Augmented Lagrangian Techniques for Solving Saddle Point Linear Systems

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

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

Submitted to SIAM Journal on Matrix Analysis and Applications
April 12, 2004

Abstract. We perform an algebraic analysis of a generalization of the augmented Lagrangian method for solution of saddle point linear systems. It is shown that in cases where the (1,1) block is singular, specifically semidefinite, a low-rank perturbation that minimizes the condition number of the perturbed matrix while maintaining sparsity is an effective approach. The vectors used for generating the perturbation are columns of the constraint matrix that form a small angle with the null-space of the original (1,1) block. Block preconditioning techniques of a similar flavor are also discussed and analyzed, and the theoretical observations are illustrated and validated by numerical results.

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

We dedicate this paper to Michael Saunders, a kind gentleman and outstanding scientist, on the occasion of his 60th birthday. 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. We thank Michael 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.

1. Introduction. We introduce a generalization of the augmented Lagrangian method for saddle point linear systems, and demonstrate its effectiveness. The discussion is algebraic, with focus on the discrete problem without relating it to any particular underlying continuous problem. Our starting point is the following 'standard' saddle point linear system:

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

The matrix A is $n \times n$, and B is $n \times m$, with $m < n$ (possibly $m \ll n$); these matrices are typically structured and sparse, attributes that could be taken advantage of when solving the linear system. Linear systems of the form (1.1) arise in numerous areas, ranging from nonlinear programming to the solution of discretized partial differential equations for fluid flow problems, elasticity problems, and many others, cf. [24] for

*Department of Computer Science, The University of British Columbia, Vancouver B.C. V6T-1Z4, Canada, greif@cs.ubc.ca. The work of this author was supported in part by the Natural Sciences and Engineering Research Council of Canada.

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

‡Professor Emeritus, Department of Computer Science, The University of British Columbia, Vancouver B.C. V6T-1Z4, Canada, varah@cs.ubc.ca. The work of this author was supported in part by the Natural Sciences and Engineering Research Council of Canada.

references. The numerical treatment of these linear systems depends strongly on the properties of the (1,1) block, A. The case of a symmetric positive definite A has been extensively investigated theoretically and experimentally in many papers and books. The case of a nonsymmetric A has also been subject to much research in the last few decades, due to the importance of applications that lead to such linear systems, a notable example being the Navier-Stokes equations, where A is real positive. The matrix A could be singular, for example when it corresponds to certain discretized operators arising from elliptic PDEs with Neumann boundary conditions. Singular A's frequently arise also in numerical optimization. The matrix B could be rank deficient as well, for example when it corresponds to a discretized gradient operator, as in the Stokes or Navier-Stokes equations, but this implies that the matrix $\mathcal{H}$ is singular. A common stabilization/regularization approach in this case involves the modification of the (2,2) block from zero to a 'convenient' nonzero matrix.

In [24] (and in the earlier technical report [23]), the first two authors showed that replacing (1.1) by the mathematically equivalent system

$$\begin{pmatrix} A + BWB^T & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c + BWd \\ d \end{pmatrix} \tag{1.2}$$

may be numerically advantageous. Here W is an $m \times m$ symmetric positive semidefinite matrix, and is to be determined. The new (1,1) block is frequently referred to in the sequel and deserves to be defined:

$$M(W) = A + BWB^T. \tag{1.3}$$

For notational convenience we also define

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

The approach of replacing (1.1) by (1.2) can be regarded as a generalization of the classical augmented Lagrangian method (also known as the method of multipliers) for saddle point linear systems, whereby the (1,1) block is given by $A + \gamma BB^T$, γ a positive scalar. With a slight abuse of notation, in the sequel we denote the saddle point matrix corresponding to this specific choice as $\mathcal{H}(\gamma)$. One obvious advantage of forming (1.2) is that even when A is singular, the associated Schur complement is well defined, whereas in (1.1) there is no Schur complement associated with a singular A. Our goal is to use (1.2) to obtain a positive definite (1,1) block, and apply block preconditioners with positive definite block elements, that can be conveniently used during the course of solving the linear system using Krylov subspace solvers like MINRES [48].

Extensive investigation of augmented Lagrangian methods has been done by numerous researchers, in areas ranging from fluid flow, electromagnetics and solid mechanics, to numerical optimization and statistics. Arrow & Solow [1], and later Hestenes [35] and Powell [51], proposed the method. Several books and papers have been written on the subject, of which we mention in particular the 1983 book of Fortin & Glowinski [19], where many interesting linear algebra aspects of the technique are discussed within the framework of discretized fluid flow problems, and the book of Bertsekas [6], that provides a thorough analysis of the method in the context of constrained optimization. The choice of the optimal parameter for the classical augmented Lagrangian method is discussed, for example, in [6, Sec. 2.2.2-2.2.4]. A

concise overview of the method and some of its features in the context of nonlinear optimization can also be found in [20].

Since our algebraic framework is first and foremost connected to numerical linear algebra aspects, an extensive overview of the augmented Lagrangian approach and its applications is beyond the scope of this paper; the interested reader can consult [6, 8, 19, 22, 45] and references therein. Those references contain many important results that are relevant to the discussion here.

In [24] we mainly considered the specific choice $W = \gamma I$. We showed that there is a range of values of γ for which the condition number of $\mathcal{H}(\gamma)$, as well as the condition numbers of its (1,1) block and the associated Schur complement, form a set that is (generally) numerically superior to the equivalent set of condition numbers associated with the original matrix, $\mathcal{H} \equiv \mathcal{H}(0)$. We experimentally observed that the choice based on scaling, $\gamma = \|A\|_2/\|B\|_2^2$, is typically effective. For large enough values of γ we also showed that the negative eigenvalues of $\mathcal{H}(\gamma)$ tend to cluster near $-1/\gamma$, hence giving rise to faster convergence of Krylov iterative solvers, but at the same time for large γ the condition number becomes larger compared to the condition number of $\mathcal{H}$. A preliminary analysis of related block preconditioning approaches was provided in the technical report [23]. The point made in [24] was that for a positive semidefine A and the choice $W = \gamma I$ there exists a range of (typically moderate) values of γ for which the spectral properties and the conditioning of the associated matrix are possibly better than those of the original matrix. The viability of such a choice of the parameter stands in line with analytical observations made in [6, 7, 19] and elsewhere, although those references do not necessarily focus on linear algebra aspects of the augmented Lagrangian method.

In this paper we devise a general technique that is not restricted to seeking a single parameter. There are desired goals like minimizing the condition number, maintaining sparsity, obtaining positive definiteness, and other considerations, not all of which can be satisfied simultaneously. We aim to obtain desired favorable spectral properties while maintaining proper sparsity and reduced computational costs in solving the linear system. This, as we show, can be done by choosing W to be a diagonal nonnegative matrix, and our approach amounts to selecting the set of nonzero entries of W so that sparse columns of B are chosen, that are as much as possible aligned with the null-space of A. The choice of the nonzero values of W is determined so as to minimize the condition number of the perturbed matrix. Here we mention that in [7], P. Bochev & Lehoucq provide helpful insight into the role of the null-space of A and low-rank perturbations. A variational analysis of the augmented Lagrangian method applied to the pure Neumann problem is provided, as well as an SVD-based algebraic analysis [7, Sec. 5.1,5.2.2].

Beyond the choice of a matrix W in the context of forming system (1.2), we explore also how this approach can be incorporated into the general framework of block preconditioners. Popular classes of preconditioners such as constraint preconditioners [19, 26, 39, 42, 50, 52] and block positive definite preconditioners [9, 13, 14, 15, 16, 40, 41, 44, 56] have been analyzed and widely used in saddle point linear solvers. Useful insights into the spectrum of the saddle point matrix can be found in [53]. Other interesting preconditioners for saddle point linear systems are based on the Hermitian/Skew-Hermitian splitting [3, 4, 54]. Finally, we mention the class of null-space and projection methods, that lead to a smaller linear system. For this class of methods there is a variety of techniques for obtaining the null-space of B^T, and various solution approaches, cf. [28, 45] for a description and several references.

The rest of this paper is organized as follows. In Section 2 we introduce our approach of low-rank perturbations to A that minimize the condition number of the (perturbed) matrix. Section 3 is devoted to block preconditioners related to the choices made in Section 2. In Section 4 we provide a set of numerical examples that validate our analysis. In Section 5 we draw some conclusions.

2. Minimization of the condition number of the (1,1) block. At the center of our discussion is the question: In which circumstances is approach (1.2) superior to solving the linear system using (1.1)? The choice of W is critical. We start by clarifying the circumstances in which positive definiteness and simultaneous diagonalization are accomplished.

The following early result was an important step towards the derivation of the augmented Lagrangian method:

LEMMA 2.1. (Debreu, 1952 [10]) *If A is symmetric matrix and if $x^TAx > 0$ for every $x \neq 0$ satisfying $B^Tx = 0$, then, for ρ sufficiently large, $A + \rho BB^T$ is positive definite.*

Debreu's result is important in that it shows that by following the augmented Lagrangian procedure it is possible to obtain a positive definite (1,1) block even if the original matrix was not positive definite. It also allows us to state the following related observation.

In the following two propositions, the following assumptions are made:

- A is a square positive semidefinite matrix with $\text{rank}(A) \geq n - m > 0$.
- B is $n \times m$ with full column rank.
- $\text{null}(A) \cap \text{null}(B^T) = \{0\}$.
- W is $m \times m$, symmetric positive semidefinite, and positive definite on the null-space of B^T.

PROPOSITION 2.2. *Under the above stated assumptions $A+BWB^T$ is symmetric positive definite.*

Proof. Since W is positive semidefinite, $W^{1/2}$ is well defined. Denote $C = BW^{1/2}$, and suppose that $C^T = U\Sigma V^T$ is the singular value decomposition of C^T. Since B^T is of rank m, and since no null vector of W belongs to the range space of B^T, it follows that CC^T is of rank m and has the same null vectors as B^T. Since $\text{null}(A)$ and $\text{null}(B^T)$ do not nontrivially intersect, the desired result follows. ∎

Next, we present the circumstances in which simultaneous diagonalization occurs. Related results can be found in [25, 38].

PROPOSITION 2.3. *Given the above stated assumptions, suppose that A, B and W are such that $M(W) = A+BWB^T$ is positive definite. Then there exists a nonsingular matrix X such that both $D_A := X^TAX$ and $D_C := X^T(BWB^T)X$ are diagonal.*

Proof. Let $M(W) = FF^T$. Then $I = F^{-1}AF^{-T} + F^{-1}BWB^TF^{-T}$. Let $F^{-1}AF^{-T} = QD_AQ^T$ be a spectral decomposition, where D_A is diagonal and Q is orthogonal, and define $X = F^{-T}Q$. Clearly, $X^TAX = D_A$. Now,

$$\begin{aligned} X^T(BWB^T)X &= Q^TF^{-1}BWB^TF^{-T}Q \\ &= Q^T\left(I - F^{-1}AF^{-T}\right)Q \\ &= I - D_A = D_B. \end{aligned}$$

∎

As mentioned in the Introduction, one of the difficulties in forming $M(W)$ is the potential loss of sparsity, which could occur if A and B have substantially different sparsity patterns. Moreover, BB^T is typically not as sparse as B, and so from a linear algebra point of view the popular choice of $W = \gamma I$ could at times result

in a computationally costly solution procedure that may void the advantages of a better spectral distribution and faster convergence (in terms of iteration counts) of iterative solvers. We propose to deal with this difficulty by setting W to be a diagonal nonnegative matrix such that $M(W)$ is sparse:

$$W = \mathrm{diag}(w_j) \quad , \quad w_j \geq 0. \tag{2.1}$$

We allow several of the entries in the diagonal of W to be 0. This means that in setting the perturbed matrix $M(W)$ we are using only some of the columns of B; dense columns could be excluded a priori. We then set the nonzero entries of W so that the condition number of the resulting matrix is minimized.

While the introduction of only a small number of parameters (namely the nonzero elements of the diagonal of W) may be helpful in maintaining the sparsity, an important requirement is that the procedure of setting W be significantly cheaper than solving the linear system itself, as this is a 'preprocessing' stage. (This requirement can be relaxed if the linear system is repeatedly solved, as part of a nonlinear solve.)

When W is diagonal we can write the (1,1) block as a sum of rank-1 perturbations of A,

$$M(W) = A + \sum_{j \in \mathcal{S}} w_j b_j b_j^T, \tag{2.2}$$

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 cannot be arbitrarily small, as the following result clarifies.

COROLLARY 2.4. *Suppose that A is symmetric positive semidefinite and its null-space does not nontrivially intersect with the null-space of B^T. A necessary and sufficient condition for the existence of a diagonal nonnegative matrix W such that $A + BWB^T$ is positive definite is that the number of nonzeros in W be equal to at least the nullity of A.*

The proof of the proposition follows from Prop. 2.2, and is based in part on the observation that each time a rank-1 perturbation is applied, a single zero eigenvalue is mapped into a nonzero one, by the Cauchy interlacing theorem. Note that this means that we must have $\mathrm{rank}(A) \geq n - m$. However, the numerical usefulness of Prop. 2.4 is limited in situations where nonzero eigenvalues of A are numerically close to zero, since in these cases some of the nonzero perturbed eigenvalues are also close to zero.

2.1. Minimizing of the condition number of a semidefinite matrix with nullity 1. Let us restrict the discussion to the case of a rank-1 perturbation of a matrix of nullity 1:

$$M(w) = A + wbb^T. \tag{2.3}$$

The following assumptions are made throughout this section, unless otherwise stated:

1. $w > 0$ is a scalar.
2. b is a *known* vector and is not orthogonal to the null vector of A.
3. A has nullity 1.
4. The largest eigenvalue of the matrix A is also simple:

$$\lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_{n-1} > \lambda_n = 0.$$

The strategy we take in selecting w is to *minimize the condition number of* $M(w)$:

$$\boxed{\min_{w>0} \kappa_2(A + wbb^T)}$$

For notational convenience, let us define

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

The problem of minimizing the condition number of a matrix or minimizing its maximal eigenvalue, and other problems of a similar flavor, have been subject to thorough investigation and analysis, notably by Overton and his collaborators, within the framework of large-scale eigenvalue optimization. See for example the comprehensive paper [47] and many related references therein. The paper [46] provides a quadratically convergent algorithm for minimizing the maximal eigenvalue of a matrix $A(x)$, an algorithm that was later used by Greenbaum & Rodrigue [30] for computing an optimal preconditioner, given a prescribed sparsity pattern. The authors showed how to find a positive definite symmetric matrix N with a given sparsity pattern that minimizes the spectral condition number of $N^{-1/2}CN^{-1/2}$, where C is a given symmetric positive matrix. The algorithm was applied to the tridiagonal matrix associated with the discretized Laplace operator in one dimension.

The problem of finding the eigenvalues of low-rank perturbations of symmetric matrices is lucidly discussed in the books [11, 25, 49, 57] and in other places. During the 1990s, it was subject to investigation in the context of a stable computation of eigenvalues of symmetric tridiagonal matrices using the divide and conquer method, see [31, 32].

Our above defined problem requires solving the nonlinear scalar equation

$$\frac{d\kappa_2(A + wbb^T)}{dw} \equiv \frac{d\kappa_2(w)}{dw} = 0. \tag{2.4}$$

Since A is assumed to be symmetric, $\kappa_2'(w)$ is a function of the largest and smallest eigenvalues, and their derivatives. The eigenvalues of the perturbed matrix are the zeros of the function f in the so-called *secular equation* [25, p. 443]: $f(w) = 1 + w\sum_i \frac{\tilde{b}_i^2}{\lambda - \lambda_i}$, where λ_i are the eigenvalues of the original (unperturbed) matrix, and $\tilde{b}_i = U^T b_i$, with U being the left singular vector of A: $A = U\Sigma V^T$.

By the interlacing property, the ith eigenvalue, $\lambda_i(w)$, of the perturbed matrix $A + wbb^T$ is bounded between $\lambda_{i-1}(A)$ and $\lambda_i(A)$, $i = 2, \ldots, n$, and $\lambda_1(w) > \lambda_1(A)$. (Note that we order the eigenvalues so that λ_1 is the largest.) We assume that the function $\kappa_2(w)$ is smooth and convex.

Suppose w is given, and

$$M(w)u(w) = \lambda(w)u(w).$$

Consider the following perturbed eigenvalue problem, where for notational convenience $M(w)$ is denoted simply by M:

$$\begin{aligned}(M + \delta w \cdot E)(u + \delta w \cdot u' + (\delta w)^2 \cdot u'' + \cdots) \\ = (\lambda + \delta w \cdot \lambda' + (\delta w)^2 \cdot \lambda'' + \cdots)(u + \delta w \cdot u' + (\delta w)^2 \cdot u'' + \cdots).\end{aligned}$$

The leading order term gives $(M-\lambda I)u' = \lambda' u - Eu$, and multiplying by u^T on the left yields an expression for λ', which is the Rayleigh Quotient $\lambda' = \frac{u^T E u}{u^T u}$. Substituting $E = bb^T$ and taking a normalized eigenvector, $\|u(w)\|_2 = 1$, we get

$$\frac{d\lambda}{dw} = [b^T u(w)]^2. \tag{2.5}$$

Minimizing the condition number $\kappa_2(w) = \dfrac{\lambda_1(w)}{\lambda_n(w)}$ implies $\frac{\lambda_1'(w)\lambda_n(w) - \lambda_n'(w)\lambda_1(w)}{\lambda_n^2(w)} = 0$.

PROPOSITION 2.5. *Suppose that A is a symmetric positive definite matrix, and let b be an arbitrary vector. Suppose that $Au_1 = \lambda_1 u_1$ and $Au_n = \lambda_n u_n$ define the largest and smallest eigenpairs, respectively, of A, and $\lambda_1 \& \lambda_n$ are simple. Then a sufficient condition for $\kappa_2'(0) < 0$ is*

$$\sqrt{\kappa_2(0)} > \frac{|b^T u_1|}{|b^T u_n|} \ . \tag{2.6}$$

This condition straightforwardly extends to the case of a matrix B in place of a vector b:

$$\sqrt{\kappa_2(0)} > \frac{\|B^T u_1\|_2}{\|B^T u_n\|_2}. \tag{2.7}$$

Proof. Write $\kappa_2'(0) = \frac{\lambda_1'(0)\lambda_n(0) - \lambda_n'(0)\lambda_1(0)}{\lambda_n^2(0)}$ The condition $\kappa_2'(0) < 0$ holds if $\frac{\lambda_1'(0)}{\lambda_1(0)} < \frac{\lambda_n'(0)}{\lambda_n(0)}$, which, together with (2.5), readily follows if (2.6) is satisfied. The equivalent result for a matrix B follows by using the same technique. □

The following result is related to the *effective* condition number of a matrix. The proof follows from the orthogonality of the eigenvectors and is omitted.

PROPOSITION 2.6. *Suppose that A is a symmetric positive semidefinite matrix 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 $\kappa_2(A + w u_n u_n^T)$ is strictly decreasing in w for $0 = \lambda_n < w < \lambda_{n-1}$, constant for $\lambda_{n-1} \le w \le \lambda_1$, strictly increasing for $w > \lambda_1$, and is explicitly given in terms of the eigenvalues of A as follows:*

$$\kappa_2(A + w u_n u_n^T) = \frac{\max(\lambda_1, w)}{\min(\lambda_{n-1}, w)}. \tag{2.8}$$

The smallest *positive* eigenvalue of a singular matrix plays an important role, and the rate of convergence of Krylov solvers is determined largely by the nonzero eigenvalues [2, Chapter 13],[55, Chapter 10]. As is pointed out in [7], the computational cost of matrix-vector products associated with a rank-1 perturbation (even if it is dense) is only very modestly higher than for the original matrix, since matrix-vector products of the form $(A + vv^T)x$ with a column vector v can be performed as a matrix vector product with A and two inner-products.

Propositions 2.5 and 2.6 lead us to the following conclusion: we should take columns of B that are well aligned (i.e. form an angle as small as possible) with the null vector(s) of A. Note that for those columns B that are strongly aligned with the direction of the null vector u_n, from orthogonality of the eigenvectors it follows that those vectors are nearly orthogonal to u_1, which may make it easier to obtain the

condition (2.6). With this observation in mind, we propose the following strategy.

The choice of a (single) nonzero diagonal entry of W:
Given (1.1) with a symmetric positive semidefinite A of nullity 1 and a normalized null vector u_n, choose the column of B, say b_k, that forms the smallest angle with u_n, and compute the scalar w for which $\kappa_2(A + wb_kb_k^T)$ is minimized.

The numerical procedure for computing the scalar that minimizes the condition number is to be discussed in Section 2.2. This approach can be generalized to a higher nullity of A and a larger number of columns of B, using concepts of angles between linear subspaces; see [25, pp. 603-604]. The choice of columns of B can be made from a set of preselected 'sufficiently sparse' columns. These issues are not further pursued here.

Example. A 500×500 dense random matrix of rank 499 was generated using the following MATLAB script, with `n` = 500:

```
A=randn(n);
A=A*A';
A=A-min(eig(A))*eye(n);
```

The matrix A has nullity 1, and the smallest positive eigenvalue was computed to be $\lambda_{n-1} = 0.0021$. The largest eigenvalue of A was found to be $\lambda_1 = 2025.2$ We then computed a series of condition numbers of the matrices $A + wbb^T$, where b is a random vector that forms a small angle with the null vector of A. Figure 2.1 shows our results. Note the logarithmic scale. The initial decline of the condition number is very rapid and by the time w becomes of order 10^{-3}, the condition number reaches its minimum.

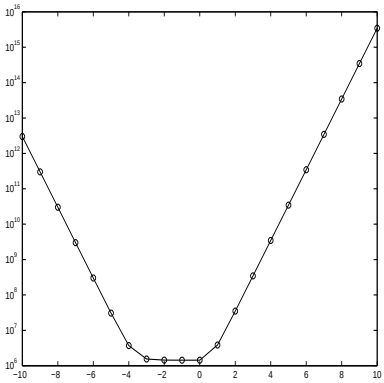


FIG. 2.1. *A graph of $\kappa(A + wbb^T)$ for the 500×500 matrix mentioned in the example. The circles correspond to the 21 values that were computed: condition numbers for $\log_{10} w = [-10 : 10]$. The x-coordinate corresponds to $\log_{10} w$ and the y-coordinate shows the condition number.*

2.2. A numerical procedure for evaluating the parameter that minimizes the condition number. Looking at Figure 2.1, we see that the condition number decreases *rapidly* with w, before starting to increase again. The rapid decrease of the condition number and the large range of possible values of the optimal

parameter for different problems, suggests that logarithmic re-scaling of the parameter w may be beneficial. We thus re-define the minimization problem accordingly:

$$\frac{d\kappa_2(A + e^{\rho(w)}bb^T)}{dw} = 0, \tag{2.9}$$

where $\rho(w) = \ln w$. In order to avoid computing second derivatives, we employ a secant scheme for solving the scalar equation. Since $\|A\|/\|B\|^2$ has experimentally proven to be an effective choice in many experimental tests, we use it as an initial guess. Since

$$\frac{d\kappa_2(A + e^{\rho}bb^T)}{dw} = e^{\rho} \cdot \frac{(b^T u_1)^2 \cdot \lambda_n - (b^T u_n)^2 \lambda_1}{\lambda_n^2}, \tag{2.10}$$

the same nonlinear equation as before is obtained:

$$f(\rho) = (b^T u_1(\rho))^2 \cdot \lambda_n(\rho) - (b^T u_n(\rho))^2 \lambda_1(\rho) = 0, \tag{2.11}$$

except now $(\lambda_1(\rho), u_1(\rho))$ and $(\lambda_n(\rho), u_n(\rho))$ are the extreme eigenpairs of the matrix $A + e^{\rho}bb^T$.

Referring to the function f defined in (2.11), our scheme is as follows:

- Set $\rho_0 = \log\|A\| - 2\log\|B\|$, compute $f(\rho_0)$ as in (2.11), and $\rho_1 = \rho_0 - \varepsilon$, $\varepsilon > 0$.

- For $i = 1, 2, \ldots$ until satisfied:
  - Evaluate the extreme eigenpairs $(\lambda_1(\rho_i), u_1(\rho_i))$ and $(\lambda_n(\rho_i), u_n(\rho_i))$ and compute $f(\rho_i)$.
  - Compute $\rho_{i+1} = \rho_i - \frac{f(\rho_i)}{(f(\rho_i) - f(\rho_{i-1}))/(\rho_i - \rho_{i-1})}$.

Upon termination, $w = \exp(\rho_i)$ is the desired parameter.

2.3. Discussion of computational costs. As mentioned before, the numerical procedure stated above must be computationally inexpensive, because it is used at the 'preprocessing' stage, before the linear system is solved. The following remarks address the issue of assuring a low computational cost:

1. *An inexact/incomplete procedure is sufficient.* There is no need to obtain an exact solution to the problem. As long as the condition number of the perturbed matrix is reasonably small, mission has been accomplished. Therefore no more than two or three iterations should be carried out, and/or a low threshold as a stopping criterion can be imposed.
2. *Rapid factorizations associated with low-rank corrections.* There are well-known inexpensive techniques to factorize a low-rank perturbation of a matrix given the factorization of the original matrix. If $\tilde{A}$ is a positive definite matrix with $\tilde{A} = FF^T$, then $\tilde{A} + wbb^T = F(I + wcc^T)F^T$, where $Fc = b$, and a variety of possibilities to proceed exist.
3. *The cost is reduced if the null-space of A is known.* In some cases the null-space of A is known, in which case no factorizations are required and the cost of the procedure reduces dramatically.
4. *Use of factorization at the linear system solve stage.* The factorization of $A + e^w bb^T$ from the last iteration in this procedure can be used later in the iterative solution of the linear system.

5. *Derivative-free algorithms.* The algorithm proposed above can be replaced by a derivative-free algorithm, i.e. one that does not require solving for the derivative of κ. If the function of the condition number is convex, an iterative minimization procedure such as the successive parabolic iteration [34, p.273] can be used. Here a condition number estimator that involves no eigenvalue computations [36] can be used.
6. *Cost of procedure compared to the overall computations.* Finally, we note that in some cases (e.g. certain instances of nonlinear optimization) the overall cost of computing the parameters involved, even if high, may be small compared to the entire solution procedure.

3. Related block preconditioning approaches. Staying within the framework of a low-rank perturbation of the (1,1) block, we now move to consider block preconditioners with the same flavor.

Notation. For the preconditioners we discuss in this section, we denote by M approximations to the (1,1) block. Also, let us define two matrices that are to be frequently referred to in the sequel. The first is given by

$$K = M^{-1/2} A M^{-1/2} \ . \tag{3.1}$$

The second matrix is an orthogonal projector onto range$(M^{-1/2}B)$, given by

$$P = C(C^T C)^{-1} C^T, \tag{3.2}$$

where

$$C = M^{-1/2} B. \tag{3.3}$$

The block preconditioners we refer to are:

- *Constraint preconditioners*, to which we attach the subscript 'C':

$$\mathcal{M}_C(M) = \begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix}. \tag{3.4}$$

- *Positive definite block diagonal preconditioners*, to which we attach the subscript 'P':

$$\mathcal{M}_P(M) = \begin{pmatrix} M & 0 \\ 0 & B^T M^{-1} B \end{pmatrix}. \tag{3.5}$$

Finally, we define the generalized eigenvalues $\{\lambda_i\}$

$$Ax = \lambda B W B^T x. \tag{3.6}$$

3.1. Constraint preconditioning. An increasingly popular technique is *constraint preconditioning* [12, 26, 39, 42, 50, 52], in which the preconditioner has the form (3.4) for some M. An attractive property of these preconditioners is that the constraints are left intact, and thus, the solution always stays on the constraint manifold. The block decomposition

$$\begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^T M^{-1} & I \end{pmatrix} \begin{pmatrix} M & 0 \\ 0 & -B^T M^{-1} B \end{pmatrix} \begin{pmatrix} I & M^{-1} B \\ 0 & I \end{pmatrix} \tag{3.7}$$

clarifies, as is pointed out in [52], that constraint preconditioning with a positive definite (1,1) block allows for using positive definite solves throughout the computation, despite the fact that the original formulation involves an indefinite preconditioner applied to an indefinite linear system.

In the context of a modified (1,1) block related to the approach discussed in this paper, the following result is stated in [24]. (See also [17],[18, Chap.12].) Suppose A is a general $n \times n$ matrix, B and C are full column rank $n \times m$ matrices ($m \leq n$), and W is a $m \times m$ matrix. If

$$\tilde{\mathcal{H}}(W) = \begin{pmatrix} A + BWC^T & B \\ C^T & 0 \end{pmatrix},$$

then for any $W \neq 0$ such that $\tilde{\mathcal{H}}(W)$ is nonsingular,

$$\tilde{\mathcal{H}}^{-1}(W) = \tilde{\mathcal{H}}^{-1}(0) - \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}. \tag{3.8}$$

The next result immediately follows.

PROPOSITION 3.1. *If $\mathcal{H}(W)$ (defined in (1.4)) is used as a preconditioner for solving the linear system (1.1) by applying a minimum residual Krylov subspace solver, convergence is obtained (if roundoff errors are ignored) within at most $m + 1$ iterations.*

Proof.

$$\mathcal{H}^{-1}(W)\mathcal{H} = \left\{\mathcal{H}^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}\right\}\mathcal{H} = \begin{pmatrix} I & 0 \\ WB^T & I \end{pmatrix}. \tag{3.9}$$

Hence all eigenvalues are equal to 1. Recall that the (1,1) block of the above matrix is of size $n \times n$; thus convergence will depend on the Jordan blocks associated with the lower blocks (m rows) of the matrix, and in any case will not exceed $m + 1$ iterations. □

Solving a linear system whose associated preconditioning matrix is $\mathcal{H}(W)$ is likely to be too computationally costly in most cases. However, the observation made in Prop. 3.1 is useful as a starting point, with the goal of generating approximations of $A + BWB^T$ for which rapid convergence can still be obtained.

For notational convenience, when there is no ambiguity with regard to M we denote below $\mathcal{M}_{\mathcal{C}}(M)$ simply by $\mathcal{M}_{\mathcal{C}}$. Expressions for the preconditioned matrix ${\mathcal{M}_{\mathcal{C}}}^{-1}\mathcal{H}$ have been obtained elsewhere (see e.g. [5, 26]), and here we perform the calculation in order to express the preconditioned matrix in terms of the matrices K and P, defined in (3.1) and (3.2) respectively. Suppose ${\mathcal{M}_{\mathcal{C}}}^{-1}\mathcal{H} = \begin{pmatrix} X & Y \\ V & Z \end{pmatrix}$, where X is $n \times n$ and Z is $m \times m$. Then

$$\begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix}\begin{pmatrix} X & Y \\ V & Z \end{pmatrix} = \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}.$$

From the expressions for the (1,2) and (2,2) blocks, we have respectively $MY + BZ = B$ and $B^TY = 0$. Multiplying the former by B^TM^{-1}, and using the latter, we immediately obtain $Z = I_m$, from which it follows that $MY = 0$, and by the nonsingularity of M we obtain $Y = 0$.

Next, we consider the expressions for X and V. For the latter, we have $B^TX = B^T$, from which it follows that $X - I_n \in \text{null}(B^T)$. While the size of $X - I_n$ is $n \times n$,

its rank cannot exceed $n-m$. The equation $MX+BV=A$ is multiplied by B^TM^{-1}, and we use $B^TX=B^T$ to obtain $V=-S^{-1}B^T(I-M^{-1}A)$, where $S=B^TM^{-1}B$. Notice that $I-M^{-1}A$ is of the classical form of an *iteration matrix* associated with the splitting $A=M-(M-A)$.

Having an expression for V, we can now obtain an explicit expression for X. From the relation $X=M^{-1}(A-BV)$ we obtain $X=M^{-1}A+M^{-1}BS^{-1}B^T(I-M^{-1}A)$. Since $P=P^2=M^{-1/2}BS^{-1}B^TM^{-1/2}$, we have $X=M^{-1}A+M^{-1/2}PM^{1/2}(I-M^{-1}A)$. The expression for X can be further simplified, as follows:

$$X=M^{-1/2}\left(M^{-1/2}AM^{-1/2}+P-PM^{-1/2}AM^{-1/2}\right)M^{1/2},$$

which can be rewritten as $X=M^{-1/2}(K+P-PK)M^{1/2}$.

We thus have

$$\mathcal{M}^{-1}\mathcal{H}=\begin{pmatrix} X & Y \\ V & Z \end{pmatrix}=\begin{pmatrix} M^{-1/2}(K+P-PK)M^{1/2} & 0 \\ -S^{-1}B^T(I-M^{-1}A) & I_m \end{pmatrix}. \tag{3.10}$$

For computing the eigenvalues of this matrix, it is now clear that it is the eigenvalues of $K+P-PK$ that we need to analyze. As we shall see in the next subsection, this establishes an interesting connection between the eigenvalues of the matrix ${\mathcal{M}_{\mathcal{C}}}^{-1}\mathcal{H}$ and those of the matrix ${\mathcal{M}_{\mathcal{P}}}^{-1}\mathcal{H}$: Both are associated with eigenvalue problems that can be expressed in terms of K and P.

It is straightforward to show that substituting $M=A+BWB^T$ in (3.10) leads back to (3.9); the algebraic details are omitted. The spectrum of the matrix $K+P-PK$ plays a central role in convergence of solvers based on constraint preconditioners, and further observations can be made in connection to the specific choice of W.

3.2. Positive definite block preconditioning. The middle matrix in (3.7) can also be used as a preconditioner, and in [9, 23] preconditioners of this form are discussed. We consider using $\mathcal{M}_P$ as defined in (3.5) as our preconditioner.

Murphy, Golub and Wathen showed in [44] that if A is nonsingular, then the preconditioner

$$\tilde{\mathcal{M}}=\begin{pmatrix} A & 0 \\ 0 & B^TA^{-1}B \end{pmatrix} \tag{3.11}$$

has the attractive property that the associated preconditioned matrix $\tilde{\mathcal{M}}^{-1}\mathcal{H}$ 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. However, inverting A and thus computing the Schur complement could often be prohibitively expensive. Moreover, if A is singular, which is often the case in applications in optimization and other areas, it is impossible to apply this procedure. Replacing A by a nonsingular matrix provides an immediate benefit.

Consider the preconditioned matrix ${\mathcal{M}_{\mathcal{P}}}^{-1}\mathcal{H}$. Let ν be an eigenvalue of this matrix with eigenvector $\begin{pmatrix} u \\ v \end{pmatrix}$. Then

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}=\nu\begin{pmatrix} M & 0 \\ 0 & B^TM^{-1}B \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}. \tag{3.12}$$

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

$$\left[\nu^2M-\nu A-B(B^TM^{-1}B)^{-1}B^T\right]u=0. \tag{3.13}$$

Denote $\tilde{u} = M^{\frac{1}{2}}u$. Then it readily follows that

$$(\nu^2 I - \nu K - P)\tilde{u} = 0, \tag{3.14}$$

where K and P are as in (3.1) and (3.2) respectively. We thus have a quadratic eigenvalue problem here, expressed in terms of K and P. Recall that for constraint preconditioning, discussed in Section 3.1, the eigenvalues of the preconditioned matrix could also be expressed in terms of those two matrices.

Note that the case $M = A$ corresponds to $K = I$, and we are back to (3.11). Suppose M is selected such that P is a projector of rank m; then it has m eigenvalues equal to 1, and $n - m$ eigenvalues equal to 0. The matrix K has eigenvalues that are equal to those of the matrix $M^{-1}A$, hence indicating that the quality of the preconditioner for the saddle point problem is strongly related to the quality of the approximation $M \approx A$. This matches observations made by Golub & Wathen in [26].

LEMMA 3.2. *For $M = A + BWB^T$, where W is symmetric positive semidefinite, the (normal) matrices P and K commute, and thus share the same eigenvectors.*

Proof. Using definitions (3.2) and (3.3) for P and C respectively, we have

$$I - CWC^T = M^{-1/2}(M - BWB^T)M^{-1/2} = M^{-1/2}AM^{-1/2} = K. \tag{3.15}$$

It immediately follows that

$$PK = C(C^TC)^{-1}C^T(I - CWC^T) = KP = P + K - I. \tag{3.16}$$

Hence K and P are commuting normal matrices and therefore [21, p. 291] share the same eigenvectors. ∎

THEOREM 3.3. *Suppose nullity$(A) \le m$, and $M = A + BWB^T$ is symmetric positive definite, with a symmetric positive semidefinite matrix W. Then the following holds:*

1. *The matrix K has an eigenvalue equal to 1 of algebraic multiplicity $n - m$. The eigenvectors are the same ones associated with the $n-m$ zero eigenvalues of P, and span the null-space of B^T.*
2. *The rest of the eigenvalues of K are all between 0 and 1, and are given explicitly in terms of the generalized eigenvalue problem (3.6), as follows.*

$$\mu_i = \frac{\lambda_i}{\lambda_i + 1}. \tag{3.17}$$

Proof. Suppose $Kx_i = \mu_i x_i$. Then μ_i is also an eigenvalue of $M^{-1}A$, with eigenvector $z_i = M^{-1/2}x_i$, and we have $Az_i = \mu_i(A+BWB^T)z_i$. If A is singular, then the matrix $M^{-1}A$ has nullity(A)) zero eigenvalues, and the corresponding eigenspace is null(A). On the other hand, for $\mu_i \neq 0$ we have $Az_i = \frac{\mu_i}{1-\mu_i}BB^Tz_i$, from which item 2 in the statement of the theorem follows.

Now, the null-space of B^T is of dimension $n - m$, and thus the set of vectors $z_i = \text{null}(B^T)$ must be eigenvectors. For those eigenvectors, the associated eigenvalues are infinite, and by (3.17) $\mu_i = 1$ is an eigenvalue of algebraic multiplicity $n - m$. ∎

In the specific case $W = \gamma I$ we can show that in addition to the mathematical observations on K and P made in Theorem 3.3 there is an explicit formula that connects P and K.

THEOREM 3.4. *Suppose $W = \gamma I$. The projection matrix P is a polynomial of degree m in the matrix K, and it is given explicitly by $P = I - f(K)$, where f is the*

Lagrange interpolant

$$f(t) = \frac{\prod_{i=1}^{m} \left(t - \frac{\lambda_i}{\lambda_i + \gamma} \right)}{\prod_{i=1}^{m} \left(1 - \frac{\lambda_i}{\lambda_i + \gamma} \right)}, \tag{3.18}$$

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

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

The polynomial connection is now established, and we can seek an explicit expression. Denote the eigenvalues of K by $\{\mu_i\}$. Since $K = M^{-1/2}AM^{-1/2}$, we have that $eig(K) = eig(M^{-1}A) = eig(A - \mu M)$. 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.

$$\lambda_i = \frac{\mu_i \gamma}{1 - \mu_i}. \tag{3.19}$$

Note that this relation is slightly different than (3.17) due to the presence of the parameter γ. Since $\text{rank}(BB^T) = m$, the $n - m$ infinite generalized eigenvalues λ_i map into $\mu_i = 1$ in (3.19). There are m finite generalized eigenvalues λ_i, and for those, (3.17) holds. The m eigenvalues of K that are not equal to 1 thus satisfy $0 \leq \mu_i < 1$.

As for the matrix P, since $n - m$ of its eigenvalues are equal to 1, we know that the function f defined in the statement of this theorem satisfies $f(1) = 1$ and $f\left(\frac{\lambda_i}{\lambda_i + \gamma}\right) = 0$, and thus, since we established a polynomial connection, f is the Lagrange interpolant of degree m, given by (3.18).

□

The explicit polynomial connection introduced in Theorem 3.4 holds if $W = \gamma I$. For any other W such a polynomial connection does not necessarily exist, (in particular if W is singular) but we can still obtain a useful result related to the clustering of the eigenvalues.

THEOREM 3.5. *If* A *is positive semidefinite, the eigenvalues of the preconditioned matrix are bounded within the two intervals:*

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

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

Proof. Since $\text{null}(B^T)$ 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, we get that $n - m$ eigenvalues of the preconditioned matrix $\mathcal{M}_{\mathcal{P}}{}^{-1}\mathcal{H}$ satisfy $\nu^2 - \nu = 0$, and assuming that the preconditioned matrix is nonsingular, we see that $\nu = 1$ are those eigenvalues.

For the rest of the eigenvalues, by (3.14) we have the connection $\lambda^2 - \lambda\mu = 1$, which leads to

$$\lambda = \frac{1}{2}\left(\mu_i \pm \sqrt{\mu_i^2 + 4}\right). \tag{3.20}$$

The zero eigenvalues of A are mapped into $\lambda = \pm 1$. 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 3.3) we can deduce that

$$1 \leq \frac{1}{2}(\mu + \sqrt{\mu^2 + 4}) \leq \frac{1}{2}(1 + \sqrt{5}) \approx 1.618,$$

and

$$-0.618 \approx \frac{1}{2}(1 - \sqrt{5}) \geq \frac{1}{2}(\mu - \sqrt{\mu^2 + 4}) \geq -1.$$

□

We conclude that all the eigenvalues of the preconditioned matrix are within two narrow intervals whose ends are fixed. The construction of this preconditioner in its current form may still be costly for large problems, and it is therefore necessary to find efficient approximations to its block elements.

3.3. Numerical properties of the Schur complement. The Sherman-Morrison formula (see, e.g. [25, p. 50]) states that for an $n \times n$ nonsingular matrix A and $n \times k$ matrices U and V the following relation holds (provided that $I + V^T A^{-1} U$ is nonsingular):

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1} U)^{-1} V^T A^{-1}.$$

Given a nonsingular matrix $\tilde{A}$, a vector b, and a scalar w, using the formula with $U = V = \sqrt{w}b$, we have

$$(\tilde{A} + wbb^T)^{-1} = \tilde{A}^{-1} - \frac{w}{1 + w \cdot b^T \tilde{A}^{-1} b} \tilde{A}^{-1} b b^T \tilde{A}^{-1}.$$

We can thus obtain an expression for the Schur Complement:

$$B^T(\tilde{A} + wbb^T)^{-1} B = S - \frac{w}{1 + w \cdot b^T \tilde{A}^{-1} b} uu^T,$$

where $S = B^T \tilde{A}^{-1} B$ and $u = B^T \tilde{A}^{-1} b$. It thus follows that the Schur Complement of the perturbed matrix is merely a low-rank perturbation of the original Schur complement, the one associated with $\tilde{A}$.

Further observations can be made for the Schur complement associated with $\mathcal{H}(W)$. Consider the matrix $A + BWB^T$, and let

$$S(W) = B^T(A + BWB^T)^{-1} B. \tag{3.21}$$

For notational convenience, denote $S(0) = B^T A^{-1} B$ simply by S. Suppose that W is symmetric positive semidefinite. Using the Sherman-Morrison formula with $U = V = BW^{1/2}$ we have

$$\begin{aligned} S(W) &= B^T(A + BWB^T)^{-1} B \\ &= B^T \left(A^{-1} - A^{-1} B W^{1/2} (I + W^{1/2} B^T A^{-1} B W^{1/2})^{-1} W^{1/2} B^T A^{-1} \right) B \\ &= S - S W^{1/2} \left(I + W^{1/2} S W^{1/2} \right)^{-1} W^{1/2} S \\ &= S \cdot \left(I - W^{1/2} (I + W^{1/2} S W^{1/2})^{-1} W^{1/2} S \right). \end{aligned}$$

PROPOSITION 3.6. *If $W^{1/2}$ and S commute, then*

$$S(W) = S \cdot \left(I + W^{1/2} S W^{1/2} \right)^{-1}. \tag{3.22}$$

Proof. If $W^{1/2}S = SW^{1/2}$ then

$$W^{1/2}S(I + W^{1/2}SW^{1/2}) = (I + W^{1/2}SW^{1/2})SW^{1/2}.$$

In this case,

$$\begin{aligned}
&\left(I - W^{1/2}(I + W^{1/2}SW^{1/2})^{-1}W^{1/2}S\right) \cdot (I + W^{1/2}SW^{1/2}) \\
&\quad = I + W^{1/2}SW^{1/2} - W^{1/2}(I + W^{1/2}SW^{1/2})^{-1}W^{1/2}S(I + W^{1/2}SW^{1/2}) \\
&\quad = I + W^{1/2}SW^{1/2} - W^{1/2}(I + W^{1/2}SW^{1/2})^{-1}(I + W^{1/2}SW^{1/2})SW^{1/2} \\
&\quad = I + W^{1/2}SW^{1/2} - W^{1/2}SW^{1/2} = I.
\end{aligned}$$

□

Proposition 3.6 leads to the following useful result.

PROPOSITION 3.7. *If S and $W^{1/2}$ commute, then S and $S(W)$ commute too.*

Proof. Express $S(W)$ as in (3.22) and $(I + W^{1/2}SW^{1/2})^{-1}$ as a Neumann series. □

Propositions 3.6 and 3.7 lead us to make the following conclusion with regard to the conditioning of the Schur complement.

THEOREM 3.8. *If S and $W^{1/2}$ commute, then the eigenvalues of $S(W)$ satisfy*

$$\lambda_i = \frac{\mu_i}{\mu_i\alpha_i + 1}, \tag{3.23}$$

where μ_i are the eigenvalues of S, and α_i are the eigenvalues of W. As a result, if S is positive definite and W is positive semidefinite, and

$$\kappa_2(S(W)) \leq \kappa_2(S). \tag{3.24}$$

Proof. Let λ be an eigenvalue of $S(W)$, μ an eigenvalue of S, and α an eigenvalue of W. From Propositions 3.6 and 3.7 it follows that those three matrices share the same eigenvectors, y, and $(\mu - \lambda\mu\alpha)y = \lambda y$. Hence (3.23) and (3.24) follow. □

One obvious case where the commuting condition holds is $W = \gamma I$. This leads to the following result.

COROLLARY 3.9. *Suppose A is symmetric positive semidefinite, and let $S(\gamma)$ be as defined in (3.22), with $W = \gamma I$. Let $S \equiv S(0) \equiv B^T A^{-1} B$ have eigenvalues $\{\mu_i\}$. Then $S(\gamma) = S\left(I + \gamma S\right)^{-1}$. Consequently the eigenvalues of $S(\gamma)$ are $\mu_i(\gamma) = \frac{\mu_i}{1+\gamma\mu_i}$, and $S(\gamma)$ is better conditioned than S for all $\gamma > 0$.*

We end this subsection with a remark that clarifies the limits of our analysis. Commuting of S and $W^{1/2}$ does not hold in most cases, including the case of a diagonal matrix W whose entries are not all equal to each other. Theorem 3.8 thus has a limited practical value. However, it is possible that even when the commuting condition does not hold, the results stated in the theorem still nearly hold, and the conditioning of $S(W)$ is better than that of S for most choices of W.

3.4. Regularization using a nonzero (2,2) block. Regularizing the system may remedy potential accuracy problems. Common regularization techniques are based on modifying the (originally zero) (2,2) block so as to obtain a better conditioned matrix. See for example [45, chap. 17] and [14].

We limit our discussion below to the case $d \equiv 0$ in (1.1) and a positive definite (1,1) block. Throughout this section we attach subscripts to the matrices associated with the regularization. Let the regularized linear system be

$$\mathcal{H}_V u_V \equiv \begin{pmatrix} A & B \\ B^T & -V \end{pmatrix} \begin{pmatrix} x_V \\ y_V \end{pmatrix} = \begin{pmatrix} c \\ 0 \end{pmatrix}. \tag{3.25}$$

Let $S_V = V + B^T A^{-1} B$. S_0 is the Schur complement of $\mathcal{H} \equiv \mathcal{H}_0$. For $\mathcal{H}^{-1} = \begin{pmatrix} X & Y \\ Y^T & Z \end{pmatrix}$. we have [26]

$$X = A^{-1} - A^{-1}B(B^T A^{-1} B)^{-1} B^T A^{-1} = A^{-1} - A^{-1} B S_0^{-1} B^T A^{-1}.$$

Since A is assumed to be symmetric and positive definite, by observing that

$$\tilde{P} = A^{-1/2} B (B^T A^{-1} B)^{-1} B^T A^{-1/2}$$

is a projection matrix, we can rewrite X in a simpler form: $X = A^{-1/2}(I - \tilde{P}^2)A^{-1/2}$. Since $x = Xc$, x can be expressed in a form that shows how far the unconstrained solution (namely $A^{-1}c$) is from the constrained one: $x = A^{-1}c - A^{-1} B S_0^{-1} B^T A^{-1} c$. From the last two equations a bound for $\|x\|$ can be obtained.

Consider now the regularized linear system (3.25). Let

$$\mathcal{H}_V^{-1} = \begin{pmatrix} X_V & Y_V \\ Y_V^T & Z_V \end{pmatrix}.$$

PROPOSITION 3.10. *The* $(1,1)$ *block of* $\mathcal{H}_V^{-1}$ *is given by*

$$X_V = A^{-1} - A^{-1} B S_V^{-1} B^T A^{-1}. \tag{3.26}$$

If $V = \mu I, \mu > 0$, *then*

$$X_V \equiv X_\mu = (A + \frac{1}{\mu} B B^T)^{-1}. \tag{3.27}$$

Proof. Write $\begin{matrix} A & B \\ B^T & -V \end{matrix} \quad \begin{matrix} X_V & Y_V \\ Y_V^T & Z_V \end{matrix} \quad = \quad \begin{matrix} I & 0 \\ 0 & I \end{matrix}$, and write down the equations for the $(1,1)$ and $(2,1)$ blocks:

$$\begin{cases} AX_V + BY_V^T = I; \\ B^T X_V - V Y_V^T = 0. \end{cases} \tag{3.28}$$

By substituting $B^T X_V = B^T A^{-1} - B^T A^{-1} B Y_V^T$ in the bottom equation of (3.28) we get $Y_V^T = S_V^{-1} B^T A^{-1}$, from which (3.26) follows.

If $V = \mu I$, we can instead proceed from (3.28) by substituting $Y_\mu^T = \frac{1}{\mu} B^T X_\mu$ in the top equation of (3.28), and (3.27) readily follows. The matrix to be inverted in that equation is guaranteed to be nonsingular, because $\mathcal{H}$ is. ∎

We note that interestingly, (3.27) is not defined for $\mu = 0$, yet the limit exists. The error in the computed solution is given by $u_V - u = (X_V - X)c$. From Proposition 3.10 it follows that

$$\begin{aligned} X_V - X &= (A^{-1} - A^{-1} B S_V^{-1} B^T A^{-1}) - (A^{-1} - A^{-1} B S_0^{-1} B^T A^{-1}) \\ &= A^{-1} B \left(S_0^{-1} - S_V^{-1} \right) B^T A^{-1}. \end{aligned}$$

In general, for any two nonsingular matrices M and N, $M^{-1} - N^{-1} = M^{-1}(N - M)N^{-1}$. Hence $S_0^{-1} - S_V^{-1} = S_0^{-1}VS_V^{-1}$, and $X_V - X = -A^{-1}BS_V^{-1}VS_0^{-1}B^TA^{-1}$. Therefore $x_V - x = -A^{-1}BS_V^{-1}VS_0^{-1}B^TA^{-1}c$, and the norm of the error is bounded by

$$\|x_\mu - x\| = \|A^{-1}BS_V^{-1}VS_0^{-1}B^TA^{-1}c\| \le \|V\|\|S_V^{-1}\| \cdot \|B^TA^{-1}\|^2\|S_0^{-1}\| \cdot \|c\|.$$

All the quantities in the expression for the error except $\|V\|\|S_V^{-1}\|$, depend on the given problem, not on the regularization parameter. The latter can be determined so as to make the error as small as desired.

PROPOSITION 3.11. *For $V = \mu I$, $\|V\|\|S_V^{-1}\| \equiv \|\mu S_\mu^{-1}\|$ tends to 0 for $\mu \ll 1$, is monotonically increasing as μ grows larger, and is bounded by 1.*

Proof. Suppose $S_V \equiv S_\mu = \mu I + B^TA^{-1}B$. Then $\mu S_\mu^{-1} = \left(I + \frac{1}{\mu}S_0\right)^{-1}$. Thus μS_μ^{-1} is positive definite and

$$\|\mu S_\mu^{-1}\|_2 = \frac{1}{\lambda_{\min}(I + \frac{1}{\mu}S_0)} = \frac{\mu}{\mu + \lambda_{\min}(S_0)}. \tag{3.29}$$

The latter is a monotonic function in μ, that tends to 0 when $\mu \to 0$, and tends to 1 as $\mu \to \infty$. ∎

We end this subsection by noting that from the error equation we can obtain

$$B^T(x_\mu - x) = B^TA^{-1}B\mu S_\mu^{-1}S_0^{-1}B^TA^{-1}c.$$

In general, for any matrix M and a scalar μ, we have $M(\mu I + M)^{-1} = (\mu I + M)^{-1}M$. Thus S_0 and S_μ^{-1} commute, and as a result we have $B^T(x_\mu - x) = \mu S_\mu^{-1}B^TA^{-1}c$. Since $B^Tx = 0$, we can also say that $B^Tx_\mu = \mu S_\mu^{-1}B^TA^{-1}c$. While the un-regularized solution component x belongs to the null-space of B^T, the perturbed vector x_μ does not, and this relation shows its distance from the null-space.

4. Numerical examples. Our experiments were done using MATLAB, on a Linux based machine with a 2.5 GHz processor. Most of the matrices (with self generated right-hand side vectors) are taken from the CUTE/CUTEr set of optimization problems [29]. However, since our interest is in numerical linear algebra aspects of the problems, we ran the experiments for the matrices as given, without reducing them as is typically done in quadratic programming, in accordance with the active and inactive constraints.

Example 4.1. The `genhs28` problem from the CUTE collection is tiny, a 28×28 saddle point matrix, where A is 10×10. According to the CUTEr documentation , this is a multi-dimensional extension of problem 28 in [37]. $A = \text{tri}[2, 4, 2]$, but with $A_{1,1} = A_{10,10} = 2$, so that the matrix is singular. B is 10×8 with values $1, 2$ and 3 along its main diagonal, first subdiagonal, and second subdiagonal, respectively. We use this example because it provides analytical knowledge that is typically not available for larger problems. In particular, A has nullity 1, and its spectrum and null vector are known.

The eigenvalues of A are

$$\lambda_i = 8\sin^2\left(\frac{\pi j}{2n}\right), \quad j = 0, \ldots, n-1.$$

Here $n = 10$ is the size of the matrix, and so $\lambda_{\min} = \lambda_{10} = 0$, $\lambda_9 = 0.1958...$, and $\lambda_{max} = \lambda_1 = 7.8042....$

The null vector of A is $v = \alpha \cdot (-1, 1, -1, 1, -1, 1, -1, 1, -1, 1)^T$. It follows that the eight columns of B form only two different angles with the null vector of A. Figure 4.1 shows that the condition number of the perturbation of the matrix A by its null vector, and the condition number of the matrix obtained by perturbing A by the fourth (arbitrary choice) column of B behave qualitatively in the same manner. The perturbation with the null vector results in a smaller condition number, yet the perturbation with a column of B maintains sparsity. For the latter, only two additional nonzero entries are generated in comparison with the nonzero structure of A. Note that the condition number of $A + wvv^T$ can be obtained analytically by the result stated in Proposition 2.6.

It is also worth noting that the range of values of w for which a nearly optimal condition number is obtained is fairly large and both $\min \kappa_2(A+wvv^T)$ and $\min \kappa_2(A+wbb^T)$ are small (recall that A is singular). The choice $w = \|A\|/\|B\|^2$ gives $\log_{10} w = -0.6367$, which is well within the range of values for which the condition numbers in the plot are close to the minimum.

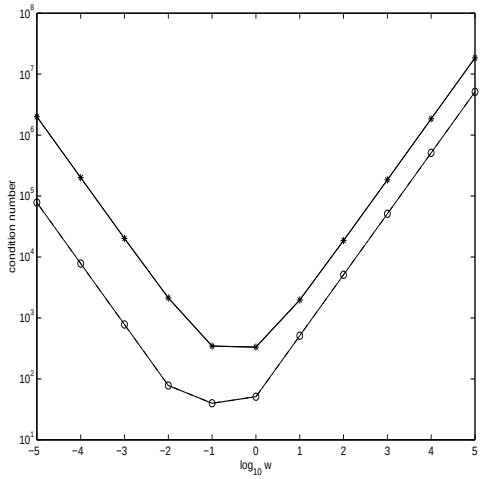


FIG. 4.1. *Example 4.1: Condition numbers of $A + wvv^T$ (curve with circles) and $A + wbb^T$ (starred curve), where v is the null vector of A and b is the fourth column of B. The matrix is* `genhs28` *from the CUTE collection.*

Example 4.2. The problem `cvxqp1` from the CUTE collection is a convex quadratic problem with linear constraints. We used here a medium sized version of the problem. (In [43] it is denoted by cvcqp1_m.) As is evident from the unusual nonzero pattern of the matrix (Figure 4.2), direct solution methods may have difficulties dealing with fill-in. A is 1000×1000, and its rank is 986. B is 1000×500. The smallest *positive* eigenvalue of A was computed to be 9.4756×10^{-7}. The rank of the 1500×1500 saddle point matrix is 1499 according to MATLAB's SVD-based `rank` command. This means that $A + BWB^T$ must have a nullity of at least 1. We generate the right-hand side vector so that the linear system is consistent.

Figure 4.3(a) provides support to the analysis of Section 2. We performed the following procedure: computed the 14 null vectors of A; computed the 14 columns of B, b_k, that form the smallest angles with the null vectors, and chose W to be a diagonal matrix whose nonzero entries correspond to those columns. In order to keep

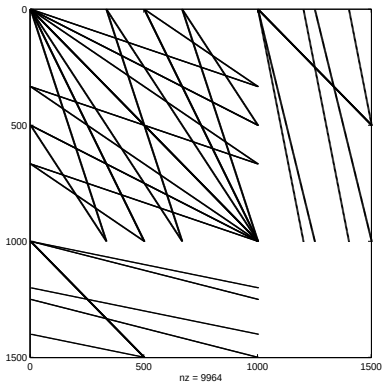
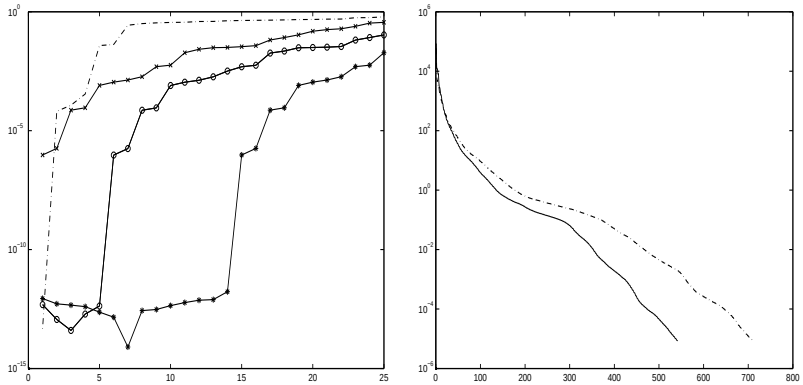


FIG. 4.2. *Example 4.2: Sparsity pattern of the* `cvxqp1` *matrix.*



(a) Eigenvalues (b) MINRES convergence

FIG. 4.3. *Example 4.2 — the CUTE matrix* `cvxqp1`. *Graph (a) depicts the 25 smallest eigenvalues of the following matrices, from bottom to top:* $A, A + BWB^T, A + VV^T, A + BB^T$. *Graph (b) shows the convergence of the preconditioned MINRES with the positive definite preconditioners* $\mathcal{M}_P(\mathrm{diag}(A))$ *(solid line) and* $\mathcal{M}_P(\mathrm{diag}(A + 156.13BB^T))$ *(broken line).*

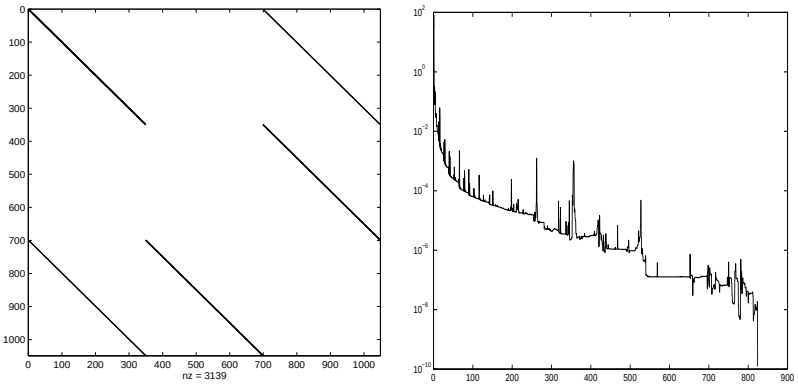
the computational cost of setting W small, the values of the nonzero entries of W are set to be the scaling factors $w_k \approx \|A\|_2/\|b_k\|_2^2$, obtained by using norm estimators.

The number of nonzeros of A is 6968, the number of nonzeros of $A + BWB^T$ is 7040, and the number of nonzeros of $A + \gamma BB^T$ is 9950. We denote by V the dense matrix of the 14 null vectors of A. As is evident from Fig. 4.3, the perturbation of A by the matrix of the null vectors or matrices associated with columns of B is an effective way of mapping zero eigenvalues of A away from zero. $A + VV^T$ is nonsingular, though dense. On the other hand, the matrix $A + BB^T$ is sparse and its nullity is 1, which is the best we can obtain given the rank of the saddle point matrix. Hence, the matrix $A + BWB^T$ is close in sparsity to A, but its nullity is substantially smaller.

Figure 4.3(b) provides experimental support to the viability of the block positive definite preconditioner analyzed in Section 3. The graph shows the convergence history of the preconditioned MINRES with the preconditioner $\mathcal{M}_P(M)$, for two choices of M. The broken curve represents the preconditioned iteration for $\mathcal{M}_P(\mathrm{diag}(A))$. (While A is singular, its diagonal is not.) The solid curve represents the iteration

for $\mathcal{M}_P(\mathrm{diag}(A+\gamma BB^T))$, where $\gamma = 156.13 \approx \|A\|_2/\|B\|_2^2$, obtained by using the norm estimator implemented in MATLAB's `normest`. The saving in iterations when the latter is used is approximately 40%, while the cost of every matrix-vector product is only marginally higher. For a random right-hand side we obtained a count of 710 iterations vs. 541.

Example 4.3. The problem `gouldqp2` is taken from the CUTE collection. The size of A is 699. The top left 349x349 block of A corresponds to a discrete one-dimensional Laplace operator with Neumann boundary conditions, and the rest of the matrix is equal to 0. B corresponds to a scaled first-order one-sided difference operator approximating the grad operator. For this problem we executed BiCG-STAB with the constraint preconditioner $\mathcal{M}_C(M)$. We used the solver with three choices of M: $M = \mathrm{diag}(A)$, $M = I$, and $M = \mathrm{diag}(A+BB^T)$. Of the three, only the latter resulted in convergence within the specified maximum iteration count of 2000 and relative residual threshold of 10^{-10}; see Figure 4.4(b). Using the diagonal of A is not effective possibly because of the many zero diagonal elements. $\mathrm{diag}(A+BB^T)$ is nonsingular, and it is seemingly more strongly related to the problem than the the choice $M = I$. While the situation described here may not necessarily be representative of what happens consistently in other problems, it shows that using $A + BB^T$ as a basis for the constraint preconditioner can be effective when other ways of selecting a (1,1) block are not obvious.



(a) Sparsity pattern (b) Convergence of BiCGStab

FIG. 4.4. *Example 4.3 — the CUTE matrix* `gouldqp2`. *Graph (a) depicts the sparsity pattern of the matrix. Graph (b) shows the convergence behavior of the preconditioned BiCG-Stab method with the constraint preconditioner* $\mathcal{M}_C(\mathrm{diag}(A+BB^T))$. *The right-hand side was constructed so that the solution is a vector of all 1s. Constraint preconditioning with* $\mathcal{M}_C(\mathrm{diag}(A))$ *or* $\mathcal{M}_C(I)$ *did not converge within the specified maximum iteration count and tolerance.*

Example 4.4. The CUTE matrix `gouldqp3` is 1048×1048 and its sparsity pattern is given in Figure 4.5. The (1,1) block A is 699×699 and its rank is 697. The smallest positive eigenvalue of A is $4.0514 \cdot 10^{-5}$, and its largest eigenvalue is 5.2360. According to the CUTEr documentation, this is a variant of a problem of optimal knot placement in a scheme for ordinary differential equations with boundary values, suggested by J. R. Kightley [27]. The null-space of A is comprised of two orthogonal vectors: a vector of all 1s (or any other nonzero constant) except the last entry, which is zero, and a vector of all 0s except the last entry, which is nonzero. Following the strategy of Section 2, we find columns of B that are aligned with the null vectors of A. The last column of B must be used, and for the other null vector there are many

choices. We picked the first and the last columns of B; let us denote them by B_1 and B_m respectively. We define $V = w_1 B_1 + w_2 B_m$, and set $M_1 = A + VV^T$, and set $M_2 = A + BB^T$. 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.

Each of the two zero eigenvalues of the matrix A is mapped into -1 and 1 in the preconditioned matrices $\mathcal{M}(M_i)^{-1}\mathcal{H}, i = 1, 2$. In both cases, as predicted by our analysis, the eigenvalues are within the interval stated in Theorem 3.5. We ran MINRES with the above two preconditioners, and then ran MINRES again with $\mathcal{M}(\text{ILU}(M_i, 0))^{-1}\mathcal{H}, i = 1, 2$. Here $\text{ILU}(M_i, 0)$ stands for the product LU of the factors of the incomplete factorization ILU(0) of M_i. In other words, instead of solving for M_1 and M_2 in each iteration, we use their incomplete factorizations. For those factorization the analysis in Section 3 does not hold. However, as can be seen in Figure 4.6, most of the eigenvalues are still inside the intervals stated in Theorem 3.5. The convergence behavior for the four preconditioners, as well as the eigenvalues of the preconditioned matrices, are given in Figures 4.6 and 4.7. In terms of timing, for the first two preconditioners an un-optimized version of the solver ran 1.33 seconds using M_1 vs. 4.57 for M_2. However, for the former the stage of computing the null vectors of A (which in this case are known and are not incorporated into the computation stage) may significantly increase the computation time.

$\text{nnz}(A)$	2092
$\text{nnz}(B)$	1047
$\text{nnz}(A + VV^T)$	2105
$\text{nnz}(A + BB^T)$	4185

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

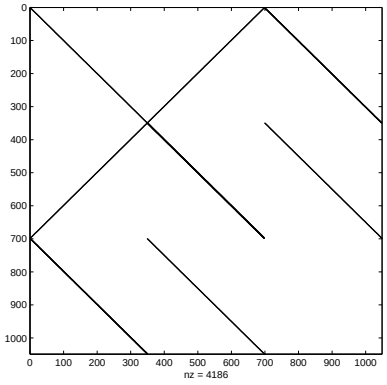


FIG. 4.5. *Sparsity pattern of the* `gouldqp3` *matrix*

.

Example 4.5. The following nonlinear geophysical inverse problem is described in [33]:

$$\min \phi(u, m) = \frac{1}{2}\|Qu - b\|^2 + \frac{\beta}{2}\|W(m - m_0)\|^2$$

$$\text{subject to} \quad A(m)u = f.$$

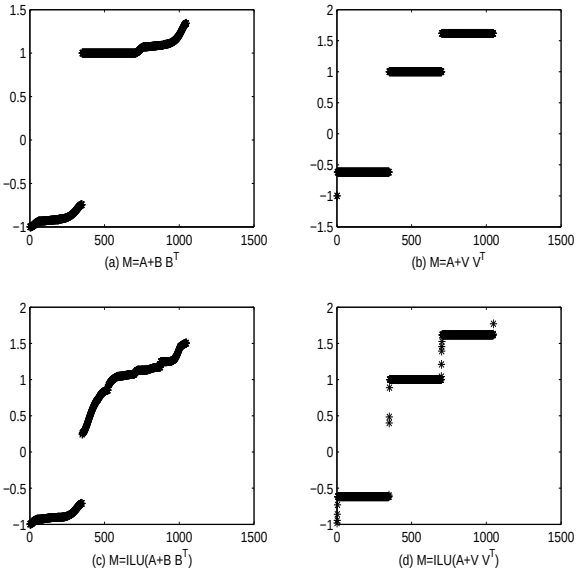


FIG. 4.6. *Example 4.4: Eigenvalues of the preconditioned matrix* $\mathcal{M_C}^{-1}\mathcal{H}$ *associated with the CUTE matrix* `gouldqp3`. *Given are four different choices, see x-label below each of the graphs and further details in the text.*

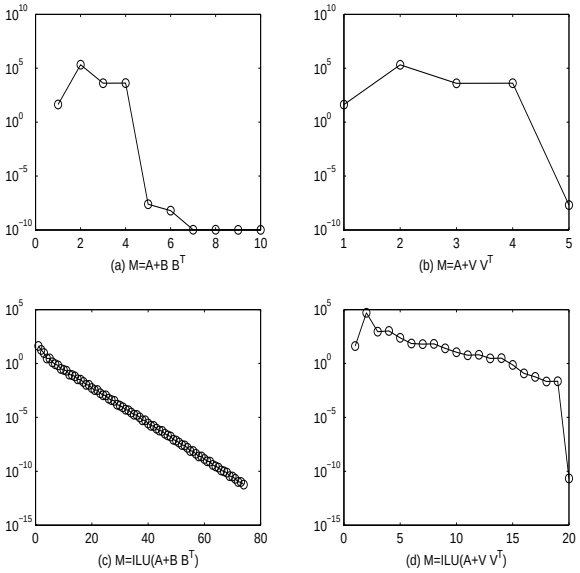


FIG. 4.7. *Convergence of preconditioned MINRES for the CUTE problem* `gouldqp3`.

The associated Lagrangian is

$$\mathcal{L}(u, m, \lambda) = \frac{1}{2}\|Qu - b\|^2 + \frac{\beta}{2}\|W(m - m_0)\|^2 + \lambda^T[A(m)u - f],$$

where λ is a vector of Lagrange multipliers. The Hessian of the Lagrangian is

$$H(m,u,\lambda) = \begin{pmatrix} Q^TQ & K(m,\lambda)^T & A(m)^T \\ K(m,\lambda) & \beta W^TW + R(m,u,\lambda) & G(m,u)^T \\ A(m) & G(m,u) & 0 \end{pmatrix},$$

and its sparsity pattern is depicted in Figure 4.8, leftmost graph. Details on the application in which this problem arises, as well as solution methods, are described in [33].

Here we have a 3×3 block matrix, and we define A to be the top left 2×2 part of $H(m,u,\lambda)$. Replacing A by $A + BB^T$ generates a modified saddle point matrix significantly less sparse than $H(m,u,\lambda)$: compare the leftmost plot of Fig. 4.8 to the rightmost plot. [We note, though, that a naive count of the nonzeros of the matrices involved does not do justice to the actual cost of forming matrix-vector products since the product $(A + \gamma BB^T)x$ can be performed as $Ax + (\gamma B(B^Tx))$.]

Sparsity is lost, but on the positive side, the (1,1) block is now positive definite, and the eigenvalues of the preconditioned matrix associated with $\mathcal{M}_P(A+BB^T)^{-1}H$ are clustered as predicted by our analysis, see Fig. 4.9. Since A has so many zero rows and columns, finding a matrix W such that $A + BWB^T$ is as sparse as possible but positive definite, in the spirit of Section 2, is practically impossible in this case.

Applying the constraint preconditioner $\mathcal{M}_C(\text{diag}(A)$ proves ineffective. In Fig. 4.10(a) the convergence behavior of GMRES(20) with this preconditioner is depicted; convergence occurs after more than 250 iterations. Attempts to use $\mathcal{M}_C(\text{diag}(A + \gamma BB^T))$ instead (for a variety of values of γ) proves futile as well.

What seems to be a more effective approach in this case is adding the third block row to the first one:

$$\tilde{H} = \begin{pmatrix} Q^TQ + A(m) & K(m,\lambda)^T + G(m,u) & A(m)^T \\ K(m,\lambda) & \beta W^TW + R(m,u,\lambda) & G(m,u)^T \\ A(m) & G(m,u) & 0 \end{pmatrix}.$$

The sparsity pattern of $\tilde{H}$ is depicted in the middle plot in Fig. 4.8. Note that $\tilde{H}$ is now *nonsymmetric*, but as we can see, the loss of sparsity is much smaller compared to the loss of sparsity with forming $A+BB^T$. Denoting the top left 2×2 part of $\tilde{H}$ by $\tilde{A}$ and its (1,3) and (2,3) blocks by $\tilde{B}$, we were able to compute an effective incomplete factorization of the (nonsymmetric) (1,1) block, and we incorporated it in a constraint preconditioner of the form $\mathcal{M}_C(\text{ILU}(\tilde{A} + \gamma\tilde{B}\tilde{B}^T))$, to obtain a nonsymmetric preconditioner. As we can see in Fig. 4.10(b), the convergence of preconditioned GMRES in this case is significantly faster compared to the previously discussed approaches.

5. Conclusions. The focus of this paper is on the solution of (1.1) by numerically solving the mathematically equivalent system (1.2), given that A is positive semidefinite. The technique is a generalization of the classical augmented Lagrangian approach, where rather than allowing for a single parameter, our formulation allows for determining up to $m(m+1)/2$ free parameters (assuming that W is to be kept symmetric). It is desirable to have W as sparse as possible in order to maintain sparsity of the (1,1) block in (1.2), and at the same time obtain positive definiteness and other desirable numerical properties. In Section 2 we discussed an approach that minimizes the condition number of $A+BWB^T$ where W is diagonal and nonnegative. We derived analytical results that led us to conclude that columns B that are aligned with the null vectors of A should be considered in the selection of nonzero entries of

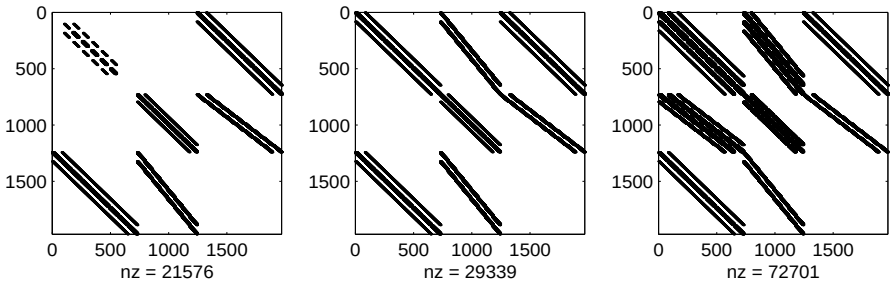


FIG. 4.8. *Sparsity patterns of the Hessian in geophysical inverse problem (left), and the two matrices discussed in this example. The middle (nonsymmetric) matrix in the figure is the one obtained by adding the third block row to the first one. The matrix on the right is* $A + BB^T$.

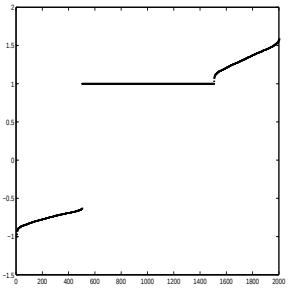


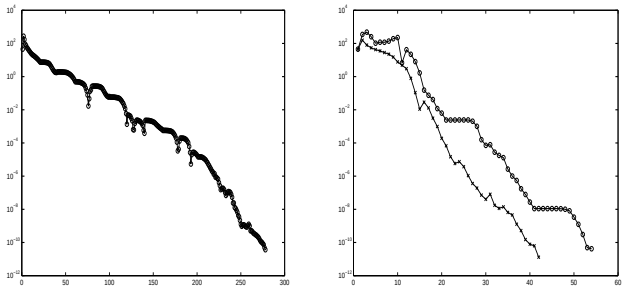
FIG. 4.9. *Eigenvalues of preconditioned matrix for the geophysical inverse problem.*

W. The analysis was derived for A with nullity 1, and we used a secant method for determining the values of the diagonal entries of W.

A drawback of this approach is that it requires computing the null-space of A, a computationally nontrivial task for large scale problems. Thus, our approach works best when knowledge on the null-space of A is available. However, even if such knowledge is not available, there are a few possible ways to keep the computational cost low. This point remains an issue for further investigation. Alternatively, it is possible to perform a 'blind' choice of several columns of B, using a scaling parameter of the form $w_i = \|A\|/\|b_i\|^2$. This involves a minimal computational effort when computationally cheap norm estimators are used. The effectiveness of using scaling factors of this form seems experimentally solid, although we do not have a full analytical justification.

Section 3 focused the application of block diagonal positive definite preconditioners whose (1,1) block is a matrix of the form $A + BWB^T$, and whose (2,2) block is the associated Schur complement. Our analysis shows that the eigenvalues of the preconditioned matrix are clustered in fixed intervals whose ends are far from the origin. Moreover, the eigenvalue 1 of the preconditioned matrix is of algebraic multiplicity at least $n - m$, and interestingly, the multiplicity goes up with the nullity of A. In order to get a complete picture of the convergence properties of Krylov solvers applied with this preconditioner, further analysis in the spirit of [39] is required, that reveals the structure of the Jordan blocks.

In practice, the matrix $A+BWB^T$ is replaced by an approximation to it, reducing the clustering but possibly making the numerical procedure significantly cheaper. Our numerical experiments demonstrate the effectiveness of this approach. With a sensible



(a) $\mathcal{M}_C(\mathrm{diag}(A)$ (b) $\mathcal{M}_C(\mathrm{diag}(\tilde{A}+\gamma\tilde{B}\tilde{B}^T)))$

FIG. 4.10. *Example 4.5: Convergence of preconditioned GMRES(20) using constraint preconditioning. The constants in graph (b) are 1 in the slower to converge case (curve with circles), and* 10^{-4} *in the faster case (curve with stars). The latter is approximately equal to the scaling value of* $\|A\|/\|B\|^2$.

choice of the columns of B, convergence does not seem to slow down in comparison with approaches that do not rely on setting $A + BWB^T$, and at the same time the advantage of making the (1,1) block positive definite is achieved.

Plans for future work include comprehensive testing of linear systems arising from interior-point optimization problems and investigation of the effectiveness of our approach in nonsymmetric cases and in cases of an indefinite (1,1) block.

Acknowledgments. Hearty thanks go to Eldad Haber for his insightful observations on Section 2 and his code and input on Example 4.5. We are indebted to Michael Friedlander for providing us with an easy-to-use interface to the CUTE test collection, along with a large set of test problems. Many thanks to Rich Lehoucq for pointing out reference [7] and its relevance to Section 2, and to Jorge Nocedal for stimulating discussions while he visited the University of British Columbia and for pointing out reference [5].

REFERENCES

[1] K. J. Arrow and R. M. Solow. *Gradient methods for constrained maxima, with weakened assumptions*. Stanford University Press, Stanford, California, 1958.
[2] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, Cambridge, 1994.
[3] M. Benzi, M. J. Gander, and G. H. Golub. Optimization of the Hermitian and Skew-Hermitian splitting iteration for saddle-point problem. *BIT*, 43:881–900, 2003.
[4] M. Benzi and G. H. Golub. A preconditioner for generalized saddle point problems. *SIAM J. Matrix Anal. Appl., to appear*.
[5] L. Bergamaschi, J. Gondzio, and G. Zilli. Preconditioning indefinite systems in interior point methods for optimization. *Computational Optimization and Applications, accepted for publication*, 2003.
[6] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, 1982.
[7] P. Bochev and R. Lehoucq. On finite element discretizations of the pure Neumann problem. *SIAM Review (accepted), and Technical report SAND2001-0733J*, 2001.
[8] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, Berlin, 1991.
[9] E. de Sturler and J. Liesen. Block-diagonal preconditioners for indefinite linear algebraic systems, part I: Theory. *Technical report UIUCDCS-R-2002-2279 or UILU-ENG-2002-1723*, 2002.
[10] G. Debreu. Definite and semidefinite quadratic forms. *Econometrica*, 20(2):295–300, April 1952.

[11] J. W. Demmel. *Applied Numerical Linear Algebra*. SIAM, 1997.
[12] N. Dyn and W. E. Ferguson. The numerical solution of equality-constrainted quadratic programming problems. *Math. Comput.*, 41:165–170, 1983.
[13] H. C. Elman. Preconditioning for the steady-state Navier-Stokes equations with low viscosity. *SIAM J. Sci. Comput.*, 20:1299–1316, 1999.
[14] H. C. Elman and D. Silvester. Fast nonsymmetric iterations and preconditioning for Navier-Stokes equations. *SIAM J. Sci. Comput.*, 17:33–46, 1996.
[15] H. C. Elman, D. J. Silvester, and A. J. Wathen. Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations. *Numer. Math.*, 90:641–664, 2002.
[16] B. Fischer, A. Ramage, D. J. Silvester, and A. J. Wathen. Minimum residual methods for augmented systems. *BIT*, 38:527–543, 1998.
[17] R. Fletcher. An ideal penalty function for constrained optimization. *AERE Report CSS 2*, 1973.
[18] R. Fletcher. *Practical Methods of Optimization*. 2nd Edition, Wiley, Chichster, 1987.
[19] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*. North-Holland, 1983.
[20] M. P. Friedlander. A Globally Convergent Linearly Constrained Lagrangian Method for Nonlinearly Constrained Optimization. *PhD thesis, Stanford University*, 2002.
[21] F. R. Gantmacher. *The Theory of Matrices*. Chelsea, 1959.
[22] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM, Philadelphia, 1989.
[23] G. H. Golub and C. Greif. Techniques for solving general KKT systems. *SCCM Technical report, number 00-05*, 2000.
[24] G. H. Golub and C. Greif. On solving block structured indefinite linear systems. *SIAM J. Sci. Comput.*, 24(6):2076–2092, 2003.
[25] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Third Edition, Johns Hopkins University Press, Baltimore, MD, 1996.
[26] G. H. Golub and A. J. Wathen. An iteration for indefinite systems and its application to the Navier-Stokes equations. *SIAM J. Sci. Comput.*, 19:530–539, 1998.
[27] N. I. M. Gould. An algorithm for large-scale quadratic programming. *IMA J. Num. Anal*, 11:299–324, problem class 3, 1991.
[28] N. I. M. Gould, M. E. Hribar, and J. Nocedal. On the solution of equality constrained quadratic programming problems arising in optimization. *Technical Report RAL-TR-1998-069, Rutherford Appleton Laboratory, UK*, 1998.
[29] 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/*.
[30] A. Greenbaum and G. H. Rodrigue. Optimal preconditioners for a given sparsity pattern. *BIT*, 29:610–634, 1990.
[31] M. Gu and S. C. Eisenstat. A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem. *SIAM J. Matrix Anal. Appl*, 15(4):1266–1276, 1994.
[32] M. Gu and S. C. Eisenstat. A divide-and-conquer algorithm for the symmetric tridiagonal eigenproblem. *SIAM J. Matrix Anal. Appl*, 16(1):172–191, 1995.
[33] E. Haber, U. M. Ascher, and D. Oldenburg. On optimization techniques for solving nonlinear inverse problems. *Inverse Problems*, 16(5):1263–1280, 2000.
[34] M. T. Heath. *Scientific Computing, An Introductory Survey, Second Edition*. McGraw-Hill, 2002.
[35] M. Hestenes. Multiplier and gradient methods. *J. Opt. Theory Applic.b*, 4:303–320, 1969.
[36] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1996.
[37] W. Hock and K. Schittkowski. *Test examples for nonlinear programming codes, Lectures Notes in Economics and Mathematical Systems 187*. Springer-Verlag, Heidelberg, 1981.
[38] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
[39] C. Keller, N. I. M. Gould, and A. J. Wathen. Constraint preconditioning for indefinite linear systems. *SIAM J. Matrix Anal. Appl.*, 4:1300–1317, 2000.
[40] A. Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(2):540–552, 1998.
[41] P. Krzyzanowski. On block preconditioners for nonsymmetric saddle point problems. *SIAM J. Sci. Comput.*, 23(1):157–169, 2001.
[42] L. Lukšan and J. Vlček. Indefinitely preconditioned inexact Newton method for large sparse equality constrained nonlinear programming problems. *Num. Lin. Algebra Appl.*, 5:219–247, 1998.
[43] I. Maros and C. Meszaros. A repository of convex quadratic programming problems. *Technical*

report, Department of Computing, Imperial College, London, 1997.

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

[45] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 1999.

[46] M. L. Overton. On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM J. Marix Anal. Appl.*, 9:256–268, 1988.

[47] M. L. Overton. Large-scale optimization of eigenvalues. *SIAM J. Optimization*, 2:88–120, 1992.

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

[49] B. N. Parlett. *The Symmetric Eigenvalue Problem*. Prentice Hall, Englewood Cliffs, N.J., 1980.

[50] I. Perugia and V. Simoncini. Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations. *Num. Lin. Alg. Appl.*, 7:585–616, 2000.

[51] M. J. D. Powell. *A method for nonlinear constraints in minimization problems*. Academic Press, 1969.

[52] M. Rozloznik and V. Simoncini. Krylov subspace methods for saddle point problems with indefinite preconditioning. *SIAM J. Matrix Analysis and Appl.*, 24(2):368–391, 2002.

[53] T. Rusten and R. Winther. A preconditioned iterative method for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 13(3):887–904, 1992.

[54] V. Simoncini and M. Benzi. Spectral properties of the hermitian and skew-hermitian splitting preconditioner for saddle point problems. *SIAM J. Matrix Anal. Appl., to appear*.

[55] H. Van Der Vorst. *Iterative Krylov Methods for Large Linear Systems*. Cambridge University Press, 2003.

[56] A. J. Wathen, B. Fischer, and D. J. Silvester. The convergence of iterative solution methods for symmetric and indefinite linear systems. *in 'Numerical Analysis 1997' (Dundee), Pitman Research Notes in Mathematics Series, Addison Wesley Longman, Harlow, England*, pages 230–243, 1997.

[57] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.