A unified analysis of AFC schemes for convection–diffusion equations

Petr Knobloch
Charles University, Prague

joint work with

Gabriel R. Barrenechea (Glasgow)

Volker John (Berlin)

Workshop Dresden–Prague on Numerical Analysis
Dresden, November 2–3, 2018

Outline

– algebraic flux correction scheme for a steady-state convection–diffusion–reaction equation

– formulation as edge-based stabilization

– theoretical analysis under general assumptions: solvability, discrete maximum principle, error estimates

– examples of limiters

– numerical results

Stabilization

Problem for a PDE containing a wide range of scales

$\Rightarrow$ Galerkin FEM fails unless all scales are resolved.

Resolution of all scales typically not affordable.

Remedy: modification of the Galerkin FEM (stabilization)

1) in the integral form

2) on the algebraic level (goal: conservation & DMP)

Algebraically stabilized schemes

Boris, Book (1973), Zalesak (1979) – basic philosophy of flux-corrected transport

Arminjon, Dervieux (1989), Selmin (1987), Löhner, Morgan, Peraire, Vahdati (1987) – FEM-FCT

Kuzmin et al. (2001–now) – algebraic flux correction

– algebraic stabilizations for linear boundary value problems

first rigorous theoretical analysis of the AFC method:

Barrenechea, John, K. (IMAJNA 2015, SINUM 2016, M3AS 2017)

Steady-state convection–diffusion–reaction equation

$$-\varepsilon\,\Delta u + \mathbf{b}\cdot\nabla u + c\,u = f \quad \text{in } \Omega\,, \qquad u = u_b \quad \text{on } \partial\Omega$$

with constant $\varepsilon > 0$ and

$$\nabla\cdot\mathbf{b} = 0\,, \qquad c \geq \sigma_0 \geq 0 \qquad \text{in } \Omega\,.$$

FE discretization

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1,\dots,N$, and

$$a(u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_h\,,$$

where

$$W_h = \{v_h \in C(\overline{\Omega})\,;\; v|_K \in P_1(K)\;\forall\,K \in \mathscr{T}_h\}\,, \qquad V_h = W_h \cap H_0^1(\Omega)\,,$$

$$a(u_h, v_h) = \varepsilon\,(\nabla u_h, \nabla v_h) + (\mathbf{b}\cdot\nabla u_h, v_h) + (c\,u_h, v_h)\,.$$

Algebraic problem

$$\sum_{j=1}^{N} a_{ij} u_j = f_i \,, \qquad i = 1, \ldots, M \,,$$

$$u_i = u_i^b \,, \qquad i = M+1, \ldots, N \,.$$

Algebraic problem

$$\sum_{j=1}^{N} a_{ij} u_j = f_i\,, \qquad i = 1, \ldots, M\,,$$

$$u_i = u_i^b\,, \qquad i = M+1, \ldots, N\,.$$

Properties: $(a_{ij})_{i,j=1}^{M}$ is positive definite,

$$\sum_{j=1}^{N} a_{ij} \geq 0 \quad \forall\, i = 1, \ldots, M$$

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^{N}$ ... FE matrix for homogeneous natural b.c.

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^{N}$ ... FE matrix for homogeneous natural b.c.

Symmetric artificial diffusion matrix $\mathbb{D}$:

$$d_{ij} = -\max\{a_{ij}, 0, a_{ji}\} \qquad \forall\, i \neq j\,, \qquad\qquad d_{ii} = -\sum_{j\neq i} d_{ij}\,.$$

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^N$ ... FE matrix for homogeneous natural b.c.

Symmetric artificial diffusion matrix $\mathbb{D}$:

$$d_{ij} = -\max\{a_{ij}, 0, a_{ji}\} \qquad \forall\, i \neq j\,, \qquad\qquad d_{ii} = -\sum_{j\neq i} d_{ij}\,.$$

$\Rightarrow$ $\mathbb{A} + \mathbb{D}$ satisfies conditions for DMP

Stabilized problem: $(\mathbb{A}\,\mathrm{U})_i + (\mathbb{D}\,\mathrm{U})_i = f_i\,, \qquad i = 1,\dots,M,$

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^{N}$ ... FE matrix for homogeneous natural b.c.

Symmetric artificial diffusion matrix $\mathbb{D}$:

$$d_{ij} = -\max\{a_{ij}, 0, a_{ji}\} \qquad \forall\, i \neq j\,, \qquad d_{ii} = -\sum_{j\neq i} d_{ij}\,.$$

$\Rightarrow$ $\mathbb{A} + \mathbb{D}$ satisfies conditions for DMP

Stabilized problem: $(\mathbb{A}\,\mathrm{U})_i + (\mathbb{D}\,\mathrm{U})_i = f_i\,, \qquad i = 1, \dots, M,$

$$(\mathbb{D}\,\mathrm{U})_i = \sum_{j\neq i} f_{ij} \quad \text{with} \quad f_{ij} = d_{ij}\,(u_j - u_i).$$

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^{N}$ ... FE matrix for homogeneous natural b.c.

Symmetric artificial diffusion matrix $\mathbb{D}$:

$$d_{ij} = -\max\{a_{ij}, 0, a_{ji}\} \qquad \forall\, i \neq j\,, \qquad\qquad d_{ii} = -\sum_{j\neq i} d_{ij}\,.$$

$\Rightarrow$ $\mathbb{A}+\mathbb{D}$ satisfies conditions for DMP

Stabilized problem: $(\mathbb{A}\,\mathrm{U})_i + (\mathbb{D}\,\mathrm{U})_i = f_i\,, \qquad i = 1,\dots,M,$

$$(\mathbb{D}\,\mathrm{U})_i = \sum_{j\neq i} f_{ij} \quad \text{with} \quad f_{ij} = d_{ij}\,(u_j - u_i).$$

Limiting: limit the diffusive fluxes f_{ij} to reduce smearing

Algebraic flux correction schemes

Aim: manipulate the algebraic system in such a way that the solution satisfies DMP and layers are not smeared.

$\mathbb{A} = (a_{ij})_{i,j=1}^N$ ... FE matrix for homogeneous natural b.c.

Symmetric artificial diffusion matrix $\mathbb{D}$:

$$d_{ij} = -\max\{a_{ij}, 0, a_{ji}\} \qquad \forall\, i \neq j\,, \qquad\qquad d_{ii} = -\sum_{j\neq i} d_{ij}\,.$$

$\Rightarrow \quad \mathbb{A} + \mathbb{D}$ satisfies conditions for DMP

Stabilized problem: $(\mathbb{A}\,\mathrm{U})_i + (\mathbb{D}\,\mathrm{U})_i = f_i\,, \qquad i = 1,\dots,M,$

$$(\mathbb{D}\,\mathrm{U})_i = \sum_{j\neq i} f_{ij} \quad \text{with} \quad f_{ij} = d_{ij}\,(u_j - u_i).$$

Limiting: limit the diffusive fluxes f_{ij} to reduce smearing

$$(\mathbb{A}\,\mathrm{U})_i + \sum_{j\neq i} \beta_{ij}\, f_{ij} = f_i\,, \qquad i = 1,\dots,M, \quad \beta_{ij} \in [0,1]\,.$$

Algebraic flux correction scheme

$$\sum_{j=1}^{N} a_{ij}\, u_j + \sum_{j=1}^{N} \beta_{ij}(\mathrm{U})\, d_{ij}\,(u_j - u_i) = f_i\,, \qquad i = 1,\dots,M\,,$$

$$u_i = u_i^b\,, \qquad i = M+1,\dots,N\,,$$

where $\beta_{ij}(\mathrm{U}) \in [0,1]$ and

$$\beta_{ij} = \beta_{ji}\,, \qquad i,j = 1,\dots,N\,.$$

Algebraic flux correction scheme

$$\sum_{j=1}^{N} a_{ij}\,u_j + \sum_{j=1}^{N} \beta_{ij}(\mathrm{U})\,d_{ij}\,(u_j - u_i) = f_i\,, \qquad i = 1,\dots,M\,,$$

$$u_i = u_i^b\,, \qquad i = M+1,\dots,N\,,$$

where $\beta_{ij}(\mathrm{U}) \in [0,1]$ and

$$\beta_{ij} = \beta_{ji}\,, \qquad i,j = 1,\dots,N\,.$$

Variational form of the AFC scheme

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1,\dots,N$, and

$$a(u_h,v_h) + d_h(u_h;u_h,v_h) = (f,v_h) \qquad \forall\, v_h \in V_h\,,$$

where $$d_h(z;v,w) = \sum_{i,j=1}^{N} \beta_{ij}(z)\,d_{ij}\,(v(x_j) - v(x_i))\,w(x_i)\,.$$

Edge-based formulation of the AFC scheme

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1, \dots, N$, and

$$a_h(u_h; u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_h\,,$$

where $\qquad a_h(z; v, w) = a(v, w) + d_h(z; v, w) \qquad$ and

$$d_h(z; v, w) = \sum_{E \in \mathcal{E}_h} \beta_E(z)\, |d_E|\, (v(x_{E,1}) - v(x_{E,2}))(w(x_{E,1}) - w(x_{E,2}))\,.$$

Edge-based formulation of the AFC scheme

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1, \ldots, N$, and

$$a_h(u_h; u_h, v_h) = (f, v_h) \qquad \forall\, v_h \in V_h\,,$$

where $\quad a_h(z; v, w) = a(v, w) + d_h(z; v, w) \quad$ and

$$d_h(z; v, w) = \sum_{E \in \mathscr{E}_h} \beta_E(z)\, |d_E|\, (v(x_{E,1}) - v(x_{E,2}))(w(x_{E,1}) - w(x_{E,2}))\,.$$

One has

$$d_h(z; v, w) = \sum_{E \in \mathscr{E}_h} \beta_E(z)\, |d_E|\, h_E\, (\nabla v \cdot \mathbf{t}_E, \nabla w \cdot \mathbf{t}_E)_E \qquad \forall\, v, w \in W_h\,.$$

Edge-based formulation of the AFC scheme

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1,\dots,N$, and

$$a_h(u_h;u_h,v_h) = (f,v_h) \qquad \forall\, v_h \in V_h\,,$$

where $\qquad a_h(z;v,w) = a(v,w) + d_h(z;v,w) \qquad$ and

$$d_h(z;v,w) = \sum_{E\in\mathcal{E}_h} \beta_E(z)\,|d_E|\,(v(x_{E,1})-v(x_{E,2}))(w(x_{E,1})-w(x_{E,2}))\,.$$

One has

$$d_h(z;v,w) = \sum_{E\in\mathcal{E}_h} \beta_E(z)\,|d_E|\,h_E\,(\nabla v\cdot\mathbf{t}_E,\nabla w\cdot\mathbf{t}_E)_E \qquad \forall\, v,w \in W_h\,.$$

Assumption (A1):

For any $E \in \mathcal{E}_h$, the function $\beta_E(u_h)(\nabla u_h)|_E \cdot \mathbf{t}_E$ is a continuous function of $u_h \in V_h$.

Edge-based formulation of the AFC scheme

Find $u_h \in W_h$ such that $u_h(x_i) = u_b(x_i)$, $i = M+1,\ldots,N$, and

$$a_h(u_h;u_h,v_h) = (f,v_h) \qquad \forall\, v_h \in V_h\,,$$

where $\quad a_h(z;v,w) = a(v,w) + d_h(z;v,w) \quad$ and

$$d_h(z;v,w) = \sum_{E\in\mathcal{E}_h} \beta_E(z)\,|d_E|\,(v(x_{E,1}) - v(x_{E,2}))(w(x_{E,1}) - w(x_{E,2}))\,.$$

One has

$$d_h(z;v,w) = \sum_{E\in\mathcal{E}_h} \beta_E(z)\,|d_E|\,h_E\,(\nabla v\cdot\mathbf{t}_E\,,\nabla w\cdot\mathbf{t}_E)_E \qquad \forall\, v,w \in W_h\,.$$

Theorem For any $\beta_E \in [0,1]$ satisfying Assumption (A1), the edge-based AFC scheme has a solution.

Discrete maximum principle

Local sets for $i = 1, \dots, M$:

$S_i = \{j \in \{1, \dots, N\} \setminus \{i\}\,;\ x_i \text{ and } x_j \text{ are endpoints of the same edge}\}$

$\Delta_i = \{K \in \mathscr{T}_h\,;\ x_i \in K\}$

Assumption (A2):

Consider any $u_h \in W_h$ and any $i \in \{1, \dots, M\}$. If $u_h(x_i)$ is a strict local extremum of u_h on Δ_i, i.e.,

$$u_h(x_i) > u_h(x) \quad \forall\, x \in \Delta_i \setminus \{x_i\}$$

or

$$u_h(x_i) < u_h(x) \quad \forall\, x \in \Delta_i \setminus \{x_i\}\,,$$

then

$$a_h(u_h; \varphi_j, \varphi_i) \leq 0 \qquad \forall\, j \in S_i\,.$$

Local discrete maximum principle

Let $u_h \in W_h$ be a solution of the AFC scheme with limiters β_E satisfying Assumption (A2). Consider any $i \in \{1,\dots,M\}$. Then

$$f \leq 0 \text{ in } \Delta_i \quad \Rightarrow \quad \max_{\Delta_i} u_h \leq \max_{\partial\Delta_i} u_h^+ ,$$

$$f \geq 0 \text{ in } \Delta_i \quad \Rightarrow \quad \min_{\Delta_i} u_h \geq \min_{\partial\Delta_i} u_h^- ,$$

where $u_h^+ = \max\{0, u_h\}$ and $u_h^- = \min\{0, u_h\}$. If, in addition, $c = 0$ in Δ_i, then

$$f \leq 0 \text{ in } \Delta_i \quad \Rightarrow \quad \max_{\Delta_i} u_h = \max_{\partial\Delta_i} u_h ,$$

$$f \geq 0 \text{ in } \Delta_i \quad \Rightarrow \quad \min_{\Delta_i} u_h = \min_{\partial\Delta_i} u_h .$$

Global discrete maximum principle

Let $u_h \in W_h$ be a solution of the AFC scheme with limiters β_E satisfying Assumptions (A1) and (A2). Then

$$f \le 0 \ \text{ in } \Omega \quad \Rightarrow \quad \max_{\overline{\Omega}} u_h \le \max_{\partial\Omega} u_h^+ \,,$$

$$f \ge 0 \ \text{ in } \Omega \quad \Rightarrow \quad \min_{\overline{\Omega}} u_h \ge \min_{\partial\Omega} u_h^- \,.$$

If, in addition, $c = 0$ in Ω, then

$$f \le 0 \ \text{ in } \Omega \quad \Rightarrow \quad \max_{\overline{\Omega}} u_h = \max_{\partial\Omega} u_h \,,$$

$$f \ge 0 \ \text{ in } \Omega \quad \Rightarrow \quad \min_{\overline{\Omega}} u_h = \min_{\partial\Omega} u_h \,.$$

A priori error estimates

Natural norm: $\|v\|_h = \left(\varepsilon\,|v|_{1,\Omega}^2 + \sigma_0\,\|v\|_{0,\Omega}^2 + d_h(u_h; v, v)\right)^{1/2}$

Theorem Let $u \in H^2(\Omega)$ and $\sigma_0 > 0$. Then

$$\|u - u_h\|_h \le C\left(\varepsilon + \sigma_0^{-1}\{\|\mathbf{b}\|_{0,\infty,\Omega}^2 + \|c\|_{0,\infty,\Omega}^2\, h^2\}\right)^{1/2} h\,|u|_{2,\Omega}$$
$$+ d_h(u_h; i_h u, i_h u)^{1/2}\,.$$

A priori error estimates

Natural norm: $$\|v\|_h = \left(\varepsilon\,|v|^2_{1,\Omega} + \sigma_0\,\|v\|^2_{0,\Omega} + d_h(u_h; v, v)\right)^{1/2}$$

Theorem Let $u \in H^2(\Omega)$ and $\sigma_0 > 0$. Then

$$\|u - u_h\|_h \leq C\big(\varepsilon + \sigma_0^{-1}\{\|\mathbf{b}\|^2_{0,\infty,\Omega} + \|c\|^2_{0,\infty,\Omega}\,h^2\}\big)^{1/2} h\,|u|_{2,\Omega} + d_h(u_h; i_h u, i_h u)^{1/2}\,.$$

Lemma Denoting

$$A_h = \max_{E \in \mathscr{E}_h} \left(|d_E|\,h_E^{2-d}\right),$$

one has

$$d_h(u_h; i_h u, i_h u) \leq C A_h\,|i_h u|^2_{1,\Omega} \qquad \forall\, u_h \in W_h,\ u \in C(\overline{\Omega})\,.$$

If, in particular, d_E are defined as at the beginning, then

$$d_h(u_h; i_h u, i_h u) \leq C\,(\varepsilon + \|\mathbf{b}\|_{0,\infty,\Omega}\,h + \|c\|_{0,\infty,\Omega}\,h^2)\,|i_h u|^2_{1,\Omega}\,.$$

An improved estimate

Assumption (A3):

The limiters β_E possess the linearity-preservation property, i.e.,

$$\beta_E(u_h) = 0 \qquad \text{if } \; u_h|_{\omega_E} \in P_1(\omega_E) \qquad \forall\, E \in \mathscr{E}_h\,.$$

Assumption (A4):

For any $E \in \mathscr{E}_h$ with endpoints x_i and x_j, the function $\beta_E(u_h)(\nabla u_h)|_E \cdot \mathbf{t}_E$ is Lipschitz continuous in the sense that

$$\Big|\beta_E(u_h)(\nabla u_h)|_E \cdot \mathbf{t}_E - \beta_E(v_h)(\nabla v_h)|_E \cdot \mathbf{t}_E\Big| \leq C \sum_{E' \in \mathscr{E}_i \cup \mathscr{E}_j} \Big|(\nabla(u_h - v_h))|_{E'} \cdot \mathbf{t}_{E'}\Big|\,.$$

An improved estimate

Assumption (A3):

The limiters β_E possess the linearity-preservation property, i.e.,

$$\beta_E(u_h) = 0 \qquad \text{if } \; u_h|_{\omega_E} \in P_1(\omega_E) \qquad \forall\, E \in \mathscr{E}_h\,.$$

Assumption (A4):

For any $E \in \mathscr{E}_h$ with endpoints x_i and x_j, the function $\beta_E(u_h)(\nabla u_h)|_E \cdot \mathbf{t}_E$ is Lipschitz continuous in the sense that

$$\left|\beta_E(u_h)(\nabla u_h)|_E \cdot \mathbf{t}_E - \beta_E(v_h)(\nabla v_h)|_E \cdot \mathbf{t}_E\right| \leq C \sum_{E' \in \mathscr{E}_i \cup \mathscr{E}_j} \left|(\nabla(u_h - v_h))|_{E'} \cdot \mathbf{t}_{E'}\right|.$$

Lemma Under Assumptions (A3) and (A4) one has

$$d_h(u_h; i_h u, i_h u) \leq \frac{\varepsilon}{2}\,|u_h - i_h u|^2_{1,\Omega} + C\,\frac{A_h^2}{\varepsilon}\,|i_h u|^2_{1,\Omega} + \varepsilon\, h^2\,|u|^2_{2,\Omega}\,.$$

Kuzmin's limiter

Zalesak (1979), Kuzmin (2007)

$$P_i^+ := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^+ , \qquad Q_i^+ := -\sum_{j=1}^{N} f_{ij}^- , \qquad R_i^+ := \min\left\{1, \frac{Q_i^+}{P_i^+}\right\} ,$$

$$P_i^- := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^- , \qquad Q_i^- := -\sum_{j=1}^{N} f_{ij}^+ , \qquad R_i^- := \min\left\{1, \frac{Q_i^-}{P_i^-}\right\} .$$

$$f_{ij} = d_{ij}\,(u_j - u_i)$$

$$\widetilde{\alpha}_{ij} := \begin{cases} R_i^+ & \text{if } f_{ij} > 0 , \\ 1 & \text{if } f_{ij} = 0 , \\ R_i^- & \text{if } f_{ij} < 0 , \end{cases} \qquad \beta_E := \begin{cases} 1 - \widetilde{\alpha}_{ij} & \text{if } a_{ji} \le a_{ij} , \\ 1 - \widetilde{\alpha}_{ji} & \text{else.} \end{cases}$$

Kuzmin's limiter

Zalesak (1979), Kuzmin (2007)

$$P_i^+ := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^+, \qquad Q_i^+ := -\sum_{j=1}^{N} f_{ij}^-, \qquad R_i^+ := \min\left\{1, \frac{Q_i^+}{P_i^+}\right\},$$

$$P_i^- := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^-, \qquad Q_i^- := -\sum_{j=1}^{N} f_{ij}^+, \qquad R_i^- := \min\left\{1, \frac{Q_i^-}{P_i^-}\right\}.$$

$$f_{ij} = d_{ij}\,(u_j - u_i)$$

$$\widetilde{\alpha}_{ij} := \begin{cases} R_i^+ & \text{if } f_{ij} > 0, \\ 1 & \text{if } f_{ij} = 0, \\ R_i^- & \text{if } f_{ij} < 0, \end{cases} \qquad \beta_E := \begin{cases} 1 - \widetilde{\alpha}_{ij} & \text{if } a_{ji} \le a_{ij}, \\ 1 - \widetilde{\alpha}_{ji} & \text{else.} \end{cases}$$

Assumption (A2) satisfied if $\min\{a_{ij}, a_{ji}\} \le 0 \quad \forall\, i, j \in \{1, \dots, N\}$

Kuzmin's limiter

Zalesak (1979), Kuzmin (2007)

$$P_i^+ := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^+ , \qquad Q_i^+ := -\sum_{j=1}^{N} f_{ij}^- , \qquad R_i^+ := \min\left\{1, \frac{Q_i^+}{P_i^+}\right\},$$

$$P_i^- := \sum_{\substack{j=1 \\ a_{ji} \le a_{ij}}}^{N} f_{ij}^- , \qquad Q_i^- := -\sum_{j=1}^{N} f_{ij}^+ , \qquad R_i^- := \min\left\{1, \frac{Q_i^-}{P_i^-}\right\}.$$

$$f_{ij} = d_{ij}\,(u_j - u_i)$$

$$\widetilde{\alpha}_{ij} := \begin{cases} R_i^+ & \text{if } f_{ij} > 0, \\ 1 & \text{if } f_{ij} = 0, \\ R_i^- & \text{if } f_{ij} < 0, \end{cases} \qquad \beta_E := \begin{cases} 1 - \widetilde{\alpha}_{ij} & \text{if } a_{ji} \le a_{ij}, \\ 1 - \widetilde{\alpha}_{ji} & \text{else.} \end{cases}$$

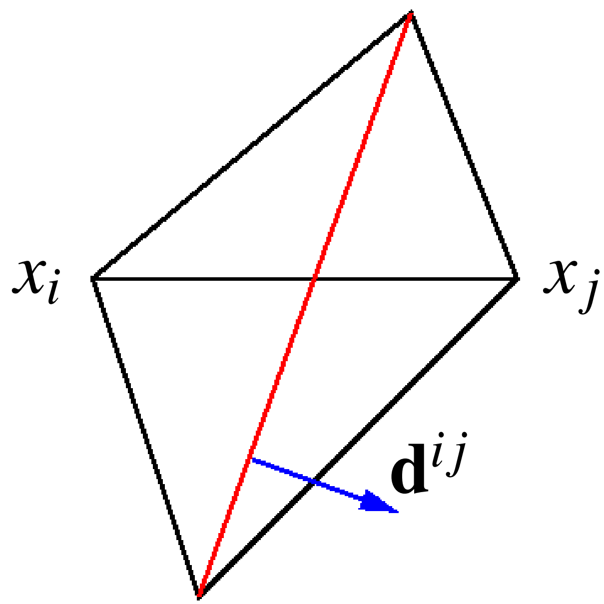
Assumption (A2) satisfied if $\min\{a_{ij}, a_{ji}\} \le 0 \quad \forall\, i, j \in \{1, \dots, N\}$

$\Rightarrow$ DMP guaranteed for Delaunay meshes for lumped react. term

... and often holds on non-Delaunay meshes!!!

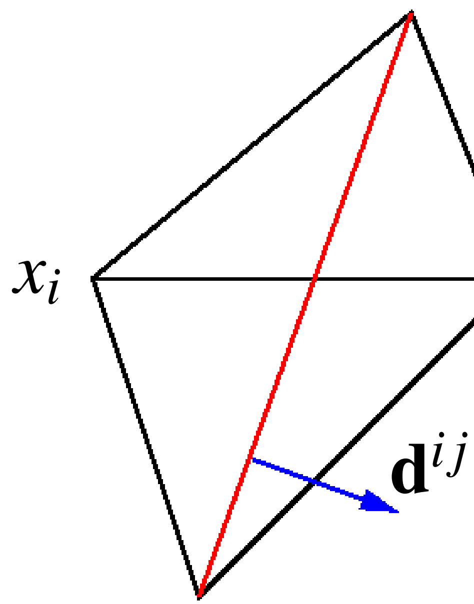
Condition $a_{ji} \leq a_{ij}$ for constant $\mathbf{b}$

$a_{ji} < a_{ij} \;\Leftrightarrow\; \mathbf{b}\cdot\mathbf{d}^{ij} > 0$ with $\mathbf{d}^{ij} = \int_\Omega \varphi_i \nabla\varphi_j \,\mathrm{d}x$



Condition $a_{ji} \leq a_{ij}$ for constant $\mathbf{b}$

$a_{ji} < a_{ij} \iff \mathbf{b} \cdot \mathbf{d}^{ij} > 0$ with $\mathbf{d}^{ij} = \int_\Omega \varphi_i \nabla \varphi_j \, \mathrm{d}x$



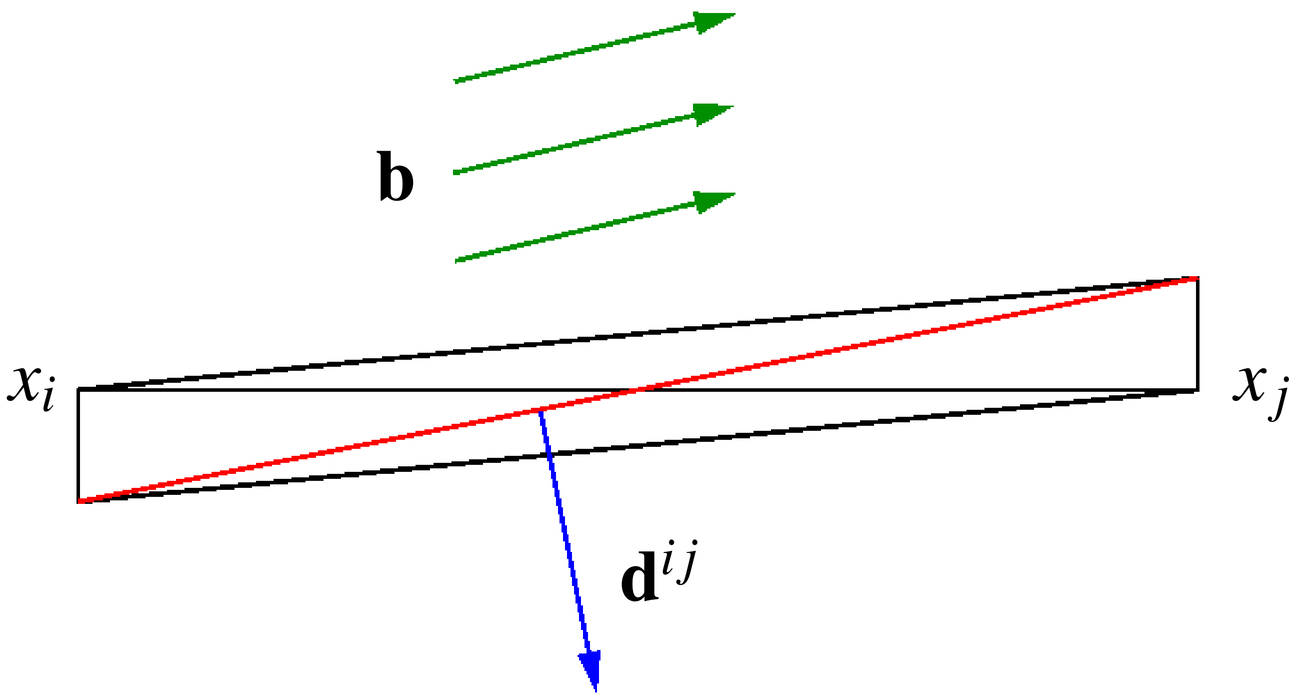
$\Rightarrow$ if $\mathbf{b} \parallel x_i x_j$ or red line $\perp x_i x_j$, then

$a_{ji} < a_{ij} \iff x_i$ is the upwind vertex.

Condition $a_{ji} \leq a_{ij}$ for constant $\mathbf{b}$

$a_{ji} < a_{ij} \;\Leftrightarrow\; \mathbf{b} \cdot \mathbf{d}^{ij} > 0$ with $\mathbf{d}^{ij} = \int_\Omega \varphi_i \nabla \varphi_j \, \mathrm{d}x$

BUT:



$a_{ji} > a_{ij}$, but x_i is the upwind vertex!

BJK limiter Kuzmin (2012), Barrenechea, John, K. (2016)

$$u_i^{\max} := \max_{j\in S_i\cup\{i\}} u_j\,, \qquad u_i^{\min} := \min_{j\in S_i\cup\{i\}} u_j\,, \qquad q_i := \gamma_i \sum_{j\in S_i} d_{ij}\,,$$

$$P_i^+ := \sum_{j\in S_i} f_{ij}^+\,, \qquad Q_i^+ := q_i\,(u_i - u_i^{\max})\,, \qquad R_i^+ := \min\left\{1, \frac{Q_i^+}{P_i^+}\right\},$$

$$P_i^- := \sum_{j\in S_i} f_{ij}^-\,, \qquad Q_i^- := q_i\,(u_i - u_i^{\min})\,, \qquad R_i^- := \min\left\{1, \frac{Q_i^-}{P_i^-}\right\},$$

$$\widetilde{\alpha}_{ij} := \begin{cases} R_i^+ & \text{if } f_{ij} > 0\,, \\ 1 & \text{if } f_{ij} = 0\,, \\ R_i^- & \text{if } f_{ij} < 0\,, \end{cases} \qquad \beta_E := 1 - \min\{\widetilde{\alpha}_{ij}, \widetilde{\alpha}_{ji}\}.$$

BJK limiter Kuzmin (2012), Barrenechea, John, K. (2016)

$$u_i^{\max} := \max_{j\in S_i\cup\{i\}} u_j\,, \qquad u_i^{\min} := \min_{j\in S_i\cup\{i\}} u_j\,, \qquad q_i := \gamma_i \sum_{j\in S_i} d_{ij}\,,$$

$$P_i^+ := \sum_{j\in S_i} f_{ij}^+\,, \qquad Q_i^+ := q_i\,(u_i - u_i^{\max})\,, \qquad R_i^+ := \min\left\{1, \frac{Q_i^+}{P_i^+}\right\},$$

$$P_i^- := \sum_{j\in S_i} f_{ij}^-\,, \qquad Q_i^- := q_i\,(u_i - u_i^{\min})\,, \qquad R_i^- := \min\left\{1, \frac{Q_i^-}{P_i^-}\right\},$$

$$\widetilde{\alpha}_{ij} := \begin{cases} R_i^+ & \text{if } f_{ij} > 0\,, \\ 1 & \text{if } f_{ij} = 0\,, \\ R_i^- & \text{if } f_{ij} < 0\,, \end{cases} \qquad \beta_E := 1 - \min\{\widetilde{\alpha}_{ij}, \widetilde{\alpha}_{ji}\}.$$

Assumption (A2) always satisfied

$\Rightarrow$ DMP guaranteed for arbitrary meshes!

BBK limiter

Barrenechea, Burman, Karakatsani (2017)

$$d_E := \gamma_0 \, h_E^{d-1} \,,$$

$$\beta_E(u_h) := \max_{x\in E} \left[\xi_{u_h}(x)\right]^p \quad , \quad p \in [1, +\infty) \,,$$

where $\xi_{u_h} \in W_h$ has the nodal values

$$\xi_{u_h}(x_i) := \begin{cases} \dfrac{\left| \sum_{j\in S_i} \left(u_h(x_i) - u_h(x_j)\right) \right|}{\sum_{j\in S_i} |u_h(x_i) - u_h(x_j)|}, & \text{if } \sum_{j\in S_i} |u_h(x_i) - u_h(x_j)| \neq 0 \,, \\ 0, & \text{otherwise} \,. \end{cases}$$

BBK limiter

Barrenechea, Burman, Karakatsani (2017)

$$d_E := \gamma_0 \, h_E^{d-1} \,,$$

$$\beta_E(u_h) := \max_{x\in E} \left[\xi_{u_h}(x)\right]^p \quad , \quad p \in [1,+\infty) \,,$$

where $\xi_{u_h} \in W_h$ has the nodal values

$$\xi_{u_h}(x_i) := \begin{cases} \dfrac{\left|\sum_{j\in S_i} (u_h(x_i) - u_h(x_j))\right|}{\sum_{j\in S_i} |u_h(x_i) - u_h(x_j)|}, & \text{if } \sum_{j\in S_i} |u_h(x_i) - u_h(x_j)| \neq 0 \,, \\ 0, & \text{otherwise} \,. \end{cases}$$

Assumption (A2) satisfied if $\gamma_0 \geq C_0 \|\mathbf{b}\|_{0,\infty,\Omega} + C_1 \, \|c\|_{0,\infty,\Omega} \, h$ and

$$(\nabla\varphi_j, \nabla\varphi_i)_\Omega \leq 0 \,, \qquad i = 1,\dots,M \,, \quad j = 1,\dots,N$$

$\Rightarrow$ DMP holds for Delaunay meshes

Validity of Assumption (A3) (linearity preservation)

Kuzmin's limiter: only for $\mathbf{b} =$ const. and special meshes (e.g., Friedrichs–Keller)

BJK limiter: for arbitrary meshes if

$$\gamma_i = \frac{\max\limits_{x_j \in \partial\Delta_i} |x_i - x_j|}{\mathrm{dist}(x_i, \partial\Delta_i^{\mathrm{conv}})}, \qquad i = 1, \dots, M$$

BBK limiter: only for symmetric patches Δ_i

Validity of Assumption (A3) (linearity preservation)

Kuzmin's limiter: only for $\mathbf{b} = \text{const.}$ and special meshes (e.g., Friedrichs–Keller)

BJK limiter: for arbitrary meshes if

$$\gamma_i = \frac{\max\limits_{x_j \in \partial\Delta_i} |x_i - x_j|}{\mathrm{dist}(x_i, \partial\Delta_i^{\mathrm{conv}})}, \qquad i = 1, \dots, M$$

BBK limiter: only for symmetric patches Δ_i

Improved error estimate on special meshes

$$\|u - u_h\|_h \le C\, h\, \|u\|_{2,\Omega} + C\, \frac{h}{\sqrt{\varepsilon}}\, |i_h u|_{1,\Omega}$$

Example 1 (polynomial solution)

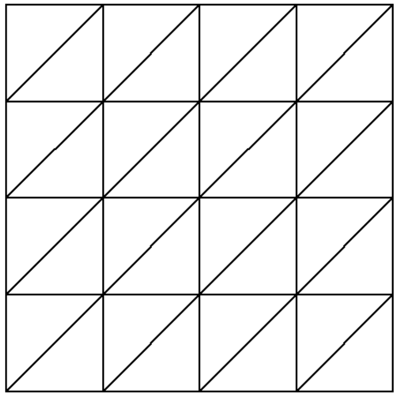
$$\Omega = (0,1)^2\,, \qquad \mathbf{b} = (3,2)\,, \qquad c = 1\,, \qquad u_b = 0\,.$$

The right-hand side f is chosen such that, for given ε,

$$u(x,y) = 100\,x^2\,(1-x)^2\,y\,(1-y)\,(1-2y)$$

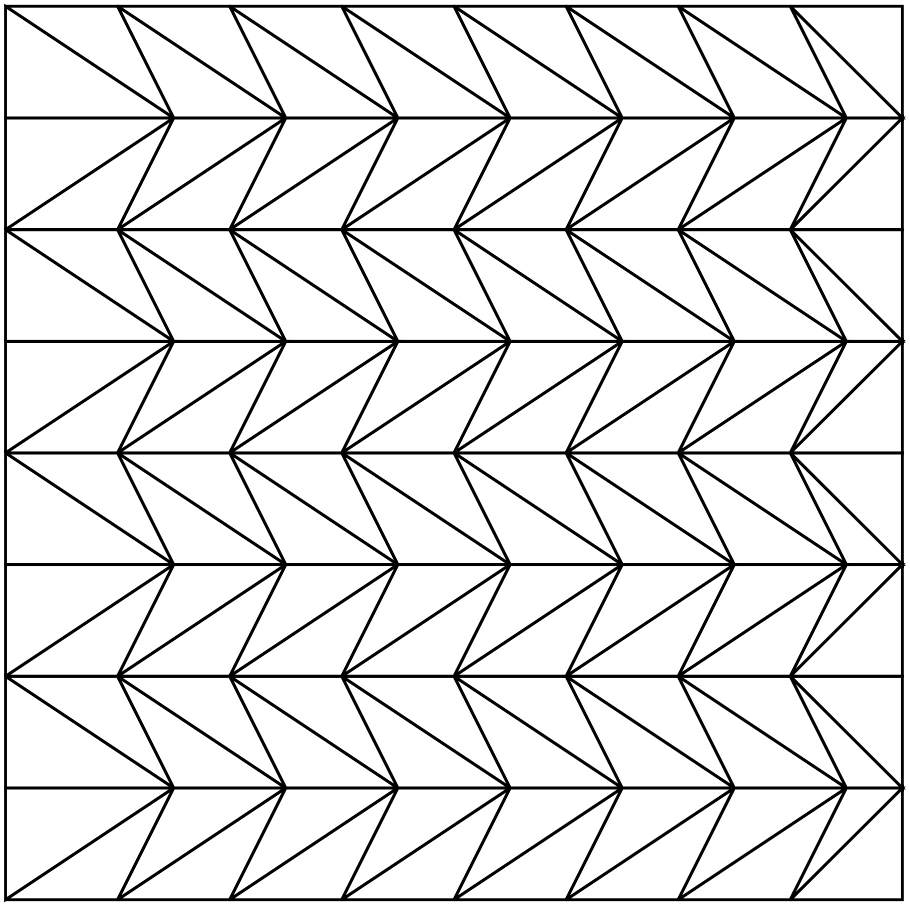
is the exact solution.

Example 1, Kuzmin's limiter, $\varepsilon = 10^{-8}$



ne	$\|e_h\|_{0,\Omega}$	ord.	$\|e_h\|_{1,\Omega}$	ord.	$d_h^{1/2}(u_h)$	ord.
32	5.457e−3	1.85	2.287e−1	1.10	1.163e−2	2.11
64	1.408e−3	1.95	1.074e−1	1.09	2.683e−3	2.12
128	3.493e−4	2.01	5.113e−2	1.07	6.410e−4	2.07
256	8.652e−5	2.01	2.546e−2	1.01	1.633e−4	1.97
512	2.152e−5	2.01	1.321e−2	0.95	4.099e−5	1.99
1024	5.357e−6	2.01	6.822e−3	0.95	1.018e−5	2.01

Non-Delaunay meshes



Example 1, Kuzmin's limiter, $\varepsilon = 10$ (non-Delaunay mesh)

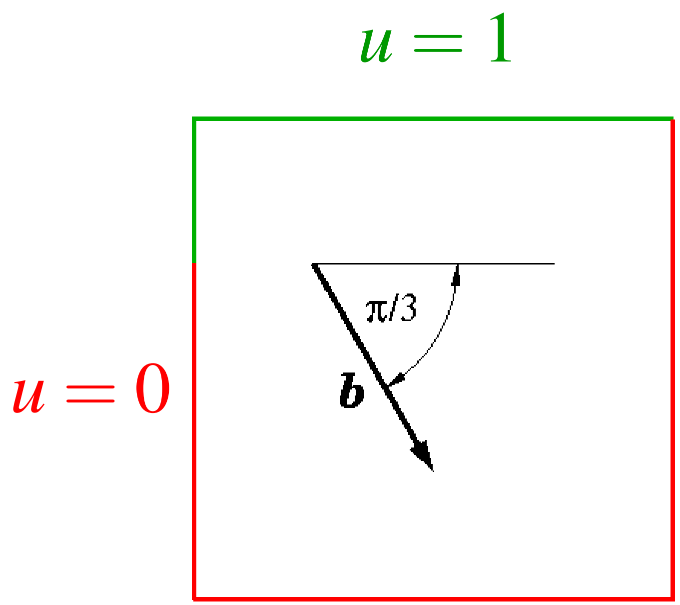
ne	$\|e_h\|_{0,\Omega}$	ord.	$\|e_h\|_{1,\Omega}$	ord.	$\|e_h\|_h$	ord.
16	5.637e−2	0.22	6.741e−1	0.41	2.626e+0	0.24
32	5.385e−2	0.07	5.908e−1	0.19	2.437e+0	0.11
64	5.332e−2	0.01	5.661e−1	0.06	2.380e+0	0.03
128	5.321e−2	0.00	5.593e−1	0.02	2.363e+0	0.01
256	5.319e−2	0.00	5.575e−1	0.00	2.358e+0	0.00
512	5.320e−2	0.00	5.570e−1	0.00	2.356e+0	0.00

Example 1, BJK limiter, $\varepsilon = 10$

(non-Delaunay mesh)

ne	$\|e_h\|_{0,\Omega}$	ord.	$\|e_h\|_{1,\Omega}$	ord.	$\|e_h\|_h$	ord.
16	1.786e−2	1.74	4.726e−1	0.87	1.522e+0	0.88
32	4.218e−3	2.08	2.404e−1	0.98	7.633e−1	1.00
64	1.016e−3	2.05	1.213e−1	0.99	3.841e−1	0.99
128	2.545e−4	2.00	6.082e−2	1.00	1.924e−1	1.00
256	6.439e−5	1.98	3.045e−2	1.00	9.632e−2	1.00
512	1.628e−5	1.98	1.524e−2	1.00	4.819e−2	1.00

Example 2 (interior layer and exponential boundary layers)



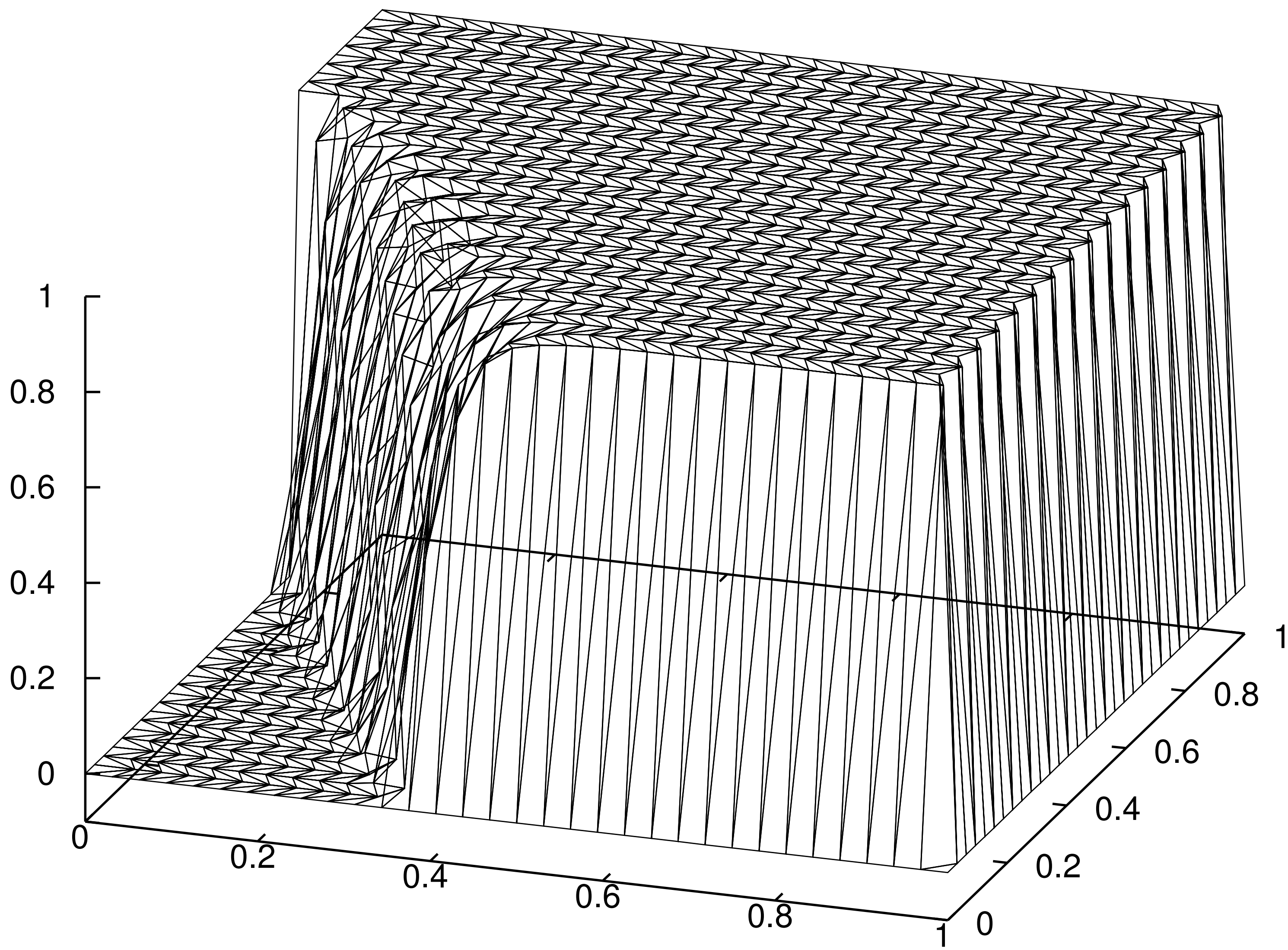
$$\varepsilon = 10^{-8}$$

$$|\mathbf{b}| = 1$$

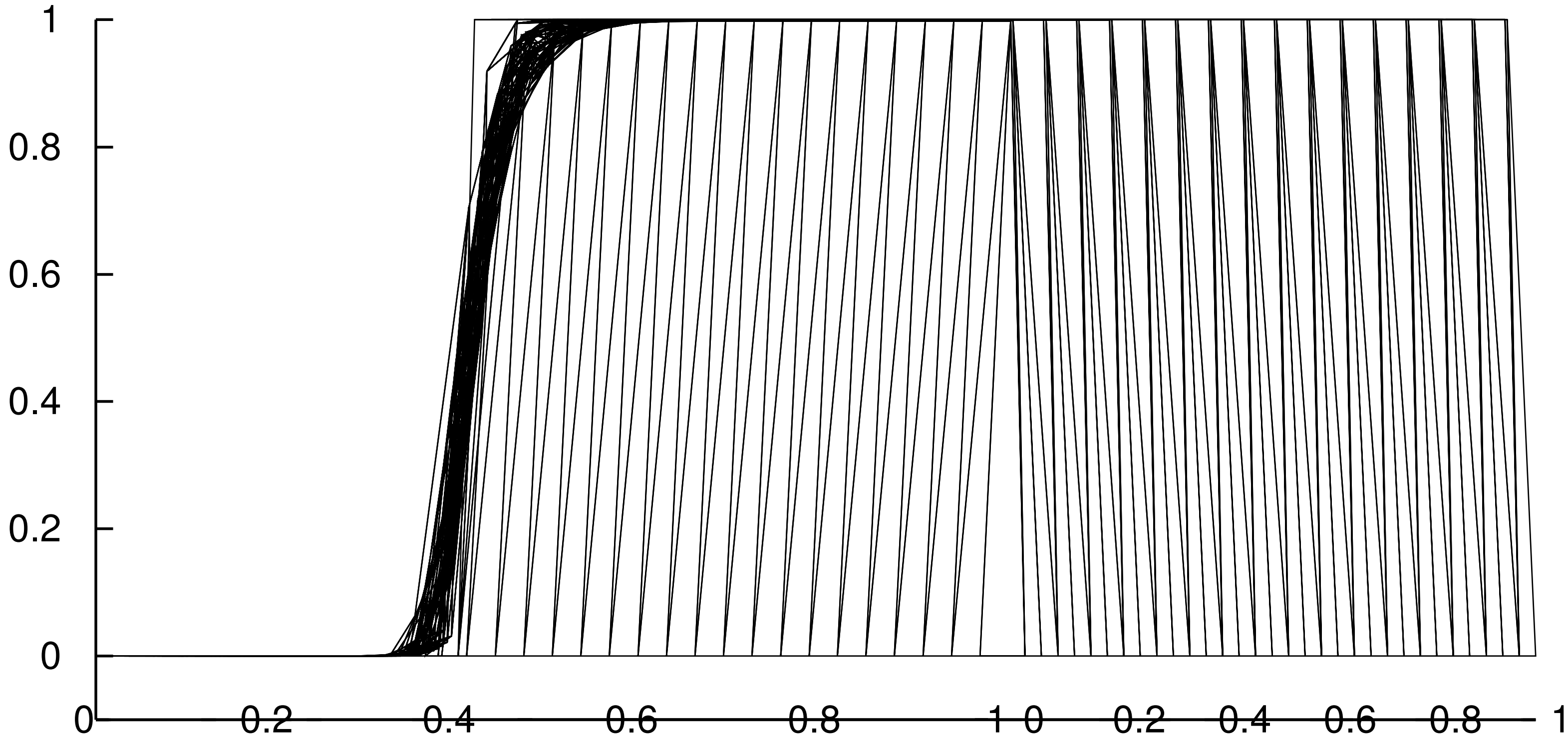
$$c = 0$$

$$f = 0$$

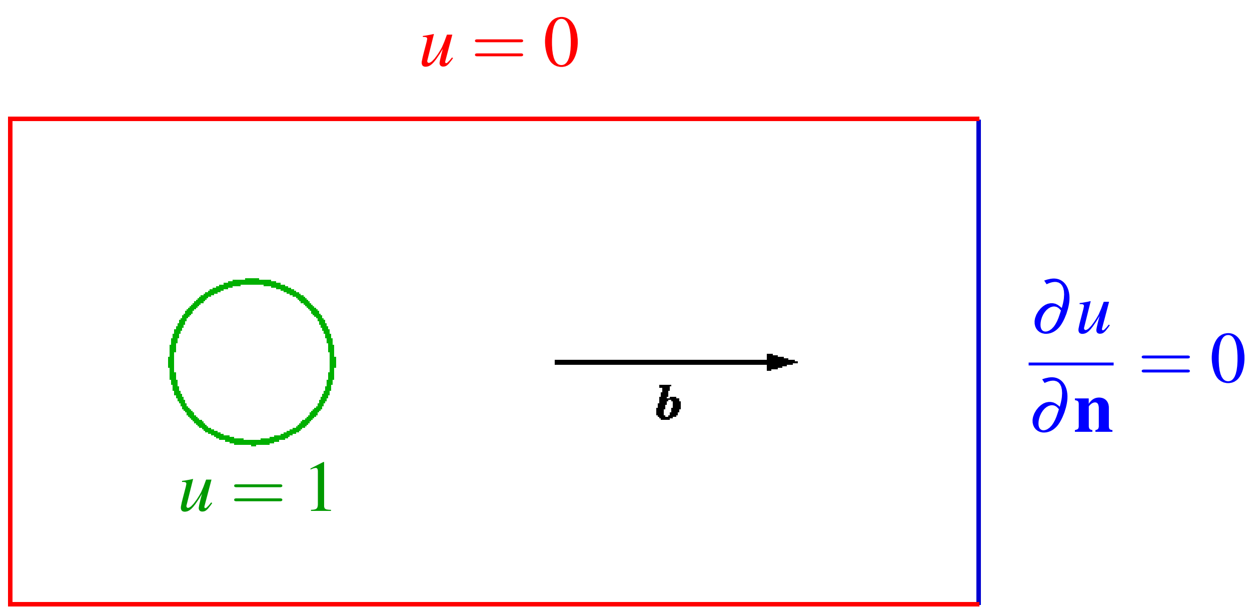
Example 2, BJK limiter



Example 2, BJK limiter



Example 3 (P. Hemker's problem)



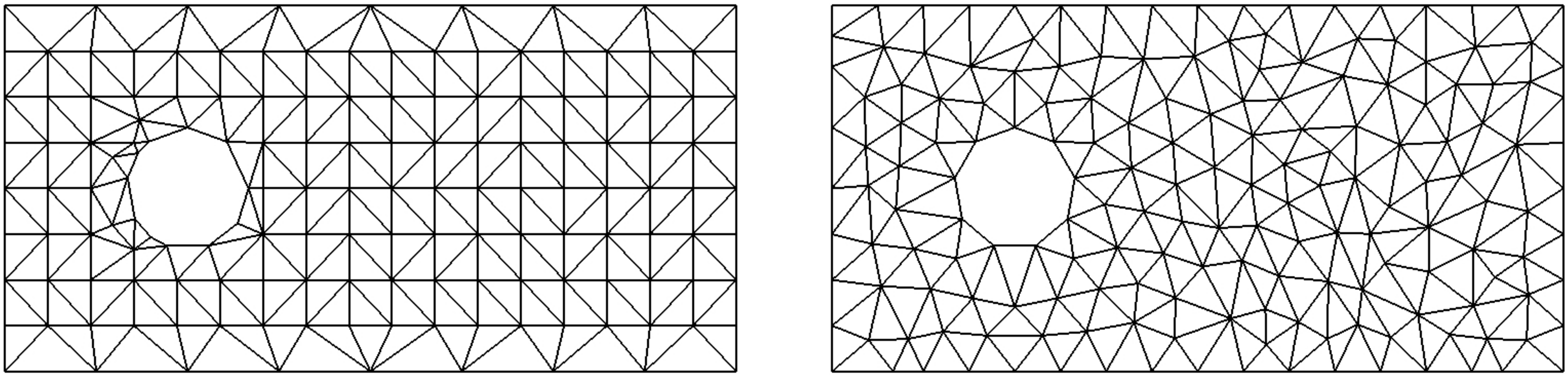
$\varepsilon = 10^{-4}$

$|\mathbf{b}| = 1$

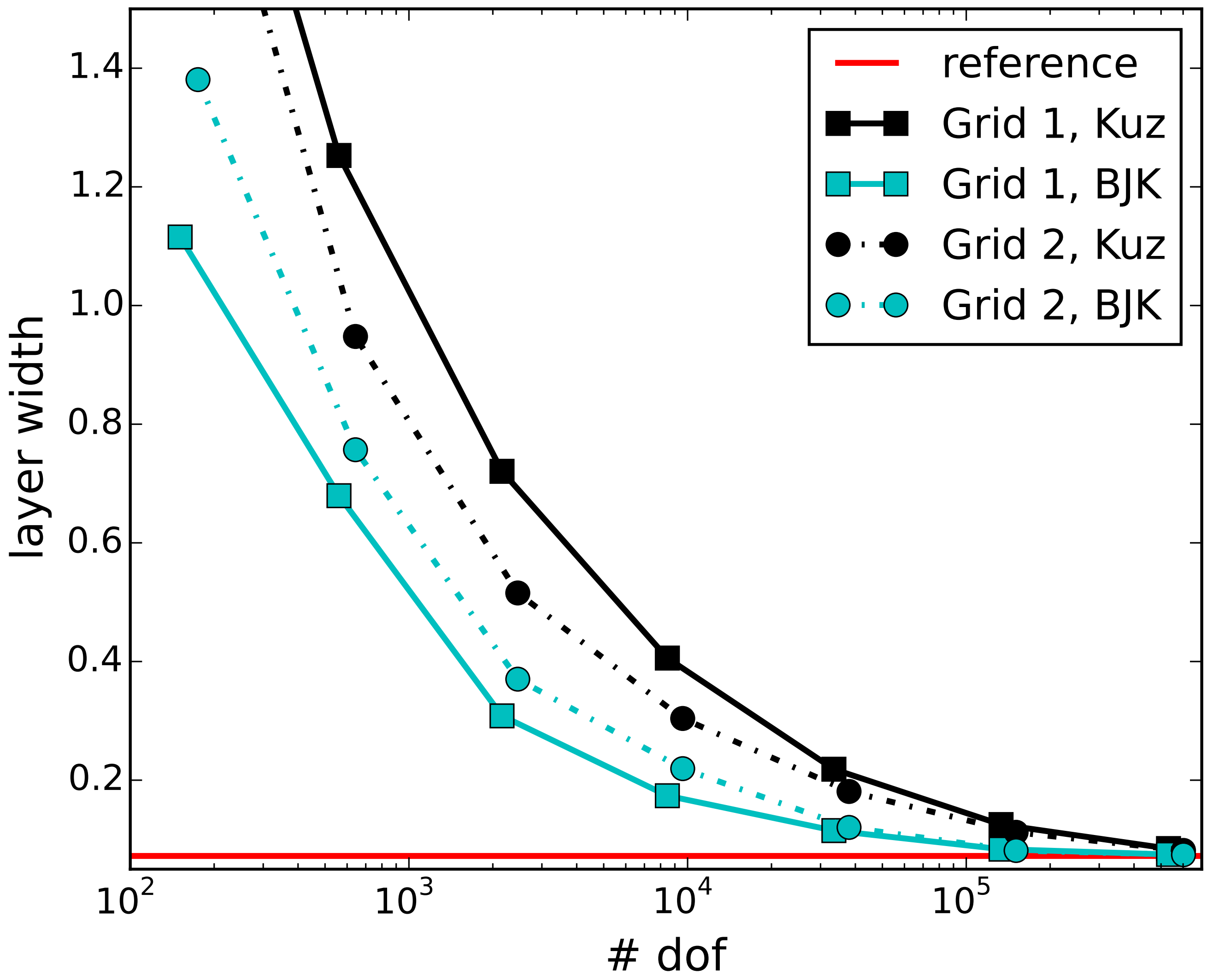
$c = 0$

$f = 0$

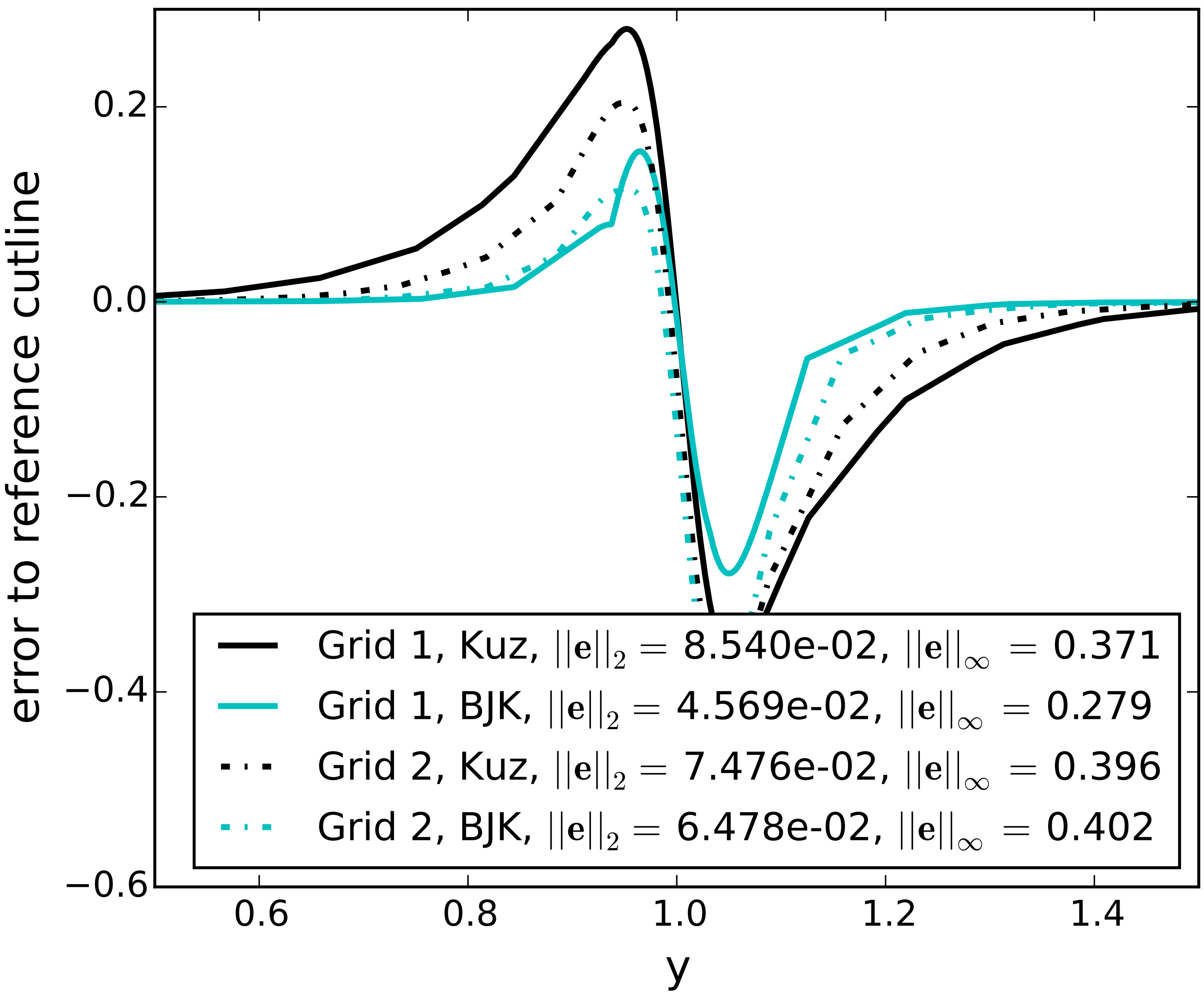
Example 3, Grid 1 (left) and Grid 2 (right), both level 0



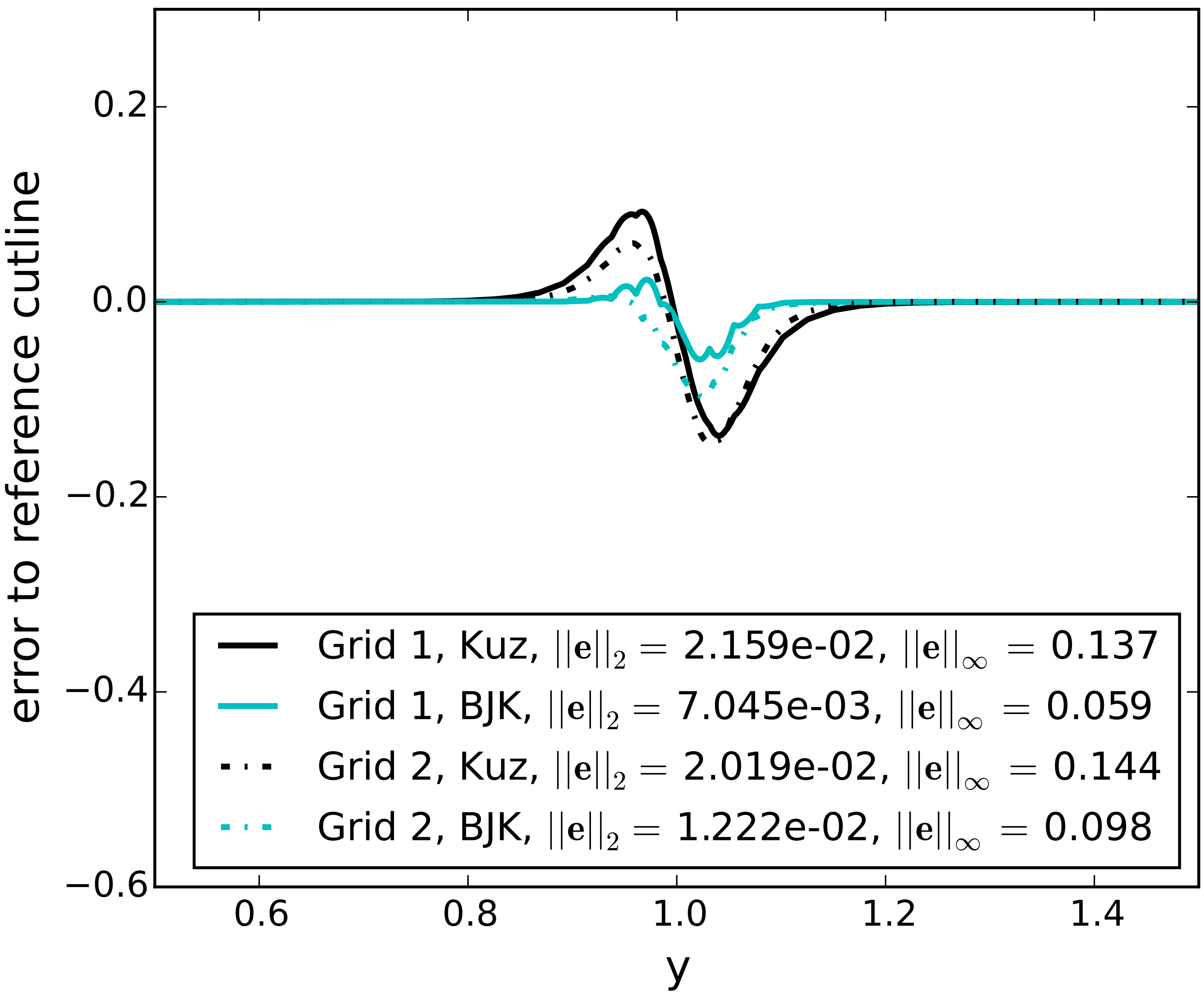
Example 3, width of the interior layer at $x = 4$



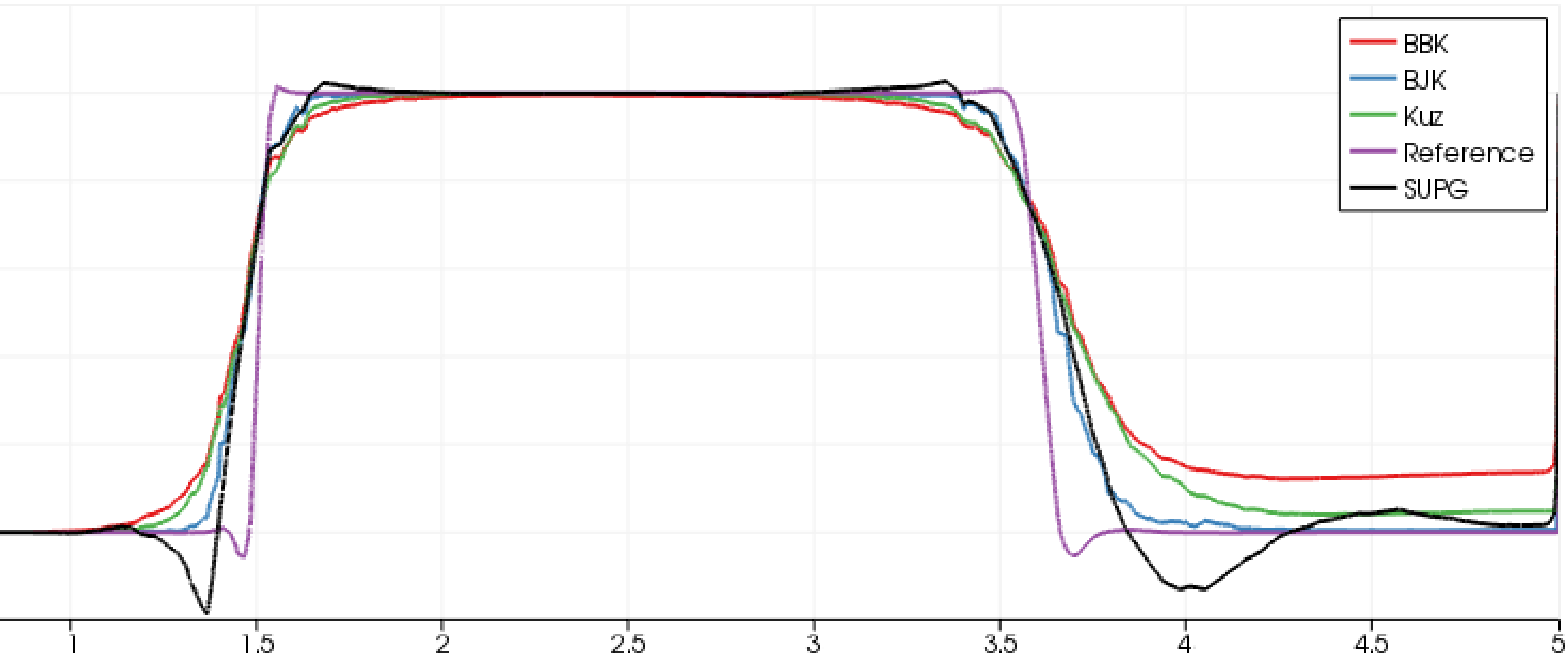
Example 3, level 3 (~ 10000 dofs), errors along $x = 4$



Example 3, level 5 (~ 150000 dofs), errors along $x = 4$



A 3D Example, approximations along a line (computed by Richard Rankin)



Conclusions

– unified theoretical analysis for algebraic flux correction schemes applied to convection–diffusion–reaction equations

– application of the theory to various limiters

– properties of the limiters illustrated by numerical results