Hybrid First-Order System Least-Squares Finite Element Methods With The Application To Stokes And Navier-Stokes Equations

by

Kuo Liu

B.S., University of Science and Technology of China, 2007

M.S., University of Colorado at Boulder, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Mathematics

2012

This thesis entitled:
Hybrid First-Order System Least-Squares Finite Element Methods With The Application To Stokes And Navier-Stokes Equations
written by Kuo Liu
has been approved for the Department of Applied Mathematics

Prof. Thomas A. Manteuffel

Prof. Stephen F. McCormick

Date ____________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Liu, Kuo (Ph.D., Applied Mathematics)

Hybrid First-Order System Least-Squares Finite Element Methods With The Application To Stokes And Navier-Stokes Equations

Thesis directed by Prof. Thomas A. Manteuffel

This thesis combines the FOSLS method with the FOSLL* method to create a Hybrid method. The FOSLS approach minimizes the error, $\mathbf{e}^h = \mathbf{u}^h - \mathbf{u}$, over a finite element subspace, $\mathcal{V}^h$, in the operator norm, $\min_{\mathbf{u}^h \in \mathcal{V}^h} \|L(\mathbf{u}^h - \mathbf{u})\|$. The FOSLL* method looks for an approximation in the range of L^*, setting $\mathbf{u}^h = L^*\mathbf{w}^h$ and choosing $\mathbf{w}^h \in \mathcal{W}^h$, a standard finite element space. FOSLL* minimizes the L^2 norm of the error over $L^*(\mathcal{W}^h)$, that is, $\min_{\mathbf{w}^h \in \mathcal{W}^h} \|L^*\mathbf{w}^h - \mathbf{u}\|$. FOSLS enjoys a locally sharp, globally reliable, and easily computable a-posterior error estimate, while FOSLL* does not.

The Hybrid method attempts to retain the best properties of both FOSLS and FOSLL*. This is accomplished by combining the FOSLS functional, the FOSLL* functional, and an intermediate term that draws them together. The Hybrid method produces an approximation, $\mathbf{u}^h$, that is nearly the optimal over $\mathcal{V}^h$ in the graph norm, $\|\mathbf{e}^h\|_{\mathcal{G}}^2 := \frac{1}{2}\|\mathbf{e}^h\|^2 + \|L\mathbf{e}^h\|^2$. The FOSLS and intermediate terms in the Hybrid functional provide a very effective a posteriori error measure.

In this dissertation we show that the Hybrid functional is coercive and continuous in a graph-like norm with modest coercivity and continuity constants, $c_0 = 1/3$ and $c_1 = 3$; that both $\|\mathbf{e}^h\|$ and $\|L\mathbf{e}^h\|$ converge with rates based on standard interpolation bounds; and that, if LL^* has full H^2-regularity, the L^2 error, $\|\mathbf{e}^h\|$, converges with a full power of the discretization parameter, h, faster than the functional norm. Letting $\tilde{\mathbf{u}}^h$ denote the optimum over $\mathcal{V}^h$ in the graph norm, we also show that if superposition is used, then $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}}$ converges two powers of h faster than the functional norm. Numerical tests on are provided to confirm the efficiency of the Hybrid method and effectiveness of the a posteriori error measure.

Dedication

To my grandmother.

Acknowledgements

Upon the finishing phase of my thesis, I can finally relax a little and recollect the scattered moments in the past five years. All I can think of is how many people have supported me wholeheartedly and how many thanks I owe them for helping me grow up and mature.

I thank my advisor Tom Manteuffel for leading me into research, for his kindness, generosity and the care he has for the well-being of his students not only academically but also personally, for the salmons he grilled so many times for us, for the bike rides, the hikes he led. I thank my advisor Steve McCormick, who can write notes for a 1.5 hour lecture in a quarter of letter paper, who always demonstrates the beauty of simplicity, who constantly reminds us how much fun a professor can have and who laughs out loud at our bad jokes. I thank John Ruge, who is so patient and explains things so clearly, who shows how smart, knowledgeable and at the same time how humble a person can be, who has the gentlest heart under a true cowboy appearance. I thank Marian Brezina, who never minds spending time to help us out of struggling in the research, who shares his love of twice-cooked pork and who treats us like an uncle. I thank my committee professor Xiao-Chuan Cai, from whom I took Numerical PDEs and had his home-grown Chinese cucumber salad. I thank Prof. Congming Li for his help on my learning of Sobolev spaces and the advice he gave. I also thank my group members: Lei, Christian, Geoff, James, Minho, Jacob, Chad, David, Toby, Chris for their willingness to share, for their unique merits respectively I have learnt from.

Finally I'd like to thank my family in China who have always been supportive in my academic endeavor, who offer me the unconditional love.

Contents

Tables

Table

Figures

Figure

Chapter 1

Introduction

The first-order system least-squares (FOSLS) finite element method has been applied to a wide class of partial differential equations (PDEs) that arise from a broad range of physics and engineering problems. For example, second order elliptic equations [19, 24], convection diffusion equations [27], Stokes [21,22], Navier-Stokes equations [11–13], linear elasticity equations [18,21,22], neutron transport equations, Helmholtz equations, linear hyperbolic equations, Burgers' equations and so on. The method has proved its success not only on prototype research problems, but also challenging real world problems, such as elliptic problems with discontinuous coefficients, that arise in the physics of flow in heterogeneous porous media, neutron transport, biophysics **etc.** [10]; modeling flow in compliant blood vessels, which involves Navier-Stokes equations on the evolving fluid domain and elasticity equations on the tissue domain [33]; complex fluid problems such as magneto-hydrodynamics (MHD) equations, which are coupled with Maxwell equations and Navier-Stokes equations that involve 16 equations and 13 unknowns [2].

The goal a FOSLS researcher wants to achieve is not only to design a successful discretization scheme, but also a scheme that leads to a nice linear system which can take full advantage of modern linear solvers, such as the multigrid method [53]. Thus, a deep understanding of modern linear solvers (stationary iterative methods: Jacobi, Gauss-Seidel, SOR **etc.**; Krylov subspace method: CG, GMRES, **etc.**, all variations of multigrid method [48]) is essential. A lot of techniques are developed, such as introducing additional equations, re-ordering variables, adding slack variables **etc.** to re-arrange and fine-tune the PDE system for this purpose. This endows FOSLS with a lot of flexibility and potential, but, at the same, presents the designers with great challenges.

FOSLS has also demonstrated its versatility by offering the user a free a-posteriori error measure that can guide the adaptive mesh refinement with little extra cost [4, 9, 49]. Based on the locally sharp, globally reliable FOSLS error indicator and the work estimate for algebraic multigrid (AMG), the user can expect an error reduction with an minimal computational cost. This methodology has been implemented both serially and parallelly and applied to MHD equations with over 15 million biquadratic elements [50].

Variations of FOSLS, such as weighted-FOSLS method and FOSLL* method, are also studied for problems with singularities [23, 38–40, 43, 44].

The general approach of standard FOSLS is as follows:

- First, reformulate the original system of PDEs into a possibly enlarged first-order differential system, $L\mathbf{u} = \mathbf{f}$. This step generally involves introducing new dependent variables and adding consistent additional equations

- Then, a least-squares L^2-norm principle is applied to this first-order system; that is, we seek the weak solution by minimizing the FOSLS functional, $\mathcal{F}(\mathbf{u}; \mathbf{f}) := \|L\mathbf{u} - \mathbf{f}\|^2$, in an appropriate Hilbert space, $\mathcal{V}$ (ideally, $\mathcal{V} \subset H^1$). The minimization problem leads to the weak problem: find $\mathbf{u} \in \mathcal{V}$, such that for all $\mathbf{v} \in \mathcal{V}$,

$$< L\mathbf{u}, L\mathbf{v} >=< \mathbf{f}, L\mathbf{v} > .$$

- Finally, we cast the minimization problem in a finite element subspace, $\mathcal{V}^h \subset \mathcal{V}$, and assemble a matrix based on the discrete weak problem: find $\mathbf{u}^h \in \mathcal{V}^h$, such that, for all $\mathbf{v}^h \in \mathcal{V}^h$,

$$< L\mathbf{u}^h, L\mathbf{v}^h >=< \mathbf{f}, L\mathbf{v}^h > .$$

  At this step, we are ready to hand the classical matrix problem, $Ax = b$, to the user's favorite linear solver.

By the name of this thesis, "hybrid-FOSLS" is a novel method that combines FOSLS and FOSLL* method to take advantage of the merits of both. The general approach of FOSLS is introduced above and the general approach of FOSLL* is as follows.

- First, based on the first-order system obtained from FOSLS (the primal system), form the adjoint system. Suppose the primal system is

$$L\mathbf{u} \;=\; \mathbf{f},$$

  the adjoint system will be

$$L^*\mathbf{w} \;=\; \hat{\mathbf{u}},$$

  where L^* is the adjoint operator of L defined by integration by parts and $\hat{\mathbf{u}}$ is the exact solution for primal system. We call $\mathbf{w}$ the adjoint variable.

- Then, similar to FOSLS method, minimize the FOSLL* functional, $\|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2$, in the adjoint space, $\mathcal{W}$, that is dependent on $\mathcal{V}$. This also induces a weak problem: find $\mathbf{w} \in \mathcal{W}$, such that for all $\mathbf{z} \in \mathcal{W}$

$$< L^*\mathbf{w}, L^*\mathbf{z} > = < \hat{\mathbf{u}}, L^*\mathbf{z} > = < \mathbf{f}, \mathbf{z} > .$$

  Note that the importance of choosing L's adjoint operator, L^*, is illustrated by the second equality above.

- Next, solve the minimization problem in the finite element subspace $\mathcal{W}^h$: find $\mathbf{w}^h \in \mathcal{W}^h$, such that, for all $\mathbf{z}^h \in \mathcal{W}^h$,

$$< L^*\mathbf{w}^h, L^*\mathbf{z}^h > = < \hat{\mathbf{u}}, L^*\mathbf{z}^h > = < \mathbf{f}, \mathbf{z}^h > .$$

- Finally, recover the numerical solution, $\mathbf{u}^h$, of the primal system by letting $\mathbf{u}^h = L^*\mathbf{w}^h$.

The advantages and the limitations of FOSLS and FOSLL* can be summarized as follows.

For FOSLS method,

its advantages include:

- since the weak form is derived from a minimization problem, as long as we can prove its weak form is well-posed (the bilinear form is continuous and coercive), the resulting linear system is symmetric and positive definite (SPD). An SPD system usually has a great advantage when it comes to solving the linear system.

- If we can further prove the FOSLS operator, L, is H^1-equivalent (the bilinear form is continuous and coercive in H^1-norm), we can show the finite element scheme has the optimal discretization convergence rate [24].

- An H^1-equivalent L also results in the existence of an optimal multigrid solver whose complexity is of $\mathcal{O}(n)$, where n is the number of unknowns in the matrix. See [24] and the reference therein.

- The Ladyzhenskaya-Babuška-Brezzi (LBB) condition is not needed for Stokes/Navier-Stokes equations when choosing conforming finite element spaces for pressure and velocity [21, 22].

- The FOSLS functional $\|L\mathbf{u} - \mathbf{f}\|^2$ serves as an excellent error indicator for adaptive mesh refinement [4, 9].

FOSLS's limitations are

- H^1-equivalence generally fails in the presence of discontinuous coefficients, when the problem is posed in an irregular domain or when the problem involves irrgular boundary conditions. In this case the convergence rate suffers and in some cases convergence may fail [39].

- The standard FOSLS method yields a loss of mass conservation for Stokes/Navier-Stokes in certain geometries with underresolved grids [34, 41].

For FOSLL* method,

its main advantages are

- Application on problem with less smoothness proves to be successful [23, 34, 40, 44].

- Previous research shows a two-stage method combining with FOSLL* can effectively enhance mass conservation in Stokes/Navier-Stokes equations.

However, FOSLL* has the limitation that:

- The a-priori error estimate is lost, since, different from FOSLS functional, the FOSLL* functional, $\|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2$, involves the exact solution, $\hat{\mathbf{u}}$, and is not computable.

Hybrid-FOSLS method combines FOSLS, FOSLL* and minimizes the Hybrid functional:

$$\mathcal{H}((\mathbf{w}, \mathbf{u}); (\hat{\mathbf{u}}, \mathbf{f})) := \|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2 + \|L\mathbf{w} - \mathbf{u}\|^2 + \|L\mathbf{u} - \mathbf{f}\|^2,$$

that involves a FOSLL* term, a FOSLS term and an intermediate term that draws the two together. The minimization problem is cast in a product space, $\mathcal{W} \times \mathcal{V}$. The weak problem of minimizing the Hybrid functional is: find $(\mathbf{w}, \mathbf{u}) \in \mathcal{W} \times \mathcal{V}$, such that, $\forall(\mathbf{z}, \mathbf{v}) \in \mathcal{W} \times \mathcal{V}$,

$$< L^*\mathbf{w}, L^*\mathbf{z} > + < L^*\mathbf{w} - \mathbf{u}, L^*\mathbf{z} - \mathbf{v} > + < L\mathbf{u}, L\mathbf{v} > = < \mathbf{f}, \mathbf{z} + L\mathbf{v} > . \tag{1.1}$$

Hybrid-FOSLS does successfully take advantage of both FOSLS and FOSLL* in that

- it keeps all the major feature FOSLS has;
- it improves performance on problems with reduced regularity;
- the sum of FOSLS and intermediate term also serves a good error indicator for adaptive mesh refinement.

The reasoning behind this is due to the fact that the homogeneous Hybrid functional, $\mathcal{H}((\mathbf{w}, \mathbf{u}); (\mathbf{0}, \mathbf{0}))$, is elliptic in a elliptic in a graph norm with mild coercivity and continuity constants. This implies that, when the Hybrid functional is reduced, L^2-norm of the error is also very well reduced. We have also shown that minimizing Hybrid functional is very close to minimizing the graph functional, $\mathcal{G}(\mathbf{u}; \hat{\mathbf{u}}, \mathbf{f}) = \frac{1}{2}\|\mathbf{u} - \hat{\mathbf{u}}\|^2 + \|L\mathbf{u} - \mathbf{f}\|^2$.

This dissertation is organized as follows: in Chapter 2, the physics background of Stokes and Navier-Stokes equations are introduced and the basics of FOSLS and FOSLL* methods are presented. We also introduce the variety of FOSLS formulations for Stokes/Navier-Stokes equation and detail the formulations we will use for our numerical tests. Our main theoretical results are developed in Chapter 3 which include proof of ellipticity in the graph norm of the bilinear form associated with the Hybrid functional, the error estimate of both the Hybrid functional and the L^2-norm of the error. The similarity of minimizing the graph functional and minimizing the Hybrid functional is also illustrated by a set of theorems that show the numerical solutions from minimizing the two functionals are converging to each other very quickly. The numerical test for Stokes equations in a longtube is carried out and behaves exactly as the theory expects. We dedicate Chapter 4 to the problem with a corner singularity and compare the mass conservation using both Hybrid and FOSLS. We also develop and apply our Hybrid version of adaptive mesh refinement to this problem. Basic background of analysis tools for problems with singularities is also introduced. In Chapter 5, we explore the application of the Hybrid method to nonlinear Navier Stokes equations. We first introduce how to combine FOSLS with Newton iteration to handle the non-linearity, then develop the Newton-Hybrid FOSLS approach. Numerical results that exhibit similarities to linear problems are presented in the end. Conclusions of this thesis are made in Chapter 6.

Chapter 2

Background

In this chapter, we introduce background that is important for the development of this thesis: the physics behind the equations, the numerical methods available and their pros and cons, the basics of FOSLS and FOSLL* methods, upon which the hybrid-FOSLS method is developed, and their advantages and limitations.

2.1 Stokes/Navier Stokes Equations

Through out this thesis, a vector is denoted by a boldface letter and is always assumed to be a column vector, a tensor (matrix) is denoted by a boldface letter with an underline and a "t" at the superscript denotes the transpose of a vector. The open domain with a certain smoothness assumption of the equations is denoted by Ω and its boundary is $\partial\Omega$.

2.1.1 Introduction

For Newtonian fluid, its defining character is that its stress tensor versus strain rate follows a linear law [46]

$$\underline{\boldsymbol{\sigma}} = \eta(\nabla \mathbf{u} + \nabla \mathbf{u}^t) + (\xi - \frac{2\eta}{3})\underline{\mathbf{I}}(\nabla \cdot \mathbf{u}), \tag{2.1}$$

where $\underline{\boldsymbol{\sigma}}$ denotes the stress tensor, $\mathbf{u}$ denotes the velocity of the fluid, $\underline{\mathbf{I}}$ denotes the unit tensor and η, ξ denote the first and second viscosity, respectively. Water, oil, salt solution are common Newtonian fluids while pasty fluids (**e.g.** yoghurt, shampoo), blood, rarefied gases are typical non-Newtonian fluids.

In this dissertation, we are particularly interested in incompressible flows, whose density $\rho(\mathbf{x}, t)$ is nearly a constant. Following conservation of momentum and conservation of mass, the incompressible Navier Stokes Equations can be derived and simplified to

$$\frac{\partial \mathbf{u}}{\partial t} - \nu\Delta\mathbf{u} + (\nabla\mathbf{u}^t)^t\mathbf{u} + \nabla p = \mathbf{f}, \tag{2.2}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{2.3}$$

where $\nu = \eta/\rho$ is the first viscosity rescaled by density. The pressure, p, and the force, f, are also rescaled and represent p/ρ, $\mathbf{f}/\rho$. Also, for clarification, for the 3D problem,

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \qquad \nabla\mathbf{u}^t = \begin{pmatrix} u_{1x} & u_{2x} & u_{3x} \\ u_{1y} & u_{2y} & u_{3y} \\ u_{1z} & u_{2z} & u_{3z} \end{pmatrix}.$$

Note that, in the literature, η is also referred as "dynamic viscosity" and ν is sometimes referred as "kinematic viscosity". The Reynolds number is defined as

$$Re := \frac{\rho\mathbf{v}L}{\eta} = \frac{\mathbf{v}L}{\nu},$$

where $\mathbf{v}$ is the reference velocity (the mean velocity of the object relative to the fluid, **e.g.** the magnitude of the inflow velocity), L is the characteristic length ($L \approx diam(\Omega)$). Without the loss of generality, $\mathbf{v}$, L are assumed to be 1. Note that ν is also called "the inverse of Reynolds number" by some authors.

By its definition, Reynolds number gives the measure of the ratio of inertial force (which measures how much a particular fluid resists any change in motion) to viscous force (which characterize the resistance of a fluid to the flow). The following table should provide a perceptual idea about Reynolds number, where the organisms are the fluids and the air or water are the flows [52].

When Re is very large ($Re >> 1$), ν is very small, the effect from the inertial part of (2.2), $(\nabla\mathbf{u}^t)^t\mathbf{u}$, is much stronger than the viscous part, $\nu\Delta\mathbf{u}$. In that case, the incompressible Navier Stokes Equations can be further reduced to incompressible Euler equations, which describes

A large whale swimming at $10m/s$	$300,000,000$
A tuna swimming at the same speed	$30,000,000$
A duck flying at $20m/s$	$300,000$
A large dragon fly going $7m/s$	$30,000$
A copepod in a speed burst of $0.2m/s$	300
Flapping wings of the smallest flying insects	30
An invertebrate larva, $0.3mm$ long, at $1mm/s$	0.3
A sea urchin sperm advancing the species at $0.2mm/s$	0.03
A bacterium, swimming at $0.01mm/s$	0.00001

Table 2.1: A spectrum of Reynolds numbers for self-propelled organisms

inviscid flows

$$\frac{\partial \mathbf{u}}{\partial t} + (\nabla \mathbf{u}^t)^t \mathbf{u} + \nabla p = \mathbf{f}, \tag{2.4}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{2.5}$$

On the other hand, when $Re << 1$, the viscous force dominates the inertial force. If we assume the inertial force is negligible, then the incompressible Navier Stokes Equations gives the Stokes Equations

$$\frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \tag{2.6}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{2.7}$$

The Stokes Equations are widely used to model low Reynolds number fluid, such as swimming micro-organisms, the flow of viscous polymers, lava *etc.* Since the steady state Stokes Equations (where we let $\nu = 1$ without the loss of generality),

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega, \end{aligned} \tag{2.8}$$

arises in a time discretization of the Navier-Stokes equations, the study of Stoke equations is also very important to Navier-Stoke equations, whose existance and smoothness in 3D still remain one of the seven most important open problems in mathematics [26].

Another note before we leave this subsection: for incompressible flows, the incompressibility constraint equation (2.3) implies that (2.1) can be simplified to

$$\underline{\boldsymbol{\sigma}} = \eta(\nabla\mathbf{u} + \nabla\mathbf{u}^t). \tag{2.9}$$

We will refer to this equation later in this section.

2.1.2 Numerical Methods

Before we start the discussion on numerical methods, we first introduce some function spaces and associated norms. First, we denote the Hilbert space of square-integrable functions on Ω by $L^2(\Omega)$, whose inner product and norm are:

$$< u, v >= \int_\Omega uv \; d\Omega, \qquad \|u\|_0 = (u, u)^{1/2}.$$

For the convenience of denotation, in the rest of the thesis, we will use $\|u\|$ interchangeably with $\|u\|_0$ without more explanation.

Next, we define the Sobolev spaces $H^k(\Omega)$ for any non-negative integer k.

$$H^k(\Omega) = \{u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega), \forall |\alpha| \leq k\},$$

where D^α is the weak derivative with index α. The norm of $H^k(\Omega)$ is defined by

$$\|u\|_k = (\sum_{0\leq|\alpha|\leq k} \|D^\alpha u\|_0^2)^{1/2},$$

which is the sum of the L^2 norm of all its derivatives up to order k and the function itself. Thus, $H^0(\Omega)$ is the same space as $L^2(\Omega)$. One of the Sobolev spaces that is of special interest is the affine space:

$$H_0^1(\Omega) := \{u \in H^1(\Omega) : u|_{\partial\Omega} = 0\},$$

where $u|_{\partial\Omega}$ should be understood as the trace of u on $\partial\Omega$. It is equipped with the $H^1(\Omega)$ norm as defined above:

$$\|u\|_1 = \left(\|u\|_0^2 + \sum_{i=1}^{d} \left\| \frac{\partial u}{\partial x_i} \right\|_0^2 \right)^{1/2},$$

where $d = 1, 2, 3,$ denotes the dimension of the problem.

Notice that the $H^1(\Omega)$ semi-norm,

$$|u|_1 := \left(\sum_{i=1}^{d} \left\| \frac{\partial u}{\partial x_i} \right\|_0^2 \right)^{1/2},$$

is equivalent to $\|u\|_1$ (there exit positive constants c_0 and c_1, such that $c_0|u|_1 \leq \|u\|_1 \leq c_1\|u\|_1$) by Poincaré-Friedrichs inequality

Theorem 1. *(Poincaré-Friedrichs Inequality) For any function $u \in H_0^1(\Omega)$, there exists a constant $c_p > 0$, depending only on the domain Ω and the dimension of the problem d, such that*

$$\|u\|_0 \leq c_p \|\nabla u\|_0. \tag{2.10}$$

Remark 1. *For the 2D problem, $c_p \approx diam(\Omega)$*

The dual space of $H_0^1(\Omega)$ is denoted by $H^{-1}(\Omega)$, which is the space of all bounded linear functional on $H_0^1(\Omega)$, and the associated norm is defined by

$$\|f\|_{-1} = \sup_{0 \neq u \in H_0^1(\Omega)} \frac{|< f, u >|}{|u|_1}.$$

The trace Sobolev spaces, $H^{k-\frac{1}{2}}(\partial\Omega)$, are traces of functions that are in $H^k(\Omega)$, with the norm

$$\|g\|_{k-\frac{1}{2},\partial\Omega} = \inf_{u \in H^k(\Omega), u|_{\partial\Omega} = g} \|u\|_k.$$

For vector functions, the according product spaces are denoted using boldface, for example,

$$\mathbf{H}_0^1(\Omega) = (H_0^1(\Omega))^d,$$

where $d = 1, 2, 3$.

Since our focus of this dissertation is to develop a novel least-squares finite element method *i.e.* hybrid-FOSLS which results in a variational problem from a minimization problem of a proper norm and leads to excellent bound on both the L^2 norm and the H^1 semi-norm of the error. For this purpose, we restrict our attention to steady state Stokes (3.58) and Navier-Stokes equations

(5.23) without any discussion on the time dicretization. The steady state Navier-Stokes equations are:

$$\begin{aligned} -\Delta \mathbf{u} + (\nabla \mathbf{u}^t)^t \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega. \end{aligned} \tag{2.11}$$

For Stokes equations with a no-slip homogeneous boundary condition,

$$-\Delta \mathbf{u} + \nabla p = \mathbf{f}, \qquad \Omega \tag{2.12}$$

$$\nabla \cdot \mathbf{u} = 0, \qquad \Omega \tag{2.13}$$

$$\mathbf{u} = 0, \qquad \partial\Omega \tag{2.14}$$

the classical Galerkin weak formulation is: for any $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$, find $\mathbf{u} \in \mathbf{H}_0^1(\Omega)$ and $p \in L_0^2(\Omega)$, such that

$$\int_\Omega (\nabla \mathbf{u} \cdot \nabla \mathbf{v} - p \nabla \cdot \mathbf{v}) \, d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{v} \, d\Omega, \qquad \forall \, \mathbf{v} \in \mathbf{H}_0^1(\Omega) \tag{2.15}$$

$$\int_\Omega q \nabla \cdot \mathbf{u} \, d\Omega = 0, \qquad \forall \, q \in L_0^2(\Omega), \tag{2.16}$$

where

$$L_0^2(\Omega) := \{p \in L^2(\Omega) : \int_\Omega p \, d\Omega = 0\}.$$

The weak problem above is also equivalent to the following constrained problem: for any $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$, find

$$\mathbf{u} \in \mathbf{Z} := \{\mathbf{u} \in \mathbf{H}_0^1(\Omega) : \; < \nabla \cdot \mathbf{u}, q >= 0, \quad \forall q \in L_0^2(\Omega)\} \tag{2.17}$$

and $p \in L_0^2(\Omega)$, such that (2.15) holds.

For linear elasticity equations,

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} + p &= g \quad \text{in } \Omega, \end{aligned} \tag{2.18}$$

which are very similar to Stokes equations but with a positive-definite weak form, using conforming finite element spaces is sufficient to ensure the existence and uniqueness of the discrete problem and an optimal finite element convergence rate. However, this is not enough for Stokes equations.

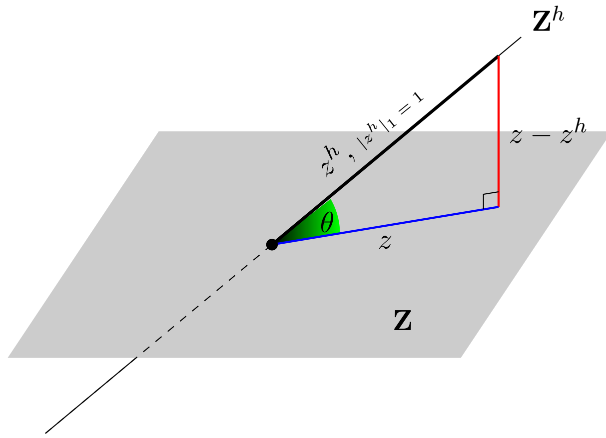


Figure 2.1: Illustration of the relation between $\mathbf{Z}$ and $\mathbf{Z}^h$, the constrained space (2.17) of velocity in Stokes equations and its discrete space. See (2.19).

The choice of conforming finite element spaces, $\mathbf{V}^h \subset \mathbf{H}_0^1(\Omega)$, for velocity and $S^h \subset L_0^2(\Omega)$ for pressure is not enough to guarantee a stable solution; that is, the resulting linear system from this discretization can be singular.

This challenge for Galerkin finite element methods is due to the fact that, using the conforming finite element spaces, the discrete version of $\mathbf{Z}$,

$$\mathbf{Z}^h := \{\mathbf{u}^h \in \mathbf{V}^h : \;\; < \nabla \cdot \mathbf{u}^h, q^h >= 0, \quad \forall q^h \in S^h\},$$

is not a subspace of $\mathbf{Z}$, in general. That is, a discretely divergence free space is not necessarily a divergence free space.

The difference of $\mathbf{Z}$ and $\mathbf{Z}^h$ can be measured by

$$\gamma = \sup_{z^h \in \mathbf{Z}^h} \inf_{z \in \mathbf{Z}} \frac{|z - z^h|_1}{|z^h|_1}. \tag{2.19}$$

The value of γ is between 0 to 1. We can think $\gamma = \sin\theta$ as the sine of the "angle" θ between $\mathbf{Z}$ and $\mathbf{Z}^h$ as in Figure (2.1.2), where $\gamma = |z - z^h|_1 = \sin(\theta)$.

To be more general, let $\nabla \cdot \mathbf{u} = g$ in place of (2.13) and define the bilinear forms

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \nu < \nabla \mathbf{u}, \nabla \mathbf{v} >, \quad \forall\, \mathbf{u},\, \mathbf{v} \in \mathbf{H}^1(\Omega), \\ b(\mathbf{v}, q) &= - < q, \nabla \cdot \mathbf{v} >, \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega) \text{ and } q \in L^2(\Omega). \end{aligned} \tag{2.20}$$

The weak problem of Stokes equations is now: find $(\mathbf{u}, p) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega)$, such that

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) &= < \mathbf{f}, \mathbf{v} >, \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega), \\ b(\mathbf{u}, q) &= < g, q >, \quad \forall q \in L_0^2(\Omega). \end{aligned} \tag{2.21}$$

To guarantee the above weak problem is well-posed, the operators $a(\cdot,\cdot)$ and $b(\cdot)$ need to satisfy the following assumptions: there exit constants $\alpha > 0$ and $\beta > 0$, such that

$$a(\mathbf{v}, \mathbf{v}) \geq \alpha \|\mathbf{v}\|_1, \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega) \tag{2.22}$$

$$\sup_{\mathbf{v} \in \mathbf{H}_0^1} \frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_1} \geq \beta \|q\|_0, \quad \forall q \in L_0^2(\Omega) \tag{2.23}$$

Notice that Eqn (2.23) is equivalent to the famous inf-sup or LBB (Ladyzhenskaya-Babuška-Brezzi) condition: there exists $\gamma > 0$, such that

$$\inf_{0 \neq q \in L_0^2(\Omega)} \sup_{0 \neq \mathbf{v} \in \mathbf{H}_0^1(\Omega)} \frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_1 \|q\|_0} > \gamma \tag{2.24}$$

Thus, a more rigorous explanation for the problem using conforming finite elements in Galerkin formulation is that using $\mathbf{V}^h \subset \mathbf{H}_0^1(\Omega)$ and $S^h \subset L_0^2(\Omega)$ doesn't guarantee the discrete version of LBB condition,

$$\inf_{0 \neq q^h \in S^h} \sup_{0 \neq \mathbf{v}^h \in \mathbf{V}^h} \frac{b(\mathbf{v}^h, q^h)}{\|\mathbf{v}^h\|_1 \|q^h\|_0} \geq \gamma, \tag{2.25}$$

will be satisfied automatically. Therefore care must be taken when it comes to choosing the right combination of finite element spaces. For example, denote P_0, P_1, P_2 the finite element spaces of piecewise constant, linear, quadratic polynomials on triangles; Q_0, Q_1, Q_2 function spaces of piecewise constant, bilinear, biquadratic polynomials on quadrilaterals.

- LBB condition is satisfied: $(\mathbf{u}^h, p) \in P_2 \times P_1$, $(\mathbf{u}^h, p) \in Q_2 \times Q_1$
- LBB condition is violated: $(\mathbf{u}^h, p) \in P_1 \times P_1$, $(\mathbf{u}^h, p) \in P_1 \times P_1$, $(\mathbf{u}^h, p) \in Q_1 \times Q_1$

In Subsection (2.3), the readers will that see for FOSLS method, the wellposedness of the discrete weak problem will be automatically satisfied with the conforming elements.

Another challenge for the Galerkin finite element method for Stokes equation is the increasing difficulty to solve the linear system as the Reynolds number increases. It is clear to see from (2.21) that the linear system after discretization is of the form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix},$$

which is known as a "saddle-point problem", since the solution, $(\underline{u}, \underline{p})$, to this indefinite matrix is the saddle point. Denote ϕ_j the basis functions for velocity. Then the entries of A are the same as from the discretization of Poisson equation: $\nu < \nabla\phi_j, \nabla\phi_i >$, and A is symmetric positive definite (SPD). When the Reynolds number is large, its inverse ν is small. This means the off-diagonal block B is relatively large, which can slow down the linear system solvers. However by rescaling $\underline{p}$, we can keep B small. Unfortunately, the rescaling trick won't work for Navier-Stokes equations due to the $\mathbf{u} \cdot \nabla \mathbf{u}$ term.

All the difficulties we have discussed in Galerkin formulation for Stokes equations also exist in Navier-Stokes equations. Below, we only present the its Galerkin weak formulation without further discussion.

For $\mathbf{f} \in \mathbf{H}^{-1}$, find $(\mathbf{u}, p) \in (\mathbf{H}_0^1(\Omega), L_0^2(\Omega))$, such that

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) + n(\mathbf{u},\mathbf{u},\mathbf{v}) + b(\mathbf{v},p) &= < \mathbf{f}, \mathbf{v} > \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega) \\ b(\mathbf{u}, q) &= < g, q > \quad \forall q \in L_0^2(\Omega), \end{aligned} \tag{2.26}$$

where,

$$n(\mathbf{w},\mathbf{u},\mathbf{v}) := \int_\Omega \mathbf{w} \cdot \nabla \mathbf{u} \cdot \mathbf{v} \; d\Omega, \qquad \forall \mathbf{u},\mathbf{v},\mathbf{w} \in \mathbf{H}^1(\Omega).$$

2.2 The Framework of FOSLS and FOSLL* Theory

2.2.1 FOSLS

"FOSLS " stands for the First-Order System Least-Squares finite element method. The name of FOSLS summaries its two most important features: the original PDEs is reformulated into a first-order differential system, $L\mathbf{u} = \mathbf{f}$; the variational problem is obtained by a least-squares procedure, *i.e.* to minimize a certain norm of the residual equations, $L\mathbf{u} - \mathbf{f}$.

In many cases, the PDEs we face are of the second order or higher. To require a first-order system implies that extra variables will be introduced, which increases the number of dependent variables. However, this does not mean an increased computational cost, as people might think. The main reason to use a first-order system is that, the algebraic system from a first-order system has condition number $\mathcal{O}(h^{-2})$, while for a second-order PDEs, the condition number is $\mathcal{O}(h^{-4})$, where h denotes the mesh size. A large condition number is problematic for algebraic solvers. On the other hand, the least-squares minimization leads to a symmetric positive definite (SPD) algebraic system, which is well-known to have great computational advantages.

To fully understand the FOSLS theory and master the skills to design a nice FOSLS system may be elusive at the start. In this section, we only present the basics of FOSLS and leave many details to later discussions.

Suppose the PDEs have already been recast into a first-order system, denote Ω as an open, connected domain in 2D or 3D with a Lipschitz boundary $\partial\Omega$. An abstract boundary value problem can be described as

$$\begin{aligned} L\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ B\mathbf{u} &= \mathbf{g} \quad \text{on } \partial\Omega, \end{aligned} \tag{2.27}$$

where $B\mathbf{u}$ can be Dirichlet, Neumann or Robin boundary conditions. Then, we denote the FOSLS functional by

$$\mathcal{F}(\mathbf{u};\mathbf{f}) = \|L\mathbf{u} - \mathbf{f}\|_{\mathcal{X}}^2$$

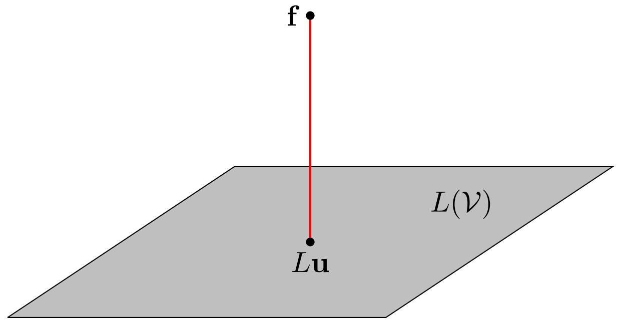
and seek the solution $\mathbf{u} \in \mathcal{V}$, such that

$$\mathbf{u} = arg \min_{\mathbf{u}\in\mathcal{V}} \|L\mathbf{u} - \mathbf{f}\|_{\mathcal{X}}^2,$$

where $\|\cdot\|_{\mathcal{X}}$ denotes some norm to be determined later. The weak problem from the minimization is thus:

Find $\mathbf{u} \in \mathcal{V}$, such that for any $\mathbf{v} \in \mathcal{V}$

$$< L\mathbf{u}, L\mathbf{v} >_{\mathcal{X}} = < \mathbf{f}, L\mathbf{v} >_{\mathcal{X}} . \tag{2.28}$$



This is obtained by taking the Frechét derivative of the FOSLS functional $\mathcal{F}(\mathbf{u};\mathbf{f})$

$$\lim_{\alpha\to 0}\frac{\mathcal{F}(\mathbf{u}+\alpha\mathbf{v};\mathbf{f})-\mathcal{F}(\mathbf{u};\mathbf{f})}{\alpha},$$

or simply by "assuming" that $(L\mathbf{u}-\mathbf{f})$ is orthogonal to the space $\mathcal{V}$ and $< L\mathbf{u}-\mathbf{f}, L\mathbf{v} >= 0$ for any $\mathbf{v}\in\mathcal{V}$.

Define the bilinear form,

$$\mathcal{B}(\mathbf{u},\mathbf{v}) =< L\mathbf{u}, L\mathbf{v} > .$$

The next important step is to prove the continuity and coercivity of $\mathcal{B}(\cdot,\cdot)$. We say the bilinear form is continuous and coercive in $\mathcal{V}$ if there exist positive constants, c_0 and c_1 such that, for every $vin\mathcal{V}$,

$$\begin{aligned}\mathcal{B}(\mathbf{u},\mathbf{v}) &\le c_1\|\mathbf{u}\|_{\mathcal{V}}\|\mathbf{v}\|_{\mathcal{V}},\\ \mathcal{B}(\mathbf{u},\mathbf{v}) &\ge c_0\|\mathbf{u}\|_{\mathcal{V}}\|\mathbf{v}\|_{\mathcal{V}},\end{aligned} \tag{2.29}$$

where c_0, c_1 are indepent of $\mathbf{u}$, $\mathbf{v}$ and are called coercivity and continuity constants, respectively. Thus, by Riesz Representation theorem (2), the weak problem is well-posed (**i.e.** the existance, uniqueness and continuous dependence on $\mathbf{f}$ of the solution).

Theorem 2. ***(Riesz)*** *If ϕ is a bounded linear functional on a Hilbert space $\mathcal{H}$, then there is a unique vector $y\in\mathcal{H}$, such that*

$$\phi(x) =< y, x >_{\mathcal{H}}, \quad \textit{for all } x\in\mathcal{H}. \tag{2.30}$$

Clearly, if $\mathcal{V}^h$ is a finite element subspace of $\mathcal{V}$, the dicrecte version of the weak problem: Find $\mathbf{u}^h\in\mathcal{V}^h$, such that for any $\mathbf{v}^h\in\mathcal{V}^h$

$$< L\mathbf{u}^h, L\mathbf{v}^h >=< \mathbf{f}, L\mathbf{v}^h > . \tag{2.31}$$

is automatically well-posed. Recall that for the Galerkin formulation of Stokes equations, the inf-sup condition cannot be automatically satisfied by merely using a finite element subspace, $\mathcal{V}^h \subset \mathcal{V}$. In other words, a conforming finite element space itself is not enough to guarantee that a well-posed continuous problem will lead to a well-posed discrete weak problem. An enormous amount of research is dedicated to finding the appropriate coupling of the mixed finite element spaces. While it may be advantageous to use mixed spaces, for a FOSLS formulation, a single continuous piecewise polynomial space is sufficient for the approximation of all unknowns.

Let $\{\phi_i\}_{i=1}^n$ denote the finite element basis functions. The numerical solution

$$\mathbf{u}^h = \sum_{i=1}^{n} u_i \phi_i,$$

is the linear combination of them, where the coefficients u_i's are to be determined. Thus, the discrete weak form can be written as

$$< L \sum_{j=1}^{n} u_j \phi_j, L\phi_i >=< \mathbf{f}, L\phi_i >$$

which is equivalent to

$$A\underline{u} = \underline{f},$$

where $\underline{u} = (u_1, \ldots, u_n)^t$, $\underline{f} = (f_1, \ldots, f_n)^t$, $f_i =< \mathbf{f}, L\phi_i >$, and the elements of the matrix A are $a_{ij} =< L\phi_j, L\phi_i >$. Again, notice that A is a symmetric matrix and is also positive-definite by the coercivity of L. A proper design of FOSLS system will yield an algebraic system with great computational advantages. We leave the details to later chapters.

When the bilinear form, $\mathcal{B}(\cdot,\cdot)$, is H^1-equivalent, that is, the $\|\cdot\|_{\mathcal{V}}$ in Eqn (2.29) is the H^1 norm in each dependent variable, we call the first-order operator associated with $\mathcal{B}(\cdot,\cdot)$ an "H^1-equivalent operator". This, in turn, leads to a optimal finite element convergence rate, using H^1-conforming finite elements such as the most common piecewise polynomials.

Assume that the domain, Ω, and right-hand-side satisfy certain smoothness constraints and the finite element space, $\mathcal{V}^h$, is defined with respect to a regular triangulation, we have the following property

Lemma 1. *(Approximation Property) For a conforming finite element subspace, $\mathcal{V}^h \subset \mathcal{V}$, where $\mathcal{V}^h$ contains the piecewise polynomials of degree $p > 0$, then for any $u \in \mathcal{V}$, there is a constant $C > 0$ independent of h, such that*

$$\|u - \mathcal{I}^h u\|_1 \leq Ch^r \|u\|_{r+1}, \quad \text{for } 0 < r \leq p, \tag{2.32}$$

where $\mathcal{I}^h u \in \mathcal{V}^h$ is the interpolant of u.

By the virtue of "Aubin-Nitsche trick" (duality argument) that is explained in the appendix, we have

$$\|u - \mathcal{I}^h u\|_0 \leq Ch^{p+1} \|u\|_{p+1}. \tag{2.33}$$

With the help of the lemma above, the error from FOSLS method can be estimated. Let $e^h = u^h - \hat{u}$ be the numerical error, recall c_0 and c_1 are coercivity and continuity constants, respectively, and suppose L is an H^1-equivalent operator. By the linearity of the operator, we have

$$\begin{aligned} \|e^h\|_1 &\leq \tfrac{1}{c_0} \|Le^h\| \leq C\|L(\mathcal{I}^h u - \hat{u})\| \\ &\leq \tfrac{c_1}{c_0} \|\mathcal{I}^h u - \hat{u}\|_1 \leq \tfrac{Cc_1}{c_0} \cdot h^p \|\hat{u}\|_{p+1}, \end{aligned} \tag{2.34}$$

where C denotes a generic positive constant that is independent of h. The inequalities above are due to coercivity of L, minimization property of FOSLS, continuity of L and the finite element approximation property, Lemma 1. This error estimate says that we can always employ higher order of elements to get the best asymptotic rate of convergence up to the smoothness order of the exact solution $\hat{u}$.

Our ultimate goal of the design of a discretization method is to have the optimal "error reduction per computation cost". This standpoint has to be twofold: we must have an approach to quantify the error reduction we can make and achieve optimal discretization convergence rate; we need the algebraic system to be amenable to linear solvers that are state-of-art. FOSLS not only does well on the first (optimal finite element convergence rate), but also the second. In fact, another advantage of the H^1-equivalence of L is that the resulting linear system will be amenable to solution by a multigrid algorithm [51, 53] in $\mathcal{O}(n)$ operations [45], where n is the dimension of the discrete space $\mathcal{V}^h$. If $\mathcal{V}$ involves products of H^1, $H(Div)$, and $H(Curl)$, effective multigrid algorithms may also exist [6–8]. Moreover, the FOSLS functional, $\|L\mathbf{u}^h - \mathbf{f}\|_\tau$, computed on each element τ is a natrual error indicator that is locally sharp and globally reliable. An adaptive mesh refinement algorithm based on FOSLS has been applied to challenge problems such as problems with singularities, magneto-hydrodynamics (MHD) equations (Maxwell equations coupled with Navier-Stokes equations) *etc.* and is proved to be very successful (see [4, 9, 49] for details).

In conclusion, FOSLS has the advantages of being able to: yield an SPD linear system with optimal multigrid convergence rate; avoid the inf-sup condition for Stokes/Navier-Stokes equations and allow the use of H^1-conforming finite elements which are widely used and relatively easy to implement; achieve optimal finite element convegence rate and provide a nice error indicator for adaptive mesh refinement. Although the development of FOSLS formulation requires tremendous analysis skills, the efforts are well paid by its advantages.

2.2.2 FOSLL*

As mentioned earlier, the hybrid-FOSLS method is essentially "hybrid" FOSLS, FOSLL* and an intermediate term that draws them together. Therefore, in order to really understand the hybrid-FOSLS method, one has to first comprehend the basic theories of the FOSLL* method. In this section, the general approach of FOSLL* is introduced in (2.2.2.1); then, a simplified but very detailed discussion of the core theories of the FOSLL* that validate this method is presented in (2.2.2.2); finally, in (2.2.2.3), the applications of FOSLL* on different numerical problems and its pros and cons are summarized.

2.2.2.1 General Approach

The FOSLL* method was first introduced in [23] by Cai et. al. in 2000. It was motivated by

- the success of the FOSLS method in its advantages of resulting the optimal finite elments convergence rate using H^1-conforming elements; yielding an SPD system with optimal ($\mathcal{O}(n)$) multigrid convergence rate, where similar experience can also apply to FOSLL*;

- the limitation of standard L^2-norm FOSLS for problems with less smoothness, such as domain with re-entrant corners and discontinuous coefficients.

Other least-squares finite element methods such as inverse-norm versions of FOSLS can overcome the limitation of standard L^2-norm FOSLS. However, it requires the relatively awkward evaluation of the negative Sobolev norm and brings much larger computation cost ([16], [18], [22]).

The general approach of FOSLL* method can be summarized in the following three steps.

(1) **Based on the primal first-order system, formulate the adjoint first-order system.**

Suppose the primal first-order problem, $L\mathbf{u} = \mathbf{f}$, is given and let $\hat{\mathbf{u}}$ denote the exact solution. The adjoint problem is, thus, formulated as

$$\begin{aligned} LL^*\mathbf{w} &= \mathbf{f}, \\ L^*\mathbf{w} &= \hat{\mathbf{u}}. \end{aligned}$$

(2) **Seek weak solution by minimizing FOSLL* functional $\|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2$.**

That is to find $\hat{\mathbf{w}} \in \mathcal{W}$, such that

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathcal{W}} \|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2. \tag{2.35}$$

At the first glance, (2.35) looks problematic since it involves the unknown $\hat{\mathbf{u}}$. However, it is the same minimizatioin problem as minimizing the functional

$$\|L^*\mathbf{w}\|^2 - 2 < \mathbf{w}, \mathbf{f} >$$

and both of them lead to the weak problem:

$$< L^*\hat{\mathbf{w}}, L^*\mathbf{z} >=< \hat{\mathbf{u}}, L^*\mathbf{z} >=< \mathbf{f}, \mathbf{z} >, \quad \forall \mathbf{z} \in \mathcal{W}. \tag{2.36}$$

Notice that the weak problem can be solved without knowledge of the exact solution $\hat{\mathbf{u}}$.

(3) **Solve discretized weak form in Finite Element subspace $\mathcal{W}^h \subset \mathcal{W}$, recover the primal solution.**

Find $\mathbf{w}^h \in \mathcal{W}^h$, **s.t.**

$$< L^*\mathbf{w}^h, L^*\mathbf{z}^h >=< \mathbf{f}, \mathbf{z}^h >, \quad \forall \mathbf{z}^h \in \mathcal{W}^h.$$

Then, recover primal solution: $\mathbf{u}^h = L^*\mathbf{w}^h$.

2.2.2.2 Core Theories

In **Subsection** 2.2.2.1, the general approach of FOSLL* is introduced with a lot of details omitted. Taking a closer look at the FOSLL* approach, one would natrually ask questions such as: how is L^* defined? How do you guarantee for any solution $\mathbf{u}$, a $\mathbf{w}$ exsits, such that $L^*\mathbf{w} = \mathbf{u}$? Is such $\mathbf{w}$ unique? etc. These details are essential in validating the FOSLL* method and the subtle treatment when implementing FOSLL* on a numerical problem. To understand these details requires a good knowledge of functional analysis. This section is dedicated to explaining the core theories of FOSLL* in a simplied problem setting.

First, let's give a formal definition of L^*.

Theorem 3. *If A is a bounded linear operator on a Hilbert space $\mathcal{H}$, then there exits a unique bounded linear operator A^*, such that*

$$< x, Ay >_{\mathcal{H}}=< A^*x, y >_{\mathcal{H}}, \quad \forall x, y \in \mathcal{H}$$

Note that the inner product, $< \cdot, \cdot >_{\mathcal{H}}$, here depends on the Hilbert space.

Since in almost all the cases, L is a bounded linear operator on some Hilbert space, L^* that is defined as above is well defined.

Denote $\mathcal{D}(A)$ and $\mathcal{R}(A)$ the domain and range of a operator A. In order for FOSLL* to stand as a valid method, it must satisfy the following two requirements:

- L^* is onto $\mathcal{D}(L)$, that is, for all $\mathbf{u} \in \mathcal{D}(L)$, there exists a $\mathbf{w} \in \mathcal{D}(L^*)$, such that $\mathbf{u} = L^*\mathbf{w}$;

- such $\mathbf{w}$ should be unique.

The focus of the whole section is to establish a set of theorems and proofs and to validate that under certain assumptions, FOSLL* does satisfy the two requirements above.

We first explore under what assumptions, L^* is onto $\mathcal{D}(L)$.

Definition 1. ***(Closed operator)*** *Let $\mathcal{V}_1$, $\mathcal{V}_2$ be Banach spaces. A linear operator $A : \mathcal{D}(A) \in \mathcal{V}_1 \to \mathcal{V}_2$ is said to be closed if: whenever every sequence $\{x_n\}_{n\in\mathcal{N}} \in \mathcal{D}(A)$ converging to $x \in \mathcal{V}_1$ such that $\{Ax_n\}_{n\in\mathcal{N}}$ converge to $y \in \mathcal{V}_2$, it holds that $x \in \mathcal{D}$ and $Ax = y$.*

With the defination of closed operator, we are now ready to prove the following lemma:

Lemma 2. *Assume $L : \mathcal{V}_1 \to \mathcal{V}_2$ is a linear operator from Banach space $\mathcal{V}_1$ to $\mathcal{V}_2$ that are equipped with the norm $\|\cdot\|_{\mathcal{V}_1}$ and $\|\cdot\|_{\mathcal{V}_2}$ respectively. Denote the domain of L by $\mathcal{D}(L) \subset \mathcal{V}_1$ and the range of L by $\mathcal{R}(L) \subset \mathcal{V}_2$. If we have:*

- *L is a closed operator,*
- *L is coercive, that is, there exists a constant $c_0 > 0$, such that*

$$c_0\|\mathbf{u}\|_{\mathcal{V}_1} \leq \|L\mathbf{u}\|_{\mathcal{V}_2}, \quad \textit{for all } \mathbf{u} \in \mathcal{D}(L),$$

then we have that $\mathcal{R}(L)$ is closed in $\mathcal{V}_2$.

Proof. To prove $\mathcal{R}(L)$ is closed in $\mathcal{V}_2$, we only need to show that: for any Cauchy sequence $\{\mathbf{v}_n\} \subset \mathcal{R}(L)$, there exsits (a unique) $\mathbf{v}^* \in \mathcal{R}(L)$, such that

$$\lim_{n\to\infty} \|\mathbf{v}_n - \mathbf{v}^*\|_{\mathcal{V}_2} = 0.$$

First, by the coercivity of L, for each n, there exists a unique $\mathbf{u}_n \in \mathcal{D}(L)$, such that $L\mathbf{u}_n = \mathbf{v}_n$. (Otherwise, if there is an $\tilde{\mathbf{u}}_n$ different from $\mathbf{u}_n$, such that $L\tilde{\mathbf{u}}_n = \mathbf{v}_n$, then $\|\tilde{\mathbf{u}}_n - \mathbf{u}_n\|_{\mathcal{V}_1} > \|L\tilde{\mathbf{u}}_n - L\mathbf{u}_n\| = 0$, which violates the coercivity.)

Second, also by the coercivity of L, the unique set $\{\mathbf{u}_n\}$ is a Cauchy sequence in $\mathcal{V}_1$. This is because for any $\epsilon > 0$

$$c_0\|\mathbf{u}_m - \mathbf{u}_n\| \leq \|L\mathbf{u}_m - L\mathbf{u}_n\| = \|\mathbf{v}_m - \mathbf{v}_n\| < \epsilon,$$

when m and n are large enough.

Finally, since $\mathcal{V}_1$ and $\mathcal{V}_2$ are Banach spaces, there exists $\mathbf{u}^* \in \mathcal{V}_1$ and $\mathbf{v}^* \in \mathcal{V}_2$, such that $\mathbf{u}_n \to \mathbf{u}^*$, $\mathbf{v}_n \to \mathbf{v}^*$ in the norm $\|\cdot\|_{\mathcal{V}_1}$ and $\|\cdot\|_{\mathcal{V}_2}$ respectively. By the defination of closed operator, $\mathbf{u}^* \in \mathcal{D}(L)$ and $L\mathbf{u}^* = \mathbf{v}^*$. Thus, $\mathbf{v}^* \in \mathcal{R}(L)$ and the lemma is proved. □

The following closed range theorem will be used in the proof of Lemma 3. It can be found in many classical functional analysis books, for example, [54].

Theorem 4. ***(Closed range theorem)*** *Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be Banach spaces, A a closed linear operator, $A : \mathcal{D}(A) \to \mathcal{V}_2$, with its domain, $\mathcal{D}(A)$, dense in $\mathcal{V}_1$. Then, the following conditions are all equivalent*

- *$\mathcal{R}(A)$ is closed;*
- *$\mathcal{R}(A^*)$ is closed;*
- *$\mathcal{R}(A) = (\mathcal{N}(A^*))^\perp := \{y \in \mathcal{V}_2 :< x^*, y >= 0, \forall x^* \in \mathcal{N}(A^*)\}$;*
- *$\mathcal{R}(A^*) = (\mathcal{N}(A))^\perp := \{x^* \in \mathcal{V}_1' :< x^*, y >= 0, \forall y \in \mathcal{N}(A)\}$.*

Lemma 3. *Assume L satisfies all the assumptions in Lemma 2. In addtion, $\mathcal{D}(L)$ is dense in $\mathcal{V}_1$, then $\mathcal{R}(L)$ is closed in $\mathcal{V}_2$ if and only if $\mathcal{R}(L^*)$ is closed in $\mathcal{V}_1$.*

Proof. The proof is a direct result from Lemma 2 and the closed range theorem. □

With the help of Lemma 3, we can now apply Theorem 5 to prove that, under mild assumptions, L^* is onto $\mathcal{V}_1$.

Theorem 5. *(Fredholm) If A is a bounded linear operator on a Hilbert space $\mathcal{H}$, then*

$$\mathcal{H} = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*),$$

where, $\mathcal{R}(A)$ denotes the range of A and $\mathcal{N}(A^*)$ denotes the kernel of A^*.

Thus, if $\mathcal{R}(A)$ is closed, then $\mathcal{H} = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$.

Theorem 6. *Assume L satisfies all the assumptions in Lemma 3, i.e., L is coercive, closed and densely defined. Also assume the adjoint operator L^* is bounded on $\mathcal{V}_2$ and $(L^*)^* = L$. Without loss of generality, let $\mathcal{V}_1 = \mathcal{D}(L)$ and let $\mathcal{V}_1$ to be a Hilbert space. Then, L^* is surjective onto $\mathcal{V}_1$. That is, for all $\mathbf{u} \in \mathcal{V}_1$, there exists a $\mathbf{w} \in \mathcal{V}_2$, such that $L^*\mathbf{w} = \mathbf{u}$.*

Proof. The coercivity of L implies that the null space of L, $\mathcal{N}(L) = \{\mathbf{0}_{\mathcal{V}_1}\}$. By Lemma 3, since $\mathcal{R}(L^*)$ is closed, $\overline{\mathcal{R}(L^*)} = \mathcal{R}(L^*)$. Consider $A = L^*$ in Theorem 5 and notice that we have assumed $(L^*)^* = L$, we have

$$\mathcal{V}_1 = \overline{\mathcal{R}(L^*)} \oplus \mathcal{N}(L) = \mathcal{R}(L^*),$$

which finishes the proof. $\square$

To make sure that such $\mathbf{w}$ is unique, notice that by (5), $\mathcal{V}_2 = \mathcal{R}(L) \oplus \mathcal{N}(L^*)$, we restrict L^* to $\mathcal{R}(L)$, or alternatively consider L^* to be defined on the quotient space $\mathcal{R}(L)/\mathcal{N}(L^*)$. Thus, L^* is coercive on $\mathcal{V}_2$; that is, there exists a constant, $c^* > 0$, such that

$$c^* \|\mathbf{u}\|_{\mathcal{V}_2} \leq \|L^*\mathbf{w}\|_{\mathcal{V}_1}, \qquad \forall \mathbf{w} \in \mathcal{W} = \mathcal{D}(L^*) \cap \mathcal{R}(L). \tag{2.37}$$

The above discussion leads to the following assumption that will hold for the remainder of the dissertation.

Assumption 1. *Let $\mathcal{V}_1$ to be a Hilbert space and the domain of the linear operator L. Denote by $\mathbf{L}^2$ the $L^2(\Omega)$ space according to some domain Ω. Assume*

- *$(L^*)^* = L$,*
- *$L : \mathcal{V}_1 \to \mathbf{L}^2$ is a coercive, closed linear operator on $\mathbf{L}^2$,*
- *L^* is a bounded linear operator on $\mathbf{L}^2$.*

Remark 2. *In the assumption above, since we consider $\mathcal{D}(L)$ is a subspace of $\mathbf{L}^2$, the adjoint operator L^* is the $\mathbf{L}^2$ adjoint of L, **i.e.** the inner product in (3) is the $\mathbf{L}^2$ inner product.*

Remark 3. *In our applications, $\mathcal{V}_1$ is always a Hilbert space, such as $H(div;\Omega)$, $H(curl;\Omega)$ (see Appendix B.1). Also, to prove the existence and uniqueness of the weak solution, we always prove L is continuous and coercive in $\mathcal{V}_1$. Thus, the assumption that L is coercive on $\mathbf{L}^2$ is always satisfied, since we always have*

$$c_0\|\mathbf{u}\| \leq c_0\|\mathbf{u}\|_{\mathcal{V}_1} \leq \|L\mathbf{u}\|.$$

The details of obtaining the adjoint operator are presented in Appendix C.

2.2.2.3 Applications

In [23, 44],the FOSLL* method has been applied to general diffusion-convection-reaction problem

$$\begin{aligned} \nabla \cdot (A\nabla p) - \mathbf{b} \cdot \nabla p - cp &= f \qquad \text{in } \Omega, \\ p &= 0 \qquad \text{on } \Gamma_D, \\ \mathbf{n} \cdot A\nabla p &= 0 \qquad \text{on } \Gamma_N, \end{aligned}$$

where Γ_D and Γ_N denote the parts of boundary that is with Dirichlet boundary condition and Neumann boundary condition respectively, and A is allowed discontinuity and the domain Ω can be irregular **e.g.** with re-rentrant corners.

In [40], 3D eddy current problem

$$\begin{aligned} \frac{\partial \mu \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} &= 0 \qquad \text{in } \Omega, \\ \nabla \times \mathbf{H} - \sigma \mathbf{E} &= 0 \qquad \text{in } \Omega, \end{aligned}$$

with two boundary conditions

$$\mathbf{n} \times \mathbf{E} = 0, \mathbf{n} \cdot \mathbf{H} = 0,$$

or,

$$\mathbf{n} \cdot \mathbf{E} = 0, \mathbf{n} \times \mathbf{H} = 0.$$

on a domain with re-entrant corner is studied using FOSLL*.

It shows a much improved convegence rate than FOSLS method. In [31,32], mass conservation for Navier-Stokes Equations is highly enhanced by a two-stage combination of FOSLS-FOSLL* method.

Some heuristic explainations of the feature of FOSLL* are as follows.

- For FOSLS, if primal solution is not in H^1, using H^1-conforming finite elements (**e.g.** piecewise polynomials), discrete solution can never converge to the exact solution.
- On the other hand, for FOSLL*, because it introduces an adjoint first-order system $L^*\mathbf{w} = \hat{\mathbf{u}}$, such that $\hat{\mathbf{u}} \in H^1$, then $\mathbf{u} = L^*\hat{\mathbf{w}}$ is with lower smoothness.

However, FOSLL* has a major limitation in adaptive mesh refinement. Recall that for FOSLS method the error indicator for refinement is $\|L\mathbf{u}^h - \mathbf{f}\|_e$, which is computable, while the analogy of FOSLL* $\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\|_e$ is not computable.

2.3 FOSLS for Stokes/Navier-Stokes Equations

As mentioned ealier, generally, the first step of FOSLS method is to introduce new variables, possibly to add new equations and boundary conditions and recast the original system into a first-order differential system. This allows a great flexiblity and questions shall be asked:

- What new variables to introduce?
- What equations/boundary conditions to add?
- What norm to use for each residual equations?

Although the design of a good first-order system may be elusive and involves lots of "fail-and-try", there are still some general principles. First, optimally, we want to design an H^1-equivalent operator, that is a first-order operator L which satisfies, for any $u \in \mathcal{D}(L)$,

$$c_0\|\mathbf{u}\|_1 \leq \|L\mathbf{u}\| \leq c_1\|\mathbf{u}\|_1.$$

This implies an optimal finite element convergence rate and an optimal multigrid convergence rate. Second, the work of M. Gunzburger and P. Bochev [11, 14, 15] shows that with the help of Agmon-Douglis-Nirenberg (ADN) elliptic theory [5], for general elliptic partial differential equations (such as Stokes/Navier-Stokes equations in this context), the norm for each residual equations can be decided to guarantee the well-posedness of the associated weak problem and the existence of the a priori estimate.

We present several popular first-order formulations for Stokes and Navier-Stokes equations in this section.

2.3.1 Velocity/Velocity-Gradient/Pressure Formulation:

The *velocity-velocity gradient-pressure* system for Stokes equations which is product H^1-equivalent for a general dimension is first developed by Cai, Manteuffel, McCormick [21]. (For the convenience, we will refer to it as *velocity-gradient* formulation from now on.) We first introduce the extra variable: *velocity-gradient*

$$\underline{\mathbf{U}} = \nabla \mathbf{u}^t = (\nabla u_1, \nabla u_2) = \begin{pmatrix} U_{11} & U_{21} \\ U_{12} & U_{22} \end{pmatrix},$$

where U_11 denotes u_{1x}, U_12 denotes u_{1y} and so on. Notice that $\underline{\mathbf{U}}$ is a 2×2 tensor matrix in 2D and 3×3 matrix in 3D. Also notice that their nonstandard numbering of the elements of $\underline{\mathbf{U}}$. The first-order system of *velocity-gradient* formulation is

$$\begin{aligned} \underline{\mathbf{U}} - \nabla \mathbf{u}^t &= \underline{\mathbf{0}}, \quad \text{in } \Omega, \\ -(\nabla \cdot \underline{\mathbf{U}})^t + \nabla p &= \mathbf{f}, \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= g, \quad \text{in } \Omega, \\ \nabla (tr\underline{\mathbf{U}}) &= \nabla g, \quad \text{in } \Omega, \\ \nabla \times \underline{\mathbf{U}} &= \underline{\mathbf{0}}, \quad \text{in } \Omega, \end{aligned} \tag{2.38}$$

where $tr\underline{\mathbf{U}} := \mathbf{U}_{11} + \mathbf{U}_{22}$, stands for the trace of the matrix $\underline{\mathbf{U}}$. The fourth and fifth equations in (2.38) are auxiliary equations derived from the first and second equations in (2.38). Notice that

$tr\underline{\mathbf{U}} = \nabla \cdot \mathbf{u}$ and $\nabla \times \nabla \mathbf{u} = 0$ (see the appendix for the details). Thus, it is an overdetermined system, but with consistent extra equations. By adding these extra equations, (2.38) becomes a H^1 equivalent first-order system [21].

According to the FOSLS formulation above, the FOSLS functional we are minimizing is

$$\begin{aligned}\mathcal{F}(\underline{\mathbf{U}}, \mathbf{u}, p; \mathbf{f}, g) &= \|\mathbf{f} + (\nabla \cdot \underline{\mathbf{U}})^t - \nabla p\|^2 + \|\underline{\mathbf{U}} - \nabla \mathbf{u}^t\|^2 + \|\nabla \times \underline{\mathbf{U}}\|^2 \\ &\quad + \|\nabla \cdot \mathbf{u} - g\|^2 + \|\nabla tr(\underline{\mathbf{U}}) - \nabla g\|^2\end{aligned}$$

It has been proved in [21] that, under general H^2-regularity assumptions and with boundary conditions that are smooth enough, $\mathcal{F}(\underline{\mathbf{U}}, \mathbf{u}, p; \mathbf{0}, 0)$ is continuous and coercive in the product H^1 norm:

$$\|\underline{\mathbf{U}}\|_1^2 + \|\mathbf{u}\|_1^2 + \|p\|_1^2.$$

2.3.2 Velocity/Vorticity/Pressure Formulation

Another popular first-order formulation for the Stokes equation make use of the vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ and the fact that

$$\nabla \times \nabla \times \mathbf{u} = -\Delta \mathbf{u} + \nabla(\nabla \cdot \mathbf{u}).$$

Several formulations of this kind have been developed for their own merits.

The first-order system

$$\begin{aligned}\nu \nabla \times \boldsymbol{\omega} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \nabla \times \mathbf{u} - \boldsymbol{\omega} &= 0 \quad \text{in } \Omega \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega\end{aligned} \tag{2.39}$$

is H^1-elliptic with nonstandard boundary condition:

Normal Velocity-Pressure Boundary Conditions

$$\text{2D: } \mathbf{u} \cdot \mathbf{n} = 0 \text{ and } p = 0, \text{ on } \partial\Omega$$

$$\text{3D: } \mathbf{u} \cdot \mathbf{n} = 0, \quad \boldsymbol{\omega} \cdot \mathbf{n} = 0, \text{ and } p = 0, \text{ on } \partial\Omega$$

or **Normal Velocity-Tangential Vorticity Boundary Conditions**

$$\text{2D: } \mathbf{u} \cdot \mathbf{n} = 0 \text{ and } \omega = 0, \text{ on } \partial\Omega$$

$$\text{3D: } \mathbf{u}\cdot\mathbf{n} = 0\text{and} \quad \boldsymbol{\omega}\times\mathbf{n} = 0, \text{ on } \partial\Omega$$

where, the associated FOSLS functional is

$$\mathcal{F}(\boldsymbol{\omega}, \mathbf{u}, p; \mathbf{f}, g) = \|\nu\nabla\times\boldsymbol{\omega} + \nabla p - \mathbf{f}\|^2 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|^2 + \|\nabla\cdot\mathbf{u}\| \tag{2.40}$$

Unfortunately, this H^1-equivalence fails for standard velocity boundary conditions: **Standard Velocity Boundary Conditions:**

$$\begin{aligned} \mathbf{u} &= \mathbf{g}, \qquad \partial\Omega \\ \int_\Omega p d\Omega &= 0 \end{aligned}$$

The functional we need to minize for this boundary condition is thus

$$\mathcal{F}(\boldsymbol{\omega}, \mathbf{u}, p; \mathbf{f}, g) = \|\nu\nabla\times\boldsymbol{\omega} + \nabla p - \mathbf{f}\|_{-1}^2 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|^2 + \|\nabla\cdot\mathbf{u}\| \tag{2.41}$$

Another vorticity formulation involves a consistent extra term $\nabla\cdot\omega = 0$ and a slack variable ϕ which makes the system to be a square system that fits the ADN setting.

$$\begin{aligned} \nu\nabla\times\boldsymbol{\omega} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \nabla\cdot\boldsymbol{\omega} &= 0 \quad \text{in } \Omega \\ \nabla\times\mathbf{u} - \boldsymbol{\omega} + \nabla\phi &= 0 \quad \text{in } \Omega \\ \nabla\cdot\mathbf{u} &= 0 \quad \text{in } \Omega \end{aligned} \tag{2.42}$$

2.3.3 Velocity/Stress/Pressure Formulation

Recall Eqn (2.9) that for imcompressible flows, the stress tensor is

$$\underline{\boldsymbol{\sigma}} = \eta(\nabla\mathbf{u} + \nabla\mathbf{u}^t).$$

Define the scaled stress tensor

$$\underline{\mathbf{T}} = \sqrt{2\nu}\boldsymbol{\sigma},$$

and notice

$$\nabla\cdot\underline{\mathbf{T}} = \frac{\sqrt{2\nu}}{2}(\Delta\mathbf{u} + \nabla(\nabla\cdot\mathbf{u})),$$

the Velocity-Stree-Pressure formulation is as follows:

$$\begin{aligned} \sqrt{2\nu}\nabla \cdot \underline{\mathbf{T}} - \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega \\ \underline{\mathbf{T}} - \sqrt{2\nu}\boldsymbol{\sigma} &= 0 \quad \text{in } \Omega \end{aligned} \tag{2.43}$$

The according functional to minize for velocity-stress-pressure formulation is

$$F(\mathbf{u}, \underline{\mathbf{T}}, p; \mathbf{f}) + \|\sqrt{2\nu}\nabla \cdot \underline{\mathbf{T}} - \nabla p - \mathbf{f}\|_{-1}^2 + \|\nabla \cdot \mathbf{u}\|_0^2 + \|\underline{\mathbf{T}} - \sqrt{2\nu}\boldsymbol{\sigma}\|_0^2 \tag{2.44}$$

For more details, please refer to [11, 14, 15] and the reference therein.

Chapter 3

Hybrid-FOSLS for Stokes Equations in a Long Tube

3.1 Motivations

The first-order system least-squares (FOSLS) finite element method has been applied to numerical solution of a wide class of partial differential equations (e.g., [15, 19, 24]). The general approach of FOSLS is as follows: first reformulate the original system of PDEs into a possibly enlarged first-order differential system, $L\mathbf{u} = \mathbf{f}$. Then, a least-squares L^2-norm principle is applied to this first-order system, that is, we seek the weak solution by minimizing the FOSLS functional, $\mathcal{F}(\mathbf{u}; \mathbf{f}) := \|L\mathbf{u} - \mathbf{f}\|^2$, in an appropriate Hilbert space $\mathcal{V}$. (Here we assume homogeneous boundary conditions.) The choice of $\mathcal{V}$ depends on the particular problem, but is generally chosen based on the pre-image of L^2 under the operator L, that is, the set of functions, $\mathbf{u}$, such that $L\mathbf{u} \in L^2$ and satisfy appropriate boundary conditions and assumptions required by the FOSLS framework. Typical examples are products of H^1, $H(Div)$ and $H(Curl)$.

Discrete approximation is accomplished by restricting the minimization to an appropriate finite dimensional subspace, $\mathcal{V}^h \subset \mathcal{V}$.

An important property for the FOSLS framework is that the homogeneous functional, $\|L\mathbf{u}\|$, be coercive and continuous in the Hilbert space $\mathcal{V}$. That is, there exist constants, $0 < c_0 \leq c_1$, such that .

$$c_0\|\mathbf{u}\|_{\mathcal{V}} \leq \|L\mathbf{u}\| \leq c_1\|\mathbf{u}\|_{\mathcal{V}} \qquad \forall \mathbf{u} \in \mathcal{V}. \tag{3.1}$$

This guarantees a unique solution of the associated weak problem via the Riesz Representation theorem. Moreover, the discrete minimization inherits these properties and, thus, involves a symmetric positive definite (SPD) linear system. This also implies that the discrete system will be stable, so the Ladyzhenskaya-Babuška-Brezzi (LBB) stability condition is automatically satisfied and the finite element spaces can be chosen independently. Continuity of the functional implies that discrete error bounds may be established by standard interpolation bounds in $\mathcal{V}$. In addition, with sufficient regularity of L^*L, the L^2-norm of the error converges at an enhanced rate [42] (see also Section 3.2.2).

If $\mathcal{V}$ is a subspace of product H^1 spaces, we say the FOSLS functional is H^1- equivalent. In this case, the resulting linear system will be amenable to solution by a multigrid algorithm in $\mathcal{O}(n)$ operations [45]. If $\mathcal{V}$ involves products of H^1, $H(Div)$, and $H(Curl)$, effective multigrid algorithms may also exist [6–8].

Another important advantage is that the FOSLS functional provides a locally sharp, globally reliable, and easily computed error estimate for local mesh refinement (see [4, 9, 49] for details).

However, the FOSLS methodology also has some limitations. When the solution of the original problem is not smooth enough due to discontinuous coefficients, non-smooth domain boundary, or certain boundary conditions, H^1-equivalence and its associated properties may be lost. For example, certain components of the solution may be in $H(Div) \cap H(Curl)$ but not in H^1 [10, 29, 30]. The use of standard H^1-conforming finite element spaces will not result in convergence of the discrete approximation. This difficulty may be overcome by a number of remedies including weighted least squares functionals [39, 43], reformulation in $H(Div)$ or $H(Curl)$ and use of appropriate finite element spaces (e.g. Raviart-Thomas or Nedelec spaces), reformulation using inverse norms [15, 16], or the use of a FOSLL* formulation described below [23, 44].

Another shortcoming, one that is a motivation for this paper, occurs when the coercivity constant, c_0 in (3.1), is very small. This may be the result of problem coefficients, the structure of the system of PDEs, the shape of the domain, or all of these combined. In this case, the discrete approximation may make the relative functional norm small while the relative L^2 norm of the error remains large. More precisely, let the error be denoted by $\mathbf{e}^h = \mathbf{u}^h - \mathbf{u}$. Assume $\|e^h\| \leq \|e^h\|_{\mathcal{V}}$;

then, (3.1) implies

$$\frac{\|\mathbf{e}^h\|}{\|\mathbf{u}\|_{\mathcal{V}}} \leq \frac{\|\mathbf{e}^h\|_{\mathcal{V}}}{\|\mathbf{u}\|_{\mathcal{V}}} \leq \frac{c_1}{c_0}\frac{\|L\mathbf{e}^h\|}{\|L\mathbf{u}\|}. \tag{3.2}$$

If $\frac{c_1}{c_0} \gg 1$, the gap may be large. Despite the fact that, with full regularity, the L^2-norm of the error converges at an enhanced rate, the grid may require excessive refinement before this factor is overcome.

An example where this occurs is with the velocity-gradient/velocity/pressure formulation of Stokes equations in a long tube, described in more detail later. If D is the length of the tube, then $c_0 \sim 1/D^3$. Using standard bilinear finite elements yields small error in the H^1-seminorm but $\mathcal{O}(1)$ error in the L^2-norm until the grid is highly refined. This shortcoming can be mitigated by use of high-order finite elements [20], careful treatment of boundary conditions [34], or alternative formulations [31, 32].

To help the readers have a more intuitive feelings on how the L^2 error affects the solutions, we show the plots of pressure solving from the steady state Stokes equations 3.58 in a long tube $[0, 16] \times [0, 1]$ with both bilinear and biquadratic elements on different mesh grids ($h = 1/4, 1/8, 1/16, 1/32, 1/64$).

3.2 Theoretical Results

3.2.1 Hybrid and Graph Functional

As defined earlier, $\hat{\mathbf{u}}$ denotes for the exact solution of the primal problem, $L\mathbf{u} = \mathbf{f}$, and $\hat{\mathbf{w}}$ denotes the exact solution of the adjoint problem, $L^*\hat{\mathbf{w}} = \hat{\mathbf{u}}$. The errors in the primal and dual problems are $\mathbf{e} = \mathbf{u} - \hat{\mathbf{u}}$ and $\varepsilon = \mathbf{w} - \hat{\mathbf{w}}$, respectively. Let $\mathcal{V} = \mathcal{D}(L)$ be the domain of L equipped with the graph norm or any convenient equivalent norm. Likewise, let $\mathcal{W} = \mathcal{D}(L^*) \cap \mathcal{N}(L^*)^\perp$ equipped with its graph norm. The Hybrid functional on the Hilbert space $\mathcal{H} := \mathcal{W} \times \mathcal{V}$ is defined as follows:

$$\mathcal{H}\left((\mathbf{w}, \mathbf{u}); (\hat{\mathbf{u}}, \mathbf{f})\right) := \|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2 + \|L^*\mathbf{w} - \mathbf{u}\|^2 + \|L\mathbf{u} - \mathbf{f}\|^2. \tag{3.3}$$

The first term on the right-hand side is the **FOSLL* term**, the second the **intermediate term**, and the last the **FOSLS term**. The Hybrid functional can be seen as a measure of the error:

$$\mathcal{H}\left((\mathbf{w},\mathbf{u});(\hat{\mathbf{u}},\mathbf{f})\right) = \mathcal{H}\left((\epsilon,\mathbf{e});(0,0)\right) = \|L^*\epsilon\|^2 + \|L^*\epsilon - \mathbf{e}\|^2 + \|L\mathbf{e}\|^2. \tag{3.4}$$

Minimizing $\mathcal{H}\left((\mathbf{w},\mathbf{u});(\hat{\mathbf{u}},\mathbf{f})\right)$ induces the following weak problem: find $(\mathbf{w},\mathbf{u}) \in \mathcal{W}\times\mathcal{V}$, such that, $\forall(\mathbf{z},\mathbf{v}) \in \mathcal{W}\times\mathcal{V}$,

$$< L^*\mathbf{w}, L^*\mathbf{z} > + < L^*\mathbf{w}-\mathbf{u}, L^*\mathbf{z}-\mathbf{v} > + < L\mathbf{u}, L\mathbf{v} > = < \mathbf{f}, \mathbf{z} + L\mathbf{v} > . \tag{3.5}$$

The bilinear form associated with the Hybrid functional is the left-hand side of (3.5):

$$\mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) := < L^*\mathbf{w}, L^*\mathbf{z} > + < L^*\mathbf{w}-\mathbf{u}, L^*\mathbf{z}-\mathbf{v} > + < L\mathbf{u}, L\mathbf{v} > . \tag{3.6}$$

Also, denoting $\mathbf{F}\left((\mathbf{z},\mathbf{v})\right) := < \mathbf{f}, \mathbf{z} + L\mathbf{v} >$, then the weak problem becomes: find $(\mathbf{w},\mathbf{u}) \in \mathcal{W}\times\mathcal{V}$ such that, for all $(\mathbf{z},\mathbf{v}) \in \mathcal{W}\times\mathcal{V}$,

$$\mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) = \mathbf{F}\left((\mathbf{z},\mathbf{v})\right) \tag{3.7}$$

For convenience of discussion, define the $\mathcal{H}$-norm on $\mathcal{W}\times\mathcal{V}$ as

$$\|(\mathbf{w},\mathbf{u})\|_{\mathcal{H}} := \left(\|L^*\mathbf{w}\|^2 + \|\mathbf{u}\|^2 + \|L\mathbf{u}\|^2\right)^{1/2}. \tag{3.8}$$

This is a norm due to the coercivity bound (2.37) and Assumption 1. We now prove that $\mathcal{B}(\cdot,\cdot)$ is elliptic in the $\mathcal{H}$-norm.

Theorem 7. *For every* $(\mathbf{w},\mathbf{u})$, $(\mathbf{z},\mathbf{v}) \in \mathcal{W}\times\mathcal{V}$, *we have the coercivity bound*

$$\frac{1}{3}\|(\mathbf{w},\mathbf{u})\|_{\mathcal{H}}^2 \le \left(\|L^*\mathbf{w}\|^2 + \|L^*\mathbf{w}-\mathbf{u}\|^2 + \|L\mathbf{u}\|^2\right) = \mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{w},\mathbf{u})\right) \tag{3.9}$$

and the continuity bound

$$\mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) \le 3\|(\mathbf{w},\mathbf{u})\|_{\mathcal{H}}\|(\mathbf{z},\mathbf{v})\|_{\mathcal{H}}. \tag{3.10}$$

Proof. The continuity bound is obtained first by applying the Cauchy-Schwarz inequality to all three terms in (3.6) and the triangle inequality to the middle term to yield

$$\mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) \leq \|L^*\mathbf{w}\|\|L^*\mathbf{z}\| + \left(\|L^*\mathbf{w}\| + \|\mathbf{u}\|\right)\left(\|L^*\mathbf{z}\| + \|\mathbf{v}\|\right) + \|L\mathbf{u}\|\|L\mathbf{v}\|. \tag{3.11}$$

Next , treating this as an inner product in $\Re^3$ and using Cauchy-Schwarz inequality again yields

$$\begin{aligned} \mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) \;\; &\leq \;\; \left(\|L^*\mathbf{w}\|^2 + (\|L^*\mathbf{w}\| + \|\mathbf{u}\|)^2 + \|L\mathbf{u}\|^2\right)^{1/2} \\ &\quad \cdot \left(\|L^*\mathbf{z}\|^2 + (\|L^*\mathbf{z}\| + \|\mathbf{v}\|)^2 + \|L\mathbf{v}\|^2\right)^{1/2}. \end{aligned} \tag{3.12}$$

The final step comes by using the inequality

$$(\|L^*\mathbf{w}\| + \|\mathbf{u}\|)^2 \leq 2(\|L^*\mathbf{w}\|^2 + \|\mathbf{u}\|^2), \tag{3.13}$$

which yields

$$\begin{aligned} \mathcal{B}\left((\mathbf{w},\mathbf{u}),(\mathbf{z},\mathbf{v})\right) \;\; &\leq \;\; \left(3\|L^*\mathbf{w}\|^2 + 2\|\mathbf{u}\|^2 + \|L\mathbf{u}\|^2\right)^{1/2} \\ &\quad \cdot \left(3\|L^*\mathbf{z}\|^2 + 2\|\mathbf{v}\|^2 + \|L\mathbf{v}\|^2\right)^{1/2}. \end{aligned} \tag{3.14}$$

The bound (3.10) now follows.

The coercivity bound is established by starting with the triangle inequality

$$\|\mathbf{u}\| \leq \|L^*\mathbf{w} - \mathbf{u}\| + \|L^*\mathbf{w}\|. \tag{3.15}$$

Squaring both sides yields

$$\|\mathbf{u}\|^2 \leq \|L^*\mathbf{w} - \mathbf{u}\|^2 + 2\|L^*\mathbf{w} - \mathbf{u}\|\|L^*\mathbf{w}\| + \|L^*\mathbf{w}\|^2 \leq 2(\|L^*\mathbf{w} - \mathbf{u}\|^2 + \|L^*\mathbf{w}\|^2). \tag{3.16}$$

Adding $\|L^*\mathbf{w}\|^2 + \|L\mathbf{u}\|^2$ to both sides yields the final result. $\square$

Remark 4. *Since $\mathcal{H}\left((\epsilon,\mathbf{e});(0,0)\right) = \mathcal{B}\left((\epsilon,\mathbf{e});(\epsilon,\mathbf{e})\right)$, this theorem implies that*

$$\frac{1}{3}\|(\epsilon,\mathbf{e})\|_{\mathcal{H}}^2 \leq \mathcal{H}\left((\mathbf{w},\mathbf{u});(\hat{\mathbf{u}},\mathbf{f})\right) \leq 3\|(\epsilon,\mathbf{e})\|_{\mathcal{H}}^2 \tag{3.17}$$

Thus, $\|(\epsilon,\mathbf{e})\|_{\mathcal{H}}^2$ bounds the Hybrid functional from below with the simple constant $1/3$, which does not depend on the mesh size or domain diameter. In fact, (3.16) shows that $\|\mathbf{e}\|^2$ bounds the functional from below with constant $1/2$. Thus, the functional cannot be small unless the L^2-norm of the error in the primal problem is also small.

Remark 5. *By scaling the terms in (3.3) and modifying the definition (3.8) the ratio* c_1/c_0 *can be reduced further. We keep this definition and the factor* $\frac{1}{2}$ *in the definition of the Graph norm (3.25) for convenience of later proofs.*

Remark 6. *If* L *is designed, such that both* L *and* L^* *are* H^1*-equivalent, the* $\mathcal{H}$*-norm is equivalent to the* H^1*-norm. Thus,* $\mathcal{B}(\cdot,\cdot)$ *is also* H^1*-elliptic, and the advantageous properties of FOSLS carry on to hybrid-FOSLS.*

To gain insight, let $\mathcal{U} = (\mathbf{w}, \mathbf{u})$ and rewrite the system implicit in (3.3) as

$$\mathcal{L}\mathcal{U} := \begin{bmatrix} L^* & \\ L^* & -I \\ & L \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{u}} \\ 0 \\ f \end{bmatrix} = \mathbf{F} \tag{3.18}$$

The formal normal equations associated with this system are

$$\mathcal{L}^*\mathcal{L}\mathcal{U} = \begin{bmatrix} 2LL^* & -L \\ -L^* & L^*L + I \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} f \\ L^*f \end{bmatrix} = \mathcal{L}^*\mathbf{F}. \tag{3.19}$$

Lemma 4. *If both* L *and* L^* *are coercive and continuous in* H^1*, then both* L^*L *and* LL^* *have full* H^2 *regularity and* $\mathcal{L}^*\mathcal{L}$ *also has full* H^2 *regularity.*

Proof. The proof is straightforward. □

3.2.2 Convergence Estimates

This section addresses convergence of finite element approximation of the Hybrid functional. We begin with the assumptions on the finite element spaces.

Assumption 2. *Finite element spaces* $\mathcal{V}^h \subset \mathcal{V}$ *and* $\mathcal{W}^h \subset \mathcal{W}$ *satisfy standard interpolation bounds: given* $\mathbf{u} \in \mathcal{V}$ *and* $\mathbf{w} \in \mathcal{W}$*,*

$$\inf_{\mathbf{u}^h \in \mathcal{V}^h} \|\mathbf{u}^h - \mathbf{u}\| \leq Ch^{r+1}\|\mathbf{u}\|_{r+1}, \tag{3.20}$$

$$\inf_{\mathbf{u}^h \in \mathcal{V}^h} \|L(\mathbf{u}^h - \mathbf{u})\| \leq Ch^{r}\|\mathbf{u}\|_{r+1}, \tag{3.21}$$

$$\inf_{\mathbf{w}^h \in \mathcal{W}^h} \|L^*(\mathbf{w}^h - \mathbf{w})\| \leq Ch^{s}\|\mathbf{w}\|_{s+1}, \tag{3.22}$$

for $0 < r \leq p$ and $0 < s \leq q$ where $C > 0$ is a generic constant, h is a mesh parameter, and p and q are the degree of the finite element polynomials for $\mathcal{V}^h$ and $\mathcal{W}^h$, respectively.

Discrete approximation of the Hybrid functional is achieved by restricting the minimization of the Hybrid functional (3.3) to $\mathcal{V}^h \times \mathcal{W}^h$. The above assumptions yield the following convergence bound. Let $(\mathbf{w}^h, \mathbf{u}^h) \in \mathcal{W}^h \times \mathcal{V}^h$ minimize the Hybrid functional. Then,

$$\mathcal{H}((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f}))^{1/2} \leq C_1 h^r \|\hat{\mathbf{w}}\|_{r+1} + C_2 h^s \|\hat{\mathbf{u}}\|_{s+1} \tag{3.23}$$

for $0 < r \leq q$, $0 < s \leq p$.

Next, enhanced convergence for the Hybrid functional in the L^2 norm is established.

Theorem 8. *Assume that both L^*L and L^*L have full H^2 regularity. If $(\mathbf{w}^h, \mathbf{u}^h)$ is the solution of discrete weak form (3.5), then*

$$\|((\mathbf{w}^h - \hat{\mathbf{w}}), (\mathbf{u}^h - \hat{\mathbf{u}}))\| \leq Ch\mathcal{H}((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f}))^{1/2}. \tag{3.24}$$

Proof. The proof follows from Lemma 4 and the Aubin-Nitsche trick (c.f. [**?**], sec 7.7). □

Remark 7. *For more precise results involving partial regularity, the reader is directed to [42].*

3.2.3 Graph Norm Estimates

To explore the relation between the Hybrid functional and graph norm, define the graph functional on $\mathcal{V}$,

$$\mathcal{G}(\mathbf{u}; \hat{\mathbf{u}}) := \frac{1}{2}\|\mathbf{u} - \hat{\mathbf{u}}\|^2 + \|L\mathbf{u} - \mathbf{f}\|^2, \tag{3.25}$$

and define graph norm on $\mathcal{V}$,

$$\|\mathbf{u}\|_{\mathcal{G}} := \left(\frac{1}{2}\|\mathbf{u}\|^2 + \|L\mathbf{u}\|^2\right)^{\frac{1}{2}} = \mathcal{G}(\mathbf{u}; 0)^{\frac{1}{2}}. \tag{3.26}$$

(The $\frac{1}{2}$ appears as a convenience to the proofs below.) Denote

$$\tilde{\mathbf{u}}^h = \arg\min_{\mathbf{u}^h \in \mathcal{V}^h} \mathcal{G}(\mathbf{u}^h, \hat{\mathbf{u}}) = \arg\min_{\mathbf{u}^h \in \mathcal{V}^h} \left(\frac{1}{2}\|(\mathbf{u}^h - \hat{\mathbf{u}})\|^2 + \|L(\mathbf{u}^h - \hat{\mathbf{u}})\|^2\right), \tag{3.27}$$

which implies, for every $\mathbf{v}^h \in \mathcal{V}^h$,

$$< L(\tilde{\mathbf{u}}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > + \frac{1}{2} < (\tilde{\mathbf{u}}^h - \hat{\mathbf{u}}), \mathbf{v}^h > = \ 0. \tag{3.28}$$

The following lemma shows that, if $\mathcal{V}^h \subseteq L^*(\mathcal{W}^h)$, then the minimizer of the Hybrid functional also minimizes the graph functional.

Lemma 5. *If, for every $\mathbf{v}^h \in \mathcal{V}^h$, there exists $\mathbf{z}^h \in \mathcal{W}^h$ such that $\mathbf{v}^h = L^*\mathbf{z}^h$, then minimizing Hybrid functional, $\mathcal{H}\left((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f})\right)$, is equivalent to minimizing graph functional, $\mathcal{G}(\mathbf{u}; \hat{\mathbf{u}})$, **i.e.** $\mathbf{u}^h = \tilde{\mathbf{u}}^h$.*

Proof. Minimizing the Hybrid functional over the space $\mathcal{W}^h \times \mathcal{V}^h \subset \mathcal{W} \times \mathcal{V}$ yields a weak problem equivalent to (3.5). That is, $(\mathbf{w}^h, \mathbf{u}^h) \in \mathcal{V}^h \times \mathcal{W}^h$ satisfies, $\forall (\mathbf{z}^h, \mathbf{v}^h) \in \mathcal{W}^h \times \mathcal{V}^h$,

$$\begin{aligned} < L^*\mathbf{w}^h - \hat{\mathbf{u}}, L^*\mathbf{z}^h > + < L^*\mathbf{w}^h - \mathbf{u}^h, L^*\mathbf{z}^h > &= 0, \\ < L\mathbf{u}^h - \mathbf{f}, L\mathbf{v}^h > + < \mathbf{u}^h - L^*\mathbf{w}^h, \mathbf{v}^h > &= 0. \end{aligned} \tag{3.29}$$

Recall $L^*\hat{\mathbf{w}} = \hat{\mathbf{u}}$ and $L\hat{\mathbf{u}} = f$ and rearrange (3.29) to obtain

$$\begin{aligned} 2 < L^*(\mathbf{w}^h - \hat{\mathbf{w}}), L^*\mathbf{z}^h > &= < \mathbf{u}^h - \hat{\mathbf{u}}, L^*\mathbf{z}^h > \\ < L(\mathbf{u}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > + < \mathbf{u}^h - \hat{\mathbf{u}}, \mathbf{v}^h > &= < L^*(\mathbf{w}^h - \hat{\mathbf{w}}), \mathbf{v}^h > . \end{aligned} \tag{3.30}$$

Now, it is easy to see if, $\forall \mathbf{v}^h \in \mathcal{V}^h$, there exists $\mathbf{z}^h \in \mathcal{W}^h$, such that $\mathbf{v}^h = L^*\mathbf{z}^h$, we can plug the left side of the first equation of (3.30) into the right side of the second equation of (3.30) and get:

$$< L(\mathbf{u}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > + \frac{1}{2} < \mathbf{u}^h - \hat{\mathbf{u}}, \mathbf{v}^h >= 0. \tag{3.31}$$

Comparing with (3.28) yields $\mathbf{u}^h = \tilde{\mathbf{u}}^h$, which proves the lemma. $\square$

Notice that the weak form induced by minimizing the graph functional involves the exact solution, $\hat{\mathbf{u}}$, which is not computable, while the weak form associated with Hybrid functional is computable.

Next, we show that, when $\mathcal{V}^h$ is not contained in $L^*(\mathcal{W}^h)$, $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}}$ converges faster than the Hybrid functional. This result requires an addition regularity assumption.

Theorem 9. *Assume that LL^* has full H^2-regularity. If $(\mathbf{w}^h, \mathbf{u}^h)$ is the solution of discrete weak form (3.29), then*

$$\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}} \leq Ch\mathcal{H}\left((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f})\right)^{\frac{1}{2}} \leq C_1 h^{r+1}\|\hat{\mathbf{w}}\|_{r+1} + C_2 h^{s+1}\|\hat{\mathbf{u}}\|_{s+1}, \tag{3.32}$$

for $0 < r \leq q$ and $0 < s \leq p$, where $C_1, C_2 > 0$ are generic constants.

Proof. From (3.30), it is easy to see that the weak problem induced by minimizing the Hybrid functional can also be written as

$$\begin{aligned} < L^*(\mathbf{w}^h - \hat{\mathbf{w}}) - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}}), L^*\mathbf{z}^h > &= 0, \quad \forall \mathbf{z}^h \in \mathcal{W}^h, & (3.33)\\ < L(\mathbf{u}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > + < \mathbf{u}^h - L^*\mathbf{w}^h, \mathbf{v}^h > &= 0, \quad \forall \mathbf{v}^h \in \mathcal{V}^h. & (3.34)\end{aligned}$$

The weak problem associated with minimizing the graph norm is:

$$< L(\tilde{\mathbf{u}}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > + \frac{1}{2} < \tilde{\mathbf{u}}^h - \hat{\mathbf{u}}, \mathbf{v}^h >= 0, \quad \forall \mathbf{v}^h \in \mathcal{V}^h \tag{3.35}$$

Subtract (3.35) from (3.34) to obtain, $\forall \mathbf{v}^h \in \mathcal{V}^h$,

$$< L(\mathbf{u}^h - \tilde{\mathbf{u}}^h), L\mathbf{v}^h > + \frac{1}{2} < \mathbf{u}^h - \tilde{\mathbf{u}}^h, \mathbf{v}^h >=< L^*\mathbf{w}^h - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}}), \mathbf{v}^h > \tag{3.36}$$

Let $\mathbf{v}^h = \mathbf{u}^h - \tilde{\mathbf{u}}^h$ and use (3.33) to obtain, $\forall \mathbf{z}^h \in \mathcal{W}^h$,

$$\begin{aligned} \|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|^2 + \frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2 &= < L^*\mathbf{w}^h - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}}), \mathbf{u}^h - \tilde{\mathbf{u}}^h > \\ &= < L^*\mathbf{w}^h - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}}), \mathbf{u}^h - \tilde{\mathbf{u}}^h - L^*\mathbf{z}^h > \\ &\leq \|L^*\mathbf{w}^h - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}})\| \cdot \|(\mathbf{u}^h - \tilde{\mathbf{u}}^h) - L^*\mathbf{z}^h\| \end{aligned}$$

Estimate the two terms of the right hand side above separately. For the first term, add and subtract $L^*\hat{\mathbf{w}} = \hat{\mathbf{u}}$ and use the triangle inequality, and the coercivity bound (3.9) , to obtain

$$\begin{aligned} \|L^*\mathbf{w}^h - \frac{1}{2}(\mathbf{u}^h + \hat{\mathbf{u}})\|^2 &= \frac{1}{4}\|(L^*\mathbf{w}^h - \hat{\mathbf{u}}) + (L^*\mathbf{w}^h - \mathbf{u}^h)\|^2 \\ &\leq \frac{1}{4}\left(\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\| + \|L^*\mathbf{w}^h - \mathbf{u}^h\|\right)^2 \\ &\leq \frac{1}{2}\mathcal{H}\left((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f})\right) \end{aligned}$$

For the second term, let $L^*\mathbf{z} = \mathbf{u}^h - \tilde{\mathbf{u}}^h$, then

$$\begin{aligned}
\inf_{\mathbf{z}^h \in \mathcal{W}^h} \|(\mathbf{u}^h - \tilde{\mathbf{u}}^h) - L^*\mathbf{z}^h\| &\leq \inf_{\mathbf{z}^h \in \mathcal{W}^h} \|L^*(\mathbf{z} - \mathbf{z}^h)\| \\
&\leq \inf_{\mathbf{z}^h \in \mathcal{W}^h} C\|\mathbf{z} - \mathbf{z}^h\|_1 \leq Ch\|\mathbf{z}\|_2 \\
&\leq Ch\|LL^*\mathbf{z}\| = Ch\|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\| \\
&\leq Ch\left(\|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|^2 + \frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2\right)^{1/2}.
\end{aligned}$$

Note that we get the inequalities above by applying continuity of L^*, finite element interpolation estimates and the full H^2 regularity of LL^*. The proof is completed by combining the estimates of the two terms and using (3.23) . □

Remark 8. *If we choose $\mathcal{W}^h$ to be of the same order polynomials as $\mathcal{V}^h$, that is let $p = q$, then $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}}$ is of convergence order $\mathcal{O}(h^{p+1})$.*

3.2.4 Convergence of Superposition with Nested Iteration

Using superposition with nested iteration (**i.e.** use the approximation from the previous coarser grid, $\mathbf{u}^{2h}$, as superposition function), will remarkably accelerate the convergence rate of $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|$. Instead of solving the Hybrid first order system

$$\begin{aligned}
L\mathbf{u}^h &= \mathbf{f} \\
L^*\mathbf{w}^h &= \mathbf{u}^h \\
L^*\mathbf{w}^h &= \hat{\mathbf{u}},
\end{aligned} \tag{3.37}$$

we solve

$$\begin{aligned}
L\mathbf{u}^h &= \mathbf{f} \\
L^*\mathbf{w}^h &= \delta\mathbf{u}^h = \mathbf{u}^h - \mathbf{u}^{2h} \\
L^*\mathbf{w}^h &= \delta\hat{\mathbf{u}} = \hat{\mathbf{u}} - \mathbf{u}^{2h}.
\end{aligned} \tag{3.38}$$

The corresponding minimization problem now becomes: find $(\mathbf{w}^h, \mathbf{u}^h) \in \mathcal{W}^h \times \mathcal{V}^h$ such that

$$(\mathbf{w}^h, \mathbf{u}^h) = arg\min \|L^*\mathbf{w}^h - (\hat{\mathbf{u}} - \mathbf{u}^{2h})\|^2 + \|L^*\mathbf{w}^h - (\mathbf{u}^h - \mathbf{u}^{2h})\| + \|L\mathbf{u}^h - \mathbf{f}\|^2. \tag{3.39}$$

The following theorem is the analytically explanation of why superposition achieves higher order convergence.

Theorem 10. *Assume that LL^* has full H^2-regularity. If $(\mathbf{w}^h, \mathbf{u}^h)$ is obtained by minimizing the Hybrid functional with superposition, (3.39), then*

$$\begin{aligned} \|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}} &\leq Ch^2\left(\tfrac{1}{2}\|\mathbf{u}^h - \hat{\mathbf{u}}\|_1 + \|\mathbf{u}^{2h} - \hat{\mathbf{u}}\|_1\right) \\ &\leq Ch^2\mathcal{H}((\mathbf{w}^h, \mathbf{u}^h);\ (\hat{\mathbf{u}}, f))^{\frac{1}{2}}. \end{aligned} \tag{3.40}$$

Proof. Similar to (3.33) and (3.34), the weak form from of (3.39) is:

$$< L^*\mathbf{w}^h - \frac{1}{2}\left((\mathbf{u}^h - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right), L^*\mathbf{z}^h >= 0, \qquad \forall \mathbf{z}^h \in \mathcal{W}^h, \tag{3.41}$$

$$< L\mathbf{u}^h - \mathbf{f}, L\mathbf{v}^h > + < \mathbf{u}^h - \mathbf{u}^{2h}, \mathbf{v}^h >=< L^*\mathbf{w}^h, \mathbf{v}^h >, \qquad \forall \mathbf{v}^h \in \mathcal{V}^h. \tag{3.42}$$

Rearranging (3.42) yields, $\forall \mathbf{v} \in \mathcal{V}^h$,

$$\begin{aligned} < L(\mathbf{u}^h - \hat{\mathbf{u}}), L\mathbf{v}^h > &+ \frac{1}{2} < \mathbf{u}^h - \hat{\mathbf{u}}, \mathbf{v}^h > \\ &= < L^*\mathbf{w}^h - \frac{1}{2}\left((\mathbf{u}^h - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right), \mathbf{v}^h > \end{aligned} \tag{3.43}$$

Subtract the weak form associated with the graph functional, (3.28), from (3.42) and apply (3.41) to obtain, $\forall \mathbf{v}^h \in \mathcal{V}^h, \mathbf{z}^h \in \mathcal{W}^h$,

$$\begin{aligned} < L(\mathbf{u}^h - \tilde{\mathbf{u}}^h), L\mathbf{v}^h > &+ \frac{1}{2} < \mathbf{u}^h - \tilde{\mathbf{u}}^h, \mathbf{v}^h > \\ &= < L^*\mathbf{w}^h - \frac{1}{2}\left((\mathbf{u}^h - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right), \mathbf{v}^h > \\ &= < L^*\mathbf{w}^h - \frac{1}{2}\left((\mathbf{u}^h - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right), \mathbf{v}^h - L^*\mathbf{z}^h > . \end{aligned} \tag{3.44}$$

Set $\mathbf{v}^h = \mathbf{u}^h - \tilde{\mathbf{u}}^h$ to get, $\forall \mathbf{z}^h \in \mathcal{W}^h$,

$$\begin{aligned} &\frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2 + \|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|^2 \\ &\qquad = < L^*\mathbf{w}^h - \frac{1}{2}\left((\mathbf{u} - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right), (\mathbf{u}^h - \tilde{\mathbf{u}}^h) - L^*\mathbf{z}^h > . \end{aligned} \tag{3.45}$$

We next consider two auxiliary problems. Find $\mathbf{w} \in \mathcal{W}$, such that

$$L^*\mathbf{w} = \frac{1}{2}\left((\mathbf{u}^h - \mathbf{u}^{2h}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\right) = \frac{1}{2}(\mathbf{u}^h - \hat{\mathbf{u}}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h}). \tag{3.46}$$

Find $\mathbf{z} \in \mathcal{W}$, such that

$$L^*\mathbf{z} = \mathbf{u}^h - \tilde{\mathbf{u}}^h \tag{3.47}$$

Equation (3.41) implies, for $0 < s \le q$,

$$\begin{aligned}
\|L^*(\mathbf{w}^h - \mathbf{w})\| &= \inf_{\mathbf{w}^h \in \mathcal{W}^h} \|L^*(\mathbf{w}^h - \mathbf{w})\| \le \|L^*(\mathcal{I}^h\mathbf{w} - \mathbf{w})\| && (3.48)\\
&\le C\|\mathcal{I}^h\mathbf{w} - \mathbf{w}\|_1 \le Ch^s\|\mathbf{w}\|_{s+1} && (3.49)\\
&= Ch^s\|\frac{1}{2}(\mathbf{u}^h - \hat{\mathbf{u}}) + (\hat{\mathbf{u}} - \mathbf{u}^{2h})\|_s && (3.50)\\
&\le Ch^s\left(\frac{1}{2}\|\mathbf{u}^h - \hat{\mathbf{u}}\|_s + \|\mathbf{u}^{2h} - \hat{\mathbf{u}}\|_s\right) && (3.51)
\end{aligned}$$

We also get

$$\begin{aligned}
\inf_{\mathbf{z}^h \in \mathcal{W}^h} \|(\mathbf{u}^h - \tilde{\mathbf{u}}^h) - L^*\mathbf{z}^h\| &= \inf_{\mathbf{z}^h \in \mathcal{W}^h} \|L^*(\mathbf{z}^h - \mathbf{z})\| \le \|L^*(\mathcal{I}\mathbf{z}^h - \mathbf{z})\| && (3.52)\\
&\le C\|\mathcal{I}^h\mathbf{z} - \mathbf{z}\|_1 \le Ch\|\mathbf{z}\|_2 && (3.53)\\
&\le Ch\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_1 && (3.54)\\
&\le Ch\left(\frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2 + \|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|^2\right)^{1/2}. && (3.55)
\end{aligned}$$

Plugging (3.48) and (3.52) into (3.45) yields

$$\left(\frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2 + \|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|^2\right)^{1/2} \le Ch^{s+1}\left(\frac{1}{2}\|\mathbf{u}^h - \hat{\mathbf{u}}\|_s + \|\mathbf{u}^{2h} - \hat{\mathbf{u}}\|_s\right). \tag{3.56}$$

Let $s = 1$ to obtain

$$\left(\frac{1}{2}\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|^2 + \|L(\mathbf{u}^h - \tilde{\mathbf{u}}^h)\|\right)^{1/2} \le Ch^2\left(\frac{1}{2}\|\mathbf{u}^h - \hat{\mathbf{u}}\|_1 + \|\mathbf{u}^{2h} - \hat{\mathbf{u}}\|_1\right). \tag{3.57}$$

The assumption of regularity implies the result. Alternatively, one may set $s = 0$ in (3.56) and use enhanced L^2 convergence from Theorem 8. □

3.3 Numerical Results

In this section, we study cases where FOSLS formulation has limitations, while hybrid-FOSLS performs well. We give both numerical results and some analysis to explain the observed numerical phenomena. Our first set of numerical tests are for steady state Stokes equations in a long thin tube with width equal to 1 as in Figure (5.1), where E, S, W, N denote east, south, west, north boundary, respectively, and D denotes the domain length.

The equations are:

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= g \quad \text{in } \Omega, \end{aligned} \tag{3.58}$$

with boundary conditions:

$$\begin{aligned} &\mathbf{n} \times \mathbf{u} = 0, \mathbf{n} \cdot \mathbf{u} = b, \quad [W]; \\ &\mathbf{n} \times \mathbf{u} = 0, \mathbf{n} \cdot \mathbf{u} = 0, \quad [N, S]; \\ &\mathbf{n} \times \mathbf{u} = 0, \mathbf{n} \cdot \sigma \mathbf{n} = 0, \quad [E]; \end{aligned} \tag{3.59}$$

where p is pressure normalized by viscosity, $\mathbf{u} = (u_1, u_2)^t$ is velocity and $\sigma := -pI + \frac{1}{2}\left(\nabla \mathbf{u} + (\nabla \mathbf{u})^t\right)$ is the stress tensor. Right-hand side $\mathbf{f}$ is a given vector function, g and b are given scalar functions, and $\mathbf{n}$ is outward unit normal vector. Throughout this paper, all vectors are column vectors; ∇ always takes a scalar function to a column vector function; $\nabla\cdot$, $\nabla\times$ always operate on column vectors.

The second order Stokes equations (3.58) are recast to an equivalent first-order system, the **velocity/velocity-gradient/pressure** formulation, by introducing an additional variable, **velocity-gradient**

$$\underline{\mathbf{U}} = \nabla \mathbf{u}^t = (\nabla u_1, \nabla u_2) = \begin{pmatrix} U_{11} & U_{21} \\ U_{12} & U_{22} \end{pmatrix}.$$

Notice that $\underline{\mathbf{U}}$ is a 2×2 tensor matrix in 2D. Notice also the nonstandard numbering of the elements of $\underline{\mathbf{U}}$.

The first-order system is

$$\begin{aligned} \underline{\mathbf{U}} - \nabla \mathbf{u}^t &= \underline{\mathbf{0}}, \quad \text{in } \Omega, \\ -(\nabla \cdot \underline{\mathbf{U}})^t + \nabla p &= \mathbf{f}, \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= g, \quad \text{in } \Omega, \\ \nabla (tr\underline{\mathbf{U}}) &= \nabla g, \quad \text{in } \Omega, \\ \nabla \times \underline{\mathbf{U}} &= \underline{\mathbf{0}}, \quad \text{in } \Omega, \end{aligned} \tag{3.60}$$

with boundary conditions

$$\begin{array}{lll} U_{22} = 0, \quad U_{12} = \frac{\partial b}{\partial y} & & [W] \\ U_{11} = 0, \quad U_{21} = 0 & & [N, S] \\ U_{22} = 0, \quad U_{11} = p & & [E] \end{array} \tag{3.61}$$

where $tr\underline{\mathbf{U}} := \mathbf{U}_{11} + \mathbf{U}_{22}$, stands for the trace of the matrix $\underline{\mathbf{U}}$. The fourth and fifth equations in (3.60) are auxiliary equations derived from the first and second equations in (3.60). Thus, they are consistent with the original system. By adding these extra equations, (3.60) becomes a H^1 equivalent first-order system [21]. Boundary conditions for $\underline{\mathbf{U}}$ and p are directly derived from original boundary conditions. From now on, we use **velocity-gradient** in short for **velocity/velocity-gradient/pressure** formulation.

According to the FOSLS formulation above, the FOSLS functional we are minimizing is

$$\begin{aligned} \mathcal{F}(\underline{\mathbf{U}}, \mathbf{u}, p; \mathbf{f}, g) &= \|\mathbf{f} + (\nabla \cdot \underline{\mathbf{U}})^t - \nabla p\|^2 + \|\underline{\mathbf{U}} - \nabla \mathbf{u}^t\|^2 + \|\nabla \times \underline{\mathbf{U}}\|^2 \\ &\quad + \|\nabla \cdot \mathbf{u} - g\|^2 + \|\nabla tr(\underline{\mathbf{U}}) - \nabla g\|^2 \end{aligned}$$

It has been proved in [21] that, under general H^2-regularity assumptions and with boundary conditions that are smooth enough, $\mathcal{F}(\underline{\mathbf{U}}, \mathbf{u}, p; \mathbf{0}, 0)$ is continuous and coercive in the product H^1 norm:

$$\|\underline{\mathbf{U}}\|_1^2 + \|\mathbf{u}\|_1^2 + \|p\|_1^2.$$

In our numerical tests, we use a two-stage velocity-gradient formulation:

Stage 1: solve for $\underline{\mathbf{U}}$ and p,

$$\begin{aligned} -(\nabla \cdot \underline{\mathbf{U}})^t + \nabla p &= \mathbf{f} \\ \nabla \times \underline{\mathbf{U}} &= \mathbf{0} \\ \nabla (tr\underline{\mathbf{U}}) &= \nabla g. \end{aligned} \tag{3.62}$$

Stage 2: solve for $\mathbf{u}$, with $\underline{\mathbf{U}}$ computed from the first stage as a known variable,

$$\begin{aligned} \nabla \mathbf{u} &= \underline{\mathbf{U}} \\ \nabla \cdot \mathbf{u} &= g. \end{aligned} \tag{3.63}$$

At the first stage, we minimize

$$\mathcal{F}^{(1)}(\underline{\mathbf{U}}, p; \mathbf{f}, g) = \|\mathbf{f} + (\nabla \cdot \underline{\mathbf{U}})^t - \nabla p\|^2 + \|\nabla \times \underline{\mathbf{U}}\|^2 + \|\nabla tr(\underline{\mathbf{U}}) - \nabla g\|^2$$

and, at the second stage, minimize the functional

$$\mathcal{F}^{(2)}(\mathbf{u}; \underline{\mathbf{U}}, g) = \|\nabla \mathbf{u}^t - \underline{\mathbf{U}}\|^2 + \|\nabla \cdot \mathbf{u} - g\|^2.$$

In fact, $\mathcal{F}^{(1)}(\underline{\mathbf{U}}, p; \mathbf{0}, 0)$ is uniformly equivalent to $\|\underline{\mathbf{U}}\|_1^2 + \|p\|_1^2$ and $\mathcal{F}^{(2)}(\mathbf{u}; \underline{\mathbf{0}}, 0)$ is uniformly equivalent to $\|\mathbf{u}\|_1^2$. This implies that solving the Stokes equations in a two-stage process yields the same optimal finite element convergence as solving them simultaneously.

To simplify the first-order system and to better enforce the conservation constraint, $tr\underline{\mathbf{U}} = U_{11} + U_{22} = \nabla \cdot \mathbf{u} = 0$, the constraint $U_{22} = -U_{11}$ is enforced throughout this paper. Thus, the simplified primal first-order system is:

$$L \begin{pmatrix} U_{11} \\ U_{12} \\ U_{21} \\ p \end{pmatrix} = \begin{bmatrix} \partial_x & \partial_y & 0 & -\partial_x \\ -\partial_y & 0 & \partial_x & -\partial_y \\ -\partial_y & \partial_x & 0 & 0 \\ -\partial_x & 0 & -\partial_y & 0 \end{bmatrix} \begin{pmatrix} U_{11} \\ U_{12} \\ U_{21} \\ p \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = \mathbf{F}, \tag{3.64}$$

with the following boundary conditions,

$$\begin{aligned} U_{11} &= 0 \quad [W, N, S, E] \\ U_{12} &= 0 \quad [W] \\ U_{21} &= 0 \quad [N, S] \\ p &= 0 \quad [E]. \end{aligned} \tag{3.65}$$

We specify exact solutions and apply the differential operator to obtain the right hand side, $\mathbf{f}$, as follows:

The adjoint first-order system is obtained from the primal first-order system by integration by parts, such that for any $\mathbf{w} \in \mathcal{W}$, $\mathbf{u} \in \mathcal{V}$, $< L^*\mathbf{w}, \mathbf{u} >=< \mathbf{w}, L\mathbf{u} >$. This implies all the boundary terms coming from integration by parts must be zero to make the equality hold. Applying

Given Solution	Right Hand Side
$U_{11} = \sin(\frac{\pi x}{D})\sin(\pi y)$,	$f1 = \pi\sin(\pi y)[\cos(\frac{\pi x}{D})/D - \sin(\frac{\pi x}{D})] + 2$,
$U_{12} = \sin(\frac{\pi x}{D})\cos(\pi y)$,	$f2 = -\pi\sin(\frac{\pi x}{D})[\cos(\pi y) + \sin(\pi y)/D]$,
$U_{21} = \cos(\frac{\pi x}{D})\sin(\pi y)$,	$f3 = \pi\cos(\pi y)[\cos(\frac{\pi x}{D})/D - \sin(\frac{\pi x}{D})]$,
$p = 2(D - x)$	$f4 = -\pi\cos(\frac{\pi x}{D})[\cos(\pi y) + \sin(\pi y)/D]$

homogeneous boundary conditions for primal variable $\mathbf{u}$ (where, in all our numerical tests, $\mathbf{u}$ here is the solution from superposition), we get boundary constraints for adjoint variable, $\mathbf{w}$. The adjoint first-order system of (3.64) is:

$$L^* \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{bmatrix} -\partial_x & \partial_y & \partial_y & \partial_x \\ -\partial_y & 0 & -\partial_x & 0 \\ 0 & -\partial_x & 0 & \partial_y \\ \partial_x & \partial_y & 0 & 0 \end{bmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} U_{11} \\ U_{12} \\ U_{21} \\ p \end{pmatrix}, \tag{3.66}$$

with boundary conditions

$$\begin{aligned} w_1 &= 0 \quad [W, N, S] \\ w_2 &= 0 \quad [W, N, S, E] \\ w_3 &= 0 \quad [E]. \end{aligned} \tag{3.67}$$

With both the primal and adjoint systems, we are ready to minimize the Hybrid functional (3.3) and solve the linear system computed from (3.5).

Consider the limitations of FOSLS in this context. The H^1-ellipticity of FOSLS only implies

$$c_0\|\mathbf{e}^h\|_1 \leq \|L\mathbf{u}^h - \mathbf{f}\| = \|L\mathbf{e}^h\| \leq c_1\|\mathbf{e}^h\|_1. \tag{3.68}$$

Ideally, if $c_0 = O(1)$, minimizing $\|L\mathbf{u}^h - \mathbf{f}\|$ over $\mathcal{V}^h$ is essentially equivalent to minimizing $\|\mathbf{e}^h\|_1$. However, if this is not the case, for example, if coercivity constant c_0 is very small, it is possible for $\|\mathbf{e}^h\|_1 \gg \|L\mathbf{u}^h - \mathbf{f}\|$. While $\|L\mathbf{u}^h - \mathbf{f}\|$ is very small, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|_1$ can still stay large.

The following numerical results are for Stokes equations in a long tube and show how the FOSLS functional and L^2 error converge with domain lengths D equal to 16, 24, 32. To exclude the affect that the the domain length has on the size of the exact solution, the relative FOSLS functional, $\|L\mathbf{u}^h - \mathbf{f}\|/\|f\|$, and relative L^2 error, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|/\|\hat{\mathbf{u}}\|$, are examined in Figure [3.7].

Figure 3.7 displays the relative FOSLS functional and relative L^2 error on a sequence of uniformly refined grids. Here, bilinear elements are used and, as expected, the FOSLS functional converges with $O(h)$. The relative L^2 error is essentially not reduced until the grid is highly refined. This phenomenon grows more pronounced as the tube becomes longer. Note that the rate of convergence of the L^2 error eventually becomes enhanced, as predicted by Theorem 8.

The question now is under what conditions will the coercivity constant, c_0, be small? How small it can be? We cannot precisely answer the question. However, for the **velocity-gradient** form of Stokes equations, the following example demonstrates $c_0 \leq \mathcal{O}(D^{-3})$.

Denote $\mathcal{U} = (U_{11}, U_{12}, U_{21}, p)^t$. Manipulating the Stokes operator L yields

$$L \begin{pmatrix} U_{11} \\ U_{12} \\ U_{21} \\ p \end{pmatrix} = \begin{bmatrix} \partial_x & \partial_y & 0 & -\partial_x \\ -\partial_y & 0 & \partial_x & -\partial_y \\ -\partial_y & \partial_x & 0 & 0 \\ -\partial_x & 0 & -\partial_y & 0 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ -1 & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ 1 & & & 1 \end{bmatrix} \begin{pmatrix} U11 \\ U12 \\ U21 \\ p \end{pmatrix}$$

$$= \begin{bmatrix} 2\partial_x & \partial_y & 0 & -\partial_x \\ 0 & 0 & \partial_x & -\partial_y \\ -\partial_y & \partial_x & 0 & 0 \\ -\partial_x & 0 & -\partial_y & 0 \end{bmatrix} \begin{pmatrix} U_{11} \\ U_{12} \\ U_{21} \\ p + U_{11} \end{pmatrix}$$

Define $\tilde{p} = p + U_{11}$ and let $f(s)$ be any smooth function. Specify $\mathcal{U}$ and apply the differential operator to obtain the right hand side, $\mathbf{F}$, as follows:

$$\begin{array}{rclcrcl} & \text{Given Solution:} & & & & \text{Right Hand Side} & \\ U_{11} & = & \frac{y(1-y)}{D^2} f''(\frac{x}{D}), & & f1 & = & 2\frac{y(1-y)}{D^3} f'''(\frac{x}{D}), \\ U_{12} & = & \frac{(1-2y)}{D} f'(\frac{x}{D}), & & f2 & = & 0, \\ U_{21} & = & 0, & & f3 & = & 0, \\ \tilde{p} & = & -2f(\frac{x}{D}) & & f4 & = & -\frac{y(1-y)}{D^3} f'''(\frac{x}{D}) \end{array}$$

Clearly, for this choice of $\mathcal{U}$,

$$c_0 \leq \frac{\|L\mathcal{U}\|}{\|\mathcal{U}\|} \leq \frac{C}{D^3}, \tag{3.69}$$

where constant $C > 0$ is independent of D.

Unlike FOSLS, the hybrid-FOSLS functional controls both the L^2 error and FOSLS functional very well. Recall that hybrid-FOSLS is elliptic in $\mathcal{H}$-norm, which implies

$$\frac{1}{3}\|(\epsilon, \mathbf{e})\|_{\mathcal{H}} \leq \mathcal{H}\left((\mathbf{w}, \mathbf{u}); (\hat{\mathbf{u}}, \mathbf{f})\right) \leq 3\|(\epsilon, \mathbf{e})\|_{\mathcal{H}},$$

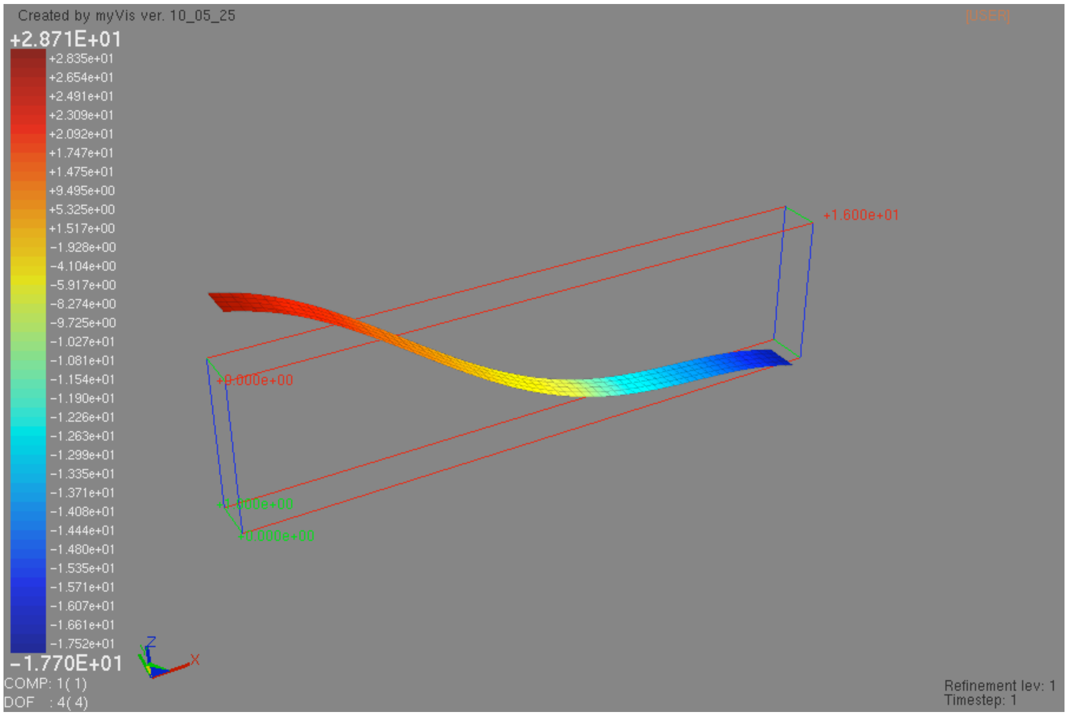
where, $\|(\epsilon, \mathbf{e})\|_{\mathcal{H}}^2 = \|L^*\epsilon\|^2 + \|\mathbf{e}\|^2 + \|L\mathbf{e}\|^2$. Because the continuity and coercivity constants are mild and independent of the size of the domain, if the hybrid-FOSLS functional is small, the L^2 error and FOSLS functional cannot be large.

This is demonstrated in Figure 3.8 where results for FOSLS alone are superimposed with results for hybrid-FOSLS. FOSLS alone yields the FOSLS functional in green and the L^2 error in red. All three terms of the hybrid-FOSLS results are displayed. Note that the FOSLS term computed by FOSLS alone (green) and computed by the hybrid-FOSLS (blue) lie almost on top of each other. However, the L^2 error computed by FOSLS alone (red) is much larger than the L^2 error computed by hybrid-FOSLS (black). This becomes more pronounced as the tube gets longer.

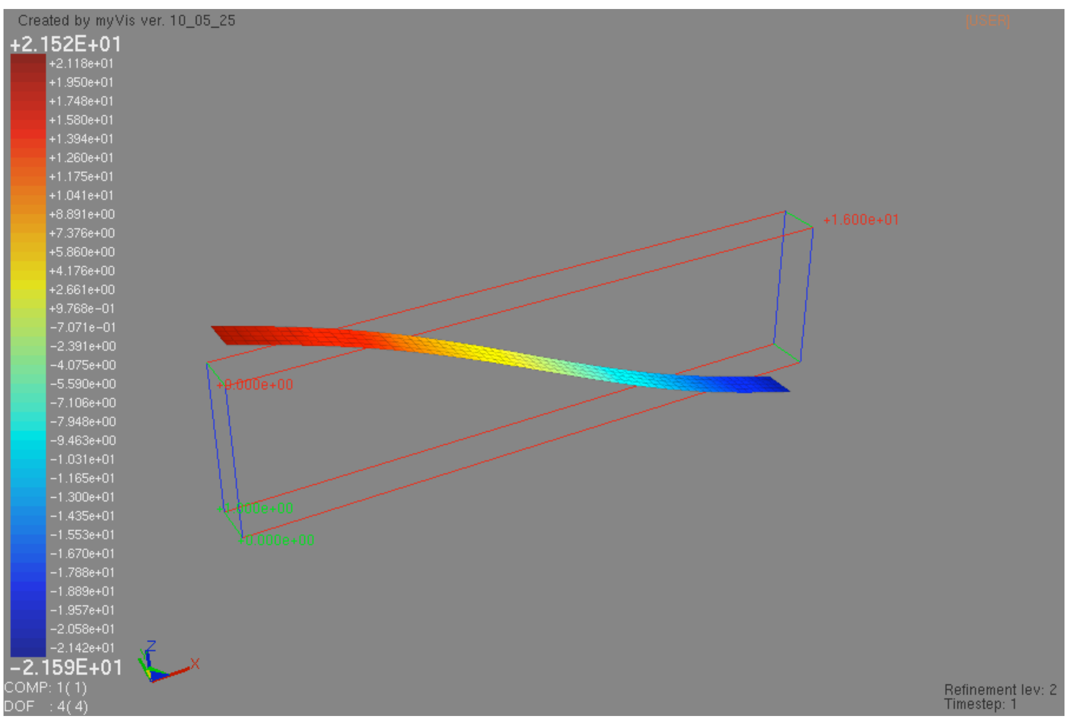
Finally, Theorem 9 predicts that, with superposition, the solution computed by hybrid-FOSLS approaches $\tilde{\mathbf{u}}^h$, the optimal finite element solution in the graph norm, much faster than the Hybrid functional converges to zero. To be precise, Theorem 10 predicts

$$\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}} \leq Ch^2 \left(\|L^*(\mathbf{w}^h - \hat{\mathbf{w}})\|^2 + \|\mathbf{u}^h - \hat{\mathbf{u}}\|^2 + \|L(\mathbf{u}^h - \hat{\mathbf{u}})\|^2\right)^{1/2}. \tag{3.70}$$

Figure 3.9 displays $\|\mathbf{u}^h - \hat{\mathbf{u}}\|_{\mathcal{G}}$ and $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_{\mathcal{G}}$ on a sequence of uniformly refined grids using biquadratic elements. The former converges with $O(h^2)$, as predicted, while the latter converges with $O(h^6)$, which is faster than the the predicted $O(h^4)$. This may be due to the very smooth exact solution, $\hat{\mathbf{u}}$.

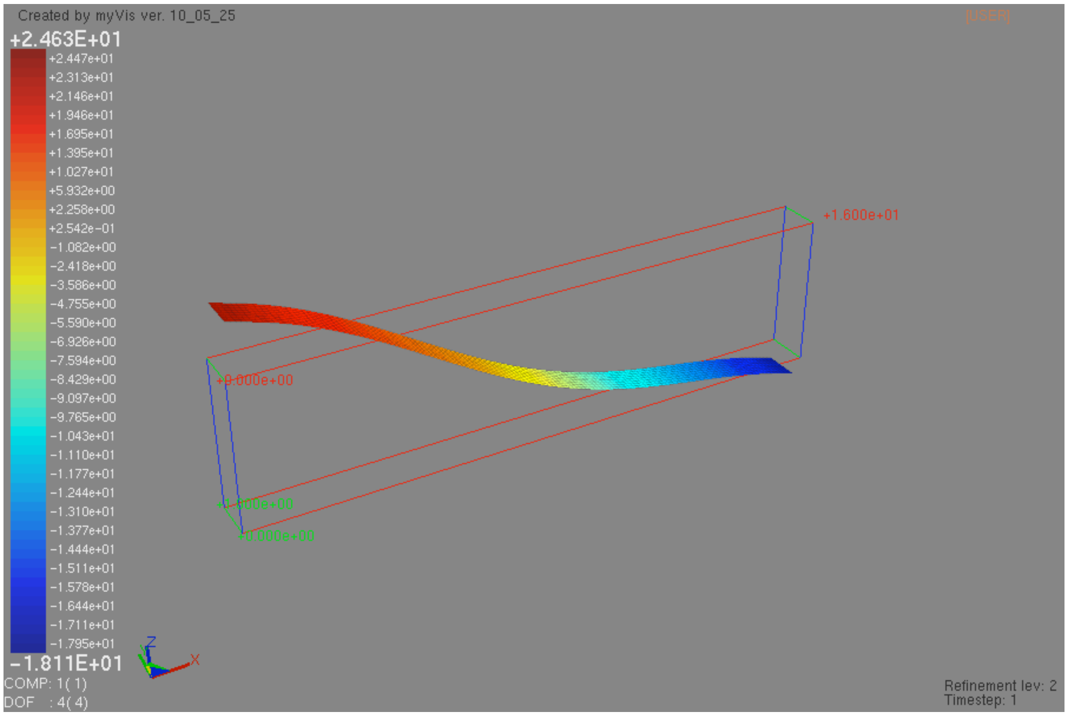


(a) bilinear

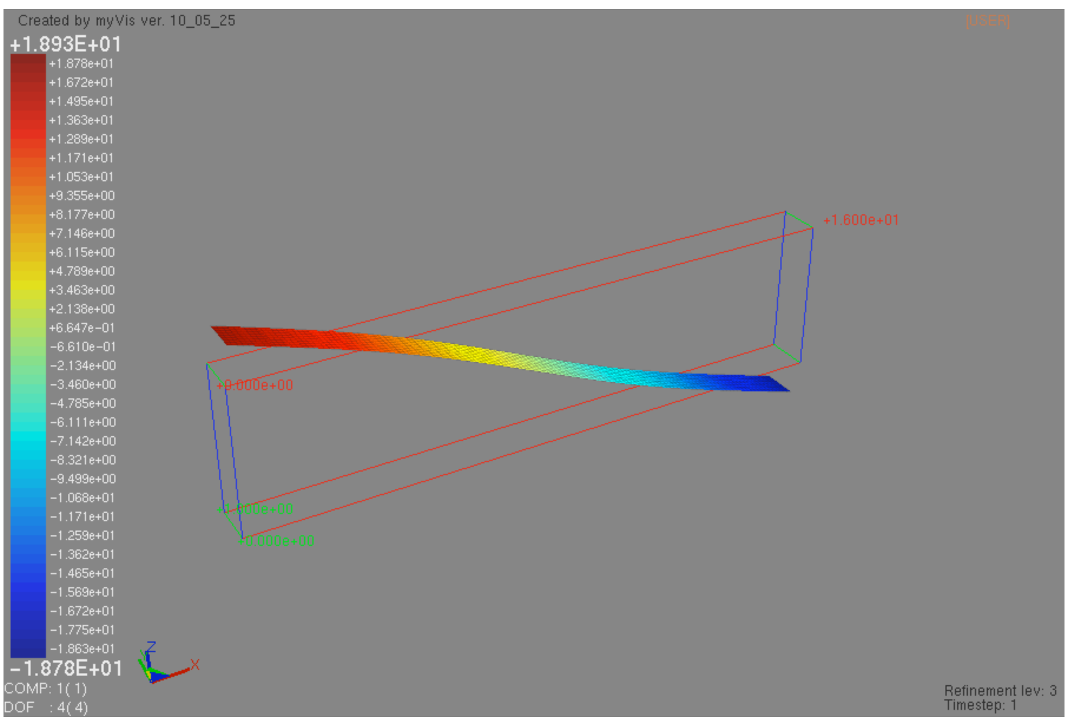


(b) biquadratic

Figure 3.1: Plot of Pressure from Stokes Equations: $D = 16$, Bilinear vs. Biquadratic Elements $h = 1/4$

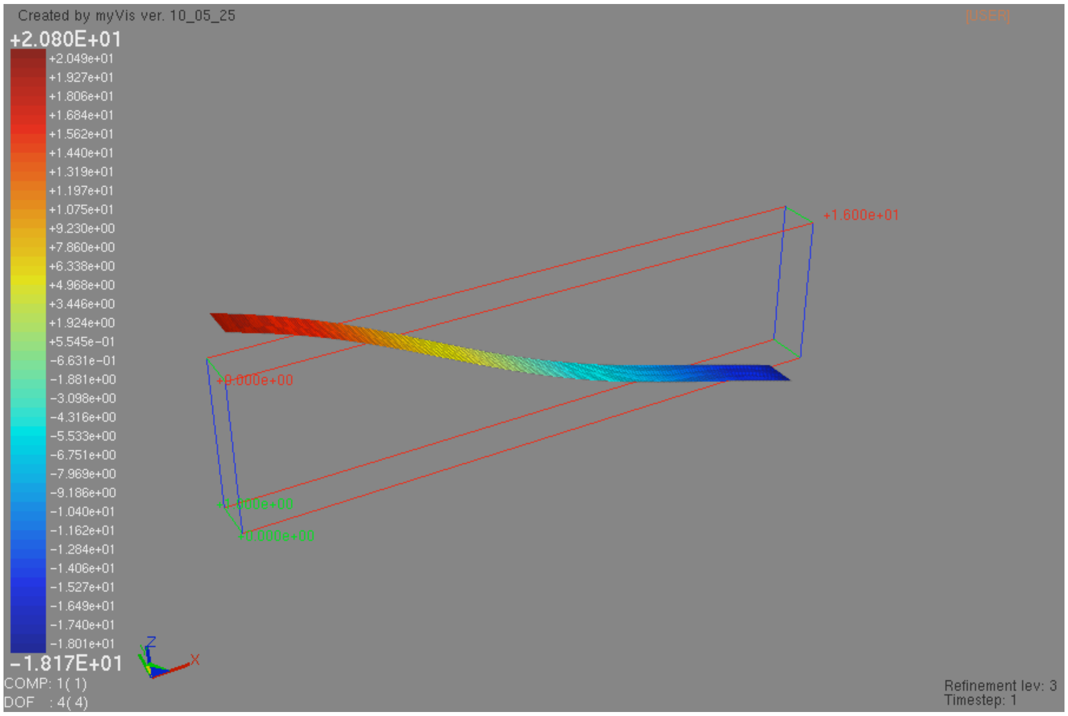


(a) bilinear

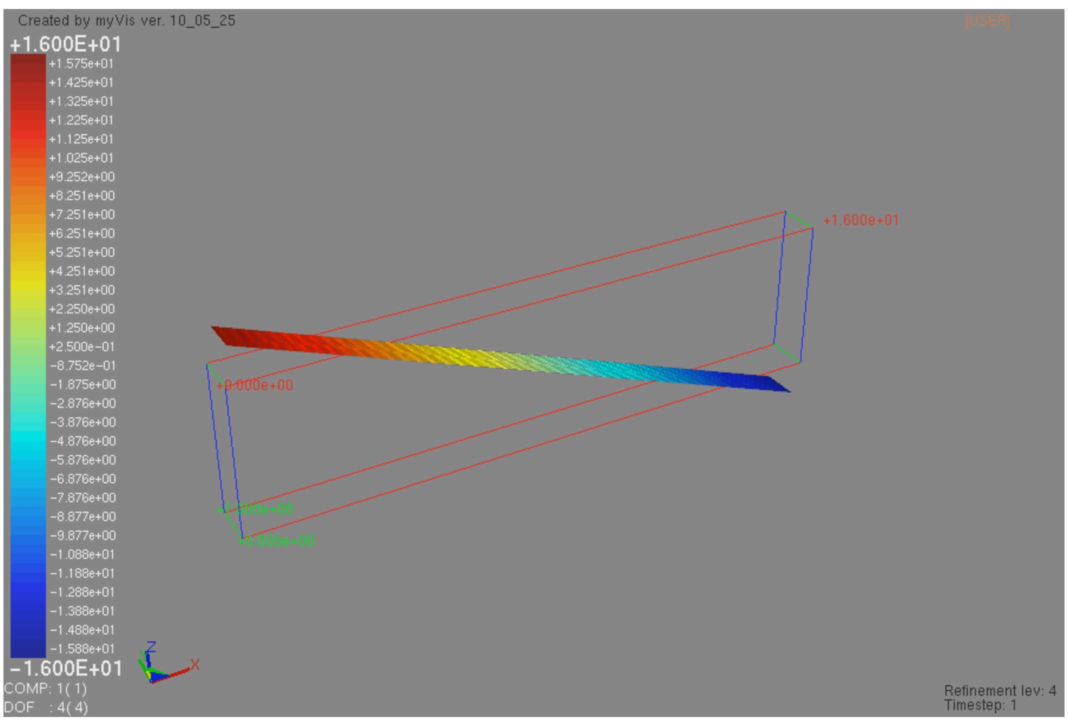


(b) biquadratic

Figure 3.2: Plot of Pressure from Stokes Equations: $D = 16$, Bilinear Elements $h = 1/8$

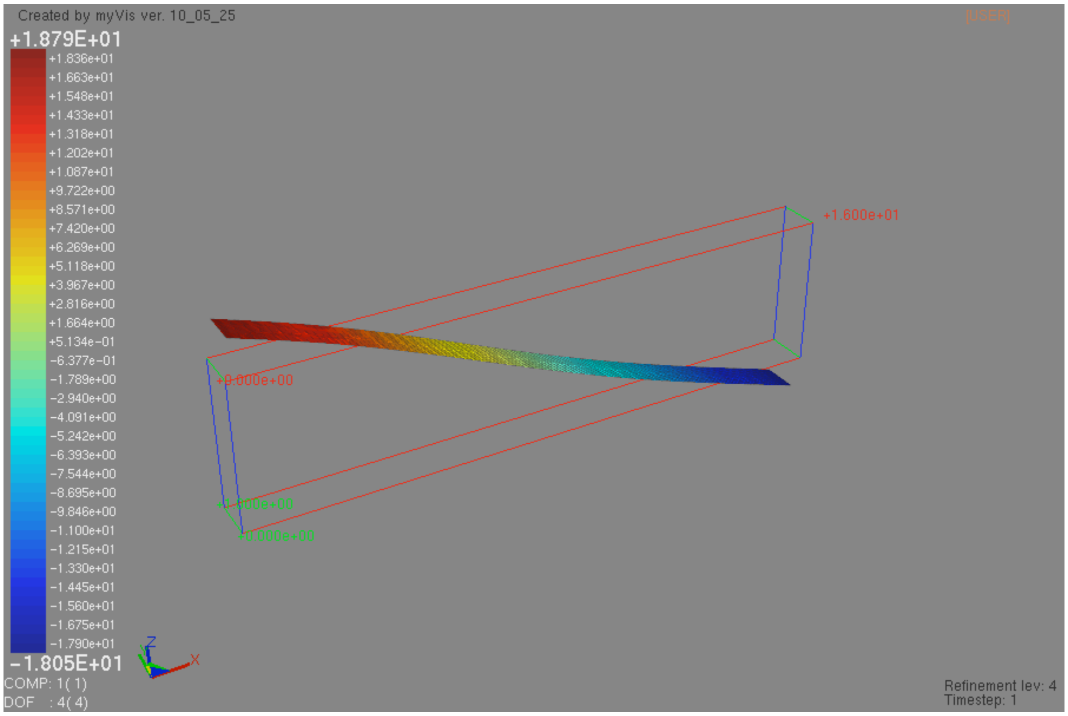


(a) bilinear

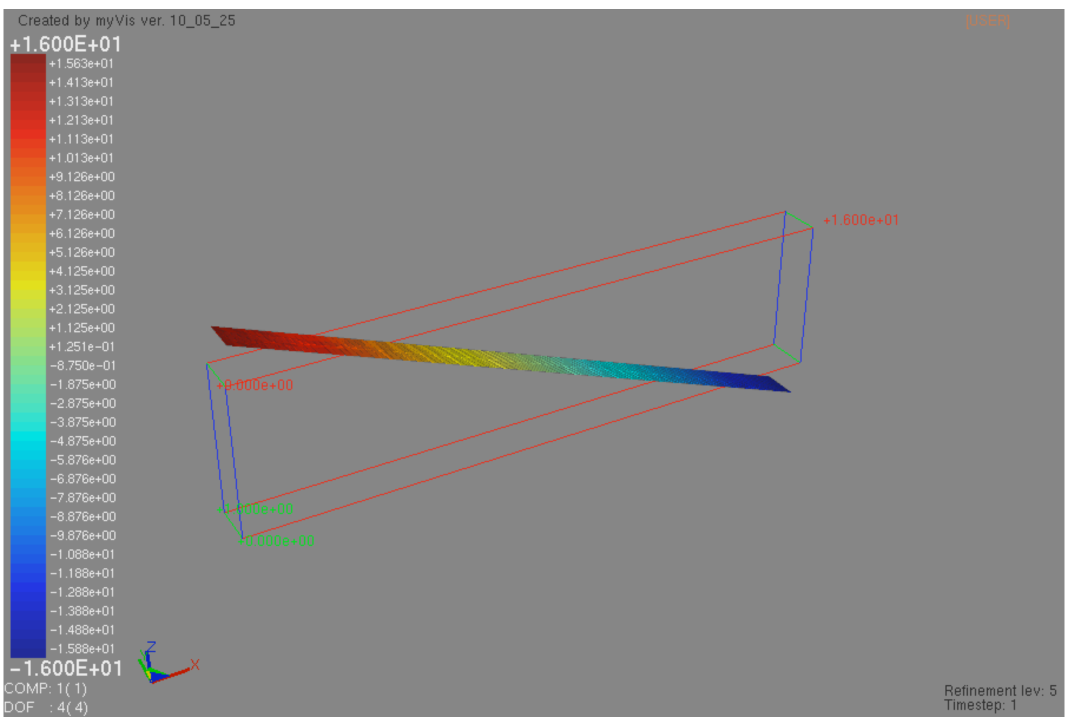


(b) biquadratic

Figure 3.3: Plot of Pressure from Stokes Equations: $D = 16$, Bilinear Elements $h = 1/16$

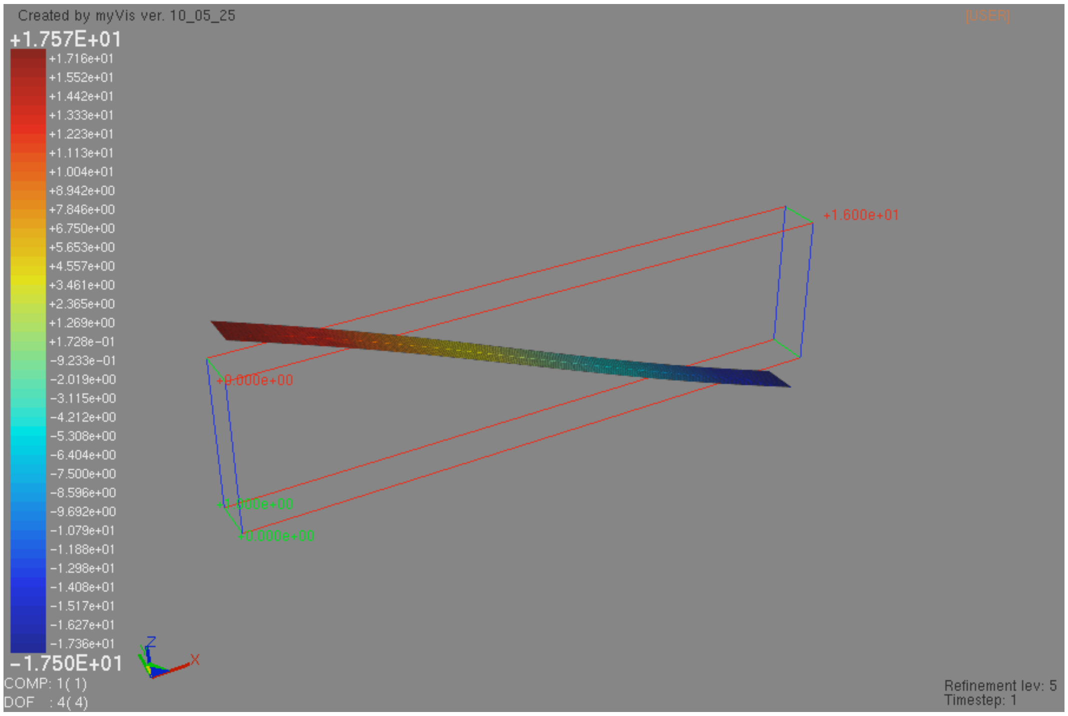


(a) bilinear

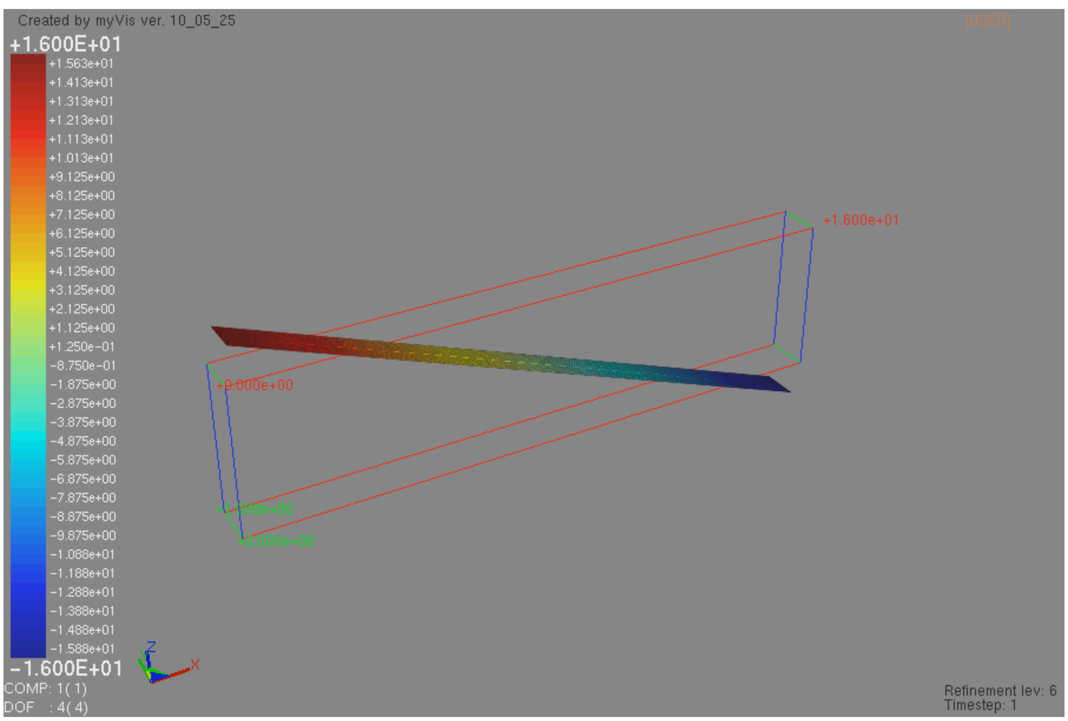


(b) biquadratic

Figure 3.4: Plot of Pressure from Stokes Equations: $D = 16$, Bilinear Elements $h = 1/32$



(a) bilinear



(b) biquadratic

Figure 3.5: Plot of Pressure from Stokes Equations: $D = 16$, Bilinear Elements $h = 1/64$

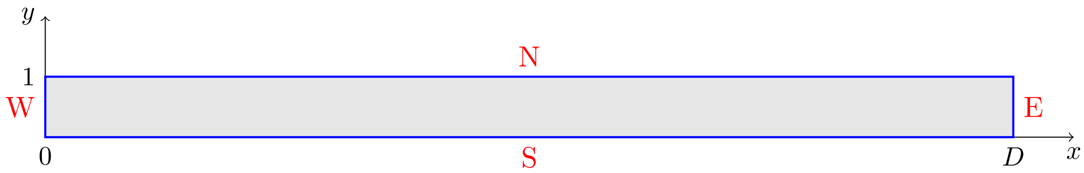
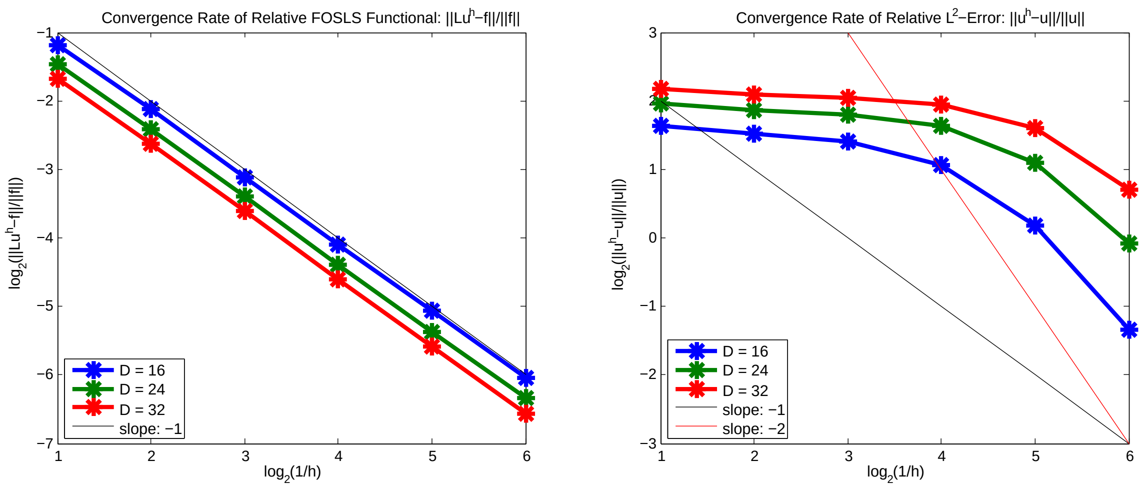


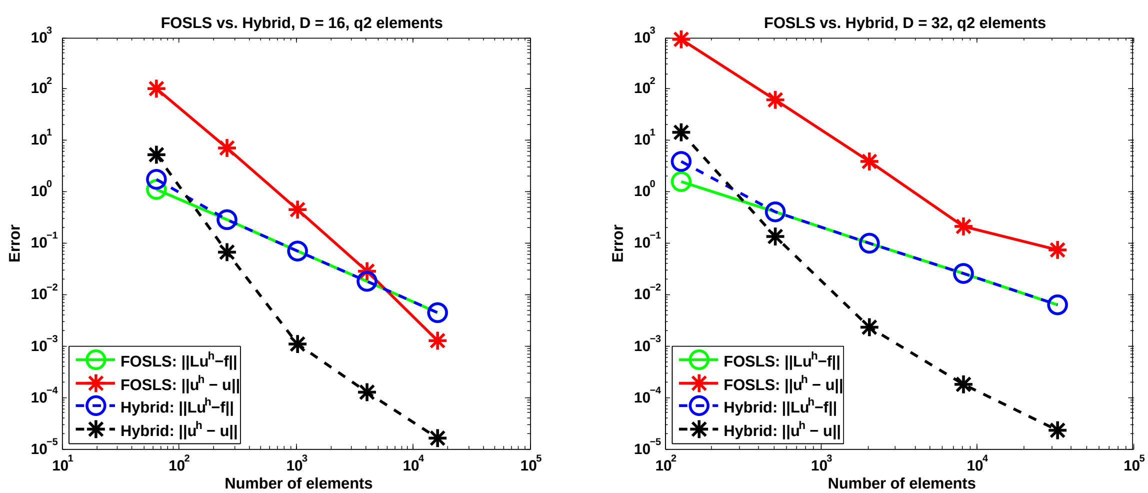
Figure 3.6: The Domain of Long Tube for Stokes Equations



(a) Relative FOSLS Functional

(b) Relative L^2 Error

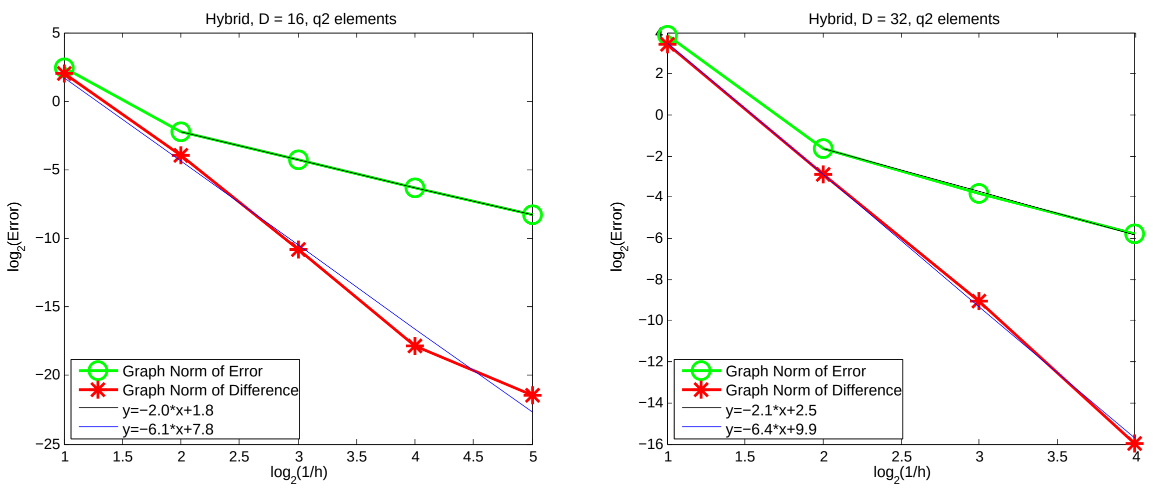
Figure 3.7: Convergence Rate of FOSLS Functional and L^2 Error, FOSLS Formulation, $q1$ Elements



(a) $D = 16$

(b) $D = 32$

Figure 3.8: Convergence Rate of FOSLS Functional and L^2 Error: FOSLS **vs.** Hybrid



(a) $D = 16$ (b) $D = 32$

Figure 3.9: Convergence Rate: $\|\mathbf{u}^h - \hat{\mathbf{u}}\|_G$ **vs.** $\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_G$, Hybrid Formulation, $q2$ Elements

Chapter 4

Hybrid-FOSLS for Stokes Equations in a Backward-step Domain

Hybrid-FOSLS method is first inspired by the success on mass conservation using a two-stage FOSLS method combined with FOSLL* [34]. Heuristically, FOSLS method minimizes $\|L\mathbf{u}^h - \mathbf{f}\|$ and when L is an H^1-equivalent operator, $\|L\mathbf{u}^h - \mathbf{f}\| \approx |\mathbf{e}|_1$; while FOSLL minimizes the adjoint equation $\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\| \approx \|\mathbf{e}\|$. Thus, if we use H^1 conforming finite elements and the solution is not in H^1, the numerical solution can never converge to the exact solution; while for FOSLL*, the exact solution $\hat{\mathbf{u}}$ is approximated by $L^*\mathbf{w}^h$, with $\mathbf{w}^h \in H^1$, thus FOSLL* allows more flexibility to find solutions with reduced smoothness.

In this chapter, we first introduce regularity analysis for PDEs on a domain with re-entrant corner, we then introduce basics of adaptive mesh refinement based on FOSLS, in the last section we present numerical results.

4.1 Issues with Corner Singularities

Singularities in PDEs solutions exist when

- Certain coefficients in the PDEs are discontinuous (*e.g.* oil reservoir with different geologic structures modeled by diffusion equation with jumping discontinuous diffusion coefficient);
- The angle within the domain where Dirichlet and Neumann boundary conditions meet at the boundary is greater than $\frac{\pi}{2}$;
- There is a re-entrant corner (*i.e.* the angle within the domain is greater than π).

A lot of research has been done in the least-squares finite elements setting, such as based on the analytical singular solutions, include discrete version of singular basis functions into the standard finite element spaces [10]; FOSLL* method which minimizes the error's L^2-norm in the range of the adjoint variable $L^*(\mathcal{W})$, thus allow a certain singular in the solutions gradient [34, 40]; weighted-FOSLS method [43] which establish the ellipticity in a weighted Sobolev norm, thus the piecewise polynomial basis functions are dense in the new space.

Our focus in this section is on singularity that is from a re-entrant corner and our approach is mainly based on the research on FOSLL* and weighted-FOSLS method for singular problem. We first introduce the analytical solutions for both Laplace and Stokes equations on a domain with re-entrant corner, then introduce theorems on weighted Sobolev spaces which serve as the theoretical base for our weighting method in the Numerical Test subsection later.

The backward step domain that is within the region $[0, 1]^2$ is shown below. Notice that the inner angle $\gamma = \frac{3\pi}{2}$, thus the point $(0,\ 0)$ is a singular point.

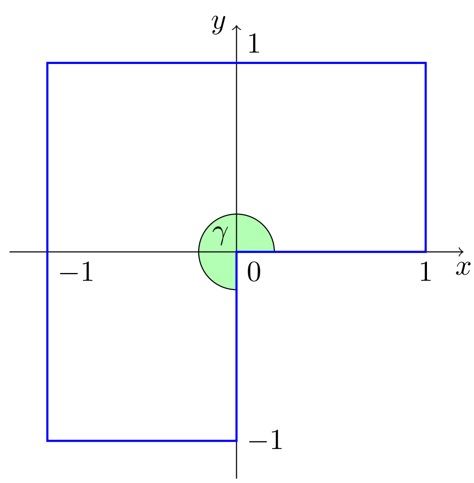


Figure 4.1: Forward Step Domain within the Region $[-1, 1]^2$, with the singular point at $(0, 0)$ and the Re-entrant Corner of degree: $\gamma = \frac{3\pi}{2}$.

4.1.1 Laplace Equation

Consider the scalar Laplace equation with Dirichlet homogeneous boundary conditions on the domain shown by Figure (4.1):

$$\begin{aligned} \Delta u &= 0, \qquad \text{in } \Omega, \\ u &= 0, \qquad \text{on } \partial\Omega. \end{aligned} \tag{4.1}$$

Behavior around corner:

To better grasp the nature of the solution in a domain like Fig 4.1, we cast the problem in a polar coordinates (r, θ) that has the origin at the re-entrant corner. Thus, conventionally, r is distance between any point in the domain to the re-entrant corner, and θ denotes the angle between the vector $\mathbf{x}$ and the vector pointing from the origin to the point. With the denotation, we let

$$u(x, y) = \phi(r, \theta).$$

We solve the homogeneous Laplace equation by the method of separation of variables, that is to assume

$$\phi(r, \theta) = M(r) \cdot N(\theta).$$

Since Laplace operator in the polar coordinates is:

$$\Delta = \left(\partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta}\right), \tag{4.2}$$

Equation 4.1 can be written as

$$\left(\partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta}\right)(M(r)N(\theta)) = 0.$$

Simplify the equation above,

$$\begin{aligned} M_{rr}N + \tfrac{M_r}{r}N + \tfrac{1}{r^2}MN_{\theta\theta} &= 0 \\ r^2 M_{rr}N + rM_rN + MN_{\theta\theta} &= 0. \end{aligned}$$

Assume M, N are not equal to 0 (since we do not care about the trivial solution $\phi(r,\theta) = 0$) and separate the terms involving different variables r, θ, we have:

$$\frac{r^2 M_{rr} + r M_r}{M} = -\frac{N_{\theta\theta}}{N} = \lambda,$$

where λ is a generic constant. Thus, we get two ordinary differential equations (ODEs):

$$r^2 M_{rr} + r M_r - \lambda M = 0, \tag{4.3}$$

$$N_{\theta\theta} + \lambda N = 0. \tag{4.4}$$

The second one is the standard homogeneous linear ODE, use its characteristic equation, it is easy to see the solutions are

$$N(\theta) = C_1 \sin(\sqrt{\lambda}\theta) + C_2 \cos(\sqrt{\lambda}\theta),$$

where C_1 and C_2 can be any constants.

The first one is the Euler's equation and can be solved through change of variable, **i.e.** let

$$t = \ln r, \qquad M(r) = w(t).$$

Since,

$$\begin{aligned} M_r &= w' \cdot \tfrac{1}{t}, \\ M_{rr} &= \tfrac{1}{r^2}(w'' - w'), \end{aligned}$$

apparently, Eqn (4.3) is equivalent to

$$w'' - \lambda w = 0.$$

Solve by using the characteristic equation again, we have

$$w = C_1 e^{\sqrt{\lambda}t} + C_2 e^{-\sqrt{\lambda}t},$$

thus,

$$M(r) = C_1 r^{\sqrt{\lambda}} + C_2 r^{-\sqrt{\lambda}}. \tag{4.5}$$

Combine the results above and notice that, without the loss of generality, we can let $\sqrt{\lambda} = \alpha$, for certain $\alpha > 0$, which leads to:

$$\phi(r,\theta) = \left(C_1 r^{\alpha} + C_2 r^{-\alpha}\right)\left(C_3 \sin(\alpha\theta) + C_4 \cos(\alpha\theta)\right). \tag{4.6}$$

At this step, it is clear that the solution is composed of its singluar part:

$$\phi_s(r,\theta) := C_2 r^{-\alpha} \left(C_3 \sin(\alpha\theta) + C_4 \cos(\alpha\theta) \right),$$

and the non-singluar part:

$$\phi_0(r,\theta) := C_1 r^{\alpha} \left(C_3 \sin(\alpha\theta) + C_4 \cos(\alpha\theta) \right).$$

$\phi_s(r,\theta)$ is a singluar solution, since when r goes to zero (the point get close to the corner), $r^{-\alpha}$ blows up to infinity. From now on we focus ourselves only on the singular solution ϕ_s. Applying the boundary condition, $\phi_s(r,\theta) = 0$ at $\theta = 0$, we have:

$$\phi_s(r,\theta) = C_2 C_4 r^{-\alpha} = 0.$$

Thus, we have $C_4 = 0$. For the simplicity, denote $C = C_2 C_4$ as a non-zero generic constant. When $\theta = \frac{3}{2}\pi$,

$$\phi_s(r,\theta) = C r^{-\alpha} \sin(\frac{3\alpha}{2}\pi) = 0.$$

Notice we have required that $\alpha > 0$, the equation above implies that $\frac{3\alpha}{2} = k$, where $k = 1, 2, 3, \ldots,$ therefore

$$\phi_s(r,\theta) = C r^{-\alpha} \sin(\alpha\theta),$$

where $\alpha = \frac{2k}{3}$, with $k = 1, 2, 3, \ldots$.

Remark 9. *We can get more information about C, α by applying other boundary conditions which do not touch the singluar point. However this is not necessary to study the solutions caused by the re-entrant corner. For example, think about that we have a domain with boundaries $\theta = 0$ and $\theta = \frac{3\pi}{2}$ and extend to whole $\mathcal{R}^2$ plane, it is apparent that boundaries that are far away from the singluar point do not affect the singluar solution.*

4.1.2 Stokes Equation

Consider Stokes Equations with homogeneous boundary condition on the domain as in 4.1:

$$-\Delta \mathbf{u} + \nabla p = 0, \quad \text{in } \Omega, \tag{4.7}$$

$$\nabla \cdot \mathbf{u} = 0, \quad \text{in } \Omega, \tag{4.8}$$

$$\mathbf{u} = \mathbf{0}, \quad \text{on } \partial\Omega. \tag{4.9}$$

Since $\nabla \cdot \mathbf{u} = 0$ and trivially, $\int_{\partial\Omega} \mathbf{u} \cdot \mathbf{n} = 0$, by Helmholtz Decomposition (Theorem 16, Appendix), there exists $\phi \in \mathcal{H}^1$ such that $\mathbf{u} = \nabla^{\perp}\phi$. The boundary condition:

$$\nabla^{\perp}\phi = \begin{pmatrix} \partial_y \phi \\ -\partial_x \phi \end{pmatrix} = \mathbf{0}, \quad \text{on } \partial\Omega$$

implies that, on the boundary $\partial\Omega$,

$$\mathbf{t} \cdot \nabla\phi = 0, \tag{4.10}$$

$$\mathbf{n} \cdot \nabla\phi = 0. \tag{4.11}$$

Eqn (4.10) implies that ϕ equals to some constant along the boundary $\partial\Omega$. Without loss of generality, we let

$$\phi = 0, \quad \text{on } \partial\Omega.$$

With the help of ϕ, (4.7) can be recast to

$$-\Delta(\nabla^{\perp}\phi) = 0,$$

which is equivalent to

$$\nabla^{\perp}(-\Delta\phi) = 0. \tag{4.12}$$

Apply $\nabla\times$ to both sides of (4.12), we have the following biharmonic boundary problem of ϕ:

$$\begin{aligned} -\Delta^2\phi &= 0, \quad \text{in } \Omega, \\ \phi &= 0, \quad \text{on } \partial\Omega, \\ \mathbf{n} \cdot \nabla\phi &= 0, \quad \text{on } \partial\Omega. \end{aligned}$$

Recall that, the Laplacian Δ in polar coordinates (r, θ) is with the following form:

$$\Delta = \partial_{rr} + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta}.$$

Assume $\phi(r,\theta) = r^\alpha N(\theta)$, use N' denote N_θ, N'' denote $N_{\theta\theta}$ and so on, we have:

$$\begin{aligned}
\Delta\phi &= \alpha(\alpha-1)r^{\alpha-2}N(\theta) + \alpha r^{\alpha-2}N(\theta) + r^{\alpha-2}N''(\theta) \\
&= r^{\alpha-2}(\alpha^2 N + N''),
\end{aligned}$$

thus,

$$\begin{aligned}
\Delta^2\phi &= (\alpha-2)^2 r^{\alpha-4}(\alpha^2 N + N'') + r^{\alpha-4}(\alpha^2 N + N'')'' \\
&= r^{\alpha-4}\left[(\alpha-2)^2\alpha^2 N + \left((\alpha-2)^2+\alpha^2\right)N'' + N''''\right]
\end{aligned}$$

$-\Delta^2\phi = 0$ implies,

$$(\alpha-2)^2\alpha^2 N + \left((\alpha-2)^2+\alpha^2\right)N'' + N'''' = 0 \tag{4.13}$$

Again, as for Laplace equation in Subsection 4.1.1, Eqn 4.13 is a standard homogeneous linear ODE. Solve by using its characteristic equation, we have the generic solution of (4.13):

$$N(\theta) = C_1\cos(\alpha\theta) + C_2\sin(\alpha\theta) + C_3\cos((\alpha-2)\theta) + C_4\sin((\alpha-2)\theta), \tag{4.14}$$

where C_1, C_2, C_3, C_4 are constants to be determined by the boundary conditions.

Apply the boundary conditions:

$$\begin{aligned}
N(0) = 0, &\qquad N(\tfrac{3\pi}{2}) = 0, \\
N'(0) = 0, &\qquad N'(\tfrac{3\pi}{2}) = 0,
\end{aligned}$$

and combine (4.14), we obtain

$$\begin{pmatrix}
\cos(0) & \sin(0) & \cos((\alpha-2)0) & \sin((\alpha-2)0) \\
-\alpha\sin(0) & \alpha\cos(0) & -(\alpha-2)\sin(0) & (\alpha-2)\cos(0) \\
\cos(\alpha\gamma) & \sin(\alpha\gamma) & \cos((\alpha-2)\gamma) & \sin((\alpha-2)\gamma) \\
-\alpha\sin(\alpha\gamma) & \alpha\cos(\alpha\gamma) & -(\alpha-2)\sin((\alpha-2)\gamma) & (\alpha-2)\cos((\alpha-2)\gamma)
\end{pmatrix}
\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{4.15}$$

Since (4.15) has a non-zero solution if and only if the above matrix is singular, we let the determinant of the left-hand-side of (4.15) equal to zero and get:

$$2\alpha(\alpha-2) - \left((\alpha-2)^2+\alpha^2\right)\sin(\alpha\gamma)\sin\left((\alpha-2)\gamma\right) - 2\alpha(\alpha-2)\cos(\alpha\gamma)\cos\left((\alpha-2)\gamma\right) = 0,$$

where $\gamma = \frac{3\pi}{2}$ as in Figure (4.1). Furthur simplification leads to

$$\alpha(\alpha-2) + \sin^2(\alpha\gamma) = 0.$$

Numerical solutions of the above equation can be easily found by finding the intersection points in Figure 4.1.2. The values of α are given in Table (4.1.2).

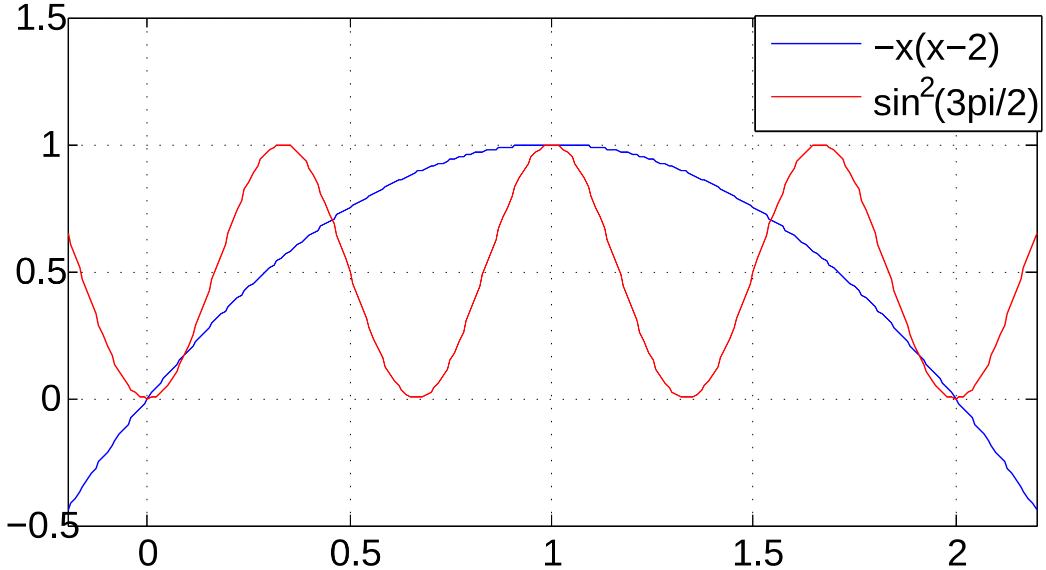


Figure 4.2: Finding Roots of α for Singular Solution of Stokes Equation: $\phi(r,\theta) = r^{\alpha}t(\theta)$ in a Domain with a Re-entrant Corner

Notice that, $\phi(r,\theta)$ is of the form, $\phi(r,\theta) = r^{\alpha}\cdot t(\theta)$, where $t(\theta)$ is a trigonometric function of θ. Thus by (C.4), $\phi \in \mathcal{H}^{\alpha-\epsilon}(\Omega)$, for all $\epsilon > 0$ that is small enough. This implies that, in general, $\mathbf{u} = \nabla^{\perp}\phi \in \mathcal{H}^{\alpha-\epsilon}(\Omega)$.

α_1	α_2	α_3	α_4	α_5	α_6
0	0.0914708	0.4555516	1.54448	1.90853	2

Table 4.1: Values of α for Singular Solution of Stokes Equation: $\phi(r, \theta) = r^{\alpha} t(\theta)$ in a Domain with a Re-entrant Corner

4.1.3 Weighted Sobolev Space and Convergence Theory

Define the weighted Sobolev space $H^{k,\beta}(\Omega)$ as

$$H^{k,\beta}(\Omega) = \{q \in L^2(\Omega) : \text{for } |j| \leq k, r^{\beta-k+j} D^j q \in L^2(\Omega)\}, \tag{4.16}$$

where $j = (j_1, j_2, j_3)$ should be understand as the index and $|j| = j_1 + j_2 + j_3$ and r is the distance to the singluar point.

$H^{k,\beta}(\Omega)$ is a Hilbert space with the norm

$$\|q\|_{k,\beta} := \left(\sum_{|j| \leq k} \|r^{\beta-k+j} D^j q\|_0^2 \right)^{1/2}. \tag{4.17}$$

For example we have the following:

$$\begin{aligned} \|q\|_{0,\beta} &= \|r^{\beta} q\|_0, \\ \|q\|_{1,\beta} &= \left(\|r^{\beta} D^1 q\|_0^2 + \|r^{\beta-1} q\|_0 \right)^{1/2}, \\ \|q\|_{2,\beta} &= \left(\|r^{\beta} D^2 q\|_0^2 + \|r^{\beta-1} D^1 q\|_0^2 + \|r^{\beta-2} q\|_0 \right)^{1/2}. \end{aligned} \tag{4.18}$$

The weighted norms above can be easily memorized by noticing that they are very similar to the non-weighted Sobolev norms with the exception that the highest differential order term is weighted by r^{β} and the the degree of r reduces as the differential order of the function reduces.

The convergence theory of weighted-FOSLS method for elliptic problems are established in [39]. The major theorem in [39] is the following:

Theorem 11. *If $\mathbf{u}^h \in \mathcal{V}^h$ is the minimizer of*

$$\mathcal{G}_w(\mathbf{u}^h; \mathbf{f}) := \|L\mathbf{u}^h - \mathbf{f}\|_{0,\beta}^2 = \|r^{\beta}(L\mathbf{u}^h - \mathbf{f})\|_0^2,$$

where $|1-\beta|<\alpha$*, then we have the following error estimates*

$$\begin{aligned} \|\hat{\mathbf{u}}-\mathbf{u}^h\|_{1,\beta} &\le Ch^{\alpha+\beta-1}\|\hat{\mathbf{u}}\|_{\alpha+\beta,\beta}, && (4.19)\\ \mathcal{G}_w(\hat{\mathbf{u}}-\mathbf{u}^h;0)^{1/2} &\le Ch^{\alpha+\beta-1}\|\hat{\mathbf{u}}\|_{\alpha+\beta,\beta}, && (4.20)\\ \|\hat{\mathbf{u}}-\mathbf{u}^h\|_0 &\le Ch^{\alpha}\left(|\hat{\mathbf{u}}|_\alpha+\|\hat{\mathbf{u}}\|_{\alpha+\beta,\beta}+\|\hat{\mathbf{u}}\|_{\alpha-\beta+2,2-\beta}\right), && (4.21) \end{aligned}$$

where $\alpha+\beta\le k+1$ *and* k *is the degree of piecewise polynomial in finite element space* $\mathcal{V}^h$*.*

The theorem above is important for it guides us to choose the right weighting order for r in order to achieve optimal finite element convergence rate based on the degree of piecewise polynomials used as basis functions.

The notations and theorems introduced here are enough for the purpose of this thesis. For a comprehensive introduction on weighted Sobolev spaces, we refer to [37].

4.2 Adaptive Mesh Refinement for Hybrid-FOSLS

The adaptive mesh refinement strategies we use for our numerical tests is the so-called "accuracy-per- computational-cost-efficiency" (ACE) method, which is based on FOSLS method in conjunction with algebraic multigrid (AMG) and in the nested iteration (NI) context. Equipped with the a posteriori error estimates FOSLS offers without extra cost, ACE estimates the AMG computational cost to solve the linear system; an "error reduction per computational cost" can thus be estimated to guide the refinement.

This concept can be summarized as:

$$\text{Solve} \;\Rightarrow\; \text{Estimate} \;\Rightarrow\; \text{Mark} \;\Rightarrow \text{Refine}$$

Besides to refine based on both error reduction and computational cost, another core idea of ACE is to achieve an "optimal" hierarchy of grids, that is to get a certain error on the whole domain with minimal possible number of elements. In 1D it has been proved that this can be achieved by equally distribute error on all elements. Numerical results in [4] suggests a similar conclusion.

The most remarkable advantage of this FOSLS-based adaptive mesh refinement strategy lie in that the local FOSLS functional $\|L\mathbf{u}-\mathbf{f}\|_\tau$ computed on each piece of element serves as the error estimate and there is no need for any additional cost to compute the estimate, which is not the case for other adaptive mesh refinement strategies.

Earlier research on adaptive mesh refinement is mainly based on the reduction of error itself with little attention on the computational cost involved. There are several algorithms,

- marking an element for refinement if the estimated error on it surpasses a certain factor of the largest error on one element at the current refine level.

- a "threshold-hold" based method, which searches a minimal set of elements, such that the sum of error-squared is bigger than a certain factor of error-squared on all elements in current level; then all the elements in this set are marked for refinement.

We introduce the basics of ACE applied with FOSLS method in this section and its Hybrid-FOSLS analog is presented in 4.3.2.

For a PDE system that is already reformulated into a first-order system $L\mathbf{u} = \mathbf{f}$, the FOSLS functional is

$$\mathcal{F}(\mathbf{u};\mathbf{f}) := \|L\mathbf{u}-\mathbf{f}\|_{0,\Omega}^2 = \sum_i \|L\mathbf{u}-\mathbf{f}\|_{0,\tau_i}^2,$$

where τ_i denotes the i'th element.

Denote

$$\mathcal{F}_i(\mathbf{u};\mathbf{f}) := \|L\mathbf{u}-\mathbf{f}\|_{0,\tau_i}^2$$

and further assume the linear operator L is continuous and coercive with respect to a certain norm $\mathcal{V}$, that is: there exist the coercivity constant c_0 and continuity constant c_1 which are independent of the mesh size h and the function $\mathbf{u}$, such that

$$c_0\|\mathbf{u}\|_{\mathcal{V},\Omega} \leq \|L\mathbf{u}-\mathbf{f}\|_{0,\Omega} \leq c_1\|\mathbf{u}\|_{\mathcal{V},\Omega}.$$

It is shown that as the error indicator, the FOSLS functional is **locally sharp** and **globally**

reliable. By the continuity of L,

$$\mathcal{F}_i(\mathbf{u}^h - \hat{\mathbf{u}}; 0)^{1/2} \le c_1 \|\mathbf{u}^h - \hat{\mathbf{u}}\|_{\mathcal{V},\tau_i}.$$

Notice that the inequality holds locally on each element and in most cases $\|\cdot\|_{\mathcal{V}}$ is $\|\cdot\|_1$. The locally sharpness implies that when the FOSLS functional is large the error's H^1 norm on the local element is large, and needs to be refined.

On the other hand, the global reliability

$$c_0 \|\mathbf{u}\|_{\mathcal{V},\Omega} \le \|L\mathbf{u} - \mathbf{f}\|_{0,\Omega},$$

implies if the FOSLS functional is small enough, the H^1 error has to be small, too. Thus, if the error indicator suggests the global error is under control, the refinement process can stop.

Noticeably, the ACE adaptive mesh refinement strategy has also been implemented for parallel machine, with nice load balancing and excellent scalability. Interested readers can refer to [36,50].

4.3 Numerical Results

4.3.1 Mass Conservation

Consider the steady state Stokes equations on a 2D backward-step domain, Ω, shown in figure (4.3). The domain is of 10×1 with a 2×0.5 rectangle removed from the corner of bottom left. Also, denote W, E, S, N, V, H as boundaries indicated in (4.3).

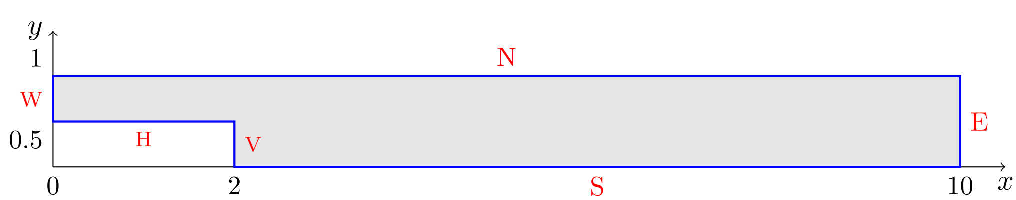


Figure 4.3: Backward Step Domain within the Region $[0, 10] \times [0, 1]$, with the Singular point at $(2, 0.5)$ and the Re-entrant Corner of Degree $\frac{3\pi}{2}$;

The simplified FOSLS (primal) system and FOSLL* (adjoint) systems are the same as (3.64) and (3.66) respectively, with slightly different boundary conditions specified as follows.

For primal variables, we give parabolic velocity for the inflow, such that its integration along the inflow boundary equals to 0.8 for the convenience. That leads to the following boundary conditions for the primal variables:

$$\begin{aligned} U_{11} &= 0, && [W,N,S,E,H,V], \\ U_{12} &= 57.6-76.8y, && [W], \\ U_{12} &= 0, && [V], \\ U_{21} &= 0, && [N,S,H], \\ p &= 0, && [E], \end{aligned} \tag{4.22}$$

and

$$\begin{aligned} w_1 &= 0, && [W,N,S,H,V], \\ w_2 &= 0, && [W,N,S,E,H,V], \\ w_3 &= 0, && [E], \end{aligned} \tag{4.23}$$

for adjoint variables.

Since the domain has a re-entrant corner, the solution has a singularity at the point $(2, 0.5)$ and the analytic solution is not available. Although we cannot compare FOSLS and hybrid-FOSLS by looking directly at the error, we can still compare them on mass conservation by measuring the the fractional change of mass flow:

$$\frac{\int_{\Gamma_i}(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_i - \int_{\Gamma_0}(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_0}{\int_{\Gamma_0}(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_0}, \tag{4.24}$$

where Γ_0 is the inflow boundary and Γ_i's are the lines, $x = x_i$, are the outflow boundaries, that are according to the red lines in Figure (4.4).

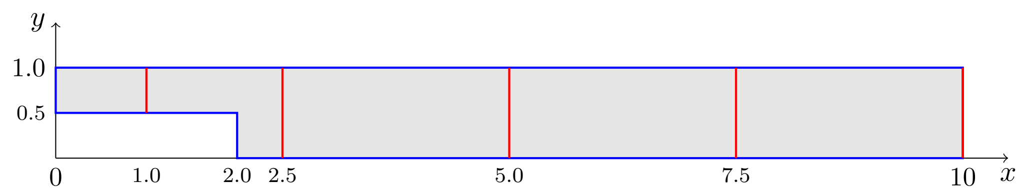


Figure 4.4: Backward Step Domain with Outflow Sections at $x = 1.0$, 2.5, 5.0, 7.5, 10.0

For problems with singularities, such as the backward-step domain problem we have here, the exact solutions may not be in H^1. Thus, in either Galerkin or least-squares formulation, using H^1-conforming finite element spaces may not lead to convergence. Numerical solutions don't approximate the exact solutions well, not only in regions near the singular points, but also regions far away from them, which is known as "pollution effect". Remedies for the loss of H^1-regularity in FOSLS can be found in [10, 23, 40, 43, 44]

By similar approach in [43], we introduce two weighting functions, w_1, w_2, with the form

$$w_i(r) = \begin{cases} r^{\alpha_i}, & \text{if } r < \epsilon/2 \\ q(r), & \text{if } \epsilon/2 \leq r < \epsilon \\ 1, & \text{if } \epsilon \leq r \end{cases} \tag{4.25}$$

where $i = 1, 2$ and α_i is to be chosen according to the regularity of exact solutions and finite element spaces.

Instead of standard FOSLS and standard hybrid-FOSLS approach, we minimize weighted FOSLS and hybrid-FOSLS functionals respectively:

$$\|w_1(L\mathbf{u}^h - \mathbf{f})\|^2, \tag{4.26}$$

for weighted FOSLS over finite element space $\mathcal{V}^h$ and

$$\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\|^2 + \|w_2(L^*\mathbf{w}^h - \mathbf{u}^h)\|^2 + \|w_1(L\mathbf{u}^h - \mathbf{f})\|^2, \tag{4.27}$$

for weighted hybrid-FOSLS over $\mathcal{W}^h \times \mathcal{V}^h$.

The reason that we are not weighting the FOSLL* term in Hybrid functional is due to the technical difficulties to weight FOSLL*. If we weight FOSLL* equation to obtain the following:

$$wL^*\mathbf{w} = w\hat{\mathbf{u}},$$

and minimize the weighted FOSLL* functional, $\|w(L^*\mathbf{w} - \hat{\mathbf{u}})\|^2$. The associated weak problem will be: find $\mathbf{w} \in \mathcal{W}$, such that

$$< wL^*\mathbf{w}, wL^*\mathbf{z} >=< w\hat{\mathbf{u}}, wL^*\mathbf{z} >, \quad \text{for all } \mathbf{z} \in \mathcal{W}. \tag{4.28}$$

We cannot use the $< \hat{\mathbf{u}}, L^*\mathbf{z} >=< \mathbf{f}, \mathbf{z} >$ trick as we did for FOSLL*, since the resulting weak form involves $L(w^2\hat{\mathbf{u}})$ which is unkown. Thus, using the same approach to weight FOSLL* as to weight FOSLS is not possible.

Analytical results in [28] show that for Stokes equations on a backward-step domain, the singular solution is of the form $Cr^s\psi(\theta)$ in polar coordinates, where C is a constant, $\psi(\theta)$ is a trigonometric function and $s \approx 1.5445$. The analysis in Appendix C.3 suggests that approximately, $\hat{\mathbf{u}} \in H^{0.5445}$. Thus, by Theorem 11, to get the optimal convergence rate, we need

$$\alpha + \beta - 1 = p,$$

where p is the degree of the finite element basis functions and $\beta \approx 0.5445$ in this case. Therefore, we choose $\alpha_1 = 3/2$ when using bilinear elements, $\alpha_1 = 5/2$ when using bi-quadratic elements and so on.

Although, unfortunately, we cannot weight the FOSLL* term, we can still use the heuristic based on the following theorem to weight the intermediate term.

Theorem 12. *Assume L, L^* are H^1-equivalent first-order operators, $\mathbf{w}^h$, $\mathbf{u}^h$ are solutions in H^1-conforming finite element spaces. Denote $\hat{\mathbf{u}}$ as the exact solution and let C be a generic positive constant that is independent of domain and mesh size. Then,*

$$\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\| \leq Ch\|L\mathbf{u}^h - \mathbf{f}\| \leq Ch^{p+1}\|\hat{\mathbf{u}}\|_{p+1}.$$

If we assume the intermediate term $\|L^*\mathbf{w}^h - \mathbf{u}^h\|$ has the similar results as the FOSLL* term $\|L^*\mathbf{w}^h - \hat{\mathbf{u}}\|$, then to make the intermediate term have the optimal convergence rate, we choose $\alpha_2 = 1/2$ when using bilinear elements, $\alpha_2 = 3/2$ when using biquadratic elements.

At the moment, our weighted hybrid-FOSLS method remains immature and how to remedy this will be left for the future research. However the numerical results below suggests our current method works well.

Nevertheless, numerical results with our weighted-hybrid-FOSLS are very encouraging as indicated in Table 4.3 and Table 4.2. These two tables compare weighted-FOSLS and weighted-

lev	ne	x= 1.0	x = 2.5	x = 5.0	x = 7.5	x = 10.0
Weighted-FOSLS with Uniform Refinement						
1	576	-37.4104%	-98.4042%	-100.8601%	-97.3274%	-95.1921%
2	2304	-32.5437%	-89.6799%	-99.8621%	-99.5952%	-97.3708%
3	9216	-22.5485%	-68.0978%	-86.2746%	-95.4464%	-100.3462%
4	36864	-10.5741%	-31.2673%	-35.0491%	-34.7209%	-33.9610%
5	147456	-3.4330%	-10.8240%	-12.5968%	-12.1150%	-10.5229%
6	589824	-1.0555%	-3.7119%	-6.4604%	-7.7685%	-7.5111%
7	2359296	-0.1724%	-0.7272%	-1.5229%	-1.9391%	-1.9741%

Table 4.2: Mass Loss on Different Sections of the Tube, Weighted-FOSLS, Steady State Stokes Equations on Backward Step Domain with q2-elements.

lev	ne	x= 1.0	x = 2.5	x = 5.0	x = 7.5	x = 10.0
Weighted-Hybrid with Uniform Refinement						
1	144	-19.7440%	-202.7332%	-0.4175%	-0.5353%	-0.4497%
2	576	-0.9627%	-98.8571%	0.3498%	-0.0207%	-0.0008%
3	2304	0.4686%	-27.9483%	-0.0111%	-0.0042%	-0.0002%
4	9216	0.0653%	-5.0429%	-0.0191%	-0.0006%	-0.0000%
5	36864	-0.0018%	-0.7265%	-0.0033%	-0.0009%	-0.0009%
6	147456	-0.0040%	-0.1827%	-0.0024%	-0.0001%	0.0000%

Table 4.3: Mass Loss on Different Sections of the Tube, Hybrid, Steady State Stokes Equations on Backward Step Domain with q2-elements.

Hybrid formulation on relative mass loss described in (4.24). In the first column are refinement levels and in the second column are number of elements.

For hybrid-FOSLS, we start with mesh size $h = 1/4$, number of elements $ne = 144$ at Level 1. Since we use uniform refinement for this set of tests, $h = 1/8$ at Level 2 and so on. To better compare FOSLS with Hybrid formulation, we carry out our FOSLS tests on IBM Blue Gene/L parallel computer, tested with 256 processors and start with $h = 1/8$, $ne = 576$ at the first level.

According to the tables, clearly, hybrid-FOSLS performs much better than FOSLS. Even though for Hybrid formulation, the linear system is doubled, the numerical benefit is much more than offset the extra computation work. For example, with the same number of elements $ne = 147456$, at $x = 2.5$, the mass loss is 0.1827% for Hybrid while 10.8240% for FOSLS, which is a $\mathcal{O}(10^2)$ improvement.

Since the singularity point is at (2, 0.5), we expect mass conservation will be the worst at section $x = 2.5$. While this is true for hybrid-FOSLS, it is not always the case for FOSLS. More noteworthy, while for FOSLS mass conservation is about the same at different sections of the domain; for Hybrid, mass conservation is much better away from the singularity. **i.e.** the pollution effect is much stronger for FOSLS than for Hybrid.

The comparison of FOSLS and Hybrid method is illustrated more clear in the following set of plots. The vertical green line is the west (inflow) boundary and the vertical red lines are the boundaries at which we measure the loss in flux, as in Eqn (4.24). The y-axis is the percentage of mass loss while the x-axis is the degrees of freedom (DOF), **i.e.** the number of unkonwns we use in our computation. Notice that, since in the Hybrid system, adjoint variables are introduced and the size of the linear system is doubled, using DOF instead of number of elements is a fair comparison measure.

We start our computation from the same mesh size. Due to the fact that, on the same grid, the Hybrid method has two times as many unknowns as the FOSLS method requires, plots below show that those points do not vertically line up together. Also because of the memory constraint of the computer, FOSLS can go to a finer mesh grids than Hybrid.

There are two interesting observations that can be drawn from the set of plots.

- First, except at the line section that is right after the singular point, Hybrid-FOSLS conserves mass very well (almost zero mass loss).

- Second, FOSLS suffers much more on mass loss near the re-entrant corner ($x = 2.5$) and the pollution effect still heavily affects its mass conservation even when the flow is far away from the singular point.

Hybrid displays more robust quality on the issue of mass conservation. Although this should not come as a surprise, given our analysis in Chapter 3, where we have shown minimizing Hybrid functional is roughly minimizing the graph norm of the error, which gives it excellent control on L^2 error and leads to good control of mass conservation; the fact that using our heuristic weighting

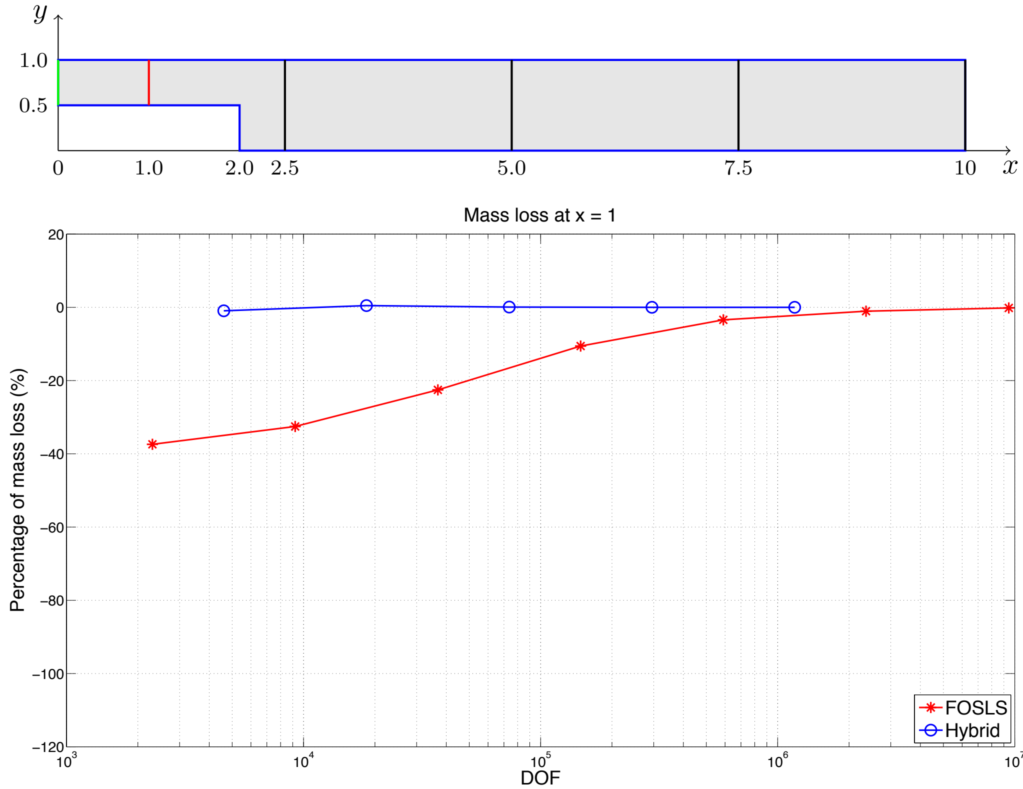


Figure 4.5: Mass conservation versus number of elements at $x = 1.0$, steady state Stokes equations in a backward-step domain.

scheme for Hybrid and leaving the FOSLL* term unweighted has led to success on irregular domain is still inspiring.

4.3.2 Adaptive Mesh Refinement

For any given element, $\tau \in \mathcal{T}$, define the local error estimate as

$$\epsilon_\tau^2 := ||L^*\mathbf{w}^h - \mathbf{u}^h||_\tau^2 + ||L\mathbf{u}^h - \mathbf{f}||_\tau^2. \tag{4.29}$$

The local sharpness of the error estimate follows directly from the continuity, i.e.,

$$\epsilon_\tau^2 \leq \mathcal{H}_\tau((\mathbf{w}^h, \mathbf{u}^h); (\hat{\mathbf{u}}, \mathbf{f})) \leq 3||(\mathbf{w}^h - \hat{\mathbf{w}}, \mathbf{u}^h - \hat{\mathbf{u}})||_\mathcal{H}. \tag{4.30}$$

However, we do not have the global reliability bound for the error estimate (4.29). In Section 3.2.2, under mild assumptions, we established the results that the numerical solution obtained

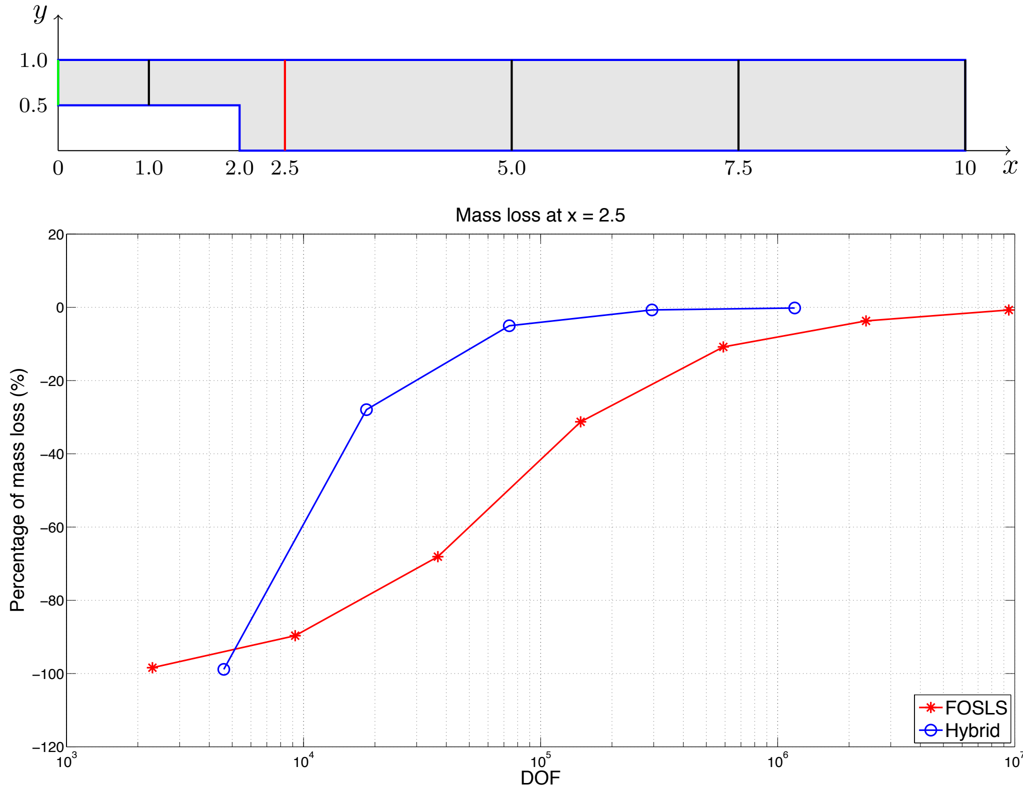


Figure 4.6: Mass conservation versus number of elements at $x = 2.5$, steady state Stokes equations in a backward-step domain.

through minimizing the Hybrid functional is very close to the numerical solution which minimizes the error in the graph functional (3.25) (see Lemma 5 and Theorems 9 and 10). Define the local error measured by the graph norm as

$$\eta_\tau^2 := \sqrt{\frac{1}{2}||\mathbf{u}^h - \hat{\mathbf{u}}||_\tau^2 + ||L\mathbf{u}^h - \mathbf{f}||_\tau^2}. \tag{4.31}$$

Write $\epsilon = \sqrt{\sum_{\tau \in \mathcal{T}} \epsilon_\tau^2}$, the global error approximated by the error estimate (4.29), and $\eta = \sqrt{\sum_{\tau \in \mathcal{T}} \eta_\tau^2}$, the global error measured in graph norm. We want to demonstrate that the local error estimate, ϵ_τ, provides a sharp approximation to the true error, η_τ, on relatively coarse grids through numerical results. To illustrate that, we list the maximum and minimum ratios between the local error estimate and true error, $\frac{\epsilon_\tau}{\eta_\tau}$, on various mesh sizes for the test problem described in section 3.3, i.e., the Stokes equations in a long tube for which exact solutions are trigonometric

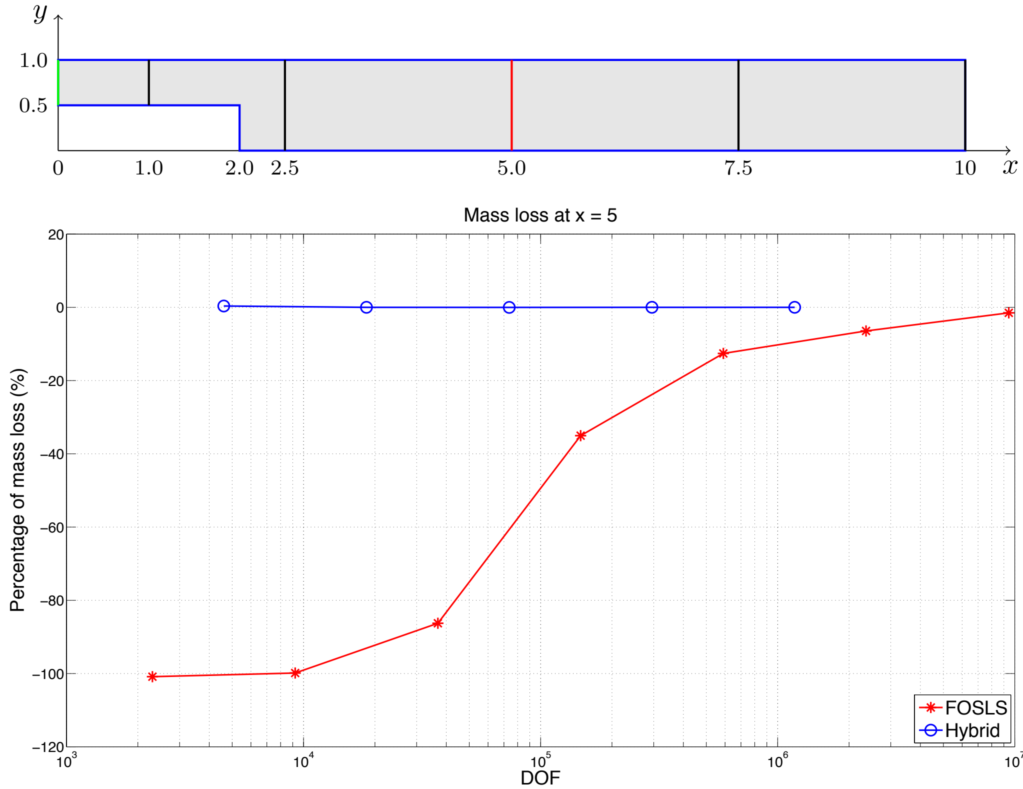


Figure 4.7: Mass conservation versus number of elements at $x = 5.0$, steady state Stokes equations in a backward-step domain.

functions; see table 4.4. The local error estimate approximates the true error very well on mesh

mesh size	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$
$\max_\mathcal{T} \frac{\epsilon_\mathcal{T}}{\eta_\mathcal{T}}$	2.421e+02	1.115e+02	7.298e+00	1.005e+00	1.036e+00
$\min_\mathcal{T} \frac{\epsilon_\mathcal{T}}{\eta_\mathcal{T}}$	1.178e+01	1.021e+00	1.001e+00	1.000e+00	1.000e+00

Table 4.4: Local error estimate versus true error for the Hybrid method applied to Stokes equations in a long tube, $[0.0, 16.0] \times [0.0, 1.0]$. Exact solutions are given in Table 3.3.

size $h = \frac{1}{16}$, with lower and upper bound 1.001 and 7.298, respectively. On the finer grid with mesh size $h = \frac{1}{32}$, the local error estimate provides almost perfect approximation to the true error. Apparently, this leads to the conclusion that the global error estimate, ϵ, approaches the true global error, η, which is stronger than the global reliability. However, to prove theoretically the local and global sharpness of the error estimate, $\epsilon_\mathcal{T}$, is rather difficult. That remains future research.

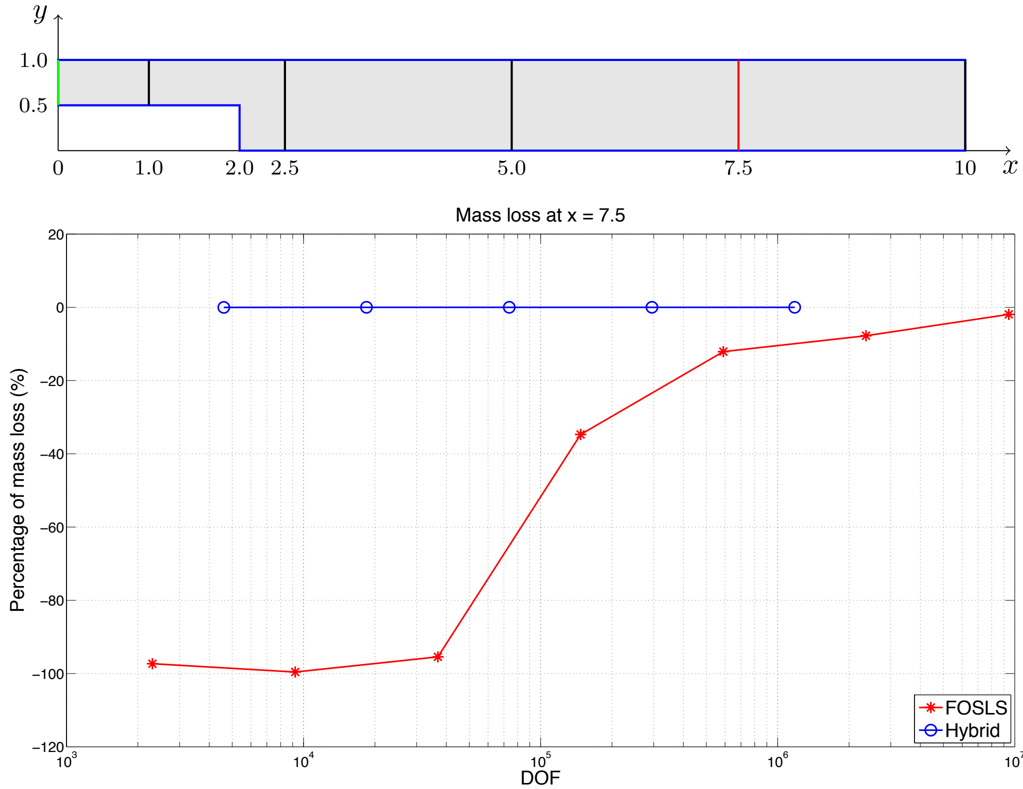


Figure 4.8: Mass conservation versus number of elements at $x = 7.5$, steady state Stokes equations in a backward-step domain.

The test problem is the steady state Stokes equations on the backward-step domain described in section 4.3.1. The refinement process starts with a uniform coarse grid containing 8 biquadratic elements per unit length, i.e., mesh size $h = \frac{1}{8}$. On each refinement level, elements are split equally into two in each dimension. Refinement decisions are made by minimizing the "accuracy-per-computational-cost" efficiency (ACE) measure as described in [4]. Refinement patterns in Figure 4.10 show that most refinements are made near the re-entrant corner, where the singularity is located. The refinements also follow the flow downstream. The mass loss at the vertical line, $x = 2.5$, which is right after the flow hits the singularity, is given in Figure 4.11. Adaptive mesh refinement based on the local error estimate (4.29) is able to reach almost zero mass loss with nearly 5,000 elements. Compared to that, uniform refinement takes close to 30,000 elements to have the same mass loss.

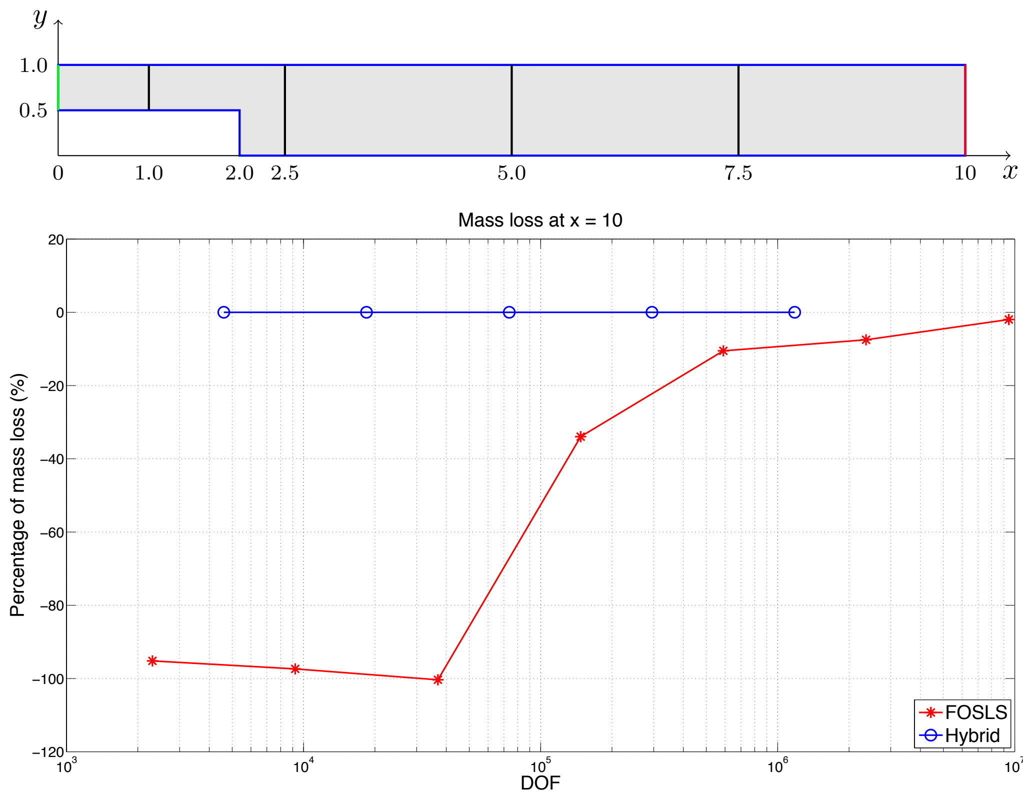
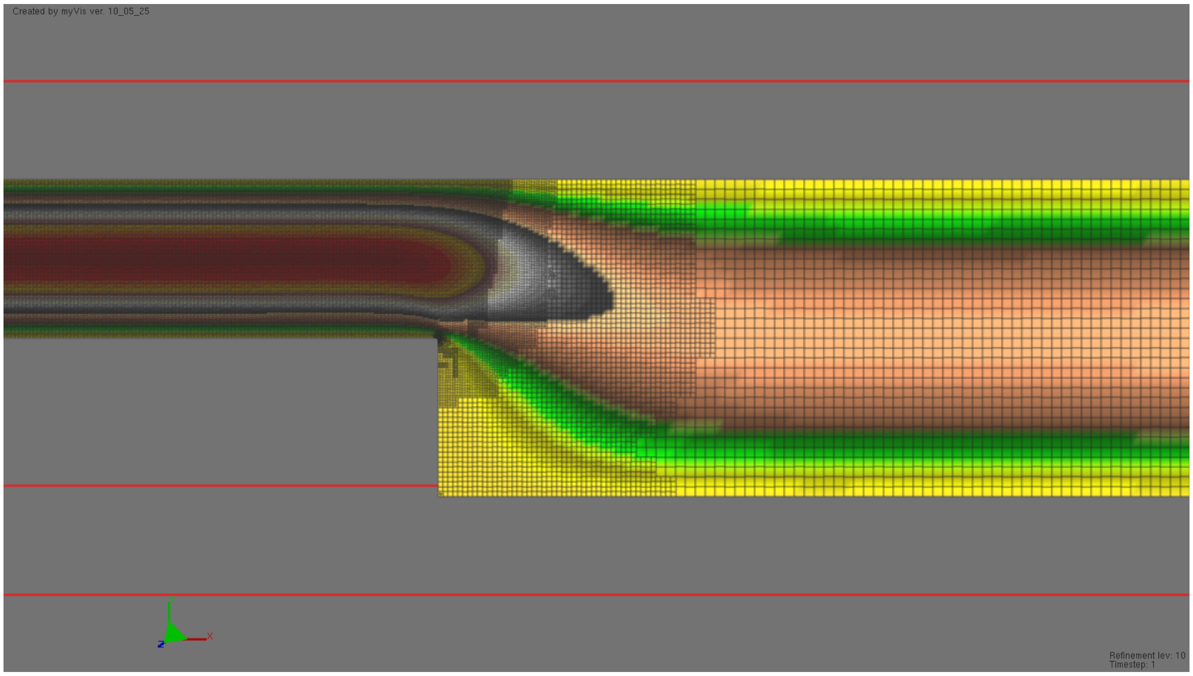
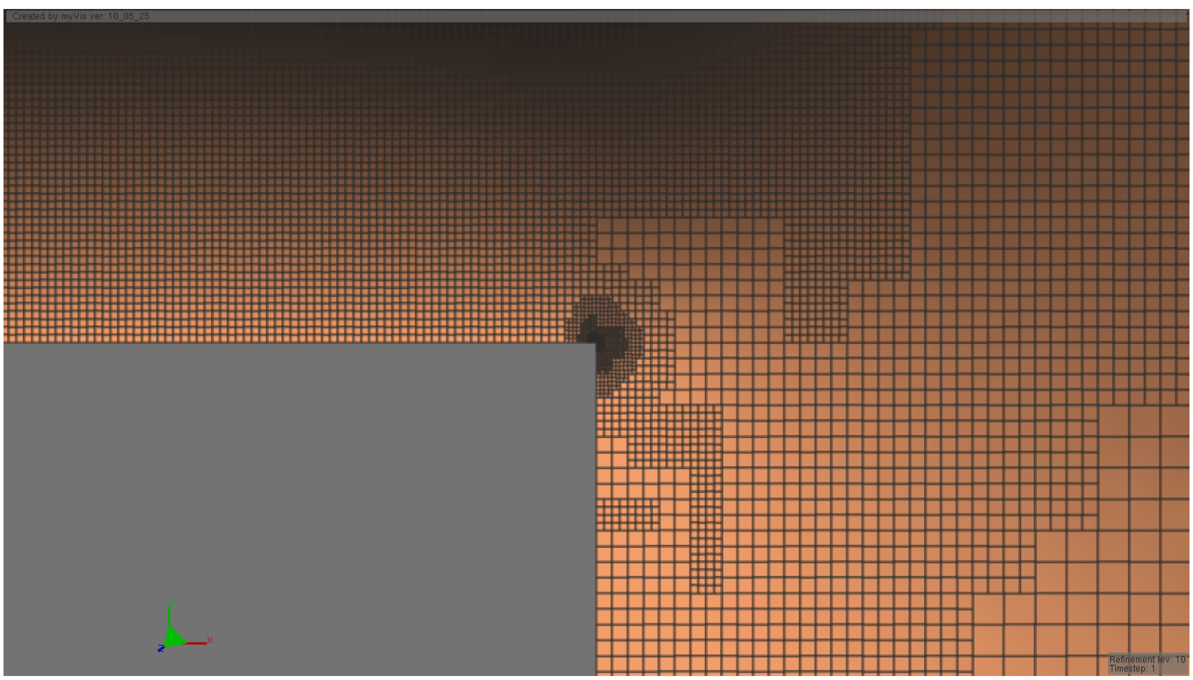


Figure 4.9: Mass conservation versus number of elements at $x = 10.0$, steady state Stokes equations in a backward-step domain.



(a) X-velocity and grid alignment



(b) Grid alignment near the corner singularity

Figure 4.10: Stokes on backward-step Domain: Computed Numerical Solution and Locally Refined Mesh

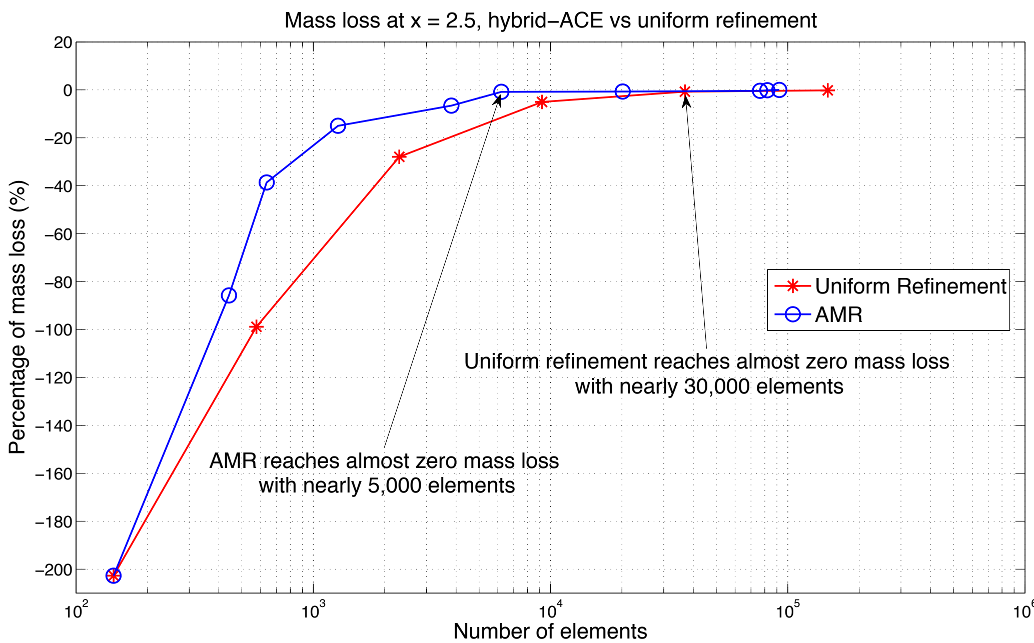


Figure 4.11: Mass conservation versus number of elements at $x = 2.5$, steady state Stokes equations in a backward-step domain.

Chapter 5

Hybrid-FOSLS for Navier-Stokes Equations in a Long Tube

5.1 FOSLS for Nonlinear PDEs

For a general nonlinear first-order PDE system:

$$\mathcal{L}(\mathbf{u}) = \mathbf{f},$$

two approaches can be applied within the FOSLS context. One is called **FOSLS-Newton**, the other is called **Newton-FOSLS**. FOSLS-Newton seeks the minimizer of the nonlinear FOSLS functional,

$$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|_{\mathcal{V}}^2,$$

by solving it with a Newton iteration; while Newton-FOSLS first linearizes the problem then casts the resulting linear problem as a standard FOSLS problem. The two methods only differ by a second Fréchet derivative term. Futhermore, when the numerical solution is close to the exact solution, the lower order derivative terms dominate and the two methods are essentially equivalent. Due to the simplicity of its numerical implementation, Newton-FOSLS is adopted for our numerical tests in later sections. For an application using Newton-FOSLS in a nested iteration context and combined with ACE adaptive refinement, we refer to [2,3], where a study on incompressible resistive magnetohydrodynamics is presented.

First, let's provide some basics of functional calculus.

Definition 2. ***(Fréchet Derivative)*** *Suppose we have Banach spaces X and Y. A functional $F : X \to Y$ is Fréchet differentiable at $x \in X$, if there exists a bounded linear functional $A : X \to Y$*

such that

$$\lim_{h\to 0}\frac{\|F(x+h)-F(x)-Ah\|_Y}{\|h\|_X}=0.$$

If such A exists, then it is unique, and it is called the first Fréchet derivative of F at x, denoted by $DF(x)=A$.

Another weaker definition of differentiability is Gâteaux differentiable.

Definition 3. *(**Gâteaux Derivative**) Suppose X and Y are Banach spaces. A function $F : X \to Y$ on a Banach space X is Gâteaux differentiable at $x \in X$, if for any $h \in X$ the limit*

$$\delta F(x;h)=\lim_{\alpha\to 0}\left(\frac{F(x+\alpha h)-F(x)}{\alpha}\right)$$

exists and $\delta F(x;h)$ is called the first Gâteaux derivative of F at x along the direction h.

Fréchet differentiable is a similar concept as "differentiable" in $\mathcal{R}^n$, while Gâteaux differentiable is a similar concept as "directional differentiable". Similar to what we have for $\mathcal{R}^n$, we have the following theorem.

Theorem 13. *If a functional $F : X \to Y$ from Banach space X to another Banach space Y is Fréchet differentiable at $x \in X$ then it is also Gâteaux differentiable at x. And for any direction $h \in X$*

$$DF(x)=\delta F(x;h).$$

In this chapter, we assume the nonlinear functional is always Fréchet differentiable and we use Gâteaux derivative to compute all the derivatives of the nonlinear functional.

The Taylor expansion of $\mathcal{L}(\mathbf{u})$ can be written as

$$\mathbf{f}=\mathcal{L}(\hat{\mathbf{u}})=\mathcal{L}(\mathbf{u})-\mathcal{L}'(\mathbf{u})[\mathbf{u}-\hat{\mathbf{u}}]+\frac{1}{2}\mathcal{L}''(\mathbf{v})[\mathbf{u}-\hat{\mathbf{u}},\mathbf{u}-\hat{\mathbf{u}}], \tag{5.1}$$

where $\mathbf{v}=\alpha\mathbf{u}+(1-\alpha)\hat{\mathbf{u}}$ with some constant $\alpha\in[0,1]$.

For any $\mathbf{w}\in\mathcal{V}$, $\mathcal{L}'(\mathbf{u})[\mathbf{w}]$ is the first Fréchet derivative along "direction" $\mathbf{w}$ and is defined as

$$\mathcal{L}'(\mathbf{u})[\mathbf{w}]=\lim_{\alpha\to 0}\frac{\mathcal{L}(\mathbf{u}+\alpha\mathbf{w})-\mathcal{L}(\mathbf{u})}{\alpha}.$$

Notice that $\mathcal{L}'(\mathbf{u})[\cdot]$ is an unary operator dependent on $\mathbf{u}$ and $\mathcal{L}''(\mathbf{v})[\cdot,\cdot]$ is a binary operator that depends on $\mathbf{v}$.

5.1.1 FOSLS-Newton and Newton-FOSLS

The FOSLS-Newton method parallels the approach we use for linear problems and seeks the weak solution by minizing some norm (usually L^2-norm) of the residual from the nonlinear equations. That is we find the solution by solving the following minimization problem

Problem 1. *For a certain Hilbert space $\mathcal{V}$, find $\mathbf{u} \in \mathcal{V}$, such that*

$$\mathbf{u} = arg \min_{\mathbf{v}\in\mathcal{V}} \|\mathcal{L}(\mathbf{v}) - \mathbf{f}\|^2$$

Generally, analogous to the operator L in the linear case, if $\mathcal{L}'(\mathbf{v})$ (note again, $\mathcal{L}'(\mathbf{v})$ is a linear operator that is dependent on $\mathbf{v}$) is also continous and coercive in $\mathcal{V}$. Therefore, the solution $\mathbf{u}$ exists and is unique.

Define the FOSLS functional

$$\mathcal{G}(\mathbf{u};\mathbf{f}) = \|\mathcal{L}(\mathbf{u}) - \mathbf{f}\|^2.$$

Since $\mathcal{G}(\mathbf{u};\mathbf{f})$ is a quadratic functional with respect to $\mathbf{u}$, we only need the first Gâteaux derivative to be zero in any "direction" $\mathbf{v}$ in the appropriate Hilbert space $\mathcal{V}$. Let

$$\lim_{\alpha\to 0} \frac{\mathcal{G}(\mathbf{u}+\alpha\mathbf{v};\mathbf{f}) - \mathcal{G}(\mathbf{u};\mathbf{f})}{\alpha} = 0 \tag{5.2}$$

and notice that

$$\begin{aligned}
\mathcal{G}(\mathbf{u}+\alpha\mathbf{v};\mathbf{f}) - \mathcal{G}(\mathbf{u};\mathbf{f}) &= \|\mathcal{L}(\mathbf{u}+\alpha\mathbf{v}) - \mathbf{f}\|^2 - \|\mathcal{L}(\mathbf{u}) - \mathbf{f}\|^2 \\
&= \|\mathcal{L}(\mathbf{u}) + \alpha\mathcal{L}'(\mathbf{u})[\mathbf{v}] + h.o.t - \mathbf{f}\|^2 - \|\mathcal{L}(\mathbf{u}) - \mathbf{f}\|^2 \\
&= \alpha^2 < \mathcal{L}'(\mathbf{u})[\mathbf{v}], \mathcal{L}'(\mathbf{u})[\mathbf{v}] > +2\alpha < \mathcal{L}(\mathbf{u}) - \mathbf{f}, \mathcal{L}'(\mathbf{u})[\mathbf{v}] >,
\end{aligned}$$

where "$h.o.t$" denotes the "higher-order-term" with respect to α, that is quadratic and higher order terms of α which go to zero when α approaches to zero. The second equation is from Taylor expansion of nonlinear functional as mentioned earlier.

The minimization of the FOSLS functional leads to the following weak problem: find $\mathbf{u} \in \mathcal{V}$, such that, for all $\mathbf{v} \in \mathcal{V}$,

$$< \mathcal{L}(\mathbf{u}) - \mathbf{f}, \mathcal{L}'(\mathbf{u})[\mathbf{v}] >= 0. \tag{5.3}$$

On the other hand, the Newton-FOSLS approach is more straight-forward. Denote $\mathbf{u}_n$ as the solution of n'th Newton iteration,

$$\mathbf{r}_n = \mathbf{f} - \mathcal{L}(\mathbf{u}_n),$$

which is the residual equation from the n'th Newton iteration. Also, denote,

$$\delta\mathbf{u}_n = \mathbf{u}_{n+1} - \mathbf{u}_n,$$

as the update equation. Ignore the second derivative term in (5.1), re-arrange (5.1), we obtain the equation to be solved:

$$\mathcal{L}'(\mathbf{u}_n)\delta\mathbf{u}_n = \mathbf{r}_n.$$

For the simplicity of the denotation, we denote $L_n = \mathcal{L}'(\mathbf{u}_n)$ as the Jacobian of n'th Newton iteration for the rest of the dissertation.

Next, we use standard FOSLS method for this linear problem: find $\delta\mathbf{u}_n \in \mathcal{V}$, such that,

$$\min_{\delta\mathbf{u}_n \in \mathcal{V}} \mathbf{f}(\delta\mathbf{u}_n, \mathbf{r}_n) := \|L_n\delta\mathbf{u}_n - \mathbf{r}_n\|^2.$$

The minimization problem induces the weak problem: find $\delta\mathbf{u}_n \in \mathcal{V}$, such that

$$< L_n\delta\mathbf{u}_n,\ L_n\mathbf{v} > \;\; = \;\; < \mathbf{r}_n, L_n\mathbf{v} >, \quad \forall\ \mathbf{v} \in \mathcal{V}. \tag{5.4}$$

It is easy to see (5.4) can also be written as

$$< L(\mathbf{u}_n) + \mathcal{L}'(\mathbf{u}_n)[\delta\mathbf{u}_n] - \mathbf{f}, \mathcal{L}'(\mathbf{u}_n)[\mathbf{v}] >= 0. \tag{5.5}$$

Notice that $L(\mathbf{u}) + \mathcal{L}'(\mathbf{u})[\delta]$ is the Taylor expansion of $\mathcal{L}(\mathbf{u})$ ignoring the higher order terms; thus, (5.5) is only different from (5.3) by higher order derivative terms.

Easy to see that the discrete analogous of 5.5 and 5.3 in finite elements implementation are:

Newton-FOSLS: weak form in finite element space:

find $\boldsymbol{\delta}^h \in \mathcal{V}^h$, such that for all $\mathbf{v}^h \in \mathcal{V}^h$

$$< L_n \delta \mathbf{u}_n^h,\, L_n \mathbf{v}^h > = < \mathbf{f} - L(\mathbf{u}_n^h),\, L_n \mathbf{v}^h >; \tag{5.6}$$

FOSLS-Newton: weak form in finite element space:

find $\boldsymbol{\delta}^h \in \mathcal{V}^h$, such that for all $\mathbf{v}^h \in \mathcal{V}^h$

$$< \mathcal{L}(\mathbf{u}_{n+1}^h) - \mathbf{f},\, L_n \mathbf{v}^h > = \ 0. \tag{5.7}$$

While the weak form of Newton-FOSLS is ready to implement, we have to carry out Newton iteration to actually solve Eqn (5.7), which brings complexities to implementation. Thus, we use Newton-FOSLS method in this dissertation for nonlinear problem,s which should give similar results as FOSLS-Newton method, as we have explained earlier.

5.1.2 Error Estimate for Nonlinear Problems

Under some mild assumptions, the finite element error estimate for the nonlinear problems turns out to be similar to the linear case. The error estimate of the nonlinear FOSLS functional itself is of the same order as the linear FOSLS functional. Moreover, with sufficient regularity, we can still carry out the Aubin-Nitsche trick and get one-order higher error estimate for the L^2 norm of the error. We present our results with more detail in the following theorem and provide proof thereafter.

Theorem 14. *Let $\mathcal{V}$ be a Hilbert space and $\mathcal{V}^h \subset \mathcal{V}$ is a shape-regular and quasi-uniform finite element space that satisfies the standard interpolation bounds. Denote $\mathbf{u}^h$ as the unique solution from either Newton-FOSLS or FOSLS-Newton method. Suppose $\mathcal{L}'(\mathbf{u})$ is a bounded linear operator for any $\mathbf{u} \in \mathcal{V}$ and assume we have done enough Newton iterations, such that the error from nonlinearity is within the order of mesh size, h, then we have*

$$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\| \leq C h^p \|\hat{\mathbf{u}}\|_{p+1}, \tag{5.8}$$

where p is the degree of finite element piecewise polynomial basis functions.

Furthermore, if we assume that $\mathcal{L}'(\mathbf{v})^*\mathcal{L}'(\mathbf{v})$ is H^2 regular for any $\mathbf{v} \in \mathcal{V}$, we have one order higher convergence for the L^2 norm of the error:

$$\|\mathbf{u}^h - \hat{\mathbf{u}}\| \leq Ch^{p+1}\|\hat{\mathbf{u}}\|_{p+1}. \tag{5.9}$$

Proof. Since we have explained earlier that Newton-FOSLS and FOSLS-Newton method are different only by higher order nonlinear terms and asymptotically the same when enough Newton iterations are done, we prove the theorem above using Newton-FOSLS formulation, **i.e.**, we denote by $\mathbf{u}^h$ the unique solution from the minimization problem, we now have:

$$\begin{aligned}
\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\| &= \|\mathcal{L}(\mathbf{u}^h) - \mathcal{L}(\hat{\mathbf{u}})\| \\
&\leq \|\mathcal{L}(\mathcal{I}^h\hat{\mathbf{u}}) - \mathcal{L}(\hat{\mathbf{u}})\| \\
&= \|\mathcal{L}'(\tilde{\mathbf{u}})(\mathcal{I}^h\hat{\mathbf{u}} - \hat{\mathbf{u}})\| \\
&\leq C\|\mathcal{I}^h\hat{\mathbf{u}} - \hat{\mathbf{u}}\|_1 \\
&\leq Ch^p\|\hat{\mathbf{u}}\|_{p+1},
\end{aligned}$$

where $\mathcal{I}^h$ is the interpolant operator to a certain grid Ω^h; $\tilde{\mathbf{u}} = \hat{\mathbf{u}} + \alpha\boldsymbol{\delta}$, with $\alpha \in [0, 1]$. That is, $\tilde{\mathbf{u}}$'s value is somewhere between the exact solution $\hat{\mathbf{u}}$ and numerical solution $\mathbf{u}^h$.

To prove the second half of the theorem, let $\mathcal{L}(\mathbf{u}^h) - \mathcal{L}(\hat{\mathbf{u}}) = \mathcal{L}'(\tilde{\mathbf{u}})[\mathbf{u}^h - \hat{\mathbf{u}}]$, where $\tilde{\mathbf{u}}$'s value is between $\hat{\mathbf{u}}$ and $\mathbf{u}$. Note that, in general, $\tilde{\mathbf{u}}$ here is different $\tilde{\mathbf{u}}$ in the proof above. Denote by $\mathbf{e}^h$ the numerical error from solving the residual equation,

$$\mathcal{L}'(\tilde{\mathbf{u}})\mathbf{e}^h = \mathcal{L}(\mathbf{u}^h) - \mathbf{f}.$$

Consider the dual problem,

$$\mathcal{L}'(\tilde{\mathbf{u}})^*\mathcal{L}'(\tilde{\mathbf{u}})\mathbf{w} = \mathbf{e}^h, \tag{5.10}$$

whose Galerkin closure is

$$< \mathcal{L}'(\tilde{\mathbf{u}})\mathbf{w}, \mathcal{L}'(\tilde{\mathbf{u}})\mathbf{v} >=< \mathbf{e}^h, \mathbf{v} >, \quad \text{for any } \mathbf{v} \in \mathcal{V}. \tag{5.11}$$

Since $\mathbf{e}^h \in \mathcal{V}^h \subset \mathcal{V}$, substituting $\mathbf{v}$ with $\mathbf{e}^h$ in (5.11), we have

$$\begin{aligned} < \mathbf{e}^h, \mathbf{e}^h > &= < \mathcal{L}'(\tilde{\mathbf{u}})\mathbf{w}, \mathcal{L}'(\tilde{\mathbf{u}})\mathbf{e}^h > \\ &= < \mathcal{L}'(\tilde{\mathbf{u}})\mathbf{w}, \mathcal{L}(\mathbf{u}^h) - \mathcal{L}(\hat{\mathbf{u}}) > \\ &= < \mathcal{L}'(\tilde{\mathbf{u}})[\mathbf{w} - \mathbf{z}^h], \mathcal{L}(\mathbf{u}^h) - \mathcal{L}(\hat{\mathbf{u}}) >, \qquad \text{for all } \mathbf{z}^h \in \mathcal{V}^h \\ &\leq \|\mathcal{L}'(\tilde{\mathbf{u}})[\mathbf{w} - \mathbf{z}^h]\| \cdot \|\mathcal{L}(\mathbf{u}^h) - \mathcal{L}(\hat{\mathbf{u}})\| \end{aligned} \tag{5.12}$$

Use the error estimate we developed earlier for the nonlinear FOSLS functional and notice that we have assumed $\mathcal{L}'(\tilde{\mathbf{u}})$ is a bounded linear operator,

$$\begin{aligned} < \mathbf{e}^h, \mathbf{e}^h > &\leq (C \cdot h^p \|\hat{\mathbf{u}}\|_{p+1}) \left(C \cdot h \|\mathbf{w} - \mathbf{z}^h\|_1\right) \\ &\leq C \cdot h^{p+1} \|\hat{\mathbf{u}}\|_{p+1} \|\mathbf{w}\|_2 \\ &\leq C h^{p+1} \|\hat{\mathbf{u}}\|_{p+1} \|\mathbf{e}^h\|. \end{aligned} \tag{5.13}$$

Note that, the last inequality is obtained by using the H^2 regularity of $\mathcal{L}'(\mathbf{v})^* \mathcal{L}'(\mathbf{v})$. □

5.2 FOSLL* for Nonlinear PDEs

The research on FOSLL* method for nonlinear PDEs is relatively new and no publication is available for the moment. In [38], the framework is laid out and the challenges are also discussed. Recall that for a linear problem, $L\mathbf{u} = \mathbf{f}$, FOSLL* method minimizes the FOSLL* functional, $\|L^*\mathbf{w} - \hat{\mathbf{u}}\|^2$, and solves the weak problem

$$< L^*\mathbf{w}, L^*\mathbf{z} >=< \mathbf{f}, \mathbf{z} >, \quad \text{for all} \mathbf{z} \in \mathcal{W}.$$

While for a nonlinear problem, at each Newton step, we minimize

$$\mathcal{G}(\delta \mathbf{w}_n^h; L_n^{-1}\mathbf{r}_n) := \|L_n^* \delta \mathbf{w}^h - L_n^{-1}\mathbf{r}_n\|^2, \tag{5.14}$$

which leads to the weak problem:

Find $\delta \mathbf{w}_n^h \in \mathcal{W}^h$, such that

$$< L_n^* \delta \mathbf{w}_n^h, L_n^* \mathbf{z}^h >=< \mathbf{r}_n, \mathbf{z}^h >, \quad \forall\ \mathbf{z}^h \in \mathcal{W}^h. \tag{5.15}$$

Thus, the solution for Newton iteration at Step $(n+1)$ is

$$\mathbf{u}_{n+1} = \mathbf{u}_n + L_n^* \delta \mathbf{w}_n^h. \tag{5.16}$$

So far, we have been vague about the finite element space $\mathcal{W}^h$ we use. Apparently, $\mathcal{W}^h$ should be a subspace of the domain of L_n^* $\mathcal{D}(L_n^*)$. However, this restriction might not be enough in general. A common piecewise polynomial finite element space is in C^0. Thus, $L_n^* \delta \mathbf{w}^h \notin C^0$, which may cause problems when the Newton iteration continues and the operator $\mathcal{L}'(\mathbf{u}_{n+1})$ needs to be computed. Several possible solutions to this problem are proposed in [38].

The first thought is to use C^1-elements. Although it can circumvent the problem, implementation of C^1-elements is generally much harder and software libraries with such elements are not widely available.

The second approach is to apply L^2-**projection**; that is, at each Newton step, after we have obtained $\delta \mathbf{w}^h$ from FOSLL*, we first compute $L_n^* \delta \mathbf{w}^h$ then solve the minimization problem:
Find $\delta \mathbf{u}_n^h \in \mathcal{V}^h$, such that

$$\min_{\delta \mathbf{u}_n^h \in \mathcal{V}^h} \|L_n^* \delta \mathbf{w}_n^h - \delta \mathbf{u}_n^h\|^2.$$

This induces a weak problem:

$$< L_n^* \delta \mathbf{w}^h - \delta \mathbf{u}^h,\ \mathbf{v}^h >= 0, \quad \forall\ \mathbf{v}^h \in \mathcal{V}^h.$$

To be more exact, let $\{\phi_i\}_{i=1}^n$ be the basis functions of finite element space $\mathcal{V}^h$; thus, $\delta \mathbf{u}^h = \sum_{i=1}^n a_i \phi_i$ and the linear system to be solved is

$$< \sum_{j=1}^n a_j \phi_j\ ,\ \phi_i > = < \delta \mathbf{w}_n^h\ ,\ \mathcal{L}'(\mathbf{u}_n^h)\phi_i > = < \mathcal{L}'(\mathbf{u}_n^h)^* \delta \mathbf{w}_n^h\ ,\ \phi_i > .$$

Another approach proposed in [38] is to combine the FOSLL* and the L^2-projection and solve the **extended system** together. We introduce a weight β and the functional

$$J(\delta \mathbf{w}, \delta \mathbf{u}) = \|L_n^* \delta \mathbf{w} - L_n^{-1} \mathbf{r}_n\|^2 + \beta \|L_n^* \delta \mathbf{w} - \delta \mathbf{u}\|^2. \tag{5.17}$$

The minimization problem is thus: find $(\delta \mathbf{w}_n^h, \delta \mathbf{u}_n^h) \in \mathcal{W}^h \times \mathcal{V}^h$, such that

$$\min_{(\delta \mathbf{w}_n^h, \delta \mathbf{u}_n^h) \in \mathcal{W} \times \mathcal{V}} J(\delta \mathbf{w}_n^h, \delta \mathbf{u}_n^h).$$

This leads to the the discrete weak problem:

Find $(\delta \mathbf{w}_n^h, \delta \mathbf{u}_n^h) \in \mathcal{W}^h \times \mathcal{V}^h$ such that, $\forall\ \mathbf{z}^h \in \mathcal{W}^h$, $\mathbf{v}^h \in \mathcal{V}^h$,

$$< L_n^* \mathbf{z}^h \ , \ L_n^* \delta \mathbf{w}_n^h - L_n^{-1} \mathbf{r}_n > + < L_n^* \mathbf{z}^h \ , \ L_n^* \delta \mathbf{w}_n^h - \delta \mathbf{u}_n^h > + < L_n^* \delta \mathbf{w}_n^h - \delta \mathbf{u}_n^h \ , \ \mathbf{v}^h >= 0, \quad (5.18)$$

which can be simplified as $\forall\ \mathbf{z}^h \in \mathcal{W}^h$, $\mathbf{v}^h \in \mathcal{V}^h$,

$$2 < L_n^* \mathbf{z}^h \ , \ L_n^* \delta \mathbf{w}_n^h > + < L_n \delta \mathbf{w}_n^h \ , \ \mathbf{v}^h > \ = \ < \mathbf{z}^h, \mathbf{r}_n > + < L_n^* \mathbf{z}^h \ , \ \delta \mathbf{u}_n^h > + < \delta \mathbf{u}_n^h, \mathbf{v}^h > . \quad (5.19)$$

5.3 Newton-Hybrid FOSLS for Nonlinear PDEs

Similar to Newton-FOSLS and Newton-FOSLL* methods, for Newton-Hybrid FOSLS, we

- first, linearize the original PDEs, carry out the Newton iteration;
- second, solve the linearized PDEs at each Newton step via Hybrid-FOSLS.

The Hybrid functional at Newton step n is

$$\begin{aligned} \mathcal{H}_n \left((\delta \mathbf{w}^h, \delta \mathbf{u}^h); (L_n^{-1} \mathbf{r}_n, \mathbf{r}_n) \right) \ &:= \ \| L_n^* \delta \mathbf{w}^h - L_n^{-1} \mathbf{r}_n \|^2 \\ &+ \ \| L_n^* \delta \mathbf{w}^h - \delta \mathbf{u}^h \|^2 + \| L_n \delta \mathbf{u}^h - \mathbf{r}_n \|^2. \end{aligned} \quad (5.20)$$

Equivalently, we can think that we have an enlarged first-order system,

$$\begin{aligned} L_n^* \delta \mathbf{w}^h &= L_n^{-1} \mathbf{r}_n, \\ L_n^* \delta \mathbf{w}^h &= \delta \mathbf{u}^h, \\ L_n \delta \mathbf{u}^h &= \mathbf{r}_n. \end{aligned} \quad (5.21)$$

The weak problem from minimizing the Hybrid functional (5.20) is: find $(\delta \mathbf{w}^h, \delta \mathbf{u}^h) \in \mathcal{W}^h \times \mathcal{V}^h$, such that, for any $(\mathbf{z}^h, \mathbf{v}^h) \in \mathcal{W}^h \times \mathcal{V}^h$,

$$< L_n^* \delta \mathbf{w}^h, L_n^* \mathbf{z}^h > + < L_n^* \mathbf{w}^h - \mathbf{u}^h, L_n^* \mathbf{z}^h - \mathbf{v}^h > + < L \mathbf{u}^h, L \mathbf{v}^h >=< \mathbf{f}, \mathbf{z}^h + L \mathbf{v}^h > . \quad (5.22)$$

5.4 Numerical Tests

Consider steady state Navier-Stokes equations:

$$\begin{aligned} \Delta \mathbf{u} - \lambda \mathbf{u} \cdot \nabla \mathbf{u} - \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= g \quad \text{in } \Omega, \\ \mathbf{n} \cdot \mathbf{u} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{5.23}$$

on the follwing domain:

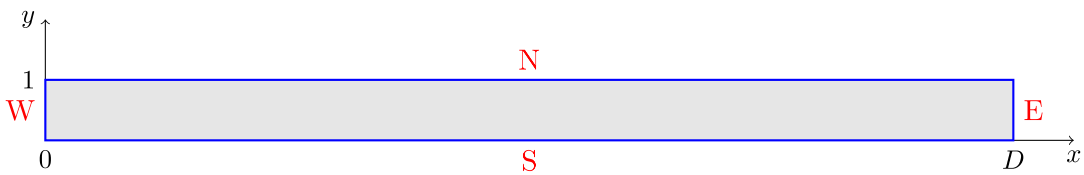


Figure 5.1: Navier-Stokes Equations in a Long Tube

Although for Stokes equations a two-stage velocity-gradient formulation has the H^1-equivalence, for Navier-Stokes equations, due to the nonlinear convection term, $\mathbf{u} \cdot \nabla \mathbf{u}$, $\mathbf{u}$ has to be included in the formulation, which means the two-stage scheme, which was used for Stokes equations in Chapter 3, cannot be used anymore. Thus, in this chapter, we apply velocity-vorticity-pressure formulation for our numerical tests.

Define vorticity $\omega := -\nabla \times \mathbf{u}$. Also, notice that the velocity field is divergence free, we have:

- in $2D$, $\nabla^{\perp}\omega = -\nabla^{\perp}(\nabla \times \mathbf{u}) = \Delta \mathbf{u} - \nabla(\nabla \cdot \mathbf{u}) = \Delta \mathbf{u}$;

- in $3D$, $\nabla^{\perp}\omega = -\nabla \times (\nabla \times \mathbf{u}) = \Delta \mathbf{u} - \nabla(\nabla \cdot \mathbf{u}) = \Delta \mathbf{u}$.

Define the total pressure $P := \frac{\lambda}{2}\mathbf{u}\cdot\mathbf{u}+p$. The **Velocity-Vorticity-Pressure (VVP)** Formulation in $2D$ is as follows:

$$
\begin{aligned}
\begin{pmatrix} 1 & \nabla\times & 0 \\ \nabla^{\perp} & 0 & -\nabla \\ 0 & \nabla\cdot & 0 \end{pmatrix}\begin{pmatrix} \omega \\ \mathbf{u} \\ P \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ -\omega u_2 \\ \omega u_1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ \mathbf{f} \\ 0 \end{pmatrix} \quad \text{in } \Omega, \\
\mathbf{n}\cdot\mathbf{u} &= 0 \quad \text{on } \partial\Omega, \\
\omega &= 0 \quad \text{on } \partial\Omega,
\end{aligned} \tag{5.24}
$$

where velocity $\mathbf{u} = (u_1,\ u_2)^t$ and λ is the Reynolds number which is set to 5 in our test.

Construct an exact solution as follows:

$$
\begin{aligned}
u_1 &= \sin(\pi x/D)\cos(\pi y), \\
u_2 &= -\tfrac{1}{D}\cos(\pi x/D)\sin(\pi y), \\
\omega &= -\pi(1+\tfrac{1}{D})\sin(\pi x/D)\sin(\pi y), \\
P &= \tfrac{\lambda}{2}(u_1^2+u_2^2) - \lambda\tfrac{1+D^2}{8D^2},
\end{aligned} \tag{5.25}
$$

where $-\lambda\frac{1+D^2}{8D^2}$ is a constant, such that the integration of total pressure, P, is 0 on Ω. By enforcing this, we eliminate the null space of any constant in P. Thus, the RHS can be computed:

$$
\begin{aligned}
f1 &= 0, \\
f2 &= -\pi^2\left(1+\tfrac{1}{D^2}\right)\cos(\pi y)\sin(\tfrac{\pi x}{D}) - \tfrac{\lambda\pi}{D}\cos(\tfrac{\pi x}{D})\sin(\tfrac{\pi x}{D}), \\
f3 &= \pi^2\left(\tfrac{1}{D^3}+\tfrac{1}{D}\right)\cos(\tfrac{\pi x}{D})\sin(\pi y) - \tfrac{\lambda\pi}{D^2}\cos(\pi y)\sin(\pi y), \\
f4 &= 0.
\end{aligned} \tag{5.26}
$$

For Eqn (5.24), the linear operator $L_n := \mathcal{L}'(\mathbf{u}_n)$ at n'th Newton iteration is

$$
L_n\begin{pmatrix} \omega \\ u_1 \\ u_2 \\ P \end{pmatrix} = \begin{pmatrix} 1 & -\partial_y & \partial_x & 0 \\ \partial_y & 0 & 0 & -\partial_x \\ -\partial_x & 0 & 0 & -\partial_y \\ 0 & \partial_x & \partial_y & 0 \end{pmatrix}\begin{pmatrix} \omega \\ u_1 \\ u_2 \\ P \end{pmatrix} + \lambda\begin{pmatrix} 0 & 0 & 0 & 0 \\ -u_2 & 0 & -\omega & 0 \\ u_1 & \omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \omega \\ u_1 \\ u_2 \\ P \end{pmatrix}. \tag{5.27}
$$

Define the adjoint variables associated with $(\omega, u_1, u_2, P)^t$ by $(z_1, z_2, z_3, z_4)^t$. Then, the adjoint operator $L_n^* := \mathcal{L}'(\mathbf{u}_n)$ at Newton Step n is

$$L_n^* = \begin{pmatrix} 1 & -\partial_y & \partial_x & 0 \\ \partial_y & 0 & 0 & -\partial_x \\ -\partial_x & 0 & 0 & -\partial_y \\ 0 & \partial_x & \partial_y & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} + \lambda \begin{pmatrix} 0 & -u_2 & u_1 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & -\omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.28}$$

Since the boundary conditions are **normal velocity - (tangential) vorticity** boundary conditions, we use (2.40) as the functional to be minimized for standard FOSLS method.

The comparison of the L^2-error, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$, and the nonlinear functional, $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, is shown in the following tables. We vary the domain length $D = 1,\ 4,\ 8,\ 16,\ 24$ and use both standard FOSLS and Hybrid methods. Note that, a little different from the quantities we measured in Chapter 3, we do not use relative functional or relative error. This is because $\|\mathbf{f}\| = 0$ in this test, which precludes the use of relative functionals. Also, since $\|\hat{\mathbf{u}}\|$ increases mildly as D increases, we should be able to expect similar results using relative L^2-error or not.

D	1	4	8	16	24
$\|\hat{\mathbf{u}}\|$	2.54648	5.41127	10.3451	20.4514	30.6108

Table 5.1: Mild increase of the L^2 norm of the exact solution $\hat{\mathbf{u}}$ as the domain length D increases, using (5.25). This suggests that we can measure the L^2 norm of the exact error, instead of the relative error

The following tables suggest that, while for the nonlinear functional both FOSLS and Hybrid achieve very similar results, for the L^2-error Hybrid achieves much better results, between $\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-3})$ better.

To take a closer look at the convergence rate of the nonlinear functional and L^2-error, we plot the cases when $D = 8,\ 16,\ 24$. Since we use $q2$ (bilinear elements) in our tests, if we assume that the error from Newton iteration can be ignored, we expect the functional converges with order $\mathcal{O}(10^{-2})$ and the L^2-error with order $\mathcal{O}(10^{-3})$. The following 2 sets of figures show that, the nonlinear functional reduction for both FOSLS and Hybrid methods are exactly as expected.

$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$					
h	1/4	1/8	1/16	1/32	1/64
FOSLS	5.3853e-01	1.4660e-01	3.7618e-02	9.4691e-03	2.3714e-03
Hybrid	5.3863e-01	1.4661e-01	3.7618e-02	9.4691e-03	2.3714e-03
$\|\hat{\mathbf{u}} - \mathbf{u}^h\|$					
h	1/4	1/8	1/16	1/32	1/64
FOSLS	5.5941e-02	7.8741e-03	1.2788e-03	3.0018e-04	7.5306e-05
Hybrid	4.4910e-02	5.0651e-03	4.4786e-04	4.3838e-05	5.0022e-06

Table 5.2: Steady state Navier-Stokes equations in a long tube, FOSLS vs. Hybrid: $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, $\|\hat{\mathbf{u}} - \mathbf{u}^h\|$. Domain length $D = 1$, q2-elements

$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$					
h	1/4	1/8	1/16	1/32	1/64
FOSLS	5.8223e-01	1.6123e-01	4.1192e-02	1.0350e-02	2.5906e-03
Hybrid	5.8507e-01	1.6132e-01	4.1193e-02	1.0350e-02	2.5906e-03
$\|\hat{\mathbf{u}} - \mathbf{u}^h\|$					
h	1/4	1/8	1/16	1/32	1/64
FOSLS	2.0155e-01	5.1986e-02	2.6125e-02	1.3389e-02	6.7799e-03
Hybrid	9.6904e-02	8.0832e-03	6.0613e-04	5.2688e-05	5.6198e-06

Table 5.3: Steady state Navier-Stokes equations in a long tube, FOSLS vs. Hybrid: $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, $\|\hat{\mathbf{u}} - \mathbf{u}^h\|$. Domain length $D = 4$, q2-elements

$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$				
h	1/4	1/8	1/16	1/32
FOSLS	8.2899e-01	2.2959e-01	5.8628e-02	1.4728e-02
Hybrid	8.3660e-01	2.2983e-01	5.8632e-02	1.4728e-02
$\|\hat{\mathbf{u}} - \mathbf{u}^h\|$				
h	1/4	1/8	1/16	1/32
FOSLS	4.8198e-01	9.7988e-02	4.5827e-02	2.3326e-02
Hybrid	1.3247e-01	1.0989e-02	8.3125e-04	7.3519e-05

Table 5.4: Steady state Navier-Stokes equations in a long tube, FOSLS vs. Hybrid: $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, $\|\hat{\mathbf{u}} - \mathbf{u}^h\|$. Domain Length $D = 8$, q2-elements

However, while using FOSLS method, the convergence for L^2-error is much slower than $\mathcal{O}(h^3)$. With Hybrid method, the convergence rate for L^2-error is, in fact, slightly faster than what is expected.

$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$				
h	1/4	1/8	1/16	1/32
FOSLS	1.1757e+00	3.2536e-01	8.3075e-02	2.0868e-02
Hybrid	1.1881e+00	3.2578e-01	8.3083e-02	2.0868e-02
$\|\hat{\mathbf{u}} - \mathbf{u}^h\|$				
h	1/4	1/8	1/16	1/32
FOSLS	1.1261e+00	1.8069e-01	7.0906e-02	3.5847e-02
Hybrid	1.7772e-01	1.4567e-02	1.1195e-03	1.0155e-04

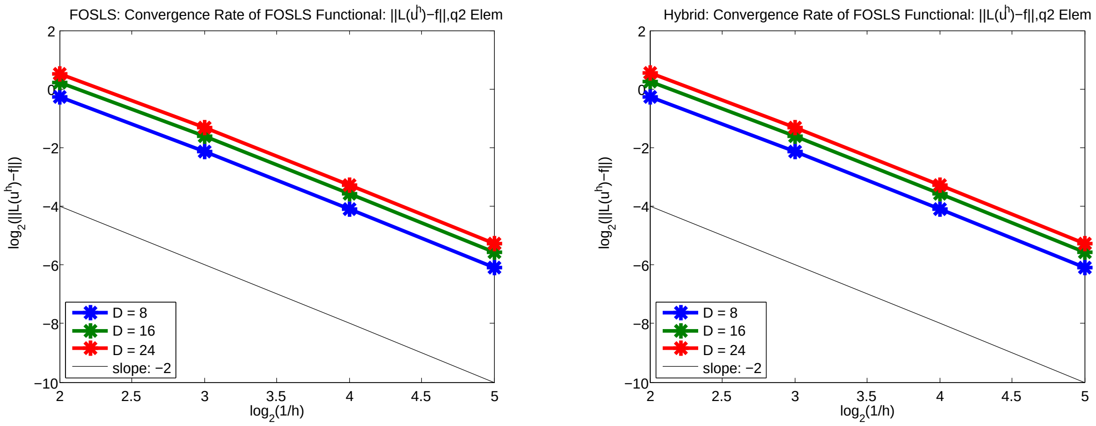
Table 5.5: Steady state Navier-Stokes equations in a long tube, FOSLS vs. Hybrid: $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, $\|\hat{\mathbf{u}} - \mathbf{u}^h\|$. Domain Length $D = 16$, q2-elements.

$\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$				
h	1/4	1/8	1/16	1/32
FOSLS	1.4425e+00	3.9864e-01	1.0178e-01	2.5567e-02
Hybrid	1.4565e+00	3.9918e-01	1.0179e-01	2.5567e-02
$\|\hat{\mathbf{u}} - \mathbf{u}^h\|$				
h	1/4	1/8	1/16	1/32
FOSLS	1.7420e+00	2.7371e-01	8.9738e-02	4.5013e-02
Hybrid	2.2869e-01	1.7482e-02	1.3507e-03	1.2349e-04

Table 5.6: Steady state Navier-Stokes equations in a long tube, FOSLS vs. Hybrid: $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, $\|\hat{\mathbf{u}} - \mathbf{u}^h\|$. Domain Length $D = 24$, q2-elements

The comparison of FOSLS and Hybrid on convergence rate of the nonlinear functional and L^2-error is more clear in the Figures (5.4) . Note that, at the coarsest grid, while using FOSLS, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$ starts at $\mathcal{O}(1)$, for Hybrid, it starts at $\mathcal{O}(10^{-2})$. Hybrid method controls the L^2-error from the beginning while FOSLS fails to do so.

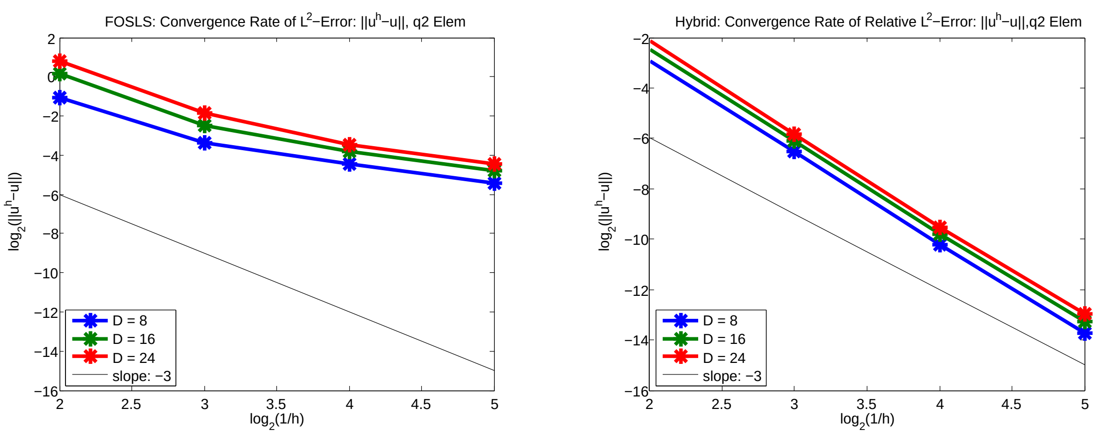
We have to also note here that, although the numerical results are in line with most of our expectations, the L^2-errors from FOSLS method should converge one order faster than FOSLS functional when the solutions get to the asymptotic region. Future work will investigate this.



(a) FOSLS

(b) Hybrid

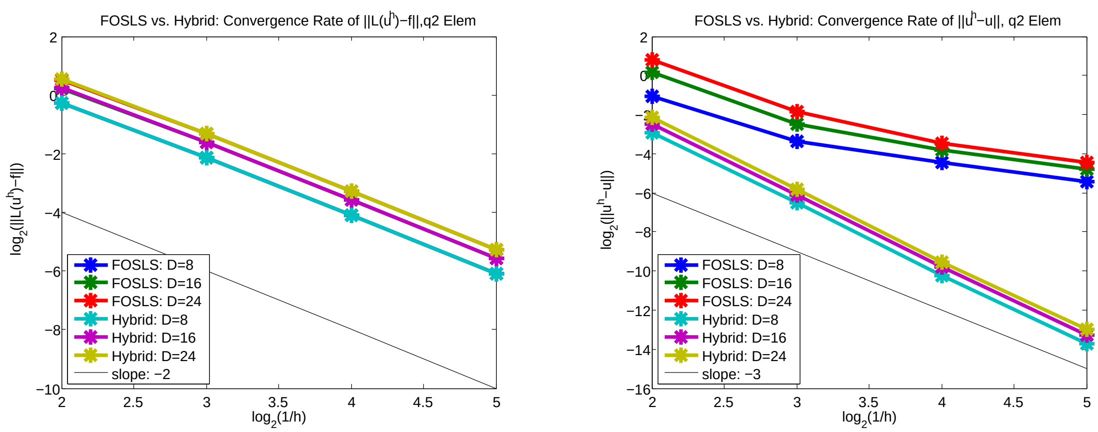
Figure 5.2: Convergence rate of nonlinear functional, $\|\mathcal{L}(\mathbf{u}^h) - \mathbf{f}\|$, using FOSLS and Hybrid, $D = 8,\ 16,\ 24$, $q2$ elements



(a) FOSLS

(b) Hybrid

Figure 5.3: Convergence rate of L^2-error, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$, using FOSLS and Hybrid, domain length $D = 8,\ 16,\ 24$, $q2$ elements



(a) FOSLS

(b) Hybrid

Figure 5.4: Convergence rate of L^2-error, $\|\mathbf{u}^h - \hat{\mathbf{u}}\|$, using FOSLS and Hybrid, domain length $D = 8,\ 16,\ 24$, $q2$ elements

Chapter 6

Conclusions and Future Research

6.1 Conclusions

In this thesis, based on the first-order system least-squares finite element method (FOSLS) and one of its variations, the FOSLL* method, we propose a novel hybrid-FOSLS method that takes advantage of both FOSLS and FOSLL*. We motivate the new method by observing that, while FOSLS and FOSLL* each has its own merits and limitations, they complement each other at the same time. After a lot of work on numerical tests, using different version of Hybrid (**e.g.** involving the intermediate term or not), testing on different problems, the numerical results suggest a good prospect for this method. To understand the mathematics behind the scene that leads to nice results, we explore both analytically and numerically and finally reach the conclusion that the excellent control over L^2-error, one of the main strengths hybrid-FOSLS has is essentially due to the fact that the Hybrid functional is elliptic with a mild coercivity constant in the $\mathcal{H}$-norm that involves the primal variable's L^2-norm. To better understand it, we introduce the graph functional and compare results from minimizing the Hybrid functional to minimizing the graph functional. To offer a rigorous theory of this new method, error estimate of the Hybrid functional's L^2-norm error and the superposition version of them are presented. Efforts are also made to accelerate the convergence, both by using the superposition technique and by fine-tuning parameters in our AMG-preconditioned-Conjugate Gradient solver. Since our standpoint is to develop a discretization scheme that can take full advantage of state-of-art linear system solvers, we studied hybrid-FOSLS in an adaptive mesh refinement setting, where, through a set of numerical tests, it demonstrates its great potential. Preliminary research of Hybrid method has also been carried out for nonlinear equations, such as Navier-Stokes equations. New challenges arise, due to the use of Newton iteration combined with hybrid-FOSLS.

6.2 Future Research

Since the hybrid-FOSLS method developed in the thesis work, and presented in [41], is the first research result in this area, there are a lot of problems to explore.

First, for Hybrid for nonlinear Navier-Stokes equations, although some theoretical numerical results are already presented in this dissertation, work should also be done to develop a rigorous theory for Newton-Hybrid FOSLS method, that should include error estimation that is parallel to what we have developed for the linear problem.

Second, we would like to take more consideration on the linear solver side and possibly modify the first-order formulation and linear solver together and make the solving of the resulting linear system more efficient. For example, we have observed that for weighted hybrid-FOSLS, adding weights into the formulation can slow the solver down and the linear solver convergence rate and becomes unsatisfactory. Also, in nonlinear problems, experience suggests that the number of Newton iterations is generally small (less than 5, on average less than 3). Especially, when it gets to a fine mesh grid, we can further reduce the number of Newton iteration. In [2], one of the core ideas is to solve the complicated MHD equations to the accuracy necessary without over-solving at each level of mesh grid. We can borrow the idea there and implement it with the Hybrid method. Currently, we use the velocity-vorticity formulation for Navier-Stokes equations. This formulation is H^1- elliptic because we use the nonstandard tangential vorticity boundary condition. With the more common velocity boundary conditions, this formulation is not H^1-elliptic. In that case, one can use the velocity gradient-velocity-pressure formulation presented in [12, 13]. However, the adjoint equations become cumbersome. Thus, new formulations for Navier-Stokes equations that are more menable to the Hybrid FOSLS methods are also a interesting direction for future research.

Third, the success of hybrid-FOSLS inspires us to hybrid not only at the functional level but also at the equation level. By careful study of the equations, we can have part of its equation as primal FOSLS equations and other parts using FOSLL* or intermediate equations. Some preliminary tests have already been carried out and to draw any confident conclusions more tests need to be done.

Bibliography

[1] R. A. Adams and J. J. F. Fournier. Sobolev Spaces, second edition. Adademic Press, 2003.

[2] J. Adler. Nested Iteration and First-Order System Least Squares for Incompressible Resistive Magnetohydrodynamics. PhD thesis, University of Colorado at Boulder, 2009.

[3] J. Adler, T. Manteuffel, S. McCormick, J. Nolting, J. Ruge, and L. Tang. An efficiency-based adaptive refinement scheme applied to incompressible, resistive magnetohydrodynamics. Leture Notes in Computer Science, 5910:1–13, 2010.

[4] J. Adler, T. Manteuffel, S. McCormick, J. Nolting, J. Ruge, and L. Tang. Efficiency-based adaptive local refinement for first-order system least-squares formulations. SIAM J. Sci. Comp., 33:1–24, 2011.

[5] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boudary for solutions of elliptic partial differential equations satisfying general boundary conditions ii. Comm. Pure Appl. Math., 17:35–92, 1964.

[6] D.N. Arnold, R.S. Falk, and R. Winther. Preconditioning in h(div) and applications. Mathematics of Computation, 66:957–984, 1997.

[7] D.N. Arnold, R.S. Falk, and R. Winther. Multigrid in h(div) and h(curl). Numerische Mathematik, 85:197–218, 2000.

[8] T. Austin, T. Manteuffel, , and S. McCormick. A robust approach to minimizing h(div) dominated functionals in an h1-conforming finite element space. J. Numer. Lin. Alg. Appl., 11:115–140, 2004.

[9] M. Berndt, T. Manteuffel, and S. McCormick. Local error estimates and adaptive refinement for first-order system least squares (fosls). E.T.N.A., 6:35–43, 1998.

[10] M. Berndt, T. Manteulffel, S. McCormick, and G. Starke. Analysis of first-order system least squares (FOSLS) for elliptic problems with discontinuous coefficients: Part I. SIAM J. Numer. Anal., 43:386–408, 2005.

[11] P. Bochev. Analysis of least squares finite element methods for the navier-stokes equations. SIAM J. Numer. Anal., 34:1817–1844, 1997.

[12] P. Bochev, Z. Cai, T. Manteuffel, and S. McCormick. Analysis of velocity-flux first-order system least-squares principles for the navier-stokes equations: Part i. SIAM J. Numer. Anal., 35:990–1009, 1998.

[13] P. Bochev, Z. Cai, T. Manteuffel, and S. McCormick. Analysis of velocity-flux first-order system least-squares principles for the navier-stokes equations: Part ii. SIAM J. Numer. Anal., 36:1125–1144, 1999.

[14] P. Bochev and M. Gunzburger. Analysis of least squares finite element methods for the stokes equations. Mathematics Of Computation, 63:479–506, 1994.

[15] Pavel B. Bochev and Max D. Gunzburger. Finite element methods of least-squares type. SIAM Review, 40:789–837, 1998.

[16] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak. A least-squares approach based on a discrete minus one inner product for first order system. Math. Comp., 66:935–955, 1997.

[17] S. C. Brenner and L. R. Scott. The Mathematical Theory of Finite Element Methods. Springer-Verlag, 1996.

[18] Z. Cai, T. A. Manteuffel andS. F. McCormick, and S. V. Parter. First-order system least-squares (fosls) for planar linear elasticity: Pure traction problem. SIAM, J. Numer. Anal., 35:320–335, 1998.

[19] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick. First-Order System Least Squares for Second-Order Partial Differential Equations. I. SIAM J. Numer. Anal., 31:1785–1799, 1994.

[20] Z. Cai, B. Lee, , and P. Wang. Least-squares methods for incompressible newtonian fluid flow: Linear stationary problems. SIAM J. Numer. Anal., 42:843–859, 2004.

[21] Z. Cai, T. A. Manteuffel, and S. F. McCormick. First-order System Least Squares for the Stokes Equations, with Application to Linear Elasticity. SIAM J. Numer. Anal., 34:1727–1741, 1997.

[22] Z. Cai, T. A. Manteuffel, and S. F. McCormick. First-order system least squares for the vorticity form of the Stokes equations, with application to linear elasticity. E.T.N.A., 3:150–159, 1997.

[23] Z. Cai, T. A. Manteuffel, and S. F. McCormick. First-order system LL* (FOSLL*): scalar elliptic partial differential equations. SIAM J. Numer. Anal., 39:1418–1445, 2001.

[24] Z. Cai, T. A. Manteulffel, and S. F. McCormick. First-Order System Least Squares for Second-Order Partial Differential Equations. II. SIAM J. Numer. Anal., 34:425–454, 1997.

[25] Lawrence C. Evans. Partial Differential Equations: Second Edition. American Mathematical Society, 1998.

[26] Charles Fefferman. Existence and Smoothness of the Navier-Stokes Equation. `http://www.claymath.org/millennium/Navier-Stokes_Equations/navierstokes.pdf`.

[27] J. M. Fiard, T. Manteuffel, and S. McCormick. First-order system least-squares(fosls) for convection-diffusion problems: numerical results. SIAM J. Sci. Comp., 19:1958–1979, 1998.

[28] J. M. Floryan and L. Czechowski. On the Numerical Treatment of Corner Singularity in the Vorticity Field. J. Comp. Phys., 118:222–228, 1995.

[29] V. Girault and P. Raviart. Finite Element Approximation of the Navier-Stokes Equations. Springer, 1979.

[30] P. Grisvard. Elliptic Problems in Nonsmooth Domains. Pitman, Boston, 1985.

[31] J. Heys, E. Lee, T. A. Manteuffel, and S. F. McCormick. On mass-conserving least-squares methods. SIAM J. Sci. Comp., 28:1675–1693, 2006.

[32] J. Heys, E. Lee, T. A. Manteuffel, and S. F. McCormick. An alternative least-squares formulation of the Navier-Stokes equations with improved mass conservation. J. Comp. Phys., 226:994–1006, 2008.

[33] J. J. Heys, C. G. DeGroff, W. W. Orlando, T. Manteuffel, and S. McCormick. First-order system least squares for elastohydrodynamics with application to flow in compliant blood vessels. Biomed. Sci. Instr., 38:277–282, 2002.

[34] J. J. Heys, T. A. Manteuffel, S. F. McCormick, and J. W. Ruge. Enhanced mass conservation in least-squares methods for Navier-Stokes equations. SIAM J. Sci. Comp., 31:2303–2321, 2009.

[35] J. K. Hunter and B. Nachtergaele. Applied Analysis. World Scientific, 2001.

[36] J.Adler, M.Brezina, T.Manteuffel, S.McCormick, J.Nolting, J.Ruge, and L.Tang. Island Coalescence using Parallel First-Order System Least-Squares on Incompressible Resistive Magnetohydrodynamics.,. SIAM J. Sci. Comput. (SISC), 2012. Submitted.

[37] Alois Kufner. Weighted Sobolev Spaces. John Wiley & Sons Ltd, 1985.

[38] E. Lee and T. Manteuffel. FOSLL* for nolinear partial differential equations. in preparation.

[39] E. Lee, T. A. Manteuffel, and C. R. Westphal. Weighted-norm First-order System Least Squares (FOSLS) For Problems with Corner Singularities. SIAM J. Numer. Anal., 44:1974–1996, 2006.

[40] E. Lee and T. A. Manteulffel. FOSLL* Method for the Eddy Current Problem with Three-dimensional Edge Singularities. SIAM J. Numer. Anal., 45:787–809, 2007.

[41] K. Liu, T. Manteuffel, S. McCormick, J. Ruge, and L. Tang. Hybrid first-order system, least squares (Hybrid-FOSLS) finite element methods with applicaiton to Stokes equations. SIAM, Numer. Anal., 2012. in review.

[42] T. A. Manteuffel, S. F. McCormick, and C. Pflaum. Improved discretization error estimates for first-order system least-squares (fosls). J. of Numer. Math., 11:163–177, 2003.

[43] T. A. Manteuffel, S. F. McCormick, J. Ruge, and J. G. Schmidt. First-order system LL* (FOSLL*) for general scalar elliptic problems in the plane. SIAM J. Num. Anal., 43:2098–2120, 2006.

[44] T. A. Manteulffel, S. F. McCormick, J. Ruge, and J. G. Schmidt. First-order system LL* (FOSLL*) for general scalar elliptic problems in the plane. SIAM J. Numer. Anal., 43:2098–2120, 2006.

[45] S. F. McCormick. Multilevel projection methods for partial differential equations. CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[46] O. Pironneau. Finite Element Methods For Fluids. Masson, Paris, 1989.

[47] J. W. Ruge. Fospack Users Manual. unpublished, 2000.

[48] Yousef Saad. Iterative Methods for Sparse Linear Systems. SIAM, 2003.

[49] H. De Sterck, T. Manteuffel, S. McCormick, J. Nolting, J. Ruge, and L. Tang. Efficiency-based h- and hp-refinement strategies for finite element methods. Numer. Lin. Alg. App., 15:84–114, 2008.

[50] L. Tang. Parallel Efficiency-based Adaptive Local Refinement. PhD thesis, University of Colorado at Boulder, 2010.

[51] U.Trottenberg, C.W.Oosterlee, and A Schüller. Multigrid. Academic Press, 2001.

[52] S. Vogel. Life in Moving Fluids: The Phsical Biology of Flow (Second Edition). Princeton University Press, 1996.

[53] W.L.Briggs, V.E.Henson, and S.F.McCormick. A Multigrid Tutorial, second edition. SIAM, 2000.

[54] K. Yosida. Functional Analysis. Springer, 1980.

Appendix A

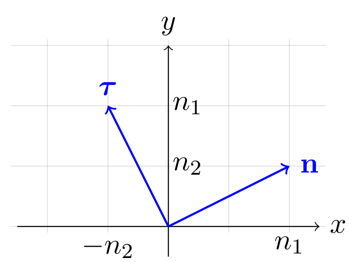
Vector Calculus

A.1 Notations

In 2D, denote:

- $\mathbf{n}$: the unit exterior normal vector along $\partial\Omega$;
- $\boldsymbol{\tau}$: the unit tangent vector.

The direction of $\boldsymbol{\tau}$ depends on the direction of $\partial\Omega$. Convetionally, $\partial\Omega$ is counter clockwise, thus, if $\mathbf{n} = (n_1, n_2)$, then $\boldsymbol{\tau} = (-n_2, n_1)$ as illustrated in the following plot.



Thus for a 2D vector $\mathbf{v} = (v_1, v_2)$,

$$\mathbf{v} \cdot \boldsymbol{\tau} = -v_1 n_2 + v_2 n_1 = \begin{vmatrix} n_1 & n_2 \\ v_1 & v_2 \end{vmatrix} = \mathbf{n} \times \mathbf{v}$$

Let the curl operator $\nabla\times$ be

$$\begin{aligned}\nabla\times\mathbf{u} &= \partial_x u_2 - \partial_y u_1, \quad \text{in 2D},\\ \nabla\times\mathbf{u} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ u_1 & u_2 & u_3 \end{vmatrix},\end{aligned}$$

where $\hat{i}$, $\hat{j}$, $\hat{k}$ are unit vectors in the x, y, z directions respectively. Denote the "Grad-perp" operator $\nabla^{\perp}$ as

$$\nabla^{\perp} q = \begin{pmatrix} \partial_y q \\ -\partial_x q \end{pmatrix}. \tag{A.1}$$

$\nabla^{\perp}$ maps a scalar function q to a vector function in 2D.

The concept of the "adjoint operator" or "dual operator" is used a lot in FOSLS theory. Recall that if $< L\mathbf{u}, \mathbf{v} > = < \mathbf{u}, L^*\mathbf{v} >$, L^* is called the adjoint (dual) of L and $< \cdot, \cdot >$ is genrally the L^2-innerproduct. The most common adjoint pairs are

- ∇ and $-\nabla\cdot$,
- $\nabla\times$ and $\nabla^{\perp}$ in 2D,
- $\nabla\times$ and $\nabla\times$ in 3D,

which are straight forward by the virtue of Green's formula (Theorem 15).

A.2 Calculus Facts

Let $\mathbf{u}$ denote to be a vector function with enough smoothness in 2D or 3D and q to be a scalar function, we have the following facts that is used multiple times in this thesis. **Zero operators:**

$$\begin{aligned}
\nabla \cdot \nabla^{\perp} &= \nabla \cdot \begin{pmatrix} \partial_y \\ -\partial_x \end{pmatrix} = \partial_{yx} - \partial_{xy} = 0 && \text{(A.2)} \\
\nabla \cdot \nabla \times &= \begin{vmatrix} \partial_x & \partial_y & \partial_z \\ \partial_x & \partial_y & \partial_z \end{vmatrix} = 0 && \text{(A.3)} \\
\nabla \times \nabla &= \begin{vmatrix} \partial_x & \partial_y \\ \partial_x & \partial_y \end{vmatrix} = 0 && \text{(A.4)}
\end{aligned}$$

Some equations:

$$\begin{aligned}
\Delta \mathbf{u} &= \nabla \cdot \nabla \mathbf{u} && \text{(A.5)} \\
\nabla^{\perp} \nabla \times \mathbf{u} &= -\Delta \mathbf{u} + \nabla \left(\nabla \cdot \mathbf{u} \right) && \text{(A.6)} \\
\nabla \times \nabla \times \mathbf{u} &= -\Delta \mathbf{u} + \nabla \left(\nabla \cdot \mathbf{u} \right) && \text{(A.7)}
\end{aligned}$$

Eqn(A.6) can be verified as follows and the readers should be able to verify the others similarly.

$$\begin{aligned}
\nabla^{\perp} \nabla \times \mathbf{u} &= \nabla^{\perp} \begin{vmatrix} \partial_x & \partial_y \\ u_1 & u_2 \end{vmatrix} = \nabla^{\perp} (\partial_x u_2 - \partial_y u_1) \\
&= \begin{pmatrix} \partial_{xy} u_2 - \partial_{yy} u_1 \\ -\partial_{xx} u_2 + \partial_{yx} u_1 \end{pmatrix} = \begin{pmatrix} -\Delta u_1 \\ -\Delta u_2 \end{pmatrix} + \begin{pmatrix} \partial_x \nabla \cdot \mathbf{u} \\ \partial_y \nabla \cdot \mathbf{u} \end{pmatrix} \\
&= -\Delta \mathbf{u} + \nabla \left(\nabla \cdot \mathbf{u} \right)
\end{aligned}$$

Appendix B

Functional Analysis Results

B.1 Function Spaces

Denote Ω as an open subset of $\mathcal{R}^d$, where $d = 2, 3$ in this dissertation. Define $\mathcal{D}(\Omega)$ to be the linear space of infinitely differentiable functions which are defined on Ω and with compact supports, and

$$\mathcal{D}(\bar{\Omega}) = \{\phi|_{\bar{\Omega}} : \phi \in \mathcal{D}(\mathcal{O})\text{for some open subset } \Omega \subset \mathcal{O} \subset \mathcal{R}^d\}.$$

Let

$$\begin{aligned} H(div;\Omega) &= \{\mathbf{v} \in L^2(\Omega)^d : \nabla \cdot \mathbf{v} \in L^2(\Omega)\}, \\ H(curl;\Omega) &= \{\mathbf{v} \in L^2(\Omega)^d : \nabla \times \mathbf{v} \in L^2(\Omega)^{2d-3}\}, \end{aligned}$$

which are Hilbert spaces with the following norms respectively

$$\begin{aligned} \|\mathbf{v}\|_{H(div;\Omega)} &:= \left(\|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla \cdot \mathbf{v}\|_{0,\Omega}^2\right)^{\frac{1}{2}}, \\ \|\mathbf{v}\|_{H(curl;\Omega)} &:= \left(\|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla \times \mathbf{v}\|_{0,\Omega}^2\right)^{\frac{1}{2}}. \end{aligned}$$

B.2 Some Theorems

Theorem 15. ***Green's Formula:***

Let $\phi \in \mathcal{D}(\bar{\Omega})$ *and* $\mathbf{v} \in \mathcal{D}(\bar{\Omega})^d$,

- ***Gradient-Divergence Form:***

$$< \mathbf{v}, \nabla\phi >= - < \nabla \cdot \mathbf{v}, \phi > + \int_{\partial\Omega} \mathbf{n} \cdot \mathbf{v}\phi, \tag{B.1}$$

- ***Curl-Curl Form:***

$$< \mathbf{v}, \nabla^{\perp}\phi > \;=\; < \nabla\times\mathbf{v}, \phi > - \int_{\partial\Omega} \mathbf{n}\times\mathbf{v}\phi, \tag{B.2}$$

$$< \mathbf{v}, \nabla\times\boldsymbol{\phi} > \;=\; < \nabla\times\mathbf{v}, \boldsymbol{\phi} > - \int_{\partial\Omega} \mathbf{n}\times\mathbf{v}\cdot\boldsymbol{\phi}. \tag{B.3}$$

Remark 10. *The theorem above can be extended to less smooth functions by interpreting functions' values on the boundary as their traces. For example, for Gradient-Divergence Form, $\mathbf{v} \in H(div;\Omega)$, $\phi \in H^1(\Omega)$; for Curl-Curl Form, $\mathbf{v} \in H(curl;\Omega)$, $\phi \in H^1(\Omega)$, $\boldsymbol{\phi} \in H^1(\Omega)^3$ are also sufficient assumptions for the theorem. For the details, please refer to Chapter 1, [29].*

Theorem 16. ***(Helmholtz)*** *Let Ω be a bounded Lipschitz domain with boundary $\partial\Omega$. There exists $\mathbf{u} \in (L^2(\Omega))^3$ satisfying*

$$\nabla\cdot\mathbf{u} = 0 \ in\ \Omega, \quad \int_{\partial Omega} \mathbf{u}\cdot\mathbf{n} = 0,$$

if and only if there exists $\mathbf{w} \in (H^1(\Omega))^3$, such that $\mathbf{u} = \nabla\times\mathbf{w}$. Furthermore, $\mathbf{w}$ can be chosen to satisfy $\nabla\cdot\mathbf{w} = 0$ and

$$\|\mathbf{w}\|_{(H^1(\Omega))^3} \leq C\|bu\|_{(L^2(\Omega))^3}.$$

Aubin-Nitsche Trick:

The Aubin-Nitsche trick (also called the Aubin-Nitsche duality argument) can be used to estimate the convergence rate of the error's L^2 norm. We have mentioned it in Chapter 3 without giving the proof. Here we provide more details on the method.

First, consider the Galerkin weak form of $L^*L\mathbf{w} = \mathbf{e}^h$, where $\mathbf{e}^h = \mathbf{u}^h - \hat{\mathbf{u}}$. Also, assume that L^*L is an H^2 regular operator.

$$< L^*L\mathbf{w}, \mathbf{v} >=< L\mathbf{w}, L\mathbf{v} >=< \mathbf{e}^h, \mathbf{v} >, \quad \forall\mathbf{v}\in\mathcal{V}. \tag{B.4}$$

Second, note that minimizing FOSLS functional, $\|L\mathbf{u}-\mathbf{f}\|^2$, leads to weak form,

$$< L\mathbf{u}^h - \mathbf{f}, L\mathbf{v}^h >=< L\mathbf{e}^h, L\mathbf{v}^h >= 0, \quad \forall\mathbf{v}^h\in\mathcal{V}^h. \tag{B.5}$$

Let $\mathbf{v} = \mathbf{e}^h$ in (B.4), note that $\mathcal{I}^h\mathbf{w} \in \mathcal{V}^h$ and use (B.5) we have

$$\begin{aligned}
\|\mathbf{e}^h\|^2 &= < L\mathbf{w}, L\mathbf{e}^h > \\
&= < L(\mathbf{w} - \mathcal{I}^h\mathbf{w}), L\mathbf{e}^h > \\
&\leq \|L(\mathbf{w} - \mathcal{I}^h\mathbf{w})\|\|L\mathbf{e}^h\| \\
&\leq c_1^2\|\mathbf{w} - \mathcal{I}^h\mathbf{w}\|_1\|\mathbf{u}^h - \hat{\mathbf{u}}\|_1 \\
&\leq c_1^2 c_I h\|\mathbf{w}\|_2\|\mathbf{u}^h - \hat{\mathbf{u}}\|,
\end{aligned}$$

where, c_1 is continuity constant of operator L, c_I is interpolation constant.

Now, assume we have H^2 regularity of $L^*L\mathbf{w} = \mathbf{e}^h$, that is $\|\mathbf{w}\|_2 \leq c_r\|L^*L\mathbf{w}\|$ for some regularity constant c_r. Also notice that from (B.4), it is easy to get $\|L^*L\mathbf{w}\| = \|\mathbf{e}^h\|$,then,

$$\|\mathbf{e}^h\|^2 \leq c_1^2 c_I c_r h\|\mathbf{e}^h\| \cdot \|\mathbf{u}^h - \hat{\mathbf{u}}\|_1$$

Eliminate $\|\mathbf{e}^h\|$ both sides and apply standard finite element estimate, we obtain the final estimate:

$$\|\mathbf{u}^h - \hat{\mathbf{u}}\| \leq Ch^{q+1}\|\hat{\mathbf{u}}\|_{q+1} \tag{B.6}$$

Appendix C

Obtaining Adjoint Operator L^*

Obtaining the adjoint operator is an important step in FOSLL* and Hybrid-FOSLS methods. It is not only to get L^* itself, but also to get boundary conditions of adjoint variables based on primal variables. One of the key points is that we always assume homogeneous boundary conditions for primal variables and recover the numerical solution of primal variables with non-homogeneous boundary conditions by superposition. In this chapter, we illustrate the solving process by an example and present a matlab script which can automate the sometimes tedious process.

C.1 An Example

Integration-by-parts formula is essential in both obtaining the L^* operator itself and the boundary conditions for adjoint variables. We will use the theorem repeatedly in the example later.

Theorem 17. ***(Integration-by-parts)*** *Let u, $v \in C^1(\overline{\Omega})$. Then*

$$\int_\Omega u_{x_i} v\, dx \;=\; -\int_\Omega u v_{x_i}\, dx + \int_{\partial\Omega} u v n_I ti\, dS, \tag{C.1}$$

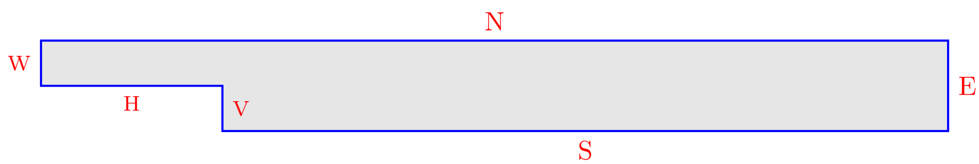
where $i = 1, \ldots, k$, unit normal vector $\mathbf{n} = (n_1, \ldots, n_k)$.

Recall that the adjoint operator L^* is defined by $< L\mathbf{u}, \mathbf{w} >=< \mathbf{u}, L^*\mathbf{w} >$, with properly assigned boundary values of $\mathbf{w}$. It is easy to see from the theorem that to obtain L^* is trival, since for a differential operator, ∂x_i, on a primal variable, we simply nagate the sign for the according adjoint variable; for a function or constant that is the coefficient of the primal variable, we keep the function(constant) unchanged for the adjoint variable. Thus L^* is transpose of L, with all its differential operators' sign negated.

Suppose we have the primal variable, $\mathbf{u} = (U_{11}, U_{12}, U_{21}, U_{22}, p, q)^t$ and the adjoint variable, $\mathbf{w} = (W_{11}, W_{12}, W_{21}, W_{22}, r, s)$. Boundary conditions for primal variables are

$$\boldsymbol{\tau} \cdot (U_{11}, U_{12}) = \boldsymbol{\tau} \cdot (U_{21}, U_{22}) = 0,$$
$$\mathbf{n} \cdot (U_{11}, U_{21}) = \mathbf{n} \cdot (U_{12}, U_{22}) = 0.$$

Given the domain Ω is as follows:



The homogeneous boundary conditions can be further written as

$$\begin{aligned}
U_{11} &= 0, \quad \text{on [N,S,E,W,H,V]},\\
U_{12} &= 0, \quad \text{on [E,W,V]},\\
U_{21} &= 0, \quad \text{on [N,S,H]},\\
U_{22} &= 0, \quad \text{on [N,S,E,W,H,V]}.
\end{aligned}$$

The primal operator L in its simplied form is:

$$L = \begin{array}{c} \\ (1)\\ (2)\\ (3)\\ (4)\\ (5)\\ (6) \end{array} \begin{array}{c} \begin{array}{cccccc} U_{11} & U_{12} & U_{21} & U_{22} & p & q \end{array} \\ \begin{pmatrix} \partial_x & \partial_y & & & -\partial_x & \\ \partial_y & -\partial_x & & & & -\partial_x \\ & & \partial_x & \partial_y & -\partial_y & \\ & & \partial_y & -\partial_x & & -\partial_y \\ \partial_x & & \partial_y & & & \\ & \partial_x & & \partial_y & & \end{pmatrix} \end{array}. \tag{C.2}$$

Easily, we have,

$$L^* = \begin{array}{c} \\ (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \begin{array}{c} \begin{array}{cccccc} W_{11} & W_{12} & W_{21} & W_{22} & r & s \end{array} \\ \begin{pmatrix} -\partial_x & -\partial_y & & & -\partial_x & \\ -\partial_y & \partial_x & & & & -\partial_x \\ & & -\partial_x & -\partial_y & -\partial_y & \\ & & -\partial_y & \partial_x & & -\partial_y \\ \partial_x & & \partial_y & & & \\ & \partial_x & & \partial_y & & \end{pmatrix} \end{array}. \tag{C.3}$$

To get boundary conditions for adjoint variables, integration-by-parts formula is needed. Boundary conditions for W_{11} is obtained using Eqn (1) of (C.2).

$$\begin{aligned} \int_\Omega (\partial_x(U_{11}) + \partial_y(U_{12}) - \partial_x(p)) \cdot W_{11} &= \int_\Omega -\partial_x(W_{11})U_{11} - \partial_y(W_{11})U_{12} + \partial_x(W_{11})p \\ &+ \int_{\partial\Omega} U_{11}W_{11}n_x + U_{12}W_{11}n_y - pW_{11}n_x, \end{aligned}$$

where the unit normal vector $\mathbf{n} = (n_x, n_y)$ (For instance, on Boundary [S], $(n_x, n_y) = (0, -1)$).

Setting the boundary term, $\int_{\partial\Omega} U_{11}W_{11}n_x + U_{12}W_{11}n_y - pW_{11}n_x = 0$ and conduct the same approach to other equations of (C.2), we have

$$\begin{aligned} W_{11} &= 0 \quad \text{on [N,S,E,W,H,V]}, \\ W_{12} &= 0 \quad \text{on [E,W,V]}, \\ W_{21} &= 0 \quad \text{on [N,S,E,W,H,V]}, \\ W_{22} &= 0 \quad \text{on [N,S,H]}. \end{aligned}$$

C.2 Matlab Script

The following matlab script that computes the boundary conditions of adjoint variables is based on the solving process described in previous section. The comments and the code should be self-explanatory.

```
%=====================================================
%  AdjointBC: compute BCs for adjoint variables
%=====================================================
%  Author:
%         Kuo Liu, University of Colorado at Boulder,
%  Date:
%         November, 2011
%=====================================================
%INPUT:
%  m:
%       number of primal/adjoint variables (they should be the same)
%  s:
%       number of sides/boundaries
%  A(m,m):
%       Stores primal operator matrix L.
%       A(i,j) = 0 : if constant in L(i,j);
%              = 1 : if dx       in L(i,j);
%              = 2 : if dy       in L(i,j)
%  B(m,s):
%       Stores BCs for primal variables.
%       B(i,j) = 0 : i'th primal variable has value on Side j
%              = 1 : otherwise
%  C(2,s):
%       Stores unit normal vector's info: |nx|,|ny|
%       C(1,j) = 0 : nx = 0 on Side j;
%              = 1 : otherwise
%       C(2,j) = 0 : ny = 0 on Side j;
%              = 1 : otherwise
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
%OUTPUT:
%  D(m,s):
%       Stores BCs for adjoint variables
%       D(i,j) ~= 0 : i'th adjoint variable = 0 on Side j
%               = 0 : no BC for the variable on Side j

A = dlmread('Ac2.mat');
B = dlmread('Bp.mat');
C = dlmread('C1.mat');

m = size(B,1);
s = size(B,2);
D = zeros(m,s);

for i=1:m
    for j = 1:m
        if A(i,j)==0
            continue;
```

```
        elseif A(i,j)==3
            D(i,:)=D(i,:) + C(1,:).*B(j,:) + C(2,:).*B(j,:);
        else
            D(i,:)=D(i,:) + C(A(i,j),:).*B(j,:);
        end
    end
end

D
```

C.3 Relation of r^α and $H^m(\Omega)$

A function, say u at the singular point can always be expanded as $u = r^\alpha \cdot$ (trig-functions). If we require $u \in H^m(\Omega)$, since $u^{(m)} = c \cdot r^{\alpha-m} \cdot$ (trig-functions), we need

$$\iint (r^{\alpha-m})^2 \cdot r\ dr\ d\theta = \frac{r^{\alpha-m+1}}{2(\alpha-m+1)}|_0^R < \infty, \tag{C.4}$$

that is, $m < \alpha + 1$.

Let $\phi(r,\theta) = r^\alpha \sin(\alpha\theta)$.

$$\begin{aligned}
\nabla\phi &= \frac{\partial(r,\theta)}{\partial(x,y)}\begin{pmatrix}\partial_r\\ \partial_\theta\end{pmatrix}\phi \\
&= \begin{pmatrix} \cos\theta & -\frac{1}{r}\sin\theta \\ \sin\theta & \frac{1}{r}\cos\theta \end{pmatrix}\begin{pmatrix}\partial_r\\ \partial_\theta\end{pmatrix}\phi \\
&= \alpha r^{\alpha-1}\begin{pmatrix}\cos\theta\sin(\alpha\theta) - \sin\theta\cos(\alpha\theta)\\ \sin\theta\sin(\alpha\theta) + \cos\theta\cos(\alpha\theta)\end{pmatrix} \\
&= \alpha r^{\alpha-1}\begin{pmatrix}\sin(\alpha-1)\theta\\ \cos(\alpha-1)\theta\end{pmatrix}
\end{aligned}$$

Therefore,

$$\iint (\nabla\phi)^2 = \alpha^2 \int_0^R r^{2(\alpha-1)} r dr \int_0^\omega \begin{pmatrix}\sin^2(\alpha-1)\theta\\ \cos^2(\alpha-1)\theta\end{pmatrix} d\theta$$

- If we require $\phi \in \mathcal{H}^1$,

$$\int_0^R (r^{2(\alpha-1)}) r dr = \int_0^R r^{2\alpha-1} dr = \left.\frac{r^{2\alpha}}{2\alpha}\right|_0^R = \frac{R^{2\alpha}}{2\alpha} < \infty$$

That is, we need $\alpha > 0$

- If we require $\phi \in \mathcal{L}^2$,

$$\int_0^R (r^\alpha)^2 r dr = \int_0^R r^{2\alpha+1} dr = \left.\frac{r^{2\alpha+2}}{2\alpha+2}\right|_0^R < \infty$$

  we need $2\alpha + 2 > 0$, that is $\alpha > -1$.

Equation C.4 offers a quick reference for the regularity of a singular solution in its expanded form. We can tell which Sobolev space the function lives in by simiply looking at the leading term of its expanded form with the help of C.4.