SIAM J. Matrix Anal. Appl.
Vol. 35, No. 1, pp. 143–173

A NEW ANALYSIS OF BLOCK PRECONDITIONERS FOR SADDLE POINT PROBLEMS[*]

YVAN NOTAY[†]

Abstract. We consider symmetric saddle point matrices. We analyze block preconditioners based on the knowledge of a good approximation for both the top left block and the Schur complement resulting from its elimination. We obtain bounds on the eigenvalues of the preconditioned matrix that depend only of the quality of these approximations, as measured by the related condition numbers. Our analysis applies to indefinite block diagonal preconditioners, block triangular preconditioners, inexact Uzawa preconditioners, block approximate factorization preconditioners, and a further enhancement of these preconditioners based on symmetric block Gauss–Seidel-type iterations. The analysis is unified and allows the comparison of these different approaches. In particular, it reveals that block triangular and inexact Uzawa preconditioners lead to identical eigenvalue distributions. These theoretical results are illustrated on the discrete Stokes problem. It turns out that the provided bounds allow one to localize accurately both real and nonreal eigenvalues. The relative quality of the different types of preconditioners is also as expected from the theory.

Key words. saddle point, preconditioning, Uzawa method, block triangular, SIMPLE, convergence analysis, linear systems, Stokes problem, PDE-constrained optimization

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

DOI. 10.1137/130911962

1. Introduction.

We consider linear systems $K\,\mathbf{u} = \mathbf{b}$ for which the system matrix K has the following saddle point structure:

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

where A is an $n \times n$ symmetric and positive definite matrix (SPD) and where C is an $m \times m$ nonnegative definite matrix. We also assume that $m \leq n$ and that B has full rank, or that C is positive definite on the null space of B^T (the case of rank deficient B with $C = 0$ is treated in the appendix).

These assumptions entail that the system is nonsingular; see, e.g., [4]. We also refer to this work for an overview of the many applications in which such linear systems arise, as well as a general introduction to the different solution methods.

Our focus in this paper is on an important class of preconditioning techniques that exploit the knowledge of a good preconditioner M_A for A, and of a good preconditioner M_S for the (negative) Schur complement

$$S = C + B\,A^{-1}B^T \; . \tag{1.2}$$

Since both A and S are SPD, we assume that M_A and M_S are SPD as well. Techniques for obtaining such preconditioners are often application dependent; see, again, [4] for examples and pointers to the literature. Here we disregard "internal" details of

[*] Received by the editors March 5, 2013; accepted for publication (in revised form) by M. Benzi September 11, 2013; published electronically February 6, 2014.
http://www.siam.org/journals/simax/35-1/91196.html

[†] Service de Métrologie Nucléaire (C.P. 165-84), Université Libre de Bruxelles, B-1050 Brussels, Belgium (ynotay@ulb.ac.be, http://homepages.ulb.ac.be/~ynotay). Yvan Notay is Research Director of the Fonds de la Recherche Scientifique – FNRS.

these preconditioners and develop an analysis of preconditioning schemes for K that depends only on the extremal eigenvalues

$$\underline{\mu} = \lambda_{\min}\left(M_A^{-1}A\right) , \qquad \overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right) , \tag{1.3}$$

$$\underline{\nu} = \lambda_{\min}\left(M_S^{-1}S\right) , \qquad \overline{\nu} = \lambda_{\max}\left(M_S^{-1}S\right) , \tag{1.4}$$

where $\lambda_{\min}(\cdot)$ (resp., $\lambda_{\max}(\cdot)$) stands for the smallest (largest) eigenvalue. Hence our results apply regardless of the application context as soon as estimates are available for these four parameters; see [15] and [33] for examples of derivation of such estimates in the contexts of Stokes and PDE-constrained optimization problems, respectively.

Our analysis applies to most "indefinite" preconditioners in 2 × 2 block form, whose indefiniteness is tailored to compensate for the indefiniteness of the system matrix, in the sense that the preconditioned matrix has only eigenvalues with positive real part. This includes indefinite block diagonal preconditioners

$$M_d = \begin{pmatrix} M_A & \\ & -M_S \end{pmatrix} , \tag{1.5}$$

block triangular preconditioners

$$M_t = \begin{pmatrix} M_A & B^T \\ & -M_S \end{pmatrix} , \tag{1.6}$$

inexact or preconditioned Uzawa preconditioners

$$M_u = \begin{pmatrix} M_A & \\ B & -M_S \end{pmatrix} , \tag{1.7}$$

block approximate factorization preconditioners

$$M_f = \begin{pmatrix} M_A & \\ B & -M_S \end{pmatrix} \begin{pmatrix} I & M_A^{-1}B^T \\ & I \end{pmatrix} , \tag{1.8}$$

and further enhancements of these preconditioners based on symmetric block Gauss–Seidel-type iterations; see section 2. Note that the SIMPLE preconditioner (e.g., [30, 43]) is a particular case of the block approximate factorization preconditioner as defined above; see also [14] for further related variants. These preconditioners are sometimes seen as symmetrized variants of block triangular or inexact Uzawa preconditioners. This framework also describes some multigrid smoothers based on "distributive relaxation"; see [4, section 11.1] for a discussion and further references.

When $M_A = A$ and $M_S = S$, it is known that all these preconditioners but M_d are such that the preconditioned matrix has all eigenvalues equal to 1 and minimal polynomial of degree at most 2 [4, 20]. On the other hand, with M_d, there are only three distinct eigenvalues when $C = 0$ [17]. However, using these "ideal" preconditioners requires exact solves with A and S, which is often impractical; just the computation of S can be prohibitive. Here we investigate the effect of using instead approximations M_A and M_S. We analyze how the eigenvalue distributions are affected by providing bounds, where "bounds," for nonreal eigenvalues, have to be understood as combinations of inequalities proving their clustering in a confined region of the complex plane.

There are very many works developing eigenvalue analyses for these types of preconditioners; see [5, 11] for block diagonal preconditioners, [39] for block triangular preconditioners, [7, 8, 16, 44] for inexact Uzawa preconditioners, and [2, 3, 44] for block approximate factorization preconditioners—to mention just a few. We refer the reader to [4] for many more references and historical remarks.

Nevertheless, as far as we know, our bounds are more accurate than previous ones, with the exception of some inequalities in [39] for nonreal eigenvalues, which, combined with ours, allow us to further restrict the area where the eigenvalues are confined. Moreover, our analysis is truly unified, and we show, seemingly for the first time, that block triangular and inexact Uzawa preconditioners lead to identical eigenvalue distributions. We also establish a clear connection between these inexact Uzawa and block triangular preconditioners and symmetrized preconditioners as in (1.8), allowing us to discriminate cases where this symmetrization can be useful and cases where it is likely not cost effective.

Some previous analyses focus on the conditions needed to have the preconditioned matrix positive definite in a nonstandard inner product, and develop related conjugate gradient–like methods; see, e.g., [7, 12, 32, 44]. Here we offer a complementary viewpoint, giving estimates that vary continuously in function of the main parameters (1.3), (1.4), without any restriction on these parameters. Moreover, whereas we reproduce the condition $\underline{\mu} \geq 1$ to have only real (and positive) eigenvalues with Uzawa [44] or block triangular [39] preconditioners, our analysis also reveals that scaling M_A to satisfy this condition often has an adverse effect on the clustering of the eigenvalues.

Note that there are several preconditioning techniques also based on approximations M_A and M_S that nevertheless do not fit with our analysis; this includes symmetric positive definite block diagonal preconditioners [2, 15, 38], which are popular because they can be combined with MINRES [29], thus avoiding the restarting associated with GMRES [35] or GCR [13, 42]. Leaving aside restarting effects, definite and indefinite block diagonal preconditioners are found in [17] to be essentially equivalent, which we further confirm independently by showing a general relation between the eigenvalues associated with both preconditioners.

Another approach that has connections with those investigated here is constraint preconditioning [19]:

$$\begin{pmatrix} M_A & B^T \\ B & -C \end{pmatrix} = \begin{pmatrix} M_A & \\ B & -C - B\,M_A^{-1}B^T \end{pmatrix} \begin{pmatrix} I & M_A^{-1}B^T \\ & I \end{pmatrix} .$$

In fact, this corresponds to block approximate preconditioning (1.8) with

$$M_S = C + B\,M_A^{-1}B^T .$$

Hence results in this paper can be applied to this preconditioner as well,[1] but specific analyses that exploit the particular form of M_S are likely more powerful; see, e.g., [4, section 10.2], [36], and the references therein. Our analysis may, however, be useful when $C + B\,M_A^{-1}B^T$ is replaced with something easier to invert (e.g., [6, 31]), the line between these inexact constraint preconditioners and block approximate factorization preconditioners being blurred.

The remainder of this paper is organized as follows. In section 2, we introduce some further variants of the preconditioners defined above. In section 3, we examine the relations that exist between the spectra associated with these different preconditioners, whereas, in section 4, we analyze the localization of the eigenvalues. These results are illustrated in section 5 on a typical example, namely the discrete Stokes

[1] For such M_S , there hold $\underline{\nu} \geq \overline{\mu}^{-1}$ and $\overline{\nu} \leq \underline{\mu}^{-1}$.

problem. Concluding remarks are given in section 6. Peculiarities associated with singular K are discussed in the appendix.

2. Further variants of block preconditioners. We first introduce a variant of the block approximate factorization preconditioners, which we call block SGS because of its close connection with block symmetric Gauss–Seidel iterations. Let

$$\widetilde{M}_A = M_A\,(2\,M_A - A)^{-1}\,M_A\ ; \tag{2.1}$$

$\widetilde{M}_A$ is in fact the preconditioner for A corresponding to the combination of two stationary iterations with M_A , as seen from the relation

$$I - \widetilde{M}_A^{-1}A = \left(I - M_A^{-1}A\right)^2\ . \tag{2.2}$$

The block SGS preconditioner is then algebraically defined by

$$M_g = \begin{pmatrix} I & \\ B\,M_A^{-1} & I \end{pmatrix} \begin{pmatrix} \widetilde{M}_A & \\ & -M_S \end{pmatrix} \begin{pmatrix} I & M_A^{-1}B^T \\ & I \end{pmatrix}\ . \tag{2.3}$$

The motivation is twofold. On the one hand, our analysis in the next section suggests that M_g can compare favorably with the block approximate factorization preconditioner (1.8). On the other hand, solving a system $M_g\mathbf{u} = \mathbf{r}$ requires only a slight modification of the algorithm that solves a system with M_f , and the extra cost is limited to one additional multiplication with A . Indeed, letting $\mathbf{u} = (\mathbf{u}_A\ ,\ \mathbf{u}_C)^T$ and $\mathbf{r} = (\mathbf{r}_A\ ,\ \mathbf{r}_C)^T$, both solves are implemented with[2]

1. $\mathbf{v}_A = M_A^{-1}\,\mathbf{r}_A\ ,$
2. $\mathbf{u}_C = M_S^{-1}\,(-\mathbf{r}_C + B\,\mathbf{v}_A)\ ,$
3. $\mathbf{u}_A = \begin{cases} \mathbf{v}_A + M_A^{-1}\,\left(-B^T\mathbf{u}_C\right) & \text{for } M_f\ , \\ 2\,\mathbf{v}_A + M_A^{-1}\,\left(-A\,\mathbf{v}_A - B^T\mathbf{u}_C\right) & \text{for } M_g\ . \end{cases}$

On the other hand, the other preconditioners can also be enhanced by using $\widetilde{M}_A$ instead of M_A , and, as will be seen, it is enlightening to explicitly include in our study the corresponding versions of block triangular and inexact Uzawa preconditioners, that is,

$$M_{t_2} = \begin{pmatrix} \widetilde{M}_A & B^T \\ & -M_S \end{pmatrix} \quad \text{and} \quad M_{u_2} = \begin{pmatrix} \widetilde{M}_A & \\ B & -M_S \end{pmatrix}\ . \tag{2.4}$$

In view of (2.1), these preconditioners represent at the algebraic level the operator used when either M_t or M_u is combined with an approximation of A^{-1} based on two stationary inner iterations with M_A . The computational cost associated with M_{t_2} and M_{u_2} is in fact the same as that associated with M_g , except that one multiplication by either B (case M_{t_2}) or B^T (case M_{u_2}) is saved.

[2] The equivalence between the algebraic definitions (1.8), (2.3) and this algorithm can be checked by observing that $\mathbf{v}_A$, $\mathbf{u}_C$, and $\mathbf{u}_A$ as defined in this algorithm satisfy $\left(\begin{smallmatrix} M_A & \\ B & -M_S \end{smallmatrix}\right)\left(\begin{smallmatrix} \mathbf{v}_A \\ \mathbf{u}_C \end{smallmatrix}\right) = \left(\begin{smallmatrix} \mathbf{r}_A \\ \mathbf{r}_C \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} F & \\ & I \end{smallmatrix}\right)\ \left(\begin{smallmatrix} I & M_A^{-1}B^T \\ & I \end{smallmatrix}\right)\left(\begin{smallmatrix} \mathbf{u}_A \\ \mathbf{u}_C \end{smallmatrix}\right) = \left(\begin{smallmatrix} \mathbf{v}_A \\ \mathbf{u}_C \end{smallmatrix}\right)$, with $F = I$ when M_f is used and, otherwise, with $F = (2\,I - M_A^{-1}A)^{-1} = M_A^{-1}\,\widetilde{M}_A$.

Of course, using either M_g or M_{t_2} , M_{u_2} makes sense only if $\widetilde{M}_A$ is positive definite. This holds if and only if $\overline{\mu} < 2$, where $\overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right)$ has already been defined in (1.3). This is also the necessary and sufficient condition for having $\rho_A < 1$, where

$$\rho_A = \rho\left(I - M_A^{-1}A\right) = \max\left(\overline{\mu} - 1\,,\, 1 - \underline{\mu}\right) \tag{2.5}$$

is the spectral radius of the iteration matrix associated with M_A .

3. Relations between the preconditioners. The following theorem highlights the connections that exist between the spectra associated with the different preconditioners. The proof of statement (3) uses an approach similar to that followed in the proof of Theorem 6 in [39], which analyzes the eigenvalues associated with block triangular preconditioners. This approach, based on a sequence of similarity transformations, is extended here to all preconditioners introduced in sections 1 and 2 and will further be used in the proof of Theorem 4.3.[3]

THEOREM 3.1. *Let*

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

be a matrix such that A is an $n \times n$ SPD matrix and C is an $m \times m$ symmetric nonnegative definite matrix with $m \leq n$. Assume that B has rank m or that C is positive definite on the null space of B^T. Let the preconditioners M_d , M_t , M_u , M_f , M_g , M_{t_2} , and M_{u_2} be defined as in, respectively, (1.5), (1.6), (1.7), (1.8), (2.3), *and* (2.4), *where M_A and M_S are SPD. Let ρ_A be defined by* (2.5), *and assume that $\rho_A < 1$ when one of M_f, M_g, M_{t_2}, or M_{u_2} is considered.*

(1) *Letting*

$$M_+ = \begin{pmatrix} M_A & \\ & M_S \end{pmatrix} , \tag{3.1}$$

the eigenvalues of $M_d^{-1}K$ and those of $M_+^{-1}K$ satisfy

$$\max_{\lambda \in \sigma(M_d^{-1}K)} |\lambda| \leq \max_{\lambda \in \sigma(M_+^{-1}K)} |\lambda| , \tag{3.2}$$

$$\min_{\lambda \in \sigma(M_d^{-1}K)} |\lambda| \geq \min_{\lambda \in \sigma(M_+^{-1}K)} |\lambda| . \tag{3.3}$$

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

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

Proof. The matrix $M_d^{-1}K$ has the same eigenvalues as $\widehat{K} = M_+^{1/2} M_d^{-1} K\, M_+^{-1/2}$. The largest of these eigenvalues in modulus is bounded above by the matrix norm $\|\widehat{K}\|_2$, which is also equal to the largest singular value of $\widehat{K}$ [40, Theorem 5.3], and thus is further equal to the square root of the largest eigenvalue of $\widehat{K}^T\widehat{K}$ [40, Theorem 5.4]. Let $\widetilde{K} = M_+^{-1/2} K\, M_+^{-1/2}$ and

$$J = \begin{pmatrix} I & \\ & -I \end{pmatrix} .$$

[3] In a preliminary draft of this paper, this approach was also used to prove statement (2); the much simpler argument given in the proof of Theorem 3.1 has been suggested independently by Artem Napov and two anonymous referees.

Because $\widehat{K} = J\widetilde{K}$, one has $\widehat{K}^T\widehat{K} = \widetilde{K}^T J^T J\widetilde{K} = \widetilde{K}^T\widetilde{K} = \widetilde{K}^2$, and the square root of the largest eigenvalue of $\widehat{K}^T\widehat{K}$ is also the largest eigenvalue in modulus of $\widetilde{K}$. This proves (3.2) since $M_+^{-1}K$ has the same eigenvalues as $\widetilde{K}$. The inequality (3.3) can be proved by applying the same reasoning to $K^{-1}M_d$, whose largest eigenvalue in modulus is the inverse of the smallest eigenvalue in modulus of $M_d^{-1}K$: $K^{-1}M_d$ has the same eigenvalues as $\widehat{K}^{-1}$, whose norm is bounded above by the square root of the largest eigenvalue in modulus of $\widehat{K}^{-1}\widehat{K}^{-T} = \widetilde{K}^{-2}$, i.e., the inverse of the smallest eigenvalue in modulus of $M_+^{-1}K$.

To prove statement (2), observe that $M_u^{-1}K$ has the same spectrum as $K\,M_u^{-1}$, which itself has the same spectrum as its transpose $(K\,M_u^{-1})^T = M_u^{-T}K = M_t^{-1}K$.

A similar argument shows that $M_{t_2}^{-1}K$ and $M_{u_2}^{-1}K$ also have the same spectrum. However, more involved developments are needed to prove that this common spectrum further coincides with the spectrum of $M_g^{-1}K$. These developments are also needed to prove Theorem 4.3 below. For this reason, we formulate them for all the preconditioners considered in this work, although only $M_g^{-1}K$, $M_{t_2}^{-1}K$, and $M_{u_2}^{-1}K$ are addressed by the remainder of this proof.

These developments require the assumption that there is no eigenvalue of $M_A^{-1}A$ that is exactly equal to 1. This is, however, no loss of generality because if there is such an eigenvalue, we can make the proof for a slightly perturbed matrix

$$K_\varepsilon = \begin{pmatrix} (1-\varepsilon)A & B^T \\ B & -C \end{pmatrix} \tag{3.4}$$

with $0 < \varepsilon < 1$. Then, since the eigenvalues continuously depend on ε, the needed results for the original matrix are obtained by considering the limit for $\varepsilon \to 0$.

Consider now the matrix

$$M = \begin{pmatrix} I_n & \\ B\,Y_A & I_m \end{pmatrix} \begin{pmatrix} \widehat{M}_A & \\ & -M_S \end{pmatrix} \begin{pmatrix} I_n & Z_A B^T \\ & I_m \end{pmatrix} .$$

Setting

$$Y_A = \begin{cases} 0 & \text{for } M_d\,,\ M_t\,,\ M_{t_2}\,, \\ M_A^{-1} & \text{for } M_u\,,\ M_f\,, M_g\,, \\ \widetilde{M}_A^{-1} & \text{for } M_{u_2}\,, \end{cases} \qquad Z_A = \begin{cases} 0 & \text{for } M_d\,,\ M_u\,,\ M_{u_2}\,, \\ M_A^{-1} & \text{for } M_t\,,\ M_f\,, M_g\,, \\ \widetilde{M}_A^{-1} & \text{for } M_{t_2}\,, \end{cases} \tag{3.5}$$

and

$$\widehat{M}_A = \begin{cases} M_A & \text{for } M_d\,,\ M_t\,,\ M_u\,,\ M_f\,, \\ \widetilde{M}_A & \text{for } M_g\,, M_{t_2}\,,\ M_{u_2} \end{cases} \tag{3.6}$$

(where $\widetilde{M}_A$ is defined in (2.1)), one sees that M can represent each of the preconditioners considered in this work.

Now let X and Λ be such that $X^T A^{1/2} M_A^{-1} A^{1/2} X = \Lambda$, with Λ diagonal and $X^T X = I$. Observe that

$$\begin{aligned} \Lambda_Y &= X^T A^{1/2} Y_A A^{1/2} X\ , \\ \Lambda_Z &= X^T A^{1/2} Z_A A^{1/2} X\ , \\ \widehat{\Lambda} &= X^T A^{1/2} \widehat{M}_A^{-1} A^{1/2} X \end{aligned}$$

are also diagonal and related to Λ via (3.5), (3.6). In particular, we have

$$\Gamma = \Lambda_Y + \Lambda_Z - \Lambda_Y \Lambda_Z = \begin{cases} 0 & \text{for } M_d\,, \\ \Lambda & \text{for } M_t\,,\ M_u\,, \\ 2\,\Lambda - \Lambda^2 & \text{for } M_f\,, M_g\,,\ M_{t_2}\,,\ M_{u_2}\,, \end{cases} \tag{3.7}$$

$$\widehat{\Lambda} = \begin{cases} \Lambda & \text{for } M_d\,,\ M_t\,,\ M_u\,,\ M_f\,, \\ 2\,\Lambda - \Lambda^2 & \text{for } M_g\,, M_{t_2}\,,\ M_{u_2}\ . \end{cases} \tag{3.8}$$

We then consider the preconditioned matrix $M^{-1}K$. It has the same eigenvalues as

$$\begin{aligned} &\begin{pmatrix} \widehat{M}_A^{-1} & \\ & -M_S^{-1} \end{pmatrix} \begin{pmatrix} I_n & \\ -B\,Y_A & I_m \end{pmatrix} \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} I_n & -Z_A\,B^T \\ & I_m \end{pmatrix} \\ &= \begin{pmatrix} \widehat{M}_A^{-1} & \\ & M_S^{-1} \end{pmatrix} \begin{pmatrix} A & (I - A\,Z_A)\,B^T \\ -B\,(I - Y_A\,A) & C + B\,(Y_A + Z_A - Y_A\,A\,Z_A)\,B^T \end{pmatrix} . \end{aligned} \tag{3.9}$$

The last matrix in (3.9) is similar to

$$\begin{aligned} &\begin{pmatrix} X^T A^{1/2} & \\ & M_S^{1/2} \end{pmatrix} \begin{pmatrix} \widehat{M}_A^{-1} & \\ & M_S^{-1} \end{pmatrix} \begin{pmatrix} A^{1/2} X & \\ & M_S^{1/2} \end{pmatrix} \begin{pmatrix} X^T A^{-1/2} & \\ & M_S^{-1/2} \end{pmatrix} \\ &\begin{pmatrix} A & (I - A\,Z_A)\,B^T \\ -B\,(I - Y_A\,A) & C + B\,(Y_A + Z_A - Y_A\,A\,Z_A)\,B^T \end{pmatrix} \begin{pmatrix} A^{-1/2} X & \\ & M_S^{-1/2} \end{pmatrix} \\ &= \begin{pmatrix} \widehat{\Lambda} & \\ & I \end{pmatrix} \begin{pmatrix} I & (I - \Lambda_Z)\,G^T \\ -G\,(I - \Lambda_Y) & \widetilde{C} + G\,(\Lambda_Y + \Lambda_Z - \Lambda_Y \Lambda_Z)\,G^T \end{pmatrix} , \end{aligned}$$

where we have set

$$\widetilde{C} = M_S^{-1/2} C\, M_S^{-1/2} \qquad \text{and} \qquad G = M_S^{-1/2} B\, A^{-1/2} X\ . \tag{3.10}$$

Now let Δ_+ , Δ_- be nonnegative diagonal matrices such that, for all $1 \le i \le n$,

$$\left(\Delta_+^2\right)_{ii} = \max\left((I - \Gamma)_{ii}\ , 0\right)\ , \qquad \left(\Delta_-^2\right)_{ii} = \max\left((\Gamma - I)_{ii}\ , 0\right)\ ,$$

where Γ is defined in (3.7). Note that this implies

$$\Delta_+^2 - \Delta_-^2 = I - \Gamma = (I - \Lambda_Y)\,(I - \Lambda_Z)\ .$$

On the other hand, our assumption that $M_A^{-1} A$ has no eigenvalue equal to 1 implies that $I - \Lambda_Y$ and $I - \Lambda_Z$ are nonsingular. Further, $\widehat{\Lambda}^{-1/2}$ exists because all entries in $\widehat{\Lambda}$ are positive; see (3.8), remembering that Λ is the diagonal matrix with the eigenvalues of $M_A^{-1} A$ on its diagonal, which are less than 2 by assumption if M_f , M_g , M_{t_2} , or M_{u_2} is considered. Hence the preconditioned matrix $M^{-1}K$ is also similar to

$$\begin{aligned} &\begin{pmatrix} (\Delta_+ + \Delta_-)\,\widehat{\Lambda}^{-1/2}\,(I - \Lambda_Z)^{-1} & \\ & I \end{pmatrix} \begin{pmatrix} \widehat{\Lambda} & \\ & I \end{pmatrix} \\ &\qquad \begin{pmatrix} I & (I - \Lambda_Z)\,G^T \\ -G\,(I - \Lambda_Y) & \widetilde{C} + G\,\Gamma G^T \end{pmatrix} \begin{pmatrix} (I - \Lambda_Y)^{-1}\,\widehat{\Lambda}^{1/2}\,(\Delta_+ - \Delta_-) & \\ & I \end{pmatrix} \\ &= \begin{pmatrix} \widehat{\Lambda} & \widehat{\Lambda}^{1/2}\,(\Delta_+ + \Delta_-)\,G^T \\ -G\,(\Delta_+ - \Delta_-)\,\widehat{\Lambda}^{1/2} & \widetilde{C} + G\,\Gamma G^T \end{pmatrix} . \end{aligned} \tag{3.11}$$

Interestingly, the matrix (3.11) resulting from the similarity transformations is the same for all preconditioners that share the same $\widehat{\Lambda}$ and Γ (hence also the same Δ_+ and Δ_-). In view of (3.7), (3.8), this concludes the proof of statement (3). □

Item (1) proves that the eigenvalue distribution associated with the positive definite block diagonal preconditioner M_+ cannot be qualitatively better than that associated with M_d. A tighter connection between both preconditioners is highlighted in [17], under the restrictive assumption that $M_A^{-1}A$ is a multiple of the identity. See also section 5 for a further comparison of both preconditioners.

On the other hand, block triangular and inexact Uzawa preconditioners are both well-established techniques that until now have been analyzed independently of each other. In item (2), we prove that they lead to identical eigenvalue distributions; hence eigenvalue bounds proved for the former are valid for the latter and vice versa.

Finally, the relation between the block SGS preconditioner and M_{t_2}, M_{u_2} seems less important. However, recall that M_{t_2} and M_{u_2} are just M_t and M_u in which one uses a closer approximation for A, based on two stationary iterations with M_A. Item (3) of Theorem 3.1 shows that using the symmetrized preconditioner M_g produces exactly the same effect, at least where the eigenvalue distribution is concerned. When it could be more interesting to use M_g instead of M_t or M_u is discussed at the end of section 4 and in section 5.

4. Eigenvalue analysis. The matrix (3.11) obtained at the end of the proof of Theorem 3.1 suggests that, at least in some cases ($\Delta_- = 0$), the eigenvalue analysis can be reduced to that of a matrix of the form

$$\widehat{K} = \begin{pmatrix} \widehat{A} & \widehat{B}^T \\ -\widehat{B} & \widehat{C} \end{pmatrix} , \tag{4.1}$$

where $\widehat{A}$ is SPD and $\widehat{C}$ is symmetric nonnegative definite. In fact, we shall see that this is true in all cases.

Such matrices are nonnegative definite in $\mathbb{R}^n$. Hence (see [5]), their eigenvalues have positive real part. Thus, if the preconditioned matrix is similar to a matrix of the form (4.1), one has gotten rid of the indefiniteness of the original matrix (1.1). Note, however, that this is at the expense of the loss of the symmetry, meaning that a portion of the eigenvalues will be in general complex.

Of course, one does not need the preconditioners introduced in sections 1 and 2 to obtain a nonsymmetric but definite linear system. As noted in, e.g., [5], it suffices to rewrite the original system $K\,\mathbf{u} = \mathbf{b}$ multiplying both sides to the left by

$$J = \begin{pmatrix} I & \\ & -I \end{pmatrix} , \tag{4.2}$$

which can also be seen as a very basic form of the block diagonal preconditioner (1.5), with $M_A = I$ and $M_S = I$. However, doing so will in general not change the magnitude of the eigenvalues by much; see item (1) of Theorem 3.1. Hence, small eigenvalues remain, entailing slow convergence of the iterative methods.

The role of the preconditioners investigated here then appears more clearly: combine the basic transformation (4.2) that makes the preconditioned matrix similar to a definite one, with further effects that improve the clustering of the eigenvalues while moving them away from the origin of the complex plane.

Now, to assess these effects, we need to be able to localize accurately the eigenvalues of matrices of the form (4.1). Our main tool in this respect is Proposition 2.12

in [5], whose main results are recalled in Theorem 4.1 below; see (4.4), (4.5), and the upper bound in (4.3). However, on their own, these inequalities (and those in [5] not reproduced here) do not provide an accurate picture of the situation. In particular, they do not allow us to show that preconditioning can be successful in moving all eigenvalues away from the origin: the lower bound for real eigenvalues is $\min(\lambda_{\min}(\widehat{A}), \lambda_{\min}(\widehat{C}))$, which vanishes when $\widehat{C}=0$. But the inverses of matrices of the form (4.1) have similar saddle point structure. Hence further inequalities can be obtained by applying the same Proposition 2.12 of [5] to the inverse of the matrix at hand. This approach is exploited in Theorem 4.1, and leads to (4.6) and the lower bound in (4.3). Thus, Theorem 4.1 combines these "new" inequalities with the "original" ones, and it turns out that nothing more is needed to obtain a satisfactory localization of all the eigenvalues.

THEOREM 4.1. *Let $\widehat{K}$ be a matrix of the form* (4.1), *where $\widehat{A}$ is an $n\times n$ SPD matrix and $\widehat{C}$ is an $m\times m$ symmetric nonnegative definite matrix with $m\le n$. Assume that $\widehat{B}$ has rank m or that $\widehat{C}$ is positive definite on the null space of $\widehat{B}^T$. Let*

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

and, if $\widehat{C}$ is positive definite,

$$S_{\widehat{A}} = \widehat{A} + \widehat{B}^T\widehat{C}^{-1}\widehat{B} \; .$$

The real eigenvalues λ of $\widehat{K}$ satisfy

$$\min\left(\lambda_{\min}\left(\widehat{A}\right),\, \lambda_{\min}\left(S_{\widehat{C}}\right)\right) \le \lambda \le \max\left(\lambda_{\max}\left(\widehat{A}\right),\, \lambda_{\max}\left(\widehat{C}\right)\right) \;, \tag{4.3}$$

and the eigenvalues λ with nonzero imaginary part are such that

$$\tfrac{1}{2}\left(\lambda_{\min}\left(\widehat{A}\right) + \lambda_{\min}\left(\widehat{C}\right)\right) \le \Re e(\lambda) \le \tfrac{1}{2}\left(\lambda_{\max}\left(\widehat{A}\right) + \lambda_{\max}\left(\widehat{C}\right)\right) \;, \tag{4.4}$$

$$|\Im m(\lambda)| \le \left(\lambda_{\max}\left(\widehat{B}\,\widehat{B}^T\right)\right)^{1/2} \;, \tag{4.5}$$

and

$$|\lambda-\zeta| \le \zeta \;, \tag{4.6}$$

where

$$\zeta = \begin{cases} \dfrac{\lambda_{\max}(S_{\widehat{A}})\,\lambda_{\max}(S_{\widehat{C}})}{\lambda_{\max}(S_{\widehat{A}})+\lambda_{\max}(S_{\widehat{C}})} & \textit{if } \widehat{C} \textit{ is positive definite,} \\ \lambda_{\max}\left(S_{\widehat{C}}\right) & \textit{otherwise .} \end{cases} \tag{4.7}$$

Proof. Inequalities (4.4) and (4.5) and the upper bound in (4.3) just translate results from [5, Proposition 2.12] in our notation. To prove the remaining inequalities, we first consider the case where $\widehat{C}$ is positive definite. Let $\widetilde{K}=J\widehat{K}$, where J is defined by (4.2). Because $\widetilde{K}$ is symmetric, its inverse is symmetric. Hence, since principal submatrices in $\widetilde{K}^{-1}$ are equal to the inverse of the Schur complements in $\widetilde{K}$ [1, p. 93], one has

$$\widetilde{K}^{-1} = \begin{pmatrix} S_{\widehat{A}}^{-1} & W^T \\ W & -S_{\widehat{C}}^{-1} \end{pmatrix} \;,$$

where W need not be known explicitly to conduct the proof. Indeed, what matters is that

$$\widehat{K}^{-1} = \widetilde{K}^{-1} J = \begin{pmatrix} S_{\widehat{A}}^{-1} & -W^T \\ W & S_{\widehat{C}}^{-1} \end{pmatrix}$$

has a structure that allows us to apply again Proposition 2.12 in [5]. For the real eigenvalues, this yields straightforwardly the lower bound in (4.3), using $\lambda_{\min}\left(S_{\widehat{A}}\right) \geq \lambda_{\min}(\widehat{A})$. For the eigenvalues λ with nonzero imaginary part, this proves

$$\Re\mathrm{e}\left(\lambda^{-1}\right) \geq \frac{1}{2}\left(\lambda_{\min}\left(S_{\widehat{A}}^{-1}\right) + \lambda_{\min}\left(S_{\widehat{C}}^{-1}\right)\right) .$$

The inequality (4.6) then follows because, for any complex number λ and real positive number ζ, $|\lambda - \zeta| \leq \zeta$ holds if and only if $\Re\mathrm{e}\left(\lambda^{-1}\right) \geq (2\,\zeta)^{-1}$.

If C is only semidefinite, we use a continuity argument: we apply the results just proved to

$$\begin{pmatrix} \widehat{A} & \widehat{B}^T \\ -\widehat{B} & \widehat{C} + \varepsilon\, I \end{pmatrix}$$

with $\varepsilon > 0$. We then let $\varepsilon \to 0$. Using $\lambda_{\max}\left(S_{\widehat{C}}\right)$ as upper bound on ζ, all quantities involved in the inequalities vary continuously with ε, as well as the eigenvalues themselves, proving the required results. □

We are now ready to state Theorem 4.3, which contains our main results in this section. For some cases (M_t and M_u when $\overline{\mu} > 1$), we need to introduce additional parameters η and $\widetilde{\nu}$ that depend on the following function:

$$f\,(\overline{\mu}\,,\,\nu) = \tfrac{1}{2}\,(\nu+1)\;\left(1 + \left(1 - \frac{4\,\nu}{\overline{\mu}\,(\nu+1)^2}\right)^{1/2}\right) \qquad (\overline{\mu} \geq 1\,,\,\nu > 0)\ . \tag{4.8}$$

It is a good idea to know how this function behaves before reading Theorem 4.3. The following lemma is helpful in this respect.

Lemma 4.2. *Let $f\,(\overline{\mu}\,,\,\nu)$ be defined by (4.8). For any $\overline{\mu} \geq 1$ and $\nu > 0$, there holds*

(4.9)

$$\max(1\,,\,\nu) \leq f\,(\overline{\mu}\,,\,\nu) \leq \begin{cases} \tfrac{1}{2}\Big(\left(1+\overline{\mu}^{\,-1/2}\right)\,\max(1\,,\,\nu) + \left(1-\overline{\mu}^{\,-1/2}\right)\,\min(1\,,\,\nu) \\ \qquad\qquad\qquad\qquad + \left(1-\overline{\mu}^{\,-1}\right)^{1/2}\,(\nu+1)\Big)\ , \\ \nu + 1 - \frac{\nu}{\overline{\mu}\,(\nu+1)}\ . \end{cases}$$

Proof. For $\overline{\mu} \geq 1$, one has

$$\begin{aligned} |1-\nu| \;&=\; (1+\nu)\,\left(1 - \tfrac{4\,\nu}{(\nu+1)^2}\right)^{1/2} \\ &\leq (1+\nu)\,\left(1 - \tfrac{4\,\nu}{\overline{\mu}(\nu+1)^2}\right)^{1/2} \;=\; 2\,f\,(\overline{\mu}\,,\,\nu) - (\nu+1) \leq (1+\nu)\,\left(1 - \tfrac{2\,\nu}{\overline{\mu}(\nu+1)^2}\right)\ , \end{aligned}$$

TABLE 1
Definitions of $\underline{\xi}$, $\overline{\xi}$, $\underline{\chi}$, $\overline{\chi}$, δ , and ζ for the different preconditioners.

	$\underline{\xi}$	$\overline{\xi}$	$\underline{\chi}$	$\overline{\chi}$
M_d	$\min(\underline{\mu}\,,\,\underline{\nu})$	$\max(\overline{\mu}\,,\,\overline{\nu})$	$\frac{1}{2}\,\underline{\mu}$	$\frac{1}{2}(\overline{\mu}+\overline{\nu})$ $\left[\frac{\overline{\mu}}{2}\text{ if } C=0\right]$
$\left.\begin{array}{r}M_u\,,\,M_t\\(\overline{\mu}\le 1)\end{array}\right\}$	$\min(\underline{\mu}\,,\,\underline{\nu})$	$\max(1\,,\,\overline{\nu})$	$\frac{1}{2}(\underline{\mu}+\underline{\nu}\,\underline{\mu})$	$\frac{1}{2}(1+\overline{\nu})$
$\left.\begin{array}{r}M_u\,,\,M_t\\(\overline{\mu}> 1)\end{array}\right\}$	$\min\left(\underline{\mu}\,,\,\eta^{-1}\,\underline{\nu}\right)$	$\max(1\,,\,\widetilde{\nu})$	$\frac{1}{2}\left(\underline{\mu}+\underline{\nu}\,\min(\eta^{-1}\,,\,\underline{\mu})\right)$	$\frac{1}{2}(1+\widetilde{\nu})$
M_f	$\min(\underline{\mu}\,,\,\underline{\nu})$	$\max(\overline{\mu}\,,\,\overline{\nu})$	$\frac{1}{2}\left(\underline{\mu}+\underline{\nu}(1-\rho_A^2)\right)$	$\frac{1}{2}(\overline{\mu}+\overline{\nu})$
$\left.\begin{array}{r}M_g\\M_{u_2}\,,\,M_{t_2}\end{array}\right\}$	$\min\left(1-\rho_A^2\,,\,\underline{\nu}\right)$	$\max(1\,,\,\overline{\nu})$	$\frac{1}{2}(1+\underline{\nu})\left(1-\rho_A^2\right)$	$\frac{1}{2}(1+\overline{\nu})$

	δ^2	ζ
M_d	$\overline{\nu}\,\overline{\mu}$	$\overline{\nu}$
$\left.\begin{array}{r}M_u\,,\,M_t\\(\overline{\mu}\le 1)\end{array}\right\}$	$\begin{cases}\overline{\nu}\,\underline{\mu}(1-\underline{\mu}) & \text{if } \underline{\mu}>\frac{1}{2},\\ \frac{1}{4}\,\overline{\nu} & \text{otherwise}\end{cases}$	$\dfrac{\overline{\nu}}{1+\overline{\nu}}$
$\left.\begin{array}{r}M_u\,,\,M_t\\(\underline{\mu}<1<\overline{\mu})\end{array}\right\}$	$\begin{cases}\overline{\nu}\,\underline{\mu}(1-\underline{\mu}) & \text{if } \underline{\mu}>\frac{1}{2},\\ \frac{1}{4}\,\overline{\nu} & \text{otherwise}\end{cases}$	$\dfrac{\widetilde{\nu}}{1+\widetilde{\nu}}$
$\left.\begin{array}{r}M_u\,,\,M_t\\(1\le\underline{\mu})\end{array}\right\}$	0	(not applicable)
M_f	$\begin{cases}\overline{\nu}\,\max\left(\frac{4}{27}\,,\,\overline{\mu}\,(1-\overline{\mu})^2\right) & \text{if } \underline{\mu}<\frac{1}{3}<\overline{\mu},\\ \overline{\nu}\,\max\Big(\underline{\mu}\,(1-\underline{\mu})^2\,, & \\ \qquad\qquad \overline{\mu}\,(1-\overline{\mu})^2\Big) & \text{otherwise}\end{cases}$	$\dfrac{\overline{\nu}}{1+\overline{\nu}\,(2-\overline{\mu})}$
$\left.\begin{array}{r}M_g\\M_{u_2}\,,\,M_{t_2}\end{array}\right\}$	$\begin{cases}\overline{\nu}\,\rho_A^2(1-\rho_A^2) & \text{if } \rho_A^2<\frac{1}{2},\\ \frac{1}{4}\,\overline{\nu} & \text{otherwise}\end{cases}$	$\dfrac{\overline{\nu}}{1+\overline{\nu}}$

from which the lower bound and the bottom upper bound (4.9) are straightforwardly deduced. On the other hand, the top upper bound follows from

$$\begin{aligned}(1+\nu)\left(1-\tfrac{4\,\nu}{\overline{\mu}(\nu+1)^2}\right)^{1/2} &= \overline{\mu}^{\,-1/2}\,\left((\nu-1)^2+(\overline{\mu}-1)\,(\nu+1)^2\right)^{1/2}\\ &\le \overline{\mu}^{\,-1/2}\,\left(|\nu-1|+\sqrt{\overline{\mu}-1}\,(\nu+1)\right)\ . \qquad \square\end{aligned}$$

THEOREM 4.3. *Let the assumptions of Theorem* 3.1 *hold, and let* $\underline{\mu}$, $\overline{\mu}$, $\underline{\nu}$, *and* $\overline{\nu}$ *be defined by* (1.3), (1.4). *For each of the preconditioners, let* $\underline{\xi}$, $\overline{\xi}$, $\underline{\chi}$, $\overline{\chi}$, δ , *and* ζ *be defined as in Table* 1, *where, when* $\overline{\mu}>1$,

$$\eta = f\,(\overline{\mu}\,,\,\underline{\nu})\ , \tag{4.10}$$

$$\widetilde{\nu} = \overline{\mu}\ f\,(\overline{\mu}\,,\,\overline{\nu})\ \ , \tag{4.11}$$

with $f\,(\overline{\mu}\,,\,\overline{\nu})$ *being defined in* (4.8).

Letting $*$ *stand for* d , t , u , f , g , t_2 , *or* u_2 , *the real eigenvalues* λ *of* $M_*^{-1}K$ *satisfy*

$$\underline{\xi} \le \lambda \le \overline{\xi}\ , \tag{4.12}$$

whereas eigenvalues with nonzero imaginary part are possible only if $\delta > 0$, *in which case they satisfy*

$$\underline{\chi} \leq \Re e(\lambda) \leq \overline{\chi} \ , \tag{4.13}$$

$$|\Im m(\lambda)| \leq \delta \ , \tag{4.14}$$

and

$$|\lambda - \zeta| \leq \zeta \ . \tag{4.15}$$

Proof. The proof is in the continuation of the proof of Theorem 3.1. The main steps are as follows. We first rewrite the matrix (3.11) obtained at the end of the earlier proof in a form that has the structure seen in (4.1), i.e., that allows us to apply Theorem 4.1. The inequalities (4.12), (4.13), (4.14), (4.15) are then deduced from, respectively, (4.3), (4.4), (4.5), (4.6). The difficulty is in the analysis of the extremal eigenvalues of the blocks $\widehat{A}$, $\widehat{C}$ and related Schur complements $S_{\widehat{A}}$, $S_{\widehat{C}}$, which needs to be done carefully to obtain bounds as accurate as possible using no other parameter than $\underline{\mu}$, $\overline{\mu}$, $\underline{\nu}$, $\overline{\nu}$.

Thus all notation and definitions introduced in the proof of Theorem 3.1 are valid here, and we also use the same continuity argument on the matrix (3.4) to handle the cases where one would have an eigenvalue of $M_A^{-1}A$ exactly equal to 1. Observe in this respect that not only the eigenvalues, but also the bounds to be proved, vary continuously with ε , at least when ε is small enough to ensure that if $\overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right) > 1$, then $(1-\varepsilon)\lambda_1 > 1$, where λ_1 is the smallest eigenvalue of $M_A^{-1}A$ that is strictly larger than 1.

Observing that, for $\lambda \in (\underline{\mu} \, , \overline{\mu})$, one has $1 - \rho_A^2 \leq 2\,\lambda - \lambda^2 \leq 1$, we further define

$$\widetilde{\mu}_{\min} = \min_i \Gamma_{ii} = \begin{cases} 0 & \text{for } M_d \, , \\ \underline{\mu} & \text{for } M_t \, , \ M_u \, , \\ 1 - \rho_A^2 & \text{for } M_f \, , M_g \, , \ M_{t_2} \, , \ M_{u_2} \, , \end{cases} \tag{4.16}$$

$$\widetilde{\mu}_{\max} = \max_i \Gamma_{ii} \leq \begin{cases} 0 & \text{for } M_d \, , \\ \overline{\mu} & \text{for } M_t \, , \ M_u \, , \\ 1 & \text{for } M_f \, , M_g \, , \ M_{t_2} \, , \ M_{u_2} \, , \end{cases} \tag{4.17}$$

whereas we observe that (3.7), (3.8) imply

$$\Gamma = \widehat{\Lambda} \qquad \text{for } M_t \, , \ M_u \ M_g \, , \ M_{t_2} \, , \ M_{u_2} \ . \tag{4.18}$$

We also note for later use that (3.10) implies $G\,G^T = M_S^{-1/2} B\,A^{-1} B^T M_S^{-1/2}$ and hence

$$\lambda_{\min}\left(\widetilde{C} + G\,G^T\right) = \underline{\nu} \ , \qquad \lambda_{\max}\left(\widetilde{C} + G\,G^T\right) = \overline{\nu} \ . \tag{4.19}$$

In the proof of Theorem 3.1, we have seen that, for each of the considered preconditioners, $M_*^{-1}K$ has the same eigenvalue as the matrix (3.11). To proceed we assume, without loss of generality, that the rows for which $(\Delta_+)_{ii}$ is positive are ordered first; i.e.,

$$\Delta_+ = \begin{pmatrix} \Delta_1 & \\ & 0 \end{pmatrix} \ , \qquad \Delta_- = \begin{pmatrix} 0 & \\ & \Delta_2 \end{pmatrix} \ .$$

We may further partition $\widehat{\Lambda}$, Γ , and G accordingly:

$$\widehat{\Lambda}=\begin{pmatrix}\widehat{\Lambda}_1 & \\ & \widehat{\Lambda}_2\end{pmatrix}\ ,\quad \Gamma=\begin{pmatrix}\Gamma_1 & \\ & \Gamma_2\end{pmatrix}\ ,\quad G=\begin{pmatrix}G_1 & G_2\end{pmatrix}\ ,\quad G^T=\begin{pmatrix}G_1^T\\ G_2^T\end{pmatrix}\ .$$

One then has $\Delta_1^2=I-\Gamma_1$ and $\Delta_2^2=\Gamma_2-I$. This allows one to rewrite the matrix (3.11) as

$$\begin{pmatrix}\widehat{\Lambda}_1 & & \widehat{\Lambda}_1^{1/2}\Delta_1 G_1^T\\ & \widehat{\Lambda}_2 & \widehat{\Lambda}_2^{1/2}\Delta_2 G_2^T\\ -G_1\Delta_1\widehat{\Lambda}_1^{1/2} & G_2\Delta_2\widehat{\Lambda}_2^{1/2} & \widetilde{C}+G_1\,\Gamma_1\,G_1^T+G_2\,\Gamma_2\,G_2^T\end{pmatrix}\ .$$

Hence we may apply Theorem 4.1 with

$$\widehat{A}=\widehat{\Lambda}_1\ ,\quad \widehat{C}=\begin{pmatrix}\widehat{\Lambda}_2 & \widehat{\Lambda}_2^{1/2}\Delta_2 G_2^T\\ G_2\Delta_2\widehat{\Lambda}_2^{1/2} & \widetilde{C}+G_1\,\Gamma_1\,G_1^T+G_2\,\Gamma_2\,G_2^T\end{pmatrix}\ ,\quad \widehat{B}=\begin{pmatrix}0\\ G_1\Delta_1\widehat{\Lambda}_1^{1/2}\end{pmatrix}\ .$$

Of course, before applying Theorem 4.1, we need to check that its assumptions are satisfied. For M_d (i.e., $\Gamma=0$, entailing $\Gamma_1=0$, $\Delta_1=I$, and that Γ_2 , Δ_2 , G_2 are trivial empty matrices), this clearly follows from the assumptions on B and C , which (see (3.10)) imply that either $\widehat{B}^T$ $\big(=\widehat{\Lambda}_1^{1/2}G_1^T\big)$ has full rank or $\widehat{C}$ $\big(=\widetilde{C}\big)$ is positive definite on its null space. Regarding all other preconditioners, we prove below (see either (4.25) (case $\widetilde{\mu}_{\max}\leq 1$) or (4.29) (case $\widetilde{\mu}_{\max}>1$)) a positive lower bound on the eigenvalues of $\widehat{C}$; hence it is positive definite, and we need not discuss further the rank of $\widehat{B}$.

Now Theorem 4.1 is actually needed only if $\widetilde{\mu}_{\min}<1$. Indeed, when $\widetilde{\mu}_{\min}\geq 1$, Δ_1 and therefore $\widehat{\Lambda}_1$ and G_1 are trivial empty matrices, and the preconditioned matrix is in fact similar to an SPD matrix. In view of (4.16), this happens only for M_t and M_u and when $\underline{\mu}\geq 1$, proving that the eigenvalues are real as claimed in this case. To be complete, this also happens for other preconditioners except M_d when $\rho_A=0$ (i.e., $M_A=A$), entailing $\delta=0$. In these cases, we have only to prove (4.12). This is done below without assuming anything specific on $\widehat{\Lambda}_1$ and G_1 , i.e., including the case where these matrices are trivial as well.

If $\widetilde{\mu}_{\min}<1$, we have, recalling (4.18) and (4.19),

$$\begin{aligned}\lambda_{\max}\left(\widehat{B}\,\widehat{B}^T\right)&=\lambda_{\max}\left(G_1\,\widehat{\Lambda}_1\Delta_1^2\,G_1^T\right)\\ &\leq\lambda_{\max}\left(G\,G^T\right)\ \max_i\left(\left(\widehat{\Lambda}_1\right)_{ii}(1-(\Gamma_1)_{ii})\right)\\ &\leq\overline{\nu}\begin{cases}\overline{\mu} & \text{if } M_*=M_d\,,\\ \displaystyle\max_{\lambda\in(\underline{\mu}\,,\,\overline{\mu})}\lambda(1-\lambda)^2 & \text{if } M_*=M_f\,,\\ \displaystyle\max_{\lambda\in(\widetilde{\mu}_{\min}\,,\,1)}(\lambda-\lambda^2) & \text{otherwise}\,.\end{cases}\end{aligned}\tag{4.20}$$

The function $g(\lambda)=\lambda(1-\lambda)^2$ is increasing for $\lambda<1/3$, decreasing for $1/3<\lambda<1$, and increasing for $\lambda>1$. Hence, if $\underline{\mu}<1/3<\overline{\mu}$, the maximum in the interval $(\underline{\mu}\,,\,\overline{\mu})$ is $\max\left(g(1/3)\,,\,g(\overline{\mu})\right)$; otherwise, the maximum is always at one of the boundaries and thus equal to $\max\left(g(\underline{\mu})\,,\,g(\overline{\mu})\right)$. On the other hand, the function $h(\lambda)=\lambda-\lambda^2$ has a unique maximum at $\lambda=1/2$. Hence its maximum over the interval $(\widetilde{\mu}_{\min}\,,\,1)$ is equal to $h(1/2)=1/4$ if $1/2$ belongs to this interval, and otherwise (i.e., when $\widetilde{\mu}_{\min}>1/2$) is always equal to $h\left(\widetilde{\mu}_{\min}\right)$. Using (4.20) and these considerations together with (4.16), the application of Theorem 4.1 then yields (4.14).

Hence we are left with the proof of (4.12), (4.13), (4.15), which requires us to bound the eigenvalues of $\widehat{A}$, $\widehat{C}$ and related Schur complements.

For $\lambda_{\max}(\widehat{A})$, we observe that if $\widehat{\Lambda}=\Gamma$, then all diagonal entries in $\widehat{\Lambda}_1$ are less than or equal to 1, since they correspond to rows for which Γ_{ii} does not exceed 1. Hence, with (4.18),

$$\lambda_{\max}\left(\widehat{A}\right)\leq\begin{cases}\max\limits_i\widehat{\Lambda}_{ii}=\overline{\mu} & \text{if } M_*=M_d \text{ or } M_*=M_f\,,\\ 1 & \text{otherwise}\,,\end{cases}\tag{4.21}$$

whereas, straightforwardly,

$$\lambda_{\min}\left(\widehat{A}\right)\geq\min_i\widehat{\Lambda}_{ii}=\begin{cases}\underline{\mu} & \text{for } M_d\,,\,M_t\,,\,M_u\,,\,M_f\,,\\ 1-\rho_A^2 & \text{for } M_g\,,M_{t_2}\,,\,M_{u_2}\,.\end{cases}\tag{4.22}$$

To analyze the Schur complement $S_{\widehat{A}}$, one first has to obtain it explicitly. One way is to consider the Schur complement in (3.11),

$$\widehat{\Lambda}^{1/2}\left(I\,+\,(\Delta_++\Delta_-)\,G^T\left(\widetilde{C}+G\Gamma G^T\right)^{-1}G\,(\Delta_+-\Delta_-)\right)\widehat{\Lambda}^{1/2}=\widehat{\Lambda}^{1/2}\;H\;\widehat{\Lambda}^{1/2}\,,$$

where the right-hand side defines the matrix H. Its inverse may be obtained by the Sherman–Morrisson–Woodbury formula:

$$\begin{aligned}H^{-1}&=I\,-\,(\Delta_++\Delta_-)\,G^T\\ &\qquad\left(\left(\widetilde{C}+G\Gamma G^T\right)\,+\,G\,(\Delta_+-\Delta_-)\,(\Delta_++\Delta_-)\,G^T\right)^{-1}\;G\,(\Delta_+-\Delta_-)\\ &=I-(\Delta_++\Delta_-)\,G^T\left(\widetilde{C}+G\,G^T\right)^{-1}G\,(\Delta_+-\Delta_-)\\ &=\begin{pmatrix}I&\\&I\end{pmatrix}\;-\begin{pmatrix}\Delta_1G_1^T\\ \Delta_2G_2^T\end{pmatrix}\left(\widetilde{C}+G\,G^T\right)^{-1}\begin{pmatrix}G_1\Delta_1 & -G_2\Delta_2\end{pmatrix}.\end{aligned}$$

The top left block of $\widehat{\Lambda}^{-1/2}H\,\widehat{\Lambda}^{-1/2}$ is the inverse of $S_{\widehat{A}}$; hence,

$$S_{\widehat{A}}^{-1}=\widehat{\Lambda}_1^{-1}-\widehat{\Lambda}_1^{-1/2}\Delta_1G_1^T\left(\widetilde{C}+G\,G^T\right)^{-1}G_1\Delta_1\widehat{\Lambda}_1^{-1/2}\;.$$

On the other hand, $G_1^T(\widetilde{C}+G\,G^T)^{-1}G_1$ and $G_1\,G_1^T(\widetilde{C}+G\,G^T)^{-1}$ have the same set of nonzero eigenvalues [27, Lemma A.1], and they are bounded by

$$\max_{\mathbf{v}}\;\frac{\mathbf{v}^T\,G_1\,G_1^T\,\mathbf{v}}{\mathbf{v}^T\left(\widetilde{C}+G\,G^T\right)\mathbf{v}}\leq\max_{\mathbf{v}}\;\frac{\mathbf{v}^T\,G_1\,G_1^T\,\mathbf{v}}{\mathbf{v}^T\left(G_1\,G_1^T+G_2\,G_2^T\right)\mathbf{v}}\leq 1\;.$$

One then finds

$$S_{\widehat{A}}^{-1}\geq\widehat{\Lambda}_1^{-1}\left(I-\Delta_1^2\right)=\widehat{\Lambda}_1^{-1}\;\Gamma_1$$

(where, here and in the following, inequalities between matrices are to be understood in the nonnegative definite sense: $Q \geq R$ if and only if $Q - R$ is nonnegative definite). Thus, $S_{\widehat{A}} \leq \widehat{\Lambda}_1 \Gamma_1^{-1}$, and hence (recalling that $\rho_A < 1 \iff \overline{\mu} < 2$)

$$\lambda_{\max}\left(S_{\widehat{A}}\right) \leq \begin{cases} (2-\overline{\mu})^{-1} & \text{if } \ M_* = M_f \,, \\ 1 & \text{if } \ \widehat{\Lambda} = \Gamma \ . \end{cases} \tag{4.23}$$

We now consider $\widehat{C}$ and $S_{\widehat{C}}$. We first consider the case where G_2 is trivial, which happens if and only if $\widetilde{\mu}_{\max} \leq 1$. One then obtains

$$\lambda_{\max}\left(\widehat{C}\right) = \lambda_{\max}\left(\widetilde{C} + G\,\Gamma G^T\right) \leq \lambda_{\max}\left(\widetilde{C} + G\,G^T\right) = \overline{\nu} \tag{4.24}$$

and ($\widetilde{\mu}_{\max} \leq 1$ implying $\widetilde{\mu}_{\min} \leq 1$)

$$\lambda_{\min}\left(\widehat{C}\right) = \lambda_{\min}\left(\widetilde{C} + G\,\Gamma G^T\right) \geq \widetilde{\mu}_{\min}\ \lambda_{\min}\left(\widetilde{C} + G\,G^T\right) = \widetilde{\mu}_{\min}\ \underline{\nu} \ . \tag{4.25}$$

Since

$$S_{\widehat{C}} = \begin{pmatrix} \widehat{\Lambda}_2 & \widehat{\Lambda}_2^{1/2} \Delta_2 G_2^T \\ G_2 \Delta_2 \widehat{\Lambda}_2^{1/2} & \widetilde{C} + G_1\, G_1^T + G_2\,\Gamma_2\, G_2^T \end{pmatrix} \ ,$$

one straightforwardly obtains, when G_2 is trivial,

$$\lambda_{\max}\left(S_{\widehat{C}}\right) = \lambda_{\max}\left(\widetilde{C} + G\,G^T\right) = \overline{\nu} \tag{4.26}$$

and

$$\lambda_{\min}\left(S_{\widehat{C}}\right) = \lambda_{\min}\left(\widetilde{C} + G\,G^T\right) = \underline{\nu} \ . \tag{4.27}$$

Note that $\widetilde{\mu}_{\max} > 1$ is not possible for M_d, M_f, M_g, M_{t_2}, M_{u_2}. Hence the above estimates are sufficient for these preconditioners, as well as for M_t and M_u when $\overline{\mu} \leq 1$. One may indeed check that, for each of these cases, (4.12), (4.13), and (4.15) are proved by combining Theorem 4.1 with the bounds in (4.21), (4.22), (4.23), (4.24), (4.25), (4.26), and (4.27). Regarding (4.23), we use $\zeta \leq \lambda_{\max}(S_{\widehat{C}})$ (i.e., the bound for $\widehat{C}$ semidefinite) in the case of M_d, where we have no valid upper bound on $\lambda_{\max}\left(S_{\widehat{A}}\right)$. On the other hand, as noted above, with M_d, one has $\widehat{C} = \widetilde{C}$ (because $\Gamma = 0$, entailing $\Gamma_1 = 0$, $\Delta_1 = I$, and that Γ_2, Δ_2, G_2 are trivial empty matrices) and therefore $\widehat{C} = 0$ when $C = 0$, hence the improved bound for $\overline{\chi}$ in this case, which is obtained by using $\lambda_{\max}(\widehat{C}) = 0$ instead of (4.24).

Now it remains to prove (4.12), (4.13), and (4.15) for M_t and M_u in the case $\overline{\mu} > 1$. Observe that we then have $\widehat{\Lambda} = \Gamma$; hence we may restrict the analysis to this situation. We first note that the matrix

$$\begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & \Gamma_2 \end{pmatrix}$$

(where each block is square with the same number of columns as G_2) can be permuted to a block diagonal form with 2 × 2 blocks:

$$\begin{pmatrix} \gamma & \delta\,\gamma^{1/2} \\ \delta\,\gamma^{1/2} & \gamma \end{pmatrix} \ ,$$

where $\gamma = (\Gamma_2)_{ii}$, $\delta = (\Delta_2)_{ii}$, and thus $\gamma = 1 + \delta^2$. It turns out that

$$\begin{pmatrix} \gamma & \delta\,\gamma^{1/2} \\ \delta\,\gamma^{1/2} & \gamma \end{pmatrix} - \tau \begin{pmatrix} \nu & \\ & 1 \end{pmatrix}$$

is nonnegative definite for τ equal to the smallest root of

$$\nu\, x^2 - (\nu + 1)\,\gamma\, x + \gamma = 0 \ , \tag{4.28}$$

which is nothing but $(f\,(\gamma\,,\,\nu))^{-1}$. Setting $\nu = \underline{\nu}$ and recalling the definition (4.10) of η , the fact that $1 \le \gamma \le \overline{\mu}$ implies $f\,(\gamma\,,\,\nu) \le f\,(\overline{\mu}\,,\,\underline{\nu}) = \eta$. Hence,

$$\begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & I + \Delta_2^2 \end{pmatrix} \ge \eta^{-1} \begin{pmatrix} \underline{\nu}\, I & \\ & I \end{pmatrix} \ .$$

Then we find (taking into account that $\eta \ge 1$; see (4.9))

$$\begin{aligned} \widehat{C} &= \begin{pmatrix} 0 & \\ & \widetilde{C} + G_1\,\Gamma_1\,G_1^T \end{pmatrix} + \begin{pmatrix} I & \\ & G_2 \end{pmatrix} \begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & \Gamma_2 \end{pmatrix} \begin{pmatrix} I & \\ & G_2^T \end{pmatrix} \\ &\ge \begin{pmatrix} 0 & \\ & \widetilde{C} + G_1\,\Gamma_1\,G_1^T \end{pmatrix} + \ \eta^{-1} \begin{pmatrix} \underline{\nu}\, I & \\ & G_2\,G_2^T \end{pmatrix} \\ &\ge \min\left(\eta^{-1}\,,\,\underline{\mu}\right) \begin{pmatrix} \underline{\nu}\, I & \\ & \widetilde{C} + GG^T \end{pmatrix} \\ &\ge \min\left(\eta^{-1}\,,\,\underline{\mu}\right)\ \underline{\nu}\ I \ ; \end{aligned}$$

that is,

$$\lambda_{\min}\left(\widehat{C}\right) \ge \min\left(\eta^{-1}\,,\,\underline{\mu}\right)\ \underline{\nu} \ . \tag{4.29}$$

Similarly, one has

$$\begin{aligned} S_{\widehat{C}} &= \begin{pmatrix} 0 & \\ & \widetilde{C} + G_1\,G_1^T \end{pmatrix} + \begin{pmatrix} I & \\ & G_2 \end{pmatrix} \begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & \Gamma_2 \end{pmatrix} \begin{pmatrix} I & \\ & G_2^T \end{pmatrix} \\ &\ge \begin{pmatrix} 0 & \\ & \widetilde{C} + G_1\,G_1^T \end{pmatrix} + \ \eta^{-1} \begin{pmatrix} \underline{\nu}\, I & \\ & G_2\,G_2^T \end{pmatrix} \\ &\ge \eta^{-1} \ \begin{pmatrix} \underline{\nu}\, I & \\ & \widetilde{C} + GG^T \end{pmatrix} \\ &\ge \eta^{-1}\,\underline{\nu}\ I \ ; \end{aligned}$$

that is,

$$\lambda_{\min}\left(S_{\widehat{C}}\right) \ge \eta^{-1}\,\underline{\nu} \ . \tag{4.30}$$

The analysis of the largest eigenvalue of $S_{\widehat{C}}$ is based on the same ideas:

$$\tau \begin{pmatrix} \nu & \\ & 1 \end{pmatrix} - \begin{pmatrix} \gamma & \delta\,\gamma^{1/2} \\ \delta\,\gamma^{1/2} & \gamma \end{pmatrix}$$

is nonnegative definite for τ equal to the largest root of (4.28), which is $\frac{\gamma}{\nu}\, f\,(\gamma\,,\,\nu)$. Setting $\nu = \overline{\nu}$ and recalling the definition (4.11) of $\widetilde{\nu}$, the fact that $1 \le \gamma \le \overline{\mu}$ implies $\frac{\gamma}{\nu}\, f\,(\gamma\,,\,\nu) \le \frac{\overline{\mu}}{\overline{\nu}}\, f\,(\overline{\mu}\,,\,\overline{\nu}) = \frac{\widetilde{\nu}}{\overline{\nu}}$. Hence,

$$\begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & I+\Delta_2^2 \end{pmatrix} \le \widetilde{\nu}\,\overline{\nu}^{-1} \begin{pmatrix} \overline{\nu} I & \\ & I \end{pmatrix} .$$

Then we find (taking into account that $\widetilde{\nu} \ge \overline{\nu}$; see (4.9))

$$\begin{aligned} S_{\widehat{C}} &= \begin{pmatrix} \; & \\ & \widetilde{C} + G_1\,G_1^T \end{pmatrix} + \begin{pmatrix} I & \\ & G_2 \end{pmatrix} \begin{pmatrix} \Gamma_2 & \Gamma_2^{1/2}\Delta_2 \\ \Delta_2\Gamma_2^{1/2} & \Gamma_2 \end{pmatrix} \begin{pmatrix} I & \\ & G_2^T \end{pmatrix} \\ &\le \begin{pmatrix} \; & \\ & \widetilde{C} + G_1\,G_1^T \end{pmatrix} + \;\widetilde{\nu}\,\overline{\nu}^{-1} \begin{pmatrix} \overline{\nu}\,I & \\ & G_2\,G_2^T \end{pmatrix} \\ &\le \widetilde{\nu}\,\overline{\nu}^{-1} \begin{pmatrix} \overline{\nu}\,I & \\ & \widetilde{C} + GG^T \end{pmatrix} \\ &\le \widetilde{\nu}\; I\;; \end{aligned}$$

that is,

$$\lambda_{\max}\left(S_{\widehat{C}}\right) \le \widetilde{\nu}\;. \tag{4.31}$$

Since $\lambda_{\max}(\widehat{C}) \le \lambda_{\max}(S_{\widehat{C}})$, we may also use $\lambda_{\max}(\widehat{C}) \le \widetilde{\nu}$. Then one may check that (4.12), (4.13), and (4.15) for M_t and M_u in the case $\overline{\mu} > 1$ indeed follow from Theorem 4.1 using this estimate and those in (4.21), (4.22), (4.23), (4.29), (4.30), and (4.31). □

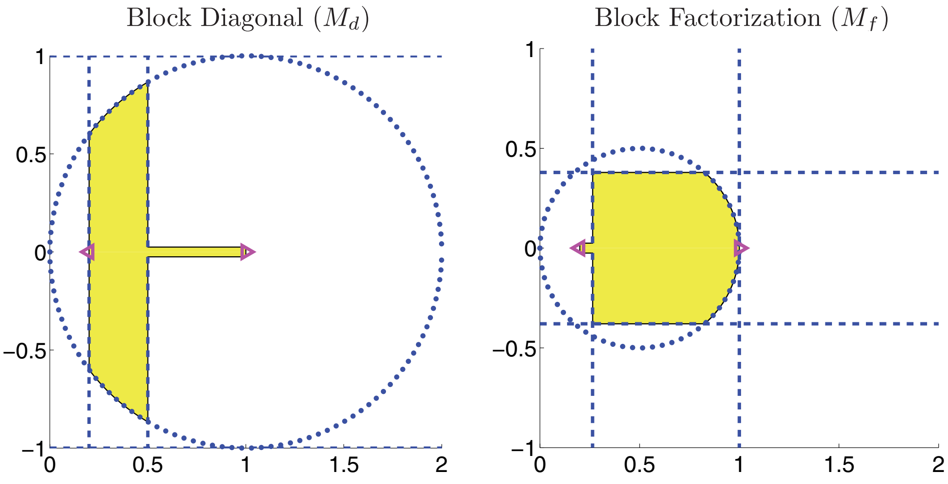
How our bounds work is illustrated in Figure 1 for two examples of preconditioners, using the values

$$\underline{\mu} = 0.4\;, \qquad \overline{\mu} = 1\;, \qquad \underline{\nu} = 0.2\;, \qquad \overline{\nu} = 1 \tag{4.32}$$

(which come from the application studied in the next section) and assuming $C = 0$ so that the more favorable value of $\overline{\chi}$ applies for the block diagonal preconditioner. The symbols ◁ and ▷ correspond to, respectively, $\underline{\xi}$ and $\overline{\xi}$; hence real eigenvalues are to be situated in between according to (4.12). Regarding nonreal eigenvalues, the dashed vertical lines correspond to $\lambda = \underline{\chi}$ (left line) and $\lambda = \overline{\chi}$ (right line), the dashed horizontal lines correspond to $\lambda = \pm i\,\delta$, and the dotted circle corresponds to $|\lambda = \zeta| = \zeta$; hence, according to (4.13) and (4.14), the nonreal eigenvalues must lie in the box delimited by the four dashed lines but also, according to (4.15), within the disk delimited by the dotted circle. In summary, they must thus be in the shaded (yellow) region delimited by solid lines, understanding that horizontal lines close to the real axis are infinitesimally close to it, with only real eigenvalues actually being permitted in the area between them.

The values in Tables 1 vary continuously in function of the four main parameters $\underline{\mu}\,,\,\overline{\mu}\,,\,\underline{\nu}\,,\,\overline{\nu}$, except possibly for M_t and M_u, where we have to distinguish different cases; however, if $\underline{\nu} \le 1 \le \overline{\nu}$ (as one expects if M_S is properly scaled), then, since $f\,(1\,,\,\nu) = \max(1\,,\,\nu)$ (see Lemma 4.2), one has $\eta \to 1$ and $\widetilde{\nu} \to \overline{\nu}$ for $\overline{\mu} \to 1$, showing that the bounds for M_t and M_u also vary continuously with $\overline{\mu}$.

Moreover, independent of the condition $\underline{\nu} \le 1 \le \overline{\nu}$, when $\underline{\mu}\,,\,\overline{\mu}\,,\,\underline{\nu}\,,\,\overline{\nu} \to 1$, then, for all preconditioners but $M_d\,,\,\underline{\xi}\,,\,\overline{\xi}\,,\,\underline{\chi}\,,\,\overline{\chi} \to 1$ and $\delta \to 0\,,\,\zeta \to \frac{1}{2}$. This means that


Block Diagonal (M_d)

Block Factorization (M_f)

Fig. 1. *Application of Theorem* 4.3 *with main parameters as in* (4.32); ◁: $\underline{\xi}$; ▷: $\overline{\xi}$; *dotted circle:* $|\lambda = \zeta| = \zeta$; *dashed vertical lines:* $\lambda = \underline{\chi}$ *and* $\lambda = \overline{\chi}$; *dashed horizontal lines:* $\lambda = \pm i\,\delta$.

both real and nonreal eigenvalues are confined in a region which converges smoothly towards the single point 1 when $M_A \to A$ and $M_S \to S$. For M_d , we then have $\underline{\xi}$, $\overline{\xi} \to 1$, whereas $\delta \to 1$, $\zeta \to 1$, $\underline{\chi} \to \frac{1}{2}$, and $\overline{\chi} \to 1$ in general, but $\overline{\chi} \to \frac{1}{2}$ when $C = 0$. Hence all real eigenvalues converge towards 1, but nonreal eigenvalues do not: their real part lies in general between $\frac{1}{2}$ and 1 , converging in particular towards $\frac{1}{2}$ when $C = 0$; on the other hand, their imaginary part may remain significant. This confirms the analysis in [17, Lemma 2.2], where it is shown that if $C = 0$, only three distinct eigenvalues remain at the limit $M_A = A$ and $M_S = S$: 1 and $\frac{1}{2}\left(1 \pm i\sqrt{3}\right)$; interestingly, these latter numbers are at the intersection of the lines $\Re\mathrm{e}(\lambda) = \frac{1}{2}$ (our dashed vertical lines, which coincide in this case) and $|\lambda - 1| \leq 1$ (our dotted circle). It is also interesting to observe that the modulus of these three remaining distinct eigenvalues is equal to 1, whereas, in the same circumstances, only the two eigenvalues ±1 remain associated with the positive definite block diagonal preconditioner (3.1) [17, Lemma 2.1]; thus, at the limit of exact preconditioning of A and S , equality is attained in both relations (3.2), (3.3) from Theorem 3.1.

Regarding real eigenvalues, it is worth noting that, when using M_d , M_f , M_t with $\overline{\mu} \leq 1$, or M_u with $\overline{\mu} \leq 1$, the bounds (4.12) then reduce to

$$\min\left(\underline{\mu}\,,\,\underline{\nu}\right) \leq \lambda \leq \max\left(\overline{\mu}\,,\,\overline{\nu}\right) \quad , \tag{4.33}$$

which is simple and appealing. With M_g , M_{t_2} , M_{u_2} , the corresponding result is

$$\min\left(1 - \rho_A^2\,,\,\underline{\nu}\right) \leq \lambda \leq \max\left(1\,,\,\overline{\nu}\right) \quad , \tag{4.34}$$

which requires only $\rho_A < 1$, i.e., $\overline{\mu} < 2$.

When using M_t or M_u with $\overline{\mu} > 1$, our estimates for real eigenvalues are somehow less favorable and indicate that scaling M_A to have the eigenvalues of $M_A^{-1}A$ be greater than 1 may have an adverse effect on the clustering of the real eigenvalues. This is better seen in an example, so consider again the values (4.32), but now add a scaling parameter so that $\underline{\mu} = 0.4\,\alpha$ and $\overline{\mu} = \alpha$ for some $\alpha \geq 1$. In Figure 2,[4] we depict the

[4] The code allowing one to reproduce the results reported in this figure is provided as supplementary material through the electronic version of the journal.

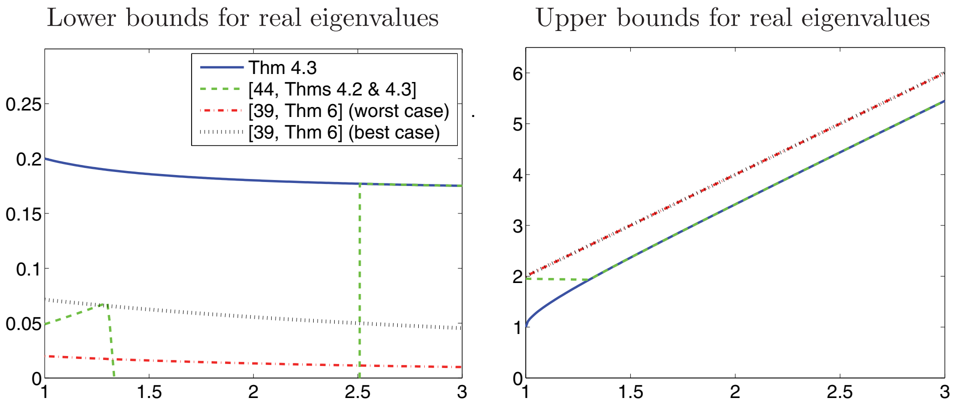


Fig. 2. *Upper and lower bounds for real eigenvalues with M_t or M_u as a function of α when $\underline{\mu} = 0.4\,\alpha$, $\overline{\mu} = \alpha$, $\underline{\nu} = 0.2$, and $\overline{\nu} = 1$ (the legend in the left plot also applies to the right one); for the bounds from* [39] *we have a "best" and a "worst" case because these bounds are expressed as functions of the extremal eigenvalues of $M_S^{-1}(C + B\,M_A^{-1}B^T)$ instead of $\underline{\nu}\,,\overline{\nu}\,$; the "best" case is obtained by setting $\lambda_{\min}(M_S^{-1}(C + B\,M_A^{-1}B^T)) = \overline{\mu}\,\underline{\nu}$ (the largest possible value) and $\lambda_{\max}(M_S^{-1}(C + B\,M_A^{-1}B^T)) = \underline{\mu}\,\overline{\nu}$ (the smallest possible value), whereas the "worst" case corresponds to $\lambda_{\min}(M_S^{-1}(C + B\,M_A^{-1}B^T)) = \underline{\mu}\,\underline{\nu}$ (the smallest possible value) and $\lambda_{\max}(M_S^{-1}(C + B\,M_A^{-1}B^T)) = \overline{\mu}\,\overline{\nu}$ (the largest possible value).*

evolution of the lower and upper bounds for real eigenvalues; these plots also illustrate how η and $\widetilde{\nu}$ vary with $\overline{\mu}$, since, in this example, $\underline{\xi} = \eta^{-1}\underline{\nu}$ and $\overline{\xi} = \widetilde{\nu}$. One sees that the η factor has only a limited impact on the lower bound, in agreement with the second upper bound (4.9) on $f\,(\overline{\mu}\,,\underline{\nu})$, which is never worse than $1 + \underline{\nu} = 1.2$. However, $\widetilde{\nu}$ grows with $\overline{\mu}$, in fact also in agreement with the same upper bound, which yields $\overline{\nu} \leq \overline{\mu}(\overline{\nu} + 1) - \overline{\nu}/(\overline{\nu} + 1) = 2\,\alpha - 1/2$. In section 5, we shall see an example where the real eigenvalues really do spread out when α increases, closely following our bounds.

In Figure 2, we also compare our bounds with bounds appearing in papers by Zulehner [44] and Simoncini [39], which analyze, respectively, inexact Uzawa and block triangular preconditioners, and contain the best previous estimates we are aware of; one sees that our analysis is significantly sharper than that of Simoncini, and more general than that of Zulehner, which is effective only if α is large enough.

Now, staying with M_t and M_u, it is well known (see [39, 44]) that scaling M_A to increase $\underline{\mu}$ has on the contrary a welcome effect on the nonreal eigenvalues, which become forbidden when $\underline{\mu} \geq 1$; this is also confirmed by our analysis, since δ decreases as $\underline{\mu}$ increases and vanishes for $\underline{\mu} \geq 1$. Whereas previous works often focused on this and accordingly suggested selecting the scaling to enforce this condition, our analysis reveals that there is in fact a tradeoff between the clustering of real and nonreal eigenvalues. This will be further illustrated in the next section.

We could also compare our bounds for nonreal eigenvalues with previous bounds. However, as seen in Figure 1, it is in fact more sensible to *combine* the different bounds than to discuss which one is the best: the more inequalities we have, the better we delimit the region that contains the eigenvalues. In particular, the bound obtained in [39, Theorem 2] for block triangular preconditioners can play a useful complementary role, and it is also appealing in its simplicity. For the sake of completeness, we recall this bound in the following theorem, noting that, by item (2) of Theorem 3.1, we extend its scope of application to inexact Uzawa preconditioners. Moreover, applying this bound to M_{t_2} allows its further extension to block SGS preconditioners, via item (3) of Theorem 3.1.

THEOREM 4.4. *Let the assumptions of Theorems* 3.1 *and* 4.3 *hold.*

(1) *Eigenvalues λ of $M_t^{-1}K$ and $M_u^{-1}K$ with nonzero imaginary part are possible only if $\underline{\mu} < 1$, in which case they satisfy*

$$|\lambda - 1| \le \sqrt{1-\underline{\mu}}\ . \tag{4.35}$$

(2) *The eigenvalues λ of $M_g^{-1}K$, $M_{t_2}^{-1}K$, and $M_{u_2}^{-1}K$ with nonzero imaginary part satisfy*

$$|\lambda - 1| \le \rho_A\ . \tag{4.36}$$

We may further combine this result with (4.33) for M_t, M_u, and with (4.34) for M_g, M_{t_2}, M_{u_2}. This straightforwardly yields the following corollary.

COROLLARY 4.5. *Let the assumptions of Theorems* 3.1 *and* 4.3 *hold, and assume that $\rho_S < 1$, where*

$$\rho_S = \ \rho\left(I - M_S^{-1}S\right) = \max\left(\overline{\nu} - 1\,,\, 1 - \underline{\nu}\right)\ .$$

(1) *If*

$$\overline{\mu} = \lambda_{\max}\left(M_A^{-1}A\right) \le 1\ , \tag{4.37}$$

then

$$\rho\left(I - M_t^{-1}K\right) = \rho\left(I - M_u^{-1}K\right) \le \max\left(\sqrt{\rho_A}\,,\, \rho_S\right)\ . \tag{4.38}$$

(2) *There holds*

$$\rho\left(I - M_g^{-1}\,K\right) = \rho\left(I - M_{t_2}^{-1}\,K\right) = \rho\left(I - M_{u_2}^{-1}\,K\right) \le \max\left(\rho_A\,,\, \rho_S\right)\ . \tag{4.39}$$

Let us stress that the assumption (4.37) is made for the sake of simplicity. If it is not satisfied, a bound on the spectral radius can still be obtained from (4.12) and (4.35). On the other hand, with this result one sees even more clearly that "symmetrized" preconditioners like M_g (and, by extension, M_f) can be cost effective when the preconditioner for S is better than that for A, or of similar quality. On the contrary, when the preconditioner for A is much better, the clustering of the spectrum essentially depends on the eigenvalues of $M_S^{-1}S$, and M_g or M_f can only bring a mitigated improvement to the block triangular preconditioners, so that the extra cost involved likely does not pay off.

As the final remark in this section, rescaling M_A is of course also possible with preconditioners other than M_t and M_u. We do not discuss this explicitly because effects are moderate and easily predicted by inserting into Table 1 rescaled estimates for $\underline{\mu}$, $\overline{\mu}$. In particular, for M_g, M_{t_2}, M_{u_2}, it is clear, at least from this theoretical viewpoint, that the best scaling is the one that minimizes ρ_A. It is also possible to rescale the whole operator $\widetilde{M}_A$. In combination with M_{t_2} or M_{u_2}, this will have the same effect, already discussed above, as rescaling M_A when using M_t or M_u; observe that the parameters in Table 1 for M_g, M_{t_2}, and M_{u_2} in fact coincide with those for

M_t and M_u when exchanging $\underline{\mu}$ for $1-\rho_A^2$ and $\overline{\mu}$ for 1. Rescaling $\widetilde{M}_A$ with M_g is more ambiguous. Letting α be the scaling parameter, one possibility is to consider

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

This thus amounts to scaling the whole spectrum obtained with unscaled M_A and $\widetilde{M}_A$, and inverse scaling applied to M_S. Since the convergence of the Krylov subspace method is independent of the global scaling of the preconditioner, this option is therefore equivalent to just applying the inverse scaling to M_S.

5. Example of application. In this section, we consider the typical example provided by the stationary Stokes problem on the unit square Ω in two dimensions. That is, finding the velocity vector $\mathbf{v} : \Omega \to \mathbb{R}^2$ and the kinematic pressure field $p : \Omega \to \mathbb{R}$ satisfying

$$\begin{aligned} -\Delta\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \ , \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega \ , \end{aligned} \tag{5.1}$$

where $\mathbf{f}$ represents a prescribed force. For the sake of simplicity we choose Dirichlet boundary conditions for the velocity.

As a general rule, the discretization of this problems yields a linear system $K\,\mathbf{u} = \mathbf{b}$, whose coefficient matrix K has the form (1.1). Here we consider more particularly finite difference discretization on a staggered grid, for which $C = 0$.

There is a technical difficulty appearing from the lack of boundary conditions for the pressure, which is determined only up to a constant. At the discrete level, this is reflected in the fact that $B^T\,\mathbf{1} = \mathbf{0}$, where $\mathbf{1}$ is the vector of all ones. Hence K is singular with null space spanned by $(\mathbf{0} \ \ \mathbf{1}^T)^T$. The case of rank deficient B with $C\,\mathbf{e} = \mathbf{0}$ for all vectors $\mathbf{e}$ in the null space of B^T is analyzed in the appendix, in light of the results in [10, 18]. It turns out that right preconditioned GMRES or GCR can be used without special treatment as long as the system is compatible, which is guaranteed in the present case by the fact that the right-hand side is zero for all pressure unknowns. The convergence is indeed the same as that of GMRES or GCR applied to a regular matrix whose eigenvalues coincide with the nonzero eigenvalues of the original preconditioned matrix. Moreover, these eigenvalues satisfy the relations and bounds proved in Theorems 3.1, 4.3, and 4.4 and Corollary 4.5, reading $\underline{\nu}$ as the smallest *nonzero* eigenvalue of $M_S^{-1}S$ (the rank deficiency of B implying that S is only semidefinite). Note that right preconditioning corresponds to the versions of GMRES and GCR that minimize the residual of the original linear system, and, regarding GCR, is equivalent to the standard preconditioning implementation in [42].

Now, for the stationary Stokes problem, it is known that the Schur complement $S = B\,A^{-1}B^T$ is spectrally equivalent to the identity when using finite difference approximations. Hence we may select

$$M_S = \omega^{-1}\,I \ ,$$

and numerical computation indeed shows that

$$\underline{\nu} \geq 0.2\,\omega \qquad \text{and} \qquad \overline{\nu} = \omega \ , \tag{5.2}$$

where $\underline{\nu}$ denotes the smallest *nonzero* eigenvalue of $M_S^{-1}S$. On the other hand, A is formed of two diagonal blocks, each of them being the five point finite difference approximation of the Laplace operator acting on one of the velocity components. Hence the conditioning of A depends on h, and a more sophisticated preconditioning approach is welcome, with multigrid methods being good candidates. For convenience, we selected the aggregation-based algebraic multigrid method (AGMG) from [26, 21, 28]. Indeed, a black box code is available with a MATLAB interface [23]; hence no further tuning or coding is needed. For relatively small matrix sizes (in the present example, as long as $h > 1/35$), the procedure uses only two levels. Then it follows from the algebraic properties of the preconditioner that

$$\overline{\mu} = 1$$

(see, e.g., [25, eq. (39)]), whereas numerical computation reveals that

$$\underline{\mu} \geq 0.4 \ .$$

For larger matrices, the preconditioner is based on the same two level method, but "inner" coarse systems are solved iteratively, in fact with the same two level method again, which is thus used recursively. Because these "inner" solves are accelerated with the "flexible" conjugate gradient method [24], the so defined preconditioner varies slightly from one application to the next. Then, the above estimates still hold, but only approximately, and should be interpreted with care since the preconditioner is on the whole a nonlinear operator.

Once M_A and M_S have been chosen, all preconditioners introduced in sections 1 and 2 are properly defined. For $h = 1/32$ and $\omega = 1$, we depict in Figure 3 the associated eigenvalue distribution. We also represent the limits provided by the theory. One sees that Theorem 4.3 accurately predicts the location of both real and nonreal eigenvalues, and one may also check the complementary role played by Theorem 2 in [39], as extended to other preconditioners in Theorem 4.4. One also sees the importance of the parameter ζ from (4.15) in controlling the imaginary extension of the eigenvalues: there are eigenvalues lying exactly on the line $|\lambda - \zeta| = \zeta$ for all preconditioners but M_f, and the improvement observed going from block diagonal to block triangular preconditioning is due largely to the decrease of ζ, the bounds on the real part remaining roughly the same.

As already discussed, the scaling of M_A plays an important role for block triangular and Uzawa preconditioners. With $\overline{\mu} \leq 1$, we have the appealing result of Corollary 4.5, but, on the other hand, if one rescales the preconditioner for A to have $\underline{\mu} \geq 1$, all eigenvalues are real, which may also be attractive, allowing us to use conjugate gradient methods in nonstandard inner products [12, 32, 44]. To investigate this, we rescaled the algebraic multigrid preconditioner by a factor α, entailing that

$$\underline{\mu} \geq 0.4\,\alpha \qquad \text{and} \qquad \overline{\mu} = \alpha\,.$$

The theory predicts that increasing α moves nonreal eigenvalues closer to the real axis until the point where they are forbidden (for $\alpha \geq 2.5$) but at the same time allows the real eigenvalues to spread over the real axis (see Figure 2). This is illustrated in Figure 4. In the left column of figures, we proceed as for Figure 3, plotting the spectrum together with the limits provided by the theory. One sees that the bounds remain accurate in all considered situations. In the right column of figures, we rescaled the spectrum to represent the situation that occurs when optimal scaling is applied

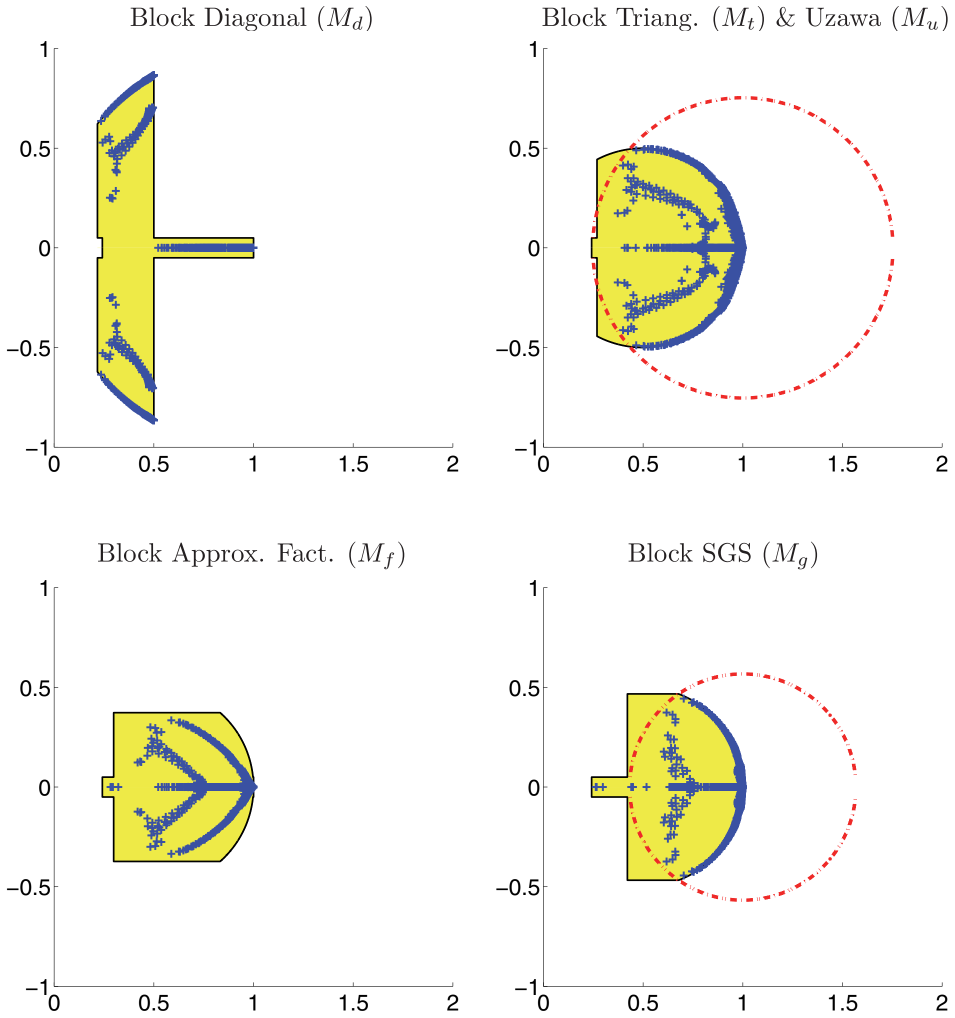


FIG. 3. +: *eigenvalues of the preconditioned matrix for* $h = 1/32$ *and* $\omega = 1$; —: *limit of the region defined by the inequalities in Theorem* 4.3 *(horizontal lines close to the real axis indicate regions where in fact only real eigenvalues are permitted);* - - -: *limit on nonreal eigenvalues provided by* [39, *Theorem* 2] *(see Theorem* 4.4*).*

to the preconditioner M_t or M_u, "optimal" meaning in such a way that the spectral radius ρ of the associated iteration matrix is minimized. We also graphically illustrate this spectral radius, plotting (with the symbol ||||) the circle of center 1 and radius ρ that contains all eigenvalues.

In Table 2,[5] we report the number of iterations actually needed to reduce the residual relative error by 10^{-6}, testing larger problem sizes and also different values of ω; results are not reported for M_g, M_{t_2}, and M_{u_2} with $\alpha = 2.5$ because the basic condition $\rho_A < 1$ is then violated, implying that $\widetilde{M}_A$ is in fact not positive definite and is therefore no longer a sensible preconditioner for A. The block approximate

[5] The code allowing one to reproduce the results reported in this table and in Table 3 is provided as supplementary material through the electronic version of the journal.

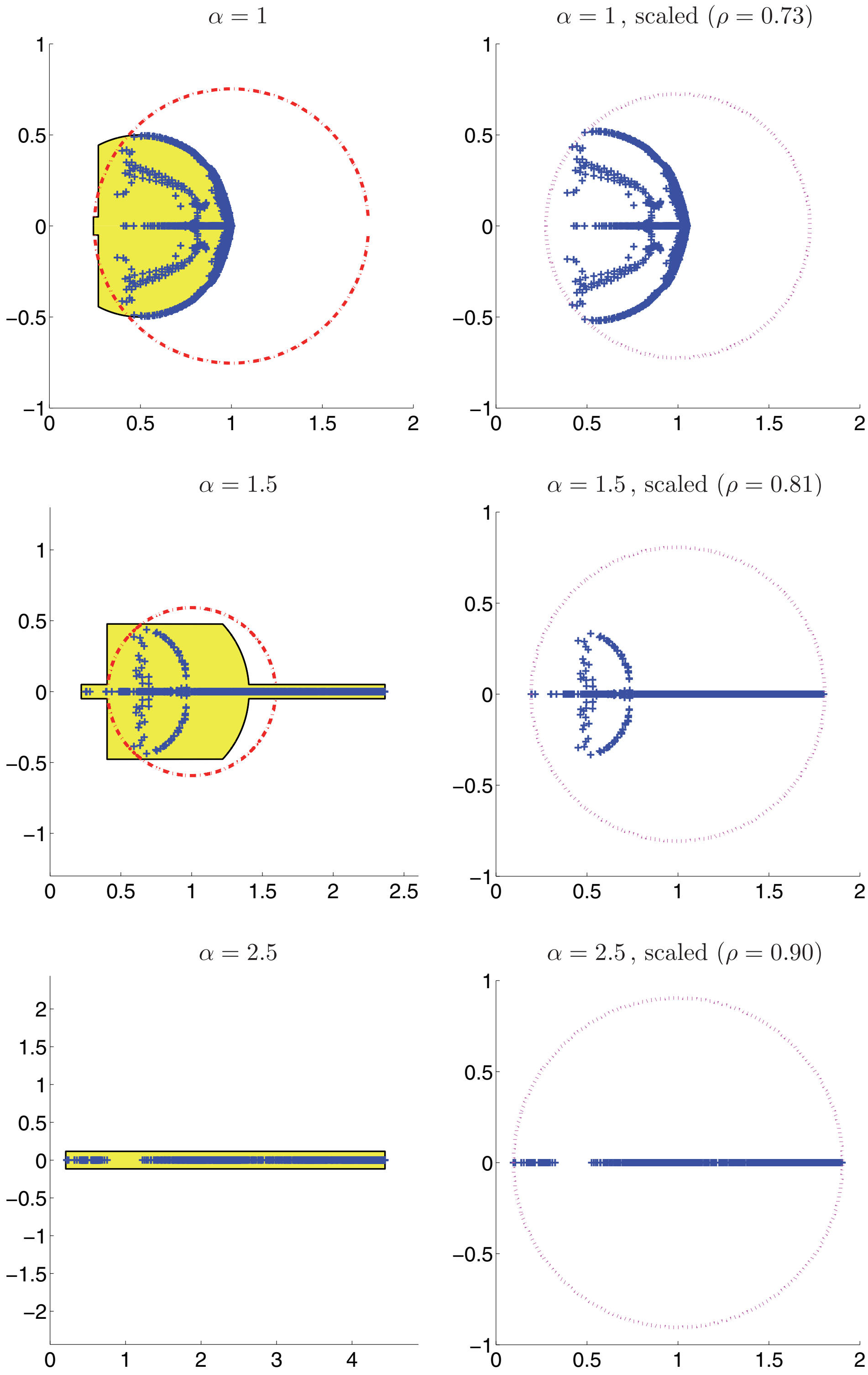


FIG. 4. *Left: Spectrum of the preconditioned matrix for block triangular and inexact Uzawa preconditioners ($h = 1/32$ and $\omega = 1$). Right: Rescaled spectrum.*

Table 2

Number of iterations needed to reduce the relative residual error by 10^{-6} *; MINRES is used for the positive definite preconditioner* M_+ *from* (3.1)*, and GCR*(15) *for all other preconditioners,* (15) *meaning that the process is restarted every* 15 *iterations.*

		$\omega = 1$			$\omega = 1.5$			$\omega = 4$		
	α	1	1.5	2.5	1	1.5	2.5	1	1.5	2.5
$h^{-1} = 32$										
Block diag.$^{(+)}$	(M_+)	44	43	42	47	44	43	53	49	47
Block diag.	(M_d)	59	58	57	60	59	57	66	59	59
Block triang.	(M_t)	27	28	35	28	27	34	34	28	33
Inexact Uzawa	(M_u)	29	30	33	29	29	31	35	28	34
Block fact.	(M_g)	20	21	50	20	20	58	23	22	130
Block SGS	(M_f)	20	19		21	19		21	20	
Block triang.$^{(2)}$	(M_{t_2})	20	19		19	19		21	19	
Inexact Uzawa$^{(2)}$	(M_{u_2})	22	20		22	20		23	21	
$h^{-1} = 512$										
Block diag.$^{(+)}$	(M_+)	62	59	57	63	62	59	71	68	64
Block diag.	(M_d)	127	89	99	152	127	75	126	157	164
Block triang.	(M_t)	54	47	56	53	46	52	59	45	46
Inexact Uzawa	(M_u)	60	61	53	59	60	57	58	56	55
Block fact.	(M_g)	33	29	89	28	28	96	32	40	169
Block SGS	(M_f)	29	23		27	23		28	24	
Block triang.$^{(2)}$	(M_{t_2})	28	23		28	23		32	30	
Inexact Uzawa$^{(2)}$	(M_{u_2})	35	26		30	24		30	24	

factorization M_f is still well defined when $\rho_A > 1$, although our analysis does not apply any longer in this case, which is reflected by the much larger number of iterations needed than with other values of α.

One sees that the hierarchy of the preconditioners is as expected from the theory, with slight differences between the variants leading to the same eigenvalue distribution, to be explained by the many other features that influence the convergence, such as "nonnormality" effects [41, Chapters 25 and 26]. Further tests show that, according to the analysis in [17], the indefinite block diagonal preconditioner M_d becomes as good as the positive definite preconditioner M_+ if the restart parameter is increased sufficiently; i.e., the advantage of M_+, as expected from item (1) of Theorem 3.1, mainly comes from the global optimality of MINRES.

The scalability of M_g, M_{t_2}, and M_{u_2} reflects well that of M_A: solving a system with A alone requires from 10 iterations for $h^{-1} = 32$ (two level variant) to 12 for $h^{-1} = 256, 512, 1024$ (multilevel "variable" preconditioner). The number of iterations increases in a bigger proportion for the triangular preconditioners M_t and M_u, displaying their greater sensitivity to the quality of the used approximation for A. This sensitivity is also expected from our analysis; compare (4.38) with (4.39) when $\overline{\mu} \leq 1$, and see how fast η and $\widetilde{\nu}$ may grow with $\overline{\mu}$ otherwise.

In Table 3, we report the results obtained on finer meshes when fixing $\alpha = 1.5$ and $\omega = 1$ (which is close to optimal in all cases). This allows us to further check the near optimality of all preconditioners except M_d and perhaps M_f. Timing results suggest that this near optimality also holds with respect to time: with about 4 times more unknowns, the elapsed time is multiplied by a factor only slightly larger than 4. We also present some results obtained when defining M_A with a classical algebraic multigrid (AMG) algorithm (along the lines of the seminal works by Brandt, McCormick, and Ruge [9] and Ruge and Stüben [34]). We selected the implementation

Table 3

Number of iterations and time needed to reduce the relative residual error by 10^{-6}*. MINRES is used for the positive definite preconditioner* M_+ *from* (3.1)*, and GCR(*15*) for all other preconditioners.* $\mathcal{T}_{\text{sol}}$ *is the time spent during the solution phase, and* $\mathcal{T}_{\text{setup}}$ *is the time needed to set up the preconditioner (essentially the time needed to build* M_A*). All these times are elapsed wall clock times reported in seconds.* $\omega = 1$ *is used in all cases, whereas the scaling factor used for* M_A *is* $\alpha = 1.5$ *in the case of AGMG and* 1 *(i.e., no scaling) in the case of AMG (IFISS).*

Preconditioner for A	AGMG				AMG (IFISS)	
h^{-1}	512		1024		512	
$\mathcal{T}_{\text{setup}}$		0.37		1.31		4476.8
	#it	$\mathcal{T}_{\text{sol}}$	#it	$\mathcal{T}_{\text{sol}}$	#it	$\mathcal{T}_{\text{sol}}$
Block diag.$^{(+)}$ (M_+)	59	12.3	62	53.6	39	22.2
Block diag. (M_d)	89	21.5	150	153.0	54	32.0
Block triang. (M_t)	47	11.7	58	61.4	18	10.8
Inexact Uzawa (M_u)	61	15.6	57	61.2	20	12.1
Block fact. (M_f)	29	12.1	37	64.0	14	15.8
Block SGS (M_g)	23	9.2	26	43.7	14	15.6
Block triang.$^{(2)}$ (M_{t_2})	23	8.8	26	42.8	13	14.3
Inexact Uzawa$^{(2)}$ (M_{u_2})	26	10.2	28	46.6	17	18.7

provided with the Incompressible Flow & Iterative Solver Software (IFISS) software package [37], both for convenience (as AGMG, it is callable from MATLAB) and because using it in combination with M_+ and MINRES is the default solver for Stokes problems in IFISS. Here we report the results obtained with $\alpha = 1$, which appear optimal or close to optimal in all cases. The setup time for $h^{-1} = 512$ is huge, and we were not able to achieve the setup for $h^{-1} = 1024$; this is clearly due to some implementation issue as, typically, the setup of classical AMG methods, although significantly more costly than that of AGMG, never requires much more time than the solution phase [22]. Leaving these implementation issues aside, as expected from the comparison in [22], the classical AMG method delivers a closer approximation to A (hence the number of iterations is reduced), but this is hardly cost effective because the computational cost per iteration step is significantly larger.

The timing results also allow us to assess the effect of the additional costs incurred when using GCR instead of a short recurrence method like MINRES: comparing the line for M_d and the line for M_+, one may check that the time per iteration is roughly 15% less with MINRES when using AGMG, and roughly 4% less when using AMG. This difference comes from the cost of the preconditioner: the higher it is, the less important is the weight of the operations associated with the Krylov subspace solver.

In fact, even with AGMG, the solution time tends to be dominated by the application of the preconditioner for A, and it is interesting to observe that, in all three reported cases (AGMG with $h^{-1} = 512$ or $h^{-1} = 1024$, and AMG with $h^{-1} = 512$), the most effective block preconditioner is finally the one that requires the least application of M_A (taking into account that M_f, M_g, M_{t_2}, and M_{u_2} involve two applications of M_A per iteration step). On the other hand, the results with AMG confirm a remark already made in section 4, that when comparing (4.38) with (4.39), additional costs incurred with M_g and related variants likely do not pay off when the preconditioner for A is much better than that for S.

6. Conclusions. We have developed the spectral analysis of a class of preconditioners for symmetric saddle point matrices. The eigenvalue bounds depend only on the extremal eigenvalues (1.3), (1.4) associated with the used approximations for the top left block A and the (negative) Schur complement S. For all the considered preconditioners (i.e., for (1.5), (1.6), (1.7), (1.8), (2.3), (2.4)), these bounds prove that the eigenvalues are located in a confined region of the right half of the complex plane. Moreover, except for the block diagonal preconditioner (defined by (1.5)), this region is clustered around 1, and its area converges smoothly towards zero when the main parameters in (1.3), (1.4) converge towards 1 (that is, when $M_A \to A$ and $M_S \to S$).

Our analysis also allows a comparison of the different types of preconditioners. First, we proved that block triangular (1.6) and inexact Uzawa preconditioners (1.7) lead to identical spectra. Next, a connection can also be made between these triangular preconditioners and the symmetrized block SGS preconditioner (2.3) via the spectral equivalence of (2.3) with "enhanced" triangular preconditioners (2.4). Finally, our bounds are accurate enough to allow an insightful discussion of the relative performances of the preconditioners.

From this viewpoint, block diagonal preconditioners appear less attractive than triangular ones, as they lead to a less clustered eigenvalue distribution for essentially the same cost. It is, however, worth noting in this respect that the indefinite variant (1.5) mainly investigated here is often outperformed by the positive definite variant (3.1). More precisely, we have shown that the eigenvalue distribution associated with (3.1) cannot really be better, but it can be used in combination with a short recurrence and globally optimal Krylov accelerator (MINRES).

On the other hand, compared with triangular preconditioners (1.6), (1.7), symmetrized preconditioners like block SGS (2.3) and block approximate factorization preconditioners (1.8) can be cost effective only when the quality of the approximation used for the top left block A is not much better than that used for the Schur complement S.

Finally, our results also give insight into the role of the scaling of the used approximation M_A for A. For the triangular preconditioners (1.6), (1.7), a scaling such that all the eigenvalues of $M_A^{-1}A$ are not less than 1 guarantees that the preconditioned matrix has only real eigenvalues, but at the expense of an adverse effect on their clustering around 1, which our bounds may help to evaluate. In contrast with this, if the scaling is such that all the eigenvalues of $M_A^{-1}A$ are not larger than 1, then the preconditioned matrix has both real and nonreal eigenvalues, but all lie within a disk around 1 whose radius satisfies an appealing bound. On the other hand, regarding symmetrized preconditioners (1.8), (2.3), it is best to choose the scaling in such a way that the spectral radius of $I - M_A^{-1}A$ is minimized, and it is mandatory to check that all the eigenvalues of $M_A^{-1}A$ remain significant below 2.

Appendix: Rank deficient B. Here we consider the case of a rank deficiency in B that is not compensated for by the positive definiteness of C, i.e., the case where there exist nontrivial vectors $\mathbf{e}$ such that $B^T\mathbf{e} = \mathbf{0}$ and $C\mathbf{e} = \mathbf{0}$. Denoting $\mathcal{N}_B$ as the null space of B^T, for the sake of clarity we restrict the analysis to the case where $C\,\mathbf{e} = \mathbf{0}$ for all $\mathbf{e} \in \mathcal{N}_B$. Then K is singular, and its null space is the subspace of all vectors of the form $(\mathbf{0}^T \;\; \mathbf{e}^T)^T$ with $\mathbf{e} \in \mathcal{N}_B$. For the model problem of section 5, these assumptions are satisfied with $\mathcal{N}_B$ equal to the one-dimensional subspace spanned by the constant vector $\mathbf{1}$ of length m.

Now, it is well known that solving a singular linear system with a Krylov subspace method requires no special treatment as long as the system is compatible; see, e.g., [15, section 8.3] for a detailed discussion in the context of the Navier–Stokes equations.

Further, an interesting analysis is developed in [18], which shows that the convergence of GMRES and GCR is actually the same as that of the same method applied to the problem orthogonally projected onto the range of the linear system. This, in particular, allows one to recover the convergence condition stated in [10], which is that the range and null spaces of the system matrix have trivial intersection: this is indeed the necessary and sufficient condition for the projected matrix to be nonsingular. Moreover, its eigenvalues coincide then with the nonzero eigenvalues of the original matrix.

Now, observe that, since K is symmetric, its range is equal to the orthogonal complement of its null space. Then let V_C be an $m \times m_r$ orthonormal matrix whose columns form a basis of $\mathcal{N}_B^{\perp}$. The system matrix orthogonally projected onto the range of K is thus

$$K_r = \begin{pmatrix} I & \\ & V_C^T \end{pmatrix} \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} I & \\ & V_C \end{pmatrix} = \begin{pmatrix} A & B^T V_C \\ V_C^T B & -V_C^T C\, V_C \end{pmatrix} . \tag{A.1}$$

However, here we are interested in preconditioning; hence we need to consider the projected *preconditioned* matrix. In the proof of Theorem 3.1, we have seen that all preconditioners considered in this work are particular instances of

$$M = \begin{pmatrix} I_n & \\ B\, Y_A & I_m \end{pmatrix} \begin{pmatrix} \widehat{M}_A & \\ & -M_S \end{pmatrix} \begin{pmatrix} I_n & Z_A B^T \\ & I_m \end{pmatrix} ,$$

where Y_A , Z_A , $\widehat{M}_A$ are related to A and M_A via (3.5), (3.6), (2.1). On the other hand, it follows from the definition of V_C that $V_C V_C^T$ is the orthogonal projector onto $\mathcal{N}_B^{\perp}$, and therefore that $I - V_C V_C^T$ is the orthogonal projector onto $\mathcal{N}_B$. Since $\mathcal{N}_B$ is included in the null space of both B^T and C , this implies $B^T(I - V_C V_C^T) = 0$ and $C(I - V_C V_C^T) = 0$, and hence $B^T = B^T V_C V_C^T$, $C = C\, V_C V_C^T$. Moreover, using right preconditioning, the preconditioned matrix has the same range as K and the projection operator to be considered is thus the same as in (A.1). The projection of the preconditioned matrix $K\, M^{-1}$ can therefore be written as

$$\begin{aligned}
&\begin{pmatrix} I & \\ & V_C^T \end{pmatrix} \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} I_n & -Z_A B^T \\ & I_m \end{pmatrix} \begin{pmatrix} \widehat{M}_A^{-1} & \\ & -M_S^{-1} \end{pmatrix} \begin{pmatrix} I_n & \\ -B\, Y_A & I_m \end{pmatrix} \begin{pmatrix} I & \\ & V_C \end{pmatrix} \\
&\quad = \begin{pmatrix} A & B^T V_C V_C^T \\ V_C^T B & -V_C^T C\, V_C V_C^T \end{pmatrix} \begin{pmatrix} I_n & -Z_A B^T V_C V_C^T \\ & I_m \end{pmatrix} \\
&\qquad\qquad \begin{pmatrix} \widehat{M}_A^{-1} & \\ & -M_S^{-1} V_C \end{pmatrix} \begin{pmatrix} I_n & \\ -V_C^T B\, Y_A & I_m \end{pmatrix} \\
&\quad = \begin{pmatrix} A & B^T V_C \\ V_C^T B & -V_C^T C\, V_C \end{pmatrix} \begin{pmatrix} I_n & -Z_A B^T V_C \\ & I_m \end{pmatrix} \\
&\qquad\qquad \begin{pmatrix} \widehat{M}_A^{-1} & \\ & -V_C^T M_S^{-1} V_C \end{pmatrix} \begin{pmatrix} I_n & \\ -V_C^T B\, Y_A & I_m \end{pmatrix} .
\end{aligned}$$

Hence the projected matrix, i.e., the matrix that governs the convergence of the iterative process, is still a matrix of the form (1.1) preconditioned with the same technique as that applied to the original singular system, but, comparing to the latter, B is exchanged for $V_C^T B$, C for $V_C^T C\, V_C$, and M_S for $\left(V_C^T M_S^{-1} V_C\right)^{-1}$, whereas A and M_A remain unchanged. Since $B^T\, V_C$ has full rank by construction, it follows that the projected matrix is indeed nonsingular; i.e., the convergence condition is met. Moreover, its eigenvalues, which coincide with the nonzero eigenvalues of $K\, M^{-1}$, can still be analyzed in light of Theorems 3.1, 4.3, and 4.4 and Corollary 4.5, only exchanging $\underline{\nu}$ and $\overline{\nu}$ for, respectively, the smallest and the largest eigenvalue of $\big(V_C^T M_S^{-1} V_C\big)\, V_C^T S\, V_C$, the Schur complement in the transformed matrix being

$$V_C^T C\, V_C + \big(V_C^T B\big)\, A^{-1} \left(B^T V_C\right) = V_C^T S\, V_C \ .$$

Further, for this smallest eigenvalue, one finds, noting that $\mathcal{R}\,(V_C) = \mathcal{N}_B^{\perp} = \mathcal{R}(S)$,

$$\begin{aligned}
\lambda_{\min}\big(\, \big(V_C^T M_S^{-1} V_C\big)\, V_C^T S\, V_C\big) &= \min_{\mathbf{v}\in\mathbb{R}^{m_r}} \frac{\mathbf{v}^T \left(V_C^T M_S^{-1} V_C\right) V_C^T S\, V_C \left(V_C^T M_S^{-1} V_C\right) \mathbf{v}}{\mathbf{v}^T \left(V_C^T M_S^{-1} V_C\right) \mathbf{v}} \\
&= \min_{\mathbf{v}\in\mathbb{R}^{m_r}} \frac{\mathbf{v}^T \left(V_C^T M_S^{-1} S\, M_S^{-1} V_C\right) \mathbf{v}}{\mathbf{v}^T \left(V_C^T M_S^{-1} V_C\right) \mathbf{v}} \\
&= \min_{\mathbf{w}\in\mathcal{R}(S)} \frac{\mathbf{w}^T M_S^{-1} S\, M_S^{-1} \mathbf{w}}{\mathbf{w}^T M_S^{-1} \mathbf{w}} \\
&= \min_{\mathbf{z}\in\mathcal{R}(M_S^{-1/2} S\, M_S^{-1/2})} \frac{\mathbf{z}^T M_S^{-1/2} S\, M_S^{-1/2} \mathbf{z}}{\mathbf{z}^T\, \mathbf{z}} \\
&= \min_{\nu\in\sigma(M_S^{-1} S)\,,\, \nu\neq 0} \nu \ . \qquad \text{(A.2)}
\end{aligned}$$

Similarly, one finds

$$\lambda_{\max}\big(\, \big(V_C^T M_S^{-1} V_C\big)\, V_C^T S\, V_C\big) = \max_{\nu\in\sigma(M_S^{-1} S)\,,\, \nu\neq 0} \nu \ = \overline{\nu} \ . \tag{A.3}$$

Hence it suffices to read $\underline{\nu}$ as the smallest *nonzero* eigenvalue of $M_S^{-1} S$, and all inequalities provided in Theorems 4.3 and 4.4 and Corollary 4.5 become valid bounds for the singular case as well.

Acknowledgments. Artem Napov and two anonymous referees are acknowledged for several useful comments and suggestions following a careful reading.

REFERENCES

[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994.

[2] O. Axelsson and M. Neytcheva, *Eigenvalue estimates for preconditioned saddle point matrices*, Numer. Linear Algebra Appl., 13 (2006), pp. 339–360.

[3] R. E. Bank, B. D. Welfert, and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math., 56 (1990), pp. 645–666.

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

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

[6] L. Bergamaschi, *On eigenvalue distribution of constraint-preconditioned symmetric saddle point matrices*, Numer. Linear Algebra Appl., 19 (2012), pp. 754–772.

[7] 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.

[8] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of the inexact Uzawa algorithm for saddle point problems*, SIAM J. Numer. Anal., 34 (1997), pp. 1072–1092.

[9] A. Brandt, S. F. McCormick, and J. W. Ruge, *Algebraic multigrid (AMG) for sparse matrix equations*, in Sparsity and Its Application, D. J. Evans, ed., Cambridge University Press, Cambridge, UK, 1984, pp. 257–284.

[10] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.

[11] E. de Sturler and J. Liesen, *Block-diagonal and constraint preconditioners for nonsymmetric indefinite linear systems. Part* I: *Theory*, SIAM J. Sci. Comput., 26 (2005), pp. 1598–1619.

[12] H. S. Dollar, N. I. M. Gould, M. Stoll, and A. J. Wathen, *Preconditioning saddle-point systems with applications in optimization*, SIAM J. Sci. Comput., 32 (2010), pp. 249–270.

[13] 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.

[14] H. Elman, V. Howle, J. Shadid, R. Shuttleworth, and R. Tuminaro, *A taxonomy and comparison of parallel block multi-level preconditioners for the incompressible Navier-Stokes equations*, J. Comput. Physics, 227 (2008), pp. 1790–1808.

[15] H. Elman, D. Silvester, and A. Wathen, *Finite Elements and Fast Iterative Solvers*, Oxford University Press, Oxford, UK, 2005.

[16] H. C. Elman and G. H. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1645–1661.

[17] B. Fischer, A. Ramage, D. Silvester, and A. Wathen, *Minimum residual methods for augmented systems*, BIT, 38 (1998), pp. 527–543.

[18] K. Hayami and M. Sugihara, *A geometric view of Krylov subspace methods on singular systems*, Numer. Linear Algebra Appl., 18 (2011), pp. 449–469.

[19] C. Keller, N. I. M. Gould, and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1300–1317.

[20] 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.

[21] A. Napov and Y. Notay, *An algebraic multigrid method with guaranteed convergence rate*, SIAM J. Sci. Comput., 34 (2012), pp. A1079–A1109.

[22] A. Napov and Y. Notay, *Algebraic Multigrid for Moderate Order Finite Elements*, Technical Report GANMN 13–02, Université Libre de Bruxelles, Brussels, Belgium, 2013; available online at http://homepages.ulb.ac.be/∼ynotay.

[23] Y. Notay, *AGMG*, software and documentation, http://homepages.ulb.ac.be/∼ynotay/AGMG (July 2, 2012).

[24] Y. Notay, *Flexible conjugate gradients*, SIAM J. Sci. Comput., 22 (2000), pp. 1444–1460.

[25] Y. Notay, *Algebraic multigrid and algebraic multilevel methods: A theoretical comparison*, Numer. Linear Algebra Appl., 12 (2005), pp. 419–451.

[26] Y. Notay, *An aggregation-based algebraic multigrid method*, Electron. Trans. Numer. Anal., 37 (2010), pp. 123–146.

[27] Y. Notay, *Algebraic analysis of two-grid methods: The nonsymmetric case*, Numer. Linear Algebra Appl., 17 (2010), pp. 73–96.

[28] Y. Notay, *Aggregation-based algebraic multigrid for convection-diffusion equations*, SIAM J. Sci. Comput., 34 (2012), pp. A2288–A2316.

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

[30] S. V. Patankar and D. B. Spalding, *A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows*, Int. J. Heat Mass Transfer, 15 (1972), pp. 1787–1806.

[31] I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numer. Linear Algebra Appl., 7 (2000), pp. 585–616.

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

[33] T. Rees and M. Stoll, *Block-triangular preconditioners for PDE-constrained optimization*, Numer. Linear Algebra Appl., 17 (2010), pp. 977–996.

[34] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, Frontiers Appl. Math. 3, S. F. McCormick, ed., SIAM, Philadelphia, 1987, pp. 73–130.

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

[36] D. Sesana and V. Simoncini, *Spectral analysis of inexact constraint preconditioning for symmetric saddle point matrices*, Linear Algebra Appl., 438 (2013), pp. 2683–2700.

[37] D. Silvester, H. Elman, and A. Ramage, *Incompressible Flow & Iterative Solver Software (IFISS), Version* 3.2, 2012, http://www.manchester.ac.uk/ifiss/.

[38] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems Part* II: *Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.

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

[40] L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[41] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press, Princeton, NJ, 2005.

[42] H. A. Van der Vorst and C. Vuik, *GMRESR: A family of nested GMRES methods*, Numer. Linear Algebra Appl., 1 (1994), pp. 369–386.

[43] C. Vuik, A. Saghir, and G. P. Boerstoel, *The Krylov accelerated SIMPLE(R) method for flow problems in industrial furnaces*, Internat. J. Numer. Methods Fluids, 33 (2000), pp. 1027–1040.

[44] W. Zulehner, *Analysis of iterative methods for saddle point problems: A unified approach*, Math. Comp., 71 (2002), pp. 479–505.