A new analysis of block preconditioners for saddle point problems

Yvan Notay*

Université Libre de Bruxelles
Service de Métrologie Nucléaire

Delft, June 14, 2013

* Supported by the Belgian FNRS
http://homepages.ulb.ac.be/∼ynotay

Outline

ULB

1. Introduction
2. Taxonomy of block preconditioners
3. Relations between the preconditioners
4. Eigenvalue analysis
5. Example of application
6. Conclusions

1. Introduction

ULB

Linear systems

$$K\,\mathbf{u} = \mathbf{b}$$

where K has saddle point structure:

$$K = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}$$

In this talk: A SPD, C sym. nonnegative definite
(all results compatible with $C = 0$)

1. Introduction

ULB

Linear systems

$$K\,\mathbf{u} = \mathbf{b}$$

where K has saddle point structure:

$$K = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}$$

In this talk: A SPD, C sym. nonnegative definite
(all results compatible with $C = 0$)

- Applications
  Stokes problems (stationnary or time dependent), PDE-constrained optimization, etc
- All results presented here can be applied to any context (purely algebraic analysis)

1. Introduction

ULB

- We are interested in iterative solution schemes combining a Krylov subspace method with block preconditioners
- A preconditioner is any procedure that provides a cheap approximation to the inverse of the system matrix K
- When the preconditioner is SPD, MINRES is a choice method

  If the preconditioner is nonsymmetric or indefinite, GMRES or GCR has to be used instead

ULB

Block preconditiners

They are rooted in / equivalent to fractional steps iterative methods: approximately solve

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_v \\ \mathbf{b}_p \end{pmatrix}$$

alternating approximates solves of

$$A\,\delta\mathbf{v} = \mathbf{b}_v - A\,\hat{\mathbf{v}} - B^T\hat{\mathbf{p}}\ ;$$

i.e.:

$$\delta\mathbf{v} = M_A^{-1}\left(\mathbf{b}_v - A\,\hat{\mathbf{v}} - B^T\hat{\mathbf{p}}\right)$$

and "pressure corrections":

$$\delta\mathbf{p} = \pm\,M_S^{-1}\left(\mathbf{b}_p - B\,\overline{\mathbf{v}} + C\,\overline{\mathbf{p}}\right)$$

1. Introduction

ULB

These operations cane be combined in a numbers of ways:

- different orders
- some may be repeated
- different possibilities to define $\hat{\mathbf{v}}$, $\hat{\mathbf{p}}$, $\overline{\mathbf{v}}$, $\overline{\mathbf{p}}$ in function of previous and current approximations

These combinations give rise to the different types of block preconditioners

1. Introduction

ULB

The schemes have in comman that they need:

- A (good) preconditioner M_A for A (possibly defined only implicitly)
- An operator M_S governing the pressure correction

  We'll see (well known fact) that it needs to be a (good) preconditioner for the Schur complement

$$S = C + B\,A^{-1}B^T$$

1. Introduction

ULB

The schemes have in comman that they need:

- A (good) preconditioner M_A for A (possibly defined only implicitly)
- An operator M_S governing the pressure correction

  We'll see (well known fact) that it needs to be a (good) preconditioner for the Schur complement

$$S = C + B\,A^{-1}B^T$$

Types of contributions in the field

- Identify relevant M_A , M_S for a given application area
- Analyse one or several schemes assuming that relevant M_A and M_S have been obtained

2. Taxonomy of block preconditioners

ULB

Block Diagonal$^+$:

$$M_+ = \begin{pmatrix} M_A & \\ & M_S \end{pmatrix}$$

Block Diagonal:

$$M_d = \begin{pmatrix} M_A & \\ & -M_S \end{pmatrix}$$

Block Triangular:

$$M_t = \begin{pmatrix} M_A & B^T \\ & -M_S \end{pmatrix}$$

Inxeact Uzawa:

$$M_u = \begin{pmatrix} M_A & \\ B & -M_S \end{pmatrix}$$

Block Approximate Factorization:

$$M_f = \begin{pmatrix} M_A & \\ B & -M_S \end{pmatrix} \begin{pmatrix} I & M_A^{-1}B^T \\ & I \end{pmatrix}$$

2. Taxonomy of preconditioners (cont.)

ULB

$\widetilde{M}_A$ such that:

$$I - \widetilde{M}_A^{-1} A \;=\; \left(I - M_A^{-1} A\right)^2$$

$\rightarrow$

Block Triangular$^{(2)}$:

$$M_{t_2} = \begin{pmatrix} \widetilde{M}_A & B^T \\ & -M_S \end{pmatrix}$$

Inxeact Uzawa$^{(2)}$:

$$M_{u_2} = \begin{pmatrix} \widetilde{M}_A & \\ B & -M_S \end{pmatrix}$$

Block SGS:

$$M_g = \begin{pmatrix} I & \\ B\,M_A^{-1} & I \end{pmatrix} \begin{pmatrix} \widetilde{M}_A & \\ & -M_S \end{pmatrix} \begin{pmatrix} I & M_A^{-1} B^T \\ & I \end{pmatrix}$$

3. Relations between the precond.

ULB

Remark

The eigenvalues of $M_+^{-1}K$ are real (positive and negative)

For all other prec., the eigenvalues have positive real part

Theorem

(1)

$$\max_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \leq \max_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

$$\min_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \geq \min_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

3. Relations between the precond.

ULB

Remark

The eigenvalues of $M_+^{-1}K$ are real (positive and negative)

For all other prec., the eigenvalues have positive real part

Theorem

(1)

$$\max_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \leq \max_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

$$\min_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \geq \min_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

(2) $M_t^{-1}K$ and $M_u^{-1}K$ have the same spectrum

3. Relations between the precond.

ULB

Remark

The eigenvalues of $M_+^{-1}K$ are real (positive and negative)

For all other prec., the eigenvalues have positive real part

Theorem

(1)

$$\max_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \leq \max_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

$$\min_{\lambda\in\sigma\left(M_d^{-1}K\right)} |\lambda| \geq \min_{\lambda\in\sigma\left(M_+^{-1}K\right)} |\lambda|$$

(2) $M_t^{-1}K$ and $M_u^{-1}K$ have the same spectrum

(3) $M_g^{-1}K$, $M_{t_2}^{-1}K$ and $M_{u_2}^{-1}K$ have the same spectrum

4. Eigenvalue analysis

ULB

The real eigenvalues λ of $M_*^{-1}K$ satisfy: $\underline{\xi} \leq \lambda \leq \overline{\xi}$

$$\underline{\mu} = \lambda_{\min}\left(M_A^{-1}A\right) \qquad \overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right)$$

$$\underline{\nu} = \lambda_{\min}\left(M_S^{-1}S\right) \qquad \overline{\nu} = \lambda_{\max}\left(M_S^{-1}S\right)$$

	$\underline{\xi}$	$\overline{\xi}$
M_d	$\min\left(\underline{\mu}\,,\,\underline{\nu}\right)$	$\max\left(\overline{\mu}\,,\,\overline{\nu}\right)$
$M_u\,,\,M_t$ $(\overline{\mu} \leq 1)$	$\min\left(\underline{\mu}\,,\,\underline{\nu}\right)$	$\max\left(1\,,\,\overline{\nu}\right)$
$M_u\,,\,M_t$ $(\overline{\mu} > 1)$	$\min\left(\underline{\mu}\,,\,\eta^{-1}\underline{\nu}\right)$	$\widetilde{\nu}$
M_f	$\min\left(\underline{\mu}\,,\,\underline{\nu}\right)$	$\max\left(\overline{\mu}\,,\,\overline{\nu}\right)$
M_g ; $M_{u_2}\,,\,M_{t_2}$	$\min\left(1-\rho_A^2\,,\,\underline{\nu}\right)$	$\max\left(1\,,\,\overline{\nu}\right)$

$\eta \approx 1$

$\widetilde{\nu} \geq \overline{\mu}\,,\,\overline{\nu}$

$\widetilde{\nu} \leq \overline{\mu}(\overline{\nu}+1)$

$\rho_A = \rho\left(I - M_A^{-1}A\right)$

4. Eigenvalue analysis

ULB

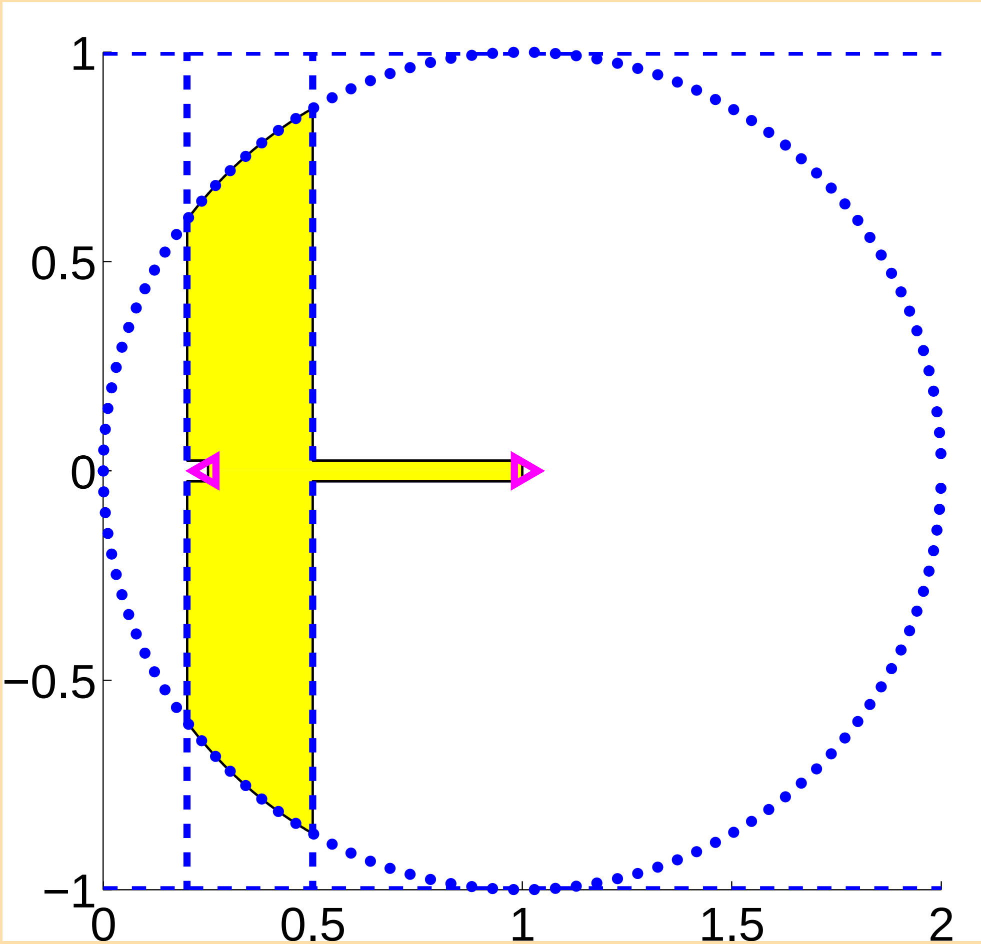
The eigenvalues with nonzero imaginary part satisfy

$$\underline{\chi} \leq \Re\mathrm{e}(\lambda) \leq \overline{\chi}, \qquad |\Im\mathrm{m}(\lambda)| \leq \delta, \qquad |\lambda - \zeta| \leq \zeta$$

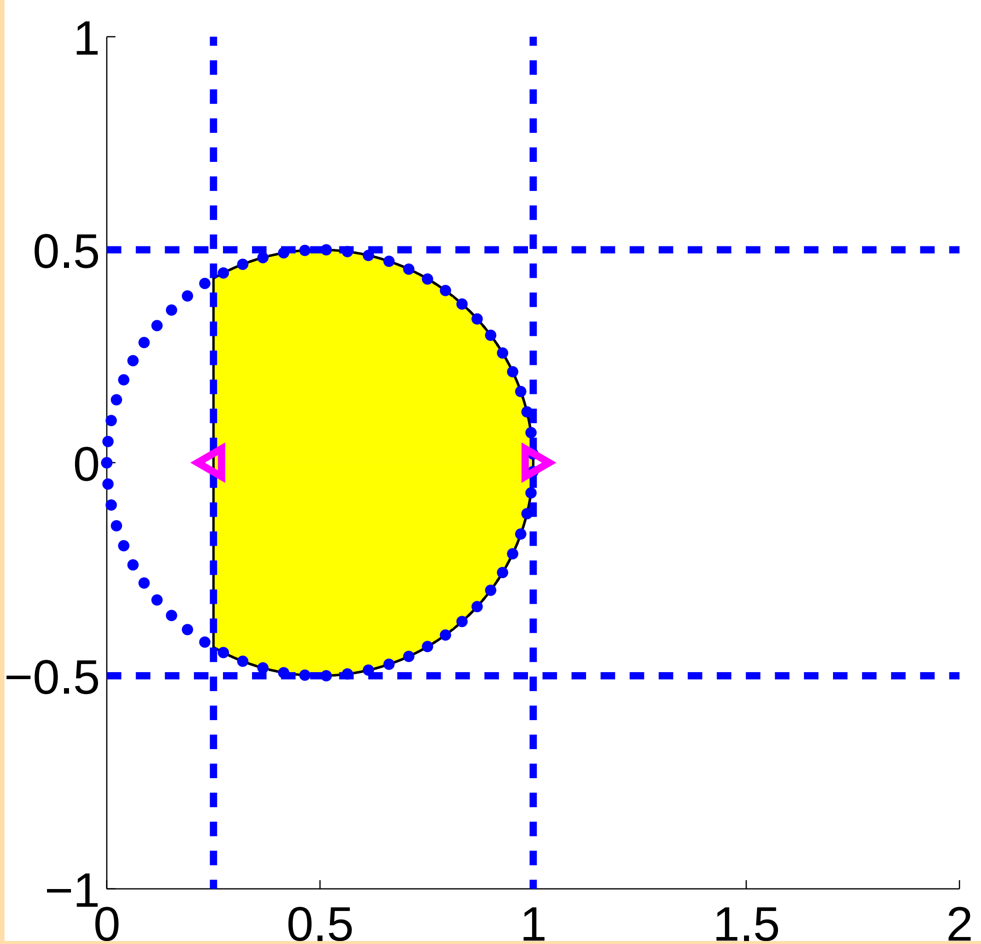
4. Eigenvalue analysis

The eigenvalues with nonzero imaginary part satisfy

$$\underline{\chi} \le \Re e(\lambda) \le \overline{\chi} \,, \qquad |\Im m(\lambda)| \le \delta \,, \qquad |\lambda - \zeta| \le \zeta$$



Block Diagonal



Block Triang. & Uzawa

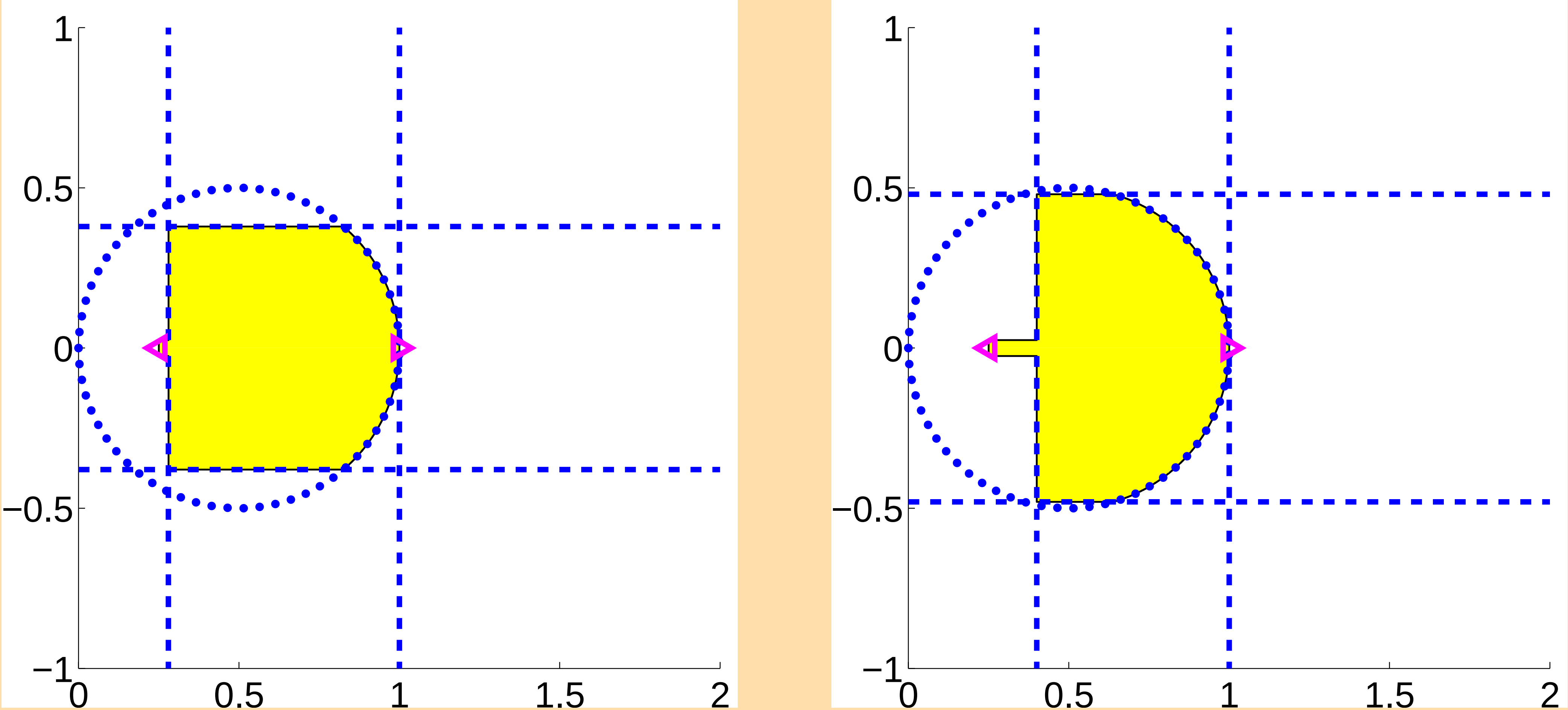
$\underline{\mu} = 0.4 \,, \ \underline{\nu} = 0.25 \,, \ \overline{\mu} = \overline{\nu} = 1$

4. Eigenvalue analysis

ULB

The eigenvalues with nonzero imaginary part satisfy

$$\underline{\chi} \leq \Re e(\lambda) \leq \overline{\chi}, \qquad |\Im m(\lambda)| \leq \delta, \qquad |\lambda - \zeta| \leq \zeta$$



Block Approx. Fact.

Block SGS

$\underline{\mu} = 0.4$, $\underline{\nu} = 0.25$, $\overline{\mu} = \overline{\nu} = 1$

4. Eigenvalue analysis

ULB

Comparison with bounds appeared previously:

- Our estimates for real eigenalues are sharper (see below)

4. Eigenvalue analysis

ULB

Comparison with bounds appeared previously:

- Our estimates for real eigenalues are sharper (see below)
- The bound in [Simoncini (2004)] for Block Triang.:

$$\Im\mathrm{m}(\lambda) \neq 0 \qquad \Rightarrow \qquad |\lambda - 1| \leq \sqrt{1-\underline{\mu}}$$

offers a useful complement.

4. Eigenvalue analysis

ULB

Comparison with bounds appeared previously:

- Our estimates for real eigenalues are sharper (see below)
- The bound in [Simoncini (2004)] for Block Triang.:

$$\Im\mathrm{m}(\lambda) \neq 0 \qquad \Rightarrow \qquad |\lambda - 1| \;\leq\; \sqrt{1-\underline{\mu}}$$

offers a useful complement.

- Ours analysis extends this result to Uzawa, and further to Block SGS, for which we get:

$$\Im\mathrm{m}(\lambda) \neq 0 \qquad \Rightarrow \qquad |\lambda - 1| \;\leq\; \rho_A$$

(via the equivalence between Block SGS and Block Triangular$^{(2)}$)

4. Eigenvalue analysis

ULB

Combining this with our bounds for real eigenvalues:

Corollary

Let

$$\rho_A = \rho\left(I - M_A^{-1}A\right), \qquad \rho_S = \rho\left(I - M_S^{-1}S\right)$$

- For Block Triang. & Uzawa, if

$$\overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right) \leq 1$$

then:

$$\rho\left(I - M_t^{-1}K\right) = \rho\left(I - M_u^{-1}K\right) \leq \max\left(\sqrt{\rho_A}\,,\,\rho_S\right)$$

- For block SGS:

$$\rho\left(I - M_g^{-1}K\right) \leq \max\left(\rho_A\,,\,\rho_S\right)$$

5. Example

ULB

Stationnary Stokes problem on the unit square Ω

Find velocity vector $\mathbf{v}$ and the kinematic pressure field p satisfying

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f} \qquad \text{in } \Omega \\ \nabla \cdot \mathbf{u} &= 0 \qquad \text{in } \Omega \end{aligned}$$

Dirichlet BC for $\mathbf{v}$ & FD MAC scheme

$\rightarrow$

$$K = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$$

Further, with $M_S = I$,

$$\underline{\nu} \approx 0.25 \qquad \text{and} \qquad \overline{\nu} = 1$$

5. Example

Preconditioner for A

- A block diagonal with each diag. block being a discrete Laplace operator

  $\rightarrow$

  well approximated by any (algebraic) multigrid method

5. Example

ULB

Preconditioner for A

- A block diagonal with each diag. block being a discrete Laplace operator
  $\rightarrow$
  well approximated by any (algebraic) multigrid method

- AGMG efficient and as simple as (in Matlab):

```
% Set up the preconditioner:
>> agmg(A,[],[],[],[],[],[],1);
% Apply prec. to x, result in y:
>> y=agmg([],x,[],[],[],[],[],3);
```

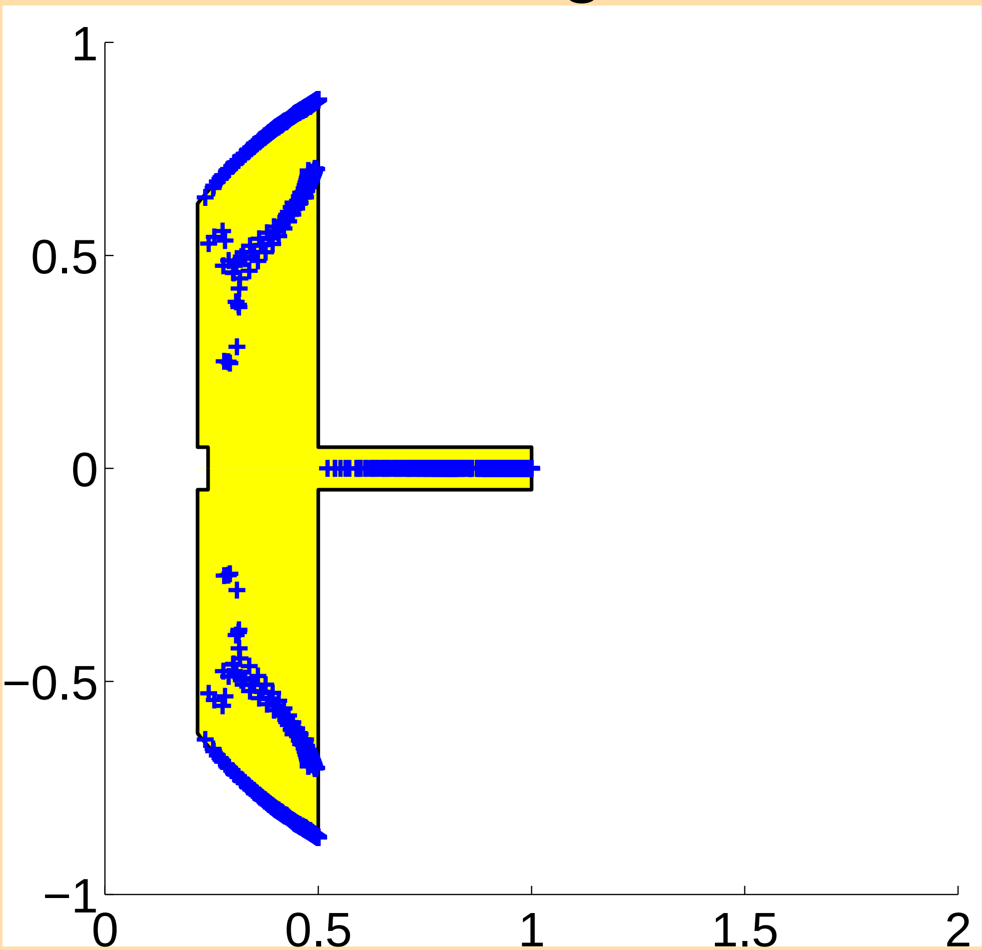
- Then:

$$\underline{\mu} \approx 0.4 \quad \text{and} \quad \overline{\mu} = 1$$

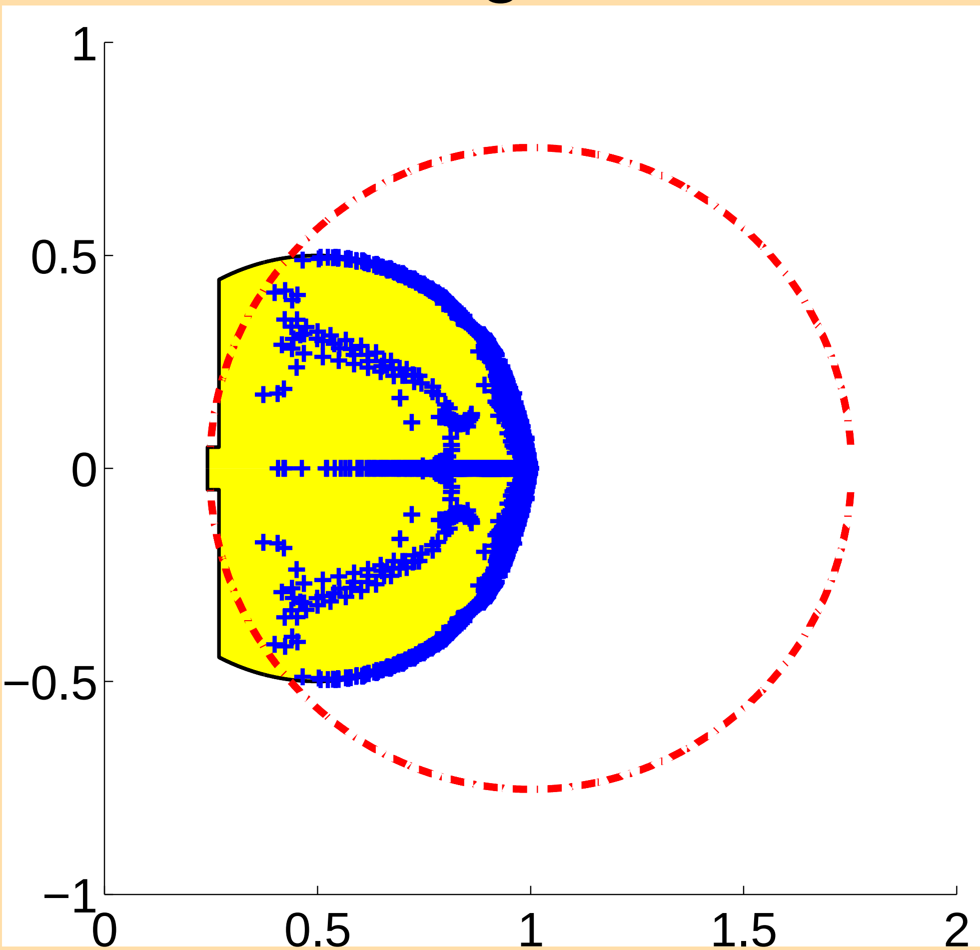
5. Example

Eigenvalue plot

Block Diagonal

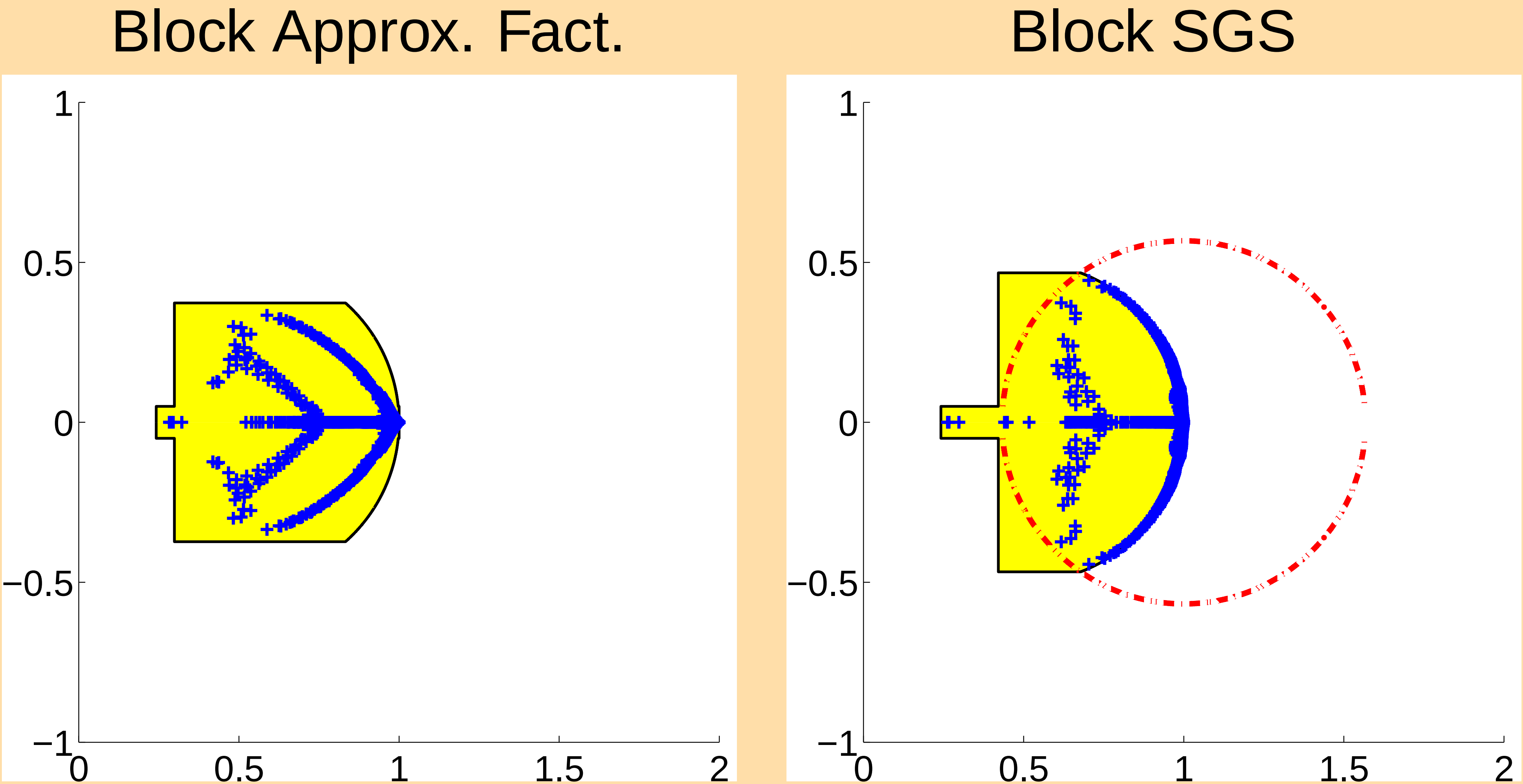


Block Triang. & Uzawa



5. Example

Eigenvalue plot



5. Example

ULB

Rescaling the preconditioner for A:

$$\underline{\mu} \approx 0.4\,\alpha \quad \text{and} \quad \overline{\mu} = \alpha$$

Makes sense for Block Triang. & Uzawa

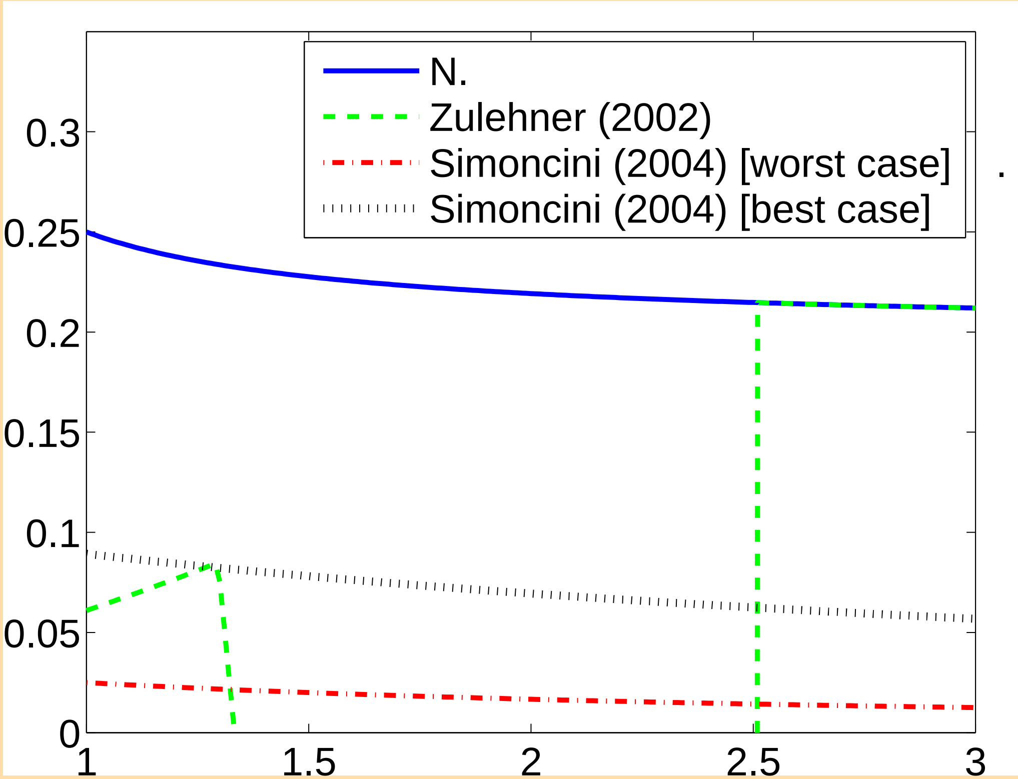
5. Example

Rescaling the preconditioner for A:

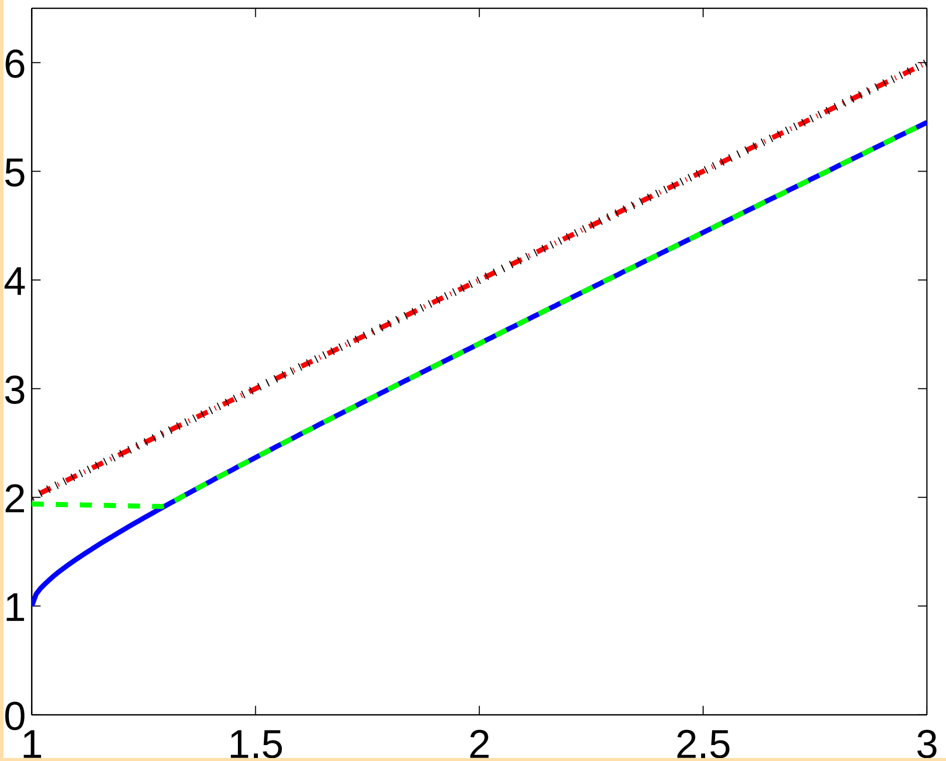
$$\underline{\mu} \approx 0.4\,\alpha \quad \text{and} \quad \overline{\mu} = \alpha$$

Makes sense for Block Triang. & Uzawa

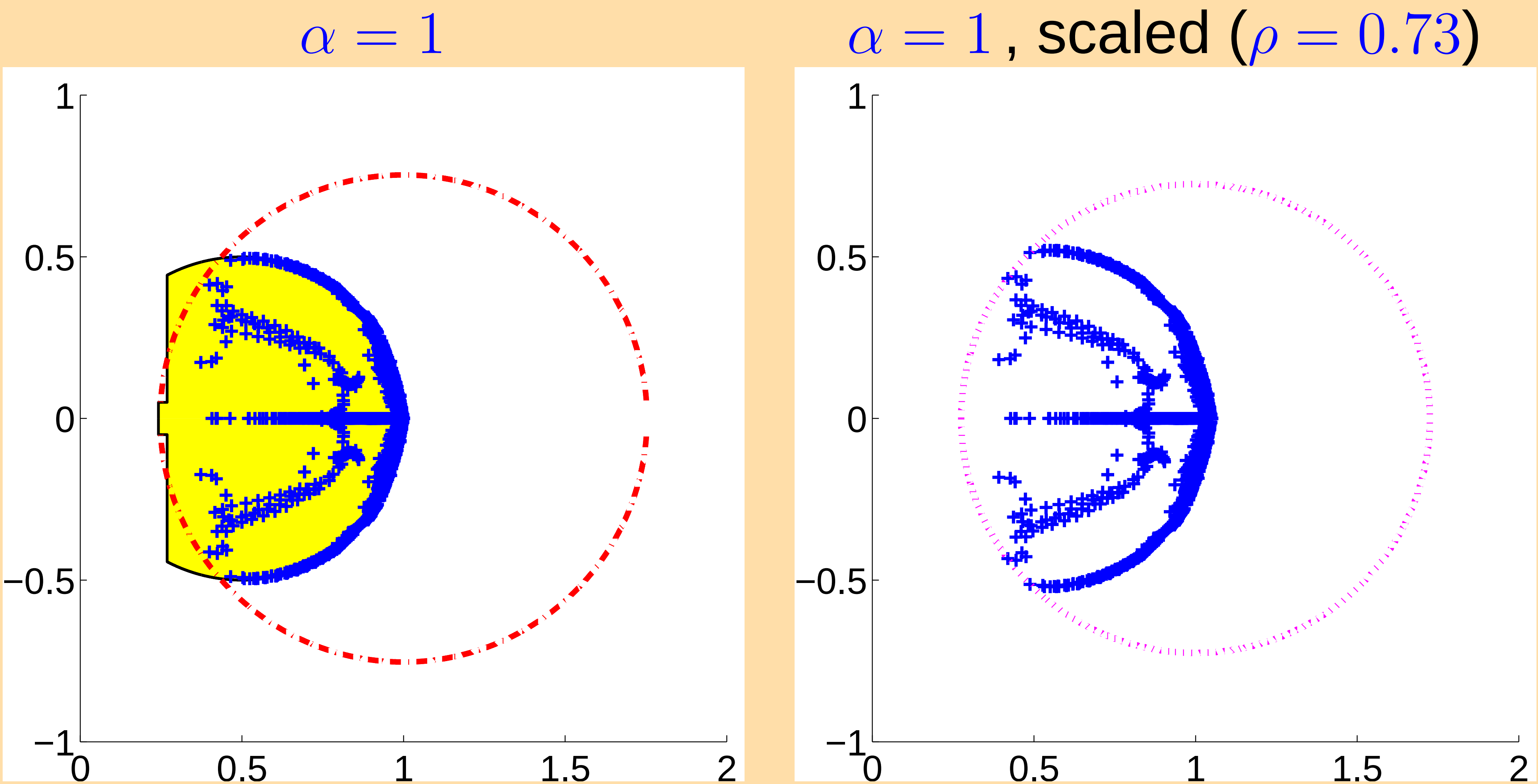
Lower bounds $= f(\alpha)$



Upper bounds $= f(\alpha)$

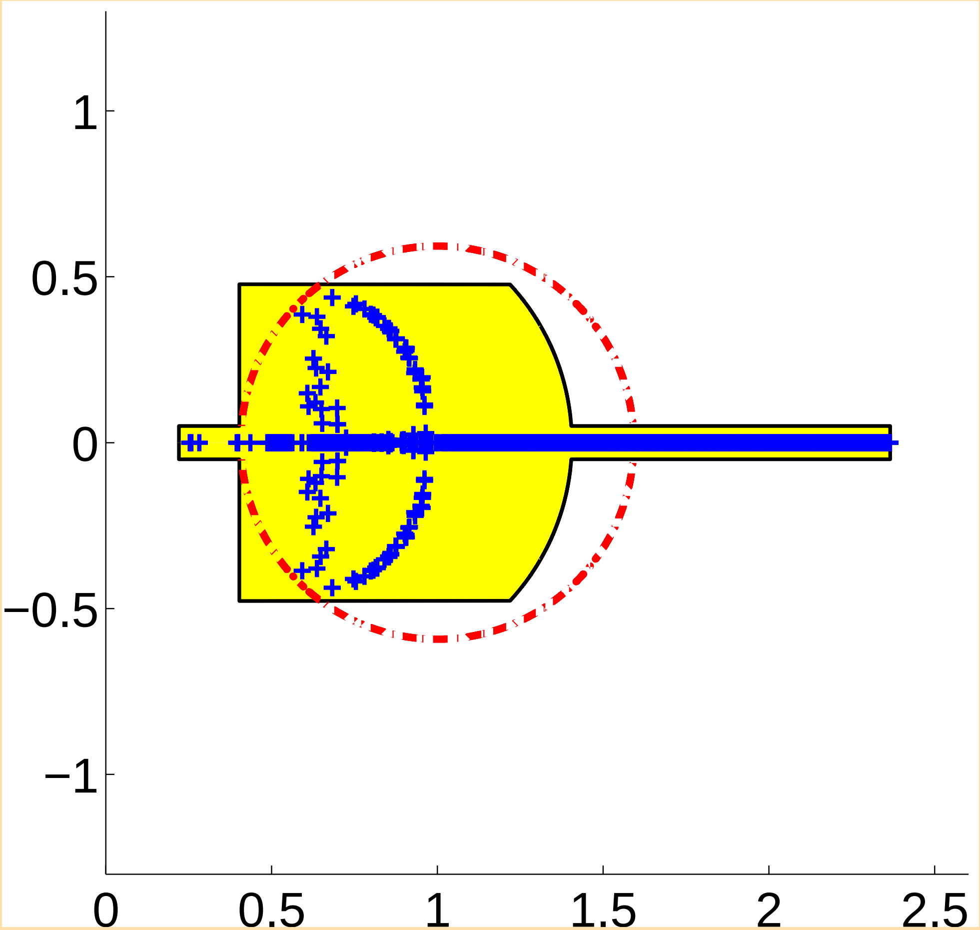


5. Example

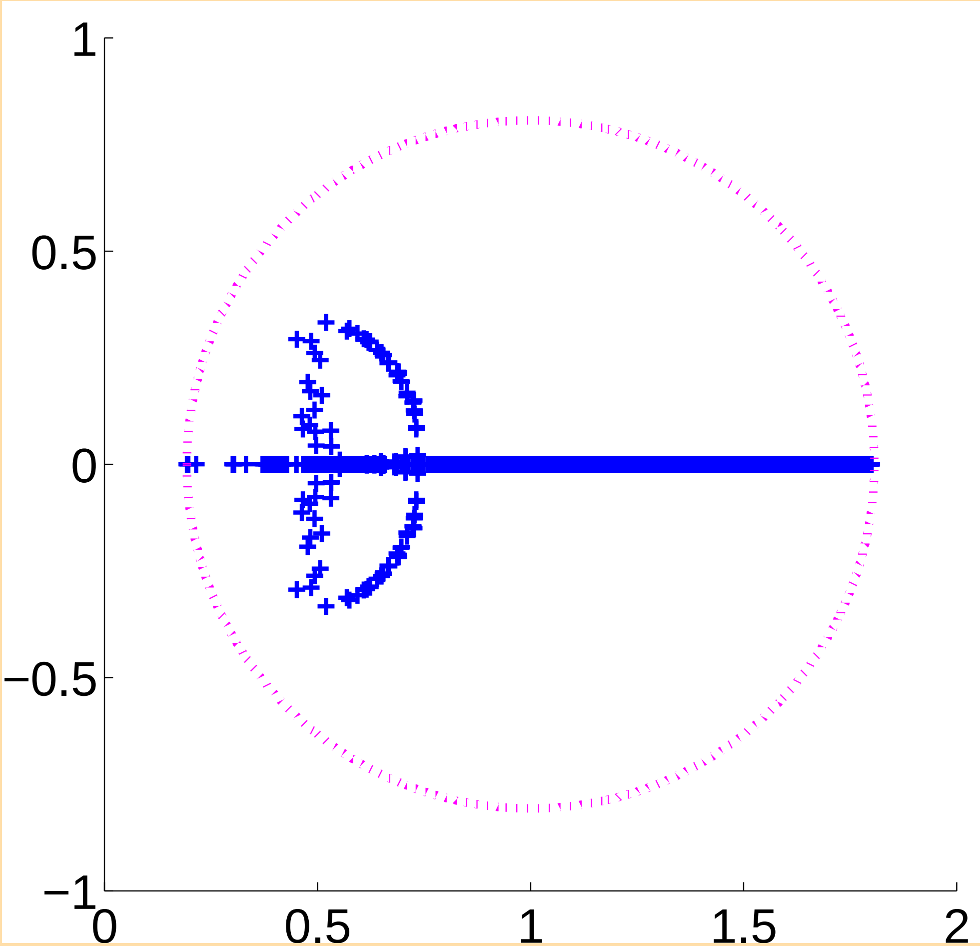


5. Example

$\alpha = 1.5$



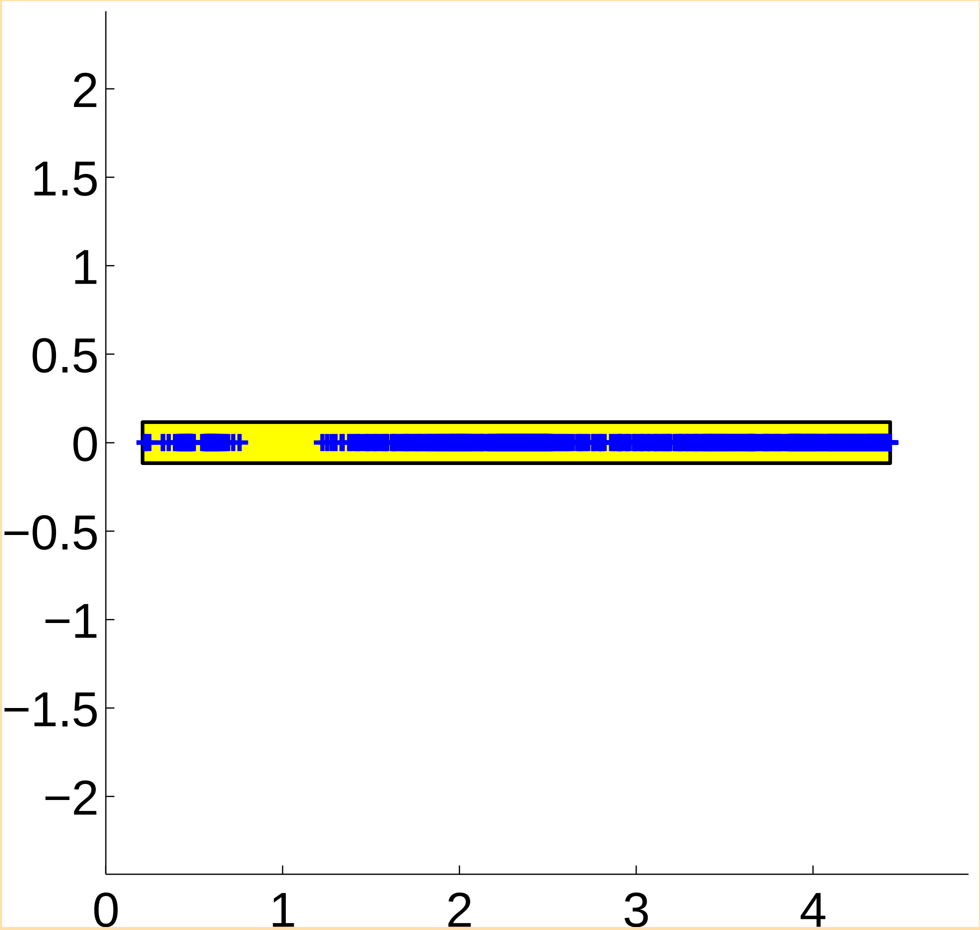
$\alpha = 1.5$, scaled ($\rho = 0.81$)



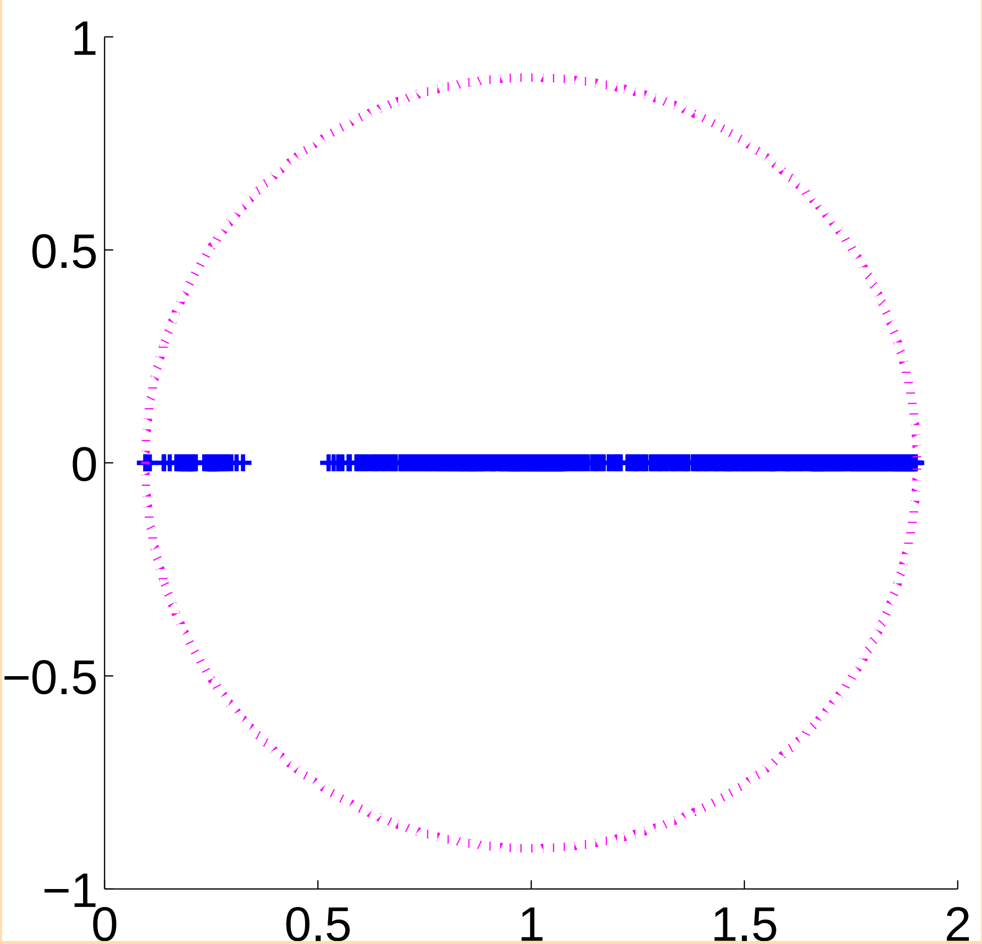
5. Example

ULB

$\alpha = 2.5$



$\alpha = 2.5$, scaled ($\rho = 0.90$)



5. Example

Number of iterations to solve the system ($\alpha = 1.5$)

h^{-1}	32	512	1024
Block Diag.$^{(+)}$ (M_+)	43	59	62
Block Diag. (M_d)	58	89	150
Block Triang. (M_t)	28	47	58
Inexact Uzawa (M_u)	30	61	57
Block Fact. (M_f)	21	29	37
Block SGS (M_g)	19	23	26
Block Triang.$^{(2)}$ (M_{t_2})	19	23	26
Inexact Uzawa$^{(2)}$ (M_{u_2})	20	26	28

Krylov:

MINRES for M_+

GCR(15) in all other cases

- AGMG becomes a variable preconditioner for $h^{-1} > 32$
- M_d performs simialrly to M_+ if larger restart is chosen

5. Example

ULB

Number of iterations to solve the system ($\alpha = 1.5$)

Preconditioner for A	AGMG			amg(ifiss 3.2)
h^{-1}	32	512	1024	512
Block Diag.$^{(+)}$ (M_+)	43	59	62	37
Block Diag. (M_d)	58	89	150	54
Block Triang. (M_t)	28	47	58	28
Inexact Uzawa (M_u)	30	61	57	28
Block Fact. (M_f)	21	29	37	25
Block SGS (M_g)	19	23	26	25
Block Triang.$^{(2)}$ (M_{t_2})	19	23	26	23
Inexact Uzawa$^{(2)}$ (M_{u_2})	20	26	28	25

5. Example

Time to solve the system ($\alpha = 1.5$)

Preconditioner for A	AGMG		amg(ifiss 3.2)
h^{-1}	512	1024	512
Block Diag.$^{(+)}$ (M_+)	16.0	17.2	28.0
Block Diag. (M_d)	26.8	49.0	40.8
Block Triang. (M_t)	15.0	19.5	21.5
Inexact Uzawa (M_u)	19.7	19.5	21.7
Block Fact. (M_f)	15.3	20.6	35.6
Block SGS (M_g)	11.8	14.0	35.3
Block Triang.$^{(2)}$ (M_{t_2})	11.2	13.5	32.0
Inexact Uzawa$^{(2)}$ (M_{u_2})	12.9	15.2	34.9

Elapsed time excluding set up in seconds per 10^6 unknowns

5. Example

Time to solve the system ($\alpha = 1.5$)

Preconditioner for A	AGMG		amg(ifiss 3.2)
h^{-1}	512	1024	512
Block Diag.$^{(+)}$ (M_+)	16.4	17.7	5717.
Block Diag. (M_d)	27.2	49.5	5730.
Block Triang. (M_t)	15.4	19.9	5711.
Inexact Uzawa (M_u)	20.1	20.0	5711.
Block Fact. (M_f)	15.7	21.1	5725.
Block SGS (M_g)	12.2	14.4	5725.
Block Triang.$^{(2)}$ (M_{t_2})	11.6	13.9	5721.
Inexact Uzawa$^{(2)}$ (M_{u_2})	13.3	15.7	5724.

Total Elapsed time (inlcuding set up) in seconds per 10^6 unknowns

6. Conclusions

ULB

- All block preconditioners work well except perhaps M_d

6. Conclusions

ULB

- All block preconditioners work well except perhaps M_d
- We are left with essentially three options:
  – postive definite block diagonal
  – block upper or lower triangular
  – block SGS

6. Conclusions

ULB

- All block preconditioners work well except perhaps M_d
- We are left with essentially three options:
  – postive definite block diagonal
  – block upper or lower triangular
  – block SGS
- What is the best scheme depends on the relative quality and cost of the preconditioners for A and S (Block SGS iterates twice on A)

6. Conclusions

ULB

- All block preconditioners work well except perhaps M_d
- We are left with essentially three options:
  – postive definite block diagonal
  – block upper or lower triangular
  – block SGS
- What is the best scheme depends on the relative quality and cost of the preconditioners for A and S (Block SGS iterates twice on A)
- There is in fact more to gain or loose in the choice of the preconditioner for the subproblems (M_A and M_S)

6. Conclusions

ULB

- All block preconditioners work well except perhaps M_d
- We are left with essentially three options:
  – postive definite block diagonal
  – block upper or lower triangular
  – block SGS
- What is the best scheme depends on the relative quality and cost of the preconditioners for A and S (Block SGS iterates twice on A)
- There is in fact more to gain or loose in the choice of the preconditioner for the subproblems (M_A and M_S)
- Be cautious with numerical results involving multigrid (especially: algebraic multigrid) and displaying only iteration counts

More details: reports available

Y. Notay, A new eigenvalue analysis of block preconditioners for saddle point problems

Francisco J. Gaspar, Y. Notay, Cornelis W. Oosterlee and Carmen Rodrigo, A simple and efficient segregated smoother for the discrete Stokes equations

More details: reports available

Y. Notay, A new eigenvalue analysis of block preconditioners for saddle point problems

Francisco J. Gaspar, Y. Notay, Cornelis W. Oosterlee and Carmen Rodrigo, A simple and efficient segregated smoother for the discrete Stokes equations

Thank you for your attention !