

Preconditioning for self-adjointness

Andy Wathen and Martin Stoll

Computing Laboratory
Oxford University
United Kingdom



A motivating example – **The Bramble-Pasciak CG**: consider saddle point problem

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \quad \text{with preconditioner} \quad \mathcal{P} = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix}$$

The (left) preconditioned matrix

$$\hat{\mathcal{A}} = \mathcal{P}^{-1} \mathcal{A} = \begin{bmatrix} A_0^{-1} A & A_0^{-1} B^T \\ B A_0^{-1} A - B & B A_0^{-1} B^T + C \end{bmatrix}$$

is **not symmetric** but is **self-adjoint** and **positive definite** when

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix}$$

defines an inner product $\langle x, y \rangle_{\mathcal{H}} := x^T \mathcal{H} y$

$\Rightarrow$ CG can be used in this inner product

Conjugate Gradient Method (*Hestenes & Stiefel (1952)*)

for $\hat{\mathcal{A}}x = b$, $\hat{\mathcal{A}}$ self-adjoint and positive definite in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$:

Choose x_0 , compute $r_0 = b - \hat{\mathcal{A}}x_0$, set $p_0 = r_0$
for $k = 0$ until convergence do

$$\alpha_k = \langle r_k, r_k \rangle_{\mathcal{H}} / \langle \hat{\mathcal{A}}p_k, p_k \rangle_{\mathcal{H}}$$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$r_{k+1} = r_k - \alpha_k \hat{\mathcal{A}}p_k$$

<Test for convergence>

$$\beta_k = \langle r_{k+1}, r_{k+1} \rangle_{\mathcal{H}} / \langle r_k, r_k \rangle_{\mathcal{H}}$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

enddo

computes iterates $\{x_k\}$ such that

$$\langle \hat{\mathcal{A}}(x - x_k), x - x_k \rangle_{\mathcal{H}} = \langle x - x_k, x - x_k \rangle_{\mathcal{H}} \hat{\mathcal{A}}$$

is minimal in the relevant Krylov subspace

Self-adjointness: assume

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

is a symmetric bilinear form or an inner product

$\mathcal{A} \in \mathbb{R}^{n \times n}$ is **self-adjoint** in $\langle \cdot, \cdot \rangle$ iff

$$\langle \mathcal{A}x, y \rangle = \langle x, \mathcal{A}y \rangle \quad \text{for all } x, y$$

Self-adjointness of $\mathcal{A}$ in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ($\langle x, y \rangle_{\mathcal{H}} = x^T \mathcal{H}y$)
thus means

$$x^T \mathcal{A}^T \mathcal{H}y = \langle \mathcal{A}x, y \rangle_{\mathcal{H}} = \langle x, \mathcal{A}y \rangle_{\mathcal{H}} = x^T \mathcal{H} \mathcal{A}y$$

for all $x, y \Rightarrow$

$$\mathcal{A}^T \mathcal{H} = \mathcal{H} \mathcal{A}$$

is the relation for **self-adjointness** of $\mathcal{A}$ in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

Basic properties:

LEMMA If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

Basic properties:

LEMMA If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

LEMMA If $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\alpha\mathcal{H}_1 + \beta\mathcal{H}_2}$ for every $\alpha, \beta \in \mathbb{R}$

Basic properties:

LEMMA If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

LEMMA If $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\alpha\mathcal{H}_1 + \beta\mathcal{H}_2}$ for every $\alpha, \beta \in \mathbb{R}$

and of relevance when preconditioning:

LEMMA For symmetric $\mathcal{A}$, $\hat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if $\mathcal{P}^{-T}\mathcal{H}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{A}}$

Basic properties:

LEMMA If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

LEMMA If $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\alpha\mathcal{H}_1 + \beta\mathcal{H}_2}$ for every $\alpha, \beta \in \mathbb{R}$

and of relevance when preconditioning:

LEMMA For symmetric $\mathcal{A}$, $\hat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if $\mathcal{P}^{-T}\mathcal{H}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{A}}$

PROOF

$$(\mathcal{P}^{-T}\mathcal{H})^T \mathcal{A} = \mathcal{H}\mathcal{P}^{-1}\mathcal{A} = (\mathcal{P}^{-1}\mathcal{A})^T \mathcal{H} = \mathcal{A}^T (\mathcal{P}^{-T}\mathcal{H})$$

also combining the above:

LEMMA If $\mathcal{P}_1$ and $\mathcal{P}_2$ are left preconditioners for the symmetric matrix $\mathcal{A}$ for which symmetric matrices $\mathcal{H}_1$ and $\mathcal{H}_2$ exist with $\mathcal{P}_1^{-1}\mathcal{A}$ self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and $\mathcal{P}_2^{-1}\mathcal{A}$ self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ and if for any α, β

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + \beta \mathcal{P}_2^{-T} \mathcal{H}_2 = \mathcal{P}_3^{-T} \mathcal{H}_3$$

for some matrix $\mathcal{P}_3$ and some symmetric matrix $\mathcal{H}_3$ then $\mathcal{P}_3^{-1}\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_3}$.

shows that if we can find such a splitting we have found a new preconditioner and a bilinear form in which the matrix is self-adjoint.

Eigenvalues when $\mathcal{A}$ self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$:

$$\mathcal{A}x = \lambda x, \quad x \neq 0$$

Multiplying from the left by $x^* \mathcal{H}$ gives

$$x^* \mathcal{H} \mathcal{A} x = \lambda x^* \mathcal{H} x$$

and $\mathcal{H} \mathcal{A} = \mathcal{A}^T \mathcal{H} \Rightarrow \lambda \in \mathbb{R}$

Eigenvalues when $\mathcal{A}$ self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$:

$$\mathcal{A}x = \lambda x, \quad x \neq 0$$

Multiplying from the left by $x^* \mathcal{H}$ gives

$$x^* \mathcal{H} \mathcal{A} x = \lambda x^* \mathcal{H} x$$

and $\mathcal{H} \mathcal{A} = \mathcal{A}^T \mathcal{H} \Rightarrow \lambda \in \mathbb{R}$

There is no symmetric bilinear form in which $\mathcal{A}$ is self-adjoint unless $\mathcal{A}$ has real eigenvalues.

LEMMA If $\mathcal{A} = R^{-1}\Lambda R$ is a diagonalization of $\mathcal{A}$ with the diagonal matrix Λ of eigenvalues being real, then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{R^T \Theta R}$ for any real diagonal matrix Θ .

PROOF self-adjointness of $\mathcal{A}$ in $\langle \cdot, \cdot \rangle_{\mathcal{H}} \Rightarrow$

$$R^T \Lambda R^{-T} \mathcal{H} = \mathcal{H} R^{-1} \Lambda R$$

clearly satisfied for $\mathcal{H} = R^T \Theta R$ whenever Θ is diagonal because then Θ and Λ commute.

We remark that this result is not of great use in practice since knowledge of the complete eigensystem of $\mathcal{A}$ is somewhat prohibitive.

Self-adjointness for saddle point systems:

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$$

General preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} X & Y^T \\ Z & W \end{bmatrix}$$

gives

$$\hat{\mathcal{A}} = \mathcal{P}^{-1} \mathcal{A} = \begin{bmatrix} XA + Y^T B & XB^T \\ ZA + WB & ZB^T \end{bmatrix}$$

For self-adjointness of

$$\hat{\mathcal{A}} = \mathcal{P}^{-1} \mathcal{A} = \begin{bmatrix} XA + Y^T B & XB^T \\ ZA + WB & ZB^T \end{bmatrix}$$

in

$$\langle \cdot, \cdot \rangle_{\mathcal{H}}, \quad \mathcal{H} = \begin{bmatrix} E & F^T \\ F & G \end{bmatrix}$$

we require

$$\begin{bmatrix} AX^T + BY^T & AZ^T + B^T W^T \\ BX^T & BZ^T \end{bmatrix} \begin{bmatrix} E & F^T \\ F & G \end{bmatrix} =$$

$$\begin{bmatrix} E & F^T \\ F & G \end{bmatrix} \begin{bmatrix} XA + Y^T B & XB^T \\ ZA + WB & ZB^T \end{bmatrix}$$

Examples: Bramble-Pasciak CG (*Bramble & Pasciak (1988)*)
widely used CG technique with preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & -I \end{bmatrix}$$

and inner product matrix

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix}$$

Examples: Bramble-Pasciak CG (*Bramble & Pasciak (1988)*)
widely used CG technique with preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & -I \end{bmatrix}$$

and inner product matrix

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix}$$

main drawback: requires

$$A_0 < A$$

involves the computation of an eigenvalue problem in the worst case!

Denote: BP^-

Examples: BP with Schur complement preconditioner
(Klawonn (1998), Meyer et al. (2001), Simoncini (2001))

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ S_0^{-1} B A_0^{-1} & -S_0^{-1} \end{bmatrix}$$

Inner product:

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & S_0 \end{bmatrix}$$

Examples: Zulehner (*Zulehner (2001), Schöberl & Zulehner (2007)*)

$$\mathcal{P} = \begin{bmatrix} A_0 & B^T \\ B & BA_0^{-1}B^T - S_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ BA_0^{-1} & I \end{bmatrix} \begin{bmatrix} A_0 & B^T \\ 0 & -S_0 \end{bmatrix}$$

gives $\mathcal{P}^{-1}\mathcal{A}$ self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$,

$$\mathcal{H} = \begin{bmatrix} A_0 - A & 0 \\ 0 & BA_0^{-1}B^T - S_0 \end{bmatrix}$$

if $A_0 > A$ and $S_0 < BA_0^{-1}B^T$

Examples: Benzi-Simoncini (*Benzi and Simoncini (2006)*)
extension of CG method of *Fischer, Ramage, Silvester & W (1998)*

$$\mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

inner product:

$$\mathcal{H} = \begin{bmatrix} A - \gamma I & B^T \\ B & \gamma I \end{bmatrix}$$

Extension for $C \neq 0$ (*Liesen (2006)*, *Liesen & Parlett (2007)*):

$$\mathcal{H} = \begin{bmatrix} A - \gamma I & B^T \\ B & \gamma I - C \end{bmatrix}$$

Example: Bramble-Pasciak⁺ method (*Stoll & W(2007)*)

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & I \end{bmatrix}$$

and inner product

$$\mathcal{H} = \begin{bmatrix} A + A_0 & 0 \\ 0 & I \end{bmatrix}$$

Note: $\mathcal{H}$ defines an inner product for any symmetric and positive definite preconditioner A_0

⇒ can always apply Lanczos, MINRES in this inner product

Example: Bramble-Pasciak⁺ method (*Stoll & W(2007)*)

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & I \end{bmatrix}$$

and inner product

$$\mathcal{H} = \begin{bmatrix} A + A_0 & 0 \\ 0 & I \end{bmatrix}$$

Note: $\mathcal{H}$ defines an inner product for any symmetric and positive definite preconditioner A_0

⇒ can always apply Lanczos, MINRES in this inner product

denote: BP⁺

Combination preconditioning

Final lemma above shows that if can find $\mathcal{P}_3$ and $\mathcal{H}_3$ with

$$\alpha \mathcal{P}_1^{-1} \mathcal{H}_1 + \beta \mathcal{P}_2^{-1} \mathcal{H}_2 = \mathcal{P}_3^{-T} \mathcal{H}_3$$

this gives a new preconditioner and symmetric bilinear form

Combine Bramble-Pasciak and Benzi-Simoncini:

$$\alpha \mathcal{P}_1^{-1} \mathcal{H}_1 + \beta \mathcal{P}_2^{-1} \mathcal{H}_2 =$$

$$\begin{bmatrix} (\alpha A_0^{-1} + \beta I)A - (\alpha + \beta\gamma)I & (\alpha A_0^{-1} + \beta I)B^T \\ -\beta B & -(\alpha + \beta\gamma)I \end{bmatrix}$$

One possibility for a splitting $\alpha \mathcal{P}_1^{-1} \mathcal{H}_1 + \beta \mathcal{P}_2^{-1} \mathcal{H}_2$ is

$$\mathcal{P}_3^{-T} = \begin{bmatrix} \alpha A_0^{-1} + \beta I & 0 \\ 0 & -\beta I \end{bmatrix}$$

and

$$\mathcal{H}_3 = \begin{bmatrix} A - (\alpha + \beta\gamma)(\alpha A_0^{-1} + \beta I)^{-1} & B^T \\ B & \frac{\alpha + \beta\gamma}{\beta} I \end{bmatrix}$$

Combination preconditioning

Final lemma above shows that if can find $\mathcal{P}_3$ and $\mathcal{H}_3$ with

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + \beta \mathcal{P}_2^{-T} \mathcal{H}_2 = \mathcal{P}_3^{-T} \mathcal{H}_3$$

this gives a new preconditioner and symmetric bilinear form

Combination preconditioning

Final lemma above shows that if can find $\mathcal{P}_3$ and $\mathcal{H}_3$ with

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + \beta \mathcal{P}_2^{-T} \mathcal{H}_2 = \mathcal{P}_3^{-T} \mathcal{H}_3$$

this gives a new preconditioner and symmetric bilinear form

Take $\mathcal{P}_1, \mathcal{H}_1$ as BP⁻ and $\mathcal{P}_2, \mathcal{H}_2$ as BP⁺

Combine BP^- and BP^+

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + (1 - \alpha) \mathcal{P}_2^{-T} \mathcal{H}_2 =$$

$$\begin{bmatrix} A_0^{-1} A + (1 - 2\alpha)I & A_0^{-1} B^T \\ 0 & (1 - 2\alpha)I \end{bmatrix}$$

Combine BP^- and BP^+

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + (1 - \alpha) \mathcal{P}_2^{-T} \mathcal{H}_2 =$$
$$\begin{bmatrix} A_0^{-1} A + (1 - 2\alpha) I & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}$$

can be split as

$$\mathcal{P}_3^{-T} = \begin{bmatrix} A_0^{-1} & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}, \mathcal{H}_3 = \begin{bmatrix} A + (1 - 2\alpha) A_0 & 0 \\ 0 & I \end{bmatrix}$$

Combine BP^- and BP^+

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + (1 - \alpha) \mathcal{P}_2^{-T} \mathcal{H}_2 =$$

$$\begin{bmatrix} A_0^{-1} A + (1 - 2\alpha) I & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}$$

can be split as

$$\mathcal{P}_3^{-T} = \begin{bmatrix} A_0^{-1} & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}, \mathcal{H}_3 = \begin{bmatrix} A + (1 - 2\alpha) A_0 & 0 \\ 0 & I \end{bmatrix}$$

Recall $\mathcal{P}_3^{-1} \mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_3}$

Combine BP^- and BP^+

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + (1 - \alpha) \mathcal{P}_2^{-T} \mathcal{H}_2 =$$

$$\begin{bmatrix} A_0^{-1} A + (1 - 2\alpha) I & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}$$

can be split as

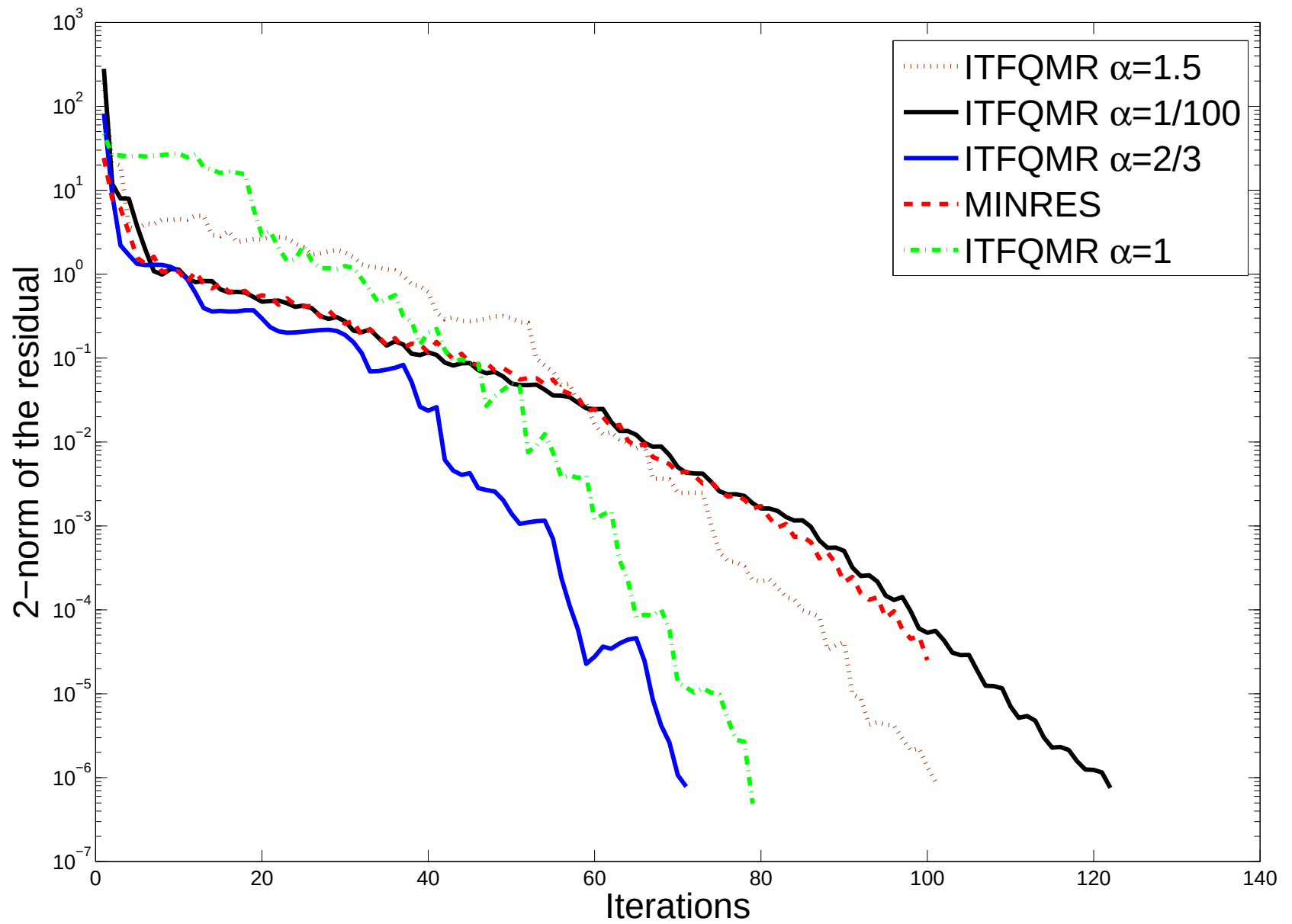
$$\mathcal{P}_3^{-T} = \begin{bmatrix} A_0^{-1} & A_0^{-1} B^T \\ 0 & (1 - 2\alpha) I \end{bmatrix}, \mathcal{H}_3 = \begin{bmatrix} A + (1 - 2\alpha) A_0 & 0 \\ 0 & I \end{bmatrix}$$

Recall $\mathcal{P}_3^{-1} \mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_3}$

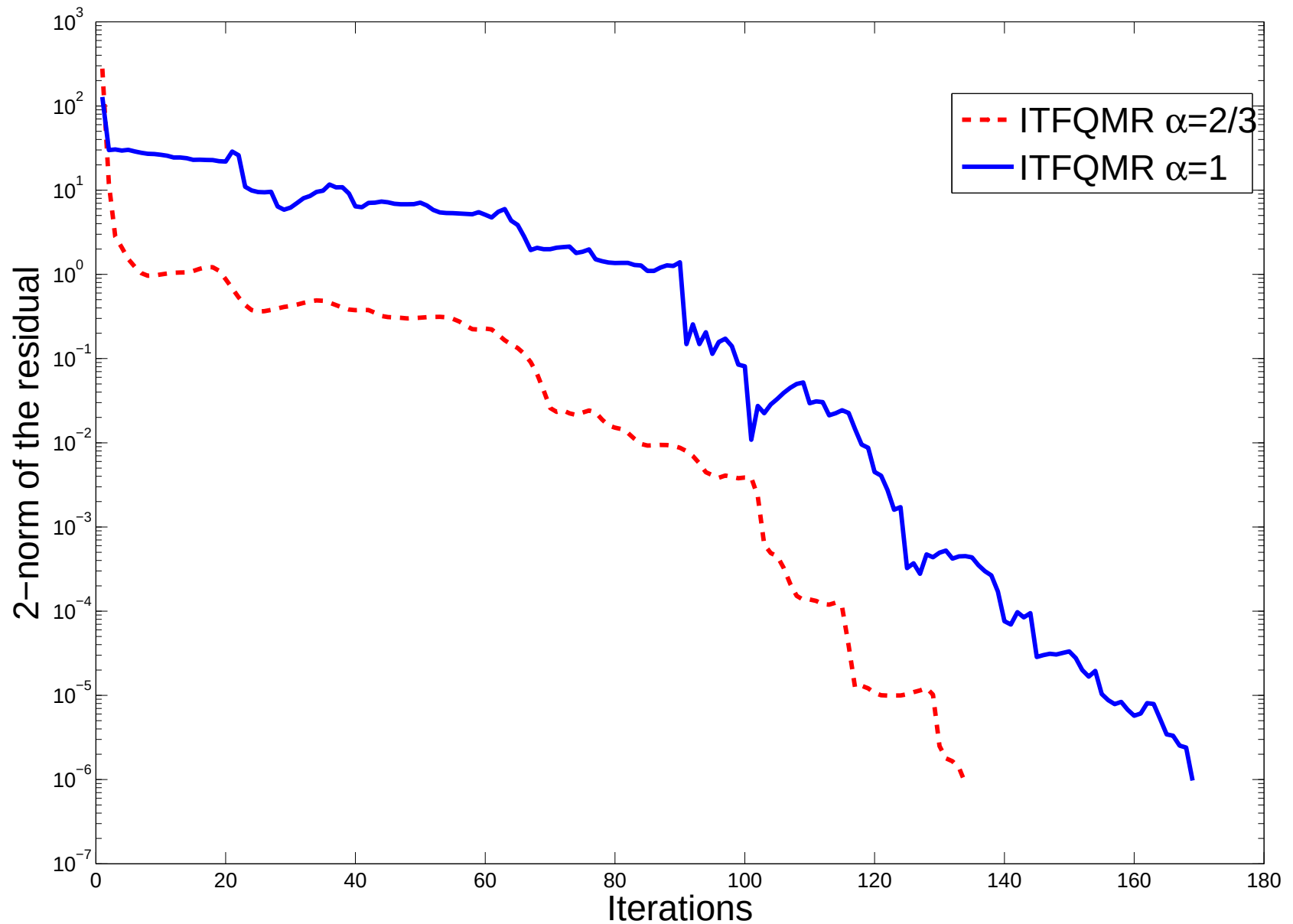
$(\alpha = 1 \leftrightarrow \text{BP}^-, \quad \alpha = 0 \leftrightarrow \text{BP}^+)$

Iterative method?

- MINRES (applicable when $\mathcal{H}_3$ is positive definite)
(*Paige & Saunders (1975)*)
- ITFQMR — ideal transpose-free Quasi Minimum Residuals
(*Freund & Nachtigal (1995)*) (applicable for symmetric nonsingular $\mathcal{H}_3$ and related to BiCG: *Rozloznik (2005)*)



Combine BP- and BP+: another IFISS Stokes problem



Summary

- iterative methods more reliable/descriptive convergence theory/know what preconditioning is trying to achieve when matrix is symmetric or self-adjoint

Summary

- iterative methods more reliable/descriptive convergence theory/know what preconditioning is trying to achieve when matrix is **symmetric or self-adjoint**
- ‘isolated’ practical examples of self-adjointness in non-standard inner products exist - in particular BP^- and the new BP^+ method

Summary

- iterative methods more reliable/descriptive convergence theory/know what preconditioning is trying to achieve when matrix is **symmetric or self-adjoint**
- ‘isolated’ practical examples of self-adjointness in non-standard inner products exist - in particular BP^- and the new BP^+ method
- theory here allow ‘interpolation’ between such examples, hence broadens the set of possible application of CG or MINRES in non-standard inner products

Summary

- iterative methods more reliable/descriptive convergence theory/know what preconditioning is trying to achieve when matrix is **symmetric or self-adjoint**
- ‘isolated’ practical examples of self-adjointness in non-standard inner products exist - in particular BP^- and the new BP^+ method
- theory here allow ‘interpolation’ between such examples, hence broadens the set of possible application of CG or MINRES in non-standard inner products
- application here to saddle-point matrices, but theory is more general