ON THE EIGENVALUES AND EIGENVECTORS OF BLOCK TRIANGULAR PRECONDITIONED BLOCK MATRICES*

JENNIFER PESTANA†

Abstract. Block lower triangular matrices and block upper triangular matrices are popular preconditioners for 2×2 block matrices. In this note we show that a block lower triangular preconditioner gives the same spectrum as a block upper triangular preconditioner and that the eigenvectors of the two preconditioned matrices are related.

Key words. block triangular preconditioner, convergence, eigenvalues, eigenvectors, iterative method, saddle point system

AMS subject classifications. 65F08, 65F10, 65F50, 65N22

1. Introduction.

Nonsingular block matrices of the form

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}, \tag{1.1}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times n}$ with $\operatorname{rank}(B) = m$, $C \in \mathbb{C}^{m\times m}$, $m \le n$ arise in a number of applications, many of which are discussed in the survey paper by Benzi, Golub and Liesen [5, Section 2]. Of particular interest are block matrices for which $C = 0$ and/or for which A is symmetric positive definite and C is symmetric positive semidefinite [5],[10, Chapters 5 and 7].

In many applications $\mathcal{A}$ in (1.1) is large and sparse, in which case linear systems with $\mathcal{A}$ as the coefficient matrix are typically solved by a preconditioned iterative method. Two popular preconditioners are the block lower triangular matrix [8, 16, 17]

$$\mathcal{P}_L = \begin{bmatrix} P_A & 0 \\ B & P_S \end{bmatrix}, \tag{1.2}$$

and block upper triangular matrix [6, 13, 15, 16, 20]

$$\mathcal{P}_U = \begin{bmatrix} P_A & B^T \\ 0 & P_S \end{bmatrix}, \tag{1.3}$$

where $P_A \in \mathbb{C}^{n\times n}$ and $P_S \in \mathbb{C}^{m\times m}$ (and, consequently, $\mathcal{P}_L$ and $\mathcal{P}_U$) are nonsingular.

When A and C are Hermitian semidefinite it is known that $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{P}_L^{-1}\mathcal{A}$ are similar [14, Remark 2]. (The case in which A is positive definite was also recently treated by Notay [16, Theorem 3.1].) For non-Hermitian matrices, Bai and Ng [3] analysed the minimal polynomials of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ when $P_A = A$ or P_S is the Schur complement, while Bai [1] obtained identical eigenvalue bounds for $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ in the more general case of inexact P_A and P_S[1]. Additionally, Bai and Ren [4] applied block triangular preconditioned GMRES [18] and BiCGStab [22] to nonsymmetric 2×2 block systems arising from discretizations of third-order ODEs. They

*This publication is based on work supported by Award No. KUK-C1-013-04, made by King Abdullah University of Science and Technology (KAUST).

†Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK (`pestana@maths.ox.ac.uk`)

[1]In fact the results of Bai and Ng [3] and Bai [1] are more general than ours since they do not assume that the (1,2) and (2,1) blocks of $\mathcal{A}$ are transposes of each other.

found that iteration counts for upper and lower triangular preconditioners were similar, while eigenvalue plots for $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ were indistinguishable. Additionally, for M-matrices arising from Markov chains Benzi and Uçar [7] noticed little difference between the performances of block lower triangular and block upper triangular preconditioners.

In this note we extend the theoretical results to the non-Hermitian case. We show that $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ have identical spectra and relate the corresponding eigenvectors. When $C = 0$ and $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ are diagonalizable, we bound the difference between the condition numbers of the eigenvector matrices; this gives some insight into when we might expect certain iterative methods to converge similarly for the block lower and block upper triangular preconditioned systems. Our results are illustrated on a numerical example.

Throughout $I_p \in \mathbb{C}^{p\times p}$ denotes the identity matrix of dimension p and $\|\cdot\|_2$ represents the Euclidean vector norm or the corresponding induced matrix norm. The conjugate transpose of a matrix E is denoted by E^*, its range by $\mathrm{range}(E)$, its nullspace by $\mathrm{null}(E)$ and its Moore-Penrose pseudoinverse by $E^\dagger$.

2. Eigenvalue, eigenvector and condition number relationships. In this section we state our main results, starting with the equivalence of the spectra of $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{P}_L^{-1}\mathcal{A}$.

PROPOSITION 1. *Let $\mathcal{A}$, $\mathcal{P}_L$ and $\mathcal{P}_U$ be invertible. Then the spectra of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ are identical, as are the spectra of $\mathcal{A}\mathcal{P}_L^{-1}$ and $\mathcal{A}\mathcal{P}_U^{-1}$.*

Proof. Let

$$Y(\lambda) = \begin{bmatrix} A - \lambda P_A & B^T \\ (1-\lambda)B & -(C + \lambda P_S) \end{bmatrix} \quad \text{and} \quad Z(\lambda) = \begin{bmatrix} A - \lambda P_A & (1-\lambda)B^T \\ B & -(C + \lambda P_S) \end{bmatrix}. \tag{2.1}$$

Then the eigenvalues λ_L of $\mathcal{P}_L^{-1}\mathcal{A}$ must be roots of $\det(Y(\lambda_L)) = 0$ while the eigenvalues λ_U of $\mathcal{P}_U^{-1}\mathcal{A}$ must satisfy $\det(Z(\lambda_U)) = 0$. Setting

$$J(\lambda) = \begin{bmatrix} I_n & 0 \\ 0 & (1-\lambda)I_m \end{bmatrix}, \tag{2.2}$$

we see that, for any $\lambda \neq 1$, $0 = \det(Y(\lambda_L)) = \det(J(\lambda_L)^{-1}Y(\lambda_L)J(\lambda_L)) = \det(Z(\lambda_L))$. Thus, the non-unit roots of $\det(Y(\lambda)) = 0$ and $\det(Z(\lambda)) = 0$ coincide. Additionally, $\det(Y(1)) = \det(Z(1)) = (-1)^m \det(A - P_A)\det(C + P_S)$. The results for right preconditioning follow from the similarity of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{P}_L^{-1}$ and of $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{P}_U^{-1}$. ∎

The above result shows that if eigenvalues alone are important, there is nothing to distinguish $\mathcal{P}_L^{-1}\mathcal{A}$, $\mathcal{A}\mathcal{P}_L^{-1}$, $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{P}_U^{-1}$. This may be the case when, for example, we precondition to achieve self-adjointness and positive definiteness with respect to a nonstandard inner product [8, 14, 17]. However, in many situations the eigenvectors will also have an effect on convergence. We relate the eigenvectors of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ in the following proposition.

PROPOSITION 2. *Let $\mathcal{A}$, $\mathcal{P}_L$ and $\mathcal{P}_U$ be nonsingular.*

1. *Suppose that $\lambda \neq 1$ is an eigenvalue of $\mathcal{P}_U^{-1}\mathcal{A}$. If $[\boldsymbol{u}_U^T, \boldsymbol{v}_U^T]^T$ is the corresponding eigenvector, then $J(\lambda)[\boldsymbol{u}_U^T, \boldsymbol{v}_U^T]^T = [\boldsymbol{u}_U^T, (1-\lambda)\boldsymbol{v}_U^T]^T$ is an eigenvector of $\mathcal{P}_L^{-1}\mathcal{A}$ corresponding to this eigenvalue.*
2. *Suppose that $\lambda = 1$ is an eigenvalue of $\mathcal{P}_L^{-1}\mathcal{A}$.*
    - *If $P_S + C$ is nonsingular then the eigenvectors of $\mathcal{P}_L^{-1}\mathcal{A}$ corresponding to λ are of the form $[\boldsymbol{u}^T, \boldsymbol{0}^T]^T$, where $\boldsymbol{u}$ is any eigenvector of the matrix*

$P_A^{-1}A$ that corresponds to its unit eigenvalue. Moreover, the eigenvectors of $\mathcal{P}_U^{-1}\mathcal{A}$ corresponding to λ are of the form $[\boldsymbol{u}^T, ((P_S + C)^{-1}B\boldsymbol{u})^T]^T$.

- *If $P_A - A$ is nonsingular then the eigenvectors of $\mathcal{P}_U^{-1}\mathcal{A}$ corresponding to λ are of the form $[\boldsymbol{0}^T, \boldsymbol{v}^T]^T$ where $\boldsymbol{v}$ is any eigenvector of $-P_S^{-1}C$ that corresponds to its unit eigenvalue. Moreover, the eigenvectors of $\mathcal{P}_L^{-1}\mathcal{A}$ corresponding to λ are of the form $[((P_A - A)^{-1}B^T\boldsymbol{v})^T, \boldsymbol{v}^T]^T$.*

Proof. The first part follows immediately from Proposition 1, since if $\lambda \neq 1$ then $Z(\lambda) = J(\lambda)^{-1}Y(\lambda)J(\lambda)$ where $J(\lambda)$ is as in (2.2).

If $\lambda = 1$, then $Y(1)\boldsymbol{w}_L = \boldsymbol{0}$ is equivalent to

$$A\boldsymbol{u}_L + B^T\boldsymbol{v}_L = P_A\boldsymbol{u}_L, \tag{2.3}$$

$$P_S\boldsymbol{v}_L + C\boldsymbol{v}_L = \boldsymbol{0}, \tag{2.4}$$

while $Z(1)\boldsymbol{w}_U = \boldsymbol{0}$ is equivalent to

$$A\boldsymbol{u}_U - P_A\boldsymbol{u}_U = \boldsymbol{0}, \tag{2.5}$$

$$B\boldsymbol{u}_U - C\boldsymbol{v}_U = P_S\boldsymbol{v}_U. \tag{2.6}$$

Note that $-P_S^{-1}C$ does not have an eigenvalue at 1 if and only if $P_S + C$ is nonsingular. In this case, (2.3) and (2.4) show that $\boldsymbol{v}_L = \boldsymbol{0}$ and that $\boldsymbol{u}_L$ must be an eigenvector of $P_A^{-1}A$ corresponding to the eigenvalue 1. Meanwhile, (2.5) and (2.6) imply $\boldsymbol{u}_U \neq 0$, since otherwise nonsingularity of $P_S + C$ means that $\boldsymbol{v}_U = 0$. Thus, $\boldsymbol{u}_U$ is also an eigenvector of $P_A^{-1}A$ corresponding to the eigenvalue 1 and $B\boldsymbol{u}_U = (C + P_S)\boldsymbol{v}_U$. The last case is proved similarly. ◻

REMARK 1. The case that $\lambda = 1$ is an eigenvalue of both $P_A^{-1}A$ and $-P_S^{-1}C$ can also be worked out but is of less interest so we omit it here.

REMARK 2. It may be preferable to use right preconditioning rather than left preconditioning, since doing so preserves the residual norm for methods such as GMRES [18]. However, if X is an eigenvector matrix of $\mathcal{P}^{-1}\mathcal{A}$ then $\mathcal{P}X$ is an eigenvector matrix of $\mathcal{A}\mathcal{P}^{-1}$ for any invertible preconditioner $\mathcal{P}$. This allows the eigenvectors of the right preconditioned block matrix to be determined. However, the eigenvector matrices of $\mathcal{A}\mathcal{P}_L^{-1}$ and $\mathcal{A}\mathcal{P}_U^{-1}$ do not have as straightforward a relationship as those of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ and we do not consider them here.

Particularly important is the case $C = 0$ in $\mathcal{A}$, for which the eigenvectors of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ are more simply related.

COROLLARY 3. *Let $\mathcal{A}$, $\mathcal{P}_L$ and $\mathcal{P}_U$ be nonsingular and let $C = 0$ in* (1.1).

1. *Suppose that $\lambda \neq 1$ is an eigenvalue of $\mathcal{P}_U^{-1}\mathcal{A}$. If $[\boldsymbol{u}_U^T, \boldsymbol{v}_U^T]^T$ is the corresponding eigenvector, then $[\boldsymbol{u}_U^T, (1-\lambda)\boldsymbol{v}_U^T]^T$ is an eigenvector of $\mathcal{P}_L^{-1}\mathcal{A}$ corresponding to this eigenvalue.*
2. *Otherwise, $\lambda = 1$ is an eigenvalue of $\mathcal{P}_L^{-1}\mathcal{A}$ with corresponding eigenvector $[\boldsymbol{u}^T, \boldsymbol{0}^T]^T$, where $\boldsymbol{u}$ is any eigenvector of the matrix $P_A^{-1}A$ that corresponds to its unit eigenvalue. Moreover, the eigenvectors of $\mathcal{P}_U^{-1}\mathcal{A}$ corresponding to λ are of the form $[\boldsymbol{u}^T, (P_S^{-1}B\boldsymbol{u})^T]^T$.*

Proof. If $C = 0$, then $P_S^{-1}C = 0$ and the last case in Proposition 2 does not apply. ◻

Of interest for Krylov methods such as GMRES is the condition number of the eigenvector matrix [18] when it is invertible.

COROLLARY 4. *Let $\mathcal{A}$, $\mathcal{P}_L$ and $\mathcal{P}_U$ be nonsingular and let $C = 0$ in $\mathcal{A}$. Assume that both $\mathcal{P}_L^{-1}\mathcal{A} = X_L\Lambda X_L^{-1}$ and $\mathcal{P}_U^{-1}\mathcal{A} = X_U\Lambda X_U^{-1}$ are diagonalizable with p* *eigenvalues equal to 1, where*

$$X_U = \begin{bmatrix} U^{(1)} & U^{(2)} \\ V^{(1)} & V^{(2)} \end{bmatrix} \text{ and } X_L = \begin{bmatrix} U^{(1)} & U^{(2)} \\ 0 & V^{(2)}(I - \Lambda^{(2)}) \end{bmatrix}, \tag{2.7}$$

with $U^{(1)} \in \mathbb{C}^{n\times p}$, $U^{(2)} \in \mathbb{C}^{n\times(n+m-p)}$, $V^{(1)} \in \mathbb{C}^{m\times p}$, $V^{(2)} \in \mathbb{C}^{m\times(n+m-p)}$ *and* $\Lambda = \mathrm{diag}(I_p, \Lambda^{(2)})$, $\Lambda^{(2)} = \mathrm{diag}(\lambda_{p+1}, \ldots, \lambda_{n+m})$. *For any matrix* E *let* $P_E = (I - EE^\dagger)$ *and* $Q_E = (I - E^\dagger E)$ *be orthogonal projectors onto* $\mathrm{null}(E)$ *and* $\mathrm{null}(E^*)$, *respectively. Then the 2-norm condition numbers* $\kappa_2(X_L)$ *of* X_L *and* $\kappa_2(X_U)$ *of* X_U *are related by*

$$\frac{1}{\alpha} \leq \frac{\kappa_2(X_L)}{\kappa_2(X_U)} \leq \alpha \tag{2.8}$$

where, when $p = 0$,

$$\alpha = \big(1 + \beta\big(1 + \|V^{(2)}(I - \Lambda^{(2)})(U^{(2)})^\dagger\|_2\big)\big)\big(1 + \beta\big(1 + \|V^{(2)}(U^{(2)})^\dagger\|_2\big)\big),$$
$$\beta = \|V^{(2)}\Lambda^{(2)}\|_2\big(\|(U^{(2)})^\dagger\|_2 + \|Q_{U^{(2)}}(V^{(2)})^* S^{-1}\|_2\big),$$

and when $p \geq 1$

$$\alpha = \big(1 + \|V^{(1)}(U^{(1)})^\dagger\|_2 + \|F\|_2\beta_1\big)\big(1 + \|V^{(1)}(U^{(1)})^\dagger\|_2 + \|F\|_2\beta_2\big),$$
$$\beta_1 = \|G^\dagger\|_2 + \|H_1^\dagger\|_2\big(1 + \|V^{(2)}(I - \Lambda^{(2)})\|_2\|G^\dagger\|_2\big),$$
$$\beta_2 = \|G^\dagger\|_2 + \|H_2^\dagger\|_2\big(1 + \|V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)}\|_2\|G^\dagger\|_2 + \|V^{(1)}(U^{(1)})^\dagger\|_2\big).$$

Here, $F = V^{(2)}\Lambda^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)}$, $G = P_{U^{(1)}}U_2$, $H_1 = V^{(2)}(I - \Lambda^{(2)})Q_G$, $H_2 = (V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)})Q_G$ *and* $S = V^{(2)}Q_{U^{(2)}}(V^{(2)})^*$.

Proof. It is clear from (2.7) that $X_U = X_L + K$, where

$$K = \begin{bmatrix} 0 & 0 \\ V^{(1)} & V^{(2)}\Lambda^{(2)} \end{bmatrix}.$$

From Ipsen [12, Corollary 3.3], we have that

$$\frac{\sigma_{max}(X_L)}{\sigma_{max}(X_U)} \leq 1 + \|KX_U^{-1}\|_2 \text{ and } \frac{\sigma_{min}(X_U)}{\sigma_{min}(X_L)} \leq 1 + \|KX_L^{-1}\|_2, \tag{2.9}$$

where $\sigma_{max}(X) = \|X\|_2$ and $\sigma_{min}(X) = 1/\|X^{-1}\|_2$ are the largest and smallest singular values of X. Combining these results gives

$$\frac{\kappa_2(X_L)}{\kappa_2(X_U)} \leq (1 + \|KX_U^{-1}\|_2)(1 + \|KX_L^{-1}\|_2). \tag{2.10}$$

We now obtain expressions for KX_L^{-1} and KX_U^{-1}, starting with the case $p = 0$. Let $X_{L,U}$ denote X_L or X_U as appropriate and let $\tilde{V}^{(2)} = V^{(2)}$ for X_U and $\tilde{V}^{(2)} = V^{(2)}(I - \Lambda^{(2)})$ for X_L. Then the inverse of $X_{L,U}$ is $X_{L,U}^{-1} = X_{L,U}^*(X_{L,U}X_{L,U}^*)^{-1}$, where

$$X_{L,U}X_{L,U}^* = \begin{bmatrix} I_n & \\ \tilde{V}^{(2)}(U^{(2)})^\dagger & I_m \end{bmatrix} \begin{bmatrix} U^{(2)}(U^{(2)})^* & \\ & S \end{bmatrix} \begin{bmatrix} I_n & (\tilde{V}(U^{(2)})^\dagger)^* \\ & I_m \end{bmatrix}$$

with $S = \tilde{V}^{(2)} Q_{U^{(2)}} (\tilde{V}^{(2)})^*$. Since $C = 0$, $\mathcal{P}_U^{-1} \mathcal{A} X_U = X_U \Lambda$ implies that $V^{(2)}\Lambda^{(2)} = P_S^{-1} B U^{(2)}$. Thus, $V^{(2)}\Lambda^{(2)} Q_{U^{(2)}} = 0$ and $S = V^{(2)} Q_{U^{(2)}} (V^{(2)})^*$. Straightforward calculation then yields

$$KX_{L,U}^{-1} = \begin{bmatrix} 0 \\ V^{(2)}\Lambda^{(2)} \end{bmatrix} \begin{bmatrix} (U^{(2)})^\dagger & Q_{U^{(2)}}(V^{(2)})^* S^{-1} \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -\tilde{V}^{(2)}(U^{(2)})^\dagger & I_m \end{bmatrix}$$

and the upper bounds are obtained by bounding $\|KX_L^{-1}\|_2$ and $\|KX_U^{-1}\|_2$. When $p \geq 1$ we use Theorem 2.1 in Tian and Takane [21]. Since $U^{(1)}$ has linearly independent columns, $Q_{U^{(1)}} = 0$, and

$$\begin{aligned} KX_L^{-1} &= \begin{bmatrix} 0 & 0 \\ V^{(1)} & V^{(2)}\Lambda^{(2)} \end{bmatrix} \begin{bmatrix} I_p & -(U^{(1)})^\dagger U^{(2)} \\ 0 & I_{n+m-p} \end{bmatrix} \begin{bmatrix} (U^{(1)})^\dagger & 0 \\ (I - H_1^\dagger V^{(2)}(I - \Lambda^{(2)}))G^\dagger & H_1^\dagger \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ V^{(1)}(U^{(1)})^\dagger + F(I - H_1^\dagger V^{(2)}(I - \Lambda^{(2)}))G^\dagger & FH_1^\dagger \end{bmatrix}. \end{aligned}$$

Similarly, letting $N = V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)}$,

$$\begin{aligned} KX_U^{-1} &= \begin{bmatrix} 0 & 0 \\ V^{(1)} & V^{(2)}\Lambda^{(2)} \end{bmatrix} \begin{bmatrix} I_p & -(U^{(1)})^\dagger U^{(2)} \\ 0 & I_{n+m-p} \end{bmatrix} \begin{bmatrix} (U^{(1)})^\dagger & 0 \\ (I - H_2^\dagger N)G^\dagger & H_2^\dagger \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -V^{(1)}(U^{(1)})^\dagger & I_m \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ V^{(1)}(U^{(1)})^\dagger + F[(I - H_2^\dagger(V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)}))G^\dagger - H_2^\dagger V^{(1)}(U^{(1)})^\dagger] & FH_2^\dagger \end{bmatrix}. \end{aligned}$$

The upper bound on the condition number again follows from bounding the norms. Then, the lower bound on the condition number is achieved by bounding $\kappa_2(X_U)/\kappa_2(X_L)$ from above. □

Although the expressions in Corollary 4 are quite complicated, they highlight that the difference between the condition numbers depends not only on the matrices that vary between X_L and X_U, namely $V^{(1)}$, $V^{(2)}$ and $V^{(2)}(I - \Lambda^{(2)})$, but also on the conditioning of $U^{(1)}$ and $U^{(2)}$. This is not surprising since $U^{(1)}$ and $U^{(2)}$ affect the conditioning of X_L and X_U.

More specifically, when $p = 0$, α is smaller when $\|(U^{(2)})^\dagger\|_2$ is small, i.e., when $U^{(2)}$ is well conditioned in the sense that its smallest (nonzero) singular value is not too small, and when the rows of $V^{(2)}$ are almost orthogonal to the rows of $U^{(2)}$, so that $\|V^{(2)}(U^{(2)})^\dagger\|_2$ is small. Also, since $Q_{U^{(2)}}(V^{(2)})^* S^{-1}$ is a right inverse of $V^{(2)}$ we expect its norm to be large when the rows of $V^{(2)}$ are almost linearly dependent. This confirms that well conditioned X_L and X_U and a small perturbation $V^{(2)}\Lambda^{(2)}$ ensure that $\kappa_2(X_L)$ and $\kappa_2(X_U)$ are close.

If $p > 0$, α depends on $\|V^{(1)}(U^{(1)})^\dagger\|_2$, $\|F\|_2$, β_1 and β_2. The term $\|V^{(1)}(U^{(1)})^\dagger\|_2$ is small when $\|V^{(1)}\|_2$ is small but the smallest singular value of $U^{(1)}$ is not. Additionally, $\|F\|_2$ is small when $\|V^{(2)}\Lambda^{(2)}\|_2$ and $\|V^{(1)}\|_2$ are small, and the columns of $U^{(1)}$ are almost orthogonal to those of $U^{(2)}$. If $\text{range}(U^{(2)}) \subset \text{range}(U^{(1)})$, then $G = 0$, $G^\dagger = 0$, $Q_G = I$ and a sufficient condition for β_1 and β_2 to be small is that $\|V^{(2)}(I - \Lambda^{(2)})\|_2$, $\|V^{(2)}\|_2$ and $\|V^{(1)}\|_2$ are small. Otherwise, $\|G^\dagger\|_2$ is not too large if the columns of $U^{(1)}$ are orthogonal to those of $U^{(2)}$ with $U^{(2)}$ well conditioned. The terms H_1 and H_2 are the most difficult to analyse. Although $H_1 = 0$ when $\text{range}((V^{(2)}(I - \lambda^{(2)}))^*) \subset \text{range}(G^*)$ and $H_2 = 0$ when $\text{range}((V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)})^*) \subset \text{range}(G^*)$, these conditions may not hold in general. Alternatively, if $V^{(1)}(U^{(1)})^\dagger U^{(2)}$ is small, a condition for H_1 and H_2 to be well conditioned is that $V^{(2)}(I - \Lambda^{(2)})$ and $V^{(2)}$ are well conditioned on $\text{null}(U^{(2)})$. Considering all these

TABLE 3.1
Left preconditioned GMRES iterations for the Stokes problem.

Grid	IC		AMG	
	$\mathcal{P}_L$	$\mathcal{P}_U$	$\mathcal{P}_L$	$\mathcal{P}_U$
4×12	132	135	65	65
8×24	194	202	85	85
16×48	313	323	89	90
32×96	553	590	91	93

conditions together, we see that again α is small when $V^{(1)}$ and $V^{(2)}\Lambda^{(2)}$ are small in norm and X_L and X_U are well conditioned. We note that additional special cases might also result in well conditioned matrices X_L and X_U and that Corollary 4 may be useful for checking these.

Of course, neither X_U and X_L, nor their condition numbers, are uniquely defined (see, for example, the discussion in Bai, Benzi and Chen [2, Remark 3.1]). In our experiment we consider the eigenvector matrices for which each eigenvector has unit norm. To this end let us fix the columns $[u_i^T, v_i^T]^T$, $i = 1, \dots, n+m$, of X_U in (2.7) to have unit length, so that $\|u_i\|_2^2 + \|v_i\|_2^2 = 1$. Then, the column scaling that transforms X_L to $\hat{X}_L$, an eigenvector matrix with unit-length columns, is given by the diagonal matrix $D = \operatorname{diag}(d_{11}, \dots, d_{n+m,n+m})$, where

$$d_{ii}^{-2} = \begin{cases} 1 - \|v_i\|^2, & i \le p, \\ 1 + (|\lambda_i|^2 - 2\Re(\lambda_i))\|v_i\|^2, & i > p. \end{cases} \tag{2.11}$$

3. Numerical example. In our experience with incompressible Stokes and Navier-Stokes examples and preconditioners in IFISS [9, 19] and with time-harmonic Maxwell equations [11] we find, similarly to [4, 7], that there is often little difference between iteration counts achieved with $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$. This seems to be true regardless of whether left or right preconditioning is used. However, there are certainly examples for which the condition numbers differ as we now show.

Our linear system comes from an incompressible Stokes problem that describes a flow over a backward facing step in two dimensions and is described in detail in Elman et al. [10, Example 5.1.2]. The equations are discretized by $Q_2 - Q_1$ finite elements in Matlab using IFISS with default parameters. We apply left-preconditioned GMRES with a zero initial guess and terminate when the preconditioned residual decreases by eight orders of magnitude—although it may be desirable to consider the unpreconditioned residual, the preconditioned residual is more closely connected with the theory of Section 2. In both of our preconditioners P_S is the diagonal of the pressure mass matrix. The first choice for P_A is a no-fill incomplete Cholesky (IC) factorization (produced by the Matlab command `ichol`) of A_0, the vector Laplace matrix obtained with natural boundary conditions. The second is the algebraic multigrid (AMG) preconditioner implemented in IFISS. The eigenvector matrix X_U is computed by the Matlab function `eig` while X_L is computed from X_U using Corollary 3. Both X_L and X_U are scaled to have unit-length vectors.

We first consider the incomplete Cholesky preconditioner. The eigenvalues of $\mathcal{P}_L^{-1}\mathcal{A}$ (and $\mathcal{P}_U^{-1}\mathcal{A}$) lie in $[-2.3, -0.046] \cup [0.19, 1.2]$. Additionally, there are 20 eigenvalues within 10^{-14} of 1 that we assume are unit eigenvalues. The iterations for different mesh sizes are given in Table 3.1, from which we see that the iteration counts are consistently lower for $\mathcal{P}_L$ than $\mathcal{P}_U$. For the problem on the 8×24 grid this

TABLE 3.2
Condition numbers of eigenvector matrices and norms of quantities in Corollary 4 for the problem on the 8×24 grid and the incomplete Cholesky preconditioner.

$\kappa_2(X_L)$	$\kappa_2(X_U)$	$\|V^{(1)}\|_2$	$\|V^{(2)}\Lambda^{(2)}\|_2$	$\|V^{(1)}(U^{(1)})^\dagger\|_2$	$\|F\|_2$
23	54	2.1	6.4	10.6	6.5
$\|G^\dagger\|_2$	$\|H_1^\dagger\|_2$	$\|H_2^\dagger\|_2$	$\|V^{(2)}(I-\Lambda^{(2)})\|_2$	$\|V^{(2)} - V^{(1)}(U^{(1)})^\dagger U^{(2)}\|_2$	α
6.9	3.1	3.1	4.7	6.6	$\alpha = 8.8 \times 10^6$

TABLE 3.3
Condition numbers of eigenvector matrices and norms of quantities in Corollary 4 for the problem on the 8×24 grid and the algebraic multigrid preconditioner.

$\kappa_2(X_L)$	$\kappa_2(X_U)$	$\|V^{(2)}\Lambda^{(2)}\|_2$	$\|V^{(2)}(U^{(2)})^\dagger\|_2$	$\|V^{(2)}(I-\Lambda^{(2)})(U^{(2)})^\dagger\|_2$
265	59	13.5	11.7	14.1
$\|S\|_2$	α			
1.6	$\alpha = 1.7 \times 10^7$			

is reflected in the condition numbers of X_L and X_U (see Table 3.2). To further investigate the disparity in these condition numbers, we also list in Table 3.2 quantities related to Corollary 4. Since $\|X_L\|_2 = 2.9$ and $\|X_U\|_2 = 5.7$, the perturbations $V^{(1)}$ and $V^{(2)}\Lambda^{(2)}$ are relatively large in norm. Moreover, $\|(U^{(1)})^\dagger\|_2 = 9.1$, so that U_1 is not so well conditioned, and as a consequence $\|V^{(1)}(U^{(1)})^\dagger\|_2$ is large. Relative to $\|V^{(1)}(U^{(1)})^\dagger\|_2$, the remaining terms in α are reasonably small and we conclude that the most significant contributions to the difference between $\kappa_2(X_L)$ and $\kappa_2(X_U)$ are $U^{(1)}$ and the perturbations $V^{(1)}$ and $V^{(2)}\Lambda^{(2)}$. We note that although the bounds in Corollary 4 are not quantitatively descriptive for this problem they give insight into why the condition numbers of X_L and X_U differ.

The eigenvalues of the AMG-preconditioned matrix lie in $[-2.5, -0.01] \cup [0.009, 58]$ and no eigenvalue is within 10^{-14} of 1. The iteration counts appear to be mesh-independent and are much lower than for the incomplete Cholesky preconditioner in spite of the wider distribution of eigenvalues (see Table 3.1). Additionally, $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{P}_L^{-1}\mathcal{A}$ give similar iteration counts, with the latter performing slightly better for larger problems. Thus, the eigenvector condition numbers are not necessarily good predictors of convergence for this problem since, at least for the 8×24 grid, $\kappa_2(X_L) > \kappa_2(X_U)$. Nevertheless, we can investigate why the condition numbers differ by applying Corollary 4. Compared to $\|X_L\|_2 = 4.5$ and $\|X_U\|_2 = 2.9$, $\|V^{(2)}\Lambda^{(2)}\|_2$ and $\|(U^{(2)})^\dagger\|_2 = 20.6$ are large. Consequently, $\|V^{(2)}(U^{(2)})^\dagger\|_2$ and $\|V^{(2)}(I-\Lambda^{(2)})(U^{(2)})^\dagger\|_2$ are large relative to the other terms in α. From this we deduce that the norm of $(U^{(2)})^\dagger$ and the size of the perturbation $V^{(2)}\Lambda^{(2)}$ are the main causes of the difference between $\kappa_2(X_L)$ and $\kappa_2(X_U)$.

For this problem the condition number bound (2.8) overestimates $\kappa_2(X_U)/\kappa_2(X_L)$. We note that there are other examples for which (2.8) is tight—these are typically problems for which $\kappa_2(X_U)$ and $\kappa_2(X_L)$ are close. When the difference between the condition numbers increases, the bound (2.8) is usually not as tight, with (2.10) overestimating $\kappa_2(X_U)/\kappa_2(X_L)$ and the subsequent bounds on $\|KX_U^{-1}\|_2$ and $\|KX_L^{-1}\|_2$ then causing (2.10) to be overestimated.

4. Conclusion. In summary, when the eigenvalues are important, or when the eigenvector matrices X_L and X_U are fairly well conditioned, there is no benefit in choosing $\mathcal{P}_L$ over $\mathcal{P}_U$ or vice versa. (Note that since the eigenvector matrix is not uniquely defined, the conditioning could be related to the particular choice of matrix.) We have shown that the eigenvectors of $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ can differ significantly, and this can affect the condition numbers of the eigenvector matrices (when $\mathcal{P}_U^{-1}\mathcal{A}$ and $\mathcal{P}_L^{-1}\mathcal{A}$ are diagonalizable). However, the convergence rate of GMRES remains weakly sensitive to these factors for the problems examined, and for others in the the literature. When $C = 0$ and $\mathcal{P}_L^{-1}\mathcal{A}$ and $\mathcal{P}_U^{-1}\mathcal{A}$ are diagonalizable, we can bound the ratio of the condition numbers of the eigenvector matrices X_L and X_U, which depend not only on difference between X_L and X_U, contained in $V^{(1)}$ and $V^{(2)}\Lambda^{(2)}$, but also on $U^{(1)}$ and $U^{(2)}$.

Acknowledgments. The author would like to thank the editor and referees, who offered many helpful comments and suggestions, and Yvan Notay, whose manuscript on symmetric saddle point matrices inspired this work.

REFERENCES

[1] Z.-Z. Bai, *Structured preconditioners for nonsingular matrices of block two-by-two structures*, Math. Comp., 75 (2006), pp. 791–815.

[2] Z.-Z. Bai, M. Benzi, and F. Chen, *On preconditioned MHSS iteration methods for complex symmetric linear systems*, Numer. Algorithms, 56 (2011), pp. 297–317.

[3] Z.-Z. Bai and M. K. Ng, *On inexact preconditioners for nonsymmetric matrices*, SIAM J. Sci. Comput., 26 (2005), pp. 1710–1724.

[4] Z.-Z. Bai and Z.-R. Ren, *Block-triangular preconditioning methods for linear third-order ordinary differential equations based on reduced-order sinc discretizations*, Japan J. Indust. Appl. Math., 30 (2013), pp. 511–527.

[5] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.

[6] M. Benzi, M. A. Olshanskii, and Z. Wang, *Modified augmented Lagrangian preconditioners for the incompressible Navier-Stokes equations*, Int. J. Numer. Meth. Fluids, 66 (2011), pp. 486–508.

[7] M. Benzi and B. Uçar, *Block triangular preconditioners for M-matrices and Markov chains*, ETNA, 26 (2007), pp. 209–227.

[8] J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50 (1988), pp. 1–17.

[9] H. C. Elman, A. Ramage, and D. J. Silvester, *Algorithm 866: IFISS, a Matlab toolbox for modelling incompressible flow*, ACM Trans. Math. Software, 33 (2007). Article 14.

[10] 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.

[11] C. Greif and D. Schötzau, *Preconditioners for saddle point linear systems with highly singular* $(1,1)$ *blocks*, Electron. Trans. Numer. Anal., 22 (2006), pp. 114–121.

[12] I. C. F. Ipsen, *Relative perturbation results for matrix eigenvalues and singular values*, Acta Numer., 7 (1998), pp. 151–201.

[13] ———, *A note on preconditioning nonsymmetric matrices*, SIAM J. Sci. Comput., 23 (2001), pp. 1050–1051.

[14] A. Klawonn, *Block-triangular preconditioners for saddle point problems with a penalty term*, SIAM J. Sci. Comput., 19 (1998), pp. 172–184.

[15] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972.

[16] Y. Notay, *A new analysis of block preconditioners for saddle point problems*, SIAM J. Matrix Anal. Appl., To appear (2013).

[17] J. Pestana and A. J. Wathen, *Combination preconditioning of saddle point systems for positive definiteness*, Numer. Linear Algebra Appl., 20 (2013), pp. 785–808.

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

[19] D. J. Silvester, H. C. Elman, and A. Ramage, *Incompressible Flow and Iterative Solver Software (IFISS) version 3.1*, January 2011. `http://www.manchester.ac.uk/ifiss/`.

[20] V. Simoncini, *Block triangular preconditioners for symmetric saddle-point systems*, Appl. Numer. Math, 49 (2004), pp. 63–80.

[21] Y. Tian and Y. Takane, *The inverse of any two-by-two monsingular partitioned matrix and three matrix inverse completion problems*, Comput. Math. Appl., 57 (2009), pp. 1294–1304.

[22] H. A. van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for*

the solution of nonsymmetric linear systems, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.