PRECONDITIONED GMRES WITH APPLICATIONS

YICONG MA, KAIBO HU, XIAOZHE HU, AND JINCHAO XU

ABSTRACT. In this paper, we analyze the preconditioned GMRES method with block triangular preconditioners. We improve the existing analysis of preconditioned GMRES method in literature in the sense of removing the scaling parameters in front of the diagonal blocks. In this paper, we first investigate the convergence theories for an abstract saddle-point problem. Then we apply this technique to two linear systems, which come from multi-physics systems after linearization and finite element discretization.

1. INTRODUCTION

Preconditioned GMRES method is one of the most widely used Krylov solvers for linear systems. No requirement on symmetric or positive definite property makes it important to solving linear systems, especially to those come from multi-physics problems after linearization and discretization [4, 3, 2, 1, 13, 14, 5, 11]. It is well-known that in terms of number of iterations, block triangular preconditioners for GMRES method always give satisfactory performance. However, not much work exists to justify this theoretically. Loghin and Wathen [8] introduced an important framework to analyze the performance of such preconditioners, which is know as Field-of-values- (FOV-) analysis.

Our analysis is motivated by that of [10]. Carrying out the analysis from the perspective of functional analysis and PDEs, we are able to improve the estimates of Loghin and Wathen in [8] in the sense that we can remove the scaling parameters in front of the diagonal blocks. By choosing appropriate norms in the analysis, we are able to get rid of these scaling parameters, which is consistent with the practical implementation and observations.

The rest of this paper is organized as follows. In §2, we consider a generic saddle-point problem and carry out convergence analysis for the preconditioned GMRES method. Then we apply the analysis technique to two linear systems, which come from the linearization and finite element discretization of multi-physics systems in §3.

2. CONVERGENCE ANALYSIS FOR THE PRECONDITIONED GMRES METHOD

In this section, we recall the abstract framework for designing the *FOV-equivalent preconditioners*, following [8]. We design block triangular preconditioners for the GMRES method, and justify their performance theoretically.

Consider a model problem $\mathcal{A}x = \boldsymbol{F}$, where $\mathcal{A}$ is a general operator. We use another general operator $\mathcal{M}_{\mathcal{L}} : \boldsymbol{H}^* \to \boldsymbol{H}$ to denote the preconditioner. Based on the inner product $(\cdot,\cdot)_{\mathcal{M}^{-1}}$ and the norm $\|\cdot\|_{\mathcal{M}^{-1}}$, we can estimate the convergence rate of the preconditioned GMRES. It is proved [6, 12] that if x^m is the m-iteration of GMRES method and x

Key words and phrases. preconditioned GMRES, field-of-values analysis.

is the exact solution, then

$$\frac{\|\mathcal{M}_{\mathcal{L}}\mathcal{A}(x-x^m)\|_{\mathcal{M}^{-1}}}{\|\mathcal{M}_{\mathcal{L}}\mathcal{A}(x-x^0)\|_{\mathcal{M}^{-1}}} \le \left(1-\frac{\gamma^2}{\Gamma^2}\right)^{m/2},$$

where

$$\gamma \le \frac{(x,\mathcal{M}_{\mathcal{L}}\mathcal{A}x)_{\mathcal{M}^{-1}}}{(x,x)_{\mathcal{M}^{-1}}}, \quad \frac{\|\mathcal{M}_{\mathcal{L}}\mathcal{A}x\|_{\mathcal{M}^{-1}}}{\|x\|_{\mathcal{M}^{-1}}} \le \Gamma. \tag{2.1}$$

According to the theory, we conclude that as long as we find an operator $\mathcal{M}_{\mathcal{L}}$ and a proper inner product $(\cdot,\cdot)_{\mathcal{M}^{-1}}$ such that condition (2.1) is satisfied with constants γ and Γ independent of the physical and discretize parameters, $\mathcal{M}_{\mathcal{L}}$ is a uniform preconditioner for the GMRES method. Such preconditioners are usually referred to as FOV-equivalent preconditioners.

We carry out the convergence analysis of the preconditioned GMRES method for a saddle-point problem. Application to other finite element discretization is discussed later. Assume that $\mathcal{A}\boldsymbol{x} = \boldsymbol{F}$ is in the form

$$\begin{pmatrix} \mathcal{A}_1 & -\mathcal{B}^* \\ \mathcal{B} & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_1 \\ \boldsymbol{F}_2 \end{pmatrix}, \tag{2.2}$$

where $\mathcal{A}_1$ is a SPD operator. Based on the partition of the system, we assume that a splitting of the Hilbert space $\boldsymbol{H}$ is $\boldsymbol{H}_1 \times \boldsymbol{H}_2$ such that $\boldsymbol{x}_1 \in \boldsymbol{H}_1$, $\boldsymbol{x}_2 \in \boldsymbol{H}_2$. Assume that the problem (2.2) is well-posed with respect to norm $\|\cdot\|_{\mathcal{M}^{-1}}$, which is induced by $\mathcal{M} = \operatorname{diag}(\mathcal{H}_1,\mathcal{H}_2)^{-1}$. And we further assume that $\mathcal{H}_1 = \mathcal{A}_1$. Therefore, the well-posedness implies that there exists a constant $\zeta > 0$, independent of physical and discretization parameters (depending on the problem) such that

$$\inf_{\boldsymbol{x}_2\in\boldsymbol{H}_2}\sup_{\boldsymbol{x}_1\in\boldsymbol{H}_1}\frac{(\mathcal{B}\boldsymbol{x}_1,\boldsymbol{x}_2)}{\|\boldsymbol{x}_1\|_{\mathcal{A}_1}\|\boldsymbol{x}_2\|_{\mathcal{H}_2}} \ge \zeta > 0. \tag{2.3}$$

Theorem 2.1. *If the condition (2.3) holds, there exist constants γ and Γ such that for all $\boldsymbol{x} \neq \boldsymbol{0}$, the operator $\mathcal{A}$ defined in (2.2) and the operator*

$$\mathcal{M}_{\mathcal{L}} = \begin{pmatrix} \mathcal{A}_1 & 0 \\ \mathcal{B} & \mathcal{H}_2 \end{pmatrix}^{-1}$$

satisfy condition (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \operatorname{diag}(\mathcal{A}_1,\mathcal{H}_2)^{-1}$.

Proof. By simple computation, we get

$$\mathcal{M}_{\mathcal{L}}\mathcal{A} = \begin{pmatrix} \mathcal{I}_1 & -\mathcal{A}_1^{-1}\mathcal{B}^* \\ 0 & \mathcal{H}_2^{-1}\mathcal{B}\mathcal{A}_1^{-1}\mathcal{B}^* \end{pmatrix}.$$

Then for any $\boldsymbol{x} = (\boldsymbol{x}_1,\boldsymbol{x}_2)^T$, we have

$$\begin{aligned}
(\boldsymbol{x},\mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} &= (\boldsymbol{x}_1 - \mathcal{A}_1^{-1}\mathcal{B}^*\boldsymbol{x}_2,\boldsymbol{x}_1)_{\mathcal{A}_1} + (\mathcal{B}\mathcal{A}_1^{-1}\mathcal{B}^*\boldsymbol{x}_2,\boldsymbol{x}_2) \\
&= \|\boldsymbol{x}_1\|_{\mathcal{A}_1}^2 - (\mathcal{B}^*\boldsymbol{x}_2,\boldsymbol{x}_1) + \|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{A}_1^{-1}}^2 \\
&\ge \|\boldsymbol{x}_1\|_{\mathcal{A}_1}^2 - \|\boldsymbol{x}_1\|_{\mathcal{A}_1}\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{A}_1^{-1}} + \|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{A}_1^{-1}}^2 \\
&= \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^T \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},
\end{aligned}$$

where $\xi_1 = \|\boldsymbol{x}_1\|_{\mathcal{A}_1}$, $\xi_2 = \|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{A}_1^{-1}}$. Since the matrix in the middle is SPD, there exists $\gamma_0 > 0$ such that

$$(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} \geq \gamma_0 \left(\|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} + \|\mathcal{B}^*\boldsymbol{x}_2\|^2_{\mathcal{A}_1^{-1}} \right).$$

Moreover,

$$\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{A}_1^{-1}} = \sup_{\boldsymbol{x}_1 \in \mathcal{H}_1} \frac{(\mathcal{B}\boldsymbol{x}_1, \boldsymbol{x}_2)}{\|\boldsymbol{x}_1\|_{\mathcal{A}_1}} \geq \zeta \|\boldsymbol{x}_2\|_{\mathcal{H}_2},$$

we get

$$(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} \geq \gamma_0 \|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} + \gamma_0 \zeta^2 \|\boldsymbol{x}_2\|^2_{\mathcal{H}_2} \geq \min\left\{\gamma_0, \gamma_0\zeta^2\right\} (\boldsymbol{x}, \boldsymbol{x})_{\mathcal{M}^{-1}},$$

which leads to the lower bound γ. The upper bound Γ follows directly from the boundedness of each term. ☐

Applying $\mathcal{M}_{\mathcal{L}}$ defined in Theorem 2.1 as preconditioner means inversing each diagonal block exactly. In implementation, we need to call direct solvers for each diagonal block, which can be expensive and time-consuming. Therefore, we replace the diagonal blocks by their spectral equivalent SPD approximations. The following theorem states that under certain assumptions, such a preconditioner is still robust.

Theorem 2.2. *If the condition (2.3) holds, there exist constants γ and Γ such that for all $\boldsymbol{x} \neq \mathbf{0}$, the operator $\mathcal{A}$ defined in (2.2) and the operator*

$$\widehat{\mathcal{M}}_{\mathcal{L}} = \begin{pmatrix} \mathcal{Q}_1^{-1} & 0 \\ \mathcal{B} & \mathcal{Q}_2^{-1} \end{pmatrix}^{-1}$$

satisfy condition (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \operatorname{diag}(\mathcal{Q}_1, \mathcal{Q}_2)$ provided that

(1) $c_{2,i}(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x}) \leq \left(\mathcal{H}_i^{-1}\boldsymbol{x}, \boldsymbol{x}\right) \leq c_{1,i}(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x})$, $i = 1$ or 2,

(2) $\|\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1\|_{\mathcal{A}_1} \leq \rho$, with $0 \leq \rho < 1$.

Proof. By simple computation, we get

$$\widehat{\mathcal{M}}_{\mathcal{L}}\mathcal{A} = \begin{pmatrix} \mathcal{Q}_1\mathcal{A}_1 & -\mathcal{Q}_1\mathcal{B}^* \\ \mathcal{Q}_2\mathcal{B}(\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1) & \mathcal{Q}_2\mathcal{B}\mathcal{Q}_1\mathcal{B}^* \end{pmatrix}.$$

Then for any $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)^T$, we have

$$\begin{aligned}(\boldsymbol{x}, \widehat{\mathcal{M}}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} &= \|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} - (\mathcal{B}^*\boldsymbol{x}_2, \boldsymbol{x}_1) + (\mathcal{B}(\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1)\boldsymbol{x}_1, \boldsymbol{x}_2) + \|\mathcal{B}^*\boldsymbol{x}_2\|^2_{\mathcal{Q}_1} \\ &= \|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} - (\mathcal{Q}_1\mathcal{A}_1\boldsymbol{x}_1, \mathcal{B}^*\boldsymbol{x}_2) + \|\mathcal{B}^*\boldsymbol{x}_2\|^2_{\mathcal{Q}_1}.\end{aligned}$$

As $\|\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1\|_{\mathcal{A}_1} \leq \rho$ implies that

$$\begin{aligned}(1-\rho)(\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{A}_1^{-1}} &\leq (\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{Q}_1} \leq (1+\rho)(\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{A}_1^{-1}}, \\ (1+\rho)^{-1}(\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{A}_1} &\leq (\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{Q}_1^{-1}} \leq (1-\rho)^{-1}(\boldsymbol{x}_1, \boldsymbol{x}_1)_{\mathcal{A}_1},\end{aligned}$$

we have

$$\begin{aligned}-(\mathcal{Q}_1\mathcal{A}_1\boldsymbol{x}_1, \mathcal{B}^*\boldsymbol{x}_2) &\leq \|\mathcal{A}_1\boldsymbol{x}_1\|_{\mathcal{Q}_1}\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{Q}_1} \leq (1+\rho)\|\mathcal{A}_1\boldsymbol{x}_1\|_{\mathcal{A}_1^{-1}}\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{Q}_1} \\ &= (1+\rho)\|\boldsymbol{x}_1\|_{\mathcal{A}_1}\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{Q}_1},\end{aligned}$$

Therefore,

$$(\boldsymbol{x}, \widehat{\mathcal{M}}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} \geq \|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} - (1+\rho)\|\boldsymbol{x}_1\|_{\mathcal{A}_1}\|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{Q}_1} + \|\mathcal{B}^*\boldsymbol{x}_2\|^2_{\mathcal{Q}_1}$$
$$= \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^T \begin{pmatrix} 1 & -(1+\rho)/2 \\ -(1+\rho)/2 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},$$

where $\xi_1 = \|\boldsymbol{x}_1\|_{\mathcal{A}_1}$, $\xi_2 = \|\mathcal{B}^*\boldsymbol{x}_2\|_{\mathcal{Q}_1}$. We can verify that the matrix in the middle is SPD when $0 \leq \rho < 1$. Therefore, there exists a constant $\gamma_0 > 0$ such that

$$(\boldsymbol{x}, \widehat{\mathcal{M}}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} \geq \gamma_0 \left(\|\boldsymbol{x}_1\|^2_{\mathcal{A}_1} + \|\mathcal{B}^*\boldsymbol{x}_2\|^2_{\mathcal{Q}_1}\right) \geq \gamma_0(1-\rho)\|\boldsymbol{x}_1\|^2_{\mathcal{Q}_1^{-1}} + \gamma_0(1-\rho)\zeta^2\|\boldsymbol{x}_2\|^2_{\mathcal{H}_2}$$
$$\geq \min\left\{\gamma_0(1-\rho), \gamma_0(1-\rho)\zeta^2 c_{1,2}^{-1}\right\}(\boldsymbol{x}, \boldsymbol{x})_{\mathcal{M}^{-1}},$$

which leads to the lower bound γ. The upper bound Γ follows directly from the fact that each term is bounded. □

To implement the preconditioner $\widehat{\mathcal{M}}_{\mathcal{L}}$, we can use iterative solvers for each diagonal block with a relative big tolerance. We further comment that the second assumption in Theorem 2.2 is reasonable as in practice we can achieve it by performing one or several steps of V-cycle multigrid method. We refer readers to [9] for detailed implementation.

3. Applications

In this section, we apply the analysis technique discussed in the previous section to some other problems.

3.1. Application to a penalty formulation of the MHD system.

In this section, we consider a penalty formulation of a MHD system. After Picard linearization and finite element discretization, the discrete problem we consider in this section is: Find $(\boldsymbol{u}_h, p_h, \boldsymbol{B}_h) \in H^1_{0,h}(\Omega)^3 \times L^2_{0,h}(\Omega) \times H^1_{n,h}(\Omega)^3$ such that for any $(\boldsymbol{v}_h, q_h, \boldsymbol{C}_h) \in H^1_{0,h}(\Omega)^3 \times L^2_{0,h}(\Omega) \times H^1_{n,h}(\Omega)^3$,

$$\left\{ \begin{aligned} & k^{-1}(\boldsymbol{u}_h, \boldsymbol{v}_h) + Re^{-1}(\nabla\boldsymbol{u}_h, \nabla\boldsymbol{v}_h) + k^{-1}(\nabla\cdot\boldsymbol{u}_h, \nabla\cdot\boldsymbol{v}_h) - (p_h, \nabla\cdot\boldsymbol{v}_h) \\ & \quad + s(\nabla\times\boldsymbol{B}_h, \boldsymbol{v}_h\times\boldsymbol{b}) = \langle \boldsymbol{f}, \boldsymbol{v}_h\rangle + k^{-1}(\boldsymbol{a}, \boldsymbol{v}_h) - (\boldsymbol{a}\cdot\nabla\boldsymbol{a}, \boldsymbol{v}_h), \\ & (\nabla\cdot\boldsymbol{u}_h, q_h) = 0, \\ & sk^{-1}(\boldsymbol{B}_h, \boldsymbol{C}_h) + \alpha(\nabla\boldsymbol{B}_h, \nabla\boldsymbol{C}_h) - s(\boldsymbol{u}_h\times\boldsymbol{b}, \nabla\times\boldsymbol{C}_h) = 0, \end{aligned} \right. \tag{3.1}$$

where k is the time step size, $\alpha = s/Rm$, $\boldsymbol{a}$ and $\boldsymbol{b}$ are the numerical solutions from the last time step. We can prove that the problem (3.1) is well-posed. That is, it satisfies the boundedness property and the inf-sup conditions under the following weighted norms

$$\|\boldsymbol{v}\|^2_{\mathcal{H}_1} = k^{-1}\|\boldsymbol{v}\|^2 + Re^{-1}\|\nabla\boldsymbol{v}\|^2 + k^{-1}\|\nabla\cdot\boldsymbol{v}\|^2,$$
$$\|q\|^2_{\mathcal{H}_2} = k\|q\|^2,$$
$$\|\boldsymbol{C}\|^2_{\mathcal{H}_3} = sk^{-1}\|\boldsymbol{C}\|^2 + \alpha\|\nabla\boldsymbol{C}\|^2,$$

when the time step size is small enough, i.e. $k \leq k_0$, where

$$k_0 = \frac{1}{8sRm}\|\boldsymbol{b}\|^{-2}_{0,\infty}, \tag{3.2}$$

Here, $\|\cdot\|_{0,\infty}$ is the L^∞ norm, which is defined by $\|v\|_{0,\infty}= \operatorname{ess\,sup}_{x\in\Omega}|v(x)|$. Moreover, $\mathcal{H}_i$ ($i = 1,2,3$) is a symmetric positive definite operator (SPD) such that $\|\boldsymbol{x}\|^2_{\mathcal{H}_i}= (\mathcal{H}_i\boldsymbol{x},\boldsymbol{x})$. The operator form of (3.1) is

$$\mathcal{A}\boldsymbol{x} = \boldsymbol{F} \Rightarrow \begin{pmatrix} \mathcal{A}_1 & -\mathrm{div}^* & -\mathcal{Z}^* \\ \mathrm{div} & 0 & 0 \\ \mathcal{Z} & 0 & \mathcal{H}_3 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ p \\ \boldsymbol{B} \end{pmatrix} = \begin{pmatrix} \boldsymbol{h}_1 \\ g \\ \boldsymbol{h}_2 \end{pmatrix}, \tag{3.3}$$

where

$$\mathcal{A}_1\boldsymbol{u} = k^{-1}\boldsymbol{u} - Re^{-1}\Delta\boldsymbol{u} + k^{-1}\mathrm{div}^*\mathrm{div}\boldsymbol{u},\ \forall\boldsymbol{u}\in H^1_{0,h}(\Omega),$$
$$\mathcal{Z}\boldsymbol{u} = s\nabla\times(\boldsymbol{b}\times\boldsymbol{u}),\ \forall\boldsymbol{u}\in H^1_{0,h}(\Omega).$$

And $\mathcal{A}_1 = \mathcal{H}_1$. The following theorems analyze the block triangular preconditioners for the $\mathcal{A}$ defined in (3.3).

Theorem 3.1. *If $k \le k_0$, which is defined by (3.2), there exists γ and Γ that are independent of the mesh size h, time step size k, and physical parameters Rm and s, such that for all $\boldsymbol{x}\neq\boldsymbol{0}$, the operator $\mathcal{A}$ defined in (3.3) and the operator*

$$\mathcal{M}_{\mathcal{L}} = \begin{pmatrix} \mathcal{A}_1 & 0 & 0 \\ \mathrm{div} & k\mathcal{I}_2 & 0 \\ \mathcal{Z} & 0 & \mathcal{H}_3 \end{pmatrix}^{-1} \tag{3.4}$$

satisfy the condition (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \mathrm{diag}(\mathcal{A}_1,\mathcal{H}_2,\mathcal{H}_3)$.

Proof. By simple computation, we get

$$\mathcal{M}_{\mathcal{L}}\mathcal{A} = \begin{pmatrix} \mathcal{I}_1 & -\mathcal{A}_1^{-1}\mathrm{div}^* & -\mathcal{A}_1^{-1}\mathcal{Z}^* \\ 0 & k^{-1}\mathrm{div}\mathcal{A}_1^{-1}\mathrm{div}^* & k^{-1}\mathrm{div}\mathcal{A}_1^{-1}\mathcal{Z}^* \\ 0 & \mathcal{H}_3^{-1}\mathcal{Z}\mathcal{A}_1^{-1}\mathrm{div}^* & \mathcal{I}_3 + \mathcal{H}_3^{-1}\mathcal{Z}\mathcal{A}_1^{-1}\mathcal{Z}^* \end{pmatrix}.$$

Noticing that

$$\|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{A}_1^{-1}}= \frac{(\mathcal{Z}^*\boldsymbol{B},\boldsymbol{u})}{\|\boldsymbol{u}\|_{\mathcal{A}_1}} \le \sqrt{sRm}\|\boldsymbol{B}\|_{\mathcal{H}_3}\frac{\|\boldsymbol{u}\times\boldsymbol{b}\|}{\|\boldsymbol{u}\|_{\mathcal{A}_1}} \le \frac{1}{2\sqrt{2}}\|\boldsymbol{B}\|_{\mathcal{H}_3},$$

for any $\boldsymbol{x} = (\boldsymbol{u},p,\boldsymbol{B})^T$, we have

$$\begin{aligned}
(\boldsymbol{x},\mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{H}} &= \|\boldsymbol{u}\|^2_{\mathcal{A}_1} - (\mathrm{div}^*p,\boldsymbol{u}) - (\mathcal{Z}^*\boldsymbol{B},\boldsymbol{u}) + \|\mathrm{div}^*p\|^2_{\mathcal{A}_1^{-1}} + 2(\mathcal{A}_1^{-1}\mathrm{div}^*p,\mathcal{Z}^*\boldsymbol{B}) \\
&\quad + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} + \|\mathcal{Z}^*\boldsymbol{B}\|^2_{\mathcal{A}_1^{-1}} \\
&\ge \|\boldsymbol{u}\|^2_{\mathcal{A}_1} - \|\mathrm{div}^*p\|_{\mathcal{A}_1^{-1}}\|\boldsymbol{u}\|_{\mathcal{A}_1} - \frac{1}{2\sqrt{2}}\|\boldsymbol{B}\|_{\mathcal{H}_3}\|\boldsymbol{u}\|_{\mathcal{A}_1} + \|\mathrm{div}^*p\|^2_{\mathcal{A}_1^{-1}} \\
&\quad - \frac{1}{\sqrt{2}}\|\mathrm{div}^*p\|_{\mathcal{A}_1^{-1}}\|\boldsymbol{B}\|_{\mathcal{H}_3} + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} \\
&= \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \begin{pmatrix} 1 & -1/2 & -1/4\sqrt{2} \\ -1/2 & 1 & -1/2\sqrt{2} \\ -1/4\sqrt{2} & -1/2\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix},
\end{aligned}$$

where $\xi_1 = \|\boldsymbol{u}\|_{\mathcal{A}_1}$, $\xi_2 = \|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}}$, $\xi_3 = \|\boldsymbol{B}\|_{\mathcal{H}_3}$. Since the matrix in the middle is SPD, there exists $\gamma_0 > 0$ such that

$$(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{H}} \geq \gamma_0 \left(\|\boldsymbol{u}\|^2_{\mathcal{A}_1} + \|\mathrm{div}^* p\|^2_{\mathcal{A}_1^{-1}} + \|\boldsymbol{B}\|^2_{\mathcal{H}_3}\right) \geq \min\left\{\gamma_0, \gamma_0\zeta^2\right\}(\boldsymbol{x}, \boldsymbol{x})_{\mathcal{H}},$$

which leads to the lower bound γ. The upper bound Γ follows directly from the fact that each term is bounded. □

To reduce the computation cost of $\mathcal{M}_{\mathcal{L}}$, we replace its diagonal blocks by their spectral equivalent SPD approximations.

Theorem 3.2. *If $k \leq k_0$, which is defined by (3.2), there exists contants γ and Γ, which are independent of the mesh size h, time step size k, and physical parameters Rm and s, such that for all $\boldsymbol{x} \neq \boldsymbol{0}$, the operator $\mathcal{A}$ defined in (3.3) and the operator*

$$\widehat{\mathcal{M}}_{\mathcal{L}} = \begin{pmatrix} \mathcal{Q}_1^{-1} & 0 & 0 \\ \mathrm{div} & \mathcal{Q}_2^{-1} & 0 \\ \mathcal{Z} & 0 & \mathcal{Q}_3^{-1} \end{pmatrix}. \tag{3.5}$$

satisfy (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \mathrm{diag}\,(\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3)$ provided that

(1) $c_{2,i}\,(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x}) \leq \left(\mathcal{H}_i^{-1}\boldsymbol{x}, \boldsymbol{x}\right) \leq c_{1,i}\,(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x})$, $i = 1, 2, 3$,
(2) $\|\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1\|_{\mathcal{A}_1} \leq \rho$, with $0 \leq \rho < 0.252$.

Proof. By simple computation, we get

$$\widehat{\mathcal{M}}_{\mathcal{L}}\mathcal{A} = \begin{pmatrix} \mathcal{Q}_1\mathcal{A}_1 & -\mathcal{Q}_1\mathrm{div}^* & -\mathcal{Q}_1\mathcal{Z}^* \\ \mathcal{Q}_2\mathrm{div}(\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1) & \mathcal{Q}_2\mathrm{div}\,\mathcal{Q}_1\mathrm{div}^* & \mathcal{Q}_2\mathrm{div}\,\mathcal{Q}_1\mathcal{Z}^* \\ \mathcal{Q}_3\mathcal{Z}(\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1) & \mathcal{Q}_3\mathcal{Z}\mathcal{Q}_1\mathrm{div}^* & \mathcal{Q}_3\mathcal{H}_3 + \mathcal{Q}_3\mathcal{Z}\mathcal{Q}_1\mathcal{Z}^* \end{pmatrix}.$$

Therefore, for any $\boldsymbol{x} = (\boldsymbol{u}, p, \boldsymbol{B})^T$,

$$\begin{aligned}(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} &= \|\boldsymbol{u}\|^2_{\mathcal{A}_1} - (\mathrm{div}^* p, \boldsymbol{u}) - (\mathcal{Z}^*\boldsymbol{B}, \boldsymbol{u}) + (\mathrm{div}(\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1)\boldsymbol{u}, p) + \|\mathrm{div}^* p\|^2_{\mathcal{Q}_1} \\ &\quad + 2(\mathcal{Z}^*\boldsymbol{B}, \mathrm{div}^* p)_{\mathcal{Q}_1} + ((\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1)\boldsymbol{u}, \mathcal{Z}^*\boldsymbol{B}) + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} + \|\mathcal{Z}^*\boldsymbol{B}\|^2_{\mathcal{Q}_1} \\ &\geq \|\boldsymbol{u}\|^2_{\mathcal{A}_1} + \|\mathrm{div}^* p\|^2_{\mathcal{Q}_1} + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} + \|\mathcal{Z}^*\boldsymbol{B}\|^2_{\mathcal{Q}_1} - (\mathcal{Q}_1\mathcal{A}_1\boldsymbol{u}, \mathrm{div}^* p) \\ &\quad + 2(\mathcal{Z}^*\boldsymbol{B}, \mathrm{div}^* p)_{\mathcal{Q}_1} - (\mathcal{Q}_1\mathcal{A}_1\boldsymbol{u}, \mathcal{Z}^*\boldsymbol{B})\end{aligned}$$

Since $\|\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1\|_{\mathcal{A}_1} \leq \rho$ implies

$$\begin{aligned}(1-\rho)(\boldsymbol{v}, \boldsymbol{v})_{\mathcal{A}_1^{-1}} &\leq (\boldsymbol{v}, \boldsymbol{v})_{\mathcal{Q}_1} \leq (1+\rho)(\boldsymbol{v}, \boldsymbol{v})_{\mathcal{A}_1^{-1}}, \\ (1+\rho)^{-1}(\boldsymbol{v}, \boldsymbol{v})_{\mathcal{A}_1} &\leq (\boldsymbol{v}, \boldsymbol{v})_{\mathcal{Q}_1^{-1}} \leq (1-\rho)^{-1}(\boldsymbol{v}, \boldsymbol{v})_{\mathcal{A}_1},\end{aligned}$$

and $\|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{A}_1^{-1}} \leq \dfrac{1}{2\sqrt{2}}\|\boldsymbol{B}\|_{\mathcal{H}_3}$, we have

$$\begin{aligned}|(\mathcal{Q}_1\mathcal{A}_1\boldsymbol{u}, \mathrm{div}^* p)| &\leq \|\mathcal{A}_1\boldsymbol{u}\|_{\mathcal{Q}_1}\|\mathrm{div}^* p\|_{\mathcal{Q}_1} \leq (1+\rho)\|\boldsymbol{u}\|_{\mathcal{A}_1}\|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}}, \\ |(\mathcal{Q}_1\mathcal{A}_1\boldsymbol{u}, \mathcal{Z}^*\boldsymbol{B})| &\leq \|\mathcal{A}_1\boldsymbol{u}\|_{\mathcal{Q}_1}\|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{Q}_1} \leq (1+\rho)\|\boldsymbol{u}\|_{\mathcal{A}_1}\|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{A}_1^{-1}} \leq \frac{1+\rho}{2\sqrt{2}}\|\boldsymbol{u}\|_{\mathcal{A}_1}\|\boldsymbol{B}\|_{\mathcal{H}_3}, \\ |(\mathcal{Z}^*\boldsymbol{B}, \mathrm{div}^* p)_{\mathcal{Q}_1}| &\leq \|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{Q}_1}\|\mathrm{div}^* p\|_{\mathcal{Q}_1} \leq (1+\rho)\|\mathcal{Z}^*\boldsymbol{B}\|_{\mathcal{A}_1^{-1}}\|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}} \\ &\leq \frac{1+\rho}{2\sqrt{2}}\|\boldsymbol{B}\|_{\mathcal{H}_3}\|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}}.\end{aligned}$$

Hence,

$$
\begin{aligned}
(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} \geq & \|\boldsymbol{u}\|^2_{\mathcal{A}_1} + (1-\rho)\|\mathrm{div}^* p\|^2_{\mathcal{A}_1^{-1}} + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} - (1+\rho)\|\boldsymbol{u}\|_{\mathcal{A}_1}\|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}} \\
& - \frac{1+\rho}{2\sqrt{2}}\|\boldsymbol{u}\|_{\mathcal{A}_1}\|\boldsymbol{B}\|_{\mathcal{H}_3} - \frac{1+\rho}{\sqrt{2}}\|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}}\|\boldsymbol{B}\|_{\mathcal{H}_3} \\
= & \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}^T \begin{pmatrix} 1 & -(1+\rho)/2 & -(1+\rho)/4\sqrt{2} \\ -(1+\rho)/2 & 1-\rho & -(1+\rho)/2\sqrt{2} \\ -(1+\rho)/4\sqrt{2} & (1+\rho)/2\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix},
\end{aligned}
$$

where $\xi_1 = \|\boldsymbol{u}\|_{\mathcal{A}_1}$, $\xi_2 = \|\mathrm{div}^* p\|_{\mathcal{A}_1^{-1}}$, and $\xi_3 = \|\boldsymbol{B}\|_{\mathcal{H}_3}$. It is easy to verify that the matrix in the middle is SPD when $0 \leq \rho < 0.252$. Therefore, there exists a constant $\gamma_0 > 0$ such that

$$
\begin{aligned}
(\boldsymbol{x}, \mathcal{M}_{\mathcal{L}}\mathcal{A}\boldsymbol{x})_{\mathcal{M}^{-1}} & \geq \gamma_0 \left(\|\boldsymbol{u}\|^2_{\mathcal{A}_1} + \|\mathrm{div}^* p\|^2_{\mathcal{A}_1^{-1}} + \|\boldsymbol{B}\|^2_{\mathcal{H}_3} \right) \\
& \geq \min\left\{ \gamma_0(1-\rho), \gamma_0\zeta^2 c_{1,2}^{-1}, \gamma_0 c_{1,3}^{-1} \right\} (\boldsymbol{x}, \boldsymbol{x})_{\mathcal{M}^{-1}},
\end{aligned}
$$

which leads to the lower bound γ. The upper bound Γ follows from the boundedness of each term.

□

3.2. **Application to an incompressible MHD system.** In this section, we consider a structure-preserving discretization [7] of an incompressible MHD system. After Picard linearization and finite element discretization, the problem we consider is: Find $(\boldsymbol{u}_h, \boldsymbol{B}_h, \boldsymbol{E}_h) \in H^1_{0,h}(\Omega)^3 \times H_{0,h}(\mathrm{div};\Omega) \times H_{0,h}(\mathrm{curl};\Omega)$ and $p_h \in L^2_{0,h}(\Omega)$ such that for any $(\boldsymbol{v}_h, \boldsymbol{C}_h, \boldsymbol{F}_h) \in H^1_{0,h}(\Omega)^3 \times H_{0,h}(\mathrm{div};\Omega) \times H_{0,h}(\mathrm{curl};\Omega)$ and $q \in L^2_{0,h}(\Omega)$,

$$
\begin{cases}
k^{-1}(\boldsymbol{u}_h, \boldsymbol{v}_h) + Re^{-1}(\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) + k^{-1}(\nabla\cdot\boldsymbol{u}_h, \nabla\cdot\boldsymbol{v}_h) - (p_h, \nabla\cdot\boldsymbol{v}_h) \\
\qquad -s((\boldsymbol{E}_h + \boldsymbol{u}_h\times\boldsymbol{b})\times\boldsymbol{b}, \boldsymbol{v}_h) = (\boldsymbol{f}_h, \boldsymbol{v}_h) + k^{-1}(\boldsymbol{a}, \boldsymbol{v}_h) - (\boldsymbol{a}\cdot\nabla\boldsymbol{a}, \boldsymbol{v}_h), \\
-k^{-1}\alpha\left(\mu_r^{-1}\boldsymbol{B}_h, \boldsymbol{C}_h\right) - \alpha\left(\mu_r^{-1}\nabla\times\boldsymbol{E}_h, \boldsymbol{C}_h\right) = -k^{-1}\alpha\left(\mu_r^{-1}\boldsymbol{b}, \boldsymbol{C}_h\right), \\
s(\boldsymbol{E}_h + \boldsymbol{u}_h\times\boldsymbol{b}, \boldsymbol{F}_h) - \alpha\left(\mu_r^{-1}\boldsymbol{B}_h, \nabla\times\boldsymbol{F}_h\right) = 0, \\
(\nabla\cdot\boldsymbol{u}_h, q_h) = 0.
\end{cases} \tag{3.6}
$$

We can prove [7] that this formulation (3.6) is well-posed. That is, it satisfies the boundedness property and the inf-sup conditions under the following weighted norms

$$
\begin{aligned}
& \|(\boldsymbol{v}, \boldsymbol{C}, \boldsymbol{F})\|^2_{\boldsymbol{X}} = \|\boldsymbol{v}\|^2_{\mathcal{H}_1} + \|\boldsymbol{C}\|^2_{\mathcal{H}_3} + \|\boldsymbol{F}\|^2_{\mathcal{H}_4}, \ \|q\|_Q = \|q\|_{\mathcal{H}_2}, \\
& \|(\boldsymbol{v}, q, \boldsymbol{C}, \boldsymbol{F})\|^2_{\mathcal{H}} = \|\boldsymbol{v}\|^2_{\mathcal{H}_1} + \|q\|^2_{\mathcal{H}_2} + \|\boldsymbol{C}\|^2_{\mathcal{H}_3} + \|\boldsymbol{F}\|^2_{\mathcal{H}_4},
\end{aligned} \tag{3.7}
$$

with

$$
\begin{aligned}
& \|\boldsymbol{v}\|^2_{\mathcal{H}_1} = k^{-1}\|\boldsymbol{v}\|^2 + Re^{-1}\|\nabla\boldsymbol{v}\|^2 + k^{-1}\|\nabla\cdot\boldsymbol{v}\|^2 + s\|\boldsymbol{v}\times\boldsymbol{B}^-\|^2_{\sigma_r}, \\
& \|q\|^2_{\mathcal{H}_2} = k\|q\|^2, \\
& \|\boldsymbol{C}\|^2_{\mathcal{H}_3} = k^{-1}\alpha\|\boldsymbol{C}\|^2_{\mu_r^{-1}} + \alpha\|\nabla\cdot\boldsymbol{C}\|^2_{\mu_r^{-1}}, \\
& \|\boldsymbol{F}\|^2_{\mathcal{H}_4} = s\|\boldsymbol{F}\|^2_{\sigma_r} + k\alpha\|\nabla\times\boldsymbol{F}\|^2_{\mu_r^{-1}},
\end{aligned}
$$

when the time step size is small enough, i.e. $k \le k_0$, where

$$k_0 = \frac{1}{8s}\|\sqrt{\sigma_r}\boldsymbol{B}^-\|_{0,\infty}^{-2}. \tag{3.8}$$

Moreover, $\mathcal{H}_i$ (i = 1, 2, 3 or 4) is a symmetric positive operator (SPD) such that $\|\boldsymbol{x}\|^2_{\mathcal{H}_i} = (\mathcal{H}_i\boldsymbol{x}, \boldsymbol{x})$. The operator form of (3.6) is

$$\mathcal{A}\boldsymbol{x} = \boldsymbol{F} \Longrightarrow \begin{pmatrix} \mathcal{A}_1 & -\mathrm{div}^* & 0 & \mathcal{F}^* \\ \mathrm{div} & 0 & 0 & 0 \\ 0 & 0 & \alpha k^{-1}\mu_r^{-1}\mathcal{I}_3 & \alpha\mu_r^{-1}\mathrm{curl} \\ \mathcal{F} & 0 & -\alpha\mathrm{curl}^*\mu_r^{-1} & s\sigma_r\mathcal{I}_4 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ p \\ \boldsymbol{B} \\ \boldsymbol{E} \end{pmatrix} = \begin{pmatrix} \boldsymbol{h}_1 \\ g \\ \boldsymbol{h}_2 \\ \boldsymbol{h}_3 \end{pmatrix}, \tag{3.9}$$

where

$$\mathcal{A}_1\boldsymbol{u} = k^{-1}\boldsymbol{u} - Re^{-1}\Delta\boldsymbol{u} + k^{-1}\mathrm{div}^*\mathrm{div}\boldsymbol{u} - s\sigma_r(\boldsymbol{u}\times\boldsymbol{B}^-)\times\boldsymbol{B}^-,\ \forall\boldsymbol{u}\in\boldsymbol{V}_h,$$
$$\mathcal{F}\boldsymbol{u} = s\sigma_r\boldsymbol{u}\times\boldsymbol{B}^-,\ \forall\boldsymbol{u}\in\boldsymbol{V}_h.$$

Notice that we have $\mathcal{H}_1 = \mathcal{A}_1$. The following theorems analyze the block triangular preconditioners for the operator $\mathcal{A}$ defined in (3.9). For the sake of briefness, we only list the results in this section. We refer the readers to [9] for detailed proof and discussion.

Theorem 3.3. *If $k \le k_0$, which is defined by (3.8), there exists γ and Γ that are independent of the mesh size h, time step size k, and physical parameters Rm, s, μ_r and σ_r, such that for all $\boldsymbol{x} \ne \boldsymbol{0}$, the operator $\mathcal{A}$ defined in (3.9) and the operator*

$$\mathcal{M}_{\mathcal{L}} = \begin{pmatrix} \mathcal{A}_1 & 0 & 0 & 0 \\ \mathrm{div} & k\mathcal{I}_2 & 0 & 0 \\ 0 & 0 & \alpha k^{-1}\mu_r^{-1}\mathcal{I}_3 & 0 \\ \mathcal{F} & 0 & -\alpha\mathrm{curl}^*\mu_r^{-1} & \mathcal{H}_4 \end{pmatrix}^{-1} \tag{3.10}$$

satisfy the condition (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \mathrm{diag}(\mathcal{A}_1, \mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4)$.

As mentioned before, we replace the diagonal blocks of $\mathcal{M}_{\mathcal{L}}$ by their spectral equivalent SPD approximations (except that of $\boldsymbol{B}$) to reduce the time and computation cost. The reason why we keep the diagonal block of $\boldsymbol{B}$ and the implementation issue are discussed in [9] in detail.

Theorem 3.4. *If $k \le k_0$, which is defined by (3.8), there exists contants γ and Γ, which are independent of the mesh size h, time step size k, and physical parameters Rm, s, μ_r and σ_r, such that for all $\boldsymbol{x} \ne \boldsymbol{0}$, the operator $\mathcal{A}$ defined in (3.9) and the operator*

$$\widehat{\mathcal{M}}_{\mathcal{L}} = \begin{pmatrix} \mathcal{Q}_1^{-1} & 0 & 0 & 0 \\ \mathrm{div} & \mathcal{Q}_2^{-1} & 0 & 0 \\ 0 & 0 & \alpha k^{-1}\mu_r^{-1}\mathcal{I}_3 & 0 \\ \mathcal{F} & 0 & -\alpha\mathrm{curl}^*\mu_r^{-1} & \mathcal{Q}_4^{-1} \end{pmatrix}^{-1} \tag{3.11}$$

satisfy (2.1) with the norm $\|\cdot\|_{\mathcal{M}^{-1}}$ induced by $\mathcal{M} = \mathrm{diag}\left(\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{H}_3^{-1}, \mathcal{Q}_4\right)$ provided that

(1) $c_{2,i}\,(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x}) \le \left(\mathcal{H}_i^{-1}\boldsymbol{x}, \boldsymbol{x}\right) \le c_{1,i}\,(\mathcal{Q}_i\boldsymbol{x}, \boldsymbol{x})$, $i = 1, 2, 3, 4$,
(2) $\|\mathcal{I}_1 - \mathcal{Q}_1\mathcal{A}_1\|_{\mathcal{A}_1} \le \rho$, *with* $0 \le \rho < 0.289$.

4. Conclusions

The major contribution of this analysis is that we improve the of FOV-analysis of block triangular preconditioners proposed in [8]. Their analysis requires scaling parameters in front of the diagonal blocks in $\widehat{\mathcal{M}}_{\mathcal{L}}$ under certain constrains, which are usually difficult to choose in practice. In our analysis, with the help of an appropriate norm $(\cdot,\cdot)_{\mathcal{M}^{-1}}$, we are able to remove those unnecessary scaling parameters, which makes the theoretical results consistent with practical implementation and observations.

References

[1] L. Chacón. An optimal, parallel, fully implicit Newton-Krylov solver for three-dimensional viscoresistive magnetohydrodynamics. *Physics of Plasmas*, 15(5), May 2008.

[2] L Chacón. Scalable parallel implicit solvers for 3D magnetohydrodynamics. *Journal of Physics: Conference Series*, 125, 2008.

[3] L. Chacón and D.A. Knoll. A 2D high-β Hall MHD implicit nonlinear solver. *Journal of Computational Physics*, 188:573–592, 2003.

[4] L. Chacón, D.A. Knoll, and J.M. Finn. An implicit , nonlinear reduced resistive MHD solver. *Journal of Computational Physics*, 178:15–36, 2002.

[5] E.C. Cyr, J.N. Shadid, R.S. Tuminaro, R.P. Pawlowski, and L. Chacón. A new approximate block factorization preconditioner for two-dimensional incompressible (reduced) resistive MHD. *SIAM Journal on Scientific Computing*, 35(3):701–730, 2013.

[6] H.C. Elman. *Iterative Methods for Large Sparse Non-symmetric Systems of Linear Equations*. PhD thesis, Yale Univeristy, 1982.

[7] K. Hu, Y. Ma, and J. Xu. Stable finite element methods preserving $\nabla \cdot \boldsymbol{B} = 0$ exactly for MHD models. Submitted to Numerische Mathematik, 2014.

[8] D. Loghin and A.J. Wathen. Analysis of preconditioners for saddle-point problems. *SIAM Journal on Scientific Computing*, 25(6):2029–2049, 2004.

[9] Y. Ma, K. Hu, X. Hu, and J. Xu. Robust preconditioners for the magnetohydrodynmaics system. Submitted to Journal of Computational Physics, 2015.

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

[11] E.G. Phillips, H.C. Elman, E.C. Cyr, J.N. Shadid, and R.P. Pawlowski. A block preconditioner for an exact penalty formulation for stationary MHD. *SIAM Journal on Scientific Computing*, 36(6):930–951, 2014.

[12] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS, Boston, 1996.

[13] J.N. Shadid, E.C. Cyr, R.P. Pawlowski, R.S. Tuminaro, L. Chacón, and P.T. Lin. Initial performance of fully-coupled AMG and approximate block factorization preconditioners for solution of implicit FE resistive MHD. In J.C.F. Pereira and A. Sequeira, editors, *V Europena Conference on Computational Fluid Dynamics*, pages 1–19, 2010.

[14] J.N. Shadid, R.P. Pawlowski, J.W. Banks, L. Chacon, P.T. Lin, and R.S. Tuminaro. Towards a scalable fully-implicit fully-coupled resistive MHD formulation with stabilized FE methods. *Journal of Computational Physics*, 229:7649–7671, 2010.

DEPARTMENT OF MATHEMATICS,THE PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16802, USA

E-mail address: yxm147@psu.edu

BEIJING INTERNATIONAL CENTER FOR MATHEMATICAL RESEARCH , PEKING UNIVERSITY, BEIJING 100871, P. R. CHINA

E-mail address: hukaibo02@gmail.com

DEPARTMENT OF MATHEMATICS, TUFTS UNIVERSITY, MEDFORD, MA 02155, USA

E-mail address: xiaozhe.hu@tufts.edu

DEPARTMENT OF MATHEMATICS,THE PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16802, USA

E-mail address: xu@math.psu.edu