SIAM J. NUMER. ANAL.
Vol. 49, No. 2, pp. 770–788

FIELD-OF-VALUES CONVERGENCE ANALYSIS OF AUGMENTED LAGRANGIAN PRECONDITIONERS FOR THE LINEARIZED NAVIER–STOKES PROBLEM*

MICHELE BENZI† AND MAXIM A. OLSHANSKII‡

Abstract. We study a block triangular preconditioner for finite element approximations of the linearized Navier–Stokes equations. The preconditioner is based on the augmented Lagrangian formulation of the problem and was introduced by the authors in [*SIAM J. Sci. Comput.*, 28 (2006), pp. 2095–2113]. In this paper we prove field-of-values type estimates for the preconditioned system which lead to optimal convergence bounds for the GMRES algorithm applied to solve the system. Two variants of the preconditioner are considered: an ideal one based on exact solves for the velocity submatrix, and a more practical variant based on block triangular approximations of the velocity submatrix.

Key words. Navier–Stokes equations, augmented Lagrangian, preconditioning, field of values, generalized minimal residual

AMS subject classifications. 65F10, 65N22, 65F50

DOI. 10.1137/100806485

1. Introduction.

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with sufficiently smooth boundary. Consider the Oseen equations in Ω,

$$\begin{aligned} \alpha \mathbf{u} - \nu\Delta\mathbf{u} + (\mathbf{a}\cdot\nabla)\,\mathbf{u} + \nabla p &= \mathbf{f}, \\ \operatorname{div}\mathbf{u} &= 0, \end{aligned} \tag{1.1}$$

together with suitable boundary conditions. Problem (1.1) arises from the linearization of the unsteady Navier–Stokes equations describing a viscous, Newtonian incompressible fluid with kinematic viscosity $\nu > 0$. The unknown quantities in (1.1) are the velocity field $\mathbf{u}$ and the pressure p; the forcing term $\mathbf{f}$, which represents external forces (such as gravity), is given. The parameter α is positive and proportional to the reciprocal of the time step for unsteady problems; it is zero in the steady case. We assume that the convection coefficient $\mathbf{a}$ is sufficiently smooth and divergence-free ($\operatorname{div}\mathbf{a} = 0$). For the purpose of theoretical analysis we will also assume, without loss of generality, that $\|\mathbf{a}\|_{L^\infty} = 1$.

Considerable effort has been devoted to the development of efficient solvers for linear systems arising from discretizations of problem (1.1). Such linear systems are of saddle point type and require specialized techniques for their efficient and robust solution. Several preconditioners for Krylov subspace methods have been proposed and analyzed; see [1] and [9] for broad overviews of this active field of research.

Recently, a promising approach based on the augmented Lagrangian (AL) formulation [10] has been proposed in [2] and further studied in [3, 4, 16]; see also [18]. The

*Received by the editors August 25, 2010; accepted for publication (in revised form) January 4, 2011; published electronically April 14, 2011.
http://www.siam.org/journals/sinum/49-2/80648.html

†Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322 (benzi@mathcs.emory.edu). The work of this author was supported in part by a grant of the University Research Committee of Emory University.

‡Department of Mechanics and Mathematics, Moscow State M. V. Lomonosov University, Moscow 119899, Russia (Maxim.Olshanskii@mtu-net.ru). The work of this author was supported in part by the Russian Foundation for Basic Research grants 08-01-00415, 09-01-00115 and by a Visiting Fellowship granted by the Emerson Center for Scientific Computation at Emory University.

preconditioners studied in these papers are of block triangular type. The extensive numerical experiments reported in these papers indicate that such preconditioners result in h-independent convergence (where h denotes the discretization parameter) and rather good robustness even for very small viscosity. Some theoretical results have been given in [2] and [3] in the form of eigenvalue bounds for the preconditioned matrices. While these bounds are useful, for instance, in the choice of the augmentation parameter (see [4]), they do not say much about the convergence behavior of preconditioned Krylov subspace methods such as GMRES [17]. Indeed, as is well known, knowledge of the eigenvalues alone is generally not sufficient to characterize the convergence of Krylov iterations applied to nonnormal matrices, and it may even be misleading in some cases.

In this paper we make use of field-of-values (FOV) estimates [14, 15] to prove the observed h-independent performance of block triangular preconditioners for the AL formulation of the discrete Oseen problem. In the "ideal case" where linear systems associated to the velocity submatrix in the preconditioner are solved exactly, we prove that preconditioned GMRES converges at a rate independent of all problem parameters h, ν, and α. For the modified (inexact) variant of the preconditioner introduced in [3] we establish h-independent convergence of GMRES.

The remainder of the paper is organized as follows. In section 2 we make some assumptions and preliminary observations on the finite element (FE) approximation of (1.1). The algebraic problem, its AL formulation, and the block triangular preconditioner are described in section 3. Some useful bounds for the Schur complement of the original saddle point system are derived in section 4. The main results are given in section 5 for the ideal version of the preconditioner and in section 6 for the modified one. Section 7 contains some closing remarks.

2. The finite element approximation. Let $\{\mathcal{T}_h\}_{h>0}$ be a family of triangulations of Ω in the sense of [6]. For the sake of analysis we assume the standard regularity condition to be satisfied by $\{\mathcal{T}_h\}$:

$$\sup_{h>0} \max_{K\in\mathcal{T}_h} \frac{\operatorname{diam}(K)}{\rho(K)} \le C < \infty,$$

where $\rho(K)$ is the radius of the ball inscribed in K. Note that we allow locally refined meshes. We assume conforming FE spaces $\mathbb{V}_h \subset \mathbf{H}_0^1(\Omega)$ and $\mathbb{Q}_h \subset L_0^2(\Omega)$, satisfying the LBB stability condition

$$\inf_{q_h\in\mathbb{Q}_h} \sup_{\mathbf{v}_h\in\mathbb{V}_h} \frac{(\operatorname{div}\mathbf{v}_h, q_h)}{\|\nabla\mathbf{v}_h\|\|q_h\|} \gtrsim 1. \tag{2.1}$$

In order to avoid the repeated use of generic but unspecified constants, here and in the remainder of the paper the binary relation $x \lesssim y$ means that there is a constant c such that $x \le c\,y$, and c does not depend on the parameters which x and y may depend on, e.g., ν, α, and h. Obviously, $x \gtrsim y$ is defined as $y \lesssim x$, and $x \simeq y$ when both $x \lesssim y$ and $y \gtrsim x$.

For $\alpha, \gamma \ge 0$ consider the following bilinear form on $(\mathbb{V}_h \times \mathbb{Q}_h) \times (\mathbb{V}_h \times \mathbb{Q}_h)$:

$$\begin{aligned} a(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) := {}& \alpha(\mathbf{u}_h, \mathbf{v}_h) + \nu(\nabla\mathbf{u}_h, \nabla\mathbf{v}_h) + \gamma(P_h\operatorname{div}\mathbf{u}_h, \operatorname{div}\mathbf{v}_h) \\ & + ((\mathbf{a}\cdot\nabla)\,\mathbf{u}_h, \mathbf{v}_h) - (p_h, \operatorname{div}\mathbf{v}_h) + (q_h, \operatorname{div}\mathbf{u}_h), \end{aligned}$$

where P_h is the L^2 orthogonal projector from L^2 into $\mathbb{Q}_h$. The FE method for the Oseen problem is based on the weak formulation: Find $\{\mathbf{u}_h, p_h\} \in \mathbb{V}_h \times \mathbb{Q}_h$ solving

$$a(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) = (\mathbf{f}, \mathbf{v}_h) \qquad \forall\ \{\mathbf{v}_h, q_h\} \in \mathbb{V}_h \times \mathbb{Q}_h. \tag{2.2}$$

Due to the obvious equalities

$$(2.3)\qquad \begin{gathered}(P_h\mathrm{div}\,\mathbf{u}_h,\mathrm{div}\,\mathbf{v}_h)=(P_h\mathrm{div}\,\mathbf{u}_h,P_h\mathrm{div}\,\mathbf{v}_h),\\ (p_h,\mathrm{div}\,\mathbf{v}_h)=(p_h,P_h\mathrm{div}\,\mathbf{v}_h),\quad (q_h,\mathrm{div}\,\mathbf{u}_h)=(q_h,P_h\mathrm{div}\,\mathbf{u}_h),\end{gathered}$$

the solution of (2.2) does not depend on γ, although the coefficient matrix of the system does.

For further study of the algebraic properties of the system we need some results from FE analysis. To this end, we introduce the following scalar products and norms (based on the symmetric part of $a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)$):

$$\begin{aligned}(\mathbf{u}_h,\mathbf{v}_h)_a&:=\alpha(\mathbf{u}_h,\mathbf{v}_h)+\nu(\nabla\mathbf{u}_h,\nabla\mathbf{v}_h)+\gamma(P_h\mathrm{div}\,\mathbf{u}_h,\mathrm{div}\,\mathbf{v}_h)\quad\text{on}\ \ \mathbb{V}_h,\\ (\{\mathbf{u}_h,p_h\};\{\mathbf{v}_h,q_h\})_a&:=(\mathbf{u}_h,\mathbf{v}_h)_a+(\nu+\gamma)(p_h,q_h)\quad\text{on}\ \ \mathbb{V}_h\times\mathbb{Q}_h.\end{aligned}$$

LEMMA 2.1. *The bilinear form $a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)$ satisfies the following stability and continuity estimates:*

$$(2.4)\qquad \inf_{\{\mathbf{u}_h,p_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\ \sup_{\{\mathbf{v}_h,q_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\frac{a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)}{\|\{\mathbf{u}_h,p_h\}\|_a\|\{\mathbf{v}_h,q_h\}\|_a}\ge C_1,$$

$$(2.5)\qquad \sup_{\{\mathbf{u}_h,p_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\ \sup_{\{\mathbf{v}_h,q_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\frac{a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)}{\|\{\mathbf{u}_h,p_h\}\|_a\|\{\mathbf{v}_h,q_h\}\|_a}\le C_2,$$

with some positive mesh-independent constants

$$C_1\gtrsim\left(1+(\nu+\gamma)\left(\nu+\gamma+\alpha+\frac{1}{\nu+\alpha}\right)\right)^{-1},\quad C_2\lesssim\frac{1}{\nu^{\frac12}(\nu^{\frac12}+\alpha^{\frac12})}.$$

Proof. Consider the auxiliary norm

$$\|\{\mathbf{u}_h,p_h\}\|_\sigma:=(\|\mathbf{u}_h\|_a^2+\sigma\|p_h\|^2)^{\frac12}\quad\text{on}\ \ \mathbb{V}_h\times\mathbb{Q}_h,$$

with a parameter $\sigma>0$ such that

$$(2.6)\qquad \sigma\simeq\left(\nu+\gamma+\alpha+(\nu+\alpha)^{-1}\right)^{-1}.$$

Due to Lemma 3.1 in [11] there exists a constant σ satisfying (2.6) such that

$$(2.7)\qquad \inf_{\{\mathbf{u}_h,p_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\ \sup_{\{\mathbf{v}_h,q_h\}\in\mathbb{V}_h\times\mathbb{Q}_h}\frac{a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)}{\|\{\mathbf{u}_h,p_h\}\|_\sigma\|\{\mathbf{v}_h,q_h\}\|_\sigma}\gtrsim 1.$$

Thus the lower bound in (2.4) follows from (2.7) and

$$\|\{\mathbf{v}_h,q_h\}\|_a\le\max\{1,(\nu+\gamma)^{\frac12}\sigma^{-\frac12}\}\|\{\mathbf{v}_h,q_h\}\|_\sigma\quad\forall\ \mathbf{v}_h,q_h\in\mathbb{V}_h\times\mathbb{Q}_h.$$

The estimate (2.5) follows from the definition of $a(\cdot,\cdot;\cdot,\cdot)$ and (2.3) through the application of the Cauchy–Schwarz and Friedrichs inequalities in a straightforward way:

$$
\begin{aligned}
a(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) &\le \alpha\|\mathbf{u}_h\|\|\mathbf{v}_h\| + \nu\|\nabla\mathbf{u}_h\|\|\nabla\mathbf{v}_h\| + \gamma\|P_h\mathrm{div}\,\mathbf{u}_h\|\|P_h\mathrm{div}\,\mathbf{v}_h\| \\
&\quad + \|\mathbf{a}\|_{L^\infty}\|\nabla\mathbf{u}_h\|\|\mathbf{v}_h\| + \|p_h\|\|P_h\mathrm{div}\,\mathbf{v}_h\| + \|q_h\|\|P_h\mathrm{div}\,\mathbf{u}_h\| \\
&\le \alpha\|\mathbf{u}_h\|\|\mathbf{v}_h\| + \nu\|\nabla\mathbf{u}_h\|\|\nabla\mathbf{v}_h\| + \gamma\|P_h\mathrm{div}\,\mathbf{u}_h\|\|P_h\mathrm{div}\,\mathbf{v}_h\| \\
&\quad + \frac{1}{\nu^{\frac12}(\nu^{\frac12}+\alpha^{\frac12})}\nu^{\frac12}\|\nabla\mathbf{u}_h\|(\alpha^{\frac12}\|\mathbf{v}_h\| + C_F\nu^{\frac12}\|\nabla\mathbf{v}_h\|) \\
&\quad + (\nu+\gamma)^{-\frac12}\|p_h\|(\gamma^{\frac12}\|P_h\mathrm{div}\,\mathbf{v}_h\| + \nu^{\frac12}\|\nabla\mathbf{v}_h\|) \\
&\quad + (\nu+\gamma)^{-\frac12}\|q_h\|(\gamma^{\frac12}\|P_h\mathrm{div}\,\mathbf{u}_h\| + \nu^{\frac12}\|\nabla\mathbf{u}_h\|) \\
&\le \frac{1}{\nu^{\frac12}(\nu^{\frac12}+\alpha^{\frac12})}\|\{\mathbf{u}_h, p_h\}\|_a\|\{\mathbf{v}_h, q_h\}\|_a,
\end{aligned}
$$

where C_F denotes the constant in the Friedrichs inequality. ▯

3. Algebraic problem and preconditioning.

Let $\{\phi_i\}_{1\le i\le n}$ and $\{\psi_j\}_{1\le j\le m}$ be bases of $\mathbb{V}_h$ and $\mathbb{Q}_h$, respectively. Define mass matrices for velocity and pressure elements,

$$(M_u)_{i,j} = (\phi_j, \phi_i), \quad (M_p)_{i,j} = (\psi_j, \psi_i),$$

and the matrices

$$D_{i,j} = (\nabla\phi_j, \nabla\phi_i), \quad N_{i,j} = ((\mathbf{a}\cdot\nabla)\,\phi_j, \phi_i), \quad B_{i,j} = -(\mathrm{div}\,\phi_j, \psi_i).$$

We shall use the following notation:

$$A := \alpha M_u + \nu D + N, \quad A_\gamma := A + \gamma B^T W^{-1} B,$$

with an auxiliary $m \times m$ matrix W, assumed to be symmetric and positive definite. With this notation and the natural ordering of unknowns, the linear algebraic system corresponding to the FE problem (2.2) takes the following form for $\gamma = 0$:

$$(3.1) \qquad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \qquad \text{or} \qquad \mathcal{A}\mathbf{x} = \mathbf{b}.$$

The augmented Lagrangian (AL) approach from [2] consists first of replacing the original system (3.1) with the equivalent one

$$(3.2) \qquad \begin{pmatrix} A + \gamma B^T W^{-1} B & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \qquad \text{or} \qquad \mathcal{A}_\gamma\mathbf{x} = \mathbf{b},$$

followed by preconditioning (3.2) with a block triangular preconditioner of the form

$$(3.3) \qquad \mathcal{P}_\gamma = \begin{pmatrix} \widehat{A}_\gamma & B^T \\ 0 & -\frac{1}{\nu+\gamma}W \end{pmatrix}.$$

Here and in what follows $\widehat{A}_\gamma$ denotes a preconditioner for the velocity block A_γ. A typical choice for W is a diagonal approximation of M_p; see [2]. Using the main diagonal of M_p works well in practice [19].

Remark 3.1. For $W = M_p$ the matrix $\mathcal{A}_\gamma$ is the coefficient matrix of the FE problem (2.2) for any γ.

As was mentioned in the introduction, eigenvalue bounds for $\mathcal{P}_\gamma^{-1}\mathcal{A}_\gamma$ have been established in [2] and [3]. The present paper is concerned with FOV-type bounds for $\mathcal{P}_\gamma^{-1}\mathcal{A}_\gamma$ which lead to rigorous convergence estimates for GMRES. For this purpose we need a few more technicalities introduced below.

For a symmetric positive definite matrix H we define the H-inner product of two vectors u and v by $\langle u, v\rangle_H := \langle Hu, v\rangle$, with the corresponding vector and induced matrix norm. Given two symmetric positive definite matrices $H_1 \in \mathbb{R}^{n_1\times n_1}$ and $H_2 \in \mathbb{R}^{n_2\times n_2}$ and a matrix $C \in \mathbb{R}^{n_2\times n_1}$, we will use the following notation for the norm of C as an operator from $\mathbb{R}^{n_1}$ to $\mathbb{R}^{n_2}$:

$$\|C\|_{H_1\to H_2} := \sup_{x\in\mathbb{R}^{n_1}} \|Cx\|_{H_2}\|x\|_{H_1}^{-1}. \tag{3.4}$$

Let $A_S := \frac{1}{2}(A_\gamma + A_\gamma^T)$. Note that $N = -N^T$, so $A_S = \alpha M_u + \nu D + \gamma B^T W^{-1} B$ is a symmetric positive definite matrix and therefore invertible. We need a few specific scalar products and corresponding norms:

$$\begin{aligned} \langle p, q\rangle_M &:= \langle M_p p, q\rangle \quad \text{on} \quad \mathbb{R}^m,\\ \langle \{u,p\},\{v,q\}\rangle_a &:= \langle A_S u, v\rangle + (\nu+\gamma)^{-1}\langle M_p p, q\rangle \quad \text{on} \quad \mathbb{R}^{n+m},\\ \langle \{u,p\},\{v,q\}\rangle_{-a} &:= \langle A_S^{-1} u, v\rangle + (\nu+\gamma)\langle M_p^{-1} p, q\rangle \quad \text{on} \quad \mathbb{R}^{n+m}. \end{aligned}$$

The corresponding norms on $\mathbb{R}^{n+m}$ are denoted by $\|\mathbf{z}\|_a = \|\{u,p\}\|_a$ and $\|\mathbf{z}\|_{-a} = \|\{u,p\}\|_{-a}$ for $\mathbf{z} = \{u,p\}$. We note that $\langle \{u,p\},\{v,q\}\rangle_a = (\{\mathbf{u}_h, p_h\};\{\mathbf{v}_h, q_h\})_a$, where u, v, p, q are vectors of coefficients for FE functions $\mathbf{u}_h, p_h, \mathbf{v}_h, q_h$ and $\|\ \|_{-a}$ is the dual norm of $\|\ \|_a$ with respect to the ℓ^2-duality.

Due to Remark 3.1, the result of Lemma 2.1 is equivalent to the following conditions for the augmented matrix $\mathcal{A}_\gamma$:

$$\inf_{\mathbf{z}_1\in\mathbb{R}^{n+m}} \sup_{\mathbf{z}_2\in\mathbb{R}^{n+m}} \frac{\langle \mathcal{A}_\gamma \mathbf{z}_1, \mathbf{z}_2\rangle}{\|\mathbf{z}_1\|_a\|\mathbf{z}_2\|_a} \ge C_1, \tag{3.5}$$

$$\sup_{\mathbf{z}_1\in\mathbb{R}^{n+m}} \sup_{\mathbf{z}_2\in\mathbb{R}^{n+m}} \frac{\langle \mathcal{A}_\gamma \mathbf{z}_1, \mathbf{z}_2\rangle}{\|\mathbf{z}_1\|_a\|\mathbf{z}_2\|_a} \le C_2, \tag{3.6}$$

with the same constants C_1 and C_2 as in (2.4) and (2.5).

For $\mathcal{B} \in \mathbb{R}^{(n+m)\times(n+m)}$ denote

$$\mu(\mathcal{B}) := \inf_{\mathbf{z}\in\mathbb{R}^{n+m}} \frac{\langle \mathcal{B}\mathbf{z}, \mathbf{z}\rangle_{-a}}{\langle \mathbf{z}, \mathbf{z}\rangle_{-a}}.$$

The convergence of the preconditioned GMRES method in the product norm can be estimated as [7]

$$\|\mathbf{r}^k\|_{-a} \le \left(1 - \mu(\mathcal{A}_\gamma\mathcal{P}^{-1})\,\mu(\mathcal{P}\mathcal{A}_\gamma^{-1})\right)^{k/2}\|\mathbf{r}^0\|_{-a}, \tag{3.7}$$

where $\mathbf{r}^k$ is the kth residual vector from $\mathbb{R}^{n+m}$.

4. Some bounds for S_0.

In this section we prove and discuss some bounds for the Schur complement of the original (nonaugmented) saddle point problem,

$$S_0 := BA^{-1}B^T \quad (A = A_\gamma \text{ for } \gamma = 0),$$

and its inverse. These results are important for our further analysis. We note that in some situations (e.g., in the case of enclosed flow) the discrete divergence operator B

does not have full row rank, making S_0 singular. In this case, S_0^{-1} is to be understood as the Moore–Penrose generalized inverse $S_0^{\dagger}$. Similarly, when taking the infimum of the generalized Rayleigh quotient $\langle S_0\, q, q\rangle / \|q\|_M^2$, we consider only vectors q in the orthogonal complement of the null space of S_0. In order to keep the notation simple, these conventions will be tacitly assumed in what follows.

We begin with the following lemma.

LEMMA 4.1. *The following bounds hold:*

$$\inf_{q\in\mathbb{R}^m} \frac{\left\langle S_0^{-1} M_p q, q \right\rangle_M}{\|q\|_M^2} \geq \nu, \tag{4.1}$$

$$\inf_{q\in\mathbb{R}^m} \frac{\langle S_0 q, q\rangle}{\|q\|_M^2} \gtrsim \frac{1}{\alpha+\nu+\nu^{-1}}. \tag{4.2}$$

Proof. First we show (4.1). For arbitrary $q \in \mathbb{R}^m$, consider $p = S_0^{-1} M_p q$. This can be rewritten as

$$BA^{-1}B^T p = M_p q \quad \Leftrightarrow \quad Bu = M_p q,\ Au = B^T p.$$

Therefore the FE counterparts $q_h \in \mathbb{Q}_h$ of q and $p_h \in \mathbb{Q}_h$ of p satisfy the following relations for all $r_h \in \mathbb{Q}_h$, $\mathbf{v}_h \in \mathbb{V}_h$, and an auxiliary $\mathbf{u}_h \in \mathbb{V}_h$:

$$\begin{aligned} &(q_h, r_h) := -(\operatorname{div} \mathbf{u}_h, r_h),\\ &\alpha(\mathbf{u}_h, \mathbf{v}_h) + \nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + ((\mathbf{a}\cdot\nabla)\mathbf{u}_h, \mathbf{v}_h) = -(p_h, \operatorname{div} \mathbf{v}_h). \end{aligned} \tag{4.3}$$

We set in (4.3) $\mathbf{v}_h = \mathbf{u}_h$ and use $((\mathbf{a}\cdot\nabla)\mathbf{u}_h, \mathbf{u}_h) = 0$ to get

$$\begin{aligned} \left\langle S_0^{-1} M_p q, q \right\rangle_M &= \langle M_p p, q\rangle = (p_h, q_h) = -(\operatorname{div} \mathbf{u}_h, p) = \nu\|\nabla \mathbf{u}_h\|^2 + \alpha\|\mathbf{u}_h\|^2\\ &\geq \nu\|\operatorname{div} \mathbf{u}_h\|^2 = \nu\|q_h\|^2 = \nu\|q\|_M^2, \end{aligned}$$

which implies (4.1).

The estimate in (4.2) is proved below with the help of the following identity [8]:

$$\frac{1}{2}\left(A^{-1} + A^{-T}\right) = A_S^{-\frac{1}{2}}(I - (A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2)^{-1} A_S^{-\frac{1}{2}}. \tag{4.4}$$

Recall that $A = A_S + N$, $A_S = \alpha M_u + \nu D$ is the symmetric part of A (discretization of $-\nu\Delta + \alpha I$), and N is the skew-symmetric part of A (discretization of $(\mathbf{a}\cdot\nabla)$). To show (4.2) we first note the estimate

$$\begin{aligned} \left\langle BA_S^{-1}B^T q, q \right\rangle &= \sup_{v\in\mathbb{R}^{n+m}} \frac{\langle Bv, q\rangle^2}{\langle A_S v, v\rangle} = \sup_{\mathbf{v}_h\in\mathbb{V}_h} \frac{(\operatorname{div} \mathbf{v}_h, q_h)^2}{\nu\|\nabla \mathbf{v}_h\|^2 + \alpha\|\mathbf{v}_h\|^2}\\ &\gtrsim \sup_{\mathbf{v}_h\in\mathbb{V}_h} \frac{(\operatorname{div} \mathbf{v}_h, q_h)^2}{(\nu+\alpha)\|\nabla \mathbf{v}_h\|^2} \gtrsim \frac{1}{\nu+\alpha}\|q_h\|^2 = \frac{1}{\nu+\alpha}\|q\|_M^2. \end{aligned} \tag{4.5}$$

With the help of (4.4) and (4.5) we obtain

$$\begin{aligned} \langle S_0 q, q\rangle &= \left\langle AB^T q, B^T q \right\rangle = \left\langle (I - (A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2)^{-1} A_S^{-\frac{1}{2}} B^T q, A_S^{-\frac{1}{2}} B^T q \right\rangle\\ &\geq \frac{\langle A_S^{-\frac{1}{2}} B^T q, A_S^{-\frac{1}{2}} B^T q\rangle}{(1 + \|(A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2\|)} = \frac{\left\langle BA_S^{-1}B^T q, q \right\rangle}{(1 + \|(A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2\|)}\\ &\gtrsim \frac{1}{(\nu+\alpha)(1 + \|(A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2\|)}\|q\|_M^2. \end{aligned} \tag{4.6}$$

Further, we estimate

$$\begin{aligned} \|(A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}})^2\| &\le \|A_S^{-\frac{1}{2}} N\, A_S^{-\frac{1}{2}}\|^2 = \max\{|\lambda|^2 \,:\, \lambda \in \mathrm{sp}(A_S^{-\frac{1}{2}} N A_S^{-\frac{1}{2}})\} \\ &= \max\{|\lambda|^2 \,:\, \lambda \in \mathrm{sp}(A_S^{-1} N)\} \\ &\le \|A_S^{-1} N\|^2_{M_u}, \end{aligned} \tag{4.7}$$

where we have used $\mathrm{sp}(\cdot)$ to denote the spectrum. For a given $w \in \mathbb{R}^n$ and $u = A_S^{-1} N\, w$ consider their FE counterparts $\mathbf{w}_h, \mathbf{u}_h \in \mathbb{V}_h$, respectively. Then it holds that

$$\nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + \alpha(\mathbf{u}_h, \mathbf{v}_h) = ((\mathbf{a}\cdot\nabla)\mathbf{w}_h, \mathbf{v}_h) \quad \forall\ \mathbf{v}_h \in \mathbb{V}_h.$$

Setting $\mathbf{v}_h = \mathbf{u}_h$ and applying the Cauchy–Schwarz and Young inequalities, we get

$$\begin{aligned} \nu\|\nabla \mathbf{u}_h\|^2 + \alpha\|\mathbf{u}_h\|^2 &= ((\mathbf{a}\cdot\nabla)\mathbf{w}_h, \mathbf{u}_h) = -((\mathbf{a}\cdot\nabla)\mathbf{u}_h, \mathbf{w}_h) \\ &\le \frac{\nu}{2}\|\nabla \mathbf{u}_h\|^2 + \nu^{-1}\|\mathbf{a}\|^2_{L^\infty}\|\mathbf{w}_h\|^2. \end{aligned}$$

Hence, by the Friedrichs inequality and $\|\mathbf{a}\|_{L^\infty} = 1$, we obtain

$$(\nu+\alpha)\|\mathbf{u}_h\|^2 \lesssim \nu\|\nabla \mathbf{u}_h\|^2 + \alpha\|\mathbf{u}_h\|^2 \le \nu^{-1}\|\mathbf{w}_h\|^2.$$

Therefore

$$\|\mathbf{u}_h\|^2 \lesssim \frac{1}{\alpha\nu + \nu^2}\|\mathbf{w}_h\|^2,$$

and we have proved

$$\|A_S^{-1} N\|^2_{M_u} \lesssim \frac{1}{\alpha\nu + \nu^2}. \tag{4.8}$$

Estimates (4.6), (4.7), and (4.8) yield the estimate in (4.2). ∎

The best constant C_S from the estimate

$$\|S_0^{-1} M_p\|_M \le C_S \tag{4.9}$$

plays an important role in our convergence analysis. Below we prove some theoretical bounds for this constant and perform a few numerical experiments to assess the actual value of C_S and check its dependence on problem parameters.

The lemma below gives the most general bound, which is valid for general convection fields, boundary conditions, and FEs; however, it can be nonoptimal in terms of ν and α.

LEMMA 4.2. *The constant C_S satisfies*

$$C_S \lesssim \nu + \alpha + \frac{1}{\nu+\alpha}. \tag{4.10}$$

Proof. For arbitrary $q \in \mathbb{R}^m$ and $p = S_0^{-1} M_p q$ it holds that

$$\frac{\|p\|_M}{\|q\|_M} = \frac{\|p_h\|}{\|q_h\|},$$

where $q_h \in \mathbb{Q}_h$ is the FE counterpart of q and $p_h \in \mathbb{Q}_h$ together with some auxiliary $\mathbf{u}_h \in \mathbb{V}_h$ satisfies (4.3) for all $r_h \in \mathbb{Q}_h$, $\mathbf{v}_h \in \mathbb{V}_h$. With the help of the LBB condition

(2.1) and the Cauchy–Schwarz inequality and using $((\mathbf{a}\cdot\nabla)\mathbf{u}_h, \mathbf{v}_h) = ((\mathbf{a}\cdot\nabla)\mathbf{v}_h, \mathbf{u}_h)$ and $\|\mathbf{a}\|_{L^\infty} = 1$, we get from the second relation in (4.3)

$$\|p_h\| \lesssim \sup_{\mathbf{v}_h\in\mathbb{V}_h} \frac{(p_h, \operatorname{div}\mathbf{v}_h)}{\|\nabla\mathbf{v}_h\|} \lesssim (\alpha+1)\|\mathbf{u}_h\| + \nu\|\nabla\mathbf{u}_h\|. \tag{4.11}$$

The LBB condition can be reformulated as follows: For a given q_h there exists $\mathbf{w}_h \in \mathbb{V}_h$ such that

$$\|\nabla\mathbf{w}_h\| \lesssim \|q_h\| \quad \text{and} \quad (\operatorname{div}\mathbf{w}_h, r_h) = (q_h, r_h) \quad \forall\ r_h \in \mathbb{Q}_h. \tag{4.12}$$

Further decompose $\mathbf{u}_h = \mathbf{w}_h + \mathbf{u}_h^0$. Relations (4.3) yield for all $r_h \in \mathbb{Q}_h$, $\mathbf{v}_h \in \mathbb{V}_h$

$$\begin{aligned} &(\operatorname{div}\mathbf{u}_h^0, r_h) = 0,\\ &\alpha(\mathbf{u}_h^0, \mathbf{v}_h) + \nu(\nabla\mathbf{u}_h^0, \nabla\mathbf{v}_h) + ((\mathbf{a}\cdot\nabla)\mathbf{u}_h^0, \mathbf{v}_h) + (p_h, \operatorname{div}\mathbf{v}_h)\\ &\qquad = \alpha(\mathbf{w}_h, \mathbf{v}_h) + \nu(\nabla\mathbf{w}_h, \nabla\mathbf{v}_h) + ((\mathbf{a}\cdot\nabla)\mathbf{w}_h, \mathbf{v}_h). \end{aligned} \tag{4.13}$$

Setting $\mathbf{v}_h = \mathbf{u}_h^0$ in the second relation of (4.13) and using the first one, we obtain

$$\alpha\|\mathbf{u}_h^0\|^2 + \nu\|\nabla\mathbf{u}_h^0\|^2 = \alpha(\mathbf{w}_h, \mathbf{u}_h^0) + \nu(\nabla\mathbf{w}_h, \nabla\mathbf{u}_h^0) + ((\mathbf{a}\cdot\nabla)\mathbf{w}_h, \mathbf{u}_h^0).$$

With the same arguments as in (4.11) and applying (4.12) and the Friedrichs inequality, we estimate

$$\begin{aligned} \alpha\|\mathbf{u}_h^0\|^2 + \nu\|\nabla\mathbf{u}_h^0\|^2 &\lesssim \alpha\|\mathbf{w}_h\|^2 + \nu\|\nabla\mathbf{w}_h\|^2 + \frac{\|\mathbf{a}\|_{L^\infty}}{\nu+\alpha}\|\nabla\mathbf{w}_h\|^2\\ &\lesssim \left(\nu + \alpha + \frac{1}{\nu+\alpha}\right)\|q_h\|^2. \end{aligned} \tag{4.14}$$

Combining (4.12) and (4.14), we get through the triangle inequality

$$\alpha\|\mathbf{u}_h\|^2 + \nu\|\nabla\mathbf{u}_h\|^2 \lesssim \left(\nu + \alpha + \frac{1}{\nu+\alpha}\right)\|q_h\|^2. \tag{4.15}$$

Now (4.11) and (4.15) give

$$\|p_h\| \lesssim \left(\nu + \alpha + \frac{1}{\nu+\alpha}\right)\|q_h\|,$$

which completes the proof. ◻

Next, we argue that in general the bound in (4.10) can be pessimistic for the case of $\alpha \to 0$. Indeed, assume that the problem admits Fourier analysis, i.e., $\mathbf{a}$ is a constant vector and periodic boundary conditions are imposed; then the action of S_0^{-1} on a given harmonic $\psi_{\mathbf{k}} = \exp(\imath\mathbf{k}\cdot\mathbf{x})$ is

$$S_0^{-1}\psi_{\mathbf{k}} = (\nu + \alpha|\mathbf{k}|^{-2} + \imath\,(\mathbf{a}\cdot\mathbf{k})|\mathbf{k}|^{-2})\psi_{\mathbf{k}}.$$

Thus, in such a setting it holds that

$$\|S_0^{-1}\|^2 = \lambda_{\max}(S_0^{-1}S_0^{-T}) = \max_{|\mathbf{k}|>0}\{(\nu + \alpha|\mathbf{k}|^{-2})^2 + (\mathbf{a}\cdot\mathbf{k})^2|\mathbf{k}|^{-4}\}.$$

Therefore, we get

$$C_S \lesssim \nu + \alpha + 1, \tag{4.16}$$

which does not blow up for $\nu \to 0$ and $\alpha \to 0$ in contrast to (4.10).

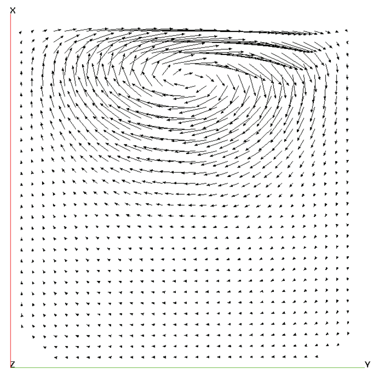

FIG. 1. *The "driven cavity" vortex given by* (4.18).

It is interesting to point out that the ν, α-dependence scenario from (4.16) can be shown to hold in somewhat more general cases than the periodic one. For example, (4.16) can be proved to hold on the continuous level in the case of mixed boundary conditions in a bounded domain: $\mathbf{u}\cdot\mathbf{n} = 0$, $(\nabla\times\mathbf{u})\times\mathbf{n} = 0$ on $\partial\Omega$. However, $\mathbf{a}$ is still assumed to be a constant vector.

Below we show results of experiments where we have computed $\|S_0^{-1}M_p\|_M$ for a set of different parameters and convection fields $\mathbf{a}$. One example of $\mathbf{a}$ is the Poiseuille flow in a channel given by

$$\mathbf{a}(x, y) = \begin{cases} 4y(1-y), \\ 0. \end{cases} \tag{4.17}$$

The second example is the "driven cavity"-type vortex [5]:

$$\mathbf{a}(x, y) = \begin{cases} \frac{r_2}{2\pi}\frac{\exp(r_2 y)}{(\exp(r_2)-1)}\sin\left(\frac{2\pi(\exp(r_2 y)-1)}{\exp(r_2)-1}\right)\left(1-\cos\left(\frac{2\pi(\exp(r_1 x)-1)}{\exp(r_1)-1}\right)\right), \\ -\frac{r_1}{2\pi}\frac{\exp(r_1 x)}{(\exp(r_1)-1)}\sin\left(\frac{2\pi(\exp(r_1 x)-1)}{\exp(r_1)-1}\right)\left(1-\cos\left(\frac{2\pi(\exp(r_2 y)-1)}{\exp(r_2)-1}\right)\right). \end{cases} \tag{4.18}$$

The position of the center of the vortex,

$$(x_0, y_0) = (r_1^{-1}\log((\exp(r_1)+1)/2),\ r_2^{-1}\log((\exp(r_2)+1)/2)),$$

is governed by the two parameters r_1 and r_2. Unless otherwise stated we set $r_1 = 4$ and $r_2 = 0.1$; the generated velocity field is shown in Figure 1.

To produce numerical results we use two FE pairs, one with continuous pressure approximation and the other with discontinuous pressure approximation. These are, respectively, the isoP2-P1 and isoP2-P0 FE pairs. Both are LBB stable in the sense of (2.1); see [13]. We use the uniform west-north triangulation of $\Omega = (0, 1)^2$. Values of h in the tables below are $\sqrt{2}$ times the radius of the pressure element, while the velocity is piecewise linear with respect to a two times finer grid. Note that A^{-1} is a dense matrix, and so is $S_0 \in \mathbb{R}^{m\times m}$. Hence, the computational cost for finding S_0^{-1} and $\|S_0^{-1}M_p\|_M$ quickly becomes prohibitive with increasing number of unknowns, restricting our calculations to rather moderate values of h. Setting $h = 1/32$ results in $m = 1089$, $n = 7938$ (isoP2-P1) and $m = 2048$, $n = 7938$ (isoP2-P0).

The values of $\|S_0^{-1}M_p\|_M$ for both choices of $\mathbf{a}$ and FEs are given in Tables 1 and 2. We put $\alpha = 0$ and vary ν from 1 down to 1e-7. For these test problems the

TABLE 1
Values of $\|S_0^{-1}M_p\|_M$ *for isoP2-P1 FEs.*

Mesh size h	Viscosity ν							
	1	0.1	0.01	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
a is the Poiseuille flow (4.17)								
1/16	10.53	1.054	0.1873	0.1802	0.3322	0.7119	3.060	30.53
1/32	10.23	1.024	0.1868	0.1801	0.1792	0.8338	3.847	21.05
a is the rotating vortex (4.18)								
1/16	10.53	1.053	0.1058	0.2912	1.358	3.666	5.597	5.040
1/32	10.23	1.023	0.1025	0.2986	1.558	5.172	6.728	6.680

TABLE 2
Values of $\|S_0^{-1}M_p\|_M$ *for isoP2-P0 FEs.*

Mesh size h	Viscosity ν							
	1	0.1	0.01	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
a is the Poiseuille flow (4.17)								
1/16	4.873	0.5003	0.1802	0.1394	0.2177	0.2957	0.2690	0.2593
1/32	4.996	0.5094	0.1848	0.1650	0.1682	0.2051	0.5532	3.600
a is the rotating vortex (4.18)								
1/16	4.873	0.4877	0.0626	0.0797	0.0508	0.0922	0.2930	0.2490
1/32	4.995	0.4998	0.0651	0.1520	0.0639	0.0597	0.1289	0.2059

results suggest that $\|S_0^{-1}M_p\|_M$ and thus C_S are of order 1 except for some growth for the case of the Poiseuille flow and extremely small viscosities. It is important to note that the results for such small values of ν are purely of academic interest, since in practice such flows no longer have an equilibrium state and time-stepping schemes have to be adopted for the computations, leading to $\alpha > 0$. Also, for large mesh Reynolds numbers a stabilization is often added in practice. We do not study how these additional mesh-dependent stabilizing terms affect the results. In general, such terms add some extra numerical diffusion; therefore we do not expect any substantial growth of $\|S_0^{-1}M_p\|_M$ in this case.

An interesting question is whether high gradients of **a** affect the values of C_S. To experiment with this we perform a series of experiments with **a** given in (4.18), for increasing values of the parameter r_1. Increasing r_1 has the effect of driving the vortex closer to the right boundary of Ω, forming a boundary layer. Since the characteristic boundary layer is known to be of $O(\nu^{\frac{1}{2}})$ thickness, we set the corresponding value of ν equal to $\|\mathbf{a}\|_{L^\infty}^{-1}(1 - r_1^{-1}\log((\exp(r_1) + 1)/2))^2$. The factor $\|\mathbf{a}\|_{L^\infty}^{-1}$ appears because we rescale the equations in such a way that the L^∞ norm of the convection field is 1. The results (the computed $\|S_0^{-1}M_p\|_M$ and reference values of ν) are given in Table 3. We note that increasing r_1 further would not make sense, since a computationally affordable mesh size h would not resolve the layer.

5. Analysis for exact AL approach. In this section we analyze the ideal case, in which $\widehat{A}_\gamma = A_\gamma$. Although not very practical, analyzing this case is useful as it provides a “best case scenario” in terms of convergence behavior of AL-based preconditioners.

Concerning the exact AL approach we prove the following main result.

TABLE 3
The values of $\|S_0^{-1}M_p\|_M$ *for the case of boundary layer in* $\mathbf{a}$, $h = 1/32$.

r_1	2	4	8	16
ν	1.37e-1	2.85e-2	3.84e-3	4.81e-4
	isoP2-P0 FE			
$\|S_0^{-1}M_p\|_M$	1.401	0.2921	0.0394	0.0208
	isoP2-P1 FE			
$\|S_0^{-1}M_p\|_M$	0.6845	0.1437	0.0311	0.0133

THEOREM 5.1. *Assume* $\gamma \lesssim \alpha + \nu + \nu^{-1}$. *For the FOV quantities in the GMRES convergence estimate* (3.7) *the following bounds hold:*

$$1 \lesssim \mu(\mathcal{P}\mathcal{A}_\gamma^{-1}),$$
$$\frac{(\gamma + C_S^2 \min\{\alpha^{-1}, \nu, \nu^{-1}\})(\gamma + \nu)}{\gamma^2 + C_S^2} \lesssim \mu(\mathcal{A}_\gamma \mathcal{P}^{-1}). \tag{5.1}$$

Proof. By the definitions of $\mu(\cdot)$ and matrices $\mathcal{A}_\gamma$ and $\mathcal{P}$, (5.1) is equivalent to

$$\inf_{[u,p]\in\mathbb{R}^{n+m}} \frac{\left\langle \begin{bmatrix} I & 0 \\ -(\nu+\gamma)^{-1}WS_\gamma^{-1}BA_\gamma^{-1} & (\nu+\gamma)^{-1}WS_\gamma^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle_{-a}}{\|[u,p]\|^2} \gtrsim 1 \tag{5.2}$$

and

$$\inf_{[u,p]\in\mathbb{R}^{n+m}} \frac{\left\langle \begin{bmatrix} I & 0 \\ BA_\gamma^{-1} & (\nu+\gamma)S_\gamma W^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle_{-a}}{\|[u,p]\|^2} \gtrsim \frac{(\gamma + C_S^2 \min\{\alpha^{-1}, \nu^{-1}, \nu\})(\gamma+\nu)}{\gamma^2 + C_S^2}. \tag{5.3}$$

The key is to show the following estimates for the Schur complement matrix of the augmented system:

$$(\nu+\gamma)\|q\|_M^2 \lesssim \left\langle S_\gamma^{-1}M_p q, q \right\rangle_M \quad \forall\, q \in \mathbb{R}^m, \tag{5.4}$$

$$\frac{(C_S^2 \min\{\alpha^{-1}, \nu^{-1}, \nu\} + \gamma)}{C_S^2 + \gamma^2}\|q\|_M^2 \lesssim \langle S_\gamma q, q\rangle \quad \forall\, q \in \mathbb{R}^m. \tag{5.5}$$

Here and below we make use of the identity (cf. [12, Proposition 2.1])

$$S_\gamma^{-1} = S_0^{-1} + \gamma\, W^{-1}. \tag{5.6}$$

Using the spectral equivalence of W and M_p, we obtain (5.4) from (5.6) and (4.1).

Next, we show (5.5). Using (5.6), we find

$$\inf_{q\in\mathbb{R}^m} \frac{\langle S_\gamma q, q\rangle}{\|q\|_M^2} = \inf_{q\in\mathbb{R}^m} \frac{\left\langle S_\gamma^{-1}M_p q, q\right\rangle_M}{\|S_\gamma^{-1}M_p q\|_M^2} \geq \inf_{q\in\mathbb{R}^m} \frac{\gamma\|q\|_M^2 + \left\langle S_0^{-1}M_p q, q\right\rangle_M}{\|S_\gamma^{-1}M_p q\|_M^2}. \tag{5.7}$$

By the triangle inequality and the spectral equivalence of W and M_p, we get

$$\|S_\gamma^{-1}M_p q\|_M^2 = \|S_0^{-1}M_p q + \gamma\, W^{-1}M_p q\|_M^2 \lesssim \|S_0^{-1}M_p q\|_M^2 + \gamma^2\|q\|_M^2. \tag{5.8}$$

Estimates (5.7) and (5.8) imply

$$
\begin{aligned}
\inf_{q\in\mathbb{R}^m} \frac{\langle S_\gamma q, q\rangle}{\|q\|_M^2} &\gtrsim \inf_{q\in\mathbb{R}^m} \frac{\gamma\|q\|_M^2 + \langle S_0^{-1}M_p q, q\rangle_M}{\gamma^2\|q\|_M^2 + \|S_0^{-1}M_p q\|_M^2} \\
&= \inf_{q\in\mathbb{R}^m} \frac{\gamma\left[\frac{\|q\|_M}{\|S_0^{-1}M_p q\|_M}\right]^2 + \frac{\langle S_0^{-1}M_p q, q\rangle_M}{\|S_0^{-1}M_p q\|_M^2}}{\gamma^2\left[\frac{\|q\|_M}{\|S_0^{-1}M_p q\|_M}\right]^2 + 1}.
\end{aligned} \tag{5.9}
$$

The second term in the numerator on the right-hand side of (5.9) can be handled (due to Lemma 4.1) as follows:

$$
\inf_{q\in\mathbb{R}^m} \frac{\langle S_0^{-1}M_p q, q\rangle_M}{\|S_0^{-1}M_p q\|_M^2} = \inf_{q\in\mathbb{R}^m} \frac{\langle S_0 q, q\rangle}{\|q\|_M^2} \gtrsim \frac{1}{\alpha+\nu+\nu^{-1}}. \tag{5.10}
$$

Now from (5.9) and (5.10) we get

$$
\inf_{q\in\mathbb{R}^m} \frac{\langle S_\gamma q, q\rangle}{\|q\|_M^2} \gtrsim \inf_{q\in\mathbb{R}^m} \frac{\gamma\left[\frac{\|q\|_M}{\|S_0^{-1}M_p q\|_M}\right]^2 + \frac{1}{\alpha+\nu+\nu^{-1}}}{\gamma^2\left[\frac{\|q\|_M}{\|S_0^{-1}M_p q\|_M}\right]^2 + 1}. \tag{5.11}
$$

Note that for $\gamma \le \alpha+\nu+\nu^{-1}$ the function $f(t) = \frac{\gamma t + (\alpha+\nu+\nu^{-1})^{-1}}{\gamma^2 t + 1}$ is monotonically increasing. Therefore, based on (4.9) we finally get

$$
\inf_{q\in\mathbb{R}^m} \frac{\langle S_\gamma q, q\rangle}{\|q\|_M^2} \gtrsim \frac{\gamma C_S^{-2} + (\alpha+\nu+\nu^{-1})^{-1}}{\gamma^2 C_S^{-2} + 1}. \tag{5.12}
$$

This proves (5.5).

Owing to the definition of the product $\langle\cdot,\cdot\rangle_{-a}$, together with the spectral equivalence of W and M_p and the estimates (5.4) and (5.5) for the Schur complement, the bounds in (5.2) and (5.3), and hence the theorem, will follow from the lower bounds

$$
\inf_{[u,p]\in\mathbb{R}^{n+m}} \frac{\left\langle \begin{bmatrix} A_S^{-1} & 0 \\ -S_\gamma^{-1}BA_\gamma^{-1} & S_\gamma^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} A_S^{-1} & 0 \\ 0 & S_\gamma^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle} \gtrsim 1 \tag{5.13}
$$

and

$$
\inf_{[u,p]\in\mathbb{R}^{n+m}} \frac{\left\langle \begin{bmatrix} A_S^{-1} & 0 \\ (\nu+\gamma)W^{-1}BA_\gamma^{-1} & (\nu+\gamma)^2 W^{-1}S_\gamma W^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} A_S^{-1} & 0 \\ 0 & (\nu+\gamma)^2 W^{-1}S_\gamma W^{-1} \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \begin{bmatrix} u \\ p \end{bmatrix} \right\rangle} \gtrsim 1. \tag{5.14}
$$

By simple eigenvalue estimates the quantities on the left-hand sides of (5.13) and (5.14) are easily shown to be greater than or equal to $\frac{1}{2}$; see [14, page 586]. ◻

COROLLARY 5.2. *For $\gamma \lesssim \alpha+\nu+\nu^{-1}$ the residual norms in the GMRES method with ideal AL preconditioner satisfy*

$$
\|\mathbf{r}^k\|_{-a} \le \left(1 - c\,\frac{(\gamma + C_S^2\min\{\nu,\nu^{-1},\alpha^{-1}\})(\gamma+\nu)}{\gamma^2 + C_S^2}\right)^{k/2} \|\mathbf{r}^0\|_{-a}
$$

for some positive constant c independent of h and other parameters. In particular, for the choice $\gamma \simeq C_S$ (this choice satisfies the restriction $\gamma \lesssim \alpha + \nu + \nu^{-1}$ due to Lemma 4.2), we get

$$\|\mathbf{r}^k\|_{-a} \le q^k \|\mathbf{r}^0\|_{-a},$$

where $q < 1$ is independent of h and other problem parameters.

The foregoing convergence result sheds some light on the optimal choice of γ. Indeed, the eigenvalue analysis in [2] shows that

$$\mathrm{sp}\,(\mathcal{P}\mathcal{A}_\gamma^{-1}) = \{1\} \bigcup \left\{ \frac{\gamma + \nu}{\gamma + \mu_i^{-1}} \,:\, 1 \le i \le m \right\}, \tag{5.15}$$

where μ_i are the generalized eigenvalues of

$$BA^{-1}B^T q = \mu M_p\, q. \tag{5.16}$$

Then, applying estimates for μ_i from [8] (for the case $\alpha = 0$), one concludes that it is sufficient to set $\gamma = O(\nu^{-1})$ to ensure that all nonunit eigenvalues of $\mathcal{P}^{-1}\widehat{\mathcal{A}}$ are contained in a box $[\mathtt{a}, \mathtt{A}] \times [\mathtt{b}, \mathtt{B}]$, $\mathtt{a} > 0$, in the complex plane, with $\mathtt{a}, \mathtt{A}, \mathtt{b}, \mathtt{B}$ independent on ν and h. However, all numerical results we are aware of suggest that $\gamma = O(1)$ is already sufficient for parameter-independent convergence. The present analysis shows that the overestimation comes from applying the eigenvalue bounds of [8], which are likely nonoptimal. Within our analysis the same choice of $\gamma = O(\nu^{-1})$ appears only if one assumes the bound of C_S from Lemma 4.2 to be optimal, which may not be true for many flows of interest, as suggested by the Fourier analysis and numerical experiments of section 4.

6. Analysis for modified AL preconditioner.

We need some further notation and definitions. Consider the natural block structure of A_γ (for simplicity we consider the two-dimensional case, although the results below hold for the three-dimensional case as well):

$$A_\gamma = \begin{pmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{pmatrix}.$$

We define the preconditioner

$$\widehat{\mathcal{P}}_\gamma = \begin{pmatrix} \widehat{A}_\gamma & B^T \\ 0 & -\frac{\rho}{\nu+\gamma} W \end{pmatrix},$$

where

$$\widehat{A}_\gamma = \begin{pmatrix} A_{11} & A_{12}^T \\ 0 & A_{22} \end{pmatrix}$$

and ρ is an additional parameter which is introduced for the sake of analysis. First we need a few auxiliary results given in the following lemmas.

LEMMA 6.1. *For all $C, D \in \mathbb{R}^{n\times n}$ with D symmetric positive definite it holds that*

$$\|C\|_D = \|C^T\|_{D^{-1}}.$$

Proof. This simple result can be shown as follows:

$$\begin{aligned}\|C\|_D &= \max_{x\in\mathbb{R}^n}\frac{\|C\,x\|_D}{\|x\|_D} = \max_{x,y\in\mathbb{R}^n}\frac{\langle D^{\frac12}C\,x,y\rangle}{\|x\|_D\|y\|} = \max_{x,y\in\mathbb{R}^n}\frac{\langle x,C^TD^{\frac12}y\rangle}{\|x\|_D\|y\|}\\ &= \max_{x,y\in\mathbb{R}^n}\frac{\langle x,C^Ty\rangle}{\|x\|_D\|y\|_{D^{-1}}} = \max_{x,y\in\mathbb{R}^n}\frac{\langle D^{-\frac12}x,C^Ty\rangle}{\|x\|\|y\|_{D^{-1}}} = \max_{x,y\in\mathbb{R}^n}\frac{\langle x,D^{-\frac12}C^Ty\rangle}{\|x\|\|y\|_{D^{-1}}}\\ &= \max_{y\in\mathbb{R}^n}\frac{\|D^{-\frac12}C^Ty\|}{\|y\|_{D^{-1}}} = \|C^T\|_{D^{-1}}. \qquad \square\end{aligned}$$

We recall that A_S is the symmetric part of A_γ. Denote $A^s_{ij} = \frac12(A_{ij} + A^T_{ij})$, $i, j \in \{1, 2\}$; then we can write

$$A_S = \begin{pmatrix} A^s_{11} & A^s_{12}\\ A^s_{12} & A^s_{22}\end{pmatrix}.$$

Note that A^s_{11} and A^s_{22} are both symmetric positive definite. With this notation and the one introduced in (3.4), the following lemma holds.

LEMMA 6.2. *For any $\nu > 0$, $\alpha \ge 0$ and any constant $\kappa \in (0, 1)$, there exists a sufficiently small $\gamma > 0$ such that*

$$\begin{aligned}&\|A^{-1}_{11}A^T_{12}\|_{A^s_{22}\to A^s_{11}} \le \kappa, \quad \|A^{-1}_{22}A_{12}\|_{A^s_{11}\to A^s_{22}} \le \kappa,\\ &\|A^{-T}_{11}A^T_{12}\|_{A^s_{22}\to A^s_{11}} \le \kappa, \quad \|A^{-T}_{22}A_{12}\|_{A^s_{11}\to A^s_{22}} \le \kappa\end{aligned} \tag{6.1}$$

hold independently of the mesh size h.

Proof. We will check the first inequality in (6.1); the second follows by similar arguments. For arbitrary $w \in \mathbb{R}^n$ consider $u = A^{-1}_{11}A^T_{12}w$. To prove the inequality $\|A^{-1}_{11}A^T_{12}\|_{A^s_{22}\to A^s_{11}} \le \kappa$, we need to show that

$$\langle A^s_{11}u, u\rangle \le \kappa^2 \langle A^s_{22}w, w\rangle. \tag{6.2}$$

Consider the FE counterparts of u and w, that is, u_h and w_h from V_h (V_h is the FE space for one component of the velocity, so $\mathbb{V}_h = V_h \times V_h$). Using the definition of the matrices involved, checking (6.2) is equivalent to showing that

$$\alpha\|u_h\|^2 + \nu\|\nabla u_h\|^2 + \gamma\|P_h(u_h)_x\|^2 \le \kappa^2(\alpha\|w_h\|^2 + \nu\|\nabla w_h\|^2 + \gamma\|P_h(w_h)_y\|^2). \tag{6.3}$$

Since $A_{11}u = A^T_{12}w$, the functions u_h and w_h satisfy the relation

$$\begin{aligned}\alpha(u_h, v_h) + \nu(\nabla u_h, \nabla v_h) + \gamma(P_h(u_h)_x, (v_h)_x) + ((\mathbf{a}\cdot\nabla)\,u_h, v_h)&\\ = \gamma(P_h(w_h)_y, (v_h)_x) \quad \forall\ v_h \in V_h.&\end{aligned} \tag{6.4}$$

Letting $u_h = v_h$ in (6.4) and using the Cauchy–Schwarz inequality and $\|P_h\| \le 1$ (since P_h is an L^2 orthogonal projector), one verifies (6.3) for any $\gamma \le c\,\nu$ for some mesh- and parameter-independent constant c.

The arguments for proving the third inequality in (6.1), i.e., $\|A^{-T}_{11}A^T_{12}\|_{A^s_{22}\to A^s_{11}} \le \kappa$, are the same with the only difference being the minus sign in front of $((\mathbf{a}\cdot\nabla)\,u_h, v_h)$ in (6.4). This sign does not matter, since the term cancels out for the choice of $v_h = u_h$. The second and fourth inequalities in (6.1) are proved similarly. $\square$

Consider now the following representations for the preconditioned velocity blocks:

$$A_\gamma\widehat{A}^{-1}_\gamma = I - \widetilde{C}, \quad \text{with } \widetilde{C} = \begin{pmatrix} 0 & 0\\ -A_{12}A^{-1}_{11} & A_{12}A^{-1}_{11}A^T_{12}A^{-1}_{22}\end{pmatrix},$$

and

$$\widehat{A}_\gamma^{-1}A_\gamma = I - \widehat{C}, \quad \text{with } \widehat{C} = \begin{pmatrix} A_{11}^{-1}A_{12}^T A_{22}^{-1}A_{12} & 0 \\ -A_{22}^{-1}A_{12} & 0 \end{pmatrix}.$$

For the matrices $\widetilde{C}$ and $\widehat{C}$ we prove the following result.

LEMMA 6.3. *For any $\nu > 0$, $\alpha \geq 0$ and any constant $\kappa \in (0,1)$, there exists a sufficiently small $\gamma > 0$ such that*

$$\|\widetilde{C}\|_{A_S^{-1}} \leq \kappa \quad \textit{and} \quad \|\widehat{C}\|_{A_S} \leq \kappa \tag{6.5}$$

hold independently of the mesh size h.

Proof. By Lemma 6.1, the first bound in (6.5) is equivalent to

$$\|\widetilde{C}^T\|_{A_S} \leq \kappa_1. \tag{6.6}$$

Consider the natural splitting:

$$\widetilde{C}^T = \widetilde{C}_1 + \widetilde{C}_2, \quad \text{with } \widetilde{C}_1 = \begin{pmatrix} 0 & -A_{11}^{-T}A_{12}^T \\ 0 & 0 \end{pmatrix}, \quad \widetilde{C}_2 = \begin{pmatrix} 0 & 0 \\ 0 & A_{22}^{-T}A_{12}A_{11}^{-T}A_{12}^T \end{pmatrix}.$$

Hence for (6.6) it is sufficient to show that

$$\|\widetilde{C}_1\|_{A_S} \leq \frac{1}{2}\kappa_1 \quad \text{and} \quad \|\widetilde{C}_2\|_{A_S} \leq \frac{1}{2}\kappa_1. \tag{6.7}$$

By the definition of A_S and $\widetilde{C}_1$, the first inequality in (6.7) is equivalent to

$$\|A_{11}^{-T}A_{12}^T\|_{A_{22}^s \to A_{11}^s} \leq \frac{1}{2}\kappa_1.$$

The latter holds by Lemma 6.2. To show the second inequality in (6.7), we apply Lemma 6.2 once again through the following decomposition:

$$\|\widetilde{C}_2\|_{A_S} = \|A_{22}^{-T}A_{12}A_{11}^{-T}A_{12}^T\|_{A_{22}^s \to A_{22}^s} \leq \|A_{22}^{-T}A_{12}\|_{A_{11}^s \to A_{22}^s}\|A_{11}^{-T}A_{12}^T\|_{A_{22}^s \to A_{11}^s}.$$

With the help of Lemma 6.2, the bound for $\widehat{C}$ in (6.5) is proved by arguments similar to those used above to show (6.6). ∎

For the analysis of the inexact AL preconditioner we apply Theorem 3.9 of [15]. For the reader's convenience we give this result below in the notation of our paper.

THEOREM 6.4 (see [15]). *Denote $\widehat{S}_\gamma = B\widehat{A}_\gamma^{-1}B^T$. Assume that*

$$\inf_{\mathbf{z}_1 \in \mathbb{R}^{n+m}} \sup_{\mathbf{z}_2 \in \mathbb{R}^{n+m}} \frac{\langle \mathcal{A}_\gamma \mathbf{z}_1, \mathbf{z}_2 \rangle}{\|\mathbf{z}_1\|_a \|\mathbf{z}_2\|_a} \geq C_1, \tag{6.8}$$

$$\sup_{\mathbf{z}_1 \in \mathbb{R}^{n+m}} \sup_{\mathbf{z}_2 \in \mathbb{R}^{n+m}} \frac{\langle \mathcal{A}_\gamma \mathbf{z}_1, \mathbf{z}_2 \rangle}{\|\mathbf{z}_1\|_a \|\mathbf{z}_2\|_a} \leq C_2, \tag{6.9}$$

$$\mu_a \leq \inf_{u \in \mathbb{R}^n} \frac{\langle A_\gamma \widehat{A}_\gamma^{-1}u, u\rangle_{A_S^{-1}}}{\|u\|_{A_S^{-1}}^2}, \quad \|A_\gamma \widehat{A}_\gamma^{-1}\|_{A_S^{-1}} \leq \Gamma_a, \tag{6.10}$$

$$\mu_s \leq (\nu+\gamma) \inf_{q \in \mathbb{R}^m} \frac{\langle \widehat{S}_\gamma W^{-1}q, q\rangle_{M_p^{-1}}}{\|q\|_{M_p^{-1}}^2}, \quad (\nu+\gamma)\|\widehat{S}_\gamma W^{-1}\|_{M_p^{-1}} \leq \Gamma_s \tag{6.11}$$

hold with some positive mesh-independent constants $C_1, C_2, \mu_a, \Gamma_a, \mu_s, \Gamma_s$. Then there exists $\rho_0 > 0$ such that for all $\rho \geq \rho_0$ it holds that

$$0 < a_1 \leq \mu(\mathcal{A}_\gamma \widehat{\mathcal{P}}^{-1}), \quad \|\widehat{\mathcal{P}} \mathcal{A}_\gamma^{-1}\|_{-a} \leq a_2$$

with some mesh-independent constants a_1, a_2, provided that

$$\|I - A_\gamma \widehat{A}_\gamma^{-1}\|_{A_S^{-1}} \leq \rho^{-1}. \tag{6.12}$$

We remark that the proof of this result in [15] gives the following sufficient lower bound on ρ_0:

$$\rho_0 \geq \frac{(C_2\Gamma_a + C_2^2 C_1^{-2}\Gamma_s \mu_a^{-1})^2 + \mu_s^2}{2\mu_a \mu_s}.$$

The main result of this section is the following theorem.

THEOREM 6.5. *For any $\nu > 0$ and $\alpha \geq 0$ and sufficiently small $\gamma > 0$ and large enough ρ, the preconditioned GMRES method with modified AL preconditioner satisfies the convergence estimate*

$$\|\mathbf{r}^k\|_{-a} \leq q^k \|\mathbf{r}^0\|_{-a},$$

with some h-independent positive constant $q < 1$.

Proof. We shall use the result of Theorem 6.4. The estimates (6.8) and (6.9) were already shown in (3.5) and (3.6) with mesh-independent constants C_1 and C_2 estimated in Lemma 2.1. Thus, it remains to show (6.10), (6.11) with some mesh-independent constants μ_a, Γ_a and μ_s, Γ_s.

First we check (6.10). Based on the splitting $A_\gamma \widehat{A}_\gamma^{-1} = I - \widetilde{C}$ and Lemma 6.3, one finds for sufficiently small $\gamma > 0$

$$\left\langle A_\gamma \widehat{A}_\gamma^{-1} u, u \right\rangle_{A_S^{-1}} = \|u\|^2_{A_S^{-1}} - \left\langle \widetilde{C} u, u \right\rangle_{A_S^{-1}} \geq (1 - \|\widetilde{C}\|_{A_S^{-1}})\|u\|^2_{A_S^{-1}} \geq (1-\kappa)\|u\|^2_{A_S^{-1}}$$

and

$$\|A_\gamma \widehat{A}_\gamma^{-1}\|_{A_S^{-1}} \leq (1 + \|\widetilde{C}\|_{A_S^{-1}}) \leq 1 + \kappa,$$

with some mesh-independent $\kappa \in (0,1)$. Thus we check that (6.10) holds with some mesh-independent constants μ_a and Γ_a.

Now we turn to proving (6.11). For the choice of $W = M_p$ the bounds in (6.11) are equivalent to

$$\mu_s \leq (\nu + \gamma) \inf_{q \in \mathbb{R}^m} \frac{\langle \widehat{S}_\gamma q, q \rangle}{\|q\|_M^2}, \quad (\nu + \gamma)\|M_p^{-1} \widehat{S}_\gamma\|_M \leq \Gamma_s. \tag{6.13}$$

Thanks to the lower bound for the Schur complement of the augmented system in (5.5), for the first inequality in (6.13) it is sufficient to show that

$$\langle S_\gamma q, q \rangle \lesssim \left\langle \widehat{S}_\gamma q, q \right\rangle \quad \forall\, q \in \mathbb{R}^m. \tag{6.14}$$

Since $S_\gamma = BA_\gamma^{-1}B^T$ and $\widehat{S}_\gamma = B\widehat{A}_\gamma^{-1}B^T$, the inequality (6.14) follows from

$$\left\langle A_\gamma^{-1}u, u\right\rangle \lesssim \left\langle \widehat{A}_\gamma^{-1}u, u\right\rangle \quad \forall\, u \in \mathbb{R}^n$$
$$\Updownarrow$$
$$\langle A_\gamma u, u\rangle \lesssim \left\langle \widehat{A}_\gamma^{-1}A_\gamma u, A_\gamma u\right\rangle \quad \forall\, u \in \mathbb{R}^n$$
$$\Updownarrow$$
$$\langle A_\gamma u, u\rangle \lesssim \left\langle \widehat{A}_\gamma \left[\widehat{A}_\gamma^{-1}A_\gamma u\right], \left[\widehat{A}_\gamma^{-1}A_\gamma u\right]\right\rangle \quad \forall\, u \in \mathbb{R}^n$$
$$\Updownarrow$$
$$\langle A_S u, u\rangle \lesssim \left\langle \frac{1}{2}(\widehat{A}_\gamma + \widehat{A}_\gamma^T)\left[\widehat{A}_\gamma^{-1}A_\gamma u\right], \left[\widehat{A}_\gamma^{-1}A_\gamma u\right]\right\rangle \quad \forall\, u \in \mathbb{R}^n.$$

Now we note that for $\gamma \simeq \nu$ it holds that $\langle A_S u, u\rangle \simeq \langle \frac{1}{2}(\widehat{A}_\gamma + \widehat{A}_\gamma^T)u, u\rangle$ for all $u \in \mathbb{R}^n$. Thus, for sufficiently small γ, estimate (6.14) follows from

$$\langle u, u\rangle_{A_S} \lesssim \left\langle \left[\widehat{A}_\gamma^{-1}A_\gamma u\right], \left[\widehat{A}_\gamma^{-1}A_\gamma u\right]\right\rangle_{A_S} \quad \forall\, u \in \mathbb{R}^n. \tag{6.15}$$

To show (6.15) we apply Lemma 6.3:

$$\begin{aligned}\left\langle \left[\widehat{A}_\gamma^{-1}A_\gamma u\right], \left[\widehat{A}_\gamma^{-1}A_\gamma u\right]\right\rangle_{A_S} &= \|(I - \widehat{C})u\|_{A_S}^2 \geq (1 - \|\widehat{C}\|_{A_S})^2\|u\|_{A_S}^2 \\ &\geq (1-\kappa)^2\|u\|_{A_S}^2,\end{aligned}$$

where κ is a mesh-independent constant less than 1. Thus, the first inequality in (6.11) is proved.

The second inequality in (6.11) is proved by the following argument. For $W = M_p$ we get

$$\begin{aligned}\|\widehat{S}_\gamma W^{-1}\|_{M_p^{-1}} &= \|M_p^{-\frac{1}{2}}\widehat{S}_\gamma M_p^{-\frac{1}{2}}\| = \|M_p^{-\frac{1}{2}}B\widehat{A}_\gamma^{-1}B^T M_p^{-\frac{1}{2}}\| \\ &= \|(M_p^{-\frac{1}{2}}BD^{-\frac{1}{2}})(D^{\frac{1}{2}}\widehat{A}_\gamma^{-1}D^{\frac{1}{2}})(D^{-\frac{1}{2}}B^T M_p^{-\frac{1}{2}})\| \\ &\leq \|D^{-\frac{1}{2}}B^T M_p^{-\frac{1}{2}}\|^2\|D^{\frac{1}{2}}\widehat{A}_\gamma^{-1}D^{\frac{1}{2}}\|. \end{aligned} \tag{6.16}$$

For the first factor on the right-hand side of (6.16) it holds that

$$\begin{aligned}\|D^{-\frac{1}{2}}B^T M_p^{-\frac{1}{2}}\| &= \sup_{q\in\mathbb{R}^m}\frac{\|D^{-\frac{1}{2}}B^T q\|}{\|q\|_M} = \sup_{q\in\mathbb{R}^m}\sup_{u\in\mathbb{R}^n}\frac{\langle D^{-\frac{1}{2}}B^T q, u\rangle}{\|q\|_M\|u\|} \\ &= \sup_{q\in\mathbb{R}^m}\sup_{u\in\mathbb{R}^n}\frac{\langle q, B\,u\rangle}{\|q\|_M\|D^{\frac{1}{2}}u\|} = \sup_{q_h\in\mathbb{Q}_h}\sup_{\mathbf{u}_h\in\mathbb{V}_h}\frac{(q_h, \operatorname{div} u_h)}{\|q_h\|\|\nabla u_h\|} \leq 1.\end{aligned} \tag{6.17}$$

It remains to estimate the second factor on the right-hand side of (6.16). First, we get by Lemma 6.1

$$\|D^{\frac{1}{2}}\widehat{A}_\gamma^{-1}D^{\frac{1}{2}}\| = \|D\widehat{A}_\gamma^{-1}\|_{D^{-1}} = \|\widehat{A}_\gamma^{-T}D\|_D.$$

Next, we compute

$$\widehat{A}_\gamma^{-T}D = \begin{pmatrix} A_{11}^{-T}D & 0 \\ -A_{22}^{-T}A_{12}A_{11}^{-T}D & A_{22}^{-T}D \end{pmatrix}.$$

Note that D is a block diagonal matrix and the D-norm of every block of the matrix $\widehat{A}_\gamma^{-T} D$ can be estimated separately. For example, let us consider the (1,1) block $A_{11}^{-T} D$. By the definition of the matrices involved, the bound $\|A_{11}^{-T} D\|_D \le C_1$ is equivalent to the estimate

$$\|\nabla u_h\| \le C_1 \|\nabla w_h\|, \tag{6.18}$$

with arbitrary $w_h \in V_h$ and where $u_h \in V_h$ solves

$$\alpha(u_h, v_h) + \nu(\nabla u_h, \nabla v_h) + \gamma(P_h(u_h)_x, (v_h)_x) - ((\mathbf{a}\cdot\nabla)\, u_h, v_h) \\ = (\nabla w_h, \nabla v_h) \quad \forall\ v_h \in V_h. \tag{6.19}$$

Setting $v_h = u_h$ in (6.19) immediately gives (6.18) with $C_1 = \nu^{-\frac{1}{2}}$. Similarly, $\|A_{22}^{-T} D\|_D \le \nu^{-\frac{1}{2}}$, and, finally, with the same arguments

$$\|A_{22}^{-T} A_{12} A_{11}^{-T} D\|_D \le \|A_{22}^{-T} A_{12}\|_D \|A_{11}^{-T} D\|_D \le \gamma\nu^{-1}.$$

This completes the proof. □

7. Conclusions. In this paper we have used field-of-values analysis to establish rigorous convergence bounds for GMRES with various AL-based preconditioners of block triangular type applied to the discrete Oseen problem. In particular, we proved that for suitable choices of the parameter γ the GMRES method with the ideal (exact) AL preconditioner converges at a rate independent of both the mesh size h and the viscosity ν. For the modified (inexact) AL preconditioner, the GMRES rate of convergence is proved to be h-independent. Thus, the theory developed in this paper provides a rigorous justification for the convergence behavior observed in practice and reported in the papers [2, 3, 4].

REFERENCES

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

[2] M. Benzi and M. A. Olshanskii, *An augmented Lagrangian-based approach to the Oseen problem*, SIAM J. Sci. Comput., 28 (2006), pp. 2095–2113.

[3] M. Benzi, M. A. Olshanskii, and Z. Wang, *Modified augmented Lagrangian preconditioners for the incompressible Navier–Stokes equations*, Internat. J. Numer. Methods Fluids, to appear; also available as DOI: 10.1002/fld.2267.

[4] M. Benzi and Z. Wang, *Analysis of augmented Lagrangian-based preconditioners for the steady incompressible Navier–Stokes equations*, SIAM J. Sci. Comput., to appear.

[5] S. Berrone, *Adaptive discretization of the Navier–Stokes equations by stabilized finite element methods*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 4435–4455.

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.

[7] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[8] H. Elman and D. Silvester, *Fast nonsymmetric iterations and preconditioning for Navier–Stokes equations*, SIAM J. Sci. Comput., 17 (1996), pp. 33–46.

[9] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2005.

[10] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, Stud. Math. Appl. 15, North–Holland, Amsterdam, 1983.

[11] T. Gelhard, G. Lube, M. A. Olshanskii, and J.-H. Starcke, *Stabilized finite element schemes with LBB-stable elements for incompressible flows*, J. Comput. Appl. Math., 177 (2005), pp. 243–267.

[12] G. H. Golub and C. Greif, *On solving block-structured indefinite linear systems*, SIAM J. Sci. Comput., 24 (2003), pp. 2076–2092.
[13] M. D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms*, Academic Press, Boston, 1989.
[14] A. Klawonn and G. Starke, *Block triangular preconditioners for nonsymmetric saddle point problems: Field-of-value analysis*, Numer. Math., 81 (1999), pp. 577–594.
[15] D. Loghin and A. J. Wathen, *Analysis of preconditioners for saddle-point problems*, SIAM J. Sci. Comput., 25 (2004), pp. 2029–2049.
[16] M. A. Olshanskii and M. Benzi, *An augmented Lagrangian approach to linearized problems in hydrodynamic stability*, SIAM J. Sci. Comput., 30 (2008), pp. 1459–1473.
[17] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[18] M. ur Rehman, C. Vuik, and G. Segal, *A comparison of preconditioners for incompressible Navier–Stokes solvers*, Internat. J. Numer. Methods Fluids, 57 (2008), pp. 1731–1751.
[19] A. J. Wathen, *Realistic eigenvalue bounds for the Galerkin mass matrix*, IMA J. Numer. Anal., 7 (1987), pp. 449–457.