Received: 7 April 2016 | Revised: 11 November 2016 | Accepted: 4 December 2016

DOI 10.1002/nla.2085

RESEARCH ARTICLE

LU factorizations and ILU preconditioning for stabilized discretizations of incompressible Navier–Stokes equations

Igor Konshin[1] | Maxim Olshanskii[2] | Yuri Vassilevski[3]

[1] Institute of Numerical Mathematics and Dorodnicyn Computing Centre FRC IC, Russian Academy of Sciences, Moscow, Russia
[2] Department of Mathematics, University of Houston, Houston, TX, USA
[3] Institute of Numerical Mathematics, Russian Academy of Sciences, and Moscow Institute of Physics and Technology, Moscow, Russia

Correspondence
Maxim Olshanskii, Department of Mathematics, University of Houston, TX 77004, USA.
Email: molshan@math.uh.edu

Funding information
Russian Science Foundation, Grant/Award Number: 14-31-00024

Summary

The paper studies numerical properties of LU and incomplete LU factorizations applied to the discrete linearized incompressible Navier–Stokes problem also known as the Oseen problem. A commonly used stabilized Petrov–Galerkin finite element method for the Oseen problem leads to the system of algebraic equations having a 2 × 2-block structure. While enforcing better stability of the finite element solution, the Petrov–Galerkin method perturbs the saddle-point structure of the matrix and may lead to less favorable algebraic properties of the system. The paper analyzes the stability of the LU factorization. This analysis quantifies the effect of the streamline upwind Petrov–Galerkin stabilization in terms of the perturbation made to a nonstabilized system. The further analysis shows how the perturbation depends on the particular finite element method, the choice of stabilization parameters, and flow problem parameters. The analysis of LU factorization and its stability helps to understand the properties of threshold ILU factorization preconditioners for the system. Numerical experiments for a model problem of blood flow in a coronary artery illustrate the performance of the threshold ILU factorization as a preconditioner. The dependence of the preconditioner properties on the stabilization parameters of the finite element method is also studied numerically.

KEYWORDS

iterative methods, finite element method, hemodynamics, Navier–Stokes equations, preconditioning, SUPG stabilization, threshold ILU factorization

1 | INTRODUCTION

The paper addresses the question of developing fast algebraic solves for finite element discretizations of the linearized Navier–Stokes equations. The Navier–Stokes equations describe the motion of incompressible Newtonian fluids. For a bounded domain $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$), with boundary $\partial\Omega$ and time interval $[0, T]$, the equations read

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} - \nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f} \text{ in } \Omega\times(0,T] \\ \operatorname{div}\ \mathbf{u} = 0 \text{ in } \Omega\times[0,T] \\ \mathbf{u} = \mathbf{g} \quad \text{on } \Gamma_0\times[0,T], -\nu(\nabla\mathbf{u})\cdot\mathbf{n} + p\mathbf{n} = \mathbf{0} \text{ on } \Gamma_{\mathrm{N}}\times[0,T] \\ \mathbf{u}(\mathbf{x},0) = \mathbf{u}_0(\mathbf{x}) \text{ in } \Omega. \end{cases} \tag{1}$$

The unknowns are the velocity vector field $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ and the pressure field $p = p(\mathbf{x}, t)$. The volume forces $\mathbf{f}$ and boundary and initial values $\mathbf{g}$ and $\mathbf{u}_0$ are given. Parameter ν is the kinematic viscosity, $\partial\Omega = \overline{\Gamma}_0 \cup \overline{\Gamma}_{\mathrm{N}}$ and $\Gamma_0 \neq \emptyset$. An important parameter of the flow is the dimensionless Reynolds number $\mathrm{Re} = \frac{UL}{\nu}$, where U and L are characteristic velocity and linear dimension. Solving Equation 1 numerically is known to get harder for higher values of Re; in particular, some special modelling of flow scales unresolved by the mesh may be needed. Implicit time discretization and linearization of the Navier–Stokes system (Equation 1) by Picard fixed-point iteration result in a sequence of (generalized) Oseen problems of the form

$$\begin{cases} \alpha\mathbf{u} - \nu\Delta\mathbf{u} + (\mathbf{w}\cdot\nabla)\mathbf{u} + \nabla p = \hat{\mathbf{f}} \text{ in } \Omega \\ \operatorname{div}\ \mathbf{u} = \hat{g} \text{ in } \Omega \\ \mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_0, -\nu(\nabla\mathbf{u})\cdot\mathbf{n} + p\mathbf{n} = \mathbf{0} \text{ on } \Gamma_{\mathrm{N},} \end{cases} \tag{2}$$

where $\mathbf{w}$ is a known velocity field from a previous iteration or time step and α is proportional to the reciprocal of the time step. Nonhomogenous boundary conditions in the nonlinear problem are accounted in the right-hand side.

Finite element (FE) methods for Equations 1 and 2 may suffer from different sources of instabilities. One is a possible incompatibility of pressure and velocity FE pairs. A remedy is a choice of FE spaces satisfying the inf-sup or Ladyzhenskaya-Babuska-Brezzi (LBB) condition[1] or the use of pressure stabilizing techniques. A major source of instabilities stems from dominating inertia terms for large Reynolds numbers. There exist several variants of stabilized FE methods, which combine stability and accuracy such as the streamline upwind Petrov–Galerkin (SUPG) method, the Galerkin/Least-squares, algebraic sub-grid scale, and internal penalty techniques, see, for example, these studies.[2–5] These methods simultaneously suppress spurious oscillations caused by both dominating advection and non-LBB-stable FE spaces. The combination of LBB-stable velocity-pressure FE pairs with advection stabilization is also often used in practice and studied in the literature, see, for example, these studies.[6,7] For numerical experiments and FE analysis in this paper, we consider a variant of the SUPG method. Details of the method are given later in this paper.

An FE spatial discretization of Equation 2 results in large, sparse systems of the form

$$\begin{pmatrix} A & \tilde{B}^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{3}$$

where u and p represent the discrete velocity and pressure, respectively, and $A \in \mathbb{R}^{n\times n}$ is the discretization of the diffusion, convection, and time-dependent terms. The matrix A accounts also for certain stabilization terms. Matrices B and $\tilde{B}^T \in \mathbb{R}^{n\times m}$ are (negative) discrete divergence and gradient. These matrices may also be perturbed due to stabilization. It is typical for the stabilized methods that $B \neq \tilde{B}$; while for a plain Galerkin method, these two matrices are the same. Matrix $C \in \mathbb{R}^{m\times m}$ results from possible pressure stabilization terms, and f and g contain forcing and boundary terms. For the LBB stable FEs, no pressure stabilization is required, and so, $C = 0$ holds. If the LBB condition is not satisfied, the stabilization matrix $C \neq 0$ is typically symmetric and positive semidefinite. For $B = \tilde{B}$ of the full rank and positive definite $A = A^T$, the solution to Equation 3 is a saddle point. Otherwise, one often refers to Equation 3 as a generalized saddle-point system, see, for example, this study.[8]

Considerable work has been done in developing efficient preconditioners for Krylov subspace methods applied to system (Equation 3) with $\tilde{B} = B$; see the comprehensive studies in these studies[8–10] of the preconditioning exploiting the block structure of the system. A common approach is based on preconditioners for block A and pressure Schur complement matrix $S = BA^{-1}\tilde{B}^T + C$, see these studies[11–13] for recent developments. Well-known block preconditioners are not completely robust with respect to variations of viscosity parameter, properties of advective velocity field $\mathbf{w}$, grid size and anisotropy ratio, and the domain geometry. The search of a more robust black-box type approach to solve algebraic system (Equation 3) stimulates an interest in developing preconditioners based on incomplete factorizations. Clearly, computing a suitable incomplete LU factorizations of Equation 3 is challenging and requires certain care for (at least) the following reasons. The matrix can be highly nonsymmetric for higher Reynolds number flows; even in symmetric case, the matrix is indefinite (both positive and negative eigenvalues occur in the spectrum); and extra stabilization terms may break the positive definiteness of A and/or of the Schur complement. Nevertheless, a progress has been recently reported in developing incomplete LU preconditioners for saddle-point matrices and generalized saddle-point matrices. Thus, the authors of these studies[14,15] studied the signed incomplete Cholesky-type preconditioners for symmetric saddle-point systems, corresponding to the Stokes problem. For the FE discretization of the incompressible Navier–Stokes equations, the authors of these studies[16,17] developed ILU preconditioners, where the fill-in is allowed based on the connectivity of nodes rather than actual nonzeros in the matrix. The papers[17,18] studied several reordering techniques for ILU factorization of Equation 3 and found that some of the resulting preconditioners are competitive with the most advanced block preconditioners. Elementwise threshold incomplete LU factorizations for nonsymmetric saddle-point matrices were developed in this study.[19] In that paper, an extension of the Tismenetsky–Kaporin variant of ILU factorization for nonsymmetric matrices is used as a preconditioner for the FE discretizations of the Oseen equations. Numerical analysis and experiments with the (nonstabilized) Galerkin methods for the incompressible Navier–Stokes equations demonstrated the robustness and efficiency of this approach. An important advantage of preconditioners based on elementwise ILU decomposition is that they are straightforward to implement in standard FE codes.

In the present paper, we extend the method and analysis from this study[19] to the system of algebraic equations resulting from the stabilized formulations of the Navier–Stokes equations. Hence, we are interested in the numerically challenging case of higher Reynolds number flows. The effect of different stabilization techniques on the accuracy of FE solutions is substantial and is well studied in the literature. However, not that much research has addressed the question of how the stabilization affects the algebraic properties of the discrete systems, see this study. [9] The present study intends to fill this gap. We analyze the stability of the (exact) LU factorization and numerical properties of a threshold ILU factorization for Equation 3. One might expect that stabilization adds to the ellipticity of matrices and hence, improves algebraic properties. This is certainly the situation in particular cases of scalar advection-diffusion equations and linear elements. However, for saddle-point problems and higher order elements, the situation appears to be more delicate. In particular, stability of the LU factorization may impose more restrictive bounds on the stabilization parameters than those satisfied by optimal parameters with respect to FE solution accuracy. We study the explicit dependence of algebraic properties of Equation 3 on flow, stabilization, and discretization parameters and show that larger values of the stabilization parameter may affect the algebraic stability. Therefore, for those fluid flow problems, which require SUPG stabilization, suitable parameters meet both restrictions: they are large enough to add necessary stability for the FE solution but not too large to guarantee stable factorizations of algebraic systems.

The remainder of the paper is organized as follows. In Section 2, we give necessary details on the FE method for the Oseen equations. Section 3 studies stability of the exact LU factorizations for Equation 3. We derive the sufficient conditions for the existence and stability of the LU factorization without pivoting. These conditions and an estimate on the entries of the resulting LU factors are given in terms of the properties of the (1,1)-block A, auxiliary Schur complement matrix $BA^{-1}B^T + C$, and the perturbation matrix $B-\tilde{B}$. In Section 4, we apply this analysis to system (Equation 3) arising from SUPG-stabilized FE discretization of the Oseen system. In Section 5, we briefly discuss the implication of our analysis of LU factorization on the stability of a two-parameter Tismenetsky–Kaporin variant of the threshold ILU factorization for nonsymmetric nondefinite problems. This factorization is used in our numerical experiments. In Section 6, we study the numerical performance of the method on the sequence of linear systems appearing in simulation of a blood flow in a right coronary artery. Section 7 collects conclusions and a few closing remarks.

2 | FE METHOD AND SUPG STABILIZATION

In this paper, we consider an inf-sup stable conforming FE method stabilized by the SUPG method. To formulate it, we first need the weak formulation of the Oseen problem. Let $\mathbf{V} := \{\mathbf{v} \in H^1(\Omega)^3 \ : \ \mathbf{v}|_{\Gamma_0} = \mathbf{0}\}$. Given $\mathbf{f} \in \mathbf{V}'$, the problem is to find $\mathbf{u} \in \mathbf{V}$ and $p \in L^2(\Omega)$ such that

$$\begin{aligned} \mathcal{L}(\mathbf{u},p;\mathbf{v},q) &= (\mathbf{f},\mathbf{v})_* + (g,q) \qquad \forall \quad \mathbf{v} \in \mathbf{V}, \ \ q \in L^2(\Omega) \ , \\ \mathcal{L}(\mathbf{u},p;\mathbf{v},q) &:= \alpha(\mathbf{u},\mathbf{v}) + \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + ((\mathbf{w}\cdot\nabla)\ \mathbf{u},\mathbf{v}) \\ &\qquad - (p, \operatorname{div}\ \mathbf{v}) + (q, \operatorname{div}\ \mathbf{u}) \ , \end{aligned} \tag{4}$$

where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product and $(\cdot,\cdot)_*$ is the duality paring for $\mathbf{V}' \times \mathbf{V}$.

We assume T_h to be a collection of tetrahedra, which is a consistent subdivision of Ω satisfying the regularity condition

$$\max_{\tau\in T_h} \operatorname{diam}(\tau)/\rho(\tau) \leqslant C_T, \tag{5}$$

where $\rho(\tau)$ is the diameter of the subscribed ball in the tetrahedron τ. A constant C_T measures the maximum anisotropy ratio for T_h. Further, we denote $h_\tau = \operatorname{diam}(\tau)$ and $h_{\min} = \min_{\tau\in T_h} h_\tau$. Given conforming FE spaces $\mathbb{V}_h \subset \mathbf{V}$ and $\mathbb{Q}_h \subset L^2(\Omega)$, the Galerkin FE discretization of Equation 2 is based on the weak formulation: Find $\{\mathbf{u}_h, p_h\} \in \mathbb{V}_h \times \mathbb{Q}_h$ such that

$$\mathcal{L}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) = (\mathbf{f},\mathbf{v}_h)_* + (g,q_h) \qquad \forall\ \mathbf{v}_h \in \mathbb{V}_h, \ \ q_h \in \mathbb{Q}_h \ . \tag{6}$$

In our experiments, we shall use P2-P1 Taylor–Hood FE pair, which satisfies the LBB compatibility condition for $\mathbb{V}_h$ and $\mathbb{Q}_h$ [1] and hence, ensures well-posedness and full approximation order for the FE linear problem.

A potential source of instabilities in Equation 6 is the presence of dominating convection terms. This necessitates stabilization of the discrete system, if the mesh is not sufficiently fine to resolve all scales in the solution. We consider below one commonly used SUPG stabilization, while more details on the family of SUPG methods can be found in, for example, these studies.[6,20,21] Using Equation 6 as the starting point, a weighted residual for the FE solution multiplied by an "advection"-depended test function is added:

$$\begin{aligned} &\mathcal{L}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) \\ &\quad + \sum_{\tau\in T_h} \sigma_\tau(\alpha\mathbf{u}_h - \nu\Delta\mathbf{u}_h + \mathbf{w}\cdot\nabla\mathbf{u}_h + \nabla p_h - \mathbf{f}, \mathbf{w}\cdot\nabla\mathbf{v}_h)_\tau \\ &= (\mathbf{f},\mathbf{v}_h) \quad \forall\ \mathbf{v}_h \in \mathbb{V}_h, \ \ q_h \in \mathbb{Q}_h \ , \end{aligned} \tag{7}$$

with $(f,g)_\tau \ := \int_\tau fg\ dx$. The second term in Equation 7 is evaluated elementwise for each element $\tau \in T_h$. Parameters σ_τ are element- and problem-dependent. To define the parameters, we introduce mesh Reynolds numbers $\mathrm{Re}_\tau \ := \|\mathbf{w}\|_{L_\infty(\tau)} h_{\mathbf{w}}/\nu$ for all $\tau \in T_h$, where $h_{\mathbf{w}}$ is the diameter of τ in direction $\mathbf{w}$. Several recipes for the particular choice of the stabilization parameters can be found in the literature. When we experiment with the stabilization, we set

$$\sigma_\tau = \begin{cases} \bar{\sigma}\dfrac{h_{\mathbf{w}}}{2\|\mathbf{w}\|_{L_\infty(\tau)}}\left(1 - \dfrac{1}{\mathrm{Re}_\tau}\right), \text{ if } \mathrm{Re}_\tau > 1, \\ \qquad\qquad\qquad 0, \text{ if } \mathrm{Re}_\tau \leqslant 1, \end{cases} \quad \text{with } 0 < \bar{\sigma} < 1. \tag{8}$$

If one enumerates velocity unknowns first and pressure unknowns next, then the resulting discrete system has the 2×2-block form (Equation 3) with $C = 0$. The stabilization alters the $(1, 2)$-block of the matrix making the latter not equal to the transpose of the $(2, 1)$-block B. In this paper, we analyze factorizations for the matrix from (Equation 3) assuming that the perturbation of B^T in the $(1, 2)$-block caused by Equation 7 is relatively small due to the choice of σ_τ. The analysis and results of numerical experiments also show that the perturbation of A caused by Equation 7 affects essentially the properties of LU and ILU decompositions.

We note that there was an intensive development of stabilized and multiscale FE methods for fluid problems over the last decade, see, for example, these studies[4,22] and references in more recent review papers.[5,23] While these methods can be more accurate and less dissipative compared to Equation 7, they add terms to the algebraic system of the same structure and similar algebraic properties as the SUPG method. The streamline diffusion stabilization as in Equation 7 is a standard (and often the only available) option in many existing computational fluid dynamics software, so we decided to consider in the present studies this more classical approach as the particular example leading to the system (Equation 3).

3 | STABILITY OF LU FACTORIZATION

The 2×2-block matrix from (Equation 3) is in general not sign definite, and if $C = 0$, its diagonal has zero entries. The algebraic framework of this section admits a generic positive semidefinite matrix C. An LU factorization of such matrices often requires pivoting (rows and columns permutations) for stability reasons. However, exploiting the block structure and the properties of blocks A and C, one readily verifies that the LU factorization

$$\mathcal{A} = \begin{pmatrix} A & \tilde{B}^T \\ B & -C \end{pmatrix} = \begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ 0 & -U_{22} \end{pmatrix} \tag{9}$$

with lower (upper) triangle matrices L_{11}, L_{22} (U_{11}, U_{22}) exists without pivoting, once $\det(\mathcal{A}) \neq 0$ and there exist LU factorizations for the (1,1)-block

$$A = L_{11} U_{11}, \tag{10}$$

and the Schur complement matrix $\tilde{S} := BA^{-1}\tilde{B}^T + C$ is factorized as

$$\tilde{S} = L_{22} U_{22}. \tag{11}$$

Decomposition (Equation 9) then holds with $U_{12} = L_{11}^{-1}\tilde{B}^T$ and $L_{21} = BU_{11}^{-1}$.

Assume A is positive definite. Then the LU factorization of A exists without pivoting. Its numerical stability (the relative size of entries in factors L_{11} and U_{11}) may depend on how large is the skew-symmetric part of A comparing to the symmetric part. To make this statement more precise, we denote $A_{\mathrm{S}} = \frac{1}{2}(A + A^T)$, $A_{\mathrm{N}} = A - A_{\mathrm{S}}$ (similar notation will be used to denote symmetric and scew-symmetric parts of other matrices) and let

$$C_A = \|A_{\mathrm{S}}^{-\frac{1}{2}} A_{\mathrm{N}} A_{\mathrm{S}}^{-\frac{1}{2}}\|. \tag{12}$$

Here and further, $\|\cdot\|$ and $\|\cdot\|_F$ denote the matrix spectral norm and the Frobenius norm, respectively; and $|M|$ denotes the matrix of absolute values of M-entries. The following bound on the size of elements of L_{11} and U_{11} holds (see Equation 3.2 in this study[19]):

$$\frac{\| |L_{11}| |U_{11}| \|_F}{\|A\|} \leqslant n \left(1 + C_A^2\right). \tag{13}$$

If $C \geqslant 0$, $\tilde{B} = B$ and matrix B^T has the full column rank, then the positive definiteness of A implies that the Schur complement matrix is also positive definite. However, this is not the case for a general block $\tilde{B} \neq B$. In the application studied in this paper, the (1, 2)-block $\tilde{B}^T$ is a perturbation of B^T. The analysis below shows that the positive definiteness of $\tilde{S}$ and the stability of its LU factorization are guaranteed if the perturbation $E = \tilde{B} - B$ is not too large. The size of the perturbation will enter our bounds as the parameter ϵ_E defined as

$$\epsilon_E := \|A_{\mathrm{S}}^{-\frac{1}{2}} E^T\|. \tag{14}$$

For the ease of analysis we introduce further notations:

$$S = BA^{-1}B^T + C, \quad \widehat{A}_N = A_{\mathrm{S}}^{-\frac{1}{2}} A_{\mathrm{N}}\, A_{\mathrm{S}}^{-\frac{1}{2}}. \tag{15}$$

We shall repeatedly make use of the following identities:

$$\begin{aligned} (A^{-1})_{\mathrm{S}} &= \frac{1}{2}\left(A^{-1} + A^{-T}\right) = A_{\mathrm{S}}^{-\frac{1}{2}} (I - \widehat{A}_N^2)^{-1} A_{\mathrm{S}}^{-\frac{1}{2}}, \\ (A^{-1})_{\mathrm{N}} &= \frac{1}{2}\left(A^{-1} - A^{-T}\right) = A_{\mathrm{S}}^{-\frac{1}{2}} (I + \widehat{A}_N)^{-1} \widehat{A}_N (I - \widehat{A}_N)^{-1} A_{\mathrm{S}}^{-\frac{1}{2}}. \end{aligned} \tag{16}$$

From the identities

$$\begin{aligned} \langle Sq, q\rangle = \langle Bv, q\rangle + \langle Cq, q\rangle = \langle v, B^T q\rangle \\ + \langle Cq, q\rangle = \langle Av, v\rangle + \langle Cq, q\rangle, \end{aligned} \tag{17}$$

which are true for $q \in \mathbb{R}^m$ and $v := A^{-1}B^T q \in \mathbb{R}^n$, we see that S is positive definite, if A is positive definite. For $\tilde{S}$, we then compute

$$\begin{aligned} \langle \tilde{S}q, q\rangle &= \langle Sq, q\rangle + \langle A^{-1}E^T q, B^T q\rangle \\ &= \langle Sq, q\rangle + \langle A_{\mathrm{S}}^{\frac{1}{2}} A^{-1} E^T q, A_{\mathrm{S}}^{-\frac{1}{2}} B^T q\rangle \\ &= \langle Sq, q\rangle + \langle A_{\mathrm{S}}^{\frac{1}{2}} A^{-1} E^T q, (I - \widehat{A}_N)(I - \widehat{A}_N)^{-1} A_{\mathrm{S}}^{-\frac{1}{2}} B^T q\rangle \\ &= \langle Sq, q\rangle + \langle \left((I + \widehat{A}_N) A_{\mathrm{S}}^{\frac{1}{2}} A^{-1} A_{\mathrm{S}}^{\frac{1}{2}} \right) A_{\mathrm{S}}^{-\frac{1}{2}} E^T q, \\ &\quad (I - \widehat{A}_N)^{-1} A_{\mathrm{S}}^{-\frac{1}{2}} B^T q\rangle. \end{aligned} \tag{18}$$

We employ identities (Equation 16) to get

$$\begin{aligned} (I + \widehat{A}_N) A_{\mathrm{S}}^{\frac{1}{2}} A^{-1} A_{\mathrm{S}}^{\frac{1}{2}} &= (I + \widehat{A}_N) A_{\mathrm{S}}^{\frac{1}{2}} ((A^{-1})_{\mathrm{S}} + (A^{-1})_{\mathrm{N}}) A_{\mathrm{S}}^{\frac{1}{2}} \\ &= (I + \widehat{A}_N)((I - \widehat{A}_N^2)^{-1} \\ &\quad + (I + \widehat{A}_N)^{-1} \widehat{A}_N (I - \widehat{A}_N)^{-1}) \\ &= (I - \widehat{A}_N)^{-1} + \widehat{A}_N (I - \widehat{A}_N)^{-1} \\ &= (I + \widehat{A}_N)(I - \widehat{A}_N)^{-1}. \end{aligned} \tag{19}$$

Noting $\|(I - \widehat{A}_N)^{-1}\| \leqslant 1$ for a skew-symmetric $\widehat{A}_N$, we estimate

$$\begin{aligned}
\langle \tilde{S}q, q\rangle &\geqslant \langle Sq, q\rangle - \|(I+\widehat{A}_N)(I-\widehat{A}_N)^{-1}\|\,\|A_{\mathrm{S}}^{-\frac{1}{2}}E^Tq\|\,\|(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\| \\
&\geqslant \langle Sq, q\rangle - \|(I+\widehat{A}_N)\|\,\|A_{\mathrm{S}}^{-\frac{1}{2}}E^T\|\,\|q\|\,\|(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\| \\
&\geqslant \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\,\|(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\| \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle (I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq, (I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq, (I+\widehat{A}_N)^{-1}(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq, (I-\widehat{A}_N^2)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle B^Tq, A_{\mathrm{S}}^{-\frac{1}{2}}(I-\widehat{A}_N^2)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle B(A^{-1})_{\mathrm{S}}B^Tq, q\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle BA^{-1}B^Tq, q\rangle^{\frac{1}{2}} \\
&= \langle Sq, q\rangle - (1+C_A)\epsilon_E\|q\|\langle Sq, q\rangle^{\frac{1}{2}} \\
&\geqslant \left(1-(1+C_A)\epsilon_E\lambda_{\min}^{-\frac{1}{2}}(S_{\mathrm{S}})\right)\langle Sq, q\rangle.
\end{aligned} \tag{20}$$

Hence, we conclude that $\tilde{S}$ is positive definite if the perturbation matrix E is sufficiently small such that it holds

$$\kappa := (1+C_A)\epsilon_E c_S^{-\frac{1}{2}} < 1, \tag{21}$$

where $c_S := \lambda_{\min}(S_{\mathrm{S}})$.

If $\tilde{S}$ is positive definite, the factorization $\tilde{S} = L_{22}U_{22}$ satisfies the stability bound similar to Equation 13:

$$\frac{\|\,|L_{22}|\,|U_{22}|\,\|_F}{\|\tilde{S}\|} \leqslant m\left(1+\left\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_{\mathrm{N}}\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right\|^2\right), \tag{22}$$

where $\tilde{S}_{\mathrm{S}} = \frac{1}{2}(\tilde{S}+\tilde{S}^T)$, $\tilde{S}_{\mathrm{N}} = \tilde{S} - \tilde{S}_{\mathrm{S}}$.

The quotients $C_A = \|A_{\mathrm{S}}^{-\frac{1}{2}}A_{\mathrm{N}}A_{\mathrm{S}}^{-\frac{1}{2}}\|$ and $\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\|$ are largely responsible for the stability of the LU factorization for Equation 3. The following lemma shows the estimate of $\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\|$ in terms of C_A, ϵ_E, and c_S.

Lemma 3.1. Let $A \in \mathbb{R}^{n\times n}$ be positive definite and (Equation 21) be satisfied, then it holds

$$\left\|\,\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right\| \leqslant \frac{\left(1+\epsilon_E c_S^{-\frac{1}{2}}\right)C_A}{1-\kappa}. \tag{23}$$

Proof. Due to the skew-symmetry of $\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}$, it holds $|\lambda| = |\mathrm{Im}(\lambda)|$ for $\lambda \in \mathrm{sp}(\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}})$, where we use $\mathrm{sp}(\cdot)$ to denote the spectrum. We apply Bendixson's theorem[24] to estimate

$$\begin{aligned}
\left\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\,\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right\| &= \max\left\{|\lambda| \;:\; \lambda \in \mathrm{sp}\left(\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right)\right\} \\
&= \max\left\{|\mathrm{Im}(\lambda)| \;:\; \lambda \in \mathrm{sp}\left(\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right)\right\} \\
&\leqslant \sup_{q\in\mathbb{C}^m}\frac{|\langle \tilde{S}_Nq, q\rangle|}{\langle \tilde{S}_{\mathrm{S}}q, q\rangle}.
\end{aligned} \tag{24}$$

Thanks to Equation 20, we estimate

$$\left\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\,\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right\| \leqslant \sup_{q\in\mathbb{C}^m}\frac{|\langle \tilde{S}_Nq, q\rangle|}{(1-\kappa)\langle S_{\mathrm{S}}q, q\rangle}. \tag{25}$$

Employing identities from Equation 16, we can write

$$\begin{aligned}
S_{\mathrm{S}} &= BA_{\mathrm{S}}^{-\frac{1}{2}}(I-\widehat{A}_N^T)^{-1}(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^T + C, \\
\tilde{S}_N &= BA_{\mathrm{S}}^{-\frac{1}{2}}(I-\widehat{A}_N^T)^{-1}\widehat{A}_N(I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}\tilde{B}^T.
\end{aligned} \tag{26}$$

With the help of the substitution $v_q = (I-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq$ in the right-hand side of Equation 25 and recalling that C is positive semidefinite, we obtain

$$\begin{aligned}
&\left\|\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\tilde{S}_N\,\tilde{S}_{\mathrm{S}}^{-\frac{1}{2}}\right\| \\
&\leqslant \sup_{q\in\mathbb{C}^m}\frac{\left|\langle \widehat{A}_Nv_q, v_q\rangle\right| + \left|\langle \widehat{A}_N(1-\widehat{A}_N)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}E^Tq, v_q\rangle\right|}{(1-\kappa)(\langle v_q, v_q\rangle + \langle Cq, q\rangle)} \\
&\leqslant \sup_{q\in\mathbb{C}^m}\frac{\|\widehat{A}_N\|\,\|v_q\|^2 + \|\widehat{A}_N\|\epsilon_E\|q\|\,\|v_q\|}{(1-\kappa)(\|v_q\|^2 + \langle Cq, q\rangle)} \\
&\leqslant \sup_{q\in\mathbb{C}^m}\frac{\|\widehat{A}_N\|\,\|v_q\|^2 + \|\widehat{A}_N\|\epsilon_E\lambda_{\min}^{-\frac{1}{2}}(S_{\mathrm{S}})\langle S_{\mathrm{S}}q, q\rangle^{\frac{1}{2}}\|v_q\|}{(1-\kappa)(\|v_q\|^2 + \langle Cq, q\rangle)} \\
&= \sup_{q\in\mathbb{C}^m}\frac{\|\widehat{A}_N\|\,\|v_q\|^2 + \|\widehat{A}_N\|\epsilon_E\lambda_{\min}^{-\frac{1}{2}}(S_{\mathrm{S}})(\|v_q\|^2 + \langle Cq, q\rangle)^{\frac{1}{2}}\|v_q\|}{(1-\kappa)(\|v_q\|^2 + \langle Cq, q\rangle)} \\
&\leqslant \frac{(1+\epsilon_E c_S^{-\frac{1}{2}})\|\widehat{A}_N\|}{1-\kappa}.
\end{aligned} \tag{27}$$

□

To estimate the entries of U_{12} and L_{21} factors in Equation 9, we repeat the arguments from this study[19] and arrive at the following bound

$$\frac{\|U_{12}\|_F + \|L_{21}\|_F}{\|U_{11}\|\|\tilde{B}\|_F + \|L_{11}\|\|B\|_F} \leqslant \frac{m(1 + C_A)}{c_A} \tag{28}$$

with $c_A := \lambda_{\min}(A_S)$.

We summarize the results of this section in the following theorem.

Theorem 3.2. Assume matrix A is positive definite, C is positive semidefinite, and the inequality Equation 21 holds with $\epsilon_E = \|A_S^{-\frac{1}{2}}(\tilde{B} - B)^T\|$, $C_A = \|A_S^{-\frac{1}{2}} A_N A_S^{-\frac{1}{2}}\|$, and $c_S = \lambda_{\min}(S_S)$, then the LU factorization for Equation 9 exists without pivoting. The entries of the block factors satisfy the following bounds:

$$\begin{aligned}\frac{\||L_{11}||U_{11}|\|_F}{\|A\|} &\leqslant n\ \left(1 + C_A^2\right),\\ \frac{\||L_{22}||U_{22}|\|_F}{\|\tilde{S}\|} &\leqslant m\left(1 + \frac{(1 + \epsilon_E c_S^{-\frac{1}{2}})C_A}{1 - \kappa}\right),\\ \frac{\|U_{12}\|_F + \|L_{21}\|_F}{\|U_{11}\|\|\tilde{B}\|_F + \|L_{11}\|\|B\|_F} &\leqslant \frac{m(1 + C_A)}{c_A}\end{aligned} \tag{29}$$

with κ from Equation 21.

The above analysis indicates that the LU factorization for Equation 3 exists if the (1, 1)-block A is positive definite and the perturbation of the (1, 2)-block is sufficiently small. The stability bounds depend on the constant C_A, which measures the ratio of skew-symmetry for A, the ellipticity constant c_A, the perturbation measure ϵ_E, and the minimal eigenvalue of the symmetric part of the unperturbed Schur complement matrix S. In Section 4 below, we estimate all these values for the discrete linearized Navier–Stokes system.

4 | PROPERTIES OF MATRICES A AND $\tilde{S}$

In this Section, we deduce the dependence of the critical constants c_A, C_A, ϵ_E, and c_S from Theorem 3.2 on the problem and discretization parameters. This analysis relies on the SUPG-FE formulation from Section 2. Starting from this section, we assume an inf-sup FE method, and so for the (2, 2)-block of Equation 3, we have $C = 0$. Let $\{\varphi_i\}_{1\leqslant i\leqslant n}$ and $\{\psi_j\}_{1\leqslant j\leqslant m}$ be the bases of $\mathbb{V}_h$ and $\mathbb{Q}_h$, respectively. For arbitrary $v \in \mathbb{R}^n$ and corresponding $\mathbf{v}_h = \sum_{i=1}^n v_i\varphi_i$, one gets the following identity from the definition of matrix A:

$$\begin{aligned}\langle Av, v\rangle = \alpha\|\mathbf{v}_h\|^2 + \nu\|\nabla\mathbf{v}_h\|^2 &\\ + \sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{v}_h\|_\tau^2 + \frac{1}{2}\int_{\Gamma_N}(\mathbf{w}\cdot\mathbf{n})|\mathbf{v}_h|^2\,ds &\\ + \frac{1}{2}\sum_{\tau\in T_h}((\mathrm{div}\ \mathbf{w})\mathbf{v}_h, \mathbf{v}_h)_\tau &\\ + \sum_{\tau\in T_h}\sigma_\tau(\alpha\mathbf{v}_h - \nu\Delta\mathbf{v}_h, \mathbf{w}\cdot\nabla\mathbf{v}_h)_\tau, &\end{aligned} \tag{30}$$

where $\mathbf{n}$ is the outward normal on Γ_N. We shall also need the velocity mass and stiffness matrices M and K: $M_{ij} = (\varphi_i, \varphi_j)$, $K_{ij} = (\nabla\varphi_i, \nabla\varphi_j)$ and the pressure mass matrix M_p: $(M_p)_{ij} = (\psi_i, \psi_j)$.

The first three terms on the right-hand side of Equation 30 are positive and contribute to the ellipticity of the block A. However, the remaining three terms are not necessarily sign definite and should be properly bounded. Although a modification of boundary conditions on Γ_N can be done to insure the resulting boundary integral is non-negative, see, for example, this study,[25] we shall use a FE trace inequality to estimate this term. We remark that this term disappears in the case of artificial outflow boundary conditions leading to Dirichlet conditions in Equation 2 on the entire boundary.[26,27] Next, $\mathbf{w}$ is typically an FE velocity field, $\mathbf{w} \in \mathbb{V}_h$, satisfying only weak divergence free constraint $(\mathrm{div}\,\mathbf{w}, q_h) = 0 \ \ \forall \ \ q_h \in \mathbb{Q}_h$. This weak divergence free equation does not imply $\mathrm{div}\mathbf{w} = 0$ pointwise for most of stable FE pairs including P2–P1 elements. Therefore, the fifth term on the right-hand side of Equation 30 should be controlled somehow. The last term in Equation 30 is due to the SUPG stabilization. The ν-dependent part of it vanishes for P1 FE velocities, but not for most of inf-sup stable discretization pressure–velocity pairs. Both analysis and numerical experiments below show that this term may significantly affect the properties of the matrix A, leading to unstable behavior of incomplete LU decomposition unless the stabilization parameters are chosen sufficiently small. We make the above statements more precise in Theorem 4.1. We need some preparation before we formulate the theorem.

First, recall the Sobolev trace inequality

$$\int_{\Gamma_N}|v|^2\,ds \leqslant C_0\|\nabla v\|^2 \ \ \forall\ \ v \in H^1(\Omega),\ \ v = 0\ \text{ on }\ \partial\Omega\setminus\Gamma_N. \tag{31}$$

For any tetrahedron $\tau\in T_h$ and arbitrary $\mathbf{v}_h \in \mathbb{V}_h$, the following FE trace and inverse inequalities hold

$$\begin{aligned}\int_{\partial\tau}\mathbf{v}_h^2\,ds \leqslant C_{\mathrm{tr}}h_\tau^{-1}\|\mathbf{v}_h\|_\tau^2,\ \ \|\nabla\mathbf{v}_h\|_\tau \leqslant C_{\mathrm{in}}h_\tau^{-1}\|\mathbf{v}_h\|_\tau,\ \ \|\Delta\mathbf{v}_h\|_\tau &\\ \leqslant \bar{C}_{\mathrm{in}}h_\tau^{-1}\|\nabla\mathbf{v}_h\|_\tau, &\end{aligned} \tag{32}$$

where the constants C_{tr}, C_{in}, $\bar{C}_{\mathrm{in}}$ depend only on the polynomial degree k and the shape regularity constant C_T from Equation 5. In addition, denote by C_{f} the constant from the Friedrichs inequality:

$$\|\mathbf{v}_h\| \leqslant C_{\mathrm{f}}\|\nabla\mathbf{v}_h\| \quad \forall \quad \mathbf{v}_h \in \mathbb{V}_h, \tag{33}$$

and let $C_{\mathbf{w}} := \|(\mathbf{w}\cdot\mathbf{n})_-\|_{L^\infty(\Gamma_N)}$.

To avoid the repeated use of generic but unspecified constants, in the remainder of the paper, the binary relation $x \lesssim y$ means that there is a constant c such that $x \leqslant c\ y$, and c does not depend on the parameters, which x and y may depend on, for example, ν, α, mesh size, and properties of $\mathbf{w}$. Obviously, $x \gtrsim y$ is defined as $y \lesssim x$.

Theorem 4.1. Assume that $\mathbf{w} \in L^\infty(\Omega)$, problem, and discretization parameters satisfy

$$\begin{cases} C_{\mathbf{w}} C_{\mathrm{tr}} h_{\min}^{-1} \leqslant \frac{\alpha}{4} \ \text{ or } \ C_{\mathbf{w}} C_0 \leqslant \frac{\nu}{4}, \\ \|\mathrm{div}\ \mathbf{w}\|_{L^\infty(\Omega)} \leqslant \frac{1}{4} \max\{\alpha, \nu \overset{-1}{\underset{f}{\mathrm{C}}}\}, \\ \sigma_\tau \leqslant \frac{1}{2}\left(\frac{h_\tau^2}{\nu \bar{C}_{\mathrm{in}}^2} + \frac{\alpha h_\tau^4}{\nu^2 \bar{C}_{\mathrm{in}}^2 C_{\mathrm{in}}^2}\right) \ \text{ and } \ \sigma_\tau \leqslant \frac{h_\tau}{4\|\mathbf{w}\|_{L^\infty(\tau)} C_{\mathrm{in}}} \ \ \forall\ \tau \in T_h, \end{cases} \tag{34}$$

with constants defined in Equations 31 to 33. Then the matrix A is positive definite, and the constants c_A,C_A,c_S, and ϵ_E can be estimated as follows:

$$\begin{aligned} c_A &\geqslant \frac{1}{4} \lambda_{\min}(\alpha M + \nu K), \\ C_A &\lesssim 1 + \frac{\|\mathbf{w}\|_{L^\infty(\Omega)}}{\sqrt{\nu\alpha} + \nu + h_{\min}\alpha}, \\ c_S &\gtrsim \frac{\lambda_{\min}(M_p)}{(\nu + \alpha + \|\mathbf{w}\|_{L^\infty(\Omega)} + \|\mathrm{div}\ \mathbf{w}\|_{L^\infty(\Omega)})(1 + C_A^2)}, \\ \epsilon_E &\leqslant \left(\frac{\bar{\sigma}}{2\nu} \lambda_{\max}(M_p)\right)^{\frac{1}{2}}. \end{aligned} \tag{35}$$

Proof. Using the Cauchy inequality and Equation 32, we bound the ν-dependent part of the last term in Equation 30 as follows:

$$\begin{aligned} &\left| \sum_{\tau \in T_h} \sigma_\tau \nu (\Delta \mathbf{v}_h, \mathbf{w} \cdot \nabla \mathbf{v}_h)_\tau \right| \\ &\leqslant \nu \left(\sum_{\tau \in T_h} \sigma_\tau \bar{C}_{\mathrm{in}}^2 h_\tau^{-2} \|\nabla \mathbf{v}_h\|_\tau^2 \right)^{\frac{1}{2}} \left(\sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 \right)^{\frac{1}{2}} \\ &\leqslant \frac{\nu^2}{2} \sum_{\tau \in T_h} \sigma_\tau \bar{C}_{\mathrm{in}}^2 h_\tau^{-2} \|\nabla \mathbf{v}_h\|_\tau^2 + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 \\ &\leqslant \frac{\nu^2}{2} \bar{C}_{\mathrm{in}}^2 \sum_{\tau \in T_h} \sigma_\tau \frac{\nu \|\nabla \mathbf{v}_h\|_\tau^2 + \alpha \|\mathbf{v}_h\|_\tau^2}{\nu h_\tau^2 + C_{\mathrm{in}}^{-2} \alpha h_\tau^4} + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 \\ &\leqslant \frac{1}{2} \sum_{\tau \in T_h} \frac{\nu^2 \sigma_\tau \bar{C}_{\mathrm{in}}^2 C_{\mathrm{in}}^2}{\nu h_\tau^2 C_{\mathrm{in}}^2 + \alpha h_\tau^4} \left(\nu \|\nabla \mathbf{v}_h\|_\tau^2 + \alpha \|\mathbf{v}_h\|_\tau^2 \right) \\ &\quad + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2. \end{aligned} \tag{36}$$

The first term in the second line of Equation 36 is bounded due to $\min\{\frac{a}{c}; \frac{b}{d}\} \leqslant \frac{a+b}{c+d}$ for $a, b, c, d > 0$. Using similar arguments, we bound the α-dependent part of the last term in Equation 30:

$$\begin{aligned} \left| \sum_{\tau \in T_h} \sigma_\tau \alpha (\mathbf{v}_h, \mathbf{w} \cdot \nabla \mathbf{v}_h)_\tau \right| &\leqslant \sum_{\tau \in T_h} \alpha \sigma_\tau \|\mathbf{w}\|_{L^\infty(\tau)} \|\mathbf{v}_h\|_\tau \|\nabla \mathbf{v}_h\|_\tau \\ &\leqslant \sum_{\tau \in T_h} \alpha \sigma_\tau \|\mathbf{w}\|_{L^\infty(\tau)} C_{\mathrm{in}} h_\tau^{-1} \|\mathbf{v}_h\|_\tau^2. \end{aligned} \tag{37}$$

Applying Equations 31, 36, and 37 in Equation 30, we deduce

$$\begin{aligned} \langle A v, v \rangle &\geqslant \sum_{\tau \in T_h} \left(1 - \frac{\nu^2 \sigma_\tau \bar{C}_{\mathrm{in}}^2 C_{\mathrm{in}}^2}{2(\nu h_\tau^2 C_{\mathrm{in}}^2 + \alpha h_\tau^4)} - \frac{\sigma_\tau \|\mathbf{w}\|_{L^\infty(\tau)} C_{\mathrm{in}}}{h_\tau} \right) \\ &\qquad \left(\nu \|\nabla \mathbf{v}_h\|_\tau^2 + \alpha \|\mathbf{v}_h\|_\tau^2 \right) \\ &\quad + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 - \frac{C_{\mathbf{w}}}{2} \int_{\Gamma_{\mathrm{N}}} |\mathbf{v}_h|^2 \, ds \\ &\quad - \frac{1}{2} \|\mathrm{div}\ \mathbf{w}\|_{L^\infty(\tau)} \|\mathbf{v}_h\|^2 \\ &\geqslant \sum_{\tau \in T_h} \left(1 - \frac{\nu^2 \sigma_\tau \bar{C}_{\mathrm{in}}^2 C_{\mathrm{in}}^2}{2(\nu h_\tau^2 C_{\mathrm{in}}^2 + \alpha h_\tau^4)} - \frac{\sigma_\tau \|\mathbf{w}\|_{L^\infty(\tau)} C_{\mathrm{in}}}{h_\tau} \right) \\ &\qquad \left(\nu \|\nabla \mathbf{v}_h\|_\tau^2 + \alpha \|\mathbf{v}_h\|_\tau^2 \right) \\ &\quad - \frac{C_{\mathbf{w}}}{2} \min\{C_0 \|\nabla \mathbf{v}_h\|^2, C_{\mathrm{tr}} h_{\min}^{-1} \|\mathbf{v}_h\|^2\} \\ &\quad + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 - \frac{1}{2} \|\mathrm{div}\ \mathbf{w}\|_{L^\infty(\tau)} \|\mathbf{v}_h\|^2. \end{aligned} \tag{38}$$

To ensure that the right-hand side is positive, we assume conditions (Equation 34) on problem parameters and coefficients. Employing conditions (Equation 34) in Equation 38, we deduce

$$\begin{aligned} \langle A v, v \rangle &\geqslant \frac{1}{4} \left(\alpha \|\mathbf{v}_h\|^2 + \nu \|\nabla \mathbf{v}_h\|_\tau^2 + \sum_{\tau \in T_h} \sigma_\tau \|\mathbf{w} \cdot \nabla \mathbf{v}_h\|_\tau^2 \right) \\ &\geqslant \frac{1}{4} \left(\alpha \langle M v, v \rangle + \nu \langle K v, v \rangle \ \right) \quad \forall \quad v \in \mathbb{R}^n \ , \end{aligned} \tag{39}$$

therefore, $c_A \geqslant \frac{1}{4} \lambda_{\min}(\alpha M + \nu K)$. Further, we estimate

$$\begin{aligned} C_A \ := \|A_{\mathrm{S}}^{-\frac{1}{2}} A_{\mathrm{N}}\ A_{\mathrm{S}}^{-\frac{1}{2}}\| &= \max\{|\lambda| \ \ : \ \ \lambda \in \mathrm{sp}(\overset{-\frac{1}{2}}{\underset{\mathrm{S}}{\mathrm{A}}} A_{\mathrm{N}} A_{\mathrm{S}}^{-\frac{1}{2}})\} \\ &= \max\{|\lambda| \ \ : \ \ \lambda \in \mathrm{sp}(\overset{-1}{\underset{\mathrm{S}}{\mathrm{A}}} A_{\mathrm{N}})\} \\ &\leqslant \|A_{\mathrm{S}}^{-1} A_{\mathrm{N}}\|_*, \end{aligned} \tag{40}$$

and for $\|\cdot\|_*$, we choose a matrix norm induced by the vector norm $\langle(\alpha M + \nu K)\cdot, \cdot\rangle^{\frac{1}{2}}$. For a given $v \in \mathbb{R}^n$ and $u = A_{\mathrm{S}}^{-1} A_{\mathrm{N}}\ v$, consider their FE counterparts $\mathbf{v}_h, \mathbf{u}_h \in \mathbb{V}_h$. Then, $A_{\mathrm{S}} u = A_{\mathrm{N}} v$ can be written in a FE form as

$$\begin{aligned} &\nu(\nabla \mathbf{u}_h, \nabla \phi_h) \\ &\quad + \alpha(\mathbf{u}_h, \phi_h) + \frac{1}{2} \int_{\Gamma_{\mathrm{N}}} (\mathbf{w} \cdot \mathbf{n}) \mathbf{u}_h \cdot \phi_h \, ds \\ &\quad + \sum_{\tau \in T_h} \sigma_\tau (\mathbf{w} \cdot \nabla \mathbf{u}_h, \mathbf{w} \cdot \nabla \phi_h)_\tau \\ &\quad + \frac{1}{2} \sum_{\tau \in T_h} ((\mathrm{div}\ \mathbf{w}) \mathbf{u}_h, \phi_h)_\tau + \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau [(\alpha \mathbf{u}_h - \nu \Delta \mathbf{u}_h, \mathbf{w} \cdot \nabla \phi_h)_\tau \\ &\quad + (\alpha \phi_h - \nu \Delta \phi_h, \mathbf{w} \cdot \nabla \mathbf{u}_h)_\tau] \\ &= \frac{1}{2} \sum_{\tau \in T_h} (1 + \alpha \sigma_\tau) [(\mathbf{w} \cdot \nabla \mathbf{v}_h, \phi_h)_\tau - (\mathbf{w} \cdot \nabla \phi_h, \mathbf{v}_h)_\tau] \\ &\quad - \frac{1}{2} \sum_{\tau \in T_h} \sigma_\tau \nu [(\Delta \mathbf{v}_h, \mathbf{w} \cdot \nabla \phi_h)_\tau - (\Delta \phi_h, \mathbf{w} \cdot \nabla \mathbf{v}_h)_\tau] \ \ \forall \phi_h \in \mathbb{V}_h. \end{aligned} \tag{41}$$

We set $\boldsymbol{\phi}_h = \mathbf{u}_h$. For the left-hand side of Equation 41, the lower bound Equation 39 holds. To estimate the right-hand side, we apply the Cauchy–Schwarz inequality, the second restriction on σ_τ from Equation 34, and FE inverse inequality:

$$\begin{aligned}
&\sum_{\tau\in T_h}(1+\alpha\sigma_\tau)[(\mathbf{w}\cdot\nabla\mathbf{v}_h,\mathbf{u}_h)_\tau-(\mathbf{w}\cdot\nabla\mathbf{u}_h,\mathbf{v}_h)_\tau]\\
&\leqslant \sum_{\tau\in T_h}(1+\frac{\alpha h_\tau}{\|\mathbf{w}\|_{L^\infty(\tau)}C_{\text{in}}})[(\mathbf{w}\cdot\nabla\mathbf{v}_h,\mathbf{u}_h)_\tau-(\mathbf{w}\cdot\nabla\mathbf{u}_h,\mathbf{v}_h)_\tau]\\
&\leqslant \|\mathbf{w}\|_{L^\infty(\Omega)}(\|\nabla\mathbf{v}_h\|\|\mathbf{u}_h\|+\|\nabla\mathbf{u}_h\|\|\mathbf{v}_h\|)\\
&\quad+\sum_{\tau\in T_h}\frac{\alpha h_\tau}{C_{\text{in}}}(\|\nabla\mathbf{v}_h\|_\tau\|\mathbf{u}_h\|_\tau+\|\nabla\mathbf{u}_h\|_\tau\|\mathbf{v}_h\|_\tau)\\
&\leqslant \|\mathbf{w}\|_{L^\infty(\Omega)}(\|\nabla\mathbf{v}_h\|\|\mathbf{u}_h\|+\|\nabla\mathbf{u}_h\|\|\mathbf{v}_h\|)+\sum_{\tau\in T_h}2\alpha\|\mathbf{v}_h\|_\tau\|\mathbf{u}_h\|_\tau\\
&\leqslant \|\mathbf{w}\|_{L^\infty(\Omega)}(\|\nabla\mathbf{v}_h\|\|\mathbf{u}_h\|+\|\nabla\mathbf{u}_h\|\|\mathbf{v}_h\|)\\
&\quad+32\alpha\|\mathbf{v}_h\|^2+\frac{\alpha}{32}\|\mathbf{u}_h\|^2.
\end{aligned}\tag{42}$$

Further, we estimate terms on the right-hand side by employing Young's, Friedrichs, and FE inverse inequalities. Thus, the product $\|\mathbf{u}_h\|\|\nabla\mathbf{v}_h\|$ one can estimate in three different ways:

$$\begin{aligned}
\|\mathbf{w}\|_{L^\infty(\Omega)}\|\mathbf{u}_h\|\|\nabla\mathbf{v}_h\| &\leqslant \frac{1}{32}\alpha\|\mathbf{u}_h\|^2\\
&\quad+8\|\mathbf{w}\|_{L^\infty(\Omega)}\frac{1}{\alpha\nu}(\nu\|\nabla\mathbf{v}_h\|^2)\\
\|\mathbf{w}\|_{L^\infty(\Omega)}\|\mathbf{u}_h\|\|\nabla\mathbf{v}_h\| &\leqslant \frac{1}{32}\nu\|\nabla\mathbf{u}_h\|^2\\
&\quad+8\|\mathbf{w}\|_{L^\infty(\Omega)}\frac{C_{\text{f}}^2}{\nu^2}(\nu\|\nabla\mathbf{v}_h\|^2)\\
\|\mathbf{w}\|_{L^\infty(\Omega)}\|\mathbf{u}_h\|\|\nabla\mathbf{v}_h\| &\leqslant \frac{1}{32}\alpha\|\mathbf{u}_h\|^2\\
&\quad+8\|\mathbf{w}\|_{L^\infty(\Omega)}\frac{C_{\text{in}}^2}{\alpha^2h_{\text{min}}^2}(\alpha\|\mathbf{v}_h\|^2).
\end{aligned}\tag{43}$$

Combining all three estimates gives

$$\begin{aligned}
&\|\mathbf{w}\|_{L^\infty(\Omega)}\|\nabla\mathbf{v}_h\|\|\mathbf{u}_h\| \leqslant \frac{1}{32}(\nu\|\nabla\mathbf{u}_h\|^2+\alpha\|\mathbf{u}_h\|^2)\\
&+8\|\mathbf{w}\|_{L^\infty(\Omega)}^2\min\left\{\frac{1}{\alpha\nu},\frac{C_{\text{f}}^2}{\nu^2},\frac{C_{\text{in}}^2}{\alpha^2h_{\text{min}}^2}\right\}(\nu\|\nabla\mathbf{v}_h\|^2+\alpha\|\mathbf{v}_h\|^2).
\end{aligned}\tag{44}$$

Using same argument to treat the second term on the right-hand side of Equation 42, we arrive at

$$\begin{aligned}
&\|\mathbf{w}\|_{L^\infty(\Omega)}\|\nabla\mathbf{u}_h\|\|\mathbf{v}_h\| \leqslant \frac{1}{32}(\nu\|\nabla\mathbf{u}_h\|^2+\alpha\|\mathbf{u}_h\|^2)\\
&+8\|\mathbf{w}\|_{L^\infty(\Omega)}^2\min\left\{\frac{1}{\alpha\nu},\frac{C_{\text{f}}^2}{\alpha^2},\frac{C_{\text{f}}^2}{\nu^2}\right\}(\nu\|\nabla\mathbf{v}_h\|^2+\alpha\|\mathbf{v}_h\|^2).
\end{aligned}\tag{45}$$

Hence, we derive using $\min\{a_1,a_2,a_3\} \leqslant 3(a_1^{-1}+a_2^{-1}+a_3^{-1})^{-1}$, the estimate for the first term on the right-hand side of Equation 41

$$\begin{aligned}
&\frac{1}{2}\sum_{\tau\in T_h}(1+\alpha\sigma_\tau)[(\mathbf{w}\cdot\nabla\mathbf{v}_h,\mathbf{u}_h)_\tau-(\mathbf{w}\cdot\nabla\mathbf{u}_h,\mathbf{v}_h)_\tau]\\
&\lesssim\left(1+\frac{\|\mathbf{w}\|_{L^\infty(\Omega)}^2}{\nu\alpha+\nu^2+h_{\text{min}}^2\alpha^2}\right)(\nu\|\nabla\mathbf{v}_h\|^2+\alpha\|\mathbf{v}_h\|^2)\\
&\quad+\frac{3}{32}(\nu\|\nabla\mathbf{u}_h\|^2+\alpha\|\mathbf{u}_h\|^2).
\end{aligned}\tag{46}$$

Now, we estimate the second term on the right-hand side of Equation 41 with the help of the third condition from Equation 34:

$$\begin{aligned}
&\sum_{\tau\in T_h}\sigma_\tau\nu[(\Delta\mathbf{v}_h,\mathbf{w}\cdot\nabla\mathbf{u}_h)_\tau-(\Delta\mathbf{u}_h,\mathbf{w}\cdot\nabla\mathbf{v}_h)_\tau]\\
&\leqslant\sum_{\tau\in T_h}[\sigma_\tau\nu\bar{C}_{\text{in}}h_\tau^{-1}\|\nabla\mathbf{v}_h\|_\tau\|\mathbf{w}\cdot\nabla\mathbf{u}_h\|_\tau\\
&\quad+\sigma_\tau\nu\bar{C}_{\text{in}}\|\mathbf{w}\|_{L^\infty(\tau)}h_\tau^{-1}\|\nabla\mathbf{u}_h\|\|\nabla\mathbf{v}_h\|_\tau]\\
&\leqslant\frac{1}{32}(\nu\|\nabla\mathbf{u}_h\|^2+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{u}_h\|_\tau^2)\\
&\quad+\sum_{\tau\in T_h}8(\sigma_\tau\nu\bar{C}_{\text{in}}^2h_\tau^{-2}+\sigma_\tau^2\bar{C}_{\text{in}}^2\|\mathbf{w}\|_{L^\infty(\tau)}^2h_\tau^{-2})\nu\|\nabla\mathbf{v}_h\|_\tau^2\\
&\lesssim\frac{1}{32}(\nu\|\nabla\mathbf{u}_h\|^2+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{u}_h\|_\tau^2)+(\nu\|\nabla\mathbf{v}_h\|^2+\alpha\|\mathbf{v}_h\|^2).
\end{aligned}\tag{47}$$

Summarizing Equations 41 to 47, we obtain

$$\begin{aligned}
&\frac{7}{8}\left(\alpha\|\mathbf{u}_h\|^2+\nu\|\nabla\mathbf{u}_h\|^2+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{u}_h\|_\tau^2\right)\\
&\quad+\frac{1}{2}\int_{\Gamma_{\text{N}}}(\mathbf{w}\cdot\mathbf{n})|\mathbf{u}_h|^2\,ds\\
&\quad-\sum_{\tau\in T_h}\sigma_\tau(\alpha\mathbf{u}_h-\nu\Delta\mathbf{u}_h,\mathbf{w}\cdot\nabla\mathbf{u}_h)_\tau+\frac{1}{2}\sum_{\tau\in T_h}((\text{div }\mathbf{w})\mathbf{u}_h,\mathbf{u}_h)_\tau\\
&\lesssim\left(1+\frac{\|\mathbf{w}\|_{L^\infty(\Omega)}^2}{\nu\alpha+\nu^2+h_{\text{min}}^2\alpha^2}\right)(\nu\|\nabla\mathbf{v}_h\|^2+\alpha\|\mathbf{v}_h\|^2).
\end{aligned}\tag{48}$$

The left-hand side of Equation 41 equals

$$\langle A_{\text{S}}u,u\rangle-\frac{1}{8}\left(\alpha\|\mathbf{u}_h\|^2+\nu\|\nabla\mathbf{u}_h\|^2+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{v}_h\|_\tau^2\right),\tag{49}$$

and due to Equation 39, it is estimated from below by $\frac{1}{2}\langle A_{\text{S}}u,u\rangle$. Recalling $4\langle A_{\text{S}}u,u\rangle \geqslant \|u\|_*^2=\nu\|\nabla\mathbf{u}_h\|^2+\alpha\|\mathbf{u}_h\|^2$, we obtain with the help of Equation 40

$$\begin{aligned}
C_A\leqslant\|A_{\text{S}}^{-1}A_{\text{N}}\|_*&=\sup_{v\in\mathbb{R}^n}\frac{\|u\|_*}{\|v\|_*}\leqslant 2\sup_{v\in\mathbb{R}^n}\frac{\langle A_{\text{S}}u,u\rangle^{\frac{1}{2}}}{\|v\|_*}\\
&\lesssim\left(1+\frac{\|\mathbf{w}\|_{L^\infty(\Omega)}}{\sqrt{\nu\alpha}+\nu+h_{\text{min}}\alpha}\right).
\end{aligned}\tag{50}$$

Denote $\tilde{c}_{\mathbf{w}} := \|\mathbf{w}\|_{L^\infty(\Omega)}$, $\hat{c}_{\mathbf{w}}=\|\text{div }\mathbf{w}\|_{L^\infty(\Omega)}$. To bound from below the ellipticity constant c_{S} for the auxiliary Schur complement matrix S, we first observe the following upper bound

$$\begin{aligned}\langle A_{\mathrm{S}}v,v\rangle &= \langle Av,v\rangle \leqslant 2(\alpha\|\mathbf{v}_h\|^2+\nu\|\nabla\mathbf{v}_h\|^2\\ &\quad+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\cdot\nabla\mathbf{v}_h\|_\tau^2)+C_0\tilde{c}_{\mathbf{w}}\|\nabla\mathbf{v}_h\|^2+\frac{1}{2}\hat{c}_{\mathbf{w}}\|\mathbf{v}_h\|^2\\ &\quad\leqslant 2(\alpha\|\mathbf{v}_h\|^2+\nu\|\nabla\mathbf{v}_h\|^2+\sum_{\tau\in T_h}\sigma_\tau\|\mathbf{w}\|^2_{L^\infty(\tau)}\|\nabla\mathbf{v}_h\|_\tau^2)\\ &\quad+C_0\tilde{c}_{\mathbf{w}}\|\nabla\mathbf{v}_h\|^2+\frac{1}{2}\hat{c}_{\mathbf{w}}\|\mathbf{v}_h\|^2\\ &\quad\leqslant 2(\alpha\|\mathbf{v}_h\|^2+\nu\|\nabla\mathbf{v}_h\|^2+\sum_{\tau\in T_h}\frac{h_\tau\|\mathbf{w}\|_{L^\infty(\tau)}}{4C_{\mathrm{in}}}\|\nabla\mathbf{v}_h\|_\tau^2)\\ &\quad+C_0\tilde{c}_{\mathbf{w}}\|\nabla\mathbf{v}_h\|^2+\frac{1}{2}\hat{c}_{\mathbf{w}}\|\mathbf{v}_h\|^2\leqslant 2(\alpha\|\mathbf{v}_h\|^2\\ &\quad+(\nu+\tilde{c}_{\mathbf{w}})\|\nabla\mathbf{v}_h\|^2)+C_0\tilde{c}_{\mathbf{w}}\|\nabla\mathbf{v}_h\|^2+\frac{1}{2}\hat{c}_{\mathbf{w}}\|\mathbf{v}_h\|^2\\ &\quad\lesssim(\nu+\alpha+\tilde{c}_{\mathbf{w}}+\hat{c}_{\mathbf{w}})\|\nabla\mathbf{v}_h\|^2.\end{aligned}\tag{51}$$

The above bound and the inf-sup stability of the FE spaces yield the following relations:

$$\begin{aligned}\langle BA_{\mathrm{S}}^{-1}B^Tq,q\rangle&=\sup_{v\in\mathbb{R}^n}\frac{\langle Bv,q\rangle^2}{\langle A_{\mathrm{S}}v,v\rangle}\gtrsim\sup_{\mathbf{v}_h\in\mathbb{V}_h}\frac{(\operatorname{div}\mathbf{v}_h,q_h)^2}{(\nu+\alpha+\tilde{c}_{\mathbf{w}}+\hat{c}_{\mathbf{w}})\|\nabla\mathbf{v}_h\|^2}\\ &\gtrsim\frac{\|q_h\|^2}{\nu+\alpha+\tilde{c}_{\mathbf{w}}+\hat{c}_{\mathbf{w}}}=\frac{\langle M_pq,q\rangle}{\nu+\alpha+\tilde{c}_{\mathbf{w}}+\hat{c}_{\mathbf{w}}}.\end{aligned}\tag{52}$$

With the help of the first identity from Equation 16 and using $C=0$ and Equation 52, we obtain

$$\begin{aligned}\langle Sq,q\rangle&=\langle A^{-1}B^Tq,B^Tq\rangle\\ &=\langle(I-(A_{\mathrm{S}}^{-\frac{1}{2}}A_{\mathrm{N}}\,A_{\mathrm{S}}^{-\frac{1}{2}})^2)^{-1}A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq,A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle\\ &\geqslant\frac{\langle A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq,A_{\mathrm{S}}^{-\frac{1}{2}}B^Tq\rangle}{1+\|(A_{\mathrm{S}}^{-\frac{1}{2}}A_{\mathrm{N}}\,A_{\mathrm{S}}^{-\frac{1}{2}})^2\|}=\frac{\langle BA_{\mathrm{S}}^{-1}B^Tq,q\rangle}{1+\|(A_{\mathrm{S}}^{-\frac{1}{2}}A_{\mathrm{N}}\,A_{\mathrm{S}}^{-\frac{1}{2}})^2\|}\\ &\gtrsim\frac{1}{(\nu+\alpha+\tilde{c}_{\mathbf{w}}+\hat{c}_{\mathbf{w}})(1+\|(A_{\mathrm{S}}^{-\frac{1}{2}}A_{\mathrm{N}}\,A_{\mathrm{S}}^{-\frac{1}{2}})\|^2)}\langle M_pq,q\rangle.\end{aligned}\tag{53}$$

The desired bound for c_S follows from Equation 53.

To estimate ϵ_E, we use similar technique. For arbitrary given $q\in\mathbb{R}^m$, let $u=A_{\mathrm{S}}^{-1}E^Tq$. We have

$$\|A_{\mathrm{S}}^{-\frac{1}{2}}E^Tq\|^2=\langle A_{\mathrm{S}}^{-1}E^Tq,E^Tq\rangle=\langle A_{\mathrm{S}}u,u\rangle.\tag{54}$$

For arbitrary $v\in\mathbb{R}^n$, it holds $\langle A_{\mathrm{S}}u,v\rangle=\langle E^Tq,v\rangle$. For corresponding FE functions, this yields

$$\begin{aligned}&\nu(\nabla\mathbf{u}_h,\nabla\mathbf{v}_h)\\ &\quad+\alpha(\mathbf{u}_h,\mathbf{v}_h)+\frac{1}{2}\int_{\Gamma_{\mathrm{N}}}(\mathbf{w}\cdot\mathbf{n})\mathbf{u}_h\cdot\mathbf{v}_h\,ds\\ &\quad+\sum_{\tau\in T_h}\sigma_\tau(\mathbf{w}\cdot\nabla\mathbf{u}_h,\mathbf{w}\cdot\nabla\mathbf{v}_h)_\tau+\frac{1}{2}\sum_{\tau\in T_h}((\operatorname{div}\mathbf{w})\mathbf{u}_h,\mathbf{v}_h)_\tau\\ &\quad+\frac{1}{2}\sum_{\tau\in T_h}\sigma_\tau[(\alpha\mathbf{u}_h-\nu\Delta\mathbf{u}_h,\mathbf{w}\cdot\nabla\mathbf{v}_h)_\tau\\ &\quad+(\alpha\mathbf{v}_h-\nu\Delta\mathbf{v}_h,\mathbf{w}\cdot\nabla\mathbf{u}_h)_\tau]=\sum_{\tau\in T_h}\sigma_\tau(\mathbf{w}\cdot\nabla\mathbf{v}_h,\nabla q_h)_\tau\\ &\quad\leqslant\sum_{\tau\in T_h}\sigma_\tau\left(\frac{1}{8}\|\mathbf{w}\cdot\nabla\mathbf{v}_h\|_\tau^2+2\|\nabla q_h\|_\tau^2\right)\\ &\quad\leqslant\sum_{\tau\in T_h}\sigma_\tau\left(\frac{1}{8}\|\mathbf{w}\cdot\nabla\mathbf{v}_h\|_\tau^2+2C_{\mathrm{in}}^2h_\tau^{-2}\|q_h\|_\tau^2\right).\end{aligned}\tag{55}$$

We set $\mathbf{v}_h=\mathbf{u}_h$ and invoke Equation 39 to conclude in the vector notation

$$\langle A_{\mathrm{S}}u,u\rangle\lesssim\max_\tau(\sigma_\tau h_\tau^{-2})\lambda_{\max}(M_p)\|q\|^2\leqslant\frac{\bar{\sigma}}{2\nu}\lambda_{\max}(M_p)\|q\|^2.\tag{56}$$

The last inequality follows from the definition of σ_τ in Equation 8 for $\mathrm{Re}_\tau>1$:

$$\begin{aligned}\sigma_\tau&=\bar{\sigma}\frac{h_{\mathbf{w}}}{2\|\mathbf{w}\|_{L_\infty(\tau)}}\left(1-\frac{1}{\mathrm{Re}_\tau}\right)\leqslant\bar{\sigma}\frac{h_{\mathbf{w}}}{2\|\mathbf{w}\|_{L_\infty(\tau)}}\mathrm{Re}_\tau\\ &=\bar{\sigma}\frac{h_{\mathbf{w}}^2}{2\nu}\leqslant\bar{\sigma}\frac{h_\tau^2}{2\nu}.\end{aligned}\tag{57}$$

Recalling the definition of ϵ_E, the inequality Equation 56 together with Equation 54 proves the last bound in Equation 35. □

The theorem shows that matrices A and $\tilde{S}$ are positive definite if conditions (Equation 34) on the parameters of the FE method are satisfied. In this case, the matrix in Equation 3 admits LU factorization without pivoting. The *first condition* in (34) is trivially satisfied with $C_{\mathbf{w}}=0$ if $\Gamma_{\mathrm{N}}\neq\emptyset$ or the entire Γ_{N} is outflow boundary. The second condition may not be restrictive, since $\mathbf{w}$ approximates velocity field of an incompressible fluid, and hence, $\|\operatorname{div}\mathbf{w}\|_{L^\infty(\Omega)}$ decreases for a refined grid. However, the $\mathbf{w}$-divergence norm depends on fluid velocity field and may be large for ν small enough. Fortunately, one can choose such small Δt that the second condition holds due to $\alpha\sim(\Delta t)^{-1}$. The third condition in Equation 34 puts an upper bound on stabilization parameters. Naturally, the same or a similar condition appears in the literature on the analysis of SUPG-stabilized methods for the linearized Navier–Stokes equations, see, for example, this study.[21] The reason is that the positive definiteness of A is equivalent to the coercivity of the velocity part of the bilinear form from Equation 7, which is crucial for deriving FE method error estimates. Therefore, stabilization parameter design suggested in the literature typically satisfies $\sigma_\tau\lesssim\frac{h_\tau^2}{\nu}$ and $\sigma_\tau\lesssim\frac{h_\tau}{\|\mathbf{w}\|_{L^\infty(\tau)}}$ asymptotically, that is, up to a scaling factor independent of discretization parameters. As follows from Equation 57, the conditions Equation 34 on the SUPG stabilization parameters Equation 8 are valid if $\bar{\sigma}\leqslant\min\{\bar{C}_{\mathrm{in}}^{-2},\ \frac{1}{2}C_{\mathrm{in}}^{-1}\}$. In practice, however, larger values of $\bar{\sigma}$ are often found optimal for FE solution accuracy. The possible reason of the inconsistency is that smooth harmonics dominate in the solution, and hence, the bounds on parameters are less tight. The situation is different when one is concerned with iterative convergence of algebraic solvers, since an algebraic solver has to reduce all possible harmonics in the decomposition of the error vector.

To get an idea about the values of the critical constants c_A, C_A, and c_S in practice and to illustrate their dependence on flow and discretization parameters, we compute these constants for a series of matrices of the FE Oseen problem. We consider a 3D analogue to the Taylor vortex problem suggested in this study[28] for the purpose of benchmarking. This

TABLE 1 Computed values of c_A, C_A, and c_S as well as the estimated values of right-hand side expressions in Equation 35 for varying α and ν

ν	α	c_A	C_A	$C_A^{\rm est}$	c_S	$c_S^{\rm est}$
.1	10	0.0090	1.0627	1.58	6.4455e-04	7.96e-04
.1	100	0.0223	0.3278	1.08	1.8526e-04	9.17e-05
.01	10	0.0021	3.7766	1.80	0.0014	8.70e-04
.01	100	0.0056	0.8133	1.09	2.0927e-04	9.23e-05
.001	10	5.8437e-04	8.1484	1.91	0.0015	9.03e-04
.001	100	0.0024	1.4904	1.09	2.1458e-04	9.24e-05

TABLE 2 Computed values of c_A, C_A, and c_S for varying stabilization scaling parameter $\bar{\sigma}$

$\bar{\sigma}$	c_A	C_A	c_S	κ
0	0.0024	1.4908	2.1458e-04	0
1/144	0.0024	1.4699	2.1451e-04	0.9362
1/96	0.0024	1.4603	2.1447e-04	1.3951
1/12	0.0025	1.3274	2.1367e-04	10.1030
1/3	0.0027	1.4161	2.1038e-04	50.3507

flow has no principle direction and shows a nontrivial vortical structure. Applying Taylor–Hood elements on a regular tetrahedral subdivision of the unit cube leads to a discrete problem with 11,802 velocity and 596 pressure unknowns. First, we experiment with nonstabilized FE method and vary ν and α. The results are reported in Table 1. We also show the computed values of the quantities $C_A^{\rm est} = 1 + \frac{\|\mathbf{w}\|_{L^\infty(\Omega)}}{\sqrt{\nu\alpha+\nu+h_{\min}\alpha}}$ and $c_S^{\rm est} = \frac{\lambda_{\min}(M_p)}{(\nu+\alpha+\|\mathbf{w}\|_{L^\infty(\Omega)}+\|\text{div }\mathbf{w}\|_{L^\infty(\Omega)})(1+C_A^2)}$ appearing in the bounds (Equation 35). We observe a very good agreement of these bounds with the computed values of C_A and c_S. We recall that the signs "$\gtrsim$" and "$\lesssim$" in Equation 35 denote upper and lower estimates up to a generic constant independent of all relevant parameters. In Table 2, we show c_A, C_A, and c_S for the same discrete problem, but now, with SUPG stabilization. The table also reports the parameter κ from Equation 21. The results indicate that the sufficient condition $\kappa < 1$ can be too pessimistic in practice, and stable factorization is done without pivoting for certain values of κ greater than 1.

5 | A TWO-PARAMETER THRESHOLD ILU FACTORIZATION

Incomplete LU factorizations of Equation 3 can be written in the form $A = LU - E$ with an error matrix E. How small is the matrix E can be ruled by the choice of a threshold parameter $\tau > 0$. The error matrix E is responsible for the quality of preconditioning, see, for example, this study[29] for estimates on generalized minimal residual method convergence written in terms of $\|E\|$ and subject to a proper prescaling of A and the diagonalizability assumption. In general, the analysis of ILU factorization is based on the following arguments. For positive definite matrices A, one can choose such a small τ that the product LU of its incomplete triangular factors L and U is also positive definite, and so, estimates from this study[30] can be applied to assess the numerical stability of the incomplete factorization: for $c_A = \lambda_{\min}(A_S)$, the sufficient condition is $\tau < c_A n^{-1}$. In practice, however, larger τ is used, and in the case of nonsymmetric matrices, nonpositive or close to zero pivots may encounter, and breakdown of an algorithm may happen. Although most of remedies were developed for the self-adjoint positive definite case,[31] some of them are applicable to nonsymmetric and nondefinite matrices. We use the matrix two-side scaling[19] in our applications.

Stability of ILU factorization for saddle point matrices with positive definite (1, 1)-block and $\tilde{B} \neq B$ deteriorates in comparison with positive definite matrices and saddle point matrices with $\tilde{B} = B$. Theorem 4.1 shows that for certain flow regimes, the ellipticity constants c_A and c_S for A and S approach zero. To ameliorate the performance of the preconditioning in such extreme situations, we consider the two-parameter Tismenetsky–Kaporin variant of the threshold ILU factorization. The factorization was introduced and first studied in this study[32–34] for symmetric positive definite matrices and recently for nonsymmetric matrices in this study.[19]

Given a matrix $A \in \mathbb{R}^{n \times n}$, the two-parameter factorization can be written as

$$A = LU + LR_u + R_\ell U - E, \tag{58}$$

where R_u and R_ℓ are strictly upper and lower triangular matrices, while U and L are upper and lower triangular matrices, respectively. Given two small parameters $0 < \tau_2 \leqslant \tau_1$, the off-diagonal elements of U and L are either zero or have absolute values greater than τ_1; the absolute values of R_ℓ and R_u entries are either zero or belong to $(\tau_2, \tau_1]$; entries of the error matrix are of order $O(\tau_2)$. We refer to Equation 58 as the ILU(τ_1, τ_2) factorization of A. Of course, a generic ILU(τ) factorization can be viewed as Equation 58 with $R_u = R_\ell = 0$ and $\tau_1 = \tau_2 = \tau$. The two-parameter ILU factorization goes over a generic ILU(τ) factorization: the fill-in of L and U is ruled by the first threshold parameter τ_1, while the quality of the resulting preconditioner is mainly defined by τ_2, once $\tau_1^2 \lesssim \tau_2$ holds. In other words, the choice $\tau_2 = \tau_1^2 := \tau^2$ may provide the fill-in of ILU(τ_1, τ_2) to be similar to that of ILU(τ), while the convergence of preconditioned Krylov subspace method is better and asymptotically (for $\tau \to 0$) can be comparable to the one with ILU(τ^2) preconditioner. For symmetric positive definite matrices, these empirical advantages of ILU(τ_1, τ_2) are rigorously explained in this study,[34] where estimates on the eigenvalues and K-condition number of $L^{-1}AU^{-1}$ were derived with $L^T = U$ and $R_\ell^T = R_u$. The price one pays is that computing L, U factors for ILU(τ_1, τ_2) is computationally more costly than for ILU(τ_1), since intermediate calculations involve the entries of R_u. However, this factorization phase of ILU(τ_1, τ_2) is still less expensive than that of ILU(τ_2). We note also that ILU(τ_1, τ_2) with $\tau_1 = \tau_2$ is similar to the ILUT(p, τ) dual parameter incomplete factorization[35] with $p = n$ (all elements passing the threshold criterion are kept in the factors). If no small pivot modifi-

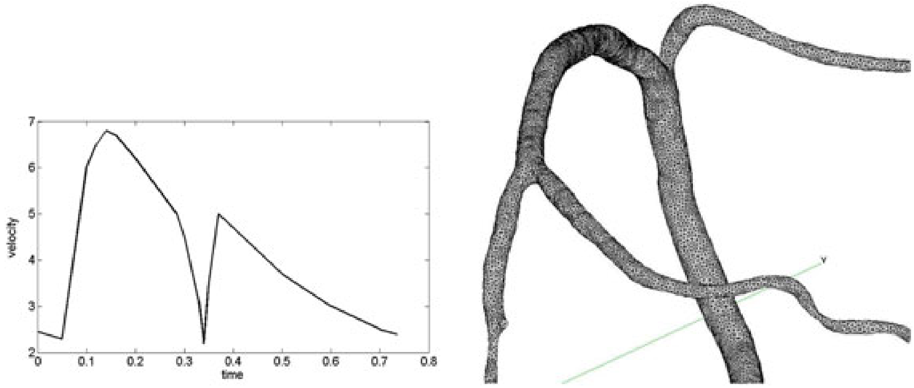

FIGURE 1 The velocity waveform on the inflow as a function of time and the coarse grid in the right coronary artery

TABLE 3 The performance of ILU($\tau_1 = 0.03, \tau_2 = 7\tau_1^2$) for right coronary artery. The number and the time of iterations accumulated for 147 time steps

Mesh	$\bar{\sigma}$	#it	T_{it}
63k	0	20908	2267.
63k	1/12	20292	2182.
120k	0	26209	6188.
120k	1/12	26446	6132.

cation is done, the only differences between the algorithms (for $\tau_1 = \tau_2$ and $p = n$) are a different scaling of pivots and a row dependent scaling of threshold values used in ILUT. A pseudocode of a row-wise ILU(τ_1, τ_2) can be found in this study.[19]

Analysis of the decomposition (Equation 58) of a general nonsymmetric matrix is limited to simple estimate (2.5) from this study[36] applied to the matrix $(L + R_\ell)(U + R_u) = A + R_\ell R_u + E$. The low bound for the pivots of the Equation 58 factorization is the following:

$$|L_{ii}U_{ii}| \geqslant \min_{v\in\mathbb{R}^n} \frac{\langle (A + R_\ell R_u + E)v, v\rangle}{\|v\|^2} \geqslant c_A - \|R_\ell R_u\| - \|E\|, \tag{59}$$

with the ellipticity constant c_A and the norms $\|R_\ell R_u\|$, $\|E\|$ proportional to τ_1^2 and τ_2, respectively. Hence, we may conclude that the numerical stability of solving for $L^{-1}x$ and $U^{-1}x$ is ruled by the second parameter and the square of the first parameter, while the fill-in in both factors is defined by τ_1 rather than τ_1^2. The Oseen problem setup may be such that the estimates from Theorem 4.1 predict that the coercitivity constant c_A and the ellipticity constant c_S are small. This increases the probability of the breakdown of ILU(τ) factorization of the saddle-point matrix, and demonstrates the benefits of ILU(τ_1, τ_2) factorization.

6 | NUMERICAL RESULTS

In this section, we demonstrate the performance of the ILU(τ) factorization for different values of discretization, stabilization, and threshold parameters. As a testbench, we simulate a blood flow in a right coronary artery within a single cardiac cycle. For numerical test, we use the implementation of ILU(τ_1,τ_2) available in the open source software.[37,38] The optimal values of ILU thresholds $\tau_1 = 0.03$, $\tau_2 = 7\tau_1^2$ are taken from this study[19] where detailed analysis of ILU(τ_1,τ_2) and ILU(τ): = ILU(τ,τ) preconditioners for the Oseen systems without stabilization is given. In all experiments, we use biconjugate gradient stabilized (BiCGstab) method with the right preconditioner defined by the ILU(τ_1,τ_2) factorization.

The geometry of the flow domain was recovered from a real patient coronary computed tomography angiography. The diameter of the inlet cross section is about 0.27 cm and is imbedded in the box 6.5cm × 6.8cm × 5cm. The ANI3D package[38] was used to generate two tetrahedral meshes; the coarse mesh is shown in Figure 1. The meshes consist of 63k and 120k tetrahedra leading to the discrete (P2-P1 FEM) Navier–Stokes system with about 300k and 600k unknowns, respectively. The Navier–Stokes system (Equation 1) was integrated in time using a semi-implicit second-order method with $\Delta t = 0.005$ and systems (Equation 3) were solved at every time step. Other model parameters are $\nu = 0.04\text{cm}^2/\text{s}$ and $\rho = 1\text{g/cm}$; one cardiac cycle period was 0.735s. The inlet velocity waveform[39] shown in Figure 1 defines the Poiseuille flow rate through the inflow cross section. The vessel walls were treated as rigid, and homogeneous Dirichlet boundary conditions for the velocity are imposed on the vessel walls. On all outflow boundaries, we set the normal component of the stress tensor equal zero.

Table 3 shows the total number of the preconditioned BiCGstab iterations, and the CPU time needed to perform all 147 time steps within a single cardial cycle. For each system, the initial residual due to the solution from the previous time step is reduced by 10 orders of magnitude. We generated sequences of the discrete Oseen problems (Equation 2) with ($\bar{\sigma} = 1/12$) and without ($\bar{\sigma} = 0$) SUPG stabilization. The choice of parameters τ_1, τ_2 leads to stable computations over the whole cardiac cycle. The total number of iterations depends on the size of the system and the mesh and appears to be very similar for both examples with and without stabilization. The total number of iterations is 20% larger for the fine grid, which should be expected for the preconditioner based on an incomplete factorization. Over the cardiac cycle, the variations of the iteration counts and CPU times per linear solve are rather modest, see the top and bottom plots in Figures 2 and 3. The difference in otherwise similar performance of liner solvers for the cases $\bar{\sigma} = 1/12$ and $\bar{\sigma} = 0$ is the following: For $\bar{\sigma} = 1/12$, when the maximum flow rate on the inlet is achieved, and the number of iterations and times needed to build preconditioner increase significantly (approximately twice as much as average). This happens over a few time steps. In these cases, when factorization is performed, several small pivots occur, and their modification is performed during the incomplete factorization.

The next series of experiments shows that restrictions (Equation 34) on σ_τ are important in practice. According to Theorems 3.2 and 4.1, exact LU factorization of without pivoting is stable if σ_τ is small enough. In particular, according to estimate (Equation 57), sufficient conditions

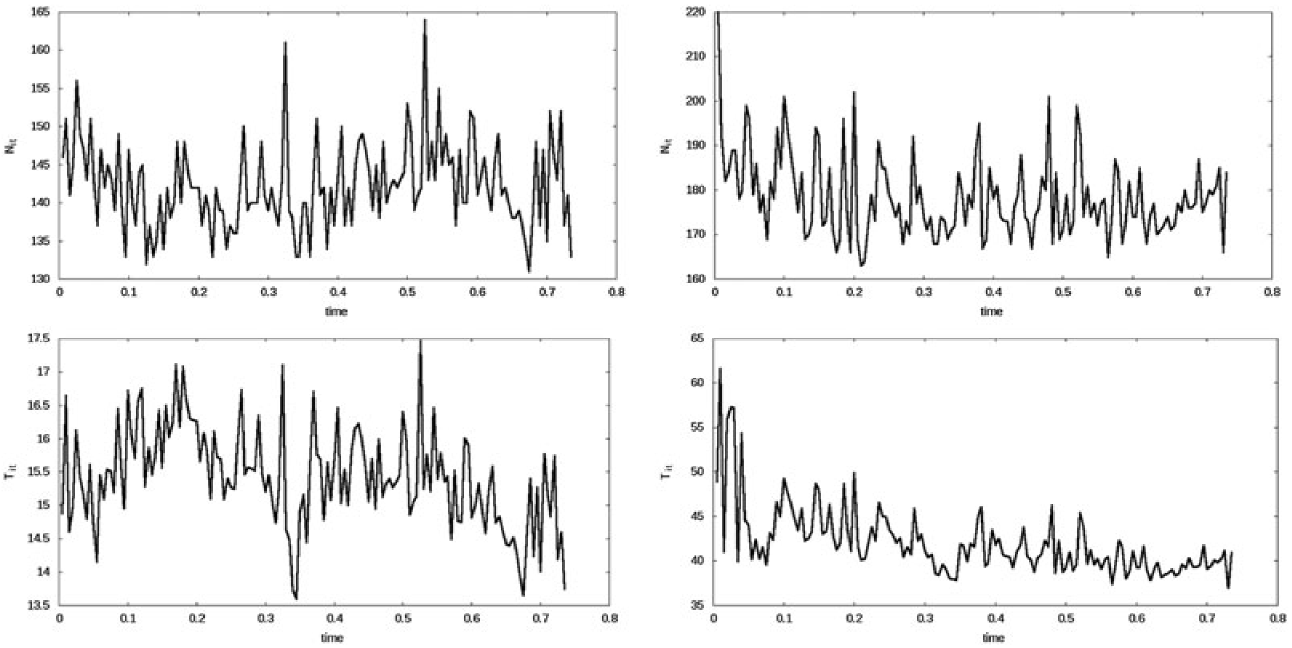

FIGURE 2 Right coronary artery, computations on grid 63k (left) and grid 120k (right) without streamline upwind Petrov–Galerkin stabilization and $\tau_1 = 0.03$: The top plots show the number of biconjugate gradient stabilized (BiCGstab) iterations, the bottom plots show the time of BiCGstab iterations at each time step

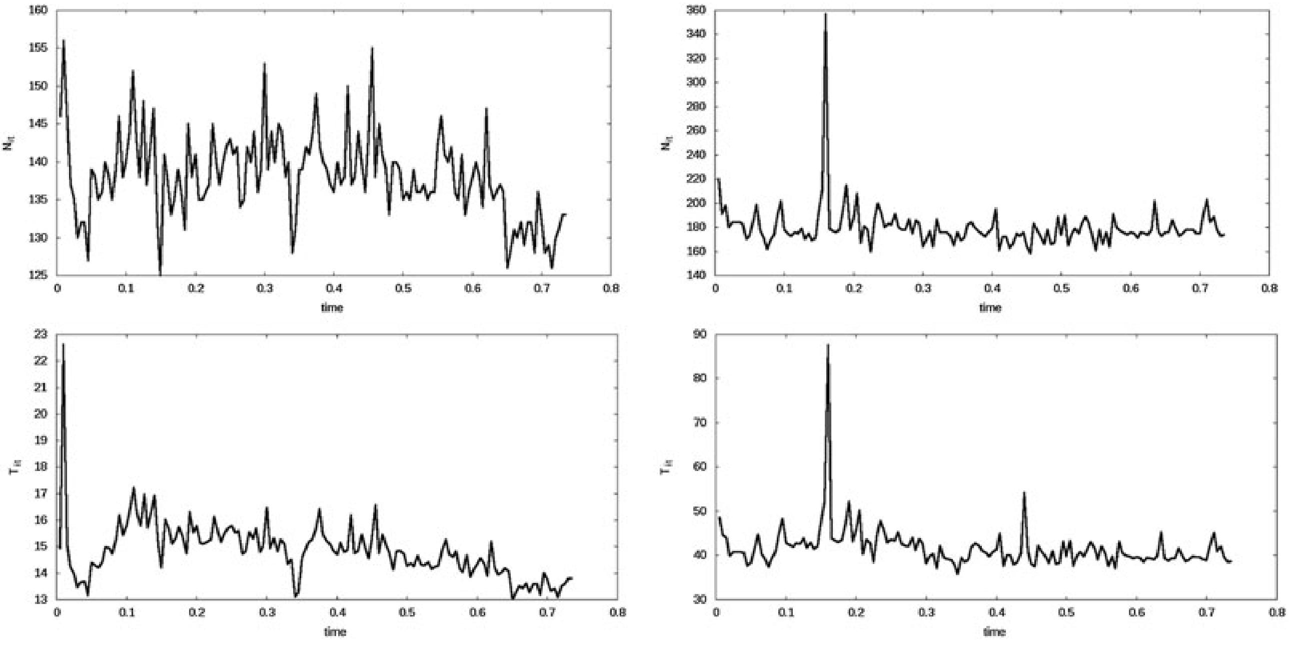

FIGURE 3 Right coronary artery, computations on grid 63k (left) and grid 120k (right), streamline upwind Petrov–Galerkin stabilization with $\bar{\sigma} = 1/12$ and $\tau_1 = 0.03$: The top plots show the number of biconjugate gradient stabilized (BiCGstab) iterations, the bottom plots show the time of BiCGstab iterations at each time step

(Equation 34) are satisfied by parameters from Equation 8 if $\bar{\sigma} \leqslant \min\{\bar{C}_{\rm in}^{-2},\ \frac{1}{2}C_{\rm in}^{-1}\}$. In this experiment, we increase $\bar{\sigma}$ two times setting $\bar{\sigma} = 1/6$. It occurs that the matrices associated with the coarse grid are more difficult to solve now. For the first threshold parameter τ_1 as small as 10^{-4}, we observe no pivot modifications, and the average number of BiCGstab iterations per linear solve is only 8. This suggests that the exact LU factorization is still stable. Such small τ_1 is nonpractical because of enormous memory demands and factorization time. However, already for τ_1 equal $3{\cdot}10^{-4}$ on two time steps, the algorithm makes 12 and 4 modifications of nearly zero pivots in order to avoid the breakdown. This caused the convergence slowdown, as many as 135 iterations for one system. Certain Oseen systems with $\bar{\sigma} = 1/6$ on the fine grid can not be solved by the ILU-preconditioned BiCGstab iterations with any values of threshold parameters that we tried. Note that for smaller $\bar{\sigma} = 1/12$, the algorithm performs without pivot modifications even for $\tau_1 = 0.03$.

Further, we decrease the viscosity of the fluid to $\nu = 0.025\mathrm{cm}^2/\mathrm{s}$ and try to run the same simulation on the coarse grid. For this value of the viscosity, the simulation without SUPG stabilization fails (solution blow-up is observed on $t = 0.23$ s). Adding SUPG stabilization allows to obtain physiologically meaningful solution, however, for large enough parameter $\bar{\sigma}$, the linear systems are harder to solve: $\bar{\sigma} = 1/6$ requires smaller threshold parameter τ_1, whereas $\bar{\sigma} = 1/3$ generates unsolvable systems, see Table 4. This experiment confirms that restrictions on $\bar{\sigma}$ come both from stability of the FE method and algebraic stability of the LU factorization.

We finally note that in experiments with varying inlet velocity, which leads to varying Reynolds number, the two-parameter ILU preconditioner demonstrated a remarkable adaptive property. The fill-in of the L and U blocks decrease or increase depending on the Reynolds number; see Figure 4 and compare to the inlet waveform in Figure 1. We will study this property of the two-parameter ILU preconditioner in more detail in a subsequent paper.

TABLE 4 The performance of ILU($\tau_1, \tau_2 = 7\tau_1^2$) for right coronary artery with less viscous blood $\nu = 0.025\mathrm{cm}^2/\mathrm{s}$. Threshold values allowing to run the entire SUPG-stabilized simulation with different stabilization parameters $\bar{\sigma}$. '⋆' means solution blow-up, '–' means untracktable systems for any applicable τ_1

$\bar{\sigma}$	0	1/96	1/48	1/24	1/12	1/6	1/3
τ_1	⋆	0.03	0.03	0.03	0.03	0.003	–

7 | CLOSING REMARKS AND CONCLUSIONS

In this paper, we studied the stability of the LU factorization for the stabilized FE formulations of the incompressible Navier–Stokes equations. Further, the two-parameter threshold ILU factorization was applied to define a preconditioner in the Krylov subspace method. Advantages and shortcomings of incomplete elementwise factorization preconditioners are well known: On the one hand, they are rather insensitive to discretization, boundary conditions for governing PDEs, domain geometry, and flow directions; on the other hand, even for discrete elliptic problems, ILU preconditioners do not scale optimally with respect to the number of unknowns. We observed such nonoptimality in the numerical experiments for generalized saddle-point problem as well. For 3D problems, when the mesh size is not too small, such dependence can be an acceptable price for other robustness properties of the preconditioner: in our experiments, the two-time increase of the number of mesh cells led only 20% increase of the iteration counts. Similar to the previous studies in this literature,[19] we found that natural **u**-p ordering of unknowns is sufficient for numerical stability of exact LU-factorization, when stabilization parameters satisfy certain bounds. In the algebraic language, this translates as the positive definiteness of the A block and the sufficiently small size of perturbation in the (1,2)-block. In this paper, the stability bounds for the factorization are rigorously formulated in terms of algebraic properties of subblocks of the original saddle-point matrix.

In general, higher Reynolds numbers lead to efficiency loss for most well-known preconditioners for Equation 3. In case of 3D blood flow in coronary arteries, the actual viscosity and velocity are such that P2-P1 stable FE discretization still provides the non-oscillatory solution on tetrahedral meshes with $\sim 10^5$ cells. However, the coronary blood flow parameters are close to the limit of non-oscillatory discretization, and SUPG stabilization may be in-demand. SUPG stabilization alters the

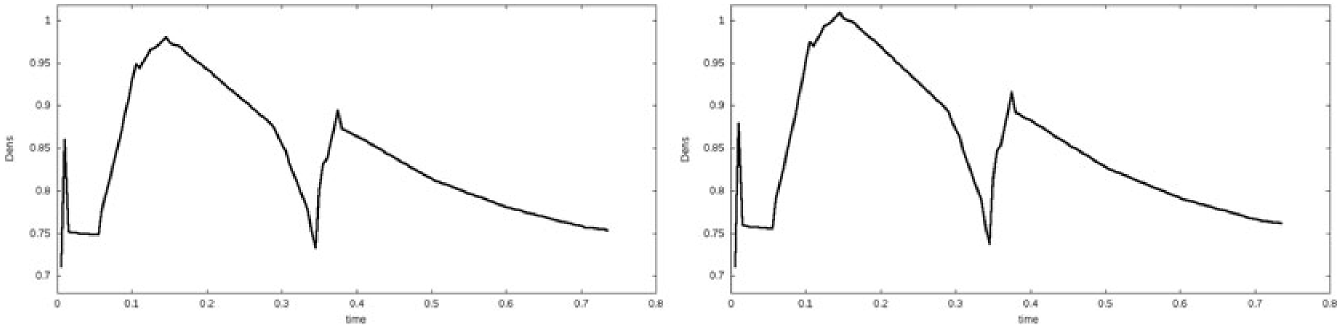

FIGURE 4 The fill-in of the LU factors for $\bar{\sigma} = 0$ (left) and $\bar{\sigma} = 1/12$ (right)

(1,1)-block and (1,2)-block of the Oseen matrix (Equation 3) and, hence, changes open new questions about the stability of factorizations. Theorem 4.1 show how the constants in the algebraic stability estimates depend on the flow and discretization parameters. This gives a certain insight into the performance of incomplete factorizations as preconditioners for flow problems. The present numerical analysis of incomplete factorizations for such nonsymmetric matrices is still limited to the lower estimate (Equation 59) of the diagonal entries of the triangular factors.

The two-parameter ILU preconditioner was applied to hemodynamic flow in a right coronary artery reconstructed from a real patient coronary computed tomography angiography. The performance of the preconditioner is good for a suitable choice of SUPG-stabilization parameters.

ACKNOWLEDGEMENTS

The authors thank Tatiana Dobroserdova and Alexander Danilov for the assistance in building tetrahedral meshes and FE systems, and Sergei Goreinov for sharing his implementation of the row-wise variant of the ILU(τ_1,τ_2) factorization.

REFERENCES

1. Girault V, Raviart P-A. Finite Element Approximation of the Navier–Stokes Equations, Lecture Notes in Mathematics, vol. 749. Berlin: Springer Verlag; 1979.
2. Franca LP, Frey SL. Stabilized finite element methods: Ii. the incompressible navier–stokes equations. Comput Meth Appl Mech Eng. 1992;99(2):209–233.
3. Roos H-G, Stynes M, Tobiska L. Robust Numerical Methods for Singularly Perturbed Differential Equations: Convection–Diffusion-Reaction and Flow Problems, vol. 24. Berlin: Springer Science & Business Media; 2008.
4. Codina R. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. Comput Meth Appl Mech Eng. 2002;191(39):4295–4321.
5. Braack M, Burman E, John V, Lube G. Stabilized finite element methods for the generalized oseen problem. Comput Meth Appl Mech Eng. 2007;196(4):853–866.
6. Turek S. Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approache, vol. 6. Berlin: Springer Science & Business Media; 1999.
7. Gelhard T, Lube G, Olshanskii MA, Starcke J-H. Stabilized finite element schemes with LBB-stable elements for incompressible flows. J Comput Appl Math. 2005;177(2):243–267.
8. Benzi M, Golub GH, Liesen J. Numerical solution of saddle point problems. Acta Numer. 2005;14(1):1–137.
9. Elman HC, Silvester D, Wathen A. Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics. Oxford: Oxford University Press; 2014.
10. Olshanskii MA, Tyrtyshnikov EE. Iterative Methods for Linear Systems: Theory and Applications. Philadelphia, PA: SIAM; 2014.
11. Vuik C, Segal G, et al. Simple-type preconditioners for the Oseen problem. Int J Numer Meth Fluids. 2009;61(4):432–452.
12. Elman HC, Tuminaro RS. Boundary conditions in approximate commutator preconditioners for the Navier–Stokes equations. Electron Trans Numer Anal. 2009;35:257–280.
13. Olshanskii MA, Vassilevski YV. Pressure Schur complement preconditioners for the discrete Oseen problem. SIAM J Sci Comput. 2007;29(6):2686–2704.
14. Scott J, Tuma M. On signed incomplete Cholesky factorization preconditioners for saddle-point systems. SIAM J Sci Comput. 2014;36(6):A2984–A3010.
15. Scott J, Tuma M. Solving symmetric indefinite systems using memory efficient incomplete factorization preconditioners: STFC Rutherford Appleton Laboratory; 2015.
16. Dahl O, Wille SØ. An ILU preconditioner with coupled node fill-in for iterative solution of the mixed finite element formulation of the 2D and 3D Navier–Stokes equations. Int J Numer Meth Fluids. 1992;15(5): 525–544.
17. Vuik C, Segal G, et al. A comparison of preconditioners for incompressible Navier–Stokes solvers. Int J Numer Meth Fluids. 2008;57(12): 1731–1751.
18. Segal A, ur Rehman M, Vuik C. Preconditioners for incompressible Navier–Stokes solvers. Numer Math: Theory, Meth and Appl. 2010;3(3):245–275.
19. Konshin I. N, Olshanskii MA, Vassilevski YV. ILU preconditioners for nonsymmetric saddle-point matrices with application to the incompressible navier–stokes equations. SIAM J Sci Comput. 2015;37(5): A2171–A2197.
20. Brooks AN, Hughes TJR. Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations. Comput Meth Appl Mech Eng. 1982;32(1):199–259.
21. Roos H-G, Stynes M, Tobiska L. Numerical Methods for Singularly Perturbed Differential Equations: Convection–Diffusion and Flow Problems. Berlin: Springer; 1996.
22. Hughes TJR, Feijóo GR, Mazzei L, Quincy J-B. The variational multiscale method – a paradigm for computational mechanics. Comput Meth Appl Mech Eng. 1998;166(1):3–24.
23. Ahmed N, Rebollo TC, John V, Rubino S. A review of variational multiscale methods for the simulation of turbulent incompressible flows. Arch Comput Meth Eng. 2015. doi: 10.1007/s11831-015-9161-0.
24. Stoer J, Bulirsch R. Introduction to Numerical Analysis. New York: Springer; 1993.
25. Braack M, Mucha PB, Zajaczkowski WM. Directional do-nothing condition for the Navier–Stokes equations. J Comput Math. 2014;32(5): 507–521.
26. Sani RL, Gresho PM. Résumé and remarks on the open boundary condition minisymposium. Int J Numer Meth Fluids. 1994;18(10):983–1008.
27. Ol'shanskii MA, Staroverov VM. On simulation of outflow boundary conditions in finite difference calculations for incompressible fluid. Int J Numer Meth Fluids. 2000;33(4):499–534.
28. Ethier CR, Steinman DA. Exact fully 3D Navier–Stokes solutions for benchmarking. Int J Numer Meth Fluids. 1994;19(5):369–375.
29. Kaporin IE. Scaling, reordering, and diagonal pivoting ILU. Russ J Numer Anal Math Model. 2007;22(4):341–375.
30. Golub GH, Loan CF. Matrix Computations. Baltimore, MD: Johns Hopkins University Press; 1996.
31. Benzi M. Preconditioning techniques for large linear systems: A survey. J Comput Phys. 2002;182(2):418–477.
32. Tismenetsky M. A new preconditioning technique for solving large sparse linear systems. Linear Algebra Appl. 1991;154:331–353.
33. Suarjana M, Law KH. A robust incomplete factorization based on value and space constraints. Int J Numer Meth Eng. 1995;38(10):1703–1719.
34. Kaporin IE. High quality preconditioning of a general symmetric positive definite matrix based on its $U^TU + U^TR + R^TU$-decomposition. Numer Linear Algebra Appl. 1998;5(6):483–509.
35. Saad Y. Iterative Methods for Sparse Linear Systems. Philadelphia, PA: SIAM; 2003.
36. Golub GH, Van Loan C. Unsymmetric positive definite linear systems. Linear Algebra Appl. 1979;28:85–97.

37. Lipnikov K, Vassilevski Y, Danilov A, et al. Advanced Numerical Instruments 2D. [accessed 2016 Dec 29]. Available from: http://sourceforge.net/projects/ani2d.

38. Lipnikov K, Vassilevski Y, Danilov A, et al. Advanced Numerical Instruments 3D. [accessed 2016 Dec 29]. Available from: http://sourceforge.net/projects/ani3d.

39. Jung J, Hassanein A, Lyczkowski RW. Hemodynamic computation using multiphase flow dynamics in a right coronary artery. Ann Biomed Eng. 2006;34(3):393–407.

How to cite this article: Konshin I, Olshanskii M, Vassilevski Y. LU factorizations and ILU preconditioning for stabilized discretizations of incompressible Navier-Stokes equations. *Numer. Linear Algebra Appl*. 2017;e2085. https://doi.org/10.1002/nla.2085