the forthcoming numerical discretization scheme could be straightforwardly applied to problem (8), provided it is well posed. □

Remark As previously mentioned, an alternative formulation of problem (1) arises from the interpretation of the boundary conditions as constraints that are accounted by means of Lagrange multipliers. Precisely, we aim to find $\mathbf{u} \in \mathcal{V}$, $p \in Q$ and N vectors $\boldsymbol{\lambda}_k \in \mathbb{R}^d$ such that,

$$\begin{cases} a(\mathbf{u}, \mathbf{v}) + b(p, \mathbf{v}) + \sum_{k=1}^{N} \langle \boldsymbol{\lambda}_k, \mathbf{v} \rangle_{\Gamma_k} = (\mathbf{f}, \mathbf{v})_{\Omega}, & \forall \mathbf{v} \in \mathcal{V}, \\ b(q, \mathbf{u}) = 0, & \forall q \in Q, \\ \langle \boldsymbol{\mu}_k, \mathbf{u} \rangle_{\Gamma_k} = \langle \boldsymbol{\mu}_k, \mathbf{U}_k \rangle_{\Gamma_k}, & \forall \boldsymbol{\mu}_k \in \mathbb{R}^d,\ k = 1, \ldots, N. \end{cases} \tag{9}$$

We refer to [2, 3] for the analysis and the numerical approximation of problem (9). This alternative problem is often addressed as *mixed* formulation for the defective boundary conditions, or also *augmented* formulation, because the

presence of the multipliers $\boldsymbol{\lambda}_k$ increases the number of unknowns of the original problem. As highlighted in [3, 4], this is an expensive but yet effective formulation to approximate problem (2) by means of a standard Galerkin discretization method. Later on, we will compare the discretization of (9) obtained by means of finite elements, with the forthcoming alternative numerical scheme. □

3 Numerical approximation

Owing to the constraints at the boundary, the definition of a finite element subspace of $\mathcal{V}_{div}^{\mathbf{0}}$ is a non trivial task. Indeed, for the numerical approximation of (2) we aim to apply standard finite elements. To this purpose, we consider a family of conforming triangulations T_h of affine simplexes K in Ω. Let T_h be shape regular and quasi-uniform and let h be the mesh characteristic parameter with the assumption $h \ll 1$. We notice that we do not strictly need T_h to be quasi-uniform, but this assumption reduces the technical aspects of the analysis without affecting the generality of the approach.

Our discrete approximation spaces for the velocity and pressure respectively are given by

$$\mathcal{V}_h^r := \{\mathbf{v}_h \in \mathcal{V} : \ \mathbf{v}_h|_K \in \mathbb{P}^r(K), \ \forall K \in T_h\},$$
$$Q_h^s := \{q_h \in Q : \ q_h|_K \in \mathbb{P}^s(K), \ \forall K \in T_h\}, \ r, s \in \mathbb{N}.$$

By exploiting standard penalization techniques for Dirichlet boundary conditions we could also remove the constraint $\mathbf{v}_h = \mathbf{0}$ on $\partial\Omega \setminus \Gamma$. This alternative formulation will be applied in the forthcoming numerical experiments. However, the corresponding changes to the discrete scheme do not affect what will be presented here.

3.1 The penalty method for defective boundary conditions

To start the derivation of the penalty method we multiply problem (1) by test functions $\mathbf{v} \in \mathcal{V}$ and apply Green's formula,

$$\int_\Omega \big(\nabla\mathbf{u} : \nabla\mathbf{v} - p\nabla\cdot\mathbf{v}\big) + \sum_{k=1}^{N} \int_{\Gamma_k} (p\mathbf{n} - \nabla\mathbf{u}\cdot\mathbf{n})\cdot\mathbf{v} = \int_\Omega \mathbf{f}\cdot\mathbf{v}, \ \forall\mathbf{v} \in \mathcal{V},$$
$$\int_{\Gamma_k} (p\mathbf{n} - \nabla\mathbf{u}\cdot\mathbf{n})\cdot\mathbf{v} = \int_{\Gamma_k} \mathbf{c}_k\cdot\mathbf{v} = \frac{1}{|\Gamma_k|}\Big(\int_{\Gamma_k} (p\mathbf{n} - \nabla\mathbf{u}\cdot\mathbf{n})\Big)\cdot\Big(\int_{\Gamma_k}\mathbf{v}\Big)$$
$$= \langle p\mathbf{n} - \nabla\mathbf{u}\cdot\mathbf{n}, \mathbf{v}\rangle_{\Gamma_k}, \ \forall\mathbf{v} \in \mathcal{V}.$$

Reminding that $\frac{1}{|\Gamma_k|}\int_{\Gamma_k}\mathbf{u}=\mathbf{U}_k$, that is $\langle\mathbf{u}-\mathbf{U}_k,\mathbf{v}\rangle_{\Gamma_k}=0$ for all $\mathbf{v}\in\mathcal{V}$, we conclude that the weak solution $\mathbf{u}$ of (1) also satisfies

$$\begin{aligned}&\int_\Omega\left(\nabla\mathbf{u}:\nabla\mathbf{v}-p\nabla\cdot\mathbf{v}\right)\\&\qquad+\sum_{k=1}^N\left[\gamma h^{-1}\langle\mathbf{u},\mathbf{v}\rangle_{\Gamma_k}+\langle p\mathbf{n}-\nabla\mathbf{u}\cdot\mathbf{n},\mathbf{v}\rangle_{\Gamma_k}+\langle q\mathbf{n}-\nabla\mathbf{v}\cdot\mathbf{n},\mathbf{u}\rangle_{\Gamma_k}\right]\\&=\int_\Omega\mathbf{f}\cdot\mathbf{v}+\sum_{k=1}^N\left[\gamma h^{-1}\langle\mathbf{U}_k,\mathbf{v}\rangle_{\Gamma_k}+\langle\mathbf{U}_k,q\mathbf{n}-\nabla\mathbf{v}\cdot\mathbf{n}\rangle_{\Gamma_k}\right],\ \forall\mathbf{v}\in\mathcal{V},q\in Q.\end{aligned}\tag{10}$$

The second term on the first row is a penalty term and γh^{-1} is a penalty parameter, that is suitably scaled with respect to h in order to ensure optimal approximation properties of the finite element method. The third term is responsible for the consistency with respect to (1), while the fourth term has been introduced artificially to maintain the symmetry of the problem. Aiming to approximate (10), we define

$$\begin{aligned}a_h(\mathbf{u}_h,\mathbf{v}_h):=&\,(\nabla\mathbf{u}_h,\nabla\mathbf{v}_h)_\Omega\\&+\textstyle\sum_{k=1}^N\left[\gamma h^{-1}\langle\mathbf{u}_h,\mathbf{v}_h\rangle_{\Gamma_k}-\langle\nabla\mathbf{u}_h\cdot\mathbf{n},\mathbf{v}_h\rangle_{\Gamma_k}-\langle\nabla\mathbf{v}_h\cdot\mathbf{n},\mathbf{u}_h\rangle_{\Gamma_k}\right],\\b_h(p_h,\mathbf{v}_h):=&-(p,\nabla\cdot\mathbf{v}_h)_\Omega+\textstyle\sum_{k=1}^N\langle p,\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k},\\F_h(\mathbf{v}_h):=&\,(\mathbf{f},\mathbf{v})_\Omega+\textstyle\sum_{k=1}^N\left[\gamma h^{-1}\langle\mathbf{U}_k,\mathbf{v}_h\rangle_{\Gamma_k}-\langle\mathbf{U}_k,\nabla\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\right],\\G_h(q_h):=&\textstyle\sum_{k=1}^N\langle\mathbf{U}_k\cdot\mathbf{n},q_h\rangle_{\Gamma_k}.\end{aligned}$$

To sum up, in order to approximate Stokes problem with defective boundary conditions we aim to find $\mathbf{u}_h\in\mathcal{V}_h^r$ and $p_h\in Q_h^s$ such that

$$\begin{cases}a_h(\mathbf{u}_h,\mathbf{v}_h)+b_h(p_h,\mathbf{v}_h)=F_h(\mathbf{v}_h), & \forall\mathbf{v}_h\in\mathcal{V}_h^r,\\ b_h(q_h,\mathbf{u}_h)=G_h(q_h), & \forall q_h\in Q_h^s.\end{cases}\tag{11}$$

Remark We notice that equation (10) is not the only possibility to set up a penalty method for net flux boundary conditions. In alternative, following [11] we might exploit the skew-symmetric formulation, where the terms $\langle q\mathbf{n}-\nabla\mathbf{v}\cdot\mathbf{n},\mathbf{u}-\mathbf{U}_k\rangle_{\Gamma_k}$ are subtracted rather than added to the bilinear form. In this case, we could basically perform the same analysis, with weaker requirements on the penalty terms. However, we would loose the symmetry of the problem, that is exploited to retrieve an a priori error estimate in the L^2 norm, but is even more important to define an efficient the numerical solution strategy for the linear system of equations corresponding to the discrete scheme. □

3.2 A priori error estimates

We now aim to establish suitable error estimates for problem (11), showing that the penalty approximation of the defective conditions does not affect the optimal convergence properties of the standard finite element method. A fundamental tool for the forthcoming analysis is given by the Jensen's inequality.

Property 3.1 *For any* $\mathbf{v} \in \mathcal{L}^2(\Sigma)$ *we have*

$$\Big(\int_\Sigma \mathbf{v} \Big)^2 \leq |\Sigma| \int_\Sigma \mathbf{v}^2 \ \textit{or equivalently} \ \langle \mathbf{v} \rangle_\Sigma^2 \leq \|\mathbf{v}\|_{0,\Sigma}^2, \ \forall \mathbf{v} \in \mathcal{L}^2(\Sigma). \tag{12}$$

The combination of (3) and (12) allows us to define a suitable norm for the analysis of the discrete problem.

Lemma 3.2 *For any fixed, quasi-uniform mesh* T_h*, the application* $\mathcal{H}^1(\Omega) \to \mathbb{R}$ *defined as,*

$$\|\mathbf{v}\|_{1,h,\Omega}^2 := |\mathbf{v}|_{1,\Omega}^2 + h^{-1} \textstyle\sum_{k=1}^N \langle \mathbf{v} \rangle_{\Gamma_k}^2 ,$$

is a mesh dependent norm on $\mathcal{H}^1(\Omega)$ *equivalent to* $\| \cdot \|_{1,\Omega}$*. More precisely, it satisfies*

$$\|\mathbf{v}\|_{1,\Omega} \lesssim \|\mathbf{v}\|_{1,h,\Omega} \lesssim h^{-\frac{1}{2}} \|\mathbf{v}\|_{1,\Omega}.$$

Here and in the sequel, the symbol $\lesssim$ denotes an inequality involving a positive constant C independent of the characteristic size of the mesh elements.

Proof The Poincaré-Friedrichs inequality implies that $\|\mathbf{v}\|_{1,\Omega}^2 \lesssim \|\mathbf{v}\|_{1,h,\Omega}^2$, while Jensen's and the standard trace inequality ensure that $\langle \mathbf{v} \rangle_{\Gamma_k}^2 \lesssim \|\mathbf{v}\|_{0,\Gamma_k}^2 \lesssim \|\mathbf{v}\|_{1,\Omega}^2$ for any $k = 1, \ldots, N$. As a consequence of that $\|\mathbf{v}\|_{1,h,\Omega}^2 \lesssim h^{-1} \|\mathbf{v}\|_{1,\Omega}^2$. □

We will also make use of the following inverse inequalities, see [12], that hold true for any $\Sigma \subseteq \partial\Omega$, provided that T_h is shape regular and quasi-uniform,

$$h \|\nabla \mathbf{v}_h\|_{0,\Omega} \lesssim \|\mathbf{v}_h\|_{0,\Omega}, \quad h^{\frac{1}{2}} \|\mathbf{v}_h\|_{0,\Sigma} \lesssim \|\mathbf{v}_h\|_{0,\Omega}, \ \forall \mathbf{v}_h \in \mathcal{V}_h^r. \tag{13}$$

First we address the basic properties of $a_h(\cdot,\cdot)$ that are summarized in the following results.

Lemma 3.3 *Choosing* γ *large enough, the bilinear form* $a_h(\cdot,\cdot)$ *is positive in the norm* $\| \cdot \|_{1,h,\Omega}$*. Precisely,*

$$\|\mathbf{v}_h\|_{1,h,\Omega}^2 \lesssim a_h(\mathbf{v}_h, \mathbf{v}_h), \ \forall \mathbf{v}_h \in \mathcal{V}_h^r.$$

Proof It is straightforward to verify that

$$a_h(\mathbf{v}_h, \mathbf{v}_h) = |\mathbf{v}_h|_{1,\Omega}^2 + \textstyle\sum_{k=1}^N \big[\gamma h^{-1} \langle \mathbf{v}_h \rangle_{\Gamma_k}^2 - 2 \langle \nabla \mathbf{v}_h \cdot \mathbf{n}, \mathbf{v}_h \rangle_{\Gamma_k} \big],$$

where the last term on the right hand side admits the following upper bound, owing to (12) and (13)

$$\begin{aligned} |\langle \nabla \mathbf{v}_h \cdot \mathbf{n}, \mathbf{v}_h \rangle_{\Gamma_k}| &\leq \epsilon h \langle \nabla \mathbf{v}_h \cdot \mathbf{n} \rangle_{\Gamma_k}^2 + (4h\epsilon)^{-1} \langle \mathbf{v}_h \rangle_{\Gamma_k}^2 \\ &\lesssim \epsilon h \|\nabla \mathbf{v}_h \cdot \mathbf{n}\|_{0,\Gamma_k}^2 + (4h\epsilon)^{-1} \langle \mathbf{v}_h \rangle_{\Gamma_k}^2 \\ &\lesssim \epsilon |\mathbf{v}_h|_{1,\Omega}^2 + (4h\epsilon)^{-1} \langle \mathbf{v}_h \rangle_{\Gamma_k}^2 . \end{aligned}$$

Combining the previous estimates we obtain,

$$a_h(\mathbf{v}_h, \mathbf{v}_h) \gtrsim (1-\epsilon)|\mathbf{v}_h|_{1,\Omega}^2 + \textstyle\sum_{k=1}^N \big(\gamma - \frac{1}{4\epsilon}\big) h^{-1} \langle \mathbf{v}_h \rangle_{\Gamma_k}^2 \gtrsim \|\mathbf{v}_h\|_{1,h,\Omega}^2, \ \forall \mathbf{v}_h \in \mathcal{V}_h^r,$$

for a sufficiently small ϵ and γ large enough. □

A suitable choice of γ will be discussed in section 4. Concerning the bilinear form $b_h(\cdot,\cdot)$, we proceed as in [13] and we assume that the discrete spaces $\mathcal{V}_h^r$ and Q_h^s are *inf-sup* compatible.

Hypothesis 3.4 *For any $q_h \in Q_h^s$ there exists $\mathbf{v}_h \in \mathcal{V}_h^r \cap \mathcal{H}_0^1(\Omega)$ such that*

$$(q_h, \nabla \cdot \mathbf{v}_h)_\Omega = \|q_h\|_{0,\Omega}^2, \quad \|\mathbf{v}_h\|_{1,\Omega} \lesssim \|q_h\|_{0,\Omega}.$$

It is well known that 3.4 is satisfied by the high-order Taylor-Hood elements, see [14], corresponding to $r \geq 2,\ s = r-1$. Since $\mathbf{v}_h \in \mathcal{H}_0^1(\Omega)$ we conclude that the additional boundary terms of $b_h(\cdot,\cdot)$, arising from the penalty formulation, do not interfere with the *inf-sup* stability of the method.

Corollary 3.5 *Under assumption 3.4, for any $q_h \in Q_h^s$ there exists $\mathbf{v}_h \in \mathcal{V}_h^r \cap \mathcal{H}_0^1(\Omega)$ such that*

$$b_h(q_h, \mathbf{v}_h) \gtrsim \|q_h\|_{0,\Omega}^2, \quad \|\mathbf{v}_h\|_{1,\Omega} = \|\mathbf{v}_h\|_{1,h,\Omega} \lesssim \|q_h\|_{0,\Omega}. \tag{14}$$

Let us now reformulate problem (11) for the couple $(\mathbf{u}_h, p_h) \in \mathcal{W}_h := \mathcal{V}_h^r \times Q_h^s$, where $\mathcal{W}_h$ is endowed with the norm $|||(\mathbf{v}_h, q_h)|||_h^2 := \|\mathbf{v}_h\|_{1,h,\Omega}^2 + \|q_h\|_{0,\Omega}^2$. For any $(\mathbf{u}_h, p_h),\ (\mathbf{v}_h, q_h) \in \mathcal{W}_h$ we also define

$$\begin{aligned} &c_h^{\pm}\big((\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h) \big) := a_h(\mathbf{u}_h, \mathbf{v}_h) + b_h(p_h, \mathbf{v}_h) \pm b_h(q_h, \mathbf{u}_h), \\ &K_h^{\pm}(\mathbf{v}_h, q_h) := F_h(\mathbf{v}_h) \pm G_h(q_h). \end{aligned}$$

Then, problem (11) is equivalent to find $(\mathbf{u}_h, p_h) \in \mathcal{W}_h$ such that

$$c_h^{\pm}\big((\mathbf{u}_h, p_h), (\mathbf{v}_h, q_h) \big) = K_h^{\pm}(\mathbf{v}_h, q_h), \ \forall (\mathbf{v}_h, q_h) \in \mathcal{W}_h. \tag{15}$$

Lemma 3.6 *Let $(\mathbf{u^0}, p) \in \mathcal{W^0}$ be the solution of problem (4), let $\mathbf{u} = (\mathbf{u^0} + \mathbf{w}) \in \mathcal{W}$ be the weak solution of (1) and let $(\mathbf{u}_h, p_h) \in \mathcal{W}_h$ be the solution of (15). Then $c_h^{\pm}\big((\mathbf{u}, p), (\mathbf{v}, q) \big) = K_h(\mathbf{v}, q)$ for all $(\mathbf{v}, q) \in \mathcal{W}$ and $c_h^{\pm}\big((\mathbf{u} - \mathbf{u}_h), (\mathbf{v}_h, q_h) \big) = 0$ for all $(\mathbf{v}_h, q_h) \in \mathcal{W}_h$.*

Proof We observe that problem (15) is conformal with (10) that is satisfied for any $(\mathbf{v}, q) \in \mathcal{W}$, provided that $(\mathbf{u}, p) \in \mathcal{W}$ is the weak solution of problem (1). □

Moreover, since $c_h^+(\cdot,\cdot)$ is symmetric and problem (4) is self-adjoint, we immediately obtain the following *adjoint consistency* property.

Corollary 3.7 *Let $(\mathbf{z_g^0}, r_g) \in \mathcal{W}^{\mathbf{0}}$ be the solution of problem (6).*
Then $c_h^+\big(\,(\mathbf{v},q)\,,(\mathbf{z_g^0}, r_g)\,\big) = (\mathbf{g},\mathbf{v})_\Omega$ for all $(\mathbf{v},q) \in \mathcal{W}$.

Lemma 3.8 *The bilinear forms $c_h^\pm(\cdot,\cdot)$ are bounded. Precisely they satisfy,*

$$c_h^\pm\big(\,(\mathbf{u}_h,p_h)\,,(\mathbf{v}_h,q_h)\,\big) \lesssim |||\mathbf{u}_h,p_h|||_h|||\mathbf{v}_h,q_h|||_h, \quad \forall(\mathbf{u}_h,p_h),\ (\mathbf{v}_h,q_h) \in \mathcal{W}_h.$$

Proof Concerning the terms of $a_h(\cdot,\cdot)$, owing to the Cauchy-Schwarz inequality we immediately get

$$(\nabla\mathbf{u}_h,\nabla\mathbf{v}_h)_\Omega+\gamma h^{-1}\textstyle\sum_{k=1}^N \langle\mathbf{u}_h,\mathbf{v}_h\rangle_{\Gamma_k} \le |\mathbf{u}_h|_{1,\Omega}|\mathbf{v}_h|_{1,\Omega}+\gamma h^{-1}\sum_{k=1}^N \langle\mathbf{u}_h\rangle_{\Gamma_k}\langle\mathbf{v}_h\rangle_{\Gamma_k},$$

while for the consistency and symmetry terms, owing to (12) and (13), we obtain

$$\begin{aligned}|\,\langle\nabla\mathbf{u}_h\cdot\mathbf{n},\mathbf{v}_h\rangle_{\Gamma_k}\,| &\le \big(h^{\frac12}\langle\nabla\mathbf{u}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\big)\big(h^{-\frac12}\langle\mathbf{v}_h\rangle_{\Gamma_k}\big)\\ &\lesssim \big(|\mathbf{u}_h|_{1,\Omega}\big)\big(h^{-\frac12}\langle\mathbf{v}_h\rangle_{\Gamma_k}\big).\end{aligned}$$

As regards $b_h(\cdot,\cdot)$ we easily see that $|\,(p_h,\nabla\cdot\mathbf{v}_h)_\Omega\,| \lesssim \|p_h\|_{0,\Omega}|\mathbf{v}_h|_{1,\Omega}$ and

$$\begin{aligned}|\,\langle p_h,\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\,| &\lesssim \big(h^{\frac12}\langle p_h\rangle_{\Gamma_k}\big)\big(h^{-\frac12}\langle\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\big)\\ &\lesssim \big(h^{\frac12}\|p_h\|_{0,\Gamma_k}\big)\big(h^{-\frac12}\langle\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\big)\\ &\lesssim \|p_h\|_{0,\Omega}\big(h^{-\frac12}\langle\mathbf{v}_h\rangle_{\Gamma_k}\big).\end{aligned}$$

The desired result follows from the combination of the previous estimates. □

Since $K_h(\mathbf{v}_h,q_h)$ is a linear and continuous functional in $\mathcal{W}_h$, the well posedness of problem (15) is ensured by the following result (see [15], Ch.2., Th. 2.22).

Theorem 3.9 *Provided that 3.4 is satisfied and under assumptions of Lemma 3.3, for any $(\mathbf{u}_h,p_h) \in \mathcal{W}_h$ there exists $(\mathbf{w}_h,s_h) \in \mathcal{W}_h$ such that*

$$c_h^-\big((\mathbf{u}_h,p_h),(\mathbf{w}_h,s_h)\big) \gtrsim |||(\mathbf{u}_h,p_h)|||_h^2, \quad |||(\mathbf{w}_h,s_h)|||_h \lesssim |||(\mathbf{u}_h,p_h)|||_h. \tag{16}$$

Proof For any $(\mathbf{u}_h,p_h) \in \mathcal{W}_h$ let us choose $(\mathbf{w}_h,s_h) = (\mathbf{u}_h,p_h)+\delta(\mathbf{v}_h,0)$ where $\mathbf{v}_h \in \mathcal{V}_h^r \cap \mathcal{H}_0^1(\Omega)$ refers to Corollary 3.5. Combining the following estimates,

$$\begin{aligned}c_h^-\big((\mathbf{u}_h,p_h),(\mathbf{u}_h,p_h)\big) &\gtrsim \|\mathbf{u}_h\|_{1,h,\Omega}^2,\\ c_h^-\big((\mathbf{u}_h,p_h),(\mathbf{v}_h,0)\big) &\gtrsim \|p_h\|_{0,\Omega}^2 - \|\mathbf{u}_h\|_{1,h,\Omega}\|\mathbf{v}_h\|_{1,h,\Omega}\\ &\gtrsim (1-\epsilon)\|p_h\|_{0,\Omega}^2 - \frac{1}{\epsilon}\|\mathbf{u}_h\|_{1,h,\Omega}^2,\end{aligned}$$

where we have exploited $\|\mathbf{v}_h\|_{1,h,\Omega} = \|\mathbf{v}_h\|_{1,\Omega}$ since $\mathbf{v}_h \in \mathcal{H}_0^1(\Omega)$. Then, we easily obtain that

$$c_h^-\big((\mathbf{u}_h, p_h), (\mathbf{w}_h, r_h)\big) \gtrsim \big(1 - C\frac{\delta}{\epsilon}\big)\|\mathbf{u}_h\|_{1,h,\Omega}^2 + \delta(1-\epsilon)\|p_h\|_{0,\Omega}^2.$$

We complete the proof observing that

$$|||(\mathbf{w}_h, s_h)|||_h^2 = \|\mathbf{u}_h\|_{1,h,\Omega}^2 + \|\mathbf{v}_h\|_{1,\Omega}^2 + \|p_h\|_{0,\Omega}^2 \lesssim \|\mathbf{u}_h\|_{1,h,\Omega}^2 + \|p_h\|_{0,\Omega}^2.$$

□

Then, we address the approximation properties of $\mathcal{W}_h$ in the norm $|||\cdot|||_h$.

Lemma 3.10 *Assume $\mathbf{u} \in \mathcal{H}^{r+1}(\Omega)$ and $p \in H^{s+1}(\Omega)$ with $r, s > 0$. Then it holds,*

$$\inf_{\mathbf{v}_h \in \mathcal{V}_h^r} \|\mathbf{u} - \mathbf{v}_h\|_{1,h,\Omega} \lesssim h^r |\mathbf{u}|_{r+1,\Omega}, \quad \inf_{q_h \in Q_h^s} \|p - q_h\|_{0,\Omega} \lesssim h^{s+1}|p|_{s+1,\Omega}. \tag{17}$$

Proof For the first estimate we remind that, owing to Jensen's inequality we have

$$\|\mathbf{v}\|_{1,h,\Omega}^2 \lesssim |\mathbf{v}|_{1,\Omega}^2 + h^{-1}\textstyle\sum_{k=1}^N \|\mathbf{v}\|_{0,\Gamma_k}^2.$$

To obtain the desired result, we combine the trace inequality with the approximation properties of $\mathcal{V}_h^r$, [12],

$$\|\mathbf{v}\|_{0,\Gamma_k}^2 \lesssim h^{-1}\|\mathbf{v}\|_{0,\Omega}^2 + h|\mathbf{v}|_{1,\Omega}^2, \quad \|\mathbf{u} - \mathbf{v}_h\|_{0,\Omega} + h|\mathbf{u} - \mathbf{v}_h|_{1,\Omega} \lesssim h^{r+1}|\mathbf{u}|_{r+1,\Omega}.$$

The second estimate is straightforward owing to the approximation properties of Q_h^s. □

Theorem 3.11 *Let $(\mathbf{u}, p) \in \mathcal{W}$ be the weak solution of (1), and let $(\mathbf{u}_h, p_h) \in \mathcal{W}_h$ be the solution of (15). Under the assumptions of Theorem 3.9 and Lemma 3.6, we obtain,*

$$|||(\mathbf{u} - \mathbf{u}_h, p - p_h)|||_h \lesssim \inf_{(\mathbf{v}_h, q_h) \in \mathcal{W}_h} |||(\mathbf{u} - \mathbf{v}_h, p - q_h)|||_h. \tag{18}$$

In particular when $r \geq 2$, $s = r - 1$ with $\mathbf{u} \in \mathcal{V} \cap H^{r+1}(\Omega)$, $p \in H^r(\Omega)$ we get,

$$|||(u - u_h, p - p_h)|||_h \lesssim h^r\big(\|\mathbf{u}\|_{r+1,\Omega} + \|p\|_{r,\Omega}\big). \tag{19}$$

Proof For the proof of (18) let us decompose the error $(\mathbf{u} - \mathbf{u}_h, p - p_h)$ in two parts

$$\begin{aligned} \mathbf{u} - \mathbf{u}_h &:= \mathbf{e}_\pi + \mathbf{e}_h := (\mathbf{u} - \mathbf{v}_h) + (\mathbf{v}_h - \mathbf{u}_h), \\ p - p_h &:= y_\pi + y_h := (p - q_h) + (q_h - p_h),\ \forall (\mathbf{v}_h, q_h) \in \mathcal{W}_h. \end{aligned}$$

Using Theorem 3.9 and Lemmas 3.6, 3.8 we get,

$$\begin{aligned}|||(\mathbf{e}_h, y_h)|||_h |||(\mathbf{w}_h, s_h)|||_h &\lesssim c_h^-((\mathbf{e}_h, y_h), (\mathbf{w}_h, s_h)) \\ &\lesssim c_h^-((\mathbf{e}_\pi, y_\pi), (\mathbf{w}_h, s_h)) \\ &\lesssim |||(\mathbf{e}_\pi, y_\pi)|||_h |||(\mathbf{w}_h, s_h)|||_h. \end{aligned} \tag{20}$$

We combine (20) with the triangle inequality,

$$|||(\mathbf{u} - \mathbf{u}_h, p - p_h)|||_h \lesssim |||(\mathbf{e}_h, y_h)|||_h + |||(\mathbf{e}_\pi, y_\pi)|||_h \lesssim |||(\mathbf{e}_\pi, y_\pi)|||_h$$

and exploiting the generality of $(\mathbf{v}_h, q_h)$, we obtain (18). Estimate (19) is recovered combining (18) with (17). □

We conclude with the analysis on the approximation properties of (11) in the $\mathcal{L}^2$ norm for the velocity.

Theorem 3.12 *Let $(\mathbf{u}, p) \in \mathcal{W}$ be the weak solution of* (1) *satisfying assumption 2.4, and let $(\mathbf{u}_h, p_h) \in \mathcal{W}_h$ be the solution of* (15)*. Under the assumptions of Theorem 3.11 we have,*

$$\|u - u_h\|_{0,\Omega} \lesssim h|||(u - u_h, p - p_h)|||_h. \tag{21}$$

Proof Let $(\mathbf{z}_g^{\mathbf{0}}, r_g) \in \big(\mathcal{H}^2(\Omega) \cap \mathcal{V}^{\mathbf{0}}\big) \times \big(H^1(\Omega) \cap Q\big)$ be the solution of (6). First, owing to the adjoint consistency, i.e. Corollary 3.7, we notice that

$$\|\mathbf{v}\|_{0,\Omega} = \sup_{\mathbf{g} \in \mathcal{L}^2(\Omega)} \frac{(\mathbf{g}, \mathbf{v})_\Omega}{\|\mathbf{g}\|_{0,\Omega}} = \sup_{\mathbf{g} \in \mathcal{L}^2(\Omega)} \frac{c_h^+\big(\,(\mathbf{v}, q)\,, \big(\mathbf{z}_g^{\mathbf{0}}, r_g\big)\,\big)}{\|\mathbf{g}\|_{0,\Omega}},\ \forall (\mathbf{v}, q) \in \mathcal{W}. \tag{22}$$

Then, we choose $\mathbf{v} = \mathbf{u} - \mathbf{u}_h$ and $q = p - p_h$ and exploiting the Galerkin orthogonality, i.e. Lemma 3.6, and the continuity of $c_h^+(\cdot, \cdot)$, we obtain

$$\begin{aligned} c_h^+\big(\,(\mathbf{u} - \mathbf{u}_h, p - p_h)\,, \big(\mathbf{z}_g^{\mathbf{0}}, r_g\big)\,\big) &= c_h^+\big(\,(\mathbf{u} - \mathbf{u}_h, p - p_h)\,, \big(\mathbf{z}_g^{\mathbf{0}} - \mathbf{z}_h, r_g - r_h\big)\,\big) \\ &\lesssim |||(\mathbf{u} - \mathbf{u}_h, p - p_h)|||_h |||(\mathbf{z}_g^{\mathbf{0}} - \mathbf{z}_h, r_g - r_h)|||_h,\ \forall (\mathbf{z}_h, r_h) \in \mathcal{W}_h. \end{aligned} \tag{23}$$

Owing to Lemma 3.10 with $r = 1$ and $s = 0$ and (5) we have,

$$\inf_{(\mathbf{z}_h, r_h) \in \mathcal{W}_h} |||(\mathbf{z}_g^{\mathbf{0}} - \mathbf{z}_h, r_g - r_h)|||_h \lesssim h|||(\mathbf{z}_g^{\mathbf{0}}, r_g)|||_* \lesssim h\|\mathbf{g}\|_{0,\Omega}. \tag{24}$$

The desired result is obtained combining (22), (23) and (24). □

3.3 Algebraic properties of the penalty method

Let $M_v = \dim(\mathcal{V}_h^r)$ be the dimension of $\mathcal{V}_h^r$ and let $\{\boldsymbol{\varphi}_i\}_{i=1}^{M_v}$ be its Lagrangian finite element basis. Proceeding similarly for $Q_h^s = \operatorname{span}\big(\{\phi_i\}_{i=1}^{M_p}\big)$, we set $\mathbf{u}_h =$

$\sum_{j=1}^{M_v} u_j \boldsymbol{\varphi}_j$, $\mathbf{U} = \{u_j\}_{j=1}^{M_v}$, $p_h = \sum_{j=1}^{M_p} p_j \phi_j$, $\mathbf{P} = \{p_j\}_{j=1}^{M_p}$ and we define the following matrices,

$$A_{h,ij} = a_h(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i), \quad B_{h,ij} = b_h(\phi_i, \boldsymbol{\varphi}_j), \quad \mathbf{F}_{h,i} = F_h(\boldsymbol{\varphi}_i), \quad \mathbf{G}_{h,i} = \mathbf{G}_h(\phi_i).$$

We easily see that problem (11) is equivalent to the following algebraic saddle point system,

$$\begin{bmatrix} A_h & B_h^T \\ B_h & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_h \\ \mathbf{G}_h \end{bmatrix}. \tag{25}$$

We aim to show that, although system system (25) is indefinite, it can be solved resorting to the Schur complement of matrix A_h, that is $R_h \mathbf{P} = \mathbf{J}_h$ where $R_h := B_h A_h^{-1} B_h^T$ and $\mathbf{J}_h = B_h A_h^{-1} \mathbf{F}_h - \mathbf{G}_h$. Let $\sigma(M)$ be the spectrum of a real valued square matrix M and let $\mathcal{R}(M)$, $\mathcal{K}(M)$ be its range and nullspace respectively. If M is symmetric positive definite, we denote with $\mathcal{X}(M)$ its spectral condition number. We also remind the following property, where d is the number of space dimensions of Ω.

Property 3.13 *Provided that T_h is shape-regular and quasi-uniform, for each $\mathbf{v}_h = \sum_{j=1}^{M_v} v_j \boldsymbol{\varphi}_j \in \mathcal{V}_h^r$, $\mathbf{V} = \{v_j\}_{j=1}^{M_v} \in \mathbb{R}^{M_v}$ endowed with the Euclidean norm $\|\mathbf{V}\|_{M_v}$, we have,*

$$h^d \|\mathbf{V}\|_{M_v}^2 \lesssim \|\mathbf{v}_h\|_{0,\Omega}^2 \lesssim h^d \|\mathbf{V}\|_{M_v}^2. \tag{26}$$

We refer to [16], Proposition 6.3.1 for a proof. Then, we are ready to prove the following fundamental properties of A_h and B_h.

Lemma 3.14 *Under assumption 3.4, the following properties are satisfied,*

$$h^d \lesssim \sigma(A_h) \lesssim h^{d-2}, \ \mathcal{X}(A_h) \lesssim h^{-2}; \quad \mathcal{R}(B_h) = \mathbb{R}^{M_p} \text{ or } \mathcal{K}(B_h^T) = \{\mathbf{0}\}. \tag{27}$$

As a result of that, matrix R_h is symmetric positive definite (SPD).

Proof For the lower bound of the spectrum of A_h we combine the positivity of $a_h(\cdot,\cdot)$ with the lower bound of Lemma 3.2,

$$h^d \|\mathbf{V}\|_{M_v}^2 \lesssim \|\mathbf{v}_h\|_{0,\Omega}^2 \lesssim \|\mathbf{v}_h\|_{1,\Omega}^2 \lesssim \|\mathbf{v}_h\|_{1,h,\Omega}^2 \lesssim a_h(\mathbf{v}_h, \mathbf{v}_h) = \mathbf{V}^T A_h \mathbf{V}.$$

Concerning the upper bound, we first observe that inequalities (12), (13) ensure that

$$\begin{aligned} \|\mathbf{v}_h\|_{1,h,\Omega}^2 &\lesssim |\mathbf{v}_h|_{1,\Omega}^2 + h^{-1} \textstyle\sum_{k=1}^N \|\mathbf{v}_h\|_{0,\Gamma_k}^2 \\ &\lesssim h^{-2} \|\mathbf{v}_h\|_{0,\Omega}^2 + h^{-1} \textstyle\sum_{k=1}^N \|\mathbf{v}_h\|_{0,\Gamma_k}^2 \lesssim h^{-2} \|\mathbf{v}_h\|_{0,\Omega}^2, \end{aligned}$$

that can be combined with the boundedness of $a_h(\cdot,\cdot)$ to complete the proof of the first part of (27),

$$\mathbf{V}^T A_h \mathbf{V} = a_h(\mathbf{v}_h, \mathbf{v}_h) \lesssim \|\mathbf{v}_h\|_{1,h,\Omega}^2 \lesssim h^{-2} \|\mathbf{v}_h\|_{0,\Omega}^2 \lesssim h^{d-2} \|\mathbf{V}\|_{M_v}^2.$$

The *inf-sup* condition (14) ensures that $\mathcal{R}(B_h) = \mathbb{R}^{M_p}$ and $\mathcal{K}(B_h^T) = \{\mathbf{0}\}$ directly follows from the Orthogonal Decomposition Theorem. Finally, since A_h is SPD, then A_h^{-1} is also SPD, that is $\mathbf{W}^T A_h^{-1}\mathbf{W} > 0$ for any $\mathbf{W} \in \mathbb{R}^{M_v}$ with $\mathbf{W} \neq \mathbf{0}$. Restricting to those $\mathbf{W}$ such that $\mathbf{W} = B_h^T\mathbf{Q}$, and observing that $\mathbf{W} \neq \mathbf{0}$ for any $\mathbf{Q} \neq \mathbf{0}$ because $\mathcal{K}(B_h^T) = \{\mathbf{0}\}$, we conclude that $\mathbf{Q}^T B_h A_h^{-1} B_h^T \mathbf{Q} > 0$ for any $\mathbf{Q} \neq \mathbf{0}$, i.e. R_h is SPD. □

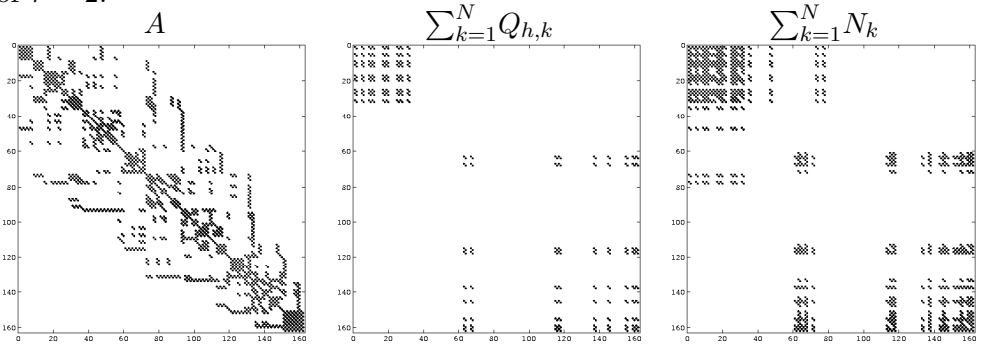
As a consequence of Lemma 3.14, matrix R_h can be solved by means of the conjugate gradient method (CG). In the case of *inf-sup* stable finite elements with standard boundary conditions this strategy is also optimal, because the Schur complement matrix is spectrally equivalent to the pressure mass matrix, whose condition number is uniformly bounded with respect to h. By consequence, the CG algorithm is an efficient strategy provided that an efficient solver for the velocity stiffness matrix, A_h, is available. However, in the case of system (25) this approach features some additional difficulties.

First, we notice that the application of the penalty method to defective boundary conditions significantly affects the sparsity pattern of matrix A_h, because the boundary terms of $a_h(\cdot,\cdot)$ that lie on Γ_k are *non local*. More precisely, let us split A_h into different parts $A_h = A + \sum_{k=1}^{N}\big(Q_{h,k} - N_k\big)$ where

$$\begin{aligned} A_{ij} &:= a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i), \\ N_{ij,k} &:= \left\langle \nabla\boldsymbol{\varphi}_j \cdot \mathbf{n}, \boldsymbol{\varphi}_i \right\rangle_{\Gamma_k} + \left\langle \nabla\boldsymbol{\varphi}_i \cdot \mathbf{n}, \boldsymbol{\varphi}_j \right\rangle_{\Gamma_k}, \\ Q_{h,k} &:= \gamma h^{-1} \left\langle \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i \right\rangle_{\Gamma_k}. \end{aligned}$$

Matrix $Q_{h,k}$ couples together all the degrees of freedom belonging to each section Γ_k, while matrix N_k couples all the degrees of freedom that belong to elements whose edges lie on Γ_k. For instance, if $\Gamma = \partial\Omega$ the bandwidth of A_h can easily be equal to M_v although A_h is still sparse, as illustrated in figure 1. As a result of that, direct methods for linear systems do not seem to be easily applicable to A_h, because of excessive fill-in. Simultaneously, since Lemma 3.14 imply that $\mathcal{X}(A_h) = \mathcal{O}(h^{-2})$, for realistic problems any iterative solver for A_h needs a preconditioner. Unfortunately, for the aforementioned reasons, the computational cost to construct matrix preconditioners derived from incomplete factorizations of A_h increases with respect to the usual case. Nevertheless, we notice that the minimum and maximum eigenvalues of A_h scale with respect to h as the ones of any finite element stiffness matrix. By consequence, for any problem with Dirichlet boundary conditions on $\partial\Omega \setminus \Gamma \neq \emptyset$, one may think to apply to A_h any preconditioner that is suitable to problem (1) complemented with homogeneous Neumann conditions on Γ and Dirichlet conditions on $\partial\Omega \setminus \Gamma$. More precisely, let A be the finite element matrix corresponding to $a(\cdot,\cdot)$ and let $P = HH^T$ be the incomplete Cholesky factorization of A. By formally replacing $A_h\mathbf{U} = \mathbf{F}$ with $H^{-1}A_hH^{-T}\mathbf{V} = H^{-1}\mathbf{F}$, $\mathbf{V} = H^T\mathbf{U}$, we apply matrix P as a preconditioner for A_h into the conjugate gradient algorithm. We notice that the sparsity pattern

Figure 1: The sparsity pattern of different parts of $A_h = A + \sum_{k=1}^{N}(Q_{h,k} - N_k)$ for $r = 2$.



of A is not affected by the drawbacks relative to A_h and the forthcoming experiments will show that this preconditioning strategy is equally effective to both A and A_h.

Second, because of the penalty technique, the condition number of the Schur complement matrix R_h depends on h^{-1}. By consequence, the number of iterations necessary to approximate $R_h\mathbf{P} = \mathbf{J}_h$ by means of the CG algorithm, will be proportional to $h^{-1/2}$. To prove this statement, we introduce the following notation. Reminding the equivalence between $\mathbf{v}_h \in \mathcal{V}_h^r$ and $\mathbf{V} \in \mathbb{R}^{M_v}$, $q_h \in Q_h^s$ and $\mathbf{Q} \in \mathbb{R}^{M_p}$ we introduce the following discrete norms,

$$\|\mathbf{U}\|_{*,M_v} = \sup_{\mathbf{V}\in\mathbb{R}^{M_v}} \frac{(\mathbf{U},\mathbf{V})_{M_v}}{\|\mathbf{v}_h\|_{1,\Omega}}, \quad \|\mathbf{U}\|_{*,h,M_v} = \sup_{\mathbf{V}\in\mathbb{R}^{M_v}} \frac{(\mathbf{U},\mathbf{V})_{M_v}}{\|\mathbf{v}_h\|_{1,h,\Omega}},$$

where $(\mathbf{U},\mathbf{V})_{M_v}$ denotes the Euclidean scalar product on $\mathbb{R}^{M_v}$. Owing to these definitions, the basic properties of $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ can be straightforwardly reinterpreted as follows.

Corollary 3.15 *Lemma 3.3 and Corollary 3.5 are respectively equivalent to*

$$\|\mathbf{v}_h\|_{1,h,\Omega} \lesssim \|A_h\mathbf{V}\|_{*,h,M_v}, \quad \|q_h\|_{0,\Omega} \lesssim \|B_h^T\mathbf{Q}\|_{*,M_v}, \quad \forall(\mathbf{v}_h, q_h) \in \mathcal{W}_h. \tag{28}$$

Lemma 3.8 is equivalent to

$$\|A_h\mathbf{V}\|_{*,M_v} \lesssim h^{-\frac{1}{2}}\|\mathbf{v}_h\|_{1,h,\Omega}, \quad \|B_h^T\mathbf{Q}\|_{*,h,M_v} \lesssim \|q_h\|_{0,\Omega}, \quad \forall(\mathbf{v}_h, q_h) \in \mathcal{W}_h. \tag{29}$$

Proof Lemma 3.3 can be reformulated as follows,

$$\|\mathbf{v}_h\|_{1,h,\Omega} \lesssim \frac{a_h(\mathbf{v}_h,\mathbf{v}_h)}{\|\mathbf{v}_h\|_{1,h,\Omega}} = \frac{(A_h\mathbf{V},\mathbf{V})_{M_v}}{\|\mathbf{v}_h\|_{1,h,\Omega}} \lesssim \sup_{\mathbf{W}\in\mathbb{R}^{M_v}} \frac{(A_h\mathbf{V},\mathbf{W})_{M_v}}{\|\mathbf{w}_h\|_{1,h,\Omega}}, \ \forall\mathbf{v}_h \in \mathcal{V}_h^r,$$

that is $\|\mathbf{v}_h\|_{1,h,\Omega} \lesssim \|A_h\mathbf{V}\|_{*,h,M_v}$. Proceeding similarly for Corollary 3.5, there exists $\mathbf{v}_h$ such that

$$\|q_h\|_{0,\Omega} \lesssim \frac{b_h(q_h,\mathbf{v}_h)}{\|\mathbf{v}_h\|_{1,\Omega}} = \frac{\left(B_h^T\mathbf{Q},\mathbf{V}\right)_{M_v}}{\|\mathbf{v}_h\|_{1,\Omega}} \lesssim \|B_h^T\mathbf{Q}\|_{*,M_v},\ \forall q_h \in Q_h^s.$$

Concerning Lemma 3.8, we observe it is equivalent to,

$$a_h(\mathbf{v}_h,\mathbf{w}_h) \lesssim \|\mathbf{v}_h\|_{1,h,\Omega}\|\mathbf{w}_h\|_{1,h,\Omega}, \quad b_h(q_h,\mathbf{v}_h) \lesssim \|q_h\|_{0,\Omega}\|\mathbf{v}_h\|_{1,h,\Omega}.$$

Then, we obtain

$$\begin{aligned}
&(A_h\mathbf{V},\mathbf{W})_{M_v} = a_h(\mathbf{v}_h,\mathbf{w}_h) \lesssim h^{-1/2}\|\mathbf{v}_h\|_{1,h,\Omega}\|\mathbf{w}_h\|_{1,\Omega},\ \forall \mathbf{v}_h,\mathbf{w}_h \in \mathcal{V}_h^r\\
&\left(B_h^T\mathbf{Q},\mathbf{V}\right)_{M_v} = b_h(q_h,\mathbf{v}_h) \lesssim \|q_h\|_{0,\Omega}\|\mathbf{v}_h\|_{1,h,\Omega},\ \forall (\mathbf{v}_h,q_h) \in \mathcal{W}_h,
\end{aligned}$$

that is $\|A_h\mathbf{V}\|_{*,M_v} \lesssim h^{-\frac{1}{2}}\|\mathbf{v}_h\|_{1,h,\Omega}$ and $\|B_h^T\mathbf{Q}\|_{*,h,M_v} \lesssim \|q_h\|_{0,\Omega}$. □

We are now ready to prove the following statement.

Lemma 3.16 *Under assumptions of Lemma 3.3, 3.8 and Corollary 3.5, the spectrum of matrix R_h satisfies the following properties,*

$$h^{d+1} \lesssim \sigma(R_h) \lesssim h^d, \quad \textit{and} \quad \mathcal{X}(R_h) \lesssim h^{-1}.$$

Proof Let $\lambda_m \in \mathbb{R}^+$ be the minimum eigenvalue of R_h and let $\mathbf{Q}_m \in \mathbb{R}^{M_p}$ be the corresponding eigenvector. We introduce $\mathbf{V} = B_h^T\mathbf{Q}_m \in \mathbb{R}^{M_v}$ and $\mathbf{W} = A_h^{-1}\mathbf{V} \in \mathbb{R}^{M_v}$ such that $\mathbf{V} = A_h\mathbf{W}$. Then, exploiting (28) and (29) we obtain,

$$\begin{aligned}
&\lambda_m\|\mathbf{Q}_m\|_{M_p}^2 = (R_h\mathbf{Q}_m,\mathbf{Q}_m)_{M_p} = \left(A_h^{-1}B_h^T\mathbf{Q}_m, B_h^T\mathbf{Q}_m\right)_{M_v} = \left(A_h^{-1}\mathbf{V},\mathbf{V}\right)_{M_v},\\
&\left(A_h^{-1}\mathbf{V},\mathbf{V}\right)_{M_v} = (\mathbf{W},A_h\mathbf{W})_{M_v} \gtrsim \|\mathbf{w}_h\|_{1,h,\Omega}^2,\\
&h^{-1}\|\mathbf{w}_h\|_{1,h,\Omega}^2 \gtrsim \|A_h\mathbf{W}\|_{*,M_v}^2 = \|\mathbf{V}\|_{*,M_v}^2.
\end{aligned}$$

Combining the previous inequalities and exploiting Property 3.13 we easily obtain,

$$\lambda_m\|\mathbf{Q}_m\|_{M_p}^2 \gtrsim h\|B_h^T\mathbf{Q}_m\|_{*,M_v}^2 \gtrsim h\|q_{h,m}\|_{0,\Omega}^2 \gtrsim h^{d+1}\|\mathbf{Q}_m\|_{M_p}^2, \quad \text{i.e.} \quad \lambda_m \gtrsim h^{d+1}.$$

Then, we consider the maximum eigenvalue $\lambda_M \in \mathbb{R}^+$ of R_h, being $\mathbf{Q}_M \in \mathbb{R}^{M_p}$ its own eigenvector. Proceeding as in the previous case, we set $\mathbf{V} = B_h^T\mathbf{Q}_M \in \mathbb{R}^{M_v}$ and $\mathbf{W} = A_h^{-1}\mathbf{V} \in \mathbb{R}^{M_v}$ and we obtain,

$$\begin{aligned}
&\lambda_M\|\mathbf{Q}_M\|_{M_p}^2 = (R_h\mathbf{Q}_M,\mathbf{Q}_M)_{M_p} = \left(A_h^{-1}B_h^T\mathbf{Q}_M, B_h^T\mathbf{Q}_M\right)_{M_v} = \left(A_h^{-1}\mathbf{V},\mathbf{V}\right)_{M_v},\\
&\left(A_h^{-1}\mathbf{V},\mathbf{V}\right)_{M_v} = (\mathbf{W},\mathbf{V})_{M_v} \lesssim \|\mathbf{V}\|_{*,h,M_v}\|\mathbf{w}_h\|_{1,h,\Omega},\\
&\|\mathbf{w}_h\|_{1,h,\Omega} \lesssim \|A_h\mathbf{W}\|_{*,h,M_v} = \|\mathbf{V}\|_{*,h,M_v},
\end{aligned}$$

from which we easily obtain,

$$\lambda_M\|\mathbf{Q}_M\|_{M_p}^2 \lesssim \|B_h^T\mathbf{Q}_M\|_{*,h,M_v}^2 \lesssim \|q_{h,M}\|_{0,\Omega}^2 \lesssim h^d\|\mathbf{Q}_M\|_{M_p}^2, \quad \text{i.e.} \quad \lambda_M \lesssim h^d.$$

Since R_h is SPD, the conclusion $\mathcal{X}(R_h) \lesssim h^{-1}$ is straightforward. □

3.4 Numerical approximation of the augmented formulation

In this paragraph we briefly review the numerical approximation and the algebraic counterpart of the augmented problem (9), with the aim to compare its computational cost with the penalty formulation (11), or equivalently (25) in algebraic form. The numerical discretization of problem (9) by means of finite elements requires to find $\mathbf{u}_h \in \mathcal{V}_h^r$, $p_h \in Q_h^s$ and N vectors $\boldsymbol{\lambda}_{h,k} \in \mathbb{R}^d$ such that,

$$\begin{cases} a(\mathbf{u}_h, \mathbf{v}_h) + b(p_h, \mathbf{v}_h) + \sum_{k=1}^N \langle \boldsymbol{\lambda}_{h,k}, \mathbf{v}_h \rangle_{\Gamma_k} = (\mathbf{f}, \mathbf{v}_h)_\Omega \,, & \forall \mathbf{v}_h \in \mathcal{V}_h^r, \\ b(q_h, \mathbf{u}_h) = 0, & \forall q_h \in Q_h^s, \\ \langle \boldsymbol{\mu}_{h,k}, \mathbf{u}_h \rangle_{\Gamma_k} = \langle \boldsymbol{\mu}_{h,k}, \mathbf{U}_k \rangle_{\Gamma_k} \,, & \forall \boldsymbol{\mu}_{h,k} \in \mathbb{R}^d, \ k = 1, N. \end{cases} \tag{30}$$

To set up the algebraic counterpart of (30), we collect the multipliers relative to each section Γ_k into a single column vector $\boldsymbol{\Lambda} := [\boldsymbol{\lambda}_{h,1}, \ldots, \boldsymbol{\lambda}_{h,N}]^T \in R^{M_\lambda}$ being $M_\lambda = Nd$. Let $\{\boldsymbol{\mu}_i\}_{i=1}^{M_\lambda}$ be a basis of R^{M_λ} and let $\boldsymbol{\mu}_{i,k} \in \mathbb{R}^d$ be the restriction of $\boldsymbol{\mu}_i$ to the degrees of freedom associated to Γ_k, such that $\boldsymbol{\mu}_i = [\boldsymbol{\mu}_{i,1}, \ldots, \boldsymbol{\mu}_{i,N}]^T$. Then, given the following matrices and vectors,

$$\begin{aligned} &A_{ij} := a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i), \quad B_{ij} := b(\phi_i, \boldsymbol{\varphi}_j), \\ &\mathbf{F}_i := (\mathbf{f}, \boldsymbol{\varphi}_i)_\Omega \,, \quad D_{ij} := \textstyle\sum_{k=1}^N \langle \boldsymbol{\mu}_{i,k}, \boldsymbol{\varphi}_j \rangle_{\Gamma_k} \,, \quad \boldsymbol{\Upsilon}_i = \textstyle\sum_{k=1}^N \langle \boldsymbol{\mu}_{i,k}, \mathbf{U}_k \rangle_{\Gamma_k} \,, \end{aligned}$$

problem (30) is equivalent to find $\mathbf{U} \in \mathbb{R}^{M_v}$, $\mathbf{P} \in \mathbb{R}^{M_p}$, $\boldsymbol{\Lambda} \in \mathbb{R}^{M_\lambda}$ such that,

$$\begin{bmatrix} A & B^T & D^T \\ B & 0 & 0 \\ D & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \\ \boldsymbol{\Lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \\ \boldsymbol{\Upsilon} \end{bmatrix}. \tag{31}$$

We observe that (31) is a saddle point problem with a nested structure featuring two sets of multipliers, namely the pressure accounting for the divergence free constraint and the vectors $\boldsymbol{\lambda}_{h,k}$ accounting for the defective boundary conditions. For the solution of system (31) we apply the reduction to the Schur complement form. To this aim, it is convenient to condense the Stokes system into the matrix S, resorting to the following equivalent reformulation of (31),

$$\begin{bmatrix} S & D^T \\ D & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{X} \\ \boldsymbol{\Lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{K} \\ \boldsymbol{\Upsilon} \end{bmatrix}; \quad S = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}; \quad \mathbf{K} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}. \tag{32}$$

In [2] it is shown that the matrix $DS^{-1}D^T$ is symmetric positive definite, by consequence we determine $\boldsymbol{\Lambda}$ solving $DS^{-1}D^T\boldsymbol{\Lambda} = DS^{-1}\mathbf{K} - \boldsymbol{\Upsilon}$ by means of the conjugate gradient algorithm. Then, $\mathbf{X}$ is given by the solution of $S\mathbf{X} = \mathbf{K} - D^T\boldsymbol{\Lambda}$. We notice that each time the matrix S^{-1} is invoked, we need to solve a generic Stokes problem $S\mathbf{X} = \mathbf{K}$. In this case, we resort again to the Schur complement matrix, determining first the vector $\mathbf{P}$ such that $BA^{-1}B^T\mathbf{P} = BA^{-1}\mathbf{F}$ and then the corresponding velocity $\mathbf{U}$ given by $A\mathbf{U} = \mathbf{F} - B^T\mathbf{P}$. Provided that assumption 3.4 is satisfied, all these subsystems can be again solved by means of the conjugate gradient method.

We finally notice that in the forthcoming numerical tests we apply a penalty method for the Dirichlet boundary conditions $\partial\Omega \setminus \Gamma$. In this case, owing to the argument described in Lemma 3.16, we observe that the condition number of the Stokes Schur complement depends on h, namely $\mathcal{X}(BA^{-1}B^T) \simeq h^{-1}$.

In the following section we will compare the computational cost of the present solution strategy, with the one coming from (25).

3.5 Extension to the time dependent case

Since net flux defective boundary conditions are particularly significant for computational hemodynamics and blood flow is typically transient, more precisely pulsatile, we aim show that the penalty method and all its fundamental properties can be straightforwardly extended to the time dependent case.

We start from the time dependent version of problem (1): given $\mathbf{f}(t)$, the flow rates $\mathbf{U}_k(t)$ and the initial state $\mathbf{u}_0$, for any $t \in (0,T)$ we aim to find $(u(t), p(t))$ and the vectors $\mathbf{c}_k(t)$ such that,

$$\begin{cases} \partial_t \mathbf{u} - \nu\nabla^2\mathbf{u} + \nabla p = \mathbf{f}(t),\ \nabla\cdot\mathbf{u} = \mathbf{0}, & \text{in } \Omega\times(0,T), \\ \mathbf{u} = \mathbf{0}, & \text{on } \big(\partial\Omega\setminus\Gamma\big)\times(0,T), \\ \mathbf{u} = \mathbf{u}_0, & \text{on } \Omega\times\{t=0\}, \\ \frac{1}{|\Gamma_k|}\int_{\Gamma_k}\mathbf{u} = \mathbf{U}_k(t),\ p\mathbf{n} - \nu\nabla\mathbf{u}\cdot\mathbf{n} = \mathbf{c}_k(t) & \text{on } \Gamma_k\times(0,T),\ k=1,N, \end{cases} \tag{33}$$

where the viscosity coefficient ν is positive and bounded away from zero, i.e. $\nu \geq \nu_0 > 0$. Setting $\mathbf{u^0}(t) = \mathbf{u}(t) - \mathbf{w}(t)$, where $\mathbf{w}^{(i)}(t) = \sum_{k=1}^{N}\mathbf{U}_k^{(i)}(t)\mathbf{w}_k^{(i)}$ for $i = 1,\dots,d$, for any $t \in (0,T)$ the weak counterpart of problem (33) requires to find $(u^{\mathbf{0}}(t), p(t))$ such that

$$\begin{cases} \big(\partial_t\mathbf{u^0},\mathbf{v}\big)_\Omega + a^\nu(\mathbf{u^0},\mathbf{v}) + b(p,\mathbf{v}) = F^\nu(t,\mathbf{v}), & \forall\mathbf{v}\in\mathcal{V}^{\mathbf{0}}, \\ b(q,\mathbf{u^0}) = 0, & \forall q\in Q, \end{cases} \tag{34}$$

where $a^\nu(\mathbf{u},\mathbf{v}) := \int_\Omega \nu\nabla\mathbf{u} : \nabla\mathbf{v}$ and $F^\nu(t,\mathbf{v}) := \int_\Omega \mathbf{f}(t)\cdot\mathbf{v} - \int_\Omega \partial_t\mathbf{w}(t)\mathbf{v} - \int_\Omega \nu\nabla\mathbf{w}(t) : \nabla\mathbf{v}$. For the space discretization of problem (34) we exploit the penalty method (11). According to the definitions of $a^\nu(\mathbf{u},\mathbf{v})$ and $F^\nu(t,\mathbf{v})$ we set,

$$\begin{aligned} a_h^\nu(\mathbf{u}_h,\mathbf{v}_h) :=& (\nu\nabla\mathbf{u}_h,\nabla\mathbf{v}_h)_\Omega \\ &+ \textstyle\sum_{k=1}^{N}\big[\gamma h^{-1}\langle\nu\mathbf{u}_h,\mathbf{v}_h\rangle_{\Gamma_k} - \langle\nu\nabla\mathbf{u}_h\cdot\mathbf{n},\mathbf{v}_h\rangle_{\Gamma_k} - \langle\nu\nabla\mathbf{v}_h\cdot\mathbf{n},\mathbf{u}_h\rangle_{\Gamma_k}\big], \\ F_h^\nu(t,\mathbf{v}_h) :=& (\mathbf{f}(t),\mathbf{v})_\Omega + \textstyle\sum_{k=1}^{N}\big[\gamma h^{-1}\langle\nu\mathbf{U}_k(t),\mathbf{v}_h\rangle_{\Gamma_k} - \langle\mathbf{U}_k(t),\nu\nabla\mathbf{v}_h\cdot\mathbf{n}\rangle_{\Gamma_k}\big], \\ G_h(t,q_h) :=& \textstyle\sum_{k=1}^{N}\langle\mathbf{U}_k(t)\cdot\mathbf{n},q_h\rangle_{\Gamma_k}. \end{aligned}$$

For the time discretization of problem (34), we apply the classical $\theta-$method. Precisely, given a sequence of times $t^n = n\tau$, being $n\in\mathbb{N}$ and $\tau > 0$ a fixed time

step, and given $\mathbf{u}_h^0 = \pi_h^r \mathbf{u}_0$ being π_h^r the finite element interpolator from $\mathcal{V}$ onto $\mathcal{V}_h^r$, the fully discrete scheme requires to find a sequence of discrete functions $(\mathbf{u}_h^n, p_h^n) \in \mathcal{W}_h$, approximating $(\mathbf{u}(t^n), p(t^n)) \in \mathcal{W}$, given by

$$\mathbf{u}_{h,\theta}^n := \theta \mathbf{u}_h^n + (1-\theta)\mathbf{u}_h^{n-1}, \quad p_{h,\theta}^n := \theta p_h^n + (1-\theta) p_h^{n-1},$$

with $0 < \theta \leq 1$, where the auxiliary variables $(\mathbf{u}_{h,\theta}^n, p_{h,\theta}^n)$ satisfy the following discrete problem,

$$\begin{cases} \frac{1}{\theta\tau}(\mathbf{u}_{h,\theta}^n, \mathbf{v}_h)_\Omega + a_h^\nu(\mathbf{u}_{h,\theta}^n, \mathbf{v}_h) + b_h(p_{h,\theta}^n, \mathbf{v}_h) = F_{h,\theta}^{n,\nu}(\mathbf{v}_h) + \frac{1}{\theta\tau}(\mathbf{u}_h^{n-1}, \mathbf{v}_h)_\Omega, \ \forall \mathbf{v}_h \in \mathcal{V}_h^r, \\ b_h(q_h, \mathbf{u}_{h,\theta}^n) = G_{h,\theta}^n(q_h), \ \forall q_h \in Q_h^s, \end{cases} \tag{35}$$

being $F_{h,\theta}^{n,\nu}(\mathbf{v}_h) := \theta F_h^\nu(t^n, \mathbf{v}_h) + (1-\theta) F_h^\nu(t^{n-1}, \mathbf{v}_h)$, $G_{h,\theta}^n(q_h) := \theta G_h(t^n, q_h) + (1-\theta) G_h(t^{n-1}, q_h)$.

It is straightforward to see that problem (35) with $\theta \neq 0$ inherits all the fundamental stability and algebraic properties that we have previously proved for (11), with the exception of Lemma 3.16. This is shown introducing the bilinear form $a_h^{\nu,\eta}(\mathbf{u}_h, \mathbf{v}_h) := \eta(\mathbf{u}_h, \mathbf{v}_h)_\Omega + a_h^\nu(\mathbf{u}_h, \mathbf{v}_h)$ with $\eta \geq 0$ and $0 < \nu_0 \leq \nu$, which satisfies the following properties,

$$a_h^{\nu,\eta}(\mathbf{v}_h, \mathbf{v}_h) \gtrsim \nu_0 \|\mathbf{v}_h\|_{1,h,\Omega}^2, \quad a_h^{\nu,\eta}(\mathbf{u}_h, \mathbf{v}_h) \lesssim \max(\eta, \nu) \|\mathbf{u}_h\|_{1,h,\Omega} \|\mathbf{v}_h\|_{1,h,\Omega}.$$

Then, proceeding as in Corollary 3.15, matrix $A_{h,ij}^{\nu,\eta} := a_h^{\nu,\eta}(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i)$ is such that

$$\nu_0 \|\mathbf{v}_h\|_{1,h,\Omega} \lesssim \|A_h^{\nu,\eta} \mathbf{V}\|_{*,h,M_v}, \quad \|A_h^{\nu,\eta} \mathbf{V}\|_{*,M_v} \lesssim \max(\eta, \nu) h^{-\frac{1}{2}} \|\mathbf{v}_h\|_{1,h,\Omega}.$$

Let us now address for simplicity the typical case of time dependent blood flow problems, i.e $\eta > 1 > \nu$ such that $\max(\eta, \nu) = \eta$. Mimicking the proof of Lemma 3.16 we get,

$$\lambda_m \gtrsim \eta^{-2} \nu_0 h^{d+1}, \quad \lambda_M \lesssim \nu_0^{-1} h^d.$$

As a result of that, the time dependent pressure matrix $R_h^{\nu,\eta} := B_h (A_h^{\nu,\eta})^{-1} B_h^T$ is such that $\mathcal{X}(R_h^{\nu,\eta}) \lesssim \eta^2 \nu_0^{-2} h^{-1}$. In conclusion, when $\nu_0 \to 0$ or $\eta \to \infty$, the time dependent pressure matrix is ill conditioned and the contribution of the coefficient h^{-1}, arising from the penalty method, is of minor importance with respect to the weight of $\eta^2 \nu_0^{-2}$. For the specific preconditioning techniques to address this drawback, we remand to [17]. Since the penalized finite element matrices A_h, B_h enjoy spectral properties that are similar to the case of standard finite elements, see Lemma 3.14, we expect that the classical preconditioners for the time dependent Stokes problem, such as the Chaouet-Chabard matrix [18], could be also effective in the present case. Concerning the approximation properties, it is well known that (35) with the choice $\theta = \frac{1}{2}$ coincides with the Cranck-Nicholson time advancing scheme. By consequence, setting $r = 2$ and $s = 1$ for the space discretization, we conclude that our discrete scheme with $\theta = \frac{1}{2}$ is second order accurate with respect to both τ and h. In what follows, we will apply the present scheme for the approximation of a classical benchmark of computational hemodynamics, i.e. the pulsatile Womersley flow.

Table 1: Approximation errors of (11) with $r = 2$ and $s = 1$, i.e. $\mathbb{P}^2 - \mathbb{P}^1$ elements for velocities and pressures respectively.

h	$\|\|\|(\mathbf{u} - \mathbf{u}_h, p - p_h)\|\|\|_h$	order(h)	$\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}$	order(h)
0.1250	1.229440e-02	−−	2.084650e-04	−−
0.0833	5.394320e-03	2.032	6.448340e-05	2.894
0.0625	3.013920e-03	2.023	2.784980e-05	2.918
0.0500	1.919490e-03	2.022	1.447930e-05	2.931

4 Numerical results and applications

We address here the numerical validation of the stability, approximation and algebraic properties of (11). To this purpose we consider different reference problems whose solutions exactly satisfy the additional condition on the stresses, namely $p\mathbf{n} - \nabla\mathbf{u} \cdot \mathbf{n} = \mathbf{c}_k$. This is easily achieved selecting steady or time dependent flow problems that satisfy the following properties $p = p(t, x)$, $\mathbf{u}_x = \mathbf{u}_x(t, y)$, $\mathbf{u}_y = 0$ on rectangular domains. However, this may lead to the false impression that the defective boundary conditions can always exactly replace the information on the full inflow profile. This is not true and in general such conditions are inexact when the normal stresses are not constant over the inflow or outflow sections.

4.1 Validation of the approximation properties of the penalty method

In order to verify the results obtained in section 3.2, we consider problem (1) on the unit square $\Omega = (0, 1) \times (0, 1)$, where we impose defective conditions on the vertical sides, namely $\Gamma_1 = \{x = 0\} \times (0, 1)$ and $\Gamma_2 = \{x = 1\} \times (0, 1)$. Setting $\mathbf{U}_1 = \mathbf{U}_2 = [2/\pi, 0]$ and $\mathbf{f} = [\pi^2 \sin(\pi y), 0]$, we obtain that $\mathbf{u}(x, y) = [\sin(\pi y), 0]$, $p(x, y) = 0$ is an exact solution of (1). Concerning the choice of γ, it is shown in [19] that a convenient value is proportional to the constants of the inverse inequalities (13) that scale as r^2 on a shape regular mesh. To ensure a reasonable stability margin we set $\gamma = 4r^2 = 16$.

First, we verify that the numerical scheme satisfies estimates (19) and (21). The corresponding results are reported in table 1 and confirm that the theoretical order of convergence is closely respected. Additional numerical simulations show that the accuracy of the augmented formulation, i.e. problem (30), is almost equivalent. In fact, only negligible differences are perceived in the corresponding errors that for this reason are not reported.

4.2 Validation of the spectral properties and of the computational costs

First, we study the spectral properties of A_h and R_h. As previously mentioned, all the linear systems corresponding to formal matrix inversions are addressed by means of the CG method up to a tolerance 10^{-8} on the relative residuals. We remind that the asymptotic rate of convergence of the CG algorithm applied to a matrix M is inversely proportional to $\sqrt{\mathcal{X}(M)}$. By consequence, denoting with $\texttt{N.iter}(M^{-1})$ the number of iterations needed by the CG algorithm to converge up to the given tolerance, we have $\texttt{N.iter}(M^{-1}) \simeq \sqrt{\mathcal{X}(M)}$.

For the present numerical experiments, we address the Poiseuille flow on the unit square with defective conditions on the vertical sides $\Gamma_1 = \{x = 0\} \times (0, 1)$ and $\Gamma_2 = \{x = 1\} \times (0, 1)$. Indeed, setting $\mathbf{U}_1 = \mathbf{U}_2 = [1/6, 0]$ and $\mathbf{f} = \mathbf{0}$, problem (1) is satisfied by the non trivial solution $\mathbf{u}(x, y) = [y(1 - y), 0]$, $p = -2x + 1$ and $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = [-1, 0]$ when the augmented formulation is applied.

Owing to Lemmas 3.16 and 3.14, we expect that $\texttt{N.iter}(R_h^{-1}) \simeq h^{-1/2}$ and $\texttt{N.iter}(A_h^{-1}) \simeq h^{-1}$. These estimates are readily verified by the results of table 2. Furthermore, we notice that the heuristic approach to exploit the inexact Cholesky factorization of A, namely $P = HH^T$ computed with drop tolerance equal to 10^{-2}, as a preconditioner for A_h turns out to be very effective. Indeed, denoting with $C_h = H^{-1}A_hH^T$ the preconditioned matrix, we notice that $\texttt{N.iter}(C_h^{-1})$ is remarkably reduced with respect to $\texttt{N.iter}(A_h^{-1})$ and $\texttt{N.iter}(C_h^{-1})$ is also less sensitive with respect to h.

Then, we aim to analyze and quantify the computational cost of systems (25) and (32). To analyze system (25) we report in table 2 the number of CG iterations needed solve the Stokes Schur complement system relative to matrix $(B_hA_h^{-1}B_h^T)$. This invokes the multiplication by matrix A_h^{-1} a number of times (denoted with $\texttt{N.call}(A_h^{-1})$), which is performed again by means of CG, involving in average `N.iter` loops. Neglecting the fixed costs needed to compute the initial residuals and to recover the final velocity and pressure, the computational cost to solve (25), quantified by the corresponding CPU time, is proportional to the indicator

$$T = \texttt{N.call}(A_h^{-1}) \times \texttt{N.iter}(A_h^{-1}) \simeq 2\texttt{N.iter}((B_hA_h^{-1}B_h^T)^{-1}) \times \texttt{N.iter}(A_h^{-1}).$$

We proceed similarly for system (32). First, we report the number of CG iterations relative to $(DS^{-1}D^T)^{-1}$. As previously mentioned, this involves the multiplication by $(BA^{-1}B^T)^{-1}$, that is invoked (`N.call`) times, while the solution of the corresponding CG algorithm requires in average (`N.iter`) loops. As illustrated in the previous case, this is translated into the repeated multiplication by A^{-1}. As a consequence of that, the computational cost of (32) is quantified

Table 2: Computational costs for the solution of system (25) corresponding to the penalty method.

(25)	$\left(B_h A_h^{-1} B_h^T\right)^{-1}$	A_h^{-1}		
h	N.iter	N.call	N.iter	CPU(s)
0.25	7	17	27.06	0.350
0.125	20	43	54.88	3.380
0.0625	33	69	108.4	34.20
0.0312	39	81	216.5	276.9
(25)	$\left(B_h A_h^{-1} B_h^T\right)^{-1}$	C_h^{-1}		
h	N.iter	N.call	N.iter	CPU(s)
0.25	7	17	8.41	0.33
0.125	22	47	12.64	1.59
0.0625	34	71	17.63	9.81
0.0312	40	83	33.17	71.89

by,

$$
\begin{aligned}
T &= \texttt{N.call}(A^{-1}) \times \texttt{N.iter}(A^{-1}) \\
&\simeq \left(\texttt{N.call}((BA^{-1}B^T)^{-1}) \times 2\texttt{N.iter}((BA^{-1}B^T)^{-1})\right) \times \texttt{N.iter}(A^{-1}) \\
&\simeq \left(2\texttt{N.iter}((DS^{-1}D^T)^{-1}) \times 2\texttt{N.iter}((BA^{-1}B^T)^{-1})\right) \times \texttt{N.iter}(A^{-1})
\end{aligned}
$$

which is expected to be directly related to the CPU time needed to solve (32). Under the assumption that the additional costs introduced by the penalty technique are negligible, this analysis anticipates that the augmented formulation is less efficient than the penalty method, because the former requires to solve the Stokes system, namely to multiply by $(BA^{-1}B^T)^{-1}$, two times for each iteration of the CG algorithm applied to $(DS^{-1}D^T)^{-1}$. In conclusion, the more iterations are needed to determine the multipliers $\mathbf{\Lambda}$, the more problem (32) is inefficient with respect to (25).

We compare the penalty with the augmented formulations in tables 2 and 3. As expected, the penalty formulation is considerably more convenient than the augmented one. Indeed, the gain in terms of CPU times (reported in seconds) can be quantified by a factor 10. This can be interpreted observing that the augmented formulation requires 3 CG iterations to solve the 4×4 system for the multipliers and it involves 9 calls to the Stokes Schur complement. Surprisingly, we also notice that the subproblems related to the augmented formulation seem to be more stiff than the ones in the penalty case. In particular, the average number of CG iterations needed to multiply by A^{-1} is considerably higher than in the case of A_h^{-1}. However, this contribution has a minor effect on the costs needed to determine the multipliers. This is confirmed by the additional tests where we apply the aforementioned incomplete Cholesky factorization of A, namely $P = HH^T$, as a preconditioner for A and A_h. This is formally equivalent to replace A^{-1} with $C^{-1} = (H^{-1}AH^{-T})^{-1}$ and A_h^{-1} with $C_h^{-1} = (H^{-1}A_hH^{-T})^{-1}$. The

Table 3: Computational costs for the solution of system (32) corresponding to the augmented formulation.

(32)	$\left(DS^{-1}D^T\right)^{-1}$	$\left(BA^{-1}B^T\right)^{-1}$		A^{-1}		
h	N.iter	N.call	N.iter	N.call	N.iter	CPU(s)
0.25	3	9	16.33	320	41.37	5.53
0.125	3	9	28.67	542	79.80	43.96
0.0625	3	9	33.89	636	152.5	318.9
0.0312	3	9	36.56	684	294.26	2188.8
(32)	$\left(DS^{-1}D^T\right)^{-1}$	$\left(BA^{-1}B^T\right)^{-1}$		C^{-1}		
h	N.iter	N.call	N.iter	N.call	N.iter	CPU(s)
0.25	3	9	16.56	324	5.81	2.40
0.125	3	9	28.56	540	8.83	10.66
0.0625	3	9	33.78	634	15.42	62.25
0.0312	3	9	36.00	674	28.77	410.82

corresponding results are reported on the bottom of tables 2 and 3. We notice that the preconditioner P is effective to both cases, but the computational cost associated to C^{-1} is now smaller than the one of C_h^{-1}. Nevertheless, the penalty method is still considerably more convenient than the augmented formulation. Indeed, our conclusions are confirmed.

4.3 Application to the transient case. Simulation of the Womersley flow.

The Womersley flow, i.e. the flow in a two dimensional or cylindrical channel with a uniform but oscillating pressure gradient, is a simple yet effective benchmark for computational hemodynamics, addressed for instance in [2, 3, 4]. In such conditions, it is possible to retrieve the exact solution of the time dependent Stokes or Navier-Stokes equations by separation of variables, see [20].

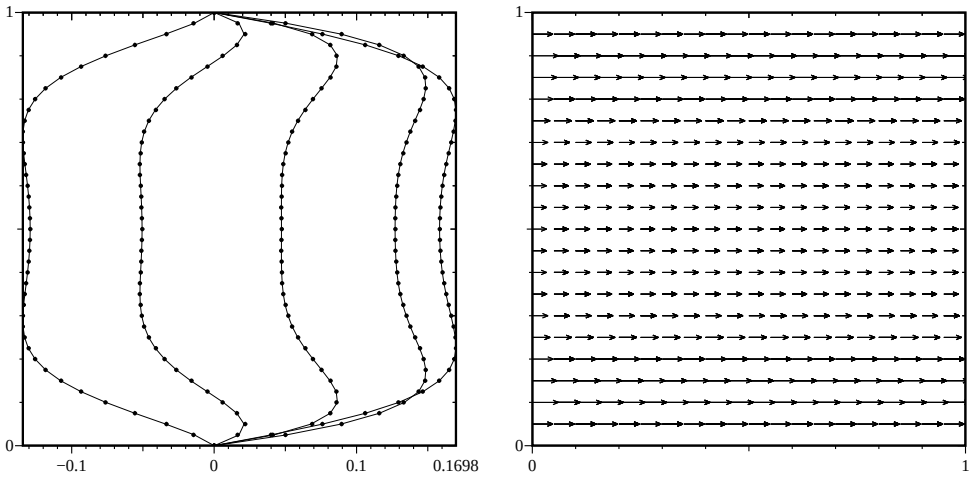
We address here the two dimensional case and we consider the unit square, $\Omega = (0,1) \times (0,1)$, with inflow and outflow sections $\Gamma_1 = \{x = 0\} \times (0,1)$ and $\Gamma_2 = \{x = 1\} \times (0,1)$, respectively. First, we notice that the classical periodic Womersley flow, adapted here to the unit square, i.e.

$$p(t,x,y) = (1-2x)\sin(\omega t), \quad \mathbf{u}_x(t,x,y) = \sum_{i=0}^{\infty} \gamma_{2i+1}(t)\sin((2i+1)\pi y), \quad \mathbf{u}_y = 0,$$

$$\gamma_i = \frac{4}{(\pi i)(\nu^2\pi^4 i^4 + \omega^2)}\left(\nu\pi^2 i^2 \sin(\omega t) - \omega\cos(\omega t)\right),$$

is a solution of the time dependent Stokes problem with prescribed inflow and outflow profiles. Exploiting the properties $p = p(t,x)$, $\mathbf{u}_x = \mathbf{u}_x(t,y)$, $\mathbf{u}_y = 0$, we can easily see that it is also a solution of (33), provided that the flow rates $\mathbf{U}_1(t) = \mathbf{U}_2(t)$ are defined accordingly.

Figure 2: Numerical simulation of the Womersley flow (solid line) compared with the exact solution (dots). The computed axial velocity profiles are reported at different times $t = 0.1, 0.2, 0.3, 0.4, 0.5$ on the left. On the right, we show the velocity field at time $t = 0.3$.



For the numerical experiments we set $\nu = 3 \times 10^{-2}$ and $\omega = 2\pi$ into the Womersley solution, such that the viscosity and the oscillation period are similar to the ones of blood flow. As initial condition for the numerical simulation we prescribe the finite element interpolate of the exact solution at $t = 0$ and we set $h = \tau = 0.05$. In figure 2 we report the axial component of the computed solution at the inflow $x = 0$ for different times (solid lines) compared with the exact solution evaluated on the finite element nodes (dots). On the left, we show the complete velocity field at time $t = 0.3$. The tangential component of the flow, namely $\mathbf{u}_y$, turns out to be equal to zero up to the computational tolerance and the numerical solution turns out to be independent of the axial coordinate, i.e. the velocity profiles do not change at different axial locations. Such results confirm the efficacy of the penalty method in the approximation of defective net flux conditions.

5 Concluding remarks

We have shown that the application of penalty techniques turns out to be an effective strategy for the discretization of the Stokes system with defective boundary conditions. Indeed, such technique allows us to set up a finite element discretization of the problem at hand with negligible additional computational costs with respect to the standard case of Dirichlet boundary conditions. Simultaneously, the accuracy of the selected finite elements is not affected. The proposed method can also be straightforwardly extended to more significant models such

as the time dependent Oseen or the Navier-Stokes equations, with interesting applications in the field of computational hemodynamics.

References

[1] J. G. Heywood, R. Rannacher, S. Turek, Artificial boundaries and flux and pressure conditions for the incompressible Navier-Stokes equations, Internat. J. Numer. Methods Fluids 22 (5) (1996) 325–352.

[2] L. Formaggia, J.-F. Gerbeau, F. Nobile, A. Quarteroni, Numerical treatment of defective boundary conditions for the Navier-Stokes equations, SIAM J. Numer. Anal. 40 (1) (2002) 376–401 (electronic).

[3] A. Veneziani, C. Vergara, Flow rate defective boundary conditions in haemodynamics simulations, Internat. J. Numer. Methods Fluids 47 (8-9) (2005) 803–816.

[4] A. Veneziani, C. Vergara, An approximate method for solving incompressible Navier-Stokes problems with flow rate conditions, Comput. Methods Appl. Mech. Engrg. 196 (9-12) (2007) 1685–1700.

[5] I. E. Vignon-Clementel, C. A. Figueroa, K. E. Jansen, C. A. Taylor, Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries, Comput. Methods Appl. Mech. Engrg. 195 (29-32) (2006) 3776–3796.

[6] L. Formaggia, A. Veneziani, C. Vergara, A new approach to numerical solution of defective boundary value problems in incompressible fluid dynamics, SIAM J. Numer. Anal. 46 (6) (2008) 2769–2794.

[7] I. Babuška, The finite element method with Lagrangian multipliers, Numer. Math. 20 (1972/73) 179–192.

[8] J. Nitsche, Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, Abh. Math. Sem. Univ. Hamburg 36 (1971) 9–15, collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[9] I. Babuška, The finite element method with penalty, Math. Comp. 27 (1973) 221–228.

[10] R. Stenberg, On some techniques for approximating boundary conditions in the finite element method, J. Comput. Appl. Math. 63 (1-3) (1995) 139–148, international Symposium on Mathematical Modelling and Computational Methods Modelling 94 (Prague, 1994).

[11] B. Rivière, M. F. Wheeler, V. Girault, A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems, SIAM J. Numer. Anal. 39 (3) (2001) 902–931 (electronic).

[12] V. Thomée, Galerkin finite element methods for parabolic problems, Vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1997.

[13] V. Girault, P.-A. Raviart, Finite element methods for Navier-Stokes equations, Vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986, theory and algorithms.

[14] F. Brezzi, M. Fortin, Mixed and hybrid finite element methods, Vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[15] A. Ern, J.-L. Guermond, Theory and practice of finite elements, Vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.

[16] A. Quarteroni, A. Valli, Numerical approximation of partial differential equations, Vol. 23 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1994.

[17] H. C. Elman, D. J. Silvester, A. J. Wathen, Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.

[18] J. Cahouet, J.-P. Chabard, Some fast 3D finite element solvers for the generalized Stokes problem, Internat. J. Numer. Methods Fluids 8 (8) (1988) 869–895.

[19] P. Hansbo, M. Larson, Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method, Comput. Methods Appl. Mech. Engrg. 191 (17-18) (2002) 1895–1908.

[20] J. R. Womersley, Oscillatory motion of a viscous liquid in a thin-walled elastic tube. I. The linear approximation for long waves, Phil. Mag. (7) 46 (1955) 199–221.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

10/2009 P. Zunino:
Numerical approximation of incompressible flows with net flux defective boundary conditions by means of penalty techniques

09/2009 E. Agostoni, S. Salsa, M. Perego, A. Veneziani:
Mathematical and Numerical Modeling of Focal Cerebral Ischemia

08/2009 P.F.Antonietti, P.Houston:
An hr-adaptive discontinuous Galerkin method for advection-diffusion problems

07/2009 M. Perego, A. Veneziani:
An efficient generalization of the Rust-Larsen method for solving electrophysiology membrane equations

06/2009 L. Formaggia, A. Veneziani, C. Vergara:
Numerical solution of flow rate boundary for incompressible fluid in deformable domains

05/2009 F. Ieva, A.M. Paganoni:
A case study on treatment times in patients with ST-Segment Elevation Myocardial Infarction

04/2009 C. Canuto, P. Gervasio, A. Quarteroni:
Finite-Element Preconditioning of G-NI Spectral Methods

03/2009 M. D'Elia, L. Dedé, A. Quarteroni:
Reduced Basis Method for Parametrized Differential Algebraic Equations

02/2009 L. Bonaventura, C. Biotto, A. Decoene, L. Mari, E. Miglio:
A couple ecological-hydrodynamic model for the spatial distribution of sessile aquatic species in thermally forced basins

01/2009 E. Miglio, C. Sgarra:
A Finite Element Framework for Option Pricing the Bates Model