Efficient FEM and Multigrid Solvers for the Least-squares Method with Application to Non-Newtonian Fluid Flows

---

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften

Der Fakultät für Mathematik der
Technischen Universität Dortmund
vorgelegt von

Masoud Nickaeen

Efficient FEM and Multigrid Solvers for the Least-squares Methods
with Application to Non-Newtonian Fluid Flows

Masoud Nickaeen

Dissertation eingereicht am: 12. 12. 2013
Tag der mündlichen Prüfung: 28. 01. 2014

Mitglieder der Prüfungskommission

Prof. Dr. Stefan Turek (1. Gutachter, Betreuer)
Prof. Dr. Christian Meyer (2. Gutachter)
Jun.-Prof. Dr. Dominik Göddeke
Prof. Dr. Matthias Röger

to my Mahbub

Acknowledgements

First of all, It would be impossible to overstate the thanks I owe to Prof. Dr. Stefan Turek, my thesis supervisor, for introducing me to the world of applied mathematics and scientific computing. I am really grateful to him believing on me and making this entire project and thesis possible. With his keen sense of the subject and with a lot of experiences in the finite element method, he has based the foundation of my knowledge in this field. I have learned a lot from him during my stay in Germany and would always appreciate it.

Second, I would like to express my sincere gratitude to Dr. Abderrahim Ouazzi for providing invaluable advice on academic and non-academic life problems during my studies at TU Dortmund. He has been a great help, introducing the basics of the least-squares minimization and the function spaces to me. He is the reason I could finish my PhD in the least possible time.

Many thanks also goes to Prof. Dr. Christian Meyer for reviewing my PhD thesis and Jun.-Prof. Dr. Dominik Göddeke and Prof. Dr. Matthias Röger for being a part of my PhD examination committee.

Further, thanks to all my colleagues at chair of applied mathematics, LSIII, for the very friendly and collaborative environment provided to me. Specifically, I would like to thank Dr. Michael Köster for introducing me to the FeatFlow-2 package and for tons of discussions around coding and the finite element theory. Thanks to Sven Buijssen from whom I have learned a lot about the unix-based platforms and technical software issues. I owe Babak Sayyid Hosseini for a lot of fruitful discussions that improved my understandings of the finite element method.

This work was supported by the Mercator Research Center Ruhr (MERCUR) grant Pr-2011-0017 and the scholarship from the graduate school of energy efficient production and logistics at TU Dortmund. I greatly appreciate the funding and support from both parties.

I would like to thank my family, my mother, my father, my brothers and my lovely sister for providing me with a lot of energy and enthusiasm to continue my studies. Thanks for all your continuous support.

And finally, thanks to my beautiful wife *Mahbub* for her limitless support and love without which I could never finish this work.

Dortmund, December, 2013

Masoud Nickaeen

Contents

1

Introduction

In this chapter, we present the motivation of the thesis by a general introduction about the least-squares finite element method. In addition, we summarize the contents of the following chapters.

1.1. Introduction and Motivation

Least-squares variational principles have been recognized as a new tool to solve elliptic partial differential equations [6, 52, 55]. Least-squares finite element methods (LSFEM) can be viewed as an attempt at retaining the advantages of the Rayleigh-Ritz setting. A Rayleigh-Ritz approximation is defined by minimizing the functional over a family of finite-dimensional subspaces. An FEM results when these spaces consist of piecewise polynomial functions defined with respect to a family of grids. When applied to problems such as linear elasticity or the Poisson equation, the Rayleigh-Ritz setting gives rise to FEMs with several advantageous features that led to their great success and popularity [55]:

1. general regions and boundary conditions are relatively easy to treat in a systematic manner
2. the conformity of the finite element spaces suffices to guarantee the stability and optimal accuracy of the approximate solutions
3. all variables can be approximated using a single type of finite element space, e.g., the same degree piecewise polynomials defined with respect to a same grid
4. the resulting linear systems are a) sparse; b) symmetric; c) positive definite.

LSFEMs possess two additional advantageous features that other FEMs, even in the Rayleigh-Ritz setting, do not possess. First, least-squares functionals provide an easily computable residual error indicator that can be used for adapting grids. Second, the treatment of general boundary conditions, including non-homogeneous ones, is greatly facilitated because boundary condition residuals can be incorporated into the least-square functional [55]. These properties can yield the following notable computational advantages and simplifications when properly accounted for in the algorithmic design of least-squares finite element methods [6]:

1. finite element spaces of equal interpolation order, defined with respect to the same triangulation, can be used for all unknowns;
2. algebraic problems can be solved using standard and robust iterative methods, such as conjugate gradient methods.

These benefits hold even for problems like the Navier-Stokes (NS) equations, which yield a saddle-point problem and indefinite linear systems for mixed finite element methods [71]. Besides the above advantages, any LSFEM must also meet additional practicality criteria [55]:

1. bases for conforming subspaces are easily constructed
2. linear systems are easily assembled
3. linear systems are relatively well conditioned.

Unfortunately, naively defined LSFEMs often fail to meet one or more of the practicality criteria. A standard framework (known as the L^2 least-squares FEM) for achieving practical finite element methods should adhere to the following guidelines:

1. Write the system in first-order form
2. Define L^2-norm least-squares functionals
3. Discretize with C^0 finite elements.

The L^2 LSFEM not only recovers all the good properties of the Rayleigh Ritz setting, but also satisfies all three practicality criteria. We can choose standard finite element spaces for which bases are easily constructed. Furthermore, since the functional only involves L^2 inner products, the assembly of the matrix system is accomplished in a standard manner. Finally, it can be shown that the condition number of the matrix system is $O(h^{-2})$.

Owing to its computational advantages, the LSFEM has been extensively used for the numerical solution of a wide range of partial differential equations (PDE), see [6, 33, 52, 55]. One important class of the problems which are studied by many investigators is the general transport equations of heat transfer and fluid flow [52]. Different first-order formulations as well as different stabilization techniques have been introduced and analyzed generally in the book by Bochev and Gunzburger [52] and specifically by other investigators in [1, 2, 15, 26, 38, 60, 65, 78, 79]. Although many theoretical analysis and error estimates have been developed for this class of problems, the investigation of the pros and cons of the proposed LSFEMs deserves a systematic computational study. We systematically split the general transport equations into three important sub-equations, namely the Poisson equation, the diffusion-reaction equation and the advection-diffusion equation. We analyze different first-order systems for each of the sub-equations. The accuracy of the different formulations along with the efficiency of the solvers such as multigrid and conjugate gradient for the solution of the discrete systems are specifically addressed in this work. We perform extensive numerical simulations which involve parametric studies with respect to the various problem-dependent parameters. Moreover, we study the effect of the different finite elements of low and higher order on the overall solution process. In addition, we study the effect of scaling (weighting) techniques on the accuracy of the results and the robustness of the solvers. Such weighting parameters have been traditionally used by many other investigators in the LSFEM literature mostly for analysis purposes [52]. In this direction, we propose a new first-order formulation for the advection-diffusion equations. We provide the coercivity and the a priori error estimates for the new formulation. We show through many numerical examples that the proposed first-order formulation performs well even for the advection-dominated flows. In other words, the stabilized LSFEM [60] which is based on the streamline diffusion or more generally the Galerkin least-squares stabilization technique, does not performs better than the scaled LSFEM for the range of applications that is used in our work. Unlike the stabilized LSFEM, the proposed LSFEM is parameter independent. We provide a complete literature review on these transport problems later in chapter 3.

Recently, there has been substantial interest in the use of least-squares finite element methods for the numerical solution of the Stokes and the Navier-Stokes equations [33, 52]. As it is mentioned earlier, the least-squares based finite element formulations offer several theoretical and computational advantages. Most notably, such formulations circumvent the inf-sup condition of

Ladyzhenskaya-Babuska-Brezzi (LBB). So, the choice of approximating spaces is not subject to any condition, and a single continuous piecewise polynomial space can be used for the approximation of all unknowns [73]. They also yield symmetric positive-definite coefficient matrix and robust iterative solvers can be used to solve resulting system of linear equations. However, these benefits do not come without some drawbacks. The following two are the most significant issues reported in the literature [34]:

- when popular C^0 finite elements are used, the number of dependent variables typically increases over the original formulation (by a factor of 2 to 3 in many cases);
- lack of mass conservation for inflow-outflow problems;
- poor approximation quality when using lower-order finite elements.

The computational advantages associated with the LSFEM are usually sufficient to overcome the first drawback, especially as problems become larger due to the scalability of multigrid solvers which are often used in conjunction with least-squares methods [34, 79]. Heys et al. studied the LSFEM solution of the Stokes equation [28] and the NS equations [34–36] with an algebraic multigrid preconditioned CG method. A geometric multigrid preconditioned CG solver was used by Ranjan and Reddy [64] for the Spectral/hp LSFEM solution of the NS equations. They demonstrated superior convergence of the multigrid solver compared to the Jacobi preconditioning. We extend the multigrid-preconditioned CG (MPCG) solver to the vorticity-based [46] and the stress-based [50, 69] NS formulations. In addition, we use a CG pre/post-smoother to obtain efficient and parameter-free smoothing sweeps. We investigate the performance of the MPCG solver for a wide range of parameters.
By changing the least-squares functional, conservation of mass can be improved at the expense of the other terms in the functional [37, 62]. An additional expense of improving mass conservation is that the performance of standard multigrid solvers may degrade [29, 79]. The use of reduced integration can result in a collocation method and a zero residual for mass conservation [13, 33], although this also results in the loss of positive definiteness of the operator, which has a strong negative impact on a standard multigrid solver [37]. Another remedy for the lack of mass conservation of the 2D problems is to use higher order finite elements [40]. We additionally use this technique to improve the approximation quality of the LSFEM results [69]. Use of the alternative first-order formulations is demonstrated to provide improved mass conservation over existing methods, see e.g. the work of Heys et al. [35] in which mass conservation is improved because of the pressure-velocity coupling along the inflow and outflow boundaries. We investigate the following techniques to analyze the mass conservation of the LSFEMs:

- weighting of the continuity equation
- use of higher order finite elements
- use of alternative first-order formulations.

We perform extensive numerical studies to investigate the effects of these methods on the mass conservation of the LSFEM solutions. In addition, we study the effects of these techniques on the accuracy of the momentum equation approximation and on the performance of the multigrid solvers.

As an important application, we extend the current least-squares finite element framework to the solution of the non-Newtonian fluid flow problems. The non-Newtonian fluids are investigated in the literature using the different first-order formulations [4, 11, 16, 20, 21, 24, 74]. We design a physically motivated weighted least-squares method to provide robust solutions for the non-Newtonian power law and the Carreau law fluid flows. We extend the multigrid-preconditioned

conjugate gradient (MPCG, see [45, 46, 50]) solver to efficiently solve the stress-based formulation of the non-Newtonian fluids. This is for the first time that a multigrid solver is used for the LSFEM solution of the non-Newtonian fluid flow problems. We study the Newton and the fixed-point linearization methods and analyze the performance of the MPCG solver for these methods.

1.2. Thesis Outline

We are mainly concerned with the LSFEM solution of the advection-diffusion-reaction equation and the incompressible NS equations for both the Newtonian and the non-Newtonian fluid flow problems. In this thesis, each of the following chapters provides its own introduction and conclusion.

In *chapter 2*, we present the basic idea of the least-squares finite element method (LSFEM) for the solution of the abstract linear and nonlinear partial differential equations (PDE). We define the general framework for the continuous and the discrete least-squares minimization problems, based on which we design the LSFEM techniques for the solution of different model PDEs in the following chapters. We use a *nonlinear basic iteration* in combination with the least-squares minimization technique for the solution of the non-linear PDEs. In addition, we briefly explain the different solvers used for the solution of the discrete least-squares systems. These solvers include specifically the conjugate gradient method, the geometric multigrid method and the multigrid-preconditioned conjugate gradient method.

In *chapter 3*, we focus on the application of the LSFEM to the general transport equations of heat transfer and fluid flow. We design efficient and robust LSFEMs for the transport equations which include the main transport mechanisms of the Navier-Stokes equations. In addition, we investigate accurate multigrid solvers for the solution of the discretized systems arising from such equations. We propose a new first-order formulation for the solution of the advection-diffusion equations and prove the uniqueness and a priori error estimates for this formulation. Moreover, we analyze the effect of the weighting techniques on the accuracy and the robustness of the LSFEM method.

In *chapter 4*, we study two first-order formulations of the stationary incompressible Navier-Stokes equations. For the discrete systems, we use a conjugate gradient (CG) solver accelerated with a geometric multigrid preconditioner for the complete system. In addition, we employ a Krylov space smoother inside of the multigrid which allows a parameter-free smoothing. Combining this linear solver with the Newton linearization, we construct a non-linear solver. We perform systematic parametric studies to investigate the robustness and efficiency of the solver. Moreover, we analyze different strategies to enhance the mass conservation of the least-squares method for the inflow-outflow problems. In order to quantitatively analyze the proposed LSFEMs, we perform extensive numerical simulations on different flow configurations of benchmarking character.

In *chapter 5*, we present the numerical simulation of the non-Newtonian power law and the Carreau law fluid flow problems. Both the shear thinning and the shear thickening fluids are investigated. We design a physically motivated weighted least-squares method to provide robust solutions for the non-Newtonian fluid flows. We extend the multigrid-preconditioned conjugate gradient solver to efficiently solve the stress-based formulation of the non-Newtonian fluids. We study the Newton and the fixed-point linearization methods and analyze the performance of the MPCG solver for these methods. We compare the accuracy of the numerical solutions with available analytic solutions.

In *chapter 6*, we summarize the LSFEM methods studied in this work and present some future works as an outlook of the current research.

2

The Least-squares Finite Element Method

After a short introduction of the notation used, we present the basic idea of the least-squares finite element method (LSFEM) for the solution of linear partial differential equations (PDE). We define the general framework for the continuous and the discrete least-squares minimization problems, based on which we design the LSFEM techniques for the solution of different model PDEs in the following chapters. We show that the least-squares finite element model yields a convex, unconstrained minimization problem with a unique minimizer that coincides with the best approximate solution to the linear PDE model in a well defined norm. In addition, we show that the resulting discrete algebraic system has a symmetric positive definite coefficient matrix. We further extend the abstract framework to the nonlinear PDEs. We use a *nonlinear basic iteration* in combination with the least-squares minimization technique. Moreover, we briefly explain the different solvers used for the solution of the discrete least-squares systems. These solvers include specifically the conjugate gradient method, the geometric multigrid method and the multigrid-preconditioned conjugate gradient method.

2.1. Notation

We consider an open bounded domain $\Omega \subset \mathbb{R}^d$ with a piecewise smooth boundary Γ, where $d = 1,2,3$, is the space dimension. For different boundary conditions the boundary segment will be distinguished by a subscript, e.g. Γ_D and Γ_N for the Dirichlet and the Neumann parts of the boundary, respectively.

We use an operator notation to display the PDE problems and the least-squares systems in the continuous form. The gradient, divergence and the Laplace operators are uniquely defined as ∇, $\nabla\cdot$ and Δ, respectively. The curl operator has two different interpretations. When applied to a scalar u, it gives a vector as follows

$$\nabla \times u = \begin{pmatrix} \frac{\partial u}{\partial y} \\ -\frac{\partial u}{\partial x} \end{pmatrix}, \tag{2.1.1}$$

and when applied to a vector $\mathbf{u} = (u_1, u_2)^T$, it gives a scalar

$$\nabla \times \mathbf{u} = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\ . \tag{2.1.2}$$

The standard notation for the Sobolev spaces is used as in [68]. The space of square integrable functions defined on the domain Ω is denoted by $L^2(\Omega)$ and has the scalar product

$$(u,v)_0 = \int_\Omega uv\, d\Omega \quad u,v \in L^2(\Omega), \tag{2.1.3}$$

which induces the following L^2-norm

$$\|u\|_0 = (u,u)_0^{1/2} = \left(\int_\Omega |u|^2 \, d\Omega \right)^{1/2} \tag{2.1.4}$$

A subspace of $L^2(\Omega)$ functions with zero-mean value is defined as

$$L_0^2(\Omega) = \left\{ q \in L^2(\Omega) \mid \int_\Omega q \, d\Omega = 0 \right\} . \tag{2.1.5}$$

For any non-negative integer m, we define the Sobolev space $H^m(\Omega)$ and the associated inner product and norm are denoted by $(\cdot,\cdot)_m$ and $\|\cdot\|_m$, respectively. Therefore, the H^1 space is defined as follows

$$H^1(\Omega) = \left\{ q \in L^2(\Omega) \mid \nabla q \in L^2(\Omega) \right\} , \tag{2.1.6}$$

which has the inner product

$$(p,q)_1 = (p,q)_0 + (\nabla p, \nabla q)_0 , \tag{2.1.7}$$

and the associated norm is given by

$$\|q\|_1^2 = \|q\|_0^2 + \|\nabla q\|_0^2 . \tag{2.1.8}$$

Under some assumptions on Ω, functions in $H^1(\Omega)$ have well-defined traces in $H^{1/2}(\Gamma)$. The norm on $H^{1/2}(\Gamma)$ can be defined as follows

$$\|\phi\|_{1/2,\Gamma} = \inf_{\psi \in H^1(\Omega), \psi = \phi \ on \ \Gamma} \|\psi\|_1 . \tag{2.1.9}$$

Further, we define the $H(\mathrm{div},\Omega)$ space

$$H(\mathrm{div},\Omega) = \left\{ \mathbf{v} \in [L^2(\Omega)]^2 \mid \nabla \cdot \mathbf{v} \in L^2(\Omega) \right\} , \tag{2.1.10}$$

which is a Hilbert space with inner product

$$(\mathbf{u},\mathbf{v})_{\mathrm{div}} = (\mathbf{u},\mathbf{v})_0 + (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v})_0 , \tag{2.1.11}$$

and its corresponding norm is defined as

$$\|\mathbf{v}\|_{\mathrm{div}}^2 = \|\mathbf{v}\|_0^2 + \|\nabla \cdot \mathbf{v}\|_0^2 . \tag{2.1.12}$$

Also we define the Hilbert space $H(\mathrm{curl},\Omega)$

$$H(\mathrm{curl},\Omega) = \left\{ \mathbf{v} \in [L^2(\Omega)]^2 \mid \nabla \times \mathbf{v} \in L^2(\Omega) \right\} , \tag{2.1.13}$$

with the inner product

$$(\mathbf{u},\mathbf{v})_{\mathrm{curl}} = (\mathbf{u},\mathbf{v})_0 + (\nabla \times \mathbf{u}, \nabla \times \mathbf{v})_0 , \tag{2.1.14}$$

and the respective norm is defined as

$$\|\mathbf{v}\|_{\mathrm{curl}}^2 = \|\mathbf{v}\|_0^2 + \|\nabla \times \mathbf{v}\|_0^2 . \tag{2.1.15}$$

We further define the following Hilbert spaces

$$H_{g,D}^1(\Omega) = \left\{ q \in H^1(\Omega) \mid q = g_D \ on \ \Gamma_D \right\} , \tag{2.1.16}$$

$$H_{g,N}(\mathrm{div},\Omega) = \left\{ \mathbf{v} \in H(\mathrm{div},\Omega) \mid \mathbf{v} \cdot \mathbf{n} = g_N \ on \ \Gamma_N \right\} , \tag{2.1.17}$$

$$H_{g,D}(\mathrm{curl},\Omega) = \{\mathbf{v} \in H(\mathrm{curl},\Omega) \mid \mathbf{v}\times\mathbf{n} = g_D \text{ on } \Gamma_D\}\,, \tag{2.1.18}$$

which are the subspaces of $H^1(\Omega)$, $H(\mathrm{div},\Omega)$ and $H(\mathrm{curl},\Omega)$, respectively. We denote by $H^{-1}(\Omega)$ the dual of $H^1(\Omega)$ associated with the norm

$$\|p\|_{-1} = \sup_{q\in H^1(\Omega)} \frac{\langle p,q\rangle_\Omega}{\|q\|_1}\,, \tag{2.1.19}$$

where $\langle\cdot,\cdot\rangle_\Omega$ denotes the dual pairing:

$$\langle p,q\rangle_\Omega = \int_\Omega p\cdot q\, d\Omega\,. \tag{2.1.20}$$

The Gâteaux derivative $\mathcal{D}$ of functional $\mathcal{F}(u)$ with respect to v is also required:

$$\mathcal{D}\mathcal{F}(u)[v] = \left.\frac{\partial\mathcal{F}(u+tv)}{\partial t}\right|_{t=0} \tag{2.1.21}$$

2.2. Continuous Least-squares Principle

We present the basic idea of the least-squares FEM in this section. The main materials are based on the book by Jiang [33], the book by Bochev and Gunzburger [52] and the book by Turek [71]. Throughout this section, $\mathcal{L}$ denotes a linear first-order differential operator that acts on functions defined on some bounded, open domain Ω and $\mathcal{R}$ denotes a linear operator which is applied to functions defined on the boundary Γ of Ω. Both $\mathcal{L}$ and $\mathcal{R}$ may depend on the spatial variable $\mathbf{x}$. We consider an abstract boundary value problem

$$\begin{cases} \mathcal{L}\mathbf{u} = \mathbf{f} & \text{in } \Omega \\ \mathcal{R}\mathbf{u} = \mathbf{g} & \text{on } \Gamma, \end{cases} \tag{2.2.1}$$

where $\mathbf{f}$ and $\mathbf{g}$ denote the data functions. Concerning (2.2.1), we make the following assumptions.

A.1 There exist Hilbert spaces $X = X(\Omega)$, $Y = Y(\Omega)$, $Z = Z(\Gamma)$ such that the mapping $\mathbf{u} \to (\mathcal{L}\mathbf{u},\mathcal{R}\mathbf{u})$ is a homeomorphism $X \to Y\times Z$.

A.2 The operator $(\mathcal{L}\mathbf{u},\mathcal{R}\mathbf{u})$ is of Fredholm type, i.e. it has a closed range and both the kernel and the co-range are finite dimensional.

These assumptions are sufficiently general to include a wide range of partial differential equation problems. An important consequence of the above assumptions is the existence of two positive constants C_1 and C_2 whose values are independent of $\mathbf{u}$ and such that

$$C_2\|\mathbf{u}\|_X \le \|\mathcal{L}\mathbf{u}\|_Y + \|\mathcal{R}\mathbf{u}\|_Z \le C_1\|\mathbf{u}\|_X\,. \tag{2.2.2}$$

The inequalities in (2.2.2) describe a relation between the solution and data of a boundary value problem that is fundamental to least-squares principles. It defines the proper balance between the solution "energy" and the residual "energy".

The continuous least-squares principle (CLSP) for (2.2.1) stems directly from the solution-data balance defined by (2.2.2). The data spaces Y and Z provide the norms for measuring the "energy" of the residuals while the solution space X serves as a trial space for candidate minimizers of the "energy" functional. Specifically, we define the artificial, quadratic, convex least-squares "energy" functional

$$\mathcal{J}(\mathbf{u};\mathbf{f},\mathbf{g}) = \|\mathcal{L}\mathbf{u}-\mathbf{f}\|_Y^2 + \|\mathcal{R}\mathbf{u}-\mathbf{g}\|_Z^2\,, \tag{2.2.3}$$

and the continuous least-squares principle for the problem

$$\text{find} \quad \mathbf{u} \in X \quad \text{such that} \quad \mathbf{u} = \underset{\mathbf{v} \in X}{\operatorname{argmin}} \; \mathcal{J}(\mathbf{v}; \mathbf{f}, \mathbf{g}) \, . \tag{2.2.4}$$

The minimizer of the problem (2.2.4) is the solution of the variational problem

$$\text{find} \quad \mathbf{u} \in X \quad \text{such that} \quad \mathcal{A}(\mathbf{u}; \mathbf{v}) = \mathcal{F}(\mathbf{v}) \quad \forall \mathbf{v} \in X \, , \tag{2.2.5}$$

where bilinear form $\mathcal{A}(\cdot;\cdot)$ and the linear functional $\mathcal{F}(\cdot)$ are given by

$$\mathcal{A}(\mathbf{u}; \mathbf{v}) = \big(\mathcal{L}\mathbf{u}, \mathcal{L}\mathbf{v}\big)_Y + \big(\mathcal{R}\mathbf{u}, \mathcal{R}\mathbf{v}\big)_Z \tag{2.2.6}$$

and

$$\mathcal{F}(\mathbf{v}) = \big(\mathbf{f}, \mathcal{L}\mathbf{v}\big)_Y + \big(\mathbf{g}, \mathcal{R}\mathbf{v}\big)_Z, \tag{2.2.7}$$

respectively. Following from the assumptions **A**.1 and **A**.2 it can be shown that the continuous least-squares principle (2.2.4) is well posed and the unique minimizer of (2.2.3) coincides with the unique solution $\mathbf{u} \in X$ of (2.2.1), see [52] for further details.

The boundary conditions can be treated in two different ways. Either the functional (2.2.3) is used directly for minimization, or the function space X, in which the solution is sought, is replaced by:

$$X_{\mathbf{g}}(\Omega) = \{\mathbf{u} \in X(\Omega) \mid \mathcal{R}\mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) \; \forall \mathbf{x} \in \Gamma\} \, . \tag{2.2.8}$$

Functions in $X_{\mathbf{g}}$ already satisfy the desired boundary conditions and hence the residuum of the boundary conditions in equation (2.2.1) will automatically be zero. In this thesis, we use the former technique for the Neumann boundary conditions, while the later technique is used for the implementation of the Dirichlet boundary conditions.

2.2.1. Operator form of the CLSP

It is worth noticing that, for sufficiently smooth $\mathbf{u}$ and $\mathbf{f}$, (2.2.5) is equivalent to

$$\text{find} \quad \mathbf{u} \in X \quad \text{such that} \quad \big(\mathcal{L}^*\mathcal{L}\mathbf{u}, \mathbf{v}\big)_Y + \big(\mathcal{R}^*\mathcal{R}\mathbf{u}, \mathbf{v}\big)_Z = \big(\mathcal{L}^*\mathbf{f}, \mathbf{v}\big)_Y + \big(\mathcal{R}^*\mathbf{g}, \mathbf{v}\big)_Z \quad \forall \mathbf{v} \in X \, , \tag{2.2.9}$$

where $\mathcal{L}^*$ and $\mathcal{R}^*$ represent the formal adjoint of $\mathcal{L}$ and $\mathcal{R}$, respectively. It is common then, to look at the formal normal operators, $\mathcal{L}^*\mathcal{L}$ and $\mathcal{R}^*\mathcal{R}$, for insight into the nature of the linear system that results from the least-squares discretization. Note that, in general, the formal normal least-squares operators are self-adjoint and nonnegative, a quality that the original continuum operators may not have.

2.3. Discrete Least-squares Principle

The discrete least-squares principle (DLSP) is obtained by restricting (2.2.5) to discrete, e.g. finite element, space X_h of the finite dimensional space X. This process leads to the discrete variational problem given by

$$\text{find} \quad \mathbf{u}_h \in X_h \quad \text{such that} \quad \mathcal{A}_h(\mathbf{u}; \mathbf{v}) = \mathcal{F}_h(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in X_h \, . \tag{2.3.1}$$

We proceed to define a discrete problem by choosing appropriate finite element spaces for each of the components of the vector valued function $\mathbf{u}$. There are no restrictive compatibility conditions on the discrete spaces, so we are able to choose the same finite element subspaces for each of the primary variables.

2.3.1. Optimality Conditions of the LSFEM

The ultimate goal is to use (2.3.1) to compute the approximate solution to (2.2.1). The least-squares functional is consistent in the sense that for sufficiently smooth data $\mathbf{f}$, $\mathbf{g}$ and smooth solutions $\mathbf{u}$ of (2.2.1), we have

$$\mathcal{J}(\mathbf{u};\mathbf{f},\mathbf{g}) = 0\,. \tag{2.3.2}$$

Furthermore, by construction, the least-squares functional is convex and positive. Which allows us to define an energy inner product

$$((\cdot,\cdot))_E \;:\; X\times X \to \mathbb{R}, \tag{2.3.3}$$

and an associated energy norm

$$\|\cdot\|_E = \mathcal{J}(\cdot;\mathbf{0},\mathbf{0})^{1/2} \;:\; X \to \mathbb{R}\,. \tag{2.3.4}$$

Having "energy" functional with the above characteristics and taking $X_h \subset X$, i.e. working with conforming finite elements, leads to the following important results

1. the variational problem (2.3.1) has a unique solution given by $u_h \in X_h$, and
2. $\mathbf{u}_h$ is the orthogonal projection of $\mathbf{u}$ with respect to the energy inner product (2.3.3), and thus represents the best approximation solution in the energy norm (2.3.4), i.e.

$$\|\mathbf{u}-\mathbf{u}_h\|_E = \inf_{\mathbf{v}_h\in X_h} \|\mathbf{u}-\mathbf{v}_h\|_E\,. \tag{2.3.5}$$

To clarify the first result above, let ϕ_i^h denotes a basis for X_h so that $\mathbf{u}_h = \sum_{i=1}^N U_i\phi_i^h$. It is clear that (2.3.1) is a linear system of algebraic equations for unknown vector $\mathbf{U}$. The matrix and the right hand side of this system are

$$A_{ij} = ((\phi_j^h,\phi_i^h))_E \quad \text{and} \quad F_i = ((\mathbf{u},\phi_i^h))_E\,.$$

Matrix A is symmetric and from the positivity assumption of the functional we conclude that A is also positive definite. As a result, the system $A\mathbf{U}=\mathbf{F}$ has a unique solution.
For the conforming finite element methods, the optimality result (2.3.5) can be deduced based on the a priori estimate (2.2.2) with the help of arguments from the standard finite element theory.

2.3.2. Practicality Conditions of the LSFEM

Another very important concern in designing LSFEMs is that of practicality. Here, we refer to obtaining a linear system of equations which is computable and whose solution is no more difficult to find than that of other numerical discretization techniques. The computability concern can generally be enforced by requiring the spaces Y and Z to be Sobolev spaces of integer order. Further, the space X should be discretized with standard finite elements, i.e. C^0 finite element spaces. Finally, the condition number of the resulting system should not grow (as a function of h) faster than that of a more traditional finite element method. All of these requirements for practicality can influence whether or not the estimate (2.2.2) can be shown to be satisfied.

A standard framework (known as the L^2 least-squares FEM) for achieving practical finite element methods should adhere to the following guidelines:

- Write the system in first-order form.
- Set the space X to be a product of H^1 spaces.

- Set the spaces Y and Z to be products of L^2 spaces.
- Discretize with C^0 finite elements.

When this framework allows estimate (2.2.2), the least-squares finite element method is optimal and practical. When the satisfaction of (2.2.2) can not be shown, modifications may need to be made to achieve optimality, and careful choices will need to be made to retain practicality. Some common changes include the following:

- Use intermediate spaces such as $H^{1/2}$ or weaker spaces such as H^{-1}.
- Reformulate the system of equations.
- Replace impractical norms by weighted L^2 norms in the discrete problem.
- Use a non-conforming discretization.

Throughout this work we always use the basic L^2 least-squares finite element framework. Although, in some cases the use of L^2-norms for the residual energy does not necessarily lead to mathematically correct energy balances and the resulting methods cannot be proved as mathematically optimal. However, thanks to the generality of the assumptions on the least-squares functionals (the consistency (2.3.2) and the positivity conditions) one can satisfy these two conditions almost automatically by any sensible definition of a least-squares functional. These methods are generally categorized as non-equivalent least-squares techniques. The non-equivalent LSFEM has been by a wide margin the most commonly used technique in the least-squares context. The reason for this is the fact that combination of first-order systems with L^2-norms to measure the residual energy leads to a very simple and easy to implement scheme.

2.4. LSFEM for Nonlinear Problems

There are two main steps in the numerical solution of every nonlinear PDE, namely the linearization and the discretization steps. Depending on the order in which these steps are used to solve the nonlinear problem, two approaches can be considered. The first approach is to discretize the nonlinear equations and then to apply the linearization method to the discrete systems. The second approach is to linearize the nonlinear differential operator on the continuous level and then to proceed with the discretization. These two approaches yield to similar system of equations in the context of the Galerkin variational principle. However, for the least-squares finite element method these approaches produce different results, see [41] for further discussions. In this work we use the second approach.

2.4.1. Nonlinear Basic Iteration

Consider an arbitrary first-order nonlinear partial differential operator $\mathcal{T}(\cdot)$

$$\mathcal{T}(\mathbf{u})\mathbf{u} = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.4.1}$$

which is assumed to be Frećhet differentiable. We solve the equation (2.4.1) with a *nonlinear basic iteration*, see Turek [71], which can be formulated as the following three-step algorithm 2.1.

In the first step, we calculate the nonlinear system residual $\mathbf{d}^n$ based on the known values $\mathbf{u}^n$ from the previous iteration superscripted by n, see equation (2.4.2). Next, we approximate the nonlinear operator $\mathcal{T}(\mathbf{u}^n)$ with $\tilde{\mathcal{T}}(\mathbf{u}^n)$ and solve the auxiliary system (2.4.3) for the solution update $\delta\mathbf{u}^n$. The

Algorithm 2.1 Basic nonlinear iteration

Given: Iterate $\mathbf{u}^n$
Perform: The following three substeps to obtain $\mathbf{u}^{n+1}$

1. Calculate the nonlinear residual $\mathbf{d}^n$,

$$\mathbf{d}^n = \mathcal{T}(\mathbf{u}^n)\mathbf{u}^n - \mathbf{f}. \tag{2.4.2}$$

2. Solve an auxiliary subproblem for $\delta\mathbf{u}^n$ with right hand side $\mathbf{d}^n$

$$\tilde{\mathcal{T}}(\mathbf{u}^n)\,\delta\mathbf{u}^n = \mathbf{d}^n\,. \tag{2.4.3}$$

3. Update $\mathbf{u}^n$ via the auxiliary solution $\delta\mathbf{u}^n$ and the relaxation parameter β, and obtain $\mathbf{u}^{n+1}$,

$$\mathbf{u}^{n+1} = \mathbf{u}^n - \beta\,\delta\mathbf{u}^n\,. \tag{2.4.4}$$

operator $\tilde{\mathcal{T}}(\mathbf{u}^n)$ represents an approximation of the Frećhet derivative of $\mathcal{T}$ with respect to the last iterate $\mathbf{u}^n$ and allows different resulting schemes. If we choose $\tilde{\mathcal{T}}(\mathbf{u}^n)$ to be the exact Frećhet derivative, we obtain the standard Newton iteration scheme

$$\tilde{\mathcal{T}}(\mathbf{u}^n)\,\delta\mathbf{u}^n = \mathcal{D}\mathcal{T}(\mathbf{u}^n)[\delta\mathbf{u}^n]\,. \tag{2.4.5}$$

This method is known to be quadratically converging. Since the convergence radius of the Newton method is small, a very good initial iterate $\mathbf{u}^0$ is required to obtain a convergent scheme. Other possibility for $\tilde{\mathcal{T}}(\mathbf{u}^n)$ leads to the so-called *fixed point* scheme. The approximate Frećhet derivative may be chosen as the operator itself

$$\tilde{\mathcal{T}}(\mathbf{u}^n) = \mathcal{T}(\mathbf{u}^n)\,. \tag{2.4.6}$$

The corresponding scheme has a larger convergence radius as compared to the Newton method, thus it is not very sensitive to the choice of the initial iterate. Moreover, the fixed point method has only linear or super-linear convergence which means more nonlinear iterations are required for convergence. Finally, we update $\mathbf{u}^n$ via the auxiliary solution $\delta\mathbf{u}^n$ to obtain the new solution $\mathbf{u}^{n+1}$, see equation (2.4.4). Here, β is the damping parameter which has to be chosen appropriately, see Turek [71] for further details.

We may stop the iteration if the maximum number of nonlinear iterations is performed, $n \geq N_{max}$, or if the nonlinear residual is sufficiently small in a certain norm,

$$\left\|\mathcal{T}(\mathbf{u}^{n+1})\mathbf{u}^{n+1} - \mathbf{f}\right\| \leq TOL\,. \tag{2.4.7}$$

In order to optimize the convergence behavior of the Newton method with respect to the initial guess, one possibility is to start the nonlinear iterations with the fixed point method, and to use the solution as an initial guess for the Newton method. Another possibility to improve the Newton method's convergence is to perform homotopy-type successive Newton runs. For example, in the solution of the Navier-Stokes equations the Reynolds number determines the degree of nonlinearity of the problem. One may start with a lower Reynolds number problem and then use the solution as an initial guess for the same problem with a slightly increased Reynolds number. This procedure will be continued until the solution of the problem at the desired Reynolds number is obtained. We employ both of the above methods in this research. For a detailed review of the least-squares researches with both methods refer to [41].

2.4.2. Least-Squares Principles for the Nonlinear Problems

The second step in the nonlinear iteration (Algorithm 2.1) involves the solution of a linear system of equations (2.4.3). We use the least-squares minimization method to solve the system. As it is previously explained for the linear PDE problems, the least-squares "energy" functional is obtained from the residual of the equation (2.4.3) as follows

$$\mathcal{J}(\delta\mathbf{u};\mathbf{d}) = \left\|\tilde{\mathcal{T}}(\mathbf{u})\,\delta\mathbf{u} - \mathbf{d}\right\|_Y . \tag{2.4.8}$$

The first variation of the above functional leads to the following variational problem in every nonlinear iteration

$$\text{find} \quad \delta\mathbf{u} \in X \quad \text{such that} \quad \mathcal{A}(\delta\mathbf{u};\mathbf{v}) = \mathcal{F}(\mathbf{v}) \quad \forall \mathbf{v} \in X , \tag{2.4.9}$$

where the bilinear form $\mathcal{A}(\cdot;\cdot)$ and the linear functional $\mathcal{F}(\cdot)$ are given by

$$\mathcal{A}(\delta\mathbf{u};\mathbf{v}) = \left(\tilde{\mathcal{T}}(\mathbf{u})\,\delta\mathbf{u}, \tilde{\mathcal{T}}(\mathbf{u})\,\mathbf{v}\right)_Y \tag{2.4.10}$$

and

$$\mathcal{F}(\mathbf{v}) = \left(\mathbf{d}, \tilde{\mathcal{T}}(\mathbf{u})\,\mathbf{v}\right)_Y . \tag{2.4.11}$$

Following the standard LSFEM procedure of the previous section, we restrict the variational problem (2.4.9) to the finite dimensional spaces. The resulting algebraic system matrix, which is obviously symmetric, will be solved in every nonlinear iteration to obtain $\delta\mathbf{u}$.

2.5. Solution of the Discrete Least-squares Systems

The discrete linear system of equations resulting from the least-squares finite element method always has a symmetric and positive definite (SPD) coefficient matrix. Therefore, it is appropriate to take full advantage of the symmetric positive definiteness by using solvers specially designed for such systems. In addition, the resulting system matrix is sparse due to the properties of the interpolation functions used in the finite element discretization. For the large sparse linear systems of equations we encounter in this thesis, direct methods become impractical. In such cases, storage space may be limited in terms of available computer memory and solving times become non-optimal in terms of CPU time. The main focus will be then on the iterative solvers. In this section we briefly discuss the conjugate gradient method as a Krylov subspace solver suitable for SPD systems, and multigrid as a highly efficient defect correction scheme for sparse linear systems arising in the discretization of (elliptic) partial differential equations. A combination of these two iterative linear solvers is then considered to improve the overall efficiency of the solution method. The main material and the notations of this chapter are based on [27, 67, 75], in which detailed theoretical results about the methods are discussed.

2.5.1. Conjugate Gradient Solver

Let us denote the linear system of equations with

$$\mathbf{A}\mathbf{x} = \mathbf{f} \tag{2.5.1}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a sparse symmetric positive definite matrix and $\mathbf{f} \in \mathbb{R}^n$ is the right hand side vector resulting from the least-squares finite element method.

The conjugate gradient (CG) method derives its name from the fact that it generates a sequence of conjugate (or orthogonal) vectors. These vectors are the residuals of the iterates. They are also

the gradients of a quadratic functional, the minimization of which is equivalent to solving the linear system. The CG is an extremely effective method when the coefficient matrix is symmetric positive definite, since storage for only a limited number of vectors is required. The popularity of the CG method is due to its attractive properties:

- the method will give the exact solution after at most n steps in the absence of roundoff errors,
- the rate of convergence can be significantly improved with various preconditioning techniques,
- the method is parameter-free, i.e. the user is not required to empirically choose parameters.

Given an initial guess $\mathbf{x}^{(0)}$ and the initial residual $\mathbf{r}^{(0)} = \mathbf{f} - \mathbf{A}\mathbf{x}^{(0)}$, the Krylov subspace $K_m(\mathbf{A}, \mathbf{r}^{(0)})$ is the m-dimensional vector space defined as

$$K_m(\mathbf{A}, \mathbf{r}^{(0)}) = \text{span}\{\mathbf{r}^{(0)}, \mathbf{A}\mathbf{r}^{(0)}, ..., \mathbf{A}^{m-1}\mathbf{r}^{(0)}\}. \tag{2.5.2}$$

The unpreconditioned conjugate gradient method constructs the i iterate $\mathbf{x}^{(i)}$ as an element of $\mathbf{x}^{(0)} + K_i(\mathbf{A}, \mathbf{r}^{(0)})$ so that $(\mathbf{x}^{(i)} - \hat{\mathbf{x}})^T \mathbf{A}(\mathbf{x}^{(i)} - \hat{\mathbf{x}})$ is minimized, where the $\hat{\mathbf{x}}$ is the exact solution of $\mathbf{A}\mathbf{x} = \mathbf{f}$. This minimum is guaranteed to exist in general only if $\mathbf{A}$ is symmetric positive definite. The preconditioned version of the method uses a different subspace for constructing the iterates, but it satisfies the same minimization property, although over this different subspace. It requires in addition that the preconditioner $\mathbf{M}$ is symmetric and positive definite. The pseudocode for the preconditioned CG method is presented in Algorithm 2.2.

Algorithm 2.2 Preconditioned conjugate gradient method

Compute $\mathbf{r}^{(0)} = \mathbf{f} - \mathbf{A}\mathbf{x}^{(0)}$ with an initial guess $\mathbf{x}^{(0)}$
for $i = 1, 2, ...$ **do**
  solve the preconditioned system $\mathbf{M}\mathbf{z}^{(i-1)} = \mathbf{r}^{(i-1)}$
  $\rho_{i-1} = \mathbf{r}^{i-1^T} \mathbf{z}^{(i-1)}$
  if $i = 1$ **then**
    $\mathbf{p}^{(1)} = \mathbf{z}^{(0)}$
  else
    $\beta_{i-1} = \rho_{i-1} / \rho_{i-2}$
    $\mathbf{p}^{(i)} = \mathbf{z}^{(i-1)} + \beta_{i-1} \mathbf{p}^{(i-1)}$
  end if
  $\mathbf{q}^{(i)} = \mathbf{A}\mathbf{p}^{(i)}$
  $\alpha_i = \rho_{i-1} / (\mathbf{p}^{(i)^T} \mathbf{q}^{(i)})$
  $\mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \alpha_i \mathbf{p}^{(i)}$
  $\mathbf{r}^{(i)} = \mathbf{r}^{(i-1)} - \alpha_i \mathbf{q}^{(i)}$
  check the convergence $\left\| \mathbf{r}^{(i)} / \mathbf{r}^{(0)} \right\| < \varepsilon$
end for

2.5.2. Multigrid Solver

Another efficient alternative for the solution of large and sparse linear systems is to use the multigrid or multilevel methods (see for instance [27, 67, 75]). These methods were mainly developed for linear systems arising from the discretization of elliptic PDEs and later on extended to handle other types of PDEs, including nonlinear ones. In the subject of numerical analysis, these methods form a group of algorithms for the solution of differential equations using a hierarchy of

discretizations. Multigrid solvers are regarded as the most efficient solvers for solving the large sparse linear systems, in particular for those arising from the discretization of PDEs where the conditions number of the system matrix deteriorates with increasing the problem size.

In contrast to the other iterative solvers these solvers converge independent of the mesh size and require only a linear amount of operations with respect to the number of unknowns. The efficiency and the robustness of these solvers crucially depend on the smoothing operator. Multigrid methods can be classified into two categories:

- Algebraic Multigrid Methods (AMG)
- Geometric Multigrid Methods (GMG)

The primary difference between the algebraic and geometric multigrid algorithms is that the AMG methods require only a single mesh information and the system matrices of the coarse grids are constructed by using algebraic operations [67]. On the other hand, the GMG generates the coarser levels by using the hierarchy of mesh refinements. The multigrid methods have become quite popular for the solution of discretized linear systems due to their convergence behavior. Throughout this thesis, we will restrict to the geometric multigrid methods (see [27, 44, 75] for more details).

We formulate the standard multigrid algorithm by considering the linear system of equations in (2.5.1) where $n = n(k)$ denotes the number of degrees of freedom on the mesh level k; $k = 1, ..., L \in \mathbb{N}$. Let

$$I_{k-1}^{k} : \mathbb{R}^{n(k-1)} \mapsto \mathbb{R}^{n(k)} \tag{2.5.3}$$

denotes the prolongation operator and

$$I_{k}^{k-1} : \mathbb{R}^{n(k)} \mapsto \mathbb{R}^{n(k-1)} \tag{2.5.4}$$

the corresponding restriction.

We assume the existence of a hierarchy of levels $k, k = 1, ..., L$ created by using standard refinement scheme [71] of a coarse mesh. On each of these levels k, we have to assemble the discrete problem matrix $\mathbf{A}^k$ and the corresponding right hand side $\mathbf{f}^k$. The right hand side $\mathbf{f}$ is specified on the finest level N only, while all other $\mathbf{f}^k$ are generated during the multigrid algorithm. Then, a standard geometric multigrid algorithm $\mathrm{MG}(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, k)$ for the solution of 2.5.1 is described in Algorithm 2.3.

Algorithm 2.3 Standard Multigrid Algorithm

Recursive linear multigrid algorithm $\mathrm{MG}(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, k)$
if $k = 1$ **then**
  $\mathrm{S}(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, \cdot, k)$
else
  $\mathrm{S}(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, \nu_1, k)$
  $\mathbf{r}^k = \mathbf{f}^k - \mathbf{A}^k \mathbf{x}^k$
  $\mathbf{f}^{k-1} = I_k^{k-1} \mathbf{r}^k$
  $\mathbf{x}^{k-1} = 0$
  for $i = 1, 2, ..., \gamma$ **do**
    $\mathrm{MG}(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, k-1)$
    $\tilde{\mathbf{x}}^{k-1} = \mathbf{x}^{k-1}$
  end for
  $\mathbf{x}^k = \mathbf{x}^k + I_{k-1}^k \mathbf{x}^{k-1}$
  $\mathrm{S}(\mathbf{x}, \mathbf{x}, \mathbf{f}, \nu_2, k)$
end if

In Algorithm 2.3, $S(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, \nu, k)$ is the smoothing step defined on each grid k. S changes an initial guess $\tilde{\mathbf{x}}^k$ into an improved approximation $\mathbf{x}^k$ with the right-hand side $\mathbf{f}^k$ by ν^k iterations with a suitable smoothing method. The number of pre/post-smoothing iterations is denoted by ν_1 and ν_2, respectively. The appropriate smoothers for the LSFEM systems will be discussed in the next sub-section. We denote the coarse-grid solution by $S(\tilde{\mathbf{x}}, \mathbf{x}, \mathbf{f}, \cdot, 1)$, where we omit the number of iterations because of the possible use of exact coarse-grid solvers. The multigrid schedule is determined by the number of coarse-grid visits in every multigrid iteration and denoted by γ in the Algorithm 2.3, see [75] for further information on different multigrid cycles.

Other components of every multigrid method are intergrid restriction and prolongation operators. We use the canonical grid transfer operators based on FEM space which treat all solution components separately. In the case of conforming Q_2 finite elements for example, the prolongation operator is constructed by using a biquadratic interpolation. The restriction operator is then set up as the adjoint of the prolongation operator, i.e., the matrices associated to prolongation and restriction are exactly transposed to each other [44].

2.5.3. Multigrid-preconditioned Conjugate Gradient solver

The rate of convergence of the CG method can be significantly improved with various preconditioning techniques [67]. Jacobi preconditioned CG solver has been widely used for the solution of the first-order system least-squares FEM, see [33]. On the other hand, the multigrid method has shown to be a rather effective preconditioner for the CG solver [34–36, 64]. For example, Ranjan and Reddy [64] has investigated a geometric multigrid preconditioned CG solver for the Spectral/hp LSFEM solution of the Navier-Stokes (NS) equations. They demonstrated that the convergence rate of the CG solver improved when replacing the Jacobi preconditioner with a multigrid solver. With due consideration of the success of the above solution techniques, we use CG as the main linear system solver and accelerate it with the multigrid preconditioning, i.e. multigrid-preconditioned conjugate gradient (MPCG) solver.

In this work, we use the geometric multigrid method, in contrast to the algebraic multigrid method which is widely used in the LSFEM community. As it is discussed in the previous section, in a geometric multigrid method a hierarchy of classically refined grids [43] are generated and the system of equations are discretized on all grids, see [27] for further details. Since the most important part of the multigrid method is the smoothing, the efficiency and robustness depend in many cases on the smoothing algorithms. Besides using the preconditioned CG method as stand-alone solver, the CG method can also be applied as smoother in the multigrid method to accelerate the convergence and to improve the robustness (see [22], [42]). We use CG as pre/post-smoother which appropriately determines the size of the solution updates at each smoothing step [76]. Therefore, the CG smoothing leads to efficient and particularly parameter-free smoothing sweeps. In addition, we accelerate the smoothing process by using a SSOR preconditioner, which in this context requires no damping parameter in case of symmetric Gauß-Seidel sweeps. It is worth noting that the possibility of using standard smoothers for the solution of the NS equations is another advantage of the LSFEM over the mixed Galerkin methods that require specially designed smoothers [71].

The MPCG solver is summarized as follows

1. Direct Gaussian elimination (UMFPACK [18]) as coarse-grid solver

2. SSOR-preconditioned CG smoother with 4 pre/post-smoothing steps

3. F-cycle as the multigrid cycle

4. Intergrid transfer and coarse grid correction based on the underlying mesh hierarchy and the properties of the chosen conforming finite elements [71].

3

Advection-Diffusion-Reaction Equations

In this chapter we focus on the application of the LSFEM to the general transport equations of heat transfer and fluid flow. Our aim in this research is then twofold. On the one hand, we analyze the performance of the LSFEM for the solution of the scalar elliptic PDEs and on the other hand, we obtain all the necessary requirements to extend our work to the mass and momentum conservation equations, i.e. Navier-Stokes equations. In other words, we design efficient and robust LSFEMs for the transport equations which include the main transport mechanisms of the Navier-Stokes equations. In addition, we investigate accurate multigrid solvers for the solution of the discretized systems arising from such equations.

3.1. The Governing Equations

Some of the most common examples of transport analysis in engineering are seen in the fields of process, chemical, and mechanical engineering, but the subject is a fundamental component of the curriculum in all disciplines involved in any way with fluid mechanics, heat transfer, and mass transfer [23]. It is now considered to be a part of the engineering discipline as much as thermodynamics [3], mechanics, and electromagnetism [32].

The generic scalar transport equation is a general partial differential equation that describes heat transfer, mass transfer, fluid dynamics (momentum transfer), etc. The conservative form of the scalar advection-diffusion-reaction equation is presented as follow

$$\begin{cases} \dfrac{\partial p}{\partial t} - \nabla \cdot (\varepsilon \nabla p - \beta p) + cp = f & \text{in } \Omega \\ p = g_D & \text{on } \Gamma_D \\ \mathbf{n} \cdot \nabla p = g_N & \text{on } \Gamma_N \\ p = p_0 & \text{in } \Omega, t = t_0 \end{cases} \tag{3.1.1}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain. The first term in equation (3.1.1) is the time variation of variable p, $-\varepsilon \nabla p$ is the diffusive flux and βp is the advective flux, ε is the diffusion coefficient, β is the advecting velocity field, $\mathbf{n}$ is the unit outward vector normal to the boundary, c is the reaction coefficient and f is the source term. The Dirichlet and Neumann part of the boundary are defined by Γ_D and Γ_N, and the g_D and g_N are the respective boundary conditions. The p_0 is the initial value specified at initial time $t = t_0$.

3.2. LSFEM for the Poisson Equation

We restrict the general advection-diffusion-reaction equation to simple diffusion (Poisson) equation. Therefore, in equation (3.1.1) we set $\varepsilon = 1$, $\beta = \mathbf{0}$, $c = 0$ and assume homogeneous Dirichlet

boundary conditions on the whole boundary Γ. We only consider the steady state diffusion equation (Poisson). The Poisson equation in a two-dimensional domain then reads

$$\begin{cases} -\Delta p = f & \text{in } \Omega \\ p = 0 & \text{on } \Gamma . \end{cases} \tag{3.2.1}$$

We further assume that f is a smooth function defined in Ω. With the help of this introductory example, we present the least-squares solution method which is described in chapter 2. Moreover, we discuss the finite element spaces and the iterative solvers which are used in this research.

The Poisson equation (3.2.1) is a second order elliptic problem. We discussed earlier in chapter 2 that one of the requirements of having a practical least-squares method is to work with first-order equations. In deed, the straightforward application of the LSFEM to the original Poisson equation (3.2.1) leads to a variational setting that requires C^1 finite elements [52]. Therefore, we transform the original Poisson equation into equivalent first order systems. A reasonable practice to obtain an equivalent first-order system is to introduce new variables based on the primary unknowns of the original PDE. These new variables may be some physically important quantities such as fluxes which may be of crucial practical interest.

3.2.1. First-order Poisson Equations

We introduce the definition of the fluxes into the Poisson equation (3.2.1) and obtain the following first-order systems:

grad−div system

grad−div−curl system.

We explain the properties of each system and use the LSFEM to solve the Poisson equation based on both first-order systems.

The grad−div *System*

The most straightforward strategy to reformulate the second order Poisson equation (3.2.1) to an equivalent first order system is to introduce the gradient of the p as a new variable. The new grad−div system then reads

$$\begin{cases} \nabla \cdot \mathbf{u} = f & \text{in } \Omega \\ \nabla p + \mathbf{u} = 0 & \text{in } \Omega , \\ p = 0 & \text{on } \Gamma \end{cases} \tag{3.2.2}$$

which is a first-order system of equations with p and $\mathbf{u}$ as unknowns.

The grad−div−curl *System*

The first order system of equations in (3.2.2) can also be augmented with additional curl constraint

$$\nabla \times \mathbf{u} = 0 \quad \text{in } \Omega \tag{3.2.3}$$

and additional boundary condition

$$\mathbf{n} \times \mathbf{u} = 0 \quad \text{on } \Gamma . \tag{3.2.4}$$

To see how these equations are obtained, we apply the curl operator $\nabla\times$ to the second equation in (3.2.2) and we obtain

$$\nabla \times (\nabla p + \mathbf{u}) = 0 . \tag{3.2.5}$$

In addition, for every sufficiently smooth scalar function p we have the following identity

$$\nabla \times \nabla p = 0 . \tag{3.2.6}$$

Using equations (3.2.5) and (3.2.6), the extra curl equation (3.2.3) is obtained. Moreover, since we have $p = 0$ on the boundary Γ then $\mathbf{n} \times \nabla p = 0$ which together with the second equation in (3.2.2) leads to the boundary condition (3.2.4). Finally, we summarize the following set of equations as the grad−div−curl system:

$$\begin{cases} \nabla \cdot \mathbf{u} = f & \text{in } \Omega \\ \nabla \times \mathbf{u} = 0 & \text{in } \Omega \\ \nabla p + \mathbf{u} = 0 & \text{in } \Omega \\ p = 0 & \text{on } \Gamma \\ \mathbf{n} \times \mathbf{u} = 0 & \text{on } \Gamma \end{cases} \tag{3.2.7}$$

Carey *et al.* [26] compared the effect of adding the curl equation (3.2.3) and the boundary condition(3.2.4) on the performance of an iterative method for the solution of non-self adjoint elliptic problems. They observed that adding these additional equations leads to optimal L^2 and H^1 error estimates for the flux variable $\mathbf{u}$. Similar conclusions have been made for the solution of the Poisson equation in the LSFEM literature, see [54] and [12] for further details. We will show that, by adding equations (3.2.3) and (3.2.4) one obtains optimal error estimates for both the scalar and the flux variables. These statements are supported by our numerical investigations in section 3.2.5 which are in accordance with the theoretical a priori error estimates too.

3.2.2. Continuous Least-squares Principle

We define the space of admissible functions for the grad−div and grad−div−curl systems based on the notation of the spaces introduced earlier in chapter 2. We define the following spaces for the flux variables $\mathbf{u}$

$$\mathbf{V}^a = H(\text{div}, \Omega) \tag{3.2.8}$$

$$\mathbf{V}^b = H(\text{div}, \Omega) \cap H_{0,D}(\text{curl}, \Omega)\,, \tag{3.2.9}$$

where $\mathbf{V}^a$ and $\mathbf{V}^b$ are the admissible spaces of the grad−div and grad−div−curl systems, respectively. In order to have a unified definition for the space of the flux variables we further set $\mathbf{V} := \mathbf{V}^{a/b}$. Next, we define the space W which represents the scalar variable p in both systems:

$$W = H^1_{0,D}(\Omega)\,. \tag{3.2.10}$$

Here, we impose the Dirichlet boundary conditions through the definition of the test spaces. Therefore, there is no need to include the boundary conditions in the least-squares energy functional.

We define the L^2-norm least-squares energy functionals $\mathcal{J}^{a/b} : \mathbf{V} \times W \to \mathbb{R}$ based on the residuals of the first-order grad−div system as follows

$$\mathcal{J}^a(\mathbf{v}, q; f) = \|\nabla \cdot \mathbf{v} - f\|_0^2 + \|\mathbf{v} + \nabla q\|_0^2 \quad \forall (v, q) \in \mathbf{V} \times W \tag{3.2.11}$$

and for the grad−div−curl system as follows

$$\mathcal{J}^b(\mathbf{v}, q; f) = \|\nabla \cdot \mathbf{v} - f\|_0^2 + \|\mathbf{v} + \nabla q\|_0^2 + \|\nabla \times \mathbf{v}\|_0^2 \quad \forall (v, q) \in \mathbf{V} \times W\,. \tag{3.2.12}$$

The least-squares minimization problem associated with the L^2-norm functionals in (3.2.11) and (3.2.12) is then: find $(\mathbf{u}, p) \in \mathbf{V} \times W$ such that

$$(\mathbf{u}, p) = \underset{(\mathbf{v}, q) \in \mathbf{V} \times W}{\text{argmin}}\ \mathcal{J}^{a/b}(\mathbf{v}, q; f)\,, \tag{3.2.13}$$

which leads to the variational problem: find $(\mathbf{u}, p) \in \mathbf{V} \times W$ such that

$$\mathcal{A}^{a/b}(\mathbf{u}, p; \mathbf{v}, q) = \mathcal{F}(\mathbf{v}, q) \quad \forall (\mathbf{v}, q) \in \mathbf{V} \times W\,, \tag{3.2.14}$$

where $\mathcal{A}^{a/b}$ is a bilinear form defined on $(\mathbf{V}\times W)\times(\mathbf{V}\times W)\to\mathbb{R}$ and $\mathcal{F}$ is a linear functional defined on $\mathbf{V}\times W\to\mathbb{R}$. The variational problem for the grad−div system is given by

$$\mathcal{A}^{a}(\mathbf{u},p;\mathbf{v},q) := (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_0+(\mathbf{u}+\nabla p,\mathbf{v}+\nabla q)_0\,, \tag{3.2.15}$$

and the variational problem related to the grad−div−curl system reads

$$\mathcal{A}^{b}(\mathbf{u},p;\mathbf{v},q) := (\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})_0+(\mathbf{u}+\nabla p,\mathbf{v}+\nabla q)_0+(\nabla\times\mathbf{u},\nabla\times\mathbf{v})_0\,. \tag{3.2.16}$$

The linear functional $\mathcal{F}$ is the same for both of the first-order systems and reads

$$\mathcal{F}(\mathbf{v},q) := (f,\nabla\cdot\mathbf{v})_0\,. \tag{3.2.17}$$

Existence and Uniqueness of the Solution

The solvability and uniqueness of the problem (3.2.14) depends on the coercivity and boundedness estimates of the bilinear form $\mathcal{A}^{a/b}$ and the boundedness of the linear form $\mathcal{F}$. Therefore, we recall the following important theorem from [54].

Theorem 3.2.1. *For all $(\mathbf{v},q)\in\mathbf{V}\times W$, there exists constants $C_1,C_2>0$ such that for the* grad−div *system*

$$C_1\left(\|q\|_1^2+\|\mathbf{v}\|_{\mathrm{div}}^2\right)\leq\mathcal{J}^{a}(\mathbf{v},q;0)\leq C_2\left(\|q\|_1^2+\|\mathbf{v}\|_{\mathrm{div}}^2\right) \tag{3.2.18}$$

and, for the grad−div−curl *system*

$$C_1\left(\|q\|_1^2+\|\mathbf{v}\|_1^2\right)\leq\mathcal{J}^{b}(\mathbf{v},q;0)\leq C_2\left(\|q\|_1^2+\|\mathbf{v}\|_1^2\right)\,. \tag{3.2.19}$$

The proof of this theorem can be found in [54] and the references therein. It should be noted that $\mathcal{J}^{a/b}(\mathbf{v},q;0)$ in these estimates is equivalent to $\mathcal{A}^{a/b}(\mathbf{v},q;\mathbf{v},q)$. Therefore, the above a priori estimates together with the Lax-Milgram lemma imply that the variational problem (3.2.14) has a unique solution $(\mathbf{u},p)\in\mathbf{V}\times W$ for the both first-order systems.

Operator form of the Problem

To analyze the properties of the least-squares problem, let us write the variational problem (3.2.14) as follows

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q)=\big(\mathsf{L}(\mathbf{u},p),\mathsf{L}(\mathbf{v},q)\big)\,, \tag{3.2.20}$$

where L is a linear operator. For the grad−div system the operator L is given by

$$\mathsf{L}=\begin{pmatrix}\nabla\cdot & 0\\ I & \nabla\end{pmatrix}\,, \tag{3.2.21}$$

and for the grad−div−curl system is given by

$$\mathsf{L}=\begin{pmatrix}\nabla\cdot & 0\\ I & \nabla\\ \nabla\times & 0\end{pmatrix}\,. \tag{3.2.22}$$

where I is the identity tensor. Restricting our test spaces to the $C_0^\infty(\Omega)$ function spaces, infinitely differentiable functions that have a compact support in Ω, we rewrite the bilinear form (3.2.20) as

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \big(\mathsf{L}^*\mathsf{L}(\mathbf{u},p),(\mathbf{v},q)\big) \tag{3.2.23}$$

where L^* is the formal adjoint of L. The least-squares operator $\mathsf{L}^*\mathsf{L}$ of the grad$-$div system is given by

$$\mathsf{L}^*\mathsf{L} \;=\; \begin{pmatrix} -\nabla & I \\ 0 & -\nabla\cdot \end{pmatrix} \begin{pmatrix} \nabla\cdot & 0 \\ I & \nabla \end{pmatrix} \;=\; \begin{pmatrix} -\nabla\nabla\cdot + I & \nabla \\ -\nabla\cdot & -\Delta \end{pmatrix}, \tag{3.2.24}$$

and the operator form of the grad$-$div$-$curl system is given by

$$\mathsf{L}^*\mathsf{L} \;=\; \begin{pmatrix} -\nabla & I & \nabla\times \\ 0 & -\nabla\cdot & 0 \end{pmatrix} \begin{pmatrix} \nabla\cdot & 0 \\ I & \nabla \\ \nabla\times & 0 \end{pmatrix} \;=\; \begin{pmatrix} -\Delta + I & \nabla \\ -\nabla\cdot & -\Delta \end{pmatrix}. \tag{3.2.25}$$

The first obvious observation about the first-order least-squares systems (3.2.24) and (3.2.25) is that both systems have symmetric and (positive) definite coefficient matrices. The use of Krylov subspace methods like conjugate gradient (CG) for the solution of the discretized systems is then justified. Moreover, the elements on the main diagonal of the system (3.2.25), corresponding to the grad$-$div$-$curl first-order system, are Laplace operators which are strongly elliptic and therefore facilitate the use of efficient and standard preconditioners like Jacobi, Gauß-Seidel or SOR (successive overrelaxation) to accelerate the iterative solution of the corresponding discrete system.

3.2.3. Discrete Least-squares Principle

We introduce the finite element counterparts of the continuous variational problems here. Let the bounded domain $\Omega \subset \mathbb{R}^2$ be partitioned by a *grid* $\mathcal{T}_h$ consisting of elements $K \in \mathcal{T}_h$ which are assumed to be open quadrilaterals such that

$$\Omega = \operatorname{int}\Big(\bigcup_{K\in\mathcal{T}_h} K\Big). \tag{3.2.26}$$

For an element $K \in \mathcal{T}_h$, we denote by h_K the diameter of the element K. The *mesh size* h of $\mathcal{T}_h$ is given by

$$h := \max_{K\in\mathcal{T}_h} h_K. \tag{3.2.27}$$

We introduce the approximation spaces $\mathbf{V}_h^a$, $\mathbf{V}_h^b$ and W_h, which consist of piecewise polynomials of order less or equal r ($r \geq 0$ is integer). Let $\mathbf{V}_h := \mathbf{V}_h^{a/b}$ and consider the approximation problem

find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ such that

$$\mathcal{A}_h^{a,b}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) = \mathcal{F}_h(\mathbf{v}_h,q_h) \quad \forall (\mathbf{v}_h,q_h) \in \mathbf{V}_h \times W_h \tag{3.2.28}$$

where $\mathcal{A}_h^{a/b}$ is a bilinear form defined on $(\mathbf{V}_h \times W_h) \times (\mathbf{V}_h \times W_h) \to \mathbb{R}$, and has the following form for the grad$-$div system

$$\mathcal{A}_h^a(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) := \sum_{K\in\mathcal{T}_h} (\mathbf{u}_h + \nabla p_h, \mathbf{v}_h + \nabla q_h)_{0,K} + \sum_{K\in\mathcal{T}_h} (\nabla\cdot\mathbf{u}_h, \nabla\cdot\mathbf{v}_h)_{0,K}, \tag{3.2.29}$$

and for the grad−div−curl system as follows

$$\begin{aligned}\mathcal{A}_h^b(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) := &\sum_{K\in\mathcal{T}_h}(\mathbf{u}_h+\nabla p_h,\mathbf{v}_h+\nabla q_h)_{0,K}+\sum_{K\in\mathcal{T}_h}(\nabla\cdot\mathbf{u}_h,\nabla\cdot\mathbf{v}_h)_{0,K}\\ &+\sum_{K\in\mathcal{T}_h}(\nabla\times\mathbf{u}_h,\nabla\times\mathbf{v}_h)_{0,K}\,.\end{aligned} \tag{3.2.30}$$

The linear functional $\mathcal{F}_h$ is defined on $\mathbf{V}_h\times W_h\to\mathbb{R}$ by

$$\mathcal{F}_h(\mathbf{v}_h,q_h) := \sum_{K\in\mathcal{T}_h}(f,\nabla\cdot\mathbf{v}_h)_{0,K}\,. \tag{3.2.31}$$

It should be noted that $(\cdot,\cdot)_{0,K}$ is the restriction of an L^2 inner product on an element K of the triangulation.

If $\mathbf{V}_h\subset\mathbf{V}$ and $W_h\subset W$ then $\mathcal{A}_h^{a/b}=\mathcal{A}^{a/b}$ and $\mathcal{F}_h=\mathcal{F}$. As a result of the coercivity estimates (3.2.18) and (3.2.19) the problem (3.2.28) has unique solution $(\mathbf{u}_h,p_h)\in\mathbf{V}_h\times W_h$.

Error Estimates in the Energy Norm

With the help of the coercivity estimates (3.2.18) and (3.2.19) we are able to derive the a priori error estimates based on the energy norms $H_0^1(\Omega)\times H(\mathrm{div},\Omega)$ for the grad−div system and $H_0^1(\Omega)\times H(\mathrm{div},\Omega)\cap H_{0,D}(\mathrm{curl},\Omega)$ for the grad−div−curl system.Therefore, we recall the following theorem from [54].

Theorem 3.2.2. *Assume that the Poisson problem has a sufficiently regular solution $(\mathbf{u},p)\in\mathbf{V}\times W\cap[H^{r+1}(\Omega)]^3$. Then, the approximate problem (3.2.28) has a unique solution $(\mathbf{u}_h,p_h)\in\mathbf{V}_h\times W_h$ and satisfies the following error estimates: there exists a constant C independent of h such that*

for the grad−div *system*

$$\|\mathbf{u}-\mathbf{u}_h\|_{\mathrm{div}}+\|p-p_h\|_1\le Ch^r(\|\mathbf{u}\|_{r+1}+\|p\|_{r+1}) \tag{3.2.32}$$

and for the grad−div−curl *system*

$$\|\mathbf{u}-\mathbf{u}_h\|_1+\|p-p_h\|_1\le Ch^r(\|\mathbf{u}\|_{r+1}+\|p\|_{r+1})\,. \tag{3.2.33}$$

Here, r is the order of the polynomial functions and we have assumed that both unknowns are approximated by equal-order polynomials. For the proof of the theorem 3.2.2 see [54] and the references therein.

Error Estimates in the L^2 norm

For equal-order continuous nodal finite element spaces, one can also prove optimal L^2-error convergence for the scalar variable p. However, showing optimal L^2-error convergence for the fluxes $\mathbf{u}$ has been only possible for the grad−div−curl first-order system. In fact, numerical evidence strongly suggests that nodal continuous flux approximations do not posses optimal L^2 accuracy. Although, optimal L^2-error rates for the fluxes can be achieved without the curl constraint, provided that one uses the div-conforming family of Brezzi-Douglas-Marini or Brezzi-Douglas-Duran-Fortin elements [54].

In order to use the duality argument, it is assumed in [54] that the solution of the problem

$$\begin{cases}-\Delta\phi=\eta & \text{in }\Omega\\ \phi=0 & \text{on }\Gamma\end{cases} \tag{3.2.34}$$

satisfies the regularity estimate

$$\|\phi\|_{s+2} \leq \|\eta\|_s \quad \text{for } s = 1,2 \quad \text{and} \quad \forall \eta \in H^s(\Omega). \tag{3.2.35}$$

The additional regularity is necessary since the L^2-error estimates are based on the duality argument [54].

The optimal L^2-error estimates for the scalar variable p can be obtained based on both first-order formulations [54]. However, as it is mentioned, the optimal L^2-error for the flux variable depends on the curl constraint [12], [54]. Therefore, the following error estimate holds only for the grad−div−curl system.

Theorem 3.2.3. *Assume that the regularity estimate (3.2.35) and the assumptions of the Theorem 3.2.2 hold. Then the approximate problem (3.2.28) has a unique solution* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ *and satisfies the following error estimate: there exists a constant C independent of h such that*

for the grad−div−curl *system*

$$\|\mathbf{u} - \mathbf{u}_h\|_0 + \|p - p_h\|_0 \leq Ch^{r+1}(\|\mathbf{u}\|_{r+1} + \|p\|_{r+1}). \tag{3.2.36}$$

Proof of this theorem can be found in [12]. In the next section, we will numerically study the error estimates presented in this section. We will clarify the differences between the two first-order formulations on issues related to the efficiency of the iterative solvers and the optimality of the a priori error estimates.

3.2.4. Finite Element Approximation

We use bilinear Q_1 and biquadratic Q_2 conforming finite elements for both scalar and vector variables. The definition of these quadrilateral elements is given below.

Bilinear Element: Q_1

The bilinear element Q_1 consists of four degrees of freedom, which coincide with the vertices of the quadrilateral as shown in the left picture in Fig. 3.1 (see [14]). To construct the local shape

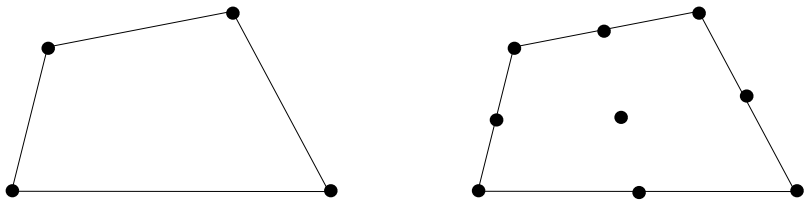


Figure 3.1: Conforming finite elements in physical coordinates; (left) bilinear Q_1 element; (right) biquadratic Q_2 element

functions on an arbitrary physical element Ω_k, we make use of a reference coordinate system. A local coordinate system (ξ, η) introduced. Let $\hat{\Omega}_k = [-1,1]^2$ be the reference element located at the center of this coordinate system (ξ, η). A one-to-one bilinear mapping between the physical and the reference element is referred to $F_k : \hat{\Omega}_k \to \Omega_k$. Once the basis functions have been defined

on the reference element in terms of reference coordinates, the inverse mapping $F_k^{-1} : \Omega_k \to \hat{\Omega}_k$ can be employed to get back to the physical space. Figure 3.1 illustrates the bilinear element in the physical coordinate system. In this case, there are four shape functions associated with the vertices. The nodal shape functions in the reference coordinate system are defined locally such that the basis function $\hat{\phi}_j$ is one at the node j and vanishes at all other nodes

$$\hat{\phi}_j(\hat{x}_i) = \delta_{ij} \tag{3.2.37}$$

where $\hat{x}_i \in \hat{\Omega}_k$ represent the ith node in $\hat{\Omega}_k$, i.e., vertices. In general, these four shape functions are bilinear polynomials, a linear combination of the following monomials

$$\{1, \xi, \eta, \xi\eta\}$$

where $-1 \leq \xi, \eta \leq 1$. Then, the space $Q_1(\Omega_k)$ on the physical element is defined as follows

$$Q_1(\Omega_k) = \left\{ q \circ F_k^{-1} : q \in \text{span}\{1, x, y, xy\} \right\}. \tag{3.2.38}$$

We refer the interested reader to [14] for more details.

Biquadratic Element: Q_2

On each quadrilateral, the biquadratic element Q_2 is defined by introducing the four additional mid-side node points, together with a ninth node at the center as shown in the right picture in figure 3.1 (see [14]). Figure 3.1 illustrates the biquadratic element in the physical coordinate system. In this case, there are nine shape functions, four associated with the vertices, four with the edge mid-points and one internal (or bubble) function. The nodal shape functions in the reference coordinate system possess the same property as described in equation 3.2.37. In general, these nine shape functions are biquadratic polynomials, a linear combination of the following monomials

$$\{1, \xi, \eta, \xi\eta, \xi^2, \eta^2, \xi^2\eta, \xi\eta^2, \xi^2\eta^2\}$$

where $-1 \leq \xi, \eta \leq 1$. Then, the space $Q_2(\Omega_k)$ on the physical element is defined as follows

$$Q_2(\Omega_k) = \left\{ q \circ F_k^{-1} : q \in \text{span}\{1, x, y, xy, x^2, y^2, x^2y, xy^2, x^2y^2\} \right\}. \tag{3.2.39}$$

We refer the interested reader to [14] for more details.

3.2.5. Numerical Results and Discussion

To investigate the LSFEMs for the solution of the Poisson equation (3.2.1), a unit square $\Omega = [0,1]^2$ is considered. Two different grid types (regular and perturbed) are used which are shown in Fig. 3.2. In the case of regular grids, the next levels are obtained by applying recursive uniform refinements starting from the coarsest grid which consists of just one element. The irregular grids however, are obtained after construction of regular grids of each level, by applying a perturbation filter based on the grid size h. Information regarding computational grids, Number of Elements (NE) and Degrees of Freedom (DoF), for both least-squares and standard finite element methods is summarized in Table (3.1).

Two different analytical functions are considered for the scalar variable p, and the convergence rates of both scalar and vector variables are investigated. The functions are defined as follows

$$p = 16xy(1-x)(1-y) \tag{3.2.40}$$

$$p = sin(\pi x)sin(\pi y). \tag{3.2.41}$$

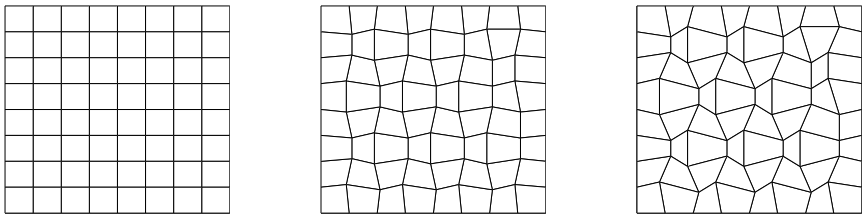


Figure 3.2: Sample regular grid (left), perturbed grid 0.1 (middle) and perturbed grid 0.2 (right)

Table 3.1: General mesh information, Poisson problem in a unit square

		Number of DoF			
		LSFEM		Standard	
Level	NE	Q_1	Q_2	Q_1	Q_2
6	1 024	3 267	12 675	1 089	4 225
7	4 096	12 675	49 923	4 225	16 641
8	16 384	49 923	198 147	16 641	66 049
9	16 384	49 923	198 147	16 641	66 049

We call these two functions, example one and example two, respectively. We extensively compare the results with those of the standard Galerkin FEM.

In order to validate the theoretical error estimate in Theorem 3.2.3, the L^2-errors along with the rate of convergence for the scalar variable p and the vector variable u are presented in Table 3.2 to 3.7. In Table 3.2 and Table 3.3 the grad−div system is solved using Q_1 elements for examples one and two, respectively. The same results are summarized for the grad−div−curl system in Table 3.4 and Table 3.5. The errors for the solution of example two based on the Q_2 element are also presented in Table 3.6 and Table 3.7 for both first-order systems. In all test cases the LSFEM results are compared with those of the Galerkin finite element method. For the case of regular grid, both the grad−div and the grad−div−curl systems perform well for the first example and optimal L^2-error convergence is achieved for all variables. However, for the second example, optimal L^2-error convergence for the vector variables is obtained just by using the grad−div−curl system. This observation is widely reported in the LSFEM literature [54]. The reason behind the optimal error in the first example is that the analytical function is inherently satisfying the curl constraint.

Increasing the mesh perturbation, all discretization errors are increased when compared to the results of the regular grids at the same level for both systems and examples. However, similar to the case of regular grids, optimal convergence rate for the vector variable is achieved after augmenting the first-order system with the curl constraint. All of the numerical results of Table 3.2 to Table 3.7 are obtained using a CG linear solver.

To investigate the effects of the additional equations (3.2.3) and (3.2.4) on the linear solver convergence, we compare the number of iterations required for the CG solver in Table 3.8. We choose example two and compare the results for the grad−div (gd-LSFEM) and the grad−div−curl (gdc-LSFEM) formulations. An additional case is also considered here by solving the set of equations in (3.2.7) without boundary condition (3.2.4). This case is denoted by LSFEM* in Table 3.8. We observe that the gd-LSFEM has the largest number of iterations for both Q_1 and Q_2 elements. The number of iterations considerably reduces for the gdc-LSFEM as compared to the gd-LSFEM.

Table 3.2: L^2-error using Q_1 element for the grad−div system, example one

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	2.435e-03	4.00	2.704e-03	4.00	2.370e-03	4.01	1.492e-01	2.00
6	6.086e-04	4.00	6.760e-04	4.00	5.921e-04	4.00	7.455e-02	2.00
7	1.521e-04	4.00	1.691e-04	4.00	1.480e-04	4.00	3.727e-02	2.00
8	3.804e-05	4.00	4.236e-05	4.00	3.700e-05	4.00	1.863e-02	2.00
perturbed grid 0.1								
5	2.810e-03	3.96	1.523e-02	1.88	2.673e-03	4.01	1.566e-01	2.00
6	7.207e-04	3.90	5.552e-03	2.74	6.874e-04	3.89	7.922e-02	1.98
7	1.800e-04	4.00	2.305e-03	2.41	1.716e-04	4.00	3.961e-02	2.00
8	4.530e-05	3.97	1.019e-03	2.26	4.320e-05	3.97	1.985e-02	2.00
perturbed grid 0.2								
5	3.869e-03	3.86	3.280e-02	1.65	3.551e-03	3.98	1.785e-01	1.97
6	1.042e-03	3.71	1.181e-02	2.78	9.634e-04	3.69	9.289e-02	1.92
7	2.601e-04	4.01	5.351e-03	2.21	2.399e-04	4.01	4.643e-02	2.00
8	6.613e-05	3.93	2.144e-03	2.50	6.103e-05	3.93	2.339e-02	1.98

However, the results of Table 3.8 show that the exclusion of the extra curl boundary condition deteriorates the convergence behavior of the CG solver. In addition, the number of iterations for the LSFEM solvers are grid-dependent and are larger than the number of iterations for the standard Galerkin solver.

In order to get grid-independent linear solver behavior, we use a multigrid solver as explained in the previous chapter, see 2.5.2. Performance of the multigrid solver is investigated for the grad−div−curl system. A CG smoother with a total of four pre-smoothing and post-smoothing steps is used at each level of the multigrid. A direct Gaussian elimination solver (UMFPACK [18]) is used for the coarse-grid solution. The number of iterations of the multigrid solver are presented in Table 3.9. The linear solver shows grid-independent convergence behavior. A second test case is also examined, in which we use a preconditioned-CG smoother for the multigrid. The number of smoothing steps are four for each level and the results are summarized in Table 3.10. A little improvement over the unpreconditioned CG smoother is observed for this solver setup.

Multigrid performance is further studied for the grad−div system in Table 3.11. In this case, a CG smoother with 4 pre/post-smoothing steps is used inside the multigrid. Performance of the multigrid solver is highly degraded and the number of iterations are increased depending on the grid level. This justifies the use of the additional curl constraints.

Table 3.3: L^2-error using Q_1 element for the grad−div system, example two

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	2.534e-03	3.78	9.208e-02	2.50	1.901e-03	4.00	1.259e-01	2.00
6	6.527e-04	3.88	3.360e-02	2.74	4.752e-04	4.00	6.295e-02	2.00
7	1.656e-04	3.94	1.198e-02	2.81	1.188e-04	4.00	3.148e-02	2.00
8	4.172e-05	3.97	4.243e-03	2.82	2.970e-05	4.00	1.574e-02	2.00
perturbed grid 0.1								
5	2.907e-03	3.69	8.702e-02	2.59	2.186e-03	3.90	1.329e-01	1.97
6	7.582e-04	3.83	3.084e-02	2.82	5.649e-04	3.87	6.734e-02	1.97
7	1.920e-04	3.95	1.044e-02	2.95	1.415e-04	3.99	3.372e-02	2.00
8	4.834e-05	3.97	3.287e-03	3.18	3.547e-05	3.99	1.688e-02	2.00
perturbed grid 0.2								
5	3.991e-03	3.54	7.535e-02	2.70	3.017e-03	3.77	1.537e-01	1.92
6	1.067e-03	3.74	2.530e-02	2.98	8.219e-04	3.67	8.015e-02	1.92
7	2.684e-04	3.98	8.568e-03	2.95	2.056e-04	4.00	4.017e-02	2.00
8	6.765e-05	3.97	2.588e-03	3.31	5.175e-05	3.97	2.015e-02	1.99

Table 3.4: L^2-error using Q_1 element for the grad−div−curl system, example one

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	3.141e-03	3.99	4.929e-03	4.29	2.370e-03	4.01	1.492e-01	2.00
6	7.859e-04	4.00	1.194e-03	4.13	5.921e-04	4.00	7.455e-02	2.00
7	1.967e-04	4.00	2.953e-04	4.04	1.480e-04	4.00	3.727e-02	2.00
8	4.949e-05	4.00	7.358e-05	4.01	3.700e-05	4.00	1.863e-02	2.00
perturbed grid 0.1								
5	3.588e-03	3.92	5.790e-03	4.07	2.673e-03	4.01	1.566e-01	2.00
6	9.180e-04	3.91	1.461e-03	3.96	6.874e-04	3.89	7.922e-02	1.98
7	2.298e-04	4.00	3.656e-04	4.00	1.716e-04	4.00	3.961e-02	2.00
8	5.770e-05	3.98	9.073e-05	4.03	4.320e-05	3.97	1.985e-02	2.00
perturbed grid 0.2								
5	4.955e-03	3.73	8.370e-03	3.63	3.551e-03	3.98	1.785e-01	1.97
6	1.323e-03	3.75	2.174e-03	3.85	9.634e-04	3.69	9.289e-02	1.92
7	3.317e-04	3.99	5.555e-04	3.91	2.399e-04	4.01	4.643e-02	2.00
8	8.379e-05	3.96	1.374e-04	4.04	6.103e-05	3.93	2.339e-02	1.98

Table 3.5: L^2-error using Q_1 element for the grad−div−curl system, example two

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	3.441e-03	3.98	6.187e-03	4.00	1.901e-03	4.00	1.259e-01	2.00
6	8.614e-04	4.00	1.547e-03	4.00	4.752e-04	4.00	6.295e-02	2.00
7	2.154e-04	4.00	3.868e-04	4.00	1.188e-04	4.00	3.148e-02	2.00
8	5.386e-05	4.00	9.670e-05	4.00	2.970e-05	4.00	1.574e-02	2.00
perturbed grid 0.1								
5	3.919e-03	3.87	7.287e-03	3.93	2.186e-03	3.90	1.329e-01	1.97
6	9.987e-04	3.92	1.838e-03	3.96	5.649e-04	3.87	6.734e-02	1.97
7	2.502e-04	3.99	4.602e-04	4.00	1.415e-04	3.99	3.372e-02	2.00
8	6.264e-05	3.99	1.151e-04	4.00	3.547e-05	3.99	1.688e-02	2.00
perturbed grid 0.2								
5	5.408e-03	3.64	1.044e-02	3.76	3.017e-03	3.77	1.537e-01	1.92
6	1.428e-03	3.79	2.666e-03	3.91	8.219e-04	3.67	8.015e-02	1.92
7	3.584e-04	3.98	6.694e-04	3.98	2.056e-04	4.00	4.017e-02	2.00
8	8.987e-05	3.99	1.675e-04	4.00	5.175e-05	3.97	2.015e-02	1.99

Table 3.6: L^2-error using Q_2 element for the grad−div system, example two

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	2.572e-05	7.96	2.068e-03	4.09	2.572e-05	7.97	3.191e-03	4.00
6	3.218e-06	7.99	5.135e-04	4.03	3.218e-06	7.99	7.979e-04	4.00
7	4.024e-07	8.00	1.282e-04	4.01	4.024e-07	8.00	1.995e-04	4.00
8	5.030e-08	8.00	3.208e-05	4.00	5.030e-08	8.00	4.987e-05	4.00
perturbed grid 0.1								
5	3.134e-05	7.89	2.257e-03	3.51	3.135e-05	7.89	3.747e-03	3.91
6	3.875e-06	8.09	6.121e-04	3.69	3.875e-06	8.09	9.315e-04	4.02
7	4.879e-07	7.94	1.330e-04	4.60	4.879e-07	7.94	2.334e-04	3.99
8	6.105e-08	7.99	3.133e-05	4.25	6.105e-08	7.99	5.840e-05	4.00
perturbed grid 0.2								
5	4.875e-05	7.66	2.813e-03	3.66	4.873e-05	7.65	5.304e-03	3.71
6	5.908e-06	8.25	6.663e-04	4.22	5.907e-06	8.25	1.310e-03	4.05
7	7.542e-07	7.83	1.364e-04	4.89	7.542e-07	7.83	3.301e-04	3.97
8	9.448e-08	7.98	2.907e-05	4.69	9.447e-08	7.98	8.261e-05	4.00

Table 3.7: L^2-error using Q_2 element for the grad−div−curl system, example two

	LSFEM				Standard FEM			
Level	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
regular grid								
5	2.572e-05	7.97	8.147e-05	8.10	2.572e-05	7.97	3.191e-03	4.00
6	3.218e-06	7.99	1.014e-05	8.04	3.218e-06	7.99	7.979e-04	4.00
7	4.024e-07	8.00	1.265e-06	8.01	4.024e-07	8.00	1.995e-04	4.00
8	5.030e-08	8.00	1.581e-07	8.00	5.030e-08	8.00	4.987e-05	4.00
perturbed grid 0.1								
5	3.140e-05	7.92	9.441e-05	8.04	3.135e-05	7.89	3.747e-03	3.91
6	3.877e-06	8.10	1.212e-05	7.79	3.875e-06	8.09	9.315e-04	4.02
7	4.880e-07	7.95	1.540e-06	7.87	4.879e-07	7.94	2.334e-04	3.99
8	6.105e-08	7.99	1.920e-07	8.02	6.105e-08	7.99	5.840e-05	4.00
perturbed grid 0.2								
5	4.897e-05	7.75	1.364e-04	8.00	4.873e-05	7.65	5.304e-03	3.71
6	5.916e-06	8.28	1.847e-05	7.38	5.907e-06	8.25	1.310e-03	4.05
7	7.545e-07	7.84	2.393e-06	7.72	7.542e-07	7.83	3.301e-04	3.97
8	9.448e-08	7.99	2.973e-07	8.05	9.447e-08	7.98	8.261e-05	4.00

Table 3.8: Number of iterations with CG linear solver, grad−div (gd-LSFEM), grad−div−curl (gdc-LSFEM) and grad−div−curl system without $\mathbf{n}\times\mathbf{u}=0$ on Γ (gdc-LSFEM*), example two

	gd-LSFEM		gdc-LSFEM		gdc-LSFEM*		Standard FEM	
Level	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2
regular grid								
5	108	512	28	78	43	198	2	3
6	252	1 034	54	139	98	549	2	3
7	484	2 266	109	200	250	1 457	2	3
8	982	3 902	175	414	621	3 407	2	3
perturbed grid 0.1								
5	201	1 075	50	113	134	493	28	84
6	409	2 158	78	214	303	1 170	48	121
7	733	4 379	152	425	550	1 758	83	251
8	1357	8 820	281	823	926	3 996	164	469
perturbed grid 0.2								
5	252	1 354	55	136	320	625	35	86
6	421	2 872	89	255	147	1 197	56	159
7	890	6 979	172	496	594	2 784	103	275
8	1 506	13 162	322	1 003	1 113	6 006	189	573

Table 3.9: Number of iterations of the multigrid solver, CG smoother with 4 pre/post-smoothing steps, grad−div−curl, example two

	W-cycle		F-cycle	
Level	Q_1	Q_2	Q_1	Q_2
5	4	5	4	5
6	4	5	4	5
7	5	5	5	5
8	5	5	4	5

Table 3.10: Number of iterations of the multigrid solver, preconditioned-CG smoother with 4 pre/post-smoothing steps, grad−div−curl, example two

	W-cycle		F-cycle	
Level	Q_1	Q_2	Q_1	Q_2
5	3	3	3	3
6	4	4	4	4
7	3	4	3	4
8	3	3	3	3

Table 3.11: Number of iterations of the multigrid solver, CG smoother with 4 pre/post-smoothing steps, grad−div, example two

	W-cycle		F-cycle	
Level	Q_1	Q_2	Q_1	Q_2
3	5	31	5	31
4	11	127	11	126
5	19	474	19	477
6	41	*	42	*
7	145	*	145	*
8	*	*	*	*

* The solver is still converging with a very low rate

3.3. LSFEM for the Diffusion-Reaction Equation

We investigate the numerical solution of the diffusion-reaction equation in this section. Therefore, we set $\varepsilon = 1$, $\beta = \mathbf{0}$, $c = k^2$ in equation (3.1.1). We assume homogeneous Dirichlet boundary conditions on the whole boundary Γ. The diffusion-reaction equation reads

$$\begin{cases} -\Delta p + k^2 p = f & \text{in } \Omega \\ p = 0 & \text{on } \Gamma \end{cases} \tag{3.3.1}$$

where k is a real constant and $\Omega \subset \mathbb{R}^2$ is a bounded domain with boundary Γ. Also we assume that f is a smooth function defined in Ω. This special problem is considered to investigate the Least-squares finite element method for the solution of the steady state diffusion-reaction equation. In addition, equation (3.3.1) can be viewed as a semi-discretized time dependent Poisson problem. In this case, the coefficient k^2 serves as the inverse of the time step which normally appears in a temporal discretization of the transient Poisson equation.

3.3.1. First-order Systems

Since the diffusion-reaction equation (3.3.1) is second order, we use the following techniques to obtain the equivalent first-order systems.

The grad−div *System*

The most straightforward strategy to reformulate the second order equation (3.3.1) to an equivalent first order system is to introduce the gradient of the p as a new variable. The new grad−div system is as follows

$$\begin{cases} \nabla \cdot \mathbf{u} + k^2 p = f & \text{in } \Omega \\ \mathbf{u} + \nabla p = 0 & \text{in } \Omega \\ p = 0 & \text{on } \Gamma. \end{cases} \tag{3.3.2}$$

The grad−div−curl *System*

In a similar manner to the Poisson equation, the first order system of equations (3.3.2) representing the diffusion-reaction equation can be augmented with the curl equation (3.2.3) and the additional boundary condition (3.2.4). So, we use the following set of equations as the grad−div−curl system

$$\begin{cases} \nabla \cdot \mathbf{u} + k^2 p = f & \text{in } \Omega \\ \nabla \times \mathbf{u} = 0 & \text{in } \Omega \\ \mathbf{u} + \nabla p = 0 & \text{in } \Omega \\ p = 0 & \text{on } \Gamma \\ \mathbf{n} \times \mathbf{u} = 0 & \text{on } \Gamma. \end{cases} \tag{3.3.3}$$

3.3.2. Continuous Least-squares Principle

We use the previous definitions, as defined in equations (3.2.8), (3.2.9) and (3.2.10), for the space of admissible functions $\mathbf{V}$ and W. We define the continuous least-squares energy functionals $\mathcal{J}^{a/b} : \mathbf{V} \times W \to \mathbb{R}$ based on the L^2-norm of the residuals of the extended first order systems. The energy functionals are defined as follows for the grad−div system (3.3.2)

$$\mathcal{J}^a(\mathbf{v}, q; f) = \left\| \alpha(\nabla \cdot \mathbf{v} + k^2 q - f) \right\|_0^2 + \| \mathbf{v} + \nabla q \|_0^2 \quad \forall (v, q) \in \mathbf{V} \times W \tag{3.3.4}$$

and for the grad−div−curl system (3.3.3)

$$\mathcal{J}^b(\mathbf{v},q;f) = \left\|\alpha(\nabla\cdot\mathbf{v}+k^2q-f)\right\|_0^2 + \|\mathbf{v}+\nabla q\|_0^2 + \|\alpha(\nabla\times\mathbf{v})\|_0^2 \quad \forall (v,q)\in\mathbf{V}\times W\,. \tag{3.3.5}$$

In the above energy functionals, we appropriately scale the equations with $\alpha = 1/k$. This will give us better energy balances and the resulting least-squares methods exhibit an interesting splitting property as explained by Bochev and Gunzburger [52].

The least-squares problem for (3.3.1) in the vector-operator setting (3.3.4, 3.3.5 is the minimization problem: find $(\mathbf{u},p)\in\mathbf{V}\times W$ such that

$$(\mathbf{u},p) = \underset{(\mathbf{v},q)\in\mathbf{V}\times W}{\operatorname{argmin}}\ \mathcal{J}^{a/b}(\mathbf{v},q;f,g)\,. \tag{3.3.6}$$

The corresponding variational problem is to find $(\mathbf{u},p)\in\mathbf{V}\times W$ such that

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \mathcal{F}(\mathbf{v},q) \quad \forall(\mathbf{v},q)\in\mathbf{V}\times W \tag{3.3.7}$$

where the bilinear form for the grad−div system is given by

$$\mathcal{A}^a(\mathbf{u},p;\mathbf{v},q) := \frac{1}{k^2}(\nabla\cdot\mathbf{u}+k^2p,\nabla\cdot\mathbf{v}+k^2q)_0 + (\mathbf{u}+\nabla p,\mathbf{v}+\nabla q)_0 \tag{3.3.8}$$

and the bilinear form for the grad−div−curl system reads

$$\mathcal{A}^b(\mathbf{u},p;\mathbf{v},q) := \frac{1}{k^2}(\nabla\cdot\mathbf{u}+k^2p,\nabla\cdot\mathbf{v}+k^2q)_0 + (\mathbf{u}+\nabla p,\mathbf{v}+\nabla q)_0 + \frac{1}{k^2}(\nabla\times\mathbf{u},\nabla\times\mathbf{v})_0\,. \tag{3.3.9}$$

The linear form of the problem for both first-order systems reads

$$\mathcal{F}(\mathbf{v},q) := \frac{1}{k^2}(\nabla\cdot\mathbf{u}+k^2p,f)_0. \tag{3.3.10}$$

Operator form of the Problem

Let us write the variational problem (3.3.7) as follows

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \big(\mathsf{L}(\mathbf{u},p),\mathsf{L}(\mathbf{v},q)\big)\,, \tag{3.3.11}$$

where L is the linear operator. Restricting our test spaces to the $C_0^\infty(\Omega)$ function spaces, infinitely differentiable functions that have a compact support in Ω, we rewrite the bilinear form (3.3.11) as

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \big(\mathsf{L}^*\mathsf{L}(\mathbf{u},p),(\mathbf{v},q)\big) \tag{3.3.12}$$

where L^* is the formal adjoint of L. The least-squares operator of the grad−div system is given by

$$\mathsf{L}^*\mathsf{L} = \begin{pmatrix} -\frac{1}{k^2}\nabla\nabla\cdot+\mathrm{I} & \mathbf{0} \\ \mathbf{0} & -\Delta+k^2 \end{pmatrix} \tag{3.3.13}$$

and for the the grad−div−curl system is given by

$$\mathsf{L}^*\mathsf{L} = \begin{pmatrix} -\frac{1}{k^2}\Delta+\mathrm{I} & \mathbf{0} \\ \mathbf{0} & -\Delta+k^2 \end{pmatrix} \tag{3.3.14}$$

where I is the identity matrix, ∇, $\nabla\cdot$ and Δ are the nabla, divergence and Laplace operators, respectively. This is a remarkable "splitting" property of the LSFEM that provides exactly the same solution for the scalar variable as that of standard Galerkin finite element method [52]. In addition, the grad−div−curl system has strong elliptic Laplace operators on the diagonals of the system matrix. This leads to better convergence properties of the iterative solvers for the grad−div−curl system as compared to the grad−div system.

3.3.3. Discrete Least-Squares Principle

The finite element approximation of the problem (3.3.7) is presented here. We introduce the approximation spaces $\mathbf{V}_h^a$, $\mathbf{V}_h^b$ and W_h. Let $\mathbf{V}_h \times W_h := \mathbf{V}_h^{a/b} \times W_h$ and consider the approximation problem

find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ such that

$$\mathcal{A}_h^{a,b}(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) = \mathcal{F}_h(\mathbf{v}_h, q_h) \quad \forall (\mathbf{v}_h, q_h) \in \mathbf{V}_h \times W_h. \tag{3.3.15}$$

A Priori Error Estimates

Let $\mathbf{V}_h \subset \mathbf{V}$ and $W_h \subset W$ be conforming approximation spaces which consist of piecewise polynomials of order less or equal r ($r \geq 0$ is integer), then $\mathcal{A}_h^{a/b} = \mathcal{A}^{a/b}$ and $\mathcal{F}_h = \mathcal{F}$. As a result of the coercivity and the boundedness of the bilinear form, the problem (3.3.7) has unique solution $(\mathbf{u}, p) \in \mathbf{V} \times W$ and $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$.

The same error estimates of the Poisson equation, Theorems 3.2.2 and 3.2.3, hold true for the diffusion-reaction equation [54].

3.3.4. Numerical Results and Discussion

We consider a unit square $\Omega = [0,1]^2$ as the computational domain in order to investigate the solution of the diffusion-reaction equation. We use conforming elements of Q_1 and Q_2 type for both the scalar and the vector variables. The same computational grids of the Poisson equation will be used in our simulation. Since we have observed the poor performance of the grad$-$div system in the Poisson problem, we only investigate the grad$-$div$-$curl system.

The following analytical functions is assumed for the scalar variable p,

$$p = sin(\pi x) sin(\pi y) \tag{3.3.16}$$

which leads to the following right hand side function f

$$f = (2\pi^2 + k^2) sin(\pi x) sin(\pi y). \tag{3.3.17}$$

Discretization errors along with the rate of convergence for the scalar variable p and the vector variable $\mathbf{u}$ are calculated based on the L^2-norm. The results for different values of k^2 are presented in Table 3.12 and Table 3.13 for Q_1 and Q_2 elements, respectively. The LSFEM results are compared with those of the standard Galerkin FEM. The results match very well with the a priori error estimates. The important remark is that the error reductions of both the scalar and the vector variables are optimal for the LSFEM. The optimality of the vector variable approximation is not achieved with the standard Galerkin FEM.

Next, the performance of the multigrid solver is investigated for this problem. A preconditioned-CG smoother is used for pre/post-smoothing inside the multigrid. The UMFPACK solver is used for the coarse grid solution. We summarize the number of linear solver iterations for both element types for a wide range values of parameter k and for different grid levels in Table 3.14. We observe a perfect grid-independent solver behavior for Q_1 and Q_2 elements. Moreover, we observe that the number of iterations is smaller for larger values of k^2.

Table 3.12: L^2-error using Q_1 element, the LSFEM based on grad−div−curl system and the standard finite element

	LSFEM				Standard			
Level	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate
k^2 =1e-4								
5	1.901e-03	4.00	5.971e-03	4.00	1.901e-03	4.00	1.259e-01	2.00
6	4.752e-04	4.00	1.493e-03	4.00	4.752e-04	4.00	6.295e-02	2.00
7	1.188e-04	4.00	3.732e-04	4.00	1.188e-04	4.00	3.148e-02	2.00
8	2.970e-05	4.00	9.330e-05	4.00	2.970e-05	4.00	1.574e-02	2.00
$k^2 = 0.0$ (Poisson equation)								
5	3.441e-03	3.98	6.187e-03	4.00	1.901e-03	4.00	1.259e-01	2.00
6	8.614e-04	4.00	1.547e-03	4.00	4.752e-04	4.00	6.295e-02	2.00
7	2.154e-04	4.00	3.868e-04	4.00	1.188e-04	4.00	3.148e-02	2.00
8	5.386e-05	4.00	9.670e-05	4.00	2.970e-05	4.00	1.574e-02	2.00
k^2 =1.0								
5	1.836e-03	4.00	5.768e-03	4.00	1.836e-03	4.00	1.259e-01	2.00
6	4.589e-04	4.00	1.442e-03	4.00	4.589e-04	4.00	6.295e-02	2.00
7	1.147e-04	4.00	3.604e-04	4.00	1.147e-04	4.00	3.148e-02	2.00
8	2.868e-05	4.00	9.011e-05	4.00	2.868e-05	4.00	1.574e-02	2.00
k^2 =1e+4								
5	1.021e-03	4.06	3.207e-03	4.05	1.021e-03	4.06	1.261e-01	2.01
6	2.543e-04	4.01	7.989e-04	4.01	2.543e-04	4.01	6.298e-02	2.00
7	6.352e-05	4.00	1.995e-04	4.00	6.352e-05	4.00	3.148e-02	2.00
8	1.588e-05	4.00	4.987e-05	4.00	1.588e-05	4.00	1.574e-02	2.00

Table 3.13: L^2-error using Q_2 element, the LSFEM based on grad−div−curl system and the standard finite element

	LSFEM				Standard			
Level	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate
k^2 =1e-4								
5	2.572e-05	7.97	8.148e-05	8.11	2.572e-05	7.97	3.191e-03	4.00
6	3.218e-06	7.99	1.014e-05	8.04	3.218e-06	7.99	7.979e-04	4.00
7	4.024e-07	8.00	1.265e-06	8.01	4.024e-07	8.00	1.995e-04	4.00
8	5.030e-08	8.00	1.581e-07	8.00	5.030e-08	8.00	4.987e-05	4.00
$k^2 = 0.0$ (Poisson equation)								
5	2.572e-05	7.97	8.147e-05	8.10	2.572e-05	7.97	3.191e-03	4.00
6	3.218e-06	7.99	1.014e-05	8.04	3.218e-06	7.99	7.979e-04	4.00
7	4.024e-07	8.00	1.265e-06	8.01	4.024e-07	8.00	1.995e-04	4.00
8	5.030e-08	8.00	1.581e-07	8.00	5.030e-08	8.00	4.987e-05	4.00
k^2 =1.0								
5	2.572e-05	7.96	8.146e-05	8.10	2.572e-05	7.96	3.191e-03	4.00
6	3.218e-06	7.99	1.014e-05	8.04	3.218e-06	7.99	7.979e-04	4.00
7	4.024e-07	8.00	1.265e-06	8.01	4.024e-07	8.00	1.995e-04	4.00
8	5.030e-08	8.00	1.581e-07	8.00	5.030e-08	8.00	4.987e-05	4.00
k^2 =1e+4								
5	2.555e-05	7.81	8.018e-05	7.79	2.555e-05	7.81	3.209e-03	4.06
6	3.213e-06	7.95	1.009e-05	7.95	3.213e-06	7.95	7.991e-04	4.02
7	4.022e-07	7.99	1.263e-06	7.99	4.022e-07	7.99	1.996e-04	4.00
8	5.029e-08	8.00	1.580e-07	8.00	5.029e-08	8.00	4.988e-05	4.00

Table 3.14: Number of iterations with multigrid solver, preconditioned-CG as smoother with 4 pre/post smoothing steps, the grad−div−curl system

	k^2 =1e-4		k^2 =1		k^2 =1e+4	
Level	Q_1	Q_2	Q_1	Q_2	Q_1	Q_2
5	2	3	2	3	1	1
6	3	3	3	3	1	1
7	3	3	3	3	1	1
8	3	3	3	3	1	1

3.4. LSFEM for the Advection-Diffusion-Reaction Equation

There is a vast amount of literature devoted to the application of the LSFEM for the solution of advection-diffusion-reaction problems. We consider the steady state form of the advection-diffusion-reaction equation 3.1.1 defined as

$$\begin{cases} -\nabla\cdot(\varepsilon\nabla p-\beta p)+cp=f & \text{in } \Omega \\ p=g_D & \text{on } \Gamma_D \\ \mathbf{n}\cdot\nabla p=g_N & \text{on } \Gamma_N\,. \end{cases} \tag{3.4.1}$$

Different first-order formulations as well as different stabilization techniques have been introduced and analyzed generally in the book by Bochev and Gunzburger [52] and specifically by other investigators in [1, 2, 15, 26, 38, 60, 65, 78, 79].

The first step of every least-squares finite element method is the reformulation of the second-order equation into a set of first-order equations. Two different formulations are generally used in the LSFEM literature for the solution of advection-diffusion-reaction equations. The first one is based on the introduction of the diffusive fluxes as new variables and the second reformulation takes the total fluxes as new variables.

Carey et al. in [1, 2, 26] and Cai et al. in [78] studied least-squares mixed finite element methods for the advection-diffusion-reaction problem based on the diffusive flux formulation. In [1, 2, 26, 78], the ellipticity and the optimal error estimates in $H(\mathrm{div},\Omega)\times H^1(\Omega)$ have been proved, while no optimal L^2-norm error estimate was provided.

Cai et al. [79] modified the least-squares energy functionals by adding a compatible curl constraint and imposing additional boundary conditions to the original n-dimensional problem. The resulting functional is proved to be elliptic in $(H^1)^{n+1}$ which implies optimal error estimates in standard finite element subspaces of $(H^1)^{n+1}$. Followed by the work of Cai et al. in [78, 79], Fiard et al. introduced an exponentially weighted LSFEM for the solution of advection-dominated flow problems [38].

Using a different formulation, Yang [15] introduced total fluxes as new variables to recast the second-order advection-diffusion equation into a set of first-order equations. An optimal L^2 error estimate is developed by Yang [15] for both the scalar and the vector variables provided that the finite element spaces of the fluxes are one of the classical mixed elements, such as Raviart-Thomas elements [51], and the L^2-norm functionals are properly scaled.

Lazarov et al. [65] introduced the streamline diffusion LSFEM to obtain more stable solutions to the advection-dominated flow problems. Hsieh and Yang [60] studied different stabilization techniques, such as Galerkin least-squares stabilization and streamline diffusion stabilization [65], combined with the LSFEM based on the diffusive flux formulation. Optimal error estimates have been proved in $H(\mathrm{div},\Omega)\times H^1(\Omega)$ and different interior and boundary layer advection-dominated flow problems have been investigated [60].

We introduce both first-order systems based on the diffusive and the total fluxes. The additional curl equation and boundary conditions are added to the first-order equations for both formulations. To the best knowledge of the author, this is the first time to add the curl equation to the total flux formulation. The analysis of this new first-order set of equations is one of the new contributions of this work. We provide the coercivity and the a priori error estimates for the new formulation.

Moreover, we study the weighted norm (scaled) LSFEM. Finally, the stabilization techniques of least-squares type as in Hsieh and Yang [60] are studied.

3.4.1. First-order Systems

In this problem, we introduce two different techniques to reformulate the second-order PDE into a set of first-order equations.

Diffusive Flux Formulation

We define the new variables, the diffusive fluxes, as follows

$$\mathbf{u} = -\varepsilon \nabla p. \tag{3.4.2}$$

Introducing equation (3.4.2) to the general advection-diffusion-reaction equation (3.4.1) leads to the following first-order system of equations

$$\begin{cases} \nabla \cdot \mathbf{u} + \nabla \cdot (\beta p) + cp = f & \text{in } \Omega \\ \mathbf{u} + \varepsilon \nabla p = \mathbf{0} & \text{in } \Omega \\ p = g_D & \text{on } \Gamma_D \\ \mathbf{u} \cdot \mathbf{n} = \varepsilon g_N & \text{on } \Gamma_N. \end{cases} \tag{3.4.3}$$

Equation (3.4.3) can be further augmented with the following equations

$$\begin{cases} \nabla \times \mathbf{u} = 0 & \text{in } \Omega \\ \mathbf{u} \times \mathbf{n} = \varepsilon \dfrac{\partial g_D}{\partial \tau} & \text{on } \Gamma_D \end{cases} \tag{3.4.4}$$

where $\frac{\partial}{\partial \tau}$ is the tangential derivative on Γ_D.

Total Flux Formulation

In this case the total fluxes, the sum of the diffusive and the advective fluxes, are considered as new variables. The new variables read

$$\mathbf{u} = -\varepsilon \nabla p + \beta p. \tag{3.4.5}$$

Using equation (3.4.5) and equation (3.4.1) one can easily obtain the following system of equations

$$\begin{cases} \nabla \cdot \mathbf{u} + cp = f & \text{in } \Omega \\ \mathbf{u} + \varepsilon \nabla p - \beta p = \mathbf{0} & \text{in } \Omega \\ p = g_D & \text{on } \Gamma_D \\ \mathbf{u} \cdot \mathbf{n} = \varepsilon g_N & \text{on } \Gamma_N. \end{cases} \tag{3.4.6}$$

Equation (3.4.6) can be further augmented with the following equations

$$\begin{cases} \nabla \times (\mathbf{u} - \beta p) = 0 & \text{in } \Omega \\ \mathbf{u} \times \mathbf{n} = \varepsilon \dfrac{\partial g_D}{\partial \tau} & \text{on } \Gamma_D. \end{cases} \tag{3.4.7}$$

3.4.2. Continuous Least-squares Principles

Let us proceed with homogeneous Dirichlet problem where $\Gamma_D \neq \emptyset$. We define the space of admissible functions for the grad−div and the grad−div−curl systems. First, we define the following spaces for the flux variables $\mathbf{u}$

$$\mathbf{V}^a = H_{0,N}(\text{div}, \Omega) \tag{3.4.8}$$

$$\mathbf{V}^b = H_{0,N}(\text{div}, \Omega) \cap H_{0,D}(\text{curl}, \Omega), \tag{3.4.9}$$

where $\mathbf{V}^a$ and $\mathbf{V}^b$ are the test spaces of the grad−div and the grad−div−curl systems, respectively. We define the space W which represents the scalar variable p in both systems:

$$W = H^1_{0,D}(\Omega). \tag{3.4.10}$$

Let $\mathbf{V} := \mathbf{V}_0^{a/b}$ for consistency. Define the L^2-norm least-squares energy functional $\mathcal{J}^{a/b} : \mathbf{V} \times W \to \mathbb{R}$ for the extended first-order system (3.4.3), (3.4.4), based on the diffusive flux, as follows

$$\mathcal{J}^a(\mathbf{v},q;f,g) = \|\nabla\cdot\mathbf{v} + \nabla\cdot(\beta q) + cq - f\|_0^2 + \|\mathbf{v} + \varepsilon\nabla q\|_0^2 \quad \forall (\mathbf{v},q) \in \mathbf{V}\times W \tag{3.4.11}$$

$$\mathcal{J}^b(\mathbf{v},q;f,g) = \|\nabla\cdot\mathbf{v} + \nabla\cdot(\beta q) + cq - f\|_0^2 + \|\mathbf{v} + \varepsilon\nabla q\|_0^2 + \|\nabla\times\mathbf{v}\|_0^2 \quad \forall (\mathbf{v},q) \in \mathbf{V}\times W \tag{3.4.12}$$

and for the first-order system (3.4.6), (3.4.7), based on the total flux, as follows

$$\mathcal{J}^a(\mathbf{v},q;f,g) = \|\nabla\cdot\mathbf{v} + cq - f\|_0^2 + \|\mathbf{v} + \varepsilon\nabla q - \beta q\|_0^2 \quad \forall (\mathbf{v},q) \in \mathbf{V}\times W \tag{3.4.13}$$

$$\mathcal{J}^b(\mathbf{v},q;f,g) = \|\nabla\cdot\mathbf{v} + cq - f\|_0^2 + \|\mathbf{v} + \varepsilon\nabla q - \beta q\|_0^2 + \|\nabla\times(\mathbf{v} - \beta q)\|_0^2 \quad \forall (\mathbf{v},q) \in \mathbf{V}\times W \tag{3.4.14}$$

The least square problem for (3.4.1) in the vector-operator setting (3.4.11, 3.4.12, 3.4.13, 3.4.14) is the minimization problem: find $(\mathbf{u},p) \in \mathbf{V}\times W$ such that

$$(\mathbf{u},p) = \underset{(\mathbf{v},q)\in\mathbf{V}\times W}{\operatorname{argmin}} \mathcal{J}^{a/b}(\mathbf{v},q;f,g) \tag{3.4.15}$$

The variational problem associated with the minimization problem (3.4.15) is to find $(\mathbf{u},p) \in \mathbf{V}\times W$ such that

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \mathcal{F}(\mathbf{v},q) \quad \forall (\mathbf{v},q) \in \mathbf{V}\times W \tag{3.4.16}$$

where $\mathcal{A}^{a/b}$ is a bilinear form defined on $(\mathbf{V}\times W)\times(\mathbf{V}\times W) \to \mathbb{R}$, and has the following form for the diffusive flux formulation

$$\begin{aligned}\mathcal{A}^a(\mathbf{u},p;\mathbf{v},q) := {}& (\nabla\cdot\mathbf{u} + \nabla\cdot(\beta p) + cp, \nabla\cdot\mathbf{v} + \nabla\cdot(\beta q) + cq)_0 \\ & + (\mathbf{u} + \varepsilon\nabla p, \mathbf{v} + \varepsilon\nabla q)_0\end{aligned} \tag{3.4.17}$$

$$\begin{aligned}\mathcal{A}^b(\mathbf{u},p;\mathbf{v},q) := {}& (\nabla\cdot\mathbf{u} + \nabla\cdot(\beta p) + cp, \nabla\cdot\mathbf{v} + \nabla\cdot(\beta q) + cq)_0 \\ & + (\nabla\times\mathbf{u}, \nabla\times\mathbf{v})_0 + (\mathbf{u} + \varepsilon\nabla p, \mathbf{v} + \varepsilon\nabla q)_0\end{aligned} \tag{3.4.18}$$

and $\mathcal{F}$ is a linear form defined, on $\mathbf{V}\times W \to \mathbb{R}$, by

$$\mathcal{F}(\mathbf{v},q) := (f, \nabla\cdot\mathbf{v} + \nabla\cdot(\beta q) + cq)_0 \tag{3.4.19}$$

The bilinear form for the total flux formulation defined, on $(\mathbf{V}\times W)\times(\mathbf{V}\times W) \to \mathbb{R}$, by

$$\begin{aligned}\mathcal{A}^a(\mathbf{u},p;\mathbf{v},q) := {}& (\mathbf{u} + \varepsilon\nabla p - \beta p, \mathbf{v} + \varepsilon\nabla q - \beta q)_0 \\ & + (\nabla\cdot\mathbf{u} + cp, \nabla\cdot\mathbf{v} + cq)_0\end{aligned} \tag{3.4.20}$$

$$\begin{aligned}\mathcal{A}^b(\mathbf{u},p;\mathbf{v},q) := {}& (\mathbf{u} + \varepsilon\nabla p - \beta p, \mathbf{v} + \varepsilon\nabla q - \beta q)_0 \\ & + (\nabla\cdot\mathbf{u} + cp, \nabla\cdot\mathbf{v} + cq)_0 \\ & + (\nabla\times(\mathbf{u} - \beta p), \nabla\times(\mathbf{v} - \beta q))_0\end{aligned} \tag{3.4.21}$$

and $\mathcal{F}$ is defined, on $\mathbf{V}\times W \to \mathbb{R}$, by

$$\mathcal{F}(\mathbf{v},q) := (f, \nabla\cdot\mathbf{v} + cq)_0. \tag{3.4.22}$$

Diffusive Flux in Operator form

Let us write $\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \left(\mathsf{L}^{a/b}(\mathbf{u},p),\mathsf{L}^{a/b}(\mathbf{v},q)\right)_0$ where $\mathsf{L}^{a/b}$ is the block form operator given by

$$\mathsf{L}^a = \begin{pmatrix} \mathsf{I} & \varepsilon\nabla \\ \nabla\cdot & \nabla\cdot\beta + c \end{pmatrix} \tag{3.4.23}$$

and

$$\mathsf{L}^b = \begin{pmatrix} \mathsf{I} & \varepsilon\nabla \\ \nabla\cdot & \nabla\cdot\beta + c \\ \nabla\times & 0 \end{pmatrix}. \tag{3.4.24}$$

Restricting to the $C_0^\infty(\Omega)$ functions, we can rewrite the bilinear form $\mathcal{A}^{a/b}$ as

$$\mathcal{A}^{a/b}(\mathbf{u},p;\mathbf{v},q) = \left(\mathsf{L}^{a/b*}\mathsf{L}^{a/b}(\mathbf{u},p),(\mathbf{v},q)\right)_0 \tag{3.4.25}$$

where $\mathsf{L}^{a/b*}$ is the formal adjoint of $\mathsf{L}^{a/b}$. The least-squares operator $\mathsf{L}^{a*}\mathsf{L}^a$ for the diffusive flux formulation is given by

$$\begin{aligned} \mathsf{L}^{a*}\mathsf{L}^a &= \begin{pmatrix} \mathsf{I} & -\nabla \\ -\varepsilon\nabla\cdot & -\beta\cdot\nabla + c \end{pmatrix}\begin{pmatrix} \mathsf{I} & \varepsilon\nabla \\ \nabla\cdot & \nabla\cdot\beta + c \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{I} - \nabla\nabla\cdot & \nabla[(\varepsilon - c)\mathsf{I} - \nabla\cdot\beta] \\ -[(\varepsilon - c)\mathsf{I} + (\beta\cdot\nabla)]\nabla\cdot & -\varepsilon^2\triangle + c(\nabla\cdot\beta - \beta\cdot\nabla) - (\beta\cdot\nabla)\nabla\cdot\beta + c^2 \end{pmatrix}. \end{aligned} \tag{3.4.26}$$

Consequently, the least-squares operator $\mathsf{L}^{b*}\mathsf{L}^b$ for the diffusive flux formulation with additional curl equations reads

$$\begin{aligned} \mathsf{L}^{b*}\mathsf{L}^b &= \begin{pmatrix} \mathsf{I} & -\nabla & (\nabla\times)^* \\ -\varepsilon\nabla\cdot & -\beta\cdot\nabla + c & 0 \end{pmatrix}\begin{pmatrix} \mathsf{I} & \varepsilon\nabla \\ \nabla\cdot & \nabla\cdot\beta + c \\ \nabla\times & 0 \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{I} - \triangle & \nabla[(\varepsilon - c)\mathsf{I} - \nabla\cdot\beta] \\ -[(\varepsilon - c)\mathsf{I} + (\beta\cdot\nabla)]\nabla\cdot & -\varepsilon^2\triangle + c(\nabla\cdot\beta - \beta\cdot\nabla) - (\beta\cdot\nabla)\nabla\cdot\beta + c^2 \end{pmatrix}. \end{aligned} \tag{3.4.27}$$

Total Flux in Operator form

The Linear Operators $\mathsf{L}^{a/b}$ for the diffusive flux are given by

$$\mathsf{L}^a_\varepsilon = \begin{pmatrix} \mathsf{I} & \varepsilon\nabla - \beta \\ \nabla\cdot & c \end{pmatrix} \tag{3.4.28}$$

and

$$\mathsf{L}^b = \begin{pmatrix} \mathsf{I} & \varepsilon\nabla - \beta \\ \nabla\cdot & c \\ \nabla\times & -\nabla\times\beta \end{pmatrix} \tag{3.4.29}$$

The least-squares operator $\mathsf{L}^{a*}\mathsf{L}^a$ for the total flux formulation is given by

$$\begin{aligned} \mathsf{L}^{a*}\mathsf{L}^a &= \begin{pmatrix} \mathsf{I} & -\nabla \\ -\varepsilon\nabla\cdot -\beta\cdot & c \end{pmatrix} \begin{pmatrix} \mathsf{I} & \varepsilon\nabla - \beta \\ \nabla\cdot & c \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{I} - \nabla\nabla\cdot & (\varepsilon - c)\nabla - \beta \\ -(\varepsilon - c)\nabla\cdot -\beta\cdot & -\varepsilon^2\triangle + \varepsilon\nabla\cdot\beta - \varepsilon\beta\cdot\nabla + \|\beta\|^2 + c^2 \end{pmatrix} \end{aligned} \tag{3.4.30}$$

and for the total flux formulation with additional curl equations is defined as

$$\begin{aligned} \mathsf{L}^{b*}\mathsf{L}^b &= \begin{pmatrix} \mathsf{I} & -\nabla & (\nabla\times)^* \\ -\varepsilon\nabla\cdot -\beta\cdot & c & T^* \end{pmatrix} \begin{pmatrix} \mathsf{I} & \varepsilon\nabla - \beta \\ \nabla\cdot & c \\ \nabla\times & T \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{I} - \triangle & (\varepsilon - c)\nabla - \beta + (\nabla\times)^* T \\ -(\varepsilon - c)\nabla\cdot -\beta\cdot + T^*\nabla\times & -\varepsilon^2\triangle + \varepsilon\nabla\cdot\beta - \varepsilon\beta\cdot\nabla + \|\beta\|^2 + c^2 + T^*T \end{pmatrix} \end{aligned} \tag{3.4.31}$$

where $T := -\nabla\times\beta$ and its adjoint operator is defined as $T^* := (-\nabla\times\beta)^*$.

We simplify the operator forms presented in (3.4.26), (3.4.27) and (3.4.30), (3.4.31) by considering the case of one dimensional advection-diffusion equation without reaction term, i.e. $c = 0$, and take the advecting vector $\beta = 1$. Therefore, the continuous form of the least-squares problem for the diffusive flux reads

$$\mathsf{L}^*\mathsf{L} = \begin{pmatrix} 1 - \frac{\partial^2}{\partial x^2} & \varepsilon\frac{\partial}{\partial x} - \frac{\partial^2}{\partial x^2} \\ -\varepsilon\frac{\partial}{\partial x} - \frac{\partial^2}{\partial x^2} & -(1+\varepsilon^2)\frac{\partial^2}{\partial x^2} \end{pmatrix}. \tag{3.4.32}$$

Similarly, the simplified continuous form of the least-squares problem for the total flux system reads

$$\mathsf{L}^*\mathsf{L} = \begin{pmatrix} 1 - \frac{\partial^2}{\partial x^2} & \varepsilon\frac{\partial}{\partial x} - 1 \\ -\varepsilon\frac{\partial}{\partial x} - 1 & -\varepsilon^2\frac{\partial^2}{\partial x^2} + 1 \end{pmatrix}. \tag{3.4.33}$$

It should be noted that the additional curl equations do not appear in the above one dimensional formulations. The operator forms show that the system is always symmetric. In addition, the total flux formulation is differentially diagonal dominant. This property helps to design efficient multigrid solvers for the solution of the discrete system with higher order finite elements [28].

3.4.3. Existence and Uniqueness of the Solution

The unique solution of the problem (3.4.16) depends on the coercivity and boundedness estimates of bilinear form $\mathcal{A}^{a/b}$. It should be noted that $\mathcal{J}^{a/b}(\mathbf{v},q;0,0) = \mathcal{A}^{a/b}(\mathbf{v},q;\mathbf{v},q)$.

Diffusive Flux

For the diffusive flux formulation (3.4.3) without the reaction term, i.e. $c = 0$, the following estimate can be found in [60]:

Theorem 3.4.1. *There exists constants $C_1, C_2 > 0$ independent of ε such that*

$$C_1 \left(\varepsilon^2 \|q\|_1^2 + \varepsilon^2 \|\mathbf{v}\|_{\text{div}}^2 \right) \leq \mathcal{J}^a(\mathbf{v}, q; 0, 0) \leq C_2 \left(\|q\|_1^2 + \|\mathbf{v}\|_{\text{div}}^2 \right) \tag{3.4.34}$$

for all $(\mathbf{v}, q) \in \mathbf{V} \times W$.

For the diffusive flux formulation (3.4.3) which is augmented by the curl equation (3.4.4) and with $c = 0$, the following estimate holds [31]:

Theorem 3.4.2. *There exists constants $C_1, C_2 > 0$ independent of ε such that*

$$C_1 \left(\varepsilon^2 \|q\|_1^2 + \varepsilon^2 \|\mathbf{v}\|_1^2 \right) \leq \mathcal{J}^b(\mathbf{v}, q; 0, 0) \leq C_2 \left(\|q\|_1^2 + \|\mathbf{v}\|_1^2 \right) \tag{3.4.35}$$

for all $(\mathbf{v}, q) \in \mathbf{V} \times W$.

For the proofs of the Theorems (3.4.1) and (3.4.2) one may refer to [60] and [31], respectively.

Total Flux

The coercivity and boundedness estimates of bilinear forms $\mathcal{A}^{a/b}$ for the total flux formulation (3.4.6) and the total flux with the curl equation (3.4.7) are discussed here.

Theorem 3.4.3. *There exist constants $C_3, C_4 > 0$ independent of ε such that*

$$\begin{aligned} C_3 \left(\varepsilon^2 \|q\|_1^2 + \varepsilon^2 \|\mathbf{v}\|_{\text{div}}^2 \right) \leq \quad & \mathcal{J}^a(\mathbf{v}, q; 0, 0) \leq C_4 \left(\|q\|_1^2 + \|\mathbf{v}\|_{\text{div}}^2 \right) \\ C_3 \left(\varepsilon^2 \|q\|_1^2 + \varepsilon^2 \|\mathbf{v}\|_1^2 \right) \leq \quad & \mathcal{J}^b(\mathbf{v}, q; 0, 0) \leq C_4 \left(\|q\|_1^2 + \|\mathbf{v}\|_1^2 \right) \end{aligned} \tag{3.4.36}$$

for all $(\mathbf{v}, q) \in \mathbf{V} \times W$.

Before we proceed with the proof of the Theorem (3.4.3), let us introduce the Poincaré-Friedrichs inequality [68]

$$\|q\|_0 \leq C \|\nabla q\|_0 \tag{3.4.37}$$

where, $C > 0$ is a constant. We also assume that there exist a constant $\gamma_0 \geq 0$ such that

$$c + \frac{1}{2} \nabla \cdot \beta \geq \gamma_0 \quad \text{in} \, \Omega. \tag{3.4.38}$$

Proof. The boundedness estimate can be easily proved. We proceed to show the validity of the coercivity, namely, the left hand sides of Theorem (3.4.3). We have

$$\mathcal{J}^a(\mathbf{v}, q; 0, 0) = \|\nabla \cdot \mathbf{v} + cq\|_0^2 + \|\mathbf{v} + \varepsilon \nabla q - \beta q\|_0^2 \tag{3.4.39}$$

$$\begin{aligned} \mathcal{J}^b(\mathbf{v}, q; 0, 0) = \|\nabla \cdot \mathbf{v} + cq\|_0^2 + \|\mathbf{v} + \varepsilon \nabla q - \beta q\|_0^2 \\ + \|\nabla \times (\mathbf{v} - \beta q)\|_0^2 \end{aligned} \tag{3.4.40}$$

lower bound with $\|q\|_1$

Let α be a positive constant that will be determined later. Applying the Poincaré-Friedrichs inequality (3.4.37),

$$\begin{aligned}
\mathcal{J}^{a/b}(\mathbf{v},q;0,0) \geq & \|\nabla\cdot\mathbf{v}+cq\|_0^2+\|\mathbf{v}+\varepsilon\nabla q-\beta q\|_0^2 \\
= & \|\nabla\cdot\mathbf{v}+cq-\alpha q\|_0^2-\alpha^2\|q\|_0^2+2\alpha c\|q\|_0^2 \\
& +2\alpha(\nabla\cdot\mathbf{v},q)_0 \\
& +\|\mathbf{v}+\varepsilon\nabla q-\beta q-\alpha\nabla q\|_0^2-\alpha^2\|\nabla q\|_0^2 \\
& +2\alpha(\mathbf{v},\nabla q)_0+2\alpha\varepsilon\|\nabla q\|_0^2-2\alpha(\beta q,\nabla q)_0 \\
= & \|\nabla\cdot\mathbf{v}+cq-\alpha q\|_0^2-\alpha^2\|q\|_0^2+2\alpha(c+\frac{1}{2}\nabla\cdot\beta)\|q\|_0^2 \\
& +\|\mathbf{v}+\varepsilon\nabla q-\beta q-\alpha\nabla q\|_0^2-\alpha^2\|\nabla q\|_0^2+2\alpha\varepsilon\|\nabla q\|_0^2 \\
\geq & -\alpha^2C^2\|\nabla q\|_0^2+2\alpha\varepsilon\|\nabla q\|_0^2-\alpha^2\|\nabla q\|_0^2
\end{aligned} \tag{3.4.41}$$

with $\alpha=\varepsilon/1+C^2$ we get

$$\mathcal{J}^{a/b}(\mathbf{v},q;0,0)\geq\frac{\varepsilon^2}{1+C^2}\|\nabla q\|_0^2 \tag{3.4.42}$$

furthermore, we have

$$\|q\|_1^2=\|q\|_0^2+\|\nabla q\|_0^2\leq(1+C^2)\|\nabla q\|_0^2 \tag{3.4.43}$$

then

$$\mathcal{J}_\varepsilon^{a/b}(\mathbf{v},q;0,0)\geq\frac{\varepsilon^2}{(1+C^2)^2}\|q\|_1^2 \tag{3.4.44}$$

lower bound with $\|\mathbf{v}\|_{\text{div}}$ ***and*** $\|\mathbf{v}\|_{\text{curl}}$

$$\begin{aligned}
\|\mathbf{v}\|_0^2 \leq & 2\left(\|\mathbf{v}+\varepsilon\nabla q-\beta q\|_0^2+\|\varepsilon\nabla q-\beta q\|_0^2\right) \\
\leq & 2\mathcal{J}^{a/b}(\mathbf{v},q;0,0)+C\|q\|_1^2\leq\frac{C}{\varepsilon^2}\mathcal{J}^{a/b}(\mathbf{v},q;0,0)
\end{aligned} \tag{3.4.45}$$

$$\begin{aligned}
\|\nabla\cdot\mathbf{v}\|_0^2 \leq & 2(\|\nabla\cdot\mathbf{v}+cq\|_0^2+\|cq\|_0^2) \\
\leq & 2\mathcal{J}^{a/b}(\mathbf{v},q;0,0)+C\|q\|_1^2\leq\frac{C}{\varepsilon^2}\mathcal{J}^{a/b}(\mathbf{v},q;0,0)
\end{aligned} \tag{3.4.46}$$

$$\begin{aligned}
\|\nabla\times\mathbf{v}\|_0^2 \leq & 2(\|\nabla\times(\mathbf{v}-\beta q)\|_0^2+\|\nabla\times(\beta q)\|_0^2) \\
\leq & 2\mathcal{J}^{b}(\mathbf{v},q;0,0)+C\|q\|_1^2\leq\frac{C}{\varepsilon^2}\mathcal{J}^{b}(\mathbf{v},q;0,0)
\end{aligned} \tag{3.4.47}$$

then, using the equivalence of the norm $\sqrt{\|\nabla\cdot\mathbf{v}\|_0^2+\|\nabla\times\mathbf{v}\|_0^2}$ and $\|\mathbf{v}\|_1$ on $H^1(\Omega)$, see [52], we get

$$\begin{aligned}
\mathcal{J}^{a}(\mathbf{v},q;0,0)\geq C\varepsilon^2\|\mathbf{v}\|_{\text{div}}^2 \\
\mathcal{J}^{b}(\mathbf{v},q;0,0)\geq C\varepsilon^2\|\mathbf{v}\|_1^2
\end{aligned} \tag{3.4.48}$$

This completes the proof. □

3.4.4. Discrete Least-squares Principles

We consider the approximation of the problem (3.4.16) with the finite element. Let the bounded domain $\Omega \subset \mathbb{R}^d$ be partitioned by a *grid* $\mathcal{T}_h$ consisting of elements $K \in \mathcal{T}_h$ which are assumed to be open quadrilaterals or hexahedrons such that $\Omega = \text{int}\big(\bigcup_{K\in\mathcal{T}_h} K\big)$. For an element $K \in \mathcal{T}_h$, we denote by h_K the diameter of the element K. The *mesh size* h of $\mathcal{T}_h$ is given by $h := \max_{K\in\mathcal{T}_h} h_K$.

We introduce the approximation spaces $\mathbf{V}_h^a$, $\mathbf{V}_h^b$ and W_h. Let $\mathbf{V}_h \times W_h := \mathbf{V}_h^{a/b} \times W_h$ and consider the approximation problem

find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ such that

$$\mathcal{A}_h^{a,b}(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) = \mathcal{F}_h(\mathbf{v}_h, q_h) \quad \forall (\mathbf{v}_h, q_h) \in \mathbf{V}_h \times W_h \tag{3.4.49}$$

where $\mathcal{A}_h^{a/b}$ is a bilinear form defined on $(\mathbf{V}_h \times W_h) \times (\mathbf{V}_h \times W_h) \to \mathbb{R}$, and has the following form for the diffusive flux formulation

$$\begin{aligned}\mathcal{A}_h^a(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) := &\sum_{K\in\mathcal{T}_h} (\mathbf{u}_h + \varepsilon\nabla p_h, \mathbf{v}_h + \varepsilon\nabla q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\cdot\mathbf{u}_h + \nabla\cdot(\beta p_h) + c p_h, \nabla\cdot\mathbf{v}_h + \nabla\cdot(\beta q_h) + c q_h)_{0,K}\end{aligned} \tag{3.4.50}$$

$$\begin{aligned}\mathcal{A}_h^b(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) := &\sum_{K\in\mathcal{T}_h} (\mathbf{u}_h + \varepsilon\nabla p_h, \mathbf{v}_h + \varepsilon\nabla q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\cdot\mathbf{u}_h + \nabla\cdot(\beta p_h) + c p_h, \nabla\cdot\mathbf{v}_h + \nabla\cdot(\beta q_h) + c q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\times\mathbf{u}_h, \nabla\times\mathbf{v}_h)_{0,K}\end{aligned} \tag{3.4.51}$$

and $\mathcal{F}_h$ is a linear form defined, on $\mathbf{V}_h \times W_h \to \mathbb{R}$, by

$$\mathcal{F}_h(\mathbf{v}_h, q_h) := \sum_{K\in\mathcal{T}_h} (f, \nabla\cdot\mathbf{v}_h + \nabla\cdot(\beta q_h) + c q_h)_{0,K}\,. \tag{3.4.52}$$

The bilinear form for the total flux formulation defined, on $(\mathbf{V}_h \times W_h) \times (\mathbf{V}_h \times W_h) \to \mathbb{R}$, by

$$\begin{aligned}\mathcal{A}_h^a(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) := &\sum_{K\in\mathcal{T}_h} (\mathbf{u}_h + \varepsilon\nabla p_h - \beta p_h, \mathbf{v}_h + \varepsilon\nabla q_h - \beta q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\cdot\mathbf{u}_h + c p_h, \nabla\cdot\mathbf{v}_h + c q_h)_{0,K}\end{aligned} \tag{3.4.53}$$

$$\begin{aligned}\mathcal{A}_h^b(\mathbf{u}_h, p_h; \mathbf{v}_h, q_h) := &\sum_{K\in\mathcal{T}_h} (\mathbf{u}_h + \varepsilon\nabla p_h - \beta p_h, \mathbf{v}_h + \varepsilon\nabla q_h - \beta q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\cdot\mathbf{u}_h + c p_h, \nabla\cdot\mathbf{v}_h + c q_h)_{0,K} \\ &+ \sum_{K\in\mathcal{T}_h} (\nabla\times(\mathbf{u}_h - \beta p_h), \nabla\times(\mathbf{v}_h - \beta q_h))_{0,K}\end{aligned} \tag{3.4.54}$$

and $\mathcal{F}_h$ for this system defined, on $\mathbf{V}_h \times W_h \to \mathbb{R}$, by

$$\mathcal{F}_h(\mathbf{v}_h, q_h) := \sum_{K\in\mathcal{T}_h} (f, \nabla\cdot\mathbf{v}_h + c q_h)_{0,K}. \tag{3.4.55}$$

In the above formulas the inner product $(\cdot,\cdot)_{0,K}$ is the restriction of the $(\cdot,\cdot)_0$ to a finite element K.

Conforming LSFEM

Let $\mathbf{V}_h \subset \mathbf{V}$ and $W_h \subset W$ be the approximating spaces consist of piecewise polynomials of order less or equal r ($r \geq 0$ is integer), then $\mathcal{A}_h^{a/b} = \mathcal{A}^{a/b}$ and $\mathcal{F}_h = \mathcal{F}$. As a result of the coercivity the problems (3.4.16) and (3.4.49) have unique solutions $(\mathbf{u}, p) \in \mathbf{V} \times W$ and $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$. Moreover, the error has the orthogonality property

$$\mathcal{A}^{a/b}(\mathbf{u} - \mathbf{u}_h, p - p_h; \mathbf{v}_h, p_h) = 0 \quad \forall (\mathbf{v}_h, q_h) \in \mathbf{V}_h \times W_h \tag{3.4.56}$$

and have the following error estimates:

Theorem 3.4.4. *Assume that the problem has a sufficiently regular solution* $(\mathbf{u}, p) \in \mathbf{V} \times W \cap (H^{r+1}(\Omega))^{d+1}$. *Then the approximate problem (3.4.49) has a unique solution* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times W_h$ *and satisfies the following error estimate: there exists a constant C independent of* ε *and h such that*

for grad−div *system*

$$\varepsilon \|\mathbf{u} - \mathbf{u}_h\|_{\text{div}} + \varepsilon \|p - p_h\|_1 \leq Ch^r (\|\mathbf{u}\|_{r+1} + \|p\|_{r+1}) \tag{3.4.57}$$

and for grad−div−curl *system*

$$\varepsilon \|\mathbf{u} - \mathbf{u}_h\|_1 + \varepsilon \|p - p_h\|_1 \leq Ch^r (\|\mathbf{u}\|_{r+1} + \|p\|_{r+1}) \tag{3.4.58}$$

Proof. From the interpolation theory we have

$$\|p - p_I\|_1 \leq Ch^r \|p\|_{r+1} \quad \text{and} \quad \|\mathbf{u} - \mathbf{u}_I\|_{\text{div}} \leq Ch^r \|\mathbf{u}\|_{r+1} \tag{3.4.59}$$

where $\mathbf{u}_I \in \mathbf{V}_h$ and $p_I \in W_h$ are standard interpolants of $\mathbf{u}$ and p. We define $\mathcal{A} := \mathcal{A}^{a/b}$ for the sake of simplicity. Based on the orthogonality property (3.4.56), we have

$$\begin{aligned}
&\mathcal{A}(\mathbf{u}_h - \mathbf{u}_I, p_h - p_I; \mathbf{u}_h - \mathbf{u}_I, p_h - p_I) \\
&= \mathcal{A}((\mathbf{u}_h - \mathbf{u}, p_h - p) + (\mathbf{u} - \mathbf{u}_I, p - p_I); \mathbf{u}_h - \mathbf{u}_I, p_h - p_I) \\
&= \mathcal{A}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{u}_h - \mathbf{u}_I, p_h - p_I) \\
&\leq \mathcal{A}^{\frac{1}{2}}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{u} - \mathbf{u}_I, p - p_I) \\
&\times \mathcal{A}^{\frac{1}{2}}(\mathbf{u}_h - \mathbf{u}_I, p_h - p_I; \mathbf{u}_h - \mathbf{u}_I, p_h - p_I)
\end{aligned} \tag{3.4.60}$$

and therefore

$$\begin{aligned}
&\mathcal{A}^{\frac{1}{2}}(\mathbf{u}_h - \mathbf{u}_I, p_h - p_I; \mathbf{u}_h - \mathbf{u}_I, p_h - p_I) \\
&\leq \mathcal{A}^{\frac{1}{2}}(\mathbf{u} - \mathbf{u}_I, p - p_I; \mathbf{u} - \mathbf{u}_I, p - p_I)
\end{aligned} \tag{3.4.61}$$

Now, using the inequality (3.4.61), the interpolation property (3.4.59) and the coercivity estimates in Theorems (3.4.1)-(3.4.3), we have

for grad−div system

$$\varepsilon \|\mathbf{u}_h - \mathbf{u}_I\|_{\text{div}} + \varepsilon \|p_h - p_I\|_1 \leq Ch^r (\|\mathbf{u}\|_{r+1} + \|p\|_{r+1}) \tag{3.4.62}$$

and for grad−div−curl system

$$\varepsilon \|\mathbf{u}_h - \mathbf{u}_I\|_1 + \varepsilon \|p_h - p_I\|_1 \leq Ch^r (\|\mathbf{u}\|_{r+1} + \|p\|_{r+1}) \tag{3.4.63}$$

We apply triangle inequality for the grad−div system

$$\begin{aligned}
&\varepsilon \|\mathbf{u} - \mathbf{u}_h\|_{\text{div}} + \varepsilon \|p - p_h\|_1 \\
&\leq \varepsilon (\|\mathbf{u} - \mathbf{u}_I\|_{\text{div}} + \|p - p_I\|_1) + \varepsilon (\|\mathbf{u}_I - \mathbf{u}_h\|_{\text{div}} + \|p_I - p_h\|_1) \\
&\leq Ch^r (\|\mathbf{u}\|_{r+1} + \|p\|_{r+1})
\end{aligned} \tag{3.4.64}$$

and for the grad−div−curl system

$$\begin{aligned}
&\varepsilon\|\mathbf{u}-\mathbf{u}_h\|_1+\varepsilon\|p-p_h\|_1\\
&\leq \varepsilon(\|\mathbf{u}-\mathbf{u}_I\|_1+\|p-p_I\|_1)+\varepsilon(\|\mathbf{u}_I-\mathbf{u}_h\|_1+\|p_I-p_h\|_1)\\
&\leq Ch^r(\|\mathbf{u}\|_{r+1}+\|p\|_{r+1})
\end{aligned} \tag{3.4.65}$$

This completes the proof. □

3.4.5. Continuous Least-Squares Principle with Scaling

In this section, the least-squares variational principles introduced earlier are slightly modified. The idea is to scale the energy functionals to obtain a more balanced variational problem [15, 78]. We introduce the following least-squares L^2-norm functionals for the diffusive flux

$$\mathcal{J}^a_\varepsilon(\mathbf{v},q;f,g)=\|\nabla\cdot\mathbf{v}+\nabla\cdot(\beta q)+cq-f\|_0^2+\frac{1}{\varepsilon}\|\mathbf{v}+\varepsilon\nabla q\|_0^2\quad\forall(\mathbf{v},q)\in\mathbf{V}\times W \tag{3.4.66}$$

$$\mathcal{J}^b_\varepsilon(\mathbf{v},q;f,g)=\|\nabla\cdot\mathbf{v}+\nabla\cdot(\beta q)+cq-f\|_0^2+\frac{1}{\varepsilon}\|\mathbf{v}+\varepsilon\nabla q\|_0^2+\|\nabla\times\mathbf{v}\|_0^2\quad\forall(\mathbf{v},q)\in\mathbf{V}\times W \tag{3.4.67}$$

and the following functionals for the total flux

$$\mathcal{J}^a_\varepsilon(\mathbf{v},q;f,g)=\|\nabla\cdot\mathbf{v}+cq-f\|_0^2+\frac{1}{\varepsilon}\|\mathbf{v}+\varepsilon\nabla q-\beta q\|_0^2\quad\forall(\mathbf{v},q)\in\mathbf{V}\times W \tag{3.4.68}$$

$$\mathcal{J}^b_\varepsilon(\mathbf{v},q;f,g)=\|\nabla\cdot\mathbf{v}+cq-f\|_0^2+\frac{1}{\varepsilon}\|\mathbf{v}+\varepsilon\nabla q-\beta q\|_0^2+\|\nabla\times(\mathbf{v}-\beta q)\|_0^2\quad\forall(\mathbf{v},q)\in\mathbf{V}\times W\,. \tag{3.4.69}$$

In the above scaled functionals, the inverse of the diffusion coefficient $\frac{1}{\varepsilon}$ is used as a wighting parameter. The associated variational problem, after minimizing the energy functional, is to find $(\mathbf{u},p)\in\mathbf{V}\times W$ such that

$$\mathcal{A}^{a/b}_\varepsilon(\mathbf{u},p;\mathbf{v},q)=\mathcal{F}_\varepsilon(\mathbf{v},q)\quad\forall(\mathbf{v},q)\in\mathbf{V}\times W \tag{3.4.70}$$

The bilinear and the linear forms are specifically denoted by a superscript ε in order to distinguish between the scaled LSFEM and the primitive LSFEM formulations. The definition of these bilinear and linear forms are very similar to the previous formulations and is not repeated here for brevity.

Diffusive Flux in Operator form

Restricting for simplicity to the $C_0^\infty(\Omega)$ functions, then we can rewrite the bilinear form $\mathcal{A}^{a/b}_\varepsilon$ as

$$\mathcal{A}^{a/b}_\varepsilon(\mathbf{u},p;\mathbf{v},q)=\left(\mathsf{L}^{a/b*}_\varepsilon\mathsf{L}^{a/b}_\varepsilon(\mathbf{u},p),(\mathbf{v},q)\right)_0 \tag{3.4.71}$$

where $\mathsf{L}^{a/b*}_\varepsilon$ is the formal adjoint of $\mathsf{L}^{a/b}_\varepsilon$ and the least-squares operator $\mathsf{L}^{a/b*}_\varepsilon\mathsf{L}^{a/b}_\varepsilon$ is given by

$$\begin{aligned}
\mathsf{L}^{a*}_\varepsilon\mathsf{L}^{a}_\varepsilon &= \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & -\nabla \\ -\sqrt{\varepsilon}\nabla\cdot & -\beta\cdot\nabla+c \end{pmatrix}\begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & \sqrt{\varepsilon}\nabla \\ \nabla\cdot & \nabla\cdot\beta+c \end{pmatrix}\\
&= \begin{pmatrix} \frac{\mathsf{I}}{\varepsilon}-\nabla\nabla\cdot & \nabla[(1-c)\mathsf{I}-\nabla\cdot\beta] \\ -[(1-c)\mathsf{I}+(\beta\cdot\nabla)]\nabla\cdot & -\varepsilon\triangle+c(\nabla\cdot\beta-\beta\cdot\nabla)-(\beta\cdot\nabla)\nabla\cdot\beta+c^2 \end{pmatrix}
\end{aligned} \tag{3.4.72}$$

and

$$\begin{aligned} \mathsf{L}_\varepsilon^{b*}\mathsf{L}_\varepsilon^{b} &= \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & -\nabla & (\nabla\times)^* \\ -\sqrt{\varepsilon}\nabla\cdot & -\beta\cdot\nabla + c & 0 \end{pmatrix} \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & \sqrt{\varepsilon}\nabla \\ \nabla\cdot & \nabla\cdot\beta + c \\ \nabla\times & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\mathsf{I}}{\varepsilon} - \triangle & \nabla[(1-c)\mathsf{I} - \nabla\cdot\beta] \\ -[(1-c)\mathsf{I} + (\beta\cdot\nabla)]\nabla\cdot & -\varepsilon\triangle + c(\nabla\cdot\beta - \beta\cdot\nabla) - (\beta\cdot\nabla)\nabla\cdot\beta + c^2 \end{pmatrix} \end{aligned} \tag{3.4.73}$$

for the diffusive flux and the diffusive flux with the additional curl equations, respectively.

Total Flux in Operator form

Similarly, the least-squares operator $\mathsf{L}_\varepsilon^{a/b*}\mathsf{L}_\varepsilon^{a/b}$ for the total flux is given by

$$\begin{aligned} \mathsf{L}_\varepsilon^{a*}\mathsf{L}_\varepsilon^{a} &= \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & -\nabla \\ -\sqrt{\varepsilon}\nabla\cdot - \frac{\beta\cdot}{\sqrt{\varepsilon}} & c \end{pmatrix} \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & \sqrt{\varepsilon}\nabla - \frac{\beta}{\sqrt{\varepsilon}} \\ \nabla\cdot & c \end{pmatrix} \\ &= \begin{pmatrix} \frac{\mathsf{I}}{\varepsilon} - \nabla\nabla\cdot & (1-c)\nabla - \frac{\beta}{\varepsilon} \\ -(1-c)\nabla\cdot - \frac{\beta\cdot}{\varepsilon} & -\varepsilon\triangle + \nabla\cdot\beta - \beta\cdot\nabla + \frac{\|\beta\|^2}{\varepsilon} + c^2 \end{pmatrix} \end{aligned} \tag{3.4.74}$$

and

$$\begin{aligned} \mathsf{L}_\varepsilon^{b^*}\mathsf{L}_\varepsilon^{b} &= \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & -\nabla & (\nabla\times)^* \\ -\sqrt{\varepsilon}\nabla\cdot - \frac{\beta\cdot}{\sqrt{\varepsilon}} & c & T^* \end{pmatrix} \begin{pmatrix} \frac{\mathsf{I}}{\sqrt{\varepsilon}} & \sqrt{\varepsilon}\nabla - \frac{\beta}{\sqrt{\varepsilon}} \\ \nabla\cdot & c \\ \nabla\times & T \end{pmatrix} \\ &= \begin{pmatrix} \frac{\mathsf{I}}{\varepsilon} - \triangle & (1-c)\nabla - \beta + (\nabla\times)^*T \\ -(1-c)\nabla\cdot - \frac{\beta\cdot}{\varepsilon} + T^*\nabla\times & -\varepsilon\triangle + \nabla\cdot\beta - \beta\cdot\nabla + \frac{\|\beta\|^2}{\varepsilon} + c^2 + T^*T \end{pmatrix}. \end{aligned} \tag{3.4.75}$$

for the total flux and the total flux with the additional curl equations, respectively. Simplifying the operator forms presented in (3.4.72)-(3.4.75) by considering the case of one dimensional advection-diffusion equation without reaction term, i.e. $c = 0$ and taking the advecting vector $\beta = 1$, we obtain the following system of equations for the diffusive flux

$$\mathsf{L}_\varepsilon^*\mathsf{L}_\varepsilon = \begin{pmatrix} \frac{1}{\varepsilon} - \frac{\partial^2}{\partial x^2} & \frac{\partial}{\partial x} - \frac{\partial^2}{\partial x^2} \\ -\frac{\partial}{\partial x} - \frac{\partial^2}{\partial x^2} & -(1+\varepsilon)\frac{\partial^2}{\partial x^2} \end{pmatrix} \tag{3.4.76}$$

and correspondingly for the total flux

$$\mathsf{L}_\varepsilon^*\mathsf{L}_\varepsilon = \begin{pmatrix} \frac{1}{\varepsilon} - \frac{\partial^2}{\partial x^2} & \frac{\partial}{\partial x} - \frac{1}{\varepsilon} \\ -\frac{\partial}{\partial x} - \frac{1}{\varepsilon} & -\varepsilon\frac{\partial^2}{\partial x^2} + \frac{1}{\varepsilon} \end{pmatrix}. \tag{3.4.77}$$

Existence and Uniqueness of the Solution

The unique solution of the problem (3.4.70) depends on the coercivity and boundedness estimates of bilinear form $\mathcal{A}_{\varepsilon}^{a/b}$.

Theorem 3.4.5. *There exist constants $C_5, C_6 > 0$ depending on ε such that*
for the grad−div *system*

$$C_5\left(\varepsilon^2\|q\|_1^2+\varepsilon^2\|\mathbf{v}\|_{\mathrm{div}}^2\right) \leq \quad \mathcal{J}_{\varepsilon}^{a/b}(\mathbf{v},q;0,0) \leq C_6\left(\|q\|_1^2+\|\mathbf{v}\|_{\mathrm{div}}^2\right) \tag{3.4.78}$$

and for the grad−div−curl *system*

$$C_5\left(\varepsilon^2\|q\|_1^2+\varepsilon^2\|\mathbf{v}\|_1^2\right) \leq \quad \mathcal{J}_{\varepsilon}^{a/b}(\mathbf{v},q;0,0) \leq C_6\left(\|q\|_1^2+\|\mathbf{v}\|_1^2\right) \tag{3.4.79}$$

for all $(\mathbf{v},q) \in \mathbf{V}\times W$.

Proof. The proof is similar to that of Theorem (3.4.3). □

3.4.6. Discrete Least-Squares Principle with Scaling

The discrete counterpart of the variational principle in (3.4.70) reads

$$\mathcal{A}_{\varepsilon,h}^{a/b}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) = \mathcal{F}_{\varepsilon,h}(\mathbf{v}_h,q_h) \quad \forall(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times W_h. \tag{3.4.80}$$

Assuming the conformity of the finite element spaces and the condition of Theorem (3.4.4), the following error estimate holds for the scaled LSFEMs:

Theorem 3.4.6. *Solution of the approximate problem (3.4.80) has a unique solution $(\mathbf{u}_h,p_h)\in \mathbf{V}_h\times W_h$ and satisfies the following error estimate: there exists a constant C_ε independent of h such that*
for grad−div *system*

$$\varepsilon\|\mathbf{u}-\mathbf{u}_h\|_{\mathrm{div}}+\varepsilon\|p-p_h\|_1 \leq C_\varepsilon h^r(\|\mathbf{u}\|_{r+1}+\|p\|_{r+1}) \tag{3.4.81}$$

and for grad−div−curl *system*

$$\varepsilon\|\mathbf{u}-\mathbf{u}_h\|_1+\varepsilon\|p-p_h\|_1 \leq C_\varepsilon h^r(\|\mathbf{u}\|_{r+1}+\|p\|_{r+1}) \tag{3.4.82}$$

Proof. The proof is similar to that of Theorem (3.4.4). We omit the details for brevity. □

Remark 3.4.7. *The constant C_ε in Theorem (3.4.6) depends on the diffusion coefficient ε which is in contrast to the constant C (ε-independent) in the a priori error estimate of the primitive LSFEM in Theorem (3.4.4). Based on this fact, we expect better results out of the scaled system for all flow regimes including highly convection-dominated flows with very small ε.*

3.4.7. Stabilized LSFEM

In this section, the idea of the stabilized LSFEM is introduced. The main goal of this setting is to achieve more stable numerical solutions for the advection-diffusion equation in highly advection-dominated flows [60]. The following materials on the stabilization are based on the work by Hsieh and Yang [60].

It should be noted that, Hsieh and Yang [60] have taken the diffusive flux formulation as the basis of their analysis; however, in this work we extend the stabilization idea to the total flux formulation as well.

Stabilization Technique

In this type of stabilization, the primitive least-squares functionals (3.4.11) and (3.4.13) are taken and some stabilization terms, as in the Galerkin Least-Squares method, are added to them. Due to the low regularity of the functions in the space $\mathbf{V}\times W$, the additional terms in the energy functional are restricted and defined on the finite element space $\mathbf{V}_h \times W_h$. Therefore, the modified least-squares functional is as follows:

for the diffusive flux formulation

$$\begin{aligned}\mathcal{J}^a_{\delta,h}(\mathbf{v}_h,q_h;f,g) &= \|\nabla\cdot\mathbf{v}_h+\nabla\cdot(\beta q_h)+cq_h-f\|_0^2+\|\mathbf{v}_h+\varepsilon\nabla q_h\|_0^2\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K\|-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h-f\|_{0,K}^2 \quad \forall(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times W_h\end{aligned} \tag{3.4.83}$$

and for the total flux formulation

$$\begin{aligned}\mathcal{J}^b_{\delta,h}(\mathbf{v}_h,q_h;f,g) &= \|\nabla\cdot\mathbf{v}_h+cq_h-f\|_0^2+\|\mathbf{v}_h+\varepsilon\nabla q_h-\beta q_h\|_0^2\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K\|-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h-f\|_{0,K}^2 \quad \forall(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times W_h\end{aligned} \tag{3.4.84}$$

where $\|\cdot\|_{0,K}$ denotes the L^2-norm restricted to the finite element K and δ_K is a stabilization parameter which should be chosen appropriately. Taking the first variation of the functionals (3.4.83) and (3.4.84), we setup the following stabilized LSFEM problem: find $(\mathbf{u}_h,p_h)\in\mathbf{V}_h\times W_h$ such that

$$\mathcal{A}^{a/b}_{\delta,h}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h)=\mathcal{F}_{\delta,h}(\mathbf{v}_h,q_h)\quad\forall(\mathbf{v}_h,q_h)\in\mathbf{V}_h\times W_h \tag{3.4.85}$$

The modified bilinear form for the diffusive flux formulation is defined as follows

$$\begin{aligned}\mathcal{A}^a_{\delta,h}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) &:= \sum_{K\in\mathcal{T}_h}(\mathbf{u}_h+\varepsilon\nabla p_h,\mathbf{v}_h+\varepsilon\nabla q_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}(\nabla\cdot\mathbf{u}_h+\nabla\cdot(\beta p_h)+cp_h,\nabla\cdot\mathbf{v}_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K(-\varepsilon\Delta p_h+\nabla\cdot(\beta p_h)+cp_h,-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}.\end{aligned} \tag{3.4.86}$$

The modified linear form for the diffusive flux formulation then reads

$$\begin{aligned}\mathcal{F}_{\delta,h}(\psi,\mathbf{v}) &:= \sum_{K\in\mathcal{T}_h}(f,\nabla\cdot\mathbf{v}_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K(f,-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}.\end{aligned} \tag{3.4.87}$$

The modified bilinear form for the total flux formulation is defined as follows

$$\begin{aligned}\mathcal{A}^b_{\delta,h}(\mathbf{u}_h,p_h;\mathbf{v}_h,q_h) &:= \sum_{K\in\mathcal{T}_h}(\mathbf{u}_h+\varepsilon\nabla p_h-\beta p_h,\mathbf{v}_h+\varepsilon\nabla q_h-\beta q_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}(\nabla\cdot\mathbf{u}_h+cp_h,\nabla\cdot\mathbf{v}_h+cq_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K(-\varepsilon\Delta p_h+\nabla\cdot(\beta p_h)+cp_h,-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}\end{aligned} \tag{3.4.88}$$

and the linear form in this case reads

$$\begin{aligned}\mathcal{F}_{\delta,h}(\psi,\mathbf{v}) &:= \sum_{K\in\mathcal{T}_h}(f,\nabla\cdot\mathbf{v}_h+cq_h)_{0,K}\\ &\quad+\sum_{K\in\mathcal{T}_h}\delta_K(f,-\varepsilon\Delta q_h+\nabla\cdot(\beta q_h)+cq_h)_{0,K}.\end{aligned} \tag{3.4.89}$$

Remark 3.4.8. *The error estimate for the stabilized LSFEM based on the diffusive flux formulation of the advection-diffusion equation, i.e. $c=0$ in equation (3.4.1), can be found in [60]. Similarly, one may derive the error estimate for the total flux formulation.*

3.4.8. Numerical Results and Discussion

In this section, we consider the solution of the advection-diffusion equation, i.e. $c = 0$ in equation (3.4.1), with different variations of the LSFEM.

In the first example, the theoretical error estimates of the Theorems 3.4.4 and 3.4.6 are investigated using a manufactured analytical solution. Both the grad−div and the grad−div−curl systems are considered.

In the next two examples, the behavior of the different LSFEM formulations for the solution of advection-dominated flows are studied. The results are graphically presented and compared with those of the standard Galerkin finite element method.

The effect of variations of the diffusion coefficient and the advecting velocity field on the behavior of the solution can be better analyzed with the local non-dimensional Peclet number (*Pe*) defined as following

$$Pe = \frac{\|\beta\| h_K}{2\varepsilon} \tag{3.4.90}$$

where $\|\cdot\|$ represents the L^2 norm and h_K is the characteristic element length. *Pe* number represents the balance of the two transport mechanisms, namely the convection transport and the diffusion transport mechanisms. In other words, the *Pe* number is the ratio of the convection to the diffusion transport phenomena. In a convection-dominated flow, for example for large convecting velocity fields or for small diffusion coefficients, the *Pe* number is very large. *Pe* number can be defined globally for the flow, in which the local element size h_K in equation 3.4.90 should be replaced with the flow length scale h. In this study, we use the local *Pe* number.

Analytic Solution: Example I

The problem domain is a unit square which is subjected to the homogeneous Dirichlet boundary conditions. We assume a constant advecting velocity field defined by $\beta = [1,1]^T$. The value of the diffusion coefficient varies in order to obtain different *Pe* numbers and hence different flow regimes. An analytic manufactured solution which is independent of ε is considered as follows

$$p = xy(1-x)(1-y)e^{-x}. \tag{3.4.91}$$

Substituting the equation (3.4.91) into the main governing equation (3.4.1), we obtain the source term f. We use equal-order finite elements for the vector and the scalar variables.

Using the Q_1 finite elements, the relative L^2-error results are presented in Table 3.15 and 3.16 for the primitive LSFEM and in Table 3.17 and 3.18 for the LSFEM with scaling. The numerical results of Table 3.15 to Table 3.18 show that both scaled LSFEM systems perform very well with respect to the scalar variable p even in the highly advection-dominated flows. The total flux results of the scaled LSFEM, Table 3.18, show optimal L^2-error convergence of the fluxes for all values of ε. For the diffusive flux formulation of the scaled LSFEM, the convergence rates of the fluxes degrade in advection-dominated flows, see Table 3.17.

As we observed for the Q_1 elements, the scaled LSFEM performs better than the primitive LSFEM. Therefore, for the Q_2 elements we only analyze the scaled formulation. We summarize the results of the scaled LSFEM for the Q_2 elements in Table 3.19 and 3.20. The approximation of the scalar and the vector variables in the total flux formulation remains optimal for different flow regimes.

Next, we investigate the performance of the stabilized LSFEM and compare the results with those of the scaled LSFEM for the Q_1 elements in Tables 3.21 and 3.22 and for the Q_2 elements in Tables 3.23 and 3.24. Different values of the stabilization parameter δ are investigated for both the diffusive and the total flux formulations. The results indicate that the scaled LSFEM performs

Table 3.15: Primitive LSFEM, diffusive flux $\mathbf{u} = -\varepsilon\nabla p$, Q_1 elements, example I

	LSFEM				Standard			
Level	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate
$\varepsilon = 1$								
6	1.221e-03	–	7.311e-04	–	1.504e-03	–	3.706e-02	–
7	3.042e-04	4.01	1.766e-04	4.14	3.762e-04	4.00	1.853e-02	2.00
8	7.599e-05	4.00	4.381e-05	4.03	9.404e-05	4.00	9.265e-03	2.00
9	1.899e-05	4.00	1.094e-05	4.00	2.351e-05	4.00	4.633e-03	2.00
$\varepsilon = 1e-2$								
6	2.742e-02	–	1.582e-01	–	2.148e-03	–	3.721e-02	–
7	1.399e-02	1.96	6.348e-02	2.49	5.364e-04	4.00	1.855e-02	2.01
8	5.888e-03	2.38	2.171e-02	2.92	1.341e-04	4.00	9.268e-03	2.00
9	1.988e-03	2.96	6.427e-03	3.38	3.352e-05	4.00	4.633e-03	2.00
$\varepsilon = 1e-6$								
6	2.790e-02	–	1.597e+03	–	3.542e-01	–	6.008e+00	–
7	1.551e-02	1.80	6.762e+02	2.36	2.234e-02	15.85	7.523e-01	7.99
8	8.135e-03	1.91	2.614e+02	2.59	1.440e-03	15.52	9.453e-02	7.96
9	4.150e-03	1.96	9.622e+01	2.72	1.026e-04	14.03	1.271e-02	7.44

The curl constraint is also added to the LSFEM equations.

much better for small diffusion coefficients and very robust for the whole range of flow regimes. The stabilized LSFEM is unable to produce optimal convergence rates for the fluxes.

Interior Layer Flow: Example II

In this example, our computational domain is a unit square. An advecting velocity $\beta = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T$ is considered, the source term is $f = 0$ and the diffusion constant is $\varepsilon = 10^{-4}$, i.e. $Pe = 78$. Equal-order finite elements and a computational mesh of 64×64 elements are used. Boundary conditions are defined as follows

$$p = 0 \text{ on } \{(x,y) : 0 < x \leq 1, y = 1\} \cup \{(x,y) : x = 0, 0 \leq y \leq 1\} \tag{3.4.92}$$

$$p = 1 \text{ on } \{(x,y) : 0 \leq x \leq 1, y = 0\} \cup \{(x,y) : x = 1, 0 \leq y < 1\}. \tag{3.4.93}$$

This is an interior layer problem due to the discontinuity of the boundary data and is highly convection-dominated with $Pe = 78$. The p contour plots for the different LSFEMs are compared with the standard Galerkin FEM in Fig. 3.3 and 3.4 for Q_1 and Q_2 elements, respectively. Since the diffusive and the total flux formulations qualitatively produce the same results, only one set of the results is depicted here. The Q_1 element result of the standard Galerkin FEM is very oscillatory as it is expected from the flow Peclet number. The primitive LSFEM is also very diffusive although there is no over/under shoot in the result. The scaled and the stabilized LSFEM methods perform much better than the primitive LSFEM and the standard Galerkin methods. Although, some over/under shoots can be observed in the results. By increasing the degree of our finite element spaces, Q_2 elements, all LSFEM variations perform well. The standard Galerkin FEM result is also reasonably good for the Q_2 elements.

Table 3.16: Primitive LSFEM, total flux $\mathbf{u} = -\varepsilon\nabla p + \beta p$, Q_1 elements, example I

	LSFEM				Standard			
Level	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate
$\varepsilon = 1$								
6	1.493e-03	–	6.222e-03	–	1.504e-03	–	3.544e-02	–
7	3.620e-04	4.13	1.999e-03	3.11	3.762e-04	4.00	1.772e-02	2.00
8	8.877e-05	4.08	6.648e-04	3.01	9.404e-05	4.00	8.858e-03	2.00
9	2.196e-05	4.04	2.269e-04	2.93	2.351e-05	4.00	4.429e-03	2.00
$\varepsilon = 1e-2$								
6	9.630e-02	–	9.863e-02	–	2.148e-03	–	2.466e-03	–
7	3.405e-02	2.83	3.493e-02	2.82	5.364e-04	4.00	8.082e-04	3.05
8	9.081e-03	3.75	9.463e-03	3.69	1.341e-04	4.00	3.305e-04	2.45
9	2.226e-03	4.08	2.392e-03	3.96	3.352e-05	4.00	1.547e-04	2.14
$\varepsilon = 1e-6$								
6	1.283e-01	–	1.294e-01	–	3.542e-01	–	3.542e-01	–
7	7.343e-02	1.76	7.355e-02	1.76	2.234e-02	15.85	2.234e-02	15.85
8	4.044e-02	1.82	4.043e-02	1.82	1.440e-03	15.52	1.440e-03	15.52
9	2.160e-02	1.87	2.159e-02	1.87	1.026e-04	14.03	1.026e-04	14.03

Table 3.17: LSFEM with scaling, diffusive flux $\mathbf{u} = -\varepsilon\nabla p$, Q_1 elements, example I

	LSFEM				Standard			
Level	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate
$\varepsilon = 1e-2$								
6	1.992e-03	–	4.388e-03	–	2.148e-03	–	3.721e-02	–
7	5.875e-04	3.39	1.257e-03	3.49	5.364e-04	4.00	1.855e-02	2.01
8	1.590e-04	3.69	3.289e-04	3.82	1.341e-04	4.00	9.268e-03	2.00
9	4.089e-05	3.89	8.338e-05	3.94	3.352e-05	4.00	4.633e-03	2.00
$\varepsilon = 1e-6$								
6	1.435e-03	–	1.497e-02	–	3.542e-01	–	6.008e+00	–
7	3.600e-04	3.99	5.051e-03	2.96	2.234e-02	15.85	7.523e-01	7.99
8	9.012e-05	3.99	1.726e-03	2.93	1.440e-03	15.52	9.453e-02	7.96
9	2.254e-05	4.00	5.935e-04	2.91	1.026e-04	14.03	1.271e-02	7.44

The curl constraint is also added to the LSFEM equations.

Table 3.18: LSFEM with scaling, total flux $\mathbf{u} = -\varepsilon\nabla p + \beta p$, Q_1 elements, example I

	LSFEM				Standard			
Level	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate	$\|p - p_h\|_0$	Rate	$\|\mathbf{u} - \mathbf{u}_h\|_0$	Rate
$\varepsilon = 1e-2$								
6	2.128e-03	–	2.192e-03	–	2.148e-03	–	2.466e-03	–
7	6.097e-04	3.49	6.339e-04	3.46	5.364e-04	4.00	8.082e-04	3.05
8	1.610e-04	3.79	1.710e-04	3.71	1.341e-04	4.00	3.305e-04	2.45
9	4.092e-05	3.93	4.529e-05	3.77	3.352e-05	4.00	1.547e-04	2.14
$\varepsilon = 1e-6$								
6	1.435e-03	–	1.435e-03	–	3.542e-01	–	3.542e-01	–
7	3.600e-04	3.99	3.600e-04	3.99	2.234e-02	15.85	2.234e-02	15.85
8	9.012e-05	3.99	9.012e-05	3.99	1.440e-03	15.52	1.440e-03	15.52
9	2.254e-05	4.00	2.254e-05	4.00	1.026e-04	14.03	1.026e-04	14.03

Table 3.19: LSFEM with scaling, diffusive flux $\mathbf{u} = -\varepsilon\nabla p$, Q_2 elements, example I

	LSFEM				Standard			
Level	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate
$\varepsilon = 1$								
6	5.906e-06	–	3.388e-06	–	5.865e-06	–	3.152e-04	–
7	7.345e-07	8.04	4.178e-07	8.11	7.333e-07	8.00	7.880e-05	4.00
8	9.170e-08	8.01	5.202e-08	8.03	9.166e-08	8.00	1.970e-05	4.00
$\varepsilon = 1e-2$								
6	1.180e-05	–	1.638e-04	–	5.865e-06	–	3.152e-04	–
7	1.073e-06	11.00	1.291e-05	12.69	7.333e-07	8.00	7.880e-05	4.00
8	1.056e-07	10.16	8.842e-07	14.60	9.166e-08	8.00	1.970e-05	4.00
$\varepsilon = 1e-6$								
6	5.874e-06	–	1.393e-03	–	2.501e-02	–	7.748e-01	–
7	7.336e-07	8.01	3.537e-04	3.94	1.572e-03	15.91	9.691e-02	8.00
8	9.168e-08	8.00	8.898e-05	3.97	1.007e-04	15.62	1.214e-02	7.98

The curl constraint is also added to the LSFEM equations.

Table 3.20: LSFEM with scaling, total flux $\mathbf{u} = -\varepsilon\nabla p + \beta p$, Q_2 elements, example I

	LSFEM				Standard			
Level	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate	$\|p-p_h\|_0$	Rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	Rate
$\varepsilon = 1$								
6	5.865e-06	–	4.858e-06	–	5.865e-06	–	3.013e-04	–
7	7.333e-07	8.00	6.072e-07	8.00	7.333e-07	8.00	7.534e-05	4.00
8	9.166e-08	8.00	7.589e-08	8.00	9.166e-08	8.00	1.884e-05	4.00
$\varepsilon = 1e-2$								
6	1.184e-05	–	1.013e-05	–	9.300e-06	–	1.444e-05	–
7	1.037e-06	11.41	9.473e-07	10.69	8.638e-07	10.77	2.758e-06	5.23
8	1.033e-07	10.05	1.007e-07	9.41	9.603e-08	9.00	6.525e-07	4.23
$\varepsilon = 1e-6$								
6	5.874e-06	–	5.873e-06	–	2.501e-02	–	2.501e-02	–
7	7.336e-07	8.01	7.336e-07	8.01	1.572e-03	15.91	1.572e-03	15.91
8	9.168e-08	8.00	9.168e-08	8.00	1.007e-04	15.62	1.007e-04	15.62

Table 3.21: Diffusive flux $\mathbf{u} = -\varepsilon\nabla p$, Q_1 elements at $\varepsilon = 1e-6$, example I

Method	$p/\mathbf{u}$	$h_K = 1/32$	$h_K = 1/64$	$h_K = 1/128$	$h_K = 1/256$	$\approx$ Rate
Scaled LSFEM	p	1.435e-03	3.600e-04	9.012e-05	2.254e-05	3.99
	$\mathbf{u}$	1.497e-02	5.051e-03	1.726e-03	5.935e-04	2.93
Stabilized LSFEM						
$\delta_K = 0$	p	2.790e-02	1.551e-02	8.135e-03	4.150e-03	1.89
	$\mathbf{u}$	1.597e+03	6.762e+02	2.614e+02	9.622e+01	2.93
$\delta_K = h_K^2$	p	2.207e-02	1.250e-02	6.596e-03	3.365e-03	1.88
	$\mathbf{u}$	1.342e+03	5.914e+02	2.327e+02	8.628e+01	2.51
$\delta_K = 0.1h_K^2$	p	2.712e-02	1.510e-02	7.926e-03	4.043e-03	1.89
	$\mathbf{u}$	1.565e+03	6.653e+02	2.577e+02	9.492e+01	2.55

Table 3.22: Total flux $\mathbf{u} = -\varepsilon\nabla p + \beta p$, Q_1 elements at $\varepsilon = 1e-6$, example I

Method	$p/\mathbf{u}$	$h_K = 1/32$	$h_K = 1/64$	$h_K = 1/128$	$h_K = 1/256$	$\approx$ Rate
Scaled LSFEM	p	1.435e-03	3.600e-04	9.012e-05	2.254e-05	3.99
	$\mathbf{u}$	1.435e-03	3.600e-04	9.012e-05	2.254e-05	3.99
Stabilized LSFEM						
$\delta_K = 0$	p	1.283e-01	7.343e-02	4.044e-02	2.160e-02	1.81
	$\mathbf{u}$	1.294e-01	7.355e-02	4.043e-02	2.159e-02	1.82
$\delta_K = h_K^2$	p	4.339e-02	2.352e-02	1.234e-02	6.360e-03	1.90
	$\mathbf{u}$	4.873e-02	2.477e-02	1.264e-02	6.431e-03	1.96
$\delta_K = 0.1h_K^2$	p	9.807e-02	5.511e-02	2.984e-02	1.574e-02	1.84
	$\mathbf{u}$	1.004e-01	5.554e-02	2.992e-02	1.576e-02	1.86

Table 3.23: Diffusive flux $\mathbf{u} = -\varepsilon\nabla p$, Q_2 elements at $\varepsilon = 1e-6$, example I

Method	$p/\mathbf{u}$	$h_K = 1/32$	$h_K = 1/64$	$h_K = 1/128$	$\approx$ Rate
Scaled LSFEM	p	5.874e-06	7.336e-07	9.168e-08	8.00
	$\mathbf{u}$	1.393e-03	3.537e-04	8.898e-05	3.95
Stabilized LSFEM					
$\delta_K = 0$	p	1.738e-04	3.661e-05	7.969e-06	4.67
	$\mathbf{u}$	1.466e+01	3.502e+00	8.572e-01	4.14
$\delta_K = h_K^2$	p	6.547e-05	1.526e-05	3.655e-06	4.24
	$\mathbf{u}$	8.281e+00	2.010e+00	4.943e-01	4.10
$\delta_K = 0.1h_K^2$	p	1.393e-04	3.056e-05	6.885e-06	4.50
	$\mathbf{u}$	1.350e+01	3.248e+00	7.972e-01	4.11

Table 3.24: Total flux $\mathbf{u} = -\varepsilon\nabla p + \beta p$, Q_2 elements at $\varepsilon = 1e-6$, example I

Method	$p/\mathbf{u}$	$h_K = 1/32$	$h_K = 1/64$	$\approx$ Rate
Scaled LSFEM	p	5.874e-06	7.336e-07	8.00
	$\mathbf{u}$	5.873e-06	7.336e-07	8.00
Stabilized LSFEM				
$\delta_K = 0$	p	3.940e-04	1.002e-04	3.93
	$\mathbf{u}$	3.274e-04	8.380e-05	3.91
$\delta_K = h_K^2$	p	1.441e-05	3.324e-06	4.33
	$\mathbf{u}$	1.229e-05	2.798e-06	4.39
$\delta_K = 0.1h_K^2$	p	1.441e-05	3.324e-06	4.33
	$\mathbf{u}$	1.229e-05	2.798e-06	4.39

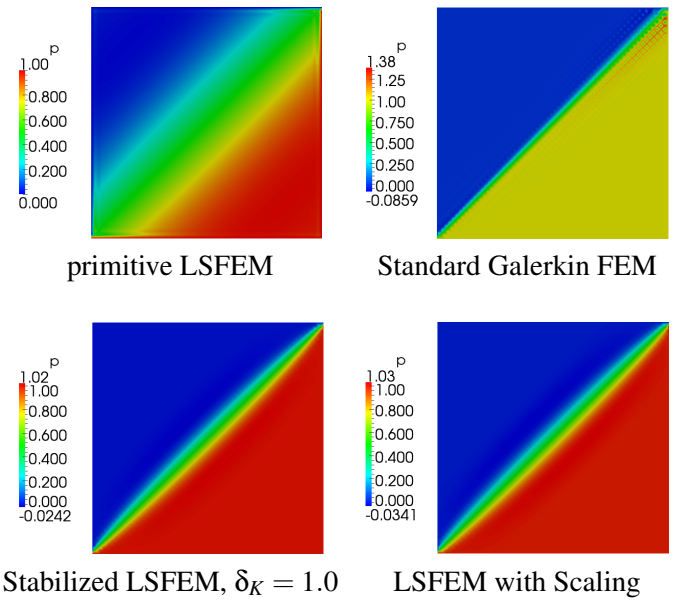


Figure 3.3: Comparison of p contour plots between the LSFEMs and the standard Galerkin FEM at $Pe = 78$ using the Q_1 elements, example II

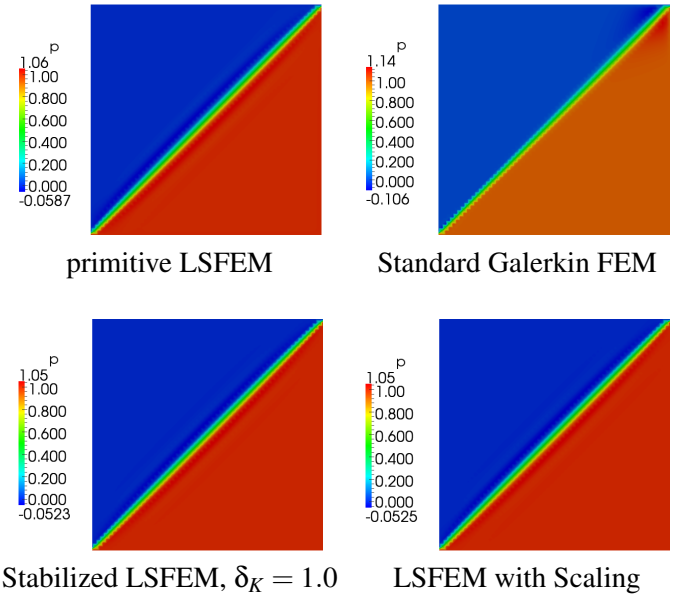


Figure 3.4: Comparison of p contour plots between the LSFEMs and the standard Galerkin FEM at $Pe = 78$ using the Q_2 elements, example II

Boundary Layer Flow: Example III

We consider the transport of the scalar variable p inside a unit square domain. A circular cylinder with non-homogeneous boundary condition is placed in the middle of the domain. The schematic flow configuration and the boundary conditions are depicted in Fig. 3.5. The constant velocity field is $\beta = [1,0]^T$, the source term is $f = 0$ and the diffusion constant varies between $\varepsilon = 10^{-4}$ and $\varepsilon = 10^{-6}$.

When the flow reaches the circular cylinder, the inhomogeneity of the boundary condition across the cylinder induces a boundary layer to the flow. This boundary layer then will be carried with the flow to the downstream. The numerical methods which have no special treatment for the convective terms, will show spurious oscillations when the flow becomes convection-dominated. We investigate the solution of this particular problem with our LSFEMs and compare the results with those obtained by the standard Galerkin FEM.

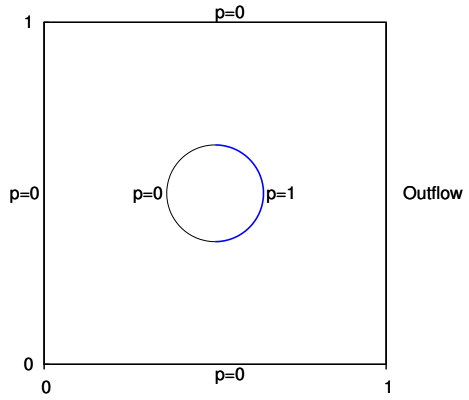


Figure 3.5: Domain configuration and the boundary conditions, transport of a scalar source term in a boundary layer flow, example III

We analyze the behavior of the scalar variable p by plotting it over the outflow boundary. The scaled (LSFEM Dif. and LSFEM Tot.) and the stabilized (LSFEM Dif. Sta. and LSFEM Tot. Sta.) LSFEMs have been compared with the standard Galerkin FEM (Std. Galerkin) for $\varepsilon = 1e-4$ and $\varepsilon = 1e-6$ in Fig. 3.6 and 3.7, respectively. Two grids have been used for every solution in order to show that the grid-independent results are obtained. The stabilization parameter is $\delta_k = 1$ for all test cases. Both the scaled and the stabilized LSFEM formulations produce similar results. The LSFEM results do not possess spurious oscillations of the standard Galerkin FEM for all values of the ε used here. However, there are some under/over shoots near the discontinuity region, very close to the circular cylinder, in the domain.

3.5. Summary and Conclusion

We have shown that, by adding the extra curl equations to the Poisson equation optimal error estimates for both the scalar and the flux variables can be obtained. These statements are supported by our numerical investigation in section 3.2.5 which are in accordance with the theoretical a priori error estimates. Robust grid-independent multigrid performance is obtained only for the grad-div-curl system.

Solving the diffusion-reaction equations, optimal error estimates for the scalar and the vector variables are achieved with the grad-div-curl system. We observed a perfect grid-independent multigrid solver behavior for the Q_1 and the Q_2 finite elements. Moreover, we observed that the number of iterations is smaller for larger values of the reaction coefficient k.

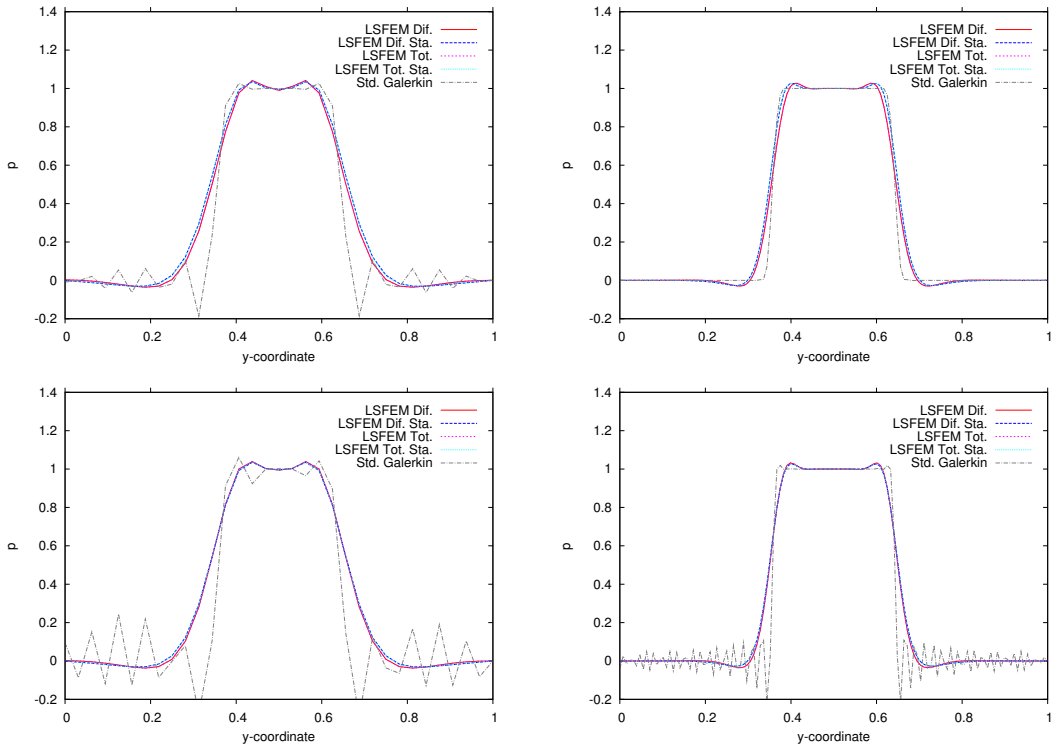


Figure 3.6: Outflow profile for the Q_1 elements at (top): $\varepsilon = 1e-4$ and (bottom): $\varepsilon = 1e-6$; on (left): a 32×32 grid; and (right): a 128×128 grid, example III

A new first-order formulation based on the total flux is introduced for the advection-diffusion equation. The uniqueness and a priori error estimates have been derived and proved. We have shown through many numerical examples that the scaled LSFEM performs well even for the advection-dominated flows. The scaled LSFEM is very robust for the whole range of flow regimes. In addition, unlike the stabilized LSFEM which its performance highly depends on the proper choice of the stabilization parameter, the scaled LSFEM is parameter-independent and performs well for all problems studied here.

Both of the scaled LSFEM systems for the advection-diffusion equation provide optimal L^2-error estimates for the scalar variable p even in the highly advection-dominated flows. The total flux results of the scaled LSFEM show optimal L^2-error convergence of the fluxes for all flow regimes. For the diffusive flux formulation of the scaled LSFEM, the convergence rates of the fluxes degraded in advection-dominated flows. These observations are valid for both the Q_1 and the Q_2 finite elements. On the other hand, the stabilized LSFEM is unable to produce optimal convergence rates for the scalar and the vector variables.

The scaled LSFEM performs reasonably good for the interior and the boundary layer problems at high Peclet numbers, further revealing the fact that no stabilization is required for the solution of the advection-diffusion flows with LSFEM.

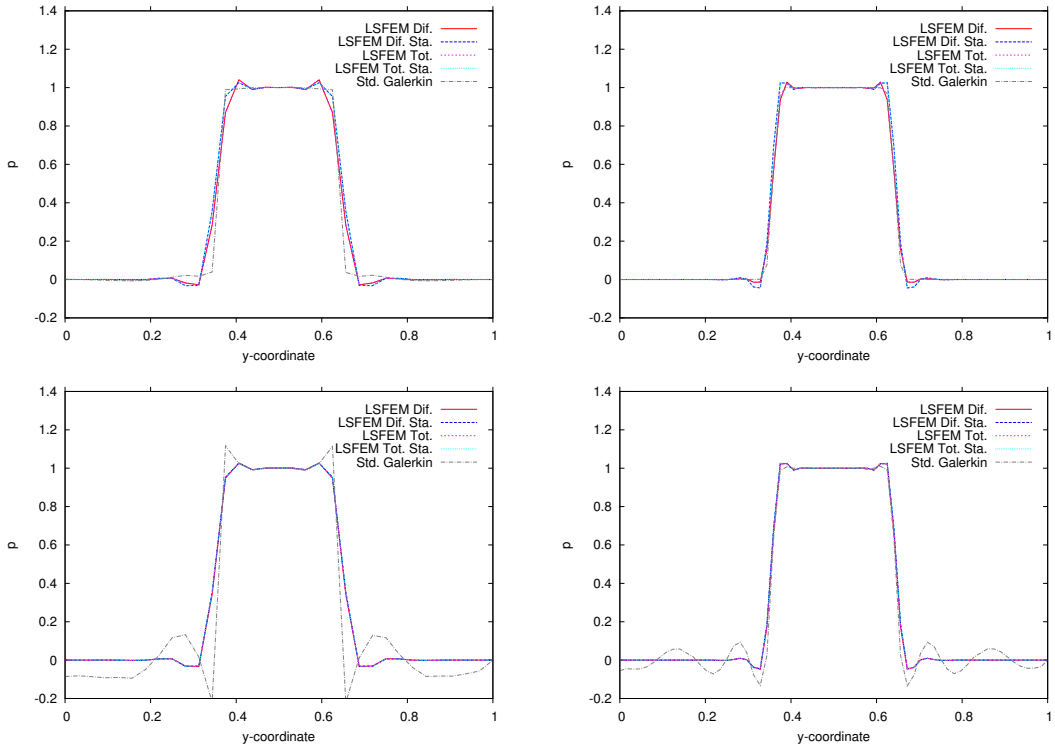


Figure 3.7: Outflow profile for the Q_2 elements at (top): $\varepsilon = 1e-4$ and (bottom): $\varepsilon = 1e-6$; on (left): a 32×32 grid; and (right): a 64×64 grid, example III

4

Incompressible Navier-Stokes Equations

We solve two first-order formulations of the stationary incompressible Navier-Stokes equations based on the least-squares finite element method. These first-order formulations have different properties and offer accurate approximations of some physical quantities of interest, e.g. the vorticity and the stress in fluid flow applications. For the discrete systems, we use a conjugate gradient (CG) solver accelerated with a geometric multigrid preconditioner for the complete system. In addition, we employ a Krylov space smoother inside of the multigrid which allows a parameter-free smoothing. Combining this linear solver with the Newton linearization, we construct a very robust and efficient solver. We use quadratic finite elements to enhance the mass conservation of the least-squares method for the inflow-outflow problems and to obtain highly accurate results. We demonstrate the advantages of using the higher order finite elements and the grid independent solver behavior through the solution of three stationary laminar flow problems of benchmarking character.

4.1. Introduction

In this chapter, we solve the incompressible Navier-Stokes (NS) equations with the LSFEM. Direct application of the LSFEM to the second-order NS equations requires the use of quite impractical C^1 finite elements [52]. Therefore, we introduce two equivalent first-order systems. The first formulation under consideration is the well-known div-curl-grad first-order velocity-vorticity-pressure (V-V-P) formulation. Here all unknowns are approximated in $H^1(\Omega)$. The second formulation is a div-grad first-order system resulting in a three-field formulation with stresses, velocities, and pressure as unknowns. This S-V-P formulation is approximated in $H^1(\Omega) \times H^1(\Omega) \times L^2(\Omega)$.

The discrete V-V-P and S-V-P LSFEM systems are symmetric and positive definite [52]. This permits the use of the conjugate gradient (CG) method and efficient multigrid solvers for the solution of the discrete systems. In order to improve the efficiency of the solution method, the multigrid and the Krylov subspace method, here CG, can be combined with two different strategies. The first strategy is to use the multigrid as a preconditioner for the Krylov method [28]. The advantage of this scheme is that the Krylov method reduces the error in eigenmodes that are not being effectively reduced by multigrid. The second strategy is to employ Krylov methods as multigrid smoother. The Krylov methods appropriately determine the size of the solution updates at each smoothing step [76]. This leads to smoothing sweeps which, in contrast to the standard SOR or Jacobi smoothing, are free from predefined damping parameters.

Heys et al. studied the LSFEM solution of the Stokes equation [28] and the NS equations [34–36] with an algebraic multigrid preconditioned CG method. A geometric multigrid preconditioned CG solver was used by Ranjan and Reddy [64] for the Spectral/hp LSFEM solution of the NS equations. They demonstrated superior convergence of the multigrid solver compared to the Jacobi preconditioning. More interestingly, Köster [42] and Wobker [76] used preconditioned BiCGStab

as smoother in a geometric multigrid method as well as an outer solver around it to solve the Poisson equation with standard Galerkin finite element method. They reported higher numerical stability and lower total costs of the solution process compared to the standalone multigrid or BiCGStab solvers.

We extend the multigrid-preconditioned CG (MPCG) solver to the V-V-P and S-V-P NS systems. In addition, we use a CG pre/post-smoother to obtain efficient and parameter-free smoothing sweeps. We investigate the performance of the MPCG solver for a wide range of parameters. We demonstrate a robust and grid independent behavior for the solution of different flow problems with both linear and quadratic finite elements for the V-V-P [46] and the S-V-P [50, 69] formulations.

Despite the advantages of the LSFEM, the lack of local mass conservation of this method is one of its drawbacks. Different strategies have been employed to overcome this deficiency. For very recent techniques and also an overview of the previous efforts we refer to the works by Bochev et al. [56, 57]. One remedy for 2D problems is to use higher order finite elements [40]. Weighting the continuity equation more strongly [37] is another well-known method to recover mass conservation. We analyze both of these methods in this work. We show, through the Poiseuille flow and the flow around cylinder problems, that quadratic finite elements satisfy the mass conservation to a great extent without the need to further weight the continuity equation.

The classical least-squares methods yield poor approximation quality when using lower-order finite elements, see e.g. [70]. The lower-order LSFEMs provide a suboptimal performance. Indeed, there are some contributions with lower-order elements with a satisfying performance in fluid dynamics or solid mechanics, see e.g. [13] and [70], but these approaches are modified least-squares formulations, which may introduce other drawbacks. An illustrative example for the poor solution of lower-order interpolation was given in [61] for the driven cavity problem. A possible solution for this problem, see e.g. [62] and [61], is the application of higher-order interpolation combinations. We show that accurate results can be obtained with the both LSFEM formulations provided that higher order finite elements are used. We demonstrate this with a quantitative analysis of the flow around cylinder and the lid-driven cavity problems.

4.2. Governing Equations

The incompressible NS equations for a stationary flow are given by

$$\begin{cases} \mathbf{u}\cdot\nabla\mathbf{u} - \nabla\cdot(2\nu\mathbf{D}(\mathbf{u})) + \nabla p = \mathbf{f} & \text{in } \Omega \\ \nabla\cdot\mathbf{u} = 0 & \text{in } \Omega \\ \mathbf{u} = \mathbf{g}_D & \text{on } \Gamma_D \\ \mathbf{n}\cdot\sigma = \mathbf{g}_N & \text{on } \Gamma_N \end{cases} \tag{4.2.1}$$

along with the zero mean pressure constraint

$$\int_\Omega p = 0 \tag{4.2.2}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain, $\mathbf{u}$ is the velocity, p is the normalized pressure $p = P/\rho$, $\nu = \mu/\rho$ is the kinematic viscosity, σ is the Cauchy stress tensor, $\mathbf{f}$ is the source term, $\mathbf{g}_D$ is the value of the Dirichlet boundary conditions on the Dirichlet boundary Γ_D, $\mathbf{g}_N$ is the prescribed traction on the Neumann boundary Γ_N, $\mathbf{n}$ is the outward unit normal on the boundary, $\Gamma = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. The kinematic viscosity and the density of the fluid are assumed to be constant. The symmetric part of the deformation tensor is defined as

$$\mathbf{D}(\mathbf{u}) = \frac{1}{2}\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right). \tag{4.2.3}$$

The first equation in (4.2.1) is the momentum equation where velocities $\mathbf{u} = [u, v]^T$ and pressure p are the unknowns and the second equation represents the continuity equation.

The straightforward application of the LSFEM to the second-order NS equations requires C^1 finite elements [52]. To avoid the practical difficulties in the implementation of such FEMs, we first recast the second-order equation to a system of first-order equations. Another important reason for not using the straightforward LSFEM is that the resulting system matrix will be ill-conditioned [55].

4.3. Vorticity-Velocity-Pressure Formulation

A common strategy to reformulate the second-order NS equations to an equivalent first-order system is to introduce the vorticity, ω, as a new variable [52]. In two-dimensional problems the vorticity is a scalar and defined as

$$\omega = \nabla \times \mathbf{u}. \tag{4.3.1}$$

Using the NS equations (4.2.1) and the vorticity equation (4.3.1) we obtain the first-order Vorticity-Velocity-Pressure (V-V-P) system of equations

$$\begin{cases} \mathbf{u}\cdot\nabla\mathbf{u} + \nabla p + \nu\nabla\times\omega = \mathbf{f} & \text{in } \Omega \\ \nabla\cdot\mathbf{u} = 0 & \text{in } \Omega \\ \omega - \nabla\times\mathbf{u} = 0 & \text{in } \Omega \\ \mathbf{u} = \mathbf{g}_D & \text{on } \Gamma_D \\ \mathbf{n}\cdot\sigma = \mathbf{g}_N & \text{on } \Gamma_N. \end{cases} \tag{4.3.2}$$

To obtain the V-V-P system (4.3.2), we use the following vector identity

$$\nabla\times\nabla\times\mathbf{u} = -\triangle\mathbf{u} + \nabla(\nabla\cdot\mathbf{u}) \tag{4.3.3}$$

and the incompressibility constraint $\nabla\cdot\mathbf{u} = 0$. It is easy to show that the first-order V-V-P equations and the NS equations are equivalent [33]. Therefore, based on the analysis provided in the book by Jiang [33] and by Ouazzi [47], we impose no extra boundary conditions for the vorticity.

The classical V-V-P formulation has been investigated by many authors, for instance by Bochev [5], Jiang [33] and Bochev and Gungburger [52] and with further modifications by Heys et al. [30, 35, 36]. We use the classical V-V-P formulation in our work and develop a multigrid-preconditioned CG solver for the solution of the discrete systems.

4.3.1. Linearization of the Convective Terms

The nonlinear convective term $\mathbf{u}\cdot\nabla\mathbf{u}$ in the momentum equations, first equation in (4.3.2), needs to be linearized. As it was discussed in chapter 2, section 2.4, we linearize the convective term before applying the least-squares technique. We adapt the *nonlinear basic iteration* algorithm 2.1 to solve the V-V-P system of equations. Therefore, the final solution is obtained from the following iterations

$$U^{n+1} = U^n - \beta\,\delta U^n \tag{4.3.4}$$

where $U = (\mathbf{u}, p, \omega)^T$, β is the damping parameter, and the superscripts n and $n+1$ denote the values corresponding to the previous and the current iterations. We obtain the solution update δU^n from the following equation

$$\tilde{\mathcal{J}}_{\text{vvp}}(U^n)\,\delta U^n = \mathbf{d}^n \tag{4.3.5}$$

where $\mathbf{d}^n$ is the residual vector defined by

$$\mathbf{d}^n = \begin{pmatrix} \mathbf{u}^n \cdot \nabla \mathbf{u}^n + \nabla p^n + \nu \nabla \times \omega^n - \mathbf{f} \\ \nabla \cdot \mathbf{u}^n \\ \omega^n - \nabla \times \mathbf{u}^n \end{pmatrix}. \tag{4.3.6}$$

Depending on the choice of the linearization technique, $\tilde{\mathcal{T}}_{\text{vvp}}(U^n)\,\delta U^n$ is approximated for the Newton method as follows

$$\mathcal{D}\mathcal{T}_{\text{vvp}}(U^n)[\delta U^n] = \begin{pmatrix} \mathbf{u}^n \cdot \nabla \delta \mathbf{u}^n + \delta \mathbf{u}^n \cdot \nabla \mathbf{u}^n + \nabla \delta p^n + \nu \nabla \times \delta \omega^n \\ \nabla \cdot \delta \mathbf{u}^n \\ \delta \omega^n - \nabla \times \delta \mathbf{u}^n \end{pmatrix}. \tag{4.3.7}$$

The linear operator for the fixed point method is simply obtained by removing the term $\delta \mathbf{u}^n \cdot \nabla \mathbf{u}^n$ from the equation (4.3.7).

Solving equation (4.3.5) for δU^n, we update the solution U^{n+1} in every nonlinear iteration. We terminate the nonlinear iterations when the relative errors, in the Euclidean norm, of the unknowns drop below a certain tolerance ε

$$\frac{\left\| U^{n+1} - U^n \right\|_2}{\left\| U^{n+1} \right\|_2} < \varepsilon. \tag{4.3.8}$$

The nonlinear iterations are further controlled by monitoring the Euclidean norm of the residual vector

$$\left\| \mathbf{d}^n \right\|_2 < \varepsilon_d \tag{4.3.9}$$

where ε_d is an error tolerance.

4.3.2. Continuous Least-Squares Principles

We define the least-squares variational problem based on the residuals of the system of equations (4.3.5). For this system, we define the following L^2-norm least-squares energy functionals

$$\begin{aligned} \mathcal{J}(\mathbf{v}, q, \xi; \mathbf{d}^n) = & \left\| \mathbf{u}^n \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}^n + \nabla q + \nu \nabla \times \xi - \mathbf{d}_1^n \right\|_0^2 \\ & + \alpha \left\| \nabla \cdot \mathbf{v} - d_2^n \right\|_0^2 + \left\| \xi - \nabla \times \mathbf{v} - d_3^n \right\|_0^2 \quad \forall (\mathbf{v}, q, \xi) \in \mathbf{V} \end{aligned} \tag{4.3.10}$$

and

$$\begin{aligned} \mathcal{J}_\nu(\mathbf{v}, q, \xi; \mathbf{d}^n) = & \frac{1}{\nu} \left\| \mathbf{u}^n \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}^n + \nabla q + \nu \nabla \times \xi - \mathbf{d}_1^n \right\|_0^2 \\ & + \alpha \left\| \nabla \cdot \mathbf{v} - d_2^n \right\|_0^2 + \left\| \xi - \nabla \times \mathbf{v} - d_3^n \right\|_0^2 \quad \forall (\mathbf{v}, q, \xi) \in \mathbf{V} \end{aligned} \tag{4.3.11}$$

where $\mathbf{V}$ is the space of admissible functions

$$\mathbf{V} = \left\{ (\mathbf{v}, q, \xi) \in \mathbf{H}^1_{g,D}(\Omega) \times H^1(\Omega) \cap L^2_0(\Omega) \times H^1(\Omega) \right\}. \tag{4.3.12}$$

and $\mathbf{d}_1^n$, d_2^n and d_3^n are the components of the residual vector $\mathbf{d}^n$. α is a scaling parameter aimed to improve the mass conservation of the LSFEM formulation [37, 58, 59]. The $\mathcal{J}_\nu$ functional, hereafter referred to as the weighted functional W-LSFEM, is obtained by scaling the momentum balance equations with the inverse kinematic viscosity. Such weighting parameters have been traditionally used by many other investigators in the LSFEM literature mostly for analysis purposes [52]. Here, our main concern is the investigation of the effects of weighting parameters on the overall accuracy and robustness of the Navier-Stokes LSFEM.

The minimization problem associated with the least-squares functionals in (4.3.10) and (4.3.11) is to find $(\delta \mathbf{u}, \delta p, \delta \omega) \in \mathbf{V}$ such that

$$(\delta \mathbf{u}, \delta p, \delta \omega) = \underset{(\mathbf{v}, q, \xi) \in \mathbf{V}}{\operatorname{argmin}} \mathcal{J}(\mathbf{v}, q, \xi; f) \tag{4.3.13}$$

where the functional $\mathcal{J}$ in equation (4.3.13) refers to both the standard and the weighted functionals.

The variational problem based on the optimality condition of the minimization problem (4.3.13) is to find $(\delta\mathbf{u}, \delta p, \delta\omega) \in \mathbf{V}$ such that

$$\mathcal{A}(\delta\mathbf{u}, \delta p, \delta\omega\,;\, \mathbf{v}, q, \xi) = \mathcal{F}(\mathbf{v}, q, \xi) \quad \forall (\mathbf{v}, q, \xi) \in \mathbf{V} \tag{4.3.14}$$

where $\mathcal{A}$ is a bilinear form defined on $\mathbf{V} \times \mathbf{V} \to \mathbb{R}$

$$\begin{aligned}\mathcal{A}(\delta\mathbf{u}, \delta p, \delta\omega\,;\, \mathbf{v}, q, \xi) :=& \alpha(\nabla\cdot\delta\mathbf{u}, \nabla\cdot\mathbf{v}) + (\delta\omega - \nabla\times\delta\mathbf{u}, \xi - \nabla\times\mathbf{v}) \\ &+ (\mathbf{u}^n\cdot\nabla\delta\mathbf{u} + \delta\mathbf{u}\cdot\nabla\mathbf{u}^n, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi) \\ &+ (\nabla\delta p + \nu\nabla\times\delta\omega, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi)\end{aligned} \tag{4.3.15}$$

and $\mathcal{F}$ is a linear form defined on $\mathbf{V} \to \mathbb{R}$

$$\mathcal{F}(\mathbf{v}, q, \xi) := (\mathbf{d}_1^n, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi) + \alpha(d_2^n, \nabla\cdot\mathbf{v}) + (d_3^n, \xi - \nabla\times\mathbf{v})\,. \tag{4.3.16}$$

For clarity, the bilinear form of the weighted functional (4.3.11) is subscripted with ν as follows

$$\begin{aligned}\mathcal{A}_\nu(\delta\mathbf{u}, \delta p, \delta\omega\,;\, \mathbf{v}, q, \xi) :=& \alpha(\nabla\cdot\delta\mathbf{u}, \nabla\cdot\mathbf{v}) + (\delta\omega - \nabla\times\delta\mathbf{u}, \xi - \nabla\times\mathbf{v}) \\ &+ \frac{1}{\nu}(\mathbf{u}^n\cdot\nabla\delta\mathbf{u} + \delta\mathbf{u}\cdot\nabla\mathbf{u}^n, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi) \\ &+ \frac{1}{\nu}(\nabla\delta p + \nu\nabla\times\delta\omega, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi)\,.\end{aligned} \tag{4.3.17}$$

The corresponding linear form for the weighted functional reads

$$\mathcal{F}_\nu(\mathbf{v}, q, \xi) := \frac{1}{\nu}(\mathbf{d}_1^n, \mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi) + \alpha(d_2^n, \nabla\cdot\mathbf{v}) + (d_3^n, \xi - \nabla\times\mathbf{v})\,. \tag{4.3.18}$$

4.3.3. Operator form of the Problem

To analyze the properties of the least-squares problem, let us write

$$\mathcal{A}(\delta\mathbf{u}, \delta p, \delta\omega\,;\, \mathbf{v}, q, \xi) = \big(\mathsf{L}(\delta\mathbf{u}, \delta p, \delta\omega), \mathsf{L}(\mathbf{v}, q, \xi)\big) \tag{4.3.19}$$

where L is the operator given by

$$\mathsf{L} = \begin{pmatrix} 0 & \nabla & \nu\nabla\times \\ \sqrt{\alpha}\nabla\cdot & 0 & 0 \\ -\nabla\times & 0 & I \end{pmatrix}. \tag{4.3.20}$$

It should be noted that the nonlinear terms are omitted for simplicity. Restricting to the $C_0^\infty(\Omega)$ functions, we rewrite the bilinear form (4.3.19) as

$$\mathcal{A}(\delta\mathbf{u}, \delta p, \delta\omega\,;\, \mathbf{v}, q, \xi) = \big(\mathsf{L}^*\mathsf{L}(\delta\mathbf{u}, \delta p, \delta\omega), (\mathbf{v}, q, \xi)\big) \tag{4.3.21}$$

where L^* is the formal adjoint of L, and the least-squares operator $\mathsf{L}^*\mathsf{L}$ is given by

$$\begin{aligned}\mathsf{L}^*\mathsf{L} &= \begin{pmatrix} 0 & -\sqrt{\alpha}\nabla & -\nabla\times \\ -\nabla\cdot & 0 & 0 \\ \nu\nabla\times & 0 & I \end{pmatrix}\begin{pmatrix} 0 & \nabla & \nu\nabla\times \\ \sqrt{\alpha}\nabla\cdot & 0 & 0 \\ -\nabla\times & 0 & I \end{pmatrix} \\ &= \begin{pmatrix} -\alpha\nabla\nabla\cdot+\nabla\times\nabla\times & 0 & -\nabla\times \\ 0 & -\nabla\cdot\nabla & 0 \\ -\nabla\times & 0 & I+\nu^2\nabla\times\nabla\times \end{pmatrix}.\end{aligned} \tag{4.3.22}$$

We repeat the same procedure (the derivation is skipped) to obtain the following least-squares operator $\mathsf{L}_\nu^*\mathsf{L}_\nu$ corresponding to the weighted formulation

$$\mathsf{L}_\nu^*\mathsf{L}_\nu = \begin{pmatrix} -\alpha\nabla\nabla\cdot+\nabla\times\nabla\times & 0 & -\nabla\times \\ 0 & -\frac{1}{\nu}\nabla\cdot\nabla & 0 \\ -\nabla\times & 0 & I+\nabla\times\nabla\times \end{pmatrix}. \tag{4.3.23}$$

The resulting system matrices, from equations (4.3.22) and (4.3.23), are symmetric and positive definite. So, after discretization, we are able to use the CG method to efficiently solve the system of equations. In addition, both of the least-squares systems are differentially diagonal dominant. This property, combined with the use of higher order finite elements, leads to efficient multigrid solver performance [35]. Our aim is to design an efficient solver which exploits the properties of the least-squares system with respect to both the CG and the multigrid methods. Therefore, we use CG as the main solver and accelerate it with the multigrid preconditioning, which is the previously mentioned MPCG solver.

Also, we use CG as pre/post-smoother which appropriately determines the size of the solution updates at each smoothing step [76]. Therefore, the CG smoothing leads to efficient and particularly parameter-free smoothing sweeps. In addition, we accelerate the smoothing process by using a SSOR preconditioner, which in this context requires no damping parameter in case of symmetric Gauß-Seidel sweeps.

It is worth noting that the possibility of using standard smoothers for the solution of the NS equations is another advantage of the LSFEM over the mixed Galerkin methods that require specially designed smoothers [71].

4.3.4. Discrete Least-Squares Principles

We introduce the approximation space $\mathbf{V_h}$, restrict our variational problem (4.3.14) to finite dimensional spaces, and consider the following approximation problem

$$\mathcal{A}_h(\delta\mathbf{u}_h,\delta p_h,\delta\omega_h\,;\,\mathbf{v}_h,q_h,\xi_h) = \mathcal{F}_h(\mathbf{v}_h,q_h,\xi_h) \quad \forall(\mathbf{v}_h,q_h,\xi_h)\in\mathbf{V_h}. \tag{4.3.24}$$

Choosing appropriate basis functions for the finite dimensional space $\mathbf{V_h}$, we obtain a discrete system of equations for the unknown LSFEM variables, namely $(\delta\mathbf{u}_h,\delta p_h,\delta\omega_h)$. Here we use conforming finite elements, therefore we set $\mathbf{V_h}\subset\mathbf{V}$. The typical LBB condition for mixed formulations leading to compatibility constraints for the involved function spaces is naturally satisfied for the LSFEM [52]. Therefore, we use equal-order finite elements in the V-V-P least-squares method.

4.4. Stress-Velocity-Pressure Formulation

In this section, we define another equivalent first-order system of equations for the incompressible NS equations. We define the Cauchy stress tensor as

$$\sigma = -p\mathbf{I} + 2\nu\mathbf{D}(\mathbf{u}), \tag{4.4.1}$$

which in combination with the NS equations (4.2.1) gives the following first-order Stress-Velocity-Pressure (S-V-P) system

$$\begin{cases} \mathbf{u}\cdot\nabla\mathbf{u} - \nabla\cdot\sigma = \mathbf{f} & \text{in } \Omega \\ \nabla\cdot\mathbf{u} = 0 & \text{in } \Omega \\ \sigma + p\mathbf{I} - 2\nu\mathbf{D}(\mathbf{u}) = 0 & \text{in } \Omega \\ \mathbf{u} = \mathbf{g}_D & \text{on } \Gamma_D \\ \mathbf{n}\cdot\sigma = \mathbf{g}_N & \text{on } \Gamma_N. \end{cases} \tag{4.4.2}$$

Compared to the V-V-P system, the S-V-P system has two more unknowns (a total of 6 unknowns in 2D). At first glance, the S-V-P system might look computationally more expensive than the V-V-P system. However, the properties of the S-V-P system matrix may differ from those of the V-V-P system. In addition, the solution of the S-V-P system directly provides the stress values which are of great practical interest.

The analysis of the Stokes equations using the S-V-P system has been considered by Bochev and Gunzburger [53] and by Chang et al. [10]. Cai et al. [9] has investigated the S-V-P formulation of the NS equations and proved optimal error estimates for the conforming finite element approximations. The numerical study of the S-V-P NS equations has been done by Ding et al. [19]. They used Jacobi-preconditioned conjugate gradient method for the solution of the discrete systems. Although, the S-V-P LSFEM has been considered in some other few works, see [41, 52], no extensive computational study has been done on the S-V-P NS system. To the best of our knowledge, the solution of the S-V-P system has never been done with the multigrid method in the literature. In this work we extend our MPCG solver for the solution of the discrete least-squares S-V-P systems [50, 69]. We extensively study the computational aspects of the system and compare the results with those of the V-V-P NS formulation.

4.4.1. Linearization of the Convective Terms

The convective flux in the momentum equations should be linearized in a similar manner to the V-V-P system linearization technique. The final solution based on the *nonlinear basic iteration* algorithm is obtained through the following iterations

$$U^{n+1} = U^n - \beta\,\delta U^n \tag{4.4.3}$$

where $U = (\mathbf{u}, p, \sigma)^T$. We obtain the solution update δU^n from the following equation

$$\tilde{\mathcal{T}}_{\text{svp}}(U^n)\,\delta U^n = \mathbf{d}^n \tag{4.4.4}$$

where $\mathbf{d}^n$ is the residual vector of the S-V-P system (4.4.2) defined by

$$\mathbf{d}^n = \begin{pmatrix} \mathbf{u}^n\cdot\nabla\mathbf{u}^n - \nabla\cdot\sigma^n - \mathbf{f} \\ \nabla\cdot\mathbf{u}^n \\ \sigma^n + p^n\mathbf{I} - 2\nu\mathbf{D}(\mathbf{u}^n) \end{pmatrix}. \tag{4.4.5}$$

Depending on the choice of the linearization technique, $\tilde{\mathcal{T}}_{\text{svp}}(U^n)\,\delta U^n$ is approximated for the Newton method as follows

$$\mathcal{D}\mathcal{T}_{\text{svp}}(U^n)[\delta U^n] = \begin{pmatrix} \mathbf{u}^n\cdot\nabla\delta\mathbf{u}^n + \delta\mathbf{u}^n\cdot\nabla\mathbf{u}^n - \nabla\cdot\delta\sigma^n \\ \nabla\cdot\delta\mathbf{u}^n \\ \delta\sigma^n + \delta p^n\mathbf{I} - 2\nu\mathbf{D}(\delta\mathbf{u}^n) \end{pmatrix}. \tag{4.4.6}$$

The linear operator for the fixed point method is simply obtained by removing the term $\delta\mathbf{u}^n \cdot \nabla\mathbf{u}^n$ from the equation (4.4.6).

Solving equation (4.4.4) for δU^n, we update the solution U^{n+1} in every nonlinear iteration. We use the same stopping criteria, see equation (4.3.8) and equation (4.3.9), for controlling the nonlinear iterations.

4.4.2. Continuous Least-Squares Principles

We define the L^2-norm least-squares energy functionals based on the residuals of the first-order system (4.4.4) as follows

$$\begin{aligned}\mathcal{J}(\mathbf{v},q,\tau;\mathbf{d}^n) = &\|\mathbf{u}^n\cdot\nabla\mathbf{v}+\mathbf{v}\cdot\nabla\mathbf{u}^n+\nabla\cdot\tau-\mathbf{d}_1^n\|_0^2 \\ &+\alpha\|\nabla\cdot\mathbf{v}-d_2^n\|_0^2+\frac{1}{\nu}\|\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})-d_3^n\|_0^2 \qquad \forall(\mathbf{v},q,\tau)\in\mathbf{V}\end{aligned} \tag{4.4.7}$$

where $\mathbf{V}$ is the space of admissible functions

$$\mathbf{V} = \left\{(\mathbf{v},q,\tau)\in\mathbf{H}^1_{g,D}(\Omega)\times H^1(\Omega)\cap L^2_0(\Omega)\times H_{g,N}(\mathrm{div},\Omega)\right\}, \tag{4.4.8}$$

The weighted functional is obtained by scaling the momentum balance equations with the inverse kinematic viscosity. The minimization problem associated with the least-squares functional (4.4.7) is to find $(\delta\mathbf{u},\delta p,\delta\sigma)\in\mathbf{V}$ such that

$$(\delta\mathbf{u},\delta p,\delta\sigma) = \underset{(\mathbf{v},q,\tau)\in\mathbf{V}}{\operatorname{argmin}}\,\mathcal{J}(\mathbf{v},q,\tau;f). \tag{4.4.9}$$

The variational problem based on the optimality condition of the minimization problem (4.4.9) is to find $(\delta\mathbf{u},\delta p,\delta\sigma)\in\mathbf{V}$ such that

$$\mathcal{A}(\delta\mathbf{u},\delta p,\delta\sigma\,;\,\mathbf{v},q,\tau) = \mathcal{F}(\mathbf{v},q,\tau) \quad \forall(\mathbf{v},q,\tau)\in\mathbf{V} \tag{4.4.10}$$

where $\mathcal{A}$ is a bilinear form defined on $\mathbf{V}\times\mathbf{V}\to\mathbb{R}$

$$\begin{aligned}\mathcal{A}(\delta\mathbf{u},\delta p,\delta\sigma\,;\,\mathbf{v},q,\tau) :=& \alpha\big(\nabla\cdot\delta\mathbf{u},\nabla\cdot\mathbf{v}\big) \\ &+\big(\mathbf{u}^n\cdot\nabla\delta\mathbf{u}+\delta\mathbf{u}\cdot\nabla\mathbf{u}^n-\nabla\cdot\delta\sigma\,,\,\mathbf{u}^n\cdot\nabla\mathbf{v}+\mathbf{v}\cdot\nabla\mathbf{u}^n-\nabla\cdot\tau\big) \\ &+\frac{1}{\nu}\big(\delta\sigma+\delta p\mathbf{I}-2\nu\mathbf{D}(\delta\mathbf{u})\,,\,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})\big)\end{aligned} \tag{4.4.11}$$

and $\mathcal{F}$ is a linear form defined on $\mathbf{V}\to\mathbb{R}$

$$\mathcal{F}(\mathbf{v},q,\tau) := \big(\mathbf{d}_1^n\,,\,\mathbf{u}^n\cdot\nabla\mathbf{v}+\mathbf{v}\cdot\nabla\mathbf{u}^n-\nabla\cdot\tau\big)+\alpha\big(d_2^n,\nabla\cdot\mathbf{v}\big)+\frac{1}{\nu}\big(d_3^n,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})\big)\,. \tag{4.4.12}$$

4.4.3. Operator form of the Problem

Restricting to the $C_0^\infty(\Omega)$ functions, the variational problem (4.4.10) can be written as

$$\mathcal{A}(\delta\mathbf{u},\delta p,\delta\sigma\,;\,\mathbf{v},q,\tau) = \big(\mathsf{L}^*\mathsf{L}(\delta\mathbf{u},\delta p,\delta\sigma),(\mathbf{v},q,\tau)\big) \tag{4.4.13}$$

where L^* is the formal adjoint of L, and the least-squares operator $\mathsf{L}^*\mathsf{L}$ after neglecting the nonlinear terms is given by

$$\begin{aligned}\mathsf{L}^*\mathsf{L} &= \begin{pmatrix} \mathbf{0} & -\sqrt{\alpha}\nabla & \frac{2\nu}{\sqrt{\nu}}\nabla\cdot \\ 0 & 0 & \frac{1}{\sqrt{\nu}}\mathbf{I}: \\ \nabla & 0 & \frac{1}{\sqrt{\nu}} \end{pmatrix} \begin{pmatrix} \mathbf{0} & 0 & -\nabla\cdot \\ \sqrt{\alpha}\nabla\cdot & 0 & 0 \\ -\frac{2\nu}{\sqrt{\nu}}\mathbf{D}(\cdot) & \frac{1}{\sqrt{\nu}}\mathbf{I} & \frac{1}{\sqrt{\nu}} \end{pmatrix} \\ &= \begin{pmatrix} -4\nu\nabla\cdot\mathbf{D}(\cdot) - \alpha\nabla\nabla\cdot & 2\nabla & 2\nabla\cdot \\ -2\nabla\cdot & \frac{2}{\nu} & \frac{1}{\nu}\mathbf{I}: \\ -2\mathbf{D}(\cdot) & \frac{1}{\nu}\mathbf{I} & \frac{1}{\nu} - \nabla\nabla\cdot \end{pmatrix}. \end{aligned} \tag{4.4.14}$$

The system matrix of the S-V-P formulation (4.4.14) is symmetric similar to the V-V-P system. However, in contrast to the vorticity-based operator forms (4.3.22) and (4.3.23) the stress-based formulation leads to the system matrices which are not differentially diagonal dominant. This property is a key factor in designing efficient multigrid solvers when higher order finite elements are used. In the numerical results section, we show that the multigrid solver for the V-V-P LSFEM outperforms that of the S-V-P LSFEM especially when the quadratic finite elements are used for all variables.

4.4.4. Discrete Least-Squares Principles

We introduce the approximation space $\mathbf{V_h}$, restrict our variational problem (4.4.10) to finite dimensional spaces, and consider the following approximation problem

$$\mathcal{A}_h(\delta\mathbf{u}_h, \delta p_h, \delta\sigma_h \,;\, \mathbf{v}_h, q_h, \tau_h) = \mathcal{F}_h(\mathbf{v}_h, q_h, \tau_h) \quad \forall(\mathbf{v}_h, q_h, \tau_h) \in \mathbf{V_h}. \tag{4.4.15}$$

Choosing appropriate basis functions for the finite dimensional space $\mathbf{V_h}$, we obtain a discrete system of equations for the unknown LSFEM variables, namely $(\delta\mathbf{u}_h, \delta p_h, \delta\sigma_h)$. Here we use conforming finite elements, therefore we set $\mathbf{V_h} \subset \mathbf{V}$. We use equal-order finite elements in the S-V-P least-squares method.

4.5. Numerical Results and Discussions

We analyze the following problems: First, we design an analytic polynomial solution to study the error estimates and the convergence rates of the least-squares primary variables, the velocity and the pressure, and the introduced vorticity and stress variables. Next, we study three steady state flow problems, i.e. the Poiseuille flow, the flow around cylinder and the lid-driven cavity flow, with both LSFEM formulations. We compare our results with the results of the mixed finite element method (MFEM) produced by FEATFLOW (see *www.featflow.de* and [17]), and with available benchmark solutions in the literature. We use the following configurations for all flow simulations unless it is specifically stated:

1. Q_1 (linear) or Q_2 (quadratic) finite elements for all unknowns
2. The MPCG solver for the solution of the linearized system of equations
3. Relative stopping criterion $\varepsilon =$ 1E-6
4. Absolute stopping criterion $\varepsilon_d =$ 1E-6.

4.5.1. Polynomial Solution

In this example, we analyze the L^2-error estimates and the convergence rates of the least-square solutions. The solution domain is a unit square $-1 \leq x, y \leq 1$. We assume no body force, i.e. $\mathbf{f} = 0$. The Reynolds number is $\mathrm{Re} = 1$. We take the following polynomial solutions for the velocities and the pressure

$$\begin{aligned} u(x,y) &= \ 2x^2(1-x)^2\left(y(1-y)^2 - y^2(1-y)\right) \\ v(x,y) &= -2y^2(1-y)^2\left(x(1-x)^2 - x^2(1-x)\right) \\ p(x,y) &= \ c(x^3 - y^3) \end{aligned} \tag{4.5.1}$$

where c is a constant. Using the definition of the velocity and the pressure functions in equation (4.5.1), the vorticity equation (4.3.1) and the stress equation (4.4.1), we obtain the following expressions for the additional variables

$$\begin{aligned} \omega(x,y) &= \ -2x^2(1-x)^2\left(6y^2 - 6y + 1\right) - 2y^2(1-y)^2\left(6x^2 - 6x + 1\right) \\ \sigma_{xx}(x,y) &= \ -c\left(x^3 - y^3\right) + 8xy(2x-1)(2y-1)(x-1)(y-1) \\ \sigma_{xy}(x,y) &= \ 2x^2(1-x)^2\left(6y^2 - 6y + 1\right) - 2y^2(1-y)^2\left(6x^2 - 6x + 1\right). \end{aligned} \tag{4.5.2}$$

We summarize the results in Table 4.1 and Table 4.2 for the V-V-P and the S-V-P formulations, respectively. We observe asymptotic convergence toward the analytical solution for all unknowns. The pressure error for the S-V-P system is not optimal as compared to the V-V-P system. In both of the LSFEM methods, the velocity error increases with the increase in the pressure norm.

Table 4.1: V-V-P LSFEM, L^2 errors for different finite elements at $Re = 1$

Level	$\|\mathbf{u} - \mathbf{u}_h\|_0$	rate	$\|p - p_h\|_0$	rate	$\|\omega - \omega_h\|_0$	rate
Q_1 elements, $c = 1$						
6	4.1338E-05	3.99	2.6736E-04	4.00	2.3374E-04	4.01
7	1.0344E-05	4.00	6.6855E-05	4.00	5.8387E-05	4.00
8	2.5867E-06	4.00	1.6715E-05	4.00	1.4593E-05	4.00
Q_1 elements, $c = 10$						
6	1.4636E-04	3.97	2.6736E-03	4.00	2.1478E-03	4.02
7	3.6664E-05	3.99	6.6855E-04	4.00	5.3632E-04	4.00
8	9.1707E-06	4.00	1.6715E-04	4.00	1.3403E-04	4.00
Q_2 elements, $c = 1$						
6	2.8081E-07	8.00	1.2459E-06	8.00	1.2868E-06	7.97
7	3.5099E-08	8.00	1.5573E-07	8.00	1.6093E-07	8.00
8	4.3873E-09	8.00	1.9467E-08	8.00	2.0113E-08	8.00
Q_2 elements, $c = 10$						
6	2.8084E-07	8.00	1.2458E-05	8.00	1.5346E-06	9.09
7	3.5100E-08	8.00	1.5573E-06	8.00	1.7708E-07	8.67
8	4.3874E-09	8.00	1.9467E-07	8.00	2.1147E-08	8.37

4.5.2. Poiseuille Flow

We study a laminar Poiseuille flow at Reynolds number $Re = 100$ in a square domain of $\Omega = [0,1] \times [0,1]$. The two horizontal solid walls have no-slip boundary conditions. The inflow velocity boundary condition is given by

$$[u, v] = [y(1-y), 0]. \tag{4.5.3}$$

Table 4.2: S-V-P LSFEM, L^2 errors for different finite elements at $Re = 1$

Level	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|p-p_h\|_0$	rate	$\|\sigma_{xx}-\sigma_{xx,h}\|_0$	rate	$\|\sigma_{xy}-\sigma_{xy,h}\|_0$	rate
Q_1 elements, $c=1$								
6	6.7409E-04	2.39	1.1854E-02	1.90	1.0189E-02	1.91	5.2672E-03	2.07
7	2.6903E-04	2.51	6.0189E-03	1.97	5.1315E-03	1.99	2.4297E-03	2.17
8	9.6392E-05	2.79	2.4956E-03	2.41	2.1471E-03	2.39	9.1563E-04	2.65
Q_1 elements, $c=10$								
6	5.9967E-03	2.61	7.7006E-02	2.57	6.1438E-02	2.53	3.7395E-02	2.62
7	1.9747E-03	3.04	2.7978E-02	2.75	2.2669E-02	2.71	1.3642E-02	2.74
8	6.0428E-04	3.27	9.9265E-03	2.82	8.0952E-03	2.80	4.9735E-03	2.74
Q_2 elements, $c=1$								
6	7.6457E-07	6.84	8.2350E-05	4.17	4.7586E-05	4.74	1.3197E-05	6.14
7	1.0041E-07	7.61	1.9762E-05	4.17	1.0558E-05	4.51	2.1300E-06	6.20
8	1.1808E-08	8.50	4.8567E-06	4.07	2.5428E-06	4.15	3.4245E-07	6.22
Q_2 elements, $c=10$								
6	4.9232E-06	7.92	2.6937E-04	5.47	2.3102E-04	5.64	1.2292E-04	5.98
7	5.9994E-07	8.21	4.8704E-05	5.53	4.0245E-05	5.74	2.0428E-05	6.02
8	6.2968E-08	9.53	9.2245E-06	5.28	7.3693E-06	5.46	3.3480E-06	6.10

The exact pressure solution for this problem is

$$p(x) = 2\nu(L-x) \tag{4.5.4}$$

where $L = 1$ is the length of the domain.

Outflow Boundary Conditions

We analyze two different boundary conditions for the outflow boundary. The first one, which comes from the exact solution for the velocity, is to set the same boundary condition as the inflow, i.e. equation (4.5.3). This is an essential boundary condition, so it is treated in a strong manner by filtering the system matrix and the right-hand side vector appropriately. The other boundary condition is to prescribe the zero-normal stress on the outflow [39, 40, 64]. Therefore the outflow boundary condition reads

$$\sigma \cdot \mathbf{n} = (-p\mathbf{I} + \nu\nabla\mathbf{u}) \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{out} \tag{4.5.5}$$

where $\mathbf{n}$ is the outward unit vector normal to the outflow boundary Γ_{out}. Considering the fact that we have a vertical outflow section, $\mathbf{n} = [1\,,\,0]^T$, equation (4.5.5) simplifies to

$$\begin{cases} -p + \nu\dfrac{\partial u}{\partial x} = 0 \quad \text{on} \quad \Gamma_{out} \\ \nu\dfrac{\partial v}{\partial x} = 0 \quad \text{on} \quad \Gamma_{out}. \end{cases} \tag{4.5.6}$$

We incorporate the boundary conditions in equation (4.5.6) into the variational problem (in a weak manner) using the L^2-norm functionals acting on the outflow boundary. Therefore, for the V-V-P formulation the energy functional (4.3.10) changes to

$$\begin{aligned} \mathcal{J}(\mathbf{v},q,\xi;\mathbf{d}^n) = {} & \|\mathbf{u}^n\cdot\nabla\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{u}^n + \nabla q + \nu\nabla\times\xi - \mathbf{d}_1^n\|_0^2 + \alpha\|\nabla\cdot\mathbf{v} - \mathbf{d}_2^n\|_0^2 \\ & + \|\xi - \nabla\times\mathbf{v} - \mathbf{d}_3^n\|_0^2 + \|(-q\mathbf{I} + \nu\nabla\mathbf{v})\cdot\mathbf{n}\|_{0,\Gamma_{out}}^2 \quad \forall(\mathbf{v},q,\xi)\in\mathbf{V}. \end{aligned} \tag{4.5.7}$$

We use a similar treatment to add the boundary conditions to the weighted V-V-P functional (4.3.11) and to the S-V-P functional (4.4.7). However, adding the boundary term to the functional (4.4.7) requires extra regularity for the stress to be well defined in the $L^2(\Omega)$. Therefore, the space of admissible functions for the S-V-P problem, see (4.4.8), is modified to the following

$$\mathbf{V} = \left\{ (\mathbf{v}, q, \tau) \in \mathbf{H}^1_{g,D}(\Omega) \times H^1(\Omega) \cap L^2_0(\Omega) \times H_{g,N}(\mathrm{div}, \Omega) \cap \mathbf{H}^s(\Omega) \right\}, \tag{4.5.8}$$

where $s \leq 1$.

Remark 4.5.1. *In the case of zero-normal stress boundary conditions, we obtain the exact pressure, equation (4.5.4), from the LSFEM simulations. In other words, the zero stress boundary condition helps to make the pressure field unique. This is due to the fact that the pressure exists in the zero-normal stress boundary condition. However, when we apply Dirichlet velocity boundary conditions on the outflow, we fix the pressure in one point to make the pressure field unique.*

The computational grid contains one quadrilateral element on the first level, level 1, and finer grids are obtained based on the multilevel grid refinement. In the multilevel refinement, every element from the coarse grid is divided into four fine elements by connecting the midpoints of the opposite edges [43]. Then we use the hierarchy of the multilevel grids in our geometric multigrid preconditioner.

In order to investigate the mass conservation, we measure the Global Mass Conservation (GMC) in terms of the fractional change of mass flow rate, defined as

$$\mathrm{GMC} = \frac{\int_{\Gamma_i} \rho\,(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_i - \int_{\Gamma_o} \rho\,(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_o}{\int_{\Gamma_i} \rho\,(\mathbf{n}\cdot\mathbf{u})\,d\Gamma_i} \times 100 \tag{4.5.9}$$

where Γ_i is the inflow boundary of the domain and Γ_o is any vertical section between the inflow and the outflow boundaries, including the outflow.

Using Q_1 elements, we present the GMC values between the inflow and the different vertical cross-sections of the domain for the LSFEM and the W-LSFEM variations of the V-V-P formulation in Table 4.3 and Table 4.4, respectively. The results show that mass conservation improves with mesh refinement and is admittedly satisfied throughout the domain at finer grids. The GMC values are further reduced when a scaling parameter, $\alpha > 1$, is employed at each level. In addition, the results of the Dirichlet outflow boundary condition show slightly better mass conservation compared to those of the zero-normal stress boundary condition. The results of Table 4.3 and Table 4.4 show that on coarse grids less mass conservation is achieved with the W-LSFEM in comparison with the standard LSFEM. However, the difference between the two formulations becomes negligible with further grid refinement.

The S-V-P LSFEM results are presented in Table 4.5. The direct comparison of the S-V-P results of Table 4.5 with those of the V-V-P LSFEM in Table 4.3 and Table 4.4 show that the mass conservation is very well satisfied using the former formulation. Further, we observe that the GMC values obtained by the S-V-P formulation do not considerably change with α. In deed, a one-by-one comparison of the values between the two LSFEM methods reveals that the GMC values of the S-V-P method with $\alpha = 1$ are almost equal to the GMC values of the V-V-P method with $\alpha = 100$. Moreover, the type of the outflow boundary condition does not play a role in the mass conservation of the S-V-P method.

In the case of Q_2 finite elements, the obtained velocity field is exact and the GMC values are zero, up to the iteration error, everywhere in the domain. This is due to the fact that the Poiseuille flow's parabolic velocity field can be fully represented by the quadratic finite element. Therefore, we skip here the Q_2 element results for the both LSFEM formulations.

Table 4.3: *V-V-P LSFEM*, absolute values of the GMC in vertical cross-sections of the Poiseuille flow at $Re = 100$ with Q_1 elements

	$\alpha = 1.0$		$\alpha = 10$		$\alpha = 100$	
Level	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$
zero-normal stress						
5	1.540063	2.435800	0.520126	0.614301	0.406549	0.416399
6	0.447549	0.722513	0.134901	0.162917	0.101985	0.104672
7	0.138748	0.238623	0.036068	0.046048	0.025626	0.026484
8	0.044002	0.082468	0.009900	0.013754	0.006470	0.006803
Dirichlet boundary condition						
5	0.967960	0.858946	0.457747	0.446324	0.398668	0.397421
6	0.283028	0.250209	0.117297	0.114153	0.099898	0.099585
7	0.082833	0.074349	0.030338	0.029532	0.025029	0.024955
8	0.023621	0.022299	0.007861	0.007731	0.006275	0.006262

Table 4.4: *V-V-P W-LSFEM*, absolute values of the GMC in vertical cross-sections of the Poiseuille flow at $Re = 100$ with Q_1 elements

	$\alpha = 1.0$		$\alpha = 10$		$\alpha = 100$	
Level	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$
zero-normal stress						
5	2.057919	3.740288	0.627447	0.881192	0.414364	0.438957
6	0.722934	0.140562	0.172234	0.258356	0.105095	0.113889
7	0.215425	0.433627	0.044964	0.069519	0.026494	0.029075
8	0.057796	0.119313	0.011419	0.017906	0.006644	0.007322
Dirichlet boundary condition						
5	1.094678	1.100044	0.490521	0.485411	0.400358	0.399665
6	0.331548	0.348048	0.126444	0.127463	0.100521	0.100554
7	0.091683	0.098458	0.031871	0.032484	0.025170	0.025223
8	0.023839	0.025854	0.007953	0.008146	0.006291	0.006309

Table 4.5: *S-V-P LSFEM*, absolute values of the GMC in vertical cross-sections of the Poiseuille flow at $Re = 100$ with Q_1 elements

	$\alpha = 1.0$		$\alpha = 10$		$\alpha = 100$	
Level	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$	$x = 0.3$	$x = 0.8$
zero-normal stress						
5	0.399466	0.409600	0.391435	0.392230	0.390704	0.390778
6	0.100250	0.103484	0.097889	0.098142	0.097678	0.097699
7	0.025112	0.025998	0.024478	0.024557	0.024419	0.024426
8	0.006282	0.006509	0.006121	0.006143	0.006105	0.006107
Dirichlet boundary condition						
5	0.393767	0.393793	0.390901	0.390909	0.390652	0.390653
6	0.098572	0.098625	0.097737	0.097740	0.097664	0.097664
7	0.024657	0.024679	0.024436	0.024438	0.024416	0.024416
8	0.006165	0.006171	0.006109	0.006110	0.006104	0.006104

4.5.3. Flow around Cylinder

In this section, we present the simulation of a laminar steady state flow passed a circular cylinder. The geometry of the benchmark configuration consists of a simple channel of length 2.2 and height 0.41. At $[x,y] = [0.2, 0.2]$ a cylinder with diameter $D = 0.1$ is placed. The kinematic viscosity of the fluid is $\nu = 0.001$ and $Re = 20$ which is defined as

$$Re = \frac{U_{mean} D}{\nu} \tag{4.5.10}$$

where U_{mean} is the average velocity of the inflow stream and defined as

$$U_{mean} = \frac{2}{3} U_{max}. \tag{4.5.11}$$

We refer to [72] and *www.featflow.de/en/benchmarks.html* for further details concerning this benchmark.

Boundary Conditions

The horizontal upper and lower walls and the cylinder have no-slip boundary conditions. The inflow velocity boundary conditions are defined as

$$[u,v] = \left[\frac{1.2y(0.41-y)}{0.41^2}, 0 \right]. \tag{4.5.12}$$

For the outflow boundary, we impose the zero-normal stress boundary condition defined in equation (4.5.5).

We present the computational mesh of the coarsest level, level 1, in Figure 4.1. Correspondingly, Table 4.6 summarizes the information regarding the number of elements and the number of degrees of freedom.

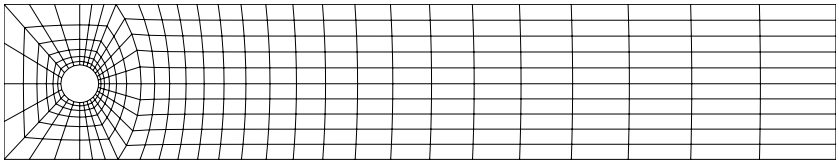


Figure 4.1: Flow around cylinder, computational grid of level 1

First, we analyze the accuracy of the LSFEM methods by calculating the lift and drag coefficients and the pressure drop across the cylinder. For the definition of these flow parameters one should refer to [72]. In the V-V-P method we obtain the stress, by post-processing, from the calculated pressure and velocities, see [45] for the definition. In the S-V-P method however, we calculate the stress as a separate variable and therefore we can directly calculate the body forces as

$$\begin{pmatrix} F_D \\ F_L \end{pmatrix} = \int_S \boldsymbol{\sigma}\mathbf{n}\, dS \tag{4.5.13}$$

where S denotes the surface of the cylinder and $\mathbf{n}$ is the normal vector of S. We present the lift and drag coefficients and the pressure drop across the cylinder at Reynolds number $\mathrm{Re} = 20$ for the V-V-P and the S-V-P formulations in Table 4.7 and Table 4.8, respectively. Here we present the W-LSFEM results of the V-V-P formulation. A comprehensive comparison between the scaled and the unscaled formulations can be found in our previous work [45]. The Q_2-element results

Table 4.6: Mesh information for the flow around cylinder problem, the number of elements (NEL) and the number of degrees of freedom.

Level	NEL	Degrees of freedom			
		V-V-P		S-V-P	
		Q_1	Q_2	Q_1	Q_2
1	346	1,564	5,896	2,346	8,844
2	1,384	5,896	22,864	8,844	34,296
3	5,536	22,864	90,016	34,296	135,024
4	22,144	90,016	357,184	135,024	535,776
5	88,576	357,184	1,422,976	535,776	2,134,464
6	354,304	1,422,976		2,134,464	

are much more accurate than the Q_1-element results, and those pertained to $\alpha = 1$ are in excellent agreement with the benchmark solutions. Departing from the $\alpha = 1$ case, the accuracy of the flow parameters is degraded especially for the Q_1 elements. The quadratic element results are less sensitive to the α variations compared to the linear element results. Using higher order finite elements, both methods show excellent convergence toward the reference solution. We observe a slight improvement for the predicted lift and drag coefficients of the S-V-P method over the V-V-P formulation.

We present the GMC values on the outflow (at $x = 2.2$) in Table 4.7 and Table 4.8. The Q_1-element V-V-P LSFEM results show that the mass conservation is not well satisfied, even for the fine grids, throughout the domain with $\alpha = 1$. Similar severe mass loss of the LSFEM, when piecewise linear finite elements are used, has been reported by Chang and Nelson [13], Deang and Gunzburger [37] and Bolton and Thatcher [58] for the Stokes flow and by Bolton and Thatcher [59] for the NS equations in the literature. The S-V-P method shows much better mass conservation as compared to the V-V-P method. However, by using $\alpha = 100$ mass conservation is also improved in the V-V-P method. Moreover, this weighting does not have a pronounced effect on the accuracy of the drag and lift coefficients and the pressure drop especially for the Q_2 element results.

Next, we analyze the performance of the MPCG solver for the solution of the scaled V-V-P and the S-V-P problems using the conforming nodal finite elements Q_1 and Q_2. A comprehensive comparison between the scaled and the unscaled V-V-P formulations can be found in our previous work [45]. We have shown that the former formulation behaves more robust than the later one with respect to the variations in α. Table 4.9 shows the number of nonlinear iterations and the corresponding averaged linear solver (MPCG solver) iterations for different levels. We observe a grid-independent convergence behavior and a constant number of iterations with grid refinement at each α for the V-V-P method. The optimal number of iterations is obtained at $\alpha = 1$. The number of the linear solver iterations increases when we depart from $\alpha = 1$. The S-V-P method shows a grid-independent solution behavior as well. However, in this case the solver requires more iterations to reach the same convergence criteria. An explanation could be that the S-V-P method, unlike the V-V-P method as has been shown by [45], does not yield a differentially diagonally dominant system matrix. It is shown in the literature that having this property is a key factor in obtaining efficient multigrid solver performance, see [35]. An important remark about the results of Table 4.9 is that the MPCG solver remains efficient for both low and high order finite elements.

Table 4.7: *V-V-P W-LSFEM*: flow parameters in the flow around cylinder at $\mathrm{Re} = 20$.

Level	Drag coefficient C_D	Lift coefficient C_L	Pressure drop $\triangle p$	GMC-value at $x = 2.2$
$Q_1, \alpha = 1$				
3	3.8910464	0.0023588	0.0771009	32.458835
4	4.8914483	0.0043424	0.1009819	12.828753
5	5.3687854	0.0086759	0.1124486	3.871370
6	5.5234496	0.0101034	0.1161682	1.025463
$Q_2, \alpha = 1$				
2	5.4639017	0.0092029	0.1148151	1.212050
3	5.5668223	0.0104928	0.1172298	0.114906
4	5.5779343	0.0106055	0.1174841	0.010779
5	5.5792792	0.0106169	0.1175144	0.001096
$Q_1, \alpha = 100$				
3	4.7949823	0.0474048	0.0893839	3.577138
4	5.2716595	0.0245279	0.1044718	1.053974
5	5.4769244	0.0132686	0.1128306	0.283655
6	5.5500584	0.0109745	0.0116144	0.073273
$Q_2, \alpha = 100$				
2	4.9640652	0.0047990	0.1050202	0.267959
3	5.4744825	0.0094397	0.1157215	0.035024
4	5.5620013	0.0104345	0.1172938	0.004049
5	5.5762393	0.0105925	0.1174909	0.000473
ref.: C_D = 5.57953523384, C_L = 0.010618948146, $\triangle p$ = 0.11752016697				

4.5.4. Lid-driven Cavity Flow

We simulate the regularized lid-driven cavity flow problem in this section. The flow domain is a unit square which has no-slip boundary conditions on the vertical and lower horizontal walls. The upper wall, the lid, has zero vertical velocity and a horizontal velocity, i.e., the lid velocity. Normally, the lid velocity is taken to be constant which leads to the singularities on the two upper corners of the domain. To remove these singularities, we prescribe a regularized horizontal velocity [8], defined as

$$\mathbf{u}_{lid} = [-16x^2(1-x)^2, 0]. \tag{4.5.14}$$

We fix the pressure in one point, $p = 0$, in the middle of the lower cavity wall.

In addition to the conventional local velocity profiles, we investigate two other global quantities as defined in [8]. The first one is the kinetic energy defined as

$$E = \frac{1}{2} \|\mathbf{u}_h\|^2_{0,\Omega} \tag{4.5.15}$$

and the other quantity is the enstrophy defined as follows

$$Z = \frac{1}{2} \|w_h\|^2_{0,\Omega} \tag{4.5.16}$$

where $\|\cdot\|_{0,\Omega}$ is the L^2 norm.

For this test case, we compare the least-squares FEM results with the reference solutions obtained from mixed finite element method, FEATFLOW solver [17]. Two pairs of LBB-stable mixed

Table 4.8: *S-V-P LSFEM*: flow parameters in the flow around cylinder at Re = 20 and α = 1.

Level	Drag coefficient C_D	Lift coefficient C_L	Pressure drop $\triangle p$	GMC-value at $x = 2.2$
Q_1				
3	4.6026594	0.0365479	0.0856446	3.896641
4	5.1716353	0.0210522	0.0103135	1.114501
5	5.4440131	0.0142939	0.1117922	0.299773
6	5.5415463	0.0117584	0.1152451	0.077866
Q_2				
2	5.4404341	0.0076287	0.1124827	0.228281
3	5.5588883	0.0101360	0.1165546	0.022791
4	5.5769755	0.0105355	0.1173265	0.003022
5	5.5792424	0.0106064	0.1174766	0.000556
ref.: C_D = 5.57953523384, C_L = 0.010618948146, $\triangle p$ = 0.11752016697				

Table 4.9: The number of nonlinear iterations and the corresponding averaged number of linear solver iterations for flow around cylinder at Re = 20, nonlinear and linear solver relative changes are kept below 1E-6 and 1E-3, respectively.

	V-V-P						S-V-P	
Level	Q_1			Q_2			Q_1	Q_2
α	1.0	10	100	1.0	10	100	-	-
3	8/4	8/4	8/6	6/5	6/6	6/10	7/19	6/12
4	8/4	8/4	8/6	6/5	6/6	6/9	7/17	6/12
5	7/4	7/4	8/6	6/5	6/6	6/9	7/17	6/12

finite elements are used for comparison, namely conforming $Q_2P_1^{disc}$ (see [7]) and nonconforming $\tilde{Q}_1Q_0$ see [63]. The discontinuous finite element P_1^{disc} consists of linear functions with the function value and both partial derivatives, located in the center of the quadrilateral, as its three local degrees of freedom. The $\tilde{Q}_1$ finite element consists of rotated bilinear shape functions. Local degrees of freedom are the mean values over the element edges. Finally, the Q_0 finite element has a piecewise constant shape function and its only degree of freedom is the function value in the center of the quadrilateral.

Here we use the same computational grid as for the Poiseuille flow problem. Table 4.10 shows the mesh information of the two LSFEM formulations with different finite element combinations.

We present the percent error of the horizontal velocity through the vertical centerline of the cavity for different *Re* in Tables 4.11, 4.12 and 4.13. We obtain the percent errors by comparing the W-LSFEM results with the converged reference solutions of the FEATFLOW solver [17] on a highly refined grid as follows

$$\text{percent error} = \frac{u_{\text{W-LSFEM}} - u_{\text{ref.}}}{u_{\text{ref.}}} \times 100\,. \tag{4.5.17}$$

At each *Re*, we use three different computational grids to ensure that grid-independent results are obtained. Results show a very good agreement with the reference solutions at all *Re*.

We compare the kinetic energy values of the V-V-P and S-V-P LSFEMs with the mixed finite element method (MFEM) results obtained from FEATFLOW for different Re. We present the V-V-

Table 4.10: Mesh information for the regularized driven cavity, the number of elements (NEL) and the number of degrees of freedom.

Level	NEL	Degrees of freedom			
		V-V-P		S-V-P	
		Q_1	Q_2	Q_1	Q_2
5	256	1,156	4,356	1,734	6,534
6	1,024	4,356	16,900	6,534	25,350
7	4,096	16,900	66,564	25,350	99,846
8	16,384	66,564	264,196	99,846	396,294
9	65,536	264,196	1,052,676	396,294	1,579,014

Table 4.11: Percent error, equation (4.5.17), of the horizontal velocity through the vertical centerline of the cavity at $Re = 1$, W-LSFEM with Q_2 elements and MFEM with $Q_2P_1^{disc}$ elements

	W-LSFEM on Lev.			MFEM [17]
Coord.	7	8	9	ref. value
0.1	0.07	0.01	0.00	4.702885E-02
0.2	0.01	0.00	0.00	8.265610E-02
0.3	0.01	0.00	0.00	1.146949E-01
0.4	0.02	0.00	0.00	1.441905E-01
0.5	0.01	0.00	0.00	1.654724E-01
0.6	0.09	0.01	0.00	1.638459E-01
0.7	0.16	0.05	0.00	1.114928E-01
0.8	0.80	0.27	0.00	-3.903038E-02
0.9	0.22	0.03	0.00	-3.658718E-01

P formulation results for $\alpha = 1$ and $\alpha = 100$ in Table 4.14. The results accurately converge towards the reference solution with mesh refinement at the different Re numbers assuring that higher order elements are used. As in the last example, accurate results are obtained by the higher-order V-V-P formulation. This clearly shows the demand for higher-order interpolations in LSFEMs. In addition, we present the nodal element S-V-P formulation results in Table 4.15. The accuracy of the results are similar to those obtained with the V-V-P formulation at $\alpha = 1$. The poor performance of the Q_1 finite elements is improved by using higher-order Q_2 elements.

We compare the accuracy of the Q_2 element results of the W-LSFEM with those of Bruneau and Saad [8], obtained with a finite difference method, and also the MFEM reference results [17] for the kinetic energy and the enstrophy. We make this comparison at $Re = 1000$ and summarize the results in Table 4.16. Both global quantities E and Z converge to the MFEM reference solution with grid refinement. Also, Table 4.16 shows that our least-squares solution is more accurate than the solution provided by Bruneau and Saad [8].

4.6. Summary

We used the least-squares FEM to solve two different formulations of the incompressible Navier-Stokes equations. The first-order system for the first formulation is introduced using the vorticity, velocity and pressure, known as the V-V-P formulation, and for the second formulation using the stress, velocity and pressure, known as the S-V-P formulation. Equal order linear and quadratic finite elements are used for the discrete systems. In combination with the Newton technique to treat

Table 4.12: Percent error, equation (4.5.17), of the horizontal velocity through the vertical centerline of the cavity at $Re = 400$, W-LSFEM with Q_2 elements and MFEM with $Q_2P_1^{disc}$ elements

	W-LSFEM on Lev.			MFEM [17]
Coord.	7	8	9	ref. value
0.1	0.75	0.05	0.00	9.447373E-02
0.2	0.50	0.03	0.00	1.817027E-01
0.3	0.11	0.01	0.00	2.335708E-01
0.4	0.30	0.03	0.00	1.967821E-01
0.5	0.68	0.05	0.00	1.040445E-01
0.6	6.35	0.81	0.05	5.466929E-03
0.7	0.68	0.04	0.00	-8.760763E-02
0.8	0.30	0.01	0.00	-1.682347E-01
0.9	0.52	0.06	0.00	-2.365439E-01

Table 4.13: Percent error, equation (4.5.17), of the horizontal velocity through the vertical centerline of the cavity at $Re = 1000$, W-LSFEM with Q_2 elements and MFEM with $Q_2P_1^{disc}$ elements

	W-LSFEM on Lev.			MFEM [17]
Coord.	7	8	9	ref. value
0.1	4.57	0.37	0.04	1.787746E-01
0.2	6.10	0.24	0.02	2.767111E-01
0.3	11.35	0.40	0.02	2.128689E-01
0.4	21.37	0.82	0.06	1.275674E-01
0.5	24.31	0.85	0.07	5.190039E-02
0.6	43.34	2.19	0.12	-2.592104E-02
0.7	28.17	1.32	0.09	-1.093997E-01
0.8	18.72	0.89	0.06	-1.985532E-01
0.9	10.37	0.50	0.04	-2.637121E-01

the nonlinearity, we developed an efficient multigrid-preconditioned CG solver for the solution of the symmetric and positive definite least-squares systems. Also, a preconditioned CG smoother is used inside of the multigrid solver to obtain parameter-free smoothing.

Three incompressible steady-state laminar flow problems are studied. In the Poiseuille flow, we studied the mass conservation of the LSFEM formulations. Different outflow boundary conditions are studied for this problem. In the flow around cylinder test case, the flow accuracy and the mass conservation of the LSFEM formulations are investigated. In the lid-driven cavity test case, the results are analyzed with the help of global quantities, namely the kinetic energy and the enstrophy. We summarize the numerical results as follows:

1. The mass conservation is investigated on one hand with respect to the different formulations and on the other hand with respect to the order of the interpolation functions. The results show that the S-V-P formulation delivers better mass conservation as compared to the V-V-P formulation. The mass conservation of the V-V-P system is enhanced with the help of an extra weighting parameter, but this comes with a negative side effect on the linear solver performance. We observe that in both formulations using higher order finite elements effectively improves the mass conser-

Table 4.14: *V-V-P W-LSFEM*: Convergence of the kinetic energy for the regularized cavity problem and comparison with MFEM results by [17].

	LSFEM				MFEM
	$\alpha = 1$		$\alpha = 100$		
Level	Q_1	Q_2	Q_1	Q_2	Q_2P_1
$\mathrm{Re} = 1$					
6	1.651709E-02	1.861622E-02	1.893533E-02	1.862734E-02	1.862452E-02
7	1.767995E-02	1.862353E-02	1.870948E-02	1.862458E-02	1.862439E-02
8	1.831566E-02	1.862432E-02	1.864614E-02	1.862439E-02	1.862438E-02
9	1.853055E-02	1.862438E-02	1.862980E-02	1.862438E-02	1.862438E-02
$\mathrm{Re} = 400$					
6	2.895209E-02	2.183378E-02	2.909173E-02	2.165057E-02	2.133880E-02
7	3.556316E-02	2.133053E-02	2.500721E-02	2.133948E-02	2.131707E-02
8	3.104720E-02	2.131581E-02	2.240941E-02	2.131695E-02	2.131547E-02
9	2.394639E-02	2.131537E-02	2.159989E-02	2.131546E-02	2.131529E-02
$\mathrm{Re} = 1000$					
6	2.927195E-02	5.338288E-02	3.602868E-02	2.692410E-02	2.289971E-02
7	1.714473E-02	2.552796E-02	3.600000E-02	2.298482E-02	2.277778E-02
8	2.962952E-02	2.287704E-02	2.939236E-02	2.278109E-02	2.276761E-02
9	3.334635E-02	2.277389E-02	2.465346E-02	2.276780E-02	2.276582E-02

vation.

2. The linear MPCG solver performs efficiently for both of the LSFEM formulations with continuous nodal finite elements. We have obtained a grid-independent solver behavior with low and higher order finite elements for the V-V-P as well as the S-V-P formulations. However, the MPCG solver outperforms for the V-V-P system in comparison to the S-V-P system. The reason could be that the V-V-P system is differentially diagonally dominant while the S-V-P system lacks this property.

We conclude that, highly accurate results are obtained with higher order finite elements. More importantly, we have obtained more accurate results with the higher-order finite elements with less number of degrees of freedom as compared to the lower-order elements. This obviously leads to less computational costs. Thus, the accuracy of the V-V-P and the S-V-P LSFEM formulations depends mainly on the order of the interpolation functions. Regarding the efficiency aspects, the MPCG solver performs efficiently for both of the LSFEM formulations with continuous nodal finite elements.

Table 4.15: *S-V-P LSFEM*: Convergence of the kinetic energy for the regularized cavity problem.

	LSFEM		MFEM	
Level	Q_1	Q_2	$\tilde{Q}_1Q_0$	$Q_2P_1^{disc}$
$\mathrm{Re} = 1$				
6	1.624891E-02	1.861560E-02	1.855267E-02	1.862449E-02
7	1.658091E-02	1.862344E-02	1.860621E-02	1.862439E-02
8	1.769969E-02	1.862426E-02	1.861982E-02	1.862438E-02
9	1.831796E-02	1.862437E-02	1.862324E-02	1.862438E-02
$\mathrm{Re} = 400$				
6	3.242921E-02	2.201041E-02	2.136749E-02	1.688675E-02
7	2.849125E-02	2.136759E-02	2.148649E-02	2.131707E-02
8	2.385371E-02	2.131884E-02	2.136484E-02	2.131547E-02
9	2.200378E-02	2.131558E-02	2.132812E-02	2.131529E-02
$\mathrm{Re} = 1000$				
6	3.754829E-02	3.222452E-02	2.768789E-02	2.290234E-02
7	3.853200E-02	2.347606E-02	2.409799E-02	2.277778E-02
8	3.431037E-02	2.281841E-02	2.305179E-02	2.276761E-02
9	2.746944E-02	2.277026E-02	2.282649E-02	2.276582E-02

Table 4.16: Convergence of the kinetic energy E and the enstrophy Z for the regularized cavity problem at $Re = 1000$, comparison between W-LSFEM with Q_2 elements, MFEM with $Q_2P_1^{disc}$ elements and the finite difference method in [8]

Method	grid	E	Z
W-LSFEM	64×64	0.025528	4.806740
	128×128	0.022877	4.827331
	256×256	0.022774	4.830225
MFEM [17]	64×64	0.022778	4.829535
	128×128	0.022768	4.830403
	256×256	0.022766	4.830499
Ref. [8]	64×64	0.021564	4.645800
	128×128	0.022315	4.771100
	256×256	0.022542	4.812300
	512×512	0.022607	4.824300

5

Non-Newtonian Fluid Flows

The numerical simulation of the non-Newtonian power law and the Carreau law fluid flow problems are considered in this chapter. Both the shear thinning and the shear thickening fluids are investigated. We design a physically motivated weighted least-squares method to provide robust solutions for the non-Newtonian fluid flows. We extend the multigrid-preconditioned conjugate gradient (MPCG, see [45, 46]) solver to efficiently solve the stress-based formulation of the non-Newtonian fluids. We study the Newton and the fixed-point linearization methods and analyze the performance of the MPCG solver for these methods.

5.1. Introduction

A non-Newtonian fluid is one whose flow curve (shear stress versus shear rate) is non-linear or does not pass through the origin, i.e. where the apparent viscosity, shear stress divided by shear rate, is not constant at a given temperature and pressure but is dependent on flow conditions such as flow geometry, shear rate, etc. and sometimes even on the kinematic history of the fluid element under consideration [66]. One important group of such materials is known as "time independent" or "generalized Newtonian fluids". They are defined as the fluids for which the rate of shear at any point is determined only by the value of the shear stress at that point at that instant [66]. We study the solution of the time independent fluids in this work.

The time independent fluids are investigated in the literature using the different first-order formulations. Vallala et al. [74] studied the non-Newtonian fluids (based on the Power-Law model) with the vorticity-based as well as the stress-based LSFEM. Both the shear thinning and the shear thickening fluids have been studied by varying the Power-Law index from 0.25 to 1.5. A least-squares finite element method is presented by Chen et al. [11] for the generalized Newtonian (Carreau fluid model) fluid flows. The numerical results indicate that with carefully chosen non-linear weighting functions, the least-squares solution achieves an optimal convergence rate in the L^2-norm for all dependent variables. Surana et al. [4, 20, 21] presented a p-version LSFEM for two-dimensional, incompressible, non-Newtonian (power-law model) fluid flows under isothermal and non-isothermal conditions. A set of first-order differential equations using pressure, velocities, temperature, non-Newtonian stresses and heat fluxes as auxiliary variables is used. The linearization is accomplished after the least-squares minimization by using Newton method with a line search. Similar three-dimensional LSFEM simulations are presented for power law model and temperature dependent viscosity by Dalimunthe and Surana [16] and for isothermal power law and Carreau-Yasuda models by Feng and Surana [24].

We develop a least-squares finite element method for the solution of the power law and the Carreau law fluid flow problems. The solution of the discrete systems is performed with the multigrid-preconditioned conjugate gradient method. To the best of our knowledge, this is for the first time that a multigrid solver is used for the LSFEM solution of the non-Newtonian fluid flow

problems.

5.2. Governing Equations

The non-Newtonian fluid equations are given by

$$\begin{cases} \mathbf{u}\cdot\nabla\mathbf{u}-\nabla\cdot\boldsymbol{\sigma}=\mathbf{f} & \text{in } \Omega \\ \nabla\cdot\mathbf{u}=0 & \text{in } \Omega \\ \boldsymbol{\sigma}+p\mathbf{I}-2\nu\left(D_{\mathrm{II}}(\mathbf{u})\right)\mathbf{D}(\mathbf{u})=0 & \text{in } \Omega \\ \mathbf{u}=\mathbf{g}_D & \text{on } \Gamma_D \\ \mathbf{n}\cdot\boldsymbol{\sigma}=\mathbf{g}_N & \text{on } \Gamma_N \end{cases} \tag{5.2.1}$$

where $\Omega\subset\mathbb{R}^2$ is a bounded domain, $\mathbf{u}$ is the velocity, p is the normalized pressure $p=P/\rho$, $\boldsymbol{\sigma}$ is the Cauchy stress tensor, $\mathbf{f}$ is the source term, $\mathbf{g}_D$ is the value of the Dirichlet boundary conditions on the Dirichlet boundary Γ_D, $\mathbf{g}_N$ is the prescribed traction on the Neumann boundary Γ_N, $\mathbf{n}$ is the outward unit normal on the boundary, $\Gamma=\Gamma_D\cup\Gamma_N$, $\Gamma_D\cap\Gamma_N=\emptyset$ and $\nu(\cdot)$ is the (nonlinear) viscosity. The symmetric part of the deformation tensor is defined as

$$\mathbf{D}(\mathbf{u})=\frac{1}{2}\left(\nabla\mathbf{u}+\nabla\mathbf{u}^T\right) \tag{5.2.2}$$

and the second invariant of the deformation rate tensor as

$$D_{\mathrm{II}}(\mathbf{u})=\frac{1}{2}\big(2\mathbf{D}(\mathbf{u}):2\mathbf{D}(\mathbf{u})\big). \tag{5.2.3}$$

Depending on the chosen viscosity function $\nu(\cdot)$ the following prototypical non-Newtonian fluid models are considered [48]:

1. Carreau law

$$\nu(z)=\nu_\infty+(\nu_0-\nu_\infty)(1+\lambda z)^{\frac{n-1}{2}} \qquad (\nu_o>\nu_\infty\geq 0,\lambda>0), \tag{5.2.4}$$

2. Power law

$$\nu(z)=\nu_o z^{\frac{n-1}{2}} \qquad (\nu_o>0). \tag{5.2.5}$$

The index n in the above equations distinguishes between different types of fluids. For $n=1$ the Newtonian fluid (constant viscosity) is recovered. The choice of $n>1$ and $n<1$ leads to the shear-thickening (viscosity increases with increase in shear rate) and shear-thinning (viscosity decreases with increase in shear rate) fluids, respectively.

5.2.1. Linearization Technique

We use the Newton and the fixed-point methods to replace the nonlinear convective term $\mathbf{u}\cdot\nabla\mathbf{u}$ and the nonlinear viscosity term $-2\nu\left(D_{\mathrm{II}}(\mathbf{u})\right)\mathbf{D}(\mathbf{u})$ in the governing equations (5.2.1) with appropriate linear forms. The effect of the linearization before and after the least-squares minimization is investigated by Vallala et al. [74]. We apply the linearization at the operator level before the minimization.

Using the *nonlinear basic iteration* algorithm, the final solution is obtained through the following iterations

$$U^{n+1}=U^n-\beta\,\delta U^n \tag{5.2.6}$$

where $U=(\mathbf{u},p,\boldsymbol{\sigma})^T$. We obtain the solution update δU^n from the following equation

$$\tilde{\mathfrak{J}}(U^n)\,\delta U^n=\mathbf{d}^n \tag{5.2.7}$$

where $\mathbf{d}^n$ is the residual vector of the system of equations (5.2.1) defined by

$$\mathbf{d}^n = \begin{pmatrix} \mathbf{u}^n \cdot \nabla \mathbf{u}^n - \nabla \cdot \sigma^n - \mathbf{f} \\ \nabla \cdot \mathbf{u}^n \\ \sigma^n + p^n \mathbf{I} - 2\nu\left(D_{II}(\mathbf{u}^n)\right) \mathbf{D}(\mathbf{u}^n) \end{pmatrix}. \tag{5.2.8}$$

For the Newton method, $\tilde{\mathcal{T}}(U^n)\,\delta U^n$ is approximated as follows [49]

$$\mathcal{DT}(U^n)[\delta U^n] = \begin{pmatrix} \mathbf{u}^n \cdot \nabla \delta \mathbf{u}^n + \delta \mathbf{u}^n \cdot \nabla \mathbf{u}^n - \nabla \cdot \delta \sigma^n \\ \nabla \cdot \delta \mathbf{u}^n \\ \delta \sigma^n + \delta p^n \mathbf{I} - 2\nu(D_{II}(\mathbf{u}^n))\,\mathbf{D}(\delta \mathbf{u}^n) - 8\nu'(D_{II}(\mathbf{u}^n))\,[\mathbf{D}(\delta \mathbf{u}^n) : \mathbf{D}(\mathbf{u}^n)]\,\mathbf{D}(\mathbf{u}^n) \end{pmatrix} \tag{5.2.9}$$

where ν' is the derivative of the viscosity with respect to a scalar function. The approximation of $\tilde{\mathcal{T}}(U^n)\,\delta U^n$ using the fixed-point method reads

$$\mathcal{T}(U^n)\delta U^n = \begin{pmatrix} \mathbf{u}^n \cdot \nabla \delta \mathbf{u}^n - \nabla \cdot \delta \sigma^n \\ \nabla \cdot \delta \mathbf{u}^n \\ \delta \sigma^n + \delta p^n \mathbf{I} - 2\nu(D_{II}(\mathbf{u}^n))\,\mathbf{D}(\delta \mathbf{u}^n) \end{pmatrix}. \tag{5.2.10}$$

Solving equation (5.2.7) for δU^n, we update the solution U^{n+1} in every nonlinear iteration. We use the same stopping criteria, see equation (4.3.8) and equation (4.3.9) in chapter 4, for controlling the nonlinear iterations.

5.3. Continuous Least-Squares Principle

We define the L^2-norm least-squares energy functionals based on the residuals of the system (5.2.7). Considering the Newton approximation (5.2.9) we define the following functional

$$\begin{aligned} \mathcal{J}(\mathbf{v}, q, \tau; \mathbf{d}) = {} & \|\mathbf{u}^n \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}^n + \nabla \cdot \tau - \mathbf{d}_1^n\|_0^2 + \alpha \|\nabla \cdot \mathbf{v} - d_2^n\|_0^2 \\ & + \frac{1}{\nu} \left\| \tau + q\mathbf{I} - 2\nu\,\mathbf{D}(\mathbf{v}) - 8\nu'\,[\mathbf{D}(\mathbf{v}) : \mathbf{D}(\mathbf{u}^n)]\,\mathbf{D}(\mathbf{u}^n) - d_3^n \right\|_0^2 \\ & \forall (\mathbf{v}, q, \tau) \in \mathbf{V} \end{aligned} \tag{5.3.1}$$

where $\mathbf{V}$ is the space of admissible functions

$$\mathbf{V} = \left\{ (\mathbf{v}, q, \tau) \in H^1_{g,D}(\Omega) \times H^1(\Omega) \cap L^2_0(\Omega) \times H_{g,N}(\mathrm{div}, \Omega) \right\}, \tag{5.3.2}$$

α is a scaling parameter to improve the mass conservation of the LSFEM formulation [37, 58], $\nu = \nu(D_{II}(\mathbf{u}^n))$ and $\nu' = \nu'(D_{II}(\mathbf{u}^n))$ are short forms used for the sake of simplicity. The third term is weighted by the inverse of the nonlinear viscosity. Such nonlinear weighting technique is used for the solution of the non-Newtonian fluid flows by Chen et al. [11]. They have shown that with carefully chosen nonlinear weighting functions, the least-squares solution achieves an optimal convergence rate in the L^2-norm for the approximation to all dependent variables. Here we use the nonlinear viscosity as the weighting parameter. The nonlinear weight is deferred using the solutions of the previous iteration (or the initial guess in the beginning of the iterations) and hence is a known quantity in every iteration.

The minimization problem associated with the functional (5.3.1) is to find $(\delta \mathbf{u}, \delta p, \delta \sigma) \in \mathbf{V}$ such that

$$(\delta \mathbf{u}, \delta p, \delta \sigma) = \underset{(\mathbf{v}, q, \tau) \in \mathbf{V}}{\mathrm{argmin}}\, \mathcal{J}(\mathbf{v}, q, \tau; f). \tag{5.3.3}$$

The variational problem based on the optimality condition of the minimization problem (5.3.3) is to find $(\delta \mathbf{u}, \delta p, \delta \sigma) \in \mathbf{V}$ such that

$$\mathcal{A}(\delta \mathbf{u}, \delta p, \delta \sigma\,;\, \mathbf{v}, q, \tau) = \mathcal{F}(\mathbf{v}, q, \tau) \quad \forall (\mathbf{v}, q, \tau) \in \mathbf{V} \tag{5.3.4}$$

where $\mathcal{A}$ is a bilinear form defined on $\mathbf{V}\times\mathbf{V}\to\mathbb{R}$

$$\begin{aligned}\mathcal{A}(\delta\mathbf{u},\delta p,\delta\sigma\,;\,\mathbf{v},q,\tau) :=& \alpha\big(\nabla\cdot\delta\mathbf{u},\nabla\cdot\mathbf{v}\big)\\ &+\big(\mathbf{u}^n\cdot\nabla\delta\mathbf{u}+\delta\mathbf{u}\cdot\nabla\mathbf{u}^n-\nabla\cdot\delta\sigma\,,\,\mathbf{u}^n\cdot\nabla\mathbf{v}+\mathbf{v}\cdot\nabla\mathbf{u}^n-\nabla\cdot\tau\big)\\ &+\frac{1}{\nu}\big(\delta\sigma+\delta p\mathbf{I}\,,\,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})-8\nu'\left[\mathbf{D}(\mathbf{v}):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\big)\\ &-\frac{1}{\nu}\big(2\nu\mathbf{D}(\delta\mathbf{u})\,,\,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})-8\nu'\left[\mathbf{D}(\mathbf{v}):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\big)\\ &-\frac{1}{\nu}\big(8\nu'\left[\mathbf{D}(\delta\mathbf{u}^n):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\,,\,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})\big)\\ &+\frac{1}{\nu}\big(8\nu'\left[\mathbf{D}(\delta\mathbf{u}^n):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\,,\,8\nu'\left[\mathbf{D}(\mathbf{v}):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\big)\end{aligned}\tag{5.3.5}$$

and $\mathcal{F}$ is a linear form defined on $\mathbf{V}\to\mathbb{R}$

$$\begin{aligned}\mathcal{F}(\mathbf{v},q,\tau) :=& \big(\mathbf{d}_1^n\,,\,\mathbf{u}^n\cdot\nabla\mathbf{v}+\mathbf{v}\cdot\nabla\mathbf{u}^n-\nabla\cdot\tau\big)+\alpha\big(d_2^n,\nabla\cdot\mathbf{v}\big)\\ &+\frac{1}{\nu}\big(d_3^n,\tau+q\mathbf{I}-2\nu\mathbf{D}(\mathbf{v})-8\nu'\left[\mathbf{D}(\mathbf{v}):\mathbf{D}(\mathbf{u}^n)\right]\mathbf{D}(\mathbf{u}^n)\big)\big)\,.\end{aligned}\tag{5.3.6}$$

5.3.1. Operator form of the Problem

We present the operator form of the least-squares problem for the non-Newtonian fluids. The least-squares operator form is defined as follows

$$\mathsf{L}=\begin{pmatrix}\mathbf{u}^n\cdot\nabla+(\nabla\mathbf{u}^n)^T & 0 & -\nabla\cdot\\ \sqrt{\alpha}\nabla\cdot & 0 & 0\\ C & \frac{1}{\sqrt{\nu}}\mathbf{I} & \frac{1}{\sqrt{\nu}}\end{pmatrix}\tag{5.3.7}$$

where

$$C(\mathbf{v})=-\frac{2\nu}{\sqrt{\nu}}\mathbf{D}(\mathbf{v})-\frac{8\nu'}{\sqrt{\nu}}\left(\mathbf{D}(\mathbf{u}^n):\mathbf{D}(\mathbf{v})\right)\mathbf{D}(\mathbf{u}^n).\tag{5.3.8}$$

To obtain the corresponding dual operator we define

$$\mathbf{A}\otimes\nabla\cdot\tau=\begin{pmatrix}a_{ij}\frac{\partial\tau_{ij}}{\partial x}\\ \\ a_{ij}\frac{\partial\tau_{ij}}{\partial y}\end{pmatrix}\tag{5.3.9}$$

and

$$\langle\langle\sigma\,,\,\tau\rangle\rangle_{\mathbf{A}}=(\mathbf{A}:\sigma\,,\,\mathbf{A}:\tau)\tag{5.3.10}$$

where $\mathbf{A}$ is a second rank tensor and a_{ij} are its components. Using (5.3.9) and (5.3.10), we obtain the dual operator C^* as follows

$$\begin{aligned}\langle C(\mathbf{u}),\tau\rangle =&\langle-\frac{2\nu}{\sqrt{\nu}}\mathbf{D}(\mathbf{u})\,,\,\tau\rangle+\langle-\frac{8\nu'}{\sqrt{\nu}}\mathbf{D}(\mathbf{u}^n):\mathbf{D}(\mathbf{u})\,,\,\mathbf{D}(\mathbf{u}^n):\tau\rangle\\ =&\frac{2\nu}{\sqrt{\nu}}\langle\mathbf{u}\,,\,\nabla\cdot\tau\rangle+\frac{8\nu'}{\sqrt{\nu}}\langle\langle\mathbf{u}\,,\,\mathbf{I}\otimes\nabla\cdot\tau\rangle\rangle_{\mathbf{D}(\mathbf{u}^n)}.\end{aligned}\tag{5.3.11}$$

The operator L^* then reads

$$\mathsf{L}^* = \begin{pmatrix} -\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T^*} & -\sqrt{\alpha}\nabla & C^* \\ 0 & 0 & \frac{1}{\sqrt{\nu}}\mathbf{I}: \\ \nabla & 0 & \frac{1}{\sqrt{\nu}} \end{pmatrix} \tag{5.3.12}$$

where

$$C^*(\tau) = \frac{2\nu}{\sqrt{\nu}}\nabla\cdot\tau + \frac{8\nu'}{\sqrt{\nu}}\mathbf{I}\otimes\nabla\cdot\tau\,. \tag{5.3.13}$$

Finally, the operator form of the least-squares system matrix (5.3.4) reads

$$\mathsf{L}^*\mathsf{L} = \begin{pmatrix} -\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T^*} & -\sqrt{\alpha}\nabla & C^* \\ 0 & 0 & \frac{1}{\sqrt{\nu}}\mathbf{I}: \\ \nabla & 0 & \frac{1}{\sqrt{\nu}} \end{pmatrix} \begin{pmatrix} \mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T} & 0 & -\nabla\cdot \\ \sqrt{\alpha}\nabla\cdot & 0 & 0 \\ C & \frac{1}{\sqrt{\nu}}\mathbf{I} & \frac{1}{\sqrt{\nu}} \end{pmatrix} \tag{5.3.14}$$

where we have

$$\begin{aligned}
(\mathsf{L}^*\mathsf{L})_{1,1} &= [-\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T^*}][\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T}](\cdot) - \alpha\nabla\nabla\cdot + C^*C \\
(\mathsf{L}^*\mathsf{L})_{1,2} &= \frac{1}{\sqrt{\nu}}C^*\mathbf{I} \\
(\mathsf{L}^*\mathsf{L})_{1,3} &= -[-\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T^*}]\nabla\cdot + \frac{1}{\sqrt{\nu}}C^* \\
(\mathsf{L}^*\mathsf{L})_{2,1} &= \frac{1}{\sqrt{\nu}}\mathbf{I}:C \\
(\mathsf{L}^*\mathsf{L})_{2,2} &= \frac{2}{\nu} \\
(\mathsf{L}^*\mathsf{L})_{2,3} &= \frac{1}{\nu}\mathbf{I}: \\
(\mathsf{L}^*\mathsf{L})_{3,1} &= \nabla[\mathbf{u}^n\cdot\nabla + (\nabla\mathbf{u}^n)^{T}](\cdot) + \frac{1}{\sqrt{\nu}}C \\
(\mathsf{L}^*\mathsf{L})_{3,2} &= \frac{1}{\nu}\mathbf{I} \\
(\mathsf{L}^*\mathsf{L})_{3,3} &= \frac{1}{\nu} - \nabla\nabla\cdot
\end{aligned} \tag{5.3.15}$$

The least-squares system matrix for the non-Newtonian fluids (5.3.15) is symmetric but not differentially diagonal dominant, which motivates the study of the multigrid solvers for these type of systems.

5.4. Discrete Least-Squares Principle

We introduce the approximation space $\mathbf{V_h}$, restrict our variational problem (5.3.4) to finite dimensional spaces, and consider the following approximation problem

$$\mathcal{A}_h(\mathbf{u}_h, p_h, \sigma_h; \mathbf{v}_h, q_h, \tau_h) = \mathcal{F}_h(\mathbf{v}_h, q_h, \tau_h) \quad \forall(\mathbf{v}_h, q_h, \tau_h) \in \mathbf{V_h}. \tag{5.4.1}$$

Choosing appropriate basis functions for the finite dimensional space $\mathbf{V_h}$, we obtain a discrete system of equations for the unknown LSFEM variables, namely $(\mathbf{u}_h, p_h, \sigma_h)$. We use conforming finite elements, therefore we set $\mathbf{V_h} \subset \mathbf{V}$.

5.5. Numerical Results: Power Law

In this section, we take the power law model for the viscosity and investigate two non-Newtonian flow problems. We consider both the shear thickening and the shear thinning fluids with varying the power law index n. For the solution of the discrete systems, we use a multigrid-preconditioned conjugate gradient (MPCG) solver. We analyze the performance of the solver with respect to the Newton and the fixed-point linearization methods. The results are compared to the analytic solutions and the grid-independency of the solutions is studied with grid refinement.

5.5.1. Fully Developed Flow between Parallel Plates

For the first test case, we solve a fully developed power law fluid flow between parallel plates with infinite width. Figure 5.1 shows the flow domain and the boundary conditions. Due to the symmetry, we consider only the upper half of the domain.

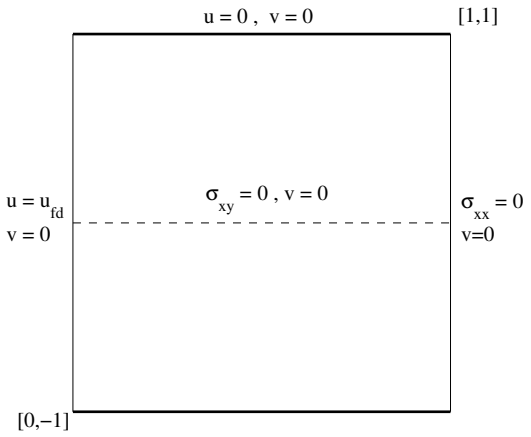


Figure 5.1: Fully developed power law fluid flow between parallel plates

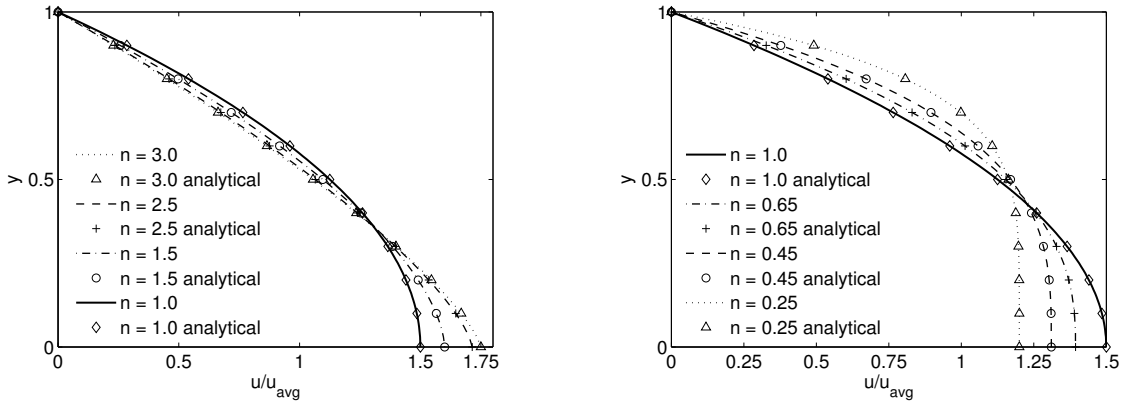
The analytical velocity profile for this flow is obtained as

$$\frac{u}{u_{avg}} = \frac{2n+1}{n+1}\left(1 - y^{\frac{n+1}{n}}\right), \quad y = [0,1] \tag{5.5.1}$$

where u_{avg} is the average velocity of the fluid and n is the power law index [66]. We impose the velocity profile defined in equation (5.5.1) for the inflow boundary condition and set the vertical velocity to zero. The upper edge has no-slip boundary condition, the lower edge is a symmetry line with zero shear-stress $\sigma_{xy} = 0$ and zero vertical velocity $v = 0$, and the outflow has a zero normal-stress boundary condition $\sigma_{xx} = 0$ and a zero vertical velocity $v = 0$.

We use the power law model (5.2.5) for the viscosity and set $\nu_0 = 1$. We solve the problem on a 32×32 grid and use biquadratic finite elements, Q_2, for all variables. We compare our numerical results with the analytical solution (5.5.1) and also the Newtonian fluid flow, i.e. power law with index $n = 1$. Figure 5.2 shows the horizontal velocity profiles along the outflow of the domain for both the shear thinning and the shear thickening fluids. The velocity field is non-dimensionalized with the average velocity u_{avg}. The results compare very well with the analytical solution for all power law indices corresponding to the shear thinning and the shear thickening fluids.

In addition to the graphical comparison of the LSFEM solution with the analytic solution, we calculate the L^2-norm of the velocity vector for different mesh levels and present the results in Table 5.1. We observe asymptotic error reduction with the grid refinement for all values of n. This



Shear thickening Shear thinning

Figure 5.2: Fully developed flow of a power law fluid between parallel plates, horizontal velocity plot along the outflow of the domain on a 32×32 grid

shows the convergence of the numerical method. Further investigations need to be done in order to confirm the a priori error estimates for the non-Newtonian fluid flow problems.

5.5.2. Flow around Circular Cylinder

We simulate a steady state flow passed a circular cylinder with the LSFEM. Figure 5.3 shows the computational mesh of the coarsest level in the multi-level grid hierarchy of the MPCG solver. We present the corresponding mesh information in Table 5.2.

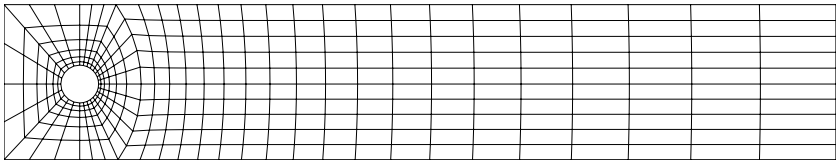


Figure 5.3: Flow around cylinder, computational grid of level 1

We use power law model (5.2.5) for the viscosity with $\nu_0 =$ 1E-3. We impose a parabolic velocity profile with average velocity $u_{avg} = 0.2$ with the vertical velocity $v = 0$ on the inflow boundary. The walls and the cylinder surface has no-slip boundary conditions. The zero normal-stress $\sigma_{xx} = 0$ and the zero vertical velocity $v = 0$ boundary conditions are imposed on the outflow.

Using a mapping between the flow domains of the previous test case and the flow around cylinder, the analytical fully developed velocity profile for this flow configuration is obtained as

$$\frac{u}{u_{avg}} = \frac{2n+1}{n+1}\left(1-(\frac{2y}{0.41}-1)^{\frac{n+1}{n}}\right), \quad y=[0.205,0.41]. \tag{5.5.2}$$

We calculate the lift and drag coefficients and the pressure drop across the cylinder and present the results in Table 5.3. The results show excellent convergence with grid refinement for different values of n. In addition, the values of the lift and drag coefficients and the pressure gradient increase with the power law index. This is true because with the increase in n the fluid becomes more viscose. Consequently, due to the additional resistance of the fluid passing the cylinder the

Table 5.1: Error analysis for the fully developed power law flow, the L^2-error of the velocity vector is calculated for different mesh levels at different power law index n

Level	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate	$\|\mathbf{u}-\mathbf{u}_h\|_0$	rate
	$n=3.0$		$n=0.25$	
5	7.524704E-04	-	5.985903E-05	-
6	1.506990E-04	4.99	6.305550E-06	9.49
7	3.074508E-05	4.90	6.613849E-07	9.53
	$n=2.5$		$n=0.45$	
5	3.164163E-04	-	1.534263E-05	-
6	6.158173E-05	5.14	2.206184E-06	6.95
7	1.202308E-05	5.12	2.810948E-07	7.85
	$n=1.5$		$n=0.65$	
5	2.946439E-05	-	9.421272E-06	-
6	5.819850E-06	5.06	1.137591E-06	8.28
7	1.195208E-06	4.87	1.446111E-07	7.87

Table 5.2: Mesh information for the flow around cylinder problem, the number of elements (NE) and the number of degrees of freedom (NDoF)

Lev.	NE	NDoF
1	346	8,844
2	1,384	34,296
3	5,536	135,024
4	22,144	535,776
5	88,576	2,134,464

hydrodynamic forces and the pressure gradient increase. The flow parameters obtained for $n=1$, corresponds to Reynolds number $\mathsf{Re}=20$ of a Newtonian fluid flow, compare very well with the reference results of the benchmark solution presented by Turek and Schäfer [72].

Moreover, we present the GMC values on the outflow boundary of the domain in Table 5.3. The GMC values decrease with grid refinement for all values of n. This shows that mass conservation is satisfied for the range of non-Newtonian fluids considered in this work.

In Table 5.4 we present the number of nonlinear iterations and the corresponding averaged linear solver (MPCG solver) iterations for both the Newton and the fixed point linearization methods. The total number of iterations for the Newton method is significantly smaller than the number of iterations required for the fixed point method. The MPCG solver for the Newton linearization method shows grid-independent behavior for different values of n. Although the number of iterations increase for the non-Newtonian fluids as compared with the Newtonian fluids. In addition, the number of iterations for the shear thickening fluids is smaller than that of the shear thinning fluids at every level. We do not achieve grid-independent solver behavior for the fixed point linearization method specially for the shear thinning fluids.

Finally, we assume that the channel length and the flow Re allow that the flow reaches a fully developed state before it leaves the domain. We study the validity of this assumption by plotting the horizontal velocity over different cross sections of the domain in Figure 5.4. We compare the

Table 5.3: Flow parameters in the power law fluid flow around cylinder

Level	Drag coefficient C_D	Lift coefficient C_L	Pressure drop $\triangle p$	GMC-value at $x = 2.2$
$n = 2.0$				
3	17.9926513	0.5489032	0.2535041	0.015534
4	18.0040212	0.5486968	0.2544319	0.001813
5	18.0058110	0.5486194	0.2545056	0.000213
$n = 1.5$				
3	9.6539394	0.1319938	0.1593145	0.011602
4	9.6600104	0.1320131	0.1599435	0.001311
5	9.6607101	0.1320204	0.1600681	0.000153
$n = 1.0$				
3	5.5588883	0.0101360	0.1165546	0.022791
4	5.5769755	0.0105355	0.1173265	0.003022
5	5.5792424	0.0106064	0.1174766	0.000556
$n = 0.5$				
3	2.9516453	-0.0289013	0.0892782	0.168400
4	3.0070574	-0.0184097	0.0932435	0.035874
5	3.0208002	-0.0152413	0.0942107	0.006660

Table 5.4: The number of nonlinear iterations and the corresponding averaged number of linear solver iterations for flow around cylinder, nonlinear and linear solver relative errors are kept below 1E-6 and 1E-3, respectively

	Newton			Fixed point		
Level \ n	0.5	1.0	1.5	0.5	1.0	1.5
2	14/15	6/12	6/10	18/12	10/8	10/16
3	14/15	6/12	6/10	18/15	10/10	10/25
4	14/15	6/12	6/10	18/17	10/16	10/25

results of level 5 with the fully developed analytical solution (5.5.2). The velocity profiles for all values of n develop toward the analytical solution as we approach the outflow boundary. The velocity profiles of the shear thickening fluids have a better match with the analytical solution. The reason is that the Reynolds number decreases with the increase in the viscosity of the shear thickening fluids and therefore the flow has a better chance to become fully developed throughout the channel. This analysis also serves as a validation of the LSFEM numerical solutions.

5.6. Summary

A least-squares finite element method is developed for the solution of the power law and the Carreau law fluid flow problems. The governing equations are described in a three-field variable form, i.e. the velocity, the pressure and the stress. Therefore, the resulting set of equations are first-order and well suited for the application of the LSFEM method. We have used conforming finite elements for the FEM approximations. The resulting discrete system is symmetric and positive

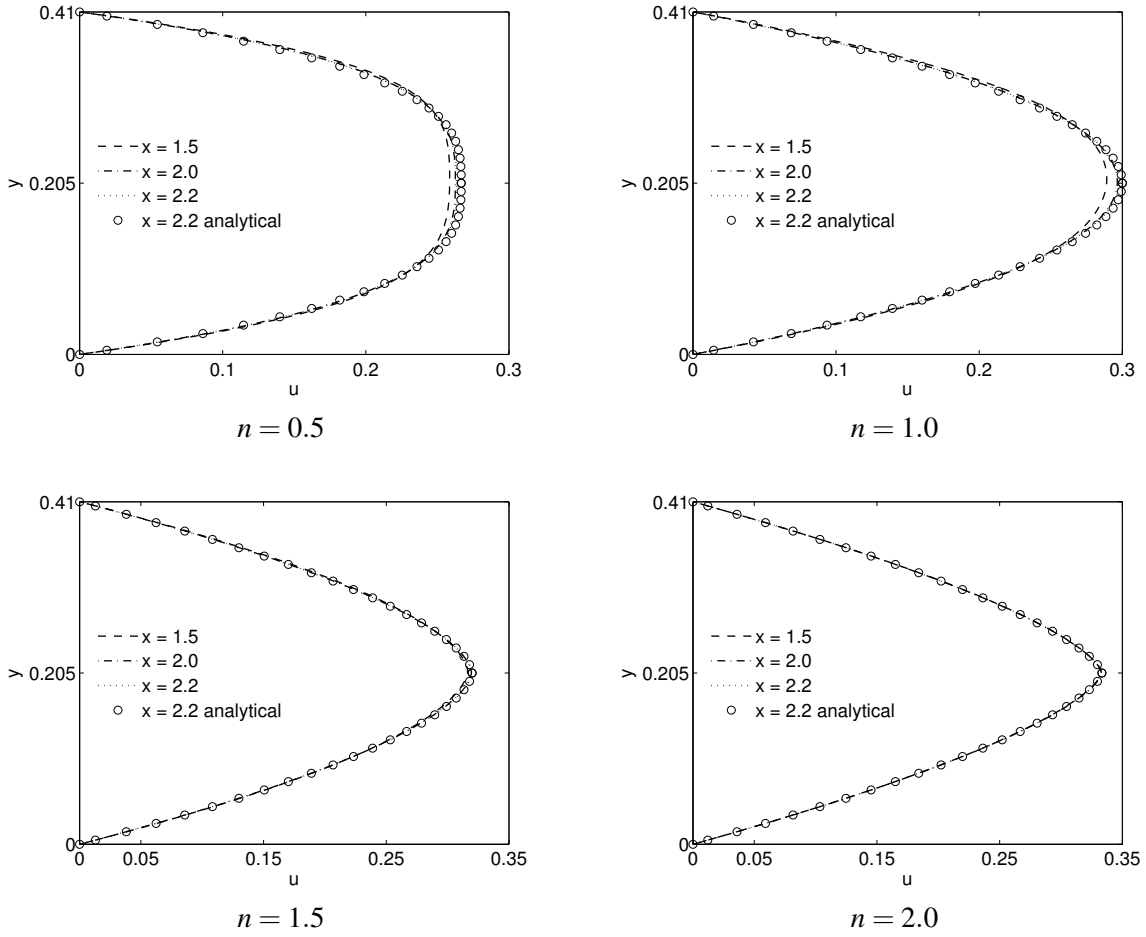


Figure 5.4: Horizontal velocity profile at different cross sections of the flow around cylinder on level 5, comparison with the fully developed analytical solution

definite. However, the system is not differentially diagonal dominant. We have extended the multigrid preconditioned conjugate gradient solver for the solution of the system of equations. Performance of the multigrid solver has been investigated with respect to the different linearization methods. Grid independent behavior is obtained for the Newton linearization for different non-Newtonian fluids. The number of iterations increase for the non-Newtonian fluids as compared with the Newtonian fluids. In addition, the number of iterations for the shear thickening fluids is smaller than that of the shear thinning fluids at every level. The accuracy of the results is confirmed by performing systematic grid refinement and also by comparing the available solutions with the analytic solutions. Further investigations need to be done in order to confirm the a priori error estimates for the non-Newtonian fluid flow problems approximated by the least-squares FEM.

6

Summary and Outlook

This chapter will summarize the findings of the previous chapters and provide a short outlook on possible future directions for the least-squares FEM applied to the advection-diffusion-reaction, Newtonian and non-Newtonian fluid flow problems.

6.1. Conclusion

We study the least-squares finite element solution of the advection-diffusion-reaction equation and the incompressible Navier-Stokes equations for both the Newtonian and the non-Newtonian fluid flow problems. The main focus of the thesis has been on the one hand on efficient multigrid techniques for the solution of the discrete least-squares systems and on the other hand on computational analysis of different first-order system of equations. We have done extensive numerical simulations for the above mentioned equations considering various scaling techniques and broad range of parameters in each case. In particular, the main results of the current work have been summarized in the following paragraphs.

We have shown that, by adding the extra curl equations to the advection-diffusion-reaction equation optimal error estimates for both the scalar and the flux variables can be obtained. These statements are supported by our extensive numerical investigations for the Poisson equation, the diffusion-reaction equation and the advection-diffusion equation which are in accordance with the theoretical a priori error estimates.

Robust grid-independent multigrid performance is obtained for the augmented grad-div-curl system for the Poisson and the diffusion-reaction equations. The advantages of using conjugate gradient smoothers for the multigrid solver are shown through systematic grid refinement studies. We have shown grid-independent multigrid behavior for the linear and quadratic conforming finite elements.

A new first-order formulation based on the total flux is introduced for the advection-diffusion equation. The uniqueness and a priori error estimates have been derived and proved. Moreover, we have used a physically motivated scaling technique to enhance the stability of the proposed formulation. The new formulation provides optimal L^2-error estimates for the scalar variable p and the flux variables even in the highly advection-dominated flows. While for the already available diffusive flux formulation in the literature, the convergence rate of the fluxes degrades in advection-dominated flows. These observations are true for both the Q_1 and the Q_2 finite elements.

We have shown through many numerical examples that the scaled LSFEM performs well even for the advection-dominated flows. In addition, unlike the stabilized LSFEM which its performance highly depends on the proper choice of the stabilization parameter, the scaled LSFEM is parameter-independent. The stabilized LSFEM however, is unable to produce optimal convergence rates for the scalar and the vector variables.

We used the least-squares FEM to solve two different formulations of the incompressible

Navier-Stokes equations. The first-order system for the first formulation is introduced using the vorticity, velocity and pressure, known as the V-V-P formulation, and for the second formulation using the stress, velocity and pressure, known as the S-V-P formulation. Equal order linear and quadratic finite elements are used for the discrete systems. In combination with the Newton technique to treat the nonlinearity, we developed an efficient multigrid-preconditioned CG solver for the solution of the least-squares systems. Also, a preconditioned CG smoother is used inside of the multigrid solver to obtain parameter-free smoothing. The linear MPCG solver performs efficiently for both LSFEM formulations with continuous nodal finite elements. We have obtained grid-independent solver behavior with low and higher order finite elements for the V-V-P as well as the S-V-P formulations. However, the MPCG solver outperforms for the V-V-P system in comparison with the S-V-P system.

The mass conservation of the V-V-P and the S-V-P formulations is investigated on one hand with respect to the weighting parameter α and on the other hand with respect to the order of the interpolation functions. The results show that the S-V-P formulation delivers better mass conservation as compared to the V-V-P formulation. We observe that using higher order finite elements effectively improves the mass conservation in both formulations. Therefore, the continuity equation weighting technique which has side effects on the solver performance is not necessary for the two-dimensional flow problems.

In addition, we conclude that highly accurate LSFEM solutions of the NS equations are obtained provided that the higher order finite elements are used.

Finally, a least-squares finite element method is developed for the solution of the power law and the Carreau law fluid flow problems. Both the shear thinning and the shear thickening fluid models are investigated. We have extended the multigrid preconditioned conjugate gradient solver for the solution of the system of equations. Performance of the multigrid solver has been investigated with respect to the different linearization methods. Grid independent behavior is obtained for the Newton linearization method for different non-Newtonian fluids. The number of iterations increase for the non-Newtonian fluids as compared with the Newtonian fluids. In addition, the number of iterations for the shear thickening fluids is smaller than that of the shear thinning fluids at every grid level. Accurate results are obtained for the fully developed flow and the flow around cylinder problems. Further investigations need to be done to obtain a priori error estimates for the non-Newtonian fluid flow equations approximated by the least-squares FEM.

6.2. Outlook

The numerical solution of the two-dimensional stationary advection-diffusion-reaction equations and the Newtonian and non-Newtonian fluid flows has been presented in this work. The main directions for the future works is listed here.

The generalization of the proposed least-squares FEMs to the three-dimensional flow problems is trivial. Given the efficiency of the proposed MPCG solver, the numerical solutions of the 3D flow problems can be easily obtained within the current framework.

In order to obtain time dependent solutions of the flow problems, the governing equations can be extended to the non-stationary PDEs. The general framework for the time-dependent LSFEM solutions has been developed by many investigators, see e.g. [33, 41], and can be easily adapted to the presented LSFEM methods.

The NS LSFEM solver can be augmented with the level-set equations for the solution of multiphase flow problems. We have shown that the LSFEM is inherently stable for the solution of the advection-dominated flows. Therefore, the LSFEM is an appropriate candidate for the solution of the level-set equation and hence the multiphase flow problems [33].

The possibility of using “arbitrary” finite elements for the approximation of the different

flow variables can be used to extend the current LSFEM methods to the non-conforming, div-conforming and discontinuous finite elements. The inclusion of these FEMs could provide the LSFEM with very uniques features. One example is the improved mass conservation of the Stokes equations using non-conforming FEMs [56, 57]. Another example could be the use of discontinuous finite elements for the solution of the two-phase flow problems. In these problems, the pressure field is discontinuous due to the jump in the pressure across the different phases. The extension of the proposed multigrid solvers to the above mentioned FEMs can also be investigated. Given the symmetry and the positive-definiteness of the discrete least-squares systems, the success of such solver studies is very promising.

In this work we have studied the "time-independent" non-Newtonian fluid flow problems. Two other important classes of the non-Newtonian fluid flows are the "time-dependent fluids" and the "visco-elastic fluids". The former class includes more complex fluids for which the relation between shear stress and shear rate depends, in addition, upon the duration of shearing and their kinematic history. The latter class describes substances exhibiting characteristics of both ideal fluids and elastic solids and showing partial elastic recovery after deformation. Numerical solution of these non-Newtonian fluids is very challenging, see [11, 25, 77], and the current LSFEM framework with its unique features could help the development and design of efficient numerical solutions.

Bibliography

[1] A. I. Pehlivanov and G. F. Carey. Error estimate for least- squares mixed finite elements, RAIRO medel. *Math. Numer. Anal.*, 5:499–516, 1994.

[2] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov. Least-squares mixed finite elements for second order elliptic problems. *SIAM J. Numer. Anal.*, 31:1368–1377, 1994.

[3] A. Bejan. *Advanced Engineering Thermodynamics*. John Wiley & Sons, Hoboken, Third edition, 2006.

[4] Brent C. Bell and Karan S. Surana. 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.

[5] P. Bochev. Analysis of least-squares finite element methods for the Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 34(5):1817–1844, 1997. DOI: 10.1137/S0036142994276001. URL http://epubs.siam.org/doi/abs/10.1137/S0036142994276001.

[6] Pavel B. Bochev, Max, and Max D. Gunzburger. Finite element methods of least-squares type. *SIAM Review*, 40:789–837, 1998.

[7] D. Boffi and L. Gastaldi. On the quadrilateral Q2-P1 element for the stokes problem. *Int. J. Numer. Meth.*, 39:1001–1011, 2002.

[8] C.-H. Bruneau and M. Saad. The 2D lid-driven cavity problem revisited. *Comput. Fluids*, 35:326–348, 2006.

[9] Z. Cai, B. Lee, and P. Wang. Least-squares methods for incompressible newtonian fluid flow: Linear stationary problems. *SIAM J. NUMER. ANAL.*, 42(2):843–859, 2009.

[10] C.L. Chang, S.Y. Yang, and C.H. Hsu. 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. ISSN 0045-7825. DOI: 10.1016/0045-7825(95)00860-7. URL http://www.sciencedirect.com/science/article/pii/0045782595008607.

[11] T.F. Chen, C.L. Cox, H.C. Lee, and K.L. Tung. Least-squares finite element methods for generalized newtonian and viscoelastic flows. *Applied Numerical Mathematics*, 60(10):1024 – 1040, 2010.

[12] C.L. Chang. Finite element approximation for GRAD-DIV type systems in the plane. *SIAM J. Numer. Anal.*, 29:452–461, 1992.

[13] C.L. Chang and J.J. Nelson. Least-squares finite element method for the Stokes problem with zero residual of mass conservation. *SIAM J. Numer. Anal.*, 34:480–489, 1997.

[14] C. Cuvelier, A. Segal, and A. A. Steenhoven. *Finite element methods and Navier-Stokes equations*. D. Reidel Publishing Company, 1986.

[15] D.-P. Yang. Analysis of least-squares mixed finite element methods for nonlinear nonstationary convection-diffusion problems. *Math. Comp.*, 69:929–963, 2000.

[16] H.S. Dalimunthe and K.S. Surana. 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.

[17] H. Damanik, J. Hron, A. Ouazzi, and S. Turek. A monolithic FEM-multigrid solver for non-isothermal incompressible flow on general meshes. *Journal of Computational Physics*, 228: 3869–3881, 2009.

[18] T.A. Davis and I.S. Duff. A combined unifrontal/multifrontal method for unsymmetric sparse matrices. *ACM Trans. Math. Softw.*, 25(1):1–20, 1999. ISSN 0098-3500. DOI: 10.1145/305658.287640. URL http://doi.acm.org/10.1145/305658.287640.

[19] Xu Ding and Tate T. H. Tsang. On first-order formulations of the least-squares finite element method for incompressible flows. *Int. J. Comput. Fluid Dynamics*, 17(3):183–197, 2003. DOI: 10.1080/1061856031000123580. URL http://www.tandfonline.com/doi/abs/10.1080/10618560310001235 80.

[20] Nathan B. Edgar and Karan S. Surana. 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.

[21] Nathan B. Edgar and Karan S. Surana. 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.

[22] Howard C. Elman, Oliver G. Ernst, and Dianne P. O'Leary. A multigrid method enhanced by krylov subspace iteration for discrete helmholtz equations. *SIAM J. Sci. Comput.*, 23 (4):1291–1315, April 2001. ISSN 1064-8275. DOI: 10.1137/S1064827501357190. URL http://dx.doi.org/10.1137/S1064827501357190.

[23] F. P. Incropera and D. P. Dewitt. *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, Fifth edition, 2002.

[24] Zhirong Feng and Karan S. Surana. p-version least squares finite element formulation for three-dimensional, isothermal, incompressible, non-newtonian fluid flow. *Computers and Structures*, 57(5):799 – 816, 1995.

[25] MarcI. Gerritsma. Direct minimization of the discontinuous least-squares spectral element method for viscoelastic fluids. *Journal of Scientific Computing*, 27(1-3):245–256, 2006. DOI: 10.1007/s10915-005-9042-z.

[26] G.F. Carey, A.I. Pehlivanov, and P.S. Vassilevski. Least-squares mixed finite element methods for non-selfadjoint elliptic problems. II: Performance of block-ILU factorization methods. *SIAM J. Sci. Comput.*, 16:1126–1136, 1995.

[27] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer-Verlag, Berlin, 1985.

[28] J. J. Heys, T. A. Manteuffel, S. F. McCormick, and L. N. Olson. Algebraic multigrid for higher-order finite elements. *J. Comput. Phys.*, 204(2):520–532, 2005. ISSN 0021-9991. DOI: 10.1016/j.jcp.2004.10.021. URL http://dx.doi.org/10.1016/j.jcp.2004.10.021.

[29] J.J. Heys, T.A. Manteuffel, S.F. McCormick, and J.W. Ruge. First-order system least squares (FOSLS) for coupled fluid-elastic problems. *J. Comput. Phys.*, 195(2):560 – 575, 2004. ISSN 0021-9991. DOI: 10.1016/j.jcp.2003.09.034. URL http://www.sciencedirect.com/science/article/pii/S0021999103005722.

[30] J.J. Heys, T.A. Manteuffel, S.F. McCormick, M. Milano, J. Westerdale, and M. Belohlavek. Weighted least-squares finite elements based on particle imaging velocimetry data. *J. Comput. Phys.*, 229(1):107 – 118, 2010.

[31] P.-W. Hsieh. *Stabilized Finite Element Methods for Solving Boundary and Interior Layer Problems*. PhD thesis, Department of Mathematics, College of Science, National Central University, Taiwan, June 2011.

[32] J. D. Jackson. *Classical electrodynamics*. Wiley, Third edition, 1999.

[33] B.-N Jiang. *The Least-squares finite element method: Theory and applications in computational fluid dynamics and electromagnetics*. Springer-Verlag, 1998.

[34] J.J. Heys, E. Lee, T.A. Manteuffel, and S.F. McCormick. On mass conserving least-squares methods. *SIAM J. Sci. Comput.*, 28:1675–1693, 2006.

[35] J.J. Heys, E. Lee, T.A. Manteuffel, and S.F. McCormick. An alternative least-squares formulation of the Navier-Stokes equations with improved mass conservation. *J. Comput. Phys.*, 226:994–1006, 2007.

[36] J.J. Heys, E. Lee, T.A. Manteuffel, S.F. McCormick, and J.W. Ruge. Enhanced mass conservation in least-squares methods for Navier-Stokes equations. *SIAM J. Sci. Comput.*, 31: 2303–2321, 2009.

[37] J.M. Deang and M.D. Gunzburger. Issues related to least-squares finite element methods for the Stokes equations. *SIAM J. Sci. Comput.*, 20:878–906, 1998.

[38] J.M. Fiard, T.A. Manteuffel, and S.F. Mccormick. First-order system least squares (FOSLS) for convection-diffusion problems: Numerical results. *SIAM J. Sci. Comput.*, 19:1958–1979, 1998.

[39] J.P. Pontaza and J.N. Reddy. Spectral/hp least-squares finite element formulation for the Navier-Stokes equations. *J. Comput. Phys.*, 190:523–549, 2003.

[40] J.P. Pontaza and J.N. Reddy. Space-time coupled spectral/hp least-squares finite element formulation for the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 197:418–459, 2004.

[41] Oliver Kayser-Herold. *Least-Squares Methods for the Solution of Fluid-Structure Interaction Problems*. PhD thesis, TU Braunschweig, 2006.

[42] M. Köster. *A Hierarchical Flow Solver for Optimisation with PDE Constraints*. PhD thesis, TU Dortmund, 2011.

[43] M. Köster, A. Ouazzi, F. Schieweck, S. Turek, and P. Zajac. New robust nonconforming finite elements of higher order. *Appl. Numer. Math.*, 62:166–184, 2012.

[44] M. Köster and S. Turek. The influence of higher order fem discretisations on multigrid convergence. *Comput. Meth. Appl. Math.*, 6(2):221–232, 2006.

[45] M. Nickaeen, A. Ouazzi, and S. Turek. Newton multigrid least–squares FEM for the V–V–P formulation of the Navier–Stokes equations. Technical report, Fakultät für Mathematik, TU Dortmund, 2013. Ergebnisberichte des Instituts für Angewandte Mathematik, Nummer 466.

[46] M. Nickaeen, A. Ouazzi, and S. Turek. Newton multigrid least-squares FEM for the V-V-P formulation of the Navier-Stokes equations. *J. Comput. Phys.*, 256:416 – 427, 2014. ISSN 0021-9991. DOI: 10.1016/j.jcp.2013.09.011. URL `http://www.sciencedirect.com/science/article/pii/S0021999113006165`.

[47] A. Ouazzi. A mixed formulation of the Stokes equation in terms of (ω, p, u). *Numer. Algorithms*, 21:343–352, 1999.

[48] A. Ouazzi. *Finite element simulation of nonlinear fluids with application to granular material and powder*. Phd thesis, TU Dortmund, Fakultät für Mathematik, Lehrstuhl 3 für Angewandte Mathematik und Numerik, December 2005.

[49] A. Ouazzi. *Finite Element Simulation of Nonlinear Fluids. Application to Granular Material and Powder*. Shaker, 2006.

[50] A. Ouazzi, M. Nickaeen, S. Turek, and M. Waseem. Newton–multigrid least–squares FEM for S–V–P formulation of the Navier–Stokes equations. Technical report, Fakultät für Mathematik, TU Dortmund, November 2013. Ergebnisberichte des Instituts für Angewandte Mathematik, Nummer 486.

[51] P. A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. *Mathematical Aspects of finite element methods, Lecture Notes in Math (606), Springer-Verlag, Berlin, New York*, pages 292–315, 1977.

[52] P. Bochev and M. Gunzburger. *Least-squares finite element methods*. Springer, 2009.

[53] P. Bochev and M. Gunzburger. Least–squares methods for the velocity–pressure–stress formulation of the Stokes equation. *Comput. Methods in Appl. Mech. Eng.*, 126:267–287, 1995.

[54] P. Bochev and M. Gunzburger. On least-squares finite element methods for the poisson equation and their connection to the Dirichlet and Kelvin principles. *SIAM J. Numer. Anal.*, 43:340–362, 2005.

[55] P. Bochev and M. Gunzburger. Least-squares finite element methods. In *Proceedings of the International Congress of Mathematicians (ICM)*, Madrid, 2006. Plenary Lectures.

[56] P. Bochev, J. Lai, and L. Olson. A locally conservative, discontinuous least-squares finite element method for the Stokes equations. *Int. J. Numer. Meth. Fluids*, 68(6):782–804, 2012.

[57] P. Bochev, J. Lai, and L. Olson. A non-conforming least-squares finite element method for incompressible fluid flow problems. *Int. J. Numer. Meth. Fluids*, 00:1–19, 2012.

[58] P. Bolton and R.W. Thatcher. On mass conservation in least-squares methods. *J. Comput. Phys.*, 203:287–304, 2005.

[59] P. Bolton and R.W. Thatcher. A least-squares finite element method for the Navier-Stokes equations. *J. Comput. Phys.*, 213:174–183, 2006.

[60] P.-W. Hsieh and S.-Y. Yang. On efficient least-squares finite element methods for convection dominated problems. *Comput. Meth. Appl. Mech. Eng.*, 199:183–196, 2009.

[61] J. P. Pontaza. *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.

[62] Michael M.J. Proot and Marc I. Gerritsma. Mass- and momentum conservation of the least-squares spectral element method for the Stokes problem. *Journal of Scientific Computing*, 27:389–401, 2006. ISSN 0885-7474. DOI: 10.1007/s10915-005-9030-3. URL `http://dx.doi.org/10.1007/s10915-005-9030-3`.

[63] R. Rannacher and S. Turek. Simple non-conforming quadrilateral Stokes element. *Numerical Methods for Partial Differential Equations*, 8:97–112, 1992.

[64] R. Ranjan and J.N. Reddy. On multigrid methods for the solution of least-squares finite element models for viscous flows. *Int. J. Comput. Fluid Dyn.*, 26:45–65, 2012. DOI: 10.1080/10618562.2011.645031.

[65] R.D. Lazarov, L. Tobiska, and P.S. Vassilevski. Streamline-diffusion least-squares mixed finite element methods for convection-diffusion problems. *East-West J. Numer. Math.*, 5(4): 249–264, 1997.

[66] R.P. Chhabra and J.F. Richardson. *Non-Newtonian Flow in the Process Industries, Fundamentals and Engineering Applications*. Butterworth-Heinemann, 1999.

[67] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, Philadelphia, 2nd edition, 2003.

[68] S.C. Brenner and L.R. Scott. *The mathematical theory of finite element methods*. Springer-Verlag, New York, 1994.

[69] A. Schwarz, M. Nickaeen, S. Serdas, A. Ouazzi, J. Schröder, and S. Turek. A comparative study of mixed least–squares fems for the incompressible Navier–Stokes equations. Technical report, Fakultät für Mathematik, TU Dortmund, August 2013. Ergebnisberichte des Instituts für Angewandte Mathematik, Nummer 478.

[70] Alexander Schwarz, Karl Steeger, and Jörg Schröder. A least-squares mixed finite element for quasi-incompressible elastodynamics. *PAMM*, 10(1):219–220, 2010. ISSN 1617-7061. DOI: 10.1002/pamm.201010102. URL `http://dx.doi.org/10.1002/pamm.201010102`.

[71] S. Turek. *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach*. Springer, Berlin, 1999.

[72] S. Turek and M. Schäfer. 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.

[73] V. Prabhakar, J.P. Pontaza, and J.N. Reddy. A collocation penalty least-squares finite element formulation for incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, 197:449–463, 2008.

[74] V.P. Vallala, J.N. Reddy, and K.S. Surana. Alternative least-squares finite element models of Navier-Stokes equations for power-law fluids. *Engineering Computations*, 28:828–852, 2011.

[75] P. Wesseling. *An Introduction to Multigrid Methods*. John Wiley & Sons, 1st edition, 1992.

[76] H. Wobker. *Efficient Multilevel Solvers and High Performance Computing Techniques for the Finite Element Simulation of Large-Scale Elasticity Problems*. PhD thesis, TU Dortmund, 2010.

[77] Z. Cai and C.R. Westphal. An adaptive mixed least-squares finite element method for viscoelastic fluids of Oldroyd type. *J. Non-Newtonian Fluid Mech.*, 159:72–80, 2009.

[78] Z. Cai, R. Lazarov, T.A. Manteuffel, and S.F. Mccormick. First-order system least squares for second-order partial differential equations: Part I. *SIAM J. Numer. Anal.*, 31:1785–1799, 1994.

[79] Z. Cai, T.A. Manteuffel, and S.F. Mccormick. First-order system least squares for second-order partial differential equations: Part II. *SIAM J. Numer. Anal.*, 34:425–454, 1997.