# SPECTRAL ANALYSIS OF INEXACT CONSTRAINT PRECONDITIONING FOR SYMMETRIC SADDLE POINT MATRICES *

DEBORA SESANA† AND VALERIA SIMONCINI‡

**Abstract.** Large symmetric linear systems in saddle point form arise in many scientific and engineering applications. Their efficient solution by means of iterative methods mainly relies on exploiting the matrix structure. Constraint preconditioners are among the most successful structure-oriented preconditioning strategies, especially when dealing with optimization problems. In this paper we provide a full spectral characterization of the constraint-based preconditioned matrix by means of the Weyr canonical form. We also derive estimates for the spectrum when the preconditioner needs to be modified to cope with possible high computational costs of its original version. Numerical experiments confirm our findings and illustrate that these theoretical results can be helpful in analyzing matrices stemming from real applications.

**Key words.** saddle point matrices, symmetric linear systems, Weyr canonical form, Jordan blocks, nonlinear perturbation.

**AMS subject classifications.** 65F10

**1. Introduction.** Large scale nonsingular linear systems whose symmetric coefficient matrix has the following saddle point structure

$$\mathcal{A} = \begin{bmatrix} A & B^\top \\ B & -C \end{bmatrix}, \qquad A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{m\times m}, m \le n, \tag{1.1}$$

arise in a wide variety of applications in science and engineering; we refer to [2] for a thorough account on the origin of the problem, and on the description of many solution strategies. Here we assume that $A$ is symmetric and positive definite, $C$ is symmetric and positive semi-definite, and that $BB^\top + C$ is nonsingular, where $B^\top$ denotes the transpose of $B$.

Conveniently exploiting the matrix block structure allows one to devise computational effective acceleration procedures that makes it possible to solve really large two and three dimensional application problems. In particular, structure-based preconditioners have become a formidable acceleration device, as they can naturally exploit a-priori information on, for example, the operators leading to the blocks $A, B$ and $C$. In this paper we investigate the spectral properties of preconditioned matrices $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$, where $\mathcal{P}_{\rm ex}$ is given by

$$\mathcal{P}_{\rm ex} = \begin{bmatrix} \widehat{A} & B^\top \\ B & -C \end{bmatrix},$$

and $\widehat{A}$ is an approximation of $A$ (in short, $\widehat{A} \approx A$). To simplify the presentation and without loss of generality, we assume that $\mathcal{A}$ was already preconditioned by means of a preprocessing (e.g., scaling or block diagonal preconditioning), so that we can

*Version of January 29, 2012.

†Dipartimento di Studi per l'Economia e l'Impresa, Università degli Studi del Piemonte Orientale - Amedeo Avogadro, Via Perrone, 18, 28100 Novara, Italy (debora.sesana@eco.unipmn.it).

‡Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, 40127 Bologna, Italy (valeria.simoncini@unibo.it).

assume that $A$ is spectrally equivalent (see, e.g., [14]) to the identity matrix. More precisely, we assume that there exist two real positive constants $\alpha_1, \alpha_2$ such that[1]

$$\alpha_1 x^\top x \le x^\top A x \le \alpha_2 x^\top x, \quad 0 \ne x \in \mathbb{R}^n.$$

Such a setting arises, for instance, when $A$ corresponds to the mass matrix, as in the mixed finite element approximation of various PDE problems (see, e.g., [14] for the general framework, and [17] for a specific application), and in the discretization of certain elliptic optimal control problems (see, e.g., [21]). In the following we shall often use the hypothesis that $A - I$ is nonsingular. This could be easily obtained by imposing, for instance, that $\alpha_1, \alpha_2$ be either both less than one, or both greater than one. Such requirement would not be restrictive, since the whole problem can be scaled so that this condition is satisfied. Moreover, the preprocessing should ensure that the spectral interval of $A$ is well clustered. Note that by assuming that, say, both scalars are strictly less than one, it also follows that all eigenvalues of $A$ are less than one.

Under these assumptions, we can take $\widehat{A}$ to be the identity matrix, so that the preconditioner $\mathcal{P}_{\rm ex}$ is now given by

$$\begin{aligned} \mathcal{P}_{\rm ex} &= \begin{bmatrix} I_n & B^\top \\ B & -C \end{bmatrix} \\ &= \begin{bmatrix} I_n & 0 \\ B & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -H \end{bmatrix} \begin{bmatrix} I_n & B^\top \\ 0 & I_m \end{bmatrix}, \end{aligned} \tag{1.2}$$

where

$$H = BB^\top + C, \tag{1.3}$$

with $H$ nonsingular, as mentioned above. In the following, $I_k$ and $0_k$ denote the identity and zero matrices of size $k$, respectively; the subscript will be omitted whenever the matrix dimension is clear from the context. The dimensions of zero rectangular matrices will also be deduced from the context.

It can be shown that when solving the transformed problem with $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ by means of an iterative method, the iterates satisfy the original constraint (given by the second block row in the associated system). This major property has made $\mathcal{P}_{\rm ex}$ very popular in the optimization community, and a lot is now known on the eigenvalue properties of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$, which to a large extent seem to guide the convergence of the iterative system solver; see, e.g., the theoretical developments and algorithmic consequences in [10], [17], [12], [7], [9], [8], [1]. As a key spectral feature, we recall that $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ has all real eigenvalues, and that $m$ Jordan blocks of size 2 corresponding to the unit eigenvalue also arise. Moreover, some eigenvector structure was analyzed in [7]. Many numerical experiments have shown that using $\mathcal{P}_{\rm ex}$ may be very competitive in various applications, and that the nonsymmetry of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ does not represent a major limitation.

The application of $\mathcal{P}_{\rm ex}^{-1}$, using the factorization (1.2), requires solving systems with $H$ (cf. (1.3)). Explicitly solving with $H$ may be unrealistic when dealing with 3D applications, so that a cheap approximation to $H$ is used, giving rise to an *inexact* constraint preconditioner; see, e.g., [17], [11]. This necessary step seems to jeopardize the whole theory of constraint preconditioning, as in general complex eigenvalues arise and may spread well away from the original values obtained when $H$ is used.

[1] Often the two positive constants are also required to be independent of some problem parameters, such as the dimension. We do not impose such condition.

This wide spreading is mainly caused by the perturbation to the multiple unit eigenvalues, which is expected to be *nonlinear* in the error made in approximating $H$. Such a phenomenon has created some concern as of the adequacy of this preconditioning procedure in the inexact case, although experimental evidence shows otherwise. Theoretical ground supporting this optimistic numerical experience has remained for long time a largely open issue. First attempts to analyze the spectral modification occurring when using an inexact form of $H$ can be found in [3], [17], but a thorough understanding is still missing. Here we aim to fill this gap. By replacing $H$ with $H_{\rm inex}$ in the preconditioner, we can thus define the preconditioning matrix

$$\mathcal{P}_{\rm inex} = \begin{bmatrix} I_n & 0 \\ B & I_m \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -H_{\rm inex} \end{bmatrix} \begin{bmatrix} I_n & B^\top \\ 0 & I_m \end{bmatrix}, \tag{1.4}$$

where $H_{\rm inex}$ is a symmetric and positive definite matrix such that $H_{\rm inex} \approx H$. Our use of the term "inexact" refers to the fact that the operator $H^{-1}$ is not applied exactly, and it is instead replaced by the hopefully computationally (much) cheaper operator $H_{\rm inex}^{-1}$. Therefore, this form of inexactness should not be confused with other inexact approaches, such as that proposed in [4].

In this paper we provide a complete spectral characterization of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ by means of the Weyr canonical form (see below), which highlights the role of the matrix blocks and of the multiple unit eigenvalue. We can thus also derive estimates for the condition number (see below for the definition) of the transformation matrix in the canonical form of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$; these estimates can be used to analyze the convergence of an optimal iterative solver such as GMRES [20]; we refer to [18] for early specialized results in this direction, for $C = 0$. Moreover, we express $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ as a perturbation of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$, and this allows us to track the perturbation of its eigenvalues.

We first discuss the case $C = 0$, and then highlight the difference occurring when $C$ is nonzero. In fact, the original Jordan form can be significantly modified for $C \neq 0$, possibly leading to more favorable properties of the modified preconditioner.

Throughout the paper we shall use the Euclidean norm for vectors and the induced norm for matrices. For a rectangular full rank matrix $X$, we define its condition number $\kappa(X)$ as the ratio of its largest and smallest (nonzero) singular values; if $X$ is square and nonsingular, then $\kappa(X) = \|X\|\,\|X^{-1}\|$. In the following, $A^*$ denotes the conjugate transpose of $A$. Moreover, whenever convenient we shall use the notation $[X; Y]$ for $[X^\top, Y^\top]^\top$, where $X, Y$ are conforming matrices. For $A$ symmetric, the notation $A > 0$ ($A \geq 0$) is used for a positive definite (semidefinite) matrix $A$; $A < 0$ is used for $A$ negative definite.

Finally, we shall make great use of a permutation of the Jordan canonical form, which is called the Weyr canonical form (see, e.g., [22]). The Weyr form may be simpler to derive for matrices already in block form, as in this context, and they may provide better structural insight than the usual Jordan form. The Weyr form amounts to a convenient permutation of the columns of the transformation matrix, so that the entries of the Jordan block are permuted. As an example, the Jordan matrix

$$\begin{bmatrix} \alpha & 1 & & \\ & \alpha & & \\ & & \alpha & 1 \\ & & & \alpha \end{bmatrix}$$

with 2 Jordan blocks of size two both associated with the eigenvalue $\alpha$, can be transformed into the following Weyr form, by permuting the second and third rows and columns:

$$\begin{bmatrix} \alpha & & 1 & \\ & \alpha & & 1 \\ & & \alpha & \\ & & & \alpha \end{bmatrix} \equiv \begin{bmatrix} \alpha I_2 & I_2 \\ & \alpha I_2 \end{bmatrix}.$$

**2. Case $C = 0$. Exact constraint preconditioner.** For $C = 0$ the nonsingularity of $H$ and of $\mathcal{A}$ follows from requiring that $B$ have full row rank. Under this condition, the eigenvalues of the preconditioned coefficient matrix $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ may be derived by considering the general eigenvalue problem

$$\begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \mathcal{P}_{\rm ex} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{2.1}$$

which can be written as

$$\begin{bmatrix} I_n & 0 \\ -B & I_m \end{bmatrix} \begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} I_n & -B^\top \\ 0 & I_m \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} I_n & 0 \\ 0 & -H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \tag{2.2}$$

where the factorization (1.2) is used, and

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} I_n & B^\top \\ 0 & I_m \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The left-hand side matrix yields

$$\begin{bmatrix} I_n & 0 \\ -B & I_m \end{bmatrix} \begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} I_n & -B^\top \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} A & (I-A)B^\top \\ B(I-A) & -B(2I-A)B^\top \end{bmatrix}. \tag{2.3}$$

After changing sign in the second block row in (2.2), we can write (2.3) as

$$\begin{aligned} \begin{bmatrix} A & (I-A)B^\top \\ -B(I-A) & B(2I-A)B^\top \end{bmatrix} &= \begin{bmatrix} (A-I) & (I-A)B^\top \\ -B(I-A) & B(I-A)B^\top \end{bmatrix} + \begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix} \\ &= \begin{bmatrix} I_n \\ B \end{bmatrix} (A-I) \begin{bmatrix} I_n, & -B^\top \end{bmatrix} + \begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix}. \end{aligned}$$

Then the eigenvalue problem (2.1) can be transformed into

$$\left( \begin{bmatrix} I_n \\ B \end{bmatrix} (A-I) \begin{bmatrix} I_n, & -B^\top \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix} = (\lambda - 1) \begin{bmatrix} I_n & 0 \\ 0 & H \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},$$

or equivalently

$$\mathcal{M}_0 w = (\lambda-1) w \iff \left( \begin{bmatrix} I_n \\ \widehat{B} \end{bmatrix} (A-I)[I_n, -\widehat{B}^\top] \right) \begin{bmatrix} u \\ \widehat{v} \end{bmatrix} = (\lambda - 1) \begin{bmatrix} u \\ \widehat{v} \end{bmatrix}, \tag{2.4}$$

where $\widehat{B} = H^{-\frac{1}{2}} B$ and $\widehat{v} = H^{\frac{1}{2}} v$.

For later results it is important to note that the matrix $\widehat{B}^\top = B^\top H^{-\frac{1}{2}}$ has orthonormal columns, so that $(I - \widehat{B}^\top \widehat{B})$ is an orthogonal projector.

The following proposition describes the Weyr canonical form of the matrix $\mathcal{M}_0$, which allows us to derive a clear block structure of the transformation matrix. The complete Weyr canonical form of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ can be then easily derived, as shown in the subsequent theorem. To the best of our knowledge, this appears to be the first time the whole canonical form of the exactly preconditioned matrix is derived.

PROPOSITION 2.1. *Let* $(A-I)(I-\widehat{B}^\top\widehat{B})\widehat{X} = \widehat{X}\Theta$ *be the partial eigenvalue decomposition of* $(A-I)(I-\widehat{B}^\top\widehat{B})$ *associated with its nonzero eigenvalues. Then the Weyr decomposition of* $\mathcal{M}_0$ *is given by*

$$\mathcal{M}_0\mathcal{X}_0 = \mathcal{X}_0 \begin{bmatrix} \Theta & & \\ & 0_m & I_m \\ & & 0_m \end{bmatrix}, \qquad \mathcal{X}_0 = \left[\begin{array}{c|c|c} \widehat{X} & \widehat{B}^\top & (A-I)^{-1}\widehat{B}^\top \\ \widehat{B}\widehat{X} & I_m & 0_m \end{array}\right], \quad (2.5)$$

*with* $\mathcal{X}_0$ *nonsingular.*

*Proof.* The decomposition can be verified by substitution. We complete the proof by explicitly deriving $\mathcal{X}_0^{-1}$. Let $\widehat{Z}_0$, $\widehat{Z}$ be the matrices of left eigenvectors of $(A-I)(I-\widehat{B}^\top\widehat{B})$ associated with zero and nonzero eigenvalues, respectively. In particular, we recall that $\widehat{B}^\top$ has orthonormal columns and we notice that $\widehat{Z}_0^*\widehat{B}^\top$ must be nonsingular, since $[\widehat{Z}_0, \widehat{Z}]$ is full rank and the columns of $\widehat{Z}$ span range$(I-\widehat{B}^\top\widehat{B})$. Then, using the full rank assumption of $\widehat{B}(A-I)^{-1}\widehat{B}^\top$, $\mathcal{X}_0^{-1}$ is given by

$$\mathcal{X}_0^{-1} = \begin{bmatrix} \Theta^{-1}(\widehat{Z}^*\widehat{X})^{-1}\widehat{Z}^*(A-I) & -\Theta^{-1}(\widehat{Z}^*\widehat{X})^{-1}\widehat{Z}^*(A-I)\widehat{B}^\top \\ (\widehat{Z}_0^*\widehat{B}^\top)^{-1}\widehat{Z}_0^*\left(I-(A-I)^{-1}\widehat{B}^\top G^{-1}\widehat{B}\right) & (\widehat{Z}_0^*\widehat{B}^\top)^{-1}\widehat{Z}_0^*(A-I)^{-1}\widehat{B}^\top G^{-1} \\ G^{-1}\widehat{B} & -G^{-1} \end{bmatrix},$$

where $G = \widehat{B}(A-I)^{-1}\widehat{B}^\top$. Explicit multiplication by $\mathcal{X}_0$ verifies the assertion. □

THEOREM 2.2. *With the notation and assumptions of Proposition 2.1, the preconditioned matrix* $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ *admits the following Weyr decomposition:*

$$\mathcal{A}\mathcal{P}_{\rm ex}^{-1}\mathcal{X} = \mathcal{X}\begin{bmatrix} I_{n-m}+\Theta & & \\ & I_m & I_m \\ & & I_m \end{bmatrix}, \quad \mathcal{X} = \left[\begin{array}{c|c|c} \widehat{X} & \widehat{B}^\top & (A-I)^{-1}\widehat{B}^\top \\ 0 & 0_m & B(A-I)^{-1}\widehat{B}^\top \end{array}\right],$$

*with* $\mathcal{X}$ *nonsingular.*

*Proof.* Let $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}z = \lambda z$. The derivation leading to (2.4) shows that

$$\mathcal{X} = \begin{bmatrix} I_n & 0 \\ B & -H^{\frac{1}{2}} \end{bmatrix}\mathcal{X}_0.$$

The result readily follows from recalling that the eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ are given by those of $\mathcal{M}_0$ plus one. □

Direct inspection shows that the inverse of $\mathcal{X}$ can be explicitly written as

$$\mathcal{Y}^* := \mathcal{X}^{-1} = \left[\begin{array}{cc} \widehat{X}^\dagger & -\widehat{X}^\dagger(A-I)^{-1}B^\top M^{-1} \\ \hline H^{\frac{1}{2}}M^{-1}B(A-I)^{-1} & -H^{\frac{1}{2}}M^{-1}B(A-I)^{-2}B^\top M^{-1} \\ \hline 0 & H^{\frac{1}{2}}M^{-1} \end{array}\right]; \quad (2.6)$$

here $M = B(A-I)^{-1}B^\top$, and $\widehat{X}^\dagger$ is the (row) portion of the inverse eigenvector matrix of $(A-I)(I-B^\top H^{-1}B)$ associated with the nonzero eigenvalues, or the properly scaled conjugate transpose left eigenvector matrix.

The subdivision of $\mathcal{X}$ in three block columns, and of $\mathcal{Y}^*$ in the corresponding block rows is used to readily describe the role of each block, which the Weyr decomposition easily emphasizes. The first column block of $\mathcal{X}$ contains all eigenvectors of $\Theta + I_{n-m}$, while the second block collects the $m$ eigenvectors associated with the unit eigenvalue with geometric multiplicity $m$. The third block is associated with the corresponding $m$ generalized eigenvectors, revealing the occurrence of $2\times 2$ Jordan blocks.

Finally, we notice that with the explicit expression of $\mathcal{X}$ and $\mathcal{X}^{-1}$ at hand, it is possible to give estimates for the condition number of $\mathcal{X}$. In particular, we see that upper bounds for this condition number depend on $\|B\|$, and on the condition number of both $\widehat{X}$ and $(A-I)$. Such estimates can be used, together with the Weyr decomposition, for bounding the residual norm of optimal Krylov subspace iterative solvers; see [23, sec. 6] and references therein.

**3. Case $C=0$. Inexact constraint preconditioner.** We start by showing that the inexactly preconditioned problem can be written as a perturbation of the exactly preconditioned one. This formulation will allow us to exploit classical perturbation theory results to derive the desired spectral perturbation bounds.

THEOREM 3.1. *Let $E = H - H_{\rm inex}$. With the notation of the previous section, it holds that*

$$\mathcal{A}\mathcal{P}_{\rm inex}^{-1} = \mathcal{A}\mathcal{P}_{\rm ex}^{-1} + \mathcal{E}, \qquad \textit{with} \qquad \mathcal{E} = -\mathcal{A}\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-1}EH_{\rm inex}^{-1}[B,\,-I_m].$$

*Proof.* We have

$$\mathcal{P}_{\rm inex} = \begin{bmatrix} I_n & B^\top \\ B & E\end{bmatrix} = \mathcal{P}_{\rm ex} + \begin{bmatrix} 0 \\ E\end{bmatrix}[0,\ I_m] = \left(I_{n+m} + \begin{bmatrix} 0 \\ E\end{bmatrix}[0,\ I_m]\mathcal{P}_{\rm ex}^{-1}\right)\mathcal{P}_{\rm ex}.$$

Therefore,

$$\mathcal{A}\mathcal{P}_{\rm inex}^{-1} = \mathcal{A}\mathcal{P}_{\rm ex}^{-1}\left(I_{n+m} + \begin{bmatrix} 0 \\ E\end{bmatrix}[0,\ I_m]\mathcal{P}_{\rm ex}^{-1}\right)^{-1}.$$

Thanks to the Sherman-Morrison formula, and using $\mathcal{P}_{\rm ex}^{-1}\begin{bmatrix} 0 \\ E\end{bmatrix} = \begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-1}E$, we obtain

$$\begin{aligned}\mathcal{A}\mathcal{P}_{\rm inex}^{-1} &= \mathcal{A}\mathcal{P}_{\rm ex}^{-1}\left(I_{n+m} - \begin{bmatrix} 0 \\ E\end{bmatrix}\left(I_m + [0,\,I_m]\mathcal{P}_{\rm ex}^{-1}\begin{bmatrix} 0 \\ E\end{bmatrix}\right)^{-1}[0,\ I_m]\mathcal{P}_{\rm ex}^{-1}\right) \\ &= \mathcal{A}\mathcal{P}_{\rm ex}^{-1} - \mathcal{A}\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-1}EH_{\rm inex}^{-1}[B,\ -I_m],\end{aligned}$$

which is the sought after relation. □

We also notice that

$$\begin{aligned}\|\mathcal{E}\| &\le \|\mathcal{A}\|\left\|\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-\frac12}\right\| \|H^{-\frac12}EH_{\rm inex}^{-1}H^{\frac12}\|\,\|H^{-\frac12}[B,\,-I_m]\| \\ &= \|\mathcal{A}\|\left\|\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-\frac12}\right\|^2 \|H^{-\frac12}EH_{\rm inex}^{-1}H^{\frac12}\|.\end{aligned}$$

By explicitly using $E = H - H_{\rm inex}$ we obtain

$$\|H^{-\frac12}EH_{\rm inex}^{-1}H^{\frac12}\| = \|H^{\frac12}H_{\rm inex}^{-1}H^{\frac12} - I\| = \max_{i=1,\ldots,m}|\lambda_i(HH_{\rm inex}^{-1}) - 1|, \qquad (3.1)$$

where $\lambda_i(HH_{\rm inex}^{-1})$, $i = 1,\dots,m$ are the eigenvalues of $HH_{\rm inex}^{-1}$. The bound for $\|\mathcal{E}\|$ thus provides a clear relation between $\|\mathcal{E}\|$ and the accuracy of $H_{\rm inex}$.

To proceed with the spectral analysis, we must distinguish between the unit and non-unit eigenvalues, since the occurrence of multiple eigenvalues with Jordan blocks requires a refined analysis.

We assume that the eigenvalues of the diagonal matrix $I_{n-m} + \Theta$ in $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ are all distinct. Therefore, we can exploit standard perturbation results to evaluate the perturbation that these simple eigenvalues undergo when $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ is perturbed by $\mathcal{E}$ [24]. For each simple eigenvalue $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1})$ there exists an eigenvalue $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1} + \mathcal{E})$ such that

$$\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1} + \mathcal{E}) = \lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \frac{y^*\mathcal{E}x}{y^*x} + O(\|\mathcal{E}\|^2), \tag{3.2}$$

where $x, y$ are the right and left eigenvectors associated with $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1})$. Since both the right and left eigenvectors are available, namely they are the columns of the first block of $\mathcal{X}$ (cf. Theorem 2.2) and of $\mathcal{Y}$ (cf. (2.6)), the first order term can be explicitly computed, emphasizing the dependence on the accuracy of $H_{\rm inex}$. This is shown in the next proposition.

PROPOSITION 3.2. *Assume the notation of Theorem 3.1 holds, and let* $(A-I)(I-\widehat{B}^\top\widehat{B})\widehat{X} = \widehat{X}\Theta$ *be the partial eigenvalue decomposition of* $(A-I)(I-\widehat{B}^\top\widehat{B})$ *associated with its nonzero eigenvalues, with* $\Theta = \mathrm{diag}(\theta_1,\dots,\theta_{n-m})$. *With the notation of Theorem 2.2, consider a simple eigenvalue* $1+\theta_i$ *of* $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$, $i \in \{1,\dots,n-m\}$, *with right eigenvector* $x_i = \mathcal{X}e_i$ *and left eigenvector* $y_i = \mathcal{Y}e_i$ *(cf. (2.6)), where* $e_i$ *is the ith column of the identity matrix of appropriate dimension. If* $A-I$ *is definite, then it holds that* $y_i^*x_i = 1$ *and*

$$|y_i^*\mathcal{E}x_i| \le |1+\theta_i|\|\widehat{X}e_i\|\,\|e_i^\top\widehat{X}^\dagger\|\,\kappa(A-I)^{\frac12}\max_{j=1,\dots,m}|\lambda_j(HH_{\rm inex}^{-1}) - 1|.$$

*Proof.* The first assertion follows from the definition of $\mathcal{Y}^*$. We notice that $e_i^\top\mathcal{Y}^*\mathcal{A} = (1+\theta_i)e_i^\top\mathcal{Y}^*\mathcal{P}_{\rm ex}$, and that $e_i^\top\mathcal{Y}^* = e_i^\top[\widehat{X}^\dagger,\, -\widehat{X}^\dagger(A-I)^{-1}B^\top M^{-1}]$, $i = 1,\dots,n-m$, where with some abuse of notation, on the right-hand side of the second equality $e_i$ has a new conforming dimension. Analogously, $\mathcal{X}e_i = [\widehat{X}e_i; 0]$. Therefore,

$$\begin{aligned}
e_i^\top\mathcal{Y}^*\mathcal{E}\mathcal{X}e_i &= -e_i^\top\mathcal{Y}^*\mathcal{A}\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-1}EH_{\rm inex}^{-1}[B,\,-I_m]\begin{bmatrix}\widehat{X}e_i \\ 0\end{bmatrix} \\
&= -(1+\theta_i)e_i^\top\mathcal{Y}^*\mathcal{P}_{\rm ex}\begin{bmatrix} B^\top \\ -I_m\end{bmatrix} H^{-1}EH_{\rm inex}^{-1}B\widehat{X}e_i \\
&= -(1+\theta_i)e_i^\top\mathcal{Y}^*\begin{bmatrix} 0 \\ BB^\top\end{bmatrix} H^{-1}EH_{\rm inex}^{-1}B\widehat{X}e_i \\
&= (1+\theta_i)e_i^\top\widehat{X}^\dagger(A-I)^{-1}B^\top M^{-1}BB^\top H^{-1}EH_{\rm inex}^{-1}B\widehat{X}e_i \\
&= (1+\theta_i)e_i^\top\widehat{X}^\dagger(A-I)^{-1}B^\top M^{-1}EH_{\rm inex}^{-1}B\widehat{X}e_i.
\end{aligned}$$

We then write

$$\begin{aligned}
&(A-I)^{-1}B^\top M^{-1}EH_{\rm inex}^{-1}B \\
&\quad= (A-I)^{-1}B^\top H^{-\frac12}\left(H^{-\frac12}B(A-I)^{-1}B^\top H^{-\frac12}\right)^{-1}H^{-\frac12}EH_{\rm inex}^{-1}H^{\frac12}H^{-\frac12}B \\
&\quad= (A-I)^{-1}\widehat{B}^\top\left(\widehat{B}(A-I)^{-1}\widehat{B}^\top\right)^{-1}H^{-\frac12}EH_{\rm inex}^{-1}H^{\frac12}\widehat{B},
\end{aligned}$$

so that, using the fact that $\|\widehat{B}\| = 1$, we obtain

$$\begin{aligned}\|(A-I)^{-1}B^\top M^{-1}EH_{\rm inex}^{-1}B\| &\le \|(A-I)^{-1}\widehat{B}^\top\left(\widehat{B}(A-I)^{-1}\widehat{B}^\top\right)^{-1}\|\,\|H^{-\frac{1}{2}}EH_{\rm inex}^{-1}H^{\frac{1}{2}}\|\\ &\le \kappa(A-I)^{\frac{1}{2}}\max_{i=1,\dots,m}|\lambda_i(HH_{\rm inex}^{-1})-1|,\end{aligned}$$

where in the last inequality we used (3.1). Moreover, for $A-I$ definite, the first factor in the last inequality can be obtained as follows. Assume first that $A-I$ is positive definite. Let $(A-I)^{-\frac{1}{2}}\widehat{B}^\top = U\Sigma V^\top$ be the thin singular value decomposition of $(A-I)^{-\frac{1}{2}}\widehat{B}^\top$; note that $(A-I)^{-\frac{1}{2}}$ may be complex. Then $(A-I)^{-1}\widehat{B}^\top\left(\widehat{B}(A-I)^{-1}\widehat{B}^\top\right)^{-1} = (A-I)^{-\frac{1}{2}}U\Sigma^{-1}V^\top$, with

$$\|\Sigma^{-1}\| = \left(\lambda_{\min}(\widehat{B}(A-I)^{-1}\widehat{B}^\top)\right)^{-\frac{1}{2}} \le \|A-I\|^{\frac{1}{2}}.$$

If $A-I$ is negative definite, then one can get the bound with $I-A$. The theorem's final bound thus follows. ◻

To analyze the perturbation of the unit eigenvalues, we first notice that some of them may not be perturbed at all, as the following theorem shows. Results in the same spirit were obtained in [6, Theorem 2.1] for $C \neq 0$ but for the special choice $E = C$ in our notation.

Theorem 3.3. *Assume that $E = H - H_{\rm inex}$ has $k \le m$ zero eigenvalues. Then $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ has $2k$ unit eigenvalues with geometric multiplicity $k$.*

*Proof.* Let $x \neq 0$ be such that $Ex = 0$. Setting $y = H^{\frac{1}{2}}x$ and using any linear combination of columns from the second block of $\mathcal{X}$, vectors of the form $[\widehat{B}^\top y; 0]$ yield

$$\begin{aligned}\mathcal{E}\begin{bmatrix}\widehat{B}^\top y\\ 0\end{bmatrix} &= -\mathcal{A}\begin{bmatrix}B^\top\\ -I_m\end{bmatrix}H^{-1}EH_{\rm inex}^{-1}H^{\frac{1}{2}}y\\ &\overset{(*)}{=} -\mathcal{A}\begin{bmatrix}B^\top\\ -I_m\end{bmatrix}H^{-\frac{1}{2}}(H^{\frac{1}{2}}H_{\rm inex}^{-1}H^{\frac{1}{2}}y - y)\\ &= -\mathcal{A}\begin{bmatrix}B^\top\\ -I_m\end{bmatrix}H^{-\frac{1}{2}}(y - H^{-\frac{1}{2}}H_{\rm inex}H^{-\frac{1}{2}}y) = 0,\end{aligned}$$

where in $(*)$ we used the fact that $Hx = H_{\rm inex}x$ yields $y = H^{-\frac{1}{2}}H_{\rm inex}H^{-\frac{1}{2}}y$. Therefore,

$$\mathcal{A}\mathcal{P}_{\rm inex}^{-1}\begin{bmatrix}\widehat{B}^\top y\\ 0\end{bmatrix} = \mathcal{A}\mathcal{P}_{\rm ex}^{-1}\begin{bmatrix}\widehat{B}^\top y\\ 0\end{bmatrix} = \begin{bmatrix}\widehat{B}^\top y\\ 0\end{bmatrix}.$$

Since the dimension of the null space of $E$ is equal to $k$, the relations above show that $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ has $k$ eigenvalues equal to one, with corresponding linearly independent eigenvectors. We also notice that the third block column of $\mathcal{X}$ is in the null space of $\mathcal{E}$, namely $\mathcal{E}[(A-I)^{-1}\widehat{B}^\top; B(A-I)^{-1}\widehat{B}^\top] = 0$, so that any $k$ columns of this block represent a set of generalized eigenvectors associated with the unit eigenvalue of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$. ◻

Theorem 3.3 shows that if $H_{\rm inex}$ is spectrally close to $H$, in the sense that their eigenstructure partially coincides, then $E$ is singular and it only partially affects the Jordan structure of the unperturbed problem.

The remaining $2(m-k)$ unit eigenvalues, with $k \geq 0$, may be perturbed by a possibly much larger quantity than for the simple eigenvalues. In particular, for a multiple eigenvalue with all Jordan blocks of size two it holds that (see, e.g., [16],[13])

$$\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1} + \mathcal{E}) = \lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \xi^{\frac{1}{2}} + o(\|\mathcal{E}\|^{\frac{1}{2}}), \tag{3.3}$$

where $\xi$ are the eigenvalues of $Y^*\mathcal{E}X$ (and both positive and negative square roots are used to compute the first order term); the columns of $X, Y$ contain all right and left eigenvectors associated with the unit eigenvalue of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$. In our case, $X$ ($Y$) is the second block column of $\mathcal{X}$ (of $\mathcal{Y}$). Once again, thanks to the explicit expression of $\mathcal{X}$ and $\mathcal{Y}$, all these quantities can be computed analytically, the way it was done in Proposition 3.2. To somewhat limit the computational aspects, we refrain from explicitly writing down these bounds.

Although we do not dwell here with the subject, we mention that [16], [13] also discuss the nonlinear perturbation of eigenvectors, by providing the zero order perturbation term explicitly. A similar procedure could be applied here; however, the results would be so technical in our setting that they might be difficult to exploit in practice.

We conclude with a result that sheds light into the type of first order perturbation induced by $Y^*\mathcal{E}X$ in (3.3).

PROPOSITION 3.4. *Let $X_2, Y_2$ be the second block columns of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Assume that $A - I$ is negative definite. Then the eigenvalues $\xi$ of $Y_2^*\mathcal{E}X_2$ are all real. Moreover, if $E = H - H_{\rm inex} \geq 0$ ($E \leq 0$) then $\xi \geq 0$ ($\xi \leq 0$).*

*Proof.* We show that $Y_2^*\mathcal{E}X_2 = W_1W_2$ with $W_1$, $W_2$ symmetric and $W_1 > 0$. This will ensure that the eigenvalues are real. The definiteness will depend on the definiteness of $W_2$.

Using the definition of $Y_2$, $\mathcal{E}$ and $X_2$, we have

$$\begin{aligned}
Y_2^*\mathcal{E}X_2 &= -H^{\frac{1}{2}}M^{-1}(BB^\top - B(A-I)^{-2}B^\top M^{-1}H)H^{-1}EH_{\rm inex}^{-1}H^{\frac{1}{2}}\\
&= -H^{\frac{1}{2}}M^{-1}(I - B(A-I)^{-2}B^\top M^{-1})EH_{\rm inex}^{-1}H^{\frac{1}{2}}\\
&= -H^{\frac{1}{2}}M^{-1}(M - B(A-I)^{-2}B^\top)M^{-1}EH_{\rm inex}^{-1}H^{\frac{1}{2}}\\
&= \left(-H^{\frac{1}{2}}M^{-1}B(A-I)^{-1}(A-2I)(A-I)^{-1}B^\top M^{-1}H^{\frac{1}{2}}\right)\left(H^{\frac{1}{2}}(H_{\rm inex}^{-1} - H^{-1})H^{\frac{1}{2}}\right)\\
&\equiv W_1W_2.
\end{aligned}$$

Since $A - 2I < A - I < 0$, it follows that $W_1$ is positive definite, while the definiteness of $W_2$ depends on that of $H_{\rm inex}^{-1} - H^{-1}$. □

The sign of the eigenvalues $\xi$ influences the type of first order perturbation of multiple eigenvalues. In particular, if $H \leq H_{\rm inex}$, then all $\xi \leq 0$, so that $\xi^{\frac{1}{2}}$ are purely imaginary. As a consequence, perturbed unit eigenvalues will all have nonzero imaginary part, and no real eigenvalues occur as first order perturbations of the unit eigenvalue with nontrivial Jordan block. This is indeed the case for the perturbed spectrum of Example 4.3 where $H_{\rm inex}$ is obtained by an Algebraic Multigrid operator.

**4. Case $C = 0$. Numerical evidence.** In this section we provide experimental evidence of our theory. We start with two simple examples with $5 \times 5$ matrices, which can be fully replicated.

EXAMPLE 4.1. We consider the matrix (1.1) with

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

| $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1})$ | $\lambda(\mathcal{A}\mathcal{P}_{\rm inex}^{-1})$ | $\lvert\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) - \lambda(\mathcal{A}\mathcal{P}_{\rm inex}^{-1})\rvert$ |
|---|---|---|
| 1 | 0.99495 - 0.14204i | 0.14213 |
| 1 | 0.99495 + 0.14204i | 0.14213 |
| 1 | 1.0051 - 0.071435i | 0.07162 |
| 1 | 1.0051 + 0.071435i | 0.07162 |
| 2 | 1.9798 | 0.02020 |

TABLE 4.1
*Example 4.1. Eigenvalues of the exactly and inexactly preconditioned problem, and their difference.*

We precondition this matrix either with $\mathcal{P}_{\rm ex}$ or with $\mathcal{P}_{\rm inex}$. In the latter case, the matrix $H_{\rm inex}$ occurring in (1.4) is defined as

$$H_{\rm inex} = B \begin{bmatrix} 0.985 & 0 & 0 \\ 0 & 0.99 & 0 \\ 0 & 0 & 0.995 \end{bmatrix} B^\top,$$

yielding a nonsingular $E = H - H_{\rm inex}$, with $\|E\| \approx 4.5\cdot 10^{-2}$ and $\|E\|/\|H\| \approx 5\cdot 10^{-3}$. The eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ and of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ are reported in Table 4.1, together with their absolute difference.

For the simple eigenvalue $\lambda = 2$ the first order perturbation in (3.2) predicts a value $\frac{y^*\mathcal{E}x}{y^*x} \approx -0.020202$, which perfectly matches the actual true perturbation. Here $x$ and $y^*$ are the first column of $\mathcal{X}$ and row in $\mathcal{Y}^*$, respectively, while $\mathcal{E}$ is as defined in Theorem 3.1.

For the multiple unit eigenvalue with 2 Jordan blocks of size 2 each, the computed quantity $\xi^{\frac{1}{2}}$ in (3.3) is given by (to the first significant digits) $\xi^{\frac{1}{2}} = 0.11120$ and $\xi^{\frac{1}{2}} = 0.11097i$. Note that here $\|\mathcal{E}\|^{\frac{1}{2}} = 0.19277$, so that higher order terms will be significantly smaller. As a consequence, the first order perturbation term $\xi^{\frac{1}{2}}$ provides a sufficiently good correction to the estimate.

EXAMPLE 4.2. In this example we consider the same data as in Example 4.1, whereas

$$H_{\rm inex} = B \begin{bmatrix} 0.985 & 0 & 0 \\ 0 & 0.99 & 0 \\ 0 & 0 & 1 \end{bmatrix} B^\top.$$

This choice yields a diagonal and singular matrix $E$, with numerically computed eigenvalues $0, 0.040$. The a-priori perturbation estimate for the simple eigenvalue is again $\frac{y^*\mathcal{E}x}{y^*x} \approx -0.020202$, while for the multiple unit eigenvalue the theory predicts that $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ has a unit eigenvalue with a Jordan block of size 2, and two non-unit eigenvalues, with first order perturbation term equal to $\xi^{\frac{1}{2}} = 0.10050$. These expectations are met in the numerical experiment, as Table 4.2 shows.

EXAMPLE 4.3. **Magnetostatic problem.** We consider a $2088 \times 2088$ linear system stemming from the mixed finite element discretization of the 2D magnetostatic problem; we refer to [17] for a detailed description of this test problem. In this example $n = 1272$, $m = 816$, and the resulting matrix $\mathcal{A}$ was properly scaled so that the matrix $(A - I)$ is nonsingular and negative definite. We approximated $H = BB^\top$ with $H_{\rm inex} = R^\top R$, where $R$ is the upper triangular factor of the incomplete Cholesky factorization of $BB^\top$ computed using the Matlab function `cholinc`

| $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1})$ | $\lambda(\mathcal{A}\mathcal{P}_{\rm inex}^{-1})$ | $\lvert\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) - \lambda(\mathcal{A}\mathcal{P}_{\rm inex}^{-1})\rvert$ |
|---|---|---|
| 1 | 1.0000 | $2.220 \cdot 10^{-16}$ |
| 1 | 1.0000 | $2.220 \cdot 10^{-16}$ |
| 1 | 1.0050 - 0.10141i | 0.10153 |
| 1 | 1.0050 + 0.10141i | 0.10153 |
| 2 | 1.9798 | 0.020198 |

TABLE 4.2

*Example 4.2. Eigenvalues of the exactly and inexactly preconditioned problem, and their difference.*

| tol | $\lVert E\rVert$ | $\lVert E\rVert/\lVert H\rVert$ | $\lVert\mathcal{E}\rVert$ | # $2\times 2$ Jordan blocks in $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ | # $2\times 2$ Jordan blocks in $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ |
|---|---|---|---|---|---|
| $1 \cdot 10^{-4}$ | 1.20 | $1.2 \cdot 10^{-4}$ | $9.88 \cdot 10^{-2}$ | 816 | 93 |
| $5 \cdot 10^{-4}$ | 9.02 | $9.1 \cdot 10^{-4}$ | $5.89 \cdot 10^{-1}$ | 816 | 65 |
| $1 \cdot 10^{-3}$ | 24.65 | $2.5 \cdot 10^{-3}$ | $1.11 \cdot 10^{0}$ | 816 | 56 |
| AMG-MI20 | 164.59 | $1.6 \cdot 10^{-2}$ | $2.68 \cdot 10^{0}$ | 816 | 145 |

TABLE 4.3

*Example 4.3. Relevant quantities for the inexactly preconditioned problem.*

with different dropping threshold [15]. Table 4.3 shows the most relevant quantities for different values of the tolerance `tol` used in `cholinc`. The table also displays the number of Jordan blocks retained after perturbation (we considered as zero all eigenvalues of $E$ less than $10^{-9}$ in modulo). This provides a feeling of the spectral quality of the incomplete Cholesky factorization (cf. Theorem 3.3). In the last table row we also report relevant information when using an Algebraic MultiGrid (AMG) preconditioner (in this experiment we used HSL_MI20 [5], with all default parameters). In spite of a larger perturbation (in norm), the spectral properties of $H$ are better reproduced, leading to a less perturbed spectrum of the preconditioned matrix; note that the number of eigenvalues of $E$ less than $10^{-5}$ in absolute value is even higher, namely 196. For the problem preconditioned with $H_{\rm inex}$ being the incomplete Cholesky factorization, in Figure 4.1 we report the true spectrum of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ ($\circ$ symbol), together with its approximation using the first order expansion in (3.2) and in (3.3) (+ symbol). As expected, we can see that the unit eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ are significantly spread both on the real line and on the complex plane. This behavior is qualitatively well captured by the first order perturbation terms (eigenvalues with + symbol) although only higher order perturbation terms would be able to capture the actual direction of the complex eigenvalues. Finally, we observe that the real eigenvalues of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ stemming from simple eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ are well estimated by the first order term in (3.3): cf. the real interval $[0.2, 0.4]$ in the plots.

Figure 4.2 reveals the special features of the Algebraic Multigrid preconditioner. Under perturbation, the unit eigenvalue of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ does not spread on the real axis, as all corresponding eigenvalues of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ ($\circ$ symbol) have nonzero imaginary part. This behavior is well captured by the first order estimate, as all values $\xi$ in (3.3) are real and negative, so that $\xi^{\frac{1}{2}}$ is purely imaginary. This phenomenon is in agreement with Proposition 3.4, and relies on well-known spectral equivalence properties of AMG which appear to hold, at least numerically, for this matrix; we refer, e.g., to [19] for a thorough discussion on AMG.

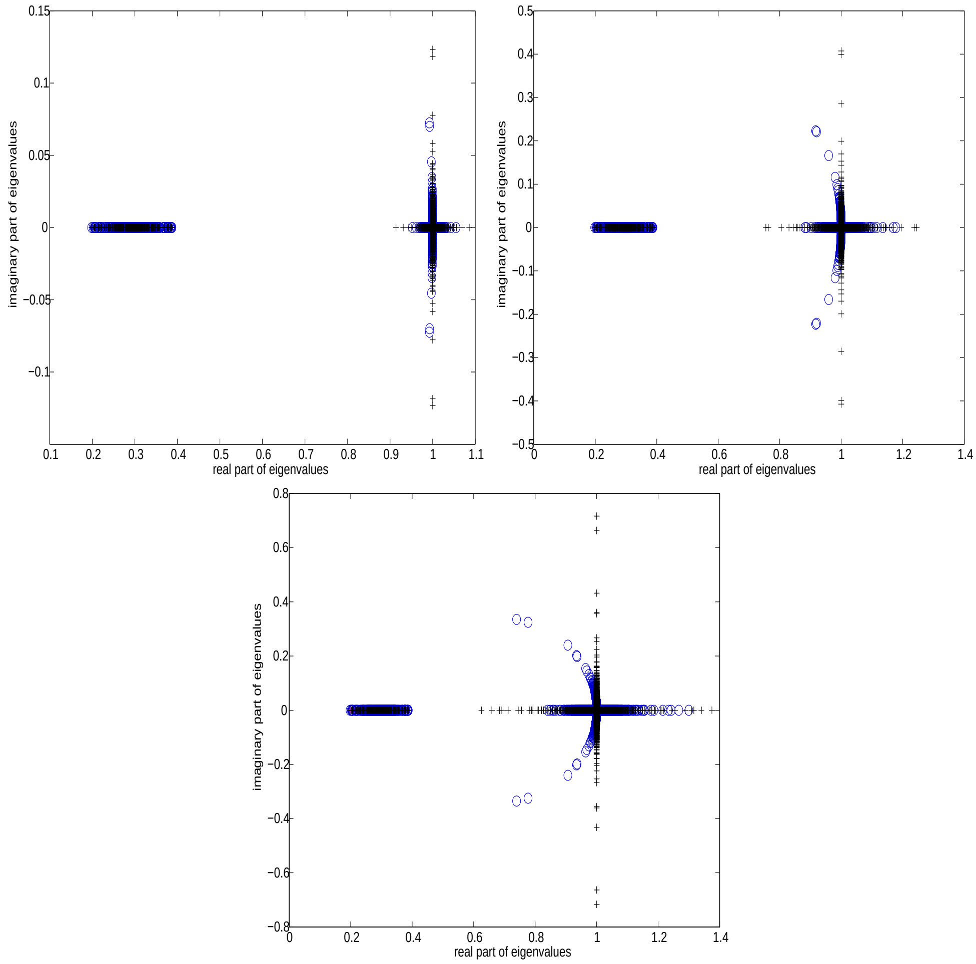


Fig. 4.1. *Example 4.3. Eigenvalues of* $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ *(*$\circ$ *symbol) and approximations (+ symbol) as either* $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \xi^{\frac{1}{2}}$ *or* $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \frac{y^*\mathcal{E}x}{y^*x}$*, as the* `cholinc` *tolerance* `tol` *varies (cf. (3.2), (3.3)). From the top left corner:* `tol` $= 10^{-4}, 5\cdot 10^{-4}, 10^{-3}$.

**5. The case of nonzero (2,2) block.** In this section we generalize our analysis to the case of nonzero (2,2) block in the matrix $\mathcal{A}$. This setting often corresponds to the case when the (1,2) block is column rank deficient, so that the original $\mathcal{A}$ would be singular. We thus assume that $0 < \mathrm{rank}(B^\top) \leq m$; in particular, we exclude the case of $B^\top$ identically zero. Therefore, we can assume that $C$ is symmetric and positive semi-definite, and such that $H$ in (1.3) is nonsingular. We thus define

$$\mathcal{P}_{\rm ex} = \begin{bmatrix} I_n & B^\top \\ B & -C \end{bmatrix}.$$

We first derive the Weyr canonical form for the exactly preconditioned problem, and then derive estimates for the inexactly preconditioned matrix by means of perturbation theory results.

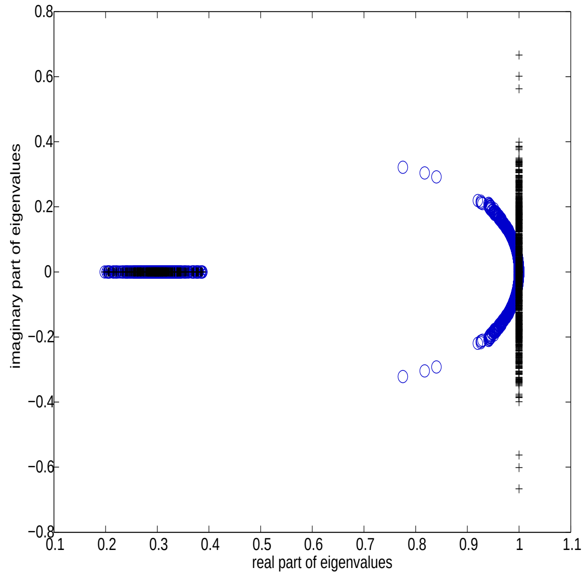


FIG. 4.2. *Example 4.3. Eigenvalues of* $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ *(○ symbol) and approximations (+ symbol) as either* $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \xi^{\frac{1}{2}}$ *or* $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1}) + \frac{y^*\mathcal{E}x}{y^*x}$*, for* $H_{\rm inex}$ *obtained as an Algebraic Multigrid operator.*

Using the same steps as in section 2 we can show that the problem

$$\begin{bmatrix} A & B^\top \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \mathcal{P}_{\rm ex} \begin{bmatrix} x \\ y \end{bmatrix},$$

can be transformed into the equivalent eigenproblem

$$\mathcal{M}_0 w = (\lambda - 1) w \;\Leftrightarrow\; \left( \begin{bmatrix} I_n \\ \widehat{B} \end{bmatrix} (A - I)[I_n, \, -\widehat{B}^\top] \right) \begin{bmatrix} u \\ \widehat{v} \end{bmatrix} = (\lambda - 1) \begin{bmatrix} u \\ \widehat{v} \end{bmatrix},$$

where $\widehat{B} = H^{-\frac{1}{2}} B$, and $u, \widehat{v}$ are related to $x$ and $y$ in the same manner. The only difference lies in the definition of $H$. Clearly, $\widehat{B}$ no longer has all orthonormal rows. The next result generalizes Proposition 2.1 to the new setting. In the following we shall assume that $(A - I)(I - \widehat{B}^\top \widehat{B})$ is diagonalizable. This is not a restrictive condition, since $\mathcal{A}$ can always be scaled to ensure that $A - I$ be definite, from which the diagonalizability follows.

PROPOSITION 5.1. *Let* $C \geq 0$ *and* $H$*, defined as in (1.3), be nonsingular. Let* $[Y_*, Y_1, Y_0]$ *be the eigenvector basis of* $\widehat{B}\widehat{B}^\top$*, whose blocks have dimension* $\ell_*, \ell_1, \ell_0 \geq 0$*, respectively, where* $Y_1$ *and* $Y_0$ *correspond to the unit and zero eigenvalues, if any. Let* $(A - I)(I - \widehat{B}^\top \widehat{B})\widehat{X} = \widehat{X}\Theta$ *be the partial eigenvalue decomposition of* $(A - I)(I - \widehat{B}^\top \widehat{B})$ *associated with its* $n - \ell_1$ *nonzero eigenvalues. Then the Weyr decomposition of* $\mathcal{M}_0$ *is given by*

$$\mathcal{M}_0 \mathcal{X}_0 = \mathcal{X}_0 \begin{bmatrix} \Theta & & & & \\ & 0_{\ell_*} & & & \\ & & 0_{\ell_1} & I_{\ell_1} & \\ & & & 0_{\ell_1} & \\ & & & & 0_{\ell_0} \end{bmatrix},$$

$$\mathcal{X}_0 = \left[\begin{array}{c|c|c|c|c} \widehat{X} & \widehat{B}^\top Y_* & \widehat{B}^\top Y_1 & (A-I)^{-1}\widehat{B}^\top Y_1 & 0 \\ \widehat{B}\widehat{X} & Y_* & Y_1 & 0 & Y_0 \end{array}\right],$$

*with $\mathcal{X}_0$ nonsingular.*

*Proof.* The identity can be easily verified. ◻

We can thus generalize Theorem 2.2 to the case $C \neq 0$. To the best of our knowledge, this is the first time a canonical form is derived for the exact case. The proof is completely analogous to that of Theorem 2.2 and is thus omitted.

THEOREM 5.2. *With the notation and assumptions of Proposition 5.1 and $C \geq 0$, the preconditioned matrix $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ admits the following Weyr decomposition:*

$$\mathcal{A}\mathcal{P}_{\rm ex}^{-1}\mathcal{X} = \mathcal{X}\begin{bmatrix} I_{n-\ell_1}+\Theta & & & & \\ & I_{\ell_*} & & & \\ & & I_{\ell_1} & I_{\ell_1} & \\ & & & I_{\ell_1} & \\ & & & & I_{\ell_0} \end{bmatrix},$$

*with the nonsingular matrix $\mathcal{X}$:*

$$\begin{aligned}\mathcal{X} &= \begin{bmatrix} I_n & 0 \\ B & -H^{\frac{1}{2}} \end{bmatrix}\mathcal{X}_0 \\ &= \left[\begin{array}{c|c|c|c|c} \widehat{X} & \widehat{B}^\top Y_* & \widehat{B}^\top Y_1 & (A-I)^{-1}\widehat{B}^\top Y_1 & 0 \\ 0 & -CH^{-\frac{1}{2}}Y_* & 0 & B(A-I)^{-1}\widehat{B}^\top Y_1 & -H^{\frac{1}{2}}Y_0 \end{array}\right].\end{aligned}$$

The decomposition in Theorem 5.2 shows that the number of Jordan blocks may be significantly low, and in particular less than $m$, if the matrix $\widehat{B}\widehat{B}^\top$ does not have unit eigenvalues, which happens, except for rare cases, if $BB^\top$ and $C$ do not complement each other (if they do[2], then $\widehat{B}^\top\widehat{B}$ is an orthogonal projection onto the range of $\widehat{B}^\top$).

To analyze the perturbation induced by an inaccurate computation of $H$ by means of some $H_{\rm inex}$, we can define $E = H - H_{\rm inex}$ and then write

$$\mathcal{P}_{\rm inex} = \begin{bmatrix} I_n & B^\top \\ B & -C+E \end{bmatrix} = \mathcal{P}_{\rm ex} + \begin{bmatrix} 0 \\ E \end{bmatrix}[0, I_m].$$

Therefore, precisely the same expression as the one in Theorem 3.1 holds for $C \geq 0$:

$$\mathcal{A}\mathcal{P}_{\rm inex}^{-1} = \mathcal{A}\mathcal{P}_{\rm ex}^{-1} + \mathcal{E}, \qquad \text{with } \mathcal{E} = -\mathcal{A}\begin{bmatrix} B^\top \\ -I_m \end{bmatrix} H^{-1}EH_{\rm inex}^{-1}[B, -I_m].$$

As already said, the Weyr decomposition of Theorem 5.2 reveals that unless $BB^\top$ and $C$ complement each other, we expect fewer than $m$ Jordan blocks to arise in general - these affect the third and forth blocks in $\mathcal{X}$ in Theorem 5.2. In terms of spectral perturbation, fewer eigenvalues will be highly perturbed when using $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ in place of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ than what we observed for $C = 0$. However, there are applications where $BB^\top$ and $C$ do complement each other: in this case $I - B^\top H^{-1}B$ is a projector, and $\ell_1 = m - \ell_0$ Jordan blocks in the exactly preconditioned matrix can be found. For these, the discussion of the previous section on their perturbation applies.

[2]In the sense that the ranges of the matrices are complementary spaces.

The number of unit eigenvalues that are left unaltered by $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ depends once again on the number of zero eigenvalues of $E = H - H_{\rm inex}$. In [6, Theorem 2.1] the authors were able to count such number for the special choice $E = C$ in our notation, and with $B$ full rank.

A result analogous to Theorem 3.3 is more difficult to obtain, since we would need to distinguish among the occurrence of unit, zero and other eigenvalues of $\widehat{B}\widehat{B}^\top$. As an example, we can easily see that if

$$\mathrm{Null}(E) \cap \mathrm{Range}\left( H_{\rm inex}^{-1}[B, -I_m] \begin{bmatrix} 0 \\ -H^{\frac{1}{2}} Y_0 \end{bmatrix} \right) \neq \{0\}, \tag{5.1}$$

and $k$ is the dimension of the intersection space, then there will be $k$ unaltered unit eigenvalues in $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ with geometric multiplicity $k$. Analogously, if

$$\mathrm{Null}(E) \cap \mathrm{Range}\left( H_{\rm inex}^{-1}[B, -I_m] \begin{bmatrix} \widehat{B}^\top Y_1 \\ 0 \end{bmatrix} \right) \neq \{0\},$$

and $k$ is the dimension of the intersection space, then there will be $2k$ unaltered unit eigenvalues in $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$ with geometric multiplicity $k$ (note that the forth block of $\mathcal{X}$ is always in the null space of $\mathcal{E}$, hence that block provides the set of generalized eigenvectors). Following the same proving strategy, more extreme cases can be obtained as follows. Recall the definition of $\ell_*, \ell_1$ and $\ell_0$ above. Assume $E = H - H_{\rm inex}$ has $k$ zero eigenvalues. Then

i) If $\ell_* = m$ (i.e. $\ell_0 = \ell_1 = 0$), then $k$ eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ remain unchanged, with geometric multiplicity $k$;

ii) If $\ell_1 = m$ (i.e. $\ell_* = \ell_0 = 0$), then $2k$ eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ remain unchanged, with geometric multiplicity $k$.

We conclude with a numerical evidence for the theory described in this section.

Example 5.3. We consider a variant of Example 4.1:

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 5 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; C = \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Here $B$ is clearly rank deficient, and depending on the values (zero vs. nonzero) of $\gamma_1, \gamma_2$ we can obtain $\ell_1 = 0, 1, 2$; in particular, notice that for $\gamma_1 = \gamma_2 = 0$ the two matrices $BB^\top$ and $C$ complement each other. At first, we consider

$$H_{\rm inex} = B \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.985 & 0 & 0 \\ 0 & 0 & 0.99 & 0 \\ 0 & 0 & 0 & 0.995 \end{bmatrix} B^\top + C.$$

For this choice of $H_{\rm inex}$, the error matrix $E$ is singular (one zero eigenvalue), irrespective of the choice of $C$. Table 5.2 summarizes information on the exact and perturbed eigenvalues. For $\gamma_1 = 1$ and $\gamma_2 = 2$ no Jordan blocks are expected even in the exactly preconditioned problem, since $I - \widehat{B}^\top \widehat{B}$ is not a projector, and it is nonsingular. Other relevant quantities are collected in Table 5.1 for these and other values in $C$. First order terms yield the following linear perturbation of the eigenvalues of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$, which matches pretty well the actual perturbation taking place:

| $\gamma_1, \gamma_2$ | $\|E\|$ | $\|E\|/\|H\|$ | $\|\mathcal{E}\|$ |
|---|---|---|---|
| 1, 2 | 0.045 | 0.00409 | 0.030879 |
| 0, 2 | 0.045 | 0.00409 | 0.044302 |
| 0, 0 | 0.045 | 0.00500 | 0.045235 |

TABLE 5.1

*Example 5.3. Norms of error and perturbation matrices, for various values of the diagonal elements in $C$.*

| $\lambda(\mathcal{A}\mathcal{P}_{\rm ex}^{-1})$ | $y^*\mathcal{E}x$ |
|---|---|
| 3.6888 | -0.003994 |
| 1.8751 | -0.040204 |
| 1.5754 | -0.023661 |
| 1.1880 | -0.015681 |

It can also be seen that (5.1) holds with $k = 1$, therefore one unaltered unit eigenvalue can also be observed.

We next chose $\gamma_1 = 0$ and $\gamma_2 = 2$, so that $C$ is singular. With these choices, $\ell_1 = 1$, so that a Jordan block in the canonical form of $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$ occurs. The theory predicts that the corresponding eigenvalue is perturbed by at least $\xi^{\frac{1}{2}} = 0.12688$ (plus the $o(\|\mathcal{E}\|^{\frac{1}{2}})$ terms), and this can be observed in practice (cf. the second block of rows in Table 5.2). Since (5.1) still holds with $k = 1$, one unaltered unit eigenvalue can also be observed. All remaining eigenvalues show an at most linear perturbation in $\|\mathcal{E}\|$.

Finally, we consider the case $\gamma_1 = \gamma_2 = 0$, so that $BB^\top$ and $C$ complement each other. In this case, $\ell_1 = 2$, so that $B^\top H^{-1}B$ is a projector onto $\mathrm{Range}(B^\top)$. The theory ensures that the eigenvalues of the 2 Jordan blocks are thus perturbed by at least $\xi^{\frac{1}{2}} = 0.13020i, 0.10944$, and this can be verified in Table 5.2. Once again, (5.1) holds with $k = 1$, therefore one unaltered unit eigenvalue can also be observed. All other eigenvalues are perturbed at most linearly.

**6. Conclusions.** In this paper we have derived a spectral canonical form of the constraint preconditioned coefficient matrix $\mathcal{A}\mathcal{P}_{\rm ex}^{-1}$. Moreover, we have analyzed in detail the spectral perturbation induced by the use of the inexact constraint preconditioner when solving large scale symmetric algebraic saddle point problems with zero and nonzero (2,2) block. Our results emphasize the role of the spectral properties of the approximation to the core matrix $H = BB^\top$ (for $C = 0$) to be able to predict the actual distribution of the spectrum of $\mathcal{A}\mathcal{P}_{\rm inex}^{-1}$. Moreover, thanks to an explicit description of the transformation matrix in the canonical form, we were able to track the linear and nonlinear perturbation of the eigenvalues. Our numerical results show that the analysis can be insightful also for data stemming from real applications.

## REFERENCES


[1] Z.-Z. Bai, M. K. Ng, and Z.-Q. Wang, *Constraint preconditioners for symmetric indefinite matrices*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 410–433.

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

[3] M. Benzi and V. Simoncini, *On the eigenvalues of a class of saddle point matrices*, Numerische Mathematik, 103 (2006), pp. 173–196.

[4] L. Bergamaschi, J. Gondzio, M. Venturin, and G. Zilli, *Inexact constraint preconditioners for linear systems arising in interior point methods*, Computational Optimization and Applications, 36 (2007), pp. 137–147.

| $\gamma_1, \gamma_2$ | $\lambda(\mathcal{A}\mathcal{P}_{\text{ex}}^{-1})$ | $\lambda(\mathcal{A}\mathcal{P}_{\text{inex}}^{-1})$ | $\lvert\lambda(\mathcal{A}\mathcal{P}_{\text{ex}}^{-1}) - \lambda(\mathcal{A}\mathcal{P}_{\text{inex}}^{-1})\rvert$ |
|---|---|---|---|
| 1, 2 | 1 | 1.0000 | $2.220 \cdot 10^{-16}$ |
| | 1.0000 | 1.0238 | 0.023763 |
| | 1.0000 | 1.0478 | 0.047758 |
| | 1.1880 | 1.1657 | 0.022336 |
| | 1.5754 | 1.5511 | 0.024283 |
| | 1.8751 | 1.8335 | 0.041568 |
| | 3.6888 | 3.6848 | 0.003960 |
| 0, 2 | 1.0000 | 1.0000 | $3.33 \cdot 10^{-16}$ |
| | 1.0000 | 1.0070 - 0.13768i | 0.13785 |
| | 1.0000 | 1.0070 + 0.13768i | 0.13785 |
| | 1.0000 | 1.0238 | 0.02377 |
| | 1.3820 | 1.3573 | 0.02468 |
| | 1.7273 | 1.6885 | 0.03878 |
| | 3.6180 | 3.6142 | 0.00381 |
| 0, 0 | 1.0000 | 1.0000 | $2.220 \cdot 10^{-16}$ |
| | 1.0000 | 1.0015 - 0.11657i | 0.11657 |
| | 1.0000 | 1.0015 + 0.11657i | 0.11657 |
| | 1.0000 | 0.9946 - 0.16696i | 0.16705 |
| | 1.0000 | 0.9946 + 0.16696i | 0.16705 |
| | 1.3820 | 1.3584 | 0.02356 |
| | 3.6180 | 3.6142 | 0.00382 |

TABLE 5.2

*Example 5.3. Eigenvalues of the exactly and inexactly preconditioned problem, and their difference. The first column shows the choice of the parameters in $C$.*


[5] J. Boyle, M. D. Mihajlović, and J. Scott, *HSL_MI20: an efficient AMG preconditioner for finite element problems in 3D*, Int. J. Numer. Meth. Engnrg,, 82 (2010), pp. 64–98.

[6] H. Dollar, N. Gould, W. Schilders, and A. Wathen, *Using constraint preconditioners with regularized saddle-point problems*, Comput Optim Appl., 36 (2007), pp. 249–270.

[7] H. S. Dollar, *Constraint-style preconditioners for regularized saddle point problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 672–684.

[8] H. S. Dollar, N. I. M. Gould, W. H. A. Schilders, and A. J. Wathen, *Implicit-factorization preconditioning and iterative solvers for regularized saddle-point systems*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 170–189.

[9] C. Durazzi and V. Ruggiero, *Indefinitely preconditioned conjugate gradient method for large sparse equality and inequality constrained quadratic problems*, Num. Linear Algebra with Appl., (2003), pp. 1–24.

[10] R. E. Ewing, R. D. Lazarov, P. Lu, and P. S. Vassilevski, *Preconditioning indefinite systems arising from mixed finite element discretization of second-order elliptic problems*, in Notes in Mathematics, Springer, 1990, pp. 28–43.

[11] J. C. Haws and C. D. Meyer, *Preconditioning KKT systems*, tech. rep., The Boeing Co., M&CT-Tech-01-021, 2003.

[12] C. Keller, N. I. M. Gould, and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1300–1317.

[13] V. B. Lidskii, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Math. and Math. Phys., 1 (1965), pp. 73–85. Also as Zh. vychisl. Mat. mat. Fiz., 6 (1965), pp. 52–60.

[14] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numerical Linear Algebra with Applications, 18 (2011), pp. 1–40.

[15] The MathWorks, Inc., *MATLAB 7*, September 2004.

[16] J. Moro, J. V. Burke, and M. L. Overton, *On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.

[17] I. Perugia and V. Simoncini, *Block–diagonal and indefinite symmetric preconditioners for*

*mixed finite element formulations*, Numerical Linear Algebra with Applications, 7 (2000), pp. 585–616.

[18] M. Rozložník and V. Simoncini, *Krylov subspace methods for saddle point problem with indefinite preconditioning*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 368–391.

[19] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, Frontiers Appl. Math., SIAM, Philadelphia, 1987.

[20] Y. Saad, *Iterative methods for sparse linear systems*, SIAM, Society for Industrial and Applied Mathematics, 2nd ed., 2003.

[21] J. Schöberl and W. Zulehner, *Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 752–773.

[22] H. Shapiro, *The Weyr characteristic*, The American Mathematical Monthly, 106 (1999), pp. 919–929.

[23] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Num. Lin. Alg. with Appl., 14 (2007), pp. 1–59.

[24] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.