Computers and Mathematics with Applications 74 (2017) 2423–2437

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Analysis on inexact block diagonal preconditioners for elliptic PDE-constrained optimization problems[☆]

Na Huang [a,b], Chang-Feng Ma [b,*]

[a] State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

[b] College of Mathematics and Informatics, Fujian Key Laboratory of Mathematical Analysis and Applications, Fujian Normal University, Fuzhou 350117, China

ARTICLE INFO

Article history:
Received 17 February 2017
Received in revised form 7 June 2017
Accepted 11 July 2017
Available online 31 July 2017

Keywords:
PDE-constrained optimization
Saddle point matrices
Preconditioner
Cholesky decomposition
Spectral bound

ABSTRACT

By using the Galerkin finite element method, the distributed control problems can be discretized into a saddle point problem with a coefficient matrix of block three-by-three, which can be reduced to a linear system with lower order. We first introduce a class of inexact block diagonal preconditioners and estimate the lower and upper bounds of positive and negative eigenvalues of the preconditioned matrices, respectively. Based on the Cholesky decomposition of the known matrices, we also analyze a lower triangular preconditioner to accelerate the minimal residual method for the reduced linear system and discuss its real and complex eigenvalues respectively. Moreover, these bounds do not rely on the regularization parameter and the eigenvalues of the matrices in the discrete system. Numerical experiments are also presented to demonstrate the effectiveness and robustness of the two new preconditioners.

1. Introduction

In this work, we shall mainly consider the efficient preconditioning techniques to the following linear elliptic distributed control problems:

$$\begin{aligned} \min_{u,f} \quad & \frac{1}{2}\|u-\hat{u}\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|f\|^2_{L^2(\Omega)}, \\ \text{s.t.} \quad & -\Delta u = f \quad \text{in } \Omega, \\ & \quad u = g \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

where Ω is a domain in $\mathbb{R}^2$ or $\mathbb{R}^3$ with boundary $\partial\Omega$, u is the state, $\hat{u}$ is a desired state, f is the control, g is given Dirichlet boundary data and β is a regularization (or Tikhonov) parameter. The problem (1.1) was originally introduced by Lions in [1]. Nowadays there are already many investigations available in literature for solving the distributed control problems of form (1.1); see [2–9]. Some other related problems, which include control constraints or state constraints, have also been studied in [2,3,10–12].

☆ The project was supported by National Postdoctoral Program for Innovative Talents (Grant No. BX201600182), China Postdoctoral Science Foundation (Grant No. 2016M600141), National Natural Science Foundation of China (Grant No. 11071041) and Fujian Natural Science Foundation (Grant Nos. 02016J01005, 2015J01578).

* Corresponding author.
E-mail address: macf@fjnu.edu.cn (C-F. Ma).

http://dx.doi.org/10.1016/j.camwa.2017.07.018
0898-1221/

By using the discretize-then-optimize approach, the PDE-constrained optimization problem (1.1) can be transformed into a linear system of the saddle point form, see, e.g., [8]. More precisely, by employing the Galerkin finite element method to Poisson equation in (1.1), we can derive the finite dimensional discrete analogue of the minimization problem [8,9]

$$\begin{aligned} \min_{u,f} \quad & \frac{1}{2}u^T Mu - u^T b + \alpha + \frac{\beta}{2} f^T Mf, \\ \text{s.t.} \quad & Ku = Mf + d, \end{aligned} \tag{1.2}$$

where $M \in \mathbb{R}^{n\times n}$ is the mass matrix, $K \in \mathbb{R}^{n\times n}$ is the stiffness matrix (the discrete Laplacian), $\alpha = \frac{1}{2}\|\hat{u}\|^2_{L^2(\Omega)}$, $b \in \mathbb{R}^n$ is the Galerkin projection of the target function $\hat{u}$, and $d \in \mathbb{R}^n$ contains the terms arising from the boundary values of the discrete solution. We should emphasize that both M and K are symmetric positive definite, so the minimization problem (1.2) is a convex optimization problem. By applying the Lagrange multiplier technique to the problem (1.2), it follows the saddle point linear systems:

$$\begin{pmatrix} \beta M & 0 & -M \\ 0 & M & K \\ -M & K & 0 \end{pmatrix} \begin{pmatrix} f \\ u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ b \\ d \end{pmatrix}, \tag{1.3}$$

where v is a vector of Lagrange multipliers, see, e.g., [13]. On account of the large and sparse of the matrices M and K, iterative methods should be more effective and efficient than direct methods. Although such linear system of form (1.3) can be seen as a special case of the standard saddle point problem:

$$\begin{pmatrix} H & D^T \\ D & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} g \\ d \end{pmatrix},$$

which has been extensively investigated in recent years, see, [14–23,30] and the relative references therein, constructing efficient solvers based on the special structures of the coefficient matrix in (1.3) is still necessary. Recently, some researchers have devoted themselves to construct and study reliable preconditioners for solving the saddle point systems of form (1.3); see, [6–9]. In [8], the authors studied the MINRES method incorporated with a block diagonal preconditioner and the projected preconditioned conjugate gradient method coupled with a constraint preconditioner to derive the numerical solution of the system (1.3) efficiently. Besides, a block counter diagonal preconditioner and a block counter tridiagonal preconditioner have been proposed in [9] to precondition the Krylov subspace methods such as GMRES. The authors in [24] designed a class of block diagonal and block triangular preconditioners with appropriate approximation blocks to solve the system (1.3).

From the first equation in (1.3), we can see that $f = \frac{v}{\beta}$. This implies that the $3n \times 3n$ linear system of form (1.3) can be reduced to

$$\mathcal{A}z =: \begin{pmatrix} M & K \\ K & -\dfrac{1}{\beta}M \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix} =: \ell, \tag{1.4}$$

whose size is $2n \times 2n$. In [25], the authors proposed several kinds of preconditioners to solve the system (1.4). As the size of the system (1.4) is smaller than that of (1.3), in this paper, we first introduce a class of inexact block diagonal preconditioners by using the preconditioners of M and the Schur complement matrix $S =: \frac{1}{\beta}M + KM^{-1}K$ and estimate the lower and upper bounds of positive and negative eigenvalues of the preconditioned matrices, respectively. Particularly, for the exact block diagonal preconditioner, all eigenvalues of the preconditioned matrix can be explicitly expressed in terms of the eigenvalues of $M^{-1}K$ after some detailed analysis by taking full advantage of its special structure and the eigenvalue decomposition of $M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$. Moreover, all these eigenvalues are contained in the interval $(-1, \frac{1-\sqrt{5}}{2}) \cup (1, \frac{1+\sqrt{5}}{2})$, which evidently does not rely on the parameter β and the eigenvalues of $M^{-1}K$. In addition, based on the Cholesky decomposition of M, K and S, we will also study a lower triangular preconditioner with bilateral preconditioning and derive the lower and upper bounds of positive and negative eigenvalues of the preconditioned system. Under some suitable conditions, these bounds will be independent of the parameter β and the eigenvalues of $M^{-1}K$ as well.

The organization of this paper is as follows. We present a class of inexact block diagonal preconditioner and derive all the eigenvalues of the preconditioned matrix in Section 2. Then we study another lower triangular preconditioner and give some explicit and sharp estimates for the spectral bounds of the preconditioned system in Section 3. Numerical experiments are presented in Section 4 to show the effectiveness of our methods.

We end this section with an introduction of some notation that will be used in the subsequent analysis. For $H \in \mathbb{R}^{n\times n}$, we shall often write H^{-1}, H^T and $\|H\|$ to denote the inverse, the transpose and the norm of H, respectively. $H_1 \sim H_2$ will be used for describing that H_1 is similar to H_2. In addition, we use $\|x\|$ to denote the norm of any vector $x \in \mathbb{R}^n$, and I a general identity matrix.

2. Block diagonal preconditioner

As it is known, the convergence rate of the Krylov subspace methods, such as MINRES and GMRES, is closely related to the eigenvalues and the eigenvectors of the coefficient matrix in the concerned linear system [26–29]. But the eigenvalues of the matrix $\mathcal{A}$ arising from control problems (1.1) are often not clustered. In this section, we shall introduce and analyze a block diagonal preconditioner based on the preconditioners of M and

$$S = \frac{1}{\beta}M + KM^{-1}K. \tag{2.1}$$

Let P_M and P_S be two symmetric positive definite matrices and the approximations of M and S, respectively. Then we construct a block diagonal preconditioner as follows

$$\mathcal{P}_1 = \begin{pmatrix} P_M & 0 \\ 0 & P_S \end{pmatrix}. \tag{2.2}$$

As the preconditioner $\mathcal{P}_1$ is a block diagonal matrix, the linear system with the coefficient matrix $\mathcal{P}_1$ can be solved relatively easily. For judging the effectiveness of our new preconditioner $\mathcal{P}_1$, we shall derive the bounds for the eigenvalues of the preconditioned matrix $\mathcal{P}_1^{-1}\mathcal{A}$, including the lower bound of the positive eigenvalues and the upper bound of the negative eigenvalues; e.g., see [26,28]. We first transform $\mathcal{P}_1^{-1}\mathcal{A}$ into a symmetric matrix by similarity transformation. It is easy to see that

$$\mathcal{P}_1^{-1}\mathcal{A} = \begin{pmatrix} P_M^{-1}M & P_M^{-1}K \\ P_S^{-1}K & -\frac{1}{\beta}P_S^{-1}M \end{pmatrix},$$

which is similar to

$$\hat{\mathcal{A}} := \begin{pmatrix} P_M^{-\frac{1}{2}}MP_M^{-\frac{1}{2}} & P_M^{-\frac{1}{2}}KP_S^{-\frac{1}{2}} \\ P_S^{-\frac{1}{2}}KP_M^{-\frac{1}{2}} & -\frac{1}{\beta}P_S^{-\frac{1}{2}}MP_S^{-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} \hat{M} & \hat{K} \\ \hat{K}^T & -\frac{1}{\beta}\hat{S} \end{pmatrix}.$$

Let λ and $(x^T, y^T)^T$ be the eigenvalue and eigenvector of $\hat{\mathcal{A}}$, then using

$$\begin{pmatrix} \hat{M} & \hat{K} \\ \hat{K}^T & -\frac{1}{\beta}\hat{S} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix},$$

we can write

$$\begin{cases} \hat{M}x + \hat{K}y = \lambda x, \\ \hat{K}^Tx - \frac{1}{\beta}\hat{S}y = \lambda y. \end{cases} \tag{2.3}$$

In the following, we just need to derive the spectral bounds of $\hat{\mathcal{A}}$ by using (2.3) sufficiently. To this end, we introduce the following basic spectral notations.

Spectral Notation 2.1.

1. $\mathrm{sp}(\hat{M}) \subset [\delta_{\hat{M}}, \Delta_{\hat{M}}]$;
2. $\Delta_{\tilde{S}}$ is the maximum eigenvalue of the matrix $\tilde{S} := P_S^{-1}S$, where S is defined as in (2.1);
3. $\delta_{\hat{S}}$ is the minimum eigenvalue of the matrix $\hat{S}$;
4. $\sigma_{\hat{K}}$ is the maximum singular value of the matrix $\hat{K}$.

Theorem 2.1. *Let $\mathcal{A}$ and $\mathcal{P}_1$ be defined as in (1.4) and (2.2) respectively, then under Spectral Notation 2.1, the eigenvalues of the matrix $\mathcal{P}_1^{-1}\mathcal{A}$ lie in the range*

$$\left[-\Delta_{\tilde{S}}, -\frac{\delta_{\hat{S}}}{\beta}\right] \cup \left[\delta_{\hat{M}}, \frac{\beta\Delta_{\hat{M}} - \delta_{\hat{S}} + \sqrt{(\beta\Delta_{\hat{M}} + \delta_{\hat{S}})^2 + 4\beta^2\sigma_{\hat{K}}^2}}{2\beta}\right].$$

Proof. If $y = 0$, then it follows from the first equation of (2.3) that $\delta_{\hat{M}} \leq \lambda \leq \Delta_{\hat{M}}$. If $y \neq 0$, it suffices to consider $\lambda \notin [\delta_{\hat{M}}, \Delta_{\hat{M}}]$. Clearly $\lambda I - \hat{M}$ is nonsingular, and it follows from (2.3) that $x = (\lambda I - \hat{M})^{-1}\hat{K}y$. Substituting into the second equality in (2.3), it leads to

$$\lambda y = \hat{K}^T(\lambda I - \hat{M})^{-1}\hat{K}y - \frac{1}{\beta}\hat{S}y.$$

Multiplying the above equality from the left with $y^T/(y^Ty)$, it yields that

$$\lambda = \frac{y^T\hat{K}^T(\lambda I - \hat{M})^{-1}\hat{K}y}{y^Ty} - \frac{1}{\beta}\frac{y^T\hat{S}y}{y^Ty}. \tag{2.4}$$

We first consider the case that $\lambda > \Delta_{\hat{M}}$. Clearly, $\lambda I - \hat{M} \geq (\lambda - \Delta_{\hat{M}})I > 0$, which follows that $(\lambda I - \hat{M})^{-1} \leq (\lambda - \Delta_{\hat{M}})^{-1}I$. Combining this with (2.4), we can get

$$\lambda \leq \frac{1}{\lambda - \Delta_{\hat{M}}}\frac{y^T\hat{K}^T\hat{K}y}{y^Ty} - \frac{1}{\beta}\frac{y^T\hat{S}y}{y^Ty} \leq \frac{1}{\lambda - \Delta_{\hat{M}}}\sigma_{\hat{K}}^2 - \frac{\delta_{\hat{S}}}{\beta},$$

which can be equivalently reformulated as

$$\lambda^2 - \left(\Delta_{\hat{M}} - \frac{\delta_{\hat{S}}}{\beta}\right)\lambda - \sigma_{\hat{K}}^2 - \frac{\delta_{\hat{S}}\Delta_{\hat{M}}}{\beta} \leq 0.$$

Solving this quadratic inequality for λ, we derive

$$\frac{\beta\Delta_{\hat{M}} - \delta_{\hat{S}} - \sqrt{(\beta\Delta_{\hat{M}} + \delta_{\hat{S}})^2 + 4\beta^2\sigma_{\hat{K}}^2}}{2\beta} \leq \lambda \leq \frac{\beta\Delta_{\hat{M}} - \delta_{\hat{S}} + \sqrt{(\beta\Delta_{\hat{M}} + \delta_{\hat{S}})^2 + 4\beta^2\sigma_{\hat{K}}^2}}{2\beta}.$$

We can directly check that

$$\frac{\beta\Delta_{\hat{M}} - \delta_{\hat{S}} - \sqrt{(\beta\Delta_{\hat{M}} + \delta_{\hat{S}})^2 + 4\beta^2\sigma_{\hat{K}}^2}}{2\beta} \leq \Delta_{\hat{M}} \leq \frac{\beta\Delta_{\hat{M}} - \delta_{\hat{S}} + \sqrt{(\beta\Delta_{\hat{M}} + \delta_{\hat{S}})^2 + 4\beta^2\sigma_{\hat{K}}^2}}{2\beta}.$$

This with $\lambda > \Delta_{\hat{M}}$ gives the desired upper bound in Theorem 2.1.

Next we consider the case that $\lambda < \delta_{\hat{M}}$. If $\lambda > 0$, then we can assert that $\lambda \geq \delta_{\hat{M}}$. Otherwise, if $0 < \lambda < \delta_{\hat{M}}$, then using $\lambda I - \hat{M} < 0$ we can know that

$$\lambda \leq \frac{y^T\hat{K}^T(\lambda I - \hat{M})^{-1}\hat{K}y}{y^Ty} \leq 0.$$

This is in contradiction with $\lambda > 0$. If $\lambda < 0$, then $\lambda I - \hat{M} \leq -\hat{M} < 0$, which follows $(\lambda I - \hat{M})^{-1} \geq -\hat{M}^{-1}$. This along with (2.4) yields that

$$\lambda \geq -\frac{y^T\hat{K}^T\hat{M}^{-1}\hat{K}y}{y^Ty} - \frac{1}{\beta}\frac{y^T\hat{S}y}{y^Ty} = -\frac{y^T\left(\hat{K}^T\hat{M}^{-1}\hat{K} + \frac{1}{\beta}\hat{S}\right)y}{y^Ty} = -\frac{y^TP_S^{-\frac{1}{2}}SP_S^{-\frac{1}{2}}y}{y^Ty} \geq -\Delta_{\hat{S}}.$$

On the other hand, using $\lambda I - \hat{M} < 0$, we have

$$\lambda \leq -\frac{1}{\beta}\frac{y^T\hat{S}y}{y^Ty} \leq -\frac{\delta_{\hat{S}}}{\beta}.$$

Summing up the above, we complete the proof. □

Incomplete Cholesky decomposition is an effective technology to produce a preconditioner of a matrix. Actually, let $\hat{L}$ and $\hat{R}$ be the incomplete Cholesky decomposition of M and S, respectively, then we can choose $P_M = \hat{L}\hat{L}^T$ and $P_S = \hat{R}\hat{R}^T$. Noticing that

$$\mathcal{P}_1 = \begin{pmatrix} \hat{L}\hat{L}^T & 0 \\ 0 & \hat{R}\hat{R}^T \end{pmatrix} = \begin{pmatrix} \hat{L} & 0 \\ 0 & \hat{R} \end{pmatrix}\begin{pmatrix} \hat{L}^T & 0 \\ 0 & \hat{R}^T \end{pmatrix} =: \hat{\mathcal{P}}_1\hat{\mathcal{P}}_1^T, \tag{2.5}$$

when using SYMMLQ or MINRES to the symmetric linear system (1.4), we can consider the following preconditioned system:

$$\hat{\mathcal{P}}_1^{-1}\mathcal{A}\hat{\mathcal{P}}_1^{-T}\hat{z} = \hat{\mathcal{P}}_1^{-1}b, \quad \hat{z} = \hat{\mathcal{P}}_1^Tz.$$

As the preconditioner $\hat{\mathcal{P}}_1$ is a block matrix, and both $\hat{L}$ and $\hat{R}$ are lower triangular matrices, the system $\hat{z} = \hat{\mathcal{P}}_1^Tz$ can be solved relatively easily.

Particularly, we can choose $P_M = M$ and $P_S = S$ with S being the form of (2.1). Then the preconditioner $\mathcal{P}_1$ has the form of

$$\mathcal{P}_e = \begin{pmatrix} M & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & R \end{pmatrix}\begin{pmatrix} L^T & 0 \\ 0 & R^T \end{pmatrix}, \tag{2.6}$$

where

$$M = LL^T, \qquad K = HH^T, \qquad S = RR^T \tag{2.7}$$

are the Cholesky decomposition of M, K and S, respectively. As the exact structure of the preconditioner $\mathcal{P}_e$ in (2.6), we can derive some more clustered bounds of the eigenvalues of $\mathcal{P}_e^{-1}\mathcal{A}$ than that in Theorem 2.1. Actually, by taking full advantage of its special structure and the eigenvalue decomposition of the symmetric matrix $M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$, we will explicitly derive all eigenvalues of the preconditioned matrix in terms of the eigenvalues of $M^{-1}K$. It is wonderful that all these eigenvalues will be contained within a bounded interval which is independent of β and $M^{-1}K$.

By using (2.6), we can know that

$$\begin{aligned}\mathcal{P}_e^{-1}\mathcal{A} &= \begin{pmatrix} M^{-1} & 0 \\ 0 & \left(\frac{1}{\beta}M + KM^{-1}K\right)^{-1}\end{pmatrix}\begin{pmatrix} M & K \\ K & -\frac{1}{\beta}M\end{pmatrix} \\ &= \begin{pmatrix} I & M^{-1}K \\ \left(\frac{1}{\beta}I + M^{-1}KM^{-1}K\right)^{-1}M^{-1}K & -\frac{1}{\beta}\left(\frac{1}{\beta}I + M^{-1}KM^{-1}K\right)^{-1}\end{pmatrix} \\ &\sim \begin{pmatrix} I & \hat{K} \\ \left(\frac{1}{\beta}I + \hat{K}^2\right)^{-1}\hat{K} & -\frac{1}{\beta}\left(\frac{1}{\beta}I + \hat{K}^2\right)^{-1}\end{pmatrix} =: \mathcal{G}_1,\end{aligned}$$

where $\hat{K} = M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$. By the eigenvalue decomposition of $\hat{K}$, i.e., $\hat{K} = PDP^T$ with $P \in \mathbb{R}^{n\times n}$ being an orthonormal matrix, $D = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)$ and $\mu_i\,(1 \le i \le n)$ the eigenvalues of $\hat{K}$, we can derive that

$$\begin{aligned}\mathcal{G}_1 &= \begin{pmatrix} I & PDP^T \\ \left(\frac{1}{\beta}I + PD^2P^T\right)^{-1}PDP^T & -\frac{1}{\beta}\left(\frac{1}{\beta}I + PD^2P^T\right)^{-1}\end{pmatrix} \\ &\sim \begin{pmatrix} I & D \\ \left(\frac{1}{\beta}I + D^2\right)^{-1}D & -\frac{1}{\beta}\left(\frac{1}{\beta}I + D^2\right)^{-1}\end{pmatrix} \\ &\sim \operatorname{diag}\left(\begin{pmatrix} 1 & \mu_1 \\ \frac{\beta\mu_1}{1+\beta\mu_1^2} & -\frac{1}{1+\beta\mu_1^2}\end{pmatrix}, \ldots, \begin{pmatrix} 1 & \mu_n \\ \frac{\beta\mu_n}{1+\beta\mu_n^2} & -\frac{1}{1+\beta\mu_n^2}\end{pmatrix}\right).\end{aligned}$$

Then it is easy to verify that all the eigenvalues of $\mathcal{G}_1$ are

$$\frac{\theta_i \pm \sqrt{\theta_i^2 + 4}}{2},$$

where $\theta_i = \frac{\beta\mu_i^2}{1+\beta\mu_i^2}$, $i = 1, 2, \ldots, n$. As the preconditioned system $\mathcal{P}_e^{-1}\mathcal{A}$ is similar to $\mathcal{G}_1$, and the matrix $M^{-1}K$ is similar to $\hat{K} = M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$, based on the above analysis, we can get the following theorem immediately.

Theorem 2.2. *Let $\mathcal{A} \in \mathbb{R}^{2n\times 2n}$ be the coefficient matrix of the linear system (1.4) and $\mathcal{P}_e$ the block diagonal preconditioner defined as in (2.6). Then all the eigenvalues of the preconditioned system $\mathcal{P}_e^{-1}\mathcal{A}$ are*

$$\frac{\theta_i \pm \sqrt{\theta_i^2 + 4}}{2},$$

with $\theta_i = \frac{\beta\mu_i^2}{1+\beta\mu_i^2}$ and $\mu_i\,(1 \le i \le n)$ being the eigenvalues of $M^{-1}K$.

For any $\beta > 0$, we have $0 < \theta_i < 1$. Then any positive eigenvalue λ of $\mathcal{P}_e^{-1}\mathcal{A}$ meets the following estimates:

$$1 < \lambda = \frac{1}{2}\left(\theta_i + \sqrt{\theta_i^2 + 4}\right) < \frac{1+\sqrt{5}}{2},$$

and for any negative eigenvalue λ of $\mathcal{P}_e^{-1}\mathcal{A}$, it holds that

$$-1 < \lambda = \frac{\theta_i - \sqrt{\theta_i^2 + 4}}{2} = -\frac{2}{\theta_i + \sqrt{\theta_i^2 + 4}} < \frac{1-\sqrt{5}}{2}.$$

Then we can get the following corollary on the spectral bounds of the preconditioned system $\mathcal{P}_e^{-1}\mathcal{A}$ immediately.

Corollary 2.1. *Under the same settings and conditions as in Theorem 2.2, all the eigenvalues of the preconditioned system $\mathcal{P}_e^{-1}\mathcal{A}$ are contained in* $(-1, \frac{1-\sqrt{5}}{2}) \cup (1, \frac{1+\sqrt{5}}{2})$.

Remark 2.1. We should emphasize that if we use the preconditioned MINRES method coupled with the preconditioner $\mathcal{P}_e$ to solve the linear system (1.4), then the convergence rate is about [28]:

$$\|r_k\| \leq \frac{1}{T_{\left[\frac{k}{2}\right]}(1.3416)}\|r_0\| = \frac{2}{2.2361^{\left[\frac{k}{2}\right]} + 0.4472^{\left[\frac{k}{2}\right]}}\|r_0\|,$$

where r_k is the residual of kth step, and $T_m(x)$ is the Chebyshev polynomial of degree m, i.e.,

$$T_m(x) = \frac{(x+\sqrt{x^2-1})^m + (x-\sqrt{x^2-1})^m}{2}.$$

It means that $\|r_k\|/\|r_0\| \leq 7.3271\text{e}{-10}$ as long as $k \geq 54$.

3. Lower triangular preconditioner

In this section, based on the Cholesky decomposition of M, K and S as done in (2.7), we propose and study the lower triangular preconditioner of the following form for solving the saddle point problem (1.4):

$$\mathcal{P}_2 = \begin{pmatrix} L & 0 \\ H & \alpha R \end{pmatrix}, \tag{3.1}$$

where $\alpha > 0$ is a given parameter. Similarly, we still consider the symmetric preconditioned system:

$$\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}\hat{z} = \mathcal{P}_2^{-1}b, \quad \hat{z} = \mathcal{P}_2^T z.$$

As the preconditioner $\mathcal{P}_2$ is a lower triangular matrix, the system $\hat{z} = \mathcal{P}_1^T z$ can be solved relatively easily as well. In the following, we shall study the spectral bounds of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$. We would emphasize that the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ considered here is much more technical than $\mathcal{P}_1^{-1}\mathcal{A}\mathcal{P}_1^{-T}$ in Section 2. Although we cannot derive all the eigenvalues of the preconditioned $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$, by some effective similarity transformations and the singular value decomposition of $L^{-1}H$, we derive some explicit and sharp estimates for the lower and upper bounds of positive and negative eigenvalues of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$. Moreover, these bounds do not rely on the parameter β and the eigenvalues of $M^{-1}K$ as long as α is sufficiently large.

Firstly, we manage to work out several effective similarity transformations. For convenience in the analysis, we introduce some matrices:

$$W = L^{-1}H, \qquad B = \frac{\alpha^2}{\beta}I + (\alpha WW^T)^2, \qquad \mathcal{L} = \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix}, \tag{3.2}$$

and

$$\widehat{\mathcal{P}}_2 = \begin{pmatrix} I & W^T \\ W & B + WW^T \end{pmatrix}, \qquad \widehat{\mathcal{A}} = \begin{pmatrix} I & WW^T \\ WW^T & -\frac{1}{\beta}I \end{pmatrix}. \tag{3.3}$$

From (1.4), (2.1) and (2.7) and (3.1)–(3.3), we can directly verify that

$$\begin{aligned}
\mathcal{P}_2\mathcal{P}_2^T &= \begin{pmatrix} L & 0 \\ H & \alpha R \end{pmatrix}\begin{pmatrix} L^T & H^T \\ 0 & \alpha R^T \end{pmatrix} = \begin{pmatrix} LL^T & LH^T \\ HL^T & HH^T + \alpha^2 RR^T \end{pmatrix} \\
&= \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} I & W^T \\ W & WW^T + \alpha^2 L^{-1}RR^T L^{-T} \end{pmatrix}\begin{pmatrix} L^T & 0 \\ 0 & L^T \end{pmatrix} \\
&= \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} I & W^T \\ W & \frac{\alpha^2}{\beta}I + (\alpha WW^T)^2 + WW^T \end{pmatrix}\begin{pmatrix} L^T & 0 \\ 0 & L^T \end{pmatrix} \\
&= \mathcal{L}\widehat{\mathcal{P}}_2\mathcal{L}^T,
\end{aligned}$$

and

$$\mathcal{A} = \begin{pmatrix} LL^T & HH^T \\ HH^T & -\frac{1}{\beta}LL^T \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix}\begin{pmatrix} I & WW^T \\ WW^T & -\frac{1}{\beta}I \end{pmatrix}\begin{pmatrix} L^T & 0 \\ 0 & L^T \end{pmatrix} = \mathcal{L}\widehat{\mathcal{A}}\mathcal{L}^T.$$

These two equalities prompt us to apply some similarity transformations to the matrix $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$. Using (3.2) and (3.3) again, we can obtain

$$\begin{aligned}
\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T} &\sim (\mathcal{P}_2\mathcal{P}_2^T)^{-1}\mathcal{A} = (\mathcal{L}\widehat{\mathcal{P}}_2\mathcal{L}^T)^{-1}\mathcal{L}\widehat{\mathcal{A}}\mathcal{L}^T = \mathcal{L}^{-T}\widehat{\mathcal{P}}_2^{-1}\widehat{\mathcal{A}}\mathcal{L}^T \sim \widehat{\mathcal{P}}_2^{-1}\widehat{\mathcal{A}} \\
&= \begin{pmatrix} I & W^T \\ W & B+WW^T \end{pmatrix}^{-1} \begin{pmatrix} I & WW^T \\ WW^T & -\frac{1}{\beta}I \end{pmatrix} \\
&= \begin{pmatrix} I & 0 \\ 0 & B^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} I & W^TB^{-\frac{1}{2}} \\ B^{-\frac{1}{2}}W & I+B^{-\frac{1}{2}}WW^TB^{-\frac{1}{2}} \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ 0 & B^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} I & WW^T \\ WW^T & -\frac{1}{\beta}I \end{pmatrix} \\
&\sim \begin{pmatrix} I & W^TB^{-\frac{1}{2}} \\ B^{-\frac{1}{2}}W & I+B^{-\frac{1}{2}}WW^TB^{-\frac{1}{2}} \end{pmatrix}^{-1} \begin{pmatrix} I & WW^TB^{-\frac{1}{2}} \\ B^{-\frac{1}{2}}WW^T & -\frac{1}{\beta}B^{-1} \end{pmatrix} \\
&= \widetilde{\mathcal{P}}_2^{-1}\widetilde{\mathcal{A}} \sim \widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}},
\end{aligned}$$

where

$$\widetilde{\mathcal{P}}_2 = \begin{pmatrix} I & W^TB^{-\frac{1}{2}} \\ B^{-\frac{1}{2}}W & I+B^{-\frac{1}{2}}WW^TB^{-\frac{1}{2}} \end{pmatrix}, \qquad \widetilde{\mathcal{A}} = \begin{pmatrix} I & WW^TB^{-\frac{1}{2}} \\ B^{-\frac{1}{2}}WW^T & -\frac{1}{\beta}B^{-1} \end{pmatrix}. \tag{3.4}$$

Therefore, the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ is similar to the symmetric matrix $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$. So they have identical eigenvalues. This implies that we just need to estimate the bounds of the eigenvalues of $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$.

Noticing that for any $0 \neq x \in \mathbb{R}^n$ and $y = \widetilde{\mathcal{P}}_2^{-\frac{1}{2}}x \neq 0$, we can write

$$\frac{x^T\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}x}{x^Tx} = \frac{x^T\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}x}{x^T\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}x} \cdot \frac{x^T\widetilde{\mathcal{P}}_2^{-1}x}{x^Tx} = \frac{y^T\widetilde{\mathcal{A}}y}{y^Ty} \cdot \frac{x^T\widetilde{\mathcal{P}}_2^{-1}x}{x^Tx},$$

from which we clearly see that some careful and sharp spectral estimates for the matrices $\widetilde{\mathcal{P}}_2$ and $\widetilde{\mathcal{A}}$ should be achieved in order to establish desired spectral estimates of the matrix $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$. It is surprising to see that the eigenvalues of $\widetilde{\mathcal{P}}_2$ and $\widetilde{\mathcal{A}}$ can be explicitly expressed in terms of the eigenvalues of $M^{-1}K$, after some detailed analysis by taking full advantage of its special structure and the singular value decomposition. In the first place, we will study the eigenvalues of $\widetilde{\mathcal{P}}_2$.

Let $W = U\Sigma V^T$ be the singular value decomposition, where U and V are $n \times n$ orthonormal matrices, $\Sigma = \mathrm{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$ with $\sigma_j > 0, j = 1, 2, \cdots, n$ being the singular values of W. Then by (3.2), we have $B = \frac{\alpha^2}{\beta}I + \alpha^2 U\Sigma^4U^T = U(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4)U^T$, which leads to

$$B^{-\frac{1}{2}} = U\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-\frac{1}{2}}U^T.$$

This together with (3.4) yields that

$$\begin{aligned}
\widetilde{\mathcal{P}}_2 &= \begin{pmatrix} I & V\Sigma\left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-\frac{1}{2}}U^T \\ U\left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-\frac{1}{2}}\Sigma V^T & I+U\left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-1}\Sigma^2U^T \end{pmatrix} \\
&\sim \begin{pmatrix} I & \Sigma\left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-\frac{1}{2}} \\ \left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-\frac{1}{2}}\Sigma & I+\left(\frac{\alpha^2}{\beta}I+\alpha^2\Sigma^4\right)^{-1}\Sigma^2 \end{pmatrix}.
\end{aligned}$$

By the same way as that in Section 2, we can derive all the eigenvalues of $\widetilde{\mathcal{P}}_2$:

$$\frac{2+\frac{\beta\sigma_i^2}{\alpha^2(1+\beta\sigma_i^4)} \pm \sqrt{\left[2+\frac{\beta\sigma_i^2}{\alpha^2(1+\beta\sigma_i^4)}\right]^2 - 4}}{2}.$$

Noticing that $WW^T = L^{-1}HH^TL^{-T} = L^{-1}KL^{-T}$ is similar to $M^{-1}K$, we know that $\sigma_i^2\,(i = 1, 2, \ldots, n)$ are the eigenvalues of $M^{-1}K$. Then we can establish the following lemma immediately.

Lemma 3.1. *Let $\widetilde{\mathcal{P}}_2 \in \mathbb{R}^{2n\times 2n}$ be defined as in (3.4) with $\alpha > 0$, let $\mu_i\,(1 \le i \le n)$ be the eigenvalues of $M^{-1}K$. Then all the eigenvalues of $\widetilde{\mathcal{P}}_2$ are*

$$\frac{2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)} \pm \sqrt{\left[2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)}\right]^2 - 4}}{2},$$

which are contained in

$$\left[\frac{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2}, \frac{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2}\right].$$

Proof. Clearly, we just need to demonstrate the second result. Since $M^{-1}K$ is similar to $M^{-\frac{1}{2}}KM^{-\frac{1}{2}}$, which is symmetric positive definite, then for any $1 \le i \le n$, we have $\mu_i > 0$. This together with $\beta > 0$ leads to

$$\frac{1}{\beta\mu_i} + \mu_i \ge 2\sqrt{\frac{\mu_i}{\beta\mu_i}} = \frac{2}{\sqrt{\beta}},$$

which implies that

$$\frac{\beta\mu_i}{1 + \beta\mu_i^2} = \frac{1}{\frac{1}{\beta\mu_i} + \mu_i} \le \frac{\sqrt{\beta}}{2}. \tag{3.5}$$

Then for any eigenvalue λ of $\widetilde{\mathcal{P}}_2$, it holds that

$$\begin{aligned}\lambda &\le \max_{1\le i\le n} \frac{2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)} + \sqrt{\left[2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)}\right]^2 - 4}}{2}\\ &\le \frac{2 + \frac{\sqrt{\beta}}{2\alpha^2} + \sqrt{\left(2 + \frac{\sqrt{\beta}}{2\alpha^2}\right)^2 - 4}}{2}\\ &= \frac{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2}.\end{aligned}$$

On the other hand, as the function $x - \sqrt{x^2 - 4}$ is monotone decreasing on $(2, +\infty)$, by (3.5) we have

$$\begin{aligned}\lambda &\ge \min_{1\le i\le n} \frac{2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)} - \sqrt{\left[2 + \frac{\beta\mu_i}{\alpha^2(1+\beta\mu_i^2)}\right]^2 - 4}}{2}\\ &\ge \frac{2 + \frac{\sqrt{\beta}}{2\alpha^2} - \sqrt{\left(2 + \frac{\sqrt{\beta}}{2\alpha^2}\right)^2 - 4}}{2}\\ &= \frac{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2},\end{aligned}$$

which follows the result. □

In the following, we shall derive the eigenvalues of $\widetilde{\mathcal{A}}$. By using the singular value decomposition of W and (3.4), we are able to obtain

$$\begin{aligned}\widetilde{\mathcal{A}} &= \begin{pmatrix} I & U\Sigma^2\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-\frac{1}{2}}U^T \\ U\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-\frac{1}{2}}\Sigma^2U^T & -\frac{1}{\beta}U\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-1}U^T \end{pmatrix}\\ &\sim \begin{pmatrix} I & \Sigma^2\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-\frac{1}{2}} \\ \left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-\frac{1}{2}}\Sigma^2 & -\frac{1}{\beta}\left(\frac{\alpha^2}{\beta}I + \alpha^2\Sigma^4\right)^{-1} \end{pmatrix},\end{aligned}$$

which shows that all the eigenvalues of $\widetilde{\mathcal{A}}$ can be expressed as

$$\frac{\alpha^2 - \frac{1}{1+\beta\sigma_i^4} \pm \sqrt{\left(\alpha^2 - \frac{1}{1+\beta\sigma_i^4}\right)^2 + 4\alpha^2}}{2\alpha^2}.$$

Noticing that $\sigma_i^2\,(i = 1, 2, \cdots, n)$ are the eigenvalues of $M^{-1}K$, we come to the following conclusion from the above analysis.

Lemma 3.2. *Let $\widetilde{\mathcal{A}} \in \mathbb{R}^{2n\times 2n}$ be defined as in (3.4), let $\mu_i\,(1 \le i \le n)$ be the eigenvalues of $M^{-1}K$. Then all the eigenvalues of $\widetilde{\mathcal{A}}$ are*

$$\frac{\alpha^2 - \frac{1}{1+\beta\mu_i^2} \pm \sqrt{\left(\alpha^2 - \frac{1}{1+\beta\mu_i^2}\right)^2 + 4\alpha^2}}{2\alpha^2},$$

which are contained in

$$\left[-\frac{1}{\alpha^2}, \frac{\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2}}{2\alpha^2}\right] \cup \left[1, \frac{\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2}}{2\alpha^2}\right].$$

Proof. It is easy to verify that

$$0 < \frac{1}{1+\beta\mu_i^2} < 1.$$

So for any $\lambda > 0$ be the positive eigenvalue of $\widetilde{\mathcal{A}}$, we can see that

$$1 < \lambda = \frac{\alpha^2 - \frac{1}{1+\beta\mu_i^2} + \sqrt{\left(\alpha^2 - \frac{1}{1+\beta\mu_i^2}\right)^2 + 4\alpha^2}}{2\alpha^2} < \frac{\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2}}{2\alpha^2},$$

and for any $\lambda < 0$ be the negative eigenvalue of $\widetilde{\mathcal{A}}$, it holds that

$$-\frac{1}{\alpha^2} < \lambda = \frac{\alpha^2 - \frac{1}{1+\beta\mu_i^2} - \sqrt{\left(\alpha^2 - \frac{1}{1+\beta\mu_i^2}\right)^2 + 4\alpha^2}}{2\alpha^2} < \frac{\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2}}{2\alpha^2},$$

which completes the proof. □

Now we come to deduce the spectral estimates of the matrix $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$.

Lemma 3.3. *Let $\widetilde{\mathcal{A}} \in \mathbb{R}^{2n\times 2n}$ and $\widetilde{\mathcal{P}}_2 \in \mathbb{R}^{2n\times 2n}$ be defined as in (3.4) with $\alpha > 0$. Then all the eigenvalues of the matrix $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$ are contained in $\mathcal{I}^- \cup \mathcal{I}^+ \subset \mathbb{R}$, with*

$$\mathcal{I}^- = \left[-\frac{4}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}, \frac{2\left(\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2}\right)}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}\right] \subset \mathbb{R}^-,$$

and

$$\mathcal{I}^+ = \left[\frac{4\alpha^2}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}, \frac{2(\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2})}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}\right] \subset \mathbb{R}^+.$$

Proof. Assume that λ is an eigenvalue of $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$. Since $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$ is symmetric, combining with Lemmas 3.1 and 3.2, it yields that

$$\begin{aligned}
\lambda &\le \max_{x\neq 0} \frac{x^T\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}\widetilde{\mathcal{A}}\widetilde{\mathcal{P}}_2^{-\frac{1}{2}}x}{x^Tx} = \max_{x\neq 0}\frac{x^T\widetilde{\mathcal{A}}x}{x^T\widetilde{\mathcal{P}}_2x} = \max_{x\neq 0}\frac{x^T\widetilde{\mathcal{A}}x}{x^Tx}\frac{x^Tx}{x^T\widetilde{\mathcal{P}}_2x} \\
&\le \max_{x\neq 0}\frac{x^T\widetilde{\mathcal{A}}x}{x^Tx}\max_{x\neq 0}\frac{x^Tx}{x^T\widetilde{\mathcal{P}}_2x} = \max_{x\neq 0}\frac{x^T\widetilde{\mathcal{A}}x}{x^Tx}\frac{1}{\min_{x\neq 0}\frac{x^T\widetilde{\mathcal{P}}_2x}{x^Tx}} \\
&\le \frac{\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2}}{2\alpha^2}\frac{4\alpha^2}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}} \\
&= \frac{2(\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2})}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}},
\end{aligned}$$

and

$$\lambda \geq \min_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2^{-\frac{1}{2}} \widetilde{\mathcal{A}} \widetilde{\mathcal{P}}_2^{-\frac{1}{2}} x}{x^T x} \geq \min_{x\neq 0} \frac{x^T \widetilde{\mathcal{A}} x}{x^T x} \max_{x\neq 0} \frac{x^T x}{x^T \widetilde{\mathcal{P}}_2 x} = \min_{x\neq 0} \frac{x^T \widetilde{\mathcal{A}} x}{x^T x} \frac{1}{\min_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2 x}{x^T x}}$$
$$\geq -\frac{1}{\alpha^2} \frac{4\alpha^2}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}} = -\frac{4}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}.$$

The rest of the proof will derive the lower bound of the positive eigenvalue and the upper bound of the negative eigenvalue, respectively. Noticing that $\frac{1}{\lambda}$ is an eigenvalue of $\widetilde{\mathcal{P}}_2^{\frac{1}{2}} \widetilde{\mathcal{A}}^{-1} \widetilde{\mathcal{P}}_2^{\frac{1}{2}}$, by using Lemmas 3.1 and 3.2, we can get

$$\frac{1}{\lambda} \leq \max_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2^{\frac{1}{2}} \widetilde{\mathcal{A}}^{-1} \widetilde{\mathcal{P}}_2^{\frac{1}{2}} x}{x^T x} = \max_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2^{\frac{1}{2}} \widetilde{\mathcal{A}}^{-1} \widetilde{\mathcal{P}}_2^{\frac{1}{2}} x}{x^T \widetilde{\mathcal{P}}_2^{\frac{1}{2}} \widetilde{\mathcal{P}}_2^{\frac{1}{2}} x} \frac{x^T \widetilde{\mathcal{P}}_2 x}{x^T x}$$
$$\leq \max_{x\neq 0} \frac{x^T \widetilde{\mathcal{A}}^{-1} x}{x^T x} \max_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2 x}{x^T x} \leq \frac{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2},$$

and

$$\frac{1}{\lambda} \geq \min_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2^{\frac{1}{2}} \widetilde{\mathcal{A}}^{-1} \widetilde{\mathcal{P}}_2^{\frac{1}{2}} x}{x^T x} \geq \min_{x\neq 0} \frac{x^T \widetilde{\mathcal{A}}^{-1} x}{x^T x} \max_{x\neq 0} \frac{x^T \widetilde{\mathcal{P}}_2 x}{x^T x}$$
$$\geq \frac{2\alpha^2}{\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2}} \frac{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{4\alpha^2}$$
$$= \frac{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}{2(\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2})}.$$

Then for any $\lambda > 0$, it follows that

$$\lambda \geq \frac{4\alpha^2}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}},$$

and for any $\lambda < 0$, we have

$$\lambda \leq \frac{2(\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2})}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}.$$

This combined with the first two results completes the proof. □

Noticing that the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ is similar to $\widetilde{\mathcal{P}}_2^{-\frac{1}{2}} \widetilde{\mathcal{A}} \widetilde{\mathcal{P}}_2^{-\frac{1}{2}}$, by Lemma 3.3, we naturally have the following theorem on the bounds of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$.

Theorem 3.1. *Let $\mathcal{A} \in \mathbb{R}^{2n\times 2n}$ be the coefficient matrix of the linear system (1.4) and $\mathcal{P}_2$ the block diagonal preconditioner defined as in (3.1) with $\alpha > 0$. Then all the eigenvalues of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ are contained in $\mathcal{I}^- \cup \mathcal{I}^+ \subset \mathbb{R}$, with*

$$\mathcal{I}^- = \left[-\frac{4}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}, \frac{2(\alpha^2 - \sqrt{\alpha^4 + 4\alpha^2})}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}\right] \subset \mathbb{R}^-,$$

and

$$\mathcal{I}^+ = \left[\frac{4\alpha^2}{4\alpha^2 + \sqrt{\beta} + \sqrt{8\alpha^2\sqrt{\beta} + \beta}}, \frac{2(\alpha^2 + \sqrt{\alpha^4 + 4\alpha^2})}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}}\right] \subset \mathbb{R}^+.$$

Remark 3.1. As we have seen in Theorem 3.1, the spectral bounds of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ are independent of the eigenvalues of $M^{-1}K$. This shows that the eigenvalues of $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ would be clustered enough even if the original problem (1.4) is ill-conditioned or the condition number of $M^{-1}K$ is very large.

We should emphasize that the bounds of the eigenvalues of $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ will also not rely on the regularization parameter β by choosing the parameter α appropriately. Actually, it is easy to verify that

$$-\frac{4}{4\alpha^2 + \sqrt{\beta} - \sqrt{8\alpha^2\sqrt{\beta} + \beta}} = -\frac{4}{\alpha^2\left(4 + \frac{\sqrt{\beta}}{\alpha^2} - \sqrt{8\frac{\sqrt{\beta}}{\alpha^2} + \frac{\beta}{\alpha^4}}\right)}, \tag{3.6}$$

Table 1
Numerical results of 16 ×16 grid.

Method	CPU	Res	Iter	CPU	Res	Iter
	$\beta = 10^{-1}$			$\beta = 10^{-2}$		
MINRES	0.0096	9.6723e−10	115	0.0127	9.4306e−10	114
MINRES($\mathcal{P}_1$)	0.0040	4.1729e−10	11	0.0040	5.8868e−10	14
MINRES($\mathcal{P}_2$[5])	0.0051	1.6840e−10	12	0.0034	5.0783e−11	12
MINRES($\mathcal{P}_2$[50])	0.0050	9.8681e−10	7	0.0037	5.7299e−10	7
MINRES($\mathcal{P}_2$[500])	0.0023	1.3063e−11	5	0.0017	1.5123e−12	5
MINRES($\mathcal{P}_d$)	0.0623	1.3060e−10	12	0.0823	1.6372e−10	18
GMRES($\mathcal{P}_{bc}$)	0.0580	5.2681e−10	10	0.0676	1.2708e−10	16
GMRES($\mathcal{P}_{sc}$)	0.1103	1.7356e−10	7	0.1145	4.7338e−11	10
GMRES($\mathcal{P}_{ct}$)	0.5221	9.0324e−10	135	0.4206	5.6199e−10	128
	$\beta = 10^{-4}$			$\beta = 10^{-8}$		
MINRES	0.0206	8.9748e−10	185	0.0024	9.3668e−10	22
MINRES($\mathcal{P}_1$)	0.0042	6.0459e−10	23	0.0019	7.1266e−13	4
MINRES($\mathcal{P}_2$[5])	0.0550	4.8946e−10	10	0.0024	2.1479e−11	3
MINRES($\mathcal{P}_2$[50])	0.0024	5.3262e−10	5	0.0027	8.8086e−16	4
MINRES($\mathcal{P}_2$[500])	0.0014	5.3650e−14	5	0.0012	1.1280e−11	2
MINRES($\mathcal{P}_d$)	0.2124	3.8728e−10	54	1.1502	9.3016e−10	304
GMRES($\mathcal{P}_{bc}$)	0.0948	3.2031e−10	37	0.2950	7.9586e−10	122
GMRES($\mathcal{P}_{sc}$)	0.2451	2.0551e−10	25	0.9758	2.0134e−08	117
GMRES($\mathcal{P}_{ct}$)	0.3136	6.2454e−10	92	0.0457	1.2617e−11	4

$$\frac{2(\alpha^2-\sqrt{\alpha^4+4\alpha^2})}{4\alpha^2+\sqrt{\beta}+\sqrt{8\alpha^2\sqrt{\beta}+\beta}} = \frac{2\left(1-\sqrt{1+\frac{4}{\alpha^2}}\right)}{4+\frac{\sqrt{\beta}}{\alpha^2}+\sqrt{8\frac{\sqrt{\beta}}{\alpha^2}+\frac{\beta}{\alpha^4}}}, \tag{3.7}$$

$$\frac{4\alpha^2}{4\alpha^2+\sqrt{\beta}+\sqrt{8\alpha^2\sqrt{\beta}+\beta}} = \frac{4}{4+\frac{\sqrt{\beta}}{\alpha^2}+\sqrt{8\frac{\sqrt{\beta}}{\alpha^2}+\frac{\beta}{\alpha^4}}}, \tag{3.8}$$

$$\frac{2(\alpha^2+\sqrt{\alpha^4+4\alpha^2})}{4\alpha^2+\sqrt{\beta}-\sqrt{8\alpha^2\sqrt{\beta}+\beta}} = \frac{2\left(1+\sqrt{1+\frac{4}{\alpha^2}}\right)}{4+\frac{\sqrt{\beta}}{\alpha^2}-\sqrt{8\frac{\sqrt{\beta}}{\alpha^2}+\frac{\beta}{\alpha^4}}}. \tag{3.9}$$

Noticing that the function $x-\sqrt{8x+x^2}$ is monotone decreasing on $(0,+\infty)$ and $x+\sqrt{8x+x^2}$ is monotone increasing on $(0,+\infty)$, then for any $\alpha > 0$ such that $\alpha^2 \geq \sqrt{\beta}$, we can get

$$4+\frac{\sqrt{\beta}}{\alpha^2}-\sqrt{8\frac{\sqrt{\beta}}{\alpha^2}+\frac{\beta}{\alpha^4}} \geq 2$$

and

$$4+\frac{\sqrt{\beta}}{\alpha^2}+\sqrt{8\frac{\sqrt{\beta}}{\alpha^2}+\frac{\beta}{\alpha^4}} \leq 8.$$

This together with (3.6)–(3.9) and Theorem 3.1 yields that all the eigenvalues of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ are contained in

$$\left[-\frac{2}{\alpha^2},\ \frac{1}{4}-\sqrt{\frac{1}{16}+\frac{1}{4\alpha^2}}\right]\cup\left[\frac{1}{2},\ 1+\sqrt{1+\frac{4}{\alpha^2}}\right].$$

Then we could summarize this by the following corollary.

Corollary 3.1. *Under the same settings and conditions as in Theorem 3.1. If $\alpha > 0$ such that $\alpha^2 \geq \sqrt{\beta}$, then all the eigenvalues of the preconditioned system $\mathcal{P}_2^{-1}\mathcal{A}\mathcal{P}_2^{-T}$ are contained in*

$$\left[-\frac{2}{\alpha^2},\ \frac{1}{4}-\sqrt{\frac{1}{16}+\frac{1}{4\alpha^2}}\right]\cup\left[\frac{1}{2},\ 1+\sqrt{1+\frac{4}{\alpha^2}}\right].$$

Table 2
Numerical results of 32 ×32 grid.

Method	CPU	Res	Iter	CPU	Res	Iter
	$\beta = 10^{-1}$			$\beta = 10^{-2}$		
MINRES	0.0486	7.5068e−10	410	0.0663	8.8715e−10	423
MINRES($\mathcal{P}_1$)	0.0202	2.1698e−10	11	0.0242	6.9490e−10	14
MINRES($\mathcal{P}_2$[5])	0.0232	4.3596e−10	12	0.0216	1.3547e−10	12
MINRES($\mathcal{P}_2$[50])	0.0160	2.9207e−11	8	0.0170	3.5839e−12	8
MINRES($\mathcal{P}_2$[500])	0.0141	3.0459e−11	5	0.0128	3.3698e−12	5
MINRES($\mathcal{P}_d$)	1.1469	1.2434e−10	12	1.5987	2.5304e−10	18
GMRES($\mathcal{P}_{bc}$)	0.1310	9.3977e−10	10	0.2060	2.7192e−10	16
GMRES($\mathcal{P}_{sc}$)	1.3186	3.2832e−10	7	1.6544	8.0846e−11	10
GMRES($\mathcal{P}_{ct}$)	1007.2681	9.9993e−10	55 111	146.3364	9.9986e−10	8068
	$\beta = 10^{-4}$			$\beta = 10^{-8}$		
MINRES	0.1002	9.9876e−10	510	0.0311	7.8285e−10	19
MINRES($\mathcal{P}_1$)	0.0572	9.2181e−10	23	0.0412	5.6319e−10	7
MINRES($\mathcal{P}_2$[5])	0.0200	5.7783e−10	10	0.0472	4.2238e−10	4
MINRES($\mathcal{P}_2$[50])	0.0160	7.5955e−10	5	0.0172	1.1178e−12	4
MINRES($\mathcal{P}_2$[500])	0.0155	7.6532e−14	5	0.0105	6.9954e−11	2
MINRES($\mathcal{P}_d$)	4.5782	8.2375e−10	55	83.1706	9.6309e−10	1079
GMRES($\mathcal{P}_{bc}$)	0.4375	1.4455e−10	41	2.0239	9.4888e−10	189
GMRES($\mathcal{P}_{sc}$)	3.9477	2.1499e−10	27	48.4909	9.6707e−10	325
GMRES($\mathcal{P}_{ct}$)	8.5521	9.6267e−10	461	0.2124	2.1172e−10	7

Table 3
Numerical results of 64 ×64 grid.

Method	CPU	Res	Iter	CPU	Res	Iter
	$\beta = 10^{-1}$			$\beta = 10^{-2}$		
MINRES	0.6098	9.9322e−10	1555	0.6729	9.2607e−10	1611
MINRES($\mathcal{P}_1$)	0.3046	1.2973e−10	11	0.3378	4.7128e−10	14
MINRES($\mathcal{P}_2$[5])	0.3123	9.7939e−10	12	0.2854	2.7932e−10	12
MINRES($\mathcal{P}_2$[50])	0.2650	5.9511e−11	8	0.2373	5.7723e−12	8
MINRES($\mathcal{P}_2$[500])	0.1999	6.5977e−11	5	0.2472	7.3843e−12	5
MINRES($\mathcal{P}_d$)	28.6988	9.7905e−11	12	42.1353	1.9389e−10	18
GMRES($\mathcal{P}_{bc}$)	0.8111	2.3079e−10	11	1.0703	4.2025e−10	16
GMRES($\mathcal{P}_{sc}$)	39.0810	5.2815e−10	7	49.3087	9.4408e−11	10
GMRES($\mathcal{P}_{ct}$)	–	–	–	–	–	–
	$\beta = 10^{-4}$			$\beta = 10^{-8}$		
MINRES	0.3478	9.9433e−10	818	0.0082	6.3027e−10	17
MINRES($\mathcal{P}_1$)	0.4332	9.4079e−10	22	0.2990	9.0045e−10	13
MINRES($\mathcal{P}_2$[5])	0.2676	9.3546e−10	10	0.2235	7.2797e−11	7
MINRES($\mathcal{P}_2$[50])	0.2241	1.5232e−10	7	0.2213	1.1674e−12	5
MINRES($\mathcal{P}_2$[500])	0.2083	1.0518e−13	5	0.1784	1.8770e−10	2
MINRES($\mathcal{P}_d$)	129.4758	8.4639e−10	58	–	–	–
GMRES($\mathcal{P}_{bc}$)	2.6193	3.1173e−10	41	–	–	–
GMRES($\mathcal{P}_{sc}$)	128.8898	4.9213e−10	27	–	–	–
GMRES($\mathcal{P}_{ct}$)	390.2654	9.9355e−10	2763	2.8963	6.4406e−10	23

4. Numerical experiments

In this section, we present some numerical results to demonstrate the effectiveness of the preconditioners introduced in Sections 2 and 3. All experiments were run on a PC with Intel(R) Core(TM) i5-5257U CPU @2.70 GHz 8.00 GB, and all programmings are implemented in MATLAB R2013b.

In our experiments, we discretize the elliptic distributed control problems (1.1) by the Q1 finite elements on some uniform grids. We apply the IFISS software package developed by Elman et al. [20] to generate the discretized linear systems for the meshes of size 16 ×16, 32 ×32, 64 ×64 and 128 ×128. For each case, we test four regularization parameters of different scale, i.e., $\beta = 10^{-1}$, $\beta = 10^{-2}$, $\beta = 10^{-4}$ and $\beta = 10^{-8}$.

Then we shall solve the saddle point problem (1.3) or the reduced system (1.4), using respectively the MINRES method without any preconditioner, the preconditioned MINRES method with our block diagonal preconditioner $\mathcal{P}_1$ defined as in (2.5), the lower triangular preconditioner $\mathcal{P}_2$ defined as in (3.1), and the block diagonal preconditioner $\mathcal{P}_d$ in [8], the preconditioned GMRES method with the constraint preconditioner $\mathcal{P}_{sc}$ in [8], the block counter tridiagonal preconditioner

Table 4
Numerical results of 128 × 128 grid.

Method	CPU	Res	Iter	CPU	Res	Iter
	$\beta = 10^{-1}$			$\beta = 10^{-2}$		
MINRES	9.6784	9.9417e−10	6033	11.6491	9.7443e−10	6370
MINRES($\mathcal{P}_1$)	5.1303	9.1972e−10	10	5.9705	2.9203e−10	14
MINRES($\mathcal{P}_2$[5])	5.6292	2.1731e−10	14	5.2662	4.7346e−10	12
MINRES($\mathcal{P}_2$[50])	4.4201	1.2641e−10	8	4.5288	1.0433e−11	8
MINRES($\mathcal{P}_2$[500])	3.9266	1.3376e−10	5	4.0855	1.5613e−11	5
MINRES($\mathcal{P}_d$)	1462.6969	1.2069e−10	12	1976.7363	1.2260e−10	18
GMRES($\mathcal{P}_{bc}$)	5.6835	4.2162e−10	11	7.3415	5.8740e−10	16
GMRES($\mathcal{P}_{sc}$)	3187.7680	2.8703e−10	7	2056.3138	9.9483e−11	10
GMRES($\mathcal{P}_{ct}$)	–	–	–	–	–	–
	$\beta = 10^{-4}$			$\beta = 10^{-8}$		
MINRES	2.2229	9.9787e−10	1245	0.0357	5.0684e−10	15
MINRES($\mathcal{P}_1$)	7.3904	4.0333e−10	22	8.5983	9.2412e−10	16
MINRES($\mathcal{P}_2$[5])	5.3056	4.2935e−11	12	4.2561	5.6252e−10	7
MINRES($\mathcal{P}_2$[50])	4.2705	2.4018e−10	7	3.7857	4.9732e−12	5
MINRES($\mathcal{P}_2$[500])	3.8167	1.7950e−13	5	3.4244	7.5713e−10	2
MINRES($\mathcal{P}_d$)	–	–	–	–	–	–
GMRES($\mathcal{P}_{bc}$)	18.6946	4.3947e−10	41	–	–	–
GMRES($\mathcal{P}_{sc}$)	5563.3185	3.3432e−10	28	–	–	–
GMRES($\mathcal{P}_{ct}$)	-	-	-	93.7074	8.4793e−10	97

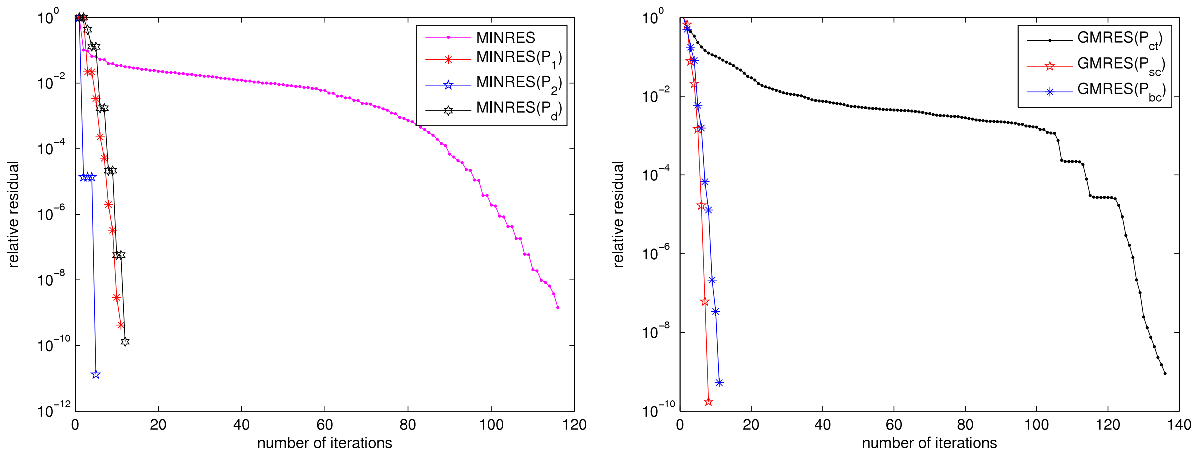

Fig. 1. The iterations curves of MINRES and GMRES with $\alpha = 500$, $\beta = 10^{-1}$ and 16 × 16 grid.

$\mathcal{P}_{ct}$ in [9] and the block counter preconditioner $\mathcal{P}_{bc}$ in [25], where

$$\mathcal{P}_d = \begin{pmatrix} \beta M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & KM^{-1}K \end{pmatrix}, \qquad \mathcal{P}_{sc} = \begin{pmatrix} 0 & 0 & -M \\ 0 & \beta KM^{-1}K & K \\ -M & K & 0 \end{pmatrix},$$

$$\mathcal{P}_{ct} = \begin{pmatrix} 0 & 0 & -M \\ 0 & M & K \\ -M & K & 0 \end{pmatrix}, \qquad \mathcal{P}_{bc} = \begin{pmatrix} 0 & K \\ K & 0 \end{pmatrix}.$$

We compare the performance of these methods by reporting the number of required iterations (denoted by "Iter"), the required CPU time (denoted by "CPU"), and the relative residuals (denoted by "Res"). Let z^k be the kth approximate solution of the saddle point problem (1.4), then the "Res" is defined by the relative residuals

$$\text{Res} := \frac{\|\ell - \mathcal{A}z^k\|_2}{\|\ell\|_2}.$$

In our implementation, we stop all considered algorithms when Res $\leq 10^{-9}$ and test the lower triangular preconditioner $\mathcal{P}_2$ in (3.1) with $\alpha = 5,\ 50,\ 500$, which we denote by MINRES($\mathcal{P}_2$[5]), MINRES($\mathcal{P}_2$[50]), MINRES($\mathcal{P}_2$[500]), respectively. All the initial guesses z^0 are taken by zero vectors. The resulting numerical results are listed in Tables 1–4 and Figs. 1–4.

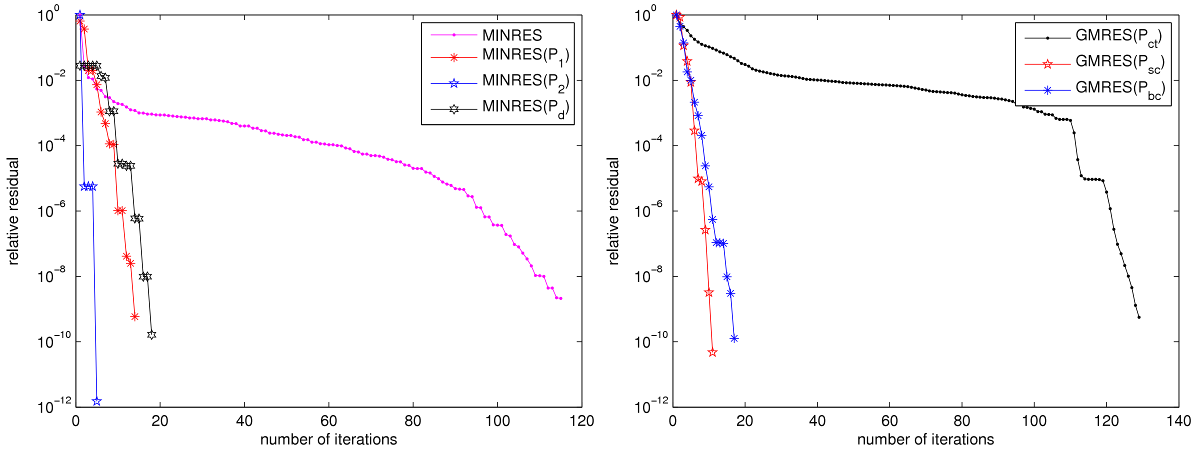

Fig. 2. The iterations curves of MINRES and GMRES with $\alpha = 500$, $\beta = 10^{-2}$ and 16×16.

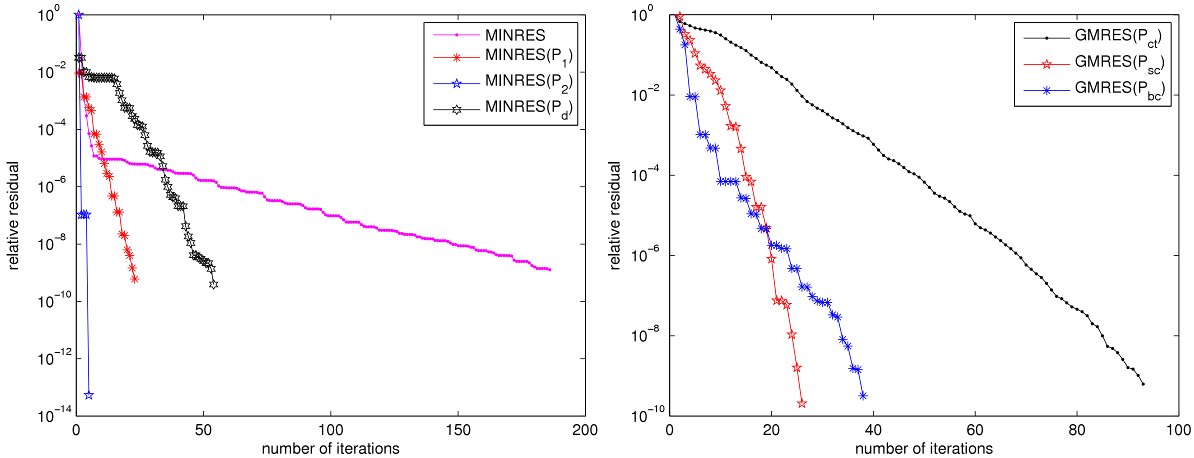

Fig. 3. The iterations curves of MINRES and GMRES with $\alpha = 500$, $\beta = 10^{-4}$ and $h = 1/16$ grid.

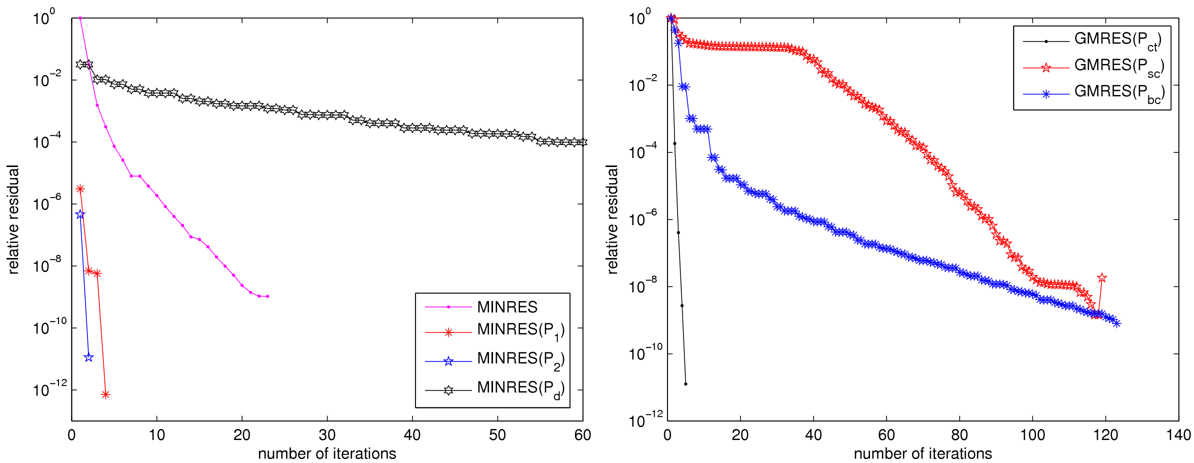

Fig. 4. The iterations curves of MINRES and GMRES with $\alpha = 500$, $\beta = 10^{-8}$ and 16×16 grid.

As can be seen in Tables 1–4 and Figs. 1–4 that our new preconditioners $\mathcal{P}_1$ and $\mathcal{P}_2$ are efficient and stabilized even for the regularization parameter β sufficiently small. The performance of the preconditioner $\mathcal{P}_2$ is not sensitive to the parameter α. In addition, it is surprising that the iterative numbers of the MINRES($\mathcal{P}_2$[500]) do not change along with the density of the grid.

References

[1] J.L. Lions, Optimal Control of Systems, Springer, New York, 1968.
[2] G. Biros, O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization, I. The Krylov-Schur solver, SIAM J. Sci. Comput. 27 (2005) 687–713.
[3] G. Biros, O. Ghattas, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization, II. The Lagrange-Newton solver and its application to optimal control of steady viscous flows, SIAM J. Sci. Comput. 27 (2005) 714–739.
[4] M. Kärcher, M.A. Grepl, A certified reduced basis method for parametrized elliptic optimal control problems, ESAIM Control Optim. Calc. Var. 20 (2014) 416–441.
[5] F. Negri, G. Rozza, A. Manzoni, A. Quarteroni, Reduced basis method for parametrized elliptic optimal control problems, SIAM J. Sci. Comput. 35 (2013) 2316–2340.
[6] J. Schöberl, W. Zulehner, Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems, SIAM J. Matrix Anal. Appl. 29 (2007) 752–773.
[7] R. Herzog, E. Sachs, Preconditioned conjugate gradient method for optimal control problems with control and state constraints, SIAM J. Matrix Anal. Appl. 31 (2010) 2291–2317.
[8] T. Rees, H.S. Dollar, A.J. Wathen, Optimal solvers for PDE-constrained optimization, SIAM J. Sci. Comput. 32 (2010) 271–298.
[9] Z.-Z. Bai, Block preconditioners for elliptic PDE-constrained optimization problems, Computing 91 (2011) 379–395.
[10] M. Bergounioux, K. Ito, K. Kunisch, Primal–dual strategy for constrained optimal control problems, SIAM J. Control Optim. 37 (1999) 1176–1194.
[11] M. Bergounioux, K. Kunisch, Primal–dual strategy for state-constrained optimal control problems, Comput. Optim. Appl. 22 (2002) 193–224.
[12] E. Casas, Control of an elliptic problem with pointwise state constraints, SIAM J. Control Optim. 31 (1993) 1297–1327.
[13] H.C. Elman, D.J. Silvester, A.J. Wathen, Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics, in: Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2014.
[14] O. Axelsson, M.G. Neytcheva, Eigenvalue estimates for preconditioned saddle point matrices, Numer. Linear Algebra Appl. 13 (2006) 339–360.
[15] D.L. Brown, R. Cortez, M.L. Minion, Accurate projection methods for the incompressible Navier-Stokes equations, J. Comput. Phys. 168 (2001) 464–499.
[16] M. Benzi, X.-P. Guo, A dimensional split preconditioner for Stokes and linearized Navier-Stokes equations, Appl. Numer. Math. 61 (2011) 66–76.
[17] Z.-Z. Bai, G.H. Golub, L.-Z. Lu, J.-F. Yin, Block triangular and skew-Hermitian splitting methods for positive-definite linear systems, SIAM J. Sci. Comput. 26 (2005) 844–863.
[18] M. Benzi, V. Simoncini, On the eigenvalues of a class of saddle point matrices, Numer. Math. 103 (2006) 173–196.
[19] H.C. Elamn, Preconditioning for the steady-state Navier-Stokes equations with low viscosity, SIAM J. Sci. Comput. 20 (1999) 1299–1316.
[20] H.C. Elman, A. Ramage, D.J. Silvester, Algorithm 866: IFISS, a MATLAB toolbox for modelling incompressible flow, ACM Trans. Math. Softw. 33 (2007) 1–18.
[21] H. Elman, D. Silvester, Fast nonsymmetric iterations and preconditioning for Navier-Stokes equations, SIAM J. Sci. Comput. 17 (1996) 33–46.
[22] H.C. Elman, D.J. Silvester, A.J. Wathen, Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations, Numer. Math. 90 (2002) 665–688.
[23] N. Huang, C.-F. Ma, Y.-J. Xie, An inexact relaxed DPSS preconditioner for saddle point problem, Appl. Math. Comput. 265 (2015) 431–447.
[24] J.W. Pearson, A.J. Wathen, A new approximation of the Schur complement in preconditioners for PDE-constrained optimization, Numer. Linear Algebra Appl. 19 (2012) 816–829.
[25] Y.P. Zeng, S.Q. Wang, H.R. Xu, S.L. Xie, Preconditioners for reduced saddle point systems arising in elliptic PDE-constrained optimization problems, J. Inequal. Appl. 355 (2015) 1–14.
[26] Z.Z. Bai, Sharp error bounds of some Krylov subspace methods for non-Hermitian linear systems, Appl. Math. Comput. 109 (2000) 273–285.
[27] C.C. Paige, M.A. Saunders, Solution of sparse indefinite systems of linear equations, SIAM J. Numer. Anal. 12 (1975) 617–629.
[28] Y. Saad, Iterative Methods for Sparse Linear Systems, Society for Industrial and Applied Mathematics, Philadelphia, 2003.
[29] Y. Saad, M.H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 7 (1986) 856–869.
[30] D.M. Young, Iterative Solution for Large Linear Systems, Academic Press, New York, 1971.