Report no. 10/07

On choice of preconditioner for minimum residual methods for nonsymmetric matrices

Jennifer Pestana Andrew J. Wathen

Existing convergence bounds for Krylov subspace methods such as GMRES for real nonsymmetric linear systems give little mathematical guidance for the choice of preconditioner. Here, we establish a desirable mathematical property of a preconditioner which indicates when convergence of a minimum residual method will essentially depend only on the eigenvalues of the preconditioned system, as is true in the symmetric case. Our theory covers the generic case of nonsymmetric coefficient matrices which are diagonalisable over $\mathbb{C}$; it does not cover matrices with nontrivial Jordan form.

Oxford University Mathematical Institute
Numerical Analysis Group
24-29 St Giles'
Oxford, England OX1 3LB
E-mail: pestana@maths.ox.ac.uk

May, 2011

1 Introduction

For linear systems of equations with a real symmetric (or complex hermitian) coefficient matrix, there are standard iterative methods of choice: for positive definite problems the Conjugate Gradient method [26] and for indefinite problems the MINRES method [39]. (The SYMMLQ method [39] is a lesser used alternative in the indefinite case.) For these methods there exist descriptive convergence theories based on eigenvalues, i.e., convergence bounds given in terms of polynomial approximation problems on the eigenvalue spectrum which are often reasonably tight [23, page 51]. A key consequence is that it is clear what one is trying to achieve to get fast convergence via preconditioning in these cases, namely well distributed eigenvalues or reduced spectral condition number.

By contrast for nonsymmetric matrices, there are many possible and competing methods. GMRES [41] is probably the most popular minimum residual method, but unless the required number of iterations is kept small, the increased work of orthogonalisation in the Arnoldi process and the storage requirements can both grow to infeasible levels. To avoid this increased work and storage, several methods based on the nonsymmetric Lanczos method are also widely employed [20, 21, 45, 46, 49], though none minimises the residual and we do not consider them here; the convergence of these methods is even less well understood than of minimum residual methods such as GMRES. For any and all of these methods, preconditioning is almost always employed in the hope of obtaining acceptably rapid convergence.

Given the importance of iterative solution methods for large and sparse or structured linear or linearised systems of equations, it remains a fundamental and largely open mathematical question as to what one is trying to achieve with preconditioning for nonsymmetric matrices: the glib answer 'fast convergence' gives no idea nor criteria for making other than heuristically motivated choices. The known results which inform the user are largely negative: Faber and Manteuffel [19] famously proved that, except for trivial cases with few distinct eigenvalues, the set of matrices for which there exist short-term recurrence methods which minimize some relevant quantity is rather small—precisely those diagonalisable matrices which have all of their eigenvalues on a straight line in the complex plane[1] (see [18] for a more recent proof), Nachtigal, Reddy and Trefethen [38] considered three of the competing nonsymmetric iterative methods available in 1992 and identified classes of real $n \times n$ matrices for which each of the methods could outperform the other two methods by a factor of a least $\sqrt{n}$ and Greenbaum, Pták and Strakoš [24] proved that given any complete set of eigenvalues and any monotonically reducing convergence curve that it is possible to construct a matrix with the given eigenvalues and an initial residual (starting iterate) such that GMRES will achieve precisely the given convergence curve[2]. Thus one can certainly achieve stagnation until the n^{th}

[1] With a different recurrence than that examined by Faber and Manteuffel it is possible to obtain a minimum residual method with short-term recurrences for matrices with concyclic eigenvalues [28, 29], although the set of matrices possessing this property is also small. For a summary of short-term recurrence methods, see [34].

[2] The Greenbaum, Pták and Strakoš result built on work in [25] and was followed by a characterisation in [1] of the set of matrices, with prescribed eigenvalues, and right-hand sides for which a given

iterate even for a matrix all of whose eigenvalues are 1 (see [25] for an example)! This is in stark contrast to the situation for real symmetric matrices.

As is well known ([23, §2.4],[16, §4.1.1]) for the linear system $Bx = b$, with $B \in \mathbb{R}^{n\times n}$ and $x, b \in \mathbb{R}^n$, a minimum residual method such as GMRES computes iterate vectors $\{x_k\}, k = 1, 2, \ldots$ from a starting guess x_0 such that $x_k - x_0$ lie in the Krylov subspaces

$$\mathcal{K}_k(B, r_0) = \text{span}\{r_0, Br_0, B^2r_0, \ldots, B^{k-1}r_0\}, \tag{1.1}$$

so that the residuals $r_k = b - Bx_k$ satisfy $r_k = p(B)r_0$ where $p \in \Pi_k$, the set of real polynomials of degree k or less, and p satisfies the consistency condition $p(0) = 1$. Assuming that $B \in \mathbb{R}^{n\times n}$ is diagonalisable,

$$B = Z\Lambda Z^{-1},$$

where Λ is a diagonal matrix of eigenvalues and Z has corresponding eigenvectors as its columns, the minimum residual property ensures that

$$\frac{\|r_k\|_2}{\|r_0\|_2} \le \|Z\|_2\|Z^{-1}\|_2 \min_{p\in\Pi_k, p(0)=1} \max_{\lambda\in\sigma(B)} |p(\lambda)| \tag{1.2}$$

where $\sigma(B)$ is the eigenvalue spectrum of B. It is clear how this allows the Greenbaum, Pták and Strakoš construction in the nonsymmetric case since even for a diagonalisable matrix B with n distinct eigenvalues arbitrarily close to 1, the smallness of the polynomial term

$$\min_{p\in\Pi_k, p(0)=1} \max_{\lambda\in\sigma(B)} |p(\lambda)| \tag{1.3}$$

in (1.2) for $k \ge 1$ can still allow for arbitrarily slow convergence for a large enough value of

$$\kappa_2(Z) = \|Z\|_2\|Z^{-1}\|_2,$$

the condition number of the eigenvector matrix. This is clearly not true in the symmetric case when Z is necessarily orthogonal and thus $\|Z\|_2 = \|Z^{-1}\|_2 = 1$, so that it is essentially the eigenvalues which determine convergence in that case.

In this paper we present some theory which aims at providing some sufficient mathematical conditions for selection of preconditioners in the nonsymmetric case. Specifically for real matrices that are diagonalisable over the complex field we identify a mathematical condition on a preconditioner which will indicate when the convergence of a minimum residual method (such as GMRES) will essentially be determined by the eigenvalues alone, without the possible huge term coming from the eigenvectors in (1.2). We comment that other bounds for GMRES convergence are available, notably based on the field of values [4, 12, 13, 14] and pseudospectra [48, Chapter 26] and improvement of these bounds might also follow from the analysis here.

Let us illustrate the point of this paper by reference to a well known example, the 5005×5005 sherman3 matrix together with its provided right hand side vector from

convergence curve is achieved.

Name	$\kappa_2(B)$	$\kappa_2(P^{-1}B)$	$\kappa_2(Z_B)$	$\kappa_2(Z_{P^{-1}B})$
sherman3	1.2×10^{19}	2.3×10^3	1.002	1.2×10^3

Table 1: Spectral norm condition numbers of the sherman3 matrix B and the ILU(0) preconditioned system $P^{-1}B$, and of the matrix of eigenvectors of each matrix (denoted by Z).

the University of Florida Sparse Matrix Collection [10]. This long standing example is certainly well within the capabilities of excellent sparse direct methods at this point [9, 11], but it is a useful (and often quoted) example in the context of iterative methods [6, 7, 30, 32, 50]. We employ in this example the well known ILU(0) preconditioning technique described by Meijerink and van der Vorst [36]. Figure 1 shows the commonly seen behaviour with such examples: without any preconditioning, the solid line shows the reduction in $\|r_k\|_2$ for increasing numbers of GMRES iterations k. For this non-symmetric example, the eigenvalues of B are in fact all real and so it is possible to approximate/bound accurately the term (1.3) by employing Chebyshev polynomials as is done in the case of symmetric matrices (see for example [16, §§2.1.1, 6.2.4]). Using such an approach, the convergence bound (1.2) in this case is given by the dashed line which is seen to be nearly horizontal. This is seen not to be such a bad upper bound for the early iterations when little progress is made by GMRES in reducing the Euclidean norm of the residual. The dot-dashed line shows the reduction in the preconditioned residuals (here plotted also in the Euclidean norm) achieved when ILU(0) preconditioning is used in conjunction with GMRES. This rapid convergence is however far from what is described by the bound (1.2) which is indicated by the dotted line on the figure and is computed similarly to the above. The eigenvectors of the original matrix are in fact quite well conditioned for this example (hence the intercept for the dashed line is near 1), however the poor eigenvalue distribution leads to slow convergence without preconditioning. The preconditioner improves the eigenvalue distribution but in fact makes the eigenvector condition number much worse as seen in the large intercept of the dotted line (and in the numbers in Table 1 where the relevant data for this problem are given). The dotted line is nowhere descriptive of the convergence as displayed in the dot-dashed line.

The main results of this paper show why the eigenvector condition number term in (1.2) can essentially be ignored for this example. In §2 we introduce the notion of self-adjointness with respect to a non-standard inner product and show how the application of minimum residual methods in this inner product results in convergence dependent only on eigenvalues. This leads in §3 to the consideration of the relationship between minimum residual methods in different inner products and norms. Our main results regarding preconditioning are given in §4 and the results of numerical tests are described in §5.

Throughout this manuscript the Parlett convention is adopted, excepting the absolutely standard use of H_k and V_k in the Arnoldi process: any matrix identified by a character with reflectional symmetry will be symmetric, otherwise nonsymmetric. Thus, $A, H, M, \cdots$ are symmetric whereas $B, C, D, \cdots$ are nonsymmetric.

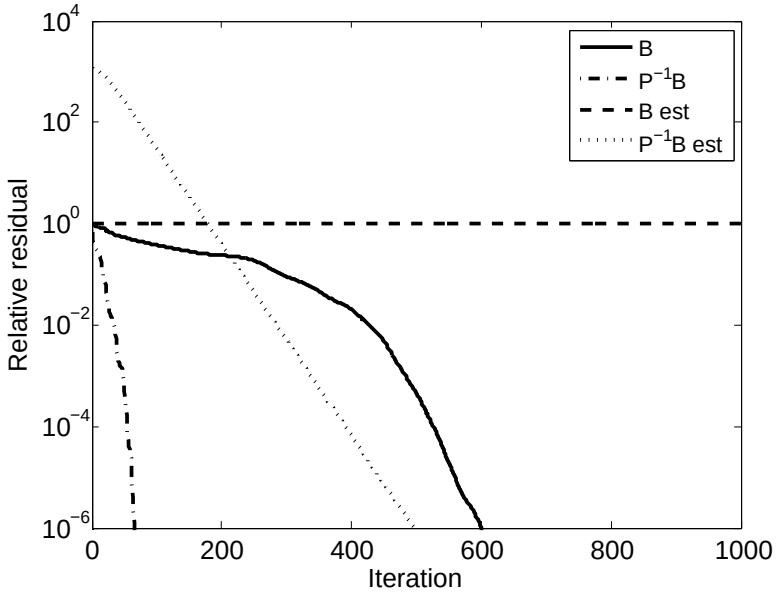


Figure 1: Convergence curves and bounds for ILU(0) preconditioned GMRES applied to sherman3.

2 Self-adjointness in nonstandard inner products

Any nonderogatory hermitian bilinear form

$$\langle x, y\rangle_H \tag{2.1}$$

on $\mathbb{C}^n$ is equivalent to x^*Hy for some invertible hermitian matrix $H \in \mathbb{C}^{n\times n}$ [22]. In the case that H is hermitian positive definite, (2.1) is an inner product and we can associate with H the vector norm

$$\|x\|_H = \sqrt{\langle x, x\rangle_H}$$

and induced matrix norm

$$||B||_H = \max_{x\neq 0} \frac{\|Bx\|_H}{\|x\|_H}.$$

The most familiar inner product is the Euclidean inner product, here denoted the I-inner product for consistency, with its induced spectral norm $\|\cdot\|_I$. Any H-matrix-norm can be related to the I-norm since, for any factorisation $H = S^*S$ of H with S invertible,

$$\|B\|_H = \|SBS^{-1}\|_I. \tag{2.2}$$

We call Z a H-unitary matrix—or a H-orthogonal matrix if H is real— if

$$Z^*HZ = I \tag{2.3}$$

and note that this differs from the definition in, for example, [22]. A matrix B is self-adjoint with respect to some symmetric bilinear form $\langle\cdot,\cdot\rangle_H$ if $\langle Bx, y\rangle_H = \langle x, By\rangle_H$ for all $x, y \in \mathbb{C}^n$ or, equivalently, if

$$HB = B^*H, \tag{2.4}$$

i.e., HB is hermitian. Moreover, every real matrix is self-adjoint with respect to some real symmetric bilinear form [22, Corollary 6.1.4]. For example, it is not difficult to show that the Jordan block $J(\lambda) \in \mathbb{R}^{n\times n}$,

$$J(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix},$$

is self-adjoint with respect to indefinite hermitian bilinear forms induced by Hankel matrices of the form

$$H = \begin{bmatrix} & & & a_1 \\ & & \iddots & a_2 \\ & \iddots & \iddots & \vdots \\ a_1 & a_2 & \dots & a_n \end{bmatrix},$$

where $a_1, \dots, a_n$ are any real scalars.

Indeed, if B is self-adjoint with respect to some hermitian bilinear form $\langle \cdot, \cdot \rangle_H$, the building blocks of H are the hermitian bilinear forms of Jordan blocks. A detailed discussion of this topic can be found in [22] but we summarise Theorem 5.1 from [22] here, following the authors' presentation. Let B have Jordan normal form $B = ZJZ^{-1}$ where Z and J may be complex. The self-adjointness of B guarantees that for every complex eigenvalue λ_k of B there exists a Jordan block of the same dimension corresponding to $\overline{\lambda}_k$ [22, Proposition 4.2.3]. Consequently, without loss of generality we assume that

$$J = \operatorname{diag}(J(\lambda_1), \dots J(\lambda_\alpha), J(\lambda_{\alpha+1}), \cdots, J(\lambda_\beta)),$$

where $\lambda_1, \dots, \lambda_\alpha$ are the real eigenvalues of B, $\lambda_{\alpha+1}, \dots, \lambda_\beta$ are the nonreal eigenvalues in the upper half-plane and $J(\lambda_k)$, $k = \alpha+1, \dots, \beta$ is a direct sum of the Jordan blocks corresponding to λ_k and $\overline{\lambda}_k$. Then H has the form

$$H = Z^{-*} Y_{\epsilon,J} Z^{-1}, \tag{2.5}$$

where

$$Y_{\epsilon_J} = \operatorname{diag}(\epsilon_1 Y_1, \dots, \epsilon_\alpha Y_\alpha, Y_{\alpha+1}, \dots, Y_\beta),$$

$\epsilon_i = \pm 1$, and each Y_i is of the same dimension as the corresponding Jordan block J_i and is of the form

$$Y_i = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}.$$

The well-known result that a real matrix B is self-adjoint with respect to an inner product if and only if it is diagonalisable with real eigenvalues is clear from this factorisation of H.

If B is diagonalisable there exist minimum residual methods in $\langle \cdot, \cdot \rangle_M$,

$$M = Z^{-*} Z^{-1}, \tag{2.6}$$

which we call M-MR methods, for which convergence depends chiefly on the eigenvalues of B. These methods have been thoroughly explored in [2, 3, 31, 47] and the references therein, although generally only in the case that B is self-adjoint with respect to an inner product. (For situations where Conjugate Gradient methods in a nonstandard inner product are applicable see [33].) This property of M-MR methods is a generalisation of the well known result for symmetric matrices and Euclidean norm minimum residual methods, or I-MR methods. It is illustrated by the following convergence bound which previously appeared, in the case that H defines an inner product, in [5]. Although we consider here only real coefficient matrices our theory is directly applicable to complex matrices; its extension to nondiagonalisable matrices is also possible but involves a polynomial in the Jordan matrix J.

Lemma 1 *Let $B = Z \Lambda Z^{-1}$ be a diagonalisation of B and let $M = Z^{-*} Z^{-1}$. Then r_k^M, the k^{th} residual vector of M-MR, satisfies*

$$\frac{\|r_k^M\|_M}{\|r_0\|_M} \leq \min_{p \in \Pi_k,\, p(0)=1} \max_{\lambda \in \sigma(B)} |p(\lambda)|. \tag{2.7}$$

Proof By definition, the M-MR residual vector is given by $r_k^M = p_k(B) r_0$, where $p_k(z)$ is the unique polynomial of degree less than or equal to k with $p(0) = 1$ that minimises $\|r_k^M\|_M$, and r_0 is the initial residual. Using (2.2), it follows that

$$\|r_k^M\|_M = \|p_k(B) r_0\|_M \leq \|Z^{-1} Z p_k(\Lambda) Z^{-1} Z\|_I \|r_0\|_M$$

from which (2.7) immediately follows. ∎

Remarks

1. M-MR methods must be applied in order that the resulting residual minimises the M-norm in the above theorem.

2. If B is real, diagonalisable and has real eigenvalues, $M = H$.

3. Lemma 1 (and all ensuing results) hold if the (possibly complex) diagonalisation of B is replaced by its real Jordan form.

4. It is possible that a similar result can be obtained for H-normal matrices, i.e., matrices satisfying $HB^*H^{-1}B = BHB^*H^{-1}$. However, this condition is more complicated than (2.4) and the above holds for any diagonalisable matrix.

5. The right-hand side vector may affect convergence significantly, as is true even of symmetric matrices (see [37, page 492] and the references therein). This effect most likely depends on the inner product used, although investigation of this aspect is beyond the scope of this manuscript. We expect, however, that the behaviour of M-MR methods applied to H-self-adjoint matrices will typically be determined by eigenvalues.

6. If B is real and diagonalisable but has complex conjugate pairs of eigenvalues, Z is complex, with complex conjugate columns of eigenvectors. Regardless, $H = Z^{-*}YZ^{-1}$ and $M = Z^{-*}Z^{-1}$ are real. This can be seen by expanding $H^{-1} = (ZY)Z^*$ (or $M^{-1} = ZZ^*$) as a sum of outer products of eigenvectors and by noting that each complex outer product uv^* in the sum is added to $\overline{uv^*}$.

7. We do not advocate solving linear systems with M-MR methods in favour of standard methods in general. Instead, this framework is adopted because it provides insight into desirable preconditioners for standard MR methods, which we discuss in the following sections. However, in the serendipitous case that M is known and inner products with M are cheap to apply M-MR can be applied using only real arithmetic.

Lemma 1 differs from eigenvalue-based convergence bounds for minimum residual methods in any other inner product. If, for example, $W = F^*F$ and W-MR is employed then by the same argument used in Lemma 1 we obtain the bound

$$\frac{\|r_k^W\|_W}{\|r_0\|_W} \leq \kappa_2(FZ) \min_{p\in\Pi_k\ p(0)=1} \max_{\lambda\in\sigma(B)} |p(\lambda)|. \tag{2.8}$$

For I-MR methods (for which $F = W = I$) this leads to the standard bound (1.2) that includes the condition number of the eigenvectors. If the eigenvector condition number is large, (1.2) may be a poor estimate of convergence.

3 Relationship between the convergence of minimum residual methods in different inner products

In the previous section we showed that diagonalisable nonsymmetric H-self-adjoint matrices enjoy M-MR convergence that essentially depends on eigenvalues. A natural consideration is whether this knowledge provides insight into the convergence of minimum residual methods in the Euclidean, or indeed any other, inner product. Progress in this direction has been made by Essai [17], who relates the convergence of GMRES (and FOM) in $\langle\cdot,\cdot\rangle_D$ to that of GMRES (and FOM) in $\langle\cdot,\cdot\rangle_I$, where D is a diagonal matrix related to the initial residual. Similar comparisons of iterative methods have been made by [2, 13, 42, 43].

Here we extend the results of Essai to address the broader question of how the convergence of two minimum residual methods, M-MR and W-MR say, in the respective inner products $\langle\cdot,\cdot\rangle_M$ and $\langle\cdot,\cdot\rangle_W$, compare. We restrict our attention to MR methods for which the basis vectors are orthogonal with respect to the associated inner product in the sense of (2.3). This still encompasses the best known implementations of minimum residual methods, MINRES and GMRES. To simplify the exposition the initial iterate is taken to be the zero vector; this loses no generality. We consider only real M and W, since these define the only inner products required in later sections, although extension to the complex case is trivial.

Let the columns of V_k^M be the M-orthogonal basis vectors of the k^{th} Krylov subspace generated by M-MR. Then the standard equation (see, for example, [23, page 38])

$$BV_k^M = V_{k+1}^M H_k^M \tag{3.1}$$

holds, where H_k^M is a $(k+1) \times k$ upper Hessenberg matrix. Similarly, if V_k^W has as its columns the W-orthogonal basis vectors of W-MR then

$$BV_k^W = V_{k+1}^W H_k^W. \tag{3.2}$$

Since V_k^W is W-orthogonal the standard relationship for I-MR methods [23, page 39] can be trivially extended to this case to give:

$$\|r_k^W\|_W = \min_{y \in \mathbb{R}^k} \| \|r_0\|_W e_1 - H_k^W y\|_2. \tag{3.3}$$

Analogously, the M-norm of the M-MR residual satisfies

$$\|r_k^M\|_M = \min_{z \in \mathbb{R}^k} \| \|r_0\|_M e_1 - H_k^M z\|_2. \tag{3.4}$$

The crucial observation is that M-MR and W-MR form bases of the same nested Krylov subspaces. Consequently, assuming that these methods do not terminate at or before the k^{th} step,

$$V_k^W = V_k^M R_k, \tag{3.5}$$

where $R_k \in \mathbb{R}^{k \times k}$ is upper triangular and nonsingular; this is a QR decomposition of V_k^W in the $\langle \cdot, \cdot \rangle_M$ inner product. From (3.1), (3.2) and (3.5) we obtain the following (see [17, Proposition 1] for details):

$$R_k = (V_k^M)^T M V_k^W, \quad R_k^{-1} = (V_k^W)^T W V_k^M, \quad H_k^W = R_{k+1}^{-1} H_k^M R_k. \tag{3.6}$$

Equations (3.5) and (3.6) allow the expression of W-MR residuals in terms of quantities associated with M-MR:

Lemma 2 *Let r_k^W be the k^{th} residual of W-MR. Then*

$$\|r_k^W\|_W = \min_{\bar{y} \in \mathbb{R}^k} \| \|r_0\|_M e_1 - H_k^M \bar{y} \|_{R_{k+1}^{-T} R_{k+1}^{-1}}, \tag{3.7}$$

where H_k^M is the upper Hessenberg matrix in (3.1) *and R_{k+1} is the upper triangular matrix in* (3.5).

Proof See [17, §4]. ∎

Comparison with the M-MR residual (3.4) shows that the quantities minimised by W-MR and M-MR are the same but the norms with respect to which they are minimised differ. This is the key relationship between W-MR methods and M-MR methods and it allows us to link M-MR and W-MR convergence in the central theorem of this section.

Theorem 3 *The k^{th} W-MR residual r_k^W and the k^{th} M-MR residual r_k^M satisfy*

$$\sqrt{\lambda_{min}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M \leq \|r_k^W\|_W \leq \sqrt{\lambda_{max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M, \tag{3.8}$$

where R_{k+1} is defined in (3.5).

Proof Let

$$z_k = \arg\min_z \|\|r_0\|_M e_1 - H_k^M z\|_2.$$

Then, from (3.7),

$$\begin{aligned}\|r_k^W\|_W &= \min_{\bar{y}\in\mathbb{R}^k} \|\|r_0\|_M e_1 - H_k^M \bar{y}\|_{R_{k+1}^{-T}R_{k+1}^{-1}} \\ &\leq \|\|r_0\|_M e_1 - H_k^M z_k\|_{R_{k+1}^{-T}R_{k+1}^{-1}} \\ &\leq \sqrt{\lambda_{max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|\|r_0\|_M e_1 - H_k^M z_k\|_2,\end{aligned}$$

where the last line follows from application of the Courant-Fischer theorem (see, for example, [27, Theorem 4.2.11]) to $R_{k+1}^{-T}R_{k+1}^{-1}$. We see from (3.4) that this is equivalent to

$$\|r_k^W\|_W \leq \sqrt{\lambda_{max}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M.$$

A similar process shows that

$$\sqrt{\lambda_{min}(R_{k+1}^{-T}R_{k+1}^{-1})}\|r_k^M\|_M \leq \|r_k^W\|_W,$$

which completes the proof. ∎

The bound (3.8) depends on the iteration number, k. The following is independent of the iteration number but—although there may be certain right-hand sides for which it is tight—it may be pessimistic in general, particularly when k is small. This topic is discussed in [43] and some examples for which the following bound is poor are given in §5.

Theorem 4 *The residuals of W-MR and M-MR satisfy*

$$\sqrt{\lambda_{min}(M^{-1}W)}\|r_k^M\|_M \leq \|r_k^W\|_W \leq \sqrt{\lambda_{max}(M^{-1}W)}\|r_k^M\|_M, \tag{3.9}$$

where $\lambda_{min}(M^{-1}W)$ and $\lambda_{max}(M^{-1}W)$ are the minimum and maximum eigenvalues of $M^{-1}W$, respectively

Proof Let the positive square root of $M = S^2$ be S. Then for any vector $y \in \mathbb{R}^n$,

$$\begin{aligned}\|y\|W^2 = y^TWy = y^TSS^{-1}WS^{-1}Sy &\leq \lambda_{max}(S^{-1}WS^{-1})\|Sy\|_2^2 \\ &= \lambda_{max}(M^{-1}W)\|y\|_M^2\end{aligned}$$

since $S^{-1}WS^{-1}$ and $M^{-1}W$ are similar. The lower bound is similar. ∎

4 Preconditioning

We have seen that when minimum residual methods in $\langle\cdot,\cdot\rangle_M$ are applied to a H-self-adjoint matrix, convergence depends primarily on its eigenvalues (except for certain right-hand side vectors for which convergence may be faster). Unless these eigenvalues are favourably distributed, convergence will be poor and preconditioning will be required. Here we investigate how the results of the previous sections may assist in the construction of effective preconditioners, P, for I-MR methods based on results for MR methods in nonstandard inner products. We concentrate on left preconditioning but show how the results may be adapted to right preconditioning.

If the preconditioned system $P^{-1}B$ is itself self-adjoint with respect to some symmetric bilinear form $\langle\cdot,\cdot\rangle_H$, the analysis of §2 implies that a minimum residual method in M can be applied, where M is related to H by (2.6). Its convergence depends on the spectrum of $P^{-1}B$ and will, therefore, be rapid if the eigenvalues are well distributed. Symmetric bilinear forms with respect to which a preconditioned matrix is self-adjoint have been characterised for certain matrices (see [44, §13] and the references therein). However, self-adjointness can *always* be preserved since if B and P are both H-self adjoint then

$$(HP)(P^{-1}B) = HB = B^TH = B^TP^{-T}P^TH = (P^{-1}B)^T(HP). \tag{4.1}$$

Thus, HP induces a symmetric bilinear form with respect to which $P^{-1}B$ is self-adjoint (cf. (2.4)). Consequently, there exists a diagonalisation $P^{-1}B = S\Lambda S^{-1}$ of $P^{-1}B$ such that $HP = S^{-*}YS^{-1}$. Furthermore, if $M = S^{-*}S^{-1}$ then by Lemma 1 the preconditioned M-MR residuals, $P^{-1}r_k^M$, satisfy

$$\frac{\|P^{-1}r_k^M\|_M}{\|P^{-1}r_0\|_M} \leq \min_{p\in\Pi_k,\, p(0)=1} \max_{\lambda\in\sigma(P^{-1}B)} |p(\lambda)|,$$

and convergence of M-MR applied to this system depends largely on eigenvalues. A similar result holds for right preconditioning: if P and B are H-self-adjoint it follows that BP^{-1} is HB^{-1}-self-adjoint. Thus, the following convergence results for left-preconditioning can be adapted to right-preconditioning by instead using $\langle\cdot,\cdot\rangle_{HB^{-1}}$. (Clearly, one would not form HB^{-1} but these results could be related to I-MR through the results of §3.)

Unfortunately, for most applications a H-self-adjoint preconditioner is not known (except when $H = I$). However, if a preconditioner is self-adjoint with respect to a symmetric bilinear form that is close to H, say $\mathcal{H}$, we might expect that we can find a matrix $\mathcal{S}$ such that $\mathcal{H}P = \mathcal{S}^{-*}Y\mathcal{S}^{-1}$ and $P^{-1}B = (\mathcal{S}-E)\Lambda(\mathcal{S}-E)^{-1}$ is a diagonalisation of $P^{-1}B$, with E a small perturbation. If this is the case then convergence of $\mathcal{M}$-MR, with $\mathcal{M} = \mathcal{S}^{-*}S^{-1}$, is still governed chiefly by eigenvalues, as the following theorem illustrates.

Theorem 5 *Let B be self-adjoint with respect to $\langle\cdot,\cdot\rangle_H$ and P be self-adjoint with respect to $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. Let*

$$\mathcal{H}P = \mathcal{S}^{-*}Y\mathcal{S}^{-1}, \tag{4.2}$$

and suppose $P^{-1}B$ *has diagonalisation*

$$P^{-1}B = (\mathcal{S} - E)\Lambda(\mathcal{S} - E)^{-1}. \tag{4.3}$$

Additionally, let $\mathcal{M} = \mathcal{S}^{-*}\mathcal{S}^{-1}$. *Then, the preconditioned residual* $P^{-1}r_k^{\mathcal{M}}$ *of any minimum residual method in* $\langle\cdot,\cdot\rangle_{\mathcal{M}}$ *satisfies*

$$\frac{\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}}}{\|P^{-1}r_0\|_{\mathcal{M}}} \leq \kappa_2(I - \mathcal{S}^{-1}E)\|p_k(\Lambda)\|_2. \tag{4.4}$$

If, moreover, $\|\mathcal{S}^{-1}E\|_2 < 1$ *then*

$$\frac{\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}}}{\|P^{-1}r_0\|_{\mathcal{M}}} \leq \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p\in\Pi_k,\, p(0)=1} \max_{\lambda\in\sigma(P^{-1}B)} |p(\lambda)|. \tag{4.5}$$

Proof We start from the polynomial representation of the residual, $P^{-1}r_k^{\mathcal{M}} = p_k(P^{-1}B)P^{-1}r_0$. Upon taking $\mathcal{M}$-norms we obtain

$$\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}} = \|p_k(P^{-1}B)P^{-1}r_0\|_{\mathcal{M}}.$$

From (2.2) and (4.3) it follows that

$$\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}} \leq \|\mathcal{S}^{-1}(\mathcal{S} - E)p_k(\Lambda)(\mathcal{S} - E)^{-1}\mathcal{S}\|_2\|P^{-1}r_0\|_{\mathcal{M}}.$$

Thus,

$$\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}} \leq \|I - \mathcal{S}^{-1}E\|_2\|(I - \mathcal{S}^{-1}E)^{-1}\|_2\|p_k(\Lambda)\|_2\|P^{-1}r_0\|_{\mathcal{M}},$$

which gives (4.4). If $\|\mathcal{S}^{-1}E\|_2 < 1$ then

$$\|(I - \mathcal{S}^{-1}E)^{-1}\|_2 \leq \frac{1}{1 - \|\mathcal{S}^{-1}E\|_2}$$

(see, for example, [27, page 301]) and

$$\frac{\|P^{-1}r_k^{\mathcal{M}}\|_{\mathcal{M}}}{\|P^{-1}r_0\|_{\mathcal{M}}} \leq \frac{1 + \|\mathcal{S}^{-1}E\|_2}{1 - \|\mathcal{S}^{-1}E\|_2} \min_{p\in\Pi_k,\, p(0)=1} \max_{\lambda\in\sigma(P^{-1}B)} |p(\lambda)|.$$

■

Remark This bound sheds some light on how to precondition a H-self-adjoint matrix. If the preconditioner P is self-adjoint with respect to a symmetric bilinear form which is close to $\langle\cdot,\cdot\rangle_H$, $\mathcal{M}$-MR convergence will essentially depend on the eigenvalues of $P^{-1}B$. If these eigenvalues have a more favourable distribution than those of the original system, convergence should be faster. This is analogous to the well understood convergence of symmetric matrices. One might expect that P will be self-adjoint with respect to such a symmetric bilinear form if P is, for example, a different discretisation of the same differential operator, or an approximation of this operator, in the context of partial differential equations. This may also explain why certain general preconditioners, such as incomplete factorisations, often perform well even when the coefficient matrix is nonsymmetric. Examples of both of these types of preconditioners are explored in the next section. Another example may be found in [8], in which the authors discuss how preconditioners for the Euler equations are improved when the eigenvectors of the preconditioner are considered.

Applying a M-MR method involves inner products with $\mathcal{M}$, which may be costly to compute even if this matrix is known. However, the analysis of §3 allows us to relate the convergence of any $\mathcal{M}$-MR method to that of any I-MR method. The results of §5 indicate that this bound can be very pessimistic; that is, that the bound (4.4) may be more descriptive of what is actually seen even when preconditioned residuals are computed with an I-MR method, such as I-GMRES, and are measured in the Euclidean norm.

5 Numerical experiments

In this section we examine some well-known preconditioners in light of the results of §4. Throughout, relative preconditioned I-MR residuals are computed by the in-built Matlab GMRES function, which is based on the unrestarted GMRES method of Saad and Schultz [41]. Relative preconditioned $\mathcal{M}$-MR residuals, with $\mathcal{M}$ as in Theorem 5, are computed by $\mathcal{M}$-GMRES. To simplify the computations, the symmetric bilinear form $\mathcal{H}P$ is given by $\mathcal{H}P = Z_P^{-*} Y Z_P^{-1}$, where Z_P has as its columns the eigenvectors of P computed by `eig` in Matlab. Accordingly, P has diagonalisation $P = \mathcal{Z}_P \Lambda_P \mathcal{Z}_P^{-1}$, where $\mathcal{Z}_P = Z_P \Lambda_P^{\frac{1}{2}}$, and $\mathcal{S}$ in Theorem 5 is equal to Z_P. We reiterate that it is in general infeasible to compute with nonstandard inner products but is done here for illustrative purposes.

Example 6 The geometric multigrid (GMG) preconditioner of Ramage [40] has been shown to be effective for advection-diffusion problems—for which the viscosity is not too small—which are approximated by streamline upwind Petrov-Galerkin (SUPG) finite elements. Here we investigate this preconditioner for two advection-diffusion problems in the context of the theory of §4.

The test problems we consider are the third and fourth examples in [16, Chapter 3], hereafter AD1 and AD2. Each is of the form

$$-\frac{1}{200}\nabla^2 u + w \cdot \nabla u = f \text{ in } \Omega = (-1,1) \times (-1,1), \quad u = g \text{ on } \delta\Omega.$$

Problem AD1 has a wind at a 30° angle to the left of vertical, so that

$$w = (-sin\frac{\pi}{6}, cos\frac{\pi}{6})^T,$$

with boundary conditions

$$\begin{aligned} u(x,-1) &= 1, & 0 < x \leq 1 \\ u(1,y) &= 1, & -1 \leq y < 1 \\ u(x,y) &= 0, & \text{elsewhere on } \delta\Omega. \end{aligned}$$

The second, AD2, is characterised by a recirculating wind,

$$w = (2y(1-x^2), -2x(1-y^2))^T,$$

	$\kappa_2(B)$	$\kappa_2(P)$	$\kappa_2(Z_B)$	$\kappa_2(Z_P)$	$\kappa_2(Z_{P^{-1}B})$	λ_{min}	λ_{max}
AD1	89.1	89.1	1.7×10^8	1.6×10^8	4.5×10^{10}	0.96	1.01
AD2	2.6×10^3	1.5×10^3	73.8	87.6	2.8×10^{10}	0.59	1.03

Table 2: Condtion numbers of B and P, and of the eigenvectors of P, B and $P^{-1}B$, as well as $\min|\lambda(P^{-1}B)|$ and $\max|\lambda(P^{-1}B)|$ for the advection-diffusion problems in Example 6.

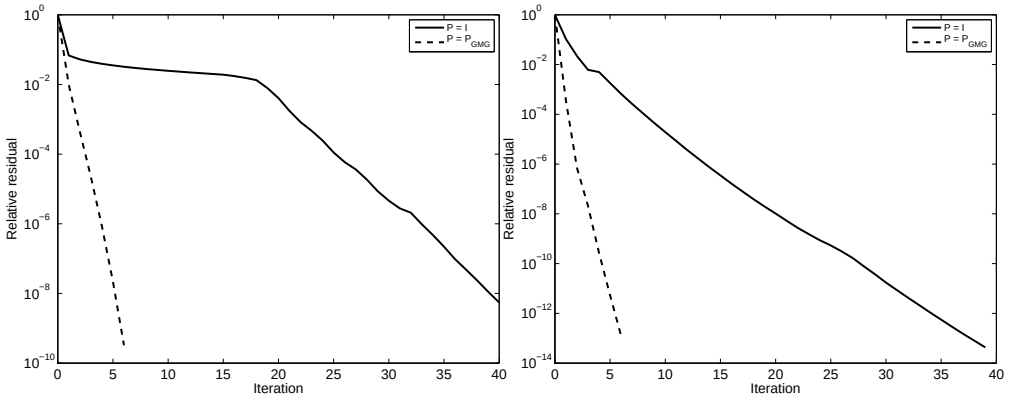
and boundary conditions

$$\begin{aligned} u(1,y) &= 1 && -1 \leq y \leq 1 \\ u(x,y) &= 0 && \text{elsewhere on } \delta\Omega. \end{aligned}$$

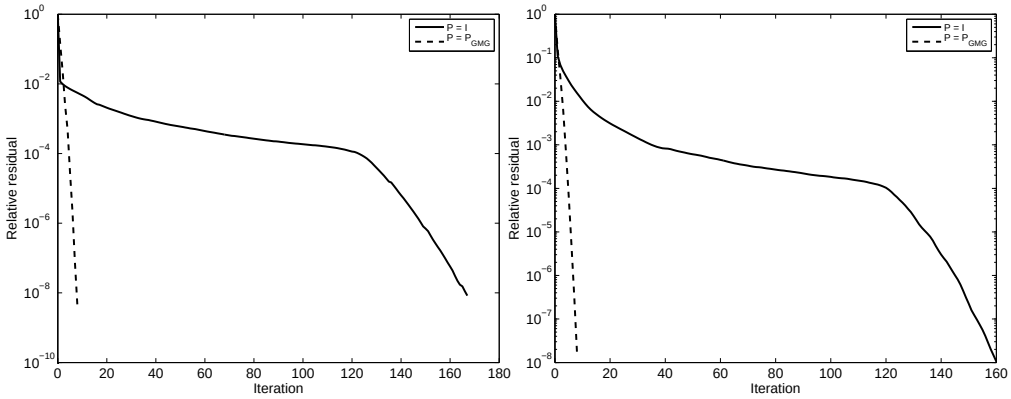
Both problems were discretised using Ifiss 3.1 [15] with the default settings of Q1 finite elements on a 16×16 grid, and SUPG with the default stabilisation parameter (defined in [16, Equation 3.44]). The multigrid preconditioner was also computed by Ifiss using the default settings of line Gauss-Seidel smoothing in 2 directions with 1 pre-smoothing step and 1 post-smoothing step. From Table 2 we see that although the eigenvalues of the preconditioned system are reasonably clustered, preconditioning makes the condition number of the eigenvectors worse. This is particularly true of AD2, for which the condition number of the eigenvectors of $P^{-1}B$ is 8 orders of magnitude larger than the condition number of the eigenvectors of B. Consequently, the standard GMRES bound (1.2) is incapable of predicting the fast convergence for both I-MR (with residuals measured in the I-norm) and $\mathcal{M}$-MR (with residuals measured in the $\mathcal{M}$-norm) which is evident in Figure 2. This fast convergence is related to the nearness of the symmetric bilinear forms with respect to which P and B are self-adjoint. We also comment on the I-MR and M-MR residuals of the unpreconditioned systems. Although the characteristic stagnation of unpreconditioned I-MR residuals [35] is clear, when the $\mathcal{M}$-orthogonal basis is employed this disappears for AD1. On the other hand, the convergence of $\mathcal{M}$-MR and I-MR are similar for the unpreconditioned problem AD2, indicating that for this problem the bound (3.9) is pessimistic.

It is clear from Table 2 that for both advection-diffusion problems the eigenvalues of $P^{-1}B$ are well-clustered. However, the GMG preconditioner achieves more than just good eigenvalues. Figure 3 shows that $|Z_P^{-1}P^{-1}BZ_P|$ is close to diagonal, indicating that $\mathcal{S} = Z_P$ is a good approximation of a matrix of eigenvectors of $P^{-1}B$. From this we infer that E in Theorem 5 is small; indeed, if $E = 0$ the matrix Z_P diagonalises $P^{-1}B$. Accordingly, for both AD1 and AD2 the convergence of $\mathcal{M}$-MR is chiefly governed by eigenvalues. Since the GMG-preconditioned I-MR residuals are qualitatively similar to the $\mathcal{M}$-MR residuals, we would expect that I-MR residuals are also determined by eigenvalues. We comment that the direct computation of E appears to be fraught with numerical difficulties.

The off-diagonal entries of $|Z_P^{-1}P^{-1}BZ_P|$, which are more prominent for AD1, are due at least in part to a nonzero E. Further, in both cases $P^{-1}B$ has many eigenvalues clustered near 1. Although this results in fast convergence of MR methods it also means that the eigenvectors are not very well-defined numerically. This may increase

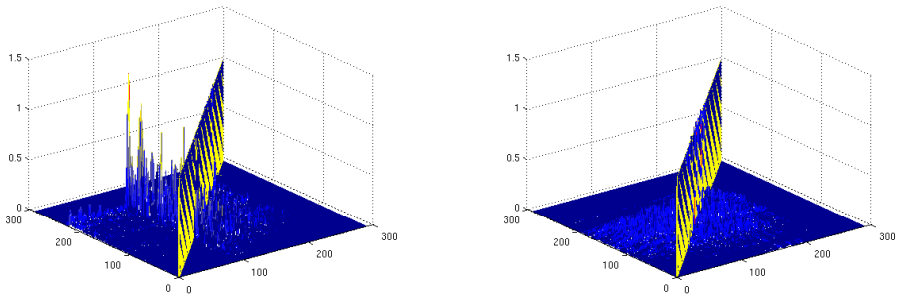


(a) AD1: I-MR (b) AD1: $\mathcal{M}$-MR



(c) AD2: I-MR (d) AD2: $\mathcal{M}$-MR

Figure 2: Relative residuals for I-MR (left column) and $\mathcal{M}$-MR convergence for AD1 and AD2 in Example 6.



(a) AD1 (b) AD2

Figure 3: $\left|Z_P^{-1}P^{-1}BZ_P\right|$ for AD1 and AD2 in Example 6.

Matrix	$\kappa_2(B)$	$\kappa_2(P)$	$\kappa_2(Z_B)$	$\kappa_2(Z_P)$	$\kappa_2(Z_{P^{-1}B})$	λ_{min}	λ_{max}
cdde2	54.8	102.3	4.3×10^{23}	4.8×10^{23}	4.7×10^{23}	0.97	1.07
comsol	1.4×10^{6}	99.4	3.0	15.5	283.6	2.9×10^{-5}	1.87
orsirr_1	7.7×10^{4}	4.2×10^{4}	5.4	5.2	277	0.041	1.68
pivtol	109.6	109.6	5.1	5.2	66.5	0.85	1.05
poisson2D	132.8	9.0	1.2	1.3	5000	0.057	1.40
sherman4	2179	70.5	2.8	2.9	26.2	0.010	1.28
steam1	2.8×10^{7}	2.8×10^{7}	3.07	1.11	4.6×10^{16}	0.99	1.01

Table 3: Condtion numbers of B, P, and of the eigenvectors of P, B and $P^{-1}B$, as well as $\min|\lambda(P^{-1}B)|$ and $\max|\lambda(P^{-1}B)|$ for the ILU(0) test problems in Example 7.

the size of off-diagonal entries further, and prevents computation of $Z_P^{-1}P^{-1}BZ_P$ for larger grids. This does not indicate a deficiency in the theory of §4 but merely highlights the well-known difficulties of obtaining eigenvalues and eigenvectors of highly nonnormal matrices.

Example 7 Incomplete LU (ILU) factorisations constitute one of the most well-known classes of general-purpose preconditioners. It is known that ILU preconditioners tend to cluster eigenvalues, but for nonnormal matrices well-clustered eigenvalues do not necessarily lead to fast GMRES convergence. We examine ILU(0) preconditioning [36], as implemented in Matlab, applied to several test matrices from the University of Florida Sparse Matrix Collection [10] and show that it is indeed the eigenvalues of the preconditioned system that lead to rapid convergence; that is, the eigenvectors are not so important for these examples. As in Example 6 the nonnormality of B, P and $P^{-1}B$ made infeasible the computation of $\mathcal{Z}P^{-1}B\mathcal{Z}$ for many of the problems from the collection. In particular, the clustering of the eigenvalues achieved by ILU(0) preconditioning caused the eigenvectors to be poorly numerically defined. However, for several examples, namely cdde2, comsol, orsirr_1, pivtol, poisson2D, sherman4 and steam1, computations were possible. We examine these in detail here.

Table 3 shows that, as in the previous example, the eigenvalues of most of the pre-

conditioned systems are clustered near 1, as is typical of ILU(0) preconditioned systems. However, in all cases the conditioning of the eigenvectors of $P^{-1}B$ is worse than the conditioning of the eigenvectors of B. This is particularly pronounced for poisson2D and steam1. Consequently, the standard bound (1.2) is unlikely to describe accurately the fast convergence of the preconditioned systems, shown in Figures 4 and 5. This fast convergence is, however, described by Theorem 5 since P and B are self-adjoint with respect to nearby symmetric bilinear forms.

In all but one case the original system has poorly distributed eigenvalues but well-conditioned eigenvectors. By contrast, preconditioning achieves well-distributed eigenvalues but makes the eigenvectors worse. However, Figure 6 shows that the matrix $|Z_P^{-1}P^{-1}BZ_P|$ is almost diagonal for each test problem, from which we infer that E in Theorem 5 is small. Consequently, convergence of $\mathcal{M}$-MR principally depends on the eigenvalues of $P^{-1}B$, which are close to 1, and convergence is much faster than for the original system. Since the convergence of I-MR is qualitatively similar to that of $\mathcal{M}$-MR for each of these problems, we assume that its convergence is also governed by eigenvalues.

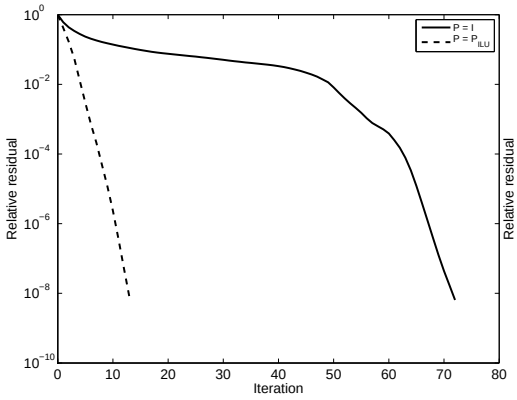
We make special mention of steam1 for which, despite $\kappa_2(Z_P)$ and E being small (see Table 3 and Figure 6), $\kappa_2(Z_{P^{-1}B})$ is large. This is not precluded by our theory, since

$$\kappa_2(Z_{P^{-1}B}) = \kappa_2(Z_P - E) \leq \kappa_2(Z_P)\kappa_2(I - Z_P^{-1}E).$$

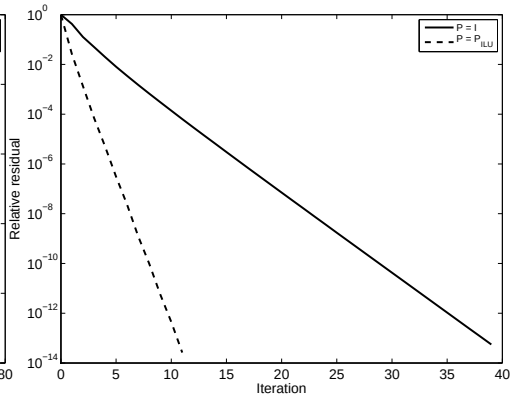
Even when Z_P is well-conditioned, $I - Z_P^{-1}E$ may be poorly conditioned. This will happen, for example, if $Z_P^{-1}E$ has as its jth row the jth unit vector.

Example 8 We next construct preconditioners for the matrices comsol and sherman4 which have the same eigenvectors as the ILU(0) preconditioners of the previous example but for which the eigenvalues differ. These examples are artificially designed to show that the eigenvalues essentially determine convergence when the symmetric bilinear forms for B and P are close, so that poor convergence follows when the eigenvalues are poorly distributed. Specifically, if $P_{ILU} = Z_P\Lambda_P Z_P^{-1}$ then the new preconditioner has diagonalisation $P_{pert} = Z_P\Lambda_{pert}Z_P^{-1}$, where the ith diagonal entry of Λ_{pert} is given by $(\Lambda_{pert})_{ii} = (1 + 10x)(\Lambda_P)_{ii}$, with $x \sim \mathcal{N}(0, 1)$ a random normal variable. Thus, P_{pert} has the same eigenvectors as P_{ILU}, and is self-adjoint with respect to the same symmetric bilinear form, but has different eigenvalues.

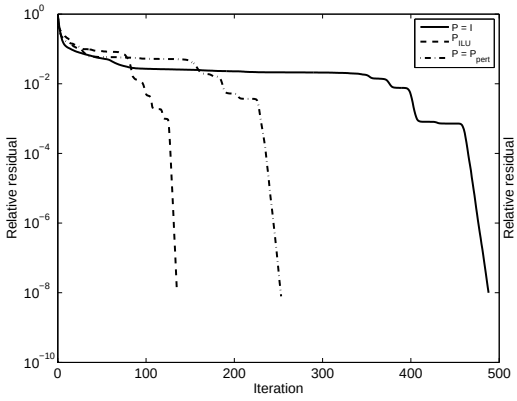
That P_{pert} and B are self-adjoint with respect to nearby symmetric bilinear forms is clear from Figure 7. For both of these problems, $|Z_P^{-1}P_{pert}^{-1}BZ_P|$ is almost diagonal, although now the diagonal entries have differing magnitudes, indicating that the eigenvalues of the preconditioned systems are not as well distributed as the eigenvalues of the systems in Example 7. The convergence of I-MR and $\mathcal{M}$-MR for these preconditioned problems is given in Figures 4 and 5. Comparison with the convergence for the systems preconditioned by the ILU(0) preconditioner show that in each case the convergence is significantly poorer, implying that the poor eigenvalues impede convergence and that, moreover, it is eigenvalues that determine the convergence of these nonsymmetric systems.



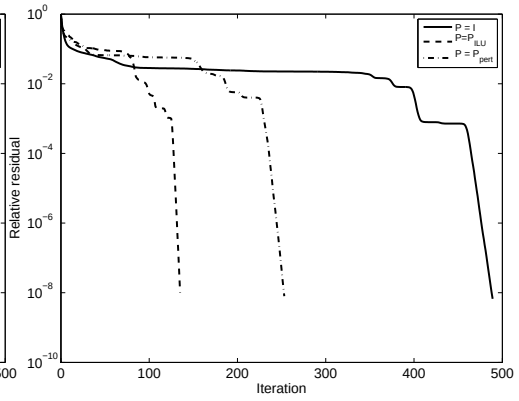
(a) cdde2: I-MR



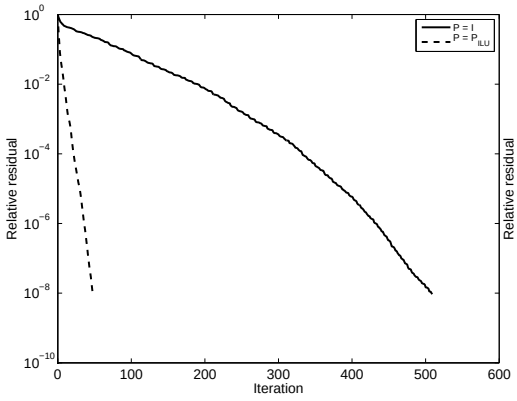
(b) cdde2: $\mathcal{M}$-MR



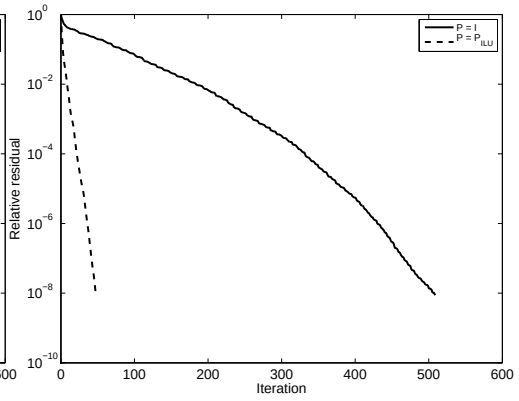
(c) comsol: I-MR



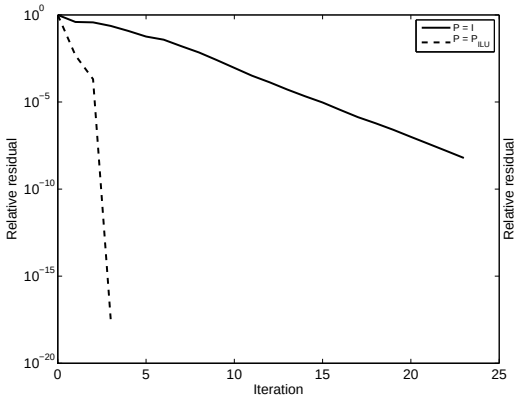
(d) comsol: $\mathcal{M}$-MR

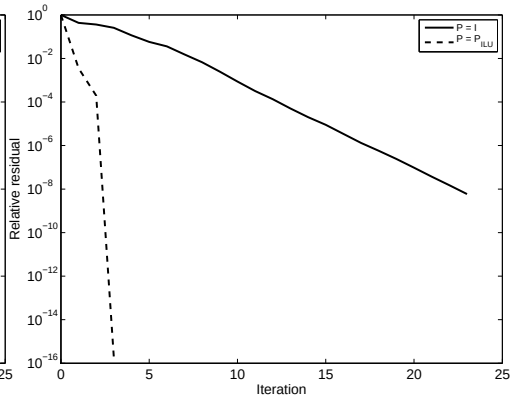


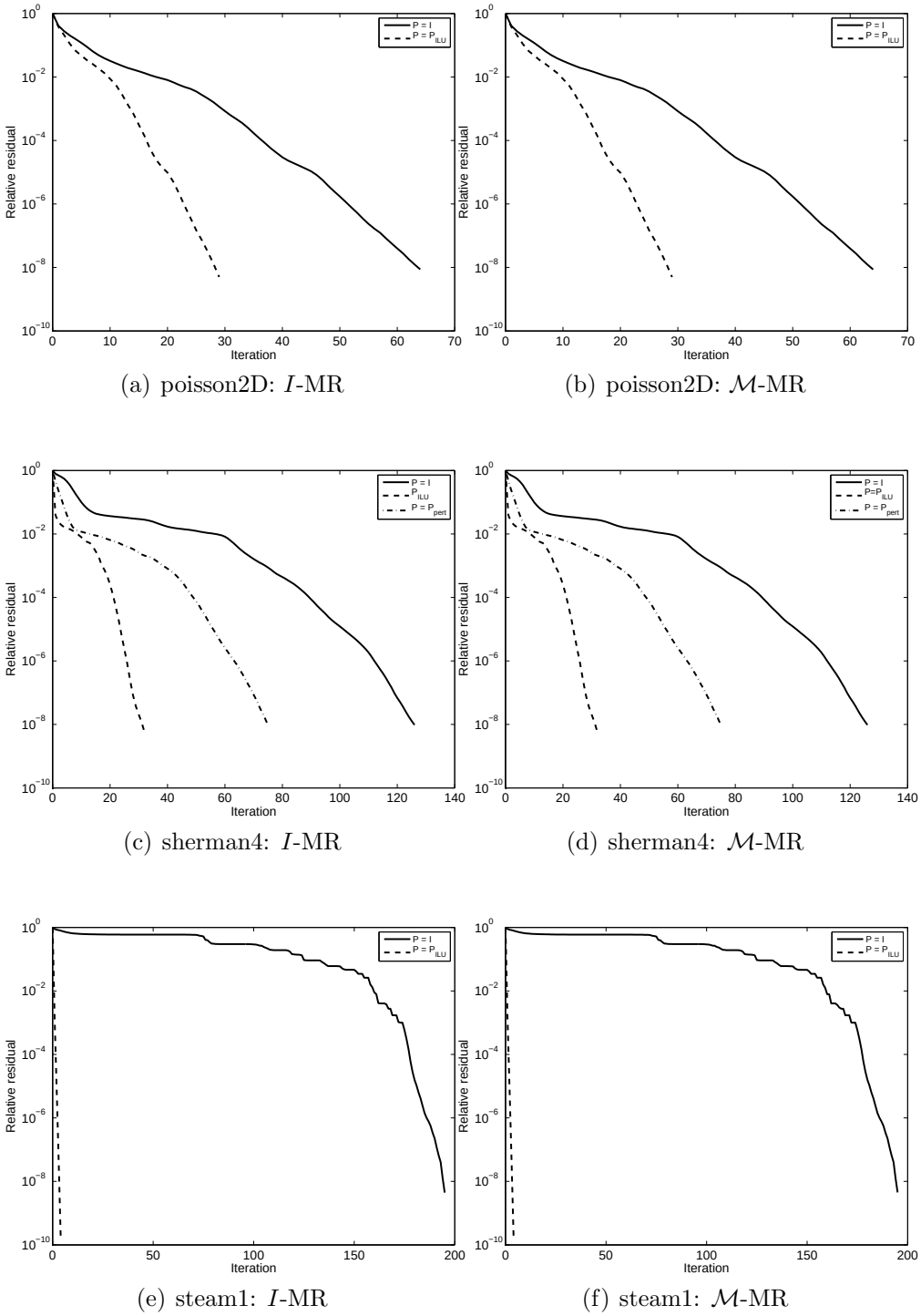
(e) orsirr_1: I-MR



(f) orsirr_1: $\mathcal{M}$-MR





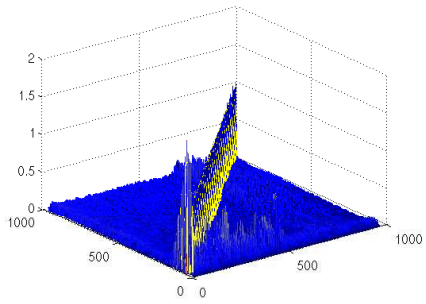


(a) poisson2D: I-MR (b) poisson2D: $\mathcal{M}$-MR

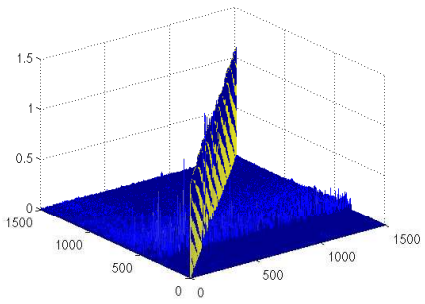
(c) sherman4: I-MR (d) sherman4: $\mathcal{M}$-MR

(e) steam1: I-MR (f) steam1: $\mathcal{M}$-MR

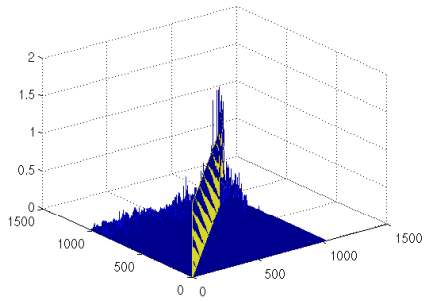
Figure 5: Relative residuals for I-MR (left column) and $\mathcal{M}$-MR (right column) for the ILU(0) test problems of Example 7. Plots (c) and (d) also show the convergence for the modified preconditioner in Example 8.



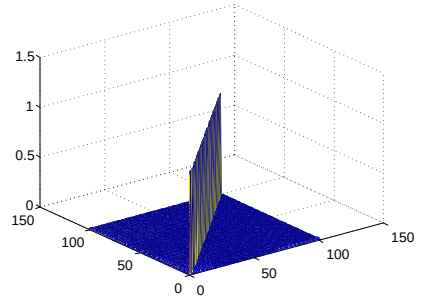
(a) cdde2



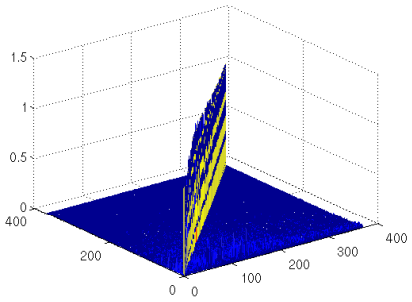
(b) comsol



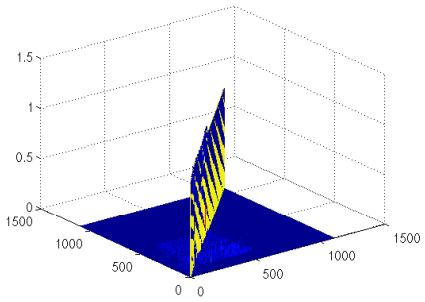
(c) orsirr_1



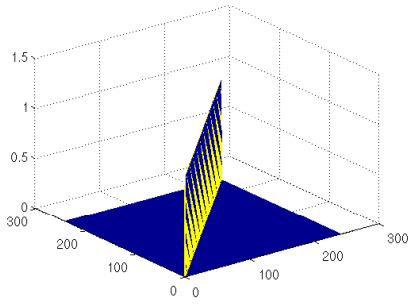
(d) pivtol

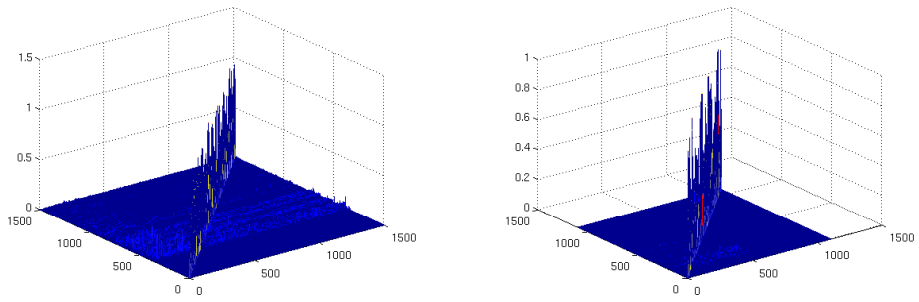


(e) poisson2D



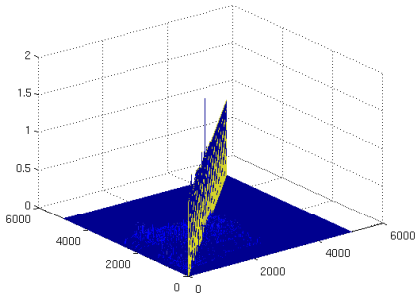
(f) sherman4





(a) comsol (b) sherman4

Figure 7: $\left|Z_P^{-1}P^{-1}BZ_P\right|$ for the modified ILU(0) test problems of Example 8.



(a)

Figure 8: (a) $|Z_P P^{-1}BZ_P|$ for sherman3 (Example 9).

Example 9 Finally, we return to the sherman3 example in §1 and explain the rapid convergence of I-GMRES applied to the ILU(0) preconditioned system shown in Figure 1. The sherman3 matrix B has real eigenvalues and so is self-adjoint with respect to an inner product $\langle\cdot,\cdot\rangle_H$. The ILU(0) preconditioner P is also self-adjoint with respect to an inner product, $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. If $\langle\cdot,\cdot\rangle_H$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ are close we would expect E in (4.3) to be small and that convergence is dominated by eigenvalues. For this matrix, $|Z_P^{-1}P^{-1}BZ_P|$ is, as in the previous examples, close to diagonal (see Figure 8); we infer that E is small. Consequently, the preceding theory indicates that the convergence of $\mathcal{M}$-MR (and, we suspect, I-MR) essentially depend on the (well-distributed) eigenvalues of the preconditioned matrix.

6 Conclusions

In the context of preconditioning for the iterative solution of nonsymmetric linear systems for diagonalisable nonsymmetric systems we have identified a mathematical property of preconditioning matrices which indicates when convergence of any minimum residual method (such as GMRES) depends essentially only on the eigenvalues of the preconditioned system. This is more consistent with convergence properties for symmetric systems.

Acknowledgment. We are indebted to two anonymous referees whose comments, suggestions and additional references greatly improved a related manuscript. In particular, we thank one referee for the proof of Theorem 4, which is much more straightforward than the original. We are extremely grateful to Michele Benzi, Jörg Liesen, Miro Rozložník, Valeria Simoncini and Zdeněk Strakoš for their helpful comments and suggestions that significantly improved a related manuscript.

References

[1] M. Arioli, V. Pták, and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.

[2] S. F. Ashby, M. J. Holst, T. A. Manteuffel, and P. E. Saylor, *The role of the inner product in stopping criteria for conjugate gradient iterations*, BIT, 41 (2001), pp. 26–52.

[3] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[4] B. Beckermann, S. A. Goreinov, and E. E. Tyrtyshinikov, *Some remards on the Elman estimate for GMRES*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 772–778.

[5] M. Benzi and V. Simoncini, *On the eigenvalues of a class of saddle point matrices*, Numer. Math., 103 (2006), pp. 173–196.

[6] M. Benzi and M. Tůma, *A comparative study of sparse approximate inverse preconditioners*, Appl. Numer. Math., 30 (1999), pp. 305–340.

[7] A. Chapman, Y. Saad, and L. Wigton, *High-order ILU preconditioners for CFD problems*, Int. J. Numer. Meth. Fluids, 33 (2000), pp. 767–788.

[8] D. L. Darmofal and P. J. Schmid, *The importance of eigenvectors for local preconditioners of the Euler equations*, J. Comput. Phys., 127 (1996), pp. 346–362.

[9] T. A. Davis, *Direct Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2006.

[10] T. A. Davis and Y. Huang, *The University of Florida sparse matrix collection*, ACM T. Math. Software, (to appear). http://www.cise.ufl.edu/research/sparse/matrices.

[11] I. S. Duff, A. M. Erisman, and J. K. Reid, *Direct Methods for Sparse Matrices*, Oxford University Press, Oxford, 1989.

[12] M. Eiermann, *Fields of values and iterative methods*, Linear Algebra Appl., 180 (1993), pp. 167–197.

[13] M. Eiermann and O. G. Ernst, *Geometric aspects in the theory of Krylov subspace methods*, Acta Numer., 10 (2001), pp. 251–312.

[14] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[15] Howard Elman, Alison Ramage, and David Silvester, *Algorithm 866: IFISS, a Matlab toolbox for modelling incompressible flow*, ACM Trans. Math. Softw., 33 (2007), pp. 2–14,.

[16] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: with applications in incompressible fluid dynamics*, Oxford University Press, Oxford, 2005.

[17] A. Essai, *Weighted FOM and GMRES for solving nonsymmetric linear systems*, Numer. Algorithms, 18 (1998), pp. 277–292.

[18] V. Faber, J. Liesen, and P. Tichý, *The Faber-Manteuffel theorem for linear operators*, SIAM J. Numer. Anal., 46 (2008), pp. 1323–1337.

[19] V. Faber and T. Manteuffel, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.

[20] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Methods Dundee, G. A. Watson, ed., Springer-Verlag, New York, 1976, pp. 73–89.

[21] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[22] I. Gohberg, P. Lancaster, and L. Rodman, *Indefinite Linear Algebra and Applications*, Birkhäuser Verlag, Basel, 2005.

[23] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

[24] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.

[25] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov varieties*, in Recent Advances in Iterative Methods, G. Golub, A. Greenbaum, and M. Luskin, eds., vol. 60 of IMA Volumes in Mathematics and its Applications, New York, 1994, Springer, pp. 95–118.

[26] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.

[27] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.

[28] C. Jagels and L. Reichel, *The isometric Arnoldi process and an application to iterative solution of large linear systems*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., Elsevier, Amsterdam, 1992, pp. 361–369.

[29] C. F. Jagels and L. Reichel, *A fast minimal residual algorithm for shifted unitary matrices*, Numer. Linear Algebra Appl., 1 (1994), pp. 555–570.

[30] Z. Jia and B. Zhu, *A power sparse approximate inverse preconditioner for large sparse linear systems*, Numer. Linear Algebra Appl., 16 (2009), pp. 259–299.

[31] W. D. Joubert and D. M. Young, *Necessary and sufficient conditions for the simplification of generalized conjugate gradient algorithms*, Linear Algebra Appl., 88/89 (1987), pp. 449–485.

[32] H.-B. Li, T.-Z. Huang, Y. Zhang, X.-P. Liu, and H. Li, *On some new approximate factorization methods for block tridiagonal matrices suitable for vector and parallel processors*, Math. Comput. Simulat., 79 (2009), pp. 2135–2147.

[33] J. Liesen and B. Parlett, *On nonsymmetric saddle point matrices that allow conjugate gradient iterations*, Numer. Math., 108 (2008), pp. 605–624.

[34] J. Liesen and Z. Strakoš, *On optimal short recurrences for generating orthogonal Krylov subspace bases*, SIAM Rev., 50 (2008), pp. 485–503.

[35] J. Liesen and Z. Strakoš, *GMRES convergence analysis for a convection-diffusion model problem*, SIAM J. Sci. Comput., 26 (2005), pp. 1989–2009.

[36] J. A. Meijerink and H. A. van der Vorst, *An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp., 31 (1977), pp. 148–162.

[37] G. Meurant and Z. Strakoš, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numer., 15 (2006), pp. 471–542.

[38] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.

[39] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[40] A. Ramage, *A multigrid preconditioner for stabilised discretisations of advection-diffusion problems*, J. Comput. Appl. Math., 110 (1999), pp. 187–203.

[41] Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. and Stat. Comput., 7 (1986), pp. 856–869.

[42] H. Sadok and D. B. Szyld, *A new look at CMRH and its relation to GMRES*, Tech. Report 10-04-25, Department of Mathematics, Temple University, Apr. 2010.

[43] V Simoncini and D. B. Szyld, *The effect of non-optimal bases on the convergence of Krylov subspace methods*, Numer. Math., 100 (2005), pp. 711–733.

[44] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.

[45] G. L. G. Sleijpen and D. R. Fokkema, *Bi-CGStab(ℓ) for linear equations involving unsymmetric matrices with complex spectrum*, Elec. Trans. Numer. Anal., 1 (1993), pp. 11–32.

[46] P. Sonneveld, *CGS, a fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 36–52.

[47] M. Stoll and A. J. Wathen, *Combination preconditioning and the Bramble-Pasciak$^+$ preconditioner*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 582–608.

[48] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behaviour of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, 2005.

[49] H. A. van der Vorst, *Bi-CGStab: A fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.

[50] H. A. van der Vorst and Q. Ye, *Residual replacement strategies for Krylov subspace iterative methods for the convergence of true residuals*, SIAM J. Sci. Comput., 22 (2000), pp. 835–852.