Analysis of Semi-Implicit DGFEM for Nonlinear Convection–Diffusion Problems on Nonconforming Meshes ⋆

V. Dolejší [a], M. Feistauer [a,*], J. Hozman [a]

[a] *Charles University Prague, Faculty of Mathematics and Physics, Sokolovská 83, 186 75 Praha, Czech Republic*

Dedicated to Professor Ivo Babuška on the occasion of his 80th birthday

Abstract

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), on general nonconforming meshes formed by possibly nonconvex elements, 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. Numerical examples demonstrate the accuracy of the method and manifest the effect of nonconvexity of elements and nonconformity of the mesh.

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

⋆ This work is a part of the research project MSM 0021620839 financed by the Ministry of Education of the Czech Republic and was partly supported by the Grant No. 201/05/0005 of the Czech Grant Agency.

* Corresponding author.
Email addresses: dolejsi@karlin.mff.cuni.cz (V. Dolejší), feist@karlin.mff.cuni.cz (M. Feistauer), jhozmi@volny.cz (J. Hozman).

1 Introduction

The numerical solution of nonlinear conservation laws, convection-diffusion problems and flow problems requires the application of efficient, robust and accurate methods allowing to overcome various difficulties, as the precise capturing and resolution of boundary layers, shock waves and contact discontinuities. It is possible to say that nowadays in computational fluid dynamics (CFD) two techniques compete: the finite volume (FV) schemes and stabilized finite element methods (FEM). A survey of FV as well as FE approaches to the numerical simulation of compressible flow can be found, e.g. in [23].

A natural generalization of the FV and FE techniques 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 allows 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 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 [3] and [5] 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 [2]). The nonsymmetric variant was also investigated in [10], [8], [9], [29] for elliptic and parabolic problems and in [19] and [20] for nonlinear convection-diffusion problems. Although this approach does not give an optimal order of convergence for elliptic problems, it leads to a coercive operator for an arbitrary positive penalty coefficient. This property is important when the DGFEM is applied to the system of the Navier-Stokes equations, where the numerical analysis is rather complicated, see [16].

There is a number of works devoted to theory and applications of the DGFEM. Let us mention, e.g., [1], [3], [5], [4], [7], [14], [15], [16], [19], [23], [24], [25], [27], [29], [31]. For a survey of various discontinuous Galerkin techniques, see, e. g. [12], [13].

In [17] and [20] we carried out a discretization of a scalar nonstationary convection-diffusion equation with nonlinear convective terms by the 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 schemes are conditionally stable and the time step is drastically limited by the CFL stability condition. In order to avoid this disadvantage, it seems suitable to apply an implicit method, which allows us 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 in the present paper, which is a continuation of [20], we propose 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. Similarly as in [20] we allow to use a nonconforming mesh formed by nonconvex star-shaped polyhedral elements. In this paper we shall be concerned with theoretical analysis of error estimates of the semi-implicit method and present several numerical experiments verifying the theoretical results. Also the effect of nonconvexity of elements and nonconformity of a mesh will be treated in numerical experiments.

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 establish 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 numerical examples demonstrating the accuracy and robustness of the DGFEM. In Section 7 we introduce some concluding remarks and formulate open problems.

2 Continuous problem

Let $\Omega \subset \mathbb{R}^d$ ($d = 2$ 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 $u : Q_T \to \mathbb{R}$ 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:

$$\begin{aligned}
&\text{a) } f_s \in C^1(I\!R),\ f_s(0) = 0,\ s = 1,\dots,d, \\
&\text{b) } \varepsilon > 0, \\
&\text{c) } g \in C([0,T]; L^2(\Omega)), \\
&\text{d) } u_D \text{ is the trace of some } u^* \in C([0,T]; H^1(\Omega)) \cap L^\infty(Q_T) \\
&\quad \text{on } \partial\Omega \times (0,T), \\
&\text{e) } u^0 \in L^2(\Omega).
\end{aligned} \tag{4}$$

We use the standard notation for function spaces (see, e. g. [26]): $L^p(\Omega)$, $L^p(Q_T)$ 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.

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.

We shall assume that problem (1) – (3) has a weak solution (cf. [20], Section 2), satisfying the regularity conditions

$$\begin{aligned}
&u \in L^\infty(0,T; H^{p+1}(\Omega)), \\
&\frac{\partial u}{\partial t} \in L^\infty(0,T; H^{p+1}(\Omega)), \\
&\frac{\partial^2 u}{\partial t^2} \in L^\infty(0,T; L^2(\Omega)),
\end{aligned} \tag{5}$$

where an integer $p \geq 1$ will denote a given degree of polynomial approximations. Such a solution satisfies problem (1) – (3) pointwise. Under (5),

$$u \in C([0,T]; H^{p+1}(\Omega)), \quad \frac{\partial u}{\partial t} \in C([0,T]; L^2(\Omega)). \tag{6}$$

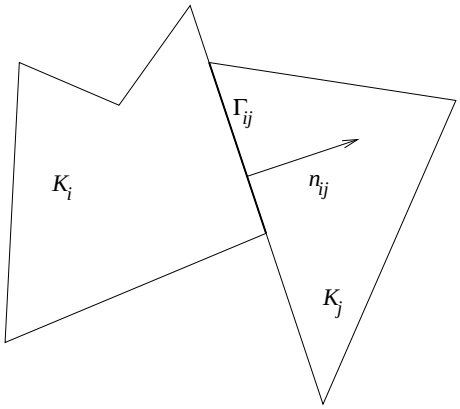


Fig. 1. Neighbouring elements $K_i,\ K_j$

3 Discretization of the problem

3.1 Triangulations

We use the same notation as in [20], Section 3.1. By T_h $(h > 0)$ we denote a partition of the closure $\overline{\Omega}$ of the domain Ω into a finite number of closed d-dimensional star-shaped polyhedra K with mutually disjoint interiors. They can be even nonconvex. All elements are numbered so that $T_h = \{K_i\}_{i\in I}$, where $I \subset Z^+ = \{0, 1, 2, \ldots\}$ is a suitable index set. We denote $h_K = \mathrm{diam}(K)$, $h = \max_{K\in T_h} h_K$, ρ_K – radius of the largest d-dimensional ball inscribed into K, $|K|$ – d-dimensional Lebesgue measure of K, $\Gamma_{ij} = \Gamma_{ji} = \partial K_i \cap \partial K_j$, provided $i \neq j$ and $K_i, K_j \in T_h$ contain a common nonempty open face. Then we call K_i and K_j neighbours. (See Figure 1, showing a possible 2D situation.) The boundary $\partial\Omega$ is formed by a finite number of faces $S_j, j \in I_b \subset \{-1, -2, \ldots\}$ of elements K_i adjacent to $\partial\Omega$. For $i \in I$ we set

$$\begin{aligned} s(i) &= \{j \in I; K_j \text{ is a neighbour of } K_i\}, \\ \gamma(i) &= \{j \in I_b; S_j \text{ is a face of } K_i\}, \\ \Gamma_{ij} &= S_j \text{ for such } K_i \in T_h \text{ that } S_j \subset \partial K_i,\ j \in I_b. \end{aligned} \tag{7}$$

If we write $S(i) = s(i) \cup \gamma(i)$, then

$$\partial K_i = \bigcup_{j\in S(i)} \Gamma_{ij}, \quad \partial K_i \cap \partial\Omega = \bigcup_{j\in\gamma(i)} \Gamma_{ij}. \tag{8}$$

Furthermore, we use the following notation: $\boldsymbol{n}_{ij} = ((n_{ij})_1, \ldots, (n_{ij})_d)$ – unit outer normal to ∂K_i on the face Γ_{ij} (see Figure 1), $|\Gamma_{ij}|$ – $(d-1)$-dimensional measure of Γ_{ij}, $d(\Gamma_{ij}) = \mathrm{diam}(\Gamma_{ij})$. We have

$$
\begin{aligned}
\boldsymbol{n}_{ij} &= -\boldsymbol{n}_{ji} \\
|K_i| &\le h_{K_i}^d \le h^d, \\
d(\Gamma_{ij}) &\le h_{K_i} \le h.
\end{aligned} \tag{9}
$$

Over the triangulation T_h we define the so-called *broken Sobolev space*

$$
H^k(\Omega, T_h) = \{v; v|_K \in H^k(K) \ \forall\, K \in T_h\} \tag{10}
$$

equipped with the norm

$$
\|v\|_{H^k(\Omega,T_h)} = \left(\sum_{K \in T_h} \|v\|^2_{H^k(K)} \right)^{1/2} \tag{11}
$$

and the seminorm

$$
|v|_{H^k(\Omega,T_h)} = \left(\sum_{K \in T_h} |v|^2_{H^k(K)} \right)^{1/2}. \tag{12}
$$

For $v \in H^1(\Omega, T_h)$, $i \in I$ and $j \in s(i)$ we denote

$$
\begin{aligned}
v|_{\Gamma_{ij}} &= \text{the trace of } v|_{K_i} \text{ on } \Gamma_{ij}, \\
v|_{\Gamma_{ji}} &= \text{the trace of } v|_{K_j} \text{ on } \Gamma_{ji}, \\
\langle v \rangle_{\Gamma_{ij}} &= \frac{1}{2} \left(v|_{\Gamma_{ij}} + v|_{\Gamma_{ji}} \right), \\
[v]_{\Gamma_{ij}} &= v|_{\Gamma_{ij}} - v|_{\Gamma_{ji}}.
\end{aligned} \tag{13}
$$

In the discretization we shall use the nonsymmetric variant of the DG representation of the diffusion terms, introduced by Oden, Baumann and Babuška, combined with *interior and boundary penalty*. We speak of a NIPG variant (see [28]). The convective terms are approximated with the aid of a *numerical flux*. This is an important ingredient in the finite volume schemes and allows to approximate convective terms in a natural way also in the framework of the DGFEM. The interior and boundary penalty terms are used here because of the following reasons. First, they stabilize the diffusion approximation and play an important role in the proof of the coercivity of the problem. (It is well-known that for schemes without penalty terms the error estimates can be proven only for elements of degree ≥ 2, see, e.g. [1], [3].) Second, the penalty terms are necessary for the control of nonlinear convective terms in the derivation of error estmates. See, e.g. estimates (38) – (40).

In [20], Section 3, the DG space semidiscretization (method of lines) was introduced. To this end, the following forms were defined: For $u, v \in H^2(\Omega, T_h)$ we set

$$(u,v) = \int_\Omega uv \mathrm{d}x, \tag{14}$$

$$\begin{aligned} a_h(u,v) = \varepsilon \sum_{i\in I} \Bigg\{ & \int_{K_i} \nabla u \cdot \nabla v \,\mathrm{d}x \\ & - \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \Big(\langle \nabla u \rangle \cdot \boldsymbol{n}_{ij}[v] - \langle \nabla v \rangle \cdot \boldsymbol{n}_{ij}[u] \Big) \,\mathrm{d}S \\ & - \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \Big((\nabla u \cdot \boldsymbol{n}_{ij})\, v - (\nabla v \cdot \boldsymbol{n}_{ij})\, u \Big) \,\mathrm{d}S \Bigg\}, \end{aligned}$$

$$\begin{aligned} b_h(u,v) = & \sum_{i\in I} \Bigg(\sum_{j\in s(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}} \,\mathrm{d}S \\ & \qquad + \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ij}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}} \,\mathrm{d}S \Bigg) \\ & - \sum_{i\in I} \int_{K_i} \sum_{s=1}^{d} f_s(u) \frac{\partial v}{\partial x_s} \,\mathrm{d}x, \quad u,\, v \in H^1(\Omega, T_h),\ u \in L^\infty(\Omega), \end{aligned} \tag{15}$$

(16)

$$J_h^\sigma(u,v) = \sum_{i\in I} \Bigg\{ \sum_{j\in s(i)} \int_{\Gamma_{ij}} \sigma [u]\,[v] \,\mathrm{d}S + \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma uv \,\mathrm{d}S \Bigg\} \tag{17}$$

$$\ell_h(v)\,(t) = (g(t), v) + \varepsilon \sum_{i\in I} \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \left(\nabla v \cdot \boldsymbol{n}_{ij}\, u_D(t) + \sigma u_D(t)\, v \right) \mathrm{d}S, \tag{18}$$

representing the $L^2(\Omega)$-scalar product, diffusion terms, convective terms, interior and boundary penalty and right-hand side. The function H is a numerical flux (also called an approximate Riemann solver). The weight function $\sigma : \bigcup_{i\in I} \bigcup_{j\in S(i)} \Gamma_{ij} \to I\!R$ from the penalty terms is defined by

$$\sigma|_{\Gamma_{ij}} = \frac{1}{d(\Gamma_{ij})}. \tag{19}$$

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.) We assume that the numerical flux has the following properties:

Assumptions (H)

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

$$|H(u, v, \boldsymbol{n}) - H(u^*, v^*, \boldsymbol{n})| \le C_1(|u - u^*| + |v - v^*|), \qquad u,\ v,\ u^*,\ v^* \in \mathbb{R},\ \boldsymbol{n} \in \boldsymbol{S}_1. \tag{20}$$

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

$$H(u, u, \boldsymbol{n}) = \sum_{s=1}^{d} f_s(u)\, n_s, \quad u \in \mathbb{R},\ \boldsymbol{n} = (n_1, \ldots, n_d) \in \boldsymbol{S}_1. \tag{21}$$

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

$$H(u, v, \boldsymbol{n}) = -H(v, u, -\boldsymbol{n}), \quad u,\ v \in \mathbb{R},\ \boldsymbol{n} \in \boldsymbol{S}_1. \tag{22}$$

In virtue of (20) and (21), the functions f_s, $s = 1, \ldots, d$, are Lipschitz-continuous with constant $L_f = 2C_1$. From (4), a) and (21) we see that

$$H(0, 0, \boldsymbol{n}) = 0 \quad \forall\, \boldsymbol{n} \in \boldsymbol{S}_1. \tag{23}$$

As an example of a numerical flux satisfying the above assumptions we can use the numerical flux from Section 6 or multidimensional variants of well-known approximate Riemann solvers for conservation laws. Let us mention, e.g. the numerical fluxes by Engquist–Osher and Lax–Friedrichs, see e.g., [21].

Now we define the space of discontinuous piecewise polynomial functions

$$S_h = S^{p,-1}(\Omega, T_h) = \{v; v|_K \in P^p(K)\ \forall\, K \in T_h\}, \tag{24}$$

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

In [20] we showed that the exact solution u with property (5) satisfies the identity

$$\left(\frac{\partial u}{\partial t}(t), v_h\right) + a_h(u(t), v_h) + b_h(u(t), v_h) + \varepsilon J_h^\sigma(u(t), v_h) = \ell_h(v_h)\,(t) \tag{25}$$

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

Now we are ready to introduce the fully discretized 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 (25) 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.

We define the *approximate solution* of problem (1) – (3) as functions u_h^k, $t_k \in [0, T]$, satisfying the conditions

$$
\begin{aligned}
&\text{a) } u_h^{k+1} \in S_h, \\
&\text{b) } \left(\frac{u_h^{k+1} - u_h^k}{\tau_k}, v_h \right) + a_h(u_h^{k+1}, v_h) + b_h(u_h^k, v_h) \\
&\qquad + \varepsilon J_h^\sigma(u_h^{k+1}, v_h) = \ell_h(v_h)\,(t_{k+1}) \quad \forall\, v_h \in S_h,\ \forall\, t_{k+1} \in (0, T], \\
&\text{c) } u_h^0 = \Pi^{L^2} u^0.
\end{aligned}
\tag{26}
$$

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

In (26), c), Π^{L^2} denotes the operator of L^2-projection on the space S_h. This means that $\Pi^{L^2} v \in S_h$ and

$$
\left(\Pi^{L^2} v, \varphi \right) = (v, \varphi) \quad \forall \varphi \in S_h. \tag{27}
$$

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

$$
(\Pi^{L^2} v, \varphi)_{L^2(K)} = (v, \varphi)_{L^2(K)} \quad \forall \varphi \in P^p(K),\ \forall K \in T_h. \tag{28}
$$

For each $t_{k+1} \in (0, T]$ problem (26), a) – b) is equivalent to a system of linear algebraic equations with a nonsymmetric, but positive definite matrix, which can be solved by a suitable solver. This implies the following result.

Lemma 1 *The discrete problem (26) a) – c) has a unique solution.*

In what follows we shall be concerned with the analysis of method (26), a) – c).

4 Some auxiliary results

In this section we summarize some important results and properties which have been proven in [19] and [20].

4.1 Geometry of the mesh

Let us consider a system $\{T_h\}_{h\in(0,h_0)}$, $h_0 > 0$, of partitions of the domain Ω, i.e. $T_h = \{K_i\}_{i\in I_h}$, $I_h \subset Z^+$. 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 $\{T_h\}_{h\in(0,h_0)}$ has the following properties:

(A1) Each element $K \in T_h,\ h \in (0, h_0)$, 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

$$\frac{\max_{x\in\partial K} |x - x_K|}{\min_{x\in\partial K} |x - x_K|} \le \kappa \quad \forall K \in T_h,\ \forall h \in (0, h_0). \tag{29}$$

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

$$K = \bigcup_{S\in\boldsymbol{S}(K)} S. \tag{30}$$

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

$$\frac{h_S}{\rho_S} \le C_2 \quad \forall S \in \boldsymbol{S}(K) \qquad \text{(shape regularity)}, \tag{31}$$

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

$$1 \le \frac{h_K}{h_S} \le \tilde{\kappa} < \infty \quad \forall S \in \boldsymbol{S}(K), \tag{32}$$

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

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

$$h_{K_i} \le C_3\, d(\Gamma_{ij}), \quad \forall i \in I,\ j \in S(i),\ h \in (0, h_0). \tag{33}$$

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 estimates can be established:

$$\|v\|^2_{L^2(\partial K)} \le C_4 \left(\|v\|_{L^2(K)} \, |v|_{H^1(K)} + h_K^{-1} \|v\|^2_{L^2(K)} \right), \qquad \forall K \in T_h, \ v \in H^1(K), \ h \in (0, h_0) \tag{34}$$

(multiplicative trace inequality),

$$|v|_{H^1(K)} \le C_5 h_K^{-1} \|v\|_{L^2(K)} \quad \forall v \in P^p(K), \ K \in T_h \tag{35}$$

(inverse inequality).

There exist a constants $C_6 > 0$ independent of v and h and a linear mapping $\Pi : H^1(K) \to P^p(K), \ p \ge 0$, such that

$$\begin{aligned} \|\Pi v - v\|_{L^2(K)} &\le C_6 h_K^{p+1} |v|_{H^{p+1}(K)}, \\ |\Pi v - v|_{H^1(K)} &\le C_6 h_K^{p} |v|_{H^{p+1}(K)}, \\ |\Pi v - v|_{H^2(K)} &\le C_6 h_K^{p-1} |v|_{H^{p+1}(K)}, \end{aligned} \tag{36}$$

for all $v \in H^{p+1}(K), \ K \in T_h$ and $h \in (0, h_0)$ (approximation properties of S_h);

The operator Π is not the L^2- projector Π^{L^2} on S_h introduced in (27). From the construction of Π in [32] it follows that $\Pi v = v$ for $v \in S_h$. Moreover,

$$\|v - \Pi^{L^2} v\|_{L^2(K)} \le \|v - \Pi v\|_{L^2(K)} \le C_6 h^{p+1} |v|_{H^{p+1}(K)}, \qquad \forall v \in H^p(K), \ \forall K \in T_h, \tag{37}$$

as follows from (36).

Under assumptions (4), a), Assumptions (H) and Assumptions (A1), (A2), the form b_h has the following properties.

$$\begin{aligned} &|b_h(u, v) - b_h(\bar{u}, v)| \\ &\le C_7 \left(J_h^\sigma(v, v)^{1/2} + |v|_{H^1(\Omega, T_h)} \right) \\ &\quad \times \left(\|u - \bar{u}\|_{L^2(\Omega)} + \left(\sum_{i \in I} h_{K_i} \|u - \bar{u}\|^2_{L^2(\partial K_i)} \right)^{1/2} \right), \\ &\qquad\qquad u, \ \bar{u}, \ v \in H^1(\Omega, T_h), \end{aligned} \tag{38}$$

$$
\begin{aligned}
&|b_h(u_h, v_h) - b_h(\bar{u}_h, v_h)| \qquad (39)\\
&\le C_8 \left(J_h^\sigma(v_h, v_h)^{1/2} + |v_h|_{H^1(\Omega,T_h)} \right) \, \|u_h - \bar{u}_h\|_{L^2(\Omega)},\\
&\qquad\qquad u_h,\ \bar{u}_h,\ v_h \in S_h,\\
&|b_h(u, v_h) - b_h(\Pi u, v_h)| \qquad (40)\\
&\le C_9 h^{p+1} \left(J_h^\sigma(v_h, v_h)^{1/2} + |v_h|_{H^1(\Omega,T_h)} \right) |u|_{H^{p+1}(\Omega)},\\
&\qquad\qquad u \in H^{p+1}(\Omega),\ v_h \in S_h,
\end{aligned}
$$

where Πu is the S_h-interpolant of u from (36).

The proof of (34) can be found in [19], for (35), see [20]. As for (36), in the case of a simplicial mesh standard results from the finite element method can be employed (see, e. g. [11] or [6]). In our case, when general nonconvex star-shaped elements are used, the approximation properties (36) are derived under the above assumptions in [32]. Estimates (38) – (40) are established in [20], Lemmas 5 and 8.

5 Error estimates

Now we shall analyze the error estimates of the approximate solution $u_h^k,\ k = 0, 1, \ldots,$ obtained by method (26) under the assumption that the exact solution u satisfies (5). 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 (36). We set

$$
\xi^k = u_h^k - \Pi u^k \in S_h, \quad \eta^k = \Pi u^k - u^k \in H^{p+1}(\Omega, T_h). \qquad (41)
$$

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

$$
e_h^k = \xi^k + \eta^k, \quad k = 0, \ldots, r. \qquad (42)
$$

Setting $v_h := \xi^{k+1}$ in (26), b), we get

$$
\begin{aligned}
&\left(u_h^{k+1} - u_h^k, \xi^{k+1}\right) + \tau \left(a_h(u_h^{k+1}, \xi^{k+1}) + b_h(u_h^k, \xi^{k+1}) \right. \qquad (43)\\
&\qquad \left. + \varepsilon J_h^\sigma(u_h^{k+1}, \xi^{k+1}) - \ell_h(\xi^{k+1})\,(t_{k+1}) \right) = 0, \quad t_k, t_{k+1} \in [0, T].
\end{aligned}
$$

Moreover, setting $t := t_{k+1}$ and $v_h := \xi^{k+1}$ in (25), we obtain

$$\begin{aligned}&\left(u'(t_{k+1}),\xi^{k+1}\right)+a_h(u^{k+1},\xi^{k+1})+b_h(u^{k+1},\xi^{k+1})\\&\quad+\varepsilon J_h^\sigma(u^{k+1},\xi^{k+1})-\ell_h(\xi^{k+1})\,(t_{k+1})=0,\quad t_k,t_{k+1}\in[0,T],\end{aligned}\tag{44}$$

where $u'=\partial u/\partial t$.

Multiplying (44) by τ and subtracting from (43), we get

$$\begin{aligned}&\left(u_h^{k+1}-u_h^k,\xi^{k+1}\right)-\tau\left(u'(t_{k+1}),\xi^{k+1}\right)\\&+\tau\left(a_h(u_h^{k+1}-u^{k+1},\xi^{k+1})+b_h(u_h^k,\xi^{k+1})-b_h(u^{k+1},\xi^{k+1})\right.\\&\qquad\left.+\varepsilon J_h^\sigma(u_h^{k+1}-u^{k+1},\xi^{k+1})\right)=0,\quad k=0,\dots,r-1.\end{aligned}\tag{45}$$

By (41) and (42), from (45) we have

$$\begin{aligned}&\left(\xi^{k+1}-\xi^k,\xi^{k+1}\right)+\tau\left(a_h(\xi^{k+1},\xi^{k+1})+\varepsilon J_h^\sigma(\xi^{k+1},\xi^{k+1})\right)\\&=\tau(u'(t_{k+1}),\xi^{k+1})-(u^{k+1}-u^k,\xi^{k+1})-(\eta^{k+1}-\eta^k,\xi^{k+1})\\&\quad+\tau\left(b_h(u^{k+1},\xi^{k+1})-b_h(u_h^k,\xi^{k+1})-a_h(\eta^{k+1},\xi^{k+1})\right.\\&\qquad\left.-\varepsilon J_h^\sigma(\eta^{k+1},\xi^{k+1})\right).\end{aligned}\tag{46}$$

In what follows, we estimate the individual terms on the right-hand side of (46).

The Cauchy inequality implies that

$$J_h^\sigma(\eta^k,\xi^k)\le\left(J_h^\sigma(\eta^k,\eta^k)\right)^{1/2}\left(J_h^\sigma(\xi^k,\xi^k)\right)^{1/2},\quad k=0,\dots,r.\tag{47}$$

Lemma 2 *Under assumptions (5), for $t_k,t_{k+1}\in[0,T]$ we have*

$$\left|(u_h^{k+1}-u_h^k,\xi^{k+1})-\tau(u'(t_{k+1}),\xi^{k+1})\right|\le C_{10}\tau^2\|\xi^{k+1}\|_{L^2(\Omega)},\tag{48}$$

$$\|u^{k+1}-u^k\|_{L^2(\Omega)}\le C_{11}\tau,\tag{49}$$

$$|u^{k+1}-u^k|_{H^1(\Omega)}\le C_{12}\tau,\tag{50}$$

$$|u^{k+1}-u^k|_{H^{p+1}(\Omega)}\le C_{13}\tau,\tag{51}$$

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

PROOF.

i) By [18], Lemma 8, we have (48) with $C_{10}=\|u''\|_{L^\infty(0,T;L^2(\Omega))}$, $u''=\partial^2u/\partial t^2$.

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

$$\|u^{k+1}-u^k\|_{L^2(\Omega)} = \left\| \int_{t_k}^{t_{k+1}} u'(t)\,\mathrm{d}t \right\|_{L^2(\Omega)} \le \tau \|u'\|_{L^\infty(0,T;L^2(\Omega))}, \tag{52}$$

which proves (49) 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

$$\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial x_l}\right) = \frac{\partial}{\partial x_l}\left(\frac{\partial u}{\partial t}\right), \qquad l = 1,\dots,d, \tag{53}$$

in the sense of distributions, we obtain

$$\begin{aligned} |u^{k+1}-u^k|_{H^1(\Omega)} &= \|\nabla u^{k+1} - \nabla u^k\|_{L^2(\Omega)} \\ &= \left\| \int_{t_k}^{t_{k+1}} \frac{\partial}{\partial t}\nabla u(t)\,\mathrm{d}t \right\|_{L^2(\Omega)} \\ &\le \int_{t_k}^{t_{k+1}} \left\| \frac{\partial}{\partial t}\nabla u(t) \right\|_{L^2(\Omega)} \mathrm{d}t = \int_{t_k}^{t_{k+1}} \|\nabla u'(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\ &= \int_{t_k}^{t_{k+1}} |u'(t)|_{H^1(\Omega)}\,\mathrm{d}t \le \tau \|u'\|_{L^\infty(0,T;H^1(\Omega))}, \end{aligned} \tag{54}$$

which is (50) with $C_{12} = \|u'\|_{L^\infty(0,T;H^1(\Omega))}$.

iv) Using a similar argumentation as in (54), we derive (51) with $C_{13} = \|u'\|_{L^\infty(0,T;H^{p+1}(\Omega))}$.

□

Lemma 3 *Under assumptions (5), for $t_k, t_{k+1} \in [0,T]$ we have*

$$|(\eta^{k+1}-\eta^k, \xi^{k+1})| \le C_{14}\tau h^{p+1} \|\xi^{k+1}\|_{L^2(\Omega)}, \tag{55}$$

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

PROOF. The Cauchy inequality, relations (41), (36) and (51) imply that

$$\begin{aligned} |(\eta^{k+1}-\eta^k, \xi^{k+1})| &\le \|\eta^{k+1}-\eta^k\|_{L^2(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &= \|\Pi(u^{k+1}-u^k) - (u^{k+1}-u^k)\|_{L^2(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &\le C_6 h^{p+1} |u^{k+1}-u^k|_{H^{p+1}(\Omega)} \|\xi^{k+1}\|_{L^2(\Omega)} \\ &\le C_6 C_{13} \tau h^{p+1} \|\xi^{k+1}\|_{L^2(\Omega)}, \end{aligned} \tag{56}$$

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

Lemma 4 *There exist constants $C_{15} > 0$ and $C_{16} > 0$ independent of u, h, k, ξ and ε such that*

$$|a_h(\eta^k, \xi^k)| \le C_{15}\varepsilon h^p |u^k|_{H^p(\Omega)} \left(J_h^\sigma(\xi^k, \xi^k)^{1/2} + |\xi^k|_{H^1(\Omega,T_h)} \right), \tag{57}$$

$$J_h^\sigma(\eta^k, \eta^k) \le C_{16}\, h^{2p} |u^k|^2_{H^p(\Omega)}, \quad h \in (0, h_0),\ t_k \in [0, T]. \tag{58}$$

PROOF. See [20], Lemma 9. □

Lemma 5 *For $h \in (0, h_0)$, $t_k, t_{k+1} \in [0, T]$ we have*

$$\begin{aligned} &\left| b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \right| \\ &\le C_{17} \left(J_h^\sigma(\xi^{k+1}, \xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)} \right) \\ &\quad \times \left(\|\xi^k\|_{L^2(\Omega)} + h^{p+1} + \tau \right), \end{aligned} \tag{59}$$

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

PROOF. We can write

$$\begin{aligned} &b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \\ &\quad = \ b_h(u^{k+1}, \xi^{k+1}) - b_h(u^k, \xi^{k+1}) \quad (=: \Psi_1) \\ &\quad + \ b_h(u^k, \xi^{k+1}) - b_h(\Pi u^k, \xi^{k+1}) \quad (=: \Psi_2) \\ &\quad + \ b_h(\Pi u^k, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) \quad (=: \Psi_3). \end{aligned} \tag{60}$$

We estimate the individual terms in (60). In virtue of (38),

$$\begin{aligned} |\Psi_1| \le & C_7 \left(J_h^\sigma(\xi^{k+1}, \xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)} \right) \\ & \times \left(\|u^{k+1} - u^k\|_{L^2(\Omega)} + \left(\sum_{i \in I} h_{K_i} \|u^{k+1} - u^k\|^2_{L^2(\partial K_i)} \right)^{1/2} \right). \end{aligned} \tag{61}$$

Using (34), (49) and (50), we find that

$$\begin{aligned} &\sum_{i \in I} h_{K_i} \|u^{k+1} - u^k\|^2_{L^2(\partial K_i)} \\ &\le C_4 \sum_{i \in I} \left(h_{K_i} \|u^{k+1} - u^k\|_{L^2(K_i)} |u^{k+1} - u^k|_{H^1(K_i)} \right. \end{aligned} \tag{62}$$

$$+\|u^{k+1}-u^k\|^2_{L^2(K_i)}\Big) \le C_{18}\tau^2,$$

where $C_{18} := C_4(C_{11}C_{12}h_0 + C_{11}^2)$. Then (61), (62) and (49) give

$$|\Psi_1| \le C_7 \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)}\right) \tau(\sqrt{C_{18}} + C_{11}). \tag{63}$$

Moreover, due to (5), we can set

$$C_{19} = \|u\|_{L^\infty(0,T;H^{p+1}(\Omega))}. \tag{64}$$

From (40) and (39) we deduce that

$$\begin{aligned} |\Psi_2| &\le C_9 h^{p+1} \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)}\right) |u^k|_{H^{p+1}(\Omega)}, \\ |\Psi_3| &\le C_8 \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)}\right) \|\Pi u^k - u_h^k\|_{L^2(\Omega)}. \end{aligned} \tag{65}$$

By (60), (41), (63), (65) and (64),

$$\begin{aligned} &\left|b_h(u^{k+1},\xi^{k+1}) - b_h(u_h^k,\xi^{k+1})\right| \\ &\le C_{17} \left(J_h^\sigma(\xi^{k+1},\xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega,T_h)}\right) \\ &\qquad \times \left(\|\xi^k\|_{L^2(\Omega)} + h^{p+1} + \tau\right), \end{aligned} \tag{66}$$

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 6 *Let assumptions (4), a) - e), (H) and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of the problem satisfying (5). 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 (26) and let $\tau \le 1/2$. Let us set*

$$\begin{aligned} e &= \{e_h^k\}_{k=0}^r = \{u_h^k - u^k\}_{k=0}^r, \\ \|e\|^2_{h,\tau,L^\infty(L^2)} &= \max_{k=0,\dots,r} \|e_h^k\|^2_{L^2(\Omega)}, \\ \|e\|^2_{h,\tau,L^2(H^1)} &= \varepsilon\tau \sum_{k=0}^r \left(|e_h^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(e_h^k,e_h^k)\right), \\ \|e\|^2_{h,\tau,L^\infty(H^1)} &= \varepsilon \max_{k=0,\dots,r} \left(|e_h^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(e_h^k,e_h^k)\right). \end{aligned} \tag{67}$$

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

$$\|e\|^2_{h,\tau,L^\infty(L^2)} \le \tilde{C}\left(h^{2p}\left(\varepsilon + h^2 + h^2/\varepsilon\right) + \tau^2\left(1 + 1/\varepsilon\right)\right), \tag{68}$$

$$\|e\|^2_{h,\tau,L^2(H^1)} \le \hat{C}\left(h^{2p}\left(\varepsilon + h^2 + h^2/\varepsilon\right) + \tau^2\left(1 + 1/\varepsilon\right)\right), \tag{69}$$

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

Moreover, provided

$$h \le C_{IS}\,\tau \tag{70}$$

with a constant C_{IS} *independent of* h *and* τ*, there exists a constant*

$$\bar{C} = O\left(\exp(2T(1 + C_{20}/\varepsilon))\right)$$

such that

$$\begin{aligned}\|e\|^2_{h,\tau,L^\infty(H^1)} \le \bar{C}\Big(&h^{2p-1}\left(1 + \varepsilon + h + h/\varepsilon + h^2 + h^2/\varepsilon + h^2/\varepsilon^2\right)\\ &+\tau\left(1 + 1/\varepsilon + 1/\varepsilon^2\right)\Big)\end{aligned} \tag{71}$$

PROOF. As in (41), we set $\xi^k = u_h^k - \Pi u^k \in S_h$, $\eta^k = \Pi u^k - u^k$, $k = 0, \ldots, r$. Then (42) holds: $e_h^k = u_h^k - u^k = \xi^k + \eta^k$. From (46) and the relations

$$a_h(\xi^{k+1}, \xi^{k+1}) = \varepsilon|\xi^{k+1}|^2_{H^1(\Omega,T_h)} \tag{72}$$

and

$$2(\xi^{k+1} - \xi^k, \xi^{k+1}) = (\|\xi^{k+1}\|^2_{L^2(\Omega)} - \|\xi^k\|^2_{L^2(\Omega)} + \|\xi^{k+1} - \xi^k\|^2_{L^2(\Omega)}), \tag{73}$$

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

$$\begin{aligned}&\|\xi^{k+1}\|^2_{L^2(\Omega)} - \|\xi^k\|^2_{L^2(\Omega)} + \|\xi^{k+1} - \xi^k\|^2_{L^2(\Omega)}\\ &+2\tau\left(\varepsilon|\xi^{k+1}|^2_{H^1(\Omega,T_h)} + \varepsilon J_h^\sigma(\xi^{k+1}, \xi^{k+1})\right)\end{aligned} \tag{74}$$

$$
\begin{aligned}
&= 2\left(\tau(u'(t_{k+1}), \xi^{k+1}) - (u^{k+1} - u^k, \xi^{k+1}) - \left(\eta^{k+1} - \eta^k, \xi^{k+1}\right)\right) \\
&\quad + 2\tau\left(b_h(u^{k+1}, \xi^{k+1}) - b_h(u_h^k, \xi^{k+1}) - a_h(\eta^{k+1}, \xi^{k+1})\right. \\
&\qquad \left. - \varepsilon J_h^\sigma(\eta^{k+1}, \xi^{k+1})\right) =: \text{RHS}.
\end{aligned}
$$

With the aid of Lemmas 2 – 5 , inequality (47), notation (64) and $C_{21} = C_{19}(C_{15} + \sqrt{C_{16}})$ we estimate the right-hand side RHS of (74):

$$
\begin{aligned}
|\text{RHS}| \le\, & 2\left(C_{10}\tau^2 + C_{14}\tau h^{p+1}\right)\|\xi^{k+1}\|_{L^2(\Omega)} \\
& + 2\tau\left(J_h^\sigma(\xi^{k+1}, \xi^{k+1})^{1/2} + |\xi^{k+1}|_{H^1(\Omega, T_h)}\right) \\
& \times \left(C_{21}\varepsilon h^p + C_{17}\left(h^{p+1} + \tau + \|\xi^k\|_{L^2(\Omega)}\right)\right).
\end{aligned} \tag{75}
$$

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

$$
q(\varepsilon, h, \tau) = 2h^{2p}\left(C_{14}^2 h^2 + 4\varepsilon C_{21}^2 + \frac{4}{\varepsilon} C_{17}^2 h^2\right) + \tau^2\left(2C_{10}^2 + \frac{C_{20}}{\varepsilon}\right), \tag{76}
$$

it follows that

$$
\begin{aligned}
&\|\xi^{k+1}\|_{L^2(\Omega)}^2 - \|\xi^k\|_{L^2(\Omega)}^2 + \tau\varepsilon\left(|\xi^{k+1}|_{H^1(\Omega, T_h)}^2 + J_h^\sigma(\xi^{k+1}, \xi^{k+1})\right) \\
&\le \tau\|\xi^{k+1}\|_{L^2(\Omega)}^2 + \tau(1 + C_{20}/\varepsilon)\|\xi^k\|_{L^2(\Omega)}^2 + \tau q(\varepsilon, h, \tau).
\end{aligned} \tag{77}
$$

Hence,

$$
\begin{aligned}
&(1 - \tau)\|\xi^{k+1}\|_{L^2(\Omega)}^2 + \tau\varepsilon\left(|\xi^{k+1}|_{H^1(\Omega, T_h)}^2 + J_h^\sigma(\xi^{k+1}, \xi^{k+1})\right) \\
&\le (1 + \tau C_{20}/\varepsilon)\|\xi^k\|_{L^2(\Omega)}^2 + \tau q(\varepsilon, h, \tau).
\end{aligned} \tag{78}
$$

Moreover, the following inequalities are valid:

$$
\begin{aligned}
\|\xi^k + \eta^k\|_{L^2(\Omega)}^2 &\le 2\left(\|\xi^k\|_{L^2(\Omega)}^2 + \|\eta^k\|_{L^2(\Omega)}^2\right), \\
|\xi^k + \eta^k|_{H^1(\Omega, T_h)}^2 &\le 2\left(|\xi^k|_{H^1(\Omega, T_h)}^2 + |\eta^k|_{H^1(\Omega, T_h)}^2\right), \\
J_h^\sigma(\xi^k + \eta^k, \xi^k + \eta^k) &\le 2(J_h^\sigma(\xi^k, \xi^k) + J_h^\sigma(\eta^k, \eta^k).
\end{aligned} \tag{79}
$$

Now we prove the error estimates (68) – (71).

i) By (78) (using assumption that $0 < \tau \le 1/2$),

$$\|\xi^{k+1}\|^2_{L^2(\Omega)} \le \frac{1+\tau C_{20}/\varepsilon}{1-\tau}\|\xi^k\|^2_{L^2(\Omega)} + \frac{\tau}{1-\tau}q(\varepsilon,h,\tau), \qquad k=0,\dots,r-1. \tag{80}$$

If we set

$$B = \frac{1+\tau C_{20}/\varepsilon}{1-\tau}, \tag{81}$$

then we derive from (80) by induction that

$$\|\xi^k\|^2_{L^2(\Omega)} \le B^k\|\xi^0\|^2_{L^2(\Omega)} + \frac{B^k-1}{B-1}\frac{\tau\, q(\varepsilon,h,\tau)}{1-\tau}, \qquad k=0,\dots,r. \tag{82}$$

By (81),

$$\frac{\tau}{(B-1)(1-\tau)} = \frac{1}{1+C_{20}/\varepsilon} \le 1. \tag{83}$$

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

$$B \le 1 + 2\tau(1+C_{20}/\varepsilon) \le \exp(2\tau(1+C_{20}/\varepsilon)). \tag{84}$$

From (82) – (84) we have

$$\|\xi^k\|^2_{L^2(\Omega)} \le \exp(2\tau k(1+C_{20}/\varepsilon))\left(\|\xi^0\|^2_{L^2(\Omega)} + q(\varepsilon,h,\tau)\right). \tag{85}$$

Further, (26) c), (27) and the Cauchy inequality imply that for each $K \in T_h$,

$$\begin{aligned}\|\xi^0\|^2_{L^2(K)} &= (\Pi^{L^2}u^0 - \Pi u^0, \xi^0)_{L^2(K)} = (u^0 - \Pi u^0, \xi^0)_{L^2(K)} \\ &\le \|u^0 - \Pi u^0\|_{L^2(K)}\|\xi^0\|_{L^2(K)}.\end{aligned} \tag{86}$$

Thus, by (6), (36) and (64),

$$\|\xi^0\|_{L^2(K)} \le \|u^0 - \Pi u^0\|_{L^2(K)} \le C_6 h_K^{p+1}|u^0|_{H^{p+1}(K)}, \tag{87}$$

and

$$\|\xi^0\|^2_{L^2(\Omega)} = \sum_{i\in I}\|\xi^0\|^2_{L^2(K_i)} \le C_6^2 h^{2(p+1)}|u^0|^2_{H^{p+1}(\Omega,T_h)} \le C_{22}h^{2(p+1)}, \tag{88}$$

where $C_{22} = C_6^2 C_{19}^2$. From (85) and (88) we get

$$\|\xi^k\|^2_{L^2(\Omega)} \le \exp(2T(1 + C_{20}/\varepsilon)) \left(C_{22} h^{2(p+1)} + q(\varepsilon, h, \tau)\right), \quad k = 0, \dots, r. \tag{89}$$

Further, by (41), (36) and (64),

$$\|\eta^k\|^2_{L^2(\Omega)} \le C_6^2 h^{2(p+1)} |u^k|^2_{H^{p+1}(\Omega)} \le C_6^2 C_{19}^2 h^{2(p+1)}. \tag{90}$$

Using (67), (79), (89) and (90), we find that

$$\begin{aligned} \|e\|^2_{h,\tau,L^\infty(L^2)} &\le 2 \max_{k=0,\dots,r} \left(\|\xi^k\|^2_{L^2(\Omega)} + \|\eta^k\|^2_{L^2(\Omega)}\right) \\ &\le 2 \exp(2T(1 + C_{20}/\varepsilon)) \left(q(\varepsilon, h, \tau) + C_{23} h^{2(p+1)}\right), \end{aligned} \tag{91}$$

with $C_{23} = C_{22} + C_6^2 C_{19}^2$, which implies estimate (68).

ii) Now let as derive (69). Summing (77) over $k = 0, \dots, r-1$, we get

$$\begin{aligned} &\|\xi^r\|^2_{L^2(\Omega)} + \tau\varepsilon \sum_{k=0}^{r-1} \left(|\xi^{k+1}|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\xi^{k+1}, \xi^{k+1})\right) \\ &\le \tau(1 + C_{20}/\varepsilon) \sum_{k=0}^{r-1} \left(\|\xi^{k+1}\|^2_{L^2(\Omega)} + \|\xi^k\|^2_{L^2(\Omega)}\right) \\ &\quad + T q(\varepsilon, h, \tau) + \|\xi^0\|^2_{L^2(\Omega)}. \end{aligned} \tag{92}$$

This, (88) and (89) imply that

$$\begin{aligned} &\tau\varepsilon \sum_{k=1}^{r} \left(|\xi^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\xi^k, \xi^k)\right) \\ &\le 2T(1 + C_{20}/\varepsilon) \exp(2T(1 + C_{20}/\varepsilon)) \left(C_{22} h^{2(p+1)} + 2q(\varepsilon, h, \tau)\right) \\ &\quad + C_{22} h^{2(p+1)}. \end{aligned} \tag{93}$$

Moreover, as $\xi^0 \in S_h$, we have from (35) and (87) the estimate

$$|\xi^0|_{H^1(K_i)} \le C_5 h_{K_i}^{-1} \|\xi^0\|_{L^2(K_i)} \le C_5 C_6 h_{K_i}^p |u^0|_{H^{p+1}(K_i)}, \quad i \in I, \tag{94}$$

and, by (64),

$$|\xi^0|^2_{H^1(\Omega,T_h)} = \sum_{i \in I} |\xi^0|^2_{H^1(K_i)} \tag{95}$$

$$\leq C_5^2 C_6^2 h^{2p} |u^0|^2_{H^{p+1}(\Omega,T_h)} \leq C_5^2 C_6^2 C_{19}^2 h^{2p}.$$

Furthermore, in virtue of (17), (19), (33), (34), (87) and (94),

$$\begin{aligned} J_h^\sigma(\xi^0,\xi^0) &\leq 4 \sum_{i\in I} \sum_{j\in S(i)} \int_{\Gamma_{ij}} \frac{1}{d(\Gamma_{ij})} |\xi^0|_{\Gamma_{ij}}|^2 \,\mathrm{d}S \\ &\leq 4C_3 \sum_{i\in I} \frac{1}{h_{K_i}} \int_{\partial K_i} |\xi^0|^2 \,\mathrm{d}S \\ &\leq 4C_3\, C_4 \sum_{i\in I} \frac{1}{h_{K_i}} \left(\|\xi^0\|_{L^2(K_i)} |\xi^0|_{H^1(K_i)} + h_{K_i}^{-1} \|\xi^0\|^2_{L^2(K_i)} \right) \\ &\leq 8C_3\, C_4\, C_5^2 \sum_{i\in I} h_{K_i}^{2p} |u^0|^2_{H^{p+1}(K_i)} \\ &\leq 8C_3\, C_4\, C_5^2\, h^{2p} |u^0|^2_{H^{p+1}(\Omega,T_h)} \\ &\leq 8C_3\, C_4\, C_5^2\, C_{19}^2\, h^{2p}. \end{aligned} \tag{96}$$

Then (93), (95) and (96) give

$$\begin{aligned} &\tau\varepsilon \sum_{k=0}^{r} \left(|\xi^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\xi^k,\xi^k) \right) \\ &\leq 2T(1+C_{20}/\varepsilon) \exp(2T(1+C_{20}/\varepsilon)) \left(C_{22} h^{2(p+1)} + 2q(\varepsilon,h,\tau) \right) \\ &\qquad + h^{2p} \left(C_{22} h^2 + \tau\varepsilon C_{24} \right), \end{aligned} \tag{97}$$

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

Taking into account (41), (36), (58) and (64), we obtain

$$\begin{aligned} &\tau\varepsilon \sum_{k=0}^{r} \left(|\eta^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\eta^k,\eta^k) \right) \\ &\leq \tau\varepsilon \left(C_6^2 h^{2p} \sum_{k=0}^{r} |u^k|^2_{H^{p+1}(\Omega,T_h)} + C_{16}\, h^{2p} \sum_{k=0}^{r} |u^k|^2_{H^{p+1}(\Omega)} \right) \\ &\leq \varepsilon C_{25} h^{2p} (T+\tau), \end{aligned} \tag{98}$$

where $C_{25} = C_{19}^2 (C_6^2 + C_{16})$. Finally, using (67), (41), (79), (97) and (98), we arrive at the estimate

$$\begin{aligned} \|e\|^2_{h,\tau,L^2(H^1)} &\leq 2\tau\varepsilon \sum_{k=0}^{r} \Big(|\xi^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\xi^k,\xi^k) \\ &\qquad + |\eta^k|^2_{H^1(\Omega,T_h)} + J_h^\sigma(\eta^k,\eta^k) \Big) \\ &\leq 4T(1+C_{20}/\varepsilon) \exp(2T(1+C_{20}/\varepsilon)) \end{aligned} \tag{99}$$

$$\times \left(2q(\varepsilon,h,\tau) + h^{2p}\left(C_{22}h^2 + \varepsilon(C_{24}+C_{25})\right)\right).$$

Now, assertion (69) of the theorem follows from (76) and (99).

iii) Finally, let $h \le \tau$. As $\tau \le 1/2$, (78) implies that

$$\begin{aligned} &\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,T_h)} + J^\sigma_h(\xi^{k+1},\xi^{k+1})\right) \\ &\quad \le \left(\frac{1}{\tau} + \frac{C_{20}}{\varepsilon}\right)\|\xi^k\|^2_{L^2(\Omega)} + q(\varepsilon,h,\tau). \end{aligned} \tag{100}$$

Using (89), we obtain

$$\begin{aligned} &\varepsilon\left(|\xi^{k+1}|^2_{H^1(\Omega,T_h)} + J^\sigma_h(\xi^{k+1},\xi^{k+1})\right) \\ &\quad \le \left(\frac{1}{\tau} + \frac{C_{20}}{\varepsilon}\right)\exp(2T(1+C_{20}/\varepsilon))\left(C_{22}h^{2(p+1)} + q(\varepsilon,h,\tau)\right) \\ &\quad +q(\varepsilon,h,\tau). \end{aligned} \tag{101}$$

Hence, in virtue of (41), (36), (58) and (64),

$$\varepsilon\left(|\eta^k|^2_{H^1(\Omega,T_h)} + J^\sigma_h(\eta^k,\eta^k)\right) \le \varepsilon C_{25}h^{2p}, \tag{102}$$

where $C_{25} = C_{19}^2(C_6^2 + C_{16})$. Now, according to (67), (41), (79), (101) and (102) we arrive at

$$\begin{aligned} \|e\|^2_{h,\tau,L^\infty(H^1)} &\le \max_{k=0,\dots,r} 2\varepsilon\left(|\xi^k|^2_{H^1(\Omega,T_h)} + J^\sigma_h(\xi^k,\xi^k)\right. \\ &\qquad \left. +|\eta^k|^2_{H^1(\Omega,T_h)} + J^\sigma_h(\eta^k,\eta^k)\right) \\ &\le 2(\varepsilon+\tau C_{20})\exp(2T(1+C_{20}/\varepsilon)) \\ &\quad \times\left(\frac{C_{22}h^{2(p+1)} + q(\varepsilon,h,\tau)}{\varepsilon\tau} + q(\varepsilon,h,\tau) + C_{25}h^{2p}\right). \end{aligned} \tag{103}$$

Finally, this, (76) and assumption (70) yield (71). □

Remark 7 *Estimate (68) implies that*

$$\|u-u_h\|_{L^\infty(0,T;L^2(\Omega))} = O(h^p+\tau) \quad \textit{for } h \to 0+. \tag{104}$$

Comparing this result with the approximation property (36) implying that

$$\|u-\Pi u\|_{L^\infty(0,T;L^2(\Omega))} = O(h^{p+1}), \tag{105}$$

we see that the error estimate (68) is suboptimal with respect to h. There is a question, whether this estimate can be improved. Numerical experiments carried out in [17] indicate that the actual order of convergence in the $L^\infty(L^2)$-norm in the case of odd degree of approximation (in [17] $p = 1$ and $p = 3$) is better than the theoretically derived estimate. (Similar experimental results can be found in [27] and [25] for stationary elliptic problems without convection.) In the next section we shall present numerical experiments with piecewise linear elements ($p = 1$) on noncomforming meshes with nonconvex elements.

It is clear that estimates (68) – (71) 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 [24].

Estimate (71) in $L^\infty(H^1)$-norm has been obtained under assumption (70), i.e. $h \le C_{IS}\,\tau$. This nonstandard "inverse stability condition" also appears in [30] and [22].

Remark 8 *Estimates (68) – (71) were derived without any restriction on the time step τ. It is possible to show that for a fixed $\varepsilon > 0$ the semi-implicit scheme is unconditionally stable. However, there is a natural question what happens, when $\varepsilon \to 0$ and in the limit we obtain an explicit scheme for a nonlinear conservation law. Its stability requires the use of a CFL condition limiting the length of the time. Our results are not in contradiction with this fact, because, due to Remark 7, the error estimate blows up and the scheme may loose the unconditional stability for $\varepsilon \to 0$. (This is also confirmed by numerical experiments in Section 6.3.)*

6 Numerical experiments

In this section we shall verify the theoretical error estimates, presented in the previous section for general grids having properties (A1) and (A2). Namely, we consider nonconforming meshes containing triangular and nonconvex quadrilateral elements constructed by the following algorithm:

(1) We start from a vertically oriented structured triangular grid, see Figure 2, a).
(2) We aplly a vertical shift to some vertices, which creates a triangular mesh with handing nodes, shown in Figure 2, b).
(3) We apply a horizontal shift to some vertices, which creates nonconvex quadrilaterals in Figure 2, c).

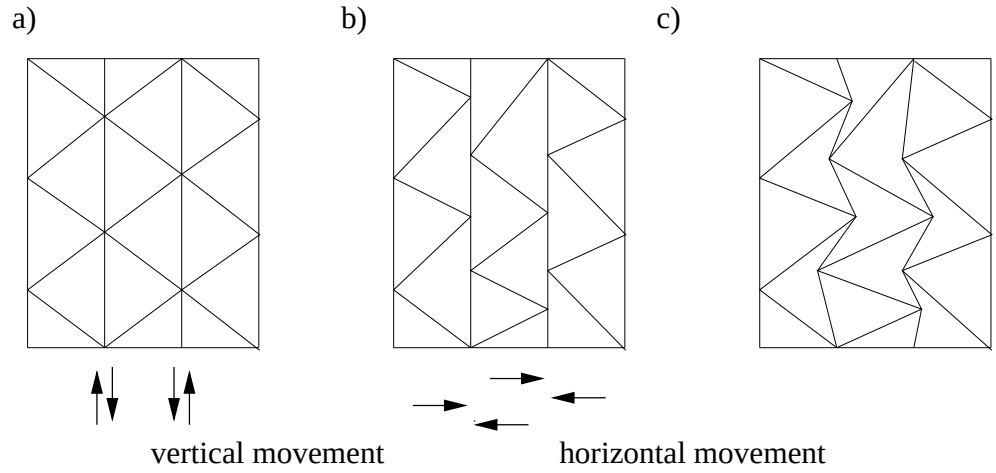


Fig. 2. Algorithm generating grids with nonconvex quadrilateral elements

This algorithm allows us to construct meshes with a prescribed constant C_3 from (33), characterizing the nonconformity of the mesh. Figure 3 shows grids with different numbers $\#T_h$ of elements and different values of C_3.

These types of grids are artificial and not used in practice, of course. We only want to demonstrate that our scheme is robust with respect to rather rough meshes. Numerical calculations on conforming triangular meshes were carried out in [20] and [17]. On the other hand, it is possible to meet grids with nonconvex elements, if the DGFEM is considered as a generalization of the so-called dual finite volumes (see, e.g. [22]). One can also meet such meshes in the process of a mesh generation, particularly in 3D.

We solve the 2D viscous Burgers equation

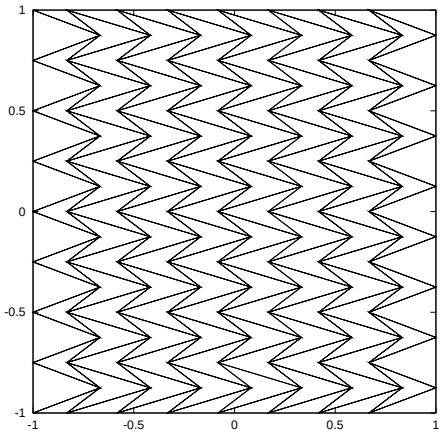
$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x_1} + u\frac{\partial u}{\partial x_2} = \varepsilon \Delta u + g \quad \text{in } \ \Omega \times (0,T), \tag{106}$$

where $\Omega = (-1,1)^2, T = 1$, equipped with the boundary condition (2) and the initial condition (3). In the definition (15) of the form b_h we use the numerical flux

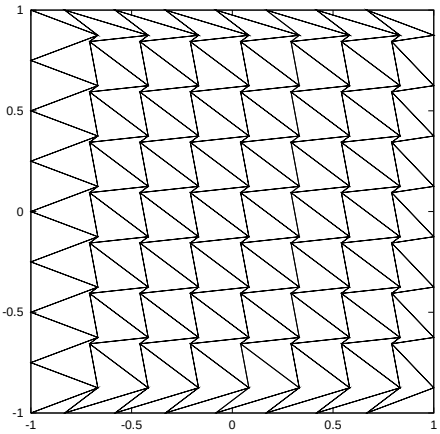
$$H(u_1,u_2,\boldsymbol{n}) = \begin{cases} \sum_{s=1}^{2} f_s(u_1)n_s, & \text{if } A > 0 \\ \sum_{s=1}^{2} f_s(u_2)n_s, & \text{if } A \le 0 \end{cases}, \tag{107}$$

where

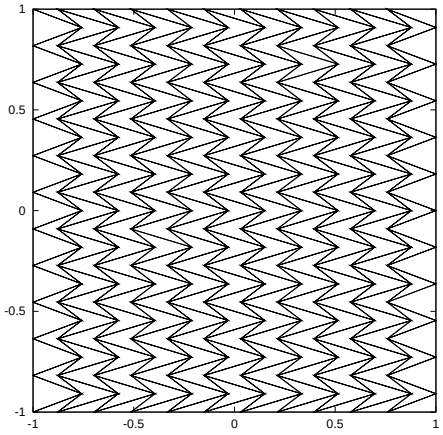
$$A = \sum_{s=1}^{2} f_s'(\bar{u})n_s, \quad \bar{u} = \frac{1}{2}(u_1+u_2) \ \text{ and } \ \boldsymbol{n} = (n_1,n_2). \tag{108}$$



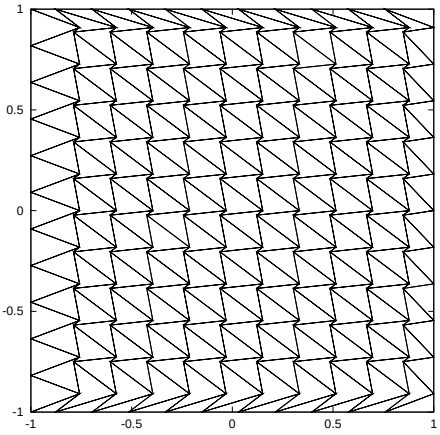
$\#T_h = 136,\ C_3 = 2.094$



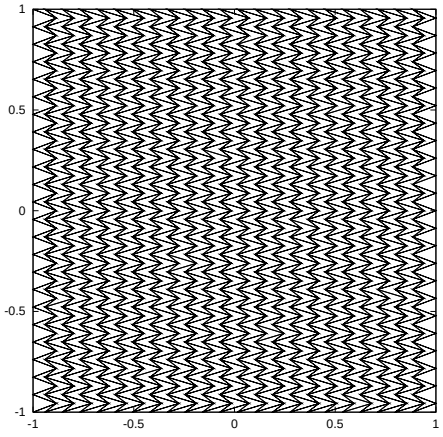
$\#T_h = 136,\ C_3 = 8.375$



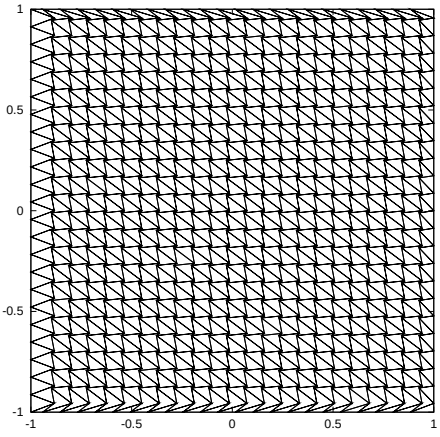
$\#T_h = 253,\ C_3 = 2.094$



$\#T_h = 253,\ C_3 = 8.375$



$\#T_h = 1081,\ C_3 = 2.094$



$\#T_h = 1081,\ C_3 = 8.375$

Fig. 3. Examples of grids formed by triangular and nonconvex quadrilateral elements with different numbers $\#T_h$ of elements and different values of C_3

One can see that H satisfies conditions (21) and (22) and is Lipschitz-continuous on any bounded subset of $\mathbb{R}^2$. For boundary edges Γ_{ij}, $j \in \gamma(i)$, in the definition of the convective form b_h, we use the treatment of boundary conditions based on upwinding. This means that the second term in (15) is replaced by

$$\sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, v|_{\Gamma_{ij}}\, \mathrm{d}S, \tag{109}$$

where

$$u|_{\Gamma_{ji}} = \begin{cases} u|_{\Gamma_{ij}}, & \text{if } \sum_{s=1}^{2} f_s'(u|_{\Gamma_{ij}})(\boldsymbol{n}_{ij})_s \geq 0, \\ u_D|_{\Gamma_{ij}}, & \text{otherwise.} \end{cases}, \tag{110}$$

Here $(\boldsymbol{n}_{ij})_s$ 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 the function u_D from the boundary condition (2) on Γ_{ij}.

All computations were carried out with piecwise linear elements and a constant time step $\tau > 0$. Volume integrals over elements $K \in T_h$ are evaluated in such a way that these elements are divided into triangles and then a quadrature formula exact for polynomials of degree ≤ 6 is applied. Line integrals are computed with the aid of a formula exact for polynomials of degree ≤ 7.

6.1 Convergence with respect to τ

First, we verify experimentally the convergence of the method in $L^2(\Omega)$-norm with respect to the time step $\tau \to 0+$. In order to restrain the discretization errors with respect to h, we use a fine mesh with 4095 triangles and $C_3 = 2.094$.

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

$$u(x_1, x_2, t) = \frac{e^{10t} - 1}{e^{10} - 1} \hat{u}(x_1, x_2), \tag{111}$$

where

$$\hat{u}(x_1, x_2) = (1 - x_1^2)^2 (1 - x_2^2)^2 \tag{112}$$

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 boundary $\partial\Omega$, see Figure 4.

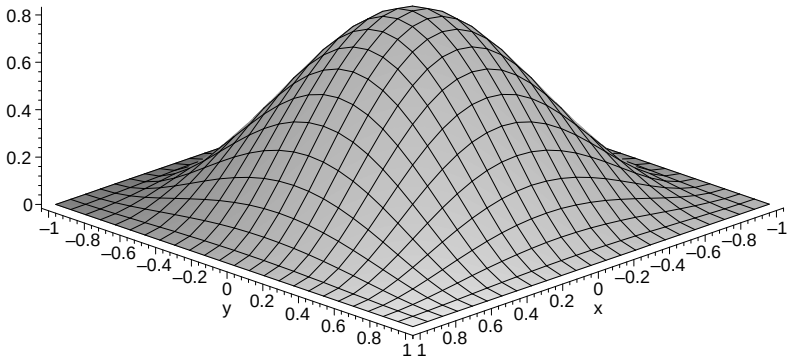


Fig. 4. Exact solution (111) at $t = 1$

The computations were carried out with 6 different time steps τ_l, see Table 1. The computational error is evaluated at time $T = 1$ in $L^2(\Omega)$–norm, i.e. we set

$$e_\tau \equiv \|u_\tau(\cdot, T) - u(\cdot, T)\|_{L^2(\Omega)}, \tag{113}$$

where $u(\cdot, T)$ is the exact solution given by (111) – (112) at time T and $u_\tau(\cdot, T)$ is the numerical solution at time T obtained by scheme (26), a)–c) with time step τ. We suppose that the error behaves according to the formula

$$e_\tau \approx D\tau^\alpha, \tag{114}$$

where $D > 0$ is a constant independent of τ and α is the *order of accuracy* of the method in $L^2(\Omega)$-norm. We define the *local experimental order of convergence* by

$$\alpha_l = \frac{\log\left(e_{\tau_l}/e_{\tau_{l-1}}\right)}{\log\left(\tau_l/\tau_{l-1}\right)}, \quad l = 2, \ldots, 6. \tag{115}$$

The *global experimental order of convergence* $\bar{\alpha}$ is obtained by the least squares method. Table 1 shows the errors e_τ, the values of α_l for $l = 2, \ldots, 6$ and $\bar{\alpha}$.

6.2 Convergence with respect to h

Now we verify numerically the convergence of errors in $L^2(\Omega)$-norm with respect to the mesh size $h \to 0+$. In order to overkill the discretization errors with respect to τ, we use the time step $\tau = 10^{-4}$. Numerical experiments indicate that this choice is sufficient. Smaller τ does not cause any further decrease of computational errors.

Table 1
Time steps $\tau_l,\ l = 1, \ldots, 6$, computational error e_τ, values of $\alpha_l,\ l = 2, \ldots, 6$, and $\bar{\alpha}$

		$L^2(\Omega)$–norm	
l	τ_l	e_τ	α_l
1	5.000E-03	2.0182E-02	-
2	2.500E-03	1.0156E-02	0.991
3	1.250E-03	5.1311E-03	0.985
4	6.250E-04	2.6506E-03	0.953
5	3.125E-04	1.3780E-03	0.944
6	1.563E-04	7.2245E-04	0.932
global order $\bar{\alpha}$			0.961

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

$$u(x_1, x_2, t) = \left(1 - \frac{e^{-t}}{2}\right) \hat{u}(x_1, x_2), \tag{116}$$

where $\hat{u}$ is given by (112) and $\varepsilon = 0.1$.

We solve the problem in consideration with the aid of piecewise linear elements. The computations were carried out on 6 grids $T_{h_l},\ l = 1, \ldots, 6$, having different number of elements and different parameter C_3, see Table 2. Some of the meshes are shown in Figure 3.

The computational error of the solution is evaluated at time $T = 1$ in $L^2(\Omega)$–norm:

$$e_h \equiv \|u_h(\cdot, T) - u(\cdot, T)\|_{L^2(\Omega)}, \tag{117}$$

where $u(\cdot, T)$ is the exact solution of equation (106) given by (116) and (112) at time T and $u_h(\cdot, T)$ is the numerical solution at time T obtained by scheme (26), a)–c). We suppose that the error behaves according to the formula

$$e_h \approx D h^\alpha, \tag{118}$$

where $h = \max_{K \in T_h} h_K$, $D > 0$ is a constant independent of h and α is the *order of accuracy* of the method. We define the *local experimental order of*

convergence by

$$\alpha_l = \frac{\log\left(e_{h_l}/e_{h_{l-1}}\right)}{\log\left(h_l/h_{l-1}\right)}, \quad l = 2, \ldots, 6, \tag{119}$$

where $h_l = \max_{K \in T_{h_l}} h_K,\ l = 1, \ldots, 6$. The *global experimental order of convergence* $\bar{\alpha}$ is obtained by the least squares method.

Table 2 shows the computational error e_h and experimental order of convergence. Moreover, Table 3 shows the error e_h obtained on meshes with 528 elements and different values of the parameter C_3. We see that the dependence of the error on C_3 is not important.

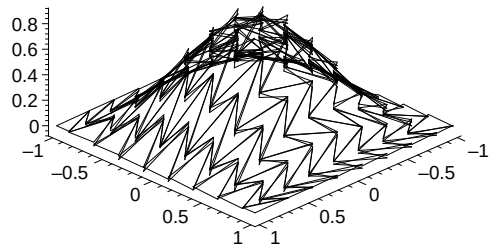
Table 2
Error e_h and experimental order of convergence for grids with $C_3 = 2.094$ and $C_3 = 8.375$

			$C_3 = 2.094$		$C_3 = 8.375$	
l	$\#T_{h_l}$	h_l	e_h	α_l	e_h	α_l
1	136	4.334E-01	1.9775E-02	-	1.5393E-02	-
2	253	3.152E-01	1.0404E-02	2.017	7.9873E-03	2.060
3	528	2.167E-01	4.9109E-03	2.004	3.6525E-03	2.088
4	1081	1.508E-01	2.3905E-03	1.986	1.7223E-03	2.073
5	2080	1.084E-01	1.2450E-03	1.976	8.8145E-04	2.029
6	4095	7.705E-02	6.1307E-04	2.075	4.3596E-04	2.062
		$\overline{\alpha}$		2.005		2.064

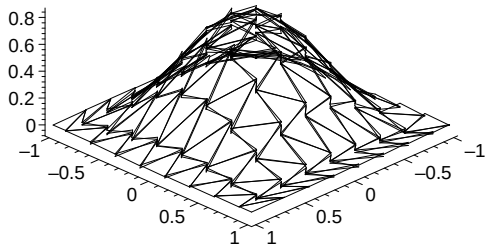
Table 3
Dependence of the computational error e_h on the value of C_3 for $\#T_h = 528$

l	C_3	e_h
1	2.094	4.9109E-03
2	4.188	3.8514E-03
3	8.375	3.6525E-03
4	16.106	3.6403E-03
5	29.911	3.6550E-03
6	52.344	3.6676E-03

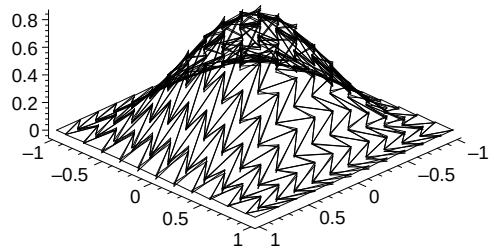
Figure 5 shows the numerical solution obtained on grids with $\#T_h = 136$, 253, 1081 and $C_3 = 2.094,\ 8.375$. We observe here the convergence of the approximate solution to the exact solution for $h \to 0$.



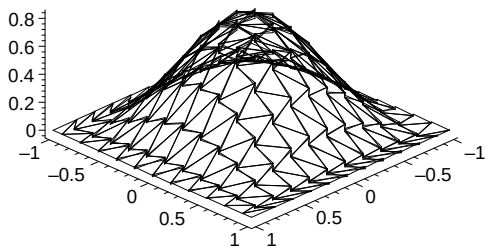
$\#T_h = 136,\ C_3 = 2.094$



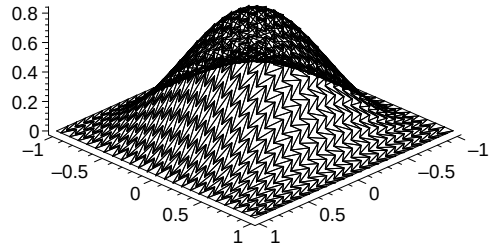
$\#T_h = 136,\ C_3 = 8.375$



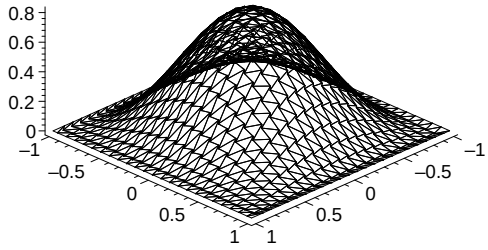
$\#T_h = 253,\ C_3 = 2.094$



$\#T_h = 253,\ C_3 = 8.375$



$\#T_h = 1081,\ C_3 = 2.094$



$\#T_h = 1081,\ C_3 = 8.375$

Fig. 5. Numerical solution on grids with nonconvex quadrilateral elements with different numbers of elements $\#T_h$ and different values of C_3

6.3 Stability of the scheme for $\varepsilon \to 0$

In virtue of Remark 8, we solve the problem from Section 6.1 with decreasing diffusion coefficient ε and investigate the stability behaviour of the semi-implicit scheme (26), a)–c). We use a fixed triangular grid having 1056 elements ($h = 6.37 \cdot 10^{-2}$) and carry out the computation with different time steps τ and different diffusion coefficients ε. Table 4 shows the computational

error $e_\tau = \|u_h(\cdot,T) - u(\cdot,T)\|_{L^2(\Omega)}$ $(T = 1)$ in depedence on τ and ε. We see that for a fixed τ the error increases with a decreasing ε, which corresponds to the obtained theoretical error estimate. On the basis of this experiment we can conclude that for a decreasing ε the unconditional stability becomes weaker and weaker and in the limit for $\varepsilon \to 0$ we get a conditionally stable method.

Table 4
Error e_τ in dependence on τ and ε

	ε			
τ	0.0001	0.001	0.01	0.1
0.01250000	divergence	divergence	0.0349157	0.0018731
0.00937500	divergence	0.0613545	0.0251048	0.0016094
0.00867187	0.0644484	0.0571331	0.0229959	0.0015743
0.00625000	0.0499788	0.0409237	0.0158801	0.0015312
0.00312500	0.0253018	0.0197805	0.0074850	0.0016554
0.00156250	0.0122522	0.0096823	0.0039342	0.0017804
0.00078125	0.0066870	0.0055137	0.0027518	0.0018550
0.00039125	0.0051225	0.0043855	0.0025410	0.0018949

7 Conclusion

We have presented an efficient numerical method for the solution of nonstationary nonlinear convection-diffusion problems, which is based on the space discretization by the discontinuous Galerkin finite element method and a semi-implicit time discretization and applied on nonstandard, nonconforming meshes. We have derived a priori error estimates of order $O(h^p + \tau)$ in $L^\infty(0,T;L^2(\Omega))$–norm , $L^2(0,T;H^1(\Omega))$–seminorm and $L^\infty(0,T;H^1(\Omega))$–seminorm. The presented numerical examples indicate a better behaviour of the experimental $L^\infty(L^2)$–order of convergence for piecewise linear approximations, in spite nonconforming meshes with nonconvex elements are used. The obtained results confirm that the DGFEM is a powerful and reliable method for the numerical solution of nonstationary nonlinear convection-diffusion problems, which is sufficiently accurate and robust even on unstructured nonconforming meshes with nonconvex elements and hanging nodes.

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,
– increase of accuracy in the time discretization,
– stability analysis of the method for $\varepsilon \to 0+$,
– development of DGFE schemes for the numerical simulation of compressible flow with a wide range of Reynolds and Mach numbers.

References

[1] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. 39 (2002) 1749–1779.

[2] I. Babuška, M. Zlámal, Nonconforming elements in the finite element method with penalty, SIAM J. Numer. Anal. 10 (1973) 863–875.

[3] I. Babuška, C. E. Baumann, J. T. Oden, A discontinuous *hp* finite element method for diffusion problems: 1-d analysis, Computers and Mathematics with Applications 37 (1999) 103–122.

[4] F. Bassi, S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations, J. Comput. Phys. 131 (1997) 267–279.

[5] C. E. Baumann, 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.

[6] S. Brenner, R. L. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.

[7] F. Brezzi, G. Manzini, D. Marini, P. Pietra, A. Russo, Discontinuous Galerkin approximations for elliptic problems, Numer. Methods Partial Differential Eq 16 (2000) 35-378.

[8] Z. Chen, On the relationship of various discontinuous finite element methods for second-order elliptic equations, East-West Numer. Math. 9 (2001) 99-122.

[9] Z. Chen, *Finite Element Methods and Their Applications*, ???? Springer, Heidelberg, 2005.

[10] H. Chen, Z. Chen, Stability and convergence of mixed discontinuous finite element methods for second-order differential problems, J. Numer. Math., 11 (2003) 253-287.

[11] P. G. Ciarlet, *The Finite Elements Method for Elliptic Problems*, North-Holland, Amsterdam, 1979.

[12] B. Cockburn, Discontinuous Galerkin methods for convection dominated problems. In T. J. Barth and H. Deconinck (eds.) *High–Order Methods*

for Computational Physics, Lecture Notes in Computational Science and Engineering 9, Springer, Berlin, 1999, pp. 69–224.

[13] B. Cockburn, G. E. Karniadakis, C.-W. Shu (eds.), *Discontinuous Galerkin Methods. Theory, Computation and Applications*. Lecture Notes in Computational Science and Engineering 11, Springer, Berlin, 2000.

[14] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM J. Numer. Anal. 35 (1998) 2440–2463.

[15] C. Dawson, V. Aizinger, *A discontinuous Galerkin method for three-dimensional shallow water equations*, J. Sci. Comput. 22-23 (2005) 245–267.

[16] V. Dolejší, On the discontinuous Galerkin method for the numerical solution of the Navier–Stokes equations, Int. J. Numer. Meth. Fluids 10 (2004) 1083–1106.

[17] V. Dolejší, M. Feistauer, Error estimates of the discontinuous Galerkin method for nonlinear nonstationary convection-diffusion problems. Numer. Funct. Anal. Optim. 26(25-26) (2005) 2709-2733.

[18] V. Dolejší, M. Feistauer, J. Felcman, 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.

[19] V. Dolejší, M. Feistauer, C. Schwab, A finite volume discontinuous Galerkin scheme for nonlinear convection–diffusion problems, Calcolo 39 (2002) 1–40.

[20] V. Dolejší, M. Feistauer, V. Sobotíková, A discontinuous Galerkin method for nonlinear convection–diffusion problems, Comput. Methods Appl. Mech. Engrg. 194 (2005) 2709–2733.

[21] M. Feistauer, *Mathematical Methods in Fluid Dynamics*, Longman Scientific & Technical, Harlow, 1993.

[22] M. Feistauer, J. Felcman, M. Lukáčová-Medvidová, 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.

[23] M. Feistauer, J. Felcman, I. Straškraba, *Mathematical and Computational Methods for Compressible Flow*, Clarendon Press, Oxford, 2003.

[24] M. Feistauer, K. Švadlenka, Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems, J. Numer. Math. 12 (2004) 97-118.

[25] P. Houston, C. Schwab, E. Süli, Discontinuous *hp*-finite element methods for advection-diffusion problems, SIAM J. Numer. Anal. 39 (2002) 2133–2163.

[26] A. Kufner, O. John, S. Fučík, *Function Spaces*, Academia, Prague, 1977.

[27] J. T. Oden, I. Babuška, C. E. Baumann, A discontinuous *hp* finite element method for diffusion problems, J. Comput. Phys 146 (1998) 491–519.

[28] S. Prudhomme, F. Pascal, J.T. Oden, A. Romkes, Review of a priori error estimation for discontinuous Galerkin methods, Technical report, Laboratoire de Mathématiques, Université Paris-Sud, 2002.

[29] B. Rivière, M. F. Wheeler, A discontinuous Galerkin method applied to nonlinear parabolic equations, In B. Cockburn (editor), *Discontinuous Galerkin Methods. Theory, Computation and Applications.*, Lecture Notes in Computational Science and Engineering 11, Springer, Berlin, 2000, pp. 231–244.

[30] H.-G. Roos, M. Stynes, L. Tobiska, *Numerical Methods for Singularly Perturbed Differential Equations*, Springer Series in Computational Mathematics 24, Springer, Berlin, 1996.

[31] J.J.W van der Vegt, H. Van der Ven, Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flow, part I. General formulation, J. Comput. Phys. 182 (2002) 546–585.

[32] R. Verfürth, A note on polynomial approximation in Sobolev spaces. M^2AN 33 (1999) 715–719.