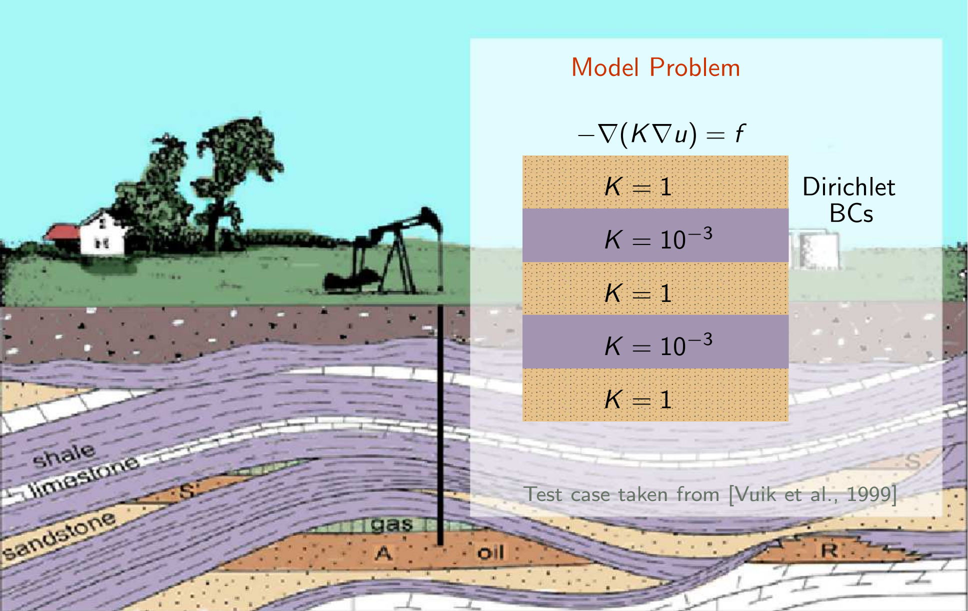
An efficient and robust Krylov method

for Discontinuous Galerkin methods

Kees Vuik   Paulien van Slingerland

Oil Reservoir Simulation

Layered structures yield challenging problems



Model Problem

$$-\nabla(K\nabla u) = f$$

$K = 1$
$K = 10^{-3}$
$K = 1$
$K = 10^{-3}$
$K = 1$

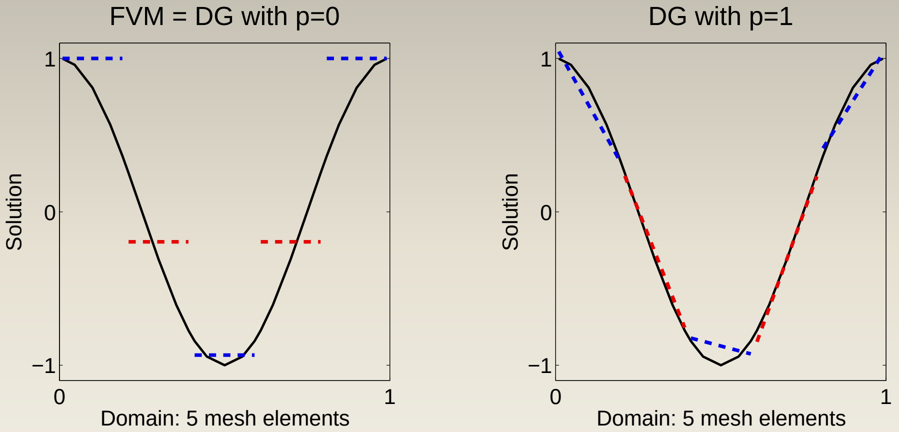
Dirichlet BCs

Test case taken from [Vuik et al., 1999]

Methods

DG Methods

DG methods are like FVMs, but then based on piecewise polynomials

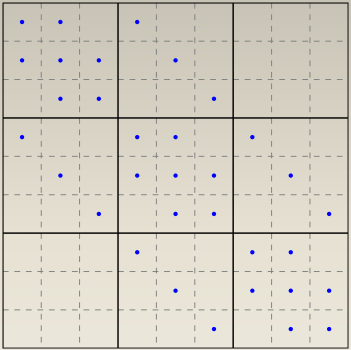


DG for elliptic problems: [Arnold et al., 2002], [Rivière, 2008]

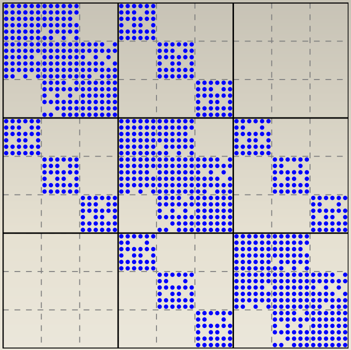
Problem

DG matrices are ill-conditioned and relatively large

2D Laplace problem with 3×3 mesh elements:



DG matrix with $p = 0$

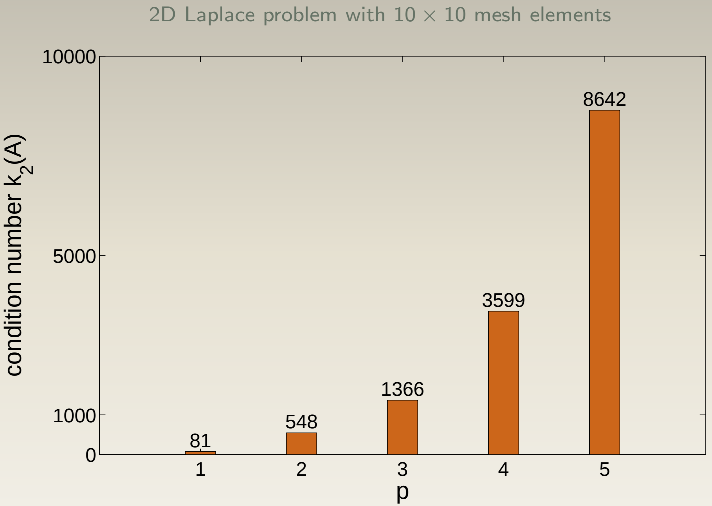


DG matrix with $p = 2$

Condition number: $\kappa(A) = O(h^{-2})$

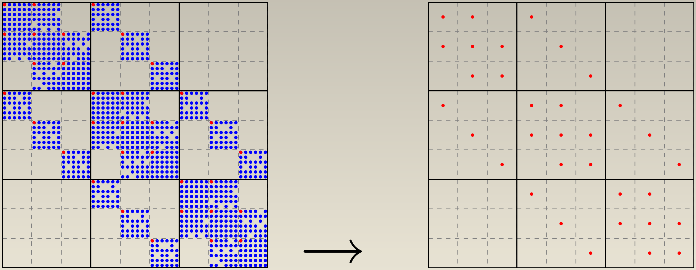
Condition number

The condition number increases with the polynomial degree p



Coarse Corrections

The main idea is to speed up CG using coarse corrections based on $p = 0$



DG matrix A with $p > 0$ DG matrix RAR^T with $p = 0$

$$A^{-1} \approx Q := \underbrace{R^T}_{\text{prolongation}} \; (RAR^T)^{-1} \; \underbrace{R}_{\text{restriction}}$$

Original idea of spectral multigrid: [Rønquist and Patera, 1987]

Two-Level Preconditioner

The coarse corrections are combined with two smoothing steps

Computing $z = P_{\mathrm{prec}} r$:

$z^{(1)} = \omega M^{-1} r$	smoothing
$z^{(2)} = z^{(1)} + Q(r - Az^{(1)})$	coarse correction
$z = z^{(2)} + \omega M^{-T}(r - Az^{(2)})$	smoothing

Requirement: $M + M^T - \omega A$ is SPD

This preconditioner yields scalable CG convergence

as was shown for $p = 1$ in [Dobrev et al., 2006] using [Falgout et al., 2005]

ADEF2 Deflation Variant

We can switch to deflation by simply skipping a smoothing step

Computing $z = P_{\mathrm{ADEF2}} r$:

$z^{(1)} = \omega M^{-1} r$	smoothing
$z^{(2)} = z^{(1)} + Q(r - Az^{(1)})$	coarse correction
~~$z = z^{(2)} + \omega M^{-T}(r - Az^{(2)})$~~	~~smoothing~~

Requirement: ~~$M + M^T - \omega A$ is SPD~~ M is SPD

This operator is not symmetric ...

(cf. next slide)

BNN deflation variant

BNN is symmetric, but more expensive due to two coarse solves

Computing $z = P_{\mathrm{BNN}} r$:

$z^{(1)} = Qr$	coarse correction
$z^{(2)} = z^{(1)} + \omega M^{-1}(r - Az^{(1)})$	smoothing
$z = z^{(2)} + Q(r - Az^{(2)})$	coarse correction

Requirement: M is SPD

ADEF2 and BNN yield the same CG iterates

assuming we preprocess the start vector: $x_0 \mapsto Qb + (I - AQ)^T x_0$ [Tang et al., 2009]

Theoretical Results

Theory

The preconditioner yields scalable CG convergence

Theorem:

Suppose that A is the SPD result of a coercive SIPG discretization for a diffusion problem $-\nabla(K\nabla u) = f$ on a regular mesh with $p \geq 1$. Then:

$$\kappa_2(P_{\mathrm{prec}}^{-1} A) \lesssim 1,$$

i.e. $\kappa_2(P_{\mathrm{prec}}^{-1} A) \leq C$ for some constant C independent of h.

Shown for $p = 1$ in [Dobrev et al., 2006] using [Falgout et al., 2005]

Assumptions:

- The diffusion and SIPG penalty parameter are bounded
- M is nonsingular and $M + M^T - \omega A$ is SPD $\quad\Rightarrow\quad$ P_{prec} is SPD
- $h^{2-d} v^T M^T (M + M^T - \omega A)^{-1} M v \lesssim v^T v$ for all vectors v

Theory

A similar result is true for BNN deflation

Theorem:

Suppose that A is the SPD result of a coercive SIPG discretization for a diffusion problem $-\nabla(K\nabla u) = f$ on a regular mesh with $p \geq 1$. Then:

$$\kappa_2(P_{\mathrm{BNN}}^{-1}A) \lesssim 1,$$

i.e. $\kappa_2(P_{\mathrm{BNN}}^{-1}A) \leq C$ for some constant C independent of h.

To be shown in this presentation

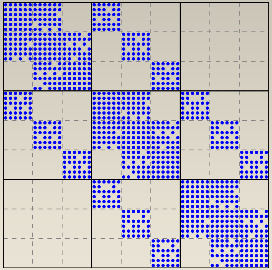
Assumptions: The diffusion and SIPG penalty parameter are bounded

M is SPD and $2M - \omega A$ is SPD $\quad \Rightarrow \quad P_{\mathrm{BNN}}$ is SPD

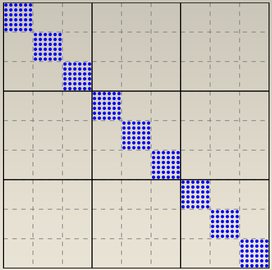
$h^{2-d} v^T M v \lesssim v^T v$ for all vectors v

Special case

The required smoothing conditions are satisfied for block Jacobi smoothing



SIPG matrix A



Block Jacobi smoother M

Assumption: the damping parameter $\omega \leq 1$ (strictly for the preconditioning variant)

Proof Overview

We show that $\kappa_2(P^{-1}A) \lesssim 1$ for BNN deflation without relaxation ($\omega = 1$)

Step 1: Writing $E = I - P^{-1}A$, we have (with $\lambda_{\text{max}}(E) < 1$):

$$\kappa_2(P^{-1}A) \leq \frac{1 - \lambda_{\text{min}}(E)}{1 - \lambda_{\text{max}}(E)}$$

Step 2: Next, there exists a constant $K \geq 1$ such that:

$$-1 \leq \lambda_{\text{min}}(E) \leq \lambda_{\text{max}}(E) \leq 1 - \frac{1}{K}$$

Step 3: The proof is completed by showing:

$$K \lesssim 1$$

Proof: Step 1

We show that $\kappa_2(P^{-1}A) \leq \frac{1-\lambda_{\min}(E)}{1-\lambda_{\max}(E)}$

- $\lambda_{\max}(E) = \max_{y\neq 0} \frac{y^TAEy}{y^TAy}, \qquad \lambda_{\min}(E) = \min_{y\neq 0} \frac{y^TAEy}{y^TAy}$

Proof: $E := I - P^{-1}A$ has the same spectrum as $A^{\frac{1}{2}}EA^{-\frac{1}{2}} = I - A^{\frac{1}{2}}P^{-1}A^{\frac{1}{2}}$ (symmetric)

$\lambda_{\max}(E) = \max_{x\neq 0} \frac{x^TA^{\frac{1}{2}}EA^{-\frac{1}{2}}x}{x^Tx} \overset{y=A^{-\frac{1}{2}}x}{=} \max_{y\neq 0} \frac{y^TAEy}{y^TAy}$

Similarly: $\lambda_{\min}(E) = \ldots = \min_{y\neq 0} \frac{y^TAEy}{y^TAy}$

Proof: Step 1

We show that $\kappa_2(P^{-1}A) \leq \frac{1-\lambda_{\min}(E)}{1-\lambda_{\max}(E)}$

- $\lambda_{\max}(E) = \max_{y\neq 0} \frac{y^T AEy}{y^T Ay}, \qquad \lambda_{\min}(E) = \min_{y\neq 0} \frac{y^T AEy}{y^T Ay}$
- $\frac{1}{1-\lambda_{\min}(E)} A \leq P \leq \frac{1}{1-\lambda_{\max}(E)} A$

Proof: For all y: $\lambda_{\min}(E)\, y^T Ay \leq y^T AEy \leq \lambda_{\max}(E)\, y^T Ay$

Shorter notation: $\lambda_{\min}(E)\, A \leq AE \leq \lambda_{\max}(E)\, A$

Rewriting yields: $\frac{1}{1-\lambda_{\min}(E)} A \leq P \leq \frac{1}{1-\lambda_{\max}(E)} A$

Proof: Step 1

We show that $\kappa_2(P^{-1}A) \le \frac{1-\lambda_{\min}(E)}{1-\lambda_{\max}(E)}$

- $\lambda_{\max}(E) = \max_{y\neq 0} \frac{y^T A E y}{y^T A y}, \qquad \lambda_{\min}(E) = \min_{y\neq 0} \frac{y^T A E y}{y^T A y}$
- $\frac{1}{1-\lambda_{\min}(E)} A \le P \le \frac{1}{1-\lambda_{\max}(E)} A$
- We conclude:

$$\kappa_2(P^{-1}A) \le \frac{1-\lambda_{\min}(E)}{1-\lambda_{max}(E)}$$

Proof: Known: $c_1 A \le P \le c_2 A \quad \Rightarrow \quad \kappa_2(P^{-1}A) \le \frac{c_2}{c_1}$ (cf. e.g. [Vassilevski, 2008])

The proof is completed using $c_1 = \frac{1}{1-\lambda_{\min}(E)}$ and $c_2 = \frac{1}{1-\lambda_{\max}(E)}$

Proof: Step 2

We show that $-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$

- Define for any SPD matrix D:

  $\Theta_D = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$ $\qquad$ $\bar{\pi}_A = A^{\frac{1}{2}} Q A^{\frac{1}{2}}$

  $K_D = \sup_{v \neq 0} \frac{\|(I - \pi_D)v\|_D^2}{\|v\|_A^2}$ $\qquad$ $\pi_D = R^T (RDR^T)^{-1} RD$

Recall: Restriction operator R: defined such that $A_0 = RAR^T$ is the SIPG matrix for $p = 0$

Coarse correction operator: $Q = R^T A_0^{-1} R$

Proof: Step 2

We show that $-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$

- Define for any SPD matrix D:

  $\Theta_D = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$ $\qquad \bar{\pi}_A = A^{\frac{1}{2}} Q A^{\frac{1}{2}}$

  $K_D = \sup_{v \neq 0} \frac{\|(I - \pi_D)v\|_D^2}{\|v\|_A^2}$ $\qquad \pi_D = R^T (R D R^T)^{-1} R D$

- E has the same spectrum as Θ_M

Proof: Known: $E = I - P^{-1}A = (I - QA)(I - M^{-1}A)(I - QA)$ [Tang et al., 2010]

At the same time: $A^{-\frac{1}{2}} \Theta_M A^{\frac{1}{2}} = (I - QA)(I - M^{-1}A)(I - QA)$

Proof: Step 2

We show that $-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$

- Define for any SPD matrix D:

  $\Theta_D = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$ $\qquad \bar{\pi}_A = A^{\frac{1}{2}} Q A^{\frac{1}{2}}$

  $K_D = \sup_{v \neq 0} \frac{\|(I - \pi_D)v\|_D^2}{\|v\|_A^2}$ $\qquad \pi_D = R^T (RDR^T)^{-1} RD$
- E has the same spectrum as Θ_M
- $0 < D_1 \leq D_2 \quad \Rightarrow \quad \Theta_{D_1} \leq \Theta_{D_2}$

Proof: $D_1 \leq D_2 \Rightarrow I - A^{\frac{1}{2}} D_1^{-1} A^{\frac{1}{2}} \leq I - A^{\frac{1}{2}} D_2^{-1} A^{\frac{1}{2}}$

$QAQ = Q \Rightarrow \bar{\pi}_A^2 = (A^{\frac{1}{2}} Q A^{\frac{1}{2}})(A^{\frac{1}{2}} Q A^{\frac{1}{2}}) = A^{\frac{1}{2}} Q A^{\frac{1}{2}} = \bar{\pi}_A$ is a symmetric projection

$\Theta_{D_1} = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D_1^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A) \leq (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D_2^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A) = \Theta_{D_2}$

Proof: Step 2

We show that $-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$

- Define for any SPD matrix D:

  $\Theta_D = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$ $\qquad$ $\bar{\pi}_A = A^{\frac{1}{2}} Q A^{\frac{1}{2}}$

  $K_D = \sup_{v \neq 0} \frac{\|(I - \pi_D)v\|_D^2}{\|v\|_A^2}$ $\qquad$ $\pi_D = R^T (RDR^T)^{-1} RD$

- E has the same spectrum as Θ_M
- $0 < D_1 \leq D_2 \quad \Rightarrow \quad \Theta_{D_1} \leq \Theta_{D_2}$
- $D - A$ SPSD $\quad \Rightarrow \quad \lambda_{\max}(\Theta_D) \leq 1 - \frac{1}{K_D}$ $\quad$ with $K_D \geq 1$

Proof: The proof is very similar to [Falgout et al., 2005]

Proof: Step 2

We show that $-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$

- Define for any SPD matrix D:

  $\Theta_D = (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} D^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$ $\qquad$ $\bar{\pi}_A = A^{\frac{1}{2}} Q A^{\frac{1}{2}}$

  $K_D = \sup_{v \neq 0} \frac{\|(I - \pi_D)v\|_D^2}{\|v\|_A^2}$ $\qquad$ $\pi_D = R^T (R D R^T)^{-1} R D$

- E has the same spectrum as Θ_M
- $0 < D_1 \leq D_2 \quad \Rightarrow \quad \Theta_{D_1} \leq \Theta_{D_2}$
- $D - A$ SPSD $\quad \Rightarrow \quad \lambda_{\max}(\Theta_D) \leq 1 - \frac{1}{K_D}$ $\quad$ with $K_D \geq 1$
- Using that $2M - A$ is SPD completes the proof (with $K_{2M} \geq 1$):

$$-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K_{2M}}$$

Proof: $2M - A > 0 \Rightarrow M \geq \frac{1}{2} A \Rightarrow \Theta_M \geq \Theta_{\frac{1}{2}A} = -(I - \bar{\pi}_A)^2 \Rightarrow \lambda_{\min}(\Theta_M) \geq -1$

$M \leq 2M \Rightarrow \Theta_M \leq \Theta_{2M} \Rightarrow \lambda_{\max}(\Theta_M) \leq 1 - \frac{1}{K_{2M}}$ (with $K_{2M} \geq 1$)

Using that E has the same spectrum as Θ_M completes the proof

Proof: Step 3

We show that $K \lesssim 1$

- For all v: $\|(I - \pi_{2M})v\|_{2M}^2 \leq h^{d-2}\|(I - \pi_I)v\|_2^2$

Proof: Recall the assumption: $h^{2-d} w^T M w \lesssim w^T w$ for all vectors w

π_D is the projection onto the coarse space Range(R^T) in the D-norm

$\|(I - \pi_{2M})v\|_{2M}^2 \leq \|(I - \pi_I)v\|_{2M}^2 = h^{d-2}\|(I - \pi_I)v\|_{h^{2-d}2M}^2 \leq h^{d-2}\|(I - \pi_I)v\|_2^2$

Proof: Step 3

We show that $K \lesssim 1$

- For all v: $\|(I - \pi_{2M})v\|_{2M}^2 \leq h^{d-2}\|(I - \pi_I)v\|_2^2$
- For all $w \in$ Range $(I - \pi_I)$: $w^T w \lesssim h^{2-d} w^T A w$

Info: This can be shown using properties of the SIPG method (we omit the proof)
The main idea is to construct a block diagonal matrix D (independent of h)
such that $w^T w \lesssim w^T D w \lesssim h^{2-d} w^T A w$ (coercivity is also used)

Proof: Step 3

We show that $K \lesssim 1$

- For all v: $\|(I - \pi_{2M})v\|_{2M}^2 \leq h^{d-2}\|(I - \pi_I)v\|_2^2$
- For all $w \in \text{Range}\ (I - \pi_I)$: $w^T w \lesssim h^{2-d} w^T A w$
- We conclude that

$$K := K_{2M} := \sup_{v \neq 0} \frac{\|(I - \pi_{2M})v\|_{2M}^2}{\|v\|_A^2} \lesssim 1$$

Info: For all v: $\|(I - \pi_{2M})v\|_{2M}^2 \leq h^{d-2}\|(I - \pi_I)v\|_2^2 \lesssim \|(I - \pi_I)v\|_A^2 \leq \|v\|_A^2$
(The last inequality follows because π_I is a projection)
Substitution into the definition of K completes the proof

Proof Overview

We show that $\kappa_2(P^{-1}A) \lesssim 1$ for BNN deflation without relaxation ($\omega = 1$)

Step 1: Writing $E = I - P^{-1}A$, we have (with $\lambda_{\max}(E) < 1$):

$$\kappa_2(P^{-1}A) \leq \frac{1 - \lambda_{\min}(E)}{1 - \lambda_{\max}(E)}$$

Step 2: Next, there exists a constant $K \geq 1$ such that:

$$-1 \leq \lambda_{\min}(E) \leq \lambda_{\max}(E) \leq 1 - \frac{1}{K}$$

Step 3: The proof is completed by showing:

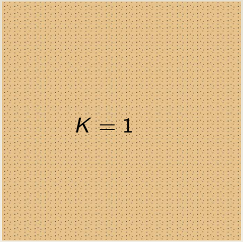
$$K \lesssim 1$$

Numerical Experiments

Poisson Problem

Both two-level methods yield fast & scalable CG convergence

degree	p=2				p=3			
mesh	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$
size A	2400	9600	38400	153600	4000	16000	64000	256000
Jacobi	301	581	1049	1644	325	576	1114	1903
block Jacobi (BJ)	205	356	676	1190	206	357	696	1183
two-level prec., 2x BJ	36	38	39	40	49	52	53	54
two-level defl., 1x BJ	32	33	33	34	36	37	37	38



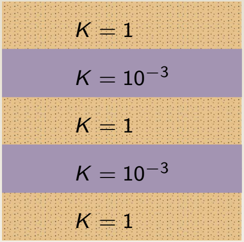
CG stopping criterion: $\frac{||b - Ax_k||_2}{||b||_2} \leq 10^{-6}$

Diagonal-scaling is applied beforehand

Layered Problem

Not so fast & scalable anymore ...

degree	p=2				p=3			
mesh	N=20^2	N=40^2	N=80^2	N=160^2	N=20^2	N=40^2	N=80^2	N=160^2
size A	2400	9600	38400	153600	4000	16000	64000	256000
Jacobi	1671	4311	9069	15923	2675	5064	9104	15657
block Jacobi (BJ)	933	2253	4996	9651	1357	2960	5660	9783
two-level prec., 2x BJ	415	1215	2534	3571	1089	2352	4709	8781
two-level defl., 1x BJ	200	414	531	599	453	591	667	698



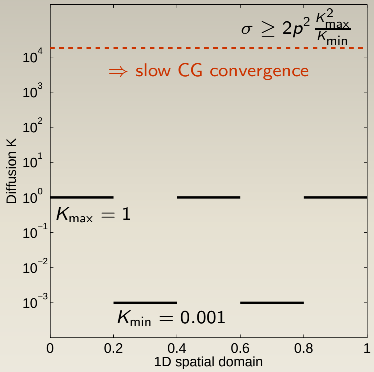
CG stopping criterion: $\frac{||b-Ax_k||_2}{||b||_2} \leq 10^{-6}$

Diagonal-scaling is applied beforehand

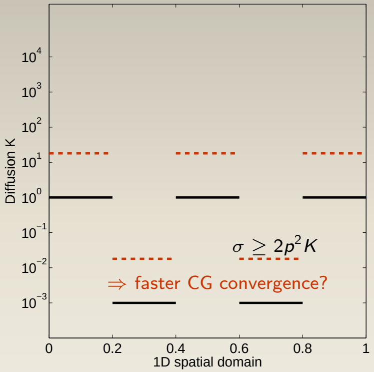
SIPG Penalty Parameter σ

Sufficiently large for stabilty, but as small as possible for proper conditioning

Constant penalty



Diffusion-dependent penalty



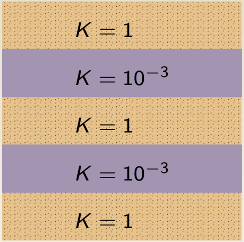
[Epshteyn and Rivière, 2007]

[Dryja, 2003]

Layered Problem Revisited

Both two-level methods now yield fast & scalable convergence

degree	p=2				p=3			
mesh	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$
size A	2400	9600	38400	153600	4000	16000	64000	256000
Jacobi	976	1264	1570	2315	1303	1490	1919	3109
block Jacobi (BJ)	243	424	788	1285	244	425	697	1485
two-level prec., 2x BJ	46	43	43	44	55	56	56	57
two-level defl., 1x BJ	43	45	45	46	47	48	48	48



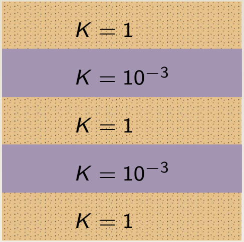
CG stopping criterion: $\frac{||b-Ax_k||_2}{||b||_2} \leq 10^{-6}$

Diagonal-scaling is applied beforehand

Weighted averages (SWIP)

The convergence can worsen dramatically

degree	p=2				p=3			
mesh	N=20^2	N=40^2	N=80^2	N=160^2	N=20^2	N=40^2	N=80^2	N=160^2
size A	2400	9600	38400	153600	4000	16000	64000	256000
Jacobi	731	2339	6908	18694	1177	3434	11151	36700
block Jacobi (BJ)	833	4050	15525	70616	344	867	2169	6529
two-level prec., 2x BJ	538	1332	3697	11948	133	225	446	1050
two-level defl., 1x BJ	219	564	1841	6068	88	142	285	570

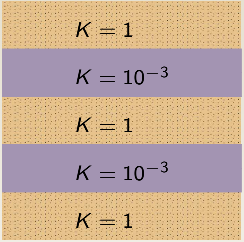


Based on harmonic averages: $\sigma = 20\frac{2K_iK_j}{K_i+K_j}$
(rather than $\sigma = 20\max\{K_i, K_j\}$)

Weighted averages (SWIP)

The convergence can worsen dramatically

degree	p=2				p=3			
mesh	N=20^2	N=40^2	N=80^2	N=160^2	N=20^2	N=40^2	N=80^2	N=160^2
size A	2400	9600	38400	153600	4000	16000	64000	256000
Jacobi	731	2339	6908	18694	1177	3434	11151	36700
block Jacobi (BJ)	833	4050	15525	70616	344	867	2169	6529
two-level prec., 2x BJ	538	1332	3697	11948	133	225	446	1050
two-level defl., 1x BJ	219	564	1841	6068	88	142	285	570



Scheme not coercive...
Matrix not positive-definite...

Based on harmonic averages: $\sigma = 20\frac{2K_iK_j}{K_i+K_j}$
(rather than $\sigma = 20\max\{K_i, K_j\}$)

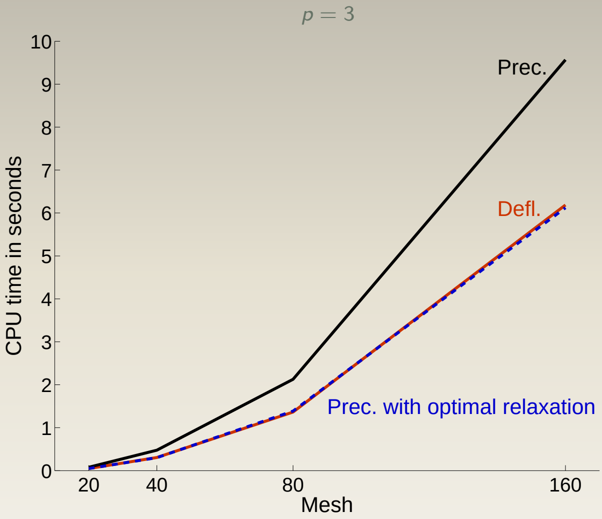
Relaxation

Using ωM^{-1} rather than M^{-1} benefits the preconditioning variant only

degree	p=2				p=3			
mesh	N=20^2	N=40^2	N=80^2	N=160^2	N=20^2	N=40^2	N=80^2	N=160^2
two-level prec. ($\omega = 1$)	**46**	**43**	**43**	**44**	**55**	**56**	**56**	**57**
($\omega = 0.9$)	34	34	34	37	38	40	40	42
($\omega = 0.8$)	32	33	34	34	36	36	37	39
($\omega = 0.7$)	**31**	**33**	**33**	**33**	**34**	**35**	**36**	**36**
($\omega = 0.6$)	32	32	33	34	35	35	36	36
($\omega = 0.5$)	34	34	34	35	37	36	37	38
two-level defl. ($\omega = 1$)	**43**	**45**	**45**	**46**	**47**	**48**	**48**	**48**
($\omega = 0.9$)	43	45	45	46	47	48	48	48
($\omega = 0.8$)	43	45	45	46	47	48	48	48
($\omega = 0.7$)	43	45	45	46	47	48	48	48
($\omega = 0.6$)	43	45	45	46	47	48	48	48
($\omega = 0.5$)	43	45	45	46	47	48	48	48

CPU Time

Deflation is still fast due to 30% lower costs per iteration



Coarse Systems

Coarse systems can be solved by applying CG again with a high tolerance

degree	p=2				p=3			
mesh	N=20^2	N=40^2	N=80^2	N=160^2	N=20^2	N=40^2	N=80^2	N=160^2
direct	43	45	45	46	47	48	48	48
$\mathrm{TOL} = 10^{-4}$	43	45	45	46	47	48	48	48
$\mathrm{TOL} = 10^{-3}$	43	45	45	46	47	48	48	48
$\mathrm{TOL} = 10^{-2}$	43	45	45	46	47	48	48	48
$\mathrm{TOL} = 10^{-1}$	55	60	81	51	48	48	54	79

Outer loop: two-level deflation, $\frac{||b - Ax_k||_2}{||b||_2} \leq 10^{-6}$

Inner loop: AMG * preconditioner, $\frac{||b - Ax_k||_2}{||b||_2} \leq \mathrm{TOL}$

*HSL, a collection of Fortran codes for large-scale scientific computation. See http://www.hsl.rl.ac.uk/

Inner iterations

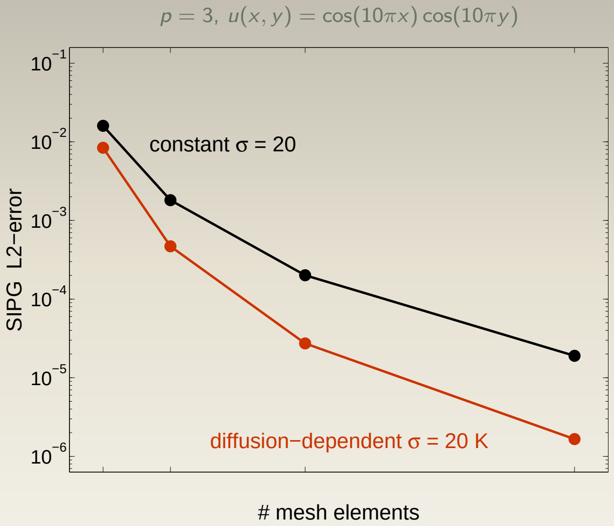
The HSL-CG solver yields fast convergence

Average # inner iterations ($\mathrm{TOL} = 10^{-2}$):

degree	p=2				p=3			
mesh	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$	$N=20^2$	$N=40^2$	$N=80^2$	$N=160^2$
$\mathrm{TOL} = 10^{-2}$	2.0	2.5	2.4	3.2	2.0	2.1	2.6	3.1

SIPG Convergence

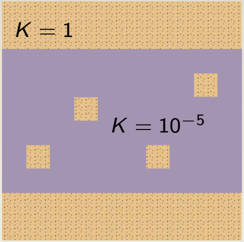
A diffusion-dependent penalty parameter yields better accuracy



Test Case II

Sand inclusions

degree	p=2				p=3			
mesh	N=40^2	N=80^2	N=160^2	N=320^2	N=40^2	N=80^2	N=160^2	N=320^2
two-level prec. ($\omega = 1$)	44	48	46	46	53	56	58	59
two-level prec. ($\omega = 0.7$)	30	30	30	31	32	34	34	34
two-level defl. ($\omega = 1$)	42	42	42	43	46	47	48	48

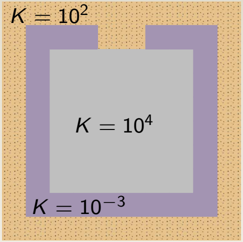


Test case taken from [Vuik et al., 2001]

Test Case III

Groundwater

degree	p=2				p=3			
mesh	N=40^2	N=80^2	N=160^2	N=320^2	N=40^2	N=80^2	N=160^2	N=320^2
two-level prec. ($\omega = 1$)	54	52	52	52	67	68	68	69
two-level prec. ($\omega = 0.7$)	38	38	38	40	41	42	42	42
two-level defl. ($\omega = 1$)	54	54	54	55	59	59	60	60



Test case taken from [Vuik et al., 2001]

Test Case IV

Anisotropic full tensor

degree	p=2				p=3			
mesh	N=40^2	N=80^2	N=160^2	N=320^2	N=40^2	N=80^2	N=160^2	N=320^2
two-level prec. ($\omega = 1$)	47	48	49	50	61	67	68	71
two-level prec. ($\omega = 0.7$)	36	37	38	38	39	39	41	42
two-level defl. ($\omega = 1$)	48	49	51	52	54	55	57	57

$$K = 10^{-3} \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

$$K = \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

This mimics a non-Cartesian mesh

Conclusion

Both two-level variants yield fast and scalable CG convergence

- The SIPG penalty parameter can best be chosen diffusion-dependent
- Coarse systems can be solved by applying AMG-CG again in an inner loop
- Without relaxation, deflation is faster due to lower smoothing costs and faster convergence
- With relaxation, the preconditioner can become equally fast

Further Research

- Compare preconditioning and deflation theoretically
- Study more challenging test cases

Contact

c.vuik@tudelft.nl

http://ta.twi.tudelft.nl/users/vuik/