

# Dual consistency and functional accuracy: a finite-difference perspective<sup>☆</sup>

J.E. Hicken<sup>a,1,\*</sup>, D.W. Zingg<sup>b,2</sup>

<sup>a</sup>*Department of Mechanical, Aerospace, and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, New York, United States*

<sup>b</sup>*Institute for Aerospace Studies, University of Toronto, Toronto, Ontario, M3H 5T6, Canada*

---

## Abstract

Consider the discretization of a partial differential equation (PDE) and an integral functional that depends on the PDE solution. The discretization is dual consistent if it leads to a discrete dual problem that is a consistent approximation of the corresponding continuous dual problem. Consequently, a dual-consistent discretization is a synthesis of the so-called discrete-adjoint and continuous-adjoint approaches. We highlight the impact of dual consistency on summation-by-parts (SBP) finite-difference discretizations of steady-state PDEs; specifically, superconvergent functionals and accurate functional error estimates. In the case of functional superconvergence, the discrete-adjoint variables do not need to be computed, since dual consistency on its own is sufficient. Numerical examples demonstrate that dual-consistent schemes significantly outperform dual-inconsistent schemes in terms of functional accuracy and error-estimate effectiveness. The dual-consistent and dual-inconsistent discretizations have similar computational costs, so dual consistency leads to improved efficiency. To illustrate the dual consistency analysis of SBP schemes, we thoroughly examine a discretization of the Euler equations of gas dynamics, including the treatment of the boundary conditions, numerical dissipation, interface penalties, and quadrature by SBP norms.

*Keywords:* dual consistency, adjoint consistency, summation-by-parts operators, functional superconvergence, adjoint-weighted residual method, differentiate-then-discretize, discretize-then-differentiate

---

## 1. Introduction

Simulations are often used to estimate the value of a functional that depends on the numerical solution of a partial differential equation (PDE). For example, in computational aerodynamics, the lift and drag forces are two functionals that are frequently computed based on the numerical

---

<sup>☆</sup>This work was supported by the Natural Sciences and Engineering Research Council (NSERC), the Canada Research Chairs program, Bombardier Aerospace, Mathematics of Information Technology and Complex Systems (MITACS), and the University of Toronto

\*corresponding author

*Email addresses:* hickej2@rpi.edu (J.E. Hicken), dwz@oddjob.utias.utoronto.ca (D.W. Zingg)

<sup>1</sup>Assistant Professor

<sup>2</sup>Professor and Director, Tier 1 Canada Research Chair in Computational Aerodynamics, J. Armand Bombardier Foundation Chair in Aerospace Flight

solution of the Navier-Stokes equations. In this context, where the functional is the quantity of primary interest, we consider an algorithm efficient if it computes the functional more accurately than a baseline algorithm using the same computational resources and time.

One way to increase the accuracy and, potentially, the efficiency of a functional estimate is to increase the accuracy of the solution itself. This is often a motivation for high-order discretization schemes. However, experience suggests that a discrete solution with a verified order of accuracy — including second-order — can produce functional estimates with erratic convergence behaviour and significantly delayed asymptotic convergence. For example, Vassberg and Jameson [1] considered subsonic and transonic flow around a modified NACA0012 airfoil. They demonstrated that several well-established solvers require a significant number of nodes ( $> 2.5 \times 10^5$  in 2D) to reach the asymptotic-convergence regime and often exhibit sub-optimal convergence rates, even for subsonic flows.

The convergence behaviour of functional estimates was also studied recently by Salas and Atkins [2]. They showed that the error due to the quadrature rule can interact with the solution error, producing singularities in the functional convergence. A similar issue arises from higher-order contributions to the solution error, i.e. when the solution itself is not yet in the region of asymptotic convergence. Indeed, their theoretical model of convergence bears a striking resemblance to several convergence results presented in [1].

In this paper we show that functional accuracy and convergence behaviour is strongly influenced by the (possibly tacit) discretization of the adjoint equations. Adjoint, or dual, problems play an increasingly important role in scientific computing. Applications that exploit the adjoint variables include optimization [3–9], error control and estimation [10–12], and mesh adaptation [13–15], to name a few.

The equations that govern the adjoint variables are typically derived in one of two ways. In the continuous-adjoint approach, associated with Jameson’s pioneering work in aerodynamic shape optimization [4], the adjoint PDE is first derived from the primal<sup>3</sup> PDE and functional; subsequently, the primal and adjoint PDEs are discretized independently. In the discrete-adjoint approach, see e.g. [16, 17], the *discretized* primal PDE and functional are used to derive the linear system for the adjoint variables. In the PDE-constrained optimization literature, the discrete-adjoint approach (resp. continuous-adjoint approach) is sometimes called the discretize-then-differentiate approach (resp. differentiate-then-discretize approach) [18].

The two approaches have distinct advantages and disadvantages. The discrete-adjoint equations are typically easier to derive, since they amount to an exercise in differentiation; thus, developers can use automatic differentiation [19, 20] or the complex-step method [21–24] to populate the (transposed) Jacobian. In addition, gradients of the discrete functional computed using the discrete-adjoint variables are accurate to machine precision, so existing gradient-based optimization algorithms can be used for design problems.

The continuous-adjoint approach leads to a discretized system that is a consistent approximation of the adjoint PDE, which has subtle but important consequences. For example, Collis and Heinkenschloss [25] found that the continuous-adjoint approach leads to control variables with better asymptotic convergence properties when high-order elements are used in a SUPG finite-element discretization. Moreover, in functional error estimation, the adjoint solution is often interpolated onto a finer grid or solution space. Adjoint solutions based on the discrete-adjoint approach can

---

<sup>3</sup>The term primal refers to the original problem of interest.

be non-smooth and oscillatory, making this interpolation process difficult, if not invalid. Solutions based on the continuous-adjoint approach are typically well behaved and straightforward to interpolate, because they converge to the infinite-dimensional solution.

The discrete- and continuous-adjoint approaches are not mutually exclusive: it is possible to discretize the primal equations such that the discrete-adjoint equations are consistent with the adjoint PDE. In this case there is no distinction between the discrete- and continuous-adjoint approaches, and one can simultaneously obtain the advantages of each. Moreover, a primal discretization that is consistent with the infinite-dimensional dual problem has some beneficial properties with respect to the behavior of the functional, such as smooth convergence and even superconvergence.

The discretization property discussed above is called adjoint, or dual, consistency<sup>4</sup>, and it is well known in the finite-element community; see, e.g., [11, 26–28]. In this paper, we show that dual consistency is an important and desirable property beyond finite-element discretizations. In particular, we demonstrate that dual consistency can significantly impact the accuracy of functionals and error estimates for high-order summation-by-parts (SBP) finite-difference discretizations. The results suggest that dual consistency offers a potential solution to the type of irregular functional convergence discussed by Salas and Atkins [2]

We begin with a review of duality in Section 2, where we also introduce the concept of dual consistency. In Section 3 we summarize the consequences of dual consistency for SBP finite-difference discretizations, namely superconvergent discrete functionals and improved functional error estimates. In Section 4, we prove that a particular SBP discretization of the Euler equations is dual consistent. We demonstrate the impact of dual consistency using several numerical experiments; these results are presented in Section 5. Finally, we conclude in Section 6 with a summary of the results and some conjectures regarding the importance of dual consistency for general discretizations.

### 1.1. Notation

Capital letters with a script type denote functions on a specified domain  $\Omega$ . For example,  $\mathcal{U}(x) \in \mathbf{H}^s(\Omega)$  is a function in the Sobolev space  $\mathbf{H}^s(\Omega)$  whose derivatives up to order  $s$  are square integrable on  $\Omega$ . Small roman letters in a serif type are used to indicate a function restricted to the grid, for example

$$u = [\mathcal{U}(x_0) \quad \mathcal{U}(x_1) \quad \cdots \quad \mathcal{U}(x_n)]^T.$$

If a subscript  $h$  appears on a vector, for example,  $u_h \in \mathbb{R}^{n+1}$ , this indicates that the vector is the solution of a difference equation on a mesh with a nominal spacing  $h$ .

We will use  $I_k$  to denote the  $k \times k$  identity matrix and  $e_i$  to denote its  $i^{\text{th}}$  column. The Kronecker product of two matrices  $A \in \mathbb{R}^{n \times m}$  and  $B \in \mathbb{R}^{q \times r}$  is denoted  $A \otimes B \in \mathbb{R}^{nq \times mr}$  and defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix}$$

The “big-O” order notation is used to indicate various bounds. We write  $F(h) = O(h^p)$  if and only if  $\exists M > 0$  and  $h_\star > 0$  such that

$$|F(h)| \leq Mh^p, \quad \forall h < h_\star.$$

---

<sup>4</sup>dual and adjoint consistency are used interchangeably throughout the paper

## 2. Duality and dual consistency

Dual consistency is a property that is rarely discussed outside the Galerkin finite-element community. Therefore, this section is devoted to a review of duality and the definition of dual, or adjoint, consistency.

### 2.1. The continuous dual problem: a simple example

We introduce the concept of duality using a linear problem with homogeneous boundary conditions. Nonlinearity is easily accommodated using Fréchet linearization; see [11] and Section 4. For a more general discussion of duality, including nonlinearity and inhomogeneous boundary conditions, see [28].

Let  $L$  be a linear differential operator on a domain  $\Omega$ , and let  $B$  be a linear-differential operator on  $\Gamma$ , the boundary of  $\Omega$ . We assume that  $L$  and  $B$  are such that the following scalar boundary-value problem (BVP) is well posed.

$$\begin{aligned} L\mathcal{U} - \mathcal{F} &= 0, & \forall x \in \Omega, \\ B\mathcal{U} &= 0, & \forall x \in \Gamma, \end{aligned} \quad (1)$$

where  $\mathcal{F} \in L^2(\Omega)$  is a spatially varying function independent of the solution  $\mathcal{U} \in W$ , with  $W$  an appropriate Sobolev space. Since (1) is well posed, we can consider a linear functional of the solution  $\mathcal{U}$ :

$$\mathcal{J}(\mathcal{U}) = \int_{\Omega} \mathcal{G}\mathcal{U} \, d\Omega, \quad (2)$$

where  $\mathcal{G} \in L^2(\Omega)$ .

Let  $L^*$  and  $B^*$  denote the adjoint operators of  $L$  and  $B$ , respectively. The adjoint operators are defined using the extended Green's identity [29]: assuming  $\mathcal{U}$  and  $\mathcal{V}$  are sufficiently differentiable, we have

$$\int_{\Omega} \mathcal{V}(L\mathcal{U}) \, d\Omega + \int_{\Gamma} (C^* \mathcal{V})(B\mathcal{U}) \, d\Gamma = \int_{\Omega} \mathcal{U}(L^* \mathcal{V}) \, d\Omega + \int_{\Gamma} (C\mathcal{U})(B^* \mathcal{V}) \, d\Gamma. \quad (3)$$

The operator  $C$  is a linear differential operator on  $\Gamma$ , and  $C^*$  is its adjoint. The identity (3) can be derived through (perhaps repeated) application of the Green-Gauss Divergence Theorem. In the present example, the operators  $C$  and  $B^*$  are not unique.

Introducing the term  $-\int_{\Omega} \mathcal{V}\mathcal{F} \, d\Omega$  into (3), we arrive at the following identity.

$$\begin{aligned} \int_{\Omega} \mathcal{V}(L\mathcal{U} - \mathcal{F}) \, d\Omega + \int_{\Gamma} (C^* \mathcal{V})(B\mathcal{U}) \, d\Gamma &= \int_{\Omega} \mathcal{U}(L^* \mathcal{V}) \, d\Omega - \int_{\Omega} \mathcal{V}\mathcal{F} \, d\Omega + \int_{\Gamma} (C\mathcal{U})(B^* \mathcal{V}) \, d\Gamma \\ &= 0 \end{aligned}$$

We can add the above (null) right-hand side to the functional  $\mathcal{J}$  without changing its value:

$$\begin{aligned} \mathcal{J}(\mathcal{U}) &= \int_{\Omega} \mathcal{G}\mathcal{U} \, d\Omega + \int_{\Omega} \mathcal{U}(L^* \mathcal{V}) \, d\Omega - \int_{\Omega} \mathcal{V}\mathcal{F} \, d\Omega + \int_{\Gamma} (C\mathcal{U})(B^* \mathcal{V}) \, d\Gamma \\ &= - \int_{\Omega} \mathcal{V}\mathcal{F} \, d\Omega + \int_{\Omega} \mathcal{U}(L^* \mathcal{V} + \mathcal{G}) \, d\Omega + \int_{\Gamma} (C\mathcal{U})(B^* \mathcal{V}) \, d\Gamma. \end{aligned} \quad (4)$$

Until now we have placed no conditions on  $\mathcal{V}$  — other than it is sufficiently differentiable — so (4) is a nonunique expression for the functional  $\mathcal{J}$ . The dual problem arises when we seek a  $\mathcal{V}$

that makes (4) independent of variations in  $\mathcal{U}$ , or, in other words, makes  $\mathcal{J}$  a functional of  $\mathcal{V}$  only. Such a  $\mathcal{V}$  is given by the solution to the adjoint BVP

$$\begin{aligned} L^* \mathcal{V} + \mathcal{G} &= 0, & \forall x \in \Omega, \\ B^* \mathcal{V} &= 0, & \forall x \in \Gamma. \end{aligned} \tag{5}$$

The solution  $\mathcal{V}$  to (5) is called the dual, costate, or adjoint variable. It follows from (4) and (5) that the adjoint variable satisfies

$$\mathcal{J}(\mathcal{U}) = \mathcal{J}(\mathcal{V}) = - \int_{\Omega} \mathcal{V} \mathcal{F} d\Omega.$$

Note that the roles of  $\mathcal{F}$  and  $\mathcal{G}$  in the primal formulation are reversed in the adjoint formulation.

## 2.2. The discrete dual problem and dual consistency

As discussed in the introduction, the adjoint PDE has a discrete analog. Let

$$L_h u_h - f = 0 \tag{6}$$

be a discretization of (1), where  $u_h \in \mathbb{R}^n$  is an element of a discrete solution space;  $u_h$  may hold basis function coefficients and/or collocation values. The matrix linear operator  $L_h : \mathbb{R}^n \rightarrow \mathbb{R}^n$  incorporates both the discretization of the PDE and the boundary conditions. Let

$$J_h(u_h) = (g, u_h)_h$$

be a discrete approximation of the functional  $\mathcal{J}$ , where  $(\cdot)_h$  denotes a discrete inner product on  $\mathbb{R}^n \times \mathbb{R}^n$ .

Using the discrete inner product, we contract the discretization (6) with an arbitrary  $v_h \in \mathbb{R}^n$  and add the result to  $J_h(u_h)$ :

$$\begin{aligned} J_h(u_h) &= (g, u_h)_h + (v_h, L_h u_h - f)_h \\ &= -(v_h, f)_h + (L_h^* v_h + g, u_h)_h, \end{aligned}$$

where  $(v_h, L_h u_h)_h = (L_h^* v_h, u_h)_h$  defines the discrete adjoint operator  $L_h^*$ . In analogy with (5), we define the discrete-adjoint equation

$$L_h^* v_h + g = 0. \tag{7}$$

What is the relationship between the above discrete system and the continuous adjoint PDE (5)?

**Definition 1 (Dual/Adjoint Consistency).** A discrete operator  $L_h$  and functional  $J_h$  are dual, or adjoint, consistent of order  $q \geq 1$  with respect to a corresponding continuous PDE and functional if

$$L_h^* v + g = O(h^q),$$

where  $v$  is the solution to the continuous dual problem projected onto the discrete solution space.

In other words, a discretization is dual consistent if it leads to a discrete dual problem that is a consistent discretization of the continuous dual PDE. This concept of dual consistency was formally introduced by Lu [27] in the context of discontinuous Galerkin discretizations; see also [28]. It is

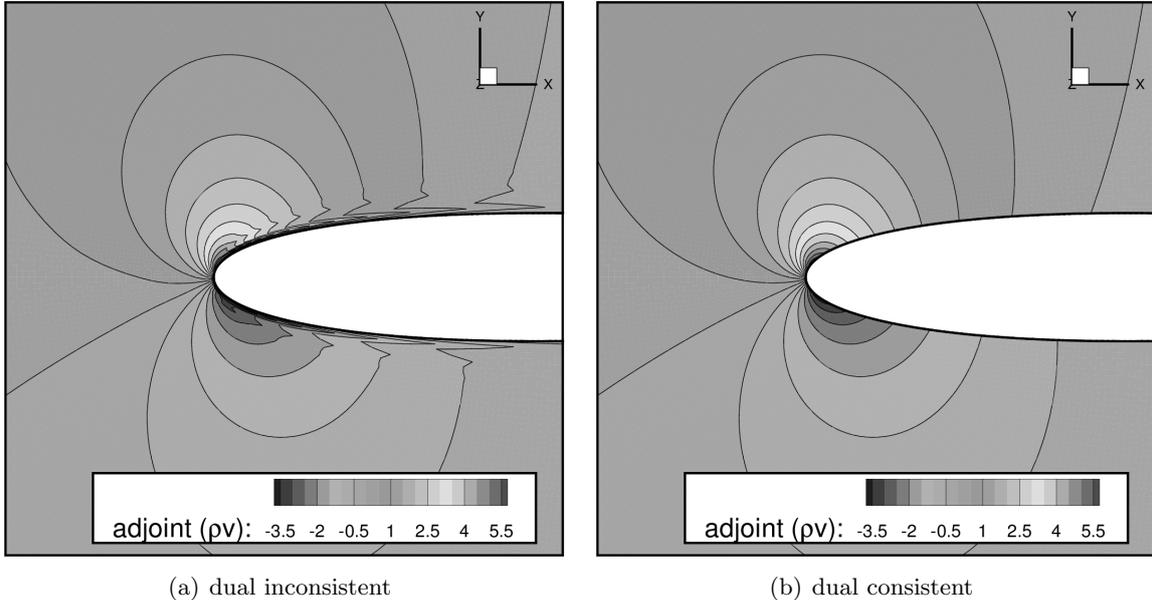


Figure 1: Contours of the third component of the drag adjoint near the leading edge of the NACA0012 airfoil in a subsonic flow.

important to recognize that dual consistency does not follow from consistency of the primal PDE discretization, in general.

What are the implications if a scheme is dual consistent? Why should we favour such schemes? Figure 1 illustrates the difference between a discrete adjoint field that is dual consistent and one that is dual inconsistent. The primal solution corresponds to an inviscid compressible flow around a NACA0012 airfoil at Mach 0.5 and an angle of attack of 1.25 degrees, and the functional is the drag-force coefficient; see Section 5.2.2 for additional details. The contours plotted are for the third component of the adjoint vector. Significant oscillations are visible near the surface of the airfoil in the dual-inconsistent results. In contrast, the dual-consistent solution is smooth. We emphasize that the primal solutions of both schemes do not exhibit any oscillations. As we shall see, the consequences of dual consistency are more than aesthetic.

### 3. Dual consistency and summation-by-parts finite-difference discretizations

This section reviews theoretical results regarding the impact of dual consistency on diagonal-norm summation-by-parts (SBP) finite-difference discretizations. SBP finite-difference operators were developed by Kreiss and Scherer [30] to mimic the stability properties of Galerkin finite-element methods. Subsequent work on weak boundary conditions [31–33], multi-block interfaces [34–36], numerical dissipation [37], and second-derivative operators [38, 39] have contributed to the theory and practice of SBP-based discretizations.

The present work focuses on the class of SBP operators with so-called “diagonal-norms.” These finite-difference operators satisfy the following definition.

**Definition 2 (Diagonal-Norm Summation-By-Parts Operator).** For discrete functions on

a uniform mesh with  $(n + 1)$  nodes, the matrix  $D \in \mathbb{R}^{(n+1) \times (n+1)}$  is a diagonal-norm summation-by-parts operator approximating the first derivative if it has the form

$$D = H^{-1}Q,$$

where  $H \in \mathbb{R}^{(n+1) \times (n+1)}$  is a diagonal positive-definite weight matrix with entries  $H_{ii} = O(h)$ , and  $Q \in \mathbb{R}^{(n+1) \times (n+1)}$  satisfies

$$Q + Q^T = \text{diag}(-1, 0, 0, \dots, 0, 1).$$

Furthermore,  $D$  is a  $2p$ -order-accurate approximation to  $d/dx$  at the interior nodes,  $\{x_k\}_{k=2p}^{n-2p}$ , and a  $p$ -order-accurate approximation at the boundary nodes,  $\{x_k\}_{k=0}^{2p-1}$  and  $\{x_k\}_{k=n-2p+1}^n$ .

Thus, an SBP operator is a first-derivative approximation that is “nearly” skew-symmetric with respect to the weight matrix  $H$ . A diagonal-norm SBP operator has  $2p$  boundary points at each end of the domain whose truncation error is  $O(h^p)$ , or better. The interior points use a centered stencil and have  $O(h^{2p})$  truncation errors.

We are primarily interested in SBP operators that have diagonal  $H$ , because Svärd [40] has shown that  $H$  must be diagonal to guarantee time stability when coordinate transformations are used. Unfortunately, to achieve both stability and high-order accuracy, diagonal-norm SBP operators must use interior stencils with twice the accuracy of their boundary stencils. For example, a sixth-order interior scheme must be paired with a third-order boundary scheme; for hyperbolic problems, this limits the global accuracy of the solution to fourth order, at best.

The price that diagonal-norm SBP operators pay for time stability is a decrease in solution accuracy for a given interior stencil size. Fortunately, in addition to enabling the derivation of time-stable schemes, the properties of SBP operators facilitate the construction of dual-consistent discretizations.

### 3.1. Dual consistency analysis of the one-dimensional advection BVP

We begin with a simple one-dimensional advection BVP, which provides an illustrative example of dual consistency and its implications for functional accuracy. Consider the first-order BVP

$$\begin{aligned} \frac{d}{dx}(\lambda \mathcal{U}) - \mathcal{F} &= 0, & \forall x \in \Omega = [0, 1] \\ \mathcal{U}(0) &= \mathcal{U}_L, \end{aligned} \tag{8}$$

where  $\lambda \in \mathbb{H}^1[0, 1]$  is the positive advection velocity. In addition, consider the linear functional

$$\mathcal{J} = \int_0^1 \mathcal{G} \mathcal{U} \, dx + \alpha (\lambda \mathcal{U})|_{x=1}. \tag{9}$$

The integral inner product of the PDE in (8) with an arbitrary  $\mathcal{V} \in \mathbb{H}^1[0, 1]$  yields zero. Similarly, the product of  $\mathcal{V}(0)$  with the boundary expression,  $\mathcal{U}(0) - \mathcal{U}_L$ , also vanishes. The sum of these two products is the Galerkin-weighted residual, which is denoted by  $\mathcal{R}(\mathcal{U}, \mathcal{V})$ :

$$\mathcal{R}(\mathcal{U}, \mathcal{V}) = \int_0^1 \mathcal{V} \left[ \frac{d}{dx}(\lambda \mathcal{U}) - \mathcal{F} \right] dx + [\lambda \mathcal{V} (\mathcal{U} - \mathcal{U}_L)]|_{x=0} = 0.$$

Note that we have inserted  $\lambda$  into the boundary-condition part of the residual to simplify the subsequent analysis; omitting this factor does not change the fundamental outcome.

Adding  $\mathcal{R}$  to  $\mathcal{J}$  and integrating by parts, we find

$$\begin{aligned}\mathcal{J} &= \int_0^1 \mathcal{G}\mathcal{U} \, dx + \alpha (\lambda\mathcal{U})|_{x=1} + \int_0^1 \mathcal{V} \left[ \frac{d}{dx} (\lambda\mathcal{U}) - \mathcal{F} \right] dx + [\lambda\mathcal{V} (\mathcal{U} - \mathcal{U}_L)]|_{x=0}, \\ &= - \int_0^1 \mathcal{V}\mathcal{F} \, dx - \mathcal{U}_L (\lambda\mathcal{V})|_{x=0} + \int_0^1 \mathcal{U} \left[ -\lambda \frac{d}{dx} (\mathcal{V}) + \mathcal{G} \right] dx + [\lambda\mathcal{U} [\mathcal{V} + \alpha]]|_{x=1}.\end{aligned}\quad (10)$$

Relating the present example to the general case in Section 2.1, we see that the primal and adjoint differential operators are given by  $L = \frac{d}{dx}(\lambda)$  and  $L^* = -\lambda \frac{d}{dx}$ , respectively. The boundary operators are given by  $B = B^* = I$  and  $C = C^* = \lambda I$ , where  $I$  is the identity operator.

The dual of the primal PDE and functional  $\mathcal{J}$  can be inferred from line (10), which can be made independent of  $\mathcal{U}$  if  $\mathcal{V}$  satisfies the adjoint PDE

$$\begin{aligned}-\lambda \frac{d}{dx} (\mathcal{V}) - \mathcal{G} &= 0, \quad \forall x \in \Omega = [0, 1], \\ \mathcal{V}(1) &= -\alpha.\end{aligned}\quad (11)$$

Consequently, the dual form of the functional is

$$\mathcal{J} = - \int_0^1 \mathcal{V}\mathcal{F} \, dx - \mathcal{U}_L (\lambda\mathcal{V})|_{x=0}.\quad (12)$$

Next, we turn to the discrete dual problem. An SBP discretization of (8) is given by

$$D\Lambda u_h - f = -H^{-1}e_0 e_0^T \Lambda(u_h - \mathcal{U}_L e_0),\quad (13)$$

where

$$\begin{aligned}\Lambda &= \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_n), \quad \lambda_i = \lambda(x_i), \quad i = 0, 1, \dots, n, \\ f &\equiv [\mathcal{F}(x_0) \quad \mathcal{F}(x_1) \quad \dots \quad \mathcal{F}(x_n)]^T.\end{aligned}\quad (14)$$

The term on the right-hand side of (13) is a penalty that imposes the boundary condition in a weak sense [31, 41]. Recall  $e_0$  is the first column of  $I_{n+1}$ , so the penalty is applied only to the first node. A discretization of the functional (9) using SBP quadrature [42] is given by

$$J_h(u) \equiv g^T H u + \alpha e_n^T \Lambda u,\quad (15)$$

where

$$g \equiv [\mathcal{G}(x_0) \quad \mathcal{G}(x_1) \quad \dots \quad \mathcal{G}(x_n)]^T.\quad (16)$$

If  $u$  is the restriction of  $\mathcal{U} \in H^{2p}(\Omega)$ , then  $J_h(u)$  is a  $2p$ -order accurate discretization of  $\mathcal{J}(\mathcal{U})$ , since the quadrature induced by  $H$  is  $2p$ -order accurate [42]. If  $u_h$  is the solution of the discrete problem (13), then the accuracy of  $j_h(u_h)$  is not as obvious. This is the subject of Theorem 1 below.

Like the functional  $\mathcal{J}$ , the continuous residual can also be discretized. We define

$$\begin{aligned}R_h(u_h, v_h) &\equiv v_h^T H [D\Lambda u_h - f + H^{-1}e_0 e_0^T \Lambda(u_h - \mathcal{U}_L e_0)] \\ &= -[\Lambda D v_h]^T H u_h - v_h^T H f + v_h^T e_n e_n^T \Lambda u_h - v_h^T \Lambda e_0 \mathcal{U}_L.\end{aligned}\quad (17)$$

The second line was obtained using the summation-by-parts property of  $D$ . The discrete residual can be shown to be a  $2p$ -order accurate discretization of its continuous counterpart [43]. Adding  $R_h(u_h, v_h) = 0$  to the discrete functional, we find

$$\begin{aligned} J_h(u_h) &= g^T H u_h + \alpha e_n^T \Lambda u_h - [\Lambda D v_h]^T H u_h - v_h^T H f + v_h^T e_n e_n^T \Lambda u_h - v_h^T \Lambda e_0 \mathcal{U}_L \\ &= -f^T H v_h - \mathcal{U}_L e_0^T \Lambda v_h + [g - \Lambda D v_h + H^{-1} e_n e_n^T \Lambda (v_h + \alpha e_n)]^T H u_h. \end{aligned}$$

To find the discrete adjoint equation, we seek  $v_h$  such that the last line above is independent of  $u_h$ , i.e. the term in brackets vanishes. This yields the linear system

$$-\Lambda D v_h + g = -H^{-1} e_n e_n^T \Lambda (v_h - \alpha e_n). \quad (18)$$

Compare the discrete adjoint equation (18) to its continuous counterpart (11). The left-hand side is obviously a consistent discretization of the adjoint PDE; indeed, it uses the same SBP operator as the primal problem, which is  $p$ -order accurate at the boundary and  $2p$ -order accurate in the interior. The term on the right-hand side of the discrete system is a (consistent) penalty for the adjoint boundary condition. Thus, we have shown that the discretization (13) and discrete functional (15) are dual consistent.

Below we describe the implications of dual consistency for this model advection problem. The theory summarized in the next two sections can be found in references [44] and [43].

### 3.2. Superconvergent discrete functionals

**Theorem 1.** *Assume the primal PDE (8) and adjoint PDE (11) are well posed with solutions  $\mathcal{U} \in \mathbf{H}^1[0, 1]$  and  $\mathcal{V} \in \mathbf{H}^1[0, 1]$ , respectively. If the discretization (13) has a solution  $u_h \in \mathbb{R}^{n+1}$  with an error that is  $O(h^{p+1})$ , then the discrete functional  $J_h(u_h)$  is a  $2p$ -order accurate approximation to  $\mathcal{J}(\mathcal{U})$ .*

See [44] for the proof. Although not obvious from the statement of the theorem, the proof relies on the dual consistency of the SBP discretization. If the difference operator did not satisfy the SBP property, or if there was a mismatch between the boundary penalty and the discrete functional definition, the functional estimate would not be superconvergent.

Thus, the dual-consistent SBP discretization leads to a superconvergent functional estimate, a result usually associated only with finite-element methods. However, we note that there are close connections between the Galerkin FEM and SBP discretizations. For example, the discrete residual (17) can be interpreted as a discrete version of a Galerkin variational statement.

### 3.3. Adjoint-based functional error estimation

The adjoint-weighted residual method has proven to be highly effective for both functional error estimation and output-based mesh adaptation; see [45] for a review. Here we describe the extension of this method to SBP finite-difference discretizations and highlight the role of dual consistency.

**Theorem 2.** *For a given uniform mesh with  $(n + 1)$  nodes, let  $D_p = H_p^{-1} Q_p$  and  $D_q = H_q^{-1} Q_q$  denote diagonal-norm SBP operators with boundary stencils of  $p$ - and  $q$ -order accuracy, respectively, where  $p < q$ . As before, assume the primal PDE (8) and adjoint PDE (11) are well posed with solutions  $\mathcal{U} \in \mathbf{H}^1[0, 1]$  and  $\mathcal{V} \in \mathbf{H}^1[0, 1]$ , respectively. Let  $R_{h,q}(u, v)$  denote the discrete residual (17) evaluated using the SBP operator  $D_q$ . Let  $u_h \in \mathbb{R}^{n+1}$  be the solution to the primal discretization*

(13) (based on  $D_p$ ) and  $v_h \in \mathbb{R}^{n+1}$  the solution to the adjoint discretization (18) (again, using  $D_p$ ). Assume that both  $u_h$  and  $v_h$  are  $O(h^{p+1})$  accurate. Then the functional error estimate

$$\delta J_{h,p} \equiv g^T (H_p - H_q) u_h - R_{h,q}(u_h, v_h) \quad (19)$$

is a  $(2p + 2)$ -order accurate approximation to the true functional error,  $\delta \mathcal{J}_{h,p} \equiv J_h(u_h) - \mathcal{J}(U)$ .

The proof can be found in [43]. This result indicates that we can recover two additional orders of accuracy in the functional for this particular problem; recall, the superconvergent functional  $J_h(u_h)$  is  $O(h^{2p})$  accurate.

While the above theorem has only been proven for a restricted class of problems, the numerical experiments in Section (5) demonstrate that it continues to hold for more complex PDEs, such as the Euler equations, *provided the discretized problem is dual consistent*.

#### 4. Dual consistency analysis of the Euler equations

If we wish to extend Theorems 1 and 2 to more complex problems, we need to prove that the corresponding discretization is dual consistent. Note that we cannot analyze a particular class of difference operator, e.g. diagonal-norm SBP, and prove it will be dual consistent for all PDEs. The operator cannot be treated in isolation, because the details of dual consistency depend on the discretization of the boundary conditions and functional, not merely that of the differential operator.

Therefore, in the present work, we have chosen to analyze the dual consistency of an SBP discretization of the Euler equations, with typical boundary conditions and the most common type of functional (a pressure-based force). It is hoped that this analysis will illustrate the procedure that needs to be followed to establish the dual consistency of other SBP discretizations.

##### 4.1. The dual problem for the steady Euler equations

For simplicity, we consider two-dimensional domains  $\Omega_x \in \mathbb{R}^2$  that are diffeomorphic to the square  $\Omega = [0, 1]^2$ . Domains that are not diffeomorphic to  $\Omega$  can often be broken into non-overlapping subdomains, each of which is individually diffeomorphic to  $\Omega$ ; in the Section 4.4, we will address the interfaces that arise in this “multi-block” setting. We have restricted the analysis to two-dimensions to keep the presentation concise. The extension to three-dimensions is straightforward.

With a suitable mapping  $\mathcal{T} : (x, y) \rightarrow (\xi, \eta)$  between  $\Omega_x$  and  $\Omega$ , the steady Euler equations take the form

$$\frac{\partial}{\partial \xi} \mathcal{F}_\xi + \frac{\partial}{\partial \eta} \mathcal{F}_\eta = 0, \quad (20)$$

where the fluxes are given by

$$\mathcal{F}_\xi = \begin{pmatrix} \rho U_\xi \\ \rho u U_\xi + p \partial_x \xi \\ \rho v U_\xi + p \partial_y \xi \\ (e + p) U_\xi \end{pmatrix}, \quad \text{and} \quad \mathcal{F}_\eta = \begin{pmatrix} \rho U_\eta \\ \rho u U_\eta + p \partial_x \eta \\ \rho v U_\eta + p \partial_y \eta \\ (e + p) U_\eta \end{pmatrix}.$$

The unknowns in (20) are the conservative variables: density,  $\rho$ ; momentum per unit volume,  $(\rho u, \rho v)^T$ , and; energy per unit volume,  $e$ . We denote the complete set of unknowns by  $\mathbf{U} = (\rho, \rho u, \rho v, e)^T$ . The ideal gas law is used to define the pressure,  $p$ , which closes the system. Finally,  $U_\xi$  and  $U_\eta$  denote the contravariant velocities:

$$U_\xi = u\partial_x\xi + v\partial_y\xi, \quad \text{and} \quad U_\eta = u\partial_x\eta + v\partial_y\eta.$$

We specify a vanishing normal momentum along  $\eta = 0$ . This inviscid-wall boundary condition is enforced using the flux in the  $\eta$  direction:

$$\mathcal{F}_\eta(\mathbf{U}) = \mathcal{F}_\eta(P\mathbf{U}) = \begin{pmatrix} 0 & p\partial_x\eta & p\partial_y\eta & 0 \end{pmatrix}^T, \quad \forall \xi \in [0, 1], \eta = 0. \quad (21)$$

The symmetric matrix  $P \in \mathbb{R}^{4 \times 4}$  projects out the normal component — in  $(x, y)$  space — of momentum from  $\mathbf{U}$ . For the subsequent analysis, note that the inviscid-wall boundary flux is in the null space of  $P$ ; consequently,

$$(I_4 - P)\mathcal{F}_\eta(P\mathbf{U}) = \mathcal{F}_\eta(P\mathbf{U}). \quad (22)$$

For the remaining three edges, we impose far-field characteristic boundary conditions. If  $\mathbf{U}_\infty$  denotes the free-stream value of the conservative variables, then the characteristic far-field boundary conditions are given by

$$\begin{aligned} \mathcal{F}_\xi(\mathbf{U}) &= A_{\xi, \infty}^- \mathbf{U} + A_{\xi, \infty}^+ \mathbf{U}_\infty, & \forall \eta \in [0, 1], \xi = 0, \\ \mathcal{F}_\xi(\mathbf{U}) &= A_{\xi, \infty}^+ \mathbf{U} + A_{\xi, \infty}^- \mathbf{U}_\infty, & \forall \eta \in [0, 1], \xi = 1, \\ \mathcal{F}_\eta(\mathbf{U}) &= A_{\eta, \infty}^+ \mathbf{U} + A_{\eta, \infty}^- \mathbf{U}_\infty, & \forall \xi \in [0, 1], \eta = 1. \end{aligned} \quad (23)$$

These boundary conditions make use of the diagonalizable flux Jacobians

$$A_\xi \equiv \frac{\partial \mathcal{F}_\xi}{\partial \mathbf{U}} = X_\xi \Lambda_\xi X_\xi^{-1} \quad \text{and} \quad A_\eta \equiv \frac{\partial \mathcal{F}_\eta}{\partial \mathbf{U}} = X_\eta \Lambda_\eta X_\eta^{-1}.$$

For the definitions of  $X_\xi, \Lambda_\xi, X_\eta$ , and  $\Lambda_\eta$  see, for example, [46]. As is common in the gas dynamics literature, we use the notation  $|A_\xi|$ , resp.  $|A_\eta|$ , to denote the matrices obtained by replacing the eigenvalues in  $A_\xi$ , resp.  $A_\eta$ , with their absolute values. In addition, we define

$$A_\xi^+ \equiv \frac{1}{2}(A_\xi + |A_\xi|) \quad \text{and} \quad A_\xi^- \equiv \frac{1}{2}(A_\xi - |A_\xi|).$$

Analogous definitions hold for  $A_\eta^+$  and  $A_\eta^-$ . In the boundary conditions (23), the subscript  $\infty$  indicates that the matrix is evaluated at the *constant* free-stream value  $\mathbf{U}_\infty$ .

For the derivation of the adjoint equations, we write the Euler equations and the boundary conditions, (21) and (23), in the Galerkin weighted-residual form:  $\forall \mathbf{V} \in [H^1(\Omega)]^4$  find  $\mathbf{U}$  such that

$\mathcal{R}(\mathbf{u}, \mathbf{v}) = 0$ , where the residual is defined by

$$\begin{aligned} \mathcal{R}(\mathbf{u}, \mathbf{v}) \equiv & \int_{\Omega} \mathbf{v}^T \left( \frac{\partial}{\partial \xi} \mathcal{F}_{\xi} + \frac{\partial}{\partial \eta} \mathcal{F}_{\eta} \right) d\eta d\xi \\ & + \int_{\Gamma_{\xi=0}} \mathbf{v}^T \left[ \mathcal{F}_{\xi}(\mathbf{u}) - A_{\xi,\infty}^- \mathbf{u} - A_{\xi,\infty}^+ \mathbf{u}_{\infty} \right] d\eta \\ & - \int_{\Gamma_{\xi=1}} \mathbf{v}^T \left[ \mathcal{F}_{\xi}(\mathbf{u}) - A_{\xi,\infty}^+ \mathbf{u} - A_{\xi,\infty}^- \mathbf{u}_{\infty} \right] d\eta. \\ & + \int_{\Gamma_{\eta=0}} \mathbf{v}^T \left[ \mathcal{F}_{\eta}(\mathbf{u}) - \mathcal{F}_{\eta}(P\mathbf{u}) \right] d\xi \\ & - \int_{\Gamma_{\eta=1}} \mathbf{v}^T \left[ \mathcal{F}_{\eta}(\mathbf{u}) - A_{\eta,\infty}^+ \mathbf{u} - A_{\eta,\infty}^- \mathbf{u}_{\infty} \right] d\xi \end{aligned}$$

In weighted-residual statements, boundary conditions are often enforced in a strong sense by restricting  $\mathbf{u}$  to lie in an admissible space where all elements of the space satisfy the boundary conditions. Here we adopt a weak imposition of the boundary conditions, due to Nitsche [47], in which the boundary conditions are incorporated directly into the residual. Using weakly imposed boundary conditions clarifies the connection between the continuous and discrete adjoint derivation for the SBP discretization. Note that we can rewrite the residual using Green's theorem as

$$\begin{aligned} \mathcal{R}(\mathbf{u}, \mathbf{v}) = & - \int_{\Omega} \left( \frac{\partial \mathbf{v}^T}{\partial \xi} \mathcal{F}_{\xi} + \frac{\partial \mathbf{v}^T}{\partial \eta} \mathcal{F}_{\eta} \right) d\eta d\xi \\ & - \int_{\Gamma_{\xi=0}} \mathbf{v}^T \left[ A_{\xi,\infty}^- \mathbf{u} + A_{\xi,\infty}^+ \mathbf{u}_{\infty} \right] d\eta + \int_{\Gamma_{\xi=1}} \mathbf{v}^T \left[ A_{\xi,\infty}^+ \mathbf{u} + A_{\xi,\infty}^- \mathbf{u}_{\infty} \right] d\eta \\ & - \int_{\Gamma_{\eta=0}} \mathbf{v}^T \mathcal{F}_{\eta}(P\mathbf{u}) d\xi + \int_{\Gamma_{\eta=1}} \mathbf{v}^T \left[ A_{\eta,\infty}^+ \mathbf{u} + A_{\eta,\infty}^- \mathbf{u}_{\infty} \right] d\xi. \end{aligned} \quad (24)$$

For our model functional, we use the force (per unit depth) exerted on the boundary  $\Gamma_{\eta=0}$  in the direction  $(t_x, t_y)^T$ . If we define the constant vector  $\mathbf{g} \equiv -(0, t_x, t_y, 0)^T$ , then the force is given by the integral

$$\mathcal{J} = \int_{\Gamma_{\eta=0}} \mathbf{g}^T \mathcal{F}_{\eta}(P\mathbf{u}) d\xi. \quad (25)$$

Note that the boundary flux in the functional is evaluated using  $P\mathbf{u}$ . This is not necessary for the present (continuous) analysis, because the boundary condition ensures  $P\mathbf{u} = \mathbf{u}$ ; however, the inclusion of  $P$  makes the comparison with the discretized case more clear.

As with the linear advection problem, the residual can be added to the functional without changing the value of  $\mathcal{J}$ :

$$\mathcal{J} = \int_{\Gamma_{\eta=0}} \mathbf{g}^T \mathcal{F}_{\eta}(P\mathbf{u}) d\xi + \mathcal{R}(\mathbf{u}, \mathbf{v}). \quad (26)$$

Unlike the linear advection problem, the Euler equations and the functional (25) are nonlinear in the solution  $\mathbf{u}$ . Thus, it is not possible to find a value for  $\mathbf{v}$  that makes (26) independent of  $\mathbf{u}$ . For nonlinear problems, the adjoint operator is defined using the linearized form of (26). In other words, we consider variations  $\delta\mathbf{u}$  about  $\mathbf{u}$  and seek a  $\mathbf{v}$  that makes the perturbation  $\delta\mathcal{J}$  independent of these variations:  $\forall \delta\mathbf{u} \in [H(\Omega)]^4$ , find  $\mathbf{v}$  such that

$$\delta\mathcal{J} \equiv \int_{\Gamma_{\eta=0}} \mathbf{g}^T A_{\eta} P \delta\mathbf{u} d\xi + \mathcal{R}'[\mathbf{u}](\delta\mathbf{u}, \mathbf{v}) = 0, \quad (27)$$

where the prime on  $\mathcal{R}$  indicates Fréchet linearization with respect to the term in square brackets. For the residual (24) we have

$$\begin{aligned}\mathcal{R}'[\mathbf{u}](\delta\mathbf{u}, \mathbf{v}) &= - \int_{\Omega} \left( \frac{\partial \mathbf{v}^T}{\partial \xi} A_{\xi} + \frac{\partial \mathbf{v}^T}{\partial \eta} A_{\eta} \right) \delta\mathbf{u} \, d\eta d\xi \\ &\quad - \int_{\Gamma_{\xi=0}} \mathbf{v}^T A_{\xi, \infty}^- \delta\mathbf{u} \, d\eta + \int_{\Gamma_{\xi=1}} \mathbf{v}^T A_{\xi, \infty}^+ \delta\mathbf{u} \, d\eta \\ &\quad - \int_{\Gamma_{\eta=0}} \mathbf{v}^T A_{\eta} P \delta\mathbf{u} \, d\xi + \int_{\Gamma_{\eta=1}} \mathbf{v}^T A_{\eta, \infty}^+ \delta\mathbf{u} \, d\eta.\end{aligned}$$

Substituting  $\mathcal{R}'[\mathbf{u}](\delta\mathbf{u}, \mathbf{v})$  into (27) and rearranging we find

$$\begin{aligned}\delta\mathcal{J} &\equiv - \int_{\Omega} \left( \frac{\partial \mathbf{v}^T}{\partial \xi} A_{\xi} + \frac{\partial \mathbf{v}^T}{\partial \eta} A_{\eta} \right) \delta\mathbf{u} \, d\eta d\xi \\ &\quad - \int_{\Gamma_{\xi=0}} \mathbf{v}^T A_{\xi, \infty}^- \delta\mathbf{u} \, d\eta + \int_{\Gamma_{\xi=1}} \mathbf{v}^T A_{\xi, \infty}^+ \delta\mathbf{u} \, d\eta \\ &\quad - \int_{\Gamma_{\eta=0}} (\mathbf{v}^T - \mathbf{g}^T) A_{\eta} P \delta\mathbf{u} \, d\xi + \int_{\Gamma_{\eta=1}} \mathbf{v}^T A_{\eta, \infty}^+ \delta\mathbf{u} \, d\eta.\end{aligned}$$

The above defines the Galerkin weighted-residual statement for the continuous adjoint. The strong form of the adjoint PDE can be inferred from the integral over  $\Omega$ , while the adjoint boundary conditions can be inferred from the integrals over  $\Gamma$ :

$$-A_{\xi}^T \frac{\partial \mathbf{v}}{\partial \xi} - A_{\eta}^T \frac{\partial \mathbf{v}}{\partial \eta} = 0, \quad \forall (\xi, \eta) \in \Omega, \quad (28)$$

subject to

$$\left( A_{\xi, \infty}^- \right)^T \mathbf{v} = 0, \quad \forall \eta \in [0, 1], \xi = 0, \quad (29)$$

$$\left( A_{\xi, \infty}^+ \right)^T \mathbf{v} = 0, \quad \forall \eta \in [0, 1], \xi = 1, \quad (30)$$

$$P A_{\eta}^T (I_4 - P) (\mathbf{v} - \mathbf{g}) = 0, \quad \forall \xi \in [0, 1], \eta = 0, \quad (31)$$

$$\left( A_{\eta, \infty}^+ \right)^T \mathbf{v} = 0, \quad \forall \xi \in [0, 1], \eta = 1. \quad (32)$$

**Remark 1.** Along the far-field boundaries, we find that the characteristic variables ( $X_{\xi, \infty}^T \mathbf{v}$ ) and ( $X_{\eta, \infty}^T \mathbf{v}$ ), are set to zero if they correspond to out-going waves in the primal problem. Therefore, the far-field adjoint boundary conditions are well-posed in the sense that they do not overdetermine  $\mathbf{v}$ .

**Remark 2.** For the inviscid-wall boundary condition along  $\Gamma_{\eta=0}$ , we have made use of the property  $(I_4 - P)A_{\eta}(P\mathbf{u}) = A_{\eta}(P\mathbf{u})$ , which follows from (22). Moreover, along the wall we can show that  $A_{\eta}^T = (\partial p / \partial \mathbf{u}) \hat{\mathbf{n}}^T$ , where the normal vector  $\hat{\mathbf{n}} = (0, \partial_x \eta, \partial_y \eta, 0)^T / \|\nabla \eta\|$ ; thus,  $A_{\eta}^T$  is a rank-one matrix along the wall. Finally, we have that  $(I_4 - P) = \hat{\mathbf{n}} \hat{\mathbf{n}}^T$ , so the boundary condition (31) enforces only one degree of freedom:  $\hat{\mathbf{n}}^T \mathbf{v} = \hat{\mathbf{n}}^T \mathbf{g}$ . This is consistent with a single (out-going) primal characteristic at the the wall.

#### 4.2. SBP discretization of the Euler equations

In this section we describe a particular summation-by-parts discretization of the Euler equations and prove that the discretization is dual consistent.

For simplicity, we assume that the domain  $\Omega$  is uniformly discretized into  $(n + 1)$  nodes in both the  $\xi$  and  $\eta$  directions. Scalar fields on this Cartesian grid are ordered first by  $\xi$  coordinate and then by  $\eta$  coordinate. Vector fields are ordered first by their components and then by  $\xi$  and  $\eta$ . For example, if the restriction of the conservative variables  $\mathbf{U}(\xi, \eta)$  to the grid is denoted by  $u \in \mathbb{R}^{4(n+1)^2}$ , then for the vertex  $(1, 0)$  we would have

$$\mathbf{U}(1, 0) = (u_{4n+1} \quad u_{4n+2} \quad u_{4n+3} \quad u_{4n+4})^T.$$

Let  $D = H^{-1}Q$  be a diagonal-norm SBP first-derivative operator that is  $O(h^p)$  accurate at the boundary and  $O(h^{2p})$  accurate in the interior. Then, a discretization of the Euler equations on  $\Omega$  is given by

$$\begin{aligned} \bar{D}_\xi f_\xi(u_h) + \bar{D}_\eta f_\eta(u_h) &= -\bar{H}^{-1} \bar{E}_{\xi,0} \left[ f_\xi(u_h) - f_\xi^b(u_h) \right] + \bar{H}^{-1} \bar{E}_{\xi,1} \left[ f_\xi(u_h) - f_\xi^b(u_h) \right] \\ &\quad - \bar{H}^{-1} \bar{E}_{\eta,0} \left[ f_\eta(u_h) - f_\eta^b(u_h) \right] + \bar{H}^{-1} \bar{E}_{\eta,1} \left[ f_\eta(u_h) - f_\eta^b(u_h) \right], \end{aligned} \quad (33)$$

where

$$\begin{aligned} \bar{D}_\xi &\equiv I_{n+1} \otimes D \otimes I_4, & \bar{D}_\eta &\equiv D \otimes I_{n+1} \otimes I_4, \\ \bar{E}_{\xi,0} &\equiv H \otimes e_0 e_0^T \otimes I_4, & \bar{E}_{\eta,0} &\equiv e_0 e_0^T \otimes H \otimes I_4, \\ \bar{E}_{\xi,1} &\equiv H \otimes e_n e_n^T \otimes I_4, & \bar{E}_{\eta,1} &\equiv e_n e_n^T \otimes H \otimes I_4, \\ \bar{H} &\equiv H \otimes H \otimes I_4. \end{aligned}$$

The left-hand side of (33) corresponds to the PDE (20), while the right-hand side consists of penalties that impose the boundary conditions weakly. The vectors  $f_\xi(u_h)$  and  $f_\eta(u_h)$  are the restriction of the fluxes  $\mathcal{F}_\xi$  and  $\mathcal{F}_\eta$  to the nodes. The vectors  $f_\xi^b(u_h)$  and  $f_\eta^b(u_h)$  are boundary flux functions, which are defined as follows.

- For nodes along the far-field boundaries, the corresponding elements in  $f_\xi^b(u_h)$  and  $f_\eta^b(u_h)$  are set to the appropriate right-hand side appearing in (23).
- For nodes along the wall boundary, the corresponding elements in  $f_\eta^b(u_h)$  are set to the right-hand side of (21).
- For all other nodes, the values in  $f_\xi^b(u_h)$  and  $f_\eta^b(u_h)$  are inconsequential and can be set to zero.

**Remark 3.** While the discretization (33) appears unwieldy, each of the terms has a simple interpretation once we form the discrete residual; see (35) below. Indeed, each term in the discrete residual can be associated with a corresponding term in the continuous residual  $\mathcal{R}(\mathbf{U}, \mathbf{V})$ . To see this, recall that, for arbitrary vectors  $u$  and  $v$ , the product  $v^T H u$  is an accurate quadrature approximating integration over  $\Omega$ , while the product  $v^T E_{\xi,0} u$  is a quadrature approximating integration over the boundary  $\Gamma_{\xi=0}$ . Analogous interpretations hold for  $v^T E_{\xi,1} u$ ,  $v^T E_{\eta,0} u$ , and  $v^T E_{\eta,1} u$ .

**Remark 4.** In practice, numerical dissipation must be included in the discretization (33). We will address dissipation in the next section.

The functional (25) is discretized using the  $2p$ -order accurate quadrature induced by the weight matrix  $H$  appearing in  $\bar{E}_{\eta,0}$  [42]:

$$J_h(u) = g^T \bar{E}_{\eta,0} f_\eta^b(u). \quad (34)$$

The use of  $\bar{E}_{\eta,0}$  for the quadrature is one of the keys to the dual consistency of the scheme. Note the choice  $f_\eta^b(u_h)$  for the boundary flux in the discrete functional matches the boundary flux used in the wall-boundary penalty; this is also important for dual consistency, as described further below.

**Theorem 3.** *The SBP discretization (33) and the discrete functional (34) produce a dual-consistent discrete adjoint equation that is  $O(h^p)$  at the boundaries and  $O(h^{2p})$  in the interior.*

PROOF. Let  $v_h \in \mathbb{R}^{4(n+1)^2}$  be a vector field on the grid. We define the following discrete residual, which is the  $\bar{H}$  inner product of  $v_h$  with the discretization (33):

$$\begin{aligned} R_h(u_h, v_h) \equiv & v_h^T \bar{H} \bar{D}_\xi f_\xi(u_h) + v_h^T \bar{H} \bar{D}_\eta f_\eta(u_h) \\ & + v_h^T \bar{E}_{\xi,0} \left[ f_\xi(u_h) - f_\xi^b(u_h) \right] - v_h^T \bar{E}_{\xi,1} \left[ f_\xi(u_h) - f_\xi^b(u_h) \right] \\ & + v_h^T \bar{E}_{\eta,0} \left[ f_\eta(u_h) - f_\eta^b(u_h) \right] - v_h^T \bar{E}_{\eta,1} \left[ f_\eta(u_h) - f_\eta^b(u_h) \right] \end{aligned} \quad (35)$$

Using the properties of SBP operators, we can show that  $\bar{H} \bar{D}_\xi = \bar{E}_{\xi,1} - \bar{E}_{\xi,0} - \bar{D}_\xi^T \bar{H}$  and, similarly,  $\bar{H} \bar{D}_\eta = \bar{E}_{\eta,1} - \bar{E}_{\eta,0} - \bar{D}_\eta^T \bar{H}$ . Therefore,

$$\begin{aligned} R_h(u_h, v_h) = & - [\bar{D}_\xi v_h]^T \bar{H} f_\xi(u_h) - [\bar{D}_\eta v_h]^T \bar{H} f_\eta(u_h) \\ & - v_h^T \bar{E}_{\xi,0} f_\xi^b(u_h) + v_h^T \bar{E}_{\xi,1} f_\xi^b(u_h) \\ & - v_h^T \bar{E}_{\eta,0} f_\eta^b(u_h) + v_h^T \bar{E}_{\eta,1} f_\eta^b(u_h). \end{aligned}$$

Clearly, if  $u_h$  satisfies (33), then  $R_h(u_h, v_h) = 0$  for all bounded  $v_h$ . Thus, we can write

$$J_h(u_h) = g^T \bar{E}_{\eta,0} f_\eta^b(u_h) + R_h(u_h, v_h).$$

Next, we linearize  $J_h$  about  $u_h$  and consider arbitrary perturbations  $du$ . To keep the notation compact, we define the block-diagonal matrices  $B_\xi$  and  $B_\eta$  that have the flux Jacobians along their diagonal:

$$B_\xi \equiv \frac{\partial f_\xi}{\partial u_h} = \text{diag}(A_\xi), \quad \text{and} \quad B_\eta \equiv \frac{\partial f_\eta}{\partial u_h} = \text{diag}(A_\eta).$$

The matrices  $B_\xi^b$  and  $B_\eta^b$  are defined similarly, with the flux Jacobians replaced with the appropriate boundary-flux Jacobians. Using these matrices, we can write the linearized perturbation in  $J_h$  as

follows.

$$\begin{aligned}
dJ_h &\equiv g^T \bar{E}_{\eta,0} \frac{\partial f_\eta^b}{\partial u_h} du + \frac{\partial R_h}{\partial u_h} du \\
&= - [\bar{D}_\xi v_h]^T \bar{H} B_\xi du - [\bar{D}_\eta v_h]^T \bar{H} B_\eta du \\
&\quad - v_h^T \bar{E}_{\xi,0} B_\xi^b du + v_h^T \bar{E}_{\xi,1} B_\xi^b du - (v_h - g)^T \bar{E}_{\eta,0} B_\eta^b du + v_h^T \bar{E}_{\eta,1} B_\eta^b du \\
&= - [B_\xi^T \bar{D}_\xi v_h + B_\eta^T \bar{D}_\eta v_h]^T \bar{H} du \\
&\quad + \left[ \bar{H}^{-1} \bar{E}_{\xi,1} \left( B_\xi^b \right)^T v_h - \bar{H}^{-1} \bar{E}_{\xi,0} \left( B_\xi^b \right)^T v_h \right]^T \bar{H} du \\
&\quad + \left[ \bar{H}^{-1} \bar{E}_{\eta,1} \left( B_\eta^b \right)^T v_h - \bar{H}^{-1} \bar{E}_{\eta,0} \left( B_\eta^b \right)^T (v_h - g) \right]^T \bar{H} du.
\end{aligned}$$

To obtain the final equality above, we have made use of the fact that  $B_\xi$  commutes with  $E_{\xi,0}$  and  $E_{\xi,1}$ ,  $B_\eta$  commutes with  $E_{\eta,0}$  and  $E_{\eta,1}$ , etc.

If we want an expression for  $dJ_h$  that is independent of the differential  $du$ , we must have

$$\begin{aligned}
-B_\xi^T \bar{D}_\xi v_h - B_\eta^T \bar{D}_\eta v_h &= \bar{H}^{-1} \bar{E}_{\xi,0} \left( B_\xi^b \right)^T v_h - \bar{H}^{-1} \bar{E}_{\xi,1} \left( B_\xi^b \right)^T v_h \\
&\quad + \bar{H}^{-1} \bar{E}_{\eta,0} \left( B_\eta^b \right)^T (v_h - g) - \bar{H}^{-1} \bar{E}_{\eta,1} \left( B_\eta^b \right)^T v_h
\end{aligned} \tag{36}$$

This is the discrete adjoint equation for the discretization (33) and functional (34). Consider each of the terms appearing in (36) in relation to the adjoint PDE (28) and boundary conditions (29)–(32).

- The term  $B_\xi^T \bar{D}_\xi v_h$  is an SBP discretization of  $A_\xi^T \partial \mathbf{V} / \partial \xi$ . The derivative operator is  $p$ -order accurate at the boundary and  $2p$ -order accurate in the interior. Similar remarks hold for  $B_\eta^T \bar{D}_\eta v_h$ .
- The adjoint inviscid-wall boundary condition is enforced by the term

$$\bar{H}^{-1} \bar{E}_{\eta,0} \left( B_\eta^b \right)^T (v_h - g). \tag{37}$$

We need to show that this term vanishes when  $v_h$  is replaced by the (grid-restricted) continuous adjoint solution. Consider node  $i$  along the wall  $\eta = 0$ . Let  $u_{h,i} \in \mathbb{R}^4$  denote the elements of  $u_h$  corresponding to the conservative variables at node  $i$ . If  $B_{\eta,i}^b \in \mathbb{R}^{4 \times 4}$  is the  $i^{\text{th}}$  diagonal block of  $B_\eta^b$ , then

$$\begin{aligned}
B_{\eta,i}^b &= \frac{\partial}{\partial \mathbf{u}} [\mathcal{F}_\eta(P\mathbf{u})] \Big|_{u_{h,i}} \\
&= \frac{\partial}{\partial \mathbf{u}} [(I_4 - P) \mathcal{F}_\eta(P\mathbf{u})] \Big|_{u_{h,i}} \\
&= (I_4 - P) A_\eta|_{u_{h,i}} P.
\end{aligned}$$

In the second line, we have used the fact that the *chosen* inviscid-wall boundary flux is in the null space of  $P$ . From the above expression, we find that the penalty term (37) at node  $i$  is proportional to

$$P A_\eta^T|_{u_{h,i}} (I_4 - P)(v_{h,i} - g_i),$$

where  $v_{h,i}$  and  $g_i$  are the four elements of  $v_h$  and  $g$  at node  $i$ . Comparing this to the boundary condition (31), we conclude that the wall-penalty term vanishes when the continuous adjoint solution replaces  $v_h$ .

- For nodes along the boundaries  $\xi = 0$  and  $\xi = 1$ , the  $4 \times 4$  blocks in  $B_\xi^b$  and  $B_\eta^b$  are equal to  $A_{\xi,\infty}^-$  and  $A_{\xi,\infty}^+$ , respectively; recall, the right-hand sides of (23) define the corresponding boundary fluxes. Similarly, the blocks of  $B_\eta^b$  are equal to  $A_{\eta,\infty}^+$  for nodes along the boundary  $\eta = 1$ . Therefore, the terms

$$\bar{H}^{-1} \bar{E}_{\xi,0} \left( B_\xi^b \right)^T v_h, \quad \bar{H}^{-1} \bar{E}_{\xi,1} \left( B_\xi^b \right)^T v_h, \quad \text{and} \quad \bar{H}^{-1} \bar{E}_{\eta,1} \left( B_\eta^b \right)^T v_h$$

contain penalties corresponding to the adjoint boundary conditions (29), (30), and (32), respectively. Clearly these penalties vanish when the continuous adjoint is substituted for  $v_h$ .

In summary, we have shown that the truncation error in (36) is order  $p$  for boundary nodes (nodes within  $2p$  of the boundary), and order  $2p$  for interior nodes.  $\square$

**Remark 5.** The dual-consistency analysis is facilitated by the summation-by-parts property built into  $D$ . For general high-order difference operators that do not obey an SBP property, establishing dual consistency may be much more difficult, and, perhaps, impossible.

**Remark 6.** Using the SBP weight matrix  $H$  to discretize the functional is critical to the proof, because the quadrature is compatible with the wall boundary penalty: it leads to a consistent penalty for the adjoint wall boundary condition. In addition, the boundary fluxes appearing in the penalty terms and functional are deliberately chosen to ensure dual consistency. In particular, the inviscid-wall boundary flux, used in both the wall penalty and functional, is a function of  $Pu_{h,i}$ . Lu’s analysis [27] shows that other choices, including the commonly used flux splitting approach, can lead to a dual-inconsistent boundary condition.

### 4.3. Dual-consistent numerical dissipation

Due to nonlinearity, numerical dissipation must be present in the discretization of the Euler equations to damp high-frequency modes. For centered finite-difference schemes, dissipation is typically introduced using a symmetric positive-definite operator that contains undivided differences of sufficient order; see, for example, [37]. In this section, we analyze the impact of dissipation operators on dual consistency.

For each first-derivative operator in the Euler equations there will be a corresponding dissipation operator. We will analyze the dissipation operator in the direction  $\xi$ , since the same analysis can be applied to  $\eta$ , and, in three dimensions,  $\zeta$ . The dissipation is defined by

$$\bar{S}_\xi u_h \equiv \kappa \bar{H}^{-1} \bar{\Delta}_\xi^T \Sigma_\xi \bar{\Delta}_\xi u_h, \tag{38}$$

where  $\kappa$  is a scaling — we typically use  $\kappa = 0.04$  — and

$$\begin{aligned} \bar{\Delta}_\xi &\equiv I_{n+1} \otimes \Delta \otimes I_4, \\ \Sigma_\xi &\equiv \text{diag} \left( \tilde{A}_\xi \right). \end{aligned}$$

The operator  $\Delta \in \mathbb{R}^{(n+1) \times (n+1)}$  is an undivided difference operator; in particular, to maintain the accuracy of the discretization scheme,  $\Delta$  is a second-order accurate discretization of  $h^{p+1} \partial^{p+1} / \partial \xi^{p+1}$ . The  $i^{\text{th}}$  block in  $\Sigma_\xi$  is given by  $\tilde{A}_\xi \equiv X \tilde{\Lambda} X^T$ , which is a modified version of  $|A_\xi|$ . The  $4 \times 4$  matrix  $\tilde{A}_\xi$  is evaluated at the nodal state  $u_{h,i}$  and defined by

$$\begin{aligned} \tilde{\Lambda}_i &\equiv \text{diag} \left( \tilde{\lambda}_{i,1}, \tilde{\lambda}_{i,2}, \tilde{\lambda}_{i,3}, \tilde{\lambda}_{i,4} \right), \\ \tilde{\lambda}_{i,k} &\equiv \max(|\lambda_{i,k}|, \epsilon \sigma_i), \quad k = 1, 2, 3, 4 \\ \sigma_i &\equiv \max_k |\lambda_{i,k}|. \end{aligned}$$

The  $\{\lambda_{i,k}\}_{k=1}^4$  are the four eigenvalues of  $A_\xi$ , and the parameter  $\epsilon$  is adjusted to avoid difficulties introduced by zero eigenvalues; typically, we use  $\epsilon = 0.025$ . The dissipation  $\bar{S}_\xi$  is an SBP implementation of matrix dissipation [48].

Recall that the discrete residual  $R_h(u_h, v_h)$  is defined by taking the  $\bar{H}$  inner product of the discretization with  $v_h$ . Thus, in  $R_h$  the dissipation operator becomes

$$\begin{aligned} v_h^T \bar{H} \bar{S}_\xi u_h &= \kappa v_h^T \bar{\Delta}_\xi^T \Sigma_\xi \bar{\Delta}_\xi u_h \\ &= [\kappa \bar{H}^{-1} \bar{\Delta}_\xi^T \Sigma_\xi^T \bar{\Delta}_\xi v_h]^T \bar{H} u_h. \end{aligned}$$

Consequently, the dissipation's contribution to  $\delta J_h$  is

$$\frac{\partial}{\partial u_h} [v_h^T \bar{H} \bar{S}_\xi u_h] du = [\kappa \bar{H}^{-1} \bar{\Delta}_\xi^T \Sigma_\xi^T \bar{\Delta}_\xi v_h]^T \bar{H} du + \left[ \kappa v_h^T \bar{\Delta}_\xi^T \left( \frac{\partial \Sigma_\xi}{\partial u_h} \right) \bar{\Delta}_\xi u_h \right] du$$

Assuming the continuous primal and adjoint solutions are sufficiently smooth, the second term on the right-hand side will be  $O(h^{2p+2})$ , so it will not impact the consistency of the discrete adjoint. The first term on the right-hand side becomes a dissipation operator in the discrete adjoint equation; it is analogous to  $\bar{S}_\xi u_h$  with  $\Sigma_\xi^T$  replacing  $\Sigma_\xi$ . Clearly this dissipation operator has the same truncation error as  $\bar{S}_\xi u_h$ , so we have proven that the discretization (33) remains dual consistent when the dissipation operator  $\bar{S}_\xi$  is present. The same conclusion clearly holds for the analogous  $\eta$ -coordinate dissipation.

**Remark 7.** The dissipation typically included in finite-difference schemes does not include a norm: the matrix  $\bar{H}^{-1}$  is absent from (38). In this case, if we repeat the analysis above, the dissipation term appearing in the discrete-adjoint equation becomes  $\kappa \bar{H}^{-1} \bar{\Delta}_\xi^T \Sigma_\xi^T \bar{\Delta}_\xi \bar{H} v_h$ . This term is not dual consistent, because the discrete field  $\bar{H} v_h$  is not the projection of a smooth (or even continuous) function onto the nodes.

#### 4.4. Dual-consistent interface treatment

In multiblock domains, SBP finite-difference discretizations typically use penalty terms to couple the domains [34–36, 49]. These penalties are similar in form to the boundary penalties in (33). Our goal in this section is to prove that these interface penalties are dual consistent. A dual-consistency analysis of interface penalties in scalar advection-diffusion can be found in [44].

We consider a two-subdomain problem, where the two square domains are

$$\begin{aligned} \Omega^L &= \{(\xi, \eta) | \xi \in [0, 1], \eta \in [0, 1]\} \\ \text{and} \quad \Omega^R &= \{(\xi, \eta) | \xi \in [1, 2], \eta \in [0, 1]\}. \end{aligned}$$

To keep notation simple, we assume both domains are uniformly discretized with  $n + 1$  nodes in both coordinate directions, and unknowns in each domain have the same relative ordering; thus, the SBP operators defined earlier can be used on both domains. The vectors  $u_h^L \in \mathbb{R}^{4(n+1)^2}$  and  $u_h^R \in \mathbb{R}^{4(n+1)^2}$  denote the discrete solutions on the left and right domains, respectively. We treat the vector of fluxes as a function,  $f_\xi : \mathbb{R}^{4(n+1)^2} \rightarrow \mathbb{R}^{4(n+1)^2}$ , so that it too can be used interchangeably on either domain.

Consider the SBP discretization of the Euler equations on  $\Omega^L + \Omega^R$ , ignoring boundary-condition penalties and fluxes in the  $\eta$  direction:

$$\begin{aligned} \bar{D}_\xi f_\xi(u_h^L) + \dots &= \bar{H}^{-1} \bar{E}_{\xi,1} [f_\xi(u_h^L) - f_\xi^I(u_h^L, u_h^R)] + \dots \\ \bar{D}_\xi f_\xi(u_h^R) + \dots &= -\bar{H}^{-1} \bar{E}_{\xi,0} [f_\xi(u_h^R) - f_\xi^I(u_h^R, u_h^L)] + \dots, \end{aligned} \quad (39)$$

where  $f_\xi^I$  is the interface flux function. Let node  $i$  and node  $j$  be two coincident nodes on the interface of  $\Omega^L$  and  $\Omega^R$ , respectively. Then the interface flux at node  $i$  is the Roe flux

$$[f_\xi^I(u_h^L, u_h^R)]_i \equiv \frac{1}{2} [\mathcal{F}_\xi(u_{h,i}^L) + \mathcal{F}_\xi(u_{h,j}^R)] + \frac{1}{2} |\bar{A}_\xi| (u_{h,i}^L - u_{h,j}^R), \quad (40)$$

where  $\bar{A}_\xi$  denotes the flux Jacobian evaluated at the Roe average of  $u_{h,i}^L$  and  $u_{h,j}^R$ . Dual-consistent interface penalties can be defined using a simple average of states from  $\Omega^L$  and  $\Omega^R$ , rather than the Roe average, but the analysis is more involved.

**Remark 8.** It is easy to show that the penalty (40) is equal to  $\bar{A}_\xi^-(u_{h,i}^L - u_{h,j}^R)$ ; hence, only characteristics entering  $\Omega^L$  at the interface are penalized against the value in  $\Omega^R$ .

The residual form of (39) is

$$\begin{aligned} R_h(u_h, v_h) &= (v_h^L)^T \bar{H} \bar{D}_\xi f_\xi(u_h^L) - (v_h^L)^T \bar{E}_{\xi,1} [f_\xi(u_h^L) - f_\xi^I(u_h^L, u_h^R)] \\ &\quad + (v_h^R)^T \bar{H} \bar{D}_\xi f_\xi(u_h^R) + (v_h^R)^T \bar{E}_{\xi,0} [f_\xi(u_h^R) - f_\xi^I(u_h^R, u_h^L)] + \dots \end{aligned}$$

It follows from the properties of SBP operators that the contribution to  $dJ_h$  is

$$\begin{aligned} \frac{\partial R_h}{\partial u_h} du &= - [B_\xi^T \bar{D}_\xi v_h^L]^T \bar{H} du^L - [B_\xi^T \bar{D}_\xi v_h^R]^T \bar{H} du^R \\ &\quad + \left[ \bar{H}^{-1} \bar{E}_{\xi,1} \left( \frac{\partial f_\xi^I}{\partial u_h^L} \right)^T (v_h^L - \langle v_h^R \rangle_L) \right]^T \bar{H} du^L \\ &\quad - \left[ \bar{H}^{-1} \bar{E}_{\xi,0} \left( \frac{\partial f_\xi^I}{\partial u_h^R} \right)^T (v_h^R - \langle v_h^L \rangle_R) \right]^T \bar{H} du^R + \dots \end{aligned}$$

where the dependence of  $B_\xi$  on  $u_h^L$  or  $u_h^R$  is implied by the vector it multiplies. The vector  $\langle v_h^R \rangle_L \in \mathbb{R}^{4(n+1)^2}$  denotes the mapping of the interface values of  $v_h^R$  onto the domain  $\Omega^L$ . Similarly,  $\langle v_h^L \rangle_R \in \mathbb{R}^{4(n+1)^2}$  is the mapping from interface values of  $v_h^L$  onto the interface nodes of  $\Omega^R$ .

**Remark 9.** To arrive at the expression above, we made use of the property

$$(w^L)^T \bar{E}_{\xi,1} \langle z^R \rangle_L = (\langle w^L \rangle_R)^T \bar{E}_{\xi,0} z^R,$$

where  $(w^L, w^R)$  and  $(z^L, z^R)$  are two arbitrary functions on the left and right domains. This property reflects the equivalence of performing a quadrature over the interface using  $E_{\xi,1}$  on the left or  $E_{\xi,0}$  on the right.

Since the differentials  $du^L$  and  $du^R$  can be varied independently in  $dJ_h$ , we arrive at the following coupled adjoint equations.

$$-B_\xi^T \bar{D}_\xi v_h^L + \dots = -\bar{H}^{-1} E_{\xi,1} \left( \frac{\partial f_\xi^I}{\partial u_h^L} \right)^T (v_h^L - \langle v_h^R \rangle_L) + \dots, \quad (41)$$

$$-B_\xi^T \bar{D}_\xi v_h^R + \dots = \bar{H}^{-1} E_{\xi,0} \left( \frac{\partial f_\xi^I}{\partial u_h^R} \right)^T (v_h^R - \langle v_h^L \rangle_R) + \dots. \quad (42)$$

Consider the penalty on the right-hand side of (41). Referring to the primal penalty (40), we find that the adjoint penalty at node  $i$  on domain  $\Omega^L$  is

$$\begin{aligned} \left[ \left( \frac{\partial f_\xi^I}{\partial u_h^L} \right)^T (v_h^L - \langle v_h^R \rangle_L) \right]_i &= \left[ \frac{1}{2} A_\xi^T (u_{i,h}^L) + \frac{1}{2} |\bar{A}_\xi^T| \right] (v_{h,i}^L - v_{h,j}^R) \\ &\quad + \frac{1}{2} (v_{h,i}^L - v_{h,j}^R)^T \left[ \frac{\partial |\bar{A}_\xi|}{\partial u_{h,i}^L} \right] (u_{h,i}^L - u_{h,j}^R) \end{aligned}$$

Clearly, the right-hand side vanishes if the continuous adjoint solution is substituted for  $v_h^L$  and  $u_h^R$ ; recall, we assume that  $\mathbf{v} \in [\mathbf{H}^1(\Omega)]^4$ , so  $\mathbf{v}$  is continuous on  $\Omega$ . Thus, this penalty is dual consistent. A similar analysis shows that (42) is also dual consistent.

**Remark 10.** If  $\|u_h - u\| = O(h^{p+1})$  and  $\|v_h - v\| = O(h^{p+1})$  then the term on the last line will be  $O(h^{2p+2})$ , and the remaining term will dominate. This remaining term is analogous to the characteristic far-field boundary penalty. At the interface  $\xi = 1$ , characteristic adjoint variables corresponding to incoming primal waves in  $\Omega^L$  are defined by the solution on  $\Omega^L$ , while variables corresponding to outgoing primal waves in  $\Omega^L$  are defined by the solution on  $\Omega^R$ .

## 5. Results

### 5.1. Examples of superconvergence

In this section we illustrate the impact of dual consistency on functional convergence. We consider three cases, ranging from a simple flow with an analytical solution to a realistic wing geometry. The results illustrate that, even when theoretical assumptions are violated (e.g., solution smoothness), the benefits of using a dual-consistent discretization remain significant.

The dual-consistent and dual-inconsistent schemes used to obtain the following results differ only in the boundary condition and functional implementation; the difference operators, interface penalties, and dissipation operators are identical for both the dual-consistent and dual-inconsistent

schemes. The different implementations amount to the addition of small  $O(h^{p+1})$  corrections to the boundary conditions and functional estimate, where  $(p+1)$  is the order of the discrete solution. For example, a dual-inconsistent discretization for the functional (25) is

$$J_h(u_h) = g^T \bar{E}_{\eta,0} f_{\eta}^{b,DI}(u_h),$$

where the boundary flux at node  $i$  is given by

$$\left[ f_{\eta}^{b,DI}(u_h) \right]_i \equiv (0 \quad p\partial_x\eta \quad p\partial_y\eta \quad 0) \Big|_{u_{h,i}}.$$

The pressure in the dual-inconsistent functional is evaluated using the state  $u_{h,i}$  and not the state  $Pu_{h,i}$ , as is done in the dual-consistent treatment. This subtle difference is sufficient to produce a dual-inconsistent discretization.

We emphasize that the results in this section do not require the solution of the adjoint equations: dual consistency or dual inconsistency is implicitly present, and this is sufficient to impact functional accuracy.

### 5.1.1. Vortex flow

Our first example is a 2-dimensional inviscid vortex whose streamlines are concentric circles about the origin. This isentropic flow has a smooth analytical solution and provides an important verification of the theory. The solution in polar coordinates is defined by

$$\rho(r) = \rho_i \left[ 1 + \frac{\gamma-1}{2} M_i^2 \left( 1 - \frac{r_i^2}{r^2} \right) \right]^{\frac{1}{\gamma-1}},$$

with the remaining variables obtained using isentropic relations. The subscript  $i$  indicates values along the radius  $r_i$ . Here we have chosen  $r_i = 1$ ,  $\rho_i = 2$ ,  $M_i = 0.95$ , and  $p_i = 1/\gamma$ , where  $\gamma$  is the ratio of heat capacities.

The geometry and block topology are illustrated in Figure 2. The grid consists of four curvilinear blocks that conform to the domain  $\Omega = \{(r, \theta) \mid r \in [1, 3], \theta \in [0, \pi/2]\}$ . An inviscid-wall boundary condition is applied along the inner radius,  $r = 1$ , while characteristic boundary conditions supplying the exact solution are applied along the remaining boundaries. For the functional, we calculate a nominal drag defined by the force in the  $x$  direction on the curve  $\Gamma = \{r = 1, \theta \in [0, \pi/2]\}$ . The exact (nondimensional) drag is given by  $D = -1/\gamma$ .

To conduct the grid refinement study, an analytical mapping is used to generate a set of 20 grids. Each block on a given grid is discretized uniformly into  $(n+1)$  nodes in the radial and angular directions, where  $n \in \{17 + 16i \mid 0 \leq i \leq 19\}$ . Thus, the finest grid consists of  $321 \times 321$  nodes on each of the four blocks, or 412 164 nodes in total.

We begin by examining the error in the density. On each of the four blocks, we define the density error as

$$E \equiv \sqrt{(u_h - u)^T \bar{H}(u_h - u)},$$

where  $\bar{H} = H \otimes H$  is the appropriate SBP quadrature, and  $u$  and  $u_h$  denote the analytical and discrete density values. The total density error is the sum of the four block-based errors and is an accurate approximation to the integral  $L^2$  error on  $\Omega$ .

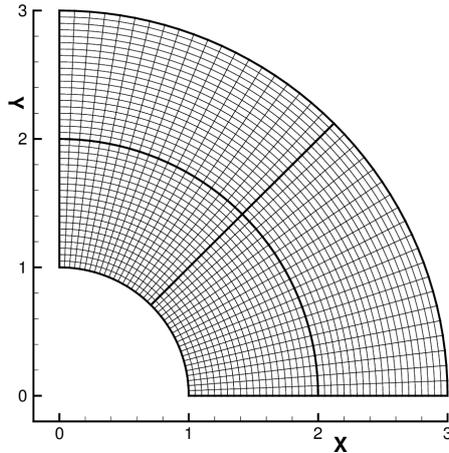


Figure 2: Geometry and block topology for the inviscid-vortex study.

Table 1: Functional convergence rate for the inviscid-vortex problem

	<b>solution design accuracy</b>		
	2nd	3rd	4th
<b>dual inconsistent</b>	2.0022	2.9219	3.9397
<b>dual consistent</b>	1.9341	4.0740	5.9095

Figure 3(a) plots the error in density versus the mesh spacing,  $h = 1/n$ . Results from second-, third-, and fourth-order accurate SBP operators are included, and both dual-consistent and dual-inconsistent treatments are shown. The discretizations obtain their design order of accuracy, and dual consistency has little impact on the solution order of accuracy.

Next, we examine the functional error,  $|J_h - \mathcal{J}|$ , plotted in Figure 3(b). Here we see a significant difference between the errors produced by the dual-consistent and dual-inconsistent discretizations. Table 1 lists the order of accuracy of the discrete functional calculated from the results on the finest two grid levels, or the finest two grid levels not affected by round-off errors in the case of the fourth-order dual-consistent results. As predicted by the theory, the third- and fourth-order dual-consistent SBP discretizations produce superconvergent functionals.

### 5.1.2. Symmetric Joukowski airfoil

For our second example, we consider the sectional drag coefficient on a symmetric Joukowski airfoil in a subsonic inviscid flow: the Mach number is 0.5 and the angle of attack is zero. For an infinite domain the drag will be zero; however, the finite domain used here has a boundary 10 chord lengths from the airfoil. Imposing the free-stream flow at the finite boundary leads to small errors in the drag coefficient that do not vanish as the grid is refined. Therefore, the grid-converged drag coefficient is not zero.

A set of four C-grids was generated for the grid refinement study. The finest grid consists of 762 032 nodes, with 1556 nodes distributed along the airfoil. The three remaining grid levels are produced by successively removing every other node from the finest grid. The coarsest grid is shown in Figure 4(a), together with a close-up view near the airfoil in Figure 4(b).

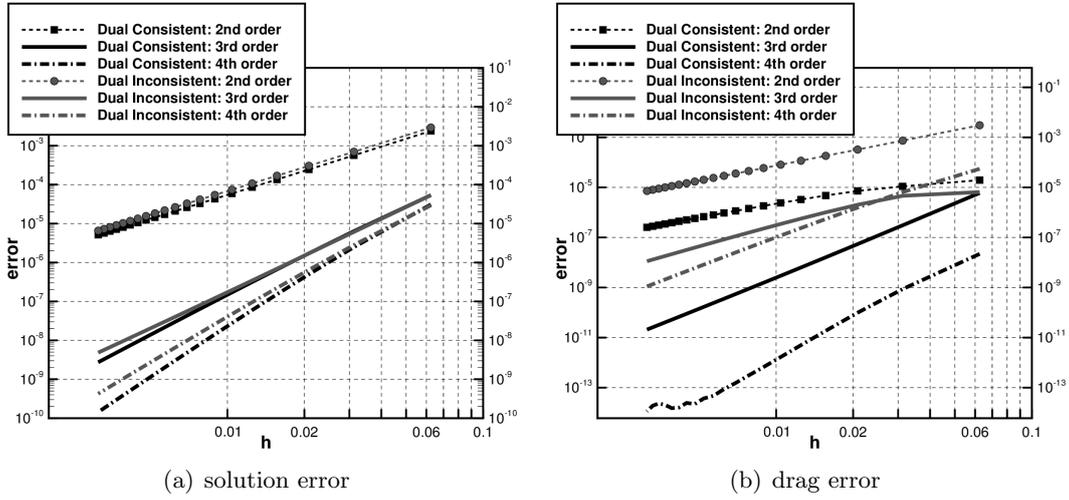


Figure 3: Solution errors in the  $L^2$  norm and the drag functional errors for the inviscid vortex flow; note the different  $y$ -axis ranges.

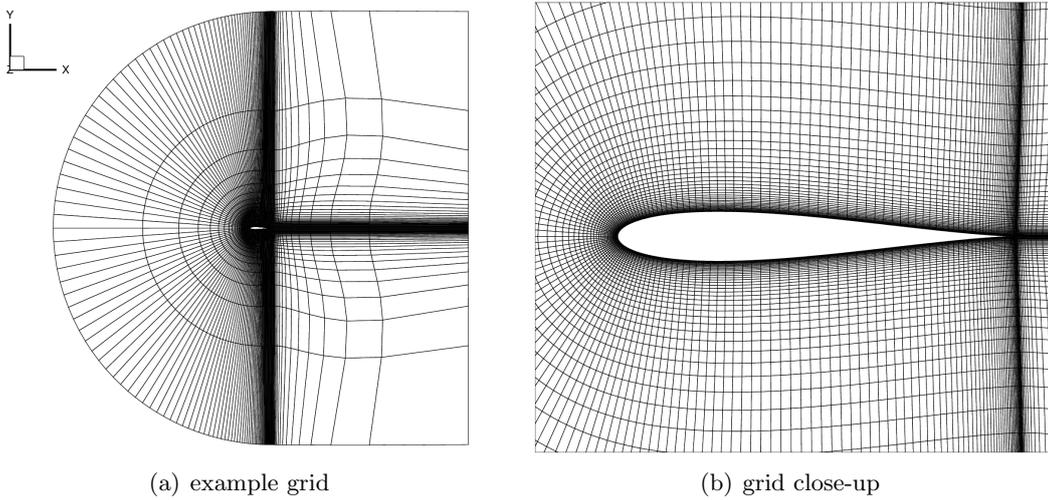


Figure 4: Coarsest grid used in the Joukowski-airfoil study, and a close-up view of the grid near the airfoil

Table 2: Estimated limiting value of  $C_d$  as  $h \rightarrow 0$ , and estimated order of accuracy for the Joukowski-airfoil example. Nonmonotone sequences or sequences that produce negative  $p$  are indicated with a dash. The grids are labeled from from coarsest, 1, to finest, 4.

		<b>dual inconsistent</b>		
		2nd	3rd	4th
<b>Limiting <math>C_d</math> Value</b>	grid set (1,2,3)	$7.588 \times 10^{-6}$	—	$2.662 \times 10^{-6}$
	grid set (2,3,4)	$2.732 \times 10^{-6}$	$2.563 \times 10^{-6}$	$2.676 \times 10^{-6}$
<b>Estimated Order</b>	grid set (1,2,3)	0.6640	—	2.6802
	grid set (2,3,4)	1.8666	1.3713	2.7762

		<b>dual consistent</b>		
		2nd	3rd	4th
<b>Limiting <math>C_d</math> Value</b>	grid set (1,2,3)	—	$2.697 \times 10^{-6}$	$2.696 \times 10^{-6}$
	grid set (2,3,4)	$2.850 \times 10^{-6}$	$2.694 \times 10^{-6}$	$2.694 \times 10^{-6}$
<b>Estimated Order</b>	grid set (1,2,3)	—	5.4262	6.3234
	grid set (2,3,4)	1.5295	2.3830	4.5682

No analytical solution is available for the flow around a Joukowski airfoil with a finite boundary, but the order of accuracy and limiting value of  $C_d$  can be estimated using the methodology of Baker [50]; see also [51]. Briefly, we use

$$C_d(h) = C_d^* + \alpha h^p$$

as an asymptotic model for the error in the drag coefficient, where  $C_d^*$  is the limiting value for the drag coefficient,  $\alpha$  is the leading-error coefficient, and  $p$  is the order of accuracy. A value of  $p$  can be estimated using the above model and data from three grid levels, provided the sequence of data points is monotone. Applying this method to the drag-coefficient functional we obtain the values listed in Table 2. The estimated order of accuracy for all the discretizations is lower than predicted by theory, based on the finest sequence of three grids. It is possible that the drag is not yet in the asymptotic regime, or that one of the assumptions of the theory is violated (e.g. the solution is not sufficiently smooth). Nevertheless, it is clear that the third- and fourth-order dual-consistent discretizations are converging at a faster rate than their dual-inconsistent counterparts.

Figure 5 shows the estimated drag-coefficient error versus a normalized mesh spacing. The error is estimated using the limiting value of the drag coefficient produced from the fourth-order dual-consistent results on the finest three grids:  $C_d^* \approx 2.694127 \times 10^{-6}$ . This choice of limiting value has a negligible impact on the dual-consistent data points in the plot, and the choice makes the dual-inconsistent data points appear more accurate than they would if the limiting value was computed from their own data.

Examining Figure 5, we find a two order of magnitude improvement going from the dual-inconsistent to the dual-consistent treatment on the three finest grids, for the third- and fourth-order methods. Indeed, the third-order dual-consistent functional using the second coarsest grid is more accurate than the third-order dual-inconsistent functional on the finest grid.

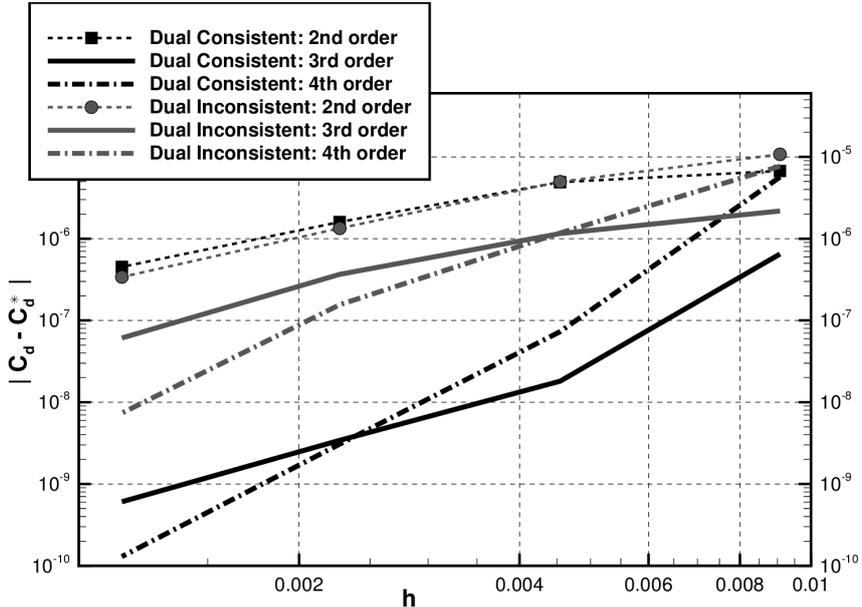


Figure 5: Joukowski study: difference between the drag coefficient and the limiting  $C_d^*$  calculated from the fourth-order dual-consistent results.

### 5.1.3. ONERA M6 wing

For the final example illustrating the impact of dual consistency on functional accuracy, we consider the ONERA M6 wing at a Mach number of 0.5 and an angle of attack of 3 degrees. In the Joukowski-airfoil study the grid sequence was generated by recursively removing every other node in each coordinate direction. This is the typical approach used for grid convergence studies of one- and two-dimensional problems. In three dimensions, these nested grids increase in size by a factor of 8 between grid levels, and this limits the number of grids that can be included in the sequence.

To overcome the above limitation, we use a multi-block grid consisting of B-spline volumes [52] that define analytical mappings; this permits any integer number of nodes to be used in a given (computational) coordinate direction. Specifically, we consider blocks with 9, 17, 25, 33, and 41 nodes in each direction. The grid topology consists of 1024 blocks, so the coarsest grid has approximately  $7.5 \times 10^5$  nodes and the finest has approximately  $7.1 \times 10^7$  nodes. The use of an analytical mapping ensures that the grids belong to the same family, which is necessary for a rigorous grid refinement study [2].

Table 3 lists the coefficients of drag computed on each grid using second- and third-order SBP schemes, with dual-inconsistent and dual-consistent formulations. The table includes the estimated limiting value of  $C_D$  as  $h \rightarrow 0$  and the estimated order of accuracy; the latter two are estimated using the 3 finest grids. The drag coefficients produced by the dual-inconsistent third-order method on the finest three grids form a nonmonotone sequence, so we cannot estimate the order of accuracy or limiting value of the functional from these grids.

Table 3 also includes the estimated percent error for each drag coefficient. The error is approximated using the limiting value from the third-order dual-consistent scheme as the “truth” value. Compared with using the limiting  $C_D$  predicted by each scheme individually, this choice

Table 3: Computed drag coefficients and their (estimated) percent error for the ONERA M6 wing study, together with the estimated limiting value of  $C_D$  as  $h \rightarrow 0$  and the estimated order of accuracy.

<b>dual inconsistent</b>				
<b>n</b>	2nd-order scheme		3rd-order scheme	
	$C_D$	% error	$C_D$	% error
8	$4.0397 \times 10^{-3}$	2.10	$4.3826 \times 10^{-3}$	10.77
16	$3.8922 \times 10^{-3}$	1.63	$3.9702 \times 10^{-3}$	0.34
24	$3.9079 \times 10^{-3}$	1.23	$3.9488 \times 10^{-3}$	0.20
32	$3.9203 \times 10^{-3}$	0.92	$3.9467 \times 10^{-3}$	0.25
40	$3.9283 \times 10^{-3}$	0.72	$3.9473 \times 10^{-3}$	0.24
<b>limiting <math>C_D</math></b>	$3.9846 \times 10^{-3}$		—	
<b>estimated order</b>	0.7252		—	

<b>dual consistent</b>				
<b>n</b>	2nd-order scheme		3rd-order scheme	
	$C_D$	% error	$C_D$	% error
8	$4.4100 \times 10^{-3}$	11.46	$4.1071 \times 10^{-3}$	3.80
16	$4.0166 \times 10^{-3}$	1.52	$3.9666 \times 10^{-3}$	0.25
24	$3.9750 \times 10^{-3}$	0.47	$3.9586 \times 10^{-3}$	0.05
32	$3.9647 \times 10^{-3}$	0.21	$3.9576 \times 10^{-3}$	0.02
40	$3.9610 \times 10^{-3}$	0.11	$3.9572 \times 10^{-3}$	0.01
<b>limiting <math>C_D</math></b>	$3.9556 \times 10^{-3}$		$3.9566 \times 10^{-3}$	
<b>estimated order</b>	2.9085		2.8564	

has a negligible impact on the dual-consistent second-order errors reported (approximately 0.01%), and the dual-inconsistent second-order errors appear smaller. Moreover, as mentioned above, no limiting value can be calculated using the third-order dual-inconsistent sequence.

As with the Joukowski airfoil, the dual-consistent schemes converge smoothly with estimated orders that are in reasonable agreement with the design order of accuracy. In contrast, the dual-inconsistent schemes converge slowly relative to their design accuracy; the second-order scheme is less than first order and the third-order scheme produces a nonmonotone sequence.

The third-order dual-consistent scheme does not appear to be asymptotically superconvergent; nevertheless, this scheme produces the most accurate functional estimates. We postulate that the degraded convergence is the result of nonsmooth features in the flow, e.g. along the trailing-edge and tip vortex. On the second coarsest grid, the estimated percent error in the dual-consistent third-order functional is 0.25%. The second-order dual-consistent scheme requires a grid approximately 8 times larger to reach this level of accuracy in  $C_D$ . The second-order dual-inconsistent scheme has not even reached this level using the finest grid, which is approximately 16 times the resolution of the second coarsest grid.

Figure 6 plots the estimated drag-coefficient error for the four schemes considered. The nominal mesh spacing is defined by  $h = 8/n$ , where  $n$  is the number of intervals in each direction on the

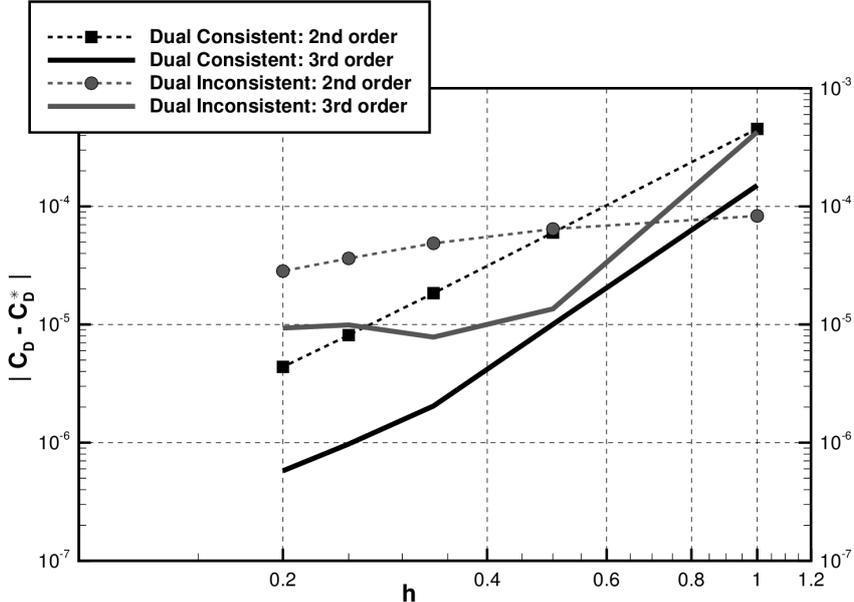


Figure 6: ONERA M6 study: difference between the  $C_D$  functional and the limiting  $C_D^*$  from the third-order dual-consistent scheme.

blocks. As with the percent errors in Table 3, the drag-coefficient error is approximated using the limiting value of the drag coefficient predicted by the third-order dual-consistent scheme.

The well-behaved convergence of the dual-consistent schemes is clearly visible in the figure. The functional computed using the third-order scheme initially converges at the theoretically predicted fourth-order rate before reducing to third order. As mentioned above, this may be caused by the trailing edge or tip vortex, which violate the smoothness assumptions required by the theory. On the finest grid, the dual-inconsistent schemes produce relatively large errors. The second-order scheme appears to yield a first-order  $C_D$  error, consistent with the estimated order in Table 3. The third-order scheme exhibits rapid convergence for large  $h$  followed by slow convergence for small  $h$ . Indeed, there appears to be a singularity of the type described by Salas and Atkins near  $h = 0.4$  in the dual-inconsistent third-order results.

### 5.2. Error estimation examples

The next set of results illustrate the impact of dual consistency on error estimates. Here, the adjoint solution must be computed explicitly, unlike the previous section. We use  $\text{GCROT}(m, k)$  [53], a variant of the truncated GCRO iterative solver [54, 55], to solve the adjoint linear system. The adjoint linear system involves the transpose of the flow Jacobian matrix; however,  $\text{GCROT}(m, k)$ , like all Krylov solvers based on Arnoldi's method, requires only products of the transposed Jacobian with arbitrary vectors. The necessary matrix-vector products are evaluated using a hand-coded version of reverse-mode algorithmic differentiation [56–58]. The adjoint system is preconditioned using a transposed variant of the parallel approximate-Schur preconditioner [59] used in the primal solution algorithm [49].

Both the primal and dual problems are solved to machine tolerance to avoid confounding effects introduced by partially converged iterative solutions; however, in practice, the tolerances should

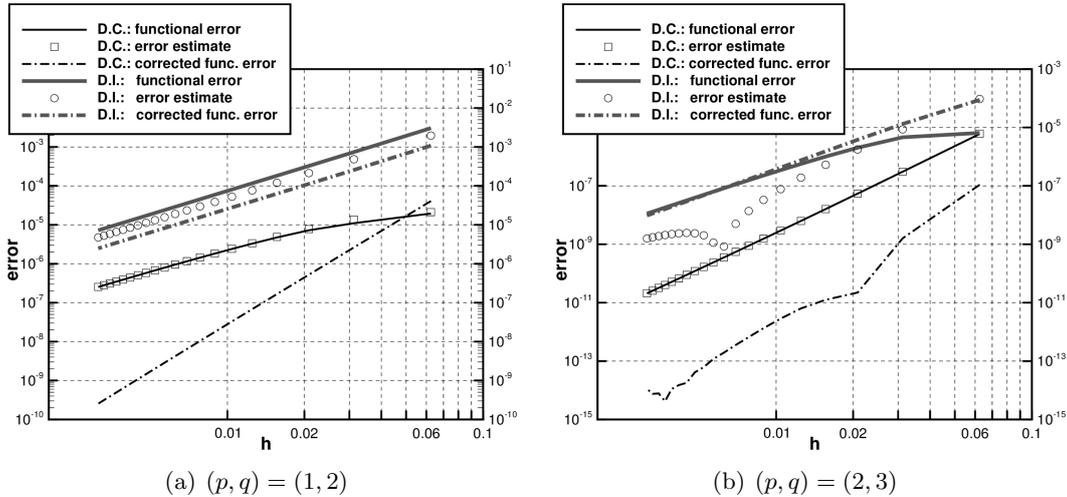


Figure 7: Results for the vortex flow functional error estimation.

be adjusted dynamically to avoid oversolving the systems when the functional error is large.

### 5.2.1. Vortex flow

As a simple illustration of the impact of dual consistency on SBP error estimates [43], we return to the inviscid vortex flow described in Section 5.1.1. The functional remains the force in the  $x$ -direction along the inner radius, and the error estimation procedure is applied on the same sequence of 20 grids.

Figure 7(a) plots the error-estimation results using a second-order discretization ( $p = 1$ ) and fourth-order residual reconstruction ( $q = 2$ ). Both the dual-consistent and dual-inconsistent schemes yield second-order accurate functionals. The error estimates from the two schemes are also second-order, but the dual-consistent error estimate is significantly more accurate. Indeed, the dual-consistent corrected functional is fourth-order accurate, as predicted by the theory. The dual-inconsistent corrected functional is three times more accurate than the baseline functional, but is only second-order accurate.

The error-estimation results for  $p = 2$  and  $q = 3$  are even more striking; see Figure 7(b). The dual-inconsistent error estimate significantly underpredicts the error for most of the grids considered; consequently, the dual-inconsistent baseline functional and corrected functional are not significantly different in terms of accuracy. Contrast this behaviour with the dual-consistent results: not only is the dual-consistent baseline functional superconvergent, its corresponding error estimate produces an asymptotically sixth-order corrected functional. On the finest grid the two error estimates are separated by six-orders of magnitude, a substantial difference considering the two flow solutions are nominally the same order of accuracy.

### 5.2.2. NACA0012 airfoil

For a more practical example of the impact of dual consistency on SBP error estimates, we consider the drag on an airfoil in subsonic flow. We use the modified NACA0012 airfoil and grid that are described in [1]. The grid is composed of 256 blocks with  $(n + 1)$  nodes in each coordinate

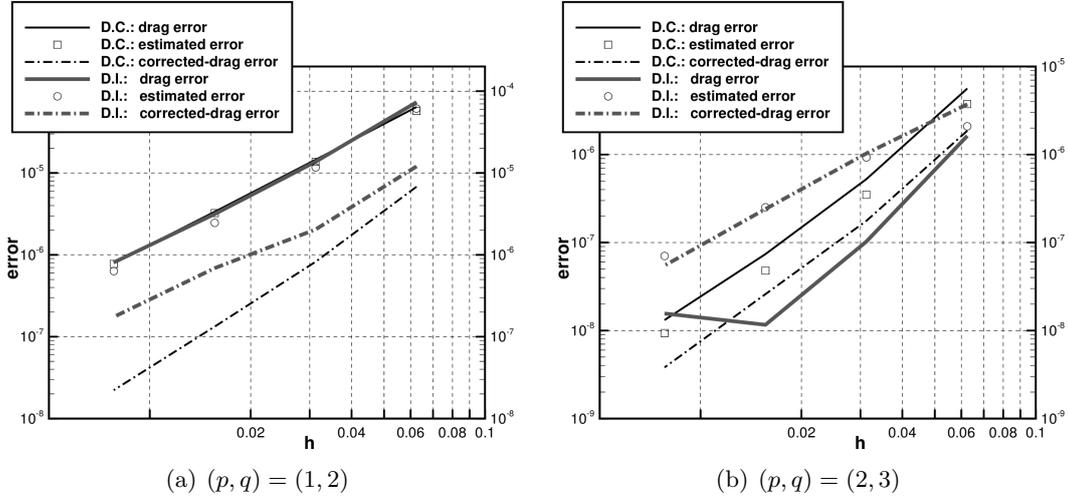


Figure 8: Results for the NACA0012 drag error estimation.

direction, where we consider  $n \in \{16, 32, 64, 128, 256\}$ . The largest grid ( $n = 256$ ) is used only in estimating the error in the computed drag.

The far-field boundary is approximately 150 chord lengths from the airfoil. The flow state used in the far-field boundary fluxes is based on a free-stream Mach number of  $M = 0.5$  and an angle of attack of 1.25 degrees. No circulation correction is applied at the far-field.

Like the Joukowski airfoil problem considered in Section 5.1.2, the present BVP has no known analytical solution. To estimate errors in the drag we require a sufficiently accurate approximation of the drag functional. As before, we use a Richardson extrapolation following the methodology of Baker [50]. The extrapolation is based on the drag values from the third-order dual-consistent scheme on the three finest grids (including  $n = 256$ ).

Figure 8(a) plots the drag error, estimated error, and corrected-drag error for the second-order discretization and fourth-order residual reconstruction, i.e.  $(p, q) = (1, 2)$ . The dual-consistent and dual-inconsistent schemes produce drag values with comparable errors; however, the dual-consistent error estimate is more accurate than the dual-inconsistent error estimate, which under predicts the baseline-drag error. The dual-consistent error estimate leads to a corrected drag that is  $O(h^{2.5})$  asymptotically and an order of magnitude more accurate than the corresponding dual-inconsistent value on the finest grid.

Results for the third-order discretization are plotted in Figure 8(b). Here, the dual-consistent error estimate systematically under predicts the baseline-drag error. Nevertheless, the corrected-drag error is approximately three times smaller than the baseline-drag error, although both have similar asymptotic behaviour. In contrast, the dual-inconsistent error estimate significantly over predicts the drag error, leading to a corrected drag that is less accurate than the baseline drag.

The baseline drag of the third-order dual-inconsistent scheme is remarkably accurate; however, we believe that this anomalous behaviour is caused by an error cancellation of the type studied in [2]. For some mesh size  $h \in [0.02, 0.03]$ , the signed errors present in the functional cancel perfectly. This type of cancellation cannot be relied upon to reduce the error on a given mesh. Indeed, it is this type of cancellation that can cause the error to increase under further refinement.

## 6. Conclusion

A dual-consistent discretization is one that leads to a set of discrete adjoint equations that converge to the continuous adjoint equations. We have shown that a dual-consistent discretization of the Euler equations using an SBP finite-difference discretization can significantly improve the accuracy of functional estimates. The results indicate that dual consistency reduces functional convergence issues of the type described by Salas and Atkins [2]. Finally, we have also demonstrated that a dual-consistent formulation leads to improved error estimates from the adjoint-weighted residual method.

The importance of dual consistency for functional accuracy in Galerkin finite-element methods is well known. However, the results presented here suggest that dual consistency, like the consistency and stability of the primal equations, is a property of fundamental importance to all methods of discretization. As we have shown, a high-order, yet dual-inconsistent, SBP discretization may not predict functionals with sufficient accuracy to outperform a second-order scheme in terms of efficiency. In contrast, high-order dual-consistent discretizations are often orders of magnitude more accurate than a second-order scheme on the same grid.

Dual consistency is a property of the discrete primal equations and the discrete functional, so a scheme does not need to solve the adjoint equations explicitly to enjoy the potential benefits<sup>5</sup>. Unfortunately, constructing a high-order dual-consistent scheme may not be straightforward in all cases. For example, it is not clear how one might construct a high-order dual-consistent finite-volume method, without satisfying an SBP property.

## Acknowledgments

The authors gratefully acknowledge financial assistance from Bombardier Aerospace, the Natural Sciences and Engineering Research Council (NSERC), the Canada Research Chairs program, Mathematics of Information Technology and Complex systems (MITACS), and the University of Toronto.

Computations were performed on the general purpose cluster at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

- [1] Vassberg, J. C. and Jameson, A., "In pursuit of grid convergence for two-dimensional Euler solutions," *AIAA Journal*, Vol. 47, No. 4, July 2010.
- [2] Salas, M. D. and Atkins, H. L., "On problems associated with grid convergence of functionals," *Computers and Fluids*, Vol. 38, No. 7, 2009, pp. 1445–1454.
- [3] Pironneau, O., "On optimum design in fluid mechanics," *Journal of Fluid Mechanics*, Vol. 64, No. 1, 1974, pp. 97–110.
- [4] Jameson, A., "Aerodynamic design via control theory," *Journal of Scientific Computing*, Vol. 3, No. 3, 1988, pp. 233–260.
- [5] Jameson, A. and Reuther, J., "Control theory based airfoil design using Euler equations," *AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Panama City Beach, Sept. 1994.
- [6] Reuther, J. J., Jameson, A., Alonso, J. J., Rimlinger, M. J., and Saunders, D., "Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers, part 1," *AIAA Journal*, Vol. 36, No. 1, Jan. 1999, pp. 51–60.

---

<sup>5</sup>Of course, the adjoint variables are required for error estimates based on the adjoint-weighted residual method.

- [7] Reuther, J. J., Jameson, A., Alonso, J. J., Rimlinger, M. J., and Saunders, D., “Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers, part 2,” *AIAA Journal*, Vol. 36, No. 1, Jan. 1999, pp. 61–74.
- [8] Anderson, W. K. and Bonhaus, D. L., “Airfoil design on unstructured grids for turbulent flows,” *AIAA Journal*, Vol. 37, No. 2, Feb. 1999, pp. 185–191.
- [9] Driver, J. and Zingg, D. W., “Numerical aerodynamic optimization incorporating laminar-turbulent transition prediction,” *AIAA Journal*, Vol. 45, No. 8, Aug. 2007, pp. 1810–1818.
- [10] Becker, R., Johnson, C., and Rannacher, R., “Adaptive error control for multigrid finite element methods,” *Journal of Computing*, Vol. 55, No. 4, 2000, pp. 271–288.
- [11] Pierce, N. A. and Giles, M. B., “Adjoint recovery of superconvergent functionals from PDE approximations,” *SIAM Review*, Vol. 42, No. 2, 2000, pp. 247–264.
- [12] Giles, M. B. and Süli, E., “Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality,” *Acta Numerica*, Vol. 11, 2002, pp. 145–236.
- [13] Venditti, D. A. and Darmofal, D. L., “Adjoint error estimation and grid adaptation for functional outputs: application to quasi-one-dimensional flow,” *Journal of Computational Physics*, Vol. 164, No. 1, 2000, pp. 204–227.
- [14] Nemec, M., Aftosmis, M. J., and Wintzer, M., “Adjoint-based adaptive mesh refinement for complex geometries,” *The 46th AIAA Aerospace Sciences Meeting and Exhibit*, No. AIAA–2008–725, Reno, Nevada, United States, Jan. 2008.
- [15] Fidkowski, K. J. and Roe, P. L., “An entropy adjoint approach to mesh refinement,” *SIAM Journal on Scientific Computing*, Vol. 32, No. 3, 2010, pp. 1261–1287.
- [16] Baysal, O. and Eleshaky, M. E., “Aerodynamic sensitivity analysis methods for the compressible Euler equations,” *Journal of Fluids Engineering*, Vol. 113, No. 4, 1991, pp. 681–688.
- [17] Frank, P. D. and Shubin, G. R., “A comparison of optimization-based approaches for a model computational aerodynamics design problem,” *Journal of Computational Physics*, Vol. 98, No. 1, Jan. 1992, pp. 74–89.
- [18] Gunzburger, M. D., *Perspectives in flow control and optimization*, Society for Industrial and Applied Mathematics, 2003.
- [19] Griewank, A., *Evaluating Derivatives*, SIAM, Philadelphia, PA, 2000.
- [20] Mohammadi, B. and Pironneau, O., “Shape Optimization in Fluid Mechanics,” *Annual Review of Fluid Mechanics*, Vol. 36, No. 1, 2004, pp. 255–279.
- [21] Squire, W. and Trapp, G., “Using complex variables to estimate derivatives of real functions,” *SIAM Review*, Vol. 40, No. 1, 1998, pp. 110–112.
- [22] Anderson, J. D., *Fundamentals of Aerodynamics*, McGraw–Hill, Inc., New York, NY, 3rd ed., 2001.
- [23] Martins, J. R. R. A., Sturdza, P., and Alonso, J. J., “The complex-step derivative approximation,” *ACM Transactions on Mathematical Software*, Vol. 29, No. 3, Sept. 2003, pp. 245–262.
- [24] Nielsen, E. J. and Kleb, B., “Efficient construction of discrete adjoint operators on unstructured grids by using complex variables,” *The 43rd AIAA Aerospace Sciences Meeting and Exhibit*, No. AIAA–2005–0324, Reno, Nevada, 2005.
- [25] Collis, S. S. and Heinkenschloss, M., “Analysis of the streamline upwind/Petrov Galerkin method applied to the solution of optimal control problems,” Tech. Rep. TR02-01, Houston, Texas, 2002.
- [26] Allmaras, S. R., “Lagrange Multiplier Implementation of Dirichlet Boundary Conditions in Compressible Navier-Stokes Finite Element Methods,” *17th AIAA Computational Fluid Dynamics Conference*, No. AIAA–2005–4714, June 2005.
- [27] Lu, J. C., *An a posteriori error control framework for adaptive precision optimization using discontinuous Galerkin finite element method*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2005.
- [28] Hartmann, R., “Adjoint Consistency Analysis of Discontinuous Galerkin Discretizations,” *SIAM Journal on Numerical Analysis*, Vol. 45, No. 6, 2007, pp. 2671–2696.
- [29] Lanczos, C., *Linear Differential Operators*, D. Van Nostrand Company, Limited, London, England, 1961.
- [30] Kreiss, H. O. and Scherer, G., “Finite element and finite difference methods for hyperbolic partial differential equations,” *Mathematical Aspects of Finite Elements in Partial Differential Equations*, edited by C. de Boor, Mathematics Research Center, the University of Wisconsin, Academic Press, 1974.
- [31] Carpenter, M. H., Gottlieb, D., and Abarbanel, S., “Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes,” *Journal of Computational Physics*, Vol. 111, No. 2, 1994, pp. 220–236.
- [32] Svård, M., Carpenter, M. H., and Nordström, J., “A stable high-order finite difference scheme for the com-

- pressible Navier-Stokes equations, far-field boundary conditions,” *Journal of Computational Physics*, Vol. 225, No. 1, July 2007, pp. 1020–1038.
- [33] Svärd, M. and Nordström, J., “A stable high-order finite difference scheme for the compressible NavierStokes equations,” *Journal of Computational Physics*, Vol. 227, No. 10, May 2008, pp. 4805–4824.
- [34] Carpenter, M. H., Nordström, J., and Gottlieb, D., “A stable and conservative interface treatment of arbitrary spatial accuracy,” *Journal of Computational Physics*, Vol. 148, No. 2, 1999, pp. 341–365.
- [35] Carpenter, M. H., Nordström, J., and Gottlieb, D., “Revisiting and extending interface penalties for multi-domain summation-by-parts operators,” *Journal of Scientific Computing*, June 2009.
- [36] Nordström, J., Gong, J., van der Weide, E., and Svärd, M., “A stable and conservative high order multi-block method for the compressible Navier-Stokes equations,” *Journal of Computational Physics*, Vol. 228, No. 24, 2009, pp. 9020–9035.
- [37] Mattsson, K., Svärd, M., and Nordström, J., “Stable and accurate artificial dissipation,” *Journal of Scientific Computing*, Vol. 21, No. 1, 2004, pp. 57–79.
- [38] Mattsson, K. and Nordström, J., “Summation by parts operators for finite difference approximations of second derivatives,” *Journal of Computational Physics*, Vol. 199, No. 2, 2004, pp. 503–540.
- [39] Mattsson, K., “Summation by Parts Operators for Finite Difference Approximations of Second-Derivatives with Variable Coefficients,” *Journal of Scientific Computing*, Vol. 51, No. 3, June 2012, pp. 650–682.
- [40] Svärd, M., “On coordinate transformations for summation-by-parts operators,” *Journal of Scientific Computing*, Vol. 20, No. 1, Feb. 2004, pp. 29–42.
- [41] Funaro, D. and Gottlieb, D., “A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations,” *Mathematics of Computation*, Vol. 51, No. 184, Oct. 1988, pp. 599–613.
- [42] Hicken, J. E. and Zingg, D. W., “Summation-by-parts operators and high-order quadrature,” *Journal of Computational and Applied Mathematics*, submitted Oct 2011: see also arXiv:1103.5182v1, pp. 18.
- [43] Hicken, J. E., “Output error estimation for summation-by-parts finite-difference schemes,” *Journal of Computational Physics*, Vol. 231, No. 9, May 2012, pp. 3828–3848.
- [44] Hicken, J. E. and Zingg, D. W., “Superconvergent functional estimates from summation-by-parts finite-difference discretizations,” *SIAM Journal on Scientific Computing*, Vol. 33, No. 2, 2011, pp. 893–922.
- [45] Fidkowski, K. J. and Darmofal, D. L., “Review of output-based error estimation and mesh adaptation in computational fluid dynamics,” *AIAA Journal*, Vol. 49, No. 4, April 2011, pp. 673–694.
- [46] Pulliam, T. H., “Efficient solution methods for the Navier-Stokes equations,” Tech. rep., Lecture Notes for the von Kármán Inst. for Fluid Dynamics Lecture Series: Numerical Techniques for Viscous Flow Computation in Turbomachinery Bladings, Rhode-Saint-Genèse, Belgium, Jan. 1986.
- [47] Nitsche, J., “Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind,” *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Vol. 36, No. 1, July 1971, pp. 9–15.
- [48] Swanson, R. C. and Turkel, E., “On central-difference and upwind schemes,” *Journal of Computational Physics*, Vol. 101, No. 2, 1992, pp. 292–306.
- [49] Hicken, J. E. and Zingg, D. W., “A parallel Newton-Krylov solver for the Euler equations discretized using simultaneous approximation terms,” *AIAA Journal*, Vol. 46, No. 11, Nov. 2008, pp. 2773–2786.
- [50] Baker, T. J., “Mesh generation: Art or science?” *Progress in Aerospace Sciences*, Vol. 41, No. 1, 2005, pp. 29–63.
- [51] Vassberg, J. C. and Jameson, A., “In pursuit of grid convergence, part 1: two-dimensional Euler solutions,” *27th AIAA Applied Aerodynamics Conference*, No. AIAA-2009-4114, San Antonio, Texas, 2009.
- [52] Hicken, J. E. and Zingg, D. W., “Aerodynamic optimization algorithm with integrated geometry parameterization and mesh movement,” *AIAA Journal*, Vol. 48, No. 2, Feb. 2010, pp. 400–413.
- [53] Hicken, J. E. and Zingg, D. W., “A simplified and flexible variant of GCROT for solving nonsymmetric linear systems,” *SIAM Journal on Scientific Computing*, Vol. 32, No. 3, 2010, pp. 1672–1694.
- [54] de Sturler, E., “Nested Krylov methods based on GCR,” *Journal of Computational and Applied Mathematics*, Vol. 67, 1996, pp. 15–41.
- [55] de Sturler, E., “Truncation strategies for optimal Krylov subspace methods,” *SIAM Journal on Numerical Analysis*, Vol. 36, No. 3, 1999, pp. 864–889.
- [56] Barth, T. J. and Linton, S. W., “An unstructured mesh Newton solver for compressible fluid flow and its parallel implementation,” *33rd AIAA Aerospace Sciences Meeting and Exhibit*, No. AIAA-95-0221, Reno, Nevada, 1995.
- [57] Anderson, W. K. and Bonhaus, D. L., “Airfoil design on unstructured grids for turbulent flows,” *AIAA Journal*, Vol. 37, No. 2, Feb. 1999, pp. 185–191.
- [58] Giles, M. B., Duta, M. C., Müller, J.-D., and Pierce, N. A., “Algorithm developments for discrete adjoint methods,” *AIAA Journal*, Vol. 41, No. 2, Feb. 2003, pp. 198–205.

- [59] Saad, Y. and Sosonkina, M., “Distributed Schur complement techniques for general sparse linear systems,” *SIAM Journal of Scientific Computing*, Vol. 21, No. 4, 1999, pp. 1337–1357.