Formulation and multigrid solution of Cauchy-Riemann optimal control problems

M. M. Butt[*]   A. Borzì[†]

Abstract

The formulation of optimal control problems governed by Cauchy-Riemann equations is presented. A distributed control mechanism through divergence and curl sources is considered with the boundary conditions of mixed type. A Lagrange multiplier framework is introduced to characterize the solution to Cauchy-Riemann optimal control problems as the solution of an optimality system of four first-order partial differential equations and two optimality conditions.

To solve the optimality system, staggered grids and multigrid methods are investigated. It results that staggered grids provide a natural collocation of the optimization variables and second-order accurate solutions are obtained. The proposed multigrid scheme is based on a coarsening by a factor of three that results in a nested hierarchy of staggered grids. On these grids a distributed-Gauss-Seidel and gradient-based smoothing scheme is employed.

Results of numerical experiments validate the proposed optimal control formulation and demonstrate the effectiveness of the staggered-grids multigrid solution procedure.

Key words: Cauchy-Riemann equations, optimal control problems, staggered grids, multigrid methods

AMS Classification: 35Q93, 49K20, 65N55.

1   Introduction

Control of systems governed by partial differential equations (PDE) is becoming a very important field of research in applied mathematics with industrial applications. The purpose of control of PDE models is manifold and it addresses the need to define ways of how to optimally change and influence

---

[*]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria (`muhammad.butt@edu.uni-graz.at`) and Department of Mathematics, GC University, Katchery Road, Lahore 54000, Pakistan.

[†]Institut für Mathematik, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany (`alfio.borzi@mathematik.uni-wuerzburg.de`).

application systems to meet a desired target. The resulting problems require to realize large-scale optimization strategies with increasing complexity and this fact motivates the development of fast iterative schemes for optimization purposes.

Within this research framework, first-order PDE models represent a much less investigated subject despite of the fact that they constitute important application models as Euler equations and Maxwell equations. We focus on Cauchy-Riemann (CR) equations that are representative of elliptic first-order PDE systems [2] and are a prototype model of all div-curl systems. In applications, our motivation to study CR equations is the fact that they govern the velocity field of subsonic compressible inviscid flows [19, 22].

We formulate CR control problems based on the optimization framework provided by the infinite-dimensional PDE optimal control theory [16, 17, 20]. In this framework, we consider a governing Cauchy-Riemann system with mixed Dirichlet and Neumann boundary conditions. We choose a control mechanism of distributed type through div- and curl-source terms and a tracking criterion defining the objective of the control and including the cost of its action. Our optimal control problem is then formulated as the minimization of the tracking objective under the constraint given by the CR equations.

While optimization with second-order partial differential equations has received much attention [16, 20], much less is known on optimization problems of the type considered in this paper; see [12] for a Galerkin-least-squares method for solving first-order elliptic control problems. Our purpose is to discuss the formulation of CR optimization problems and to present a multigrid strategy to solve the corresponding optimality systems.

In the next two sections, we introduce the functional setting and formulate a div-curl problem in two-dimensional domains. We discuss well-posedness of the model problem illustrating existence and uniqueness of solutions to the CR equations. In Section 4, we formulate Cauchy-Riemann optimal control problems and prove existence and uniqueness of optimal solutions. In particular, we discuss the characterization of these solutions as solutions to optimality systems consisting of four first-order PDE equations and two optimality conditions. In Section 5, we present a second-order approximation to the CR optimality system by staggered grids on non-uniform meshes. In our approach, a natural collocation of the optimization variables is obtained thus avoiding any interpolation of variables in the formulation of the discretized CR optimality system. In Section 6, we illustrate our multigrid approach for solving the optimality system. First, we notice that a coarsening by a factor of three is advantageous to solve the optimality system on a staggered grid. In fact, in contrast to the standard approach of using a factor of two also in combination with staggered grids, in our scheme a nested hierarchy of staggered grids is obtained that results in more accurate and easy to implement intergrid transfer operators. Further, we focus on the development of distributed Gauss-Seidel relaxation schemes where the control functions are updated using a gradient-based approach. The resulting smoothing scheme is embedded in

2

a non-linear multigrid procedure which can also accommodate possible non-linearities and constraints on the controls. In Section 7, we report results of numerical experiments to validate our CR optimal control formulation and the staggered-grid multigrid solution procedure. These results demonstrated second-order accuracy of the optimal solutions and the ability of the multigrid scheme to solve the optimality system for different optimization parameters with mesh-independent convergence factors. A section of conclusion completes this work.

2   Model Problem

We consider a computational domain Ω in $\mathbb{R}^2$ with (x, y) coordinates and piecewise smooth boundary and $\mathbf{n} = (n_x, n_y)$ is the unit outward normal vector to the boundary. Referring to [1], the function space $H^s(\Omega)$ denotes the Sobolev space of order $s \geq 0$ of real valued functions defined on Ω. The space $H^s(\Omega)$ is equipped with the Sobolev norm $\| \cdot \|_s$ and associated seminorm $| \cdot |_s$. In particular, $s = 0$ signifies the $L^2(\Omega)$ space and corresponding norm induced by the $L^2(\Omega)$ inner product $(\cdot, \cdot)$. A similar notation is used to denote norms in the product spaces $\mathbf{H}^s(\Omega) = H^s(\Omega) \times H^s(\Omega)$. The inner product in $\mathbf{L}^2(\Omega) = L^2(\Omega) \times L^2(\Omega)$ is also denoted by $(\cdot, \cdot)$.

Let $f_1, f_2 \in L^2(\Omega)$, and $P(f_1, f_2)$ denotes the following div-curl system in two-dimensions

$$\operatorname{div} \mathbf{w} = f_1 \quad \text{in} \quad \Omega \tag{1}$$

$$\operatorname{curl} \mathbf{w} = f_2 \quad \text{in} \quad \Omega, \tag{2}$$

where $\mathbf{w} = (u, v)$ are the state variables, $\mathbf{f} = (f_1, f_2)$ are the divergence-source and curl-source terms, and we have

$$\operatorname{div} \mathbf{w} := \partial_x u + \partial_y v \qquad \text{and} \qquad \operatorname{curl} \mathbf{w} := \partial_x v - \partial_y u,$$

subject to the following boundary conditions

$$\mathbf{w} \times \mathbf{n} = 0 \quad \text{on} \quad \Gamma_D, \tag{3}$$

$$\mathbf{w} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_N = \partial\Omega \setminus \overline{\Gamma}_D. \tag{4}$$

That is, $u\,n_y - v\,n_x = 0$ on Γ_D and $u\,n_x + v\,n_y = 0$ on Γ_N. Here, Γ_D and Γ_N denote Dirichlet and Neumann boundaries, respectively.

In the following, it is also convenient to consider a general three-dimensional setting. In this case, we write (1) and (2) in a differential vector notation as follows

$$\nabla \cdot \mathbf{w} = f_1 \quad \text{in} \quad \Omega \tag{5}$$

$$\nabla \times \mathbf{w} + \nabla \phi = \mathbf{f}_2 \quad \text{in} \quad \Omega \tag{6}$$

where $\mathbf{w} = (u, v, t)$ and ϕ is a scalar function that must be added to have a well-defined set of four first-order partial differential equations [12]. Additional

conditions of ϕ are required that depend on the boundary condition settings. We postpone this discussion to the section on the optimal control formulation where we show that ϕ is actually a zero function.

We consider boundary conditions of mixed type as follows

$$\mathbf{w} \times \mathbf{n} = 0 \quad \text{on} \quad \Gamma_D, \tag{7}$$
$$\mathbf{w} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_N = \partial\Omega \setminus \overline{\Gamma}_D, \tag{8}$$

where $\mathbf{n} = (n_x, n_y, n_z)$. Notice that the three-dimensional setting reduces to the two-dimensional case assuming that all variables do not depend on the third z-coordinate in $\Omega \subset \mathbb{R}^3$ and ϕ does not appear.

For more insight in the boundary condition formulation, consider a square domain $\Omega = (0, 1) \times (0, 1)$, where $\Gamma_D = (0, 1) \times \{0\}$ and $\Gamma_N = \partial\Omega \setminus \overline{\Gamma}_D$. This corresponds to the following

$$
\begin{aligned}
u &= 0 \quad \text{on} \quad \Gamma_1 = (0, 1) \times \{0\} \\
u &= 0 \quad \text{on} \quad \Gamma_2 = \{0\} \times (0, 1) \\
u &= 0 \quad \text{on} \quad \Gamma_3 = \{1\} \times (0, 1) \\
v &= 0 \quad \text{on} \quad \Gamma_4 = (0, 1) \times \{1\}
\end{aligned}
$$

Interpreting $\mathbf{w} = (u, v)$ as a velocity vector field in Ω, we can see that the Dirichlet boundary conditions correspond to the assignment of the tangential velocity on the boundary, while on the Neumann boundary the normal component of the velocity is assigned.

It is not a restriction to consider homogeneous boundary conditions since a non homogeneous problem can be converted to an homogeneous one; see [24].

3 Well-posedness of the model problem

To discuss well-posedness of the model problem (1)-(4), we remark that the solution to the $P(f_1, f_2)$ problem can be obtained as the sum of the solution to the $P(f_1, 0)$ and the solution to the $P(0, f_2)$ problems.

Consider the problem $P(f_1, 0)$ and define the 'potential function' φ, such that

$$\partial_x \varphi = u, \quad \partial_y \varphi = v \quad \text{in} \quad \Omega.$$

Based on (1)-(4), we have that φ satisfies the following boundary value problem

$$
\begin{aligned}
\Delta\varphi &= f_1 \quad \text{in} \quad \Omega, \\
\varphi &= c \quad \text{on} \quad \Gamma_D, \\
\frac{\partial\varphi}{\partial n} &= 0 \quad \text{on} \quad \Gamma_N,
\end{aligned}
\tag{9}
$$

where c stands for a real constant. According to [11], a solution φ to this problem exists in $H^2(\Omega)$ and is unique. Therefore $P(f_1, 0)$ has a unique solution $\mathbf{w_1} \in \mathbf{H}^1(\Omega)$.

4

Next, we consider the problem $P(0, f_2)$ and define the 'stream function' ψ by
$$\partial_x \psi = v, \quad \partial_y \psi = -u \quad \text{in} \quad \Omega.$$
From (1)-(4) it follows that ψ satisfies
$$
\begin{aligned}
\Delta \psi &= f_2 \quad \text{in} \quad \Omega, \\
\frac{\partial \psi}{\partial n} &= 0 \quad \text{on} \quad \Gamma_D, \\
\psi &= c \quad \text{on} \quad \Gamma_N,
\end{aligned}
\tag{10}
$$
where c stands for a real constant. It follows that there exist a unique solution to (10) in $H^2(\Omega)$. Therefore $P(0, f_2)$ has a unique solution $\mathbf{w}_2 \in \mathbf{H}^1(\Omega)$ and hence $P(f_1, f_2)$ has a unique solution $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 \in \mathbf{H}^1(\Omega)$.

The setting of the potential and stream function problems show that $P(f_1, f_2)$ is not well-posed in the case of boundary conditions of a unique type. In fact in the case of Neumann boundary conditions, $\mathbf{w} \cdot \mathbf{n} = 0$ on $\partial\Omega$, we have that the problem is well posed only if the following compatibility condition is satisfied
$$\int_\Omega f_1 dx = 0. \tag{11}$$
In the case of Dirichlet boundary conditions, $\mathbf{w} \times \mathbf{n} = 0$ on $\partial\Omega$, we have that the condition
$$\int_\Omega f_2 dx = 0 \tag{12}$$
must be satisfied in order to have a well-defined problem.

For all well-posed CR problems that we have discussed, the following holds
$$\|\mathbf{w}\|_{\mathbf{H}^1(\Omega)} \leq c \|\mathbf{f}\|_{\mathbf{L}^2(\Omega)} \tag{13}$$
where c is a real constant. Moreover, the map
$$\mathbf{L}^2(\Omega) \ni \mathbf{f} \mapsto \mathbf{w}(\mathbf{f}) \in \mathbf{H}^1(\Omega),$$
defined by the unique solution of the div-curl system is continuous and Fréchet differentiable [12].

4   Formulation of Cauchy - Riemann optimal control problems

We consider the following CR optimal control problem
$$\min J(u, v, f, g) := \frac{1}{2}\|\mathbf{w} - \mathbf{W}\|^2_{\mathbf{L}^2(\Omega)} + \frac{\alpha}{2}\|f\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|g\|^2_{L^2(\Omega)}, \tag{14}$$
subject to $(f, g) \in \mathbf{L}^2(\Omega)$ and
$$
\begin{aligned}
\operatorname{div} \mathbf{w} &= f \tag{15} \\
\operatorname{curl} \mathbf{w} &= g \tag{16}
\end{aligned}
$$

and mixed boundary conditions given by

$$
\begin{aligned}
\mathbf{w} \times \mathbf{n} &= 0 \quad \text{on} \quad \Gamma_D, & (17) \\
\mathbf{w} \cdot \mathbf{n} &= 0 \quad \text{on} \quad \Gamma_N = \partial\Omega \setminus \overline{\Gamma}_D. & (18)
\end{aligned}
$$

We denote with $\mathbf{W} = (U, V) \in \mathbf{L}^2(\Omega)$ the target function and $\alpha, \beta > 0$ are the weights of the costs of the controls.

Notice that the objective function J is convex and the constraints are linear. Thus existence of a unique solution to (14)-(18) and its characterization can be shown based on classical techniques [16, 17, 20]. For this purpose, we consider a three-dimensional setting and define the Lagrangian

$$
L(u, v, p, q, \phi, f, g) = J(u, v, f, g) + (\nabla \cdot \mathbf{w} - f, p) + (\nabla \times \mathbf{w} + \nabla\phi - \mathbf{g}, \mathbf{q}). \quad (19)
$$

In this formulation, the unique solution to the optimal control problem satisfies the extrema conditions for the Lagrangian, that are the first-order optimality conditions. Therefore, we have to evaluate the Fréchet derivatives $\nabla_p L$, $\nabla_q L$, $\nabla_u L$, $\nabla_v L$, $\nabla_\phi L$, $\nabla_f L$, and $\nabla_g L$. We easily obtain

$$
\begin{aligned}
(\nabla_p L, \delta p) &= (\nabla \cdot \mathbf{w} - f, \delta p), \\
(\nabla_q L, \delta \mathbf{q}) &= (\nabla \times \mathbf{w} + \nabla\phi - \mathbf{g}, \delta \mathbf{q}), \\
(\nabla_f L, \delta f) &= (\alpha f - p, \delta f), \\
(\nabla_g L, \delta \mathbf{g}) &= (\beta \mathbf{g} - \mathbf{q}, \delta \mathbf{g}).
\end{aligned}
$$

We also have that

$$
(\nabla_\phi L, \delta\phi) = (\nabla\delta\phi, \mathbf{q}) = -\int_\Omega (\nabla \cdot \mathbf{q}) \, \delta\phi \, dx + \int_{\partial\Omega} \delta\phi \, (\mathbf{q} \cdot \mathbf{n}) \, ds. \quad (20)
$$

Next, consider the computation of $(\nabla_\mathbf{w} L, \delta \mathbf{w})$. We have

$$
\begin{aligned}
(\nabla_\mathbf{w} L, \delta\mathbf{w}) &= (\mathbf{w} - \mathbf{W}, \delta\mathbf{w}) + (\nabla \cdot \delta\mathbf{w}, p) + (\nabla \times \delta\mathbf{w}, \mathbf{q}) \\
&= (\mathbf{w} - \mathbf{W}, \delta\mathbf{w}) - (\nabla p, \delta\mathbf{w}) + \int_{\partial\Omega} p \, (\delta\mathbf{w} \cdot \mathbf{n}) \, ds \\
&\quad + (\nabla \times \mathbf{q}, \delta\mathbf{w}) + \int_{\partial\Omega} (\mathbf{q} \times \mathbf{n}) \cdot \delta\mathbf{w} \, ds \quad (21)
\end{aligned}
$$

We are now ready to formulate the optimality conditions. First, we have

$$
\begin{aligned}
\nabla \cdot \mathbf{w} - f &= 0, \\
\nabla \times \mathbf{w} + \nabla\phi - \mathbf{g} &= 0, \\
\alpha f - p &= 0, \\
\beta \mathbf{g} - \mathbf{q} &= 0.
\end{aligned}
$$

Requiring $(\nabla_\phi L, \delta\phi) = 0$ for any $\delta\phi \in H^1(\Omega)$, the following equation must hold

$$
\nabla \cdot \mathbf{q} = 0 \qquad \text{in } \Omega. \quad (22)
$$

6

Further, we require $(\nabla_{\mathbf{w}} L, \delta\mathbf{w}) = 0$ for all $\delta\mathbf{w} \in \mathbf{H}^1(\Omega)$ that satisfy the homogeneous boundary conditions (17)-(18). Because $\delta\mathbf{w} \cdot \mathbf{n} = 0$ on Γ_N, we must have

$$p = 0 \quad \text{on} \quad \Gamma_D. \tag{23}$$

Similarly, because $\delta\mathbf{w} \times \mathbf{n} = 0$ on Γ_D, we obtain

$$\mathbf{q} = 0 \quad \text{on} \quad \Gamma_N. \tag{24}$$

Using this last result in (20), we obtain the following boundary condition, $\delta\phi = 0$ on Γ_D. Moreover, taking the divergence of the curl equation gives $\Delta\phi = 0$ because from (22) we have $\nabla \cdot g = \nabla \cdot \mathbf{q}/\beta = 0$. Therefore we obtain $\phi = 0$ in Ω.

As a result, we have the equation

$$\nabla \times \mathbf{q} - \nabla p + \mathbf{w} - \mathbf{W} = 0 \qquad \text{in } \Omega. \tag{25}$$

Now, let us consider the two-dimensional case. We have that the first two components of $\mathbf{q} = (q_1, q_2, q_3)$ are differentiated with respect to z and therefore they do not appear in (25). However, from (22) we have $\operatorname{div}(q_1, q_2) = 0$, which is satisfied by $q_1 = 0$ and $q_2 = 0$. For this reason and for notational convenience, we take $\mathbf{q} = (0, 0, -q)$. It follows that

$$(\nabla \times \mathbf{q}, \delta\mathbf{w}) - (\nabla \mathbf{p}, \delta\mathbf{w}) = -\int_\Omega \delta u \frac{\partial q}{\partial y}\, dx + \int_\Omega \delta v \frac{\partial q}{\partial x}\, dx - \int_\Omega \delta u \frac{\partial p}{\partial x}\, dx - \int_\Omega \delta v \frac{\partial p}{\partial y}\, dx.$$

In addition, we have $(\mathbf{w} - \mathbf{W}, \delta\mathbf{w}) = (u - U, \delta u) + (v - V, \delta v)$.

Thus we can write the adjoint div-curl system in two dimensions as follows

$$\partial_x p + \partial_y q = (u - U) \quad \text{in} \quad \Omega \tag{26}$$
$$\partial_x q - \partial_y p = -(v - V) \quad \text{in} \quad \Omega \tag{27}$$

with boundary conditions given by

$$\begin{cases} p &=& 0 \quad \text{on} \quad \Gamma_D \\ q &=& 0 \quad \text{on} \quad \Gamma_N \end{cases} \tag{28}$$

Summarizing, the solution of the Cauchy-Riemann optimal control problem is characterized as the solution to the following CR optimality system

$$\begin{array}{rcll}
\partial_x u + \partial_y v &=& f & \text{in} \quad \Omega \\
\partial_x v - \partial_y u &=& g & \text{in} \quad \Omega \\
u\, n_y - v\, n_x &=& 0 & \text{on} \quad \Gamma_D \\
u\, n_x + v\, n_y &=& 0 & \text{on} \quad \Gamma_N \\
\partial_x p + \partial_y q &=& (u - U) & \text{in} \quad \Omega \\
\partial_x q - \partial_y p &=& -(v - V) & \text{in} \quad \Omega \\
p &=& 0 & \text{on} \quad \Gamma_D \\
q &=& 0 & \text{on} \quad \Gamma_N \\
\alpha f - p &=& 0 & \text{in} \quad \Omega \\
\beta g + q &=& 0 & \text{in} \quad \Omega
\end{array}$$

Theorem 1 Assume existence and uniqueness of solution to (15)-(18) and define the gradients $\nabla_f \hat{J}(f, g) := \alpha\, f - p$ and $\nabla_g \hat{J}(f, g) := \beta\, g + q$, where p and q are the solution to the adjoint CR system (26)-(28). Then the control problem (14)-(18) has a unique solution in $\mathbf{L}^2(\Omega)$ if and only if $\nabla_f \hat{J}(f, g) = 0$ and $\nabla_g \hat{J}(f, g) = 0$. Therefore, the optimal solution is characterized as the solution of the first-order optimality system.

Proof. Since the forward Cauchy-Riemann problem (15)-(18) has a unique solution $\mathbf{w} = (u, v)$ for given $\mathbf{f} = (f, g)$, we denote this dependence by $\mathbf{w} = \mathbf{w}(\mathbf{f})$. Therefore, to discuss existence and characterization of the unique solution to (14)-(18), we can introduce the so-called reduced cost functional $\hat{J}$ ([16, 17]) given by

$$\hat{J}(\mathbf{f}) = J(\mathbf{w}(\mathbf{f}), \mathbf{f}), \tag{29}$$

In fact, the optimal solution corresponds to the unique minimizer of $\hat{J}(\mathbf{f})$. Recall that the map $\mathbf{L}^2(\Omega) \ni \mathbf{f} \mapsto \mathbf{w}(\mathbf{f}) \in \mathbf{H}^1(\Omega)$ is affine and continuous and Fréchet differentiable. Let us denote its first derivative at $\mathbf{f}$ in the direction $\delta\mathbf{f} = (\delta f, \delta g)$ by $\mathbf{w}'(\mathbf{f})\, \delta\mathbf{f}$. It is characterized as the solution to

$$\operatorname{div}(\mathbf{w}'(\mathbf{f})\, \delta\mathbf{f}) = \delta f \tag{30}$$
$$\operatorname{curl}(\mathbf{w}'(\mathbf{f})\, \delta\mathbf{f}) = \delta g \tag{31}$$

subject to mixed boundary conditions. The second derivative of $\mathbf{f} \mapsto \mathbf{w}(\mathbf{f})$ is zero, due to the linear dependence of the governing CR model on the controls.

Next, from (14) we find for the second derivative of $\mathbf{f} \to \hat{J}(\mathbf{f})$ the following

$$\hat{J}''(\mathbf{f})(\delta\mathbf{f}, \delta\mathbf{f}) = \|\mathbf{w}'(\mathbf{f})\, \delta\mathbf{f}\|^2_{\mathbf{L}^2(\Omega)} + \alpha \,\|\delta f\|^2_{L^2(\Omega)} + \beta \,\|\delta g\|^2_{L^2(\Omega)},$$

and thus $\mathbf{f} \to \hat{J}(\mathbf{f})$ is uniformly convex. This implies existence of a unique solution $\mathbf{f}^*$ to the optimization problem; we denote this solution with *. Moreover, the minimum is characterized by $\hat{J}'(\mathbf{f}^*; \delta\mathbf{f}) = 0$ for all $\delta\mathbf{f}$ and consequently

$$\hat{J}'(\mathbf{f}^*; \delta\mathbf{f}) = (\mathbf{w}^* - \mathbf{W}, \mathbf{w}'(\mathbf{f}^*)\, \delta\mathbf{f})_{\mathbf{L}^2(\Omega)} + \alpha(f^*, \delta f)_{L^2(\Omega)} + \beta(g^*, \delta g)_{L^2(\Omega)} = 0, \tag{32}$$

where $\mathbf{w}^* = \mathbf{w}(\mathbf{f}^*)$. To remove $\mathbf{w}'$ from this equation, we introduce the Lagrange variables $(p^*, q^*) \in \mathbf{H}^1(\Omega)$ as the unique solution to the adjoint equations (26)-(28). Since

$$(u^* - U) = \partial_x p^* + \partial_y q^*$$
$$-(v^* - V) = \partial_x q^* - \partial_y p^*,$$

we replace $\mathbf{w}^* - \mathbf{W}$ in (32) and using (30)-(31) with the the Gauss-Green theorem [9], we obtain

$$\hat{J}'(\mathbf{f}^*; \delta\mathbf{f}) = -(p^*, \delta f)_{L^2(\Omega)} + \alpha(f^*, \delta f)_{L^2(\Omega)} + (q^*, \delta g)_{L^2(\Omega)} + \beta(g^*, \delta g)_{L^2(\Omega)}$$
$$= (\alpha\, f^* - p^*, \delta f)_{L^2(\Omega)} + (\beta\, g^* + q^*, \delta g)_{L^2(\Omega)} = 0, \tag{33}$$

for all $\delta\mathbf{f} \in \mathbf{L}^2(\Omega)$. Thus, we have obtained the gradient of the reduced cost functional, $\nabla_f \hat{J}(f, g) := \alpha\, f - p$ and $\nabla_g \hat{J}(f, g) := \beta\, g + q$, and formally $\nabla \hat{J}(\mathbf{f}^*) = 0$ constitutes the necessary and sufficient (because of convexity) optimality condition for the CR optimal control problem. □

5 Discretization of the CR optimality system

In this section, we illustrate the discretization of the CR optimality system by finite difference approximations on staggered grids [4]. This discretization framework is typical in the approximation of first-order partial differential systems [13]. We provide implementation details and notice the advantageous collocation of the optimization variables.

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

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

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 vertices, cell centers, and cell edges that can be horizontal and vertical. We denote these sets of grid points with Ω_h^s, $s \in \{v, c, eh, ev\}$ with the obvious meaning of the alphabetic index. Notice that within the same set the grid points are spaced with h_x and h_y intervals in the x- and y-direction, respectively.

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

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

with associated norm $\|v^h\|_{L_h^2(\Omega_h^s)} = (v^h, v^h)_{L_h^2(\Omega_h^s)}^{1/2}$. We require as well the discrete H^1-norm given by

$$\|v^h\|_{H_h^1(\Omega_h^s)} = \left(\|v^h\|_{L_h^2(\Omega_h^s)}^2 + \|\partial_x^h v^h\|_{L_h^2(\Omega_h^s)}^2 + \|\partial_y^h v^h\|_{L_h^2(\Omega_h^s)}^2 \right)^{1/2},$$

where ∂_x^h and ∂_y^h denote the following difference quotients in the x- and y-direction, respectively. We have

$$\partial_x^h v(x, y) = \frac{v(x + h_x, y) - v(x, y)}{h_x} \qquad \partial_y^h v(x, y) = \frac{v(x, y + h_y) - v(x, y)}{h_y}.$$

Here v^h is extended by 0 on grid points outside of Ω; see [15]. The spaces $L_h^2(\Omega_h^s)$ and $H_h^1(\Omega_h^s)$ consist of the sets of grid functions v^h defined on Ω_h^s endowed with $\|v^h\|_{L_h^2(\Omega_h^s)}$, respectively $\|v^h\|_{H_h^1(\Omega_h^s)}$, as norm. In the following, we denote with $\mathcal{U}_h$, $\mathcal{V}_h$, $\mathcal{P}_h$, and $\mathcal{Q}_h$ the space of the grid functions u^h, v^h, p^h, and q^h, that approximate the state and adjoint variables, respectively.

We first discuss the discretization of the Cauchy – Riemann state model. In the staggered grid the variable u is defined on Ω_h^{ev} while v is defined on Ω_h^{eh}; see Figure 1.

We consider a unique set of grid indices (i, j), $i = 1, \ldots, N_x + 1$, $j = 1, \ldots, N_y$ that index all grid points including the boundaries in a lexicographic order starting from the lowest-left corner $i = 1$, $j = 1$. The vertices coordinates

are given by $x_i = (i-1)\,h_x$ and $y_j = (j-1)\,h_y$. With $v_{i+1/2,j+1/2}$ we mean the discrete counterpart to $v(x_i + h_x/2, y_j + h_y/2)$. The set of CR equations is given by

$$\frac{u_{i+1,j} - u_{i,j}}{h_x} + \frac{v_{i+1/2,j+1/2} - v_{i+1/2,j-1/2}}{h_y} \;=\; f_{i+1/2,j}, \qquad \text{on } \Omega_h^c \quad (34)$$

$$\frac{v_{i+1/2,j+1/2} - v_{i-1/2,j+1/2}}{h_x} - \frac{u_{i,j+1} - u_{i,j}}{h_y} \;=\; g_{i,j+1/2}, \qquad \text{on } \Omega_h^v \quad (35)$$

where the divergence equation (34) is centered at all internal cell centers and the curl equation is centered at all internal cell vertices. With this discretization, second-order accurate solutions are obtained [10].

Next, we notice that the optimality conditions $\alpha\,f - p = 0$ and $\beta\,g + q = 0$ suggest that the staggered grid for the adjoint CR system should be such that p, similarly to f, is defined on Ω_h^c, and q, similarly to g, is defined on Ω_h^v. In this way, we implement a direct coupling between state, adjoint, and control variables without the need of interpolation. Therefore, we have the discrete adjoint CR system as follows

$$\frac{p_{i+1/2,j} - p_{i-1/2,j}}{h_x} + \frac{q_{i,j+1/2} - q_{i,j-1/2}}{h_y} \;=\; (u - U)_{i,j}, \qquad \text{on } \Omega_h^{ev} \quad (36)$$

$$\frac{q_{i+1,j+1/2} - q_{i,j+1/2}}{h_x} - \frac{p_{i+1/2,j+1} - p_{i+1/2,j}}{h_y} \;=\; -(v - V)_{i+1/2,j+1/2}, \qquad \text{on } \Omega_h^{eh} \quad (37)$$

In Figure 1, we depict the staggered grid for the optimality system, including the allocation of all variables. Notice that with our approach based on the optimality conditions, the variables f, p, and the variables g, q, share the same location and as a by-product they are appropriately located on the boundaries in order to implement the boundary conditions.

Summarizing, the discrete Cauchy – Riemann optimality system results as follows

$$
\begin{aligned}
\frac{u_{i+1,j} - u_{i,j}}{h_x} + \frac{v_{i+1/2,j+1/2} - v_{i+1/2,j-1/2}}{h_y} &= f_{i+1/2,j}, & \text{on } \Omega_h^c \\
\frac{v_{i+1/2,j+1/2} - v_{i-1/2,j+1/2}}{h_x} - \frac{u_{i,j+1} - u_{i,j}}{h_y} &= g_{i,j+1/2}, & \text{on } \Omega_h^v \\
\frac{p_{i+1/2,j} - p_{i-1/2,j}}{h_x} + \frac{q_{i,j+1/2} - q_{i,j-1/2}}{h_y} &= (u - U)_{i,j}, & \text{on } \Omega_h^{ev} \\
\frac{q_{i+1,j+1/2} - q_{i,j+1/2}}{h_x} - \frac{p_{i+1/2,j+1} - p_{i+1/2,j}}{h_y} &= -(v - V)_{i+1/2,j+1/2}, & \text{on } \Omega_h^{eh} \\
\alpha\,f_{i+1/2,j} - p_{i+1/2,j} &= 0 & \text{in } \Omega_h^c \\
\beta\,g_{i,j+1/2} + q_{i,j+1/2} &= 0 & \text{in } \Omega_h^v,
\end{aligned}
\quad (38)
$$

with the following boundary conditions

$$
\begin{aligned}
u_{i,j} &= 0 & & \text{for } i = 1,\ldots,N_x+1,\; j=1;\; i=1, N_x+1,\; j=1,\ldots,N_y \\
v_{i+1/2,j+1/2} &= 0 & & \text{for } i = 1,\ldots,N_x,\; j=N_y \\
p_{i+1/2,j} &= 0 & & \text{for } i = 1,\ldots,N_x,\; j=1 \\
q_{i,j+1/2} &= 0 & & \text{for } i = 2,\ldots,N_x,\; j=N_y;\; i=1, N_x+1,\; j=1,\ldots,N_y-1.
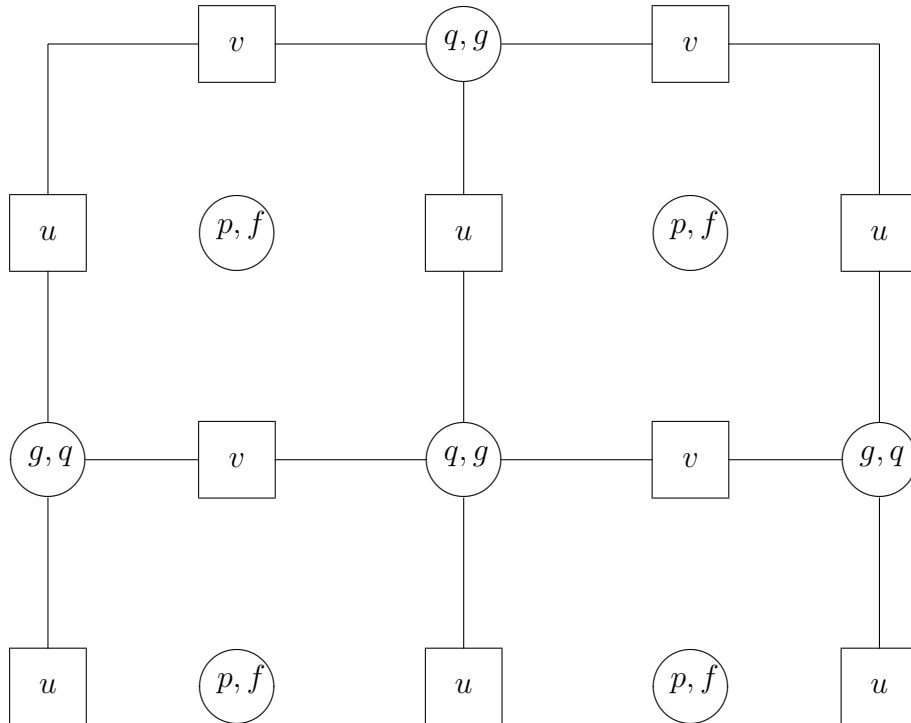\end{aligned}
\quad (39)
$$

Figure 1: Staggered grid for the CR optimality system.

Notice that following [6, 10], it is possible to prove second-order accuracy of the solution to the optimality system. To solve (38)-(39), in the next sections we discuss a multigrid framework that allows to demonstrate second-order accurate approximation.

6   A multigrid framework

We develop a multigrid scheme to solve the optimality system (38)-(39) defined on staggered grids as the one depicted in Figure 1 (coarsest grid). Our work extends previous developments on multigrid methods for CR systems [4, 7, 18] to the case of CR optimality systems. This development faces some difficulties due to the nature of the staggered grids and of the structure of the coupled forward and adjoint CR equations.

Concerning the grid, we anticipate the fact that a multigrid scheme uses different discretization grids on which the differential problem is represented. Usually, such grids are obtained by refining a starting coarse grid, as the one depicted in Figure 1. This refinement process consists in, e.g., halving the mesh size in some or all directions [21]. However, this procedure results in a non-nested hierarchy of grids such that the same type of variables are placed on

different spatial coordinates corresponding to different levels of the hierarchy. This fact is well known and requires additional effort in the construction of the required intergrid transfer operators; see [4, 18]. To alleviate these difficulties and recover the advantages of a nested sequence of grids, we notice that starting from the given coarse grid a nested sequence of grids is obtained by tripling the mesh size. This remark seems novel in the staggered-grid context and it has important consequences in the development of multigrid schemes on staggered grids. See [8] for a coarsening strategy by a factor of three in the context of cell-centered discretizations of second-order PDE systems.

Another important concern in the development of multigrid schemes is the construction of appropriate smoothers. This is also a challenging problem with staggered grids and first-order PDEs because there is no one-to-one correspondence between variables and equations. That is, there is no one-to-one correspondence between residual of equations and variable updates in a smoothing process. To address this difficulty, different block smoothing strategies have been suggested as the distributed relaxation proposed in [4] and the coupled relaxation first proposed in [23] for solving the Navier-Stokes equations on staggered grids; see [18] for a review. Concerning iterative schemes to solve (38)-(39), our experience with coupled relaxation has been less successful while we have developed an appropriate smoothing scheme based on the distributed relaxation strategy.

In the following, we illustrate our smoothing scheme and the other components of our multigrid method. We define a sequence of nested grids (also referred to as levels) Ω_k of mesh size $h_{xk} = h_{x1}/3^{(k-1)}$ and $h_{yk} = h_{y1}/3^{(k-1)}$, $k = 1, \ldots, L$, where $k = L$ is the finest level and h_{x1} and h_{y1} are the mesh sizes of the coarsest grid in the x and y direction, respectively. In the following, denote all operators and functions defined on Ω_k in terms of the index k. Notice that with this setting a variable ϕ_{IJ}^{k-1} at the grid point (I, J) of the coarse grid Ω_{k-1} has the same spatial placement as the variable ϕ_{ij}^k at the grid point (i, j) of the fine grid Ω_k as follows

- u_{IJ}^{k-1} corresponds to u_{ij}^k, with $i = 3I - 2$ and $j = 3J - 2$;

- $v_{I+1/2, J+1/2}^{k-1}$ corresponds to $v_{i+1/2, j+1/2}^k$, with $i = 3I - 1$ and $j = 3J - 1$;

- $p_{I+1/2, J}^{k-1}$ corresponds to $p_{i+1/2, j}^k$, with $i = 3I - 1$ and $j = 3J - 2$;

- $q_{I, J+1/2}^{k-1}$ corresponds to $q_{i, j+1/2}^k$, with $i = 3I - 2$ and $j = 3J - 1$.

6.1 A distributed Gauss-Seidel scheme

Most pointwise iterative schemes are based on a one-to-one correspondence between equations and unknowns. That is, a step of the iteration on a single (grid-point) variable is defined with the purpose to satisfy the algebraic equation associated to that variable. Such one-to-one correspondence is natural in second-order elliptic problems but not on first-order systems. For this reason

12

distributive relaxation [4] was proposed, that aims at satisfying at each step the set of discrete equations by distributing changes to several unknowns.

As discussed in [4], to derive a natural distributive scheme we note that neither the divergence equation nor the curl equations are elliptic by themselves. However, combining them by differentiation one can obtain two Poisson equations to which a classical Gauss-Seidel scheme can be applied. By reverse engineering this procedure, one arrives at a distributed relaxation applied directly to the div-curl equations. However, in the case of systems of coupled div-curl systems this reverse engineering approach is difficult to apply because of the many different stencils involved and the fact that the coupling between the corresponding 'elliptic problems' appears through convective terms. For this reason, we decide to formulate a distributed relaxation technique which applies sequentially to the forward and adjoint div-curl equations and updates the control variables through a gradient step. See [3] for a similar procedure to solve second-order PDE control problems.

Let $(u^h, v^h, p^h, q^h, f^h, g^h)$ be the current approximation to the numerical solution. We define an update to this approximation by an sequence of iterative steps.

We start with the update of the control functions. For this purpose, we perform a gradient update, that is,

$$
\begin{aligned}
f^h &:= f^h - t\,\nabla_f \hat{J}(f^h, g^h) \\
g^h &:= g^h - t\,\nabla_g \hat{J}(f^h, g^h)
\end{aligned}
$$

where $\nabla_f \hat{J}(f^h, g^h) = \alpha f^h - p^h$ and $\nabla_g \hat{J}(f^h, g^h) = \beta g^h + q^h$ are the gradients in the control spaces and $t \in (0, 1]$ is the steplegth used in the gradient update. Numerical experience shows that the choice $t = 1$ always provides good convergence rates.

Next, we implement the iterative step for the divergence equation defined at the cell center $(i + 1/2, j)$, $i = 1, \ldots, N_x$, $j = 2, \ldots, N_y$, where we compute the following 'dynamic' residual [4]. We have

$$
r_1^h = f^h - \partial_x^h u^h - \partial_y^h v^h. \tag{40}
$$

We call r_1 a dynamic residual because it is the residual at the cell center just before the update step. This step is as follows

$$
\begin{aligned}
u_{i+1,j} &= u_{i+1,j} + h_x h_y^2 \delta_1 \\
u_{i,j} &= u_{i,j} - h_x h_y^2 \delta_1 \\
v_{i+1/2,j+1/2} &= v_{i+1/2,j+1/2} + h_y h_x^2 \delta_1 \\
v_{i+1/2,j-1/2} &= v_{i+1/2,j-1/2} - h_y h_x^2 \delta_1
\end{aligned}
$$

where

$$
\delta_1 = \frac{1}{2(h_x^2 + h_y^2)}\, r_1^h.
$$

13

With this distributed update, we have that the residual (40) becomes zero while the residuals of the curl equation at all vertices remain unchanged. In the case where one of the updates at the boundaries are not allowed because of boundary conditions, we need to modify δ_1 as follows

$$\delta_1 = \frac{1}{2(h_x^2 + h_y^2) - d} r_1^h$$

where $d \in \{h_x^2, h_y^2, h_x^2 + h_y^2\}$ depending upon which boundary is touched. That is, $d = h_y^2$ whenever one of the u updates is not performed and $d = h_x^2$ if one of the v updates is not performed. We have $d = h_x^2 + h_y^2$ in the case when both u and v are not updated at the boundary (corner).

The update step for the curl equation at $(i, j + 1/2)$, $i = 2, \ldots, N_x$, $j = 1, \ldots, N_y - 1$, is made in a similar way. Consider the dynamic residual

$$r_2^h = g^h + \partial_y^h u^h - \partial_x^h v^h, \tag{41}$$

at $(i, j + 1/2)$ before the update. Define

$$\delta_2 = \frac{1}{2(h_x^2 + h_y^2)} r_2^h.$$

The update step for the curl equation is given by

$$
\begin{aligned}
u_{i,j+1} &= u_{i,j+1} - h_y h_x^2 \delta_2 \\
u_{i,j} &= u_{i,j} + h_y h_x^2 \delta_2 \\
v_{i+1/2,j+1/2} &= v_{i+1/2,j+1/2} + h_x h_y^2 \delta_2 \\
v_{i-1/2,j+1/2} &= v_{i-1/2,j+1/2} - h_x h_y^2 \delta_2
\end{aligned}
$$

This distributed update is such that the residuals of the divergence equation at all centers remains unchanged while the residual (41) at (x, y) becomes zero. Also in the relaxation of the curl equation, the term δ_2 becomes

$$\delta_2 = \frac{1}{2(h_x^2 + h_y^2) - d} r_2^h,$$

where $d \in \{h_x^2, h_y^2, h_x^2 + h_y^2\}$ depending upon which boundary is touched. That is, $d = h_x^2$ whenever one of the u updates is not performed and $d = h_y^2$ if one of the v updates is not performed. We have $d = h_x^2 + h_y^2$ in the case when both u and v are not updated at the boundary.

Next, we discuss the distributed relaxation for the adjoint system. We consider the adjoint divergence equation defined at (i, j), $i = 2, \ldots, N_x$, $j = 2, \ldots, N_y$, where the following 'dynamic' residual is computed

$$r_3^h = (u - U)^h - \partial_x^h p^h - \partial_y^h q^h. \tag{42}$$

14

Based on this dynamic residual, we have the following relaxation step for the first equation of the adjoint system

$$
\begin{aligned}
p_{i+1/2,j} &= p_{i+1/2,j} + h_x h_y^2 \, \delta_3 \\
p_{i-1/2,j} &= p_{i-1/2,j} - h_x h_y^2 \, \delta_3 \\
q_{i,j+1/2} &= q_{i,j+1/2} + h_y h_x^2 \, \delta_3 \\
q_{i,j-1/2} &= q_{i,j-1/2} - h_y h_x^2 \, \delta_3
\end{aligned}
$$

where the term δ_3 is given by

$$
\delta_3 = \frac{1}{2(h_x^2 + h_y^2) - d} \, r_3^h,
$$

where d is defined as in the forward divergence equation.

The adjoint curl equation is defined at $(i + 1/2, j + 1/2)$, $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y - 1$, we compute the corresponding residual as follows

$$
r_4^h = -(v - V)^h - \partial_x^h q^h + \partial_y^h p^h. \tag{43}
$$

The distributed relaxation results in the following update step for the second equation of the adjoint system

$$
\begin{aligned}
p_{i+1/2,j+1} &= p_{i+1/2,j+1} - h_y h_x^2 \, \delta_4 \\
p_{i+1/2,j} &= p_{i+1/2,j} + h_y h_x^2 \, \delta_4 \\
q_{i+1,j+1/2} &= q_{i+1,j+1/2} + h_x h_y^2 \, \delta_4 \\
q_{i,j+1/2} &= q_{i,j+1/2} - h_x h_y^2 \, \delta_4
\end{aligned}
$$

where the term δ_4 is given by

$$
\delta_4 = \frac{1}{2(h_x^2 + h_y^2) - d} \, r_4^h,
$$

where d is defined as in the forward curl equation.

The above steps update the controls, the state variables, and the adjoint variables.

Notice that all relaxation steps can be performed in any ordering and we choose lexicographic ordering. The effectiveness of the smoothing procedure could be improved by applying a linesearch to determine the steplength t in the controls update at each grid point where the controls are defined. However, this computation would make the smoothing step very CPU expensive.

6.2   Inter-grid transfer operators

It is well known [14] that since the governing equations are only first-order differential equations, it is admissible to have a coarse-to-fine first-order piecewise interpolation. Indeed, higher-order interpolation is advantageous but difficult

to implement in a non-nested staggered grid. On the other hand, using a coarsening by a factor of three, a nested staggered grid is obtained and the implementation of bilinear interpolation results easier.

In order to illustrate the interpolation procedure, consider the space $\mathcal{U}_k$ of $u^k : \Omega_k^{ev} \to \mathbb{R}$, $k = 1, \ldots, L$. Among two grids Ω_k and Ω_{k-1}, we define a prolongation operator, $I_{k-1}^k : \mathcal{U}_{k-1} \to \mathcal{U}_k$, that is consistent with the assumption of bilinear finite elements on each rectangular partition of the discretization. That is, on each rectangular partition $[x_I, x_{I+1}] \times [y_J, y_{J+1}]$ of $\overline{\Omega}_{k-1}$, define $s_x = x_{I+1} - x_I$ and $s_y = y_{J+1} - y_J$, and the piecewise bilinear function which interpolates u is given by

$$\tilde{u}(x, y) = \frac{(x_{I+1} - x)(y_{J+1} - y)}{s_x \; s_y} u_{IJ} + \frac{(x - x_I)(y_{J+1} - y)}{s_x \; s_y} u_{I+1\,J}$$
$$+ \frac{(x_{I+1} - x)(y - y_J)}{s_x \; s_y} u_{I\,J+1} + \frac{(x - x_I)(y - y_J)}{s_x \; s_y} u_{I+1\,J+1}.$$

Notice that the prolongation of u on a nested coarse-grid point at one of the vertices of the rectangle of reference gives the value of u on the fine grid corresponding to that grid point location.

In order to transfer residuals and solution functions from the fine to the coarse grids, we use the straight injection operator $I_k^{k-1} : \mathcal{U}_k \to \mathcal{U}_{k-1}$. The same operators apply also to the other function spaces.

6.3 The multigrid algorithm

The multigrid algorithm presented in this section is based on the full approximation storage (FAS) framework. This is a natural choice in the treatment of optimization problems [5], because the state, the adjoint, and the control functions, and not their errors, are available at all levels. This is essential when solving nonlinear problems and control problems with constraints on the control.

To illustrate the multigrid method, consider the optimality system (38)-(39) at the discretization level k for the unknown variables $\Phi_k = (u_k, v_k, p_k, q_k, f_k, g_k)$. We write this system with a compact notation as follows

$$A_k(\Phi_k) = F_k. \tag{44}$$

Denote with $\Phi_k^{(l)} = S_k\,(\Phi_k^{(l-1)}, F_k)$, the result of our smoothing scheme given in Section 6.1. Suppose that we apply m_1-times this iteration to (44) starting with the current approximation $\Phi_k^{(0)}$ to obtain the approximate solution $\tilde{\Phi}_k = \Phi_k^{(m_1)}$.

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

$$A_k(\tilde{\Phi}_k + e_k) - A_k(\tilde{\Phi}_k) = r_k, \tag{45}$$

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

Next, we represent the problem (45) on the coarser grid Ω_{k-1}. To represent $\tilde{\Phi}_k + e_k$ on this coarse grid we write

$$\Phi_{k-1} := I_k^{k-1}\tilde{\Phi}_k + e_{k-1}. \qquad (46)$$

We have that $I_k^{k-1}\tilde{\Phi}_k$ and $\tilde{\Phi}_k$ should represent the same function 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 (45) on the coarse grid replace $A_k(\cdot)$ by $A_{k-1}(\cdot)$, $\tilde{\Phi}_k$ by $I_k^{k-1}\tilde{\Phi}_k$, and r_k by restriction $I_k^{k-1}r_k = I_k^{k-1}(F_k - A_k(\tilde{\Phi}_k))$. We get the following (FAS) equation

$$A_{k-1}(\Phi_{k-1}) = I_k^{k-1}(F_k - A_k(\tilde{\Phi}_k)) + A_{k-1}(I_k^{k-1}\tilde{\Phi}_k). \qquad (47)$$

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

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

The term τ_k^{k-1} is the fine-to-coarse defect or residual correction such that at convergence the solution to (47) coincides with the fine grid solution in the sense that $\Phi_{k-1} = I_k^{k-1}\Phi_k$. With Φ_{k-1} obtained solving (47) and from (46) we have

$$e_{k-1} = \Phi_{k-1} - I_k^{k-1}\tilde{\Phi}_k.$$

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

$$\Phi_k = \tilde{\Phi}_k + I_{k-1}^k(\Phi_{k-1} - I_k^{k-1}\tilde{\Phi}_k). \qquad (48)$$

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 (48) is followed by m_2-times smoothing iteration.

The procedure just described corresponds to one multigrid cycle that is repeated iteratively until a given tolerance of the numerical approximation is obtained. This multigrid cycle is summarized in the following algorithm.

Algorithm 2 Multigrid MG(m_1, m_2) method for solving $A_k(\Phi_k) = F_k$.

1. *If $k = 1$ solve $A_k(\Phi_k) = F_k$ exactly.*

2. *Pre-smoothing steps on the fine grid: $\Phi_k^{(l)} = S_k(\Phi_k^{(l-1)}, F_k)$, $l = 1, \ldots, m_1$;*

3. *Computation of the residual: $r_k = F_k - A_k(\Phi_k^{(m_1)})$;*

4. *Restriction of the residual: $r_{k-1} = I_k^{k-1}r_k$;*

5. *Set $\Phi_{k-1} = I_k^{k-1}\Phi_k^{(m_1)}$;*

6. *Set $F_{k-1} = r_{k-1} + A_{k-1}(\Phi_{k-1})$*

7. *Call m times MG(m_1, m_2) to solve $A_{k-1}(\Phi_{k-1}) = F_{k-1}$;*

8. *Coarse-grid correction:* $\Phi_k^{(m_1+1)} = \Phi_k^{(m_1)} + I_{k-1}^k(\Phi_{k-1} - I_k^{k-1}\Phi_k^{(m_1)})$;

9. *Post-smoothing steps on the fine grid:* $\Phi_k^{(l)} = S_k(\Phi_k^{(l-1)}, F_k)$, $l = m_1 + 2, \ldots, m_1 + m_2 + 1$;

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].

7 Numerical experiments

In this section, we present a numerical investigation of the proposed CR optimal control formulation using staggered-grid discretization and the multigrid solution process. For this purpose, we consider a setting for which the exact solution is available. This fact allows us to evaluate the accuracy of the discretization scheme used.

Consider the CR optimal control problem $(14) - (18)$ on the rectangular domain $\Omega = (0,1) \times (0,3/2)$ depicted in Figure 1, where $\Gamma_D = (0,1) \times \{0\}$ and $\Gamma_N = \partial\Omega \setminus \overline{\Gamma}_D$. An exact solution to this problem is obtained with the following setting. Take

$$
\begin{aligned}
f(x,y) &:= 0 \\
g(x,y) &:= -2\pi \sin(\pi x)\cos(\pi y) \\
U(x,y) &:= (1 + 2\beta\pi^2)\sin(\pi x)\sin(\pi y) \\
V(x,y) &:= (1 + 2\beta\pi^2)\cos(\pi x)\cos(\pi y)
\end{aligned}
$$

With this choice, the exact solution is given by

$$
\begin{aligned}
u(x,y) &= \sin(\pi x)\sin(\pi y) \\
v(x,y) &= \cos(\pi x)\cos(\pi y) \\
p(x,y) &= 0 \\
q(x,y) &= 2\beta\pi \sin(\pi x)\cos(\pi y)
\end{aligned}
$$

The boundary conditions for the forward problem are given by

$$
\begin{aligned}
u &= 0 \quad \text{on} \quad \Gamma_D = (0,1) \times \{0\} \\
u &= 0 \quad \text{on} \quad \Gamma_1 = \{0\} \times (0,3/2) \\
u &= 0 \quad \text{on} \quad \Gamma_2 = \{1\} \times (0,3/2) \\
v &= 0 \quad \text{on} \quad \Gamma_3 = (0,1) \times \{3/2\}
\end{aligned}
$$

and for adjoint system we have

$$
\begin{aligned}
p &= 0 \quad \text{on} \quad \Gamma_D = (0,1) \times \{0\} \\
q &= 0 \quad \text{on} \quad \Gamma_1 = \{0\} \times (0,3/2) \\
q &= 0 \quad \text{on} \quad \Gamma_2 = \{1\} \times (0,3/2) \\
q &= 0 \quad \text{on} \quad \Gamma_3 = (0,1) \times \{3/2\}.
\end{aligned}
$$

To validate the accuracy of the discretization scheme, we use our multigrid method to solve the discrete optimality system (38)-(39) with the data given by the exact solution above. In all experiments, we employ W-cycles and $m_1 = m_2 = 2$. We find it advantageous to choose $t = 1$. Moreover, in the multigrid scheme we implement a stopping criteria with a given tolerance on the norm of the gradients as follows

$$\|\alpha\, f^h - p^h\|_{L^2} \leq 10^{-8} \qquad \text{and} \qquad \|\beta\, g^h + q^h\|_{L^2} \leq 10^{-8}.$$

In the Tables 1 - 5, we report results of experiments to validate the accuracy of our staggered-grids discretization scheme. For this purpose, the L^2-norm of solution errors are given corresponding to different meshes and different values of the optimization parameters α and β. We obtain second-order accuracy of the numerical solution. In fact, as we refine the mesh by a factor of three, we have a reduction of the norm of the errors of a factor 3^2.

Table 1: L^2-norm of errors with $\alpha = \beta = 1$.

$N_x \times N_y$	$\|u - u^h\|_{L^2}$	$\|v - v^h\|_{L^2}$	$\|p - p^h\|_{L^2}$	$\|q - q^h\|_{L^2}$
6×5	3.41E-02	3.41E-02	6.84E-02	9.94E-02
18×14	3.71E-03	3.71E-03	7.36E-03	1.10E-02
54×41	4.11E-04	4.11E-04	8.14E-04	1.23E-03
162×122	4.57E-05	4.57E-05	9.04E-05	1.36E-04

Table 2: L^2-norm of errors with $\alpha = \beta = 10^{-1}$.

$N_x \times N_y$	$\|u - u^h\|_{L^2}$	$\|v - v^h\|_{L^2}$	$\|p - p^h\|_{L^2}$	$\|q - q^h\|_{L^2}$
6×5	2.34E-02	2.34E-02	6.73E-03	3.41E-03
18×14	2.58E-03	2.58E-03	7.34E-04	3.97E-04
54×41	2.87E-04	2.87E-04	8.14E-05	4.44E-05
162×122	3.18E-05	3.18E-05	9.04E-06	4.93E-06

Table 3: L^2-norm of errors with $\alpha = \beta = 10^{-2}$.

$N_x \times N_y$	$\|u - u^h\|_{L^2}$	$\|v - v^h\|_{L^2}$	$\|p - p^h\|_{L^2}$	$\|q - q^h\|_{L^2}$
6×5	5.64E-03	5.64E-03	6.54E-04	7.42E-04
18×14	6.39E-04	6.39E-04	7.32E-05	8.19E-05
54×41	7.11E-05	7.11E-05	8.14E-06	9.09E-06
162×122	7.89E-06	7.89E-06	9.04E-07	1.00E-06

Table 4: L^2-norm of errors with $\alpha = \beta = 10^{-3}$.

$N_x \times N_y$	$\|u - u^h\|_{L^2}$	$\|v - v^h\|_{L^2}$	$\|p - p^h\|_{L^2}$	$\|q - q^h\|_{L^2}$
6×5	6.57E-04	6.57E-04	6.49E-05	1.05E-04
18×14	7.50E-05	7.50E-05	7.31E-06	1.17E-05
54×41	8.33E-06	8.33E-06	8.14E-07	1.30E-06
162×122	9.07E-07	9.05E-07	9.04E-08	1.37E-07

Table 5: L^2-norm of errors with $\alpha = \beta = 10^{-4}$.

$N_x \times N_y$	$\|u - u^h\|_{L^2}$	$\|v - v^h\|_{L^2}$	$\|p - p^h\|_{L^2}$	$\|q - q^h\|_{L^2}$
6×5	6.68E-05	6.68E-05	6.48E-06	1.08E-05
18×14	7.61E-06	7.61E-06	7.31E-07	1.21E-06
54×41	8.26E-07	8.84E-07	8.14E-08	1.27E-07
162×122	2.05E-07	1.11E-07	9.05E-09	6.16E-09

In Table 6, we report the observed values of the asymptotic convergence factors of the multigrid scheme for different meshes and different values of the optimization parameters. We obtain multigrid convergence factors that are independent on the mesh size. However, these convergence factors are less competitive compared to textbook multigrid Poisson efficiency. Notice that the observed convergence performance deteriorates taking smaller values of the optimization parameters.

Table 6: Convergence factors ρ_u, ρ_v, ρ_p, ρ_q for u, v, p, q, respectively.

$\alpha = \beta$	$N_x \times N_y$	ρ_u	ρ_v	ρ_p	ρ_q
10^{-1}	54×41	3.04E-01	3.49E-01	3.26E-01	3.04E-01
	162×122	2.89E-01	3.35E-01	2.85E-01	2.89E-01
10^{-2}	54×41	5.49E-01	6.19E-01	5.37E-01	5.49E-01
	162×122	5.08E-01	5.86E-01	4.89E-01	5.08E-01
10^{-3}	54×41	6.28E-01	6.59E-01	6.44E-01	6.28E-01
	162×122	5.88E-01	6.74E-01	5.82E-01	5.88E-01
10^{-4}	54×41	6.31E-01	6.55E-01	6.48E-01	6.31E-01
	162×122	5.98E-01	6.71E-01	5.95E-01	5.98E-01

In Table 7, we report tracking errors for different values of the optimization parameters. We obtain better tracking for smaller values of these parameters. Notice that the improvement in tracking is less sensitive to the value of α because in this case the exact solution for the corresponding control function is zero.

Table 7: L^2-norm of the tracking errors.

α	β	$\|u - U\|_{L^2}$	$\|u - V\|_{L^2}$
10^{-1}	10^{-1}	1.21E-00	1.21E-00
	10^{-3}	1.22E-02	1.20E-02
	10^{-5}	3.35E-04	8.79E-05
10^{-3}	10^{-1}	1.21E-00	1.21E-00
	10^{-3}	1.21E-02	1.21E-02
	10^{-5}	2.85E-04	1.56E-04
10^{-5}	10^{-1}	1.21E-00	1.21E-00
	10^{-3}	1.21E-02	1.21E-02
	10^{-5}	2.83E-04	1.58E-04

8   Conclusion

The investigation of distributed optimal control problems governed by Cauchy-Riemann equations was presented. In particular, existence and uniqueness of optimal solutions was proved. To compute these solutions, an optimality system of four first-order partial differential equations and two optimality conditions was formulated. In order to solve the optimality system, staggered grids and multigrid methods were developed. The multigrid scheme was based on a coarsening by a factor of three that resulted in a nested hierarchy of staggered grids. A smoothing algorithm consisting of a distributed-Gauss-Seidel scheme and a gradient-based control update was employed.

Results of numerical experiments demonstrated second-order accuracy of the optimality solutions and the ability of the multigrid scheme to solve the optimality system for different optimization parameters with mesh-independent convergence factors. We noticed that these convergence factors are not typical textbook multigrid. In fact, due to the tripling of the mesh size the spectrum of 'high-frequencies' that must be damped by the smoothing iteration is much larger and thus a less efficient scheme was obtained.

Acknowledgment

References

[1] R. Adams, Sobolev Spaces, Academic Press, New York, 1975

[2] S. Agmon, A. Douglis, and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II, Comm. Pure Appl. Math., 17 (1964), 35-92.

[3] E. Arian and S. Ta'asan, Multigrid one shot methods for optimal control problems: Infinite dimensional control, ICASE Report No. 94-52, 1994.

[4] A. Brandt, N. Dinar, Multigrid Solutions to Elliptic Flow Problems, Numerical Methods for PDEs, (1979) 53-147.

[5] A. Borzì and V. Schulz, Multigrid methods for PDE optimization, SIAM Review, 51 (2009), 361-395.

[6] A. Borzì, K. Kunisch, and D. Y. Kwak, Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system, SIAM J. Control Optim., 41 (2002), 1477-1497.

[7] A. Borzì, K.W. Morton, E. Süli, and M. Vanmaele, Multilevel solution of cell vertex Cauchy-Riemann equations, SIAM J. Sci. Comput., 18 (1997), 441-459.

[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), 577-598.

[9] L.C. Evans, Partial Differential Equations, Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, vol. 19, 2002.

[10] M. Ghil and R. Balgovind, A fast Cauchy-Riemann solver, Math. Comp., 33 (1979), 585-635.

[11] P. Grisvard, Singularities in Boundary value Problems, Springer-verlag, Berlin, 1992.

[12] M. D. Gunzburger and H.-C. Lee, Analysis and approximation of optimal control problems for first-order elliptic systems in three dimensions, Applied Mathematics and Computation, 100 (1999), 49-70.

[13] G. Fix and M. Rose, A comparative study of finite element and finite difference methods for Cauchy-Riemann type equations, SIAM J. Numer. Anal., 22 (1985), 250-261.

[14] W. Hackbusch, Multi-grid Methods and Applications, Springer-Verlag, New York, 1985.

[15] W. Hackbusch, Elliptic Differential Equations, Springer-Verlag, New York, 1992.

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

[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, Computing in Science & Engineering, 8 (2006), 34-43.

[19] S. Ta'asan, Canonical forms of multidimensional steady inviscid flows, Report No. 93-34, ICASE, Hampton, VA, 1993.

[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] D. Sidilkover, Factorizable schemes for the equations of fluid flow, Appl. Numer. Math., 41 (2002), 423-436.

[23] S.F. Vanka, Block-implicit multigrid solution of Navier-Stokes equations in primitive variables, J. Comp. Physics, 65 (1986), 138-158.

[24] M. Vanmaele, K.W. Morton, E. Süli, and A. Borzì, Analysis of the cell vertex finite volume method for the Cauchy-Riemann equations, SIAM J. Numer. Anal., 34 (1997), 2043-2062.