Numerical Algorithms for Algebraic Stabilizations of Scalar Convection-Dominated Problems

Abhinav Jha, PhD

Applied and Computational Mathematics

Lunch Seminar
26^{th} Jan. 2021

Outline

- Steady-state convection-diffusion-reaction equation

$$\begin{aligned} -\varepsilon\Delta u + \boldsymbol{b}\cdot\nabla u + cu &= f \quad \text{in } \Omega \\ u &= u^b \quad \text{on } \Gamma, \end{aligned} \tag{1}$$

  - Ω – bounded polyhedral Lipschitz domain in $\mathbb{R}^d$, $d \in \{2,3\}$
  - Assume

$$\left(c(x) - \frac{1}{2}\nabla\cdot\boldsymbol{b}(x)\right) \geq \sigma > 0$$

  - Interested in convection-dominated regime, $\varepsilon \ll \|\boldsymbol{b}\|_{L^\infty(\Omega)}$

- Ideal discretization
  1. Sharp layers
     - Many discretizations satisfy this property, e.g., SUPG[1]
     - Reasonably well for AFC schemes

[1] Brooks, Hughes: CMAME (32), 199-259, 1982

- Ideal discretization
  1. Sharp layers
     - Many discretizations satisfy this property, e.g., SUPG[1]
     - Reasonably well for AFC schemes
  2. Physically consistent results (no spurious oscillations)
     - Most discretizations violate this property, e.g., SUPG, SOLD schemes
     - Satisfied for AFC schemes

[1] Brooks, Hughes: CMAME (32), 199-259, 1982

- Ideal discretization

  1. Sharp layers
     - Many discretizations satisfy this property, e.g., SUPG[1]
     - Reasonably well for AFC schemes
  2. Physically consistent results (no spurious oscillations)
     - Most discretizations violate this property, e.g., SUPG, SOLD schemes
     - Satisfied for AFC schemes
  3. Efficient computation of the solutions
     - Satisfied for linear discretizations
     - Usually not satisfied for nonlinear discretizations, like AFC schemes

[1] Brooks, Hughes: CMAME (32), 199-259, 1982

- Ideal discretization
  1. Sharp layers
     - Many discretizations satisfy this property, e.g., SUPG[1]
     - Reasonably well for AFC schemes
  2. Physically consistent results (no spurious oscillations)
     - Most discretizations violate this property, e.g., SUPG, SOLD schemes
     - Satisfied for AFC schemes
  3. Efficient computation of the solutions
     - Satisfied for linear discretizations
     - Usually not satisfied for nonlinear discretizations, like AFC schemes
- Because of 2^{nd} property: AFC schemes very well suited for applications

[1] Brooks, Hughes: CMAME (32), 199-259, 1982

- Hemker Problem

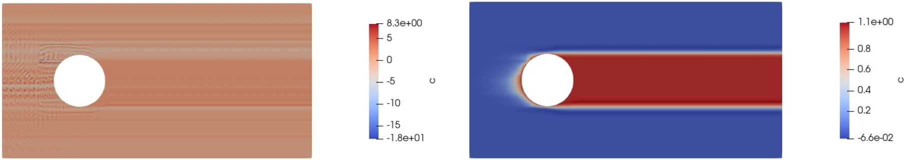


Figure 1: Galerkin Solution and SUPG Solution

Figure 2: AFC Solution

- Derivation
  - Galerkin FEM (Algebraic form)

$$\begin{aligned} \sum_{j=1}^{N} a_{ij}u_j &= f_i, \quad i = 1,\ldots,M, \\ u_i &= u_i^b, \quad i = M+1,\ldots,N \end{aligned}$$

- Derivation
  - Galerkin FEM (Algebraic form)

$$\begin{aligned} \sum_{j=1}^{N} a_{ij} u_j &= f_i, \quad i = 1, \ldots, M, \\ u_i &= u_i^b, \quad i = M+1, \ldots, N \end{aligned}$$

  - Artificial diffusion matrix D

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

  - Anti-diffusive fluxes

$$f_{ij} = d_{ij}(u_j - u_i), \quad f_{ij} = -f_{ji}, \quad i, j = 1, \ldots, N$$

- Derivation (cont.)
  - Solution-dependent coefficients

$$\alpha_{ij} = \alpha_{ji}, \quad i,j = 1,\ldots,N$$

with

$$\alpha_{ij} \in [0,1]$$

- Derivation (cont.)
  - Solution-dependent coefficients

$$\alpha_{ij} = \alpha_{ji}, \quad i,j = 1,\ldots,N$$

with

$$\alpha_{ij} \in [0,1]$$

  - Final scheme

$$\begin{aligned} \sum_{j=1}^{N} a_{ij}u_j + \sum_{j=1}^{N}(1-\alpha_{ij})d_{ij}(u_j - u_i) &= f_i, \quad i = 1,\ldots,M, \\ u_i &= u_i^b, \quad i = M+1\ldots,N \end{aligned}$$

- Limiters
  - Kuzmin limiter [1,2,3]
  - BJK limiter [4]
  - Analytical properties [2,3]
  - BJK limiter in general more accurate[4]

[1] Kuzmin: in Proc. Int. Conf. Comput. Meth. for Coupled Problems in Science and Engineering, CIMNE, 2007

[2] Barrenechea, John, Knobloch: SINUM 54, 2427–2451, 2016

[3] Barrenechea, John, Knobloch, Rankin: SeMA Journal 75, 655-685, 2018

[4] Barrenechea, John, Knobloch: M3AS 27, 525–548, 2017

- Fixed point iteration with changing matrix

$$\begin{aligned} \sum_{j=1}^{N} a_{ij}\, u_j^{(m+1)} + \sum_{j=1}^{N} \left(1 - \alpha_{ij}^{(m)}\right) d_{ij} \left(u_j^{(m+1)} - u_i^{(m+1)}\right) &= f_i, \\ u_i^{(m+1)} &= u_i^b \end{aligned}$$

- Fixed point iteration with fixed matrix: using

$$\sum_{j=1}^{N}(1-\alpha_{ij})d_{ij}(u_j-u_i) = \sum_{j=1}^{N} d_{ij}u_j - u_i \underbrace{\sum_{j=1}^{N} d_{ij}}_{=0} - \sum_{j=1}^{N} \alpha_{ij}d_{ij}(u_j-u_i),$$

gives

$$\begin{aligned} \sum_{j=1}^{N}(a_{ij}+d_{ij})u_j^{(m+1)} &= f_i + \sum_{j=1}^{N} \alpha_{ij}^{(m)} f_{ij}^{(m)}, \quad & i = 1,\ldots,M, \\ u_i^{(m+1)} &= u_i^b, & i = M+1,\ldots,N \end{aligned}$$

- Fixed point iterations
  - Fixed point iteration with fixed matrix (FPR)
    - Matrix is M-matrix
    - With direct sparse solver: factorization is needed only once
  - Fixed point iteration with changing matrix (FPM)
    - More implicit approach, hope for better convergence properties
  - General fixed point iteration by linear combination

$$\begin{aligned} \sum_{j=1}^{N} (a_{ij} + d_{ij})\, u_j^{(m+1)} - \omega_{\mathrm{fp}} \sum_{j=1}^{N} \alpha_{ij}^{(m)} d_{ij} \left(u_j^{(m+1)} - u_i^{(m+1)} \right) & \\ = \quad f_i + (1 - \omega_{\mathrm{fp}}) \sum_{j=1}^{N} \alpha_{ij}^{(m)} f_{ij}^{(m)}, \quad i = 1, \ldots, M, & \\ u_i^{(m+1)} \quad = \quad u_i^b, \quad i = M+1, \ldots, N & \end{aligned}$$

- Formal Newton method
  - Formal derivation of Jacobian

$$
DF\left(u^{(m)}\right)_{ij} = \begin{cases} a_{ij} + d_{ij} - \alpha_{ij}^{(m)} d_{ij} - \sum_{k=1}^{N} \frac{\partial \alpha_{ik}^{(m)}}{\partial u_j} d_{ik} \left(u_k^{(m)} - u_i^{(m)}\right) & \text{if } i \neq j, \\ a_{ii} + d_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^{N} \alpha_{ij}^{(m)} d_{ij} - \sum_{k=1}^{N} \frac{\partial \alpha_{ik}^{(m)}}{\partial u_i} d_{ik} \left(u_k^{(m)} - u_i^{(m)}\right) & \text{if } i = j \end{cases}
$$

- Formal Newton method: how to deal with non-smooth cases?
- Discussion only for Kuzmin limiter
  - Involves maxima and minima of two arguments, one of them is constant
  1. Non-regularized approach
     - Take one-sided derivative w.r.t. constant, i.e., set value to zero
  2. Regularized approach
     - Replace maximum for some $\sigma_N > 0$ by[1]

$$\max_{\sigma_N(x,y)} = \frac{1}{2}\left(x + y + \sqrt{(x-y)^2 + \sigma_N}\right)$$

     - We did not regularize the limiter in the equation, only in the iteration matrix

[1] Badia, Bonilla: CMAME 313, 133–158, 2017

- General form of the matrix

$$\underbrace{\underbrace{\underbrace{a_{ij} + d_{ij}}_{\text{FPR, const. matrix}} - \omega_{\text{fp}} \alpha_{ij} d_{ij}}_{\text{FPM, changing matrix}} + \omega_{\text{jac}} (\text{term with der. of } \alpha_{ij})}_{\text{formal Newton}}, \quad i \neq j$$

  - Similar for diagonal entries
  - Some modifications for regularized Newton approach
- Iteration

$$u^{(m+1)} = u^{(m)} + \omega \left(u^{(m+1)} - u^{(m)} \right)$$

- Algorithmic components
  - Adaptive choice of damping parameter [1]
  - Anderson acceleration [2]
  - Projection to admissible values [3]
  - Selection of initial iterate [4]

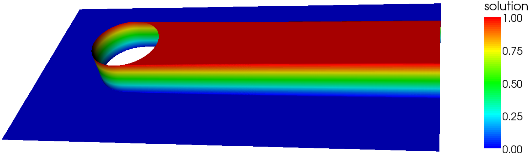
[1] John, Knobloch: CMAME 197, 1997–2014, 2008

[2] Walker, Ni: SINUM 49(4), 1715-1735, 2011

[3] Badia, Bonilla: CMAME 313, 133–158, 2017

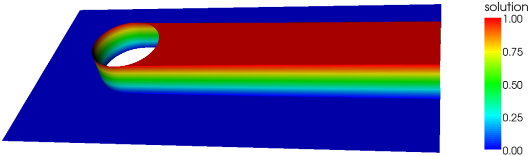
[4] J., John: BAIL 2018- Boundary and Interior Layers, (in press)

- Hemker problem[1]
- Various values of $\varepsilon = 10^{-6}$, $\boldsymbol{b} = (1,0)^T$, $c = f = 0$



[1] Hemker: JCAM 76, 277-285, 1996

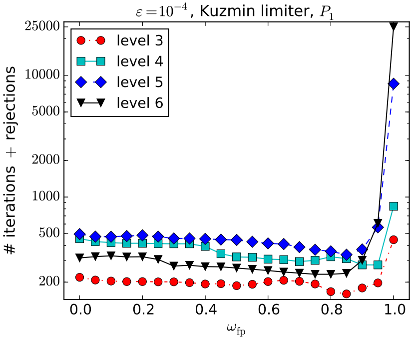
- Hemker problem[1]
- Various values of $\varepsilon = 10^{-6}$, $\boldsymbol{b} = (1,0)^T$, $c = f = 0$

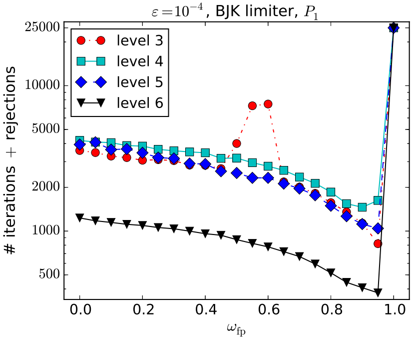


- P_1 finite elements
- Stopping criteria
    - $\|\text{residual}\|_2 \leq \sqrt{\text{\# dof}}\, 10^{-10}$
    - 25000 iterations
- Direct sparse solver (UMFPACK)

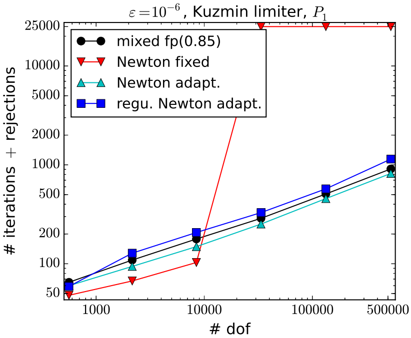
[1] Hemker: JCAM 76, 277-285, 1996

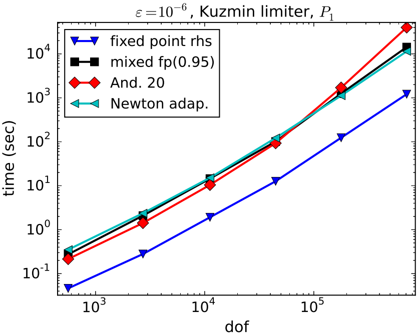
- Hemker problem[1]





[1] Hemker: JCAM 76, 277-285, 1996





- FPR most efficient in 3d

- Variational problem of AFC scheme reads
  Find $u_h \in V_h$ such that

$$a_h(u_h, v_h) + d_h(u_h; u_h, v_h) = \langle f, v_h \rangle \quad \forall\, v_h \in V_h$$

  - V_h- finite element space with homogeneous Dirichlet boundary conditions $(V_h \subset V)$
  - stabilization

$$d_h(w_h; z_h, v_h) = \sum_{i,j=1}^{N} (1 - \alpha_{ij}(w_h)) d_{ij}(z_j - z_i) v_i \quad \forall\, w_h, v_h, z_h \in V_h$$

[1] Barrenechea, John, Knobloch, Rankin: SeMA Journal 75, 655–685, 2018

- Variational problem of AFC scheme reads
  Find $u_h \in V_h$ such that

$$a_h(u_h, v_h) + d_h(u_h; u_h, v_h) = \langle f, v_h \rangle \quad \forall\, v_h \in V_h$$

  - $V_h -$ finite element space with homogeneous Dirichlet boundary conditions $(V_h \subset V)$
  - stabilization

$$d_h(w_h; z_h, v_h) = \sum_{i,j=1}^{N} (1 - \alpha_{ij}(w_h)) d_{ij}(z_j - z_i) v_i \quad \forall\, w_h, v_h, z_h \in V_h$$

- Another representation of stabilization for $w_h, v_h, z_h \in V_h$,[1]

$$d_h(w_h; z_h, v_h) = \sum_{E \in \mathcal{E}_h} (1 - \alpha_E(w_h)) d_E h_E \left(\nabla z_h \cdot \mathbf{t}_E, \nabla v_h \cdot \mathbf{t}_E\right)$$

[1] Barrenechea, John, Knobloch, Rankin: SeMA Journal 75, 655–685, 2018

- AFC norm

$$\|u_h\|_{\text{AFC}}^2 = \|u_h\|_a^2 + d_h(u_h, u_h, u_h) \quad \forall u_h \in V_h$$

  - where $\|u_h\|_a^2 = \varepsilon|u_h|_1^2 + \sigma\|u_h\|_0^2$
- Let $I_h u$ denote an interpolation operator. Galerkin orthogonality arguments

$$\begin{aligned} \|u - u_h\|_{\text{AFC}}^2 &= \langle f, u - I_h u\rangle + \langle g, u - I_h u\rangle_{\Gamma_N} - a(u_h, u - I_h u) \\ &\quad + d_h(u_h; u, I_h u - u_h) \end{aligned}$$

[1] John, Novo: CMAME 255, 289–305, 2013

- AFC norm

$$\|u_h\|_{\text{AFC}}^2 = \|u_h\|_a^2 + d_h(u_h, u_h, u_h) \quad \forall u_h \in V_h$$

  - where $\|u_h\|_a^2 = \varepsilon |u_h|_1^2 + \sigma \|u_h\|_0^2$
- Let $I_h u$ denote an interpolation operator. Galerkin orthogonality arguments

$$\begin{aligned} \|u - u_h\|_{\text{AFC}}^2 &= \langle f, u - I_h u\rangle + \langle g, u - I_h u\rangle_{\Gamma_N} - a(u_h, u - I_h u) \\ &\quad + d_h(u_h; u, I_h u - u_h) \end{aligned}$$

- Standard residual posteriori error bound [1]

$$\begin{aligned} &\langle f, u - I_h u\rangle + \langle g, u - I_h u\rangle_{\Gamma_N} - a(u_h, u - I_h u) \\ &= \sum_{K \in \mathcal{T}_h} (R_K(u_h), u - I_h u)_K + \sum_{F \in \mathcal{F}_h} \langle R_F(u_h), u - I_h u\rangle_F \end{aligned}$$

[1] John, Novo: CMAME 255, 289–305, 2013

with

$$R_K(u_h) := f + \varepsilon\Delta u_h - \boldsymbol{b}\cdot\nabla u_h - cu_h|_K,$$

$$R_F(u_h) := \begin{cases} -\varepsilon[|\nabla u_h \cdot \mathbf{n}_F|]_F & \text{if } F \in \mathcal{F}_{h,\Omega}, \\ g - \varepsilon(\nabla u_h \cdot \mathbf{n}_F) & \text{if } F \in \mathcal{F}_{h,N}, \\ 0 & \text{if } F \in \mathcal{F}_{h,D} \end{cases}$$

Residual Based Approach

with

$$R_K(u_h) := f + \varepsilon\Delta u_h - \boldsymbol{b}\cdot\nabla u_h - cu_h|_K,$$

$$R_F(u_h) := \begin{cases} -\varepsilon[|\nabla u_h\cdot \mathbf{n}_F|]_F & \text{if } F\in\mathcal{F}_{h,\Omega}, \\ g-\varepsilon(\nabla u_h\cdot\mathbf{n}_F) & \text{if } F\in\mathcal{F}_{h,N}, \\ 0 & \text{if } F\in\mathcal{F}_{h,D} \end{cases}$$

- Using interpolation estimates, Cauchy-Schwarz and Young's inequality

$$\begin{aligned} &\|u-u_h\|_a^2 + \frac{C_Y}{C_Y-1}d_h(u_h;u-u_h,u-u_h) \\ \leq\ & \frac{C_Y^2}{2(C_Y-1)}\sum_{K\in\mathcal{T}_h}\min\left\{\frac{C_I^2}{\sigma},\frac{C_I^2h_K^2}{\varepsilon}\right\}\|R_K(u_h)\|_{L^2(K)}^2 \\ &+\frac{C_Y^2}{2(C_Y-1)}\sum_{F\in\mathcal{F}_h}\min\left\{\frac{C_F^2h_F}{\varepsilon},\frac{C_F^2}{\sigma^{1/2}\varepsilon^{1/2}}\right\}\|R_F(u_h)\|_{L^2(F)}^2 \\ &+\frac{C_Y}{C_Y-1}d_h(u_h;u,I_hu-u_h) \end{aligned}$$

- Linearity of $d_h(\ ;\ ,\ )$,

$$d_h(u_h; u, I_h u - u_h) = d_h(u_h; u - u_h, I_h u - u_h) + d_h(u_h; u_h, I_h u - u_h)$$

- Linearity of $d_h(\ ;\ ,\ )$,

$$d_h(u_h;u,I_hu-u_h) = d_h(u_h;u-u_h,I_hu-u_h) + d_h(u_h;u_h,I_hu-u_h)$$

- Using interpolation estimates, Cauchy-Schwarz, trace inequality, inverse estimate, and Young's inequality

$$\begin{aligned} d_h(u_h;u_h,I_hu-u_h) \quad &\leq \quad \frac{C_Y}{2}\sum_{E\in\mathcal{E}_h}\min\left\{\frac{\kappa_1 h_E^2}{\varepsilon},\frac{\kappa_2}{\sigma}\right\}(1-\alpha_E)^2|d_E|^2 h_E^{1-d} \\ &\quad \times\|\nabla u_h\cdot\mathbf{t}_E\|_{L^2(E)}^2 + \frac{1}{C_Y}\|u-u_h\|_a^2, \end{aligned}$$

where

$$\begin{aligned} \kappa_1 &= C_{\text{edge,max}}(1+(1+C_I)^2), \\ \kappa_2 &= C_{\text{inv}}^2 C_{\text{edge,max}}(1+(1+C_I)^2) \end{aligned}$$

Theorem [Global a posteriori error estimate] A global a posteriori error estimate for the energy norm is given by

$$\|u - u_h\|_a^2 \leq \eta_1^2 + \eta_2^2 + \eta_3^2,$$

where

$$\eta_1^2 = \sum_{K \in \mathcal{T}_h} \min\left\{\frac{4C_I^2}{\sigma}, \frac{4C_I^2 h_K^2}{\varepsilon}\right\} \|R_K(u_h)\|_{L^2(K)}^2,$$

$$\eta_2^2 = \sum_{F \in \mathcal{F}_h} \min\left\{\frac{4C_F^2 h_F}{\varepsilon}, \frac{4C_F^2}{\sigma^{1/2}\varepsilon^{1/2}}\right\} \|R_F(u_h)\|_{L^2(F)}^2,$$

$$\eta_3^2 = \sum_{E \in \mathcal{E}_h} \min\left\{\frac{4\kappa_1 h_E^2}{\varepsilon}, \frac{4\kappa_2}{\sigma}\right\} (1 - \alpha_E)^2 |d_E|^2 h_E^{1-d} \|\nabla u_h \cdot \mathbf{t}_E\|_{L^2(E)}^2.$$

- Local lower bound for a mesh cell K

$$\eta_K^2 = \eta_{\mathrm{Int},\mathrm{K}}^2 + \sum_{F \in \mathcal{F}_h(K)} \eta_{\mathrm{Face},F}^2 + \sum_{E \in \mathcal{E}_h(K)} \eta_{d_h,E}^2$$

- Local lower bound for a mesh cell K

$$\eta_K^2 = \eta_{\text{Int,K}}^2 + \sum_{F\in\mathcal{F}_h(K)} \eta_{\text{Face},F}^2 + \sum_{E\in\mathcal{E}_h(K)} \eta_{d_h,E}^2$$

where

$$\begin{aligned}
\eta_{\text{Int},K}^2 &= \min\left\{\frac{4C_I^2}{\sigma}, \frac{4C_I^2 h_K^2}{\varepsilon}\right\} \|R_{K,h}(u_h)\|_{L^2(K)}^2, \\
\eta_{\text{Face},F}^2 &= \frac{1}{2}\min\left\{\frac{4C_F^2 h_F}{\varepsilon}, \frac{4C_F^2}{\sigma_0^{1/2}\varepsilon^{1/2}}\right\} \|R_F(u_h)\|_{L^2(F)}^2, \\
\eta_{d_h,E}^2 &= \min\left\{\frac{4\kappa_1 h_E^2}{\varepsilon}, \frac{4\kappa_2}{\sigma}\right\}(1-\alpha_E)^2|d_E|^2 h_E^{1-d}\|\nabla u_h\cdot \mathbf{t}_E\|_{L^2(E)}^2
\end{aligned}$$

- Using standard bubble function arguments

$$\eta_{\mathrm{Int},K} \leq C\Bigg(\max\left\{C_K^2 + \frac{C_K h_K}{\varepsilon}\|\mathbf{b}\|_{L^\infty(K)}, \frac{C_K}{\sigma}\|c\|_{L^\infty(K)}\right\}\|u-u_h\|_{a,K}$$
$$+\frac{h_K}{\varepsilon^{1/2}}C_K\Big(\|f-f_h\|_{0,K} + \|(\mathbf{b}-\mathbf{b}_h)\cdot\nabla u_h\|_{0,K} + \|(c-c_h)u_h\|_{0,K}\Big)\Bigg)$$

- Using standard bubble function arguments

$$\eta_{\text{Int},K} \leq C\Bigg(\max\left\{C_K^2 + \frac{C_K h_K}{\varepsilon}\|\mathbf{b}\|_{L^\infty(K)}, \frac{C_K}{\sigma}\|c\|_{L^\infty(K)}\right\}\|u-u_h\|_{a,K}$$
$$+\frac{h_K}{\varepsilon^{1/2}}C_K\Big(\|f-f_h\|_{0,K} + \|(\mathbf{b}-\mathbf{b}_h)\cdot\nabla u_h\|_{0,K} + \|(c-c_h)u_h\|_{0,K}\Big)\Bigg)$$

and

$$\eta_{\text{Face},F} \leq C\Bigg(\max\left\{C_F + \frac{C_F h_F\|\mathbf{b}\|_{L^\infty(\omega_F)}}{\varepsilon}, \frac{C_F h_F\|c\|_{L^\infty(\omega_F)}}{\varepsilon^{1/2}\sigma^{1/2}}\right\}\|u-u_h\|_{a,\omega_F}$$
$$+\delta_{F\in\mathcal{F}_{h,N}}\frac{h_F^{1/2}}{\varepsilon^{1/2}}\|g-g_h\|_{L^2(F)}$$
$$+\sum_{K\in\omega_F}\Big[\eta_{\text{Int},K} + \frac{h_K}{\varepsilon^{1/2}}\big(\|f-f_h\|_{0,K} + \|(\mathbf{b}-\mathbf{b}_h)\cdot\nabla u_h\|_{0,K}$$
$$+\|(c-c_h)u_h\|_{0,K}\big)\Big]\Bigg)$$

- For the stabilization term, from [1] we get

$$|d_E| \leq C\left(\varepsilon + \|\mathbf{b}\|_{L^\infty(\Omega)} h + \|c\|_{L^\infty(\Omega} h^2\right) h_E^{d-2}$$

[1] Barrenechea, John, Knobloch, Rankin: SeMA Journal 75, 655–685, 2018

- For the stabilization term, from [1] we get

$$|d_E| \leq C \left(\varepsilon + \|\mathbf{b}\|_{L^\infty(\Omega)} h + \|c\|_{L^\infty(\Omega} h^2\right) h_E^{d-2}$$

Hence,

$$\eta_{d_h,E} \leq C \sum_{E \in \mathcal{E}_h} (1-\alpha_E) \left(\varepsilon + \|\mathbf{b}\|_{L^\infty(\Omega)} h + \|c\|_{L^\infty(\Omega)} h^2\right) \times \frac{h_E^{(3-d)/2}}{\varepsilon^{1/2}} \|\nabla u_h \cdot \mathbf{t}_E\|_{L^2(E)}$$

[1] Barrenechea, John, Knobloch, Rankin: SeMA Journal 75, 655–685, 2018

Theorem [Local lower bound] There exists a constant $C > 0$, independent of the size of elements of $\mathcal{T}$, such that, for every $K \in \mathcal{T}$, the following local lower bound holds

$$\eta_{\text{Int},K} + \sum_{K \in \mathcal{F}_h(K)} \eta_{\text{Face},F} + \sum_{E \in \mathcal{E}_h(K)} \eta_{d_h,E}$$

$$\leq \max\left\{ C_K^2 + \frac{C_K h_K}{\varepsilon} \|\mathbf{b}\|_{L^\infty(K)}, \frac{C_K}{\sigma_0} \|c\|_{L^\infty(K)} \right\} \|u - u_h\|_{a,\omega_K}$$

$$+ C \sum_{K \in \omega_K} \frac{h_K}{\varepsilon^{1/2}} \left(\|f - f_h\|_{0,K} + \|(\mathbf{b} - \mathbf{b}_h) \cdot \nabla u_h\|_{0,K} + \|(c - c_h) u_h\|_{0,K} \right)$$

$$+ C \sum_{F \in \mathcal{F}_h(K)} \delta_{F \in \mathcal{F}_{h,N}} \frac{h_F^{1/2}}{\varepsilon^{1/2}} \|g - g_h\|_{L^2(F)}$$

$$+ \sum_{E \in \mathcal{E}_h(K)} h^{1-d/2} \frac{h^{1/2}}{\varepsilon^{1/2}} \left(\varepsilon + \|b\|_{L^\infty(\Omega)} h + \|c\|_{L^\infty(\Omega)} h^2 \right) \|\nabla u_h \cdot t_E\|_{L^2(E)}.$$

- The initial solution for the nonlinear loop is the SUPG solution [1]
  - $u_{\text{AFC}} :=$ AFC solution
  - $u_{\text{SUPG}} :=$ SUPG solution
- By triangle inequatliy

$$\|u - u_{\text{AFC}}\|_a^2 \leq 2\left(\|u - u_{\text{SUPG}}\|_a^2 + \|u_{\text{SUPG}} - u_{\text{AFC}}\|_a^2\right)$$
$$\leq 2\left(\|u - u_{\text{SUPG}}\|_{\text{SUPG}}^2 + \|u_{\text{SUPG}} - u_{\text{AFC}}\|_a^2\right)$$

[1] J.John, BAIL 2018 (135), 2020

[2] John, Novo: CMAME, 255, 289-305, 2013

AFC-SUPG Approach

- The initial solution for the nonlinear loop is the SUPG solution [1]
  - $u_{\text{AFC}} :=$ AFC solution
  - $u_{\text{SUPG}} :=$ SUPG solution
- By triangle inequatliy

$$\begin{aligned}\|u - u_{\text{AFC}}\|_a^2 &\leq 2\left(\|u - u_{\text{SUPG}}\|_a^2 + \|u_{\text{SUPG}} - u_{\text{AFC}}\|_a^2\right)\\ &\leq 2\left(\|u - u_{\text{SUPG}}\|_{\text{SUPG}}^2 + \|u_{\text{SUPG}} - u_{\text{AFC}}\|_a^2\right)\end{aligned}$$

- Using estimators from [2]

$$\|u - u_{\text{SUPG}}\|_{\text{SUPG}}^2 \leq \eta_{\text{SUPG}}^2$$

[1] J.John, BAIL 2018 (135), 2020
[2] John, Novo: CMAME, 255, 289-305, 2013

AFC-SUPG Approach

- The initial solution for the nonlinear loop is the SUPG solution [1]
  - $u_{\mathrm{AFC}} :=$ AFC solution
  - $u_{\mathrm{SUPG}} :=$ SUPG solution
- By triangle inequatliy

$$\begin{aligned}\|u - u_{\mathrm{AFC}}\|_a^2 &\leq 2\left(\|u - u_{\mathrm{SUPG}}\|_a^2 + \|u_{\mathrm{SUPG}} - u_{\mathrm{AFC}}\|_a^2\right)\\ &\leq 2\left(\|u - u_{\mathrm{SUPG}}\|_{\mathrm{SUPG}}^2 + \|u_{\mathrm{SUPG}} - u_{\mathrm{AFC}}\|_a^2\right)\end{aligned}$$

- Using estimators from [2]

$$\|u - u_{\mathrm{SUPG}}\|_{\mathrm{SUPG}}^2 \leq \eta_{\mathrm{SUPG}}^2$$

  and denoting

$$\eta_{\mathrm{AFC-SUPG}}^2 := \|u_{\mathrm{SUPG}} - u_{\mathrm{AFC}}\|_a^2$$

$\Rightarrow$

$$\|u - u_h\|_a^2 \leq 2\left(\eta_{\mathrm{SUPG}}^2 + \eta_{\mathrm{AFC-SUPG}}^2\right)$$

[1] J.John, BAIL 2018 (135), 2020

[2] John, Novo: CMAME, 255, 289-305, 2013

- Standard strategy for solving

SOLVE $\rightarrow$ **ESTIMATE** $\rightarrow$ **MARK** $\rightarrow$ **REFINE**

- Effectivity index for the estimator

$$\eta_{\text{eff}} = \frac{\eta}{\|u - u_h\|_a}$$

- Standard strategy for solving

SOLVE → ESTIMATE → MARK → REFINE

- Effectivity index for the estimator

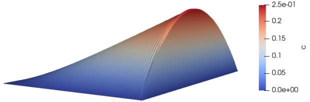
$$\eta_{\text{eff}} = \frac{\eta}{\|u - u_h\|_a}$$

- Comparison of results:
  - On effectivity index
  - Adaptive grid refinement
  - Behavior of η_{d_h}
  - Behavior of η_{SUPG} and $\eta_{\text{AFC-SUPG}}$

Numerical Studies

- $\varepsilon = 10^{-3}$, $\boldsymbol{b} = (2,1)^T$, $c = 1$ and f such that

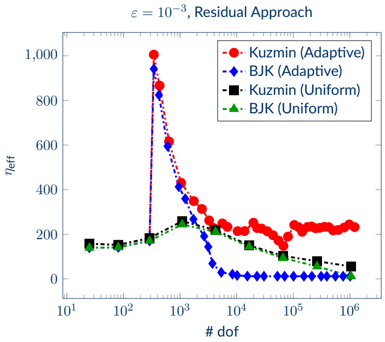
$$u(x,y) = y(1-y)\left(x - \frac{e^{(x-1)/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}}\right)$$

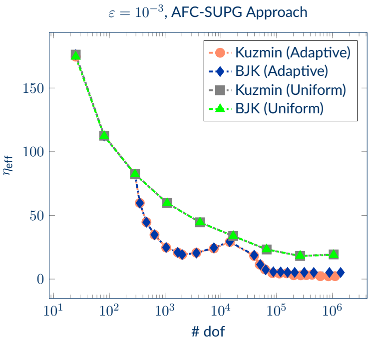


- proposed in [1]
- P_1 finite elements
- stop of the adaptive algorithm
  - $\eta \leq 10^{-3}$
  - #nDOFs $\approx 10^6$

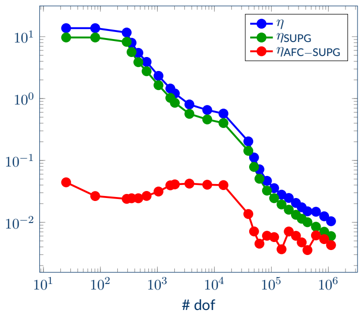
[1] Allendes et. al, SISC, 39(5):A1903-A1927, 2017

- Effectivity index

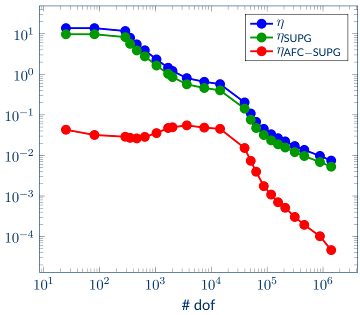




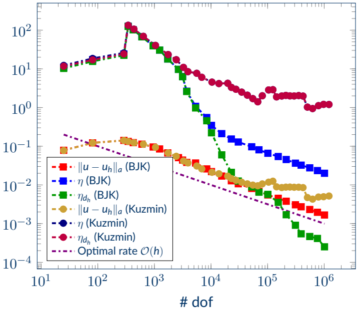
$\varepsilon = 10^{-3}$, Kuzmin limiter



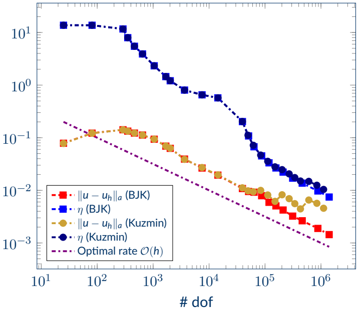
$\varepsilon = 10^{-3}$, BJK limiter



$\varepsilon = 10^{-3}$, Residual Approach[1]



$\varepsilon = 10^{-3}$, AFC-SUPG Approach



[1] Barrenechea, John, Knobloch: SINUM 54, 2427–2451, 2016

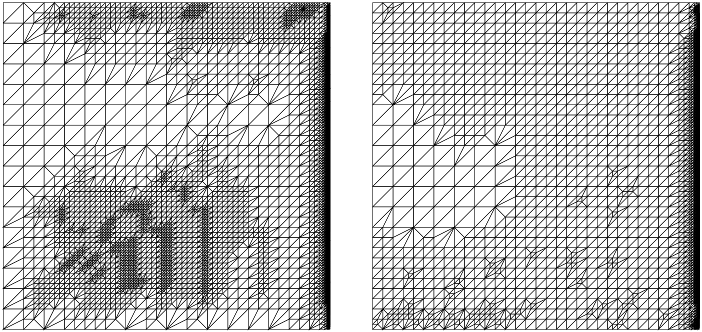


Figure 3: 14^{th} adaptively refined grid with residual based approach. Kuzmin limiter (#nDOFs $= 22962$) (left) and BJK limiter (#nDOFs $= 23572$) (right)

- Hanging nodes
  - Preserve angles after red-refinement
  - Avoid prism and pyramids in 3D mesh refinement
- Certain stabilized schemes rely on the properties of triangulation[1]

[1] Xu, Zikatanov, MC, 68(228):1429-1446, 1999

[2] John, Knobloch, CMAME, 196(17-20):2197-2215, 2007

- Hanging nodes
  - Preserve angles after red-refinement
  - Avoid prism and pyramids in 3D mesh refinement
- Certain stabilized schemes rely on the properties of triangulation[1]
- Discrete Maximum Principle

$$\mathrm{Var}(u_h) = u_h^{\max} - u_h^{\min}$$

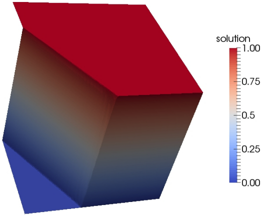
- Thickness of Interior Layer[2]

[1] Xu, Zikatanov, MC, 68(228):1429-1446, 1999

[2] John, Knobloch, CMAME, 196(17-20):2197-2215, 2007

- Example with interior and boundary layer[1]
- $\varepsilon = 10^{-4}$, $\boldsymbol{b} = (\cos(-\pi/3), \sin(-\pi/3))^T$, $c = f = 0$

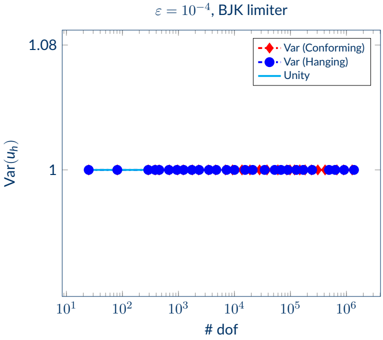
$$u_b = \begin{cases} 1 & (y = 1 \wedge x > 0) \text{ or } (x = 0 \wedge y > 0.7), \\ 0 & \text{else.} \end{cases}$$

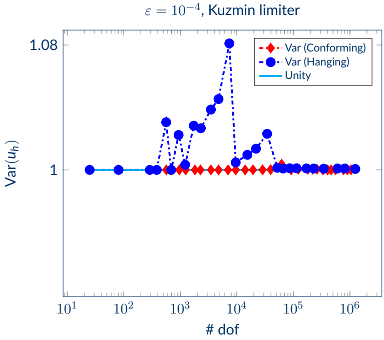


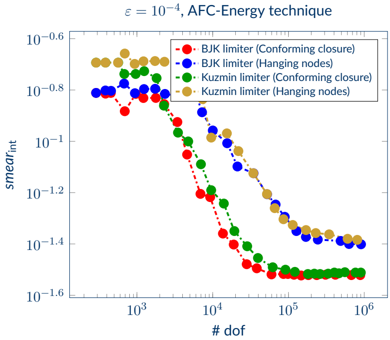
- P_1 finite elements

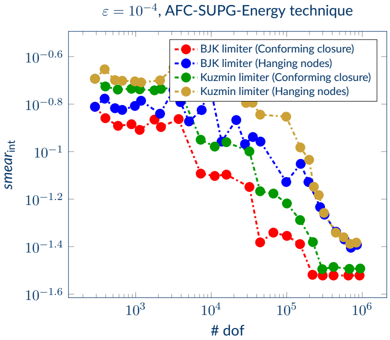
[1] Hughes, Mallet, Mizukami, CMAME, 54(3):341-345, 1986

- $\mathrm{Var}(u_h)$






$\varepsilon = 10^{-4}$, AFC-Energy technique
$10^{-0.6}$
$10^{-0.8}$
10^{-1}
$10^{-1.2}$
$10^{-1.4}$
$10^{-1.6}$
$smear_{int}$
10^3
10^4
10^5
10^6
dof
BJK limiter (Conforming closure)
BJK limiter (Hanging nodes)
Kuzmin limiter (Conforming closure)
Kuzmin limiter (Hanging nodes)


$\varepsilon = 10^{-4}$, AFC-SUPG-Energy technique
$10^{-0.6}$
$10^{-0.8}$
10^{-1}
$10^{-1.2}$
$10^{-1.4}$
$10^{-1.6}$
$smear_{int}$
10^3
10^4
10^5
10^6
dof
BJK limiter (Conforming closure)
BJK limiter (Hanging nodes)
Kuzmin limiter (Conforming closure)
Kuzmin limiter (Hanging nodes)

- Iterative Solvers[1,2]
    - With appropriate parameters some gain w.r.t. number of iterations
    - Easier to solve problems for Kuzmin limiter than for BJK limiter
    - With an appropriate solver: FPR is usually most efficient

---

[1] J., John: CAMWA 78, 3117-3138, 2019

[2] J., John: BAIL 2018, Springer 113-128, 2020

[3] J.: arXiv , 2005.02938, 2020

[4] J.: arXiv , 2007.08405 , 2020

Conclusions

- Iterative Solvers[1,2]
  - With appropriate parameters some gain w.r.t. number of iterations
  - Easier to solve problems for Kuzmin limiter than for BJK limiter
  - With an appropriate solver: FPR is usually most efficient
- A Posteriori Error Analysis[3]
  - Effectivity index not robust with AFC-Energy technique
  - For the AFC-SUPG-Energy technique, the effectivity index was better
  - With adaptive grid refinement, problem could become locally diffusion-dominated. Then use BJK limiter
  - For a small diffusion coefficient, use Kuzmin limiter

---

[1] J., John: CAMWA 78, 3117-3138, 2019

[2] J., John: BAIL 2018, Springer 113-128, 2020

[3] J.: arXiv , 2005.02938, 2020

[4] J.: arXiv , 2007.08405 , 2020

Conclusions

- Iterative Solvers[1,2]
  - With appropriate parameters some gain w.r.t. number of iterations
  - Easier to solve problems for Kuzmin limiter than for BJK limiter
  - With an appropriate solver: FPR is usually most efficient
- A Posteriori Error Analysis[3]
  - Effectivity index not robust with AFC-Energy technique
  - For the AFC-SUPG-Energy technique, the effectivity index was better
  - With adaptive grid refinement, problem could become locally diffusion-dominated. Then use BJK limiter
  - For a small diffusion coefficient, use Kuzmin limiter
- Hanging Nodes[4]
  - Kuzmin limiter fails to satisfy the DMP on grids with hanging nodes
  - BJK limiter satisfies the DMP on grids with hanging nodes
  - Layers were better resolved on grids with conforming closure
  - Numerical results on grids with hanging nodes not satisfactory

---

[1] J., John: CAMWA 78, 3117-3138, 2019

[2] J., John: BAIL 2018, Springer 113-128, 2020

[3] J.: arXiv , 2005.02938, 2020

[4] J.: arXiv , 2007.08405 , 2020

Future Work

- Development of robust estimators
- Extending the analysis for local lower bound
- Stability and convergence analysis for Flux Corrected Transport (FCT) schemes for time-dependent convection-diffusion-reaction equations
- Efficient solution of the nonlinear problems in (FCT) schemes