Optimality of error estimates in the combined finite volume-finite element and discontinuous Galerkin methods for nonlinear convection-diffusion problems *

Miloslav Feistauer
Charles University Prague
Faculty of Mathematics and Physics

*Presented at the conference FVCA08, Aussois 2008

1994-2001 M.F., J. Felcman, M. Lukáčová, V. Dolejší, P. Angot, G. Warnecke and A. Kliková developed combined FV-FE method for the solution of nonlinear convection-diffusion problems and compressible viscous flow

several combinations:

a) dual finite volumes over a triangular mesh - conforming P^1 finite elements

b) barycentric finite volumes over a triangular mesh - Crouzeix-Raviart finite elements

c) triangular FV's-conforming FE's

Theoretical analysis obtained in the cases a), b) in 1995-2001

c) gives best computational results for compressible viscous flow $\times$ analysis unsuccessful till 2007 - M.F., M. Bejček, T. Gallouët, R. Herbin, J. Hájek

Continuous problem

$\Omega \subset I\!R^2$ - **bounded polygonal domain**

$(0, T)$, $T > 0$, - **time interval**

Initial-boundary value problem:

$$\frac{\partial u}{\partial t} + \sum_{s=1}^{2} \frac{\partial f_s(u)}{\partial x_s} = \varepsilon\, \Delta u + g \quad \textbf{in } Q_T = \Omega \times (0, T), \qquad (1)$$

initial and boundary conditions

$$u(x, 0) = u^0(x), \;\; x \in \Omega, \qquad u|_{\partial\Omega \times (0,T)} = 0. \qquad (2)$$

Assumptions on data:

a) $f_s \in C^1(I\!R)$, $f_s(0) = 0$, $|f_s'| \leq C_{f'}$ $s = 1, 2$,

b) $\varepsilon > 0$,

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

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

Notation:

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

$$a(u,v) = \varepsilon \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x \ ,$$

$$|u|_{H^1(\Omega)} = \left(\int_\Omega |\nabla u|^2 \mathrm{d}x \right)^{1/2}.$$

Combined FV-FE method

$\mathcal{T}_h = \{K_i\}_{i\in I}$ and $\mathcal{D}_h = \{D_i\}_{i\in J}$ ($I, J \subset Z^+ = \{0, 1, 2, \dots\}$ - suitable index sets): two triangulations of the domain Ω satisfying the standard assumptions from the FE method

Triangles from $\mathcal{T}_h$ = finite elements

Triangles from $\mathcal{D}_h$ = finite volumes

If two finite volumes $D_i, D_j \in \mathcal{D}_h$ have a common side, we call them neighbours.

Then we use the notation $\Gamma_{ij} = \Gamma_{ji} = \partial D_i \cap \partial D_j$

$s(i) = \{j \in J; j \neq i, D_j$ is a neighbour of $D_i\}$.

$|\Gamma_{ij}|$ - **length of the side** Γ_{ij}

$\boldsymbol{n}_{ij}$ - **unit outer normal to** ∂D_i **on the side** Γ_{ij}

For $k \in Z^+, K \in \mathcal{T}_h$ **we denote by** $P^k(K)$ **the space of all polynomials on** K **of degree** $\leq k$**.**

Finite element spaces:

$$X_h = \{v_h \in C(\overline{\Omega});\ v_h|_K \in P^1(K)\ \forall K \in \mathcal{T}_h\},$$

$$V_h = \{v_h \in X_h;\ v_h|_{\partial\Omega} = 0\}$$

Finite volume space:

$$Y_h = \{v_h \in L^2(\Omega);\ v_h|_{D_i} \in P^0(D_i)\ \forall i \in J\}$$

The relation between the FE and FV spaces - *lumping operator* L_h:

$$L_h v|_{D_i} = \frac{1}{|D_i|} \int\limits_{D_i} v \, \mathrm{d}x, \qquad i \in J. \tag{3}$$

Discrete problem

Let u be an exact sufficiently regular solution. Then

$$\left(\frac{\partial u}{\partial t}, v\right) + \sum_{i \in J} \int_{D_i} \sum_{s=1}^{2} \frac{\partial f_s(u)}{\partial x_s} v \,\mathrm{d}x + a(u,v) = (g,v) \;\forall v \in V_h. \tag{4}$$

The terms with fluxes f_s approximated with the aid of numerical flux H:

$$b_h(u,v) = \sum_{i \in J} L_h v|_{D_i} \sum_{j \in s(i)} H(L_h u|_{D_i}, L_h u|_{D_j}, \boldsymbol{n}_{ij})\,|\Gamma_{ij}|. \tag{5}$$

1) $H(u,v,\boldsymbol{n})$ **is defined for** $u,v \in I\!R$ **and** $\boldsymbol{n} \in B_1 = \{\boldsymbol{n} \in I\!R^2; |\boldsymbol{n}| = 1\}$**, and** *Lipschitz-continuous* **with respect to** u, v.

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

$H(u,u,\boldsymbol{n}) = \sum_{s=1}^{2} f_s(u)\, n_s, \quad u \in I\!R, \;\; \boldsymbol{n} = (n_1, n_2) \in B_1$.

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

$H(u,v,\boldsymbol{n}) = -H(v,u,-\boldsymbol{n})$,

$u,\ v \in I\!R, \;\; \boldsymbol{n} \in B_1$.

Approximate solution:

$u_h \in C^1([0,T];\ V_h)$,

$$a) \quad \left(\frac{\partial u_h}{\partial t}, v_h\right) + b_h(u_h, v_h) + a(u_h, v_h) = (g, v_h) \quad (6)$$

$$\forall v_h \in V_h,$$

$$b) \quad u_h(0) = u_h^0 = \Pi_h u^0 = X_h - \text{interpolation}$$

Theoretical analysis

Consider systems $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ of finite element meshes and $\{\mathcal{D}_h\}_{h\in(0,h_0)}$ of finite volume meshes of the domain Ω, with $h_0 > 0$

Notation:

$h_K = \mathrm{diam}(K)$, $h = \max_{K\in\mathcal{T}_h} h_K$

$\rho_K =$ **the radius of the largest circle inscribed into the element $K \in \mathcal{T}_h$**

Let

$$\frac{h_K}{\rho_K} \le C_R, \qquad h \le C_I h_K,$$
$$\mathrm{diam}(D_i) \le C_D\, h,$$
$$K \in \mathcal{T}_h, \;\; h \in (0, h_0), \;\; i \in J,$$

with constants $C_R, C_I, C_D > 0$ independent of K, h and i.

Let us set

$\omega(D_i) = \cup \{K \in \mathcal{T}_h;\ \ K \cap D_i \neq \emptyset\}$

For a given element $K \in \mathcal{T}_h$, let R_K be the number of sets $\omega(D_i)$ containing the element K.

We assume that there exists $R < +\infty$, independent of h, such that $R_K \leq R$ for any $K \in \mathcal{T}_h$.

$\Longrightarrow$ each element $K \in \mathcal{T}_h$ intersects at most R finite volumes D_i. Then

$$\sum_{i \in J} |v_h|^2_{H^1(\omega(D_i))} \leq R\, |v_h|^2_{H^1(\Omega)}. \tag{7}$$

The analysis of the FV-FE method carried out under the assumption

$\frac{\partial u}{\partial t} \in L^2(0,T; H^2(\Omega))$.

Under the above assumptions the error of the method $e_h = u - u_h$ satisfies the estimate

$$\max_{t\in[0,T]} \|e_h\|_{L^2(\Omega)} \leq C\, h, \;\; \sqrt{\varepsilon} \sqrt{\int_0^T |e_h(\vartheta)|^2_{H^1(\Omega)} \mathrm{d}\vartheta} \leq C\, h. \tag{8}$$

Optimality of the error estimates

Verification on the example of the 2D viscous Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x_1} + u\frac{\partial u}{\partial x_2} - \varepsilon \Delta u = g, \tag{9}$$

considered in the space-time domain $Q_T = \Omega \times (0,1)$, $\Omega = (-1,1)^2$,

equipped with such data that the exact solution has the form

$u = (1 - e^{-2t})(1 - x_1^2)^2(1 - x_2^2)^2$ **and** $\varepsilon = 0.1$

The time discretization is carried out by a semi-implicit Euler scheme

$$\left(\frac{u_h^k - u_h^{k-1}}{\tau}, v_h\right) + b_h(u_h^{k-1}, v_h) + a_h(u_h^k, v_h) = (g^{k-1}, v_h), \tag{10}$$

with overkill in time, i.e. with very small time step.

Numerical flux:

$H(u_1, u_2, \boldsymbol{n}) = \sum_{s=1}^{2} f_s(u_1) n_s$, if $A > 0$

and

$H(u_1, u_2, \boldsymbol{n}) = \sum_{s=1}^{2} f_s(u_2) n_s$, if $A \leq 0$,

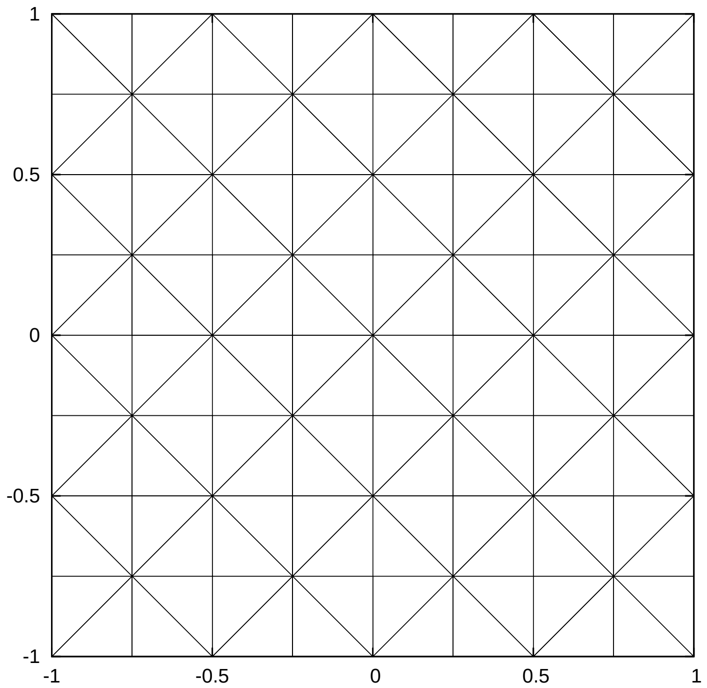
where

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

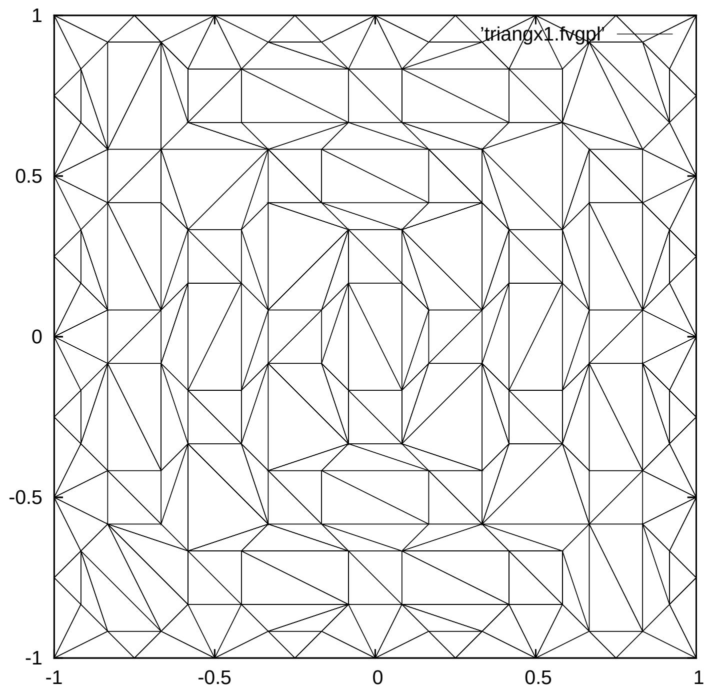
We consider successively refined FE meshes $\mathcal{T}_{h_i}$, $i = 1, \ldots, 6$, and two types of secondary FV meshes

a) $\mathcal{D}_h^1 = \mathcal{T}_h$,

b) $\mathcal{D}_h^2$ defined as the Delaunay triangulation with vertices equal to the barycenters of the FE triangles from $\mathcal{T}_h$ and the FE boundary vertices. In Figures the triangulations $\mathcal{T}_{h_1}$ and $\mathcal{D}_{h_1}^2$ are shown.



Triangulation $\mathcal{T}_{h_1}$



Triangulation $\mathcal{D}_{h_1}^2$

Tables:

computational results obtained with the aid of the FV meshes $\mathcal{D}^1_{h_i}$ and $\mathcal{D}^2_{h_i}$

$e_{h,L^\infty(L^2)}$ - computational errors evaluated in the $L^\infty(L^2)$-norm

$e_{h,L^2(H^1)}$ - computational errors evaluated in the $L^2(H^1)$-norm

$\mathrm{EOC}_{L^\infty(L^2)}$, $\mathrm{EOC}_{L^2(H^1)}$ - corresponding experimental orders of convergence

$\#I$	h	$e_{h,L^\infty(L^2)}$	$\mathrm{EOC}_{L^\infty(L^2)}$	$e_{h,L^2(H^1)}$	$\mathrm{EOC}_{L^2(H^1)}$
128	3.54E-01	6.57E-02	-	1.09E-01	-
512	1.77E-01	2.95E-02	1.16	5.58E-02	0.97
2048	8.84E-02	1.40E-02	1.08	2.81E-02	0.99
8192	4.42E-02	6.87E-03	1.03	1.41E-02	0.99
32768	2.21E-02	3.40E-03	1.02	7.05E-03	1.00
131072	1.11E-02	1.69E-03	1.01	3.53E-03	1.00
Average			1.06		0.99

Method with the FV meshes $\mathcal{D}^1_h = \mathcal{T}_h$

$\#I$	h	$e_{h,L^\infty(L^2)}$	$\mathrm{EOC}_{L^\infty(L^2)}$	$e_{h,L^2(H^1)}$	$\mathrm{EOC}_{L^2(H^1)}$
128	3.54E-01	7.50E-02	-	1.13E-01	-
512	1.77E-01	4.57E-02	0.71	6.18E-02	0.87
2048	8.84E-02	1.78E-02	1.36	3.01E-02	1.04
8192	4.42E-02	1.18E-02	0.59	1.62E-02	0.89
32768	2.21E-02	4.37E-03	1.43	7.56E-03	1.10
131072	1.11E-02	2.99E-03	0.55	4.12E-03	0.88
Average			0.93		0.96

Method with the FV meshes $\mathcal{D}_h^2$

Conclusion: the method is of the first order both in $L^2(0,T;H^1(\Omega))$-norm as well as in $L^\infty(0,T;L^2(\Omega))$-norm.

DG finite element method

A natural extension of the FV method for obtaining higher order schemes:
the discontinuous Galerkin method
piecewise polynomial approximation on suitable meshes without any requirement of the continuity between neighbouring elements

Consider again problem $(1) - (2)$.

$\mathcal{T}_h$ - triangulation of Ω with standard properties

$\Gamma_{ij}, s(i), |\Gamma_{ij}|$, $\boldsymbol{n}_{ij}$ as above

use the symbol Γ_{ij} for sides of K_i which are parts of $\partial\Omega$ and set $\gamma(i) = \{j;\ \Gamma_{ij} \subset \partial K_i \cap \partial\Omega\}$.

Over the triangulation $\mathcal{T}_h$ we introduce the **broken Sobolev space** $H^1(\Omega, \mathcal{T}_h) = \{v;\ v|_K \in H^1(K)\ \forall K \in \mathcal{T}_h\}$ with seminorm

$$|v|_{H^1(\Omega,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} |v|^2_{H^1(K)}\right)^{1/2}, \qquad v \in H^1(\Omega, \mathcal{T}_h). \tag{11}$$

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

$v|_{\Gamma_{ij}} =$ **trace of** $v|_{K_i}$ **on** Γ_{ij}**,**

$v|_{\Gamma_{ji}} =$ **trace of** $v|_{K_j}$ **on** Γ_{ji}**,**

$\langle v \rangle_{\Gamma_{ij}} = \frac{1}{2}\left(v|_{\Gamma_{ij}} + v|_{\Gamma_{ji}} \right)$**,**

$[v]_{\Gamma_{ij}} = v|_{\Gamma_{ij}} - v|_{\Gamma_{ji}}$

The approximate solution is sought in the space of discontinuous piecewise polynomial functions S_h **defined by**

$$S_h = S^{p,-1}(\Omega, \mathcal{T}_h) = \{v; v|_K \in P^p(K)\ \forall K \in \mathcal{T}_h\}. \quad (12)$$

DGFE formulation

We introduce the forms for $u, \varphi \in H^2(\Omega, \mathcal{T}_h)$

$$a_h(u,\varphi) = \sum_{i\in I} \int_{K_i} \varepsilon \, \nabla u \cdot \nabla \varphi \, \mathrm{d}x$$

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

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

$$J_h^\sigma(u,\varphi) = \sum_{i\in I} \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \sigma [u] \, [\varphi] \, \mathrm{d}S + \sum_{i\in I} \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma \, u \, \varphi \, \mathrm{d}S,$$

$$\ell_h(\varphi)(t) = \int_\Omega g(t),$$

$$b_h(u,\varphi) = -\sum_{i\in I}\int_{K_i}\sum_{s=1}^{d} f_s(u)\,\frac{\partial\varphi}{\partial x_s}\,\mathrm{d}x$$
$$+\sum_{i\in I}\sum_{j\in s(i)}\int_{\Gamma_{ij}} H\left(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij}\right)\varphi|_{\Gamma_{ij}}\,\mathrm{d}S.$$

H - numerical flux with above properties

Weight σ: $\sigma|_{\Gamma_{ij}} = C_W/d(\Gamma_{ij})$,

$C_W > 0$ - sufficiently large constant.

The forms a_h and J_h^σ represent here the symmetric discretization of the diffusion term and the interior penalty - SIPG version of the DG approximation

Discrete problem:

$u_h^0 \in S_h$ - $L^2(\Omega)$**-projection of** u^0 **onto** S_h

u_h - **DGFE solution, if**

a) $u_h \in C^1([0,T]; S_h)$,

b) $\left(\frac{\partial u_h(t)}{\partial t}, \varphi_h\right) + b_h(u_h(t), \varphi_h) + a_h(u_h(t), \varphi_h) + \varepsilon J_h^\sigma(u_h(t), \varphi_h)$

$= \ell_h(\varphi_h)(t) \quad \forall \varphi_h \in S_h, \ \forall t \in (0,T),$

c) $u_h(0) = u_h^0.$

(13)

Assumptions:

$$u_t = \frac{\partial u}{\partial t} \in L^2(0, T;\, H^{p+1}(\Omega)), \qquad (14)$$

$\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, **is a regular system of triangulations:**

there exists a constant $C_R > 0$ **such that**

$h_K/\rho_K \le C_R$ **for all** $K \in \mathcal{T}_h$ **and all** $h \in (0, h_0)$.

Analysis: **M.F., V. Dolejší, V. Sobotíková**

Under the above assumptions, the following error estimate was obtained:

$$
\begin{aligned}
& \max_{t\in[0,T]} \|e_h(t)\|^2_{L^2(\Omega)} \qquad (15)\\
+ \quad & \varepsilon \int_0^T \Big(|e_h(\vartheta)|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h(\vartheta), e_h(\vartheta)) \Big) \,\mathrm{d}\vartheta \le Ch^{2p}.
\end{aligned}
$$

This estimate is optimal in the $L^2(H^1)$-norm, but suboptimal in the $L^\infty(L^2)$-norm.

Goal: to derive an optimal error estimate in the $L^\infty(L^2)$-norm

carried out by M.F., V. Dolejší, V. Kučera, V. Sobotíková

with the aid of the **Aubin-Nitsche technique** based on the use of the dual problem:

$$-\Delta\psi = z \quad \text{in } \Omega,\ \psi|_{\partial\Omega} = 0 \tag{16}$$

Assumption: Ω is **convex**

Then the weak solution $\psi \in H^2(\Omega)$ and there exists a constant $C > 0$, independent of z, such that

$$\|\psi\|_{H^2(\Omega)} \leq C\|z\|_{L^2(\Omega)}. \tag{17}$$

Under the above assumptions the error $e_h = u - u_h$ satisfies the estimate

$$\|e_h\|_{L^\infty(0,T;L^2(\Omega))} \le Ch^{p+1}, \tag{18}$$

with a constant $C > 0$ independent of h.

Numerical experiments

Verification of estimate (18) by the evaluation of error and experimental order of convergence for 2D viscous Burgers equation

Data:

$\Omega = (0,1)^2$, $T = 10$, $\varepsilon = 0.1$ **and** g **such that the exact solution:** $u(x_1, x_2, t) = \left(1 - \mathrm{e}^{-10t}\right) \hat{u}(x_1, x_2)$,

where

$\hat{u}(x_1, x_2) = 2r^{\alpha} x_1 x_2 (1 - x_1)(1 - x_2)$, $r \equiv (x_1^2 + x_2^2)^{1/2}$

$\alpha \in \mathbb{R}$ - **constant which determines the regularity of the solution** u, **by Babuška**

$$\hat{u} \in H^{\beta}(\Omega) \quad \forall \beta \in (0, \alpha + 3) \tag{19}$$

Table:

computational errors in the $L^2(\Omega)$-norm at time $t = T = 10$

experimental orders of convergence (EOC) for $\alpha = 4$

$\Longrightarrow$ **we have the optimal order of convergence $O(h^{p+1})$ for $p = 1, 2, 3$**

mesh	h	P^1		P^2		P^3	
		$\|e_h\|_{L^2(\Omega)}$	EOC	$\|e_h\|_{L^2(\Omega)}$	EOC	$\|e_h\|_{L^2(\Omega)}$	EOC
1	0.177E+00	0.3943E-02	–	0.1948E-03	–	0.1099E-04	–
2	0.118E+00	0.1843E-02	1.88	0.6046E-04	2.89	0.2330E-05	3.83
3	0.884E-01	0.1060E-02	1.92	0.2621E-04	2.91	0.7689E-06	3.85
4	0.589E-01	0.4811E-03	1.95	0.7999E-05	2.93	0.1593E-06	3.88
5	0.442E-01	0.2733E-03	1.97	0.3427E-05	2.95	0.5173E-07	3.91
6	0.295E-01	0.1226E-03	1.98	0.1032E-05	2.96	0.1054E-07	3.92
7	0.221E-01	0.6927E-04	1.98	0.4385E-06	2.97	0.3494E-08	3.84

Computational errors and the corresponding orders of convergence of P^1, P^2 and P^3 approximations for $\alpha = 4$.

A small decrease of EOD P^3 approximation - caused by rounding off errors.

Conclusion

The combined FV-FE method simple, but first-order accurate

Natural extension: DGFEM - allows to construct higher-order accurate method

Open problems:

- Analysis of the problem with mixed Dirichlet-Neumann boundary conditions
- Analysis of optimal error estimates in nonconvex domains

References:

Bejček M., Feistauer M., Gallouët, T., Hájek J., Herbin R., Combined triangular FV – triangular FE method for nonlinear convection-diffusion problems, ZAMM, 87 (2007), 499-517.

Dolejší V., Feistauer M., Kučera V., Sobotíková V., An optimal $L^\infty(L^2)$-error estimate for the discontinuous Galerkin approximation of a nonlinear nonstationary convection-diffusion problem, IMA J. Numer. Anal., in press.