USING CONSTRAINT PRECONDITIONERS WITH REGULARIZED SADDLE-POINT PROBLEMS

H. S. DOLLAR[*†], N. I. M. GOULD[‡], W. H. A. SCHILDERS[§¶], AND A. J. WATHEN[*]

Abstract. The problem of finding good preconditioners for the numerical solution of a certain important class of indefinite linear systems is considered. These systems are of a 2 by 2 block structure in which the (2,2) block (denoted by $-C$) is assumed to be nonzero.

In *Constraint preconditioning for indefinite linear systems,* SIAM J. Matrix Anal. Appl., 21 (2000), Keller, Gould and Wathen introduced the idea of using constraint preconditioners that have a specific 2 by 2 block structure for the case of C being zero. We shall give results concerning the spectrum and form of the eigenvectors when a preconditioner of the form considered by Keller, Gould and Wathen is used but the system we wish to solve may have $C \neq 0$. Numerical results to validate our conclusions are also presented.

1. Introduction. The solution of systems of the form

$$\underbrace{\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}}_{\mathcal{A}} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} c \\ d \end{bmatrix}}_{b}, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$ are symmetric and $B \in \mathbb{R}^{m\times n}$, are often required in optimization and other various fields. We shall assume that $0 < m \leq n$ and B is of full rank. Dollar and Wathen [6] recently proposed a class of incomplete factorizations for saddle-point problems where $C = 0$. Dollar *et al.* [5] then extend this work to problems where C is not necessarily equal to 0. In particular, preconditioners of the form

$$\mathcal{P} = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \tag{1.2}$$

are produced, where $G \in \mathbb{R}^{n\times n}$ is some symmetric matrix.

When $C = 0$, (1.2) is commonly known as a constraint preconditioner [2, 13]. In practice C is often positive semi-definite (and frequently diagonal).

Example 1.1 (Nonlinear Programming). Consider the convex nonlinear optimization problem

$$\text{minimize } f(x) \text{ such that } c(x) \geq 0,$$

where $x \in \mathbb{R}^n$, and $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $-c : \mathbb{R}^n \mapsto \mathbb{R}^{\widehat{m}}$ are convex and twice differentiable. Primal-dual interior point methods [17] for this problem aim to track solutions to the (perturbed) optimality conditions

$$\nabla f(x) = B^T(x)y \text{ and } Yc(x) = \mu e, \tag{1.3}$$

[*]Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford, OX1 3QD, U.K. (hsd@comlab.ox.ac.uk, wathen@comlab.ox.ac.uk).

[†]This authors work was supported by the O.U.C.L. Doctoral Training Account.

[‡]Computational Science and Engineering Department, Rutherford Appleton Laboratory, Chilton, Oxfordshire, OX11 0QX, England, UK. (n.i.m.gould@rl.ac.uk)

[§]Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands. (wil.schilders@philips.com)

[¶]Also, Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, PO Box 513, 5600 MB Eindhoven, The Netherlands.

where y are Lagrange multipliers (dual variables), e is the vector of ones,

$$B(x) = \nabla c(x) \text{ and } Y = \operatorname{diag}\{y_1, y_2, \ldots, y_{\widehat{m}}\},$$

as the positive scalar parameter μ is decreased to zero. The Newton correction $(\Delta x, \Delta y)$ to the solution estimate (x, y) of (1.3) satisfy the equation [3]:

$$\begin{bmatrix} A(x,y) & -B^T(x) \\ YB(x) & C(x) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} -\nabla f(x) + B^T(x)y \\ -Yc(x) + \mu e \end{bmatrix},$$

where

$$A(x,y) = \nabla_{xx} f(x) - \sum_{i=1}^{\widehat{m}} y_i \nabla_{xx} c_i(x) \text{ and } C(x) = \operatorname{diag}\{c_1(x), c_2(x), \ldots, c_{\widehat{m}}(x)\}.$$

It is common to eliminate the variables Δy from the Newton system. Since this may introduce unwarranted ill conditioning, it is often better [9] to isolate the effects of poor conditioning by partitioning the constraints so that the values of those indexed by $\mathcal{I}$ are "large" while those indexed by $\mathcal{A}$ are "small", and instead to solve

$$\begin{bmatrix} A + B_{\mathcal{I}}^T C_{\mathcal{I}}^{-1} Y_{\mathcal{I}} B_{\mathcal{I}} & B_{\mathcal{A}}^T \\ B_{\mathcal{A}} & -C_{\mathcal{A}} Y_{\mathcal{A}}^{-1} \end{bmatrix} \begin{bmatrix} \Delta x \\ -\Delta y_{\mathcal{A}} \end{bmatrix} = \begin{bmatrix} -\nabla f + B_{\mathcal{A}}^T y_{\mathcal{A}} + \mu B_{\mathcal{I}}^T C_{\mathcal{I}}^{-1} e \\ -c_{\mathcal{A}} + \mu Y_{\mathcal{A}}^{-1} e \end{bmatrix}$$

where, for brevity, we have dropped the dependence on x and y. The matrix $C_{\mathcal{A}} Y_{\mathcal{A}}^{-1}$ is symmetric and positive definite; as the iterates approach optimality, the entries of this matrix become small. The entries of $B_{\mathcal{I}}^T C_{\mathcal{I}}^{-1} Y_{\mathcal{I}} B_{\mathcal{I}}$ also become small when close to optimality.

Example 1.2 (Stokes). Mixed finite element (and other) discretisation of the Stokes equations

$$\begin{aligned} -\nabla^2 \vec{u} + \nabla p &= \vec{f} \quad \text{in } \Omega \\ \nabla \cdot \vec{u} &= 0 \quad \text{in } \Omega, \end{aligned}$$

for the fluid velocity $\vec{u}$ and pressure p in the domain $\Omega \subset \mathbb{R}^2$ or $\mathbb{R}^3$ yields linear systems in the saddle-point form (1.1) (for derivation and the following properties of this example see [7]). The symmetric block A arises from the diffusion terms $-\nabla^2 \vec{u}$ and B^T represents the discrete gradient operator whilst B represents its adjoint, the (negative) divergence. When (inf-sup) stable mixed finite element spaces are employed, $C = 0$, however for equal order and other spaces which are not inherently stable, stabilised formulations yield symmetric and positive semi-definite matrices C which typically have a large-dimensional kernel - for example the famous $\boldsymbol{Q}_1$–$\boldsymbol{P}_0$ element which has piecewise bilinear velocities and piecewise constant pressures in 2-dimensions, C typically has a kernel of dimension $m/4$.

In Section 2, we shall give an overview of the known spectral properties for $\mathcal{P}^{-1}\mathcal{A}$. In interior-point methods a sequence of such problems are solved with the entries in C generally becoming small as the optimization iteration progresses. That is, the regularization is successively reduced as the optimizer gets closer to the minimum. For the Stokes problem, the entries of C are generally small since they scale with the underlying mesh size and so reduce for finer grids. This motivates us to look at the spectral properties of $\widetilde{\mathcal{P}}^{-1}\mathcal{A}$, where

$$\widetilde{\mathcal{P}} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}, \tag{1.4}$$

but $C \neq 0$, Section 3.

The obvious advantage in being able to use such a constraint preconditioner is as follows: if B remains constant in each system of the form (1.1), and we choose G in our preconditioner to remain constant, then the preconditioner $\widetilde{\mathcal{P}}$ will be unchanged. Any factorizations required to carry out the preconditioning steps in a iterative method will only need to be done once and then used during each execution of the iterative method of choice, instead of carrying out the factorizations at the beginning of each iterative method.

For symmetric (and in general normal) matrix systems, the convergence of an applicable iterative method is determined by the distribution of the eigenvalues of the coefficient matrix. It is often desirable for the number of distinct eigenvalues to be small so that the rate of convergence is rapid. For non-normal systems the convergence is not so readily described, see [12, page 6].

2. Spectral properties of $\mathcal{P}^{-1}\mathcal{A}$. The spectral properties of $\mathcal{P}^{-1}\mathcal{A}$ for the case $C = 0$ where analyzed by Keller, Gould, and Wathen [13]. The proof of the following theorem can be found in [13].

THEOREM 2.1. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$ be a symmetric and indefinite matrix of the form*

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

where $A \in \mathbb{R}^{n\times n}$ is symmetric and $B \in \mathbb{R}^{m\times n}$ is of full rank. Assume Z is an $n \times (n-m)$ basis for the nullspace of B. Preconditioning $\mathcal{A}$ by a matrix of the form

$$\mathcal{P} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix},$$

where $G \in \mathbb{R}^{n\times n}$ is symmetric, and $B \in \mathbb{R}^{m\times n}$ is as above, implies that

- *the matrix $\mathcal{P}^{-1}\mathcal{A}$ has*
  1. *an eigenvalue at 1 with multiplicity $2m$, and*
  2. *$n - m$ eigenvalues λ which are defined by the generalized eigenvalue problem $Z^TAZx_z = \lambda Z^TGZx_z$,*
- *the dimension of the Krylov subspace $\mathcal{K}(\mathcal{P}^{-1}\mathcal{A}, b)$ is at most $n - m + 2$.*

Keller, Gould and Wathen [13] also define the form of the eigenvectors for such preconditioned systems.

THEOREM 2.2. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$ be a symmetric and indefinite matrix of the form*

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

where $A \in \mathbb{R}^{n\times n}$ is symmetric and $B \in \mathbb{R}^{m\times n}$ is of full rank. Assume the preconditioner $\mathcal{P}$ is defined by a matrix of the form

$$\mathcal{P} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix},$$

where $G \in \mathbb{R}^{n\times n}$ is symmetric, and $B \in \mathbb{R}^{m\times n}$ is as above. Let Z denote an $n \times (n-m)$ basis for the nullspace of B and suppose that Z^TGZ is positive definite. The preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ has $n + m$ eigenvalues as defined by Theorem 2.1 and $m + i + j$ linearly independent eigenvectors. There are

1. *m eigenvectors of the form $\begin{bmatrix} 0^T & 0^T & y^T \end{bmatrix}^T$ that correspond to the case $\lambda = 1$;*
2. *i $(0 \le i \le n)$ eigenvectors of the form $\begin{bmatrix} x_z^T & x_y^T & y^T \end{bmatrix}^T$ arising from $Aw = \sigma G w$ with $w = \begin{bmatrix} x_y^T & x_z^T \end{bmatrix}^T$ linearly independent, $\sigma = 1$, and $\lambda = 1$; and*
3. *j $(0 \le j \le n-m)$ eigenvectors of the form $\begin{bmatrix} x_z^T & 0^T & y^T \end{bmatrix}^T$ that correspond to the case $\lambda \neq 1$.*

If either Z^TAZ or Z^TGZ are positive definite, then the indefinite preconditioner $\mathcal{P}$ applied to the indefinite saddle point matrix $\mathcal{A}$ with $C = 0$ yields a preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ which has real eigenvalues, [13]. If both Z^TAZ and Z^TGZ are positive definite, then we can use a projected preconditioned conjugate gradient method to find x and y, see [10].

Analogous results for the case $C \neq 0$ can be found in [4]: in particular, the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ is shown to have at least $2m - \operatorname{rank}(C)$ unit eigenvalues.

3. An alternative preconditioner for the case $C \neq 0$. Suppose that instead of preconditioning $\mathcal{A}$ by $\mathcal{P}$, we precondition $\mathcal{A}$ by $\widetilde{\mathcal{P}}$, where $\widetilde{\mathcal{P}}$ is defined in (1.4). The decision to investigate this form of preconditioner is motivated in Section 1.

THEOREM 3.1. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$ be a symmetric and indefinite matrix of the form*

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

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$ are symmetric and $B \in \mathbb{R}^{m\times n}$ is of full rank. We shall assume that C has rank p and is factored as EDE^T, where $E \in \mathbb{R}^{m\times p}$ has orthogonal columns and $D \in \mathbb{R}^{p\times p}$ is non-singular, $Z \in \mathbb{R}^{n\times(n-m)}$ is a basis for the nullspace of B and $Y \in \mathbb{R}^{n\times m}$ is such that $\begin{bmatrix} Y & Z \end{bmatrix}$ spans $\mathbb{R}^n$. Preconditioning $\mathcal{A}$ by a matrix of the form

$$\widetilde{\mathcal{P}} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix},$$

where $G \in \mathbb{R}^{n\times n}$ is symmetric, and $B \in \mathbb{R}^{m\times n}$is as above, implies that the matrix $\widetilde{\mathcal{P}}^{-1}\mathcal{A}$ has at most $j+k+1$ distinct eigenvalues as defined below:

- *at least $2(m-p)$ eigenvalues at 1,*
- *at most $n-m$ eigenvalues defined by the generalized eigenvalue problem*

$$Z^TAZx_z = \lambda Z^TGZx_z$$

subject to there existing some y_f such that $[Y^THZ - \lambda Y^TGZ]x_z = (\lambda - 1)RFy_f$. Of these eigenvalues, j $(0 \le j \le n-m)$ are non-unit,

- *at most $n-m+2p$ eigenvalues defined by the generalized eigenvalue problem*

$$0 = \lambda^2 B^TED^{-1}E^TBw - \lambda(G + 2B^TED^{-1}E^TB)w + (A + B^TED^{-1}E^TB)w,$$

where $w = \begin{bmatrix} x_y^T & x_z^T \end{bmatrix}^T$, subject to $E^TBYx_y \neq 0$. Of these, k $(0 \le k \le n-m+2p)$ are not equal to 1,

- *$0 \le j+k \le n-m+2p$.*

Proof. Let $QR = \begin{bmatrix} Y & Z \end{bmatrix}\begin{bmatrix} R^T & 0^T \end{bmatrix}^T$ be an orthogonal factorization of B^T, where $R \in \mathbb{R}^{m\times m}$ is upper triangular, $Y \in \mathbb{R}^{n\times m}$, and $Z \in \mathbb{R}^{n\times(n-m)}$ is a basis for the nullspace of B. We can therefore write any $x \in \mathbb{R}^n$ as $x = Zx_z + Yx_y$, where

$x_z \in \mathbb{R}^{n-m}$, $x_y \in \mathbb{R}^m$ are unique vectors. Similarly, by our assumptions, any $y \in \mathbb{R}^m$ can be written as $y = Fy_f + Ey_e$ where $y_f \in \mathbb{R}^{m-p}$, $y_e \in \mathbb{R}^p$ are unique vectors. Premultiplying the generalized eigenvalue problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.1}$$

by the nonsingular and square matrix

$$\begin{bmatrix} Z^T & 0 \\ Y^T & 0 \\ 0 & F^T \\ 0 & E^T \end{bmatrix}$$

and using the substitution above gives

$$\begin{bmatrix} Z^TAZ & Z^TAY & 0 & 0 \\ Y^TAZ & Y^TAY & RF & RE \\ 0 & F^TR^T & 0 & 0 \\ 0 & E^TR^T & 0 & -D \end{bmatrix} \begin{bmatrix} x_z \\ x_y \\ y_f \\ y_e \end{bmatrix} = \lambda \begin{bmatrix} Z^TGZ & Z^TGY & 0 & 0 \\ Y^TGZ & Y^TGY & RF & RE \\ 0 & F^TR^T & 0 & 0 \\ 0 & E^TR^T & 0 & 0 \end{bmatrix} \begin{bmatrix} x_z \\ x_y \\ y_f \\ y_e \end{bmatrix}, \tag{3.2}$$

where we made use of the the equalities $BZ = 0$, $CF = 0$, $E^TE = I$ and $R = (BY)^T$. Expanding out the general eigenvalue problem (3.2) yields

$$\begin{aligned} Z^TAZx_z + Z^TAYx_y &= \lambda[Z^TGZx_z + Z^TGYx_y], && (3.3) \\ Y^TAZx_z + Y^TAYx_y + RFy_f + REy_e &= \lambda[Y^TGZx_z + Y^TGYx_y \\ &\quad +RFy_f + REy_e], && (3.4) \\ F^TR^Tx_y &= \lambda F^TR^Tx_y, && (3.5) \\ E^TR^Tx_y - Dy_e &= \lambda E^TR^Tx_y. && (3.6) \end{aligned}$$

From (3.5), it may be deduced that either $\lambda = 1$ or $R^Tx_y \in Null(F^T)$. In the former case, (3.6) implies that $y_e = 0$, whilst (3.3) and (3.4) simplify to

$$\begin{aligned} Z^TAZx_z + Z^TAYx_y &= Z^TGZx_z + Z^TGYx_y, \\ Y^TAZx_z + Y^TAYx_y &= Y^TGZx_z + Y^TGYx_y, \end{aligned}$$

which can consequently be written as

$$Q^TAQw = Q^TGQw, \tag{3.7}$$

where $Q = \begin{bmatrix} Y & Z \end{bmatrix}$ and $w = \begin{bmatrix} x_y^T & x_z^T \end{bmatrix}^T$. Since Q is orthogonal, the general eigenvalue problem (3.7) is equivalent to considering

$$Aw = \sigma Gw, \tag{3.8}$$

where $w \neq 0$ if and only if $\sigma = 1$. There are $m - p$ linearly independent eigenvectors $\begin{bmatrix} 0^T & 0^T & y_f^T & 0^T \end{bmatrix}^T$ corresponding to $w = 0$, and a further i $(1 \leq i \leq n)$ linearly independent eigenvectors (corresponding to eigenvalues $\sigma = 1$ of (3.8)).

Now, suppose that $\lambda \neq 1$, in which case $R^Tx_y \in Null(F^T)$. Equation (3.6) also implies that

$$(1 - \lambda)E^TR^Tx_y = Dy_e. \tag{3.9}$$

We have two cases, either $Cy = 0$ or $Cy \neq 0$. If the former holds, then $y_e = 0$ and (3.9) implies that $R^T x_y \in Null(E^T)$, as well as $R^T x_y \in Null(F^T)$. Hence, $x_y = 0$.

From (3.3) and (3.4) we obtain

$$Z^T AZx_z = \lambda Z^T GZx_z, \tag{3.10}$$

$$Y^T AZx_z + RFy_f = \lambda[Y^T GZx_z + RFy_f]. \tag{3.11}$$

The generalized eigenvalue problem (3.10) defines $n-m$ eigenvalues, where j ($1 \leq j \leq n-m$) of these are not equal to 1 and for which two cases have to be distinguished. If $x_z \neq 0$, y_f must satisfy

$$[Y^T AZ - \lambda Y^T GZ]x_z = (\lambda - 1)RFy_f,$$

which follows that the corresponding eigenvectors are defined by $\begin{bmatrix} x_z^T & 0^T & y_f^T & 0^T \end{bmatrix}^T$. If $x_z = 0$, then from (3.11) we deduce that $y_f = 0$ since $\lambda \neq 1$. As $\begin{bmatrix} x_z^T & x_y^T & y_f^T & y_e^T \end{bmatrix}^T = 0$ in this case, no extra eigenvalues arise.

Suppose that $Cy \neq 0$, then any y satisfying this can be written as $y = Fy_f + Ey_e$, where $y_e \neq 0$. The fact that the matrix D is non-singular along with (3.9) implies that

$$y_e = (1 - \lambda)D^{-1}E^T R^T x_y. \tag{3.12}$$

Equations (3.3) and (3.4), along with $\lambda \neq 1$ imply that $y_f = 0$. Substituting this and (3.12) into (3.4), and rearranging, gives

$$\begin{aligned} 0 = {} & \lambda^2 \begin{bmatrix} Y^T B^T ED^{-1}E^T BY & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_y \\ x_z \end{bmatrix} \\ & -\lambda \begin{bmatrix} Y^T(G + 2B^T ED^{-1}E^T B)Y & Y^T GZ \\ Z^T GY & Z^T GZ \end{bmatrix} \begin{bmatrix} x_y \\ x_z \end{bmatrix} \\ & + \begin{bmatrix} Y^T(A + B^T ED^{-1}E^T B)Y & Y^T AZ \\ Z^T AY & Z^T AZ \end{bmatrix} \begin{bmatrix} x_y \\ x_z \end{bmatrix}. \end{aligned} \tag{3.13}$$

Using the fact that $Q = \begin{bmatrix} Y & Z \end{bmatrix}$, $BZ = 0$ we can show that

$$Q^T B^T ED^{-1}E^T BQ = \begin{bmatrix} Y^T B^T ED^{-1}E^T BY & 0 \\ 0 & 0 \end{bmatrix}.$$

We therefore obtain the quadratic eigenvalue problem

$$\begin{aligned} 0 = {} & \lambda^2 Q^T B^T ED^{-1}E^T BQw - \lambda Q^T(G + 2B^T ED^{-1}E^T B)Qw \\ & + Q^T(A + B^T ED^{-1}E^T B)Qw, \end{aligned} \tag{3.14}$$

where $w = \begin{bmatrix} x_y^T & x_z^T \end{bmatrix}^T$. Once again, the orthogonality of Q implies that the quadratic eigenvalue problem (3.13) is equivalent to

$$\begin{aligned} 0 = {} & \lambda^2 B^T ED^{-1}E^T Bw - \lambda(G + 2B^T ED^{-1}E^T B)w \\ & + (A + B^T ED^{-1}E^T B)w. \end{aligned} \tag{3.15}$$

The generalized quadratic eigenvalue problem (3.15) defines at most $n - m + 2p$ eigenvalues for which $R^T x_y \in Null(F^T)$, and $Cy \neq 0$ are also satisfied, but at most p linearly independent eigenvectors. Of these, k ($0 \leq k \leq n - m + 2p$) correspond

to the case $\lambda \neq 1$. It follows that the corresponding eigenvectors are defined by $[\ x_z^T \quad x_y^T \quad 0^T \quad y_e^T\ ]^T$. Note that (3.10) is a subproblem of (3.13) when (3.11) is only satisfied with $y_f = 0$. But this would then correspond to $x_y = 0$ and $y_e = 0$, so $0 \leq j + k \leq n - m + 2p$.
□

The definition of some of the eigenvalues through a quadratic eigenvalue problem is very interesting and shall be examined in more detail later on.

THEOREM 3.2. *Let $\mathcal{A}, \widetilde{P} \in \mathbb{R}^{(n+m)\times(n+m)}$ and their sub-blocks be as defined in Theorem 3.1 (using the same notation and assumptions). Suppose that Z^TGZ is positive definite. Then the matrix $\widetilde{P}^{-1}\mathcal{A}$ has $n+m$ eigenvalues and $m-p+i+j+k$ linearly independent eigenvectors. There are*

1. *$m-p$ eigenvectors of the form $[\ 0^T \quad 0^T \quad y_f^T \quad 0^T\ ]^T$ that correspond to the case $\lambda = 1$;*
2. *i $(0 \leq i \leq n)$ eigenvectors of the form $[\ x_z^T \quad x_y^T \quad y_f^T \quad 0^T\ ]^T$ arising from $Hw = \sigma Gw$ with $w = [\ x_y^T \quad x_z^T\ ]^T$ linearly independent, $\sigma = 1$, and $\lambda = 1$;*
3. *j $(0 \leq j \leq n-m)$ eigenvectors of the form $[\ x_z^T \quad 0^T \quad y_f^T \quad 0^T\ ]^T$ that correspond to the case $\lambda \neq 1$ and $Cy = 0$ with $y = Fy_f + Ey_e$;*
4. *k $(0 \leq k \leq n-m+p)$ eigenvectors of the form $[\ x_z^T \quad x_y^T \quad 0^T \quad y_e^T\ ]^T$ that correspond to the case $\lambda \neq 1$ and $Cy \neq 0$ with $y = Fy_f + Ey_e$;*
5. *$0 \leq j + k \leq n - m + p$.*

Proof. We need only prove that the $m-p+i+j+k$ eigenvectors of $\widetilde{P}^{-1}\mathcal{A}$ defined in the proof of Theorem 3.1 are linearly independent.

We need to show that

$$\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ y_{f1}^{(1)} & \cdots & y_{f(m-p)}^{(1)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_{m-p}^{(1)} \end{bmatrix} + \begin{bmatrix} x_{z1}^{(2)} & \cdots & x_{zi}^{(2)} \\ x_{y1}^{(2)} & \cdots & x_{yi}^{(2)} \\ y_{f1}^{(2)} & \cdots & y_{fi}^{(2)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(2)} \\ \vdots \\ a_i^{(2)} \end{bmatrix} \tag{3.16}$$
$$+ \begin{bmatrix} x_{z1}^{(3)} & \cdots & x_{zj}^{(3)} \\ 0 & \cdots & 0 \\ y_{f1}^{(3)} & \cdots & y_{fj}^{(3)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(3)} \\ \vdots \\ a_j^{(3)} \end{bmatrix} + \begin{bmatrix} x_{z1}^{(4)} & \cdots & x_{zk}^{(4)} \\ x_{y1}^{(4)} & \cdots & x_{yk}^{(4)} \\ 0 & \cdots & 0 \\ y_{e1}^{(4)} & \cdots & y_{ek}^{(4)} \end{bmatrix} \begin{bmatrix} a_1^{(4)} \\ \vdots \\ a_k^{(4)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

implies that the vectors $a^{(l)}$ $(l = 1, \ldots, 4)$ are zero vectors. Multiplying (3.16) by $\mathcal{A}$ and $\widetilde{\mathcal{P}}^{-1}$, and recalling that in the previous equation the first matrix arises from $\lambda_l = 1$ $(l = 1, \ldots, m)$, the second matrix from the case that $\lambda_l = 1$ and $\omega_l = 1$ $(l = 1, \ldots, i)$, the third matrix from $\lambda_l \neq 1$ $(l = 1, \ldots, j)$ and $Cy = 0$, and the last matrix from $\lambda_l \neq 1$ $(l = 1, \ldots, k)$ and $Cy \neq 0$, gives

$$\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ y_{f1}^{(1)} & \cdots & y_{f(m-p)}^{(1)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_{m-p}^{(1)} \end{bmatrix} + \begin{bmatrix} x_{z1}^{(2)} & \cdots & x_{zi}^{(2)} \\ x_{y1}^{(2)} & \cdots & x_{yi}^{(2)} \\ y_{f1}^{(2)} & \cdots & y_{fi}^{(2)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(2)} \\ \vdots \\ a_i^{(2)} \end{bmatrix} \tag{3.17}$$
$$+ \begin{bmatrix} x_{z1}^{(3)} & \cdots & x_{zj}^{(3)} \\ 0 & \cdots & 0 \\ y_{f1}^{(3)} & \cdots & y_{fj}^{(3)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1^{(3)} a_1^{(3)} \\ \vdots \\ \lambda_j^{(3)} a_j^{(3)} \end{bmatrix} + \begin{bmatrix} x_{z1}^{(4)} & \cdots & x_{zk}^{(4)} \\ x_{y1}^{(4)} & \cdots & x_{yk}^{(4)} \\ 0 & \cdots & 0 \\ y_{e1}^{(4)} & \cdots & y_{ek}^{(4)} \end{bmatrix} \begin{bmatrix} \lambda_1^{(4)} a_1^{(4)} \\ \vdots \\ \lambda_k^{(4)} a_k^{(4)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Subtracting (3.16) from (3.17) gives

$$\begin{bmatrix} x_{z1}^{(3)} & \cdots & x_{zj}^{(3)} \\ 0 & \cdots & 0 \\ y_{f1}^{(3)} & \cdots & y_{fj}^{(3)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} (\lambda_1^{(3)}-1)a_1^{(3)} \\ \vdots \\ (\lambda_j^{(3)}-1)a_j^{(3)} \end{bmatrix} + \begin{bmatrix} x_{z1}^{(4)} & \cdots & x_{zk}^{(4)} \\ x_{y1}^{(4)} & \cdots & x_{yk}^{(4)} \\ 0 & \cdots & 0 \\ y_{e1}^{(4)} & \cdots & y_{ek}^{(4)} \end{bmatrix} \begin{bmatrix} (\lambda_1^{(4)}-1)a_1^{(4)} \\ \vdots \\ (\lambda_k^{(4)}-1)a_k^{(4)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.18}$$

The linear independence of $x_{yl}^{(4)}$ $(l = 1, \ldots, k)$ in (3.15) gives rise to $(\lambda_l^{(4)} - 1)a_l^{(4)} = 0$ $(l = 1, \ldots, k)$. The eigenvalues $\lambda_l^{(4)}$ $(l = 1, \ldots, j)$ are non-unit which implies that $a_l^{(4)} = 0$ $(l = 1, \ldots, j)$. The assumption that $Z^T GZ$ is positive definite implies that $x_{zl}^{(3)}$ $(l = 1, \ldots, j)$ in (3.18) are linearly independent, and hence $a_l^{(3)} = 0$ $(l = 1, \ldots, j)$.

We also have linear independence of $[\ x_{zl}^{(2)T} \quad x_{yl}^{(2)T}\ ]^T$ $(l = 1, \ldots, i)$, and thus $a_l^{(2)} = 0$ $(l = 1, \ldots, i)$. Equation 3.16 simplifies to

$$\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ y_{f1}^{(1)} & \cdots & y_{f(m-p)}^{(1)} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_{(m-p)}^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

However, $y_{fl}^{(1)}$ $(l = 1, \ldots, m-p)$ are linearly independent giving $a_l^{(1)} = 0$.
□

REMARK 3.3. *$\widetilde{\mathcal{P}}^{-1}\mathcal{A}$ has at least $2(m-p)$ unit eigenvalues, but there is no guarantee that the associated eigenvectors are all linearly independent. However, we can divide these eigenvectors into two groups such that all the eigenvectors in a group are linearly independent and each group has at least $m-p$ members.*

3.1. Analysis of the quadratic eigenvalue problem. We note that the quadratic eigenvalue problem (3.15) can have negative and complex eigenvalues. The following theorem gives sufficient conditions for general quadratic eigenvalue problems to have real and positive eigenvalues.

THEOREM 3.4. *Consider the quadratic eigenvalue problem*

$$\left(\lambda^2 K - \lambda L + M\right) x = 0, \tag{3.19}$$

where $M, L, K\mathbb{R}^{n\times n}$. are symmetric positive definite, $K \in \mathbb{R}^{n\times n}$ is symmetric positive semidefinite. Define $\gamma(M, L, K)$ to be

$$\gamma(M, L, K) = \min\left\{(x^T Lx)^2 - 4(x^T Mx)(x^T Kx) : \|x\|_2 = 1\right\}.$$

If M, L are symmetric positive definite, K is symmetric positive semidefinite and $\gamma(M, L, K) > 0$, then the eigenvalues λ are real and positive.

Proof. From [16, Section 1] we know that under our assumptions the quadratic eigenvalue problem

$$\left(\mu^2 M + \mu L + K\right) x = 0$$

has real and negative eigenvalues. Suppose we divide this equation by μ^2 and set $\lambda = -1/\mu$. The quadratic eigenvalue problem (3.19) is obtained, and since μ is real and negative, λ is real and positive. □

We would like to be able to use the above theorem to show that, under suitable assumptions, all the eigenvalues of $\widetilde{K}^{-1}\mathcal{H}$ are real and positive. Let

$$\widetilde{D} = B^T E D^{-1} E^T B, \tag{3.20}$$

where D and E are as defined in Theorem 3.1. If we assume that $A + \widetilde{D}$ is positive definite, then we may write $A + \widetilde{D} = R^T R$ for some nonsingular matrix R. The quadratic eigenvalue (3.15) is similar to

$$\left(\lambda^2 R^{-T}\widetilde{D}R^{-1} - \lambda R^{-T}(G + 2\widetilde{D})R^{-1} + I\right) z = 0,$$

where $z = Rw$. Thus, if we assume that $A + \widetilde{D}$ and $G + 2\widetilde{D}$ are positive definite, and can show that

$$\gamma(I, R^{-T}(G + 2\widetilde{D})R^{-1}, R^{-T}\widetilde{D}R^{-1}) > 0,$$

where $\gamma(\cdot,\cdot,\cdot)$ is as defined in Theorem 3.4, then we can apply the above theorem to show that (3.15) has real and positive eigenvalues.

Let us assume that $\|z\|_2 = 1$, then

$$\begin{aligned}
&\left(z^T R^{-T}\left(G + 2\widetilde{D}\right)R^{-1}z\right)^2 - 4z^T z z^T R^{-T}\widetilde{D}R^{-1}z \\
&= \left(z^T R^{-T}GR^{-1}z + 2z^T R^{-T}\widetilde{D}R^{-1}z\right)^2 - 4z^T R^{-T}\widetilde{D}R^{-1}z \\
&= \left(z^T R^{-T}GR^{-1}z\right)^2 + 4z^T R^{-T}\widetilde{D}R^{-1}z\left(z^T R^{-T}GR^{-1}z + z^T R^{-T}\widetilde{D}R^{-1}z - 1\right) \\
&= \left(w^T Gw\right)^2 + 4w^T\widetilde{D}w\left(w^T Gw + w^T\widetilde{D}w - 1\right),
\end{aligned} \tag{3.21}$$

where $1 = \|z\|_2 = \|Rw\|_2 = \|w\|_{A+\widetilde{D}}$. Clearly, we can guarantee that (3.21) is positive if

$$w^T Gw + w^T\widetilde{D}w > 1 \quad \text{for all } w \text{ such that } \|w\|_{A+\widetilde{D}} = 1,$$

that is

$$\frac{w^T Gw + w^T\widetilde{D}w}{w^T\left(A + \widetilde{D}\right)w} > \frac{w^T\left(A + \widetilde{D}\right)w}{w^T\left(A + \widetilde{D}\right)w} \quad \text{for all } w \neq 0.$$

Rearranging we find that we require

$$w^T Gw > w^T Aw$$

for all $w \neq 0$. Thus we need only scale any positive definite G such that $\frac{w^T Gw}{w^T w} > \|A\|_2^2$ for all $w \neq 0$ to guarantee that (3.21) is positive for all w such that $\|w\|_{A+\widetilde{D}} = 1$. For example, we could choose $G = \alpha I$, where $\alpha > \|A\|_2^2$.

If $G + 2\widetilde{D}$ and $A + \widetilde{D}$ are positive definite, then $Z^T GZ$ and $Z^T AZ$ are positive definite. Using the above in conjunction with Theorem 3.1 we obtain:

THEOREM 3.5. *Suppose that A, B, C, D, E, G, and Z are as defined in Theorem 3.1 and $\widetilde{D}$ is as defined in (3.20). Further, assume that $A + \widetilde{D}$ and $G + 2\widetilde{D}$ are symmetric positive definite, $\widetilde{D}$ is symmetric positive semidefinite and*

$$\min\left\{(z^T Gz)^2 + 4(z^T\widetilde{D}z)(z^T Gz + z^T\widetilde{D}z - 1) : \|z\|_{A+\widetilde{D}} = 1\right\} > 0,$$

then all the eigenvalues of $\widetilde{P}^{-1}\mathcal{A}$ are real and positive.

4. Convergence. In the context of this paper, the convergence of an iterative method under preconditioning is not only influenced by the spectral properties of the coefficient matrix, but also by the relationship between m, n and p. We can determine an upper bound on the number of iterations of an appropriate Krylov subspace method by considering minimum polynomials of the coefficient matrix.

DEFINITION 4.1. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$. The monic polynomial f of minimum degree such that $f(\mathcal{A}) = 0$ is called the minimum polynomial of $\mathcal{A}$.*

Krylov subspace theory states that iteration with any method with an optimality property, e.g. GMRES, will terminate when the degree of the minimum polynomial is attained, [15]. In particular, the degree of the minimum polynomial is equal to the dimension of the corresponding Krylov subspace (for general b), [14, Proposition 6.1].

THEOREM 4.2. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$ be a symmetric and indefinite matrix of the form*

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

where $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$ are symmetric and $B \in \mathbb{R}^{m\times n}$ is of full rank. We shall assume that C has rank p. Let the preconditioner $\widetilde{\mathcal{P}}$ be defined by a matrix of the form

$$\widetilde{\mathcal{P}} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix},$$

where $G \in \mathbb{R}^{n\times n}$ is symmetric, $G \neq A$, and $B \in \mathbb{R}^{m\times n}$ is as above. Suppose that Z^TGZ is positive definite. The dimension of the Krylov subspace $\mathcal{K}(\widetilde{\mathcal{P}}^{-1}\mathcal{A}, b)$ is at most $\min\{n - m + 2p + 2, n + m\}$.

Proof. As in the proof to Theorem 3.1, the generalized eigenvalue problem can be written as

$$\begin{bmatrix} Z^TAZ & Z^TAY & 0 & 0 \\ Y^TAZ & Y^TAY & RF & RE \\ 0 & F^TR^T & 0 & 0 \\ 0 & E^TR^T & 0 & -D \end{bmatrix} \begin{bmatrix} x_z \\ x_y \\ y_f \\ y_e \end{bmatrix} = \lambda \begin{bmatrix} Z^TGZ & Z^TGY & 0 & 0 \\ Y^TGZ & Y^TGY & RF & RE \\ 0 & F^TR^T & 0 & 0 \\ 0 & E^TR^T & 0 & 0 \end{bmatrix} \begin{bmatrix} x_z \\ x_y \\ y_f \\ y_e \end{bmatrix}, \tag{4.1}$$

where E, F, R, Y and Z are also defined in Theorem 3.1. Performing a simultaneous sequence of row and column interchanges on both matrices in (4.1) reveals two block triangular matrices

$$\widehat{\mathcal{A}} = \begin{bmatrix} Y^TAY & Y^TAZ & RE & RF \\ Z^TAY & Z^TAZ & 0 & 0 \\ E^TR^T & 0 & -D & 0 \\ F^TR^T & 0 & 0 & 0 \end{bmatrix}, \quad \widehat{\mathcal{P}} = \begin{bmatrix} Y^TGY & Y^TGZ & RE & RF \\ Z^TGY & Z^TGZ & 0 & 0 \\ E^TR^T & 0 & 0 & 0 \\ F^TR^T & 0 & 0 & 0 \end{bmatrix},$$

and, hence, the preconditioned matrix $\widetilde{\mathcal{P}}^{-1}\mathcal{A}$ is similar to

$$\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} = \begin{bmatrix} \Theta_1 & 0 \\ \Theta_2 & I \end{bmatrix}, \tag{4.2}$$

where the precise forms of $\Theta_1 \in \mathbb{R}^{(n+p)\times(n+p)}$ and $\Theta_2 \in \mathbb{R}^{(m-p)\times(n+p)}$ are irrelevant for the argument that follows.

From the eigenvalue derivation in Section 3, it is evident that the characteristic polynomial of the preconditioned linear system (4.2) is

$$(\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - I)^{2(m-p)} \prod_{i=1}^{n-m+2p} (\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - \lambda_i I).$$

In order to prove the upper bound on the Krylov subspace dimension, we need to show that the order of the minimum degree polynomial is less than or equal to $\min\{n - m + 2p + 2, n + m\}$.

Expanding the polynomial $(\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - I)\prod_{i=1}^{n-m+2p}(\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - \lambda_i I)$ of degree $n - m + 2p + 1$, we obtain

$$\left[\begin{array}{cc} (\Theta_1 - I)\prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) & 0 \\ \Theta_2 \prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) & 0 \end{array} \right].$$

The eigenvalues of Θ_1 are 1 (with multiplicity $m-p$) and $\{\lambda_i\}$, $i = 1, \ldots, n-m+2p$. Since Θ_1 has a full set of linearly independent eigenvectors, Θ_1 is diagonalizable. Hence, $(\Theta_1 - I)\prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) = 0$. We therefore obtain

$$(\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - I) \prod_{i=1}^{n-m+2p} (\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - \lambda_i I) = \left[\begin{array}{cc} 0 & 0 \\ \Theta_2 \prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) & 0 \end{array} \right]. \tag{4.3}$$

If $\Theta_2 \prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) = 0$, then the order of the minimum polynomial of $\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}}$ is less than or equal to $\min\{n-m+2p+1, n+m\}$. If $\Theta_2 \prod_{i=1}^{n-m+2p}(\Theta_1 - \lambda_i I) \neq 0$, then the dimension of $\mathcal{K}(\widetilde{\mathcal{P}}^{-1}\mathcal{A}, b)$ is at most $\min\{n-m+2p+2, n+m\}$ since multiplication of (4.3) by another factor $(\widehat{\mathcal{P}}^{-1}\widehat{\mathcal{A}} - I)$ gives the zero matrix.
□

4.1. Clustering of eigenvalues when $\|C\|$ is small. When using interior-point methods to solve optimization problems, the matrix C is generally diagonal and of full rank. In this case, Theorem 4.2 would suggest that there is little advantage of using a constraint preconditioner of the form $\widetilde{\mathcal{P}}$ over any other preconditioner. However, in interior-point methods the entries of C also become small as we grow close to optimality and, hence, $\|C\|$ is small. In the following we shall assume that the norm considered is the ℓ_2 norm, but the results can be generalized to other norms.

THEOREM 4.3. *Let $\zeta > 0$, $\delta \geq 0$, $\varepsilon \geq 0$ and $\delta^2 + 4\zeta(\delta - \varepsilon) \geq 0$ then the roots of the quadratic function*

$$\lambda^2\zeta - \lambda(\delta + 2\zeta) + \varepsilon + \zeta = 0$$

satisfy

$$\lambda = 1 + \frac{\delta}{2\zeta} \pm \mu, \quad \mu \leq \sqrt{2}\max\left\{\frac{\delta}{2\zeta}, \sqrt{\frac{|\delta - \varepsilon|}{\zeta}}\right\}$$

Proof. The roots of the quadratic equation satisfies

$$\begin{aligned}\lambda &= \frac{\delta + 2\zeta \pm \sqrt{(\delta+2\varepsilon)^2 - 4(\varepsilon+\zeta)}}{2\zeta} \\ &= 1 + \frac{\delta}{2\zeta} \pm \frac{\sqrt{\delta^2 + 4\zeta(\delta-\varepsilon)}}{2\zeta} \\ &= 1 + \frac{\delta}{2\zeta} \pm \sqrt{\left(\frac{\delta}{2\zeta}\right)^2 + \frac{\delta-\varepsilon}{\zeta}}\end{aligned}$$

If $\frac{\delta-\varepsilon}{\zeta} \geq 0$, then

$$\begin{aligned}\sqrt{\left(\frac{\delta}{2\zeta}\right)^2 + \frac{\delta-\varepsilon}{\zeta}} &\leq \sqrt{2\max\left\{\left(\frac{\delta}{2\zeta}\right)^2, \frac{\delta-\varepsilon}{\zeta}\right\}} \\ &= \sqrt{2}\max\left\{\frac{\delta}{2\zeta}, \sqrt{\frac{\delta-\varepsilon}{\zeta}}\right\}.\end{aligned}$$

If $\frac{\delta-\varepsilon}{\zeta} \leq 0$, then the assumption $\delta^2 + 4\zeta(\varepsilon-\delta) \geq 0$ implies that

$$\left(\frac{\delta}{2\zeta}\right)^2 \geq \frac{\delta-\varepsilon}{\zeta} \geq 0.$$

Hence,

$$\begin{aligned}\sqrt{\left(\frac{\delta}{2\zeta}\right)^2 + \frac{\delta-\varepsilon}{\zeta}} &\leq \frac{\delta}{2\zeta} \\ &< \sqrt{2}\max\left\{\frac{\delta}{2\zeta}, \sqrt{\frac{\varepsilon-\delta}{\zeta}}\right\}.\end{aligned}$$

□

REMARK 4.4. *If $\zeta \gg \delta$ and $\zeta \gg \varepsilon$, then $\lambda \approx 1$ in Theorem 4.3.*

THEOREM 4.5. *Let $\mathcal{A}, \widehat{P} \in \mathbb{R}^{(n+m)\times(n+m)}$ and their sub-blocks be as defined in Theorem 3.1 (using the same notation and assumptions). We shall assume that A, B, and G remain fixed, but C may change so long as E also remains fixed. Further, assume that $A + \widetilde{D}$ and $G + 2\widetilde{D}$ are symmetric positive definite, $\widetilde{D}$ is symmetric positive semidefinite and*

$$\min\left\{(z^TGz)^2 + 4(z^T\widetilde{D}z)(z^TGz + z^T\widetilde{D}z - 1) : \|z\|_{A+\widetilde{D}} = 1\right\} > 0,$$

then all the eigenvalues of $\widetilde{P}^{-1}\mathcal{A}$ are real and positive.

The eigenvalues λ of (3.15) subject to $E^TBYx_y \neq 0$, will also satisfy

$$|\lambda - 1| \leq \mathcal{O}(\|C\|).$$

Proof. That the eigenvalues of $\widetilde{P}^{-1}\mathcal{A}$ are real and positive follows directly from Theorem 3.5.

Suppose that $C = EDE^T$ is a reduced singular value decomposition of C, where the columns of $E \in \mathbb{R}^{m\times p}$ are orthogonal and $D \in \mathbb{R}^{p\times p}$ is diagonal with entries d_j that are non-negative and in non-increasing order.

In the following, $\|.\| = \|.\|_2$, so that

$$\|C\| = \|D\| = d_1.$$

Premultiplying the quadratic eigenvalue problem (3.15) by w^T gives

$$\begin{aligned} 0 = \lambda^2 w^T \widetilde{D} w - \lambda(w^T G w + 2w^T \widetilde{D} w) \\ +(w^T A w + w^T \widetilde{D} w). \end{aligned} \tag{4.4}$$

Assume that $v = E^T Bw$ and $\|v\| = 1$, where w is an eigenvalue of the above quadratic eigenvector problem, then

$$\begin{aligned} w^T \widetilde{D} w &= v^T D^{-1} v \\ &= \frac{v_1^2}{d_1} + \frac{v_2^2}{d_2} + \ldots + \frac{v_m^2}{d_m} \\ &\geq \frac{v^T v}{d_1} \\ &= \frac{1}{\|C\|}. \end{aligned}$$

Hence,

$$\frac{1}{w^T \widetilde{D} w} \leq \|C\| .$$

Let $\zeta = w^T \widetilde{D} w$, $\delta = w^T G w$ and $\varepsilon = w^T A w$, then (4.4) becomes

$$\lambda^2 \zeta - \lambda(\delta + 2\zeta) + \varepsilon + \zeta = 0.$$

From Lemma 4.3, λ must satisfy

$$\lambda = 1 + \frac{\delta}{2\zeta} \pm \mu, \quad \mu \leq \sqrt{2} \max\left\{ \frac{\delta}{2\zeta}, \sqrt{\frac{|\delta - \varepsilon|}{\zeta}} \right\}.$$

Now $\delta \leq c\|G\|$, $\varepsilon \leq c\|G\|$, where c is an upper bound on $\|w\|$ and $w = [\ x_y^T \quad x_z^T\ ]^T$ are eigenvectors of (3.15) subject to $E^T BYx_y \neq 0$ and $\left\|E^T Bw\right\| = 1$. Hence, the eigenvalues of (3.15) subject to $E^T BYx_y \neq 0$ satisfy

$$|\lambda - 1| = \mathcal{O}(\|C\|).$$

□

This clustering of part of the spectrum of $\widetilde{P}^{-1}\mathcal{A}$ will often translate into a speeding up of the convergence of a selected Krylov subspace method, [1, Section 1.3].

4.2. Numerical Examples. We shall verify our theoretical results by considering some simple saddle point systems.

EXAMPLE 4.6 (C nonsingular). *Consider the matrices*

$$\mathcal{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}, \quad \widetilde{P} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

so that $m = p = 1$ and $n = 2$. The preconditioned matrix $\widetilde{P}^{-1}\mathcal{A}$ has eigenvalues at $\frac{1}{2}$, $2 - \sqrt{2}$ and $2 + \sqrt{2}$. The corresponding eigenvectors are $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$, $\begin{bmatrix} 1 & 0 & (\sqrt{2}-1) \end{bmatrix}^T$ and $\begin{bmatrix} 1 & 0 & -(\sqrt{2}+1) \end{bmatrix}^T$ respectively. The preconditioned system $\widetilde{P}^{-1}\mathcal{A}$ has all non-unit eigenvalues, but this does not go against Theorem 3.1 because $m - p = 0$. The generalized eigenvalue problem $Z^T A Z x_z = \lambda Z^T G Z x_z$, becomes

$$x_z = \lambda 2 x_z,$$

thus defining the eigenvalue $\frac{1}{2}$ of the preconditioned system $\widetilde{P}^{-1}\mathcal{A}$. With our choices of $\mathcal{A}$ and $\widetilde{P}$, and setting $D = I$ and $E = I$ ($C = EDE^T$), the quadratic eigenvalue problem (3.15) is

$$\left(\lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_y \\ x_z \end{bmatrix} = 0.$$

This quadratic eigenvalue problem has three finite eigenvalues, of which two correspond to the case $E^T BY x_y \neq 0$. These are $\lambda = 2 - \sqrt{2}$ and $\lambda = 2 + \sqrt{2}$; the corresponding eigenvectors have $x_z = 0$.

EXAMPLE 4.7 (C semidefinite). *Consider the matrices*

$$\mathcal{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \quad \widetilde{P} = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

so that $m = 2$, $n = 2$ and $p = 1$. The preconditioned matrix $\widetilde{P}^{-1}\mathcal{A}$ has two unit eigenvalues and a further two at $\lambda = 2 - \sqrt{2}$ and $\lambda = 2 + \sqrt{2}$. There is just one linearly independent eigenvector associated with the unit eigenvector; specifically this is $\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$. For the non-unit eigenvalues, the eigenvectors are $\begin{bmatrix} 0 & 1 & 0 & (\sqrt{2}-1) \end{bmatrix}^T$ and $\begin{bmatrix} 0 & 1 & 0 & -(\sqrt{2}+1) \end{bmatrix}^T$ respectively.

Since $2(m-p) = 2$, we correctly expected there to be at least two unit eigenvalues, Theorem 3.1. The same theorem and the fact that $n-m = 0$ implies that the remaining eigenvalues will be defined by the quadratic eigenvalue problem (3.15):

$$\left(\lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \right) \begin{bmatrix} x_{y1} \\ x_{y2} \end{bmatrix} = 0,$$

where $D = [1]$ and $E = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ are used as factors of C. This quadratic eigenvalue problem has three finite eigenvalues, of which two correspond to the case $E^T BY x_y \neq 0$, *i.e.* $x_{y2} \neq 0$. *These are* $\lambda = 2 - \sqrt{2}$ *and* $\lambda = 2 + \sqrt{2}$; *the corresponding eigenvectors have* $x_{y1} = 0$.

EXAMPLE 4.8 (C with small entries). *Suppose that* $\mathcal{A}$ *and* $\widetilde{P}$ *are as in Example 4.6, but* $C = [10^{-a}]$ *for some positive real number* a. *The generalized eigenvalue problem* $Z^T AZx_z = \lambda Z^T GZx_z$ *is unchanged, so one of the eigenvalues will take the value* $\frac{1}{2}$. *Setting* $D = 10^{-a}I$ *and* $E = I$ ($C = EDE^T$), *the quadratic eigenvalue problem (3.15) is*

$$\left(\lambda^2 \begin{bmatrix} 10^a & 0 \\ 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 2+2\times 10^a & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1+10^a & 0 \\ 0 & 1 \end{bmatrix}\right) \begin{bmatrix} x_y \\ x_z \end{bmatrix} = 0.$$

This quadratic eigenvalue problem has three finite eigenvalues, but just two of these have associated eigenvectors with $E^T BY x_y \neq 0$. *These two eigenvalues are defined by*

$$\lambda = 1 + 10^{-a} \pm 10^{-a}\sqrt{1+10^a}.$$

For large values of a, $\lambda \approx 1 + 10^{-a} \pm 10^{-\frac{a}{2}}$; *the eigenvalues will be close to 1.*

The CUTEr test set [11] provides a set of quadratic programming problems. We shall use the problem CVXQP2_S in the following examples. This problem is very small with $n = 100$ and $m = 25$. "Barrier" penalty terms (in this case 1.1) are added to the diagonal of A to simulate systems that might arise during and iteration of an interior-point method for such problems. We shall set $G = diag(A)$, and $C = \alpha \times diag(0, \ldots, 0, 1, \ldots, 1)$, where α is a positive, real parameter that we will change.

All tests were performed on a dual Intel Xeon 3.20GHz machine with hyper-threading and 2GiB of RAM. It was running Fedora Core 2 (Linux kernel 2.6.8) with MATLAB® 7.0. The linear systems were solved using the Simplified Quasi-Minimal Residual Algorithm (SQMR) [8] – MATLAB® code for SQMR can be obtained from the MATLAB® Central File Exchange at `http://www.mathworks.fr/matlabcentral/`. We terminate the iteration when the value of residual is reduced by at least a factor of 10^{-8}.

In Figure 4.1 we compare the performance (in terms of iteration count) between using a preconditioner of the form $\widetilde{\mathcal{P}}$ and one of the form $\mathcal{P}$, Equations (1.4) and (1.2) respectively. The matrix C used in this set of results takes the form αI. As α becomes smaller, we shall expect the difference between the number of iterations required to become less between the two preconditioners. We observe that, in this example, once $\alpha \leq 10^{-3}$ there is little benefit in reproducing C in the preconditioner.

In Figure 4.2 we also compare the performance (in terms of iteration count) between using a preconditioner of the form $\widetilde{\mathcal{P}}$ and one of the form $\mathcal{P}$, Equations (1.4) and (1.2) respectively. However, we have now set $C = \alpha \times diag(0, \ldots, 0, 1, \ldots, 1)$, where $\operatorname{rank} C = \lfloor m/2 \rfloor$. We observe that the convergence is faster in the second figure - this is as we would expect because of there now being a guarantee of at least 24 unit eigenvalues in the preconditioned system compared to the possibility of none.

5. Conclusion. In this paper, we have investigated a new class of preconditioner for indefinite linear systems that incorporate the (1,2) and (2,1) blocks of the original matrix. These blocks are often associated with constraints. We have shown that if C has rank $p > 0$, then the preconditioned system has at least $2(m-p)$ unit eigenvalues, regardless of the structure of G. In addition, we have shown that if the entries of C are very small, then we will expect an additional $2p$ eigenvalues to be clustered around

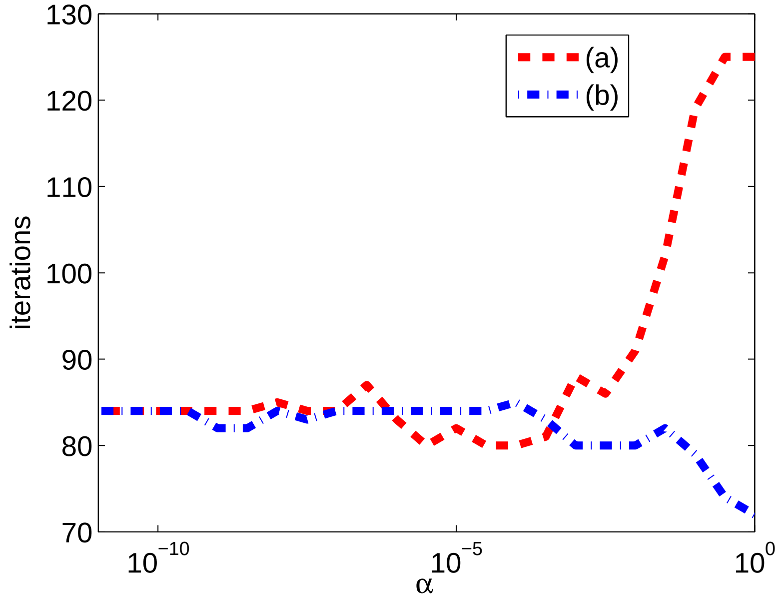

FIG. 4.1. *Number of SQMR iterations when either* (a) $\widetilde{\mathcal{P}}$ *or* (b) $\mathcal{P}$ *are used as preconditioners for* $C = \alpha I$.

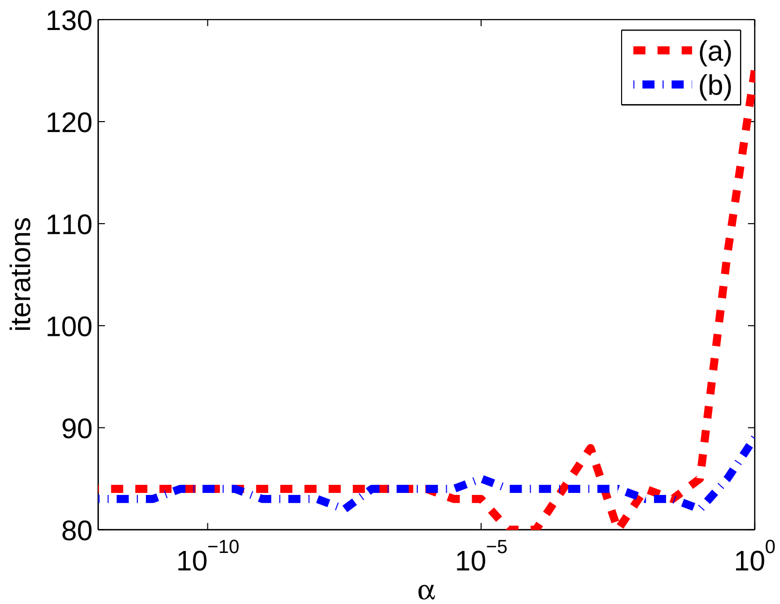

FIG. 4.2. *Number of SQMR iterations when either* (a) $\widetilde{\mathcal{P}}$ *or* (b) $\mathcal{P}$ *are used as preconditioners for* $C = \alpha \times diag(0, \ldots, 0, 1, \ldots, 1)$, *where* $\operatorname{rank} C = \lfloor m/2 \rfloor$.

1 and, hence, for the number of iterations required by our chosen Krylov subspace method to be dramatically reduced.

REFERENCES

[1] O. Axelsson and V. A. Barker, *Finite element solution of boundary value problems*, vol. 35 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Theory and computation, Reprint of the 1984 original.

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

[3] L. Bergamaschi, J. Gondzio, and G. Zilli, *Preconditioning indefinite systems in interior point methods for optimization*, Comput. Optim. Appl., 28 (2004), pp. 149–171.

[4] H. S. Dollar, *Extending constraint preconditioners for saddle point problems*, Tech. Report NA-05/02, Oxford University Computing Laboratory, 2005. (Submitted to SIAM J. Matrix Anal. Appl.).

[5] H. S. Dollar, N. I. M. Gould, and A. J. Wathen, *On implicit-factorization constraint preconditioners*, Tech. Report RAL-TR-2004-036, Rutherford Appleton Laboratory, 2004. (To appear in *Large Scale Nonlinear Optimization*, G. Di Pillo and M. Roma, eds., Springer Verlag).

[6] H. S. Dollar and A. J Wathen, *Incomplete factorization constraint preconditioners for saddle-point problems*, Tech. Report 04-01, Oxford University Computing Laboratory, Oxford, England, 2004. (To appear in SIAM J. Sci. Comput.).

[7] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: with applications in Incompressible Fluid Dynamics*, Oxford University Press, Oxford, 2005.

[8] R. W. Freund and N. M. Nachtigal, *A new Krylov-subspace method for symmetric indefinite linear systems*, in Proceedings of the 14th IMACS World Congress on Computational and Applied Mathematics, W. F. Ames, ed., IMACS, 1994, pp. 1253–1256.

[9] N. I. M. Gould, *On the accurate determination of search directions for simple differentiable penalty functions*, IMA Journal of Numerical Analysis, 6 (1986), pp. 357–372.

[10] N. I. M. Gould, M. E. Hribar, and J. Nocedal, *On the solution of equality constrained quadratic programming problems arising in optimization*, SIAM J. Sci. Comput., 23 (2001), pp. 1376–1395.

[11] N. I. M. Gould, D. Orban, and P. L. Toint, *CUTEr (and SifDec), a constrained and unconstrained testing environment, revisited*, ACM Transactions on Mathematical Software, 29 (2003).

[12] A. Greenbaum, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[13] 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.

[14] Y. Saad, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2003.

[15] 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.

[16] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.

[17] S. J. Wright, *Primal-dual interior-point methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.