

Variational stabilization of convection-dominated diffusion with optimal test functions

VM Calo

Applied Geology, Curtin University, Perth, Australia
Minerals, Commonwealth Scientific and Industrial Research Organisation
(CSIRO), Australia

Collaborators: A Romkes, E Valseth

Ritz-Galerkin Method

Best Approximation Property (Projection Principle)

- If a symmetric, real-valued bilinear form & *elliptic* on Hilbert space U

$$a(\mathbf{v}, \mathbf{v}) \geq \gamma \|\mathbf{v}\|_U^2 \quad \forall \mathbf{v} \in U$$

it defines an inner product and a norm in U :

$$\|\mathbf{v}\| = \sqrt{a(\mathbf{v}, \mathbf{v})} \quad \forall \mathbf{v} \in U$$

Ritz-Galerkin Method

Best Approximation Property (Projection Principle)

- If a symmetric, real-valued bilinear form & *elliptic* on Hilbert space U

$$a(\mathbf{v}, \mathbf{v}) \geq \gamma \|\mathbf{v}\|_U^2 \quad \forall \mathbf{v} \in U$$

it defines an inner product and a norm in U :

$$\|\mathbf{v}\| = \sqrt{a(\mathbf{v}, \mathbf{v})} \quad \forall \mathbf{v} \in U$$

- Then, the Ritz-Galerkin approximation \mathbf{u}_h of $\mathbf{u} \in U$ in a finite-dimensional subspace $U_h \subset U$ satisfies

$$a(\mathbf{u}_h, \mathbf{v}_h) = a(\mathbf{u}, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in U_h$$

which is the *orthogonal projection* of \mathbf{u} onto U_h :

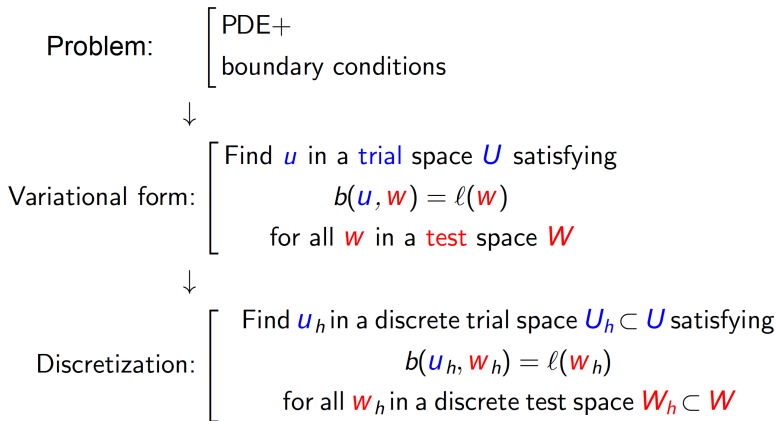
$$a(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in U_h$$

- Thus, it is the *best approximation* of \mathbf{u} in U_h :

$$\|\mathbf{u} - \mathbf{u}_h\| = \min_{\mathbf{v}_h \in U_h} \|\mathbf{u} - \mathbf{v}_h\|$$

“Petrov-Galerkin” schemes (PG)

PG schemes define different **trial** and **test** (Hilbert) spaces



For PG schemes, $U_h \neq W_h$ in general

Optimal Petrov-Galerkin Method

- Consider a general variational problem

$$\text{Find } \mathbf{u} \in U \quad \text{s.t.} \quad b(\mathbf{u}, \mathbf{w}) = l(\mathbf{w}) \quad \forall \mathbf{w} \in W$$

and let $U_h \subset U$ be a finite-dimensional trial subspace

$$U_h = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$$

Optimal Petrov-Galerkin Method

- Consider a general variational problem

$$\text{Find } \mathbf{u} \in U \quad \text{s.t.} \quad b(\mathbf{u}, \mathbf{w}) = l(\mathbf{w}) \quad \forall \mathbf{w} \in W$$

and let $U_h \subset U$ be a finite-dimensional trial subspace

$$U_h = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$$

- The corresponding *optimal test space* is defined as

$$W_h^{\text{opt}} = \text{span}\{\mathbf{T}\mathbf{e}_1, \dots, \mathbf{T}\mathbf{e}_N\} \subset W$$

where the trial-to-test map $\mathbf{T} : U \rightarrow W$ is defined through

$$(\mathbf{T}\mathbf{u}, \delta\mathbf{w})_W = b(\mathbf{u}, \delta\mathbf{w}) \quad \forall \delta\mathbf{w} \in W$$

- Let each $\mathbf{w}_h \in W_h^{\text{opt}}$ be $\mathbf{w}_h = \mathbf{T}\mathbf{v}_h$ for some $\mathbf{v}_h \in U_h$, then

$$b(\mathbf{u}_h, \mathbf{w}_h) = b(\mathbf{u}_h, \mathbf{T}\mathbf{w}_h) = (\mathbf{T}\mathbf{u}_h, \mathbf{T}\mathbf{w}_h)_W \stackrel{\text{def}}{=} a(\mathbf{u}_h, \mathbf{v}_h)$$

$$l(\mathbf{w}_h) = l(\mathbf{T}\mathbf{v}_h) \stackrel{\text{def}}{=} Q(\mathbf{v}_h)$$

Optimal Petrov-Galerkin Method (cont.)

- Optimal PG delivers best approximation in generalized energy norms:

$$\|u\|^2 = a(u, u) = (\mathbf{T}u, \mathbf{T}u)_W = \|\mathbf{T}u\|_W^2 = \left\{ \sup_{\mathbf{w} \in W} \frac{|b(u, \mathbf{w})|}{\|\mathbf{w}\|_W} \right\}^2$$

- Ellipticity of $a(u, u)$ induces the inf-sup condition on $b(u, v)$

Optimal Petrov-Galerkin Method (cont.)

- Optimal PG delivers best approximation in generalized energy norms:

$$\| \mathbf{u} \|_{\mathbf{u}}^2 = a(\mathbf{u}, \mathbf{u}) = (\mathbf{T}\mathbf{u}, \mathbf{T}\mathbf{u})_W = \| \mathbf{T}\mathbf{u} \|_W^2 = \left\{ \sup_{\mathbf{w} \in W} \frac{|b(\mathbf{u}, \mathbf{w})|}{\| \mathbf{w} \|_W} \right\}^2$$

- Ellipticity of $a(\mathbf{u}, \mathbf{u})$ induces the inf-sup condition on $b(\mathbf{u}, \mathbf{v})$
- Q1: How to determine optimal test function space in a practice?
- A1: Computed (almost) *automatically* within DPG framework!

Optimal Petrov-Galerkin Method (cont.)

- Optimal PG delivers best approximation in generalized energy norms:

$$|||\mathbf{u}|||^2 = a(\mathbf{u}, \mathbf{u}) = (\mathbf{T}\mathbf{u}, \mathbf{T}\mathbf{u})_W = \|\mathbf{T}\mathbf{u}\|_W^2 = \left\{ \sup_{\mathbf{w} \in W} \frac{|b(\mathbf{u}, \mathbf{w})|}{\|\mathbf{w}\|_W} \right\}^2$$

- Ellipticity of $a(\mathbf{u}, \mathbf{u})$ induces the inf-sup condition on $b(\mathbf{u}, \mathbf{v})$
- Q1: How to determine optimal test function space in a practice?
- A1: Computed (almost) *automatically* within DPG framework!
- Q2: What if $|||\cdot|||$ has harmful parameter dependencies?
- A2: We can have $|||\cdot||| = \|\cdot\|_U$ if we select

$$\|\mathbf{w}\|_W = |||\mathbf{w}|||_{W,\text{opt}} = \sup_{\mathbf{u} \in U} \frac{|b(\mathbf{u}, \mathbf{w})|}{\|\mathbf{u}\|_U}$$

Motivation

Stabilized FEMs resolve the numerical instability issue.

- 1 Least squares FEM (LSFEM)
- 2 Galerkin method with least squares (G/LS)
- 3 Streamlined-upwind Petrov-Galerkin (SUPG) method
- 4 Variational multi-scale (VMS) method
- 5 Discontinuous Petrov-Galerkin method (DPG)

1 – Overly diffuse solutions on coarse meshes, limits $f \in L^2(\Omega)$

2-4 – Requires fine-tuning of penalty/stabilization parameters

5 – Introduces additional degrees of freedom (DOFs)

Target Problem

Convection Dominated Diffusion Problems

Find u such that:

$$-\nabla \cdot \kappa \nabla u + \mathbf{a} \cdot \nabla u = f, \text{ in } \Omega \subset \mathbb{R}^{2,3}$$

Boundary conditions:

$$u = 0, \text{ on } \Gamma_D,$$

$$\kappa \nabla u \cdot \mathbf{n} = g, \text{ on } \Gamma_N.$$

- $\xi \cdot \kappa \xi > 0, \quad \forall \xi \neq 0$
- $\|\kappa\|_{L^\infty(\Omega)} \ll \|\mathbf{a}\|_{L^\infty(\Omega)}$

Proposed Approach

- Derive integral statement on broken Hilbert spaces:
 - ▶ Test functions in $L^2(\Omega)$ with local higher regularity
⇒ Reduced local regularity requirements on f
- Compute 'optimal' discontinuous test functions that automatically deliver discrete stability
- Reduce DOF number in discrete approximation
 - ▶ Apply Petrov-Galerkin framework
 - ▶ Employ classical $C^0(\Omega)$ trial/solution basis functions
 - ▶ Incorporate piecewise **discontinuous** test functions
- Use first order system description (mixed form)
⇒ enforces normal flux continuity in heterogeneous media

Derivation of the Weak Statement

Step 1: Rewrite BVP in mixed form

Find $(u, \boldsymbol{\sigma})$ such that:

$$\boldsymbol{\kappa} \nabla u - \boldsymbol{\sigma} = \mathbf{0}, \text{ in } \Omega$$

$$-\nabla \cdot \boldsymbol{\sigma} + \mathbf{a} \cdot \nabla u = f, \text{ in } \Omega$$

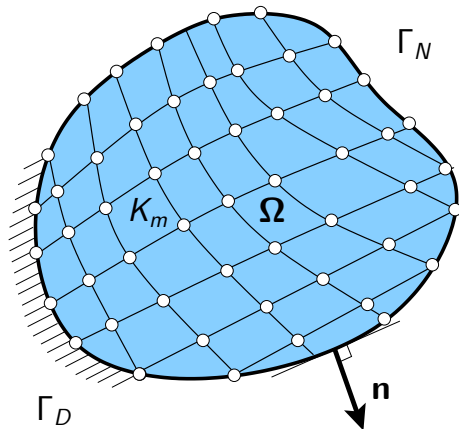
Boundary conditions:

$$u = 0, \text{ on } \Gamma_D$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = g, \text{ on } \Gamma_N$$

Derivation of the Weak Statement

Domain Partition $\mathcal{P}_h = \{K_m\}_{m=1}^{N_{\text{elm}}}$



Derivation of the Weak Statement

Step 2: On each element K_m , enforce PDE weakly, i.e.,

Find $(u, \boldsymbol{\sigma})$ such that, $\forall (v, \mathbf{w}) \in L^2(K_m) \times [L^2(K_m)]^2$

$$\int_{K_m} \left\{ [\boldsymbol{\kappa} \nabla u - \boldsymbol{\sigma}] \cdot \mathbf{w} - [\nabla \cdot \boldsymbol{\sigma} + (\mathbf{a} \cdot \nabla u)] v \right\} d\mathbf{x} = \int_{K_m} \{f v\} d\mathbf{x}$$

Derivation of the Weak Statement

Step 3: Apply Green's Identity

Find $(u, \boldsymbol{\sigma})$ such that, $\forall (v, \mathbf{w}) \in H^1(K_m) \times [L^2(K_m)]^2$

$$\int_{K_m} \left\{ (\kappa \nabla u - \boldsymbol{\sigma}) \cdot \mathbf{w} + \boldsymbol{\sigma} \cdot \nabla v - (\mathbf{a} \cdot \nabla v) u \right\} dx \\ + \oint_{\partial K_m} \left\{ (\mathbf{a} \cdot \mathbf{n}) u v - (\boldsymbol{\sigma} \cdot \mathbf{n}) v \right\} ds = \int_{K_m} \{ f v \} dx$$

Derivation of the Weak Statement

Step 4: Apply boundary & continuity conditions, sum local statements:

Find $(u, \boldsymbol{\sigma}) \in U(\mathcal{P}_h)$ such that:

$$B((u, \boldsymbol{\sigma}), (v, \mathbf{w})) = F(v), \quad \forall (v, \mathbf{w}) \in V(\mathcal{P}_h)$$

where

$$\begin{aligned} B((u, \boldsymbol{\sigma}), (v, \mathbf{w})) &= \sum_{K_m \in \mathcal{P}_h} \left[\int_{K_m} \left\{ (\boldsymbol{\kappa} \nabla u - \boldsymbol{\sigma}) \cdot \mathbf{w} + (\boldsymbol{\sigma} - u \mathbf{a}) \cdot \nabla v \right\} dx \right. \\ &\quad \left. - \int_{\partial K_m \setminus \partial \Omega} \left\{ (\mathbf{a} \cdot \mathbf{n}) u_{\text{neigh}} v - (\boldsymbol{\sigma}_{\text{neigh}} \cdot \mathbf{n}) v \right\} ds \right] \\ F(v) &= \sum_{K_m \in \mathcal{P}_h} \left[\int_{K_m} \{ f v \} dx + \int_{\partial K_m \cap \Gamma_N} \{ g v \} ds \right] \end{aligned}$$

Derivation of the Weak Statement

Discrete Spaces

$$U(\mathcal{P}_h) \stackrel{\text{def}}{=} \left\{ (u, \boldsymbol{\sigma}) \in H^1(\mathcal{P}_h) \times [L^2(\Omega)]^2 : \boldsymbol{\sigma}|_{K_m} \in H(\text{div}, K_m) \right. \\ \left. \wedge \gamma^0(u|_{K_m})|_{\partial K_m \cap \Gamma_D} = 0, \forall K_m \in \mathcal{P}_h \right\}$$

$$V(\mathcal{P}_h) \stackrel{\text{def}}{=} \left\{ (v, \mathbf{w}) \in H^1(\mathcal{P}_h) \times [L^2(\Omega)]^2 : \right. \\ \left. \gamma^0(v|_{K_m})|_{\partial K_m \cap \Gamma_D} = 0, \forall K_m \in \mathcal{P}_h \right\}$$

where:

$$H^1(\mathcal{P}_h) \stackrel{\text{def}}{=} \left\{ v \in L^2(\Omega) : v_m \in H^1(K_m), \forall K_m \in \mathcal{P}_h \right\}$$

FE Discretization: Trial Space Proposition

Find $(u_h, \boldsymbol{\sigma}_h) \in U^h(\mathcal{P}_h)$ such that:

$$B((u_h, \boldsymbol{\sigma}_h), (v_h, \mathbf{w}_h)) = F(v_h), \quad \forall (v_h, \mathbf{w}_h) \in V^*(\mathcal{P}_h)$$

Trial space $U^h(\mathcal{P}_h) \subset U(\mathcal{P}_h)$ consists of piecewise **continuous** polynomials of degree p , that is:

$$u_h(\mathbf{x}) = \sum_{i=1}^N u_i^h e_i(\mathbf{x}), \quad \boldsymbol{\sigma}_h(\mathbf{x}) = \begin{Bmatrix} q_x^h(\mathbf{x}) \\ q_y^h(\mathbf{x}) \end{Bmatrix} = \sum_{j=1}^N \begin{Bmatrix} q_{x_j}^h \varepsilon_{x_j}(\mathbf{x}) \\ q_{y_j}^h \varepsilon_{y_j}(\mathbf{x}) \end{Bmatrix}$$

FE Discretization: Test Space Construction

For each trial function $e_i, \varepsilon_{x_j}, \varepsilon_{y_k}$, compute 'optimal' test functions (v_i, \mathbf{w}_i) , (v_j, \mathbf{w}_j) , & (v_k, \mathbf{w}_k) , resp.:

$$((p, \mathbf{r}), (v_i, \mathbf{w}_i))_{V(\mathcal{P}_h)} = B((e_i, \mathbf{0}), (p, \mathbf{r})), \quad \forall (p, \mathbf{r})$$

$$((p, \mathbf{r}), (v_j, \mathbf{w}_j))_{V(\mathcal{P}_h)} = B((0, (\varepsilon_{x_j}, 0)), (p, \mathbf{r})), \quad \forall (p, \mathbf{r})$$

$$((p, \mathbf{r}), (v_k, \mathbf{w}_k))_{V(\mathcal{P}_h)} = B((0, (0, \varepsilon_{y_k})), (p, \mathbf{r})), \quad \forall (p, \mathbf{r})$$

LHS is **inner product** on $V(\mathcal{P}_h) \times V(\mathcal{P}_h)$, i.e.,

$$((v, \mathbf{w}), (p, \mathbf{r}))_{V(\mathcal{P}_h)} \stackrel{\text{def}}{=} \sum_{K_m \in \mathcal{P}_h} \left\{ \int_{K_m} [h_m^2 \nabla v \nabla p + v p + \mathbf{w} \cdot \mathbf{r}] \, d\mathbf{x} \right\}$$

FE Discretization

Remarks:

- Integral statements governing optimal test functions are infinite-dimensional problems \Rightarrow **test functions computed numerically**
- C^0 trial functions imply that support of test functions is identical to trial functions and solved locally, i.e., **element-by-element solution**
- Optimal test functions span subspace $V^*(\mathcal{P}_h) \subset V(\mathcal{P}_h)$
 \Rightarrow used in FE computation of (u_h, σ_h) .
- (DPG argument) Optimal test functions imply discrete stability

$$B_{ij} = B((e_i, \mathbf{0}), (v_j, \mathbf{w}_j)) = ((v_j, \mathbf{w}_j), (v_i, \mathbf{w}_i))_{V(\mathcal{P}_h)}.$$

FE Discretization: Summary

Continuous-discontinuous Petrov-Galerkin (cDPG) method

- Introduction of a hybrid FE method:
 - ▶ [Petrov-Galerkin Framework]
 - ★ Continuous trial functions (Classical C^0 piecewise polynomials)
 - ★ Discontinuous test functions
 - ▶ [DPG] Optimal test functions \Rightarrow unconditional discrete stability
 - ▶ [Mixed FEMs/LSFEM] First order systems
- Discrete stability guaranteed without calibrating coefficients
- Comparison to LSFEM
 - ▶ Weaker regularity required on source and Neumann data
 - ▶ Sufficient for f to be in the dual of $H^1(\mathcal{P}_h)$ and $g \in H^{-1/2}(\Gamma_N)$

Preliminary Numerical Results

- Test functions may use the same local polynomial degree of approximation as trial/solution functions
- Optimal h -convergence rates for $\|u\|_{H^1(\Omega)}$, $\|u\|_{L^2(\Omega)}$ & $\|\sigma\|_{L^2(\Omega)}$
- Comparison alternative methods
 - ▶ Convection-dominated diffusion problem
 - ▶ 'Shock' problem

Comparison Study - cDPG vs. Other Methods

Model problem: convection diffusion problem on the unit square:

Find u such that:

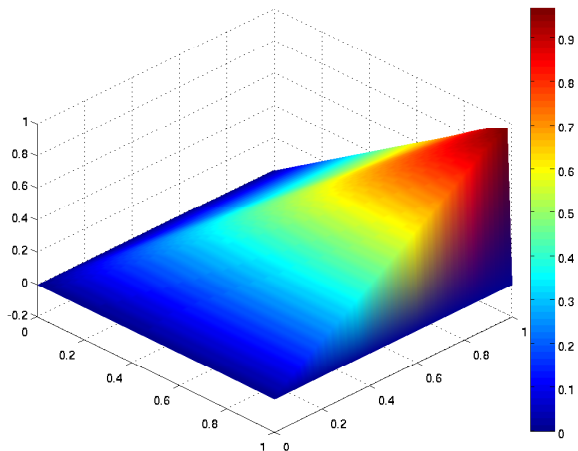
$$\begin{aligned} -\frac{1}{Pe} \Delta u + \mathbf{a} \cdot \nabla u &= 1 \text{ in } \Omega = (0, 1) \times (0, 1), \\ u &= 0, \text{ on } \partial\Omega, \end{aligned}$$

Focus on **convection dominated** scenario: $Pe = 10^6$ & $\mathbf{a} = \{1, 1\}^T$

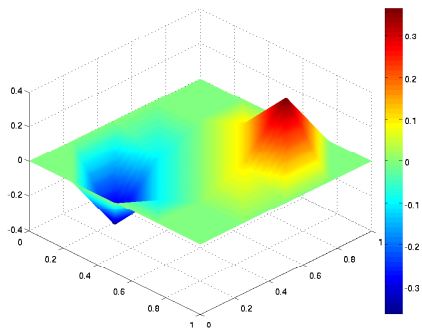
Thus, convection with $\pi/4$ angle, boundary layers along top/right edges
(width $\sim \frac{1}{Pe}$)

Convection Dominated Diffusion Problem

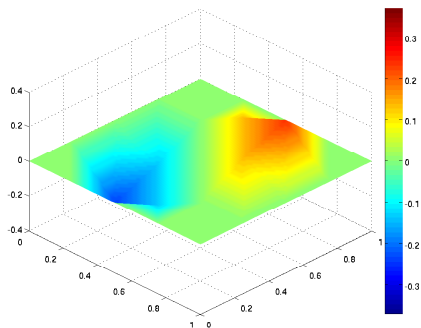
Reference Solution:



Uniform 4×4 Coarse Mesh

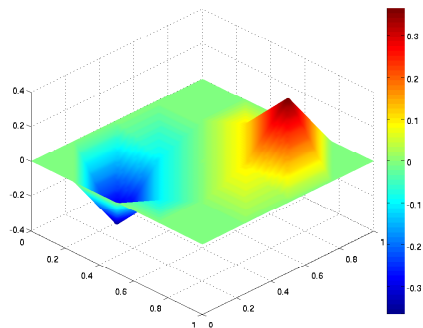


cDPG

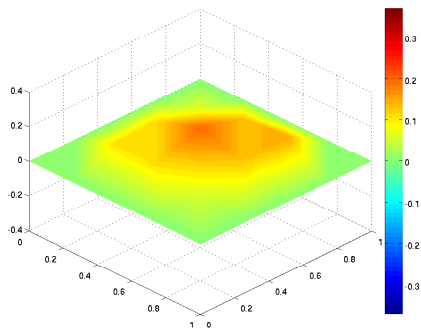


LSFEM

Uniform 4×4 Coarse Mesh



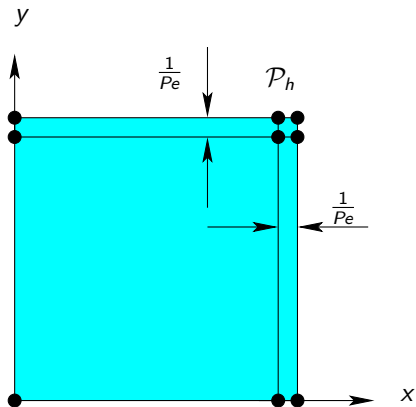
cDPG



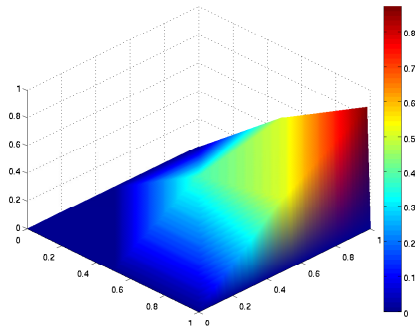
SUPG, VMS, G/LS

Graded Mesh

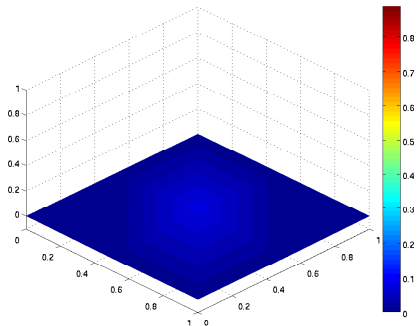
Numerical results for $Pe = 10^6$, starting with graded four-element mesh, $\Omega = (0, 1) \times (0, 1)$, $p = 2$, with subsequent uniform refinements



4×4 Graded Mesh

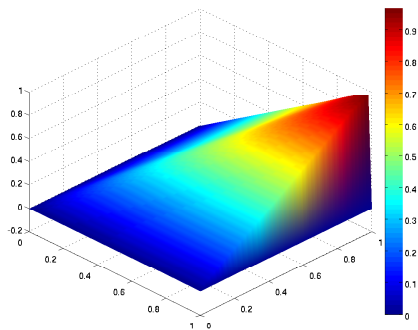


cDPG

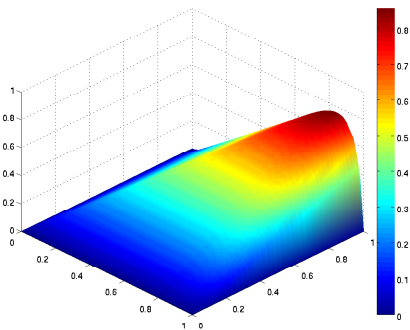


SUPG, VMS, and G/LS

Refined Graded Mesh



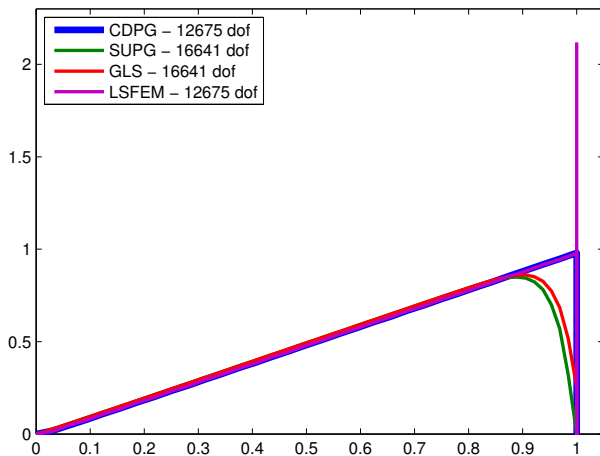
cDPG, LSFEM
(12k dofs)



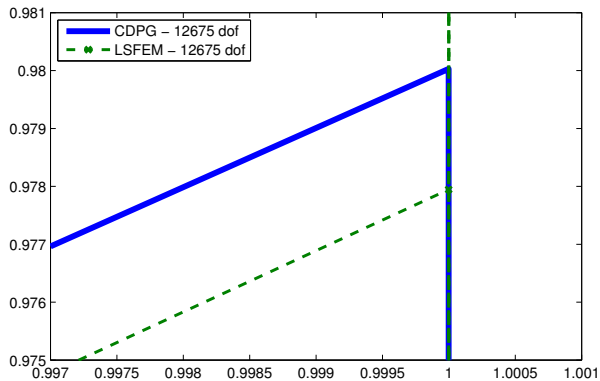
SUPG, G/LS
(16k dofs)

VMS produced highly oscillatory solutions

Distribution of u Along the Diagonal



(Zoomed in) Distribution of u at the Corner



Comparison Study - 'Shock' Problem

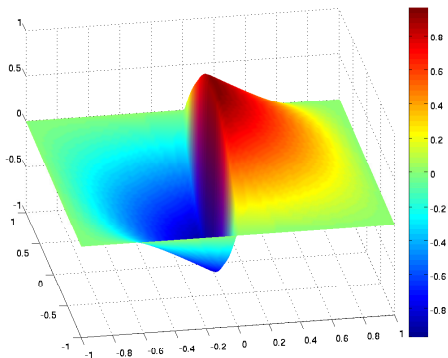
Convection-diffusion problem on the square $\Omega = (-1, 1) \times (-1, 1)$:

Find u such that:

$$\begin{aligned} -\frac{1}{Pe} \Delta u - \begin{Bmatrix} x \\ 0 \end{Bmatrix} \cdot \nabla u &= x(1-y^2) \frac{2x}{Pe}, \quad \text{in } \Omega, \\ u &= 0, \quad \text{on } \partial\Omega, \end{aligned}$$

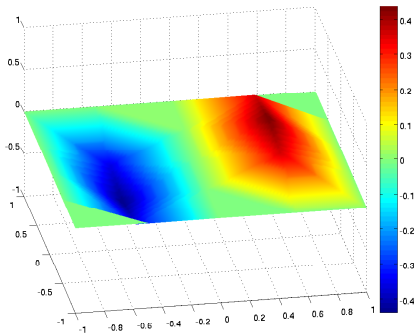
Comparison Study - 'Shock' Problem

Reference Solution:

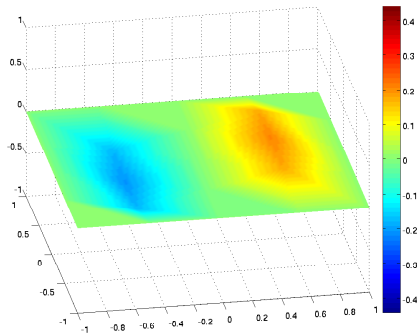


Numerical results for $Pe = 10^9$, uniform meshes, $p = 1$ for cDPG and LSFEM, and $p = 2$ for SUPG, VMS, G/LS.

16×16 Uniform Coarse Mesh

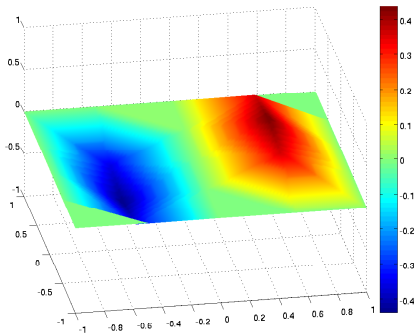


cDPG
($p = 1$)

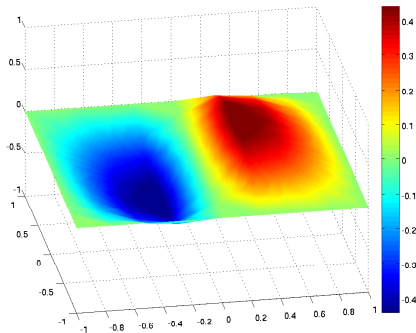


LSFEM
($p = 1$)

16×16 Uniform Coarse Mesh

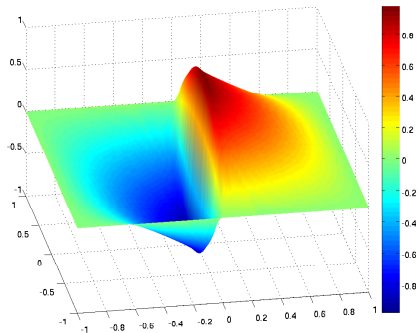


cDPG
($p = 1$)

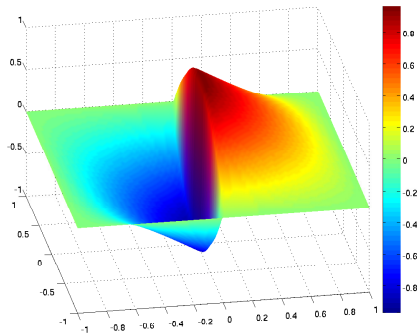


SUPG, VMS, G/LS
($p = 2$)

Uniform Refined Mesh

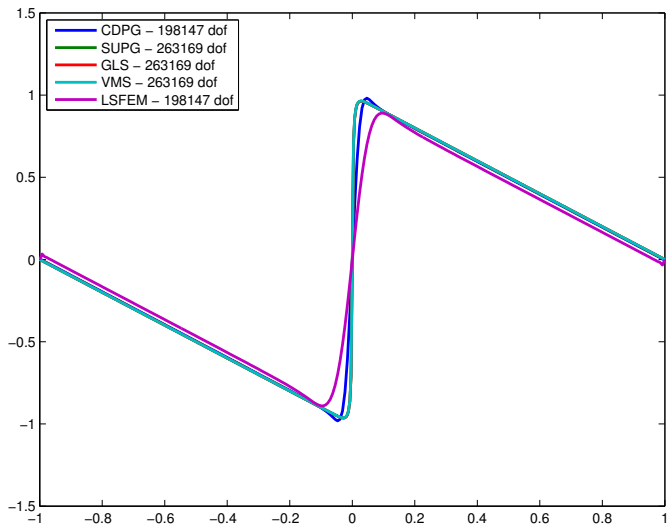


cDPG
($p = 1$, 198k dofs)



SUPG, G/LS
($p = 2$, 263k dofs)

Distribution of u Along the Centerline $x = 0$



Preliminary Results - Heterogeneous Domain

Convection-diffusion problem

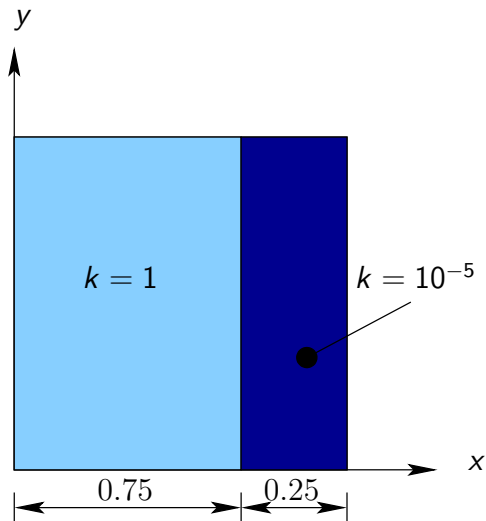
Find u such that:

$$-\nabla \cdot (k(\mathbf{x})\mathbf{l})\nabla u + \mathbf{a} \cdot \nabla u = f, \text{ in } \Omega = (0, 1) \times (0, 1),$$

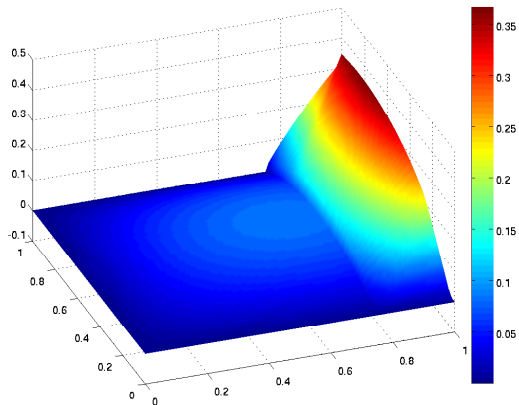
$$u = 0, \text{ on } \partial\Omega,$$

- $\mathbf{a} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$, $f = 1$, and $\kappa = k\mathbf{l}$
- Solution space built on uniform meshes with polynomial order $p = 2$
- Test functions solved using $p + 1 = 3$ on the same mesh

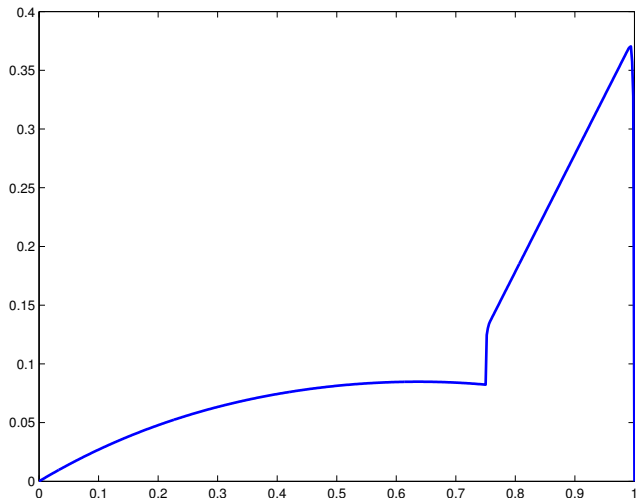
Preliminary Results - Heterogeneous Domain



Preliminary Results - Heterogeneous Domain



Distribution of u along $y = \frac{1}{2}$



Concluding Remarks

- Introduced hybrid continuous-discontinuous Petrov-Galerkin method.
 - ▶ Solution/trial functions are piecewise continuous
 - ▶ Weight/test functions are piecewise discontinuous.
- DPG approach: test functions computed automatically to establish numerically stable FE approximations
- Support of each discontinuous test function is identical to its corresponding continuous trial function
- Local test-function contribution computed locally (i.e. decoupled)
 - ▶ Sufficient accuracy by using corresponding p -level of trial function

Concluding Remarks

- Formulation allows for lower regularity of f (in dual of $H^1(\mathcal{P}_h)$) compared to LSFEM ($f \in L^2(\Omega)$)
- Numerical solutions do not show overly diffuse LSFEM solutions
- Numerical results compete with (SUPG, VMS, G/LS)
 - ▶ Show no oscillations at boundary layers.

Remark:

Weaker trial functions for σ than C^0 , i.e., use trial functions with continuous normal fluxes discontinuous tangential fluxes across element boundaries (\sim Raviart-Thomas/Brezzi-Douglas-Marini approach) deliver similar results to those reported herein