Monolithic Weighted Least-Squares Finite Element Method for non-Newtonian Fluids with non-isothermal Effects

---

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

Der Fakultät für Mathematik der
Technischen Universität Dortmund
vorgelegt von

Muhammad Waseem

September 2020

Monolithic Weighted Least-Squares Finite Element Method for non-Newtonian Fluids with non-isothermal Effects.

Fakultät für Mathematik
Technische Universität Dortmund

Erstgutachter: Prof. Dr. Stefan Turek

Zweitgutachter: Prof. Dr. Dmitri Kuzmin

Tag der mündlichen Prüfung: 23.02.2022

Abstract

We study the monolithic finite element method, based on the least-squares minimization principles for the solution of non-Newtonian fluids with non-isothermal effects. The least-squares functionals are balanced by the linear and nonlinear weighted functions and the residuals comprised of $\mathcal{L}^2$-norm only. The weighted functions are the function of viscosities and proved significant for optimal results. The lack of mass conservation is an important issue in LSFEM and is studied extensively for the diverse range of weighted functions. Therefore, we consider only inflow/outflow problems. We use the Krylov subspace linear solver, i.e. conjugate gradient method, with a multigrid method as a preconditioner. The SSOR-PCG is used as smoother for the multigrid method. The Gauss-Newton and fixed point methods are employed as nonlinear solvers. The LSFEM is investigated for two main types of fluids, i.e. Newtonian and non-Newtonian fluids.

The stress-based first-order systems, named SVP formulations, are employed to investigate the Newtonian fluids. The different types of quadratic finite elements are used for the analysis purposes. The nonlinear Navier-Stokes problem is investigated for two mesh configurations for flow around cylinder problem. The coefficients of lift/drag, pressure difference, global mass conservation are analyzed. The comparison of linear and nonlinear solvers, based on iterations, is presented as well. The analysis of non-Newtonian fluids is divided into two parts, i.e. isothermal and non-isothermal. For the non-Newtonian fluids, we consider only Q_2 finite elements for the discretization of unknown variables. The isothermal non-Newtonian fluids are investigated with SVP-based formulations. The power law and Cross law viscosity models are considered for investigations with different nonlinear weighted functions. We study the flow parameters for flow around cylinder problem along with the mass conservation for shear thinning and shear thickening fluids. To study the non-isothermal non-Newtonian fluids, we introduced a new first-order formulation which includes temperature and named it as SVPT formulation. The non-isothermal effects are obtained due to the additional viscous dissipation in the fluid flow and from the preheated source as well. The flow around cylinder problem is analyzed for a variety of flow parameters for Cross law fluids. It is shown that the MPCG solver generates very accurate results for the coupled and highly complex problems.

Key words: Monolithic, Weighted LSFEM, Krylov subspace, Multigrid-PCG, SVP, SVPT, Navier-Stokes equation, non-Newtonian fluids, non-isothermal

To my parents,
my wife,
and lovely son Ibaad.

Acknowledgment

I would like to offer my sincere thanks to my advisor Prof. Dr. Stefan Turek for his support, trust, guidance and encouragement during the span of my research work. I am grateful that he provided me an opportunity to explore the field of computational mathematics under his kind supervision. Despite his busy routine, he was always available for discussions and provided his invaluable advises. I am also grateful to Prof. Dr. Dmitri Kuzmin for being a reviewer of my thesis. I am truly grateful to Prof. Dr. Christian Kreuzer and Prof. Dr. Lorenz Schwachhöfer for being the part of my examination committee.

I would like to express my special gratitude to Dr. Abdurrahim Ouazzi for introducing the Least-squares FEM to me. I am thankful for the discussions and guidance that he provided me to understand the theoretical aspect regarding my work. I would like to express my appreciation to all the members at the Institut für Angewandte Mathematik und Numerik LS III, TU Dortmund, for providing the friendly environment. I am extremely grateful to Dr. Mudassar Razzaq, Dr. Absaar Ul Jabbar, Dr. Ramzan Ali, and, Abida Begum for sharing scientific and non-scientific thoughts during this period. My sincere thanks are due to Mr. Ibrar Aslam for being supportive and helpful to me in difficult phases of my life and for making my stay memorable in Dortmund.

I would like to express my gratitude to the Higher Education Commission (HEC) of Pakistan for providing their financial support and the Deutscher Akademischer Austauschdienst (DAAD) for the management and distribution of funds.

In addition, I would like to thank my parents for their wise counsel and support. Finally, I could not have completed this thesis without the support of my wife and son who provided invigorating discussions as well as blissful distractions to rest my mind outside of my research.

Muhammad Waseem

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this thesis we are concerned with the development of new numerical method, based on finite element method in least-squares manner, to analyze the non-isothermal effects on incompressible Navier-Stokes equations for Newtonian and non-Newtonian fluid flows.

In classical physics, a fluid is one of the phases of matter that can be defined as a substance with deforming (flow) properties under some externally applied forces. Mankind encounters the fluid flows in everyday life, whether it is inside the human body as blood or outside as oil, water etc. Hence, the importance of fluids can not be neglected. Considering fluids importance, they are investigated extensively for the prosperity of human race. Fluid mechanics is one of the branches of applied sciences that study the fluid flow behaviors governed by partial differential equations (PDEs) which obey the conservation laws. The partial differential equations, known as Navier-Stokes equations, are the most developed mathematical framework used to represent the fluid flow phenomena. Interestingly, the exact analytical solution of the Navier-Stokes problem is remaining an open problem. The invention of computer was a real sigh of relief for researchers as it opened door to viable possibilities. As a consequence, the computational fluid dynamics (CFD) is appeared as a new field of computational mathematics in which complex partial differential equations are replaced by algebraic systems and approximated with suitable numerical techniques. The most popular numerical discretization methods employed in literature are namely finite volume method, finite element method and finite difference method. In this thesis, finite element method is employed in least-square manner to treat the incompressible non-isothermal Navier-Stokes equations for Newtonian and non-Newtonian fluid flows.

The Finite Element Methods (FEM) are the numerical arrangements for estimating the approximate solutions of mathematical phenomena that are normally constructed in order to represent some physical reality with highest precision. The finite element method is truly a quasi-projection scheme. Therefore, the variational principle and the

closed subspace are essential requirements for their construction. In this context, this method is a combination of the variational principle and the closed finite-dimensional subspace. The variational principle is specifically given in terms of a weak form of the considered partial differential equations. Hence, the quasi-projections of the exact weak solutions onto the closed subspace are characterized as approximate solutions. The favorable outcome of the finite element method relies upon wisely culling approximation spaces determined by piecewise polynomials constructed on elementary geometrical structures.

Although the approximate spaces (i.e. finite element spaces) and variational principles are equally significant in amelioration of finite element methods, but the fundamental properties of these methods are always determined by variational principles and lead to distinct kinds of quasi-projections. Instead of naturally occurring quasi-projections, the primary characteristic of the quasi-projection aims towards the search of suitable externally defined variational principles in the finite element method. This leads to a pair of well-known existing schemes to get more suitable quasi-projections. The first type of the finite element methods is based on the modification of naturally occurring variational principles. The second type of method is based on the exploration of artificial energy functional which is externally defined. The main aim of our work is to focus particularly on second type of methods called least-squares finite element methods (LSFEM).

1.1 Methods based on variational principles

The finite element method can be divided into three major variational settings. The Rayleigh-Ritz method, Galerkin method, and the least-squares method. The Rayleigh-Ritz method works on the principle to find minimizer of the total potential energy functional, i.e. the global minimization of an unconstrained positive functional. It provides the most suitable setting for a finite element method, i.e. the finite element solution defined in this case is a true projection with a suitable inner product in some Hilbert space. Remarkably, the true inner product projections are acceptable for the vast range of approximating spaces. The Rayleigh-Ritz method is proved to be extremely effective in the exploration of heat conduction problems for elliptic diffusion-type equations and mainly in the field of solid mechanics. The Rayleigh-Ritz principle induced finite element methods with many invaluable features that prompted their incredible achievement and prominence. The Rayleigh-Ritz variational procedure always results in a sparse, symmetric and positive definite[1] (SPD) system. Also, the

[1]The matrix $\mathbf{K} \in \mathbb{R}^{nxn}$ is symmetric, if $\mathbf{K} = \mathbf{K^T}$.
and $\mathbf{K}$ is a positive definite matrix, if for vector $u \in \mathbb{R}^n$ following relation holds

$$u^T \mathbf{K} u > 0, \quad \text{for all } u \neq \mathbf{0}.$$

conditions on boundaries in general regions are easy to tackle and the freedom to use conforming finite element spaces seal the optimality and stability of solutions. Moreover, we can employ identical piece-wise polynomials for the approximation of all unknown variables. Due to these practical features, many commercially developed codes for the finite element method are based on the Rayleigh-Ritz principle and they are generally in practice for the analysis of problems arise in engineering.

The affiliation among finite element methods and Rayleigh-Ritz variational principles played a key role in the development of finite element methods. However, this fact was initially neglected by researchers and they pursued their research beyond the class of problems based on Rayleigh-Ritz settings which eventually resulted in violation of some basic approximation principles. But, with the fast theoretical as well as practical advancement of finite element methods, the understanding was soon established over the causes of primary complications. Later, two entirely distinct types of variational principles were developed for differential equations not associated with unconstrained minimization principles. The first type dealt with the constraint minimization problem in which the resulted quasi-projection is an outcome of the search carried out for a stationary point of an indefinite functional. This procedure led to a symmetric, indefinite algebraic system and required additional compatibility constraints to satisfy on the spaces, namely Ladyzhenskaya-Babuska-Brezzi (LBB) condition, for the stable and optimal approximations. The mixed-Galerkin method employed to handle these types of problems.

The second type of principle was based on very general settings that could be used to obtain variational form of partial differential equations. The idea is known as the Galerkin method in applied mathematics. The standard Galerkin method is a weighted residual method and quasi-projections are obtained by residual orthogonal procedures. The idea is applicable to almost any partial differential equation regardless if it is an optimization problem or not. In the event if such an optimization problem exists, at that point, Galerkin methods do recuperate the related optimality framework. On account of this comprehensiveness, Galerkin strategy has been a natural option for expanding finite element methods beyond differential equations issues related to minimization problems. Like a mixed-Galerkin method, the standard Galerkin method may also need to satisfy some stability constraints (inf-sup conditions) on the choices of the spaces and eventually lead to a non-symmetric algebraic system.

In both situations as discussed above, the corresponding finite element schemes work in considerably less beneficial settings in comparison to the settings adopted for the Rayleigh-Ritz method because the variational settings of these schemes do not lead to a true inner product projections. In first case, the mathematical formulation ends up with an indefinite algebraic system and the second situation is confronted by the

problem of solving the non-symmetric algebraic system.

1.2 Motivation behind LSFEM

There have been numerous endeavors for the development of finite element methods that share the maximum achievable mathematical properties of the Rayleigh-Ritz method for issues not associated with unconstrained variational principles. The least-squares finite element method have been perceived as another tool to solve elliptic partial differential equations. The approach is based on an artificially defined energy functional for partial differential equation by totally relinquishing the naturally existing variational principles. The term "least-squares" is used with the finite element methods because the artificially defined energy functional generally appears as a residual minimization problem in an appropriate Hilbert space. Like orthogonal minimization scheme, i.e. Galerkin method, the residual minimization method is applicable to almost any partial differential equation in physics but with distinct fundamental properties. The major difference among both approaches, unlike the Galerkin method, is that the least-squares finite element method has the potential to recover almost all the precious features of the Rayleigh-Ritz method. The least-squares finite element methods obtain few additional properties which make them more superior from other finite element methods. The boundary conditions can be evaluated inside the defined energy functional in a weak manner as residuals. Also, the least-squares functionals give an effectively calculable residual error indicator that can be utilized for adapted grids techniques.

In the early 70s, the straight forward implementation of the least-squares technique to the partial differential equations experienced severe issues that seriously constrained their allure. For example, in many situations C^1 smoothness or higher ordered discrete approximation spaces were required and prompted issues with higher condition number. Additionally, such techniques frequently requested higher solution regularity to get convergence and not easily preconditioned. Later, the remarkable idea of transforming partial differential equations to first order systems was a game-changer in the development of least-squares finite element methods. The transformations to the first order linear system guarantee to work in C^0 finite element spaces with low condition number. The least-squares finite element method is considered to be practical if the linear first order system is discretized to conforming subspaces at ease and is well-conditioned. Also, it is based on a simple structure for implementation. Initially, the standard and straightforward least-squares finite element method was defined in $\mathcal{L}^2$-norm which has the potential to recover all the characteristics of the Rayleigh-Ritz method. Later, the idea of norm equivalence of least-squares functionals was floated and raised as an essential requirement to recuperate all the Rayleigh-Ritz settings.

However, eventually it is substantiated that norm-equivalence is frequently in strife with practicality. For the development and implementation point of view, norm equivalence is often relinquished over practicality as it is an intransigent circumscription.

1.3 Contributions

A handsome portion of theoretical analysis and numerical research on least-squares finite element methods is associated with the theory of elliptic partial differential equations. That shows the potential and provides the solid grounds to explore it further in this area. Therefore, we consider elliptic systems for our study purpose. The major development in this field was the consequences of work from Agmon, Douglas and Nirenberg ([2], [3]). They build up the foundation for ADN-theory which was utilized in numerous research productions. Soon after, another idea was floated by Wendland [96] which applied complex function theory to produce several results in 2D spaces for elliptic systems. He demonstrated that the straightforward use of a least-squares principle with appropriate weights may prompt optimal convergence. Furthermore, Aziz et al. [6] presented more general ideas for elliptic systems by using ADN-theory for least-squares finite element methods. These studies are the basic foundation for the development of least-squares finite element methods for elliptic systems and many researchers have followed these ideas for analysis and applications of least-squares finite element methods. The comprehensive analysis on continuous and discrete least-squares principles was provided by Bochev and Gunzburger [15]. They highlighted the conflicting aspects between continuous least-squares principles (i.e. mathematically well-posed) and discrete least-squares principles (i.e. feasibility setting for implementation) and proposed some remedies as well.

The real essences of our study is to constitute an agreement among the mathematical analysis and algorithmic development for defining the least-squares finite element methods. Throughout, the straightforward $\mathcal{L}^2$-norm based functionals are employed for analysis of least-squares finite element methods. Furthermore, the functionals are balanced by different viscosity dependent weighted functions. These methods comfortably fulfill all the practicality issues and recover the Rayleigh-Ritz settings as well. The major advantage of practicing $\mathcal{L}^2$-norm based functionals is the assembly of the matrix framework, which can be cultivated in a standard way with the condition number of order $O\left(h^{-2}\right)$ of system. There are many forms of first order transformations of least-squares finite element methods used for study purposes in the literature. Some of the most commonly first order transformations use the Velocity Vorticity Pressure, Stress Velocity Pressure, Velocity Velocity-Gradient Pressure, that we name as (VVP), (SVP) and ($\mathrm{VV_GP}$) formulations.

In literature, the velocity vorticity pressure and velocity velocity-gradient pressure

are considered most often for the least-squares finite element analysis. The stress velocity pressure formulation is not extensively used for the analysis and this provide an opportunity to dig deep in that domain. The stress velocity pressure first order transformation is employed for least-squares finite element methods throughout our study. In the SVP formulation "stresses" are utilized as a new external variables and are beneficial for the problems where additional information over stresses is needed. We introduce a new first order transformation named stress velocity pressure temperature (SVPT) formulation to analyze the coupled system of equations. Moreover, the inflow-outflow fluid problems are engaged. We can divide our work into two major types of fluid problems, i.e. Newtonian and non-Newtonian fluid problems. In the case of Newtonian fluids, the steady incompressible Stokes and Navier-Stokes type equations are considered for investigations. The SVP transformation is rarely used for Stokes problem. Therefore, we are keen to implement Stokes problem for our code validations. The outcomes of linear problem are used to understand the behavior of nonlinear solvers for nonlinear problem.

The utilization of least-squares finite element methods for Navier-Stokes problem is considerably boosted due to some of its computational advantages over Galerkin method. Such that, the selection of approximation spaces is not liable to any inf-sup (i.e. Ladyzhenskaya-Babuska-Brezzi) condition as discussed earlier. The resulting matrix from system is always symmetric positive definite and robust iterative methods can be employed to solve huge and challenging systems. These leverages allow us to use the conjugate gradient method as our frontline linear Krylov solver and multigrid method as its preconditioner. The Gauss-Newton method and fixed point methods are used as nonlinear solvers. The linear and nonlinear solvers are discussed in detail later. We define stress based least-squares functional [77] which is balanced by weighted functions for Navier-Stokes problem. The main theme of our work is to study least-squares finite element method for different variations of weighted functions. Furthermore, we show that along with other factors, the viscosity dependent weighted functions are equally important for the better performance of the solvers. Both the lower as well as higher order finite elements are used for computational purpose. We notice that higher order finite elements are more effective to get optimal results.

The non-Newtonian fluids had caught the attention and used successively for exploration of least-squares finite element methods. In similar fashion, we investigate, improve and expand the recent work done by Nickaeen [75] on least-squares finite element methods for the power law model based non-Newtonian fluids. We restrict our study only to stress velocity pressure first order formulation and mainly focus on different variations of nonlinear weighted functions used to balance the artificial externally defined energy functionals. In the literature very few people have investigated these types of weights. The nonlinear weights are dependent exclusively on the viscosity

functions but not on the mesh size of treated geometries. Next, we extend our study to non-Newtonian fluids based on Cross law model which represent more complex fluid problems. We investigate Cross law fluid model of nonlinear viscosity functions for different types of dependent variables i.e. shear, pressure, temperature. The nonlinear viscosity, depends on shear rate and pressure, is employed to investigate the least-squares finite element method for Navier-Stokes equations. Number of nonlinear weighted functions are tested for numerical analysis. Later, the SVPT first order formulation is introduced for coupled system of equations. The non-isothermal effects of flow combined with Navier-Stokes equations are considered for the exploration. These types of problems were studied by Damanik [37] for the Galerkin method. We suggest a new least-squares energy functional and balance it with different nonlinear weighted functions. To the best of our knowledge, no one has so far used, the Cross law based, nonlinear weighted least-squares techniques for the solution of Navier-Stokes equations and for the coupled problem as well.

Throughout our analysis, we deal with a monolithic approach for the subsequent nonlinear and corresponding discrete linear frameworks. As a nonlinear solver, we hired the Gauss-Newton (GN) and fixed point (FP) methods while multigrid-preconditioned conjugate gradient is engaged as linear solver. The goal of our research is to build a methodology which is robust, highly accurate and handle complex problems in size and nonlinearities. Therefore, the analysis is carried out for a huge variations of parametric values for Newtonian and non-Newtonian fluids. The actual focus is to study the importance as well as the effects of linear and nonlinear weighted functions on the least-squares finite element methods.

Besides many advantages of least-squares finite element methods there are some controversies as well. One of the major challenges is the mass conservation in inflow-outflow problems. Therefore, to study such issues, the flow around cylinder (inflow-outflow) problems [91], [75] are under considerations. In general, the lack of local mass conservation appears in all the finite element methods but for least-squares finite element methods even the mass is not conserved globally. The reason for that is the continuity equation, which guarantee the divergence free velocity field, presented as another component of least-squares residual functional. Consequently, if some other component influence the least-squares functional then mass conservation can turn out to be very frail. For this reason, the higher order finite element spaces are used to overcome these issues. Also, the fluid flow is monitored on critical points for global mass conservation. The nonlinear weighted functions are used with different variations to develop least-squares finite element methods for highly efficient results and for better mass conservation.

1.4 Thesis outline

The objective of our study is to develop least-squares finite element method, which is balanced by linear and nonlinear weighted functions depending on viscosity, for the solutions of Newtonian and non-Newtonian fluid flow problems. The thesis is organized as follow.

In the chapter 2, the basic concepts of least-squares theory are emphasized and the basic notations used in this work is defined. The elementary framework and least-squares principles for continuous and discrete problems are explained. Some practicality issues of least-squares principles are discussed briefly. The straight forward application of least-squares principle on the second order PDEs leads to complications. Therefore, the first order formulations of different types are used to transform the problem into linear first order system. Some first order formulations for Stokes problem are presented. Finally, the linearization strategies for the nonlinear problem are studied in detail. In the chapter 3, the nonlinear and linear solvers are explained. The application of least-squares principle to discrete linear system of equations leads to symmetric positive definite coefficient matrix. This permit us to employe efficient Krylov solvers for SPD system. Therefore, we use a multigrid preconditioned conjugate gradient method for discrete problem. Furthermore, the preconditioned conjugate gradient method is employed also as a smoother for multigrid solver. The Gauss-Newton and fixed point method are used as nonlinear solvers. A brief introduction of finite elements is presented as well. The flow around cylinder problem is discussed with two mesh configurations.

In the chapter 4, the steady incompressible Navier-Stokes problem is modeled for linear viscosity i.e. the Newtonian fluid. The SVP first order formulation is employed to transform the problem to linear system of equations. A linear weighted LSFEM is proposed for Newtonian fluids. The MPCG solver is tested for various flow parameters of the flow around cylinder problem. The lower and higher order finite element discretizations are used for numerical investigations. The global mass conservation is computed as well. The iteration based comparison is performed for linear and nonlinear solvers. In the Chapter 5, the nonlinear weighted LSFEM is proposed for non-Newtonian fluids. The power law and Cross law fluids are investigated for vast range of parameters. The Cross law model is never been investigated for least-squares FEM and we are the first to initiate research by employing it. We use only higher order finite element discretization for computational purposes. The unbounded viscosities are investigated for power law viscosity model and the bounded viscosities for Cross law model. The shear thinning and shear thickening effects of the fluids are studied. The comparison for solvers is performed as well. In the Chapter 6, the complexity of the problem is increased. As the fluid considered is non-Newtonian with non-isothermal properties. The nonlinear weighted LSFEM is proposed for the coupled

system based on a new SVPT formulation. The SVPT is constructed by introducing two auxiliary unknown variables, stress and temperature gradient, in the problem's model. The Cross law viscosity model is employed for numerical simulations. The model depends on shear rate, pressure and temperature and it is tested for different combinations. The solvers perform very well and produce accurate results with grid independent behavior. The mass conservation is obtained even for very hard problems. In the Chapter 7, the research is concluded overall and a brief overview of future plan is presented.

Chapter 2

Least-squares Finite Element Method

This chapter consists of fundamental theory regarding least-squares finite element method for solving the system of elliptic linear partial differential equations. The relation between partial differential equations and optimization problems are discussed. This discourse establishes the framework for the least-squares principle as a general methodology that connects an optimization problem with a system of given partial differential equations.

2.1 Introduction

In this work, the SVP formulation and its extension SVPT formulation are employed for the numerical investigations of Newtonian and non-Newtonian fluids. The theoretical aspect of LSFEM are explored in upcoming sections. The basic notations and definitions, for better understanding of the least-squares theory, are addressed in the upcoming sections of this chapter. We present a concise synopsis of the fundamental theory emerged for the investigation of numerous partial differential equations. For an extensive analysis of theory regarding partial differential equations, one can consult the following references [1], [5], [51], [53], [68], [84], [90], [96] . The major contribution in theoretical analysis of LSFEM on elliptic PDEs is done in [2], [3]. Later, the complex function theory is used in [96] to develop least-squares methods for elliptic system of PDEs. The need of weighted functions for optimal results was emphasized for LSFEM. A more general theory for the LSFEM was given in [6] to established a priori estimates for elliptic systems to get optimal convergence rates. The comprehensive analysis on continuous and discrete least-squares principles were provided in [15]. They highlighted the conflicts between continuous least-squares principles (i.e. mathematically well-posed) and discrete least-squares principles and proposed some

remedies as well. The Stress based first order formulations are studied in [27], [30], [43], [55], [15], [17], [77].

2.2 Basic notation

Throughout this thesis an open bounded domain Ω is under consideration which lies in the real coordinate space $\mathbb{R}^n$ of n-dimensions, where $n = 1, 2, 3$. The boundary region $\hat{\Omega} = \partial\Omega$ of domain is considered to be sufficiently piecewise smooth. In the event that diverse boundary conditions need to be employed on various sections of the boundary of a given domain then these particular segments of the boundary are represented as Dirichlet boundary part $\hat{\Omega}_D$ and Neumann boundary part $\hat{\Omega}_N$ respectively. We also assume a pair of additional characteristic associated with boundary region $\hat{\Omega}$, as

$$\overline{\hat{\Omega}} = \overline{\hat{\Omega}_D \cup \hat{\Omega}_N} \quad \text{and} \quad \hat{\Omega}_D \cap \hat{\Omega}_N = \emptyset.$$

We define a multi-index representation of partial derivatives for functions of multi-variables as

$$\mathcal{D}^{\wp} = \left(\frac{\partial}{\partial x_1}\right)^{\wp_1} \left(\frac{\partial}{\partial x_2}\right)^{\wp_2} \left(\frac{\partial}{\partial x_3}\right)^{\wp_3} \cdots \left(\frac{\partial}{\partial x_l}\right)^{\wp_l}, \tag{2.1}$$

where $\wp = (\wp_1, \wp_2, \wp_3, ..., \wp_l) \in \mathbb{N}_l$. The Functions under consideration have continuous derivative to the order $\hat{r}$ and belong to the space $\mathcal{C}^{\hat{r}}$, such that

$$|\wp| = |\wp_1| + |\wp_2| + |\wp_3| + \cdots + |\wp_l| \leq \hat{r}. \tag{2.2}$$

Now, some differential operators namely gradient ∇, divergence $\nabla\cdot$, and Laplacian Δ are defined which are very helpful to show the system of partial differential equations to an operators structure.

Let $a = x\hat{\imath} + y\hat{\jmath} \in \mathbb{R}^2$ be a position vector with length $|a| = \sqrt{x^2 + y^2}$ and its scalar field given as $\mathfrak{s}(a) = \mathfrak{s}(x, y) \in \mathbb{R}$ and vector filed as $\mathfrak{v}(a) = u\hat{\imath} + v\hat{\jmath} \in \mathbb{R}^2$, where $\hat{\imath}, \hat{\jmath}$ are unit vectors. Then gradient is a vector valued operator given as

$$\nabla\mathfrak{s} = \frac{\partial\mathfrak{s}}{\partial x}\hat{\imath} + \frac{\partial\mathfrak{s}}{\partial y}\hat{\jmath}, \tag{2.3}$$

and formal vector representation can be shown as

$$\nabla = \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath}. \tag{2.4}$$

Similarly, divergence is a scalar valued operator provided as under

$$\begin{aligned} \nabla \cdot \mathfrak{v} &= \left(\frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath}\right) \cdot (u\hat{\imath} + v\hat{\jmath}) \\ &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \end{aligned} \tag{2.5}$$

$$= \quad div\ \mathfrak{v}$$

The divergence of a gradient operator produce Laplacian such that

$$div \nabla \mathfrak{s} = \nabla \cdot \nabla \mathfrak{s} = \nabla^2 \mathfrak{s} = \Delta \mathfrak{s}. \tag{2.6}$$

By using these definitions we can represent any partial differential equation in operator form. The modern theories for partial differential equations depend intensively on the Sobolev spaces ([1], [25]) which contain generalized or weak solutions. We provide a brief introduction of Hilbert spaces and their connection with different differential operators. The notations for inner products and norms are $\langle \cdot, \cdot \rangle_{\mathrm{X}}$ and $\| \cdot \|_{\mathrm{X}}$ on the Hilbert space X, respectively.

Let $1 \leq p < \infty$ and $k \geq 0$ be non-negative integers. The space $\mathcal{L}^p$ of p-th power integrable functions can be defined on the domain Ω as

$$\mathcal{L}^p(\Omega) := \left\{ \mathfrak{u} \mid \|\mathfrak{u}\|_{0,\Omega}^p < \infty \right\}, \tag{2.7}$$

and the norm

$$\|\mathfrak{u}\|_{0,\Omega} = \left(\int_\Omega |\mathfrak{u}|^p \, d\Omega \right)^{\frac{1}{p}}. \tag{2.8}$$

As a special case for $p = 2$, we can define the space $\mathcal{L}^2$ of all square integrable functions as

$$\mathcal{L}^2(\Omega) := \left\{ \mathfrak{u} \mid \|\mathfrak{u}\|_{0,\Omega}^2 < \infty \right\}, \tag{2.9}$$

with the inner product

$$\langle \mathfrak{u}, \mathfrak{v} \rangle_{0,\Omega} = \int_\Omega \mathfrak{u} \cdot \mathfrak{v} \ d\Omega, \tag{2.10}$$

and the induced $\mathcal{L}^2$-norm as

$$\|\mathfrak{u}\|_{0,\Omega} = \left(\langle \mathfrak{u}, \mathfrak{u} \rangle_{0,\Omega} \right)^{\frac{1}{2}} = \left(\int_\Omega |\mathfrak{u}|^2 \, d\Omega \right)^{\frac{1}{2}}. \tag{2.11}$$

These definitions lead to well known Sobolev spaces $\mathbb{W}_p^k$

$$\mathbb{W}_p^k(\Omega) := \{ \mathfrak{u} \in \mathcal{L}^p(\Omega) \mid \mathcal{D}^\wp \mathfrak{u} \in \mathcal{L}^p(\Omega) \quad \forall \ |\wp| \leq k \}, \tag{2.12}$$

with the norm

$$\|\mathfrak{u}\|_{k,\Omega} = \sum_{|\wp| \leq k} \left(\int_\Omega |\mathcal{D}^\wp \mathfrak{u}|^p \, d\Omega \right)^{\frac{1}{p}}. \tag{2.13}$$

where $\mathcal{D}^\wp$ represents the derivatives of $\wp$ order.

Now we discuss some special cases of Sobolev spaces for $p = 2$ and if the considered domain is sufficiently smooth and bounded, we get

$$\mathcal{H}^k(\Omega) = \mathbb{W}_2^k := \left\{\mathfrak{u} \in \mathcal{L}^2(\Omega) \mid \mathcal{D}^{\wp}\mathfrak{u} \in \mathcal{L}^2(\Omega) \quad \forall \ |\wp| \leq k\right\}, \tag{2.14}$$

with the norm

$$\|\mathfrak{u}\|_{k,\Omega} = \sum_{|\wp|\leq k} \left(\int_{\Omega} |\mathcal{D}^{\wp}\mathfrak{u}|^2 \, d\Omega \right)^{\frac{1}{2}}. \tag{2.15}$$

For $k = 0$, we get $\mathcal{H}^0 = \mathcal{L}^2$ the space of all square integrable function as defined in (2.9) . For $k = 1$, we have the space $\mathcal{H}^1$ of all square integrable functions whose first order derivatives are square integrable as well

$$\mathcal{H}^1(\Omega) := \left\{\mathfrak{u} \in \mathcal{L}^2(\Omega) \mid \|\mathfrak{u}\|_{1,\Omega}^2 < \infty\right\}, \tag{2.16}$$

and

$$\begin{aligned} \|\mathfrak{u}\|_{1,\Omega} &= \sum_{|\wp|\leq 1} \left(\int_{\Omega} |\mathcal{D}^{\wp}\mathfrak{u}|^2 \, d\Omega \right)^{\frac{1}{2}}, \\ &= \left(\int_{\Omega} |\mathfrak{u}|^2 \, d\Omega \right)^{\frac{1}{2}} + \left(\int_{\Omega} |\mathcal{D}\mathfrak{u}|^2 \, d\Omega \right)^{\frac{1}{2}}, \\ &= \|\mathfrak{u}\|_{0,\Omega} + \|\nabla\mathfrak{u}\|_{0,\Omega}. \end{aligned} \tag{2.17}$$

The spaces $\mathcal{L}^2$ and $\mathcal{H}^1$ are very significant in the development of least-squares finite element theory. Some important subspaces stem from these spaces are discussed. Let's assume a nonempty subset B of boundary $\hat{\Omega}$ of a open bounded domain Ω. We define the subspace of all square integrable with zero mean value as

$$\mathcal{L}_0^2 := \left\{ \mathfrak{u} \in \mathcal{L}^2 \mid \int_{\Omega} \mathfrak{u} \, d\Omega = 0 \right\}, \tag{2.18}$$

and the subspace of $\mathcal{H}^1(\Omega)$ with zero trace on B as

$$\mathcal{H}_B^1 := \left\{\mathfrak{u} \in \mathcal{H}^1 \mid \mathfrak{u} = 0 \text{ on } B\right\}, \tag{2.19}$$

similarly

$$\mathcal{H}_{\hat{\Omega}}^1 = \mathcal{H}_0^1 := \left\{\mathfrak{u} \in \mathcal{H}^1 \mid \mathfrak{u} = 0 \text{ on } \hat{\Omega}\right\}. \tag{2.20}$$

We also define the space $\mathcal{H}(\text{div}, \Omega)$ as

$$\mathcal{H}(\text{div}, \Omega)^n := \left\{\mathfrak{u} \in \left[\mathcal{L}^2(\Omega)\right]^n \mid \nabla \cdot \mathfrak{u} \in \mathcal{L}^2(\Omega)\right\}, \tag{2.21}$$

with the corresponding inner product and norm defined as under

$$\langle \mathfrak{u}, \mathfrak{v} \rangle_{div} = \langle \mathfrak{u}, \mathfrak{v} \rangle_{0,\Omega} + \langle \nabla \cdot \mathfrak{u}, \nabla \cdot \mathfrak{v} \rangle_{0,\Omega} , \tag{2.22}$$

and

$$\|\mathfrak{u}\|_{div} = \|\mathfrak{u}\|_{0,\Omega} + \|\nabla \cdot \mathfrak{u}\|_{0,\Omega} . \tag{2.23}$$

We further define the following Hilbert spaces

$$\mathcal{H}^1_{g,D}(\Omega) := \left\{ \mathfrak{q} \in \mathcal{H}^1(\Omega) \mid \mathfrak{q} = g_{\hat{\Omega}_D} \text{ on } \hat{\Omega}_D \right\}, \tag{2.24}$$

$$\mathcal{H}_{g,N}(\text{div}, \Omega) := \left\{ \mathfrak{v} \in \mathcal{H}(\text{div}, \Omega) \mid n \cdot \mathfrak{u} = g_{\hat{\Omega}_N} \text{ on } \hat{\Omega}_N \right\}, \tag{2.25}$$

which are subspaces of $\mathcal{H}^1(\Omega)$, $\mathcal{H}(\text{div}, \Omega)$.

2.3 Minimization problems and PDEs

In modeling many physical phenomena one can define suitable functionals, for which the physically significant solutions of the model are described as minimizers. Normally, these functionals illustrate the energy of the framework and the states of the framework are described by some minimization principle. Hence, a variational problem is promptly feasible for the possible choice of minimizers. The variational problem is always constituted on a Hilbert space. Therefore, the system of partial differential equations as well as the necessary boundary conditions are affiliated with the variational problem, can easily be restored by following the basic rule of integration by parts. To begin with variational formulation, an approximation technique for minimizers can be defined on a finite-dimensional discrete subspace of considered Hilbert space which contains the required solutions. This discrete form of variational formulation can be seen as necessary condition for the defined discrete energy functional by the virtue of Euler-Lagrange equation. The existence and uniqueness of approximate solutions are then the matters of discussion regarding continuous and its discrete variational problems.

In case of energy functional, the bilinear form[1] is coercive[2] and symmetric for the corresponding variational problem. Eventually the partial differential equation associated with such a problem is elliptic and self-adjoint. The existence and uniqueness

[1] A function $\mathbf{B} : \mathcal{X} \times \mathcal{X} \to \mathbf{S}$ is called a bilinear form if

$$\begin{aligned} \mathbf{B}(\alpha u + \beta v, w) &= \alpha \mathbf{B}(u, w) + \beta \mathbf{B}(v, w) , \\ \mathbf{B}(w, \alpha u + \beta v) &= \alpha \mathbf{B}(w, u) + \beta \mathbf{B}(w, v) , \end{aligned}$$

where $\mathcal{X}$ on $\left(\mathcal{L}^2 \text{ or } \mathcal{H}^1\right)$ is a vector space and $\mathbf{S}$ is field of scalars.

[2] The bilinear form $\mathbf{B}$ is coercive on $\mathcal{H}^1$, if for a constant $\alpha > 0$, the following relation hold

$$|\mathbf{B}(w, w)| \geq \alpha \|w\|_1^2 , \quad \forall w \in \mathcal{H}^1.$$

of solution is affirmed for both continuous and discrete problems. Due to the coercivity and the symmetric properties of bilinear form, one can define equivalent energy norm and inner product on approximation space. Consequently, the discrete minimizer can be distinguished as a true projection with an equivalent inner product onto the finite-dimensional space. We illustrate the above discussed scenario by the help of an example.

Let us consider the quadratic energy functional

$$\mathcal{E}\left(\mathfrak{u};\mathfrak{f}\right)=\int\limits_{\Omega}\left|\nabla\mathfrak{u}\right|^{2}\ d\Omega-2\int\limits_{\Omega}\mathfrak{f}\mathfrak{u}\ d\Omega, \tag{2.26}$$

the natural choice of space for the minimization of this particular functional is $\mathcal{H}_0^1$ and the minimization principle associated with the functional is given as

$$\mathfrak{u}=\arg\min_{u\in\mathcal{H}_0^1}\mathcal{E}\left(\mathfrak{u};\mathfrak{f}\right). \tag{2.27}$$

Thus, the necessary condition for the minimization of variational problem is as: Find $\mathfrak{u}\in\mathcal{H}_0^1$, such that

$$\lim_{t\to0}\frac{d}{dt}\mathcal{E}\left(\mathfrak{u}+t\mathfrak{v}\right)=0,\qquad\qquad\mathfrak{v}\in\mathcal{H}_0^1, \tag{2.28}$$

then

$$2\left[\int\limits_{\Omega}\nabla\mathfrak{u}\cdot\nabla\mathfrak{v}\ d\Omega-\int\limits_{\Omega}\mathfrak{f}\mathfrak{v}\ d\Omega\right]=0, \tag{2.29}$$

$$\int\limits_{\Omega}\nabla\mathfrak{u}\cdot\nabla\mathfrak{v}\ d\Omega=\int\limits_{\Omega}\mathfrak{f}\mathfrak{v}\ d\Omega, \tag{2.30}$$

the variational equation can be given as:

$$\begin{cases}\text{Find }\mathfrak{u}\in\mathcal{H}_0^1\text{ such that,}\\ \mathfrak{B}\left(\mathfrak{u};\mathfrak{v}\right)=\mathfrak{L}\left(\mathfrak{v}\right),\quad\forall\ \mathfrak{v}\in\mathcal{H}_0^1,\end{cases} \tag{2.31}$$

where $\mathfrak{B}$ and $\mathfrak{L}$ are bilinear and linear forms respectively.

The relation between minimization problem (2.27) and partial differential equation is obtain by using integration by parts rule in equation (2.31). We get, assuming the required regularity of the solution,

$$\begin{aligned}\int\limits_{\Omega}\left(\nabla\mathfrak{u}\cdot\nabla\mathfrak{v}-\mathfrak{f}\mathfrak{v}\right)\ d\Omega &= 0,\\ -\int\limits_{\Omega}\mathfrak{v}\left(\Delta\mathfrak{u}+\mathfrak{f}\right)\ d\Omega &= 0,\end{aligned} \tag{2.32}$$

since $\mathfrak{v}$ is arbitrary, it follows that every sufficiently smooth minimizer of functional $\mathcal{E}(\cdot;\mathfrak{f})$ is a solution of well known Poisson problem.

$$\begin{cases} -\Delta\mathfrak{u} = \mathfrak{f}, \text{ in } \Omega, \\ \quad \mathfrak{u} = 0, \text{ on } \hat{\Omega}. \end{cases} \tag{2.33}$$

The bilinear form in (2.31) is continuous: Thus for some constant $\mathfrak{C}_1 > 0$, the following inequality holds

$$|\mathfrak{B}(\mathfrak{u};\mathfrak{v})| \leq \mathfrak{C}_1 \|\mathfrak{u}\|_1 \|\mathfrak{v}\|_1 \qquad \forall\, \mathfrak{u},\mathfrak{v} \in \mathcal{H}_0^1, \tag{2.34}$$

and coercive, for constant $\mathfrak{C}_2 > 0$, the following condition holds

$$|\mathfrak{B}(\mathfrak{u};\mathfrak{u})| \geq \mathfrak{C}_2 \|\mathfrak{u}\|_1^2 \qquad \forall\, \mathfrak{u} \in \mathcal{H}_0^1. \tag{2.35}$$

Similarly, the existence and uniqueness of minimizer of (2.26) follows from Lax-Milgram theorem [78].

The Poisson problem in (2.33) requires two continuous derivatives of $\mathfrak{u}$. Therefore, it seems logical for $\mathfrak{u} \in \mathcal{H}_0^1$ in (2.27) that vanish on boundary $\hat{\Omega}$. This reveals an attractive characteristic of unconstraint minimization problem that each classical solution of Poisson problem (2.33) is also a minimizer of the minimization problem (2.27). We introduce some non-classical solutions which are known as weak solutions. An enormous number of physical processes is represented by minimization principles.

One of the major advantage of energy minimization principle is that an equivalent energy norm on $\mathcal{H}_0^1$ can be defined by the functional such as

$$\mathcal{E}(\mathfrak{v};0) = \int_\Omega |\nabla\mathfrak{v}|^2 \; d\Omega = \|\mathfrak{v}\|_1^2 . \tag{2.36}$$

Subsequently, the bilinear form define an equivalent inner product on $\mathcal{H}_0^1$, such that

$$\mathfrak{B}(\mathfrak{u};\mathfrak{v}) = \langle \mathfrak{u};\mathfrak{v} \rangle . \tag{2.37}$$

The computational interests for this setting can be explored by finite element method. Consider, a discrete finite dimensional space $\mathcal{X}^h$ in $\mathcal{H}_0^1$ and let $\mathfrak{u}^h$ be a finite approximate solution of weak solutions $\mathfrak{u}$. The approximation principle, that is a restriction of problem (2.31) to $\mathcal{X}^h$, is given as

$$\begin{cases} \text{Find } \mathfrak{u}^h \in \mathcal{X}^h \text{ such that,} \\ \mathfrak{B}\left(\mathfrak{u}^h;\mathfrak{v}^h\right) = \mathfrak{L}\left(\mathfrak{v}^h\right), \quad \forall\, \mathfrak{v}^h \in \mathcal{X}^h. \end{cases} \tag{2.38}$$

We note that $\mathcal{X}_h \subset \mathcal{H}_0^1$ for all h, (2.34) and (2.35) hold for all functional in $\mathcal{H}_0^1$. This implies that $\mathfrak{u}^h$ is an orthogonal projection of $\mathfrak{u}$ onto $\mathcal{X}^h$ w.r.t. inner product $\langle \cdot,\cdot \rangle$.

Since, the exact solution satisfy the discrete problem and (2.38) it follows that $\mathfrak{u}^h$ approximates the energy norm of error such that

$$\left|\left|\left|\mathfrak{u}-\mathfrak{u}^h\right|\right|\right| = \inf_{v^h\in\mathcal{X}_h}\left|\left|\left|\mathfrak{u}-\mathfrak{v}^h\right|\right|\right|. \tag{2.39}$$

Together with coercivity and continuity this bound provides an error estimate in the norm of $\mathcal{H}_0^1$, such as

$$\left\|\mathfrak{u}-\mathfrak{u}^h\right\|_1 \leq \mathfrak{C}\inf_{v^h\in\mathcal{X}_h}\left\|\mathfrak{u}-\mathfrak{v}^h\right\|_1. \tag{2.40}$$

The linear algebraic system associated to this problem for given basis $\{\mathfrak{u}_i\}_{i=1}^N$ has the following structure

$$\mathfrak{A}\mathfrak{u}^h = \mathfrak{F}, \tag{2.41}$$

where

$$\begin{aligned}\mathfrak{A}_{ij} &= \mathfrak{B}\left(\mathfrak{u}_i;\mathfrak{u}_j\right),\\ \mathfrak{F}_i &= \mathfrak{L}\left(\mathfrak{u}_i\right).\end{aligned}$$

It comply from (2.31) , (2.34) and (2.35) that system matrix in (2.41) is symmetric and positive definite which is a key factor for the construction of a powerful preconditioner for system (2.41). These type of settings refer to the classical Rayleigh-Ritz method [38].

Generally, the mathematical phenomena can be presented straight away by a system of partial differential equations or simply by an equation. Then, by following the above prescribed procedure in a reverse manner, one can affiliate a functional including some mandatory conditions, i.e. a minimization principle, with the provided system of equations. However, It is not necessary that this procedure prompt an optimization problem with a unique solution.

For more clarification, we consider Stokes problem as follows

$$\left.\begin{aligned}-\Delta\mathfrak{u}+\nabla\mathfrak{p} &= \mathfrak{f}\\ \nabla\cdot\mathfrak{u} &= 0\end{aligned}\right\}\ \text{in } \Omega, \qquad \left.\begin{aligned}\mathfrak{u} &= 0\end{aligned}\right\}\ \text{on } \hat{\Omega}. \tag{2.42}$$

This problem gives rise to an important class of mixed variational problem. Let us consider the above problem for a couple of Hilbert spaces $\mathcal{H}_0^1$ and $\mathcal{L}_0^2$ with continuous bilinear forms $\mathfrak{a}(\cdot,\cdot)$ on $\mathcal{H}_0^1\times\mathcal{H}_0^1$ and $\mathfrak{b}(\cdot,\cdot)$ on $\mathcal{H}_0^1\times\mathcal{L}_0^2$, respectively. The problem is define as

$$\begin{cases}\text{Find } (\mathfrak{u},\mathfrak{p})\in\mathcal{H}_0^1\times\mathcal{L}_0^2 \text{ such that}\\ \mathfrak{a}(\mathfrak{u},\mathfrak{v})+\mathfrak{b}(\mathfrak{v},\mathfrak{p}) = f(\mathfrak{v}) \quad \forall\ \mathfrak{v}\in\mathcal{H}_0^1,\\ \mathfrak{b}(\mathfrak{u},\mathfrak{q}) = g(\mathfrak{q}) \quad \forall\ \mathfrak{q}\in\mathcal{L}_0^2,\end{cases} \tag{2.43}$$

where

$$\mathfrak{a}\left(\mathfrak{u},\mathfrak{v}\right)=\int_{\Omega}\nabla\mathfrak{u}\cdot\nabla\mathfrak{v}\;d\Omega, \tag{2.44}$$

and

$$\mathfrak{b}\left(\mathfrak{u},\mathfrak{q}\right)=-\int_{\Omega}\mathfrak{q}\left(\nabla\cdot\mathfrak{u}\right)\;d\Omega, \tag{2.45}$$

along with linear functional f and g on $\mathcal{H}_0^1$ and $\mathcal{L}_0^2$. The formulation in (2.43) is also well-known as saddle point problem, as opposed to a minimization problem. It is clear from (2.44) that the bilinear form $\mathfrak{a}\left(\mathfrak{u},\mathfrak{v}\right)$ is symmetric and based on quadratic functional. The bilinear form in (2.45) represents constraint's weak form. For the well-posedness of problem (2.43), the following conditions [8], [26], [51] must be satisfied,

$$\sup_{u\in\mathcal{Z}}\frac{\mathfrak{a}\left(\mathfrak{u},\mathfrak{v}\right)}{\left\|\mathfrak{u}\right\|_1}\geq\beta_{_a}\left\|\mathfrak{v}\right\|_1\qquad\forall\;\mathfrak{u}\in\mathcal{Z}, \tag{2.46}$$

Since $\mathfrak{a}\left(\cdot,\cdot\right)$ is symmetric

$$\mathfrak{a}\left(\mathfrak{v},\mathfrak{v}\right)\geq\alpha\left\|\mathfrak{v}\right\|^2,$$

and

$$\sup_{v\in\mathcal{H}_0^1}\frac{\mathfrak{b}\left(\mathfrak{u},\mathfrak{q}\right)}{\left\|\mathfrak{v}\right\|_1}\geq\beta_{_b}\left\|\mathfrak{q}\right\|_0\qquad\forall\;\mathfrak{q}\in\mathcal{L}_0^2, \tag{2.47}$$

where the space of divergence-free functions is given by

$$\mathcal{Z}=\left\{\mathfrak{z}\in\mathcal{H}_0^1\mid\mathfrak{b}\left(\mathfrak{z},\mathfrak{q}\right)=0\quad\forall\;\mathfrak{q}\in\mathcal{L}_0^2\right\}. \tag{2.48}$$

Now we consider the problem on discrete level, let $\mathcal{X}^h$ and $\mathcal{Y}^h$ be two finite element subspaces in the function spaces $\mathcal{H}_0^1$ and $\mathcal{L}_0^2$, respectively. We then restrict problem (2.43) to these spaces to obtain the discrete problem,

$$\begin{cases}\text{Find }\;\left(\mathfrak{u}^h,\mathfrak{p}^h\right)\in\mathcal{X}^h\times\mathcal{Y}^h\;\text{ such that}\\ \mathfrak{a}\left(\mathfrak{u}^h,\mathfrak{v}^h\right)+\mathfrak{b}\left(\mathfrak{v}^h,\mathfrak{p}^h\right)=f\left(\mathfrak{v}^h\right)\quad\forall\;\mathfrak{v}^h\in\mathcal{X}^h,\\ \qquad\qquad\mathfrak{b}\left(\mathfrak{u}^h,\mathfrak{q}^h\right)=g\left(\mathfrak{q}^h\right)\quad\forall\;\mathfrak{q}^h\in\mathcal{Y}^h.\end{cases} \tag{2.49}$$

The linear algebraic block representation of the system matrix is as following

$$\bar{\mathfrak{A}}\mathbf{u}^h=\begin{bmatrix}A_{_{uu}}&B_{_{up}}\\B_{_{up}}^{^T}&0\end{bmatrix}\begin{bmatrix}\mathfrak{u}^h\\\mathfrak{p}^h\end{bmatrix}=\begin{bmatrix}f\\g\end{bmatrix}=\bar{\mathfrak{F}}. \tag{2.50}$$

The matrix $\bar{\mathfrak{A}}$ in (2.50) is symmetric and indefinite and it makes the system quite hard to solve in comparison to matrix in the system (2.41). Since the problem (2.49) is discrete saddle point problem and solvable as long as matrix (2.50) is nonsingular, therefore solution of such problem depends on certain conditions. Above all, the well-posedness of (2.49) depends on the so-called Ladyzhenskaya-Babuska-Brezzi (LBB) or inf-sup condition on discrete spaces $\mathcal{X}^h$ and $\mathcal{Y}^h$, see [7], [26], [69]. The following inequality holds for $\mathfrak{b}\left(\cdot,\cdot\right)$ and $\beta_{_b}^h>0$, such that

$$\inf_{q^h \in \mathcal{Y}^h} \sup_{v^h \in \mathcal{X}^h} \frac{\mathfrak{b}\left(\mathfrak{v}^h, \mathfrak{q}^h\right)}{\|\mathfrak{v}^h\|_1 \|\mathfrak{q}^h\|_0} \geq \beta_b^h, \tag{2.51}$$

along with the coercive bilinear form $\mathfrak{a}(\cdot,\cdot)$ on $\mathcal{Z}^h \times \mathcal{Z}^h$, where $\mathcal{Z}^h \subset \mathcal{X}^h$ is given in discrete form as

$$\mathcal{Z}^h = \left\{ \mathfrak{z}^h \in \mathcal{X}^h \mid \mathfrak{b}\left(\mathfrak{z}^h, \mathfrak{q}^h\right) = 0 \quad \forall\ \mathfrak{q}^h \in \mathcal{Y}^h \right\}. \tag{2.52}$$

This is also another contradiction to Rayleigh-Ritz settings. One of the major drawbacks of inf-sup condition is the exclusion of equal order finite element spaces for velocity and pressure.

There exists another methodology that can be employed to develop variational settings for partial differential equation. Such type of technique is recognized as Galerkin method or weighted residual method and is truly based on principle of residual orthogonalization. This technique can be applied to any partial differential equation regardless of whether a minimization principle is associated with or not. But in case of existence of such an optimization problem, the method has the potential to redeem the optimality system. This makes the method a popular choice for broadening finite elements theory beyond minimization principles connected to partial differential equations. For instance, if we consider the problem (2.42), the standard process in Galerkin technique is to multiply the first and second equations with the weighted (test) functions $\mathfrak{v}$,vanish on boundary $\hat{\Omega}$, and $\mathfrak{q}$, respectively and integrate over the domain Ω. The following procedure leads to same variational problem as we got in (2.43). This shows that, if the partial differential equation is related to an optimization principle then Galerkin approach leads to the optimality system. Beside all these qualities there are still some uncertainties associated with the Galerkin approach such as

- For a given problem, the process can lead to more than one variational problem.
- In general, the associated bilinear form is not coercive and/or symmetric.
- The bilinear form does not define an equivalent inner product on the finite element space.
- The bilinear form must satisfy the inf-sup condition for the well-posedness of the problem.
- The analysis and the implementation of method is more difficult than the energy minimization principle based method.

However, the versatility of the method can be achieved, but at some cost. it is clear from the above discussion that, by following standard Galerkin procedure one can easily construct the variational problem for almost any type of partial differential

equation. The method did not demand any prior information about the naturally occurring minimization principle.

2.4 Least-squares principle for PDEs

The Rayleigh-Ritz settings have so far provided ideal ground for the exploration of finite element method. The key features of Rayleigh-Ritz setting can be summarized as follow

- The considered differential equation can be associated to an equivalent naturally occurring variational problem.
- The bilinear form is coercive, symmetric and positive definite.
- An equivalent inner product on the considered finite element space can be defined.
- No extra condition, i.e. inf-sup condition, required for the existence and uniqueness of solution of the problem.
- A single finite element space can be used for all unknowns for discrete minimization problem.

It is shown in previous section that if the variational principle does not satisfy few/all of Rayleigh-Ritz properties then it must suffer analytical and computational difficulties onwards. Therefore, due to the above-mentioned properties of Rayleigh-Ritz method, there has been made numerous efforts to achieve these qualities for the problems where it does not appear naturally. Now, an important question arises that is it feasible for a given problem, i.e. system of partial differential equations, to define an unconstrained minimization principle such that the finite element method can be benefited from the properties acquired by Rayleigh-Ritz method.

In literature efforts have been progressed in many ways. One of them is established on the use of artificial energy functionals which can be defined externally. This technique is known as least-squares finite element method which has the ability to capture most of the Rayliegh-Ritz settings. The least-squares finite element method is truly established on the concept of residual minimization. It can be applied nearly to any partial differential equation because of the concept of residual minimization is as broad as the residual orthogonalization used in Galerkin method. But, in contrast to residual orthogonalization, one can associate the inner product projections with residual minimization regardless of the given problem is an optimization problem or not. Due to such a valuable potential of least-squares approach, a suitable energy functional, as a square sum of all possible residual equations, can be defined and minimized under some appropriate norm. The subsequent artificial energy functional, as a general

rule, is not practically meaningful, however it provides a tool to converts any partial differential equation problem into a corresponding minimization problem.

2.5 Elementary framework

The main focus in this section is to present the basic concept of least-squares principle. The fundamental data or evidence are taken mainly from text books referenced [15] and [63]. In the upcoming discussions, we consider the system of linear partial differential equations represented by a differential operator $\mathcal{S}_{_L}$ in a domain $\Omega \subset \mathbb{R}^2$ with sufficiently smooth boundary and appropriate boundary conditions denoted by operator $\mathcal{T}_{_B}$ on the boundary region $\hat{\Omega}$. We consider the abstract problem form as

$$\begin{cases} \mathcal{S}_{_L}\mathfrak{u} = \mathfrak{f} & \text{in } \Omega, \\ \mathcal{T}_{_B}\mathfrak{u} = \mathfrak{g} & \text{on } \hat{\Omega}, \end{cases} \tag{2.53}$$

where $\mathfrak{f}$ and $\mathfrak{g}$ are given vector-valued functions in and on the boundary, respectively. The boundary value problem (2.53) also oblige a couple of hypothesis. First, if the Hilbert spaces $\mathcal{X}(\Omega)$ and $\mathcal{Y}(\Omega) \times \mathcal{Y}(\hat{\Omega})$ exist such that $(\mathcal{S}_{_L}, \mathcal{T}_{_B})$ is a homeomorphism, i.e. the mapping $\mathfrak{u} \longrightarrow (\mathcal{S}_{_L}\mathfrak{u}, \mathcal{T}_{_B}\mathfrak{u})$ is a homeomorphism on $\mathcal{X}(\Omega) \longrightarrow \mathcal{Y}(\Omega) \times \mathcal{Y}(\hat{\Omega})$. Moreover, the operator $(\mathcal{S}_{_L}, \mathcal{T}_{_B})$ has a closed range and the kernel as well as the co-range are finite dimensional. These conditions provide sufficient grounds to consider wide range of partial differential equations problems. The above assumptions ensure the well-posedness of boundary value problem (2.53). The fundamental relation among solution and the boundary value problem can be described by inequality as follows,

$$\mathfrak{C}_1 \left\|\mathfrak{u}\right\|_{\mathcal{X}(\Omega)} \leq \left\|\mathcal{S}_{_L}\mathfrak{u}\right\|_{\mathcal{Y}(\Omega)} + \left\|\mathcal{T}_{_B}\mathfrak{u}\right\|_{\mathcal{Y}(\hat{\Omega})} \leq \mathfrak{C}_2 \left\|\mathfrak{u}\right\|_{\mathcal{X}(\Omega)}, \tag{2.54}$$

where $\mathfrak{C}_1$and $\mathfrak{C}_2$ are two positive constants and $\|\cdot\|$ is some norm defined on respective Hilbert spaces. One can also elaborate inequality (2.54) as it illustrates propitious stability between the residual energy and the solution energy.

Remark 1 *It is important to note that there may be several combinations of data and solution spaces present for any given boundary value problem for which the problem is well-posed and, in particular, the inequality* (2.54) *holds for such problem.*

Now, we discuss some fundamental principles, namely continuous and discrete, that are experienced in the early developments of least-squares theory.

2.6 Continuous principles for LSFEM

In this section, we focus mainly on continuous least-squares principles. A continuous least-squares principle can be explained as an ideal arrangement between the solution

norm and artificially defined energy functional, i.e. the residual, is mathematically true. For a given problem (2.53) , a norm equivalent functional can be associated with partial differential equations and it can also be shown that this function is achieved by a priori estimates for the PDE problem, see [15].

Let us consider the Hilbert spaces $\mathcal{Y}_{\Omega}$ and $\mathcal{Y}_{\hat{\Omega}}$ as data spaces that give norms for residual energy while the solution space $\mathcal{X}_{\Omega}$ provides the minimizer for energy functional and serves as trial space. Now the quadratic artificial energy functional for problem (2.53) can be defined as

$$\mathcal{E}\left(\mathfrak{u};\mathfrak{f},\mathfrak{g}\right)=\frac{1}{2}\left(\left\|\mathcal{S}_{L}\mathfrak{u}-\mathfrak{f}\right\|_{\mathcal{Y}_{\Omega}}^{2}+\left\|\mathcal{T}_{B}\mathfrak{u}-\mathfrak{g}\right\|_{\mathcal{Y}_{\hat{\Omega}}}^{2}\right), \tag{2.55}$$

and the continuous least-squares principle for the problem (2.55) is given by

$$\begin{cases} \text{Find } \mathfrak{u}\in\mathcal{X}_{\Omega} \text{ such that} \\ \mathcal{E}\left(\mathfrak{u};\mathfrak{f},\mathfrak{g}\right)\leq\mathcal{E}\left(\mathfrak{v};\mathfrak{f},\mathfrak{g}\right) \quad \forall\ \mathfrak{v}\in\mathcal{X}_{\Omega}. \end{cases} \tag{2.56}$$

We can represent energy functional (2.55) as $\mathcal{E}\left(\mathfrak{u}\right)$, if the data functions $\mathfrak{f}$ and $\mathfrak{g}$ are identical zeros.

Theorem 2 ([15]) *Assume that the Hilbert spaces $\mathcal{X}_{\Omega}$ and $\mathcal{Y}_{\Omega}\times\mathcal{Y}_{\hat{\Omega}}$ exist such that the operator $(\mathcal{S}_{L},\mathcal{T}_{B})$ is a homeomorphism and is of Fredholm type. Then,*

I. the functional (2.55) *is norm-equivalent in the sense that*

$$\frac{1}{4}\mathfrak{C}_{2}^{2}\left\|\mathfrak{u}\right\|_{\mathcal{X}_{\Omega}}^{2}\leq\mathcal{E}\left(\mathfrak{u}\right)\leq\frac{1}{2}\mathfrak{C}_{1}^{2}\left\|\mathfrak{u}\right\|_{\mathcal{X}_{\Omega}}^{2} \qquad \forall\ \mathfrak{u}\in\mathcal{X}_{\Omega}, \tag{2.57}$$

II. there exists a unique minimizer $\mathfrak{u}\in\mathcal{X}_{\Omega}$ of (2.55) moreover, the unique minimizer $\mathfrak{u}$ depends continuously on the data, i.e. $\mathfrak{u}$ satisfies

$$\left\|\mathfrak{u}\right\|_{\mathcal{X}_{\Omega}}\leq\mathfrak{C}\left(\left\|\mathfrak{f}\right\|_{\mathcal{Y}_{\Omega}}+\left\|\mathfrak{g}\right\|_{\mathcal{Y}_{\hat{\Omega}}}\right), \tag{2.58}$$

where $\mathfrak{C}$ is a constant whose value is independent of $\mathfrak{f}$, $\mathfrak{g}$ and $\mathfrak{u}$.

III. $\mathfrak{u}$ is the unique minimizer of (2.55) *if and only if $\mathfrak{u}$ is the unique solution of problem* (2.53) .

The above theorem describes the well-posedness of continuous least-squares principle (2.56) and the unique minimizer of problem (2.55) corresponds to the unique solution of boundary value problem (2.53) . By using standard calculus of variation, one can show that the minimizers $\mathfrak{u}\in\mathcal{X}_{\Omega}$ of (2.55) must satisfy the Euler-Lagrange equation such as

$$\mathcal{D}\mathcal{E}\left(\mathfrak{u}\right)=\lim_{t\to0}\frac{d\mathcal{E}\left(\mathfrak{u}+t\mathfrak{v}\right)}{dt}=0, \qquad \mathfrak{v}\in\mathcal{X}_{\Omega}, \tag{2.59}$$

By using basic maths, it can be determined that an equivalent variational problem can be developed such as

$$\begin{cases} \text{Find } \mathfrak{u} \in \mathcal{X}_{\Omega} \text{ such that,} \\ \mathfrak{B}(\mathfrak{u};\mathfrak{v}) = \mathfrak{L}(\mathfrak{v}), \quad \forall\ \mathfrak{v} \in \mathcal{X}_{\Omega}, \end{cases} \tag{2.60}$$

where, the bilinear form $\mathfrak{B}(\cdot;\cdot)$ is given as

$$\mathfrak{B}(\mathfrak{u};\mathfrak{v}) = \langle \mathcal{S}_{L}\mathfrak{u}, \mathcal{S}_{L}\mathfrak{v}\rangle_{\mathcal{Y}_{\Omega}} + \langle \mathcal{T}_{B}\mathfrak{u}, \mathcal{T}_{B}\mathfrak{v}\rangle_{\mathcal{Y}_{\hat{\Omega}}}, \tag{2.61}$$

and the linear form $\mathfrak{L}(\cdot)$ is provided as

$$\mathfrak{L}(\mathfrak{v}) = \langle \mathfrak{f}, \mathcal{S}_{L}\mathfrak{v}\rangle_{\mathcal{Y}_{\Omega}} + \langle \mathfrak{g}, \mathcal{T}_{B}\mathfrak{v}\rangle_{\mathcal{Y}_{\hat{\Omega}}}. \tag{2.62}$$

Remark 3 *From the above discussion, a mathematically well-posed structure is built in a way that the boundary value problem* (2.53)*, i.e. linear PDEs problem, can be replaced by an equivalent unconstraint minimization problem which is well-posed as well. The construction process of such minimization problems is completely characterized as the functional space* $\mathcal{X}_{\Omega}$ *and the norm equivalent functional* $\mathcal{E}(\cdot)$ *being minimized over same space* $\mathcal{X}_{\Omega}$*, i.e. the pair* $\{\mathcal{X}_{\Omega}, \mathcal{E}\}$*. The set of all such pairs that can be associated with a given linear PDE problem constitutes the class of its continuous least-squares principles (CLSPs).*

2.6.1 Treatment for various boundary conditions

The boundary conditions can be served in a couple of distinct manners, see [89] for further details.

i) To minimize the energy functional (2.55) straight away as the boundary residual is a part of that functional.

ii) Implementation of boundary condition in such a way that the contribution of boundary residual $\|\mathcal{T}_{B}\mathfrak{u} - \mathfrak{g}\|$, in the problem (2.53), becomes zero, by restriction of solution space $\mathcal{X}_{\Omega}$. Let us define a functional space $\tilde{\mathcal{X}}_{\Omega}$, as a restriction of solution space $\mathcal{X}_{\Omega}$, such that

$$\tilde{\mathcal{X}}_{\Omega} = \left\{ \mathfrak{u} \in \mathcal{X}_{\Omega} \mid \mathcal{T}_{B}\mathfrak{u}(z) = \mathfrak{g}(z) \quad \forall\ z \in \hat{\Omega} \right\},$$

hence, the residual $\|\mathcal{T}_{B}\mathfrak{u} - \mathfrak{g}\|$ in the system (2.53) will naturally vanish for all those functions which belong to space $\tilde{\mathcal{X}}_{\Omega}$.

We use strategy (ii) to deal with the boundary conditions and consider even natural boundary condition as essential boundary condition by restriction of solution spaces (See for details [15], Chapter 12).

2.6.2 An operator form of least-squares principle

In this thesis, more specifically, we studied the least-squares method based on the minimization of residuals (2.55) defined in $\mathcal{L}^2$-norm. Suppose that in problem (2.53), $\mathfrak{f} \in \mathcal{L}^2$ and without loss of generality we assume that the vector $\mathfrak{g}$ is null. We choose an appropriate subspace $\mathcal{X}_\Omega$ of the Hilbert space $\mathcal{L}^2$. The functions in $\mathcal{X}_\Omega$ satisfy the boundary condition:

$$\mathcal{T}_B \mathfrak{v} = 0 \quad \text{on } \hat{\Omega}. \tag{2.63}$$

Then the variational problem (2.60) can be written as

$$\begin{cases} \text{Find } \mathfrak{u} \in \mathcal{X}_\Omega \text{ such that,} \\ \mathfrak{B}(\mathfrak{u}; \mathfrak{v}) = \mathfrak{L}(\mathfrak{v}), \quad \forall\ \mathfrak{v} \in \mathcal{X}_\Omega, \end{cases} \tag{2.64}$$

where, the bilinear form $\mathfrak{B}(\cdot;\cdot)$ is given as

$$\mathfrak{B}(\mathfrak{u}; \mathfrak{v}) = \langle \mathcal{S}_L \mathfrak{u}, \mathcal{S}_L \mathfrak{v} \rangle, \tag{2.65}$$

and the linear form $\mathfrak{L}(\cdot)$ is provided as

$$\mathfrak{L}(\mathfrak{v}) = \langle \mathfrak{f}, \mathcal{S}_L \mathfrak{v} \rangle. \tag{2.66}$$

Let us now formulate the Euler-Lagrange equation corresponding to the variational formulation (2.64). Now using the well-known rule of integration by parts on problem (2.64), we obtain:

$$\left\langle \mathcal{S}_L^* \mathcal{S}_L \mathfrak{u} - \mathcal{S}_L^* \mathfrak{f}, \mathfrak{v} \right\rangle + \langle \mathcal{S}_L \mathfrak{u} - \mathfrak{f}, \mathfrak{v} \rangle_{\hat{\Omega}} = 0. \tag{2.67}$$

where $\mathcal{S}_L^*$ is the adjoint operator of $\mathcal{S}_L$ with its range in $\mathcal{L}^2$, and $\langle \cdot, \cdot \rangle_{\hat{\Omega}}$ is the so-called bilinear form that contains boundary terms generated by the integration by parts rule. Hence we have the following Euler-Lagrange equation

$$\mathcal{S}_L^* \mathcal{S}_L \mathfrak{u} = \mathcal{S}_L^* \mathfrak{f} \quad \text{in } \Omega. \tag{2.68}$$

The boundary condition $\mathcal{T}_B \mathfrak{v} = 0$ serves as an essential boundary condition for (2.68). The original first-order equation serves as a natural boundary condition which can be obtained from the boundary term $\langle \mathcal{S}_L \mathfrak{u} - \mathfrak{f}, \mathfrak{v} \rangle_{\hat{\Omega}} = 0$ for all admissible $\mathfrak{v}$. Therefore, the least-squares method for first-order system is formally equivalent to the Galerkin method for second-order system with the differential equations given in (2.68). The operator $\mathcal{S}_L^* \mathcal{S}_L$ in (2.67) is self-adjoint and nonnegative, even though $\mathcal{S}_L$ itself is not self-adjoint.

Now we understand the advantage of the least-squares method that it converts a difficult non-self-adjoint first-order system into a relatively easy self-adjoint second-order system. For this reason, it is not surprising that the condition number of the resulting matrix produced by LSFEM is $O(h^{-2})$ which is the same as that for the Galerkin method for second-order equations.

2.7 Discrete principles for LSFEM

The continuous least-squares principle $\{\mathcal{X}_\Omega, \mathcal{E}\}$, in section 2.6, propose an alternative, external variational formulation of boundary value problem (2.53). With the help of pair $\{\mathcal{X}_\Omega, \mathcal{E}\}$, we can define finite element methods with Rayleigh-Ritz solver like properties. A discrete least-squares principle (DLSP) can be represented by the pair $\left\{\mathcal{X}_\Omega^h, \mathcal{E}^h\right\}$, where $\mathcal{X}_\Omega^h$ denotes a discrete finite dimensional space (space of piecewise polynomial functions), parameterized by mesh-size parameter h and $\mathcal{E}^h : \mathcal{X}_\Omega^h \to \mathbb{R}$ denotes the energy functional. The discrete least-squares principle under these assumptions can be defined as:

$$\begin{cases} \text{Find } \mathfrak{u}^h \in \mathcal{X}_\Omega^h \text{ such that} \\ \mathcal{E}^h\left(\mathfrak{u}^h; \mathfrak{f}, \mathfrak{g}\right) \leq \mathcal{E}^h\left(\mathfrak{v}^h; \mathfrak{f}, \mathfrak{g}\right) \quad \forall\ \mathfrak{v}^h \in \mathcal{X}_\Omega^h. \end{cases} \tag{2.69}$$

2.7.1 LSFEM and optimal approximations

To answer this question, we need a couple of more assumptions to establish a connection among discrete problem (2.69) and boundary value problem (2.53).

Condition 4 *The least-squares functional is consistent in the sense that for all smooth data* $\mathfrak{f}$ *and* $\mathfrak{g}$ *and all smooth solutions* $\mathfrak{u}$ *of* (2.53)*,*

$$\mathcal{E}^h\left(\mathfrak{u}; \mathfrak{f}, \mathfrak{g}\right) = 0. \tag{2.70}$$

Condition 5 *The least-squares functional is positive, i.e.,*

$$\mathcal{E}^h(\mathfrak{v}^h; \mathfrak{f}, \mathfrak{g}) > 0 \quad \forall\ 0 \neq \mathfrak{v}^h \in \mathcal{X}_\Omega^h. \tag{2.71}$$

From the hypothesis (2.71), we can define discrete energy norm $\|\cdot\|_h : \mathcal{X}_\Omega^h \to \mathbb{R}$ such that

$$\|\cdot\|_h = \mathcal{E}^h(\cdot; 0, 0)^{\frac{1}{2}}, \tag{2.72}$$

Hence, we can also associate an energy inner product $\langle\cdot,\cdot\rangle_h : \mathcal{X}_\Omega^h \times \mathcal{X}_\Omega^h \to \mathbb{R}$ defined as

$$||\mathfrak{v}^h||_h = \sqrt{\langle \mathfrak{v}^h, \mathfrak{v}^h\rangle_h} \qquad \forall\ \mathfrak{v}^h \in \mathcal{X}_\Omega^h. \tag{2.73}$$

Let us now state a theorem which is important to show that the above-discussed conditions are sufficient to solve discrete problem (2.69).

Theorem 6 ([15]) *Suppose that assumptions* (2.70) *and* (2.71) *hold for the DLSP* (2.69) *and let* $\mathfrak{u} \in \mathcal{X}_\Omega$ *denote a sufficiently smooth solution. Then,*

- *The problem* (2.69) *has a unique minimizer* $\mathfrak{u}^h \in \mathcal{X}_\Omega^h$.
- $\mathfrak{u}^h$ *is the orthogonal projection of* $\mathfrak{u}$ *w.r.t. the discrete energy inner product* (2.73).

As an immediate consequence of above theorem, we have following corollary:

Corollary 7 ([15]) *Let the hypothesis of theorem 6 hold. Then, the least-squares solution* $\mathfrak{u}^h$ *minimizes the discrete energy norm error, that is*

$$||\mathfrak{u} - \mathfrak{u}^h||_h = \inf_{v^h \in \mathcal{X}^h_\Omega} ||\mathfrak{u} - \mathfrak{v}^h||_h. \tag{2.74}$$

It is quite easy to define a discrete equivalent variational problem of (2.69) as follows

$$\begin{cases} \text{Find } \mathfrak{u}^h \in \mathcal{X}^h_\Omega \text{ such that,} \\ \mathfrak{B}\left(\mathfrak{u}^h; \mathfrak{v}^h\right) = \mathfrak{L}\left(\mathfrak{v}^h\right), \quad \forall\ \mathfrak{v}^h \in \mathcal{X}^h_\Omega, \end{cases} \tag{2.75}$$

where

$$\mathfrak{B}\left(\cdot;\cdot\right) = \left\langle \cdot, \cdot \right\rangle_h \text{ and } \mathfrak{L}\left(\cdot\right) = \left\langle \mathfrak{u}, \cdot \right\rangle_h. \tag{2.76}$$

It is clear that (2.75) is a linear system of algebraic equations. To represent a matrix-vector form of this problem, let the vector function $\mathfrak{u}^h \in \mathcal{X}^h_\Omega$ be represented as $\mathfrak{u}^h = \sum_{j=1}^{N} \vec{w}_j \psi^h_j$, where $\vec{w} = (w_1, \cdots, w_N)^T$ denotes the coefficient vector and $\left\{\psi^h_j\right\}$ denotes a basis of $\mathcal{X}^h_\Omega$. Then a matrix-vector is given by

$$\mathfrak{A}_{ij} = \left\langle \psi^h_i, \psi^h_j \right\rangle_h \tag{2.77}$$

$$\mathfrak{F}_i = \left\langle \mathfrak{u}, \psi^h_j \right\rangle_h, \tag{2.78}$$

respectively. The problem (2.75) is then equivalent to the linear system of algebraic equations

$$\mathfrak{A}\vec{w} = \mathfrak{F}. \tag{2.79}$$

with the coefficient vector $\vec{w}$ corresponding to $\mathfrak{u}^h$. It is clear that the matrix $\mathfrak{A}$ is symmetric, positive definite and the system (2.79) has a unique solution.

2.8 LSFEM and practicality issues

In this particular section, we discuss briefly some practicality issues regarding least-squares principles. Although, we can not deny the importance of mathematical well-posedness of the considered problem but the primary objective is to construct a worthy computational algorithm. Hence, the proposed techniques must also be practical to implement. The proposed least-squares finite element method must satisfy the practicality measures such that

- The standard discrete finite element spaces should be chosen in a manner that they must be easy to handle.

- The constructed matrices and vectors should be computed without any difficulties.
- The condition number of discrete problem should be regulated conveniently.

2.8.1 Least-squares principle for second order PDEs

The straightforward application of least-squares principle to second order partial differential equation makes the problem impractical in least-squares sense. For example, the equivalent variational equation of the problem (2.33) in section 2.3 must contain the term

$$\int_{\Omega} \triangle \mathfrak{u} \cdot \triangle \mathfrak{v} \; d\Omega, \tag{2.80}$$

with corresponding discrete term

$$\int_{\Omega} \triangle \mathfrak{u}^h \cdot \triangle \mathfrak{v}^h \; d\Omega. \tag{2.81}$$

The terms are well-defined within the finite element space, as it consist of piecewise polynomial functions but the complications arise at the boundary region of considered domain. The terms are not well-defined on boundary except for the finite element spaces which are continuously differentiable in higher order. For multidimensional problem such spaces are not easy to handle. Subsequently, any technique that adopts such terms, including the strategies presented here, is unrealistic.

Another disadvantage of the methods based on the discrete variational problem, containing the term (2.81), is an abrupt growth of order $O\left(h^{-4}\right)$ of the condition number w.r.t. mesh size, regardless of whether or not we utilize smooth finite element spaces. This ought to be appeared differently for the problem (2.33) for which the condition number is of order $O\left(h^{-2}\right)$ for discrete problem in relation to the Rayleigh-Ritz approach. In this regard, the least-squares finite element approach abandons one of the practicality conditions. The discrete term (2.81) also provides the information that a strong regularity is needed for the weak solutions of problem. As it contains two second order square integrable derivative terms which are double as compared to the Galerkin approach. Due to these flaws, the least-squares finite element method did not catch the attention of researchers in early stages of developments.

2.8.2 Least-squares principle for first order linear systems

On the basis of observation discussed above, the researchers came up with a brilliant idea of transformation of second order partial differential equation to an equivalent system of linear first order partial differential equations. The development of a practical and mathematically well-defined least-squares finite element methods requires

more than just picking the most evident functionals which are not necessarily physical quantities. The functional is based on the first order system. In this manner, we overcome majority of the practicality issues of least-squares finite element methods as listed below.

- It allow us to employ simple discrete finite element spaces.
- The weak variational problem contains first order square integrable derivative terms and requires only C^0 regularity to solution space.
- The matrix constructed from linear algebraic system is symmetric positive definite and easy to compute.
- The condition number is manageable as it is of order $O\left(h^{-2}\right)$.

The transformation of second order partial differential equation, in primitive variables, to first order linear system can be done in many ways, as it is not a uniquely defined process. There are three first order linear formulations employed extensively that are associated with least-squares finite element methods, which are as follow:

i) The Vorticity Velocity Pressure (VVP) Formulation

ii) The Stress Velocity Pressure (SVP) Formulation

iii) The Velocity Gradient Velocity Pressure (VGVP) Formulation

The transformation process of higher order partial differential equations to the first order linear system must include some new variables in the resulting system. In this thesis, we are mainly studying the Stokes and Navier-Stokes problems therefore the problem (2.42) can be used for better knowing of the transformation process. The first formulation (i) employs the vorticity ω as a new variable

$$\omega = \nabla \times \mathfrak{u}, \tag{2.82}$$

which is the pseudo vector of the skew-symmetric part $\frac{1}{2}\left(\nabla\mathfrak{u} - \nabla\mathfrak{u}^T\right)$ of the velocity gradient tensor $\nabla\mathfrak{u}$. Therefore the transformation process leads to a vorticity-based first-order system.

$$\left.\begin{aligned} \eta_\nu \nabla \times \omega + \nabla p &= \mathfrak{f} \\ \nabla \times \mathfrak{u} - \omega &= 0 \\ \nabla \cdot \mathfrak{u} &= 0 \end{aligned}\right\} \text{ in } \Omega, \qquad \left.\begin{aligned} \mathfrak{u} &= 0 \\ \mathbf{n} \cdot \omega &= 0 \end{aligned}\right\} \text{ on } \hat{\Omega}. \tag{2.83}$$

In formulation (ii), we work with the symmetric part $\mathsf{D}(\mathfrak{u}) = \frac{1}{2}\left(\nabla\mathfrak{u} + \nabla\mathfrak{u}^T\right)$ of the velocity gradient tensor $\nabla\mathfrak{u}$ and we use stress tensor as

$$\boldsymbol{\sigma} = -\mathfrak{p}\mathsf{I} + 2\eta_\nu \mathsf{D}(\mathfrak{u}). \tag{2.84}$$

Hence the process gives rise to a stress-based first order system.

$$\left.\begin{aligned} -\nabla\cdot\boldsymbol{\sigma} &= \mathfrak{f} \\ \boldsymbol{\sigma} + \mathfrak{p}\mathsf{I} - 2\eta_\nu\mathsf{D}(\mathfrak{u}) &= 0 \\ \nabla\cdot\mathfrak{u} &= 0 \end{aligned}\right\} \text{ in } \Omega, \qquad \left.\begin{aligned} \mathfrak{u} &= 0 \\ \mathbf{n}\cdot\boldsymbol{\sigma} &= 0 \end{aligned}\right\} \text{ on } \hat{\Omega}. \tag{2.85}$$

The formulation (iii) uses the whole velocity gradient tensor $\nabla\mathfrak{u}$ as a new dependent variable of velocity flux. We set velocity flux as

$$\mathbf{U} = (\nabla\mathfrak{u})^T, \tag{2.86}$$

so that $\mathbf{U}_{ij} = \frac{\partial u_i}{\partial x_j}$; $i, j = 1, 2$ then $\nabla\cdot\mathbf{U} = \triangle\mathfrak{u}$ and the transformation process leads to velocity gradient based first order system.

$$\left.\begin{aligned} -\eta_\nu\nabla\cdot\mathbf{U} + \nabla\mathfrak{p} &= \mathfrak{f} \\ \mathbf{U} - (\nabla\mathfrak{u})^T &= 0 \\ \nabla\cdot\mathfrak{u} &= 0 \end{aligned}\right\} \text{ in } \Omega, \qquad \left.\begin{aligned} \mathfrak{u} &= 0 \\ \mathbf{n}\cdot\mathbf{U} &= \mathbf{0} \end{aligned}\right\} \text{ on } \hat{\Omega}. \tag{2.87}$$

for all these formulation the zero mean pressure constraint is used i.e.

$$\int_\Omega \mathfrak{p}\, d\Omega = 0. \tag{2.88}$$

In literature, more research has been done on VVP formulation in [13], [59], [60], [61], [63], [74], [76]. The VGVP based first order systems presented in [28], [71]. The SVP system is used in this work and has been investigated by many researchers, see section 2.1. We further introduce a first order formulation SPVT in chapter 6 for solution of non-Newtonian fluids with non-isothermal properties.

2.8.3 Conflict aspects of norm equivalence and practicality

A least-squares finite element method can be constructed in a way that it fulfills all the required features of the Rayleigh–Ritz settings but it does not necessarily be practical. In general, it is not a difficult task to define a norm-equivalent linear least-squares functional $\mathcal{E}(\cdot;\cdot)$ for any given partial differential equation problem defined under some

Hilbert spaces $\mathcal{X}_\Omega$, $\mathcal{Y}_\Omega$ and $\mathcal{Y}_{\hat{\Omega}}$ such that the following inequality (2.54) holds. Also it can be shown that for the same problem one can define several norm-equivalent linear functionals but not every functional is implementable, see [15] for more detail. Therefore a linear functional generated by a straightforward application of least-squares principle does not necessarily lead to a practical method because sometime it involves the norms that are not easy to implement mathematically. These functionals may contain impractical fractional or negative norms.

In this regard, it is an important task to construct a practical least-squares finite element method. Here, we allude to a standard $\mathcal{L}^2$-least-squares finite element framework enforced to get a first order linear system of equations whose solution is easily computable in comparison to other numerical techniques. To obtain a practical scheme, the choice of spaces is a very significant factor. The space $\mathcal{X}_\Omega$ is set to be a product of $\mathcal{H}^1$ Hilbert space and discretized with standard C^0 finite elements. The spaces $\mathcal{Y}_\Omega$ and $\mathcal{Y}_{\hat{\Omega}}$ are the product of $\mathcal{L}^2$ Sobolev space. In addition to the discussed requirements, if the framework also satisfies the estimate (2.54) then it produces optimal results but if the framework does not fulfill the estimate (2.54) then some mandatory adjustments are required to handle optimality issues. The treatment to tackle such situations is described as under:

- Reconstitution of governing system.
- Employment of weaker spaces $\mathcal{H}^{-1}$ or intermediate spaces $\mathcal{H}^{\frac{1}{2}}$, see [75].
- Use of weighted $\mathcal{L}^2$ norms to replace impractical norms in the discrete problem.
- Appropriate choice of spaces.
- Work with non-conforming discretization

Another strategy is to define a non-equivalent least-squares functional for discrete problem. The main idea is to define least-squares functional in standard $\mathcal{L}^2$-norm for residual energy. In this work, we utilize non-equivalent least-squares finite element method based on $\mathcal{L}^2$-least-squares framework. The energy functional may not be mathematically correct but it is easy to implement. Due to the mathematically incorrectness of energy balance, nothing can be said about the optimality of the method. However, the general assumptions, namely consistency (2.70) and positivity (2.71), established for discrete least-squares principle can easily be satisfied by any sensible definition of a least-squares functional. The non-equivalent least-squares finite element method has been the most frequently exploited technique in the least-squares context. The reason for this is the fact that combination of first-order systems with $\mathcal{L}^2$-norms to measure the residual energy leads to a very simple and easy to implement scheme.

2.8.4 Mass conservation of least-squares method

In many standard finite element methods, one of the major drawback appeared to be the lack of local mass conservation for the incompressible Stokes and Navier-Stokes equations (cf. for further details [52]). Various procedures have been developed in literature to overcome this inadequacy. One can reestablish the local mass conservation of a problem by using discretely divergence-free finite elements [91] or utilizing some distinctive unifications of element [52]. But in context of least-squares finite element method even global mass conservation is not preserved for currently existing formulations. The continuity equation, which provides the divergence-free velocity field, is one of the component of proposed least-squares functional for minimization. The functional can be written as a sum of squared residuals of system equations and if an equation other than continuity equation dominates the residuum, can cause a lack of global mass conservation. The Stokes problem was studied in [29] with weighted least-squares finite element method for flow around cylinder. The author shown the lack of mass conservation at narrow channels around cylinder, as more mass flow into the domain then pass by the cylinder and that is a clear indication of mass inadequacy. Later, Chang introduced an additional constraint to system of equations to overcome mass conservation issue. Unfortunately, his suggested idea resulted in a saddle point problem. The proposed method was impractical as it violated the basic characteristics of least-squares finite element method.

Furthermore the authors in [42] worked in similar directions and tried to reproduce the numerical experiments done in [29]. They associated a scaler weight to the equation (i.e. continuity equation) which establish the mass conservation of the system. According to the article, small weights can be helpful in the improvement of mass conservation. Another treatment to deal with such an issue for 2D problems is the use of higher order finite elements [80]. For very recent techniques and also an overview of the previous efforts we refer to the works of Bochev et al. [21], [22].The treatment and calculation of global mass conservation is discussed in Section 3.4.5 for our work.

2.9 Least-squares principle for nonlinear problem

The nonlinear problem, in the context of least-squares finite element method, is a bit complicated as compared to the Galerkin approach. To solve any nonlinear partial differential equations, there are two main stages which can be considered i.e. discretization and linearization. The usual approach, which is commonly used together with the Galerkin variational principle is to discretized the fully nonlinear problem. For further elaboration, we consider an arbitrary nonlinear partial differential operator $\hat{L}$ as

$$\hat{L}\mathfrak{u} = \mathfrak{f} \text{ in } \Omega, \tag{2.89}$$

and the residual of system is given by

$$\mathcal{R}(\mathfrak{u}) = \hat{L}\mathfrak{u} - \mathfrak{f} = \mathbf{0}, \tag{2.90}$$

the discretization yields a nonlinear algebraic system:

$$\bar{K}(\mathfrak{u})\,\mathfrak{u} = \mathbf{0}, \tag{2.91}$$

where the finite dimensional vector $\mathfrak{u}$ represents the approximating finite element function. Usually this algebraic system is then solved with some appropriate method for nonlinear systems of equations.

The second way is to linearize the differential nonlinear operator $\hat{L}$ on the continuous level. Subsequently, a mapping should be defined as $\hat{L} : X_\Omega \to Y_\Omega$, where X_Ω, Y_Ω are subspaces in $W^{k,p}(\Omega)^n$ and the operator $\hat{L}$ is considered to be Fréchet differentiable. The nonlinear problem satisfies the exact solution $\mathfrak{u}^* \in X_\Omega$ in a weak sense such that

$$\mathcal{R}(\mathfrak{u}^*) = \hat{L}\mathfrak{u}^* - \mathfrak{f} = \mathbf{0}. \tag{2.92}$$

To find a direct solution $\mathfrak{u}^*$ for the problem (2.92) is not possible most of the time. Therefore, we consult iterative methods with some initial guess $\mathfrak{u}^k$. The Newton's iterative method solves the problem as

$$\mathcal{DR}\left(\mathfrak{u}^k\right)[\delta\mathfrak{u}] = -\mathcal{R}\left(\mathfrak{u}^k\right), \tag{2.93}$$

where the new solution refinement can be given as $\mathfrak{u}^{k+1} = \mathfrak{u}^k + \delta\mathfrak{u}$. Undoubtedly, the refined term $\mathfrak{u}^{k+1}$ is not the exact solution of system (2.92), but under few sufficiently strong conditions and with proper choice of initial guess $\mathfrak{u}_0$ the sequence $\mathfrak{u}^k$ eventually converge to the exact $\mathfrak{u}^*$(cf. [87]). Furthermore, the Galerkin variational principle applied in both approaches, discussed in the Section 2.3, lead to same system of equations. This mean that the order of discretization and linearization do not matter and can be exchanged in case of Galerkin method.

For the least-squares finite element method, the above observations do not hold and yield distinct results. The difference among these approaches can be shown by deriving abstract variational forms for both approaches.

Least-squares variation principle before linearization

In this section we study the behavior of nonlinear least-squares functional as given by

$$\mathcal{E}(\mathfrak{u}) = \frac{1}{2}\left\|\mathcal{R}(\mathfrak{u})\right\|_0^2, \tag{2.94}$$

and minimization problem associated with functional is given as $\min_{u \in X_\Omega} \mathcal{E}(\mathfrak{u})$. The necessary condition or first variant of (2.94) is given as

$$\mathcal{DE}(\mathfrak{u})[\mathfrak{v}] = (\mathcal{R}(\mathfrak{u}))(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]) = 0, \tag{2.95}$$

since $\mathcal{DE}(\mathfrak{u})[\mathfrak{v}]$ is nonlinear in $\mathfrak{u}$. To apply the Newton method, we need to calculate the first variant of (2.95), such that

$$\begin{aligned}\mathcal{D}^2\mathcal{E}(\mathfrak{u})[\mathfrak{v},\delta\mathfrak{u}] &= (\mathcal{R}(\mathfrak{u}))\left(\mathcal{D}^2\mathcal{R}(\mathfrak{u})[\mathfrak{v},\delta\mathfrak{u}]\right) \\ &\quad + (\mathcal{DR}(\mathfrak{u})[\delta\mathfrak{u}])(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]).\end{aligned} \tag{2.96}$$

One Newton iteration can be written as

$$\begin{aligned}(\mathcal{R}(\mathfrak{u}))\left(\mathcal{D}^2\mathcal{R}(\mathfrak{u})[\mathfrak{v},\delta\mathfrak{u}]\right) + (\mathcal{DR}(\mathfrak{u})[\delta\mathfrak{u}])(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]) &= \\ -(\mathcal{R}(\mathfrak{u}))(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}])\ \forall\ \mathfrak{v} &\in X_\Omega.\end{aligned} \tag{2.97}$$

Least-squares variational principle after linearization

In this approach, the least-squares variational principle applies to the functional which is obtained from the Newton linearization of the nonlinear terms in equation (2.89). The functional we get is

$$\mathcal{E}(\delta\mathfrak{u}) = \frac{1}{2}\left\|\mathcal{DR}(\mathfrak{u})[\delta\mathfrak{u}] + \mathcal{R}(\mathfrak{u})\right\|_0^2, \tag{2.98}$$

then the first variation of (2.98) gives the following iterative scheme

$$\mathcal{DE}(\delta\mathfrak{u})[\mathfrak{v}] = 0, \tag{2.99}$$

$$(\mathcal{DR}(\mathfrak{u})[\delta\mathfrak{u}])(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]) + (\mathcal{R}(\mathfrak{u}))(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]) = 0,$$

$$(\mathcal{DR}(\mathfrak{u})[\delta\mathfrak{u}])(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}]) = -(\mathcal{R}(\mathfrak{u}))(\mathcal{DR}(\mathfrak{u})[\mathfrak{v}])\ \forall\ \mathfrak{v} \in X_\Omega. \tag{2.100}$$

It can be observed from the equation (2.97) that it contains an extra term with second order Fréchet derivative, in comparison to equation (2.100).

Remark 8 *The order of the application of least-squares variational principle and linearization influences the nonlinear problem and leads to contrary outcomes.*

In literature, both techniques have been employed and numerically examined. The research analysis for first approach can be found in [12], [14], [16], [20], where the Navier Stokes equation based on different least-squares formulations had been studied for the nonlinear least-squares functional. Whereas in least-squares sense, when the nonlinear problem is initiated by linearization, then the least-squares principle is exercised directly to linear problem and proposed theory of [63] is applicable. The detailed analysis of these two approaches was done in [33] and concluded that near solution the first order derivative terms dominate over second order derivative terms. That means, if the radius of convergent is smaller than the first approach behave same as second approach. Also the calculation of second order derivative term is a difficult task and it increases the regularity of the solution's space. Therefore, in this work the second approach is used for our research purpose.

Chapter 3

Numerical Methods And Geometries

If modeling the physical problem is a challenging task then solving it numerically is not that easy at all and in fact is even harder. The solution strategies for fluid flow problems have been investigated from many decades, this chapter sorts it out with references therein. Furthermore, we explain the general differences between the direct and iterative solvers. The core ideas of iterative solvers like Krylov subspace solver and multigrid solvers strategies are explained thoroughly with associated references.

3.1 Methods for nonlinear problem

In this section, the numerical techniques for solution of nonlinear stationary incompressible problems are the main focus of our discussion.

3.1.1 Nonlinear basic Iterative methods

The methods are formulated by the following type of abstract nonlinear problem given by equation (2.89) and proposed to be stationary. The nonlinear basic iteration, cf. [91], is divided into three stages to solve a nonlinear problem. Let w^k be the known iteration and the residual d^k of nonlinear problem is calculated by this known value. The new solution update δw is approximated as following

Provided: iterate the given value w^k
Execute: The following steps to produce w^{k+1}
Step-I: Compute the residual

$$d^k = \hat{L}\left(w^k\right) w^k - f,$$

Step-II: Solve a subproblem for δw with residual d^k on right hand side

$$\tilde{\mathcal{R}}\left(w^k\right) \delta w = d^k,$$

Step-III: Update the solution w^k via δw and relaxation parameter ϖ to get w^{k+1}

$$w^{k+1} = w^k - \varpi \delta w.$$

Algorithm 3.1: Nonlinear Basic Iteration

the operator $\tilde{\mathcal{R}}\left(w^k\right)$ is the Fréchet derivative of nonlinear operator $\hat{L}$ w.r.t. to last approximate solution w^k. The value ϖ is an extra damping parameter which has to be selected properly. The nonlinear iteration scheme may be terminated if a maximum number of iterations $N_{\max}$ have been achieved, i.e. $k+1 \geq N_{\max}$, or if the residual of considered system is smaller enough to some tolerance parameter TOL set by user and it is given in a certain norm, such as

$$\left\| \hat{L}\left(w^{k+1}\right) w^{k+1} - \mathfrak{f} \right\| < TOL. \tag{3.1}$$

In the least-squares context, we approximate the squares of residual (2.94) of the given problem. Then, the method mentioned in Algorithm 3.1 take a different form due to the application of least-squares principle as shown in Algorithm 3.2.

If we consider $\tilde{\mathcal{R}} = \|\mathcal{DR}\|_0^2 = \langle \mathcal{DR}, \mathcal{DR} \rangle$ and $\mathcal{R}_{def} = \langle \mathcal{R}, \mathcal{DR} \rangle$ then the following nonlinear iterative scheme is achieved,

$$w^{k+1} = w^k - \varpi \left[\left\langle \mathcal{DR}\left(w^k\right), \mathcal{DR}\left(w^k\right) \right\rangle \right]^{-1} \left[\left\langle \mathcal{R}\left(w^k\right), \mathcal{DR}\left(w^k\right) \right\rangle \right], \tag{3.2}$$

or for simplicity

$$w^{k+1} = w^k - \varpi \left[\mathcal{DR}\left(w^k\right)^* \mathcal{DR}\left(w^k\right) \right]^{-1} \left[\mathcal{R}\left(w^k\right)^* \mathcal{DR}\left(w^k\right) \right]. \tag{3.3}$$

where $\mathcal{DR}$ represents the Fréchet derivatives of operator $\hat{L}$. This scheme is called Gauss-Newton method and known to have super linear/linear convergence. The convergence radius of the Gauss-Newton method is very narrow and it requires a very good initial guess to acquire a convergent scheme. The stopping criterion for Gauss-Newton method is given as

$$\left\| \left\langle \mathcal{R}\left(w^{k+1}\right), \mathcal{DR}\left(w^{k+1}\right) \right\rangle \right\| < TOL. \tag{3.4}$$

Provided: iterate the given value w^k,
We define a least squares functional as

$$\mathcal{E}\left(w^k\right) = \frac{1}{2}\left\|\mathcal{R}\left(w^k\right)\right\|_0^2 = \frac{1}{2}\left\|\hat{L}\left(w^k\right)w^k - \mathfrak{f}\right\|_0^2.$$

Execute: The following steps to produce w^{k+1},
Step-I: Calculate the residual $\mathcal{R}_{def}^k$ for LS problem as

$$\mathcal{R}_{def}^k = \left\langle \mathcal{R}\left(w^k\right), \mathcal{D}\mathcal{R}\left(w^k\right)\right\rangle.$$

Step-II: Solve a subproblem for auxiliary solution δw with defect residual $\mathcal{R}_{def}^k$ on right hand side

$$\tilde{\mathcal{R}}\left(w^k\right)\delta w = \mathcal{R}_{def}^k.$$

Step-III: To get w^{k+1}, update the solution w^k with auxiliary solution δw and relaxation parameter ϖ as

$$w^{k+1} = w^k - \varpi\delta w.$$

Algorithm 3.2: Nonlinear Basic Iteration for LS

We use another iterative scheme so-called fixed point iterative method for the operator $\tilde{\mathcal{R}}$. In such type of method, the nonlinear operator $\hat{L}$ may be chosen itself as an approximate Fréchet derivative or only its linear part. Since, the least-squares principle leads to the following minimization system

$$\left\|\mathcal{R}\left(w^k\right)\right\|^2 = \left\langle \mathcal{R}\left(w^k\right), \mathcal{R}\left(w^k\right)\right\rangle = 0, \tag{3.5}$$

Therefore, we only consider linear part in the residual of nonlinear system $\mathcal{R}$ to develop fixed point method. The iterative scheme can be given as

$$w^{k+1} = w^k - \varpi\left[\left\langle \mathcal{R}\left(w^k\right), \mathcal{R}\left(w^k\right)\right\rangle\right]^{-1}\left[\left\langle \mathcal{R}\left(w^k\right), \mathcal{R}\left(w^k\right)\right\rangle\right], \tag{3.6}$$

or

$$w^{k+1} = w^k - \varpi\left[\mathcal{R}\left(w^k\right)^*\mathcal{R}\left(w^k\right)\right]^{-1}\left[\mathcal{R}\left(w^k\right)^*\mathcal{R}\left(w^k\right)\right], \tag{3.7}$$

the general convergence behavior of such scheme is linear or some time super linear as well. The stopping criterion is similar as inequality (3.4), but the term $\mathcal{D}\mathcal{R}\left(w^{k+1}\right)$ is replaced by the term $\mathcal{R}\left(w^{k+1}\right)$. In many cases it is helpful to generate a good initial guess for Gauss-Newton method.

3.2 Methods for sparse linear systems

A linear system of equations of form given as

$$\mathbf{K}u = \mathbf{b}, \tag{3.8}$$

with sparse and large coefficient matrix $\mathbf{K}$ can be constructed either from discretization of linear problem or linearization and discretization of a nonlinear problem. The modelling and simulation process for such sparse and large problem is a challenging task, as it may be done at huge computational cost. A huge amount of research efforts have been made to develop efficient algorithms for problem (3.8). In general, these methods split into two important branches, the direct methods and iterative methods. In the forthcoming subsections, we provide the brief description of the methods which are adopted for our research problems and citations are presented additionally for further details.

3.2.1 Direct methods

The direct methods can be seen as an efficient development of Gaussian elimination method and these methods generate the exact solution after performing limited numbers of operations to a linear system in the absence of round off error. These methods are only applicable when $\mathbf{K}$ is a square matrix in the problem (3.8) . In our case, we employed the direct Gaussian elimination solver (UMFPACK [39]) as a coarse grid solver, where the small sparse systems are under consideration in 2D. Also it gives a very robust linear scheme, despite the fact that the memory and CPU time requirements are too high when the provided linear system is large. For further detail on sparse direct solvers one can see [40], [41], [45].

3.2.2 Iterative methods

The iterative methods refer to the schemes that adopt iterative procedures to obtain the accurate approximation towards the solution of a large linear system at each iteration. The solvers are subdivided into two major categories, namely the Krylov Subspace solvers and the Multigrid solver. In the following subsections general description of these methods is documented and the solution algorithms to solve our problem is incorporated. One may go through [85] for comprehensive discernment on the iterative methods.

Krylov space methods

In the twentieth century, the Krylov space methods are considered among the top rated [32] iterative methods to deal with large sparse linear systems. The distinct

approximation properties and modest storage requirements make them popular. In this thesis, we encounter large sparse linear system, due to the characteristic of the interpolation functions utilized in the finite element discretization, and direct methods are not practical in this scenario. The Krylov methods are considered to be very effective schemes for large linear problems with suitable choice of preconditioners and it obtain approximate solution from a finite m-dimension subspace $u_0 + \mathcal{K}_m$, where u_0 represents an arbitrary initial guess and $\mathcal{K}_m$ is the Krylov subspace defined as:

$$\mathcal{K}_m\left(\mathbf{K}, r_0\right) = span\ \left\{r_0, \mathbf{K}r_0, \mathbf{K}^2 r_0, \cdots, \mathbf{K}^{m-1} r_0\right\}, \tag{3.9}$$

with $r_0 = \mathbf{b} - \mathbf{K}\mathfrak{u}_0$. For simplicity $\mathcal{K}_m\left(\mathbf{K}, r_0\right)$ is denoted by $\mathcal{K}_m$ later on. A common practice is to proceed with initial guess $\mathfrak{u}_0 = 0$ which permits $r_0 = \mathbf{b}$, and the corresponding Krylov subspace is $\mathcal{K}_m\left(\mathbf{K}, \mathbf{b}\right)$ achieved by the given linear system.

The Krylov space schemes are based on projection processes, both orthogonal and incline, onto Krylov subspaces which are subspaces spanned by vectors of the form $p(\mathbf{K})v$, where p is a polynomial. More precisely, theses techniques approximate $\mathbf{K}^{-1}\mathbf{b}$ by $p(\mathbf{K})\mathbf{b}$, where p is unique minimal polynomial [62] of smallest degree for which $p(\mathbf{K}) = 0$. It can be easily seen that $\mathbf{K}^{-1}\mathbf{b}$ is a member of Krylov subspace and approximation obtained is of following form

$$\mathbf{K}^{-1}\mathbf{b} \approx q_{k-1}\left(\mathbf{K}\right) b, \tag{3.10}$$

where q_{k-1} is a minimal polynomial of degree $k-1$. In simple words the linear system $\mathbf{K}^{-1}\mathbf{b}$ is approximated by a polynomial $q_{k-1}\left(\mathbf{K}\right) b$. The convergence of Krylov method is depends on the degree of minimal polynomial, if it is small the Krylov method converge faster. Because, the corresponding Krylov subspace has small dimension which contain the solution vector.

Conjugate gradient method The Conjugate Gradient (CG) method is well known to solve the linear problem (3.8) by using concepts of orthogonal or conjugate vectors. This class of methods was first introduced by Hestenes and Stiefel [56] in early 50s of twentieth century. The CG method is very efficient to deal with symmetric positive definite systems. Then, the minimization problem $\min \mathbf{Q}$ of a quadratic form

$$\mathbf{Q}(\mathfrak{u}) = \frac{1}{2}\mathfrak{u}^T\mathbf{K}\mathfrak{u} - \mathfrak{u}^T\mathbf{b}, \tag{3.11}$$

is an equivalent problem of finding solution for the linear system, if

$$\mathcal{D}\mathbf{Q}(\mathfrak{u}) = \text{grad}\ \mathbf{Q}\left(\mathfrak{u}\right) = \mathbf{K}\mathfrak{u} - \mathbf{b} = 0. \tag{3.12}$$

So, the CG method can be seen as an iterative scheme to find the minimizer $\mathfrak{u}_k \in \mathfrak{u}_0 + \mathcal{K}_k\left(\mathbf{K}, r_0\right)$, at kth-iteration, of form (3.11) and eventually an iterative solver of

linear system (3.8). The working mechanism of CG method is given in the Algorithm 3.3 below.

Compute: $r_0 := \mathbf{b} - \mathbf{K}u_0$ with initial values $u_0, v_0 := r_0$.
If r_0 is sufficiently small, then return u_0 as result and exit.
FOR $k := 0, 1, \cdots$, till convergence, DO
$w_k := \mathbf{K}v_k$
$\alpha_k := \frac{(r_k, r_k)}{(w_k, v_k)}$
$u_{k+1} := u_k + \alpha_k v_k$
$r_{k+1} := r_k - \alpha_k w_k$
$\beta_k := \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)}$
$v_{k+1} := r_{k+1} + \beta_k v_k$
If r_{k+1} is sufficiently small, then exit the loop.
END DO
Output: Return u_{k+1}.

Algorithm 3.3: Conjugate Gradient Method

It can be observed from the algorithm that the vector $\mathfrak{v}_k$ which gives the search direction is highly significant for optimality of scheme. If the iterative solution at $\mathfrak{u}_{k+1}$ is optimal w.r.t. $\mathfrak{v}_k \neq \mathbf{0}$, then the next iteration $\mathfrak{u}_{k+2}$ must be optimal and it is possible only when the following condition hold, $\mathbf{K}\mathfrak{v}_{k+1} \perp \mathfrak{v}_k$, i.e.

$$\left(\mathbf{K}\mathfrak{v}_{k+1}, \mathfrak{v}_k\right) = \mathbf{0}, \tag{3.13}$$

where the direction vectors $\mathfrak{v}_{k+1}$ and $\mathfrak{v}_k$ are orthogonal or pair wise conjugate to each other. As each new search direction is generated by the remaining residual and conjugate to the prior search direction, it is also conjugate to all previously generated search directions. Therefore a system of conjugate directions is developed or an equivalent system of orthogonal residuals.

Theorem 9 (For proof see [70]) *Suppose the CG method applied to a SPD system* $\mathbf{K}\mathfrak{u} = \mathbf{b}$. *Then, the following inequality for errors, at kth iteration, will hold*

$$\frac{\|\mathfrak{u} - \mathfrak{u}_k\|_{\mathbf{K}}}{\|\mathfrak{u} - \mathfrak{u}_0\|_{\mathbf{K}}} \leq 2\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k ; \text{ for } k \geq 0, \tag{3.14}$$

where κ *is 2-norm condition number*[1] *of matrix* $\mathbf{K}$.

[1]The Condition Number $\kappa(\mathbf{A})$ is defined as

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}},$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are maximum and minimum eigen values of SPD matrix $\mathbf{A}$.

From above theorem, It is observed that the convergence rate of CG scheme is linear and depends on the spectral condition number of matrix $\mathbf{K}$. Although, with help of proper super linear preconditioning techniques one can achieve a super linear convergence for CG method. For more discussion on the convergence of CG method cf. [25], [34] and [95].

Preconditioning

It is an absolute truth that lack of robustness/stability are common flaws of iterative methods, if compared to direct methods. But still we can not fully ignore them as they are applicable to large linear systems. To counter such situations, the proper use of suitable techniques, for instant precoditioning, can substantially improve the efficiency and robustness of iterative schemes. Earlier in this section, we have discussed briefly that the convergence rate of Krylov subspace solvers for symmetric linear systems is strongly dependent on the spectral properties of the system matrix. Particularly, the conjugate gradient method produces best results when the condition number of matrix $\kappa(\mathbf{K})$ is small or eigenvalues cumulate around one [11].

Preconditioning techniques are usually exercised to the systems with built-in deficiencies as discussed above. The main idea of preconditioning is to reconstruct the new systems having common solution as the original systems but the properties of new systems are favorable for the iterative solvers. For the linear system (3.8), the preconditioner $\mathbf{P}$ can be defined as

$$\mathbf{P}^{-1}\mathbf{K}\mathfrak{u} = \mathbf{P}^{-1}\mathbf{b}, \tag{3.15}$$

where $\mathbf{P}$ is an invertible matrix[2]. This system is expected to be solved easily by the Krylov space methods and has same solution as system (3.8). A preconditioner is considered to be good when the choice of matrix $\mathbf{P}$ is close to coefficient matrix $\mathbf{K}$. In this regard, the condition number of preconditioned matrix is closed to one and the method will converge fast.

Multigrid method

The geometric multigrid method (GMG) is another renown alternative approach to solve large and sparse linear systems generated from discretization of partial differential equations. Initially, multigrid scheme was developed only for elliptic problems but later on it was also used for other type of PDEs problem. For general introduction on the

[2] A nonsingular square martix $\mathbf{A}$ is said to be invertible if and only if there exists a square matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I},$$

where $\mathbf{I}$ is identity matrix.

method, one can consults the book [54]. From numerical point of view, the multigrid method can be described as a procedure that combines a set of algorithms which solve the discrete problem (3.8) of differential equations on different hierarchical mesh levels. This makes it one of the successful and fastest method for CFD problems [97].

To explain the working cycle of geometric multigrid method, we assume that there exists the hierarchical mesh levels $l+1$; $j = 0, \cdots .l$,where l is the finest mesh level and the index c represents any coarse mesh level such that $0 \leq c < l$. The linear problem (3.8) on finest level can be written as

$$\mathbf{K}_l \mathfrak{u}_l = \mathbf{b}_l. \tag{3.16}$$

The multigrid procedure for one coarse grid correction is given as follow:

Pre-Smoothing: The actual vector as initial guess at level l is $\mathfrak{u}_l^k$. On this level, few iterations of a basic iterative scheme are performed to remove highly oscillating errors presented in the solution vector $\mathfrak{u}_l^k$ and obtain a smoothed error field. These schemes are known as smoothers and process on this specific level is called pre-smoothing. The improved solution $\mathfrak{u}_l^{k+\frac{1}{3}}$ obtained after smoothing process with step size ϵ_1 and corresponding residual vector $r_l^{k+\frac{1}{3}}$ is shown as under

$$\left[\mathfrak{u}_l^{k+\frac{1}{3}}, r_l^{k+\frac{1}{3}}\right] := \mathbf{Smoother}\left(\mathbf{K}_l, \mathfrak{u}_l^k, \mathbf{b}_l, \epsilon_1\right). \tag{3.17}$$

Restriction: Now multigrid perform a transformation $\mathbb{I}_l^c := \mathbb{R}^{n_l} \rightarrow \mathbb{R}^{n_c}$ of residual vector from fine mesh level to coarse mesh level. This procedure is named as restriction process and given as

$$r_c^{k+\frac{1}{3}} := \mathbb{I}_l^c r_l^{k+\frac{1}{3}}. \tag{3.18}$$

Coarse level correction: Now the smoothed error is approximated efficiently on the coarse level for a smaller linear system

$$\mathbf{K}_c \mathfrak{u}_c^e := r_c^{k+\frac{1}{3}}, \tag{3.19}$$

and get a potential correction $\mathfrak{u}_c^e$ for the solution at finest level.

Prolongation: Now a reverse transformation $\mathbb{I}_c^l := \mathbb{R}^{n_c} \rightarrow \mathbb{R}^{n_l}$ from coarse level to fine level is performed by prolongation or interpolation process

$$\mathfrak{u}_l^e := \mathbb{I}_c^l \mathfrak{u}_c^e, \tag{3.20}$$

and a significant correction is made to solution vector on finest level

$$\mathfrak{u}_l^{k+\frac{2}{3}} := \mathfrak{u}_l^{k+\frac{1}{3}} + \mathfrak{u}_l^e. \tag{3.21}$$

Post-Smoothing: Few more iterations on the obtained solution $\mathfrak{u}_l^{k+\frac{2}{3}}$ by a smoother lead to the refined approximate solution vector $\mathfrak{u}_l^{k+1}$ with corresponding residual r_l^{k+1}

$$\left[\mathfrak{u}_l^{k+1}, r_l^{k+1}\right] := \mathbf{Smoother}\left(\mathbf{K}_l, \mathfrak{u}_l^{k+\frac{2}{3}}, \mathbf{b}_l, \epsilon_2\right), \tag{3.22}$$

where ϵ_2 is step size for pre-smoothing process. This is a complete two level process for the multigrid method.

The main goal in multigrid method is to obtain a small sparse linear system on coarse grid level and common exercise to solve such system is by employing an appropriate direct solver. However, if the system matrix is not small on coarse level then the choice of solver is very crucial. The direct solver is not a good option for such situation as it requires more memory and eventually is very expensive for large linear system. In this case, the multigrid method will shift the problem to more coarsest grid level and this hierarchical process will continue until the requirements for coarse grid solver meet. This hierarchical grid process in multigrid is multigrid cycle. The most common and famous multigrid cycles are V-cycle,W-cycle and F-cycle. The type of the multigrid cycle along with other multigrid components play a vital role in the convergence behavior of MG method. We investigate F-multigrid cycle for study purposes.

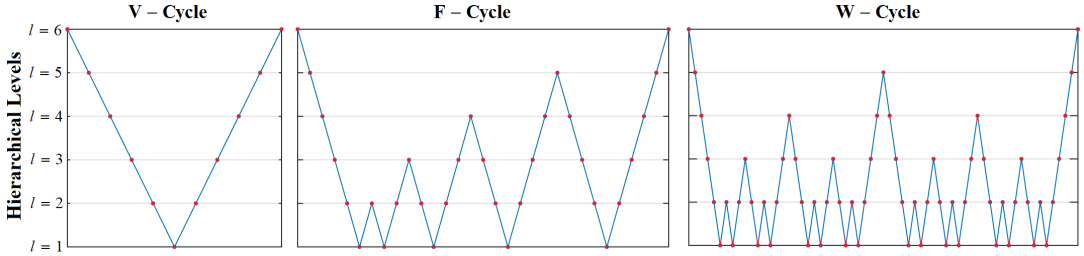

Figure 3.1: Multigrid V, F and W cycles with (\) restriction, (/) prolongation and (•) pre/post smoothing operations

Multigrid-preconditioned conjugate gradient solver

The CG method, as discussed in previous sections, is highly dependent upon the condition number of the considered linear system. However, the convergence rate of CG solver can be reasonably enhanced by employing suitable preconditioning approaches [85]. For the solution of least-squares finite element method, a lot of research has to be done on preconditioned conjugate gradient (PCG) method. One of the most employed candidate for this purpose is PCG based on Jacobi method (JPCG) [85]. Poles apart from Jacobi solver, the multigrid scheme is another powerful successor for PCG method used in [58], [59] and [60]. The LSFEM for Navier-Stokes equations is studied in [82] with geometric multigrid preconditioned CG method and observed that MPCG's convergence rate is better than the JPCG method. The Algorithm 3.4 of MPCG solver is described below.

Compute: $r_0 := \mathbf{b} - \mathbf{K}u_0$ with initial values u_0, $M_G z_0 := r_0$ and $v_0 := z_0$.
If r_0 is sufficiently small, then return u_0 as result and exit.
FOR $k := 0, 1, \cdots$, till convergence, DO
$w_k := \mathbf{K}v_k$
$\alpha_k := \frac{(r_k, z_k)}{(w_k, v_k)}$
$u_{k+1} := u_k + \alpha_k v_k$
$r_{k+1} := r_k - \alpha_k w_k$
$M_G z_{k+1} := r_{k+1}$
$\beta_k := \frac{(r_{k+1}, z_{k+1})}{(r_k, z_k)}$
$v_{k+1} := z_{k+1} + \beta_k v_k$
If r_{k+1} is sufficiently small, then exit the loop.
END DO
Output: Return u_{k+1}.

Algorithm 3.4: Multigrid Preconditioned Conjugate Gradient Method

In this work, we use geometric multigrid scheme as preconditioner of CG method which serve as main solver for linear system. Apart from that, we also use CG method as a pre/post smoother for multigrid method to enhance convergence and robustness, see [48] and [66] for more detail. Since, at each smoothing step CG smoother determines the size of the solution updates accordingly, therefore it leads to productive and specially parameter-free smoothing sweeps. Additionally, the smoothing process is further accelerated by SSOR preconditioner. At the coarsest level of the problem, the direct Gaussian elimination [39] procedure is followed.

3.3 Finite element approximation

The Discrete variational problem for least-squares finite element scheme has been discussed already in Section 2.7. The construction of finite elements approximation is explained briefly in this section.

3.3.1 Basic idea

The basic concept behind finite elements approximation is to break the considered domain $\Omega \subset \mathbb{R}^n$ into small partitions Ω_{e_k} such that the problem becomes easy to approximate. For the problems in 1-dimension these partitions are referred to as intervals, for 2-dimensional problems the partitions are picked as triangles or quadrilaterals and similarly tetrahedron or hexahedron are chosen for 3-dimensional cases. To find an approximate solution $\mathfrak{u}^h$ of discrete variational problem (2.75), the computational domain is divided into N_e numbers of finite partitions Ω_{e_k}, which are called elements,

with following properties

$$\bar{\Omega} := \bigcup_{k=1}^{N_e} \bar{\Omega}_{e_k}, \quad \Omega_{e_k} \bigcap \Omega_{e_l} = \emptyset;\ k \neq l.$$

The finite element spaces are described by the range of basis functions characterized on the partitions studied in [44]. The approximations are considered as follow

$$\mathfrak{u}^h := \mathfrak{u}^h_{\hat{\Omega}_D} + \sum_{j=1}^{N_e} \mathfrak{u}_j \phi_j, \tag{3.23}$$

where $\mathfrak{u}^h_{\hat{\Omega}_D}$ is discrete value on Dirichlet boundary $\hat{\Omega}_D$ and

$$\mathfrak{v}^h := \sum_{i=1}^{N_e} \mathfrak{v}_i \psi_i, \tag{3.24}$$

where N_e is the total number of elements, the piecewise polynomials ϕ_j and ψ_i are called basis functions, which vanish on boundary $\hat{\Omega}_D$. The approximate $\mathfrak{u}^h$ in (3.23) in some finite-dimensional trial space $\mathcal{X}^h_{\hat{\Omega}_D} \subset \mathcal{H}^1_{\hat{\Omega}_D}$ is referred to as a trial function. Similarly, an approximate $\mathfrak{v}^h$ in the finite-dimensional test space $\mathcal{X}^h_0 \subset \mathcal{H}^1_0$ is defined as test function. For more information one can consult [25], [50], [65], [72] and [73].

Remark 10 *From equation* (3.23) *it is very clear that the Dirichlet boundary conditions are essentially contained in the considered spaces, hence, they are sometimes referred to as essential boundary conditions. The Neumann boundary conditions are named natural boundary conditions occasionally, because they are automatically built in the variational form, and are not explicitly imposed in trial and test spaces.*

Construction of shape functions

In practice, the finite element implementation is allocated element-wise using the local basis termed as shape functions ϕ_j. Therefore, the finite element construction critically depends on these shape functions. A local approximate solution $\mathfrak{u}^h_{local}$ inside an element Ω_{e_k} can be represented as a combination of local basis functions and nodal approximate values $\mathfrak{u}_j$ on the local degrees of freedom n_{dof} of an element and defined as

$$\mathfrak{u}^h_{local}(\mathrm{w}) := \sum_{j=1}^{n_{dof}} \mathfrak{u}_j \phi^k_j(\mathrm{w}), \quad \forall\ \mathrm{w} \in \Omega_{e_k}. \tag{3.25}$$

Now, we can expand this concept for all the elements in the domain Ω. So, a global approximate solution $\mathfrak{u}^h$ is constructed by summing up all local basis functions ϕ_j over the domain with total N_{dof} degree of freedom in region Ω.

$$\mathfrak{u}^h(\mathrm{w}) := \sum_{j=1}^{N_{dof}} \mathfrak{u}_j \phi_j(\mathrm{w}), \quad \forall\ \mathrm{w} \in \Omega. \tag{3.26}$$

The shape functions $\phi_j : \bar{\Omega} \to \mathbb{R}$ are enforced to have the following general properties,

- The shape function at any node w_i must have value one at that specific nodal point and value zero for all other nodes of domain at the same time. It can be given as

$$\phi_j(\mathrm{w}_i) := \delta_{ij} = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{when } i \neq j, \end{cases} \tag{3.27}$$

  where δ_{ij} is referred to as Kronecker delta function.

- The summation of all local basis functions ϕ_j^k on nodal points should be equivalent to one for each element $\Omega_{e_k} \subset \Omega$, that is

$$\sum_{j=1}^{N_{dof}} \phi_j(\mathrm{w}) = 1. \tag{3.28}$$

- The summation of derivatives of all basis functions should be zero in each element, such that

$$\sum_{j=1}^{N_{dof}} \nabla\phi_j(\mathrm{w}) = 0. \tag{3.29}$$

- The local basis function should vanish on any element point other then nodal points for computational efficiency. Consequently, a compact local support established for the global basis functions to achieve an approximation which tends to sparse linear system.

Now in next section, we discuss some particular elements used in this work.

3.3.2 Types of finite element

The finite elements can be divided into two groups, one with continuous pressure and the other with discontinuous pressure. In this thesis, we deal with two-dimensional problems and use quadrilateral finite elements for the problem's discretization. Therefore, we present brief description for few quadrilateral finite elements below.

Q_1-Finite element

The Q_1-Finite Element is also known as bilinear element. It is composed of 4 nodes, which are named as degrees of freedom, by including all the 4 corners of quadrilateral as presented in the Figure 3.2. For computational point of view, the shape functions are constructed on some reference element $\Omega_{e_{ref}}$ with smooth shape rather than directly on physical element Ω_{e_k}. Let (ξ_x, ξ_y) be a local coordinate system and a reference element $\Omega_{e_{ref}} := [-1, 1] \times [-1, 1]$ is defined at the origin coordinates of the system.

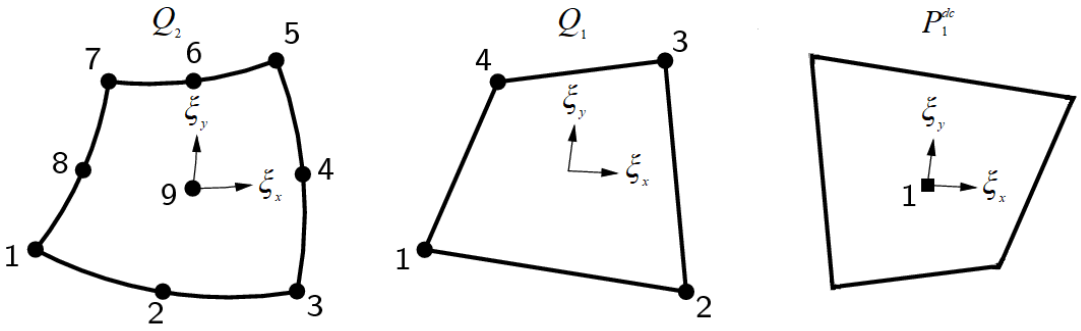

Figure 3.2: Quadrilateral finite elements

A bilinear mapping $\psi_{_\Omega} : \Omega_{e_{ref}} \rightarrow \Omega_{e_k}$ transforms the reference element to physical element. In similar fashion, one can define an inverse mapping $\psi_{_\Omega}^{-1} : \Omega_{e_k} \rightarrow \Omega_{e_{ref}}$ which map physical element to reference element.

In general, the bilinear shape functions on reference element are obtained by the tensor product of 1D linear Lagrange polynomials[3]. So, the local basis for $Q_{_1}(\Omega_{e_k})$ is defined as

$$Q_{_1}(\Omega_{e_k}) := \left\{p \circ \psi_{_\Omega}^{-1} : p \in \text{span } \left\{1, \xi_x, \xi_y, \xi_x\xi_y\right\}\right\}, \tag{3.30}$$

and

$$\begin{aligned} Q_{_1}(\Omega_{e_k}) \quad = \quad \text{span} \Bigg\{ & \left(\frac{1}{4}(1-\xi_x)(1-\xi_y)\right), \left(\frac{1}{4}(1-\xi_x)(1+\xi_y)\right) \\ & \left(\frac{1}{4}(1+\xi_x)(1-\xi_y)\right), \left(\frac{1}{4}(1+\xi_x)(1+\xi_y)\right)\Bigg\}, \end{aligned} \tag{3.31}$$

with four degrees of freedom at corners of the quadrilateral. For more details cf. [35].

$Q_{_2}$- Finite element

The $Q_{_2}$-Element is also known as biquadratic element which consists of nine nodes or degrees of freedom of a quadrilateral, specifically with four corners or vertices, four mid-points of edges and one center point see Figure 3.2. In general, the biquadratic shape functions on reference element are obtained by the tensor product of 1D quadratic

[3]The linear Lagrange polynomials in 1D are defined by [73] as

$$p_1 = \frac{1-\xi}{2},\ p_2 = \frac{1+\xi}{2},\ -1 \leq \xi \leq 1.$$

Lagrange polynomials[4]. So, the local basis for $Q_2(\Omega_{e_k})$ is defined as

$$Q_2(\Omega_{e_k}) := \left\{p \circ \psi_\Omega^{-1} : p \in \text{span } \left\{1, \xi_x, \xi_y, \xi_x\xi_y, \xi_x^2, \xi_y^2, \xi_x^2\xi_y, \xi_x\xi_y^2, \xi_x^2\xi_y^2\right\}\right\}, \tag{3.32}$$

and

$$\begin{aligned} Q_2(\Omega_{e_k}) \;=\; \text{span}\Bigg\{ & \left(\left(1-\xi_x^2\right)\left(1-\xi_y^2\right)\right), \left(-\frac{1}{2}\left(1-\xi_x^2\right)\left(\xi_y-\xi_y^2\right)\right), \\ & \left(\frac{1}{2}\left(1-\xi_x^2\right)\left(\xi_y+\xi_y^2\right)\right), \left(-\frac{1}{2}\left(\xi_x-\xi_x^2\right)\left(1-\xi_y^2\right)\right), \\ & \left(\frac{1}{4}\left(\xi_x-\xi_x^2\right)\left(\xi_y-\xi_y^2\right)\right), \left(-\frac{1}{4}\left(\xi_x-\xi_x^2\right)\left(\xi_y+\xi_y^2\right)\right), \\ & \left(\frac{1}{2}\left(\xi_x+\xi_x^2\right)\left(1-\xi_y^2\right)\right), \left(-\frac{1}{4}\left(\xi_x+\xi_x^2\right)\left(\xi_y-\xi_y^2\right)\right), \\ & \left(\frac{1}{4}\left(\xi_x+\xi_x^2\right)\left(\xi_y+\xi_y^2\right)\right)\Bigg\}, \end{aligned} \tag{3.33}$$

with nine degrees of freedom of the quadrilateral. For more details cf. [35].

P_1^{dc}-Finite element

The P_1^{dc}-Element is a discontinuous bilinear element which consists of one node with the function value and both partial derivatives as its three local degrees of freedom located at the center of the quadrilateral, see Figure 3.2. The basis function for this element is linear (discontinuous) polynomial on the reference element and the local basis for $P_1^{dc}(\Omega_{e_k})$ is defined as following

$$P_1^{dc}(\Omega_{e_k}) := \left\{p \circ \psi_\Omega^{-1} : p \in \text{span } \left\{1, \xi_x, \xi_y\right\}\right\}, \tag{3.34}$$

and on each element the space $P_1^{dc}(\Omega_{e_k})$ considered in [4], [83] is as following

$$P_1^{dc}(\Omega_{e_k}) = \text{span}\left\{1, \xi_x, \xi_y\right\}. \tag{3.35}$$

3.4 Geometry of computational domain

In this section, we discuss the geometry of computational domain which is flow around cylinder in our case. Also a couple of mesh variations are explained as well.

[4]The quadratic Lagrange polynomials in 1D, for $-1 \le \xi \le 1$ we have

$$p_1 = \frac{-\xi(1-\xi)}{2},\ p_2 = \frac{\xi(1+\xi)}{2},\ p_3 = 1-\xi^2.$$

3.4.1 Flow around cylinder

The benchmark "flow around cylinder (FAC)" is established on the research outcome provided in [86] and [91]. In two dimensional case, the geometry of domain Ω is defined as a rectangular channel

$$\Omega := \{(x, y) : 0 \leq x \leq 2.2,\ 0 \leq y \leq 0.41\} \setminus B_r\,(c)\,, \tag{3.36}$$

where $c = (0.2, 0.2)$ is the center and $r := \frac{D_B}{2} = \frac{0.1}{2} = 0.05\ m$ refer to the radius of a circular cylinder $C_r\,(c)$ excluded from domain as shown in Figure 3.3. The incompress-

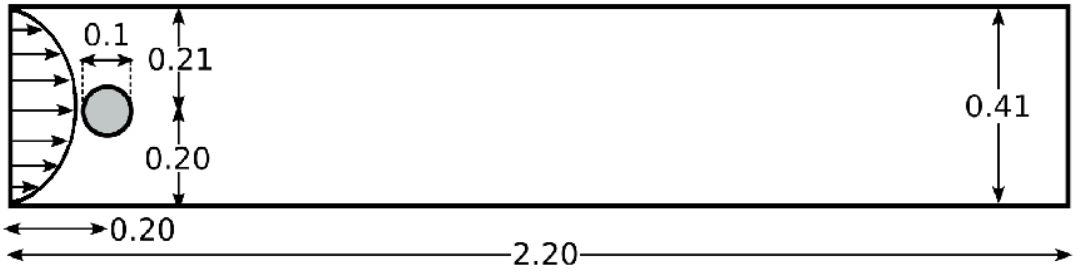

Figure 3.3: Geometry of flow arround cylinder problem in 2D

ible Newtonian fluids can be used to describe the domain for more details. Consider the fluid with velocity $\mathbf{u}$= $(u\,(x, y)\,, v\,(x, y))$ is flowing through FAC domain. The kinematic viscosity is taken as $\eta_{_\nu} = 0.001\ m^2/s$ and the density as $\rho = 1.0\ kg/m^3$. When the fluid with maximum velocity $\mathbf{u}_{\max}$ of 0.3 m/s passes through inlet of rectangular channel then the characteristic or mean velocity $\mathbf{u}_{mean}$ of parabolic flow profile can be obtained as

$$\mathbf{u}_{mean} = \frac{2}{3}\mathbf{u}_{\max} = \frac{2}{3}\,(0.3) = 0.2\ m/s.$$

The Reynolds number **Re** correlated to stationary laminar flow is calculated as

$$\mathbf{Re} = \frac{\mathbf{u}_{mean} D_{_B}}{\eta_{_\nu}} = \frac{(0.2)\,(0.1)}{0.001} = 20.$$

Boundary conditions

The boundary conditions associated with domain are described as following:

- The boundary condition on lower wall $\hat{\Omega}_{E_2} := \{0 \leq x \leq 2.2,\ y = 0\}$, upper wall $\hat{\Omega}_{E_4} := \{0 \leq x \leq 2.2,\ y = 0.41\}$ and circular cylinder $\hat{\Omega}_{\partial B_r}$ with center at $(0.2, 0.2)$ is defined as no-slip boundary such that $u = v = 0$.
- The inflow boundary condition on left wall or at inlet

$$\hat{\Omega}_{E_1} := \{x = 0,\ 0 \leq y \leq 0.41\}\,,$$

of domain is exercised as a parabolic profile and given as

$$\begin{aligned} u\,(0,y) &= \left(\frac{4\mathbf{u}_{\max} y\,(H_y - y)}{(H_y)^2}\right), \\ v\,(0,y) &= 0, \end{aligned}$$

where H_y is the total width, which is 0.41 m of rectangular domain along y-axis.

- The outflow boundary condition on right wall

$$\hat{\Omega}_{E_3} := \{x = 2.2,\ 0 \leq y \leq 0.41\},$$

of domain is described as stress-free or do-nothing condition.

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \left(\eta_{\nu} \nabla \mathfrak{u} - \mathfrak{pl}\right) \cdot \mathbf{n}, \tag{3.37}$$

where $\mathbf{n}$ is the outward unit vector normal to the outflow boundary.

The quadrilateral finite elements of different types, as discussed in previous section 3.3.2, are employed to discretize the domain. A couple of computational meshes are used for the analysis purpose and also discussed briefly as under.

3.4.2 Computational mesh configuration-I

The first FAC mesh configuration consists of 130 quadrilateral elements at coarse level one as shown in the Figure 3.4. The fine meshes are generated by linking the mid segment points of the opposite line segments of each element of the immediate coarse level mesh which is also called mesh refinement.

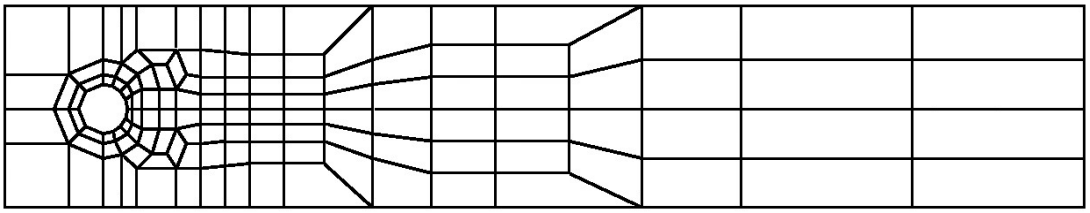

Figure 3.4: The computational mesh-I at coarse grid level

The hierarchy of uniform mesh refinements is presented in the Table 3.1.

Level	NEL	NVT	NMP
1	130	156	286
2	520	572	1,092
3	2,080	2,184	4,264
4	8,320	8,528	16,848
5	33,280	33,696	66,976
6	1,33,120	1,33,952	2,67,072

Table 3.1: Computational mesh-I information

where NEL is the number of total elements, NVT is the number of vertices and NMP is the number of midpoints for respective coarse (top) to finer (bottom) mesh levels.

3.4.3 Computational mesh configuration-II

The second FAC mesh configuration consists of 346 quadrilateral elements at coarse level one as shown in Figure 3.5.

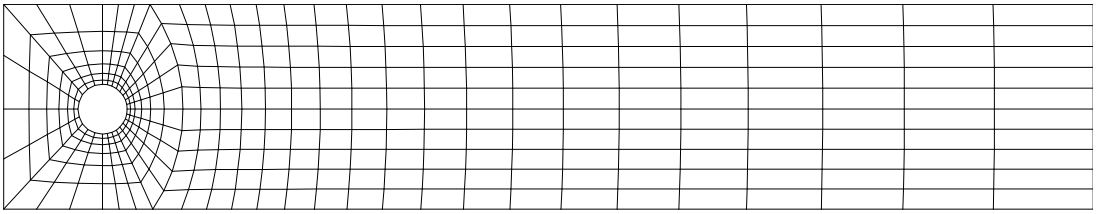

Figure 3.5: The computational mesh-II at coarse grid level

The hierarchy of fine meshes generated by uniform mesh refinements is presented in the Table 3.2.

Level	NEL	NVT	NMP
1	346	391	737
2	1, 384	1, 474	2, 858
3	5, 536	5, 716	11, 252
4	22, 144	22, 504	44, 648
5	88, 576	89, 296	1, 77, 872
6	3, 54, 304	3, 55, 744	7, 10, 048

Table 3.2: Computational mesh-II information

where NEL, NVT and NMP are the number of elements, vertices and midpoints for different refinement levels respectively.

3.4.4 Benchmark values for numerical comparisons

The performance of LSFEM methods is examined by computing some well-known quantities across the surface S_c of the circular cylinder, named as the lift $C_{_L}$, drag $C_{_D}$ coefficients and the pressure drop $\triangle \mathfrak{p}$. The comprehensive detail of these quantities can be seen in [93]. The SVP first order formulation is employed throughout our work. Therefore, we compute the stress $\boldsymbol{\sigma}$ as a separate variable and the body forces can be computed directly as

$$(F_{_D}, F_{_L})^T = \int_{S_c} \boldsymbol{\sigma}\mathbf{n}\, dS_c, \tag{3.38}$$

where stress tensor $\boldsymbol{\sigma} = \nu\nabla\mathfrak{u} - \mathfrak{p}\mathbf{I}$ and $\mathbf{n}$ is the normal vector of circle's surface S_c. By simple maths one can obtain lift and drag coefficients for $\mathbf{Re} = 20$, as following

$$\begin{aligned} C_{_D} &= \frac{2}{\mathbf{u}^2_{mean} D_{_{B_r}}} F_{_D} = 500 F_{_D}, \\ C_{_L} &= \frac{2}{\mathbf{u}^2_{mean} D_{_{B_r}}} F_{_L} = 500 F_{_L}. \end{aligned} \tag{3.39}$$

3.4.5 Global mass conservation in LSFEM

The lack of mass conservation generally presents in all finite element methods for the incompressible Stokes and Navier-Stokes problems. But in case of LSFEM, the mass conservation is not preserved globally. Because, the continuity equation is computed as another component of minimized energy functional. In case, if another equation of the system have more influence than the continuity equation in the functional. Then, this leads to loss of mass in the system. As, the continuity equation is one which established that the velocity field is divergence-free. The global mass conservation was investigated in [29] for LSFEM. He observed the mass loss in narrow channels when flow passed around the circular cylinder. Later, the global mass conservation was explored in [42] and introduced some weights with continuity equation component in the functional to improve mass conservation.

We study the traditional mass conservation problems of the LSFEM through the numerical investigations of the flow around cylinder problem. The lack of global mass conservation is one of the significant issues for least-squares finite element method. Therefore, we calculate the global mass conservation (m^{GMC}) in terms of the fractional change of mass flow rate, defined as

$$m^{GMC} = \frac{\int_{\hat{\Omega}_i} \rho(\mathbf{n} \cdot \mathfrak{u}) d\hat{\Omega}_i - \int_{\hat{\Omega}_0} \rho(\mathbf{n} \cdot \mathfrak{u}) d\hat{\Omega}_o}{\int_{\hat{\Omega}_i} \rho(\mathbf{n} \cdot \mathfrak{u}) d\hat{\Omega}_i} \times 100, \tag{3.40}$$

where $\hat{\Omega}_i$ is the inflow boundary of the domain and $\hat{\Omega}_o$ is any vertical section between the inflow and the outflow boundaries, including the outflow, $\mathbf{n}$ is normal vector and $\mathfrak{u}$ represents the velocity.

3.5 Newtonian and non-Newtonian fluids

In this work, we study two prominent types of fluids, namely Newtonian and non-Newtonian fluids. The details are given briefly as under.

If a real fluid which obeys the Newton's law of viscosity, i.e. the shear stress τ is directly proportional to the strain rate $\mathsf{D}(\mathfrak{u})$, then it is known as the Newtonian fluid.

$$\tau = \eta_\nu \mathsf{D}(\mathfrak{u}), \tag{3.41}$$

where η is the constant viscosity of fluid and $\mathsf{D}(\mathfrak{u}) = \frac{1}{2}\left(\nabla\mathfrak{u} + \nabla\mathfrak{u}^T\right)$. It is therefore very clear that viscosity is independent of strain rate.

On the other hand, the fluid that does not follow Newton's law of fluid is called the non-Newtonian fluid, i.e. equation (3.41) does not hold and $\tau \neq \eta_\nu \mathsf{D}(\mathfrak{u})$. But, if a fluid holds equation (3.41) and viscosity η_ν depends on strain rate $\mathsf{D}(\mathfrak{u})$ then it is called generalized Newtonian fluid.

For the incompressible fluids, the second invariant or shear rate $\dot{\gamma}_{_{II}}$ is defined as $\dot{\gamma}_{_{II}} = \sqrt{\mathrm{tr}\mathsf{D}^2} = \sqrt{\mathsf{D} : \mathsf{D}}$. Hence, the constitutive equation for generalized Newtonian fluid can be given as

$$\tau = \eta_\nu\left(\dot{\gamma}_{_{II}}\right)\mathsf{D}(\mathfrak{u}). \tag{3.42}$$

Fluid flow models

There are many fluid models, for nonlinear viscosity $\eta_\nu\left(\dot{\gamma}_{_{II}}\right)$, have been given in literature. In our work, we employ a couple of models for analysis purposes. A brief discussion of these models is provided below.

Power law fluid model The power law fluid model is one of the common choices to investigate some important properties of non-Newtonian fluids. It can be defined as

$$\eta_\nu(\dot{\gamma}_{_{II}}) = \eta_0\left(\epsilon + \dot{\gamma}_{_{II}}^2\right)^{\frac{r}{2}-1}, \tag{3.43}$$

where r is the power law index. The values of index r are very significant and used to represent different types of fluids:

$\frac{r}{2} < 1$: Shear thinning fluids.

$\frac{r}{2} = 1$: Newtonian fluids (Constant viscosity).

$\frac{r}{2} > 1$: Shear thickening fluids.

In this model, the viscosity may be unbounded during the numerical computation such that it no longer has a physical meaning. Therefore, we consider another model for bounded viscosity.

Cross law fluid model This model bounds the viscosity between η_1 and η_0 and defined as

$$\eta_\nu(\dot{\gamma}_{_{II}}, \mathfrak{p}, \boldsymbol{\theta}) = \eta_0 + \frac{\eta_1 - \eta_0}{\left(1 + \left(\lambda\dot{\gamma}_{_{II}}\right)^{r_1}\right)^r}\exp\left(\alpha_{_p}\mathfrak{p}\right)\ e^{a_1 + \frac{a_2}{a_3 + \boldsymbol{\theta}}}, \tag{3.44}$$

where η_0 is the minimum shear viscosity, $\boldsymbol{\theta}$ is the temperature, r is the Power law index to generalize the model to cover both shear-thickening and shear-thinning. r_1 is dimensionless and known as Cross Rate Constant. It is used to measures the degree of the dependency of viscosity on the shear rate in the shear-thickening or shear-thinning region. λ is the Cross Time Constant. Here, we set the values to $r_1 = 1$ and $\lambda = 8.2$.

Chapter 4

Newtonian Fluid Flows

The stationary incompressible Navier-Stokes equations are solved by least-squares finite element method based on first-order formulations, which are the outcome of suitable substitutions of some physical quantities in the governing equations. These physical quantities represent different characteristics of fluid flow. In this work, we build the first-order formulation by substitution of physical fields present in the fluid flow. In the upcoming sections, we present a brief introduction and discuss the problem formulation for Navier-Stokes equations in details. The numerical validation of proposed method is given for linear Stokes problem. Afterwards, the nonlinear problem is considered for further analysis by using quadrilateral finite elements of different orders.

4.1 Introduction

The least-squares finite element method is a true variational method used to solve second order elliptic partial differential equations. But, the straight forward application of least-squares principle to second order PDEs gives rise to some practically issues. To avoid these problems, the second order PDEs can be transformed into first order system of equations by substituting the new auxiliary unknown variables i.e. vorticity, velocity gradient, stresses etc. Therefore, we use stress as new auxiliary variable and the formulation based on such variable is called stress-velocity-pressure or SVP formulation. The least-squares principle on SVP formulation always leads to symmetric positive definite system [15]. This allows us to employ a robust Krylov subspace solver named conjugate gradient method. We use the combination of conjugate gradient and multigrid method for the solution of discrete linear system. The conjugate gradient method was preconditioned by multigrid solver [57], as the Krylov method uses eigenmode for error reduction. Moreover, the preconditioned conjugate gradient method is used as smoother of multigrid solver, as the size of solution updates is suitably determine by Krylov methods at each smoothing step [98]. Also, it led to parameter-free smoothing sweeps. The algebraic MPCG was studied for the solutions

of Stokes problem [57] and NS problem [58], [59], [60] using LSFEM. The geometric MPCG was employed in [82] for solution of NS problem using Spectral/hp LSFEM and established fast convergence than Jacobi PCG solver. We employed the multigrid preconditioned conjugate gradient solver to the discrete linear system based on SVP formulation. Furthermore, the SSOR preconditioned conjugate gradient solver is used as smoother for multigrid method to get better and parameter-free smoother.

We conduct various numerical tests for the variety of parametric values to analyze the efficiency of MPCG method. The flow around cylinder domain is employed for a couple of mesh configurations and investigated for the solutions of NS problem with lower and higher order finite element discretizations. The studies [29], [79], [88] on LSFEM had shown poor performance and drawbacks for lower order finite elements. The use of higher order finite elements is a remedy to such problem as illustrated in [79] and [81]. Therefore, we use both types of finite elements for the numerical investigations. The lack of mass conservation is another topic of discussion in the study of LSFEM. To overcome this deficiency, numerous research has been conducted in [21], [22], [23], [42], [80] and many others. We use the weighted function depending on the viscosity of fluid to get overall optimal results. The results produced by our weighted LSFEM are very accurate and robust. The grid independent behavior is shown by the solver for flow around cylinder domain.

4.2 LSFEM for Stokes equations

In the following section, we demonstrate the numerical validity of our weighted least-squares finite element method. The results are presented for the Stokes problem (2.85) with SVP formulation. The investigation is made with $Q_1Q_1Q_1$, $Q_2P_1^{dc}Q_2$ and $Q_2Q_2Q_2$ finite element discretizations for two mesh configurations (3.4.2) and (3.4.3) of flow around cylinder problem.

4.2.1 Numerical validation

At lev. 6	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Mesh Configuration $-$ *I*				
$Q_1Q_1Q_1$	$3.158050E+00$	$3.076306E-02$	$4.453144E-02$	$1.192327E-01$
$Q_2P_1^{dc}Q_2$	$3.142273E+00$	$3.019508E-02$	$4.556511E-02$	$9.569492E-05$
$Q_2Q_2Q_2$	$3.142273E+00$	$3.019508E-02$	$4.556508E-02$	$9.563460E-05$
Mesh Configuration $-$ *II*				
$Q_1Q_1Q_1$	$3.153172E+00$	$2.916881E-02$	$4.535571E-02$	$4.318738E-02$
$Q_2P_1^{dc}Q_2$	$3.142408E+00$	$3.019559E-02$	$4.557725E-02$	$9.904366E-06$
$Q_2Q_2Q_2$	$3.142408E+00$	$3.019559E-02$	$4.557659E-02$	$9.902297E-06$
Ref [37] : *Drag* $C_D = 3.142292E+00$, *Lift* $C_L = 3.019366e-02$				

Table 4.1: Numerical validation of the proposed solver

The outcomes presented in Table 4.1 are very accurate for the higher order finite elements in comparison to the lower order finite elements and got on mesh refinement level 6. The results calculated by configuration-I (3.4.2) are more closer to reference solution [37] but the global mass conservation is more promising for mesh configuration-II (3.4.3) .

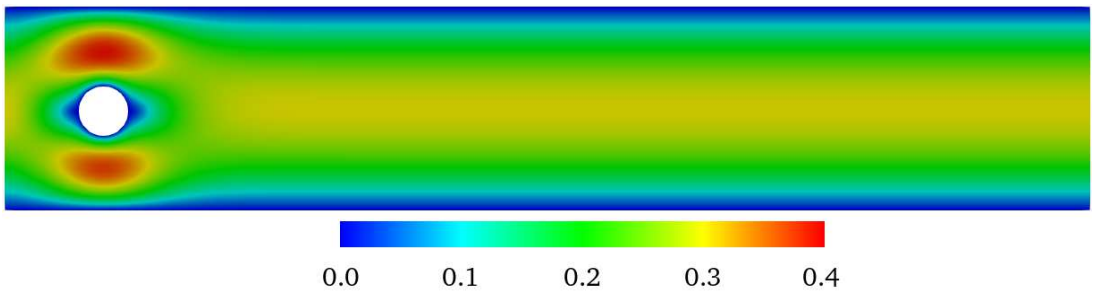

Figure 4.1: Velocity profile for Stokes problem

4.3 LSFEM for the Navier-Stokes equations

The most of the work in this section is from our research paper cited as [77]. In addition to that we present approximate solutions of Navier-Stokes equations for the lower and higher order finite elements for different nonlinear schemes.

4.3.1 Governing equations

The incompressible Navier-Stokes (NS) equations on a bounded domain $\Omega \subset \mathbb{R}^2$ for a stationary flow are given by

$$\left.\begin{aligned} \mathfrak{u} \cdot \nabla \mathfrak{u} + \nabla \mathfrak{p} - \eta_{\nu} \Delta \mathfrak{u} &= \mathfrak{f} \\ \nabla \cdot \mathfrak{u} &= 0 \end{aligned}\right\} \text{ in } \Omega, \tag{4.1}$$

with the boundary conditions

$$\left.\begin{aligned} \mathfrak{u} &= f_{\hat{\Omega}_D} \\ \mathbf{n} \cdot \left(\eta_{\nu} \nabla \mathfrak{u} - \mathfrak{p}\mathfrak{l}\right) &= f_{\hat{\Omega}_N} \end{aligned}\right\} \text{ on } \hat{\Omega}, \tag{4.2}$$

and along with the zero mean pressure constraint as given by (2.88). The notation used for the system is described below as

$\mathfrak{u}\,(u, v)$	: is the velocity of the fluid passing through domain Ω,
$\mathfrak{p} = \frac{P}{\rho}$	: stands for the normalize pressure term,
$\eta_{\nu} = \frac{\mu}{\rho}$	: expresses the kinematic viscosity of fluid,
$\mathfrak{f}$	: act as the external source term,
$f_{\hat{\Omega}_D}$	: be the value of Dirichlet boundary conditions on the boundary region $\hat{\Omega}_D$,
$f_{\hat{\Omega}_N}$	: represents the value of Neumann boundary conditions on the boundary region $\hat{\Omega}_N$,
$\mathbf{n}$	: represents the outward unit normal on the boundary regions $\hat{\Omega} = \partial\Omega = \hat{\Omega}_D \cup \hat{\Omega}_N$ and $\hat{\Omega}_D \cap \hat{\Omega}_N = \emptyset$.

The kinematic viscosity and the density of the fluid are assumed to be constant. In system (4.1) the first equation represents the momentum equation where velocity $\mathfrak{u} = [u, v]^T$ has two components, which are further functions of x, y and pressure $\mathfrak{p}$ are the unknowns to be calculated. Whereas, the second equation represents the divergence of velocity field of fluid flow in the domain and also known as continuity equation.

4.4 Stress-based first order system

As it is discussed previously that the straightforward application of least-squares principle on the second-order equations require C^1 finite elements for approximation [15]. The practical implementation of such finite elements is an arduous task. To overcome these difficulties, we first reduce the second-order equations into first-order system of equations. Another consequences of straightforward application of least-squares principle to the second-order Navier-Stokes equation is that it results into ill-conditioned

system matrices [19]. Therefore the transformation of second-order system into the first-order system is an essential requirement to develop a practical scheme.

To construct the stress based formulation also named Stress-Velocity-Pressure (SVP) formulation, we introduce a physical quantity called Cauchy stress tensor $\boldsymbol{\sigma}$ as a new variable defined as

$$\begin{aligned} \boldsymbol{\sigma} &= \mathbf{v}_{\boldsymbol{\sigma}} - \mathbf{p}_{\boldsymbol{\sigma}}, \\ &= 2\eta_{\nu}\mathsf{D}(\mathfrak{u}) - \mathfrak{p}\mathsf{I}, \end{aligned} \tag{4.3}$$

where $\mathbf{p}_{\boldsymbol{\sigma}} = \mathfrak{p}\mathsf{I}$ is the inviscid reactive component and $\mathbf{v}_{\boldsymbol{\sigma}} = 2\eta_{\nu}\mathsf{D}(\mathfrak{u})$ is the active viscous component of the Cauchy stress tensor $\boldsymbol{\sigma}$, which is the symmetric part of the gradient of velocity $\mathsf{D}(\mathfrak{u}) = \frac{1}{2}\left(\nabla\mathfrak{u} + \nabla\mathfrak{u}^T\right)$. Now by using the Navier Stokes equations (4.1) and the Cauchy stress equation (4.3), we obtain the following first-order SVP system of equations

$$\left.\begin{aligned} \mathfrak{u}\cdot\nabla\mathfrak{u} - \nabla\cdot\boldsymbol{\sigma} &= \mathfrak{f} \\ \nabla\cdot\mathfrak{u} &= 0 \\ \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}} - \mathbf{v}_{\boldsymbol{\sigma}} &= 0 \end{aligned}\right\} \text{ in } \Omega, \tag{4.4}$$

along with general homogeneous boundary conditions

$$\left.\begin{aligned} \mathfrak{u} &= 0 \\ \mathbf{n}\cdot\boldsymbol{\sigma} &= 0 \end{aligned}\right\} \text{ on } \hat{\Omega}. \tag{4.5}$$

4.5 Linearization of nonlinear terms

Due to the presence of the convective term $\mathfrak{u}\cdot\nabla\mathfrak{u}$ in the momentum equation, the system (4.4) is nonlinear. Now before the application of least-squares principle the Newton's linearization is applied to the nonlinear convection term $\mathfrak{u}\cdot\nabla\mathfrak{u}$ such that

$$\mathfrak{u}\cdot\nabla\mathfrak{u} \approx \mathfrak{u}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u}.$$

Let the sum of residuals for the system of equations (4.4) be denoted by $\mathcal{R}$. To approximate the nonlinear residuals, the nonlinear iteration is updated with the correction $\mathbf{u}^{k+1} = \mathbf{u}^k + \delta\mathbf{u}$. Then, the residual approximation given by Taylor series expansion as

$$\begin{aligned} \mathcal{R}\big(\mathbf{u}^{k+1}\big) &= \mathcal{R}\big(\mathbf{u}^k + \delta\mathbf{u}\big), \\ &\simeq \mathcal{R}\big(\mathbf{u}^k\big) + \big[\mathcal{D}\mathcal{R}\big(\mathbf{u}^k\big)\big]\delta\mathbf{u}, \end{aligned} \tag{4.6}$$

where

$$\mathcal{R}(\mathbf{u}^k) = \begin{bmatrix} \mathcal{R}_1^k \\ \mathcal{R}_2^k \\ \mathcal{R}_3^k \end{bmatrix} = \begin{bmatrix} \mathfrak{u}^k \cdot \nabla \mathfrak{u}^k - \nabla \cdot \boldsymbol{\sigma}^k - \mathfrak{f} \\ \nabla \cdot \mathfrak{u}^k \\ \boldsymbol{\sigma}^k + \mathbf{p}_{\boldsymbol{\sigma}}(\mathbf{u}^k) - \mathbf{v}_{\boldsymbol{\sigma}}(\mathbf{u}^k) \end{bmatrix}, \tag{4.7}$$

and the system of equations after substitution, $\delta\mathbf{u} = (\mathfrak{u}, \mathfrak{p}, \boldsymbol{\sigma})$ is as following

$$\mathcal{DR}(\mathbf{u}^k)\delta\mathbf{u} = \begin{bmatrix} \mathfrak{u} \cdot \nabla \mathfrak{u}^k + \mathfrak{u}^k \cdot \nabla \mathfrak{u} - \nabla \cdot \boldsymbol{\sigma} - \mathfrak{f} \\ \nabla \cdot \mathfrak{u} \\ \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}} - \mathbf{v}_{\boldsymbol{\sigma}} \end{bmatrix}. \tag{4.8}$$

By following the procedure discussed in Algorithm 3.2, i.e. the nonlinear basic iteration for least-squares scheme, one can obtain a solution update $\delta\mathbf{u}$ as

$$\tilde{\mathcal{R}}(\mathbf{u}^k)\delta\mathbf{u} = \mathcal{R}_{def}(\mathbf{u}^k), \tag{4.9}$$

- For the Gauss-Newton method, the values of $\tilde{\mathcal{R}}$ and $\mathcal{R}_{def}$ are consider to be

$$\tilde{\mathcal{R}}(\mathbf{u}^k) = \mathcal{DR}(\mathbf{u}^k)^* \mathcal{DR}(\mathbf{u}^k), \tag{4.10}$$

  and

$$\mathcal{R}_{def}(\mathbf{u}^k) = -\mathcal{R}(\mathbf{u}^k)^* \mathcal{DR}(\mathbf{u}^k). \tag{4.11}$$

- For the fixed point method, the value of $\mathcal{DR}$ is simply replaced by $\mathcal{R}$ in above equations (4.10) and (4.11).

4.6 Continuous least-squares principle

We introduce the spaces $\mathcal{X}_\Omega$ of admissible functions based on the residuals of the first-order systems (4.4)

$$\mathcal{X}_\Omega := \mathcal{H}^1_{0,D}(\Omega) \times \mathcal{L}^2_0(\Omega) \times \mathcal{H}_{0,N}(\mathrm{div}, \Omega). \tag{4.12}$$

The quadratic linearized functional $\mathcal{E}_{svp}$ in term of the $\mathcal{L}^2$-norm is obtained by using least-squares principles on linearized residual (4.6) as

$$\mathcal{E}_{svp}(\delta\mathbf{u};0) = \frac{1}{2}\left\|\mathcal{R}(\mathbf{u}^k) + \left[\mathcal{DR}(\mathbf{u}^k)\right]\delta\mathbf{u}\right\|_0^2, \tag{4.13}$$

and the $\mathcal{L}^2$-norm based S-V-P energy functionals can be expanded as follows

$$\begin{aligned} \mathcal{E}_{svp}(\delta\mathbf{u};0) = \; & \tfrac{1}{2}\big(\left\|\mathcal{R}_1^k + \mathfrak{u} \cdot \nabla \mathfrak{u}^k + \mathfrak{u}^k \cdot \nabla \mathfrak{u} - \nabla \cdot \boldsymbol{\sigma}\right\|_0^2 + \left\|\mathcal{R}_2^k + \nabla \cdot \mathfrak{u}\right\|_0^2 \\ & + \left\|\mathcal{R}_3^k + \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}} - \mathbf{v}_{\boldsymbol{\sigma}}\right\|_0^2\big) \;\forall\; (\mathfrak{u}, \mathfrak{p}, \boldsymbol{\sigma}) \in \mathcal{X}_\Omega, \end{aligned} \tag{4.14}$$

and weighted functional associated with the weighted function is

$$\begin{aligned}\mathcal{E}_{svp}^{w_{\eta_\nu}}\big(\delta\mathbf{u};0\big) = \quad & \tfrac{1}{2}\big(\big\|\mathcal{R}_1^k + \mathfrak{u}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} - \nabla\cdot\boldsymbol{\sigma}\big\|_0^2 + \alpha_m\big\|\mathcal{R}_2^k + \nabla\cdot\mathfrak{u}\big\|_0^2 \\ & + w_{\eta_\nu}^n\big\|\mathcal{R}_3^k + \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}} - \mathbf{v}_{\boldsymbol{\sigma}}\big\|_0^2\big)\ \forall\ \big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_\Omega,\end{aligned} \tag{4.15}$$

where $w_{\eta_\nu}^n$ is the weighted function multiplied only with residual part of additional Cauchy stress variable. Here, to improve mass conservation, a scaling parameter α_m is used for the least-squares formulation [42], [23], [24].

The minimization problem associated with the least-squares formulation in (4.15) is to find $\delta\mathbf{u} = \big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big) \in \mathcal{X}_\Omega$ such that

$$\delta\mathbf{u} = \underset{\mathbf{v}\in\mathcal{X}_\Omega}{\arg\min}\mathcal{E}_{svp}^{w_{\eta_\nu}}\big(\mathbf{v};0\big). \tag{4.16}$$

The first variant of the quadratic linearized functional (4.16) leads to an equivalent problem of finding $\delta\mathbf{u}$ such that

$$\big[\mathcal{R}\big(\mathbf{u}^k\big) + \mathcal{DR}\big(\mathbf{u}^k\big)\delta\mathbf{u}\big]\mathcal{DR}\big(\mathbf{u}^k\big)\mathbf{v} = 0,\quad \forall\ \mathbf{v}\in\mathcal{X}_\Omega. \tag{4.17}$$

The variational problem based on the optimality condition (4.17) of the minimization problem (4.16) is given by

$$\begin{cases}\text{Find }\ \delta\mathbf{u} = \big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_\Omega \quad \text{such that,}\\ \mathbf{K}\big(\delta\mathbf{u},\mathbf{v}\big) = \mathbf{F}\big(\mathbf{v}\big)\ \forall\ \mathbf{v} = \big(\mathfrak{v},\mathfrak{q},\boldsymbol{\tau}\big)\in\mathcal{X}_\Omega.\end{cases} \tag{4.18}$$

where $\mathbf{K}$ is the bilinear form defined on $\mathcal{X}_\Omega\times\mathcal{X}_\Omega\to\mathbb{R}$ as follows

$$\begin{aligned}\mathbf{K}\left(\delta\mathbf{u},\mathbf{v}\right) &= \big\langle\mathcal{DR}\big(\mathbf{u}^k\big)\delta\mathbf{u},\mathcal{DR}\big(\mathbf{u}^k\big)\mathbf{v}\big\rangle \\ &= \big\langle\big(\mathfrak{u}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} - \nabla\cdot\boldsymbol{\sigma}\big),\big(\mathfrak{v}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{v} - \nabla\cdot\boldsymbol{\tau}\big)\big\rangle \\ &+ w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big) - \mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big),\boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big) - \mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle \\ &+ \alpha_m\big\langle\big(\nabla\cdot\mathfrak{u}\big),\big(\nabla\cdot\mathfrak{v}\big)\big\rangle,\end{aligned} \tag{4.19}$$

and the linear form $\mathbf{F}$ is defined on $\mathcal{X}_\Omega\to\mathbb{R}$ as follows

$$\begin{aligned}\mathbf{F}(\mathbf{v}) &= -\big\langle\mathcal{R}\big(\mathbf{u}^k\big),\mathcal{DR}\big(\mathbf{u}^k\big)\mathbf{v}\big\rangle \\ &= -\big[\big\langle\big(\mathfrak{u}^k\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\sigma}^k\big),\big(\mathfrak{v}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{v} - \nabla\cdot\boldsymbol{\tau}\big)\big\rangle \\ &+ w_{\eta_\nu}^n\big\langle\big(\boldsymbol{\sigma}^k + \mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}^k\big) - \mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}^k\big)\big),\big(\boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big) - \mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle \\ &+ \alpha_m\big\langle\big(\nabla\cdot\mathfrak{u}^k\big),\big(\nabla\cdot\mathfrak{v}\big)\big\rangle\big].\end{aligned} \tag{4.20}$$

For detailed analysis of the least-square problem, one can build the coefficient matrix $\mathbf{K}$ as follows

$$\mathbf{K}\left(\delta\mathbf{u},\mathbf{v}\right)=\begin{bmatrix}\mathbf{K}_{vu} & \mathbf{K}_{vp} & \mathbf{K}_{v\boldsymbol{\sigma}}\\ \mathbf{K}_{qu} & \mathbf{K}_{qp} & \mathbf{K}_{q\boldsymbol{\sigma}}\\ \mathbf{K}_{\boldsymbol{\tau}u} & \mathbf{K}_{\boldsymbol{\tau}p} & \mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\sigma}}\end{bmatrix}, \tag{4.21}$$

where

$$\begin{aligned}
\mathbf{K}_{vu} &= \big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),-\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle+\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u},\mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle+\alpha_m\big\langle\nabla\cdot\mathfrak{u},\nabla\cdot\mathfrak{v}\big\rangle \\
&\quad+\boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k,\mathfrak{v}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle}+\boxed{\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u},\mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle}. \\
\mathbf{K}_{vp} &= w^{n}_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),-\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle, \\
\mathbf{K}_{v\sigma} &= w^{n}_{\eta_\nu}\big\langle \boldsymbol{\sigma},-\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle+\big\langle -\nabla\cdot\boldsymbol{\sigma},\mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle+\boxed{\big\langle -\nabla\cdot\boldsymbol{\sigma},\mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle}, \\
\mathbf{K}_{qu} &= w^{n}_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle, \\
\mathbf{K}_{qp} &= w^{n}_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle, \\
\mathbf{K}_{q\sigma} &= w^{n}_{\eta_\nu}\big\langle \boldsymbol{\sigma},\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle, \\
\mathbf{K}_{\tau u} &= w^{n}_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),\boldsymbol{\tau}\big\rangle+\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u},-\nabla\cdot\boldsymbol{\tau}\big\rangle+\boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k,-\nabla\cdot\boldsymbol{\tau}\big\rangle}, \\
\mathbf{K}_{\tau p} &= w^{n}_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),\boldsymbol{\tau}\big\rangle, \\
\mathbf{K}_{\tau\sigma} &= w^{n}_{\eta_\nu}\big\langle \boldsymbol{\sigma},\boldsymbol{\tau}\big\rangle+\big\langle -\nabla\cdot\boldsymbol{\sigma},-\nabla\cdot\boldsymbol{\tau}\big\rangle.
\end{aligned} \tag{4.22}$$

Here, the terms in the boxes have contributions in the coefficient matrix $\mathbf{K}$ due to the nonlinear terms. These terms will be add to coefficient matrix $\mathbf{K}$ after defect calculation. The resulting system matrix (4.19) is symmetric and positive definite. However, the stress-based SVP formulation leads to the system matrix (4.22) which is not differentially diagonal dominant. These properties are very significant when we design a multigrid solver for higher order finite elements, see [57] [59].

4.7 Discrete least-squares principle

The discrete form of the problem (4.18) with approximation spaces

$$\mathcal{X}^h_\Omega:=\Big\{\delta\mathbf{u}^h\in\mathcal{H}^{1,h}_{0,D}(\Omega)\times\mathcal{L}^2_0(\Omega)\times\mathcal{H}^h_{0,N}(\mathrm{div},\Omega)\Big\}, \tag{4.23}$$

where $\delta\mathbf{u}^h{=}\big(\mathfrak{u}_h,\mathfrak{p}_h,\boldsymbol{\sigma}_h\big)$ can be written as

$$\begin{cases}\text{Find }\ \delta\mathbf{u}^h\in\mathcal{X}^h_\Omega\quad\text{s.t.}\\ \mathbf{K}^h\big(\delta\mathbf{u}^h,\mathbf{v}^h\big)=\mathbf{F}^h\big(\mathbf{v}^h\big),\end{cases} \tag{4.24}$$

where $\mathbf{v}^h{=}\big(\mathfrak{v}_h,\mathfrak{q}_h,\boldsymbol{\tau}_h\big)\in\mathcal{X}^h_\Omega$ and $\mathbf{K}^h$ is the approximate bilinear form defined on $\mathcal{X}^h_\Omega\times\mathcal{X}^h_\Omega\to\mathbb{R}$.

We are able to use different combinations of finite element approximations since least-squares formulation allows a free choice of finite element spaces [19]. We restrict

our study here only to Q_1, Q_2 and P_1^{dc} elements. Our aim is to develop an efficient solver which takes full advantage of the sparse and symmetric positive definite systems. We use conjugate gradient method as a Krylov subspace solver and accelerate it with multigrid preconditioning. The SSOR-preconditioned conjugate gradient as smoother, with various range of smoothing steps, is used for MPCG. The direct Gaussian elimination (UMFPACK [39]) is employed as coarse-grid solver. In order to demonstrate the solver flexibility with respect to different flow problems, an 'F-cycle' of multigrid is used with various smoothing steps in the numerical tests. The Gauss-Newton and fixed point methods are considered as nonlinear solvers

4.8 Numerical results and analysis

In this section, we analyze the steady state flow around cylinder problem for two different configurations. The main aim of this analysis is to understand the importance of the choice of finite elements and the weighted functions to achieve the desired results. We use three distinct combinations of finite elements for velocity, pressure and stress variables. We employ both lower and higher order conforming finite elements for our numerical investigations. The Navier-Stokes equations are solved for Reynold number $\mathbf{Re} = 20$ with constant viscosity $\eta_\nu = 0.001$ and $\rho = 1.0$. The weighted functions used in least-squares functional (4.15) are given as

$$w_{\eta_\nu}^{n} = \big(1.0/\eta_\nu\big)^n; \quad n = 1.0 \;\text{ and }\; \alpha_m = 1.0$$

For the reference solutions, we use the well known benchmark results [91] and results from [37] as given below.

References:	$Drag\ C_{_D}$	$Lift\ C_{_L}$	Δp
Turek [91]	$5.579535E+00$	$1.061894E-02$	$1.175201E-01$
Damanik [37]	$5.579313E+00$	$1.061503E-02$	

4.8.1 Numerical results for mesh configuration-I

The numerical investigation in this section is motivated by the well known flow around cylinder benchmark [64], [91], [92], [93]. The computational grid (3.4.2) contains 130 quadrilateral elements on the coarse grid level and further finer levels are achieved by the multilevel grid refinements. In the refinement process, each element from the coarse grid or level one is divided into four fine elements by joining the mid-points of the opposite edges [67]. Then we use the hierarchy of the multilevel grids, as shown by Table 3.1, in our multigrid preconditioner. The inflow/outflow details can be found

in section 3.4.1.The bilinear finite element Q_1 is used for all unknowns i.e. velocity, pressure and stress.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	312	156	468	936
2	1144	572	1716	3432
3	4368	2184	6552	13104
4	17056	8528	25584	51168
5	67392	33696	101088	202176
6	267904	133952	401856	803712

Table 4.2: Total unknowns for $Q_1Q_1Q_1$-element in mesh-I

The total number of degrees of freedom computed on each level for the stress-velocity-pressure problem are presented by Table 4.2.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$4.708725E+00$	$7.654299E-02$	$8.102117E-02$	$2.833422E+00$
5	$5.258712E+00$	$4.092649E-02$	$9.982358E-02$	$8.433074E-01$
6	$5.480722E+00$	$2.119097E-02$	$1.097471E-01$	$2.267643E-01$
Fixed Point Method				
4	$4.789604E+00$	$4.644177E-02$	$8.295211E-02$	$2.387085E+00$
5	$5.298038E+00$	$2.177232E-02$	$1.010048E-01$	$6.962395E-01$
6	$5.495277E+00$	$1.312701E-02$	$1.102622E-01$	$1.858159E-01$

Table 4.3: Flow parameters for $Q_1Q_1Q_1$-element in mesh-I

Now, we use biquadratic finite elements Q_2 for our numerical investigation for mesh configuration 3.4.2. The Q_2 element is employed for all the unknown variables.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	1144	572	1716	3432
2	4368	2184	6552	13104
3	17056	8528	25584	51168
4	67392	33696	101088	202176
5	267904	133952	401856	803712
6	1068288	534144	1602432	3204864

Table 4.4: Total unknowns for $Q_2Q_2Q_2$-elements in mesh-I

The total number of unknowns on each level for the stress-velocity-pressure problem are given by Table 4.4.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.556446E+00$	$1.115572E-02$	$1.162045E-01$	$3.618464E-02$
5	$5.576758E+00$	$1.066844E-02$	$1.172607E-01$	$4.701878E-03$
6	$5.579248E+00$	$1.062216E-02$	$1.174601E-01$	$5.885212E-04$
Fixed Point Method				
4	$5.556770E+00$	$1.028484E-02$	$1.162086E-01$	$2.663399E-02$
5	$5.576749E+00$	$1.057478E-02$	$1.172595E-01$	$3.416974E-03$
6	$5.579242E+00$	$1.061386E-02$	$1.174598E-01$	$4.229598E-04$

Table 4.5: Flow parameters for $Q_2Q_2Q_2$-element in mesh-I

Next, the flow parameters are computed by employing Q_2-element for velocity and stress variables in the problem and the discontinuous pressure element P_1^{dc} is considered for pressure variable.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	1144	390	1716	3250
2	4368	1560	6552	12480
3	17056	6240	25584	48880
4	67392	24960	101088	193440
5	267904	99840	401856	769600
6	1068288	399360	1602432	3070080

Table 4.6: Total unknownss for $Q_2P_1^{dc}Q_2$-elements in mesh-I

The total number of unknowns on each level for the stress-velocity-pressure problem are presented by Table 4.6.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − NewtonMethod				
4	$5.556063E+00$	$1.103601E-02$	$1.163489E-01$	$3.695583E-02$
5	$5.576732E+00$	$1.066008E-02$	$1.172852E-01$	$4.754437E-03$
6	$5.579246E+00$	$1.062159E-02$	$1.174647E-01$	$5.920616E-04$
Fixed Point Method				
4	$5.556490E+00$	$1.020864E-02$	$1.163587E-01$	$2.710671E-02$
5	$5.576730E+00$	$1.056945E-02$	$1.172843E-01$	$3.448818E-03$
6	$5.579241E+00$	$1.061348E-02$	$1.174644E-01$	$4.251011E-04$

Table 4.7: Flow parameters for $Q_2P_1^{dc}Q_2$-element in mesh-I

The comparison of nonlinear (N) and linear (L) iterations for the solvers is presented in the Table 4.8 for different combinations of finite elements. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Level	$Q_1Q_1Q_1\,(N/L)$	$Q_2Q_2Q_2\,(N/L)$	$Q_2P_1^{dc}Q_2\,(N/L)$
Gauss − Newton Method			
2	7/3	9/8	9/12
3	7/7	5/12	5/13
4	7/10	4/17	4/17
5	5/14	4/22	4/22
Fixed Point Method			
2	8/2	14/7	14/10
3	9/5	10/9	10/10
4	11/7	9/11	9/10
5	11/10	8/21	8/18

Table 4.8: Comparison of nonlinear (N) versus linear (L) iterations

4.8.2 Numerical results for mesh configuration-II

In this section, the numerical investigation is performed by using computational mesh (3.4.3) for the solution of Navier Stokes equations. The computational mesh is discussed in detail in section 3.4. It consists of 346 quadrilateral elements. For the numerical investigation, we perform the same tests as we did in the above section and this work is taken from our research done in [77].

We employ the bilinear finite element Q_1 for the unknown variables such as velocity, stress and pressure. At first, we calculate and present the total number of degrees of freedom N_{DoF} for all the variable to be computed as given in the Table 4.9.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	782	391	1173	2346
2	2948	1474	4422	8844
3	11432	5716	17148	34296
4	45008	22504	67512	135024
5	178592	89296	267888	535776
6	711488	355744	1067232	2134464

Table 4.9: Total unknowns for $Q_1Q_1Q_1$-elements in mesh-II

Also, the flow parameters such as coefficients of drag, lift, pressure difference and global mass conservation at $x = 2.2$ are presented in the Table 4.10 below.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.171635E+00$	$2.105224E-02$	$1.031347E-01$	$1.114499E+00$
5	$5.444013E+00$	$1.429395E-02$	$1.117922E-01$	$2.997717E-01$
6	$5.541546E+00$	$1.175843E-02$	$1.152451E-01$	$7.786550E-02$
Fixed Point Method				
4	$5.257355E+00$	$1.575412E-02$	$1.055235E-01$	$9.842843E-01$
5	$5.479349E+00$	$1.186275E-02$	$1.127035E-01$	$2.621342E-01$
6	$5.552244E+00$	$1.074011E-02$	$1.155112E-01$	$6.741351E-02$

Table 4.10: Flow parameters for $Q_1Q_1Q_1$-element in mesh-II

Now, the biquadratic finite element Q_2 is used to solve the problem 4.4 for all the variables. The total number of unknowns or degrees of freedom generated for Q_2 element on each level for the velocity, pressure and stress variables are presented by Table 4.11.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	2948	1474	4422	8844
2	11432	5716	17148	34296
3	45008	22504	67512	135024
4	178592	89296	267888	535776
5	711488	355744	1067232	2134464
6	2840192	1420096	4260288	8520576

Table 4.11: Total unknowns for $Q_2Q_2Q_2$-elements in mesh-II

The coefficients of drag and lift, pressure difference and global mass conservation are displayed below.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.576975E+00$	$1.053552E-02$	$1.173265E-01$	$4.246499E-03$
5	$5.579242E+00$	$1.060640E-02$	$1.174766E-01$	$5.550170E-04$
6	$5.579503E+00$	$1.061714E-02$	$1.175098E-01$	$6.810928E-05$
Fixed Point Method				
4	$5.576955E+00$	$1.049599E-02$	$1.173236E-01$	$3.021125E-03$
5	$5.579238E+00$	$1.060355E-02$	$1.174762E-01$	$3.952429E-04$
6	$5.579503E+00$	$1.061696E-02$	$1.175098E-01$	$4.815527E-05$

Table 4.12: Flow parameters for $Q_2Q_2Q_2$-element in mesh-II

In the last test, we employ biquadratic element Q_2 for velocity, stress and discontinuous pressure element P_1^{dc} for pressure.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	2948	1038	4422	8408
2	11432	4152	17148	32732
3	45008	16608	67512	129128
4	178592	66432	267888	512912
5	711488	265728	1067232	2044448
6	2840192	1062912	4260288	8163392

Table 4.13: Total unknowns for $Q_2P_1^{dc}Q_2$-elements in mesh-II

The total number of unknowns calculated on each level for the stress-velocity-pressure problem are presented by Table 4.13. The fluid flow quantities are given in Table 4.14.

Lev.	*Drag* C_D	*Lift* C_L	Δp	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.576955E+00$	$1.053593E-02$	$1.173867E-01$	$4.266894E-03$
5	$5.579240E+00$	$1.060643E-02$	$1.174910E-01$	$5.564637E-04$
6	$5.579503E+00$	$1.061715E-02$	$1.175130E-01$	$6.821582E-05$
Fixed Point Method				
4	$5.576939E+00$	$1.049607E-02$	$1.173839E-01$	$3.033199E-03$
5	$5.579237E+00$	$1.060354E-02$	$1.174907E-01$	$3.961265E-04$
6	$5.579503E+00$	$1.061699E-02$	$1.175130E-01$	$4.821435E-05$

Table 4.14: Flow parameters for $Q_2P_1^{dc}Q_2$-element in mesh-II

The iterative comparison of solvers is illustrated in table 4.15. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

level	$Q_1Q_1Q_1\,(N/L)$	$Q_2Q_2Q_2\,(N/L)$	$Q_2P_1^{dc}Q_2\,(N/L)$
Gauss − Newton Method			
2	8/4	7/7	7/7
3	7/7	4/15	4/15
4	6/12	4/21	4/21
5	5/21	5/23	5/23
Fixed Point Method			
2	9/3	14/6	14/6
3	9/5	9/9	9/7
4	11/10	8/20	8/17
5	10/21	6/23	6/23

Table 4.15: Comparison of solvers with nonlinear and averaged linear iterations

From Tables 4.8 and 4.15, it can be seen that linear solver take more and more iterations with respective mesh refinements and it effects also the performance of nonlinear solvers. As it is explained in the section 4.6 that the coefficient matrix 4.19 generated by SVP formulation is not differentially diagonal dominant [59]. Therefore, the smoother has slow convergence rate and ultimately effects the multigrid solver. In the Table 4.16, the numerical test is performed for pre/post smoothing steps $SmSt$ 2, 4, 8 and 16 using $Q_2Q_2Q_2$ finite elements.

level	$SmSt-2$	$SmSt-4$	$SmSt-8$	$SmSt-16$
Gauss − Newton Method				
2	7/12	7/7	7/4	7/2
3	4/20	4/15	4/7	4/4
4	6/24	4/21	3/14	3/7
5	8/25	5/23	3/19	3/15
Fixed Point Method				
2	14/10	14/6	14/4	14/2
3	9/15	9/9	9/4	9/2
4	8/24	8/20	8/9	8/4
5	9/25	6/23	6/22	6/11

Table 4.16: Iterative comparison of solvers for vaious smoothin steps

It is observed that for the higher smoothing steps the performance of linear can be enhanced. The fixed point method is outperformed by Gauss-Newton method. From

Table 4.16 it is shown that the efficiency of the solvers can be improved in term of nonlinear iterations but at the cost of some additional computations.

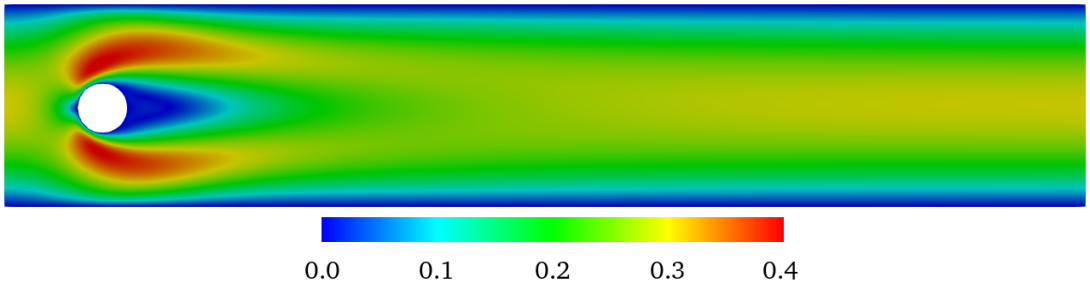

Figure 4.2: Velocity profile for Navier-Stokes problem

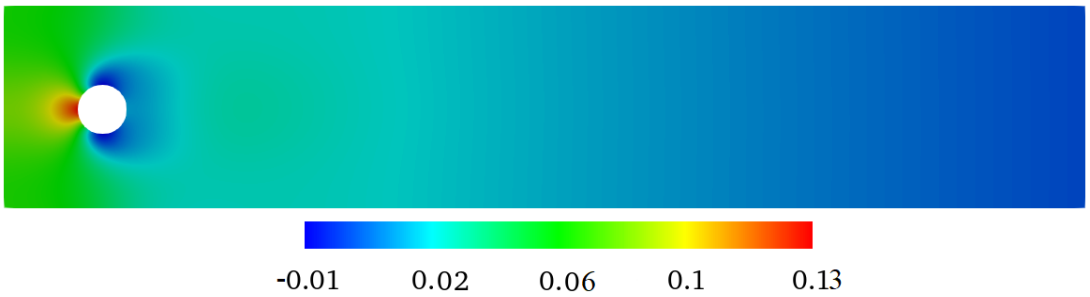

Figure 4.3: Pressure Difference for Navier-Stokes problem

4.9 Summary

We have presented a numerical study to develop an efficient linear weighted least-squares finite element method (L_w-LSFEM) for Stokes and Navier-Stokes problems using stress-velocity-pressure first order systems. We have used two mesh configurations (3.4.2) and (3.4.3) for the the numerical investigations of flow around cylinder problem and computed the flow parameters such as coefficients of drag and lift, pressure difference across cylinder at Reynolds number $\mathbf{Re} = 20$. The SVP formulation is rarely used for Stokes problem because it appears to have more unknowns and consequently more computational cost. But, we are working only with SVP type formulations therefore we employed it to validate our solver for both lower and higher order finite element discretizations. It is also observed that for a the linear problem we got equal approximations for nonlinear solvers i.e. Gauss-Newton and fixed point methods. Because, the defect computed in both methods for the linear problem is equivalent. In case of Navier-Stokes problem or nonlinear problem the results are slightly different. The reason behind it, is the application of least-squares principle after the linearization of nonlinear terms in the given problem. We got different defect

for each method. The application of least-squares principle on system of equations, based on SVP formulation, leads to a coefficient matrix which is not differentially diagonal dominant. Therefore, the linear solver show slow convergence and the optimality of linear solver is an open problem.

We use three types of finite element discretizations $Q_1Q_1Q_1$, $Q_2P_1^{dc}Q_2$ and $Q_2Q_2Q_2$ for analysis in both mesh configurations referred above. The overall results computed for all respective discretizations using mesh-II (3.4.3) are better than those for mesh-I (3.4.3). Also, the results for higher order finite element discretizations are much more accurate than the results for lower order finite element discretizations. The results obtained from mesh-II for higher order finite element discretizations are more promising toward benchmark results [91] than the outcomes produced by mesh-I. The convergence behavior is also mesh independent for both cases.

One of the main objectives of the study is to find the proper weighted function $w^n_{\eta_\nu}$ used in least-squares energy functional to develop an accurate and robust solver. The weighted function used for the analysis is a function depending on the viscosity of fluid and the viscosity used in this chapter is constant $\eta_\nu = 0.001$. It is observed from numerical tests that the choice of weighed function played an important role in the convergence of our solver. The numerical test carried out only for the weighted function $w^n_{\eta_\nu} = (1.0/\eta_\nu)^n = (1.0/0.001)^n$ in this chapter. We got the optimal results only for the parametric value $n = 1.0$ in weighted function $w^n_{\eta_\nu}$.

The mass conservation is another important factor in the study of least-squares finite element methods. The aim of our research is to get the global mass conservation $m^{GMC}_{x=2.2}$ on outflow at $x = 2.2$ of flow domain without activating weight or scaling parameter α_m associated with continuity equation. Therefore, we use $\alpha_m = 1.0$ in our tests so that the scaling parameter has no effect in computation of global mass conservation. The only weighted function employed in our least-squares functional is $w^n_{\eta_\nu}$ and we used only this weight to study overall behavior of L_w-LSFEM. The lack of global mass conservation has been obtained in numerical tests when it is carried out for lower order finite element discretization but for higher order finite element discretizations it is outperformed, especially for Q_2 elements for mesh-II in Table 4.12 and 4.14.

We can conclude that for higher order finite elements the highly accurate results are achieved. The mesh independent convergence behavior is observed for all the performed test. The global mass conservation is achieved with higher order finite element discretizations. We got excellent results for mesh configuration 3.4.3 over the mesh 3.4.2 but at some computational cost.

Chapter 5

Isothermal non-Newtonian Fluid Flows

In this chapter, the non-Newtonian fluid flow problems are simulated numerically for the power law and the Cross law viscosity models. We develop the least-squares finite element method balanced by different nonlinear weighted functions as suitable physical quantities in the flow model. The numerical investigation of the power law fluids is followed by the analysis of Cross law fluid flow problems. We designed special nonlinear weighted least-squares finite element techniques for stationary incompressible Navier-Stokes 2D problem for nonlinear models approximated with higher order finite elements. We used nonlinear viscosity as nonlinear weighted function and obtained the optimal results for a vast range of parameters. The shear thickening and thinning behavior of the fluids are examined for power law viscosity model. The Cross law model is used to study complex fluids with bounded viscosities. We employ conjugate gradient scheme preconditioned by multigrid solver to approximate the systems based on SVP formulation with accuracy and efficiently.

5.1 Introduction

The general motivation behind using least-squares finite element methods is to achieve the providential properties of Rayleigh-Ritz methods. For instance, in contrast to standard mixed finite element methods (MFEM), the selection of finite element spaces is independent of compatibility condition (i.e. the LBB stability condition) and the linear system is always symmetric and positive definite [15]. This allows to build an efficient MPCG solver to approximate linear systems which is the key task in this work. The non-Newtonian fluid is an important class of fluids and used extensively for research activities. We study the solution of the time independent fluids, which are also known as "generalized Newtonian fluid". We consider two types of nonlinear viscosity

models namely power law model and Cross law model for the problem formations and numerical investigations.

In Literature, the non-Newtonian fluids have been investigated for many first order formulations and nonlinear viscosity models in LSFEM. The stress based formulation is used in [94] to investigate the power law fluids and the results for shear thinning and shear thickening fluids for the variations of power law index were presented. The power law fluids explored for p-version least-squares finite element method in [10], [46], [47] and optimal results were presented. The least-squares finite element was developed for Carreau fluid model in [31]. In 3D spaces, the power law fluids were studied in [36] and the Carreau-Yasuda model was solved in [49] using LSFEM. For the flow around cylinder domain, the NS problem was investigated for power law model in [37] (MFEM), [75] (LSFEM) and for Cross law fluids in [37] using MFEM. We use these results as reference for our investigations and study the NS problem for nonlinear weighted LSFEM for both power law and Cross law models. The Cross law model is employed first time to investigate least-squares finite element method for flow around cylinder domain.

5.2 System of governing equations

The system of equations considered for the steady problem of generalized Newtonian flow on a bounded domain $\Omega \subset \mathbb{R}^2$ is given as

$$\left.\begin{aligned} \mathfrak{u}\cdot\nabla\mathfrak{u} + \nabla\mathfrak{p} - \nabla\cdot\big(2\eta_{\nu}\big(\dot{\gamma}_{_{II}},\mathfrak{p}\big)\mathsf{D}\big(\mathfrak{u}\big)\big) &= \mathfrak{f} \\ \nabla\cdot\mathfrak{u} &= 0 \end{aligned}\right\} \text{ in } \Omega \tag{5.1}$$

where $\mathfrak{u}$ is the velocity of fluid, $\mathfrak{p} = P/\rho$ the normalized pressure, $\mathfrak{f}$ is the external source term, $\eta_{\nu}(\cdot)$ is the nonlinear viscosity depending on shear rate or pressure or both.

5.3 Formulation of first order system

To formulate the first order system, we consider the Cauchy stress tensor $\boldsymbol{\sigma}$ as a new externally defined variable as follows

$$\begin{aligned} \boldsymbol{\sigma} &= \mathbf{v}_{\boldsymbol{\sigma}} - \mathbf{p}_{\boldsymbol{\sigma}} \\ &= 2\eta_{\nu}\big(\dot{\gamma}_{_{II}},\mathfrak{p}\big)\mathsf{D}\big(\mathfrak{u}\big) - \mathfrak{p}\mathsf{I} \end{aligned} \tag{5.2}$$

where $\mathbf{p}_{\boldsymbol{\sigma}} = \mathfrak{p}\mathsf{I}$ is called the inviscid reactive component of Cauchy stress tensor and $\mathbf{v}_{\boldsymbol{\sigma}} = 2\eta_{\nu}\big(\dot{\gamma}_{_{II}},\mathfrak{p}\big)\mathsf{D}\big(\mathfrak{u}\big)$ is the active viscous component of the Cauchy stress tensor, which is a function of deformation gradient of velocity as $\mathsf{D}\big(\mathfrak{u}\big)$ and the second

invariant of the shear rate tensor as $\dot{\gamma}_{II}$ as defined in section 3.5, but the only difference is in the active viscous part $\mathbf{v}_{\boldsymbol{\sigma}}$ of the fluid. Here, the viscosity of the fluid in this case is nonlinear.

Now the first order stress-velocity-pressure (SVP) formulation for the steady problems of generalized Newtonian flow has the following form.

$$\left.\begin{aligned} \mathfrak{u}\cdot\nabla\mathfrak{u} - \nabla\cdot\boldsymbol{\sigma} &= \mathfrak{f}, \\ \nabla\cdot\mathfrak{u} &= 0, \\ \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}} - \mathbf{v}_{\boldsymbol{\sigma}} &= 0, \end{aligned}\right\} \text{ in } \Omega, \tag{5.3}$$

The Dirichlet boundary conditions $f_{\hat{\Omega}_D}$ on the Dirichlet boundary region $\hat{\Omega}_D$ and the Neumann boundary condition $f_{\hat{\Omega}_N}$ on the outward unit normal $\mathbf{n}$ of Neumann boundary region $\hat{\Omega}_D$ is given as

$$\left.\begin{aligned} \mathfrak{u} &= f_{\hat{\Omega}_D} \\ \mathbf{n}\cdot\boldsymbol{\sigma} &= f_{\hat{\Omega}_N} \end{aligned}\right\} \text{ on } \hat{\Omega} \tag{5.4}$$

and it also satisfy the zero mean pressure constraint (2.88).

5.4 Linearization of nonlinear problem

To overcome the computational difficulties, we first linearize the nonlinear terms present in the considered problem. In the given problem there are two nonlinear terms, the convection term $\mathfrak{u}\cdot\nabla\mathfrak{u}$ and the viscus nonlinear term $\mathbf{v}_{\boldsymbol{\sigma}} = 2\eta_\nu\big(\dot{\gamma}_{II},\mathfrak{p}\big)\mathsf{D}\big(\mathfrak{u}\big)$. We apply the Newton's linearization to the nonlinear terms before the application of least-squares principle to the problem. Therefore, the following substitutions are made for the nonlinear term

$$\mathfrak{u}\cdot\nabla\mathfrak{u} \approx \mathfrak{u}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} \tag{5.5}$$

and

$$\mathbf{v}_{\boldsymbol{\sigma}} \approx \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big) + \mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big) + \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big) \tag{5.6}$$

$$\begin{aligned} 2\eta_\nu\big(\dot{\gamma}_{II},\mathfrak{p}\big)\mathsf{D}\big(\mathfrak{u}\big) \;\approx\; & 2\frac{\partial\eta_\nu}{\partial\mathfrak{p}}\mathsf{D}\big(\mathfrak{u}^k\big)\mathfrak{p} + 2\eta_\nu\big(\dot{\gamma}^k_{II},\mathfrak{p}^k\big)\mathsf{D}\big(\mathfrak{u}\big) \\ & + \frac{2}{\dot{\gamma}_{II}}\frac{\partial\eta_\nu}{\partial\dot{\gamma}_{II}}\Big[\mathsf{D}\big(\mathfrak{u}\big) : \mathsf{D}\big(\mathfrak{u}^k\big)\Big]\,\mathsf{D}\big(\mathfrak{u}^k\big). \end{aligned} \tag{5.7}$$

Let $\mathcal{R}$ be the sum of residuals for the system of equations (5.3). To approximate the residuals, the nonlinear iteration is updated with the correction $\mathbf{u}^{k+1} = \mathbf{u}^k + \delta\mathbf{u}$. Then, the residual approximation is obtained by Taylor series expansion as

$$\begin{aligned}\mathcal{R}\big(\mathbf{u}^{k+1}\big) &= \mathcal{R}\big(\mathbf{u}^{k}+\delta\mathbf{u}\big),\\ &\simeq \mathcal{R}\big(\mathbf{u}^{k}\big)+\big[\mathcal{D}\mathcal{R}\big(\mathbf{u}^{k}\big)\big]\delta\mathbf{u},\end{aligned} \tag{5.8}$$

where

$$\mathcal{R}\big(\mathbf{u}^{k}\big)=\begin{bmatrix}\mathcal{R}_1^k\\ \mathcal{R}_2^k\\ \mathcal{R}_3^k\end{bmatrix}=\begin{bmatrix}\mathfrak{u}^k\cdot\nabla\mathfrak{u}^k-\nabla\cdot\boldsymbol{\sigma}^k-\mathfrak{f}\\ \nabla\cdot\mathfrak{u}^k\\ \boldsymbol{\sigma}^k+\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{u}^k\big)-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}^k\big)\end{bmatrix}, \tag{5.9}$$

and the system of equations after substitution is as following

$$\mathcal{D}\mathcal{R}\big(\mathbf{u}^{k}\big)\delta\mathbf{u}=\begin{bmatrix}\mathfrak{u}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{u}-\nabla\cdot\boldsymbol{\sigma}-\mathfrak{f}\\ \nabla\cdot\mathfrak{u}\\ \boldsymbol{\sigma}+\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)\end{bmatrix}. \tag{5.10}$$

By following the procedure discussed in Algorithm 3.2, i.e. the nonlinear basic iteration for least-squares scheme, one can obtain a solution update $\delta\mathbf{u}$ as

$$\tilde{\mathcal{R}}\big(\mathbf{u}^{k}\big)\delta\mathbf{u}=\mathcal{R}_{_{def}}\big(\mathbf{u}^{k}\big), \tag{5.11}$$

and for different choices of $\tilde{\mathcal{R}}$ and $\mathcal{R}_{_{def}}$, as discussed in section 4.5, one can obtain Newton and fixed point approximation schemes.

5.5 Continuous least-squares principle

We introduce the spaces $\mathcal{X}_{_\Omega}$ of admissible functions based on the residuals of the first-order systems (5.3)

$$\mathcal{X}_{_\Omega}:=\mathcal{H}^1_{0,D}\big(\Omega\big)\times\mathcal{L}^2_0\big(\Omega\big)\times\mathcal{H}_{0,N}\big(\mathrm{div},\Omega\big). \tag{5.12}$$

The linearized functional $\mathcal{E}_{_{svp}}$ in term of the $\mathcal{L}^2$-norm is obtained by using least-squares principles on linearized residual (5.8) as

$$\mathcal{E}_{_{svp}}\big(\delta\mathbf{u};0\big)=\frac{1}{2}\big\|\mathcal{R}\big(\mathbf{u}^{k}\big)+\big[\mathcal{D}\mathcal{R}\big(\mathbf{u}^{k}\big)\big]\delta\mathbf{u}\big\|^2_0. \tag{5.13}$$

and the $\mathcal{L}^2$-norm based SVP energy functionals can be expanded as following

$$\begin{aligned}\mathcal{E}_{_{svp}}\big(\delta\mathbf{u};0\big) &= \tfrac{1}{2}\Big(\big\|\mathcal{R}_1^k+\mathfrak{u}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{u}-\nabla\cdot\boldsymbol{\sigma}\big\|^2_0+\big\|\mathcal{R}_2^k+\nabla\cdot\mathfrak{u}\big\|^2_0\\ &+\big\|\mathcal{R}_3^k+\boldsymbol{\sigma}+\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)-\ \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)\big\|^2_0\Big),\\ &\qquad\qquad\forall\ \big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_{_\Omega}.\end{aligned} \tag{5.14}$$

and weighted functional associated with the nonlinear weighted function $w^{^n}_{\eta_\nu}$ is

$$\begin{aligned}\mathcal{E}_{svp}^{w_{\eta_\nu}}\big(\delta\mathbf{u};0\big) &= \tfrac{1}{2}\,\Big(\big\|\mathcal{R}_1^k+\mathfrak{u}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{u}-\nabla\cdot\boldsymbol{\sigma}\big\|_0^2+\alpha_m\big\|\mathcal{R}_2^k+\nabla\cdot\mathfrak{u}\big\|_0^2\\ &+ w_{\eta_\nu}^n\big\|\mathcal{R}_3^k+\boldsymbol{\sigma}+\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)-\ \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)\big\|_0^2\Big),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\forall\ \big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_\Omega.\end{aligned}\tag{5.15}$$

where the nonlinear weighted functions are defined as $w_{\eta_\nu}^n=\big(1.0/\eta_\nu\big)^n$; $n\geq 1$ and these functions applied only to the residual part of additional Cauchy stress equation. Here, to improve mass conservation, a scaling parameter α_m is used for the least-squares formulation [42], [23], [24].

The minimization problem associated with the least-squares formulation in (5.15) is to find $\delta\mathbf{u}=\big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_\Omega$ such that

$$\delta\mathbf{u}=\underset{\mathbf{v}\in\mathcal{X}_\Omega}{\arg\min}\ \mathcal{E}_{svp}^{w_{\eta_\nu}}\big(\mathbf{v};0\big).\tag{5.16}$$

The first variant of the quadratic linearized functional (5.16) leads to an equivalent problem of finding $\delta\mathbf{u}$ such that

$$\big[\mathcal{R}\big(\mathbf{u}^k\big)+\mathcal{D}\mathcal{R}\big(\mathbf{u}^k\big)\delta\mathbf{u}\big]\mathcal{D}\mathcal{R}\big(\mathbf{u}^k\big)\mathbf{v}=0,\quad\forall\ \mathbf{v}\in\mathcal{X}_\Omega.\tag{5.17}$$

The variational problem based on the optimality condition (5.17) of the minimization problem (4.16) is given by

$$\begin{cases}\text{Find }\ \delta\mathbf{u}=\big(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma}\big)\in\mathcal{X}_\Omega\quad\text{such that,}\\ \mathbf{K}\big(\delta\mathbf{u},\mathbf{v}\big)=\mathbf{F}\big(\mathbf{v}\big)\quad\forall\ \mathbf{v}=\big(\mathfrak{v},\mathfrak{q},\boldsymbol{\tau}\big)\in\mathcal{X}_\Omega.\end{cases}\tag{5.18}$$

where $\mathbf{K}$ is the bilinear form defined on $\mathcal{X}_\Omega\times\mathcal{X}_\Omega\to\mathbb{R}$ as following

$$\begin{aligned}\mathbf{K}\ :\ &=\big\langle\mathfrak{u}^k\cdot\nabla\mathfrak{u}+\mathfrak{u}\cdot\nabla\mathfrak{u}^k-\nabla\cdot\boldsymbol{\sigma},\mathfrak{u}^k\cdot\nabla\mathfrak{v}+\mathfrak{v}\cdot\nabla\mathfrak{u}^k-\nabla\cdot\boldsymbol{\tau}\big\rangle\\ &+\alpha_m\big\langle\nabla\cdot\mathfrak{u},\nabla\cdot\mathfrak{v}\big\rangle+w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma},\boldsymbol{\tau}\big\rangle\\ &+w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma},\mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)\big\rangle+w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma},-\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle\\ &+w_{\eta_\nu}^n\big\langle\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big),\boldsymbol{\tau}+\mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)\big\rangle\\ &+w_{\eta_\nu}^n\big\langle\mathbf{p}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big),-\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle\\ &+w_{\eta_\nu}^n\big\langle-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big),\boldsymbol{\tau}+\mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)\big\rangle\\ &+w_{\eta_\nu}^n\big\langle-\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big),-\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle\end{aligned}\tag{5.19}$$

and the linear forms $\mathbf{F}$ is defined on $\mathcal{X}_\Omega\to\mathbb{R}$ as following

$$\begin{aligned}\mathbf{F}\ :\ &=\big\langle\mathcal{R}_1^k,\mathfrak{v}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{v}-\nabla\cdot\boldsymbol{\tau}\big\rangle+\alpha_m\big\langle\mathcal{R}_2^k,\nabla\cdot\mathfrak{v}\big\rangle\\ &+w_{\eta_\nu}^n\big\langle\mathcal{R}_3^k,\boldsymbol{\tau}+\mathbf{p}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{q}\big)-\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}\big(\mathfrak{v}\big)\big\rangle\end{aligned}\tag{5.20}$$

To analyze the properties of the least-square problem the operator form is as following

$$\begin{aligned}\mathbf{K} &= \big[\mathcal{DR}(\mathbf{u}^k)\big]^*\big[\mathcal{DR}(\mathbf{u}^k)\big] \\ &= \begin{bmatrix} \mathbf{K}_{vu} & \mathbf{K}_{vp} & \mathbf{K}_{v\boldsymbol{\sigma}} \\ \mathbf{K}_{qu} & \mathbf{K}_{qp} & \mathbf{K}_{q\boldsymbol{\sigma}} \\ \mathbf{K}_{\boldsymbol{\tau}u} & \mathbf{K}_{\boldsymbol{\tau}p} & \mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\sigma}} \end{bmatrix},\end{aligned} \tag{5.21}$$

where

$$\begin{aligned}
\mathbf{K}_{vu} &= w^n_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + \big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle + \alpha_m\big\langle\nabla\cdot\mathfrak{u}, \nabla\cdot\mathfrak{v}\big\rangle \\
&\quad + \boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k, \mathfrak{v}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle} + \boxed{\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle}, \\
\mathbf{K}_{vp} &= w^n_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle}, \\
\mathbf{K}_{v\boldsymbol{\sigma}} &= w^n_{\eta_\nu}\big\langle \boldsymbol{\sigma}, -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \boldsymbol{\sigma}, -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + \big\langle -\nabla\cdot\boldsymbol{\sigma}, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle \\
&\quad + \boxed{\big\langle -\nabla\cdot\boldsymbol{\sigma}, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle}, \\
\mathbf{K}_{qu} &= w^n_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}, \\
\mathbf{K}_{qp} &= w^n_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}, \\
\mathbf{K}_{q\boldsymbol{\sigma}} &= w^n_{\eta_\nu}\big\langle \boldsymbol{\sigma}, \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \boldsymbol{\sigma}, -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}, \\
\mathbf{K}_{\boldsymbol{\tau}u} &= w^n_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \boldsymbol{\tau}\big\rangle + \big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, -\nabla\cdot\boldsymbol{\tau}\big\rangle + \boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k, -\nabla\cdot\boldsymbol{\tau}\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \boldsymbol{\tau}\big\rangle}, \\
\mathbf{K}_{\boldsymbol{\tau}p} &= w^n_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), \boldsymbol{\tau}\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), \boldsymbol{\tau}\big\rangle}, \\
\mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\sigma}} &= w^n_{\eta_\nu}\big\langle \boldsymbol{\sigma}, \boldsymbol{\tau}\big\rangle + \big\langle -\nabla\cdot\boldsymbol{\sigma}, -\nabla\cdot\boldsymbol{\tau}\big\rangle,
\end{aligned}$$

Here, the terms shown in the boxes have contributions in the coefficient matrix $\mathbf{K}$ due to the nonlinear terms. These terms will be add to coefficient matrix $\mathbf{K}$ after defect calculation The resulting system matrix (5.21) is symmetric and positive definite but the matrix is not differentially diagonal dominant. We are able to use the conjugate gradient method to efficiently solve the system of equations.

5.6 Discrete least-squares principle

The discrete form of the problem (5.18) with approximation spaces $\mathcal{X}_{\Omega}^{h}$ is defined as

$$\mathcal{X}_{\Omega}^{h} = \left\{ \delta\mathbf{u}^{h} \in \mathcal{H}_{0,D}^{1,h}(\Omega) \times \mathcal{L}_{0}^{2,h}(\Omega) \times \mathcal{H}_{0,N}^{h}(\mathrm{div}, \Omega) \right\}. \tag{5.22}$$

where $\delta\mathbf{u}^{h} = \left(\mathfrak{u}_{h}, \mathfrak{p}_{h}, \boldsymbol{\sigma}_{h}\right)$ can be written as

$$\begin{cases} \text{Find } \delta\mathbf{u}^{h} \in \mathcal{X}_{\Omega}^{h} \quad \text{such that,} \\ \mathbf{K}^{h}\left(\delta\mathbf{u}^{h}, \mathbf{v}^{h}\right) = \mathbf{F}^{h}\left(\mathbf{v}^{h}\right), \end{cases} \tag{5.23}$$

where, for all $\mathbf{v}^{h} = (\mathfrak{v}_{h}, \mathfrak{q}_{h}, \boldsymbol{\tau}_{h}) \in \mathcal{X}_{\Omega}^{h}$, $\mathbf{K}^{h}$ is the approximated bilinear form defined on $\mathcal{X}_{\Omega}^{h} \times \mathcal{X}_{\Omega}^{h} \to \mathbb{R}$ and the linear forms $\mathbf{F}^{h}$ is defined on $\mathcal{X}_{\Omega}^{h} \to \mathbb{R}$.

5.7 Numerical results and analysis

In this section, we analyze the nonlinear weighted least-squares finite element method (N_{w}-LSFEM) for the solution of stationary incompressible flow problems of non-Newtonian fluids by using SVP first order formulation. The power law and Cross law models are examined for unbounded and bounded nonlinear viscosities, respectively. The Q_2 biquadratic finite element is used for the discretization of all the unknown variables, which are velocity, stress and pressure. The multigrid preconditioned conjugate gradient method is employed as a linear solver with multigrid solver as a preconditioner. For the nonlinear solver, the Gauss-Newton and the fixed point like methods are used. The efficiency of the method is measured by its performance and accuracy. The results are compared with the reference solutions calculated by Damanik [37] and Nickaeen [75].

We computed and analyzed the numerical outcomes for a variety of nonlinear weighted functions $w_{\eta_\nu}^{n}$ used in the least-squares energy functionals. The main objective of the study is to find the significant impact of nonlinear weights on the solver's performance. The computational mesh configuration (3.4.2) is used for numerical tests. The total number of unknowns computed for Q_2 element discretization are presented in the Table 5.1 for the stress-velocity-pressure problem, where unknowns are computed in hierarchical order from coarser(top) to finer(bottom) levels.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	*Total* N_{DoF}
1	1144	572	1716	3432
2	4368	2184	6552	13104
3	17056	8528	25584	51168
4	67392	33696	101088	202176
5	267904	133952	401856	803712
6	1068288	534144	1602432	3204864

Table 5.1: Total number of unknows using $Q_2Q_2Q_2$ discretization

5.7.1 Power law model

The power law viscosity model (3.43) is used for variants of power law index r to represent the shear thinning and shear thickening fluids. The tests are carried out for nonlinear weight $w^{n}_{\eta_\nu}$; $n = 1.0$ which is a function dependent on nonlinear viscosity and other useful parametric values are $\epsilon = 0.1$, $\eta_0 = 10^{-3}$.

Shear thinning fluid flow for $w^{1.0}_{\eta_\nu}$

The nonlinear viscosity η_ν in power law model (3.43) depends on shear rate $\dot{\gamma}_{_{II}}$ of the fluid. The shear thinning effects of the flow are obtained for the power law index value $r = 1.5$ with weighted function $w^{1.0}_{\eta_\nu}$ and is shown in Table 5.2.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$3.246358E+00$	$-1.149702E-02$	$9.382544E-02$	$1.103180E-01$
5	$3.269706E+00$	$-1.318620E-02$	$9.610286E-02$	$2.514482E-02$
6	$3.276482E+00$	$-1.331922E-02$	$9.664751E-02$	$5.122436E-03$
Fixed Point Method				
4	$3.207203E+00$	$-1.332608E-02$	$9.314482E-02$	$8.133711E-02$
5	$3.260655E+00$	$-1.335087E-02$	$9.594829E-02$	$1.747758E-02$
6	$3.275091E+00$	$-1.332572E-02$	$9.663882E-02$	$3.140688E-03$
[37]	$3.27833E+00$	$-1.332E-02$		
[75]	$3.02080E+00$	$-1.52413E-02$	$9.42107E-02$	

Table 5.2: Shear thinning effects for $r = 1.5$ with weighted function $w^{1.0}$

Newtonian fluid flow for $w^{1.0}_{\eta_\nu}$

For power law index value $r = 2.0$, the fluid behaves as Newtonian and flow parameters computed for flow around cylinder problem are shown in Table 5.3. As it can be seen that global mass is conserved very well for both methods.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.556446E+00$	$1.115572E-02$	$1.162045E-01$	$3.618464E-02$
5	$5.576758E+00$	$1.066844E-02$	$1.172607E-01$	$4.701878E-03$
6	$5.579248E+00$	$1.062216E-02$	$1.174601E-01$	$5.885206E-04$
Fixed Point Method				
4	$5.556770E+00$	$1.028484E-02$	$1.162086E-01$	$2.663399E-02$
5	$5.576749E+00$	$1.057478E-02$	$1.172595E-01$	$3.416974E-03$
6	$5.579242E+00$	$1.061386E-02$	$1.174598E-01$	$4.229640E-04$
[37]	$5.579313E+00$	$1.061503E-02$		
[75]	$5.579242E+00$	$1.060640E-02$	$1.174766E-01$	

Table 5.3: Newtonian fluid flow for $r = 2.0$ with weighted function $w^{1.0}$

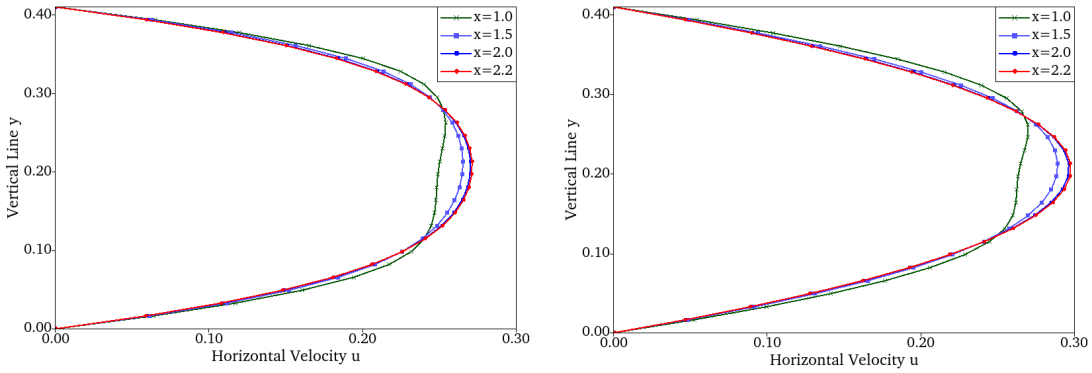

Figure 5.1: Horizontal velocity profile u at different cross sections x of the shear thinning and Newtonian flow for $r = 1.5$ and $r = 2.0$

In Figure 5.2, the comparison of the Gauss-Newton and fixed point method for the global mass conservation in the Tables 5.2 and 5.3 is presented on the outflow cross-section at $x = 2.2$. It is observed that the mass is conserved globally for both methods.

Shear thickening fluid flow for $w^{1.0}_{\eta_\nu}$

The shear thickening effects of the fluid flow are simulated for the power law index $r = 2.5$ and the results are shown in Table 5.4.

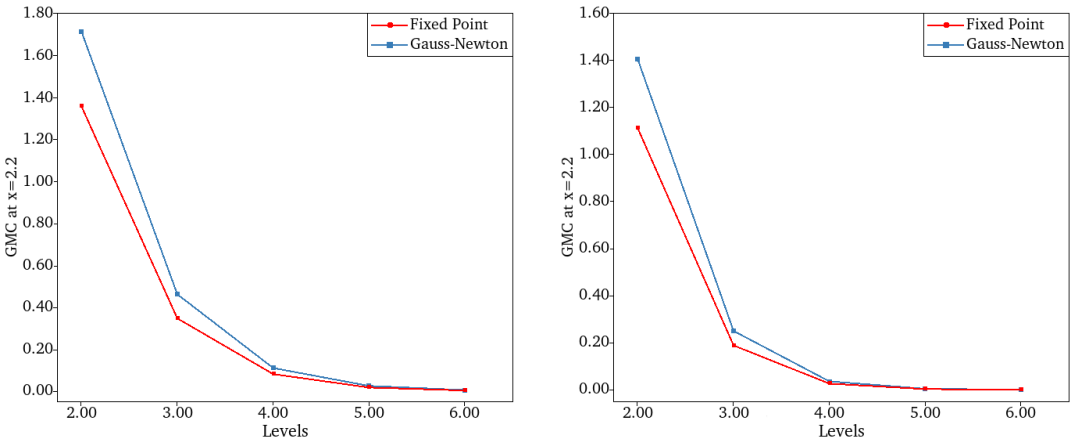

Figure 5.2: Comparison of nonlinear solvers for global mass conservation at outflow cross section of flow at $x = 2.2$ with corresponding levels for shear thinning flow (left) and Newtonian fluid flow (right).

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$8.694916E+00$	$9.399892E-02$	$1.479648E-01$	$1.621211E-02$
5	$8.704340E+00$	$9.405548E-02$	$1.489464E-01$	$1.854955E-03$
6	$8.705402E+00$	$9.406192E-02$	$1.491260E-01$	$2.208340E-04$
Fixed Point Method				
4	$8.694395E+00$	$9.352406E-02$	$1.479553E-01$	$1.227045E-02$
5	$8.704326E+00$	$9.398930E-02$	$1.489491E-01$	$1.376196E-03$
6	$8.705408E+00$	$9.405171E-02$	$1.491274E-01$	$1.626526E-04$
[75]	$9.660710E+00$	$13.20204E-02$	$1.600681E-01$	

Table 5.4: Shear thickening effect for $r = 2.5$ with weighted function $w^{1.0}$

Now we increase the power law index value to $r = 3.0$, which made the fluid thicker, hard to flow and challenging to solve. Even for this configuration our solver performs very well and generates accurate results, which are mesh independent as well.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$1.381088E+01$	$3.525638E-01$	$2.034762E-01$	$2.211654E-02$
5	$1.382478E+01$	$3.525774E-01$	$2.048400E-01$	$3.103754E-03$
6	$1.382730E+01$	$3.527192E-01$	$2.049490E-01$	$4.511015E-04$
Fixed Point Method				
4	$1.381416E+01$	$3.510574E-01$	$2.035904E-01$	$1.265699E-02$
5	$1.382558E+01$	$3.526187E-01$	$2.048686E-01$	$1.503496E-03$
6	$1.382746E+01$	$3.528160E-01$	$2.049549E-01$	$1.833629E-04$
[37]	$1.382715E+01$	$3.5294E-01$		
[75]	$1.800581E+01$	$5.486194E-01$	$2.545056E-01$	

Table 5.5: Shear thickening effect for $r = 3.0$ with weighted function $w^{1.0}$

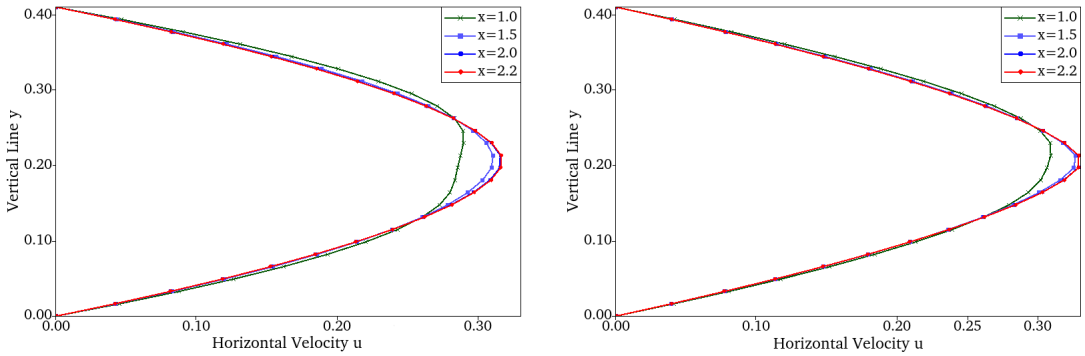

Figure 5.3: Horizontal velocity profile u at different cross sections x of the shear thickening flow for $r = 2.5$ and $r = 3.0$

In the following Table 5.6, we make a comparison of total nonlinear iterations and the corresponding averaged number of linear solver iterations for flow around cylinder. The relative errors for nonlinear and linear solvers are retained 1E-6 and 1E-3, respectively. It can be clearly seen from the table below that the Gauss-Newton worked very well in comparison to fixed point technique.

Level	$r = 1.5$	$r = 2.0$	$r = 2.5$	$r = 3.0$
Gauss − Newton Method				
2	15/12	9/8	7/6	9/6
3	12/23	5/12	5/10	5/11
4	10/24	4/17	4/19	5/22
5	6/25	4/22	4/23	5/25
Fixed Point Method				
2	23/8	14/7	14/6	94/5
3	20/12	10/9	14/11	242/12
4	22/18	9/11	12/19	628/20
5	19/23	8/20	10/24	1211/25

Table 5.6: Solvers comperison based on nonlinear/linear iterations

The shear thickening fluid for $r = 3.0$ at higher refinement levels the Gauss-Newton method converges in very few number of iterations but the fixed point method takes too many iterations to get approximate solutions. We observe an increase in the linear iterations as the level of mesh refinement is increased. We employ preconditioned conjugate gradient as a linear solver therefore the increase in the iterations is due to the increase in the condition number of the matrices generated on different levels. The

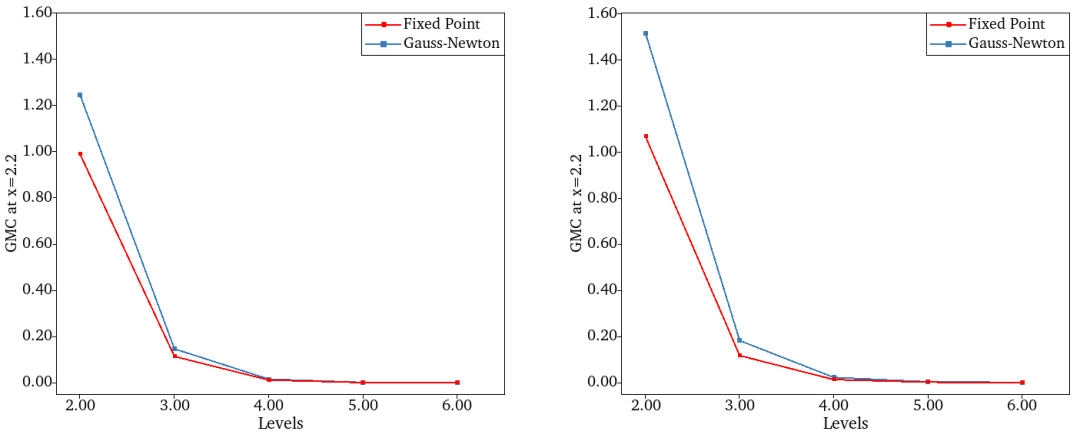

Figure 5.4: Comparison of nonlinear solvers for global mass conservation at outflow cross section of flow at $x = 2.2$ with corresponding levels for shear thickening flows at $r = 2.5$ (left) and $r = 3.0$ (right).

global mass computed in the Tables 5.4 and 5.5 is conserved for the shear thickening fluid flows and the comparison of nonlinear solvers on different levels is shown in the Figure 5.4.

5.7.2 Cross law model

The nonlinear viscosity $\eta_\nu(\dot{\gamma}_{_{II}}, \mathfrak{p})$, depends on shear rate $\dot{\gamma}_{_{II}}$ and pressure $\mathfrak{p}$, is employed for the numerical tests. The Cross law model is used for different parametric values of shear rate and pressure. The flow parameters for flow around cylinder problem with two different nonlinear weighted functions $w^{1.0}_{\eta_\nu}$, $w^{1.5}_{\eta_\nu}$ with different variants of the viscosity function are simulated. The weighting parameter associated with continuity equation is kept constant i.e. $\alpha_m = 1.0$. The tests are performed for the bounded viscosities range between $\eta_0 = 10^{-3}$ and $\eta_1 = 10^{-2}$.

Viscosity depends on shear rate with $w^{1.0}_{\eta_\nu}$

For the first test, the nonlinear viscosity function $\eta_\nu(\dot{\gamma}_{_{II}})$ depends on shear rate only. The results are obtained for the following set of parameters

$$r = 1,\ \alpha_p = 0,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2},$$

and the nonlinear weighted function is defined as

$$w^{1.0}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{_{II}})\right)^{1.0}.$$

The Table 5.7 presents the solutions for above mentioned parameters for flow around cylinder problem.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$6.302004E+00$	$2.508582E-02$	$1.262176E-01$	$3.631986E-02$
5	$6.324114E+00$	$2.510212E-02$	$1.273364E-01$	$4.498997E-03$
6	$6.326834E+00$	$2.510995E-02$	$1.275341E-01$	$5.338835E-04$
Fixed Point Method				
4	$6.305141E+00$	$2.472376E-02$	$1.263343E-01$	$2.707871E-02$
5	$6.324500E+00$	$2.506052E-02$	$1.273604E-01$	$3.323894E-03$
6	$6.326877E+00$	$2.510415E-02$	$1.275495E-01$	$3.925826E-04$
Ref [37] : *Drag* $C_{_D} = 6.32691E+00$, *Lift* $C_{_L} = 2.510E-02$				

Table 5.7: Shear dependent viscosity with weighted function $w^{1.0}$

Viscosity depends on pressure with $w^{1.0}_{\eta_\nu}$

Now, the numerical test is conducted for the nonlinear viscosity $\eta_\nu(\mathfrak{p})$ which is a function of pressure variable. The results are approximated for the parametric values as given by

$$r = 0,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2},$$

and nonlinear weighted function is defined as

$$w_{\eta_\nu}^{1.0} = \big(1/\eta_\nu(\mathfrak{p})\big)^{1.0},$$

and given in the Table 5.8.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m_{x=2.2}^{GMC}$
Gauss − Newton Method				
4	$3.330691E+01$	$8.216384E-01$	$4.870698E-01$	$1.673726E-02$
5	$3.331938E+01$	$8.222597E-01$	$4.927083E-01$	$1.574403E-03$
6	$3.332084E+01$	$8.223553E-01$	$4.942405E-01$	$1.589661E-04$
Fixed Point Method				
4	$3.330757E+01$	$8.216474E-01$	$4.870709E-01$	$1.407247E-02$
5	$3.331943E+01$	$8.222604E-01$	$4.927076E-01$	$1.356518E-03$
6	$3.332085E+01$	$8.223554E-01$	$4.942397E-01$	$1.400759E-04$
Ref [37] : *Drag* $C_{_D} = 3.331069E+01$, *Lift* $C_{_L} = 8.2217E-01$				

Table 5.8: Pressure dependent viscosity with weighted function $w^{1.0}$

Viscosity depends on shear rate and pressure with $w_{\eta_\nu}^{1.0}$

The nonlinear viscosity $\eta_\nu\big(\dot\gamma_{_{II}}, \mathfrak{p}\big)$ is a function of shear rate and pressure variables and the outcomes are generated for the following values

$$r = 1,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2},$$

and the weighted function is

$$w_{\eta_\nu}^{1.0} = \big(1/\eta_\nu(\dot\gamma_{_{II}}, \mathfrak{p})\big)^{1.0}.$$

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m_{x=2.2}^{GMC}$
Gauss − Newton Method				
4	$6.306605E+00$	$2.515882E-02$	$1.262980E-01$	$3.639816E-02$
5	$6.328733E+00$	$2.517751E-02$	$1.274168E-01$	$4.507904E-03$
6	$6.331455E+00$	$2.518569E-02$	$1.276143E-01$	$5.349343E-04$
Fixed Point Method				
4	$6.309764E+00$	$2.479951E-02$	$1.264156E-01$	$2.708711E-02$
5	$6.329122E+00$	$2.513629E-02$	$1.274413E-01$	$3.325034E-03$
6	$6.331499E+00$	$2.517993E-02$	$1.276302E-01$	$3.927158E-04$
Ref [37] : *Drag* $C_{_D} = 6.33151E+00$, *Lift* $C_{_L} = 2.518E-02$				

Table 5.9: Shear and pressure dependent viscosity with weighted function $w^{1.0}$

The global mass conservation is calculated at each level and it is worth noting that in all the cases mentioned above the global mass conservation of the system is very well-preserved. It can be seen from the Table 5.10 that the performance of Gauss-Newton method in term of nonlinear iterations is better than that of fixed point method. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Lev.	$r = 1,\ \alpha_p = 0$	$r = 0,\ \alpha_p = 0.1$	$r = 1,\ \alpha_p = 0.1$
Gauss − Newton Method with $w^{1.0}_{\eta_\nu}$			
2	9/7	5/7	9/6
3	6/15	4/20	6/15
4	5/21	4/23	5/21
5	6/23	7/25	6/23
Fixed Point Method with $w^{1.0}_{\eta_\nu}$			
2	15/6	7/6	15/6
3	11/10	5/13	11/10
4	10/18	5/23	10/18
5	9/23	7/25	9/23

Table 5.10: Comparison of nonlinear and linear iterrations

Now, we continue with the same configuration of parameters and repeat the performed test again for a different nonlinear weighted function. Here, the nonlinear weighted function under consideration is $w^{1.5}_{\eta_\nu}$. The purpose of doing this is to observe the effects of the weighted functions on the approximation of results and convergence of solver.

Viscosity depends on shear rate with $w^{1.5}_{\eta_\nu}$

The Table 5.11 represents the results of shear dependent nonlinear viscosity $\eta_\nu(\dot{\gamma}_{II})$ function for the following set of considered values

$$r = 1,\ \alpha_p = 0,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2}.$$

The nonlinear weighted function is defined as

$$w^{1.5}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{II})\right)^{1.5}.$$

which is a function of the shear dependent viscosity.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$6.303997E+00$	$2.511087E-02$	$1.265522E-01$	$5.829458E-02$
5	$6.324761E+00$	$2.512897E-02$	$1.274333E-01$	$6.330799E-03$
6	$6.326923E+00$	$2.511514E-02$	$1.275873E-01$	$7.225744E-04$
Fixed Point Method				
4	$6.308014E+00$	$2.457609E-02$	$1.266852E-01$	$4.383738E-02$
5	$6.325208E+00$	$2.507243E-02$	$1.275129E-01$	$4.713574E-03$
6	$6.326975E+00$	$2.510679E-02$	$1.276384E-01$	$5.328366E-04$
Ref [37] : *Drag* $C_D = 6.32691E+00$, *Lift* $C_L = 2.510E-02$				

Table 5.11: Shear dependent viscosity with weighted function $w^{1.5}$

Viscosity depends on pressure with $w^{1.5}_{\eta_\nu}$

The nonlinear viscosity function $\eta_\nu(\mathfrak{p})$ depends on the pressure variable and the parameters are as under

$$r = 1,\ \alpha_p = 0,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2},$$

and the weighted function is proposed as

$$w^{1.5}_{\eta_\nu} = \big(1/\eta_\nu(\mathfrak{p})\big)^{1.5}.$$

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$3.330658E+01$	$8.215899E-01$	$4.915623E-01$	$2.620138E-02$
5	$3.331923E+01$	$8.222500E-01$	$4.939690E-01$	$2.623409E-03$
6	$3.332081E+01$	$8.223538E-01$	$4.945947E-01$	$2.828488E-04$
Fixed Point Method				
4	$3.330765E+01$	$8.215850E-01$	$4.915738E-01$	$2.105093E-02$
5	$3.331933E+01$	$8.222498E-01$	$4.939700E-01$	$2.124295E-03$
6	$3.332082E+01$	$8.223538E-01$	$4.945948E-01$	$2.307281E-04$
Ref [37] : *Drag* $C_D = 3.331069E+01$, *Lift* $C_L = 8.2217E-01$				

Table 5.12: Pressure dependent viscosity with weighted function $w^{1.5}$

Viscosity depends on shear rate and pressure with $w^{1.5}_{\eta_\nu}$

The nonlinear viscosity function $\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p})$ depends on shear rate and pressure. The numerical test is performed for the set of parameters as

$$r = 1,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-2}$$

and the respective nonlinear weighted function is defined as

$$w^{1.5}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p})\right)^{1.5},$$

and the numerical results are presented in Table 5.13.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$6.308623E+00$	$2.518586E-02$	$1.266326E-01$	$5.845167E-02$
5	$6.329383E+00$	$2.520474E-02$	$1.275142E-01$	$6.347563E-03$
6	$6.331545E+00$	$2.519086E-02$	$1.276685E-01$	$7.245894E-04$
Fixed Point Method				
4	$6.312650E+00$	$2.465365E-02$	$1.267679E-01$	$4.385262E-02$
5	$6.329831E+00$	$2.514840E-02$	$1.275960E-01$	$4.715183E-03$
6	$6.331596E+00$	$2.518252E-02$	$1.277205E-01$	$5.330296E-04$
Ref [37] : *Drag* $C_D = 6.33151E+00$, *Lift* $C_L = 2.518E-02$				

Table 5.13: Shear and pressure dependent viscosity with weighted function $w^{1.5}$

The numerical tests are performed for the same set of parameters but with two different types of nonlinear weighted functions. It is observed that the change in the weighted function can slightly improved the results. The tables generated by nonlinear weighted functions $w^{1.5}_{\eta_\nu}$ give better approximation solutions than those produced by the weighted function $w^{1.0}_{\eta_\nu}$. It can be seen from the Table 5.14 that the change in the weighted functions also improve the performance of nonlinear and linear solvers. Also, the weighted function $w^{1.5}_{\eta_\nu}$ is helpful in production of more accurate approximations. The global mass conservation is computed for each numerical test and it is preserved as well. The comparison of solvers by their iterations is provided in Table 5.14. It is shown that the Gauss-Newton method approximate the problem more efficiently than that of the fixed point method. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Lev.	$r = 1,\ \alpha_p = 0$	$r = 0,\ \alpha_p = 0.1$	$r = 1,\ \alpha_p = 0.1$
Gauss − Newton Method with $w^{1.5}_{\eta_\nu}$			
2	10/4	5/3	10/4
3	6/5	3/9	6/5
4	4/8	3/18	4/9
5	3/16	3/21	3/16
Fixed Point Method with $w^{1.5}_{\eta_\nu}$			
2	15/2	7/3	15/2
3	12/3	5/6	12/3
4	10/5	5/13	10/5
5	8/11	4/22	8/11

Table 5.14: Comparison of nonlinear and averaged linear iterations of solvers

From Table 5.10 and 5.14, it is clearly seen that weighted functions improve the linear solver but still it is not satisfactory. This is because of the SVP formulation which does not lead to the differentially diagonal dominant matrix. Therefore, the smoother, which is SSOR-PCG, lacks to provide efficient smoothing sweeps and results into more linear iterations with mesh refinements. The Table 5.15 shows that the performance of linear solver can be improved by increasing the pre/post smoothing steps in the linear solver.

$r = 1,\ \alpha_p = 0.1,\ w^{1.5}_{\eta_\nu}$				
Lev	$SmSt - 4$	$SmSt - 8$	$SmSt - 16$	$SmSt - 32$
Gauss − Newton Method				
2	10/4	10/2	10/1	10/1
3	6/5	6/2	6/2	6/1
4	4/9	4/4	4/2	4/1
5	3/16	3/9	3/5	3/3
Fixed Point Method				
2	15/2	15/1	15/1	15/1
3	12/3	12/1	12/1	12/1
4	10/5	10/2	10/1	10/1
5	8/11	8/5	8/2	8/1

Table 5.15: Comparison of solvers for different smoothing steps

We have observed from previous conducted test cases for $\eta_1 = 10^{-2}$ that the results generated by weighted function $w^{1.5}_{\eta_\nu}$ are more optimal, therefore we use it for further work. Now, for further investigation we make our problem more complex and harder by considering the bounded viscosities between $\eta_1 = 10^{-3}$ and $\eta_1 = 10^{-1}$.

Viscosity depends on shear rate with $w^{1.5}_{\eta_\nu}$

The results are generated for the viscosity function depends on shear rate $\eta_\nu(\dot{\gamma}_{II})$ for the set of values

$$r = 1,\ \alpha_p = 0,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

and

$$w^{1.5}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{II})\right)^{1.5},$$

is the nonlinear weight.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$1.492521E+01$	$1.378197E-01$	$2.609283E-01$	$6.584484E-02$
5	$1.494282E+01$	$1.377123E-01$	$2.608543E-01$	$7.300716E-03$
6	$1.494457E+01$	$1.376976E-01$	$2.600338E-01$	$8.549437E-04$
Fixed Point Method				
4	$1.492936E+01$	$1.372102E-01$	$2.615617E-01$	$7.991103E-02$
5	$1.494332E+01$	$1.376342E-01$	$2.612626E-01$	$8.184891E-03$
6	$1.494463E+01$	$1.376915E-01$	$2.602365E-01$	$8.776746E-04$
Ref [37] : *Drag* $C_D = 1.494420E+01$, *Lift* $C_L = 1.3769E-01$				

Table 5.16: Shear dependent viscosity with weighted function $w^{1.5}$

Viscosity depends on pressure with $w^{1.5}_{\eta_\nu}$

The results nonlinear viscosity is pressure dependent $\eta_\nu(\mathfrak{p})$ and test performed for the parameters

$$r = 0,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

and

$$w^{1.5}_{\eta_\nu} = \left(1/\eta_\nu(\mathfrak{p})\right)^{1.5},$$

is nonlinear weighted function. The flow parameters are computed for flow around cylinder problem are shown in Table 5.17.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
3	$4.212870E+02$	$4.392329E+00$	$5.054963E+00$	$2.572146E+01$
4	$5.164601E+02$	$6.171659E+00$	$7.005304E+00$	$4.579241E+00$
5	$5.365891E+02$	$6.560752E+00$	$7.745609E+00$	$4.195456E-01$
Fixed Point Method				
3	$4.738928E+02$	$5.138614E+00$	$5.597942E+00$	$6.469141E+00$
4	$5.302715E+02$	$6.402106E+00$	$7.174572E+00$	$7.764225E-01$
5	$5.379536E+02$	$6.584248E+00$	$7.763042E+00$	$6.322954E-02$
Ref [37] : *Drag* $C_D = 5.358480E+02$, *Lift* $C_L = 6.56621E+00$				

Table 5.17: Pressure dependent viscosity with weighted function $w^{1.5}$

Viscosity depends on shear rate and pressure with $w^{1.5}_{\eta_\nu}$

The considered nonlinear viscosity function $\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p})$ is shear rate and pressure-dependent. and the simulations are performed for parameters

$$r = 1,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

and nonlinear weighted function in this case is defined as

$$w^{1.5}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p})\right)^{1.5},$$

The Table 5.18 represents the approximated fluid flow values for different parameters.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$1.517453E+01$	$1.400115E-01$	$2.651726E-01$	$6.914758E-02$
5	$1.519248E+01$	$1.398618E-01$	$2.650040E-01$	$7.618641E-03$
6	$1.519426E+01$	$1.398432E-01$	$2.641678E-01$	$8.866902E-04$
Fixed Point Method				
4	$1.517883E+01$	$1.393550E-01$	$2.657012E-01$	$8.128097E-02$
5	$1.519299E+01$	$1.397783E-01$	$2.654335E-01$	$8.319800E-03$
6	$1.519432E+01$	$1.398367E-01$	$2.643736E-01$	$8.908413E-04$
Ref [37] : *Drag* $C_D = 1.519262E+01$, *Lift* $C_L = 1.3982E-01$				

Table 5.18: Shear and pressure dependent viscosity with weighted function $w^{1.5}$

In term of nonlinear iterations the Gauss-Newton method perform better than that of fixed point method as given in the Table 5.19. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Level	$r = 1,\ \alpha_p = 0$	$r = 0,\ \alpha_p = 0.1$	$r = 1,\ \alpha_p = 0.1$
Gauss − Newton Method with $w^{1.5}_{\eta_\nu}$			
2	11/4	15/18	11/4
3	7/9	10/24	7/9
4	5/21	9/24	5/21
5	6/23	18/25	6/23
Fixed Point Method with $w^{1.5}_{\eta_\nu}$			
2	39/2	8/10	39/2
3	29/4	9/23	29/4
4	24/8	10/24	24/9
5	19/18	25/25	19/18

Table 5.19: Comparison of nonlinear and linear iterations

Now, we continue with the same configuration of parameters and repeat the test for $w^{1.5}_{\eta_\nu}$ again with a different nonlinear weighted function. Here, the nonlinear weighted function under consideration is $w^{2.0}_{\eta_\nu}$. The purpose of this numerical investigations is to observe the effects of the weighted functions on the approximation of results and convergence of the solvers.

Viscosity depends on shear rate with $w^{2.0}_{\eta_\nu}$

Now we simulate the problem with the nonlinear viscosity function $\eta_\nu(\dot{\gamma}_{II})$ for the set of given values

$$r = 1,\ \alpha_p = 0,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

and the nonlinear weighted function

$$w^{2.0}_{\eta_\nu} = \left(1/\eta_\nu(\dot{\gamma}_{II})\right)^{2.0}.$$

The aim of test is to observe the behavior of solvers for different nonlinear weight. The Table 5.20 represent the results for shear dependent viscosity and nonlinear weight $w^{2.0}_{\eta_\nu}$.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$1.487019E+01$	$1.370977E-01$	$2.644543E-01$	$2.728245E-01$
5	$1.493569E+01$	$1.377148E-01$	$2.617246E-01$	$3.441961E-02$
6	$1.494370E+01$	$1.376949E-01$	$2.604787E-01$	$4.233133E-03$
Fixed Point Method				
4	$1.487101E+01$	$1.354280E-01$	$2.618610E-01$	$3.857796E-01$
5	$1.493608E+01$	$1.375135E-01$	$2.615512E-01$	$4.427182E-02$
6	$1.494378E+01$	$1.376783E-01$	$2.605034E-01$	$5.023333E-03$
Ref [37] : *Drag* $C_D = 1.494420E+01$, *Lift* $C_L = 1.3769E-01$				

Table 5.20: Shear dependent viscosity with weighted function $w^{2.0}$

Viscosity depends on pressure with $w^{2.0}_{\eta_\nu}$

The outcomes for pressure dependent viscosity function $\eta_\nu(\mathfrak{p})$ under the given parameters

$$r = 0,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

and nonlinear weighted function

$$w^{2.0}_{\eta_\nu} = \big(1/\eta_\nu(\mathfrak{p})\big)^{2.0},$$

are presented in the Table 5.21.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
3	$4.142216E+02$	$4.259511E+00$	$4.986134E+00$	$2.784193E+01$
4	$5.153436E+02$	$6.149515E+00$	$7.054644E+00$	$4.888113E+00$
5	$5.365108E+02$	$6.559204E+00$	$7.767027E+00$	$4.422426E-01$
Fixed Point Method				
3	$4.686562E+02$	$5.025892E+00$	$5.538630E+00$	$7.232882E+00$
4	$5.298441E+02$	$6.391330E+00$	$7.232002E+00$	$8.396302E-01$
5	$5.379280E+02$	$6.583574E+00$	$7.784692E+00$	$6.758960E-02$
Ref [37] : *Drag* $C_D = 5.358480E+02$, *Lift* $C_L = 6.56621E+00$				

Table 5.21: Pressure dependent viscosity with weighted function $w^{2.0}$

Viscosity depends on shear rate and pressure with $w^{2.0}_{\eta_\nu}$

The results for shear and pressure dependent viscosity function $\eta_\nu\big(\dot{\gamma}_{_{II}}, \mathfrak{p}\big)$ and for given values

$$r = 1,\ \alpha_p = 0.1,\ \eta_0 = 10^{-3},\ \eta_1 = 10^{-1},$$

with nonlinear weight function

$$w^{2.0}_{\eta_\nu} = \big(1/\eta_\nu\big(\dot{\gamma}_{_{II}}, \mathfrak{p}\big)\big)^{2.0}.$$

are presented in Table 5.22.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$1.511920E+01$	$1.393265E-01$	$2.687517E-01$	$2.897741E-01$
5	$1.518534E+01$	$1.398707E-01$	$2.659321E-01$	$3.621753E-02$
6	$1.519338E+01$	$1.398408E-01$	$2.646239E-01$	$4.423254E-03$
Fixed Point Method				
4	$1.512010E+01$	$1.376272E-01$	$2.660693E-01$	$3.912867E-01$
5	$1.518572E+01$	$1.396613E-01$	$2.657554E-01$	$4.480156E-02$
6	$1.519347E+01$	$1.398238E-01$	$2.646467E-01$	$5.074837E-03$
Ref [37] : *Drag* $C_{_D} = 1.519262E+01$, *Lift* $C_{_L} = 1.3982E-01$				

Table 5.22: Shear and pressure dependent viscosity with weighted function $w^{2.0}$

From above discussed cases, it is observed that the weighted function $w^{2.0}_{\eta_\nu}$ play an important role to get approximated results and better convergence behavior. The Table 5.23 represents a comparison of nonlinear iterations and averaged linear iterations taken by solvers. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Level	$r=1, \alpha_p=0$	$r=0, \alpha_p=0.1$	$r=1, \alpha_p=0.1$
Gauss − Newton Method with $w^{2.0}_{\eta_\nu}$			
2	14/8	13/11	14/8
3	13/7	10/23	12/7
4	5/15	8/24	5/15
5	5/22	19/24	5/22
Fixed Point Method with $w^{2.0}_{\eta_\nu}$			
2	40/3	8/7	40/3
3	31/3	9/19	32/3
4	27/4	9/24	26/4
5	20/8	19/24	21/7

Table 5.23: Comparison of nonlinear/linear iterations

From Table 5.19 and 5.23, it is clearly observed that the nonlinear weighted functions $w^{1.5}_{\eta_\nu}$ and $w^{2.0}_{\eta_\nu}$ improve the linear solver but still it lack in performance. The reason for this issue is the SVP formulation which does not leads to the differentially diagonal dominant matrix. Therefore, the smoother, which is SSOR-PCG, needs more smoothing steps to provide efficient smoothing sweeps and results into more linear iterations with mesh refinements. In the following Table 5.24, the results for different pre and post smoothing steps are calculated. It shows that the linear solver's convergence can be improved by increasing smoothing step size.

$r=1,\ \alpha_p=0.1,\ \eta_1=10^{-1},\ w^{2.0}_{\eta_\nu}$				
Level	$SmSt-4$	$SmSt-8$	$SmSt-16$	$SmSt-32$
Gauss − Newton Method				
2	14/8	14/4	14/2	15/1
3	12/7	12/4	12/2	13/1
4	5/15	5/8	5/4	5/2
5	5/22	5/19	5/11	5/6
Fixed Point Method				
2	40/3	40/1	40/1	40/1
3	32/3	31/2	32/1	32/1
4	26/4	27/2	26/1	26/1
5	21/7	20/3	20/2	20/1

Table 5.24: Comparison of nonlinear/linear iterations taken by the solvers for smoothing steps 4, 8, 16 and 32

5.8 Summary

The nonlinear weighted least-squares finite element method (N_w-LSFEM) has been proposed for the solutions of power law fluids and Cross law fluids. The stress, velocity and pressure (SVP) based first order system formulation is employed for the transformation of second order systems. From the previous chapter, it is observed that lower order finite elements are less accurate and lack global mass conservation. Therefore, the higher order finite element Q_2 is used for the discretizations of all unknown variables. The least-squares principle leads to symmetric and positive definite matrix. The conjugate gradient method is the best choice to solve linear system in such a situation. Additionally, the conjugate gradient is preconditioned by multigrid solver. Furthermore, the SSOR preconditioned conjugate gradient method is employed as smoother for multigrid method. But, the performance of linear solver needs more improvement because the coefficient matrix is not differentially diagonal dominant.

The numerical discussion is divided into two parts. The first part of discussion is composed of results related to power law viscosity model and the second part deals with the Cross law viscosity model. The power law model (3.43) is employed to study the effects of unbounded viscosities for non-Newtonian fluids. The power law viscosity model depends only on shear rate of the fluid flow. Therefore, the shear thinning and shear thickening effects are investigated for different values of power law index r. The fluid flow quantities, such as drag, lift and pressure gradient, are obtained with great accuracy for flow around cylinder problem. It is observed that with parametric value $r = 1.5$ the fluid becomes less viscous and shear thinning features are being seen. The value $r = 2.0$ corresponds to the Newtonian fluid and for $r = 2.5$, 3.0 the fluid has behaved as shear thickening fluid as it becomes more viscous. Moreover, the nonlinear weighted functions $w_{\eta_\nu}^{1.0}$ are observed to be mandatory for better convergence of the solutions. The grid independent convergence has been achieved for both nonlinear solvers. The comparison of nonlinear and averaged linear iterations taken by the solvers is presented as well. The performance of Gauss-Newton method is outstanding verses fixed point method.

The second part of our investigation is based on the Cross law viscosity model (3.44). We are the first to use this model for numerical investigations of least-squares finite element method. The Cross law viscosity model is used to discuss the fluid flow with bounded viscosities. The nonlinear viscosities employed in this chapter depend on shear rate and pressure only. The flow parameters for flow around cylinder problem are explored for bounded viscosities between $\eta_0 = 0.001$ and $\eta_1 = 0.01$, 0.1. The numerical outcomes are compared with the solutions of Damanik [37]. A variety of nonlinear weighted functions $w_{\eta_\nu}^{n}$; $n = 1.0$, 1.5, 2.0, which depend on nonlinear viscosities $\eta_\nu(\dot{\gamma}_{II})$, $\eta_\nu(\mathfrak{p})$ and $\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p})$, are used for the calculations of flow parameters. We have observed that the parametric values n in the weighted functions increase with the

increase in the complexity of the problem. The solvers have shown gird independent behavior towards the solutions.

The mass conservation of the inflow-outflow systems is another important topic in the study of least-squares finite element method. The lack of global mass conservation in the flow problem is the source of motivation for us to consider flow around cylinder problem for our numerical investigations. Therefore, we have computed global mass conservation, at the outflow cross-section $x = 2.2$ of considered domain, in almost each test case. In the power law fluids the global mass conservation is achieved in each test case with higher refinement levels. For the Cross law fluids, the global mass conservation for fluids, in which the viscosities depend on shear rate $\eta_\nu\big(\dot{\gamma}_{_{II}}\big)$, shear rate and pressure $\eta_\nu\big(\dot{\gamma}_{_{II}}, \mathfrak{p}\big)$, is very promising. But in the case when viscosity depends only on pressure $\eta_\nu\big(\mathfrak{p}\big)$ the global mass conservation is achieved at a slow pace even on higher refinement levels. The iterative comparison is studied for the nonlinear and linear solvers and the independence of the solver w.r.t. mesh refinement are observed.

We can conclude that the nonlinear weighted least-squares finite element method performs very well overall for both nonlinear viscosity models. The convergence behavior of Gauss-Newton method is very promising over fixed point method. The linear solver has performed very well and the grid independent convergence is obtained for our numerical investigations with very good accuracy and robustness.

Chapter 6

Non-isothermal non-Newtonian Fluid Flows

The non-Newtonian fluids with non-isothermal effects, for the Cross law (3.44) fluid flow problems, are considered in this chapter. We develop a physically motivated nonlinear weighted least-squares finite element method to generate accurate and robust solutions for the non-Newtonian fluid flows. We implement a MPCG solver to efficiently approximate the newly proposed SVPT formulation for the non-Newtonian fluid flow models.

6.1 Introduction

The non-Newtonian fluids have been explored for last few decades and most of the research work have been carried out for power law fluids with nonlinear viscosity. The p-version LSFEM was developed by [10], [46], [47] for two dimensional incompressible non-Newtonian fluids along with isothermal and non-isothermal conditions. The viscosity function dependent on shear rate and temperature with power law index between $0.25 \leq n \leq 1.5$ was used for numerical investigations of Newtonian and generalized Newtonian fluid flow.

We designed special nonlinear weighted least-squares finite element techniques for the solution of stationary incompressible NS equations coupled with energy equation. The Cross law fluids with nonlinear viscosity depending on shear rate, pressure and temperature is employed for numerical investigations. The first order formulation SVPT is the composite of two additional auxiliary variables stress and temperature gradient along with velocity-pressure-temperature. The MPCG solver is tested for discrete linear system discretized with higher equal order finite elements.

6.2 System of governing equations

In this section, we consider the system of equations on a bounded domain $\Omega \subset \mathbb{R}^2$ for the steady problem of generalized Newtonian flow which is similar to the system studied in section 5.2. But the main difference here is the nonlinear viscosity depends also on temperature of the fluid. The non-isothermal effects of the flow are given in terms of the Boussinesq approximation where the coupling with the Navier Stokes problem is written as

$$\left.\begin{aligned} \mathfrak{u}\cdot\nabla\mathfrak{u}+\nabla\mathfrak{p}-\nabla\cdot\Big(2\eta_{\nu}(\dot{\gamma}_{_{II}},\mathfrak{p},\boldsymbol{\theta})\mathsf{D}\left(\mathfrak{u}\right)\Big)&=\mathfrak{f}\\ \nabla\cdot\mathfrak{u}&=0\\ \big(\mathfrak{u}\cdot\nabla\big)\boldsymbol{\theta}-k_1\nabla^2\boldsymbol{\theta}-k_2\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big]&=0\end{aligned}\right\}\ \text{in }\Omega, \tag{6.1}$$

where $\mathfrak{u}$ is the velocity of the flowing fluid, $\mathfrak{p}$ is the normalized pressure,$\boldsymbol{\theta}$ is the temperature of the flowing fluid, $\mathfrak{f}$ is the external source term, $\eta_{\nu}\left(\cdot\right)$ is the nonlinear viscosity depending on shear rate, pressure and temperature, $\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big]$ is the additional viscous dissipation term due to fluid friction among flowing particles, k_1 and k_2 are thermal diffusively and friction parameter.

6.3 First order system formulation

To derive the stress-velocity-pressure-temperature (SVPT) formulation, we introduce the Cauchy stress tensor $\boldsymbol{\sigma}$ as a new variable defined as

$$\begin{aligned}\boldsymbol{\sigma} &= \mathbf{v}_{\boldsymbol{\sigma}}-\mathbf{p}_{\boldsymbol{\sigma}},\\ &= 2\eta_{\nu}(\dot{\gamma}_{_{II}},\mathfrak{p},\boldsymbol{\theta})\mathsf{D}\left(\mathfrak{u}\right)-\mathfrak{p}\mathsf{I},\end{aligned} \tag{6.2}$$

where $\mathbf{p}_{\boldsymbol{\sigma}}=\mathfrak{p}\mathsf{I}$ is called the inviscid reactive part and $\mathbf{v}_{\boldsymbol{\sigma}}=2\eta_{\nu}(\dot{\gamma}_{_{II}},\mathfrak{p},\theta)\mathsf{D}(\mathfrak{u})$ is the active viscous part of the Cauchy stress tensor $\boldsymbol{\sigma}$, which is a function of deformation gradient of velocity $\mathsf{D}(\mathfrak{u})$ and the second invariant of the shear rate tensor $\dot{\gamma}_{_{II}}$ defined in section 5.3.

Also, we introduce temperature gradient $\boldsymbol{\Theta}$ as another new variable which is defined as

$$\boldsymbol{\Theta}=k_1\nabla\boldsymbol{\theta}. \tag{6.3}$$

Using equations (6.2) and (6.3) in (6.1) , the first order Stress-Velocity-Pressure-Temperature (SVPT) system of equations obtained on a bounded domain $\Omega \subset \mathbb{R}^2$ is given as

$$\left.\begin{aligned} \mathfrak{u}\cdot\nabla\mathfrak{u}-\nabla\cdot\boldsymbol{\sigma} &= \mathfrak{f}\\ \nabla\cdot\mathfrak{u} &= 0\\ \boldsymbol{\sigma}+\mathbf{p}_{\boldsymbol{\sigma}}-\mathbf{v}_{\boldsymbol{\sigma}} &= 0\\ \big(\mathfrak{u}\cdot\nabla\big)\boldsymbol{\theta}-\nabla\cdot\boldsymbol{\Theta}-k_2\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big] &= 0\\ \boldsymbol{\Theta}-k_1\nabla\boldsymbol{\theta} &= 0 \end{aligned}\right\} \text{ in } \Omega, \tag{6.4}$$

The Dirichlet boundary conditions $f_{_{u,\hat{\Omega}_D}}$, $f_{_{\boldsymbol{\theta},\hat{\Omega}_D}}$ on the Dirichlet boundary region $\hat{\Omega}_{_D}$

$$\left.\begin{aligned} \mathfrak{u} &= f_{_{u,\hat{\Omega}_D}}\\ \boldsymbol{\theta} &= f_{_{\boldsymbol{\theta},\hat{\Omega}_D}} \end{aligned}\right\} \text{ on } \hat{\Omega}_{_D}, \tag{6.5}$$

and the Neumann boundary conditions $f_{_{u,\hat{\Omega}_N}}$, $f_{_{\boldsymbol{\theta},\hat{\Omega}_N}}$ on the outward unit normal $\mathbf{n}$ of Neumann boundary region $\hat{\Omega}_{_D}$ is given as

$$\left.\begin{aligned} \mathbf{n}\cdot\boldsymbol{\sigma} &= f_{_{u,\hat{\Omega}_N}}\\ \mathbf{n}\cdot\boldsymbol{\Theta} &= f_{_{\boldsymbol{\theta},\hat{\Omega}_N}} \end{aligned}\right\} \text{ on } \hat{\Omega}_{_N}. \tag{6.6}$$

6.4 Linearization of system

We first linearize the nonlinear terms present in the considered system (6.4) . In the given problem there are four nonlinear terms, first is the convection term $\mathfrak{u}\cdot\nabla\mathfrak{u}$, the second is due to the nonlinear viscosity in the term $\mathbf{v}_{\boldsymbol{\sigma}}=2\eta_\nu\big(\dot{\gamma}_{_{II}},\mathfrak{p},\boldsymbol{\theta}\big)\mathsf{D}\big(\mathfrak{u}\big)$, the third is coupled term $\big(\mathfrak{u}\cdot\nabla\big)\boldsymbol{\theta}$ and Fourth nonlinear term is friction term $k_2\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big]$. We apply the Newton's linearization to the nonlinear terms before the application of least-squares principle to the problem. Therefore, the following substitutions are made for the nonlinear terms

$$\mathfrak{u}\cdot\nabla\mathfrak{u}\approx\mathfrak{u}\cdot\nabla\mathfrak{u}^k+\mathfrak{u}^k\cdot\nabla\mathfrak{u}, \tag{6.7}$$

and

$$\mathbf{v}_{\boldsymbol{\sigma}}\approx\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{p}\big)+\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\boldsymbol{\theta}\big)+\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big)+\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}\big(\mathfrak{u}\big), \tag{6.8}$$

$$\begin{aligned} 2\eta_\nu\big(\dot{\gamma}_{_{II}},\mathfrak{p},\boldsymbol{\theta}\big)\mathsf{D}\big(\mathfrak{u}\big) \;\approx\; & 2\frac{\partial\eta_\nu}{\partial\mathfrak{p}}\mathsf{D}\big(\mathfrak{u}^k\big)\mathfrak{p}+2\frac{\partial\eta_\nu}{\partial\boldsymbol{\theta}}\mathsf{D}\big(\mathfrak{u}^k\big)\boldsymbol{\theta}+2\eta_\nu\big(\dot{\gamma}^k_{_{II}},\mathfrak{p}^k,\boldsymbol{\theta}^k\big)\mathsf{D}\big(\mathfrak{u}\big)\\ & +\frac{2}{\dot{\gamma}_{_{II}}}\frac{\partial\eta_\nu}{\partial\dot{\gamma}_{_{II}}}\Big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}^k\big)\Big]\,\mathsf{D}\big(\mathfrak{u}^k\big), \end{aligned} \tag{6.9}$$

and

$$\big(\mathfrak{u}\cdot\nabla\big)\boldsymbol{\theta}\approx\big(\mathfrak{u}\cdot\nabla\big)\boldsymbol{\theta}^k+\big(\mathfrak{u}^k\cdot\nabla\big)\boldsymbol{\theta}, \tag{6.10}$$

and

$$k_2\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big]\approx k_2\big[\mathsf{D}\big(\mathfrak{u}\big)\!:\!\mathsf{D}\big(\mathfrak{u}^k\big)\big]+k_2\big[\mathsf{D}\big(\mathfrak{u}^k\big)\!:\!\mathsf{D}\big(\mathfrak{u}\big)\big]. \tag{6.11}$$

Let $\mathcal{R}$ be the sum of residuals for the system of equations (6.4). To approximate the residuals, the nonlinear iteration is updated with the correction $\mathbf{u}^{k+1} = \mathbf{u}^k + \delta\mathbf{u}$. Then, the residual approximation is obtained by Taylor series expansion as

$$\begin{aligned} \mathcal{R}(\mathbf{u}^{k+1}) &= \mathcal{R}(\mathbf{u}^k + \delta\mathbf{u}), \\ &\simeq \mathcal{R}(\mathbf{u}^k) + \big[\mathcal{D}\mathcal{R}(\mathbf{u}^k)\big]\delta\mathbf{u}, \end{aligned} \tag{6.12}$$

where

$$\mathcal{R}(\mathbf{u}^k) = \begin{bmatrix} \mathcal{R}_1^k \\ \mathcal{R}_2^k \\ \mathcal{R}_3^k \\ \mathcal{R}_4^k \\ \mathcal{R}_5^k \end{bmatrix} = \begin{bmatrix} \mathfrak{u}^k \cdot \nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\sigma}^k - \mathfrak{f} \\ \nabla\cdot\mathfrak{u}^k \\ \boldsymbol{\sigma}^k + \mathbf{p}_{\boldsymbol{\sigma}}(\mathbf{u}^k) - \mathbf{v}_{\boldsymbol{\sigma}}(\mathbf{u}^k) \\ (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}^k - \nabla\cdot\boldsymbol{\Theta}^k - k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u}^k)\big] \\ \boldsymbol{\Theta}^k - k_1\nabla\boldsymbol{\theta}^k = 0 \end{bmatrix}, \tag{6.13}$$

and the system of equations after substitution is as following

$$\mathcal{D}\mathcal{R}(\mathbf{u}^k)\delta\mathbf{u} = \begin{bmatrix} \mathfrak{u}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} - \nabla\cdot\boldsymbol{\sigma} - \mathfrak{f} \\ \nabla\cdot\mathfrak{u} \\ \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}) - \mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}) \\ (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k + (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta} - \nabla\cdot\boldsymbol{\Theta} - k_2\big[\mathsf{D}(\mathfrak{u}):\mathsf{D}(\mathfrak{u}^k)\big]\cdot\cdots \\ \qquad\qquad - k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u})\big] \\ \boldsymbol{\Theta} - k_1\nabla\boldsymbol{\theta} \end{bmatrix}. \tag{6.14}$$

By following the procedure discussed in Algorithm 3.2, i.e. the nonlinear basic iteration for least-squares scheme, one can obtain a solution update $\delta\mathbf{u}$ as

$$\tilde{\mathcal{R}}(\mathbf{u}^k)\delta\mathbf{u} = \mathcal{R}_{def}(\mathbf{u}^k), \tag{6.15}$$

and for different choices of $\tilde{\mathcal{R}}$ and $\mathcal{R}_{def}$, as discussed in section 4.5, one can obtain different approximation schemes. We use Gauss-Newton and Fixed point scheme in this thesis.

6.5 Continuous least-squares principle

In this section, the $\mathcal{L}^2$-norm least-squares energy functional based on the residual (6.12) is defined. Based on the Newton approximation, we define the functional as

$$\mathcal{E}_{svpt}(\delta\mathbf{u};0) = \frac{1}{2}\big\|\mathcal{R}(\mathbf{u}^k) + \big[\mathcal{D}\mathcal{R}(\mathbf{u}^k)\big]\delta\mathbf{u}\big\|_0^2, \tag{6.16}$$

and the unweighted functional for the SVPT problem has following form

$$\begin{aligned}\mathcal{E}_{svpt} &:= \tfrac{1}{2}\big[\big\|\mathcal{R}_1^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} + \mathfrak{u}\cdot\nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\sigma}\big\|_0^2 + \big\|\mathcal{R}_2^k + \nabla\cdot\mathfrak{u}\big\|_0^2 \\ &+\big\|\mathcal{R}_3^k + \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}) - \mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u})\big\|_0^2 \\ &+\big\|\mathcal{R}_4^k + (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta} + (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k - \nabla\cdot\boldsymbol{\Theta} - k_2\big[\mathsf{D}(\mathfrak{u}):\mathsf{D}(\mathfrak{u}^k)\big] \\ &-k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u})\big]\big\|_0^2 + \big\|\mathcal{R}_5^k + \boldsymbol{\Theta} - k_1\nabla\boldsymbol{\theta}\big\|_0^2\big] \quad \forall\ (\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta}) \in \mathcal{X}_\Omega\end{aligned} \tag{6.17}$$

and the weighted form of the least-squares functional is

$$\begin{aligned}\mathcal{E}_{svpt}^{w_{\eta_\nu}} &:= \tfrac{1}{2}\big[\big\|\mathcal{R}_1^k + \mathfrak{u}^k\cdot\nabla\mathfrak{u} + \mathfrak{u}\cdot\nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\sigma}\big\|_0^2 + \alpha_m\big\|\mathcal{R}_2^k + \nabla\cdot\mathfrak{u}\big\|_0^2 \\ &+w_{\eta_\nu}^n\big\|\mathcal{R}_3^k + \boldsymbol{\sigma} + \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}) - \mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u})\big\|_0^2 \\ &+\big\|\mathcal{R}_4^k + (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta} + (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k - \nabla\cdot\boldsymbol{\Theta} - k_2\big[\mathsf{D}(\mathfrak{u}):\mathsf{D}(\mathfrak{u}^k)\big] \\ &-k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u})\big]\big\|_0^2 + w_{\eta_\nu}^n\big\|\mathcal{R}_5^k + \boldsymbol{\Theta} - k_1\nabla\boldsymbol{\theta}\big\|_0^2\big] \quad \forall\ (\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta}) \in \mathcal{X}_\Omega\end{aligned} \tag{6.18}$$

where nonlinear weight

$$w_{\eta_\nu}^n = \big(1/\eta_\nu(\dot{\gamma}_{_{II}},\mathfrak{p},\boldsymbol{\theta})\big)^n;\ n \geq 2.0,$$

is a function of nonlinear viscosity and $\mathcal{X}_\Omega$ is the space of admissible functions

$$\mathcal{X}_\Omega := \Big\{\mathcal{H}_{0,D}^1(\Omega)\times\mathcal{L}_0^2(\Omega)\times\mathcal{H}_{0,N}(\mathrm{div},\Omega)\times\mathcal{H}_{0,D}^1(\Omega)\times\mathcal{H}_{0,N}(\mathrm{div},\Omega)\Big\}. \tag{6.19}$$

The minimization problem associated with the weighted functional (6.18) is defined, such that

$$\begin{cases}\text{Find } \delta\mathbf{u} = (\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta}) \in \mathcal{X}_\Omega \text{ such that,} \\ (\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta}) = \underset{(v,q,\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}})\in\mathcal{X}_\Omega}{\arg\min}\ \mathcal{E}_{svpt}^{w_{\eta_\nu}}(\mathfrak{v},\mathfrak{q},\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}}),\end{cases} \tag{6.20}$$

where $\mathbf{v} = (\mathfrak{v},\mathfrak{q},\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}})$.Then, the variational problem at each nonlinear iteration step based on the optimality condition of the problem(6.20) is to find $(\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta}) \in \mathcal{X}_\Omega$ such that

$$\mathbf{K}\big((\mathfrak{u},\mathfrak{p},\boldsymbol{\sigma},\boldsymbol{\theta},\boldsymbol{\Theta});(\mathfrak{v},\mathfrak{q},\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}})\big) = \mathbf{F}(\mathfrak{v},\mathfrak{q},\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}}) \quad \forall\ (\mathfrak{v},\mathfrak{q},\boldsymbol{\tau},\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\Theta}}) \in \mathcal{X}_\Omega, \tag{6.21}$$

where $\mathbf{K}$ is the bilinear form defined on $\mathcal{X}_\Omega\times\mathcal{X}_\Omega \to \mathbb{R}$ as follows

$$\begin{aligned}\mathbf{K} &:= \big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u} + \mathfrak{u}\cdot\nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\sigma},\ \mathfrak{u}^k\cdot\nabla\mathfrak{v} + \mathfrak{v}\cdot\nabla\mathfrak{u}^k - \nabla\cdot\boldsymbol{\tau}\big\rangle \\ &+ \alpha_m\big\langle\nabla\cdot\mathfrak{u},\nabla\cdot\mathfrak{v}\big\rangle + w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma},\boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle \\ &+ w_{\eta_\nu}^n\big\langle\boldsymbol{\sigma},-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) - \mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + w_{\eta_\nu}^n\big\langle\mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),\boldsymbol{\tau}\big\rangle \\ &+ w_{\eta_\nu}^n\big\langle\mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w_{\eta_\nu}^n\big\langle\mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle\end{aligned} \tag{6.22}$$

$$
\begin{aligned}
&+ w_{\eta_\nu}^n \big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle + w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), \boldsymbol{\tau} \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}), \boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) - \mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle + w_{\eta_\nu}^n \big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) - \mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) - \mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle + \big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, (\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}} \big\rangle \\
&+ \big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k \big\rangle + \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, (\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}} \big\rangle \\
&+ \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k \big\rangle + \big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, -\nabla\cdot\tilde{\boldsymbol{\Theta}} \big\rangle \\
&+ \big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle + \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, -\nabla\cdot\tilde{\boldsymbol{\Theta}} \big\rangle \\
&+ \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle + \big\langle -\nabla\cdot\boldsymbol{\Theta}, (\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}} \big\rangle \\
&+ \big\langle -\nabla\cdot\boldsymbol{\Theta}, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k \big\rangle + \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], (\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}} \big\rangle \\
&+ \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k \big\rangle + \big\langle -\nabla\cdot\boldsymbol{\Theta}, -\nabla\cdot\tilde{\boldsymbol{\Theta}} \big\rangle \\
&+ \big\langle -\nabla\cdot\boldsymbol{\Theta}, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle \\
&+ \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], -\nabla\cdot\tilde{\boldsymbol{\Theta}} \big\rangle \\
&+ \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big], -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle \\
&+ \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle + w_{\eta_\nu}^n \big\langle \boldsymbol{\Theta}, \tilde{\boldsymbol{\Theta}} \big\rangle \\
&+ w_{\eta_\nu}^n \big\langle \boldsymbol{\Theta}, -k_1\nabla\tilde{\boldsymbol{\theta}} \big\rangle + w_{\eta_\nu}^n \big\langle -k_1\nabla\boldsymbol{\theta}, \tilde{\boldsymbol{\Theta}} \big\rangle + w_{\eta_\nu}^n \big\langle -k_1\nabla\boldsymbol{\theta}, -k_1\nabla\tilde{\boldsymbol{\theta}} \big\rangle.
\end{aligned}
$$

and the linear form $\mathbf{F}$ is defined on $\mathcal{X}_\Omega \to \mathbb{R}$ as

$$
\begin{aligned}
\mathbf{F} \quad : \quad & = \big\langle \mathcal{R}_1^k, \mathfrak{v}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{v} - \nabla\cdot\boldsymbol{\tau} \big\rangle + \alpha_m \big\langle \mathcal{R}_2^k, \nabla\cdot\mathfrak{v} \big\rangle \\
& + w_{\eta_\nu}^n \big\langle \mathcal{R}_3^k, \boldsymbol{\tau} + \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}}) - \mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) - \mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v}) \big\rangle \\
& + \big\langle \mathcal{R}_4^k, (\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}} + (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k - \nabla\cdot\tilde{\boldsymbol{\Theta}} - k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] \\
& - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big] \big\rangle + w_{\eta_\nu}^n \big\langle \mathcal{R}_5^k, \tilde{\boldsymbol{\Theta}} - k_1\nabla\tilde{\boldsymbol{\theta}} \big\rangle.
\end{aligned} \tag{6.23}
$$

To analyze the properties of the least-square problem the operator form is as following

$$
\mathbf{K} \quad = \quad \Big[\mathcal{D}\mathcal{R}\Big(\mathbf{u}^k\Big)\Big]^* \Big[\mathcal{D}\mathcal{R}\Big(\mathbf{u}^{kk}\Big)\Big] \tag{6.24}
$$

$$= \begin{bmatrix} \mathbf{K}_{vu} & \mathbf{K}_{vp} & \mathbf{K}_{v\boldsymbol{\sigma}} & \mathbf{K}_{v\boldsymbol{\theta}} & \mathbf{K}_{v\Theta} \\ \mathbf{K}_{qu} & \mathbf{K}_{qp} & \mathbf{K}_{q\boldsymbol{\sigma}} & \mathbf{K}_{q\boldsymbol{\theta}} & \mathbf{0} \\ \mathbf{K}_{\boldsymbol{\tau}u} & \mathbf{K}_{\boldsymbol{\tau}p} & \mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\sigma}} & \mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\theta}} & \mathbf{0} \\ \mathbf{K}_{\bar{\boldsymbol{\theta}}u} & \mathbf{K}_{\bar{\boldsymbol{\theta}}p} & \mathbf{K}_{\bar{\boldsymbol{\theta}}\boldsymbol{\sigma}} & \mathbf{K}_{\bar{\boldsymbol{\theta}}\boldsymbol{\theta}} & \mathbf{K}_{\bar{\boldsymbol{\theta}}\Theta} \\ \mathbf{K}_{\bar{\Theta}u} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{\bar{\Theta}\boldsymbol{\theta}} & \mathbf{K}_{\bar{\Theta}\Theta} \end{bmatrix},$$

where

$$\begin{aligned}
\mathbf{K}_{vu} &= w^n_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + \big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle + \alpha_m\big\langle\nabla\cdot\mathfrak{u}, \nabla\cdot\mathfrak{v}\big\rangle \\
&\quad + \boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k, \mathfrak{v}\cdot\nabla\mathfrak{u}^k + \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle} + \boxed{\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + \boxed{\big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k\big\rangle} \\
&\quad + \boxed{\big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big]\big\rangle} \\
&\quad + \boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big], (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k\big\rangle} + \boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big], (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k\big\rangle} \\
&\quad + \boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}){:}\mathsf{D}(\mathfrak{u}^k)\big], -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big] - k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big]\big\rangle} \\
&\quad + \boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big]\big\rangle} \\
&\quad + \big\langle -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{u})\big], -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big]\big\rangle, \\
\mathbf{K}_{vp} &= w^n_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle}, \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle}, \\
\mathbf{K}_{v\boldsymbol{\sigma}} &= w^n_{\eta_\nu}\big\langle \boldsymbol{\sigma}, -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \boldsymbol{\sigma}, -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + \big\langle -\nabla\cdot\boldsymbol{\sigma}, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle \\
&\quad + \boxed{\big\langle -\nabla\cdot\boldsymbol{\sigma}, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle}, \\
\mathbf{K}_{v\theta} &= w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}), -\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{v})\big\rangle} \\
&\quad + \boxed{\big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k\big\rangle} + \boxed{\big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big]\big\rangle} \\
&\quad + \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta}, -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big]\big\rangle, \\
\mathbf{K}_{v\Theta} &= \boxed{\big\langle -\nabla\cdot\boldsymbol{\Theta}, (\mathfrak{v}\cdot\nabla)\boldsymbol{\theta}^k\big\rangle} + \boxed{\big\langle -\nabla\cdot\boldsymbol{\Theta}, -k_2\big[\mathsf{D}(\mathfrak{v}){:}\mathsf{D}(\mathfrak{u}^k)\big]\big\rangle} \\
&\quad + \big\langle -\nabla\cdot\boldsymbol{\Theta}, -k_2\big[\mathsf{D}(\mathfrak{u}^k){:}\mathsf{D}(\mathfrak{v})\big]\big\rangle,
\end{aligned}$$

$$\begin{aligned}
\mathbf{K}_{qu} &= w^n_{\eta_\nu}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} \\
&\quad + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle} + w^n_{\eta_\nu}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}, \\
\mathbf{K}_{qp} &= w^n_{\eta_\nu}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), \mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle + w^n_{\eta_\nu}\boxed{\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}), -\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}
\end{aligned}$$

$$
\begin{aligned}
&+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}),\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle},\\
\mathbf{K}_{q\boldsymbol{\sigma}} &= w_{\eta_\nu}^{n}\big\langle \boldsymbol{\sigma},\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle+w_{\eta_\nu}^{n}\boxed{\big\langle \boldsymbol{\sigma},-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle},\\
\mathbf{K}_{q\boldsymbol{\theta}} &= w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}),\mathbf{p}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle}+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\mathfrak{q})\big\rangle},
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{\boldsymbol{\tau}u} &= w_{\eta_\nu}^{n}\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),\boldsymbol{\tau}\big\rangle+\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u},-\nabla\cdot\boldsymbol{\tau}\big\rangle+\boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k,-\nabla\cdot\boldsymbol{\tau}\big\rangle}\\
&\quad +w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),\boldsymbol{\tau}\big\rangle},\\
\mathbf{K}_{\boldsymbol{\tau}p} &= w_{\eta_\nu}^{n}\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),\boldsymbol{\tau}\big\rangle+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}),\boldsymbol{\tau}\big\rangle},\\
\mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\sigma}} &= w_{\eta_\nu}^{n}\big\langle \boldsymbol{\sigma},\boldsymbol{\tau}\big\rangle+\big\langle -\nabla\cdot\boldsymbol{\sigma},-\nabla\cdot\boldsymbol{\tau}\big\rangle,\\
\mathbf{K}_{\boldsymbol{\tau}\boldsymbol{\theta}} &= w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}),\boldsymbol{\tau}\big\rangle},
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{\hat{\boldsymbol{\theta}}u} &= w_{\eta_\nu}^{n}\boxed{\big\langle -\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle}+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{u}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle}\\
&\quad +\boxed{\big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k,(\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}}\big\rangle}+\boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}):\mathsf{D}(\mathfrak{u}^k)\big],(\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}}\big\rangle}\\
&\quad +\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u})\big],(\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}}\big\rangle,\\
\mathbf{K}_{\hat{\theta}p} &= w_{\eta_\nu}^{n}\boxed{\big\langle \mathbf{p}_{\boldsymbol{\sigma}}(\mathfrak{p}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle}+w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\mathfrak{p}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle},\\
\mathbf{K}_{\hat{\theta}\boldsymbol{\sigma}} &= w_{\eta_\nu}^{n}\boxed{\big\langle \boldsymbol{\sigma},-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle},\\
\mathbf{K}_{\hat{\theta}\theta} &= w_{\eta_\nu}^{n}\boxed{\big\langle -\mathcal{D}\mathbf{v}_{\boldsymbol{\sigma}}(\boldsymbol{\theta}),-\mathcal{D}\mathbf{v}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{\theta}})\big\rangle}+\big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta},(\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}}\big\rangle\\
&\quad +w_{\eta_\nu}^{n}\big\langle -k_1\nabla\boldsymbol{\theta},-k_1\nabla\tilde{\boldsymbol{\theta}}\big\rangle,\\
\mathbf{K}_{\hat{\theta}\Theta} &= \big\langle -\nabla\cdot\boldsymbol{\Theta},(\mathfrak{u}^k\cdot\nabla)\tilde{\boldsymbol{\theta}}\big\rangle+w_{\eta_\nu}^{n}\big\langle \boldsymbol{\Theta},-k_1\nabla\tilde{\boldsymbol{\theta}}\big\rangle,
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{\hat{\Theta}u} &= \boxed{\big\langle (\mathfrak{u}\cdot\nabla)\boldsymbol{\theta}^k,-\nabla\cdot\tilde{\boldsymbol{\Theta}}\big\rangle}+\boxed{\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}):\mathsf{D}(\mathfrak{u}^k)\big],-\nabla\cdot\tilde{\boldsymbol{\Theta}}\big\rangle}\\
&\quad +\big\langle -k_2\big[\mathsf{D}(\mathfrak{u}^k):\mathsf{D}(\mathfrak{u})\big],-\nabla\cdot\tilde{\boldsymbol{\Theta}}\big\rangle,\\
\mathbf{K}_{\hat{\Theta}\theta} &= \big\langle (\mathfrak{u}^k\cdot\nabla)\boldsymbol{\theta},-\nabla\cdot\tilde{\boldsymbol{\Theta}}\big\rangle+w_{\eta_\nu}^{n}\big\langle -k_1\nabla\boldsymbol{\theta},\tilde{\boldsymbol{\Theta}}\big\rangle,\\
\mathbf{K}_{\hat{\Theta}\Theta} &= \big\langle -\nabla\cdot\boldsymbol{\Theta},-\nabla\cdot\tilde{\boldsymbol{\Theta}}\big\rangle+w_{\eta_\nu}^{n}\big\langle \boldsymbol{\Theta},\tilde{\boldsymbol{\Theta}}\big\rangle,
\end{aligned}
$$

Here, the terms shown in the boxes have contributions in the coefficient matrix $\mathbf{K}$ due to the nonlinear terms. These terms will be add to coefficient matrix $\mathbf{K}$ after defect calculation. The resulting system matrix (6.22) is symmetric and positive definite. It can be observed from system matrix that it is not differentially diagonal dominant which can effect the performance of linear solver. We are able to use the Krylov method to efficiently solve the system of equations. As, the system matrix is SPD matrix.

6.6 Discrete least-squares principle

Now we approximate the problem (6.21) with the finite element method to finite dimension space $\mathcal{X}_{\Omega}^{h}$ defined as

$$\mathcal{X}_{\Omega}^{h} := \Big\{ \delta\mathbf{u}^{h} \in \mathcal{H}_{0,D}^{1,h}(\Omega) \times \mathcal{L}_{0}^{2,h}(\Omega) \times \mathcal{H}_{0,N}^{h}(\mathrm{div},\Omega) \times \mathcal{H}_{0,D}^{1,h}(\Omega) \times \mathcal{H}_{0,N}^{h}(\mathrm{div},\Omega) \Big\}, \tag{6.25}$$

where $\delta\mathbf{u}^{h} = \big(\mathfrak{u}_{h}, \mathfrak{p}_{h}, \boldsymbol{\sigma}_{h}, \boldsymbol{\theta}_{h}, \boldsymbol{\Theta}_{h}\big)$ and consider the approximation form of variational problem as

$$\begin{cases} \text{Find } \delta\mathbf{u}^{h} \in \mathcal{X}_{\Omega}^{h} \text{ such that} \\ \mathbf{K}^{h}\big(\delta\mathbf{u}^{h}; \mathbf{v}^{h}\big) = \mathbf{F}^{h}\big(\mathbf{v}^{h}\big), \end{cases} \tag{6.26}$$

for all $\mathbf{v}^{h} = \big(\mathfrak{v}_{h}, \mathfrak{q}_{h}, \boldsymbol{\tau}_{h}, \tilde{\boldsymbol{\theta}}_{h}, \tilde{\boldsymbol{\Theta}}_{h}\big) \in \mathcal{X}_{\Omega}^{h}$ where $\mathbf{K}^{h}$ and $\mathbf{F}^{h}$ are bilinear and linear forms respectively.

6.7 Numerical discussions

In this section, we employ the Cross law model (3.44) with nonlinear viscosities dependent on velocity, pressure and temperature. We simulate steady state flow passed by a circular cylinder (3.4.2) with the nonlinear weighted LSFEM and use higher order finite element Q_2 for discretization of all the unknown variables. For the approximation of the discrete system, we employ a multigrid-preconditioned conjugate gradient solver and also compare the performance of the nonlinear solvers which are the Gauss-Newton method and the fixed-point linearization method respectively. We further investigate the nonlinear weighted functions dependent on nonlinear viscosity for different parametric values and show their importance to obtain optimal results.

Now our problem is composed of nine unknowns which are velocity $\mathfrak{u} = (u, v)$, stress $\boldsymbol{\sigma} = (\sigma_{xx}, \sigma_{xy}, \sigma_{yy})$ pressure $\mathfrak{p}$, temperature $\boldsymbol{\theta}$ and gradient of temperature $\boldsymbol{\Theta} = (\Theta_x, \Theta_y)$. We named this formulation as stress-velocity-pressure-temperature (SVPT) first order formulation. The SVPT formulation has 1.6 million (at finest level) more unknowns than SVP formulation, i.e. 50% more unknowns on each level repetitively, as referred to in Table 6.1 and Table 5.1. If we do a comparison with the reference results of Damanik [37], the SVPT formulation has 2.8 million unknowns more than the unknowns computed by VTP formulation (see Table V) of [37], i.e. 140% more unknowns.

Level	N_{DoF} $\mathfrak{u}$	N_{DoF} $\mathfrak{p}$	N_{DoF} $\boldsymbol{\sigma}$	N_{DoF} $\boldsymbol{\theta}$	N_{DoF} $\boldsymbol{\Theta}$	*Total* N_{DoF}
1	1144	572	1716	572	1144	5148
2	4368	2184	6552	2184	4368	19656
3	17056	8528	25584	8528	17056	76752
4	67392	33696	101088	33696	67392	303264
5	267904	133952	401856	133952	267904	1205568
6	1068288	534144	1602432	534144	1068288	4807296

Table 6.1: Total number of equations for the SVPT problem

6.7.1 Case-I: Temperature with viscous dissipation effects

For the SVPT formulation we consider two different configurations to study the temperature effects on the fluid flow. In the first case we include an additional viscous dissipation term in our computational formulation and calculate results for $\eta_0 = 10^{-3}$, $k_1 = k_2 = 10^{-2}$, $a_1 = 0$, $a_3 = 1$.

Viscosity depends on temperature

The results for temperature dependent nonlinear viscosity $\eta_\nu(\boldsymbol{\theta})$ for given parameters

$$r = 0,\ \alpha_p = 0,\ \eta_1 = 10^{-2},\ a_2 = 1,$$

and the nonlinear weighted function

$$w^{2.0}_{\eta_\nu} = \left(1/\eta_\nu(\boldsymbol{\theta})\right)^{2.0},$$

are shown in Table 6.2.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$7.440319E+01$	$1.319147E+00$	$1.082454E+00$	$9.511158E-02$
5	$7.447729E+01$	$1.321487E+00$	$1.088388E+00$	$1.430703E-02$
6	$7.448320E+01$	$1.321575E+00$	$1.089755E+00$	$3.731521E-03$
Fixed Point Method				
4	$7.442628E+01$	$1.319479E+00$	$1.082738E+00$	$9.314405E-02$
5	$7.448006E+01$	$1.321521E+00$	$1.088419E+00$	$1.354745E-02$	
6	$7.448367E+01$	$1.321577E+00$	$1.089759E+00$	$3.353438E-03$
Ref [37] : *Drag* $C_D = 7.447830E+01$, *Lift* $C_L = 1.32129E+00$				

Table 6.2: Temperature dependent viscosity with weighted function $w^{2.0}$

Viscosity depends on shear rate and temperature

The outcomes of temperature and shear dependent nonlinear viscosity $\eta_\nu\big(\dot{\gamma}_{_{II}}, \boldsymbol{\theta}\big)$ for the given values

$$r = 0.1,\ \alpha_p = 0,\ \eta_1 = 10^{-2},\ a_2 = 1,$$

and the weighted function

$$w^{2.0}_{\eta_\nu} = \big(1/\eta_\nu\big(\dot{\gamma}_{_{II}}, \boldsymbol{\theta}\big)\big)^{2.0},$$

are given in Table 6.3.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.387592E+01$	$1.057640E+00$	$7.964457E-01$	$1.020152E-01$
5	$5.392096E+01$	$1.059151E+00$	$7.997985E-01$	$1.906031E-02$
6	$5.392301E+01$	$1.059041E+00$	$8.005498E-01$	$6.101257E-03$
Fixed Point Method				
4	$5.388962E+01$	$1.057827E+00$	$7.966298E-01$	$1.030678E-01$
5	$5.392294E+01$	$1.059162E+00$	$7.998315E-01$	$1.821827E-02$
6	$5.392346E+01$	$1.059038E+00$	$8.005572E-01$	$5.512035E-03$
Ref [37] : *Drag* $C_{_D} = 5.391770E+01$, *Lift* $C_{_L} = 1.05863E+00$				

Table 6.3: Temperature and shear dependent viscosity with weighted function $w^{2.0}$

Viscosity depends on shear rate, pressure and temperature

The Table 6.4 is generated for temperature, shear rate and pressure dependent nonlinear viscosity $\eta_\nu\big(\dot{\gamma}_{_{II}}, \mathfrak{p}, \boldsymbol{\theta}\big)$ for provided parametric values

$$r = 0.1,\ \alpha_p = 10^{-3},\ \eta_1 = 10^{-2},\ a_2 = 1,$$

and the weighted function

$$w^{2.0}_{\eta_\nu} = \big(1/\eta_\nu\big(\dot{\gamma}_{_{II}}, \mathfrak{p}, \boldsymbol{\theta}\big)\big)^{2.0},$$

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$5.390952E+01$	$1.058079E+00$	$7.969481E-01$	$1.022693E-01$
5	$5.395465E+01$	$1.059592E+00$	$8.003066E-01$	$1.908287E-02$
6	$5.395670E+01$	$1.059483E+00$	$8.010592E-01$	$6.101391E-03$
Fixed Point Method				
4	$5.392328E+01$	$1.058267E+00$	$7.971330E-01$	$1.030582E-01$
5	$5.395663E+01$	$1.059604E+00$	$8.003397E-01$	$1.821093E-02$
6	$5.395715E+01$	$1.059480E+00$	$8.010666E-01$	$5.508308E-03$
Ref [37] : *Drag* $C_D = 5.395115E+01$, *Lift* $C_L = 1.05907E+00$				

Table 6.4: Temperature, shear and pressure dependent viscosity with weighted function $w^{2.0}$

In the Table 6.5, the nonlinear and linear iterations are compared for both the solvers. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively. The Gauss-Newton perform better in terms of nonlinear iterations. The SVPT formulation does not produce differentially diagonal dominant system matrix. So, the linear solver does not generate satisfactory outcomes for 4 pre/post smoothing steps.

Level	$r=0,\ \alpha_p=0$	$r=0.1,\ \alpha_p=0$	$r=0.1, \alpha_p=10^{-3}$
Gauss − Newton Method			
2	7/4	7/3	7/3
3	4/8	4/6	4/6
4	4/18	4/15	4/15
5	5/23	4/22	4/22
Fixed Point Method			
2	8/3	9/1	9/2
3	7/5	7/1	8/3
4	6/11	7/1	7/7
5	5/23	6/2	6/15

Table 6.5: Comparison of nonlinear/linear iterations

In the following Table 6.6 results are produced for different values of pre/post smoothing steps. It can be clearly seen that increase in smoothing steps can improve the solvers.

$r = 0.1,\ \alpha_p = 10^{-3}, \eta_1 = 10^{-2},\ a_2 = 1,\ w_{\eta_\nu}^{2.0}$				
Lev	$SmSt-4$	$SmSt-8$	$SmSt-16$	$SmSt-32$
Gauss − Newton Method				
2	7/3	7/1	7/1	7/1
3	4/6	4/3	4/2	4/1
4	4/15	4/8	4/4	4/2
5	4/22	3/15	3/9	3/5
Fixed Point Method				
2	9/2	9/1	9/1	9/1
3	8/3	8/2	7/1	7/1
4	7/7	7/4	7/2	7/1
5	6/15	6/8	6/5	6/2

Table 6.6: Nonlinear/linear iterations for smoothing steps

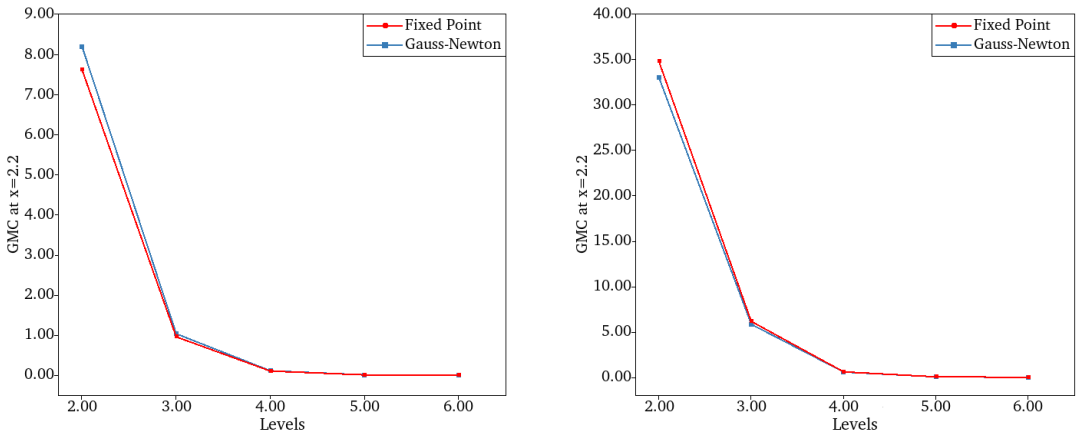

Figure 6.1: Comparison of nonlinear solvers for global mass conservation at outflow cross section of flow at $x = 2.2$ with corresponding levels for the nonlinear weight $w_{\eta_\nu}^{2.0}$ and bounded viscosities $\eta_1 = 10^{-2}$(left) and $\eta_1 = 10^{-1}$(right).

The Figure 6.2 shows the temperature generated by additional viscous dissipation term in the SVPT system (6.4). The boundary setting for the temperature in this case is set as Dirichlet data (prescribed as $\boldsymbol{\theta} = 0$) for inflow section. On the side walls and outflow section, temperature is set to natural boundary condition (prescribed as $\mathbf{n}.\nabla\boldsymbol{\theta} = 0$).

6.7.2 Case-II: Pre-heated cylinder with viscous dissipation

In the second case the temperature effects on fluid flow due to a pre-heated cylinder along with viscous dissipation are studied. The nonlinear viscosity in this case is

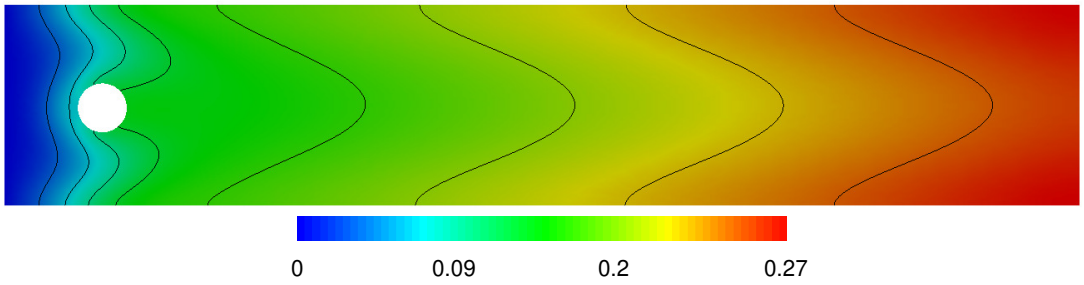

Figure 6.2: Temperature field generated by viscous dissipation

also temperature, shear and pressure dependent. The constant parametric values are $\eta_0 = 10^{-3}$, $r = 0.1$, $\alpha_p = 10^{-3}$, $a_1 = 0$, $a_2 = 1$, $a_3 = 1$.

The outcomes for

$$k_1 = k_2 = 10^{-2},\ \eta_1 = 10^{-2},$$

and

$$w_{\eta_\nu}^{2.0} = \left(1/\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p}, \boldsymbol{\theta})\right)^{2.0},$$

are given in the Table 6.7.

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m_{x=2.2}^{GMC}$
Gauss − Newton Method				
4	$4.532036E+01$	$9.040044E-01$	$6.981777E-01$	$1.145043E-01$
5	$4.537452E+01$	$9.062238E-01$	$7.007059E-01$	$1.297251E-02$
6	$4.538099E+01$	$9.065056E-01$	$7.012602E-01$	$1.523507E-03$
Fixed Point Method				
4	$4.532380E+01$	$9.037922E-01$	$6.982330E-01$	$1.068142E-01$
5	$4.537488E+01$	$9.062024E-01$	$7.007234E-01$	$1.216337E-02$
6	$4.538103E+01$	$9.065032E-01$	$7.012658E-01$	$1.434071E-03$
Ref [37] : *Drag* $C_D = 4.537988E+01$, *Lift* $C_L = 9.0649E-01$				

Table 6.7: Temperature, shear and pressure dependent viscosity with weighted function $w^{2.0}$

The results in Table 6.8 are generated for

$$k_1 = k_2 = 10^{-2},\ \eta_1 = 10^{-1},$$

and

$$w_{\eta_\nu}^{2.0} = \left(1/\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p}, \boldsymbol{\theta})\right)^{2.0}.$$

Lev.	*Drag* C_D	*Lift* C_L	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$4.619743E+02$	$4.847925E+00$	$6.895207E+00$	$5.810764E-01$
5	$4.649844E+02$	$4.921026E+00$	$7.071875E+00$	$4.694081E-02$
6	$4.652347E+02$	$4.928061E+00$	$7.112045E+00$	$3.761342E-03$
Fixed Point Method				
4	$4.623893E+02$	$4.846103E+00$	$6.903091E+00$	$5.993212E-01$
5	$4.650186E+02$	$4.921043E+00$	$7.072412E+00$	$4.824062E-02$
6	$4.652373E+02$	$4.928093E+00$	$7.112023E+00$	$3.866474E-03$
Ref [37] : *Drag* $C_D = 4.652119E+02$, *Lift* $C_L = 4.92828E+00$				

Table 6.8: Temperature, shear and pressure dependent viscosity with weighted function $w^{2.0}$

It is shown in the Tables 6.7 and 6.8 that the mass is conserved very well as we increase the refinement levels. The Table 6.9 represent the iterations comparison of Gauss-Newton and fixed point method for two cases of η_1. For $\eta_1 = 10^{-2}$ and $k_1 = k_2 = 10^{-2}$ the Gauss-Newton performs better than the fixed point method in terms of nonlinear iterations. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Level	$\eta_1 = 10^{-2}$	$\eta_1 = 10^{-1}$
Gauss − Newton Method		
2	6/3	13/13
3	4/5	7/23
4	3/11	7/25
5	3/19	12/25
Fixed point Method		
2	8/2	8/9
3	6/3	7/19
4	6/6	8/24
5	5/11	12/25

Table 6.9: Comparison of nonlinear/linear iterations

The outcomes for

$$k_1 = k_2 = 10^{-3},\ \eta_1 = 10^{-2},$$

and nonlinear weighted function

$$w^{2.0}_{\eta_\nu} = \big(1/\eta_\nu(\dot{\gamma}_{II}, \mathfrak{p}, \boldsymbol{\theta})\big)^{2.0},$$

are shown in the Table 6.10.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$4.963237E+01$	$9.611518E-01$	$7.722800E-01$	$1.211410E-01$
5	$4.970489E+01$	$9.636684E-01$	$7.752118E-01$	$1.389537E-02$
6	$4.971342E+01$	$9.639941E-01$	$7.758310E-01$	$1.643185E-03$
Fixed Point Method				
4	$4.964555E+01$	$9.608884E-01$	$7.725296E-01$	$1.193354E-01$
5	$4.970625E+01$	$9.636444E-01$	$7.752499E-01$	$1.366896E-02$
6	$4.971356E+01$	$9.639918E-01$	$7.758384E-01$	$1.617713E-03$
Ref [37] : *Drag* $C_{_D} = 4.971239E+01$, *Lift* $C_{_L} = 9.6399E-01$				

Table 6.10: Temperature, shear and pressure dependent viscosity with weighted function $w^{2.0}$

Now, we consider a bit harder case as

$$\eta_1 = 10^{-1},\ k_1 = k_2 = 10^{-3}.$$

The optimal convergence is obtained for the value of nonlinear weighted function

$$w^{4.0}_{\eta_\nu} = \big(1/\eta_\nu(\dot{\gamma}_{_{II}}, \mathfrak{p}, \boldsymbol{\theta})\big)^{4.0},$$

as shown in the Table 6.11.

Lev.	*Drag* $C_{_D}$	*Lift* $C_{_L}$	$\Delta\mathfrak{p}$	$m^{GMC}_{x=2.2}$
Gauss − Newton Method				
4	$4.688078E+02$	$4.430503E+00$	$7.099451E+00$	$6.881007E+00$
5	$5.072187E+02$	$5.253643E+00$	$7.800688E+00$	$9.329903E-01$
6	$5.132120E+02$	$5.381100E+00$	$7.925989E+00$	$1.154749E-01$
Fixed Point Method				
4	$4.780946E+02$	$4.404470E+00$	$7.289815E+00$	$7.657795E+00$
5	$5.094555E+02$	$5.268389E+00$	$7.846948E+00$	$9.884661E-01$
6	$5.135090E+02$	$5.383273E+00$	$7.932421E+00$	$1.205632E-01$
Ref [37] : *Drag* $C_{_D} = 5.140215E+02$, *Lift* $C_{_L} = 5.39856E+00$				

Table 6.11: Temperature, shear and pressure dependent viscosity with weighted function $w^{4.0}$

It is observed from Table 6.11 that our solvers are struggling as mass conservation is not very well preserved. The Gauss-Newton method performs better as fixed point

in comparison to nonlinear iterations for $\eta_1 = 10^{-2}$ and $k_1 = k_2 = 10^{-3}$. For more challenging case at $\eta_1 = 10^{-1}$ both solvers worked in a similar fashion. The nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively.

Level	$\eta_1 = 10^{-2},\ w_{\eta_\nu}^{2.0}$	$\eta_1 = 10^{-1},\ w_{\eta_\nu}^{4.0}$
Gauss − Newton Method		
2	7/3	15/11
3	4/5	19/16
4	3/11	10/22
5	3/19	9/23
Fixed Point Method		
2	9/2	10/7
3	7/3	8/10
4	6/7	7/20
5	5/12	10/23

Table 6.12: Comparison of solvers based on nonlinear/linear iterations

In the following Table 6.13, the results are produced for different values of pre/post smoothing steps. It can be clearly seen that increase in smoothing steps can improve the solvers.

$r = 0.1,\ \alpha_p = 10^{-3}, \eta_1 = 10^{-1},\ a_2 = 1,\ w_{\eta_\nu}^{4.0}$				
Lev	$SmSt - 4$	$SmSt - 8$	$SmSt - 16$	$SmSt - 32$
Gauss − Newton Method				
2	15/11	15/6	14/4	15/3
3	19/16	19/10	19/7	19/5
4	10/22	10/18	10/11	10/6
5	9/23	5/21	5/19	5/12
Fixed Point Method				
2	10/7	10/4	10/2	11/2
3	8/10	8/6	9/3	9/3
4	7/20	7/10	8/6	8/3
5	10/23	7/20	7/12	7/7

Table 6.13: Nonlinear/linear iterations for varios smoothing steps

In this particular case heat is not only generated by the viscous dissipation but is also coming from the heated cylinder. The inflow boundary setting for the temperature is set to be zero, but prescribing no-slip condition at walls and inhomogeneous Dirichlet data $\boldsymbol{\theta} = 1$ at the cylinder.

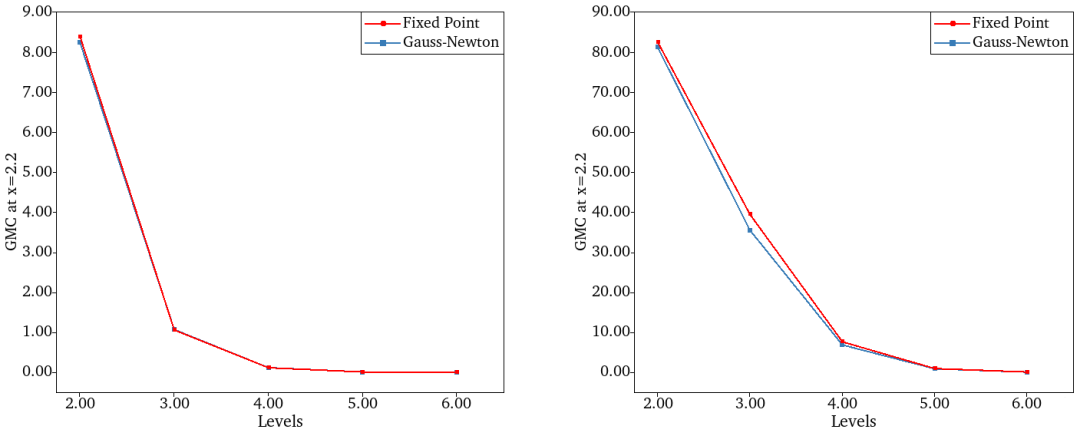

Figure 6.3: Comparison of nonlinear solvers for global mass conservation calculated at outflow cross section of flow at $x = 2.2$ with corresponding levels with bounded viscosities and nonlinear weights $\eta_1 = 10^{-2}$, $w_{\eta_\nu}^{2.0}$ (left) and $\eta_1 = 10^{-1}$, $w_{\eta_\nu}^{4.0}$ (right).

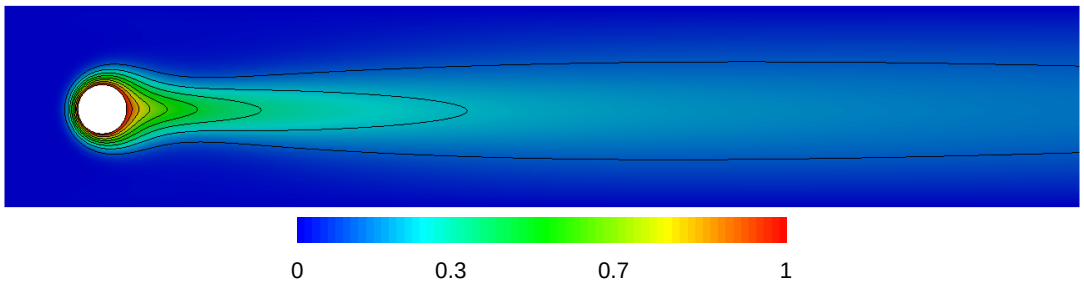

Figure 6.4: Temperature generated by the heated cylinder and viscous dissipation

6.8 Summary

In this chapter, we have presented an analysis on the non-Newtonian fluid flows with different levels of complexities. These complexities are determined by the choices of nonlinear viscosity functions. Again the problem under consideration is fluid flow passed circular cylinder (3.4.2) satisfying Cross law model (3.44). In addition, the fluid flow holds some thermal properties as given in problem model (6.1). Then, the second order problem is transformed into first order problem (6.4) by suitable substitutions and named as SVPT formulation. We are the first ones to introduce such a formulation for least-squares finite element method. Furthermore, the nonlinear weighted functions are employed to get more stable solver. The monolithic Newton-multigrid like solver is proposed to efficiently solve the huge problems. The MPCG solver is employed as linear solver because of SPD system. The SSOR preconditioned conjugate gradient (SSOR-PCG) method is used as smoother for multigrid method to get parametric free smoothing. But it requires to improve the performance of linear solver as SVP formulation leads to coefficient matrix which is not differentially diagonal dominant. The Newton linearization is performed to the nonlinear system to apply the least-squares principle. To building a robust and accurate linear solver is a challenging task in such scenarios. The Gauss-Newton and fixed point methods are the considered nonlinear solvers.

We have investigated two types of flow cases for numerical tests. In first case, the flow parameters are computed and the temperature effects are obtained from viscous dissipation or friction in the fluid flow. In second case, we have performed the numerical tests for pre-heated cylinder with additional temperature produced by viscous dissipation in the flow. The nonlinear viscosity functions are based on shear rate, pressure and temperature which are used with various combinations. The nonlinear weighted functions $w^n_{\eta_\nu}$; $n = 2.0, 4.0$, with different nonlinear viscosity functions are employed to develop the stable solver for complex problems. Without the weighted function, we are unable to get accurate and robust results. The solvers have shown grid independent convergence behaviors except few cases. The Gauss-Newton method performs better than the fixed point method.

The global mass conservation is computed on outflow cross-section of the domain. It is observed from Table 6.11 that the increase in the difficulty level of a problem leads to lack in global mass conservation. It also effects the performance of the solvers, thus the nonlinear weighted function $w^{4.0}_{\eta_\nu}$ is used in such situation. Overall, the global mass conservation is achieved for all other cases.

Chapter 7

Summary and Outlook

In this chapter, we sum up the tasks briefly and present the final thoughts on the conducted research work. The possible scope of LSFEM is projected for future work as well.

Conclusion

In this thesis, an alternative monolithic method based on least-squares finite element formulations is proposed for Newtonian and non-Newtonian fluids. The least-squares finite element method is based on the minimization of energy functional. Since, it is considered as one of the true variational methods among weighted residual schemes. Therefore, it has the capability to generate the 'best' approximation to the exact solution of the modeled systems. The Least-Squares method provides a more general, adaptable and robust process for system formulation than weak form-based Galerkin finite element methods. Moreover, it has several theoretical and computational advantages. The least-squares framework has the capability to avoid the 'inf-sup' stability condition of LBB and make it possible to employ equal order interpolation for all the unknowns. The least-squares principle on a discrete model leads to a symmetric and positive definite coefficient matrix. Hence, it opens the door for the implementation of robust and fast iterative schemes to approximate the system of equations.

The above features of the least-squares method captured our attention for developing a weighted least-squares finite element method. The main objective of our work is to establish an effective and powerful monolithic technique equipped with multigrid features for solving the discrete system of equations. Another aim of this study is to emphasis the importance of weighted functions for the stability of least-squares energy functional. Therefore, we have employed both linear and nonlinear weighted functions for numerical investigations of Newtonian and non-Newtonian fluids. We have applied our methodology on wide range of problem complexity from easy to hard or complex fluid flow problems. The constant viscosity is used for Newtonian fluids and nonlinear

viscosity models, namely power law viscosity model and Cross law viscosity model for unbounded and bounded viscosities cases, are used for non-Newtonian fluid flow problems. We consider only an inflow-outflow flow around cylinder problem for numerical computation. The reason to stick to the inflow-outflow problem is to emphasis on the lack of global mass conservation issues. The calculations associated with least-squares based method encounters lack of global mass conservation in inflow-outflow problems. Therefore, we focused on developing a weighted least-squares finite element method that can provide robust and accurate approximations without affecting the global mass conservation of the system.

We have presented the introduction, motivation and brief overview of the work in the first chapter. The theoretical aspects of the least-squares method are covered in the second chapter. The least-squares principles for continuous and discrete problems are given with practicality issues are explained in detail. The second-order elliptic PDEs problem in least-squares sense is quite difficult to handle, therefore we use different transformation procedures to convert the problem to the first-order formulations. Some first-order formulations are discussed and we also introduced the SVP, SVPT formulations which are used in this thesis for research purposes. The nonlinear problem in least-squares sense is elaborated as well. The linearization techniques for the nonlinear problem are studied in detail. We have used the linearization before the application of the least-squares principle to formulate our systems. In the following chapter, we presented the method used for numerical simulations which included both linear and nonlinear solvers. The multigrid preconditioned conjugate gradient method is employed to compute the discrete linear problems. The coefficient matrix generated by the application of least-squares principle based on SVP formulation is not differentially diagonal dominant. This lack of differentially diagonal dominance property directly effects the performance of linear solver. Therefore, the linear solver is an open problem and still requires to improve. The Gauss-Newton and fixed point methods have been used as nonlinear solvers. The linear and quadric finite elements are employed. The flow around cylinder domain is examined and two mesh configurations (3.4.2) and (3.4.3) are used for the numerical investigations.

The second part of the thesis is concerned with the implementation, numerical testing and discussions of different fluid flow models. The numerical discussions are followed on by a similar pattern in chapter four, five and six. We have calculated the flow parameters such as coefficients drag/lift, pressure difference of across the cross-sections of the circular domain and global mass conservation on cross-section of FAC domain at $x = 2.2$ as well. Then, the solvers are compared based on their performance. The nonlinear iterations performed by nonlinear solvers and averaged linear iterations taken by linear solver are compared as well. The weighted functions depending on the viscosity of the fluids are employed in the energy functionals for

a more stable least-squares method. Since, we have used SVP transformation for the model's formulation which consists of stress, velocity and pressure as unknown variables which is rarely used for the Stokes problem. So, we have discussed the Stokes problem for numerical validations of the solver. The Navier-Stokes problem has been modeled by SVP formulation and the variational problem has been defined by the application of the straight forward least-squares principle in L_2-norm. Further, the least-squares energy functional has been balanced by linear weighted function. The flow parameters are computed for two mesh configurations using lower and higher-order finite element discretizations. It has been observed the solutions for higher-order finite element discretization are more accurate and global mass is well preserved. Moreover, the L_w-LSFEM has shown robust and grid-independent convergence.

The major contribution of this thesis is the results obtained for non-Newtonian fluids with additional viscous dissipation and non-isothermal effects. In chapter five, we have developed the least-squares based method i.e. N_w-LSFEM for the solutions of power law and Cross law fluid problems. The power law fluids are relied on the shear rate of fluid viscosity which is unbounded. The shear thinning and shear thickening effects of fluids have been computed for different values of power law index. To study the bounded viscosities, we employed Cross law viscosity model and produced results for different ranges of viscosity. The analysis has been made on the viscosities depending on shear rate, pressure and correspondingly weighted functions. We have concluded that the weighted functions are important for both the optimal convergence and accuracy of the N_w-LSFEM. The solver has shown grid-independent behavior for a huge range of conducted tests. Finally, a nonlinear weighted least-squares finite element method has been developed for non-Newtonian fluids with non-isothermal effects. We have introduced a new first-order formulation named SVPT formulation for the coupled system of equations. The temperature effects are computed for two test cases of fluid flow. In the first case, the heat is generated by additional viscous dissipation of flow while in the second case, the heat is obtained from a pre-heated cylinder along with friction term. The viscosity taken for the study is stress, velocity, pressure and temperature dependent. The nonlinear weighted functions with corresponding viscosity functions have been used for the stable solver. We have observed that the choice of weighted functions varied with the complexity of the problem. The proposed monolithic solver has shown great accuracy, robustness, and efficiency for the complex fluids flow problem. We have obtained accurate results for flow around cylinder problem and solver has generated the grid-independent results.

Future work

A numerical study of Newtonian and non-Newtonian fluids have been presented for flow around cylinder problem. The performance of the linear and nonlinear solvers have

been tested for variety of complex problems. The recommendations for the future work is outlined in this section.

3D Extension: The numerical solutions of 2D NS problems are discussed in this research work. The 3D extension for the current least square finite element framework is the next milestone to be achieved. This will make the LSFEM more practical to handle real life situations.

Viscoelastic Problem: Another aim is to develop the LSFEM for the equations governing generalized Newtonian and viscoelastic flows such as those occurring in polymer processes. The Giesekus viscoelastic constitutive equation will be considered as

$$(\mathfrak{u} \cdot \nabla)\,\boldsymbol{\tau} - \nabla\mathfrak{u}\boldsymbol{\tau} - \boldsymbol{\tau}\nabla\mathfrak{u}^{\mathrm{T}} + \frac{1}{\Lambda}\left(\boldsymbol{\tau} - \mathbf{I} + \boldsymbol{\alpha}\,(\boldsymbol{\tau} - \mathbf{I})^2\right) = 0,$$

where the parameter $\boldsymbol{\alpha}$ is Giesekus mobility factor, $\boldsymbol{\tau}$ is additional stress known as conformation tensor, and Λ is the relaxation time of the polymer. The parameter $\boldsymbol{\alpha} = \mathbf{0}$ yields to the Oldroyd-B model, i.e.

$$(\mathfrak{u} \cdot \nabla)\,\boldsymbol{\tau} - \nabla\mathfrak{u}\boldsymbol{\tau} - \boldsymbol{\tau}\nabla\mathfrak{u}^{\mathrm{T}} + \frac{1}{\Lambda}\,(\boldsymbol{\tau} - \mathbf{I}) = 0.$$

The total stress is consist of viscus and elastic part as well. This will increase the numerical complexity of problem and will be very challenging to compute.

Time Dependent Study: The straightforward application of least-squares principle on the time dependent flow problems for Newtonian fluids can easily be adapted, see e.g. [55], [63]. In this work, the non-Newtonian fluids, without time dependency constant, have been studied. The time dependent non-Newtonian fluids lead to highly complex problems as the relation between shear stress and shear rate depend on the duration of shearing and their kinematic history. The solution to the time dependent non-Newtonian fluid problems will be the challenging task of proposed future work.

First Order Formulations and Geometries: The stress based formulation has been used throughout this work for inflow/outflow problem. The coefficient matrix generated by the application of least-squares principle on system based on SVP formulation is not differentially diagonal dominant. This effects the linear solver and gives slow convergence toward the solution. Many other first order formulation haven used by researchers in literature. Therefore, we will investigate the LSFEM for different formulations and will consider different geometries for our future research work.

Appendix A

Block Matrices

Calculation of blocks in coefficient matrix

The NS problem 4.4 for SVP formulation is studied in chapter 4. The application of least-squares principle leads to a SPD matrix which is called a coefficient matrix $\mathbf{K}$. The coefficient matrix for the NS problem is given by the equation (4.21) and presented in block form as under

$$\mathbf{K}\big(\delta\mathbf{u},\mathbf{v}\big) = \begin{bmatrix} \boxed{\mathbf{K_{vu}}} & \mathbf{K_{vp}} & \mathbf{K_{v\sigma}} \\ \mathbf{K_{qu}} & \mathbf{K_{qp}} & \mathbf{K_{q\sigma}} \\ \mathbf{K_{\tau u}} & \mathbf{K_{\tau p}} & \mathbf{K_{\tau\sigma}} \end{bmatrix}, \tag{A.1}$$

where

$$\begin{aligned} \mathbf{K_{vu}} \;=\; & \boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle} + \boxed{\big\langle \mathfrak{u}\cdot\nabla\mathfrak{u}^k, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle} + \boxed{\big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{v}\cdot\nabla\mathfrak{u}^k\big\rangle} \\ & + \big\langle \mathfrak{u}^k\cdot\nabla\mathfrak{u}, \mathfrak{u}^k\cdot\nabla\mathfrak{v}\big\rangle + \big\langle \nabla\cdot\mathfrak{u}, \nabla\cdot\mathfrak{v}\big\rangle + \big\langle -\mathbf{v_\sigma}\big(\mathfrak{u}\big), -\mathbf{v_\tau}\big(\mathfrak{v}\big)\big\rangle. \end{aligned} \tag{A.2}$$

We consider only the first diagonal block matrix $\mathbf{K_{vu}}$ of the coefficient matrix $\mathbf{K}$ and it consists of velocity terms only. In the following sections, we are going to discuss the construction of matrix in detail on minor levels.

The trail and test functions of velocity

The trail and test functions for the velocity field are defined as

$$\begin{aligned} \delta\mathbf{u} \;&=\; \sum_{j=1}^{N_{dof}} \left[\begin{bmatrix} \varphi_j \\ 0 \end{bmatrix} u_j + \begin{bmatrix} 0 \\ \varphi_j \end{bmatrix} v_j \right], \\ \mathbf{v} \;&=\; \begin{bmatrix} \varphi_i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \varphi_i \end{bmatrix}, \end{aligned} \tag{A.3}$$

The construction of sub-block matrices

The matrix $\mathbf{K_{vu}}$ can be shown as the sum of six block matrices because of the six terms in equation (A.2) . Each term contributes further to sub-matrices.

$$\mathbf{K_{vu}} = \mathbf{K^1_{vu}} + \mathbf{K^2_{vu}} + \mathbf{K^3_{vu}} + \mathbf{K^4_{vu}} + \mathbf{K^5_{vu}} + \mathbf{K^6_{vu}} = \sum_{i=1}^{6} \mathbf{K}^i_{\mathbf{vu}} \tag{A.4}$$

The integrals are calculated term by term.

$$\int_\Omega \big(\big(\mathfrak{u} \cdot \nabla \mathfrak{u}^k \big) \cdot \big(\mathfrak{v} \cdot \nabla \mathfrak{u}^k \big) \big) d\Omega = \begin{bmatrix} \mathbf{K}^1_{uu} & \mathbf{K}^1_{uv} \\ \mathbf{K}^1_{vu} & \mathbf{K}^1_{vv} \end{bmatrix} \tag{A.5}$$

where

$$\begin{aligned}
\mathbf{K}^1_{uu} &= \left[\big(\frac{\partial u^k}{\partial x}\big)^2 + \big(\frac{\partial v^k}{\partial x}\big)^2 \right] \varphi_i \varphi_j, \\
\mathbf{K}^1_{uv} &= \left[\big(\frac{\partial u^k}{\partial x}\frac{\partial u^k}{\partial y}\big) + \big(\frac{\partial v^k}{\partial x}\frac{\partial v^k}{\partial y}\big) \right] \varphi_i \varphi_j, \\
\mathbf{K}^1_{vu} &= \left[\big(\frac{\partial u^k}{\partial x}\frac{\partial u^k}{\partial y}\big) + \big(\frac{\partial v^k}{\partial x}\frac{\partial v^k}{\partial y}\big) \right] \varphi_i \varphi_j, \\
\mathbf{K}^1_{vv} &= \left[\big(\frac{\partial u^k}{\partial y}\big)^2 + \big(\frac{\partial v^k}{\partial y}\big)^2 \right] \varphi_i \varphi_j,
\end{aligned}$$

$$\int_\Omega \big(\big(\mathfrak{u} \cdot \nabla \mathfrak{u}^k \big) \cdot \big(\mathfrak{u}^k \cdot \nabla \mathfrak{v} \big) \big) d\Omega = \begin{bmatrix} \mathbf{K}^2_{uu} & \mathbf{K}^2_{uv} \\ \mathbf{K}^2_{vu} & \mathbf{K}^2_{vv} \end{bmatrix} \tag{A.6}$$

where

$$\begin{aligned}
\mathbf{K}^2_{uu} &= \left[\big(u^k \frac{\partial u^k}{\partial x}\big)\big(\frac{\partial \varphi_i}{\partial x}\varphi_j\big) + \big(v^k \frac{\partial u^k}{\partial x}\big)\big(\frac{\partial \varphi_i}{\partial y}\varphi_j\big) \right], \\
\mathbf{K}^2_{uv} &= \left[\big(u^k \frac{\partial u^k}{\partial y}\big)\big(\frac{\partial \varphi_i}{\partial x}\varphi_j\big) + \big(v^k \frac{\partial u^k}{\partial y}\big)\big(\frac{\partial \varphi_i}{\partial y}\varphi_j\big) \right], \\
\mathbf{K}^2_{vu} &= \left[\big(u^k \frac{\partial v^k}{\partial x}\big)\big(\frac{\partial \varphi_i}{\partial x}\varphi_j\big) + \big(v^k \frac{\partial v^k}{\partial x}\big)\big(\frac{\partial \varphi_i}{\partial y}\varphi_j\big) \right], \\
\mathbf{K}^2_{vv} &= \left[\big(u^k \frac{\partial v^k}{\partial y}\big)\big(\frac{\partial \varphi_i}{\partial x}\varphi_j\big) + \big(v^k \frac{\partial v^k}{\partial y}\big)\big(\frac{\partial \varphi_i}{\partial y}\varphi_j\big) \right],
\end{aligned}$$

$$\int_\Omega \big(\big(\mathfrak{u}^k \cdot \nabla \mathfrak{u} \big) \cdot \big(\mathfrak{v} \cdot \nabla \mathfrak{u}^k \big) \big) d\Omega = \begin{bmatrix} \mathbf{K}^3_{uu} & \mathbf{K}^3_{uv} \\ \mathbf{K}^3_{vu} & \mathbf{K}^3_{vv} \end{bmatrix} \tag{A.7}$$

where

$$\begin{aligned}
\mathbf{K}^3_{uu} &= \left[\big(u^k \frac{\partial u^k}{\partial x}\big)\big(\varphi_i \frac{\partial \varphi_j}{\partial x}\big) + \big(v^k \frac{\partial u^k}{\partial x}\big)\big(\varphi_i \frac{\partial \varphi_j}{\partial y}\big) \right], \\
\mathbf{K}^3_{uv} &= \left[\big(u^k \frac{\partial v^k}{\partial x}\big)\big(\varphi_i \frac{\partial \varphi_j}{\partial x}\big) + \big(v^k \frac{\partial v^k}{\partial x}\big)\big(\varphi_i \frac{\partial \varphi_j}{\partial y}\big) \right],
\end{aligned}$$

$$
\begin{aligned}
\mathbf{K}^3_{vu} &= \Big[\big(u^k\frac{\partial u^k}{\partial y}\big)\big(\varphi_i\frac{\partial \varphi_j}{\partial x}\big) + \big(v^k\frac{\partial u^k}{\partial y}\big)\big(\varphi_i\frac{\partial \varphi_j}{\partial y}\big)\Big], \\
\mathbf{K}^3_{vv} &= \Big[\big(u^k\frac{\partial v^k}{\partial y}\big)\big(\varphi_i\frac{\partial \varphi_j}{\partial x}\big) + \big(v^k\frac{\partial v^k}{\partial y}\big)\big(\varphi_i\frac{\partial \varphi_j}{\partial y}\big)\Big],
\end{aligned}
$$

$$
\int_\Omega \big(\big(\mathfrak{u}^k\cdot\nabla\mathfrak{u}\big)\cdot\big(\mathfrak{u}^k\cdot\nabla\mathfrak{v}\big)\big)d\Omega = \begin{bmatrix} \mathbf{K}^4_{uu} & \mathbf{K}^4_{uv} \\ \mathbf{K}^4_{vu} & \mathbf{K}^4_{vv} \end{bmatrix} \tag{A.8}
$$

where

$$
\begin{aligned}
\mathbf{K}^4_{uu} &= \Big[\big(u^k\big)^2\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) + u^kv^k\Big\{\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big) + \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big)\Big\} \\
&\quad + \big(v^k\big)^2\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big)\Big], \\
\mathbf{K}^4_{uv} &= \mathbf{0}, \\
\mathbf{K}^4_{vu} &= \mathbf{0}, \\
\mathbf{K}^4_{vv} &= \Big[\big(u^k\big)^2\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) + u^kv^k\Big\{\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big) + \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big)\Big\} \\
&\quad + \big(v^k\big)^2\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big)\Big],
\end{aligned}
$$

$$
\int_\Omega \big(\nabla\cdot\mathfrak{u}\big)\cdot\big(\nabla\cdot\mathfrak{v}\big)d\Omega = \begin{bmatrix} \mathbf{K}^5_{uu} & \mathbf{K}^5_{uv} \\ \mathbf{K}^5_{vu} & \mathbf{K}^5_{vv} \end{bmatrix} \tag{A.9}
$$

where

$$
\begin{aligned}
\mathbf{K}^5_{uu} &= \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big), \\
\mathbf{K}^5_{uv} &= \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big), \\
\mathbf{K}^5_{vu} &= \big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big), \\
\mathbf{K}^5_{vv} &= \big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big),
\end{aligned}
$$

$$
\int_\Omega \big(-2\eta_\nu\mathbf{D}\big(\mathfrak{u}\big)\big)\cdot\big(-2\eta_\nu\mathbf{D}\big(\mathfrak{v}\big)\big)d\Omega = \begin{bmatrix} \mathbf{K}^6_{uu} & \mathbf{K}^6_{uv} \\ \mathbf{K}^6_{vu} & \mathbf{K}^6_{vv} \end{bmatrix} \tag{A.10}
$$

where

$$
\begin{aligned}
\mathbf{K}^6_{uu} &= \big(-2\eta_\nu\big)^2\Big\{\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) + \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big)\Big\}, \\
\mathbf{K}^6_{uv} &= \big(-2\eta_\nu\big)^2\Big\{\frac{1}{2}\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big)\Big\},
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}^6_{vu} &= \big(-2\eta_\nu\big)^2 \left\{ \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big) \right\}, \\
\mathbf{K}^6_{vv} &= \big(-2\eta_\nu\big)^2 \left\{ \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) + \big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big) \right\},
\end{aligned}
$$

The matrix $\mathbf{K_{vu}}$ in equation (A.4) can be written as

$$
\mathbf{K_{vu}} = \begin{bmatrix} \mathbf{K}_{uu} & \mathbf{K}_{uv} \\ \mathbf{K}_{vu} & \mathbf{K}_{vv} \end{bmatrix} = \sum_{i=1}^{6} \begin{bmatrix} \mathbf{K}^i_{uu} & \mathbf{K}^i_{uv} \\ \mathbf{K}^i_{vu} & \mathbf{K}^i_{vv} \end{bmatrix} \tag{A.11}
$$

where

$$
\begin{aligned}
\mathbf{K}_{uu} &= \mathbf{K}^1_{uu} + \mathbf{K}^2_{uu} + \mathbf{K}^3_{uu} + \mathbf{K}^4_{uu} + \mathbf{K}^5_{uu} + \mathbf{K}^6_{uu} \\
&= \left[\big(\frac{\partial u^k}{\partial x}\big)^2 + \big(\frac{\partial v^k}{\partial x}\big)^2 \right] \varphi_i\varphi_j + \left[\big(u^k\frac{\partial u^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial x}\varphi_j\big) + \big(v^k\frac{\partial u^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial y}\varphi_j\big) \right] \\
&\quad + \left[\big(u^k\frac{\partial u^k}{\partial x}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial x}\big) + \big(v^k\frac{\partial u^k}{\partial x}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial y}\big) \right] + \left[\big(u^k\big)^2\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) \right. \\
&\quad \left. + u^kv^k \left\{ \big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big) + \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big) \right\} + \big(v^k\big)^2\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big) \right] + \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) \\
&\quad + \big(-2\eta_\nu\big)^2 \left\{ \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x}\big) + \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y}\big) \right\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{uv} &= \mathbf{K}^1_{uv} + \mathbf{K}^2_{uv} + \mathbf{K}^3_{uv} + \mathbf{K}^4_{uv} + \mathbf{K}^5_{uv} + \mathbf{K}^6_{uv} \\
&= \left[\big(\frac{\partial u^k}{\partial x}\frac{\partial u^k}{\partial y}\big) + \big(\frac{\partial v^k}{\partial x}\frac{\partial v^k}{\partial y}\big) \right] \varphi_i\varphi_j + \left[\big(u^k\frac{\partial u^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial x}\varphi_j\big) + \big(v^k\frac{\partial u^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial y}\varphi_j\big) \right] \\
&\quad + \left[\big(u^k\frac{\partial v^k}{\partial x}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial x}\big) + \big(v^k\frac{\partial v^k}{\partial x}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial y}\big) \right] + \mathbf{0} + \big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big) \\
&\quad + \big(-2\eta_\nu\big)^2 \left\{ \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big) \right\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{vu} &= \mathbf{K}^1_{vu} + \mathbf{K}^2_{vu} + \mathbf{K}^3_{vu} + \mathbf{K}^4_{vu} + \mathbf{K}^5_{vu} + \mathbf{K}^6_{vu} \\
&= \left[\big(\frac{\partial u^k}{\partial x}\frac{\partial u^k}{\partial y}\big) + \big(\frac{\partial v^k}{\partial x}\frac{\partial v^k}{\partial y}\big) \right] \varphi_i\varphi_j + \left[\big(u^k\frac{\partial v^k}{\partial x}\big)\big(\frac{\partial\varphi_i}{\partial x}\varphi_j\big) + \big(v^k\frac{\partial v^k}{\partial x}\big)\big(\frac{\partial\varphi_i}{\partial y}\varphi_j\big) \right] \\
&\quad + \left[\big(u^k\frac{\partial u^k}{\partial y}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial x}\big) + \big(v^k\frac{\partial u^k}{\partial y}\big)\big(\varphi_i\frac{\partial\varphi_j}{\partial y}\big) \right] + \mathbf{0} + \big(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x}\big) \\
&\quad + \big(-2\eta_\nu\big)^2 \left\{ \frac{1}{2}\big(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y}\big) \right\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{K}_{vv} &= \mathbf{K}^1_{vv} + \mathbf{K}^2_{vv} + \mathbf{K}^3_{vv} + \mathbf{K}^4_{vv} + \mathbf{K}^5_{vv} + \mathbf{K}^6_{vv} \\
&= \left[\big(\frac{\partial u^k}{\partial y}\big)^2 + \big(\frac{\partial v^k}{\partial y}\big)^2 \right] \varphi_i\varphi_j + \left[\big(u^k\frac{\partial v^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial x}\varphi_j\big) + \big(v^k\frac{\partial v^k}{\partial y}\big)\big(\frac{\partial\varphi_i}{\partial y}\varphi_j\big) \right]
\end{aligned}
$$

$$
\begin{aligned}
&+\left[(u^k\frac{\partial v^k}{\partial y})(\varphi_i\frac{\partial\varphi_j}{\partial x})+(v^k\frac{\partial v^k}{\partial y})(\varphi_i\frac{\partial\varphi_j}{\partial y})\right]+\left[(u^k)^2(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x})\right.\\
&\left.+u^kv^k\left\{(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial x})+(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial y})\right\}+(v^k)^2(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y})\right]+(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y})\\
&+(-2\eta_\nu)^2\left\{\frac{1}{2}(\frac{\partial\varphi_i}{\partial x}\frac{\partial\varphi_j}{\partial x})+(\frac{\partial\varphi_i}{\partial y}\frac{\partial\varphi_j}{\partial y})\right\}
\end{aligned}
$$

Bibliography

[1] ADAMS, R.: *Sobolev Spaces.* Academic Press, New York, 1975.

[2] AGMON, S., DOUGLIS, A., NIRENBERG, L.: *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I.* Comm. Pure Appl. Math. 12, 623–727, 1959.

[3] AGMON, S., DOUGLIS, A., NIRENBERG, L.: *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II.* Comm. Pure Appl. Math. 17, 35–92, 1964.

[4] ARNOLD, D. N., BOFFI, D. and FALK, R. S.: *Approximation by quadrilateral finite element.* Math. Comput., 71:909–922, 2002.

[5] AZIZ, A.: *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations.* Academic Press, New York, 1972.

[6] AZIZ, A., KELLOG, R., STEPHENS, A.: *least-squares methods for elliptic systems.* Math. Comp. 44, 53–70, 1985.

[7] BABUSKA, I.: *Error bounds for finite element method*, Numer. Math. 16, pp. 322-333, 1971.

[8] BABUSKA, I.: *The finite element method with Lagrange multipliers*, Numer. Math., 20, 179–192, 1973.

[9] BATTETT et al, R.: *Templates for the solution of linear systems: Building blocks for iterative methods.* Philadelphia: SIAM, 1994.

[10] BELL, B. C. and SURANA, K. S.: *p-version least-squares finite element formulation for two-dimensional, incompressible, non-Newtonian isothermal and non-isothermal fluid flow.* International Journal for Numerical Methods in Fluids, 18(2):127–162, 1994.

[11] BENZI, M.: *Preconditioning Techniques for Large Linear Systems: A Survey.* In: Journal of Computational Physics 182.2, pp. 418–477, 2002.

[12] BOCHEV, P. B.: *Least Squares Finite Element Methods for the Stokes and Navier-Stokes Equations.* PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, July 1994.

[13] BOCHEV, P. B.: *Analysis of least-squares finite element methods for the Navier-Stokes equations.* SIAM Journal on Numerical Analysis, 34(5):1817–1844, 1997.

[14] BOCHEV, P. B.: *Experiences with negative norm least-square methods for the Navier-Stokes equations.* Elect. Trans. Numer. Anal., 6:44–62, 1997.

[15] BOCHEV, P., GUNZBURGER, M.: *Least-squares finite element methods.* Applied Mathematical Sciences 166, Springer, 2009.

[16] BOCHEV, P. B., CAI, Z., MANTEUFFEL, T. A. and McCORMICK, S. F.: *Analysis of Velocity-Flux First-Order System Least-Squares Principles for the Navier-Stokes Equations*: Part I. SIAM J. Numer. Anal., 35(3):990–1009, 1998.

[17] BOCHEV, P., GUNZBURGER, M.: *Least–squares methods for the velocity–pressure–stress formulation of the Stokes equation.* Comput. Methods in Appl. Mech. Eng., 126:267–287, 1995.

[18] BOCHEV, P., GUNZBURGER, M.: *Finite Element Methods of Least Squares Type.* SIAM Rev., 40(4):789–837, 1998.

[19] BOCHEV, P., GUNZBURGER, M.: *Least-squares finite element methods.* In Proceedings of the International Congress of Mathematicians (ICM), Madrid, 2006.

[20] BOCHEV, P. B., MANTEUFFEL, T. A. and McCORMICK, S. F.: *Analysis of Velocity-Flux Least-Squares Principles for the Navier-Stokes Equations*: Part II. SIAM J. Numer. Anal., 36(4):1125–1144, 1999.

[21] BOCHEV, P. B., LAI, J. and OLSON, L.: *A locally conservative, discontinuous least-squares finite element method for the Stokes equations.* Int. J. Numer. Meth. Fluids, 68(6):782–804, 2012.

[22] BOCHEV, P. B., LAI, J. and OLSON, L.: *A non-conforming least-squares finite element method for incompressible fluid flow problems.* Int. J. Numer. Meth. Fluids, 00:1–19, 2012.

[23] BOLTON, P., THATCHER, R. W.: *On mass conservation in least-squares methods.* J. Comput. Phys., 203:287–304, 2005.

[24] BOLTON, P., THATCHER, R. W.: *A least-squares finite element method for the Navier-Stokes equations.* J. Comput. Phys., 213:174–183, 2006.

[25] BRENNER, S. C., SCOTT, L. R.: *The mathematical theory of finite element methods.* 3^{rd} Edition, Springer , New York, 2008.

[26] BREZZI, F.: *On existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers*, RAIRO Model. Math. Anal. Numer., 21, 129–151, 1974.

[27] CAI, Z., LEE, B. and WANG, P.: *Least-squares methods for incompressible Newtonian fluid flow: Linear stationary problems.* SIAM J. NUMER. ANAL., 42(2):843–859, 2009.

[28] CAI, Z., MANTEUFFEL, T. A. and McCORMICK, S. F.: *First-order system least-squares for the Stokes equations, with applications to linear elasticity.* SIAM J. Numer. Anal. 34, pp. 1727-1741, 1997.

[29] CHANG, C. L. and JOHN, L. N.: *Least-Squares Finite Element Method for the Stokes Problem with Zero Residual Mass Conservation.* SIAM J. Numer. Anal., 34(2):480–489, 1997.

[30] CHANG, C. L., YANG, S. Y. and HSU, C. H.: *A least-squares finite element method for incompressible flow in stress-velocity-pressure version.* Comput. Methods in Appl. Mech. Eng., 128(1-2):1–9, 1995.

[31] CHEN, T. F., COX, C. L., LEE, H. C. and TUNG, K. L.: *Least-squares finite element methods for generalized Newtonian and viscoelastic flows.* Applied Numerical Mathematics, 60(10):1024–1040, 2010.

[32] CIPRA, B. A.: *The Best of the 20th Century: Editors Name Top 10 Algorithms.* In: SIAM News 33.4, 2000.

[33] CODD, A. L.: *Elasticity-Fluid Coupled Systems and Elliptic Grid Generation (EGG) based on First-Order System Least Squares (FOSLS).* PhD thesis, University of Colorado, Department of Applied Mathematics, 2001.

[34] CONCUS, P., GOLUB, G. H. and LEARY, D. P. O.: *A Generalized Conjugate Gradient Method for the Numerical Solution of Elliptic Partial Differential Equations.* In: Sparse Matrix Computations, J. Bunch and D. Rose, eds., pp. 309–332, 1976.

[35] CUVELIER, C., SEGAL, A. and VAN STEENHOVEN, A. A.: *Finite element methods and Navier-Stokes equations.* Vol. 22. Springer Science & Business Media, 1986.

[36] DALIMUNTHE, H. S. and SURANA, K. S.: *p-version least-squares finite element formulation for three-dimensional, incompressible, non-isothermal, non-newtonian fluid flow.* Computers and Structures, $58(1) : 85 - 105, 1996$.

[37] DAMANIK, H., HRON, J., OUAZZI, A., TUREK, S.: *Monolithic Newton-multigrid solution techniques for incompressible nonlinear flow models,* International Journal for Numerical Methods in Fluids, Vol: 71, 208-222, 2012.

[38] DAVIES, A. J.: *The Finite Element Method: A First Approach*, Oxford Applied Mathematics and Computing Science Series, 1980.

[39] DAVIS, T. A. and DUFF, I. S.: *A combined unifrontal/multifrontal method for unsymmetric sparse matrices.* ACM Trans. Math. Softw., 25(1): 01–20, 1999.

[40] DAVIS, T. A.: *Direct methods for sparse linear system.* Philadelphia, PA: SIAM, 2006.

[41] DAVIS, T. A., RAJAMANICKAM, S. and SID-LAKHDAR, W. M.: *A survey of direct methods for sparse linear systems.* In: Acta Numerica 25, pp. 383–566, 2016.

[42] DEANG, J. M., GUNZBURGER, M. D.: *Issues related to least-squares finite element methods for the Stokes equations.* SIAM J. Sci. Comput., 20:878–906, 1998.

[43] DING, X. and TSANG, T. H.: *On first-order formulations of the least-squares finite element method for incompressible flows.* Int. J. Comput. Fluid Dynamics, 17(3):183–197, 2003.

[44] DONEA, J. and HUERTA, A.: *Finite element methods for flow problems.* John Willey and Sons, England, 2003.

[45] DUFF, I. S., ERISMAN, I. and REID, J.: *Direct methods for sparse matrices.* London, England: Oxford University Press, 1986.

[46] EDGAR, N. B. and SURANA, K. S.: *p-version least-squares finite element formulation for axisymmetric incompressible non-Newtonian fluid flow.* Computer Methods in Applied Mechanics and Engineering, 113(3-4):271 – 300, 1994.

[47] EDGAR, N. B. and SURANA, K. S.: *p-version least-squares finite-element formulation for axisymmetric incompressible Newtonian and non-Newtonian fluid flow with heat transfer.* Numerical Heat Transfer, Part B: Fundamentals, 27(2):213–235, 1995.

[48] ELMAN, H. C., ERNST, O. G. and O'LEARY, D. P.: *A multigrid method enhanced by krylov subspace iteration for discrete Helmholtz equations.* SIAM J. Sci. Comput., 23 (4):1291–1315, April 2001.

[49] FENG, Z. and SURANA, K. S.: *p-version least-squares finite element formulation for three-dimensional, isothermal, incompressible, non-Newtonian fluid flow.* Computers and Structures, 57(5):799 – 816, 1995.

[50] FLAHERTY, J. E.: *Lecture Notes: Finite element analysis.* CSCI-6860 /MATH-6860.

[51] GIRAULT, V., RAVIART, P.: *Finite Element Methods for Navier-Stokes Equations.* Springer, Berlin, 1986.

[52] GRESHO, P. and SANI, R. *Incompressible Flow and the Finite Element Method.* John Wiley and Sons, 2000.

[53] GRISVARD, P.: *Elliptic Problems in Non-smooth Domains.* Pitman, Boston, 1985.

[54] HACKBUSCH, W.: *Multi-Grid Methods and Applications.* Springer-Verlag, Berlin, ISBN: 0-387-12761-5, 1985.

[55] KAYSER-HEROLD, O.: *Least-Squares Methods for the Solution of Fluid-Structure Interaction Problems.* PhD thesis, TU Braunschweig, 2006.

[56] HESTENES, M. R. and STIEFEL, E.: *Methods of Conjugate Gradients for Solving Linear Systems.* In: Journal of Research of the National Bureau of Standards 49.6, pp. 409–436, 1952.

[57] HEYS, J. J., MANTEUFFEL, T. A., McCORMICK, S. F. and OLSON, L. N.: *Algebraic multigrid for higher-order finite elements.* J. Comput. Phys., 204(2):520–532, 2005.

[58] HEYS, J. J., LEE, E., MANTEUFFEL, T. A. and McCORMICK, S. F.: *On mass conserving least-squares methods.* SIAM J. Sci. Comput., 28:1675–1693, 2006.

[59] HEYS, J. J., LEE, E., MANTEUFFEL, T. A. and McCORMICK, S. F.: *An alternative least-squares formulation of the Navier-Stokes equations with improved mass conservation.* J. Comput. Phys., 226:994–1006, 2007.

[60] HEYS, J. J., LEE, E., MANTEUFFEL, T. A., McCORMICK, S. F. and RUGE, J. W.: *Enhanced mass conservation in least-squares methods for Navier-Stokes equations.* SIAM J. Sci. Comput., 31: 2303–2321, 2009.

[61] HEYS, J. J., MANTEUFFEL, T. A., McCORMICK, S. F., MILANO, M., WESTERDALE, J. and BELOHLAVEK, M.: *Weighted least-squares finite elements based on particle imaging velocimetry data.* J. Comput. Phys., 229(1):107 – 118, 2010.

[62] IPSEN, I. C. F. and MEYER, C. D.: *The Idea behind Krylov Methods.* In: The American Mathematical Monthly 105.10, pp. 889–899, 1998.

[63] JIANG, B.: *The Least-squares finite element method: Theory and applications in computational fluid dynamics and electromagnetics.* Springer-Verlag, 1998.

[64] JOHN, V.: *Higher order finite element methods and multigrid solvers in a benchmark problem for the 3-D Navier-Stokes equations*, Int. J. Numer. Meth. Fluids 40, 775–798, 2002.

[65] JOHNSON, C.: *Numerical solution of partial differential equations by the finite element method.* Cambridge Uni. Press, 1994.

[66] KÖSTER, M.: *A Hierarchical Flow Solver for Optimization with PDE Constraints.* PhD thesis, TU Dortmund, 2011.

[67] KÖSTER, M., OUAZZI, A., SCHIEWECK, F., TUREK, S. and ZAJAC, P.: *New robust nonconforming finite elements of higher order.* Appl. Numer. Math., 62: 166–184, 2012.

[68] KREYSZIG, E.: *Introductory Functional Analysis with Applications.* John Wiley & Sons, New York, 1989.

[69] LADYZHENSKAYA, O.: *The Mathematical Theory of Viscous Incompressible Flows*, Gordon and Breach, New York, 1969.

[70] LUENBERGER, D. G.: *Introduction to Linear and Nonlinear Programming.* Newyork: Addison-Wesley, 1973.

[71] LYNN, P. and ARYA, S.: *Use of least-squares criterion in the finite element method.* Int. J. Numer. Meth. Engng. 6, pp. 75-88, 1973.

[72] MÜLLER, M.: *Adaptive high-resolution finite element schemes.* PhD thesis. Technische Universität Dortmund, 2010.

[73] MÜLLER, M.: *Lecture Notes: Mathematical and practical aspects of finite elements.* 2010.

[74] NICKAEEN, M., OUAZZI, A. and TUREK, S.: *Newton multigrid least-squares FEM for the V-V-P formulation of the Navier-Stokes equations.* J. Comput. Phys. 256, 416–427, 2014.

[75] NICKAEEN, M.: *Efficient FEM and multigrid solvers for the least-squares method with application to non-Newtonian fluid flows.* Phd thesis, TU Dortmund, Fakultät für Mathematik, Lehrstuhl III für Angewandte Mathematik und Numerik, January 2014.

[76] OUAZZI, A.: *A mixed formulation of the Stokes equation in terms of* (ω, p, u). Numer. Algorithms, 21:343–352, 1999.

[77] OUAZZI, A., NICKAEEN, M., TUREK, S., WASEEM, M.: *Newton-Multigrid Least-Squares FEM for S-V-P Formulation of the Navier-Stokes Equations*, Numerical Mathematics and Advanced Applications - ENUMATH2013, Springer v103, 651-659, 2015.

[78] PEDERSEN, M.: *Functional Analysis in Applied Mathematics and Engineering.* Studies in advanced mathematics. Chapman & Hall/CRC, Boca Raton, 2000.

[79] PONTAZA, J. P.: *Least-squares variational principles and the finite element method: theory, form, and model for solid and fluid mechanics.* PhD thesis, Texas A&M University, 2003.

[80] PONTAZA, J. P. and REDDY, J. N.: *Space-time coupled spectral/hp least-squares finite element formulation for the incompressible Navier-Stokes equations.* J. Comput. Phys., 197:418–459, 2004.

[81] PROOT, M. M. J. and GERRITSMA, M. I.: *Mass and momentum conservation of the least-squares spectral element method for the Stokes problem.* Journal of Scientific Computing, 27:389–401, 2006.

[82] RANJAN, R. and REDDY, J. N.: *On multigrid methods for the solution of least-squares finite element models for viscous flows.* Int. J. Comput. Fluid Dyn., 26:45–65, 2012.

[83] RANNACHER, R. and TUREK, S.: *A simple nonconforming quadrilateral stokes element. Numer. Methods Partial Differential Equations.*, 8:97–111, 1992.

[84] RENARDY, M., ROGERS, R.: *Introduction to Partial Differential Equations.* Number 13 in Texts in Applied Mathematics. Springer-Verlag, Berlin, 1993.

[85] SAAD, Y.: *Iterative Methods for Sparse Linear Systems.* SIAM, 2nd edition, 2000.

[86] SCHÄFER et al, M.: *Benchmark computations of laminar flow around a cylinder.* In: Flow simulation with high-performance computers II. Springer, pp. 547–566, 1996.

[87] SCHWARZ, H. R.: *Numerische Mathematik.* Teubner, Stuttgart, Leipzig, 1997.

[88] SCHWARZ, A., STEEGER, K. and SCHRÖDER, J.: *A least-squares mixed finite element for quasi-incompressible elastodynamics.* PAMM, 10(1):219–220, 2010.

[89] STARKE, G.: *Multilevel boundary functionals for least-squares mixed finite element methods.* SIAM J. Numer. Anal., 36:1065–1077, 1999.

[90] TAYLOR, M.: *Partial Differential Equations. Basic Theory.* Springer, New York, 1999.

[91] TUREK, S.: *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach.* Springer-Verlag, 1999.

[92] TUREK, S., HRON, J.: *A monolithic FEM solver for an ALE formulation of fluid–structure interaction with configuration for numerical benchmarking*, In: P. Wesseling, E. Onate, J. Periaux (Eds.), Books of Abstracts European Conference on Computational Fluid Dynamics, nn, 2006, p. 176, Eccomas CFD 2006.

[93] TUREK, S. and SCHÄFER, M.: *Benchmark computations of laminar flow around cylinder.* In E.H. Hirschel, editor, Flow Simulation with High-Performance Computers II, volume 52 of Notes on Numerical Fluid Mechanics, pages 547–566. Vieweg, 1996.

[94] VALLALA, V. P., REDDY, J. N. and SURANA, K. S.: *Alternative least-squares finite element models of Navier-Stokes equations for Power law fluids.* Engineering Computations, 28:828-852, 2011.

[95] VAN DER SLUIS, A. and VAN DER VORST, H. A.: *The rate of convergence of conjugate gradients.* In: Numerische Mathematik 48.5, pp. 543– 560, 1986.

[96] WENDLAND, W. L.: *Elliptic Systems in the Plane.* Pitman, Boston, 1979.

[97] WESSELING, P. and OOSTERLEE, C. W.: *Geometric multigrid with applications to computational dynamics.* Journal of Computational and Applied Mathematics, 128(2):311–334, 2001.

[98] WOBKER, H.: *Efficient Multilevel Solvers and High Performance Computing Techniques for the Finite Element Simulation of Large-Scale Elasticity Problems.* PhD thesis, TU Dortmund, 2010.