

## **Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Zhen Wang

---

Date

Preconditioning Techniques for the Incompressible Navier–Stokes Equations

By

Zhen Wang  
Doctor of Philosophy

Mathematics

---

Michele Benzi, Ph.D.  
Advisor

---

James G. Nagy, Ph.D.  
Committee Member

---

Alessandro Veneziani, Ph.D.  
Committee Member

Accepted:

---

Lisa A. Tedesco, Ph.D.  
Dean of the James T. Laney School of Graduate Studies

---

Date

Preconditioning Techniques for the Incompressible Navier–Stokes Equations

By

Zhen Wang  
M.S., Tsinghua University, 2007

Advisor: Michele Benzi, Ph.D.

An abstract of  
A dissertation submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
in Mathematics  
2011

## Abstract

Preconditioning Techniques for the Incompressible Navier–Stokes Equations  
By Zhen Wang

We study different preconditioning techniques for the incompressible Navier–Stokes equations in two and three space dimensions. Both steady and unsteady problems are considered.

First we analyze different variants of the augmented Lagrangian-based block triangular preconditioner. The preconditioners are used to accelerate the convergence of the Generalized Minimal Residual (GMRES) method applied to both stable and stabilized finite element and MAC discretizations of the Oseen problem. We study the eigenvalues of the preconditioned matrices obtained from Picard linearization, and we devise a simple and effective method for the choice of the augmentation parameter  $\gamma$  based on Fourier analysis. Numerical experiments on a wide range of model problems demonstrate the robustness of these preconditioners, yielding fast convergence independent of mesh size and only mildly dependent on viscosity on both uniform and stretched grids. Good results are also obtained on linear systems arising from Newton linearization. We also show that performing inexact preconditioner solves with an algebraic multigrid algorithm results in excellent scalability. Comparisons of the modified augmented Lagrangian preconditioners with other state-of-the-art techniques show the competitiveness of our approach. Implementation on parallel architectures is also considered.

Moreover, we study a Relaxed Dimensional Factorization (RDF) preconditioner for saddle point problems. Properties of the preconditioned matrix are analyzed and compared with those of the closely related Dimensional Splitting preconditioner. Numerical results for a variety of finite element discretizations of both steady and unsteady incompressible flow problems indicate very good behavior of the RDF preconditioner with respect to both mesh size and viscosity.

Preconditioning Techniques for the Incompressible Navier–Stokes Equations

By

Zhen Wang  
M.S., Tsinghua University, 2007

Advisor: Michele Benzi, Ph.D.

A dissertation submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
in Mathematics  
2011

## Acknowledgments

First and foremost, I want to express my sincere gratitude to my advisor, Dr. Michele Benzi for his enthusiasm, patience, encouragement and inspiration. I appreciate all his contributions of time, ideas, advice and funding to support my Ph.D. study and research.

I also thank the other members of my thesis committee: Dr. James G. Nagy and Dr. Alessandro Veneziani, for their invaluable suggestions and insightful questions.

My sincere thanks also go to the faculty, staff, postdoctoral fellows and graduate students in the Department of Mathematics and Computer Science of Emory University for creating a friendly and stimulating environment and assisting me in many different ways.

I am indebted to my other collaborators Dr. Michael Ng, Dr. Qiang Niu, Dr. Maxim Olshanskii and Dr. Leo Rebholz. Their originality, perspicacity and diligence greatly improved several publications. I wish to thank Dr. Chao Yang at Lawrence Berkeley National Laboratory, for offering me the summer internship opportunity to work on his project on image processing.

I would like to thank all the teachers and friends in my life, for helping me grow into a mature young man and making my life colorful. I also owe a huge debt of gratitude to Dr. Zhongxiao Jia at Tsinghua University, who led me to and guided me in the path of becoming a researcher.

Last but not the least, I would like to thank my family for their love and support. For my parents, they bore me, raised me, and are supportive throughout my life. For my wife Qing, without her, it would have been definitely harder to finish my Ph.D. For my daughter Ruiya, her birth made the most beautiful page in my life. To them I dedicate this thesis.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Augmented Lagrangian-based preconditioners</b>	<b>14</b>
2.1	Problem formulation . . . . .	14
2.2	Ideal AL preconditioner for stable finite elements . . . . .	15
2.3	Numerical experiments . . . . .	20
2.4	Modified AL preconditioner . . . . .	26
2.4.1	Eigenvalue analysis . . . . .	28
2.4.2	Estimation of the optimal augmentation parameter $\gamma$ using Fourier analysis . . . . .	35
2.5	Numerical experiments . . . . .	39
2.5.1	Empirical rule for choosing $\gamma$ . . . . .	39
2.5.2	Fourier analysis-based approach for choosing $\gamma$ . . . . .	41
2.5.3	Comparison with other preconditioners . . . . .	46
2.5.4	Inexact solves . . . . .	49
2.5.5	Parallel results . . . . .	51
2.5.6	Unsteady problems . . . . .	52
2.5.7	Extension to 3D problems . . . . .	55
2.5.8	Numerical experiments: 3D examples . . . . .	57
2.5.9	Parallel results: 3D examples . . . . .	60

<b>3</b>	<b>The stabilized case</b>	<b>63</b>
3.1	Ideal AL preconditioners . . . . .	63
3.1.1	Problem formulation . . . . .	64
3.1.2	Augmented linear systems and ideal AL preconditioners	65
3.1.3	Eigenvalue analysis . . . . .	68
3.2	Numerical experiments . . . . .	76
3.3	Modified AL preconditioner . . . . .	78
3.4	Numerical experiments . . . . .	80
3.4.1	Empirical rule for choosing $\gamma$ . . . . .	80
3.4.2	Fourier analysis-based approach for choosing $\gamma$ . . . . .	80
3.4.3	Comparison with other preconditioners . . . . .	85
3.4.4	Inexact solves . . . . .	86
<b>4</b>	<b>A relaxed dimensional factorization preconditioner</b>	<b>88</b>
4.1	Dimensional splitting preconditioner . . . . .	88
4.2	Relaxed dimensional factorization preconditioner . . . . .	91
4.2.1	Practical implementation of the RDF preconditioner . . . . .	95
4.2.2	Estimation of the optimal relaxation parameter $\alpha$ using Fourier analysis . . . . .	97
4.3	Numerical experiments . . . . .	98
4.3.1	Comparison with other preconditioners . . . . .	105
4.4	The stabilized case . . . . .	106
4.5	Numerical experiments . . . . .	112
<b>5</b>	<b>Conclusions</b>	<b>114</b>
	<b>Bibliography</b>	<b>117</b>

# List of Figures

2.1	Streamline plot (left) and pressure plot (right) for lid driven cavity problem ( $\nu = 0.001$ ) using Q2-Q1 approximation on $128 \times 128$ grid. . . . .	21
2.2	Streamline plot (above) and pressure plot (below) for backward facing step problem ( $\nu = 0.005$ ) using Q2-Q1 approximation on $64 \times 192$ grid. . . . .	22
2.3	Plots of the eigenvalues of the Oseen matrix with ideal AL preconditioner (cavity, Q2-Q1, uniform $32 \times 32$ grid). Left: $\nu = 0.1$ . Right: $\nu = 0.001$ . . . . .	24
2.4	Number of iterations vs. parameter $\gamma$ (cavity, Q2-Q1, uniform grids). Left: $\nu = 0.01$ . Right: $\nu = 0.001$ . . . . .	44
2.5	Plots of the eigenvalues of the Oseen matrix with modified AL preconditioner (cavity, Q2-Q1, uniform $32 \times 32$ grid, $\nu = 0.01$ ). Left: with $\gamma$ chosen by Fourier analysis. Right: with optimal $\gamma$ . . . . .	44
4.1	Plots of the eigenvalues of the preconditioned Oseen matrix (cavity, Q2-Q1, uniform $32 \times 32$ grid, $\nu = 0.01$ and experimentally optimal $\alpha$ ). Top: $\tilde{\mathcal{A}}\mathcal{B}^{-1}$ . Bottom: $\tilde{\mathcal{A}}\mathcal{M}^{-1}$ . . . . .	93
4.2	Plots of the eigenvalues of the preconditioned Oseen matrix (cavity, Q2-Q1, uniform $32 \times 32$ grid, $\nu = 0.001$ and experimentally optimal $\alpha$ ). Top: $\tilde{\mathcal{A}}\mathcal{B}^{-1}$ . Bottom: $\tilde{\mathcal{A}}\mathcal{M}^{-1}$ . . . . .	94

4.3 Number of iterations vs. parameter  $\alpha$  (cavity, Q2-Q1, uniform grids). Left:  $\nu = 0.01$ . Right:  $\nu = 0.001$ . . . . . 102

## List of Tables

2.1	Eigenvalue bounds with the ideal AL preconditioner. . . . .	19
2.2	GMRES(50) iterations with ideal AL preconditioner (cavity, Q2-Q1, uniform grids). . . . .	23
2.3	GMRES(50) iterations with ideal AL preconditioner (cavity, Q2-Q1, stretched grids). . . . .	23
2.4	GMRES(50) iterations with ideal AL preconditioner (step, Q2-Q1, uniform grids). . . . .	24
2.5	Inexact inner solvers. GMRES(50) iterations and timings with ideal AL preconditioner (cavity, Q2-Q1, uniform grids). . . . .	25
2.6	The correspondence between the operators and their symbols.	36
2.7	Values of $\gamma$ obtained by empirical rule vs. optimal values (cav- ity, Q2-Q1, uniform grids). . . . .	40
2.8	GMRES(50) iterations with modified AL preconditioner (cav- ity, Q2-Q1, uniform grids). . . . .	40
2.9	GMRES(50) iterations with modified AL preconditioner (cav- ity, Q2-Q1, stretched grids). . . . .	40
2.10	GMRES(50) iterations with modified AL preconditioner (step, Q2-Q1, uniform grids). . . . .	41
2.11	Values of $\gamma$ obtained by Fourier analysis vs. optimal values (cavity, Q2-Q1, uniform grids). . . . .	42

2.12	GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, uniform grids). . . . .	42
2.13	GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, stretched grids). . . . .	42
2.14	GMRES(50) iterations with modified AL preconditioner (step, Q2-Q1, uniform grids). . . . .	43
2.15	GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q2-Q1, uniform grids). The asterisk means that the FA values of $\gamma$ are the ones found for the Oseen problem. . . . .	45
2.16	GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q2-Q1, stretched grids). The asterisk means that the FA values of $\gamma$ are the ones found for the Oseen problem. . . . .	45
2.17	GMRES(50) iterations with modified AL preconditioner (step, Newton, Q2-Q1, uniform grids). The asterisk means that the FA values of $\gamma$ are the ones found for the Oseen problem. . . . .	46
2.18	Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Q2-Q1, uniform grids). . . . .	47
2.19	Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Q2-Q1, stretched grids). . . . .	47
2.20	Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Newton, Q2-Q1, uniform grids). . . . .	47
2.21	Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Newton, Q2-Q1, stretched grids). . . . .	48
2.22	Comparison of exact and inexact inner solvers. GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q2-Q1, uniform grids, $\nu = 0.005$ ). . . . .	49
2.23	GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q2-Q1, uniform $256 \times 256$ grid, $\nu = 0.01$ ) . . . . .	51

2.24	GMRES(50) iterations with modified AL preconditioner (unsteady cavity, Q2-Q1, uniform grids, improved $\widehat{S}^{-1}$ ).	53
2.25	GMRES(50) iterations with modified AL preconditioner (unsteady cavity, Q2-Q1, stretched grids, improved $\widehat{S}^{-1}$ ).	53
2.26	GMRES(50) iterations with modified AL preconditioner (unsteady step, Q2-Q1, uniform grids, improved $\widehat{S}^{-1}$ ).	53
2.27	GMRES iterations with ideal and modified AL preconditioners (3D Stokes, MAC).	57
2.28	GMRES iterations with ideal and modified AL preconditioners (3D Oseen in convection form, MAC).	58
2.29	GMRES iterations with ideal and modified AL preconditioners (3D Oseen in rotation form, MAC).	59
2.30	GMRES(50) iterations with ideal and modified AL preconditioners (unsteady 3D Oseen in convection form, MAC).	59
2.31	GMRES(50) iterations with ideal and modified AL preconditioners (unsteady 3D Oseen in rotation form, MAC).	60
2.32	GMRES(50) iterations and timings with modified AL preconditioner (3D Oseen, MAC, $\nu = 0.01$ ).	61
2.33	GMRES(50) iterations and timings with modified AL preconditioner (3D cavity, P2-P1, $\nu = 0.05$ ).	62
3.1	GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-P0, uniform and stretched grids).	76
3.2	GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-Q1, uniform grids).	77
3.3	GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-Q1, stretched grids).	77
3.4	GMRES(50) iterations with ideal AL preconditioner (step, Q1-Q1, uniform grids).	77

3.5	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	80
3.6	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	81
3.7	GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	81
3.8	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ). . . . .	81
3.9	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ). . . . .	82
3.10	GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ). . . . .	82
3.11	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	83
3.12	GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	83
3.13	GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids, $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ). . . . .	83
3.14	GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q1-Q1, uniform grids). The asterisk means that the FA values of $\gamma$ are the ones found for the Oseen problem. . . . .	84
3.15	GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q1-Q1, stretched grids). The asterisk means that the FA values of $\gamma$ are the ones found for the Oseen problem. . . . .	84
3.16	Full GMRES iterations with PCD, LSC preconditioners (cavity, Q1-Q1, uniform and stretched grids). . . . .	85
3.17	Full GMRES iterations with PCD, LSC preconditioners (cavity, Newton, Q1-Q1, uniform and stretched grids). . . . .	86

3.18	Comparison of exact and inexact inner solvers. GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q1-Q1, uniform grids, $\nu = 0.005$ ).	87
4.1	GMRES(50) iterations with RDF preconditioner (cavity, Q2-Q1, uniform grids).	99
4.2	GMRES(50) iterations with RDF preconditioner (cavity, Q2-Q1, stretched grids).	99
4.3	GMRES(50) iterations with RDF preconditioner (step, Q2-Q1, uniform grids).	100
4.4	GMRES(50) iterations with RDF preconditioner (unsteady cavity, Q2-Q1, uniform grids).	100
4.5	GMRES(50) iterations with RDF preconditioner (unsteady cavity, Q2-Q1, stretched grids).	100
4.6	GMRES(50) iterations with RDF preconditioner (unsteady step, Q2-Q1, uniform grids).	101
4.7	GMRES(50) iterations with RDF preconditioner of the first five Picard iterations (Cavity, Q2-Q1, uniform $128 \times 128$ grid, different viscosities).	102
4.8	GMRES(50) iteration counts and timings with RDF preconditioner (3D Oseen, MAC).	103
4.9	GMRES(50) iteration counts and timings with RDF preconditioner (unsteady 3D Oseen, MAC).	104
4.10	GMRES(50) iterations with diagonal scaling and modified AL preconditioner (cavity, Q2-Q1, stretched grids).	106
4.11	Comparison of RDF and modified AL preconditioners. GMRES(50) iterations and timings (unsteady step, Q2-Q1, uniform grids).	107
4.12	GMRES(50) iterations with RDF preconditioner (cavity, Q1-Q1, uniform grids).	112

4.13 GMRES(50) iterations with RDF preconditioner (cavity, Q1- Q1, stretched grids). . . . .	113
4.14 GMRES(50) iterations with RDF preconditioner (step, Q1- Q1, uniform grids). . . . .	113

# Chapter 1

## Introduction

In this thesis, we consider the solution of the incompressible Navier–Stokes equations governing the flow of viscous Newtonian fluids [72]. For an open bounded domain  $\Omega \subset \mathbb{R}^d (d = 2, 3)$  with boundary  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$  where  $\partial\Omega_D$  and  $\partial\Omega_N$  are disjoint, time interval  $[0, T]$ , and a given external force field  $\mathbf{f}$ , Dirichlet boundary data  $\mathbf{g}$  and initial condition  $\mathbf{u}_0$ , the goal is to find the velocity vector field  $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$  and pressure scalar field  $p = p(\mathbf{x}, t)$  satisfying the following system of partial differential equations:

$$\frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} \quad \text{on } \Omega \times (0, T], \quad (1.1)$$

$$\operatorname{div} \mathbf{u} = 0 \quad \text{on } \Omega \times [0, T], \quad (1.2)$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial\Omega_D \times [0, T], \quad (1.3)$$

$$\nu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} - \mathbf{n} p = 0 \quad \text{on } \partial\Omega_N \times [0, T], \quad (1.4)$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \quad \text{on } \Omega, \quad (1.5)$$

where  $\nu > 0$  is the kinematic viscosity (inversely proportional to the Reynolds number),  $\Delta$  is the vector Laplacian,  $\nabla$  is the gradient,  $\operatorname{div}$  the divergence and  $\mathbf{n}$  the outward normal to the boundary.

Linearization of the Navier–Stokes system (1.1)–(1.5) by Picard fixed-point iteration [37, Section 7.2.2] or time lagging results in a sequence of (general-

ized) Oseen problems of the form

$$\sigma \mathbf{u} - \nu \Delta \mathbf{u} + (\mathbf{v} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} \quad \text{on } \Omega, \quad (1.6)$$

$$\operatorname{div} \mathbf{u} = 0 \quad \text{on } \Omega, \quad (1.7)$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial\Omega_D, \quad (1.8)$$

$$\nu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} - \mathbf{n}p = 0 \quad \text{on } \partial\Omega_N, \quad (1.9)$$

where  $\mathbf{v}$  is a known velocity field from a previous Picard step or time step (the ‘wind’) and  $\sigma$  is a function of the reciprocal of the time step ( $\sigma = 0$  for a steady problem). When  $\mathbf{v} = \mathbf{0}$  we have a (generalized) Stokes problem.

In this thesis we focus mostly on the steady-state Navier–Stokes equations because they present challenges for even the best preconditioning techniques in literature. The steady Navier–Stokes equations are

$$-\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} \quad \text{on } \Omega, \quad (1.10)$$

$$\operatorname{div} \mathbf{u} = 0 \quad \text{on } \Omega \quad (1.11)$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial\Omega_D, \quad (1.12)$$

$$\nu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} - \mathbf{n}p = 0 \quad \text{on } \partial\Omega_N. \quad (1.13)$$

The weak formulation of (1.10)–(1.13) is as follows:

Find  $\mathbf{u} \in \mathbf{H}^1$  and  $p \in L_2(\Omega)$  such that

$$\nu \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} + \int_{\Omega} (\mathbf{u} \cdot \nabla \mathbf{u}) \cdot \mathbf{v} - \int_{\Omega} p(\nabla \cdot \mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \quad \text{for all } \mathbf{v} \in \mathbf{H}_0^1, \quad (1.14)$$

$$\int_{\Omega} q(\nabla \cdot \mathbf{u}) = 0 \quad \text{for all } q \in L^2(\Omega). \quad (1.15)$$

Here  $\nabla \mathbf{u} : \nabla \mathbf{v}$  represents componentwise scalar product, i.e.,  $\nabla u_1 \cdot \nabla v_1 + \nabla u_2 \cdot \nabla v_2$  in 2D ( $\mathbf{u} = (u_1, u_2)$  and  $\mathbf{v} = (v_1, v_2)$ ) and the function spaces are

defined as follows (2D case):

$$\begin{aligned} L_2(\Omega) &=: \left\{ p : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} p^2 < \infty \right\}, \\ \mathcal{H}^1(\Omega) &=: \left\{ u : \Omega \rightarrow \mathbb{R} \mid u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \in L^2(\Omega) \right\}, \\ \mathbf{H}^1 &=: \{ \mathbf{u} \in \mathcal{H}^1(\Omega)^d \mid \mathbf{u} = \mathbf{g} \text{ on } \partial\Omega_D \}, \\ \mathbf{H}_0^1 &=: \{ \mathbf{v} \in \mathcal{H}^1(\Omega)^d \mid \mathbf{v} = \mathbf{0} \text{ on } \partial\Omega_D \}, \end{aligned}$$

The Navier–Stokes system is nonlinear due to the presence of the convective term  $(\mathbf{u} \cdot \nabla)\mathbf{u}$  in (1.10) ( $\int_{\Omega}(\mathbf{u} \cdot \nabla\mathbf{u}) \cdot \mathbf{v}$  in (1.14)). Two widely used linearization schemes are Newton’s method and Picard’s method. Given an approximate solution  $(\mathbf{u}_k, p_k)$  at the  $k$ th step, the nonlinear residual of the weak formulation (1.14)–(1.15) is

$$\begin{aligned} R_k(\mathbf{v}) &= \int_{\Omega} \mathbf{f} \cdot \mathbf{v} - \nu \int_{\Omega} \nabla \mathbf{u}_k : \nabla \mathbf{v} - \int_{\Omega} (\mathbf{u}_k \cdot \nabla \mathbf{u}_k) \cdot \mathbf{v} + \int_{\Omega} p_k (\nabla \cdot \mathbf{v}), \\ r_k(q) &= - \int_{\Omega} q (\nabla \cdot \mathbf{u}_k) \end{aligned}$$

for any  $\mathbf{v} \in \mathbf{H}_0^1$  and  $q \in L^2(\Omega)$ . Writing  $\mathbf{u} = \mathbf{u}_k + \delta\mathbf{u}_k$  and  $p = p_k + \delta p_k$  where the corrections  $\delta\mathbf{u}_k \in \mathbf{H}_0^1$  and  $\delta p_k \in L^2(\Omega)$ , we obtain

$$\begin{aligned} &\nu \int_{\Omega} \nabla \delta\mathbf{u}_k \cdot \nabla \mathbf{v} + \int_{\Omega} (\delta\mathbf{u}_k \cdot \nabla \delta\mathbf{u}_k) \cdot \mathbf{v} + \int_{\Omega} (\mathbf{u}_k \cdot \nabla \delta\mathbf{u}_k) \cdot \mathbf{v} \\ &\quad + \int_{\Omega} (\delta\mathbf{u}_k \cdot \nabla \mathbf{u}_k) \cdot \mathbf{v} - \int_{\Omega} \delta p_k (\nabla \cdot \mathbf{v}) = R_k(\mathbf{v}), \\ &\quad \int_{\Omega} q \nabla \cdot \delta\mathbf{u}_k = r_k(q). \end{aligned}$$

Dropping the quadratic term  $\int_{\Omega} (\delta\mathbf{u}_k \cdot \nabla \delta\mathbf{u}_k) \cdot \mathbf{v}$ , we arrive at the Newton’s linearization:

Find  $\delta\mathbf{u}_k \in \mathbf{H}^1$  and  $\delta p_k \in L_2(\Omega)$  such that

$$\begin{aligned} \nu \int_{\Omega} \nabla \delta\mathbf{u}_k \cdot \nabla \mathbf{v} + \int_{\Omega} (\mathbf{u}_k \cdot \nabla \delta\mathbf{u}_k) \cdot \mathbf{v} + \int_{\Omega} (\delta\mathbf{u}_k \cdot \nabla \mathbf{u}_k) \cdot \mathbf{v} - \int_{\Omega} \delta p_k (\nabla \cdot \mathbf{v}) &= R_k(\mathbf{v}), \\ \int_{\Omega} q \nabla \cdot \delta\mathbf{u}_k &= r_k(q) \end{aligned}$$

for all  $\mathbf{v} \in \mathbf{H}_0^1$  and  $q \in L^2(\Omega)$ . After solving the linearized system for  $\mathbf{u}_k$  and  $\delta p_k$ , the new approximate solutions are  $\delta \mathbf{u}_{k+1} = \mathbf{u}_k + \delta \mathbf{u}_k$  and  $p_{k+1} = p_k + \delta p_k$ . An advantage of Newton's method is its locally quadratic convergence rate provided the initial guess is 'sufficiently' close to the exact solution and the associated Jacobian matrix is nonsingular. However, the radius of the ball of convergence is usually proportional to the viscosity. Therefore, for the steady Navier–Stokes equations with small viscosity, it is essential to run a few steps of Picard's iteration to provide a sufficiently good initial guess for Newton's method.

The weak formulation for Picard's method is identical to that of Newton's method except that another term  $\int_{\Omega} (\delta \mathbf{u}_k \cdot \nabla \mathbf{u}_k) \cdot \mathbf{v}$  is dropped. Hence the linearized problem is:

Find  $\delta \mathbf{u}_k \in \mathbf{H}^1$  and  $\delta p_k \in L_2(\Omega)$  such that

$$\begin{aligned} \nu \int_{\Omega} \nabla \delta \mathbf{u}_k \cdot \nabla \mathbf{v} + \int_{\Omega} (\mathbf{u}_k \cdot \nabla \delta \mathbf{u}_k) \cdot \mathbf{v} - \int_{\Omega} \delta p_k (\nabla \cdot \mathbf{v}) &= R_k(\mathbf{v}), \\ \int_{\Omega} q \nabla \cdot \delta \mathbf{u}_k &= r_k(q) \end{aligned}$$

for all  $\mathbf{v} \in \mathbf{H}_0^1$  and  $q \in L^2(\Omega)$ . The advantage of Picard's methods is that the radius of the ball of convergence is very large, so it is typical to use Picard's iteration to obtain good approximate solution and then switch to quickly convergent Newton's method to take advantage of these two methods. See [37, Section 7.2.2] for more details on Newton and Picard linearization methods.

Spatial discretization of (1.6)–(1.9) using finite differences or finite volumes or that of (1.14)–(1.15) using finite elements results in large, sparse saddle point systems of the form

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \quad \text{or} \quad \mathcal{A}x = b. \quad (1.16)$$

Here  $u$  and  $p$  represent the discrete velocity and pressure, respectively,  $A \in$

$\mathbb{R}^{n \times n}$  is the discretization of the diffusion, convection, and time-dependent terms,  $B^T \in \mathbb{R}^{n \times m}$  is the discrete gradient,  $B$  the (negative) discrete divergence,  $C \in \mathbb{R}^{m \times m}$  is a stabilization matrix, and  $f$  and  $g$  contain forcing and boundary terms. If the discretization satisfies the Ladyzhenskaya-Babuška-Brezzi (LBB, or ‘inf-sup’) stability condition [37], no pressure stabilization is required and we can take  $C = 0$ . If the LBB condition is not satisfied, the stabilization matrix  $C \neq 0$  is usually symmetric and positive semi-definite and the actual choice of  $C$  depends on the particular finite element pair being used; see, e.g., [37, Section 5.3.2].

The linear system (1.16) can also be written as

$$\begin{pmatrix} A & B^T \\ -B & C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ -g \end{pmatrix}, \quad \text{or} \quad \tilde{\mathcal{A}}x = \tilde{b}. \quad (1.17)$$

Note that writing the constraint equation in the form  $-Bu + Cp = -g$  instead of the more frequently used form  $Bu - Cp = g$  in (1.16) leads to a coefficient matrix with eigenvalue spectrum entirely contained in the right half-plane. In contrast, the more symmetric-looking system (1.16) is highly indefinite, in the sense that its eigenvalues surround the origin in the complex plane. See [9, 16] for a discussion of these issues.

Linear systems of the form (1.16) are examples of generalized saddle point problems. In the past few decades, tremendous effort has been invested in the development of fast solution methods for (1.16); see [9] for a comprehensive survey, and [37, 74, 75] for thorough discussions of solvers tailored to the finite element discretizations of Stokes and Navier–Stokes equations. Recent works on sparse direct methods for symmetric saddle point problems include [23, 31]. While highly reliable, direct methods usually require extensive resources in terms of computing time and memory. This is true for three-dimensional (3D) problems (see, e.g., [18]), and also for certain unsteady two-dimensional (2D) problems requiring additional spatial resolution in order to resolve very fine scale features of the solution.

Hence, the efficient solution of systems of the form (1.16) necessitates rapidly convergent iterative methods. The two main approaches available in literature are preconditioned Krylov subspace methods [67] and multigrid methods [73, 74, 78]. The two approaches can be combined by using one or more multigrid cycles as preconditioners for Krylov methods; see [48]. Here we focus on preconditioned Krylov subspace methods, specifically, preconditioned GMRES [67, 68].

Krylov subspace methods are a class of iterative methods for solving large linear systems. For brevity, we will only describe the methods without any preconditioner. Suppose that  $x_0$  is an initial guess for solving (1.16). Denoting the initial residual  $b - \mathcal{A}x_0$  by  $r_0$ , the  $k$ th Krylov subspace is defined as

$$\mathcal{K}_k(\mathcal{A}, r_0) =: \text{span}\{r_0, \mathcal{A}r_0, \mathcal{A}^2r_0, \dots, \mathcal{A}^{k-1}r_0\}, \quad k = 1, 2, \dots$$

Imposing the Petrov-Galerkin condition

$$b - \mathcal{A}x_k \perp \mathcal{L}_k,$$

where  $\mathcal{L}_k$  is a given subspace, we can obtain a  $k$ th approximate solution  $x_k$  lying in the subspace  $x_0 + \mathcal{K}_k(\mathcal{A}, r_0)$ . Two widely used choices are  $\mathcal{L}_k = \mathcal{K}_k(\mathcal{A}, r_0)$  and  $\mathcal{L}_k = \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0)$ . The former applies to the case when  $\mathcal{A}$  is symmetric positive definite, while the latter is used for general nonsingular  $\mathcal{A}$ . Because we mainly deal with (asymmetric) Oseen problems, GMRES, which is built upon the latter choice, is used as the linear solver.

One can see that the Krylov subspaces form a nested sequence

$$\mathcal{K}_1(\mathcal{A}, r_0) \subset \mathcal{K}_2(\mathcal{A}, r_0) \subset \dots \subset \mathcal{K}_d(\mathcal{A}, r_0) = \dots = \mathcal{K}_{n+m}(\mathcal{A}, r_0),$$

where  $d =: \dim \mathcal{K}_{n+m}(\mathcal{A}, b) \leq n + m$ . In exact arithmetic, the exact solution to (1.16) must be obtained when the dimension of the Krylov subspace is  $d$ . However, in practice, one is more interested in computing an approximate solution satisfying some criterion. For instance, in the context of numerical

PDEs, a solution with error at the level of discretization error is usually sufficient. Therefore Krylov subspace methods can be stopped once this level is reached, and it is often the case that the iteration number needed is much less than the size of the matrix  $\mathcal{A}$  provided effective preconditioner is adopted. For more details on Krylov subspace methods as well as preconditioning, see [67].

The convergence rate of Krylov subspace methods is mainly determined by the preconditioner. Generally speaking, the convergence rate with an ideal preconditioner should be independent of problem parameters, such as the mesh size  $h$ , the viscosity  $\nu$ , the particular finite element scheme used, etc. Moreover, the cost per iteration should be linear in the number of unknowns. In spite of considerable progress in recent years, especially in terms of  $h$ -independence, there is still plenty of room for improvements. Specifically, steady-state problems with low viscosity (high Reynolds numbers) and the use of stretched grids pose significant challenges to state-of-the-art solvers.

An important class of preconditioners is the one based on the block LU factorization of the coefficient matrix; see papers [3, 5, 36, 40, 55, 58, 64, 66], the survey [9], the monograph [37] and the many references therein. This class includes a variety of block diagonal and block triangular preconditioners. The crucial ingredient in all these methods is an approximation to the Schur complement  $BA^{-1}B^T + C$ . Examples of preconditioners of this type include the pressure convection diffusion (PCD) preconditioner, the least squares commutator (LSC) preconditioner, and their variants [33, 37, 38]; see also the review in [63]. These preconditioners are fairly robust with respect to grid size and viscosity, but the original PCD preconditioner necessitates the construction of an artificial convection-diffusion operator on the finite element space for the pressure, which is typically not available to the end user, while LSC-type schemes may not yield grid-independent convergence rates for small  $\nu$  and stretched grids; see the discussion in [63] and the exper-

imental results in Section 2.5.3. More recent contributions include the papers [13,14,61,15,17] on preconditioning based on the augmented Lagrangian (AL) approach [40]; this method is also studied in [19]. Other relevant work includes the development of ILU-type preconditioners for saddle point problems [76] and SIMPLE-type block preconditioners [77]. In these papers, these preconditioners are analyzed and compared with other preconditioners; see also [30] and [75], where AL-based preconditioners were found to compare favorably with other approaches. Promising results using AL-type preconditioners have also been reported in the solution of saddle point problems from other application areas; see, e.g., [46].

In this thesis we describe and analyze the AL preconditioners for the incompressible Navier–Stokes equations. These preconditioners can be motivated in terms of the block LU factorization of the augmented linear system. The difference is that due to the presence of augmentation, approximating the pressure Schur complement is relatively easy and the main issue becomes the approximate solution of linear systems associated with the augmented (1,1) block (the ‘primal’ Schur complement); see Chapter 2. It was shown in [13] that the AL-based approach results in preconditioners that are independent of the mesh size  $h$  and fairly insensitive to the viscosity  $\nu$ , resulting in a remarkably robust and nearly optimal solver for the steady Oseen problem up to Reynolds numbers of about 10,000. It was further shown in [61] that the preconditioner performs quite well also for challenging linear systems arising from the linear stability analysis of linearized (Newton) solutions of the incompressible Navier–Stokes equations. Moreover, it was proved in [14] that GMRES with exact application of the AL preconditioner is convergent independent of problem parameters, such as mesh size  $h$ , viscosity  $\nu$  and  $\sigma$ . As shown in [13,61], the crucial ingredient for the AL-based preconditioner was an efficient multigrid cycle used as an approximate solver for the velocity subproblem associated with the (1,1) block of the preconditioner. For

stable discretizations of the steady Oseen problem, excellent results were obtained in [13] with a geometric multigrid scheme based on the method in [69]. However, this sophisticated coupled geometric multigrid scheme may be difficult to implement for general discretizations and geometries, particularly if unstructured grids are used.

To overcome this difficulty, a variant of the AL-based block triangular preconditioner, which we refer to as the *modified AL preconditioner* was introduced in [15], which can be more readily implemented for general discretizations and problem geometries using off-the-shelf algebraic multigrid solvers for scalar elliptic problems. In particular, state-of-the-art parallel algebraic multilevel solvers can be used to solve the subsystems arising in the application of the preconditioner. Though promising results have been obtained in [15], nevertheless, several important questions need further investigation. Firstly, a spectral analysis of the modified AL preconditioner is needed. Secondly, a systematic procedure for estimating the optimal value of the augmentation parameter  $\gamma$  is necessary (only rules of thumb have been provided in [15]). In [17], we presented an eigenvalue analysis of the modified AL preconditioners. We prove that the preconditioned coefficient matrix has 1 as an eigenvalue of algebraic multiplicity at least  $n$  (recall that  $n$  is the number of velocity degrees of freedom). The remaining  $m$  eigenvalues cannot generally be all close to 1 for any value of the augmentation parameter  $\gamma$ ; however, we show how the latter can be chosen so as to approximately minimize the average distance of these  $m$  eigenvalues from 1. We do this by means of a Fourier analysis, following an approach similar to that used in [7, 33, 48]. As we shall see, this approach gives remarkably accurate estimates of the optimal parameter value and results in excellent convergence behavior. Furthermore, numerical results show that the convergence rate of GMRES with the modified AL preconditioner is independent of  $h$ , which has also been proved in [14], and only mild dependent on  $\nu$ .

Furthermore, comparisons on standard benchmark problems show that the modified AL preconditioner often outperform preconditioners based on pressure Schur complement approximations like PCD and LSC, particularly for low viscosities and stretched grids.

In addition, we consider 3D Oseen problems discretized by the Marker-and-Cell (MAC) [49] finite difference method and 3D lid driven cavity problem discretized by P2-P1 stable finite elements in LifeV [1]. The staggered grid in MAC makes it a stable discretization method, so the ideal and modified AL preconditioners for stable finite elements can be applied directly. Our numerical experiments show that in 3D one can expect similar results to those obtained in 2D using stable finite elements. Numerical results are also reported on a computer cluster showing good scalability of the modified AL preconditioner as the number of cores grows.

Moreover, we extend the ideal and modified AL preconditioners to saddle point systems arising from stabilized finite element methods, resulting in  $C \neq 0$  in (1.16) [15]. This requires a different approach to augment the linear system before constructing and applying the preconditioner. We analyze the eigenvalues of the matrix preconditioned by the AL-based preconditioners for stabilized finite elements, and obtain similar spectral properties to the results for stable finite elements [15, 17]. We also use Fourier analysis to choose the augmentation parameter  $\gamma$ , and present the results of numerical experiments [17]. These results show that the convergence rate obtained with the ideal AL-based preconditioner is independent of problem parameters, including grid size, viscosity, and non-uniform meshes, and that with the modified AL preconditioner it is independent of grid size and non-uniform meshes, and only mildly dependent to viscosity. Furthermore, comparisons on standard benchmark problems show that these techniques often outperform preconditioners based on pressure Schur complement approximations like PCD and LSC, particularly for low viscosities.

We also note that in [45], a block diagonal preconditioner based on the augmentation of the (1,1) block, has been introduced and analyzed. However, this technique is different from the AL-based approach discussed above because the preconditioner is applied to the original (non-augmented) system. A similar augmentation block triangular preconditioner has been proposed in [65], and generalized to nonsymmetric saddle point systems in [26]. In the block triangular case, a parameter appears in the off-diagonal block; analysis of this type of preconditioners as well as parameters have been studied in [25, 28, 53]. Similar to the AL preconditioner, the augmentation may increase the cost of applying the preconditioner; some variants have been established in [28, 44, 46, 65] to address this issue. However, for incompressible flow problems, we have found that applying the AL preconditioners to the *augmented* linear system leads to much faster convergence than the original one, so we will only focus on the AL-based approach.

Other types of preconditioners for (possibly nonsymmetric) saddle point problems include those based on the Hermitian and skew-Hermitian Splitting (HSS) and on the Dimensional Splitting (DS) of the coefficient matrix  $\tilde{\mathcal{A}}$ ; see, respectively, [4, 8, 11, 27, 48] and [10]. These preconditioners have been shown to be effective on a wide range of problems. However, HSS is difficult to implement efficiently for the Oseen problem (except as a smoother, or for the rotation form of the Navier–Stokes equations, cf. [48, 11]), and DS preconditioning has difficulties dealing with low-viscosity problems on stretched grids.

In this thesis, we build on the DS preconditioner introduced in [10] and develop a technique which will be referred to as the *Relaxed Dimensional Factorization* (RDF) preconditioner [12] for solving linear system (1.17). The idea for this method comes from explicitly forming the (parameter-dependent) DS preconditioner, originally given in factorized form, and comparing the DS preconditioner with the original coefficient matrix  $\tilde{\mathcal{A}}$ . This

reveals that certain diagonal terms of the DS preconditioner can be neglected without adversely affecting the quality of the approximation; indeed, dropping some of these terms actually leads to a *better* approximation of  $\tilde{\mathcal{A}}$ , suggesting the possibility of improvements in the performance of the preconditioner. This intuition is indeed confirmed in many cases both by numerical experiments and by comparing the clustering effect of RDF preconditioning with that of DS on the spectrum of the preconditioned matrices. We derive some simple results for the spectrum of RDF-preconditioned matrices. Furthermore, we apply a Fourier analysis (more complete than the one in [12]) to guide in the choice of the RDF parameter. We present the result of numerical experiments, including comparisons with other preconditioners, using test problems generated from discretizations of the 2D Oseen equations by Q2-Q1 stable finite elements. We also present a few results for 3D problems. Additionally, we propose a generalization of the RDF preconditioner to linear systems obtained by stabilized finite element discretization, and its effectiveness is demonstrated by numerical experiments.

The thesis is organized as follows. In Chapter 2 we focus on the augmented Lagrangian-based preconditioners. We first review the ideal AL preconditioner [13] and analyze its spectral properties in Section 2.2. Next in Section 2.4, we describe the modified AL preconditioner, analyze the eigenvalues of the modified AL-preconditioned saddle point matrix in the case of LBB-stable finite elements discretizations, and show how to use Fourier analysis to guide in the choice of  $\gamma$ . Numerical experiments on various standard reference problems and comparison with other state-of-the-art preconditioners show the effectiveness of our approach. We also investigate the effect of inexact solves in the application of the preconditioner. Extension of the modified AL preconditioner to 3D problems is discussed in Section 2.5.7. Results for Oseen problems discretized by stable MAC and finite elements methods on a parallel computer are also presented. In Sections 3.1 and 3.3 we consider

the ideal and modified AL preconditioners for stabilized finite elements, respectively. The spectrum of the coefficient matrix preconditioned with the ideal and modified AL preconditioners is analyzed, Fourier analysis is applied to determine  $\gamma$ , and numerical experiments are presented. In Chapter 3 we recall the RDF preconditioner [12], show how Fourier analysis can be used to select the parameter in RDF, present the results of numerical experiments demonstrating the performance of the RDF preconditioner, and discuss a generalization to linear systems discretized by stabilized finite elements.

# Chapter 2

## Augmented Lagrangian-based preconditioners

In this chapter we begin by recalling the ideal AL-based preconditioner described in [13]. Then we present the modified AL-based block triangular preconditioner for the Oseen problem discretized by stable finite element pairs, such as Q2-Q1 or Q2-P1; see, e.g., [37]. Subsequently, we study the spectrum of the preconditioned saddle point matrix using this preconditioner for the steady 2D Oseen problem. Moreover, we describe an approach based on Fourier analysis for choosing the augmentation parameter  $\gamma$ , and we report on the results of numerical experiments demonstrating the effectiveness of this approach. The extension of the modified AL preconditioner to the 3D problems is given.

### 2.1 Problem formulation

Here we consider solving the steady-state Oseen equations ( $\sigma = 0$  in (1.6)) for simplicity. Unless otherwise stated, we will focus on the steady case in the sequel. An LBB-stable finite element discretization gives rise to the following system:

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \quad \text{or} \quad \mathcal{A}x = b. \quad (2.1)$$

From  $Bu = g$  it follows that

$$\gamma B^T W^{-1} B u = \gamma B^T W^{-1} g.$$

Adding the above equation to  $Au + B^T p = f$  gives

$$(A + \gamma B^T W^{-1} B)u + B^T p = f + \gamma B^T W^{-1} g,$$

which leads to the equivalent augmented Lagrangian (AL) formulation [40]

$$\begin{pmatrix} A_\gamma & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f_\gamma \\ g \end{pmatrix}, \quad \text{or} \quad \hat{\mathcal{A}}x = \hat{b}, \quad (2.2)$$

where  $A_\gamma =: A + \gamma B^T W^{-1} B$ ,  $f_\gamma =: f + \gamma B^T W^{-1} g$ ,  $W$  is symmetric positive definite, and  $\gamma > 0$ . A good choice of  $W$  is the pressure mass matrix  $M_p$ ; in practice, we use the main diagonal of  $M_p$  instead, in order to maintain the sparsity in  $A_\gamma$ . The choice of  $\gamma$  is important and will be discussed below.

The use of the AL formulation (2.2) instead of the original one (2.1) can be justified in various ways; see for instance the discussion in [13, 61]. Here we justify this choice by the observation that preconditioning (2.2) allows us to circumvent the delicate issue of finding good approximations for the pressure Schur complement  $BA^{-1}B^T$  or its inverse, which is crucial when constructing block diagonal or block triangular preconditioners for the non-augmented system (2.1).

## 2.2 Ideal AL preconditioner for stable finite elements

An ideal preconditioner for problem (2.2) is given by the block triangular matrix [13]

$$\mathcal{P} = \begin{pmatrix} A_\gamma & B^T \\ 0 & \hat{S} \end{pmatrix}, \quad (2.3)$$

where

$$\widehat{S}^{-1} =: -\nu\widehat{M}_p^{-1} - \gamma W^{-1}. \quad (2.4)$$

The spectral properties associated with the preconditioner above have been analyzed in the following theorem.

**Theorem 2.1** ([13]). *Setting  $W = M_p$  (the pressure mass matrix), for  $\gamma > 0$ ,  $\mathcal{P}^{-1}\widehat{A}$  has the eigenvalue 1 of algebraic multiplicity  $n$  (size of  $A$ ), and the remaining  $m$  (size of  $W$ ) eigenvalues are contained in a rectangle in the right half-plane independent of the mesh size  $h$ . Moreover,  $\gamma$  can be chosen such that the rectangle does not depend on the viscosity  $\nu$ . When  $\gamma \rightarrow \infty$ , all the eigenvalues tend to 1.*

Moreover, using field-of-values analysis, it has been proved that GMRES with this preconditioner is convergent at a rate independent of problems parameters, such as  $h$ ,  $\nu$  and  $\sigma$ .

**Theorem 2.2** ([14]). *Define the norm*

$$\langle \{u, p\}, \{v, q\} \rangle_{-a} =: \langle A_S^{-1}u, v \rangle + (\nu + \gamma)\langle M_p^{-1}p, q \rangle,$$

where  $A_S =: \frac{1}{2}(A_\gamma + A_\gamma^T)$ . For  $\nu < 1$ , if  $\gamma = \|(BA^{-1}B^T)^{-1}M_p\|_{M_p}$ , the residual norms in GMRES with the ideal AL preconditioner satisfy

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

where  $r^k$  is the GMRES residual of the  $k$ -th step, and  $\tau < 1$  is independent of problem parameters  $h$ ,  $\nu$  and  $\sigma$ .

Here the term ‘ideal’ refers to the fact that the analysis in [13, 14] assumes that the action of the preconditioner is computed exactly, i.e., linear systems associated with the sub-matrices  $A_\gamma$  and  $\widehat{S}$  are solved exactly, which is

generally not feasible in practice. This approach is different from the ‘modified’ preconditioner discussed in Section 2.4. Another simple choice of  $\widehat{S}$  is  $\widehat{S} = -\gamma^{-1}W$ . Because of the identity

$$\mathcal{P}^{-1} = \begin{pmatrix} A_\gamma^{-1} & \gamma A_\gamma^{-1} B^T W^{-1} \\ 0 & -\gamma W^{-1} \end{pmatrix} = \begin{pmatrix} A_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} I_n & B^T \\ 0 & -I_m \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & \gamma W^{-1} \end{pmatrix},$$

applying  $\mathcal{P}^{-1}$  to a vector mainly requires one solve with  $W$  and one with  $A_\gamma$ . The right-preconditioned matrix is

$$\widehat{\mathcal{A}}\mathcal{P}^{-1} = \begin{pmatrix} I_n & 0 \\ BA_\gamma^{-1} & \gamma BA_\gamma^{-1} B^T W^{-1} \end{pmatrix}, \quad (2.5)$$

showing that  $\lambda = 1$  is an eigenvalue of algebraic multiplicity  $n$ . Additionally, the other  $m$  eigenvalues  $\lambda$  satisfy the generalized eigenvalue problem

$$BA_\gamma^{-1} B^T p = \lambda \left( \frac{1}{\gamma} W \right) p.$$

Lemma 4.1 in [13] states that if all the relevant matrices are invertible, it holds

$$\begin{aligned} & (BA_\gamma^{-1} B^T)^{-1} \\ &= (B(A + \gamma B^T W^{-1} B)^{-1} B^T)^{-1} \\ &= (BA^{-1} B^T)^{-1} + \gamma W^{-1}. \end{aligned} \quad (2.6)$$

Setting  $W = M_p$  (the pressure mass matrix), we thus have

$$\begin{aligned} \frac{1}{\lambda} p &= \frac{1}{\gamma} (BA_\gamma^{-1} B^T)^{-1} M_p p \\ &= \frac{1}{\gamma} ((BA^{-1} B^T)^{-1} + \gamma M_p^{-1}) M_p p \\ &= \frac{1}{\gamma} (BA^{-1} B^T)^{-1} M_p p + p \\ &= \frac{1}{\gamma \mu} p + p, \end{aligned}$$

where  $\mu$  satisfies the generalized eigenproblem

$$BA^{-1}B^Tq = \mu M_p q.$$

Hence, we obtain

$$\lambda = \frac{\gamma\mu}{1 + \gamma\mu}. \quad (2.7)$$

Next, we derive bounds for the real and imaginary parts of  $\lambda$ . Writing  $\lambda = a_\lambda + ib_\lambda$  and  $\mu = a_\mu + ib_\mu$ , where  $i = \sqrt{-1}$ , we have

$$a_\lambda + ib_\lambda = \frac{\gamma(a_\mu + ib_\mu)}{1 + \gamma(a_\mu + ib_\mu)} = \frac{\gamma a_\mu(1 + \gamma a_\mu) + \gamma^2 b_\mu^2 + i\gamma b_\mu}{(1 + \gamma a_\mu)^2 + \gamma^2 b_\mu^2},$$

which implies

$$a_\lambda = \frac{\gamma a_\mu(1 + \gamma a_\mu) + \gamma^2 b_\mu^2}{(1 + \gamma a_\mu)^2 + \gamma^2 b_\mu^2} \quad \text{and} \quad b_\lambda = \frac{\gamma b_\mu}{(1 + \gamma a_\mu)^2 + \gamma^2 b_\mu^2}.$$

Following the argument in [34], it can be shown that the  $\mu$ 's are contained in a rectangle which lies in the positive half-plane  $\text{Re}(z) > 0$  and which does not depend on  $h$ . Easy manipulations give

$$0 < \min_{\mu} \frac{\gamma a_\mu}{1 + \gamma a_\mu} \leq a_\lambda \leq 1, \quad (2.8)$$

$$|b_\lambda| \leq \max_{\mu} \min \left\{ \gamma |b_\mu|, \frac{1}{\gamma |b_\mu|} \right\} \leq 1. \quad (2.9)$$

Note that the real part of  $\lambda$  is bounded away from zero uniformly in  $h$ , for all fixed  $\gamma > 0$ . Likewise, the imaginary part of  $\lambda$  is bounded uniformly in  $h$  by 1, for all fixed  $\gamma > 0$ . Furthermore, for  $\gamma \rightarrow \infty$  all eigenvalues tend to 1; cf. Theorem 2.1.

In Table 2.1 we show the maximum and minimum of the real parts and the maximum imaginary parts of  $\mu$  and  $\lambda$  corresponding to a uniform Q2-Q1 discretization of the regularized lid driven cavity problem described in [37, Section 7.1] (see also [47]). The value  $\gamma = 1$  is used in the ideal AL preconditioner. Notice that 0 is also an eigenvalue (corresponding to the hydrostatic

Table 2.1: Eigenvalue bounds with the ideal AL preconditioner.

Grid	$\max a_\mu$	$\min a_\mu$	$\max  b_\mu $	$\max a_\lambda$	$\min a_\lambda$	$\max  b_\lambda $
Viscosity	0.1					
$16 \times 16$	15.677	1.259	2.274	0.9411	0.5573	0.0127
$32 \times 32$	19.355	1.277	4.323	0.9519	0.5608	0.0121
$64 \times 64$	21.147	1.278	4.973	0.9553	0.5610	0.0185
Viscosity	0.01					
$16 \times 16$	132.77	9.16	38.22	0.9925	0.9016	0.0275
$32 \times 32$	159.89	11.57	64.60	0.9938	0.9204	0.0292
$64 \times 64$	192.85	12.65	88.61	0.9948	0.9267	0.0255
Viscosity	0.001					
$16 \times 16$	1279.6	2.3	148.9	0.9992	0.6961	0.0586
$32 \times 32$	1477.7	2.2	301.3	0.9993	0.6914	0.0529
$64 \times 64$	1584.5	2.3	452.2	0.9994	0.6968	0.0270

pressure mode), which does not affect the convergence of preconditioned GMRES and can be excluded from further consideration (see [37, Section 2.3]). One can clearly see the independence of  $\lambda$  with respect to  $h$ , and the near-independence with respect to  $\nu$ . The real and imaginary parts of  $\mu$  show a weak dependence on  $h$ , which according to the theory must disappear in the limit of  $h \rightarrow 0$ ; note that the minimum of the real part of  $\mu$  is already  $h$ -independent even for these rather coarse grids. On the other hand, there is a strong dependence of  $\mu$  on the viscosity  $\nu$ , as already observed in [34].

Because of the expensive solve associated with the velocity sub-matrix  $A_\gamma$ , the ‘ideal’ preconditioner  $\mathcal{P}$  in (2.3) is not practical. It is necessary to replace

the exact solves with inexact ones, leading to preconditioners of the form

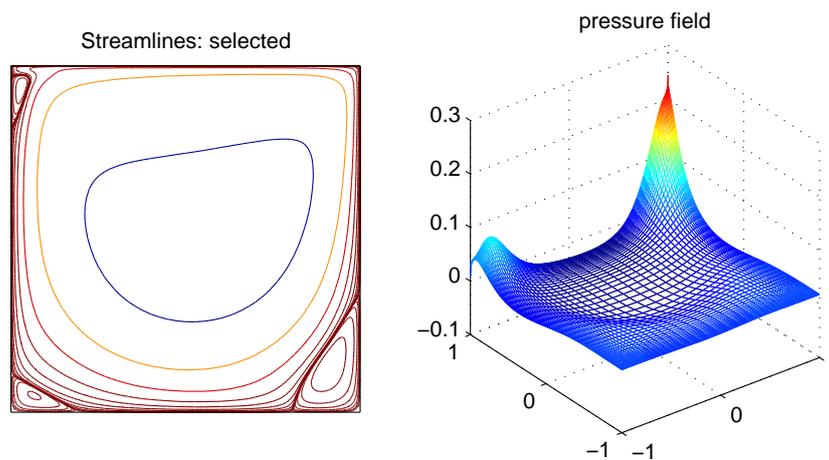
$$\mathcal{P} = \begin{pmatrix} \widehat{A}_\gamma & B^T \\ 0 & \widehat{S} \end{pmatrix}, \quad (2.10)$$

where  $\widehat{A}_\gamma \approx A + \gamma B^T W^{-1} B$  and  $\widehat{S}$  are implicitly defined via the action of their inverses on vectors. In [13],  $\widehat{A}_\gamma^{-1}$  was implemented as a W-cycle of a non-standard geometric multigrid method based on [69]; for  $\widehat{S}^{-1}$ , a few Richardson iterations preconditioned with the diagonal  $\widehat{M}_p$  of  $M_p$  were used to approximately solve linear systems associated with  $M_p$ . (In practice, another acceptable choice is to simply use  $\widehat{S}^{-1} = -\gamma \widehat{M}_p^{-1}$ .) Theory and numerical experiments in [13] show that the preconditioner (2.10) is nearly optimal, meaning that the rate of convergence of Krylov subspace methods with this AL preconditioner is independent of the grid and almost insensitive to viscosity.

## 2.3 Numerical experiments

In this section we present the results of numerical experiments with the ideal AL preconditioner. The test problems are generated by IFISS [35, 70]. The linearized Oseen system is the one in the first step of the Picard iteration immediately following the initial one (which reduces to a Stokes problem) and is discretized by Q2-Q1 stable finite elements. For the ideal AL preconditioner, we use  $\gamma = 1$  for all the experiments; little is gained in tuning  $\gamma$ . The basic Krylov solver used in all our experiments is GMRES (50) [67], starting from a zero initial guess. The iteration is stopped when the relative residual norm is reduced below  $10^{-6}$ . Also, we set  $W = \widehat{M}_p = \text{diag}(M_p)$  and  $\widehat{S}^{-1} = -\gamma \widehat{M}_p^{-1}$ . Right preconditioning is used in all cases; the results for left preconditioning are similar. An exact solve is performed on the subproblem involving  $A_\gamma$  by means of sparse LU factorizations preceded by an approximate minimum degree column reordering [2] (using the MATLAB function

Figure 2.1: Streamline plot (left) and pressure plot (right) for lid driven cavity problem ( $\nu = 0.001$ ) using Q2-Q1 approximation on  $128 \times 128$  grid.



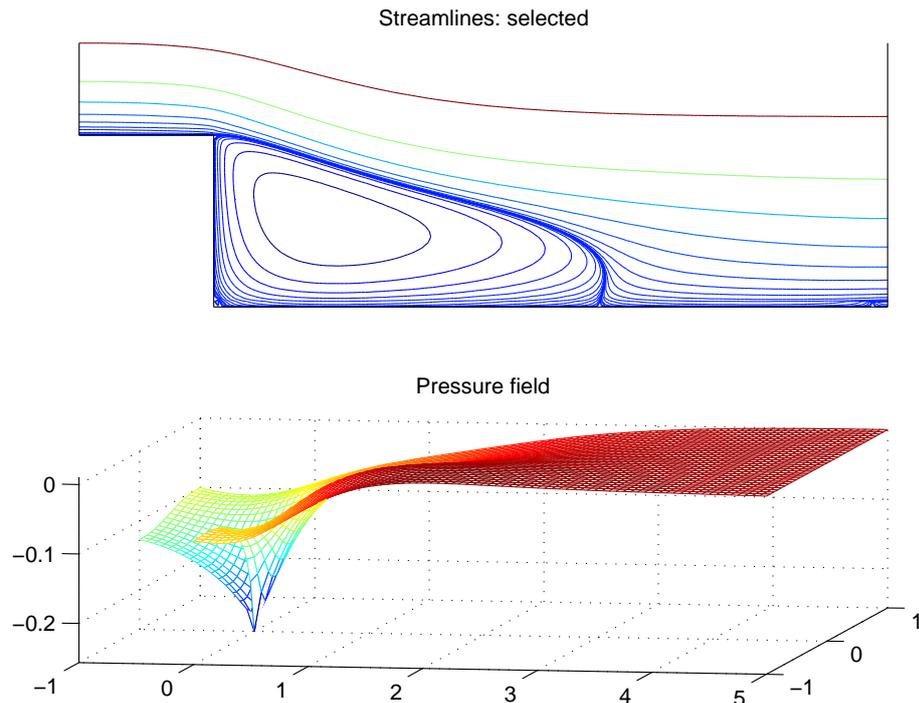
colamd) of the matrix  $A_\gamma$ . All computations are performed in MATLAB on a Sun Microsystems SunFire with 4 dual-core AMD Opteron processors and 32 GB of memory.

The numerical experiments are based on three test problems. The first one is the regularized lid driven cavity problem discretized on a uniform grid. The flow is enclosed in a square with no flow conditions on the bottom, left and right boundaries, and the following regularized condition:

$$u_1 = 1 - x^4, \quad u_2 = 0$$

represents the moving lid. The second one is the same problem but discretized on a stretched grid to investigate the influence of non-uniform elements; the

Figure 2.2: Streamline plot (above) and pressure plot (below) for backward facing step problem ( $\nu = 0.005$ ) using Q2-Q1 approximation on  $64 \times 192$  grid.



stretch factors used are the default ones in IFISS, namely 1.2712 for the  $16 \times 16$  grid, 1.1669 for the  $32 \times 32$  grid, 1.0977 for the  $64 \times 64$  grid, and 1.056 for the  $128 \times 128$  grid. The stretching is done in both the horizontal and vertical directions starting at the center of the domain, resulting in rather fine grids near the boundaries. The third problem is the backward facing step problem; we include this problem because it is a standard benchmark and because we are interested in seeing the effect of a non-square domain. For this problem the smallest value of the viscosity used is  $\nu = 0.005$ , since the flow would be

Table 2.2: GMRES(50) iterations with ideal AL preconditioner (cavity, Q2-Q1, uniform grids).

Viscosity	0.1	0.01	0.005	0.001
Grid	$\gamma = 1$			
$16 \times 16$	6	4	5	5
$32 \times 32$	5	4	4	4
$64 \times 64$	5	3	3	4
$128 \times 128$	4	3	3	3

Table 2.3: GMRES(50) iterations with ideal AL preconditioner (cavity, Q2-Q1, stretched grids).

Viscosity	0.1	0.01	0.005	0.001
Grid	$\gamma = 1$			
$16 \times 16$	5	4	5	5
$32 \times 32$	4	3	4	4
$64 \times 64$	4	3	3	4
$128 \times 128$	3	3	3	3

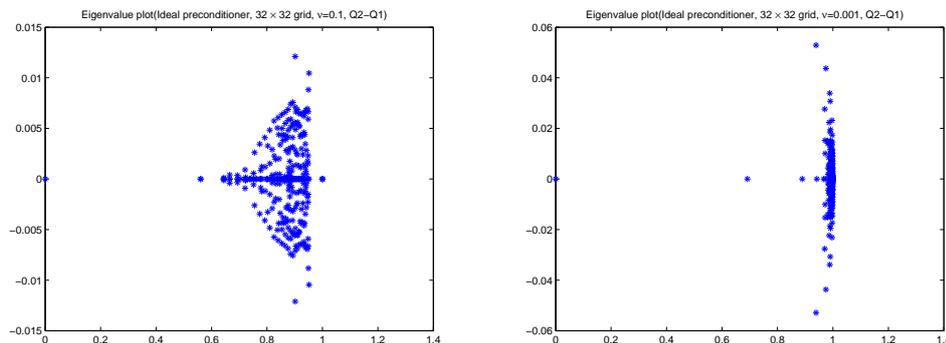
turbulent for  $\nu < 0.001$  in this case [37, Section 7.1] and consequently it does not make sense to compute a steady solution. Streamline and pressure plots of the approximate solutions of the Navier–Stokes equations (computed with IFISS) for the lid driven cavity and backward facing step problems are shown in Figures 2.1–2.2, respectively. We refer to the IFISS documentation [35] and to [37] for a detailed description of these test problems.

GMRES iteration counts with the ideal AL preconditioner are shown in Tables 2.2, 2.3 and 2.4. From these results we can see that the performance of the ideal AL preconditioner is independent of both the mesh size  $h$  and the viscosity  $\nu$ . The eigenvalues of the preconditioned matrices corresponding to the lid driven cavity problem with  $\nu = 0.1$  and  $\nu = 0.001$  discretized by Q2-

Table 2.4: GMRES(50) iterations with ideal AL preconditioner (step, Q2-Q1, uniform grids).

Viscosity	0.1	0.01	0.005
Grid	$\gamma = 1$		
$16 \times 16$	8	7	7
$32 \times 32$	7	6	6
$64 \times 64$	6	5	5
$128 \times 128$	6	5	5

Figure 2.3: Plots of the eigenvalues of the Oseen matrix with ideal AL preconditioner (cavity, Q2-Q1, uniform  $32 \times 32$  grid). Left:  $\nu = 0.1$ . Right:  $\nu = 0.001$ .



Q1 finite elements on uniform  $32 \times 32$  grid are plotted on the two figures in Figure 2.3. The non-zero eigenvalues are bounded away from 0, and cluster near 1 (note the different scales used for the horizontal and vertical axes). The only 0 eigenvalue comes from the hydrostatic pressure mode, which makes the saddle point system singular. Note that for  $\nu$  smaller, the imaginary parts of the eigenvalues are slightly larger, as already shown in Table 2.1. The strong clustering of the eigenvalues is in agreement with the fast convergence of the

Table 2.5: Inexact inner solvers. GMRES(50) iterations and timings with ideal AL preconditioner (cavity, Q2-Q1, uniform grids).

Viscosity	0.1	0.01	0.005	0.1	0.01	0.005
	MI20			ML		
$16 \times 16$	43	279	300*	300*	300*	300*
Setup time	0.24	0.05	-	-	-	-
Iter time	0.11	0.68	-	-	-	-
Total time	0.35	0.73	-	-	-	-
$32 \times 32$	45	298	300*	30	60	105
Setup time	0.12	0.04	-	0.24	0.22	0.24
Iter time	0.33	2.26	-	0.53	1.87	2.27
Total time	0.45	2.30	-	0.77	2.09	2.52
$64 \times 64$	45	253	300*	83	163	300*
Setup time	0.48	0.21	-	1.20	1.09	-
Iter time	6.09	11.78	-	10.47	19.88	-
Total time	6.57	11.99	-	11.67	20.98	-
$128 \times 128$	44	201	300*	-	-	-
Setup time	0.90	2.57	-	-	-	-
Iter time	8.24	72.30	-	-	-	-
Total time	9.14	74.87	-	-	-	-

preconditioned GMRES iteration.

In spite of the extreme robustness of the ideal AL preconditioner, solving  $A_\gamma$  exactly is not feasible even for medium-sized problems due to memory constraint and long computing time. Therefore we make an attempt to use algebraic multigrid methods (AMG) since these solvers are memory-efficient and much faster, and moreover, they are easier to apply to general discretizations and geometries than the geometric multigrid method used in [13]. Here we consider two implementations: MI20 described in [20, 21] and ML [42] in

the Trilinos package [51]. The former is written in Fortran while the latter is in C++; both can be called through a MATLAB interface. We perform one application of AMG with default settings of MI20, and default parameters with slight modification (full multigrid cycle with both pre- and post-smoothing and 2 sweeps for each) of ML to solve linear system with  $A_\gamma$  inexactly. The iteration counts and timings are presented in Table 2.5. The ‘Setup time’ includes the construction of the preconditioner, the ‘Iter time’ is the iterative phase of preconditioned GMRES, and the ‘Total time’ is the sum of the preceding two. All times are in seconds. The notation ‘300\*’ means that the total number of 300 GMRES iterations (6 cycles of GMRES(50)) is performed without reaching the desired stopping tolerance, and no report on iteration count means that MATLAB is crashed. We can observe that even for mild viscosity  $\nu = 0.1$  and  $0.01$ , MI20 and ML already have difficulty in solving  $A_\gamma$ : They may require a large number of iterations, or deteriorate as the grid becomes larger. For smaller viscosity  $0.005$ , neither converges within 300 iterations except for ML on the  $32 \times 32$  grid. We emphasize that these algebraic multigrid methods are designed for scalar PDEs rather than coupled system of PDEs, so it is not surprising that they perform poorly on  $A_\gamma$ , a set of two (three for 3D problems) coupled scalar PDEs. This motivates the proposal of the modified AL preconditioner in Section 2.4, which takes advantage of AMG and delivers excellent performance for a variety of problems.

## 2.4 Modified AL preconditioner

In the ideal AL-based preconditioner, linear systems associated with the augmented velocity sub-matrix  $A_\gamma$  were solved inexactly by means of a sophisticated geometric multigrid method for coupled systems of PDEs [13] similar to the one described in [69]. This solver is difficult to implement for problems

discretized using unstructured grids in complex geometries. This observation motivated the introduction in [15] of the *modified* AL preconditioner which could be implemented using standard algebraic multilevel solvers for scalar elliptic PDEs. The modified AL preconditioner can be described as follows. Recall that in 2D we have  $A = \text{diag}(A_1, A_2)$ , with each block  $A_i$  ( $i = 1, 2$ ) square and of order  $n/2$ , and  $B = (B_1, B_2)$ . Thus

$$\begin{aligned} A_\gamma &= A + \gamma B^T W^{-1} B \\ &= \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} + \gamma \begin{pmatrix} B_1^T \\ B_2^T \end{pmatrix} W^{-1} \begin{pmatrix} B_1 & B_2 \end{pmatrix} \\ &= \begin{pmatrix} A_1 + \gamma B_1^T W^{-1} B_1 & \gamma B_1^T W^{-1} B_2 \\ \gamma B_2^T W^{-1} B_1 & A_2 + \gamma B_2^T W^{-1} B_2 \end{pmatrix} \\ &=: \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \end{aligned}$$

Approximating  $A_\gamma$  with its block upper triangular part

$$\tilde{A}_\gamma = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

we define the modified AL preconditioner to be the block triangular matrix

$$\tilde{\mathcal{P}} = \begin{pmatrix} \tilde{A}_\gamma & B^T \\ 0 & \hat{S} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & B_1^T \\ 0 & A_{22} & B_2^T \\ 0 & 0 & \hat{S} \end{pmatrix}. \quad (2.11)$$

We note in passing that although the original construction of the modified AL preconditioner relied on the block diagonal structure of  $A$ , this approximation can be generalized to cases where  $A$  does not have block diagonal structure, as in the case of Newton linearization (in a way, it is even more natural for such problems).

### 2.4.1 Eigenvalue analysis

The block upper triangular structure of  $\tilde{\mathcal{P}}$  yields the following factorization of  $\tilde{\mathcal{P}}^{-1}$ :

$$\tilde{\mathcal{P}}^{-1} = \begin{pmatrix} \tilde{A}_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} I_n & B^T \\ 0 & -I_m \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & -\hat{S}^{-1} \end{pmatrix}.$$

Most of the work in computing the action of  $\tilde{\mathcal{P}}^{-1}$  on a vector lies in the solution of the two linear systems with coefficient matrices  $A_{11}$  and  $A_{22}$ . Observing that  $A_{ii} = A_i + \gamma B_i^T W^{-1} B_i$  ( $i = 1, 2$ ), in which the  $A_i$ 's represent discrete scalar convection-diffusion operators and the  $B_i$ 's are discretizations of the partial derivatives with respect to  $x$  and  $y$ , we immediately see that the  $A_{ii}$ 's can be interpreted as discrete scalar anisotropic convection-diffusion operators with diffusion anisotropy ratio  $\approx 1 + \gamma/\nu$ . Thus, applying  $A_{11}^{-1}$  and  $A_{22}^{-1}$  requires solving two scalar anisotropic convection-diffusion problems (three in 3D). These subsystems can be solved exactly or inexactly. For 2D problems of moderate sizes, sparse direct solvers can be used to solve the subsystems efficiently. For 3D problems, however, direct solvers become prohibitively expensive for sufficiently fine meshes, and algebraic multigrid methods should be used instead.

Applying right preconditioning to the coefficient matrix in (2.2) yields

$$\begin{aligned} \hat{\mathcal{A}}\tilde{\mathcal{P}}^{-1} &= \begin{pmatrix} A_\gamma & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \tilde{A}_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} I_n & B^T \\ 0 & -I_m \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & -\hat{S}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} A_\gamma \tilde{A}_\gamma^{-1} & -(A_\gamma \tilde{A}_\gamma^{-1} - I_n) B^T \hat{S}^{-1} \\ B \tilde{A}_\gamma^{-1} & -B \tilde{A}_\gamma^{-1} B^T \hat{S}^{-1} \end{pmatrix}. \end{aligned} \tag{2.12}$$

A simple but tedious calculation gives the result stated in the lemma below.

**Lemma 2.3.** *The right-preconditioned matrix has the following block struc-*

ture:

$$\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1} = \begin{pmatrix} I_{n/2} & 0 & 0 \\ A_{12}^T A_{11}^{-1} & I_{n/2} - A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} & T_{23} \\ B_1 A_{11}^{-1} & -B_1 A_{11}^{-1} A_{12} A_{22}^{-1} + B_2 A_{22}^{-1} & T_{33} \end{pmatrix}, \quad (2.13)$$

where  $T_{23} = (-A_{12}^T A_{11}^{-1} B_1^T + A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1}$  and  $T_{33} = -(B_1 A_{11}^{-1} B_1^T + B_2 A_{22}^{-1} B_2^T - B_1 A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1}$ .

*Proof.* First of all, we observe that the inverse of  $\widetilde{A}_\gamma$  is

$$\widetilde{A}_\gamma^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix}.$$

Thus, noticing that  $A_{12}^T = A_{21}$ , the (1,1) block of the right-hand side of (2.12) is

$$A_\gamma \widetilde{A}_\gamma^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ A_{12}^T A_{11}^{-1} & I - A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} \end{pmatrix}.$$

(Here and thereafter we omit the subscripts for the identity matrices for notational brevity when they are clear from the context.) The (1,2), (2,1) and (2,2) blocks of the right-hand side of (2.12) are, respectively,

$$\begin{aligned} -(A_\gamma \widetilde{A}_\gamma^{-1} - I) B^T \widehat{S}^{-1} &= \begin{pmatrix} 0 & 0 \\ -A_{12}^T A_{11}^{-1} & A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} \end{pmatrix} \begin{pmatrix} B_1^T \\ B_2^T \end{pmatrix} \widehat{S}^{-1} \\ &= \begin{pmatrix} 0 \\ (-A_{12}^T A_{11}^{-1} B_1^T + A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1} \end{pmatrix}, \\ B \widetilde{A}_\gamma^{-1} &= \begin{pmatrix} B_1 & B_2 \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} B_1 A_{11}^{-1} & -B_1 A_{11}^{-1} A_{12} A_{22}^{-1} + B_2 A_{22}^{-1} \end{pmatrix}, \\ -B \widetilde{A}_\gamma^{-1} B^T \widehat{S}^{-1} &= -\begin{pmatrix} B_1 & B_2 \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix} \begin{pmatrix} B_1^T \\ B_2^T \end{pmatrix} \widehat{S}^{-1} \\ &= -(B_1 A_{11}^{-1} B_1^T + B_2 A_{22}^{-1} B_2^T - B_1 A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1}. \end{aligned}$$

Plugging the above expressions into (2.12) gives (2.13).  $\square$

From (2.13) it follows immediately that  $\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1}$  has 1 as an eigenvalue of algebraic multiplicity at least  $n/2$ ; the remaining eigenvalues can be analyzed by focusing on the (2,2), (2,3), (3,2), (3,3) blocks of the right-hand side of (2.13). In order to proceed further, we need the following two lemmas.

**Lemma 2.4.** *The following two identities hold:*

$$\gamma B_1 A_{11}^{-1} B_1^T W^{-1} = I_m - (I_m + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}, \quad (2.14)$$

$$\gamma B_2 A_{22}^{-1} B_2^T W^{-1} = I_m - (I_m + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1}. \quad (2.15)$$

*Proof.* Applying the Sherman–Morrison–Woodbury matrix identity [43, page 50]:

$$(Y + UZV)^{-1} = Y^{-1} - Y^{-1}U(Z^{-1} + VY^{-1}U)^{-1}VY^{-1} \quad (2.16)$$

to  $A_{11}^{-1} = (A_1 + \gamma B_1^T W^{-1} B_1)^{-1}$  gives

$$A_{11}^{-1} = A_1^{-1} - A_1^{-1} B_1^T (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} B_1 A_1^{-1}.$$

Thus, we have

$$\begin{aligned} B_1 A_{11}^{-1} B_1^T &= B_1 (A_1^{-1} - A_1^{-1} B_1^T (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} B_1 A_1^{-1}) B_1^T \\ &= B_1 A_1^{-1} B_1^T (I - (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} B_1 A_1^{-1} B_1^T) \\ &= B_1 A_1^{-1} B_1^T (I - (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} (B_1 A_1^{-1} B_1^T + \gamma^{-1} W - \gamma^{-1} W)) \\ &= B_1 A_1^{-1} B_1^T (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} \gamma^{-1} W \\ &= (B_1 A_1^{-1} B_1^T + \gamma^{-1} W - \gamma^{-1} W) (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} \gamma^{-1} W \\ &= \gamma^{-1} W - \gamma^{-1} W (\gamma^{-1} W + B_1 A_1^{-1} B_1^T)^{-1} \gamma^{-1} W. \end{aligned}$$

Post-multiplying by  $\gamma W^{-1}$ , we obtain

$$\gamma B_1 A_{11}^{-1} B_1^T W^{-1} = I - (I + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}.$$

Similarly, one can show that

$$\gamma B_2 A_{22}^{-1} B_2^T W^{-1} = I - (I + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1}.$$

We complete the proof by observing that all the necessary inverses exist. Indeed, in the Oseen problem the sub-matrices  $A_1$  and  $A_2$  are positive definite, in the sense that they have positive definite symmetric part. Therefore,  $B_1 A_1^{-1} B_1^T$  and  $B_2 A_2^{-1} B_2^T$  are positive semi-definite. Since  $W$  is SPD and the product of an SPD matrix and a positive semi-definite one has eigenvalues with nonnegative real parts, matrices like  $I + \gamma B_1 A_1^{-1} B_1^T W^{-1}$  and  $\gamma^{-1} W + B_2 A_2^{-1} B_2^T$  are necessarily invertible.  $\square$

Applying (2.14) and (2.15) to (2.13), we obtain the result below.

**Lemma 2.5.** *Letting*

$$\begin{aligned} D &= \gamma B_2^T W^{-1} (I_m - (I_m + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}), \\ E &= (I_m + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1} B_2 A_2^{-1}, \\ F &= (I_m + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}, \\ G &= (I_m + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1}, \end{aligned}$$

*the right-preconditioned matrix can be written as*

$$\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1} = \begin{pmatrix} I_{n/2} & 0 & 0 \\ A_{12}^T A_{11}^{-1} & I_{n/2} - DE & -\gamma^{-1} D G W \widehat{S}^{-1} \\ B_1 A_{11}^{-1} & FE & -\gamma^{-1} (I_m - FG) W \widehat{S}^{-1} \end{pmatrix}. \quad (2.17)$$

*Proof.* First consider the (2,2) block  $I - A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1}$  of the right-hand side of (2.13). Using (2.14), we obtain

$$\begin{aligned} A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} &= \gamma B_2^T W^{-1} B_1 A_{11}^{-1} \gamma B_1^T W^{-1} B_2 A_{22}^{-1} \\ &= \gamma B_2^T W^{-1} (\gamma B_1 A_{11}^{-1} B_1^T W^{-1}) B_2 A_{22}^{-1} \\ &= \gamma B_2^T W^{-1} (I - (I + \gamma B_1 A_{11}^{-1} B_1^T W^{-1})^{-1}) B_2 A_{22}^{-1}. \end{aligned} \quad (2.18)$$

Applying again the Sherman–Morrison–Woodbury matrix identity to  $A_{22}^{-1}$  gives

$$\begin{aligned}
B_2 A_{22}^{-1} &= B_2 (A_2^{-1} - A_2^{-1} B_2^T (\gamma^{-1} W + B_2 A_2^{-1} B_2^T)^{-1} B_2 A_2^{-1}) \\
&= (I - B_2 A_2^{-1} B_2^T (\gamma^{-1} W + B_2 A_2^{-1} B_2^T)^{-1}) B_2 A_2^{-1} \\
&= (I - (-\gamma^{-1} W + \gamma^{-1} W + B_2 A_2^{-1} B_2^T) (\gamma^{-1} W + B_2 A_2^{-1} B_2^T)^{-1}) B_2 A_2^{-1} \\
&= \gamma^{-1} W (\gamma^{-1} W + B_2 A_2^{-1} B_2^T)^{-1} B_2 A_2^{-1} \\
&= (I + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1} B_2 A_2^{-1}.
\end{aligned} \tag{2.19}$$

Now, (2.18) and (2.19) together imply

$$\begin{aligned}
&A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} \\
&= \gamma B_2^T W^{-1} (I - (I + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}) (I + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1} B_2 A_2^{-1} \\
&= DE.
\end{aligned} \tag{2.20}$$

Second, for the (2,3) block in (2.13), we have

$$\begin{aligned}
&(-A_{12}^T A_{11}^{-1} B_1^T + A_{12}^T A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1} \\
&= (-A_{12}^T A_{11}^{-1} B_1^T + A_{12}^T A_{11}^{-1} (\gamma B_1^T W^{-1} B_2) A_{22}^{-1} B_2^T) \widehat{S}^{-1} \\
&= -\gamma B_2^T W^{-1} B_1 A_{11}^{-1} B_1^T (I - \gamma W^{-1} B_2 A_{22}^{-1} B_2^T) \widehat{S}^{-1} \\
&= \gamma B_2^T W^{-1} (\gamma B_1 A_{11}^{-1} B_1^T W^{-1}) (I - \gamma B_2 A_{22}^{-1} B_2^T W^{-1}) (-\gamma W^{-1})^{-1} \widehat{S}^{-1} \\
&= \gamma B_2^T W^{-1} (I - (I + \gamma B_1 A_1^{-1} B_1^T W^{-1})^{-1}) (I + \gamma B_2 A_2^{-1} B_2^T W^{-1})^{-1} (-\gamma W^{-1})^{-1} \widehat{S}^{-1} \\
&= DG (-\gamma W^{-1})^{-1} \widehat{S}^{-1}.
\end{aligned} \tag{2.21}$$

Next, using (2.14) and (2.19), the (3,2) block in (2.13) is

$$\begin{aligned}
& -B_1 A_{11}^{-1} A_{12} A_{22}^{-1} + B_2 A_{22}^{-1} \\
&= -B_1 A_{11}^{-1} \gamma B_1^T W^{-1} B_2 A_{22}^{-1} + B_2 A_{22}^{-1} \\
&= (I - \gamma B_1 A_{11}^{-1} B_1^T W^{-1}) B_2 A_{22}^{-1} \\
&= (I + \gamma B_1 A_{11}^{-1} B_1^T W^{-1})^{-1} (I + \gamma B_2 A_{22}^{-1} B_2^T W^{-1})^{-1} B_2 A_{22}^{-1} \\
&= FE.
\end{aligned} \tag{2.22}$$

Finally, let us investigate the (3,3) block of (2.13). We have

$$\begin{aligned}
& -(B_1 A_{11}^{-1} B_1^T + B_2 A_{22}^{-1} B_2^T - B_1 A_{11}^{-1} A_{12} A_{22}^{-1} B_2^T) \widehat{S}^{-1} \\
&= -(B_1 A_{11}^{-1} B_1^T + B_2 A_{22}^{-1} B_2^T - B_1 A_{11}^{-1} \gamma B_1^T W^{-1} B_2 A_{22}^{-1} B_2^T) (\gamma W^{-1}) (\gamma W^{-1})^{-1} \widehat{S}^{-1} \\
&= (I - (I - \gamma B_1 A_{11}^{-1} B_1^T W^{-1}) (I - \gamma B_2 A_{22}^{-1} B_2^T W^{-1})) (-\gamma W^{-1})^{-1} \widehat{S}^{-1} \\
&= (I - (I + \gamma B_1 A_{11}^{-1} B_1^T W^{-1})^{-1} (I + \gamma B_2 A_{22}^{-1} B_2^T W^{-1})^{-1}) (-\gamma W^{-1})^{-1} \widehat{S}^{-1} \\
&= (I - FG) (-\gamma W^{-1})^{-1} \widehat{S}^{-1}.
\end{aligned} \tag{2.23}$$

Substituting the expressions in (2.20), (2.21), (2.22) and (2.23) into (2.13) we arrive at (2.17).  $\square$

The foregoing lemma suggests replacing the choice of  $\widehat{S}^{-1}$  given in (2.4) with  $\widehat{S}^{-1} = -\gamma W^{-1}$ , leading to the following result.

**Theorem 2.6.** *Taking  $\widehat{S}^{-1} = -\gamma W^{-1}$ , we obtain*

$$\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1} = \begin{pmatrix} I_{n/2} & 0 & 0 \\ A_{12}^T A_{11}^{-1} & I_{n/2} - DE & DG \\ B_1 A_{11}^{-1} & FE & I_m - FG \end{pmatrix}. \tag{2.24}$$

Then  $\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1}$  has 1 as an eigenvalue of algebraic multiplicity at least  $n$ . The remaining eigenvalues are the non-unit eigenvalues of the matrix

$$\begin{pmatrix} I_{n/2} - DE & DG \\ FE & I_m - FG \end{pmatrix}.$$

*Proof.* Equality (2.24) immediately follows from (2.17). Furthermore, the spectrum of  $\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1}$  consists of the eigenvalue 1 of algebraic multiplicity  $n/2$ , plus the spectrum of the matrix

$$\begin{pmatrix} I_{n/2} - DE & DG \\ FE & I_m - FG \end{pmatrix} = \begin{pmatrix} I_{n/2} & 0 \\ 0 & I_m \end{pmatrix} - \begin{pmatrix} DE & -DG \\ -FE & FG \end{pmatrix}.$$

Observing that  $E = GB_2A_2^{-1}$ , we have

$$\begin{pmatrix} DE \\ -FE \end{pmatrix} = - \begin{pmatrix} -DG \\ FG \end{pmatrix} B_2A_2^{-1}.$$

Therefore, we obtain

$$\text{rank} \begin{pmatrix} DE & -DG \\ -FE & FG \end{pmatrix} \leq \text{rank} \begin{pmatrix} -DG \\ FG \end{pmatrix} \leq m,$$

because  $\begin{pmatrix} -DG \\ FG \end{pmatrix}$  has  $m$  columns. This implies that the matrix

$$\begin{pmatrix} I_{n/2} - DE & DG \\ FE & I_m - FG \end{pmatrix}$$

has the eigenvalue 1 of algebraic multiplicity at least  $n/2$ . Therefore,  $\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1}$  has 1 as an eigenvalue of algebraic multiplicity at least  $n$ .  $\square$

**Remark 2.7.** For the choice  $\widehat{S}^{-1} = -\gamma W^{-1}$ , we have

$$I_{n+m} - \widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ -A_{12}^T A_{11}^{-1} & DE & -DG \\ -B_1 A_{11}^{-1} & -FE & FG \end{pmatrix}.$$

Hence, a sufficient condition for all the eigenvalues of  $\widehat{\mathcal{A}}\widetilde{\mathcal{P}}^{-1}$  to be clustered around 1 is that  $\|D\|$ ,  $\|E\|$ ,  $\|F\|$  and  $\|G\|$  be as small as possible. For the

ideal AL preconditioner (2.3), all the eigenvalues of the preconditioned matrix cluster around 1 as  $\gamma \rightarrow \infty$ , cf. Theorem 2.1. However, this is not generally the case for the modified variant (2.11). Recalling the definitions of  $D$ ,  $E$ ,  $F$  and  $G$  in Lemma 2.5, we find that

$$\begin{aligned} \lim_{\gamma \rightarrow 0^+} \|D\| = 0, \quad \lim_{\gamma \rightarrow 0^+} \|E\| = \|B_2 A_2^{-1}\|, \quad \lim_{\gamma \rightarrow 0^+} \|F\| = 1, \quad \lim_{\gamma \rightarrow 0^+} \|G\| = 1; \\ \lim_{\gamma \rightarrow \infty} \|D\| = \infty, \quad \lim_{\gamma \rightarrow \infty} \|E\| = 0, \quad \lim_{\gamma \rightarrow \infty} \|F\| = 0, \quad \lim_{\gamma \rightarrow \infty} \|G\| = 0. \end{aligned}$$

Hence, letting  $\gamma \rightarrow 0^+$  or  $\gamma \rightarrow \infty$  does not make  $\|D\|$ ,  $\|E\|$ ,  $\|F\|$  and  $\|G\|$  simultaneously small. Furthermore, the above limits suggest that if an optimal  $\gamma$  exists for the modified AL preconditioner, it will likely be neither very large nor very small.

Furthermore, using field-of-values analysis, it has been proved that GMRES with the modified AL preconditioner converges independently of  $h$ .

**Theorem 2.8** ([14]). *For any  $\nu > 0$ ,  $\sigma \geq 0$ , sufficiently small  $\gamma > 0$  and  $\|I - A_\gamma \tilde{A}_\gamma^{-1}\|_{A_\gamma^{-1}}$ , the preconditioned GMRES with the modified AL preconditioner satisfies the convergence estimate*

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

where  $r^k$  is the GMRES residual of the  $k$ -th step, and  $\tau < 1$  is independent of  $h$ .

#### 2.4.2 Estimation of the optimal augmentation parameter $\gamma$ using Fourier analysis

The simple argument in Remark 2.7 suggests that setting  $\gamma$  to be some arbitrarily small or large number will likely result in sub-optimal performance of the preconditioner. In [15] an empirical rule based on numerical evidence has been stated and discussed. Details about it and corresponding results will

Table 2.6: The correspondence between the operators and their symbols.

Operator	Symbol
$L_x$	$2 - e^{i2\pi h\theta_x} - e^{-i2\pi h\theta_x}$
$L_y$	$2 - e^{i2\pi h\theta_y} - e^{-i2\pi h\theta_y}$
$N_x$	$e^{i2\pi h\theta_x} - e^{-i2\pi h\theta_x}$
$N_y$	$e^{i2\pi h\theta_y} - e^{-i2\pi h\theta_y}$
$S_x$	$h(1 - e^{-i2\pi h\theta_x})$
$S_y$	$h(1 - e^{-i2\pi h\theta_y})$
$W$	$h^2$

be presented in Section 2.5.1. Here, we resort to a Fourier analysis (FA) for guiding in the choice of  $\gamma$ ; see, e.g., [79]. As usual with this technique, some rather drastic simplifications and assumptions on the problem are needed. We assume that the Oseen problem (1.6)–(1.8) has constant coefficients, is defined on the unit square with periodic boundary conditions, and is discretized on a uniform  $l \times l$  grid with  $h = 1/l$ . Moreover, the matrices  $A_1$ ,  $A_2$ ,  $B_1$ ,  $B_2$  and  $W = \widehat{M}_p$  are assumed to be all square (of the same order) and to commute with one another. Though these assumptions are virtually never met in practice, they are ‘almost’ true (at least locally) away from the boundary and, as we shall see, they have remarkable heuristic value. Indeed, the value of  $\gamma$  obtained by Fourier analysis can be used even in problems that are defined in non-square domains, are discretized on stretched grids and do not have constant coefficients and periodic boundary conditions; see also [33] for a discussion in a similar context.

Under the assumptions above, we can replace  $A_1$ ,  $A_2$ ,  $B_1$  and  $B_2$  with the symbols of the corresponding operators. Indeed,  $A_1$  and  $A_2$  represent copies of the discrete 2D convection-diffusion operator

$$\mathcal{A} = I_l \otimes (\nu L_x + N_x) + (\nu L_y + N_y) \otimes I_l,$$

where  $\otimes$  denotes the Kronecker (tensor) product,  $I_l$  is the  $l \times l$  identity matrix,  $L_x$  and  $L_y$  are discrete one-dimensional (1D) Laplacians and  $N_x$  and  $N_y$  are discrete 1D convection operators in the  $x$  and  $y$  directions, respectively. Similarly, denoting the discretizations of the ordinary derivatives  $\frac{d}{dx}$  and  $\frac{d}{dy}$  by  $S_x$  and  $S_y$ , respectively, the matrices  $B_1$  and  $B_2$  represent discrete partial derivatives with respect to  $x$  and  $y$ , i.e.,  $B_1 = I_l \otimes S_x$  and  $B_2 = S_y \otimes I_l$ . Finally, let  $\theta_x, \theta_y$  denote integers running over the values  $1, 2, \dots, l$ . Then discretizing diffusion and convection terms by centered differences and the divergence (and gradient) by one-sided differences and observing that  $W = \widehat{M}_p$  scales as  $h^2$  give the correspondence between the operators and their symbols reported in Table 2.6. Note that the symbols respect the scaling of the matrices discretized by finite element methods. Then  $L_x, L_y, N_x, N_y, S_x$  and  $S_y$  can be expressed as diagonal matrices, whose diagonal entries are the corresponding eigenvalues. Hence,  $A_1, A_2, B_1$  and  $B_2$  can be represented by the symbols in Table 2.6 as well. More specifically, we express  $B_1 A_1^{-1} B_1^T W^{-1}$ ,  $B_2 A_2^{-1} B_2^T W^{-1}$ ,  $B_1^T W^{-1}$  and  $B_2 A_2^{-1}$  as diagonal matrices  $D_1 = \text{diag}(d_1)$ ,  $D_2 = \text{diag}(d_2)$ ,  $D_3 = \text{diag}(d_3)$  and  $D_4 = \text{diag}(d_4)$ . For ease of notation, we use  $d_i$  to denote the generic diagonal entry of  $D_i$  ( $i = 1 : 4$ ) though there are actually  $l^2$  such entries. Then the block matrix  $\begin{pmatrix} DE & -DG \\ -FE & FG \end{pmatrix}$  can be represented by

$$H =: \begin{pmatrix} (I - (I + \gamma D_1)^{-1})(I + \gamma D_2)^{-1} \gamma D_2 & \gamma D_3 (I - (I + \gamma D_1)^{-1})(I + \gamma D_2)^{-1} \\ (I + \gamma D_1)^{-1}(I + \gamma D_2)^{-1} D_4 & (I + \gamma D_1)^{-1}(I + \gamma D_2)^{-1} \end{pmatrix}.$$

All the 4 blocks of  $H$  are diagonal, so the spectrum of  $H$  can be represented by the eigenvalues of the  $2 \times 2$  matrix symbol

$$\widetilde{H} =: \begin{pmatrix} (1 - (1 + \gamma d_1)^{-1})(1 + \gamma d_2)^{-1} \gamma d_2 & \gamma d_3 (1 - (1 + \gamma d_1)^{-1})(1 + \gamma d_2)^{-1} \\ (1 + \gamma d_1)^{-1}(1 + \gamma d_2)^{-1} d_4 & (1 + \gamma d_1)^{-1}(1 + \gamma d_2)^{-1} \end{pmatrix}.$$

Observing that  $d_3d_4 = d_2$ , the eigenvalues of  $\tilde{H}$  are 0 and

$$\lambda = \lambda(\gamma) = \frac{1 + \gamma^2 d_1 d_2}{1 + \gamma d_1 + \gamma d_2 + \gamma^2 d_1 d_2}.$$

Here the  $d_i$ 's ( $i = 1 : 4$ ) run over all the diagonal entries of the  $D_i$ 's ( $i = 1 : 4$ ). Moreover,  $d_1$  and  $d_2$  are the symbols of  $B_1 A_1^{-1} B_1^T W^{-1}$  and  $B_2 A_2^{-1} B_2^T W^{-1}$ , so they depend on  $\theta_x$  and  $\theta_y$ . Recall now that for a clustered spectrum of the preconditioned matrix, we want the eigenvalues of  $H$  to be as small as possible. This suggests that the augmentation parameter  $\gamma$  should be chosen as the solution of the following (non-convex, non-smooth) optimization problem:

$$\begin{aligned} \min_{\gamma > 0} \quad & \text{mean} |(\lambda(\gamma; \theta_x, \theta_y))| \\ \text{where} \quad & \theta_x, \theta_y = 1, 2, \dots, l, \end{aligned}$$

where mean represents the average value. Since the expression of  $\lambda$  as a function of  $\gamma$  is very complicated and the optimization problem has several local minima, we find an approximate global minimum as follows. We sample  $\gamma$  over the range of values from 0.001 to 1 with step 0.001 (numerical experiments show that the optimal  $\gamma$  lies in the interval  $[0.001, 1]$  for all interesting problem parameters), and simply select the one that minimizes the arithmetic mean of the values of  $|\lambda|$  for  $\theta_x, \theta_y = 1, 2, \dots, l$ . Note that this approach imposes little overhead compared with the cost of solving the Oseen problem. In practice we also include  $L$ , the length of the domain where the problem is defined, to obtain the 1D convection-diffusion operators  $\nu L_x + L N_x$  and  $\nu L_y + L N_y$  in order to take into account the fact that the problem is generally not posed on the unit square.

## 2.5 Numerical experiments

In this section we present the results of numerical experiments with the modified AL preconditioner. In particular, we evaluate the empirical rule (discovered in [15] and discussed in details below) and the Fourier analysis-based approach for choosing  $\gamma$  by comparing GMRES iterations preconditioned by the modified AL preconditioners with the value of  $\gamma$  obtained by the empirical rule, by Fourier analysis and with the optimal  $\gamma$  (determined experimentally). The test problems are those described in Section 2.3. We set  $W = \widehat{M}_p = \text{diag}(M_p)$ . Right preconditioning is used in all cases; the results for left preconditioning are similar. All exact solves are performed on the subproblems involving  $A_{11}$  and  $A_{22}$  by means of sparse LU factorizations preceded by an approximate minimum degree column reordering [2] of the matrices  $A_{11}$ ,  $A_{22}$ .

### 2.5.1 Empirical rule for choosing $\gamma$

We now turn to the empirical rule [15] of for choosing  $\gamma$  when using the modified AL preconditioner. The empirical rule is as follows: when  $\nu = 0.1$  the modified AL preconditioner is not sensitive to  $\gamma$ , and taking  $\gamma$  as a constant turns out to be a good choice; for other values of the viscosity, the experimentally found optimal  $\gamma$  for the coarsest grid is divided by a factor of about  $\sqrt{2}$  when the mesh size  $h$  is divided by 2. Hence, for fixed  $\nu$ , letting  $\gamma_0$  be the  $\gamma$  for the coarsest grid, we use  $\gamma_0/\sqrt{2}$ ,  $\gamma_0/2$  and  $\gamma_0/2\sqrt{2}$  for each refined grid. In Table 2.7 we compare the values of  $\gamma$  obtained by the empirical rule with the optimal ones for the Q2-Q1 discretization of the cavity problem on uniform grids. As one can see, the empirical values are fairly close to the optimal ones.

We present GMRES iterations with the modified AL preconditioner in Tables 2.8, 2.9 and 2.10 for the three test problems mentioned in Section 2.3.

Table 2.7: Values of  $\gamma$  obtained by empirical rule vs. optimal values (cavity, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	Em	Opt	Em	Opt	Em	Opt	Em	Opt
$16 \times 16$	0.45	0.45	0.085	0.085	0.068	0.068	0.063	0.063
$32 \times 32$	0.45	0.38	0.060	0.050	0.048	0.043	0.045	0.035
$64 \times 64$	0.45	0.32	0.043	0.045	0.034	0.032	0.031	0.022
$128 \times 128$	0.45	0.28	0.030	0.046	0.024	0.032	0.022	0.017

Table 2.8: GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	Em	Opt	Em	Opt	Em	Opt	Em	Opt
$16 \times 16$	9	9	12	12	15	15	23	23
$32 \times 32$	10	9	12	11	14	14	30	29
$64 \times 64$	10	9	10	11	13	13	30	27
$128 \times 128$	9	9	10	10	13	12	26	24

Table 2.9: GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, stretched grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	Em	Opt	Em	Opt	Em	Opt	Em	Opt
$16 \times 16$	9	9	11	11	13	13	20	20
$32 \times 32$	9	9	11	11	14	14	23	23
$64 \times 64$	10	8	11	11	14	14	25	25
$128 \times 128$	9	7	11	11	13	13	26	26

‘Em’ denotes the  $\gamma$  chosen by the empirical rule, while ‘Opt’ stands for the optimal  $\gamma$ . It is clear that the empirical rule works remarkably well in all

Table 2.10: GMRES(50) iterations with modified AL preconditioner (step, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
Grid	Em	Opt	Em	Opt	Em	Opt
$16 \times 48$	12	12	16	16	19	19
$32 \times 96$	12	12	17	17	20	20
$64 \times 192$	12	11	16	16	19	19
$128 \times 384$	12	11	15	15	19	19

these cases.

Although the empirical rule is quite successful in predicting the values of  $\gamma$  for all the test problems, it could be impractical since the problem being solved may not be discretized on a coarse grid. Therefore we need some more convenient and efficient method, such as the Fourier analysis-based approach since it only requires an estimate of the mesh size and is much faster.

### 2.5.2 Fourier analysis-based approach for choosing $\gamma$

In Table 2.11 we compare the values of  $\gamma$  obtained by Fourier analysis with the optimal ones for the Q2-Q1 discretization of the cavity problem on uniform grids. As one can see, the Fourier estimates are fairly close to the optimal ones when the viscosity is not too small; for smaller viscosities, the estimated values tend to approach the optimal value as the mesh is refined.

In Tables 2.12, 2.13 and 2.14, we present preconditioned GMRES iteration counts for the three test problems already mentioned. ‘FA’ denotes the  $\gamma$  chosen by Fourier analysis, while ‘Opt’ stands for the optimal  $\gamma$ . For the Oseen problem on stretched grids (Table 2.13), we estimate the value of  $\gamma$  by Fourier analysis using an ‘average’ mesh size  $h$  defined as  $h =: 2/m$  where  $m$  is the number of grid points in the x-direction (or y-direction; recall that

Table 2.11: Values of  $\gamma$  obtained by Fourier analysis vs. optimal values (cavity, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	0.42	0.45	0.075	0.085	0.270	0.068	0.220	0.063
$32 \times 32$	0.29	0.38	0.056	0.050	0.098	0.043	0.067	0.035
$64 \times 64$	0.32	0.32	0.055	0.045	0.032	0.032	0.037	0.022
$128 \times 128$	0.28	0.28	0.036	0.046	0.022	0.032	0.020	0.017

Table 2.12: GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	9	9	12	12	26	15	42	23
$32 \times 32$	10	9	11	11	20	14	37	29
$64 \times 64$	9	9	11	11	13	13	33	27
$128 \times 128$	9	9	10	10	13	12	25	24

Table 2.13: GMRES(50) iterations with modified AL preconditioner (cavity, Q2-Q1, stretched grids).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	9	9	11	11	21	13	35	20
$32 \times 32$	9	9	11	11	17	14	31	23
$64 \times 64$	8	8	11	11	14	14	29	25
$128 \times 128$	8	7	11	11	14	13	26	26

here  $\Omega = [-1, 1] \times [-1, 1]$ ), which is actually the same as the value on the corresponding uniform grid. This strategy turns out to work quite well in

Table 2.14: GMRES(50) iterations with modified AL preconditioner (step, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
Grid	FA	Opt	FA	Opt	FA	Opt
$16 \times 48$	15	12	46	16	59	19
$32 \times 96$	12	12	24	17	38	20
$64 \times 192$	12	11	17	16	26	19
$128 \times 384$	11	11	15	15	19	19

practice.

Except for the case of coarse grids and very small viscosity, which is in any case irrelevant since physical solutions cannot be computed on such coarse grids, the iteration counts for  $\gamma$  chosen by FA are very close to those with optimal  $\gamma$ , no matter which problem is solved and how small  $\nu$  is, demonstrating a wide applicability. We also observe that the convergence rate of GMRES with either choice of  $\gamma$  is almost grid-independent and is only mildly dependent on  $\nu$ , confirming the robustness of the method. An especially noteworthy feature of the preconditioner is the excellent behavior on stretched meshes, in some cases even better than for uniform meshes (see Table 2.13). This could be due to the fact that stretched meshes lead to more accurate approximations for problems with small viscosity, hence they better reflect the underlying physics.

The dependence of the number of iterations on the value of  $\gamma$  is shown in Figure 2.4 for the case of the lid driven cavity Oseen problem with  $\nu = 0.01$  and  $\nu = 0.001$  discretized with Q2-Q1 elements on two uniform grids. From these plots we can see that the rate of convergence is more sensitive to the value of  $\gamma$  when the viscosity is small. In Figure 2.5, we display the eigenvalues of the preconditioned Oseen matrix for the lid driven cavity problem (with  $\nu = 0.01$ ) discretized with Q2-Q1 elements on a uniform  $32 \times 32$  grid.

Figure 2.4: Number of iterations vs. parameter  $\gamma$  (cavity, Q2-Q1, uniform grids). Left:  $\nu = 0.01$ . Right:  $\nu = 0.001$ .

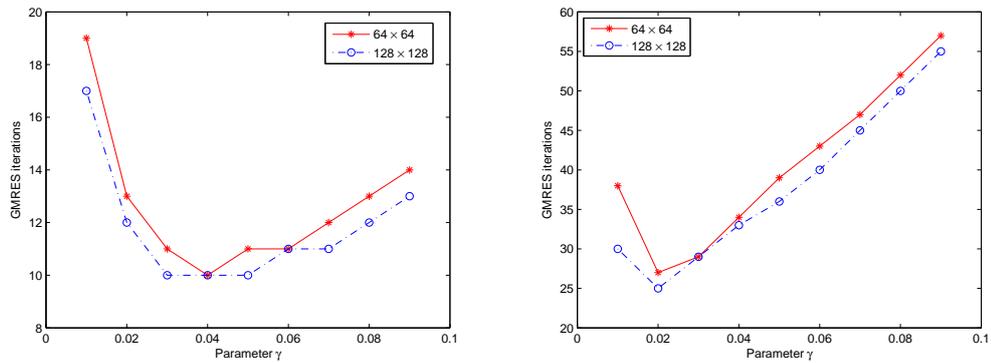
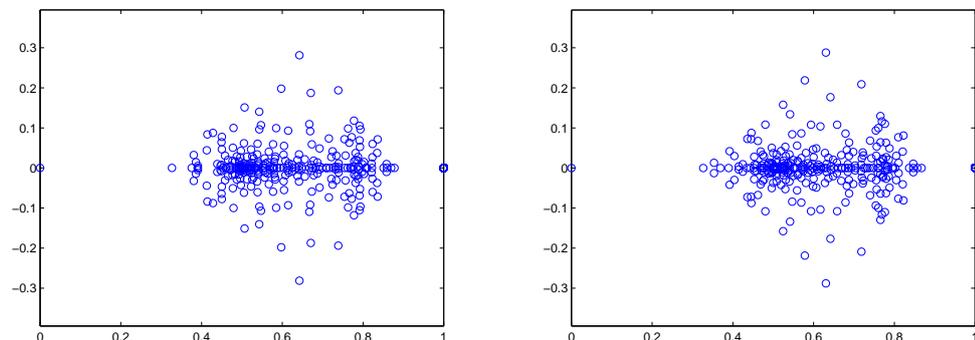


Figure 2.5: Plots of the eigenvalues of the Oseen matrix with modified AL preconditioner (cavity, Q2-Q1, uniform  $32 \times 32$  grid,  $\nu = 0.01$ ). Left: with  $\gamma$  chosen by Fourier analysis. Right: with optimal  $\gamma$ .



On the left we show the eigenvalues for the value of  $\gamma$  chosen by Fourier analysis, on the right those for the optimal  $\gamma$ . These values are, respectively,  $\gamma = 0.05$  and  $\gamma = 0.056$ . The spectra are almost identical, and nicely clustered away from the origin except for the zero eigenvalue corresponding to the hydrostatic pressure mode, which has no effect on the convergence of GMRES. Note also the eigenvalue at 1 (of algebraic multiplicity 2178).

Table 2.15: GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q2-Q1, uniform grids). The asterisk means that the FA values of  $\gamma$  are the ones found for the Oseen problem.

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA*	Opt	FA*	Opt	FA*	Opt	FA*	Opt
$16 \times 16$	13	13	22	21	39	29	138	50
$32 \times 32$	14	13	23	23	37	33	103	92
$64 \times 64$	14	14	23	23	35	35	98	98
$128 \times 128$	15	14	25	23	39	35	99	98

Table 2.16: GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q2-Q1, stretched grids). The asterisk means that the FA values of  $\gamma$  are the ones found for the Oseen problem.

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA*	Opt	FA*	Opt	FA*	Opt	FA*	Opt
$16 \times 16$	13	13	21	21	30	25	99	62
$32 \times 32$	14	14	23	23	31	30	71	60
$64 \times 64$	14	14	24	23	35	33	84	72
$128 \times 128$	15	14	26	23	40	34	95	82

Additionally, we consider the solution of Newton linearization for the three test problems. The purpose of these experiments is to show that although the modified AL-based preconditioner was designed having in mind the Picard form of linearized equations, it performs well also for the Newton linearization. When solving the lid driven cavity problem, for  $\nu = 0.005$ , we initialize the Newton iteration with the solution computed by one Picard iteration; for  $\nu = 0.001$ , three Picard iterations are needed to provide Newton's method with a sufficiently good initial guess. The results are shown in Table 2.15, 2.16 and 2.17. Here, we use the same value of  $\gamma$  chosen by FA for the Pi-

Table 2.17: GMRES(50) iterations with modified AL preconditioner (step, Newton, Q2-Q1, uniform grids). The asterisk means that the FA values of  $\gamma$  are the ones found for the Oseen problem.

Viscosity	0.1		0.01		0.005	
Grid	FA*	Opt	FA*	Opt	FA*	Opt
$16 \times 48$	21	18	63	27	81	30
$32 \times 96$	18	18	42	30	63	34
$64 \times 192$	18	19	35	32	52	37
$128 \times 384$	19	19	34	33	40	38

card linearization. From the GMRES iteration counts, we can observe that Fourier analysis still gives very good estimates for the optimal parameter values. Although with  $\gamma$  chosen by FA  $h$ -independent convergence is not retained in some cases for the cavity problem on stretched grids, we still get good convergence rates (it should be kept in mind that the linear systems from Newton linearization are considerably harder to solve than the Oseen problem.) Moreover, with the optimal  $\gamma$ , grid-independent convergence rate is maintained (except for  $\nu = 0.001$  in Table 2.16, where some deterioration occurs), again demonstrating the robustness of the augmented Lagrangian-based approach.

### 2.5.3 Comparison with other preconditioners

Next, we compare the modified AL preconditioner with some of the best preconditioners available in the literature, namely, the least-squares commutator (LSC) preconditioner, the pressure convection-diffusion preconditioner (PCD) [37], and its modified variant (mPCD) introduced in [38]. For the experiments, we use the implementations of these methods found in IFISS. It should be kept in mind that in IFISS, no restarting is used with GMRES,

Table 2.18: Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Q2-Q1, uniform grids).

Viscosity	0.005			0.001		
Grid	PCD	LSC	mPCD	PCD	LSC	mPCD
$16 \times 16$	40	27	37	83	69	83
$32 \times 32$	41	31	35	106	85	97
$64 \times 64$	40	29	34	108	91	93
$128 \times 128$	36	33	34	92	67	71

Table 2.19: Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Q2-Q1, stretched grids).

Viscosity	0.005			0.001		
Grid	PCD	LSC	mPCD	PCD	LSC	mPCD
$16 \times 16$	36	28	51	67	49	81
$32 \times 32$	38	38	37	89	77	95
$64 \times 64$	38	54	38	101	109	93
$128 \times 128$	36	80	35	96	154	98

Table 2.20: Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Newton, Q2-Q1, uniform grids).

Viscosity	0.005			0.001		
Grid	PCD	LSC	mPCD	PCD	LSC	mPCD
$16 \times 16$	66	49	64	84	81	84
$32 \times 32$	74	53	64	215	178	196
$64 \times 64$	73	53	64	239	187	199
$128 \times 128$	71	57	64	226	169	180

therefore the results presented here are for full GMRES. Moreover, exact solves are done by backslash, which is rather inefficient since the correspond-

Table 2.21: Full GMRES iterations with PCD, LSC and modified PCD preconditioners (cavity, Newton, Q2-Q1, stretched grids).

Viscosity	0.005			0.001		
Grid	PCD	LSC	mPCD	PCD	LSC	mPCD
$16 \times 16$	56	42	64	83	76	82
$32 \times 32$	65	60	62	182	147	160
$64 \times 64$	67	89	62	226	216	208
$128 \times 128$	69	128	62	235	286	206

ing matrices are factored anew at each solve, rather than reusing the triangular factors computed after the first application of the preconditioners. For these reasons, we do not include timings for these preconditioners. For Oseen problems at small Reynolds numbers discretized on uniform grids, we found that all these methods display convergence rates comparable to those obtained with modified AL preconditioning. Hence, here we focus on harder problems. Results for the Oseen problem with small viscosity ( $\nu = 0.005$  and  $0.001$ ) are reported in Tables 2.18 and 2.19 for both uniform and stretched grids. Results for the Newton linearization with two values of the viscosity are shown in Tables 2.20 and 2.21.

Comparing these results with those for the modified AL preconditioner suggests that the latter is to be preferred when solving difficult problems with small viscosities. The exceptional robustness of the augmented Lagrangian-based approach, and particularly its ability to effectively cope with stretched grids, give it a clear advantage over the existing methods. We observe that a comparison in terms of iteration counts is meaningful, since all these preconditioners have comparable cost per iteration.

Table 2.22: Comparison of exact and inexact inner solvers. GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q2-Q1, uniform grids,  $\nu = 0.005$ ).

Grid	Picard		Newton	
	Exact	Inexact	Exact	Inexact
16 × 16	26	35	39	71
Setup time	0.03	0.02	0.02	0.01
Iter time	0.09	0.19	0.10	0.43
Total time	0.12	0.21	0.12	0.44
32 × 32	20	33	37	43
Setup time	0.15	0.15	0.15	0.07
Iter time	0.18	0.77	0.38	1.12
Total time	0.33	0.92	0.54	1.19
64 × 64	13	15	35	36
Setup time	1.93	0.27	1.95	0.26
Iter time	0.62	1.42	1.75	3.29
Total time	2.55	1.69	3.70	3.55
128 × 128	13	16	39	41
Setup time	34.90	1.23	34.34	1.20
Iter time	4.44	7.55	10.94	15.82
Total time	39.34	8.78	45.28	17.02
256 × 256	13	15	43	44
Setup time	856.74	5.17	673.29	5.26
Iter time	40.22	25.79	85.84	94.28
Total time	896.96	30.96	759.12	99.54

#### 2.5.4 Inexact solves

We also report on the effect of inexact solves within the modified AL preconditioner. Instead of using an exact sparse LU factorization for  $A_{11}$  and

$A_{22}$ , we now solve the corresponding linear systems inexactly using a single iteration of the algebraic multigrid method (AMG) implemented in the code MI20 and described in [20, 21]. For the smoother we use symmetric Gauss–Seidel. The parameter  $\gamma$  is again chosen by FA, ignoring the fact that the solves are inexact.

In Table 2.22 we show iteration counts, setup time for constructing the different preconditioners, iteration time for preconditioned GMRES, and total time (that is, the sum of the preceding two). All timings are in seconds. We do not include the time for estimating the optimal  $\gamma$  with Fourier analysis, which is small compared to the overall solution costs (e.g., under 2 seconds for the  $128 \times 128$  grid, with a simple code that has not been optimized). A few observations are in order. First, except for the two coarsest grids the iteration counts are almost unchanged when exact preconditioner solves are replaced with inexact ones. Hence, a single iteration of AMG with the chosen smoother is enough. In particular, the preconditioned iteration with inexact inner solves retains the very good convergence behavior of the exact variant as the mesh is refined. Second, and most importantly, the timings associated with exact solves (based on sparse LU factorization with approximate minimum degree reordering) are not scalable, whereas those with inexact solves show excellent scalability. Hence, the exact variant of the preconditioner is outperformed by the inexact variant already for a  $64 \times 64$  grid.

We note that in the inexact case, the preconditioner construction time for AMG could be further reduced by reusing the same setup for several Picard or Newton iterations. We performed some experiments to see how ‘freezing’ the AMG preconditioner affects the convergence of preconditioned GMRES, and we found that for  $\nu = 0.01$  or larger, the number of iterations increases only slightly (if at all), resulting in considerable savings in terms of time to solution. However, this strategy does not work well with smaller viscosities. In this case it seems to be necessary to update the AMG preconditioner at

Table 2.23: GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q2-Q1, uniform  $256 \times 256$  grid,  $\nu = 0.01$ )

No. of cores	2	4	8	16	32	64
Iterations	16	15	15	14	15	15
Setup time	5.72	2.91	1.57	0.90	0.77	0.44
Iter time	6.07	2.95	1.55	0.78	0.64	0.31
Total time	11.79	5.86	3.12	1.68	1.41	0.75

each Picard or Newton step.

### 2.5.5 Parallel results

In this subsection we show the results of a few numerical experiments on a computer cluster. The test problem is the steady 2D Oseen problem for the lid driven cavity. The value of the viscosity is  $\nu = 0.01$ , and  $\gamma$  is chosen by FA. Using Q2-Q1 finite elements on uniform  $256 \times 256$  grid results in a  $148,739 \times 148,739$  saddle point matrix with 60,341,640 nonzero entries. The preconditioner is the modified AL preconditioner, where the linear systems associated with the diagonal sub-blocks  $A_{ii}$  are solved by the smoothed aggregation AMG preconditioner ML. The parameters for ML are based on the default parameters for non-symmetric smoothed aggregation with some modifications. The linear solver is GMRES implemented in the Trilinos package AztecOO [50].

The experiments are performed on a cluster consisting of 32 nodes and 128 processor cores. Each node has 2 dual-core AMD 2.2 GHz Opteron CPUs and 4 GB RAM. The program is compiled and run with Open MPI. GMRES iteration counts and timings using 2, 4, 8, 16, 32 and 64 cores are shown in Table 2.23. The ‘Setup time’ includes matrix multiplication and addition for  $A_{11}$  and  $A_{22}$  as well as AMG setup time for them, and ‘Iter time’ is the

iterative phase of preconditioned GMRES. The ‘Total time’ is the sum of the previous two. The iteration numbers keep almost constant as the number of cores grows, and the timings show fairly good scalability up to 64 cores.

### 2.5.6 Unsteady problems

Here we present some results for unsteady (generalized) Oseen problems. In this type of problem the (1,1) block  $A$  of the saddle point matrix contains an additional term of the form  $\sigma M_u$  where  $\sigma$  is the reciprocal of the time step  $\Delta t$  and  $M_u$  is the velocity mass matrix. Linear systems of this type tend to be easier to solve than the ones arising in the steady case, since the presence of the additional positive definite term  $\sigma M_u$  makes the (1,1) block more diagonally dominant, especially when  $\Delta t$  is sufficiently small. On the other hand, many such systems (one for each time step) have to be solved during a simulation, so fast solvers are absolutely indispensable.

We have performed numerical experiments for the lid driven cavity problem. In our experiments we let  $\sigma = h^{-1}$  where  $h$  is the mesh size. The parameter  $\gamma$  is chosen by Fourier analysis as for steady problems. Similar to the well-known Cahouet–Chabard [24] preconditioner for the unsteady Stokes problem, we modify the (2,2) block of the preconditioner (implicitly defined by its inverse) by using

$$\widehat{S}^{-1} = -\gamma \widehat{M}_p^{-1} - \sigma (B \widehat{M}_u^{-1} B^T)^{-1} \quad (2.25)$$

instead of  $\widehat{S}^{-1} = -\gamma \widehat{M}_p^{-1}$ . Here  $\widehat{M}_u$  denotes the diagonal of the velocity mass matrix; moreover, the action of the inverse of the scaled discrete Laplacian  $B \widehat{M}_u^{-1} B^T$  is implemented exactly by backslash in MATLAB.

Iteration counts are shown in Tables 2.24 and 2.25 for the lid driven cavity discretized on uniform and stretched grids and Table 2.26 for the backward facing step problem on uniform grids. This preconditioner with either choice of  $\gamma$  is now quite robust and stable with respect to both  $\nu$  and  $h$ .

Table 2.24: GMRES(50) iterations with modified AL preconditioner (unsteady cavity, Q2-Q1, uniform grids, improved  $\widehat{S}^{-1}$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	9	8	5	5	6	6	6	6
$32 \times 32$	9	8	5	5	4	4	5	5
$64 \times 64$	9	8	6	5	4	4	4	4
$128 \times 128$	8	8	5	5	4	4	3	3

Table 2.25: GMRES(50) iterations with modified AL preconditioner (unsteady cavity, Q2-Q1, stretched grids, improved  $\widehat{S}^{-1}$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	8	8	7	8	6	6	5	5
$32 \times 32$	9	8	8	8	9	6	4	4
$64 \times 64$	8	8	6	6	6	6	5	4
$128 \times 128$	6	6	5	5	4	4	3	3

Table 2.26: GMRES(50) iterations with modified AL preconditioner (unsteady step, Q2-Q1, uniform grids, improved  $\widehat{S}^{-1}$ ).

Viscosity	0.1		0.01		0.005	
Grid	FA	Opt	FA	Opt	FA	Opt
$16 \times 48$	10	9	9	9	9	9
$32 \times 96$	10	9	9	9	11	9
$64 \times 192$	9	9	8	8	8	8
$128 \times 384$	9	9	8	8	8	8

It is important to point out that in actual unsteady flow calculations the number of iterations can be expected to be significantly less than reported

above, since one takes the solution from the previous time step as the initial guess. Indeed, using the solution at the previous time step typically results in an initial residual that is much smaller in norm than the residual corresponding to the zero initial guess, therefore leading to fewer iterations required to satisfy the same reduction in the relative size of the residual norm.

We make a few remarks about the performance of the classical Cahouet–Chabard preconditioner [24] applied to the unsteady Oseen problem. It turns out that without augmentation (that is, working with the original system (2.1) rather than the augmented one), this block triangular preconditioner is almost always more efficient than the AL-based approach, except for some cases with small viscosity and large time steps. This is mainly due to the fact that with augmentation, the linear systems associated with the velocity unknowns that must be solved (exactly or inexactly) at each application of the preconditioner become more complicated. Therefore, even though the Cahouet–Chabard preconditioner typically requires more iterations than the AL-based one, it is actually faster in terms of solution time.

It should be pointed out, however, that there are important situations where the augmentation term arises naturally, in which case the AL-based preconditioner will be a good candidate. We note that the algebraic augmentation is closely related to the so-called grad-div stabilization [62] of the Galerkin method for the incompressible Navier–Stokes equations. The stabilization is a commonly used one for those problems which require additional subgrid pressure modeling or enhanced mass conservation [57, 59]. It is also an important ingredient of some turbulence models [56]. The least square term added in such models gives rise to a matrix  $G$  with elements  $g_{ij} = \gamma \int_{\Omega} \operatorname{div} \psi_i \operatorname{div} \psi_j \, d\mathbf{x}$ , where the  $\psi_i$  are basis functions of a discrete (e.g., finite element) space. This matrix is added to the velocity (1,1) block and possesses a block structure and algebraic properties similar to those of  $\gamma B^T W^{-1} B$  (algebraic augmentation). Therefore, the solvers studied in this

thesis should be useful for handling algebraic systems resulting from such stabilized Galerkin discretizations. In [19], it was found that for a 2D Oseen problem with recirculating convection, Taylor–Hood finite element discretization with grad-div stabilization results in a sparser velocity matrix, whose number of nonzero entries is only approximately 1/7 of that obtained from explicit augmentation. Furthermore, a sparser matrix requires less setup times for, e.g.,  $\mathcal{H}$ -LU factorization [19] and AMG (see Section 2.5.9). Hence the ‘first-augment-then-discretize’ approach is preferred because of its efficiency. Yet another example is the formulation of the Navier–Stokes equations in the context of fluid-structure interaction problems, where it is essential that the velocity deformation tensor  $(\nabla \mathbf{u} + \nabla \mathbf{u}^T)/2$  is retained in its entirety in the equations in order to ensure good momentum conservation properties during the numerical solution. This point of view is adopted in, e.g., [39, Chapter 9]. We present numerical results on problems with the full tensor in Section 2.5.9.

## 2.5.7 Extension to 3D problems

In this subsection we evaluate the performance of the ideal and modified AL preconditioners for Stokes and Oseen problems in 3D. Besides the usual (convective) form of the Oseen equations, we also consider linearizations of the Navier–Stokes equations in rotation form; see, e.g., [11, 60] and references therein for details.

For a stable discretization, the saddle point system is again of the form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \quad \text{or} \quad \mathcal{A}x = b,$$

where for 3D problems  $A = \text{diag}(A_1, A_2, A_3)$  and  $B = (B_1, B_2, B_3)$ . Therefore, for the convection form the coefficient matrix of the equivalent aug-

mented Lagrangian formulation is

$$\begin{aligned}
A_\gamma &= A + \gamma B^T W^{-1} B \\
&= \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix} + \gamma \begin{pmatrix} B_1^T \\ B_2^T \\ B_3^T \end{pmatrix} W^{-1} \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} \\
&= \begin{pmatrix} A_1 + \gamma B_1^T W^{-1} B_1 & \gamma B_1^T W^{-1} B_2 & \gamma B_1^T W^{-1} B_3 \\ \gamma B_2^T W^{-1} B_1 & A_2 + \gamma B_2^T W^{-1} B_2 & \gamma B_2^T W^{-1} B_3 \\ \gamma B_3^T W^{-1} B_1 & \gamma B_3^T W^{-1} B_2 & A_3 + \gamma B_3^T W^{-1} B_3 \end{pmatrix} \\
&=: \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}.
\end{aligned}$$

The ideal AL preconditioner is again given by (2.3). In the modified variant we replace  $A_\gamma$  with the block triangular approximation

$$\begin{aligned}
\tilde{A}_\gamma &= \begin{pmatrix} A_1 + \gamma B_1^T W^{-1} B_1 & \gamma B_1^T W^{-1} B_2 & \gamma B_1^T W^{-1} B_3 \\ 0 & A_2 + \gamma B_2^T W^{-1} B_2 & \gamma B_2^T W^{-1} B_3 \\ 0 & 0 & A_3 + \gamma B_3^T W^{-1} B_3 \end{pmatrix} \\
&= \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix}.
\end{aligned}$$

Note that in the 3D case we drop three blocks: the (2,1), (3,1) and (3,2) blocks of  $A_\gamma$ , so the performance could be affected more than in the 2D case. As in the 2D case, each diagonal block  $A_{ii}$  ( $i = 1, 2, 3$ ) represents a discrete scalar convection-diffusion operator. For the Stokes problem and for the rotation form of the Navier–Stokes equations, no convective term is present and each sub-block  $A_{ii}$  ( $i = 1, 2, 3$ ) is symmetric and positive definite (SPD). On the other hand, in the rotation form each off-diagonal block  $A_{ij}$  (with  $i \neq j$ ) contains additional coupling terms not present in the standard form

Table 2.27: GMRES iterations with ideal and modified AL preconditioners (3D Stokes, MAC).

Grid	Ideal	Modified
$8 \times 8 \times 8$	8	11
$16 \times 16 \times 16$	9	13
$24 \times 24 \times 24$	10	14

of the Oseen problem. Finally, for unsteady problems an additional reaction term (also SPD) is present in each diagonal sub-block.

### 2.5.8 Numerical experiments: 3D examples

We use a Marker-and-Cell (MAC) scheme [49] to discretize the Stokes and Oseen problems on the unit cube  $\Omega = [0, 1] \times [0, 1] \times [0, 1]$ . This scheme is known to be div-stable, hence no pressure stabilization is needed and the (2,2) block  $C$  in the saddle point problem (1.16) is zero. Homogeneous Dirichlet boundary conditions are imposed on the velocity components.

In the first experiment, we compare the ideal and the modified AL preconditioners on a steady Stokes problem. We use  $\gamma = 1$  for both preconditioners for simplicity, and because it gives good results. Iteration counts for three grids of increasing size are shown in Table 2.27. Using a complete sparse Cholesky factorization on the sub-matrix  $A_\gamma$  makes the ideal AL preconditioner unfeasible already for rather coarse grids in 3D. On the other hand, the modified AL preconditioner is able to handle larger problems, since only the diagonal blocks  $A_{ii}$  ( $i = 1, 2, 3$ ) of  $A_\gamma$  need to be factored.

Our results indicate that for this problem, the rate of convergence with the ideal and modified AL preconditioners is essentially independent of the mesh size.

Table 2.28: GMRES iterations with ideal and modified AL preconditioners (3D Oseen in convection form, MAC).

Viscosity	0.1		0.01	
Grid	Ideal	Modified	Ideal	Modified
$8 \times 8 \times 8$	6	10 (0.4)	5	16 (0.06)
$16 \times 16 \times 16$	6	11 (0.4)	5	17 (0.06)
$24 \times 24 \times 24$	6	12 (0.4)	5	17 (0.06)
$32 \times 32 \times 32$	-	12 (0.4)	-	18 (0.06)

Next, we proceed to the more challenging case of 3D Oseen problems, in both convection and rotation forms. The divergence-free wind function ( $\mathbf{v}$  in (1.6)) is taken to be

$$\mathbf{v} = \begin{pmatrix} (2y - 1)x(1 - x) \\ (2x - 1)y(1 - y) \\ -2z(1 - 2x)(2y - 1) \end{pmatrix}.$$

The results are shown in Tables 2.28 and 2.29. The use of LU factorization makes the ideal preconditioner unfeasible on  $32 \times 32 \times 32$  or larger grids due to exceeded memory limits, which is shown as a '-'. We set  $\gamma = 1$  for the ideal preconditioner; for the modified one, the values of  $\gamma$  chosen by FA are too small, so we find optimal values experimentally, and show them in parentheses after the number of GMRES iterations.

While the ideal AL preconditioner is independent of the mesh size and viscosity, the modified preconditioner appears to be independent of  $h$  but shows a degradation of convergence rate when  $\nu$  becomes very small. Also notice that  $\gamma$  must be taken smaller when  $\nu = 0.01$ . Also for small  $\nu$  the modified preconditioner is less effective for the rotation form than for the convection form. This is due to the presence, in the rotation form, of relatively stronger

Table 2.29: GMRES iterations with ideal and modified AL preconditioners (3D Oseen in rotation form, MAC).

Viscosity	0.1		0.01	
Grid	Ideal	Modified	Ideal	Modified
$8 \times 8 \times 8$	7	12 (0.4)	5	23 (0.08)
$16 \times 16 \times 16$	7	14 (0.4)	5	22 (0.08)
$24 \times 24 \times 24$	8	14 (0.4)	5	22 (0.08)
$32 \times 32 \times 32$	-	15 (0.4)	-	23 (0.08)

Table 2.30: GMRES(50) iterations with ideal and modified AL preconditioners (unsteady 3D Oseen in convection form, MAC).

Viscosity	0.1		0.01	
Grid	Ideal	Modified	Ideal	Modified
$8 \times 8 \times 8$	7	12	7	12
$16 \times 16 \times 16$	8	13	7	11
$24 \times 24 \times 24$	9	13	8	11
$32 \times 32 \times 32$	-	13	-	11

coupling terms in the off-diagonal blocks  $A_{ij}$  ( $i > j$ ) that are neglected when forming the preconditioner.

Next, we consider unsteady Oseen problems with  $\sigma = h^{-1}$ . Because of the dominance of the block diagonal part of the (1,1) block of the saddle point system, we can expect good performance of the modified AL preconditioner. For the ideal preconditioner, we set  $\gamma = 1$ ; for the modified variant, FA gives good heuristic values of  $\gamma$ . The (2,2) block is given by

$$\widehat{S}^{-1} = -\gamma \widehat{M}_p^{-1} - \sigma (B \widehat{M}_u^{-1} B^T)^{-1}$$

Table 2.31: GMRES(50) iterations with ideal and modified AL preconditioners (unsteady 3D Oseen in rotation form, MAC).

Viscosity	0.1		0.01	
Grid	Ideal	Modified	Ideal	Modified
$8 \times 8 \times 8$	8	13	6	12
$16 \times 16 \times 16$	9	15	8	10
$24 \times 24 \times 24$	10	15	8	10
$32 \times 32 \times 32$	-	16	-	10

as for 2D problems. Iteration counts are shown in Tables 2.30 (convection form) and 2.31 (rotation form).

The performance of the modified AL preconditioner is quite satisfactory. Indeed, the iteration number is essentially independent of the mesh size and viscosity.

### 2.5.9 Parallel results: 3D examples

In this subsection we first show parallel results for the 3D Oseen problem in convection form discretized by MAC with viscosity  $\nu = 0.01$ . GMRES iteration counts and timings are shown in Table 2.32 when the problem is discretized on  $64 \times 64 \times 64$  and  $128 \times 128 \times 128$  grids. The augmentation parameter  $\gamma = 0.06$  is determined experimentally to minimize the GMRES iteration counts.

On the  $64 \times 64 \times 64$  grid, the coefficient matrix is  $1,036,288 \times 1,036,288$  with 8,442,624 nonzero elements. The ‘Setup time’ includes matrix multiplication and addition, that is, explicit formation of  $A_{11}$ ,  $A_{22}$  and  $A_{33}$ , and ML setup phase of these matrices. One can observe that the iteration counts do not increase and timings show almost perfect scalability. For the larger  $128 \times$

Table 2.32: GMRES(50) iterations and timings with modified AL preconditioner (3D Oseen, MAC,  $\nu = 0.01$ )

No. of cores	2	4	8	16	32	64
$64 \times 64 \times 64$	19	19	19	19	19	27
Setup time	8.68	5.03	3.82	1.94	1.66	1.60
Iter time	22.91	12.74	7.19	4.14	2.22	1.66
Total time	31.59	17.77	11.01	6.08	3.88	3.26
$128 \times 128 \times 128$	19	19	19	19	19	21
Setup time	78.70	44.96	23.65	14.30	9.23	8.25
Iter time	217.11	141.78	64.73	36.73	18.90	12.03
Total time	295.81	186.74	88.38	51.03	28.13	20.28

$128 \times 128$  grid, the coefficient matrix is now  $8,339,456 \times 8,339,456$  with 99,290,112 nonzero entries. Very good scalability is still achieved up to 32 cores. When using 64 cores, some improvement is obtained although not perfect.

Next we present GMRES iteration counts and timings in Table 2.33 for a linearized steady 3D lid driven cavity problem discretized by P2-P1 finite elements in the finite element library LifeV [1]. The viscosity  $\nu$  is 0.05. Here we retain the full tensor  $(\nabla \mathbf{u} + \nabla \mathbf{u}^T)/2$  so as to circumvent the explicit construction of  $A_{ii}$  ( $i = 1, 2, 3$ ). In fact, using the full tensor could save up to 5/6 of total time. Firstly, we observe that the iteration counts do not change essentially as the number of cores increase or the size of the mesh becomes larger, demonstrating the effectiveness of the modified AL preconditioner and the ML preconditioner for subproblems. In fact, larger grids provide better approximations to the continuous problem, and the number of iterations decreases. Secondly, good scalability is achieved for the  $32 \times 32 \times 32$  grid, but not for smaller grids when using less than 8 cores. Finally, we observe quite good weak scalability.

Table 2.33: GMRES(50) iterations and timings with modified AL preconditioner (3D cavity, P2-P1,  $\nu = 0.05$ )

No. of cores	2	4	8	16	32	64
$16 \times 16 \times 16$	25	25	24	25	23	23
Setup time	1.85	1.57	1.68	1.16	0.89	0.72
Iter time	5.35	3.68	3.06	1.94	0.95	0.67
Total time	7.20	5.25	4.74	3.10	1.84	1.39
$24 \times 24 \times 24$	24	23	23	23	23	22
Setup time	7.53	5.98	6.34	3.82	2.83	1.82
Iter time	18.80	11.61	10.32	6.14	4.11	2.04
Total time	25.33	17.59	16.66	9.96	6.94	3.86
$32 \times 32 \times 32$	22	21	20	20	20	19
Setup time	37.71	16.53	18.09	9.38	6.27	4.00
Iter time	67.88	29.24	25.41	12.88	8.11	4.33
Total time	105.59	45.77	43.50	22.26	14.38	8.33

# Chapter 3

## The stabilized case

In this chapter, we consider the Oseen problem discretized by stabilized finite elements, e.g., Q1-Q1 elements. We generalize the ideal and modified AL preconditioners to stabilized elements, aiming to achieve robustness with respect to  $h$  and  $\nu$ . We analyze the eigenvalues of the matrix preconditioned by the AL-based preconditioners for stabilized finite elements, use Fourier analysis to choose the augmentation parameter  $\gamma$ , and present numerical results.

### 3.1 Ideal AL preconditioners

From this section, we consider the Oseen problem discretized by stabilized finite elements, e.g., Q1-Q1 or Q1-P0 elements. We generalize the ideal and modified AL preconditioners to stabilized elements, aiming to achieve robustness with respect to  $h$  and  $\nu$ . We further analyze the eigenvalues of the matrix preconditioned by the AL-based preconditioners. As in Section 2.4.2, we use Fourier analysis to choose the augmentation parameter  $\gamma$ , and present the results of numerical experiments.

### 3.1.1 Problem formulation

Equal order finite element pairs, like Q1-Q1, are extensively used in the engineering community due to their ease of implementation and other advantages. However, this choice of finite element spaces does not meet the LBB condition, and pressure stabilization is required; see, e.g., [54]. In this case, the (2,2) block of the saddle point matrix is no longer zero; it is replaced by  $-C$  with some symmetric positive semi-definite matrix  $C = \delta \widehat{C}$ . Here  $\delta$  is a stabilization parameter, and  $\widehat{C}$  is a stabilization matrix for the corresponding Stokes problem.

For Q1-Q1 finite elements, pressure stabilization in [32] is used, and for Q1-P0 elements, the pressure jumps over all internal edges of the triangulation are penalized [22]. In either case the spectral properties of the matrix  $\widehat{C}$  are somewhat similar to those of a scaled Laplacian discretization. For more details on the construction of  $\widehat{C}$ , see [37]. We note that different pressure stabilization methods can be applied, including those based on residual-free bubbles [41], local projection [6], as well as the method in [54]. In general they lead to matrices  $\widehat{C}$  with similar algebraic properties.

We use the following choice of the stabilization parameter (see [29]):

$$\delta = \frac{\beta}{\nu + h\|\mathbf{v}\|},$$

where  $h$  is the mesh size,  $\mathbf{v}$  is the wind function, and  $\beta = 1$  for the Q1-Q1 finite element pair and  $1/4$  for Q1-P0. The resulting linear system reads

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \quad (3.1)$$

### 3.1.2 Augmented linear systems and ideal AL preconditioners

Owing to the presence of a nonzero (2,2) block, the augmentation of system (3.1) must be done differently than in the case of stable finite elements. As before, let  $\gamma > 0$  and let  $W$  be a symmetric positive definite matrix. Then from  $Bu - Cp = g$  it follows that

$$\gamma B^T W^{-1} Bu - \gamma B^T W^{-1} Cp = \gamma B^T W^{-1} Bg.$$

Adding the above equation to  $Au + B^T p = f$  gives

$$(A + \gamma B^T W^{-1} B)u + (B^T - \gamma B^T W^{-1} C)p = f + \gamma B^T W^{-1} Bg.$$

Therefore, the first (non-symmetric) augmented linear system is

$$\begin{pmatrix} A_\gamma & B_\gamma^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f_\gamma \\ g \end{pmatrix}, \quad \text{or} \quad \hat{A}x = \hat{b}, \quad (3.2)$$

where  $A_\gamma = A + \gamma B^T W^{-1} B$ ,  $B_\gamma^T = B^T - \gamma B^T W^{-1} C$  and  $f_\gamma = f + \gamma B^T W^{-1} Bg$ .

Notice that in (3.2), the (1,2) block  $B_\gamma^T$  is not equal to the transpose of the (2,1) block  $B$ . To get a more ‘symmetric’ augmented linear system, we can obtain from  $Bu - Cp = g$  the equation

$$-\gamma CW^{-1} Bu + \gamma CW^{-1} Cp = -\gamma CW^{-1} Bg.$$

Then, combining this equation with  $Bu - Cp = g$ , we have

$$(B - \gamma CW^{-1} B)u - (C - \gamma CW^{-1} C)p = g - \gamma CW^{-1} Bg.$$

Letting  $C_\gamma = C - \gamma CW^{-1} C$  and  $g_\gamma = g - \gamma CW^{-1} Bg$ , we obtain the second (symmetrized) augmented system

$$\begin{pmatrix} A_\gamma & B_\gamma^T \\ B_\gamma & -C_\gamma \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f_\gamma \\ g_\gamma \end{pmatrix}. \quad (3.3)$$

Numerical experiments suggest that applying the AL preconditioner to (3.2) produces almost the same results as to (3.3). Therefore, all results of numerical experiments will be shown for (3.2) in the sequel.

Similar to the stable finite elements case we build the block triangular preconditioner in the form

$$\mathcal{P} = \begin{pmatrix} A_\gamma & 0 \\ B & \widehat{S} \end{pmatrix} \quad (3.4)$$

for (3.2). Here in order to simplify the action of  $\mathcal{P}^{-1}$  we use a block lower triangular AL preconditioner so as to have  $B$  in the (2,1) block rather than the more cumbersome  $B_\gamma^T$  in the (1,2) block. For the ‘symmetrized’ system (3.3),  $B$  should be replaced by  $B_\gamma$ ; the corresponding AL preconditioner is

$$\mathcal{P} = \begin{pmatrix} A_\gamma & 0 \\ B_\gamma & \widehat{S} \end{pmatrix}. \quad (3.5)$$

Both lower and upper triangular preconditioners are essentially equivalent in this case.

For LBB-stable elements, in Section 2.2 we set  $\widehat{S} = -\gamma^{-1}W = -\gamma^{-1}\widehat{M}_p$ . As we will see, for the case of  $C \neq 0$  the choice of  $\widehat{S}$  and  $W$  is more delicate. From (2.5) one notices that  $-\widehat{S}$  intends to approximate the pressure Schur complement of the augmented system, i.e. the matrices  $S_\gamma =: B_\gamma A_\gamma^{-1} B_\gamma^T + C_\gamma$  for (3.3) and  $\widetilde{S}_\gamma =: B A_\gamma^{-1} B_\gamma^T + C$  for (3.2). Recall the notation for the pressure Schur complement matrix of the non-augmented problem:  $S = B A^{-1} B^T + C$ . The following result, which extends the representation in (2.6) to the case of  $C \neq 0$ , will help us to set  $W$ , build the preconditioner  $\widehat{S}$ , and analyze the spectrum of the preconditioned system.

**Lemma 3.1** ([15]). *Assuming all the relevant matrices are invertible, it holds*

$$S_\gamma^{-1} = S^{-1} + \gamma(W - \gamma C)^{-1}, \quad (3.6)$$

$$\widetilde{S}_\gamma^{-1} = S^{-1}(I - \gamma C W^{-1}) + \gamma W^{-1}. \quad (3.7)$$

*Proof.* The matrix  $X =: -S_\gamma^{-1}$  is the (2,2) block of the inverse of the coefficient matrix in (3.3). Denoting by  $Y$  the (1,2) block of this inverse matrix we get the following system of matrix equations:

$$(A + \gamma B^T W^{-1} B)Y + B^T (I - \gamma W^{-1} C)X = 0, \quad (3.8)$$

$$(I - \gamma C W^{-1})BY - (I - \gamma C W^{-1})CX = I. \quad (3.9)$$

From (3.9) we get  $BY = (I - \gamma C W^{-1})^{-1} + CX$ . Substituting this into (3.8) and applying  $A^{-1}$  lead to

$$Y = -\gamma A^{-1} B^T W^{-1} (I - \gamma C W^{-1})^{-1} - A^{-1} B^T X.$$

Now substituting  $Y$  to (3.9) gives, after simple manipulations,

$$-(BA^{-1}B^T + C)X(I - \gamma C W^{-1}) = I + \gamma BA^{-1}B^T W^{-1}.$$

By straightforward computations one verifies that the last equation is solved by matrix  $X = -(BA^{-1}B^T + C)^{-1} - \gamma(W - \gamma C)^{-1}$ . Thus (3.6) is proved. The result in (3.7) follows from the obvious identity  $(I - \gamma C W^{-1})\tilde{S}_\gamma = S_\gamma$ .  $\square$

The expressions (3.6) and (3.7) suggest that the auxiliary matrix  $W$  should be such that  $W - \gamma C$  is positive definite. Below we consider the following two choices of  $W$  satisfying this constraint:

$$W_1 =: M_p + \gamma C,$$

$$W_2 =: M_p \quad \text{with} \quad 0 < \gamma \leq (2\|M_p^{-1}C\|)^{-1}.$$

Similar to the LBB-stable case, in practical computations  $M_p$  is replaced by its diagonal approximation  $\widehat{M}_p$ . Let us briefly comment on both choices of  $W$ .

**Remark 3.2.** *Setting  $W = W_1$  will lead to a simple choice of preconditioner  $\widehat{S}$  such that the preconditioned system enjoys the same eigenvalue bounds*

as in the LBB-stable case, i.e. (2.8)–(2.9). At the same time,  $W = W_1$  involves the Laplacian-type matrix  $C$ . Hence the inverse  $W^{-1}$  may become an (almost) full matrix, so that  $A_\gamma = A + B^T W^{-1} B$  is an (almost) full matrix and consequently making the solution of linear systems with  $A_\gamma$  much more difficult. This happens for example with Q1-Q1 elements. For Q1-P0 elements, however, the matrix  $C$  has a special block-diagonal structure, which leads to a relatively cheap solve with  $A_\gamma$ . In practice we will use  $W_1$  only with Q1-P0 elements.

The choice  $W = W_2$  preserves the sparsity of  $A_\gamma$ . However, the restriction on  $\gamma$  yields a decrease of  $\gamma$  when  $\nu$  is small and  $h$  tends to zero, since for small  $\nu$  it holds  $\|M_p^{-1} C\| = O(h^{-1})$ ; see the discussion on matrix  $C$  in Section 3.1.1. Thus less augmentation is introduced and the performance of the solver becomes more sensitive to the variation in  $\nu$  and  $h$ .

In the next subsection we present the eigenvalue analysis and show the corresponding choices of  $\widehat{S}$  for both cases  $W = W_1$  and  $W = W_2$ . We shall also discuss a third (intermediate) alternative of setting the augmentation and preconditioning, which is not covered by our analysis, but shows stable and  $\nu$ - and  $h$ -independent convergence behavior while keeping the matrix  $A_\gamma$  sparse.

### 3.1.3 Eigenvalue analysis

To analyze the spectrum of the preconditioned matrix, we consider the following generalized eigenvalue problem:

$$\begin{pmatrix} A_\gamma & B_\gamma^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} A_\gamma & 0 \\ B & \widehat{S} \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix}. \quad (3.10)$$

For the symmetrized system, matrix  $B$  in the (2,1) block is replaced by  $B_\gamma$  and matrix  $C$  in the (2,2) block by  $C_\gamma$ . As in the case of stable finite elements,

we consider the eigenvalue problem

$$Sq = \mu M_p q, \quad (3.11)$$

where  $S = BA^{-1}B^T + C$ , and obtain bounds on  $\lambda$  in terms of  $\mu$ .

For  $W = W_1$  one immediately gets from (3.6) and (3.7)

$$\begin{aligned} S_\gamma^{-1} &= S^{-1} + \gamma M_p^{-1}, \\ \tilde{S}_\gamma^{-1} &= S^{-1} M_p (M_p + \gamma C)^{-1} + \gamma (M_p + \gamma C)^{-1}. \end{aligned}$$

Therefore, setting

$$\hat{S} =: -\gamma^{-1} M_p - C \quad \text{for (3.2)} \quad \text{or} \quad \hat{S} := -\gamma^{-1} M_p \quad \text{for (3.3)}, \quad (3.12)$$

we obtain with the same arguments as for the case of  $C = 0$  in Section 2.2 that all non-unit eigenvalues of (3.10) satisfy  $S^{-1}p + \gamma M_p^{-1}p = \lambda^{-1} \gamma M_p^{-1}p$ , where  $p \neq 0$ , and thus

$$\lambda = \frac{\gamma \mu}{1 + \gamma \mu}.$$

This representation is identical to the one in (2.7). Therefore, we obtain the following theorem.

**Theorem 3.3** ([15]). *Assume  $W = W_1$  and  $\hat{S}$  is defined as in (3.12). The preconditioned matrix  $\mathcal{P}^{-1} \hat{\mathcal{A}}$  has the eigenvalue 1 of algebraic multiplicity at least  $n$ . All other (nonunit) eigenvalues satisfy the following bounds:*

$$0 < \min_{\mu} \frac{\gamma a_{\mu}}{1 + \gamma a_{\mu}} \leq a_{\lambda} \leq 1, \quad |b_{\lambda}| \leq \max_{\mu} \min \left\{ \gamma |b_{\mu}|, \frac{1}{\gamma |b_{\mu}|} \right\} \leq 1,$$

where  $\lambda = a_{\lambda} + ib_{\lambda}$  and  $\mu = a_{\mu} + ib_{\mu}$ .

We noted already that the choice  $W = W_1$  is not always practical. The next theorem shows eigenvalue bounds for the case  $W = W_2$  in terms of the bounds given by the Bendixson's Theorem [71] for the generalized eigenvalue problem (3.11):

$$\alpha_{\mu} =: \min_{p \neq 0} \frac{p^T D p}{p^T M_p p} \leq a_{\mu}, \quad |b_{\mu}| \leq \beta_{\mu} =: \max_{p \neq 0} \frac{|p^T R p|}{p^T M_p p}, \quad (3.13)$$

where  $\mu = a_\mu + ib_\mu$ ,  $D = B \frac{A^{-1} + A^{-T}}{2} B^T + C$  is the symmetric part of  $S$ , and  $R = B \frac{A^{-1} - A^{-T}}{2} B^T$  is its skew-symmetric part.

**Theorem 3.4** ([15]). *Assume  $W = W_2$ ,  $0 < \gamma \leq (2\|M_p^{-1}C\|)^{-1}$  and  $\widehat{S} = -\gamma^{-1}M_p$ . The preconditioned matrix  $\mathcal{P}^{-1}\widehat{A}$  has the eigenvalue 1 of algebraic multiplicity at least  $n$ . All other (nonunit) eigenvalues satisfy the following bounds for the non-symmetric augmentation (3.2),*

$$0 < \frac{\gamma\alpha_\mu}{1 + \gamma\alpha_\mu} \leq a_\lambda \leq 1, \quad |b_\lambda| \leq \max_\mu \min \{\gamma\beta_\mu, 1\} \leq 1, \quad (3.14)$$

where  $\lambda = a_\lambda + ib_\lambda$ .

*Proof.* From (3.10), we immediately get that  $\lambda = 1$  is eigenvalue of algebraic multiplicity (at least)  $n$  and any vector  $[u; 0]$  with  $u \neq 0$  is a corresponding eigenvector. The remaining eigenvalues  $\lambda$  satisfy

$$\widetilde{S}_\gamma p = -\lambda \widehat{S} p.$$

For  $W = M_p$ ,  $\widehat{S} = -\gamma^{-1}M_p$ , using representation (3.7), we obtain

$$Sp = \frac{\lambda}{1 - \lambda} \left( \frac{1}{\gamma} M_p - C \right) p. \quad (3.15)$$

For brevity, we let  $\eta = \frac{\lambda}{1 - \lambda}$  and  $Q = \frac{1}{\gamma} M_p - C$ . It follows from Bendixson's Theorem that

$$\min_{p \neq 0} \frac{p^T D p}{p^T Q p} \leq a_\eta \leq \max_{p \neq 0} \frac{p^T D p}{p^T Q p}, \quad |b_\eta| \leq \max_{p \neq 0} \frac{|p^T R p|}{p^T Q p},$$

where  $\eta = a_\eta + ib_\eta$ . Using (3.15) we shall obtain bounds for  $\lambda$  in terms of  $\mu$  from (3.11). Since  $\gamma$  satisfies  $0 < \gamma \leq (2\|M_p^{-1}C\|)^{-1}$ , it holds

$$\frac{1}{2\gamma} M_p \leq \frac{1}{\gamma} M_p - C \leq \frac{1}{\gamma} M_p.$$

Therefore, we have

$$\gamma \min_{p \neq 0} \frac{p^T D p}{p^T M_p p} \leq \min_{p \neq 0} \frac{p^T D p}{p^T Q p} \leq a_\eta \leq \max_{p \neq 0} \frac{p^T D p}{p^T Q p} \leq 2\gamma \max_{p \neq 0} \frac{p^T D p}{p^T M_p p}, \quad (3.16)$$

$$|b_\eta| \leq \max_{p \neq 0} \frac{|p^T R p|}{p^T Q p} \leq 2\gamma \max_{p \neq 0} \frac{|p^T R p|}{p^T M_p p}. \quad (3.17)$$

Applying (3.13) yields

$$a_\eta \geq \gamma\alpha_\mu > 0, \quad |b_\eta| \leq 2\gamma\beta_\mu. \quad (3.18)$$

Solving  $\eta = \frac{\lambda}{1-\lambda}$  for  $a_\lambda$  and  $b_\lambda$ , we have

$$a_\lambda = \frac{a_\eta(1+a_\eta) + b_\eta^2}{(1+a_\eta)^2 + b_\eta^2}, \quad b_\lambda = \frac{b_\eta}{(1+a_\eta)^2 + b_\eta^2}.$$

From this and (3.18) the result in (3.14) follows.  $\square$

The bounds for  $\lambda$  in Theorems 3.3 and 3.4 are written in terms of bounds for the eigenvalues  $\mu$  from (3.11). Following the same argument as in [34], we can prove that  $\alpha_\mu$  and  $\beta_\mu$  from (3.14) and hence the smallest real and the largest imaginary parts of  $\mu$  are independent of  $h$ , but depend on  $\nu$ . The resulting eigenvalue bounds for  $\lambda$  are very similar to those for the LBB-stable case ( $C = 0$ ) from Section 2.2 and [13]. This suggests that the choice  $\gamma = O(1)$  leads to a method which is essentially insensitive to variations of parameters  $\nu$  and  $h$ . However, for the practical choice of  $W = \widehat{M}_p$  we have the restriction on  $\gamma$ . Numerical experiments [15] show that with the setting of  $W$  and  $\widehat{S}$  from Theorem 3.4 the restriction is indeed important and prohibits the choice  $\gamma = O(1)$  for all values  $\nu$  and  $h$  of interest. The situation looks better from the numerical viewpoint if one sets (for the non-symmetrized case (3.2))

$$W = \widehat{M}_p, \quad \widehat{S} =: -\gamma^{-1}\widehat{M}_p - C. \quad (3.19)$$

With this choice, solving linear systems with coefficient matrix  $\widehat{S}$  is inexpensive: either a direct sparse Cholesky solver or (better yet) one or two iterations of multigrid will suffice in practice. This combination of the augmentation and preconditioning, which is intermediate between those in Theorems 3.3 and 3.4, is not covered by the eigenvalue analysis above. However, numerical experiments [15] show that the convergence rate of GMRES preconditioned by (3.4) with the above-defined  $W$  and  $\widehat{S}$  and  $\gamma = 1$  is insensitive

to the variation of grid size and viscosity. Furthermore, with  $\gamma$  chosen by Fourier analysis (discussed later in this subsection), the above AL preconditioner can yield  $h$ - and  $\nu$ -independent convergence, as will be shown in numerical experiments.

Next we use a different approach to analyze the eigenvalues of the coefficient matrices preconditioned by the ideal AL-based preconditioners, which arrives at the same conclusion as in Theorem 3.3, and, moreover, provides a natural way to incorporate Fourier analysis to choose the augmentation parameter  $\gamma$ .

We start by analyzing the non-symmetric augmented system (3.2) and the corresponding preconditioner (3.4). Using the following factorization:

$$\mathcal{P}^{-1} = \begin{pmatrix} I_n & 0 \\ 0 & -\widehat{S}^{-1} \end{pmatrix} \begin{pmatrix} I_n & 0 \\ B & -I_m \end{pmatrix} \begin{pmatrix} A_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix},$$

the preconditioned augmented matrix is

$$\begin{aligned} \mathcal{P}^{-1}\widehat{\mathcal{A}} &= \begin{pmatrix} I_n & 0 \\ 0 & -\widehat{S}^{-1} \end{pmatrix} \begin{pmatrix} I_n & 0 \\ B & -I_m \end{pmatrix} \begin{pmatrix} A_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A_\gamma & B^T \\ B & -C \end{pmatrix} \\ &= \begin{pmatrix} I_n & A_\gamma^{-1}B_\gamma^T \\ 0 & -\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C) \end{pmatrix}. \end{aligned} \quad (3.20)$$

Clearly,  $\mathcal{P}^{-1}\widehat{\mathcal{A}}$  has 1 as an eigenvalue of algebraic multiplicity at least  $n$ . The remaining eigenvalues are the non-unit eigenvalues of  $-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C)$ . Applying Lemma 2.4 to  $\gamma BA_\gamma^{-1}B^TW^{-1}$ , we obtain

$$\begin{aligned} BA_\gamma^{-1}B_\gamma^T &= BA_\gamma^{-1}(B^T - \gamma B^TW^{-1}C) \\ &= \gamma^{-1}(\gamma BA_\gamma^{-1}B^TW^{-1})W(I_m - \gamma W^{-1}C) \\ &= (I_m - (I_m + \gamma BA^{-1}B^TW^{-1})^{-1})(\gamma^{-1}W - C). \end{aligned}$$

Plugging the above expression into  $-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C)$  and using the iden-

tity (2.16) in the Appendix yield

$$\begin{aligned}
& -\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C) \\
&= -\widehat{S}^{-1}(\gamma^{-1}W - (I_m + \gamma BA^{-1}B^T W^{-1})^{-1}(\gamma^{-1}W - C)) \\
&= -\widehat{S}^{-1}\left(\gamma^{-1}W - (I_m - \gamma CW^{-1} + \gamma(BA^{-1}B^T + C)W^{-1})^{-1}(\gamma^{-1}W - C)\right) \\
&= -\widehat{S}^{-1}\left(\gamma^{-1}W - (\gamma W^{-1} + \gamma(\gamma^{-1}W - C)^{-1}(BA^{-1}B^T + C)W^{-1})^{-1}\right) \\
&= -\widehat{S}^{-1}\left(\gamma^{-1}W - \gamma^{-1}W(I_m + (\gamma^{-1}W - C)^{-1}(BA^{-1}B^T + C))^{-1}\right) \\
&= -\widehat{S}^{-1}\left(\gamma^{-1}W - \gamma^{-1}W\left(I_m - (I_m + (BA^{-1}B^T + C)^{-1}(\gamma^{-1}W - C))^{-1}\right)\right) \\
&= -\widehat{S}^{-1}\gamma^{-1}W(I_m + (BA^{-1}B^T + C)^{-1}(\gamma^{-1}W - C))^{-1}.
\end{aligned} \tag{3.21}$$

**Remark 3.5.** *In the above derivation we have assumed that  $\gamma^{-1}C - W$  is nonsingular, as is always the case in practice.*

**Remark 3.6.** *If we let*

$$W = M_p + \gamma C \quad \text{and} \quad \widehat{S} = -\gamma^{-1}W,$$

*then*

$$\begin{aligned}
-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C) &= (I_m + (BA^{-1}B^T + C)^{-1}\gamma^{-1}M_p)^{-1} \\
&= I_m - (I_m + \gamma M_p^{-1}(BA^{-1}B^T + C))^{-1},
\end{aligned}$$

*in agreement with the analysis of the ideal AL preconditioner in Theorem 3.3.*

**Remark 3.7.** *If  $C = 0$  as in the case of stable finite elements and if we choose*

$$W = M_p \quad \text{and} \quad \widehat{S}^{-1} = -\nu W^{-1} - \gamma W^{-1},$$

*then*

$$\begin{aligned}
-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T) &= \frac{\nu + \gamma}{\gamma} \left( (I_m + (BA^{-1}B^T)^{-1}\gamma^{-1}M_p)^{-1} \right) \\
&= \frac{\nu + \gamma}{\gamma} (I_m - (I_m + \gamma M_p^{-1}BA^{-1}B^T)^{-1}),
\end{aligned}$$

in agreement with the analysis of the ideal AL preconditioner for stable finite elements in Theorem 2.1.

Next let us turn to the symmetric augmented system (3.3) and the preconditioner (3.5). Applying the preconditioner leads to

$$\begin{pmatrix} A_\gamma^{-1} & 0 \\ -\widehat{S}^{-1}B_\gamma A_\gamma^{-1} & \widehat{S}^{-1} \end{pmatrix} \begin{pmatrix} A_\gamma & B_\gamma^T \\ B_\gamma & -C_\gamma \end{pmatrix} = \begin{pmatrix} I_n & A_\gamma^{-1}B_\gamma \\ 0 & -\widehat{S}^{-1}B_\gamma A_\gamma^{-1}B_\gamma^T - \widehat{S}^{-1}C_\gamma \end{pmatrix}. \quad (3.22)$$

Again, we observe that 1 is the eigenvalue of the preconditioned matrix with algebraic multiplicity  $n$ . Similarly, the (2,2) block of the above matrix on the right-hand side can be expressed as

$$\begin{aligned} & -\widehat{S}^{-1}B_\gamma A_\gamma^{-1}B_\gamma^T - \widehat{S}^{-1}C_\gamma \\ &= -\widehat{S}^{-1}((I_m - \gamma CW^{-1})B(A + \gamma B^T W^{-1}B)^{-1}B^T(I_m - \gamma W^{-1}C) \\ & \quad + (C - \gamma CW^{-1}C)) \\ &= -\widehat{S}^{-1}((I_m - \gamma CW^{-1})((BA^{-1}B^T)^{-1} + \gamma W^{-1})^{-1}(I_m - \gamma W^{-1}C) \\ & \quad + (C - \gamma CW^{-1}C)) \\ &= -\widehat{S}^{-1}((I_m - \gamma CW^{-1})(\gamma^{-1}W - \gamma^{-1}W(BA^{-1}B^T + \gamma^{-1}W)^{-1}\gamma^{-1}W) \\ & \quad (I_m - \gamma W^{-1}C) + (C - \gamma CW^{-1}C)) \\ &= -\widehat{S}^{-1}((I_m - \gamma CW^{-1})\gamma^{-1}W(I_m - \gamma W^{-1}C) \\ & \quad - (I_m - \gamma CW^{-1})\gamma^{-1}W(BA^{-1}B^T + \gamma^{-1}W)^{-1}\gamma^{-1}W(I_m - \gamma W^{-1}C) \\ & \quad + (C - \gamma CW^{-1}C)) \\ &= -\widehat{S}^{-1}((\gamma^{-1}W - C) - (\gamma^{-1}W - C)(BA^{-1}B^T + C + \gamma^{-1}W - C)^{-1} \\ & \quad (\gamma^{-1}W - C)) \\ &= -\widehat{S}^{-1}((BA^{-1}B^T + C)^{-1} + (\gamma^{-1}W - C)^{-1}). \end{aligned}$$

Therefore, letting  $\widehat{S} = -\gamma^{-1}M_p$  and  $W = M_p + \gamma C$ , we obtain

$$\begin{aligned}
& -\widehat{S}^{-1}B_\gamma A_\gamma^{-1}B_\gamma^T - \widehat{S}^{-1}C_\gamma \\
&= -\widehat{S}^{-1}((BA^{-1}B^T + C)^{-1} + (\gamma^{-1}W - C)^{-1})^{-1} \\
&= (\gamma^{-1}M_p)^{-1}((BA^{-1}B^T + C)^{-1} + (\gamma^{-1}M_p)^{-1})^{-1} \\
&= ((BA^{-1}B^T + C)^{-1}\gamma^{-1}M_p + I_m)^{-1} \\
&= I_m - (I_m + \gamma M_p^{-1}(BA^{-1}B^T + C))^{-1},
\end{aligned}$$

from which we can draw the same conclusion as for the non-symmetric case, namely,

$$\lambda = \frac{\gamma\mu}{1 + \gamma\mu},$$

where  $\lambda$  is the nonunit eigenvalue of the matrix in (3.22) and  $\mu$  is as in (3.11).

When applying Fourier analysis to the Schur complement  $-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C)$  in (3.21) to choose  $\gamma$  (recall that  $C = \delta\widehat{C}$ ), the following inequality from [37, page 245]:

$$\frac{p^T\widehat{C}p}{p^TM_pp} \leq 1, \quad p \in \mathbb{R}^m,$$

along with

$$O(h^2) \leq \frac{p^T\widehat{C}p}{p^TM_pp}$$

verified through numerical experiments, shows that the eigenvalues of  $\widehat{C}$  are between  $O(h^4)$  and  $O(h^2)$  except the 0 eigenvalue. In our experiments, the same as  $M_p$ , the symbol of  $\widehat{C}$  is taken to be  $h^2$  as well; using other values, say 0,  $h^4$  or  $h^3$ , produces almost identical results. For all other matrices we use the same symbols as in Table 2.6. From (3.20) we can see that we need to choose  $\gamma$  so as to cluster the spectrum of  $-\widehat{S}^{-1}(BA_\gamma^{-1}B_\gamma^T + C)$  about 1. Thus, we pick the value of  $\gamma$  that minimizes the average distance of the eigenvalues from 1.

**Remark 3.8.** *If  $C = 0$  as in the case of stable finite element discretizations,*

Table 3.1: GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-P0, uniform and stretched grids).

Viscosity	0.1	0.01	0.005	0.001	0.1	0.01	0.005	0.001
Grid	Uniform				Stretched			
$16 \times 16$	5	5	5	5	5	5	5	5
$32 \times 32$	5	4	4	5	5	4	4	5
$64 \times 64$	5	4	4	5	5	4	4	5
$128 \times 128$	5	4	4	4	5	4	4	4

*Fourier analysis shows that the larger  $\gamma$  is, the more clustered the eigenvalues are about 1. This agrees with the analysis in Theorem 2.1.*

## 3.2 Numerical experiments

In this section, we study the numerical behavior of the ideal AL preconditioners for linear system discretized by stabilized Q1-P0 and Q1-Q1 finite elements.

The first set of experiments is to use the ideal AL preconditioner to solve the lid driven cavity problem discretized by Q1-P0 finite elements. With this choice of elements one can set  $W = \widehat{M}_p + \gamma C$ ,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$  for the AL preconditioner, cf. Remark 3.2. In this case no restriction on  $\gamma$  applies (see Theorem 3.3), so we set  $\gamma = 1$ . The results are shown in Table 3.1 for the lid driven cavity problem on both uniform and stretched grids. The  $h$ -independence and  $\nu$ -independence of the AL preconditioner are obvious from the data.

We also present numerical results for the three test problems obtained from Q1-Q1 finite element discretization in Tables 3.2, 3.3 and 3.4. In this case, we set  $W = \widehat{M}_p$  to maintain the sparsity in  $A_\gamma$  and  $\widehat{S} = -\gamma^{-1}\widehat{M} - C$  as in

Table 3.2: GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	7	7	7	7	7	7	8	8
$32 \times 32$	7	7	7	7	7	7	8	7
$64 \times 64$	6	6	7	6	7	6	7	7
$128 \times 128$	6	6	6	6	7	6	7	6

Table 3.3: GMRES(50) iterations with ideal AL preconditioner (cavity, Q1-Q1, stretched grids).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	7	7	7	7	7	7	7	7
$32 \times 32$	6	6	7	6	7	7	7	7
$64 \times 64$	6	6	6	6	6	6	7	7
$128 \times 128$	5	5	6	5	6	6	7	6

Table 3.4: GMRES(50) iterations with ideal AL preconditioner (step, Q1-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
	FA	Opt	FA	Opt	FA	Opt
$16 \times 48$	9	9	8	8	9	8
$32 \times 96$	9	9	8	8	9	9
$64 \times 192$	9	8	8	8	9	9
$128 \times 384$	8	8	9	8	9	9

(3.19). These results show that the FA-based strategy for choosing  $\gamma$  gives very accurate estimates for the optimal  $\gamma$ . The experiments also show that

the convergence rates with both sets of  $\gamma$  do not depend on the grid size and viscosity, demonstrating the robustness of the ideal AL preconditioner with respect to different problems with various parameters.

### 3.3 Modified AL preconditioner

Unfortunately, the ideal AL preconditioner (3.4) is not feasible in practice for large problems due to the high cost of exactly solving linear systems associated with  $A_\gamma$ . Instead, we consider its modified version for stabilized finite elements given by

$$\tilde{\mathcal{P}} = \begin{pmatrix} \tilde{A}_\gamma & 0 \\ B & \hat{S} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & 0 \\ 0 & A_{22} & 0 \\ B_1 & B_2 & \hat{S} \end{pmatrix}, \quad (3.23)$$

where  $\tilde{A}_\gamma = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ . Here we retain an upper triangular approximation  $\tilde{A}_\gamma$  for  $A_\gamma$  to simplify the analysis and in order to reuse parts of the code already written for the stable case; however, similar results are obtained by using a block lower triangular approximation to  $A_\gamma$ . Then applying  $\tilde{\mathcal{P}}^{-1}$  to the augmented matrix  $\hat{\mathcal{A}}$  gives rise to

$$\begin{aligned} \tilde{\mathcal{P}}^{-1} \hat{\mathcal{A}} &= \begin{pmatrix} I_n & 0 \\ 0 & -\hat{S}^{-1} \end{pmatrix} \begin{pmatrix} I_n & 0 \\ B & -I_m \end{pmatrix} \begin{pmatrix} \tilde{A}_\gamma^{-1} & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A_\gamma & B^T \\ B & -C \end{pmatrix} \\ &= \begin{pmatrix} \tilde{A}_\gamma^{-1} A_\gamma & \tilde{A}_\gamma^{-1} B_\gamma^T \\ -\hat{S}^{-1} B (\tilde{A}_\gamma^{-1} A_\gamma - I) & -\hat{S}^{-1} (B \tilde{A}_\gamma^{-1} B_\gamma^T + C) \end{pmatrix}. \end{aligned} \quad (3.24)$$

Carrying out calculations similar to those found in Section 2.4.1, we obtain the following result.

**Theorem 3.9.** *Letting  $\widehat{S}^{-1} = -\gamma W^{-1}$ , we have*

$$\widetilde{\mathcal{P}}^{-1}\widehat{\mathcal{A}} = \begin{pmatrix} I_{n/2} - DE & 0 & DG(I_m - \gamma W^{-1}C) \\ A_{22}^{-1}A_{21} & I_{n/2} & A_{22}^{-1}B_2^T(I_m - \gamma W^{-1}C) \\ FE & 0 & I_m - FG(I_m - \gamma W^{-1}C) \end{pmatrix}. \quad (3.25)$$

Hence,  $\widetilde{\mathcal{P}}^{-1}\widehat{\mathcal{A}}$  has 1 as an eigenvalue of algebraic multiplicity at least  $n$ . The remaining eigenvalues are the non-unit eigenvalues of the matrix

$$\begin{pmatrix} I_{n/2} - DE & DG(I_m - \gamma W^{-1}C) \\ FE & I_m - FG(I_m - \gamma W^{-1}C) \end{pmatrix}.$$

*Proof.* Letting  $\widehat{S}^{-1} = -\gamma W^{-1}$ , we have (with  $I = I_m$ )

$$\begin{aligned} & I - \widehat{S}^{-1} \left( (-\gamma W^{-1})^{-1}FG + \gamma^{-1}W + (C + \widehat{S})(I - \gamma W^{-1}C)^{-1} \right) (I - \gamma W^{-1}C) \\ &= I + \gamma W^{-1} \left( -\gamma^{-1}WFG + \gamma^{-1}W + (C - \gamma^{-1}W)(I - \gamma W^{-1}C)^{-1} \right) (I - \gamma W^{-1}C) \\ &= I - FG(I - \gamma W^{-1}C), \end{aligned}$$

from which (3.25) immediately follows. The same derivation as in Theorem 2.6 gives that the algebraic multiplicity of the eigenvalue 1 is at least  $n$ .  $\square$

Again, Fourier analysis is used to guide in the choice of  $\gamma$ . The procedure followed is similar to that described in Section 2.4. In the interest of brevity, we omit the details.

Motivated by the same reason for the choice of  $\widehat{S}$  in the ideal AL preconditioner (3.4), we also consider the choice

$$\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$$

in the modified version. The presence of  $C$  complicates the analysis, so we will only investigate it through numerical experiments in the next section.

Table 3.5: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	Em	Opt	Em	Opt	Em	Opt	Em	Opt
$16 \times 16$	10	10	15	15	18	18	29	29
$32 \times 32$	10	9	13	12	16	15	29	28
$64 \times 64$	10	8	11	10	13	13	26	26
$128 \times 128$	9	8	9	9	12	12	24	23

## 3.4 Numerical experiments

### 3.4.1 Empirical rule for choosing $\gamma$

As before, we compare the number of GMRES iterations preconditioned by the modified AL preconditioner with  $\gamma$  chosen by the empirical rule and with experimentally determined optimal values. We test the preconditioner with  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ . The results are shown in Tables 3.5, 3.6 and 3.7. The iteration counts show the excellent behavior of the modified AL preconditioner, with independence on the mesh size and a mild dependence on  $\nu$ . In addition, the empirical rule gives almost optimal convergence rate in terms of GMRES iteration counts.

### 3.4.2 Fourier analysis-based approach for choosing $\gamma$

In the following, we evaluate  $\gamma$  chosen by the Fourier analysis and experimentally determined optimal values. First we test the preconditioner with approximate Schur complement given by  $\widehat{S}^{-1} = -\gamma W^{-1}$  (with  $W = \widehat{M}_p$ ), as suggested by Theorem 3.9. We consider both uniform and stretched grids. The results are shown in Tables 3.8, 3.9 and 3.10.

As for stable finite elements, the value of  $\gamma$  determined by FA gives nearly

Table 3.6: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	Em	Opt	Em	Opt	Em	Opt	Em	Opt
$16 \times 16$	9	9	14	14	16	16	25	25
$32 \times 32$	9	9	12	12	15	14	25	24
$64 \times 64$	9	8	11	11	14	14	25	25
$128 \times 128$	9	7	10	10	13	13	26	25

Table 3.7: GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005	
Grid	Em	Opt	Em	Opt	Em	Opt
$16 \times 48$	13	13	18	17	20	20
$32 \times 96$	13	12	18	17	21	20
$64 \times 192$	12	12	17	16	22	20
$128 \times 384$	12	11	18	14	25	20

Table 3.8: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	10	9	16	14	38	17	92	27
$32 \times 32$	9	8	13	12	24	14	50	27
$64 \times 64$	8	8	11	10	13	13	36	26
$128 \times 128$	8	7	9	9	13	12	25	24

optimal results in almost all cases; the only exceptions occur for low viscosity

Table 3.9: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	9	9	14	13	32	15	82	24
$32 \times 32$	9	8	12	11	22	14	43	24
$64 \times 64$	9	8	12	11	14	14	35	25
$128 \times 128$	8	7	10	10	13	13	26	25

Table 3.10: GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p$ ).

Viscosity	0.1		0.01		0.005	
Grid	FA	Opt	FA	Opt	FA	Opt
$16 \times 48$	16	12	67	16	97	18
$32 \times 96$	13	11	40	17	57	20
$64 \times 192$	12	11	20	16	35	20
$128 \times 384$	11	11	17	17	27	23

problems with coarse grids, and we have already observed that these problems are not physically meaningful. What is important is that the gap between the iteration counts corresponding to the two choices of  $\gamma$  narrows as the mesh is refined.

Again, we observe iteration counts that are essentially  $h$ -independent and only mildly dependent on the viscosity. One exception appears to be the step problem with viscosity  $\nu = 0.005$ , for which the iteration count with the optimal  $\gamma$  shows a mild dependence on  $h$ . We address this problem by modifying the approximate Schur complement in the preconditioner; letting  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$  in the modified AL preconditioner, we obtain the results

Table 3.11: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	10	10	16	15	18	18	64	29
$32 \times 32$	9	9	13	12	18	15	48	28
$64 \times 64$	8	8	10	10	13	13	35	26
$128 \times 128$	8	8	9	9	13	12	24	23

Table 3.12: GMRES(50) iterations with modified AL preconditioner (cavity, Q1-Q1, stretched grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005		0.001	
	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	9	9	14	14	16	16	55	25
$32 \times 32$	9	9	12	12	17	14	42	24
$64 \times 64$	9	8	11	11	14	14	34	25
$128 \times 128$	7	7	10	10	14	13	26	25

Table 3.13: GMRES(50) iterations with modified AL preconditioner (step, Q1-Q1, uniform grids,  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ ).

Viscosity	0.1		0.01		0.005	
	FA	Opt	FA	Opt	FA	Opt
$16 \times 48$	13	13	24	17	49	20
$32 \times 96$	13	12	24	17	42	20
$64 \times 192$	12	12	19	16	28	20
$128 \times 384$	11	11	15	14	20	20

presented in Tables 3.11, 3.12 and 3.13. Now we see that GMRES iteration

Table 3.14: GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q1-Q1, uniform grids). The asterisk means that the FA values of  $\gamma$  are the ones found for the Oseen problem.

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA*	Opt	FA*	Opt	FA*	Opt	FA*	Opt
$16 \times 16$	14	13	25	23	31	31	121	63
$32 \times 32$	14	13	23	23	35	35	107	77
$64 \times 64$	14	13	23	23	34	34	100	81
$128 \times 128$	15	13	25	23	41	35	97	96

Table 3.15: GMRES(50) iterations with modified AL preconditioner (cavity, Newton, Q1-Q1, stretched grids). The asterisk means that the FA values of  $\gamma$  are the ones found for the Oseen problem.

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA*	Opt	FA*	Opt	FA*	Opt	FA*	Opt
$16 \times 16$	13	13	24	21	28	25	107	47
$32 \times 32$	13	13	23	22	32	30	96	62
$64 \times 64$	14	14	24	24	33	33	87	81
$128 \times 128$	14	14	25	24	40	34	89	77

counts with the optimal  $\gamma$  are essentially  $h$ -independent for all values of  $\nu$ ; indeed, the rate of convergence tends to improve as the mesh is refined. Again, the iteration counts for  $\gamma$  chosen by FA are very close, and even identical to those with optimal  $\gamma$  on fine grids. This choice of  $\widehat{S}$  is only slightly more expensive than the previous one; in practice, including  $C$  in  $\widehat{S}$  may be recommended, as it results in increased robustness of the solver. We also emphasize that the preconditioner does not have any difficulties handling stretched grids.

We also performed some experiments on linear systems generated from

Table 3.16: Full GMRES iterations with PCD, LSC preconditioners (cavity, Q1-Q1, uniform and stretched grids).

Viscosity	0.005		0.001		0.005		0.001	
	PCD	LSC	PCD	LSC	PCD	LSC	PCD	LSC
Grid	Uniform				Stretched			
$16 \times 16$	54	43	144	122	50	38	118	98
$32 \times 32$	54	42	155	147	51	48	143	126
$64 \times 64$	41	31	154	149	50	63	145	154
$128 \times 128$	35	28	110	104	47	84	142	188

Newton linearization using the FA values of  $\gamma$  found for the Oseen equations and with the choice  $\widehat{S} = -\gamma^{-1}\widehat{M}_p - C$ . The iteration numbers are shown in Table 3.14 and 3.15. Similar remarks to those in Section 2.4.2 apply.

### 3.4.3 Comparison with other preconditioners

Next, we briefly compare the modified augmented Lagrangian preconditioner with some of the best existing methods for the class of problems of interest. In [33], the LSC preconditioner has been generalized to deal with stabilized finite element discretizations of the Navier–Stokes equations, like Q1-Q1. Comparing our results with those reported in [33] as well as Table 3.16, we observe that the modified AL preconditioner is much less sensitive to variations in  $h$  and  $\nu$ , and performs much better and more consistently on stretched grids. We also tested the LSC and PCD preconditioners on linear systems from Newton linearization; results for two values of the viscosity using uniform and stretched grids are shown in Table 3.17. Comparing these results with those in Tables 3.14 and 3.15, we see that the modified AL preconditioner clearly outperforms the other preconditioners on these problems.

Table 3.17: Full GMRES iterations with PCD, LSC preconditioners (cavity, Newton, Q1-Q1, uniform and stretched grids).

Viscosity	0.005		0.001		0.005		0.001	
	PCD	LSC	PCD	LSC	PCD	LSC	PCD	LSC
Grid	Uniform				Stretched			
$16 \times 16$	69	62	122	94	63	53	125	112
$32 \times 32$	83	73	311	290	77	75	244	231
$64 \times 64$	77	65	240	204	82	111	248	246
$128 \times 128$	72	55	256	199	80	168	251	298

### 3.4.4 Inexact solves

Finally, we perform experiments analogous to those reported in Table 2.22 to compare the use of exact and inexact solves for the linear systems associated with  $A_{11}$  and  $A_{22}$ . Again, exact solves are obtained with a sparse LU factorization with column AMD reordering. Inexact solves are now obtained by performing three AMG iterations with symmetric Gauss–Seidel as the smoother. The reason while three inner AMG iterations are performed is that for the larger meshes, we observed a slight increase in the number of iterations with inexact solves if only one AMG iteration is used. The results for both Picard and Newton linearization are given in Table 3.18. By comparing iteration counts, we observe almost identical rates of convergence for the exact and inexact approaches; in one case ( $128 \times 128$  grid, Newton linearization) the exact variant actually requires one more iteration than the inexact one. We see that while the exact variant is faster for small and moderate-size problems, the inexact variant becomes considerably faster for large problems. Moreover, the timings show that the inexact solver yields a preconditioner with very good scalability, especially for the Picard linearization.

Table 3.18: Comparison of exact and inexact inner solvers. GMRES(50) iterations and timings with modified AL preconditioner (cavity, Q1-Q1, uniform grids,  $\nu = 0.005$ ).

Grid	Picard		Newton	
Timings	Exact	Inexact	Exact	Inexact
$16 \times 16$	18	18	31	31
Setup time	0.01	0.06	0.01	0.01
Iter time	0.04	0.13	0.11	0.21
Total time	0.05	0.19	0.12	0.22
$32 \times 32$	18	18	35	35
Setup time	0.10	0.05	0.10	0.05
Iter time	0.16	0.57	0.34	0.89
Total time	0.26	0.62	0.44	0.94
$64 \times 64$	13	13	34	37
Setup time	0.96	0.31	0.89	0.29
Iter time	0.52	2.17	1.33	4.98
Total time	1.49	2.48	2.22	5.27
$128 \times 128$	13	14	41	40
Setup time	19.52	1.34	19.14	1.33
Iter time	3.24	13.13	10.88	30.51
Total time	22.76	14.47	30.02	31.84
$256 \times 256$	12	13	42	42
Setup time	158.62	5.43	158.97	5.32
Iter time	15.93	38.95	63.55	149.36
Total time	174.55	44.38	222.52	154.68

## Chapter 4

# A relaxed dimensional factorization preconditioner

In this chapter we study a relaxed dimensional factorization (RDF) preconditioner for saddle point problems. Properties of the preconditioned matrix are analyzed and compared with those of the closely related dimensional splitting (DS) preconditioner introduced in [10]. Numerical results for a variety of discretization methods of both steady and unsteady incompressible flow problems indicate very good behavior of the RDF preconditioner with respect to both mesh size and viscosity. Finally we discuss a generalization of the RDF preconditioner for linear system obtained by using stabilized finite elements. Additional results can be found in [12].

### 4.1 Dimensional splitting preconditioner

We begin with a brief description of DS preconditioning; for further details, see [10].

We first consider the saddle point system (1.17) for the 2D Oseen problem with stable finite elements discretization; in this case,  $C = 0$  and  $\tilde{\mathcal{A}}$  has the

block structure

$$\tilde{\mathcal{A}} = \begin{pmatrix} A_1 & 0 & B_1^T \\ 0 & A_2 & B_2^T \\ -B_1 & -B_2 & 0 \end{pmatrix}, \quad (4.1)$$

i.e.,  $A = \text{diag}(A_1, A_2)$  and  $B = (B_1, B_2)$  in (1.17). Each diagonal block matrix  $A_i \in \mathbb{R}^{n/2 \times n/2}$  is a discrete scalar convection-diffusion-reaction operator

$$A_i = \sigma M + \nu L + N_i, \quad (4.2)$$

with  $M$  being the diagonal block of the velocity mass matrix  $M_u$  ( $M_u = \text{diag}(M, M)$  for 2D),  $L$  the discrete (negative) Laplacian, and  $N_i$  the convective terms; the parameter  $\sigma \geq 0$  is typically proportional to the reciprocal of the time step, and is zero in the steady case. Moreover,  $B_1^T \in \mathbb{R}^{n/2 \times m}$ ,  $B_2^T \in \mathbb{R}^{n/2 \times m}$  are discretizations of the partial derivatives  $\frac{\partial}{\partial x}$ ,  $\frac{\partial}{\partial y}$ , respectively.

The *dimensional splitting* (DS) preconditioner proposed in [10] is of the form

$$\mathcal{B} = \frac{1}{\alpha} \begin{pmatrix} A_1 + \alpha I & 0 & B_1^T \\ 0 & \alpha I & 0 \\ -B_1 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} \alpha I & 0 & 0 \\ 0 & A_2 + \alpha I & B_2^T \\ 0 & -B_2 & \alpha I \end{pmatrix}, \quad (4.3)$$

and is suggested by splitting  $\tilde{\mathcal{A}}$  as follows:

$$\tilde{\mathcal{A}} = \tilde{\mathcal{A}}_1 + \tilde{\mathcal{A}}_2 = \begin{pmatrix} A_1 & 0 & B_1^T \\ 0 & 0 & 0 \\ -B_1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & A_2 & B_2^T \\ 0 & -B_2 & 0 \end{pmatrix}. \quad (4.4)$$

(Here and thereafter we omit the subscripts for the identity matrices for notational brevity when they are clear from the context.) The alternating (alternating direction iteration-like) stationary iteration corresponding to the DS splitting is

$$\begin{cases} (\alpha I + \tilde{\mathcal{A}}_1)x^{(k+\frac{1}{2})} = (\alpha I - \tilde{\mathcal{A}}_2)x^{(k)} + \tilde{b}, \\ (\alpha I + \tilde{\mathcal{A}}_2)x^{(k+1)} = (\alpha I - \tilde{\mathcal{A}}_1)x^{(k+\frac{1}{2})} + \tilde{b}, \end{cases} \quad (4.5)$$

(with  $k = 0, 1, \dots$ , and  $x^{(0)}$  arbitrary). It is obtained alternating between the following two splittings of  $\tilde{\mathcal{A}}$ :

$$\tilde{\mathcal{A}} = (\alpha I + \tilde{\mathcal{A}}_1) - (\alpha I - \tilde{\mathcal{A}}_2)$$

and

$$\tilde{\mathcal{A}} = (\alpha I + \tilde{\mathcal{A}}_2) - (\alpha I - \tilde{\mathcal{A}}_1).$$

In [10], it is shown that the iteration (4.5) is convergent for all  $\alpha > 0$  to the unique solution of  $\tilde{\mathcal{A}}x = \tilde{b}$ , provided that  $A + A^T$  is positive definite and  $B$  has full rank. We should point out that the factor  $1/\alpha$  used in (4.3) originally appears as  $1/(2\alpha)$  in [10]. Since this factor has no effect on the preconditioned system, we use  $1/\alpha$  in this chapter just for analysis purpose.

By performing the matrix multiplication on the right-hand side of (4.3), it follows that  $\mathcal{B}$  has the block structure

$$\mathcal{B} = \begin{pmatrix} \alpha I + A_1 & -\alpha^{-1}B_1^T B_2 & B_1^T \\ 0 & \alpha I + A_2 & B_2^T \\ -B_1 & -B_2 & \alpha I \end{pmatrix}. \quad (4.6)$$

From (4.1) and (4.6), we can see that the difference between the preconditioner  $\mathcal{B}$  and the coefficient matrix  $\tilde{\mathcal{A}}$  is given by

$$\mathcal{B} - \tilde{\mathcal{A}} = \begin{pmatrix} \alpha I & -\alpha^{-1}B_1^T B_2 & 0 \\ 0 & \alpha I & 0 \\ 0 & 0 & \alpha I \end{pmatrix}. \quad (4.7)$$

Equation (4.7) shows that as  $\alpha$  tends to zero, the weight of the three diagonal blocks in the difference matrix decreases, whereas the weight of the nonzero off-diagonal block becomes unbounded. Hence, the choice of  $\alpha$  requires a balancing act. The size of  $\alpha$  depends on the scaling of the equations in the linear system.

## 4.2 Relaxed dimensional factorization preconditioner

Based on the previous observations, an improved variant of the DS preconditioner was proposed in [12]. The new preconditioner is defined as follows:

$$\mathcal{M} = \begin{pmatrix} A_1 & -\alpha^{-1}B_1^T B_2 & B_1^T \\ 0 & A_2 & B_2^T \\ -B_1 & -B_2 & \alpha I \end{pmatrix}. \quad (4.8)$$

By comparing the preconditioner  $\mathcal{M}$  defined in (4.8) with the DS preconditioner  $\mathcal{B}$  defined in (4.6), we can see that the new preconditioner no longer contains the shift terms  $\alpha I$  appearing in the (1,1) and (2,2) blocks of  $\mathcal{B}$ . It is important to note that the preconditioner  $\mathcal{M}$  can be written in factorized form as

$$\mathcal{M} = \frac{1}{\alpha} \begin{pmatrix} A_1 & 0 & B_1^T \\ 0 & \alpha I & 0 \\ -B_1 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} \alpha I & 0 & 0 \\ 0 & A_2 & B_2^T \\ 0 & -B_2 & \alpha I \end{pmatrix} =: \frac{1}{\alpha} \mathcal{M}_1 \mathcal{M}_2.$$

Note that  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are invertible provided that  $A_1, A_2$  have positive definite symmetric part, hence in this case the preconditioner itself is non-singular. This (sufficient) condition is satisfied for both Stokes and Oseen problems. In the particular case  $\alpha = 1$ , the preconditioner  $\mathcal{M}$  reduces to

$$\mathcal{M} = \begin{pmatrix} A_1 & 0 & B_1^T \\ 0 & I & 0 \\ -B_1 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & A_2 & B_2^T \\ 0 & -B_2 & I \end{pmatrix}. \quad (4.9)$$

By analogy with the concept of *dimensional splitting* (4.3), it follows that the preconditioner  $\mathcal{M}$  given by (4.9) can be regarded as a *dimensional factorization* preconditioner, and hence the preconditioner  $\mathcal{M}$  given by (4.8) is referred to as the *Relaxed Dimensional Factorization* (RDF) preconditioner.

By comparing (4.8) with (4.1), we can see that the difference between  $\mathcal{M}$  and  $\tilde{\mathcal{A}}$  is given by

$$\mathcal{R} = \mathcal{M} - \tilde{\mathcal{A}} = \begin{pmatrix} 0 & -\alpha^{-1}B_1^T B_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha I \end{pmatrix}. \quad (4.10)$$

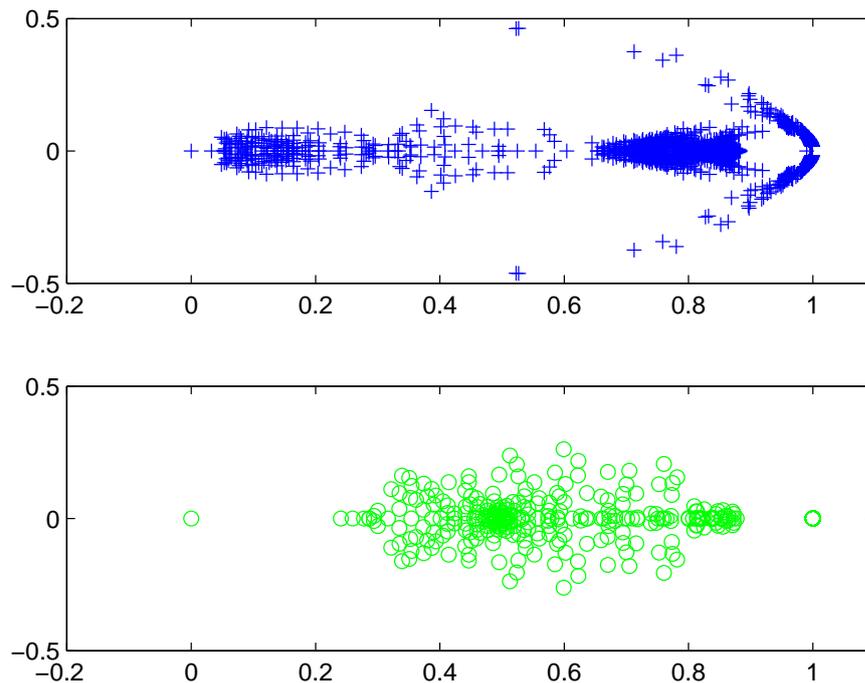
Hence, compared to DS preconditioning, now only one of the three diagonal blocks (the smallest one in size) is nonzero, while the nonzero off-diagonal block is the same for both RDF and DS preconditioning. This observation suggests that  $\mathcal{M}$  may be a better preconditioner than  $\mathcal{B}$ , since it gives a better approximation of the system matrix  $\tilde{\mathcal{A}}$  for the same value of  $\alpha$ . Furthermore, the structure of (4.10) somewhat facilitates the analysis of the eigenvalue distribution of the preconditioned matrix. We should remark that the RDF preconditioner  $\mathcal{M}$  no longer relates to an alternating direction iteration like (4.5). Clearly, this fact is of no consequence when  $\mathcal{M}$  is used as a preconditioner for a Krylov subspace method like GMRES [67].

We have the following result from [12].

**Theorem 4.1** ([12]). *The preconditioned matrix  $\tilde{\mathcal{A}}\mathcal{M}^{-1}$  has an eigenvalue at 1 with algebraic multiplicity  $n$ . The remaining eigenvalues are the eigenvalues of the matrix  $Z_\alpha = \alpha^{-1}(S_1 + S_2) - 2\alpha^{-2}S_1S_2$ , where  $\hat{A}_1 = A_1 + \alpha^{-1}B_1^T B_1$ ,  $S_1 = B_1\hat{A}_1^{-1}B_1^T$ ,  $\hat{A}_2 = A_2 + \alpha^{-1}B_2^T B_2$  and  $S_2 = B_2\hat{A}_2^{-1}B_2^T$ .*

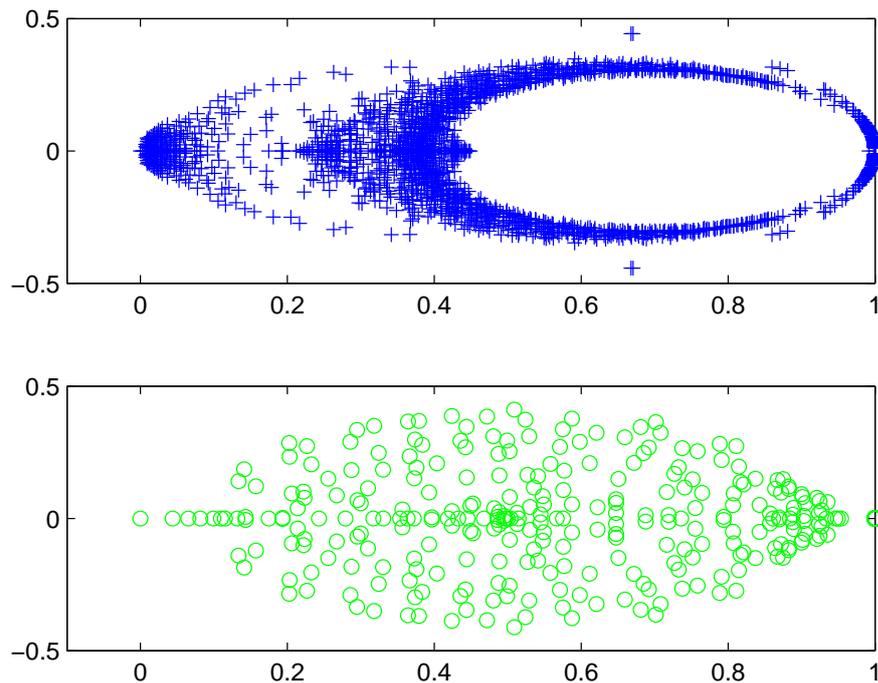
Eigenvalue plots of the preconditioned matrices obtained with DS and RDF preconditioners (with optimal values of  $\alpha$ ) are displayed in Figures 4.1–4.2 [12]. These two plots confirm that for both DS and RDF preconditioning, the eigenvalues of the preconditioned matrices are confined to a rectangular region in the half-plane  $\text{Re}(z) \geq 0$ ; note that the appearance of a zero eigenvalue is due to the singularity of the saddle point system (1.17), which is caused by the hydrostatic pressure mode [37]. As already mentioned, this zero eigenvalue is harmless in practice and can be ignored [37, Section 2.3]. In

Figure 4.1: Plots of the eigenvalues of the preconditioned Oseen matrix (cavity, Q2-Q1, uniform  $32 \times 32$  grid,  $\nu = 0.01$  and experimentally optimal  $\alpha$ ). Top:  $\tilde{\mathcal{A}}\mathcal{B}^{-1}$ . Bottom:  $\tilde{\mathcal{A}}\mathcal{M}^{-1}$ .



these two examples, corresponding to the viscosities  $\nu = 0.01$  and  $\nu = 0.001$ , it is clear that RDF produces a much more favorable eigenvalue distribution than DS. Indeed, in these examples the DS preconditioner fails to force many of the eigenvalues away from zero (especially in the case of  $\nu = 0.001$ ), which may cause the GMRES method preconditioned by DS preconditioner to converge more slowly. In contrast, the RDF preconditioner is able to cluster most of the eigenvalues at 1. Indeed, according to Theorem 4.1, there are at least 2178 eigenvalues equal to 1 in this case, and the plots show that

Figure 4.2: Plots of the eigenvalues of the preconditioned Oseen matrix (cavity, Q2-Q1, uniform  $32 \times 32$  grid,  $\nu = 0.001$  and experimentally optimal  $\alpha$ ). Top:  $\tilde{\mathcal{A}}\mathcal{B}^{-1}$ . Bottom:  $\tilde{\mathcal{A}}\mathcal{M}^{-1}$ .



the remaining nonzero eigenvalues are well separated from the origin. The clustering of the spectrum obtained with DS preconditioning can be greatly improved by diagonally scaling  $\tilde{\mathcal{A}}$  prior to applying the DS preconditioner (see [10]), but it is interesting to see that for these examples RDF achieves excellent clustering without the need for scaling.

The following result provides additional information about the non-unit eigenvalues of the RDF-preconditioned matrices.

**Theorem 4.2** ([12]). *The eigenvalues  $\mu_i$  of  $Z_\alpha$  are of the form*

$$\mu_i = \frac{\alpha \xi_i}{1 + \alpha \xi_i},$$

where the  $\xi_i$ 's satisfy the generalized eigenvalue problem

$$BA^{-1}B^T \phi_i = \xi_i(\alpha^2 I + \widehat{S}_1 \widehat{S}_2) \phi_i, \quad \text{with} \quad \widehat{S}_k = B_k A_k^{-1} B_k^T \quad (k = 1, 2).$$

The foregoing theorem can be used to obtain estimates on the magnitude of the non-unit eigenvalues of the preconditioned matrix; for example, it can be used to show that they go to zero like  $O(\alpha)$  for  $\alpha \rightarrow 0^+$ , and like  $O(\alpha^{-1})$  for  $\alpha \rightarrow \infty$ . Indeed, let  $(\xi, \phi)$  be a generalized eigensolution of  $BA^{-1}B^T \phi = \xi(\alpha^2 I + \widehat{S}_1 \widehat{S}_2) \phi$ , with  $\|\phi\|_2 = 1$ . Then

$$\xi = \frac{\phi^* BA^{-1} B^T \phi}{\alpha^2 + \phi^* \widehat{S}_1 \widehat{S}_2 \phi},$$

where  $\phi^*$  is the conjugate transpose of  $\phi$ . It follows that

$$\mu = \frac{\frac{\phi^* BA^{-1} B^T \phi}{\alpha^2 + \phi^* \widehat{S}_1 \widehat{S}_2 \phi}}{\frac{1}{\alpha} + \frac{\phi^* BA^{-1} B^T \phi}{\alpha^2 + \phi^* \widehat{S}_1 \widehat{S}_2 \phi}}.$$

Taking the limits for  $\alpha \rightarrow 0^+$  and for  $\alpha \rightarrow \infty$  we see that the non-unit eigenvalues of the preconditioned matrix  $\widetilde{\mathcal{A}}\mathcal{M}^{-1}$  tend to 0 like  $O(\alpha)$  and to  $\infty$  like  $O(\alpha^{-1})$ , respectively.

### 4.2.1 Practical implementation of the RDF preconditioner

In this subsection we outline the practical implementation of the RDF preconditioner in a Krylov subspace iterative method, such as GMRES. The main step is applying the preconditioner, i.e., solving linear systems with

coefficient matrix  $\mathcal{M}$ . The RDF preconditioner can be factorized as follows:

$$\begin{aligned} \mathcal{M} &= \begin{pmatrix} A_1 & 0 & \alpha^{-1}B_1^T \\ 0 & I & 0 \\ -B_1 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & A_2 & B_2^T \\ 0 & -B_2 & \alpha I \end{pmatrix} \\ &= \begin{pmatrix} I & 0 & \alpha^{-1}B_1^T \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} \widehat{A}_1 & 0 & 0 \\ 0 & I & 0 \\ -B_1 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & \widehat{A}_2 & B_2^T \\ 0 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & -\alpha^{-1}B_2 & I \end{pmatrix}, \end{aligned}$$

showing that the preconditioner requires solving two linear systems at each step, with coefficient matrices  $\widehat{A}_1 = A_1 + \alpha^{-1}B_1^T B_1$  and  $\widehat{A}_2 = A_2 + \alpha^{-1}B_2^T B_2$ . Note that these systems can be interpreted as discretizations of anisotropic scalar elliptic boundary value problems of convection-diffusion type (for the unsteady case, the equations are of reaction-convection-diffusion type). Several different approaches are available for solving linear systems involving  $\widehat{A}_1$  and  $\widehat{A}_2$ . We defer the discussion of these to Section 4.3.

In [10] it was pointed out that the performance of DS preconditioning can be significantly improved by diagonal scaling. We found that scaling can be beneficial for RDF as well. Unless otherwise specified, we perform a preliminary symmetric scaling of the system  $\widetilde{\mathcal{A}}x = \widetilde{b}$  in the form  $\mathcal{D}^{-\frac{1}{2}}\widetilde{\mathcal{A}}\mathcal{D}^{-\frac{1}{2}}y = \mathcal{D}^{-\frac{1}{2}}\widetilde{b}$  with  $y = \mathcal{D}^{\frac{1}{2}}x$ , and

$$\mathcal{D} = \begin{pmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & I \end{pmatrix},$$

where (for 2D problems)  $\text{diag}(D_1, D_2)$  is the main diagonal of the velocity mass matrix  $M_u$ .

### 4.2.2 Estimation of the optimal relaxation parameter $\alpha$ using Fourier analysis

In this section we describe an inexpensive technique for approximating the optimal value of the parameter  $\alpha$  using Fourier analysis as we did for the modified AL preconditioner in Section 2.4. In [12], the 2D differential operators are replaced by 1D surrogates; in this thesis, we use the original 2D operators.

Following the symbols in Table 2.6, we use  $a_1, a_2, b_1, b_2$  to denote the generic eigenvalues of  $A_1, A_2, B_1, B_2$ . Furthermore, from  $S_1 = B_1(A_1 + \alpha^{-1}B_1^T B_1)^{-1}B_1^T$ ,  $S_2 = B_2(A_2 + \alpha^{-1}B_2^T B_2)^{-1}B_2^T$  and  $Z_\alpha = \alpha^{-1}(S_1 + S_2) - 2\alpha^{-2}S_1S_2$ , it follows that the eigenvalues of  $S_1, S_2$  and  $Z_\alpha$  can also be represented in terms of  $a_1, a_2, b_1$  and  $b_2$ . To be more specific, the eigenvalues of  $Z_\alpha$  can be expressed as

$$z(\alpha) = \frac{1}{\alpha}(s_1 + s_2) - \frac{2}{\alpha^2}s_1s_2,$$

where  $s_1, s_2$  are the eigenvalues of  $S_1, S_2$  and are given by

$$s_1 = \frac{b_1^2}{a_1 + \frac{1}{\alpha}b_1^2}, \quad \text{and} \quad s_2 = \frac{b_2^2}{a_2 + \frac{1}{\alpha}b_2^2}.$$

Note that the diagonal scaling strategy described in the previous subsection leaves  $s_1$  and  $s_2$ , and thus  $z(\alpha)$ , unchanged.

From Theorem 4.1 we expect that clustering the eigenvalues of  $Z_\alpha$  around 1 could lead to fast convergence of the RDF-preconditioned iteration. This can be achieved by choosing  $\alpha$  so as to cluster the values of  $z(\alpha)$  around 1. Therefore, for any fixed  $h$  and  $\nu$ , choosing the value of  $\alpha$  that maximizes the clustering of the  $z(\alpha)$  around 1 is equivalent to solving the following optimization problem:

$$\min_{\alpha > 0} \quad \text{mean} |z(\alpha; \theta_x, \theta_y) - 1|$$

$$\text{where} \quad \theta_x, \theta_y = 1, 2, \dots, l$$

to determine the value of  $\alpha$ , where mean represents the average value. Unfortunately, finding a closed form solution to this problem appears to be rather difficult. In practice, we compute an approximate solution to the optimization problem as follows. We sample  $\alpha$  over the range of  $[0.0001, 2]$  (with step 0.0001) for 2D finite element discretizations, and over the range  $[0.1, 30]$  (with step 0.1) for 3D finite differences, since numerical tests indicate that the best results are obtained for  $\alpha$  in these intervals. These values are then substituted into the expression for  $z(\alpha)$ ; we then pick the  $\alpha$  that minimizes the average of the values of  $|z(\alpha; \theta_x, \theta_y) - 1|$ . This computation is inexpensive and imposes almost no overhead compared with the cost of solving the Stokes (or Oseen) problem. If desired, the results can be given in the form of a table giving the approximate value of  $\alpha_{opt}$  for each value of  $h$  and  $\nu$ .

Finally, we note that when dealing with unsteady problems we also have to take into account the mass matrices. For the purpose of the Fourier analysis, we use as the corresponding symbol  $h^2$  in 2D finite elements and 1 in 3D finite differences discretizations, to account for the different scaling used in two types of discretizations.

### 4.3 Numerical experiments

In this section we present the results of numerical experiments on saddle point systems arising from linearization and discretization of the test problems in Section 2.3. We consider only Oseen problems; for results on Stokes problems as well as the comparison between DS and RDF, see [12]. Note that the reported results for steady cases in [12] correspond to the linear system occurring at the 5th Picard iteration, while in this section we solve the linear system at the 1st Picard iteration, i.e., the one immediately after Stokes solve, as what was done for the AL preconditioners. In fact, the iteration count almost keeps the same during the Picard iteration, cf. Table 4.7. For

Table 4.1: GMRES(50) iterations with RDF preconditioner (cavity, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	11	11	12	12	32	14	59	23
$32 \times 32$	11	11	13	11	27	12	47	28
$64 \times 64$	11	11	13	9	13	10	40	21
$128 \times 128$	10	6	9	8	10	10	22	13

Table 4.2: GMRES(50) iterations with RDF preconditioner (cavity, Q2-Q1, stretched grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	14	14	15	14	31	16	43	23
$32 \times 32$	21	21	21	21	31	20	50	27
$64 \times 64$	19	17	25	16	26	18	65	28
$128 \times 128$	17	10	22	12	24	14	56	21

unsteady cases, we present results for a linear system occurring in the course of a simulation, before reaching the steady-state. We let  $\sigma = h^{-1}$  in (4.2); using a larger time step (say, five times as large) typically leads to a small increase in the number of iterations but the overall behavior of the tested preconditioner is the same.

Experimental results for the steady case are displayed in Tables 4.1–4.3; the middle table is for stretched grids (with the default parameter setting in IFISS), the remaining ones for uniform grids. For stretched grids (Table 4.2), we use the same strategy as in Section 2.5.2: Define the average mesh size  $h =: 2/m$  where  $m$  is the number of grid points in the x-direction (recall that here  $\Omega = [-1, 1] \times [-1, 1]$ ), and use it for Fourier analysis. This strategy

Table 4.3: GMRES(50) iterations with RDF preconditioner (step, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
Grid	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	21	15	62	14	81	15
$32 \times 32$	13	13	34	13	56	17
$64 \times 64$	12	12	22	14	34	16
$128 \times 128$	10	10	12	12	19	19

Table 4.4: GMRES(50) iterations with RDF preconditioner (unsteady cavity, Q2-Q1, uniform grids).

Viscosity	0.1	0.01	0.005	0.001
Grid	FA	FA	FA	FA
$16 \times 16$	11	10	10	8
$32 \times 32$	11	11	10	8
$64 \times 64$	10	10	8	4
$128 \times 128$	10	5	4	2

Table 4.5: GMRES(50) iterations with RDF preconditioner (unsteady cavity, Q2-Q1, stretched grids).

Viscosity	0.1	0.01	0.005	0.001
Grid	FA	FA	FA	FA
$16 \times 16$	13	11	11	9
$32 \times 32$	17	12	11	10
$64 \times 64$	14	9	5	2
$128 \times 128$	6	2	2	2

turns out to give good heuristic values in practice. Results for the unsteady problems are reported in Tables 4.4–4.6. Symmetric diagonal scaling is used

Table 4.6: GMRES(50) iterations with RDF preconditioner (unsteady step, Q2-Q1, uniform grids).

Viscosity	0.1	0.01	0.005
Grid	FA	FA	FA
$16 \times 16$	17	15	18
$32 \times 32$	21	26	20
$64 \times 64$	25	14	10
$128 \times 128$	14	8	7

for RDF.

For all problems we observe convergence rates that are independent of  $h$ , and  $\nu$ -independent convergence is obtained for unsteady problems. Note that for the lid driven cavity problem on the  $128 \times 128$  grid with very small viscosity ( $\nu = 0.005$  and  $0.001$ ), 4 or even 2 iterations are required to satisfy the desired tolerance. Fourier analysis accurately estimates the optimal  $\alpha$  only if the mesh is fine enough to yield physical solutions. Note that the results for the value of  $\alpha$  obtained by Fourier analysis ('FA estimate' in Tables 4.1 and 4.3) are essentially optimal, except for the cavity problem on stretched grids (Table 4.2).

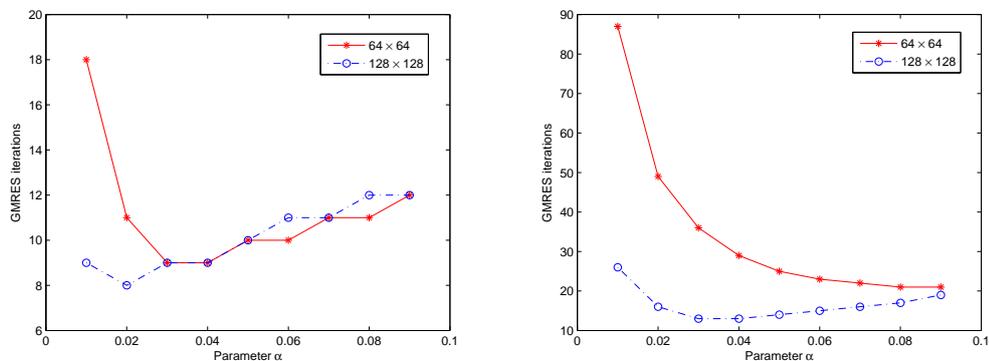
An important property of the RDF preconditioner is that both the optimal  $\alpha$  and the performance of the preconditioner remain virtually unchanged throughout the solution of the Navier–Stokes equation by Picard iteration. For the Q2-Q1 discretization of the lid driven cavity problem on the  $128 \times 128$  grid, this phenomenon is illustrated in Table 4.7, which displays the optimal value of  $\alpha$  and the number of linear iterations required at each of the first five Picard steps needed to solve the Navier–Stokes equations with  $\nu = 0.1, 0.01, 0.005$  and  $\nu = 0.001$ .

The dependence of the RDF preconditioned GMRES on the parameter  $\alpha$  is illustrated in Figure 4.3; these results are from tests on two representative

Table 4.7: GMRES(50) iterations with RDF preconditioner of the first five Picard iterations (Cavity, Q2-Q1, uniform  $128 \times 128$  grid, different viscosities).

Viscosity	0.1		0.01		0.005		0.001	
Picard	$\alpha_{opt}$	<i>its</i>	$\alpha_{opt}$	<i>its</i>	$\alpha_{opt}$	<i>its</i>	$\alpha_{opt}$	<i>its</i>
1	0.04	6	0.007	9	0.012	10	0.03	13
2	0.04	6	0.007	9	0.012	10	0.03	16
3	0.04	6	0.007	8	0.012	10	0.03	11
4	0.04	6	0.007	8	0.012	10	0.03	15
5	0.04	6	0.007	8	0.012	10	0.03	14

Figure 4.3: Number of iterations vs. parameter  $\alpha$  (cavity, Q2-Q1, uniform grids). Left:  $\nu = 0.01$ . Right:  $\nu = 0.001$ .



steady Oseen problems with  $\nu = 0.01$  and  $\nu = 0.001$ , using the  $64 \times 64$  and  $128 \times 128$  grids. We can see that the RDF preconditioner is not overly sensitive to the value of the parameter  $\alpha$ , in the sense that the iteration count does not change dramatically near the experimental optimal  $\alpha$ . We observe that there is a fairly wide range of values of the parameter  $\alpha$  that produce similar convergence results.

Table 4.8: GMRES(50) iteration counts and timings with RDF preconditioner (3D Oseen, MAC).

Viscosity	0.1	0.01	0.005
Grid	$its(\alpha_{opt})$	$its(\alpha_{opt})$	$its(\alpha_{opt})$
$16 \times 16 \times 16$	12 (1.6-3.6)	18 (13.1-14)	27 (21.6-21.9)
Setup time	0.11	0.48	0.13
Iter time	0.58	1.32	2.30
Total time	0.69	1.80	2.43
$32 \times 32 \times 32$	12 (2.3-2.4)	19 (12.3-14)	26 (20.5 - 22)
Setup time	1.60	1.89	2.45
Iter time	10.87	17.73	36.27
Total time	12.47	19.62	38.72
$48 \times 48 \times 48$	13 (1.7-3.0)	19 (12.6-14)	26 (20.4 - 22)
Setup time	7.04	9.43	8.61
Iter time	45.92	88.39	125.19
Total time	52.96	97.82	133.80
$64 \times 64 \times 64$	13 (2.1-2.6)	19 (12.7-14)	26 (20.6 - 22)
Setup time	18.24	23.46	25.22
Iter time	112.06	219.17	333.51
Total time	130.30	242.63	358.73

Next, we solve 3D steady and unsteady Oseen problems discretized by Marker-and-Cell [49] using GMRES(50) with the RDF preconditioner, which takes the form

$$\begin{pmatrix} A_1 & 0 & 0 & B_1^T \\ 0 & \alpha I & 0 & 0 \\ 0 & 0 & \alpha I & 0 \\ -B_1 & 0 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} \alpha I & 0 & 0 & 0 \\ 0 & A_2 & 0 & B_2^T \\ 0 & 0 & \alpha I & 0 \\ 0 & -B_2 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} \alpha I & 0 & 0 & 0 \\ 0 & \alpha I & 0 & 0 \\ 0 & 0 & A_3 & B_3^T \\ 0 & 0 & -B_3 & \alpha I \end{pmatrix}.$$

Table 4.9: GMRES(50) iteration counts and timings with RDF preconditioner (unsteady 3D Oseen, MAC).

Viscosity	0.1	0.01	0.005
Grid	$its(\alpha_{opt})$	$its(\alpha_{opt})$	$its(\alpha_{opt})$
$16 \times 16 \times 16$	12 (1.0 -1.6)	18(3.0 - 5.7)	20 (4.6-4.7)
Setup time	0.10	0.06	0.07
Iter time	0.61	0.67	0.76
Total time	0.71	0.73	0.83
$32 \times 32 \times 32$	14 (0.9-1.2)	21 (2.3 - 5.2)	25 (2.9-7.0)
Setup time	1.73	1.19	0.99
Iter time	13.92	18.44	17.96
Total time	15.65	19.63	18.95
$48 \times 48 \times 48$	15 (0.8-1.0)	22 (2.0 - 2.8)	27 (2.7-3.4)
Setup time	6.26	4.43	4.24
Iter time	54.60	71.35	92.70
Total time	60.86	75.78	96.94
$64 \times 64 \times 64$	16 (0.6-1.0)	22 (1.8-2.3)	28 (2.4 - 3)
Setup time	18.83	12.11	9.54
Iter time	156.69	175.91	177.72
Total time	175.52	188.02	187.26

The subproblems associated with  $\hat{A}_i = A_i + \alpha^{-1}B_i^T B_i$  ( $i = 1, 2, 3$ ) are solved inexactly via a single V-cycle of algebraic multigrid (AMG) with symmetric Gauss–Seidel as the smoother. The AMG implementation is the one provided by the Fortran code MI20; see [20, 21]. No diagonal scaling is applied here since the velocity mass matrix is the identity matrix. In Tables 4.8 and 4.9 we report results obtained using the optimal  $\alpha$  for the steady and unsteady case, respectively. When a range of values is reported for the optimal  $\alpha$ , the same number of iterations was observed for all values of  $\alpha$  in that range. Note

that the optimal  $\alpha$ 's are now much larger than in the case of finite elements, due to the different scaling of the matrix entries. Indeed, the stiffness matrix entries now scale as  $O(h^{-2})$  rather than being  $O(1)$  as in the case of finite elements. (The mass matrices now are just identity matrices, whereas the finite element mass matrices scale as  $O(h^2)$  for 2D and  $O(h^3)$  for 3D.) As in the 2D case, we set the time step to be the mesh size  $h$ . Once again, our results show that RDF is able to achieve convergence rates that are  $h$ -independent and only moderately dependent on  $\nu$ . The smallest value of the viscosity we consider is  $\nu = 0.005$ , since smaller values would require finer grids to give physically meaningful solutions.

We report setup times, iteration times and total solution times; note the very small setup times achieved by MI20 for these problems. The scalability with respect to problem size, while not perfect, appears to be quite good overall.

### 4.3.1 Comparison with other preconditioners

In this subsection we briefly compare RDF preconditioning with PCD, the modified PCD, LSC and the modified AL preconditioners. First comparing the results in the previous subsection and the Tables 2.18 and 2.19, all the PCD, the modified variant and LSC preconditioners require a significantly higher number of iterations than RDF.

To compare the RDF and the modified AL preconditioners, we apply the same diagonal scaling to the linear systems and then the preconditioners. The results for the steady lid driven cavity with Q2-Q1 finite element discretizations on stretched grids are presented in Table 4.10. The value of  $\gamma$  is chosen by Fourier analysis since it give almost optimal convergence rate. Comparing the iteration counts in Table 4.2, the modified AL preconditioner outperforms RDF clearly even with optimal  $\alpha$ . Furthermore, the iteration

Table 4.10: GMRES(50) iterations with diagonal scaling and modified AL preconditioner (cavity, Q2-Q1, stretched grids).

Viscosity	0.1	0.01	0.005	0.001
Grid	FA	FA	FA	FA
$16 \times 16$	9	9	19	32
$32 \times 32$	9	9	15	26
$64 \times 64$	8	9	11	22
$128 \times 128$	8	8	10	14

counts are smaller than those in Table 2.13, especially on the largest grid.

We also show GMRES(50) iterations and timings in Table 4.11 for unsteady step problem discretized by Q2-Q1 finite elements. The subproblems associated with  $\widehat{A}_1$  and  $\widehat{A}_2$  in RDF and  $A_{11}$ ,  $A_{22}$  and the compatible Laplacian in the modified AL preconditioner are performed inexactly by MI20. First, using inexact solves does not affect the convergence rate significantly, especially for problems with small viscosity. Next, very good scalability is achieved except for the case with  $256 \times 256$  grid and  $\nu = 0.1$ . Finally, the RDF and modified AL preconditioners give similar performance in terms of both GMRES iteration counts and timings.

## 4.4 The stabilized case

In this section we discuss a straightforward generalization of the RDF preconditioner to saddle point systems obtained by using stabilized finite element discretizations. The spectrum of the preconditioned matrix is analyzed, the technique based on Fourier analysis for estimating the relaxation parameter is described, and numerical results show the competitiveness of the RDF preconditioner.

Table 4.11: Comparison of RDF and modified AL preconditioners. GMRES(50) iterations and timings (unsteady step, Q2-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
Grid	RDF	Mod AL	RDF	Mod AL	RDF	Mod AL
$16 \times 16$	19	17	19	20	22	24
Setup time	0.15	0.03	0.14	0.39	0.18	0.41
Iter time	0.25	0.21	0.21	0.24	0.23	0.29
Total time	0.40	0.24	0.35	0.63	0.41	0.70
$32 \times 32$	23	25	31	30	24	26
Setup time	0.16	0.19	0.15	0.22	0.14	0.21
Iter time	1.44	1.87	1.60	2.72	1.23	2.28
Total time	1.60	2.06	1.75	2.94	1.37	2.49
$64 \times 64$	29	33	16	13	12	13
Setup time	0.72	0.82	0.74	1.00	0.78	1.04
Iter time	8.96	9.78	5.03	5.17	4.40	5.34
Total time	9.68	10.60	5.77	6.17	5.18	6.38
$128 \times 128$	20	23	9	10	9	10
Setup time	3.96	3.63	4.23	4.11	3.89	4.43
Iter time	30.84	31.50	16.23	16.29	16.32	19.11
Total time	34.80	35.13	20.46	20.40	20.21	23.54
$256 \times 256$	31	29	8	8	8	8
Setup time	15.02	15.00	17.15	15.36	19.07	17.18
Iter time	179.75	182.26	59.82	51.77	58.29	65.49
Total time	194.77	197.26	76.97	67.13	77.36	82.67

Using stabilized finite element discretizations, a pressure stabilization ma-

trix  $C$  is required; thus we consider solving

$$\begin{pmatrix} A & B^T \\ -B & C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ -g \end{pmatrix}, \quad \text{or} \quad \tilde{\mathcal{A}}x = \tilde{b}, \quad (4.11)$$

Note that the coefficient matrix in (4.11) has a spectrum entirely contained in the right half-plane.

The RDF preconditioner for (4.11) is the same as for linear system discretized by stable finite elements; see Section 4.2. By comparing the preconditioner (4.8) with (4.11), we can see that the difference between  $\mathcal{M}$  and  $\tilde{\mathcal{A}}$  is given by

$$\begin{aligned} \mathcal{R} &= \mathcal{M} - \tilde{\mathcal{A}} \\ &= \begin{pmatrix} 0 & -\alpha^{-1}B_1^T B_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha I - C \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\alpha^{-1}B_1^T B_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I - \alpha^{-1}C \end{pmatrix}. \end{aligned} \quad (4.12)$$

Following the same approach used in Section 4.2, we can prove the lemma below.

**Lemma 4.3.** *Recall  $\hat{A}_1 = A_1 + \alpha^{-1}B_1^T B_1$ ,  $S_1 = B_1 \hat{A}_1^{-1} B_1^T$ ,  $\hat{A}_2 = A_2 + \alpha^{-1}B_2^T B_2$  and  $S_2 = B_2 \hat{A}_2^{-1} B_2^T$ . Define*

$$T_{22} = \begin{pmatrix} \alpha^{-2} \hat{A}_2^{-1} B_2^T S_1 B_2 & (-\hat{A}_2^{-1} B_2^T + \alpha^{-1} \hat{A}_2^{-1} B_2^T S_1)(I - \alpha^{-1}C) \\ -\alpha^{-2} S_1 B_2 + \alpha^{-3} S_2 S_1 B_2 & \alpha^{-2}(\alpha I - S_2)(\alpha I - S_1)(I - \alpha^{-1}C) \end{pmatrix}.$$

*Then the eigenvalues of  $T_{22}$  are given by 0 with algebraic multiplicity at least  $n/2$ , and the remaining eigenvalues are  $1 - \mu_i$ , where  $\mu_i$  are the eigenvalues of the  $m \times m$  matrix  $Z_\alpha =: \alpha^{-1}(S_1 + S_2) - 2\alpha^{-2}S_1 S_2 + \alpha^{-3}(\alpha I - S_1)C(\alpha I - S_2)$ .*

*Proof.* Firstly, we observe that

$$\begin{aligned} T_{22} &= \alpha^{-2} \begin{pmatrix} -\widehat{A}_2^{-1} B_2^T \\ \alpha^{-1}(\alpha I - S_2) \end{pmatrix} \begin{pmatrix} -S_1 B_2 & \alpha(\alpha I - S_1)(I - \alpha^{-1}C) \end{pmatrix} \\ &= \alpha^{-2} X Y^T, \end{aligned} \quad (4.13)$$

where

$$X = \begin{pmatrix} -\widehat{A}_2^{-1} B_2^T \\ \alpha^{-1}(\alpha I - S_2) \end{pmatrix} \in \mathcal{R}^{(n/2+m) \times m}$$

and

$$Y^T = \begin{pmatrix} -S_1 B_2 & \alpha(\alpha I - S_1)(I - \alpha^{-1}C) \end{pmatrix} \in \mathcal{R}^{m \times (n/2+m)}.$$

Hence, from (4.13) we can see that  $T_{22}$  is a matrix with rank at most  $m$ . Therefore,  $T_{22}$  has an eigenvalue 0 of algebraic multiplicity at least  $n/2$ . By a well known result [52, Theorem 1.3.20], the remaining eigenvalues are the eigenvalues of the matrix

$$\begin{aligned} \alpha^{-2} Y^T X &= \alpha^{-2} (S_1 B_2 \widehat{A}_2^{-1} B_2^T + (\alpha I - S_1)(I - \alpha^{-1}C)(\alpha I - S_2)) \\ &= I + 2\alpha^{-2} S_1 S_2 - \alpha^{-1} (S_1 + S_2) - \alpha^{-3} (\alpha I - S_1) C (\alpha I - S_2) \\ &= I - Z_\alpha. \end{aligned}$$

□

Based on Lemma 1 [12] and Lemma 4.3, we have the following result.

**Theorem 4.4.** *The preconditioned matrix  $\mathcal{T} = \widetilde{\mathcal{A}}\mathcal{M}^{-1}$  has an eigenvalue at 1 with algebraic multiplicity at least  $n$ . The remaining eigenvalues are the eigenvalues  $\mu_i$  of the matrix  $Z_\alpha$ .*

*Proof.* First of all, from  $\widetilde{\mathcal{T}} =: \mathcal{M}^{-1}(\widetilde{\mathcal{A}}\mathcal{M}^{-1})\mathcal{M} = \mathcal{M}^{-1}\widetilde{\mathcal{A}}$  we see that the right-preconditioned matrix  $\mathcal{T}$  is *similar* to the left-preconditioned one  $\widetilde{\mathcal{T}}$ , so

$\mathcal{T}$  and  $\tilde{\mathcal{T}}$  have the same eigenvalues. We have

$$\begin{aligned}
\tilde{\mathcal{T}} &= I_{n+m} - \mathcal{M}^{-1}\mathcal{R} \\
&= I_{n+m} - \alpha\mathcal{M}_2^{-1}\mathcal{M}_1^{-1}\mathcal{R} \\
&= I_{n+m} - \begin{pmatrix} \alpha^{-1}I & 0 & 0 \\ 0 & \hat{A}_2^{-1} & -\alpha^{-1}\hat{A}_2^{-1}B_2^T \\ 0 & \alpha^{-1}B_2\hat{A}_2^{-1} & \alpha^{-1}I - \alpha^{-2}S_2 \end{pmatrix} \begin{pmatrix} \alpha\hat{A}_1^{-1} & 0 & -\hat{A}_1^{-1}B_1^T \\ 0 & I & 0 \\ B_1\hat{A}_1^{-1} & 0 & I - \alpha^{-1}S_1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} 0 & -\alpha^{-1}B_1^TB_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I - \alpha^{-1}C \end{pmatrix} \\
&= I_{n+m} - \begin{pmatrix} 0 & -\alpha^{-1}\hat{A}_1^{-1}B_1^TB_2 & -\hat{A}_1^{-1}B_1^T \\ 0 & \alpha^{-2}\hat{A}_2^{-1}B_2^TS_1B_2 & -\hat{A}_2^{-1}B_2^T + \alpha^{-1}\hat{A}_2^{-1}B_2^TS_1 \\ 0 & -\alpha^{-2}S_1B_2 + \alpha^{-3}S_2S_1B_2 & \alpha^{-2}(\alpha I - S_2)(\alpha I - S_1) \end{pmatrix} \\
&\quad \times \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I - \alpha^{-1}C \end{pmatrix} \\
&= I_{n+m} - \begin{pmatrix} 0 & T_{12} \\ 0 & T_{22} \end{pmatrix}, \tag{4.14}
\end{aligned}$$

where

$$T_{12} = \begin{pmatrix} -\alpha^{-1}\hat{A}_1^{-1}B_1^TB_2 & -\hat{A}_1^{-1}B_1^T(I - \alpha^{-1}C) \end{pmatrix}$$

and

$$T_{22} = \begin{pmatrix} \alpha^{-2}\hat{A}_2^{-1}B_2^TS_1B_2 & (-\hat{A}_2^{-1}B_2^T + \alpha^{-1}\hat{A}_2^{-1}B_2^TS_1)(I - \alpha^{-1}C) \\ -\alpha^{-2}S_1B_2 + \alpha^{-3}S_2S_1B_2 & \alpha^{-2}(\alpha I - S_2)(\alpha I - S_1)(I - \alpha^{-1}C) \end{pmatrix}.$$

According to Lemma 4.3, the eigenvalues of  $T_{22}$  are given by 0 with algebraic multiplicity  $n$  and  $1 - \mu_i (i = 1, 2, \dots, m)$ . Therefore, from (4.14) we can see that the eigenvalues of  $\mathcal{T}$  are given by 1 with algebraic multiplicity at least  $n$  and by the  $\mu_i$ 's.  $\square$

Additional information on the nonunit eigenvalues of the Oseen matrix with RDF preconditioner is given in the theorem below, which generalizes Theorem 2 in [12] to the stabilized case.

**Theorem 4.5.** *The eigenvalues  $\mu_i$  of  $Z_\alpha$  are of the form*

$$\mu_i = \frac{\alpha\xi_i}{1 + \alpha\xi_i},$$

where the  $\xi_i$ 's satisfy the generalized eigenvalue problem

$$(BA^{-1}B^T + C)\phi_i = \xi_i(\alpha^2 I + \widehat{S}_1 \widehat{S}_2 - \alpha C)\phi_i, \quad \text{with } \widehat{S}_k = B_k A_k^{-1} B_k^T \quad (k = 1, 2).$$

*Proof.* In the stabilized case, the matrix  $Z_\alpha$  has an additional term  $\alpha^{-3}(\alpha I - S_1)C(\alpha I - S_2)$ . Using the identity  $S_k = \alpha \widehat{S}_k(\alpha I + \widehat{S}_k)^{-1}$  in [12], we get

$$\begin{aligned} & \alpha^{-3}(\alpha I - S_1)C(\alpha I - S_2) \\ &= \alpha^{-3}(\alpha I - \alpha \widehat{S}_1(\alpha I + \widehat{S}_1)^{-1})C(\alpha I - \alpha \widehat{S}_2(\alpha I + \widehat{S}_2)^{-1}) \\ &= \alpha^{-1}(I - \widehat{S}_1(\alpha I + \widehat{S}_1)^{-1})C(I - \widehat{S}_2(\alpha I + \widehat{S}_2)^{-1}) \\ &= \alpha^{-1}(\alpha I(\alpha I + \widehat{S}_1)^{-1})C(\alpha I(\alpha I + \widehat{S}_2)^{-1}) \\ &= \alpha(\alpha I + \widehat{S}_1)^{-1}C(\alpha I + \widehat{S}_2)^{-1}. \end{aligned}$$

Therefore using equation (21) in [12], we obtain

$$Z_\alpha = \alpha(\alpha I + \widehat{S}_1)^{-1}(\widehat{S}_1 + \widehat{S}_2 + C)(\alpha I + \widehat{S}_2)^{-1}.$$

Since  $\widehat{S}_1 + \widehat{S}_2 + C = BA^{-1}B^T + C$ ,  $\mu_i$ 's are the eigenvalues of the following generalized eigenvalue problem:

$$\alpha(BA^{-1}B^T + C)\varphi_i = \mu_i(\alpha(\widehat{S}_1 + \widehat{S}_2) + \alpha^2 I + \widehat{S}_1 \widehat{S}_2)\varphi_i, \quad (4.15)$$

which can be rewritten as

$$(BA^{-1}B^T + C)\phi_i = \xi_i(\alpha^2 I + \widehat{S}_1 \widehat{S}_2 - \alpha C)\phi_i,$$

with  $\xi_i = \frac{\mu_i}{\alpha(1-\mu_i)}$ . It immediately follows that  $\mu_i = \frac{\alpha\xi_i}{1+\alpha\xi_i}$ , as claimed.  $\square$

Table 4.12: GMRES(50) iterations with RDF preconditioner (cavity, Q1-Q1, uniform grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	12	11	21	16	22	18	46	30
$32 \times 32$	10	10	15	12	16	15	28	23
$64 \times 64$	8	8	9	9	13	11	22	21
$128 \times 128$	7	5	9	8	12	11	14	14

Similar to the argument in [12], the previous theorem suggests that the non-unit eigenvalues of the preconditioned matrix  $\tilde{\mathcal{A}}\mathcal{M}^{-1}$  tend to 0 like  $O(\alpha)$  and to  $\infty$  like  $O(\alpha^{-1})$ , respectively.

Based on the previous results, Fourier analysis can be used to choose  $\alpha$ . The only difference to the case of stable finite elements is the presence of  $C$  in  $Z_\alpha$ , which is treated the same as we did in Section 3.1.3.

## 4.5 Numerical experiments

In Tables 4.12, 4.13 and 4.14 we present the results of numerical experiments with the RDF preconditioner for the three test problems discretized by stabilized Q1-Q1 finite elements. It is clear that the convergence rate of RDF-preconditioned GMRES is independent of grids and largely insensitive to viscosity. Furthermore, the values of  $\alpha$  chosen by the Fourier analysis-based approach are fairly close to the optimal values, especially on the finest grids.

Comparing with the results in Section 3.4, we can see that the RDF preconditioner performs similarly to the modified AL preconditioner, and RDF is even better for the lid driven cavity and step problems on uniform grids

Table 4.13: GMRES(50) iterations with RDF preconditioner (cavity, Q1-Q1, stretched grids).

Viscosity	0.1		0.01		0.005		0.001	
Grid	FA	Opt	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	28	26	34	26	31	28	64	63
$32 \times 32$	28	28	22	22	25	22	37	30
$64 \times 64$	26	21	21	18	20	20	40	31
$128 \times 128$	22	9	17	13	17	16	29	25

Table 4.14: GMRES(50) iterations with RDF preconditioner (step, Q1-Q1, uniform grids).

Viscosity	0.1		0.01		0.005	
Grid	FA	Opt	FA	Opt	FA	Opt
$16 \times 16$	15	13	36	16	50	20
$32 \times 32$	12	12	19	13	24	17
$64 \times 64$	11	11	13	13	18	16
$128 \times 128$	11	11	16	13	17	15

with small viscosity. Looking at the results of RDF preconditioner for stable finite elements in Tables 4.1, 4.2 and 4.3, one can observe that the two sets of results are quite similar.

# Chapter 5

## Conclusions

In this thesis, we have studied various preconditioning techniques for the incompressible Navier–Stokes equations discretized by stable and stabilized finite elements as well as the Marker-and-Cell finite difference method. We considered both steady and unsteady problems in 2D and 3D domains. Numerical results on parallel architecture were also presented. Comparison with other state-of-the-art preconditioners shows the competitiveness of our preconditioners.

We first studied the ideal and modified augmented Lagrangian-based preconditioners for Oseen-type problems. We have also extended the AL-based approach for stable finite elements to the case of stabilized finite element pairs, and analyzed the spectral properties of saddle point matrices preconditioned with such techniques. One advantage of the modified AL preconditioners is that they can be readily implemented using standard off-the-shelf algebraic multilevel solvers developed for elliptic PDEs, in particular parallel AMG-type solvers. Our theory, together with a form of Fourier analysis, provides an inexpensive way to select the augmentation parameter  $\gamma$  in a nearly optimal way for the modified variants.

The preconditioner performance has been thoroughly investigated on a variety of benchmark 2D problems. Our numerical experiments show excellent performance of the modified AL preconditioner for a wide range of problem parameters. The preconditioner is able to handle small viscosities and

stretched grids, and is found to be generally superior to (and more robust than) some of the best existing preconditioners. Numerical experiments on linear systems arising from the Newton linearization show that using the modified AL preconditioner with the same values of the parameter  $\gamma$  found using Fourier analysis for the Oseen problem gives surprisingly good results.

We have also investigated the use of inexact solves for the velocity subsystems arising in the application of the preconditioner. We found that for sufficiently large problems, replacing exact solves with one iteration of an algebraic multigrid method with an appropriate smoother yields rates of convergence that are nearly as good as those obtained with exact solves while significantly reducing total computing times. This holds for Q2-Q1 discretizations and for both the Oseen (Picard) and Newton linearizations. The resulting solver shows excellent scalability in terms of solution times.

Furthermore, the modified AL preconditioner has been implemented in the framework of Trilinos, and further incorporated into the finite element library LifeV. Parallel results on various 2D and 3D problems discretized by finite elements and finite differences show good performance of this preconditioner, which is an effective approach for the parallel solution of the incompressible Navier–Stokes equations.

In addition, we have studied a novel relaxed dimensional factorization preconditioner for solving saddle point systems. Although the preconditioner can be applied to rather general linear systems in saddle point form, in this thesis we have focused on discretizations of systems of PDEs arising in incompressible fluid flow simulations. Some results on the eigenvalues of the preconditioned matrices have been obtained, and an inexpensive technique for estimating the relaxation parameter has been described based on Fourier analysis.

Numerical experiments on a variety of test cases indicate very fast convergence of RDF-preconditioned GMRES independent of mesh size in the case

of uniform grids. The convergence rate is only moderately affected by the viscosity  $\nu$ , and appears to be much less sensitive to it than other current approaches (including the closely related DS preconditioner). The convergence behavior is also quite good for problems posed on stretched grids. In spite of some deterioration of the preconditioner quality in the steady case for very low viscosity values, RDF appears to be quite competitive when compared to some of the best existing methods.

Efficient implementation of the RDF preconditioner in 3D requires the use of inexact (inner) solves. Our experiments indicate that the excellent convergence properties of the ‘exact’ RDF preconditioner are retained even when the inner solves are performed with low accuracy. In our 3D tests we have used a single AMG V-cycle (with symmetric Gauss–Seidel smoothing) for the inner solves. The resulting solver appears to scale reasonably well.

# Bibliography

- [1] The LifeV project. <http://www.lifev.org>.
- [2] P. R. Amestoy, T. A. Davis, and I. S. Duff. Algorithm 837: AMD, an approximate minimum degree ordering algorithm. *ACM Trans. Math. Software*, 30(3):381–388, 2004.
- [3] Z.-Z. Bai. Structured preconditioners for nonsingular matrices of block two-by-two structures. *Math. Comp.*, 75(254):791–815 (electronic), 2006.
- [4] Z.-Z. Bai, G. H. Golub, and M. K. Ng. Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. *SIAM J. Matrix Anal. Appl.*, 24(3):603–626, 2003.
- [5] Z.-Z. Bai and M. K. Ng. On inexact preconditioners for nonsymmetric matrices. *SIAM J. Sci. Comput.*, 26(5):1710–1724 (electronic), 2005.
- [6] R. Becker and M. Braack. A finite element pressure gradient stabilization for the Stokes equations based on local projections. *Calcolo*, 38(4):173–199, 2001.
- [7] M. Benzi, M. J. Gander, and G. H. Golub. Optimization of the Hermitian and skew-Hermitian splitting iteration for saddle-point problems. *BIT*, 43:881–900, 2003.

- [8] M. Benzi and G. H. Golub. A preconditioner for generalized saddle point problems. *SIAM J. Matrix Anal. Appl.*, 26(1):20–41, 2004.
- [9] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.
- [10] M. Benzi and X.-P. Guo. A dimensional split preconditioner for Stokes and linearized Navier-Stokes equations. *Appl. Numer. Math.*, 61(1):66–76, 2011.
- [11] M. Benzi and J. Liu. An efficient solver for the incompressible Navier–Stokes equations in rotation form. *SIAM J. Sci. Comput.*, 29(5):1959–1981, 2007.
- [12] M. Benzi, M. Ng, Q. Niu, and Z. Wang. A relaxed dimensional factorization preconditioner for the incompressible Navier–Stokes equations. *J. Comput. Phys.*, 230(16):6185–6202, 2011.
- [13] M. Benzi and M. A. Olshanskii. An augmented Lagrangian-based approach to the Oseen problem. *SIAM J. Sci. Comput.*, 28(6):2095–2113, 2006.
- [14] M. Benzi and M. A. Olshanskii. Field-of-values convergence analysis of augmented Lagrangian preconditioners for the linearized Navier–Stokes problem. *SIAM J. Numer. Anal.*, 49(2):770–778, 2011.
- [15] M. Benzi, M. A. Olshanskii, and Z. Wang. Modified augmented Lagrangian preconditioners for the incompressible Navier–Stokes equations. *Internat. J. Numer. Methods Fluids*, 66(4):486–508, 2011.
- [16] M. Benzi and V. Simoncini. On the eigenvalues of a class of saddle point matrices. *Numer. Math.*, 103(2):173–196, 2006.

- [17] M. Benzi and Z. Wang. Analysis of augmented Lagrangian-based preconditioners for the steady incompressible Navier–Stokes equations. Technical Report Math/CS TR-2010-011, Department of Mathematics and Computer Science, Emory University, June 2010. To appear in *SIAM J. Sci. Comput.*
- [18] M. Benzi and A. J. Wathen. Some preconditioning techniques for saddle point problems. In *Model Order Reduction: Theory, Research Aspects and Applications*, volume 13 of *Math. Ind.*, pages 195–211. Springer, Berlin, 2008.
- [19] S. Börm and S. Le Borne.  $\mathcal{H}$ -LU factorization in preconditioners for augmented Lagrangian and grad-div stabilized saddle point systems. *Internat. J. Numer. Methods Fluids*, DOI: 10.1002/fld.2495, 2010.
- [20] J. Boyle, M. Mihajlović, and J. Scott. HSL\_MI20: an efficient AMG preconditioner. Technical Report RAL-TR-2007-021, Rutherford Appleton Laboratory, Chilton, Oxfordshire, UK, December 2007.
- [21] J. Boyle, M. Mihajlović, and J. Scott. HSL\_MI20: an efficient AMG preconditioner for finite element problems in 3D. *Internat. J. Numer. Methods Engrg.*, 82(1):64–98, 2010.
- [22] F. Brezzi and J. Pitkäranta. On the stabilization of finite element approximations of the Stokes equations. In *Efficient Solutions of Elliptic Systems (Kiel, 1984)*, volume 10 of *Notes Numer. Fluid Mech.*, pages 11–19. Vieweg, Braunschweig, 1984.
- [23] R. Bridson. KKTDirect: A direct solver package for saddle-point (KKT) matrices. <http://www.cs.ubc.ca/~rbridson/kktdirect/>.

- [24] J. Cahouet and J.-P. Chabard. Some fast 3D finite element solvers for the generalized Stokes problem. *Internat. J. Numer. Methods Fluids*, 8(8):869–895, 1988.
- [25] Y. Cao, M.-Q. Jiang, and Y.-L. Zheng. A splitting preconditioner for saddle point problems. *Numer. Linear Algebra Appl.*, DOI: 10.1002/nla.772, 2011.
- [26] Z.-H. Cao. Augmentation block preconditioners for saddle point-type matrices with singular  $(1, 1)$  blocks. *Numer. Linear Algebra Appl.*, 15(6):515–533, 2008.
- [27] L. C. Chan, M. K. Ng, and N. K. Tsing. Spectral analysis for HSS preconditioners. *Numer. Math. Theory Methods Appl.*, 1(1):57–77, 2008.
- [28] G.-H. Cheng, T.-Z. Huang, and S.-Q. Shen. Block triangular preconditioners for the discretized time-harmonic Maxwell equations in mixed form. *Comput. Phys. Comm.*, 180(2):192–196, 2009.
- [29] R. Codina. A stabilized finite element method for generalized stationary incompressible flows. *Comput. Methods Appl. Mech. Engrg.*, 190(20-21):2681–2706, 2001.
- [30] A. C. de Niet and F. W. Wubs. Two preconditioners for saddle point problems in fluid flows. *Internat. J. Numer. Methods Fluids*, 54(4):355–377, 2007.
- [31] A. C. de Niet and F. W. Wubs. Numerically stable  $LDL^T$ -factorization of  $\mathcal{F}$ -type saddle point matrices. *IMA J. Numer. Anal.*, 29(1):208–234, 2009.
- [32] C. R. Dohrmann and P. B. Bochev. A stabilized finite element method for the Stokes problem based on polynomial pressure projections. *Internat. J. Numer. Methods Fluids*, 46(2):183–201, 2004.

- [33] H. Elman, V. E. Howle, J. Shadid, D. Silvester, and R. Tuminaro. Least squares preconditioners for stabilized discretizations of the Navier–Stokes equations. *SIAM J. Sci. Comput.*, 30(1):290–311, 2007.
- [34] H. Elman and D. Silvester. Fast nonsymmetric iterations and preconditioning for Navier–Stokes equations. *SIAM J. Sci. Comput.*, 17(1):33–46, 1996.
- [35] H. C. Elman, A. Ramage, and D. J. Silvester. Algorithm 886: IFISS, a Matlab toolbox for modelling incompressible flow. *ACM Trans. Math. Software*, 33(2):2–14, 2007.
- [36] H. C. Elman, D. J. Silvester, and A. J. Wathen. Performance and analysis of saddle point preconditioners for the discrete steady-state Navier–Stokes equations. *Numer. Math.*, 90(4):665–688, 2002.
- [37] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005.
- [38] H. C. Elman and R. S. Tuminaro. Boundary conditions in approximate commutator preconditioners for the Navier–Stokes equations. *Electron. Trans. Numer. Anal.*, 35:257–280, 2009.
- [39] L. Formaggia, A. Quarteroni, and A. Veneziani (Eds.). *Cardiovascular Mathematics*. Springer-Verlag, New York, 2009.
- [40] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems*, volume 15 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1983.

- [41] L. P. Franca and A. Russo. Approximation of the Stokes problem by residual-free macro bubbles. *East-West J. Numer. Math.*, 4(4):265–278, 1996.
- [42] M. W. Gee, C. M. Siefert, J. J. Hu, R. S. Tuminaro, and M. G. Sala. ML 5.0 smoothed aggregation user’s guide. Technical Report SAND2006-2649, Sandia National Laboratories, 2006.
- [43] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
- [44] C. Greif and M. L. Overton. An analysis of low-rank modifications of preconditioners for saddle point systems. *Electron. Trans. Numer. Anal.*, 37:307–320, 2010.
- [45] C. Greif and D. Schötzau. Preconditioners for saddle point linear systems with highly singular  $(1, 1)$  blocks. *Electron. Trans. Numer. Anal.*, 22:114–121 (electronic), 2006.
- [46] C. Greif and D. Schötzau. Preconditioners for the discretized time-harmonic Maxwell equations in mixed form. *Numer. Linear Algebra Appl.*, 14(4):281–297, 2007.
- [47] P. M. Gresho and R. L. Sani. *Incompressible Flow and the Finite Element Method. Volume 2: Isothermal Laminar Flow*. John Wiley and Sons, New York, 1998.
- [48] S. Hamilton, M. Benzi, and E. Haber. New multigrid smoothers for the Oseen problem. *Numer. Linear Algebra Appl.*, 17(2-3):557–576, March 2010.

- [49] F. H. Harlow and J. E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids*, 8(12):2182–2189, 1965.
- [50] M. A. Heroux. AztecOO user’s guide. Technical Report SAND2004-3796, Sandia National Laboratories, 2004.
- [51] M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. M. Willenbring, A. Williams, and K. S. Stanley. An overview of the Trilinos project. *ACM Trans. Math. Software*, 31(3):397–423, 2005.
- [52] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.
- [53] T.-Z. Huang, L.-T. Zhang, T.-X. Gu, and X.-Y. Zuo. New block triangular preconditioner for linear systems arising from the discretized time-harmonic Maxwell equations. *Comput. Phys. Comm.*, 180(10):1853–1859, 2009.
- [54] T. J. R. Hughes, L. P. Franca, and M. Balestra. A new finite element formulation for computational fluid dynamics. V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Comput. Methods Appl. Mech. Engrg.*, 59(1):85–99, 1986.
- [55] I. C. F. Ipsen. A note on preconditioning nonsymmetric matrices. *SIAM J. Sci. Comput.*, 23(3):1050–1051 (electronic), 2001.
- [56] V. John and A. Kindl. Numerical studies of finite element variational multiscale methods for turbulent flow simulations. *Comput. Methods Appl. Mech. Engrg.*, 199(13-16):841–852, 2010.

- [57] A. Linke. Collision in a cross-shaped domain—A steady 2d Navier–Stokes example demonstrating the importance of mass conservation in CFD. *Comput. Methods Appl. Mech. Engrg.*, 198(41-44):3278–3286, 2009.
- [58] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 21(6):1969–1972 (electronic), 2000.
- [59] M. Olshanskii, G. Lube, T. Heister, and J. Löwe. Grad-div stabilization and subgrid pressure models for the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 198(49-52):3975–3988, 2009.
- [60] M. A. Olshanskii. An iterative solver for the Oseen problem and numerical solution of incompressible Navier–Stokes equations. *Numer. Linear Algebra Appl.*, 6(5):353–378, 1999.
- [61] M. A. Olshanskii and M. Benzi. An augmented Lagrangian approach to linearized problems in hydrodynamic stability. *SIAM J. Sci. Comput.*, 30(3):1459–1473, 2008.
- [62] M. A. Olshanskii and A. Reusken. Grad-div stabilization for Stokes equations. *Math. Comp.*, 73(248):1699–1718 (electronic), 2004.
- [63] M. A. Olshanskii and Y. V. Vassilevski. Pressure Schur complement preconditioners for the discrete Oseen problem. *SIAM J. Sci. Comput.*, 29(6):2686–2704 (electronic), 2007.
- [64] I. Perugia and V. Simoncini. Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations. *Numer. Linear Algebra Appl.*, 7(7-8):585–616, 2000.

- [65] T. Rees and C. Greif. A preconditioner for linear systems arising from interior point optimization methods. *SIAM J. Sci. Comput.*, 29(5):1992–2007 (electronic), 2007.
- [66] T. Rusten and R. Winther. A preconditioned iterative method for saddlepoint problems. *SIAM J. Matrix Anal. Appl.*, 13(3):887–904, 1992.
- [67] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
- [68] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.
- [69] J. Schöberl. Multigrid methods for a parameter dependent problem in primal variables. *Numer. Math.*, 84(1):97–119, 1999.
- [70] D. J. Silvester, H. C. Elman, and A. Ramage. IFISS: Incompressible Flow Iterative Solution Software, January 2011. <http://www.manchester.ac.uk/ifiss/>.
- [71] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1993.
- [72] R. Temam. *Navier–Stokes Equations*. AMS Chelsea Publishing, Providence, RI, 2001.
- [73] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press Inc., San Diego, CA, 2001.

- [74] S. Turek. *Efficient Solvers for Incompressible Flow Problems*, volume 6 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 1999.
- [75] M. ur Rehman, C. Vuik, and G. Segal. A comparison of preconditioners for incompressible Navier–Stokes solvers. *Internat. J. Numer. Methods Fluids*, 57(12):1731–1751, 2008.
- [76] M. ur Rehman, C. Vuik, and G. Segal. Preconditioners for the steady incompressible Navier–Stokes problem. *IAENG Int. J. Appl. Math.*, 38(4):223–232, 2008.
- [77] M. ur Rehman, C. Vuik, and G. Segal. SIMPLE-type preconditioners for the Oseen problem. *Internat. J. Numer. Methods Fluids*, 61(4):432–452, 2009.
- [78] P. Wesseling. *Principles of Computational Fluid Dynamics*, volume 29 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2001.
- [79] R. Wienands and W. Joppich. *Practical Fourier Analysis for Multigrid Methods*, volume 4 of *Numerical Insights*. Chapman & Hall/CRC, Boca Raton, FL, 2005.