# Backward error analysis and stopping criteria for Krylov space method

Mario Arioli

m.arioli@rl.ac.uk

CCLRC-Rutherford Appleton Laboratory

# Outline

- Backward error and norms
- Krylov methods and stopping criteria
- Elliptic problems
- Conclusions and future work

# Linear systems

$$\mathbf{Au} = \mathbf{b}$$

$\mathbf{A} \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}$ positive definite

# Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

# Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

- Existence and uniqueness:

$$\begin{aligned} a(\mathbf{v}, \mathbf{w}) &\leq C_1 \|\mathbf{v}\|_{\mathcal{H}} \|\mathbf{w}\|_{\mathcal{H}} \\ \sup_{\mathbf{v} \in \mathcal{H} \setminus \{0\}} \frac{a(\mathbf{v}, \mathbf{w})}{\|\mathbf{v}\|_{\mathcal{H}}} &\geq C_2 \|\mathbf{w}\|_{\mathcal{H}} \end{aligned}$$

# Linear systems: variational framework

- Find $\mathbf{u} \in \mathcal{H}$ such that for all $\mathbf{v} \in \mathcal{H}$

$$a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}) \qquad (L(\cdot) \in \mathcal{H}' \text{ dual space of } \mathcal{H})$$

- Existence and uniqueness:

$$\begin{aligned} a(\mathbf{v}, \mathbf{w}) &\leq C_1 \|\mathbf{v}\|_{\mathcal{H}} \|\mathbf{w}\|_{\mathcal{H}} \\ \sup_{\mathbf{v} \in \mathcal{H} \setminus \{0\}} \frac{a(\mathbf{v}, \mathbf{w})}{\|\mathbf{v}\|_{\mathcal{H}}} &\geq C_2 \|\mathbf{w}\|_{\mathcal{H}} \end{aligned}$$

- $\mathcal{H} = (\mathbf{R}^N, \|\cdot\|_H)$ and $\mathcal{H}' = (\mathbf{R}^N, \|\cdot\|_{H^{-1}})$ $H$ SPD

# Backward error

We have the following equivalence:

$$\left.\begin{array}{l}\exists b(\cdot,\cdot) \in \mathcal{BL}(\mathcal{H}), \exists \delta L \in \mathcal{H}' \text{ such that:} \\ a(\tilde{\mathbf{u}},\mathbf{v}) + b(\tilde{\mathbf{u}},\mathbf{v}) = (L+\delta L)(\mathbf{v}), \\ \forall \mathbf{v} \in \mathcal{H}V, \text{ and} \\ \|b(\cdot,\cdot)\|_{\mathcal{BL}(\mathcal{H})} \le \alpha, \|\delta L\|_{\mathcal{H}'} \le \beta \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \|\rho_{\tilde{\mathbf{u}}}\|_{\mathcal{H}'} \le \alpha \|\tilde{\mathbf{u}}\|_{\mathcal{H}} + \beta \\ \text{where } \rho_{\tilde{\mathbf{u}}} \in \mathcal{H}' \text{ is defined by} \\ \langle \rho_{\tilde{\mathbf{u}}}, \mathbf{v}\rangle_{\mathcal{H}',\mathcal{H}} = a(\tilde{\mathbf{u}},\mathbf{v}) - L(\mathbf{v}) \\ \forall \mathbf{v} \in \mathcal{H} \end{array}\right.$$

(Rigal and Gaches, A., Noulard, and Russo)

# The symmetric case: conjugate gradient method

$A$ symmetric positive definite

$\mathcal{H} = (\mathbf{R}^N, \|\cdot\|_A)$ and $\mathcal{H}' = (\mathbf{R}^N, \|\cdot\|_{A^{-1}})$

At each step $k$ the conjugate gradient method minimizes the energy norm of the error $\delta\mathbf{u}^{(k)} = \mathbf{u} - \mathbf{u}^{(k)}$ on a Krylov space $\mathbf{u}^{(0)} + \mathcal{K}_k$:

$$\min_{\mathbf{u}^{(k)} \in \mathbf{u}^{(0)} + \mathcal{K}_k} \|\delta\mathbf{u}\|_A^2.$$

$$\|\delta\mathbf{u}\|_A = \|\rho_{\mathbf{u}^{(k)}}\|_{\mathcal{H}} = \|r^{(k)}\|_{A^{-1}}$$

$$r^{(k)} = \mathbf{b} - A\mathbf{u}^{(k)}$$

# The symmetric case: stopping criteria

- Classical Criterion

$$\text{IF } \|A\mathbf{u}^{(k)} - \mathbf{b}\|_2 \leq \sqrt{\varepsilon}\|\mathbf{b}\|_2 \text{ THEN STOP },$$

## The symmetric case: stopping criteria

- Classical Criterion

$$\text{IF } \|A\mathbf{u}^{(k)} - \mathbf{b}\|_2 \leq \sqrt{\varepsilon}\|\mathbf{b}\|_2 \text{ THEN STOP ,}$$

- New Criterion

$$\text{IF } \|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}} \leq \eta\|\mathbf{b}\|_{A^{-1}} \text{ THEN STOP ,}$$

with $\eta < 1$ an a-priori threshold fixed by the user

# The symmetric case: stopping criteria

- Classical Criterion

$$\text{IF } \|A\mathbf{u}^{(k)} - \mathbf{b}\|_2 \leq \sqrt{\varepsilon}\|\mathbf{b}\|_2 \text{ THEN STOP },$$

- New Criterion

$$\text{IF } \|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}} \leq \eta\|\mathbf{b}\|_{A^{-1}} \text{ THEN STOP },$$

with $\eta < 1$ an a-priori threshold fixed by the user The choice of $\eta$ will depend on the properties of the problem that we want to solve, and, in the practical cases, $\eta$ can be frequently much larger than $\varepsilon$, the roundoff unit of the computer finite precision arithmetic.

CCLRC
Rutherford Appleton Laboratory

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}} \qquad ?$$

CCLRC Rutherford Appleton Laboratory

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}} \qquad ?$$

$$\|\mathbf{b}\|_{A^{-1}} \qquad ?$$

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichy, 2002)
  numerically stable

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichy, 2002) numerically stable
- Gauss quadrature rules (Golub and Meurant, 1997)

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichy, 2002) numerically stable
- Gauss quadrature rules (Golub and Meurant, 1997)
  - Gauss equivalent to Hestenes-Stiefel rule (Strakoš and Tichy). The Gauss quadrature does not require any a-priori knowledge of the smallest and the biggest eigenvalues and computes a lower bound of $\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$.

# The symmetric case: stopping criteria cont.

$$\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$$

- Hestenes-Stiefel rule (1952) (see Strakoš and Tichy, 2002) numerically stable
- Gauss quadrature rules (Golub and Meurant, 1997)
  - Gauss equivalent to Hestenes-Stiefel rule (Strakoš and Tichy). The Gauss quadrature does not require any a-priori knowledge of the smallest and the biggest eigenvalues and computes a lower bound of $\|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}$.
  - Gauss-Lobatto and Gauss-Radau. They compute lower and upper bounds using the extremes eigenvalues of $A$.

# The symmetric case: Hestenes-Stiefel rule

The conjugate gradient iterates compute:

$$
\begin{aligned}
\mathbf{u}^{(k)} &= \mathbf{u}^{(k-1)} + \alpha_{k-1} p^{(k-1)}, \quad \alpha_{k-1} = \frac{r^{(k-1)T} r^{(k-1)}}{p^{(k-1)T} A p^{(k-1)}}\,, \\
r^{(k)} &= r^{(k-1)} - \alpha_{k-1} A p^{(k-1)}\,, \\
p^{(k)} &= r^{(k)} + \beta_{k-1} p^{(k-1)}, \quad \beta_{k-1} = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}},
\end{aligned}
$$

where $u^{(0)} = 0$ and $r^{(0)} = p^{(0)} = b$.

# The symmetric case: Hestenes-Stiefel rule

The conjugate gradient iterates compute:

$$
\begin{aligned}
\mathbf{u}^{(k)} &= \mathbf{u}^{(k-1)} + \alpha_{k-1} p^{(k-1)}, \quad \alpha_{k-1} = \frac{r^{(k-1)T} r^{(k-1)}}{p^{(k-1)T} A p^{(k-1)}}\,, \\
r^{(k)} &= r^{(k-1)} - \alpha_{k-1} A p^{(k-1)}\,, \\
p^{(k)} &= r^{(k)} + \beta_{k-1} p^{(k-1)}, \quad \beta_{k-1} = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}},
\end{aligned}
$$

where $u^{(0)} = 0$ and $r^{(0)} = p^{(0)} = b$.

$p^{(j)T} A p^{(i)} = 0, \qquad i \neq j$ and

# The symmetric case: Hestenes-Stiefel rule

The conjugate gradient iterates compute:

$$\begin{aligned}
\mathbf{u}^{(k)} &= \mathbf{u}^{(k-1)} + \alpha_{k-1} p^{(k-1)}, \quad \alpha_{k-1} = \frac{r^{(k-1)T} r^{(k-1)}}{p^{(k-1)T} A p^{(k-1)}}\,, \\
r^{(k)} &= r^{(k-1)} - \alpha_{k-1} A p^{(k-1)}\,, \\
p^{(k)} &= r^{(k)} + \beta_{k-1} p^{(k-1)}, \quad \beta_{k-1} = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}},
\end{aligned}$$

where $u^{(0)} = 0$ and $r^{(0)} = p^{(0)} = b$.

$p^{(j)T} A p^{(i)} = 0, \qquad i \neq j$ and $\mathbf{u} = \sum_{j=1}^{N} \alpha_j p^{(j)}$ ,

# The symmetric case: Hestenes-Stiefel rule

The conjugate gradient iterates compute:

$$
\begin{aligned}
\mathbf{u}^{(k)} &= \mathbf{u}^{(k-1)} + \alpha_{k-1} p^{(k-1)}, \quad \alpha_{k-1} = \frac{r^{(k-1)T} r^{(k-1)}}{p^{(k-1)T} A p^{(k-1)}}\,, \\
r^{(k)} &= r^{(k-1)} - \alpha_{k-1} A p^{(k-1)}\,, \\
p^{(k)} &= r^{(k)} + \beta_{k-1} p^{(k-1)}, \quad \beta_{k-1} = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}},
\end{aligned}
$$

where $u^{(0)} = 0$ and $r^{(0)} = p^{(0)} = b$.

$p^{(j)T} A p^{(i)} = 0,\qquad i \neq j$ and $\mathbf{u} = \sum_{j=1}^{N} \alpha_j p^{(j)}$ ,

$$
\|\delta \mathbf{u}^{(k)}\|_A^2 = \|A\mathbf{u}^{(k)} - \mathbf{b}\|_{A^{-1}}^2 = e_A^{(k)} = \sum_{j=k+1}^{N} \alpha_j r^{(j)T} r^{(j)}.
$$

# The symmetric case: Hestenes-Stiefel rule

Under the assumption that $e_A^{(k+d)} << e_A^{(k)}$, where the integer $d$ denotes a suitable delay, the Hestenes and Stiefel estimate $\xi_k$ will be

$$\xi_k = \sum_{j=k+1}^{k+d} \alpha_j r^{(j)T} r^{(j)}.$$

# The symmetric case: Hestenes-Stiefel rule

Under the assumption that $e_A^{(k+d)} << e_A^{(k)}$, where the integer $d$ denotes a suitable delay, the Hestenes and Stiefel estimate $\xi_k$ will be

$$\xi_k = \sum_{j=k+1}^{k+d} \alpha_j r^{(j)T} r^{(j)}.$$

The choice of a value for $d$ depends on preconditioner and ill-conditioning.

# $\mathbf{b}^T A^{-1}\mathbf{b}$

From

$$r^{(k)T}\mathbf{v} = 0, \qquad \forall \mathbf{v} \in \mathcal{K}_k \ .$$

we prove

$$\mathbf{b}^T A^{-1}\mathbf{b} = \mathbf{u}^T A\mathbf{u} \geq \mathbf{u}^{(k)T} r^{(0)} + \mathbf{b}^T \mathbf{u}^{(0)},$$

(the right-hand side will converge monotonically to $||\mathbf{u}||_A^2$).

Therefore, we use the following stopping criterion

IF $\quad \xi_k \leq \eta^2 (\mathbf{u}^{(k)T} r^{(0)} + \mathbf{b}^T \mathbf{u}^{(0)}) \quad$ THEN STOP.

# Preconditioning

The dual norm of the preconditioned residual is equal to the dual norm of the original residual.

CCLRC Rutherford Appleton Laboratory

# Continuous problem

$$a(u,v) = \int_\Omega \mathfrak{K}(\mathbf{x})\nabla u \cdot \nabla v d\mathbf{x}, \quad \forall u, v \in H_0^1(\Omega)$$

$\forall u, v \in H_0^1(\Omega)$, $\exists \gamma \in \mathbb{R}_+$ and $\exists M \in \mathbb{R}_+$ such that

$$\begin{aligned} \gamma ||u||_{1,\Omega}^2 \leq \; & a(u,u) \\ & a(u,v) \; \leq M||u||_{1,\Omega}||v||_{1,\Omega} \; , \end{aligned}$$

$L(v) = \int_\Omega fv d\mathbf{x}$, $L(v) \in H^{-1}(\Omega)$.

$$(P) \begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that} \\ a(u,v) = L(v), \;\; \forall v \in H_0^1(\Omega), \end{cases}$$ has a unique solution.

# Finite-element approximation

- Weak formulation

$$\text{Find } u_h(\mathbf{x}) \in \mathcal{H}_h \text{ such that for all } v_h(\mathbf{x}) \in \mathcal{H}_h$$

$$a_h(u_h, v_h) = L_h(v_h),$$

Finite element methods choose $\mathcal{H}_h$ to be a space of functions $v_h$ defined on a subdivision $\Omega_h$ of $\Omega$ into simplices $T$ of diameter $h_T$; $h$ denotes a piecewise constant function defined on $\Omega_h$ via $h|_T = h_T$. $\mathcal{H}_h \subset \mathcal{H}$.

# Finite-element approximation

- Weak formulation

$$\text{Find } \; u_h(\mathbf{x}) \in \mathcal{H}_h \; \text{ such that for all } v_h(\mathbf{x}) \in \mathcal{H}_h$$

$$a_h(u_h, v_h) = L_h(v_h),$$

Finite element methods choose $\mathcal{H}_h$ to be a space of functions $v_h$ defined on a subdivision $\Omega_h$ of $\Omega$ into simplices $T$ of diameter $h_T$; $h$ denotes a piecewise constant function defined on $\Omega_h$ via $h|_T = h_T$.

$$\mathcal{H}_h \subset \mathcal{H}.$$

- Existence and uniqueness:

$$\begin{aligned} a_h(v_h, w_h) &\le C_1 \|v_h\|_{\mathcal{H}_h} \|w_h\|_{\mathcal{H}_h} \\ \sup_{v_h \in \mathcal{H}_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{\|v_h\|_{\mathcal{H}_h}} &\ge C_2 \|w_h\|_{\mathcal{H}_h} \end{aligned}$$

- Error estimate: $\|u - u_h\|_{\mathcal{H}_h} \le C(h)$.

# Finite element framework

Solve

$$A\mathbf{u}_h = \mathbf{b}$$

given

$$\sup_{\mathbf{w}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \sup_{\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \frac{\mathbf{w}^T A\mathbf{v}}{\|\mathbf{v}\|_H \|\mathbf{w}\|_H} \leq C_1 \qquad \text{(sup-sup)}$$

$$\inf_{\mathbf{w}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \sup_{\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \frac{\mathbf{w}^T A\mathbf{v}}{\|\mathbf{v}\|_H \|\mathbf{w}\|_H} \geq C_2 \qquad \text{(inf-sup)}$$

Note: $\|v_h\|_{\mathcal{H}_h} = \|\mathbf{v}_h\|_A$.

# Finite element framework

Finally, assuming that $h < 1$ and $t > 0$, we have

$$\|u(\mathbf{x}) - u_h^{(k)}\|_{\mathcal{H}} \leq C^*(h^t)\|u\|_{\mathcal{H}} + 2\,\|u - u_h\|_{\mathcal{H}} \leq C(h)$$

where

- $u$ is the exact solution of variational problem,
- $u_h = \sum_{i=1}^{N} \mathbf{u}_i^{(k)} \phi_i(\mathbf{x})$ is the exact solution of the approximate problem,
- $\mathbf{u}_h^{(k)}(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{u}_i^{(k)} \phi_i(\mathbf{x})$ is the approximate solution at step $k$

($\phi_i(\mathbf{x})$ are the basis functions)

# Test problems

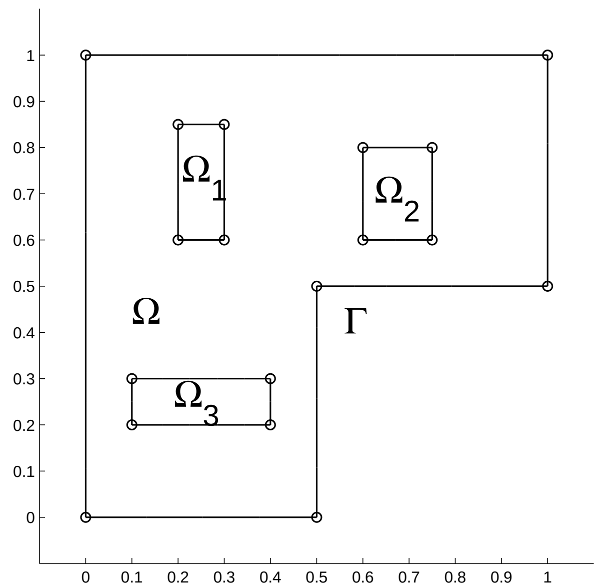
Problem 1

$$\mathfrak{K}(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega \setminus \{\Omega_1 \cup \Omega_2 \cup \Omega_3\} \\ 10^{-6} & \mathbf{x} \in \Omega_1 \\ 10^{-4} & \mathbf{x} \in \Omega_2 \\ 10^{-2} & \mathbf{x} \in \Omega_3 \end{cases}$$

Problem 2

$$\mathfrak{K}(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega \setminus \{\Omega_1 \cup \Omega_2 \cup \Omega_3\} \\ 10^{6} & \mathbf{x} \in \Omega_1 \\ 10^{4} & \mathbf{x} \in \Omega_2 \\ 10^{2} & \mathbf{x} \in \Omega_3 \end{cases}$$

$$L(\mathbf{v}) = \int_\Omega 10\mathbf{v}d\mathbf{x},\ \forall \mathbf{v} \in H_0^1(\Omega)$$



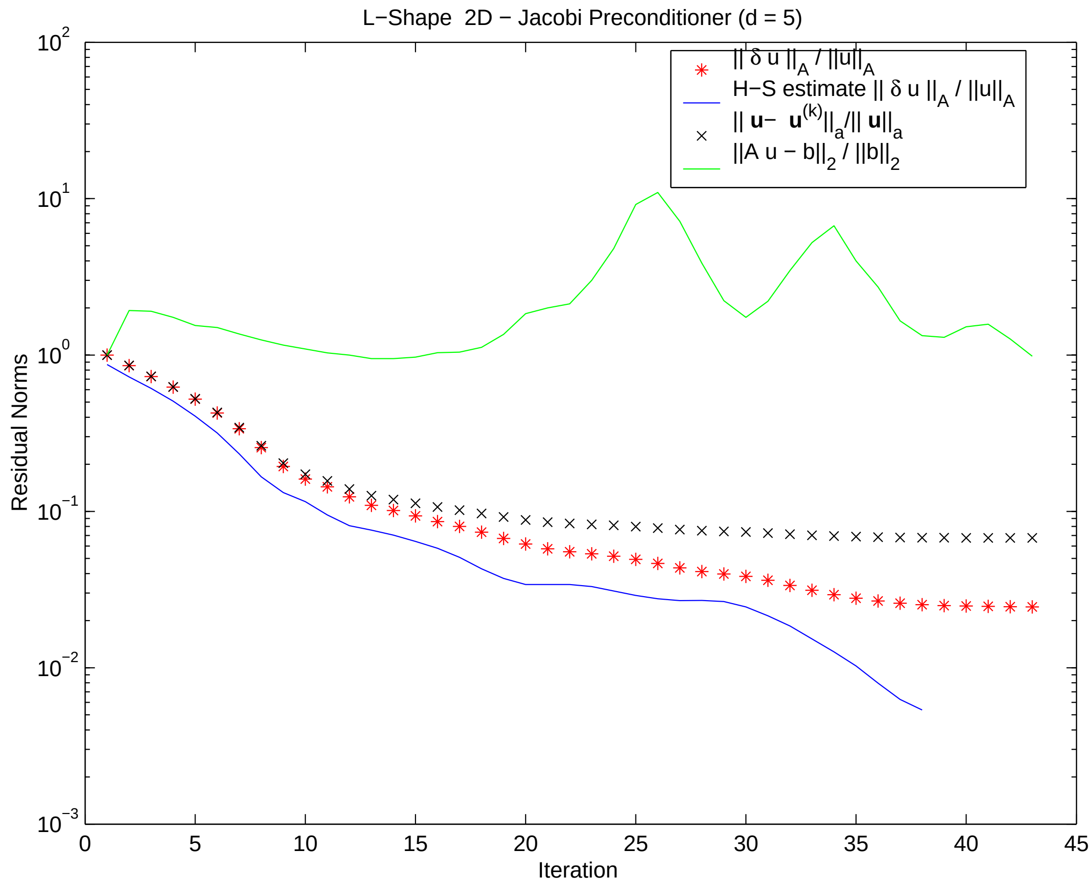
# Preconditioners: estimates for $\kappa(M^{-1}A)$.

| $M$ | Problem 1 | Problem 2 |
|---|---|---|
| $I$ | 3.6 $10^8$ | 1.8 $10^{10}$ |
| Jacobi | 2.4 $10^4$ | 1.5 $10^9$ |
| Inc. Cholesky(0) | 7.2 $10^3$ | 4.3 $10^8$ |

$\eta^2 = 3.44305 \times 10^{-5}$ and $N = 29619$

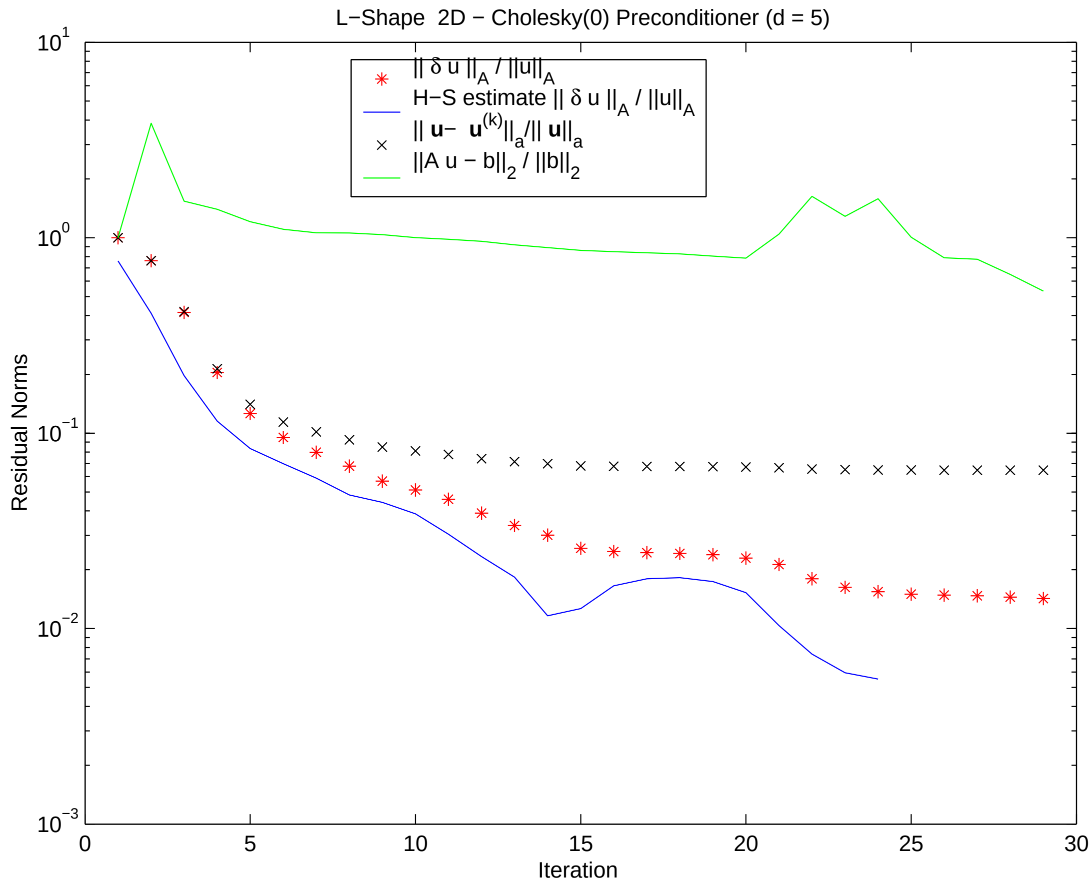
The condition numbers of the preconditioned matrices $M^{-1}A$ for the second problem are are still very high, and only the incomplete Cholesky preconditioner with drop tolerance $10^{-2}$ is an effective choice.

# Example: Problem 1



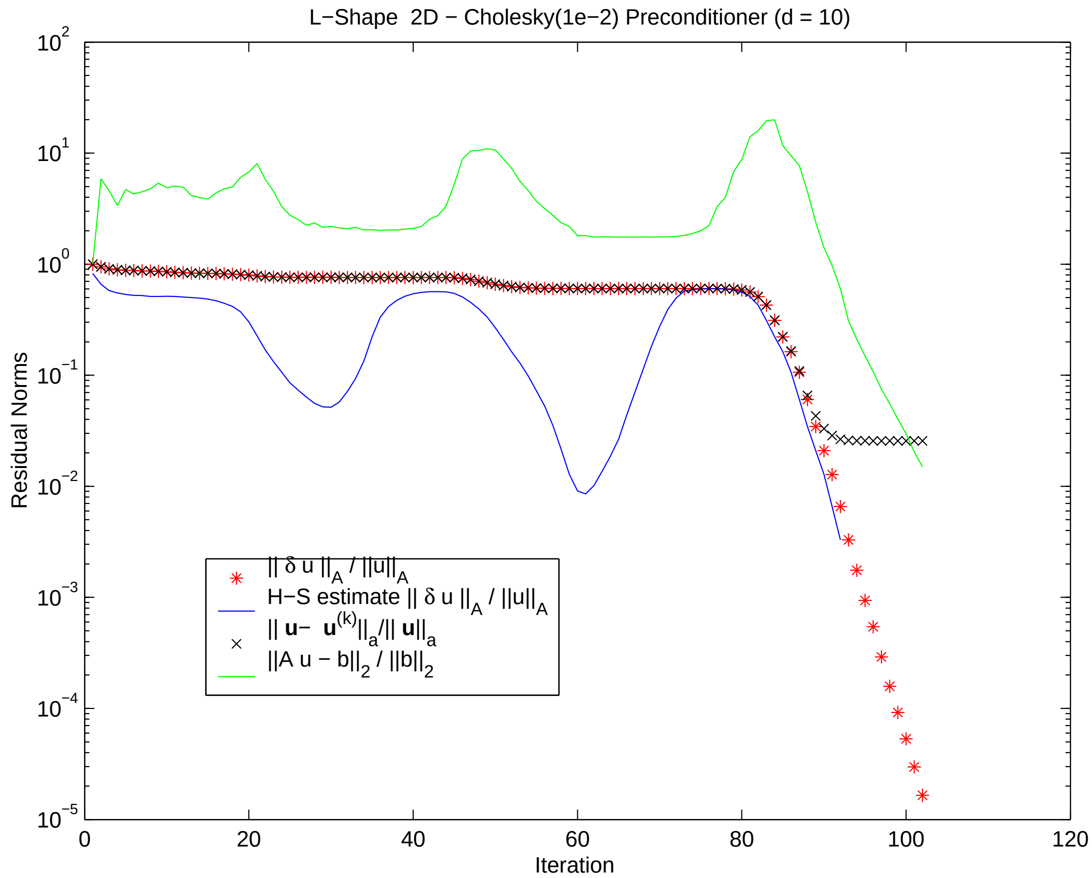

Behaviour of the norms of the residual for the Jacobi preconditioner.

# Example: Problem 1



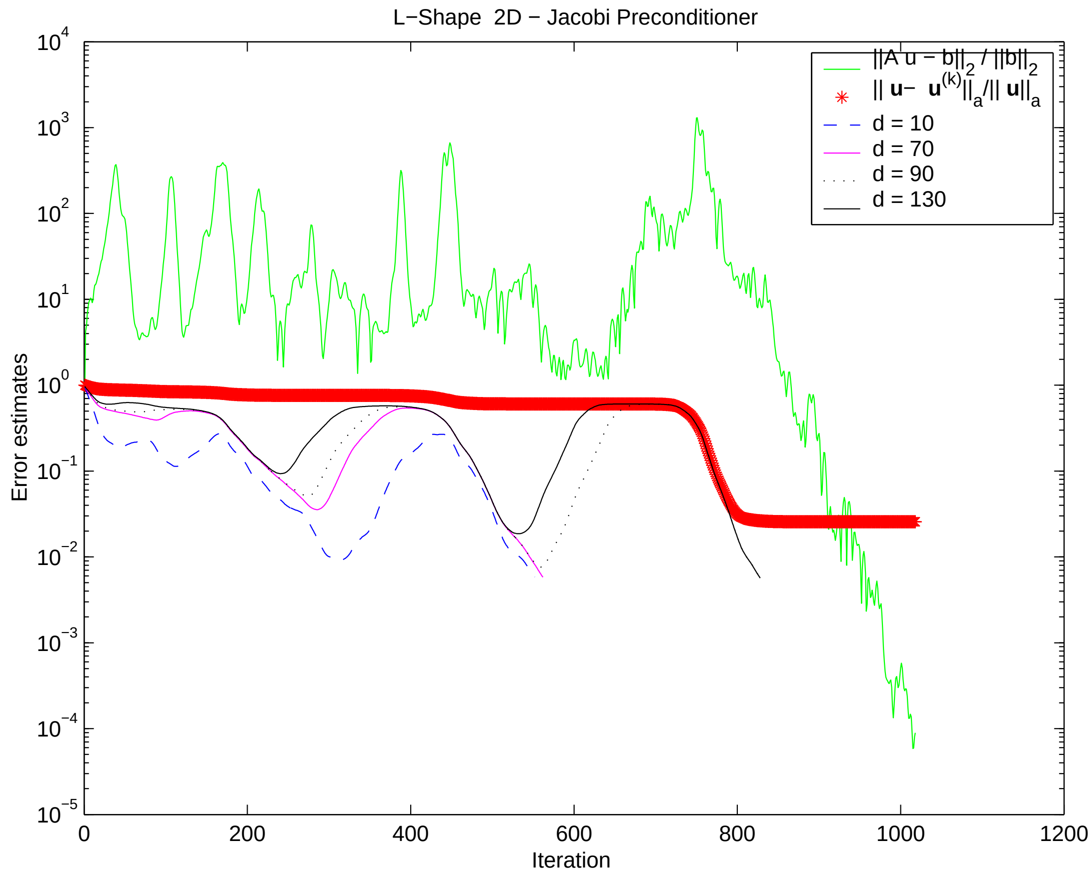

Behaviour of the norms of the residual for the incomplete Cholesky preconditioner.

# Example: Problem 2




Behaviour of the norms of the residual for the incomplete Cholesky preconditioner with drop tolerance $10^{-2}$ and $d = 10$.

# Example: Problem 2




Comparison of several estimates of the energy error for $d = 10, 70, 90, 130$ in Problem 2.

# The positive definite problem

- $a(\mathbf{u}, \mathbf{v}) \neq a(\mathbf{v}, \mathbf{u})$
- $A$ asymmetric but positive
- $H = \frac{1}{2}(A^T + A)$ SPD

How to calculate $\|\mathbf{r}^k\|_{H^{-1}}$?

- Solve preconditioned system

$$H^{-1/2} A H^{-1/2} \hat{\mathbf{u}} = H^{-1/2} \mathbf{f}$$

  - $\|\hat{\mathbf{r}}^{(k)}\|_{l_2} = \|\mathbf{r}^{(k)}\|_{H^{-1}}$
  - 3-term recurrence.

- Approximate it from Krylov subspace information.

See A., Login, and Wathen RAL-TR-2003-009

Current reports available by anonymous ftp to ftp.numerical.rl.ac.uk in directory pub/reports.

# One crime

Replace

$$\|u - u_h\|_{\mathcal{H}_h} \leq C(h)$$

with

$$\|u - u_h^k\|_{\mathcal{H}_h} \leq C(h)$$

# One crime

Replace

$$\|u - u_h\|_{\mathcal{H}_h} \leq C(h)$$

with

$$\|u - u_h^k\|_{\mathcal{H}_h} \leq C(h)$$

Sufficient condition

$$\|u - u_h\|_{\mathcal{H}_h} + \|u_h - u_h^k\|_{\mathcal{H}_h} \sim O(C(h))$$

$$\Downarrow$$

$$\|u_h - u_h^k\|_{\mathcal{H}_h} \sim O(C(h))$$

CCLRC Rutherford Appleton Laboratory

# Stopping criteria

A general stopping criterion:

$$\|u_h - u_h^{(k)}\|_{\mathcal{H}_h} = \|\mathbf{u} - \mathbf{u}^{(k)}\|_H \leq C(h)$$

# Stopping criteria

A general stopping criterion:

$$\|u_h - u_h^{(k)}\|_{\mathcal{H}_h} = \|\mathbf{u} - \mathbf{u}^{(k)}\|_H \leq C(h)$$

Residual equation

$$\mathbf{r}^{(k)} = A(\mathbf{u} - \mathbf{u}^{(k)})$$

$$\Downarrow$$

$$\|\mathbf{u} - \mathbf{u}^{(k)}\|_H = \|A^{-1}\mathbf{r}^{(k)}\|_H = \|\mathbf{r}^{(k)}\|_{A^{-T}HA^{-1}} \leq C(h)$$

# Stopping criteria

**Lemma** *Let (inf-sup) hold. Then*

$$\|\mathbf{r}^{(k)}\|_{A^{-T}HA^{-1}} \leq C_2^{-1}\|\mathbf{r}^{(k)}\|_{H^{-1}}.$$

# Stopping criteria

**Lemma** *Let (inf-sup) hold. Then*

$$\|\mathbf{r}^{(k)}\|_{A^{-T}HA^{-1}} \leq C_2^{-1}\|\mathbf{r}^{(k)}\|_{H^{-1}}.$$

New stopping criterion

$$\|\mathbf{r}^{(k)}\|_{H^{-1}} \leq C_2 C(h)\|\mathbf{u}^{(k)}\|_{H}.$$

# One more crime

**Theorem** *Let* $\tilde{\mathbf{u}}$ *satisfy*

$$\frac{\|\mathbf{f} - A\tilde{\mathbf{u}}\|_{H^{-1}}}{\|\tilde{\mathbf{u}}\|_H} \leq C(h)C_2.$$

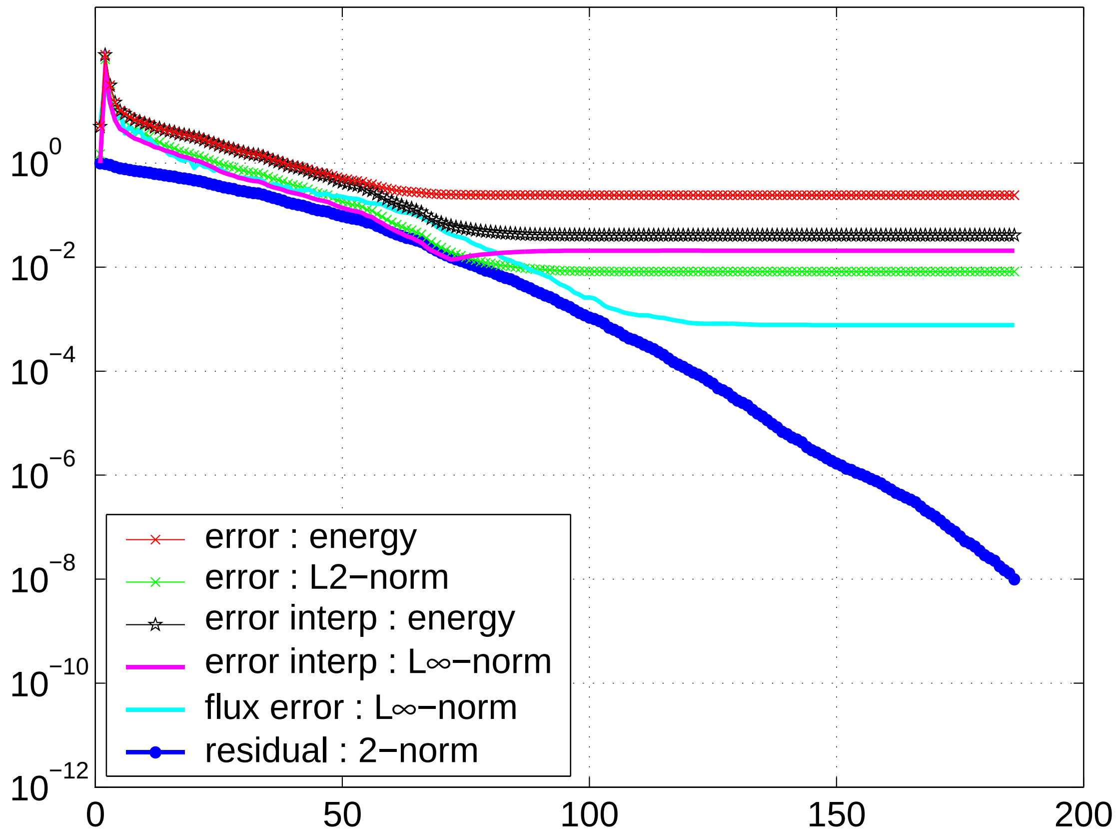
*Then there exists* $\Delta A$ *such that*

$$(A + \Delta A)\tilde{\mathbf{u}} = \mathbf{f}$$

*and*

$$\|\Delta A\|_{H,H^{-1}} \leq C(h)C_2$$

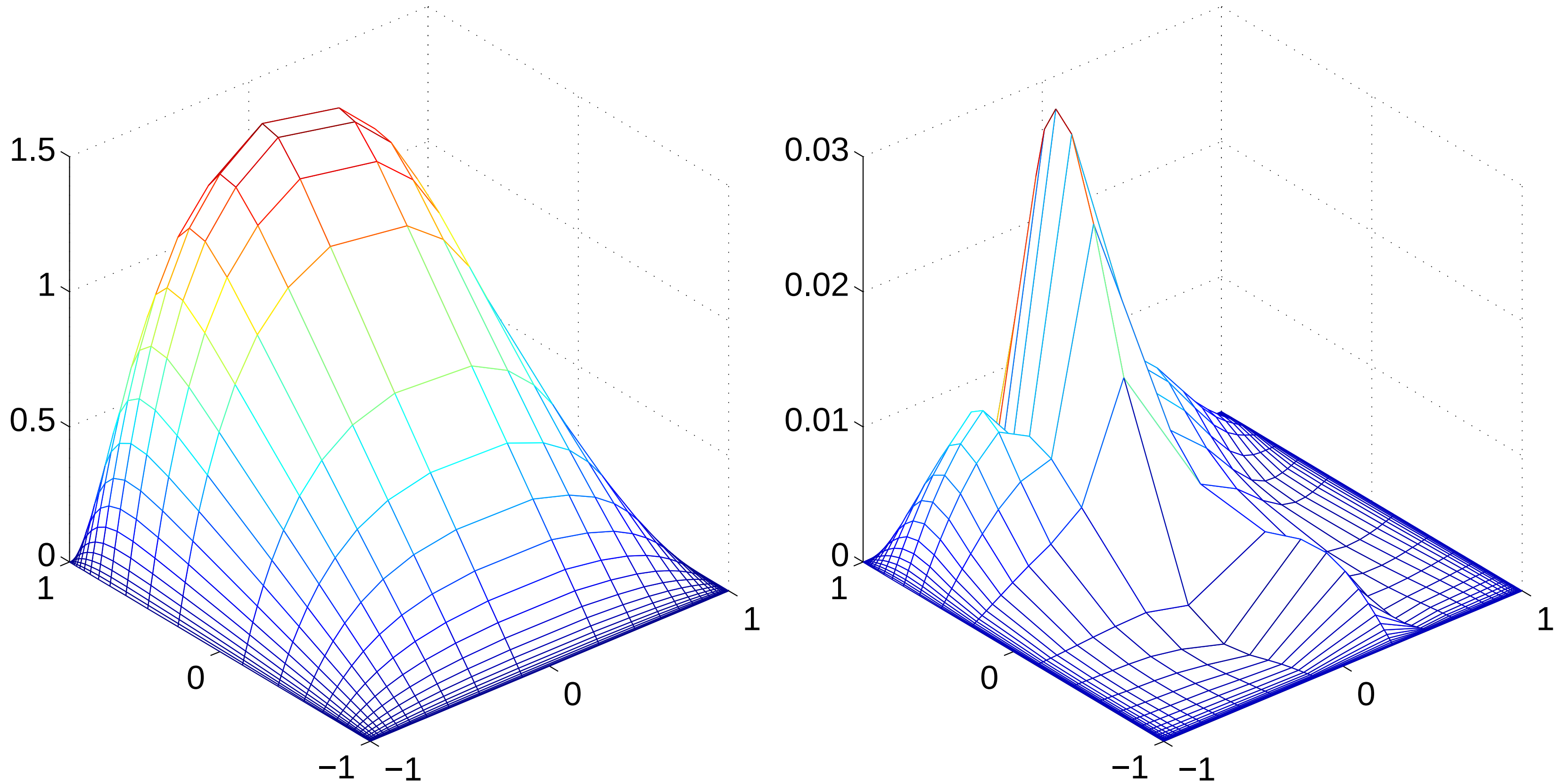
# Convection-diffusion example




$$\frac{\|\nabla(u-u_h^k)\|_{L^2(\Omega)}}{\|\nabla u_h^k\|_{L^2(\Omega)}} \quad \frac{\|u-u_h^k\|_{L^2(\Omega)}}{\|u_h^k\|_{L^2(\Omega)}} \quad \frac{\max_\Omega \left|u^I-u_h^k\right|}{\max_\Omega \left|u_h^k\right|} \quad \frac{\max_\Omega \left|\nabla(u^I-u_h^k)\right|}{\max_\Omega \left|\nabla u_h^k\right|} \quad \frac{\int_{\mathcal{C}} \left|\mathbf{n}\cdot\nabla(u-u_h^k)\right|}{\int_{\mathcal{C}} \left|\mathbf{n}\cdot\nabla u_h^k\right|}$$

# Convection-diffusion example

Solution and interpolation error



# Examples

Elliptic problems in $\mathbb{R}^2$

$$
\begin{aligned}
-\nabla \cdot (\mathbf{a}(\mathbf{x})\nabla u) + \mathbf{b}(\mathbf{x}) \cdot \nabla u + c(\mathbf{x})u &= f \qquad \text{in } \Omega \\
u &= 0 \qquad \text{on } \Gamma.
\end{aligned}
$$

where

$$(\mathbf{a})_{ij},\ (\mathbf{b})_i,\ c \in \mathbf{L}^\infty(\Omega), \quad i, j = 1, 2,$$

$$k_2(\mathbf{x})\,|\boldsymbol{\xi}|^2 \leq \boldsymbol{\xi}^t \mathbf{a}(\mathbf{x})\boldsymbol{\xi} \leq k_1(\mathbf{x})\,|\boldsymbol{\xi}|^2\,,$$

$$c(\mathbf{x}) - \frac{1}{2}\nabla \cdot \mathbf{b}(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \Omega.$$

# Examples

$$a(w,v) = (\mathbf{a}\cdot\nabla w, \nabla v) + (\mathbf{b}\cdot\nabla w, v) + (cw, v),$$

is continuous and coercive with

$$C_1 = \|k_1\|_{L^\infty(\Omega)} + \|\mathbf{b}\|_{L^\infty(\Omega)} + C(\Omega)\|c\|_{L^\infty(\Omega)},$$

$$C_2 = \min_{\mathbf{x}\in\Omega} k_2(\mathbf{x}),$$

wrt $\|\cdot\|_{\mathcal{H}} = |\cdot|_{H_0^1(\Omega)} := |\cdot|_1.$

# Examples

Error estimate:

$$|u - u_h|_1 \leq Ch^{s-1}\|u\|_s, \qquad 1 \leq s \leq 2.$$

# Examples

Error estimate:

$$|u - u_h|_1 \leq C h^{s-1} \|u\|_s, \qquad 1 \leq s \leq 2.$$

Issues

- What is $h$?
- How to approximate $\|u\|_s$?

# Numerical experiments

- Discretization:
  linear elements on uniform & adaptive meshes

# Numerical experiments

- Discretization:
  linear elements on uniform & adaptive meshes
- Estimation of parameters

$$h \sim \frac{\|\mathbf{u}^k\|_M}{\|\mathbf{u}^k\|_{l_2}}, \quad \|u\|_s \sim \|A\mathbf{u}^k\|_{l_2}$$

# Numerical experiments

Stopping criteria and estimates

- Residual dual norm: $\|\mathbf{r}^k\|_{H^{-1}}$

# Numerical experiments

Stopping criteria and estimates

- Residual dual norm: $\|\mathbf{r}^k\|_{H^{-1}}$
- Energy estimate $\|\mathbf{u}^k - \mathbf{u}^{k-1}\|_H \leq C_2 h^2 \|A\mathbf{u}^k\|_{l_2}$

# Advection-diffusion problem

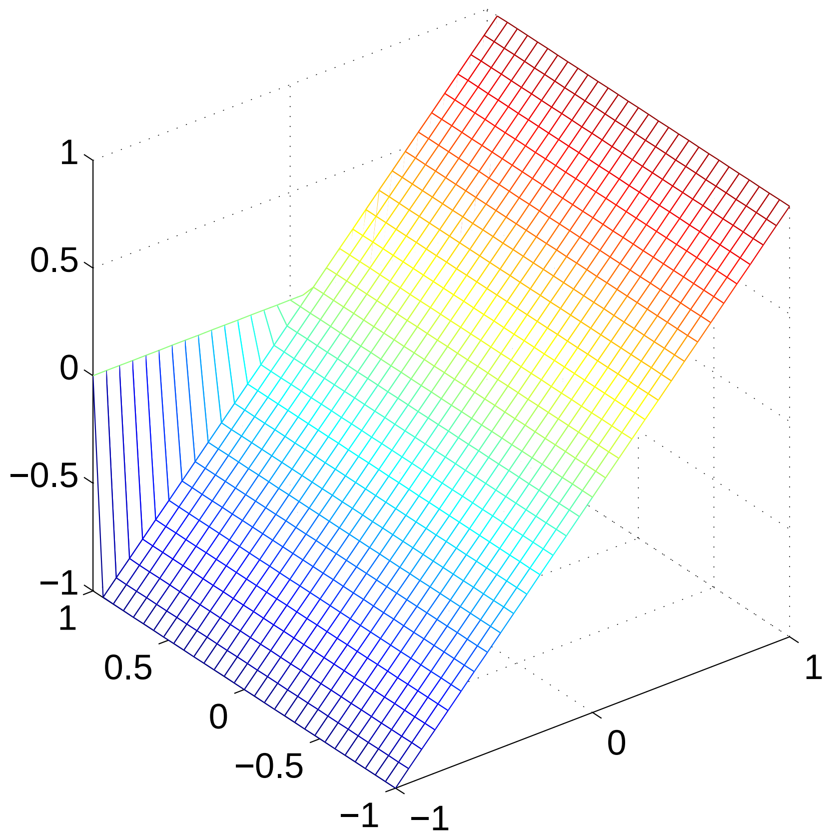
$$\begin{aligned} -\varepsilon \Delta u + \mathbf{b} \cdot \nabla u &= f \qquad \text{in } \Omega \\ u &= g \qquad \text{on } \Gamma. \end{aligned}$$

$$\mathbf{b} = (2y(1-x^2), -2x(1-y^2)),$$

$$u(x,y) = x\left(\frac{1-e^{\frac{y-1}{\varepsilon}}}{1-e^{-\frac{2}{\varepsilon}}}\right),$$

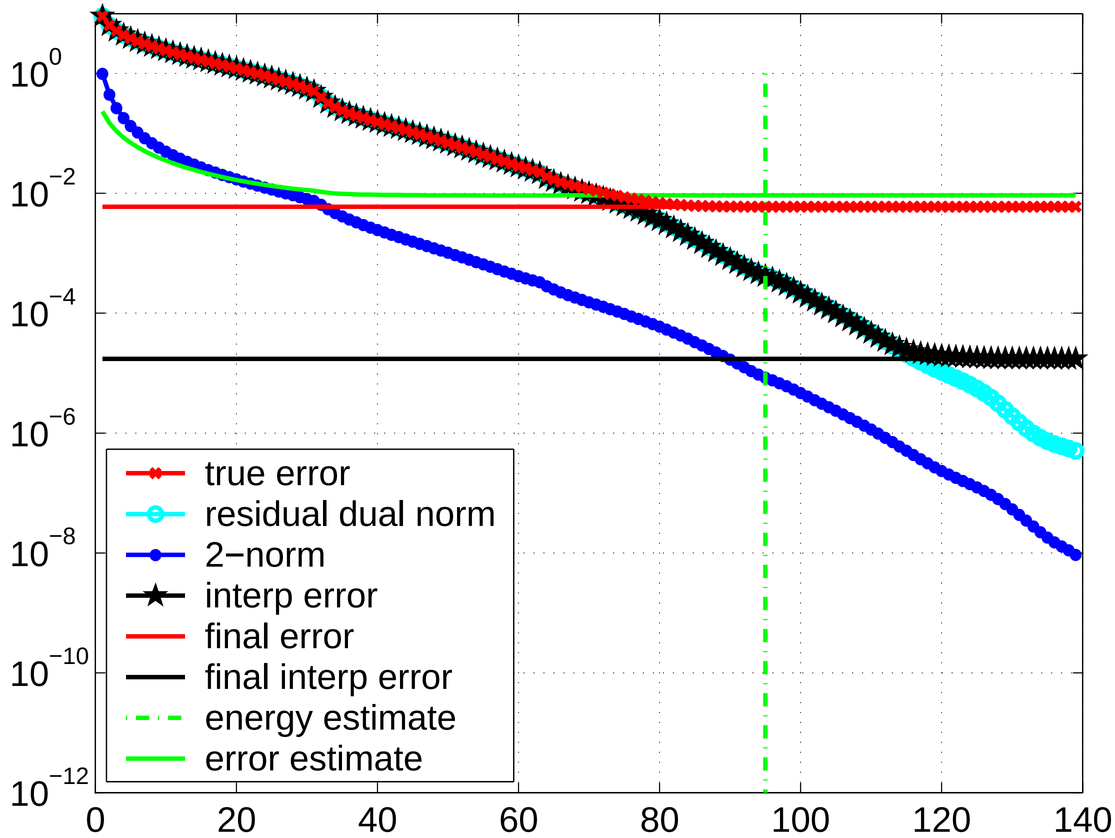
$$\|v_h\|^2_{\mathcal{H}_h} = \varepsilon |v_h|_1^2 + \sum_{T \in \mathcal{T}^h} \delta_T \|\mathbf{b}\cdot\nabla v_h\|^2_{0,T}$$

# Advection-diffusion problem



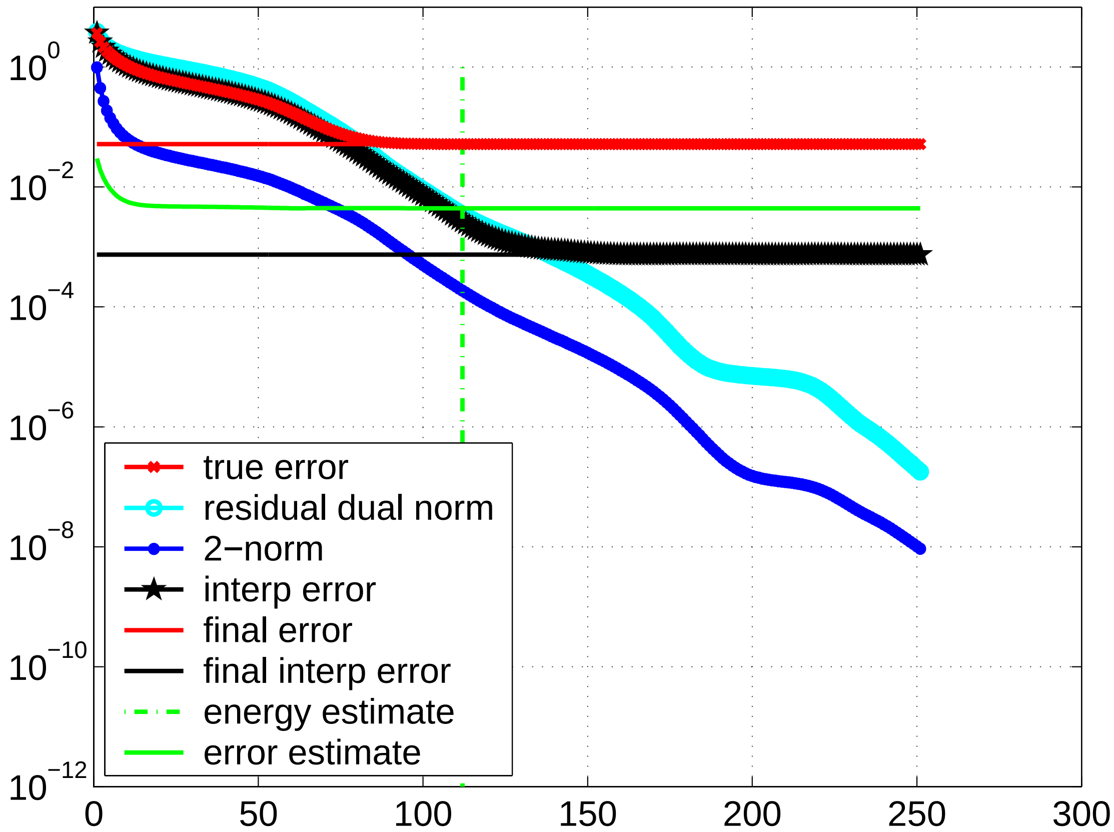

$$\varepsilon = 10^{-2}$$

# Advection-diffusion problem



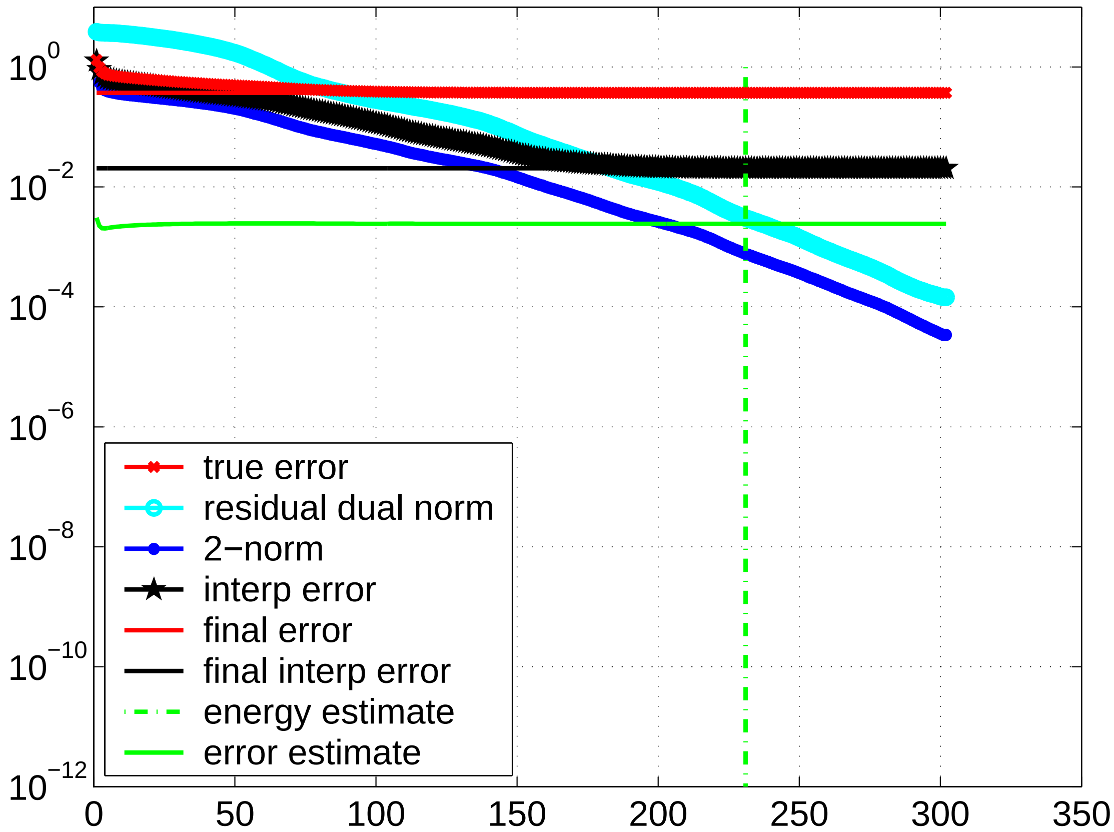

Uniform mesh; $\varepsilon = 1$

# Advection-diffusion problem




Uniform mesh; $\varepsilon = 10^{-1}$

# Advection-diffusion problem




Uniform mesh; $\varepsilon = 10^{-2}$

## Conclusions and future work

- The natural norm for FEM problems is the residual dual norm;
- Estimation is possible;
- For non-symmetric problems, preconditioning with the norm is useful;
- Reliable estimates for finite element error are required;
- Application to indefinite/saddle-point problems;
- Application to non linear problems.