SIAM J. MATRIX ANAL. APPL.
Vol. 27, No. 3, pp. 779–792

AN ALGEBRAIC ANALYSIS OF A BLOCK DIAGONAL PRECONDITIONER FOR SADDLE POINT SYSTEMS*

GENE H. GOLUB†, CHEN GREIF‡, AND JAMES M. VARAH‡

Dedicated to Michael Saunders on the occasion of his 60th birthday

Abstract. We consider a positive definite block preconditioner for solving saddle point linear systems. An approach based on augmenting the (1,1) block while keeping its condition number small is described, and algebraic analysis is performed. Ways of selecting the parameters involved are discussed, and analytical and numerical observations are given.

Key words. saddle point linear systems, augmented Lagrangian, block preconditioning, positive semidefinite matrix, null space

AMS subject classifications. 65F10, 15A12

DOI. 10.1137/04060679X

1. Introduction.

Consider the large and sparse saddle point linear system

$$\mathcal{H}u = b \equiv \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}, \tag{1.1}$$

where A is $n \times n$ and B is $n \times m$, with $m \leq n$ (possibly $m \ll n$). Systems of the form (1.1) arise in many areas; for a recent comprehensive survey of numerical solution methods see [3]. The numerical treatment of these systems strongly depends on the properties of the (1,1) block, A, and we assume throughout (unless otherwise stated) that it is positive semidefinite. We also assume that $\mathcal{H}$ is nonsingular, which implies that B has full column rank.

It is possible to replace (1.1) by

$$\begin{pmatrix} A + \gamma BB^T & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c + \gamma Bd \\ d \end{pmatrix}, \tag{1.2}$$

where γ is a positive scalar. Such an approach has been considered in various applications (see, e.g., [6]). In the context of optimization, it is related to forming an augmented Lagrangian function [14]; for example, the method of multipliers [2] amounts to applying the Uzawa algorithm [1] to (1.2).

Equation (1.1) can also be reformulated as a linear system that is more general than (1.2), as follows (see, for example, [5]):

$$\begin{pmatrix} A + BWB^T & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c + BWd \\ d \end{pmatrix}, \tag{1.3}$$

*Received by the editors April 15, 2004; accepted for publication (in revised form) by A. Wathen March 24, 2005; published electronically January 27, 2006.
http://www.siam.org/journals/simax/27-3/60679.html

†SCCM Program, Gates 2B, Stanford University, Stanford, CA 94305-9025 (golub@sccm.stanford.edu). The work of this author was supported in part by the US Department of Energy.

‡Department of Computer Science, The University of British Columbia, Vancouver BC V6T-1Z4, Canada (greif@cs.ubc.ca, varah@cs.ubc.ca). The work of these authors was supported in part by the Natural Sciences and Engineering Research Council of Canada.

where W is an $m \times m$ matrix. We will assume that W is symmetric positive semidefinite yet $A + BWB^T$ is positive definite. Let us define

$$M(W) = A + BWB^T; \qquad \mathcal{H}(W) = \begin{pmatrix} M(W) & B \\ B^T & 0 \end{pmatrix}. \tag{1.4}$$

Note that $\mathcal{H}$ defined in (1.1) is equivalent to $\mathcal{H}(0)$ in (1.4).

An advantage of forming (1.2) or the more general form (1.3) is that the Schur complement is well defined whereas in (1.1) there is no Schur complement associated with a singular A. Thus, one could solve the system by using the decomposition

$$\begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^T M^{-1} & I \end{pmatrix} \begin{pmatrix} M & 0 \\ 0 & -B^T M^{-1} B \end{pmatrix} \begin{pmatrix} I & M^{-1}B \\ 0 & I \end{pmatrix}, \tag{1.5}$$

which holds for any symmetric nonsingular matrix M, including $M = M(W)$ defined in (1.4). Using (1.5), solution methods for positive definite systems can be applied throughout the iterative process of solving for the indefinite matrix $\mathcal{H}(W)$. Other possible numerical advantages of solving (1.2) or (1.3) instead of (1.1) are discussed in [7] and in other places. However, choosing W to make $M(W)$ easy to solve for may hurt the conditioning of $\mathcal{H}(W)$ and the choice of this parameter is delicate. For example, for $W = \gamma I$, as γ grows the conditioning of the saddle point matrix of (1.2) typically deteriorates at a rate proportional to γ^2 whereas the condition number of its associated (1,1) block deteriorates at a rate typically proportional to γ. It may be useful, then, to consider an approach of using a block preconditioner based on $M(W)$ rather than (or in addition to) modifying system (1.1) to (1.2) or (1.3).

Motivated by the above, our goal in this paper is to perform an algebraic study of the block diagonal positive definite preconditioner

$$\mathcal{M}(W) = \begin{pmatrix} A + BWB^T & 0 \\ 0 & B^T(A + BWB^T)^{-1}B \end{pmatrix}. \tag{1.6}$$

Forming the preconditioner (1.6) may be computationally expensive (in particular setting up the Schur complement and solving for it) unless A and B have a very special structure, and in practice cheaper alternatives must be sought. Nevertheless, understanding the spectral properties of (1.6) is useful since it can illuminate the behavior of preconditioners based on approximations of the components of $\mathcal{M}(W)$. For example, finite volume discretization of the Maxwell equation yields a semidefinite (1,1) block corresponding to the curl-curl operator, and in certain applications this operator is replaced by an operator of the form $A + BDB^T$ with D a diagonal scaling matrix [11]. This is not necessarily done in a saddle point system context, but such a formulation is also viable.

Our focus throughout this paper is on the analysis, and we stay within the algebraic framework without associating the problem with any specific application. In section 2 we provide analytical results that characterize the spectral distribution of the preconditioned matrix and indicate what can be expected in terms of convergence of minimum residual Krylov solvers such as MINRES [16]. Section 3 is devoted to the question of how to choose the parameter(s) involved, and we provide a perturbation analysis which may help in the choice of W. The idea here is that aiming to minimize the condition number of the augmented (1,1) block (and in practice look for cheap computational alternatives that keep it small) may prove effective. A few numerical examples are provided in section 4, and in section 5 we draw some conclusions.

2. The block diagonal preconditioner. In [13] it is shown that if A is nonsingular, then a matrix that in our notation here is equal to $\mathcal{M}(0)$ has the attractive property that the associated preconditioned matrix $\mathcal{M}^{-1}(0)\mathcal{H}$ has at most four distinct eigenvalues: $0, 1, \frac{1}{2} \pm \frac{\sqrt{5}}{2}$. Thus, when this matrix is nonsingular, a minimum residual Krylov solver terminates within three iterations (if roundoff errors are ignored). In the context of the linear system we discuss, if (1.1) is first reformulated as (1.3) and then $\mathcal{M}(W)$ is applied as a preconditioner, then we again obtain a preconditioned matrix with three distinct nonzero eigenvalues. However, it is possible that a particular choice of W that works well in solving for $\mathcal{M}(W)$ will not be the ideal choice in terms of the conditioning or the spectral distribution of $\mathcal{H}(W)$. Using $\mathcal{M}(W)$ as a preconditioner applied to (1.1) rather than (1.3) enhances flexibility in that it allows for a choice of W that makes the solve for $\mathcal{M}(W)$ effective without directly having to balance it with considerations related to the conditioning of the saddle point matrix.

Let us begin this part of the analysis by showing that if the matrix of (1.3) were to be used as a *preconditioner* for system (1.1), then a minimal residual Krylov solver converges almost immediately. In [7] it is shown that for general matrices A of size $n \times n$, B and C of size $n \times m$ with full column rank ($m \le n$), and W of size $m \times m$, if $\tilde{\mathcal{H}}(W) = \left(\begin{smallmatrix} A+BWC^T & B \\ C^T & 0 \end{smallmatrix}\right)$ is nonsingular, then

$$\tilde{\mathcal{H}}^{-1}(W) = \tilde{\mathcal{H}}^{-1}(0) - \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}.$$

The same result, for symmetric matrices, can be found in [5, Chap. 12]. From this we can obtain the following result.

PROPOSITION 2.1. *If $\mathcal{H}(W)$ (defined in* (1.4)*) is used as a preconditioner for solving the linear system* (1.1)*, and a minimum residual Krylov subspace solver is applied, convergence is obtained (in the absence of roundoff errors) within two iterations.*

Proof. Since

$$\mathcal{H}^{-1}(W)\mathcal{H} = \left\{\mathcal{H}^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}\right\}\mathcal{H} = \begin{pmatrix} I & 0 \\ WB^T & I \end{pmatrix},$$

it follows that all eigenvalues are equal to 1 and the minimal polynomial is $p_2(z) = (1-z)^2$. ∎

From Proposition 2.1 it follows that a constraint preconditioner (see, e.g., [12] or [17] for definition and spectral analysis) based on taking $M(W)$ as its (1,1) block behaves like a direct solver. Clearly, solving for $A + BWB^T$ may be a nontrivial computational task, and so the effectiveness of this approach depends on the numerical properties of the preconditioning matrix. The preconditioner $\mathcal{M}(W)$ defined in (1.6) is the middle matrix in the decomposition (1.5) applied to $\mathcal{H}(W)$. If it is to be used, then MINRES (which requires a positive definite preconditioner) could be applied.

2.1. Spectral properties of the preconditioned matrix. Consider the matrix $\mathcal{M} = \left(\begin{smallmatrix} M & 0 \\ 0 & B^T M^{-1} B \end{smallmatrix}\right)$, where M is a positive definite matrix. Let ν be an eigenvalue of $\mathcal{M}^{-1}\mathcal{H}$ with eigenvector $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$. Then

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \nu \begin{pmatrix} M & 0 \\ 0 & B^T M^{-1} B \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Substituting $v = \frac{1}{\nu}(B^T M^{-1} B)^{-1} B^T u$, we get a quadratic expression (in ν) for u:

$$\left[\nu^2 M - \nu A - B(B^T M^{-1} B)^{-1} B^T\right] u = 0.$$

Denoting $\tilde{u} = M^{\frac{1}{2}} u$ and $C = M^{-1/2} B$, the equation can be written as

$$(\nu^2 I - \nu K - P)\tilde{u} = 0, \tag{2.1}$$

where

$$K = M^{-1/2} A M^{-1/2}; \qquad P = C(C^T C)^{-1} C^T.$$

Note that if A is positive definite and $M = A$, then $K = I$ and the result of [13] applies. P is an orthogonal projector of rank m, thus it has m eigenvalues equal to 1 and $n - m$ eigenvalues equal to 0. The matrix K has eigenvalues equal to those of the matrix $M^{-1}A$.

So far we have not relied on specific properties, and (2.1) holds for any positive definite M. We now discuss the specific choice $M(W)$, with W positive semidefinite. We start with $W = \gamma I$, for which we can show that there is an explicit formula that connects P and K.

PROPOSITION 2.2. *Suppose $W = \gamma I$. The projection matrix P is a polynomial of degree m in K, given explicitly by $P = I - p(K)$, where p is the Lagrange interpolant*

$$p(x) = \frac{\displaystyle\prod_{i=1}^{m} \left(x - \frac{\lambda_i}{\lambda_i + \gamma}\right)}{\displaystyle\prod_{i=1}^{m} \left(1 - \frac{\lambda_i}{\lambda_i + \gamma}\right)} \tag{2.2}$$

and $\{\lambda_i\}$ are the eigenvalues of the generalized eigenvalue problem $Ax = \lambda BB^T x$.

Proof. Let $L = \frac{1}{\gamma}(I - K)$. Then $L = CC^T = C(I)C^T$, and $L^p = C(C^T C)^{p-1} C^T$ for any positive integer p. Since $P = C(C^T C)^{-1} C^T$, and since by the Cayley–Hamilton theorem $(C^T C)^{-1}$ is a polynomial of degree m in $C^T C$, it follows that P is a polynomial of degree m in L and hence a polynomial of degree m in K.

The polynomial connection is now established, and we can seek an explicit expression. Denote the eigenvalues of K by $\{\mu_i\}$. We have $eig(K) = eig(M^{-1}A) = eig(A - \mu M)$. Thus $\{\mu_i\} = eig(A - \mu(A + \gamma BB^T)) = eig((1 - \mu)A - \mu\gamma BB^T)$, and the eigenvalues of K are related to the eigenvalues of the generalized eigenvalue problem $(A - \lambda BB^T)x = 0$ as follows: $\lambda_i = \frac{\mu_i \gamma}{1 - \mu_i}$. Since $\mathrm{rank}(BB^T) = m$, the $n - m$ infinite generalized eigenvalues λ_i map onto $\mu_i = 1$. There are m finite generalized eigenvalues λ_i; the m eigenvalues of K that are not equal to 1 thus satisfy $0 \le \mu_i < 1$.

As for the matrix P, since $n - m$ of its eigenvalues are equal to 1, we know that the function p defined in the statement of this theorem satisfies $p(1) = 1$ and $p(\frac{\lambda_i}{\lambda_i + \gamma}) = 0$, and thus p is the Lagrange interpolant of degree m given by (2.2). ∎

If $W \neq \gamma I$, a polynomial connection does not necessarily exist (in particular if W is singular), but we can still make an observation on the clustering of the eigenvalues.

LEMMA 2.3. *For $M = M(W)$, P and K share the same eigenvectors.*

Proof. We have

$$I - CWC^T = M^{-1/2}(M - BWB^T)M^{-1/2} = M^{-1/2} A M^{-1/2} = K,$$

from which it follows that

$$PK = C(C^TC)^{-1}C^T(I - CWC^T) = KP = P + K - I.$$

Since K and P are commuting normal matrices, they share the same eigenvectors. ∎

THEOREM 2.4. *Suppose A and W are symmetric positive semidefinite and $M(W)$ is symmetric positive definite. Then the eigenvalues of K, $\{\mu_i\}$, are given as follows.*

1. *$n-m$ eigenvalues are equal to 1; the corresponding eigenvectors span the null space of B^T.*
2. *nullity(A) eigenvalues are equal to 0; the corresponding eigenvectors span the null space of A.*
3. *The rest of the eigenvalues are all strictly between 0 and 1, and are given explicitly in terms of the generalized eigenvalue problem $Ax = \lambda BWB^Tx$ as follows:*

$$\mu_i = \frac{\lambda_i}{\lambda_i + 1}, \qquad i = 1, \ldots, 2m - \text{nullity}(A).$$

Proof. Suppose $Kx = \mu x$. Then μ is also an eigenvalue of $M^{-1}A$, with eigenvector $\tilde{x} = M^{-1/2}x$, and we have

$$A\tilde{x} = \mu(A + BWB^T)\tilde{x}. \tag{2.3}$$

Proof of statement 1. If z is a vector in the null space of B^T, namely $B^Tz = 0$, then from nonsingularity of $\mathcal{H}$ it follows that $Az \neq 0$. Thus z is an eigenvector of (2.3), with eigenvalue $\mu = 1$. Since B has full column rank, this eigenvalue is of multiplicity $n - m$.

Proof of statement 2. Suppose A is semidefinite, and let $z \neq 0$ be a null vector. Since $A+BWB^T$ is positive definite, it follows that $BWB^Tz \neq 0$. From this it follows that $\mu = 0$ is a corresponding eigenvalue with multiplicity equal to nullity(A).

Proof of statement 3. If $\mu \neq 0, 1$ we have, by (2.3), $A\tilde{x} = \frac{\mu}{1-\mu}BWB^T\tilde{x}$. Since A and BWB^T are positive semidefinite, the generalized eigenvalues $\lambda = \frac{\mu}{1-\mu}$ must be nonnegative and hence $0 < \mu = \frac{\lambda}{\lambda+1} < 1$. ∎

THEOREM 2.5. *If A is positive semidefinite, the eigenvalues of the preconditioned matrix $\mathcal{M}^{-1}(W)\mathcal{H}$ are bounded within the two intervals*

$$\left[-1, \frac{1-\sqrt{5}}{2}\right] \bigcup \left[1, \frac{1+\sqrt{5}}{2}\right].$$

The eigenvalue 1 is of algebraic multiplicity $n - m +$ nullity(A). Also, nullity(A) *eigenvalues are equal to -1.*

Proof. Since null(B^T) is the space of eigenvectors associated with the $n - m$ zero eigenvalues of P and the $n-m$ eigenvalues of K that are equal to 1, and since K and P commute by Lemma 2.3, we get that $n - m$ eigenvalues of the preconditioned matrix $\mathcal{M}^{-1}(W)\mathcal{H}$ satisfy $\nu^2 - \nu = 0$, and when the preconditioned matrix is nonsingular those eigenvalues are all equal to 1.

For the rest of the eigenvalues, by (2.1) we have the connection $\nu^2 - \nu\mu = 1$, which leads to

$$\nu_i = \frac{1}{2}\left(\mu_i \pm \sqrt{\mu_i^2 + 4}\right).$$

The zero eigenvalues of A are mapped onto $\nu = \pm 1$. On the other hand, from what we know about the eigenvalues of K (namely, that they are all bounded between 0 and 1; see Theorem 2.4) we can deduce that

$$1 \le \frac{1}{2}\left(\mu_i + \sqrt{\mu_i^2 + 4}\right) \le \frac{1}{2}(1 + \sqrt{5}) \approx 1.618$$

and

$$-0.618 \approx \frac{1}{2}(1 - \sqrt{5}) \ge \frac{1}{2}\left(\mu_i - \sqrt{\mu_i^2 + 4}\right) \ge -1. \qquad \square$$

All the eigenvalues of the preconditioned matrix are thus within two narrow intervals whose ends are fixed. We note that some of the bounds presented above can be obtained by applying [18, Lemma 2.1] or [17, Proposition 2] to the matrix $\mathcal{M}^{-1/2}\mathcal{H}\mathcal{M}^{-1/2}$; however at least one of the bounds obtained here is tighter, since we use the special connection between K and P, observed in Lemma 2.3, which does not hold in more general cases.

2.2. Numerical properties of the Schur complement. The preconditioner we are studying involves forming the Schur complement

$$-S(W) = -B^T(A + BWB^T)^{-1}B.$$

The same Schur complement arises if (1.3) is to be solved using a block Gaussian elimination procedure. An analysis of the Schur complement is useful for evaluating the merits of the general approach, even if in practice one may seek cheaper alternatives.

When A is positive definite $S(0) = B^T A^{-1} B$ is defined; for simplicity of notation let us denote it simply by S. If A is semidefinite, consider a small perturbation of A, e.g., $A(\varepsilon) \equiv A + \varepsilon BB^T$ for some small ε, so that the Schur complement is still defined, and in this case we refer to $A(\varepsilon)$ and $S(\varepsilon)$ simply as A and S, and both are symmetric positive definite.

Using the Sherman–Morrison formula [8, p. 50] in its block version, we have

$$\begin{aligned} S(W) &= B^T(A + BWB^T)^{-1}B \\ &= B^TA^{-1}B - B^TA^{-1}B(I + WB^TA^{-1}B)^{-1}WB^TA^{-1}B \\ &= B^TA^{-1}B(I + WB^TA^{-1}B)^{-1}(I + WB^TA^{-1}B - WB^TA^{-1}B) \\ &= S(I + WS)^{-1}, \end{aligned}$$

and we have the following commuting relation.

PROPOSITION 2.6. *If S and W commute, then S and $S(W)$ commute.*

Proof. If $WS = SW$, then we have

$$\begin{aligned} S \cdot S(W) &= S^2(I + WS)^{-1} = S(S^{-1} + WSS^{-1})^{-1} \\ &= S(S^{-1} + S^{-1}SW)^{-1} = S(I + WS)^{-1}S = S(W) \cdot S. \qquad \square \end{aligned}$$

This leads to the following conclusion with regard to the conditioning of the Schur complement.

PROPOSITION 2.7. *Suppose that S and W commute, and let $\{\varphi_i\}$ be the eigenvalues of $S(W)$, $\{\beta_i\}$ the eigenvalues of S, and $\{\alpha_i\}$ the eigenvalues of W. Then*

$$\varphi_i = \frac{\beta_i}{\beta_i\alpha_i + 1}. \tag{2.4}$$

Proof. From Proposition 2.6 it follows that S, $S(W)$, and W share the same eigenvectors. Ordering the eigenvalues so that the ith eigenvalue of each of the matrices corresponds to the same eigenvector, (2.4) readily follows. ∎

COROLLARY 2.8. *Let $S(\gamma)$ denote $S(W)$ with $W = \gamma I$. Then $\kappa_2(S(\gamma)) \leq \kappa_2(S)$ for any $\gamma > 0$.*

Proof. By Proposition 2.7 we have $S(\gamma) = S(I + \gamma S)^{-1}$ and the eigenvalues of $S(\gamma)$ are $\varphi_i = \frac{\beta_i}{1+\gamma\beta_i}$. The desired result readily follows, by monotonicity. ∎

Proposition 2.7 has limited practical value, since S and W rarely commute (even if W is a simple nonconstant diagonal matrix there is no commuting). However, even when the commuting condition does not hold, it may still nearly hold.

3. Minimizing the condition number of the augmented (1,1) block. Suppose that

$$W = \operatorname{diag}(w_j), \qquad w_j \geq 0.$$

The (1,1) block can be expressed as a sum of rank-1 perturbations of A as follows:

$$M(W) = A + \sum_{j \in \mathcal{S}} w_j b_j b_j^T,$$

where b_j is the jth column of B, and $\mathcal{S}$ is the set of indices j for which $w_j \neq 0$. The number of nonzero entries of W is equal to at least the nullity of A if positive definiteness of the modified (1,1) block is to be obtained.

A possible way of selecting W is to base it on condition number minimization of $A+BWB^T$. See [15] and the references therein for a survey of eigenvalue optimization problems; see also [10] for minimizing the condition number given a prescribed sparsity pattern.

Since $M(W)$ defined above is a sum of A and a sequence of rank-1 contributions, we can simplify the discussion by considering a single rank-1 perturbation:

$$M(w) = A + wbb^T,$$

where A is positive definite and $w \geq 0$.

PROPOSITION 3.1. *Suppose that A is symmetric positive semidefinite with nullity 1, and suppose that $(0, u_n)$ and (λ_1, u_1) are the eigenpairs with the smallest and the largest eigenvalues of A (known to be simple), and $\lambda_{n-1} > 0$ is the smallest positive eigenvalue of A. Then the function $\kappa_2(A + wu_nu_n^T)$ is strictly decreasing in w for $0 = \lambda_n < w < \lambda_{n-1}$, is constant for $\lambda_{n-1} \leq w \leq \lambda_1$, is strictly increasing for $w > \lambda_1$, and is explicitly given in terms of the eigenvalues of A as follows:*

$$\kappa_2(A + wu_nu_n^T) = \frac{\max(\lambda_1, w)}{\min(\lambda_{n-1}, w)}. \tag{3.1}$$

The proof follows straightforwardly from orthogonality of the eigenvectors and is omitted. In [4, Theorem 5.3] it is shown that for saddle point systems rank-1 perturbations of a (1,1) block with nullity 1 change the zero eigenvalue to a positive one and perturb the eigenvectors, and the norm of the perturbation is related to the angle formed between the null vector of the matrix and the vector used for the rank-1 perturbation. The angle between a column of B and the null vectors of A is relevant in our discussion here, in which the condition number is considered.

Let the eigenvalues of A be $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n > 0$, and let us define

$$\kappa(w) \equiv \kappa_2(A + wbb^T); \qquad \lambda_i(w) \equiv \lambda_i(A + wbb^T).$$

Also, we denote by $u_i(w)$ the eigenvectors of $A + wbb^T$. The eigenvalues of $M(w)$ and A interlace:

$$\lambda_n \leq \lambda_n(w) \leq \lambda_{n-1} \leq \cdots \leq \lambda_1 \leq \lambda_1(w).$$

In general, each $\lambda_k(w)$ is a piecewise smooth function, with "corners" possible at multiple eigenvalues. In order to avoid discontinuities with the extremal eigenvalues λ_1 and λ_n, we assumed above that they were *simple* eigenvalues of A; then $\lambda_1(w)$, $\lambda_n(w)$, and hence $\kappa(w) = \lambda_1(w)/\lambda_n(w)$ are all smooth functions. Note also that only $\lambda_1(w)$ increases without limit as $w \to \infty$; the other eigenvalues of $A + wbb^T$ are all bounded by the next eigenvalue of A. For $A = D = \mathrm{diag}(\lambda_k)$, the limiting values are the roots of $\sum_{k=1}^{n} \frac{b_k^2}{(\lambda - \lambda_k)} = 0$ [8, section 8.5.3].

We are interested in the behavior of $\kappa(w)$ for $w > 0$, in particular whether it has a minimum. To this end, we need expressions for the derivatives of the eigenvalues, which can be derived from perturbation theory.

PROPOSITION 3.2. *Let $\lambda_k(w)$ be a simple eigenvalue of $M(w)$. Then*

$$\lambda_k'(w) = [b^T u_k(w)]^2,$$

where $u_k(w)$ is the corresponding (normalized) eigenvector.

Proof. A simple perturbation expansion [19] for $\lambda_k(w)$ about w gives this immediately, since our perturbation matrix is bb^T. ∎

Applying Proposition 3.2 to $\lambda_1(w)$ and $\lambda_n(w)$ gives

$$\frac{\kappa'(w)}{\kappa(w)} = \frac{\lambda_1'(w)}{\lambda_1(w)} - \frac{\lambda_n'(w)}{\lambda_n(w)} = \frac{\left(b^T u_1(w)\right)^2}{\lambda_1(w)} - \frac{\left(b^T u_n(w)\right)^2}{\lambda_n(w)}. \tag{3.2}$$

In general, $\kappa(w)$ does not necessarily have a minimum for $w > 0$. However, we can state when it does. Below we will again (as we did in section 2.2) assume that A is positive definite, and when it is not we will consider a small perturbation $A(\varepsilon)$ that is positive definite as the matrix for which the following result is stated.

PROPOSITION 3.3. *When λ_1 and λ_n are simple, a necessary and sufficient condition for $\kappa(A + wbb^T)$ to have at least one minimum for $w > 0$ is*

$$\frac{|b^T u_1|}{|b^T u_n|} < \sqrt{\kappa_2(0)}.$$

Proof. Since $\kappa(w) \to \infty$ as $w \to \infty$, it will have at least one minimum if and only if $\kappa'(0) < 0$. From (3.2), this holds if and only if $|b^T u_1(0)|^2 < \kappa(0)|b^T u_n(0)|^2$. ∎

From Proposition 3.3 it follows that it may be desirable that b form a small angle with the eigenvector of A that corresponds to the smallest eigenvalue (a null vector if A is rank deficient). This is a reasonable condition, which can be concluded also from the analysis in [4, section 5]. We can in fact show more, as follows.

THEOREM 3.4. *When λ_1 and λ_n are simple, $\kappa(A + wbb^T)$ has at most one minimum for $w > 0$.*

Proof. We shall show that $\kappa(w)$ is convex: $\kappa''(w) > 0$. To see this, consider (3.2). Multiplying by $\kappa(w)$ on both sides and differentiating with respect to w yield

$$\kappa'' = \kappa' \cdot \left(\frac{\lambda_1'}{\lambda_1} - \frac{\lambda_n'}{\lambda_n}\right) + \kappa \cdot \left(\frac{\lambda_1\lambda_1'' - (\lambda_1')^2}{\lambda_1^2} - \frac{\lambda_n\lambda_n'' - (\lambda_n')^2}{\lambda_n^2}\right).$$

Simplifying, we get

$$\frac{\kappa''}{\kappa} = \frac{\lambda_1''}{\lambda_1} - \frac{\lambda_n''}{\lambda_n} + 2\frac{\lambda_n'}{\lambda_n}\left(\frac{\lambda_n'}{\lambda_n} - \frac{\lambda_1'}{\lambda_1}\right).$$

From [19, p. 20] we can get an expression for the second derivative. For a general perturbation $A + wB$,

$$\lambda_k'' = 2\sum_{i\neq k}\frac{(u_i^T B u_k)^2}{\lambda_k - \lambda_i}.$$

Here we have $B = bb^T$, so

$$\lambda_1'' = 2(b^T u_1)^2 \sum_{i=2}^{n}\frac{(b^T u_i)^2}{\lambda_1 - \lambda_i}, \qquad \lambda_n'' = 2(b^T u_n)^2 \sum_{i=1}^{n-1}\frac{(b^T u_i)^2}{\lambda_n - \lambda_i}.$$

Now define $c_k = b^T u_k$. Then

$$\frac{\kappa''}{2\kappa} = \frac{c_1^2}{\lambda_1}\sum_{i=2}^{n}\frac{c_i^2}{\lambda_1 - \lambda_i} - \frac{c_n^2}{\lambda_n}\sum_{i=1}^{n-1}\frac{c_i^2}{\lambda_n - \lambda_i} + \frac{c_n^2}{\lambda_n}\left(\frac{c_n^2}{\lambda_n} - \frac{c_1^2}{\lambda_i}\right).$$

Notice that the terms in the first sum are positive and in the second sum are negative. Collecting terms, we have

$$\frac{\kappa''}{2\kappa} = \frac{c_1^2}{\lambda_1}\sum_{i=2}^{n}\frac{c_i^2}{\lambda_1 - \lambda_i} + \frac{c_n^2}{\lambda_n}\left[c_1^2\left(\frac{1}{\lambda_1 - \lambda_n} - \frac{1}{\lambda_1}\right) + \frac{c_n^2}{\lambda_n} + \sum_{i=2}^{n-1}\frac{c_i^2}{\lambda_i - \lambda_n}\right],$$

and all terms of the right are positive. □

As stated in Proposition 3.1, for $A + wu_n u_n^T$ any choice of the parameter w between the minimal positive eigenvalue and the maximal eigenvalue of A will obtain the minimized condition number. The upper bound of that range corresponds to taking $w = \|A\|_2 = \|A\|_2/\|u_n\|_2^2$. Given b, if it forms a small angle with the null vector of A we could make the choice $w = \|A\|_2/\|b\|_2^2$. This choice is relatively easy to compute or estimate. For the choice $A + wBB^T$ we can apply a similar strategy and set $w = \|A\|_2/\|B\|_2^2$, and if 2-norms are difficult to calculate, then norm estimators could be used. This establishes a computationally inexpensive way of choosing the parameter w in the general case.

Finally, we mention that other choices of W are possible. For example, setting $W = (B^T B)^{-1}$ (see, for example, [5]) has the advantage that the matrix $B(B^T B)^{-1}B^T$ is an orthogonal projector onto the range of B, which is orthogonal to the null space of B^T, and since the null space of A does not intersect with the null space of B^T either, such a choice of W is viable.

4. Numerical examples. We have tested with a set of matrices, including a few from the CUTE/CUTEr collection [9], with self-generated right-hand-side vectors. The matrices were used merely for testing purposes, without considering the type of constraints in the underlying optimization problem.

Example 4.1. This example validates some of the analytical observations made in section 3. The `genhs28` problem from the CUTE collection is a tiny 28×28 saddle point matrix, where A is 10×10. A is tridiagonal, almost Toeplitz, with

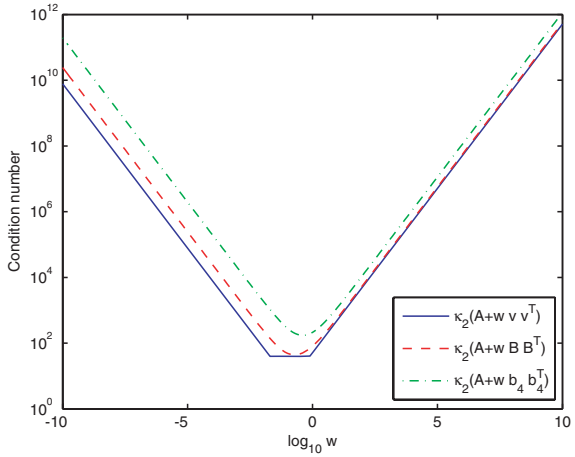

FIG. 4.1. *Example* 4.1*: Condition numbers of* $A + wvv^T$*,* $A + wBB^T$*, and* $A + wb_4b_4^T$ *for the* `genhs28` *matrix. Here* v *is the null vector of* A *and* b_4 *is the fourth column of* B*.*

$2, 4, 2$ along its superdiagonal, main diagonal, and subdiagonal, respectively, except $A_{1,1} = A_{10,10} = 2$. The matrix is thus singular with nullity 1, and its null vector is $v = \alpha \cdot (-1, 1, -1, 1, -1, 1, -1, 1, -1, 1)^T$. B is 10×8 with values 1, 2, and 3 along its main diagonal, first subdiagonal, and second subdiagonal, respectively.

The eigenvalues of A are $8\sin^2(\frac{\pi j}{2n})$, $j = 0, \ldots, 9$, and so $\lambda_{\min} = 0$, $\lambda_{\max} = 7.8042\ldots$, and the smallest positive eigenvalue is $0.1958\ldots$.

The inner products of the eight columns of B with v are equal in magnitude and alternate in sign. Figure 4.1 shows the condition numbers of the perturbation of the matrix A by multiples of vv^T, the matrix obtained by perturbing A by a multiple of $b_4b_4^T$, where b_4 is the fourth column of B (arbitrary choice), and the matrix obtained by adding a multiple of BB^T to A. The range of values of w for which a nearly optimal condition number is obtained is fairly large and is similar for the three matrices. The condition numbers were computed for $\log_{10} w = -10 : 0.1 : 10$ (201 values for each of the matrices). We obtained $\min_w \kappa_2(A + wbb^T) = 172.3\ldots$, $\min_w \kappa_2(A + wBB^T) = 44.2\ldots$, and $\min_w \kappa_2(A + wvv^T) = 39.9\ldots$. The choice $w = \|A\|/\|B\|^2$ gives $\log_{10} w = -0.6367\ldots$, which is well within the range of values for which the condition numbers in the plot are close to the minimum, and for this value we obtained $\kappa_2(A+wBB^T) = 44.6\ldots$, which is extremely close to the minimum stated above. Finally, we mention that $\kappa_2(A + B(B^TB)^{-1}B^T) = 25.5\ldots$, which is superior to all the other choices and may justify this as a viable choice, as briefly discussed at the end of section 3.

Example 4.2. We took a version of the CUTE matrix `cvxqp1` of size 1500×1500; A is 1000×1000 with rank 986, and B is 1000×500. The smallest positive eigenvalue of A is $9.4756\cdots \times 10^{-7}$. The rank of the saddle point matrix is 1499, and thus $A + BWB^T$ is rank deficient with nullity of (at least) 1 for any choice of W. We set $W = (\|A\|_2/\|B\|_2^2)I$ and generated the right-hand side vector so that the solution was a vector of all 1's.

Figure 4.2 shows the history of residual norms using MINRES with preconditioners $\mathcal{M}$, for five choices of the (1,1) block M. In the top two curves M is the identity matrix and the diagonal of A. The bottom three are related to the shift by a scalar multiple of BB^T: we set $w = \|A\|/\|B\|^2 = 156.13\ldots$, and take $A + wBB^T$,

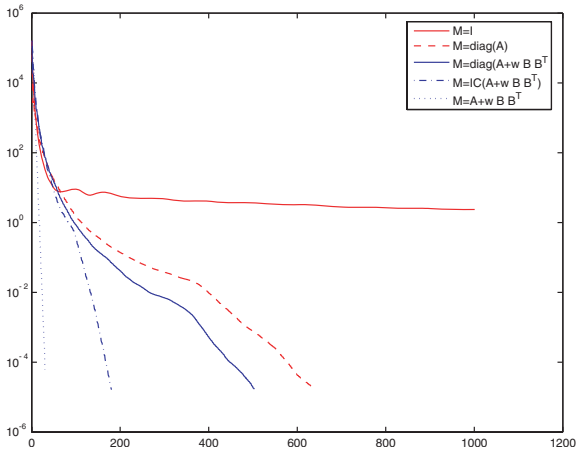

FIG. 4.2. *Example* 4.2*: Convergence of the preconditioned MINRES with various positive definite block diagonal preconditioners.*

its diagonal part, and its incomplete Cholesky decomposition as our choices for M. The number of nonzeros of A is 6968, whereas $A + wBB^T$ has 9950 nonzeros; thus matrix-vector products and linear solves with the latter are more computationally expensive. However the reduced iteration count more than compensates for that, and the overall computational work is reduced when $A+wBB^T$ or its approximations are used. We note that A is sparse but not narrow banded, and is not structured in a way that allows for easily exploiting its nonzero pattern.

Example 4.3. The CUTE matrix `gouldqp3` is 1048×1048, with A of size 699×699 and rank 697. The smallest positive eigenvalue of A is approximately $4.0514\cdot10^{-5}$, and its largest eigenvalue is approximately 5.236. The rank-2 null space of A is spanned by a vector of all 1's (or any other nonzero constant) except the last entry, which is zero, and a vector of all 0's except the last, nonzero entry. Since this null space is explicitly available, the observations of section 3 can be used. We find the columns of B that form the smallest angle with the null vectors of A. The first and the last columns of B are valid (though not unique) choices; let us denote them by b_1 and b_m, respectively. We define $V = w_1b_1+w_2b_m$, with $w_1 = \|A\|/\|b_1\|^2$ and $w_2 = \|A\|/\|b_2\|^2$. We then set $M_1 = A + VV^T$ and $M_2 = A + wBB^T$, with $w = \|A\|/\|B\|^2$.

TABLE 4.1
Number of nonzeros in A, B, $A + VV^T$*, and* $A + BB^T$ *for the* `gouldqp3` *matrix.*

nnz(A)	2092
nnz(B)	1047
nnz($A + VV^T$)	2105
nnz($A + BB^T$)	4185

The number of nonzeros of all the matrices involved is given in Table 4.1. As is evident, M_1 is very close in sparsity to A. The advantage of forming a rank-2 perturbation is that the eigenvalues are all nearly equal to the three eigenvalues of the case described in [13], due to interlacing and the high algebraic multiplicity of each of the eigenvalues. This is illustrated in Figure 4.3. For the choice $A + wBB^T$ the analysis of section 2 holds, and all eigenvalues are within the intervals stated in

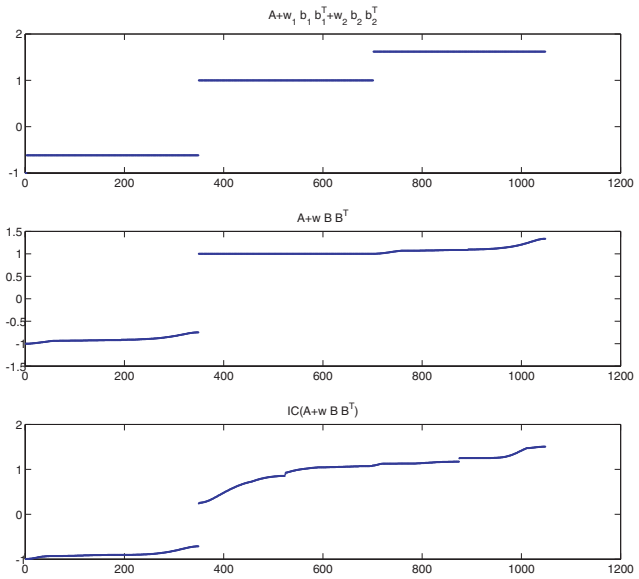

FIG. 4.3. *Example* 4.3*: Eigenvalue distribution of the preconditioned matrix* `gouldqp3` *using various positive definite block diagonal preconditioners.*

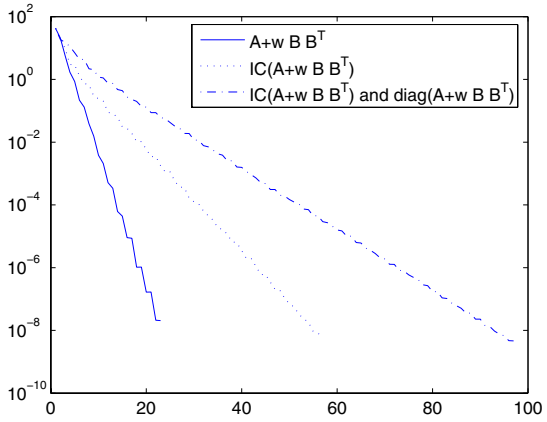

FIG. 4.4. *Example* 4.3*: Convergence of preconditioned MINRES with various positive definite block diagonal preconditioners. The dot-dashed line refers to a preconditioner whose* (1,1) *block is the incomplete Cholesky factorization of* $A + wBB^T$ *and whose* (2,2) *block is* $B^T \operatorname{diag}(A + wBB^T)^{-1}B$.

Theorem 2.5, with the expected algebraic multiplicities. Finally, for the incomplete Cholesky factorization the result of the theorem does not hold, but the eigenvalues of the preconditioned matrix are still fairly close to the eigenvalues in the intervals stated in Theorem 2.5, and none of them is close to zero. Numerical experiments show that the eigenvalue structure well predicts the convergence behavior of preconditioned MINRES, as demonstrated in Figure 4.4.

5. Conclusions. We have performed an algebraic analysis for solving (1.1) by using the block diagonal positive definite preconditioner (1.6), whose (1,1) block is $A + BWB^T$. While the preconditioner in that precise form may be computationally expensive, it forms a basis for a preconditioning approach (based on approximations of the components of (1.6)) that may be useful in practice.

Our analysis shows that the eigenvalues of the preconditioned matrix are clustered in fixed intervals whose ends are fixed far from the origin. Moreover, the eigenvalue 1 is of algebraic multiplicity at least $n - m$, and interestingly, the multiplicity goes up with the nullity of A. In section 3 we discussed an approach of minimizing the condition number of $A + BWB^T$. Our analysis shows that columns of B that form a small angle with the null vectors of A should be considered in the selection of nonzero entries of W. In practice obtaining the null space of A is a computationally nontrivial task for large scale problems, and we have provided analytical justification for using all the columns of B and choosing a scaling parameter of the form $w = \|A\|/\|B\|^2$. This involves a minimal computational effort when cheap norm or extremal eigenvalue estimators are used.

Acknowledgments. We are grateful to the two referees, whose helpful suggestions have substantially improved this paper. We dedicate this paper to Michael Saunders, a kind gentleman and outstanding scientist, for the tireless devotion and unflagging patience with which he spends countless hours with colleagues, helping to improve the quality of their work and writing. Among his many contributions to numerical linear algebra and optimization, Michael has played a pivotal role in the software implementation of robust algorithms for solving the type of indefinite linear systems discussed in this paper.

REFERENCES

[1] K. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Nonlinear Programming*, Stanford University Press, Stanford, CA, 1958.
[2] K. J. Arrow and R. M. Solow, *Gradient methods for constrained maxima, with weakened assumptions*, in Studies in Linear and Nonlinear Programming, K. J. Arrow, L. Hurwicz, and H. Uzawa, eds., Stanford University Press, Stanford, CA, 1958, pp. 166–176.
[3] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle-point systems*, in Acta Numerica, Acta Numer. 14, Cambridge University Press, Cambridge, UK, 2005, pp. 1–137.
[4] P. Bochev and R. B. Lehoucq, *On the finite element solution of the pure Neumann problem*, SIAM Rev., 47 (2005), pp. 50–66.
[5] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., Wiley, Chichster, UK, 1987.
[6] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, North–Holland, Amsterdam, 1983.
[7] G. H. Golub and C. Greif, *On solving block-structured indefinite linear systems*, SIAM J. Sci. Comput., 24 (2003), pp. 2076–2092.
[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[9] N. I. M. Gould, D. Orban, and Ph. L. Toint, *CUTEr, a Constrained and Unconstrained Testing Environment, Revisited*, http://cuter.rl.ac.uk/cuter-www/.
[10] A. Greenbaum and G. H. Rodrigue, *Optimal preconditioners for a given sparsity pattern*, BIT, 29 (1990), pp. 610–634.
[11] R. Hiptmair, *Finite elements in computational electromagnetism*, in Acta Numerica, Acta Numer. 11, Cambridge University Press, Cambridge, UK, 2002, pp. 237–339.
[12] 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.
[13] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972.
[14] J. Nocedal and S. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.
[15] M. L. Overton, *Large-scale optimization of eigenvalues*, SIAM J. Optim., 2 (1992), pp. 88–120.

[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] I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numer. Linear Algebra Appl., 7 (2000), pp. 585–616.
[18] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
[19] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.