Two block preconditioners for a class of double saddle point linear systems

Fariba Bakrani Balani, Masoud Hajarian* and Luca Bergamaschi †

Abstract

Two different block preconditioners are introduced and studied for solving a class of double saddle point linear systems. We provide an eigenvalue analysis for the preconditioned matrices, with special focus on their inexact variants, and give some numerical experiments to illustrate the performance of the studied preconditioning techniques.

AMS classification 2010: 65F10, 65F50, 56F08.

Keywords: Double saddle point problems. Splitting. Preconditioning. Krylov subspace methods.

1. Introduction

The aim of this paper is to introduce two new preconditioners for a class of double saddle point problems. More exactly, let us consider the solution of linear system of the type

$$\mathcal{A}z \equiv \begin{pmatrix} A & B^T & 0 \\ -B & 0 & -C^T \\ 0 & C & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} f \\ g \\ h \end{pmatrix} \equiv b, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite (SPD) matrix, $B \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{p\times m}$ have full row rank, $f \in \mathbb{R}^n$, $g \in \mathbb{R}^m$ and $h \in \mathbb{R}^p$ are given vectors. This kind of linear systems may arise in many applications such as constrained least squares problems [1], constrained quadratic programming [2], liquid crystal problems [3, 4] and so on; see, e.g. [5–8]. Thus, efficient strategies for the solution of the linear system (1.1) are often needed. According to the following partitioning structure

$$\left(\begin{array}{cc|c} A & B^T & 0 \\ -B & 0 & -C^T \\ \hline 0 & C & 0 \end{array}\right), \qquad \left(\begin{array}{c|cc} A & B^T & 0 \\ \hline -B & 0 & -C^T \\ 0 & C & 0 \end{array}\right), \tag{1.2}$$

it is easy to see that the matrix $\mathcal{A}$ in the linear system of equations (1.1) is of classical saddle point form

$$H = \begin{pmatrix} E & F^T \\ -F & 0 \end{pmatrix} \quad \text{with} \quad E = \begin{pmatrix} A & B^T \\ -B & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & -C \end{pmatrix}, \tag{1.3}$$

or generalized saddle point form

$$K = \begin{pmatrix} A & G^T \\ -G & D \end{pmatrix} \quad \text{with} \quad D = \begin{pmatrix} 0 & -C^T \\ C & 0 \end{pmatrix}, \quad G = \begin{pmatrix} B \\ 0 \end{pmatrix}. \tag{1.4}$$

*Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran. E-mail: farbakrani@gmail.com (F.B. Balani), m_hajarian@sbu.ac.ir (M. Hajarian).

†Department of Civil Environmental and Architectural Engineering, University of Padua, Italy. E-mail: luca.bergamaschi@unipd.it

We can see that the submatrices in (1.3) and (1.4) for this problem have different properties than those of the standard block two by two saddle point problems [9, 10]. Due to the special structure of the matrix $\mathcal{A}$, it is important to study some viable numerical solutions to solve the linear system (1.1). Since such these saddle point matrices are typically large and sparse, the efficient solution by the iterative methods is crucial for the overall performance. Among all the iterative methods currently available, Krylov subspace methods are successfully used for solving large and sparse linear systems. However, computational efficiency and high-quality of this class of methods is related to the choice of an appropriate preconditioning strategies.

There have been several effective preconditioners for the linear system (1.1) presented and studied in the literature. An exact block diagonal preconditioner $\mathcal{P}_D$ and its inexact version $\widehat{\mathcal{P}}_D$ have been studied in [11, 12], which are defined as

$$\mathcal{P}_D = \begin{pmatrix} A & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & CS^{-1}C^T \end{pmatrix}, \qquad \widehat{\mathcal{P}}_D = \begin{pmatrix} \hat{A} & 0 & 0 \\ 0 & \hat{S} & 0 \\ 0 & 0 & C\hat{S}^{-1}C^T \end{pmatrix}. \tag{1.5}$$

Here $S = BA^{-1}B^T$, $\hat{A}$ and $\hat{S}$ are SPD approximations of A and S, respectively. Although the exact form of the block diagonal preconditioner in (1.5) has nice eigenvalue properties, constructing the preconditioner $\mathcal{P}_D$ can be a time consuming process.

The shift-splitting preconditioner, originally proposed by Bai et al. [13] for non-Hermitian positive definite linear systems, was later used to precondition the saddle point problems [14] and the generalized saddle point problems [15]. For linear system (1.1), the suggestion of Cao [16] is a shift-splitting preconditioner $\mathcal{P}_{SS}$ and a relaxed version of the shift-splitting preconditioner $\mathcal{P}_{RSS}$ which are structured as follows:

$$\mathcal{P}_{SS} = \frac{1}{2}\begin{pmatrix} \alpha I + A & B^T & 0 \\ -B & \alpha I & -C^T \\ 0 & C & \alpha I \end{pmatrix}, \qquad \mathcal{P}_{RSS} = \frac{1}{2}\begin{pmatrix} A & B^T & 0 \\ -B & \alpha I & -C^T \\ 0 & C & \alpha I \end{pmatrix}, \tag{1.6}$$

where α is a positive constant and I is the identity matrix with proper size. The preconditioner $\mathcal{P}_{SS}$ is obtained based on a splitting of double saddle point problem (1.1) which results in an unconditionally convergent stationary iterative method. In [17], the equivalent symmetric linear system of (1.1) with coefficient matrix

$$\mathcal{B} = \begin{pmatrix} A & B^T & 0 \\ B & 0 & C^T \\ 0 & C & 0 \end{pmatrix}, \tag{1.7}$$

is considered and three exact block preconditioners are introduced. It is also shown in [17] that the preconditioned matrices have only at most three distinct eigenvalues. By introducing two variable parameters, Huang [18] presented a variant of Uzawa iterative method for the problem (1.7). Furthermore, Huang et al. [19] extended the well-known Uzawa method for solving the linear system with coefficient matrix $\mathcal{B}$ of (1.7). They also proposed the inexact Uzawa method and showed that the proposed inexact method is more effective than the exact one. However, finding a practical way of choosing parameters is often difficult in this method. More recently, Wang and Li [20] proposed an exact parameterized block SPD preconditioner to solve the double saddle point linear system (1.1). In addition, the inexact version of the proposed preconditioner has been examined.

The purpose of this work is to propose two block preconditioners for the solution of linear system of equations (1.1) which can be effective for the Krylov subspace methods. Motivated by the preconditioner suggested in [21], we construct a splitting preconditioner for solving (1.1) and show that the corresponding iterative method is unconditionally convergent to the solution of linear system (1.1). Moreover, a new block preconditioner based on the block diagonal preconditioner $\mathcal{P}_D$ and relaxed shift-splitting preconditioner $\mathcal{P}_{RSS}$, is also constructed.

The sketch of the remaining sections is as follows. After introducing two block preconditioners in section 2, we give the implementation details of the application of the preconditioners in conjunction with a Krylov subspace method, such as GMRES. In section 3 we provide the convergence property of the splitting iterative method and characterize the eigenvalues of the corresponding preconditioned matrices. In section 4 we show some numerical results, and in section 5 we give our conclusions.

2. Block preconditioners and their implementation details

This section is structured as follows. In section 2.1, the splitting iterative method and the implementation of the preconditioner extracted from the splitting are described. In section 2.2, we introduce another block preconditioner which can be considered for the system (1.1) in order to increase the convergence rate of the Krylov subspace methods.

2.1 The splitting preconditioner

Motivated by the idea of [21], we are interested in the use of a preconditioner of the form

$$\mathcal{P} = \begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 2C & I \end{pmatrix}, \tag{2.1}$$

in an iterative solution method for (1.1). Specially, we are interested in an iteration of the form

$$\begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 2C & I \end{pmatrix} \begin{pmatrix} u^{(k+1)} \\ v^{(k+1)} \\ w^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & C^TC & C^T \\ 0 & C & I \end{pmatrix} \begin{pmatrix} u^{(k)} \\ v^{(k)} \\ w^{(k)} \end{pmatrix} + \begin{pmatrix} f \\ g \\ h \end{pmatrix}, \tag{2.2}$$

which is associated to the splitting

$$\mathcal{A} = \begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 2C & I \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & C^TC & C^T \\ 0 & C & I \end{pmatrix} = \mathcal{P} - \mathcal{R}. \tag{2.3}$$

Equivalently, the iterative method (2.2) can be expressed in closed form as follows

$$z^{(k+1)} = \mathcal{G}z^{(k)} + d, \tag{2.4}$$

where

$$\mathcal{G} = \mathcal{P}^{-1}\mathcal{R} = \begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 2C & I \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & C^TC & C^T \\ 0 & C & I \end{pmatrix}, \tag{2.5}$$

is the iteration matrix and $d = \mathcal{P}^{-1}b$. In general case we can say that any linear iterative method of the form (2.4) can be utilized to define a preconditioner for matrix $\mathcal{A}$ and thus be embedded in a Krylov subspace method; see [22] for further details. We now give the detail of the implementation of the splitting preconditioner $\mathcal{P}$. In order to apply the preconditioner $\mathcal{P}$ combined with a Krylov subspace method, we need to solve the linear system $\mathcal{P}z = r$, i.e.,

$$\begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 2C & I \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}. \tag{2.6}$$

Therefore, in algorithm 1, we describe the application of this preconditioner to a vector:

Algorithm 1 Computation of $z = \mathcal{P}^{-1}r$.

1. Solve $At = r_1$ for t;
2. Solve $(BA^{-1}B^T + C^TC)z_2 = r_2 + Bt$ for z_2;
3. Solve $Az_1 = r_1 - B^Tz_2$ for z_1;
4. Set $z_3 = r_3 - 2Cz_2$.

An application of this preconditioner amounts to one solve with $M = BA^{-1}B^T + C^TC$ and two solves with A. Since the matrices A and M are SPD, we can use the sparse Cholesky factorization or the preconditioned conjugate gradient (PCG) method to solve these linear systems. In general, solving the linear system with the coefficient matrix M is not a trivial task due to the practical difficulty to form the Schur complement matrix $BA^{-1}B^T$. This is generally dense and hence cannot be found explicitly, especially for large problem size. In practice, one possibility is to consider the matrix $\widehat{M} = B\hat{A}^{-1}B^T + C^TC$ with $\hat{A}$ being an approximation of A and solve the related linear system either exactly or inexactly by using some inner iterative methods. Here we may expect that the linear system with the coefficient matrix $\widehat{M}$ can be practically (if not cheaply) solved.

2.2 The new block preconditioner

In this section, to solve the linear system of equations (1.1), based on the block diagonal preconditioner $\mathcal{P}_D$ and relaxed shift-splitting preconditioner $\mathcal{P}_{RSS}$, we construct the new block preconditioner

$$\mathcal{Q} = \begin{pmatrix} A & B^T & 0 \\ 0 & S & -C^T \\ 0 & C & \alpha I \end{pmatrix}, \tag{2.7}$$

with a parameter $\alpha > 0$. We next give the implementation of the preconditioner $\mathcal{Q}$. The application of the preconditioner $\mathcal{Q}$ with a Krylov subspace method requires solving a linear system of the following form

$$\begin{pmatrix} A & B^T & 0 \\ 0 & S & -C^T \\ 0 & C & \alpha I \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}. \tag{2.8}$$

As a result, we summarize the implementation of this preconditioner in the form of the following algorithm.

Algorithm 2 Computation of $z = \mathcal{Q}^{-1}r$.

1. Solve $(BA^{-1}B^T + \frac{1}{\alpha}C^TC)z_2 = r_2 + \frac{1}{\alpha}C^Tr_3$ for z_2;
2. Solve $Az_1 = r_1 - B^Tz_2$ for z_1;
3. Set $z_3 = \frac{1}{\alpha}(r_3 - Cz_2)$.

In the implementation based on the above algorithm, the application of matrices $N = BA^{-1}B^T + \frac{1}{\alpha}C^TC$ and A are required to apply this preconditioner. As we already discussed in previous section, the solution of linear system with coefficient matrix N can be problematic in this preconditioner. However, we can choose an approximation to the matrix N such that the linear system is easy to be solved. Note that in actual computations, if α is too small, then the matrix system becomes ill-conditioned.

3. Convergence analysis and preconditioning properties

In this section, we present the unconditional convergence analysis of the splitting iterative method (2.4). Moreover, we study results concerning the eigenvalues of the preconditioned matrices $\mathcal{P}^{-1}\mathcal{A}$ and $\mathcal{Q}^{-1}\mathcal{A}$. For the general theory of iterative methods, it is well-known that the vector iterates (2.4) are convergent if and only if the spectral radius of the iteration matrix $\mathcal{G}$ is less than unity.

We now state the following lemma that will be used in the convergence analysis of the splitting iterative method. The proof of the lemma is straightforward computations, and is omitted.

Lemma 1. Let

$$\mathcal{P}_1 = \begin{pmatrix} A & B^T & 0 \\ -B & C^TC & 0 \\ 0 & 0 & I \end{pmatrix}, \qquad \mathcal{P}_2 = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 2C & I \end{pmatrix}. \tag{3.1}$$

Then we have $\mathcal{P} = \mathcal{P}_1\mathcal{P}_2$ and

$$\mathcal{P}^{-1} = \mathcal{P}_2^{-1}\mathcal{P}_1^{-1} = \begin{pmatrix} A^{-1} - A^{-1}B^TQ^{-1}BA^{-1} & -A^{-1}BQ^{-1} & 0 \\ Q^{-1}BA^{-1} & Q^{-1} & 0 \\ -2CQ^{-1}BA^{-1} & -2CQ^{-1} & I \end{pmatrix}, \tag{3.2}$$

with $Q = S + C^TC$ and $S = BA^{-1}B^T$.

Based on Lemma 1, we obtain the following result.

Theorem 1. Suppose that $A \in \mathbb{R}^{n\times n}$ is an SPD matrix, $B \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{p\times m}$ are matrices with full row rank. Then the spectral radius of the iteration matrix $\mathcal{G}$ satisfies $\rho(\mathcal{G}) < 1$, meaning that the splitting iterative method (2.4) converges unconditionally to the unique solution of the linear system of equations (1.1).

Proof. In light of the relation (3.2), we get

$$\begin{aligned} \mathcal{G} = \mathcal{P}^{-1}\mathcal{R} &= \begin{pmatrix} A^{-1} - A^{-1}B^TQ^{-1}BA^{-1} & -A^{-1}BQ^{-1} & 0 \\ Q^{-1}BA^{-1} & Q^{-1} & 0 \\ -2CQ^{-1}BA^{-1} & -2CQ^{-1} & I \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & C^TC & C^T \\ 0 & C & I \end{pmatrix} \\ &= \begin{pmatrix} 0 & -A^{-1}BQ^{-1}C^TC & -A^{-1}BQ^{-1}C^T \\ 0 & Q^{-1}C^TC & Q^{-1}C^T \\ 0 & -2CQ^{-1}C^TC + C & I - 2CQ^{-1}C^T \end{pmatrix}. \end{aligned} \tag{3.3}$$

It can be easily shown that

$$\mathcal{G} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & C & I \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & -A^{-1}BQ^{-1}C^T \\ 0 & 0 & Q^{-1}C^T \\ 0 & 0 & I - CQ^{-1}C^T \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & C & I \end{pmatrix}, \tag{3.4}$$

which is similar to

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & -A^{-1}BQ^{-1}C^T \\ 0 & 0 & Q^{-1}C^T \\ 0 & 0 & I - CQ^{-1}C^T \end{pmatrix}. \tag{3.5}$$

Since the iteration matrix $\mathcal{G}$ is similar to the matrix $\mathcal{H}$, we can compute the eigenvalues of $\mathcal{G}$ by computing the eigenvalues of $\mathcal{H}$. From (3.5), we see that the matrix $\mathcal{H}$ has $n + m$ eigenvalues at zero and the remaining ones come from the eigenvalues of matrix $K = I - CQ^{-1}C^T$. To proceed we must study the eigenvalues of the matrix K. The matrix K can be written as follows

$$\begin{aligned} K = I - CQ^{-1}C^T &= I - C(S^{\frac{1}{2}}S^{\frac{1}{2}} + C^TC)^{-1}C^T \\ &= I - (CS^{-\frac{1}{2}})\Big(I + (CS^{-\frac{1}{2}})^T(CS^{-\frac{1}{2}})\Big)^{-1}(CS^{-\frac{1}{2}})^T. \end{aligned} \tag{3.6}$$

Because $\text{rank}(C) = p$ and S is SPD, we know that $\text{rank}(CS^{-\frac{1}{2}}) = p$. Denote the singular value decomposition of the matrix $CS^{-\frac{1}{2}}$ by

$$CS^{-\frac{1}{2}} = U(\Sigma, 0)V^T, \tag{3.7}$$

where $U \in \mathbb{R}^{p\times p}$ and $V \in \mathbb{R}^{m\times m}$ are orthogonal matrices and $\Sigma = \text{diag}(\nu_1, \nu_2, \ldots, \nu_p) \in \mathbb{R}^{p\times p}$ is a diagonal matrix with $\nu_1 \geq \nu_2 \geq \ldots \geq \nu_p > 0$ being the nonzero singular values of $CS^{-\frac{1}{2}}$. Therefore, the matrix K can be computed as follows

$$\begin{aligned} K &= I_p - U(\Sigma, 0)\begin{pmatrix} I_p + \Sigma^2 & 0 \\ 0 & I_{m-p} \end{pmatrix}^{-1}\begin{pmatrix} \Sigma \\ 0 \end{pmatrix} U^T \\ &= U\text{diag}\Big(\frac{1}{1+\nu_1^2}, \frac{1}{1+\nu_2^2}, \ldots, \frac{1}{1+\nu_p^2}\Big)U^T. \end{aligned} \tag{3.8}$$

From (3.8), we deduce that the nonzero eigenvalues of the matrix $\mathcal{H}$ are $\frac{1}{1+\nu_i^2}$ $(i = 1, \ldots, p)$. We therefore obtain

$$\rho(\mathcal{G}) = \rho(\mathcal{H}) = \max_{1\leq i\leq p} \frac{1}{1+\nu_i^2} = \frac{1}{1+\nu_p^2} < 1, \tag{3.9}$$

which completes the proof of theorem. □

Theorem 1 leads to the following result on the eigenvalues of the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$.

Theorem 2. Suppose that $A \in \mathbb{R}^{n\times n}$ is an SPD matrix, $B \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{p\times m}$ are matrices with full row rank. Then the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ has $n + m$ eigenvalues at one and the rest of eigenvalues are expressed by $\frac{\nu_i^2}{1+\nu_i^2}$ $(i = 1, \ldots, p)$, where ν_i is the singular value of $CS^{-\frac{1}{2}}$.

Proof. From Theorem 1, it follows that $n + m$ eigenvalues of $\mathcal{P}^{-1}\mathcal{R}$ are zero and the other ones are of the form $\frac{1}{1+\nu_i^2}$ $(i = 1, \ldots, p)$. Since

$$\mathcal{P}^{-1}\mathcal{A} = I - \mathcal{P}^{-1}\mathcal{R}, \tag{3.10}$$

hence the $n + m$ eigenvalues of the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ are equal to one and the nonunit eigenvalues are obtained from $\frac{\nu_i^2}{1+\nu_i^2}$ $(i = 1, \ldots, p)$. □

Remark 1. From the earlier eigenvalues derivation, it is not difficult to see that the minimal polynomial of $\mathcal{P}^{-1}\mathcal{A}$ takes the form

$$p(x) = (x - 1)\prod_{i=1}^{p}\Big(x - \frac{\nu_i^2}{1+\nu_i^2}\Big).$$

Therefore, the order of the minimum polynomial of $\mathcal{P}^{-1}\mathcal{A}$ is $p + 1$ and from [23], we find that the dimension of the Krylov subspace $\mathscr{K}(\mathcal{P}^{-1}\mathcal{A}, b)$ is $p+1$. Hence, termination of the Krylov subspace methods is guaranteed in at most $p + 1$ iterations.

In the remainder of this section, we discuss the eigenvalues of the preconditioned matrix $\mathcal{Q}^{-1}\mathcal{A}$.

Theorem 3. Let $A \in \mathbb{R}^{n\times n}$ be an SPD matrix, and let $B \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{R}^{p\times m}$ be matrices with full row rank. Then the preconditioned matrix $\mathcal{Q}^{-1}\mathcal{A}$ has $n+m$ eigenvalues at 1, and the remaining p eigenvalues are given by $\frac{\mu_i}{\alpha+\mu_i}$ $(i = 1, 2, \ldots, p)$, where μ_i are the eigenvalues of $CS^{-1}C^T$ and therefore $\mu_i = \nu_i^2$ $(i = 1, \ldots, p)$.

Proof. We know that the eigenvalues of $\mathcal{Q}^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{Q}^{-1}$ are equal, therefore we study the spectrum results of

$$\mathcal{A}\mathcal{Q}^{-1} = \begin{pmatrix} A & B^T & 0 \\ -B & 0 & -C^T \\ 0 & C & 0 \end{pmatrix}\begin{pmatrix} A & B^T & 0 \\ 0 & S & -C^T \\ 0 & C & \alpha I \end{pmatrix}^{-1}. \tag{3.11}$$

It can be directly verified that

$$\mathcal{A}\mathcal{Q}^{-1} = \begin{pmatrix} I & 0 & 0 \\ -BA^{-1} & I & 0 \\ 0 & \Phi & CS^{-1}C^TG^{-1} \end{pmatrix}, \tag{3.12}$$

where $G = \alpha I + CS^{-1}C^T$. As the precise form of Φ is irrelevant for the argument, we do not write it here. It is clear that $n+m$ eigenvalues of $\mathcal{A}\mathcal{Q}^{-1}$ are unit ($\lambda_i = 1$), and the other eigenvalues are determined by its $(3,3)$ block. These p eigenvalues satisfy the generalized eigenvalue problem

$$CS^{-1}C^Tw = \lambda_i(\alpha I + CS^{-1}C^T)w, \tag{3.13}$$

and from which we obtain

$$\lambda_i = \frac{\mu_i}{\alpha + \mu_i} = \frac{\nu_i^2}{\alpha + \nu_i^2}. \tag{3.14}$$

Therefore, we have proved the theorem. □

Remark 2. From Theorems 2 and 3, we deduce that all nonunit eigenvalues of $\mathcal{P}^{-1}\mathcal{A}$ and $\mathcal{Q}^{-1}\mathcal{A}$ lie inside the open complex disk centered at (1,0) with radius 1, which may improve the convergence rate of the Krylov subspace methods.

Remark 3. The eigenvalue distribution previously characterized shows that the two proposed preconditioners are expected to behave in a similar way, when $\alpha = 1$.

3.1 Approximate preconditioner

The application of two proposed preconditioners in the previous sections requires the solution of linear systems with matrices $M = BA^{-1}B^T + C^TC$ or $N = BA^{-1}B^T + \frac{1}{\alpha}C^TC$. Implementation of these tasks rests on the:

1. Explicit computation of A^{-1}
2. Explicit construction of M or N.
3. Solution of a linear system with either M or N, at each iteration of the Krylov solver of choice.

Sub-task 1. will possibly destroy the sparsity in matrices M and N and make the proposed preconditioners completely inefficient. Hence in the numerical experiments, we may work with approximate preconditioner $\hat{A}$; for example it can be approximated by its diagonal/tridiagonal part. This approximation will make the application of the preconditioner more effective. At the same time the nice eigenvalue distribution proved in the previous theorems will not hold anymore. To characterize the eigenvalue distribution in the approximate case, we consider the preconditioner $\widehat{\mathcal{Q}}$ defined as

$$\widehat{\mathcal{Q}} = \begin{pmatrix} \hat{A} & B^T & 0 \\ 0 & \hat{S} & -C^T \\ 0 & C & \alpha I \end{pmatrix}, \tag{3.15}$$

where $\hat{A}$ is an SPD (sparse) approximation of A and $\hat{S} = B\hat{A}^{-1}B^T$ is computed accordingly. Let us define a measure of the error in the $(1,1)$ block approximation as the matrix $\Delta_A = A\hat{A}^{-1} - I$.

The preconditioned matrix reads in this case

$$\begin{aligned}\mathcal{A}\widehat{\mathcal{Q}}^{-1} &= \left[\begin{pmatrix}\hat{A} & B^T & 0\\ -B & 0 & -C^T\\ 0 & C & 0\end{pmatrix} + \begin{pmatrix}A-\hat{A} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\right]\begin{pmatrix}\hat{A} & B^T & 0\\ 0 & \hat{S} & -C^T\\ 0 & C & \alpha I\end{pmatrix}^{-1}\\ &= \begin{pmatrix}I & 0 & 0\\ -B\hat{A}^{-1} & I & 0\\ 0 & \Psi & C\hat{S}^{-1}C^TG^{-1}\end{pmatrix}\\ &+ \begin{pmatrix}\Delta_A & -\Delta_A B^T\hat{S}^{-1}(I-C^TG^{-1}C\hat{S}^{-1}) & -\Delta_A B^T\hat{S}^{-1}C^TG^{-1}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\end{aligned} \tag{3.16}$$

where now $G = \alpha I + C\hat{S}^{-1}C^T$. The eigenvalues of the $(1,1)$ block in (3.16) are exactly characterized by Theorem 3. The second matrix will produce a perturbation of these eigenvalues whose size depends on the norm of the matrix Δ_A.

The next theorem will characterize the eigenvalues of the preconditioned matrix $\mathcal{A}\widehat{\mathcal{Q}}^{-1}$. The real eigenvalues are characterized by the reasoning as in [24], in terms of the eigenvalues of the matrix $\tilde{A} = \hat{A}^{-1/2}A\hat{A}^{-1/2}$, which is similar to the matrix $A\hat{A}^{-1}$.

Notation. We denote with $\gamma_{\min}$ and $\gamma_{\max}$ the extremal eigenvalues of $\tilde{A}$ and γ a generic value of its Rayleigh Quotient: $\gamma = \dfrac{z^T\tilde{A}z}{z^Tz}$ for a given nonzero vector z.

Theorem 4. The real eigenvalues of $\mathcal{A}\widehat{\mathcal{Q}}^{-1}$ lie in the interval $[\lambda_0, 1+\gamma_{\max}]$, where $\lambda_0 > 0$. Moreover, the truly complex eigenvalues lie in a circle of the Gauss plane, with center the real number 1 and radius 1.

Proof. The eigenvalues of $\mathcal{A}\widehat{\mathcal{Q}}^{-1}$ must satisfy the following equations

$$\begin{aligned}Ax + B^Ty &= \lambda(\hat{A}x + B^Ty),\\ -Bx - C^Tz &= \lambda(\hat{S}y - C^Tz),\\ Cy &= \lambda(Cy + \alpha z),\end{aligned} \tag{3.17}$$

which can be expressed in matrix form as $\mathcal{A}\psi = \lambda\widehat{\mathcal{Q}}\psi$:

$$\begin{pmatrix}A & B^T & 0\\ -B & 0 & -C^T\\ 0 & C & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \lambda\begin{pmatrix}\hat{A} & B^T & 0\\ 0 & B\hat{A}^{-1}B^T & -C^T\\ 0 & C & \alpha I\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}. \tag{3.18}$$

Now we define the block diagonal matrix

$$\mathcal{D} = \text{blkdiag}(\hat{A}^{-1/2}, \hat{S}^{-1/2}, I),$$

then the previous generalized eigenvalue problem has the same eigenvalues as $\mathcal{D}\mathcal{A}\mathcal{D}w = \lambda\mathcal{D}\widehat{\mathcal{Q}}\mathcal{D}w$:

$$\begin{pmatrix}\tilde{A} & R^T & 0\\ -R & 0 & -K^T\\ 0 & K & 0\end{pmatrix}\begin{pmatrix}u\\ v\\ t\end{pmatrix} = \lambda\begin{pmatrix}I & R^T & 0\\ 0 & I & -K^T\\ 0 & K & \alpha I\end{pmatrix}\begin{pmatrix}u\\ v\\ t\end{pmatrix}, \tag{3.19}$$

where

$$R = \hat{S}^{-1/2}B\hat{A}^{-1/2}, \quad K = C\hat{S}^{-1/2} \quad \text{and} \quad \psi = \mathcal{D}w.$$

Note that $K^TK = S^{-1/2}C^TCS^{-1/2}$ and $RR^T = I$. The rest of the proof is divided into two parts:

Real eigenvalues. If $\lambda \in \mathbb{R}$ and $\lambda \notin [\gamma_{\min}, \gamma_{\max}]$, then from the first equation of (3.19) we obtain

$$u = (\tilde{A} - \lambda I)^{-1}(\lambda - 1)R^Tv, \tag{3.20}$$

and from the third equation of (3.19) we have

$$t = \frac{1-\lambda}{\lambda\alpha} Kv. \tag{3.21}$$

Substituting the above relations into the second equation of (3.19) yields

$$(1-\lambda)R(\tilde{A}-\lambda I)^{-1}R^T v - \lambda v - \frac{(1-\lambda)^2}{\lambda\alpha}K^T K v = 0. \tag{3.22}$$

Now premultiplying (3.22) with $\frac{v^*}{v^*v}$ results in

$$\frac{v^*(1-\lambda)R(\tilde{A}-\lambda I)^{-1}R^T v}{v^*v} - \lambda - \frac{(1-\lambda)^2}{\lambda\alpha}\frac{v^*K^TKv}{v^*v} = 0. \tag{3.23}$$

Since $\tilde{A}-\lambda I$ is an SPD or minus SPD, there exists an SPD matrix such that $\tilde{A}-\lambda I = J^2$ (or $-J^2$). Defining $s = J^{-1}p$ and $p = R^T v$, we rewrite the first term of (3.23) as

$$(1-\lambda)\frac{s^*s}{p^*p}\frac{p^*p}{v^*v} = (1-\lambda)\frac{s^*s}{s^*(\tilde{A}-\lambda I)s}\frac{v^*RR^Tv}{v^*v} = (1-\lambda)\frac{s^*s}{s^*(\tilde{A}-\lambda I)s} = \frac{1-\lambda}{\gamma-\lambda}. \tag{3.24}$$

After denoting the quantity $\frac{v^*K^TKv}{v^*v}$ by μ, we rewrite (3.23) as

$$q(\lambda) \equiv \frac{1-\lambda}{\gamma-\lambda} - \lambda - \frac{(1-\lambda)^2}{\lambda\alpha}\mu = 0. \tag{3.25}$$

The real eigenvalues of the preconditioned matrix not lying in $[\gamma_{\min}, \gamma_{\max}]$ are the zeros of equation (3.25). To obtain an upper bound for the real eigenvalue we observe that if $\lambda > \gamma_{\max} + 1$, then

$$\frac{1-\lambda}{\gamma-\lambda} - \lambda = \frac{\lambda^2 - (1+\gamma)\lambda + 1}{\gamma-\lambda} < 0, \tag{3.26}$$

since the numerator is positive for $\lambda > \gamma + 1$. Then it follows that $q(\lambda) < 0, \ \forall\, \lambda > \gamma_{\max} + 1$. To obtain a lower bound we consider $\gamma < 1$ and observe that

$$\lim_{\lambda\to 0^+} q(\lambda) = -\infty, \qquad \lim_{\lambda\to\gamma^-} q(\lambda) = +\infty,$$

so there is a zero of $q(\lambda)$, $\lambda_0 \in (0,\gamma)$, which is bounded away from zero.

Complex Eigenvalues. The equations (3.19) can be written as follows

$$\tilde{A}u + R^T v = \lambda(u + R^T v), \tag{3.27}$$
$$-Ru - K^T t = \lambda(v - K^T t), \tag{3.28}$$
$$Kv = \lambda(Kv + \alpha t). \tag{3.29}$$

Multiplying (3.27) by u^* on the left, the transposed conjugate of (3.28) by v on the right and (3.29) by t^* on the left, we derive

$$u^*\tilde{A}u - \lambda\|u\|^2 = (\lambda-1)u^*R^Tv, \tag{3.30}$$
$$-u^*R^Tv + (\bar{\lambda}-1)t^*Kv = \bar{\lambda}\|v\|^2, \tag{3.31}$$
$$(1-\lambda)t^*Kv = \lambda\alpha\|t\|^2, \tag{3.32}$$

We assume that $\lambda \neq 1$. From (3.32), it can be seen that

$$t^*Kv = \frac{\lambda\alpha\|t\|^2}{1-\lambda},$$

and substituting the above relation into (3.31), we get

$$u^* R^T v = -\bar{\lambda}\|v\|^2 + (\bar{\lambda} - 1)\frac{\lambda\alpha\|t\|^2}{1-\lambda}. \tag{3.33}$$

Inserting (3.33) into equation (3.30), we obtain

$$u^*\tilde{A}u - \lambda\|u\|^2 = (\lambda - |\lambda|^2)\,\alpha\|t\|^2 + (\bar{\lambda} - |\lambda|^2)\|v\|^2. \tag{3.34}$$

Let $\lambda = a + ib$. By exploiting the real and imaginary parts of (3.34), we have

$$-b\|u\|^2 = b\alpha\|t\|^2 - b\|v\|^2, \tag{3.35}$$

$$u^*\tilde{A}u - a\|u\|^2 = (a - a^2 - b^2)(\alpha\|t\|^2 + \|v\|^2). \tag{3.36}$$

For truly complex eigenvalues, (3.35) yields ($b \neq 0$)

$$\|v\|^2 = \|u\|^2 + \alpha\|t\|^2. \tag{3.37}$$

Equation (3.36) tells that we can not have complex eigenvalues with negative real part. Moreover, in view of the identity: $|\lambda - 1|^2 - 1 = a^2 + b^2 - 2a$ it is equivalent to

$$u^*\tilde{A}u - a(\|u\|^2 - \alpha\|t\|^2 - \|v\|^2) = -(|\lambda - 1|^2 - 1)(\alpha\|t\|^2 + \|v\|^2), \tag{3.38}$$

and in view of relation (3.37), we write (3.38) as

$$u^*\tilde{A}u + 2a\alpha\|t\|^2 = -(|\lambda - 1|^2 - 1)(\alpha\|t\|^2 + \|v\|^2). \tag{3.39}$$

It follows that

$$|\lambda - 1|^2 = 1 - \frac{u^*\tilde{A}u + 2a\alpha\|t\|^2}{\alpha\|t\|^2 + \|v\|^2} = 1 - \rho,$$

with

$$\rho = \frac{u^*\tilde{A}u + 2a\alpha\|t\|^2}{\alpha\|t\|^2 + \|v\|^2} > 0,$$

and the thesis holds. □

4. Numerical experiments

This section considers the performance of the GMRES and the flexible GMRES (FGMRES) methods with two inexact proposed preconditioners, and compares the performance of these preconditioners with the preconditioners $\widehat{\mathcal{P}}_D$ and $\mathcal{P}_{RSS}$ for solving the linear system (1.1). Here the right-preconditioned GMRES method is utilized. In the tables, the number of iterations and the elapsed CPU time in seconds are denoted by IT and CPU, respectively. The initial guess is taken to be zero vector and the iterations will be stopped whenever

$$\|b - \mathcal{A}z^{(k)}\|_2/\|b\|_2 < \text{tol}.$$

We also report the accuracy of the methods under ERR $= \|z^{(k)} - z^{(*)}\|_2/\|z^{(*)}\|_2$. In the tables, the symbol (–) means that the iterations do not reach the given tolerance or the direct method for solving blocks is out of memory. In all tests, the right-hand side vector b is to be adjusted such that the exact solution of (1.1) is $z^* = (1, 1, \ldots, 1)^T \in \mathbb{R}^{n+m+p}$.

As already mentioned, the computation of the coefficient matrices M and N in Algorithms 1 and 2 is expensive. So, we replace M and N with their approximations $\widehat{M} = B\hat{A}^{-1}B^T + C^TC$ and $\widehat{N} = B\hat{A}^{-1}B^T + \frac{1}{\alpha}C^TC$ that are easy to implement. Here we consider $\hat{A} = \text{diag}(A)$, the diagonal part of A. The parameter of the preconditioner $\mathcal{P}_{RSS}$ is taken as $\alpha = 0.01$.

The numerical experiments presented in this work have been carried out on a computer with the configuration: Intel Core i7-8550U CPU @ 1.80GHz processor and 12 GB RAM using MATLAB 2022a.

Example 1. ([16,17]) As the first example, we consider the linear system of equations (1.1) where the block matrices in this problem are defined as

$$A = \begin{pmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{pmatrix} \in \mathbb{R}^{2p^2 \times 2p^2}, \quad B = \begin{pmatrix} I \otimes F & F \otimes I \end{pmatrix} \in \mathbb{R}^{p^2 \times 2p^2},$$

$$C = E \otimes F \in \mathbb{R}^{p^2 \times p^2},$$

with

$$T = \frac{1}{h^2} \cdot \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{p \times p}, \quad F = \frac{1}{h} \cdot \text{tridiag}(0, 1, -1) \in \mathbb{R}^{p \times p},$$

$$E = \text{diag}(1, p+1, \ldots, p^2 - p + 1),$$

in which the symbol $\otimes$ denotes Kronecker product and $h = \frac{1}{p+1}$ is the discretization size.

In Table 1, we report the numerical results for the GMRES method with the proposed block preconditioners, $\widehat{\mathcal{P}}_D$ and $\mathcal{P}_{RSS}$. In the inexact block diagonal preconditioner $\widehat{\mathcal{P}}_D$, the matrix $\hat{S}$ is taken as $\hat{S} = \text{diag}(B\hat{A}^{-1}B^T)$. The linear subsystems involved in these preconditioners are solved exactly via the Cholesky factorization of the coefficient matrices.

From Table 1, we can conclude that the preconditioners $\widehat{\mathcal{P}}_D$ and $\mathcal{P}_{RSS}$ require high computational cost for solving the linear system of equations (1.1) and the preconditioners $\mathcal{P}$ and $\mathcal{Q}$ display a good performance to speed up the convergence rate of Krylov subspace methods. Apparently, the preconditioner $\mathcal{P}_{RSS}$ is more efficient as it takes less iterations. However, its application cost is not comparable with that of $\mathcal{P}$ and $\mathcal{Q}$ and this is clearly revealed by the CPU times.

Since the Cholesky factorization is unreasonable for large scale problems, it is important to solve the linear subsystems with the coefficient matrices $\widehat{M}$ and $\widehat{N}$ inexactly. Therefore, in Table 2, we report the results corresponding to FGMRES with the block preconditioners $\mathcal{P}$ and $\mathcal{Q}$ for different values of α. We used the PCG method (with $\text{tol}_{\text{PCG}} = 1e-03$) to solve subsystems as a inner iteration in the outer preconditioned FGMRES method. In addition, the incomplete Cholesky factorization of matrices $\widehat{M}/\widehat{N}$ with drop tolerance $\tau = 1e-03$ are employed as preconditioner in the inner PCG iteration. The total number of inner PCG iterations are reported in parenthesis. We observe that the block preconditioners with FGMRES give reasonable numerical results in terms of the number of FGMRES iterations required and CPU time taken to solve this problem, especially for large scale problems.

Table 1: Numerical results of preconditioned GMRES method for Example 1 (tol = 1e − 7).

	$\widehat{\mathcal{P}}_D$		$\mathcal{P}_{RSS}$		$\mathcal{P}$		$\mathcal{Q}$ $(\alpha = 10)$	
Size	IT	CPU	IT	CPU	IT	CPU	IT	CPU
1024	35	0.02	2	0.08	6	0.004	9	0.005
4096	38	0.16	2	0.13	6	0.01	8	0.02
16384	40	1.52	2	2.60	5	0.11	7	0.10
65536	44	88.51	-	-	4	0.45	6	0.40

Table 2: Numerical results of preconditioned FGMRES method for Example 1 (tol = 1e − 10).

	$\mathcal{P}$		$\mathcal{Q}$ $(\alpha = 0.1)$		$\mathcal{Q}$ $(\alpha = 1)$		$\mathcal{Q}$ $(\alpha = 10)$	
Size	IT	CPU	IT	CPU	IT	CPU	IT	CPU
16384	26 (48)	0.19	5 (5)	0.05	8 (8)	0.06	12 (12)	0.07
65536	25 (45)	0.98	5 (5)	0.27	7 (7)	0.35	11 (11)	0.34
262144	16 (28)	5.01	4 (4)	1.73	6 (6)	2.36	10 (10)	2.31
1048576	23 (40)	54.12	4 (4)	33.03	5 (5)	34.96	9 (9)	36.44

We have also tested different values of the parameter α for preconditioner $\mathcal{Q}$ and results are listed in Table 3. It is clear from Theorem 3 that a small value of α makes the eigenvalue of $\mathcal{Q}^{-1}\mathcal{A}$ approaches 1, but the preconditioned GMRES method converges to a low accuracy solution. In this paper, we consider α to be naturally moderate. A similar observation can be found in [21].

Table 3: Numerical results of preconditioned GMRES method for Example 1 with different α.

		$\mathcal{Q}$-GMRES		
Size		$\alpha = 1$	$\alpha = 10$	$\alpha = 20$
16384	IT	4	7	9
	CPU	0.0505	0.0585	0.0665
	ERR	5.4×10^{-4}	1.5×10^{-4}	3.7×10^{-5}
65536	IT	3	6	8
	CPU	0.3227	0.3838	0.4435
	ERR	2.6×10^{-3}	6.4×10^{-4}	1.6×10^{-4}

Regarding the eigenvalue distribution, we plot in Figures 1, 2 and 3 the spectrum of matrices $\mathcal{A}$, $\widehat{\mathcal{P}}_D^{-1}\mathcal{A}$, $\mathcal{P}^{-1}\mathcal{A}$, $\mathcal{Q}^{-1}\mathcal{A}$ (exact preconditioner $\mathcal{Q}$) and $\widehat{\mathcal{Q}}^{-1}\mathcal{A}$ (approximate preconditioner $\mathcal{Q}$) for two different values of α. These Figures show that all the eigenvalues of the preconditioned matrices are well-clustered. From the Figures 2 and 3, we notice how the inexactness in matrix A makes the real eigenvalues spread in the interval $[0, 2]$ and the perturbation transforms the unit eigenvalues into complex ones (yet close to the point $(1, 0)$ in the complex plane).

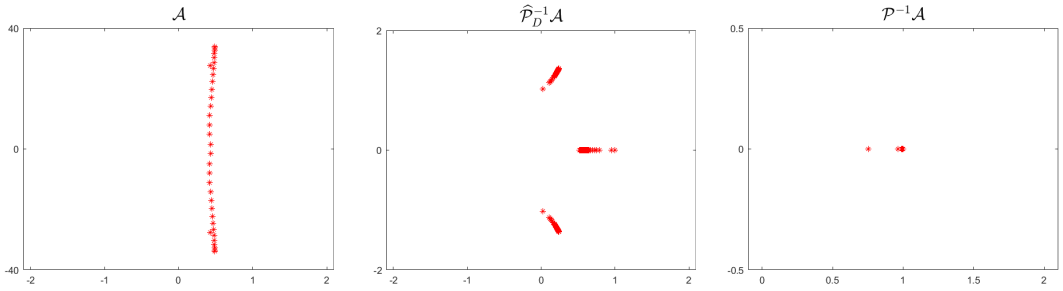

Figure 1: Eigenvalue distribution on the complex plane of the matrix $\mathcal{A}$ (left), preconditioned matrices $\widehat{\mathcal{P}}_D^{-1}\mathcal{A}$ (center) and $\mathcal{P}^{-1}\mathcal{A}$ (right).

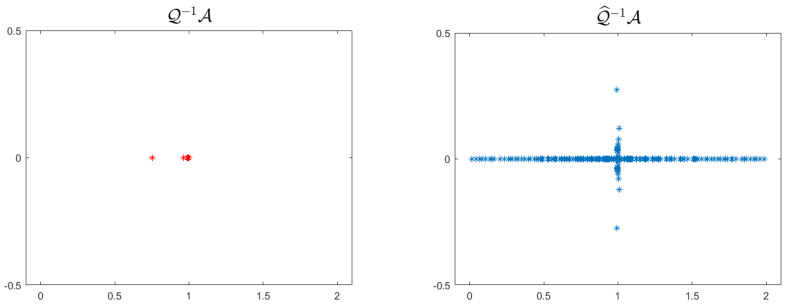

Figure 2: Eigenvalue distribution on the complex plane of the preconditioned matrix with exact $\mathcal{Q}$ (left) and with approximate $\mathcal{Q}$ (right) $\alpha = 1$.

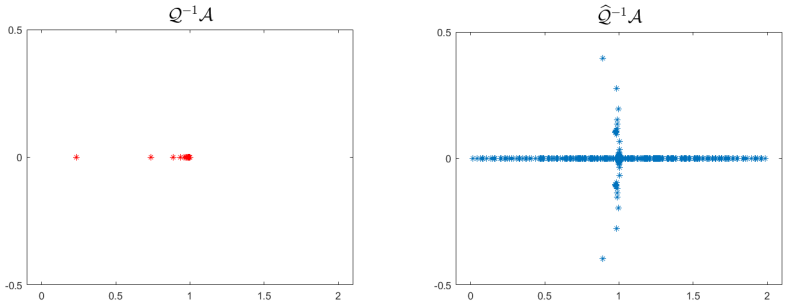

Figure 3: Eigenvalue distribution on the complex plane of the preconditioned matrix with exact $\mathcal{Q}$ (left) and with approximate $\mathcal{Q}$ (right) $\alpha = 10$.

Example 2. ([11,17]) In this example, we consider the linear system of equations (1.1) for which

$$A = \text{diag}\left(2W^TW + D_1, D_2, D_3\right) \in \mathbb{R}^{n\times n},$$

is a block diagonal matrix;

$$B = [E, -I_{2P_1}, I_{2p_2}] \in \mathbb{R}^{m\times n}, \quad \text{and} \qquad C = E^T \in \mathbb{R}^{l\times m},$$

are both full row rank matrices where $p_1 = p^2$, $p_2 = p(p+1)$; $W = (w_{ij}) \in \mathbb{R}^{p_2\times p_2}$ with $w_{ij} = e^{-2((i/3)^2+(j/3)^2)}$; $D_1 = I_{p_2}$ is an identity matrix; $D_i = \text{diag}(d_j^{(i)}) \in \mathbb{R}^{2p_1\times 2p_1}$, $(i = 2, 3)$ are diagonal matrices with

$$d_j^2 = \begin{cases} 1, & \text{for} \quad 1 \le j \le p_1, \\ 10^{-5}(j - p_1^2), & \text{for} \quad p_1 + 1 \le j \le 2p_1, \end{cases}$$
$$d_j^3 = 10^{-5}(j + p_1^2), \qquad \text{for} \quad 1 \le j \le 2p_1;$$

and

$$E = \begin{pmatrix} E_1 \otimes I_p \\ I_p \otimes E_1 \end{pmatrix}, \quad \text{with} \qquad E_1 = \begin{pmatrix} 2 & -1 & & \\ & 2 & -1 & \\ & & \ddots & \ddots \\ & & & 2 \quad -1 \end{pmatrix} \in \mathbb{R}^{p\times(p+1)}.$$

The numerical results of preconditioned GMRES method for tested example are listed in Table 6. In the inexact block diagonal preconditioner $\widehat{\mathcal{P}}_D$, we consider $\hat{A} = \text{diag}(A)$ and $\hat{S} = B\hat{A}^{-1}B^T$. It can be seen numerically that the inexact block diagonal preconditioner $\widehat{\mathcal{P}}_D$, the preconditioners $\mathcal{P}$ and $\mathcal{Q}$ are more convenient than the RSS preconditioner in terms of both iteration number and CPU time. Computing the exact factorization of the Schur complement matrices is time and memory consuming, and make it impossible to grow further the size of the problem. This issue is circumvented by the use of preconditioned iterative solvers in combination with the FGMRES method (as done in the previous example).

Table 5 report the corresponding results which show that we can increase the dimension up to almost 10^7 unknowns and that the proposed preconditioners perform much better than the preconditioners $\widehat{\mathcal{P}}_D$ and $\mathcal{P}_{RSS}$, especially for the large problems.

In a nutshell, the preconditioners $\mathcal{P}$ and $\mathcal{Q}$ in the considered numerical experiments compare favorably with some existing ones. In fact, the proposed preconditioners can be effective and display high convergence rates for the Krylov subspace methods like (F)GMRES.

Table 4: Numerical results of preconditioned GMRES method for Example 2 (tol = 1e − 10).

	$\widehat{\mathcal{P}}_D$		$\mathcal{P}_{RSS}$		$\mathcal{P}$		$\mathcal{Q}$ $(\alpha = 1)$	
Size	IT	CPU	IT	CPU	IT	CPU	IT	CPU
2080	56	0.06	76	0.29	19	0.27	19	0.24
8256	71	0.76	76	4.10	15	0.84	15	0.79
32896	75	5.65	70	77.37	12	1.99	13	1.90
131328	62	69.72	–	–	10	33.05	11	40.26

Table 5: Numerical results of preconditioned FGMRES method for Example 2 (tol = 1e − 10).

	$\widehat{\mathcal{P}}_D$			$\mathcal{P}$			$\mathcal{Q}$ $(\alpha = 1)$		
Size	IT		CPU	IT		CPU	IT		CPU
	outer	inner		outer	inner		outer	inner	
32896	83	332	0.98	12	27	0.11	16	32	0.14
131328	70	279	2.84	10	21	0.28	13	25	0.38
524800	47	188	7.91	8	16	0.98	10	20	1.33
2098176	33	131	20.64	6	13	3.09	6	14	3.25
8390656	18	68	52.42	3	6	8.56	4	8	10.80

As the last experiment, we considered a further approximation in the application of the proposed preconditioners which avoids the presence of two nested iteration. We then modified Algorithms 1 and 2 as described below.

Algorithm 3 Computation of $z = \mathcal{P}^{-1}r$.

1. Solve $At = r_1$ for t;
2. Compute $\hat{A}$ a (cheaply invertible) approximation of A.
3. Compute an incomplete Cholesky factorization L_M of $\widehat{M} = B\hat{A}^{-1}B^T + C^TC$
4. Solve $(L_M L_M^T)z_2 = r_2 + Bt$ for z_2;
5. Solve $Az_1 = r_1 - B^Tz_2$ for z_1;
6. Set $z_3 = r_3 - 2Cz_2$.

Algorithm 4 Computation of $z = \mathcal{Q}^{-1}r$.

1. Compute $\hat{A}$ a (cheaply invertible) approximation of A.
2. Compute an incomplete Cholesky factorization L_N of $\widehat{N} = B\hat{A}^{-1}B^T + \frac{1}{\alpha}C^TC$
3. Solve $(L_N L_N^T)z_2 = r_2 + \frac{1}{\alpha}C^Tr_3$ for z_2;
4. Solve $Az_1 = r_1 - B^Tz_2$ for z_1;
5. Set $z_3 = \frac{1}{\alpha}(r_3 - Cz_2)$.

The results of this last run are reported in Table 6. Inspecting this table we notice a further reduction of the CPU time for the three preconditioners. This is justified by the fact that the outer iteration number only slightly increase with respect to the FGMRES inner/outer solver, but the cost for the preconditioner application is substantially reduced. Moreover these results confirm the superiority of the proposed preconditioner.

Table 6: Numerical results of preconditioned FGMRES method for Example 2 (tol = 1e − 10). The application of the preconditioners is implemented following Algorithms 3 and 4.

	$\widehat{\mathcal{P}}_D$		$\mathcal{P}$		$\mathcal{Q}$ ($\alpha = 1$)	
Size	IT	CPU	IT	CPU	IT	CPU
32896	84	0.75	14	0.08	17	0.09
131328	70	2.31	11	0.21	13	0.23
524800	49	7.13	8	0.60	10	0.65
2098176	33	14.42	6	1.83	7	1.96
8390656	18	38.83	4	8.18	4	7.13

5. Conclusions

In this work we have introduced two block preconditioners for the solution of double saddle point linear systems (1.1). We first showed that the splitting iterative method is unconditionally convergent. Then, the eigenvalues of the corresponding preconditioned matrices were studied. Numerical experiments obtained reveal that the proposed preconditioned (F)GMRES method results in improved computational efficiency for solving double saddle-point problem (1.1) as compared to preconditioners described in the literature.

Future work is aimed at generalizing this work to provide the eigenvalue distribution of more general double saddle-point matrices, in particular those with nonzero $(2, 2)$ and $(3, 3)$ blocks, and to test them on a wide number of realistic applications, such as, e.g., coupled poromechanical models [25], and the coupled Stokes-Darcy equation [3, 26].

Acknowledgements

The authors would thank the referees for their valuable comments and suggestions which improved the quality of the paper. The first author is deeply grateful to the Department of Civil, Environmental and Architectural Engineering at the University of Padua for completing this work. LB gratefully acknowledges the support of the IndAM-GNCS project *Advanced Numerical Linear Algebra and Optimization Techniques for Evolutionary Problems*.

References

[1] J.-Y. Yuan, Numerical methods for generalized least squares problems, J. Comput. Appl. Math., 66 (1996), pp. 571-584.

[2] D.-R. Han, X.-M. Yuan, Local linear convergence of the alternating direction method of multipliers for quadratic programs, SIAM J. Numer. Anal., 51 (2013), pp. 3446-3457.

[3] A. Ramage, E.C. Gartland Jr, A preconditioned nullspace method for liquid crystal director modeling, SIAM J. Sci. Comput., 35 (2013), pp. B226-B247.

[4] F.P.A. Beik, M. Benzi, Iterative methods for double saddle point systems, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 902-921.

[5] Z.-M. Chen, Q. Du, J. Zou, Finite element methods with matching and nonmatching meshes for Maxwell equations with discontinuous coefficients, SIAM J. Numer. Anal., 37 (2000), pp. 1542-1570.

[6] P. Monk, Analysis of a finite element method for Maxwell's equations, SIAM J. Numer. Anal., 29 (1992), pp. 714-729.

[7] P. Degond, P.A. Raviart, An analysis of the Darwin model of approximation to Maxwell's equations, Forum Math., 4 (1992), pp. 13-44.

[8] K. E. Holter, M. Kuchta, K. A. Mardal, Robust preconditioning for coupled Stokes-Darcy problems with the Darcy problem in primal form, Comput. Math. Appl., 91 (2021), pp. 53-66.

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

[10] Z.-Z. Bai, Structured preconditioners for nonsingular matrices of block two-by-two structures, Math. Comp., 75 (2006), pp. 791-815.

[11] N. Huang, C.-F. Ma, Spectral analysis of the preconditioned system for the 3×3 block saddle point problem, Numer. Algor., 81 (2019), pp. 421-444.

[12] S. Bradley, C. Greif, Eigenvalue bounds for double saddle-point systems, arXiv:2110.13328 [math.NA], 2021.

[13] Z.-Z. Bai, J.-F. Yin, Y.-F. Su, A shift-splitting preconditioner for non-Hermitian positive definite matrices, J. Comput. Math., 24 (2006), pp. 539-552.

[14] Y. Cao, J. Du, Q. Niu, Shift-splitting preconditioners for saddle point problems, J. Comput. Appl. Math., 272 (2014), pp. 239-250.

[15] Q.-Q. Shen, Q. Shi, Generalized shift-splitting preconditioners for nonsingular and singular generalized saddle point problems, Comput. Math. Appl., 72 (2016), pp. 632-641.

[16] Y. Cao, Shift-splitting preconditioners for a class of block three-by-three saddle point problems, Appl. Math. Lett., 96 (2019), pp. 40-46.

[17] X. Xie, H.-B. Li, A note on preconditioning for the 3×3 block saddle point problem, Comput. Math. Appl., 79 (2020), pp. 3289-3296.

[18] N. Huang, Variable parameter Uzawa method for solving a class of block three-by-three saddle point problems, Numer. Algor., 85 (2020), pp. 1233-1254.

[19] N. Huang, Y.-H. Dai, Q.-Y. Hu, Uzawa methods for a class of block three-by-three saddle-point problems, Numer. Linear Algebra Appl., 26 (2019), e2265.

[20] N.-N. Wang, J.-C. Li, On parameterized block symmetric positive definite preconditioners for a class of block three-by-three saddle point problems, Comput. Appl. Math., 405 (2022), 113959.

[21] Y. Cao, M.-Q. Jiang, Y.-L. Zheng, A splitting preconditioner for saddle point problems, Numer. Linear Algebra Appl., 18 (2011), pp. 875-895.

[22] M. A. Olshanskii, E. E. Tyrtyshnikov, Iterative methods for linear systems: theory and applications, SIAM, Philadelphia, 2014.

[23] Y. Saad, Iterative methods for sparse linear systems, second ed., SIAM, Philadelphia, 2003.

[24] L. Bergamaschi, Eigenvalue distribution of constraint-preconditioned symmetric saddle point matrices, Numer. Linear Algebra Appl., 19 (2012), pp. 754-772.

[25] M. Frigo, N. Castelletto, M. Ferronato, Enhanced relaxed physical factorization preconditioner for coupled poromechanics, Comput. Math. Appl., 106 (2022), pp. 27–39.

[26] F.P.A. Beik, M. Benzi., Preconditioning techniques for the coupled Stokes–Darcy problem: spectral and field-of-values analysis. Numer. Math. 150 (2022), pp. 257–298.