THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

A stabilised finite element method for the convection-diffusion-reaction equation in mixed form

Citation for published version:
Barrenechea, GR, Poza, AH & Yorston, H 2018, 'A stabilised finite element method for the convection-diffusion-reaction equation in mixed form', *Computer Methods in Applied Mechanics and Engineering*, vol. 339, pp. 389-415. https://doi.org/10.1016/j.cma.2018.04.019

Digital Object Identifier (DOI):
10.1016/j.cma.2018.04.019

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Computer Methods in Applied Mechanics and Engineering

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 17. Aug. 2021

A STABILISED FINITE ELEMENT METHOD FOR THE CONVECTION-DIFFUSION-REACTION EQUATION IN MIXED FORM

GABRIEL R. BARRENECHEA*, ABNER H. POZA†, AND HEATHER YORSTON‡

Abstract. This paper is devoted to the approximation of the convection-diffusion-reaction equation using a mixed, first-order, formulation. We propose, and analyse, a stabilised finite element method that allows equal order interpolations for the primal and dual variables. This formulation, reminiscent of the Galerkin least-squares method, is proven stable and convergent. In addition, a numerical assessment of the numerical performance of different stabilised finite element methods for the mixed formulation is carried out, and the different methods are compared in terms of accuracy, stability, and sharpness of the layers for two different classical test problems.

1. Introduction. Despite the large amount of work that has been devoted to the numerical approximation of convection dominated problems, there is still the open question of finding a method that 'ticks all the boxes'. By this, we mean a method that provides stable results while not smearing the sharp layers appearing in the solution. For example, the SUPG method (cf. [10, 34]) has been accepted as an efficient method that produces sharp layers, but at the cost of producing over- and undershoots in the regions close to them. In order to avoid these non-physical oscilations, several methods have been proposed over the years, including Continuous Interior Penalty (e.g. [11]), LPS methods (e.g. [28]), or using shock-capturing related ideas (see, e.g., [26, 27] for a review, and [22, 2, 4, 3] for more recent developments). Several alternatives were compared in the relatively recent paper [1], and the conclusion was that, up to that date, no method could be considered to be completely satisfactory.

Alternatively, some attempts have been made to approximate this problem by first rewriting it as a first-order system. To the best of our knowledge, the first papers that addressed this possibilty were [16, 17]. Different first-order formulations were tried in these papers, and the discretisation was carried out by means of Raviart-Thomas finite element methods (cf. [33]). Nevertheless, two issues remain that are not covered by those papers. Firstly, the numerical stability of the resulting scheme was only proven when the mesh discretisation parameter was small enough, which limits the applicability of such a discretisation to the diffusion-dominated case. Secondly, since the discretisation did not include any form of stabilisation, the same instabilities from the plain Galerkin scheme are to be expected for this mixed method. With the aim of addressing that issue, in [35] the author proposes a new method, which also uses Raviart-Thomas spaces, but adds an upwind-based stabilisation. Nevertheless, the resulting method is only applicable to higher order discretisations. A more modern approach, including a posteriori error estimation and different choices for finite element spaces, can be found in [14].

Several works have tried to address the points raised in the previous paragraph. For example, one possibility is to consider a least-squares method, such as FOSLS. This

*Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, Scotland (gabriel.barrenechea@strath.ac.uk).

†Departamento de Matemática y Física Aplicadas, Facultad de Ingeniería, Universidad Católica de la Santísima Concepción, Casilla 297, Concepción, Chile (apoza@ucsc.cl). This author was partially funded by Dirección de Investigación e Innovación of the Universidad Católica de la Santísima Concepción through project DINREG 04/2017.

‡Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, Scotland (heather.yorston@strath.ac.uk).

leads to an elliptic problem, thus freeing the choice of the finite element spaces, see, e.g., [12, 13, 19] and the references therein, or [6] for more general least-squares methods and an extensive review. One disadvantage of this sort of approach is that it leads to fairly diffusive layers, thus, again, making its interest for convection-dominated problems limited; see [24] for a discussion on this issue, [25] for the possibility of using a FOSLS method combined with an enrichment of the finite element space with bubble functions, or [29] for a streamline-based FOSLS method. To address this issue, in [15] a weighted FOSLS method was proposed, combined with a weak imposition of the boundary conditions. Alternatively, some finite volume-inspired methods have been proposed in conjunction with Raviart-Thomas elements (see, e.g. [7]). However, their performance for problems that contain strong layers is still to be tested. Other approaches to stabilise this mixed problem include the hybridized discontinuous Galerkin methods (see, e.g., [32]), the discontinuous Petrov-Galerkin method with optimal test functions (see, e.g. [9, 8]), and augmented formulations (see, e.g., [21, 5]). It is interesting to remark that almost none of the references just quoted use Lagrangian elements for both variables. In fact, several of them make use of the Raviart-Thomas' space for the vector variable, even in the case the final formulation is driven by an elliptic bilinear form.

In this work we pursue a different approach. Our interest is to approximate the convection-diffusion-reaction equation using a mixed, first-order formulation, but using standard Lagrangian elements in both variables. Thus, stabilisation is needed in order to prove stability and convergence. As far as we are aware, the only method that has been proposed with this purpose is the one presented in [31], which is a modification of the method proposed in [30] for the Darcy equation. In the work [31] no stability, or error estimates, are proven. Our first aim is to bridge this gap. In the process of trying to analyse the method from [31], the need to modify its definition appeared. Thus, in this work we propose a new stabilised mixed finite element method for the first order writing of the convection-diffusion-reaction equation, which can be proven to be stable and convergent. To assess the performance of the new method, we have also carried out intensive comparisons with several previously existing alternative methods. More precisely, by means of two standard test cases for the convection-diffusion equation we have compared the new method to the original method from [31] and two variants of the FOSLS approach. As a reference, we have also considered the results provided by the SUPG method.

The rest of this manuscript is organised as follows. In Section 2 the problem of interest and the main notations are introduced. The stabilised finite element method is presented in Section 3, and its stability is proven. In Section 4 error estimates are shown, and these are corroborated numerically in Section 5. In Section 6 some alternative finite element methods for the mixed formulation of the convection-diffusion equation are reviewed, and then a detailed comparison of the performance of these alternatives with the present approach is given.

2. Notation and preliminaries. We consider $\Omega \subseteq \mathbb{R}^d, d = 2, 3$, an open, bounded, polyhedral domain with Lipschitz boundary Γ. Standard notations for Sobolev spaces and their corresponding norms are used throughout. For $D \subseteq \Omega$, the inner product in $L^2(D)$, or $L^2(D)^d$, is denoted by $(\cdot,\cdot)_D$. In the case $D = \Omega$ the subscript will be dropped. The norm and semi-norm in $W^{m,p}(D)$ will be denoted by $\|\cdot\|_{m,p,D}$ and $|\cdot|_{m,p,D}$, respectively, with the convention $\|\cdot\|_{m,D} = \|\cdot\|_{m,2,D}$, where $H^m(D) = W^{m,2}(D)$ and $L^2(D) = H^0(D)$. We also introduce the subspace of $L^2(\Omega)^d$:

$$H(\mathrm{div}\,;\Omega) = \left\{ \boldsymbol{w} \in L^2(\Omega)^d : \nabla \cdot \boldsymbol{w} \in L^2(\Omega) \right\}.$$

Our problem of interest is the following convection–diffusion–reaction equation: find p such that

$$\left\{ \begin{array}{rcll} -\varepsilon\Delta p \,+\, \boldsymbol{\alpha}\cdot\nabla p + \mu\, p & = & f & \text{in } \Omega, \\ p & = & 0 & \text{on } \Gamma, \end{array} \right. \tag{2.1}$$

where $\boldsymbol{\alpha} \in W^{1,\infty}(\Omega)^d$ is a convective field such that $\nabla\cdot\boldsymbol{\alpha} = 0$ in Ω, $\varepsilon > 0$ is a diffusion coefficient, $\mu > 0$ is a reaction coefficient, and $f \in L^2(\Omega)$.
To rewrite (2.1) as a first-order mixed problem we define the total flux by $\boldsymbol{v} := -\varepsilon\nabla p + \boldsymbol{\alpha} p$ as an independent variable, so that (2.1) becomes

$$\left\{ \begin{array}{rl} \dfrac{1}{\varepsilon}\,\boldsymbol{v} + \nabla p - \dfrac{1}{\varepsilon}\,\boldsymbol{\alpha} p = \boldsymbol{0} & \text{in } \Omega, \\ \nabla\cdot\boldsymbol{v} + \mu\, p = f & \text{in } \Omega, \\ p = 0 & \text{on } \Gamma. \end{array} \right. \tag{2.2}$$

Following a standard approach, we obtain the following weak formulation of (2.2): find $(\boldsymbol{v}, p) \in \boldsymbol{V}\times Q := H(\mathrm{div}\,;\Omega)\times L^2(\Omega)$ such that

$$\frac{1}{\varepsilon}\,(\boldsymbol{v},\boldsymbol{w}) - (p, \nabla\cdot\boldsymbol{w}) - \frac{1}{\varepsilon}\,(\boldsymbol{\alpha}\, p, \boldsymbol{w}) + (\nabla\cdot\boldsymbol{v}, q) + \mu\,(p,q) = (f,q)\,, \tag{2.3}$$

for all $(\boldsymbol{w}, q) \in \boldsymbol{V}\times Q$.
Remark 1. *An alternative formulation arises if, instead of the total flux, the diffusive flux $\boldsymbol{v}_d = -\varepsilon\nabla p$ is introduced as an extra unknown. This has been done in [16, 35, 15]. In this case the first-order system for* (2.1) *becomes*

$$\left\{ \begin{array}{rl} \dfrac{1}{\varepsilon}\,\boldsymbol{v}_d + \nabla p = \boldsymbol{0} & \textit{in } \Omega, \\ \nabla\cdot\boldsymbol{v}_d + \nabla p\cdot\boldsymbol{\alpha} + \mu\, p = f & \textit{in } \Omega, \\ p = 0 & \textit{on } \Gamma. \end{array} \right. \tag{2.4}$$

The weak variational form for (2.4) *reads: find $(\boldsymbol{v}_d, p) \in \boldsymbol{V}\times Q$ such that*

$$\frac{1}{\varepsilon}\,(\boldsymbol{v}_d,\boldsymbol{w}) - (p, \nabla\cdot\boldsymbol{w}) + (\nabla\cdot\boldsymbol{v}_d, q) - \frac{1}{\varepsilon}\,(\boldsymbol{\alpha}\cdot\boldsymbol{v}_d, q) + \mu\,(p,q) = (f,q)\,, \tag{2.5}$$

for all $(\boldsymbol{w}, q) \in \boldsymbol{V}\times Q$.

Remark 2. *Using the Lax-Milgram Lemma, it can be proven that* (2.1) *has a unique weak solution $p \in H^1_0(\Omega)$; see [18] for details. Thus, the existence and uniqueness of solution of the problem* (2.3)*, or* (2.5)*, follows from the fact that a solution of either of these problems is a weak solution of* (2.1)*, and vice–versa.*
Remark 3. *The restriction imposed on $\boldsymbol{\alpha}$, namely, its solenoidal character, appears to make the derivation of* (2.3) *from* (2.1) *clearer. The introduction of the diffusive flux $\boldsymbol{v}_d$ leading to* (2.5)*, does not need this restriction, which has made some authors (especially the ones that have presented the FOSLS methods introduced in Section 6.1.3 below) favor the latter alternative. Nevertheless, in order to show the existence of solutions both formulations need the restriction $-\frac{\nabla\cdot\boldsymbol{\alpha}}{2} + \mu \geq 0$, since under this condition,* (2.1) *can be proven to have one weak solution. In addition, it is important to remark that this restriction does not play any role in the proposal and analysis of the stabilised finite element method presented in the next section.*

Let $\{\mathcal{T}_h\}_{h>0}$ be a family of regular triangulations of Ω, built up using simplices T with diameter $h_T := \operatorname{diam}(T)$, and $h := \max\{h_T : T \in \mathcal{T}_h\}$. For a polynomial order $k \geq 1$, we introduce the finite element space for the flux variable as

$$\boldsymbol{H}_h := \left\{\boldsymbol{\varphi} \in C^0(\overline{\Omega})^d : \boldsymbol{\varphi}|_T \in \mathcal{P}_k(T)^d \quad \forall T \in \mathcal{T}_h\right\}, \tag{2.6}$$

and the discrete subspace for the scalar variable p as

$$Q_h^0 := Q_h \cap H_0^1(\Omega) \ \text{ where } \ Q_h := \left\{q_h \in C^0(\overline{\Omega}) : q_h|_T \in \mathcal{P}_k(T)\ ,\ \forall T \in \mathcal{T}_h\right\}. \tag{2.7}$$

We denote by $\boldsymbol{\Pi}_h$ the L^2–orthogonal projection onto $\boldsymbol{H}_h$ defined by

$$(\boldsymbol{\Pi}_h(\boldsymbol{v}), \boldsymbol{w}_h) = (\boldsymbol{v}, \boldsymbol{w}_h) \qquad \forall \boldsymbol{w}_h \in \boldsymbol{H}_h. \tag{2.8}$$

We will need the following properties of this operator in the sequel.

LEMMA 2.1. *There exists a positive constant C, independent of h, such that*

$$\|\boldsymbol{\Pi}_h(\boldsymbol{v})\|_{0,\Omega} \leq \|\boldsymbol{v}\|_{0,\Omega} \quad \forall \boldsymbol{v} \in L^2(\Omega)^d, \tag{2.9}$$

$$\|\boldsymbol{v} - \boldsymbol{\Pi}_h(\boldsymbol{v})\|_{0,\Omega} \leq C\, h\, |\boldsymbol{v}|_{1,\Omega} \quad \forall \boldsymbol{v} \in H^1(\Omega)^d. \tag{2.10}$$

Proof. See Lemma 1.131 in [18]. □

We finally recall the following inverse inequality, which will be used throughout, and whose proof is a direct consequence of classical inverse inequalities for polynomial functions (see, e.g., [18, Lemma 1.138]): There exists $C_k > 0$, depending only on k and the regularity of the mesh, such that, for all $\boldsymbol{w}_h \in \boldsymbol{H}_h$:

$$h_T\, \|\nabla \cdot \boldsymbol{w}_h\|_{0,T} \leq C_k \|\boldsymbol{w}_h\|_{0,T} \qquad \forall\, T \in \mathcal{T}_h\,. \tag{2.11}$$

3. The stabilised finite element method. As mentioned in the introduction, our method is a modification of the one from [31] (see Section 6.1.1 later for details). More precisely, the stabilised finite element method studied in this work reads: find $(\boldsymbol{v}_h, p_h) \in \boldsymbol{H}_h \times Q_h^0$ such that

$$B((\boldsymbol{v}_h, p_h), (\boldsymbol{w}_h, q_h)) = (f, q_h) + \sum_{T \in \mathcal{T}_h} \delta_{div}^T (f, \nabla \cdot \boldsymbol{w}_h + \mu\, q_h)_T, \tag{3.1}$$

for all $(\boldsymbol{w}_h, q_h) \in \boldsymbol{H}_h \times Q_h^0$, where the bilinear form $B(\cdot,\cdot)$ is given by

$$\begin{aligned} & B((\boldsymbol{v}_h, p_h), (\boldsymbol{w}_h, q_h)) \\ := & \frac{1}{\varepsilon}\,(\boldsymbol{v}_h, \boldsymbol{w}_h) - (p_h, \nabla \cdot \boldsymbol{w}_h) + (\nabla \cdot \boldsymbol{v}_h, q_h) - \frac{1}{\varepsilon}\,(\boldsymbol{\alpha}\, p_h, \boldsymbol{w}_h) + \mu\,(p_h, q_h) \\ & - \frac{\varepsilon}{2}\left(\frac{1}{\varepsilon}\,\boldsymbol{v}_h + \nabla p_h - \frac{1}{\varepsilon}\,\boldsymbol{\alpha}\, p_h, \frac{1}{\varepsilon}\,\boldsymbol{w}_h - \nabla q_h + \frac{1}{\varepsilon}\,\boldsymbol{\alpha}\, q_h\right) \\ & + \sum_{T \in \mathcal{T}_h} \delta_{div}^T (\nabla \cdot \boldsymbol{v}_h + \mu\, p_h, \nabla \cdot \boldsymbol{w}_h + \mu\, q_h)_T, \end{aligned} \tag{3.2}$$

and the stabilisation parameter δ_{div} is defined as

$$\delta_{div}^T := \delta \min\left\{h_T, \frac{h_T^2}{4\varepsilon}\right\} \quad \text{where } \delta > 0 \text{ is arbitrary.} \tag{3.3}$$

In what follows we will denote $\delta_{div} := \max_{T\in\mathcal{T}_h} \delta_{div}^T$.

REMARK 4. *Although of similar shape, Method* (3.1) *and Masud-Kwack's method [31] contain significant differences. The first is the addition of the convective term in the test function for the stabilising term. This is added to make the analysis possible (in fact, to the best of our knowledge, there is no analysis for the original method from [31]). Moreover, the div-div term added to the formulation improves the numerical results significantly.*

The stability and error analysis will be carried out using the following mesh-dependent norm:

$$\|(\boldsymbol{w},q)\|_h := \left\{ \frac{1}{\varepsilon}\,\|\boldsymbol{w}-\boldsymbol{\Pi}_h(\boldsymbol{\alpha}\,q)\|_{0,\Omega}^2 + \varepsilon\,|q|_{1,\Omega}^2 + \mu\,\|q\|_{0,\Omega}^2 + \sum_{T\in\mathcal{T}_h} \delta_{div}^T\,\|\nabla\cdot\boldsymbol{w}+\mu\,q\|_{0,T}^2 \right\}^{1/2}. \tag{3.4}$$

Using this norm, we present the main result about stability of the method.

THEOREM 3.1. *Let $B(\cdot,\cdot)$ be the bilinear form given by* (3.2)*. Then, there exists a positive constant C, independent of ε,μ,h, and $\boldsymbol{\alpha}$, such that*

$$\sup_{(\boldsymbol{w}_h,q_h)\in\boldsymbol{H}_h\times Q_h^0} \frac{B\left((\boldsymbol{v}_h,p_h),(\boldsymbol{w}_h,q_h)\right)}{\|(\boldsymbol{w}_h,q_h)\|_h} \geq C\,\|(\boldsymbol{v}_h,p_h)\|_h, \tag{3.5}$$

for all $(\boldsymbol{v}_h,p_h)\in\boldsymbol{H}_h\times Q_h^0$. Thus, (3.1) *is well-posed.*

Proof. Let $(\boldsymbol{v}_h,p_h)\in\boldsymbol{H}_h\times Q_h^0$. First, using the definition of $B(\cdot,\cdot)$, and Cauchy–Schwarz and Young inequalities we arrive at

$$\begin{aligned} B((\boldsymbol{v}_h,p_h),(\boldsymbol{v}_h,p_h)) =& \frac{1}{2\varepsilon}\,\|\boldsymbol{v}_h\|_{0,\Omega}^2 - \frac{1}{\varepsilon}\,(\boldsymbol{\alpha}\,p_h,\boldsymbol{v}_h) + \frac{\varepsilon}{2}\,|p_h|_{1,\Omega}^2 + \frac{1}{2\varepsilon}\,\|\boldsymbol{\alpha}\,p_h\|_{0,\Omega}^2 \\ &+ \mu\,\|p_h\|_{0,\Omega}^2 + \sum_{T\in\mathcal{T}_h} \delta_{div}^T\,\|\nabla\cdot\boldsymbol{v}_h+\mu\,p_h\|_{0,T}^2 \\ \geq& \frac{\varepsilon}{2}\,|p_h|_{1,\Omega}^2 + \mu\,\|p_h\|_{0,\Omega}^2 + \sum_{T\in\mathcal{T}_h} \delta_{div}^T\,\|\nabla\cdot\boldsymbol{v}_h+\mu\,p_h\|_{0,T}^2\,. \end{aligned} \tag{3.6}$$

Let now $\boldsymbol{w}_h\in\boldsymbol{H}_h\times Q_h^0$. The definition of $B(\cdot,\cdot)$ and integration by parts give

$$\begin{aligned} & B((\boldsymbol{v}_h,p_h),(\boldsymbol{w}_h,0)) \\ =& \frac{1}{\varepsilon}\,(\boldsymbol{v}_h,\boldsymbol{w}_h) - (p_h,\nabla\cdot\boldsymbol{w}_h) - \frac{1}{\varepsilon}\,(\boldsymbol{\alpha}\,p_h,\boldsymbol{w}_h) - \frac{\varepsilon}{2}\left(\frac{1}{\varepsilon}\,\boldsymbol{v}_h + \nabla p_h - \frac{1}{\varepsilon}\,\boldsymbol{\alpha}\,p_h, \frac{1}{\varepsilon}\,\boldsymbol{w}_h\right) \\ &+ \sum_{T\in\mathcal{T}_h} \delta_{div}^T\,(\nabla\cdot\boldsymbol{v}_h+\mu\,p_h,\nabla\cdot\boldsymbol{w}_h)_T \\ =& \frac{1}{2\varepsilon}\,(\boldsymbol{v}_h-\boldsymbol{\alpha}\,p_h,\boldsymbol{w}_h) + \frac{1}{2}\,(\nabla p_h,\boldsymbol{w}_h) + \sum_{T\in\mathcal{T}_h} \delta_{div}^T\,(\nabla\cdot\boldsymbol{v}_h+\mu\,p_h,\nabla\cdot\boldsymbol{w}_h)_T\,. \end{aligned}$$

Thus, using (3.3), (2.11), taking $\tilde{\boldsymbol{w}}_h := \boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\,p_h)$, and using the Cauchy-Schwarz,

Young, and inverse inequalities we obtain

$$
\begin{aligned}
&B((\boldsymbol{v}_h, p_h), (\tilde{\boldsymbol{w}}_h, 0)) \\
=&\frac{1}{2\varepsilon}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega}^2 + \frac{1}{2}\left(\nabla p_h, \boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right) \\
&\quad + \sum_{T\in\mathcal{T}_h} \delta_{div}^T \left(\nabla\cdot\boldsymbol{v}_h + \mu\, p_h, \nabla\cdot(\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h))\right)_T \\
\geq&\frac{1}{4\varepsilon}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega}^2 - \frac{\varepsilon}{4}|p_h|_{1,\Omega}^2 \\
&\quad - \sum_{T\in\mathcal{T}_h}\left\{\frac{\delta_{div}^T C_k^2\delta}{2}\left\|\nabla\cdot\boldsymbol{v}_h + \mu\, p_h\right\|_{0,T}^2 + \frac{\delta_{div}^T}{2C_k^2\delta}\left\|\nabla\cdot(\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h))\right\|_{0,T}^2\right\} \\
\geq&\frac{1}{4\varepsilon}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega}^2 - \frac{\varepsilon}{4}|p_h|_{1,\Omega}^2 \\
&\quad - \sum_{T\in\mathcal{T}_h}\left\{\frac{\delta_{div}^T C_k^2\delta}{2}\left\|\nabla\cdot\boldsymbol{v}_h + \mu\, p_h\right\|_{0,T}^2 + \frac{\delta_{div}^T}{2\delta h_T^2}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,T}^2\right\} \\
\geq&\frac{1}{8\varepsilon}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega}^2 - \frac{\varepsilon}{4}|p_h|_{1,\Omega}^2 - \sum_{T\in\mathcal{T}_h}\frac{\delta_{div}^T C_k^2\delta}{2}\left\|\nabla\cdot\boldsymbol{v}_h + \mu\, p_h\right\|_{0,T}^2. \qquad (3.7)
\end{aligned}
$$

Adding (3.6) and (3.7), and defining $\gamma := \min\{1, (\delta C_k^2)^{-1}\}$, the following holds

$$
\begin{aligned}
&B\left((\boldsymbol{v}_h, p_h), (\boldsymbol{v}_h + \gamma\tilde{\boldsymbol{w}}_h, p_h)\right) \\
&\geq \frac{\varepsilon(4-\gamma)}{8}|p_h|_{1,\Omega}^2 + \mu\left\|p_h\right\|_{0,\Omega}^2 + \frac{\gamma}{8\varepsilon}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega}^2 \\
&\quad + \sum_{T\in\mathcal{T}_h}\delta_{div}^T\left(1 - \frac{C_k^2\delta\gamma}{2}\right)\left\|\nabla\cdot\boldsymbol{v}_h + \mu\, p_h\right\|_{0,T}^2 \\
&\geq C\left\|(\boldsymbol{v}_h, p_h)\right\|_h^2. \qquad (3.8)
\end{aligned}
$$

Finally, from (3.3), (2.11), and using that $\gamma \leq 1$, it follows that

$$
\begin{aligned}
\left\|(\boldsymbol{v}_h + \gamma\tilde{\boldsymbol{w}}_h, p_h)\right\|_h &\leq \left\{\left\|(\boldsymbol{v}_h, p_h)\right\|_h + \frac{1}{\varepsilon^{1/2}}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,\Omega} + \left(\sum_{T\in\mathcal{T}_h}\delta_{div}^T\left\|\nabla\cdot(\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h))\right\|_{0,T}^2\right)^{\frac{1}{2}}\right\} \\
&\leq C\left\{\left\|(\boldsymbol{v}_h, p_h)\right\|_h + \left(\sum_{T\in\mathcal{T}_h}\frac{\delta_{div}^T C_k^2}{h_T^2}\left\|\boldsymbol{v}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, p_h)\right\|_{0,T}^2\right)^{\frac{1}{2}}\right\} \\
&\leq \tilde{C}\left\|(\boldsymbol{v}_h, p_h)\right\|_h,
\end{aligned}
$$

where $\tilde{C}$ is independent of ε, μ, h and $\boldsymbol{\alpha}$. Hence, from (3.8) the discrete inf–sup condition

$$
\begin{aligned}
\sup_{(\boldsymbol{w}_h, q_h)\in\boldsymbol{H}_h\times Q_h^0}\frac{B\left((\boldsymbol{v}_h, p_h), (\boldsymbol{w}_h, q_h)\right)}{\left\|(\boldsymbol{w}_h, q_h)\right\|_h} &\geq \frac{B\left((\boldsymbol{v}_h, p_h), (\boldsymbol{v}_h + \gamma\tilde{\boldsymbol{w}}_h, p_h)\right)}{\left\|(\boldsymbol{v}_h + \gamma\tilde{\boldsymbol{w}}_h, p_h)\right\|_h} \\
&\geq C\left\|(\boldsymbol{v}_h, p_h)\right\|_h,
\end{aligned}
$$

follows, which concludes the proof. □

4. Error analysis. Let $k \geq 1$. We introduce the Scott-Zhang interpolation operators $\boldsymbol{\mathcal{I}}_h : H^1(\Omega)^d \longrightarrow \boldsymbol{H}_h$ and $\mathcal{J}_h : H_0^1(\Omega) \longrightarrow Q_h^0$. These interpolation operators satisfy (see, e.g., [18])

$$|\eta^{\boldsymbol{v}}|_{m,\Omega} := |\boldsymbol{v} - \boldsymbol{\mathcal{I}}_h \boldsymbol{v}|_{m,\Omega} \leq C h^{s-m} \, |\boldsymbol{v}|_{s,\Omega} \qquad \forall \boldsymbol{v} \in H^s(\Omega)^d, \tag{4.1}$$

$$|\eta^p|_{m,\Omega} := |p - \mathcal{J}_h p|_{m,\Omega} \leq C h^{s-m} \, |p|_{s,\Omega} \qquad \forall p \in H^s(\Omega) \cap H_0^1(\Omega), \tag{4.2}$$

for $0 \leq m \leq 2$ and $\max\{m, 1\} \leq s \leq k+1$.
The main error estimate for Method (3.1) is stated next.

THEOREM 4.1. *Let $(\boldsymbol{v}, p) \in H^{k+1}(\Omega)^d \times \left[H^{k+1}(\Omega) \cap H_0^1(\Omega)\right]$ be the solution of* (2.3) *and $(\boldsymbol{v}_h, p_h) \in \boldsymbol{H}_h \times Q_h^0$ the solution of* (3.1)*. Then, there exists a positive constant C, independent of ε, μ, and h, such that*

$$\|(\boldsymbol{v} - \boldsymbol{v}_h, p - p_h)\|_h \leq C h^k \Big(M_1 \, |\boldsymbol{v}|_{k+1,\Omega} + M_2 \, |p|_{k+1,\Omega} \Big), \tag{4.3}$$

where

$$M_1 = C_1 \frac{h}{\varepsilon} \quad , \quad M_2 = \mu^{1/2} h + \mu h^{3/2} + C_1 \left(\frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega} h}{\varepsilon} + 1 \right),$$

and

$$C_1 = \min \left\{ \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\mu^{1/2}}, \frac{\|\boldsymbol{\alpha}\|_{1,\infty,\Omega} h}{\mu^{1/2}} \right\} + \varepsilon^{1/2}. \tag{4.4}$$

Proof. First, using the definition of $\|\cdot\|_h$, the triangle inequality and estimates (4.1)-(4.2), we obtain

$$\begin{aligned}
&\|(\eta^{\boldsymbol{v}}, \eta^p)\|_h \\
&\leq \left\{ \frac{1}{\varepsilon^{1/2}} \|\eta^{\boldsymbol{v}}\|_{0,\Omega} + \frac{1}{\varepsilon^{1/2}} \|\boldsymbol{\Pi}_h(\boldsymbol{\alpha}\, \eta^p)\|_{0,\Omega} + \varepsilon^{1/2} \, |\eta^p|_{1,\Omega} + \mu^{1/2} \, \|\eta^p\|_{0,\Omega} + \delta_{div}^{1/2} \, \|\nabla \cdot \eta^{\boldsymbol{v}}\|_{0,\Omega} + \delta_{div}^{1/2} \mu \, \|\eta^p\|_{0,\Omega} \right\} \\
&\leq \left\{ \frac{1}{\varepsilon^{1/2}} \|\eta^{\boldsymbol{v}}\|_{0,\Omega} + \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\varepsilon^{1/2}} \|\eta^p\|_{0,\Omega} + \varepsilon^{1/2} \, |\eta^p|_{1,\Omega} + \mu^{1/2} \, \|\eta^p\|_{0,\Omega} + \delta_{div}^{1/2} \, |\eta^{\boldsymbol{v}}|_{1,\Omega} + \delta_{div}^{1/2} \mu \, \|\eta^p\|_{0,\Omega} \right\} \\
&\leq C h^k \left\{ \frac{h}{\varepsilon^{1/2}} \, |\boldsymbol{v}|_{k+1,\Omega} + \left[\varepsilon^{1/2} \left(\frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega} h}{\varepsilon} + 1 \right) + \mu^{1/2} h + \mu \, h^{3/2} \right] \, |p|_{k+1,\Omega} \right\}.
\end{aligned} \tag{4.5}$$

Next, let $(\boldsymbol{w}_h, q_h) \in \boldsymbol{H}_h \times Q_h^0$. Then, applying (2.9) to $id - \boldsymbol{\Pi}_h$ (where id denotes the identity operator) we get

$$\|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq \|\boldsymbol{\alpha}\|_{0,\infty,\Omega} \, \|q_h\|_{0,\Omega} \leq \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\mu^{1/2}} \, \|(\boldsymbol{w}_h, q_h)\|_h \, . \tag{4.6}$$

Alternatively, if we use a discrete commutator property (see Lemma 1.137 in [18]) we obtain

$$\|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq C \, h \|\boldsymbol{\alpha}\|_{1,\infty,\Omega} \, \|q_h\|_{0,\Omega} \leq C \, \frac{\|\boldsymbol{\alpha}\|_{1,\infty,\Omega} h}{\mu^{1/2}} \, \|(\boldsymbol{w}_h, q_h)\|_h. \tag{4.7}$$

So, from (4.6) and (4.7), we get

$$\|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq C \, \min \left\{ \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\mu^{1/2}}, \frac{h \|\boldsymbol{\alpha}\|_{1,\infty,\Omega}}{\mu^{1/2}} \right\} \, \|(\boldsymbol{w}_h, q_h)\|_h. \tag{4.8}$$

Thus, using the triangle inequality and (4.8) we arrive at

$$\|\boldsymbol{w}_h - \boldsymbol{\alpha} q_h\|_{0,\Omega} \leq \|\boldsymbol{w}_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} + \|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq C\, C_1\, \|(\boldsymbol{w}_h, q_h)\|_h, \tag{4.9}$$

where C_1 is given by (4.4), for all $(\boldsymbol{w}_h, q_h) \in \boldsymbol{H}_h \times Q_h^0$.
Using the definition of B and integration by parts, we arrive at

$$\begin{aligned}
&B((\eta^{\boldsymbol{v}}, \eta^p), (\boldsymbol{w}_h, q_h)) \\
&= \frac{1}{2\varepsilon}(\eta^{\boldsymbol{v}} - \boldsymbol{\alpha}\eta^p, \boldsymbol{w}_h) + \frac{1}{2}\,(\nabla\eta^p, \boldsymbol{w}_h - \boldsymbol{\alpha} q_h) - \frac{1}{2\varepsilon}\,(\eta^{\boldsymbol{v}} - \boldsymbol{\alpha}\eta^p, \boldsymbol{\alpha} q_h) \\
&\quad + \frac{\varepsilon}{2}(\nabla\eta^p, \nabla q_h) - \frac{1}{2}\,(\eta^{\boldsymbol{v}} + \boldsymbol{\alpha}\eta^p, \nabla q_h) + \mu\,(\eta^p, q_h) + \delta_{div}\,(\nabla\cdot\eta^{\boldsymbol{v}} + \mu\,\eta^p, \nabla\cdot\boldsymbol{w}_h + \mu\, q_h) \\
&= \frac{1}{2}\left(\frac{1}{\varepsilon}\eta^{\boldsymbol{v}} - \frac{1}{\varepsilon}\,\boldsymbol{\alpha}\eta^p + \nabla\eta^p, \boldsymbol{w}_h - \boldsymbol{\alpha} q_h\right) + \frac{1}{2}\,(\varepsilon\,\nabla\eta^p - \eta^{\boldsymbol{v}} - \boldsymbol{\alpha}\eta^p, \nabla q_h) + \mu\,(\eta^p, q_h) \\
&\quad + \delta_{div}\,(\nabla\cdot\eta^{\boldsymbol{v}} + \mu\,\eta^p, \nabla\cdot\boldsymbol{w}_h + \mu\, q_h) \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{4.10}$$

We bound the expression above term by term. First, I_1 is bounded using Cauchy–Schwarz inequality, estimate (4.1)-(4.2) and (4.8) as follows

$$\begin{aligned}
I_1 &\leq \left\{\frac{1}{\varepsilon}\,\|\eta^{\boldsymbol{v}}\|_{0,\Omega} + \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\varepsilon}\|\eta^p\|_{0,\Omega} + |\eta^p|_{1,\Omega}\right\}\, \|\boldsymbol{w}_h - \boldsymbol{\alpha} q_h\|_{0,\Omega} \\
&\leq C\, C_1 h^k \left\{\frac{h}{\varepsilon}\,|\boldsymbol{v}|_{k+1,\Omega} + \left(\frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega} h}{\varepsilon} + 1\right)\,|p|_{k+1,\Omega}\right\}\, \|(\boldsymbol{w}_h, q_h)\|_h.
\end{aligned} \tag{4.11}$$

Using the Cauchy-Schwarz inequality and (4.1)-(4.2), I_2 is bounded as follows

$$\begin{aligned}
I_2 &= \frac{\varepsilon}{2}\,(\nabla\eta^p, \nabla q_h) - \frac{1}{2}\,(\eta^{\boldsymbol{v}}, \nabla q_h) - \frac{1}{2}\,(\boldsymbol{\alpha}\eta^p, \nabla q_h) \\
&\leq C h^k \left\{\frac{h}{\varepsilon^{1/2}}|\boldsymbol{v}|_{k+1,\Omega} + \varepsilon^{1/2}\left(1 + \frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega} h}{\varepsilon}\right)\,|p|_{k+1,\Omega}\right\}\, \|(\boldsymbol{w}_h, q_h)\|_h.
\end{aligned} \tag{4.12}$$

For the third term in (4.10), we have

$$I_3 \leq C\,\mu\,\|\eta^p\|_{0,\Omega}\|q_h\|_{0,\Omega} \leq C\mu^{1/2}\, h^{k+1}|p|_{k+1,\Omega}\, \|(\boldsymbol{w}_h, q_h)\|_h. \tag{4.13}$$

Finally, the last term in (4.10) is bounded as follows

$$\begin{aligned}
I_4 &\leq C h^k \left\{\delta_{div}^{1/2}\,|\eta^{\boldsymbol{v}}|_{1,\Omega} + \delta_{div}^{1/2}\mu\,\|\eta^p\|_{0,\Omega}\right\}\, \|(\boldsymbol{w}_h, q_h)\|_h \\
&\leq C h^k \left\{\frac{h}{\varepsilon^{1/2}}\,|\boldsymbol{v}|_{k+1,\Omega} + \mu h^{3/2}\,|p|_{k+1,\Omega}\right\}\, \|(\boldsymbol{w}_h, q_h)\|_h.
\end{aligned} \tag{4.14}$$

Thus, defining $e_h^{\boldsymbol{v}} := \boldsymbol{v}_h - \boldsymbol{\mathcal{I}}_h\boldsymbol{v}$ and $e_h^p := p_h - \mathcal{J}_h p$, using the consistency of the scheme, (3.5), and combining (4.11)-(4.14) with (4.10), we arrive at

$$\begin{aligned}
&\|(e_h^{\boldsymbol{v}}, e_h^p)\|_h \leq C \sup_{(\boldsymbol{w}_h, q_h)\in\boldsymbol{H}_h\times Q_h^0} \frac{B((e_h^{\boldsymbol{v}}, e_h^p), (\boldsymbol{w}_h, q_h))}{\|(\boldsymbol{w}_h, q_h)\|_h} \\
&= C \sup_{(\boldsymbol{w}_h, q_h)\in\boldsymbol{H}_h\times Q_h^0} \frac{B((\eta^{\boldsymbol{v}}, \eta^p), (\boldsymbol{w}_h, q_h))}{\|(\boldsymbol{w}_h, q_h)\|_h} \\
&\leq C h^k \left\{\frac{C_1 h}{\varepsilon}\,|\boldsymbol{v}|_{k+1,\Omega} + \left[C_1\left(\frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega} h}{\varepsilon} + 1\right) + \mu^{1/2} h + \mu h^{3/2}\right]\,|p|_{k+1,\Omega}\right\}.
\end{aligned} \tag{4.15}$$

Using then the triangle inequality we arrive at

$$\|(\boldsymbol{v}-\boldsymbol{v}_h, p-p_h)\|_h \leq \|(\eta^{\boldsymbol{v}}, \eta^p)\|_h + \|(e_h^{\boldsymbol{v}}, e_h^p)\|_h,$$

and the result follows using (4.15) and (4.5). □

REMARK 5. *If we suppose $\boldsymbol{\alpha} \in W^{2,\infty}(\Omega)^d$ then a further use of the discrete commutator property gives*

$$\|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq C\, h^2 \|\boldsymbol{\alpha}\|_{2,\infty,\Omega}\, |q_h|_{1,\Omega} \leq C\, \frac{\|\boldsymbol{\alpha}\|_{2,\infty,\Omega} h^2}{\varepsilon^{1/2}}\, \|(\boldsymbol{w}_h, q_h)\|_h\,. \quad (4.16)$$

Thus, combining this estimate with (4.7) *we obtain*

$$\|\boldsymbol{\alpha} q_h - \boldsymbol{\Pi}_h(\boldsymbol{\alpha} q_h)\|_{0,\Omega} \leq C\, C_1\, \|(\boldsymbol{w}_h, q_h)\|_h, \quad (4.17)$$

but now with $C_1 := \min\left\{\frac{\|\boldsymbol{\alpha}\|_{0,\infty,\Omega}}{\mu^{1/2}}, \frac{\|\boldsymbol{\alpha}\|_{1,\infty,\Omega} h}{\mu^{1/2}}, \frac{\|\boldsymbol{\alpha}\|_{2,\infty,\Omega} h^2}{\varepsilon^{1/2}}\right\} + \varepsilon^{1/2}$ *in Theorem 4.1.*

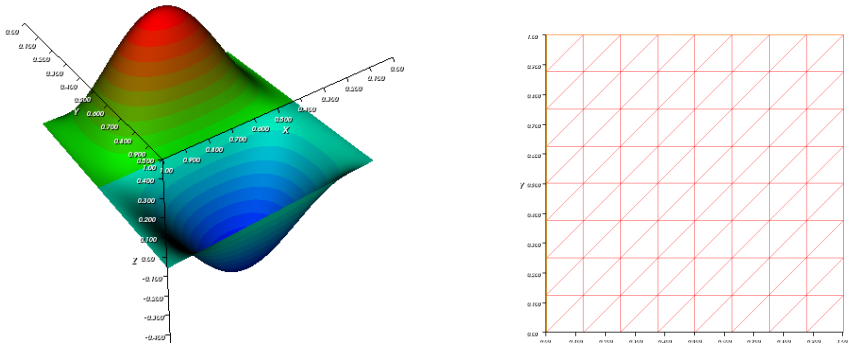
REMARK 6. *It is important to remark that neither the stability analysis nor the convergence analysis of the stabilised method* (3.1) *uses the fact that $\boldsymbol{\alpha}$ is solenoidal. As was mentioned earlier, this restriction is only used to derive* (2.3)*. If $\boldsymbol{\alpha}$ is not assumed to be solenoidal, then, starting from* (2.3) *the method can still be proposed and analysed, but the relation between* (2.3) *and* (2.1) *is less clear. Alternatively, a stabilised method similar to* (3.1) *can be proposed starting from* (2.5) *instead. The implications of this are not clear at the moment, and will be the subject of future research.*

5. Convergence studies for Method (3.1).

5.1. A two-dimensional problem with a smooth analytical solution. We start testing the numerical performance of Method (3.1) by considering a two-dimensional example with a smooth, known solution. More significant tests with singular solutions will be considered afterwards. For this, and all subsequent, numerical experiments the value of δ has been set to 1. We consider $\Omega = (0,1)^2$, $\boldsymbol{\alpha} = [y, -x]^T$, $\mu = 0$, and test with different values of ε ranging from 10^{-5} to 1. Both f and the boundary conditions are chosen such that the solution of (2.1) is given by $p(x,y) = 100x^2(1-x)^2 y(1-y)(1-2y)$, depicted in Figure 5.1a. Structured Friedrichs–Keller–type meshes are used in these computations as shown in Figure 5.1b, where N is the number of nodes along one side.

In Figure 5.2 we depict the errors obtained on implementing method (3.1) in a sequence of uniformily refined meshes obtained increasing the value of N. The first two plots correspond to the results obtained by using $\mathcal{P}_1$ elements for both variables, p and $\boldsymbol{v}$, with $\varepsilon = 10^{-3}$ (Figure 5.2a) and $\varepsilon = 10^{-5}$ (Figure 5.2b) . We observe that all the errors tend to zero with a ratio which is consistent with the results of Section 3. The same comments are applicable to the cases depicted in the Figures 5.2c and 5.2d, where quadratic $\mathcal{P}_2$ elements are considered for both variables.

Finally, to justify our choice of stabilisation parameter δ we carry out the following experiment. We fix a mesh, of the type depicted in Figure 5.1b with $N = 2^6$, $\varepsilon = 10^{-3}$, and compute the errors for the method using a range of values for δ, spanning from 10^{-2} to 10^2. The results are depicted in Figure 5.3. For this smooth solution all the errors, except for the one associated to the divergence of $\boldsymbol{v}$ (which is multiplied by $\delta^{\frac{1}{2}}$), show a fairly robust behavior with respect to δ in this range (they do deteriorate for more extreme choices). This will not be the case for a problem presenting boundary and interior layers, as it will be the subject of a numerical test presented later.



(a) Exact solution for convergence test. (b) mesh for convergence test with $N = 9$.

Fig. 5.1: Exact solution and sample mesh.

5.2. A three-dimensional numerical experiment. In this section we perform a convergence test for the Method (3.1) for a smooth solution in a three-dimensional domain. We consider $\Omega = (0,1)^3$, $\varepsilon = 10^{-3}$, $\mu = 0$, $\boldsymbol{\alpha} = [1,2,1]^T$, and f is chosen such that the exact solution is given by

$$u(x,y,z) = \sin(2\pi x)\sin(2\pi y)\sin(2\pi z)\,. \tag{5.1}$$

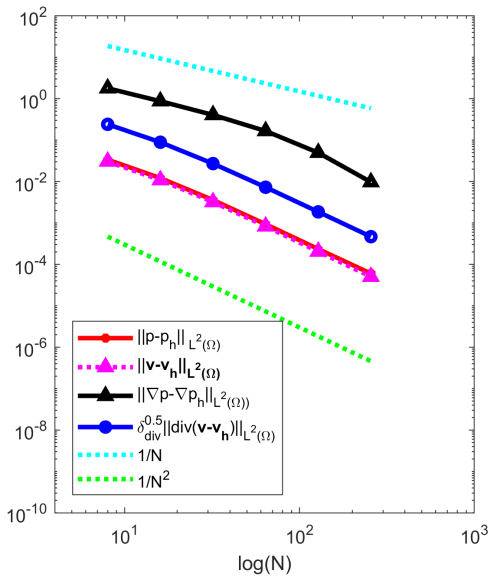
The domain is partitioned by dividing each side of the unit cube into N segments of equal length. This generates a structured mesh of each face of the unit cube, which is then propagated inside the domain (for details, see the Freefem documentation, or [23]). We have measured the errors in the different norms, and the results are depicted in Figure 5.4, where we can see that they follow orders that are in accordance with the theoretical results.

6. A numerical assessment of different stabilised mixed methods. In this section we review different alternative mixed discretisations of (2.1), and carry out two series of numerical experiments to evaluate them, along with a further numerical assessment of the performance of Method (3.1).

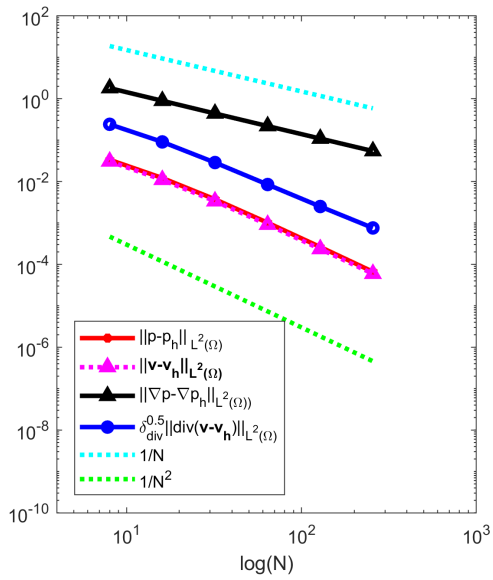
6.1. Previous mixed methods for (2.1)**.** We now review some existing stabilised mixed finite element methods for (2.1). Our presentation is restricted to $d = 2$ for simplicity.

6.1.1. Masud and Kwack method. For the method proposed in [31], continuous Lagrangian elements of order $k \geq 1$ were used to approximate both variables. The method reads as follows: find $(\boldsymbol{v}_h, p_h) \in \boldsymbol{H}_h \times Q_h^0$, such that

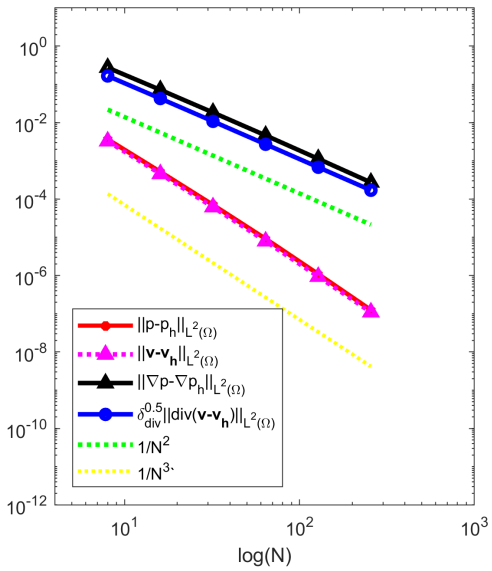
$$\begin{aligned}\frac{1}{\varepsilon}(\boldsymbol{v_h}, \boldsymbol{w_h}) - (p_h, \nabla\cdot\boldsymbol{w_h}) - \frac{1}{\varepsilon}(\boldsymbol{\alpha} p_h, \boldsymbol{w_h}) + (\nabla\cdot\boldsymbol{v_h}, q_h) + \mu(p_h, q_h)\\ -(\tau(\boldsymbol{v_h} - \boldsymbol{\alpha} p_h + \varepsilon\nabla p_h), \frac{1}{\varepsilon}\boldsymbol{w_h} - \nabla q_h) = (f, q_h)\,,\end{aligned} \tag{6.1}$$



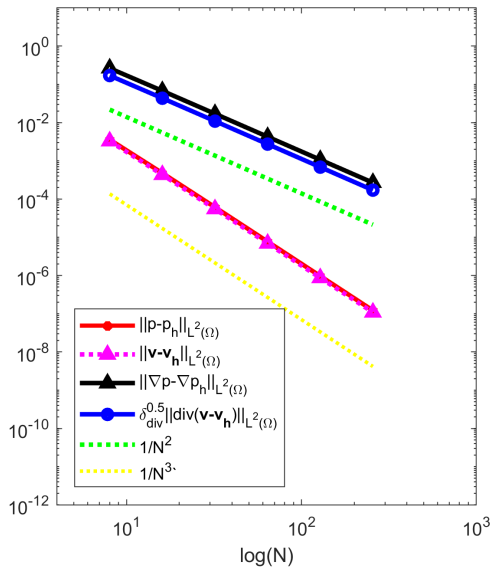
(a) $\mathcal{P}_1\mathcal{P}_1$ Convergence study $\varepsilon = 10^{-3}$.



(b) $\mathcal{P}_1\mathcal{P}_1$ Convergence study $\varepsilon = 10^{-5}$.

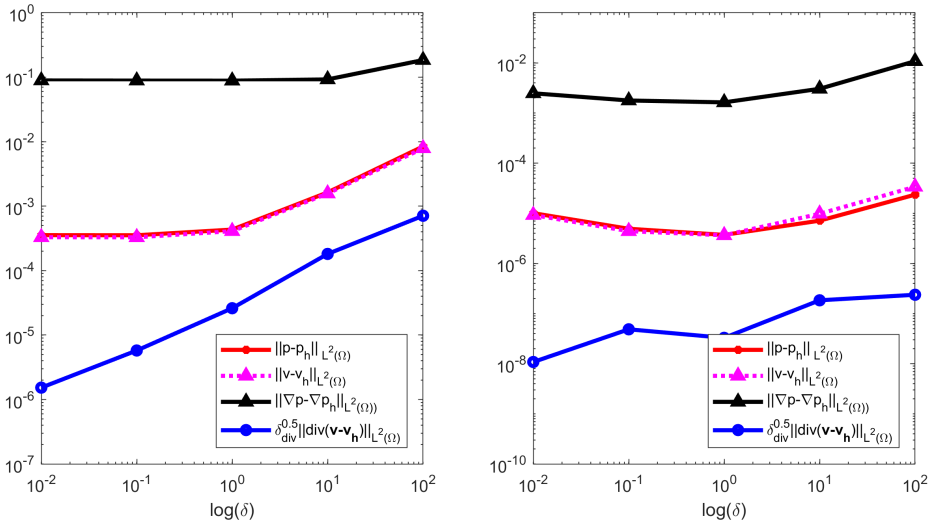


(c) $\mathcal{P}_2\mathcal{P}_2$ Convergence study $\varepsilon = 10^{-3}$.



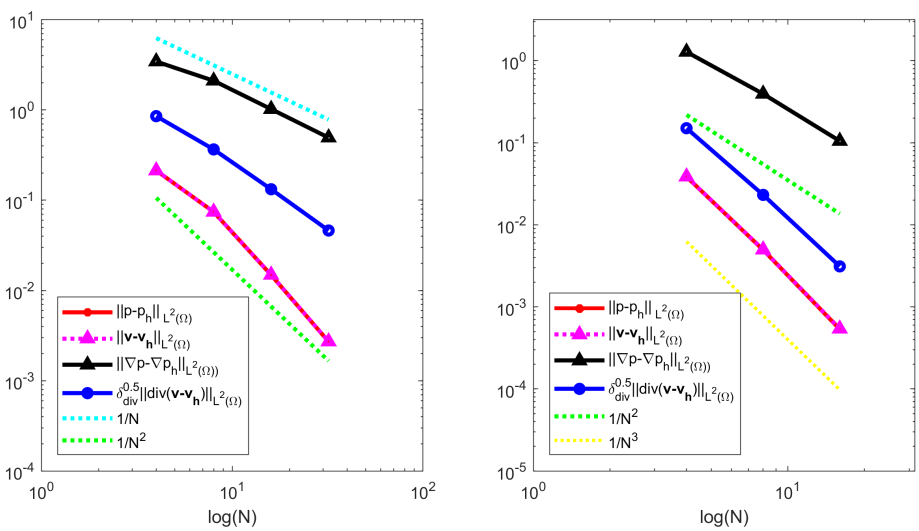
(d) $\mathcal{P}_2\mathcal{P}_2$ Convergence study $\varepsilon = 10^{-5}$.

Fig. 5.2: Convergence studies for the Present Method.



(a) Present method using $\mathcal{P}_1\,\mathcal{P}_1$ elements. (b) Present method using $\mathcal{P}_2\,\mathcal{P}_2$ elements.

Fig. 5.3: Errors for the present method for $\varepsilon = 10^{-3}$, and different values for δ.



(a) $3D - \mathcal{P}_1\mathcal{P}_1$ Convergence study $\varepsilon = 10^{-3}$. (b) $3D - \mathcal{P}_2\mathcal{P}_2$ Convergence study $\varepsilon = 10^{-3}$.

Fig. 5.4: Three-dimensional case with $\varepsilon = 10^{-3}$: convergence of the method.

for all $(\boldsymbol{w}_h, q_h) \in \boldsymbol{H}_h \times Q_h^0$. This method is referred to as MK in our numerical results. The value of τ was estimated in [31] from calculations using bubble functions.

REMARK 7. *The differences between* (6.1) *and* (3.1) *appear in a more explicit manner. In fact, we see that* (3.1) *includes a term involving* $\boldsymbol{\alpha} q_h$ *that made it possible to show the inf-sup condition, and an extra div-div term that enhances the stability, thus improving the numerical results greatly.*

6.1.2. Raviart–Thomas based mixed methods. The Raviart-Thomas pair of finite elements introduced in [33] is one of the most popular discrete inf-sup stable pairs for first-order mixed problems. For a simplex $T \in \mathcal{T}_h$ the $\mathcal{RT}$ space of order $k \geq 0$, is defined as

$$\mathcal{RT}_k(T) = \mathcal{P}_k(T)^2 + \boldsymbol{x}\mathcal{P}_k(T) \ . \tag{6.2}$$

Then the associated global Raviart-Thomas space is given by

$$\mathcal{RT}_k(\Omega) = \{\boldsymbol{v}_h \in H(\mathrm{div}, \Omega) : \boldsymbol{v}_h|_T \in \mathcal{RT}_k(T), \ \forall T \in \mathcal{T}_h\} \quad . \tag{6.3}$$

The primal variable p is approximated using the space of discontinuous piecewise polynomial function of degree $k \geq 0$ given by

$$\mathcal{P}_k^{dc}(\Omega) = \{q_h \in L^2(\Omega) : q_h|_T \in \mathcal{P}_k(T), \forall T \in \mathcal{T}_h\} \, . \tag{6.4}$$

The first mixed discretisations of (2.1) using the pair $\mathcal{RT}_k(\Omega) \times \mathcal{P}_k^{dc}(\Omega)$ are presented in [16, 17]. However, since those papers deal with non-stabilised methods, the mixed formulations suffer from the same instabilities as the plain Galerkin method, and our numerical experiments confirm that fact. Therefore, we have not included this version in our numerical comparison. As was mentioned in the introduction, in [35] a stabilised finite element method was proposed for one of the weak forms from [16]. In that work, the imposition of essential boundary conditions is done weakly, without adding any extra inforcement of them. Then, although the method does produce sharp layers, this (very) weak imposition of the boundary conditions (especially at entry) leads to inaccurate results (see Figure 6.2 for details in the first case tested). Thus, we have also not included the method from [35] in our study.

6.1.3. Weakly imposed boundary conditions and a weighted FOSLS approach. As was mentioned in the introduction, FOSLS methods usually show diffusive results for problems involving sharp layers. To improve this, Chen et al. [15] proposed the following first-order formulation for (2.4).

$$\boldsymbol{v} + \varepsilon^{1/2}\nabla p = 0 \qquad \text{in } \Omega \ , \tag{6.5a}$$

$$\varepsilon^{1/2}\nabla \cdot \boldsymbol{v} + \boldsymbol{\alpha} \cdot \nabla p + \mu p = f \qquad \text{in } \Omega \ . \tag{6.5b}$$

The solution is sought in the finite element space $\mathcal{U}_h := \mathcal{RT}_k(\Omega) \times Q_h$, where Q_h is defined in (2.7), using $k \geq 1$. The method proposed in [15] reads as follows: Find $(\boldsymbol{v}_h, p_h) \in \mathcal{U}_h$ such that

$$\begin{aligned}(\boldsymbol{v}_h + \varepsilon^{1/2}\nabla p_h \ , \ \boldsymbol{w}_h + \varepsilon^{1/2}\nabla q_h) + \sum_{F \in \xi_h^\partial} h_F^{-1}\Big\langle (\varepsilon + \max(-\boldsymbol{\alpha}\cdot\boldsymbol{n}, 0))p_h, q_h \Big\rangle_F \\ +(\varepsilon^{1/2}\nabla \cdot \boldsymbol{v}_h + \boldsymbol{\alpha}\cdot\nabla p_h + \mu p_h \ , \ \varepsilon^{1/2}\nabla\cdot\boldsymbol{w}_h + \boldsymbol{\alpha}\cdot\nabla q_h + \mu q_h) \\ = (f \ , \ \varepsilon^{1/2}\nabla\cdot\boldsymbol{w}_h + \boldsymbol{\alpha}\cdot\nabla q_h + \mu q_h) + \sum_{F \in \xi_h^\partial} h_F^{-1}\Big\langle (\varepsilon + \max(-\boldsymbol{\alpha}\cdot\boldsymbol{n}, 0))g, q_h \Big\rangle_F \ ,\end{aligned} \tag{6.6}$$

for all $(\boldsymbol{w}_h, q_h) \in \mathcal{U}_h$. Here, ξ_h^∂ is the set of edges of the triangulation (denoted by F) that lie in the boundary Γ, $h_F = |F|$, $\langle .,. \rangle_F$ stands for the inner product in $L^2(F)$, $\boldsymbol{n}$ denotes the unit normal vector outward to Γ, and the term in the right-hand side involving g is present to cover the possibly more general case in which the boundary condition is $p = g$ on Γ. This method will be referred to as FOSLS in our numerical experiments. As an alternative weak imposition of the boundary conditions in [15, Remark 2.2], the following method is proposed: find $(\boldsymbol{v}_h, p_h) \in \mathcal{U}_h$ such that

$$\begin{aligned}(\boldsymbol{v}_h + \varepsilon^{1/2}\nabla p_h \ , \ \boldsymbol{w}_h + \varepsilon^{1/2}\nabla q_h) + \sum_{F\in\xi_h^\partial} \Big\langle (h_F^{-1}\varepsilon + \max(-\boldsymbol{\alpha}\cdot\boldsymbol{n}, 0))p_h, q_h \Big\rangle_F \\ +(\varepsilon^{1/2}\nabla\cdot\boldsymbol{v}_h + \boldsymbol{\alpha}\cdot\nabla p_h + \mu p_h \ , \ \varepsilon^{1/2}\nabla\cdot\boldsymbol{w}_h + \boldsymbol{\alpha}\cdot\nabla q_h + \mu q_h) \\ = (f \ , \ \varepsilon^{1/2}\nabla\cdot\boldsymbol{w}_h + \boldsymbol{\alpha}\cdot\nabla q_h + \mu q_h) + \sum_{F\in\xi_h^\partial} \Big\langle (h_F^{-1}\varepsilon + \max(-\boldsymbol{\alpha}\cdot\boldsymbol{n}, 0))g, q_h \Big\rangle_F \ , \quad (6.7)\end{aligned}$$

for all $(\boldsymbol{w}_h, q_h) \in \mathcal{U}_h$. This alternative will be referred to as FOSLSb in our experiments that follow.

6.2. Advection skew to the mesh test. This test is a slight variation of the test proposed in [10]. The advective velocity is chosen as $\boldsymbol{\alpha} = \frac{1}{\sqrt{5}}[1,2]^T$ (giving $|\boldsymbol{\alpha}| = 1$), and the same family of meshes from Figure 5.1b is used on the unit square domain of $\Omega = (0,1)^2$ with $f = 0, \mu = 0$, and ε in a range of values from 10^{-5} to 1. We impose Dirichlet boundary conditions for p on the whole boundary, given by

$$p(x,y) = \begin{cases} 1 \text{ on } \{(0,y) : 0 \le y \le 1\} \cup \{(x,1) : 0 \le x \le 1\} \\ 0 \text{ on } \{(1,y) : 0 \le y \le 1\} \cup \{(x,0) : 0 \le x \le 1\}\,. \end{cases}$$

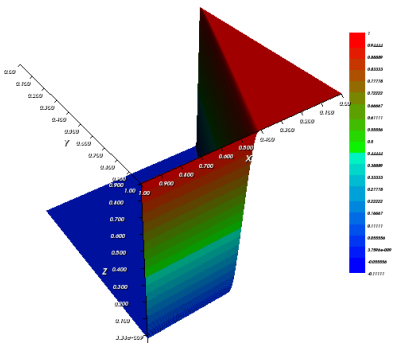
The analytical solution to this problem is not known. Therefore, we have computed a reference solution using the SUPG method on a highly refined mesh using $N = 2^{11}$ (giving 8,388,608 triangles), and quadratic ($\mathcal{P}_2$) elements. Elevations and cross-sections of the reference solution are depicted in Figure 6.1. The SUPG method has been implemented using the definition of the stabilisation parameter given in [20], that is

$$\tau_T = \frac{h_T}{2|\boldsymbol{\alpha}|}\min\left\{1, \frac{m_k h_T|\boldsymbol{\alpha}|}{2\varepsilon}\right\} \qquad \text{for } T \in \mathcal{T}_h\,, \tag{6.8}$$

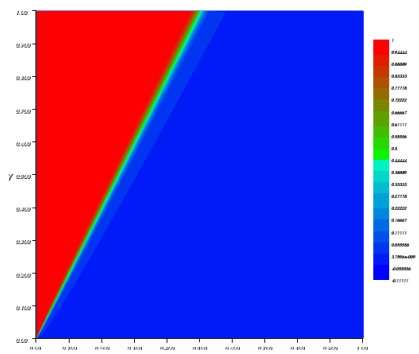
and m_k is a constant appearing in an inverse inequality related to (2.11) (for details, see [20]).

To justify the non-inclusion of the method presented in [35] in our detailed study, in Figure 6.2 we show the cross-section along the line $y = 0.5$ of the solution given by it, compared to the reference, and the FOSLS method (both using $\mathcal{RT}_1(\Omega)$ elements for $\boldsymbol{v}$). As was stated, the weak imposition of essential boundary conditions makes the results inaccurate. This can be seen in the figure, where the value of p_h misses the boundary condition by a margin too large to be deemed acceptable.

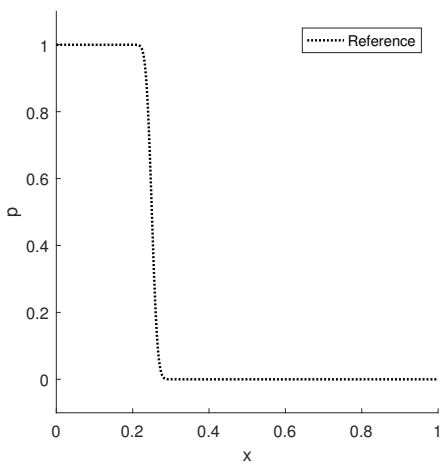
Before moving onto more detailed comparisons, we further justify our choice of the stabilisation parameter δ by performing a sensitivity test. We fix $\varepsilon = 10^{-4}$ and the mesh using $N = 2^6$, and solve the problem for a variety of values for δ ranging from 10^{-2} to 10^2. The results are depicted in Figure 6.3, where we can observe that too small a value for δ does not add enough stability to the method, resulting in the presence of significant oscillations in the discrete solution. On the other hand, too



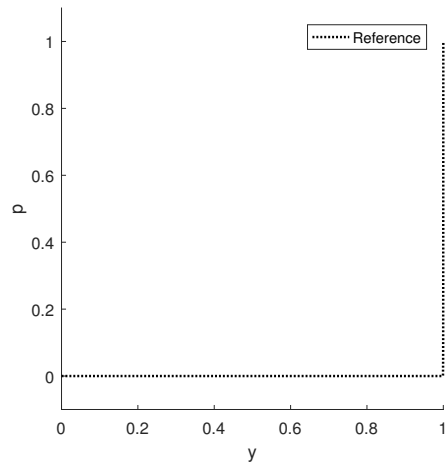
(a) 3-D visualisation of the reference solution.



(b) 2-D visualisation of the reference solution.



(c) Reference cross-section at $y = 0.5$.



(d) Reference cross-section at $x = 0.7$.

Fig. 6.1: SUPG $\mathcal{P}_2$ reference solution, $N = 2^{11}$, $\varepsilon = 10^{-4}$, for the advection skew to the mesh test.

large a value of δ results in a numerical solution that is too diffusive to be considered of practical interest. Based on these results (and others not shown in this manuscript due to space restrictions) we conclude that $\delta = 1$ is an appropriate value for the stabilisation parameter for this method.

In Figure 6.4 we depict elevations of the discrete solutions obtained using the different methods described in the last section. A more detailed comparison, using cross-sections along the lines $y = 0.5$ and $x = 0.7$, is carried out in Figure 6.5 (for linear elements) and Figure 6.6 (for quadratic elements). We also have included the solution obtained using the SUPG method on the same mesh, and the reference solution. We observe that the MK method exhibits oscillations near the outflow layer and that, although reduced, these are not eliminated when the mesh is refined. For quadratic elements we also include the results given by methods (6.6) and (6.7), since these are second order methods.

The same oscillations that appear for linear elements are also present for the MK

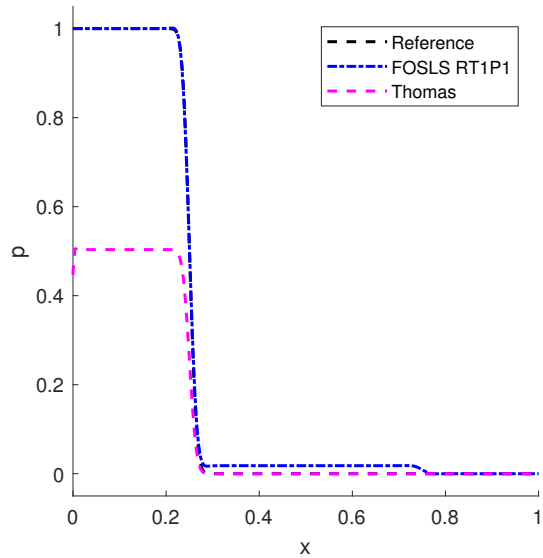


Fig. 6.2: Cross-section at $y = 0.5$. The method from [35] does not satisfy the inflow condition.

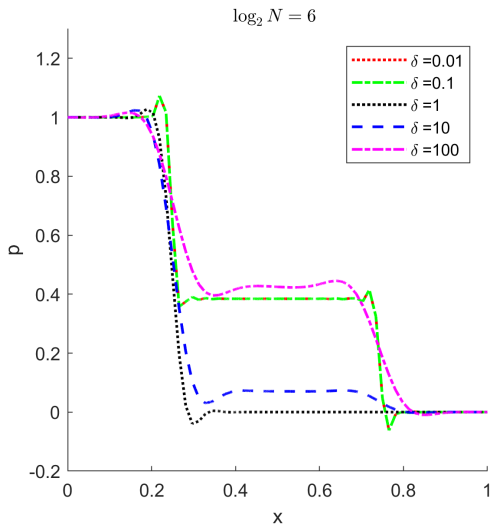
solution using quadratic elements. These are not present for FOSLS methods thanks to the weak imposition of the boundary conditions. However, this comes at the price of the FOSLS solutions not capturing the outflow boundary layer, unless the mesh is extremely refined. This can be observed in Figure 6.7a where we zoom in on the plots with all the solutions (except [31]) for $N = 2^7$. Here we observe that SUPG and the present method capture the outflow boundary layer, while the FOSLS methods do not.

We continue by examining the over- and undershoots produced by all of the methods. These are computed as follows:

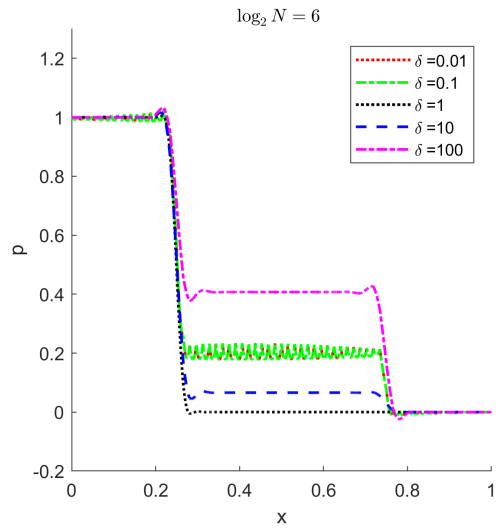
$$p_{max} = \max_{x \in \bar{\Omega}} p_h(x) - 1,$$
$$p_{min} = \min_{x \in \bar{\Omega}} p_h(x) \ .$$

For quadratic elements, we have approximated these values by using the maximum and minimum over the degrees of freedom. These results are depicted for different levels of mesh refinment in Figure 6.8. The over- and undershoots given by the present method show a comparable behaviour to SUPG, with both outperforming the results given by both FOSLS methods and the MK method, especially for small values of ε (note that some of the results from the MK method lie outside the plots).

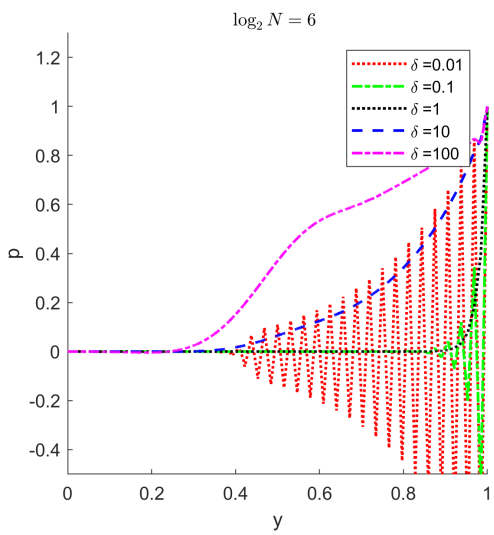
Finally, we compare the layer thickness of both the internal and outflow layers. In the graphs in Figure 6.9, the width of the interior layer is defined as the width of the interval taken for the value of the discrete solution p_h along the line $y = 0.5$ to decrease from 0.9 to 0.1. Similarly, the width of the outflow layer is defined as the width of the interval taken for the value of the discrete solution p_h along the line $x = 0.7$ to increase from 0.1 to 0.9. We observe that, among the mixed alternatives, the present method provides the best results. The internal layer width provided by



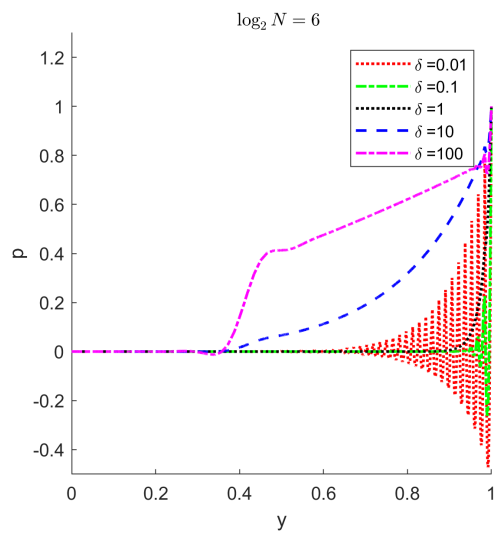
(a) Cross-section at $y = 0.5$. $\mathcal{P}_1\,\mathcal{P}_1$ elements.



(b) Cross-section at $y = 0.5$. $\mathcal{P}_2\,\mathcal{P}_2$ elements.



(c) Cross-section at $x = 0.7$. $\mathcal{P}_1\,\mathcal{P}_1$ elements.

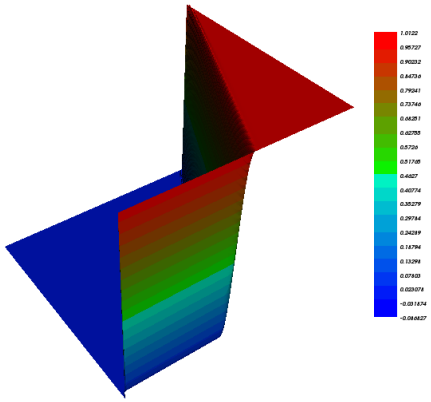


(d) Cross-section at $x = 0.7$. $\mathcal{P}_2\,\mathcal{P}_2$ elements.

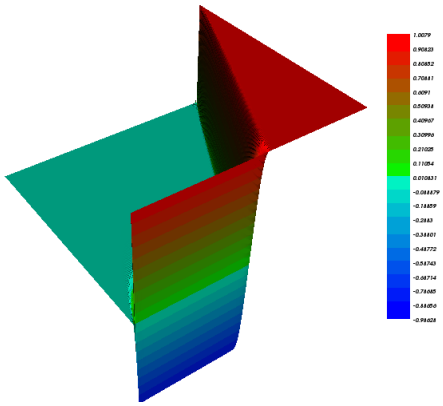
Fig. 6.3: Cross-sections of the solutions for different values of δ.

(3.1) is comparable to the one given by the SUPG method, and the outflow layer is larger than the one given by SUPG for linear elements, and comparable to it for quadratic elements. The instabilities of the method in [31] led us to remove its results when considering $\varepsilon = 10^{-5}$. It is worth mentioning that the increase of the interior layer width with increasing refinement in both FOSLS methods is due to the fact that the weak imposition of the outflow boundary conditions makes the method only

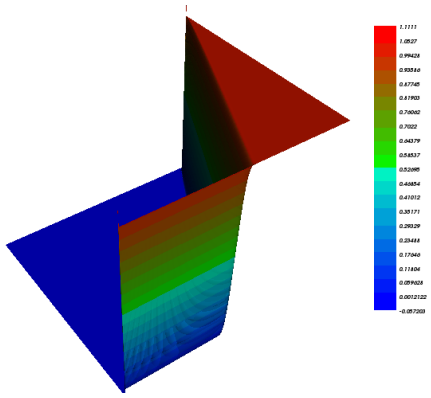
capture the outflow layer if the mesh is refined enough. To illustrate this, in Figure 6.11 we plot elevations of the discrete solution given by both FOSLS methods with $\varepsilon = 10^{-3}$ and $N = 2^8$, where we can see that they present an outflow layer that was absent when $N = 2^7$ was used.



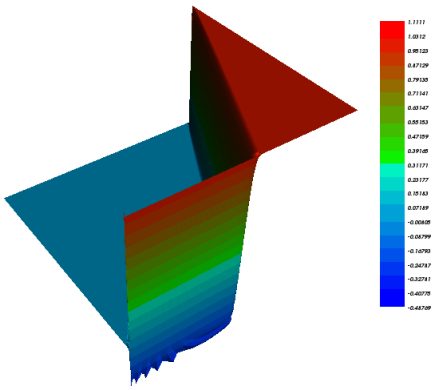
(a) Present method $\mathcal{P}_1\mathcal{P}_1$.



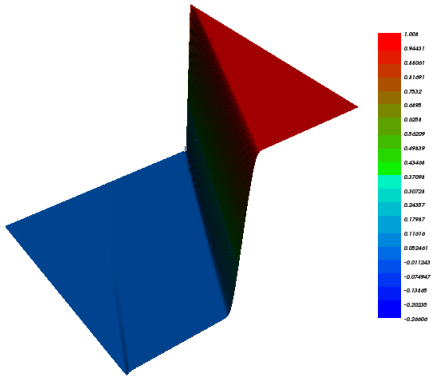
(b) MK method $\mathcal{P}_1\mathcal{P}_1$.



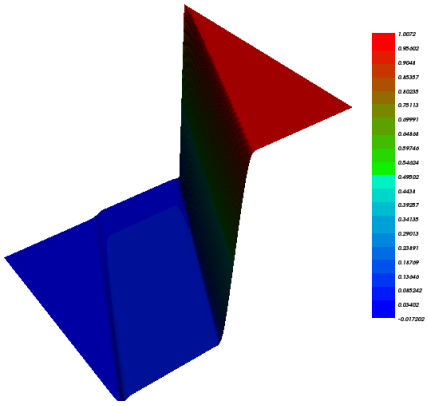
(c) Present method $\mathcal{P}_2\mathcal{P}_2$.



(d) MK method $\mathcal{P}_2\mathcal{P}_2$.

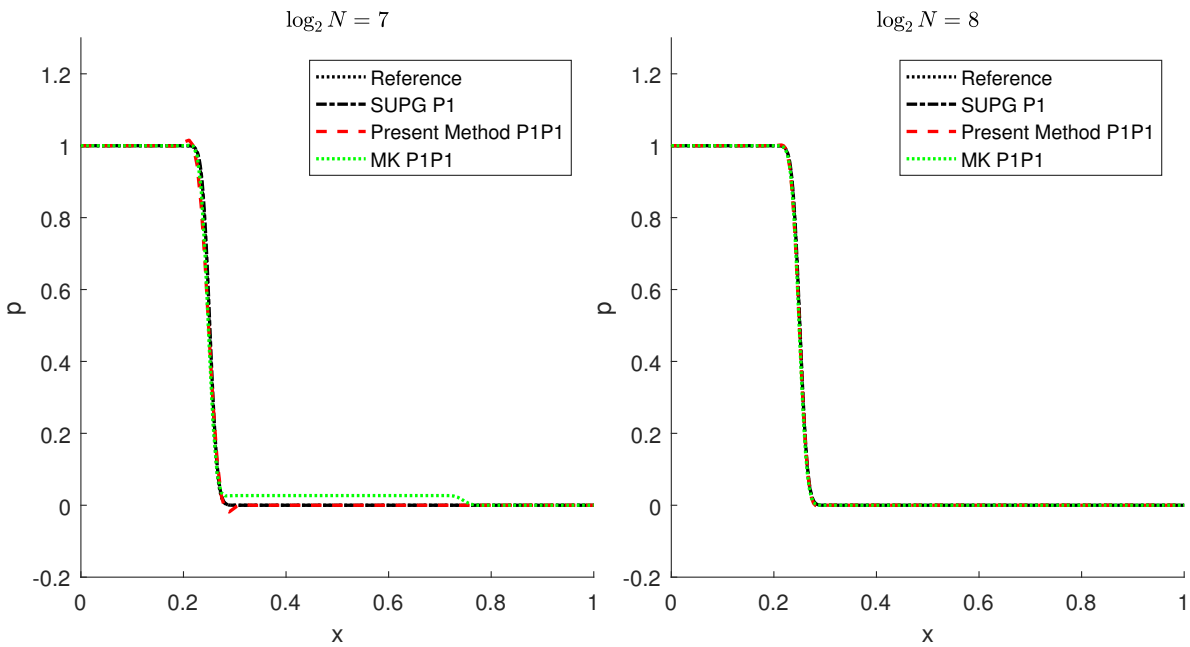


(e) FOSLS method RT1P1.



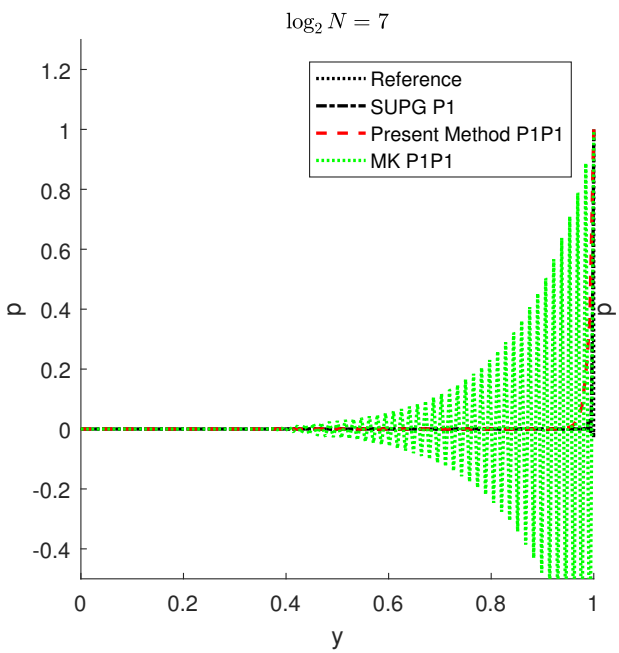
(f) FOSLSb method RT1P1.

Fig. 6.4: Advection skew to the mesh: Elevations using $N = 2^8$ and $\varepsilon = 10^{-4}$.

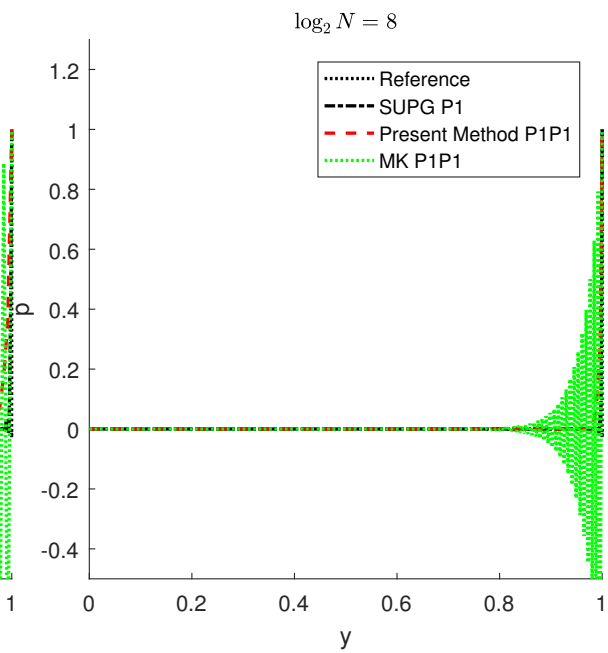


(a) cross-section at $y = 0.5$.

(b) cross-section at $y = 0.5$.

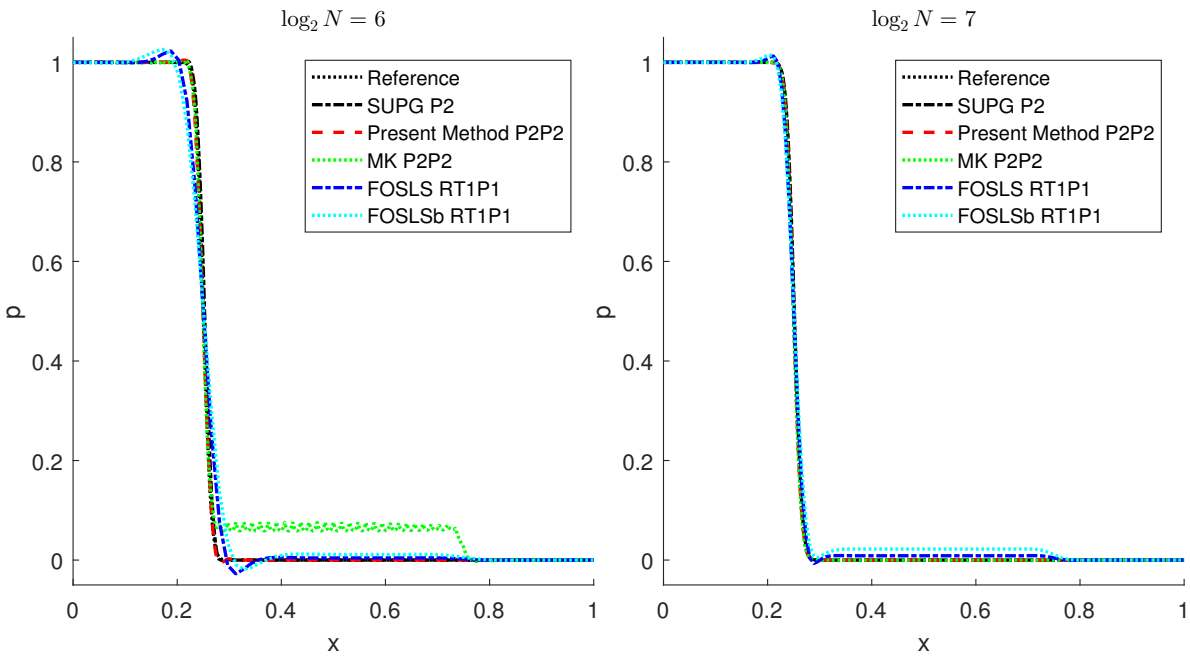


(c) cross-section at $x = 0.7$.



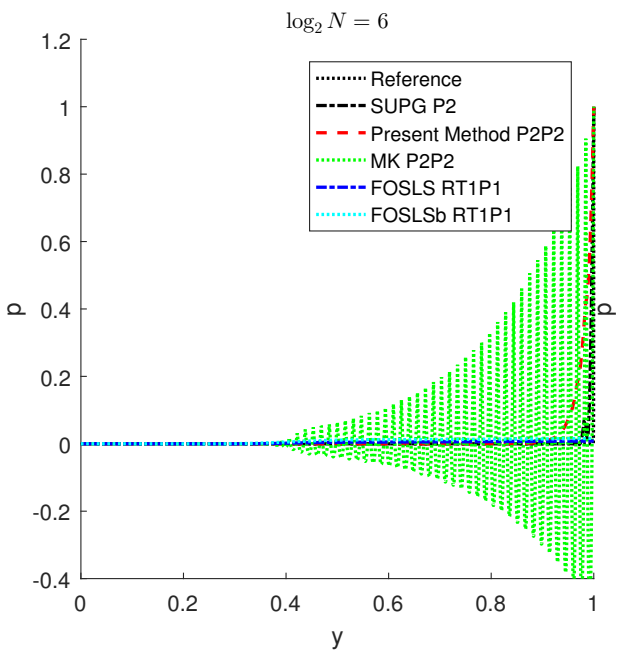
(d) cross-section at $x = 0.7$.

Fig. 6.5: Advection skew to the mesh, $\varepsilon = 10^{-4}$: Cross-sections of the different methods considered using linear elements.

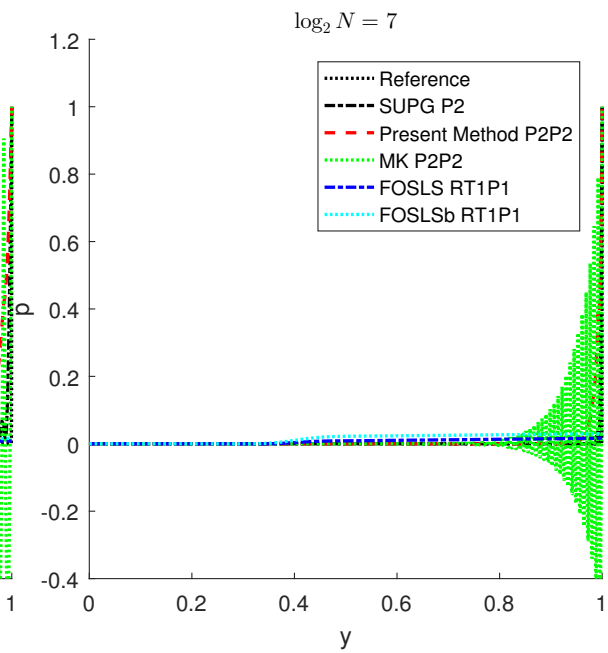


(a) cross-section at $y = 0.5$.

(b) cross-section at $y = 0.5$.

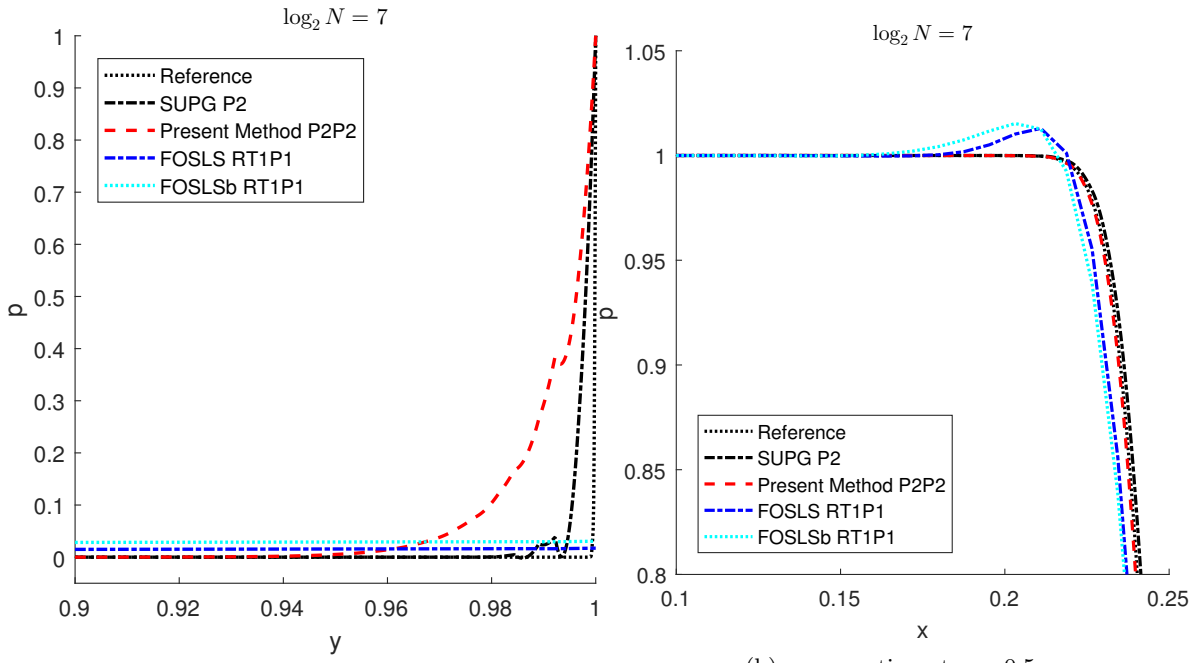


(c) cross-section at $x = 0.7$.



(d) cross-section at $x = 0.7$.

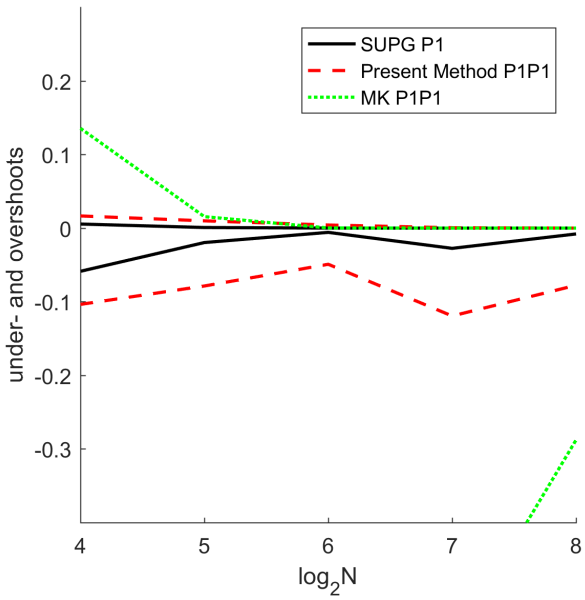
Fig. 6.6: Advection skew to the mesh, $\varepsilon = 10^{-4}$: Cross-sections of the different methods considered using quadratic elements.



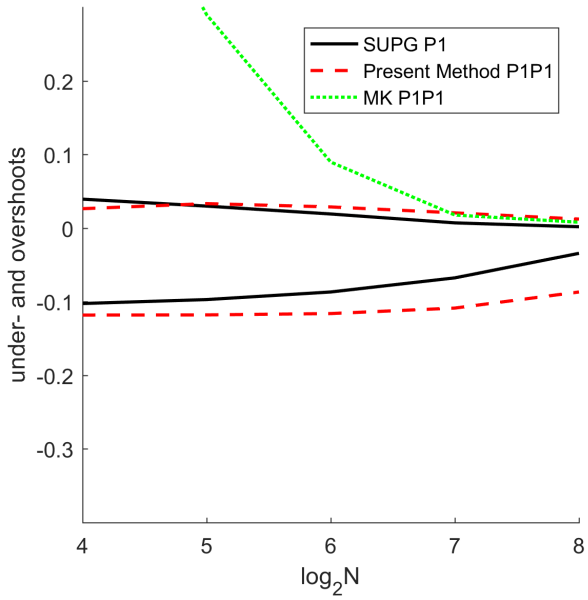
(a) cross-section at $x = 0.7$.

(b) cross-section at $y = 0.5$.

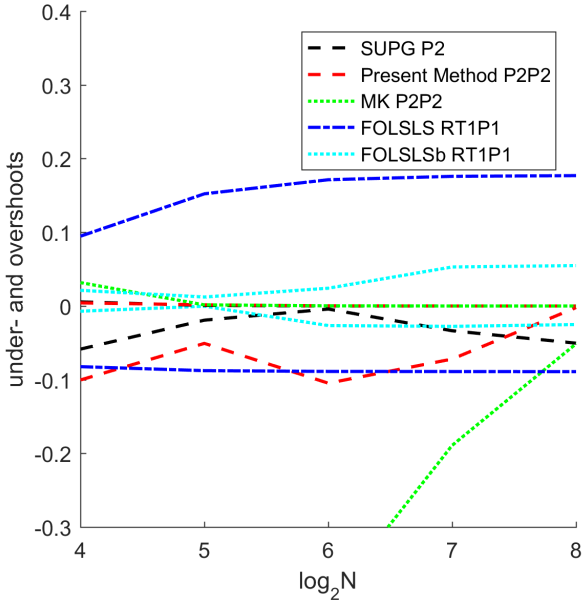
Fig. 6.7: Advection skew to the mesh, $\varepsilon = 10^{-4}$: Close-up of cross-sections of the different methods considered using quadratic elements.



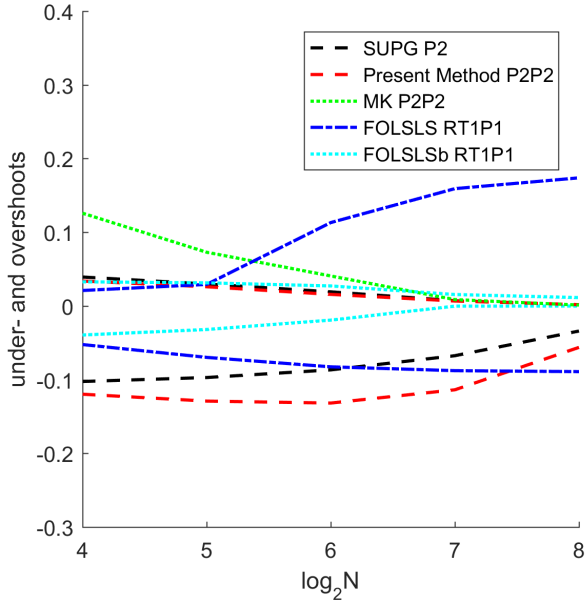
(a) Linear elements over- and undershoots, $\varepsilon = 10^{-3}$.



(b) Linear elements over- and undershoots, $\varepsilon = 10^{-4}$.

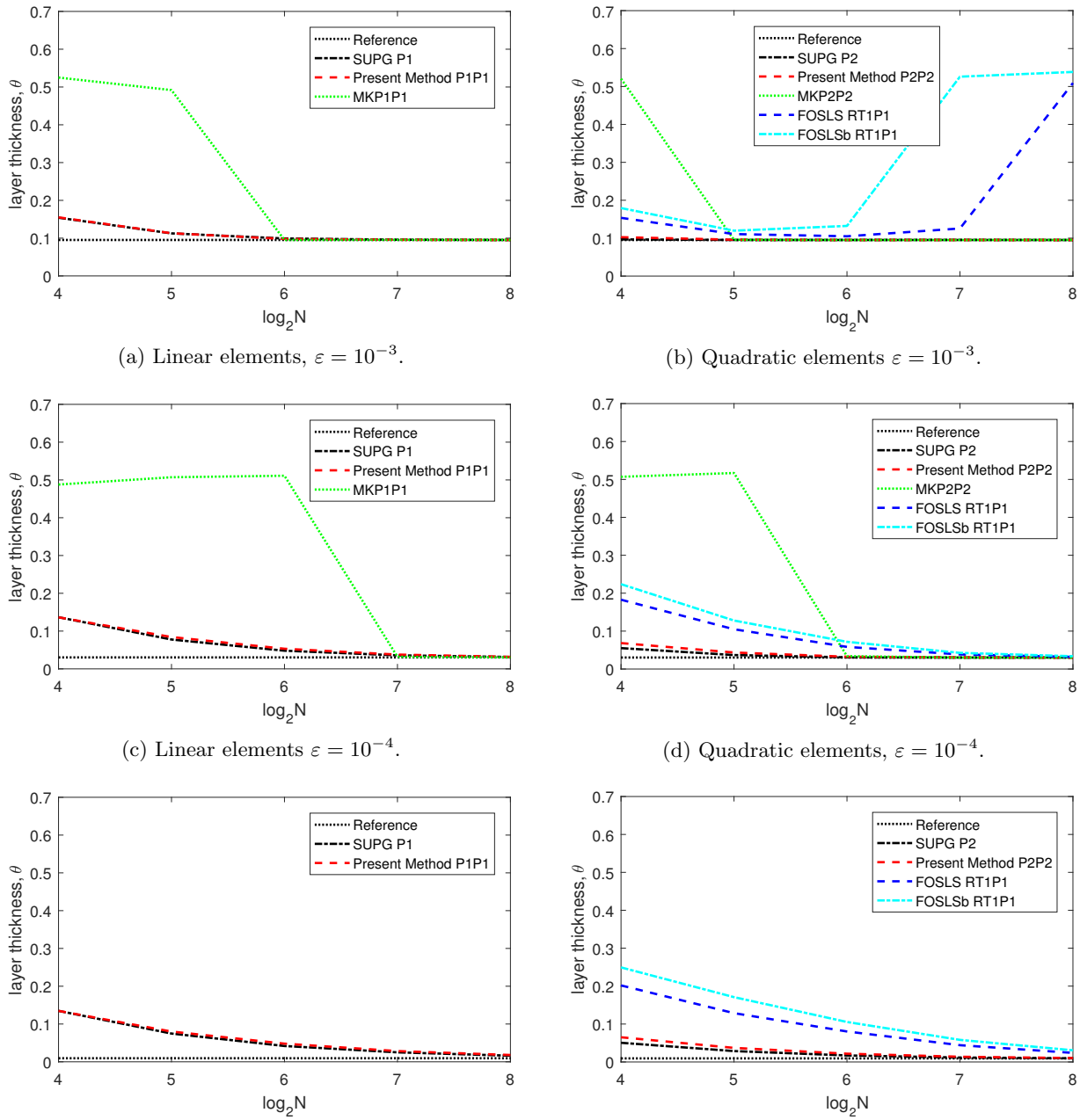


(c) Quadratic elements over- and undershoots, $\varepsilon = 10^{-3}$.



(d) Quadratic elements over- and undershoots, $\varepsilon = 10^{-4}$.

Fig. 6.8: Advection skew to the mesh, different values for ε: Over- and undershoots for the different methods.



(a) Linear elements, $\varepsilon = 10^{-3}$.

(b) Quadratic elements $\varepsilon = 10^{-3}$.

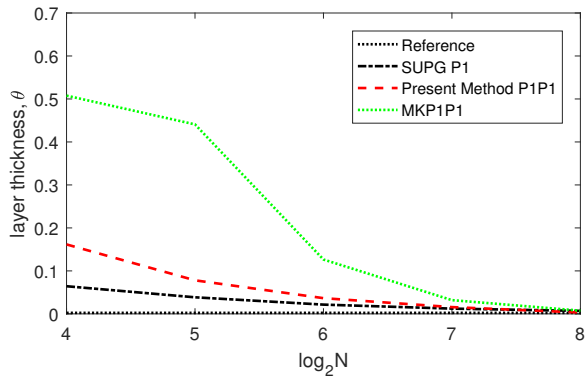
(c) Linear elements $\varepsilon = 10^{-4}$.

(d) Quadratic elements, $\varepsilon = 10^{-4}$.

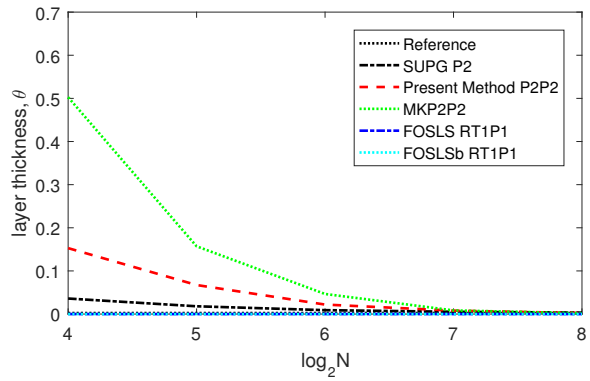
(e) Linear elements, $\varepsilon = 10^{-5}$.

(f) Quadratic elements, $\varepsilon = 10^{-5}$.

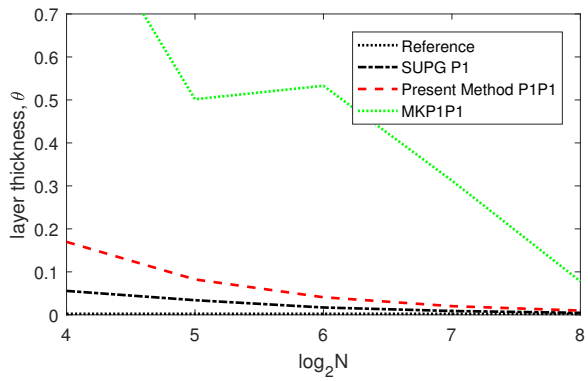
Fig. 6.9: Advection skew to the mesh, different values for ε: Internal layer thickness, θ, for $0.1 < p(x, 0.5) < 0.9$ with respect to the refinement level.



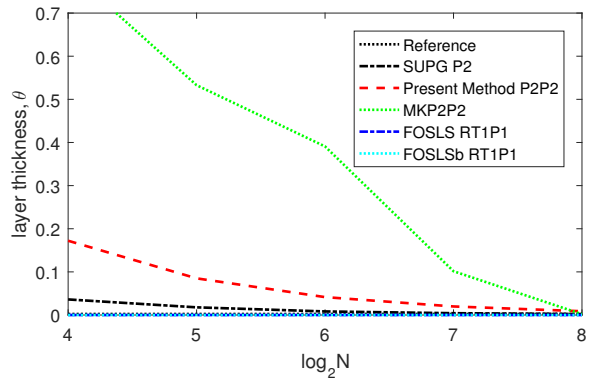
(a) Linear elements, $\varepsilon = 10^{-3}$.



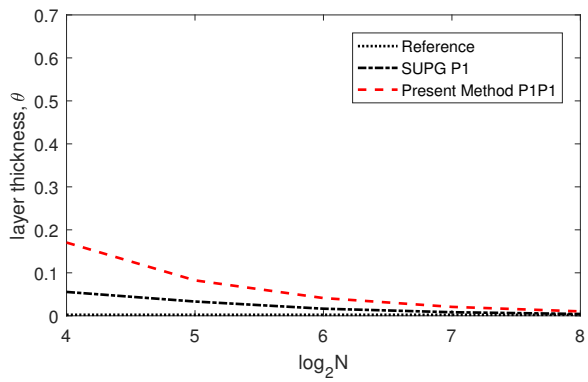
(b) Quadratic elements, $\varepsilon = 10^{-3}$.



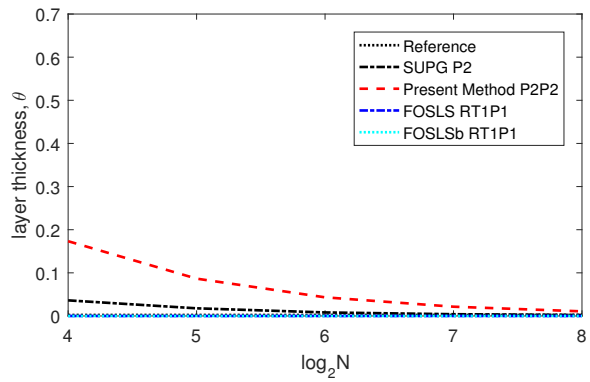
(c) Linear elements, $\varepsilon = 10^{-4}$.



(d) Quadratic elements, $\varepsilon = 10^{-4}$.

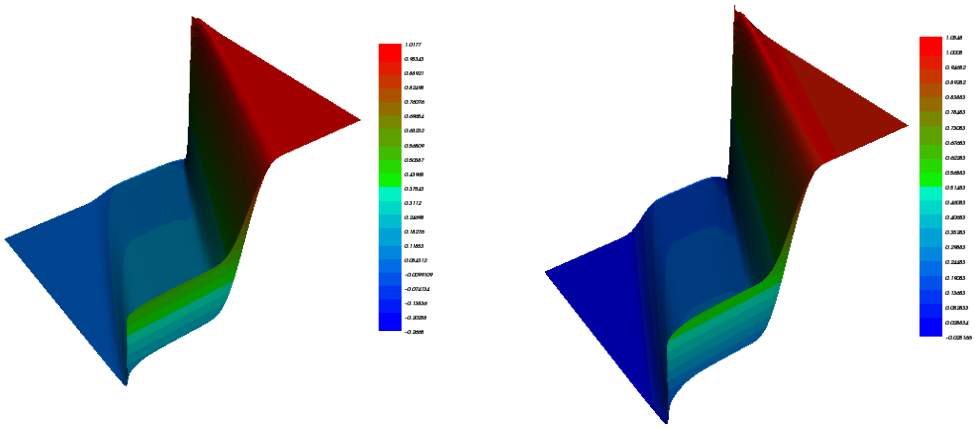


(e) Linear elements, $\varepsilon = 10^{-5}$.



(f) Quadratic elements, $\varepsilon = 10^{-5}$.

Fig. 6.10: Advection skew to the mesh, different values for ε: Outflow layer thickness, θ, for $0.1 < p(0.7, y) < 0.9$ with respect to the refinement level.



(a) FOSLS method.

(b) FOSLSb method.

Fig. 6.11: Elevations for FOSLS methods, $N = 2^8$, $\varepsilon = 10^{-3}$.

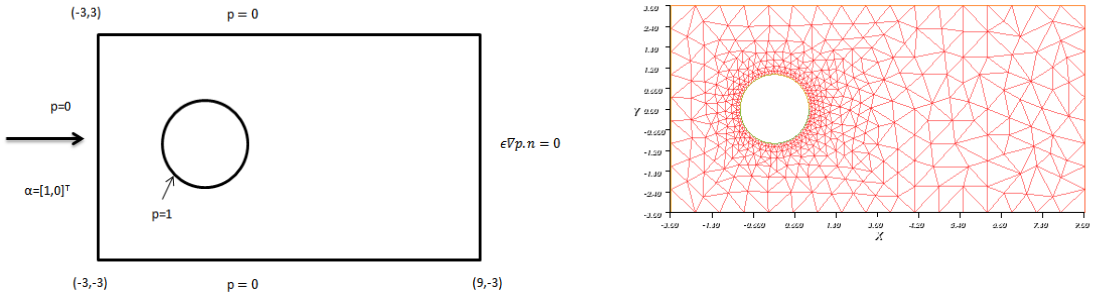
Table 6.1: Details of Hemker meshes

level	No of Triangles	No of Vertices	SUPG $\mathcal{P}_2$ DOFs	Present DOFs $\mathcal{P}_2\mathcal{P}_2$	FOSLS DOFs $RT1P1$	h_{min}
0	978	549	2076	4152	5559	0.098
1	3918	2079	8076	16152	21909	0.047
2	15522	8001	31524	63048	86091	0.023
3	61494	31227	123948	247896	339657	0.011
4	247542	124731	497004	994008	1364361	0.0056
5	988588	496214	1981016	3962032	5442994	0.0026
6	3951688	1979624	7910816	–	–	0.0012

6.3. The Hemker problem. The Hemker test has been used in numerous works as an example of a convection-dominated problem (see, e.g., [1]). The geometry and boundary conditions for this test case are depicted in Figure 6.12a, and we have used $f = \mu = 0$, $\boldsymbol{\alpha} = [1, 0]^T$, and $\varepsilon = 10^{-4}$. The convective field points towards the right of the domain. As a consequence, a boundary layer appears on the left-hand side of the circle, while two characteristic (interior) layers start from the top and bottom of the circle in the direction of the convection stretching out to the right-hand side.

The meshes used were generated from an initial unstructured grid depicted in Figure 6.12b. Successive refinements led to meshes whose characteristics are detailed in Table 6.1. When using linear elements we used meshes up to level 6 and with quadratic elements we used meshes up to level 5. For this problem we have not included a comparison with the MK method since several plots lie outside the scale of the plots shown. A reference solution for the Hemker problem with $\varepsilon = 10^{-4}$, obtained using the SUPG method with quadratic elements on a very fine mesh (level 6), is depicted in Figure 6.13a.

In Figure 6.13, we depict elevations of solutions given by the present method using



(a) Hemker test: geometry and boundary conditions.

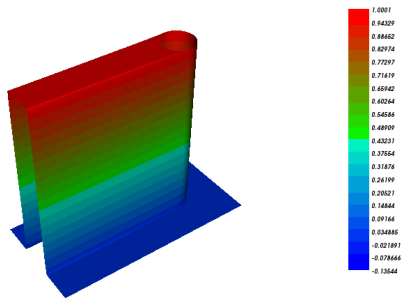
(b) Mesh for Hemker test, level 0.

Fig. 6.12: Hemker test details and initial mesh.

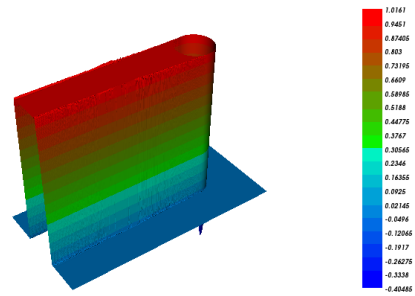
both linear and quadratic elements and also the solutions provided by both the FOSLS methods.

A more detailed comparison is shown in Figure 6.14, where we depict the cross-sections of the solutions along the lines $y = 1$ and $x = 4$. In this figure we use linear elements and also include both the reference solution and the SUPG solution on the same mesh using $\mathcal{P}_1$ elements. We can observe that the solutions provided by (3.1) and the one provided by the SUPG method are close to each other. We repeat this process for the quadratic elements and in Figure 6.15 we depict cross-sections of the present method, the reference solution, the solution given by the different versions of the FOSLS methods presented in subsection 6.1.3, and SUPG solutions on the same mesh. We observe that FOSLS fails to provide sharp layers. Close-ups of the regions near the layer on the left-hand side of the circle are shown in Figure 6.16. (Note that the FOSLS solutions do not appear on certain plots as they lie outside the range of the plot.)

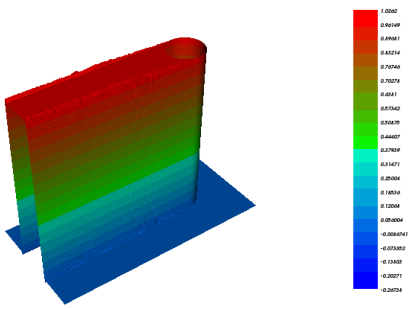
To carry out more quantitative comparisons, in Figure 6.17 we depict the error of the computed solution with respect to that of the reference solution on level 4 along the lines $y = 1$ and $x = 4$. We observe that the present method's results are comparable to the ones given by SUPG. The results of the other methods have been excluded since in some cases they lie outside the scale of the plot. The layer thickness using quadratic elements for all methods are depicted in Figure 6.18, where we confirm that the present method provides steeper layers than the other mixed approaches. Finally, in Figure 6.19, we plot the over- and undershoots of all methods tested. The lower undershoots that occur in FOSLS are consistent with the wider, more diffuse layers.



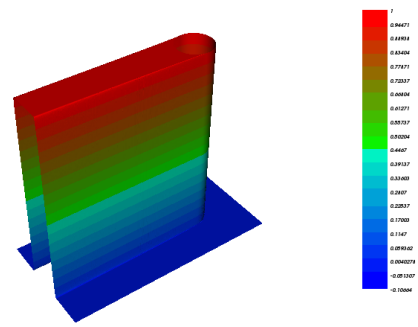
(a) SUPG Reference $\mathcal{P}_2$ solution.



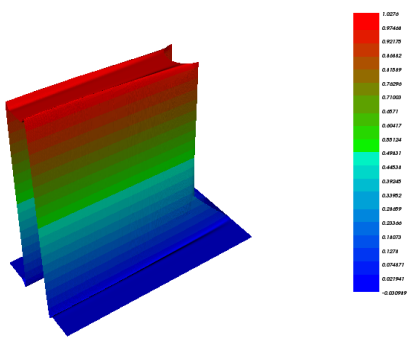
(b) SUPG $\mathcal{P}_1$ solution.



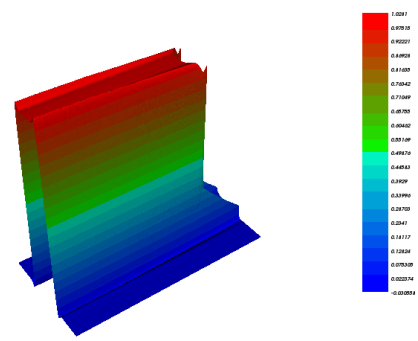
(c) Present Method $\mathcal{P}_1\mathcal{P}_1$ solution.



(d) Present Method $\mathcal{P}_2\mathcal{P}_2$ solution.

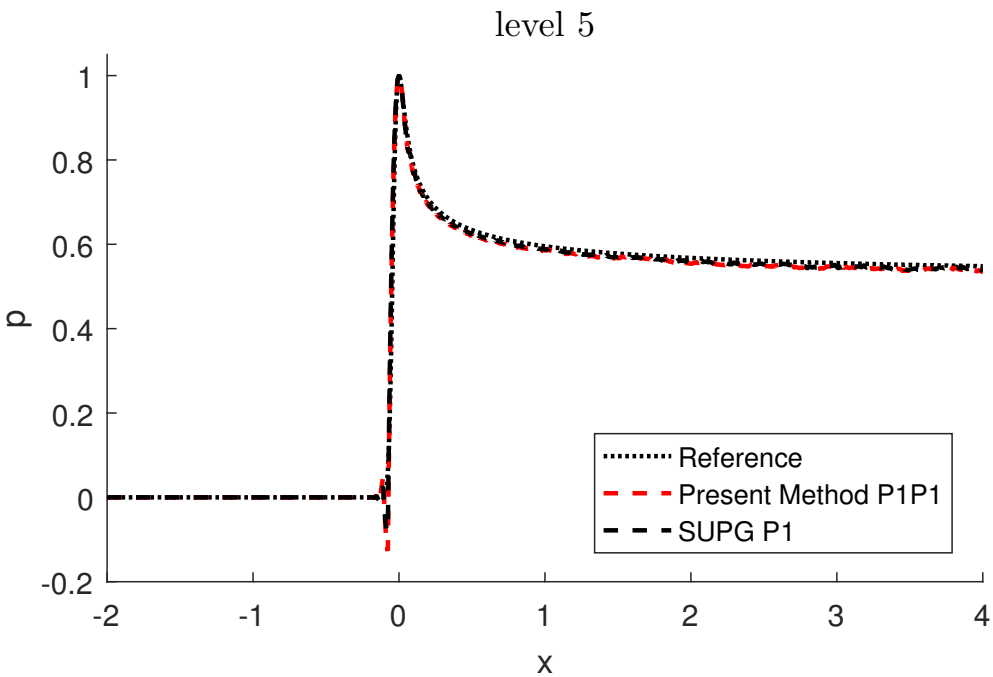


(e) FOSLS RT1P1 solution.

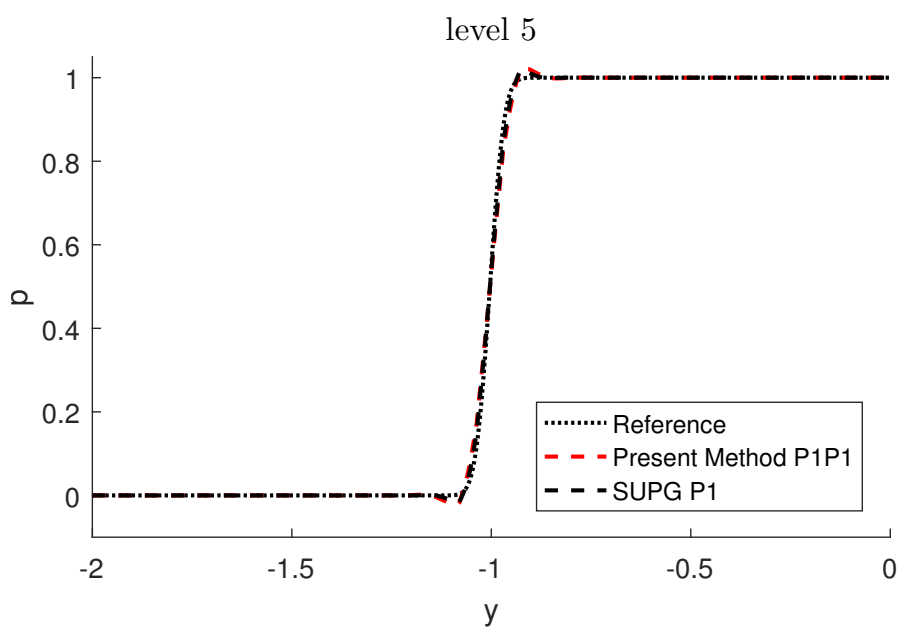


(f) FOSLSb RT1P1 solution.

Fig. 6.13: Hemker problem: Discrete solutions for level 5.

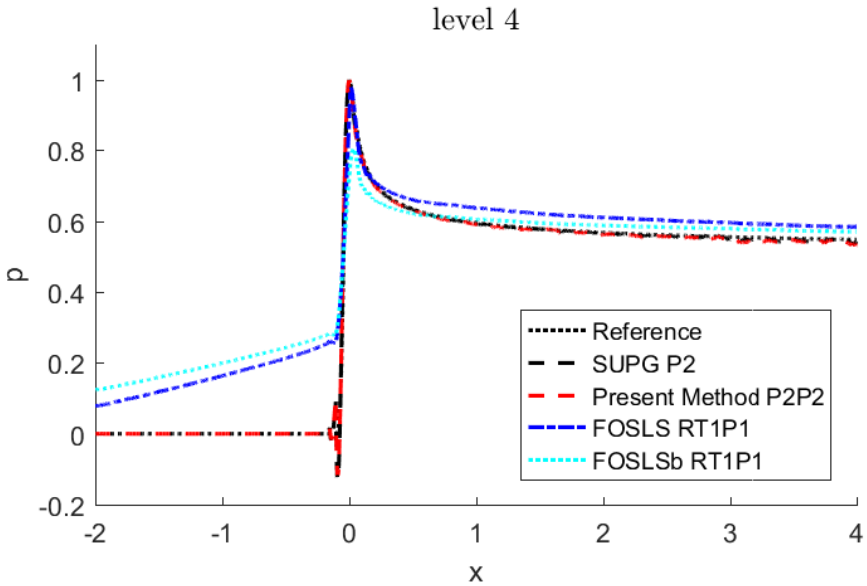


(a) x cross-section at $y = 1$.

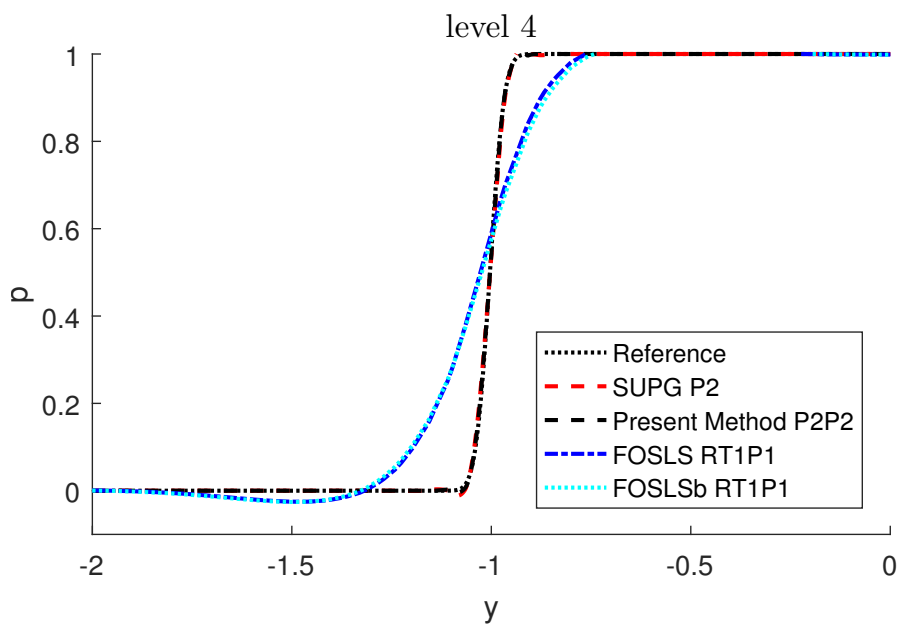


(b) y cross-section at $x = 4$.

Fig. 6.14: Hemker problem: Cross-sections using linear elements, level 5.

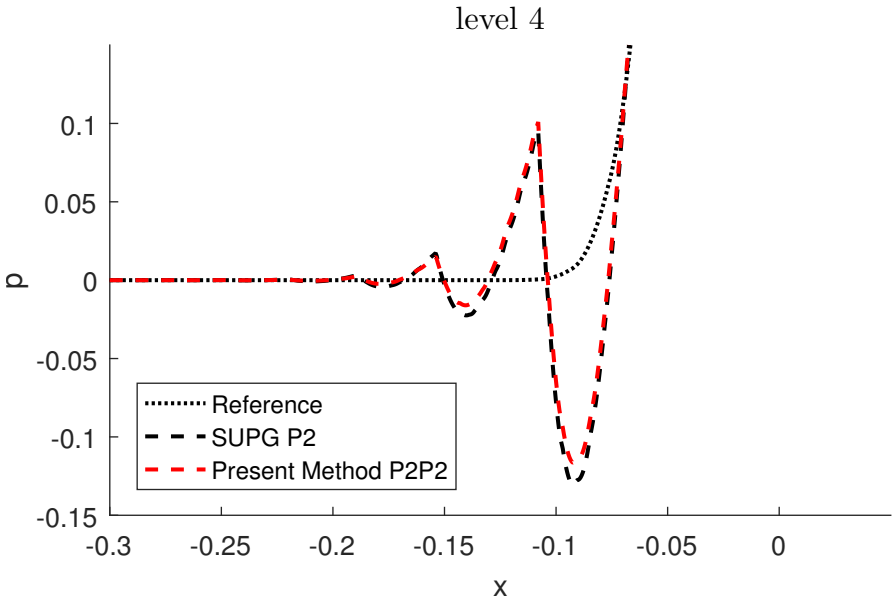


(a) x cross-section at $y = 1$.

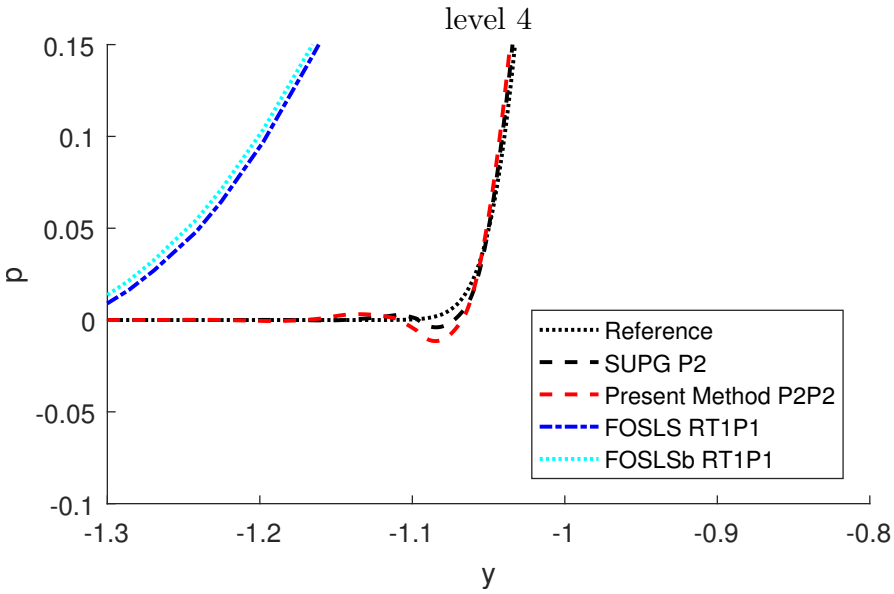


(b) y cross-section at $x = 4$.

Fig. 6.15: Hemker problem: Cross-sections using quadratic elements, level 4.

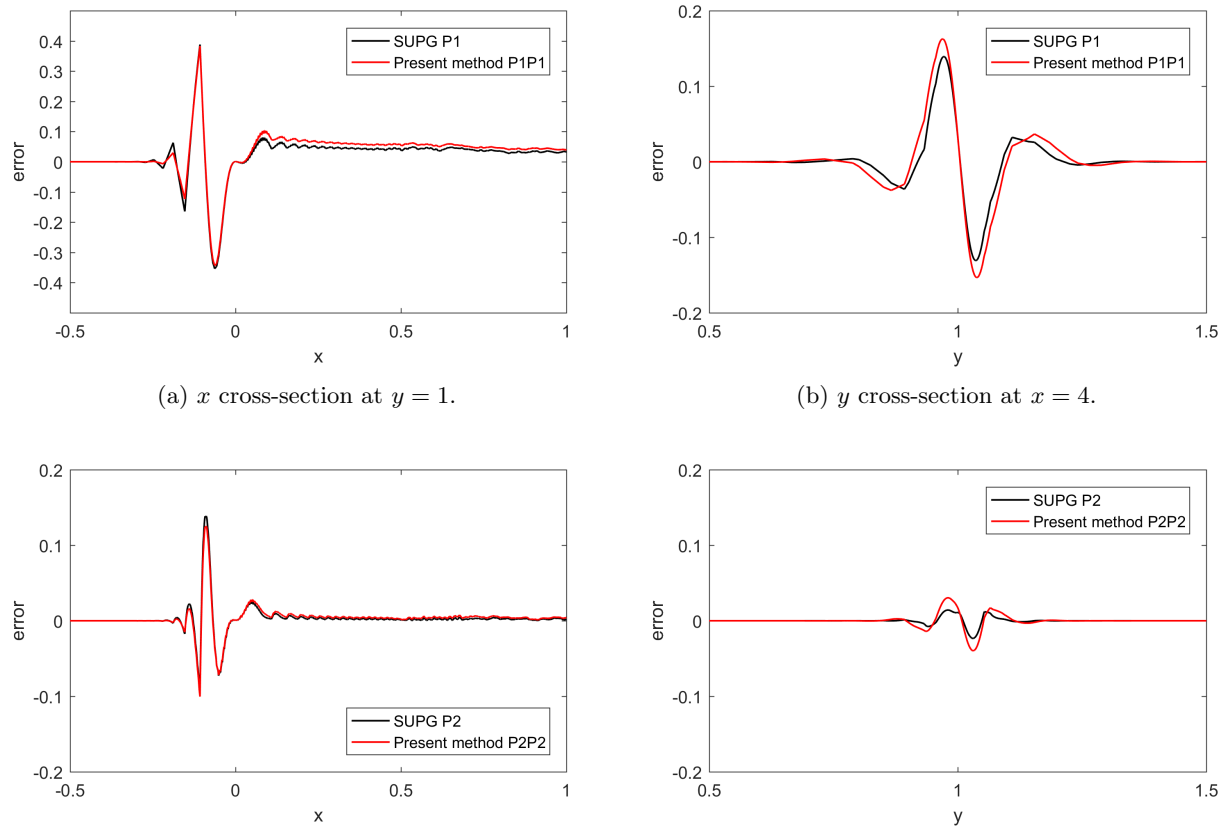


(a) x cross-section at $y = 1$.



(b) y cross-section at $x = 4$.

Fig. 6.16: Hemker problem: Close up of the cross-sections using quadratic elements, level 4.



(a) x cross-section at $y = 1$.

(b) y cross-section at $x = 4$.

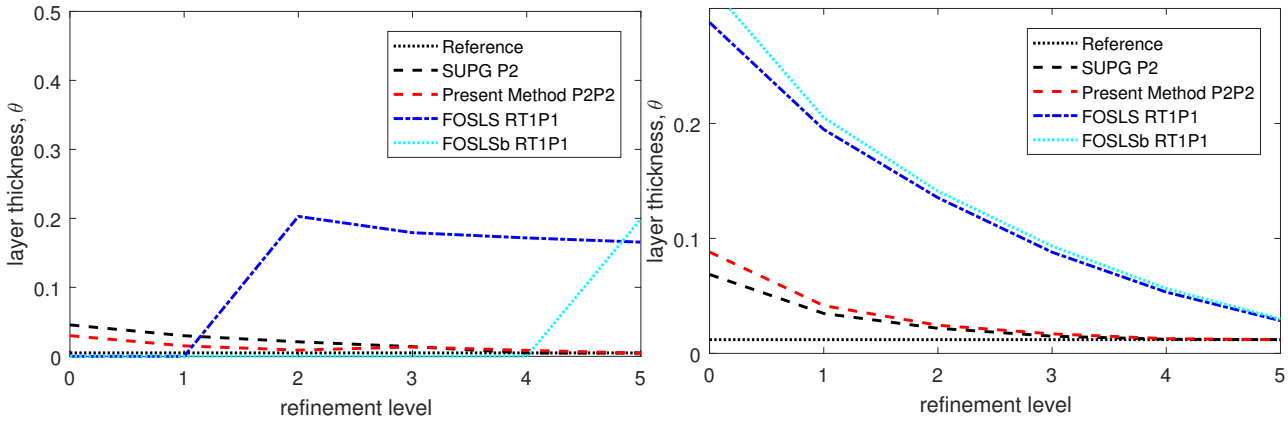
(c) x cross-section at $y = 1$.

(d) y cross-section at $x = 4$.

Fig. 6.17: Hemker problem: Error with respect to the reference solution for linear elements (top) and quadratic elements (bottom).

7. Conclusion. In this work a modified version of the Masud-Kwack method for the mixed form of the convection-diffusion-reaction equation was proposed. The main motivation for this modification was the possibility of performing its error analysis, but the modifications thus introduced led to a significant improvement in the quality of the numerical results. To stress this fact, a review of different mixed finite element methods for the convection-diffusion equation was presented, and their numerical performance was assessed using two classical benchamark problems. The conclusion is that the present method emerges as the one that produces the best numerical results amongst the mixed methods reviewed in this work.

A drawback of the present method is the fact that the outer boundary layer seems to be more diffused than the one provided by SUPG (the inner layers appear to be of comparable width). Thus, further investigations will include possible local adaptations of the stabilisation parameter in order to deal with this. In addition, the fact that the stability and convergence of the present method has been analysed opens the door to perform more refined numerical analyses, such as local convergence



(a) x cross-section at $y = 1$.

(b) y cross-section at $x = 4$.

Fig. 6.18: Hemker problem: Layer thickness, θ, using quadratic elements for solution with $0.1 < p < 0.9$.

analysis, or the proposal of a posteriori error estimators. These will be the topics of future research.

REFERENCES

[1] M. Augustin, A. Caiazzo, A. Fiebach, J. Fuhrmann, V. John, A. Linke, and R. Umla, *An assessment of discretizations for convection-dominated convection-diffusion equations*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 3395–3409.

[2] S. Badia and A. Hierro, *On monotonicity-preserving stabilized finite element approximations of transport problems*, SIAM J. Sci. Comput., 36 (2014), pp. A2673–A2697.

[3] G. R. Barrenechea, E. Burman, and F. Karakatsani, *Edge-based nonlinear diffusion for finite element approximations of convection-diffusion equations and its relation to algebraic flux-correction schemes*, Numer. Math., 135 (2017), pp. 521–545.

[4] G. R. Barrenechea, V. John, and P. Knobloch, *An algebraic flux correction scheme satisfying the discrete maximum principle and linearity preservation on general meshes*, Math. Models Methods Appl. Sci., 27 (2017), pp. 525–548.

[5] T. P. Barrios, J. M. Cascón, and M. González, *Augmented mixed finite element method for the Oseen problem: A priori and a posteriori error analyses*, Comput. Methods Appl. Mech. Engrg., 313 (2017), pp. 216 – 238.

[6] P. B. Bochev and M. D. Gunzburger, *Least-squares finite element methods*, Springer Science & Business Media, 2009.

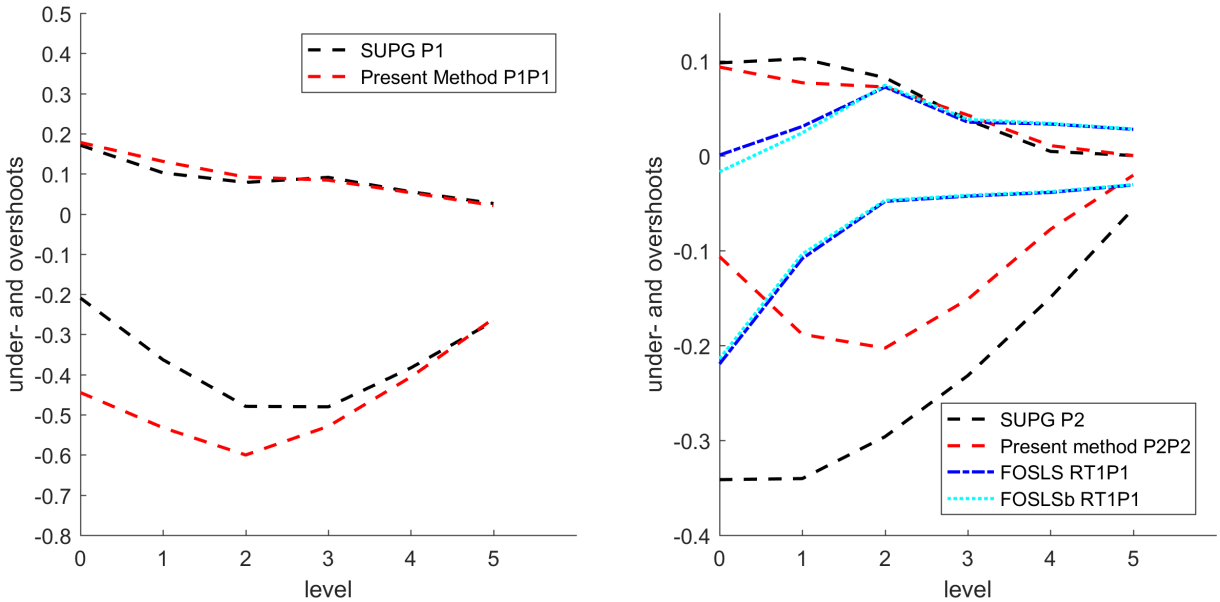
[7] F. Brezzi, L. D. Marini, S. Micheletti, P. Pietra, and R. Sacco, *Stability and error analysis of mixed finite-volume methods for advection dominated problems*, Comput. Math. Appl., 51 (2006), pp. 681–696.

[8] D. Broersen and R. Stevenson, *A robust Petrov-Galerkin discretisation of convection-diffusion equations*, Comput. Math. Appl., 68 (2014), pp. 1605–1618.

[9] D. Broersen and R. P. Stevenson, *A Petrov-Galerkin discretization with optimal test space of a mild-weak formulation of convection-diffusion equations in mixed form*, IMA J. Numer. Anal., 35 (2015), pp. 39–73.

[10] A. N. Brooks and T. J. R. Hughes, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259. FENOMECH '81, Part I (Stuttgart, 1981).

[11] E. Burman and P. Hansbo, *Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 1437–



(a) Linear elements.

(b) Quadratic elements.

Fig. 6.19: Hemker problem: Over- and undershoots.

1453.

[12] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations. I*, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.

[13] Z. Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations. II*, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[14] C. Carstensen, A. K. Dond, N. Nataraj, and A. K. Pani, *Error analysis of nonconforming and mixed FEMs for second-order linear non-selfadjoint and indefinite elliptic problems*, Numer. Math., 133 (2016), pp. 557–597.

[15] H. Chen, G. Fu, J. Li, and W. Qiu, *First order least squares method with weakly imposed boundary condition for convection dominated diffusion problems*, Comput. Math. Appl., 68 (2014), pp. 1635–1652.

[16] J. Douglas and J. E. Roberts, *Mixed Finite Element Methods for Second Order Elliptic Problems*, Mat. Aplic. Comp, 1 (1982), pp. 91–103.

[17] ———, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39–52.

[18] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.

[19] J. Fiard, T. Manteuffel, and S. McCormick, *First-order system least squares (FOSLS) for convection-diffusion problems: Numerical results*, SIAM J. Sci. Comput., 19 (1998), pp. 1958–1979.

[20] L. P. Franca, S. L. Frey, and T. J. R. Hughes, *Stabilized finite element methods. I. Application to the advective-diffusive model*, Comput. Methods Appl. Mech. Engrg., 95 (1992), pp. 253–276.

[21] M. González, J. Jansson, and S. Korotov, *A posteriori error analysis of a stabilized mixed FEM for convection-diffusion problems*, Discrete Contin. Dyn. Syst., (2015), pp. 525–532.

[22] J.-L. Guermond, M. Nazarov, B. Popov, and Y. Yang, *A second-order maximum principle preserving Lagrange finite element technique for nonlinear scalar conservation equations*,

SIAM J. Numer. Anal., 52 (2014), pp. 2163–2182.

[23] F. Hecht, *New development in FreeFem++*, J. Numer. Math., 20 (2012), pp. 251–265.

[24] P.-W. Hsieh and S.-Y. Yang, *On efficient least-squares finite element methods for convection-dominated problems*, Comput. Methods Appl. Mech. Engrg., 199 (2009), pp. 183–196.

[25] ———, *A novel least-squares finite element method enriched with residual-free bubbles for solving convection-dominated problems*, SIAM J. Sci. Comput., 32 (2010), pp. 2047–2073.

[26] V. John and P. Knobloch, *On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. I. A review*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 2197–2215.

[27] ———, *On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. II. Analysis for P_1 and Q_1 finite elements*, Comput. Methods Appl. Mech. Engrg., 197 (2008), pp. 1997–2014.

[28] P. Knobloch, *A generalization of the local projection stabilization for convection-diffusion-reaction equations*, SIAM J. Numer. Anal., 48 (2010), pp. 659–680.

[29] R. D. Lazarov, L. Tobiska, and P. S. Vassilevski, *Streamline diffusion least-squares mixed finite element methods for convection-diffusion problems*, East-West J. Numer. Math., 5 (1997), pp. 249–264.

[30] A. Masud and T. J. R. Hughes, *A stabilized mixed finite element method for Darcy flow*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4341–4370.

[31] A. Masud and J. Kwack, *A stabilized mixed finite element method for the first-order form of advection-diffusion equation*, Internat. J. Numer. Methods Fluids, 57 (2008), pp. 1321–1348.

[32] N. C. Nguyen, J. Peraire, and B. Cockburn, *An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations*, J. Comput. Phys., 228 (2009), pp. 3232–3254.

[33] P.-A. Raviart and J.-M. Thomas, *A mixed finite element method for 2-nd order elliptic problems*, in Mathematical aspects of finite element methods, Springer, 1977, pp. 292–315.

[34] H.-G. Roos, M. Stynes, and L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, vol. 24 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2008. Convection-diffusion-reaction and flow problems.

[35] J.-M. Thomas, *Mixed Finite Elements Methods for Convection-Diffusion Problems*, in Numerical Approximation of Partial Differential Equations –Selection of Papers Presented at the International Symposium on Numerical Analysis held at the Polytechnic University of Madrid, E. L. Ortiz, ed., vol. 133 of North-Holland Mathematics Studies, North-Holland, 1987, pp. 241 – 250.