Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2017.m1637

Vol. **10**, No. 3, pp. 639-655
August 2017

A Full Multigrid Method for Distributed Control Problems Constrained by Stokes Equations

M. M. Butt[1,*] and Y. Yuan[2]

[1] *Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing 100080, China; and Higher Education Department, Government of the Punjab, Lahore 54000, Pakistan*
[2] *Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing 100080, China*

Received 7 September 2016; Accepted (in revised version) 5 January 2017

Abstract. A full multigrid method with coarsening by a factor-of-three to distributed control problems constrained by Stokes equations is presented. An optimal control problem with cost functional of velocity and/or pressure tracking-type is considered with Dirichlet boundary conditions. The optimality system that results from a Lagrange multiplier framework, form a linear system connecting the state, adjoint, and control variables. We investigate multigrid methods with finite difference discretization on staggered grids. A coarsening by a factor-of-three is used on staggered grids that results nested hierarchy of staggered grids and simplified the inter-grid transfer operators. A distributive-Gauss-Seidel smoothing scheme is employed to update the state- and adjoint-variables and a gradient update step is used to update the control variables. Numerical experiments are presented to demonstrate the effectiveness and efficiency of the proposed multigrid framework to tracking-type optimal control problems.

AMS subject classifications: 35Q93, 65N55, 49K20

Key words: PDE-constrained optimization, stokes equations, multigrid, staggered grids, finite difference.

1. Introduction

For the last decade, there is a growing interest among computation community to devise fast and efficient solution methods for large-scale distributed optimal control problems constrained by partial differential equations (PDEs). Optimal control problems constrained by the Stokes system form a stepping stone in the natural progression from the - now classical-Poisson-constrained test problem to problems constrained by more specialized and complex PDE systems modeling fluid flow such as Navier-Stokes, non-Newtonian Stokes, or the shallow water equations. Optimal control problems constrained by such PDE

*Corresponding author. *Email addresses:* mmunirbutt@gmail.com (M. M. Butt), yyx@lsec.cc.ac.cn (Y. Yuan)

models play important roles in real-world applications, such as modeling of ice sheets or data assimilation for ocean flows and weather models.

How to effectively solve large scale algebraic systems arising from the discretization of PDEs is a fundamental question in scientific and engineering computing. For the positive definite linear systems corresponding to elliptic boundary value problems, multigrid methods are proven to be one of the most efficient algorithms [4, 13, 23]. However, it is much more challenging for saddle-point systems [2]. In the last few decades, much attention has been given to numerical solution of Stokes and/or Navier-Stokes systems, e.g., a distributive Gauss-Seidel relaxation based on the least squares commutator is devised for the saddle-point systems arising from the discretized Stokes equations in [22]. In [1], a distributed relaxation method for the incompressible Stokes problem has been advertised. A large selection of solution methods for linear systems in saddle point form is presented with an emphasis on iterative methods for large and sparse problems in [2].

As a motivation of this work, suppose that we have a flow that satisfies the Stokes equations in some domain with some given boundary condition, and then we have some mechanism (e.g., a magnetic field application) to change the forcing variable of the PDE. Suppose we have given functions, the so-called desired states. Then the question is how do we choose the forcing term, while satisfying the Stokes equations? One way of formulating such problem is by minimizing a cost functional of tracking-type with constrained PDE as Stokes equations, is presented in this article.

This paper aims at to construct an efficient multigrid scheme without any preconditioners, on staggered grids to solve (velocity and/or pressure tracking-type) distributed optimal control problem constrained by the Stokes equations. A significant amount of work has been devoted to develop multigrid methods for optimal control problems in the recent years, for example, see a review article [3] and the references therein. However, less attention has been given specifically to optimal control problems constrained by the Stokes system. In [9], multigrid preconditioners to accelerate the solution process of a distributed optimal control problem constrained by the Stokes equations are constructed. Recently, [19] proposed a robust all-at-once multigrid method for the Stokes control problem. This article extends the work [7], where a formulation and multigird solution for Cauchy-Riemann optimal control problems has been presented to control problems constrained by Stokes equations. Adopting a coarsening by factor-of-three strategy on staggered grids to these distributed control problems has the potential advantage of simplifying the inter-grid transfer operators, coarsening more quickly, and ultimately reducing the number of levels and parallel computations.

The paper is organized as follows. In the next Section 2, an optimal control problem with a cost functional of tracking-type is considered and its solution is characterized as an optimality system, in a two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$. Discretization with finite difference on staggered grids, on non-uniform meshes, is presented in Section 3. In Section 4, a full multigrid scheme is discussed for solving the optimality system. It is shown that a coarsening by a factor-of-three of the mesh sizes is advantageous, easy to implement inter-grid transfer operators, and a nested hierarchy of staggered grids is obtained. A distributive-Gauss-Seidel relaxation scheme with inter-grid transfer operators is explained

in Section 4.1. Numerical results are presented in Section 5 to illuminate our full multigrid solution procedure to tracking-type control problems. A section of conclusions is given in the last section.

2. Control problem

In this work, we consider an optimal control problem with a cost functional of (velocity and/or pressure) tracking-type, in a two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$:

$$\min J(\mathbf{u},p,\mathbf{f}) := \frac{\alpha_{\mathbf{u}}}{2}\|\mathbf{u}-\mathbf{U}\|^2_{\mathbf{L}^2(\Omega)} + \frac{\alpha_p}{2}\|p-P\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|\mathbf{f}\|^2_{L^2(\Omega)} \tag{2.1}$$

subject to the Stokes equations

$$-\triangle\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \tag{2.2a}$$
$$-\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega, \tag{2.2b}$$
$$\mathbf{u} = 0 \quad \text{on } \partial\Omega, \tag{2.2c}$$

where $\mathbf{U} \in \mathbf{L}^2(\Omega)$, $P \in L^2_0(\Omega)$ denotes the target velocity, respectively, target pressure and $\alpha > 0$ being a fixed positive constant represents the weight of the cost of control $\mathbf{f}$. The constants $\alpha_{\mathbf{u}}$, α_p are nonnegative but not both zero.

Here and in the following, $L^2(\Omega)$ and $H^1(\Omega)$ denote the standard Lebesque and Sobolev spaces with $\|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{H^1(\Omega)}$, respectively, as associated standard norms. The usual inner product associated with $L^2(\Omega)$ will be denoted by $(\cdot,\cdot)$. Throughout this paper, we will assume Ω to be convex and use bold script to denote vectors and product spaces and

$$L^2_0(\Omega) = \left\{ q \in L^2(\Omega) : \int_\Omega q\,dx = 0 \right\}, \qquad \mathbf{H}^1_0(\Omega) = \left\{ \mathbf{w} \in \mathbf{H}^1(\Omega) : \mathbf{w} = 0 \text{ on } \partial\Omega \right\}.$$

The weak solution

$$(\mathbf{u},p) \in \mathbf{H}^1_0(\Omega) \times L^2_0(\Omega)$$

of (2.2a)-(2.2c) is the solution of

$$a(\mathbf{u},\mathbf{w}) + b(\mathbf{w},p) = (\mathbf{f},\mathbf{w}) \qquad \forall\, \mathbf{w} \in \mathbf{H}^1_0(\Omega),$$
$$b(\mathbf{u},q) = 0 \qquad \forall\, q \in L^2_0(\Omega),$$

where $a : \mathbf{H}^1_0(\Omega) \times \mathbf{H}^1_0(\Omega) \to \mathbb{R}$ and $b : \mathbf{H}^1_0(\Omega) \times L^2_0(\Omega)$ are the bilinear forms defined as

$$a(\mathbf{u},\mathbf{w}) = \sum_{i=1}^{2} \int_\Omega \nabla\mathbf{u}_i \cdot \nabla\mathbf{w}_i, \qquad b(\mathbf{w},p) = -\int_\Omega p\,\nabla\cdot\mathbf{w}.$$

For $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$ the problem has a unique solution [11]. Moreover, if Ω is a convex polygon and $\mathbf{f} \in \mathbf{L}^2(\Omega)$, then $\mathbf{u} \in \mathbf{H}^2(\Omega)$, $p \in H^1(\Omega)$ [12] and there exists $C = C(\Omega) > 0$ satisfying

$$\|\mathbf{u}\|_{\mathbf{H}^2(\Omega)} + \|\nabla p\|_{L^2(\Omega)} \le C\|\mathbf{f}\|_{\mathbf{L}^2(\Omega)}.$$

The solution to the problem (2.1)-(2.2c) is characterized by the Lagrange framework [15, 17, 20], and the Lagrangian functional is given by

$$L(\mathbf{u}, p, \mathbf{f}, \lambda, q) := J(\mathbf{u}, p, \mathbf{f}) + (-\Delta\mathbf{u} + \nabla p - \mathbf{f}, \lambda) - (\nabla \cdot \mathbf{u}, q)$$

where $\lambda \in \mathbf{H}_0^1(\Omega)$ and $q \in L_0^2(\Omega)$ are the Lagrange variables, which leads to the following optimality system

$$-\Delta\mathbf{u} + \nabla p = \mathbf{f}, \qquad -\nabla \cdot \mathbf{u} = 0 \ \text{ in } \ \Omega, \qquad \mathbf{u} = 0 \ \text{ on } \ \partial\Omega, \tag{2.3a}$$

$$-\Delta\lambda + \nabla q = \alpha_{\mathbf{u}}(\mathbf{U} - \mathbf{u}), \qquad -\nabla \cdot \lambda = \alpha_p(P - p) \ \text{ in } \ \Omega, \quad \lambda = 0 \ \text{ on } \ \partial\Omega, \tag{2.3b}$$

$$\alpha\mathbf{f} - \lambda = 0 \ \text{ in } \ \Omega. \tag{2.3c}$$

This system characterizes the solution

$$(\mathbf{u}, p, \mathbf{f}) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \mathbf{L}^2(\Omega)$$

of the optimal control problem (2.1)-(2.2c) with Lagrange variables

$$(\lambda, q) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega).$$

Moreover, (2.3a) represents the state system, (2.3b) the adjoint system and (2.3c) the optimality conditions, respectively;

$$\mathbf{u} = (u, v), \qquad \lambda = (\lambda, \mu),$$
$$\mathbf{f} = (f, g), \qquad \mathbf{U} = (U, V).$$

In the following Theorem 2.1, we present the first-order optimality conditions for the purpose of completeness.

Theorem 2.1. *Assume existence and uniqueness of solution to the Stokes (state) system* (2.2a)-(2.2c) *and define the gradient* $\nabla\hat{J}(\mathbf{f}) := \alpha\mathbf{f} - \lambda$ *where* λ *and* q *are the solution to the adjoint system* (2.3b). *Then the control problem* (2.1)-(2.2c) *has a unique solution in* $\mathbf{L}^2(\Omega)$ *if and only if* $\nabla\hat{J}(\mathbf{f}) = 0$. *That is, the optimal solution is characterized as the solution of the first-order optimality system.*

Proof. Since the Stokes system (2.2a)-(2.2c) has a unique solution $(\mathbf{u}, p)$ for given $\mathbf{f}$, therefore we can write $(\mathbf{u}, p) = \big(\mathbf{u}(\mathbf{f}), p(\mathbf{f})\big)$ to show this dependence.

To discuss existence and uniqueness of the solution to (2.1)-(2.2c), we introduced a so-called reduced cost functional $\hat{J}$ [15, 17, 20] given by

$$\hat{J}(\mathbf{f}) = J\left(\mathbf{u}(\mathbf{f}), p(\mathbf{f}), \mathbf{f}\right). \tag{2.4}$$

In fact, the unique minimizer of $\hat{J}(\mathbf{f})$ gives the optimal solution. Recall that the map

$$\mathbf{L}^2(\Omega) \ni \mathbf{f} \mapsto \big(\mathbf{u}(\mathbf{f}), p(\mathbf{f})\big) \in \Big(\mathbf{H}_0^1(\Omega) \bigcap \mathbf{H}^2(\Omega), L_0^2(\Omega)\Big)$$

is Fréchet differentiable. Let its first derivative be denoted by $\left(\mathbf{u}'(\mathbf{f})\,\delta\mathbf{f}, p'(\mathbf{f})\,\delta\mathbf{f}\right)$ at $\mathbf{f}$ in the direction $\delta\mathbf{f}$, and is characterized as the solution to (the so-called tangent equation)

$$-\Delta(\mathbf{u}'(\mathbf{f})\,\delta\mathbf{f}) + \nabla(p'(\mathbf{f})\,\delta\mathbf{f}) = \delta\mathbf{f}, \tag{2.5a}$$

$$-\nabla\cdot(\mathbf{u}'(\mathbf{f})\,\delta\mathbf{f}) = 0, \tag{2.5b}$$

subject to Dirichlet boundary condition. Because of linear dependence of the governing Stokes system on the control, the second derivative of $\mathbf{f}\mapsto(\mathbf{u}(\mathbf{f}), p(\mathbf{f}))$ is zero.

Using (2.1) we get the second derivative of $\mathbf{f}\to\hat{J}(\mathbf{f})$ as

$$\hat{J}''(\mathbf{f})(\delta\mathbf{f},\delta\mathbf{f}) = \alpha_{\mathbf{u}}\|\mathbf{u}'(\mathbf{f})\,\delta\mathbf{f}\|^2_{\mathbf{L}^2(\Omega)} + \alpha_p\|p'(\mathbf{f})\,\delta\mathbf{f}\|^2_{L^2(\Omega)} + \alpha\,\|\delta\mathbf{f}\|^2_{L^2(\Omega)},$$

that results that $\mathbf{f}\to\hat{J}(\mathbf{f})$ is uniformly convex. This implies existence of a unique solution $\mathbf{f}^*$ to the optimal control problem, denoted by *. Furthermore, the minimum is characterized by $\hat{J}'(\mathbf{f}^*)(\delta\mathbf{f}) = 0$ for all $\delta\mathbf{f}\in L^2(\Omega)$ and thus

$$\hat{J}'(\mathbf{f}^*)(\delta\mathbf{f}) = \alpha_{\mathbf{u}}\left(\mathbf{u}^*-\mathbf{U}, \mathbf{u}'(\mathbf{f}^*)\,\delta\mathbf{f}\right)_{\mathbf{L}^2(\Omega)} + \alpha_p\left(p^*-P, p'(\mathbf{f}^*)\,\delta\mathbf{f}\right)_{L^2(\Omega)} + \alpha(\mathbf{f}^*,\delta\mathbf{f})_{L^2(\Omega)}, \tag{2.6}$$

where $\mathbf{u}^* = \mathbf{u}(\mathbf{f}^*)$. To remove $\mathbf{u}'$ and p' from this equation, we introduce the Lagrange variables $(\lambda^*, q^*)\in\mathbf{H}^1_0(\Omega)\times L^2_0(\Omega)$ as the unique solution to the adjoint system (2.3b). Since

$$-\alpha_{\mathbf{u}}(\mathbf{u}^*-\mathbf{U}) = -\Delta\lambda^* + \nabla q^*,$$
$$-\alpha_p(p^*-P) = -\nabla\cdot\lambda^*,$$

we replace $\alpha_{\mathbf{u}}(\mathbf{u}^*-\mathbf{U})$ and $\alpha_p(p^*-P)$ in (2.6) and using (2.5a)-(2.5b) with the the Gauss-Green theorem [10], we obtain

$$\hat{J}'(\mathbf{f}^*)(\delta\mathbf{f}) = -(\lambda^*,\delta\mathbf{f})_{L^2(\Omega)} + \alpha(\mathbf{f}^*,\delta\mathbf{f})_{L^2(\Omega)} = (\alpha\,\mathbf{f}^*-\lambda^*,\delta\mathbf{f})_{L^2(\Omega)} = 0, \tag{2.7}$$

for all $\delta\mathbf{f}\in\mathbf{L}^2(\Omega)$. Thus, we have obtained the gradient of the reduced cost functional, $\nabla\hat{J}(\mathbf{f}) := \alpha\,\mathbf{f}-\lambda$ and formally $\nabla\hat{J}(\mathbf{f}^*) = 0$ constitutes the necessary and sufficient optimality condition (due to convexity) for the Stokes optimal control problem. □

3. Discretization

In this section, the discretization of the optimality system (2.3a)-(2.3b) by finite difference approximations on staggered grids is illustrated. We provide implementation details and notice the advantageous collocation of the optimization variables.

Consider a sequence of grids $\{\Omega_h\}_{h>0}$ defined by

$$\Omega_h = \{\mathbf{x}\in\mathbb{R}^2 : x_i = i\,h_x,\ y_j = j\,h_y,\ i,j\in\mathbb{Z}\}\cap\Omega.$$

We assume that Ω is a rectangular domain and that the values of h_x and h_y are chosen such that the boundaries of Ω coincide with grid lines. On staggered grids, variables may

be placed on cell edge that can be vertical or horizontal and on cell centers. We denote these sets of grid points with Ω_h^s, $s \in \{ev, eh, c\}$. For example, Ω_h^{ev} means grid points defined on center of cell edge-vertical. Notice that within the same set the grid points are spaced with h_x and h_y intervals in $x-$ and $y-$direction, respectively.

For grid functions u^h and v^h defined on the same set Ω_h^s, we introduce the discrete L^2-scalar product

$$\left(u^h, v^h\right)_{L_h^2(\Omega_h^s)} = h_x h_y \sum_{\mathbf{x}\in\Omega_h^s} u^h(\mathbf{x})\, v^h(\mathbf{x}),$$

with associated norm

$$\|u^h\|_{L_h^2(\Omega_h^s)} = \left(u^h, u^h\right)^{1/2}_{L_h^2(\Omega_h^s)}.$$

We require as well the discrete H^1-product given by

$$\|u^h\|_{H_h^1(\Omega_h^s)} = \left(\|u^h\|^2_{L_h^2(\Omega_h^s)} + \sum_{i=1}^{2} \|\partial_i u^h\|^2_{L_h^2(\Omega_h^s)} \right)^{1/2},$$

where ∂_i denotes the difference quotient in the $\hat{ß}$-direction

$$\partial_i^h u(\mathbf{x}) = \frac{u(\mathbf{x} + \hat{ß}\, h_i) - u(\mathbf{x})}{h_i}.$$

Here u^h is extended by 0 on grid points outside of Ω; see [14]. The spaces $L_h^2(\Omega_h^s)$ and $H_h^1(\Omega_h^s)$ consist of the sets of grid functions u^h defined on Ω_h^s endowed with $\|u^h\|_{L_h^2(\Omega_h^s)}$, respectively $\|u^h\|_{H_h^1(\Omega_h^s)}$, as norm. In the following, we denote with $\mathscr{U}_h$, $\mathscr{V}_h$ and $\mathscr{P}_h$ the space of the grid functions $u^h(\lambda^h)$, $v^h(\mu^h)$ and $p^h(q^h)$, respectively that approximate the state (adjoint) variables.

We first discuss the discretization of the state system. On the staggered grid the variable u is defined on Ω_h^{ev}, v is defined on Ω_h^{eh} and p is defined at cell centers Ω_h^c; see Fig. 1.

$$\begin{aligned} -\Delta^h u^h + \partial_x^h p^h &= f^h \qquad \text{in } \Omega_h^{ev}, \\ -\Delta^h v^h + \partial_y^h p^h &= g^h \qquad \text{in } \Omega_h^{eh}, \\ -(\partial_x^h u^h + \partial_y^h v^h) &= 0 \qquad \text{in } \Omega_h^{c}, \end{aligned}$$

where the discrete approximation Δ^h to Laplace operator is the usual 5-point approximation. However, for a point $\mathbf{x} = (x, y)$ near a boundary, $\Delta^h \mathbf{u}^h(\mathbf{x})$ may involve an exterior value. This value is defined by quadratic extrapolation; see [5].

In the following, we use a lexicographic order starting from the lowest-left corner, and consider a unique set of grid indices (i, j) where $i = 1, \cdots, N_x + 1$ and $j = 1, \cdots, N_y + 1$ that index all grid points including the boundaries. The vertices coordinate are given by $x_i = (i-1)h_x$ and $y_j = (j-1)h_y$. With $u_{i+1/2,j+1/2}$ we mean the discrete counterpart to

$u(x_i + h_x/2, y_j + h_y/2)$. Thus by finite difference approximations, we have the following discretized state system

$$-\left(\frac{u_{i-1,j+1/2} - 2u_{i,j+1/2} + u_{i+1,j+1/2}}{h_x^2} + \frac{u_{i,j-1/2} - 2u_{i,j+1/2} + u_{i,j+3/2}}{h_y^2}\right) + \frac{p_{i+1/2,j+1/2} - p_{i-1/2,j+1/2}}{h_x} = f_{i,j+1/2} \quad \text{on } \Omega_h^{ev}, \tag{3.1a}$$

$$-\left(\frac{v_{i-1/2,j} - 2v_{i+1/2,j} + v_{i+3/2,j}}{h_x^2} + \frac{v_{i+1/2,j-1} - 2v_{i+1/2,j} + v_{i+1/2,j+1}}{h_y^2}\right) + \frac{p_{i+1/2,j+1/2} - p_{i+1/2,j-1/2}}{h_y} = g_{i+1/2,j} \quad \text{on } \Omega_h^{eh}, \tag{3.1b}$$

$$-\left(\frac{u_{i+1,j+1/2} - u_{i,j+1/2}}{h_x} + \frac{v_{i+1/2,j+1} - v_{i+1/2,j}}{h_y}\right) = 0_{i+1/2,j+1/2} \quad \text{on } \Omega_h^{c}, \tag{3.1c}$$

where the first momentum Eq. (3.1a) is centered at all internal cell edge-vertical Ω_h^{ev}, the second momentum Eq. (3.1b) is centered at all internal cell edge-horizontal Ω_h^{eh} and the continuity Eq. (3.1c) is centered at all internal cell centers Ω_h^c.

Next, we note that the optimality conditions

$$\alpha f - \lambda = 0, \quad \alpha g - \mu = 0$$

suggest that the staggered grid for adjoint system should be such that u, λ and f share the same location and thus defined on Ω_h^{ev}; v, μ and g is defined on Ω_h^{eh} and p and q are defined on Ω_h^c. In this way, we require no interpolation, i.e., we implement a direct coupling between state, adjoint and control variables. Therefore, we have the discretized adjoint system given by

$$-\left(\frac{\lambda_{i-1,j+1/2} - 2\lambda_{i,j+1/2} + \lambda_{i+1,j+1/2}}{h_x^2} + \frac{\lambda_{i,j-1/2} - 2\lambda_{i,j+1/2} + \lambda_{i,j+3/2}}{h_y^2}\right) + \frac{q_{i+1/2,j+1/2} - q_{i-1/2,j+1/2}}{h_x} = \alpha_{\mathbf{u}}(U-u)_{i,j+1/2} \quad \text{on } \Omega_h^{ev}, \tag{3.2a}$$

$$-\left(\frac{\mu_{i-1/2,j} - 2\mu_{i+1/2,j} + \mu_{i+3/2,j}}{h_x^2} + \frac{\mu_{i+1/2,j-1} - 2\mu_{i+1/2,j} + \mu_{i+1/2,j+1}}{h_y^2}\right) + \frac{q_{i+1/2,j+1/2} - q_{i+1/2,j-1/2}}{h_y} = \alpha_{\mathbf{u}}(V-v)_{i+1/2,j} \quad \text{on } \Omega_h^{eh}, \tag{3.2b}$$

$$-\left(\frac{\lambda_{i+1,j+1/2} - \lambda_{i,j+1/2}}{h_x} + \frac{\mu_{i+1/2,j+1} - \mu_{i+1/2,j}}{h_y}\right) = \alpha_p(P-p)_{i+1/2,j+1/2} \quad \text{on } \Omega_h^{c}. \tag{3.2c}$$

Next, the optimality condition and boundary conditions after discretization gives

$$\alpha f_{i,j+1/2} - \lambda_{i,j+1/2} = 0 \quad \text{on } \Omega_h^{ev}, \tag{3.3a}$$
$$\alpha g_{i+1/2,j} - \mu_{i+1/2,j} = 0 \quad \text{on } \Omega_h^{eh} \tag{3.3b}$$

and

$$u_{i,j+1/2} = 0 \quad \text{for} \quad i = 1, N_x+1, \;\; j = 1, \cdots, N_y, \tag{3.4a}$$
$$v_{i+1/2,j} = 0 \quad \text{for} \quad j = 1, N_y+1, \;\; i = 1, \cdots, N_x, \tag{3.4b}$$
$$\lambda_{i,j+1/2} = 0 \quad \text{for} \quad i = 1, N_x+1, \;\; j = 1, \cdots, N_y, \tag{3.4c}$$
$$\mu_{i+1/2,j} = 0 \quad \text{for} \quad j = 1, N_y+1, \;\; i = 1, \cdots, N_x. \tag{3.4d}$$

Summarizing, the discretized state (3.1a)-(3.1c) and adjoint system (3.2a)-(3.2c) with the optimality conditions (3.3) and boundary conditions (3.4) constitute the discrete optimality system.

4. A full multigrid framework

In this section, we present a full multigrid scheme to solve the discrete optimality system (3.1a)-(3.4), which faces some difficulties due to the nature of staggered grids and of coupled state and adjoint systems.

Here, we anticipate the fact that a multigrid scheme uses different discretization grids. On such grids a refinement is obtained by refining a starting coarse grid, e.g., halving the mesh size [21] that results in a non-nested hierarchy of grids. This fact requires additional effort in the construction of the required inter-grid transfer operators; see [5, 18]. To overcome this, we notice that a nested sequence of grids is obtained by tripling the mesh size. This remark seems novel in the staggered-grid and it has important consequences in the development of multigrid staggered grid schemes and thus to the proposed tracking-type control problems, i.e., distributed optimal control problems with cost functional of velocity and/or pressure tacking-type. A coarsening strategy by a factor-of-three for cell-centered discretization to second-order PDEs in given by [8] and for first-order elliptic optimal control problems in [7], respectively.

Next, we define a sequence of nested grids (levels) Ω_k of mesh size

$$h_{xk} = h_{x1}/3^{(k-1)}, \qquad h_{y_k} = h_{y_1}/3^{(k-1)}, \qquad k = 1, \cdots, L,$$

where $k = L$ is the finest level and h_{x1} and h_{y_1} are the mesh sizes of the coarsest grid in the x- and y- direction, respectively. In the numerical experiments we choose

$$h_{x1} = h_{y_1} = \frac{1}{2}.$$

We denote all operators and functions defined on Ω_k in terms of the index k. Notice that with this setting a variable X_{IJ}^{k-1} at the grid point (I,J) of the coarse grid Ω_{k-1} has the

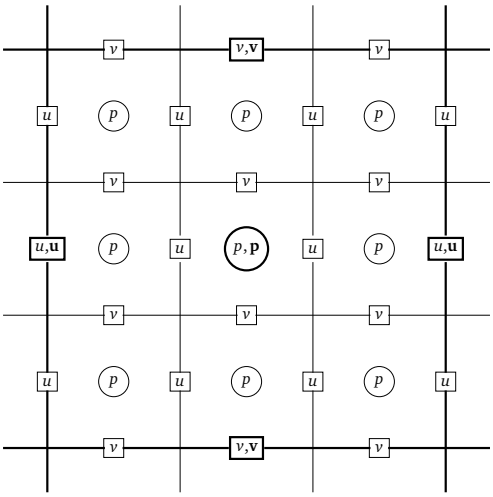

Figure 1: Illustration of the coarsening by a factor-of-three on a single mesh, where the bold-lines represent the coarse grid and fine-lines depict the fine grid, respectively.

same spatial placement as the variable X_{ij}^k at the grid point (i,j) of the fine grid Ω_k as follows; see Fig. 1.

$$
\begin{aligned}
&u_{I,J+1/2}^{k-1} \longleftrightarrow u_{ij}^k \ \ (i=3I-2, j=3J-1),\\
&v_{I+1/2,J}^{k-1} \longleftrightarrow v_{ij}^k \ \ (i=3I-1, j=3J-2),\\
&p_{I+1/2,J+1/2}^{k-1} \longleftrightarrow p_{ij}^k \ (i=3I-1, j=3J-1),\\
&\lambda_{I,J+1/2}^{k-1} \longleftrightarrow \lambda_{ij}^k \ \ (i=3I-2, j=3J-1),\\
&\mu_{I+1/2,J}^{k-1} \longleftrightarrow \mu_{ij}^k \ \ (i=3I-1, j=3J-2),\\
&q_{I+1/2,J+1/2}^{k-1} \longleftrightarrow q_{ij}^k \ \ (i=3I-1, j=3J-1).
\end{aligned}
$$

4.1. A smoothing scheme and intergrid transfer operators

Here, we illustrate a smoothing scheme called *distributive relaxation* [5] to solve the state and adjoint systems and updates the control variables through a gradient step. Furthermore, intergrid transfer operators are also presented in this section.

Let $(u^h, v^h, p^h, \lambda^h, \mu^h, q^h, f^h, g^h)$ be the current approximation to the numerical solution to the optimality system (3.1a)-(3.4). We define an update to this approximation by a sequence of iterative steps.

In the start, we update the control variables by performing a gradient update step as follows

$$f^h := f^h - t\,\nabla_f \hat{J}(f^h, g^h), \tag{4.1a}$$

$$g^h := g^h - t\,\nabla_g \hat{J}(f^h, g^h), \tag{4.1b}$$

where

$$\nabla_f \hat{J}(f^h, g^h) = \alpha f^h - \lambda^h, \qquad \nabla_g \hat{J}(f^h, g^h) = \alpha g^h - \mu^h$$

are the gradients in the control spaces, and $t \in (0, 1]$ is the step length.

Next, we implement the iterative step to the state system by *distributive relaxation* scheme [5, 6] as follows. Let (u^h, v^h, p^h) be the current approximation to the discretized state system (3.1a)-(3.1c). The state momentum Eqs. (3.1a)-(3.1b) are elliptic, therefore for each fix equation, we relax (by point-wise Gauss-Seidel) all the interior points where u^h and v^h is defined. Now, we have to relax the state continuity Eq. (3.1c) to make one complete iterative step for the state system. For this, let $(i + \frac{1}{2}, j + \frac{1}{2})$ be the current cell center and let

$$r_p^h = 0 - (\partial_x^h u^h + \partial_y^h v^h) \tag{4.2}$$

be the *prior residual* at the current cell center, i.e., the residual at cell center just before relaxing there. Then the relaxation step for state continuity equation (3.1c) at current cell center is made up of the following nine changes

$$\begin{aligned}
&u_{i+1/2,j} \leftarrow u_{i+1/2,j} + \delta_p/h_x, && v_{i,j+1/2} \leftarrow v_{i,j+1/2} + \delta_p/h_y,\\
&u_{i-1/2,j} \leftarrow u_{i-1/2,j} - \delta_p/h_x, && v_{i,j-1/2} \leftarrow v_{i,j-1/2} - \delta_p/h_y,\\
&p_{i,j} \leftarrow p_{i,j} + 2(\delta_p/h_x^2 + \delta_p/h_y^2), && p_{i+1,j} \leftarrow p_{i+1,j} - \delta_p/h_x^2,\\
&p_{i,j+1} \leftarrow p_{i,j+1} - \delta_p/h_y^2, && p_{i-1,j} \leftarrow p_{i-1,j} - \delta_p/h_x^2,\\
&p_{i,j-1} \leftarrow p_{i,j-1} - \delta_p/h_y^2,
\end{aligned}$$

where

$$\delta_p = \frac{r_p^h}{2(1/h_x^2 + 1/h_y^2)}. \tag{4.3}$$

Above changes and δ_p are such that, after changing, r_p^h vanishes. The pressure changes are such that the momentum equations residuals

$$r_u^h = f^h + \Delta^h u^h - \partial_x^h p^h, \tag{4.4a}$$

$$r_v^h = g^h + \Delta^h v^h - \partial_y^h p^h \tag{4.4b}$$

at all points remain unchanged, in the approximation sense. We need to modify δ_p because of boundary conditions as

$$\delta_p = \frac{r_p^h}{2(1/h_x^2 + 1/h_y^2) - d}, \tag{4.5}$$

where $d \in \{1/h_x^2, 1/h_y^2, 1/h_x^2 + 1/h_y^2\}$, i.e., $d = 1/h_x^2$ whenever one of the u updates is not performed and $d = 1/h_y^2$ if one of the v updates is not performed. We have $d = 1/h_x^2 + 1/h_y^2$ in the case when both u and v are not updated at the boundary (corner). This completes one iterative step to the state system.

Next, let (λ^h, μ^h, q^h) be the current approximation to the discretized adjoint system (3.2a)-(3.2c). We relax (by point-wise Gauss-Seidel) the residuals of Stokes momentum

Eqs. (3.2a)-(3.2b) at all the interior points where λ^h and μ^h is defined. After completing this, we relax, q^h, the adjoint continuity Eq. (3.2c) by *distributive relaxation* as described in the state system relaxation step. This completes one iterative (relaxation) step for the optimality system.

In the following, we discuss the prolongation (interpolation) and restriction operator needed for the multigrid scheme. A coarsening by factor of three gives a nested sequence of staggered grids and the implementation of bilinear interpolation becomes easier. Therefore we use bilinear interpolation, e.g, consider the space $\mathscr{U}_k$ of $u^k : \Omega_k^{ev} \to \mathbb{R}$, $k = 1, \cdots, L$. Among two grids Ω_k and Ω_{k-1}, we define a prolongation operator, $I_{k-1}^k : \mathscr{U}_{k-1} \to \mathscr{U}_k$, that is consistent with the assumption of bilinear finite elements on each rectangular partition of the discretization.

Note that in coarsening by factor-of-three on staggered grids, the coarse-grid points spatial location is the same as for the fine-grid points; see Fig. 1. Therefore, we use only straight injection operator $I_k^{k-1} : \mathscr{U}_k \to \mathscr{U}_{k-1}$ for transfer of residuals and solution functions from fine to coarse grid.

4.2. The multigrid algorithms

In this section, we discuss the full approximation scheme (FAS) together with the proposed full multigrid method (FMG). The FAS is a nonlinear multigrid method that allows to accommodate non-linearities like the inequalities representing the optimality conditions. The full multigrid method allows to improve the computational complexity of the FAS. The full multigrid method is obtained combining a nested iteration strategy with the FAS that we discuss next.

Consider the optimality system (3.1a)-(3.4) at the discretization level k for the unknown variables $X_k = (u_k, v_k, p_k, \lambda_k, \mu_k, q_k, f_k, g_k)$. We write this system with a compact notation as follows

$$A_k(X_k) = F_k. \tag{4.6}$$

Denote with $X_k^{(l)} = S_k(X_k^{(l-1)}, F_k)$, the result of our smoothing scheme given in Sec. 4.1. Suppose to apply m_1-times this iteration to (4.6) starting with the current approximation $X_k^{(0)}$ to obtain the approximate solution $\tilde{X}_k = X_k^{(m_1)}$.

It is clear that the desired correction e_k to $\tilde{X}_k$ is defined by $A_k(\tilde{X}_k + e_k) = F_k$. This correction can be defined as the solution to

$$A_k(\tilde{X}_k + e_k) - A_k(\tilde{X}_k) = r_k, \tag{4.7}$$

where $r_k = F_k - A_k(\tilde{X}_k)$ is the residual associated to $\tilde{X}_k$.

Next, assume to represent the problem (4.7) on the coarser grid Ω_{k-1}. To represent $\tilde{X}_k + e_k$ on this coarse grid we write

$$X_{k-1} := I_k^{k-1}\tilde{X}_k + e_{k-1}. \tag{4.8}$$

Since $I_k^{k-1}\tilde{X}_k$ and $\tilde{X}_k$ should represent the same function but on different grids. We can think of representing e_k by a coarse function e_{k-1} because e_k is smooth due to the action of S_k.

Now to formulate (4.7) on the coarse grid replace $A_k(\cdot)$ by $A_{k-1}(\cdot)$, $\tilde{X}_k$ by $I_k^{k-1}\tilde{X}_k$, and r_k by restriction $I_k^{k-1}r_k = I_k^{k-1}\Big(F_k - A_k(\tilde{X}_k)\Big)$. We get the following (FAS) equation

$$A_{k-1}(X_{k-1}) = I_k^{k-1}\Big(F_k - A_k(\tilde{X}_k)\Big) + A_{k-1}(I_k^{k-1}\tilde{X}_k). \tag{4.9}$$

This equation is also written in the form $A_{k-1}(X_{k-1}) = I_k^{k-1}F_k + \tau_k^{k-1}$ where

$$\tau_k^{k-1} = A_{k-1}\Big(I_k^{k-1}\tilde{X}_k\Big) - I_k^{k-1}A_k(\tilde{X}_k).$$

The term τ_k^{k-1} is the fine-to-coarse defect or residual correction such that at convergence the solution to (4.9) coincides with the fine grid solution in the sense that $X_{k-1} = I_k^{k-1}X_k$. With X_{k-1} obtained solving (4.9) and from (4.8) we have

$$e_{k-1} = X_{k-1} - I_k^{k-1}\tilde{X}_k.$$

Therefore we can obtain a correction to the fine-grid approximation as follows

$$X_k = \tilde{X}_k + I_{k-1}^k\Big(X_{k-1} - I_k^{k-1}\tilde{X}_k\Big), \tag{4.10}$$

where I_{k-1}^k is the coarse-to-fine interpolation operator. To damp possible high-frequency errors arising through the entire coarse-grid correction process, the correction step (4.10) is followed by m_2-times smoothing iteration.

The full approximation scheme (FAS) cycle is summarized in the following algorithm.

Notice that we can perform m two-grid iterations at each working level. For $m = 1$ we have a V-cycle and for $m = 2$ we have a W-cycle; m is called the cycle index [21].

Next, the full mutigrid algorithm is explained as follows. Suppose to start the solution process from a coarse working level $K < M$ where the discretized problem $A_l(X_l) = F_l$ with $l = K$ is easily solved. The idea is to interpolate this solution to the next finer working level as initial approximation for the iterative process to solve $A_{l+1}(X_{l+1}) = F_{l+1}$ as $X_{l+1} = I_l^{l+1}X_l$. Thereafter the FAS algorithm at level $l+1$ is applied. The idea of using a coarse grid approximation as a first guess for the solution process on a finer grid is known as nested iteration. The algorithm obtained by combining the multigrid scheme with nested iteration is called full multigrid (FMG) method. Because of the improvement on the initial solution at each starting level, the FMG scheme results to be more efficient than the iterative application of the multigrid cycle without FMG initialization. An algorithm describing the FMG method is given in the following algorithm.

5. Numerical experiments

In this section, we present some numerical results that illustrate the application of the proposed multigrid scheme to distributed optimal control problems (2.1)-(2.2c). We consider a rectangular domain $\Omega = (0,1)^2$, a family of uniform rectangular grids with mesh

Algorithm 4.1 FAS(m_1, m_2) for solving $A_k(X_k) = F_k$.

1. If $k = 1$ solve $A_k(X_k) = F_k$ exactly.
2. Pre-smoothing steps on the fine grid:
$$X_k^{(l)} = S_k(X_k^{(l-1)}, F_k), \ \ l = 1, \cdots, m_1.$$
3. Computation of the residual:
$$r_k = F_k - A_k(X_k^{(m_1)}).$$
4. Restriction of the residual:
$$r_{k-1} = I_k^{k-1} r_k.$$
5. Set $X_{k-1} = I_k^{k-1} X_k^{(m_1)}$.
6. Set $F_{k-1} = r_{k-1} + A_{k-1}(X_{k-1})$.
7. Call m times MG(m_1, m_2) to solve:
$$A_{k-1}(X_{k-1}) = F_{k-1}.$$
8. Coarse-grid correction:
$$X_k^{(m_1+1)} = X_k^{(m_1)} + I_{k-1}^k \left(X_{k-1} - I_k^{k-1} X_k^{(m_1)} \right).$$
9. Post-smoothing steps on the fine grid:
$$X_k^{(l)} = S_k(X_k^{(l-1)}, F_k), \quad l = m_1 + 2, \cdots, m_1 + m_2 + 1.$$

size h, and run our experiments on a laptop i7, 1.86GHz, 4GB RAM, using Matlab 8.5.0 (R2015a).

In the following, we consider the tracking-type control problem (2.1)-(2.2c) and take the following target velocity [9]

$$U(x, y) := -2x^2 y(1-x)^2(1-3y+2y^2),$$
$$V(x, y) := 2xy^2(1-y)^2(1-3x+2x^2)$$

and the target pressure

$$P(x, y) := cos(\pi x) cos(\pi y).$$

To solve this problem, we use our full multigrid scheme to the optimality system (3.1a)-(3.4). We employ W-cycles with $m_1 = m_2 = 2$, i.e., 2-pre and 2-post smoothing steps.

Algorithm 4.2 FMG for solving $A_L(X_L) = F_L$.

1. For $l = K < L$ set initial approximation u_l.
2. If $l < L$ then interpolate to the next finer working level:
$$\tilde{X}_{l+1} = I_l^{l+1} X_l.$$
3. Apply FAS to solve $A_{l+1}(X_{l+1}) = F_{l+1}$, starting with $\tilde{X}_{l+1}$.
4. Set $l := l + 1$. If $l < L$ go to step 2, else stop.

Further, for the control variable update step, we find advantages by choosing $t = 1$ in the numerical experiments. In the FMG scheme, we implement a stopping criteria with a given tolerance on the discrete L^2-norm of the residuals as

$$\max\left\{\|r_i^h\|_{L^2} : i \in \{u, v, p, \lambda, \mu, q\}\right\} < tol = 10^{-6}.$$

In Table 1, we report maximum number of iterations (W-cycles) required to reach the given tol on finest level. This shows the computational performance of the proposed FMG scheme and the tracking ability to velocity control only ($\alpha_{\mathbf{u}} = 1, \alpha_p = 0$) and to mixed velocity-pressure control ($\alpha_{\mathbf{u}} = 1, \alpha_p \neq 0$). In Table 1, it is evident that our proposed multigrid scheme gives good convergence results to the proposed control problem formulation for small values of optimization parameter α, and as expected number of iterations increases as $\alpha \downarrow 0$ due to the fact that very small values of α correspond to more stiff optimality system. Moreover, a decrease in the number of iterations with mesh size can be observed in Table 1.

The choice of values for α_p in mixed velocity-pressure control is justified by the data in Table 2, where we report relative tracking errors for velocity

$$E_{\mathbf{u}} = \|\mathbf{u}^* - \mathbf{U}\|_{\mathbf{L}^2} / \|\mathbf{U}\|_{\mathbf{L}^2}$$

and pressure

$$E_p = \|p^* - P\|_{L^2} / \|P\|_{L^2}$$

on 54×54 mesh for velocity control only, and mixed velocity-pressure control. From Table 2, it is evident that for velocity control only, the pressure is not recovered at all (i.e., $E_p \approx 1$). For the mixed control both velocity and pressure are being recovered, i.e., both $E_{\mathbf{u}}$, E_p decrease with $\alpha \downarrow 0$. Furthermore, for $\alpha_p = 10^{-5}$ the relative velocity tracking error $E_{\mathbf{u}}$ is smaller than E_p, i.e., velocity is recovered better than pressure, whereas for $\alpha_p = 10^{-3}$ the situation is reversed, i.e., the relative pressure error E_p is better recovered than $E_{\mathbf{u}}$.

It is important to remark that the dependence of our proposed full multigrid method to optimization (or regularisation) parameter α, see Table 1, can be improved by *line search method* (or by changing the choice $t = 1$) during update step of the control variables in (4.1a)-(4.1b), e.g., see [6]. On the other hand, preconditioning iterative methods can be used to improve this phenomena.

Table 1: Iteration history for velocity control only and mixed (velocity-pressure) control.

$N_x \times N_y$	6×6	18×18	54×54
$\alpha_{\mathbf{u}} = 1, \alpha_p = 0$			
$\alpha = 1$	4	2	2
$\alpha = 10^{-1}$	9	5	3
$\alpha = 10^{-2}$	99	46	22
$\alpha = 10^{-3}$	933	468	188
$\alpha = 10^{-4}$	6239	3930	863
$\alpha = 10^{-5}$	37511	5423	6458
$\alpha = 10^{-7}$	20248	1171	440
$\alpha_{\mathbf{u}} = 1, \alpha_p = 10^{-5}$			
$\alpha = 1$	4	2	2
$\alpha = 10^{-1}$	9	5	3
$\alpha = 10^{-2}$	99	46	22
$\alpha = 10^{-3}$	932	468	187
$\alpha = 10^{-5}$	31875	5415	4601
$\alpha = 10^{-7}$	319758	14256	2684
$\alpha_{\mathbf{u}} = 1, \alpha_p = 10^{-3}$			
$\alpha = 1$	6	4	4
$\alpha = 10^{-1}$	8	6	4
$\alpha = 10^{-2}$	87	51	28
$\alpha = 10^{-3}$	814	466	249
$\alpha = 10^{-5}$	11446	19244	11527
$\alpha = 10^{-7}$	16565	5332	1978

Table 2: Relative tracking errors for velocity $E_{\mathbf{u}} = \|\mathbf{u}^* - \mathbf{U}\|_{\mathbf{L}^2} / \|\mathbf{U}\|_{\mathbf{L}^2}$ and pressure $E_p = \|p^* - P\|_{L^2} / \|P\|_{L^2}$.

$\alpha_{\mathbf{u}}$	α_p	$E_{\mathbf{u}}$	E_p	$E_{\mathbf{u}}$	E_p	$E_{\mathbf{u}}$	E_p
α		10^{-3}		10^{-5}		10^{-7}	
1	0	$7.3592e-1$	≈ 1	$4.0563e-2$	≈ 1	$6.0637e-3$	$8.1557e-1$
1	10^{-5}	$7.3585e-1$	$9.9941e-1$	$4.1252e-2$	$9.4753e-1$	$5.4686e-3$	$1.7919e-1$
1	10^{-3}	$7.2719e-1$	$9.4138e-1$	$1.2297e-1$	$1.4859e-1$	$7.3012e-3$	$9.7044e-3$

6. Conclusions

In this article, a full multigrid method for distributed optimal control problem with cost functional of tracking-type, constrained by Stokes system with Dirichlet boundary conditions, was investigated. We discretize the optimality system using finite differences on staggered grids and a coarsening by a factor-of-three is used. A distributive-Gauss-Seidel smoothing scheme with gradient update step was employed. Results of numerical experiments demonstrate the effectiveness and efficiency of proposed multigrid staggered

grid framework to tracking-type control problems.

The proposed multigrid scheme can also be extended to constrained control problems, e.g., see [6]. This work also provides a first step towards the solution of the tracking-type problems constrained by Navier-Stokes system using the proposed multigrid staggerd grid procedure, which is the focus of our current research.

Acknowledgments This work is supported by CAS President's International Fellowship Initiative (PIFI). The authors are thankful to the anonymous referees for their careful reading and valuable remarks.

References

[1] C. Bacuta, P. Vassilevski and S. Zhang, *A new approach for solving Stokes systems arising from a distributive relaxation method*, Numer. Methods Partial Differ. Equ., 27 (2011), pp. 898–914.

[2] M. Benzi, G. H., Golub and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.

[3] A. Borzì and V. Schulz, *Multigrid methods for PDE optimization*, SIAM Review, 51 (2009), pp. 361–395.

[4] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. 31 (1977), pp. 333–390.

[5] A. Brandt and N. Dinar, *Multigrid solutions to elliptic flow problems*, in S. V. Parter (ed.), Numerical Methods for PDEs, Academic Press, New York, 1979, pp. 53–147.

[6] M. M. Butt, *A multigrid solver for stokes control problems*, Int. J. Comput. Math.

[7] M. M. Butt and A. Borzì, *Formulation and multigrid solution of Cauchy-Riemann optimal control problems*, Computing and Visualization in Science, 14 (2011), pp. 79–90.

[8] J. E. Dendy Jr and J. D. Moulton, *Black Box Multigrid with coarsening by a factor of three*, Numer. Linear Algebra Appl., 17 (2010), pp. 577–598.

[9] A. Drgnescu and A. M. Soane, *Multigrid solution of a distributed optimal control problem constrained by the Stokes equations*, Appl. Math. Comput., 219 (2013), pp. 5622–5634.

[10] L.C. Evans, *Partial Differential Equations, Graduate Studies in Mathematics*, American Mathematical Society, 128(1-2) (2002), pp. 55–82.

[11] V. Girault and P. A. Raviart, Finite Element Methods for NavierâĂŞStokes Equations: Theory and Algorithms, Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 5 (1986).

[12] R. B. Kellogg and J. E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, J. Funct. Anal., 21 (1976), pp. 397–431.

[13] W. Hackbusch, Multi-grid Methods and Applications, Springer-Verlag, 960(4) (1985), pp. 558–575.

[14] W. Hackbusch, Elliptic Differential Equations, Springer, New York, 1992.

[15] K. Ito and K. Kunisch, Lagrange Multiplier Approach to Variational Problems and Applications, SIAM, 2008.

[16] M. Kollmann and W. Zulehner, *A Robust Preconditioner for Distributed Optimal Control for Stokes Flow with Control Constraints*, In Andrea Cangiani, Ruslan L. Davidchack, Numerical Mathematics and Advanced Applications 2011, Springer Berlin Heidelberg, (2013), pp. 771–779.

[17] J. L. LIONS, Optimal Control of Systems Governed by Partial Differential Equations, Springer, Berlin, 1971.
[18] C. W. OOSTERLEE AND F. J. GASPAR, *Multigrid methods for the stokes system*, Comput. Sci. Eng., 8 (2006), pp. 34–43.
[19] S. TAKACS, *A robust all-at-once multigrid method for the Stokes control problem*, Numerische Mathematik, 130 (2015), pp. 517–540
[20] F. TRÖLTZSCH, *Optimal control of partial differential equations: theory, methods and applications*, AMS, 2010.
[21] U. TROTTENBERG, C. OOSTERLEE AND A. SCHÜLLER, Multigrid, Academic Press, London, 2001.
[22] M. WANG AND L. CHEN, *Multigrid methods for the stokes equations using distributive Gauss–seidel relaxations based on the least squares commutator*, J. Sci. Comput. 56 (2013), pp. 409–431.
[23] G. WITTUM, *Multi-grid methods for Stokes and Navier-Stokes equations*, Numerische Mathematik, 54 (1989), pp. 543–563.