Ideas on Schwarz Smoothers in Efficient Multigrid Solvers

Daniel Arndt
Ryan Grove Julius Witte
Guido Kanschat

7th International Conference on HPSC

Hanoi University of Science and Technology, Vietnam

March 22, 2018

deal.II

IWR
Interdisciplinary Center for Scientific Computing

Table of Contents

Problem

A typical finite element problem reads in weak form

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

Discretization $\Rightarrow Ax = b$.

Solving large problems is restricted by

- Memory limitations
  - $\Rightarrow$ Matrix-Free methods
- Time limitations (Scalability)
  - $\Rightarrow$ Parallelization
  - $\Rightarrow$ Geometric Multigrid

Matrix-Free

Iterative solvers only require a matrix-vector product.
$\Rightarrow$ Idea:

- Storing the matrix not required
- "Assemble" the matrix-vector product directly, when solving.[1]

Matrix-free algorithm:

- v=0
- loop over cells
  1. Extract local vector values on cell: $u_K = P_K u$
  2. Apply operation locally on cell: $v_K = A_K u_K$ (without forming A_K)
  3. Sum results into the global solution vector: $v = v + P_K^T v_K$

[1] Martin Kronbichler and Katharina Kormann. "A generic interface for parallel cell-based finite element operator application". In: *Computers & Fluids* 63 (2012), pp. 135 –147. ISSN: 0045-7930

Matrix-Free

Matrix-Based

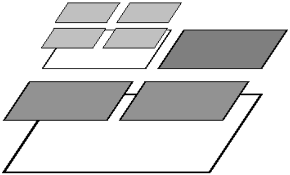
1. Build matrix $A^K_{i,j} = \int_K a(\phi_j, \phi_i)\, dx$
2. Solve using iterative method (CG) $\rightarrow$ compute $v_i = (Au)_i$.

Matrix-Free

Compute matrix-vector product directly

$$v_i = (Au)_i = \sum_K \int_K a(u, \phi_i)\, dx$$

Parallel Adaptive Multigrid



- Only method known that is $\mathcal{O}(N)$
- Geometric multigrid on adaptively refined meshes with local smoothing[2]
- `deal.II` implements parallel geometric multigrid for both continuous and discontinuous elements.

[1] Bärbel Janssen and Guido Kanschat. "Adaptive Multilevel Methods with Local Smoothing for H^1- and H^{curl}-Conforming High Order Finite Element Methods". In: *SIAM Journal on Scientific Computing* 33.4 (2011), pp. 2095–2114

Parallel Adaptive Multigrid - V-Cycle

1. Pre-smoothing:

$$u^{(k+1)} = u^{(k)} - R_l^{-1}(A_l u^{(k)} - f_l), \qquad 0 \leq k < m_{pre}$$

2. Coarse grid correction:

$$\begin{aligned} f_{l-1} &= \Pi_{l-1}^T (f_l - A_l u^{(m_{pre})}) \\ v^{(k+1)} &= MG_{l-1}(v^{(k)}, f_{l-1}), \qquad 0 \leq k < m_{coarse} \\ w^{(0)} &= u^{(m_{pre})} + v^{(m_{coarse})} \end{aligned}$$

3. Post-smoothing:

$$w^{(k+1)} = w^{(k)} - R_l^{-1}(A_l w^{(k)} - f_l), \qquad 0 \leq k < m_{post}$$

4. Assign: $MG(u^{(0)}, f_l) = w^{(m_{post})}$

Coarse grid solver $MG_0(u(0), f) = A_0^{-1} f_0$

Parallel Adaptive Multigrid - V-Cycle

1. Pre-smoothing:

$$u^{(k+1)} = u^{(k)} - R_l^{-1}(A_l u^{(k)} - f_l), \qquad 0 \leq k < m_{pre}$$

2. Coarse grid correction:

$$\begin{aligned} f_{l-1} &= \Pi_{l-1}^T (f_l - A_l u^{(m_{pre})}) \\ v^{(k+1)} &= MG_{l-1}(v^{(k)}, f_{l-1}), \qquad 0 \leq k < m_{coarse} \\ w^{(0)} &= u^{(m_{pre})} + v^{(m_{coarse})} \end{aligned}$$

3. Post-smoothing:

$$w^{(k+1)} = w^{(k)} - R_l^{-1}(A_l w^{(k)} - f_l), \qquad 0 \leq k < m_{post}$$

4. Assign: $MG(u^{(0)}, f_l) = w^{(m_{post})}$

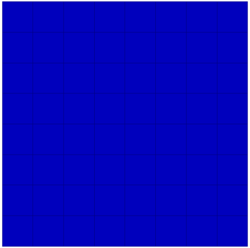
Coarse grid solver $MG_0(u(0), f) = A_0^{-1} f_0$

Additive Schwarz Smoother

Hermann Amandus Schwarz

- Take the local structure of the problem into account
- Use local problems for preconditioning

$$R_l = \sum_{K \in \mathcal{T}_l} P_K A_K^{-1}$$



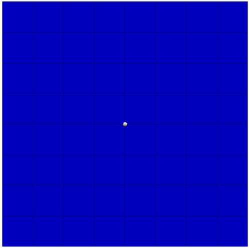
Additive Schwarz Smoother

Hermann Amandus Schwarz

- Take the local structure of the problem into account
- Use local problems for preconditioning

$$R_l = \sum_{K \in \mathcal{T}_l} P_K A_K^{-1}$$

- Point patches ⇒ Jacobi



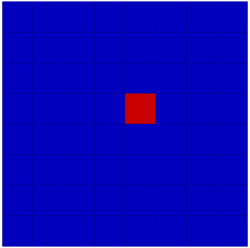
Additive Schwarz Smoother

Hermann Amandus Schwarz

- Take the local structure of the problem into account
- Use local problems for preconditioning

$$R_l = \sum_{K \in \mathcal{T}_l} P_K A_K^{-1}$$

- Cell patches ⇒ BlockJacobi



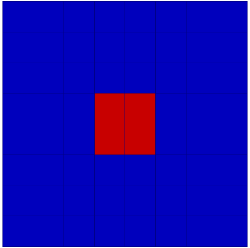
Additive Schwarz Smoother

Hermann Amandus Schwarz

- Take the local structure of the problem into account
- Use local problems for preconditioning

$$R_l = \sum_{K \in \mathcal{T}_l} P_K A_K^{-1}$$

- Vertex patches $\Rightarrow$ BlockJacobi



Abstract Schwarz Theory

PDE problem

Let V be Hilbert space, $f \in V'$. Find $u \in V$ such that

$$a(u, v) = f(v) \quad \forall v \in V,$$

where $a(.,.)$ is a symmetric, positive definite (s.p.d.) bilinear form.

Schwarz Operators

Assume V admits decomposition: $V = R_0^T V_0 + \sum_{j=1}^{J} R_j^T V_j$. Then, Schwarz operators $P_j \equiv R_j^T \tilde{P}_j$ are defined by projection-like operators $\tilde{P}_j : V \to V_j$:

$$\tilde{a}_j(\tilde{P}_j u, v_j) = a(u, R_j^T v_j) \quad \forall v_j \in V_j,$$

where $\tilde{a}_j(.,.)$ are local bilinear forms (s.p.d.).

Convergence Theory

I. Stable Decomposition

It exists C_0 such that every $u \in V$ has a decomposition $u = \sum_{j=0}^{J} R_j^T u_j$ that satisfies $\sum_{j=0}^{J} \tilde{a}_j(u_j, u_j) \leq C_0^2 a(u, u)$.

II. Strengthened Cauchy-Schwarz Inequality

It exist $0 \leq \epsilon_{ij} \leq 1, 1 \leq i, j \leq J$, such that

$$|a(R_i^T u_i, R_j^T u_j)| \leq \epsilon_{ij} a(R_i^T u_i, R_i^T u_i)^{1/2} a(R_j^T u_j, R_j^T u_j)^{1/2},$$

for $u_i \in V_i$ and $u_j \in V_j$. Define $\mathcal{E} = (\epsilon_{ij})$.

III. Local Stability

There exists $\omega > 0$, such that

$$a(R_j^T u_j, R_j^T u_j) \leq \omega \tilde{a}_j(u_j, u_j), \; u_j \in range(\tilde{P}_j) \subset V_j, \; 0 \leq j \leq J.$$

Convergence Theory

Theorem

Let Assumptions I. - III. be satisfied. Then

$$\kappa(P_{ad}) \leq C_0^2 \omega(\rho(\mathcal{E}) + 1).$$

- Schwarz operator: $P_j = R_j^T \tilde{A}_j^{-1} R_j A, \ j = 0, \dots, J$
- additive preconditioner:

$$P_{\text{ad}} = A_{\text{ad}}^{-1} A = \sum_{j=0}^{J} \left(R_j^T \tilde{A}_j^{-1} R_j A \right)$$

- multiplicative preconditioner:

$$P_{\text{mu}} = A_{\text{mu}}^{-1} A = I - \prod_{j=0}^{J} \left(I - R_j^T \tilde{A}_j^{-1} R_j A \right)$$

- mixtures possible

The Stokes problem

Stokes

Consider the Stokes problem

$$\nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (p, \nabla \cdot \boldsymbol{v}) + (\nabla \cdot \boldsymbol{u}, q) + \tau_{gd}(\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})$$
$$\nabla \cdot \boldsymbol{u} = 0$$

with the grad-div stabilization term $\tau_{gd}(\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v})$.

This leads to the following structure of the system matrix:

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ p \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix}$$

Hence, we have to solve a symmetric , but indefinite problem.

Raviart-Thomas Elements

Original result from Kanschat and Mao[3] using Raviart-Thomas elements.

Key assumption

$$\nabla \cdot V_h = Q_h \qquad V_{h,0}^{div} \subset \ldots \subset V_{h,L}^{div}$$

where

$$V_{h,l}^{div} := \{v_h \in V_l : (\nabla \cdot u_h, q_h) = 0 \quad \forall q_h \in Q_h\}$$

Can this assumption be weakened and the result be applied to other inf-sup stable elements?

[3] Guido Kanschat and Youli Mao. "Multigrid methods for Hdiv-conforming discontinuous Galerkin methods for the Stokes equations". In: *Journal of Numerical Mathematics* 23.1 (2015), pp. 51–66

Raviart-Thomas Elements

Theorem[4]

The multilevel iteration $I - \mathcal{B}_L \mathcal{A}_L$ for the Stokes problem

- with the variable V-cycle operator $\mathcal{B}_L$
- employing the smoother $\mathcal{R}_l$ with suitably small scaling factor η

is a contraction with contraction number independent of the level L.

[4] Guido Kanschat and Youli Mao. "Multigrid methods for Hdiv-conforming discontinuous Galerkin methods for the Stokes equations". In: *Journal of Numerical Mathematics* 23.1 (2015), pp. 51–66

Stokes, Perturbed Problem

Idea: Eliminate the pressure by considering a perturbed formulation

$$\nu(\nabla \boldsymbol{u}_l, \nabla \boldsymbol{v}_h) + \tau_{gd}(\nabla \cdot \boldsymbol{u}_h - \epsilon p_h, \nabla \cdot \boldsymbol{v}_h - \epsilon q_h)$$
$$-(p_l, \nabla \cdot \boldsymbol{v}_h) + (\nabla \cdot \boldsymbol{u}_l, q_h) - \epsilon(p_l, q_h) = (\boldsymbol{f}, \boldsymbol{v}_h)$$

Defining the operator $\mathcal{A}_l : \boldsymbol{V}_l \times Q_l \to (\boldsymbol{V}_l \times Q_l)^*$ by

$$\begin{aligned}\mathcal{A}_l((\boldsymbol{u}_l, p_l), (\boldsymbol{v}_h, q_h)) :=& \nu(\nabla \boldsymbol{u}_l, \nabla \boldsymbol{v}_h) \\ &+ \tau_{gd}(\nabla \cdot \boldsymbol{u}_l - \epsilon p_l, \nabla \cdot \boldsymbol{v}_h - \epsilon q_h) \\ &+ (p_l, \nabla \cdot \boldsymbol{v}_h) + (\nabla \cdot \boldsymbol{u}_l, q_h) - \epsilon(p_l, q_h).\end{aligned}$$

this problem can be written as $\mathcal{A}_l((\boldsymbol{u}_l, p_l), (\boldsymbol{v}_h, q_h)) = (\boldsymbol{f}, \boldsymbol{v}_h)$ for all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_l \times Q_l$.

Stokes, Perturbed Problem

For $\epsilon > 0$, the Stokes problem can be rewritten as

$$\begin{aligned}\mathcal{A}_l(\boldsymbol{u}_l, \boldsymbol{v}_h) := \nu(\nabla \boldsymbol{u}_l, \nabla \boldsymbol{v}_h) \\ + \tau_{gd}(\pi_{Q_h}^{\perp}(\nabla \cdot \boldsymbol{u}_l), \pi_{Q_h}^{\perp}(\nabla \cdot \boldsymbol{v}_h)) \\ + \frac{1}{\epsilon}(\pi_{Q_h}(\nabla \cdot \boldsymbol{u}_l), \pi_{Q_h}(\nabla \cdot \boldsymbol{v}_h)).\end{aligned}$$

$$\mathcal{A}_l(\boldsymbol{u}_l, \boldsymbol{v}_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_l$$

Lemma

Let $(\boldsymbol{u}_l, p_l)$ be the solution to the perturbed problem in two variables and $\widetilde{\boldsymbol{u}}_l$ the solution to the perturbed problem in one variable. Then it holds

$$\widetilde{\boldsymbol{u}}_l = \boldsymbol{u}_l \qquad \epsilon p_l = \pi_{Q_h}(\nabla \cdot \boldsymbol{u}_l) = \pi_{Q_h}(\nabla \cdot \widetilde{\boldsymbol{u}}_l)$$

Convergence of the Perturbation

Let $(\boldsymbol{u}, p)$ be the solution to the continuous Stokes problem and $(\boldsymbol{u}_h, p_h)$ the solution to the discretized (perturbed) problem.

Lemma

It holds

$$\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_0^2 + \tau_{gd}\|\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}_h)\|_0^2 + \|p-p_h\|_0^2$$
$$\lesssim \epsilon + h^{2k_p+2} + h^{2k_u}.$$

Convergence Result

Convergence

If $\mathcal{R}_l$ satisfies for all $\boldsymbol{w} \in \boldsymbol{V}_l$

$$\mathcal{A}_l((\mathcal{I}_l - \mathcal{R}_l\mathcal{A}_l)\boldsymbol{w}, \boldsymbol{w}) \geq 0 \tag{1}$$

$$(\mathcal{R}_l^{-1}[\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}, [\mathcal{I}_l - \mathcal{P}_{l-1})\boldsymbol{w}) \leq \beta_l\mathcal{A}_l([\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}, [\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}), \tag{2}$$

then it holds

$$0 \leq \mathcal{A}_l([\mathcal{I}_l - \mathcal{B}_l\mathcal{A}_l)\boldsymbol{w}, \boldsymbol{w}) \leq \delta\mathcal{A}_l(\boldsymbol{w}, \boldsymbol{w}), \quad \forall \boldsymbol{w} \in \boldsymbol{V}_l$$

where $\delta < 1$ is independent of the number of levels.

Lemma

Let $\eta \leq 2^{-dim}$*, then*

$$\mathcal{A}_l((\mathcal{I}_l - \mathcal{R}_l\mathcal{A}_l)\boldsymbol{w}, \boldsymbol{w}) \geq 0, \quad \forall \boldsymbol{w} \in \boldsymbol{V}_l.$$

Stable Decomposition

Lemma

It holds

$$\eta \sum_v (\mathcal{R}_l^{-1} \boldsymbol{u}, \boldsymbol{u}) = \inf_{\substack{\boldsymbol{u}_v \in \boldsymbol{V}_{l,v} \\ \sum_v \mathcal{I}_{l,v} \boldsymbol{u}_v = \boldsymbol{u}}} \sum_v \mathcal{A}_l(\mathcal{I}_{l,v} \boldsymbol{u}_v, \mathcal{I}_{l,v} \boldsymbol{u}_v)$$

Essentially, we only need to find a decomposition $(\boldsymbol{u}_v)_v$ of $[\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}$, i.e.

$$\boldsymbol{u} := [\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w} = \sum_v \mathcal{I}_{l,v} \boldsymbol{u}_v.$$

such that

$$\sum_v (\mathcal{A}_l \mathcal{I}_{l,v} \boldsymbol{u}_v, \mathcal{I}_{l,v} \boldsymbol{u}_v) \leq \beta_l \mathcal{A}_l([\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}, [\mathcal{I}_l - \mathcal{P}_{l-1}]\boldsymbol{w}).$$

Helmholtz-like decomposition

Denote the bilinear form a_l corresponding to the weak Laplace operator by

$$a_l(\boldsymbol{u}, \boldsymbol{v}) := \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v})$$

For $\boldsymbol{u}_l \in \boldsymbol{V}_l$ define $\boldsymbol{u}_l^0 \in \boldsymbol{V}_{h,l}^{div}$ as projection of $\boldsymbol{u}_l$ onto $\boldsymbol{V}_{h,l}^{div}$ with respect to a_l, i.e.

$$a_l(\boldsymbol{u}_l^0, \boldsymbol{v}_l) = a_l(\boldsymbol{u}_l, \boldsymbol{v}_l) \quad \forall \boldsymbol{v}_l \in \boldsymbol{V}_{h,l}^{div}.$$

Then define $\boldsymbol{u}_l^\perp$ by $\boldsymbol{u}_l^\perp := \boldsymbol{u}_l - \boldsymbol{u}_l^0$.

Lemma

$$\frac{\alpha}{d^2}\|\nabla \cdot \boldsymbol{u}_l^\perp\|_0^2 \leq a_l(\boldsymbol{u}_l^\perp, \boldsymbol{u}_l^\perp) \leq \frac{\nu}{\gamma^2}\|\pi_{Q_h}(\nabla \cdot \boldsymbol{u}_l^\perp)\|_0^2$$

Stable Decomposition

Theorem

For any $\mathbf{v}_l \in \mathbf{V}_l$ there exists a decomposition $\mathbf{v}_{l,j}$ such that

$$\sum_{j=0}^{J} \mathcal{A}_l(\mathbf{v}_{l,j}, \mathbf{v}_{l,j}) \lesssim \mathcal{A}_l(\mathbf{v}_l, \mathbf{v}_l)$$

provided $\tau_{gd} \lesssim \min\{\nu, \epsilon^{-1}\}$.

Assumption

$$\sum_{v} a_l(\mathbf{u}_v^\perp, \mathbf{u}_v^\perp) \leq C a_l(\mathbf{u}_l^\perp, \mathbf{u}_l^\perp)$$

This holds for discontinuous, divergence-free elements.
What about $\mathbb{Q}_k/\mathbb{Q}_{k-1}$, $\mathbb{Q}_k/\mathbb{P}^-_{k-1}$ or other inf-sup stable element pairs?

Numerical Results - Test Problem

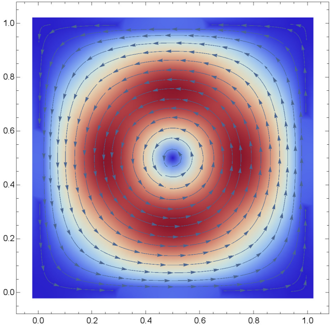
We consider the test problem

$$-\nu\Delta \boldsymbol{u} + \nabla p = -\nu\Delta \boldsymbol{u}_{ref} + \nabla p_{ref}$$
$$\nabla \cdot \boldsymbol{u} = 0$$

with the reference solution

$$\boldsymbol{u}(x,y) = \begin{pmatrix} \sin^2(\pi x)\sin(2\pi y) \\ -\sin^2(\pi y)\sin(2\pi x) \end{pmatrix}$$
$$p(x,y) = \sin(\pi x)\cos(\pi y).$$

in 2D.



Observe for $\nu = 10^{-6}$

- errors
- iteration counts (error reduction by 10^{-6}).

Numerical Results - Error Estimates

Theorem

For a given $\mathbf{f} \in H^{-1}(\Omega)^d$, *let* $(\boldsymbol{u}, p)$ *be the continuous problem, and let* $(\boldsymbol{u}_h, p_h)$ *be the solution to the discrete solution. Then, the error in the energy norm for the velocity is bounded by*

$$\nu\|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_0^2 \leq \inf_{\boldsymbol{w}_h \in [H_0^1(\Omega)]^d} \left(4\nu\|\nabla(\boldsymbol{u} - \boldsymbol{w}_h)\|_0^2 + 2\tau_{gd}\|\nabla \cdot \boldsymbol{w}_h\|_0^2\right) + \frac{2}{\tau_{gd}} \inf_{q_h \in Q_h} \|p - q_h\|_0^2.$$

[4] Eleanor W Jenkins et al. "On the parameter choice in grad-div stabilization for the Stokes equations". In: *Advances in Computational Mathematics* 40.2 (2014), pp. 491–516

Numerical Results - Optimal Stabilization Parameter

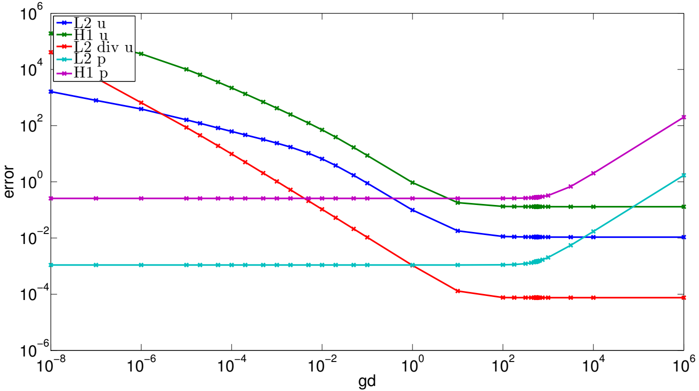


Figure: $\mathbb{Q}_2/\mathbb{Q}_1$ elements, optimal $\tau_{gd} = 544.917$

Numerical Results - Optimal Stabilization Parameter

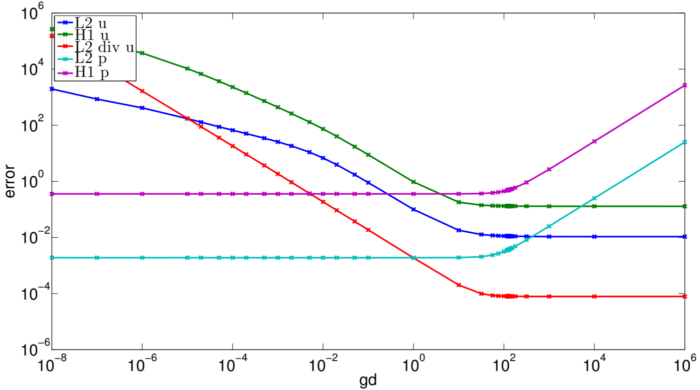


Figure: $\mathbb{Q}_2/\mathbb{P}_1^-$ elements, optimal $\tau_{gd} = 128.93$

Numerical Results

Element	Refinement 1	2	3	4	5	6
$\mathbb{Q}_2/\mathbb{P}_1^-$	$1 \cdot 10^0$	$3 \cdot 10^1$	$2 \cdot 10^1$	$5 \cdot 10^1$	$1 \cdot 10^2$	$9 \cdot 10^1$
$\mathbb{Q}_2/\mathbb{Q}_1$	$9 \cdot 10^0$	$9 \cdot 10^1$	$9 \cdot 10^1$	$2 \cdot 10^2$	$5 \cdot 10^2$	$4 \cdot 10^2$
$\mathbb{Q}_3/\mathbb{P}_2^-$	$1 \cdot 10^5$	$5 \cdot 10^3$	$4 \cdot 10^0$	$1 \cdot 10^1$	$5 \cdot 10^2$	$9 \cdot 10^{-1}$
$\mathbb{Q}_3/\mathbb{Q}_2$	$1 \cdot 10^5$	$8 \cdot 10^2$	$6 \cdot 10^1$	$4 \cdot 10^1$	$4 \cdot 10^2$	$5 \cdot 10^{-1}$

Table: Optimal stabilization parameter

#Levels	0	1	2	3	4
RT1, 2D	3	9	10	11	13
RT2, 2D	3	9	10	11	11

Table: Iteration counts for Raviart-Thomas elements

Numerical Results - Iteration Counts - $\mathbb{Q}_2/\mathbb{P}_1^-$ - $\eta = \frac{1}{4}$ - 2D

	τ_{gd}				
#Levels	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}
0	2	2	2	2	2
1	24	23	27	22	17
2	70	65	57	43	25
3	236	159	93	48	27
4	459	247	105	49	28

	τ_{gd}				
#Levels	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
0	2	2	2	2	2
1	14	12	13	13	13
2	19	18	18	18	18
3	19	19	19	19	19
4	19	18	19	20	20

Numerical Results - Iteration Counts - $\mathbb{Q}_2/\mathbb{Q}_1$ - $\eta = \frac{1}{8}$ - 2D

	τ_{gd}				
#Levels	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}
0	2	2	2	2	2
1	35	31	36	28	22
2	137	98	85	59	31
3	454	294	159	71	37
4	-	610	190	76	38

	τ_{gd}				
#Levels	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
0	2	2	2	2	2
1	17	17	17	17	17
2	23	24	24	25	25
3	28	31	34	35	35
4	28	33	38	39	39

Numerical Results - Iteration Counts - $\mathbb{Q}_3/\mathbb{P}_2^-$ - $\eta = \frac{1}{4}$ - 2D

	τ_{gd}				
#Levels	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}
0	2	2	2	2	2
1	12	12	13	14	15
2	19	19	19	20	17
3	30	30	29	24	16
4	39	38	35	24	16

	τ_{gd}				
#Levels	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
0	2	2	3	3	2
1	15	16	16	16	16
2	15	17	18	18	18
3	15	17	18	18	18
4	14	16	18	18	18

Numerical Results - Iteration Counts - $\mathbb{Q}_3/\mathbb{Q}_2$ - $\eta = \frac{1}{8}$ - 2D

	τ_{gd}				
#Levels	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}
0	2	2	2	2	2
1	19	19	20	20	21
2	31	31	31	27	25
3	36	35	33	32	25
4	50	50	50	36	25

	τ_{gd}				
#Levels	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
0	2	2	2	2	2
1	21	22	22	22	22
2	25	28	31	31	31
3	26	30	33	34	34
4	27	32	37	37	38

Test case - Jumping Coefficients

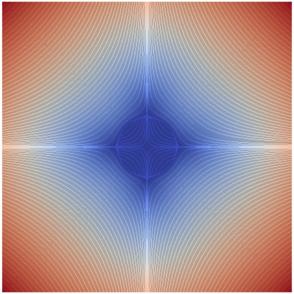


Figure: Reference velocity

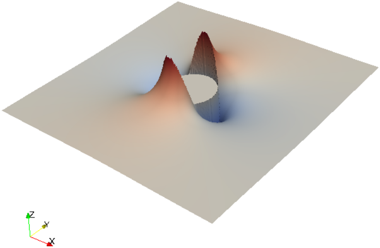


Figure: Reference pressure

$$\nu(r) = \begin{cases} \eta & \text{for } r^2 < r_i^2 \\ 1 & \text{for } r^2 \geq r_i^2 \end{cases}$$

Test case - Jumping Coefficients

$$A = \frac{\eta - 1}{\eta + 1} \qquad z = x + \mathrm{i}y$$

$$\phi = \begin{cases} 0 & \text{for } r^2 < r_i^2 \\ -2Ar_i^2/z & \text{for } r^2 \geq r_i^2 \end{cases} \qquad \psi = \begin{cases} -4z\frac{\eta}{1+\eta} & \text{for } r^2 < r_i^2 \\ \frac{-2(z+Ar_i^4}{z^3} & \text{for } r^2 \geq r_i^2 \end{cases}$$

$$\nu(r) = \begin{cases} \eta & \text{for } r^2 < r_i^2 \\ 1 & \text{for } r^2 \geq r_i^2 \end{cases}$$

$$\boldsymbol{u} = \begin{pmatrix} \Re(\phi - z\overline{\phi'(z)} - \overline{\psi(z)}) \\ \Im(\phi - z\overline{\phi'(z)} - \overline{\psi(z)}) \end{pmatrix} /2\nu(r)$$

$$p = -2\Re(\phi'(z));$$

Test case - Jumping Coefficients

	η					
#level	10		100		1000	
	$\mathbb{Q}_2/\mathbb{P}_1^-$	$\mathbb{Q}_3/\mathbb{P}_2^-$	$\mathbb{Q}_2/\mathbb{P}_1^-$	$\mathbb{Q}_3/\mathbb{P}_2^-$	$\mathbb{Q}_2/\mathbb{P}_1^-$	$\mathbb{Q}_3/\mathbb{P}_2^-$
1	1	1	1	1	1	1
2	1	3	1	3	1	3
3	10	12	10	13	12	16
4	14	14	15	14	21	20
5	15	14	18	15	29	20
6	16	15	19	17	33	25
7	16	15	22	18	40	33
8	16	14	22	19	45	33

Table: #iterations

Dictionary of Inverses

- Theory: We don't have to solve exact local problems!
- Idea: Store only a few local inverses that lead to similar behavior as an exact Schwarz smoother

Check Inverse Similarity

Store only local inverses that differ significantly,

$$\frac{\|A_i^{-1}A_j - I\|_{Frob}}{N_j} < \text{TOL} \qquad \Rightarrow \qquad A_j^{-1} \approx A_i^{-1},$$

where N_j is the amount of DoFs on j-th subdomain.

- use linear elements to check similarity
- generic

Test Settings

- SIPG discretization for diffusion term
- uniform refinement: $c(\Omega_0)4^\ell$ cells on level ℓ
  - unit square: $c = 1$
  - unit ball: $c = 5$
- standard V-cycle with 1 pre- & post-smoothing step
- non-overlapping, additive Schwarz smoother
- residual reduction of 10^{-6}

Laplace - Iterations 2D

$$-\Delta u = 1 \text{ in } (0,1)^2, \quad g_D \equiv 0.$$

	$k = 10$			$k = 7$		
#levels	1-Inv	BJ	Dict(0.01)	1-Inv	BJ	Dict(0.01)
5	19	19	19	18	18	18
6	19	19	19	18	18	18
7	20	20	20	19	19	19
8	21	21	21	19	19	19
9	22	21	21	20	20	20
10	22	22	22	20	20	20

CG–iterations: $\mathbb{Q}_k^-$ elements; Dictionary (4x corner, 4x side, 1x interior); 1-Inverse.

Laplace - Iterations 2D - Unit Ball

$$-\Delta u = 1 \text{ in } B(0,1), \quad g_D \equiv 0.$$

	$k=4$			$k=7$				
#levels	BJ	Dict(0.05)	Dict(0.01)	BJ	Dict(0.05)	(%)	Dict(0.01)	(%)
5	22	25	25	32	36		33	
6	24	28	26	34	43		35	
7	25	32	27	36	49		37	
8	26	34	28	37	54	(0.2)	38	(4)
9	28	37	29	39	61	(0.05)	40	(1)

(*relative inverses on finest level*)

CG–iterations: $\mathbb{Q}_k^-$ elements; 1-Inverse does not converge.

Diffusion - Iterations 2D

$$-\nabla \cdot (D\nabla u) = 1 \text{ in } (0,1)^2,\ g_D \equiv 0,\ D(x) = \frac{1}{0.01+100|x|^2}$$

	BJ	Dict(0.1)		Dict(0.01)		Dict(0.001)	
#levels	$k=10$	$k=10$	(%)	$k=10$	(%)	$k=10$	(%)
5	22	27		24		24	
6	24	27		24		24	
7	23	27		23		23	
8	23	28		23		23	
9	23	28	(0.05)	23	(0.7)	23	(6.1)
10	24	29	(0.01)	24	(0.2)	24	(2.3)

(*relative inverses on finest level*)

CG–iterations: $\mathbb{Q}_k^-$ elements. (Scaled) 1-Inverse does not converge.

Diffusion - Iterations 2D

$$-\nabla \cdot (D \nabla u) = 1 \text{ in } (0,1)^2,\ g_D \equiv 0,\ D(x) = \frac{1}{\alpha + \beta |x|^2}$$

	$\alpha = 0.001, \beta = 100$					$\alpha = 0.001, \beta = 10000$				
#levels	BJ	Dict(0.1)	(%)	Dict(0.05)	(%)	BJ	Dict(0.1)	(%)	Dict(0.05)	(%)
5	31	40		36		31	39		29	
6	27	33		29		30	38		31	
7	26	32		28		36	44		40	
8	26	30		27		39	47		36	
9	24	30	(0.08)	28	(0.33)	32	39	(0.08)	33	(0.35)
10	26	31	(0.02)	29	(0.09)	30	38	(0.02)	33	(0.1)

(*relative inverses on finest level*)

CG–iterations: $\mathbb{Q}_{10}^{-}$ elements. (Scaled) 1-Inverse does not converge.

Separable Operators

Let $\Omega \subset \mathbb{R}^2$ be a domain, $H = H_1 \otimes H_2$ be a tensor product of Hilbert spaces and $\mathcal{L} : H \to H$ a separable, differential operator,

$$\mathcal{L}u = \left(\mathcal{L}_1 \otimes \mathrm{id} + \mathrm{id} \otimes \mathcal{L}_2\right) u_1 \otimes u_2.$$

Let a, a_1, a_2 denote bilinear forms associated to $\mathcal{L}, \mathcal{L}_1, \mathcal{L}_2$,

$$a(u, v) = a_1(u_1, v_1)\,(u_2, v_2)_{L^2} + (u_1, v_1)_{L^2}\, a_2(u_2, v_2).$$

A FEM discretization leads to a sum of Kronecker products of one dimensional stiffness matrices

$$A = L^{(1)} \otimes M^{(2)} + M^{(1)} \otimes L^{(2)}.$$

Fast Diagonalization Method (FDM)[5]

For convenience, we assume $\mathcal{L}$ to be symmetric.
Compute one dimensional generalized eigenvalue problems

$$Q^T L^{(1)} Q = \Lambda^{(1)}, \qquad Q^T M^{(1)} Q = I,$$
$$P^T L^{(2)} P = \Lambda^{(2)}, \qquad P^T M^{(2)} P = I,$$

A tensor-product basis transformation of A into the eigenvector spaces results in a diagonal matrix,

$$(Q \otimes P)^T A (Q \otimes P) = \Lambda^{(1)} \otimes I + I \otimes \Lambda^{(2)}.$$

Therefore it is sufficient to solve small generalized eigenvalue problems to obtain the inverse of A,

$$A^{-1} = Q \otimes P \left[\Lambda^{(1)} \otimes I + I \otimes \Lambda^{(2)}\right]^{-1} Q^T \otimes P^T.$$

[1] R.E. Lynch, J.R. Rice, and D.H. Thomas. "Tensor product analysis of partial difference equations". In: *Bulletin of the American Mathematical Society* 70.3 (1964), pp. 378–384

FDM - Reduced Complexities

Let d denote the spatial dimension and n the number of degrees of freedom in one direction per element.

inverting	we requires $\mathcal{O}(dn^3)$ operations to invert A
storage	d matrices of eigenvectors and d vectors of eigenvalues
vmult	we exploit the Kronecker decompositions of the basis transformation matrices for sum-factorized matrix-vector multiplications

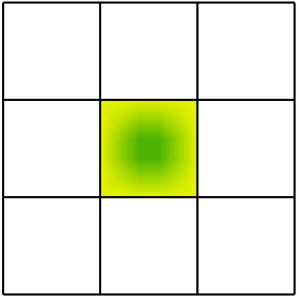
	standard	*tensor-product*
inverting	$\mathcal{O}(n^{3d})$	$\mathcal{O}(dn^3)$
storage	$\mathcal{O}(n^{2d})$	$\mathcal{O}(dn^2)$
vmult	$\mathcal{O}(n^{2d})$	$\mathcal{O}(dn^{d+1})$

Table: Complexity comparison to standard inverting

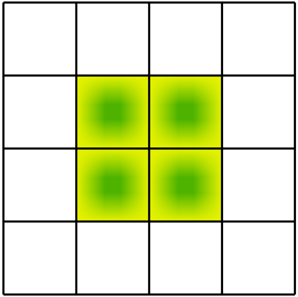
Schwarz Smoothers

Local Homogeneous Poisson Problem

Solve $-\Delta u = 1$ in Ω_S, $u = 0$ on $\partial\Omega_S$, using a local SIPG discretization.



(a) *cell* patch Ω_S



(b) regular *vertex* patch Ω_S

Additive Cell-Patch

Levels	Convergence steps			
2D	$k = 3$	$k = 4$	$k = 7$	$k = 10$
5	15	15	19	21
6	15	14	18	20
7	15	14	18	20
8	14	14	17	19
9	14	13	17	19
10	14	13	17	19
3D	$k = 3$	$k = 4$	$k = 7$	$k = 10$
2	16	16	22	25
3	18	17	22	25
4	18	17	23	25
5	18	17	22	25
6	17	17	22	24

Table: $\mathbb{Q}_k^-$ elements. Relaxation $\omega = 0.7$.

Additive Vertex-Patch

Levels	Convergence steps			
2D	$k = 3$	$k = 4$	$k = 7$	$k = 10$
5	18	18	21	23
6	17	18	21	23
7	17	18	20	23
8	17	18	20	22
9	17	18	19	21
10	17	18	19	20
3D	$k = 3$	$k = 4$	$k = 7$	$k = 10$
2	21	21	21	21
3	26	26	28	29
4	28	30	32	36
5	28	32	37	39
6	28	32	39	42

Table: $\mathbb{Q}_k^-$ elements. Relaxation $\omega = 0.25$ (0.15 in 3D).

Subdomain Coloring

Graph Coloring Algorithm [6]

First partition the mesh into zones, then colorize with respect to a given conflict type (DSATUR algorithm) and finally gather colors among zones.[7]

cell-patch	red-black coloring, 2 colors in 2D & 3D
vertex-patch (FC)	shared face defines a conflict, minimal number of colors is 8 in 2D and 16 in 3D
vertex-patch (CC)	shared cell defines a conflict, minimal number of colors is 4 in 2D and 8 in 3D

Table: optimal numbers of colors on a uniformly refined unit square/cube

[2] B. Turcksin, M. Kronbichler, and W. Bangerth. “*WorkStream* – a design pattern for multicore-enabled finite element computations”. In: *ACM Transactions on Mathematical Software* 43.1 (2016), pp. 2/1–2/29

[3] Algorithm is tuned to schedule tasks equally, not to minimize the number of colors.

Multiplicative Cell-Patch

Levels	Convergence steps				Colors
2D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
5	7.5	7.8	10.4	11.4	2
6	7.5	7.7	10.3	11.3	2
7	7.5	7.6	10.3	11.2	2
8	7.4	7.5	10.2	10.8	2
9	7.3	7.4	9.9	10.6	2
10	7.2	7.3	9.8	10.5	2
3D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
2	8.1	8.4	11.8	13.5	2
3	8.3	8.5	11.9	13.5	2
4	8.4	8.8	11.9	13.3	2
5	8.4	8.9	11.8	13.3	2
6	8.4	8.8	11.7	12.8	2

Table: $\mathbb{Q}_k^-$ elements. Relaxation $\omega = 1.0$ (1.1 in 3D).

Multiplicative Vertex-Patch (FC)

Levels	Convergence steps				Colors
2D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
5	2.7	3.0	2.0	2.0	16
6	2.7	3.0	2.0	2.0	17
7	2.8	3.0	2.0	2.0	17
8	2.6	2.8	2.1	2.0	17
9	2.5	2.6	2.1	2.0	17
10	2.5	2.5	2.0	2.0	17
3D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
2	2.0	2.0	2.0	1.5	1
3	2.1	2.2	2.0	2.0	19
4	2.2	2.3	2.0	2.0	36
5	2.3	2.4	2.0	2.0	50
6	2.2	2.4	2.0	2.0	52

Table: $\mathbb{Q}_k^-$ elements. Relaxation $\omega = 1.0$.

Multiplicative Vertex-Patch (CC)

Levels	Convergence steps				Colors
2D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
5	2.0	2.2	2.0	2.0	6
6	2.0	2.2	2.0	2.0	6
7	2.0	2.2	2.0	2.0	6
8	2.0	2.1	2.0	2.0	6
9	2.0	2.1	2.0	2.0	6
10	2.0	2.1	2.0	2.0	6
3D	$k = 3$	$k = 4$	$k = 7$	$k = 10$	
2	2.0	2.0	2.0	2.0	1
3	2.0	2.1	2.0	2.0	14
4	2.0	2.1	2.0	2.0	16
5	2.0	2.1	2.0	2.0	16
6	2.0	2.1	2.0	2.0	16

Table: $\mathbb{Q}_k^-$ elements. Relaxation $\omega = 1.0$.

Convergence Comparison 2D

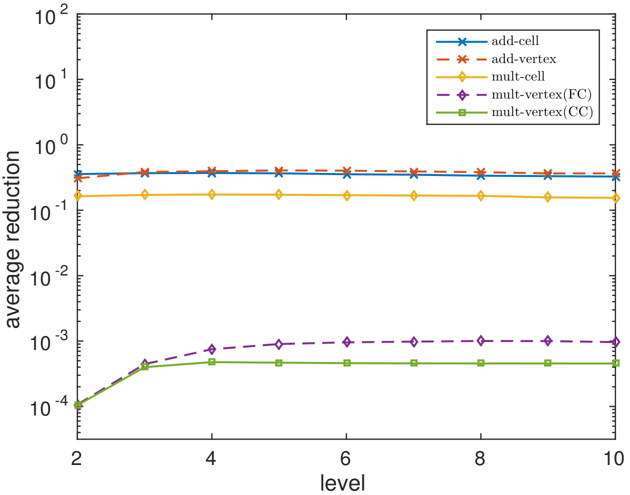


Figure: Average reduction for GMRES and $\mathbb{Q}_7^-$ elements

Convergence Comparison 3D

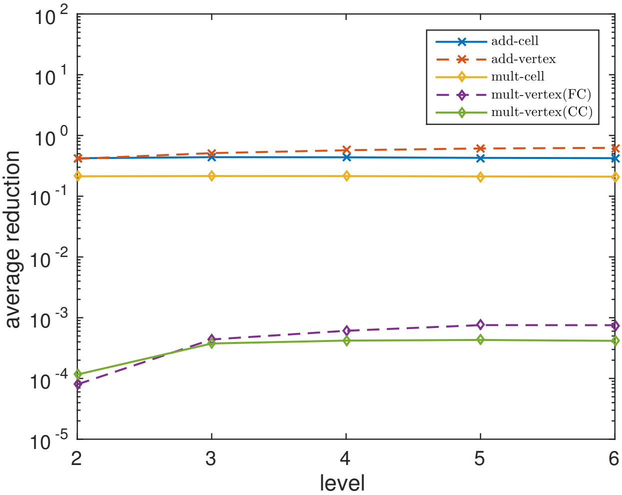


Figure: Average reduction for GMRES and $\mathbb{Q}_7^-$ elements

Summary

Results

- Local Schwarz smoothers applicable for inf-sup stable conforming elements
- Comparable results to Raviart-Thomas elements
- $\mathbb{Q}_{k+1}/\mathbb{P}_k^-$ elements perform much better than $\mathbb{Q}_{k+1}/\mathbb{Q}_k$ elements
- Analysis requires $\tau_{gd} \lesssim \nu$; sharpness confirmed by numerical results (at least for continuous pressure)
- Positive effect of stabilization especially for $\mathbb{Q}_{k+1}/\mathbb{Q}_k$ elements
- Dictionary approach/FDM shows promising results for different problems

Outlook/Challenges:

- Consider also convection dominated problems (Oseen, Navier-Stokes)
- Lift the restriction $\tau_{gd} \lesssim \nu$?

Thank you for your attention!

References I

- Bärbel Janssen and Guido Kanschat. "Adaptive Multilevel Methods with Local Smoothing for H^1- and H^{curl}-Conforming High Order Finite Element Methods". In: *SIAM Journal on Scientific Computing* 33.4 (2011), pp. 2095–2114.
- Eleanor W Jenkins et al. "On the parameter choice in grad-div stabilization for the Stokes equations". In: *Advances in Computational Mathematics* 40.2 (2014), pp. 491–516.
- Guido Kanschat and Youli Mao. "Multigrid methods for Hdiv-conforming discontinuous Galerkin methods for the Stokes equations". In: *Journal of Numerical Mathematics* 23.1 (2015), pp. 51–66.

References II

- Martin Kronbichler and Katharina Kormann. "A generic interface for parallel cell-based finite element operator application". In: *Computers & Fluids* 63 (2012), pp. 135 –147. ISSN: 0045-7930.
- R.E. Lynch, J.R. Rice, and D.H. Thomas. "Tensor product analysis of partial difference equations". In: *Bulletin of the American Mathematical Society* 70.3 (1964), pp. 378–384.
- B. Turcksin, M. Kronbichler, and W. Bangerth. "*WorkStream* – a design pattern for multicore-enabled finite element computations". In: *ACM Transactions on Mathematical Software* 43.1 (2016), pp. 2/1–2/29.