SIAM J. MATRIX ANAL. APPL.
Vol. 39, No. 2, pp. 902–921

ITERATIVE METHODS FOR DOUBLE SADDLE POINT SYSTEMS*

FATEMEH PANJEH ALI BEIK† AND MICHELE BENZI‡

Abstract. We consider the iterative solution of a class of linear systems with double saddle point structure. Several block preconditioners for Krylov subspace methods are described and analyzed. We derive some bounds for the eigenvalues of preconditioned matrices and present results of numerical experiments using test problems from two different applications: the potential fluid flow problem and the modeling of liquid crystals directors.

Key words. saddle point problems, preconditioning, Krylov methods, finite elements, potential fluid flow, liquid crystals

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

DOI. 10.1137/17M1121226

1. Introduction.

In this paper we consider iterative methods for solving large, sparse, linear systems of equations of the form

$$\mathscr{A}u \equiv \begin{bmatrix} A & B^T & C^T \\ B & 0 & 0 \\ C & 0 & -D \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \equiv b, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite (SPD) and $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, and $D \in \mathbb{R}^{p\times p}$ are symmetric positive semidefinite (SPS) and possibly zero. Throughout the paper we assume that $n \geq m + p$.

Linear systems of the form (1.1) arise frequently from mixed and mixed-hybrid formulations of second-order elliptic equations [7, sect. 7.2], [12] and elasticity [7, sect. 9.3.1] problems. Numerical methods in constrained optimization [13, 14] and liquid crystal modeling [18] also lead to sequences of linear systems of the type (1.1). We further mention that finite element models of certain incompressible flow problems arising in the analysis of non-Newtonian fluids and in geophysics lead to large linear systems with coefficient matrices of the form

$$\mathscr{B} = \begin{bmatrix} A & C^T & B^T \\ C & -D & 0 \\ B & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathscr{C} = \begin{bmatrix} -D & C & 0 \\ C^T & A & B^T \\ 0 & B & 0 \end{bmatrix};$$

see, e.g., [3] and [8], respectively. It is easy to see that both $\mathscr{B}$ and $\mathscr{C}$ can be brought into the same form as matrix $\mathscr{A}$ in (1.1) by means of symmetric permutations (row and column interchanges).

It is important to observe that matrix $\mathscr{A}$ can be regarded as a 2×2 block matrix in two different ways, according to which of the following partitioning strategies are used:

*Received by the editors March 16, 2017; accepted for publication (in revised form) February 9, 2018; published electronically May 24, 2018.
http://www.siam.org/journals/simax/39-2/M112122.html
Funding: The work of the second author was supported in part by NSF grant DMS-1418889.
†Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran (f.beik@vru.ac.ir).
‡Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322 (mbenzi@emory.edu).

$$\mathscr{A} = \left[\begin{array}{c|cc} A & B^T & C^T \\ \hline B & 0 & 0 \\ C & 0 & -D \end{array}\right] \quad \text{or} \quad \mathscr{A} = \left[\begin{array}{cc|c} A & B^T & C^T \\ B & 0 & 0 \\ \hline C & 0 & -D \end{array}\right]. \tag{1.2}$$

The first partitioning highlights the fact that problem (1.1) can in principle be treated as a "standard" saddle point problem, possibly stabilized (or regularized) when $D \neq 0$; see, e.g., [6]. On the other hand, the second partitioning shows that (1.1) can also be regarded as having a *double* saddle point structure, since the (1,1) block is itself the coefficient matrix of a saddle point problem; see, e.g., [18]. While in this paper we make use of both partitionings, we are especially interested in studying solvers and preconditioners that make explicit use of the 3×3 block structure of $\mathscr{A}$.

The paper is organized as follows. In section 2 we give a detailed discussion of conditions that ensure the unique solvability of (1.1). Block preconditioners for Krylov-type methods are discussed and analyzed in section 3. Illustrative numerical experiments are presented in section 4. Section 5 contains brief concluding remarks.

2. Solvability conditions. In this section we investigate the solvability of (1.1) under various assumptions on the blocks A, B, C, and D. Invertibility conditions for the coefficient matrix $\mathscr{A}$ in (1.1) under different assumptions on the blocks can be found scattered in the literature; see, for instance, [6], [7, Chapter 3], as well as [4] and [10] for eigenvalue bounds. While our results overlap in part with known ones, we find it useful to collect all the needed statements with complete proofs here, also in order to make the paper self-contained. In the following, for a real square matrix A we write $A \succ 0$ ($A \succcurlyeq 0$) if A is SPD (respectively, SPS) and $A \succ B$ ($A \succcurlyeq B$) if A and B are real symmetric matrices such that $A - B$ is SPD (respectively, SPS). Moreover, we write $(x; y; z)$ to denote the vector $(x^T, y^T, z^T)^T$.

The following theorem provides a necessary and sufficient condition for the invertibility of the matrix $\mathscr{A}$ in the case that the $(1,1)$ and $(3,3)$ blocks are both SPD.

PROPOSITION 2.1. *Assume that $A \succ 0$ and $D \succ 0$. Then matrix $\mathscr{A}$ is invertible if and only if B has full row rank.*

Proof. Let B have full row rank and assume that $\mathscr{A}u = 0$ for $u = (x; y; z)$, i.e.,

$$Ax + B^T y + C^T z = 0, \tag{2.1}$$

$$Bx = 0, \tag{2.2}$$

$$Cx - Dz = 0. \tag{2.3}$$

If $x = 0$, then (2.3) implies $z = 0$ (since $D \succ 0$) and thus from (2.1) we conclude that $y = 0$, since B^T has full column rank. Hence, $u = 0$. If $z = 0$, then from (2.1) and (2.2) we obtain $0 = Bx = -BA^{-1}B^T y$ and thus $y = 0$ since $BA^{-1}B^T$ is SPD. Hence, $x = 0$ and thus again it must be $u = 0$. Let us assume now that both the vectors x and z are nonzero. Multiplying (2.1) by x^T from the left, we find

$$x^T Ax + x^T B^T y + x^T C^T z = 0. \tag{2.4}$$

From (2.3), it can be seen that $z^T Cx = z^T Dz$. Substituting $z^T Cx = z^T Dz$ and (2.2) into (2.4), we have

$$x^T Ax = -z^T Dz. \tag{2.5}$$

In view of the positive definiteness of A and D, the preceding equality implies that $x = 0$ and $z = 0$ which shows that $u = 0$.

Conversely, suppose that $\mathscr{A}$ is nonsingular. Let $y \in \mathbb{R}^m$ be such that $B^T y = 0$. Setting $u = (0; y; 0)$, we obtain $\mathscr{A}u = 0$. In view of the invertibility of $\mathscr{A}$, we conclude that $y = 0$. This completes the proof. ∎

Next, we consider relaxing the assumptions of Proposition 2.1 so that either $A \succcurlyeq 0$ or $D \succcurlyeq 0$. In the following theorem we establish sufficient conditions which guarantee the nonsingularity of $\mathscr{A}$. We further show that some of these conditions are also necessary.

THEOREM 2.2. *Let A and $D \neq 0$ be SPS matrices. Assume that at least one of them is positive definite and B has full row rank. Then the following statements hold*

Case 1. *Suppose that $A \succ 0$ and $D \succcurlyeq 0$.*

- *If $\ker(C^T) \cap \ker(D) = \{0\}$ and $\mathrm{range}(B^T) \cap \mathrm{range}(C^T) = \{0\}$, then $\mathscr{A}$ is nonsingular.*
- *If $\mathscr{A}$ is nonsingular then $\ker(C^T) \cap \ker(D) = \{0\}$.*

Case 2. *Suppose that $A \succcurlyeq 0$ and $D \succ 0$. Then $\mathscr{A}$ is nonsingular if and only if $\ker(A) \cap \ker(B) \cap \ker(C) = \{0\}$.*

Proof. For clarity we divide the proof into two steps. In the first step, we show the validity of the stated sufficient conditions for the invertibility of $\mathscr{A}$ for both cases. In the second step, it is proved that in each case one of the conditions is also necessary.

Step I. Let $u = (x; y; z)$ be an arbitrary vector such that $\mathscr{A}u = 0$. We recall from the proof of Proposition 2.1 that relation (2.5) must hold true.

Let us first consider the case that $A \succ 0$. From (2.5), it can be seen that $x = 0$, hence $Dz = 0$ from (2.3). Note that $B^T y + C^T z = 0$ together with the assumption $\mathrm{range}(B^T) \cap \mathrm{range}(C^T) = \{0\}$ imply that $C^T z = 0$ and $B^T y = 0$. Since B^T has full column rank, $B^T y = 0$ implies $y = 0$. From $z \in \ker(C^T)$ and $Dz = 0$, we may immediately conclude from the assumption that $z = 0$, hence $u = 0$ and thus $\mathscr{A}$ is nonsingular.

For the second case, assume that $D \succ 0$. From (2.5), we can see that $z = 0$ since $A \succeq 0$. In addition $x^T Ax = 0$ which implies that $Ax = 0$, i.e., $x \in \ker(A)$. Since $\mathscr{A}u = 0$, we have $Bx = 0$ and $Cx = 0$, i.e., $x \in \ker(B)$ and $x \in \ker(C)$. Consequently, we deduce that $x = 0$ and therefore $y = 0$ in view of the fact that B^T has full column rank. Hence, $u = (x; y; z)$ is the zero vector, which shows the invertibility of $\mathscr{A}$.

Step II. Suppose that $\mathscr{A}$ is a nonsingular matrix.

Consider the case that $A \succ 0$. Assume there exists a nonzero vector $z \in \ker(C^T) \cap \ker(D)$. Then letting $u = (0; 0; z)$, we get $\mathscr{A}u = 0$, which is a contradiction. Hence, $\ker(C^T) \cap \ker(D) = \{0\}$ is a necessary condition for the invertibility of $\mathscr{A}$.

Finally, let us consider Case 2 and show that $\ker(A) \cap \ker(B) \cap \ker(C) = \{0\}$ is a necessary condition for the invertibility of $\mathscr{A}$. If there exists a nonzero vector $x \in \ker(A) \cap \ker(B) \cap \ker(C)$, then for $u = (x; 0; 0)$, we have $\mathscr{A}u = 0$, which is again a contradiction. Therefore, $\ker(A) \cap \ker(B) \cap \ker(C) = \{0\}$. ∎

It is worth noting that the sufficient condition $\mathrm{range}(B^T) \cap \mathrm{range}(C^T) = \{0\}$ given in Case 1 of Theorem 2.2 is not a necessary condition for $\mathscr{A}$ to be invertible in the case that $D \neq 0$. We illustrate this fact with the following simple example in which $\mathscr{A}$ is nonsingular and $\mathrm{range}(B^T) \cap \mathrm{range}(C^T) \neq \{0\}$. Consider the 8×8 matrix

$$\mathscr{A} = \begin{bmatrix} I_4 & B^T & C^T \\ B & 0 & 0 \\ C & 0 & -D \end{bmatrix},$$

where I_n stands for the $n \times n$ identity matrix, and the matrices B, C, and D are given as follows:

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.6}$$

Then matrix $\mathscr{A}$ is invertible but $\text{range}(B^T) \cap \text{range}(C^T) \neq \{0\}$.

The following proposition addresses the case where D is a zero matrix. We begin by noting that, in this case, a necessary condition for $\mathscr{A}$ to be invertible is that C has full row rank. Indeed, if there exists a nonzero vector z such that $C^T z = 0$, then $\mathscr{A}u = 0$ for $u = (0; 0; z) \neq 0$ and thus $\mathscr{A}$ cannot be invertible.

PROPOSITION 2.3. *Let $A \succ 0$ and assume that B and C have full row rank. Consider the linear system* (1.1) *with $D = 0$. Then* $\text{range}(B^T) \cap \text{range}(C^T) = \{0\}$ *is a necessary and sufficient condition for the coefficient matrix $\mathscr{A}$ to be invertible.*

Proof. As seen in the proof of Theorem 2.2, $\text{range}(B^T) \cap \text{range}(C^T) = \{0\}$ is a sufficient condition for invertibility of $\mathscr{A}$. Therefore we only need to show that it is also a necessary condition when $D = 0$ in (1.1). To this end, suppose that there exists a nonzero vector $v \in \text{range}(B^T) \cap \text{range}(C^T)$. As a result, $v = B^T y$ and $v = C^T z$ for some nonzero vectors y and z, and letting $u = (0; y; -z)$, we get $\mathscr{A}u = 0$, contradicting the invertibility of $\mathscr{A}$. Hence, it must be $\text{range}(B^T) \cap \text{range}(C^T) = \{0\}$. ∎

Remark 2.4. We stress that in the case $D = 0$, both B and C must have full row rank for $\mathscr{A}$ to be invertible. In contrast, in the case that $D \succeq 0$ and $D \neq 0$, only the matrix B is required to have full row rank while the matrix C can be rank deficient.

In the remainder of the paper we will always assume that $\mathscr{A}$ is nonsingular.

3. Preconditioning techniques. In this section we develop and analyze several block preconditioners to be used in conjunction with Krylov subspace methods to solve linear systems of equations of the form (1.1). The section is divided into two subsections which correspond to the two main cases $D = 0$ and $D \neq 0$, respectively.

3.1. Block preconditioners of the first type. In this part we discuss the eigenvalue distribution of the preconditioned matrices corresponding to the following block diagonal and block triangular preconditioners for solving systems of the form (1.1) with $D = 0$:

$$\mathscr{P}_{D} = \begin{bmatrix} A & 0 & 0 \\ 0 & BA^{-1}B^T & 0 \\ 0 & 0 & CA^{-1}C^T \end{bmatrix}, \quad \mathscr{P}_{T} = \begin{bmatrix} A & B^T & C^T \\ 0 & -BA^{-1}B^T & 0 \\ 0 & 0 & -CA^{-1}C^T \end{bmatrix},$$

$$\mathscr{P}_{GD} = \begin{bmatrix} A & 0 & 0 \\ 0 & BA^{-1}B^T & BA^{-1}C^T \\ 0 & CA^{-1}B^T & CA^{-1}C^T \end{bmatrix}, \quad \mathscr{P}_{GT,1} = \begin{bmatrix} A & 0 & 0 \\ B & -BA^{-1}B^T & -BA^{-1}C^T \\ C & -CA^{-1}B^T & -CA^{-1}C^T \end{bmatrix}, \tag{3.1}$$

and

$$\mathscr{P}_{GT,2} = \begin{bmatrix} A & B^T & 0 \\ B & 0 & 0 \\ C & 0 & -\bar{S} \end{bmatrix}, \tag{3.2}$$

where

$$\bar{S} = C(A^{-1} - A^{-1}B^T S_B^{-1} BA^{-1})C^T. \tag{3.3}$$

These preconditioners can be regarded as extensions or generalizations of "standard" block diagonal and block triangular preconditioners for saddle point problems (see, e.g., [6] and [9] for extensive treatments). We note that the two block triangular preconditioners $\mathscr{P}_{GT,1}$ and $\mathscr{P}_{GT,2}$ correspond to the two natural possible partitionings of the matrix $\mathscr{A}$ shown in (1.2). We also remark that all these preconditioners are examples of "ideal" preconditioners, in the sense that in general the matrices $S_B = BA^{-1}B^T$, $S_C = CA^{-1}C^T$, $BA^{-1}C^T$ (or $CA^{-1}B^T$), and $\bar{S}$ will be full and therefore cannot be formed explicitly. In practice, they (or their inverses) will have to be approximated, possibly by some iterative process; the same applies to the action of A^{-1} when solving the systems associated with the preconditioners.[1] Hence, in practice, the preconditioners will have to be applied "inexactly," possibly necessitating the use of a flexible Krylov subspace method. Nevertheless, the spectral analysis for the ideal case is still useful as it provides insight on the performance of the inexact preconditioners, at least for "sufficiently accurate" inexact solves.

We also mention that one can just as well adopt block upper triangular variants of the preconditioners $\mathscr{P}_{GT,1}$ and $\mathscr{P}_{GT,2}$. It has been shown in [17] that the difference between employing block lower and upper preconditioners should not be very significant, with the block upper triangular versions often working slightly better in practice. Nevertheless, in our numerical experiments we opted for $\mathscr{P}_{GT,1}$ instead of the block upper triangular version as the subsystem corresponding to the submatrix

$$\mathscr{S} = \begin{bmatrix} BA^{-1}B^T & BA^{-1}C^T \\ CA^{-1}B^T & CA^{-1}C^T \end{bmatrix} \tag{3.4}$$

is solved inexactly by an inner iteration, while the subsystem associated with coefficient matrix A is solved "exactly." Hence, using forward substitution leads to a more accurate application of the preconditioner. For consistency we also chose to adopt the lower triangular form for $\mathscr{P}_{GT,2}$.

Our first result concerns the block diagonal preconditioner $\mathscr{P}_D$. It is obvious that $\mathscr{P}_D$ is invertible (indeed, SPD) if and only if $A \succ 0$ and B and C have full rank. Under these assumptions, $\mathscr{P}_D$ can be used to precondition the minimal residual (MINRES) method [16]. Here and thereafter, $\Lambda(\cdot)$ is used to denote the spectrum of a matrix.

THEOREM 3.1. *Suppose matrix $\mathscr{A}$ in (1.1) is nonsingular, with $A \succ 0$, B and C of full row rank, and $D = 0$. Then*

$$\Lambda(\mathscr{P}_D^{-1}\mathscr{A}) \subset \left(-1, \frac{1-\sqrt{1+4\gamma_*}}{2}\right] \cup \{1\} \cup \left[\frac{1+\sqrt{1+4\gamma_*}}{2}, 2\right), \tag{3.5}$$

with

$$0 < \gamma_* = \min \frac{x^T(B^T S_B^{-1} B + C^T S_C^{-1} C)x}{x^T A x} < 2, \tag{3.6}$$

where the minimum is taken over all $x \in \mathbb{R}^n$, $x \notin \ker(B) \cap \ker(C)$, such that $(x; y; z)$ is an eigenvector of $\mathscr{P}_D^{-1}\mathscr{A}$. In particular, the set $\{1\} \cup [\frac{1+\sqrt{1+4\gamma_*}}{2}, 2)$ contains n eigenvalues and the negative interval $(-1, \frac{1-\sqrt{1+4\gamma_*}}{2}]$ contains $m+p$ eigenvalues. Furthermore, if $\lambda \neq 1$ is an eigenvalue of $\mathscr{P}_D^{-1}\mathscr{A}$, then $1-\lambda$ is also an eigenvalue.

[1] See section 4.1, however, for an example in which some of these matrices remain sparse and can be formed explicitly.

Proof. Since $\mathscr{A}$ is symmetric and $\mathscr{P}_D$ is SPD, all the eigenvalues and corresponding eigenvectors are real. Let λ be an arbitrary eigenvalue of $\mathscr{P}_D^{-1}\mathscr{A}$, then there exists a vector $(x;y;z) \neq (0;0;0)$ such that

$$Ax + B^T y + C^T z = \lambda Ax, \tag{3.7}$$

$$Bx = \lambda BA^{-1}B^T y, \tag{3.8}$$

$$Cx = \lambda CA^{-1}C^T z. \tag{3.9}$$

Note that it must be $x \neq 0$; otherwise $y = 0$ and $z = 0$ by (3.8) and (3.9). If $\ker(B) \cap \ker(C) \neq \{0\}$, then $\lambda = 1$ is an eigenvalue, since any vector $(x;0;0)$ with $x \neq 0$, $x \in \ker(B) \cap \ker(C)$ will be a corresponding eigenvector of $\mathscr{A}$. Conversely, any eigenvector corresponding to $\lambda = 1$ is necessarily of this form.

Assume now that $\lambda \neq 1$. We compute $y = \frac{1}{\lambda}(BA^{-1}B^T)^{-1}Bx \equiv \frac{1}{\lambda}S_B^{-1}Bx$ and $z = \frac{1}{\lambda}(CA^{-1}C^T)^{-1}Cx \equiv \frac{1}{\lambda}S_C^{-1}Cx$ from (3.8) and (3.9), respectively. Substituting the computed y and z into (3.7) and premultiplying by x^T, we obtain the following quadratic equation:

$$\lambda^2 - \lambda - \gamma = 0, \tag{3.10}$$

where

$$\gamma = \frac{x^T \left(B^T S_B^{-1} B + C^T S_C^{-1} C\right) x}{x^T A x} > 0.$$

The roots of (3.10) are given by

$$\lambda_+ = \frac{1 + \sqrt{1+4\gamma}}{2} \quad \text{and} \quad \lambda_- = \frac{1 - \sqrt{1+4\gamma}}{2}, \tag{3.11}$$

which shows that $\lambda_\pm = 1 - \lambda_\mp$. Since

$$\frac{x^T B^T S_B^{-1} B x}{x^T A x} \leq \lambda_{\max}(A^{-1}B^T S_B^{-1} B) = 1$$

and, in a similar way,

$$\frac{x^T C^T S_C^{-1} C x}{x^T A x} \leq 1,$$

we obtain that $\gamma \in (0, 2]$. We now show that in fact $\gamma \in (0, 2)$ since $\lambda = 2$ cannot be an eigenvalue. Indeed, if $\lambda = 2$ then (3.7) implies that

$$B^T y + C^T z = Ax.$$

It follows that

$$Bx = BA^{-1}B^T y + BA^{-1}C^T z,$$
$$Cx = CA^{-1}B^T y + CA^{-1}C^T z.$$

From (3.8) and (3.9), we get $Bx = 2BA^{-1}B^T$ and $Cx = 2CA^{-1}C^T$. Therefore,

$$\mathscr{S}_s(y;z) = (0;0),$$

where

$$\mathscr{S}_s = \begin{bmatrix} BA^{-1}B^T & -BA^{-1}C^T \\ -CA^{-1}B^T & CA^{-1}C^T \end{bmatrix}.$$

But $\mathscr{S}_s = (I_m; -I_p)\mathscr{S}(I_m; -I_p)$, where $\mathscr{S}$ is given by (3.4). From this and the nonsingularity of $\mathscr{S}$, it follows that y and z both must be zero and therefore $x = 0$, which is contrary to the assumption that $(x; y; z)$ is an eigenvector. Therefore, $\gamma \in (0, 2)$ and thus $-1 < \lambda_- \leq \frac{1-\sqrt{1+4\gamma_*}}{2} < 0$ and $1 < \frac{1+\sqrt{1+4\gamma_*}}{2} \leq \lambda_+ < 2$, proving (3.5).

Finally, recalling that $\mathscr{A}$ has n positive and $m+p$ negative eigenvalues (see, e.g., [6, sect. 3.4]) and observing that $\mathscr{P}_D^{-1}\mathscr{A}$ is similar to $\mathscr{P}_D^{-\frac{1}{2}}\mathscr{A}\mathscr{P}_D^{-\frac{1}{2}}$, we conclude by Sylvester's law of inertia that there are exactly n eigenvalues that are either 1 or lie in the positive interval in (3.5), and exactly $m+p$ eigenvalues lying in the negative interval, counted with their multiplicities. ∎

Remark 3.2. It is clear from the foregoing proof that for any positive eigenvalue of the form λ_+, there must be a corresponding negative eigenvalue $\lambda_- = 1 - \lambda_+$; see (3.11). On the other hand, we also showed that $\mathscr{P}_D^{-1}\mathscr{A}$ must have n positive and $m+p$ negative eigenvalues, and in general $n > m+p$. This is true whether $\lambda = 1$ is an eigenvalue or not. This apparent contradiction can be explained by observing that the multiplicity of λ_+ as an eigenvalue of $\mathscr{P}_D^{-1}\mathscr{A}_D$ will generally be different from that of the corresponding λ_-. Indeed, there may be a different number of eigenvectors of the form $(x; y; z)$ corresponding to λ_+ and λ_-, all with the same x (and thus the same γ) but different y or z. Hence, while the negative and positive intervals must contain the same number of *distinct* nonunit eigenvalues, the multiplicities of the positive and negative eigenvalues must add up to n and $m+p$, respectively.

Remark 3.3. While Theorem 3.1 shows that the positive eigenvalues are nicely bounded (between 1 and 2), as it stands it does not provide any useful information on the rightmost negative eigenvalue, since γ_*, while always strictly greater than zero, can in principle be arbitrarily small. Nevertheless, in special cases, given additional assumptions on the blocks A, B, and C, it is possible to say something about the value of γ_* and thus on the condition number of the preconditioned matrix. To see this, let $A = LL^T$ be the Cholesky factorization of A and let $\hat{B} = BL^{-T}$ and $\hat{C} = CL^{-T}$. Note that $P_{\hat{B}} := \hat{B}^T S_B^{-1}\hat{B}$ and $P_{\hat{C}} := \hat{C}^T S_C^{-1}\hat{C}$ are the orthogonal projectors onto range$(\hat{B}^T)$ and range$(\hat{C}^T)$, respectively. Letting $v = L^T x$, we can rewrite γ_* as

$$\gamma_* = \min \frac{v^T(P_{\hat{B}} + P_{\hat{C}})v}{\|v\|_2^2},$$

where the minimum is taken over all vectors v of the form $v = L^T x$ where $x \in \mathbb{R}^n$ is subject to the restrictions stated in Theorem 3.1. If the subspaces range$(\hat{B}^T)$ and range$(\hat{C}^T)$ happened to be mutually orthogonal we would have $P_{\hat{B}} + P_{\hat{C}} = I$ and therefore $\gamma_* = 1$. In this case the preconditioned matrix would have precisely three distinct eigenvalues: 1 (provided that $\ker(\hat{B}) \cap \ker(\hat{C}) \neq \{0\}$) and $\frac{1\pm\sqrt{5}}{2}$. (This is also a simple consequence of the fact that under these assumptions, the double saddle point problem decouples into two independent saddle point problems, and the preconditioned system into two independent saddle point matrices each preconditioned with the ideal block diagonal preconditioner in [15].) This argument suggests that the preconditioned matrix $\mathscr{P}_D^{-1}\mathscr{A}$ will be well conditioned if the transformed constrained matrices $\hat{B}^T$ and $\hat{C}^T$ have "nearly orthogonal" ranges; this happens, for example, if the constraints represented by B and C are only weakly coupled, and if A is either (block) diagonal or has entries that decay rapidly away from the main diagonal.

On the other hand, if the two subspaces range$(\hat{B})$ and range$(\hat{C})$ are "nearly collinear" and if for some x the vector $v = L^T x$ happened to be (nearly) orthogonal to either of these two subspaces, we would have that $\gamma_* \approx 0$, and the preconditioned

matrix will have at least one eigenvalue close to zero. Precise bounds on γ_* might be possible in terms of angles between the two subspaces, but the usefulness of such bounds would be limited in view of the difficulty of procuring information on such angles in practice.

Next, we prove a result concerning the spectrum of matrices preconditioned with the block triangular preconditioner $\mathscr{P}_{_T}$. We note that since this preconditioner is nonsymmetric, it cannot be used with MINRES. Note that $\mathscr{P}_{_T}$ is guaranteed to be nonsingular when $A \succ 0$ and B and C have full rank.

THEOREM 3.4. *Under the assumptions of Theorem* 3.1, $\Lambda(\mathscr{P}_{_T}^{-1}\mathscr{A}) \subset (0,2)$ *with* $\lambda = 1$ *being an eigenvalue of multiplicity at least* n. *Moreover, the spectrum of* $\mathscr{P}_{_T}^{-1}\mathscr{A}$ *is symmetric with respect to* $\lambda = 1$, *i.e., if* $\lambda_1 \neq 1$ *and* $\lambda_2 \neq 1$ *are two eigenvalues of* $\mathscr{P}_{_T}^{-1}\mathscr{A}$, *then* $\lambda_1 + \lambda_2 = 2$.

Proof. Suppose that λ is an arbitrary eigenvalue of $\mathscr{P}_{_T}^{-1}\mathscr{A}$ with the corresponding eigenvector $(x; y; z)$, i.e.,

$$Ax + B^T y + C^T z = \lambda(Ax + B^T y + C^T z), \tag{3.12}$$

$$Bx = -\lambda B A^{-1} B^T y, \tag{3.13}$$

$$Cx = -\lambda C A^{-1} C^T z. \tag{3.14}$$

Notice that $x \neq 0$; otherwise, in view of the fact that B^T and C^T are full column rank, $x = 0$ implies $(x; y; z) = (0; 0; 0)$ in contradiction with the fact that $(x; y; z)$ is an eigenvector.

Clearly, $\lambda = 1$ is an eigenvalue of $\mathscr{P}_{_T}^{-1}\mathscr{A}$ with corresponding eigenvector of the form $(x; -S_B^{-1}Bx; -S_C^{-1}Cx)$. The multiplicity of this eigenvalue is therefore at least n. Assume now that $\lambda \neq 1$. From (3.12), we deduce that

$$Ax + B^T y + C^T z = 0. \tag{3.15}$$

Similar to the proof of Theorem 3.1, we compute y and z from (3.13) and (3.14) in terms of λ and x, respectively. Substituting the derived values of y and z into (3.15), we get

$$\lambda = \frac{x^* \left(B^T S_B^{-1} B + C^T S_C^{-1} C\right) x}{x^* A x}. \tag{3.16}$$

(Note that since λ is real, the corresponding eigenvector can also be chosen to be real and therefore x^* in (3.16) can be replaced by x^T.) Hence, λ has the same expression as γ in the proof of Theorem 3.1, and by the same argument given there we conclude that $\lambda \in (0, 2)$.

Next, recall that $\Lambda(\mathscr{A}\mathscr{P}_{_T}^{-1}) = \Lambda(\mathscr{P}_{_T}^{-1}\mathscr{A})$. Straightforward computations reveal that

$$\begin{aligned}
\mathscr{A}\mathscr{P}_{_T}^{-1} &= \begin{bmatrix} A & B^T & C^T \\ B & 0 & 0 \\ C & 0 & 0 \end{bmatrix} \begin{bmatrix} A^{-1} & A^{-1}B^T S_B^{-1} & A^{-1}C^T S_C^{-1} \\ 0 & -S_B^{-1} & 0 \\ 0 & 0 & -S_C^{-1} \end{bmatrix} \\
&= \begin{bmatrix} I & 0 & 0 \\ BA^{-1} & I & BA^{-1}C^T S_C^{-1} \\ CA^{-1} & CA^{-1}B^T S_B^{-1} & I \end{bmatrix}.
\end{aligned}$$

The above relation, incidentally, confirms that the number of eigenvalues which are equal to one cannot be less than n, the order of the $(1,1)$-block. In addition, it can

be seen that the remaining $m+p$ eigenvalues of $\mathscr{P}_{_T}^{-1}\mathscr{A}$ are the eigenvalues of $I+\hat{\mathscr{S}}$, where

$$\hat{\mathscr{S}} = \begin{bmatrix} 0 & BA^{-1}C^TS_C^{-1} \\ CA^{-1}B^TS_B^{-1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & BA^{-1}C^T \\ CA^{-1}B^T & 0 \end{bmatrix} \begin{bmatrix} S_B^{-1} & 0 \\ 0 & S_C^{-1} \end{bmatrix}.$$

To conclude the proof, we only need to show that the distribution of the eigenvalues of $\hat{\mathscr{S}}$ is symmetric with respect to zero. Hence, all the eigenvalues of $\mathscr{P}_{_T}^{-1}\mathscr{A}$ must lie in the interval $(0,2)$. In view of the fact that $S_B \succ 0$ and $S_C \succ 0$, matrix $\hat{\mathscr{S}}$ is similar to

$$\check{\mathscr{S}} = \begin{bmatrix} S_B^{-1/2} & 0 \\ 0 & S_C^{-1/2} \end{bmatrix} \begin{bmatrix} 0 & BA^{-1}C^T \\ CA^{-1}B^T & 0 \end{bmatrix} \begin{bmatrix} S_B^{-1/2} & 0 \\ 0 & S_C^{-1/2} \end{bmatrix}$$

and therefore the two matrices have the same eigenvalues. Evidently,

$$\check{\mathscr{S}} = \begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}$$

with $X = S_B^{-1/2}BA^{-1}C^TS_C^{-1/2}$. It is well known that the eigenvalues of a matrix of the above form are given by $\pm\sigma_i(X)$, where $\sigma_i(X)$ stands for the ith singular value of X. This shows the symmetric distribution of the eigenvalues of $\hat{\mathscr{S}}$ with respect to zero. ∎

Remark 3.5. Similar to Remark 3.3, we note that additional assumptions on the matrices A, B, and C are required in order to obtain a lower bound on the eigenvalues of $\mathscr{P}_T^{-1}\mathscr{A}$.

We conclude this section with a few brief remarks on the preconditioners $\mathscr{P}_{GD}$, $\mathscr{P}_{GT,1}$, and $\mathscr{P}_{GT,2}$. We observe that the first two are just special cases of the "ideal" block diagonal and block (lower) triangular preconditioners for saddle point problems based on the first of the two partitionings in (1.2); the third one is the ideal block (lower) triangular preconditioner based on the second partitioning of $\mathscr{A}$ in (1.2). The spectral properties of preconditioned saddle point matrices with any of these block preconditioners are well known; see [15] or [6, sects. 10.1.1–10.1.2]. In particular, $\mathscr{P}_{GD}^{-1}\mathscr{A}$ has only three distinct eigenvalues and is diagonalizable, while $\mathscr{P}_{GT,1}^{-1}\mathscr{A}$ and $\mathscr{P}_{GT,2}^{-1}\mathscr{A}$ have all the eigenvalues equal to 1 and are nondiagonalizable but have minimum polynomial of degree 2. Hence, MINRES and the generalized minimum residual (GMRES) method [20] will reach the exact solution in at most three and two steps, respectively. As before, these ideal block preconditioners may be prohibitively expensive to construct and apply; in practice, they are usually replaced by inexact variants.

3.2. Block preconditioners of the second type. In this part the eigenvalue distributions of the preconditioned matrices are discussed for the case that the coefficient matrix $\mathscr{A}$ has nonzero $(3,3)$-block. We consider the two following types of block triangular preconditioners:

$$\tilde{\mathscr{P}}_{_T} = \begin{bmatrix} A & B^T & C^T \\ 0 & -BA^{-1}B^T & 0 \\ 0 & 0 & -(D+CA^{-1}C^T) \end{bmatrix} \tag{3.17}$$

and

$$\hat{\mathscr{P}}_{_T} = \begin{bmatrix} A & B^T & C^T \\ 0 & -BA^{-1}B^T & -BA^{-1}C^T \\ 0 & 0 & -(D + CA^{-1}C^T) \end{bmatrix}. \tag{3.18}$$

We note that these preconditioners will be nonsingular if $A \succ 0$, B has full row rank, $D \succeq 0$, and $\ker(D) \cap \ker(C^T) = \{0\}$. From Theorem 2.2, these conditions also guarantee the invertibility of $\mathscr{A}$.

For ease of exposition, we present the analysis in several steps. Our first result is the following.

THEOREM 3.6. *Assume that $A \succ 0$, B has full rank, $D \succeq 0$, and $\ker(D) \cap \ker(C^T) = \{0\}$. Then all the eigenvalues of $\mathscr{A}\tilde{\mathscr{P}}_{_T}^{-1}$ are real and nonzero. Moreover, $\lambda = 1$ is an eigenvalue of algebraic multiplicity at least n.*

Proof. Under the stated assumptions, both $\mathscr{A}$ and $\tilde{\mathscr{P}}_{_T}$ are nonsingular. We have

$$\begin{aligned} \mathscr{A}\tilde{\mathscr{P}}_{_T}^{-1} &= \begin{bmatrix} A & B^T & C^T \\ B & 0 & 0 \\ C & 0 & -D \end{bmatrix} \begin{bmatrix} A^{-1} & A^{-1}B^T S_B^{-1} & A^{-1}C^T \tilde{S}_C^{-1} \\ 0 & -S_B^{-1} & 0 \\ 0 & 0 & -\tilde{S}_C^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 & 0 \\ BA^{-1} & I & BA^{-1}C^T \tilde{S}_C^{-1} \\ CA^{-1} & CA^{-1}B^T S_B^{-1} & I \end{bmatrix}, \end{aligned}$$

where $\tilde{S}_C = D + CA^{-1}C^T$. Similar to the proof of Theorem 3.4, we find that the number of eigenvalues of $\mathscr{A}\tilde{\mathscr{P}}_{_T}^{-1}$ which are equal to one is at least n, the order of the $(1,1)$-block, with the remaining eigenvalues being those of the matrix $I + \tilde{\mathscr{S}}_1$, where

$$\begin{aligned} \tilde{\mathscr{S}}_1 &= \begin{bmatrix} 0 & BA^{-1}C^T \tilde{S}_C^{-1} \\ CA^{-1}B^T S_B^{-1} & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & BA^{-1}C^T \\ CA^{-1}B^T & 0 \end{bmatrix} \begin{bmatrix} S_B^{-1} & 0 \\ 0 & \tilde{S}_C^{-1} \end{bmatrix}. \end{aligned}$$

Since $\tilde{\mathscr{S}}_1$ is the product of two symmetric matrices, one of which is positive definite, its eigenvalues are all real and the result is proved. ∎

Next, we present bounds on the eigenvalues of the preconditioned matrices $\tilde{\mathscr{P}}_T^{-1}\mathscr{A}$ and $\hat{\mathscr{P}}_T^{-1}\mathscr{A}$. To this end, we make use of the Cholesky factorization of the $(1,1)$-block of $\mathscr{A}$, i.e., $A = LL^T$. Consider the lower triangular matrix $\mathscr{L}$ defined by

$$\mathscr{L} = \begin{bmatrix} L & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}.$$

We define $\hat{\mathscr{A}} = \mathscr{L}^{-1}\mathscr{A}\mathscr{L}^{-T}$, which has the following structure:

$$\hat{\mathscr{A}} = \begin{bmatrix} I & \hat{B}^T & \hat{C}^T \\ \hat{B} & 0 & 0 \\ \hat{C} & 0 & D \end{bmatrix},$$

where again we set $\hat{B} = BL^{-T}$ and $\hat{C} = CL^{-T}$. Now we consider the following two block triangular preconditioners for $\hat{\mathscr{A}}$:

$$\tilde{\hat{\mathscr{P}}}_T = \begin{bmatrix} I & \hat{B}^T & \hat{C}^T \\ 0 & -\hat{B}\hat{B}^T & 0 \\ 0 & 0 & -(D + \hat{C}\hat{C}^T) \end{bmatrix} \tag{3.19}$$

and

$$\hat{\hat{\mathscr{P}}}_T = \begin{bmatrix} I & \hat{B}^T & \hat{C}^T \\ 0 & -\hat{B}\hat{B}^T & -\hat{B}\hat{C}^T \\ 0 & 0 & -(D + \hat{C}\hat{C}^T) \end{bmatrix}. \tag{3.20}$$

It is not difficult to check that the following two relations hold:

$$\tilde{\mathscr{P}}_T^{-1}\mathscr{A} = \mathscr{L}^{-T}\tilde{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}}\mathscr{L}^T \quad \text{and} \quad \hat{\mathscr{P}}_T^{-1}\mathscr{A} = \mathscr{L}^{-T}\hat{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}}\mathscr{L}^T,$$

which reveal that $\Lambda(\tilde{\mathscr{P}}_T^{-1}\mathscr{A}) = \Lambda(\tilde{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}})$ and $\Lambda(\hat{\mathscr{P}}_T^{-1}\mathscr{A}) = \Lambda(\hat{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}})$.

THEOREM 3.7. *Under the same assumptions of Theorem* 3.6, $\Lambda(\tilde{\mathscr{P}}_T^{-1}\mathscr{A}) = \Lambda(\tilde{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}}) \subset (0, 1-\sqrt{\xi}] \cup \{1\} \cup [1+\sqrt{\xi}, 2]$ *where*

$$\xi = \frac{\bar{\sigma}_{\min}^2(\hat{C})}{\lambda_{\max}(D) + \bar{\sigma}_{\min}^2(\hat{C})}. \tag{3.21}$$

Here $\bar{\sigma}_{\min}(\hat{C})$ *denotes the smallest nonzero singular value of* $\hat{C}$.

Proof. The equality $\Lambda(\tilde{\mathscr{P}}_T^{-1}\mathscr{A}) = \Lambda(\tilde{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}})$ has already been noted. From Theorem 3.6 we already know that the spectrum is real and that $\lambda = 1$ is an eigenvalue of algebraic multiplicity at least n. Assume now that $\lambda \neq 1$ is an eigenvalue of $\tilde{\hat{\mathscr{P}}}_T^{-1}\hat{\mathscr{A}}$. There exists a (real) nonzero vector $(x; y; z)$ such that

$$x + \hat{B}^T y + \hat{C}^T z = \lambda(x + \hat{B}^T y + \hat{C}^T z), \tag{3.22}$$

$$\hat{B}x = -\lambda\hat{B}\hat{B}^T y, \tag{3.23}$$

$$\hat{C}x - Dz = -\lambda(D + \hat{C}\hat{C}^T)z. \tag{3.24}$$

Notice that $x \neq 0$, otherwise $x = 0$ implies that y and z are both zero in contradiction with the fact that $(x; y; z)$ is an eigenvector.

From (3.22), we get

$$x + \hat{B}^T y + \hat{C}^T z = 0$$

and therefore

$$\hat{B}x = -(\hat{B}\hat{B}^T y + \hat{B}\hat{C}^T z)$$

and

$$\hat{C}x = -(\hat{C}\hat{B}^T y + \hat{C}\hat{C}^T z).$$

Substituting the preceding two relations into (3.23) and (3.24), respectively, we get

$$(\lambda - 1)\hat{B}\hat{B}^T y = \hat{B}\hat{C}^T z \tag{3.25}$$

and

$$(\lambda - 1)(D + \hat{C}\hat{C}^T)z = \hat{C}\hat{B}^T y. \tag{3.26}$$

We observe that the vectors y and z must both be nonzero. Indeed, our assumptions imply that both $\hat{B}\hat{B}^T$ and $D + \hat{C}\hat{C}^T$ are positive definite, and this fact, together with (3.25) and (3.26), implies that $y = 0$ if and only if $z = 0$. Notice that $\hat{C}^T z \neq 0$, otherwise (3.25) implies that $\lambda = 1$ which is contrary to our assumption. By computing y from (3.25) and then substituting it into (3.26), we obtain

$$(\lambda - 1)^2 = \frac{z^T \hat{C} P \hat{C}^T z}{z^T (D + \hat{C}\hat{C}^T) z}, \tag{3.27}$$

where $P = \hat{B}^T (\hat{B}\hat{B}^T)^{-1} \hat{B}$. Note that we can use z^T instead of z^* (since the eigenvalues and therefore the eigenvectors are necessarily real) and that P is an orthogonal projector, i.e., $P^2 = P$ and $P = P^T$. Using the fact that $\|Pv\|_2 \leq \|v\|_2$ for any vector v, we obtain as a consequence of (3.27) that $|\lambda - 1| \leq 1$, which is equivalent to say that $\lambda \in (0, 2]$ since $0 \notin \Lambda(\tilde{\mathscr{P}}_{\mathscr{T}}^{-1} \mathscr{A})$.

Finally, we apply the Rayleigh–Ritz theorem [11, Thm. 4.2.2] to obtain

$$\frac{1}{\lambda_{\max}(D)/\bar{\lambda}_{\min}(\hat{C}\hat{C}^T) + 1} \leq \frac{z^T \hat{C} P \hat{C}^T z}{z^T (D + \hat{C}\hat{C}^T) z} \leq \frac{1}{\lambda_{\min}(D)/\lambda_{\max}(\hat{C}\hat{C}^T) + 1}, \tag{3.28}$$

where $\bar{\lambda}_{\min}(\hat{C}\hat{C}^T)$ denotes the smallest nonzero eigenvalue of $\hat{C}\hat{C}^T$. This shows that $|\lambda - 1| \geq \sqrt{\xi}$. The proof is complete. ∎

Remark 3.8. It can be seen from (3.27) that if either $D \succ 0$ or $\text{range}(B^T) \cap \text{range}(C^T) = \{0\}$ then $\lambda = 2$ cannot be an eigenvalue of $\tilde{\mathscr{P}}_{\mathscr{T}}^{-1} \mathscr{A}$.

Remark 3.9. The previous theorem is rather unsatisfactory since the lower bound for the eigenvalues of $\tilde{\mathscr{P}}_{\mathscr{T}}^{-1} \mathscr{A}$ is 0. This bound, unfortunately, cannot be improved without making further assumptions on the problem. Also, it may seem odd that the matrix B does not appear in the expression of the quantity ξ, in terms of which the bounds are formulated. However, using (3.25) and (3.26) we can see that any nonunit eigenvalue satisfies

$$\lambda = 1 + \frac{y^T \hat{B}\hat{C}^T z}{y^T \hat{B}\hat{B}^T y} = 1 + \frac{y^T \hat{B}\hat{C}^T z}{z^T (D + \hat{C}\hat{C}^T) z}, \tag{3.29}$$

where y and z are components of the eigenvector $u = (x; y; z)$ associated with λ. In general, the fractions in (3.29) could be arbitrarily close to -1; for example, it follows from the first equality in (3.29) that $\lambda \to 0$ if $C^T z \to -B^T y$. On the other hand, the expressions in (3.29) can be useful if information is available on, e.g., the principal angles between $\text{range}(\hat{B}^T)$ and $\text{range}(\hat{C}^T)$, or on the relative "size" of $\hat{C}$ with respect to either $\hat{B}$ or D. For instance, if $\sigma_{\max}(\hat{C})$ becomes very small as the mesh size $h \to 0$ as compared to either $\sigma_{\min}(\hat{B})$ or $\lambda_{\min}(D)$, then the eigenvalues of $\tilde{\mathscr{P}}_{\mathscr{T}}^{-1} \mathscr{A}$ would cluster around 1 as the mesh is refined.

We conclude this section with a result on the preconditioner $\hat{\mathscr{P}}_{\mathscr{T}}$.

THEOREM 3.10. *Assume that $A \succ 0$, B has full row rank, and $D \succ CA^{-1}C^T$. Then $\Lambda(\hat{\mathscr{P}}_{\mathscr{T}}^{-1} \mathscr{A}) = \Lambda(\hat{\tilde{\mathscr{P}}}_{\mathscr{T}}^{-1} \hat{\mathscr{A}}) \subset (\frac{1}{2}, 1]$.*

Proof. First, we note that $\hat{\mathscr{P}}_{_T}^{-1}\mathscr{A}$ is invertible and $\Lambda(\hat{\mathscr{P}}_{_T}^{-1}\mathscr{A}) = \Lambda(\hat{\tilde{\mathscr{P}}}_{_T}^{-1}\hat{\mathscr{A}})$, where $\hat{\tilde{\mathscr{P}}}_{_T}$ is given in (3.20). Let λ be an eigenvalue of $\hat{\tilde{\mathscr{P}}}_{_T}^{-1}\hat{\mathscr{A}}$ with corresponding eigenvector $(x; y; z)$. We have

$$x + \hat{B}^T y + \hat{C}^T z = \lambda(x + \hat{B}^T y + \hat{C}^T z), \tag{3.30}$$

$$\hat{B}x = -\lambda(\hat{B}\hat{B}^T y + \hat{B}\hat{C}^T z), \tag{3.31}$$

$$\hat{C}x - Dz = -\lambda(D + \hat{C}\hat{C}^T)z. \tag{3.32}$$

If C^T (and therefore $\hat{C}^T$) does not have full column rank, we observe that $\lambda = 1$ is an eigenvalue with corresponding eigenvectors of the form $(0; 0; z)$, where $0 \neq z \in \ker(C^T)$. Hence, the multiplicity of $\lambda = 1$ is at least equal to $p-r$, where $r = \operatorname{rank}(C)$.

Let us now assume that $\hat{C}$ has full row rank, and let $x \in \mathbb{R}^n$ be any nonzero vector. It is then easy to see that $\lambda = 1$ is an eigenvalue of $\hat{\tilde{\mathscr{P}}}_{_T}^{-1}\hat{\mathscr{A}}$ with corresponding eigenvector

$$\left(x; -(\hat{B}\hat{B}^T)^{-1}(\hat{B} - \hat{B}\hat{C}^T(\hat{C}\hat{C}^T)^{-1}\hat{C})x; -(\hat{C}\hat{C}^T)^{-1}\hat{C}x\right).$$

Since there are n linearly independent vectors of this form, $\lambda = 1$ is an eigenvalue of multiplicity at least n of $\hat{\mathscr{P}}_{_T}^{-1}\mathscr{A}$.

In the sequel we assume that $\lambda \neq 1$. From (3.30) we obtain

$$x + \hat{B}^T y + \hat{C}^T z = 0.$$

It follows that

$$\hat{B}x = -(\hat{B}\hat{B}^T y + \hat{B}\hat{C}^T z), \tag{3.33}$$

$$\hat{C}x = -(\hat{C}\hat{B}^T y + \hat{C}\hat{C}^T z). \tag{3.34}$$

Substituting (3.33) and (3.34) into (3.31) and (3.32), respectively, we get

$$(\lambda - 1)(\hat{B}\hat{B}^T y + \hat{B}\hat{C}^T z) = 0, \tag{3.35}$$

$$(\lambda - 1)(D + \hat{C}\hat{C}^T)z = \hat{C}\hat{B}^T y. \tag{3.36}$$

From (3.36) it can be deduced that $y = 0$ if (and only if) $z = 0$, in which case $x = 0$ in contradiction with the assumption that $(x; y; z)$ is an eigenvector. Keeping in mind that $\lambda \neq 1$, the vector y can be computed from (3.35) as $y = -(\hat{B}\hat{B}^T)^{-1}\hat{B}\hat{C}^T z$. In order to complete the proof, we first substitute y in (3.36), and then multiply both sides of the resulting relation by z^T (again, we can actually use z^T in place of z^* since the eigenvalues are necessarily real). Thus,

$$\lambda = 1 - \frac{z^T\hat{C}P\hat{C}^T z}{z^T(D + \hat{C}\hat{C}^T)z},$$

where $P = \hat{B}^T(\hat{B}\hat{B}^T)^{-1}\hat{B}$. As pointed before, the matrix P is an orthogonal projector. The result immediately follows from the inequality

$$\frac{z^T\hat{C}P\hat{C}^T z}{z^T(D + \hat{C}\hat{C}^T)z} < \frac{1}{2},$$

which follows from the assumption that $D - CA^{-1}C^T \succ 0$. ∎

Remark 3.11. It is interesting to note that simply including in the preconditioner the additional term $-BA^{-1}C^T$ leads to much better bounds for the eigenvalues of the preconditioned matrix than in the case of $\tilde{\mathscr{P}}_T$. This, however, does not imply that the *performance* of $\hat{\mathscr{P}}_T$ will be necessarily better; rather, adding the extra term in the preconditioner simply leads to a much more satisfactory lower bound. Also, it is well known that eigenvalue information alone does not suffice, in general, to predict the convergence behavior of nonsymmetric Krylov subspace methods like GMRES. Nevertheless, experience shows that in many cases of practical interest convergence can be expected to be fast when the spectrum is real, positive, and contained in an interval of modest length bounded away from zero. This behavior is also observed when the "ideal" preconditioners are replaced with inexact versions, as long as the preconditioner is applied with a reasonable degree of accuracy.

4. Numerical experiments. In this section, we present a selection of numerical tests aimed at illustrating the performance of some of the proposed preconditioners. Due to space limitations, we present detailed results only for some of the methods analyzed in the theoretical sections, and comment briefly on the remaining ones. We focus on two sets of problems of the type (1.1) arising from two very different applications, one with $D = 0$ and the other with $D \neq 0$. All of the reported numerical results were performed on a 64-bit 2.45 GHz core i7 processor and 8.00 GB RAM using MATLAB version 8.3.0532. In all the experiments we have used right-hand sides corresponding to random solution vectors, performing ten runs, and then averaging the CPU times. The iteration counts reported in the tables (under "Iter") are also averages (rounded to the nearest integer).

All the methods require repeated solution (whether "exact" or inexact) of SPD linear systems as subtasks. These are either solved by sparse Cholesky factorization with symmetric approximate minimum degree reordering or by the preconditioned conjugate gradient (PCG) method. When using PCG, unless otherwise specified, the preconditioner used is a drop tolerance-based incomplete Cholesky factorization [5, 20] computed using the MATLAB function "`ichol(.,opts)`", where

- `opts.type = 'ict'`,
- `opts.droptol = 1e-2`.

We comment that with the inexact variants, performing inner iterations is essential. Just replacing the exact block solves with an incomplete Cholesky factorization leads to a degradation of convergence rates for the outer Krylov solver.

In all the numerical tests below, the initial guess is taken to be the zero vector. For the MINRES, GMRES, and Flexible GMRES (FGMRES) methods the iterations are stopped once

$$\|b - \mathscr{A}(x^{(k)}; y^{(k)}; z^{(k)})\|_2 < 10^{-10}\|b\|_2.$$

For the inner PCG iterations (whenever applicable), the stopping tolerances used are specified below.

4.1. Saddle point systems from potential fluid flow modeling. Here we consider linear systems of equations of the form

$$\begin{bmatrix} A & B^T & C^T \\ B & 0 & 0 \\ C & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{4.1}$$

arising from a low-order Raviart–Thomas mixed-hybrid finite element approximation [7] of Darcy's law and continuity equation describing the three-dimensional (3D) po-

tential fluid flow problem in porous media. The continuous problem reads

$$\mathbf{A}\mathbf{u} = -\nabla p, \quad \nabla \cdot \mathbf{u} = q,$$

where $\mathbf{u}$ is the fluid velocity, p is the piezometric potential (fluid pressure), $\mathbf{A}$ is the symmetric and uniformly positive definite second-rank tensor of the hydraulic resistance of the medium with $[\mathbf{A}(\mathbf{x})]_{ij} \in L^\infty(\Omega)$ for $i, j = 1, 2, 3$, and q represents the density of potential sources in the medium. The underlying spatial domain Ω is cubic, and the boundary conditions are given by

$$p = p_D \quad \text{on} \quad \partial\Omega_D, \qquad \mathbf{u} \cdot \mathbf{n} = \mathbf{u}_N \quad \text{on} \quad \partial\Omega_N,$$

where $\partial\Omega = \overline{\partial\Omega_D} \cup \overline{\partial\Omega_N}$ with $\partial\Omega_D \neq \emptyset$, $\partial\Omega_D \cap \partial\Omega_N = \emptyset$, and $\mathbf{n}$ is the outward normal vector defined (a.e.) on $\partial\Omega$. We refer to [12] for details of the problem and its discretization. The solution vectors x and y in (4.1) correspond to velocity and pressure degrees of freedom (respectively), while z is a vector of Lagrange multipliers. For this problem we have that $A \succ 0$, B and C have full row rank and that $\mathscr{A}$ is nonsingular. Details on the dimensions of sub-blocks A, B, and C and further information can be found in [12, Table 1].

For this test problem, the SPD matrix A is block diagonal with small blocks, and linear systems associated with it can be solved very cheaply by means of Cholesky factorization. Likewise, the Schur complements $S_B = BA^{-1}B^T$, $S_C = CA^{-1}C^T$, $\bar{S}$ (see (3.3)) and the matrix $BA^{-1}C^T$ are still relatively sparse matrices which can be formed explicitly at low expense.[2] Concerning the block preconditioners, the best results were obtained with $\mathscr{P}_{GD}$ and $\mathscr{P}_{GT,1}$ in (3.1) and $\mathscr{P}_{GT,2}$ in (3.2). The block diagonal preconditioner $\mathscr{P}_{GD}$ was used with MINRES and FGMRES, while the two block triangular preconditioners $\mathscr{P}_{GT,1}$ and $\mathscr{P}_{GT,2}$ were used with both GMRES and FGMRES.

Apart from the inexpensive solves associated with A, the implementation of $\mathscr{P}_{GD}$ and $\mathscr{P}_{GT,1}$ requires solving linear systems associated with the Schur complement $\mathscr{S}$ given in (3.4). In spite of the sparsity of $\mathscr{S}$, solution by sparse Cholesky factorization is expensive (recall that this is a 3D problem). Thus, we solve such systems with the PCG method with a very stringent stopping criterion (inner relative residual norm less than $tol = 10^{-15}$) for MINRES and GMRES and a looser one ($tol = 10^{-4}$) for FGMRES.

The application of the preconditioner $\mathscr{P}_{GT,2}$, on the other hand, requires solving at each step a linear system of the form $\mathscr{P}_{GT,2}(w_1; w_2; w_3) = (r_1; r_2; r_3)$. This amounts to solving a saddle point problem of size $(n+m) \times (n+m)$ of the form

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \tag{4.2}$$

followed by solution of a linear systems with the coefficient matrix $\bar{S}$ (see (3.3)). The solution of (4.2) can be obtained in two steps as follows

- **Step I.** Solve $S_B w_2 = BA^{-1}r_1 - r_2$, to find w_2.
- **Step II.** Set $w_1 = A^{-1}(r_1 - B^T w_2)$.

We recall that for this particular test problem, A is block diagonal (with small blocks) and S_B is just a scalar multiple of the identity, so the above solution process is

[2]The Schur complement $BA^{-1}B^T$ for this problem turns out to be a scalar multiple of the $m \times m$ identity matrix.

TABLE 1
Results for the block diagonal preconditioner $\mathscr{P}_{GD}$, potential fluid flow problem.

	Method				
	MINRES		FGMRES		
Size	Iter	CPU time	Iter	Iter_{pcg}	CPU time
2125	3	0.0125	7	41	0.0115
17000	3	0.0947	7	73	0.0746
57375	3	0.4829	7	107	0.3390
136000	3	1.6226	7	137	1.0253
265625	3	3.9002	7	165	2.5793
459000	3	8.8899	7	196	5.8504

TABLE 2
Results for GMRES with block triangular preconditioners, potential fluid flow problem.

	Preconditioner			
	$\mathscr{P}_{GT,1}$		$\mathscr{P}_{GT,2}$	
Size	Iter	CPU time	Iter	CPU time
2125	2	0.0191	2	0.0180
17000	2	0.1284	2	0.1180
57375	2	0.5247	2	0.4516
136000	2	1.5425	2	1.2936
265625	2	3.6811	2	3.1080
459000	2	7.9861	2	6.8368

TABLE 3
Results for FGMRES with block triangular preconditioners, potential fluid flow problem.

	Preconditioner					
	$\mathscr{P}_{GT,1}$			$\mathscr{P}_{GT,2}$		
Size	Iter	Iter_{pcg}	CPU time	Iter	Iter_{pcg}	CPU time
2125	5	25	0.0085	5	25	0.0073
17000	6	47	0.0575	6	53	0.0534
57375	6	66	0.2361	6	72	0.2265
136000	6	87	0.7480	6	95	0.6563
265625	6	108	1.8190	6	112	1.5220
459000	6	134	4.2658	5	117	3.0442

extremely cheap within our GMRES and FGMRES iterative methods. As mentioned earlier, in addition to solving (4.2), to apply $\mathscr{P}_{GT,2}$ we also need to solve $\bar{S}w_3 = -r_3 + Cw_1$, where $\bar{S}$ is defined by (3.3). As pointed out, in this problem $\bar{S}$ can be formed explicitly as it is a sparse matrix. To solve $\bar{S}w_3 = -r_3 + Cw_1$ (we observe that $\bar{S}$ is SPD) the PCG method was used where the inner stopping tolerances were chosen as before as 10^{-15} and 10^{-4} depending on whether GMRES or FGMRES is used, respectively.

In Tables 1, 2, and 3 we report the results for the preconditioned MINRES, GMRES, and FGMRES iterative methods. The total number $n+m+p$ of unknowns is reported under "size". As expected, MINRES/GMRES with the "ideal" block diagonal/triangular preconditioners require exactly three and two steps to converge, independent of problem size. In Tables 1 and 3, the cumulative number of inner PCG iterations required is reported under "Iter_{pcg}."

These results show that for this particular example, the best results are obtained with the inexact block triangular preconditioners $\mathscr{P}_{GT,1}$ and $\mathscr{P}_{GT,2}$; of these two, the latter one (based on the second of the two partitionings (1.2)) appears to be slightly better in this particular case. We note the satisfactory scaling in terms of CPU time

for sufficiently small h, especially for FGMRES with the inexact $\mathscr{P}_{GT,2}$ preconditioner. As for the other two preconditioners, $\mathscr{P}_D$ and $\mathscr{P}_T$, their performance was generally inferior, with worsening iteration counts for increasing problem sizes. The observed behavior appears to be due to the fact that for this problem, some of the eigenvalues of the preconditioned matrices corresponding to $\mathscr{P}_D$ and $\mathscr{P}_T$ approach zero as the mesh is refined. Still, these preconditioners may well be useful in solving saddle point systems arising from other applications.

4.2. Saddle point systems from liquid crystal directors modeling. Continuum models for the orientational properties of liquid crystals require the minimization of free energy functionals of the form

$$\mathscr{F}[u,v,w,U] = \frac{1}{2}\int_0^1 \left[(u_z^2+v_z^2+w_z^2) - \alpha^2(\beta+w^2)U_z^2\right] dz, \tag{4.3}$$

where u, v, w, and U are functions of $z \in [0,1]$ subject to suitable end-point conditions, $u_z = \frac{du}{dz}$ (etc.), α and β are positive prescribed parameters. Approximation via a uniform piecewise-linear finite element scheme with $k+1$ cells using nodal quadrature and the prescribed boundary conditions leads to replacing the functional $\mathscr{F}$ with a function f of $4k$ variables:

$$\mathscr{F}[u,v,w,U] \approx f(u_1,\dots,u_k,v_1,\dots,v_k,w_1,\dots,w_k,U_1,\dots,U_k);$$

see [18, eq. (2.4)] for the precise form of f.

Minimization of the free energy (4.3) must be carried out under the so-called *unit vector constraint*, which at the discrete level can be expressed by imposing that the solution components u_j, v_j, and w_j satisfy

$$u_j^2+v_j^2+w_j^2 = 1, \quad j = 1,\dots,k.$$

Introducing Lagrange multipliers $\lambda_1,\dots,\lambda_k$, the problem reduces to finding the critical points of the Lagrangian function

$$L = f + \frac{1}{2}\sum_{j=1}^{k}\lambda_j(u_j^2+v_j^2+w_j^2-1).$$

Imposing the first-order conditions results in the system of $5k$ nonlinear equations $\nabla L(\mathbf{x}) = \mathbf{0}$, where the unknown vector $\mathbf{x} \in \mathbb{R}^{5k}$ collects the values (u_j, v_j, w_j) $(j = 1,\dots,k)$, $(\lambda_1,\dots,\lambda_k)$, and $(U_1,\dots,U_k)$ (in this order). Solving this nonlinear system with Newton's method leads to a linear system of the form

$$\nabla^2 L(\mathbf{x}^{(\ell)})\,\delta\mathbf{x}^{(\ell)} = -\nabla L(\mathbf{x}^{(\ell)}) \tag{4.4}$$

at each step ℓ, where $\nabla^2 L(\mathbf{x}^{(\ell)})$ denotes the Hessian of L evaluated at $\mathbf{x}^{(\ell)}$. As shown in [18], the Hessian has the following structure:

$$\nabla^2 L = \begin{bmatrix} A & B^T & C^T \\ B & 0 & 0 \\ C & 0 & -D \end{bmatrix},$$

where A is $n \times n$, B is $m \times n$, C is $p \times n$, and $D \neq 0$ is $p \times p$ with $n = 3k$ and $m = p = k$. Therefore, it is necessary to solve a system of the form (1.1) within each Newton step. Details on the structure of the blocks A, B, C, and D can be found

TABLE 4
Numerical results for preconditioner $\tilde{\mathscr{P}}_T$, liquid crystal problem.

	Method				
	GMRES		FGMRES		
Size	Iter	CPU	Iter	Iter_{pcg}	CPU
5115	10	0.3535	9	40	0.0406
10235	10	0.6563	9	41	0.0667
20475	10	1.2761	9	39	0.1128
40955	9	2.5051	9	37	0.2093
81915	9	5.4725	9	36	0.4072
163835	9	12.4417	9	35	0.8647
327675	9	29.099	9	29	1.7507

in [18]. Here we note that A is SPD, B has full row rank and is such that BB^T is diagonal (and indeed $BB^T = I_m$ if the unit vector constraints are satisfied exactly), C is rank deficient, and D is tridiagonal and SPD. Hence, $\mathscr{A}$ is nonsingular.[3] Moreover, the condition $D \succ CA^{-1}C^T$ in Theorem 3.10 is satisfied. We also mention that in our experiments we used the following values of the parameters α and β appearing in (4.3): $\alpha = 0.5\alpha_c$ and $\beta = 0.5$, where $\alpha_c \approx 2.721$ is known as the *critical switching value*. For further details we refer the reader to [18]; see also [1, 2].

We present a few results obtained with the block triangular preconditioners $\tilde{\mathscr{P}}_T$ and $\hat{\mathscr{P}}_T$ given in (3.17)–(3.18). The application of these two preconditioners can be performed "exactly" or inexactly, via block backsubstitution. Both versions of the preconditioners require the solution ("exact" and approximate) of linear systems with SPD coefficient matrices $\tilde{S}_C = D + CA^{-1}C^T$, $S_B = BA^{-1}B^T$, and A at each (outer) GMRES or FGMRES iteration. The first two systems are solved via the PCG method, which does not require forming the (full) matrices $D + CA^{-1}C^T$ and $BA^{-1}B^T$ explicitly. The preconditioner used for $D + CA^{-1}C^T$ is the (sparse) matrix $D + CD_A^{-1}C^T$, where D_A is the diagonal part of A. Application of the preconditioner is accomplished via a sparse Cholesky factorization. As for the subsystems with coefficient matrix $BA^{-1}B^T$, the observation that B has (nearly) orthogonal rows suggests that the matrix BAB^T would be a good approximate inverse of $BA^{-1}B^T$, and indeed it was found to be a very effective preconditioner. Note that only sparse matrix-vector products are required for its application, and there is no construction overhead. Within GMRES, we used a rather stringent convergence tolerance (10^{-12}) on the relative residual norm to terminate the inner PCG iterations. For FGMRES, we used looser tolerances, namely, 10^{-1} for systems with $\tilde{S}_C$ and 10^{-3} for systems with S_B. Finally, the matrix A can be factored inexpensively and therefore we used sparse Cholesky factorization to solve all linear systems associated with A.

The results for preconditioners $\tilde{\mathscr{P}}_T$ and $\hat{\mathscr{P}}_T$ are shown in Tables 4 and 5, respectively. In all cases we observe mesh-independent convergence rates, with no deterioration when using the inexact variants of the block preconditioners in place of the exact ones; indeed, in several cases FGMRES even requires one less iteration than GMRES with the "exact" preconditioner. The CPU timings are clearly much better for the inexact variants, especially for larger problems. Overall, the fastest solution times are obtained with FGMRES preconditioned by the inexact variant of the block preconditioner $\hat{\mathscr{P}}_T$. With this method, solution times exhibit almost linear scaling behavior.

[3] We are assuming here that the Hessian is being evaluated away from bifurcation points and turning points; again, see [18].

TABLE 5
Numerical results for preconditioner $\hat{\mathscr{P}}_T$, liquid crystal problem.

	Method				
	GMRES		FGMRES		
Size	Iter	CPU	Iter	Iter_{pcg}	CPU
5115	6	0.2236	6	25	0.0267
10235	6	0.4181	6	24	0.0449
20475	6	0.8210	6	22	0.0806
40955	6	1.6339	6	22	0.1563
81915	6	3.5996	6	19	0.2986
163835	6	8.3431	6	19	0.6427
327675	6	19.459	6	19	1.4221

An indirect comparison with the results reported in [18, Table 6.6] indicates that the proposed approaches are competitive with existing methods, such as preconditioned MINRES with constraint preconditioning.

5. Conclusions. In this paper we have introduced and analyzed several block preconditioners for the solution of sparse linear systems with double saddle point structure. While "standard" techniques for saddle point problems are certainly applicable to systems of the form (1.1), several of the methods investigated in this paper and their analysis make specific use of the 3×3 block structure of the coefficient matrix. Furthermore, different block partitionings (see (1.2)) lead to different solvers with distinct theoretical and practical properties.

Numerical experiments on test problems arising from two distinct application domains show that some of the proposed solvers can be very efficient in situations of practical interest, resulting in rapid convergence (independent of problem size) and scalable behavior. Of course, the performance of each method is highly problem-dependent, and specific information on the spectral properties of the problem at hand may be needed in order to make a good choice. We stress that it is quite possible that some of the methods that were found to be not competitive for the two test problems considered here may well turn out to be useful on other problems and, conversely, some of the methods found to be effective here may well perform poorly on other problems.

In our analysis we assumed that the various preconditioners were implemented exactly. Numerical experiments, however, showed that the rates of convergence do not necessarily deteriorate when inexact solves are used instead, often leading to significantly faster solution times relative to the "exact" versions. This is consistent with previous experience for block preconditioners; see, e.g., [6] or [9].

We conclude by mentioning that there are applications leading to 3×3 block linear systems in which all three diagonal blocks are nonzero:

$$\mathscr{A}u \equiv \begin{bmatrix} A & B^T & C^T \\ B & -E & 0 \\ C & 0 & -D \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \equiv b, \tag{5.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, and $D \in \mathbb{R}^{p\times p}$ are as before, and $E \in \mathbb{R}^{m\times m}$ is SPS see, for example, [19]. When E and D are both nonzero, we can think of this as a "fully stabilized" double saddle point system, whereas the case $E = 0$, $D \neq 0$ corresponds to a "partially stabilized" problem. Block preconditioners for (5.1) with $B = C$ have been developed in [19]. Similar methods should also work well in applications where $B \neq C$.

Acknowledgments. We would like to thank Alison Ramage and Miroslav Tůma for providing the test problems used in the numerical experiments, and Davod Khojasteh Salkuyeh and Maryam Rahimian for helpful comments on a previous draft of the paper. Thanks also to Andy Wathen and two anonymous referees for their careful reading of the manuscript and helpful suggestions. The first author is grateful for the hospitality of the Department of Mathematics and Computer Science at Emory University, where this work was completed.

REFERENCES

[1] J. H. Adler, T. J. Atherton, T. R. Benson, D. B. Emerson, and S. P. MacLachlan, *Energy minimization for liquid crystal equilibrium with electric and flexoelectric effects*, SIAM J. Sci. Comput., 37 (2015), pp. S157–S176.

[2] J. H. Adler, D. B. Emerson, S. P. MacLachlan, and T. A. Manteuffel, *Constrained optimization for liquid crystal equilibria*, SIAM J. Sci. Comput., 38 (2016), pp. B50–B76.

[3] A. Aposporidis, E. Haber, M. A. Olshanskii, and A. Veneziani, *A mixed formulation of the Bingham fluid flow problem: Analysis and numerical solution*, Comput. Methods Appl. Mech. Eng., 200 (2011), pp. 2434–2446.

[4] Z.-Z. Bai, *Eigenvalue estimates for saddle point matrices of Hermitian and indefinite leading blocks*, J. Comput. Appl. Math., 237 (2013), pp. 295–306.

[5] M. Benzi, *Preconditioning techniques for large linear systems: A survey*, J. Comput. Phys., 182 (2002), pp. 417–477.

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

[7] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Ser. Comput. Math., Springer-Verlag, New York, 2013.

[8] P. Chidyagway, S. Ladenheim, and D. B. Szyld, *Constraint preconditioning for the coupled Stokes–Darcy system*, SIAM J. Sci. Comput., 38 (2016), pp. A668–A690.

[9] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers. With Applications in Incompressible Fluid Dynamics*, 2nd ed., Oxford University Press, Oxford, UK, 2014.

[10] N. I. M. Gould and V. Simoncini, *Spectral analysis of saddle point matrices with indefinite leading blocks*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1152–1171.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[12] J. Maryška, M. Rozložník, and M. Tůma, *Schur complement systems in the mixed-hybrid finite element approximation of the potential fluid flow problem*, SIAM J. Sci. Comput., 22 (2000), pp. 704–723.

[13] B. Morini, V. Simoncini, and M. Tani, *Spectral estimates for unreduced symmetric KKT systems arising from Interior Point methods*, Numer. Linear Algebra Appl., 23 (2016), pp. 776–800.

[14] B. Morini, V. Simoncini, and M. Tani, *A comparison of reduced and unredued KKT systems arising from interior point methods*, Comput. Optim. Appl., 68 (2017), pp. 1–27.

[15] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 1969–1972.

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

[17] J. Pestana, *On the eigenvalues and eigenvectors of block triangular preconditioned block matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 517–525.

[18] A. Ramage and E. C. Gartland, Jr., *A preconditioned nullspace method for liquid crystal director modeling*, SIAM J. Sci. Comput., 35 (2013), pp. B226–B247.

[19] S. Rhebergen, G. N. Wells, A. J. Wathen, and R. F. Katz, *Three-field block preconditioners for models of coupled magma/mantle dynamics*, SIAM J. Sci. Comput., 37 (2015), pp. A2270–A2294.

[20] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.