SIAM J. Matrix Anal. Appl.
Vol. 36, No. 2, pp. 367–384

ON NONSINGULAR SADDLE-POINT SYSTEMS WITH A MAXIMALLY RANK DEFICIENT LEADING BLOCK*

RON ESTRIN† AND CHEN GREIF†

Abstract. We consider nonsingular saddle-point matrices whose leading block is maximally rank deficient, and show that the inverse in this case has unique mathematical properties. We then develop a class of indefinite block preconditioners that rely on approximating the null space of the leading block. The preconditioned matrix is a product of two indefinite matrices, but under certain conditions the conjugate gradient method can be applied and is rapidly convergent. Spectral properties of the preconditioners are observed and validated by numerical experiments.

Key words. saddle-point systems, null space, approximate inverse, preconditioning, iterative methods

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

DOI. 10.1137/140989996

1. Introduction.

Consider the saddle-point linear system

$$\underbrace{\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}}_{\mathcal{K}} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$ with $m < n$ (the case $m = n$ is also covered by our analysis but in practice it is not of much interest), $u, f \in \mathbb{R}^n$, and $p, g \in \mathbb{R}^m$. We will mostly focus on the case of a symmetric positive semidefinite matrix A, though some of our analysis is applicable or can be generalized to any A. We denote the block 2×2 matrix of (1.1) by $\mathcal{K}$ and assume that it is nonsingular, which implies that B has full row rank.

Saddle-point systems are prominently featured in numerous applications, including computational fluid dynamics, constrained optimization, electromagnetism, image processing, parameter identification, and inverse problems. See, for example, [1, section 2] for a comprehensive list of applications and a detailed description of some problems of interest. In the last couple of decades a plethora of numerical solution methods and preconditioning techniques for such systems have been proposed in a large body of literature [1, 5, 21, 22].

If A is symmetric positive semidefinite, then a necessary and sufficient condition for $\mathcal{K}$ to be nonsingular is $\ker(A) \cap \ker(B) = \{0\}$, and in fact it is sufficient to require that A be positive (or negative) definite on $\ker(B)$; see [1, Theorem 3.2]. If A is a general matrix, sufficient conditions for the nonsingularity of $\mathcal{K}$ are a bit more involved and concern the intersection of the symmetric part of A with the null space of B; see [1, Theorem 3.4].

*Received by the editors October 3, 2014; accepted for publication (in revised form) by S. Le Borne January 28, 2015; published electronically April 2, 2015.
http://www.siam.org/journals/simax/36-2/98999.html
†Department of Computer Science, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (ronestrin756@gmail.com, greif@cs.ubc.ca). The research of the second author was supported in part by Discovery grant 261539 from the Natural Sciences and Engineering Research Council of Canada (NSERC).

From the above it readily follows that if $\mathcal{K}$ is nonsingular, the rank of A must be at least $n - m$. We can thus say that if the nullity of A, namely the dimension of its null space, satisfies

$$\text{null}(A) = m,$$

then A is *maximally rank deficient*, and we mean it in the sense that a higher nullity than m necessarily implies that $\mathcal{K}$ is singular.

Since there is a trivial intersection of the null spaces of A and B, the dimensions of those spaces must add up to n, and we have the Helmholtz-type decomposition

$$\ker(A) \oplus \ker(B) = \mathbb{R}^n. \tag{1.2}$$

The maximal rank deficiency situation arises, for example, in the numerical solution of time-harmonic Maxwell's equations [12, 14, 18]. It may also occur in underdetermined norm-minimization problems and other applications; see section 2. This particular setting has not received nearly as much attention as other situations, for example, the case of a symmetric positive definite $(1, 1)$ block, but as we show, when the leading block is indeed maximally rank deficient, a rich mathematical structure can be revealed, which gives rise to potentially new solution methods.

In section 2 we give a few examples of linear systems of the type that is the focus of our paper. In section 3 we develop formulas for the inverse of $\mathcal{K}$. We show that in general, if the (1,1) block of $\mathcal{K}$, namely A, has nullity r, the rank of the (2,2) block of $\mathcal{K}^{-1}$ is equal to $m - r$ and therefore, remarkably, the (2,2) block of the inverse is zero if the leading block A is maximally rank deficient (i.e., has nullity $r = m$); that is, the inverse itself has a saddle-point structure. In section 4 we use the formulas of the preceding section to derive a new family of block preconditioners. Our approach requires the knowledge of the null space of A or a sparse approximation of it, and the practicality of the preconditioner depends on the application. In section 5 we numerically illustrate the merits of our preconditioning approach. In section 6 we discuss a few relevant issues, map out directions for future investigation, and draw some conclusions.

2. Examples. In this section we present a few examples of saddle-point systems that have a leading block with a high nullity, and we include details on the availability of the null space.

Time-harmonic Maxwell. The time-harmonic Maxwell equations with constant coefficients in lossless media with perfectly conducting boundaries are given by

$$\begin{aligned} \nabla \times \nabla \times u - k^2 u + \nabla p &= f \qquad \text{in } \Omega, \\ \nabla \cdot u &= 0 \qquad \text{in } \Omega, \\ u \times n &= 0 \qquad \text{on } \partial\Omega, \\ p &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Here, u is an electric vector field and p is a scalar multiplier. The wave number is given by $k^2 = \omega^2 \epsilon \mu$, where ω is the temporal frequency, and ϵ and μ are permittivity and permeability parameters. The case $k = 0$ is physically relevant and it gives rise to linear systems of the kind we focus on.

The Lagrange multiplier p is not physically meaningful in this setting, and in fact the divergence constraint does not need to be presented explicitly; it is possible to

decouple the unknowns and solve directly for u [14]. But the saddle-point formulation is stable and well-posed [4, section 3], and it gives rise to effective numerical solution methods.

After discretization, using Nédélec finite elements for the electric vector field and nodal elements for the multiplier [12, section 2.1], we have a saddle-point system with a discrete curl-curl operator as the leading block, A, and a discrete divergence operator, B. The matrix A satisfies $\text{null}(A) = m$ and its null space is explicitly available—it is the space of gradient functions. Taking the gradients of the basis functions of the Nédélec vector polynomials, within $H_0(\text{curl})$ (see [12, section 2.2]), allows for constructing a sparse matrix C of dimensions $n \times m$ that satisfies $AC = 0$.

Norm minimization and constrained optimization problems. Another example arises in norm minimization problems with equality constraints; see, for example, [2]. Consider the problem

$$\begin{aligned} &\min_x \frac{1}{2}\|Ex - b\|_2^2 \\ &\text{s.t.} \quad Fx = d. \end{aligned}$$

The Lagrangian is given by $\mathcal{L}(x, \lambda) = \frac{1}{2}\|Ex-b\|_2^2+\lambda^T(Fx-d)$, and the corresponding optimality conditions, obtained by equating the gradient of the Lagrangian to zero, are given by

$$\begin{pmatrix} E^TE & F^T \\ F & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} E^Tb \\ d \end{pmatrix}.$$

If $F \in \mathbb{R}^{m\times n}$, $E \in \mathbb{R}^{\ell\times n}$ with $\ell < n$ and it has full row rank, then E^TE is $n \times n$ and has rank ℓ. Therefore, in the case where $\ell = n - m$, the nullity of the (1,1) block is m and it is the maximal nullity that still allows for the saddle-point matrix to be nonsingular.

It is possible to rewrite the above problem in *augmented form*: introduce a residual variable $r = Ex - b$ and insert this relation into the set of block equations, to obtain a 3×3 block system of the form (see also [1, section 2.2])

$$\begin{pmatrix} I_\ell & 0 & E \\ 0 & 0 & F \\ E^T & F^T & 0 \end{pmatrix} \begin{pmatrix} r \\ \lambda \\ -x \end{pmatrix} = \begin{pmatrix} -b \\ -d \\ 0 \end{pmatrix}.$$

We can now think of $\left(\begin{smallmatrix} I_\ell & 0 \\ 0 & 0 \end{smallmatrix}\right)$ as the leading block of a standard saddle-point matrix. In the context of this paper, this generates a trade-off: Now the leading block is not maximally rank deficient, but its null space is trivial to construct.

Interior point methods for constrained optimization problems give rise to a potential set of saddle-point problems with a rank deficient (1,1) block, although here the nullity depends on the number of active constraints at the solution. As the iterations approach the solution, the barrier parameter is set closer to zero, and the 2×2 saddle-point formulation of the problem becomes increasingly ill-conditioned [10, 20]. In the linear programming case the (1,1) block is diagonal. During the iterations the diagonal matrix is nonsingular but it becomes increasingly ill-conditioned, with some of its entries tending to zero. Given the diagonality of the leading block, its "near null space" (i.e., the eigenvectors that correspond to the small eigenvalues, that is, the small diagonal entries) is easy to construct.

Geophysical inverse problems. Geophysical inverse problems often have a saddle-point structure with a high nullity of the leading block. A relevant example is described in [3, 13], and it concerns the inversion of electromagnetic data to recover three-dimensional distributions of physical properties from observed electric and magnetic fields. The model problem in regularized form is given by

$$\begin{aligned} &\min \frac{1}{2}\|P^T u - d\|^2_{W_d} + \frac{\alpha}{2}\|Lm\|^2 \\ &\text{s.t.} \qquad F(m)u + Gm = f, \end{aligned}$$

where m is a model, L is a regularization operator (typically a discretized second order differential operator), and the forward problem $F(m)u+Gm = f$ is typically a second order differential equation; G is an averaging operator. P is an observation matrix, and W_d is the standard deviation. If Gauss–Newton is used, the resulting saddle-point linear systems that need to be solved throughout the nonlinear iterations have the form

$$\begin{pmatrix} PW_dP^T & 0 & F^T \\ 0 & \alpha L^T L & G^T \\ F & G & 0 \end{pmatrix} \begin{pmatrix} \delta u \\ \delta m \\ \delta\lambda \end{pmatrix} = - \begin{pmatrix} r_u \\ r_m \\ r_\lambda \end{pmatrix}.$$

Strong undersampling means that many rows of P are zero, and it is easy to obtain the null-space of PW_dP^T in that case. The leading block is thus highly rank deficient, though its nullity is not maximal.

3. Formulas for the inverse of the saddle-point matrix. The survey paper [1] presents two well-known formulas for inverses of saddle-point matrices. One of them [1, equation (3.4)] assumes that A is nonsingular and uses Schur complements:

$$\mathcal{K}^{-1} = \begin{pmatrix} A^{-1} - A^{-1}B^T S^{-1} BA^{-1} & A^{-1}B^T S^{-1} \\ S^{-1}BA^{-1} & -S^{-1} \end{pmatrix}, \tag{3.1}$$

where $S = BA^{-1}B^T$. (The formula written here is slightly different than [1, equation (3.4)] because our S differs by sign from the analogous matrix S defined there.)

Another formula [1, equation (3.8)] does not require A to be invertible and resorts to using expressions related to using $Z^T AZ$, where Z is an $n \times (n-m)$ matrix whose columns form a basis for the null space of B:

$$\mathcal{K}^{-1} = \begin{pmatrix} V & (I - VA)\, B^T \left(BB^T\right)^{-1} \\ \left(BB^T\right)^{-1} B\,(I - AV) & -\left(BB^T\right)^{-1} B\,(A - AVA)\, B^T \left(BB^T\right)^{-1} \end{pmatrix}, \tag{3.2}$$

where

$$V = Z\left(Z^T AZ\right)^{-1} Z^T.$$

Remark 3.1. The result (3.2) can be easily generalized to *nonsymmetric* nonsingular saddle-point matrices of the form

$$\mathcal{K} = \begin{pmatrix} A & B_1^T \\ B_2 & 0 \end{pmatrix}.$$

The matrix V in this case takes on the slightly more involved form

$$V = Z_2\left(Z_1^T AZ_2\right)^{-1} Z_1^T,$$

where Z_1 and Z_2 are, respectively, the null-space matrices of B_1 and B_2, and the inverse is given by

$$\mathcal{K}^{-1} = \begin{pmatrix} V & (I - VA)\, B_1^T \left(B_2 B_1^T\right)^{-1} \\ \left(B_2 B_1^T\right)^{-1} B_2 \left(I - AV\right) & -\left(B_2 B_1^T\right)^{-1} B_2 \left(A - AVA\right) B_1^T \left(B_2 B_1^T\right)^{-1} \end{pmatrix}.$$

We now highlight a connection between the nullity of the (1,1) block of $\mathcal{K}$ and the rank of the (2,2) block of $\mathcal{K}^{-1}$.

THEOREM 3.2. *Let A satisfy $\mathrm{null}(A) = r \leq m$ and suppose $\mathcal{K}$ is invertible. Then if we write its inverse as*

$$\mathcal{K}^{-1} = \begin{pmatrix} * & * \\ * & X \end{pmatrix},$$

where $X \in \mathbb{R}^{m \times m}$, the rank of X is

$$\mathrm{rank}(X) = m - r.$$

Proof. Since $\mathcal{K}$ is invertible, $\ker(A) \cap \ker(B) = \{0\}$. Suppose $C \in \mathbb{R}^{n \times r}$ and $Z \in \mathbb{R}^{n \times (n-m)}$ are the matrices whose columns span the null spaces of A and B, respectively. Then $[C|Z] \in \mathbb{R}^{n \times (n-m+r)}$ has linearly independent columns. There exist $m-r$ additional vectors, which we collect into a matrix that we call F, such that $[C|Z|F] \in \mathbb{R}^{n \times n}$ forms a basis for $\mathbb{R}^n$. Thus, any vector $x \in \mathbb{R}^n$ can be represented as

$$x = [C|Z|F] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = Cx_1 + Zx_2 + Fx_3$$

for some $x_1 \in \mathbb{R}^r$, $x_2 \in \mathbb{R}^{n-m}$, $x_3 \in \mathbb{R}^{m-r}$.

From (3.2), we know that

$$X = -\left(BB^T\right)^{-1} B \left(A - AZ \left(Z^T AZ\right)^{-1} Z^T A\right) B^T (BB^T)^{-1}.$$

Since B has full row rank, all we need to show is that

$$\mathrm{rank}\left(A - AZ \left(Z^T AZ\right)^{-1} Z^T A\right) = m - r.$$

Any pair of matrices M and N of the same dimensions satisfies $\mathrm{rank}(M) \leq \mathrm{rank}(M - N) + \mathrm{rank}(N)$ [15, section 0.4.5(d)]. Therefore,

$$\mathrm{rank}(A) \leq \mathrm{rank}(A - AZ \left(Z^T AZ\right)^{-1} Z^T A) + \mathrm{rank}(AZ \left(Z^T AZ\right)^{-1} Z^T A)$$

and hence $n - r \leq \mathrm{rank}(A - AZ \left(Z^T AZ\right)^{-1} Z^T A) + (n - m)$. We therefore have $\mathrm{rank}(X) = \mathrm{rank}(A - AZ \left(Z^T AZ\right)^{-1} Z^T A) \geq m - r$. On the other hand, it is possible to show that $\mathrm{rank}(X) \leq m - r$ by showing that

$$\mathrm{null}\left(A - AZ \left(Z^T AZ\right)^{-1} Z^T A\right) \geq n - m + r.$$

Indeed,

$$\begin{aligned} \left(A - AZ \left(Z^T AZ\right)^{-1} Z^T A\right) Cx_1 &= ACx_1 - AZ \left(Z^T AZ\right)^{-1} Z^T ACx_1 \\ &= 0 \end{aligned}$$

and

$$\begin{aligned}\left(A - AZ\left(Z^TAZ\right)^{-1}Z^TA\right)Zx_2 &= AZx_2 - AZ\left(Z^TAZ\right)^{-1}Z^TAZx_2 \\ &= AZx_2 - AZx_2 \\ &= 0.\end{aligned}$$

We thus have shown that $\text{rank}(X) = \text{rank}(A - AZ\left(Z^TAZ\right)^{-1}Z^TA) = m - r$, as required. ∎

In the maximally rank deficient case the following result readily follows from Theorem 3.2.

COROLLARY 3.3. *Let $A \in \mathbb{R}^{n\times n}$ satisfy $\text{null}(A) = m$ and suppose $\mathcal{K}$ is invertible. Then the inverse of $\mathcal{K}$ is itself also a saddle-point matrix, that is, it is of the form*

$$\mathcal{K}^{-1} = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix}.$$

Corollary 3.3 reveals a surprising structure and is in fact useful for further simplifying the formula for $\mathcal{K}^{-1}$. To that end, we now recall the following result; see [7, Proposition 2.1] or [6].

LEMMA 3.4. *Let*

$$\mathcal{K}(W) = \begin{pmatrix} A + B^TW^{-1}B & B^T \\ B & 0 \end{pmatrix},$$

where $W \in \mathbb{R}^{m\times m}$. Then, if $\mathcal{K}(W)$ and $\mathcal{K}(0)$ are invertible, we have

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

Note that $\mathcal{K}(0)$ is just the standard saddle-point matrix denoted throughout this paper by $\mathcal{K}$. We note also that this result generalizes to nonsymmetric matrices and a singular $m \times m$ weight matrix, in which case one should write W rather than W^{-1} on the right-hand sides of the above two formulas.

Corollary 3.3 and Lemma 3.4 can be combined to produce the following result, related to the Schur complement of $\mathcal{K}(W)$.

THEOREM 3.5. *Suppose $\text{null}(A) = m$ and let $W \in \mathbb{R}^{m\times m}$ be an invertible matrix. Then*

$$B\left(A + B^TW^{-1}B\right)^{-1}B^T = W.$$

Proof. Since $A + B^TW^{-1}B$ is invertible, we may apply (3.1) to $\mathcal{K}(W)$, and hence the $(2,2)$ block of $(\mathcal{K}(W))^{-1}$ is given by $-\left(B(A + B^TW^{-1}B)^{-1}B^T\right)^{-1}$. On the other hand, since A has nullity m, Corollary 3.3 applies and the (2,2) block of $(\mathcal{K}(0))^{-1} \equiv \mathcal{K}^{-1}$ is 0. The desired result now readily follows in light of Lemma 3.4. ∎

Theorem 3.5 allows us to develop a new formula for the inverse of $\mathcal{K}$. We continue to assume that $\text{null}(A) = m$. By Corollary 3.3 and Lemma 3.4 it follows that the inverse of $\mathcal{K} \equiv \mathcal{K}(0)$ has a zero (2,2) block and is otherwise equal in the other three blocks to the corresponding blocks of $(\mathcal{K}(W))^{-1}$. Now, applying Theorem 3.5, if we denote

$$\widehat{A} = A + B^TW^{-1}B,$$

then (3.1) becomes

$$\mathcal{K}^{-1} = \begin{pmatrix} \widehat{A}^{-1} - \widehat{A}^{-1}B^TW^{-1}B\widehat{A}^{-1} & \widehat{A}^{-1}B^TW^{-1} \\ W^{-1}B\widehat{A}^{-1} & 0 \end{pmatrix}. \tag{3.3}$$

Note that the leading block of (3.3) can be written in multiplicative form, using the following relation:

$$\begin{aligned} \widehat{A}^{-1} - \widehat{A}^{-1}B^TW^{-1}B\widehat{A}^{-1} &= \widehat{A}^{-1}\widehat{A}\widehat{A}^{-1} - \widehat{A}^{-1}B^TW^{-1}B\widehat{A}^{-1} \\ &= \widehat{A}^{-1}\left(\widehat{A} - B^TW^{-1}B\right)\widehat{A}^{-1} \\ &= \widehat{A}^{-1}A\widehat{A}^{-1}. \end{aligned}$$

Formula (3.3) is interesting for two reasons. First, as opposed to (3.1), it does not contain Schur complements of the form $B(A+B^TW^{-1}B)^{-1}B^T$ (note there is no Schur complement of the form $BA^{-1}B^T$ since A is singular). Second, we have obtained a unique expression for the inverse which nonetheless contains a free matrix parameter, namely W. We now show that the mathematical reason for this lies in the relationship between W and different bases for the null space of A. Mathematically, since the inverse is unique, it does not matter what choice of W we make, but numerically, a judicious choice of W may increase the stability of the computation of the inverse or approximations of it.

PROPOSITION 3.6. *Given $A \in \mathbb{R}^{n\times n}$ whose rank is $n-m$, define $C \in \mathbb{R}^{n\times m}$ as follows:*

$$C = (A + B^TW^{-1}B)^{-1}B^T, \tag{3.4}$$

where $W \in \mathbb{R}^{m\times m}$ is nonsingular. Then, C is a null-space matrix of A. Conversely, if C is a given null-space matrix of A, then it can be written in the form (3.4) *with $W = BC$.*

Proof. Suppose C is given by (3.4), where W is an arbitrary $m \times m$ nonsingular matrix. Then we have

$$\begin{aligned} AC &= \left(A + B^TW^{-1}B - B^TW^{-1}B\right)\left(A + B^TW^{-1}B\right)^{-1}B^T \\ &= B^T - B^TW^{-1}B\left(A + B^TW^{-1}B\right)^{-1}B^T \\ &= B^T - B^TW^{-1}W \\ &= 0, \end{aligned}$$

where in the third row above we used Theorem 3.5. This proves the first part of the statement of the proposition. To show the converse, for a given C, we first show that $W = BC$ is necessarily nonsingular. Indeed, suppose $x \in \text{null}(BC)$, namely $BCx = B(Cx) = 0$. Then $Cx \in \text{null}(A) \cap \text{null}(B)$ and hence x must be 0 by the Helmholtz decomposition (1.2). It is now enough to observe that

$$(A + B^T(BC)^{-1}B)C = B^T$$

for *any* matrix C that satisfies $AC = 0$. ∎

Given the special role that the weight matrix BC plays, from this point onward let us denote it by L, to distinguish it from W, which has so far mostly referred to an *arbitrary* weight matrix:

$$L = BC.$$

Remark 3.7. If C is given, it is always possible to generate a symmetric positive definite weight matrix by defining $\widehat{C} = C(BC)^T$ and setting $L = B\widehat{C}$. Indeed, in this case $B\widehat{C}$ is symmetric, and it is positive definite since the null spaces of A and B do not intersect except at the zero vector.

Given Remark 3.7, from this point onward we will assume that $L = BC$ is symmetric positive definite.

Proposition 3.6 is useful because it gives rise to the simplification

$$\begin{aligned}\widehat{A}^{-1}B^T &= (A + B^T W^{-1} B)^{-1} B^T \\ &= C.\end{aligned}$$

Indeed, the matrix on the left-hand side of the above equation appears four times in formula (3.3) for the inverse. Thus, if C is given we obtain the following simplified expression for the inverse:

$$\mathcal{K}^{-1} = \begin{pmatrix} \left(A + B^T L^{-1} B\right)^{-1} - CL^{-1}C^T & CL^{-1} \\ L^{-1}C^T & 0 \end{pmatrix}. \tag{3.5}$$

An alternative form with a multiplicative leading block, which will come in handy when we move to derive preconditioners in section 4, is

$$\mathcal{K}^{-1} = \begin{pmatrix} \left(A + B^T L^{-1} B\right)^{-1} (I - B^T L^{-1} C^T) & CL^{-1} \\ L^{-1}C^T & 0 \end{pmatrix}. \tag{3.6}$$

In applications where the null space of A is sufficiently easy to compute or approximate, such as the Maxwell problem, one could resort to solution methods based on (3.5) or (3.6).

We note that several of the formulas derived in this section apply to a nonsymmetric A. That said, fully generalizing to the nonsymmetric case entails some additional assumptions on the nonsingularity of $A + B^T W^{-1} B$ and the need to possibly distinguish between left and right null space matrices of A.

4. Block preconditioning. Next, we consider using the formulas for the inverse that were derived in section 3 for developing a new family of block preconditioners. We will seek preconditioners of the form

$$\mathcal{P}^{-1} \approx \mathcal{K}^{-1},$$

based on approximating the new formula (3.6).

In practice, inverting matrices of the form $A + B^T L^{-1} B$ is computationally costly. In most applications the matrix $B^T L^{-1} B$ is dense and cannot be formed. In the iterative context, inversion of $A + B^T L^{-1} B$ is typically done implicitly using inner iterations, but the cost may be high. It is thus necessary to find an effective sparse approximation, and we will seek a sparse matrix $R \approx B^T L^{-1} B$.

To motivate the choice of R, we present the following ideal case.

THEOREM 4.1. *Suppose Z is a null matrix of B, namely $BZ = 0$, and C is a null matrix of A. Then, $R = B^T L^{-1} B$ if and only if* (i) $RZ = 0$, (ii) $RC = B^T$, *and* (iii) $R = R^T$.

Proof. It is straightforward to confirm that $R = B^T L^{-1} B$ satisfies the three conditions (i)–(iii). We now prove that the converse is true too. If $RZ = 0$, then $R = XB$, where X is any full column rank $n \times m$ matrix. Using $R = R^T$, we write

$RZ = R^T Z = B^T X^T Z = 0$. Therefore $X^T = W^{-T}B$, where W is any arbitrary $m \times m$ nonsingular matrix, and $R = B^T W^{-1} B$. Now, $RC = B^T W^{-1} BC = B^T$, which is satisfied if and only if $W = BC = L$. ∎

Based on Theorem 4.1 we can now devise a wish list for designing a preconditioner. The recipe we will seek involves finding a *sparse* matrix $R \in \mathbb{R}^{n \times n}$ such that $A + R$ is easy to invert and the conditions of the theorem are satisfied to the extent possible, so that $A + R \approx A + B^T L^{-1} B$. Clearly, satisfying all these conditions simultaneously is impractical, as it yields the exact inverse. We will therefore settle for satisfying the conditions partially.

4.1. Two block triangular preconditioners. Using R in place of $B^T L^{-1} B$, we can modify $\mathcal{K}^{-1}$ in (3.6) to the first of two block triangular preconditioners that we are proposing:

$$\mathcal{P}_1^{-1} = \begin{pmatrix} (A+R)^{-1}\left(I - B^T L^{-1} C^T\right) & CL^{-1} \\ L^{-1} C^T & 0 \end{pmatrix}. \tag{4.1}$$

It is straightforward to see that if we take $\mathcal{P}_1^{-1}$ and multiply it by $\mathcal{K}$, the term $B^T L^{-1} C^T$ in the leading block of $\mathcal{P}_1^{-1}$ will be multiplied by A and hence it has no effect, because $C^T A = 0$. This leads to the following choice for a second preconditioner, which is obtained by dropping $B^T L^{-1} C^T$ from the leading term of $\mathcal{P}_1^{-1}$:

$$\mathcal{P}_2^{-1} = \begin{pmatrix} (A+R)^{-1} & CL^{-1} \\ L^{-1} C^T & 0 \end{pmatrix}. \tag{4.2}$$

The preconditioned matrices $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ are given respectively by

$$\mathcal{P}_1^{-1}\mathcal{K} = \begin{pmatrix} (A+R)^{-1}A + CL^{-1}B & 0 \\ 0 & I \end{pmatrix}$$

and

$$\mathcal{P}_2^{-1}\mathcal{K} = \begin{pmatrix} (A+R)^{-1}A + CL^{-1}B & (A+R)^{-1}B^T \\ 0 & I \end{pmatrix}.$$

As can be observed, $\mathcal{P}_1^{-1}\mathcal{K}$ is block diagonal, whereas $\mathcal{P}_2^{-1}\mathcal{K}$ is block upper triangular. While the eigenvalues of the two matrices are identical and have the same algebraic multiplicities, their geometric multiplicities and the degrees of the minimal polynomials associated with these preconditioned matrices are different, and in fact this has an implication on the choice of solution methods, as we will soon see.

4.2. Spectral analysis. From the structure of $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$, it is evident that we need to concern ourselves with only the diagonal blocks to determine their eigenvalues. We immediately obtain m eigenvalues equal to 1 from the identity matrix in the (2,2) block, and so we need only focus on the leading block for the remaining eigenvalues.

Since A is maximally rank deficient, any vector $u \in \mathbb{R}^n$ can be decomposed into $u = u_A + u_B$, where $u_A \in \ker(A)$ and $u_B \in \ker(B)$, as per (1.2). Notice that for vectors in the null space of A, i.e., those of the form $u_A = Cy$,

$$\begin{aligned} \left((A+R)^{-1}A + CL^{-1}B\right) Cy &= (A+R)^{-1} ACy + CL^{-1}BCy \\ &= Cy. \end{aligned} \tag{4.3}$$

Therefore, since there are m linearly independent vectors in the null space of A, we immediately obtain m eigenvalues equal to 1. We then aim to choose R so that the remaining $n-m$ eigenvalues of the (1,1) block of $\mathcal{P}_1^{-1}\mathcal{K}$ or $\mathcal{P}_2^{-1}\mathcal{K}$ are well distributed.

Let us collect the above observations into a statement that clarifies what multiplicity the eigenvalue 1 has for the preconditioned matrices that we are considering.

PROPOSITION 4.2. *The matrices $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ have eigenvalue 1 of algebraic multiplicity at least $2m$. In $\mathcal{P}_1^{-1}\mathcal{K}$, the eigenvalue 1 has geometric multiplicity at least $2m$ with eigenvectors of the form (Cy, u), where $u, y \in \mathbb{R}^m$. In $\mathcal{P}_2^{-1}\mathcal{K}$ the eigenvalue 1 has geometric multiplicity at least m with eigenvectors of the form $(Cy, 0)$.*

We now consider the remaining eigenvalues of the leading block of $\mathcal{P}_1^{-1}\mathcal{K}$ or $\mathcal{P}_2^{-1}\mathcal{K}$. By (1.2), suppose $u = u_A + u_B$ is an eigenvector, $\|u\| = 1$, with $u_A \in \ker(A)$ and $u_B \in \ker(B)$. Then we have

$$\begin{aligned}\lambda u &= \left((A+R)^{-1}A + C(BC)^{-1}B\right)u \\ &= (A+R)^{-1}Au_B + C(BC)^{-1}Bu_A \\ &= \left(I - (A+R)^{-1}R\right)u_B + u_A \\ &= u - (A+R)^{-1}Ru_B. \qquad (4.4)\end{aligned}$$

Therefore

$$|1-\lambda| = \|(A+R)^{-1}Ru_B\|.$$

It follows that if R satisfies condition (i) of Theorem 4.1, then $\mathcal{P}_1^{-1} = \mathcal{K}^{-1}$. This is obviously a case that yields stronger results than what is guaranteed in Proposition 4.2: the preconditioned matrix is trivially diagonalizable and the multiplicity of the eigenvalue 1 is $n+m > 2m$.

Constructing R such that its null space is equal to that of B is easy, because any choice of the form $R = XB$ (including the specific choice $R = B^TW^{-1}B$) will give condition (i). However, this would practically mean that we are computing the exact inverse of $\mathcal{K}$, and hence it is impractical. We therefore consider relaxing condition (i) of Theorem 4.1, while still aiming to achieve a good eigenvalue distribution.

PROPOSITION 4.3. *If R satisfies $\|(A+R)^{-1}Ru_B\| \le \alpha$ for $\alpha < 1$ and $u_B \in \ker(B)$, then the eigenvalues λ of $\mathcal{P}_i^{-1}\mathcal{K}, i = 1, 2$, satisfy $1-\alpha \le \lambda \le 1+\alpha$.*

Remark 4.4. Proposition 4.3 allows for a simple choice of R. For example, if $R = \delta I$, for $\delta < \sigma_{n-m}$, then the eigenvalues λ of $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ satisfy $|1-\lambda| < \frac{\delta}{\sigma_{n-m}-\delta}$, where σ_{n-m} is the smallest nonzero singular value of A.

Continuing on from (4.4), we observe that an equivalent condition can be stated on R, as follows. Multiplying by $A+R$ on both sides and rearranging we get $(1-\lambda)(A+R)u = Ru_B$, which yields

$$\frac{\lambda}{1-\lambda}u_B^TRu_B - u_B^TRu_A = u_B^TAu_B.$$

From this we conclude that if R is chosen so that

$$\frac{|u_B^TRu_B|}{|u_B^TAu_B|} \le \beta \quad \text{and} \quad |u_B^TRu_A| \ll 1,$$

where $\beta \ll 1$, then $\lambda \approx 1$. This is an algebraically analogous condition to [12, equation (9)], which eventually leads to [12, Theorem 3.3].

Remark 4.5. If $RC = B^T$, then we automatically have the R-orthogonality relation $u_B^TRu_A = 0$, which amounts to weakening the conditions of Proposition 4.3

and requiring that there be $\beta \ll 1$ such that $\frac{|u_B^T R u_B|}{|u_B^T A u_B|} \leq \beta$, in order to ensure a good eigenvalue distribution.

4.3. Using conjugate gradient with the indefinite preconditioners $\mathcal{P}_1$ and $\mathcal{P}_2$. We now move to consider preconditioned Krylov subspace solvers. Both $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ have the same leading block, which upon using $AC = 0$ simplifies to

$$\begin{aligned}(A+R)^{-1}A + CL^{-1}B &= (A+R)^{-1}\left(A + (A+R)CL^{-1}B\right)\\ &= (A+R)^{-1}(A + RCL^{-1}B).\end{aligned}$$

Let us assume that condition (ii) of Theorem 4.1 holds, namely $RC = B^T$. Then, the (1,1) blocks of $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ are both equal to $(A+R)^{-1}(A + B^T L^{-1}B)$, and $R \approx B^T L^{-1}B$ is likely to accomplish the goal of having a well-distributed spectrum. It then follows that $\mathcal{P}_1^{-1}\mathcal{K}$ is a block diagonal matrix whose diagonal blocks are $(A + R)^{-1}(A + B^T L^{-1}B)$ and I. Therefore, if $A + R$ is symmetric positive definite, a surprising prospect arises of solving the preconditioned linear system associated with $\mathcal{P}_1^{-1}\mathcal{K}$ using the conjugate gradient (CG) method. In other words, CG may be used despite the fact that both $\mathcal{K}$ and $\mathcal{P}_1$ are indefinite. A necessary condition is thus that R is symmetric and is positive definite on the null space of A.

On the other hand, $\mathcal{P}_2^{-1}\mathcal{K}$ simplifies as follows:

$$\mathcal{P}_2^{-1}\mathcal{K} = \begin{pmatrix}(A+R)^{-1}A + CL^{-1}B & C\\ 0 & I\end{pmatrix}.$$

The off-diagonal block is now C rather than $(A+R)^{-1}B^T$, by condition (ii) of Theorem 4.1. However, in order to use CG here too, we will have to require the right-hand side to satisfy a divergence-free condition, as follows.

Suppose that we have a zero initial guess. Then the initial residual is equal to

$$r_0 = \begin{pmatrix}f\\ g\end{pmatrix}.$$

Therefore, if $C^T f = 0$, then

$$\mathcal{P}_1^{-1}r_0 = \begin{pmatrix}(A+R)^{-1}f + CL^{-1}g\\ 0\end{pmatrix} = \mathcal{P}_2^{-1}r_0.$$

We note that if a nonzero initial guess is to be used, the initial residual may not have a divergence-free component.

We now observe that

$$\begin{aligned}(\mathcal{P}_2^{-1}\mathcal{K})(\mathcal{P}_2^{-1}r_0) &= \begin{pmatrix}(A+R)^{-1}A + CL^{-1}B & C\\ 0 & I\end{pmatrix}\begin{pmatrix}(A+R)^{-1}f + CL^{-1}g\\ 0\end{pmatrix}\\ &= (\mathcal{P}_1^{-1}\mathcal{K})(\mathcal{P}_1^{-1}r_0).\end{aligned}$$

Noticing that the off-diagonal block term C in $\mathcal{P}_2^{-1}\mathcal{K}$ is multiplied by zero, we conclude that for higher powers, $j > 1$, we also have $(\mathcal{P}_2^{-1}\mathcal{K})^j(\mathcal{P}_2^{-1}r_0) = (\mathcal{P}_1^{-1}\mathcal{K})^j(\mathcal{P}_1^{-1}r_0)$. Therefore, the preconditioned Krylov subspaces associated with $\mathcal{P}_i^{-1}\mathcal{K}$ and the preconditioned residual $\mathcal{P}_i^{-1}r_0$, which are

$$\operatorname{span}(\mathcal{P}_i^{-1}r_0, \mathcal{P}_i^{-1}\mathcal{K}\mathcal{P}_i^{-1}r_0, \ldots, (\mathcal{P}_i^{-1}\mathcal{K})^{k-1}\mathcal{P}_i^{-1}r_0), \quad i = 1, 2,$$

are identical.

We conclude that if $RC = B^T$ and f is divergence-free, then $\mathcal{P}_2$ can be used as a preconditioner for CG. If on the other hand f is not divergence-free but $RC = B^T$, then it may be useful to use $\mathcal{P}_1$ in conjunction with CG.

At first sight it seems that matrix-vector products with $\mathcal{P}_1^{-1}\mathcal{K}$ require an extra L-solve compared to matrix-vector products with $\mathcal{P}_2^{-1}\mathcal{K}$. But in fact the two are almost identical in terms of computational cost. This is because for an arbitrary vector $\binom{x}{y}$,

$$\mathcal{P}_1^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (A+R)^{-1}(I - B^T L^{-1} C^T)x + CL^{-1}y \\ L^{-1}C^T x \end{pmatrix},$$

that is, the extra term $L^{-1}C^T x$ is repeated, and hence the only overhead compared to forming $\mathcal{P}_2^{-1}\binom{x}{y}$ is a multiplication by B^T and a vector subtraction operation.

Altogether, a preconditioner solve with either $\mathcal{P}_1$ or $\mathcal{P}_2$ entails one inversion of $A + R$ and two L-solves. This is a reasonable cost, assuming that C is available.

Satisfying $RC = B^T$ is thus greatly beneficial, since it allows for using CG on the preconditioned operator $\mathcal{P}_1^{-1}\mathcal{K}$ without any additional conditions, and on $\mathcal{P}_2^{-1}\mathcal{K}$ with the additional condition that $C^T f = 0$. However, as previously noted, it may be difficult to find R that satisfies $RC = B^T$. In that case CG cannot be used and we may resort to a standard nonsymmetric solver. We note also that our analysis relies on exact solves with $A + R$. If it is difficult to solve linear systems with $A + R$, then we may need to find an approximation for $A + R$ or apply inexact solves, which would amount to solving linear systems with matrices of the form $A + R + E$. In that case CG may not work, because $(R + E)C \neq B^T$ or $A + (R + E)CL^{-1}B$ is likely to be nonsymmetric. One way to overcome this difficulty is to continue to use $A + R$ in the preconditioner $\mathcal{P}_1^{-1}$ but apply inner iterations in which the linear systems with $A + R$ are solved by using preconditioned iterations.

If CG cannot be used, the strategy of choosing R may be affected, as it may be more useful to satisfy a relaxed version of condition (i) of Theorem 4.1; see Proposition 4.3. Here using $\mathcal{P}_2$ as a preconditioner may be beneficial, in particular if C is not the exact null-space matrix.

Inspecting the conditions of Theorem 4.1, we conclude that the condition $RZ = 0$ and its relaxed form in Proposition 4.3 primarily affect the distribution of the eigenvalues, whereas the conditions $RC = B^T$ and $R = R^T$ are needed in order to be able to potentially use CG. In this case we need to require A to be semidefinite and R to be positive definite on the null space of A. If A is indefinite or nonsymmetric, then CG cannot be easily used (using CG may require finding a special inner product and may not be feasible [17]). In such situations there is no point in seeking R that satisfies $RC = B^T$. Instead, one should aim for finding R that satisfies Proposition 4.3, and a standard nonsymmetric solver or a specialized version of MINRES [8] can be used.

4.4. Relation to block diagonal preconditioners. $\mathcal{P}_2^{-1}$ is connected to the block diagonal preconditioner used in [12] in the following way. We use the standard block LDL decomposition of a saddle-point matrix—see, for example, [1, equation (3.1)]—and exploit the fact that $\mathcal{P}_2^{-1}$ has a saddle-point structure. If $RC = B^T$ and $R = R^T$, then from the symmetry of A and $C^T A = AC = 0$ we have the relation $C^T(A + R) = B$, and we can form the block LDL decomposition

$$\begin{aligned}\mathcal{P}_2^{-1} &= \begin{pmatrix} I & 0 \\ L^{-1}B & I \end{pmatrix}\begin{pmatrix} (A+R)^{-1} & 0 \\ 0 & -L^{-1}C^T(A+R)CL^{-1} \end{pmatrix}\begin{pmatrix} I & B^T L^{-1} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ L^{-1}B & I \end{pmatrix}\begin{pmatrix} (A+R)^{-1} & 0 \\ 0 & -L^{-1} \end{pmatrix}\begin{pmatrix} I & B^T L^{-1} \\ 0 & I \end{pmatrix}.\end{aligned}$$

Since the condition $RC = B^T$ is satisfied if $R = B^T L^{-1} B$, the above result holds also for a leading (1,1) block of the form $(A + B^T L^{-1} B)^{-1}$, which is in fact the ideal block diagonal preconditioner studied and used in [11, 12], except the sign of the (2,2) block is positive so as to obtain a block diagonal positive definite preconditioner.

5. Numerical experiments. We now illustrate the numerical merits of the proposed preconditioners. We consider two of the problems described in section 2, one of them with a maximally rank deficient leading block and one with a leading block whose nullity is not maximal. In both cases our preconditioning approach can be utilized in a way that yields rapid convergence.

5.1. Time-harmonic Maxwell. The time-harmonic Maxwell problem, as described in section 2, gives rise to a scenario that can be handled by our preconditioner. The leading block is maximally rank deficient, and its null space is known and sparse—the space of gradient functions. As mentioned in section 2, it is possible to construct a matrix C of gradients within the finite element spaces that satisfies $AC = 0$, where A is a discrete curl-curl operator. Denoting the vector mass matrix by M and the scalar Laplacian by L, it can be shown that $MC = B^T$ and $L = BC$ [12, Proposition 2.2]. These properties lead to the observation that M approximates $B^T(BC)^{-1}B \equiv B^T L^{-1} B$ in a spectral equivalence sense [12, Theorem 3.3], making it useful for preconditioning. Hence, the matrix M is the ideal choice for playing the role of the matrix R discussed in section 4.

Using spectral equivalence and inf-sup stability results, it is shown in [12] that a block diagonal symmetric positive definite preconditioner whose $(1, 1)$ block is equal to $A + M$ and whose $(2, 2)$ block is the scalar Laplacian $L = BC$, used within a MINRES iteration, is a highly effective preconditioner with a large number of eigenvalues of the preconditioned matrix clustered: the multiplicities of the eigenvalues 1 and -1 are provably m for each. The $n - m$ additional positive eigenvalues are bounded away from zero, and in typical settings most of them appear (experimentally) to be near 1.

In the context of the preconditioning approach proposed in this paper, we set R as the vector mass matrix and use the scalar Laplacian $L = BC$ to set up the preconditioner $\mathcal{P}_1$ defined in (4.1), which allows for applying CG for any right-hand-side vector. This preconditioner entails a computational cost per iteration which is almost identical to that associated with $\mathcal{P}_2$, and it has the additional advantage that the preconditioned matrix $\mathcal{P}_1^{-1}\mathcal{K}$ is symmetric and may yield slightly faster convergence compared to solving the linear system associated with $\mathcal{P}_2^{-1}\mathcal{K}$, due to the degree of its minimal polynomial.

We show in Table 1 (outer) iteration counts, using preconditioned CG with $\mathcal{P}_1$ as a preconditioner, for the model problem described in [12, section 6.1], defined on the unit square with a quasi-uniform grid. The iteration counts show full scalability. They are a bit higher than the iteration counts of using the diagonal block preconditioner in [12], but in the current setting we can use CG rather than MINRES, which gives rise to more economical iterations. Table 2 shows iterations for an adaptively refined L-shaped domain, described in [12, section 6.2]. Again, we observe full scalability.

In Figure 1 we show the eigenvalue distribution of $\mathcal{P}_1^{-1}\mathcal{K}$. It is evident that almost all eigenvalues are strongly clustered around 1, and the ones that are smaller than 1 are bounded away from 0.

TABLE 1

Iteration counts for a Maxwell problem on the unit square with quasi-uniform meshes, using $\mathcal{P}_1$ as a preconditioner for various right-hand sides, $\begin{pmatrix} f \\ g \end{pmatrix}$. The iteration was stopped once the initial relative residual was reduced by a factor of 10^{-10}. RfRg = random f, random g; Rf0g = random f, zero g; DfRg = div-free f, random g; Df0g = div-free f, zero g.

Grid	$n+m$	RfRg	Rf0g	DfRg	Df0g
G1	113	8	8	7	6
G2	481	7	8	7	6
G3	1,985	7	8	7	6
G4	8,065	7	8	7	6
G5	32,513	7	7	7	6

TABLE 2

Iteration counts for a Maxwell problem on an L-shaped domain with locally refined meshes, using $\mathcal{P}_1$ as a preconditioner for various right-hand sides, $\begin{pmatrix} f \\ g \end{pmatrix}$. The iteration was stopped once the initial relative residual was reduced by a factor of 10^{-10}.

Grid	$n+m$	RfRg	Rf0g	DfRg	Df0g
L1	451	6	6	5	5
L2	813	6	6	5	5
L3	2,608	6	6	5	5
L4	9,927	6	6	5	5
L5	37,882	6	6	5	5

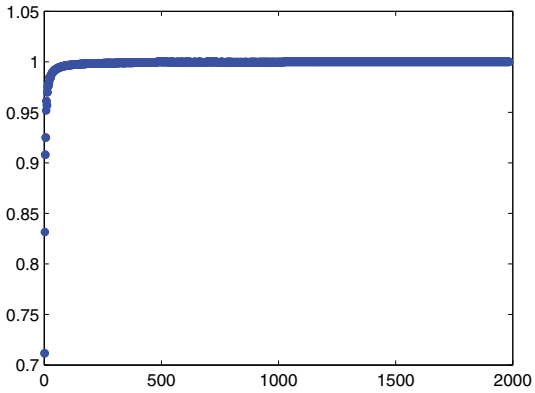

FIG. 1. *Eigenvalue distribution of the preconditioned matrix $\mathcal{P}_1^{-1}\mathcal{K}$ for the time-harmonic Maxwell problem with grid G3.*

5.2. Geophysical inverse problem. The saddle-point matrix for the geophysical inverse problem [3, 13] briefly described in section 2 can be written as follows:

$$\mathcal{K} = \begin{pmatrix} D & 0 & F^T \\ 0 & \alpha N & G^T \\ F & G & 0 \end{pmatrix},$$

where $D \in \mathbb{R}^{m\times m}$, $N \in \mathbb{R}^{(n-m)\times(n-m)}$, $F \in \mathbb{R}^{m\times m}$, and $G \in \mathbb{R}^{m\times(n-m)}$. In the notation of this paper, the blocks A and B of (1.1) are given by

$$A = \begin{pmatrix} D & 0 \\ 0 & \alpha N \end{pmatrix}; \qquad B = \begin{pmatrix} F & G \end{pmatrix}.$$

TABLE 3
Information on the dimensions of the problems solved, and the rank of the data.

Problem	n	m	*DataRank*
P1	1729	1000	199
P2	4941	2744	323
P3	14859	8000	375

The matrix D is extremely sparse and the matrix N is typically nonsingular. Let us denote the nullity of D (and hence of A) by r. We have a situation whereby r is relatively large but smaller than m. To apply our preconditioner, we need to assume maximal rank deficiency and the availability of an $n \times m$ null space matrix C such that $AC = 0$, but in this case only up to r of the columns of a matrix C with m columns could possibly be null vectors of A. We make the convenient choice

$$C = \begin{pmatrix} I_m \\ 0 \end{pmatrix},$$

which provides us with a set of r linearly independent null vectors of A, complemented by $m-r$ standard basis vectors that are not in the null space. From a purely spectral point of view, a better choice than those $m-r$ vectors for complementing the columns of C would be the eigenvectors of N that correspond to its smallest eigenvalues, but those eigenvectors are hard to compute and are dense. We thus use the above defined C, which is sparse and involves no computation, as an approximate null space of A for the purpose of constructing our preconditioner.

For the matrix R we make the computationally convenient choice $R = \delta I$ and experiment with the preconditioner $\mathcal{P}_1$. It is possible to simplify the formula for the preconditioner as follows:

$$\begin{aligned}
\mathcal{P}_1^{-1} &= \begin{pmatrix} \begin{pmatrix} D+\delta I & 0 \\ 0 & \alpha N + \delta I \end{pmatrix}^{-1} \left(I - \begin{pmatrix} F^T \\ G^T \end{pmatrix} F^{-T} \begin{pmatrix} I & 0 \end{pmatrix} \right) & \begin{pmatrix} F^{-T} \\ 0 \end{pmatrix} \\ \begin{pmatrix} F^{-1} & 0 \end{pmatrix} & 0 \end{pmatrix} \\
&= \begin{pmatrix} \begin{pmatrix} D+\delta I & 0 \\ 0 & \alpha N + \delta I \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ G^T F^{-T} & I \end{pmatrix} & \begin{pmatrix} F^{-T} \\ 0 \end{pmatrix} \\ \begin{pmatrix} F^{-1} & 0 \end{pmatrix} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & F^{-T} \\ -(\alpha N + \delta I)^{-1} G^T F^{-T} & (\alpha N + \delta I)^{-1} & 0 \\ F^{-1} & 0 & 0 \end{pmatrix}.
\end{aligned}$$

The results we provide next demonstrate the robustness of our solution approach. We use a MATLAB code that solves a basic version of the model problem described in [13]; a more elaborate implementation based on ocTree has recently been developed [3]. Table 3 provides information on experiments with three problem dimensions, which we call P1, P2, and P3. The column *DataRank* specifies the rank of the matrix D. The nullity of the leading block is thus $r = m - \mathit{DataRank}$. For example, in P1 the matrix C contains 801 null vectors of A, whereas 199 columns of C are not null vectors.

Table 4 shows that there is fast convergence for all cases. GMRES is used, and the tolerance for convergence was set as 10^{-3}, which is a reasonable stopping criterion for this type of problem. It is to be expected that small values of the regularization

TABLE 4
Iteration counts for P1, P2, P3 for various values of the regularization parameter α.

α	P1	P2	P3
1.0E-02	8	8	8
1.0E-04	10	10	10
1.0E-06	11	11	11
1.0E-08	13	13	13
1.0E-10	15	15	15

parameter cause difficulty in terms of convergence, and indeed the iteration counts go up as α becomes smaller, but the solver performs well and the iteration counts seem independent of the problem size.

We note that there are other possible solution methods, which do not necessarily rely on the saddle-point formulation. In [3], a method based on projecting the problem onto the constraint manifold is applied. The matrix $(Z = \begin{smallmatrix} F^{-1}G \\ -I \end{smallmatrix})$ satisfies $BZ = 0$, and consequently the reduced Hessian is given by $Z^TAZ = G^TF^{-T}DF^{-1}G + \alpha N$. Preconditioned CG is used with N as a preconditioner. The cost per iteration of the reduced Hessian approach is identical to the cost of our preconditioning approach in terms of the matrix-vector products involved. A potential advantage of the reduced Hessian approach is that the iterations satisfy the constraints by design. A disadvantage is that for the linear system to be solved to a high accuracy, F basically needs to be inverted exactly. In contrast, our approach allows for seamlessly using inexact F-solves, and this flexibility may prove extremely useful in large-scale settings.

6. Discussion and concluding remarks. The preconditioner that we have derived relies on a reworked formula for the inverse of the saddle-point matrix. As we have shown, if the $(1,1)$ block of a saddle-point matrix has nullity equal to the number of rows of the $(2,1)$ block, the inverse has unique and interesting properties, which can be utilized for the design of new preconditioning techniques.

A limitation of our approach is that knowledge of the null space matrix of A is required. In the context of preconditioned iterative methods, it is desirable for the null space to be easily computable and sparse. In some situations this does not form a difficulty. The case of the time-harmonic Maxwell equations is amenable to our approach, since the leading block in that case (a discrete curl-curl operator) has a null space of discrete gradients that are available and sparse. Other situations where the null space may be easily available and sparse are when the leading block has a simple structure. Some examples were given in section 2.

Even if the leading block of the saddle-point matrix is known to be highly rank-deficient, the null space is often not easily computable and is not sparse, and we would like to be able to apply the preconditioning approach in such cases too. This is similar to the situation with Schur complement-based preconditioning approaches, which also rely on approximating matrices that are not sparse and are not easily available; the key is to find effective sparse approximations. When it is not easy to directly find an effective approximation for the null-space matrix C that satisfies $AC = 0$, Proposition 3.6, and specifically (3.4), may allow for obtaining an approximation in terms of A, B, and a free matrix parameter W. The choice of the matrix R described in section 4 is problem-dependent, and practical application of the preconditioner should also take into account the properties of R and the implementation of inexact inner iterations.

In addition, it is desirable to be able to handle situations where the nullity (or rank deficiency) is not maximal, i.e., it is not equal to the number of the constraints. If the nullity is sufficiently close to m, then we may still be able to obtain good convergence properties using a matrix C that comprises an approximation to the known null space, complemented by a convenient choice of sparse additional columns. We have experimentally demonstrated the viability of such an approach for a geophysical inverse problem in section 5.

Another issue that should be considered is the possibility of decoupling. If C is explicitly known, then by multiplying the first block-row of (1.1) by C^T we can immediately decouple the system and solve the system $C^T B^T p = C^T f$. Symmetry of BC further simplifies the setting, and if $C^T f = 0$ as we assumed in one of the scenarios considered in section 4, then it readily follows that $p = 0$ and what is left is to compute u without solving a saddle-point system. However, keeping the saddle-point structure allows for additional flexibility: first, an inexact application of the C matrix may still work on the preconditioning level, whereas the exact C is required for decoupling. Second, $\mathcal{P}_1^{-1}$ and $\mathcal{P}_2^{-1}$ in (4.1) and (4.2), respectively, can be used as a starting point, and further approximations of the individual blocks may be used. This is similar to the situation with other block preconditioners. For example, for nonsingular A the preconditioner analyzed in [19] has A as its (1,1) block and $BA^{-1}B^T$ as its (2,2) block as a starting point, but in practice those operators are not explicitly formed or inverted; rather, approximations to the operation of their inverses are utilized.

The ability to use the preconditioned CG method for $\mathcal{P}_1^{-1}\mathcal{K}$ and $\mathcal{P}_2^{-1}\mathcal{K}$ is interesting, because both the preconditioner and the coefficient matrix are indefinite and their product is nonsymmetric. In this regard the proposed preconditioner reveals properties that are similar to those of constraint preconditioners, where projected CG can be used [9]. For $\mathcal{P}_2^{-1}\mathcal{K}$ the eigenvalue 1 seems to follow a pattern that the eigenvalue 1 has in constraint preconditioning, namely that its algebraic multiplicity is $2m$, whereas its geometric multiplicity is only m [16].

Acknowlegments. We thank Kristofer Davis and Eldad Haber for their code for solving the geophysical inverse problem described in sections 2 and 5. We also thank two anonymous referees for their helpful comments.

REFERENCES

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

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[3] K. Davis, E. Haber, and D. Oldenburg, *Large-scale magnetic inversion using differential equations and octrees*, in Proceedings of ASEG-PESA, 23rd International Geophysical Conference and Exhibition, Melbourne, Australia, 2013.

[4] L. Demkowicz and L. Vardapetyan, *Modeling of electromagnetic absorption/scattering problems using hp-adaptive finite elements*, Comput. Methods Appl. Mech. Engrg., 152 (1998), pp. 103–124.

[5] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2005.

[6] R. Fletcher, *An ideal penalty function for constrained optimization*, in Nonlinear Programming 2 (Proc. Sympos. Special Interest Group on Math. Programming, Univ. Wisconsin, Madison, Wis., 1974), Academic Press, New York, 1974, pp. 121–163.

[7] G. H. Golub and C. Greif, *On solving block-structured indefinite linear systems*, SIAM J. Sci. Comput., 24 (2003), pp. 2076–2092.

[8] N. I. M. Gould, D. Orban, and T. Rees, *Projected krylov methods for saddle-point systems*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1329–1343.
[9] N. I. M. Gould, M. E. Hribar, and J. Nocedal, *On the solution of equality constrained quadratic programming problems arising in optimization*, SIAM J. Sci. Comput., 23 (2001), pp. 1376–1395.
[10] C. Greif, E. Moulding, and D. Orban, *Bounds on eigenvalues of matrices arising from interior-point methods*, SIAM J. Optim., 24 (2014), pp. 49–83.
[11] C. Greif and D. Schötzau, *Preconditioners for saddle point linear systems with highly singular* $(1,1)$ *blocks*, Electron. Trans. Numer. Anal., 22 (2006), pp. 114–121.
[12] C. Greif and D. Schötzau, *Preconditioners for the discretized time-harmonic Maxwell equations in mixed form*, Numer. Linear Algebra Appl., 14 (2007), pp. 281–297.
[13] E. Haber, U. M. Ascher, and D. Oldenburg, *On optimization techniques for solving nonlinear inverse problems*, Inverse Problems, 16 (2000), pp. 1263–1280.
[14] R. Hiptmair, *Finite elements in computational electromagnetism*, Acta Numer., 11 (2002), pp. 237–339.
[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.
[16] 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.
[17] J. Liesen and B. N. Parlett, *On nonsymmetric saddle point matrices that allow conjugate gradient iterations*, Numer. Math., 108 (2008), pp. 605–624.
[18] P. Monk, *Finite Element Methods for Maxwell's Equations*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2003.
[19] 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.
[20] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 2006.
[21] Y. Notay, *A new analysis of block preconditioners for saddle point problems*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 143–173.
[22] J. Pestana and A. J. Wathen, *Natural preconditioning and iterative methods for saddle point systems*, SIAM Rev., 57 (2015), pp. 71–91.