Numer Algor (2013) 62:655–675
DOI 10.1007/s11075-013-9696-9

ORIGINAL PAPER

Additive block diagonal preconditioning for block two-by-two linear systems of skew-Hamiltonian coefficient matrices

Zhong-Zhi Bai · Fang Chen · Zeng-Qi Wang

Received: 13 December 2011 / Accepted: 23 January 2013 /
Published online: 22 February 2013

Abstract For a class of block two-by-two systems of linear equations with certain skew-Hamiltonian coefficient matrices, we construct additive block diagonal preconditioning matrices and discuss the eigen-properties of the corresponding preconditioned matrices. The additive block diagonal preconditioners can be employed to accelerate the convergence rates of Krylov subspace iteration methods such as MINRES and GMRES. Numerical experiments show that MINRES preconditioned by the exact and the inexact additive block diagonal preconditioners are effective, robust and scalable solvers for the block two-by-two linear systems arising from the Galerkin finite-element discretizations of a class of distributed control problems.

Supported by The National Natural Science Foundation for Creative Research Groups (No. 11021101), The Hundred Talent Project of Chinese Academy of Sciences, The National Basic Research Program (No. 2011CB309703), and The National Natural Science Foundation (No. 91118001 and No. 11001175), P.R. China.

Z.-Z. Bai
School of Mathematics and Computer Science, Guizhou Normal University,
Guiyang 550001, People's Republic of China

Z.-Z. Bai (✉)
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, P.O. Box 2719, Beijing 100190, People's Republic of China
e-mail: bzz@lsec.cc.ac.cn

F. Chen
Key Laboratory of Environmental Mechanics, Institute of Mechanics, Chinese Academy of Sciences,
Beijing 100190, People's Republic of China
e-mail: chenfreesky@gmail.com

Z.-Q. Wang
Department of Mathematics, Shanghai Jiaotong University,
Shanghai 200240, People's Republic of China
e-mail: wangzengqi@sjtu.edu.cn

Keywords Block two-by-two matrices · PMHSS iteration · Block diagonal matrix · Preconditioning · Spectral properties · PDE-constrained optimization

Mathematics Subject Classifications (2010) 65F10 · 65F50 · CR · G1.3

1 Introduction

Consider block two-by-two systems of linear equations of the form

$$\mathbf{A}\mathbf{x} \equiv \begin{pmatrix} \mathbf{W} & -\mathbf{T} \\ \mathbf{T} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} \equiv \mathbf{g}, \tag{1.1}$$

where $\mathbf{W}, \mathbf{T} \in \mathbb{R}^{n\times n}$ are real, symmetric and positive semidefinite matrices satisfying $\mathrm{null}(\mathbf{W}) \cap \mathrm{null}(\mathbf{T}) = \{0\}$, where $\mathrm{null}(\cdot)$ represents the null space of the corresponding matrix. Note that the matrix $\mathbf{A} \in \mathbb{R}^{2n\times 2n}$ is nonsingular. Hence the linear system (1.1) has a unique solution. This class of linear systems can be formally regarded as a special case of the generalized saddle-point problem [4, 5, 16] and the skew-Hamiltonian linear system [25, 32]. It frequently arises from finite-element discretization and first-order linearization of the two-phase flow problems based on Cahn-Hilliard equation [3, 17, 18], order-reduction and sinc discretization of the third-order linear ordinary differential equations [30], finite-element discretizations of elliptic PDE-constrained optimization problems such as the distributed control problems [6, 23, 24, 29], real equivalent formulations of complex symmetric linear systems [2, 15, 20], linear quadratic control problems [22, 25], and H_∞ control problems [28, 32, 34].

The matrix $\mathbf{A}$ naturally possesses the *Hermitian and skew-Hermitian splitting* (**HSS**)[1]

$$\mathbf{A} = \mathbf{H} + \mathbf{S},$$

with

$$\mathbf{H} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) = \begin{pmatrix} \mathbf{W} & 0 \\ 0 & \mathbf{W} \end{pmatrix} \quad \text{and} \quad \mathbf{S} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) = \begin{pmatrix} 0 & -\mathbf{T} \\ \mathbf{T} & 0 \end{pmatrix}$$

being the symmetric and the skew-symmetric parts and $\mathbf{A}^T$ being the transpose of the matrix $\mathbf{A}$; see [1, 14]. By modifying and preconditioning the HSS iteration method [10, 11, 13], recently Bai, Benzi, Chen and Wang [9] proposed and discussed a class of *preconditioned modified HSS* (**PMHSS**) iteration methods for solving the block two-by-two linear system (1.1); see also [7, 8].

The PMHSS iteration method introduced in [9] naturally leads to a preconditioning matrix

$$\mathbf{F}(\alpha) = (\alpha+1)\mathbf{P}(\alpha) \begin{pmatrix} \alpha\mathbf{W} + \mathbf{T} & 0 \\ 0 & \alpha\mathbf{W} + \mathbf{T} \end{pmatrix}, \text{ with } \quad \mathbf{P}(\alpha) = \frac{1}{2\alpha} \begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{pmatrix}, \tag{1.2}$$

for the block two-by-two matrix $\mathbf{A}$, where α is a given positive parameter and $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix. Theoretical analysis has shown that the eigenvalues of

[1] In the real case this becomes the symmetric and skew-symmetric splitting.

the preconditioned matrix $\mathbf{F}(\alpha)^{-1}\mathbf{A}$ are clustered within a complex disk centered at 1 with radius $\frac{\sqrt{\alpha^2+1}}{\alpha+1}$, and the matrix of the corresponding eigenvectors is unitary.

Because $\mathbf{F}(\alpha)$ is essentially a scaled product of the orthogonal and the symmetric positive definite matrices

$$\frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix},$$

in practical applications its action can be computed by solving only two systems of linear equations with the same coefficient matrix $\alpha\mathbf{W}+\mathbf{T}$, which is symmetric and positive definite. However, noticing that $\mathbf{F}(\alpha)$ is, in general, a nonsymmetric matrix, it can be only employed to precondition Krylov subspace iteration methods such as GMRES, which may require considerable memory and storage in computations.

In this paper, by further simplifying and modifying the PMHSS preconditioning matrix $\mathbf{F}(\alpha)$ in (1.2) we construct an *additive block diagonal* (**ABD**) preconditioning matrix

$$\mathbf{B}(\alpha) = \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix} \in \mathbb{R}^{2n\times 2n} \tag{1.3}$$

for the block two-by-two matrix $\mathbf{A}$. Note that the matrix $\mathbf{B}(\alpha)$ is obtained by directly dropping $\mathbf{P}(\alpha)$ in the matrix $\mathbf{F}(\alpha)$. Since the matrix $\mathbf{B}(\alpha)$ is symmetric positive definite, it can be used to precondition the symmetric form, denoted by $\mathbf{A}_s$, of the matrix $\mathbf{A}$ as well, where

$$\mathbf{A}_s = \begin{pmatrix} \mathbf{W} & \mathbf{T} \\ \mathbf{T} & -\mathbf{W} \end{pmatrix} \in \mathbb{R}^{2n\times 2n}. \tag{1.4}$$

In this case, instead of GMRES we can use a symmetric Krylov subspace iteration solver of orthogonal and minimal properties such as MINRES or SYMMLQ (see [26]), with considerable savings in terms of both computation and storage. Hence, the ABD preconditioner $\mathbf{B}(\alpha)$ can be applied to a larger spectrum of Krylov subspace iteration methods than the PMHSS preconditioner $\mathbf{F}(\alpha)$, though the computational cost of applying either of them is about the same.

We derive the expressions for the eigenvalues and eigenvectors of the preconditioned matrices $\mathbf{B}(\alpha)^{-1}\mathbf{A}$ and $\mathbf{B}(\alpha)^{-1}\mathbf{A}_s$, showing that the eigenvalues of $\mathbf{B}(\alpha)^{-1}\mathbf{A}$ are clustered within a complex disk centered at $(1, 0)$ with radius $\frac{\sqrt{\alpha^2+1}}{\alpha+1}$, those of $\mathbf{B}(\alpha)^{-1}\mathbf{A}_s$ are clustered within a real interval, and the matrices of the corresponding eigenvectors of both matrices are orthogonal or unitary. These immediately lead to theoretical estimates about the convergence rates of the ABD-preconditioned GMRES and the ABD-preconditioned MINRES methods for solving the block two-by-two linear systems with the coefficient matrices $\mathbf{A}$ in (1.1) and $\mathbf{A}_s$ in (1.4), respectively; see [31].

The exact and the inexact additive block diagonal preconditioners are applied to precondition a class of KKT (Karush-Kuhn-Tucker) linear systems arising from a finite-element discretization of a class of distributed control problems [6, 23, 24, 29]. Numerical results show that the exact and the inexact ABD-preconditioned MINRES or GMRES methods [19] lead to rapid convergence and outperform the exact and the inexact PMHSS-preconditioned GMRES methods, respectively; see [9].

The remainder of the paper is organized as follows. In Section 2, we analyze the preconditioning properties of the additive block diagonal preconditioning matrix with respect to both $\mathbf{A}$ and its symmetric form $\mathbf{A}_s$. In Section 3, we derive the additive block diagonal preconditioner for a class of block two-by-two linear systems arising from the distributed control problems. Numerical results are given in Section 4 to show the effectiveness of the additive block diagonal preconditioner. Finally, in Section 5 we end the paper with some conclusions and remarks.

2 Analyses of preconditioning properties

In this section we analyze the spectral properties of the preconditioning matrix $\mathbf{B}(\alpha)$ for both symmetric and nonsymmetric matrices $\mathbf{A}_s$ and $\mathbf{A}$, respectively. These results are precisely stated in the following theorem.

Theorem 2.1 *Let $\boldsymbol{A}_s, \boldsymbol{A} \in \mathbb{R}^{2n\times 2n}$ be the block two-by-two matrices defined in* (1.4) and (1.1)*, respectively, with $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ being symmetric positive semidefinite matrices satisfying* $\text{null}(\boldsymbol{W}) \cap \text{null}(\boldsymbol{T}) = \{0\}$*, and let α be a positive constant. Define $\boldsymbol{Z}^{(\alpha)} = (\alpha\boldsymbol{W} + \boldsymbol{T})^{-\frac{1}{2}}(\boldsymbol{W} - \alpha\boldsymbol{T})(\alpha\boldsymbol{W} + \boldsymbol{T})^{-\frac{1}{2}}$. Denote by $\mu_1^{(\alpha)}, \mu_2^{(\alpha)}, \ldots, \mu_n^{(\alpha)}$ the eigenvalues of the symmetric matrix $\boldsymbol{Z}^{(\alpha)} \in \mathbb{R}^{n\times n}$, and by $\boldsymbol{q}_1^{(\alpha)}, \boldsymbol{q}_2^{(\alpha)}, \ldots, \boldsymbol{q}_n^{(\alpha)}$ the corresponding $(\alpha\boldsymbol{W} + \boldsymbol{T})^{-1}$-orthonormal eigenvectors. Set $\boldsymbol{x}_j^{(\alpha)} = (\alpha\boldsymbol{W}+\boldsymbol{T})^{-\frac{1}{2}}\boldsymbol{q}_j^{(\alpha)}$ $(j = 1, 2, \ldots, n)$ and $\boldsymbol{X}^{(\alpha)} = (\boldsymbol{x}_1^{(\alpha)}, \boldsymbol{x}_2^{(\alpha)}, \ldots, \boldsymbol{x}_n^{(\alpha)}) \in \mathbb{R}^{n\times n}$. Then*

(i) the eigenvalues of the matrix $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}_s$ are given by

$$\lambda_{\pm}^{(\alpha,j)} = \pm\sqrt{\frac{\left(\mu_j^{(\alpha)}\right)^2 + 1}{\alpha^2 + 1}}, \qquad j = 1, 2, \ldots, n,$$

and the corresponding orthonormal eigenvectors are

$$\widetilde{\boldsymbol{x}}_j^{(\alpha)} = \begin{pmatrix} -\boldsymbol{x}_j^{(\alpha)}\phi_j^{(\alpha)} \\ \boldsymbol{x}_j^{(\alpha)}\psi_j^{(\alpha)} \end{pmatrix} \quad \text{and} \quad \widetilde{\boldsymbol{x}}_{n+j}^{(\alpha)} = \begin{pmatrix} \boldsymbol{x}_j^{(\alpha)}\psi_j^{(\alpha)} \\ \boldsymbol{x}_j^{(\alpha)}\phi_j^{(\alpha)} \end{pmatrix}, \qquad j = 1, 2, \ldots, n,$$

where

$$\phi_j^{(\alpha)} = \frac{\alpha + \mu_j^{(\alpha)} + \delta_j^{(\alpha)}}{\sqrt{2\delta_j^{(\alpha)}\left(\alpha + \mu_j^{(\alpha)} + \delta_j^{(\alpha)}\right)}} \quad \text{and} \quad \psi_j^{(\alpha)} = \frac{\alpha\mu_j^{(\alpha)} - 1}{\sqrt{2\delta_j^{(\alpha)}\left(\alpha + \mu_j^{(\alpha)} + \delta_j^{(\alpha)}\right)}},$$

with $\delta_j^{(\alpha)} = \sqrt{(\alpha^2 + 1)\left(\left(\mu_j^{(\alpha)}\right)^2 + 1\right)}$;

(ii) the eigenvalues of the matrix $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}$ are given by

$$\lambda_{\pm}^{(\alpha,j)} = \frac{(\alpha \pm \mathrm{i})\left(1 \mp \mathrm{i}\mu_j^{(\alpha)}\right)}{\alpha^2 + 1}, \qquad j = 1, 2, \ldots, n,$$

and the corresponding unitary eigenvectors are

$$\widetilde{\boldsymbol{x}}_j^{(\alpha)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{x}_j^{(\alpha)} \\ \mathrm{i}\,\boldsymbol{x}_j^{(\alpha)} \end{pmatrix}, \quad \widetilde{\boldsymbol{x}}_{n+j}^{(\alpha)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{x}_j^{(\alpha)} \\ -\mathrm{i}\,\boldsymbol{x}_j^{(\alpha)} \end{pmatrix}, \qquad j = 1, 2, \ldots, n,$$

where $\mathrm{i} = \sqrt{-1}$ *denotes the imaginary unit.*

Therefore, we have $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}_s = \widetilde{\boldsymbol{X}}^{(\alpha)} \Lambda^{(\alpha)} \widetilde{\boldsymbol{X}}^{(\alpha)^{-1}}$ *with*

$$\widetilde{\boldsymbol{X}}^{(\alpha)} = \begin{pmatrix} -\boldsymbol{X}^{(\alpha)}\Phi^{(\alpha)} & \boldsymbol{X}^{(\alpha)}\Psi^{(\alpha)} \\ \boldsymbol{X}^{(\alpha)}\Psi^{(\alpha)} & \boldsymbol{X}^{(\alpha)}\Phi^{(\alpha)} \end{pmatrix}$$

being an orthogonal matrix so that $\kappa_2\left(\widetilde{\boldsymbol{X}}^{(\alpha)}\right) = 1$, *and* $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A} = \widetilde{\boldsymbol{X}}^{(\alpha)} \Lambda^{(\alpha)} \widetilde{\boldsymbol{X}}^{(\alpha)^{-1}}$ *with*

$$\widetilde{\boldsymbol{X}}^{(\alpha)} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{X}^{(\alpha)} & \boldsymbol{X}^{(\alpha)} \\ \mathrm{i}\,\boldsymbol{X}^{(\alpha)} & -\mathrm{i}\,\boldsymbol{X}^{(\alpha)} \end{pmatrix}$$

being a unitary matrix so that $\kappa_2\left(\widetilde{\boldsymbol{X}}^{(\alpha)}\right) = 1$, *where* $\kappa_2(\cdot)$ *represents the condition number in the Euclidean norm,*

$$\Lambda^{(\alpha)} = \begin{pmatrix} \Lambda_-^{(\alpha)} & 0 \\ 0 & \Lambda_+^{(\alpha)} \end{pmatrix}, \qquad \Lambda_\pm^{(\alpha)} = \mathrm{diag}\left(\lambda_\pm^{(\alpha,1)}, \lambda_\pm^{(\alpha,2)}, \ldots, \lambda_\pm^{(\alpha,n)}\right),$$

and

$$\Phi^{(\alpha)} = \mathrm{diag}\left(\phi_1^{(\alpha)}, \phi_2^{(\alpha)}, \ldots, \phi_n^{(\alpha)}\right), \quad \Psi^{(\alpha)} = \mathrm{diag}\left(\psi_1^{(\alpha)}, \psi_2^{(\alpha)}, \ldots, \psi_n^{(\alpha)}\right).$$

Proof Define matrices

$$\mathbf{Q}^{(\alpha)} = \left(\mathbf{q}_1^{(\alpha)}, \mathbf{q}_2^{(\alpha)}, \ldots, \mathbf{q}_n^{(\alpha)}\right) \in \mathbb{R}^{n\times n}$$

and

$$\Xi^{(\alpha)} = \mathrm{diag}\left(\mu_1^{(\alpha)}, \mu_2^{(\alpha)}, \ldots, \mu_n^{(\alpha)}\right) \in \mathbb{R}^{n\times n}.$$

Then it holds that

$$\mathbf{Z}^{(\alpha)}\mathbf{Q}^{(\alpha)} = \mathbf{Q}^{(\alpha)}\Xi^{(\alpha)} \quad \text{and} \quad \mathbf{X}^{(\alpha)} = (\alpha\mathbf{W} + \mathbf{T})^{-\frac{1}{2}}\mathbf{Q}^{(\alpha)}.$$

Let

$$\widehat{\mathbf{X}}^{(\alpha)} = \begin{pmatrix} \mathbf{X}^{(\alpha)} & 0 \\ 0 & \mathbf{X}^{(\alpha)} \end{pmatrix}, \quad \widehat{\Psi}^{(\alpha)} = \begin{pmatrix} \mathbf{I} & \Xi^{(\alpha)} \\ -\Xi^{(\alpha)} & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{J}_s^{(\alpha)} = \begin{pmatrix} \alpha\mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\alpha\mathbf{I} \end{pmatrix}, \quad \mathbf{J}^{(\alpha)} = \begin{pmatrix} \alpha\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \alpha\mathbf{I} \end{pmatrix}.$$

Then we have

$$\left(\mathbf{J}_s^{(\alpha)}\right)^{-1} = \frac{1}{\alpha^2+1}\mathbf{J}_s^{(\alpha)} \quad \text{and} \quad \left(\mathbf{J}^{(\alpha)}\right)^{-1} = \frac{1}{\alpha^2+1}\mathbf{J}^{(\alpha)}.$$

Based on the identities

$$\mathbf{A}_s\mathbf{J}_s^{(\alpha)} = \begin{pmatrix} \mathbf{W} & \mathbf{T} \\ \mathbf{T} & -\mathbf{W} \end{pmatrix} \begin{pmatrix} \alpha\mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\alpha\mathbf{I} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & \mathbf{W}-\alpha\mathbf{T} \\ \alpha\mathbf{T}-\mathbf{W} & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}$$

and

$$\mathbf{A}\mathbf{J}^{(\alpha)} = \begin{pmatrix} \mathbf{W} & -\mathbf{T} \\ \mathbf{T} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \alpha\mathbf{I} & \mathbf{I} \\ -\mathbf{I} & \alpha\mathbf{I} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & \mathbf{W}-\alpha\mathbf{T} \\ \alpha\mathbf{T}-\mathbf{W} & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix},$$

we obtain

$$\begin{aligned}
\mathbf{B}(\alpha)^{-1}\mathbf{A}_s &= \frac{1}{\alpha^2+1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{-1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & \mathbf{W}-\alpha\mathbf{T} \\ \alpha\mathbf{T}-\mathbf{W} & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix} \mathbf{J}_s^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{-\frac{1}{2}} \begin{pmatrix} \mathbf{I} & \mathbf{Z}^{(\alpha)} \\ -\mathbf{Z}^{(\alpha)} & \mathbf{I} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{\frac{1}{2}} \mathbf{J}_s^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \begin{pmatrix} \mathbf{X}^{(\alpha)} & 0 \\ 0 & \mathbf{X}^{(\alpha)} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \Xi^{(\alpha)} \\ -\Xi^{(\alpha)} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(\alpha)} & 0 \\ 0 & \mathbf{X}^{(\alpha)} \end{pmatrix}^{-1} \mathbf{J}_s^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \widehat{\mathbf{X}}^{(\alpha)} \widehat{\Psi}^{(\alpha)} \widehat{\mathbf{X}}^{(\alpha)^{-1}} \mathbf{J}_s^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \widehat{\mathbf{X}}^{(\alpha)} \widehat{\Psi}^{(\alpha)} \mathbf{J}_s^{(\alpha)} \widehat{\mathbf{X}}^{(\alpha)^{-1}}
\end{aligned} \tag{2.1}$$

and

$$\begin{aligned}
\mathbf{B}(\alpha)^{-1}\mathbf{A} &= \frac{1}{\alpha^2+1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{-1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & \mathbf{W}-\alpha\mathbf{T} \\ \alpha\mathbf{T}-\mathbf{W} & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix} \mathbf{J}^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{-\frac{1}{2}} \begin{pmatrix} \mathbf{I} & \mathbf{Z}^{(\alpha)} \\ -\mathbf{Z}^{(\alpha)} & \mathbf{I} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \alpha\mathbf{W}+\mathbf{T} & 0 \\ 0 & \alpha\mathbf{W}+\mathbf{T} \end{pmatrix}^{\frac{1}{2}} \mathbf{J}^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \begin{pmatrix} \mathbf{X}^{(\alpha)} & 0 \\ 0 & \mathbf{X}^{(\alpha)} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \Xi^{(\alpha)} \\ -\Xi^{(\alpha)} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{X}^{(\alpha)} & 0 \\ 0 & \mathbf{X}^{(\alpha)} \end{pmatrix}^{-1} \mathbf{J}^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \widehat{\mathbf{X}}^{(\alpha)} \widehat{\Psi}^{(\alpha)} \widehat{\mathbf{X}}^{(\alpha)^{-1}} \mathbf{J}^{(\alpha)} \\
&= \frac{1}{\alpha^2+1} \widehat{\mathbf{X}}^{(\alpha)} \widehat{\Psi}^{(\alpha)} \mathbf{J}^{(\alpha)} \widehat{\mathbf{X}}^{(\alpha)^{-1}}.
\end{aligned} \tag{2.2}$$

Now, it follows from straightforward computations that

$$\widehat{\Psi}_s^{(\alpha)} := \widehat{\Psi}^{(\alpha)} \mathbf{J}_s^{(\alpha)} = \begin{pmatrix} \alpha\mathbf{I} + \Xi^{(\alpha)} & \mathbf{I} - \alpha\,\Xi^{(\alpha)} \\ \mathbf{I} - \alpha\,\Xi^{(\alpha)} & -\left(\alpha\mathbf{I} + \Xi^{(\alpha)}\right) \end{pmatrix}.$$

Note that this matrix is real symmetric and has the eigenvalues

$$\pm\sqrt{(\alpha^2+1)\left(\left(\mu_j^{(\alpha)}\right)^2+1\right)}, \qquad j=1,2,\ldots,n.$$

Hence, from (2.1) we see that the eigenvalues of the matrix $\mathbf{B}(\alpha)^{-1}\mathbf{A}_s$ are given by

$$\lambda_{\pm}^{(\alpha,j)} = \pm\sqrt{\frac{\left(\mu_j^{(\alpha)}\right)^2+1}{\alpha^2+1}}, \qquad j=1,2,\ldots,n.$$

In addition, by letting

$$\mathbf{V}^{(\alpha)} = \begin{pmatrix} -\Phi^{(\alpha)} & \Psi^{(\alpha)} \\ \Psi^{(\alpha)} & \Phi^{(\alpha)} \end{pmatrix},$$

we know that $\mathbf{V}^{(\alpha)} \in \mathbb{R}^{2n\times 2n}$ is an orthogonal matrix. By straightforward computations we have

$$\mathbf{V}^{(\alpha)^T}\widehat{\Psi}_s^{(\alpha)}\mathbf{V}^{(\alpha)} = \begin{pmatrix} \sqrt{\alpha^2+1}\left(\mathbf{I}+\left(\Xi^{(\alpha)}\right)^2\right)^{\frac{1}{2}} & 0 \\ 0 & -\sqrt{\alpha^2+1}\left(\mathbf{I}+\left(\Xi^{(\alpha)}\right)^2\right)^{\frac{1}{2}} \end{pmatrix}.$$

It readily follows from (2.1) again that

$$\begin{aligned}\mathbf{B}(\alpha)^{-1}\mathbf{A}_s &= \frac{1}{\sqrt{\alpha^2+1}}\widehat{\mathbf{X}}^{(\alpha)}\mathbf{V}^{(\alpha)}\begin{pmatrix} -\left(\mathbf{I}+\left(\Xi^{(\alpha)}\right)^2\right)^{\frac{1}{2}} & 0 \\ 0 & \left(\mathbf{I}+\left(\Xi^{(\alpha)}\right)^2\right)^{\frac{1}{2}} \end{pmatrix}\mathbf{V}^{(\alpha)^T}\widehat{\mathbf{X}}^{(\alpha)^{-1}} \\ &= \widetilde{\mathbf{X}}^{(\alpha)}\Lambda^{(\alpha)}\widetilde{\mathbf{X}}^{(\alpha)^{-1}},\end{aligned}$$

where

$$\widetilde{\mathbf{X}}^{(\alpha)} = \widehat{\mathbf{X}}^{(\alpha)}\mathbf{V}^{(\alpha)} \quad \text{and} \quad \Lambda_{\pm}^{(\alpha)} = \pm\frac{1}{\sqrt{\alpha^2+1}}\left(\mathbf{I}+\left(\Xi^{(\alpha)}\right)^2\right)^{\frac{1}{2}}.$$

Moreover, as $\mathbf{Q}^{(\alpha)} \in \mathbb{R}^{n\times n}$ is an $(\alpha\mathbf{W}+\mathbf{T})^{-1}$-orthogonal matrix, we know that $\mathbf{X}^{(\alpha)}$ and, hence $\widehat{\mathbf{X}}^{(\alpha)}$, are orthogonal matrices. It follows immediately from the orthogonality of $\mathbf{V}^{(\alpha)} \in \mathbb{R}^{2n\times 2n}$ that $\widetilde{\mathbf{X}}^{(\alpha)} \in \mathbb{R}^{2n\times 2n}$ is orthogonal. Therefore, $\kappa_2\left(\widetilde{\mathbf{X}}^{(\alpha)}\right) = 1$. This demonstrates the validity of (i).

We now turn to prove (ii). To this end, we let

$$\mathbf{U}^{(\alpha)} = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathrm{i}\,\mathbf{I} & -\mathrm{i}\,\mathbf{I} \end{pmatrix},$$

which is evidently a unitary matrix. Then by straightforward computations we have

$$\mathbf{U}^{(\alpha)^*}\widehat{\Psi}^{(\alpha)}\mathbf{U}^{(\alpha)} = \begin{pmatrix} \mathbf{I}+\mathrm{i}\,\Xi^{(\alpha)} & 0 \\ 0 & \mathbf{I}-\mathrm{i}\,\Xi^{(\alpha)} \end{pmatrix}$$

and

$$\mathbf{U}^{(\alpha)^*}\mathbf{J}^{(\alpha)}\mathbf{U}^{(\alpha)} = \begin{pmatrix} (\alpha - \mathrm{i})\mathbf{I} & 0 \\ 0 & (\alpha + \mathrm{i})\mathbf{I} \end{pmatrix}.$$

Here and in the sequel, we use $(\cdot)^*$ to denote the conjugate transpose of the corresponding matrix. Now, it follows from (2.2) that

$$\begin{aligned}\mathbf{B}(\alpha)^{-1}\mathbf{A} &= \frac{1}{\alpha^2+1}\widehat{\mathbf{X}}^{(\alpha)}\mathbf{U}^{(\alpha)}\begin{pmatrix} \mathbf{I}+\mathrm{i}\,\Xi^{(\alpha)} & 0 \\ 0 & \mathbf{I}-\mathrm{i}\,\Xi^{(\alpha)} \end{pmatrix}\begin{pmatrix} (\alpha - \mathrm{i})\mathbf{I} & 0 \\ 0 & (\alpha+\mathrm{i})\mathbf{I} \end{pmatrix}\mathbf{U}^{(\alpha)^*}\widehat{\mathbf{X}}^{(\alpha)^{-1}} \\ &= \widetilde{\mathbf{X}}^{(\alpha)}\Lambda^{(\alpha)}\widetilde{\mathbf{X}}^{(\alpha)^{-1}},\end{aligned}$$

where

$$\widetilde{\mathbf{X}}^{(\alpha)} = \widehat{\mathbf{X}}^{(\alpha)}\mathbf{U}^{(\alpha)} \quad \text{and} \quad \Lambda_{\pm}^{(\alpha)} = \frac{\alpha \pm \mathrm{i}}{\alpha^2+1}\left(\mathbf{I} \mp \mathrm{i}\,\Xi^{(\alpha)}\right).$$

Moreover, as $\mathbf{Q}^{(\alpha)} \in \mathbb{R}^{n\times n}$ is an $(\alpha\mathbf{W}+\mathbf{T})^{-1}$-orthogonal matrix, we know that $\mathbf{X}^{(\alpha)}$ and, hence, $\widehat{\mathbf{X}}^{(\alpha)}$, are orthogonal matrices. It follows immediately from the fact that $\mathbf{U}^{(\alpha)} \in \mathbb{C}^{2n\times 2n}$ is unitary that $\widetilde{\mathbf{X}}^{(\alpha)} \in \mathbb{C}^{2n\times 2n}$ is unitary, too. Therefore, $\kappa_2\big(\widetilde{\mathbf{X}}^{(\alpha)}\big) = 1$. □

Based on Theorem 2.1 we can further derive bounds for the eigenvalues of the preconditioned matrices $\mathbf{B}(\alpha)^{-1}\mathbf{A}_s$ and $\mathbf{B}(\alpha)^{-1}\mathbf{A}$.

Theorem 2.2 *Let $\boldsymbol{A}_s, \boldsymbol{A} \in \mathbb{R}^{2n\times 2n}$ be the block two-by-two matrices defined in* (1.4) and (1.1)*, respectively, with $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ being symmetric positive semidefinite matrices satisfying* $\mathrm{null}(\boldsymbol{W}) \cap \mathrm{null}(\boldsymbol{T}) = \{0\}$*, and let α be a positive constant. Define $\boldsymbol{Z}^{(\alpha)} = (\alpha\boldsymbol{W}+\boldsymbol{T})^{-\frac{1}{2}}(\boldsymbol{W}-\alpha\boldsymbol{T})(\alpha\boldsymbol{W}+\boldsymbol{T})^{-\frac{1}{2}}$ and denote by $\mu_1^{(\alpha)}, \mu_2^{(\alpha)}, \ldots, \mu_n^{(\alpha)}$ the eigenvalues of the symmetric matrix $\boldsymbol{Z}^{(\alpha)} \in \mathbb{R}^{n\times n}$. Assume that*

$$\mathrm{sp}\left(\boldsymbol{Z}^{(\alpha)}\right) \subseteq \left[\mu_{\min}^{(\alpha)}, \quad \mu_{\max}^{(\alpha)}\right]$$

and

$$\mu^{(\alpha)} = \max\left\{\left|\mu_{\min}^{(\alpha)}\right|, \quad \left|\mu_{\max}^{(\alpha)}\right|\right\}, \quad \eta^{(\alpha)} = \max\left\{\left|\alpha\mu_{\max}^{(\alpha)} - 1\right|, \quad \left|\alpha\mu_{\min}^{(\alpha)} - 1\right|\right\}.$$

Then

(a) the eigenvalues of the matrix $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}_s$ are bounded as

$$\mathrm{sp}\left(\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}_s\right) \subseteq \left[-\sqrt{\tfrac{(\mu^{(\alpha)})^2+1}{\alpha^2+1}}, \quad -\tfrac{1}{\sqrt{\alpha^2+1}}\right] \bigcup \left[\tfrac{1}{\sqrt{\alpha^2+1}}, \quad \sqrt{\tfrac{(\mu^{(\alpha)})^2+1}{\alpha^2+1}}\right];$$

(b) the eigenvalues of the matrix $\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}$ are bounded as

$$\mathrm{sp}(\boldsymbol{B}(\alpha)^{-1}\boldsymbol{A}) \subseteq \left[\frac{\alpha+\mu_{\min}^{(\alpha)}}{\alpha^2+1}, \quad \frac{\alpha+\mu_{\max}^{(\alpha)}}{\alpha^2+1}\right] \times \left[-\frac{\eta^{(\alpha)}}{\alpha^2+1}, \quad \frac{\eta^{(\alpha)}}{\alpha^2+1}\right].$$

Proof The bounds in (a) and (b) straightforwardly follow from the expressions for the eigenvalues of matrices $\mathbf{B}(\alpha)^{-1}\mathbf{A}_s$ and $\mathbf{B}(\alpha)^{-1}\mathbf{A}$ given in Theorem 2.1, and from the bounds assumed for the eigenvalues $\mu_j^{(\alpha)}$, $j = 1, 2, \ldots, n$, of the matrix $\mathbf{Z}^{(\alpha)}$. □

Remark 2.1 *If* $\alpha = 1$, *then Theorems 2.1 and 2.2 show that when*

$$\boldsymbol{B} := \begin{pmatrix} \boldsymbol{W}+\boldsymbol{T} & 0 \\ 0 & \boldsymbol{W}+\boldsymbol{T} \end{pmatrix}$$

is used to precondition the matrices $\boldsymbol{A}_s$ *and* $\boldsymbol{A}$, *the eigenvalues of the preconditioned matrix* $\boldsymbol{B}^{-1}\boldsymbol{A}_s$ *are clustered within the real interval* $\left[-1, -\frac{\sqrt{2}}{2}\right] \cup \left[\frac{\sqrt{2}}{2}, 1\right]$, *and those of the matrix* $\boldsymbol{B}^{-1}\boldsymbol{A}$ *are located within the complex rectangle* $(0, 1] \times [-1, 1]$. *In addition, both* $\boldsymbol{B}^{-1}\boldsymbol{A}_s$ *and* $\boldsymbol{B}^{-1}\boldsymbol{A}$ *are orthogonally and unitarily diagonalizable with the matrices* $\widetilde{\boldsymbol{X}}^{(1)}$ *formed by their eigenvectors, respectively. Hence, the ABD-preconditioned MINRES can be expected to converge very rapidly when it is employed to solve the symmetric form of the block two-by-two linear system* (1.1) *with respect to the coefficient matrix* $\boldsymbol{A}_s$ *defined in* (1.4). *However, when employed to solve the block two-by-two linear system* (1.1), *the ABD-preconditioned GMRES may converge very slowly, as the real parts of the eigenvalues of the preconditioned matrix* $\boldsymbol{B}^{-1}\boldsymbol{A}$ *could be very close to 0, especially for large matrices.*

3 Applications to the numerical solution of distributed control problems

Consider the distributed control problem

$$\min_{u,f} \frac{1}{2}\|u - u_*\|_2^2 + \beta\|f\|_2^2, \tag{3.1}$$

$$\text{subject to} \quad -\nabla^2 u = f \quad \text{in} \quad \Omega, \tag{3.2}$$

$$\text{with} \quad u = \upsilon_1 \quad \text{on} \quad \partial\Omega_1 \quad \text{and} \quad \frac{\partial u}{\partial \vec{n}} = \upsilon_2 \quad \text{on} \quad \partial\Omega_2, \tag{3.3}$$

where Ω is a domain in $\mathbb{R}^2$ or $\mathbb{R}^3$, $\partial\Omega$ is the boundary of Ω, $\partial\Omega_1$ and $\partial\Omega_2$ are two parts of $\partial\Omega$ satisfying $\partial\Omega_1 \cup \partial\Omega_2 = \partial\Omega$ and $\partial\Omega_1 \cap \partial\Omega_2 = \emptyset$, and $\vec{n}$ is the outward normal of Ω. Such problems, introduced by Lions in [24], consist of a cost functional (3.1) to be minimized subject to a *partial differential equation* (**PDE**) problem (3.2)–(3.3) posed on the domain Ω. Here, the function u_* (the "desired state") is known, and we want to find u which satisfies the PDE problem and is as close to u_* as possible in the L_2-norm sense. For recent references on this topic, see, e.g., [6, 23].

When the PDE-constrained optimization problem (3.1)–(3.3) is treated with the discretize-then-optimize approach [29] through application of a Galerkin finite-element method to its weak formulation, we obtain a KKT system in the following saddle-point form:

$$\begin{pmatrix} 2\beta M & 0 & -M \\ 0 & M & K^T \\ -M & K & 0 \end{pmatrix} \begin{pmatrix} f \\ u \\ \phi \end{pmatrix} = \begin{pmatrix} 0 \\ b \\ d \end{pmatrix}, \tag{3.4}$$

where $M \in \mathbb{R}^{m\times m}$ is the mass matrix, $K \in \mathbb{R}^{m\times m}$ is the stiffness matrix (the discrete Laplacian), $\beta > 0$ is the regularization parameter, $d \in \mathbb{R}^m$ contains the terms coming from the boundary values of the discrete solution, and $b \in \mathbb{R}^m$ is the Galerkin projection of the discrete state u_*. In addition, ϕ is a vector of Lagrange multipliers.

After eliminating the Lagrange multiplier ϕ, we can equivalently rewrite the saddle-point linear system (3.4) as the block two-by-two linear system

$$Ax \equiv \begin{pmatrix} \frac{1}{2\beta}M & K^T \\ -K & M \end{pmatrix}\begin{pmatrix} u \\ f \end{pmatrix} = \begin{pmatrix} \frac{1}{2\beta}b \\ -d \end{pmatrix} \equiv g. \tag{3.5}$$

Recall that $M \in \mathbb{R}^{m\times m}$ is the mass matrix and is, thus, symmetric positive definite. Therefore, the matrix $A \in \mathbb{R}^{2m\times 2m}$ is positive real, i.e., its symmetric part is positive definite.

In the remainder of this section, we assume that the stiffness matrix K is symmetric and positive semidefinite. Through symmetric block-scaling by the diagonal matrix

$$D = \begin{pmatrix} -\sqrt{2\beta}I & 0 \\ 0 & I \end{pmatrix},$$

where $I \in \mathbb{R}^{m\times m}$ represents the identity matrix, we can reformulate the block two-by-two linear system (3.5) into the form of (1.1), with

$$\mathbf{A} = \begin{pmatrix} M & -\sqrt{2\beta}K \\ \sqrt{2\beta}K & M \end{pmatrix}, \qquad \text{i.e.,} \quad \mathbf{W} = M \quad \text{and} \quad \mathbf{T} = \sqrt{2\beta}K, \tag{3.6}$$

and

$$\begin{cases} \mathbf{x} = \begin{pmatrix} -\frac{1}{\sqrt{2\beta}}u \\ f \end{pmatrix}, & \text{i.e., } \mathbf{y} = -\frac{1}{\sqrt{2\beta}}u \quad \text{and } \mathbf{z} = f, \\ \mathbf{g} = \begin{pmatrix} -\frac{1}{\sqrt{2\beta}}b \\ -d \end{pmatrix}, & \text{i.e., } \mathbf{p} = -\frac{1}{\sqrt{2\beta}}b \quad \text{and } \mathbf{q} = -d. \end{cases}$$

Now from (1.3) we know that the additive block diagonal preconditioning matrix $\mathbf{B}(\alpha)$ of the matrix $\mathbf{A}$ in (3.6) (see also (1.1)) is defined by

$$\mathbf{B}(\alpha) = \begin{pmatrix} \alpha\mathbf{W} + \mathbf{T} & 0 \\ 0 & \alpha\mathbf{W} + \mathbf{T} \end{pmatrix} = \begin{pmatrix} \alpha M + \sqrt{2\beta}K & 0 \\ 0 & \alpha M + \sqrt{2\beta}K \end{pmatrix}.$$

Therefore, the additive block diagonal preconditioning matrix for the block two-by-two matrix A in (3.5) is as follows:

$$B(\alpha) = \begin{pmatrix} \frac{1}{2\beta}(\alpha M + \sqrt{2\beta}K) & 0 \\ 0 & \alpha M + \sqrt{2\beta}K \end{pmatrix}. \tag{3.7}$$

Of course, the matrix $B(\alpha)$ can also be employed to precondition the symmetric form of the linear system (3.5) defined as follows:

$$A_s x \equiv \begin{pmatrix} \frac{1}{2\beta}M & K \\ K & -M \end{pmatrix}\begin{pmatrix} u \\ f \end{pmatrix} = \begin{pmatrix} \frac{1}{2\beta}b \\ d \end{pmatrix} \equiv g_s, \tag{3.8}$$

We recall that the PMHSS preconditioning matrix for the block two-by-two matrix A in (3.5) is given by

$$F(\alpha) = (\alpha+1)P(\alpha;\beta)\begin{pmatrix} \alpha M + \sqrt{2\beta}K & 0 \\ 0 & \alpha M + \sqrt{2\beta}K \end{pmatrix}, \qquad (3.9)$$

where

$$P(\alpha;\beta) = \frac{1}{4\alpha\beta}\begin{pmatrix} I & \sqrt{2\beta}I \\ -\sqrt{2\beta}I & 2\beta I \end{pmatrix};$$

see [9] for a detailed derivation.

Note that the matrix $\alpha M + \sqrt{2\beta}K$ is symmetric positive definite. Hence, in implementations the actions of the preconditioning matrices $B(\alpha)$ and $F(\alpha)$ can be effectively accomplished either exactly by Cholesky factorization or inexactly by some conjugate gradient or multigrid scheme; see [12].

In particular, if $\alpha = 1$ we further have

$$B := \begin{pmatrix} \frac{1}{2\beta}(M + \sqrt{2\beta}K) & 0 \\ 0 & M + \sqrt{2\beta}K \end{pmatrix} \qquad (3.10)$$

and

$$F := P\begin{pmatrix} M + \sqrt{2\beta}K & 0 \\ 0 & M + \sqrt{2\beta}K \end{pmatrix}, \qquad (3.11)$$

with

$$P = \frac{1}{2\beta}\begin{pmatrix} I & \sqrt{2\beta}I \\ -\sqrt{2\beta}I & 2\beta I \end{pmatrix}.$$

Below we state several remarks about the ABD preconditioning matrices $B(\alpha)$ and B.

Remark 3.1 *We stress that our solution approach is not limited to the special (and rather simple) model problem* (3.1)–(3.3)*, and that it can handle any kind of distributed control problem leading to saddle-point linear systems of the form* (3.4) *with K symmetric and positive definite (or semidefinite). Hence, a broad class of elliptic PDE constraints can be accommodated besides Poisson's equation.*

Remark 3.2 *In* [35]*, by making use of certain nonstandard norms the author derived a general form of block two-by-two preconditioning matrices for the saddle-point matrices, which specifically lead to the preconditioning matrix B defined in* (3.10) *for the symmetric form A_s of the block two-by-two matrix A in* (3.8) and (3.5)*, respectively.*

Remark 3.3 *Of course, we have realized that reducing the block three-by-three KKT system* (3.4) *to the block two-by-two linear system* (3.5) *through eliminating the Lagrange multiplier is not always possible, as the matrix M could be singular or rectangular in some cases. For example, if the control and state are discretized by different finite elements, then the matrix M will be rectangular. We refer to* [27] *for several related preconditioners based on Schur complement approximations* *for the saddle-point matrix resulting from the discretize-then-optimize approach to the distributed control problem* (3.1)–(3.3)*, which may effectively treat the cases mentioned above.*

4 Numerical results

In this section, we use the following example to examine the numerical behavior of the additive block diagonal preconditioning matrix and the corresponding preconditioned Krylov subspace iteration methods:

Example 4.1 [29] *Let* $\Omega = [0, 1]^2$ *be a unit square and consider the distributed control problem* (3.1)–(3.3)*, with* $\partial\Omega_2 = \emptyset$*,* $\upsilon_1 = u_*$ *and*

$$u_* = \begin{cases} (2x-1)^2(2y-1)^2, & \text{if } (x, y) \in \left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right], \\ 0, & \text{otherwise}. \end{cases}$$

To this end, we solve the system of linear equations (3.5) by the GMRES method preconditioned with the ABD preconditioning matrices $B(\alpha)$ and B, defined by (3.7)

Table 1 Experimental optimal parameters α_{opt} for $B(\alpha)$- and $B^{(\text{app})}(\alpha)$-MINRES methods

β	h	α_{opt} for $B(\alpha)$-MINRES	α_{opt} for $B^{(\text{app})}(\alpha)$-MINRES
10^{-2}	2^{-2}	$[0.07, 0.12] \cup [0.15, 0.24]$	$[0.07, 0.12] \cup [0.15, 0.24]$
	2^{-3}	$[0.06, 0.08] \cup [0.16, 0.21]$	$[0.03, 0.16]$
	2^{-4}	$[0.03, 0.07] \cup [0.13, 0.22]$	$[0.01, 0.16]$
	2^{-5}	$[0.01, 0.25]$	0.01
	2^{-6}	$[0.01, 0.30]$	$[0.01, 0.87]$
10^{-4}	2^{-2}	$[0.58, 0.76] \cup [0.91, 1.02]$	$[0.58, 0.76] \cup [0.91, 1.02]$
	2^{-3}	$[0.43, 0.47] \cup [0.52, 0.63]$	$[0.43, 0.47] \cup [0.52, 0.63]$
	2^{-4}	$[0.51, 0.59]$	$[0.49, 0.54]$
	2^{-5}	$[0.24, 0.89]$	$[0.01, 0.19]$
	2^{-6}	$[0.23, 0.86]$	$[0.01, 0.03]$
10^{-6}	2^{-2}	$[1.97, 3.00]$	$[1.97, 3.00]$
	2^{-3}	$[1.31, 2.79]$	$[1.31, 2.79]$
	2^{-4}	$[0.94, 1.32]$	$[0.94, 1.32]$
	2^{-5}	$[0.90, 1.10]$	$[0.90, 1.10]$
	2^{-6}	$[0.86, 1.14]$	$[0.54, 0.86]$
10^{-8}	2^{-2}	$[1.27, 3.00]$	$[1.27, 3.00]$
	2^{-3}	$[0.99, 3.00]$	$[0.99, 3.00]$
	2^{-4}	$[1.88, 3.00]$	$[1.88, 3.00]$
	2^{-5}	$[1.28, 2.20]$	$[1.28, 2.20]$
	2^{-6}	$[0.99, 1.18]$	$[0.99, 1.18]$

Table 2 Experimental optimal parameters α_{opt} for $B(\alpha)$- and $B^{(\text{app})}(\alpha)$-GMRES methods

β	h	α_{opt} for $B(\alpha)$-GMRES	α_{opt} for $B^{(\text{app})}(\alpha)$-GMRES
10^{-2}	2^{-2}	$[0.01,\ 1.52]$	$[0.01,\ 1.52]$
	2^{-3}	$[0.01,\ 0.11]$	$[0.01,\ 0.11]$
	2^{-4}	$[0.01,\ 0.03]$	$[0.01,\ 0.59]$
	2^{-5}	$[0.01,\ 0.02] \cup [0.66,\ 0.70]$	$[0.01,\ 0.55] \cup [1.41,\ 2.34]$
	2^{-6}	$[0.01,\ 0.02] \cup [0.65,\ 0.70]$	$[0.01,\ 0.69]$
10^{-4}	2^{-2}	$[0.01,\ 3.00]$	$[0.01,\ 3.00]$
	2^{-3}	$[0.08,\ 0.20]$	$[0.08,\ 0.20]$
	2^{-4}	$[0.02,\ 0.03] \cup [0.28,\ 0.29]$	0.26
	2^{-5}	$[0.13,\ 0.16] \cup [0.39,\ 0.41]$	$[0.01,\ 0.05] \cup \{0.33\}$
	2^{-6}	$[0.04,\ 0.11] \cup [0.25,\ 0.30] \cup [0.49,\ 0.52]$	$[0.01,\ 0.10]$
10^{-6}	2^{-2}	$[0.88,\ 3.00]$	$[0.88,\ 3.00]$
	2^{-3}	$[2.53,\ 2.55]$	$[2.53,\ 2.55]$
	2^{-4}	$[0.80,\ 0.82]$	$[0.80,\ 0.82]$
	2^{-5}	$[0.58,\ 0.65] \cup [0.74,\ 0.87] \cup [1.01,\ 1.02]$	$[0.58,\ 0.65] \cup [0.74,\ 0.87] \cup [1.01,\ 1.02]$
	2^{-6}	$\{0.61,\ 0.77\}$	$[0.53,\ 0.54]$
10^{-8}	2^{-2}	$[1.53,\ 1.99]$	$[1.53,\ 1.99]$
	2^{-3}	$[0.98,\ 1.02]$	$[0.98,\ 1.02]$
	2^{-4}	$[0.76,\ 1.03] \cup [2.07,\ 3.00]$	$[0.76,\ 1.03] \cup [2.07,\ 3.00]$
	2^{-5}	$[0.83,\ 0.98] \cup [1.33,\ 1.47]$	$[0.83,\ 0.98] \cup [1.33,\ 1.47]$
	2^{-6}	$[0.73,\ 0.78] \cup [0.98,\ 1.02]$	$[0.73,\ 0.78] \cup [0.98,\ 1.02]$

and (3.10), with the PMHSS preconditioning matrices $F(\alpha)$ and F, defined by (3.9) and (3.11), and with their inexact variants $B^{(\text{app})}(\alpha)$, $B^{(\text{app})}$ and $F^{(\text{app})}(\alpha)$, $F^{(\text{app})}$. In addition, we also solve the symmetric form (3.8) of the linear system (3.5) by the MINRES method preconditioned with the ABD preconditioning matrices $B(\alpha)$ and

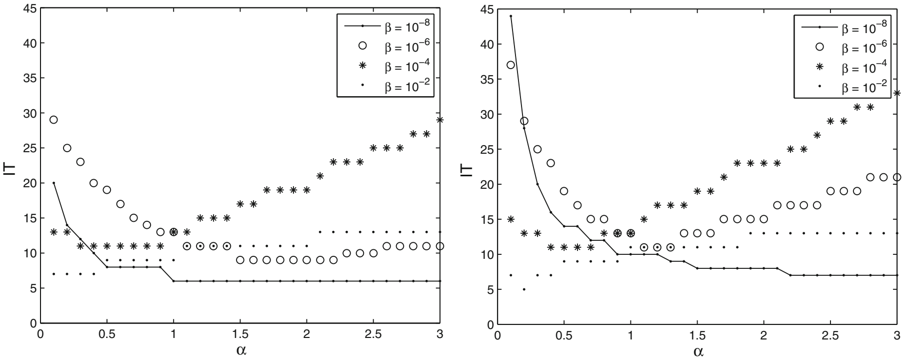

Fig. 1 Number of iteration steps versus α for $B(\alpha)$-MINRES when $\beta = 10^{-2}$, 10^{-4}, 10^{-6} and 10^{-8}, with $h = 2^{-3}$ (*left*) and $h = 2^{-4}$ (*right*)

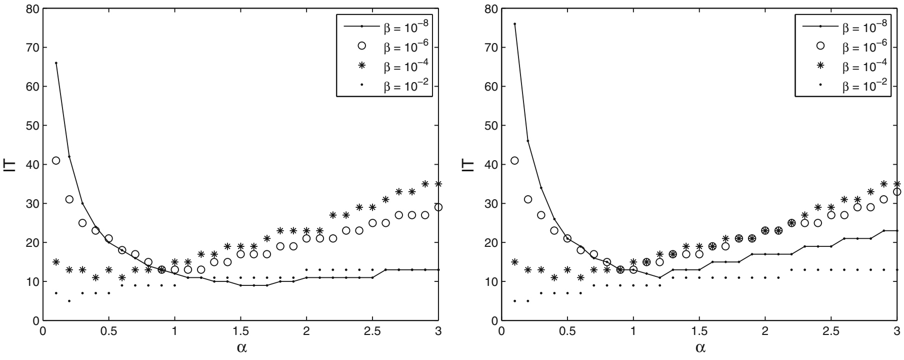

Fig. 2 Number of iteration steps versus α for $B(\alpha)$-MINRES when $\beta = 10^{-2}$, 10^{-4}, 10^{-6} and 10^{-8}, with $h = 2^{-5}$ (*left*) and $h = 2^{-6}$ (*right*)

B, and their inexact variants $B^{(\mathrm{app})}(\alpha)$ and $B^{(\mathrm{app})}$. In computing the actions of the inverses of $B^{(\mathrm{app})}(\alpha)$, $B^{(\mathrm{app})}$ and $F^{(\mathrm{app})}(\alpha)$, $F^{(\mathrm{app})}$, the inverses of the approximations $G^{(\mathrm{app})}(\alpha)$ and $G^{(\mathrm{app})}$ corresponding to the matrices $G(\alpha) := \alpha M + \sqrt{2\beta}K$ and $G := M + \sqrt{2\beta}K$ are implemented by 20 steps of Chebyshev semi-iteration approximation; see [29, 33]. Here, the bounds of the eigenvalues of the matrices $G(\alpha)$ and G are approximately given by those of the matrices αM and M, respectively; this is reasonable especially when the regularization parameter β is small.

In our implementations, all iteration processes are terminated once the Euclidean norms of the current residuals are reduced by a factor of 10^4 from those of the initial residuals, and the optimal iteration parameters α_{opt} adopted in the exact and the inexact ABD preconditioners $B(\alpha)$ and $B^{(\mathrm{app})}(\alpha)$ are experimentally found ones that minimize the number of total iteration steps of the corresponding iteration processes; see Tables 1 and 2. The principle adopted for determining such an optimal value of the iteration parameter α is as follows: Given a stopping criterion, say, 10^{-4}, we can obtain a number of iteration steps for the designated iteration method at each mesh point with respect to α in the interval (0, 100] with the meshsize 0.01. Then the values

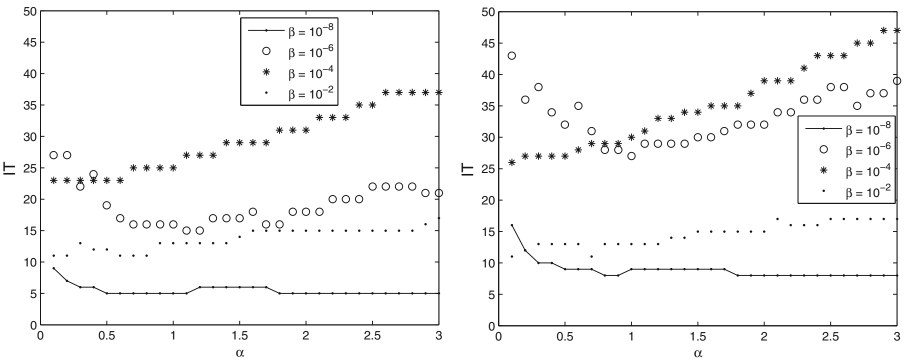

Fig. 3 Number of iteration steps versus α for $B(\alpha)$-GMRES when $\beta = 10^{-2}$, 10^{-4}, 10^{-6} and 10^{-8}, with $h = 2^{-3}$ (*left*) and $h = 2^{-4}$ (*right*)

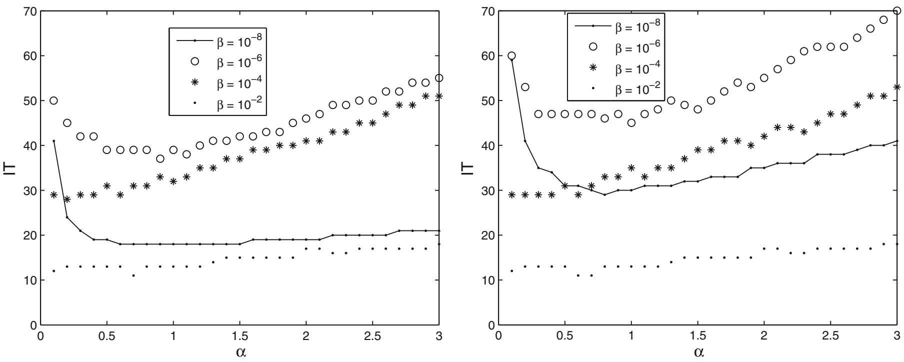

Fig. 4 Number of iteration steps versus α for $B(\alpha)$-GMRES when $\beta = 10^{-2}$, 10^{-4}, 10^{-6} and 10^{-8}, with $h = 2^{-5}$ (*left*) and $h = 2^{-6}$ (*right*)

which lead to the minimal number of iteration steps are the 'optimal' values of the iteration parameter α.

From Tables 1–2 and Figs. 1, 2, 3 and 4 we see that in most cases the optimal iteration parameter α_{opt} forms an interval or a union of intervals, but not a single value. We do not pick up a single value as the optimal value of the iteration parameter because a single optimal value may be changed with the adopted stopping criterion. However, for all optimal values of the iteration parameter α in the same interval or the same union of intervals given in Tables 1 and 2, the iteration methods can almost achieve different orders of the minimal ratios of the Euclidean norms of the current and the initial residuals, e.g., 10^{-3} and 10^{-4}. In other words, for different stopping criteria the ranges of the optimal iteration parameter α_{opt} are about the same; see Fig. 5.

In Tables 3 and 4, we list the number of iteration steps, the computing time (in parentheses) and the speed-up (in brackets) with respect to the ABD preconditioner and its approximate variants, which are employed to precondition the MINRES and

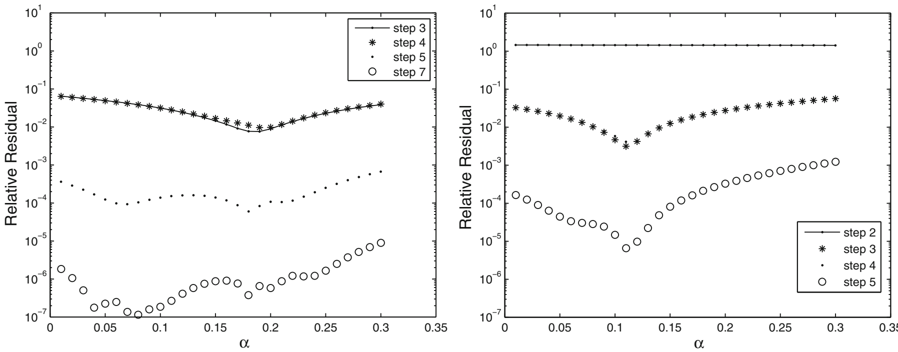

Fig. 5 Relative residual versus α for ABD-preconditioned MINRES (*left*) and approximate ABD-preconditioned MINRES (*right*) when $\beta = 10^{-2}$ and $h = 2^{-3}$

Table 3 Iteration steps, computing times (in parentheses) and speed-ups (in brackets) for additive block diagonal preconditioners

Method		MINRES for (3.8)						GMRES for (3.5)					
β	h	$B(\alpha_{opt})$			B			$B(\alpha_{opt})$			B		
10^{-2}	2^{-2}	5	(1.86e-2)	[1.37]	9	(2.07e-3)	[12.30]	11	(1.86e-2)	[1.37]	11	(2.54e-3)	[10.01]
	2^{-3}	5	(2.10e-2)	[1.17]	9	(6.40e-3)	[3.84]	11	(2.35e-2)	[1.04]	13	(1.01e-2)	[2.43]
	2^{-4}	5	(3.28e-2)	[1.30]	9	(2.26e-2)	[1.88]	11	(6.49e-2)	[0.66]	13	(3.94e-2)	[1.08]
	2^{-5}	5	(5.65e-2)	[2.04]	9	(6.46e-2)	[1.78]	11	(1.39e-1)	[0.83]	13	(1.26e-1)	[0.92]
	2^{-6}	5	(2.30e-1)	[1.77]	9	(3.61e-1)	[1.13]	11	(5.90e-1)	[0.69]	13	(6.45e-1)	[0.63]
10^{-4}	2^{-2}	8	(1.85e-2)	[1.33]	8	(1.60e-3)	[15.37]	13	(1.86e-2)	[1.33]	13	(2.90e-3)	[8.49]
	2^{-3}	9	(2.35e-2)	[1.12]	11	(7.88e-3)	[3.35]	21	(3.52e-2)	[0.75]	25	(2.19e-2)	[1.20]
	2^{-4}	10	(4.30e-2)	[1.20]	12	(3.04e-2)	[1.70]	23	(8.69e-2)	[0.59]	29	(8.33e-2)	[0.62]
	2^{-5}	11	(9.43e-2)	[1.49]	13	(9.56e-2)	[1.47]	25	(2.49e-1)	[0.57]	31	(2.99e-1)	[0.47]
	2^{-6}	11	(4.77e-1)	[1.24]	13	(5.51e-1)	[1.08]	27	(1.31e+0)	[0.45]	31	(1.51e+0)	[0.39]
10^{-6}	2^{-2}	6	(2.08e-2)	[1.10]	10	(2.01e-3)	[11.38]	8	(1.68e-2)	[1.36]	8	(1.65e-3)	[13.84]
	2^{-3}	8	(2.49e-2)	[1.14]	10	(7.98e-3)	[3.56]	16	(3.02e-2)	[0.94]	18	(1.59e-2)	[1.79]
	2^{-4}	10	(4.51e-2)	[1.14]	10	(2.70e-2)	[1.90]	26	(9.61e-2)	[0.53]	27	(7.72e-2)	[0.66]
	2^{-5}	12	(1.06e-1)	[1.20]	12	(8.35e-2)	[1.52]	34	(3.49e-1)	[0.36]	35	(3.41e-1)	[0.37]
	2^{-6}	13	(5.72e-1)	[1.04]	13	(5.38e-1)	[1.11]	38	(1.86e+0)	[0.32]	41	(1.95e+0)	[0.30]
10^{-8}	2^{-2}	4	(1.87e-2)	[1.28]	6	(1.29e-3)	[18.59]	4	(1.59e-2)	[1.50]	5	(9.35e-4)	[25.63]
	2^{-3}	6	(2.12e-2)	[1.25]	6	(4.01e-3)	[6.61]	5	(1.88e-2)	[1.41]	5	(3.30e-3)	[8.03]
	2^{-4}	6	(3.26e-2)	[1.47]	10	(2.53e-2)	[1.89]	9	(4.82e-2)	[0.99]	9	(2.65e-2)	[1.80]
	2^{-5}	8	(7.80e-2)	[1.70]	10	(6.97e-2)	[1.91]	19	(1.90e-1)	[0.70]	21	(1.96e-1)	[0.68]
	2^{-6}	10	(4.36e-1)	[1.36]	10	(4.14e-1)	[1.44]	28	(1.38e+0)	[0.43]	28	(1.34e+0)	[0.44]

Table 4 Iteration steps, computing times (in parentheses) and speed-ups (in brackets) for inexact additive block diagonal preconditioners

Method		MINRES for (3.8)						GMRES for (3.5)					
β	h	$B^{(\text{app})}(\alpha_{\text{opt}})$			$B^{(\text{app})}$			$B^{(\text{app})}(\alpha_{\text{opt}})$			$B^{(\text{app})}$		
10^{-2}	2^{-2}	5	(7.71e-3)	[1.25]	9	(6.45e-3)	[1.50]	11	(1.10e-2)	[0.88]	11	(8.26e-3)	[1.17]
	2^{-3}	5	(1.02e-2)	[1.04]	9	(1.06e-2)	[1.00]	11	(1.54e-2)	[0.69]	13	(1.47e-2)	[0.72]
	2^{-4}	9	(2.70e-2)	[0.96]	11	(2.55e-2)	[1.02]	13	(4.34e-2)	[0.60]	15	(4.23e-2)	[0.61]
	2^{-5}	21	(1.93e-1)	[1.13]	23	(2.06e-1)	[1.06]	25	(3.01e-1)	[0.72]	26	(2.98e-1)	[0.73]
	2^{-6}	63	(1.50e+0)	[0.55]	65	(1.47e+0)	[0.56]	69	(2.04e+0)	[0.40]	70	(2.05e+0)	[0.40]
10^{-4}	2^{-2}	8	(9.47e-3)	[1.01]	8	(5.40e-3)	[1.76]	13	(1.22e-2)	[0.78]	13	(9.17e-3)	[1.04]
	2^{-3}	9	(1.43e-2)	[0.97]	11	(1.15e-2)	[1.20]	21	(2.96e-2)	[0.47]	25	(3.20e-2)	[0.43]
	2^{-4}	9	(2.52e-2)	[1.09]	12	(2.78e-2)	[0.99]	23	(6.43e-2)	[0.43]	29	(7.93e-2)	[0.35]
	2^{-5}	13	(1.23e-1)	[1.20]	19	(1.67e-1)	[0.88]	29	(3.38e-1)	[0.44]	33	(3.85e-1)	[0.38]
	2^{-6}	49	(1.15e+0)	[0.46]	55	(1.29e+0)	[0.41]	63	(1.93e+0)	[0.28]	67	(2.04e+0)	[0.26]
10^{-6}	2^{-2}	6	(8.98e-3)	[1.20]	10	(6.80e-3)	[1.59]	8	(8.35e-3)	[1.29]	8	(5.47e-3)	[1.98]
	2^{-3}	8	(1.44e-2)	[1.11]	10	(1.26e-2)	[1.27]	16	(2.35e-2)	[0.68]	18	(2.45e-2)	[0.65]
	2^{-4}	10	(2.74e-2)	[1.07]	10	(2.30e-2)	[1.27]	26	(7.05e-2)	[0.42]	27	(7.15e-2)	[0.41]
	2^{-5}	12	(1.10e-1)	[1.20]	12	(1.07e-1)	[1.24]	34	(4.02e-1)	[0.33]	35	(4.20e-1)	[0.32]
	2^{-6}	11	(2.72e-1)	[1.01]	13	(3.12e-1)	[0.88]	37	(1.11e+0)	[0.25]	40	(1.17e+0)	[0.24]
10^{-8}	2^{-2}	4	(7.48e-3)	[1.15]	6	(4.45e-3)	[1.94]	4	(5.60e-3)	[1.54]	5	(3.57e-3)	[2.42]
	2^{-3}	6	(1.04e-2)	[1.20]	6	(6.74e-3)	[1.86]	5	(8.32e-3)	[1.51]	5	(5.23e-3)	[2.40]
	2^{-4}	6	(1.98e-2)	[1.42]	10	(2.37e-2)	[1.19]	9	(2.98e-2)	[0.94]	9	(2.50e-2)	[1.12]
	2^{-5}	8	(7.58e-2)	[1.86]	10	(8.80e-2)	[1.60]	19	(2.17e-1)	[0.65]	21	(2.34e-1)	[0.60]
	2^{-6}	10	(2.41e-1)	[1.17]	10	(2.39e-1)	[1.18]	28	(8.22e-1)	[0.34]	28	(8.15e-1)	[0.35]

the GMRES methods, respectively. The speed-up is defined as the ratio of the computing times with respect to $F(\alpha_{\text{opt}})$ (or $F^{(\text{app})}(\alpha_{\text{opt}})$ in the approximate variants) and the referred preconditioner when they are used in proper iterative solvers. The optimal iteration parameters α_{opt} used in these two tables are chosen from Tables 1 and 2 such that they are, if not single points, the ones closest to 1.0.

From Table 3 we see that the numbers of iteration steps of the $B(\alpha_{\text{opt}})$- and B-preconditioned MINRES methods (briefly written as $B(\alpha_{\text{opt}})$- and B-MINRES) are roughly independent of the discretization meshsize $h = \frac{1}{\sqrt{m+1}}$ for all tested values of β, those of the $B(\alpha_{\text{opt}})$- and B-preconditioned GMRES methods (briefly written as $B(\alpha_{\text{opt}})$- and B-GMRES) are, however, increasing distinctly as h decreases from 2^{-2} to 2^{-6}. This phenomenon is well illustrated by the eigenvalue bounds given in Remark 2.1, that is, the preconditioned matrix $B^{-1}A$ is very ill-conditioning, as the real and the imaginary parts of some of its eigenvalues are very close to 0 and ± 1, respectively; see Fig. 6.

Also, Table 3 shows that the speed-ups of $B(\alpha_{\text{opt}})$- and B-MINRES are always larger than 1.0. They can even reach 2.04 for $B(\alpha_{\text{opt}})$-MINRES when $h = 2^{-5}$ and $\beta = 10^{-2}$, and 18.59 for B-MINRES when $h = 2^{-2}$ and $\beta = 10^{-8}$. The speed-ups of $B(\alpha_{\text{opt}})$- and B-GMRES are, however, less than 1.0 for many cases of β especially when h is very small. This phenomenon indicates that the MINRES method for solving the symmetric saddle-point linear system (3.8) is more effective than the GMRES method for solving the nonsymmetric saddle-point linear system (3.5) when both of them are preconditioned by the additive block diagonal preconditioning matrix. Moreover, for most cases, $B(\alpha_{\text{opt}})$ shows better preconditioning property than B in terms of iteration steps, computing times and speed-ups. But there are some exceptions for which the preconditioner B leads to much smaller CPU times, e.g., for $h = 2^{-2}$. Because the numerical results with respect to $B(\alpha_{\text{opt}})$ are comparable to those with respect to B, counting the elapsed times for numerically searching the optimal iteration parameters we could simply take the iteration parameter α to be 1.0 in implementations of the ABD preconditioning matrix $B(\alpha)$, resulting in a parameter-free method.

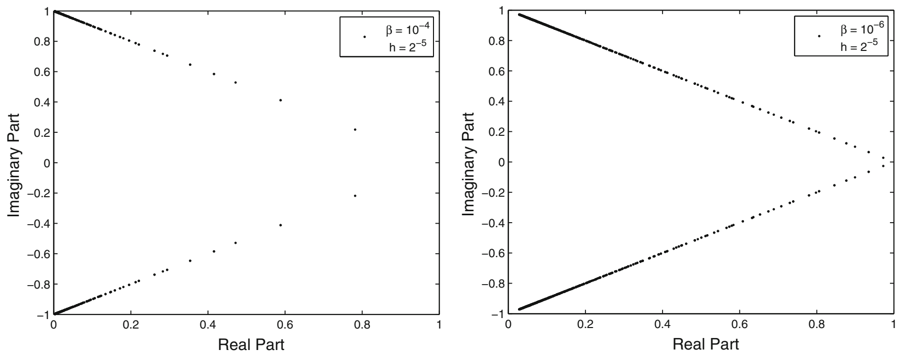

Fig. 6 Eigenvalue distribution of the preconditioned matrix $B^{-1}A$

Comparing the results in Tables 3 and 4 we observe that the approximated ABD preconditioners $B^{(\mathrm{app})}(\alpha_{\mathrm{opt}})$ and $B^{(\mathrm{app})}$ yield the same number of iteration steps but smaller computing times than the exact ABD preconditioners $B(\alpha_{\mathrm{opt}})$ and B when $\beta = 10^{-6}$ and 10^{-8}. Moreover, from the speed-ups we see that the MINRES methods preconditioned with $B^{(\mathrm{app})}(\alpha_{\mathrm{opt}})$ and $B^{(\mathrm{app})}$ outperform the GMRES methods preconditioned with $F^{(\mathrm{app})}(\alpha_{\mathrm{opt}})$ and $F^{(\mathrm{app})}$ in most cases. Because the numerical results with respect to $B^{(\mathrm{app})}(\alpha_{\mathrm{opt}})$ are comparable to those with respect to $B^{(\mathrm{app})}$, counting the elapsed times for numerically searching the optimal iteration parameters we could simply take the iteration parameter α to be 1.0 in implementations of the approximated ABD preconditioning matrix $B^{(\mathrm{app})}(\alpha)$, resulting in a parameter-free method, too.

5 Concluding remarks

We have constructed and analyzed a class of additive block diagonal preconditioning matrices which can be regarded as a simplified variant of the PMHSS preconditioner, proving that it enjoys similar theoretical properties to PMHSS when used to precondition the block two-by-two matrix $\mathbf{A}$ as well as its symmetric form $\mathbf{A}_s$.

In general, the block two-by-two matrix $\mathbf{A} \in \mathbb{R}^{2n\times 2n}$ is nonsingular and Theorem 2.1 holds true even when the matrix $\alpha\mathbf{W} + \mathbf{T}$ is symmetric, and is either positive or negative definite, which may allow the real symmetric matrix $\mathbf{W}$ or $\mathbf{T}$ to be indefinite. Indeed, if the matrix $\alpha\mathbf{W}+\mathbf{T}$ is symmetric negative definite, we may simply multiply -1 through both sides of the linear system (1.1), obtaining a new block two-by-two linear system satisfying the requirement that $\alpha\mathbf{W}+\mathbf{T}$ is a symmetric positive definite matrix. The eigenvalue bounds for the preconditioned matrices given in Theorem 2.2 are valid, however, only when each of the matrices $\mathbf{W}$ and $\mathbf{T}$ is either positive or negative semidefinite and when $\mathrm{null}(\mathbf{W}) \cap \mathrm{null}(\mathbf{T}) = \{0\}$. These observations equally apply to the special case[2] $\mathbf{V} = \mathbf{W}$ of the PMHSS iteration methods introduced and discussed in [8, 9], i.e., under these weaker conditions the PMHSS-preconditioned matrices $\mathbf{F}(\alpha)^{-1}\mathbf{A}$ are unitarily diagonalizable. For more recent discussions on this topic, we refer to [21] and the references therein.

Acknowledgments The authors are very much indebted to the referees for providing very useful comments and constructive suggestions, which greatly improved the original manuscript of this paper.

References

1. Axelsson, O., Bai, Z.-Z., Qiu, S.-X.: A class of nested iteration schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part. Numer. Algorithms **35**, 351–372 (2004)

[2] Actually, we may take $\mathbf{V} = \mathbf{W}$ or $\mathbf{V} = \mathbf{T}$ (or $\mathbf{V} = -\mathbf{W}$ or $\mathbf{V} = -\mathbf{T}$) regarding to $\mathbf{W}$ or $\mathbf{T}$ being positive (or negative) semidefinite.

2. Axelsson, O., Kucherov, A.: Real valued iterative methods for solving complex symmetric linear systems. Numer. Linear Algebra Appl. **7**, 197–218 (2000)
3. Badalassi, V.E., Ceniceros, H.D., Banerjee, S.: Computation of multiphase systems with phase field models. J. Comput. Phys. **190**, 371–397 (2003)
4. Bai, Z.-Z.: Construction and analysis of structured preconditioners for block two-by-two matrices. J. Shanghai Univ. (English Edition) **8**, 397–405 (2004)
5. Bai, Z.-Z.: Structured preconditioners for nonsingular matrices of block two-by-two structures. Math. Comput. **75**, 791–815 (2006)
6. Bai, Z.-Z.: Block preconditioners for elliptic PDE-constrained optimization problems. Computing **91**, 379–395 (2011)
7. Bai, Z.-Z., Benzi, M., Chen, F.: Modified HSS iteration methods for a class of complex symmetric linear systems. Computing **87**, 93–111 (2010)
8. Bai, Z.-Z., Benzi, M., Chen, F.: On preconditioned MHSS iteration methods for complex symmetric linear systems. Numer. Algorithms **56**, 297–317 (2011)
9. Bai, Z.-Z., Benzi, M., Chen, F., Wang, Z.-Q.: Preconditioned MHSS iteration methods for a class of block two-by-two linear systems with applications to distributed control problems. IMA J. Numer. Anal. **33**, 343–369 (2013)
10. Bai, Z.-Z., Golub, G.H., Li, C.-K.: Convergence properties of preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite matrices. Math. Comput. **76**, 287–298 (2007)
11. Bai, Z.-Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. SIAM J. Matrix Anal. Appl. **24**, 603–626 (2003)
12. Bai, Z.-Z., Golub, G.H., Ng, M.K.: On inexact Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. Linear Algebra Appl. **428**, 413–440 (2008)
13. Bai, Z.-Z., Golub, G.H., Pan, J.-Y.: Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems. Numer. Math. **98**, 1–32 (2004)
14. Bai, Z.-Z., Qiu, S.-X.: Splitting-MINRES methods for linear systems with the coefficient matrix with a dominant indefinite symmetric part. Math. Numer. Sinica **24**, 113–128 (2002). (In Chinese)
15. Benzi, M., Bertaccini, D.: Block preconditioning of real-valued iterative algorithms for complex linear systems. IMA J. Numer. Anal. **28**, 598–618 (2008)
16. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems. Acta Numerica **14**, 1–137 (2005)
17. Berti, A., Bochicchio, I.: A mathematical model for phase separation: A generalized Cahn-Hilliard equation. Math. Meth. Appl. Sci. **34**, 1193–1201 (2011)
18. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. J. Chem. Phys. **28**, 258–267 (1958)
19. Chen, K.: Matrix Preconditioning Techniques and Applications. Cambridge University Press, Cambridge (2005)
20. Day, D.D., Heroux, M.A.: Solving complex-valued linear systems via equivalent real formulations. SIAM J. Sci. Comput. **23**, 480–498 (2001)
21. Guo, X.-X., Wang, S.: Modified HSS iteration methods for a class of non-Hermitian positive-definite linear systems. Appl. Math. Comput. **218**, 10122–10128 (2012)
22. Kailath, T.: Linear Systems. Prentice-Hall, Englewood Cliffs (1980)
23. Lass, O., Vallejos, M., Borzì, A., Douglas, C.C.: Implementation and analysis of multigrid schemes with finite elements for elliptic optimal control problems. Computing **84**, 27–48 (2009)
24. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer-Verlag, Berlin (1968)
25. Mehrmann, V.L.: The Autonomous Linear Quadratic Control Problem. Springer-Verlag, Berlin (1991)
26. Paige, C.C., Saunders, M.A.: Solution of sparse indefinite systems of linear equations. SIAM J. Numer. Anal. **12**, 617–629 (1975)
27. Pearson, J.W., Wathen, A.J.: A new approximation of the Schur complement in preconditioners for PDE-constraint optimization. Numer. Linear Algebra Appl. **19**, 816–829 (2012)
28. Petersen, I.R.: Disturbance attenuation and H^{∞}-optimization: A design method based on the algebraic Riccati equation. IEEE Trans. Automat. Control **32**, 427–429 (1987)
29. Rees, T., Dollar, H.S., Wathen, A.J.: Optimal solvers for PDE-constrained optimization. SIAM J. Sci. Comput. **32**, 271–298 (2010)

30. Ren, Z.-R.: On Preconditioned Iterative Methods for Sinc Systems of Linear Third-Order Ordinary Differential Equations, Ph.D. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing (2011)
31. Saad, Y., Schultz, M.H.: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. **7**, 856–869 (1986)
32. van der Schaft, A.: L_2-Gain and Passivity Techniques in Nonlinear Control, 2nd edn. Springer-Verlag, London (2000)
33. Wathen, A.J., Rees, T.: Chebyshev semi-iteration in preconditioning for problems including the mass matrix. Electron. Trans. Numer. Anal. **34**, 125–135 (2009)
34. Zhou, K.-M., Khargonekar, P.P.: An algebraic Riccati equation approach to H^∞ optimization. Syst. Control Lett. **11**, 85–91 (1988)
35. Zulehner, W.: Nonstandard norms and robust estimates for saddle point problems. SIAM J. Matrix Anal. Appl. **32**, 536–560 (2011)