

A practical discrete-adjoint method for high-fidelity compressible turbulence simulations



Ramanathan Vishnampet^a, Daniel J. Bodony^b, Jonathan B. Freund^{a,b,*}

^a Department of Mechanical Science and Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA

^b Department of Aerospace Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA

ARTICLE INFO

Article history:

Received 1 July 2014

Received in revised form 27 December 2014

Accepted 6 January 2015

Available online 14 January 2015

Keywords:

Adjoint methods

Compressible turbulence

Aeroacoustic optimization

High-order methods

ABSTRACT

Methods and computing hardware advances have enabled accurate predictions of complex compressible turbulence phenomena, such as the generation of jet noise that motivates the present effort. However, limited understanding of underlying physical mechanisms restricts the utility of such predictions since they do not, by themselves, indicate a route to design improvements. Gradient-based optimization using adjoints can circumvent the flow complexity to guide designs, though this is predicated on the availability of a sufficiently accurate solution of the forward and adjoint systems. These are challenging to obtain, since both the chaotic character of the turbulence and the typical use of discretizations near their resolution limits in order to efficiently represent its smaller scales will amplify any approximation errors made in the adjoint formulation. Formulating a practical exact adjoint that avoids such errors is especially challenging if it is to be compatible with state-of-the-art simulation methods used for the turbulent flow itself. Automatic differentiation (AD) can provide code to calculate a nominally exact adjoint, but existing general-purpose AD codes are inefficient to the point of being prohibitive for large-scale turbulence simulations. Here, we analyze the compressible flow equations as discretized using the same high-order workhorse methods used for many high-fidelity compressible turbulence simulations, and formulate a practical space–time discrete-adjoint method without changing the basic discretization. A key step is the definition of a particular discrete analog of the continuous norm that defines our cost functional; our selection leads directly to an efficient Runge–Kutta-like scheme, though it would be just first-order accurate if used outside the adjoint formulation for time integration, with finite-difference spatial operators for the adjoint system. Its computational cost only modestly exceeds that of the flow equations. We confirm that its accuracy is limited by computing precision, and we demonstrate it on the aeroacoustic control of a mixing layer with a challengingly broad range of turbulence scales. For comparison, the error from a corresponding discretization of the continuous-adjoint equations is quantified to potentially explain its limited success in past efforts to control jet noise. The differences are illuminating: the continuous-adjoint is shown to suffer from exponential error growth in (reverse) time even for the best-resolved largest turbulence scales. Implications for jet noise reduction and turbulence control in general are discussed.

© 2015 Elsevier Inc. All rights reserved.

* Corresponding author at: Department of Mechanical Science and Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA.
E-mail address: jbfreund@illinois.edu (J.B. Freund).

1. Introduction

Adjoint-based methods are attractive and have been used for several purposes in conjunction with simulations of fluid flow, including optimal flow control [1], aerodynamic shape design [2,3], goal-oriented mesh refinement [4], sensitivity-based uncertainty quantification [5], and stability and global mode analysis [6,7]. Our specific efforts are currently directed towards control optimization for turbulent aeroacoustic flows [8], and our demonstration simulations are motivated by this application, which we discuss in some detail. The present work develops an exact space-time adjoint formulation that is practical in that it incurs a computational cost comparable to the flow equations or a direct discretization of their continuous-adjoint. It is also compatible with high-order finite-difference schemes that are the “workhorse” methods for a wide range of compressible flow simulations, including turbulent free-shear flows [9–11], boundary-layer flows [12], and combustion [13]. We show that an exact adjoint is surprisingly important for accurately predicting the sensitivity at all turbulence scales, and that our formulation is computationally efficient and extensible.

The control of sound generation by turbulence provides a particularly challenging target application due to its complexity and the relative subtlety of its noise-generation mechanisms. In aeroacoustic simulations, the computational challenge of representing a range of turbulence scales and the relatively low-amplitude fluctuations is well documented [14,15]. Without a reduced model of noise generation mechanisms to provide guidance, predictions by themselves do not indicate routes to design improvement. Adjoint-based optimization methods used in conjunction with predictive simulations have been proposed and demonstrated to circumvent this complexity by providing gradient information that can be harnessed to achieve noise reductions [8]. In essence, the adjoint carries sensitivity information with respect to an arbitrarily large number of control parameters, which enables the optimization of noise-controlling actuation or geometries. Building on efforts in aerodynamic optimization [16] and incompressible turbulence control [1], Wei and Freund [17] achieved 11 dB noise reduction of far-field sound with internal energy control of a direct numerical simulation (DNS) of a two-dimensional compressible shear layer. Though the method has been demonstrated to also work in genuinely turbulent three-dimensional flows [18], it has not been so remarkably successful in this case. For example, with an approximate adjoint, Kim et al. [19] reduced the noise from a Mach 1.3 large-eddy simulation turbulent jet by 3.5 dB, which would seem to be a relatively modest amount given the flexibility of their actuation. It is not possible in this case to decouple the limits of control from numerical imprecision.

A straightforward approach to computing the adjoint is to start at the PDE level, derive the analytic adjoint PDE, and discretize the two systems independently. However, this is understood to introduce a particular kind of truncation error since the calculus identities used in the derivation of the adjoint are only approximated by their discrete counterparts. The most obvious approximation is the product rule of differentiation used in the formulations, but in general, boundary treatments and geometric transformations also introduce approximations that lead to an inconsistent adjoint. Thus, the predicted gradient is not fully compatible with the numerical solution; it no longer provides an exact gradient of the discrete system. This issue was previously identified for aerodynamics applications by Nadarajah and Jameson [20], who compare this continuous approach with a discrete-adjoint method, designed to be compatible with the discrete forward model. However, their discretization would not perform well for turbulence simulations. Carnarius et al. [21] showed that discretization errors in the continuous-adjoint-based gradient led to an over-prediction in the minimum drag coefficient for flow around a rotating cylinder. For the chaotic Lorenz system, Lea et al. [22] showed that the adjoint-based gradient suffers from cumulative error growth when the cost functional is a time-averaged quantity over an interval longer than the predictability time scales of the system. This is expected to be an important factor for the turbulent flow we consider, both because turbulence is chaotic and because efficient simulations usually employ resolutions close to the limits of the discretization. The resulting truncation errors will thus be relatively large to start with, before they amplify via the chaotic character of the system. In such situations, the discrete-adjoint method is expected to significantly accelerate and improve optimization. Developing a practical discrete-adjoint method, in the sense that it does not require operation counts or memory much beyond direct discretization of the continuous-adjoint, and demonstrating it on a challenging large-scale turbulent flow simulation are our goals. An inherent limitation of conventional continuous- and discrete-adjoint based formulations is their inability to overcome the chaos of turbulence indefinitely. Fortunately, they can be useful for finite-time-horizon control, and means of reformulating the overall problem to overcome this limitation is a subject of ongoing investigation [5].

In this paper, we derive an exact adjoint of the three-dimensional compressible flow equations discretized with the same high-order schemes commonly used in aeroacoustic simulations, and verify that it provides an exact (aside from finite-precision errors) gradient of the cost functional for the sound radiated by a compressible turbulent mixing layer. These schemes are not locally conservative, but are attractive and effective for the class of flows we consider, particularly because of their resolution. To do this, we make some key advances that facilitate implementation for such discretizations. This is an important step beyond the algorithms that have been developed to do this for aerodynamic simulations. For example, Rumpfkeil and Zingg [23] derived the discrete-adjoint equation for unsteady flows governed by the Euler equations integrated using implicit Euler and second-order backward difference (BDF2) methods; Yamaleev et al. [24] formulated a discrete-adjoint method for optimization using the three-dimensional Reynolds-averaged Navier–Stokes (RANS) equations on dynamic unstructured grids; Wang et al. [25] derived the discrete-adjoint equation for the two-dimensional compressible Euler equations with high-order discontinuous Galerkin discretization and an implicit fourth-order Runge–Kutta scheme; Roth and Ulbrich [26] developed a discrete-adjoint approach based on the sparse forward mode of automatic differentiation for a shape optimization problem incorporating various turbulence models; and Nielsen and Diskin [27] tested the

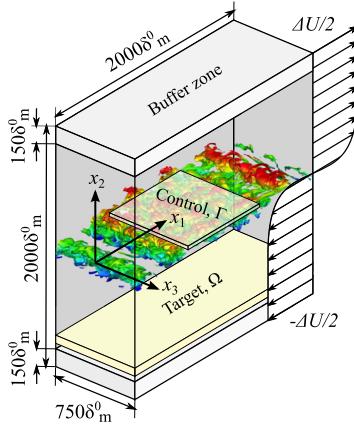


Fig. 1. The turbulent temporally developing mixing layer at $t = t_0$ with turbulence visualized by vorticity magnitude isosurfaces colored by the value of x_2 .

accuracy of the unsteady discrete-adjoint for the incompressible Navier–Stokes equations discretized on composite grids using a complex-variable approach and demonstrated its utility for shape optimization of a wind turbine geometry. None of the discretizations used in these examples are commonly used for aeroacoustic predictions of compressible turbulence because their resolution is often deemed to be inadequate [14,15]. The motivation for choosing discretization schemes in these examples has been to facilitate a straightforward derivation of the discrete-adjoint; we instead employ the same high-resolution numerics commonly used for aeroacoustic and compressible turbulence simulations. de Pando et al. [28] outlined an algorithm for efficiently evaluating the semi-discrete-adjoint operator from a modular implementation of a high-order nonlinear flow solver. While this modular approach is attractive, there is lost opportunity for combining operations in the adjoint evaluation which we exploit to save on computational cost. We shall also see that our approach requires even fewer modifications to the code than their formulation. Moreover, our formulation can provide a gradient that is compatible with the space–time discrete equations at no additional computational cost. Its accuracy is therefore not limited by either spatial or temporal resolution. The implementation is built directly from operations already available for solving the flow equations with important but remarkably few changes required to the algorithm.

Though it is well known that tools are available for automatic differentiation of discrete systems [29,30], these tools have rarely been used in flow simulations because they do not yield code with sufficient efficiency for use with large-scale simulations. Moreover, the underlying linear systems can be large and can become stiff or otherwise ill-conditioned on stretched grids or in the presence of complex source terms [31]. A fully discrete adjoint method has not yet been reported for this class of high-resolution schemes, perhaps due to their formidable complexity. The simplifications we identify are essential for its practical implementation.

In the following section we introduce a specific large-scale turbulence and sound simulation that will provide a challenging demonstration of the methods. This includes full details of the flow configuration, governing equations, boundary conditions, and discretization. It also includes a demonstration gradient calculation using the formulations introduced in detail in Section 3, with some of the more intricate details relegated to [Appendix A](#). The scheme is fully demonstrated in Section 4, where we demonstrate its accuracy and with it examine the controllability of free-shear turbulence. Implications for jet noise control are included in the discussion of conclusions in Section 5.

2. Simulation details

2.1. Flow configuration

We demonstrate our discrete-adjoint method on the turbulent temporally developing mixing layer shown in [Fig. 1](#), which is identical to the “ML2” mixing layer studied by Kleinman and Freund [11]. Full details of the flow are provided elsewhere [11,18]. The mixing layer is homogeneous in the x_1 - and x_3 -directions and inhomogeneous in the x_2 -direction. The initial mean flow is

$$\bar{u}_1 = \frac{\Delta U}{2} \tanh\left(\frac{x_2}{2\delta_m^0}\right), \quad \bar{u}_2 = \bar{u}_3 = 0 \quad \text{and} \quad \bar{p} = p_\infty,$$

with velocity difference $\Delta U = 0.9a_\infty$, where a_∞ is the ambient speed of sound. The initial mean density is obtained using the Crocco–Busemann relation assuming isothermal upper and lower streams,

$$\bar{\rho} = \rho_\infty \left[1 + \frac{\gamma - 1}{2a_\infty^2} \left(\frac{\Delta U^2}{4} - \bar{u}_1^2 \right) \right]^{-1}.$$

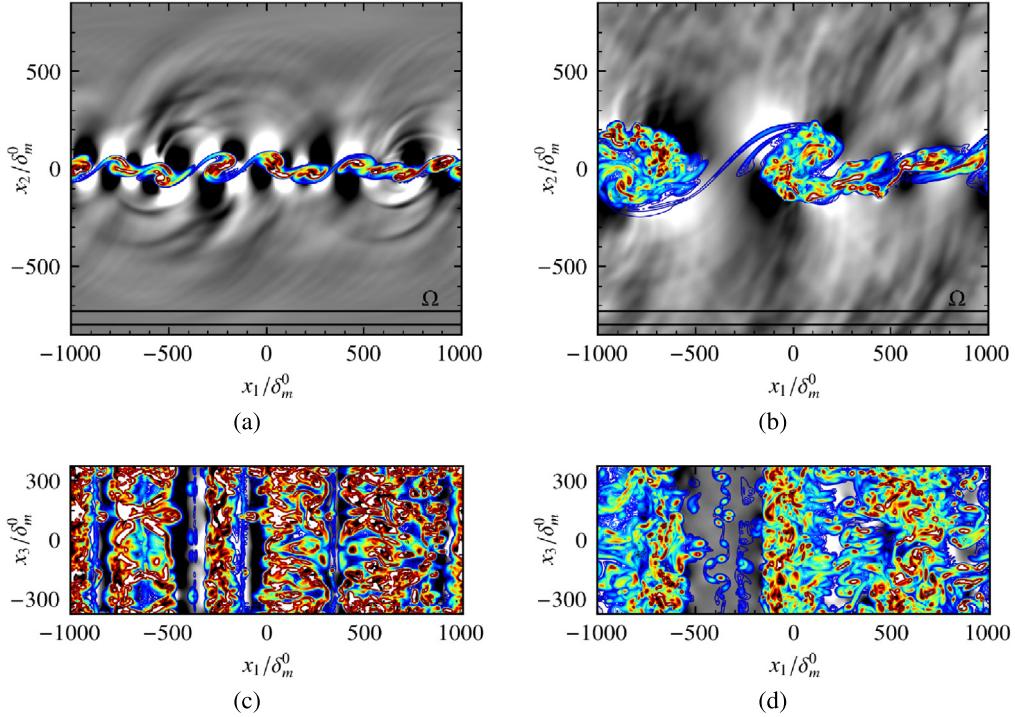


Fig. 2. Visualization showing contours of vorticity and dilatation along $x_3/\delta_m^0 = 0$ (top) and $x_2/\delta_m^0 = 0$ (bottom) at $t = t_0$ (left) and $t = t_1$ (right). The solid black lines in (a) and (b) are the boundaries of finite W_Ω defined in (9), where the radiated sound is measured.

The momentum thickness is used throughout as a length scale, and is defined as

$$\delta_m(t) = \frac{1}{\rho_\infty \Delta U^2} \int_{-\infty}^{\infty} \bar{\rho} \left(\frac{1}{4} \Delta U^2 - \tilde{u}_1^2 \right) dx_2,$$

where $\bar{\cdot}$ and $\tilde{\cdot}$ denote Reynolds and Favre averages, respectively [32]. The initial momentum thickness is defined as $\delta_m^0 = \delta_m(t=0)$. For the present flow configuration, mean quantities $\bar{\varphi}(x_2, t)$ are obtained by averaging $\varphi(\mathbf{x}, t)$ in the x_1 and x_3 directions. The corresponding Favre average is $\tilde{\varphi} = \bar{\rho}\bar{\varphi}/\bar{\rho}$. Perturbations from the Reynolds and Favre averages are denoted as φ' and φ'' , respectively.

As described by Kleinman and Freund [11], an auxiliary simulation (denoted in that reference as “ML3”) provides the initial condition for the present mixing layer. Starting from this initial condition, the flow is simulated from $t = 0$ to $t = t_0 \equiv 900\delta_m^0/a_\infty$ to allow the flow to evolve away from the imposed initial condition. The control we will introduce in Section 2.5 starts at $t = t_0$, though for aeroacoustic flow control, its effect on the sound field is delayed due to the finite sound speed. To compute adjoint coefficients that depend on the time-evolving flow field, an initial simulation must be performed without control, which also provides a baseline for comparison. Starting from $t = t_0$, the flow is simulated up to $t = t_1 \equiv 2900\delta_m^0/a_\infty$. Fig. 2 shows visualizations of the flow at $t = t_0$ and $t = t_1$. A realistic turbulence field appears to have developed. Reynolds stress statistics support this, and are reported in full elsewhere [33], as do the turbulence spectra presented in Section 4.2.

2.2. Governing equations

The equations governing the flow of a compressible, constant-viscosity Newtonian perfect gas are written in an operator form following Wei and Freund [17] and using the same non-dimensionalization, providing a concise notation for formulating the control problem. We take the state variable to be composed of the conservative flow quantities

$$Q = [\rho \quad \rho u_1 \quad \rho u_2 \quad \rho u_3 \quad e]^T,$$

where the total energy is $e = \frac{p}{\gamma-1} + \frac{1}{2}\rho u_i u_i$, the equation of state is $p = (\gamma-1)\rho T/\gamma$, and $\gamma = 1.4$ is the ratio of specific heats. The governing equations are written as $\mathcal{N}[Q](\mathbf{x}, t) = 0$, and can be expanded as

$$\mathcal{N}_1[Q](\mathbf{x}, t) = \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) \quad (1a)$$

$$\mathcal{N}_{i+1}[Q](\mathbf{x}, t) = \frac{\partial}{\partial t}(\rho u_i) + \frac{\partial}{\partial x_j}(\rho u_i u_j + p \delta_{ij}) - \frac{\partial \tau_{ij}}{\partial x_j}, \quad \text{for } i = 1, 2, 3 \quad (1b)$$

$$\mathcal{N}_5[Q](\mathbf{x}, t) = \frac{\partial e}{\partial t} + \frac{\partial}{\partial x_i}[u_i(e + p)] - \frac{\gamma}{\text{Re} \Pr(\gamma - 1)} \frac{\partial^2}{\partial x_i \partial x_i} \left(\frac{p}{\rho} \right) - \frac{\partial}{\partial x_i}(\tau_{ij} u_j), \quad (1c)$$

where

$$\tau_{ij} = \frac{1}{\text{Re}} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right).$$

2.3. Discretization

For our demonstration simulation, the governing equations are discretized on a grid with $N_1 \times N_2 \times N_3 = 680 \times 425 \times 168$ points in the stream-wise (x_1), cross-stream (x_2) and span-wise (x_3) directions, respectively. The computational domain extends from $-1000\delta_m^0$ to $1000\delta_m^0$ in the stream-wise and cross-stream directions, and from $-375\delta_m^0$ to $375\delta_m^0$ in the span-wise direction. The mesh is uniform in the x_1 and x_3 directions. The mesh in the x_2 -direction is stretched according to a mapping function $x_2 = g(s)$. Derivatives in the x_2 -direction are computed as

$$\frac{\partial f}{\partial x_2} = \frac{1}{g'} \frac{\partial f}{\partial s} \quad \text{and} \quad \frac{\partial^2 f}{\partial x_2^2} = \frac{1}{g'^2} \frac{\partial^2 f}{\partial s^2} - \frac{g''}{g'^3} \frac{\partial f}{\partial s}. \quad (2)$$

The specific form of $g(s)$ we use is $g(s) = 1000\delta_m^0 \hat{g}(s)/\hat{g}(1)$, where

$$\hat{g}(s) = (s - 0.5)(1 + 2b) + b\sigma \left(\int_{-\frac{c}{\sigma}}^{\frac{s-0.5-c}{\sigma}} \text{erf}(x) dx - \int_{\frac{c}{\sigma}}^{\frac{s-0.5+c}{\sigma}} \text{erf}(x) dx \right), \quad (3)$$

$b = 12$, $c = 0.6$, and $\sigma = 0.21$, which gives a minimum spacing $\Delta x_{2,\min} = 2.29\delta_m^0$ and a maximum point-to-point change in spacing of 1.97% at $x_2 = \pm 578\delta_m^0$.

Explicit narrow-stencil SBP finite-difference schemes [34] are used to approximate the first- and second-derivatives in the x_1 and x_2 directions. This choice is made even in the x_1 direction, which is periodic and therefore would be amenable to Fourier methods, because it facilitates efficient parallelization and better illustrates the methods as they would be used more generally. The schemes are fourth-order accurate with respect to a diagonal norm. We will use the same norm, which defines a sixth-order accurate quadrature rule, to define our discrete cost functional in Section 3.2. Mixed-derivatives are computed using a repeated application of two discrete first-derivative operators. The derivatives in the x_3 -direction are computed using a discrete Fourier transform. A Fourier filter removes the contribution of the highest 15% of wavenumbers in the x_3 -direction to suppress aliasing errors.

We will be analyzing and manipulating both continuous variables $f(\mathbf{x}, t)$ and their discretized-in-space approximations $\vec{f}(t)$, where the components of \vec{f} are the approximations of $f(\mathbf{x}, t)$ at the respective grid points. Thus, \vec{f} is a semi-discrete vector with $N_1 N_2 N_3$ elements. Using the Kronecker product notation described in Section 4.1 of Ref. [35], the finite-difference operators that approximate the j -th derivative along each direction are

$$\mathbf{D}_{jx_1} = D_{jx_1} \otimes I_{N_2} \otimes I_{N_3}$$

$$\mathbf{D}_{jx_2} = I_{N_1} \otimes D_{jx_2} \otimes I_{N_3}$$

$$\mathbf{D}_{jx_3} = I_{N_1} \otimes I_{N_2} \otimes D_{jx_3},$$

where I_n denotes an $n \times n$ identity matrix, and D_{jx_i} is the finite-difference approximation of $\partial^j f / \partial x_i^j$ represented in the usual way as a matrix operator. The Laplacian operator is $\mathbf{L} = \sum_i \mathbf{D}_{2x_i}$ and the discrete mixed-derivative operator is

$$\Delta_{ij} = \begin{cases} \mathbf{D}_{1x_i} \mathbf{D}_{1x_j} & i \neq j \\ \mathbf{D}_{2x_i} & i = j \end{cases}$$

Additionally, the operator \circ denotes an element-wise multiplication. Element-wise division is written as \vec{f}/\vec{g} . Using this notation, spatial discretization of (1) leads to

$$\vec{\mathbf{N}}[\vec{Q}](t) = 0, \quad (4)$$

which can be expanded as

$$\vec{\mathbf{N}}_1[\vec{Q}](t) = \frac{d\vec{\rho}}{dt} + \mathbf{D}_{1x_i} \rho \vec{u}_i, \quad (5a)$$

$$\vec{\mathbf{N}}_{i+1}[\vec{Q}](t) = \frac{d\rho \vec{u}_i}{dt} + \mathbf{D}_{1x_j} (\rho \vec{u}_i \circ \vec{u}_j + \vec{p} \delta_{ij}) - \frac{1}{\text{Re}} \left(\mathbf{L} \vec{u}_i - \frac{1}{3} \Delta_{ij} \vec{u}_j \right), \quad \text{for } i = 1, 2, 3 \quad (5b)$$

$$\vec{N}_5[\vec{Q}](t) = \frac{d\vec{e}}{dt} + \mathbf{D}_{1x_i}[\vec{u}_i \circ (\vec{e} + \vec{p})] - \frac{\gamma}{\text{Re} \Pr(\gamma-1)} \mathbf{L} \frac{\vec{p}}{\rho} - \frac{1}{\text{Re}} \vec{u}_i \circ \left(\mathbf{L} \vec{u}_i + \frac{1}{3} \Delta_{ij} \vec{u}_j \right) - \frac{1}{\text{Re}} \vec{\phi}, \quad (5c)$$

where

$$\vec{\phi} = (\mathbf{D}_{1x_j} \vec{u}_i) \circ (\mathbf{D}_{1x_i} \vec{u}_j) + (\mathbf{D}_{1x_i} \vec{u}_j) \circ (\mathbf{D}_{1x_j} \vec{u}_i) - \frac{2}{3} (\mathbf{D}_{1x_i} \vec{u}_i) \circ (\mathbf{D}_{1x_j} \vec{u}_j).$$

2.4. Boundary conditions

Periodic boundaries are employed in the x_1 and x_3 directions to model a spatially extensive shear layer. Absorbing buffer zones [36] of size $150\delta_m^0$ are used near the boundaries in the x_2 -direction, where the solution is damped towards the known base state $\vec{Q}_b(\mathbf{x})$, which is identical to the initial mean flow. This is achieved by adding a damping term

$$\vec{N}_{\text{damping}}[\vec{Q}, \vec{Q}_b](t) = -\vec{\xi} \circ \{\vec{Q}(t) - \vec{Q}_b\}, \quad (6)$$

to the left-hand-side of (4). $\vec{\xi}$ is the semi-discrete representation of the buffer strength $\xi(\mathbf{x})$, which only has support in the buffer zones, and varies quadratically from 0 at $x_2 = \pm 850\delta_m^0$ to $0.6a_\infty/\delta_m^0$ at $x_2 = \pm 1000\delta_m^0$. At the far edge of the buffer zones, we also use a simultaneous approximation term (SAT) to impose non-reflecting characteristic boundary conditions [35]. The specific form of the SAT we use is

$$\vec{N}_{\text{SAT}}[\vec{Q}, \vec{Q}_b](t) = -\vec{\Xi}[\vec{Q}_b]\{\vec{Q}(t) - \vec{Q}_b\}, \quad (7)$$

where

$$\vec{\Xi}[\vec{Q}_b] \equiv \sigma^I \sum_{i=1}^3 \mathbf{P}_i^{-1} (\mathbf{E}_{1i} \boldsymbol{\Theta}_i [\vec{Q}_b] \boldsymbol{\Lambda}_i^- [\vec{Q}_b] \boldsymbol{\Theta}_i^{-1} [\vec{Q}_b] - \mathbf{E}_{0i} \boldsymbol{\Theta}_i [\vec{Q}_b] \boldsymbol{\Lambda}_i^+ [\vec{Q}_b] \boldsymbol{\Theta}_i^{-1} [\vec{Q}_b]),$$

with $\sigma^I = 1$. Svärd et al. [35] define the matrices \mathbf{P}_i , \mathbf{E}_{0i} and \mathbf{E}_{1i} : \mathbf{P}_i is a quadrature scheme for approximating integrals over the two surfaces normal to the x_i -direction, and \mathbf{E}_{0i} and \mathbf{E}_{1i} provide a compact support over these surfaces. $\boldsymbol{\Theta}_i[\vec{Q}_b]$ and $\boldsymbol{\Lambda}^\pm[\vec{Q}_b]$ are semi-discrete representations of matrices derived by Pulliam and Chaussee [37] from an eigenvalue decomposition of the Jacobian of the inviscid fluxes.

Thus, including the boundary terms the spatially discretized unforced flow equations are

$$\vec{N}[\vec{Q}](t) + \vec{N}_{\text{damping}}[\vec{Q}, \vec{Q}_b](t) + \vec{N}_{\text{SAT}}[\vec{Q}, \vec{Q}_b](t) = 0.$$

2.5. Control implementation

The control for this demonstration is a source $f_i(\mathbf{x}, t)$ in the governing equation,

$$\mathcal{M}_i[Q, f](\mathbf{x}, t) \equiv \mathcal{N}_i[Q](\mathbf{x}, t) - W_\Gamma^2(\mathbf{x}, t) f_i(\mathbf{x}, t) = 0, \quad \text{for } i = 1, 2, \dots, 5, \quad (8)$$

for $\mathbf{x} \in \mathbb{D}$ and $t_0 \leq t \leq t_1$, where \mathbb{D} is the computational domain. The $W_\Gamma^2(\mathbf{x}, t)$ factor sets the support for the control forcing, which is restricted to $\Gamma \subset \mathbb{D}$ and goes to zero smoothly near the boundary $\partial\Gamma$ of the control region (see Fig. 1). The cost functional is denoted as $\mathcal{J}[Q, f]$ and is defined as the norm of a function whose support is restricted to a target region $\Omega \subset \mathbb{D}$ (see Fig. 1). This compactness is enforced through a factor $W_\Omega^2(\mathbf{x}, t)$, defined in the same manner as $W_\Gamma^2(\mathbf{x}, t)$. The support functions W_Ω^2 and W_Γ^2 are expressed as squared quantities to emphasize that they preserve the positive-definiteness of subsequently crafted integrands. In our noise suppression demonstrations, an appropriate cost functional for aeroacoustic noise suppression is

$$\mathcal{J}[Q, f] = \| [0 \ 0 \ 0 \ 0 \ p'(\mathbf{x}, t)]^T \circ W_\Omega(\mathbf{x}, t) \|^2, \quad (9)$$

where $p'(\mathbf{x}, t)$ is the pressure perturbation. The control and target regions for the turbulent mixing layer simulations are shown in Fig. 1. The specific control and target mollifying support functions are defined such that

$$\begin{aligned} W_{\Omega,5}^2(\mathbf{x}, t) &\sim B_{0,2}(x_2^\Omega) r_\Omega(t) \\ W_{\Gamma,5}^2(\mathbf{x}, t) &\sim B_{0,2}(x_2^\Gamma) \{ \tanh[60(x_1^\Gamma - 0.1)] - \tanh[60(x_1^\Gamma - 0.9)] \} r_\Gamma(t), \end{aligned} \quad (10)$$

where $r_\Omega(t) = r_\Gamma(t) = 0.5 \{ \tanh[1000(\frac{t-t_0}{t_1-t_0} - 0.01)] + \tanh[1000(\frac{t_1-t}{t_1-t_0} - 0.01)] \}$,

$$\begin{aligned} x_2^\Omega &= \begin{cases} \frac{1}{70}(\frac{x_2}{\delta_m^0} + 800), & |\frac{x_2}{\delta_m^0} + 765| \leq 35 \\ 0, & \text{otherwise,} \end{cases} \\ x_1^\Gamma &= \begin{cases} \frac{1}{800}(\frac{x_1}{\delta_m^0} + 400), & |\frac{x_1}{\delta_m^0}| \leq 400 \\ 0, & \text{otherwise,} \end{cases} \quad x_2^\Gamma = \begin{cases} \frac{1}{32}(\frac{x_2}{\delta_m^0} + 16), & |\frac{x_2}{\delta_m^0}| \leq 16 \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

and $B_{0,2}$ is the B-spline basis function of degree 2, which provides compact support and blends smoothly to zero at the cross-stream boundaries of the control and target regions. The mollifying support functions in (10) are further scaled by the inverse of their norm based on the inner product (13). The target region extends fully across the computational domain in the periodic x_1 and x_3 directions. $r_\Gamma(t)$ and $r_\Omega(t)$ ensure that the control forcing in (8) and the target forcing in the adjoint equations (see Section 3) increase gradually from zero. With these specifications, we see in Fig. 2(a) & (b) that the target region is located in the acoustic field of the mixing layer since that is where we are interested in reducing the aeroacoustic noise. The control region is located in the shear layer, so that the control will alter the turbulence that makes the sound. We emphasize that we are not designing an “anti-sound” acoustic cancellation control. Wei and Freund [17] explored different types of controls for a two-dimensional compressible mixing layer and achieved maximum noise reduction using an internal energy control, which we use in all of our simulations.

2.6. Sensitivity gradient and its error

The adjoint solution $Q^\dagger(\mathbf{x}, t)$ provides a sensitivity gradient $\mathcal{G}[Q^\dagger](\mathbf{x}, t)$, the precise formulation for which is provided in Section 3. Here, in our demonstrations for compressible turbulence and sound, we use it to update the control forcing $f(\mathbf{x}, t)$. Solving (8) with $f^{(0)}(\mathbf{x}, t) = 0$ provides a baseline solution $Q^{(0)}(\mathbf{x}, t)$. We consider the initial search direction $-\mathcal{G}[Q^{(0)}](\mathbf{x}, t)$, along which better controls are sought by determining an α_{\min} such that

$$f^{(1)}(\mathbf{x}, t) = f^{(0)}(\mathbf{x}, t) - \alpha_{\min} \mathcal{G}[Q^{(0)}](\mathbf{x}, t), \quad \alpha_{\min} \in \mathbb{R}^+, \quad (11)$$

minimizes $\mathcal{J}[Q, f^{(1)}]$.

As discussed in Section 1, it is understood that the discretization of the continuous-adjoint formulation will have inherent inconsistencies that limits its accuracy. To quantify these errors in order to assess the quality of gradient predictions and validate our implementation of a consistent adjoint, we compare the adjoint-based sensitivity with a finite-difference derived estimate of the gradient. To measure these errors, we consider the variation in the control forcing, which is thus related to the gradient as

$$\delta f(\mathbf{x}, t) \equiv f(\mathbf{x}, t) - f^{(0)}(\mathbf{x}, t) = -\alpha \mathcal{G}[Q^{(0)}](\mathbf{x}, t),$$

where $\alpha \in \mathbb{R}^+$ is a distance measure in the control space. Expanding the cost functional using a Taylor series, we obtain

$$\mathcal{J}[Q, f + \delta f] = \mathcal{J}[Q^{(0)}, f] + \left\langle \frac{\delta \mathcal{J}}{\delta f} [Q^{(0)}, f^{(0)}; \delta f], \delta f \right\rangle + \mathcal{O}(\delta^2 f),$$

or, equivalently, $\mathcal{J}[Q, f + \delta f] - \mathcal{J}[Q^{(0)}, f] = -\alpha \|\mathcal{G}[Q^{(0)}]\|^2 + \mathcal{O}(\alpha^2)$. This leads to an error measure

$$\mathcal{E} = \left| \frac{\mathcal{J}[Q, f + \delta f] - \mathcal{J}[Q^{(0)}, f]}{\alpha} + \|\mathcal{G}[Q^{(0)}]\|^2 \right|, \quad (12)$$

where all quantities on the right-hand side (except α) are computed by numerically solving the flow or adjoint equations. Therefore, $\mathcal{E} = \mathcal{O}_1(\alpha) + \mathcal{O}_2(\Delta x_i^a, \Delta t^b)$, where $\mathcal{O}_1(\alpha)$ is the error due to a first-order finite-difference approximation of the gradient embodied in (12), and $\mathcal{O}_2(\Delta x_i^a, \Delta t^b)$ is due to numerical discretization reflecting the order of the discretization for how the solutions for $Q(\mathbf{x}, t)$ and $Q^\dagger(\mathbf{x}, t)$ are obtained. Specifically, the exponents a and b in the numerical discretization error depend on the accuracy of the numerical differentiation and quadrature schemes used. However, since it is not in general possible to anticipate how these errors accumulate for a turbulent flow it is not possible to anticipate where $\mathcal{O}_1(\alpha)$ or $\mathcal{O}_2(\Delta x_i^a, \Delta t^b)$ will dominate, which is important to understand since it directly affects the utility of the adjoint solution in providing an approximate gradient. Since turbulence is well understood to be chaotic, even having the nominally exact adjoint in its discrete formulation does not guarantee its utility for any control objective. We quantify this for our model turbulent flow in Section 4.4 and discuss additional implications in Section 5. Nadarajah and Jameson [38] used an alternate error measure based on a complex-step derivative of the cost functional. We do not use this approach since it is computationally three times more expensive than the finite-difference approach, and its advantages over the current approach do not adequately justify the added computational cost for simulations of this scale.

At this point, we only make an initial assessment based on (12) of the accuracy of the continuous- and discrete-adjoint formulations to be developed in the following section. Fig. 3 shows the gradient error \mathcal{E} from (12) for the three-dimensional mixing layer. As $\alpha \rightarrow 0$, the continuous-adjoint formulation leads to approximately constant \mathcal{E} , indicating that the dominant contribution to the gradient error is the $\mathcal{O}_2(\Delta x_i^a, \Delta t^b)$ term for our selected resolution. For the discrete-adjoint simulations, $\mathcal{E} \sim \alpha$ up to an apparent round-off level, which will be confirmed in Section 4.1. The formulation that provides this accuracy in a practical way is developed in the following section.

3. Adjoint formulations

3.1. Continuous-adjoint formulation

For subsequent comparisons and to provide context, we start with the continuous-adjoint formulation. Our discrete-adjoint formulation will employ a corresponding notation. We first define an inner product,

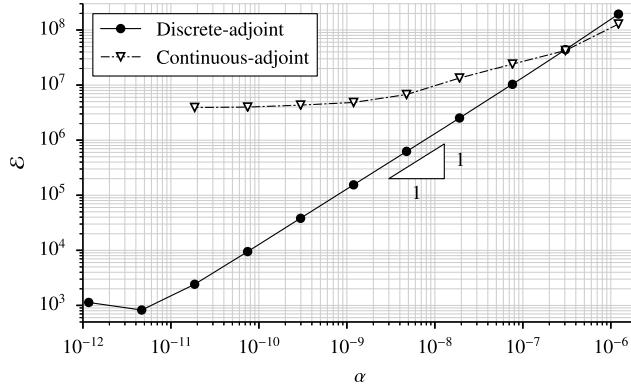


Fig. 3. Comparison of gradient accuracy from the discrete- and continuous-adjoint formulations for the three-dimensional turbulent flow configuration of Fig. 1.

$$\langle f_1, f_2 \rangle = \int_{t_0}^{t_1} \int_{\mathbb{D}} f_1(\mathbf{x}, t)^T f_2(\mathbf{x}, t) d^3 \mathbf{x} dt, \quad (13)$$

and a corresponding norm $\|f_1\|^2 = \langle f_1, f_1 \rangle$, where f_1 and f_2 are sufficiently smooth vector-valued square-integrable functions of \mathbf{x} and t , with $\mathbf{x} \in \mathbb{D} \subset \mathbb{R}^3$ and $t_0 \leq t \leq t_1$. The (weak) derivatives of the cost functional with respect to $Q(\mathbf{x}, t)$ and $f(\mathbf{x}, t)$ are

$$\frac{\partial \mathcal{J}}{\partial Q}[Q, f](\mathbf{x}, t) = W_Q^2(\mathbf{x}, t) \circ g_Q(\mathbf{x}, t) \quad \text{and} \quad \frac{\partial \mathcal{J}}{\partial f}[Q, f](\mathbf{x}, t) = 0, \quad (14)$$

where $g_Q(\mathbf{x}, t)$ is the target forcing to be developed from (9). The gradient \mathcal{G} of the cost functional is defined using

$$\delta \mathcal{J}[Q, f; \delta Q, \delta f] = \langle \mathcal{G}, \delta f \rangle. \quad (15)$$

It is well understood that this definition of the gradient is useful since it relates the total variation of the cost functional to a variation in the control forcing $f(\mathbf{x}, t)$ and does not involve the variation in the state variable $Q(\mathbf{x}, t)$. Thus, if this gradient were known, an iterative minimization algorithm has to update only $f(\mathbf{x}, t)$ in order to accelerate convergence to a local minimum. To find this gradient, we formulate the control problem using a Lagrange multiplier to enforce adherence to the governing equations (8) as a constraint. The Lagrangian is

$$\mathcal{L}[Q, Q^\dagger, f] = \mathcal{J}[Q, f] - \langle Q^\dagger, \mathcal{M}[Q, f] \rangle, \quad (16)$$

with $\delta \mathcal{L}[Q, Q^\dagger, f; \delta Q, \delta f] = \delta \mathcal{J}[Q, f; \delta Q, \delta f]$ since $\delta \mathcal{M}[Q, f; \delta Q, \delta f](\mathbf{x}, t) = 0$, and its variation is

$$\begin{aligned} \delta \mathcal{L}[Q, Q^\dagger, f; \delta Q, \delta f] &= \delta \mathcal{J}_Q[Q, f; \delta Q] - \langle Q^\dagger, \delta \mathcal{M}_Q[Q, f; \delta Q] \rangle \\ &\quad + \delta \mathcal{J}_f[Q, f; \delta f] - \langle Q^\dagger, \delta \mathcal{M}_f[Q, f, \delta f] \rangle, \end{aligned} \quad (17)$$

where the subscripts on $\delta \mathcal{J}$ and $\delta \mathcal{M}$ denote the function with respect to which the functional variation is calculated. If $Q^\dagger(\mathbf{x}, t)$, which serves as a Lagrange multiplier in (16), is chosen to satisfy the continuous-adjoint equations

$$\mathcal{M}^\dagger[Q, Q^\dagger, g_Q] \equiv \mathcal{N}^\dagger[Q, Q^\dagger](\mathbf{x}, t) - W_Q^2(\mathbf{x}, t) \circ g_Q(\mathbf{x}, t) = 0, \quad (18)$$

where $\mathcal{N}^\dagger[Q, Q^\dagger](\mathbf{x}, t)$ is obtained from

$$\langle Q^\dagger, \delta \mathcal{N}[Q, \delta Q] \rangle = \langle \mathcal{N}^\dagger[Q, Q^\dagger], \delta Q \rangle, \quad (19)$$

then the term $\delta \mathcal{J}_Q[Q, f; \delta Q] - \langle Q^\dagger, \delta \mathcal{M}_Q[Q, f; \delta Q] \rangle$ in (17) is eliminated, which leads to the desired explicit expression for the gradient in terms of the adjoint variable Q^\dagger

$$\mathcal{G}[Q^\dagger](\mathbf{x}, t) = W_Q^2(\mathbf{x}, t) \circ Q^\dagger(\mathbf{x}, t). \quad (20)$$

Were the control strength in (8) somehow penalized, as it often is in some applications, $\frac{\partial \mathcal{J}}{\partial f}$ in (14) is non-zero and appears on the right-hand side of (20).

From (14) and our definition of $\mathcal{J}[Q, f]$ for our control objective in (9), the source term in the adjoint equation is

$$g_Q(\mathbf{x}, t) = [0 \ 0 \ 0 \ 0 \ 2p'(\mathbf{x}, t)]^T.$$

Using the governing equations (1) and integration-by-parts within (19) leads to the continuous-adjoint equations for the compressible flow equations,

$$\begin{aligned} \mathcal{N}_1^\dagger[Q, Q^\dagger] = & \frac{\partial Q_1^\dagger}{\partial t} + \frac{\gamma - 1}{2} u_j u_j \frac{\partial Q_{i+1}^\dagger}{\partial x_i} - u_i u_j \frac{\partial Q_{i+1}^\dagger}{\partial x_j} + \left[\frac{\gamma - 1}{2} u_i u_j u_j - \frac{u_i}{\rho} (e + p) \right] \frac{\partial Q_5^\dagger}{\partial x_i} \\ & - \frac{1}{\text{Re}} \left(\frac{u_i}{\rho} \frac{\partial^2 Q_{i+1}^\dagger}{\partial x_j \partial x_j} + \frac{1}{3} \frac{u_j}{\rho} \frac{\partial^2 Q_{i+1}^\dagger}{\partial x_i \partial x_j} \right) + \frac{u_j}{\rho} \tau_{ij} \frac{\partial Q_5^\dagger}{\partial x_i} \\ & - \frac{1}{\text{Re}} \left[\frac{u_i}{\rho} \frac{\partial}{\partial x_j} \left(u_j \frac{\partial Q_5^\dagger}{\partial x_i} \right) + \frac{u_j}{\rho} \frac{\partial}{\partial x_i} \left(u_j \frac{\partial Q_5^\dagger}{\partial x_i} \right) \right] + \frac{2}{3 \text{Re}} \frac{u_j}{\rho} \frac{\partial}{\partial x_j} \left(u_i \frac{\partial Q_5^\dagger}{\partial x_i} \right) \\ & + \frac{\gamma}{\text{Re} \Pr \rho^2} \left(\frac{1}{2} \rho u_j u_j - \frac{p}{\gamma - 1} \right) \frac{\partial^2 Q_5^\dagger}{\partial x_i \partial x_j}, \end{aligned} \quad (21a)$$

$$\begin{aligned} \mathcal{N}_{i+1}^\dagger[Q, Q^\dagger] = & \frac{\partial Q_{i+1}^\dagger}{\partial t} + \frac{\partial Q_1^\dagger}{\partial x_i} - (\gamma - 1) u_i \frac{\partial Q_{j+1}^\dagger}{\partial x_j} + u_j \left(\frac{\partial Q_{j+1}^\dagger}{\partial x_i} + \frac{\partial Q_{i+1}^\dagger}{\partial x_j} \right) \\ & - (\gamma - 1) u_i u_j \frac{\partial Q_5^\dagger}{\partial x_j} + \frac{1}{\rho} (e + p) \frac{\partial Q_5^\dagger}{\partial x_i} + \frac{1}{\text{Re}} \left(\frac{1}{\rho} \frac{\partial^2 Q_{i+1}^\dagger}{\partial x_j \partial x_j} + \frac{1}{\rho} \frac{\partial^2 Q_{j+1}^\dagger}{\partial x_i \partial x_j} \right) \\ & - \frac{2}{3 \text{Re}} \frac{1}{\rho} \frac{\partial^2 Q_{j+1}^\dagger}{\partial x_j \partial x_i} - \frac{1}{\rho} \tau_{ji} \frac{\partial Q_5^\dagger}{\partial x_j} + \frac{1}{\text{Re}} \left[\frac{1}{\rho} \frac{\partial}{\partial x_j} \left(u_j \frac{\partial Q_5^\dagger}{\partial x_i} \right) + \frac{1}{\rho} \frac{\partial}{\partial x_j} \left(u_i \frac{\partial Q_5^\dagger}{\partial x_j} \right) \right] \\ & - \frac{2}{3 \text{Re}} \frac{1}{\rho} \frac{\partial}{\partial x_i} \left(u_j \frac{\partial Q_5^\dagger}{\partial x_j} \right) - \frac{\gamma u_i}{\text{Re} \Pr \rho} \frac{\partial^2 Q_5^\dagger}{\partial x_j \partial x_i}, \quad i = 1, 2, 3, \end{aligned} \quad (21b)$$

$$\mathcal{N}_5^\dagger[Q, Q^\dagger] = \frac{\partial Q_5^\dagger}{\partial t} + (\gamma - 1) \frac{\partial Q_{i+1}^\dagger}{\partial x_i} + \gamma u_i \frac{\partial Q_5^\dagger}{\partial x_i} + \frac{\gamma}{\text{Re} \Pr} \frac{1}{\rho} \frac{\partial^2 Q_5^\dagger}{\partial x_i \partial x_i}. \quad (21c)$$

Note, in deriving (21), we have used a causality condition in time and neglected boundary terms. The latter are strictly zero only when $\mathbb{D} = \mathbb{R}^3$ in (13), which is modeled using a radiation condition (6). A similar condition is used for the adjoint, taking care of the time-reversal with a negative damping strength, as has been described before [17]. The boundary terms are problem-specific, and can be included as necessary in general. These equations are identical to those reported by Kleinman [33], which includes non-reflecting adjoint boundary conditions that are compatible with the radiation condition for the adjoint.

Although the formulation does not require this, to numerically solve (21), we discretize the derivatives using the same schemes as for the flow equations discussed in Section 2.3 including the same fourth-order Runge–Kutta scheme for time-integration. The boundary conditions for Q^\dagger are implemented using an SAT similar to (7).

3.2. Discrete-adjoint formulation

The discrete-adjoint formulation is specific to the discretization, and it is clearest if we consider the time discretization at the outset. While our approach for formulating the discrete-adjoint remains generic to the broad class of popular high-order explicit multi-stage methods, we will demonstrate our approach using the standard fourth-order Runge–Kutta (RK4) scheme. First, we write the semi-discrete approximation to the governing equations (8)

$$\vec{\mathbf{M}}[\vec{Q}, \vec{f}](t) \equiv \frac{d\vec{Q}}{dt} - \vec{\mathbf{R}}[\vec{Q}](t) - \vec{F}_\Gamma(t) = 0, \quad (22)$$

where $\vec{F}_\Gamma(t) = \vec{W}_\Gamma^2(t) \circ \vec{f}(t)$ is the semi-discrete form of the control forcing term, and $\vec{\mathbf{R}}[\vec{Q}](t)$ represents the discretized right-hand side of the compressible flow equations including SATs that enforce boundary conditions. For (22), we take initial condition $\vec{Q}(t_0) = \vec{Q}^0$. For convenience of presentation, we assume a uniform time step such that $t^n = t_0 + \Delta t(n - 1)$, where $\Delta t = (t_1 - t_0)/M$ with $n = 1, 2, \dots, M + 1$ for $M + 1$ time steps. Integrating (22) using the most common RK4 scheme leads to

$$\vec{\mathbf{M}}^{n,1} = \frac{2\vec{Q}^{n,1} - 2\vec{Q}^{n-1,4}}{\Delta t} - \vec{\mathbf{R}}^{n-1,4} - \vec{F}_\Gamma^{n,1} \quad (23a)$$

$$\vec{\mathbf{M}}^{n,2} = \frac{2\vec{Q}^{n,2} - 2\vec{Q}^{n-1,4}}{\Delta t} - \vec{\mathbf{R}}^{n,1} - \vec{F}_\Gamma^{n,2} \quad (23b)$$

$$\vec{\mathbf{M}}^{n,3} = \frac{\vec{Q}^{n,3} - \vec{Q}^{n-1,4}}{\Delta t} - \vec{\mathbf{R}}^{n,2} - \vec{F}_\Gamma^{n,3} \quad (23c)$$

$$\vec{M}^{n,4} = \frac{6\vec{Q}^{n,4} + 2\vec{Q}^{n-1,4} - 2\vec{Q}^{n,1} - 4\vec{Q}^{n,2} - 2\vec{Q}^{n,3}}{\Delta t} - \vec{R}^{n,3} - \vec{F}_T^{n,4}, \quad (23d)$$

each stage of which has a single unknown \vec{Q} , which can be simply solved for in the usual way for explicit RK schemes given the $\vec{M}[\vec{Q}, \vec{f}] = 0$ condition of (22). This particular form of expression is useful for our objectives. In (23), $\vec{Q}^{0,4}$ is the known initial state \vec{Q}^0 with $\vec{R}^{0,4} = \vec{R}[\vec{Q}^0](t_0)$. For $n > 1$, $\vec{Q}^{n-1,4}$ is the numerical approximation to $\vec{Q}(t^n)$ and the sub-step values are $\vec{Q}^{n,s}$ for $n = 1, \dots, M$ and $s = 1, 2, 3$. $\vec{R}^{n,s}$ denotes $\vec{R}[\vec{Q}^{n,s}](t_n + c_s \Delta t)$, where $c_1 = c_2 = 1/2$ and $c_3 = c_4 = 1$.

To derive the discrete analog of (19), we define an inner product

$$\langle \underline{f}_1, \underline{f}_2 \rangle_{\mathbf{P}} = \sum_{n=1}^M \sum_{s=1}^4 \beta^{n,s} \Delta t (\underline{f}_1^{n,s})^T \mathbf{P} \underline{f}_2^{n,s}, \quad (24)$$

and the corresponding norm $\|\underline{f}_1\|_{\mathbf{P}}^2 = \langle \underline{f}_1, \underline{f}_1 \rangle_{\mathbf{P}}$, where \underline{f}_1 denotes a discrete vector with components $\vec{f}_1^{1,1}, \dots, \vec{f}_1^{M,4}$. In total, \underline{f}_1 consists of $20MN_1N_2N_3$ elements. The values of $\beta^{n,s}$ are developed in the following such that (24) is a consistent approximation of (13). In (24), $\mathbf{P} = I_5 \otimes P_1 \otimes P_2 \otimes P_3$, where each P_i is a diagonal matrix representing a quadrature rule for the x_i -direction. In the mixing layer simulations, the stream-wise and span-wise directions are periodic with uniform mesh spacings Δx_1 and Δx_3 , so $P_1 = \Delta x_1 \text{diag}(\mathbb{1}_{N_1})$ and $P_3 = \Delta x_3 \text{diag}(\mathbb{1}_{N_3})$ are natural choices, where $\mathbb{1}_n$ denotes a vector with elements all equal to unity. In the cross-stream direction, a numerical quadrature scheme is used. A natural choice for P_2 is based on the diagonal norm used for constructing the SBP operator D_{1x_2} [34], which is a symmetric and positive-definite matrix, and defines a sixth-order accurate quadrature rule [39]. This leads to

$$P_2 = \text{diag}\left(\left[\frac{13649}{43200} \quad \frac{12013}{8640} \quad \frac{2711}{4320} \quad \frac{5359}{4320} \quad \frac{7877}{8640} \quad \frac{43801}{43200} \quad 1 \quad \dots \quad \frac{13649}{43200}\right] \circ \left[\frac{g'(0)}{N_2 - 1} \quad \dots \quad \frac{g'(1)}{N_2 - 1}\right]\right),$$

and accounts for mesh stretching per (3).

The counterparts to (8), (18), (19) and (20) from the continuous-adjoint formulation can now be developed. To start with, we write (23) compactly as

$$\underline{M}[\underline{Q}, \underline{f}] \equiv \underline{N}[\underline{Q}] - \underline{W}_T^2 \circ \underline{f} = 0.$$

The corresponding discrete-adjoint equations are

$$\underline{M}^\dagger[\underline{Q}, \underline{Q}^\dagger, \underline{g}_\Omega] \equiv \underline{N}^\dagger[\underline{Q}, \underline{Q}^\dagger] - \underline{W}_\Omega^2 \circ \underline{g}_\Omega = 0,$$

and are obtained using

$$\langle \underline{Q}^\dagger, \delta \underline{N}[\underline{Q}, \delta \underline{Q}] \rangle_{\mathbf{P}} = \langle \underline{N}^\dagger[\underline{Q}, \underline{Q}^\dagger], \delta \underline{Q} \rangle_{\mathbf{P}}. \quad (25)$$

The gradient of the cost functional is

$$\underline{G}[\underline{Q}^\dagger] = \underline{W}_T^2 \circ \underline{Q}^\dagger. \quad (26)$$

Using (23) and (24), we get

$$\begin{aligned} \vec{M}^{\dagger n,4} &= \frac{6\beta^{n,4}\vec{Q}^{\dagger n,4} + 2\beta^{n+1,4}\vec{Q}^{\dagger n+1,4} - \beta^{n+1,3}\vec{Q}^{\dagger n+1,3} - 2\beta^{n+1,2}\vec{Q}^{\dagger n+1,2} - 2\beta^{n+1,1}\vec{Q}^{\dagger n+1,1}}{\beta^{n,4}\Delta t} \\ &\quad - \frac{\beta^{n+1,1}}{\beta^{n,4}} \mathbf{P}^{-1} \mathbf{T}^{n,4} \mathbf{P} \vec{Q}^{\dagger n+1,1} - \vec{G}_\Omega^{\dagger n,4} \end{aligned} \quad (27a)$$

$$\vec{M}^{\dagger n,3} = \frac{\beta^{n,3}\vec{Q}^{\dagger n,3} - 2\beta^{n,4}\vec{Q}^{\dagger n,4}}{\beta^{n,3}\Delta t} - \frac{\beta^{n,4}}{\beta^{n,3}} \mathbf{P}^{-1} \mathbf{T}^{n,3} \mathbf{P} \vec{Q}^{\dagger n,4} - \vec{G}_\Omega^{\dagger n,3} \quad (27b)$$

$$\vec{M}^{\dagger n,2} = \frac{2\beta^{n,2}\vec{Q}^{\dagger n,2} - 4\beta^{n,4}\vec{Q}^{\dagger n,4}}{\beta^{n,2}\Delta t} - \frac{\beta^{n,3}}{\beta^{n,2}} \mathbf{P}^{-1} \mathbf{T}^{n,2} \mathbf{P} \vec{Q}^{\dagger n,3} - \vec{G}_\Omega^{\dagger n,2} \quad (27c)$$

$$\vec{M}^{\dagger n,1} = \frac{2\beta^{n,1}\vec{Q}^{\dagger n,1} - 2\beta^{n,4}\vec{Q}^{\dagger n,4}}{\beta^{n,1}\Delta t} - \frac{\beta^{n,2}}{\beta^{n,1}} \mathbf{P}^{-1} \mathbf{T}^{n,1} \mathbf{P} \vec{Q}^{\dagger n,2} - \vec{G}_\Omega^{\dagger n,1} \quad (27d)$$

for $n = 1, \dots, M-1$, where $\vec{G}_\Omega^{\dagger n,s} = \vec{W}_\Omega^2(t_n + c_s \Delta t) \circ \vec{g}_\Omega^{\dagger n,s}$ and $\mathbf{T}^{n,s}$ is a matrix operator defined implicitly using

$$\delta \vec{R}^{\dagger n,s} = (\mathbf{T}[\vec{Q}^{\dagger n,s}])^T \delta \vec{Q}^{\dagger n,s}. \quad (28)$$

For $n = M$, (27) holds except $\vec{Q}^{\dagger M,4} = (\Delta t/6)\vec{G}_\Omega^{\dagger M,4}$. $\mathbf{T}[\vec{Q}^{\dagger n,s}]$ is the adjoint coefficient matrix and can be determined by linearizing (5). Appendix A includes a synopsis of this procedure and the result. The superscripts on the adjoint coefficient matrix in (27) indicate that it is evaluated at the physical time at the end of the time-reversed sub-step, which is not restrictive since it only depends on the state variable available from the forward simulation. Hence, it is known at all times $t_0 \leq t \leq t_1$, though it may pose a data management challenge in practice.

As formulated, the coefficients $\beta^{n,s}$ provide a family of self-consistent adjoint equations per the norm defined in (24). However, to be useful for the control problem that we have formulated (or any corresponding objective), the discrete norm defined in (24) should provide a consistent approximation of an objective function such as (9) with inner product such as (13). Note that in applications where it is more suitable to choose a discrete cost functional, this step is unnecessary. For our mixing layer demonstration, the control and target mollifying support functions defined in (10) and their derivatives in time are zero at the start and end of the control interval. Hence, we choose $\beta^{1,s} = \beta^{2,s} = \dots = \beta^{M,s}$ while retaining enough flexibility to ensure that the discrete norm is an accurate approximation of (9). With this assumption, for the discrete norm to be consistent, $\sum_{s=1}^4 \beta^{n,s} = 1$. Additionally, we require $\beta^{n,s} > 0$ to retain positive-definiteness.

Within the subset of discrete-adjoint time-marching schemes that satisfy the constraints imposed on the coefficients $\beta^{n,s}$, there is only one choice that is at least an $\mathcal{O}(\Delta t)$ approximation of the continuous-adjoint equations: $\beta^{n,1} = \beta^{n,4} = 1/6$ and $\beta^{n,2} = \beta^{n,3} = 1/3$. In general, this choice leads to an $\mathcal{O}(\Delta t)$ quadrature rule, but for the essentially homogeneous end conditions imposed by the specific control and target mollifying support functions in (10), its accuracy is effectively the same as pseudo-spectral integration. The resulting discrete-adjoint time-marching scheme is

$$\vec{M}^{\dagger n,4} = \frac{6\vec{Q}^{\dagger n,4} + 2\vec{Q}^{\dagger n+1,4} - 2\vec{Q}^{\dagger n+1,3} - 4\vec{Q}^{\dagger n+1,2} - 2\vec{Q}^{\dagger n+1,1}}{\Delta t} - [\mathbf{P}^{-1}\mathbf{T}^{n,4}\mathbf{P}]\vec{Q}^{\dagger n+1,1} - \vec{G}_{\Omega}^{\dagger n,4} \quad (29a)$$

$$\vec{M}^{\dagger n,3} = \frac{2\vec{Q}^{\dagger n,3} - 2\vec{Q}^{\dagger n,4}}{\Delta t} - [\mathbf{P}^{-1}\mathbf{T}^{n,3}\mathbf{P}]\vec{Q}^{\dagger n,4} - [\mathbf{2}\vec{G}_{\Omega}^{\dagger n,3}] \quad (29b)$$

$$\vec{M}^{\dagger n,2} = \frac{2\vec{Q}^{\dagger n,2} - 2\vec{Q}^{\dagger n,4}}{\Delta t} - [\mathbf{P}^{-1}\mathbf{T}^{n,2}\mathbf{P}]\vec{Q}^{\dagger n,3} - \vec{G}_{\Omega}^{\dagger n,2} \quad (29c)$$

$$\vec{M}^{\dagger n,1} = \frac{\vec{Q}^{\dagger n,1} - \vec{Q}^{\dagger n,4}}{\Delta t} - [\mathbf{P}^{-1}\mathbf{T}^{n,1}\mathbf{P}]\vec{Q}^{\dagger n,2} - [\frac{1}{2}\vec{G}_{\Omega}^{\dagger n,1}] \quad (29d)$$

for $n = 1, \dots, M$ except $\vec{Q}^{\dagger M,4} = [\Delta t/6]\vec{G}_{\Omega}^{M,4}$

Some terms are boxed in (29) to emphasize the difference from approximating the time derivative in (18) with the RK4 method directly,

$$\vec{M}^{\dagger n,4} = \frac{6\vec{Q}^{\dagger n,4} + 2\vec{Q}^{\dagger n+1,4} - 2\vec{Q}^{\dagger n+1,3} - 4\vec{Q}^{\dagger n+1,2} - 2\vec{Q}^{\dagger n+1,1}}{\Delta t} - \mathbf{T}^{n+1,1}\vec{Q}^{\dagger n+1,1} - \vec{G}_{\Omega}^{\dagger n,4} \quad (30a)$$

$$\vec{M}^{\dagger n,3} = \frac{2\vec{Q}^{\dagger n,3} - 2\vec{Q}^{\dagger n,4}}{\Delta t} - \mathbf{T}^{n,4}\vec{Q}^{\dagger n,4} - \vec{G}_{\Omega}^{\dagger n,3} \quad (30b)$$

$$\vec{M}^{\dagger n,2} = \frac{2\vec{Q}^{\dagger n,2} - 2\vec{Q}^{\dagger n,4}}{\Delta t} - \mathbf{T}^{n,3}\vec{Q}^{\dagger n,3} - \vec{G}_{\Omega}^{\dagger n,2} \quad (30c)$$

$$\vec{M}^{\dagger n,1} = \frac{\vec{Q}^{\dagger n,1} - \vec{Q}^{\dagger n,4}}{\Delta t} - \mathbf{T}^{n,2}\vec{Q}^{\dagger n,2} - \vec{G}_{\Omega}^{\dagger n,1} \quad (30d)$$

for $n = 1, \dots, M$ except $\vec{Q}^{\dagger M,4} = 0$. We anticipate that these differences can be eliminated by using an SBP time integration scheme [40]. Our choice of the coefficients $\beta^{n,s}$ in (27) facilitates a straightforward implementation, requiring only these changes from an RK4 discretization of the continuous-adjoint equations. In addition, for centered finite-difference schemes, the transpose that leads to the adjoint discrete derivative operators in (A.1) only require a modification near the boundaries. This similarity is a key advance of our formulation.

A practical matter when implementing (29) involves data management. The conserved variables appear in the adjoint equations (A.1) as coefficients and must be stored when the forward problem is solved. For large problem sizes, some sort of checkpointing approach is necessary to provide sufficient storage space [41]. The explicit formulation for the flow state coefficients at the sub-steps is a key result; there is no such specification in the continuous adjoint formulation. The RK4 scheme (for example) does not provide a fourth-order estimate of the state at fractional sub-steps such as $t_{n+1/2}$. In this case, a consistent approach is to regard the stored values as available data points and interpolate to obtain the sub-step values of the state variable. For our demonstrations, when solving the continuous-adjoint equations, $\vec{Q}(t_n)$, $\vec{Q}(t_n + \Delta t)$, $\frac{d\vec{Q}}{dt}(t_n)$, and $\frac{d\vec{Q}}{dt}(t_n + \Delta t)$ are interpolated with cubic Hermite polynomials.

The computational cost of the discrete-adjoint method is connected to the choice of discretization of the forward model. Our discrete-adjoint method is modestly more computationally expensive than the continuous-adjoint method, which we quantify for discussion here based upon the number of field derivatives since these constitute most of the computational effort. For a three-dimensional problem, there are 72 derivative operations in (A.1) compared to a discretization of the continuous-adjoint equations, which requires only 54. The additional terms are not in the corresponding continuous formulation because continuous first- and second-derivative operators are skew-Hermitian and Hermitian, respectively, which affords cancellation of certain combinations of terms that result from the integration-by-parts procedure. The same cancellation is not exact for the discrete operators, which leads to (A.1), and a modest increase in required operations. This expression has been simplified to exploit the linearity and distributive property of adjoint discrete derivative operators.

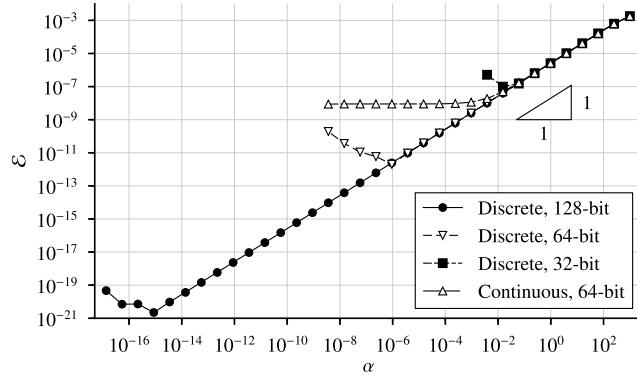


Fig. 4. Gradient accuracy per (12) for different arithmetic precisions for the acoustic cancellation test [17].

It can be further simplified in some obvious ways to exploit the properties of SBP operators, but we do not pursue these in detail since they are for the special case of a uniform mesh.

4. Results and discussion

4.1. Assessment of gradient accuracy for a two-dimensional model flow

We first confirm the efficacy of the discrete-adjoint formulation based on the gradient error measure defined in (12) for an acoustic cancellation test case [17]. That reference includes the full details of a noise source and the optimization problem formulated to cancel sound on a target. This is effectively a deterministic solution of the wave equation, so we do not expect the amplification of errors to occur via the chaotic mechanisms available in the nonlinear and turbulent flows we consider subsequently.

Fig. 4 shows the gradient error \mathcal{E} versus α from (12) for a continuous-adjoint simulation of the anti-sound test case performed with 64-bit (double-precision) floating-point arithmetic, and discrete-adjoint simulations performed with 32-, 64- and 128-bit floating-point arithmetic. As $\alpha \rightarrow 0$, the continuous-adjoint shows approximately constant \mathcal{E} for larger α than the corresponding 64-bit discrete-adjoint indicating the point where the dominant contribution to the gradient error is from numerical discretization. For the discrete-adjoint simulations, $\mathcal{E} \sim \alpha$ up to a limit that is proportional to the computing precision.

We next consider a relatively deterministic, but nonlinear flow: the aeroacoustic control of the two-dimensional spatially developing mixing layer shown schematically in Fig. 5, the control of which has been studied extensively using the continuous-adjoint method by Wei and Freund [17]. It is a mixing layer between two streams of velocities $0.9a_\infty$ and $0.2a_\infty$ along the stream-wise (x_1) direction, excited at randomly selected frequencies around the nominally most unstable mode. It is simulated on a mesh with 960×640 grid points in the x_1 - and x_2 -directions respectively. The control and target mollifying support functions from (8) and (18) are

$$\begin{aligned} W_{\Gamma,4}^2(\mathbf{x}, t) &\sim B_{0,2}(x_1^\Gamma) B_{0,2}(x_2^\Gamma) r_\Gamma(t) \\ W_{\Omega,4}^2(\mathbf{x}, t) &\sim B_{0,2}(x_2^\Omega) [\tanh(120(x_1^\Omega - 0.05)) - \tanh(120(x_1^\Omega - 0.95))] r_\Omega(t), \end{aligned}$$

where $r_\Gamma(t) = 0.5[1 + \tanh(2000(\frac{t-t_0}{t_1-t_0} - 0.01))]$ and $r_\Omega(t) = 0.5[1 + \tanh(2000(\frac{t_1-t}{t_1-t_0} - 0.01))]$. Full details of this flow are available [17,42], and are not reported here.

We provide an assessment of the accuracy of the gradient for this problem using both the continuous- and discrete-adjoint methods. First, the flow is simulated until time $t_0 = 840\delta_w^0/a_\infty$ which marks the beginning of the control interval. Following this, 25 000 time steps are taken with $\Delta t = 0.025\delta_w^0/a_\infty$. All results for this flow are computing using 64-bit arithmetic.

Fig. 6 shows contours of $Q_5^\dagger(\mathbf{x}, t_0)$ obtained using the continuous- and discrete-adjoint methods. Clearly the two fields are nearly the same for this high-resolution simulation, which explains the success of the continuous-adjoint for control [17], despite the limitations we show in Section 4.4 for true turbulence. However, some differences are noted, such as the appearance of high spatial wavenumber features in the discrete-adjoint. Similar behavior has been reported before [43], and has been attributed to the sensitivity gradient corresponding to computational modes supported by the forward discretization. It is characteristic of a dual-inconsistent discretization of the flow equations [44].

Fig. 7 shows a comparison between the accuracy of the gradients obtained using the continuous- and discrete-adjoint methods. The continuous-adjoint method is consistent for about 3 decades of change in α . As $\alpha \rightarrow 0$, the discrete formulation provides about a factor of 40 further reduction in the gradient truncation error per (12) for this flow. A relatively close agreement is anticipated in this relatively well-resolved and relatively deterministic flow. We shall see that a truly

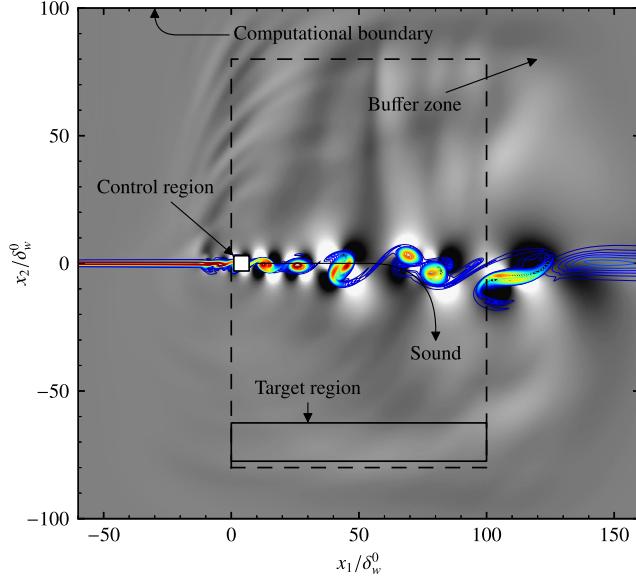


Fig. 5. The two-dimensional spatially developing mixing layer visualized with vorticity magnitude $|\nabla \times \mathbf{u}|$ (colors) and dilatation $\nabla \cdot \mathbf{u}$ (grays) at $t = t_0$. The extent of the control region is $[\delta_w^0, 7\delta_w^0] \times [-3\delta_w^0, 3\delta_w^0]$ and that of the target region is $[0, 100\delta_w^0] \times [-77.5\delta_w^0, -62.5\delta_w^0]$.

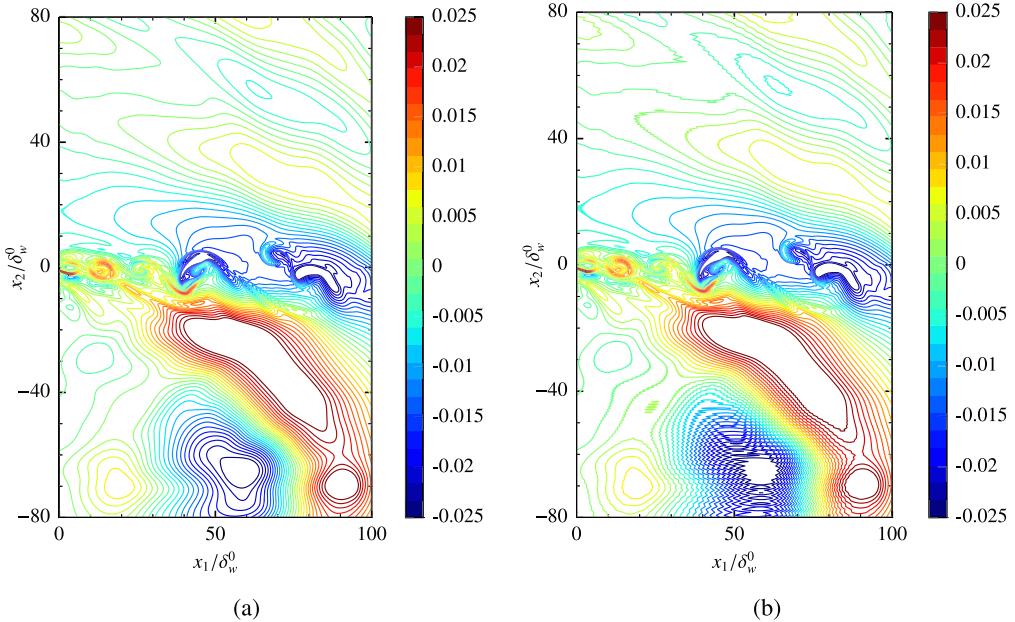


Fig. 6. Contours of $Q_5^\dagger(\mathbf{x}, t_0)$ for a two-dimensional spatially developing mixing layer using the (a) continuous-, and (b) discrete-adjoint formulations.

turbulent, three-dimensional flow will lead to a very different conclusion. Based upon validation for the anti-sound configuration, we can anticipate that the discrete-adjoint provides accuracy up to the point where finite-precision effects become significant. Though such discrepancies are expected to grow relatively slowly in this relatively deterministic two-dimensional flows, a precise assessment of this is not possible.

4.2. Baseline turbulent mixing layer

We now revisit the three-dimensional mixing layer introduced in Section 2. Before we apply the control, the mixing layer spreads by a factor of 19, which affords it time to evolve away from the initial condition, and it spreads again by a factor of 2.8 during our control horizon. Mean flow and Reynolds stress turbulence statistics confirm that this time is sufficient for it to achieve an approximately self-similar state of realistic turbulence for our study. We are interested in

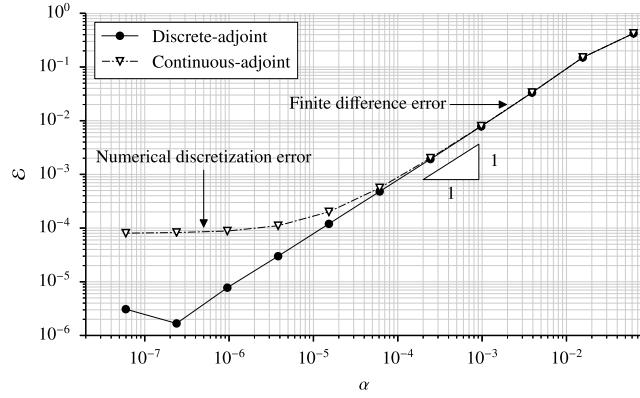


Fig. 7. Gradient accuracy for the two-dimensional spatially developing mixing layer shown in Fig. 5.

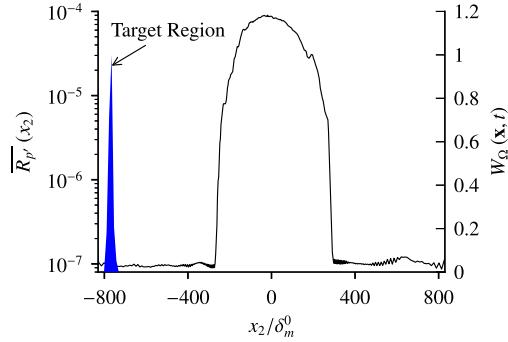


Fig. 8. Residual of the linearized Euler equation applied to the pressure computed from the DNS data at $t = t_1$. The target mollifier $W_Q(\mathbf{x}, t)$ is also shown.

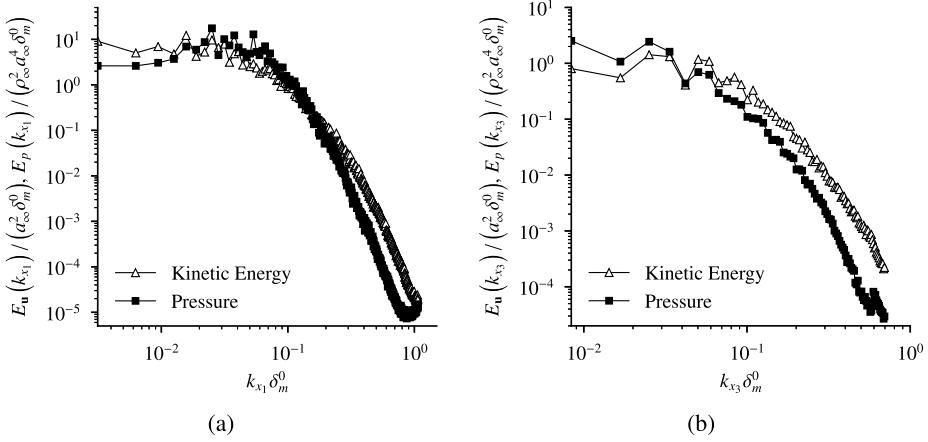


Fig. 9. One-dimensional spectra at $x_2 = 0$ at control onset time t_0 : (a) stream-wise, and (b) span-wise.

controlling the radiated sound, so $W_Q^2(\mathbf{x}, t)$ in (8) is selected to have support only in the acoustic field. To confirm that its support is sufficiently far from the turbulence to achieve this, we compute the residual of the linearized Euler equations for the uniform free-stream flow, which should become small in the acoustic field. Fig. 8 shows this is so even at the end of the control interval when the mixing layer has spread furthest. From its peak value at the mixing layer, the residual drops significantly suggesting that nonlinearity is indeed negligible in the target region around $x_2 = -765\delta_m^0$.

Fig. 9 shows the kinetic energy and pressure spectra at the beginning of the control interval. Our goal was to challenge the control by applying it on a realistically broad-banded turbulent flow. The stream-wise and span-wise spectra show significant energy over a broad range of wavenumbers, which drops at high wavenumbers due to viscosity, though a distinct $k \propto -5/3$ inertial range is not expected for $\text{Re}_{\delta_m} = 926$.

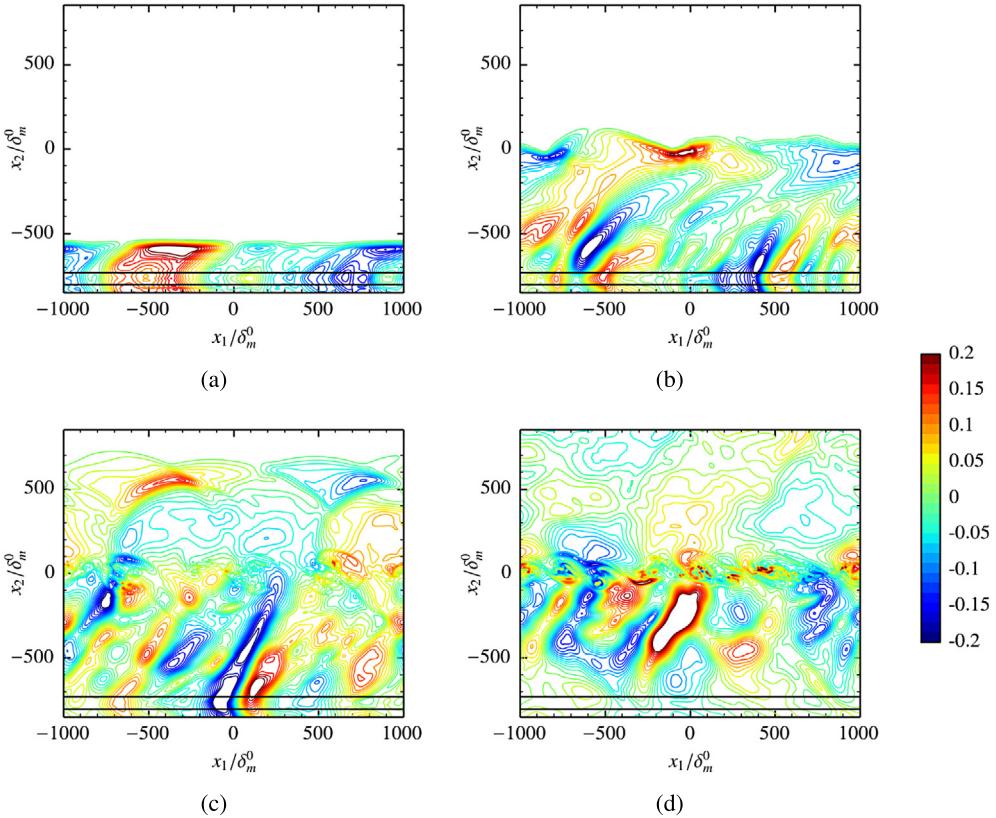


Fig. 10. Visualization of $Q_5^\dagger(\mathbf{x}, t)$ from the discrete-adjoint method for $x_3/\delta_m^0 = 0$ at times $(t - t_0)a_\infty/\delta_m^0 =$: (a) 1800, (b) 1200, (c) 600 and (d) 0. The straight lines outline the boundary of finite W_{Ω} .

4.3. Comparison between continuous- and discrete-adjoint solutions

Fig. 10 shows a visualization of $Q_5^\dagger(\mathbf{x}, t)$ from the discrete-adjoint formulation. The equation for $Q_5^\dagger(\mathbf{x}, t)$ in (29) is forced in the target region, which is in the acoustic field. Since the flow equations are self-adjoint in the acoustic limit, this quantity is expected to behave approximately like a sound wave until it interacts with the mixing layer.

Given that the shape of the cost functional in the control variable space is in general unknown, it is difficult to anticipate in detail the implications of an erroneous gradient for the convergence to the correct optimal control of the discrete model system. At best, errors will slow convergence by providing an erroneous amplitude in approximately the correct direction, but in general they will lead to the wrong solution, which is more consequential. We thus specifically assess error in the gradient direction via

$$\varphi(t) = \cos^{-1} \left(\frac{\langle \vec{G}_{\text{continuous}}, \vec{G}_{\text{discrete}} \rangle_{\mathbf{P}}}{\|\vec{G}_{\text{continuous}}\|_{\mathbf{P}} \|\vec{G}_{\text{discrete}}\|_{\mathbf{P}}} \right), \quad (31)$$

where $\vec{G}_{\text{discrete}}$ is from (26), and $\vec{G}_{\text{continuous}}$ is from the discretization of (20). Fig. 11 shows a plot of this angle as a function of time. At around $(t - t_0)a_\infty/\delta_m^0 = 1150$, the adjoint encounters the mixing layer and excites instabilities. Until this time, the magnitude of the sensitivity gradient, which has support in the control region located in the mixing layer, is trivially small. For $(t - t_0)a_\infty/\delta_m^0 < 1150$, we observe that the error in the direction of the sensitivity gradient increases until it reaches its maximum value of about 10° at $t = t_0$. For the two-dimensional spatial mixing layer demonstration in Section 4.1, we determined the maximum value of $\varphi(t)$ to be 1.2° , which supports the rapid convergence of the continuous-adjoint-based noise optimization achieved for this flow by Wei and Freund [17].

It is expected that any gradient errors will increase in time due to the usual sensitivity to initial conditions of a chaotic system. However, just how this occurs and how it affects the utility of any formulation for flow control is unclear. Sound generation is primarily by the energetic large scales in a flow, and controlling these can be anticipated to be most important for its mitigation. So a gradient that is accurate for only low wavenumbers might still be able to provide the essential information for control optimization. Without a fundamental theory of turbulence, it is unlikely that general conclusions can be drawn in this regard, but we can make an empirical assessment for our model turbulent flow. Since $\vec{G}_{\text{continuous}}$ is obtained

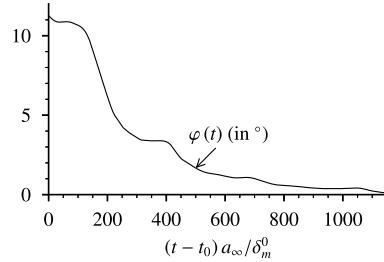


Fig. 11. Time-dependent gradient direction error (31).

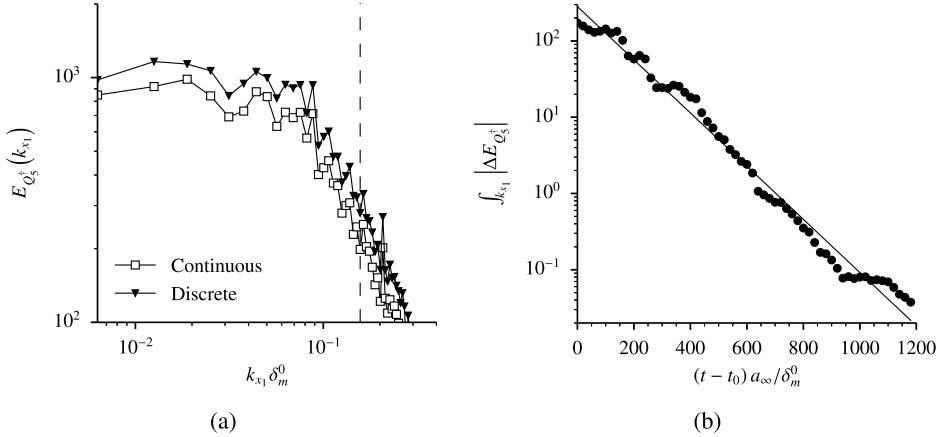


Fig. 12. Results for the three-dimensional, turbulent mixing layer. (a) One-dimensional stream-wise energy spectra of Q_5^\dagger at $x_2/\delta_m^0 = 0$ at $t = t_0$, and (b) difference in the magnitude of the same quantity averaged between the lowest wavenumber and the wavenumber shown by the dashed line in (a) as a function of time.

from an accurate discretization of (20), errors will be most pronounced at small scales. By comparison with the exact adjoint of the discretized system $\bar{G}_{\text{discrete}}$, we can measure how these grow in time, as expected for a chaotic flow, and more importantly how they affect the accuracy of the arguably more important large-scale components of the gradient. Of course, this is an exercise rather than a direct concern since the discrete-adjoint we derive will be preferable for this flow. However, not all models will be so amenable to an exact discrete-adjoint formulation, so this specific question may be a practical concern in some circumstances. It is unclear, for example, that a general sub-grid-scale model of viscous dissipation or scalar transport for large-eddy simulation will necessarily be so easily expressed such that its discrete adjoint is efficient or even tractable. Thus, we evaluate the accuracy of the continuous-adjoint against the discrete-adjoint to anticipate challenges when the discrete-adjoint is unavailable. The discussion in Section 5 will also revisit past efforts that used the continuous-adjoint formulation.

To quantify the scale-dependence of the adjoint solution errors, the energy spectra of the continuous- and discrete-adjoint Q_5^\dagger are shown in Fig. 12(a). The magnitude of this difference averaged between wavelengths $40\delta_m^0$ and $2000\delta_m^0$ increases exponentially in reverse-time (Fig. 12(b)). It is clear that despite the initiation of errors at small scales, over the course of this simulation forcing of even the largest turbulence scales would be affected by continuous-adjoint errors (note that based on Fig. 3, the error in the discrete-adjoint should be significantly smaller than any discretization errors). By separating turbulent fluctuations in flow quantities in (21) from their mean values, it can be shown that adjoint variables are transported like a passive scalar by the turbulent mean flow. This suggests the existence of an energy cascade mechanism for the energy norm for Q_5^\dagger , which is transported by the mean flow velocity. Since the adjoint equations are integrated in reverse-time, we should expect the cascade to transport energy from small to large spatial scales as the adjoint simulation proceeds. This is evident from Fig. 12(a), which shows that the energy norm for the discrete-adjoint Q_5^\dagger is consistently higher than the continuous-adjoint, and the difference is significant even at the largest spatial scales, against the nominal expectation that the exact sensitivity of the largest spatial scales are accurately captured by the continuous-adjoint method. Thus, even if the sound generation and its control is expected to be relatively large-scale, the accumulation of discretization errors will potentially hinder efficacy, and this in turn is potentially sensitive to Reynolds numbers. For longer simulation times, we can anticipate that the finite numerical precision errors will also similarly propagate to all scales, though this would require longer simulation times than this particular flow is set up for.

Table 1
Noise reduction after one line search based upon a single gradient.

	Cost functional $\mathcal{J} \times 10^5$ from (9)
Baseline (no control)	11.0
Continuous-adjoint control	10.2
Discrete-adjoint control	9.23

4.4. Controlled mixing layer

Though the continuous-adjoint is clearly less accurate, its full effect cannot be completely anticipated because of the complexity of the turbulence. We therefore also analyze the beginning of a control optimization. A steepest descent line-search was performed with the sensitivity gradient obtained from the continuous- and discrete-adjoint methods to determine the effect of the error in sensitivity gradient on noise reduction. To find the local minimum along the descent direction, a minimum bracket is evaluated followed by a modification of Brent's method [45] that uses derivatives to isolate the minimum to a fractional precision of 0.01. Table 1 shows the values of the baseline cost functional, and the cost functional with a control corresponding to the local minimum obtained using this procedure with the continuous- and discrete-adjoint methods for one line-search. The gradient errors quantified in the previous section are indeed important by this measure. The discrete-adjoint method achieves twice the reduction.

5. Conclusions

A principal result is the discrete-adjoint formulation we derived for the three-dimensional compressible flow equations. The potential advantages of such a formulation are well understood, but the advantage of this new formulation is that it provides the exact (precision-limited) gradient for the highly accurate workhorse methods commonly used for simulating compressible turbulence, particularly in aeroacoustic simulations. The formulation is crafted such that it exploits the properties of adjoint discrete derivative operators in algebraic combinations that produce an efficient computational code. Such opportunities are generally missed by current automatic differentiation tools because they involve combinations that are not easy for such tools to identify based upon a numerical implementation. At the same time, the operations mirror those that would be followed for a comparable discretization of the continuous-adjoint equations, and thus do not demand any intricate operations beyond those already available for the flow solution or its adjoint if a discretization of that is already available.

Although our demonstrations are oriented around aeroacoustic applications, the utility of the adjoint formulation (A.1) extends to other high-fidelity applications, since it is independent of the choice of cost functional. We have presented the method by contrasting it with a discretization of the continuous-adjoint method to highlight the fact that it can be practically implemented with only a few changes and that the formulation is modestly ($\sim 1.5 \times$) computationally more expensive than the continuous-adjoint method but more effective in controlling aeroacoustic noise ($\sim 2.2 \times$ for our model mixing layer flow). It is similarly only $\sim 1.85 \times$ more expensive than computing the flow solution.

The fully discrete adjoint method we have derived is compatible with flexible multi-stage temporal discretizations of the flow equations (23), such as the fourth-order Runge–Kutta scheme we employ in our demonstrations. The modular procedure that has been proposed [28] for the evaluation of the semi-discrete-adjoint does not seem so easily extendable to include time discretization. However, the time-exactness in the present formulation makes it more tied to the specific flow equations. Thus, if used only to provide an exact semi-discrete-adjoint for other equations (i.e., not the compressible flow equations), the modular approach would be more readily adapted.

Practical implementation of our discrete-adjoint formulation relies on the fact that the discrete operators in (5) are local, and the bandwidth of the interior stencils do not change when transposed as in (A.1) (the discrete-Fourier operators we use in x_3 are an exception, since their transpose yields also a discrete-Fourier operator). Thus (A.1) is useful for finite-differences, artificial dissipation based on high-order even derivatives, polynomial interpolation between overset grids and so on. However, spatially nonlocal operators such as dynamic large-eddy simulation (LES) models are not of this form. In this case, (A.1) will be most useful in a hybrid adjoint approach [46], in which a continuous formulation can be used for nonlocal operators, while retaining the discrete approach for the remaining terms in lieu of its compatibility with the discretized flow equations. Such an implementation would be straightforward since our choice of coefficients $\beta^{n,s}$ in (27) provides a fully-discrete adjoint (29) that is consistent with an RK4 discretization of the continuous-adjoint equations (30). Alternatively, for complicated nonlinear but relatively local operations, (A.1) can be used effectively in conjunction with automatic differentiation, which is less burdensome for such small sub-problems.

The error in the sensitivity gradient predicted by the continuous-adjoint method has been assessed in three different ways. First, we defined an error measure \mathcal{E} (12) that quantifies the difference between the sensitivity gradient predicted by our adjoint methods and a finite-difference approximation of it. As $\alpha \rightarrow 0$, we observed that the gradient error $\mathcal{E} \sim \alpha$ up to round-off errors using the discrete-adjoint method, which is the expected behavior for the exact sensitivity gradient. The actual error is the result of finite computing precision plus a degree of amplification via the nonlinearity in the chaotic flow.

However, its sensitivity to arithmetic precision was ascertained for a two-dimensional flow, for which case the gradient precision was shown to be proportional to the precision of the underlying arithmetic.

Second, we quantified the error in a generalized gradient direction φ , which is anticipated to be particularly important for seeking a local optimum. In our compressible turbulent mixing layer demonstration, we showed that φ increases throughout the duration of the control interval for which the adjoint \vec{Q}^\dagger waves interact with the mixing layer, and reaches a maximum value of about 10° for the time simulated. This time horizon was practical for our turbulence simulation since it corresponds to the time over which the turbulence could be realistically simulated in the computational box. For longer simulations of other flows, there is every expectation that this error would continue to increase due to the chaotic character of the flow.

Finally, we considered the spatial scales of the error. The difference between the discrete- and continuous-adjoint formulations is of the order of the truncation error of the numerical approximation, and will therefore for any particular field be larger for shorter and therefore more challenging-to-resolve wavelengths. However, our results for the evolution equations show that the error in the sensitivity gradient becomes large across the entire range of scales. The broad-banded difference is not a direct consequence of nonlinearity, since the adjoint equations are linear, but is a manifestation of the nonlinearity of the flow equations, which appears in the adjoint as space and time dependent coefficients with a similar effect of coupling the adjoint solution across all scales. Further, the difference grows exponentially in reverse-time. Many efforts in optimal flow control have focused on laminar flows or two-dimensional flows, where high accuracy is relatively easy to achieve and the flows are often relatively deterministic because they lack of all the mechanical features of true three-dimensional turbulence. How such errors in a continuous-adjoint formulation contaminate the large-scale components of the gradient is important for two reasons. The first is physical: the large scales are the more likely points of flow control. The second involves modeling and the challenge of developing sub-grid-scale (or other) turbulence models for which practical exact adjoints are available. If errors are incurred, say, by using a hybrid formulation [46], they can be anticipated to contaminate the solution at many scales. This study shows what effects might be anticipated by the inherent approximations.

Our results have implications for some recent flow control simulations. Using a continuous-adjoint formulation, Kim et al. [19] found that even with a general control, the overall noise reduction achieved for a turbulent Mach 1.3 jet was significantly lower than in a previous two-dimensional study [17]. Taking our plane, temporally developing turbulent mixing layer as a model for the initial shear layers of a turbulent jet, we can estimate the deleterious effects of the continuous-adjoint used in that study. We see a significant error for the continuous formulation as the momentum thickness increases from about $19\delta_m^0$ to $53\delta_m^0$, a factor of 2.8, during the control interval. This growth is comparable to the momentum thickness growth of $2.75\times$ along the potential core length of their simulated turbulent jet. This suggests that the instantaneous continuous-adjoint-based sensitivity for a nozzle control will entail a similar 10° error by $x/D \sim 6$, which is the region of peak noise generation. Given the reduced effectiveness even in the present simpler flow, we can anticipate that the errors in the gradient might be more significant in the jet. It should also be noted that convergence was slow and not pursued through to completion by Kim et al. [19]. Significant slowing of convergence, relative to the earlier two-dimensional flows [17], might have also masked the potential success of the control.

An important result here is that the discrete-adjoint formulation was found to be sufficiently accurate (e.g., Fig. 3) for the control horizons needed for our aeroacoustic objectives. However, for longer times the fidelity would eventually succumb to the accumulation of round-off errors, amplified by the chaotic character of the turbulence. The amplification of the difference between the discrete- and continuous-adjoints in Fig. 12 shows how this will occur. These round-off errors are complicated functions of the coefficients of our finite-difference schemes and the adjoint coefficient matrix, and are neither uncorrelated, nor uniformly distributed on either side of zero on the real line. However, they are roughly proportional to the number of floating point operations and will undoubtedly accumulate to the point of overwhelming the exactness of the discrete formulation for sufficiently long times. Estimating or designing numerical methods to reduce or eliminate such round-off errors is beyond the scope of this work.

Acknowledgements

This work is supported [in part] by the Office of Naval Research grant number N00014-11-1-0756. The computations were performed on Stampede HPC of the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number TG-CTS090004. This material is also based [in part] upon work supported by the Department of Energy, National Nuclear Security Administration, under Award Number(s) DE-NA0002374.

Appendix A. Discrete-adjoint of the compressible flow equations

Here, we outline the procedure to derive the adjoint coefficient matrix defined in (28) for the compressible flow equations. The result is expressed in terms of an operator $\vec{R}^\dagger[\vec{Q}, \vec{Q}^\dagger](t)$ defined implicitly using $\vec{R}^\dagger[\vec{Q}, \vec{Q}^\dagger](t) = -T[\vec{Q}](t)\vec{Q}^\dagger(t)$. In (29), the adjoint coefficient matrix always appears within the product $\mathbf{P}^{-1}\mathbf{T}[\vec{Q}](t)\mathbf{P}\vec{Q}^\dagger(t)$, which is equivalent to $-\mathbf{P}^{-1}\vec{R}^\dagger[\vec{Q}, \mathbf{P}\vec{Q}^\dagger](t)$.

Linearizing (5) and using (25) leads to

$$\begin{aligned} \vec{R}_1^{\dagger}[\vec{Q}, \vec{Q}^{\dagger}](t) = & \left(\delta_{ij} \frac{\gamma - 1}{2} \vec{u}_k \circ \vec{u}_k - \vec{u}_i \circ \vec{u}_j \right) \circ (\mathbf{D}_{1x_j}^T \vec{Q}_{i+1}^{\dagger}) \\ & + \left[\frac{\gamma - 1}{2} \vec{u}_i \circ (\vec{u}_k \circ \vec{u}_k) - \frac{\vec{u}_i}{\bar{\rho}} \circ (\vec{e} + \vec{p}) \right] \circ (\mathbf{D}_{1x_i}^T \vec{Q}_5^{\dagger}) \\ & + \frac{1}{\text{Re}} \frac{\vec{u}_i}{\bar{\rho}} \circ (\mathbf{L}^T \vec{Q}_{i+1}^{\dagger}) + \frac{1}{3 \text{Re}} \frac{\vec{u}_j}{\bar{\rho}} \circ (\Delta_{ij}^T \vec{Q}_{i+1}^{\dagger}) + \frac{1}{\text{Re}} \frac{\vec{u}_i}{\bar{\rho}} \circ [\mathbf{L}^T (\vec{u}_i \circ \vec{Q}_5^{\dagger})] \\ & + \frac{2}{\text{Re}} \frac{\vec{u}_i}{\bar{\rho}} \circ \{ \mathbf{D}_{1x_j}^T [(\mathbf{D}_{1x_i} \vec{u}_j + \mathbf{D}_{1x_j} \vec{u}_i) \circ \vec{Q}_5^{\dagger}] \} \\ & + \frac{1}{3 \text{Re}} \frac{\vec{u}_j}{\bar{\rho}} \circ [\mathbf{D}_{ij}^T (\vec{u}_i \circ \vec{Q}_5^{\dagger})] - \frac{4}{3 \text{Re}} \frac{\vec{u}_i}{\bar{\rho}} \circ \{ \mathbf{D}_{1x_i}^T [(\mathbf{D}_{1x_j} \vec{u}_j) \circ \vec{Q}_5^{\dagger}] \} \\ & - \frac{\gamma}{\text{Re} \Pr(\gamma - 1)} \left(\frac{\gamma - 1}{2} \frac{\vec{u}_k \circ \vec{u}_k}{\bar{\rho}} - \frac{\vec{p}}{\bar{\rho}^2} \right) \circ (\mathbf{L}^T \vec{Q}_5^{\dagger}) \\ & + \frac{1}{\text{Re}} \frac{\vec{u}_i}{\bar{\rho}} \circ \left[\left(\mathbf{L} \vec{u}_i + \frac{1}{3} \Delta_{ij} \vec{u}_j \right) \circ \vec{Q}_5^{\dagger} \right] + \vec{N}_{1,\text{SAT}}^{\dagger}[\vec{Q}^{\dagger}, \vec{Q}_b](t) \end{aligned} \quad (\text{A.1a})$$

$$\begin{aligned} \vec{R}_{i+1}^{\dagger}[\vec{Q}, \vec{Q}^{\dagger}](t) = & \mathbf{D}_{1x_i}^T \vec{Q}_1^{\dagger} - (\gamma - 1) \vec{u}_i \circ (\mathbf{D}_{1x_j}^T \vec{Q}_{j+1}^{\dagger}) + \vec{u}_j \circ (\mathbf{D}_{1x_i}^T \vec{Q}_{j+1}^{\dagger} + \mathbf{D}_{1x_j}^T \vec{Q}_{i+1}^{\dagger}) \\ & + \frac{\vec{e} + \vec{p}}{\bar{\rho}} \circ (\mathbf{D}_{1x_i}^T \vec{Q}_5^{\dagger}) - (\gamma - 1) (\vec{u}_i \circ \vec{u}_j) \circ (\mathbf{D}_{1x_j}^T \vec{Q}_5^{\dagger}) \\ & - \frac{1}{\text{Re}} \frac{\mathbf{L}^T (\vec{Q}_{i+1}^{\dagger} - \vec{u}_i \circ \vec{Q}_5^{\dagger})}{\bar{\rho}} + \frac{\gamma}{\text{Re} \Pr} \frac{\vec{u}_i}{\bar{\rho}} \circ (\mathbf{L}^T \vec{Q}_5^{\dagger}) \\ & - \frac{1}{\text{Re}} \left(\frac{\mathbf{L} \vec{u}_i + \frac{1}{3} \Delta_{ij} \vec{u}_j}{\bar{\rho}} \right) \circ \vec{Q}_5^{\dagger} - \frac{1}{3 \text{Re}} \frac{\Delta_{ji}^T (\vec{Q}_{j+1}^{\dagger} + \vec{u}_j \circ \vec{Q}_5^{\dagger})}{\bar{\rho}} \\ & + \frac{4}{3 \text{Re}} \frac{\mathbf{D}_{1x_i}^T [(\mathbf{D}_{1x_j} \vec{u}_j) \circ \vec{Q}_5^{\dagger}]}{\bar{\rho}} - \frac{2}{\text{Re}} \frac{\mathbf{D}_{1x_j}^T [(\mathbf{D}_{1x_i} \vec{u}_j + \mathbf{D}_{1x_j} \vec{u}_i) \circ \vec{Q}_5^{\dagger}]}{\bar{\rho}} \\ & + \vec{N}_{i+1,\text{SAT}}^{\dagger}[\vec{Q}^{\dagger}, \vec{Q}_b](t) \end{aligned} \quad (\text{A.1b})$$

$$\vec{R}_5^{\dagger}[\vec{Q}, \vec{Q}^{\dagger}](t) = (\gamma - 1) \mathbf{D}_{1x_i}^T \vec{Q}_{i+1}^{\dagger} + \gamma \vec{u}_i \circ (\mathbf{D}_{1x_i}^T \vec{Q}_5^{\dagger}) - \frac{\gamma}{\text{Re} \Pr} \frac{\mathbf{L}^T \vec{Q}_5^{\dagger}}{\bar{\rho}} + \vec{N}_{5,\text{SAT}}^{\dagger}[\vec{Q}^{\dagger}, \vec{Q}_b](t), \quad (\text{A.1c})$$

where

$$\vec{N}_{\text{SAT}}^{\dagger}[\vec{Q}^{\dagger}, \vec{Q}_b](t) = \boldsymbol{\Xi}^T[\vec{Q}_b] \vec{Q}^{\dagger}(t),$$

with

$$\boldsymbol{\Xi}^T[\vec{Q}_b] = -\sigma^I \sum_{i=1}^3 \mathbf{P}_i^{-1} (\mathbf{E}_{1i} \boldsymbol{\Theta}_i^{-1}[\vec{Q}_b] \boldsymbol{\Lambda}_i^-[\vec{Q}_b] \boldsymbol{\Theta}_i[\vec{Q}_b] - \mathbf{E}_{0i} \boldsymbol{\Theta}_i^{-1}[\vec{Q}_b] \boldsymbol{\Lambda}_i^+[\vec{Q}_b] \boldsymbol{\Theta}_i[\vec{Q}_b]).$$

For $\sigma^I = 1$, Hicken and Xingg [47] showed that combining $\vec{N}_{\text{SAT}}^{\dagger}[\vec{Q}^{\dagger}, \vec{Q}_b](t)$ with boundary terms in an SBP expansion of first-derivative operators in (A.1) results in an SAT that consistently enforces adjoint characteristic far-field boundary conditions for \vec{Q}^{\dagger} .

References

- [1] T.R. Bewley, P. Moin, R. Temam, DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms, *J. Fluid Mech.* 447 (2001) 179–225, <http://dx.doi.org/10.1017/S0022112001005821>.
- [2] J.R.R.A. Martins, J.J. Alonso, J.J. Reuther, High-fidelity aerostructural design optimization of a supersonic business jet, *J. Aircr.* 41 (2004) 523–530, <http://dx.doi.org/10.2514/1.11478>.
- [3] A. Jameson, J.J. Alonso, J.J. Reuther, L. Martinelli, J.C. Vassberg, Aerodynamic shape optimization techniques based on control theory, in: *AIAA 29th Fluid Dynamics Conference*, Albuquerque, NM, 1998.
- [4] R. Becker, R. Rannacher, A feed-back approach to error control in finite element methods: basic analysis and examples, *East-West J. Numer. Math.* 4 (1996) 237–264.
- [5] Q. Wang, Forward and adjoint sensitivity computation of chaotic dynamical systems, *J. Comput. Phys.* 235 (2013) 1–13, <http://dx.doi.org/10.1016/j.jcp.2012.09.007>.

- [6] P. Luchini, A. Bottaro, Adjoint equations in stability analysis, *Annu. Rev. Fluid Mech.* 46 (1) (2014) 493–517, <http://dx.doi.org/10.1146/annurev-fluid-010313-141253>.
- [7] D.C. Hill, Adjoint systems and their role in the receptivity problem for boundary layers, *J. Fluid Mech.* 292 (1995) 183–204, <http://dx.doi.org/10.1017/S0022112095001480>.
- [8] J.B. Freund, Adjoint-based optimization for understanding and suppressing jet noise, *J. Sound Vib.* 330 (17) (2011) 4114–4122, <http://dx.doi.org/10.1016/j.jsv.2011.02.009>.
- [9] J.B. Freund, Numerical simulation of a Mach 1.92 turbulent jet and its sound field, *AIAA J.* 38 (11) (2000) 2023–2031, <http://dx.doi.org/10.2514/2.889>.
- [10] C. Pantano, S. Sarkar, A study of compressibility effects in the high-speed turbulent shear layer using direct simulation, *J. Fluid Mech.* 451 (2002) 329–371, <http://dx.doi.org/10.1017/S0022112001006978>.
- [11] R.R. Kleiman, J.B. Freund, The sound from mixing layers simulated with different ranges of turbulence scales, *Phys. Fluids* 20 (2008) 101503, <http://dx.doi.org/10.1063/1.3005823>.
- [12] X. Zhong, High-order finite-difference schemes for numerical simulation of hypersonic boundary-layer transition, *J. Comput. Phys.* 144 (1998) 662–709, <http://dx.doi.org/10.1006/jcph.1998.6010>.
- [13] E.R. Hawkes, R. Sankaran, J.C. Sutherland, J.H. Chen, Direct numerical simulation of turbulent combustion: fundamental insights towards predictive models, *J. Phys. Conf. Ser.* 16 (1) (2005) 65–79.
- [14] T. Colonius, S.K. Lele, Computational aeroacoustics: progress on nonlinear problems of sound generation, *Prog. Aerosp. Sci.* 40 (2004) 345–416, <http://dx.doi.org/10.1016/j.paerosci.2004.09.001>.
- [15] M. Wang, J.B. Freund, S.K. Lele, Computational prediction of flow-generated sound, *Annu. Rev. Fluid Mech.* 38 (2006) 483–512, <http://dx.doi.org/10.1146/annurev.fluid.38.050304.092036>.
- [16] A. Jameson, Aerodynamic shape optimization using the adjoint method, in: *VKI Lecture Series on Aerodynamic Drag Prediction and Reduction*, 2003.
- [17] M. Wei, J.B. Freund, A noise-controlled free shear flow, *J. Fluid Mech.* 546 (2006) 123–152, <http://dx.doi.org/10.1017/S0022112005007093>.
- [18] R.R. Kleiman, J.B. Freund, Adjoint-based control of the noise from a turbulent mixing layer, in: *12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustic Conference)*, Cambridge, MA, 2006.
- [19] J. Kim, D.J. Bodony, J.B. Freund, Adjoint-based control of loud events in a turbulent jet, *J. Fluid Mech.* 741 (2014) 28–59, <http://dx.doi.org/10.1017/jfm.2013.654>.
- [20] S.K. Nadarajah, A. Jameson, A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization, in: *38th Aerospace Sciences Meeting and Exhibit*, Reno, NV, 2000.
- [21] A. Carnarius, F. Thiele, E. Özkaya, N.R. Gauger, Adjoint approaches for optimal flow control, *AIAA Paper 2010-5088*, in: *5th Flow Control Conference*, Chicago, IL, 2010.
- [22] D.J. Lea, M.R. Allen, T.W.N. Haine, Sensitivity analysis of the climate of a chaotic system, *Tellus A* 52 (2000) 523–532.
- [23] M.P. Rumpfkeil, D.W. Zingg, A general framework for the optimal control of unsteady flows with applications, in: *45th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, 2007.
- [24] N.K. Yamaeleev, B. Diskin, E.J. Nielsen, Adjoint-based methodology for time-dependent optimization, in: *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, British Columbia, 2008, <http://dx.doi.org/10.2514/6.2008-5857>.
- [25] L. Wang, D.J. Mavriplis, K.W. Anderson, Unsteady discrete adjoint formulation for high-order discontinuous Galerkin discretizations in time-dependent flow problems, in: *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, Orlando, FL, 2010.
- [26] R. Roth, S. Ulbrich, A discrete adjoint approach for the optimization of unsteady turbulent flows, *Flow Turbul. Combust. (ISSN 1386-6184)* (2013), <http://dx.doi.org/10.1007/s10494-012-9439-3>, 21 pp.
- [27] E.J. Nielsen, B. Diskin, Discrete adjoint-based design for unsteady turbulent flows on dynamic overset unstructured grids, *AIAA J.* 51 (2013) 1355–1373, <http://dx.doi.org/10.2514/1.J051859>.
- [28] M.F. de Pando, D. Sipp, P.J. Schmid, Efficient evaluation of the direct and adjoint linearized dynamics from compressible flow solvers, *J. Comput. Phys.* 231 (2012) 7739–7755, <http://dx.doi.org/10.1016/j.jcp.2012.06.038>.
- [29] M.B. Giles, M.C. Duta, Algorithm developments for discrete adjoint methods, *AIAA J.* 41 (2) (2003) 198–205, <http://dx.doi.org/10.2514/2.1961>.
- [30] C.A. Mader, J.R.R.A. Martins, J.J. Alonso, E. van der Weide, ADjoint: an approach for the rapid development of discrete adjoint solvers, *AIAA J.* 46 (2008) 863–873, <http://dx.doi.org/10.2514/1.29123>.
- [31] K. Duraisamy, J. Alonso, Adjoint based techniques for uncertainty quantification in turbulent flows with combustion, in: *42nd AIAA Fluid Dynamics Conference and Exhibit*, 2012.
- [32] S.K. Lele, Compressibility effects on turbulence, *Annu. Rev. Fluid Mech.* 26 (1994) 211–254, <http://dx.doi.org/10.1146/annurev.fl.26.010194.001235>.
- [33] R.R. Kleiman, On the turbulence-generated sound and control of compressible mixing layers, PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL, 2010.
- [34] K. Mattsson, J. Nordström, Summation by parts operators for finite difference approximations of second derivatives, *J. Comput. Phys.* 199 (2) (2004) 503–540, <http://dx.doi.org/10.1016/j.jcp.2004.03.001>.
- [35] M. Svärd, M.H. Carpenter, J. Nordström, A stable high-order finite difference scheme for the compressible Navier-Stokes equations, far-field boundary conditions, *J. Comput. Phys.* 225 (1) (2007) 1020–1038, <http://dx.doi.org/10.1016/j.jcp.2007.01.023>.
- [36] J.B. Freund, *Compressibility effects in a turbulent annular mixing layer*, PhD thesis, Stanford University, Stanford, CA, 1997.
- [37] T.H. Pulliam, D.S. Chaussee, A diagonal form of an implicit approximate-factorization algorithm, *J. Comput. Phys.* 39 (1981) 347–363.
- [38] S. Nadarajah, A. Jameson, Studies of the continuous and discrete adjoint approaches to viscous automatic aerodynamic shape optimization, in: *15th AIAA Computational Fluid Dynamics Conference*, 2001.
- [39] J.E. Hicken, D.W. Zingg, Summation-by-parts operators and high-order quadrature, *J. Comput. Appl. Math.* 237 (2013) 111–125, <http://dx.doi.org/10.1016/j.cam.2012.07.015>.
- [40] J. Nordström, T. Lundquist, Summation-by-parts in time, *J. Comput. Phys.* 251 (2013) 487–499, <http://dx.doi.org/10.1016/j.jcp.2013.05.042>.
- [41] A. Griewank, G.F. Corliss (Eds.), *Automatic Differentiation of Algorithms: Theory, Implementation and Application*, SIAM Proc. Ser., Society for Industrial and Applied Mathematics, 1991.
- [42] M. Wei, Jet noise control by adjoint-based optimization, PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL, 2004.
- [43] Z. Sirkes, E. Tziperman, Finite difference of adjoint or adjoint of finite difference?, *Mon. Weather Rev.* 125 (1997) 3373–3378.
- [44] J.E. Hicken, D.W. Zingg, Dual consistency and functional accuracy: a finite-difference perspective, *J. Comput. Phys.* 256 (2014) 161–182, <http://dx.doi.org/10.1016/j.jcp.2013.08.014>.
- [45] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 2007.
- [46] T.W.R. Taylor, F. Palacios, K. Duraisamy, J.J. Alonso, A hybrid adjoint approach applied to turbulent flow simulations, in: *21st AIAA Computational Fluid Dynamics Conference*, 2013.
- [47] J.E. Hicken, D.W. Xingg, Superconvergent functional estimates from summation-by-parts finite-difference discretizations, *SIAM J. Sci. Comput.* 33 (2) (2011) 893–922, <http://dx.doi.org/10.1137/100790987>.