AIAA 2001-0821

Towards Adjoint-Based Methods for Aeroacoustic Control

S. Scott Collis, Kaveh Ghayour,
Matthias Heinkenschloss
Rice University
Houston, Texas 77005-1892

Michael Ulbrich, Stefan Ulbrich
Technische Universität München
D–80290 München, Germany

39th Aerospace Sciences Meeting & Exhibit
January 8–11, 2001
Reno, Nevada

AIAA 2001-0821

Towards Adjoint-Based Methods for Aeroacoustic Control

S. Scott Collis,* Kaveh Ghayour,†
Matthias Heinkenschloss‡
Rice University
Houston, Texas 77005-1892

Michael Ulbrich,§ Stefan Ulbrich§
Technische Universität München
D–80290 München, Germany

This paper is concerned with the numerical solution of optimal boundary control problems governed by the unsteady two-dimensional compressible Navier–Stokes equations. Specifically, results are presented for a model problem consisting of two counter-rotating viscous vortices above an infinite wall which, due to the self-induced velocity field, propagate downward and interact with the wall. The wall boundary control is the temporal and spatial distribution of wall-normal velocity which is used to minimize cost functionals of interest. The motivation for this work is on–blade control of aeroacoustic noise generated by blade-vortex interaction. We discuss some problem formulation issues, especially the choice of regularization terms; we outline our adjoint computations; and we present results from optimal control calculations using two different objectives and different control regularizations.

Introduction

The coupling of accurate computational fluid dynamics analyses with optimal control theory holds the promise for modifying a wide-range of fluid flows to achieve enhancement of desirable flow characteristics. Reduction of skin-friction drag, separation suppression, and increased lift to drag ratios for airfoils are examples of the types of optimization that such an approach enables.

The present investigation is part of our effort to put forward an optimal control framework for the control of aeroacoustic noise where the acoustic source is predicted by the unsteady, compressible, Navier-Stokes equations. The practical problem of interest is the control of the sound arising from Blade-Vortex Interaction (BVI) that can occur for rotorcraft in low speed, descending flight conditions, such as on approach to landing. When BVI occurs, tip vortices shed by a preceding blade interact with subsequent blades resulting in a high amplitude, impulsive noise that can dominate other rotorcraft noise sources. Reduction of the noise generated by this mechanism can alleviate restrictions on civil rotorcraft use near city centers and thereby enhance community acceptance. High frequency loading associated with this phenomenon also causes fatigue and hence reductions in BVI can have a direct impact on maintenance costs associated with blade failure in fatigue mode.

While several strategies for alleviating BVI have been examined, such as the use of a porous leading edge (Ref. 1), passive blade design (Ref. 2), and actively controlled trailing edge flaps (Ref. 3), an optimal control setting in which the mean flow and the noise component are computed by a high fidelity model of the flow physics has the potential to identify new strategies for BVI noise control.

In this investigation, the unsteady two-dimensional compressible Navier–Stokes equations are used to model the near-field flow as well as the propagation of small amplitude acoustic waves. Various controls can be incorporated within the optimal control framework but as an initial test, we consider unsteady wall-normal suction and blowing. The coupling of the near-field Navier–Stokes to far-field acoustic equations using a Kirchhoff type method and optimal control of the coupled system will be considered in forthcoming work. Since the compressible Navier-Stokes equations are an essential component of aeroacoustic simulation and since actuators for noise control will be located on the blade surface, i.e., on the boundary of the near-field, the computational tools developed for the present study and the observations made are immediately applicable to the coupled problem. While this investigation is motivated by aeroacoustic noise control, our findings extend to other control applications involving compressible Navier–Stokes equations.

The success of an optimal control approach pivots around two critical issues: the accuracy of the acoustic

*Assistant Professor, Department of Mechanical Engineering and Materials Science, collis@rice.edu

†Postdoctoral Associate, Department of Computational and Applied Mathematics and Department of Mechanical Engineering and Materials Science, kghayour@rice.edu

‡Associate Professor, Department of Computational and Applied Mathematics, heinken@rice.edu

§Assistant Professor, Lehrstuhl für Angewandte Mathematik und Mathematische Statistik, mulbrich@ma.tum.de,sulbrich@ma.tum.de

field prediction and the accuracy of the gradient for updating the control. Acoustic waves are non-dissipative and non-dispersive and, as such, high-order accurate numerical schemes with minimal dissipation and dispersion errors are essential for computational aeroacoustics. Our discretization is spatially sixth-order accurate using centered finite-difference stencils and a fourth-order accurate explicit Runge-Kutta scheme is used to advance the solution in time. In addition to accurate interior methods, accurate nonreflecting boundary conditions are also required in order to allow acoustic waves and spurious numerical waves to leave the computational domain without significant reflections. A variety of far-field boundary conditions ranging from a one dimensional Riemann invariant treatment to the application of dampers that parabolize the equations in the vicinity of the outflow boundary and buffer domains have been tried and tested extensively (Ref. 4). In the present study we use a Riemann extrapolation based inviscid far-field boundary condition (see Ref. 4).

In the optimization procedure, gradient information is computed using adjoint methods, which are widely used in flow control and design. References 5–11 present a small selection of papers in which adjoint based gradient methods are used to control unsteady incompressible Navier–Stokes flows of varying complexity. Adjoint based gradient methods for optimal design problems governed by the steady state compressible Navier–Stokes equations are discussed in Refs. 12, 13. To the best of the authors' knowledge, optimal control based on the unsteady, two-dimensional compressible Navier–Stokes equations has not been discussed in the literature to date.

There are two main approaches to derive adjoint equations for gradient computation. One approach is to derive adjoint and gradient equations in function spaces, i.e., on the partial differential equation level and then discretize the resulting equations. The other approach is to first discretize the objective function and Navier–Stokes equations and then perform gradient computations for the discrete problem, possibly with the aid of automatic differentiation (see Ref. 14). The first approach is referred to as 'optimize-then-discretize', and the latter approach as 'discretize-then-optimize'. Both approaches and hybrids of the two are used in the literature. As mentioned earlier, accuracy of the gradient is important for optimal control. However what 'accuracy' means in this context is not always correctly described. Often when the term 'accurate gradient' is used in the literature, it is used to mean that the gradient approximation computed for the discretized optimal control problem is a good approximation of the derivative of the discrete objective function. Such a property is necessary to ensure convergence, up to approximation error, of the iterates generated by the gradient based optimization method applied to the discretized optimal control problem. However, this meaning of 'accurate gradient' is incomplete. In fact, there are simple examples, see, e.g., section 6 in Ref. 15, which illustrate that exact solutions of the discretized optimal control problem may have nothing to do with the optimal control of the continuous problem. Therefore, it is also important that the discrete adjoint equations and discrete gradient equations (computed by the optimize-then-discretize approach, the discretize-then-optimize approach, or a hybrid) converge towards the adjoint equations and gradient equations, respectively, of the original infinite dimensional problem as the discretization is refined. This aspect of gradient accuracy requires a more comprehensive view of the problem that integrates well posedness of the infinite dimensional problem, existence of adjoint equations and gradient equations, and properties of the discretization. For incompressible Navier–Stokes equations, treatments of the infinite dimensional problem can be found, e.g., in Refs. 11, 16 and a general convergence analysis of discretizations with applications to the control of steady-state incompressible Navier–Stokes flow is given in Ref. 17.

Currently the mathematical foundation for optimal control problems governed by the unsteady compressible Navier–Stokes equations is not sufficiently developed to allow a rigorous and comprehensive study of gradient and adjoint accuracy in the previous sense. Even mathematical existence theories for the unsteady compressible Navier–Stokes equations are less developed than for the incompressible case. However to gain some confidence in our calculations we have implemented and compared adjoints and gradients computed using the optimize-then-discretize approach and using the discretize-then-optimize approach. Due to space restrictions the present study only contains results computed by the discretize-then-optimize approach. The optimize-then-discretize approach is described in Ref. 18. In the current study, adjoints at fixed times are computed using the automatic differentiation tool TAMC (Ref. 19) because it supports a sufficiently large subset of the Fortran90 language used in our flow solver. In this paper the role of the control regularization term in the objective function is also investigated. In particular, we show that inappropriate regularization terms can lead to highly oscillatory controls, where the oscillations are likely non-physical and due to an improper problem formulation. More detailed and comprehensive studies including comparisons between the two adjoint formulations will be presented in a forthcoming paper.

To make the current discussion concrete, we present results for a model problem consisting of two counter-rotating viscous vortices above an infinite wall which, due to the self-induced velocity field, propagate downward and interact with the wall. For this problem, the control is the temporal and spatial distribution of wall-normal velocity which is used to minimize cost functionals of interest. This problem is a reasonable, although highly idealized, test case for the blade-vortex interaction problem and it can potentially highlight the possible computational difficulties with the state and adjoint solutions for optimal control applied to the unsteady compressible Navier–Stokes equations.

Problem Formulation

The spatial domain occupied by the fluid is $\Omega = \{\mathbf{x} \in I\!R^2 : x_2 > 0\}$ and Γ denotes its spatial boundary. The portion of the boundary on which suction and blowing is applied is the controlled boundary

$$\Gamma_c = \{\mathbf{x} = (x_1, 0) : x_1 \in [a, b]\}.$$

Let

$$\mathbf{u} = (\rho, v_1, v_2, T)^T.$$

denote the primitive flow variables, where $\rho(t, \mathbf{x})$ is the density, $v_i(t, \mathbf{x})$ denotes the velocity in the x_i-direction, $i = 1, 2$, $\mathbf{v} = (v_1, v_2)^T$, and $T(t, \mathbf{x})$ denotes the temperature. The pressure p and the total energy per unit mass E are given by

$$p = \frac{\rho T}{\gamma \mathsf{M}^2}, \quad E = \frac{T}{\gamma(\gamma - 1)\mathsf{M}^2} + \frac{1}{2}\mathbf{v}^T\mathbf{v},$$

respectively, where γ is the ratio of specific heats and M is the reference Mach number. We write the conserved variables as functions of the primitive variables,

$$\mathbf{q}(\mathbf{u}) = (\rho, \rho v_1, \rho v_2, \rho E)^T$$

and we define the inviscid flux terms

$$\mathbf{F}^1(\mathbf{u}) = \begin{pmatrix} \rho v_1 \\ \rho v_1^2 + p \\ \rho v_2 v_1 \\ (\rho E + p) v_1 \end{pmatrix}, \quad \mathbf{F}^2(\mathbf{u}) = \begin{pmatrix} \rho v_2 \\ \rho v_1 v_2 \\ \rho v_2^2 + p \\ (\rho E + p) v_2 \end{pmatrix},$$

and the viscous flux terms

$$\mathbf{G}^i(\mathbf{u}) = \frac{1}{\mathsf{Re}} \begin{pmatrix} 0 \\ \tau_{1i} \\ \tau_{2i} \\ \tau_{1i} v_1 + \tau_{2i} v_2 + \dfrac{\kappa}{\mathsf{Pr}\mathsf{M}^2(\gamma - 1)} T_{x_i} \end{pmatrix},$$

$i = 1, 2$, where τ_{ij} are the elements of the stress tensor $\tau = \mu(\nabla \mathbf{v} + \nabla \mathbf{v}^T) + \lambda(\nabla \cdot \mathbf{v})I$. Here μ, λ are first and second coefficients of viscosity, κ is the thermal conductivity, Pr is the reference Prandtl number, and Re is the reference Reynolds number.

The two-dimensional compressible Navier–Stokes equations for the time interval $[0, t_f]$ can now be written as

$$\mathbf{q}(\mathbf{u})_t + \sum_{i=1}^{2} \left(\mathbf{F}^i(\mathbf{u})_{x_i} - \mathbf{G}^i(\mathbf{u})_{x_i}\right) = \mathbf{0} \tag{1}$$

in $(t_0, t_f) \times \Omega$ with boundary conditions

$$\mathbf{B}(\mathbf{u}, \mathbf{g}) = \mathbf{0} \tag{2}$$

on $(t_0, t_f) \times \Gamma$, and initial conditions $\mathbf{u}(t_0, \mathbf{x}) = \mathbf{u}_0(\mathbf{x})$ in Ω. The function $\mathbf{g}$ in the boundary conditions (2) acts as the control, which is taken to be suction and blowing in wall normal direction on Γ_c, and is modeled by

$$\mathbf{v} = \mathbf{b} + \mathbf{g} \quad \text{on } \Gamma_c. \tag{3}$$

Here $\mathbf{b}$ is a given boundary velocity that satisfies the compatibility condition $\mathbf{v}(t_0, \mathbf{x}) = \mathbf{b}(t_0, \mathbf{x})$ for $\mathbf{x} \in \Gamma$. Since $\Gamma_c \subset \{\mathbf{x} : x_2 = 0\}$, we have $\mathbf{g} = (0, g_2)^T$.

Abstractly, all optimal control problems treated in this paper are of the form

$$\min_{\mathbf{g} \in \mathcal{G}_{ad}} J(\mathbf{g}) \stackrel{\text{def}}{=} J_{\text{Obs}}(\mathbf{u}(\mathbf{g})) + J_{\text{Reg}}(\mathbf{g}), \tag{4}$$

where $\mathbf{g}$ is the control, $\mathcal{G}_{ad}$ is the set of admissible controls, $\mathbf{u}(\mathbf{g})$ is the solution of the compressible Navier–Stokes, J_{Obs} is the part of the cost function that represents our flow control objective (the 'observed part') and $J_{\text{Reg}}(\mathbf{g})$ is a regularization term, typically some weighted norm of the control. The concrete choices of the set of admissible controls $\mathcal{G}_{ad}$ and of the objective function J depend on the physical system we want to control (e.g., how much suction and blowing can the actuators provide, what part of the flow do we want to influence), but they also need to be chosen to ensure the existence of a solution $\mathbf{g}_*$ of the optimal control problem (4).

In this paper we consider two objectives. The first objective is the minimization of kinetic energy in $\Omega_0 \subset \Omega$ at the final time t_f,

$$J_{\text{Obs}}(\mathbf{u}(\mathbf{g})) = \frac{1}{2}\int_{\Omega_0} \rho(t_f, \mathbf{x}) \|\mathbf{v}(t_f, \mathbf{x})\|_2^2 d\mathbf{x} \tag{5}$$

and the second is the minimization of heat transfer over a subset Γ_0 of the bottom wall,

$$J_{\text{Obs}}(\mathbf{u}(\mathbf{g})) = \frac{1}{2}\int_{t_0}^{t_f} \int_{\Gamma_0} \omega(\mathbf{x}) \left(\frac{\partial}{\partial \mathbf{n}} T(t, \mathbf{x})\right)^2 d\mathbf{x}. \tag{6}$$

Here $\mathbf{n} = (n_1, n_2)^T$ is the outward unit normal and ω is a non–negative weighting function.

Given J_{Obs} we must choose $\mathcal{G}_{ad}$ and the regularization term $J_{\text{Reg}}(\mathbf{g})$ so that (4) is well-posed. General existence results for optimal control problems of the type considered here can be found in, e.g., Chapter 1 of Ref. 20. Their application to our problem is difficult, because general existence results for the compressible Navier–Stokes equations suitable for optimal control are still missing. We refer to Refs. 21–23 for recent results and further references on the existence of solutions of the compressible Navier–Stokes equations. Therefore our treatment of the infinite dimensional problem is formal. It follows the mathematical theory of optimal control of incompressible Navier–Stokes flow, especially Refs. 11, 20, 24. Our space of controls is either

$$\mathcal{G}_{ad} = \Big\{\mathbf{g} : \begin{array}{l} \mathbf{g} \in L^2((t_0, t_f); H^1(\Gamma_c)), \\ \frac{\partial}{\partial t}\mathbf{g} \in L^2((t_0, t_f); L^2(\Gamma_c)), \\ \mathbf{g}(t, \mathbf{x}) = 0 \text{ on } (t_0, t_f) \times \partial\Gamma_c, \\ \mathbf{g}(t_0, \mathbf{x}) = 0 \text{ in } \Gamma_c \end{array}\Big\}. \tag{7}$$

or

$$\mathcal{G}_{ad} = \Big\{ \mathbf{g} : \quad \mathbf{g} \in L^2((t_0,t_f);H^2(\Gamma_c)), \\ \tfrac{\partial}{\partial t}\mathbf{g} \in L^2((t_0,t_f);L^2(\Gamma_c)), \\ \mathbf{g}(t,\mathbf{x}) = 0, \nabla\mathbf{g}(t,\mathbf{x}) = 0 \text{ on } (t_0,t_f)\times\partial\Gamma_c, \\ \mathbf{g}(t_0,\mathbf{x}) = 0 \text{ in } \Gamma_c \Big\}. \tag{8}$$

Our notation for function spaces follows, e.g., Ref. 25. In (7), (8) $\nabla\mathbf{g}$ denotes the gradient of $\mathbf{g}$ on the boundary, which in our test problem is given by $\nabla\mathbf{g} = (0,(g_2)_{x_1})^T$. The first two conditions in (7) are smoothness conditions on the admissible controls in space and in time, respectively. The third condition ensures that there are no jumps in $\mathbf{b}+\mathbf{g}$ at $(t_0,t_f)\times\partial\Gamma_c$. The fourth condition enforces compatibility between initial velocity field $\mathbf{v}_0$ and boundary data $\mathbf{b}+\mathbf{g}$. A similar interpretation is valid for the conditions in (8), except that the require spatial regularity of the controls is one order higher. If our control space is (7), the regularization term J_{Reg} in (4) is chosen to be

$$J_{\mathrm{Reg}}(\mathbf{g}) \quad \stackrel{\mathrm{def}}{=} \quad \int_{t_0}^{t_f}\int_{\Gamma_c} \Big(\frac{\alpha_1}{2}\|\mathbf{g}_t\|_2^2 + \frac{\alpha_2}{2}\|\mathbf{g}\|_2^2 + \frac{\alpha_3}{2}\|\nabla\mathbf{g}\|_2^2 \Big) d\mathbf{x}dt. \tag{9}$$

Otherwise, if our control space is (8), the regularization term J_{Reg} in (4) is chosen to be

$$J_{\mathrm{Reg}}(\mathbf{g}) \quad \stackrel{\mathrm{def}}{=} \quad \int_{t_0}^{t_f}\int_{\Gamma_c} \Big(\frac{\alpha_1}{2}\|\mathbf{g}_t\|_2^2 + \frac{\alpha_2}{2}\|\mathbf{g}\|_2^2 + \frac{\alpha_3}{2}\|\nabla\mathbf{g}\|_2^2 + \frac{\alpha_4}{2}\|\Delta\mathbf{g}\|_2^2 \Big) d\mathbf{x}dt. \tag{10}$$

Here $\alpha_1,\alpha_2,\alpha_3,\alpha_3 \geq 0$, $\nabla\mathbf{g}$ is the gradient of $\mathbf{g}$ on the boundary, in our case $\nabla\mathbf{g} = (0,(g_2)_{x_1})^T$, and $\Delta\mathbf{g}$ is the Laplacian of $\mathbf{g}$ on the boundary, in our case $\Delta\mathbf{g} = (0,(g_2)_{x_1x_1})^T$. We ran experiments without regularization of the derivatives of $\mathbf{g}$, i.e., with $\alpha_1 = \alpha_3 = \alpha_4 = 0$ and $\alpha_2 > 0$. In these cases the optimization algorithm produced strongly oscillatory controls, which frequently led to a blow-up in the flow variables generated by our compressible Navier–Stokes solver. Therefore we use $\alpha_2,\alpha_3 > 0$ in all our computations reported here, only $\alpha_1,\alpha_4 \geq 0$ are allowed to be zero. If $\alpha_1 = 0$ we do not enforce smoothness of the control in time. In particular the requirements $\frac{\partial}{\partial t}\mathbf{g} \in L^2((t_0,t_f);L^2(\Gamma_c))$ and $\mathbf{g}(t_0,\mathbf{x}) = 0$ are dropped from (7), (8). As shown in the results section, the computed control can exhibit strong oscillations in time if $\alpha_1 = 0$.

Spatial Discretization

Our compressible Navier–Stokes flow solver is based on a conservative extension of the method described in Ref. 4. The Navier–Stokes equations are mapped to the computational space; a unit square divided into an equally spaced grid system. This transformation allows clustering of grid points in regions of high gradients, simplifies the implementation of the boundary conditions, and extends finite-difference techniques to moderately complex geometries. Spatial derivatives are approximated using sixth-order accurate central differences in the inner domain with third-order biased and one-sided differences used at boundaries that are designed to enhance stability when used with explicit time advancement methods (see Ref. 26).

Central-schemes are non-dissipative and require the addition of explicit artificial dissipation to suppress the growth of high-frequency error modes. These error modes stem from a variety of sources including approximate boundary treatments, mesh stretching, and conflicting initial and boundary conditions. In this study, a fourth-order artificial dissipation term acting of the numerical fluxes is added to the right hand side of the discretized equations to damp out error modes. This dissipation term is computed using sixth-order accurate finite-differences and the dissipation parameter, ϵ, is chosen judiciously to damp out the error modes while avoiding excessive dissipation in the solution as established through numerical experimentation.

The Semi-Discrete Problem

We consider the optimal control problem (4) with state equation (1) and (2) after a discretization in space has been performed. We allow the possibility that the objective function depends on the value of the flow variables at the final time and on the flow variables distributed in time. The final time contribution to the objective is denoted by $l^f(\mathbf{u}(t_f))$ and the distributed contribution is denoted by $\int_{t_0}^{t_f} l(\mathbf{u}(t),\mathbf{g}(t))dt$. This enables us to treat both objective functions (5) and (6) simultaneously. After a discretization in space, the optimal control problem can be written as

$$\min J(\mathbf{g}) = l^f(\mathbf{u}(t_f)) + \int_{t_0}^{t_f} l(\mathbf{u}(t),\mathbf{g}(t))dt + \int_{t_0}^{t_f} \frac{\alpha_1}{2}\mathbf{g}_t(t)^T\mathbf{L}_1\mathbf{g}_t(t) + \frac{\alpha_2}{2}\mathbf{g}(t)^T\mathbf{L}_2\mathbf{g}(t)dt \tag{11}$$

subject to the compatibility condition

$$\mathbf{g}(t_0) = \mathbf{0} \tag{12}$$

on the control, where $\mathbf{u}$ is the solution of the semi-discretized Navier–Stokes equations

$$\mathbf{q}(\mathbf{u}(t))_t + \mathbf{N}(\mathbf{u}(t),\mathbf{g}(t),t) = 0, \quad t\in[t_0,t_f], \tag{13}$$

$$\mathbf{u}(t_0) = \mathbf{u}_0. \tag{14}$$

The states $\mathbf{u}$ and controls $\mathbf{g}$ are vector valued functions with

$$\mathbf{u} : [t_0,t_f] \to I\!R^{n_u}, \quad \mathbf{g} : [t_0,t_f] \to I\!R^{n_g}.$$

The boundary conditions $\mathbf{g}(t,x) = 0$, $x\in\partial\Gamma_c$,:q

and, if (8) is used, $\nabla\mathbf{g}(t,x) = 0$, $x\in\partial\Gamma_c$, on the control are incorporated into the spatial discretization. We do not introduce different notation, such as superscript h to denote the semi-discretized states, controls, etc. The meaning of these variables and functions should be clear from the context.

The function $\mathbf{q} : I\!R^{n_u} \to I\!R^{n_u}$ represents the discretized conserved variables and $\mathbf{N} : I\!R^{n_u}\times I\!R^{n_g}\times I\!R \to I\!R^{n_u}$

represents the inviscid and viscid flux terms as well as boundary conditions. In the objective function (11),

$$l^f : I\!R^{n_u} \to I\!R, \quad l : I\!R^{n_u} \times I\!R^{n_g} \to I\!R,$$

and $\mathbf{L}_1, \mathbf{L}_2 \in I\!R^{n_g \times n_g}$ are symmetric matrices. The matrix $\mathbf{L}_1$ is such that $\mathbf{g}_t(t)^T \mathbf{L}_1 \mathbf{g}_t(t)$ is a discretization of $\int_{\Gamma_c} \|\mathbf{g}_t(t,\mathbf{x})\|_2^2 d\mathbf{x}$. The matrix $\mathbf{L}_2$ is such that $\mathbf{g}(t)^T \mathbf{L}_2 \mathbf{g}(t)$ is a discretization of $\int_{\Gamma_c} \|\mathbf{g}(t,\mathbf{x})\|_2^2 + \|\nabla \mathbf{g}(t,\mathbf{x})\|_2^2 d\mathbf{x}$ or of $\int_{\Gamma_c} \|\mathbf{g}(t,\mathbf{x})\|_2^2 + \|\nabla \mathbf{g}(t,\mathbf{x})\|_2^2 + \|\Delta \mathbf{g}(t,\mathbf{x})\|_2^2 d\mathbf{x}$, depending on whether (9) or (10) is used. The matrices $\mathbf{L}_1$, $\mathbf{L}_2$ also incorporate the boundary conditions $\mathbf{g}(t,x) = 0$, $x \in \partial\Gamma_c$. If the control space (8) and regularization term (10) are used, $\mathbf{L}_1$, $\mathbf{L}_2$ also incorporate the boundary conditions $\nabla \mathbf{g}(t,x) = 0$, $x \in \partial\Gamma_c$. Thus, $\mathbf{L}_1$ is a mass matrix on the boundary and $\mathbf{L}_2$ is the sum of a mass matrix and a stiffness matrix on the boundary. To simplify the presentation, we have assumed that $\alpha_2 = \alpha_3$ and that if $\alpha_4 > 0$, then $\alpha_2 = \alpha_3 = \alpha_4$.

For the numerical solution of (11)–(14) we need the gradient of the objective function J. Given an inner product $\langle \cdot, \cdot \rangle$ on the space of controls, the gradient (with respect to this inner product) is a function denoted by $(\nabla J(\mathbf{g}))(t)$ such that the Fréchet-derivative DJ of J can be represented as

$$DJ(\mathbf{g})\delta\mathbf{g} = \langle \nabla J(\mathbf{g}), \delta\mathbf{g} \rangle \tag{15}$$

for all admissible variations $\delta\mathbf{g}$ of the control. The choice of the inner product partly determines the conditioning of the problem and it can significantly influence the performance of gradient based optimization algorithms (see Ref. 27). However, often little attention is paid to the choice of the inner product and in many cases it is chosen to be $\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = \int_{t_0}^{t_f} \mathbf{g}_1(t)^T \mathbf{g}_2(t) dt$. The proper choice of the inner product depends on the control space $\mathcal{G}_{ad}$ which for our problem is given by (7) or (8). If we impose a regularization on the time derivative of the control, i.e., if $\alpha_1 > 0$, a proper inner product is

$$\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = \int_{t_0}^{t_f} (\mathbf{g}_1)_t(t)^T \mathbf{L}_1 (\mathbf{g}_2)_t(t) + \mathbf{g}_1(t)^T \mathbf{L}_2 \mathbf{g}_2(t) dt. \tag{16}$$

If $\alpha_1 = 0$, then a proper inner product is

$$\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = \int_{t_0}^{t_f} \mathbf{g}_1(t)^T \mathbf{L}_2 \mathbf{g}_2(t) dt. \tag{17}$$

Notice that the regularization terms (9), (10) in (4) correspond to weighted inner products. This is no coincidence. If no explict control norm-constraints are imposed in $\mathcal{G}_{ad}$, then typically a weighted norm is added to the objective function to ensure mathematical well-posedness.

Case $\alpha_1 > 0$.

If a regularization of the time-derivative of the control is imposed and if the control $\mathbf{g}$ satisfies (12), then one can show that the gradient of J with respect to the scalar product (16) is given by

$$\nabla J(\mathbf{g}) = \mathbf{k}, \tag{18}$$

where $\mathbf{k}$ solves

$$\begin{aligned} &-\mathbf{L}_1 \mathbf{k}_{tt}(t) + \mathbf{L}_2 \mathbf{k}(t) \\ &\quad = -\alpha_1 \mathbf{L}_1 \mathbf{g}_{tt}(t) + \alpha_2 \mathbf{L}_2 \mathbf{g}(t) + \nabla_{\mathbf{g}} l(\mathbf{u}(t), \mathbf{g}(t)) \\ &\qquad + \mathbf{N}_{\mathbf{g}}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t), \quad t \in [t_0, t_f], \\ &\mathbf{k}(t_0) = \mathbf{0}, \\ &\mathbf{k}_t(t_f) = \alpha_1 \mathbf{g}_t(t_f). \end{aligned} \tag{19}$$

Here $\boldsymbol{\lambda}$ is the solution of the adjoint equation

$$\begin{aligned} &-\mathbf{M}(\mathbf{u}(t))^T \boldsymbol{\lambda}_t(t) \\ &\quad = \quad -\nabla_{\mathbf{u}} l(\mathbf{u}(t), \mathbf{g}(t)) - \mathbf{N}_{\mathbf{u}}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t), \end{aligned} \tag{20}$$

$t \in [t_0, t_f]$, where

$$\mathbf{M}(\mathbf{u}(t)) = \frac{d}{d\mathbf{u}} \mathbf{q}(\mathbf{u}(t)), \tag{21}$$

with final conditions

$$\mathbf{M}(\mathbf{u}(t))^T \boldsymbol{\lambda}(t_f) = -\nabla_{\mathbf{u}} l^f(\mathbf{u}(t_f), t_f). \tag{22}$$

For details we refer to Ref. 18.

In the numerical solution of the discretized optimal control problem we want to avoid the solution of (19). Moreover, we need to ensure that the discretization of the objective function (11) is consistent with the time-discretization of (13). Therefore we introduce the additional variables

$$\mathbf{h}(t) = \mathbf{g}_t(t)$$

and

$$\begin{aligned} w(t) \quad = \quad & \int_{t_0}^{t} \Big(l(\mathbf{u}(\tau), \mathbf{g}(\tau)) + \tfrac{\alpha_1}{2} \mathbf{h}(\tau)^T \mathbf{L}_1 \mathbf{h}(\tau) \\ & + \tfrac{\alpha_2}{2} \mathbf{g}(\tau)^T \mathbf{L}_2 \mathbf{g}(\tau) \Big) d\tau. \end{aligned}$$

Since the time derivative is applied to the conserved variables $\mathbf{q}$ we view the primitive variables $\mathbf{u}$ as functions of the conserved variables $\mathbf{q}$. Thus, we rewrite (11)–(14) as

$$\min \widetilde{J}(\mathbf{h}) = l^f(\mathbf{u}(\mathbf{q}(t_f)), t_f) + w(t_f) \tag{23}$$

with state equations

$$\begin{aligned} \mathbf{z}_t(t) \quad &= \quad \mathbf{r}(\mathbf{z}(t), \mathbf{h}(t), t), \quad t \in [t_0, t_f], \\ \mathbf{z}(t_0) \quad &= \quad \mathbf{z}_0, \end{aligned} \tag{24}$$

where

$$\begin{aligned} &\mathbf{r}(\mathbf{z}(t), \mathbf{h}(t), t)) \\ &= \begin{pmatrix} l(\mathbf{u}(\mathbf{q}(t)), \mathbf{g}(t)) + \frac{\alpha_1}{2} \mathbf{h}(t)^T \mathbf{L}_1 \mathbf{h}(t) + \frac{\alpha_2}{2} \mathbf{g}(t)^T \mathbf{L}_2 \mathbf{g}(t) \\ -\mathbf{N}(\mathbf{u}(\mathbf{q}(t)), \mathbf{g}(t), t) \\ \mathbf{h}(t) \end{pmatrix} \end{aligned} \tag{25}$$

and

$$\mathbf{z}(t) = \begin{pmatrix} w(t) \\ \mathbf{q}(t) \\ \mathbf{g}(t) \end{pmatrix}, \quad \mathbf{z}_0 = \begin{pmatrix} 0 \\ \mathbf{q}(\mathbf{u}_0) \\ \mathbf{0} \end{pmatrix}. \tag{26}$$

In the reformulation (23)–(26) of (11)–(14) $\mathbf{z}$ plays the role of the state and $\mathbf{h}$ plays the role of the control. We define

$$f(\mathbf{z}(t_f), t_f) = l^f(\mathbf{u}(\mathbf{q}(t_f)), t_f) + w(t_f) \tag{27}$$

so that the objective (23) can be written as

$$\min \widetilde{J}(\mathbf{h}) = f(\mathbf{z}(t_f), t_f). \tag{28}$$

The Lagrangian function for (24) and (28) is given by

$$\begin{array}{l} L(\mathbf{z}, \mathbf{h}, \boldsymbol{\pi}, \boldsymbol{\pi}_0) \\ \quad = \quad f(\mathbf{z}(t_f), t_f) + \boldsymbol{\pi}_0^T (\mathbf{z}(t_0) - \mathbf{z}_0) \\ \qquad + \int_{t_0}^{t_f} \boldsymbol{\pi}(t)^T (\mathbf{z}_t(t) - \mathbf{r}(\mathbf{z}(t), \mathbf{h}(t), t))\, dt. \end{array} \tag{29}$$

The adjoint equations are

$$\begin{array}{rcl} -\boldsymbol{\pi}_t(t) & = & \mathbf{r}_\mathbf{z}(\mathbf{z}(t), \mathbf{h}(t), t))^T \boldsymbol{\pi}(t), \quad t \in [t_0, t_f], \\ \boldsymbol{\pi}(t_f) & = & -\nabla_\mathbf{z} f(\mathbf{z}(t_f), t_f), \end{array} \tag{30}$$

where $\mathbf{r}_\mathbf{z}$ is the partial Jacobian of $\mathbf{r}$ with respect to $\mathbf{z}$. The gradient of $\widetilde{J}$ is computed with respect to the scalar product

$$\langle \mathbf{h}_1, \mathbf{h}_2 \rangle = \int_{t_0}^{t_f} \mathbf{h}_1(t)^T \mathbf{L}_1 \mathbf{h}_2(t) dt. \tag{31}$$

In this case it can be computed by differentiation of the Lagrangian (29) with respect to $\mathbf{h}$ and is given by

$$(\nabla \widetilde{J}(\mathbf{h}))(t) = -\mathbf{r}_\mathbf{h}(\mathbf{z}(t), \mathbf{h}(t), t))^T \boldsymbol{\pi}(t), \tag{32}$$

where $\mathbf{r}_\mathbf{h}$ is the partial Jacobian of $\mathbf{r}$ with respect to $\mathbf{h}$.

If we set $\boldsymbol{\pi} = (\theta, \boldsymbol{\lambda}, \boldsymbol{\psi})^T$, $\mathbf{u}(t) = \mathbf{u}(\mathbf{q}(t))$, and use (25) and (27), then the adjoint equation (30) is

$$\begin{array}{rcl} -\theta_t(t) & = & 0, \\ -\mathbf{M}(\mathbf{u}(t))^T \boldsymbol{\lambda}_t(t) & = & \nabla_\mathbf{u} l(\mathbf{u}(t), \mathbf{g}(t))\theta(t) \\ & & -\mathbf{N}_\mathbf{u}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t), \\ -\boldsymbol{\psi}_t(t) & = & \alpha_2 \mathbf{L}_2 \mathbf{g}(t)\theta(t) + \nabla_\mathbf{g} l(\mathbf{u}(t), \mathbf{g}(t))\theta(t) \\ & & -\mathbf{N}_\mathbf{g}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t), \end{array} \tag{33}$$

for $t \in [t_0, t_f]$ with final conditions

$$\begin{pmatrix} \theta(t_f) \\ \boldsymbol{\lambda}(t_f) \\ \boldsymbol{\psi}(t_f) \end{pmatrix} = \begin{pmatrix} -1 \\ -\mathbf{M}(\mathbf{u}(t))^{-T} \nabla_\mathbf{u} l^f(\mathbf{u}(t_f), t_f) \\ \mathbf{0} \end{pmatrix}. \tag{34}$$

The gradient (32) is given by

$$(\nabla \widetilde{J}(\mathbf{h}))(t) = -\alpha_1 \theta(t) \mathbf{L}_1 \mathbf{h}(t) - \boldsymbol{\psi}(t). \tag{35}$$

The relations between (33)–(35) and (18)–(22), are explored in Ref. 18.

Case $\alpha_1 = 0$.

If no regularization is imposed on the time derivative of the control, the adjoint equations for (11)–(14) are still (20) with final conditions (22). However, the gradient of J in (11) in the inner product (17) is given by $\nabla J(\mathbf{g}) = \mathbf{k}$, where $\mathbf{k}$ solves $\mathbf{L}_2 \mathbf{k}(t) = \alpha_2 \mathbf{L}_2 \mathbf{g}(t) + \nabla_\mathbf{g} l(\mathbf{u}(t), \mathbf{g}(t)) + \mathbf{N}_\mathbf{g}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t)$, i.e.,

$$\begin{array}{rcl} \nabla J(\mathbf{g}) & = & \mathbf{L}_2^{-1} \big[\alpha_2 \mathbf{L}_2 \mathbf{g}(t) + \nabla_\mathbf{g} l(\mathbf{u}(t), \mathbf{g}(t)) \\ & & + \mathbf{N}_\mathbf{g}(\mathbf{u}(t), \mathbf{g}(t), t)^T \boldsymbol{\lambda}(t)\big]. \end{array} \tag{36}$$

To ensure consistency between discretization of the objective function and the time-discretization of the state equation, w is defined as before and we set

$$\begin{array}{l} \mathbf{r}(\mathbf{z}(t), \mathbf{g}(t), t)) \\ = \begin{pmatrix} l(\mathbf{u}(\mathbf{q}(t)), \mathbf{g}(t)) + \frac{\alpha_2}{2} \mathbf{g}(t)^T \mathbf{L}_2 \mathbf{g}(t) \\ -\mathbf{N}(\mathbf{u}(\mathbf{q}(t)), \mathbf{g}(t), t) \end{pmatrix} \end{array} \tag{37}$$

and $\mathbf{z} = (w, \mathbf{q})^T$, $\mathbf{z}_0 = (0, \mathbf{q}(\mathbf{u}_0))^T$. In this abstract setting our control problem can again be formulated as

$$\min \widetilde{J}(\mathbf{g}) = f(\mathbf{z}(t_f), t_f), \tag{38}$$

where f is defined in (27) and where $\mathbf{z}$ solves

$$\begin{array}{rcl} \mathbf{z}_t(t) & = & \mathbf{r}(\mathbf{z}(t), \mathbf{g}(t), t)), \quad t \in [t_0, t_f], \\ \mathbf{z}(t_0) & = & \mathbf{z}_0. \end{array} \tag{39}$$

Thus, using the abstract notation allows us to treat both cases simultaneously. We only have to substitute (37) for (25) and $\mathbf{g}$ for $\mathbf{h}$. For example the adjoint equations for (38)–(39) are

$$\begin{array}{rcl} -\boldsymbol{\pi}_t(t) & = & \mathbf{r}_\mathbf{z}(\mathbf{z}(t), \mathbf{g}(t), t))^T \boldsymbol{\pi}(t), \quad t \in [t_0, t_f], \\ \boldsymbol{\pi}(t_f) & = & -\nabla_\mathbf{z} f(\mathbf{z}_{N_t}, t_f). \end{array} \tag{40}$$

The gradient of $\widetilde{J}(\mathbf{g})$ in (38) with *respect to the inner product* $\langle \mathbf{h}_1, \mathbf{h}_2 \rangle = \int_{t_0}^{t_f} \mathbf{h}_1(t)^T \mathbf{h}_2(t) dt$ is

$$-\mathbf{r}_\mathbf{g}(\mathbf{z}(t), \mathbf{g}(t), t))^T \boldsymbol{\pi}(t), \tag{41}$$

where $\mathbf{r}_\mathbf{g}$ is the partial Jacobian of $\mathbf{r}$ with respect to $\mathbf{g}$. However, we want the gradient of $\widetilde{J}(\mathbf{g})$ in (39) with respect to the inner product (17). It is obtained by scaling (41) by $\mathbf{L}_2^{-1}$ and it is given by

$$(\nabla \widetilde{J}(\mathbf{h}))(t) = -\mathbf{L}_2^{-1} \mathbf{r}_\mathbf{g}(\mathbf{z}(t), \mathbf{g}(t), t))^T \boldsymbol{\pi}(t). \tag{42}$$

Not surprisingly, (42) and (36) are equivalent (see Ref. 18).

Time–Discretization

For the discretization of the state equation we use a Runge–Kutta method with coefficients

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ \vdots & \vdots & \vdots & & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} \tag{43}$$

and time steps $t_0 < t_1 < \ldots < t_{N_t} = t_f$. We set $\Delta t_n = t_{n+1} - t_n$. In our computational experiments we use the classical fourth order explicit Runge–Kutta method. In this case $s = 4$ and only $a_{21}, a_{32}, a_{43} \neq 0$. This structure, as well as the structure in the right hand side (25) or (37) of the state equation is exploited in our implementation. However, for our discussion of the time discretization, this is not important and therefore we treat the general case. More details are given in Ref. 18.

Remark 1 *Our abstract notation allows us to treat both cases, $\alpha_1 > 0$ and $\alpha_1 = 0$, simultaneously. The notation used in this section follows* (23)*,* (24)*. In particular the control is denoted by* $\mathbf{h}$*. If we consider the case $\alpha_1 = 0$ we only have to substitute* (37) *for* (25) *and* $\mathbf{g}$ *for* $\mathbf{h}$*.*

The Runge–Kutta method introduces intermediate time locations

$$t_{ni} = t_n + c_i \Delta t_n.$$

We let $\mathbf{z}_n$ be approximations of the states at t_n, $n = 1, \ldots, N_t$, and we introduce controls $\mathbf{h}_{ni}$, $i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$.

The discretized optimal control problem (23)–(24) is given by

$$\min \widetilde{J}(\mathbf{h}) = f(\mathbf{z}_{N_t}, t_f), \tag{44}$$

where

$$\mathbf{z}_{n+1} = \mathbf{z}_n + \Delta t_n \sum_{i=1}^{s} b_i \mathbf{r}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni}) \tag{45}$$

$$\mathbf{Z}_{ni} = \mathbf{z}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \mathbf{r}(\mathbf{Z}_{nj}, \mathbf{h}_{nj}, t_{nj}), \tag{46}$$

$i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$.

In the fully discretized control problem, we view (45) and (46) as the state equations. Thus, $\mathbf{z}_n$, $n = 0, \ldots, N_t$, and $\mathbf{Z}_{ni}$, $i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$, are the states and $\mathbf{h}_{ni}$, $i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$, are the controls.

Remark 2 *We point out that the discrete controls $\mathbf{h}_{ni}$ can, and often are poor approximations of the control $\mathbf{h}$ at t_{ni}. Actually, for some Runge–Kutta methods such as the classical Runge Kutta method, $t_{ni} = t_{nj}$ for some $i \neq j$, but in general one has $\mathbf{h}_{ni} \neq \mathbf{h}_{nj}$. Hager (Ref. 15) shows how to compute an approximation $\mathbf{h}_n$ of the control $\mathbf{h}$ at t_n. The 'discrete controls' $\mathbf{h}_{ni}$ should be viewed as artificial variables needed to obtain good approximations of the states and good approximations of the adjoints from which good approximations of the controls can be computed, see Remark* 3 *below.*

The Adjoint Equation and the Gradient for the Fully Discretized Problem

The Lagrangian function for (44)–(46) is given by

$$\begin{aligned} & L(\mathbf{z}, \mathbf{h}, \mathbf{Z}, \boldsymbol{\pi}, \boldsymbol{\Pi}) \\ & = f(\mathbf{z}_{N_t}, t_f) \\ & \quad + \sum_{n=0}^{N_t-1} \boldsymbol{\pi}_{n+1}^T \left[\mathbf{z}_{n+1} - \mathbf{z}_n - \Delta t_n \sum_{i=1}^{s} b_i \mathbf{r}_{ni} \right] \\ & \quad + \sum_{n=0}^{N_t-1} \sum_{i=1}^{s} \boldsymbol{\Pi}_{ni}^T \left[\mathbf{Z}_{ni} - \mathbf{z}_n - \Delta t_n \sum_{j=1}^{s} a_{ij} \mathbf{r}_{nj} \right], \end{aligned} \tag{47}$$

where $\mathbf{r}_{ni} = \mathbf{r}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})$

The adjoint equations are obtained by differentiating the Lagrangian function (47) with respect to $\mathbf{z}_n$, $n = 1, \ldots, N_t$, $\mathbf{Z}_{ni}$, $i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$, and setting the derivatives to zero. This gives the adjoint equations

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_{n+1} + \sum_{i=1}^{s} \boldsymbol{\Pi}_{ni}, \tag{48}$$

$n = 1, \ldots, N_t - 1$, and

$$\begin{aligned} \boldsymbol{\Pi}_{ni} &= \mathbf{r}_{\mathbf{z}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T \\ &\quad \times \Delta t_n [b_i \boldsymbol{\pi}_{n+1} + \textstyle\sum_{j=1}^{s} a_{ji} \boldsymbol{\Pi}_{nj}], \end{aligned} \tag{49}$$

$i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$, with

$$\boldsymbol{\pi}_{N_t} = -\nabla_{\mathbf{z}} f(\mathbf{z}_{N_t}, t_{N_t}). \tag{50}$$

The ni-th partial derivative of $\widetilde{J}(\mathbf{h})$ in (44) is obtained by differentiating the Lagrangian function (47) with respect to $\mathbf{h}_{ni}$. This yields

$$\begin{aligned} \tfrac{d}{d\mathbf{h}_{ni}} \widetilde{J}(\mathbf{h}) &= -\mathbf{r}_{\mathbf{h}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T \\ &\quad \times \Delta t_n [b_i \boldsymbol{\pi}_{n+1} + \textstyle\sum_{j=1}^{s} a_{ji} \boldsymbol{\Pi}_{nj}], \end{aligned} \tag{51}$$

$i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$. These partial derivatives are the components of the gradient with respect to the Euclidean inner product $\langle \mathbf{h}_1, \mathbf{h}_2 \rangle = \mathbf{h}_1^T \mathbf{h}_2$, where $\mathbf{h}_j$ is a vectors of components $\mathbf{h}_{ni}^{(j)}$, $i = 1, \ldots, s$, $n = 0, \ldots, N_t - 1$. This is commonly used in practice. However, this gradient is highly mesh dependent and can lead to artificial ill-conditioning in the problem. Instead, we use discrete versions of the inner products (31) and (17).

Gradient computation in the case $\alpha_1 > 0$:
The discretized version of the inner product (31) is

$$\langle \mathbf{h}_1, \mathbf{h}_2 \rangle = \sum_{n=0}^{N_t-1} \sum_{i=0}^{s} \Delta t_n b_i (\mathbf{h}_{ni}^{(j)})^T \mathbf{L}_1 \mathbf{h}_{ni}^{(j)} \tag{52}$$

and the gradient of (44) with respect to this inner product is the vector

$$\nabla \widetilde{J}(\mathbf{h}) = \left(\frac{1}{\Delta t_n b_i} \mathbf{L}_1^{-1} \frac{d}{d\mathbf{h}_{ni}} \widetilde{J}(\mathbf{h}) \right)_{\substack{n=0,\ldots,N_t-1 \\ i=1,\ldots,s}}, \tag{53}$$

where $\frac{d}{d\mathbf{h}_{ni}}\widetilde{J}(\mathbf{h})$ is given by (51). We have assumed that $b_1, \ldots, b_s \neq 0$, which is the case for the classical fourth order Runge Kutta method.

Gradient computation in the case $\alpha_1 = 0$:
We replace $\mathbf{h}$ by $\mathbf{g}$, see Remark 1. The discretized version of the inner product (17) is

$$\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = \sum_{n=0}^{N_t-1} \sum_{i=0}^{s} \Delta t_n b_i (\mathbf{g}_{ni}^{(j)})^T \mathbf{L}_2 \mathbf{g}_{ni}^{(j)}, \qquad (54)$$

and the gradient of (44) with respect to this inner product is the vector

$$\nabla \widetilde{J}(\mathbf{g}) = \left(\frac{1}{\Delta t_n b_i} \mathbf{L}_2^{-1} \frac{d}{d\mathbf{h}_{ni}} \widetilde{J}(\mathbf{g}) \right)_{\substack{n=0,\ldots,N_t-1 \\ i=1,\ldots,s}}, \qquad (55)$$

where $\frac{d}{d\mathbf{h}_{ni}}\widetilde{J}(\mathbf{g})$ is again given by (51) with $\mathbf{h}$ replaced by $\mathbf{g}$.

Remark 3 *As we have pointed out in Remark* 2*, the discrete 'controls' $\mathbf{h}_{ni}$ can be poor approximations of the controls $\mathbf{h}(t)$ at $t = t_{ni}$. They are needed to compute good approximations $\mathbf{z}_n$, $n = 1, \ldots, N_t$, of the states and good approximations $\boldsymbol{\pi}_n$, $n = 1, \ldots, N_t$, of the adjoints. After a local minimum of the fully discretized problem* (44)–(46) *has been found and the corresponding states $\mathbf{z}_n^*$, $n = 1, \ldots, N_t$, and adjoints $\boldsymbol{\pi}_n^*$, $n = 1, \ldots, N_t$, have been computed from* (45)*,* (46)*,* (48)*,* (49)*, approximations $\mathbf{h}_n^*$ of the controls $\mathbf{h}^*(t)$ at $t = t_n$ can be computed by solving the minimization problems*

$$\min_{\mathbf{h}_n} (\boldsymbol{\pi}_n^*)^T \mathbf{r}(\mathbf{z}_n^*, \mathbf{h}_n, t_n), \quad n = 0, \ldots, N_t - 1. \qquad (56)$$

Case $\alpha_1 > 0$:
If we set $\boldsymbol{\pi}_n^* = (\theta_n^*, \boldsymbol{\lambda}_n^*, \boldsymbol{\psi}_n^*)^T$ and use the structure of $\mathbf{r}$ in (25)*, then the appropriate approximation of the optimal control $\mathbf{h}^*(t)$ at $t = t_n$ is*

$$\mathbf{h}_n^* = -\frac{1}{\alpha_1 \theta_n^*} \mathbf{L}_1^{-1} \boldsymbol{\psi}_n^*, \quad n = 0, \ldots, N_t - 1. \qquad (57)$$

Case $\alpha_1 = 0$:
We replace $\mathbf{h}$ by $\mathbf{g}$, see Remark 1*, and we set $\boldsymbol{\pi}_n^* = (\theta_n^*, \boldsymbol{\lambda}_n^*)^T$. If l in* (11) *is independent of $\mathbf{g}$ and if $\mathbf{N}(\mathbf{u}, \mathbf{g}, t)$ is linear in $\mathbf{g}$, which is both the case in all our examples, then we can use the structure of $\mathbf{r}$ in* (37) *to solve* (56) *explicitly. In this case the appropriate approximation of the optimal control $\mathbf{g}^*$ at t_n is*

$$\mathbf{g}_n^* = \frac{1}{\alpha_2 \theta_n^*} \mathbf{L}_2^{-1} \mathbf{N}_{\mathbf{g}}(\mathbf{u}_n^*, \mathbf{g}_n, t_n)^T \boldsymbol{\lambda}_n^*, \qquad (58)$$

$n = 0, \ldots, N_t - 1$, where $\mathbf{u}_n^* = \mathbf{u}(\mathbf{q}_n^*)$. Note that because of our linearity assumption, $\mathbf{N}_{\mathbf{g}}(\mathbf{u}, \mathbf{g}, t)$ does not depend on $\mathbf{g}$.

In Ref. 15 it is proven that under suitable assumptions on (23)–(24) *the errors* $\max_{0 \leq n \leq N_t} \|\mathbf{h}_n^* - \mathbf{h}^*(t_n)\|$, $\max_{0 \leq n \leq N_t} \|\mathbf{z}_n^* - \mathbf{z}^*(t_n)\|$, $\max_{0 \leq n \leq N_t} \|\boldsymbol{\pi}_n^* - \boldsymbol{\pi}^*(t_n)\|$ *are of the same order. The order of approximation depends on the order of the Runge–Kutta method* (43) *applied to the* optimal control problem *and the regularity in time of the optimal control of* (23)*,* (24)*. The order conditions for the Runge–Kutta method* (43) *applied to the* optimal control problem *include the order conditions for the Runge–Kutta method* (43) *applied to the state equation* (24)*. However, they also include other conditions that need not be met if the Runge–Kutta method is just applied to solve a single state equation* (24)*. Order conditions for Runge–Kutta methods* (43) *applied to the* optimal control problem *are listed in Ref. 15. They are satisfied by the explicit classical fourth order Runge–Kutta method used in our numerical examples.*

For theoretical and also implementation reasons it is more favorable to transform the adjoint equations (48)–(50) so that the transformed adjoint equations admit the form of Runge–Kutta method. This transformation is well known, see, e.g., the papers 15, 28, 29, and we will review it in the next section. We remark that if one has a Fortran or C/C++ code for the solution of (45), (46) and the subsequent evaluation of the objective (44), and if one uses reverse mode automatic differentiation (AD) (Ref. 14) to compute gradients, then the AD tool will generate and solve (48)–(50), and it will compute the partial derivatives from (51). In our case the time discretization poses few challenges when computing adjoints. We use AD only to compute derivative information at a fixed time t. Specifically, we use reverse mode AD to compute products $\mathbf{r}_{\mathbf{z}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T v$ or $\mathbf{r}_{\mathbf{h}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T v$ for a given vector v (cf. (49), (51)).

Transformation of the Adjoint Equation and the Gradient for the Fully Discretized Problem

The following transformation of the adjoint system (48)–(50) can be found, e.g., in Refs. 15, 28, 29. If we define

$$\widetilde{\boldsymbol{\Pi}}_{ni} = \boldsymbol{\pi}_{n+1} + \sum_{j=1}^{s} \frac{a_{ji}}{b_i} \boldsymbol{\Pi}_{nj}. \qquad (59)$$

for $i, \ldots, s$ and $n = 0, \ldots, N_t - 1$, then the adjoint equations (48)–(50) can be equivalently written as

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_{n+1} + \Delta t_n \sum_{i=1}^{s} b_i \mathbf{r}_{\mathbf{z}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T \widetilde{\boldsymbol{\Pi}}_{ni} \qquad (60)$$

and

$$\widetilde{\boldsymbol{\Pi}}_{ni} = \boldsymbol{\pi}_{n+1} + \Delta t_n \sum_{j=1}^{s} \frac{b_j a_{ji}}{b_i} \mathbf{r}_{\mathbf{z}}(\mathbf{Z}_{nj}, \mathbf{h}_{nj}, t_{nj})^T \widetilde{\boldsymbol{\Pi}}_{nj}, \qquad (61)$$

$i, \ldots, s$, $n = 0, \ldots, N_t - 1$, with

$$\boldsymbol{\pi}_{N_t} = -\nabla_{\mathbf{z}} f(\mathbf{z}_{N_t}), t_{N_t}), \qquad (62)$$

The equations (60)–(62) are called the transformed adjoint equations.

Using the transformed adjoint system the ni-th partial derivative (51) can be written as

$$\frac{d}{d\mathbf{h}_{ni}}\widetilde{J}(\mathbf{h}) = -\mathbf{r}_{\mathbf{h}}(\mathbf{Z}_{ni}, \mathbf{h}_{ni}, t_{ni})^T \Delta t_i b_i \widetilde{\mathbf{\Pi}}_{ni}, \qquad (63)$$

see (59).

For our application the special structure of $\mathbf{r}$ in (25), (37) allows a significant reduction in the transformed adjoint equations. For details see Ref. 18. This is taken advantage of in our implementation.

Optimization Algorithm

Our numerical results are produced using a nonlinear conjugate gradient algorithm (e.g., Ref. 30) for the solution of the discretized problem (44)–(46). However, we use the inner products (52) or (54) instead of the standard Euclidean inner product and our gradients are computed from (53) or (55), respectively. We noticed that even in the case $\alpha_1 > 0$ use of the inner product (54) instead of (52) was more effective in the nonlinear conjugate gradient method. The following computations use (54) and the corresponding gradient (55) in both cases, $\alpha_1 > 0$ and $\alpha_1 = 0$. The inner product acts as a preconditioner within the nonlinear conjugate gradient algorithm. Thus, it effects the convergence behavior of the algorithms, but it does not alter the solution of the problem. The choices (52) or (54) minimize the mesh-dependent behavior of the algorithms and avoid artificial ill-conditioning due to discretization.

In addition to the nonlinear conjugate gradient algorithm, we have also experimented with the limited memory BFGS methods (e.g., Ref. 30), and even extended the latter algorithm to handle bound constraints on the control. Those results will be presented elsewhere.

To cope with the very large size of state and control variables, we apply a simple storage management technique. More sophisticated storage management techniques, which will be considered in the future, are discussed in Refs. 5,31.

All computations are performed in parallel using OpenMP on the SGI Origin 2000. We remark that the solution of one adjoint equation, where spatial adjoints are computed using the AD tool TAMC is about four times as expensive as the solution of one state equation (compressible Navier–Stokes equation).

Results

The optimal control formulation described above is now applied to the vortex-wall interaction problem. The initial flow field is determined by the superposition of two compressible Oseen vortices (Ref. 32). The initial vortex core radii are L^* and the maximum azimuthal velocity $v^*_{\theta_m}$ at the edge of the viscous core of each vortex at the initial time are the length and velocity scales chosen for the nondimensionalization of the Navier–Stokes equations. Thus, the reference Reynolds and Mach numbers are defined as

$$\mathsf{Re} = \frac{\rho^*_\infty L^* v^*_{\theta_m}}{\mu^*_\infty}, \quad \mathsf{M} = \frac{v^*_{\theta_m}}{\sqrt{\gamma R^* T^*_\infty}},$$

where ρ^*_∞, μ^*_∞, and T^*_∞ are the density, viscosity, and temperature far from the vortices; R^* is the gas constant; and "$*$" denotes dimensional quantities.

Our computational domain is $[-15, 15] \times [0, 15]$ in nondimensional units and at time $t = 0$ the vortices are centered at $(\pm 2, 7.5)$. Although each individual Oseen vortex is a solution to the Navier–Stokes equations, the superposed field does not satisfy the equations of motion due to the nonlinearity of the convective term and it also fails to satisfy the wall no-slip and thermal boundary conditions. Hence, the flow is advanced a number of time steps until time t_0 to let the transients associated with these incompatibilities leave the domain through the top and sides where a nonreflecting boundary treatment is used. The resulting field at time t_0 is taken as the initial condition to our problem. The Mach, Reynolds, and Prandtl numbers are $\mathsf{M} = 0.5, \mathsf{Re} = 25, \mathsf{Pr} = 1$, respectively. For the demonstration problems presented here, constant Prandtl number and fluid properties (viscosities and thermal conductivity) are assumed along with Stokes hypothesis for the second coefficient of viscosity, $\lambda = -2\mu/3$. Our formulation also supports variable fluid properties and these effects will be included in future calculations.

In all cases, the control $\mathbf{g}$ is the wall-normal velocity which is a function of the wall coordinate and time. Due to the geometry of our problem $\mathbf{g} = (0, g_2)$ so that positive g_2 represents injection (blowing) of fluid into the domain while negative $\mathbf{g}$ corresponds to suction of fluid out of the domain [cf. equation (3)]. It is assumed that the injected/ingested fluid is always in thermal equilibrium with the solid surface.

Two different control objectives are considered as given by (4) with J_{Obs} defined either as (5), referred to as Terminal Kinetic Energy (TKE) control, or (6) which is denoted as heat transfer control. Before presenting the control results, we first document the uncontrolled dynamics by considering the interaction of the vortex pair with a no-slip, adiabatic wall with inviscid characteristic based far-field boundary treatments used on the top and sides. In the uncontrolled flow, the vortex pair convects downward due to the self induced velocity eventually interacting with the wall at the lower boundary. Since the Reynolds number is rather low, the vortices decay as shown in Fig. 1 by the evolution of integrated kinetic energy $\int_\Omega \rho \mathbf{v}^T \mathbf{v}/2 \, d\Omega$ and integrated enstrophy $\int_\Omega \boldsymbol{\omega}^T \boldsymbol{\omega}/2 \, d\Omega$, where $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ is the vorticity vector. The integrated kinetic energy undergoes a monotonic decay as dissipation converts kinetic energy into internal energy. While the integrated enstrophy also generally decays, when the vortices approach the wall, secondary vorticity is generated through the no-slip condition leading to a local maximum in integrated enstrophy at approximately $t = 22$. In all controlled runs, the integrated enstrophy peak of the uncontrolled flow lies well within the optimization time window. The next two sections present detailed results for the TKE control and heat transfer control flows.

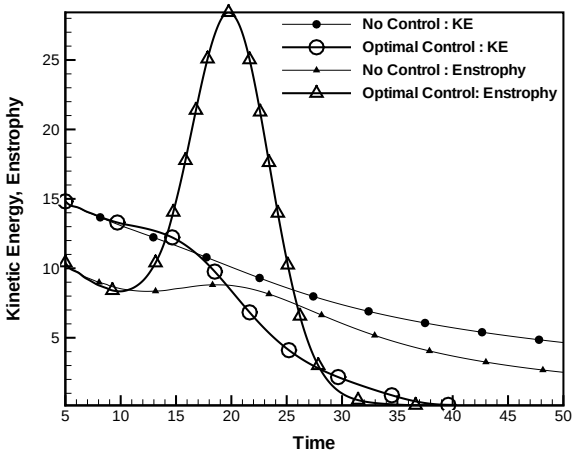


Fig. 1 The evolution of integrated kinetic energy and enstrophy with time for uncontrolled and TKE control, run (a).

Test Case 1: Optimal Control of Terminal Kinetic Energy

The objective functional for this test case is given by (4) with (5) and $\Omega_0 = [-15, 15] \times [0, 15]$ (the entire computational domain). The regularization term in (4) is given by (9). The computations are performed on a 128×128 uniform mesh with a fixed time step $\Delta t = 0.05$. The optimization time window is 700 time steps from $t_0 = 5$ to $t_f = 40$ and the enstrophy peak of the uncontrolled flow lies well within this time window, see Figure 1. The lower wall is adiabatic while an inviscid characteristic based farfield boundary condition is imposed on the other boundaries. The injected/ingested fluid is at the local wall temperature.

We performed three runs – two include a regularization of the time derivative of the control, the third does not. In all runs the control space (7) and the regularization term (9) is used, i.e., $\alpha_4 = 0$ throughout. The coefficients α_j in the regularization term for the two runs are:

Run	α_1	α_2	α_3	α_4
a	0.5	0.005	0.005	0
b	0.05	0.005	0.005	0
c	0	0.005	0.005	0

In all cases the optimization is started with zero control. The optimal wall-normal velocity distributions g_2 for runs (a)–(b) are plotted in Figures 2(a) to 2(c), respectively. These plots clearly show the effect of the regularization term $\int \frac{\alpha_1}{2} \|\mathbf{g}_t\|_2^2$. Without it, the control starts to oscillate in time in the second half of $[t_0, t_f]$ and also exhibits a large jump at t_0. As mentioned in the 'Problem' section, setting $\alpha_1 = \alpha_2 = 0$ produces strong spatial and temporal oscillations in the control which frequently led to a failure in the compressible Navier–Stokes solver. Such jumps and oscillations are also physically undesirable since they would place severe demands on actuator performance.

There are only slight differences between the controls for runs (a) and (b) which both include time regularization. In both of these runs, the control is localized in the center of the domain near the region of vortex/wall interaction and consists primarily of suction although there are small amounts of blowing visible at late times near $x = \pm 7$. The main effect of the control is to absorb the vortices through the wall as described in more detail below.

The following table shows the values of the objective functional (4) with (5) at the initial iterate, i.e., zero control (J_0); at the final control iterate (J_{final}); and the integrated terminal kinetic energy (5) at the final control iterate (TKE_{final}). We see that because of the large regularization parameter α_1, the terminal kinetic energy reduction in run (a) is less than that for run (b) although the smaller $\alpha_1 > 0$ of run (b) is still sufficient to maintain temporal smoothness in the controls. However, both run (a) and (b) achieve a dramatic reduction of TKE of approximately 97%. Although run (c), without time regularization, leads to slightly greater reduction in TKE_{final} (98%), the added benefit is more than offset by the physically and numerically undesirable control distributions.

Run	J_0	J_{final}	TKE_{final}
a	12.43	0.48	0.42
b	12.43	0.37	0.33
c	12.43	0.24	0.20

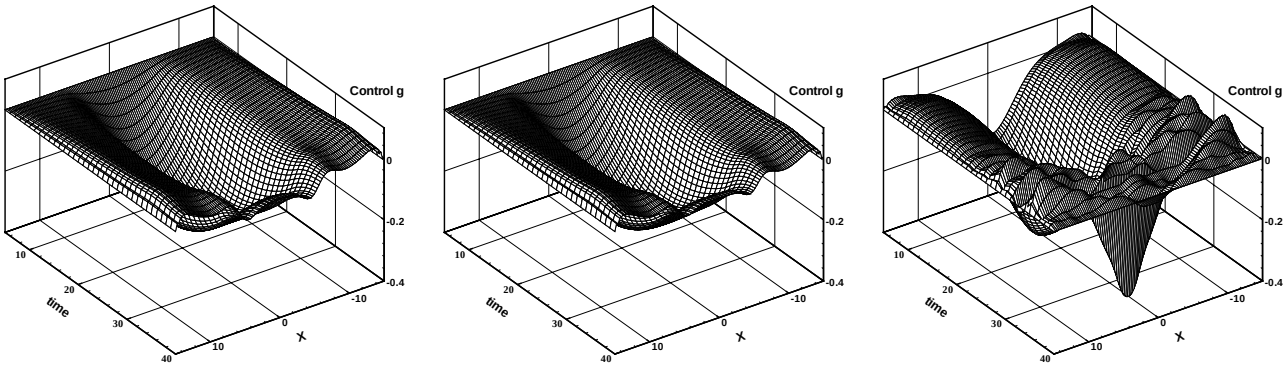
Figure 1 illustrates the effect of the applied control on the temporal evolution of integrated kinetic energy and enstrophy. The control initially introduces a large amount of vorticity near the wall which increases the integrated enstrophy peak almost three fold compared to the uncontrolled flow. This is accompanied by a slight decrease in the kinetic energy decay rate as energy is added to the flow via the control input. As $t \to t_f$, the rate of decrease of both kinetic energy and enstrophy increases considerably and both quantities approach their terminal values smoothly.

Figure 3 shows contours of kinetic energy for the uncontrolled flow and the TKE controlled flow for run (a). In the uncontrolled flow, the vortices propagate toward the wall and, upon interacting with the no-slip wall, eventually reach an quasi equilibrium position (see Fig. 3(c)) where they slowly decay due to viscous diffusion. The effect of the optimal control is to almost completely absorb the vortices through the wall so that the kinetic energy at the final time is reduced by two orders of magnitude.

Test Case 2: Optimal Control of Heat Transfer

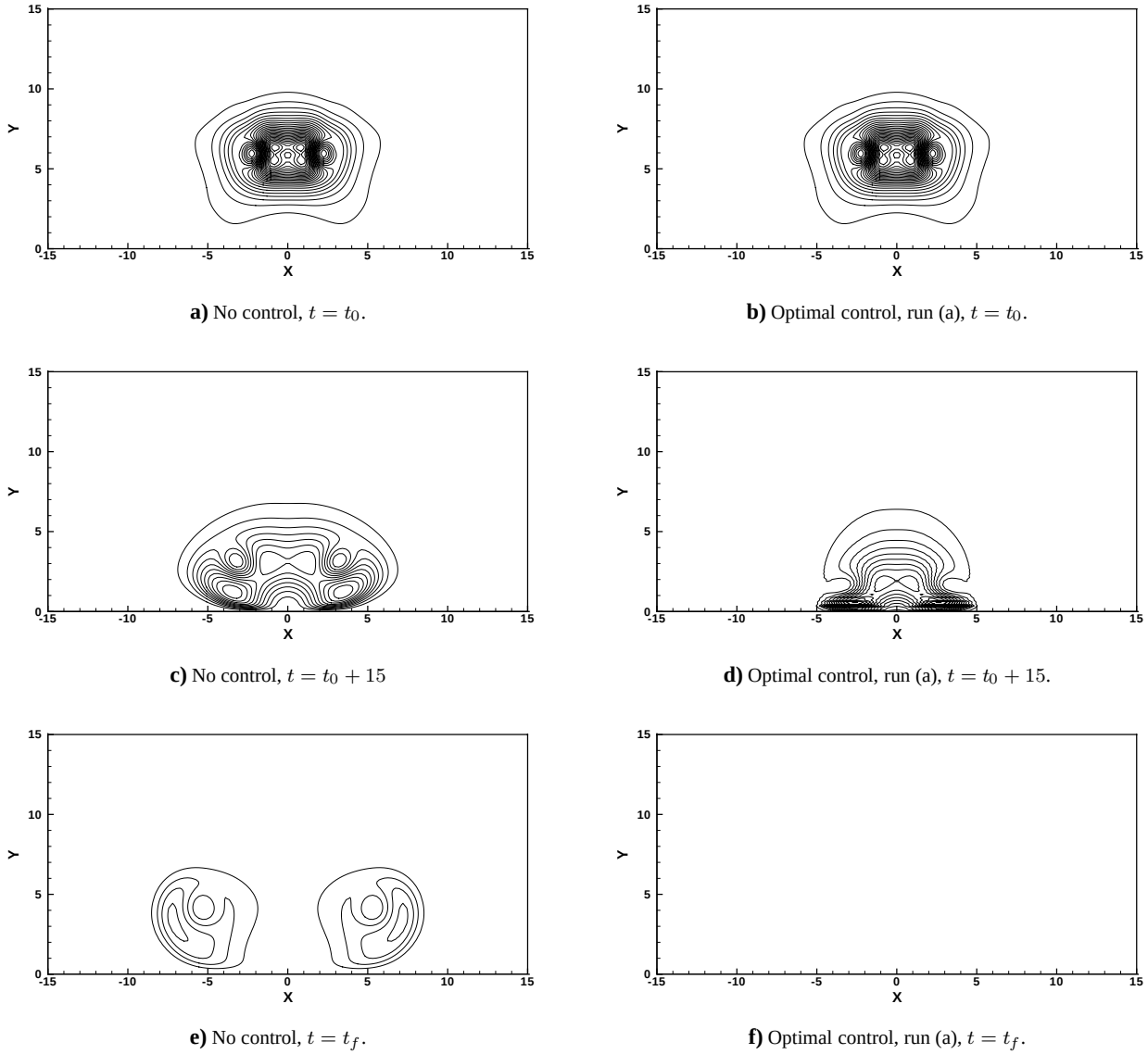
The motivation for this test case is to investigate the possibility of controlling the heat transfer across an isothermal solid wall in a time interval of interest by applying unsteady wall-normal velocity control. This test problem demonstrates how the coupling of thermodynamic and kinematic flow variables through the compressible Navier–Stokes equations can be used to effectively control one with the other.

The objective functional is now given by (4) with (6). The regularization term in (4) is given by (9) or by (10). Suction/blowing is applied only on the portion $\Gamma_c = \{\mathbf{x} = (x_1, 0) : x_1 \in [-11.25, 11.25]\}$ of the lower wall $\Gamma = \{\mathbf{x} = (x_1, 0) : x_1 \in [-15, 15]\}$. Moreover, the normal derivative is only controlled over the same part of the



a) $\alpha_1 = 0.5$, $\alpha_2 = \alpha_3 = 0.005$, $\alpha_4 = 0.$ **b)** $\alpha_1 = 0.05$, $\alpha_2 = \alpha_3 = 0.005$, $\alpha_4 = 0.$ **c)** $\alpha_1 = 0$, $\alpha_2 = \alpha_3 = 0.005$, $\alpha_4 = 0.$

Fig. 2 Optimal wall-normal velocity distributions for TKE control, runs (a)-(c).



a) No control, $t = t_0$.

b) Optimal control, run (a), $t = t_0$.

c) No control, $t = t_0 + 15$

d) Optimal control, run (a), $t = t_0 + 15$.

e) No control, $t = t_f$.

f) Optimal control, run (a), $t = t_f$.

Fig. 3 Contours of kinetic energy for uncontrolled and TKE controlled flow, run (a). There are 25 evenly spaced contours from 0.0 to 1.6.

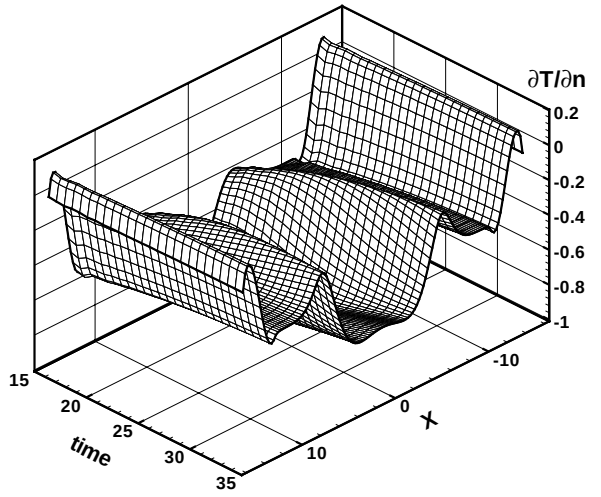


Fig. 4 Distribution of wall normal temperature gradient, $\partial T/\partial n$, at the wall for the uncontrolled flow.

lower boundary, i.e., $\Gamma_0 = \Gamma_c$ in (6). The weighting function ω is a smooth function which is zero at $x_1 = \pm 11.25$ and one on $[-9, 9]$. The term $\omega \partial T/\partial n$ in the objective contributes to adjoint boundary conditions on $\Gamma_0 \subset \Gamma$, but not outside Γ_0. If the weighting were $\omega \equiv 1$, then the adjoint boundary conditions can and usually do exhibit jumps at $\mathbf{x} = (\pm 11.25, 0)$. The addition of the smooth weighting function with $\omega(\pm 11.25) = 0$ avoids these jumps in the adjoint boundary conditions. For details we refer to Ref. 18.

The lower wall is isothermal while an inviscid characteristic based far-field boundary condition is imposed on the other boundaries. The temperature boundary condition at the lower wall is $T(\mathbf{x}, t) = T_b(\mathbf{x})$, $\mathbf{x} \in \Gamma_c$, where T_b is a continuously differentiable function with $T_b(\mathbf{x}) = 1.5$, $\mathbf{x} \in \Gamma_c$ and $T_b(\mathbf{x}) = 2$, $\mathbf{x} = (\pm 15, 0)$. This temperature profile ensures compatibility with the far-field temperature at $(\pm 15, 0)$.

In all computations for this test case, we use a mesh with 128 grid points in the x- and y-directions. A uniform mesh is used in the x-direction while the mesh is clustered in the y-direction near the wall using an algebraic stretching function in order to better capture the wall-normal temperature gradient. The resulting mesh size Δy at the wall is 0.37 times that of a uniform mesh, i.e., $\Delta y = 0.37 * 15/128$. The time step is $\Delta t = 0.025$. To compute the initial conditions, the flow equations are advanced for 600 time steps to minimize effects resulting from the startup transient. The optimal controls are computed over 800 time steps, i.e., $t_0 = 15$ to $t_f = 35$ and the enstrophy peak of the uncontrolled flow lies within this time window, see Figure 1.

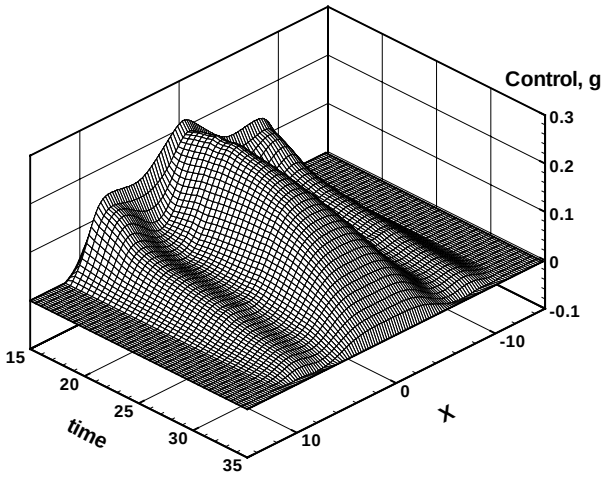
In our first four runs we use the control space (7) and the regularization term (9). The last two runs use the control space (8) and the regularization term (10) The first two runs do not include a regularization of the time derivative of the control, the other four do. The coefficients α_j in the regularization term for the runs are

Run	α_1	α_2	α_3	α_4
a	0	0.5	0.5	0
b	0	0.05	0.05	0
c	0.5	0.5	0.5	0
d	0.5	0.05	0.5	0
e	0.5	0.5	0.5	0.05
f	0.5	0.05	0.5	0.05

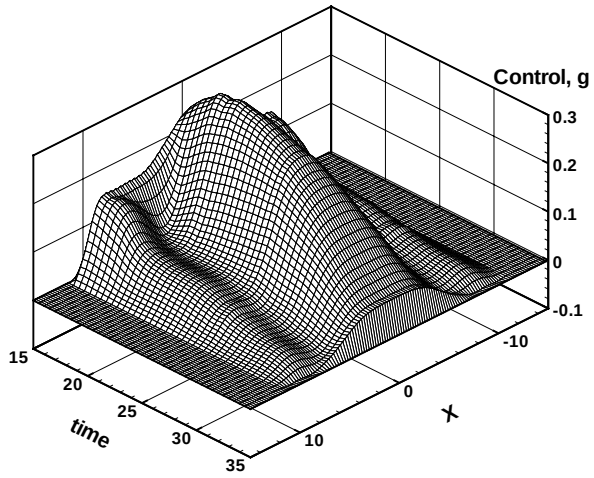
The uncontrolled flow is given in Figure 4 which shows the wall-normal temperature gradient $\partial T/\partial n = -\partial T/\partial x_2$ on the bottom wall that is proportional to the heat transfer from the fluid to the wall, $\dot{q} = -\kappa \partial T/\partial n$. Since $\mathbf{n}$ is the outward unit normal, $\partial T/\partial n > 0$ denotes heat transfer from the wall to the fluid. Since the uncontrolled wall is cool, the nominal heat transfer is from the flow to the wall as indicated in figure 4 by the predominately negative $\partial T/\partial n$. Since the vortex cores are lower in temperature than the ambient fluid, there is a local reduction in heat transfer near the vortex core locations which can be easily tracked in Figure 4.

Heat transfer results for the controlled flows are shown in Figure 6. Keep in mind that we only control $\frac{\partial T}{\partial n}$ over $\Gamma_c = \{\mathbf{x} = (x_1, 0) : x_1 \in [-11.25, 11.25]\}$. In this region of the wall, $(\frac{\partial T}{\partial n})^2$ is reduced nicely by all controls. Only for runs (a) and (b), where no time-derivative regularization term ($\alpha_1 = 0$) is enforced, $(\frac{\partial T}{\partial n})^2$ is large for a small time near t_0. Actually in that time region for runs (a) and (b), the controlled $(\frac{\partial T}{\partial n})^2$ is larger than in the no-control flow. Nevertheless, the control substantially reduces the integral (6) in all cases (see below). These sudden changes in the wall-normal temperature gradient for the runs without time regularization are directly related to sudden changes in the controls for these flows.

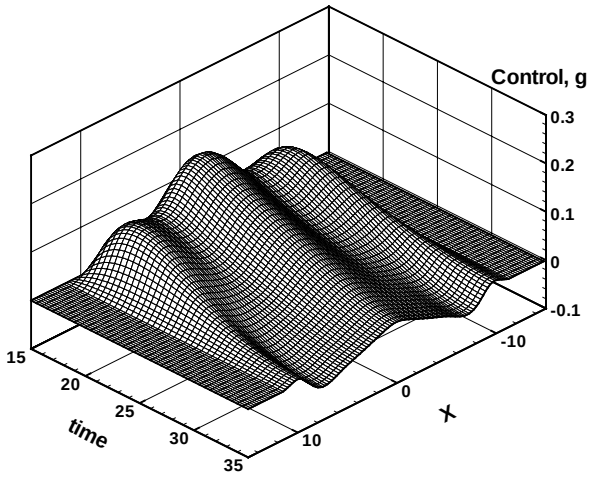
The computed controls are plots (a)-(d) in Figure 5. Since control is only applied over Γ_c the wall-normal velocity is zero elsewhere. We see that the structure of the optimal controls computed using a time derivative regularization ($\alpha_1 > 0$) is fundamentally different from those computed without time derivative regularization ($\alpha_1 = 0$). This follows already from the optimality conditions $\nabla J(\mathbf{g}) = 0$ using (18)–(20) for the case $\alpha_1 > 0$ and (20), (36) for the case $\alpha_1 = 0$. The computed controls for $\alpha_1 = 0$ change rapidly in time near t_0 and near t_f and they exhibit slight oscillations elsewhere in time. This can be better seen in Figure 7, where a slice of the controls at $x = 0$ is plotted. As in the TKE control case, the controls without time regularization can exhibit large temporal derivatives which may not only cause numerical difficulties but which would also place undue demands on the design of physical actuators for problems of this type. A regularization of the time derivative of the controls enforces temporal smoothness. However, close expection of the optimal controls computed with time regularization in runs (c) and (d) reveal slight spatial oscillations in these optimal controls. See plots (c) and (d) in Figure 5. Moreover, all optimal controls may increase sharply near the endpoints $x_1 = \pm 11.25$ of the control boundary Γ_c. To supress oscillations and to avoid a sharp increases in the horizontal velocity near the bound-



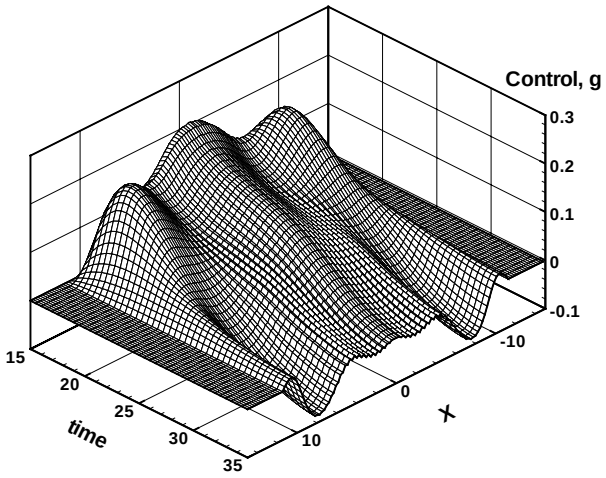
a) $\alpha_1 = 0, \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0$



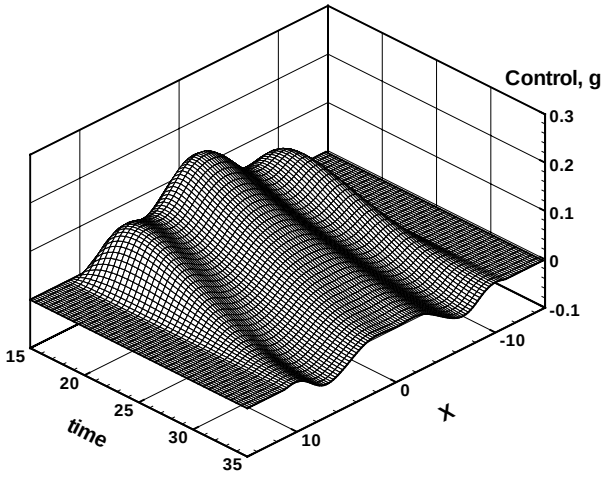
b) $\alpha_1 = 0, \alpha_2 = \alpha_3 = 0.05, \alpha_4 = 0$



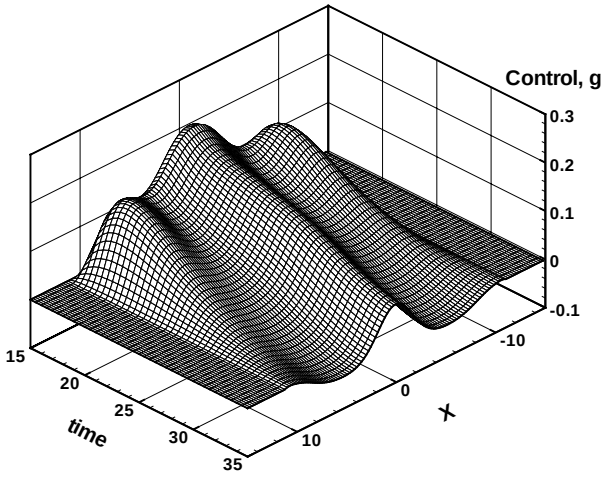
c) $\alpha_1 = \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0$



d) $\alpha_1 = \alpha_3 = 0.5, \alpha_2 = 0.05, \alpha_4 = 0$

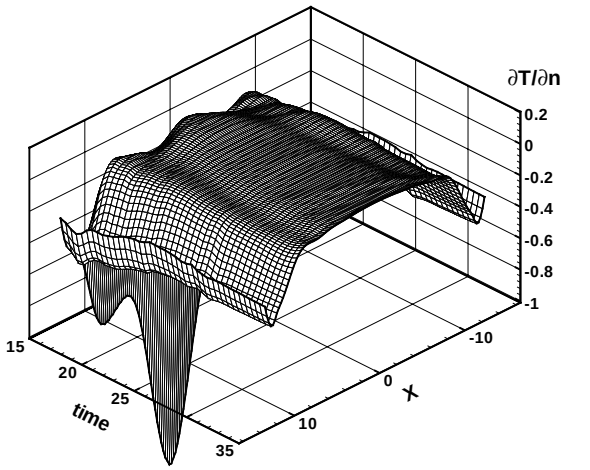


e) $\alpha_1 = \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0.05$

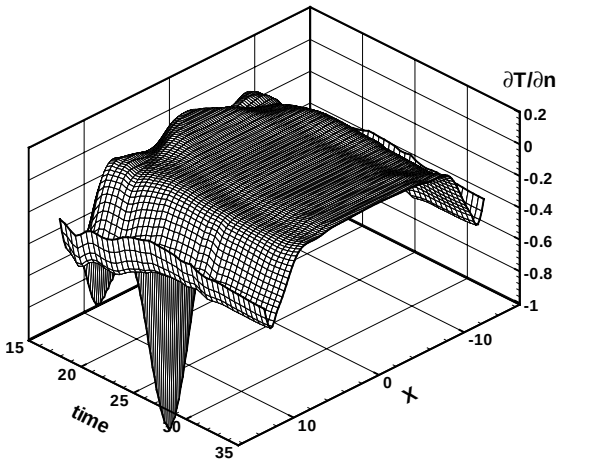


f) $\alpha_1 = \alpha_3 = 0.5, \alpha_2 = 0.05, \alpha_4 = 0.05$

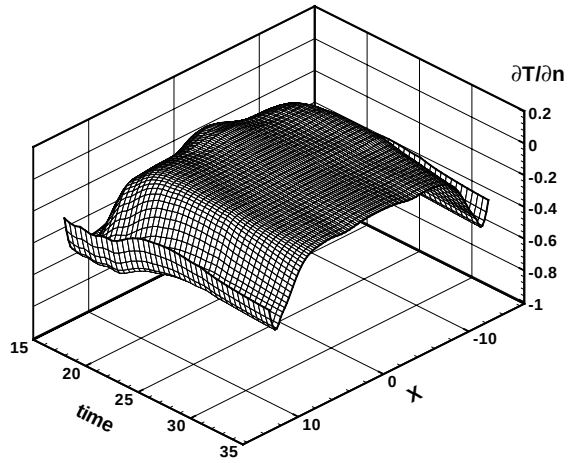
Fig. 5 Optimal control distributions for heat transfer control.



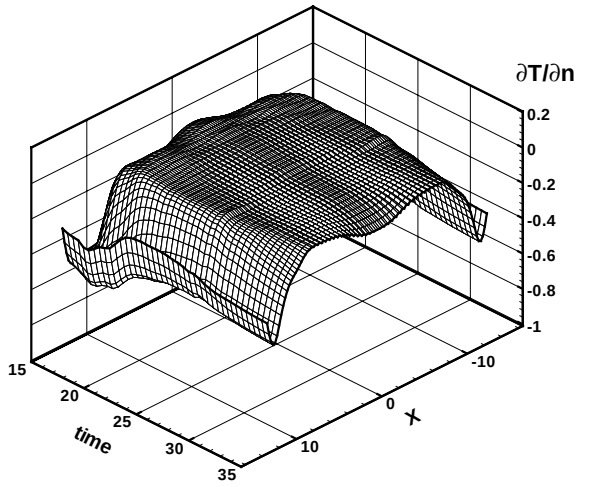
a) $\alpha_1 = 0, \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0$



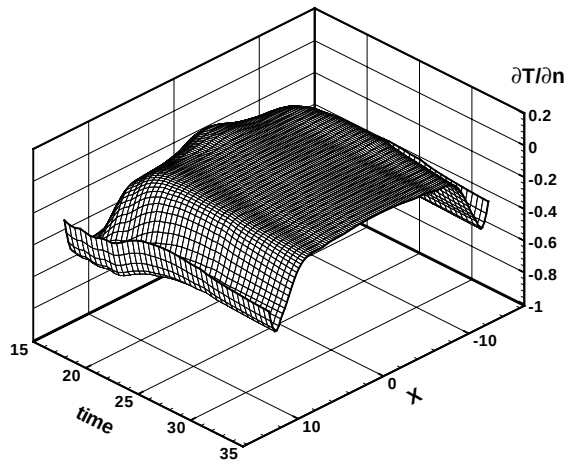
b) $\alpha_1 = 0, \alpha_2 = \alpha_3 = 0.05, \alpha_4 = 0$



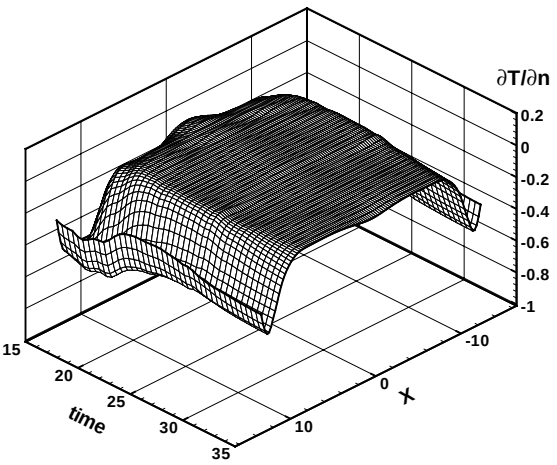
c) $\alpha_1 = \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0$



d) $\alpha_1 = \alpha_3 = 0.5, \alpha_2 = 0.05, \alpha_4 = 0$



e) $\alpha_1 = \alpha_2 = \alpha_3 = 0.5, \alpha_4 = 0.05$



f) $\alpha_1 = \alpha_3 = 0.5, \alpha_2 = 0.05, \alpha_4 = 0.05$

Fig. 6 Distribution of wall normal temperature gradient, $\partial T/\partial n$, at the wall for flows with heat transfer control.

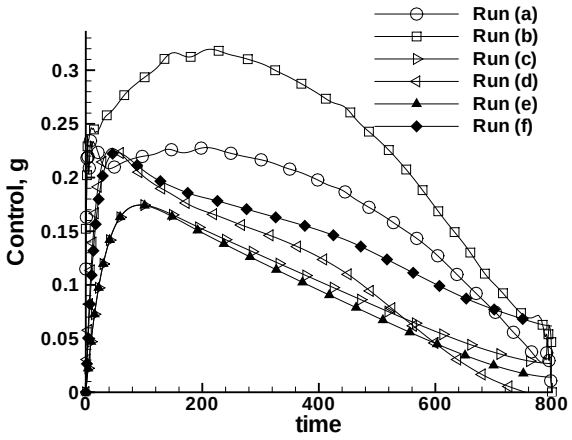


Fig. 7 Comparison of optimal controls at $x = 0$ for heat transfer control.

ary points $(x_1, x_2) = \pm(11.25, 0)$ we switch to the control space (8) and the regularization term (10). The regularization parameters for the two additional runs (e) and (f) are listed in the table above. The optimal controls for these runs are plots (e) and (f) in Figure 5. We see that spatial oscillations are suppressed and the spatial transition from the no-control to the control boundary is smoother than in runs (c) and (d). The additional control regularization has little effect on the total heat transfer as can be seen in Figure 6.

The following tables show the values of the objective functional (4) with (6) at the initial iterate, i.e., zero control (J_0) and at the final control iterate, (J_{final}), in addition to the values of the (6) and those for the individual regularization terms at the final control.

Run	J_0	J_{final}	$\frac{1}{2}\int \omega(\frac{\partial T}{\partial \mathbf{n}})^2$
a	53.970	3.650	2.560
b	53.970	2.807	2.609
c	53.970	4.781	4.159
d	53.970	3.435	3.055
e	53.970	4.844	4.236
f	53.970	3.656	3.354

Run	$\frac{\alpha_1}{2}\int \|\mathbf{g}_t\|_2^2$	$\frac{\alpha_2}{2}\int \|\mathbf{g}\|_2^2$	$\frac{\alpha_3}{2}\int \|\nabla \mathbf{g}\|_2^2$	$\frac{\alpha_4}{2}\int \|\Delta \mathbf{g}\|_2^2$
a	-	1.034	0.057	-
b	-	0.188	0.011	-
c	0.054	0.537	0.031	-
d	0.202	0.101	0.077	-
e	0.053	0.527	0.028	0.002
f	0.156	0.089	0.056	0.004

Overall, the RMS heat transfer is typically reduced by a factor of 3.5 or approximately 93% for the controlled flows. The slight additional reduction in heat transfer for the flows without time regularization is likely due to the fact that these flows have large controls near t_0, however the sudden jump in control at t_0 not only causes oscillations in the numerical solutions, but would also require unrealistically high actuator frequency response.

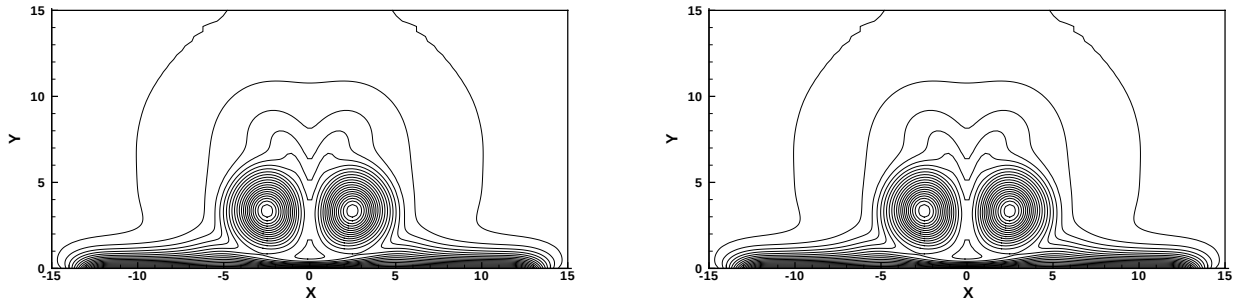
Temperature contours for the uncontrolled flow and controlled flow, run (c), are plotted in Figure 8. Figure 5 shows that in all cases fluid is primarily blown into the domain such that the optimal control takes the form of film cooling. Since the fluid blown into the domain is at the wall temperature $T = 1.5$, this creates a small layer of fluid with almost constant temperature $T = 1.5$ near Γ_c, see Figure 8 d–e, thereby reducing $\frac{\partial T}{\partial n}$. The temperature contour plots for the controlled flows in runs (b) to (f) are very similar to those shown in Figure 8 d–f for run (a) and are not given here.

Conclusions

This work, represents to the Authors' best knowledge, the first application of optimal control theory to problems governed by the unsteady, compressible Navier–Stokes equations. This research is motivated by the potential to develop novel and effective flow control strategies for inherently compressible phenomena including aeroacoustics and heat transfer by utilizing optimal control theory. We began by describing a detailed problem formulation for two-dimensional unsteady problems using wall-normal velocity as control with observation of either terminal kinetic energy or heat transfer at the wall and this formulation can be readily extended to more general configurations. An important aspect of this research involves the role of control regularization. Motivated by the mathematical theory for the boundary control incompressible flows we consider not only regularization of control magnitude but also regularization of spatial and temporal derivatives. In formulating the optimal control problem, a novel approach is utilized where the time regularization is incorporated by augmenting the state and control variables. The fully discrete optimal control problem is solved using a nonlinear conjugate gradient algorithm where the gradient is obtained via the adjoint equations. Adjoints at a fixed time t are computed using automatic differentiation. The optimal control formulation is demonstrated for a simple wall/vortex interaction problem and results are presented for both control of terminal kinetic energy as well as control of heat transfer. In both controlled problems, regularization of spatial and temporal derivatives is required to ensure physically realistic control distributions and with sufficient regularization, realizable optimal controls were found that lead to dramatic improvements in the control objectives. This work represents our initial foray into optimal control of unsteady compressible flows and the formulation and simulation software described here are now available to study a variety of problems ranging from fundamental issues in optimization to flow control applications.

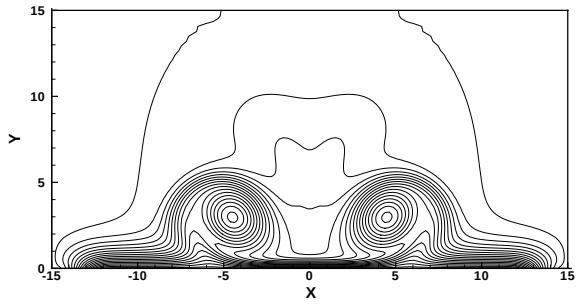
Acknowledgments

The work of S. Collis, K. Ghayour and M. Heinkenschloss was supported by Texas ATP grant 003604–0001, 1999. M. Ulbrich and S. Ulbrich were supported by DFG grants UL157/3-1 and UL158/2-1, respectively, and by CRPC grant CCR-9120008. Computations were performed on an SGI Origin 2000 which was purchased with the aid of NSF SCREMS grant 98–72009.

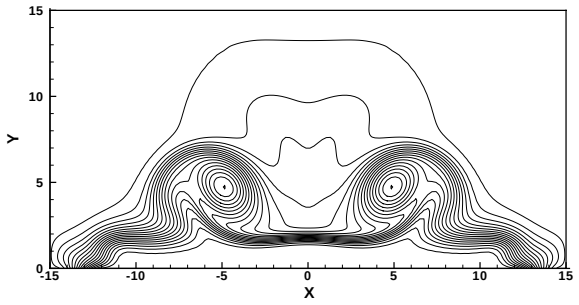


a) No control, $t = t_0$.

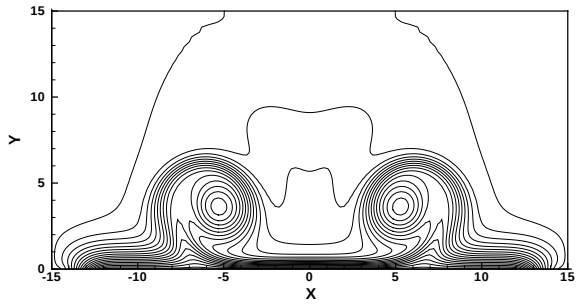
b) Optimal control, run (a), $t = t_0$.



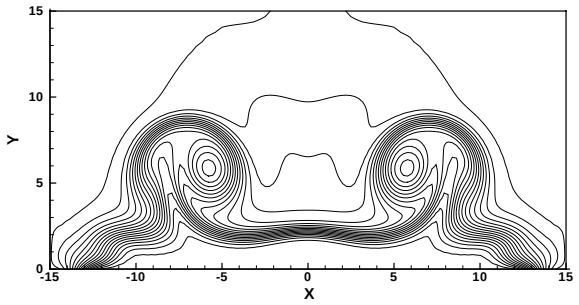
c) No control, $t = t_0 + 12.5$



d) Optimal control, run (a), $t = t_0 + 12.5$.



e) No control, $t = t_f$.



f) Optimal control, run (a), $t = t_f$.

Fig. 8 Temperature contours for the uncontrolled flow and heat transfer control flow, run (a). There are 25 evenly spaced contours from 1.4 to 2.1.

References

[1]Lee, S., "Reduction of Blade-Vortex Interaction Noise Through Porous Leading Edge," *AIAA Journal*, Vol. 32, No. 3, 1994, pp. 480–488.

[2]Baeder, J. D., "Passive Design for Isolated Blade-Vortex Interaction Noise Reduction," *Proceedings of the 53rd Annual Forum of the American Helicopter Society, Part 1*, 1997.

[3]Wang, B., Baeder, J. D., and Singh, R. K., "A Computational Study of Trailing-Edge Flap Aerodynamics and Acoustics," *55th Annual Forum of the American Helicopter Society*, May 1999.

[4]Collis, S. S. and Lele, S. K., "A Computational Investigation of Receptivity in High-Speed Flow Near a Swept Leading-Edge," Tech. Rep. TF-71, Flow Physics and Computation Division, Department of Mechanical Engineering, Stanford University, April 1997.

[5]Berggren, M., "Numerical Solution of a Flow-Control Problem: Vorticity Reduction by Dynamic Boundary Action," *SIAM J. Scientific Computing*, Vol. 19, 1998, pp. 829–860.

[6]Bewley, T. and Moin, P., "Optimal Control of Turbulent Channel Flows," *Active Control of Vibration and Noise, ASME-DE*, Vol. 75, 1994.

[7]Bewley, T. R., Moin, P., and Teman, R., "DNS-based predictive control of turbulence: an optimal target for feedback algorithms," *Submitted to J. Fluid Mech.*, 2000.

[8]Collis, S. S., Chang, Y., Kellogg, S., and Prabhu, R. D., "Large Eddy Simulation and Turbulence Control," *AIAA 2000-2564*, 2000.

[9]Collis, S. S. and Chang, Y., "On the Use of LES with a Dynamic Subgrid Scale Model for the Optimal Control of Wall Bounded Turbulence," *Recent Advances in DNS and LES*, edited by D. Knight and L. Sakell, Kluwer Academic Publishers, 1999, pp. 99–110.

[10]Chang, Y., *Reduced Order Methods for Optimal Control of Turbulence*, Ph.D. thesis, Rice University, 2000.

[11]Gunzburger, M. D. and Manservisi, S., "The velocity tracking problem for Navier–Stokes flows with boundary control," *SIAM J. Control Optimization*, Vol. 39, 2000, pp. 594–634.

[12]Soemarwoto, B., "The Variational Method for Aerodynamic Optimization Using the Navier-Stokes Equations," Tech. Rep. 97-71, ICASE, December 1997.

[13]Jameson, A., Martinelli, L., and Pierce, N. A., "Optimum Aerodynamic Design using the Navier–Stokes equations," *Theor. Comp. Fluid Dyn.*, Vol. 10, No. 1–4, 1998, pp. 213–237.

[14]"Argonne National Laboratory Computational Differentiation Web Pages," http://www.mcs.anl.gov/autodiff/.

[15]Hager, W. W., "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Math-*

ematik, 2000, To appear. Electronic version is available at http://www.math.ufl.edu/ hager/papers/rk.ps.

[16]Abergel, F. and Temam, R., "On some Control Problems in Fluid Mechanics," *Theoretical and Computational Fluid Dynamics*, Vol. 1, 1990, pp. 303–325.

[17]Gunzburger, M. D. and Hou, L. S., "Finite dimensional approximation of a class of constrained nonlinear optimal control problems," *SIAM J. Control Optimization*, Vol. 34, 1996, pp. 1001–1043.

[18]Collis, S. S., Ghayour, K., Heinkenschloss, M., Ulbrich, M., and Ulbrich, S., "Optimal Control of the Compressible Navier–Stokes Equations using Adjoint Based Methods," Tech. Rep. in preparation, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892, 2000.

[19]Giering, R. and Kaminski, T., "Recipes for Adjoint Code Construction," *ACM Transactions on Mathematical Software*, Vol. 24, 1998, pp. 437–474.

[20]Fursikov, A. V., *Optimal Control of Distributed Systems. Theory and Applications*, Translation of Mathematical Monographs, Vol. 187, American Mathematical Society, Providence, Rhode Island, 2000.

[21]Hoff, D., "Discontinuous Solutions of the Navier–Stokes Equations for Multidimensional Flows of Heat–Conducting Fluids," *Arch. Rational Mech. Anal.*, Vol. 139, 1997, pp. 303–354.

[22]Lions, P. L., *Mathematical Topics in Fluid Mechanics. Volume 2, Compressible Models*, Oxford Lecture Series in Mathematics and its Applications, Vol. 10, Claredon Press, Oxford, 1998.

[23]Valli, A., "Mathematical Results for Compressible Flows," *Mathematical Topics in Fluid Mechanics*, edited by J. F. Rodriguez and A. Sequeira, Pitman Research Notes Mathematics, Vol. 274, Longman Scientific and Technical, Essex, 1992, pp. 193–229.

[24]Fursikov, A. V., Gunzburger, M. D., and Hou, L. S., "Boundary value Problems and Optimal Boundary Control for the Navier–Stokes Systems: The Two–Dimensional Case," *SIAM J. Control Optimization*, Vol. 36, 1998, pp. 852–894.

[25]Renardy, M. and Rogers, R. C., *An Introduction to Partial Differential Equations*, Texts in Applied Mathematics, Vol. 13, Springer Verlag, Berlin, Heidelberg, New-York, 1993.

[26]Carpenter, M. H., Gottlieb, D., and Abarbanel, S., "Stable and Accurate Boundary Treatments for Compact, High-Order Finite Difference Schemes," *Appl. Num. Math.*, Vol. 12, No. 1–3, 1993, pp. 55–87.

[27]Heinkenschloss, M. and Vicente, L. N., "An Interface Between Optimization and Application for the Numerical Solution of Optimal Control Problems," *ACM Transactions on Mathematical Software*, Vol. 25, 1999, pp. 157–190.

[28]Hager, W. W., "Rates of Convergence for Discrete Approximations to Unconstrained Control Problems," *SIAM J. Numer. Analysis*, Vol. 13, 1976, pp. 449–472.

[29]Bock, H. G., "Randwertprobleme zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen," Preprint Nr. 442, Universität Heidelberg, Institut für Angewandte Mathematik, SFB 123, D–6900 Heidelberg, Germany, 1988.

[30]Nocedal, J. and Wright, S. J., *Numerical Optimization*, Springer Verlag, Berlin, Heidelberg, New York, 1999.

[31]Restrepo, J. M., Leaf, G. K., and Griewank, A., "Circumventing Storage Limitations in Variational Data Assimilation Studies," *SIAM J. Scientific Computing*, Vol. 19, 1998, pp. 1586–1605.

[32]Colonius, T., Lele, S. K., and Moin, P., "The Free Compressible Viscous Vortex," *J. Fluid Mech.*, Vol. 230, 1991, pp. 45–73.