Technische Universität Graz

Adaptive least-squares space-time finite element methods for convection-diffusion problems

C. Köthe, O. Steinbach

Berichte aus dem Institut für Angewandte Mathematik

Bericht 2025/6

Technische Universität Graz

Adaptive least-squares space-time finite element methods for convection-diffusion problems

C. Köthe, O. Steinbach

Berichte aus dem Institut für Angewandte Mathematik

Bericht 2025/6

Technische Universität Graz
Institut für Angewandte Mathematik
Steyrergasse 30
A 8010 Graz

WWW: `http://www.applied.math.tugraz.at`

Adaptive least-squares space-time finite element methods for convection-diffusion problems

Christian Köthe, Olaf Steinbach

Institut für Angewandte Mathematik, TU Graz,
Steyrergasse 30, 8010 Graz, Austria

In memoriam Raytcho D. Lazarov (1943–2024)

Abstract

In this paper we formulate and analyse adaptive (space-time) least-squares finite element methods for the solution of convection-diffusion equations. The convective derivative $\boldsymbol{v} \cdot \nabla u$ is considered as part of the total time derivative $\frac{d}{dt} u = \partial_t u + \boldsymbol{v} \cdot \nabla u$, and therefore we can use a rather standard stability and error analysis for related space-time finite element methods. For stationary problems we restrict the ansatz space $H^1_0(\Omega)$ such that the convective derivative is considered as an element of the dual $H^{-1}(\Omega)$ of the test space $H^1_0(\Omega)$, which also allows unbounded velocities $\boldsymbol{v}$. While the discrete finite element schemes are always unique solvable, the numerical solutions may suffer from a bad approximation property of the finite element space when considering convection dominated problems, i.e., small diffusion coefficients. Instead of adding suitable stabilization terms, we aim to resolve the solutions by using adaptive (space-time) finite element methods. For this we introduce a least-squares approach where the discrete adjoint defines local a posteriori error indicators to drive an adaptive scheme. Numerical examples illustrate the theoretical considerations.

Key words: convection-diffusion, least-squares methods, space-time FEM, adaptivity
AMS subject classifications: 65M60, 65M12, 65M50, 65N30, 65N12, 65N50

1 Introduction

As documented by Bochev and Gunzburger [7], least-squares finite element methods are a well established approach for the numerical solution of second order partial differential equations. In most cases, the partial differential equation is rewritten as first order system, and the residuals of both equations are minimized, using appropriate norms and weights. For early contributions of R. Lazarov using the $L^2(\Omega)$ norm for both residuals, see [15, 36],

and [10, 11] when considering the residual of the equilibrium equation in $H^{-1}(\Omega)$. Later on, see [31], a least-squares formulation in $L^2(\Omega)$ for a first-order system for a convection dominated convection-diffusion equation was analyzed, for a related minimization approach in $H^{-1}(\Omega)$, see [30]. For more recent contributions on first order least-squares systems we refer to, e.g., [23, 24, 38].

For a stable and accurate numerical solution of diffusion-convection-reaction equations there exists a huges amount of literature, here we mention, e.g., [2, 4, 5, 12, 17, 19, 22, 37], just to name a few. In the particular context of space-time variational formulations we refer to, e.g., [6, 13, 42].

In our recent work [29] we have formulated and analysed a least-squares approach for the numerical solution of rather general operator equations, including elliptic, parabolic, and hyperbolic partial differential equations, for related work on parabolic evolution equations, see also [1, 18, 33, 41]. Instead of rewriting a second order partial differential equation as first order system, which is then solved by a least-squares approach, we consider a least-squares approach for the original equation minimizing the residual in appropriate norms which are induced by elliptic and self-adjoint operators. This requires the solution of a saddle point variational formulation, where we can use standard arguments as known in mixed finite element methods [8]. While the adjoint variable turns out to be zero in the continuous setting, its discrete approximation can be used to define local a posteriori error indicators to drive an adaptive scheme.

In this work we extend the approach of [29] to solve stationary and instationary convection-diffusion equations by using adaptive (space-time) finite element methods. For parabolic evolution equations we combine the convective derivative with the partial time derivative and introduce the total time derivative. The resulting and standard space-time finite element method can be analyzed as in [25, 40], but here we present a more general proof of surjectivity when a convective derivative appears. For an adaptive space-time finite element scheme we then apply the least-squares approach as in [29]. In the case of an elliptic convection-diffusion equation, and motivated by the parabolic case, we consider the first order convective term in the dual $H^{-1}(\Omega)$ of the variational test space $H^1_0(\Omega)$. This allows to consider even unbounded velocities. For a direct finite element discretization we present a related numerical analysis showing the relations between the size of the diffusion coefficient, and the finite element mesh size which has to be sufficiently small in the case of convection dominated problems. Instead of using appropriate stabilization techniques, here we aim to construct accurate solutions by using an adaptive least-squares approach.

2 Stationary convection-diffusion problems

For a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, we consider, as a model problem, the Dirichlet boundary value problem for the convection-diffusion equation,

$$-\mathrm{div}[\alpha(x)\nabla u(x)] + \boldsymbol{v}(x) \cdot \nabla u(x) = f(x) \quad \text{for } x \in \Omega, \quad u(x) = 0 \quad \text{for } x \in \partial\Omega, \qquad (2.1)$$

where $f(x)$, $x \in \Omega$, is given. We assume that the diffusion coefficient $\alpha(x)$ is bounded and strictly positive, i.e.,

$$0 < \underline{\alpha} \leq \alpha(x) \leq \overline{\alpha} \quad \text{for all } x \in \Omega. \tag{2.2}$$

Assumptions on the given velocity field $\boldsymbol{v}$ will be specified later on. Note that in (2.1) we may add some reaction term $c(x)u(x)$ with some non-negative function $c(x)$, $x \in \Omega$, and we may consider inhomogeneous Dirichlet boundary conditions, or boundary conditions of mixed type, including Neumann and Robin type boundary conditions.

2.1 Variational formulation

The variational formulation of the Dirichlet boundary value problem (2.1) is to find $u \in X$ such that

$$b(u,q) := \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla q(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, q(x)\, dx = \int_\Omega f(x) q(x)\, dx \tag{2.3}$$

is satisfied for all $q \in Y := H^1_0(\Omega)$, where the related energy norm of the test space Y is given by

$$\|q\|^2_Y := \int_\Omega \alpha(x)\, |\nabla q(x)|^2\, dx,$$

and for which we have the norm equivalence inequalities

$$\sqrt{\underline{\alpha}}\, \|\nabla q\|_{L^2(\Omega)} \leq \|q\|_Y \leq \sqrt{\overline{\alpha}}\, \|\nabla q\|_{L^2(\Omega)} \quad \text{for all } q \in Y = H^1_0(\Omega). \tag{2.4}$$

According to the variational formulation (2.3) we assume $f \in H^{-1}(\Omega) := [H^1_0(\Omega)]^*$ with the norm

$$\|f\|_{Y^*} := \sup_{0 \neq q \in Y} \frac{\langle f, q \rangle_\Omega}{\|q\|_Y}, \tag{2.5}$$

using the duality pairing $\langle f, q \rangle_\Omega$ for $f \in Y^*$ and $q \in Y$ as extension of the $L^2(\Omega)$ inner product. While the most standard choice for the ansatz space is $X = H^1_0(\Omega)$, here we will use a slightly different approach. We introduce the velocity dependent ansatz space $X_v := \{u \in Y : \boldsymbol{v} \cdot \nabla u \in Y^*\} \subset Y$, with the graph norm

$$\|u\|_{X_v} := \sqrt{\|u\|^2_Y + \|\boldsymbol{v} \cdot \nabla u\|^2_{Y^*}} = \sqrt{\|u\|^2_Y + \|w_u\|^2_Y},$$

where the Riesz representant $w_u \in Y$ is the unique solution of the variational problem

$$a(w_u, q) := \int_\Omega \alpha(x)\, \nabla w_u(x) \cdot \nabla q(x)\, dx = \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, q(x)\, dx \quad \text{for all } q \in Y. \tag{2.6}$$

Depending on the regularity of the given velocity $\boldsymbol{v}$, the condition $\boldsymbol{v} \cdot \nabla u \in Y^*$ may result in the fact that $X_v \subset Y$ is a real subspace, i.e., $X_v \neq Y$. At this time we only assume that

X_v is not empty. By the definition of the underlying function spaces we then conclude

$$\begin{aligned} |b(u,q)| &= \left| \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla q(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, q(x)\, dx \right| \\ &= \left| \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla q(x)\, dx + \int_\Omega \alpha(x)\, \nabla w_u(x) \cdot \nabla q(x)\, dx \right| \\ &\leq \Big[\|u\|_Y + \|w_u\|_Y \Big]\, \|q\|_Y \\ &\leq \sqrt{2}\, \sqrt{\|u\|_Y^2 + \|w_u\|_Y^2}\, \|q\|_Y \;=\; \sqrt{2}\, \|u\|_{X_v} \|q\|_Y \end{aligned} \tag{2.7}$$

for all $(u,q) \in X_v \times Y$, i.e., boundedness of the bilinear form $b(u,q)$. A first rather standard assumption on the given velocity $\boldsymbol{v}$ is as follows.

Lemma 2.1 *Assume* $\operatorname{div} \boldsymbol{v}(x) \leq 0$ *for almost all* $x \in \Omega$. *Then,*

$$\int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, u(x)\, dx \geq 0 \quad \textit{for all } u \in X_v \subset Y = H_0^1(\Omega). \tag{2.8}$$

Proof. For $u \in X_v \subset Y$ we first note that the left hand side in (2.8) is bounded. Using integration by parts for $u \in X_v \subset Y = H_0^1(\Omega)$,

$$\begin{aligned} \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, u(x)\, dx &= \sum_{k=1}^n \int_\Omega \partial_{x_k} u(x)\, v_k(x) u(x)\, dx \\ &= -\sum_{k=1}^n \int_\Omega u(x)\, \partial_{x_k} [v_k(x) u(x)]\, dx \\ &= -\sum_{k=1}^n \int_\Omega \Big([u(x)]^2 \partial_{x_k} v_k(x) + u(x)\, v_k(x)\, \partial_{x_k} u(x) \Big)\, dx, \end{aligned}$$

i.e.,

$$2 \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, u(x)\, dx = -\int_\Omega [u(x)]^2 \operatorname{div} \boldsymbol{v}(x)\, dx \;\geq\; 0.$$

■

In order to ensure uniqueness for the solution of the variational formulation (2.3) we need to have the following result.

Lemma 2.2 *Assume* $\operatorname{div} \boldsymbol{v}(x) \leq 0$ *for almost all* $x \in \Omega$. *Then the bilinear form* $b(\cdot,\cdot)$ *as defined in* (2.3) *satisfies the inf-sup stability condition*

$$\|u\|_{X_v} \leq \sup_{0 \neq q \in Y} \frac{b(u,q)}{\|q\|_Y} \quad \textit{for all } u \in X_v. \tag{2.9}$$

Proof. For an arbitrary but fixed $u \in X_v \subset Y$ and the solution $w_u \in Y$ of the variational formulation (2.6) we define $q_u := u + w_u \in Y$, and we obtain

$$\begin{aligned} b(u, q_u) &= \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla q_u(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, q_u(x)\, dx \\ &= \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla q_u(x)\, dx + \int_\Omega \alpha(x)\, \nabla w_u(x) \cdot \nabla q_u(x)\, dx \\ &= \int_\Omega \alpha(x)\, \nabla [u(x) + w_u(x)] \cdot \nabla q_u\, dx \\ &= \int_\Omega \alpha(x)\, |\nabla q_u(x)|^2\, dx \;=\; \|q_u\|_Y^2 \,. \end{aligned}$$

On the other hand we have, using (2.8),

$$\begin{aligned} \|q_u\|_Y^2 &= \|u + w_u\|_Y^2 \\ &= \int_\Omega \alpha(x)\, |\nabla_x [u(x) + w_u(x)]|^2\, dx \\ &= \int_\Omega \alpha(x)\, |\nabla_x u(x)|^2\, dx + \int_\Omega \alpha(x)\, |\nabla_x w_u(x)|^2\, dx + 2 \int_\Omega \alpha(x)\, \nabla_x w_u(x) \cdot \nabla u(x)\, dx \\ &= \|u\|_Y^2 + \|w_u\|_Y^2 + 2 \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, u(x)\, dx \\ &\geq \|u\|_{X_v}^2 . \end{aligned}$$

Hence, we have shown the inf-sup stability condition

$$\|u\|_{X_v} \leq \frac{b(u, q_u)}{\|q_u\|_Y} \leq \sup_{0 \neq q \in Y} \frac{b(u, q)}{\|q\|_Y} \quad \text{for all } u \in X_v .$$

■

It remains to prove solvability of the variational formulation (2.3), i.e., surjectivity of the related operator $B : X_v \to Y^*$ which is defined by

$$\langle Bu, q \rangle_\Omega := b(u, q) \quad \text{for all } (u, q) \in X_v \times Y .$$

Recall that the definition of X_v depends on the regularity of the velocity field $\boldsymbol{v}$. If the velocity is bounded, which is a standard assumption in many applications, i.e., $\|\boldsymbol{v}(x)\|_2 \leq c_v$ for almost all $x \in \Omega$, we then obtain

$$\|\boldsymbol{v} \cdot \nabla u\|_{Y^*} = \sup_{0 \neq q \in Y} \frac{\langle \boldsymbol{v} \cdot \nabla u, q \rangle_\Omega}{\|q\|_Y} \;\leq\; c_v \, \|\nabla u\|_{L^2(\Omega)} \sup_{0 \neq q \in Y} \frac{\|q\|_{L^2(\Omega)}}{\|q\|_Y} \;\leq\; c \, \|u\|_Y ,$$

where we have used (2.4), and Friedrich's inequality. In this case we conclude that $\|u\|_Y$ defines an equivalent norm in X_v, and $X_v = Y = H_0^1(\Omega)$ follows for all bounded velocities $\boldsymbol{v}$. In any case, the assumption on the boundedness of $\boldsymbol{v}$ can be relaxed as follows.

Lemma 2.3 *Assume*

$$c_v := \sup_{0 \neq q \in H_0^1(\Omega)} \frac{\|q\,\boldsymbol{v}\|_{L^2(\Omega)}}{\|\nabla q\|_{L^2(\Omega)}} < \infty. \tag{2.10}$$

Then,

$$\|\boldsymbol{v} \cdot \nabla u\|_{Y^*} \leq \frac{c_v}{\underline{\alpha}} \|u\|_Y \quad \textit{for all } u \in X_v. \tag{2.11}$$

Proof. When using the definition (2.5) we have

$$\begin{aligned} \|\boldsymbol{v} \cdot \nabla u\|_{Y^*} &= \sup_{0 \neq q \in Y} \frac{\langle \boldsymbol{v} \cdot \nabla u, q \rangle_\Omega}{\|q\|_Y} = \sup_{0 \neq q \in Y} \frac{\langle \nabla u, q\,\boldsymbol{v} \rangle_\Omega}{\|q\|_Y} \\ &\leq \|\nabla u\|_{L^2(\Omega)} \sup_{0 \neq q \in Y} \frac{\|q\,\boldsymbol{v}\|_{L^2(\Omega)}}{\|q\|_Y} \\ &\leq \frac{1}{\underline{\alpha}} \|u\|_Y \sup_{0 \neq q \in Y} \frac{\|q\,\boldsymbol{v}\|_{L^2(\Omega)}}{\|\nabla q\|_{L^2(\Omega)}} = \frac{c_v}{\underline{\alpha}} \|u\|_Y . \end{aligned}$$

■

As a consequence of (2.11) we conclude the norm equivalence inequalities

$$\|u\|_Y^2 \leq \|u\|_{X_v}^2 = \|u\|_Y^2 + \|\boldsymbol{v} \cdot \nabla u\|_{Y^*}^2 \leq \left(1 + \frac{c_v^2}{\underline{\alpha}^2}\right) \|u\|_Y^2$$

for all $u \in X_v$, i.e., $\|u\|_Y$ defines an equivalent norm in X_v, and hence, $X_v = Y = H_0^1(\Omega)$ follows when (2.10) is satisfied. With this we are now in the position to state the surjectivity of B.

Lemma 2.4 *Assume* $\operatorname{div} \boldsymbol{v}(x) \leq 0$ *for almost all* $x \in \Omega$ *and* (2.10) *to be satisfied. For any* $p \in Y \backslash \{0\}$ *there exists a* $u_p \in X_v$ *such that*

$$b(u_p, p) > 0 .$$

Proof. Due to (2.10) we have $0 \neq p \in Y = X_v$, and hence we can choose $u_p = p \in X_v$ to conclude

$$b(u_p, p) = \int_\Omega \alpha(x)\, \nabla p(x) \cdot \nabla p(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla p(x)\, p(x)\, dx \geq \|p\|_Y^2 > 0.$$

■

Now we are in a position to state the unique solvability result for the solution of the variational formulation (2.3).

Lemma 2.5 *Assume* (2.10) *and* $\operatorname{div} \boldsymbol{v}(x) \leq 0$ *for almost all* $x \in \Omega$. *For any* $f \in Y^*$ *there exists a unique solution* $u \in X_v$ *of the variational formulation* (2.3) *satisfying*

$$\|u\|_{X_v} \leq \|f\|_{Y^*} .$$

Proof. Since the bilinear $b(u,q)$ is bounded for all $(u,q) \in X_v \times Y$, satisfies the inf-sup stability condition (2.9), and is surjective, all assumptions of the Babuška–Nečas theory, e.g., [3, 9, 21, 34], are satisfied, and unique solvability of (2.3) follows. Moreover, using (2.9) we have

$$\|u\|_{X_v} \leq \sup_{0 \neq q \in Y} \frac{b(u,q)}{\|q\|_Y} = \sup_{0 \neq q \in Y} \frac{\langle f, q \rangle_\Omega}{\|q\|_Y} = \|f\|_{Y^*} .$$

■

Example 2.1 *For $n = 1$, $\Omega = (0,1)$ and $\boldsymbol{v}(x) = v(x)$ we consider the Dirichlet boundary value problem*

$$-u''(x) + v(x)u'(x) = f(x) \quad \text{for } x \in (0,1), \quad u(0) = u(1) = 0,$$

and the variational formulation to find $u \in X_v$ such that

$$b(u,q) := \int_0^1 u'(x)q'(x)\,dx + \int_0^1 v(x)u'(x)q(x)\,dx = \int_0^1 f(x)q(x)\,dx \tag{2.12}$$

is satisfied for all $q \in Y = H_0^1(0,1)$. For $q \in H_0^1(0,1)$ we can write

$$q(x) = \int_0^x q'(s)\,ds,$$

and

$$\begin{aligned} |v(x)\,q(x)|^2 &= \left| v(x) \int_0^x q'(s)\,ds \right|^2 \\ &\leq [v(x)]^2 \int_0^x ds \int_0^x [q'(s)]^2\,ds \;\leq\; x\,[v(x)]^2 \int_0^1 [q'(s)]^2\,ds, \end{aligned}$$

and integration over $x \in (0,1)$ gives

$$\|v\,q\|^2_{L^2(0,1)} \leq \int_0^1 x\,[v(x)]^2 dx\; \|q'\|^2_{L^2(0,1)},$$

i.e.,

$$c_v \leq \left(\int_0^1 x\,[v(x)]^2\,dx \right)^{1/2} . \tag{2.13}$$

Hence we conclude unique solvability of the variational formulation (2.12) for velocities $v(x)$ satisfying (2.13). For example, we can consider

$$v(x) = \frac{1}{\sqrt{x}}, \quad c_v \leq 1, \quad \operatorname{div} \boldsymbol{v}(x) = \frac{d}{dx}\frac{1}{\sqrt{x}} = -\frac{1}{2}\,x^{-3/2} < 0 \quad \text{for } x \in (0,1).$$

But for $v(x) = 1/x$ we can not apply the above estimates, although we have

$$\operatorname{div} \boldsymbol{v}(x) = \frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2} < 0 \quad \textit{for } x \in (0,1).$$

While the inf-sup condition (2.9) remains valid for $v(x) = 1/x$, we have to prove surjectivity in a different way: For $0 \neq p \in H_0^1(0,1)$ we define $u_p(x) := x\,p(x)$ with $u_p \in Y$. It remains to consider

$$\|v\,u_p'\|_{Y^*} = \sup_{0\neq q\in Y} \frac{\langle v\,u_p', q\rangle_{(0,1)}}{\|q\|_Y}\,,$$

where we have

$$\begin{aligned}\langle v\,u_p', q\rangle_{(0,1)} &= \int_0^1 \frac{1}{x}[p(x) + xp'(x)]q(x)\,dx \\ &= \int_0^1 \frac{1}{x}\,p(x)\,q(x)\,dx + \int_0^1 p'(x)q(x)\,dx\,.\end{aligned}$$

Note that the second term can be bounded as

$$\begin{aligned}\left|\int_0^1 p'(x)q(x)\,dx\right| &\leq \left(\int_0^1 [p'(x)]^2 dx\right)^{1/2}\left(\int_0^1 [q(x)]^2 dx\right)^{1/2} \\ &= \|p\|_Y \left(\int_0^1 \left[\int_0^x q'(s)\,ds\right]^2 dx\right)^{1/2} \leq \|p\|_Y \left(\int_0^1 \int_0^x 1^2 ds \int_0^x [q'(s)]^2 ds\,dx\right)^{1/2} \\ &\leq \frac{1}{2}\,\|p\|_Y \|q\|_Y\,.\end{aligned}$$

In a similar way we have, recall $p, q \in H_0^1(0,1)$,

$$\begin{aligned}\left|\int_0^1 \frac{1}{x}\,p(x)\,q(x)\,dx\right| &= \left|\int_0^1 \frac{1}{\sqrt{x}}\int_0^x p'(s)\,ds\,\frac{1}{\sqrt{x}}\int_0^x q'(s)\,ds\,dx\right| \\ &\leq \left(\int_0^1 \frac{1}{x}\left[\int_0^x p'(s)\,ds\right]^2 dx\right)^{1/2}\left(\int_0^1 \frac{1}{x}\left[\int_0^x q'(s)\,ds\right]^2 dx\right)^{1/2} \\ &\leq \left(\int_0^1 \frac{1}{x}\int_0^x 1^2 ds \int_0^x [p'(s)]^2\,ds\,dx\right)^{1/2}\left(\int_0^1 \frac{1}{x}\int_0^x 1^2 ds \int_0^x [q'(s)]^2\,ds\,dx\right)^{1/2} \\ &\leq \|p\|_Y \|q\|_Y\,.\end{aligned}$$

Hence we conclude

$$\|v\,u_p'\|_{Y^*} \leq \frac{3}{2}\,\|p\|_Y,$$

i.e., $u_p \in X_v$. *Then we obtain*

$$\begin{aligned} b(u_p,p) &= \int_0^1 u_p'(x)p'(x)\,dx + \int_0^1 \frac{1}{x}\,u_p'(x)\,p(x)\,dx \\ &= \int_0^1 [p(x)+xp'(x)]p'(x)\,dx + \int_0^1 \frac{1}{x}\,[p(x)+xp'(x)]p(x)\,dx \\ &= 2\int_0^1 p(x)p'(x)\,dx + \int_0^1 x\,[p'(x)]^2\,dx + \int_0^1 \frac{1}{x}\,[p(x)]^2 dx \\ &= \int_0^1 x\,[p'(x)]^2\,dx + \int_0^1 \frac{1}{x}\,[p(x)]^2 dx \;>\; 0, \end{aligned}$$

i.e., surjectivity and therefore unique solvability of (2.12) *follows also for* $v(x) = 1/x$.

Remark 2.1 *One possibility to prove the surjectivity result for the bilinear form* $b(u,q)$ *is to assume* (2.10) *to be able to conclude* $X_v = Y$, *which corresponds to the standard setting. As we have seen in the previous example,* (2.10) *is not necessary to establish surjectivity. At this time it remains open to generalize this result to more general situations. In any case, for the numerical realisation, surjectivity follows from injectivity, and assumption* (2.10) *is not needed.*

2.2 Finite element discretization

Let $X_H = Y_H = S_H^1(\Omega) \cap H_0^1(\Omega) = \mathrm{span}\{\varphi_k\}_{k=1}^M$ be a standard finite element space of, e.g., piecewise linear continuous basis functions φ_k which are defined with respect to some admissible decomposition of Ω into shape regular simplicial finite elements τ_ℓ of local mesh size H_ℓ, $\ell = 1,\ldots,N$. For any given $u \in X_v$, and similar as in (2.6), we define $w_{u,H} \in Y_H$ as the unique solution of the variational problem

$$\int_\Omega \alpha(x)\,\nabla w_{u,H}(x)\cdot\nabla q_H(x)\,dx = \int_\Omega \boldsymbol{v}(x)\cdot\nabla u(x)\,q_H(x)\,dx \quad \text{for all } q_H \in Y_H. \tag{2.14}$$

Hence, we can introduce the discrete norm

$$\|u\|_{X_v,H} := \sqrt{\|u\|_Y^2 + \|w_{u,H}\|_Y^2} \le \sqrt{\|u\|_Y^2 + \|w_u\|_Y^2} = \|u\|_{X_v} \quad \text{for all } u \in X_v.$$

With respect to this discrete norm and as in the continuous case, see Lemma 2.5, we can prove a discrete inf-sup stability condition:

Lemma 2.6 *Assume* $\operatorname{div}\boldsymbol{v}(x) \le 0$ *for almost all* $x \in \Omega$. *Let* $X_H \subset X_v$ *and* $Y_H \subset Y$ *be conforming finite element spaces satisfying* $X_H \subseteq Y_H$. *Then there holds the discrete inf-sup stability condition*

$$\|u_H\|_{X_v,H} \le \sup_{0\neq q_H\in Y_H} \frac{b(u_H,q_H)}{\|q_H\|_Y} \quad \textit{for all } u_H \in X_H. \tag{2.15}$$

Proof. For $u_H \in X_H$ we define $w_{u_H,H} \in Y_H$ as the unique solution of the variational formulation

$$\int_\Omega \alpha(x)\, \nabla w_{u_H,H}(x) \cdot \nabla q_H(x)\, dx = \int_\Omega \boldsymbol{v}(x) \cdot \nabla u_H(x)\, q_H(x)\, dx \quad \text{for all } q_H \in Y_H. \tag{2.16}$$

As in the continuous case we now define $q_{u_H,H} := u_H + w_{u_H,H} \subset Y_H$ to conclude

$$\begin{aligned} b(u_H, q_{u_H,H}) &= \int_\Omega \alpha(x)\, \nabla u_H(x) \cdot \nabla q_{u_H,H}(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u_H(x)\, q_{u_H,H}(x)\, dx \\ &= \int_\Omega \alpha(x)\, \nabla u_H(x) \cdot \nabla q_{u_H,H}(x)\, dx + \int_\Omega \alpha(x)\, \nabla w_{u_H,H}(x) \cdot \nabla q_{u_H,H}(x)\, dx \\ &= \int_\Omega \alpha(x)\, \nabla [u_H(x) + w_{u_H,H}(x)] \cdot \nabla q_{u_H,H}\, dx \\ &= \int_\Omega \alpha(x)\, |\nabla q_{u_H,H}(x)|^2\, dx \;=\; \|q_{u_H,H}\|_Y^2\,, \end{aligned}$$

and

$$\begin{aligned} \|q_{u_H,H}\|_Y^2 &= \|u_H + w_{u_H,H}\|_Y^2 \;=\; \int_\Omega \alpha(x)\, |\nabla_x [u_H(x) + w_{u_H,H}(x)]|^2\, dx \\ &= \int_\Omega \alpha(x)\, |\nabla_x u_H(x)|^2\, dx + \int_\Omega \alpha(x)\, |\nabla_x w_{u_H,H}(x)|^2\, dx \\ &\qquad + 2 \int_\Omega \alpha(x)\, \nabla_x w_{u_H,H}(x) \cdot \nabla u_H(x)\, dx \\ &= \|u_H\|_Y^2 + \|w_{u_H,H}\|_Y^2 + 2 \int_\Omega \boldsymbol{v}(x) \cdot \nabla u_H(x)\, u_H(x)\, dx \;\geq\; \|u_H\|_{X_v,H}^2, \end{aligned}$$

i.e., the assertion follows. ∎

Let $u_H \in X_H$ be the unique Galerkin solution related to the variational formulation (2.3), satisfying

$$\int_\Omega \alpha(x)\, \nabla u_H(x) \cdot \nabla q_H(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u_H(x)\, q_H(x)\, dx = \int_\Omega f(x)\, q_H(x)\, dx \tag{2.17}$$

for all $q_H \in Y_H$. The Galerkin variational formulation (2.17) is equivalent to an algebraic system of linear equations, $B_H \underline{u} = \underline{f}$, where the stiffness matrix is given by its entries

$$B_H[j,k] = \int_\Omega \alpha(x)\, \nabla \varphi_k(x) \cdot \nabla \varphi_j(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla \varphi_k(x)\, \varphi_j(x)\, dx$$

for $k, j = 1, \ldots, M$. Invertibility of B_H follows from the discrete inf-sup condition (2.15). In fact, for any given $z \in X_v$ we can define the Galerkin projection $z_H = G_H z \in X_H$ as the unique solution of the variational formulation

$$b(G_H z, q_H) = b(z, q_H) \quad \text{for all } q_H \in Y_H,$$

and when combining the discrete inf-sup stability condition (2.15) and the boundedness result (2.7) of the bilinear form $b(\cdot,\cdot)$ we immediately conclude

$$\|G_H z\|_{X_v,H} \leq \sqrt{2}\,\|z\|_{X_v} \quad \text{for all } z \in X_v. \tag{2.18}$$

With (2.18) and using [16, 43] we then can formulate Cea's lemma,

$$\|u - u_H\|_{X_v,H} \leq \sqrt{2} \inf_{z_H \in X_H} \|u - z_H\|_{X_v}. \tag{2.19}$$

Lemma 2.7 *Let $u_H \in X_H$ be the unique solution of the Galerkin variational formulation (2.17), where we assume $u \in H^2(\Omega)$. When assuming (2.10) and $\operatorname{div} \boldsymbol{v}(x) \leq 0$ for almost all $x \in \Omega$ there holds the error estimate*

$$\|u - u_H\|_Y^2 \leq c \sum_{\ell=1}^{N} H_\ell^2 \left[\sup_{x\in\tau_\ell} \alpha(x) + \frac{c_v^2}{\underline{\alpha}}\right] |u|^2_{H^2(\tau_\ell)}. \tag{2.20}$$

Moreover, when $\boldsymbol{v} \in [L^\infty(\Omega)]^n$ and $\operatorname{div} \boldsymbol{v}(x) = 0$ for almost all $x \in \Omega$ is satisfied, the error estimate

$$\|u - u_H\|_Y^2 \leq c \sum_{\ell=1}^{N} \left[H_\ell^2 \sup_{x\in\tau_\ell} \alpha(x) + H_\ell^4 \sup_{x\in\tau_\ell} \frac{|\boldsymbol{v}(x)|^2}{\alpha(x)}\right] |u|^2_{H^2(\tau_\ell)} \tag{2.21}$$

follows.

Proof. From (2.19) we first have

$$\|u - u_H\|_Y^2 \;\leq\; \|u - u_H\|^2_{X_v,H} \leq \|u - I_H u\|^2_{X_v} \;=\; \|u - I_H u\|_Y^2 + \|\boldsymbol{v} \cdot \nabla(u - I_H u)\|_{Y^*},$$

where $I_H u$ is the piecewise linear nodal interpolation of the solution $u \in H^2(\Omega)$. Using standard local interpolation error estimates we obtain

$$\begin{aligned}
\|u - I_H u\|_Y^2 &= \int_\Omega \alpha(x)\,|\nabla[u(x) - I_H u(x)]|^2\,dx \\
&= \sum_{\ell=1}^{N} \int_{\tau_\ell} \alpha(x)\,|\nabla[u(x) - I_H u(x)]|^2\,dx \;\leq\; c \sum_{\ell=1}^{N} \sup_{x\in\tau_\ell} \alpha(x)\, H_\ell^2\, |u|^2_{H^2(\tau_\ell)}.
\end{aligned}$$

For the remaining part, and using the dual norm (2.5), we have to consider

$$\begin{aligned}
\|\boldsymbol{v} \cdot \nabla(u - I_H u)\|_{Y^*} &= \sup_{0\neq q\in Y} \frac{\langle \boldsymbol{v} \cdot \nabla(u - I_H u), q\rangle_\Omega}{\|q\|_Y} = \sup_{0\neq q\in Y} \frac{\langle \nabla(u - I_H u), q\,\boldsymbol{v}\rangle_\Omega}{\|q\|_Y} \\
&\leq \|\nabla(u - I_H u)\|_{L^2(\Omega)} \sup_{0\neq q\in Y} \frac{\|q\,\boldsymbol{v}\|_{L^2(\Omega)}}{\|q\|_Y} \\
&\leq \frac{1}{\sqrt{\underline{\alpha}}}\,\|\nabla(u - I_H u)\|_{L^2(\Omega)} \sup_{0\neq q\in Y} \frac{\|q\,\boldsymbol{v}\|_{L^2(\Omega)}}{\|\nabla q\|_{L^2(\Omega)}} \\
&= \frac{c_v}{\sqrt{\underline{\alpha}}}\,\|\nabla(u - I_H u)\|_{L^2(\Omega)}\,,
\end{aligned}$$

where we have used (2.10), i.e., we obtain

$$\|\boldsymbol{v}\cdot\nabla(u-I_Hu)\|_{Y^*}^2 \le c\,\frac{c_v^2}{\underline{\alpha}}\,\sum_{\ell=1}^N H_\ell^2\,|u|_{H^2(\tau_\ell)}^2\,,$$

and (2.20) follows.

In the case $\operatorname{div}\boldsymbol{v}(x)=0$ for almost all $x\in\Omega$ and $\boldsymbol{v}\in[L^\infty(\Omega)]^n$ we can use integration by parts to conclude

$$\begin{aligned}
|\langle\boldsymbol{v}\cdot\nabla(u-I_Hu),q\rangle_\Omega| &= \left|\int_\Omega \boldsymbol{v}(x)\cdot\nabla[u(x)-I_Hu(x)]\,q(x)\,dx\right| \\
&= \left|-\int_\Omega [u(x)-I_Hu(x)]\operatorname{div}[q(x)\,\boldsymbol{v}(x)]\,dx\right| \\
&= \left|-\int_\Omega [u(x)-I_Hu(x)]\,\boldsymbol{v}(x)\cdot\nabla q(x)\,dx\right| \\
&\le \sum_{\ell=1}^N \left|\int_{\tau_\ell} [u(x)-I_Hu(x)]\,\boldsymbol{v}(x)\cdot\nabla q(x)\,dx\right| \\
&\le \sum_{\ell=1}^N \|u-I_Hu\|_{L^2(\tau_\ell)}\left(\int_{\tau_\ell}[\boldsymbol{v}(x)\cdot\nabla q(x)]^2\,dx\right)^{1/2} \\
&\le c\sum_{\ell=1}^N H_\ell^2\,|u|_{H^2(\tau_\ell)}\sup_{x\in\tau_\ell}\frac{|\boldsymbol{v}(x)|}{\sqrt{\alpha(x)}}\left(\int_{\tau_\ell}\alpha(x)\,|\nabla q(x)|^2\,dx\right)^{1/2} \\
&\le c\left(\sum_{\ell=1}^N H_\ell^4\sup_{x\in\tau_\ell}\frac{|\boldsymbol{v}(x)|^2}{\alpha(x)}\,|u|_{H^2(\tau_\ell)}^2\right)^{1/2}\left(\sum_{\ell=1}^N\int_{\tau_\ell}\alpha(x)\,|\nabla q(x)|^2dx\right)^{1/2} \\
&= c\left(\sum_{\ell=1}^N H_\ell^4\sup_{x\in\tau_\ell}\frac{|\boldsymbol{v}(x)|^2}{\alpha(x)}\,|u|_{H^2(\tau_\ell)}^2\right)^{1/2}\|q\|_Y\,.
\end{aligned}$$

Hence we conclude

$$\|\boldsymbol{v}\cdot\nabla(u-I_Hu)\|_{Y^*}^2 \le c^2\sum_{\ell=1}^N H_\ell^4\sup_{x\in\tau_\ell}\frac{|\boldsymbol{v}(x)|^2}{\alpha(x)}\,|u|_{H^2(\tau_\ell)}^2,$$

and (2.21) follows. ∎

Example 2.2 *As a first example, we consider the convection-diffusion equation* (2.1) *in the particular case* $\alpha(x)=1$ *and* $\boldsymbol{v}(x)=v$ *for all* $x\in\Omega=(0,1)$, *i.e.,* $\underline{\alpha}=\overline{\alpha}=1$. *For a globally quasi-uniform mesh with* $H_\ell\sim H$ *for all* $\ell=1,\ldots,N$ *we can write the finite element error estimate* (2.21) *as*

$$\|\nabla(u-u_H)\|_{L^2(\Omega)} \le c\left(H^2+v\,H^4\right)^{1/2}|u|_{H^2(\Omega)}, \tag{2.22}$$

when assuming $u \in H^2(\Omega)$. *As in* [37, Example 1.2] *we consider the solution*

$$u(x) = x - \frac{\exp(x-1) - \exp(-1)}{1 - \exp(-1)} \quad \text{for } x \in (0,1) \tag{2.23}$$

and compute the right hand side accordingly. For the discretization we use piecewise linear basis functions which are defined with respect to some uniform decompostion of $\Omega = (0,1)$ *into* N *finite elements of mesh size* $H = 1/N$. *The numerical results are given in Table 1 which confirm linear convergence asymptotically, as expected. On the other hand, for* $v = 1000$, *we observe some initial higher order convergence as indicated in the error estimate* (2.22), *see also the discussion on related observations in* [35, 39].

	$v = 10$		$v = 100$		$v = 1000$	
N	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc
2	1.567e-01		5.283e-01		5.073e+00	
4	7.554e-02	1.053	9.943e-02	2.410	6.511e-01	2.962
8	3.759e-02	1.007	3.877e-02	1.359	8.967e-02	2.860
16	1.877e-02	1.002	1.890e-02	1.036	2.155e-02	2.057
32	9.385e-03	1.000	9.400e-03	1.008	9.563e-03	1.172
64	4.692e-03	1.000	4.694e-03	1.002	4.714e-03	1.021
128	2.346e-03	1.000	2.346e-03	1.000	2.349e-03	1.005
256	1.173e-03	1.000	1.173e-03	1.000	1.173e-03	1.001

Table 1: Convection-diffusion equation (2.1) in $\Omega = (0,1)$, $\alpha \equiv 1$, $\boldsymbol{v} = v \in \{10, 100, 1000\}$, error and estimated order of convergence (eoc) for the solution u as given in (2.23).

Example 2.3 *In this example, we consider the convection-diffusion equation* (2.1) *for* $\Omega = (0,1)$, $\alpha(x) = 1$ *for* $x \in (0,1)$, *and a velocity field which depends on the spatial variable, i.e,* $\boldsymbol{v}(x) = v(x)$. *As exact solution we consider* (2.23) *and compute the right hand side accordingly to the related chosen velocity fields. Firstly, we choose* $v(x) = 1/\sqrt{x}$ *which satisfies* (2.10) *and* $\operatorname{div} v(x) \le 0$ *for* $x \in (0,1)$. *Hence, for a globally quasi-uniform mesh the error estimate* (2.20) *gives*

$$\|\nabla(u - u_H)\|_{L^2(\Omega)} \le cH\,|u|_{H^2(\Omega)}$$

when assuming $u \in H^2(\Omega)$. *Secondly, we choose* $v(x) = 1/x$ *with* $\operatorname{div} v(x) \le 0$ *for* $x \in (0,1)$ *but we can not use* (2.13). *Although we could prove surjectivity in a different way, we can not apply the error estimates as given in Lemma 2.7. The numerical results for both choices of the velocity field can be seen in Tab. 2. In both cases we observe a linear convergence rate which confirms the a priori error estimate* (2.20).

	$v(x) = \frac{1}{\sqrt{x}}$		$v(x) = \frac{1}{x}$	
N	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc
2	1.485e-01	0.000	1.494e-01	0.000
4	7.487e-02	0.988	7.503e-02	0.994
8	3.751e-02	0.997	3.754e-02	0.999
16	1.876e-02	0.999	1.877e-02	1.000
32	9.383e-03	1.000	9.384e-03	1.000
64	4.692e-03	1.000	4.692e-03	1.000
128	2.346e-03	1.000	2.346e-03	1.000
256	1.173e-03	1.000	1.173e-03	1.000

Table 2: Convection-diffusion equation (2.1) in $\Omega = (0,1)$, $\alpha \equiv 1$, $\boldsymbol{v}(x) = v(x) \in \left\{\frac{1}{\sqrt{x}}, \frac{1}{x}\right\}$, error and estimated order of convergence (eoc) for the solution u as given in (2.23).

Example 2.4 *We consider $\alpha(x) = \varepsilon \ll 1$, i.e., a singularly perturbed problem with dominating convection. Using again a globally quasi-uniform mesh, the error estimate* (2.21) *now gives*

$$\varepsilon \, \|\nabla(u-u_H)\|^2_{L^2(\Omega)} \leq c \left[H^2 \, \varepsilon + \|\boldsymbol{v}\|^2_{[L^\infty(\Omega)]^n} \, H^4 \, \varepsilon^{-1} \right] |u|^2_{H^2(\Omega)},$$

when assuming $u \in H^2(\Omega)$, i.e.,

$$\|\nabla(u-u_H)\|_{L^2(\Omega)} \leq c \left(H^2 + H^4 \, \varepsilon^{-2} \right)^{1/2} |u|_{H^2(\Omega)} \, .$$

As in [37, Example 1.2] *we consider $\Omega = (0,1)$ and $\boldsymbol{v}(x) = v = 1$ with the exact solution to be*

$$u(x) = x - \frac{\exp\left(-\frac{1-x}{\varepsilon}\right) - \exp\left(-\frac{1}{\varepsilon}\right)}{1 - \exp\left(-\frac{1}{\varepsilon}\right)}, \quad x \in (0,1), \tag{2.24}$$

which has a boundary layer at $x = 1$. The convergence behaviour of the numerical solutions u_H for $\varepsilon \in \{10^{-2}, 10^{-4}, 10^{-5}\}$ are depicted in Tab. 3 and Fig. 1. We see that we obtain linear convergence if the mesh size H is of the order of the singular perturbation parameter ε, i.e., $H \sim \varepsilon$. Before, the numerical solutions u_H obtain oscillations due to an insufficient resolution of the boundary layer, see also Fig. 2. However, when using a globally quasi-uniform mesh the choice $H \sim \varepsilon$ is impractical in applications as this leads to an unacceptably large number of grid points. This motivates to consider an adaptive approach, which works for all values of the singular perturbation parameter ε.

$\varepsilon = 10^{-2}$			$\varepsilon = 10^{-4}$			$\varepsilon = 10^{-5}$		
N	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc	N	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc	N	$\|\nabla(u-u_H)\|_{L^2(0,1)}$	eoc
128	1.567e+00		4096	4.280e+01		65536	9.237e+01	
256	7.938e-01	0.981	8192	2.389e+01	0.841	131072	4.843e+01	0.931
512	3.982e-01	0.995	16384	1.233e+01	0.955	262144	2.452e+01	0.982
1024	1.993e-01	0.999	32768	6.213e+00	0.988	524288	1.230e+01	0.995
2048	9.966e-02	1.000	65536	3.113e+00	0.997	1048576	6.154e+00	0.999

Table 3: Convection-diffusion equation (2.1) in $\Omega = (0,1)$, $\alpha(x) = \varepsilon \in \{10^{-2}, 10^{-4}, 10^{-5}\}$, $\boldsymbol{v} = 1$, error and estimated order of convergence (eoc) for the solution u as given in (2.24).

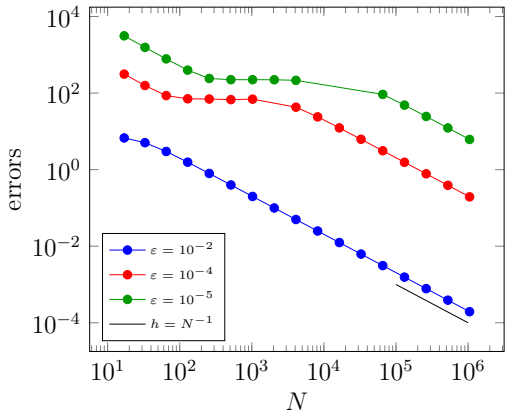
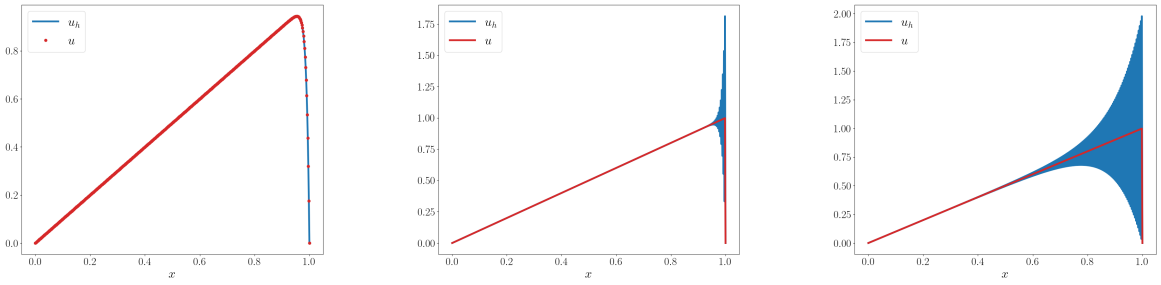


Figure 1: Convergence behaviour of the error $\|\nabla(u-u_H)\|_{L^2(0,1)}$ for a uniform refinement strategy in case of the function u as given in (2.24).



(a) $\varepsilon = 10^{-2}$ (b) $\varepsilon = 10^{-4}$ (c) $\varepsilon = 10^{-5}$

Figure 2: Numerical solutions u_H to Example 2.4 in case of $N = 512$ elements.

2.3 Adaptive least squares finite element method

The bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ as defined in (2.6) and (2.3) imply bijective operators $A : Y \to Y^*$ and $B : X_v \to Y^*$ satisfying

$$\langle Ap, q\rangle_\Omega = \int_\Omega \alpha(x)\, \nabla p(x) \cdot \nabla q(x)\, dx \quad \text{for all } p, q \in Y,$$

and

$$\langle Bu, q\rangle_\Omega = b(u, q) \quad \text{for all } (u, q) \in X_v \times Y.$$

Hence, instead of the operator equation $Bu = f$ in Y^*, i.e., of the variational formulation (2.3), we now consider the equivalent problem to minimize the quadratic functional

$$\mathcal{J}(z) = \frac{1}{2}\, \|Bz - f\|^2_{A^{-1}} \quad \text{for } z \in X_v. \tag{2.25}$$

The minimizer $u \in X_v$ of (2.25) is given as the unique solution of the gradient equation

$$B^* A^{-1}(Bu - f) \;=\; 0, \tag{2.26}$$

and using the adjoint $p = A^{-1}(f - Bu) \in Y$ we have to solve a coupled system to find $(u, p) \in X_v \times Y$ such that

$$Ap + Bu = f, \quad B^* p = 0\,,$$

i.e.,

$$\int_\Omega \alpha(x)\nabla p(x) \cdot \nabla q(x)dx + \int_\Omega \alpha(x)\nabla u(x) \cdot \nabla q(x)dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u(x)\, q(x)dx = \int_\Omega f(x)q(x)dx \tag{2.27}$$

for all $q \in Y$, and

$$\int_\Omega \alpha(x)\nabla z(x) \cdot \nabla p(x)dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla z(x)\, p(x)dx = 0 \quad \text{for all } z \in X_v. \tag{2.28}$$

Unique solvability of the gradient equation (2.26) and therefore of the equivalent mixed variational formulation (2.27) and (2.28) follows as in [29], since $S := B^* A^{-1} B : X_v \to X_v^*$ is bounded and elliptic. Note that in the continuous case we have $p \equiv 0$, since B and therefore B^* are bijective.

For the Galerkin finite element discretization of the mixed variational system (2.27) and (2.28) we will use again the conforming finite element space $X_H = \text{span}\{\varphi_k\}_{k=1}^M \subset X_v$, but we introduce a finite element space $Y_h = \text{span}\{\phi_i\}_{i=1}^{\overline{M}} \subset Y$ of piecewise linear continuous basis functions ϕ_i which are defined with respect to a possibly refined decomposition of Ω into $\overline{N}$ finite elements $\overline{\tau}_\ell$ of local mesh size h_ℓ, e.g., starting from X_H we use at least one additional refinement to construct Y_h, i.e., $h_\ell = H_\ell/2$ when $\overline{\tau}_\ell \subset \tau_\ell$. This definition ensures

$X_H \subset Y_h$, but even the choice $Y_h = X_H$ would be allowed for a stable discretization. In both cases, (2.15) implies the discrete inf-sup condition

$$\|u_H\|_{X_v,H} \leq \sup_{0 \neq q_h \in Y_h} \frac{b(u_H, q_h)}{\|q_h\|_Y} \quad \text{for all } u_H \in X_H. \tag{2.29}$$

The mixed finite element discretization of (2.27) and (2.28) is to find $(p_h, u_H) \in Y_h \times X_H$ such that

$$\int_\Omega \alpha(x) \nabla p_h(x) \cdot \nabla q_h(x) dx + \int_\Omega \alpha(x) \nabla u_H(x) \cdot \nabla q_h(x) dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla u_H(x)\, q_h(x) dx = \int_\Omega f(x) q_h(x) dx \tag{2.30}$$

for all $q_h \in Y_h$, and

$$\int_\Omega \alpha(x) \nabla z_H(x) \cdot \nabla p_h(x) dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla z_H(x)\, p_h(x) dx = 0 \quad \text{for all } z_H \in X_H. \tag{2.31}$$

This is equivalent to a coupled system of linear algebraic equations

$$\begin{pmatrix} A_h & B_h \\ B_h^\top & \end{pmatrix} \begin{pmatrix} \underline{p} \\ \underline{u} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix} \tag{2.32}$$

where the entries of the stiffness matrix are now given by

$$\begin{aligned} A_h[j,i] &= \int_\Omega \alpha(x)\, \nabla \phi_i(x) \cdot \nabla \phi_j(x)\, dx, \\ B_h[j,k] &= \int_\Omega \alpha(x)\, \nabla \varphi_k(x) \cdot \nabla \phi_j(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla \varphi_k(x)\, \phi_j(x)\, dx \end{aligned}$$

for $i, j = 1, \ldots, \overline{M}$, $k = 1, \ldots, M$. Unique solvability of (2.32) follows from (2.29). Since A_h is invertibe, we can also consider the Schur complement system

$$B_h^\top A_h^{-1} B_h \underline{u} = B_h^\top A_h^{-1} \underline{f}$$

which is the discrete counter part of the gradient equation (2.26). In the particular case $Y_h = X_H$ we have $B_h = B_H$ to be invertible, and $\underline{p} = \underline{0}$ follows in this case. Hence, to have a built in error estimation we now consider the case $X_H \subset Y_h$, $X_H \neq Y_h$ only. As in (2.14) we now define $w_{u,h} \in Y_h$, and introduce the discrete norm $\|u\|_{X_v,h} = \sqrt{\|u\|_Y^2 + \|w_{u,h}\|_Y^2}$. As in [29, Lemma 2.4, equations (2.16), (2.31)] we then conclude the a priori error estimate

$$\|u - u_H\|_{X_v,h} \leq c \inf_{z_H \in X_H} \|u - z_H\|_{X_v}, \tag{2.33}$$

and we finally obtain the error estimates as already given in Lemma 2.7. In order to derive an a posteriori error indicator we first have the following result:

Lemma 2.8 *Let $(p_h, u_H) \in Y_h \times X_H$ be the unique solution of the variational formulations* (2.30) *and* (2.31)*. Then there holds*

$$\frac{1}{\sqrt{2}} \|p_h\|_Y \leq \|u - u_H\|_{X_v,h} \,. \tag{2.34}$$

Proof. When subtracting (2.30) from (2.27) for the test function $q_h = p_h \in Y_h$ this gives the Galerkin orthogonality, recall $p \equiv 0$,

$$\begin{aligned}
\|p_h\|_Y^2 &= \int_\Omega \alpha(x)\, \nabla p_h(x) \cdot \nabla p_h(x)\, dx \\
&= \int_\Omega \alpha(x)\, \nabla [u(x) - u_H(x)] \cdot \nabla p_h(x)\, dx + \int_\Omega \boldsymbol{v}(x) \cdot \nabla [u(x) - u_H(x)]\, p_h(x)\, dx \\
&= \int_\Omega \alpha(x)\, \nabla [u(x) - u_H(x)] \cdot \nabla p_h(x)\, dx + \int_\Omega \nabla w_{u-u_H,h}(x) \cdot \nabla p_h(x)\, dx \\
&\leq \Big[\|u - u_H\|_Y + \|w_{u-u_H,h}\|_Y \Big] \|p_h\|_Y \leq \sqrt{2} \Big(\|u - u_H\|_Y^2 + \|w_{u-u_H,h}\|_Y^2 \Big)^{1/2} \|p_h\|_Y \\
&= \sqrt{2}\, \|u - u_H\|_{X_v,h} \|p_h\|_Y,
\end{aligned}$$

i.e., the assertion follows. ∎

It remains to prove that the error estimator $\|p_h\|_Y$ is also reliable. For this we consider the variational formulations (2.30) and (2.31), using $X_h = Y_h$, to find $(\overline{p}_h, u_h) \in Y_h \times X_h$ such that

$$\langle A\overline{p}_h, q_h \rangle_\Omega + \langle B u_h, q_h \rangle_\Omega = \langle f, q_h \rangle_\Omega, \quad \langle B z_h, \overline{p}_h \rangle_\Omega = 0 \tag{2.35}$$

is satisfied for all $(z_h, q_h) \in Y_h \times X_h$. As in (2.15) we have the discrete inf-sup condition

$$\|u_h\|_{X_v,h} \leq \sup_{0 \neq q_h \in Y_h} \frac{b(u_h, q_h)}{\|q_h\|_Y} \quad \text{for all } u_h \in X_h, \tag{2.36}$$

from which unique solvability of (2.35) follows.

Lemma 2.9 *Let $(p_h, u_H) \in Y_h \times X_H$ and $(\overline{p}_h, u_h) \in X_h \times Y_h$ be the unique solutions of the variational formulations* (2.30)–(2.31) *and* (2.35)*, respectively. Assume the saturation condition*

$$\|u - u_h\|_{X_v,h} \leq \eta \, \|u - u_H\|_{X_v,h} \quad \textit{for some } \eta \in (0,1). \tag{2.37}$$

Then there holds the reliability estimate

$$\|u - u_H\|_{X_v,h} \leq \frac{1}{1-\eta} \|p_h\|_Y \,. \tag{2.38}$$

Proof. Since the variational formulation (2.35) is considered for $X_h = Y_h$, $\overline{p}_h \equiv 0$ follows. Hence we obtain, using (2.30),

$$\langle B u_h, q_h \rangle_\Omega = \langle f, q_h \rangle_\Omega = \langle A p_h, q_h \rangle_\Omega + \langle B u_H, q_h \rangle_\Omega \quad \text{for all } q_h \in Y_h,$$

i.e.,

$$\langle B(u_h - u_H), q_h\rangle_\Omega = \langle Ap_h, q_h\rangle_\Omega \quad \text{for all } q_h \in Y_h.$$

Now, using the discrete inf-sup condition (2.36) for $u_h - u_H \in X_h$ this gives

$$\|u_h - u_H\|_{X_v,h} \leq \sup_{0\neq q_h\in Y_h} \frac{\langle B(u_h - u_H), q_h\rangle_\Omega}{\|q_h\|_Y} = \sup_{0\neq q_h\in Y_h} \frac{\langle Ap_h, q_h\rangle_\Omega}{\|q_h\|_Y} \leq \|p_h\|_Y \,.$$

Then, using the triangle inequality and the saturation assumption (2.37), we obtain

$$\|u - u_H\|_{X_v,h} \leq \|u - u_h\|_{X_v,h} + \|u_h - u_H\|_{X_v,h} \leq \eta \,\|u - u_H\|_{X_v,h} + \|p_h\|_Y \,,$$

i.e., the assertion follows. ∎

Remark 2.2 *The numerical results indicate that the saturation condition* (2.37) *is satisfied for the choice $h = H/2$, i.e., one additional refinement to define Y_h when starting from X_H. Otherwise, one may use some more additional refinement steps when required.*

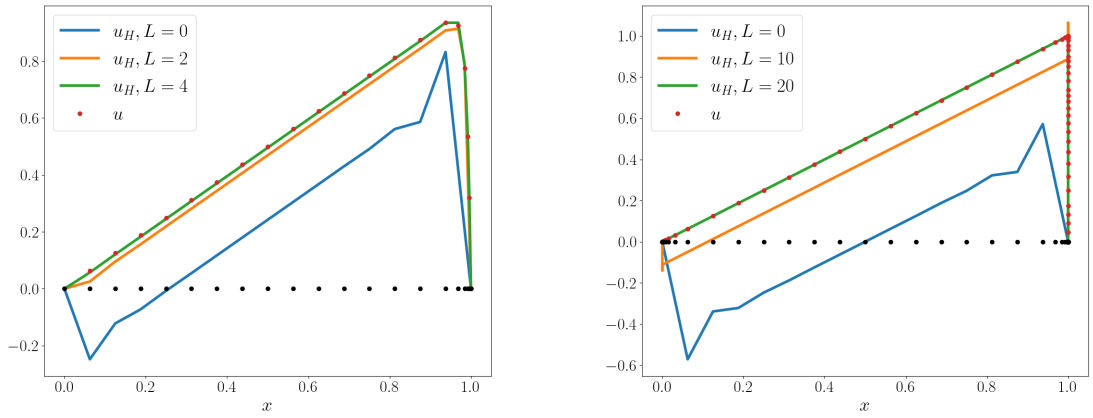
Example 2.5 *In this example we apply the least-squares approach* (2.25) *to the singularly perturbed problem already considered in Example 2.4. We use the finite element spaces $X_H = S^1_H(\mathcal{T}_H) \cap X_v$ and $Y_h = S^2_H(\mathcal{T}_H) \cap Y$, which are defined with respect to an admissible and locally quasi-uniform decomposition $\mathcal{T}_H = \{\tau_\ell\}_{\ell=1}^N$ of the interval $\Omega = (0, 1)$ into finite elements τ_ℓ with $\widetilde{M} = N + 1$ nodes. We use the global error estimator*

$$\eta_H^2 = \|p_h\|_Y^2 = \varepsilon \int_0^1 \nabla p_h(x) \cdot \nabla p_h(x)\, dx = \sum_{\ell=1}^N \eta_\ell^2 \tag{2.39}$$

with the local error indicators

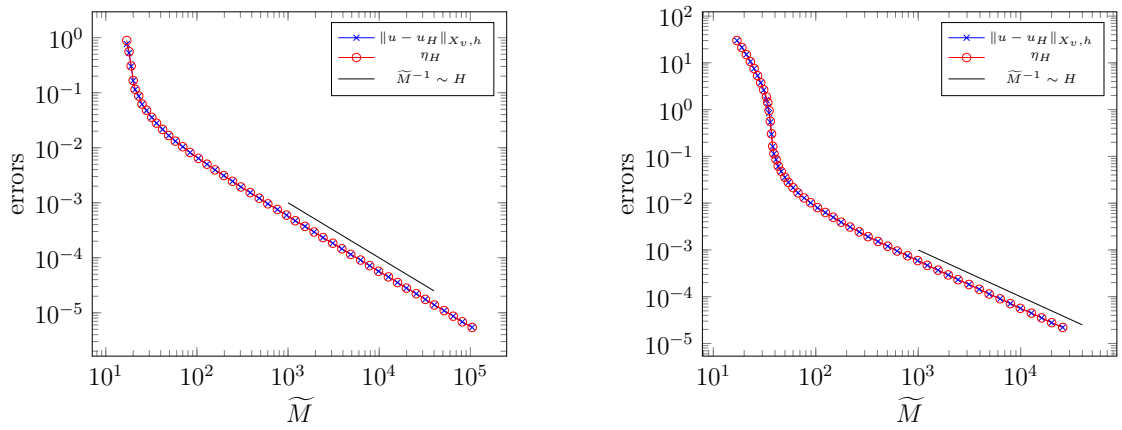
$$\eta_\ell^2 = \varepsilon \int_{\tau_\ell} \nabla p_h(x) \cdot \nabla p_h(x)\, dx \tag{2.40}$$

to drive an adaptive refinement scheme with a Dörfler marking strategy [20]. The sparse direct solver `Pardiso` *is used to solve the resulting linear systems. The convergence behaviour of the error and the estimator are given in Fig. 4. As expected, we observe a linear rate. The numerical solutions obtained on different refinement levels are provided in Fig. 3. We see that the sequence of iterates from the adaptive refinement process converges to the physical true solution. However, the first few iterates obtain a constant shift from the true solution and some minor oscillations at $x = 0$ and at $x = 1$, respectively. Note that this behaviour was also observed in* [5, 17]*. Further, we see that the numerical solutions on the first few refinement levels have negative values even though the true solution is not negative. Therefore, the discrete maximum principle is not satisfied for all the solutions from the adaptive refinement process. However, the inbuilt error estimator detects how many degrees of freedom need to be added in order to obtain a numerical solution u_H, which is in good accordance with the physical correct reference solution u. Finally, we want to mention*



(a) $\varepsilon = 10^{-2}$ (b) $\varepsilon = 10^{-5}$

Figure 3: Numerical solutions u_H to Example 2.5 in the case of the adaptive refinement scheme.



(a) $\varepsilon = 10^{-2}$ (b) $\varepsilon = 10^{-5}$

Figure 4: Convergence behaviour of the errors and the estimators.

that in the case of $\varepsilon = 10^{-5}$ *the numerical solution* u_H *on* $L = 20$ *was computed on a mesh with* 60 *vertices (indicated in black) and* 179 *degrees of freedom (dof) for the corresponding saddle point system. This already lead to a satisfactory result. In comparison, the direct approach needed about* 100000 *dofs to give a satisfactory approximation to the solution, see Fig. 1 and Tab. 3.*

Example 2.6 ([28, Example 3.3]) *This problem deals with a non-constant convection field* $\boldsymbol{v}$*. In particular, we consider* $\Omega = (0,1)^2$ *with* $\boldsymbol{v} = (-y, x)^\top$, $\alpha(x) = \varepsilon = 10^{-5}$, *and* $f = 0$*. We prescribe homogeneous Dirichlet boundary conditions on the boundaries* $\{1\} \times [0,1]$ *and* $[0,1] \times \{1\}$*, i.e., on the right and top boundary. At the inlet boundary*

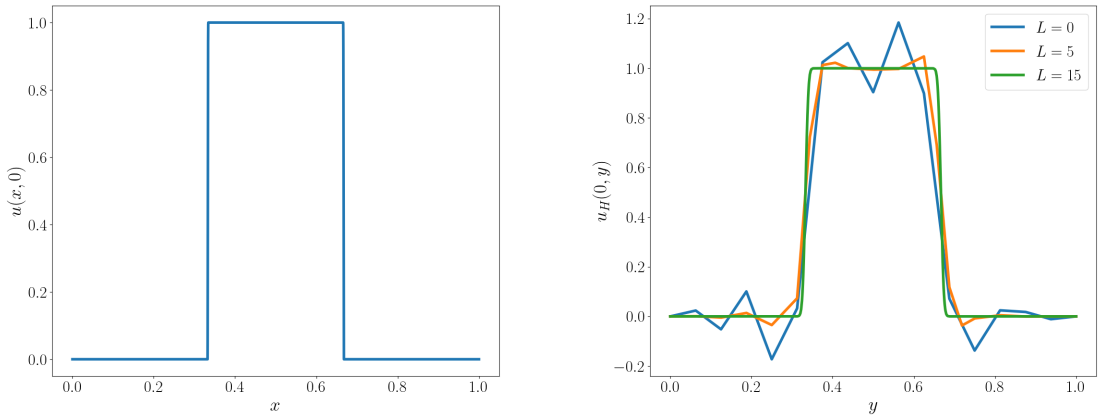


Figure 5: Left: Inhomogeneous boundary condition (2.41) prescribed on inlet boundary. Right: Numerical solution u_H at the outlet boundary for different refinement levels.

$[0,1]\times\{0\}$ *we consider the inhomogeneous boundary condition given by, with* $\xi = 10^{-3}$,

$$u(x,0) = \begin{cases} 1 - \frac{1}{4}\left(1 - \cos\left(\frac{1/3+\xi-x}{2\xi}\pi\right)\right)^2 & \text{for } x \in \left[\frac{1}{3}-\xi, \frac{1}{3}+\xi\right], \\ 1 & \text{for } x \in \left(\frac{1}{3}+\xi, \frac{2}{3}-\xi\right), \\ 1 - \frac{1}{4}\left(1 - \cos\left(\frac{x-2/3+\xi}{2\xi}\pi\right)\right)^2 & \text{for } x \in \left[\frac{2}{3}-\xi, \frac{2}{3}+\xi\right], \\ 0 & \text{else.} \end{cases} \tag{2.41}$$

On the remaining outlet boundary $\{0\}\times(0,1)$ *we prescribe homogeneous Neumann boundary conditions. In our implementation we use piecewise linear and piecewise quadratic basis functions to define the finite element spaces* X_H *and* Y_h*, respectively. Furthermore, we use the inbuilt error estimator to drive an adaptive refinement scheme with a Dörfler marking strategy* [20]*. All resulting linear systems are solved with the sparse direct solver* **`Pardiso`**. *In order to study the satisfaction of the global discrete maximum principle (DMP) we evaluate as in* [27] *the quantity*

$$\mathrm{osc}_{max}(u_H) = \max_{(x,y)\in\overline{\Omega}} u_H(x,y) - 1 - \min_{(x,y)\in\overline{\Omega}} u_H(x,y). \tag{2.42}$$

In order to assess the accuracy of the numerical solution three characteristic values of the solution at the outflow boundary are provided in [28]*. The reference values read:*

- *width of the lower layer:* 0.01439869,
- *width of the upper layer:* 0.01439637,
- *outflow profile width:* 0.3482541.

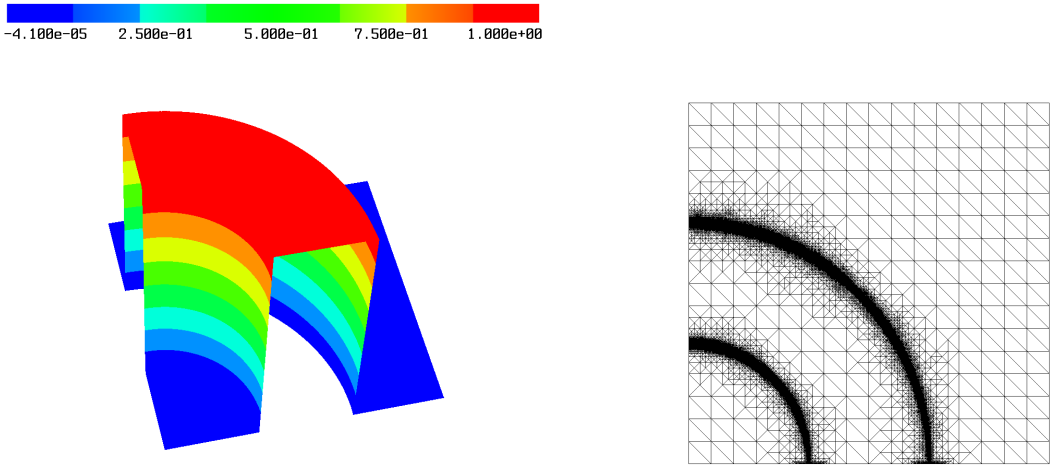


Figure 6: Left: Numerical solution u_H on $L = 15$. Right: Adaptive mesh on $L = 15$ with 148968 dofs.

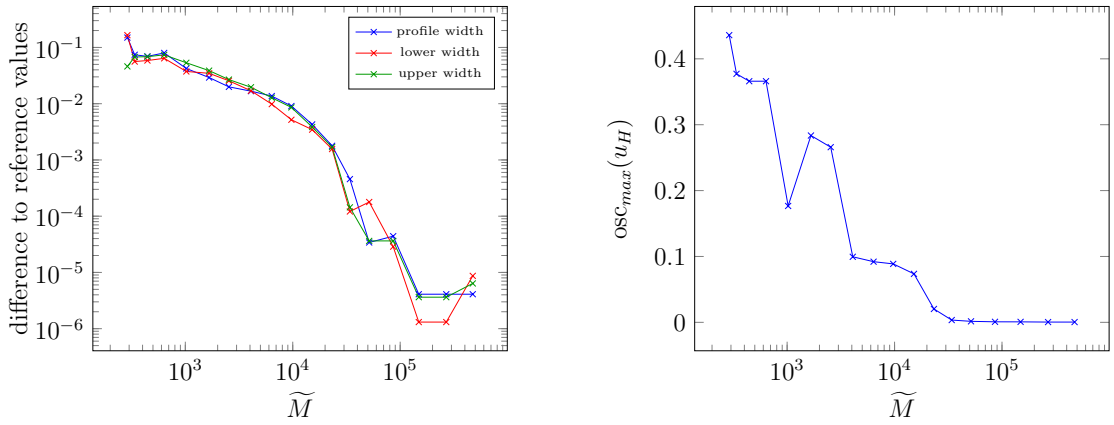


Figure 7: Left: Comparison of the computed characteristic values to the reference values. Right: Satisfaction of the global DMP on different refinement levels.

L	$\widetilde{M}$	lower layer	upper layer	profile	$\mathrm{osc}_{max}(u_H)$
0	289	0.17989000	0.06040999	0.49811000	0.43594821
5	1666	0.04909000	0.05312000	0.37743000	0.28357896
10	15160	0.01785999	0.01824999	0.35253999	0.07359097
13	51480	0.01422000	0.01436000	0.34822000	1.523e-03
15	148968	0.01440000	0.01440000	0.34825000	7.373e-04
17	473779	0.01439000	0.01439000	0.34825000	4.033e-04

Table 4: Computed characteristic values and evaluation of (2.42).

In Fig. 5 we provide a plot of the inhomogeneous boundary condition (2.41) *on the inlet boundary as well as the numerical solution* u_H *at the outlet boundary for different refinement levels. In Fig. 6 the numerical solution* u_H *as well as the adaptive mesh generated on refinement level* $L = 15$ *are depicted. The adaptive mesh was generated from a structured initial mesh with* 16×16 *elements. Further, in Tab. 4 and Fig. 7 we provide a comparison of the computed characteristic values to the reference values and the satisfaction of the global DMP. We see that for earlier refinement levels over and under shoots are visible, but they are almost vanishing on higher refinement levels. This can be also seen from the quantity* (2.42) *in Table 4 and Fig. 7, which show a good satisfaction of the global DMP for higher refinements. In addition to that Fig. 7 shows that the difference of the computed characteristic quantities to the reference values is in the order of floating point precision for higher refinement levels, i.e., we have a good agreement of the computed reference values with the characteristic values.*

Example 2.7 (Hemker problem) *In this example we consider the Hemker problem* [26]*, which is a standard benchmark problem for steady state convection-diffusion problems. The domain is given by* $\Omega = (-3, 9) \times (-3, 3) \setminus \{(x, y) : x^2 + y^2 \leq 1\}$*. The velocity is given by* $\boldsymbol{v} = (1, 0)^\top$*, and the right-hand side* f *is set equal to zero. Further we have the boundary conditions*

$$u(x,y) = \begin{cases} 0 & x = -3, \\ 1 & x^2 + y^2 = 1, \\ \alpha(x)\nabla u \cdot \mathbf{n} = 0 & x = 9 \ \vee \ y = -3 \ \vee \ y = 3 \end{cases} .$$

As in [2, 27, 28] *we consider the diffusion coefficient to be* $\alpha(x) = \varepsilon = 10^{-4}$*. In order to assess the accuracy of the numerical solution a value for the width of the interior layer at* $x = 4$ *was provided in* [2]*. This width is defined to be the length of the interval, where* $u(4, y) \in [0.1, 0.9]$*. In* [2] *the reference value* 0.0723 *is provided for the upper layer, i.e., where* $y \geq 0$*. Furthermore, we evaluate the quantity* (2.42) *to measure the satisfaction of the global DMP. In our numerical experiments we use piecewise linear finite elements for the trial space* X_H *and piecewise quadratic finite elements for the test space* Y_h*. Further we use the local error indicators* (2.40) *to drive an adaptive refinement scheme. As a marking strategy we use the Dörfler criterion with parameter* $\theta = 0.5$*. The marked elements are then refined using newest vertex bisection. The resulting linear systems are solved with the sparse direct solver* `Pardiso`*.*

In Fig. 8 the initial mesh and the adaptive mesh obtained on level $L = 17$ *are depicted. We see stronger refinements at the boundary layer around the circle, which at the top and bottom of the circle passes into an interior layer that spread into the direction of the convection. The refinements at the left boundray may be explained in terms of the over- and undershoots that occur for lower refinements. In Fig. 9 we depict the numerical solution* u_H *obtained on refinement level* $L = 17$*. In Fig. 10 we provide plots of the cut lines* $u_H(x, 1)$ *for* $-3 \leq x \leq 9$ *and* $u_H(4, y)$ *for* $-3 \leq y \leq 3$*. The cut lines possess some oscillations for lower refinement levels which get significantly reduced for higher refinement levels. In Fig. 11 and Tab. 5 we provide a comparison of the computed characteristic value*

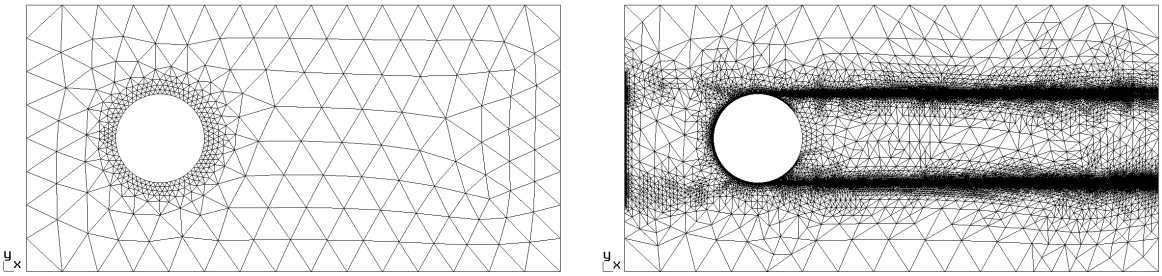


Figure 8: Left: Initial mesh with 438 dofs. Right: Adaptive mesh on $L = 17$ with 676072 dofs.

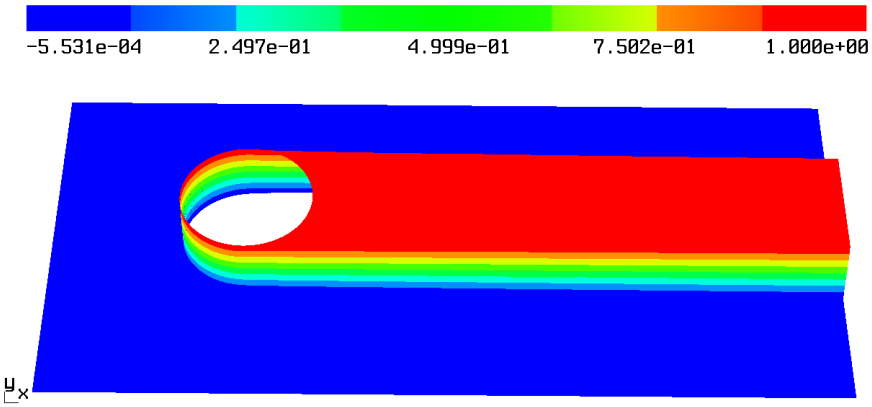


Figure 9: Numerical solution u_H on $L = 17$.

L	$\widetilde{M}$	upper layer	$\mathrm{osc}_{max}(u_H)$
0	438	0.8087	0.9742
8	12048	0.4873	0.6315
12	82453	0.1603	0.1909
15	227728	0.0801	0.0165
17	676072	0.0728	1.187e-03
18	1181853	0.0727	6.823e-04

Table 5: Computed characteristic value and evaluation of (2.42) for the Hemker problem.

with the reference value as well as a plot of the satisfaction of the global DMP. We see that the computed value converges to the reference value given for the upper layer width. Furthermore, we see a reduction of the quantity (2.42) *for higher refinements.*

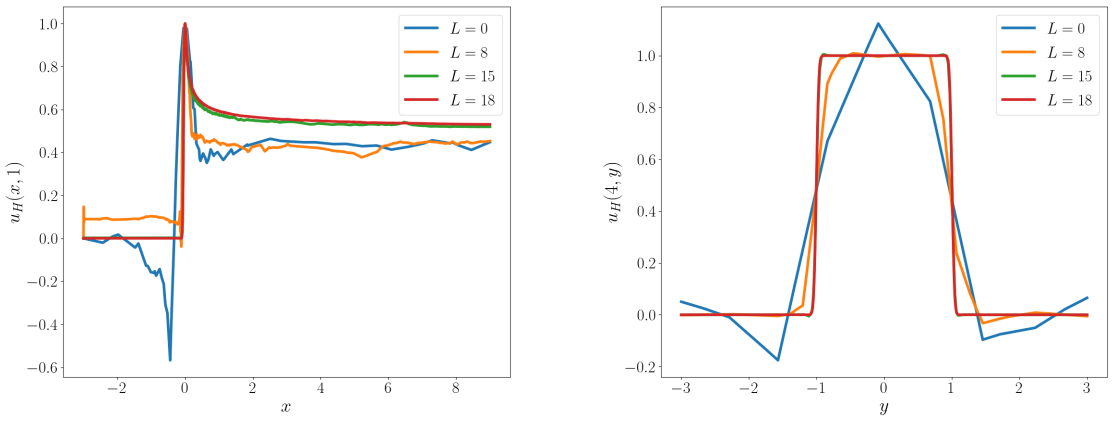


Figure 10: Cut lines of the numerical solution on different refinement levels. Left: $u_H(x, 1)$. Right: $u_H(4, y)$.

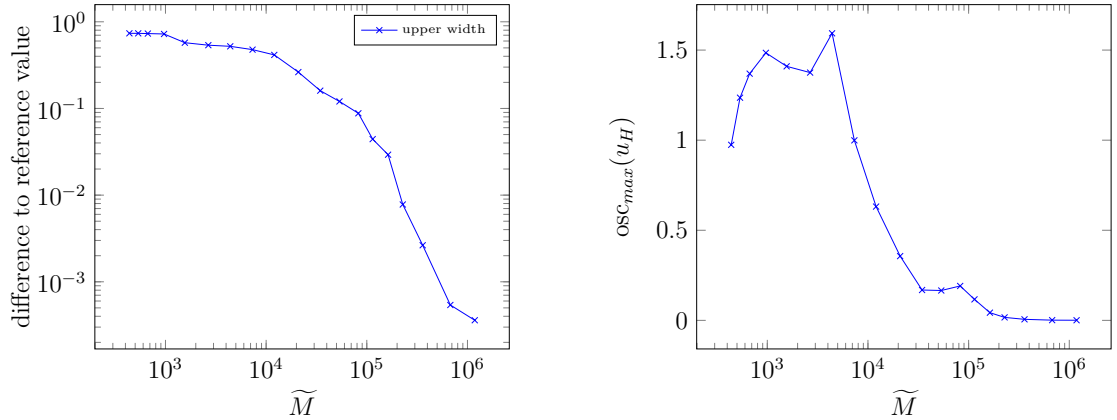


Figure 11: Left: Comparison of the computed characteristic values to the reference values. Right: Satisfaction of the global DMP on different refinement levels.

3 Instationary convection-diffusion problems

Instead of (2.1) we now consider the time-dependent Dirichlet boundary value problem for the convection-diffusion equation,

$$\begin{aligned} \partial_t u(y,t) + \boldsymbol{v}(y,t)\cdot\nabla_y u(y,t) - \mathrm{div}_y[\alpha(y)\,\nabla_y u(y,t)] &= f(y,t) && \text{for } (y,t)\in Q,\\ u(y,t) &= 0 && \text{for } (y,t)\in\Sigma,\\ u(x,0) &= 0 && \text{for } x\in\Omega, \end{aligned} \tag{3.1}$$

where the space-time domain Q is given by

$$Q := \Big\{(y,t)\in\mathbb{R}^{n+1} : y=\varphi(t,x), x\in\Omega\subset\mathbb{R}^n, t\in(0,T)\Big\},$$

with the lateral boundary

$$\Sigma := \Big\{(y,t)\in\mathbb{R}^{n+1} : y=\varphi(t,x), x\in\partial\Omega, t\in(0,T)\Big\}.$$

Here, $T>0$ is a given time horizon, and $\boldsymbol{v}(y,t)=\frac{d}{dt}y(t)$ is the velocity along the trajectory $y(t)=\varphi(t,x)\in\mathbb{R}^n$ for a reference point $x\in\Omega$. We assume that the deformation φ is bijective and sufficient regular for all $t\in(0,T)$, satisfying $\varphi(0,x)=x$ for all $x\in\Omega$, and $\mathrm{div}_y\boldsymbol{v}(y,t)=0$. Then we can write Reynold's transport theorem as

$$\frac{d}{dt}\int_{\Omega(t)} u(y,t)\,dy \;=\; \int_{\Omega(t)}\Big[\partial_t u(y,t)+\boldsymbol{v}(y,t)\cdot\nabla_y u(y,t)\Big]\,dy \;=\; \int_{\Omega(t)}\frac{d}{dt}u(y,t)\,dy. \tag{3.2}$$

As in the stationary case we assume that the diffusion coefficient $\alpha(y)$ is bounded and strictly positive, see (2.2).

3.1 Variational formulation

Similar as in the stationary case we now define the Bochner spaces $Y := L^2(0,T;H_0^1(\Omega(t)))$ with the norm

$$\|q\|_Y^2 := \int_0^T\int_{\Omega(t)}\alpha(y)\,|\nabla_y q(y,t)|^2\,dy\,dt,$$

and

$$X_v := \Big\{u\in Y : \frac{d}{dt}u = \partial_t u+\boldsymbol{v}\cdot\nabla_y u\in Y^*, u(x,0)=0 \text{ for } x\in\Omega\Big\}.$$

A norm in X_v is given by the graph norm

$$\|u\|_{X_v} := \sqrt{\|u\|_Y^2+\|\partial_t u+\boldsymbol{v}\cdot\nabla_y u\|_{Y^*}^2} = \sqrt{\|u\|_Y^2+\|w_u\|_Y^2},$$

where $w_u\in Y$ is the unique solution of the variational problem such that

$$\begin{aligned} a(w_u,q) := \int_0^T\int_{\Omega(t)}\alpha(y)\,\nabla_y w_u(y,t)\cdot\nabla_y q(y,t)\,dy\,dt \\ = \int_0^T\int_{\Omega(t)}\Big[\partial_t u(y,t)+\boldsymbol{v}(y,t)\cdot\nabla_y u(y,t)\Big]\,q(y,t)\,dy\,dt \end{aligned} \tag{3.3}$$

is satisfied for all $q \in Y$. The space-time variational formulation of (3.1) then reads to find $u \in X_v$ such that

$$\begin{aligned} b(u,q) \;:=\; & \int_0^T \int_{\Omega(t)} \Big[\partial_t u(y,t) + \boldsymbol{v}(y,t)\cdot\nabla_y u(y,t)\Big] q(y,t)\,dy\,dt \\ & + \int_0^T \int_{\Omega(t)} \alpha(y)\,\nabla_y u(y,t)\cdot\nabla_y q(y,t)\,dy\,dt \;=\; \int_0^T \int_{\Omega(t)} f(y,t) q(y,t)\,dy\,dt \end{aligned} \tag{3.4}$$

is satisfied for all $q \in Y$. The boundedness of the bilinear form $b(\cdot,\cdot) : X \times Y \to \mathbb{R}$ follows as in the proof of Lemma 2.5, i.e.,

$$|b(u,q)| \le \sqrt{2}\,\|u\|_{X_v}\|q\|_Y \quad \text{for all } (u,q) \in X_v \times Y.$$

Moreover, choosing $q_u := u + w_u \in Y$, we conclude

$$\begin{aligned} b(u,q_u) &= \langle \partial_t u + \boldsymbol{v}\cdot\nabla_y u, u + w_u\rangle_Q \;+\; \langle \nabla_y u, \nabla_y(u+w_u)\rangle_{L^2(Q)} \\ &= 2\,\langle \nabla_y w_u, \nabla_y u\rangle_Q \;+\; \langle \nabla_y w_u, \nabla_y w_u\rangle_{L^2(Q)} \;+\; \langle \nabla_y u, \nabla_y u\rangle_{L^2(\Omega)} \\ &= \langle \nabla_y(u+w_u), \nabla_y(u+w_u)\rangle_{L^2(Q)} \;=\; \|q_u\|_Y^2 \,. \end{aligned}$$

Using Reynold's transport theorem (3.2) and $u(x,0) = 0$ for $x \in \Omega$, this gives

$$\begin{aligned} & 2\,\langle \partial_t u + \boldsymbol{u}\cdot\nabla_y u, u\rangle_Q = 2\int_0^T \int_{\Omega(t)} \frac{d}{dt} u(y,t)\,u(y,t)\,dy\,dt \\ & = \int_0^T \int_{\Omega(t)} \frac{d}{dt}[u(y,t)]^2\,dy\,dt \;=\; \int_0^T \frac{d}{dt}\int_{\Omega(t)} [u(y,t)]^2\,dy\,dt \;=\; \int_{\Omega(T)} [u(y,T)]^2 dy \;\ge\; 0. \end{aligned}$$

Hence we obtain

$$\begin{aligned} \|q_u\|_Y^2 &= \|u+w_u\|_Y^2 \;=\; \|u\|_Y^2 + \|w_u\|_Y^2 + 2\,\langle \nabla_y w_u, \nabla_y u\rangle_{L^2(Q)} \\ &= \|u\|_Y^2 + \|w_u\|_Y^2 + 2\,\langle \partial_t u + \boldsymbol{u}\cdot\nabla_y u, u\rangle_Q \\ &\ge \|u\|_Y^2 + \|w_u\|_Y^2 \;=\; \|u\|_{X_v}^2, \end{aligned}$$

and as in (2.9) we therefore conclude the inf-sup stability condition

$$\|u\|_{X_v} \le \sup_{0\neq q\in Y} \frac{b(u,q)}{\|q\|_Y} \quad \text{for all } u \in X_v. \tag{3.5}$$

In order to apply the Babuška–Nečas theory it remains to prove surjectivity.

Lemma 3.1 *Consider the gradient of the velocity $\boldsymbol{v}(y,t)$,*

$$A_v = \begin{pmatrix} \dfrac{\partial}{\partial y_1} v_1 & \dfrac{\partial}{\partial y_1} v_2 & \dfrac{\partial}{\partial y_1} v_3 \\ \dfrac{\partial}{\partial y_2} v_1 & \dfrac{\partial}{\partial y_2} v_2 & \dfrac{\partial}{\partial y_2} v_3 \\ \dfrac{\partial}{\partial y_3} v_1 & \dfrac{\partial}{\partial y_3} v_2 & \dfrac{\partial}{\partial y_3} v_3 \end{pmatrix}$$

which is assumed to be positive semi-definite. Then, for all $q \in Y\backslash\{0\}$ there exists a $u_q \in X_v$ such that $b(u_q, q) > 0$.

Proof. For any given $q \in Y\backslash\{0\}$ we define $u_q \in X_v$,

$$u_q(y,t) = u_q(\varphi(t,x),t) = \int_0^t q(\varphi(s,x),s)\,ds, \quad \frac{d}{dt}u_q(y,t) = q(y,t).$$

Hence we have

$$b(u_q,q) = \langle \frac{d}{dt}u_q, q\rangle_Q + \langle \nabla_y u_q, \nabla_y q\rangle_{L^2(Q)} = \|q\|^2_{L^2(Q)} + \langle \nabla_y u_q, \nabla_y \frac{d}{dt}u_q\rangle_{L^2(Q)},$$

and it is sufficient to prove that the second summand is non-negative. A direct computation gives

$$\nabla_y u_q \cdot \nabla_y \frac{d}{dt}u_q \;=\; \frac{1}{2}\,\frac{d}{dt}\,|\nabla_y u_q|^2 + (A_v \nabla_y u_q, \nabla_y u_q),$$

and therefore,

$$\begin{aligned}\langle \nabla_y u_q, \nabla_y \frac{d}{dt}u_q\rangle_{L^2(Q)} \;&=\; \int_0^T \int_{\Omega(t)} \left[\frac{1}{2}\,\frac{d}{dt}\,|\nabla_y u_q|^2 + (A_v \nabla_y u_q, \nabla_y u_q)\right] dydt \\ &= \frac{1}{2}\int_{\Omega(T)} |\nabla_y u_q(y,T)|^2\,dy + \int_0^T \int_{\Omega(t)} (A_v \nabla_y u_q, \nabla_y u_q)\,dy\,dt \;\geq\; 0\end{aligned}$$

follows. This concludes the proof. ∎

Example 3.1 *Consider the deformation $y = \varphi(t,x)$ to be a planar rotation, i.e., using polar coordinates we have*

$$y(t) = r\begin{pmatrix} \cos(\phi+\alpha t) \\ \sin(\phi+\alpha t)\end{pmatrix}, \quad \boldsymbol{v}(y,t) = \frac{d}{dt}y(t) = \begin{pmatrix} -y_2(t) \\ y_1(t)\end{pmatrix}.$$

In this case we conclude

$$A_v = \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}, \quad (A_v \nabla_y u, \nabla_y u) = 0\,.$$

This particular case was already considered in [25].

Example 3.2 *For any velocity field $\boldsymbol{v}(y,t) = \boldsymbol{v}(t)$ we obviously have $A_v = 0$, and the assumption of Lemma 3.1 is trivially satisfied.*

As in Lemma 2.5 we now conclude unique solvability of the variational formulation (3.4), since all assumptions of the Babuška–Nečas theory are satisfied.

3.2 Adaptive least-squares space-time finite element method

Instead of the bilinear form $b(u,q)$ as defined in (2.3) we now consider the definition of $b(u,q)$ as used in (3.4) for $(u,q) \in X_v \times Y$, and instead of the bilinear form $a(w_u,q)$ as defined in (2.6) we now use the definition as given in (3.3). With these definitions we consider the minimization of the quadratic functional (2.25) whose minimizer $u \in X_v$ is given as the unique solution of the gradient equation (2.26). Instead of (2.27) and (2.28) we therefore have to solve the coupled system to find $(u,p) \in X_v \times Y$ such that

$$\begin{aligned} &\int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y p(y,t) \cdot \nabla_y q(y,t)\, dy\, dt \\ &\quad + \int_0^T \int_{\Omega(t)} \Big[\partial_t u(y,t) + \boldsymbol{v}(y,t) \cdot \nabla_y u(y,t)\Big] q(y,t)\, dy\, dt \\ &\quad + \int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y u(y,t) \cdot \nabla_y q(y,t)\, dy\, dt \;=\; \int_0^T \int_{\Omega(t)} f(y,t) q(y,t)\, dy\, dt \end{aligned} \tag{3.6}$$

is satisfied for all $q \in Y$, and

$$\begin{aligned} &\int_0^T \int_{\Omega(t)} \Big[\partial_t z(y,t) + \boldsymbol{v}(y,t) \cdot \nabla_y z(y,t)\Big] p(y,t)\, dy\, dt \\ &\qquad + \int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y z(y,t) \cdot \nabla_y p(y,t)\, dy\, dt \;=\; 0 \end{aligned} \tag{3.7}$$

is satisfied for all $z \in X_v$. Unique solvability of the mixed variational formulation (3.6) and (3.7) follows again as in [29], in particular when using the inf-sup stability condition (3.5), and the ellipticity of the bilinear form $a(q,q) = \|q\|_Y^2$ for all $q \in Y$. Since $B : X_v \to Y^*$ is surjective, $p \equiv 0$ follows.

For the space-time finite element discretization of the mixed variational formulation (3.6) and (3.7) we use the conforming finite element space $X_H = \operatorname{span}\{\varphi_k\}_{k=1}^M \subset X_v$ of piecewise linear continuous basis functions φ_k which are defined with respect to some admissible locally quasi-uniform decomposition of the space-time domain Q into simplicial space-time finite elements τ_ℓ of mesh size H_ℓ. In addition, we introduce a space-time finite element space $Y_h = \operatorname{span}\{\phi_i\}_{i=1}^{\overline{M}} \subset Y$ of piecewise linear continuous basis functions ϕ_i which are defined with respect to a possibly refined decomposition of Q into simplicial finite elements $\overline{\tau}_\ell$ of local mesh size h_ℓ. As in the stationary case we may use one additional refinement, i.e., $h_\ell = H_\ell/2$ when $\overline{\tau}_\ell \subset \tau_\ell$.

According to (3.3) we define $w_{u,h} \in Y_h$ as unique solution of the variational formulation such that

$$\begin{aligned} &\int_0^T \int_{\Omega(t)} \alpha(y)\, \nabla_y w_{u,h}(y,t) \cdot \nabla_y q_h(y,t)\, dy\, dt \\ &\qquad = \int_0^T \int_{\Omega(t)} \Big[\partial_t u(y,t) + \boldsymbol{v}(y,t) \cdot \nabla_y u(y,t)\Big]\, q_h(y,t)\, dy\, dt \end{aligned} \tag{3.8}$$

is satisfied for all $q_h \in Y_h$. Hence we can define the discrete norm

$$\|u\|_{X_v,h} := \sqrt{\|u\|_Y^2 + \|w_{u,h}\|_Y^2} \leq \|u\|_{X_v} \quad \text{for all } u \in X_v,$$

and as in (3.5) we can establish the discrete inf-sup stability condition

$$\|u_H\|_{X_v,h} \leq \sup_{0 \neq q_h \in Y_h} \frac{b(u_H, q_h)}{\|q_h\|_Y} \quad \text{for all } u_H \in X_H, \tag{3.9}$$

which remains true for all $X_H \subseteq Y_h$. Hence we conclude unique solvability of the space-time finite element discretization of (3.6) and (3.7) to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\begin{aligned} &\int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y p_h(y,t) \cdot \nabla_y q_h(y,t) \, dy \, dt \\ &\quad + \int_0^T \int_{\Omega(t)} \Big[\partial_t u_H(y,t) + \boldsymbol{v}(y,t) \cdot \nabla_y u_H(y,t) \Big] q_h(y,t) \, dy \, dt \\ &\qquad + \int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y u_H(y,t) \cdot \nabla_y q_h(y,t) \, dy \, dt \; = \; \int_0^T \int_{\Omega(t)} f(y,t) q_h(y,t) \, dy \, dt \end{aligned} \tag{3.10}$$

is satisfied for all $q_h \in Y_h$, and

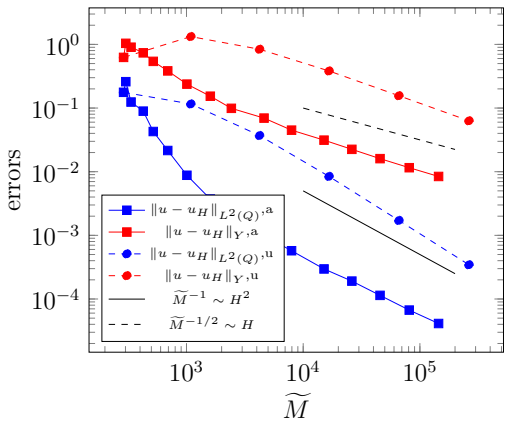
$$\begin{aligned} &\int_0^T \int_{\Omega(t)} \Big[\partial_t z_H(y,t) + \boldsymbol{v}(y,t) \cdot \nabla_y z_H(y,t) \Big] p_h(y,t) \, dy \, dt \\ &\qquad + \int_0^T \int_{\Omega(t)} \alpha(y) \nabla_y z_H(y,t) \cdot \nabla_y p_h(y,t) \, dy \, dt \; = \; 0 \end{aligned} \tag{3.11}$$

is satisfied for all $z_H \in X_H$. Moreover, and as in [25, 29, 40], we can derive estimates for the space-time finite element error $\|u - u_H\|_Y \leq c\, H\, |u|_{H^2(Q)}$ when assuming $u \in H^2(Q)$. While for $Y_h = X_H$ we have $p_h \equiv 0$, in the more general case $X_h \subset Y_h$ but $Y_h \neq X_H$ and assuming an saturation condition such as in (2.37), we can use $\|p_h\|_Y$ as a posteriori error indicator for $\|u - u_H\|_{X_v,h}$ to drive an adaptive scheme.

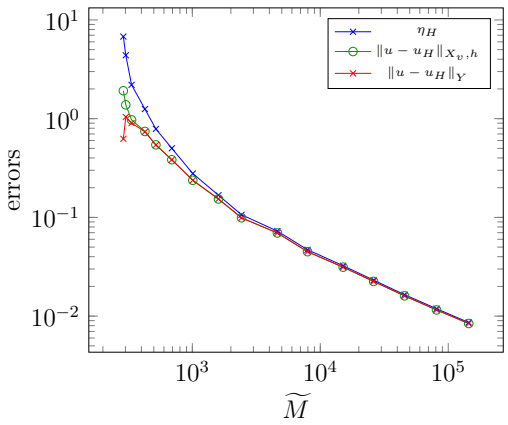
Example 3.3 *As a first example, we consider the one dimensional domain $\Omega = (0,1)$, and the time horizon $T = 1$, i.e., $Q = (0,1)^2$. Further, we consider $\alpha(x) = \varepsilon = 10^{-2}$ and $\boldsymbol{v} = 1$. As exact solution, and similar as in* [14]*, we choose the smooth function*

$$u(x,t) := \left(1 - e^{-t/\varepsilon}\right) \left(\frac{e^{(x-1)/\varepsilon} - 1}{e^{-1/\varepsilon} - 1} + x - 1 \right), \tag{3.12}$$

and we compute $f = \partial_t u - \varepsilon \Delta_x u + \boldsymbol{v} \cdot \nabla_x u$ accordingly. The smooth function (3.12) *exhibits a spatial (at $x = 1$) and a temporal (at $t = 0$) boundary layer. The numerical results for both a uniform and an adaptive refinement strategy are shown in Fig. 12(a). We observe a rate of $\mathcal{O}(H)$ for the error in the energy norm and $\mathcal{O}(H^2)$ for the L^2 error, as expected. In Fig. 12(b) we present a comparison between the errors $\|u - u_H\|_Y$, $\|u - u_H\|_{X_v,h}$ and the error estimator $\eta_H = \|p_h\|_Y$. One can see that the error indicator is effective and that the error in the norm $\|\cdot\|_{X_v,h}$ is mainly driven by the spatial part of the norm. Finally, in Fig. 13 we present the related finite element mesh and the numerical solution u_H.*

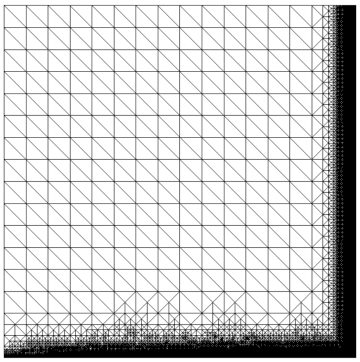


(a) Errors $\|u-u_H\|_Y$ and $\|u-u_H\|_{L^2(Q)}$ for uniform and adaptive refinement strategies.

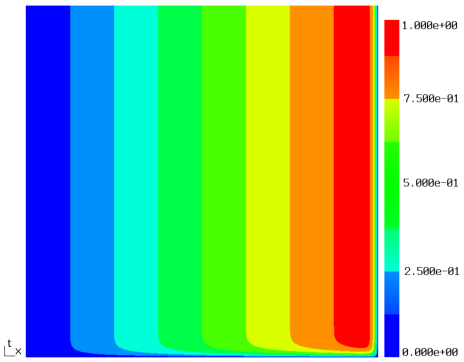


(b) Comparison between estimator and true error for an adaptive refinement strategy.

Figure 12: Convergence results in the case of a smooth solution for a nonstationary convection-diffusion equation.

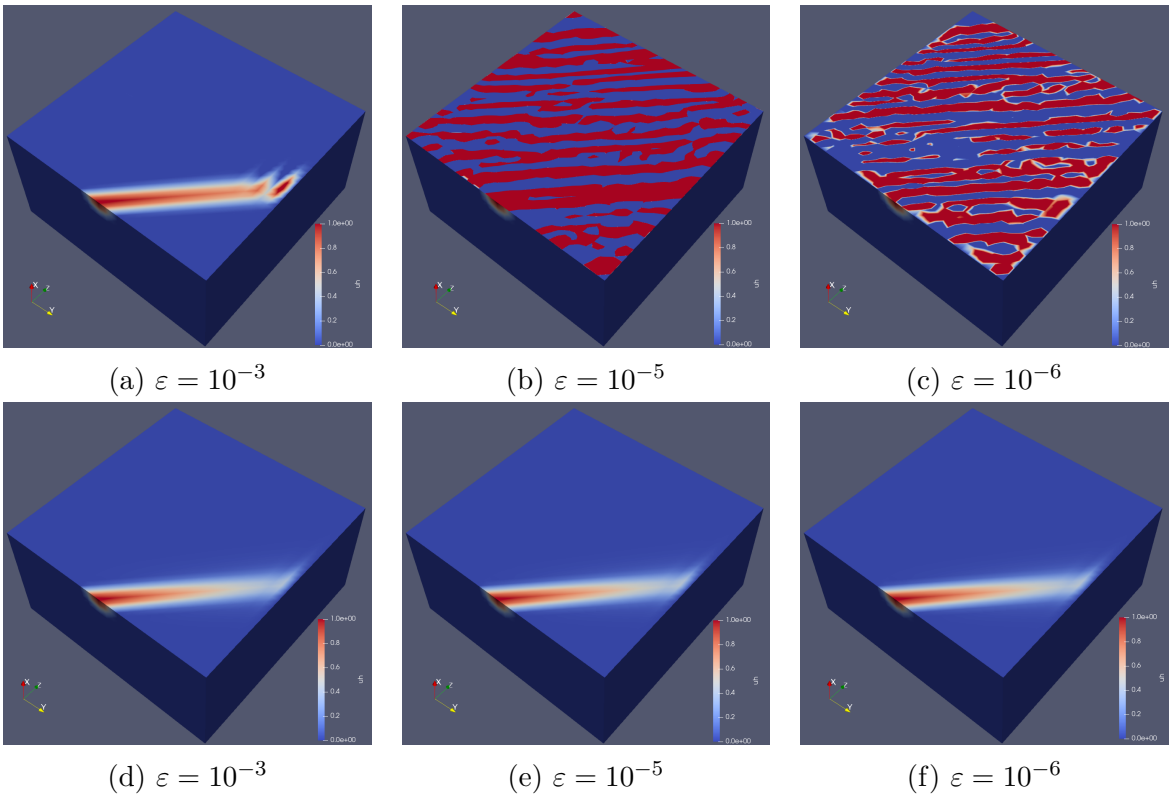


(a) Adaptive mesh on $L = 15$, 144563 dofs



(b) Solution u_H on adaptive mesh

Figure 13: Simulation results for the adaptive refinement process.



(a) $\varepsilon = 10^{-3}$ (b) $\varepsilon = 10^{-5}$ (c) $\varepsilon = 10^{-6}$

(d) $\varepsilon = 10^{-3}$ (e) $\varepsilon = 10^{-5}$ (f) $\varepsilon = 10^{-6}$

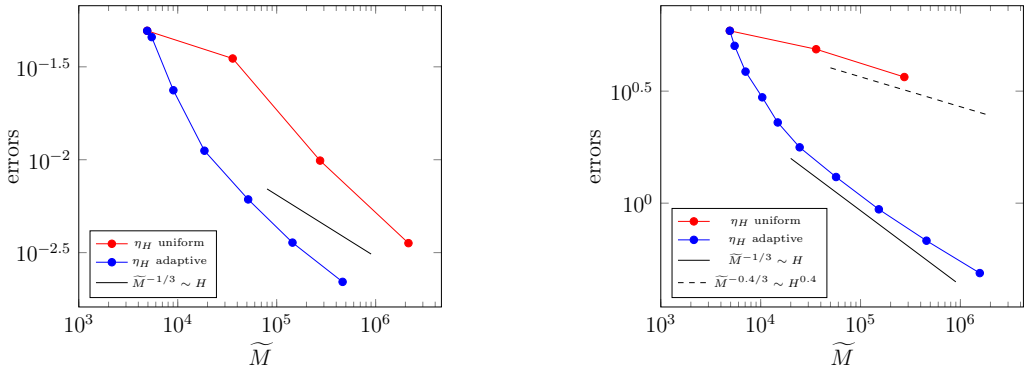
Figure 14: Numerical results for $\boldsymbol{v} = (0,1)^T$ on a mesh with $32 \times 32 \times 32$ elements. Top: no stabilization via direct formulation [40], bottom: stabilization via developed least-squares formulation.

Example 3.4 *As a second example we consider the two-dimensional domain* $\Omega = (0,1)^2$ *and the time horizon* $T = 1$, *i.e.,* $Q = (0,1)^3$. *As initial state* $u(x,0) = u_0(x)$, $x \in \Omega$, *we consider similar as in* [32] *the function*

$$u_0(x) := \psi(10\,\|x - x_0\|_2), \quad \psi(r) := \begin{cases} (1-r^2)^2, & \text{for } r \le 1, \\ 0, & \text{for } r > 1, \end{cases} \quad x_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

We compute numerical solutions to (3.1) *for the velocity field* $\boldsymbol{v} = (0,1)^\top$ *and without a source term, i.e.* $f \equiv 0$. *Furthermore,* $\alpha(x) = \varepsilon \in \{10^{-3}, 10^{-5}, 10^{-6}\}$. *The results for a mesh with* $32 \times 32 \times 32$ *elements (*35937 *dofs) can be seen in Fig. 14. In the top row one can see the numerical solution* u_H *computed by solving* (3.4) *with the space-time finite element method described in* [40]. *This leads to oscillations in the solution as the mesh size is not sufficiently small. In the bottom row one can see the solution using the developed least-squares formulation with piecewise linear trial and piecewise quadratic test functions. This formulation leads to stable results. In a further step we use the inbuilt error estimator to drive an adaptive refinement scheme for the parameters* $\varepsilon = 10^{-3}$, $\boldsymbol{v} = (0, 0.3)^\top$, *and* $\varepsilon = 10^{-6}$, $\boldsymbol{v} = (0,1)^\top$. *In Fig. 15 the convergence rate of the error estimator in case of a*

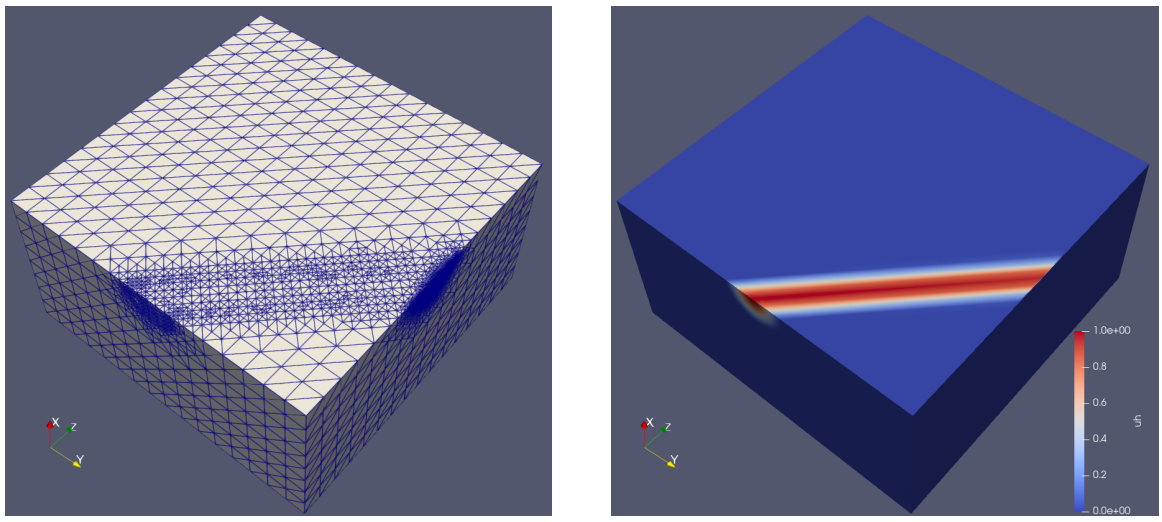
uniform and an adaptive refinement strategy for both sets of parameters is depicted. In the case $\varepsilon = 10^{-3}$, $\boldsymbol{v} = (0, 0.3)^\top$, *we observe a linear rate* $\mathcal{O}(H)$ *for both refinement strategies. In the case* $\varepsilon = 10^{-6}$, $\boldsymbol{v} = (0, 1)^\top$ *we observe a reduced rate of* $\mathcal{O}(H^{0.4})$ *in the uniform case. However, we can recover the full rate of* $\mathcal{O}(H)$ *in the adaptive case. The obtained adaptive meshes as well as the corresponding numerical solutions are depicted in Fig. 16 and 17. Note that we obtain a mesh which is fully unstructured in space and time.*



(a) $\varepsilon = 10^{-3}$, $\boldsymbol{v}$= $(0, 0.3)^\top$ (b) $\varepsilon = 10^{-6}$, $\boldsymbol{v}$= $(0, 1)^\top$

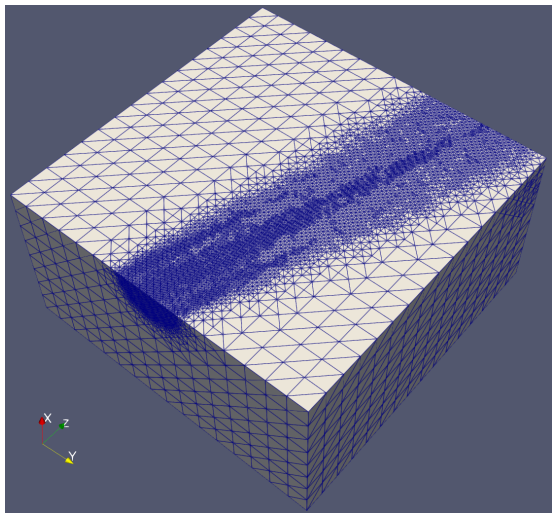
Figure 15: Error estimator $\eta_H = \|p_h\|_Y$ in case of an adaptive and uniform refinement strategy for the second example.

Example 3.5 *As a third example, we consider again the unit cube in the space-time domain, i.e.,* $Q = (0, 1)^3$. *We choose* $u_0 = 0$, $\alpha(x) = \varepsilon = 10^{-2}$ *and the source term to be*

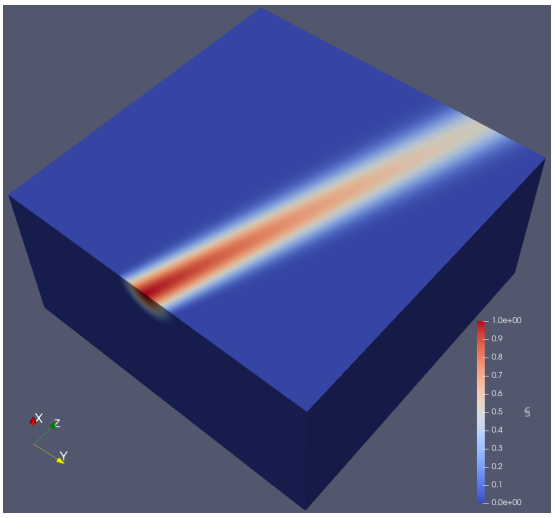


(a) Adaptive mesh on $L = 7$, 152513 dofs (b) Solution u_H on the adaptive mesh

Figure 16: Obtained results for $\varepsilon = 10^{-6}$ and $\boldsymbol{v}$= $(0, 1)^\top$ after the adaptive refinement process.



(a) Adaptive mesh on $L = 6$, 463696 dofs



(b) Solution u_H on adaptive the mesh

Figure 17: Obtained results for $\varepsilon = 10^{-3}$ and $\boldsymbol{v}= (0, 0.3)^\top$ after the adaptive refinement process.

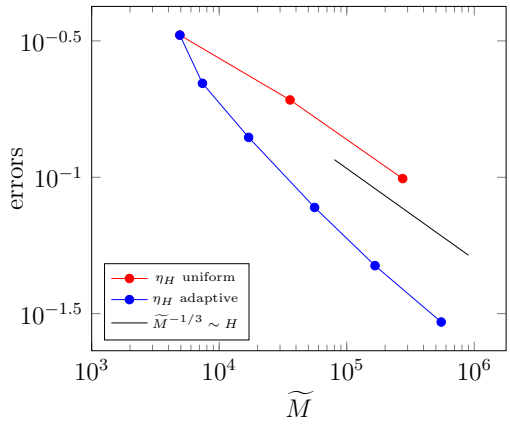
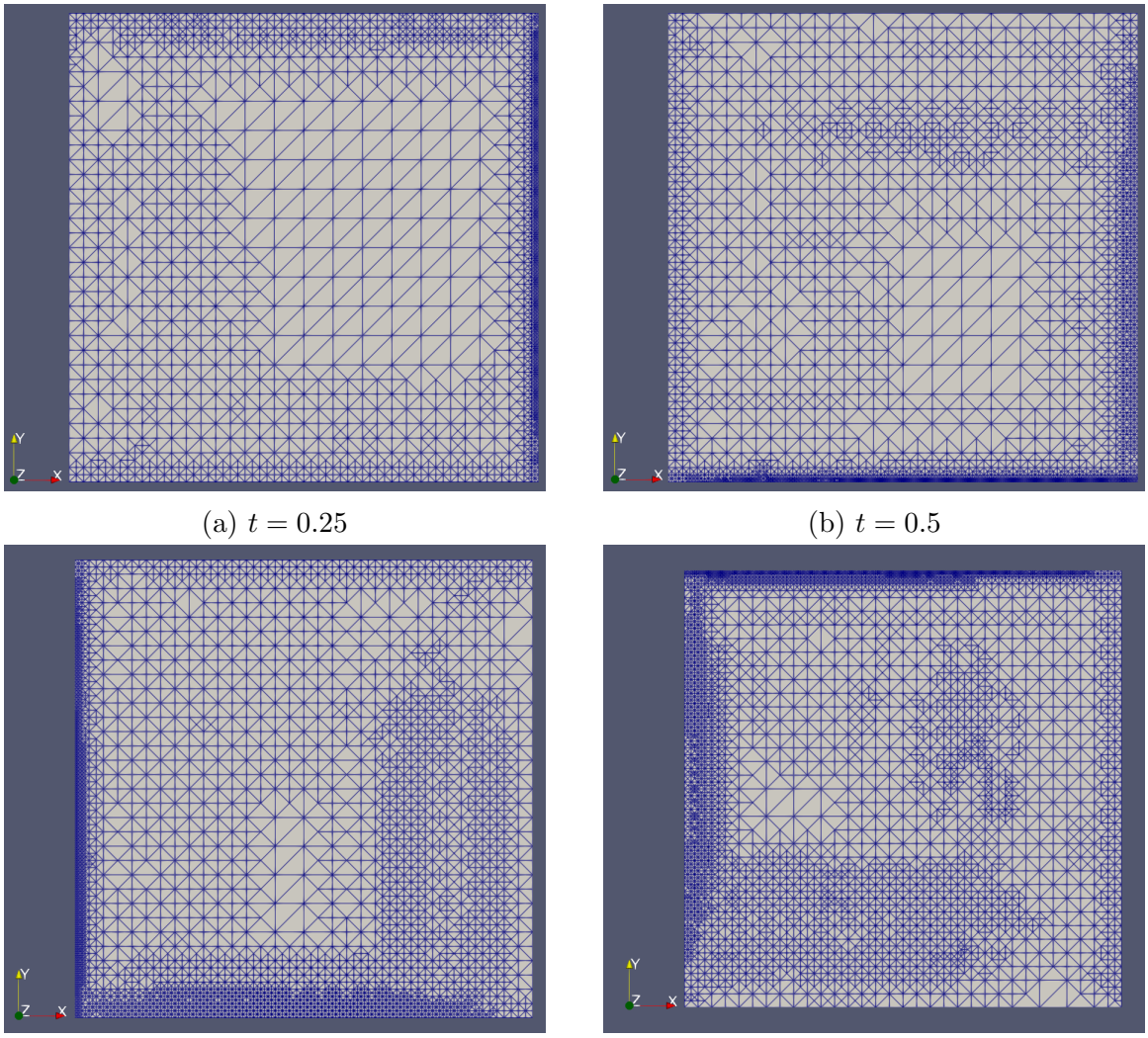


Figure 18: Error estimator $\eta_H = \|p_h\|_Y$ in case of an adaptive and uniform refinement strategy for the third example.

$f = 1$. The velocity field is a time dependent function with $\boldsymbol{v}(y,t) = (\sin(2\pi t), \cos(2\pi t))^\top$ for $(y,t) \in Q$. Thus, the solution u has a boundary layer whose location depends on time. Note that a similar example is considered in [14, Example 4]*. In Fig. 18 a comparison of the error estimator in case of an adaptive and uniform refinement strategy is depicted. We observe a convergence rate of $\mathcal{O}(H)$ in both cases. The generated grids in the adaptive case at different fixed times can be seen in Fig. 19. The circular movement of the boundary layer in time is visible.*



(a) $t = 0.25$ (b) $t = 0.5$

(c) $t = 0.75$ (d) $t = 1$

Figure 19: Generated mesh on refinement level $L = 5$ at different times t in case of the time dependent velocity field $\boldsymbol{v}(x,t) = (\sin(2\pi t), \cos(2\pi t))^\top$.

Acknowledgement: This work is supported by the joint DFG/FWF Collaborative Research Centre CREATOR (DFG: Project-ID 492661287/TRR361; FWF: 10.55776/F90) at TU Darmstadt, TU Graz and JKU Linz.

References

[1] R. Andreev: Stability of sparse space-time finite element discretizations of linear parabolic evolution equations. IMA J. Numer. Anal. 33 (2013) 242–260.

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

[3] I. Babuška: Error-bounds for finite element method. Numer. Math. 16 (1971) 322–333.

[4] C. Bacuta, C. Bacuta, D. Hayes: Comparison of variational discretizations for a convection-diffusion problem. Rev. Roumaine Math. Pures Appl. 69 (2024) 327–351.

[5] C. Bacuta, D. Hayes, T. O'Grady: Saddle point least squares discretization for convection-diffusion. Appl. Anal. 103 (2024) 2241–2268.

[6] R. E. Bank, P. S. Vassilevski, L. T. Zikatanov: Arbitrary dimension convection-diffusion schemes for space-time discretizations. J. Comput. Appl. Math. 310 (2017) 19–31.

[7] P. B. Bochev, M. D. Gunzburger: Least-squares finite element methods. Applied Mathematical Sciences, vol. 166, Springer, New York, 2009.

[8] D. Boffi, F. Brezzi, M. Fortin: Mixed finite element methods and applications. Springer Series in Computational Mathematics, vol. 44, Springer, Heidelberg, 2013.

[9] D. Braess: Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics. Cambridge University Press, Cambridge, UK, 2007.

[10] J. H. Bramble, R. D. Lazarov, J. E. Pasciak: A least-squares approach based on a discrete minus one inner product for first order systems. Math. Comput. 66 (1997) 935–955.

[11] J. H. Bramble, R. D. Lazarov, J. E. Pasciak: Least-squares for second-order elliptic systems. Comput. Methods Appl. Mech. Engrg. 152 (1998) 195–210.

[12] A. N. Brooks, 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) 199–259.

[13] E. Burman, A. Ern: Implicit-explicit Runge–Kutta schemes and finite elements with stabilization for advection-diffusion equations. ESAIM Math. Model. Numer. Anal. 46 (2012) 681–707.

[14] A. Cangiani, E. H. Georgoulis, S. Metcalfe: Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems. IMA J. Numer. Anal. 34 (2014) 1578–1597.

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

[16] C. Carstensen: A note on the quasi-best approximation constant. Comput. Methods Appl. Math. 25 (2025) 581–585.

[17] A. Cohen, W. Dahmen, G. Welper: Adaptivity and variational stabilization for convection-diffusion equations. ESAIM Math. Model. Numer. Anal. 46 (2012) 1247–1273.

[18] W. Dahmen, H. Monsuur, R. Stevenson: Least squares solvers for ill-posed PDEs that are conditionally stable. ESAIM Math. Model. Numer. Anal. 57 (2023) 2227–2255.

[19] L. Demkowicz, N. Heuer: Robust DPG method for convection-dominated diffusion problems. SIAM J. Numer. Anal. 51 (2013) 2514–2537.

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

[21] A. Ern, J.–L. Guermond: Theory and Practice of Finite Elements. Springer, New York, 2004.

[22] L. P. Franca, G. Hauke, A. Masud: Revisiting stabilized finite element methods for the advective-diffusive equation. Comput. Methods Appl. Mech. Engrg. 195 (2006) 1560–1572.

[23] T. Führer: On a mixed FEM and a FOSLS with H^{-1} loads. Comput. Methods Appl. Math. 24 (2024) 355–370.

[24] T. Führer, N. Heuer, M. Karkulik: MINRES for second-order PDEs with singular data. SIAM J. Numer. Anal. 60 (2022) 1111–1135.

[25] P. Gangl, M. Gobrial, O. Steinbach: A space-time finite element method for the eddy current approximation of rotating electric machines. Comput. Methods Appl. Math. 25 (2025) 441–457.

[26] P. W. Hemker: A singularly perturbed model problem for numerical computation. J. Comput. Appl. Math. 76 (1996) 277–285.

[27] A. Jha, V. John, P. Knobloch: Adaptive grids in the context of algebraic stabilizations for convection-diffusion-reaction equations. SIAM J. Sci. Comput. 45 (2023) B564–B589.

[28] V. John, P. Knobloch, O. Pártl: A numerical assessment of finite element discretizations for convection-diffusion-reaction equations satisfying discrete maximum principles. Comput. Methods Appl. Math. 23 (2023) 969–988.

[29] C. Köthe, R. Löscher, O. Steinbach: Adaptive least-squares space-time finite element methods. arXiv:2309.14300, 2023.

[30] R. D. Lazarov, P. S. Vassilevski: Least-squares streamline diffusion finite element approximations to singularly perturbed convection-diffusion problems. Technical Report 1999/2, Institute for Scientific Computation, Texas A&M University, College Station, 1999.

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

[32] M. Łoś, P. Sepúlveda, M. Sikora, M. Paszyński: Solver algorithm for stabilized space-time formulation of advection-dominated diffusion problem. Comput. Math. Appl. 152 (2023) 67–80.

[33] H. Monsuur, R. Stevenson, J. Storn: Minimal residual methods in negative or fractional Sobolev norms. Math. Comp. 93 (2024) 1027–1052.

[34] J. Nečas, Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. Ann. Sc. Norm. Super. Pisa, 16 (1962) 305–326.

[35] D. R. Q. Pacheco, O. Steinbach: On the initial higher-order pressure convergence in equal-order finite element discretizations of the Stokes system. Comput. Math. Appl. 109 (2022) 140–145.

[36] A. I. Pehlivanov, G. F. Carey, R. D. Lazarov: Least-squares mixed finite elements for second-order elliptic problems. SIAM J. Numer. Anal. 31 (1994) 1368–1377.

[37] H. G. Roos, M. Stynes, L. Tobiska: Robust Numerical Methods for Singularly Perturbed Differential Equations: Convection-Diffusion-Reaction and Flow Problems. Springer Series in Computational Mathematics, vol. 24, Springer, Berlin, Heidelberg, 2008.

[38] A. Schafelner, P. S. Vassilevski: Numerical results for adaptive (negative norm) constrained first order system least squares formulations. Comput. Math. Appl. 95 (2021) 256–270.

[39] O. Steinbach: A note on initial higher order convergence results for boundary element methods with approximated boundary conditions. Num. Methods Part. Diff. Eq. 16 (2000) 581–588.

[40] O. Steinbach: Space-time finite element methods for parabolic problems. Comput. Methods Appl. Math. 15 (2015) 551–566.

[41] R. Stevenson, J. Westerdiep: Stability of Galerkin discretizations of a mixed space-time variational formulation of parabolic evolution equations. IMA J. Numer. Anal. 41 (2021) 28–47.

[42] I. Toulopoulos: SUPG space-time scheme on anisotropic meshes for general parabolic equations. J. Numer. Math., published online, 2025.

[43] J. Xu, L. Zikatanov: Some observations on Babuška and Brezzi theories. Numer. Math., 94 (2003) 195–202.