AIAA AVIATION Forum
17-21 June 2019, Dallas, Texas
AIAA Aviation 2019 Forum

10.2514/6.2019-2952

Generalized Summation-By-Parts Methods: Coordinate Transformations, Quadrature Accuracy, and Functional Superconvergence

David A. Craig Penner* and David W. Zingg†
University of Toronto Institute for Aerospace Studies, Toronto, Ontario, M3H 5T6

We investigate coordinate transformations, quadrature accuracy, and functional superconvergence for diagonal-norm tensor-product generalized summation-by-parts operators. We show that projection operators of degree $r \geq 2p$ are required to preserve quadrature accuracy, and therefore functional superconvergence, in curvilinear coordinates when: (1) the Jacobian of the coordinate transformation is approximated by the same generalized summation-by-parts operator that is used to approximate the flux terms and (2) the degree of the generalized summation-by-parts operator is lower than the degree of the polynomial used to represent the geometry of interest. Legendre-Gauss-Lobatto and Legendre-Gauss element-type operators are considered. When the aforementioned condition (2) is violated for the Legendre-Gauss operators, there is an even-odd quadrature convergence pattern that is explained by the cancellation of the leading truncation error terms for the projection operators that correspond to the odd-degree Legendre-Gauss operators.

I. Introduction

This paper examines some of the practical issues associated with using generalized summation-by-parts (SBP) schemes to solve aerodynamic flows around complex geometries on curvilinear domains. In combination with simultaneous approximation terms (SATs) [1–3], which allow boundary conditions to be enforced in a stable manner, SBP methods [4, 5] provide a provably stable, conservative, and consistent way to numerically solve a wide class of linear and nonlinear partial differential equations (PDEs) [6–8]. For classical diagonal-norm SBP operators, if the discretization is dual consistent, then the underlying quadrature and solution functionals converge at the same rate as the order of the interior operator [9, 10]. The objective of the present work is to develop the conditions necessary to preserve quadrature accuracy and superconvergence on curvilinear domains discretized using generalized SBP operators that do not include one or both boundary nodes.

II. Notation and Definitions

The notation is adapted from [6, 9, 11]. Upper case script letters, e.g., $\mathcal{U}$, denote continuous functions, while lower case bold letters, e.g., $\boldsymbol{u}$, indicate the restriction of these continuous functions onto a set of nodes. A sans-serif capital letter, e.g., H, represents a matrix. Let $\Omega \subset \mathbb{R}^d$ be a d-dimensional Lipschitz domain. The inner product and norm are defined for two square-integrable real-valued functions, $\mathcal{U}, \mathcal{V} \in L^2(\Omega)$, as

$$(\mathcal{U}, \mathcal{V}) \equiv \int_\Omega \mathcal{U}\mathcal{V}\, d\Omega, \quad \text{and} \quad \|\mathcal{U}\|^2 \equiv \int_\Omega \mathcal{U}^2\, d\Omega, \tag{1}$$

which are approximated by the discrete inner product, $(\boldsymbol{u}, \boldsymbol{v})_\mathsf{H} \equiv \boldsymbol{u}^\mathrm{T}\mathsf{H}\boldsymbol{v}$, and norm, $\|\boldsymbol{u}\|_\mathsf{H}^2 \equiv \boldsymbol{u}^\mathrm{T}\mathsf{H}\boldsymbol{u}$, respectively. We use the following definition of a generalized SBP operator [11].

Definition 1. *Generalized summation-by-parts operator for the first derivative: A matrix operator, $\mathsf{D} \in \mathbb{R}^{(n+1)\times(n+1)}$, is an SBP operator that approximates the derivative $\frac{\partial}{\partial x}$, on the nodal distribution $\Omega_{\boldsymbol{x}} \in [a, b]$ having $n+1$ nodes, of degree p if*

*Ph.D. Candidate, Institute for Aerospace Studies, 4925 Dufferin Street, and AIAA Student Member (david.craigpenner@mail.utoronto.ca).
†University of Toronto Distinguished Professor of Computational Aerodynamics and Sustainable Aviation, Director, Centre for Research in Sustainable Aviation, Director, Centre for Computational Science and Engineering, Institute for Aerospace Studies, 4925 Dufferin Street, and AIAA Associate Fellow (dwz@oddjob.utias.utoronto.ca).

1) $\mathsf{D}\boldsymbol{x}^k = \mathsf{H}^{-1}\mathsf{Q}\boldsymbol{x}^{k-1} = k\boldsymbol{x}^{k-1}, \quad k = 0, 1, \ldots, p;$

2) H*, the norm matrix, is symmetric and positive definite; and*

3) $\mathsf{Q} + \mathsf{Q}^{\mathrm{T}} = \mathsf{E}$*, where* $\left(\boldsymbol{x}^i\right)^{\mathrm{T}} \mathsf{E}\boldsymbol{x}^j = b^{i+j} - a^{i+j}, \quad i, j, = 0, 1, \ldots r, r \geq p.$

From Definition 1, we see that the accuracy of an SBP operator is expressed in terms of the maximum degree of monomial for which it is exact. For operators constructed according to Definition 1 on tensor-product domains, it is common to decompose E as

$$\mathsf{E} = \boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}} - \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}, \quad \text{where} \quad \boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{x}^k = a^k, \quad \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{x}^k = b^k, \quad k = 0, 1, \ldots, r. \tag{2}$$

Throughout this work we refer to $\boldsymbol{t}_{\mathrm{R}}$ and $\boldsymbol{t}_{\mathrm{L}}$ as projection operators.

We define two classes of SBP operators: classical SBP operators and generalized SBP operators [11]. Classical SBP operators are constructed with a repeated interior stencil on a uniform nodal distribution that includes both boundary nodes. Generalized SBP operators can be constructed on nonuniform nodal distributions that do not include one or both boundary nodes. A classical diagonal-norm SBP operator of degree p is associated with a degree $\tau = 2p - 1$ quadrature rule, while a generalized diagonal-norm SBP operator of degree p is associated with a degree $\tau \geq 2p - 1$ quadrature rule [11]. The order of a quadrature is equal to $\tau + 1$. Note that classical SBP operators are a subset of generalized SBP operators.

III. Analysis

Solving PDEs on complex geometries normally requires the use of a curvilinear coordinate transformation which relates points in the physical domain to points in a reference space. For classical SBP operators, the impact that the geometric terms introduced by the coordinate transformation have on the accuracy of diagonal-norm SBP quadrature was studied in [12]. It was found that classical diagonal-norm SBP quadrature retains its order $2p$ theoretical accuracy in curvilinear coordinates when the Jacobian of the transformation is constructed using the SBP derivative operator associated with the quadrature (i.e., the norm, H). For tensor-product domains, besides retaining quadrature accuracy for classical SBP operators, the motivation for constructing the Jacobian using the SBP derivative operator associated with the norm arises from studies by, for example, [13], that imply that the derivative operators used to approximate the fluxes should also be used to compute the metrics to satisfy the metric invariants.

To see the effect of curvilinear transformations on the quadrature accuracy of generalized SBP schemes that do not include one or both boundary nodes, we first formalize the coordinate transformation in a one-dimensional setting, similar to [12]. We decompose the physical domain, $\Omega_{\boldsymbol{x}} \in [a, b]$, into n_e non-overlapping elements K_i, such that $\overline{\Omega_{\boldsymbol{x}}} = \cup_{i=1}^{n_e} \overline{K_i}$, $K_i \cap K_j = \emptyset$, $i \neq j$. Suppose there exists an invertible transformation that maps the projection of the physical domain onto the element K_i, $(\Omega_{\boldsymbol{x}})_{K_i}$, to the reference domain, $\Omega_{\boldsymbol{\xi}} \in [0, 1]$. The change of variable theorem gives

$$\int_a^b \mathcal{U}\, dx = \sum_{i=1}^{n_e} \int_0^1 \mathcal{U}_{K_i} \mathcal{J}\, d\xi, \tag{3}$$

where $\mathcal{J} = \frac{dx_{K_i}}{d\xi}$ is the Jacobian of the inverse transformation, and $\mathcal{U}_{K_i}$ is $\mathcal{U}$ on the element K_i. The right-hand side of Eq. (3) can be approximated as

$$\sum_{i=1}^{n_e} \boldsymbol{u}_{K_i}^{\mathrm{T}} \mathsf{H}\mathsf{D}_{\boldsymbol{\xi}}\boldsymbol{x}_{K_i} = \sum_{i=1}^{n_e} \boldsymbol{u}_{K_i}^{\mathrm{T}} \mathsf{Q}_{\boldsymbol{\xi}}\boldsymbol{x}_{K_i}. \tag{4}$$

Here, $\mathsf{D}_{\boldsymbol{\xi}}$ and $\mathsf{Q}_{\boldsymbol{\xi}}$ are defined on the reference domain. For classical SBP operators, to show that Eq. (4) is a $2p$-order accurate approximation to the right-hand side of Eq. (3), [12] introduces and proves the following theorem.

Theorem 1. *Let* $\mathsf{D} = \mathsf{H}^{-1}\mathsf{Q}$ *be an SBP operator of degree p approximating the first derivative. Then*

$$(\boldsymbol{z}, \mathsf{D}\boldsymbol{u})_{\mathsf{H}} = \boldsymbol{z}^{\mathrm{T}}\mathsf{Q}\boldsymbol{u}$$

is a $2p$*-order accurate approximation to the integral*

$$\int_0^1 \mathcal{Z}\frac{d\mathcal{U}}{dx}\, dx,$$

where $\mathcal{Z}\frac{d\mathcal{U}}{dx} \in C^{2p-1}[0, 1]$.

Remark 1. *The first derivative SBP operator in Theorem 1 is a classical SBP operator constructed on a uniform domain that includes both boundary nodes.*

Remark 2. *The accuracy of the discrete quadrature, $(z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}}$, in Theorem 1 is defined in terms of order rather than degree. In this case, the quadrature rule associated with the norm matrix, H, is a degree $\tau = 2p - 1$ quadrature rule. For classical diagonal-norm SBP operators that include both boundary nodes, if $(z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}}$ is a $2p$-order approximation to the integral*

$$\int_0^1 \mathcal{Z} \frac{d\mathcal{U}}{dx}\, dx,$$

this means that

$$\int_0^1 \mathcal{Z} \frac{d\mathcal{U}}{dx}\, dx = (z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}} + O(h^{2p}),$$

where h is the uniform mesh spacing. Thus, the order of the quadrature is equal to the degree of the quadrature plus one.

For generalized SBP operators, we can prove an analogous theorem, which is essentially the same as Theorem 3.4 in [14], where the proof is similar to that given in [12] for Theorem 1 above. The theorem and proof follow.

Theorem 2. *Let $\mathsf{D} = \mathsf{H}^{-1}\mathsf{Q}$ be a generalized SBP operator of degree p approximating the first derivative operator, as in Definition 1. Then*

$$(z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}} = z^{\mathsf{T}}\mathsf{Q}\boldsymbol{u} = \int_0^1 \mathcal{Z} \frac{d\mathcal{U}}{dx} dx, \quad i, j \leq r, i + j \leq 2p,$$

where $\mathcal{Z}\frac{d\mathcal{U}}{dx} \in C^{2p-1}[0, 1]$, $z = \boldsymbol{x}^i$, and $\boldsymbol{u} = \boldsymbol{x}^j$.

Proof. Let $\boldsymbol{u}'$ and $\boldsymbol{z}'$ be the exact derivatives of $\mathcal{U}$ and $\mathcal{Z}$ evaluated at the nodes, respectively. Due to the accuracy of H, the result will follow if we can show that

$$(z, \boldsymbol{u}')_{\mathsf{H}} = (z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}}, \quad i, j \leq r, i + j \leq 2p. \tag{5}$$

First take $j \leq p$, this gives

$$\mathsf{D}\boldsymbol{u} = j\boldsymbol{x}^{j-1} = \boldsymbol{u}', \tag{6}$$

which means $(z, \boldsymbol{u}')_{\mathsf{H}} = (z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}}$ for $j \leq p$. Next, we consider $j \geq p$, which means that $i \leq p$ and $\mathsf{D}z = i\boldsymbol{x}^{i-1} = \boldsymbol{z}'$. Using $\mathsf{D}z = \boldsymbol{z}'$ along with the SBP property, $\mathsf{Q} + \mathsf{Q}^{\mathsf{T}} = \mathsf{E}$, gives

$$\begin{aligned}
(z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}} &= z^{\mathsf{T}}\mathsf{H}\mathsf{D}\boldsymbol{u} \\
&= z^{\mathsf{T}}(\mathsf{E} - \mathsf{Q}^{\mathsf{T}})\boldsymbol{u} \\
&= z^{\mathsf{T}}\mathsf{E}\boldsymbol{u} - z^{\mathsf{T}}\mathsf{Q}^{\mathsf{T}}\boldsymbol{u} \\
&= z^{\mathsf{T}}\mathsf{E}\boldsymbol{u} - (\boldsymbol{u}, \mathsf{D}z)_{\mathsf{H}} \\
&= z^{\mathsf{T}}\mathsf{E}\boldsymbol{u} - (\boldsymbol{u}, \boldsymbol{z}')_{\mathsf{H}} \\
&= z^{\mathsf{T}}\mathsf{E}\boldsymbol{u} - \int_0^1 \mathcal{U}\frac{d\mathcal{Z}}{dx}\, dx,
\end{aligned}$$

or, taking $i, j \leq r$ and using the accuracy condition on E gives,

$$\begin{aligned}
(z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}} &= z^{\mathsf{T}}\mathsf{E}\boldsymbol{u} - \int_0^1 \mathcal{U}\frac{d\mathcal{Z}}{dx}\, dx \\
&= \int_0^1 \frac{d(\mathcal{U}\mathcal{Z})}{dx}\, dx - \int_0^1 \mathcal{U}\frac{d\mathcal{Z}}{dx}\, dx \\
&= \int_0^1 \mathcal{Z}\frac{d\mathcal{U}}{dx}\, dx.
\end{aligned}$$

Therefore, $(z, \boldsymbol{u}')_{\mathsf{H}} = (z, \mathsf{D}\boldsymbol{u})_{\mathsf{H}}$ for $j \geq p$ and $i, j \leq r$, with $i + j \leq 2p$. We have thus shown the desired result. □

From Theorems 1 and 2, we can see that, for generalized SBP operators, $\boldsymbol{z}^{\mathrm{T}}\mathsf{Q}\boldsymbol{u}$ is at least a $2p$-order approximation of $\int_0^1 \mathcal{Z}\frac{d\mathcal{U}}{dx}\,dx$ if and only if $r \geq 2p$. This is a more stringent condition compared to Definition 1 where the accuracy requirement on r is $r \geq p$. The implication of Theorem 2 when considered along with the preceding one-dimensional example is that quadrature accuracy in curvilinear coordinates is decreased for generalized SBP operators if: (1) $r < 2p$, (2) the Jacobian of the transformation is approximated by the same SBP operator that is associated with the norm, and (3) an element uses a higher degree representation of the geometry compared to the degree of the SBP operator associated with that element.

The practical implication of this theoretical decrease in quadrature accuracy under the preceding conditions is the loss of superconvergent functionals in curvilinear coordinates under the above conditions. To appreciate this, we take a small detour and examine functional superconvergence *without* a curvilinear transformation for tensor-product SBP operators. We build upon [9], which explains functional superconvergence for classical tensor-product SBP discretizations, and [15], which extends the theory of superconvergent linear functionals to generalized SBP time-marching methods.

Consider the one-dimensional problem

$$\begin{aligned} \frac{d\mathcal{U}(x)}{dx} &= \mathcal{F}(x) \quad \forall x \in \Omega = [0,1], \\ \mathcal{U}(x=0) &= \mathcal{U}_{\mathrm{L}}, \end{aligned} \tag{7}$$

and the associated discretization

$$\mathsf{D}\boldsymbol{u}_h = \boldsymbol{f} - \overbrace{\mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{L}}(\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{u}_h - \mathcal{U}_{\mathrm{L}})}^{\text{SAT}}. \tag{8}$$

We introduce the linear functional,

$$\mathcal{I}(\mathcal{U}) = \int_\Omega \mathcal{G}\mathcal{U}\,dx + \alpha(\mathcal{U})|_{x=1}, \tag{9}$$

and the discrete approximation of the preceding functional,

$$\mathcal{I}_h(\boldsymbol{u}_h) = \boldsymbol{g}^{\mathrm{T}}\mathsf{H}\boldsymbol{u}_h + \alpha\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h. \tag{10}$$

Here, α is a scalar constant. The dual problem associated with this PDE and linear functional is [9]

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

The question we are interested in is: how well does $\mathcal{I}_h(\boldsymbol{u}_h)$ approximate $\mathcal{I}(\mathcal{U})$? To answer this question we require the following two lemmas.

Lemma 1. *The term $\boldsymbol{t}_R^{\mathrm{T}}\boldsymbol{u}_h$ is a degree τ and order $\tau+1$ approximation of $\mathcal{U}|_{x=1}$.*

Proof. We have

$$\begin{aligned} \mathcal{U}|_{x=1} - \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h &= \int_\Omega \frac{d\mathcal{U}}{dx}\,dx + \mathcal{U}|_{x=0} - \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h && \text{(expand } \mathcal{U}|_{x=1}\text{)} \\ &= \int_\Omega \mathcal{F}\,dx + \mathcal{U}_{\mathrm{L}} - \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h && \text{(definition of PDE)} \\ &= \mathbf{1}^{\mathrm{T}}\mathsf{H}\boldsymbol{f} + \mathbf{1}^{\mathrm{T}}\boldsymbol{t}_{\mathrm{L}}\mathcal{U}_{\mathrm{L}} - \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h + O(h^{\tau+1}), && \text{(insert H)} \end{aligned}$$

since $\mathbf{1}^{\mathrm{T}}\boldsymbol{t}_{\mathrm{L}} = 1$ and, using the accuracy of H,

$$\mathbf{1}^{\mathrm{T}}\mathsf{H}\boldsymbol{f} = \int_\Omega \mathcal{F}\,dx + O(h^{\tau+1}).$$

Also, from the definition of the discretization of the PDE we have

$$\mathsf{H}\boldsymbol{f} + \boldsymbol{t}_{\mathrm{L}}\mathcal{U}_{\mathrm{L}} = (\mathsf{Q} + \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}})\boldsymbol{u}_h$$

and using the summation-by-parts (SBP) property gives

$$\mathsf{H}\boldsymbol{f} + \boldsymbol{t}_\mathrm{L}\mathcal{U}_\mathrm{L} = (-\mathsf{Q}^\mathrm{T} + \boldsymbol{t}_\mathrm{R}\boldsymbol{t}_\mathrm{R}^\mathrm{T})\boldsymbol{u}_h.$$

Therefore,

$$\begin{aligned}
\mathcal{U}|_{x=1} - \boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}_h &= \mathbf{1}^\mathrm{T}(\mathsf{H}\boldsymbol{f} + \boldsymbol{t}_\mathrm{L}\mathcal{U}_\mathrm{L}) - \boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}_h + O(h^{\tau+1}) && \text{(collect terms)}\\
&= \mathbf{1}^\mathrm{T}(\mathsf{Q} + \boldsymbol{t}_\mathrm{L}\boldsymbol{t}_\mathrm{L}^\mathrm{T})\boldsymbol{u}_h - \boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}_h + O(h^{\tau+1}) && \text{(definition of discrete PDE)}\\
&= \mathbf{1}^\mathrm{T}(-\mathsf{Q}^\mathrm{T} + \boldsymbol{t}_\mathrm{R}\boldsymbol{t}_\mathrm{R}^\mathrm{T})\boldsymbol{u}_h - \boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}_h + O(h^{\tau+1}), && \text{(SBP property)}
\end{aligned}$$

or, since $\mathbf{1}^\mathrm{T}\boldsymbol{t}_\mathrm{R} = 1$ and $\mathsf{Q}\mathbf{1} = \mathbf{0}$,

$$\mathcal{U}|_{x=1} - \boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}_h = O(h^{\tau+1}).$$

This concludes the proof. □

Remark 3. *Lemma 1 is essentially a variant of Theorem 4 in [15].*

Lemma 2. *Suppose we have $\psi\frac{d\mathcal{U}}{dx} \in C^{\min(2p,\tau)}$, where ψ and $\mathcal{U}$ are the dual and primal solutions, respectively, associated with the PDE and functional defined by Eq. (7) and Eq. (9), respectively. Then,*

$$\boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_L(\boldsymbol{t}_L^\mathrm{T}\boldsymbol{u} - \mathcal{U}_L)) - \alpha\boldsymbol{t}_R^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_R$$

is a degree $\min(2p, \tau)$ *and order* $\min(2p+1, \tau+1)$ *approximation of*

$$\int_\Omega \psi\frac{d\mathcal{U}}{dx}\,dx,$$

i.e.,

$$\boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_L(\boldsymbol{t}_L^\mathrm{T}\boldsymbol{u} - \mathcal{U}_L)) - \alpha\boldsymbol{t}_R^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_R = \int_\Omega \psi\frac{d\mathcal{U}}{dx}\,dx + O(h^{\min(2p+1,\tau+1)}).$$

Proof. We approach this proof in the same manner as the proof of Theorem 2. From the accuracy of H we have

$$\int_\Omega \psi\frac{d\mathcal{U}}{dx}\,dx = (\boldsymbol{\psi}, \boldsymbol{u}')_\mathsf{H} + O(h^{\tau+1}). \tag{12}$$

Here, $\boldsymbol{u}'$ is $\frac{d\mathcal{U}}{dx}$ at the nodes. Based on the accuracy of H, we would like to show

$$(\boldsymbol{\psi}, \boldsymbol{u}')_\mathsf{H} = \boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_\mathrm{L}(\boldsymbol{t}_\mathrm{L}^\mathrm{T}\boldsymbol{u} - \mathcal{U}_\mathrm{L})) - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R} + O(h^{\min(2p+1,\tau+1)}). \tag{13}$$

Using a similar argument as in [12], it is sufficient to show that the preceding equation is exact for polynomial integrands of degree less than $2p+1$. To this end, we consider $\boldsymbol{\psi} = \boldsymbol{p}_k$ and $\boldsymbol{u} = \boldsymbol{p}_m$ as degree k and m polynomials, respectively, where $k + m \le 2p+1$ defines the highest permissible degree of the combined integrand. We begin by taking $m \le p$, which gives

$$\boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_\mathrm{L}(\boldsymbol{t}_\mathrm{L}^\mathrm{T}\boldsymbol{u} - \mathcal{U}_\mathrm{L})) - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R} = \boldsymbol{\psi}^\mathrm{T}\mathsf{H}\boldsymbol{u}' = (\boldsymbol{\psi}, \boldsymbol{u}')_\mathsf{H}.$$

Next, we reverse the situation and take $m > p$, which means we must have $k < p+1$ due to the condition that $k + m \le 2p+1$. For this case we have

$$\begin{aligned}
\boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_\mathrm{L}(\boldsymbol{t}_\mathrm{L}^\mathrm{T}\boldsymbol{u} - \mathcal{U}_\mathrm{L})) - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R} &= \boldsymbol{\psi}^\mathrm{T}(\mathsf{E} - \mathsf{Q}^\mathrm{T})\boldsymbol{u} + \boldsymbol{\psi}^\mathrm{T}\boldsymbol{t}_\mathrm{L}(\boldsymbol{t}_\mathrm{L}^\mathrm{T}\boldsymbol{u} - \mathcal{U}_\mathrm{L}) - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R} && \text{(SBP property)}\\
&= \boldsymbol{\psi}^\mathrm{T}(\boldsymbol{t}_\mathrm{R}\boldsymbol{t}_\mathrm{R}^\mathrm{T} - \mathsf{Q}^\mathrm{T})\boldsymbol{u} - \boldsymbol{\psi}^\mathrm{T}\boldsymbol{t}_\mathrm{L}\mathcal{U}_\mathrm{L} - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R} && \text{(definition of E)}\\
&= \cancel{\alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}} - \boldsymbol{\psi}^\mathrm{T}\mathsf{Q}^\mathrm{T}\boldsymbol{u} - \psi_\mathrm{L}\mathcal{U}_\mathrm{L} - \cancel{\alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u}} + \psi_\mathrm{R}\mathcal{U}_\mathrm{R} && (\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{\psi} = \psi_\mathrm{R} = \alpha)\\
&= -(\mathsf{D}\boldsymbol{\psi})^\mathrm{T}\mathsf{H}\boldsymbol{u} + \psi_\mathrm{R}\mathcal{U}_\mathrm{R} - \psi_\mathrm{L}\mathcal{U}_\mathrm{L}\\
&= -(\boldsymbol{\psi}')^\mathrm{T}\mathsf{H}\boldsymbol{u} + \psi\mathcal{U}|_{x=0}^{x=1}\\
&= -\int_\Omega \frac{d\psi}{dx}\mathcal{U}\,dx + \psi\mathcal{U}|_{x=0}^{x=1} + O(h^{\tau+1}) && \text{(accuracy of H)}\\
&= \int_\Omega \psi\frac{d\mathcal{U}}{dx}\,dx + O(h^{\tau+1}) && \text{(integ. by parts)}
\end{aligned}$$

Since $k + m \le 2p+1$, we see that $\boldsymbol{\psi}^\mathrm{T}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_\mathrm{L}(\boldsymbol{t}_\mathrm{L}^\mathrm{T}\boldsymbol{u} - \mathcal{U}_\mathrm{L})) - \alpha\boldsymbol{t}_\mathrm{R}^\mathrm{T}\boldsymbol{u} + \alpha\mathcal{U}_\mathrm{R}$ is an order $\min(2p+1, \tau+1)$ approximation of $\int_\Omega \psi\frac{d\mathcal{U}}{dx}\,dx$, which concludes the proof. □

Remark 4. *Lemma 2 is similar to Lemma 12 in [15], however here we consider general α rather than $\alpha = 0$.*

We now return to our initial question and state the following theorem as an answer.

Theorem 3. *Let $\boldsymbol{u}_h$ be the discrete solution that satisfies Eq. (8). Then $\mathcal{I}_h(\boldsymbol{u}_h)$ (defined by Eq. (10)) is an order* $\min(2p+1, \tau+1)$ *approximation of $\mathcal{I}(\mathcal{U})$ (defined by Eq. (9)).*

Proof. Consider

$$\begin{aligned}
\mathcal{I}(\mathcal{U}) - \mathcal{I}_h(\boldsymbol{u}_h) &= \int_\Omega \mathcal{G}\mathcal{U}\, dx + \alpha(\mathcal{U})|_{x=1} - \boldsymbol{g}^{\mathrm{T}}\mathsf{H}\boldsymbol{u}_h - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h \\
&= \int_\Omega \mathcal{G}\mathcal{U}\, dx - \boldsymbol{g}^{\mathrm{T}}\mathsf{H}\boldsymbol{u}_h + \alpha((\mathcal{U})|_{x=1} - \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h) \\
&= \int_\Omega \mathcal{G}\mathcal{U}\, dx - \boldsymbol{g}^{\mathrm{T}}\mathsf{H}\boldsymbol{u}_h + O(h^{\tau+1}) && \text{(using Lemma 1)} \\
&= \boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + O(h^{\tau+1}). && \text{(accuracy of H)}
\end{aligned}$$

Next, we must determine the order of the term $\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h)$. To do this, we first note that the discretization of the dual problem associated with the primal PDE is

$$-\mathsf{D}\boldsymbol{\psi}_h = \boldsymbol{g} - \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}}(\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\psi}_h - \alpha). \tag{14}$$

Based on our PDEs, we can define the truncation error associated with the primal and dual problems, respectively, as

$$\begin{aligned}
e_u &= \mathsf{D}\boldsymbol{u} - \boldsymbol{u}' + \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{L}}(\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{u} - \mathcal{U}_{\mathrm{L}}) \\
e_\psi &= -\mathsf{D}\boldsymbol{\psi} + \boldsymbol{\psi}' + \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}}(\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\psi} - \alpha).
\end{aligned}$$

Multiplying e_u by H and rearranging gives

$$\begin{aligned}
\mathsf{H}e_u - (\mathsf{Q} + \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}})\boldsymbol{u} + (\mathsf{H}\boldsymbol{f} + \boldsymbol{t}_{\mathrm{L}}\mathcal{U}_{\mathrm{L}}) &= \boldsymbol{0} \\
\mathsf{H}e_u - \mathsf{A}(\boldsymbol{u} - \boldsymbol{u}_h) &= \boldsymbol{0},
\end{aligned}$$

where we have introduced $\mathsf{A} \equiv \mathsf{Q} + \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}$ (note that from the definition of the discretization of the primal problem we can write $\mathsf{A}\boldsymbol{u}_h = \mathsf{H}\boldsymbol{f} + \boldsymbol{t}_{\mathrm{L}}\mathcal{U}_{\mathrm{L}}$). Adding $\boldsymbol{\psi}_h^{\mathrm{T}}\boldsymbol{0} = 0 = \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u - \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{A}(\boldsymbol{u} - \boldsymbol{u}_h)$ to the discrete integral equation, $\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h)$, gives

$$\begin{aligned}
\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) &= \boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u - \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{A}(\boldsymbol{u} - \boldsymbol{u}_h) \\
&= \boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) - \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{A}\mathsf{H}^{-1}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u && \text{(insert } \mathsf{I} = \mathsf{H}^{-1}\mathsf{H}) \\
&= (\boldsymbol{g}^{\mathrm{T}} - \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{A}\mathsf{H}^{-1})\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u \\
&= (\boldsymbol{g} - \mathsf{H}^{-1}\mathsf{A}^{\mathrm{T}}\boldsymbol{\psi}_h)^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u \\
&= (\boldsymbol{g} - \mathsf{H}^{-1}(-\mathsf{Q} + \boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}})\boldsymbol{\psi}_h)^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u && (\mathsf{A}^{\mathrm{T}} = -\mathsf{Q} + \boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}) \\
&= (\mathsf{D}\boldsymbol{\psi}_h + \boldsymbol{g} - \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\psi}_h)^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u \\
&= (\mathsf{D}\boldsymbol{\psi}_h + \boldsymbol{g} - \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}}(\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\psi}_h - \alpha))^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) - \alpha(\mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}})^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u \\
&= \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u - \alpha(\mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}})^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h),
\end{aligned}$$

where the first term in the final line above is zero due to the definition of the discretization of the dual problem. Continuing by adding and subtracting $\boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u$, we have

$$\begin{aligned}
\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) &= -\alpha(\mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{R}})^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{\psi}_h^{\mathrm{T}}\mathsf{H}e_u + \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u - \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u \\
&= (\boldsymbol{\psi}_h - \boldsymbol{\psi})^{\mathrm{T}}\mathsf{H}e_u + \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u - \alpha\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u} - \boldsymbol{u}_h).
\end{aligned}$$

But, note that $\mathsf{A}^{\mathrm{T}}(\boldsymbol{\psi} - \boldsymbol{\psi}_h) = \mathsf{H}e_\psi$, which means

$$\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \boldsymbol{u}_h) = -(\mathsf{A}^{-\mathrm{T}}\mathsf{H}e_\psi)^{\mathrm{T}}\mathsf{H}e_u + \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u - \alpha\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u} - \boldsymbol{u}_h).$$

Furthermore, making the assumption that $\left\|\mathsf{A}^{-\mathrm{T}}\mathsf{H}\right\|_\infty \leq C$, where C is a constant, gives

$$\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u}-\boldsymbol{u}_h) = \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_h) + O(h^{\tau+1}).$$

Alternatively, we can substitute e_u into the above and write,

$$\begin{aligned}
\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u}-\boldsymbol{u}_h) &= \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}e_u - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_h) + O(h^{\tau+1}) \\
&= \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{L}}(\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{u} - \mathcal{U}_{\mathrm{L}})) - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_h) - \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}\boldsymbol{u}' + O(h^{\tau+1}) \\
&= \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{L}}(\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{u} - \mathcal{U}_{\mathrm{L}})) - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{u}_h) - \int_\Omega \psi \frac{d\mathcal{U}}{dx}\, dx + O(h^{\tau+1}).
\end{aligned}$$

Furthermore, using Lemma 1, we can substitute $\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u}_h = \mathcal{U}_{\mathrm{R}} + O(h^{\tau+1})$ into the above and write

$$\begin{aligned}
\boldsymbol{g}^{\mathrm{T}}\mathsf{H}(\boldsymbol{u}-\boldsymbol{u}_h) &= \boldsymbol{\psi}^{\mathrm{T}}\mathsf{H}(\mathsf{D}\boldsymbol{u} + \mathsf{H}^{-1}\boldsymbol{t}_{\mathrm{L}}(\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{u} - \mathcal{U}_{\mathrm{L}})) - \alpha \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{u} + \alpha\mathcal{U}_{\mathrm{R}} - \int_\Omega \psi \frac{d\mathcal{U}}{dx}\, dx + O(h^{\tau+1}) \\
&= O(h^{\min(2p+1,\tau+1)}) + O(h^{\tau+1}) \qquad \text{(Lemma 2)} \\
&= O(h^{\min(2p+1,\tau+1)})
\end{aligned}$$

Therefore,

$$\mathcal{I}(\mathcal{U}) - \mathcal{I}_h(\mathcal{U}) = O(h^{\min(2p+1,\tau+1)}), \tag{15}$$

which concludes the proof. □

To summarize, thus far we have seen that we can prove the superconvergence of linear functionals *without* a curvilinear transformation for the PDE defined by Eq. (7) without invoking Theorem 2. This changes in Section 4 of [9] when Theorem 1 of the present paper (Lemma 3 of [9] and Theorem 2 of [12]) is invoked to prove Theorem 6 (involving a curvilinear transformation) of [9], which concerns the accuracy of quadratures of the form

$$\mathcal{I}_h(\boldsymbol{u}) = \boldsymbol{u}^{\mathrm{T}}(\mathsf{H}\otimes\mathsf{H})\boldsymbol{J} \approx \iint_{\Omega_{\boldsymbol{x}}} \mathcal{U}\, dx\, dy, \tag{16}$$

based on the uniform discretization of the computational domain. The proof of Theorem 6 is in Appendix A of [9], and it involves the repeated application of Theorem 1 to show that

$$\mathcal{I}_h(\boldsymbol{u}) = \boldsymbol{u}^{\mathrm{T}}(\mathsf{H}\otimes\mathsf{H})\boldsymbol{J} \tag{17}$$

is an order $2p$ approximation of

$$\mathcal{I}(\mathcal{U}) = \iint_{\Omega_{\boldsymbol{x}}} \mathcal{U}\, dx\, dy. \tag{18}$$

Because Theorem 1 does not apply to generalized SBP operators with $r < 2p$, we do not expect superconvergent functionals when: (1) $r < 2p$, (2) the Jacobian of the transformation is approximated by the same SBP operator that is associated with the norm, and (3) an element uses a higher degree representation of the geometry compared to the degree of the SBP operator associated with that element.

IV. Results

We can confirm the quadrature accuracy of different generalized SBP operators numerically by examining a two-dimensional quadrature on a curvilinear domain. We take the test problem from Section 4.2 of [12] and restate it here for completeness. Consider the domain

$$\Omega = \{(x,y) \in \mathbb{R}^2 \mid 1 \leq xy \leq 3, 1 \leq x^2 - y^2 \leq 4\},$$

and the integral

$$\begin{aligned}
\ell &= \iint_\Omega (x^2+y^2) e^{\frac{1-x^2+y^2}{3}} \sin\left(\frac{xy-1}{2}\right) dx\, dy \\
&= 3(1-e^{-1})(1-\cos(1)).
\end{aligned} \tag{19}$$

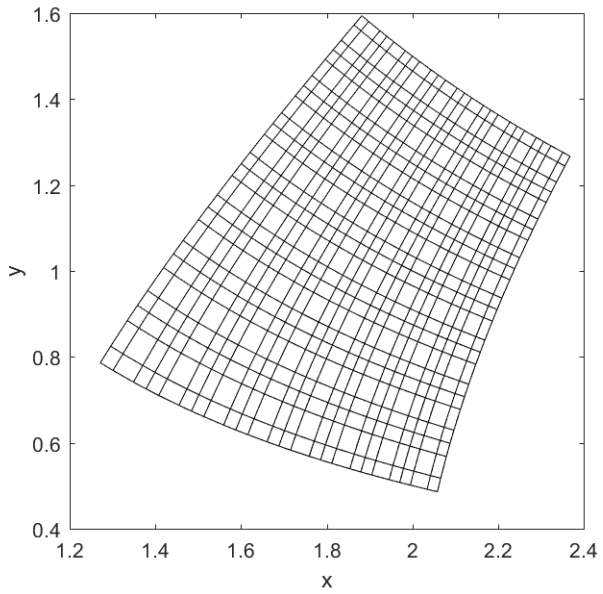

Fig. 1 Example grid for Ω with $n_e = 8$ non-overlapping elements in the $\tilde{x}$ and $\tilde{y}$ directions. The $p = 3$ LGL nodes are used in each element.

We can compute this integral numerically by introducing the global mapping coordinates

$$\tilde{x} = \frac{x^2 - y^2 - 1}{3} \quad \text{and} \quad \tilde{y} = \frac{xy - 1}{2}.$$

For a given number of non-overlapping elements, n_e, with $n + 1$ nodes in the $\tilde{x}$ and $\tilde{y}$ directions in each element, we partition the square $[0, 1]^2$ with n_e elements in the $\tilde{x}$ and $\tilde{y}$ directions. Figure 1 shows the grid for $n_e = 8$ using the $p = 3$ Legendre-Gauss-Lobatto (LGL) nodes in each element. We note here that although $\tilde{x} = \tilde{x}(x, y)$ and $\tilde{y} = \tilde{y}(x, y)$ are polynomial functions, $x = x(\tilde{x}, \tilde{y})$ and $y = y(\tilde{x}, \tilde{y})$ are not. This means that the geometry representation in each element is not a polynomial; however the geometry in each element corresponds to the analytical geometry.

On the i^{th} element, the Jacobian of the transformation is constructed as

$$\boldsymbol{J} = [(\mathsf{I} \otimes \mathsf{D})\boldsymbol{x}_i] \circ [(\mathsf{D} \otimes \mathsf{I})\boldsymbol{y}_i] - [(\mathsf{I} \otimes \mathsf{D})\boldsymbol{y}_i] \circ [(\mathsf{D} \otimes \mathsf{I})\boldsymbol{x}_i], \tag{20}$$

where $\otimes$ denotes the Kronecker product and $\circ$ denotes the Hadamard product. For a given n_e, the approximation of Eq. (19) is summed over all elements to obtain

$$\ell_{n_e} = \sum_{i=1}^{n_e} \sum_{j=1}^{n_e} \boldsymbol{J}^{\mathrm{T}} (\mathsf{H} \otimes \mathsf{H}) \boldsymbol{f}, \tag{21}$$

where $\boldsymbol{f}$ is the integrand of Eq. (19) computed using the $\boldsymbol{x}$ and $\boldsymbol{y}$ coordinates for each element. The error associated with the quadrature approximation is computed as $E_{n_e} = |\ell - \ell_{n_e}|$. Table 1 lists the generalized SBP operators that are used throughout this paper. Note that the LGL nodal distributions include the boundary nodes; and therefore the projection operators associated with this class of nodal distributions are exact (i.e., $r = \infty$). Details regarding the construction of the operators listed in Table 1 can be found in [11].

Figure 2 plots E_{n_e} as a function of $(\text{DOF})^{-1}$ for the operators listed in Table 1, and Table 2 gives the associated convergence rates. Here, $\text{DOF} = n_e(n + 1)$, i.e., DOF is the square root of the total number of degrees of freedom used to discretize the domain $\Omega_{\boldsymbol{x}}$. The convergence rates were computed using data points from the three finest grid levels, except for the LGLp4 operator. The data point on the finest grid level for the LGLp4 operator was ignored due to round-off error and the slope was instead computed using the three preceding data points.

As expected based on Theorem 2, the quadratures computed using the LGL operators, which have $r = \infty > 2p$, all converge at a rate of approximately order $\tau + 1$, where $\tau \geq 2p - 1$ denotes the degree of the quadrature rule associated

Operator	*Nodal distribution*	*Operator degree, p*	*Projection degree, r*	*Quadrature degree, τ*
LGLp1	Legendre-Gauss-Lobatto	1	∞	1
LGLp2	Legendre-Gauss-Lobatto	2	∞	3
LGLp3	Legendre-Gauss-Lobatto	3	∞	5
LGLp4	Legendre-Gauss-Lobatto	4	∞	7
LGp1	Legendre-Gauss	1	1	3
LGp2	Legendre-Gauss	2	2	5
LGp3	Legendre-Gauss	3	3	7
LGp4	Legendre-Gauss	4	4	9

Table 1 Some generalized summation-by-parts operators that satisfy Definition 1.

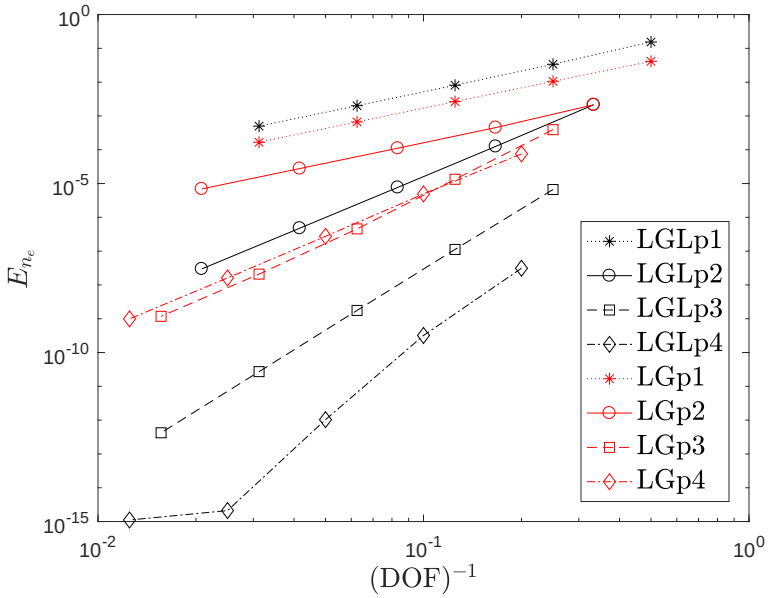

Fig. 2 Convergence of the error when using the generalized SBP operators in Table 1 to approximate Eq. (19).

with each operator. These results agree with [11]. In contrast, the LG operators all converge at rates less than $2p$, which is also expected since for these operators $r = p < 2p$, the Jacobian is computed using the same SBP operator associated with the norm, and the geometry representation is non-polynomial. Specifically, there is an even-odd convergence pattern associated with the LG operators. The even-degree LG operators converge at a rate of p while the odd-degree LG operators converge at a rate of approximately $p + 1$. This even-odd quadrature convergence behaviour can be explained by considering the interactions between the leading truncation error terms associated with the respective even- and odd-degree LG projection operators. Consider decomposing E in terms of the projection operators $\boldsymbol{t}_\mathrm{L}$ and tr

$$\mathsf{E} = \boldsymbol{t}_\mathrm{R}\boldsymbol{t}_\mathrm{R}^\mathrm{T} - \boldsymbol{t}_\mathrm{L}\boldsymbol{t}_\mathrm{L}^\mathrm{T}. \tag{22}$$

Next, recall the accuracy condition on E from Definition 1. Namely, for some $\Omega_{\boldsymbol{x}} \in [a, b]$, we have

$$\left(\boldsymbol{x}^i\right)^\mathrm{T} \mathsf{E}\boldsymbol{x}^j = b^{i+j} - a^{i+j}, \quad i, j = 0, 1, \ldots, r, r \geq p. \tag{23}$$

With Eq. (22) and Eq. (23) established, we take $\Omega_{\boldsymbol{x}} \in [-1, 1]$. Substituting $\Omega_{\boldsymbol{x}}$, i.e., $a = -1$ and $b = 1$, into Eq. (23) and moving all the terms to the right-hand side gives us an expression for the error in E, E_E, for different values of i and j.

$$E_\mathsf{E} = \left(\boldsymbol{x}^i\right)^\mathrm{T} \mathsf{E}\boldsymbol{x}^j - (1)^{i+j} + (-1)^{i+j}. \tag{24}$$

Operator	*Convergence Rate*
LGLp1	2.00994
LGLp2	4.00502
LGLp3	6.01197
LGLp4	8.60466
LGp1	1.99889
LGp2	1.99984
LGp3	4.29535
LGp4	4.05604

Table 2 Convergence rates when using the generalized SBP operators in Table 1 to approximate Eq. (19).

LGp1	i	0	1	2					
	j	3	2	1					
	i+j	$\boxed{3}$	$\boxed{3}$	$\boxed{3}$					
LGp2	i	0	1	2	3	4			
	j	3	4	3	0	1			
	i+j	$\boxed{3}$	5	5	$\boxed{3}$	5			
LGp3	i	0	1	2	3	4	5	6	
	j	5	4	5	4	1	0	1	
	i+j	$\boxed{5}$	$\boxed{5}$	7	7	$\boxed{5}$	$\boxed{5}$	7	
LGp4	i	0	1	2	3	4	5	6	7	8
	j	5	6	5	6	5	0	1	0	1
	i+j	$\boxed{5}$	7	7	9	9	$\boxed{5}$	7	7	9

Table 3 Values of i, j, and $i + j$ when the first nonzero value of E_{E} occurs. The values of i and j in Eq. (24) are increased with j running first.

For LG operators, based on Eq. (23), Eq. (24) will be equal to zero for $i, j \leq r$. Therefore, we are interested in the behaviour of Eq. (24) when $i, j > r$. Table 3 gives the values of i, j, and $i + j$ when the first nonzero value of E_{E} (i.e., Eq. (24)) occurs. For each operator, the minimum value of $i + j$ is $\boxed{\text{boxed}}$. For the odd-degree LG operators, the minimum value of $i + j$ is equal to $p + 2$. For the even-degree LG operators, the minimum value of $i + j$ is equal to $p + 1$. To understand why this pattern occurs, we can examine Eq. (24) for the minimum values of $i + j$ for each LG operator, as reported in Table 3.

From Table 3, the minimum value of $i + j$ occurs for each operator when $i = 0$. Therefore, we substitute $i = 0$ and Eq. (22) into Eq. (24), this gives

$$\begin{aligned} E_{\mathsf{E}}|_{i=0} &= \left(\boldsymbol{x}^0\right)^{\mathrm{T}} (\boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}} - \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}})\boldsymbol{x}^j - (1)^{0+j} + (-1)^{0+j} \\ &= \mathbf{1}^{\mathrm{T}}(\boldsymbol{t}_{\mathrm{R}}\boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}} - \boldsymbol{t}_{\mathrm{L}}\boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}})\boldsymbol{x}^j - (1)^j + (-1)^j \\ &= \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{x}^j - \boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{x}^j - 1 + (-1)^j, \end{aligned}$$

where we have used $\mathbf{1}$ to denote a vector of ones. Note also that $\mathbf{1}^{\mathrm{T}}\boldsymbol{t}_{\mathrm{R}} = \mathbf{1}^{\mathrm{T}}\boldsymbol{t}_{\mathrm{L}} = 1$ since the projection operators interpolate the constant function exactly. We now introduce two additional error metrics. Let

$$E_{\boldsymbol{t}_{\mathrm{R}}} = \boldsymbol{t}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{x}^j - 1, \tag{25}$$

$$E_{\boldsymbol{t}_{\mathrm{L}}} = \boldsymbol{t}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{x}^j - (-1)^j, \tag{26}$$

Operator	j	$E_{\mathsf{E}}\|_{i=0}$	$E_{\boldsymbol{t}_{\mathrm{R}}}$	$E_{\boldsymbol{t}_{\mathrm{L}}}$
LGp1	0	$0.00000e+00$	$-2.22045e-16$	$-2.22045e-16$
$j=r$	1	$-1.11022e-15$	$-2.22045e-16$	$2.22045e-16$
$j=2p$	2	$0.00000e+00$	$\boxed{-6.66667e-01}$	$\boxed{-6.66667e-01}$
	3	$\boxed{-1.33333e+00}$	$-6.66667e-01$	$6.66667e-01$
LGp2	0	$0.00000e+00$	$0.00000e+00$	$0.00000e+00$
	1	$8.88178e-16$	$0.00000e+00$	$0.00000e+00$
$j=r$	2	$0.00000e+00$	$2.22045e-16$	$2.22045e-16$
	3	$\boxed{-8.00000e-01}$	$\boxed{-4.00000e-01}$	$\boxed{4.00000e-01}$
$j=2p$	4	$0.00000e+00$	$-4.00000e-01$	$-4.00000e-01$
LGp3	0	$0.00000e+00$	$0.00000e+00$	$0.00000e+00$
	1	$0.00000e+00$	$-5.55112e-16$	$5.55112e-16$
	2	$0.00000e+00$	$0.00000e+00$	$0.00000e+00$
$j=r$	3	$0.00000e+00$	$-4.44089e-16$	$4.44089e-16$
	4	$0.00000e+00$	$\boxed{-2.28571e-01}$	$\boxed{-2.28571e-01}$
	5	$\boxed{-4.57143e-01}$	$-2.28571e-01$	$2.28571e-01$
$j=2p$	6	$0.00000e+00$	$-4.24490e-01$	$-4.24490e-01$
LGp4	0	$2.22045e-16$	$0.00000e+00$	$2.22045e-16$
	1	$0.00000e+00$	$0.00000e+00$	$2.22045e-16$
	2	$1.11022e-16$	$-2.22045e-16$	$-2.22045e-16$
	3	$-4.44089e-16$	$-2.22045e-16$	$4.44089e-16$
$j=r$	4	$1.11022e-16$	$-1.11022e-16$	$-1.11022e-16$
	5	$\boxed{-2.53968e-01}$	$\boxed{-1.26984e-01}$	$\boxed{1.26984e-01}$
	6	$1.11022e-16$	$-1.26984e-01$	$-1.26984e-01$
	7	$-5.36155e-01$	$-2.68078e-01$	$2.68078e-01$
$j=2p$	8	$1.11022e-16$	$-2.68078e-01$	$-2.68078e-01$

Table 4 Projection error associated with LG operators.

be the error associated with the projection operators $\boldsymbol{t}_{\mathrm{R}}$ and $\boldsymbol{t}_{\mathrm{L}}$, respectively. Note that we can recast $E_{\mathsf{E}}|_{i=0}$ in terms of $E_{\boldsymbol{t}_{\mathrm{R}}}$ and $E_{\boldsymbol{t}_{\mathrm{L}}}$ as

$$E_{\mathsf{E}}|_{i=0} = E_{\boldsymbol{t}_{\mathrm{R}}} - E_{\boldsymbol{t}_{\mathrm{L}}}. \tag{27}$$

Table 4 numerically tabulates these error terms for the LG operators listed in Table 1 for different j.

First note that all error terms are zero (or machine zero) up to $j = r$. This is expected from the accuracy condition on E associated with Definition 1. Therefore, we expect that $E_{\boldsymbol{t}_{\mathrm{R}}}$ and $E_{\boldsymbol{t}_{\mathrm{L}}}$ will be non-zero for $j > r$. The first non-zero (or non-machine-zero) values of $E_{\boldsymbol{t}_{\mathrm{R}}}$ and $E_{\boldsymbol{t}_{\mathrm{L}}}$ that appear when increasing j from 0 to $2p$ are $\boxed{\text{boxed}}$, and this occurs for each operator when $j = r + 1$. Similarly, the first non-zero value of $E_{\mathsf{E}}|_{i=0}$ for each operator is $\boxed{\text{boxed}}$. For the even-degree operators, the first non-zero values of $E_{\boldsymbol{t}_{\mathrm{R}}}$ and $E_{\boldsymbol{t}_{\mathrm{L}}}$ are equal and opposite in sign, which results in the first non-zero value of $E_{\mathsf{E}}|_{i=0}$ occurring at $j = r + 1$. In contrast, the first non-zero values of $E_{\boldsymbol{t}_{\mathrm{R}}}$ and $E_{\boldsymbol{t}_{\mathrm{L}}}$ are equal and of the same sign, which causes them to cancel when $j = r + 1$. As a result of this cancellation, the first non-zero value of $E_{\mathsf{E}}|_{i=0}$ for the odd-degree operators occurs at $j = r + 2$, i.e., one value of j higher than for the even-degree operators. This explains the even-odd quadrature convergence behaviour observed in Figure 2 and Table 2 with the even-degree LG operators converging at a rate of p and the odd-degree LG operators converging at a rate of $p + 1$. Essentially, in curvilinear coordinates, when the Jacobian is approximated using the same SBP derivative operator associated with the norm and the geometry representation is either non-polynomial or of a degree higher than that of the SBP operator, the

Polynomial Degree	*Mesh Function*	*Abbreviation*
1	ξ	MFD1
2	$\frac{1}{2}\xi^2 + \frac{1}{2}\xi$	MFD2
3	$\frac{1}{3}\xi^3 + \frac{1}{3}\xi^2 + \frac{1}{3}\xi$	MFD3
4	$\frac{1}{4}\xi^4 + \frac{1}{4}\xi^3 + \frac{1}{4}\xi^2 + \frac{1}{4}\xi$	MFD4
5	$\frac{1}{5}\xi^5 + \frac{1}{5}\xi^4 + \frac{1}{5}\xi^3 + \frac{1}{5}\xi^2 + \frac{1}{5}\xi$	MFD5
non-polynomial	$(\exp(4\xi) - 1)/(\exp(4) - 1)$	MFNP

Table 5 Mesh functions.

LG quadrature accuracy is limited by the accuracy of the projection operators, and the even-odd quadrature convergence behaviour is associated with the leading truncation error cancellation of the projection operators that occurs for the odd-degree LG operators.

To examine functional superconvergence, consider the steady one-dimensional linear convection equation with unit wave speed on the domain $\Omega_{\boldsymbol{x}} \in [0, 1]$

$$\frac{\partial \mathcal{U}}{\partial x} = \mathcal{S}(x), \tag{28}$$

where $\mathcal{S}(x)$ is a source term. The source term

$$\mathcal{S}(x) = \frac{\pi e^x}{e-1} \cos\left(\frac{\pi e^x - \pi + e - 1}{e-1}\right) \tag{29}$$

gives the steady-state solution

$$\mathcal{U}(x) = \sin\left(\frac{\pi(e^x - 1)}{e-1} + 1\right). \tag{30}$$

We use the mesh functions, $x(\xi)$, in Table 5 to introduce a non-constant metric Jacobian into integral functionals on $\Omega_{\boldsymbol{x}}$, which simulates the effect of a curvilinear coordinate transformation.

Consider the integral functional

$$\mathcal{I}(\mathcal{U}) = \int_0^1 \mathcal{U} dx = \int_0^1 \mathcal{U} \frac{dx}{d\xi} d\xi, \tag{31}$$

discretized as

$$\mathcal{I}_{n_e}(\boldsymbol{u}) = \sum_{i=1}^{n_e} \boldsymbol{J}^{\mathrm{T}} \mathsf{H} \boldsymbol{u}. \tag{32}$$

Table 6 gives the convergence results when using Eq. (32) to discretize Eq. (31) with the mesh functions in Table 5. Figure 3 shows the visual results when using the non-polynomial mesh function. The convergence rates were all computed using data from the asymptotic region prior to round-off error. Upwind SATs were used.

From Table 6, as for the two-dimensional quadrature example, the LGL operators converge at a rate of approximately the quadrature degree, τ, plus one, i.e., $\tau + 1$. In contrast, the LG operators converge at a rate of $\tau + 1$ only when the mesh function is a polynomial whose degree is less than or equal to the degree of the SBP operator used to compute the Jacobian. When a mesh function is used that does not satisfy this condition, the odd- and even-degree LG operators converge at rates of $p + 1$ and p, respectively. These convergence rates are boxed in Table 6. As before, the even-odd convergence behaviour can be explained by considering the error cancellation associated with the leading truncation error terms of the odd-degree LG projection operators.

V. Conclusion

We have shown that, for tensor-product generalized SBP operators, projection operators of degree $r \geq 2p$ are required to preserve quadrature accuracy and therefore superconvergent functionals in curvilinear coordinates when (1)

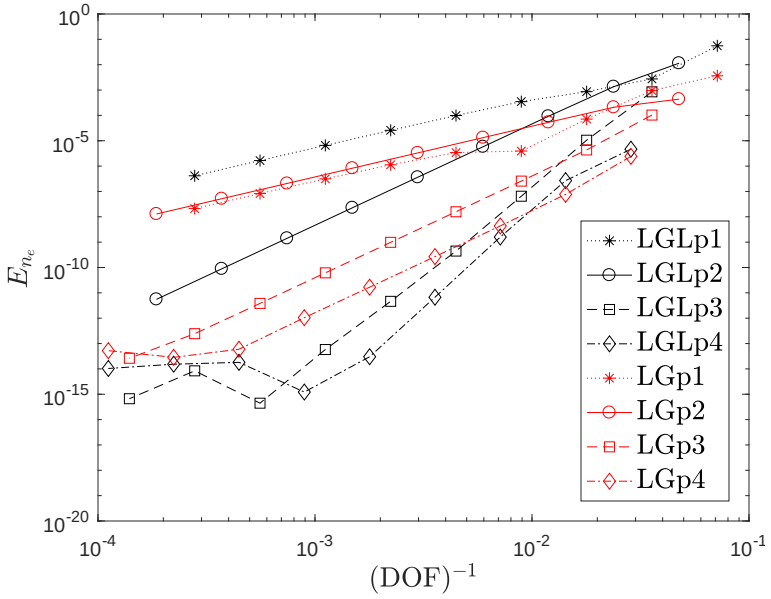

Fig. 3 Convergence of the error when using the generalized SBP operators in Table 1 to approximate Eq. (31) via Eq. (32) using the non-polynomial mesh function.

Operator	MFD1	MFD2	MFD3	MFD4	MFD5	MFNP
LGLp1	2.00306	2.00912	1.97688	1.98138	1.98822	1.99188
LGLp2	4.00483	4.19932	4.00088	4.00030	4.00005	3.99982
LGLp3	6.00458	5.98032	6.05277	5.79955	5.65710	6.88729
LGLp4	8.40074	8.63669	8.01160	7.91618	7.82547	7.84233
LGp1	4.00515	2.00105	1.98979	1.97327	1.97019	1.94877
LGp2	6.00316	5.95041	1.99924	2.00003	2.00003	2.00004
LGp3	8.69240	7.96493	8.22565	3.96046	4.01958	4.01082
LGp4	9.48193	9.38194	10.39338	10.74542	3.99982	4.01261

Table 6 Convergence rates when using the generalized SBP operators in Table 1 to approximate Eq. (31) via Eq. (32) with the mesh functions defined in Table 5.

the Jacobian of the transformation is approximated by the same SBP operator that is associated with the norm and (2) when a higher degree representation of the geometry is used compared to the degree of the SBP operator. Furthermore, when the geometry condition is violated for the LG SBP operators, which have $r = p < 2p$, there is an even-odd quadrature convergence behaviour that can be explained by considering the cancellation of the leading truncation error terms for the LG projection operators associated with the odd-degree LG operators.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada and the University of Toronto. D. A. Craig Penner thanks Dr. Pieter D. Boom for helpful discussions regarding some of the proofs in his Ph.D. thesis ([15]).

References

[1] 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, 1988, pp. 599–613.

[2] 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, 1994, pp. 220–236.

[3] Svärd, M., Carpenter, M. H., and Nordström, J., "A stable high-order finite difference scheme for the compressible Navier-Stokes equations, far-field boundary conditions," *Journal of Computational Physics*, Vol. 225, No. 1, 2007, pp. 1020–1038.

[4] 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*, Academic Press, New York/London, 1974, pp. 195–212.

[5] Strand, B., "Summation by Parts for Finite Difference Approximations for d/dx," *Journal of Scientific Computing*, Vol. 20, No. 1, 1994, pp. 47–67.

[6] Del Rey Fernández, D. C., Hicken, J. E., and Zingg, D. W., "Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations," *Computers & Fluids*, Vol. 95, 2014, pp. 171–196.

[7] Svärd, M., and Nordström, J., "Review of summation-by-parts schemes for initial–boundary-value problems," *Journal of Computational Physics*, Vol. 268, 2014, pp. 17–38.

[8] Crean, J., Hicken, J. E., Del Rey Fernández, D. C., Zingg, D. W., and Carpenter, M. H., "Entropy-Stable Summation-By-Parts Discretization of the Euler Equations on General Curved Elements," *Journal of Computational Physics*, Vol. 356, 2018, pp. 410–438.

[9] 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.

[10] Hicken, J. E., and Zingg, D. W., "Dual consistency and functional accuracy: a finite-difference perspective," *Journal of Computational Physics*, Vol. 256, 2014, pp. 161–182.

[11] Del Rey Fernández, D. C., Boom, P. D., and Zingg, D. W., "A generalized framework for nodal first derivative summation-by-parts operators," *Journal of Computational Physics*, Vol. 266, 2014, pp. 214–239.

[12] Hicken, J. E., and Zingg, D. W., "Summation-by-parts operators and high-order quadrature," *Journal of Computational and Applied Mathematics*, Vol. 237, No. 1, 2013, pp. 111–125.

[13] Thomas, P. D., and Lombard, C. K., "Geometric conservation law and its application to flow computations on moving grids," *AIAA Journal*, Vol. 17, No. 10, 1979, pp. 1030–1037.

[14] Hicken, J. E., Del Rey Fernández, D. C., and Zingg, D. W., "Multidimensional summation-by-parts operators: General theory and application to simplex elements," *SIAM Journal on Scientific Computing*, Vol. 38, No. 4, 2016, pp. A1935–A1958.

[15] Boom, P. D., "High-order implicit time-marching methods for unsteady fluid flow simulation," Ph.D. thesis, University of Toronto, 2015.