Vít Dolejší Miloslav Feistauer

Analysis of Semi-Implicit DGFEM for Nonlinear Convection-Diffusion Problems

Preprint No. MATH-knm-2004/1

30.8.2004

Submitted *Numerische Mathematik*

Department of Numerical Mathematics, Faculty of Mathematics and Physics,
Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic

http://www.karlin.mff.cuni.cz/ms-preprints/

Cover print by Reprostředisko, Malostranské nám. 25, Praha 1 - Malá Strana
Design by Mirko Rokyta, picture of snowflake by Ondřej Kalenda
Current series administrator: V. Dolejsi

1

Numerische Mathematik manuscript No.
(will be inserted by the editor)

Analysis of Semi-Implicit DGFEM for Nonlinear Convection-Diffusion Problems

Vít Dolejší, Miloslav Feistauer **

Charles University Prague, e-mail: {dolejsi,feist}@karlin.mff.cuni.cz

Received: date / Revised version: date

Summary The paper deals with the numerical analysis of a scalar nonstationary nonlinear convection-diffusion equation. The space discretization is carried out by the discontinuous Galerkin finite element method (DGFEM) with nonsymmetric treatment of stabilization terms and interior and boundary penalty. The time discretization is carried out by a semi-implicit Euler scheme, in which the diffusion and stabilization terms are treated implicitly whereas, the nonlinear convective terms are treated explicitly. We derive a priori asymptotic error estimates in the discrete $L^\infty(L^2)$–norm, $L^2(H^1)$–seminorm and $L^\infty(H^1)$–seminorm with respect to the mesh size h and time step τ. Several numerical examples demonstrating the accuracy of piecewise linear, quadratic and cubic finite elements are presented.

Key words nonlinear convection-diffusion equation – discontinuous Galerkin finite element method – nonsymmetric treatment of stabilization terms – interior and boundary penalty – semi-implicit scheme – a priori error estimates – experimental order of convergence

* This work was supported under the Grant No. 201/02/0684 of the Czech Grant Agency and by the Grant No. MSM 113200007 of the Ministry of Education of the Czech Republic. The authors acknowledge this support.

** *Present address:* Charles University Prague, Faculty of Mathematics and Physics, Sokolovská 83, 186 75 Praha 8, Czech Republic

1 Introduction

The investigation of convection-diffusion problems is a very topical subject. These problems play an important role, e. g. in fluid dynamics, hydrology, heat and mass transfer, environmental protection on one side, but also financial mathematics, image processing on other side. There is an extensive literature, devoted mainly to linear convection-diffusion problems, represented by monographs [31], [37], [38], [25], and a number of papers. (Let us mention, e. g. [3], [26], [29], [30], [33], [41].)

The main difficulty which must be overcome in the numerical analysis of convection-diffusion problems is the precise resolution of the so-called boundary layers. If the equation in consideration represents a nonlinear conservation law with a small dissipation, then beside boundary layers also shock waves appear (slightly smeared due to dissipation). This is particularly the case of the system of the Navier-Stokes equations describing viscous compressible flow.

Our aim is to developed a sufficiently robust and accurate numerical method for the solution of nonlinear convection-diffusion equations and compressible flow. To this end, we derived and analyzed combined finite volume (FV) – finite element (FE) methods ([1], [12], [18], [19], [21]). We obtained a priori error estimates and applied these methods to the solution of the compressible Navier-Stokes equations (see [16], [17], [23]). The main drawback of this approach was a low order of convergence.

A natural generalization of the FV and FE methods is the *discontinuous Galerkin finite element method* (DGFEM), which appears to be very suitable for problems with solutions containing discontinuities and/or steep gradients. The DGFEM is based on piecewise polynomial but discontinuous approximations. It uses advantages of the FV as well as FE methods. Similarly as in the finite volume method, the DGFEM uses discontinuous approximations and boundary fluxes are evaluated with the aid of a numerical flux, which allow a precise capturing of discontinuities and steep gradients. Similarly as in the finite elements method, the DGFEM uses higher degree polynomial approximations of solutions, which produces an accurate resolution in regions, where the solution is smooth. There is a number of works devoted to theory and applications of the DGFEM. Let us mention, e.g., [2], [5], [6], [11], [14], [22], [24], [32], [36]. For a survey of various discontinuous Galerkin techniques, see, e. g. [9], [10].

In this paper we shall be concerned with the analysis of the DGFEM for the numerical solution of a nonlinear convection-diffusion initial-boundary value problem. There are several variants of the

DGFEM for the solution of problems containing diffusion terms. It is possible to use primitive variables or a mixed method. The method can be stabilized with the aid of a symmetric or nonsymmetric treatment of diffusion terms, often combined with an interior and boundary penalty. We consider here the nonsymmetric variant with the interior and boundary penalty (denoted as NIPG method). This stabilization technique was proposed in [5] and [6] and represents the generalization of the boundary penalty by Babuška and Zlámal allowing to impose the Dirichlet boundary condition in a weak sense instead of building it in the finite element space (see [4]). The nonsymmetric variant was also investigated in [36] for elliptic and parabolic problems and in [14] and [15] for nonlinear convection-diffusion problems.

Although this approach does not give an optimal order of convergence for elliptic problems (see [2], [36]), it leads to a coercive operator for an arbitrary positive penalty coefficient σ, see Lemma 1 of this paper. This property is important when the DGFEM is applied to the system of the Navier-Stokes equations, where a numerical analysis is very complicated, see [11].

In [13] and [15] we carried out a discretization of a scalar convection-diffusion equation by DGFEM with respect to space variables (the method of lines) and derived a priori error estimates. The time discretization can be carried out by the (explicit) Runge-Kutta methods, which are simple for implementation, but the resulting scheme is conditionally stable and the time step is drastically limited. In order to avoid this disadvantage, it seems suitable to apply an implicit method, which allows to use a much longer time step. However, a fully implicit DGFEM leads to a large, strongly nonlinear algebraic system, whose solution is rather complicated. This is the reason that we propose here a semi-implicit scheme, which appears quite efficient and robust. The linear diffusion and stabilization terms are treated implicitly, whereas the nonlinear convective terms explicitly. In this paper we shall be concerned with theoretical analysis of error estimates of this method and present several numerical experiments verifying the theoretical results.

The contents of the paper is the following. In Section 2, the initial-boundary value problem for a scalar nonlinear convection-diffusion equation is formulated. In Section 3, we carry out the discretization of the problem by the semi-implicit DGFEM and prove the existence and uniqueness of the numerical solution. Section 4 contains some auxiliary results, namely assumptions on the space discretization (allowing even nonconforming grids with nonconvex star-shaped elements) and some important inequalities and estimates. These results are used in Section 5, where error estimates in the discrete $L^\infty(L^2)$–norm, $L^2(H^1)$–seminorm and $L^\infty(H^1)$–seminorm are proven. We obtain also estimates of the error in the penalty terms. In Section 6 we present several numerical examples with piecewise linear, quadratic and cubic approximations, demonstrating the efficiency of the DGFEM. In Section 7 we introduce some concluding remarks and formulate open problems.

2 Continuous problem

Let $\Omega \subset I\!R^d$ ($d = 2$ or 3) be a bounded polyhedral domain and $T > 0$. (For $d = 2$ under the concept of a polyhedral domain we mean a polygonal domain.) We set $Q_T = \Omega \times (0,T)$. By $\overline{\Omega}$ and $\partial\Omega$ we denote the closure and boundary of Ω, respectively. Let us consider the following *initial-boundary value problem*: Find $u : Q_T \to I\!R$ such that

$$\frac{\partial u}{\partial t} + \sum_{s=1}^{d} \frac{\partial f_s(u)}{\partial x_s} = \varepsilon\, \Delta u + g \quad \text{in } Q_T, \tag{1}$$

$$u\Big|_{\partial\Omega\times(0,T)} = u_D, \tag{2}$$

$$u(x,0) = u^0(x), \quad x \in \Omega. \tag{3}$$

We assume that the data satisfy the following conditions:

$$\text{a) } f_s \in C^1(I\!R),\ f_s(0) = 0,\ s = 1,\dots,d, \tag{4}$$

b) $\varepsilon > 0$,

c) $g \in C([0,T]; L^2(\Omega))$,

d) u_D is the trace of some $u^* \in C([0,T]; H^1(\Omega)) \cap L^\infty(Q_T)$ on $\partial\Omega \times (0,T)$,

e) $u^0 \in L^2(\Omega)$.

We use the standard notation for function spaces (see, e. g. [27]): $L^p(\Omega)$, $L^p(Q_T)$ denote the Lebesgue spaces, $W^{k,p}(\Omega)$, $H^k(\Omega) = W^{k,2}(\Omega)$ are the Sobolev spaces, $L^p(0,T;X)$ is the Bochner space of functions p-integrable over the interval $(0,T)$ with values in a Banach space X, $C([0,T];X)$ ($C^1([0,T];X)$) is the space of continuous (continuously differentiable) mappings of the interval $[0,T]$ into X. By $H^1_0(\Omega)$ we denote the subspace of all functions from $H^1(\Omega)$ with zero traces on $\partial\Omega$.

The assumption that $f_s(0) = 0$, $s = 1, \dots, d$, does not cause any loss of generality, as can be seen from equation (1). The functions f_s, called fluxes, represent convective terms, $\varepsilon > 0$ is the diffusion coefficient. The diffusion term can be, in general, more complicated (in some cases even nonlinear as in [36]). It is also possible to consider mixed Dirichlet–Neumann boundary conditions, see [36], [22], [13]. For simplicity we prescribe the Dirichlet condition on the whole boundary.

A sufficiently regular function satisfying (1) – (3) pointwise is called a *classical solution*. It is suitable to introduce the concept of a weak solution. To this end we use the following *notation*:

$$(u, v) = \int_\Omega uv \, \mathrm{d}x, \quad u,\, v \in L^2(\Omega) \tag{5}$$

(scalar product in $L^2(\Omega)$),

$$\|u\|_{L^2(\Omega)} = (u, u)^{1/2} \tag{6}$$

(norm in $L^2(\Omega)$),

$$a(u, v) = \varepsilon \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x, \quad u,\, v \in H^1(\Omega), \tag{7}$$

$$b(u, v) = \int_\Omega \sum_{s=1}^{d} \frac{\partial f_s(u)}{\partial x_s} v \, \mathrm{d}x, \quad u \in H^1(\Omega) \cap L^\infty(\Omega), \tag{8}$$

$$v \in L^2(\Omega),$$

$$\|u\|_{H^1(\Omega)} = \left(\int_\Omega \left(|u|^2 + |\nabla u|^2 \right) \mathrm{d}x \right)^{1/2}, \quad u \in H^1(\Omega), \tag{9}$$

(norm in $H^1(\Omega)$),

$$|u|_{H^1(\Omega)} = \left(\int_\Omega |\nabla u|^2 \, \mathrm{d}x \right)^{1/2}, \quad u \in H^1(\Omega), \tag{10}$$

(seminorm in $H^1(\Omega)$). It is known that $|\cdot|_{H^1(\Omega)}$ is a norm on $H^1_0(\Omega)$ equivalent to $\|\cdot\|_{H^1(\Omega)}$.

Definition 1 *We say that a function u is a* weak solution *of (1) – (3), if the following conditions are satisfied*

a) $\quad u - u^* \in L^2(0, T; H^1_0(\Omega)), \quad u \in L^\infty(Q_T),$ (11)

b) $\quad \dfrac{\mathrm{d}}{\mathrm{d}t}(u(t), v) + b(u(t), v) + a(u(t), v) = (g(t), v)$

for all $v \in H^1_0(\Omega)$ in the sense of distributions on $(0, T)$,

c) $\quad u(0) = u_0 \quad$ *in Ω.*

By $u(t)$ we denote the function on Ω such that $u(t)(x) = u(x,t)$, $x \in \Omega$.

With the aid of techniques from [28] and [35], it is possible to prove that there exists a unique weak solution. Moreover, it satisfies the condition $\partial u/\partial t \in L^2(Q_T)$. Then (11), b) can be rewritten as

$$\left(\frac{\partial u(t)}{\partial t}, v\right) + b(u(t),v) + a(u(t),v) = (g(t),v) \tag{12}$$
$$\text{for all } v \in H_0^1(\Omega) \text{ and almost every } t \in (0,T).$$

We shall assume that the weak solution u is sufficiently regular, namely,

$$\begin{aligned} &u \in L^\infty(0,T;H^{p+1}(\Omega)), \\ &\frac{\partial u}{\partial t} \in L^\infty(0,T;H^{p+1}(\Omega)), \\ &\frac{\partial^2 u}{\partial t^2} \in L^\infty(0,T;L^2(\Omega)), \end{aligned} \tag{13}$$

where an integer $p \geq 1$ will denote a given degree of polynomial approximations. Such a solution satisfies problem (1) – (3) pointwise. Under (13),

$$u \in C([0,T];H^{p+1}(\Omega)), \quad \frac{\partial u}{\partial t} \in C([0,T];L^2(\Omega)). \tag{14}$$

3 Discretization of the problem

3.1 Triangulations

Let $\mathcal{T}_h$ $(h > 0)$ be a partition of the closure $\overline{\Omega}$ of the domain Ω into a finite number of closed d-dimensional (convex or nonconvex) polyhedra K with mutually disjoint interiors. We call $\mathcal{T}_h$ a triangulation of $\overline{\Omega}$ and *do not* require the conforming properties from the finite element method. In 2D problems we choose usually $K \in \mathcal{T}_h$ as triangles or quadrilaterals. In 3D, $K \in \mathcal{T}_h$ can be, e. g., tetrahedra, pyramids or hexahedra, but we can construct even more general elements K, as dual finite volumes from [19]. In the analysis carried out in Section 4 it is important to assume that all elements $K \in \mathcal{T}_h$ are star-shaped.

In our further considerations we shall use the following *notation*. We set $h_K = \mathrm{diam}(K)$, $h = \max_{K\in\mathcal{T}_h} h_K$. By ρ_K we denote the radius of the largest d-dimensional ball inscribed into K and by $|K|$ we denote the d-dimensional Lebesgue measure of K. All elements of $\mathcal{T}_h$ will be numbered so that $\mathcal{T}_h = \{K_i\}_{i\in I}$, where $I \subset Z^+ = \{0,1,2,\dots\}$

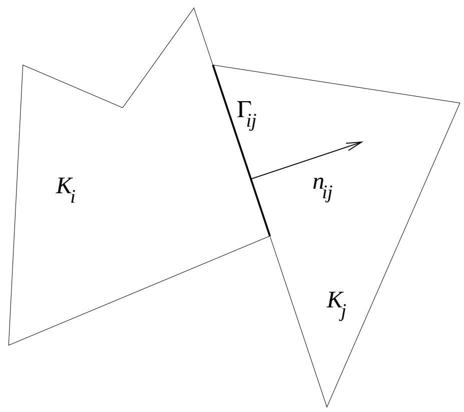


Fig. 1. Neighbouring elements $K_i,\ K_j$

is a suitable index set. If two elements K_i, $K_j \in \mathcal{T}_h$, $i \neq j$, contain a nonempty common open face, we call them *neighbours*. In this case we set

$$\Gamma_{ij} = \partial K_i \cap \partial K_j. \tag{15}$$

(See Figure 1, showing a possible 2D situation.) For $i \in I$ we set

$$s(i) = \{j \in I; K_j \text{ is a neighbour of } K_i\}. \tag{16}$$

The boundary $\partial\Omega$ is formed by a finite number of faces of elements K_i adjacent to $\partial\Omega$. We denote all these boundary faces by S_j, where $j \in I_b \subset Z^- = \{-1, -2, \dots\}$ and set

$$\begin{aligned} \gamma(i) &= \{j \in I_b; S_j \text{ is a face of } K_i\}, \\ \Gamma_{ij} &= S_j \text{ for } K_i \in \mathcal{T}_h \text{ such that } S_j \subset \partial K_i,\ j \in I_b. \end{aligned} \tag{17}$$

For K_i not containing any boundary face S_j we set $\gamma(i) = \emptyset$. Obviously, $s(i) \cap \gamma(i) = \emptyset$ for all $i \in I$. Now, if we write $S(i) = s(i) \cup \gamma(i)$, we have

$$\partial K_i = \bigcup_{j \in S(i)} \Gamma_{ij}, \quad \partial K_i \cap \partial\Omega = \bigcup_{j \in \gamma(i)} \Gamma_{ij}. \tag{18}$$

Furthermore, we use the following notation: $\boldsymbol{n}_{ij} = ((n_{ij})_1, \dots, (n_{ij})_d)$ – unit outer normal to ∂K_i on the face Γ_{ij} (see Figure 1), $|\Gamma_{ij}|$ – $(d-1)$-dimensional measure of Γ_{ij} (i. e., the length of Γ_{ij}, if $d = 2$, and the area of Γ_{ij}, if $d = 3$), $d(\Gamma_{ij}) = \mathrm{diam}(\Gamma_{ij})$. Obviously, $\boldsymbol{n}_{ij} = -\boldsymbol{n}_{ji}$ and, from the point of view of point sets, we have $\Gamma_{ij} = \Gamma_{ji}$, if $i \in I$, $j \in s(i)$ (and, hence, $j \in I$, $i \in s(j)$).

It is obvious that for $i \in I$ and $j \in S(i)$ we have

$$|K_i| \leq h_{K_i}^d \leq h^d, \tag{19}$$

$$d(\Gamma_{ij}) \leq h_{K_i} \leq h. \tag{20}$$

3.2 Broken Sobolev spaces

Over the triangulation $\mathcal{T}_h$ we define the so-called *broken Sobolev space*

$$H^k(\Omega, \mathcal{T}_h) = \{v; v|_K \in H^k(K)\ \forall\, K \in \mathcal{T}_h\} \tag{21}$$

and define there the norm

$$\|v\|_{H^k(\Omega,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} \|v\|^2_{H^k(K)} \right)^{1/2} \tag{22}$$

and the seminorm

$$|v|_{H^k(\Omega,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} |v|^2_{H^k(K)} \right)^{1/2}. \tag{23}$$

For $v \in H^1(\Omega, \mathcal{T}_h)$, $i \in I$ and $j \in s(i)$ we introduce the following notation:

$$\begin{aligned} v|_{\Gamma_{ij}} &= \text{the trace of } v|_{K_i} \text{ on } \Gamma_{ij}, \\ v|_{\Gamma_{ji}} &= \text{the trace of } v|_{K_j} \text{ on } \Gamma_{ji}, \\ \langle v\rangle_{\Gamma_{ij}} &= \frac{1}{2}\left(v|_{\Gamma_{ij}} + v|_{\Gamma_{ji}}\right), \\ [v]_{\Gamma_{ij}} &= v|_{\Gamma_{ij}} - v|_{\Gamma_{ji}}. \end{aligned} \tag{24}$$

Obviously, $\langle v\rangle_{\Gamma_{ij}} = \langle v\rangle_{\Gamma_{ji}}$, but $[v]_{\Gamma_{ij}} = -[v]_{\Gamma_{ji}}$. On the other hand, $[v]_{\Gamma_{ij}}\boldsymbol{n}_{ij} = [v]_{\Gamma_{ji}}\boldsymbol{n}_{ji}$.

3.3 Space-time discretization

In order to carry out the space discretization, we start from the weak solution u satisfying (13), multiply equation (1) by an arbitrary $v \in H^2(\Omega, \mathcal{T}_h)$, integrate over each $K_i \in \mathcal{T}_h$ and apply Green's theorem. Summing over all $K_i \in \mathcal{T}_h$, using (18), (2 and the relations

$$[u]_{\Gamma_{ij}} = 0, \quad \langle \nabla u\rangle_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ji}}, \tag{25}$$

$$\int_{\Gamma_{ij}} \langle \nabla v\rangle \cdot \boldsymbol{n}_{ij}\, [u]\, \mathrm{d}S = 0, \int_{\Gamma_{ij}} [u][v]\, \mathrm{d}S = 0 \quad \forall j \in s(i)\ \forall i \in I, \tag{26}$$

we obtain the identity

$$\begin{aligned}
\left(\frac{\partial u}{\partial t}(t), v\right) \ &+ \sum_{i\in I}\left\{\sum_{j\in S(i)}\int_{\Gamma_{ij}}\sum_{s=1}^{d} f_s(u(t))\,(n_{ij})_s\, v|_{\Gamma_{ij}}\,\mathrm{d}S\right. && (27)\\
&-\int_{K_i}\sum_{s=1}^{d} f_s(u(t))\,\frac{\partial v}{\partial x_s}\,\mathrm{d}x + \varepsilon\int_{K_i}\nabla u(t)\cdot\nabla v\,\mathrm{d}x\\
&-\varepsilon\sum_{\substack{j\in s(i)\\ j<i}}\int_{\Gamma_{ij}}\Big(\langle\nabla u(t)\rangle\cdot \boldsymbol{n}_{ij}\,[v] - \langle\nabla v\rangle\ \cdot\boldsymbol{n}_{ij}\,[u(t)]\Big)\,\mathrm{d}S\\
&-\varepsilon\sum_{j\in\gamma(i)}\int_{\Gamma_{ij}}\Big(\nabla u(t)\cdot\boldsymbol{n}_{ij}\,v - \nabla v\cdot\boldsymbol{n}_{ij}\,u\Big)\,\mathrm{d}S\\
&\left.+\varepsilon\sum_{j\in s(i)}\int_{\Gamma_{ij}}\sigma[u]\,[v]\,\mathrm{d}S + \sum_{j\in\gamma(i)}\int_{\Gamma_{ij}}\sigma\,u\,v\,\mathrm{d}S\right\}\\
&= \int_{\Omega} g(t)\,v\,\mathrm{d}x + \varepsilon\sum_{i\in I}\sum_{j\in\gamma(i)}\int_{\Gamma_{ij}}\Big(\nabla v\cdot\boldsymbol{n}_{ij}\,u_D + \sigma\,u_D\,v\Big)\,\mathrm{d}S,
\end{aligned}$$

where the weight function $\sigma : \bigcup_{i\in I}\bigcup_{j\in S(i)}\Gamma_{ij} \to I\!R$ is defined by

$$\sigma|_{\Gamma_{ij}} = \frac{1}{d(\Gamma_{ij})}. \tag{28}$$

We see that due to the boundary condition (2), the terms

$$\varepsilon\int_{\Gamma_{ij}}\left(\nabla v\cdot\boldsymbol{n}_{ij}\,u(t) + \sigma\,u(t)\,v\right)\,\mathrm{d}S$$

and

$$\varepsilon\int_{\Gamma_{ij}}\left(\nabla v\cdot\boldsymbol{n}_{ij}\,u_D(t) + \sigma\,u_D(t)\,v\right)\,\mathrm{d}S,$$

$i\in I,\ j\in\gamma(i)$, on the left-hand side and the right-hand side of (27), respectively, are equal. Above and in the sequel, in the integrals $\int_{\Gamma_{ij}}$, under $\langle\cdot\rangle$ and $[\cdot]$ we understand $\langle\cdot\rangle_{\Gamma_{ij}}$ and $[\cdot]_{\Gamma_{ij}}$, respectively.

Let us note that in (27) the *nonsymmetric variant* of DG representation of the diffusion terms is used. This is called NIPG variant (see, [34]).

The above considerations lead us to the definition of the following forms. For $u, v \in H^2(\Omega, \mathcal{T}_h)$ we set

$$a_h(u,v) = \varepsilon\sum_{i\in I}\left\{\int_{K_i}\nabla u\cdot\nabla v\,\mathrm{d}x\right. \tag{29}$$

$$- \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \Big(\langle \nabla u \rangle \cdot \boldsymbol{n}_{ij} [v] - \langle \nabla v \rangle \cdot \boldsymbol{n}_{ij} [u] \Big)\, \mathrm{d}S$$

$$- \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \Big((\nabla u \cdot \boldsymbol{n}_{ij})\, v - (\nabla v \cdot \boldsymbol{n}_{ij})\, u \Big)\, \mathrm{d}S \Bigg\},$$

$$\tilde{b}_h(u,v) = \sum_{i\in I} \Bigg\{ \sum_{j\in S(i)} \int_{\Gamma_{ij}} \sum_{s=1}^{d} f_s(u)\, (n_{ij})_s\, v|_{\Gamma_{ij}}\, \mathrm{d}S \tag{30}$$

$$- \sum_{i\in I} \int_{K_i} \sum_{s=1}^{d} f_s(u)\, \frac{\partial v}{\partial x_s}\, \mathrm{d}x \Bigg\},$$

$$\ell_h(v)\,(t) = (g(t), v) + \varepsilon \sum_{i\in I} \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \left(\nabla v \cdot \boldsymbol{n}_{ij}\, u_D(t) + \sigma u_D(t)\, v \right)\, \mathrm{d}S, \tag{31}$$

$$J_h^\sigma(u,v) = \sum_{i\in I} \Bigg\{ \sum_{j\in s(i)} \int_{\Gamma_{ij}} \sigma [u]\, [v]\, \mathrm{d}S + \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma u v\, \mathrm{d}S \Bigg\}. \tag{32}$$

In view of (27), the exact solution with properties (13) satisfies the identity

$$\left(\frac{\partial u}{\partial t}(t), v \right) + a_h(u(t), v) + \tilde{b}_h(u(t), v) + \varepsilon J_h^\sigma(u(t), v) = \ell_h(v)\,(t)$$
$$\forall v \in H^2(\Omega, \mathcal{T}_h)\ \forall t \in (0,T). \tag{33}$$

Now we shall approximate the form $\tilde{b}_h$ similarly as in the finite volume method. Thus, the fluxes $\int_{\Gamma_{ij}} \sum_{s=1}^{d} f_s(u)\, n_s\, v\, \mathrm{d}S$ are approximated with the aid of the so-called *numerical flux* $H(u, u', \boldsymbol{n})$ by the expression $\int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}}\, \mathrm{d}S$, for $i \in I$, $j \in S(i)$. Of course, if $j \in \gamma(i)$, then $\Gamma_{ij} \subset \partial\Omega$ and it is necessary to specify the meaning of $u|_{\Gamma_{ji}}$. Here we use the extrapolation, i.e. we set $u|_{\Gamma_{ji}} := u|_{\Gamma_{ij}}$. (In practical computations, usually more sophisticated boundary conditions depending on the behavious of u are used, see Section 6.) In this way we obtain the approximation $b_h(u,v)$ of the convection form $\tilde{b}_h(u,v)$:

$$b_h(u,v) = \sum_{i\in I} \Bigg(\sum_{j\in s(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}}\, \mathrm{d}S \tag{34}$$

$$+ \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ij}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}}\, \mathrm{d}S \Bigg)$$

$$- \sum_{i \in I} \int_{K_i} \sum_{s=1}^{d} f_s(u) \frac{\partial v}{\partial x_s} \, \mathrm{d}x, \quad u, \, v \in H^1(\Omega, \mathcal{T}_h), \; u \in L^\infty(\Omega).$$

We shall assume that the numerical flux has the following properties:

Assumptions (H):

1. $H(u, v, \boldsymbol{n})$ is defined in $I\!R^2 \times \boldsymbol{S}_1$, where $\boldsymbol{S}_1 = \{\boldsymbol{n} \in I\!R^d; |\boldsymbol{n}| = 1\}$, and *Lipschitz-continuous* with respect to u, v: there exists a constant $C_1 > 0$ such that

$$|H(u, v, \boldsymbol{n}) - H(u^*, v^*, \boldsymbol{n})| \leq C_1(|u - u^*| + |v - v^*|), \quad (35)$$
$$u, \, v, \, u^*, \, v^* \in I\!R, \; \boldsymbol{n} \in \boldsymbol{S}_1.$$

2. $H(u, v, \boldsymbol{n})$ is *consistent*:

$$H(u, u, \boldsymbol{n}) = \sum_{s=1}^{d} f_s(u) \, n_s, \quad u \in I\!R, \; \boldsymbol{n} = (n_1, \ldots, n_d) \in \boldsymbol{S}_1. (36)$$

3. $H(u, v, \boldsymbol{n})$ is *conservative*:

$$H(u, v, \boldsymbol{n}) = -H(v, u, -\boldsymbol{n}), \quad u, \, v \in I\!R, \; \boldsymbol{n} \in \boldsymbol{S}_1. \quad (37)$$

In virtue of (35) and (36), the functions f_s, $s = 1, \ldots, d$, are Lipschitz-continuous with constant $L_f = 2C_1$. ¿From (4), a) and (36) we see that

$$H(0, 0, \boldsymbol{n}) = 0 \quad \forall \, \boldsymbol{n} \in \boldsymbol{S}_1. \quad (38)$$

By (30), (34) and (36),

$$b_h(u, v) = \tilde{b}_h(u, v) \quad \forall u \in H^1(\Omega) \cap L^\infty(\Omega), \; \forall v \in H^1(\Omega, \mathcal{T}_h). \quad (39)$$

Now we define the space of discontinuous piecewise polynomial functions

$$S_h = S^{p,-1}(\Omega, \mathcal{T}_h) = \{v; v|_K \in P^p(K) \; \forall \, K \in \mathcal{T}_h\}, \quad (40)$$

where $P^p(K)$ denotes the space of all polynomials on K of degree $\leq p$, where the integer $p \geq 1$ is a given degree of approximation.

Using (33) and (39), we find that the exact solution u with property (13) satisfies the identity

$$\left(\frac{\partial u}{\partial t}(t), v_h \right) + a_h(u(t), v_h) + b_h(u(t), v_h) + \varepsilon J_h^\sigma(u(t), v_h) = \ell_h(v_h) \, (t) \quad (41)$$

for all $v_h \in S_h$ and all $t \in (0, T)$.

Now we are ready to introduce the discrete problem. To this end, we consider a partition $0 = t_0 < t_1 < \ldots$ of the time interval $[0, T]$, set $\tau_k = t_{k+1} - t_k$ for $k = 0, 1, \ldots$, the exact solution $u(t_k)$ will be approximated by an element $u^k \in S_h$, as test functions v we shall use functions $v_h \in S_h$ and the time derivative in (33) will be approximated by the backward difference. In order to obtain a stable and efficient scheme, the forms a_h, J_h^σ and ℓ_h will be treated implicitly, whereas the nonlinear form b_h will be treated explicitly. In this way we arrive at the following method.

Definition 2 *We define the* approximate solution *of problem (11) a) – c) as functions* u_h^k, $t_k \in [0, T]$, *satisfying the conditions*

$$
\begin{aligned}
&a)\ u_h^{k+1} \in S_h, && (42)\\
&b)\ \left(\frac{u_h^{k+1} - u_h^k}{\tau_k}, v_h\right) + a_h(u_h^{k+1}, v_h) + b_h(u_h^k, v_h)\\
&\qquad + \varepsilon J_h^\sigma(u_h^{k+1}, v_h) = \ell_h(v_h)\,(t_{k+1}) \quad \forall\, v_h \in S_h,\ \forall\, t_{k+1} \in (0, T],\\
&c)\ (u_h^0, v_h) = (u^0, v_h) \quad \forall\, v_h \in S_h.
\end{aligned}
$$

The function u_h^k *is called the approximate solution at time* t_k*.*

The approximate solution u_h^0 at $t = 0$ given by (42), c) is the L^2-projection of the function u^0 from the initial condition (3). If we define the operator $\Pi^{L^2} : L^2(\Omega) \to S_h$ of the $L^2(\Omega)$-projection on S_h, i.e. for $v \in L^2(\Omega)$ we assume that $\Pi^{L^2} v \in S_h$ and

$$\left(\Pi^{L^2} v, \varphi\right) = (v, \varphi) \quad \forall \varphi \in S_h, \tag{43}$$

then

$$u_h^0 = \Pi^{L^2} u^0. \tag{44}$$

It is obvious that $(\Pi^{L^2} v)|_K \in P^p(K)$ and

$$(\Pi^{L^2} v, \varphi)_{L^2(K)} = (v, \varphi)_{L^2(K)} \quad \forall \varphi \in P^p(K),\ \forall K \in \mathcal{T}_h,\ \forall v \in L^2(\Omega). \tag{45}$$

The discrete problem (42), a) – c) is equivalent to a system of linear algebraic equations for each $t_{k+1} \in (0, T]$, which can be solved by a suitable solver. In what follows we shall be concerned with the analysis of method (42), a) – c).

Lemma 1 *The discrete problem (42) a) – c) has a unique solution.*

Proof Problem (42) a) – c) can be rewritten in the following way. Given $u_h^k \in S_h$ and $\tau_k > 0$ for $t_{k+1} \in (0, T]$, we seek $u_h^{k+1} \in S_h$ such that

$$\boldsymbol{A}_h^k(u_h^{k+1}, v_h) = \boldsymbol{f}_h^k(v_h) \qquad \forall v_h \in S_h, \tag{46}$$

where

$$\begin{aligned}\boldsymbol{A}_h^k(u_h^{k+1}, v_h) &= (u_h^{k+1}, v_h) + \tau_k \left(a_h(u_h^{k+1}, v_h) + \varepsilon J_h^\sigma(u_h^{k+1}, v_h) \right), \\ \boldsymbol{f}_h^k(v_h) &= (u_h^k, v_h) + \tau_k \left(\ell_h(v_h)\,(t_{k+1}) - b_h(u_h^k, v_h) \right). \end{aligned} \tag{47}$$

Using the definition of forms (29) – (32) and (34), it is easy to observe that $\boldsymbol{A}_h^k$ is a bilinear form on the finite dimensional space S_h and $\boldsymbol{f}_h^k(v_h)$ is a linear functional on S_h. Moreover, the form $\boldsymbol{A}_h^k$ is coercive, since

$$\begin{aligned}\boldsymbol{A}_h^k(v_h, v_h) &= (v_h, v_h) + \tau_k \left(a_h(v_h, v_h) + \varepsilon J_h^\sigma(v_h, v_h) \right) \\ &\geq \|v_h\|_{L^2(\Omega)}^2 \quad \forall v_h \in S_h. \end{aligned} \tag{48}$$

Hence, equation (46) has a unique solution $u_h^{k+1} \in S_h$. $\square$

4 Some auxiliary results

In this section we summarize some important results and properties which have been proven in [14] and [15].

4.1 Geometry of the mesh

Let us consider a system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, of partitions of the domain Ω, i.e. $\mathcal{T}_h = \{K_i\}_{i\in I_h}$, $I_h \subset Z^+$. For the sake of simplicity, we shall write I instead of I_h $(h \in (0, h_0))$ and the dependence of index sets $I, I_b, s(i), \gamma(i)$ and $S(i)$ on h will not be emphasized by the notation.

In what follows, by the symbol $C_i, i = 1, 2, \ldots$, we shall denote constants always independent of h and τ.

Let us assume that the system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ has the following properties:

(A1) Each element $K \in \mathcal{T}_h$, $h \in (0, h_0)$, is a *star-shaped* domain with respect to at least one point $x_K = (x_{K1}, \ldots, x_{Kd}) \in K^\circ$, where K° is the interior of K. We assume:
   i) There exists a constant $\kappa > 0$ independent of K and h such that

$$\frac{\max_{x\in\partial K} |x - x_K|}{\min_{x\in\partial K} |x - x_K|} \leq \kappa \quad \forall K \in \mathcal{T}_h, \ \forall h \in (0, h_0). \tag{49}$$

ii) The element K can be divided into a finite number of closed simplexes:

$$K = \bigcup_{S \in \boldsymbol{S}(K)} S. \tag{50}$$

There exists a positive constant C_2 independent of K, S and h such that

$$\frac{h_S}{\rho_S} \le C_2 \quad \forall S \in \boldsymbol{S}(K) \qquad \text{(shape regularity)}, \tag{51}$$

where h_S is the diameter of S, ρ_S is the radius of the largest d-dimensional ball inscribed into S and, moreover,

$$1 \le \frac{h_K}{h_S} \le \tilde{\kappa} < \infty \quad \forall S \in \boldsymbol{S}(K), \tag{52}$$

where $\tilde{\kappa}$ is a constant independent of K, S and h.

(A2) There exists a constant $C_3 > 0$ such that

$$h_{K_i} \le C_3\, d(\Gamma_{ij}), \quad \forall i \in I,\ j \in S(i),\ h \in (0, h_0). \tag{53}$$

Let us note that these properties can be verified, e.g. in the case of dual finite volumes constructed over a regular simplicial mesh.

4.2 Some important inequalities and estimates

Under the above assumptions, the following results can be established.

Lemma 2 (Multiplicative trace inequality) *There exists a constant $C_4 > 0$ independent of v, h and K such that*

$$\|v\|^2_{L^2(\partial K)} \le C_4 \left(\|v\|_{L^2(K)}\, |v|_{H^1(K)} + h_K^{-1} \|v\|^2_{L^2(K)} \right), \tag{54}$$
$$\forall K \in \mathcal{T}_h,\ v \in H^1(K),\ h \in (0, h_0).$$

Proof See [15]. $\square$

Lemma 3 (Inverse inequality) *There exists a constant $C_5 > 0$ independent of v, h and K such that*

$$|v|_{H^1(K)} \le C_5 h_K^{-1} \|v\|_{L^2(K)} \quad \forall v \in P^p(K),\ K \in \mathcal{T}_h. \tag{55}$$

Proof See [15]. $\square$

Lemma 4 (Approximation properties of S_h) *There exist a constants $C_6 > 0$ independent of v and h and a linear mapping $\Pi : H^1(K) \to P^p(K),\ p \geq 0$ such that*

$$\begin{aligned} \|\Pi v - v\|_{L^2(K)} &\leq C_6 h_K^{p+1} |v|_{H^{p+1}(K)}, \\ |\Pi v - v|_{H^1(K)} &\leq C_6 h_K^{p} |v|_{H^{p+1}(K)}, \end{aligned} \tag{56}$$

for all $v \in H^{p+1}(K),\ K \in \mathcal{T}_h$ and $h \in (0, h_0)$.

Proof If $\mathcal{T}_h$ is a simplicial mesh, then standard results from the finite element method can be employed (see, e. g. [8] or [7]). In our case, when general nonconvex elements are used, the approximation properties (56) are derived under assumptions (A1), (A2) in [40]. □

Remark 1 The operator Π is not the L^2- projector Π^{L^2} on S_h introduced in (43). From the construction of Π in [40] it follows that $\Pi v = v$ for $v \in S_h$. Moreover,

$$\|v - \Pi^{L^2} v\|_{L^2(K)} \leq \|v - \Pi v\|_{L^2(K)} \leq C_6 h^{p+1} |v|_{H^{p+1}(K)}, \qquad \forall v \in H^p(K),\ \forall K \in \mathcal{T}_h, \tag{57}$$

as follows from (56).

4.3 Properties of the form b_h

Let assumptions (4), a), Assumptions (H) and Assumptions (A1), (A2) be satisfied. Then the form b_h has the following properties.

Lemma 5 *There exist constants $C_7 > 0$, $C_8 > 0$ and $C_9 > 0$ independent of $h \in (0, h_0)$ such that*

$$\begin{aligned} &|b_h(u, v) - b_h(\bar{u}, v)| \\ &\leq C_7 \left(J_h^\sigma(v, v)^{1/2} + |v|_{H^1(\Omega, \mathcal{T}_h)} \right) \\ &\quad \times \left(\|u - \bar{u}\|_{L^2(\Omega)} + \left(\sum_{i \in I} h_{K_i} \|u - \bar{u}\|^2_{L^2(\partial K_i)} \right)^{1/2} \right), \\ &\qquad u,\ \bar{u},\ v \in H^1(\Omega, \mathcal{T}_h), \end{aligned} \tag{58}$$

$$\begin{aligned} &|b_h(u_h, v_h) - b_h(\bar{u}_h, v_h)| \\ &\leq C_8 \left(J_h^\sigma(v_h, v_h)^{1/2} + |v_h|_{H^1(\Omega, \mathcal{T}_h)} \right) \|u_h - \bar{u}_h\|_{L^2(\Omega)}, \\ &\qquad u_h,\ \bar{u}_h,\ v_h \in S_h, \end{aligned} \tag{59}$$

$$\begin{aligned} &|b_h(u, v_h) - b_h(\Pi u, v_h)| \\ &\leq C_9 h^{p+1} \left(J_h^\sigma(v_h, v_h)^{1/2} + |v_h|_{H^1(\Omega, \mathcal{T}_h)} \right) |u|_{H^{p+1}(\Omega)}, \\ &\qquad u \in H^{p+1}(\Omega),\ v_h \in S_h, \end{aligned} \tag{60}$$

where Πu is the S_h-interpolant of u from Lemma 4.

Proof In [15], Lemma 5 gives (58) and (59) and Lemma 8 proves (60). □

5 Error estimates

Our goal is now to analyze the error estimates of the approximate solution $u_h^k,\ k = 0, 1, \ldots,$ obtained by the method (42) under the assumptions that the exact solution u satisfies (13). For simplicity, we consider a uniform partition $t_k = k\tau,\ k = 0, 1, \ldots, r,$ of the time interval $[0, T]$ with time step $\tau = T/r$, where $r > 1$ is an integer.

Let Πu^k be the S_h-interpolation of $u^k = u(t_k)$ $(k = 0, \ldots, r)$ from Lemma 4. We set

$$\xi^k = u_h^k - \Pi u^k \in S_h, \quad \eta^k = \Pi u^k - u^k \in H^{p+1}(\Omega, \mathcal{T}_h). \tag{61}$$

Then the error $e_h^k = u_h^k - u^k$ can be expressed as

$$e_h^k = \xi^k + \eta^k, \quad k = 0, \ldots, r. \tag{62}$$

Setting $v_h := \xi^{k+1}$ in (42), b), we get

$$\begin{aligned} &\left(u_h^{k+1} - u_h^k, \xi^{k+1}\right) + \tau \left(a_h(u_h^{k+1}, \xi^{k+1}) + b_h(u_h^k, \xi^{k+1})\right. \\ &\quad \left. + \varepsilon J_h^\sigma(u_h^{k+1}, \xi^{k+1}) - \ell_h(\xi^{k+1})\,(t_{k+1})\right) = 0, \quad t_k, t_{k+1} \in [0, T]. \end{aligned} \tag{63}$$

Moreover, setting $t := t_{k+1}$ and $v_h := \xi^{k+1}$ in (41), we obtain

$$\begin{aligned} &\left(u'(t_{k+1}), \xi^{k+1}\right) + a_h(u^{k+1}, \xi^{k+1}) + b_h(u^{k+1}, \xi^{k+1}) \\ &\quad + \varepsilon J_h^\sigma(u^{k+1}, \xi^{k+1}) - \ell_h(\xi^{k+1})\,(t_{k+1}) = 0, \quad t_k, t_{k+1} \in [0, T], \end{aligned} \tag{64}$$

where $u'(t_k) = \partial u/\partial t\,(t_k),\ t_k \in [0, T]$.

Multiplying (64) by τ, subtracting from (63) and using the linearity of the forms a_h and J_h^σ, we get

$$\begin{aligned} &\left(u^{k+1} - u^k, \xi^{k+1}\right) - \tau \left(u'(t_{k+1}), \xi^{k+1}\right) \\ &+ \tau \left(a_h(u_h^{k+1} - u^{k+1}, \xi^{k+1}) + b_h(u_h^k, \xi^{k+1}) - b_h(u^{k+1}, \xi^{k+1})\right. \\ &\quad \left. + \varepsilon J_h^\sigma(u_h^{k+1} - u^{k+1}, \xi^{k+1})\right) = 0, \quad k = 0, \ldots, r-1. \end{aligned} \tag{65}$$

By relations (61) and (62), from (65) we have

$$\begin{aligned}
&\left(\xi^{k+1}-\xi^k,\xi^{k+1}\right)+\tau\left(a_h(\xi^{k+1},\xi^{k+1})+\varepsilon J_h^\sigma(\xi^{k+1},\xi^{k+1})\right) \quad (66)\\
&=\tau(u'(t_{k+1}),\xi^{k+1})-(u^{k+1}-u^k,\xi^{k+1})-(\eta^{k+1}-\eta^k,\xi^{k+1})\\
&\quad+\tau\left(b_h(u^{k+1},\xi^{k+1})-b_h(u_h^k,\xi^{k+1})-a_h(\eta^{k+1},\xi^{k+1})\right.\\
&\qquad\left.-\varepsilon J_h^\sigma(\eta^{k+1},\xi^{k+1})\right).
\end{aligned}$$

In what follows, we estimate the individual terms on the right-hand side of (66).

The Cauchy inequality implies that

$$J_h^\sigma(\eta^k,\xi^k)\le (J_h^\sigma(\eta^k,\eta^k))^{1/2}\,(J_h^\sigma(\xi^k,\xi^k))^{1/2},\quad k=0,\ldots,r. \quad (67)$$

Lemma 6 *Under assumptions (13), for $t_k,t_{k+1}\in[0,T]$ we have*

$$\left|(u_h^{k+1}-u_h^k,\xi^{k+1})-\tau(u'(t_{k+1}),\xi^{k+1})\right|\le C_{10}\tau^2\|\xi^{k+1}\|_{L^2(\Omega)}, \quad (68)$$

$$\|u^{k+1}-u^k\|_{L^2(\Omega)}\le C_{11}\tau, \quad (69)$$

$$|u^{k+1}-u^k|_{H^1(\Omega)}\le C_{12}\tau, \quad (70)$$

$$|u^{k+1}-u^k|_{H^{p+1}(\Omega)}\le C_{13}\tau, \quad (71)$$

with C_{10}, C_{11}, C_{12} and C_{13} depending on u, but independent of k and τ.

Proof i) By [12], Lemma 8, we have (68) with
$C_{10}=\|u''\|_{L^\infty(0,T;L^2(\Omega))}$.

ii) Since $u'\in L^\infty(0,T;H^{p+1}(\Omega))\subset L^\infty(0,T;L^2(\Omega))$, we can write

$$\|u^{k+1}-u^k\|_{L^2(\Omega)}=\left\|\int_{t_k}^{t_{k+1}}u'(t)\,\mathrm{d}t\right\|_{L^2(\Omega)}\le\tau\|u'\|_{L^\infty(0,T;L^2(\Omega))}, \quad (72)$$

which proves (69) with $C_{11}=\|u'\|_{L^\infty(0,T;L^2(\Omega))}$.

iii) Since $u'\in L^\infty(0,T;H^{p+1}(\Omega))\subset L^\infty(0,T;H^1(\Omega))$ and

$$\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial x_l}\right)=\frac{\partial}{\partial x_l}\left(\frac{\partial u}{\partial t}\right),\qquad l=1,\ldots,d, \quad (73)$$

in the sense of distributions, we obtain

$$\begin{aligned}
&|u^{k+1}-u^k|_{H^1(\Omega)}=\|\nabla u^{k+1}-\nabla u^k\|_{L^2(\Omega)} \quad (74)\\
&=\left\|\int_{t_k}^{t_{k+1}}\frac{\partial}{\partial t}\nabla u(t)\,\mathrm{d}t\right\|_{L^2(\Omega)}\\
&\le\int_{t_k}^{t_{k+1}}\left\|\frac{\partial}{\partial t}\nabla u(t)\right\|_{L^2(\Omega)}\mathrm{d}t=\int_{t_k}^{t_{k+1}}\|\nabla u'(t)\|_{L^2(\Omega)}\,\mathrm{d}t\\
&=\int_{t_k}^{t_{k+1}}|u'(t)|_{H^1(\Omega)}\,\mathrm{d}t\le\tau\|u'\|_{L^\infty(0,T;H^1(\Omega))},
\end{aligned}$$

which is (70) with $C_{12} = \|u'\|_{L^\infty(0,T;H^1(\Omega))}$.

iv) Using a similar argumentation as in (74), we derive (71) with $C_{13} = \|u'\|_{L^\infty(0,T;H^{p+1}(\Omega))}$.
□

Lemma 7 *Under assumptions (13), for $t_k, t_{k+1} \in [0,T]$ we have*

$$|(\eta^{k+1} - \eta^k, \xi^{k+1})| \le C_{14} \tau h^{p+1} \|\xi^{k+1}\|_{L^2(\Omega)}, \tag{75}$$

with $C_{14} = C_{14}(u)$.

Proof The Cauchy inequality, relations (61), (56) and (71) imply that

$$\begin{aligned} |(\eta^{k+1} - \eta^k, \xi^{k+1})| &\le \|\eta^{k+1} - \eta^k\|_{L^2(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &= \|\Pi(u^{k+1} - u^k) - (u^{k+1} - u^k)\|_{L^2(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &\le C_6 h^{p+1} |u^{k+1} - u^k|_{H^{p+1}(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &\le C_6 C_{13} \tau h^{p+1} \|\xi^{k+1}\|_{L^2(\Omega)}, \end{aligned} \tag{76}$$

which proves the lemma with $C_{14} := C_6 C_{13}$. □

Lemma 8 *There exist constants $C_{15} > 0$ and $C_{16} > 0$ independent of u, h, k, ξ and ε such that*

$$|a_h(\eta^k, \xi^k)| \le C_{15} \varepsilon h^p |u^k|_{H^p(\Omega)} \left(J_h^\sigma(\xi^k, \xi^k)^{1/2} + |\xi^k|_{H^1(\Omega, \mathcal{T}_h)} \right), \tag{77}$$

$$J_h^\sigma(\eta^k, \eta^k) \le C_{16}\, h^{2p} |u^k|^2_{H^p(\Omega)}, \quad h \in (0, h_0),\ t_k \in [0,T]. \tag{78}$$

Proof See [15], Lemma 9. □

Lemma 9 *For $h \in (0, h_0)$, $t_k, t_{k+1} \in [0,T]$ we have*

$$\begin{aligned} &\left| b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \right| \\ &\le C_{17} \left(J_h^\sigma(\xi^{k+1}, \xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega, \mathcal{T}_h)} \right) \\ &\quad \times \left(\|\xi^k\|_{L^2(\Omega)} + h^{p+1} + \tau \right), \end{aligned} \tag{79}$$

where $C_{17} = C_{17}(u)$ is independent of h,τ, k, ξ.

Proof We can write

$$\begin{aligned} &b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \\ &\quad = \ b_h(u^{k+1}, \xi^{k+1}) - b_h(u^k, \xi^{k+1}) \quad (=: \Psi_1) \\ &\quad + b_h(u^k, \xi^{k+1}) - b_h(\Pi u^k, \xi^{k+1}) \quad (=: \Psi_2) \\ &\quad + b_h(\Pi u^k, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \quad (=: \Psi_3). \end{aligned} \tag{80}$$

We estimate the individual terms in (80). In virtue of (58),

$$|\Psi_1| \le C_7 \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)} \right) \times \left(\|u^{k+1}-u^k\|_{L^2(\Omega)} + \left(\sum_{i\in I} h_{K_i} \|u^{k+1}-u^k\|^2_{L^2(\partial K_i)} \right)^{1/2} \right). \tag{81}$$

Using (54), (69) and (70) we find that

$$\begin{aligned} &\sum_{i\in I} h_{K_i} \|u^{k+1}-u^k\|^2_{L^2(\partial K_i)} \\ &\le C_4 \sum_{i\in I} \Big(h_{K_i} \|u^{k+1}-u^k\|_{L^2(K_i)} |u^{k+1}-u^k|_{H^1(K_i)} \\ &\qquad + \|u^{k+1}-u^k\|^2_{L^2(K_i)} \Big) \le C_{18}\tau^2, \end{aligned} \tag{82}$$

where $C_{18} := C_4(C_{11}C_{12}h_0 + C_{11}^2)$. Then (81), (82) and (69) give

$$|\Psi_1| \le C_7 \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)} \right) \tau(\sqrt{C_{18}} + C_{11}). \tag{83}$$

Moreover, due to (13), we can set

$$C_{19} = \|u\|_{L^\infty(0,T;H^{p+1}(\Omega))}. \tag{84}$$

From (60) and (59) we deduce that

$$\begin{aligned} |\Psi_2| &\le C_9 h^{p+1} \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)} \right) |u^k|_{H^{p+1}(\Omega)}, \\ |\Psi_3| &\le C_8 \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)} \right) \|\Pi u^k - u_h^k\|_{L^2(\Omega)}. \end{aligned} \tag{85}$$

By (80), (61), (83), (85) and (84),

$$\begin{aligned} &\left| b_h(u^{k+1},\xi^{k+1}) - b_h(u_h^k,\xi^{k+1}) \right| \\ &\le C_{17} \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)} \right) \\ &\qquad \times \left(\|\xi^k\|_{L^2(\Omega)} + h^{p+1} + \tau) \right), \end{aligned} \tag{86}$$

with $C_{17} := \max(C_8,\ C_9C_{19},\ C_7(C_{11} + \sqrt{C_{18}})$, which proves the lemma. □

Now we shall formulate the *main result*.

Theorem 1 *Let assumptions (4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (11) satisfying (13). Let $t_k = k\tau,\ k = 0, 1, \dots, r,\ \tau = T/r$, be a time partition of $[0, T]$ and let $u_h^k,\ k = 0, \dots, r$, be the approximate solution defined by (42) and let $\tau \le 1/2$. Let us set*

$$
\begin{aligned}
e &= \{e_h^k\}_{k=0}^r = \{u_h^k - u^k\}_{k=0}^r, \qquad (87)\\
\|e\|^2_{h,\tau,L^\infty(L^2)} &= \max_{k=0,\dots,r} \|e_h^k\|^2_{L^2(\Omega)}\\
\|e\|^2_{h,\tau,L^2(H^1)} &= \varepsilon\tau \sum_{k=0}^{r} \left(|e_h^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^k, e_h^k) \right)\\
\|e\|^2_{h,\tau,L^\infty(H^1)} &= \varepsilon \max_{k=0,\dots,r} \left(|e_h^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^k, e_h^k) \right)
\end{aligned}
$$

Then there exist constants $\tilde{C} = O\left(\exp(2T(1 + C_{20}/\varepsilon))\right)$ and $\hat{C} = O\left(\exp(2T(1 + C_{20}/\varepsilon))\right)$ such that

$$
\|e\|^2_{h,\tau,L^\infty(L^2)} \le \tilde{C} \left(h^{2p} \left(\varepsilon + h^2 + h^2/\varepsilon \right) + \tau^2 \left(1 + 1/\varepsilon\right) \right), \quad (88)
$$

$$
\|e\|^2_{h,\tau,L^2(H^1)} \le \hat{C} \left(h^{2p} \left(\varepsilon + h^2 + h^2/\varepsilon \right) + \tau^2 \left(1 + 1/\varepsilon\right) \right), \quad (89)
$$

where $C_{20} = 8C_{17}^2$.

Moreover, provided $h \le \tau$, there exists a constant $\bar{C} = O\left(\exp(2T(1 + C_{20}/\varepsilon))\right)$ such that

$$
\begin{aligned}
\|e\|^2_{h,\tau,L^\infty(H^1)} \le \bar{C} &\left(h^{2p-1} \left(1 + \varepsilon + h + h/\varepsilon + h^2 + h^2/\varepsilon + h^2/\varepsilon^2 \right) \right.\\
&\left. + \tau \left(1 + 1/\varepsilon + 1/\varepsilon^2 \right) \right) \qquad (90)
\end{aligned}
$$

Proof As in (61), we set $\xi^k = u_h^k - \Pi u^k \in S_h$, $\eta^k = \Pi u^k - u^k$, $k = 0, \dots, r$. Then (62) holds: $e_h^k = u_h^k - u^k = \xi^k + \eta^k$. From (66) and the relations

$$
a_h(\xi^{k+1}, \xi^{k+1}) = \varepsilon |\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)} \qquad (91)
$$

and

$$
2(\xi^{k+1} - \xi^k, \xi^{k+1}) = (\|\xi^{k+1}\|^2_{L^2(\Omega)} - \|\xi^k\|^2_{L^2(\Omega)} + \|\xi^{k+1} - \xi^k\|^2_{L^2(\Omega)}), \qquad (92)
$$

for $k = 0, \dots, r-1$ we get

$$
\begin{aligned}
&\|\xi^{k+1}\|^2_{L^2(\Omega)} - \|\xi^k\|^2_{L^2(\Omega)} + \|\xi^{k+1} - \xi^k\|^2_{L^2(\Omega)} \qquad (93)\\
&+2\tau \left(\varepsilon |\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)} + \varepsilon J_h^\sigma(\xi^{k+1}, \xi^{k+1}) \right)\\
= {} & 2 \left(\tau(u'(t_{k+1}), \xi^{k+1}) - (u^{k+1} - u^k, \xi^{k+1}) - \left(\eta^{k+1} - \eta^k, \xi^{k+1} \right) \right)\\
&+2\tau \left(b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) - a_h(\eta^{k+1}, \xi^{k+1}) \right.\\
&\left. -\varepsilon J_h^\sigma(\eta^{k+1}, \xi^{k+1}) \right) =: \mathrm{RHS}.
\end{aligned}
$$

With the aid of Lemmas 6 – 9 , inequality (67), notation (84) and $C_{21} = C_{19}(C_{15} + \sqrt{C_{16}})$ we estimate the right-hand-side RHS of (93):

$$|\mathrm{RHS}| \leq 2\left(C_{10}\tau^2 + C_{14}\tau h^{p+1}\right)\|\xi^{k+1}\|_{L^2(\Omega)} \tag{94}$$
$$+2\tau\left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,\mathcal{T}_h)}\right)$$
$$\times\left(C_{21}\varepsilon h^p + C_{17}\left(h^{p+1} + \tau + \|\xi^k\|_{L^2(\Omega)}\right)\right).$$

From Young's inequality, under the notation $C_{20} = 8C_{17}^2$ and

$$q(\varepsilon, h, \tau) = 2h^{2p}\left(C_{14}^2 h^2 + 4\varepsilon C_{21}^2 + \frac{4}{\varepsilon}C_{17}^2 h^2\right) + \tau^2\left(2C_{10}^2 + \frac{C_{20}}{\varepsilon}\right), \tag{95}$$

it follows that

$$\|\xi^{k+1}\|^2_{L^2(\Omega)} - \|\xi^k\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{k+1},\xi^{k+1})\right)$$
$$\leq \tau\|\xi^{k+1}\|^2_{L^2(\Omega)} + \tau(1 + C_{20}/\varepsilon)\|\xi^k\|^2_{L^2(\Omega)} + \tau q(\varepsilon, h, \tau). \tag{96}$$

Hence,

$$(1-\tau)\|\xi^{k+1}\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{k+1},\xi^{k+1})\right) \tag{97}$$
$$\leq (1 + \tau C_{20}/\varepsilon)\|\xi^k\|^2_{L^2(\Omega)} + \tau q(\varepsilon, h, \tau).$$

Moreover, the following inequalities are valid:

$$\|\xi^k + \eta^k\|^2_{L^2(\Omega)} \leq 2\left(\|\xi^k\|^2_{L^2(\Omega)} + \|\eta^k\|^2_{L^2(\Omega)}\right), \tag{98}$$
$$|\xi^k + \eta^k|^2_{H^1(\Omega),\mathcal{T}_h)} \leq 2\left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)} + |\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)}\right),$$
$$J_h^\sigma(\xi^k + \eta^k, \xi^k + \eta^k) \leq 2(J_h^\sigma(\xi^k,\xi^k) + J_h^\sigma(\eta^k,\eta^k).$$

Now we prove the error estimates (88) – (90).

i) By (97),

$$\|\xi^{k+1}\|^2_{L^2(\Omega)} \leq \frac{1 + \tau C_{20}/\varepsilon}{1-\tau}\|\xi^k\|^2_{L^2(\Omega)} + \frac{\tau}{1-\tau}q(\varepsilon, h, \tau), \quad k = 0, \dots, r-1. \tag{99}$$

If we set

$$B = \frac{1 + \tau C_{20}/\varepsilon}{1-\tau}, \tag{100}$$

then we derive from (99) by induction that

$$\|\xi^k\|^2_{L^2(\Omega)} \leq B^k\|\xi^0\|^2_{L^2(\Omega)} + \frac{B^k - 1}{B-1}\frac{\tau\, q(\varepsilon, h, \tau)}{1-\tau}, \quad k = 0, \dots, r. \tag{101}$$

By (100),

$$\frac{\tau}{(B-1)(1-\tau)} = \frac{1}{1+C_{20}/\varepsilon} \le 1. \tag{102}$$

As $\tau \le 1/2$, then $1-\tau \ge 1/2$ and

$$B \le 1 + 2\tau(1 + C_{20}/\varepsilon) \le \exp(2\tau(1 + C_{20}/\varepsilon)). \tag{103}$$

From (101) – (103) we have

$$\|\xi^k\|^2_{L^2(\Omega)} \le \exp(2\tau k(1 + C_{20}/\varepsilon)) \left(\|\xi^0\|^2_{L^2(\Omega)} + q(\varepsilon, h, \tau) \right). \tag{104}$$

Further, (44), (42) c), (43) and the Cauchy inequality imply that for each $K \in \mathcal{T}_h$,

$$\begin{aligned} \|\xi^0\|^2_{L^2(K)} &= (\Pi^{L^2} u^0 - \Pi u^0, \xi^0)_{L^2(K)} = (u^0 - \Pi u^0, \xi^0)_{L^2(K)} \\ &\le \|u^0 - \Pi u^0\|_{L^2(K)} \|\xi^0\|_{L^2(K)}. \end{aligned} \tag{105}$$

Thus, by (14), (56) and (84),

$$\|\xi^0\|_{L^2(K)} \le \|u^0 - \Pi u^0\|_{L^2(K)} \le C_6 h_K^{p+1} |u^0|_{H^{p+1}(K)}, \tag{106}$$

and

$$\|\xi^0\|^2_{L^2(\Omega)} = \sum_{i \in I} \|\xi^0\|^2_{L^2(K_i)} \le C_6^2 h^{2(p+1)} |u^0|^2_{H^{p+1}(\Omega, \mathcal{T}_h)} \le C_{22} h^{2(p+1)}, \tag{107}$$

where $C_{22} = C_6^2 C_{19}^2$. ¿From (104) and (107) we get

$$\|\xi^k\|^2_{L^2(\Omega)} \le \exp(2T(1 + C_{20}/\varepsilon)) \left(C_{22} h^{2(p+1)} + q(\varepsilon, h, \tau) \right), \quad k = 0, \dots, r. \tag{108}$$

Further, by (61), (56) and (84),

$$\|\eta^k\|^2_{L^2(\Omega)} \le C_6^2 h^{2(p+1)} |u^k|^2_{H^{p+1}(\Omega)} \le C_6^2 C_{19}^2 h^{2(p+1)}. \tag{109}$$

Now, using (87), (98), (108) and (109), we find that

$$\begin{aligned} \|e\|^2_{h,\tau,L^\infty(L^2)} &\le 2 \max_{k=0,\dots,r} \left(\|\xi^k\|^2_{L^2(\Omega)} + \|\eta^k\|^2_{L^2(\Omega)} \right) \\ &\le 2 \exp(2T(1 + C_{20}/\varepsilon)) \left(q(\varepsilon, h, \tau) + C_{23} h^{2(p+1)} \right), \end{aligned} \tag{110}$$

with $C_{23} = C_{22} + C_6^2 C_{19}^2$, which implies estimate (88).

ii) Now let as derive (89). Summing (96) over $k = 0, \dots, r-1$, we get

$$\begin{aligned} &\|\xi^r\|^2_{L^2(\Omega)} + \tau\varepsilon \sum_{k=0}^{r-1} \Big(|\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{k+1},\xi^{k+1}) \Big) \qquad (111) \\ &\le \tau(1 + C_{20}/\varepsilon) \sum_{k=0}^{r-1} \Big(\|\xi^{k+1}\|^2_{L^2(\Omega)} + \|\xi^k\|^2_{L^2(\Omega)} \Big) \\ &\quad + Tq(\varepsilon,h,\tau) + \|\xi^0\|^2_{L^2(\Omega)}. \end{aligned}$$

This, (107) and (108) imply that

$$\begin{aligned} &\tau\varepsilon \sum_{k=1}^{r} \Big(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^k,\xi^k) \Big) \qquad (112) \\ &\le 2T(1 + C_{20}/\varepsilon) \exp(2T(1 + C_{20}/\varepsilon)) \Big(C_{22} h^{2(p+1)} + 2q(\varepsilon,h,\tau) \Big) \\ &\quad + C_{22} h^{2(p+1)}. \end{aligned}$$

Moreover, as $\xi^0 \in S_h$, we have from (55) and (106) the estimate

$$|\xi^0|_{H^1(K_i)} \le C_5 h_{K_i}^{-1} \|\xi^0\|_{L^2(K_i)} \le C_5 C_6 h_{K_i}^p |u^0|_{H^{p+1}(K_i)}, \quad i \in I, \qquad (113)$$

and, by (84),

$$\begin{aligned} |\xi^0|^2_{H^1(\Omega,\mathcal{T}_h)} &= \sum_{i\in I} |\xi^0|^2_{H^1(K_i)} \qquad (114) \\ &\le C_5^2 C_6^2 h^{2p} |u^0|^2_{H^{p+1}(\Omega,\mathcal{T}_h)} \le C_5^2 C_6^2 C_{19}^2 h^{2p}. \end{aligned}$$

Furthermore, in virtue of (32), (28), (53), (54), (106) and (113),

$$\begin{aligned} J_h^\sigma(\xi^0,\xi^0) &\le 4 \sum_{i\in I} \sum_{j\in S(i)} \int_{\Gamma_{ij}} \frac{1}{d(\Gamma_{ij})} |\xi^0|_{\Gamma_{ij}}|^2 \, \mathrm{d}S \qquad (115) \\ &\le 4C_3 \sum_{i\in I} \frac{1}{h_{K_i}} \int_{\partial K_i} |\xi^0|^2 \, \mathrm{d}S \\ &\le 4C_3\, C_4 \sum_{i\in I} \frac{1}{h_{K_i}} \Big(\|\xi^0\|_{L^2(K_i)} |\xi^0|_{H^1(K_i)} + h_{K_i}^{-1} \|\xi^0\|^2_{L^2(K_i)} \Big) \\ &\le 8C_3\, C_4\, C_5^2 \sum_{i\in I} h_{K_i}^{2p} |u^0|^2_{H^{p+1}(K_i)} \\ &\le 8C_3\, C_4\, C_5^2\, h^{2p} |u^0|^2_{H^{p+1}(\Omega,\mathcal{T}_h)} \\ &\le 8C_3\, C_4\, C_5^2\, C_{19}^2\, h^{2p}. \end{aligned}$$

Then (112), (114) and (115) give

$$\tau\varepsilon\sum_{k=0}^{r}\left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^k,\xi^k)\right) \tag{116}$$
$$\le 2T(1+C_{20}/\varepsilon)\exp(2T(1+C_{20}/\varepsilon))\left(C_{22}h^{2(p+1)}+2q(\varepsilon,h,\tau)\right)$$
$$+h^{2p}\left(C_{22}h^2+\tau\varepsilon C_{24}\right),$$

where $C_{24}=C_5^2C_6^2C_{19}^2+8C_3\,C_4\,C_5^2\,C_{19}^2$.

Taking into account (61), (56), (78) and (84), we obtain

$$\tau\varepsilon\sum_{k=0}^{r}\left(|\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\eta^k,\eta^k)\right) \tag{117}$$
$$\le \tau\varepsilon\left(C_6^2h^{2p}\sum_{k=0}^{r}|u^k|^2_{H^{p+1}(\Omega,\mathcal{T}_h)}+C_{16}\,h^{2p}\sum_{k=0}^{r}|u^k|^2_{H^{p+1}(\Omega)}\right)$$
$$\le \varepsilon C_{25}h^{2p}(T+\tau),$$

where $C_{25}=C_{19}^2(C_6^2+C_{16})$. Finally, using (87), (61), (98), (116) and (117), we arrive at the estimate

$$\|e\|^2_{h,\tau,L^2(H^1)}\le 2\tau\varepsilon\sum_{k=0}^{r}\left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^k,\xi^k)\right. \tag{118}$$
$$\left.+|\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\eta^k,\eta^k)\right)$$
$$\le 4T(1+C_{20}/\varepsilon)\exp(2T(1+C_{20}/\varepsilon))$$
$$\times\left(2q(\varepsilon,h,\tau)+h^{2p}\left(C_{22}h^2+\varepsilon(C_{24}+C_{25})\right)\right).$$

Now, assertion (89) of the theorem follows from (95) and (118).

iii) Finally, let $h\le\tau$. As $\tau\le 1/2$, (97) implies that

$$\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{k+1},\xi^{k+1})\right) \tag{119}$$
$$\le\left(\frac{1}{\tau}+\frac{C_{20}}{\varepsilon}\right)\|\xi^k\|^2_{L^2(\Omega)}+q(\varepsilon,h,\tau).$$

Using (108), we obtain

$$\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{k+1},\xi^{k+1})\right) \tag{120}$$
$$\le\left(\frac{1}{\tau}+\frac{C_{20}}{\varepsilon}\right)\exp(2T(1+C_{20}/\varepsilon))\left(C_{22}h^{2(p+1)}+q(\varepsilon,h,\tau)\right)$$
$$+q(\varepsilon,h,\tau).$$

Hence, in virtue of (61), (56), (78) and (84),

$$\varepsilon\left(|\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^k,\eta^k)\right) \le \varepsilon C_{25} h^{2p}, \tag{121}$$

where $C_{25} = C_{19}^2(C_6^2 + C_{16})$. Now, according to (87), (61), (98), (120) and (121) we arrive at

$$\begin{aligned} \|e\|^2_{h,\tau,L^\infty(H^1)} \le & \max_{k=0,\dots,r} 2\varepsilon\left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^k,\xi^k)\right. \\ & \left. + |\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^k,\eta^k)\right) \\ \le & \, 2(\varepsilon + \tau C_{20}) \exp(2T(1 + C_{20}/\varepsilon)) \\ & \times \left(\frac{C_{22}h^{2(p+1)} + q(\varepsilon,h,\tau)}{\varepsilon\tau} + q(\varepsilon,h,\tau) + C_{25}h^{2p}\right). \end{aligned} \tag{122}$$

Finally, this, (95) and the assumption that $h \le \tau$ yield (90). $\square$

Remark 2 Estimate (88) implies that

$$\|u - u_h\|_{L^\infty(0,T;L^2(\Omega))} = O(h^p + \tau) \quad \text{for } h \to 0+. \tag{123}$$

Comparing this result with the approximation property (56) implying that

$$\|u - \Pi u\|_{L^\infty(0,T;L^2(\Omega))} = O(h^{p+1}), \tag{124}$$

we see that the error estimate (88) is suboptimal with respect to h. There is a question, whether this estimate can be improved. Numerical experiments carried out in the next section indicate that the actual order of convergence in the $L^\infty(L^2)$-norm in the case of odd order of approximation (here $p = 1$ and $p = 3$) is better than the theoretically derived estimate. (This fact was also documented experimentally for stationary elliptic problems without convection in [32] and [24].)

It is clear that estimates (88) – (90) cannot be used for $\varepsilon \to 0+$, because they blow up exponentially with respect to $1/\varepsilon$. This is a consequence of the application of Young's inequality and Gronwall's lemma, necessary for overcoming the nonlinearity of the convective terms. This nonlinearity represents a serious obstacle for obtaining a uniform error estimate with respect to $\varepsilon \to 0+$, as in [22].

Estimate (90) in $L^\infty(H^1)$-norm has been obtained under the assumption that $h \le \tau$. This nonstandard "inverse stability condition" also appears in [37] and [19].

6 Numerical experiments

In this section we shall verify the theoretical error estimates, presented in the previous section, in the case of the 2D viscous Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x_1} + u\frac{\partial u}{\partial x_2} = \varepsilon \Delta u + g \quad \text{in} \quad \Omega \times (0,T), \tag{125}$$

where $\Omega = (0,1)^2, T = 1$, equipped with the boundary condition (2) and the initial condition (3). In the definition (34) of the form b_h we use the numerical flux

$$H(u_1, u_2, \boldsymbol{n}) = \begin{cases} \sum_{s=1}^{2} f_s(u_1) n_s, & \text{if } A > 0 \\ \sum_{s=1}^{2} f_s(u_2) n_s, & \text{if } A \le 0 \end{cases}, \tag{126}$$

where

$$A = \sum_{s=1}^{2} f_s'(\bar{u}) n_s, \quad \bar{u} = \frac{1}{2}(u_1 + u_2) \ \text{ and } \ \boldsymbol{n} = (n_1, n_2). \tag{127}$$

One can see that H satisfies conditions (36) and (37) and is Lipschitz-continuous on any bounded subset of $\mathbb{R}^2$. For boundary edges $\Gamma_{ij},\ j \in \gamma(i)$ we use more sophisticated boundary conditions in the definition of the convective form b_h (based on upwinding). The second term in (34) is replaced by

$$\sum_{j \in \gamma(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, \bar{u}|_{\Gamma_{ij}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}} \, \mathrm{d}S, \tag{128}$$

where

$$\bar{u}|_{\Gamma_{ij}} = \begin{cases} u|_{\Gamma_{ij}}, & \text{if } \sum_{s=1}^{2} f_s'(u|_{\Gamma_{ij}}) (\boldsymbol{n}_{ij})_s > 0, \\ u_D|_{\Gamma_{ij}}, & \text{otherwise.} \end{cases}, \tag{129}$$

where $(\boldsymbol{n}_{ij})_s$ is the s-th component of outer normal $\boldsymbol{n}_{ij}$ to $\partial\Omega$ on Γ_{ij} and $u_D|_{\Gamma_{ij}}$ is the restriction of boundary condition (2) on Γ_{ij}.

All computations were carried out with a constant time step $\tau = T/r$, where $r > 1$ is a chosen integer. Volume integrals and line integrals are evaluated with the aid of quadrature formulae exact for polynomials of degree ≤ 6 and ≤ 7, respectively.

The systems of linear algebraic equations were solved by the bi-conjugate gradient method with SSOR preconditioning with the aid of LASPACK package [39].

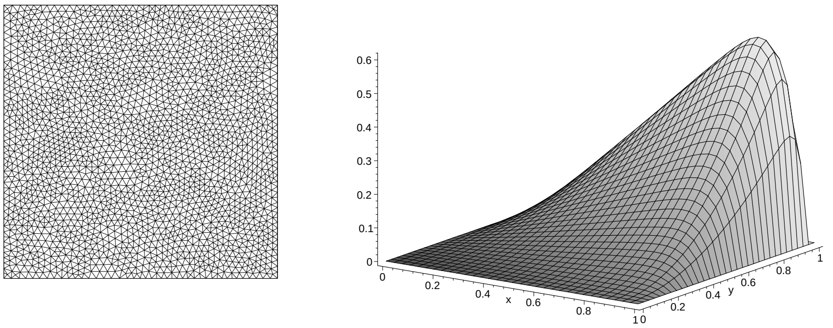


Fig. 2. Computational mesh $\mathcal{T}_{h_6}$ (left) and the exact solution (130) at $t = 1$ for $\varepsilon = 0.1$ (right)

6.1 Convergence with respect to τ

First, we verify numerically the convergence of the method in L^2-norm and H^1-seminorm with respect to the time step τ. In order to restrain the discretization errors with respect to h, we use a fine mesh with 4219 triangles (see Figure 2, left) and piecewise cubic approximation, i.e. $p = 3$ in (40).

We define the function g and the initial and boundary conditions in such a way that the exact solution has the form

$$u(x_1, x_2, t) = \frac{e^{10t} - 1}{e^{10} - 1} \hat{u}(x_1, x_2), \tag{130}$$

where

$$\begin{aligned} \hat{u}(x_1, x_2) = x_1 x_2^2 - x_2^2 \exp\left(\frac{2(x_1 - 1)}{\varepsilon}\right) \\ -x_1 \exp\left(\frac{3(x_2 - 1)}{\varepsilon}\right) + \exp\left(\frac{2x_1 + 3x_2 - 5}{\varepsilon}\right), \end{aligned} \tag{131}$$

and $\varepsilon = 0.1$. The solution u is equal to zero at $t = 0$ and converges exponentialy to $\hat{u}$ for $t \to 1$. The function $\hat{u}$ vanishes on the part of the boundary $\partial\Omega$, where $x_1 = 1$ or $x_2 = 1$ and $|\hat{u}| \leq \exp(-20)$ for $x_1 = 0$ or $x_2 = 0$. It contains two boundary layers along the right and top sides of the computational domain, see Figure 2 (right).

The computations were carried out with 6 different time steps τ_l, see Table 6.1. The computational error is evaluated at time $T = 1$ in $L^2(\Omega)$–norm and $H^1(\Omega, \mathcal{T}_h)$–seminorm:

$$\begin{aligned} e_\tau^0 &\equiv \|u_\tau(\cdot, T) - u(\cdot, T)\|_{L^2(\Omega)}, \\ e_\tau^1 &\equiv |u_\tau(\cdot, T) - u(\cdot, T)|_{H^1(\Omega, \mathcal{T}_h)}, \end{aligned} \tag{132}$$

Table 1. Time steps $\tau_l,\ l = 1,\dots,5$, computational errors in L^2–norm and H^1–seminorm, values of $\alpha_l^n,\ l = 2,\dots,5$ and $\bar{\alpha}^n, n = 0,1$

		L^2–norm		H^1–seminorm	
l	τ_l	e_τ^0	α_l^0	e_τ^1	α_l^1
1	4.830 E-02	3.221 E-02	–	2.051 E-01	–
2	2.415 E-02	1.667 E-02	0.950	1.075 E-01	0.932
3	1.208 E-02	8.837 E-03	0.916	5.830 E-02	0.882
4	6.038 E-03	4.436 E-03	0.995	2.941 E-02	0.987
5	3.109 E-03	2.248 E-03	0.981	1.506 E-02	0.965
6	1.509 E-03	1.129 E-03	0.993	7.753 E-03	0.958
global order $\bar{\alpha}^n$			0.967		0.946

where $u(\cdot,T)$ is the exact solution given by (130) – (131) at time T and $u_\tau(\cdot,T)$ is the numerical solution at time T obtained by scheme (42), a)–c) with time step τ. We suppose that the errors behave according to the formula

$$e_\tau^n \approx D_n \tau^{\alpha^n}, \quad n = 0,1, \tag{133}$$

where $D_n > 0,\ n = 0,1$, are constants independent of τ and $\alpha^n,\ n = 0,1$, are the *orders of accuracy* of the method in the norms considered. We define the *local experimental order of convergence* by

$$\alpha_l^n = \frac{\log\left(e_{\tau_l}^n / e_{\tau_{l-1}}^n\right)}{\log\left(\tau_l/\tau_{l-1}\right)}, \quad l = 2,\dots,5,\ n = 0,1. \tag{134}$$

The *global experimental order of convergence* $\bar{\alpha}_n,\ n = 0,1$, is obtained by the least squares method. Table 6.1 shows the errors in L^2-norm and H^1-seminorm, the values of α_l^n for $l = 2,\dots,6$ and $\bar{\alpha}^n$. The orders of convergence can be seen in Figure 3, where also expected optimal error slopes are indicated. Figure 4 shows the graphs of the numerical solutions along the line $x_2 = 0.5$ for various time steps and indicates the convergence to the exact solution.

6.2 Convergence with respect to h

Here we verify numerically the convergence of errors in L^2-norm and H^1-seminorm with respect to the mesh size h. In order to restrain the discretization errors with respect to τ, we use sufficiently small time step τ and apply the scheme with a modified three-step time discretization in the form

$$\begin{aligned} &\text{a)}\ u_h^0 \in S_h, \quad (u_h^0, v_h) = (u^0, v_h) \quad \forall v_h \in S_h, \\ &\text{b)}\ u_h^{k+1} \in S_h, \quad k = 0,1,\dots r-1, \end{aligned} \tag{135}$$

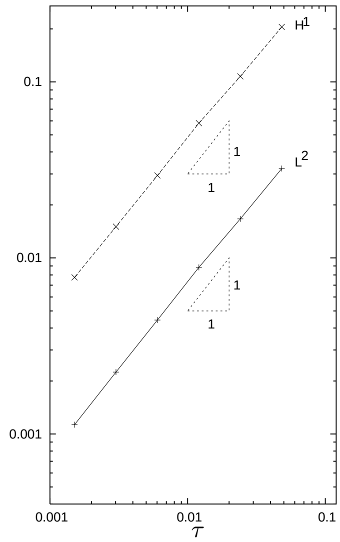


Fig. 3. The convergence of error in L^2–norm and H^1–seminorm with respect to τ (in logarithmic scale)

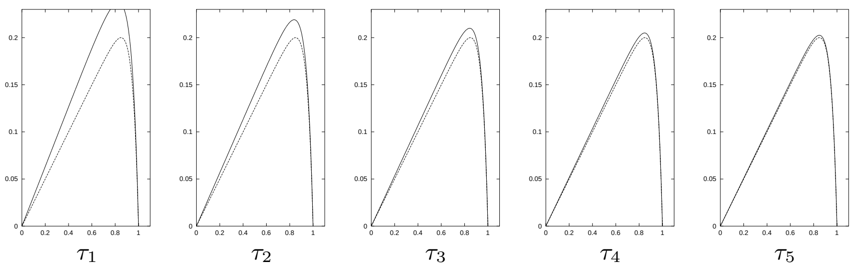


Fig. 4. Graphs of the numerical solutions (full lines) along the line $x_2 = 0.5$ for time steps $\tau_1, \ldots, \tau_5$ in comparison with the exact solution (dashed lines)

$$\text{c)}\quad \left(\frac{11u_h^{k+1} - 18u_h^k + 9u_h^{k-1} - 2u_h^{k-2}}{6\tau}, v_h\right)$$
$$+a_h(u_h^{k+1}, v_h) + b_h(\hat{u}_h^{k+1}, v_h) + \varepsilon J_h^\sigma(u_h^{k+1}, v_h) = \ell_h(v_h)\,(t_{k+1})$$
$$\forall\, v_h \in S_h,\ k = 2, 3, \ldots r-1,$$

where $\hat{u}_h^{k+1} = 3\hat{u}_h^k - 3\hat{u}_h^{k-1} + \hat{u}_h^{k-2}$ for $k = 2, \ldots, r-1$. The values u_h^1 and u_h^2 are obtained by the fourth-order Runge-Kutta scheme from [20], Section 3.2.2. The above scheme has formally third-order of accuracy with respect to time. The extension of the algorithmization from scheme (42), a)–c) to (135), a)–c) does not represents any difficulties and any increase of CPU time.

Table 2. Number of elements $\#\mathcal{T}_{h_l}$, maximal diameter h and the parameter h_l for triangulations $\mathcal{T}_{h_l},\ l = 1, \dots, 7$

l	$\#\mathcal{T}_{h_l}$	h	h_l
1	148	1.733E-01	1.898E-01
2	289	1.277E-01	1.358E-01
3	591	9.008E-02	9.500E-02
4	1056	6.371E-02	7.110E-02
5	2360	4.510E-02	4.754E-02
6	4219	3.194E-02	3.555E-02
7	9372	2.262E-02	2.386E-02

We define the function g and the initial and boundary conditions in such a way that the exact solution has the form

$$u(x_1, x_2, t) = \left(1 - e^{-10t}\right) \hat{u}(x_1, x_2), \tag{136}$$

where $\hat{u}$ is given by (131) with $\varepsilon = 0.1$. The solution u is equal to zero at $t = 0$ and for $t = 1$ is nearly identical with the steady-state solution $\hat{u}$.

We solve the problem in consideration with the aid of piecewise linear (P^1), quadratic (P^2) and cubic (P^3) elements. The computations were carried out on 7 unstructured regular triangular meshes ($\mathcal{T}_{h_l},\ l = 1, \dots, 7$), see Figures 2 (left) and 5. Since we use unstructured meshes, it is suitable to investigate the dependence of computational errors not on $h = \max_{K \in \mathcal{T}_h} h_K$, but to associate it with the number of elements of of the triangulation. Therefore, we define

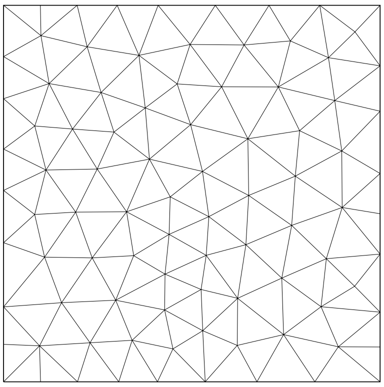
$$h_l = \frac{2}{\sqrt{\sqrt{3}\#\mathcal{T}_{h_l}}}, \quad l = 1, \dots, 7, \tag{137}$$

where $\#\mathcal{T}_{h_l}$ is the number of elements of $\mathcal{T}_h$. Relation (137) follows from the fact that the area of an equilateral triangle with sides of the length equal to h is $\sqrt{3}h^2/4$. The values of h and h_l are shown in Table 6.2.

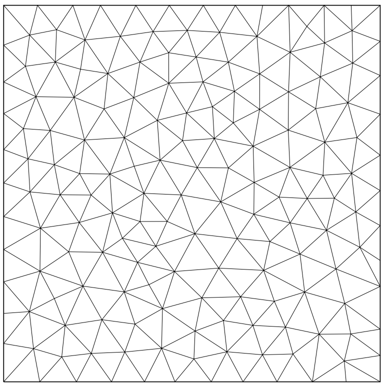
The time step τ was chosen according to the formula

$$\tau \le \frac{1}{75}\, h, \tag{138}$$

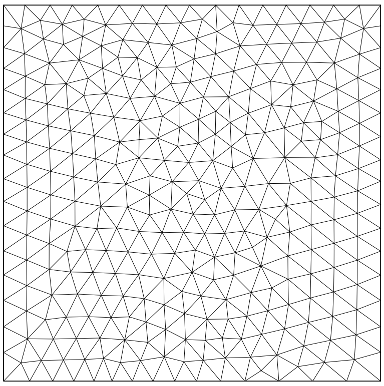
where h is the space mesh size. Numerical experiments indicate, that this choice is sufficient. Smaller τ does not cause any further decrease of computational errors.



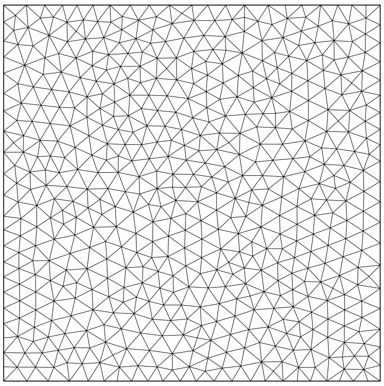
$\mathcal{T}_{h_1}$



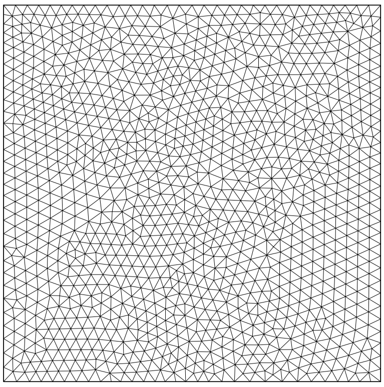
$\mathcal{T}_{h_2}$



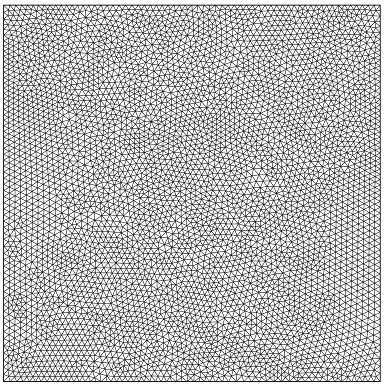
$\mathcal{T}_{h_3}$



$\mathcal{T}_{h_4}$



$\mathcal{T}_{h_5}$



$\mathcal{T}_{h_7}$

Fig. 5. Triangulations $\mathcal{T}_{h_l},\ l = 1, \ldots, 5, 7$

The computational error of the solution is evaluated at time $T = 1$ in $L^2(\Omega)$–norm and in $H^1(\Omega, \mathcal{T}_h)$–seminorm:

$$e_h^0 \equiv \|u_h(\cdot, T) - u(\cdot, T)\|_{L^2(\Omega)}, \tag{139}$$
$$e_h^1 \equiv |u_h(\cdot, T) - u(\cdot, T)|_{H^1(\Omega,\mathcal{T}_h)},$$

where $u(\cdot, T)$ is the exact solution given by (136) and (131) of equation (125) at time T and $u_h(\cdot, T)$ is the numerical solution at time T obtained by scheme (42), a)–c). We suppose that the errors behave according to the formula

$$e_h^n \approx D_n h^{\alpha^n}, \quad n = 0, 1, \tag{140}$$

where $D_n > 0,\ n = 0, 1$, are constants independent of h and $\alpha^n,\ n = 0, 1$, are the *orders of accuracy* of the method. We define the *local experimental order of convergence* by

$$\alpha_l^n = \frac{\log\left(e_{h_l}^n / e_{h_{l-1}}^n\right)}{\log\left(h_l / h_{l-1}\right)}, \quad l = 2, \ldots, 7,\ n = 0, 1. \tag{141}$$

The *global experimental order of convergence* $\bar{\alpha}^n,\ n = 0, 1$, is obtained by the least squares method.

Tables 6.2, 6.2 and 6.2 show the errors in L^2-norm and H^1-seminorm, the values of α_l^n, for $l = 2, \ldots, 7$ and $\bar{\alpha}^n$ for P^1, P^2 and P^3 approximations, respectively. The orders of convergence in comparison with expected optimal error slopes are shown in Figure 6. Figures 7 – 9 show the graphs of the piecewise linear, quadratic and cubic approximate solutions along the line $x_2 = 0.5$, obtained on the meshes $\mathcal{T}_{h_1}$, $\mathcal{T}_{h_4}$ and $\mathcal{T}_{h_7}$, in comparison with the exact solution. We observe here the convergence of the approximate solution to the exact solution for $h \to 0+$ in depedence on increasing p. (It is possible to speak about *hp*-version of the DGFEM.) The above results also show an improvement of the quality of the numerical solution in dependence on p. For $p = 1$, spurious overshoots appear in boundary layers, but they disappear when p is increased.

One can see some discrepancies between teoretically derived error estimates and expected optimal error estimates on one side and numerical experiments on the other side. This is caused by the fact that in numerical computations there are additional sources of errors beside the discretization. The experimental results can be influenced by numerical integration errors, round-off errors and errors caused by the iterative solution of the linear discrete problem on each time level. Moreover, the accuracy is also influenced by the fact that the computational meshes in numerical experiments are nonuniform and

Table 3. Computational errors in L^2–norm and H^1–seminorm, values of α_l^n, $l = 2, \dots, 7$, and $\bar{\alpha}^n, n = 0, 1$, for P^1 elements

		L^2–norm		H^1–seminorm	
l	h_l	e_h^2	α_l^2	e_h^1	α_l^1
1	1.898E-01	3.466 E-01	–	3.139 E-00	–
2	1.358E-01	1.858 E-01	1.864	2.208 E-00	1.051
3	9.500E-02	8.277 E-02	2.260	1.231 E-00	1.633
4	7.110E-02	5.232 E-02	1.580	9.605 E-01	0.855
5	4.754E-02	2.334 E-02	2.008	5.398 E-01	1.433
6	3.555E-02	1.580 E-02	1.342	4.436 E-01	0.675
7	2.386E-02	6.264 E-03	2.319	2.524 E-01	1.413
global order $\bar{\alpha}^n$			1.902		1.210

Table 4. Computational errors in L^2–norm and H^1–seminorm, values of α_l^n, $l = 2, \dots, 7$, and $\bar{\alpha}^n, n = 0, 1$, for P^2 elements

		L^2–norm		H^1–seminorm	
l	h_l	e_h^2	α_l^2	e_h^1	α_l^1
1	1.898E-01	4.012 E-02	–	7.348 E-01	–
2	1.358E-01	1.975 E-02	2.119	4.261 E-01	1.628
3	9.500E-02	9.573 E-03	2.024	1.942 E-01	2.197
4	7.110E-02	5.694 E-03	1.791	1.280 E-01	1.437
5	4.754E-02	2.571 E-03	1.977	5.618 E-02	2.048
6	3.555E-02	1.622 E-03	1.585	3.727 E-02	1.413
7	2.386E-02	6.355 E-04	2.348	1.493 E-02	2.292
global order $\bar{\alpha}^n$			1.955		1.858

Table 5. Computational errors in L^2–norm and H^1–seminorm, values of α_l^n, $l = 2, \dots, 7$, and $\bar{\alpha}^n, n = 0, 1$, for P^3 elements

		L^2–norm		H^1–seminorm	
l	h_l	e_h^2	α_l^2	e_h^1	α_l^1
1	1.898E-01	8.290 E-03	–	1.835 E-01	–
2	1.358E-01	2.929 E-03	3.110	9.704 E-02	1.905
3	9.500E-02	5.899 E-04	4.480	3.250 E-02	3.058
4	7.110E-02	2.621 E-04	2.795	1.494 E-02	2.679
5	4.754E-02	5.463 E-05	3.900	4.783 E-03	2.832
6	3.555E-02	2.596 E-05	2.561	2.126 E-03	2.792
7	2.386E-02	4.191 E-06	4.570	6.040 E-04	3.153
global order $\bar{\alpha}^n$			3.611		2.785

non-nested. One has to take into account that the derived estimates are asymptotic, valid for $h \to 0+$. Also nonlinearity may play a role and cause an oscillatory behaviour of the experimental order of convergence. On the other hand, the numerical experiments demonstrate well that the discontinuous Galerkin method is suitable for the numerical solution of nonlinear convection-diffusion problems.

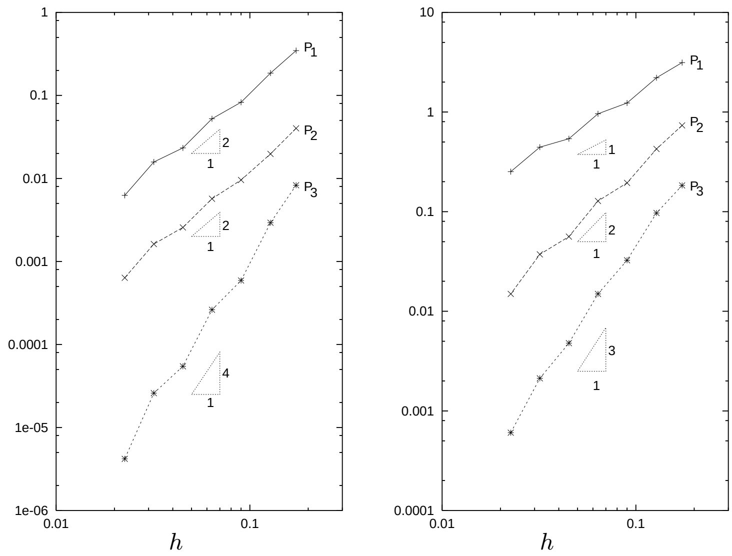


Fig. 6. The convergence of error in L^2–norm (left) and H^1–seminorm (right) with respect to h (in logarithmic scale) and expected optimal error slopes

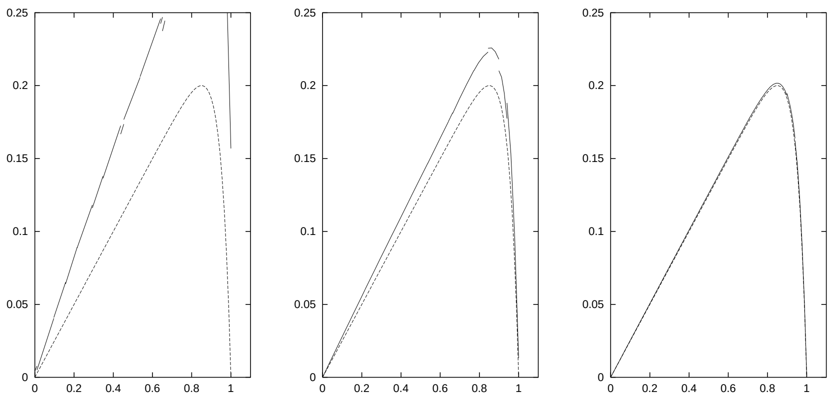


Fig. 7. Graphs of the numerical solutions (full lines) along the line $x_2 = 0.5$ for P^1 (left), P^2 (centre) and P^3 (right) elements on the mesh $\mathcal{T}_{h_1}$ in comparison with the exact solution (dashed lines)

7 Conclusion

We have presented an efficient numerical method for the solution of nonstacionary nonlinear convection-diffusion problems, which is based on the space discretization by the discontinuous Galerkin finite element method and a semi-implicit time discretization. We have derived a priori error estimates, namely $O(h^p+\tau)$ in $L^\infty(0,T;L^2(\Omega))$–

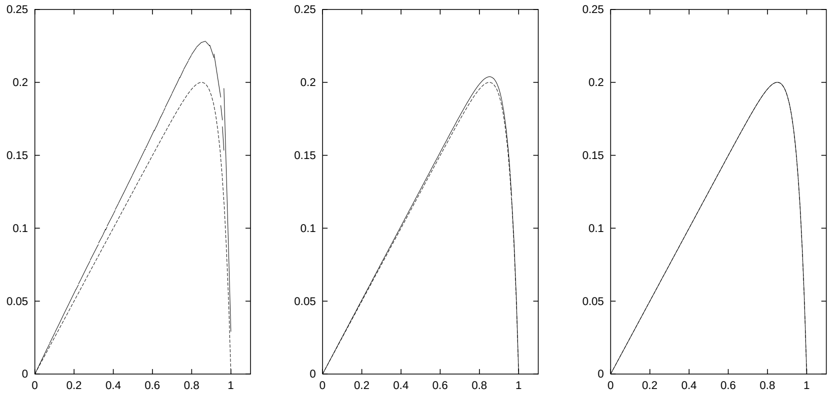


Fig. 8. Graphs of the numerical solutions (full lines) along the line $x_2 = 0.5$ for P^1 (left), P^2 (centre) and P^3 (right) elements on the mesh $\mathcal{T}_{h_4}$ in comparison with the exact solution (dashed lines)

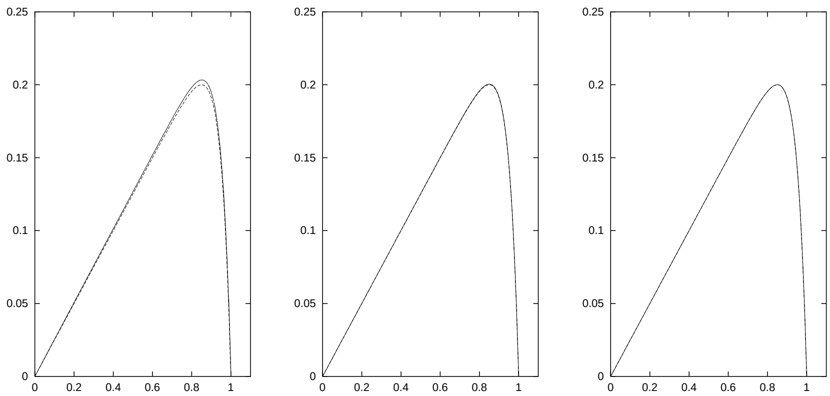


Fig. 9. Graphs of the numerical solutions (full lines) along $x_2 = 0.5$ for P^1 (let), P^2 (centre) and P^3 (right) elements on the mesh $\mathcal{T}_{h_7}$ in comparison with the exact solution (dashed lines)

norm , $L^2(0,T;H^1(\Omega))$–seminorm and $L^\infty(0,T;H^1(\Omega))$–seminorm. On the other hand, numerical examples indicate a better behaviour of the experimental $L^\infty(L^2)$–order of convergence, which is expected to be asymptotically $O(h^{p+1}+\tau)$ for p odd and $O(h^p+\tau)$ for p even (in our case for $p = 1, 2, 3$). The obtained results show that the DGFEM is a challenging method for the numerical solution of nonstationary nonlinear convection-diffusion problems. It is sufficiently accurate and robust and yields sufficiently precise approximations of solutions with steep gradients.

There are several items for the future work:

– derivation of optimal error estimates,
– avoiding the blow up behaviour of estimates with respect to $\varepsilon \to 0+$,
– development of efficient a posteriori error estimates,
– increas of accuracy in time discretization,
– application of the method to the solution of compressible flow, with the goal to work out efficient, accurate and robust method for the solution of flow problems with a wide range of Reynolds and Mach numbers.

References

1. Ph. Angot, V. Dolejší, M. Feistauer, and J. Felcman, Analysis of a combined barycentric finite volume – nonconforming finite element method for nonlinear convection-diffusion problems, Appl. Math. **43**, (1998) 263–310.
2. D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. **39**, (2002) 1749–1779.
3. K. Baba and M. Tabata, On a conservative upwind finite element scheme for convective diffusin equations, RAIRO Anal. Numér. **15**, (1981) 3–25.
4. I. Babuška and M. Zlámal, Nonconforming elements in the finite element method with penalty, SIAM J. Numer. Anal. **10**, (1973) 863–875.
5. I. Babuška, C. E. Baumann, and J. T. Oden, A discontinuous *hp* finite element method for diffusion problems: 1-d analysis, Computers and Mathematics with Applications **37**, (1999) 103–122.
6. C. E. Baumann and J. T. Oden, A discontinuous *hp* finite element method for the Euler and Navier-Stokes equations, Int. J. Numer. Methods Fluids **31**, (1999) 79–95.
7. S. Brenner and R. L. Scott, *The Mathematical Theory of Finite Element Methods* (Spriger, New York, 1994).
8. P. G. Ciarlet, *The Finite Elements Method for Elliptic Problems* (North-Holland, Amsterdam, New York, Oxford, 1979).
9. B. Cockburn, Discontinuous Galerkin methods for convection dominated problems. In T. J. Barth and H. Deconinck, editors, *High–Order Methods for Computational Physics*, Lecture Notes in Computational Science and Engineering 9, (Springer, Berlin, 1999) 69–224.
10. B. Cockburn, G. E. Karniadakis, and C.-W. Shu (eds.), Discontinuous Galerkin methods. In *Lecture Notes in Computational Science and Engineering 11.* Springer, Berlin, 2000.
11. V. Dolejší, On the discontinuous Galerkin method for the numerical solution of the Navier–Stokes equations, Int. J. Numer. Methods Fluids, (to appear).
12. V. Dolejší, M. Feistauer, J. Felcman, and A. Kliková, Error estimates for barycentric finite volumes combined with nonconforming finite elements applied to nonlinear convection–diffusion problems, Appl. Math. **47**, (2002) 301–340.
13. V. Dolejší, M. Feistauer, Error estimates of the discontinuous Galerkin method for nonlinear nonstationary convection-diffusion problems. Numer. Funct. Anal. Optimiz. (submitted).

14. V. Dolejší, M. Feistauer, and C. Schwab, A finite volume discontinuous Galerkin scheme for nonlinear convection–diffusion problems, Calcolo **39**, (2002) 1–40.
15. V. Dolejší, M. Feistauer, and V. Sobotíková, A discontinuous Galerkin method for nonlinear convection–diffusion problems, Comput. Methods Appl. Mech. Eng. (submitted).
16. M. Feistauer and J. Felcman, Theory and applications of numerical schemes for nonlinear convection-diffusion problems and compressible Navier–Stokes equations, In J. R. Whitemen, editor, *The Mathematics of Finite Element and Applications*, (John Wiley & Sons, 1996) 175–194.
17. M. Feistauer, J. Felcman, and V. Dolejší, Numerical simulation of compresssible viscous flow through cascades of profiles, ZAMM **76**, (1996) 297–300.
18. M. Feistauer, J. Felcman, and M. Lukáčová-Medvidová, On the convergence of a combined finite volume – finite element method for nonlinear convection – diffusion problems, Numer. Methods Partial Differ. Equations **13**, (1997) 163–190.
19. M. Feistauer, J. Felcman, M. Lukáčová-Medvidová, and G. Warnecke, Error estimates of a combined finite volume – finite element method for nonlinear convection – diffusion problems, SIAM J. Numer. Anal. **36**, (1999) 1528–1548.
20. M. Feistauer, J. Felcman, I. Straškraba, *Mathematical and Computational Methods for Compressible Flow* (Clarendon Press, Oxford, 2003).
21. M. Feistauer, J. Slavík, P. Stupka, On the convergence of a combined finite volume-finite element method for nonlinear convection-diffusion problems. Explicit schemes, Numer. Methods Partial Differ. Equations **15**, (1999) 215-235.
22. M. Feistauer, K. Švadlenka, Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems, J. Numer. Math. **12**, (2004) 97-118.
23. J. Felcman and V. Dolejší, Numerical solution of compresssible flow, ZAMM **77**, (1997) 473–476.
24. P. Houston, C. Schwab, and E. Süli, Discontinuous *hp*-finite element methods for advection-diffusion problems, SIAM J. Numer. Anal. **39**, (2002) 2133–2163.
25. T. Ikeda, Maximum principle in finite element models for convection–diffusion phenomena, In *Lecture Notes in Numerical and Applied Analysis*, volume 4 of *Mathematics Studies 76*. (North-Holland, Amsterdam – New York – Oxford, 1983).
26. C. Johnson, Finite element methods for convection – diffusion problems. In R. Glowinski and J. L. Lions, editors, *Computing Methods in Engineering and Applied Sciences V.*, (North–Holland, Amsterdam, 1982) 311–323.
27. A. Kufner, O. John, and S. Fučík, *Function Spaces* (Academia, Prague, 1977).
28. P. L. Lions, *Mathematical Topics in Fluid Mechanics* (Oxford Science Publications, 1996).
29. J. M. Melenk and C. Schwab, An *hp* finite element method for convection–diffusion problems, East–West J. Numer. Math **7**, (1999) 31–60.
30. J. J. H. Miller and S. Wang, A new non-conforming Petrov–Galerkin finite-element method with triangular elements for a singularly perturbed advection-diffusion problem, IMA J. Numer. Anal. **14**, (1994) 257–276.
31. K. W. Morton, *Numerical Solution of Convection – Diffusion Problems* (Chapman & Hall, London, 1996).

32. J. T. Oden, I. Babuška, and C. E. Baumann, A discontinuous *hp* finite element method for diffusion problems. J. Comput. Phys **146** (1998) 491–519.
33. K. Ohmori and T. Ushijima, A technique of upstream type applied to a linear nonconforming finite element approximation of convective diffusion equations, RAIRO, Anal. Numér. **18**, (1984) 309–322.
34. S. Prudhomme, F. Pascal, J.T. Oden, and A. Romkes, Review of a priori error estimation for discontinuous Galerkin methods, Technical report, Laboratoire de Mathématiques, Université Paris-Sud, 2002.
35. K. Rektorys, *The Method of Discretization in Time and Partial Differential Equations* (Reidel, Dodrecht, 1982).
36. B. Rivière and M. F. Wheeler, A discontinuous Galerkin method applied to nonlinear parabolic equations, In B. Cockburn, editor, *Discontinuous Galerkin methods. Theory, computation and applications.*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, (Berlin, Springer, 2000) 231–244.
37. H.-G. Roos, M. Stynes, and L. Tobiska, Numerical Methods for Singularly Perturbed Differential Equation, Springer Series in Computational Mathematics **24**, (Springer-Verlag, Berlin, 1996).
38. A. A. Samarskii and P. N. Vabishchevich, *Numerical Methods for the Solution of Convection-Diffusion Problems* (Publisher URSS, Moscow, 1999), (in Russian).
39. T. Skalický, LASPACK, Dresden University of Technology, version 1.12.2 edition, 1995. www.netlib.org/laspack.
40. R. Verfürth, A note on polynomial approximation in Sobolev spaces. $\mathrm{M}^2\mathrm{AN}$ **33**, (1999) 715–719.
41. G. Zhou, A local L^2-error analysis of the streamline diffusion method for nonstationary convection–diffusion systems. M^2AN **29** (1995) 577–603.