# A Generalized Framework for Nodal First Derivative Summation-By-Parts Operators

David C. Del Rey Fernández[a,*], Pieter D. Boom[a,*], David W. Zingg[a,**]

[a] *Institute for Aerospace Studies, University of Toronto, Toronto, Ontario, M3H 5T6, Canada*

## Abstract

A generalized framework is presented that extends the classical theory of finite-difference summation-by-parts (SBP) operators to include a wide range of operators, where the main extensions are i) non-repeating interior point operators, ii) nonuniform nodal distribution in the computational domain, iii) operators that do not include one or both boundary nodes. Necessary and sufficient conditions are proven for the existence of nodal approximations to the first derivative with the SBP property. It is proven that the positive-definite norm matrix of each SBP operator must be associated with a quadrature rule; moreover, given a quadrature rule there exists a corresponding SBP operator, where for diagonal-norm SBP operators the weights of the quadrature rule must be positive. The generalized framework gives a straightforward means of posing many known approximations to the first derivative as SBP operators; several are surveyed, such as discontinuous Galerkin discretizations based on the Legendre-Gauss quadrature points, and shown to be SBP operators. Moreover, the new framework provides a method for constructing SBP operators by starting from quadrature rules; this is illustrated by constructing novel SBP operators from known quadrature rules. To demonstrate the utility of the generalization, the Legendre-Gauss and Legendre-Gauss-Radau quadrature points are used to construct SBP operators that do not include one or both boundary nodes.

*Keywords:* First Derivative, Summation-by-Parts, Energy method, Quadrature, Simultaneous-Approximation-Terms, Finite-Difference

[*] Ph.D. Candidate

[**] Professor and Director, Tier 1 Canada Research Chair in Computational Aerodynamics, J. Armand Bombardier Foundation Chair in Aerospace Flight

Methods, Pseudo-Spectral, Discontinuous-Galerkin, Partial-Differential Equations
*2000 MSC:* 65D25,
*2000 MSC:* 65D30
*2000 MSC:* 6532

## 1. Introduction

The use of computers to solve partial differential equations (PDEs) numerically has reached a sufficient level of maturity where both academic and industrial codes are routinely applied to real-world problems, for example in computational fluid dynamics (CFD) [31]. However, despite exponential improvements in computational power, solution of industrially-relevant PDEs remains a time intensive endeavour. In most settings turnaround time is of paramount importance; thus computational efficiency remains a primary concern. In the early 1970s, Kreiss and Oliger [29] and Swartz and Wendroff [50] demonstrated that substantial efficiency gains can be made by use of higher-order (HO) methods. In the asymptotic region, HO methods have a local truncation error of order $O([\Delta x]^p)$, where $p \geq 3$, and $\Delta x$ is the mesh spacing. Thus, for a given accuracy, HO methods require coarser mesh spacing relative to lower-order methods. Nevertheless, in many fields, such as CFD, second-order methods are prevalent. The potential of HO methods to provide substantial gains in computational efficiency motivate their further study and development.

One of the difficulties in constructing a well-posed mathematical model for a physical process is determining appropriate boundary and initial conditions. Loosely speaking, a well-posed mathematical model is one for which a unique solution exists and small perturbations to the data, including the initial and boundary conditions and forcing function, lead to small perturbations in the solution, or alternatively, the solution depends continuously on the data, see [28, 16, 15]. Similarly, a discrete model is stable if small perturbations in the data lead to small perturbations in the solution [14]. Convergence to the true solution of the PDE requires a well-posed mathematical model as well as a stable and consistent numerical scheme. The main difficulty in developing HO methods is associated with developing numerical schemes for the implementation of boundary conditions and inter-element coupling that are efficient and stable. For finite-difference (FD) methods this has been addressed through the use of summation-by-parts (SBP) operators [30, 45, 36, 7, 35, 37, 33] with boundary conditions and inter-block coupling weakly imposed using simultaneous-approximation-terms (SATs) [10, 4, 18, 5, 39, 47, 49, 32, 41]. There are several attractive properties to the SBP-SAT approach: they lead to provably stable discretizations for linearized problems, for example the linearized Navier-Stokes equations [47, 49, 41]; they lead naturally to multi-block schemes that have constant and, more importantly, low communication overhead, which is advantageous for parallel computations. This results from the fact that only $C^0$ continuity needs to be maintained between blocks and, regardless of the order of the scheme, the same amount of information is passed between blocks. Moreover, Hicken and Zingg have shown that if the formulation is dual consistent [24], then HO-FD SBP-SAT discretizations benefit from superconvergence of functionals [22]. Other recent areas of exploration for the FD-SBP-SAT framework include ENO/WENO SBP-SAT formulations [52, 53, 8, 2]. In addition, Kitson, McLachlan and Robidoux [25] examined the existence and properties of diagonal normed SBP operators on periodic one-dimensional domains.

Despite these advantageous properties, the SBP-SAT approach has been primarily developed in the context of FD schemes, with some notable exceptions. There has been work on constructing SBP-SAT methods for finite-volume discretizations [38, 40, 48, 17] and collocated-pseudo-spectral implementations [3, 20, 19, 18]. Chiu, Hu and Jameson [6] developed an algorithm for constructing SBP operators for mesh-free schemes where meshes are replaced with point clouds on the interior of the domain and point distributions at the boundaries of the domain. Also of interest is the extension of the FD-SBP-SAT method by Reichert, Heath and Bodony [44] to overset grid methods.

The purpose of this paper is to construct a framework that explicitly lays out the necessary and the sufficient conditions for the existence of nodal SBP operators (i.e. the unknowns exist at nodes in physical space in contrast with modal methods where the expansion coefficients exist in the frequency domain). We extend the FD-SBP theory of Kreiss and Scherer [30] to include operators that 1) do not have a repeated interior point operator, 2) have a nonuniform nodal distribution in the computational domain, and 3) do not include one or both boundary nodes. By doing so it is possible to unify a wide array of operators as SBP operators, enabling the reinterpretation of various discretization methods under one cohesive framework and providing a natural and advantageous means of constructing numerical schemes for boundary conditions and inter-element/block coupling that lead to provably stable semi-discrete forms. For example, Gassner [11] has shown that a class of DG methods are SBP operators and the imposition of boundary conditions can be seen as SATs. Here we prove that a wider array of DG methods can be seen as SBP operators. Moreover, our framework enables FD operators without a repeated interior point operator to be interpreted as operators with subcell resolution similar to DG schemes. In contrast to classical FD schemes, this leads to the notion of elements with a prescribed internal node distribution such that $h$ refinement is carried out by increasing the number of elements rather than the number of nodes within an element.

The focus in this paper is on one-dimensional nodal first-derivative SBP operators that can be extended to multiple dimensions using Kronecker products. The framework developed in this paper follows the seminal papers on FD-SBP operators by Kreiss and Scherer [30] and the extension to block norms by Strand [45]. We generalize the ideas of Hicken and Zingg [23], who proved that the norms of classical FD-SBP operators are associated with Gregory type quadrature rules. We are thus not the first to note the deep link between SBP operators, their norms, and quadrature rules. This insight was implicit in using SATs and collocated-pseudo-spectral methods in [3, 20, 19, 18], since these methods have implicit quadrature rules associated with the nodal distributions. One of the main contributions of this paper is making this link explicit, proving that the norm must be associated with a quadrature for nodal SBP operators, and conversely that given a quadrature rule, an SBP operator can be constructed. A second important contribution consists of relaxing the need to include boundary nodes. This allows us to recast a wider array of known methods, such as collocated-pseudo-spectral and DG methods on Legendre-Gauss quadrature points, as SBP operators. The objective of the new framework is to provide a unification that facilitates an improved understanding of a range of methods as well as a generalization that enables the development of new operators.

## 2. Notation and definitions

We extend the concept of SBP operators to nodal distributions with variable node spacing that may or may not include the boundary nodes. The general nodal distribution for the domain $[\alpha, \beta]$ is given as $\mathbf{x} = [x_1, \dots, x_2]^T$, where the node locations $x_i$ are required only to obey the ordering property $\alpha \leq x_1 < x_2 < \cdots < x_n \leq \beta$. In other words, the nodal distribution may or may not include the boundary points; the only restrictions are that nodes do not overlap and a natural ordering with increasing $x$ coordinate.

Given that the nodal distribution is non-uniform, it is more natural to discuss the accuracy of the SBP operators in terms of the degree of the polynomial for which they are exact. Throughout the paper monomials are used in proving the degree of various operators. These are represented by $\mathbf{x}^i = \left[x_1^i, \ldots, x_n^i\right]^T$, with the convention that $\mathbf{x}^{-1} = 0$.

Capital letters with script type are used to denote continuous functions on a specified domain $x \in [\alpha, \beta]$. As an example, $\mathcal{U}(x) \in C^\infty[\alpha, \beta]$ denotes a function that is infinitely differentiable over the domain $x \in [\alpha, \beta]$. Lower case bold font is used to denote the restriction of such functions onto the nodes; for example the restriction of $\mathcal{U}$ onto the nodes $\mathbf{x}$ is given by:

$$\mathbf{u} = \left[\mathcal{U}(x_1), \ldots, \mathcal{U}(x_n)\right]^T.$$

Vectors with a subscript d, for example $\mathbf{u}_\mathrm{d} \in \mathbb{R}^{n\times 1}$, represent the solution to a system of discrete or semi-discrete equations.

## 3. Preliminaries

In this paper we extend the FD-SBP theory in three directions. First, classical FD-SBP operators are constructed around centred-difference approximations that are repeated on the interior of the operator, with the SBP property enforced through special treatment at nodes close to the boundaries. The theory is extended to accommodate operators that do not have this property. The resultant subset of SBP operators now have a fixed number of nodes; it therefore makes sense to interpret the operators as being cell based with the interior nodes providing subcell resolution. Secondly, FD-SBP operators have traditionally been defined on equi-spaced nodal distributions in the computational domain. Other SBP operators exist, collocated-pseudo-spectral and DG methods being prime examples. Our objective is to construct a framework that captures a broad class of SBP operators that includes FD operators as well as some collocated-pseudo-spectral and DG methods.

Hence we include nonuniform nodal distributions within cells. The final extension is accommodating operators that have nodal distributions that do not include nodes at the cell or element boundary. This extension allows, for example, framing DG operators based on Legendre-Gauss quadrature points as SBP operators. To make the presentation cleaner, the theory is initially developed with only the first two extensions. Subsequently it is proven that the theory applies to operators where one or both boundary nodes are not included.

One means of determining well-posed boundary and initial conditions is through the use of the energy method. In the energy method the PDE is multiplied by the solution and integrated in space. Then integration by parts is used to convert volume integrals to surface integrals, thereby allowing the introduction of boundary conditions. If the boundary conditions are homogeneous, this is usually sufficient to draw conclusions by showing that the time rate of change of the norm of the solution is zero or negative, thereby bounding the solution. For more general boundary conditions, a further integration in time is carried out. The challenge is to determine the restrictions on the boundary conditions such that an estimate on the solution, called an energy estimate, in terms of the data exists. For more information regarding the energy method see [28, 16, 15].

The key component of the energy method is integration by parts:

$$\int_{\alpha}^{\beta} \mathcal{V}\frac{\partial \mathcal{U}}{\partial x}\mathrm{d}x = \mathcal{U}\mathcal{V}|_{\alpha}^{\beta} - \int_{\alpha}^{\beta} \mathcal{U}\frac{\partial \mathcal{V}}{\partial x}\mathrm{d}x. \tag{1}$$

Our interest is in being able to apply the energy method to determine suitable boundary conditions that lead to schemes that are stable when applied to well-posed problems. To apply the energy method in the continuous case the following definitions of the inner product and norm are useful:

$$(\mathcal{U},\mathcal{V}) = \int_{\alpha}^{\beta} \mathcal{U}\mathcal{V}\mathrm{d}x, \quad \|\mathcal{U}\|^2 = \int_{\alpha}^{\beta} \mathcal{U}^2\mathrm{d}x. \tag{2}$$

With these definitions, (1) can be written as

$$\left(\mathcal{V},\frac{\partial \mathcal{U}}{\partial x}\right) = \mathcal{U}\mathcal{V}|_{\alpha}^{\beta} - \left(\mathcal{U},\frac{\partial \mathcal{V}}{\partial x}\right). \tag{3}$$

SBP operators are constructed to mimic (3) discretely. To construct a discrete analogue to (3), consider a nodal distribution defined by $\mathbf{x}^T = [x_1, \ldots, x_n]$. Suppose that the domain of interest is $x \in [\alpha, \beta]$, where admissible nodal distributions are limited to having the following ordering $\alpha = x_1 <, \ldots, < x_n = \beta$, where the assumption that $x_1 = \alpha$ and $x_n = \beta$ is temporary and will be removed in Section 5. A discrete analogue to (3) requires a first derivative operator, $D_1$, which is defined here by its degree, $q$, the maximum degree of the polynomial for which it is exact, i.e.

$$D_1 \mathbf{x}^j = j\mathbf{x}^{j-1}, \quad j \in [0, q]. \tag{4}$$

For vector spaces, a general inner product and norm have the form:

$$(\mathbf{u}, \mathbf{v})_H = \mathbf{u}^T H \mathbf{v}, \quad ||\mathbf{u}||_H^2 = \mathbf{u}^T H \mathbf{u}, \tag{5}$$

where $H$ must be symmetric and positive definite (PD). The discrete analogue of integration by parts, summation by parts, has the form:

$$\mathbf{v}^T H D_1 \mathbf{u} = \mathbf{v}^T \tilde{E} \mathbf{u} - \mathbf{u}^T H D_1 \mathbf{v}, \tag{6}$$

where, for now $\tilde{E} = \mathrm{diag}\left[-1, 0, \ldots, 0, 1\right]$[1]. Not all $D_1$ satisfy (6), and the conditions under which $D_1$ can satisfy (6) need to be determined. To do so, taking the transpose of (6), adding it to (6), and rearranging gives

$$\mathbf{v}^T \left[H D_1 + D_1^T H\right] \mathbf{u} + \mathbf{u}^T \left[D_1^T H + H D_1\right] \mathbf{v} = \mathbf{v}^T \tilde{E} \mathbf{u} + \mathbf{u}^T \tilde{E} \mathbf{v}. \tag{7}$$

Let $\Theta = H D_1$ (since $H$ is invertible, $D_1 = H^{-1}\Theta$); (7) becomes

$$\mathbf{v}^T \left[\Theta + \Theta^T\right] \mathbf{u} + \mathbf{u}^T \left[\Theta^T + \Theta\right] \mathbf{v} = \mathbf{v}^T \tilde{E} \mathbf{u} + \mathbf{u}^T \tilde{E} \mathbf{v}, \tag{8}$$

and it is concluded that

[1] In Section 5 nodal distributions that do not include one or both boundary nodes are considered, and $\tilde{E}$ has a more general form.

$$\Theta + \Theta^T = \tilde{E}. \tag{9}$$

To summarize, the following classical definition is given:

**Definition 1. Summation-by-parts operator:** *An operator is an approximation to the first derivative of degree $q$ with the SBP property if*

i) $D_1\mathbf{x}^j = H^{-1}\Theta\mathbf{x}^j = j\mathbf{x}^{j-1},\ j \in [0, q]$,

ii) *$H$ is a PD symmetric matrix,*

iii) $\Theta + \Theta^T = \tilde{E}$*, where for now* $\tilde{E} = diag\,[-1, 0, \ldots, 0, 1]$*, and a more general form is given in Section 5.*

## 4. A generalized theory for SBP operators

In this section necessary and sufficient conditions for the existence of operators satisfying Definition 1 are proven. Before proceeding let us clarify the intent: the approach will be to determine the conditions on the norm $H$, given that $\Theta + \Theta^T = \tilde{E}$, such that the resultant derivative operator $D_1 = H^{-1}\Theta$ is an SBP operator exact for polynomials of up to degree $q$. It is proven that a necessary condition is that the norm matrix of an SBP operator must be associated with a quadrature of at least degree $q-1$. Then, $H$ is temporarily restricted to be diagonal. Under this restriction it is proven that given a quadrature rule of degree $\tau$ with positive weights, an associated SBP operator of degree $q = \min\left(\lceil\frac{\tau}{2}\rceil, n-1\right)$ with diagonal norm having the quadrature weights along its diagonal is guaranteed to exist[2]. For dense-norm SBP operators, we first prove that they exist up to degree $n-1$ and then that they can be constructed from known quadrature rules, even if the weights are negative. In Section 5, it is proven that the theory presented in the current section can be applied to operators that do not include one or both boundary nodes.

[2] The ceiling operator $\lceil a \rceil$ gives the smallest integer greater than or equal to $a$.

Now that we have a road map, the first step is to derive the necessary conditions on $H$ to ensure that $D_1$ satisfies Definition 1. The accuracy requirements are:

$$D_1\mathbf{x}^j = H^{-1}\Theta\mathbf{x}^j = j\mathbf{x}^{j-1},\ j \in [0,q] \tag{10}$$

Multiplying (10) by $H$ we find

$$\Theta\mathbf{x}^j = jH\mathbf{x}^{j-1},\ j \in [0,q]. \tag{11}$$

Multiplying (11) by $\left(\mathbf{x}^i\right)^T$ gives

$$\left(\mathbf{x}^i\right)^T \Theta\mathbf{x}^j = j\left(\mathbf{x}^i\right)^T H\mathbf{x}^{j-1},\ i,j \in [0,q]. \tag{12}$$

Swapping indices in (12) gives

$$\left(\mathbf{x}^j\right)^T \Theta\mathbf{x}^i = i\left(\mathbf{x}^j\right)^T H\mathbf{x}^{i-1},\ i,j \in [0,q]. \tag{13}$$

Adding (12) and (13) results in

$$\left(\mathbf{x}^i\right)^T \Theta\mathbf{x}^j + \left(\mathbf{x}^j\right)^T \Theta\mathbf{x}^i = j\left(\mathbf{x}^i\right)^T H\mathbf{x}^{j-1} + i\left(\mathbf{x}^j\right)^T H\mathbf{x}^{i-1},\ i,j \in [0,q]. \tag{14}$$

However, all terms in (14) are scalars so we find

$$\left(\mathbf{x}^j\right)^T \Theta\mathbf{x}^i = \left(\left(\mathbf{x}^j\right)^T \Theta\mathbf{x}^i\right)^T = \left(\mathbf{x}^i\right)^T \Theta^T\mathbf{x}^j. \tag{15}$$

Substitution into (14) gives

$$\left(\mathbf{x}^i\right)^T \left[\Theta + \Theta^T\right]\mathbf{x}^j = j\left(\mathbf{x}^i\right)^T H\mathbf{x}^{j-1} + i\left(\mathbf{x}^j\right)^T H\mathbf{x}^{i-1},\ i,j \in [0,q]. \tag{16}$$

Using the condition that $\Theta + \Theta^T = \tilde{E} = \operatorname{diag}\left[-1,0,\ldots,0,1\right]$, this results in the necessary equations that $H$ must satisfy such that $D_1$ is an SBP operator of degree $q$; referred to here as the compatibility equations:

$$j\left(\mathbf{x}^{i}\right)^{T}H\mathbf{x}^{j-1}+i\left(\mathbf{x}^{j}\right)^{T}H\mathbf{x}^{i-1}=\left(\mathbf{x}^{i}\right)^{T}\tilde{E}\mathbf{x}^{j}=\beta^{i+j}-\alpha^{i+j}\quad i,j\in[0,q]. \tag{17}$$

The implication of (17) is that the degree of $D_1$, $q$, depends on the number of compatibility equations that are satisfied, which in turn depends on $H$.

The following Theorem can now be proven:

**Theorem 1.** *The norm matrix, $H$, of an SBP operator of degree $q$ that satisfies Definition 1 must be associated with a quadrature rule of at least degree $q-1$.*

*Proof.* Taking $i=0$ in (17) results in

$$j\left(\mathbb{1}\right)^{T}H\mathbf{x}^{j-1}=\left(\mathbb{1}\right)^{T}\tilde{E}\mathbf{x}^{j}\quad j\in[0,q]. \tag{18}$$

where $\mathbb{1}=[1,\ldots,1]^{T}$. Expanding and using the definition of $\tilde{E}$ gives

$$\textstyle\sum_{k=1}^{n}jx_{k}^{j-1}\sum_{p=1}^{n}H_{k,p}=\beta^{j}-\alpha^{j},\quad j\in[0,q], \tag{19}$$

taking $\tilde{H}_{k}=\sum_{p=1}^{n}H_{k,p}$, noting that the $j=0$ condition is automatically satisfied, and rearranging gives

$$\textstyle\sum_{k=1}^{n}\tilde{H}_{k}x_{k}^{j-1}=\frac{\beta^{j}-\alpha^{j}}{j},\quad j\in[1,q], \tag{20}$$

which are the conditions on a quadrature, $\int_{\alpha}^{\beta}f(x)\mathrm{d}x\approx\sum_{k=1}^{n}\tilde{H}_{k}f(x_{k})$, of at least degree $q-1$. □

### *4.1. Diagonal-norm SBP operators*

Theorem 1 states that a necessary condition for $D_1=H^{-1}\Theta$ to be an SBP operator is that $H$ be associated with a quadrature rule of at least degree $q-1$. Now the question is whether such operators exist and if so what degree can be achieved. We first examine the case where $H$ is diagonal. Expanding

(17) for a diagonal $H$

$$\textstyle\sum_{v=1}^{n} jx_v^i H_{vv} x_v^{j-1} + ix_v^j H_{vv} x_v^{i-1} \quad = \quad \left(\mathbf{x}^j\right)^T \tilde{E}\mathbf{x}^i,\ \ i,j \in [0,q]. \tag{21}$$

Using the definition of $\tilde{E}$ and simplifying

$$(i+j)\textstyle\sum_{v=1}^{n} H_{vv} x_v^{i+j-1} \quad = \quad \beta^{i+j} - \alpha^{i+j},\ \ i,j \in [0,q]. \tag{22}$$

Finally, noting that the condition $i = j = 0$ is automatically satisfied, we substitute $\sigma = i + j$ and rearrange to find

$$\textstyle\sum_{v=1}^{n} H_{vv} x_v^{\sigma-1} = \frac{\beta^\sigma - \alpha^\sigma}{\sigma}, \quad \sigma \in [1, 2q]. \tag{23}$$

The system of equations (23) are the accuracy equations for a quadrature of at least degree $\tau = 2q - 1$.

Furthermore, an upper bound on the degree of the first derivative exists with respect to the number of nodes. To determine the upper bound, consider the accuracy equations (10), which can be recast as

$$D_1 X = G \tag{24}$$

where $G = [\mathbf{0}, \mathbf{x}^0, \ldots, (n-1)\mathbf{x}^{n-2}]$. The matrix $X = [\mathbf{x}^0, \ldots, \mathbf{x}^{n-1}]$ is the Vandermonde matrix and is invertible, therefore $D_1$ has a unique solution:

$$D_1 = GX^{-1}. \tag{25}$$

The implication of (25) is that the system of equations is fully determined and therefore it is not possible to construct a $D_1$ of greater degree since the system would then become over determined, and therefore the upper bound on the degree of the first derivative is $n - 1$.

The following theorem is now proven:

**Theorem 2.** *A quadrature rule of degree $\tau$ with positive weights for a nodal distribution $\mathbf{x}$ is necessary and sufficient for the existence of a diagonal-norm SBP approximation to the first derivative, $D_1 = H^{-1}\Theta$, that is exact for polynomials of degree $q \le \min\left(\lceil \frac{\tau}{2} \rceil, n-1\right)$, where $n \ge 2$ is the size of $D_1$.*

*Proof.* A necessary condition on an SBP operator $D_1$ with a diagonal-norm H is that it satisfy the compatibility equations (22). By (23) this means that H must be associated with a quadrature rule of at least degree $2q-1$. Therefore, $q \leq \frac{\tau+1}{2}$, but since $q$ must be an integer, $q \leq \lceil \frac{\tau}{2} \rceil$. Such an H is readily constructed from the given quadrature rule with positive weights by putting the weights along its diagonal. Now we need to prove that there exist $\Theta$ matrices that lead to first derivative operators of degree $q$, that is they satisfy (11). First decompose $\Theta$ into its symmetric $\Theta_S$ and anti-symmetric $\Theta_A$ components:

$$\Theta = \Theta_S + \Theta_A. \tag{26}$$

We obtain $\Theta + \Theta^T = 2\Theta_S = \tilde{E}$; therefore,

$$\Theta = \frac{1}{2}\tilde{E} + \Theta_A. \tag{27}$$

Now (11) becomes,

$$\Theta_A \mathbf{x}^j = jH\mathbf{x}^{j-1} - \frac{1}{2}\tilde{E}\mathbf{x}^j = r_j, \ j \in [0, q]. \tag{28}$$

It is sufficient to prove that $\Theta_A$ exist for $q = n-1$, the maximum degree attainable for $D_1$. For $q = n-1$, (28) can be compactly written as

$$\Theta_A X = R, \tag{29}$$

where $X$ is the Vandermonde matrix, and $R = [r_0, r_1, \ldots, r_{n-1}]$. Therefore, $\Theta_A$ has solution

$$\Theta_A = RX^{-1}. \tag{30}$$

To complete the proof we must show that $RX^{-1}$ is anti-symmetric. To do this we use the properties of the compatibility equations. Consider the expansion

$$RX^{-1} = \left(X^{-1}\right)^T X^T R X^{-1}, \tag{31}$$

if $X^TR$ is anti-symmetric then so too is $\left(X^{-1}\right)^T X^TRX^{-1}$ and hence $RX^{-1}$; this follows since if the matrix $S$ is anti-symmetric so is $B^TSB$. Rearranging the compatibility equations (17) results in

$$j\left(\mathbf{x}^i\right)^T H\mathbf{x}^{j-1} - \frac{1}{2}\mathbf{x}^i\tilde{E}\mathbf{x}^j = -\left[i\left(\mathbf{x}^j\right)^T H\mathbf{x}^{i-1} - \frac{1}{2}\mathbf{x}^j\tilde{E}\mathbf{x}^i\right]\ i,j\in[0,q], \tag{32}$$

which can be recast in terms of the right-hand-side of (28) as

$$\left(\mathbf{x}^i, r_j\right) = -\left(\mathbf{x}^j, r_i\right), \tag{33}$$

where $(\mathbf{u},\mathbf{v}) = \mathbf{u}^T\mathbf{v}$ is the usual dot product. Expanding $X^TR$ results in

$$X^TR = \begin{bmatrix} (\mathbf{x}^0, r_0) & (\mathbf{x}^0, r_1) & \dots & (\mathbf{x}^0, r_{n-1}) \\ (\mathbf{x}^1, r_0) & (\mathbf{x}^1, r_1) & \dots & (\mathbf{x}^1, r_{n-1}) \\ \vdots & \vdots & \dots & \vdots \\ (\mathbf{x}^{n-1}, r_0) & (\mathbf{x}^{n-1}, r_1) & \dots & (\mathbf{x}^{n-1}, r_{n-1}) \end{bmatrix}. \tag{34}$$

Using (33) gives

$$X^TR = \begin{bmatrix} 0 & (\mathbf{x}^0, r_1) & \dots & (\mathbf{x}^0, r_{n-1}) \\ -(\mathbf{x}^0, r_1) & 0 & \dots & (\mathbf{x}^1, r_{n-1}) \\ \vdots & \vdots & \dots & \vdots \\ -(\mathbf{x}^0, r_{n-1}) & -(\mathbf{x}^1, r_{n-1}) & \dots & 0 \end{bmatrix}, \tag{35}$$

and we conclude that $\Theta_A = RX^{-1}$ is anti-symmetric, as required. □

As a simple example, consider the Newton-Cotes quadrature rule on four equally spaced nodes which has positive weights and is of degree 3. By Theorem 2 an SBP operators exists with maximal degree 2 and a PD norm:

$$H = h\begin{bmatrix} \frac{3}{8} & 0 & 0 & 0 \\ 0 & \frac{9}{8} & 0 & 0 \\ 0 & 0 & \frac{9}{8} & 0 \\ 0 & 0 & 0 & \frac{3}{8} \end{bmatrix}, \tag{36}$$

where $h$ is the spacing between nodes. Setting up the accuracy equations (11) and solving gives an SBP operator of degree 2:

$$D_1 = \tfrac{1}{h}\begin{bmatrix} -\frac{4}{3} & \frac{3}{2} & 0 & -\frac{1}{6} \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{6} & 0 & -\frac{3}{2} & \frac{4}{3} \end{bmatrix}, \tag{37}$$

where

$$\Theta = \begin{bmatrix} -\frac{1}{2} & \frac{9}{16} & 0 & -\frac{1}{16} \\ -\frac{9}{16} & 0 & \frac{9}{16} & 0 \\ 0 & -\frac{9}{16} & 0 & \frac{9}{16} \\ \frac{1}{16} & 0 & -\frac{9}{16} & \frac{1}{2} \end{bmatrix}. \tag{38}$$

Using the classical FD-SBP approach on four nodes it is only possible to obtain the second-order operator, which is

$$D_1 = \frac{1}{h}\begin{bmatrix} -1 & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & -1 & 1 \end{bmatrix}. \tag{39}$$

The quadrature rule associated with (39) is the composite trapezoidal rule, which is of degree one and leads to

$$H = h\begin{bmatrix} \frac{1}{2} & & & \\ & 1 & & \\ & & 1 & \\ & & & \frac{1}{2} \end{bmatrix}. \tag{40}$$

In contrast with (37), the operator (39), is of degree one. The implications of this will be further discussed in Section 8.

### *4.2. Dense-norm SBP operators*

For classical FD-SBP operators, using block-norm $H$ leads to increased degree relative to diagonal-norm operators. It is therefore natural to ask whether using dense $H$, constructed with the the same quadrature rule of degree $\tau$ as the diagonal-norm case treated in Theorem 2, the resultant $D_1$ can be of degree higher than the diagonal-norm operator. Moreover, the diagonal-norm case was restricted to quadrature rules with positive weights, otherwise $H$ is no longer a norm. In the dense-norm case we will show that that restriction can be lifted while still retaining a PD $H$. In this paper we take dense to mean any $H$ that is not strictly diagonal, with the exception of classical FD-SBP operators where we will continue to use the terminology of block-norm, so that it is clear that we are referring those operators.

For dense norms there is no reduction in the compatibility equations as for diagonal norms. Instead the full compatibility equations (17) must be dealt with. One can visualize the difference between associated diagonal and dense-norm operators by setting up a matrix, $P_{ij}$, where $\times$ is inserted for $i, j$ combinations of the compatibility equations that are satisfied. The diagonal-norm operator presented in the previous subsection based on a quadrature rule of degree 3 ($\tau = 3$) has the following $P$:

$$P = \begin{array}{c|ccccc} i/j & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & \times & \times & \times & \times & \times \\ 1 & \times & \times & \times & \times & \\ 2 & \times & \times & \times & & \\ 3 & \times & \times & & & \\ 4 & \times & & & & \end{array}.$$

The above form of the $P$ matrix arises from the reduction in independent compatibility equations for diagonal-norm operators. The reduction is such that all $i + j = g$ combinations lead to identical equations for a given $g$. In the present example, the compatibility equations associated with $i, j = [0, 2]$ must be satisfied for an SBP operator of degree 2, so the maximum value of $i + j$ is 4.

Again for $\tau = 3$, in order to achieve an SBP operator of degree 3 a dense norm must be used that satisfies a greater number of compatibility relations, with a $P$ matrix of the following form:

$$P = \begin{array}{c|ccccc} i/j & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & \times & \times & \times & \times & \times \\ 1 & \times & \times & \times & \times & \\ 2 & \times & \times & \times & \times & \\ 3 & \times & \times & \times & \times & \\ 4 & \times & & & & \end{array}.$$

The entries highlighted in white are the compatibility equations that are satisfied by restricting $H$ such that $\left(\mathbb{1}^T H\right)_i = (H_{\text{diag}})_{ii}$, where $H_{\text{diag}}$ is the diagonal norm (36). The entries highlighted in gray are the remaining compatibility equations that $H$ must satisfy such that the resultant operator is of degree 3.

We show below that such a dense norm can be found, and it is thus possible to satisfy a greater number of compatibility equations without increasing the degree of the associated quadrature rule.

In contrast to the diagonal-norm case we first start by proving that dense-norm operators exist, and then we show how to construct them from specific quadrature rules. We state and prove the following theorem:

**Theorem 3.** *Given a nodal distribution* $\mathbf{x}$*, there exists an SBP operator* $D_1 = H^{-1}\Theta$ *of degree* $q \leq n-1$*, with a dense-norm* $H$ *and an associated quadrature rule* $\mathbf{W} = [w_1, \ldots, w_n]$ *of degree* $\tau \geq q-1$ *such that* $\int_{x_1}^{x_n} \mathcal{F} \mathrm{d}x \approx \sum_{k=1}^{n} w_k f_k$*, for* $n \geq 2$.

*Proof.* We start with the case $q = n-1$ and so $\tau \geq n-2$ per Theorem 1. Using the definition of the discrete inner product, $(\mathbf{v}, \mathbf{u})_H = \mathbf{v}^T H \mathbf{u}$, the compatibility equations (17) can be recast as

$$j\left(\mathbf{x}^i, \mathbf{x}^{j-1}\right)_H + i\left(\mathbf{x}^j, \mathbf{x}^{i-1}\right)_H = \left(\mathbf{x}^i\right)^T \tilde{E}\mathbf{x}^j = \beta^{i+j} - \alpha^{i+j} \quad i, j \in [0, q]. \tag{41}$$

The compatibility equations are satisfied if the following equations, referred to as the norm equations, are satisfied:

$$\left(\mathbf{x}^j, \mathbf{x}^i\right)_H = \left(\mathbf{x}^j\right)^T H \mathbf{x}^i = \frac{\beta^{i+j+1} - \alpha^{i+j+1}}{i+j+1} = M_{ij}, \; i, j \in [0, q]. \tag{42}$$

To prove this, substitute (42) into (41) to obtain

$$j\frac{(\beta^{i+j}-\alpha^{i+j})}{i+j}+i\frac{(\beta^{i+j}-\alpha^{i+j})}{i+j}=\beta^{i+j}-\alpha^{i+j},\ i,j\in[0,q], \tag{43}$$

which is an identity; therefore $H$ that satisfy (42) satisfy the compatibility equations (41). The norm equations for $q=n-1$ are given as

$$X^T H X = M, \tag{44}$$

where $X=[\mathbf{x}^0,\dots,\mathbf{x}^{n-1}]$. Since $X$ is a Vandermonde matrix, it is invertible and therefore

$$H=(X^{-1})^T M X^{-1}. \tag{45}$$

Thus we have constructed an $H$ that satisfies the compatibility equations for an SBP operator of degree $q=n-1$ and have also satisfied sufficient norm equations such that the associated quadrature rule is of degree $n-1$. Shortly we will show that this is not necessary, in line with Theorem 1. By the above form, if $M$ is PD, then so too is $H$. To prove that $M$ is PD, we adapt the classical proof that a Hilbert matrix is PD; we must show that

$$\mathbf{v}^T M\mathbf{v} > 0. \tag{46}$$

Expanding the left-hand side of (46) (notice that for convenience we have shifted the indexes) gives

$$\sum_{p=1}^{n}\sum_{m=1}^{n} v_p v_m \frac{\beta^{p+m-1}-\alpha^{p+m-1}}{p+m-1}. \tag{47}$$

We also have that

$$\frac{\beta^{p+m-1}-\alpha^{p+m-1}}{p+m-1}=\int_\alpha^\beta y^{p+m-2}\mathrm{d}y. \tag{48}$$

Substituting (48) into (47) gives

$$\sum_{p=1}^{n}\sum_{m=1}^{n} v_p v_m \int_{\alpha}^{\beta} y^{p+m-2}\mathrm{d}y, \tag{49}$$

and thus

$$\int_{\alpha}^{\beta}\sum_{p=1}^{n}\sum_{m=1}^{n} v_p v_m y^{p+m-2}\mathrm{d}y. \tag{50}$$

This can be recast as

$$\int_{\alpha}^{\beta} \mathbf{v}^T\mathbf{y}\mathbf{y}^T\mathbf{v}\mathrm{d}y, \tag{51}$$

where $\mathbf{y}^T = [y^0, \dots, y^{n-1}]$. Making the substitution $p(y) = \mathbf{y}^T\mathbf{v}$ results in

$$\int_{\alpha}^{\beta} p^2(y)\mathrm{d}y. \tag{52}$$

However, the integral of a nonnegative function must be nonnegative; therefore:

$$\int_{\alpha}^{\beta} p^2(y)\mathrm{d}y \geq 0. \tag{53}$$

The equality in (53) implies that $p(y) = 0$, which cannot be the case, unless the monomials, $[y^0, \dots, y^{n-1}]$ are linearly dependent. However, on a finite interval, the monomials are linearly independent, and we conclude that $p^2(y) > 0$ if $\mathbf{v} \neq 0$ and finally that $M$ is PD.

To summarize, we have proven that there exists a PD $H$ that satisfies the compatibility equations for $i, j \in [0, n-1]$. Moreover, by the norm equations (42), $H$ is associated with a quadrature rule $w_i = \left(\mathbf{1}^T H\right)_i$ of degree $q = n-1$. This is one degree greater than required by Theorem 1. In setting up $M$ we are effectively solving more norm equations than necessary to solve the required compatibility equations, namely the last column and row of $M$.

We would like to show that instead of the above, it is possible to construct $H$

associated with a quadrature rule of degree $\tau = q - 1$. To do so we examine the case of $q = n - 1$ but set $\tau = n - 2$, that is we only satisfy the norm equations associated with $i, j \in [0, q - 1]$. This same procedure can then be used to prove the case for $q < n - 1$. Consider the following system

$$X^T H X = \begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^T & c \end{bmatrix}, \tag{54}$$

where $M$ is an $(n-1) \times (n-1)$ matrix constructed from (42) for $i, j \in [0, n-2]$. We now prove that it is always possible to choose $\mathbf{b}$ and $c$ such that the RHS of (54) is PD, if $M > 0$. The trivial case is $\mathbf{b} = \mathbf{0}$ and $c > 0$. The PD condition is

$$[\tilde{\mathbf{v}}^T, v_n] \begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^T & c \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}} \\ \\ v_n \end{bmatrix} > 0, \tag{55}$$

where $\tilde{\mathbf{v}}^T = [v_1, \dots, v_{n-1}]$. Expanding gives,

$$\tilde{\mathbf{v}}^T M \tilde{\mathbf{v}} + 2 v_n \mathbf{b}^T \tilde{\mathbf{v}} + c v_n^2 > 0. \tag{56}$$

The matrix $M$ is PD and has decomposition $M = L^T \Lambda L$, where $L$ is unitriangular and therefore invertible and $\Lambda > 0$ is a diagonal matrix with the eigenvalues of $M$. With these definitions we get

$$\tilde{\mathbf{v}}^T L^T \Lambda L \tilde{\mathbf{v}} + 2 v_n \mathbf{b}^T \tilde{\mathbf{v}} + c v_n^2 > 0. \tag{57}$$

With $\hat{\mathbf{v}} = L \tilde{\mathbf{v}}$, which gives $\tilde{\mathbf{v}} = L^{-1} \hat{\mathbf{v}}$, we get

$$\hat{\mathbf{v}}^T \Lambda \hat{\mathbf{v}} + 2 v_n \mathbf{b}^T L^{-1} \hat{\mathbf{v}} + c v_n^2 > 0. \tag{58}$$

Taking $\hat{\mathbf{b}} = \mathbf{b}^T L^{-1}$ and expanding (58) gives

$$\sum_{i=1}^{n} \lambda_i \hat{v}_i^2 + 2 \tilde{v} \hat{\mathbf{b}}_i \hat{v}_i + c v_n^2 > 0. \tag{59}$$

Now we complete the square using $\lambda_i\left(\hat{v}_i+\Gamma_i v_n\right)^2$, with $\Gamma_i=\frac{\hat{b}_i}{\lambda_i}$, to obtain

$$\sum_{i=1}^{n}\lambda_i\left(\hat{v}_i+\Gamma_i v_n\right)^2+\left(c-\sum_{i=1}^{n}\lambda_i\Gamma_i^2\right)v_n^2>0. \tag{60}$$

Therefore, a sufficient condition is that $b_i$ and $c$ satisfy $\left(c-\sum_{i=1}^{n}\frac{\hat{b}_i^2}{\lambda_i}\right)\geq 0$, or more specifically that $c\geq\sum_{i=1}^{n}\frac{\hat{b}_i^2}{\lambda_i}$. This process can be applied one row and column at a time to construct a PD $H$ for $q<n-1$, where the restriction $n\geq 2$ comes from requiring the SBP operator be of at least degree 1.

By the arguments in the proof of Theorem 2 there exist $\Theta$ such that an operator of degree $q$ exists, and again the associated quadrature is constructed by $w_i=\left(\mathbb{1}^T H\right)_i$ and is of degree $\geq q-1$. □

Now we are in a position to prove the following:

**Theorem 4.** *A quadrature rule* $\mathbf{W}=[w_1,\ldots,w_n]^T$ *of degree* $\tau$ *on a nodal distribution* $\mathbf{x}$*, such that* $\int_{x_1}^{x_n}\mathcal{F}\mathrm{d}x\approx\sum_{k=1}^{n}w_k f_k$*, is necessary and sufficient for the existence of a dense PD norm* $H$ *that satisfies* $\mathbb{1}^T H\mathbf{f}=\mathbf{W}^T\mathbf{f}$ *and an associated SBP operator,* $D_1=H^{-1}\Theta$ *of degree* $q=\min\left(\tau+1,n-1\right)$*.*

*Proof.* We have already proven that each dense-norm SBP operator has a norm $H$ associated with some quadrature rule and so a quadrature rule is a necessary condition; now we prove that a quadrature rule is a sufficient condition for the existence of a dense-norm SBP operator. We consider the case where $q=n-1$ and $\tau=q-1=n-2$. By the arguments in the previous theorem we have

$$X^T H X=\begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^T & c \end{bmatrix}, \tag{61}$$

where $\mathbf{b}$ and $c$ are to be determined. We need to show that we can choose $\mathbf{b}$ and $c$ such that $H$ is associated with the quadrature rule $\mathbf{W}$ and is PD. The first requirement means that

$$H\mathbb{1}=\mathbf{W}. \tag{62}$$

Solving for $H$ in (61) and inserting into (62) results in

$$H\mathbb{1} = \left(X^{-1}\right)^{T} \begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^{T} & c \end{bmatrix} X^{-1}\mathbb{1} = \mathbf{W}. \tag{63}$$

Pre multiplying by $X^{T}$ gives

$$\begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^{T} & c \end{bmatrix} X^{-1}\mathbb{1} = X^{T}\mathbf{W}. \tag{64}$$

Now $X^{-1}\mathbb{1} = \mathbf{e}_0$ and by definition

$$X^{T}\mathbf{W} = \begin{bmatrix} \frac{\beta^{1}-\alpha^{1}}{1} \\ \\ \vdots \\ \\ \frac{\beta^{n-1}-\alpha^{n-1}}{n-1} \\ \\ \left(\mathbf{x}^{n-1}\right)^{T}\mathbf{W} \end{bmatrix}. \tag{65}$$

On the other hand we have that

$$\begin{bmatrix} M & \mathbf{b} \\ \\ \mathbf{b}^{T} & c \end{bmatrix} \mathbf{e}_0 = \begin{bmatrix} M_{0,0} \\ \\ \vdots \\ \\ M_{n-2,0} \\ b_1 \end{bmatrix} = \begin{bmatrix} \frac{\beta^{1}-\alpha^{1}}{1} \\ \\ \vdots \\ \\ \frac{\beta^{n-1}-\alpha^{n-1}}{n-1} \\ b_1 \end{bmatrix} \tag{66}$$

and we conclude that for $H$ to be associated with the quadrature rule $\mathbf{W}$ we must have $b_1 = \left(\mathbf{x}^{n-1}\right)^{T}\mathbf{W}$. The remaining free parameters in $\mathbf{b}$ and $c$ are chosen so that the resultant matrix is PD; for example, a sufficient condition is that $c \geq \sum_{i=1}^{n} \frac{\hat{b}_i^2}{\lambda_i}$, where the various quantities are defined in the proof of Theorem 3. □

For dense-norm SBP operators, stability cannot be proven on curvilinear grids [46]. However, recently, Mattsson [34] introduced a boundary stabilization operator that has been numerically shown to lead to stable discretizations. Theorem 4 provides a means of increasing the degree of the resultant

SBP operator constructed from a given quadrature rule beyond that achievable with a diagonal $H$ by using the degrees of freedom from a dense $H$ so as to satisfy more of the compatibility equations. For example, consider the diagonal-norm case given in (36) and (37), where $\tau = 3$ and $q = 2$, using the Newton-Cotes quadrature rule on four equally spaced nodes. The diagonal norm can be converted to a dense norm (see Section 7 for more details) and the accuracy equations solved, giving an operator with degree 3, consistent with Theorem 4 . The following operator is obtained:

$$H = \begin{bmatrix} \frac{1}{4} & \frac{1}{8} & 0 & 0 \\ \frac{1}{8} & \frac{5}{4} & -\frac{1}{4} & 0 \\ 0 & -\frac{1}{4} & \frac{5}{4} & \frac{1}{8} \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} \end{bmatrix}, \tag{67}$$

$$D_1 = \begin{bmatrix} -\frac{11}{6} & 3 & -\frac{3}{2} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{2} & 1 & -\frac{1}{6} \\ \frac{1}{6} & -1 & \frac{1}{2} & \frac{1}{3} \\ -\frac{1}{3} & \frac{3}{2} & -3 & \frac{11}{6} \end{bmatrix}, \tag{68}$$

with

$$\Theta = \begin{bmatrix} -\frac{1}{2} & \frac{11}{16} & -\frac{1}{4} & \frac{1}{16} \\ -\frac{11}{16} & 0 & \frac{15}{16} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{15}{16} & 0 & \frac{11}{16} \\ -\frac{1}{16} & \frac{1}{4} & -\frac{11}{16} & \frac{1}{2} \end{bmatrix}. \tag{69}$$

## 5. Generalization to nodal distributions that do not include one or both boundary nodes

In this section the theory is proven to hold for nodal distributions that do not include one or both boundary nodes of the domain. In the classical description [30], the nodal distribution contains the endpoints of the domain and $\tilde{E} = \text{diag}[-1, 0, \dots, 0, 1]$. Thus

$$\mathbf{v}^T \tilde{E} \mathbf{u} = \mathcal{V}(x_n)\mathcal{U}(x_n) - \mathcal{V}(x_1)\mathcal{U}(x_1). \tag{70}$$

However, if one or both boundary nodes are not included in the nodal distribution, then it is not possible to satisfy (70). To extend the SBP concept requires a generalization of (70). Consider two functions $\mathcal{U}(x)$ and $\mathcal{V}(x)$ on the domain $x \in [\alpha, \beta]$, and a nodal distribution $\mathbf{x} = [x_1, \ldots, x_n]^T$ that has the following ordering property $\alpha < x_1 < \cdots < x_n < \beta$. The restriction of the two functions onto the nodes is given by $\mathbf{u} = [\mathcal{U}(x_1), \ldots, \mathcal{U}(x_n)]^T$ and $\mathbf{v} = [\mathcal{V}(x_1), \ldots, \mathcal{V}(x_n)]^T$. Instead of requiring (70), the SBP property is extended by requiring

$$\mathbf{v}^T \tilde{E} \mathbf{u} \approx \mathcal{V}\mathcal{U}|_{\alpha}^{\beta}, \tag{71}$$

which is quantified by

$$\left(\mathbf{x}^i\right)^T \tilde{E} \mathbf{x}^j = \left(\beta^{i+j} - \alpha^{i+j}\right), \quad i, j \in [0, r]. \tag{72}$$

We restrict the theory to SBP operators for which $r \geq q$. SBP operators that have $r < q$ can be constructed; however, as will be demonstrated in Section 6 on imposition of boundary and interface conditions, the degree of the SAT term is $r$. If the boundary conditions are enforced with terms of degree $r$, and $r < q$, then the imposition of the boundary conditions represents the largest error in the discretization. In fact, we face a further problem; for operators constructed such that $r < q$ we are unsure how to construct SATs that lead to consistent, conservative and stable discretizations.

In order to prove that Theorems 1 through 4 hold for nodal distributions that exclude boundary nodes, under the restriction that $r \geq q$, one must prove that $\tilde{E}$ matrices exist that satisfy (72). For $n$ distinct nodes it is possible to construct a one-dimensional interpolant of degree $n-1$. Evaluating the interpolant of $\mathcal{U}$ at the boundaries yields

$$t_{\alpha}^T \mathbf{u} = \tilde{u}_{\alpha} \approx \mathcal{U}(\alpha), \quad t_{\beta}^T \mathbf{u} = \tilde{u}_{\beta} \approx \mathcal{U}(\alpha), \tag{73}$$

where $t_{\alpha}$ and $t_{\beta}$ have properties,

$$t_{\alpha}^T \mathbf{x}^j = \alpha^j, \quad t_{\beta}^T \mathbf{x}^j = \beta^j, \quad j \in [0, n-1]. \tag{74}$$

which can be combined to form the matrix operator

$$T = e_1 t_{\alpha}^T + e_n t_{\beta}^T \tag{75}$$

where $e_1 = [1, 0, \dots, 0]^T$ and $e_n = [0, \dots, 0, 1]^T$. Taking $\tilde{E} = T^T E T$, where $E = \operatorname{diag}[-1, 0, \dots, 0, 1]$, gives the required property. The case where only one boundary node is excluded follows identical logic. Now we state an extended definition of an SBP operator:

**Definition 2. Summation-by-parts operator:** *An operator $D_1$ is an approximation to the first derivative of degree $q$ with the SBP property if*

i) $D_1 \mathbf{x}^j = H^{-1}\Theta\mathbf{x}^j = j\mathbf{x}^{j-1},\ j \in [0, q]$,

ii) *$H$ is a PD symmetric matrix,*

iii) $\Theta + \Theta^T = \tilde{E}$, *where* $\left(\mathbf{x}^i\right)^T \tilde{E}\mathbf{x}^j = \beta^{i+j} - \alpha^{i+j}$, $i, j \in [0, r]$, $r \geq q$.

This definition includes the case where $\tilde{E} = \operatorname{diag}[-1, 0, \dots, 0, 1]$ for $r = \infty$.

## 6. Time stability of generalized SBP/SAT discretizations

SBP operators do not include boundary conditions, and in fact are singular, mimetic of the continuous case where in the absence of boundary conditions systems of PDEs are ill-posed. Here SATs are used to impose boundary conditions weakly. In this section SATs are derived for the one-dimensional convection equation for the generalized definition of SBP operators. The PDE is

$$\tfrac{\partial \mathcal{U}}{\partial t} = -a\tfrac{\partial \mathcal{U}}{\partial x},\quad x \in [\alpha, \beta],\ t > 0 \tag{76}$$

with real $a > 0$, initial condition $\mathcal{U}(x, 0) = f(x)$, and homogeneous boundary condition $\mathcal{U}(\alpha, t) = g_\alpha(t) = 0$. To gain insight in to what needs to be done to construct SATs for the semi-discrete case it is instructive to first analyze the continuous case. The objective is to show that the chosen boundary conditions lead to a well-posed problem, specifically, that the solution is bounded, where it is assumed that the solution exists and is unique. To do so the energy method is employed, see [15]: multiply (76) by $\mathcal{U}$ and integrate in space,

$$\int_\alpha^\beta \mathcal{U}\frac{\partial \mathcal{U}}{\partial t}\mathrm{d}x = -a\int_\alpha^\beta \mathcal{U}\frac{\partial \mathcal{U}}{\partial x}\mathrm{d}x, \tag{77}$$

using the fact that $\mathcal{U}\frac{\partial \mathcal{U}}{\partial \xi} = \frac{1}{2}\frac{\partial \mathcal{U}^2}{\partial \xi}$ with Leibniz's rule on the LHS gives

$$\frac{\mathrm{d}||\mathcal{U}(\cdot,t)||^2}{\mathrm{d}t} = -a\left(\mathcal{U}^2(\beta,t) - \mathcal{U}^2(\alpha,t)\right). \tag{78}$$

Inserting the boundary conditions we find

$$\frac{\mathrm{d}||\mathcal{U}(\cdot,t)||^2}{\mathrm{d}t} = -a\mathcal{U}^2(\beta,t). \tag{79}$$

Integrating in time, inserting the initial condition, and rearranging gives

$$||\mathcal{U}(\cdot,t)||_x^2 + a\|\mathcal{U}(\beta,\cdot)\|_t^2 = \|f(\cdot)\|_x^2, \tag{80}$$

where $\|\mathcal{U}(\cdot,t)\|_x^2 = \int_\alpha^\beta \mathcal{U}^2(x,t)\mathrm{d}x$ and $\|\mathcal{U}(x,\cdot)\|_t^2 = \int_0^t \mathcal{U}^2(x,\tau)\mathrm{d}\tau$. Equation (80) demonstrates that the solution is bounded by the data, so the continuous problem is well-posed.

We have shown in (75) that a decomposition in the form $\tilde{E} = T^T E T$ always exists; there may be other possibilities, but here we restrict our interest to such a decomposition. This allows us to construct consistent, conservative, stable schemes with SATs. Application of a spatial discretization to (76) using SBP operators with boundary conditions enforced with SATs gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{d} = -aD_1\mathbf{u}_\mathrm{d} + \sigma_\alpha H^{-1} t_\alpha (t_\alpha^T \mathbf{u}_\mathrm{d} - g_\alpha(t)), \tag{81}$$

We need to prove that the semi-discrete equations are consistent and stable. Definition 2 ensures that in the absence of the SATs (81) is consistent. What remains is to prove that the SATs are consistent. Rearrange (81) to obtain

$$\sigma_\alpha H^{-1} t_\alpha (t_\alpha^T \mathbf{u}_\mathrm{d} - g_\alpha(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{d} + aD_1\mathbf{u}_\mathrm{d}. \tag{82}$$

Multiplying both sides by $\frac{1}{\sigma_\alpha}H$ gives

$$t_\alpha (t_\alpha^t \mathbf{u}_\mathrm{d} - g_\alpha(t)) = \frac{1}{\sigma_\alpha} H \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}_\mathrm{d} + aD_1\mathbf{u}_\mathrm{d}\right). \tag{83}$$

Taking the limit $h \to 0$, the RHS of (83) goes to zero because $H$ is an order one function of $h$, resulting in,

$$\lim_{h\to 0} t_\alpha (t_\alpha^T \mathbf{u}_\mathrm{d} - g_\alpha(t)) = 0. \tag{84}$$

Since $t_\alpha \neq 0$, we obtain

$$\lim_{h\to 0} (t_\alpha^T \mathbf{u}_\mathrm{d} - g_\alpha(t)) = 0. \tag{85}$$

By definition $\lim_{h\to 0} t_\alpha^T \mathbf{u} = \mathcal{U}(\alpha, t)$ resulting in

$$\mathcal{U}(\alpha, t) - g_\alpha(t) = 0, \tag{86}$$

which is the boundary condition in the continuous case.

Now the energy method is used to determine the conditions for which the proposed SATs lead to stable a scheme. Multiply (81) by $\mathbf{u}_\mathrm{d}^T H$ to obtain

$$\mathbf{u}_\mathrm{d}^T H \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{u}_\mathrm{d} = -a \mathbf{u}_\mathrm{d}^T \Theta \mathbf{u}_\mathrm{d} + \sigma_\alpha \mathbf{u}_\mathrm{d}^T t_\alpha (t_\alpha^T \mathbf{u}_\mathrm{d} - g_\alpha(t)). \tag{87}$$

Adding (87) to its transpose gives,

$$\begin{aligned} \mathbf{u}_\mathrm{d}^T H \frac{\mathrm{d}\mathbf{u}_\mathrm{d}}{\mathrm{d}t} + \mathbf{u}_\mathrm{d}^T H \frac{\mathrm{d}\mathbf{u}_\mathrm{d}}{\mathrm{d}t} = & -a \mathbf{u}_\mathrm{d}^T \left( \Theta + \Theta^T \right) \mathbf{u}_\mathrm{d} \\ & + 2\sigma_\alpha \mathbf{u}_\mathrm{d}^T t_\alpha t_\alpha^T \mathbf{u}_\mathrm{d} - 2\sigma_\alpha \mathbf{u}_\mathrm{d}^T t_\alpha g_\alpha(t). \end{aligned} \tag{88}$$

With $\mathbf{u}_\mathrm{d}^T \left( \Theta + \Theta^T \right) \mathbf{u}_\mathrm{d} = \tilde{u}_\beta^2 - \tilde{u}_\alpha^2$ and $t_\alpha^T \mathbf{u}_\mathrm{d} = \tilde{u}_\alpha$, we find

$$\frac{\mathrm{d}||\mathbf{u}_\mathrm{d}||_H}{\mathrm{d}t} = -a(\tilde{u}_\beta^2 - \tilde{u}_\alpha^2) + 2\sigma_\alpha \tilde{u}_\alpha^2 - 2\sigma_\alpha \tilde{u}_a g_\alpha(t). \tag{89}$$

Rearranging, and considering the homogenous case $g_\alpha(t) = 0$, we obtain

$$\frac{\mathrm{d}||\mathbf{u}_\mathrm{d}||_H}{\mathrm{d}t} + a\tilde{u}_\beta^2 = (a + 2\sigma_\alpha)\tilde{u}_\alpha^2. \tag{90}$$

For a stable scheme we require $(a + 2\sigma_\alpha) \leq 0$ or $\sigma_\alpha \leq -\frac{a}{2}$. To retain the same energy estimate as the continuous case we set $\sigma_\alpha = -\frac{a}{2}$. Integrating in time, inserting the initial condition and rearranging gives

$$\|\mathbf{u}_{\mathrm{d}}\|_H + a\|\tilde{u}_\beta\|_t^2 = \|f\|_H^2. \tag{91}$$

Thus an energy estimate exists, and the scheme is stable with the proposed SATs.

Next we consider the interface between two elements or blocks. The other boundaries are neglected assuming without loss of generality that suitable SATs have been specified. The goal is to determine the SAT parameters such that the resultant scheme is stable and conservative. The solution in the left and right domain are denoted with subscript $L$ and $R$ giving

$$\frac{\mathrm{d}\mathbf{u}_{\mathrm{d,L}}}{\mathrm{d}t} = -aD\mathbf{u}_{\mathrm{d},L} + \sigma_L t_L \left(t_L^T \mathbf{u}_{\mathrm{d,L}} - t_R^T \mathbf{u}_{\mathrm{d,R}}\right), \tag{92}$$

$$\frac{\mathrm{d}\mathbf{u}_{\mathrm{d,R}}}{\mathrm{d}t} = -aD\mathbf{u}_{\mathrm{d,R}} + \sigma_R t_R \left(t_R^T \mathbf{u}_{\mathrm{d,R}} - t_L^T \mathbf{u}_{\mathrm{d,L}}\right). \tag{93}$$

where

$$t_L^T \mathbf{u}_{\mathrm{d,L}} = \tilde{u}_{L,\delta}, \quad t_R^T \mathbf{u}_{\mathrm{d,R}} = \tilde{u}_{R,\delta}, \tag{94}$$

and $x = \delta$ is the location of the interface between the two elements. Premultiply (92) by $\mathbb{1}^T H_L$, to obtain

$$\mathbb{1}^T H_L \frac{d\mathbf{u}_{\mathrm{d,L}}}{dt} = -a\mathbb{1}^T \Theta_L \mathbf{u}_{\mathrm{d,L}} + \sigma_L \mathbb{1}^T t_L \left(t_L^T \mathbf{u}_{\mathrm{d,L}} - t_R^T \mathbf{u}_{\mathrm{d,R}}\right). \tag{95}$$

The extended SBP property is $\Theta + \Theta^T = \tilde{E}$; thus $\Theta = \tilde{E} - \Theta^T$. Substituting gives

$$\mathbb{1}^T H_L \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,L}}}{\mathrm{d}t} = -a\mathbb{1}^T \left(\tilde{E}_L - \Theta_L^T\right) \mathbf{u}_{\mathrm{d,L}} + \sigma_L \mathbb{1}^T t_L \left(t_L^T \mathbf{u}_L - t_R^T \mathbf{u}_{\mathrm{d,R}}\right). \tag{96}$$

Furthermore $\mathbb{1}$ is in the null space of $\Theta$; thus

$$\mathbb{1}^T H_L \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,L}}}{\mathrm{d}t} = -a\mathbb{1}^T \left(\tilde{E}_L\right) \mathbf{u}_{\mathrm{d,L}} + \sigma_L \mathbb{1}^T t_L \left(t_L^T \mathbf{u}_{\mathrm{d,L}} - t_R \mathbf{u}_{\mathrm{d,R}}\right) \tag{97}$$

Finally, because we are ignoring the left boundary of the left element and the right boundary of the right element, and $t_L^T \mathbb{1} = 1$, we get for the left element

$$\mathbb{1}^T H_L \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,R}}}{\mathrm{d}t} = -a\tilde{u}_{L,\delta} + \sigma_L \left(\tilde{u}_{L,\delta} - \tilde{u}_{R,\delta}\right), \tag{98}$$

and similarly for the right element

$$\mathbb{1}^T H_R \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,L}}}{\mathrm{d}t} = a\tilde{u}_{R,\delta} + \sigma_R \left(\tilde{u}_{R,\delta} - \tilde{u}_{L,\delta}\right). \tag{99}$$

Adding (98) to (99) gives

$$\frac{\mathrm{d}\left[\mathbb{1}^T H_L \mathbf{u}_{\mathrm{d,L}} + \mathbb{1}^T H_R \mathbf{u}_{\mathrm{d,R}}\right]}{\mathrm{d}t} = \tilde{u}_{L,\delta}\left[-a + \sigma_L - \sigma_R\right] + \tilde{u}_{R,\delta}\left[\sigma_R - \sigma_L + a\right]. \tag{100}$$

For conservation the RHS of (100) must be zero, and we conclude that we must have $\sigma_R = \sigma_L - a$, again an identical result to the classical SBP derivation [13].

Now we consider the stability of the semi-discrete form. Left multiplying (92) by $U_L^T H_L$ and (93) by $U_R^T H_R$ gives

$$\mathbf{u}_{\mathrm{d,L}}^T H_L \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,L}}}{\mathrm{d}t} = -a\mathbf{u}_{\mathrm{d,L}}^T \Theta_L \mathbf{u}_{\mathrm{d,L}} + \sigma_L \mathbf{u}_{\mathrm{d,L}}^T t_L \left(t_L^T \mathbf{u}_{\mathrm{d,L}} - t_R^T \mathbf{u}_{\mathrm{d,R}}\right) \tag{101}$$

$$\mathbf{u}_{\mathrm{d,R}}^T H_R \frac{\mathrm{d}\mathbf{u}_{\mathrm{d,R}}}{\mathrm{d}t} = -a\mathbf{u}_{\mathrm{d,R}}^T \Theta_R \mathbf{u}_{\mathrm{d,R}} + \sigma_R \mathbf{u}_{\mathrm{d,R}}^T t_R \left(t_R^T \mathbf{u}_{\mathrm{d,R}} - t_L^T \mathbf{u}_{\mathrm{d,L}}\right). \tag{102}$$

Adding the transpose, using the SBP property, simplifying, and ignoring the boundary conditions, we find

$$\frac{\mathrm{d}\left[\mathbf{u}_{\mathrm{d,L}}^{T} H_L \mathbf{u}_{\mathrm{d,L}}\right]}{\mathrm{d}t} = -a\tilde{u}_{L,\delta}^2 + 2\sigma_L\left(\tilde{u}_{L,\delta}^2 - \tilde{u}_{L,\delta}\tilde{u}_{R,\delta}\right), \tag{103}$$

$$\frac{\mathrm{d}\left[\mathbf{u}_{\mathrm{d,R}}^{T} H_R \mathbf{u}_{\mathrm{d,R}}\right]}{\mathrm{d}t} = a\tilde{u}_{R,\delta}^2 + 2\sigma_R\left(\tilde{u}_{R,\delta}^2 - \tilde{u}_{R,\delta}\tilde{u}_{L,\delta}\right). \tag{104}$$

Adding (103) to (104) and using the condition on the penalty parameters so that the discretization is conservative, $\sigma_R = \sigma_L - a$, gives

$$\frac{\mathrm{d}\left[2\mathbf{u}_{\mathrm{d,L}}^{T} H_L \mathbf{u}_{\mathrm{d,L}} + \mathbf{u}_{\mathrm{d,R}}^{T} H_R \mathbf{u}_{\mathrm{d,R}}\right]}{\mathrm{d}t} = \left(2\sigma_L - a\right)\left(\tilde{u}_{L,\delta} - \tilde{u}_{R,\delta}\right)^2. \tag{105}$$

To have stability, the RHS of (105) must be $\leq 0$ thus, $\sigma_L \leq \frac{a}{2}$, as in the classical SBP case [13].

## 7. Derivation of SBP operators

We begin this section with a brief review of what we have shown thus far. Theorem 1 states that the norm matrix of an SBP operator of degree $q$ must correspond to a quadrature rule of at least degree $q-1$. Such a quadrature is necessary but not sufficient. Theorem 2 states that given a quadrature rule of degree $\tau$ with positive weights for a nodal distribution $\mathbf{x}$, we can find a diagonal-norm SBP operator for the first derivative of degree $q = \min\left(\lceil\frac{\tau}{2}\rceil, n-1\right)$. Theorem 3 proves the existence of dense-norm SBP operators up to degree $n-1$. Theorem 4 proves that a quadrature rule of degree $\tau$ is necessary and sufficient for the existence of a dense-norm SBP operator of degree $\min\left(\tau+1, n-1\right)$, relaxing the requirement for positive quadrature weights. In contrast to the classical FD-SBP approach, the operator need not have a repeating interior point operator nor a uniform nodal distribution, though the nodal distribution is assumed to include the boundary nodes. The generalization of Theorems 2 and 4 to nodal distributions that do not include the boundary nodes is presented in Section 5. The significance of these Theorems is that the existence of a quadrature is necessary and sufficient for the existence of an SBP operator, and that the degree of the resulting operator is linked to the degree of the underlying quadrature. Furthermore, this extends the theory of the classical FD-SBP approach to a much broader class of first-derivative operators.

Next we consider some examples of how the theory in this article can be applied in the derivation of SBP operators. Classical FD-SBP operators are briefly discussed, followed by collocated-pseudo-spectral operators, which are used in some Discontinuous-Galerkin approaches, and ending with the derivation of novel SBP schemes based on Newton-Cotes quadrature, Chebyshev-polynomial-based quadratures, and barycentric rational interpolation.

### *7.1. Classical finite-difference SBP operators*

Classical FD-SBP operators were originally constructed without any knowledge of an underlying quadrature. The form of the individual operators is set *a priori* to have a uniform nodal distribution in the computational domain including the boundary nodes and a repeating interior point operator. Boundary closures are derived in order to satisfy Definition 1. The interior point operators are centered-difference formulae of degree $2p$ and the boundary closures are formed by a minimum of $2p$ biased-difference formulae [45]. As an example, consider the case where $p = 2$. The block norm has form

$$H = h\begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} & & \\ H_{12} & H_{22} & H_{23} & H_{24} & & \\ H_{13} & H_{23} & H_{33} & H_{34} & & \\ H_{14} & H_{24} & H_{34} & H_{44} & & \\ & & & & 1 & \\ & & & & & \ddots \end{bmatrix}, \tag{106}$$

while $\Theta$ has the form

$$\Theta = \begin{bmatrix} -\frac{1}{2} & \theta_{12} & \theta_{13} & \theta_{14} & & & & \\ -\theta_{12} & 0 & \theta_{23} & \theta_{24} & & & & \\ -\theta_{13} & -\theta_{23} & 0 & \theta_{34} & -\frac{1}{12} & & & \\ -\theta_{14} & -\theta_{24} & -\theta_{34} & 0 & \frac{2}{3} & -\frac{1}{12} & & \\ & & \frac{1}{12} & -\frac{2}{3} & 0 & \frac{2}{3} & -\frac{1}{12} & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{107}$$

Note that these operators are constructed to be invariant under the transformation from $x \rightarrow -x$ implying that $H$ is bisymmetric and $\Theta$ is nearly anti-bisymmetric, such that $\Theta + \Theta^T = E = \text{diag}[-1, 0, \ldots, 0, 1]$. Bisymmetric matrices have the following two forms:

$$(n\,\text{even}) : \begin{bmatrix} M & B \\ B^T & PMP \end{bmatrix}, \; (n\,\text{odd}) : \begin{bmatrix} M & C & B \\ C^T & d & C^T P \\ & & \\ B^T & PC & PMP \end{bmatrix}, \tag{108}$$

where $M$ is a symmetric matrix. Anti-bisymmetric matrices have the following two forms:

$$(n\text{even}) : \begin{bmatrix} M & B \\ & \\ -B^T & -PMP \end{bmatrix}, \; (n\text{odd}) : \begin{bmatrix} M & C & B \\ -C^T & d & C^T P \\ & & \\ -B^T & -PC & -PMP \end{bmatrix}, \tag{109}$$

where $M$ is anti-symmetric. Thus, we present only the left boundary closure with the understanding that the right boundary closure follows immediately. The coefficients are determined from the accuracy equations given by

$$D_1 \mathbf{x}^j = H^{-1} \Theta \mathbf{x}^j = j \mathbf{x}^{j-1}, \; j \in [0, p]. \tag{110}$$

However, to avoid solving nonlinear equations, it is easier to multiply through by $H$ and solve the following

$$\Theta \mathbf{x}^j = j H \mathbf{x}^{j-1}, \; j \in [0, p]. \tag{111}$$

### *7.2. Pseudo-spectral collocation SBP operators*

This section highlights how the theory presented can be applied to derive a wide array of pseudo-spectral SBP operators based on Legendre polynomials, as well as the associated SATs that lead to consistent, conservative and stable discretizations. This is not a new idea; recently, Gassner [11] examined the Legendre-Gauss-Lobatto (LGL) quadrature points and showed that one can construct the various portions of an SBP operator with a diagonal norm. In an earlier paper, Carpenter and Gottlieb [3], dealt more generally with collocated-pseudo-spectral methods on the LGL quadrature points and defined a dense norm. Despite the norms being different, the $\Theta$ and $D_1$ operators are identical. Here, we present a different approach starting from the quadrature rule itself. Furthermore, we demonstrate how the theory applies to quadratures that do not include the end-points of the domain, which is the case with Legendre-Gauss (LG) and Legendre-Gauss-Radau (LGR) quadratures.

The following is an outline of the approach taken here to derive SBP operators. Given a quadrature rule $W = [w_1, w_2, \ldots, w_n]$ of degree $\tau$ defined on the nodal distribution $\mathbf{x}$ with $x_i \in [\alpha, \beta]$:

- Restrict the coefficients of $H$ such that $\mathbb{1}^T H = W$;
- Construct $t_\alpha$, $t_\beta$, $T$ and $\tilde{E}$ from the nodal distribution $\mathbf{x}$ using equations (74);
- Restrict the coefficients of $\Theta$ such that $\Theta + \Theta^T = \tilde{E}$; and
- Solve accuracy equations (10) or (11) for the remaining free coefficients in $H$ and $\Theta$.

To demonstrate these steps, we will apply them to the construction of SBP operators on the LGR quadrature points. These points were chosen since they are asymmetric in the domain $[\alpha, \beta] = [-1, 1]$, including one boundary point, but not the other.

Consider the three-point LGR quadrature of degree $\tau = 4$ with nodal distribution $\mathbf{x}^T = \left[-1, \frac{1}{5} - \frac{1}{5}\sqrt{6}, \frac{1}{5} + \frac{1}{5}\sqrt{6}\right]$ and associated quadrature weights $W = \left[\frac{2}{9}, -\frac{5}{18}\,\frac{\sqrt{6}\left(-2+3\sqrt{6}\right)}{-6+\sqrt{6}}, \frac{5}{18}\,\frac{\sqrt{6}\left(2+3\sqrt{6}\right)}{6+\sqrt{6}}\right]$. Since all of the quadrature weights are positive, a diagonal $H$ can be constructed as $(H)_{ii} = w_i$. Alternatively, one must satisfy $\left(\mathbb{1}^T H\right)_i = w_i$ for a dense $H$.

The simplest means of constructing $t_\alpha$ and $t_\beta$ is to use the Lagrange interpolant through $\mathbf{x}$, which is unique and of degree $n-1$, and evaluate it at $\alpha$ and $\beta$, which gives

$$\begin{aligned} t_\alpha^T &= \left[l_0(\alpha), \ldots, l_{n-1}(\alpha)\right], \text{ and} \\ t_\beta^T &= \left[l_0(\beta), \ldots, l_{n-1}(\beta)\right], \end{aligned} \tag{112}$$

where $l_j(x)$ are the Lagrangian basis functions and are defined by

$$l_j(x) = \prod_{\substack{0 \le m \le n-1 \\ m \ne j}} \frac{x - x_m}{x_j - x_m}. \tag{113}$$

$T$ is then constructed as $T = e_1 t_\alpha^T + e_n t_\beta^T$ and $\tilde{E} = T^T E T$. In the present case

$$\begin{aligned} t_\alpha^T &= [1, 0, \ldots, 0]\,, \text{ and} \\ t_\beta^T &= \left[\tfrac{1}{3}, \tfrac{5}{3}\,\tfrac{2\sqrt{6}-3}{-6+\sqrt{6}}, \tfrac{5}{3}\,\tfrac{2\sqrt{6}+3}{6+\sqrt{6}}\right], \end{aligned} \tag{114}$$

from which $T$ and $\tilde{E}$ can be constructed. The SBP property is enforced by solving the equations $\Theta + \Theta^T = \tilde{E}$. Finally, the accuracy equations (11) are solved, giving a $D_1$ of degree 2:

$$D_1 = \begin{bmatrix} -2 & 1 + \frac{7}{12}\sqrt{6} & 1 - \frac{7}{12}\sqrt{6} \\ -3/5\,\frac{\sqrt{6}(-6+\sqrt{6})(-1/9 - \frac{8}{27}\sqrt{6})}{-2+3\sqrt{6}} & -3/5\,\frac{\sqrt{6}(-6+\sqrt{6})(\frac{29}{36} - 1/6\sqrt{6})}{-2+3\sqrt{6}} & -3/5\,\frac{\sqrt{6}(-6+\sqrt{6})(-\frac{25}{36} + \frac{25}{54}\sqrt{6})}{-2+3\sqrt{6}} \\ 3/5\,\frac{\sqrt{6}(6+\sqrt{6})(-1/9 + \frac{8}{27}\sqrt{6})}{2+3\sqrt{6}} & 3/5\,\frac{\sqrt{6}(6+\sqrt{6})(-\frac{25}{36} - \frac{25}{54}\sqrt{6})}{2+3\sqrt{6}} & 3/5\,\frac{\sqrt{6}(6+\sqrt{6})(\frac{29}{36} + 1/6\sqrt{6})}{2+3\sqrt{6}} \end{bmatrix}, \tag{115}$$

with

$$\Theta = \begin{bmatrix} -\frac{4}{9} & \frac{2}{9} + \frac{7}{54}\sqrt{6} & \frac{2}{9} - \frac{7}{54}\sqrt{6} \\ -\frac{1}{9} - \frac{8}{27}\sqrt{6} & \frac{29}{36} - \frac{1}{6}\sqrt{6} & -\frac{25}{36} + \frac{25}{54}\sqrt{6} \\ -\frac{1}{9} + \frac{8}{27}\sqrt{6} & -\frac{25}{36} - \frac{25}{54}\sqrt{6} & \frac{29}{36} + \frac{1}{6}\sqrt{6} \end{bmatrix}. \tag{116}$$

Section 8 presents some simple numerical results obtained with SBP operators derived from the LGL, LGR, and LG quadratures.

### *7.3. SBP operators based on Newton-Cotes quadrature*

Theorem 2 states that positive quadrature weights are required for the construction of diagonal-norm operators. However, it is well known that the quadrature weights of closed Newton-Cotes formulae are only positive up to 10 points with the exception of the 9-point quadrature rule [43]. In the case of negative weights, the degree of the quadrature can be reduced, freeing coefficients which can then be used to satisfy the PD requirement of the norm. The resulting quadrature is no longer a Newton-Cotes rule, but extends this form of evenly spaced quadrature rules for diagonal norms beyond 10 points and degree $\tau = 9$.

For example, consider a quadrature of degree $\tau = 11$. This can be achieved with an 11-point Newton-Cotes quadrature; however, it has negative weights. Noting that the quadrature is invariant under the transformation from $x \rightarrow -x$, we only show the first $\lceil \frac{n}{2} \rceil$ weights:

$$W_{1\ldots 6} = \left[ \begin{array}{cccccc} \frac{16067}{299376} & \frac{26575}{74844} & -\frac{16175}{99792} & \frac{5675}{6237} & -\frac{4825}{5544} & \frac{17807}{12474} \end{array} \right]. \tag{117}$$

To obtain a positive quadrature of degree $\tau = 11$ on an equally spaced nodal distribution requires at least 14 points, 3 more than the classical Newton-Cotes quadrature. An example of such a quadrature, where only the first $\lceil \frac{n}{2} \rceil$ weights are shown, is as follows:

$$W_{1\ldots 7} = \left[ \begin{array}{ccccccc} \frac{834231029}{18968463360} & \frac{2098059869}{8622028800} & \frac{20497297}{878169600} & \frac{573325999}{2155507200} & \frac{269917811}{1724405760} & \frac{14097547}{319334400} & \frac{12500}{56133} \end{array} \right]. \tag{118}$$

This yields an SBP operator of degree $q = 6$. Such operators are not unique, containing free parameters in both $H$ and $\Theta$ that can be used for a particular purpose, such as, reducing the truncation error, or optimizing the spectral properties of the scheme. Here we present one such possible scheme. The resulting nearly anti-bisymmetric $\Theta$, using (109), can be defined to 5 decimal places by

$$M_\Theta = \begin{bmatrix} -1/2 & 0.72588 & -0.15737 & -0.42132 & 1.23717 & -1.84246 & 0.97891 \\ -0.72588 & 0 & 0.15104 & 3.45601 & -10.52502 & 16.19056 & -8.93953 \\ 0.15737 & -0.15104 & 0 & -8.41563 & 35.24034 & -59.48605 & 34.91258 \\ 0.42132 & -3.45601 & 8.41563 & 0 & -50.66990 & 112.70838 & -73.99512 \\ -1.23717 & 10.52502 & -35.24034 & 50.66990 & 0 & -99.88748 & 86.51115 \\ 1.84246 & -16.19056 & 59.48605 & -112.70838 & 99.88748 & 0 & -44.01227 \\ -0.97891 & 8.93953 & -34.91258 & 73.99512 & -86.51115 & 44.012268 & 0 \end{bmatrix}, \tag{119}$$

$$B_\Theta = \begin{bmatrix} 1.18785 & -2.99269 & 2.79045 & -0.70437 & -0.89422 & 0.77656 & -0.18440 \\ -11.23900 & 28.84350 & -28.24144 & 12.15592 & 1.02788 & -2.93059 & 0.77656 \\ 46.21080 & -121.60357 & 124.80838 & -67.78605 & 15.97922 & 1.02788 & -0.89422 \\ -106.27225 & 289.87817 & -312.23200 & 191.53628 & -67.78605 & 12.15592 & -0.70437 \\ 144.40016 & -416.38168 & 473.51506 & -312.23200 & 124.80838 & -28.24144 & 2.79045 \\ -107.61800 & 341.56949 & -416.38168 & 289.87817 & -121.60357 & 28.84350 & -2.99269 \\ 28.78616 & -107.61800 & 144.40016 & -106.27225 & 46.21080 & -11.23900 & 1.18785 \end{bmatrix}. \tag{120}$$

An alternative approach is to use a dense norm. Theorem 4 guarantees the existence of a dense-norm SBP operator given any quadrature rule, independent of the sign of the quadrature weights. This can indeed be done in the case of Newton-Cotes formulae with negative quadrature weights. Continuing with the example, an 11-point dense bisymmetric norm of degree $\tau = 11$ can be derived and, using (108), is defined by the following matrices:

$$M_H = \begin{bmatrix} \frac{9306678962671}{299360989287360} & \frac{3749461876015}{59872197857472} & -\frac{90050785535}{782643109248} & \frac{958663149805}{4989349821456} & -\frac{695541722335}{2851057040832} \\ \frac{3749461876015}{59872197857472} & \frac{28401948875}{62890964136} & -\frac{3438675839125}{6652466428608} & \frac{545920266625}{623668727682} & -\frac{1612158311125}{1425528520416} \\ -\frac{90050785535}{782643109248} & -\frac{3438675839125}{6652466428608} & \frac{9684135875}{8122669632} & -\frac{928686293875}{554372202384} & \frac{53761503625}{24368008896} \\ \frac{958663149805}{4989349821456} & \frac{545920266625}{623668727682} & -\frac{928686293875}{554372202384} & \frac{321998785250}{103944787947} & -\frac{469112800375}{118794043368} \\ -\frac{695541722335}{2851057040832} & -\frac{1612158311125}{1425528520416} & \frac{53761503625}{24368008896} & -\frac{469112800375}{118794043368} & \frac{49210500875}{8485288812} \end{bmatrix}, \tag{121}$$

$$C_H = \begin{bmatrix} \frac{542627709523}{2375880867360} \\ \frac{64770208025}{59397021684} \\ -\frac{116196053525}{52797352608} \\ \frac{20356362725}{4949751807} \\ -\frac{9861191675}{1616245488} \end{bmatrix}, \; d_H = \frac{16897031273}{2357024670}, \tag{122}$$

and

$$B_H = \begin{bmatrix} -\frac{440685034315}{2851057040832} & \frac{28025138965}{383796140112} & -\frac{73892199595}{3326233214304} & \frac{186312802135}{59872197857472} & -\frac{542627709523}{598721978574720} \\ -\frac{1105351670125}{1425528520416} & \frac{14695762625}{36686395746} & -\frac{20188214125}{135764620992} & \frac{323851040125}{7484024732184} & \frac{186312802135}{59872197857472} \\ \frac{522020627875}{316784115648} & -\frac{511439062375}{554372202384} & \frac{580980267625}{1478325873024} & -\frac{20188214125}{135764620992} & -\frac{73892199595}{3326233214304} \\ -\frac{386465405875}{118794043368} & \frac{203563627250}{103944787947} & -\frac{511439062375}{554372202384} & \frac{14695762625}{36686395746} & \frac{28025138965}{383796140112} \\ \frac{49305958375}{9697472928} & -\frac{386465405875}{118794043368} & \frac{522020627875}{316784115648} & -\frac{1105351670125}{1425528520416} & -\frac{440685034315}{2851057040832} \end{bmatrix}, \tag{123}$$

which yields an SBP operator of degree $q = 10$, with nearly anti-bisymmetric $\Theta$, defined by:

$$M_\Theta = \begin{bmatrix} -1/2 & \frac{1500841000535}{1247337455364} & -\frac{6118774216495}{2956651746048} & \frac{168786631855}{48915194328} & -\frac{309925589065}{70396470144} \\ -\frac{1500841000535}{1247337455364} & 0 & \frac{1167290099375}{369581468256} & -\frac{1020412823125}{207889575894} & \frac{2969081958125}{475176173472} \\ \frac{6118774216495}{2956651746048} & -\frac{1167290099375}{369581468256} & 0 & \frac{28600922500}{11549420883} & -\frac{4647044375}{1552863312} \\ -\frac{168786631855}{48915194328} & \frac{1020412823125}{207889575894} & -\frac{28600922500}{11549420883} & 0 & \frac{6375383125}{3046001112} \\ \frac{309925589065}{70396470144} & -\frac{2969081958125}{475176173472} & \frac{4647044375}{1552863312} & -\frac{6375383125}{3046001112} & 0 \end{bmatrix}, \tag{124}$$

$$C_\Theta = \begin{bmatrix} \frac{54906115193}{13199338152} \\ -\frac{2794260905}{471404934} \\ \frac{3724177855}{1257079824} \\ -\frac{3493410145}{1649917269} \\ \frac{6890202535}{3771239472} \end{bmatrix}, \; d_\Theta = 0, \tag{125}$$

and

$$B_\Theta = \begin{bmatrix} -\frac{2718536777135}{950352346944} & \frac{165568121195}{118794043368} & -\frac{4438316555}{9662260608} & \frac{83679973235}{831558303576} & -\frac{1277065708991}{79829597143296} \\ \frac{650128240625}{158392057824} & -\frac{422385469375}{207889575894} & \frac{4189781875}{5866372512} & -\frac{12243513125}{47974517514} & \frac{83679973235}{831558303576} \\ -\frac{463398116875}{211189410432} & \frac{57651340625}{46197683532} & -\frac{656406923125}{985550582016} & \frac{4189781875}{5866372512} & -\frac{4438316555}{9662260608} \\ \frac{73541879375}{39598014456} & -\frac{967607500}{679377699} & \frac{57651340625}{46197683532} & -\frac{422385469375}{207889575894} & \frac{165568121195}{118794043368} \\ -\frac{80903830625}{45254873664} & \frac{73541879375}{39598014456} & -\frac{463398116875}{211189410432} & \frac{650128240625}{158392057824} & -\frac{2718536777135}{950352346944} \end{bmatrix}. \tag{126}$$

The degree of the quadrature is the same as the diagonal norm (118); however, three fewer nodes are required and the resulting operator is exact for polynomials 4 degrees higher.

As seen from the example, besides enabling the use of quadrature rules with negative weights, a further advantage of using dense norms with Newton-Cotes quadratures is the increased degree that is possible for the derivative operators, where for the diagonal-norm case $q = \lceil \frac{n-1}{2} \rceil$, the dense-norm operators achieve $q = n - 1$.

Some simple numerical results of Newton-Cotes based SBP operators are presented in Section 8.

### *7.4. Chebyshev-polynomial-based SBP operators*

Chebyshev-polynomial-based quadrature rules have many attractive characteristics. Firstly, Chebyshev polynomials have been shown to be particularly well suited for approximating functions on finite domains, minimizing the Runge phenomenon and providing an efficient alternative to the optimal minmax polynomial approximation [12]. Furthermore, despite being lower order than LG quadrature rules, they often exhibit similar convergence rates for certain classes of problems [42, 51].

Clenshaw-Curtis quadrature points are the extrema of the Chebyshev polynomials, which include the boundary nodes. The related Fejér quadratures of the first and second kind use the roots of the Chebyshev polynomial of the first and second kind which do not include the boundary nodes. All three quadratures have degree $\tau = n - 1$ for even $n$ and $\tau = n$ for odd $n$.

Similar to Newton-Cotes rules, both dense and diagonal-norm operators can be derived for each of the above mentioned quadrature rules, with a similar reduction in the degree for the diagonal-norm operators. Numerical results are presented for all three quadratures with both dense and diagonal norms in Section 8.

### *7.5. SBP operators based on barycentric rational interpolation*

There exists a vast literature on construction of interpolants and quadratures. In this section we highlight the ease of constructing SBP operators given an interpolant or a quadrature rule. We discuss the barycentric rational interpolants first proposed by Floater and Hormann [9]. These have been investigated for application to quadrature as well as the construction of approximations to the derivative, see for example [1, 26, 27]. Our purpose is not to investigate the practical aspects of these methods as applied to PDEs but rather to demonstrate how one constructs an SBP operator starting from either a quadrature rule or an interpolant. Floater and Hormann [9] give the following prescription for constructing their barycentric rational interpolant on an $n$ point nodal distribution:

$$f(x) \approx r(x) = \frac{\sum_{i=0}^{n-1-d} \lambda_i(x) p_i(x)}{\sum_{i=0}^{n-1-d} \lambda_i(x)}, \tag{127}$$

where

$$\lambda_i(x) = \frac{(-1)^i}{(x - x_i) \dots (x - x_{i+d})}, \tag{128}$$

and the $p_i(x)$ are the unique Lagrange interpolants of degree at most $d$ of $f$ at the $d+1$ points $x \in [x_i, x_{i+d}]$.

For the purpose of this demonstration $d$ is chosen to be 3 and the nodal distribution equally spaced. These parameters are not necessary, but chosen to minimize the size of the family of methods generated. The operators are generated with dense norms and satisfy $\Theta + \Theta^T = \tilde{E} = \text{diag}[-1, 0, \dots, 0, 1]$.

As an example, consider the six node case ($n = 6$). First, the Lagrange basis functions are defined using sets of $d + 1 = 4$ points, namely $\tilde{\mathbf{x}}_0 = \left[-1, -\frac{3}{5}, -\frac{1}{5}, \frac{1}{5}\right]$, $\tilde{\mathbf{x}}_1 = \left[-\frac{3}{5}, -\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right]$, and $\tilde{\mathbf{x}}_2 = \left[-\frac{1}{5}, \frac{1}{5}, \frac{3}{5}, 1\right]$ for the domain $[-1, 1]$. These basis functions are then used to construct the Lagrange interpolants $p_0(x)$, $p_1(x)$, and $p_2(x)$, respectively. For example

$$\begin{aligned} p_0(x) &= \textstyle\sum_{j=1}^{d+1=4} l_{0,j}(x) f_j \\ &= -\tfrac{1}{48} \left(5\,x + 3\right) \left(25\,x^2 - 1\right) f_1 \\ &\quad + \tfrac{5}{16} \left(x + 1\right) \left(25\,x^2 - 1\right) f_2 \\ &\quad - \tfrac{5}{16} \left(5\,x - 1\right) \left(5\,x + 3\right) \left(x + 1\right) f_3 \\ &\quad + \tfrac{5}{48} \left(5\,x + 1\right) \left(5\,x + 3\right) \left(x + 1\right) f_4 \end{aligned} \tag{129}$$

where $f_j = f(\tilde{x}_{0,j})$. Now the global interpolant $r(x)$ is constructed and integrated to obtain the quadrature rule of degree $\tau = 3$; in the present case (to five decimal places):

$$W = \left[\begin{array}{cccccc} 0.13917 & 0.49914 & 0.36168 & 0.36168 & 0.49914 & 0.13917 \end{array}\right]. \tag{130}$$

Now, with the quadrature known, the rest of the operator can be formed following the process described in Section 7.2. With a dense-norm $H$, an operator of degree $q = 3$ can be constructed. One instance of such an operator is

$$D_1 = \begin{bmatrix} -\frac{61}{12} & 10 & -\frac{35}{4} & \frac{35}{6} & -\frac{5}{2} & \frac{1}{2} \\ -\frac{5}{8} & -\frac{215}{96} & \frac{35}{8} & -\frac{35}{16} & \frac{5}{6} & -\frac{5}{32} \\ \frac{5}{28} & -\frac{10}{7} & -\frac{55}{84} & \frac{5}{2} & -\frac{5}{7} & \frac{5}{42} \\ -\frac{5}{42} & \frac{5}{7} & -\frac{5}{2} & \frac{55}{84} & \frac{10}{7} & -\frac{5}{28} \\ \frac{5}{32} & -\frac{5}{6} & \frac{35}{16} & -\frac{35}{8} & \frac{215}{96} & \frac{5}{8} \\ -\frac{1}{2} & \frac{5}{2} & -\frac{35}{6} & \frac{35}{4} & -10 & \frac{61}{12} \end{bmatrix} \tag{131}$$

with bisymmetric norm, evaluated to 5 decimal places,

$$M = \begin{bmatrix} 0.08330 & 0.08675 & -0.06173 \\ 0.08675 & 0.50214 & -0.11338 \\ -0.06173 & -0.11338 & 0.53056 \end{bmatrix}, \tag{132}$$

and

$$B = \begin{bmatrix} 0.03543 & -0.01114 & 0.00656 \\ 0.09105 & -0.05629 & -0.01114 \\ -0.12025 & 0.09105 & 0.03543 \end{bmatrix}. \tag{133}$$

## 8. Numerical simulations

This section presents numerical results that illustrate the concepts set forth in this paper. No attempt is made to quantify the relative efficiency of the different schemes. Results for classical FD-SBP schemes are included for comparison.

### *8.1. Governing equation*

Consider the one-dimensional linear advection equation with unit wave speed. Using the method of manufactured solutions, a source term is added such that a steady-state solution exists:

$$\frac{\partial \mathcal{U}}{\partial t} + \frac{\partial \mathcal{U}}{\partial x} = \mathcal{S}(x), \tag{134}$$

with

$$x \in [\alpha, \beta],\ t \in [0, \infty),\ \mathcal{U}(x, 0) = \mathcal{I}(x),\ \mathcal{U}(\alpha, t) = \mathcal{G}_\alpha(t). \tag{135}$$

The source term is

$$\begin{aligned}\mathcal{S}(x) = 1024\,\mathrm{e}^{-4\,(2\,x-1)^2}\left(-\frac{25}{256}\,\pi^2 + \frac{7}{32} + x^2 - x\right)\sin\left(10\,\pi\,x\right) \\ - 320\,\mathrm{e}^{-4\,(2\,x-1)^2}\left(2\,x - 1\right)\cos\left(10\,\pi\,x\right)\pi,\end{aligned} \tag{136}$$

giving the following steady-state solution:

$$\mathcal{U}(x) = 1 + \left(\left(-32\,x + 16\right)\sin\left(10\,\pi\,x\right) + 10\,\cos\left(10\,\pi\,x\right)\pi\right)\mathrm{e}^{-4\,(2\,x-1)^2}. \tag{137}$$

The steady-state discrete form of (134) is

$$D_1\mathbf{u}_{\mathrm{d}} + \mathbf{SAT}_{\mathrm{BC}} + \mathbf{SAT}_{\mathrm{I}} = \mathbf{s} \tag{138}$$

where $D_1$ is the SBP derivative operator, $\mathbf{u}_{\mathrm{d}}$ is the solution vector, $\mathbf{SAT}_{\mathrm{BC}}$ and $\mathbf{SAT}_{\mathrm{I}}$ are the boundary and interface SATs respectively, and $\mathbf{s}$ is the forcing function projected onto the nodes. The solution domain chosen for this exercise is $[\alpha, \beta] = [0, 1]$, and the SAT coefficients are chosen such that the discretizations are equivalent to characteristic boundary conditions and are dual consistent [24].

The primary results are convergence rates based on the solution error

$$e_{\mathcal{U}} = \|\mathbf{u}_{\mathrm{d}} - \mathbf{u}\|_H = \sqrt{(\mathbf{u}_d - \mathbf{u})^T H (\mathbf{u}_d - \mathbf{u})}, \tag{139}$$

where $H$ is the norm consistent with the discretization and $\mathbf{u}$ is the exact solution projected onto the nodes. In addition, the convergence of a linear functional is investigated. Here, the integral of the solution is used as the linear functional, and convergence rates are based on the functional error:

$$e_{\mathcal{F}} = \left|\mathbb{1}^T H \mathbf{u}_d - \mathcal{F}(\mathcal{U})\right|, \tag{140}$$

where $H$ is the norm consistent with the discretization and $\mathcal{F}(\mathcal{U}) = \int_\alpha^\beta \mathcal{U}(x)dx = \sin\left(10\,\pi\,x\right)\mathrm{e}^{-16\,(x-1/2)^2} + x + 1\Big|_0^1 = 1$.

*8.2. Results*

In this section we present the solution and functional errors using examples of several of the operators discussed. In particular, results are presented for

- FD-SBP operators applied both in the traditional manner with repeating interior point operators or as elements;
- Operators based on LGL, LG, and LGR quadratures;
- Operators based on Newton-Cotes quadratures;
- Operators based on Clenshaw-Curtis and Fejér quadratures;
- Barycentric rational interpolation SBP operators.

To make the presentation concise, Table 1 lists the various SBP operators, their abbreviations, and their general properties.

The convergence of the solution error, $e_{\mathcal{U}}$, is summarized in Table 2. It can be seen that all operators have convergence rates of $q+1$, which is in line with the arguments made by Gustaffsson [14]. In particular, we numerically demonstrate that the generalized definition presented in this paper for SBP operators, in conjunction with the proposed SATs, leads to valid discretizations. For example, the SBP operators based on LG and LGR quadrature nodes obtain the expected rate of convergence. Table 2 also summarizes the convergence rates of the error in the computed functional, $e_{\mathcal{F}}$, which are approximately $\tau+1$, the degree of the quadrature plus one.

| SBP Scheme | Abbreviation | $q$ | $\tau$ |
|---|---|---|---|
| Classical Finite-Difference | | | |
| → Traditional single-block | T-FD$_{\text{Diag.}}$ | - | - |
| | T-FD$_{\text{Block}}$ | - | - |
| → Traditional 5-block | T5-FD$_{\text{Diag.}}$ | - | - |
| | T5-FD$_{\text{Block}}$ | - | - |
| → Element–based | EB-FD$_{\text{Diag.}}$ | $\frac{n-1}{4}$ | $\frac{n-1}{2}$ |
| | EB-FD$_{\text{Block}}$ | $\frac{n-1}{2}$ | $\frac{n-1}{2}$ |
| Legendre-Gauss-Lobatto | LGL | $n-1$ | $2n-3$ |
| Legendre-Gauss-Radau | LGR | $n-1$ | $2n-2$ |
| Legendre-Gauss | LG | $n-1$ | $2n-1$ |
| Newton-Cotes | NC$_{\text{Diag.}}$ | $\lceil\frac{n-1}{2}\rceil$ | $n-1$ |
| | NC$_{\text{Dense}}$ | $n-1$ | $n-1$ |
| Clenshaw-Curtis | CC$_{\text{Diag.}}$ | $\lceil\frac{n-1}{2}\rceil$ | $n-1$ |
| | CC$_{\text{Dense}}$ | $n-1$ | $n-1$ |
| Fejér (type 1) | F1$_{\text{Diag.}}$ | $\lceil\frac{n-1}{2}\rceil$ | $n-1$ |
| | F1$_{\text{Dense}}$ | $n-1$ | $n-1$ |
| Fejér (type 2) | F2$_{\text{Diag.}}$ | $\lceil\frac{n-1}{2}\rceil$ | $n-1$ |
| | F2$_{\text{Dense}}$ | $n-1$ | $n-1$ |
| Barycentric Rational Interpolation ($d = 3$) | BCRI | | |
| → Odd $n$ | | 4 | 5 |
| → Even $n$ | | 3 | 3 |

Table 1: Summary of SBP operators, their associated abbreviations and general properties. Notes: 1) the degrees $q$ and $\tau$ of traditional implementations of classical FD-SBP operators are not dependent of the number of nodes in the block/element and are therefore not reported; 2) the general properties of diagonal SBP operators based on Newton-Cotes quadrature hold only for the case of positive quadrature weights; 3) NC/CC/F1/F2 quadratures with an odd number of points, $n$, achieve one degree higher than reported in the table; 4) diagonal-norm SBP operators based on NC/CC/F1/F2 quadratures with odd numbers of points, $n$, achieve one additional degree in the operator as well; 5) the value for $q$ given for EB-FD$_{\text{Diag.}}$ applies only to $q \geq 2$.

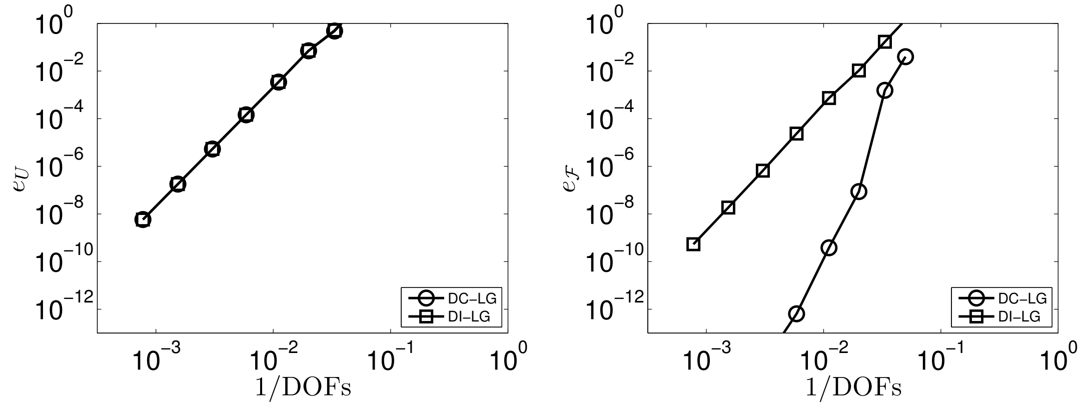
| SBP Scheme | degree | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $q = 1$ | | | $q = 2$ | | | $q = 3$ | | | $q = 4$ | | |
| | $n$ | $\tau$ | $p_{\mathcal{U}}/p_{\mathcal{F}}$ | $n$ | $\tau$ | $p_{\mathcal{U}}/p_{\mathcal{F}}$ | $n$ | $\tau$ | $p_{\mathcal{U}}/p_{\mathcal{F}}$ | $n$ | $\tau$ | $p_{\mathcal{U}}/p_{\mathcal{F}}$ |
| T-FD$_{\text{Diag.}}$ | - | 1 | 2.0084 / 2.0049 | - | 3 | 3.0720 / 4.2598 | - | 5 | 3.9984 / 6.5159 | - | 7 | 4.8175 / 10.5074 |
| T-FD$_{\text{Block}}$ | - | 1 | 2.0080 / 2.0048 | - | - | - | - | 3 | 4.0020 / 3.7822 | - | - | - |
| T5-FD$_{\text{Diag.}}$ | - | 1 | 2.0084 / 2.0232 | - | 3 | 3.0720 / 4.6398 | - | 5 | 3.9984 / 6.9676 | - | 7 | 5.0656 / 8.7761 |
| T5-FD$_{\text{Block}}$ | - | 1 | 2.0080 / 2.0208 | - | - | - | - | 3 | 4.0020 / 5.5638 | - | - | - |
| EB-FD$_{\text{Diag.}}$ | 3 | 1 | 2.0283 / 1.9943 | 9 | 3 | 2.9864 / 3.9948 | 13 | 5 | 3.9564 / 5.9838 | 17 | 7 | 4.9121 / 8.1041 |
| EB-FD$_{\text{Block}}$ | 3 | 1 | 2.0283 / 1.9943 | - | - | - | 9 | 3 | 4.0213 / 4.0743 | - | - | - |
| LGL | 2 | 1 | 1.9981 / 1.9973 | 3 | 3 | 2.9956 / 4.0253 | 4 | 5 | 3.9946 / 6.0180 | 5 | 7 | 4.9908 / 8.1091 |
| LGR | 2 | 2 | 1.9998 / 2.9989 | 3 | 4 | 2.9981 / 4.9905 | 4 | 6 | 3.9920 / 7.3133 | 5 | 8 | 4.9851 / 8.7358 |
| LG | 2 | 3 | 2.0000 / 4.0032 | 3 | 5 | 2.9994 / 6.1134 | 4 | 7 | 3.9952 / 7.9023 | 5 | 9 | 4.9803 / 10.1264 |
| NC$_{\text{Diag.}}$ | 2 | 1 | 1.9981 / 1.9973 | 3 | 3 | 2.9956 / 4.0253 | 5 | 5 | 3.9964 / 6.0252 | 7 | 7 | 4.9899 / 8.3111 |
| NC$_{\text{Dense}}$ | 2 | 1 | 1.9981 / 1.9973 | 3 | 3 | 2.9956 / 4.0253 | 4 | 3 | 3.9955 / 4.0226 | 5 | 5 | 4.9996 / 6.0235 |
| CC$_{\text{Diag.}}$ | 2 | 1 | 1.9981 / 1.9973 | 3 | 3 | 2.9984 / 4.0093 | 5 | 5 | 4.0047 / 6.1612 | 7 | 7 | 5.0527 / 8.2994 |
| CC$_{\text{Dense}}$ | 2 | 1 | 1.9981 / 1.9973 | 3 | 3 | 3.0034 / 4.0093 | 4 | 3 | 4.0096 / 4.0525 | 5 | 5 | 5.0082 / 6.1612 |
| F1$_{\text{Diag.}}$ | 2 | 1 | 2.0000 / 1.9974 | 3 | 3 | 2.9991 / 4.0107 | 5 | 5 | 4.0902 / 6.1281 | 7 | 7 | 4.9296 / 8.3169 |
| F1$_{\text{Dense}}$ | 2 | 1 | 2.0004 / 1.9954 | 3 | 3 | 3.0010 / 4.0107 | 4 | 3 | 3.9980 / 4.0282 | 5 | 5 | 4.9868 / 6.1280 |
| F2$_{\text{Diag.}}$ | 2 | 1 | 1.9998 / 1.9995 | 3 | 3 | 2.9990 / 4.0084 | 5 | 5 | 3.9647 / 6.1443 | 7 | 7 | 4.9827 / 8.3239 |
| F2$_{\text{Dense}}$ | 2 | 1 | 1.9999 / 1.9983 | 3 | 3 | 2.9995 / 4.0084 | 4 | 3 | 3.9942 / 4.0397 | 5 | 5 | 4.9827 / 6.1443 |
| BCRI | - | - | - | - | - | - | 4 | 3 | 3.9986 / 4.0163 | 7 | 5 | 4.9948 / 6.0227 |

Table 2: Convergence rates, $p_{\mathcal{U}}$ of the solution error, $e_{\mathcal{U}}$, and $p_{\mathcal{F}}$ of the functional error, $e_{\mathcal{F}}$. The degree of the SBP operator is given by $q$, $n$ is the number of nodes in each element, and $\tau$ is the degree of the associated quadrature rule. The convergence rates were computed using a line of best fit. For T-FD$_{\text{Diag./Block}}$ and T5-FD$_{\text{Diag./Block}}$ operators the number of nodes is not stated but the minimum is given by the EB-FD$_{\text{Diag./Block}}$ operators.

One of the disadvantages of using diagonal-norm SBP operators is that in many cases the degree of the operators is less than what can be achieved using dense norms. However, diagonal norms are required for stability in curvilinear coordinates [46]. In the FD community, the reduction in degree is mitigated by a dual-consistent implementation that leads to superconvergence of functionals i.e. the functional converges at a higher rate than the solution [21]. For the present set of simulations, including the novel operators presented, the boundary conditions have been implemented using dual-consistent SATs and the functional integrated with the quadrature associated with the norm of the discretization. The results numerically demonstrate that using such an implementation leads to superconvergence of functionals.

To demonstrate the effect of implementing SATs that are not dual consistent, the SBP operators for the LG quadrature points were used to solve (138) with dual-inconsistent SATs. Specifically, the penalty parameter for the boundary SATs was chosen as $\sigma = -\frac{3}{4}a$, but the functional was still integrated with the quadrature associated with the norm of the discretization. This choice results in a scheme that is stable but not dual-consistent. Examining Figure 1 we can see that although the convergence rate of the solution error remains the same, the convergence rate of the functional error is substantially reduced using dual-inconsistent SATs.

The maximum degree an operator approximating the first derivative can attain is $n - 1$, and there is therefore no difference in terms of the convergence rate of the solution error between the Legendre-Gauss based pseudospectral SBP operators: LGL, LG, and LGR. However, if one is concerned with functionals, and the discretization is dual-consistent, there is a potential advantage in using operators that do not include boundary points. With a dual-consistent discretization the functional error converges at the rate of approximately $\tau+1$ and so the higher the degree of the quadrature the higher the convergence rate of functionals. This means that although the pseudospectral methods that do not include boundary points (LG,LGR), have the same convergence rate in the solution error as methods that include the boundary points, the convergence rate of the functional error is improved. However, the required SATs are dense matrices that fully couple adjacent elements. This necessarily increases the bandwidth of the discretization and the computational cost.



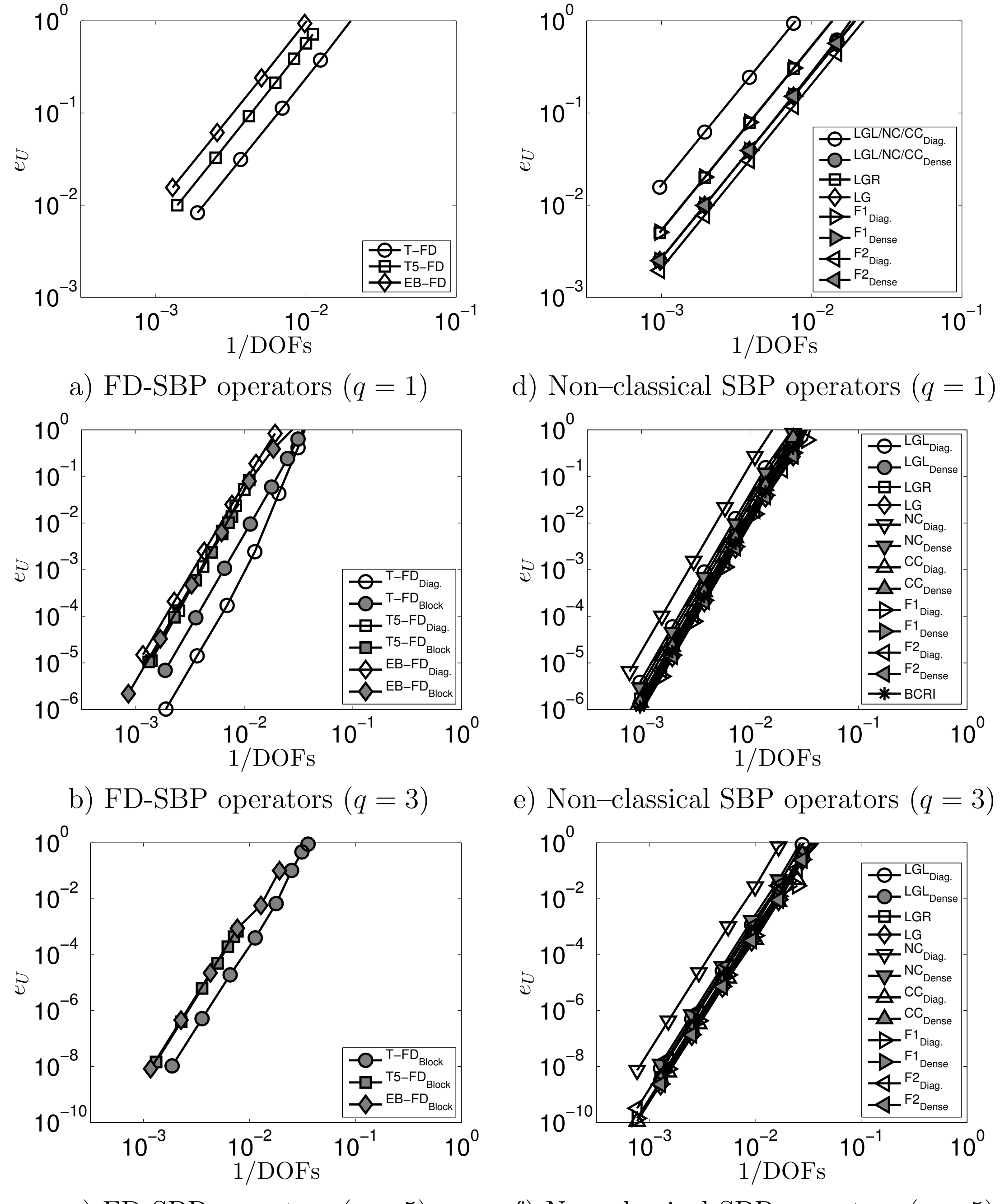
a) Solution error

b) Functional error

Figure 1: Comparison of a) solution error and b) functional error using dual-consistent (DC) and dual-inconsistent (DI) SATs for 5-node LG-quadrature-based SBP discretization.

Figure 2 (a-c) compares the convergence rates of the solution error for traditional and element-based implementations of classical FD-SBP operators for various $q$, while Figure 3 (a-c) displays the convergence rate of the computed functional, arranged in terms of $\tau$. The motivation for organizing Figure 3 based on $\tau$, rather than $q$, is to highlight the connection of the associated quadrature rule to the superconvergence of the computed functional. Similarly, results for non-classical SBP operators are presented in Figure 2 (d-f), for solution error, and Figure 3 (d-f), for the error in the computed functional.

Traditionally, classical FD-SBP operators are implemented on complex domains by decomposing the domain into a set of simple domains, each of which can be mapped onto a line, square, or cube, for one, two, or three dimensions respectively. For an operator of degree $q$, the error in a given simulation is decreased by adding more interior nodes. The generalized framework allows for the derivation of element based FD-SBP operators. It is therefore interesting to see the effect of treating classical FD-SBP operators as elements. For the purpose of this study, classical FD-SBP operators are used in three different ways: 1) using a single block (T-FD$_{\text{Diag./Dense}}$), 2) 5 blocks (T5-FD$_{\text{Diag./Dense}}$), and 3) as elements (EB-FD$_{\text{Diag./Dense}}$), where each element



a) FD-SBP operators ($q = 1$)

d) Non–classical SBP operators ($q = 1$)

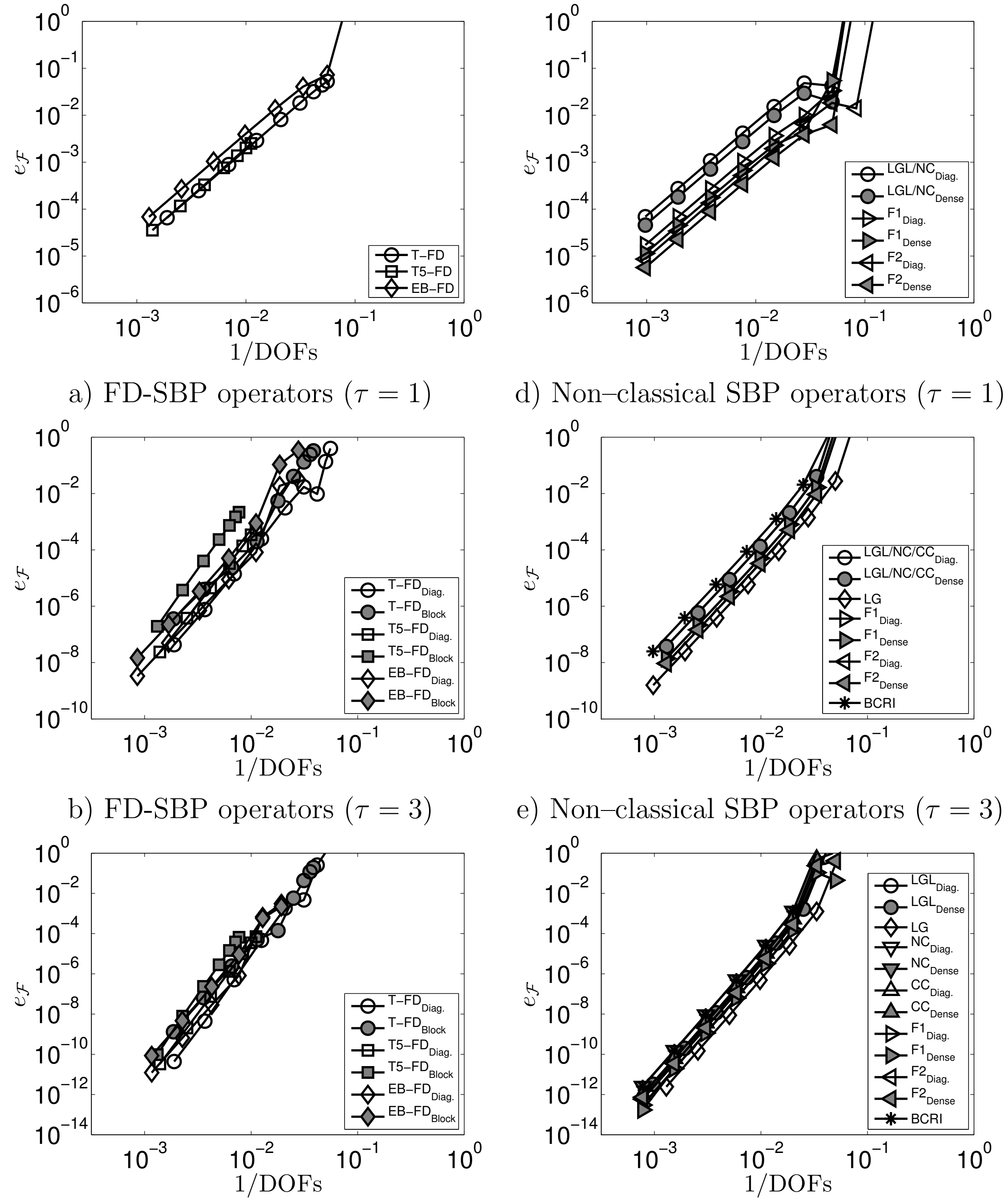
b) FD-SBP operators ($q = 3$)

e) Non–classical SBP operators ($q = 3$)

c) FD-SBP operators ($q = 5$)

f) Non–classical SBP operators ($q = 5$)

Figure 2: Convergence of solution error $e_{\mathcal{U}}$. Note that the absence of a subscript Diag. or Dense indicates that the diagonal and dense operators operators are in fact the same.



a) FD-SBP operators ($\tau = 1$)

d) Non–classical SBP operators ($\tau = 1$)

b) FD-SBP operators ($\tau = 3$)

e) Non–classical SBP operators ($\tau = 3$)

c) FD-SBP operators ($\tau = 5$)

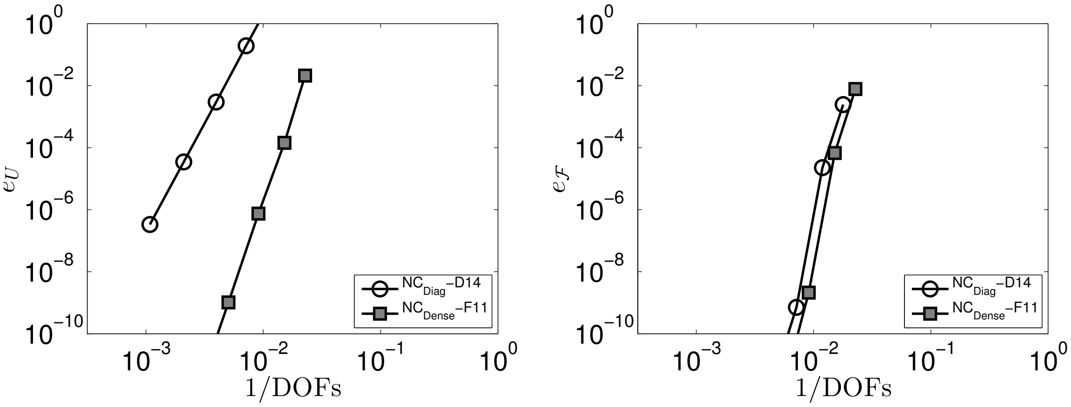
f) Non–classical SBP operators ($\tau = 5$)

Figure 3: Convergence of the functional error $e_{\mathcal{F}}$. Note that the absence of a subscript Diag. or Dense indicates that the diagonal and dense operators operators are in fact the same.

has a sufficient number of nodes so that one node has the interior point operator. The first two cases represent the traditional method of implementing FD-SBP operators where accuracy is increased by increasing the number of nodes with a fixed number of blocks, while for the third case the number of elements is increased with a constant $n$.

Figure 3 (a-c) displays the convergence of the error in the computed functional for classical FD-SBP operators. The interior point operators of the diagonal and block norm SBP operators are the same, but the degree of the resultant operator is different, with the block-norm operators having higher degree. For norms associated with quadrature rules of degree $\tau > 1$, the choice of diagonal or block norm and discretization strategy, 1 block, 5 block, or element-based, has a large effect. Overall the results displayed in Figures 2 and 3 are consistent with the theory presented.

Finally, Figure 4 displays the difference in both solution error and functional error for the diagonal and dense-norm Newton-Cotes rule based SBP operators discussed in Section 7.3. As anticipated, the dense-norm operator has a better rate of solution convergence, relative to the diagonal-norm operator, and a lower solution error. However, because of the dual-consistent implementation, both operators have similar functional error.



a) Solution error

b) Functional error

Figure 4: Comparison of a) solution error and b) functional error using a 14-point diagonal-norm equally spaced SBP operator ($q = 6$, $\tau = 11$) and an 11-point-dense-norm Newton-Cotes SBP operator ($q = 10$, $\tau = 11$).

## 9. Conclusions and future work

This paper has extended the FD-SBP theory to a more general class of operators including: i) non-repeating interior point operators, ii) nonuniform nodal distributions in the computational domain, and iii) operators that do not include one or both boundary nodes. The approach has been to determine the necessary and sufficient conditions for the existence of nodal SBP operators. We found that SBP operators are intimately tied to quadrature rules and proved that given a quadrature rule an SBP operator is guaranteed to exist. Conversely, we proved that the norm of an SBP operator must be associated with a quadrature. The extension of the FD-SBP theory to operators that do not necessarily include boundary nodes required generalizing the concept of the definition of an SBP operator. The SBP property by itself is sufficient to guarantee stability for Cauchy problems. However, for initial-boundary-value problems, the SBP property is insufficient. We derived SATs for the generalized SBP operators for the imposition of boundary conditions and inter-block/element coupling that lead to consistent, conservative, and stable numerical algorithms.

The extensions proven here allow for a large class of operators to be considered within the definition of SBP operators and enables the rigorous development of SATs for such operators. The examples considered in this paper include the following: Legendre-Gauss, Legendre-Gauss-Radau, Newton-Cotes, Clenshaw-Curtis, Fejér, Gauss-Chebyshev-quadratures, and Barycentric rational interpolation. A selection of these operators was used to solve the steady linear convection equation with a source term. The boundary conditions and block/element interfaces were numerically implemented using dual-consistent SATs. It was found that the solution error converged at rate of $q+1$, where $q$ is the degree of the operator. The convergence rate of the computed error of a simple functional, the integral of the solution, was shown to display superconvergence of $\tau+1$, where $\tau$ is the degree of the underlying quadrature of the norm matrix. We also demonstrated that without dual-consistent SATs this superconvergence is lost. This result implies that the concept of dual-consistency applies to the generalized SBP operators presented in this paper. Extending the theory of dual-consistent classical FD-SBP operators to generalized SBP operators that include both boundary nodes should be straightforward; for generalized SBP operators that do not include one or both boundary nodes, we also do not immediately see any impediments. Although operators that do not include boundary nodes have the potential to have higher degree quadratures, the resultant SATs are dense matrices that couple adjacent elements and the implications of this on efficiency require investigation.

Finally, the traditional implementation of classical FD-SBP operators was contrasted with an element-based approach. The latter opens up the use of $p$-refinement with FD-SBP operators in a similar manner to, for example, DG schemes.

The theory presented is based on one-dimensional operators that are extended to multiple dimensions through Kronecker products. Further generalizations are possible to multidimensional operators that can be applied, for example, to simplex elements, or multidimensional meshless methods; the theory herein suggests that the starting point to prove the necessary and sufficient conditions for existence of SBP operators will be a quadrature rule.

## References

[1] J.P. Berrut, M.S. Floater, G. Kelin, Convergence rates of derivatives of a family of barycentric rational interpolants, Applied Numerical Mathematics 61 (2011) 989–1000.

[2] M.H. Carpenter, T.C. Fisher, N.K. Yamaleev, Boundary closures for sixth-order energy-stable weighted essentially non-oscillatory finite-difference schemes, in: Advances in Applied Mathematics, Modeling, and Computational Science, Volume 66 of *Fields Institute Communications*, Springer US, 2013, pp. 117–160.

[3] M.H. Carpenter, D. Gottlieb, Spectral methods on arbitrary grids, Journal of Computational Physics 129 (1996) 74–86.

[4] M.H. Carpenter, D. Gottlieb, S. Abarbanel, Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes, Journal of Computational Physics 111 (1994) 220–236.

[5] M.H. Carpenter, J. Nordström, D. Gottlieb, A stable and conservative interface treatment of arbitrary spatial accuracy, Journal of Computational Physics 148 (1999) 341–365.

[6] E.K.Y. Chiu, Q. Wang, R. Hu, A. Jameson, A conservative mesh-free scheme and generalized framework for conservation laws, SIAM Journal on Scientific Computing 34 (2012) A2896–A2916.

[7] P. Diener, E.N. Dorband, E. Schnetter, M. Tiglio, Optimized high-order derivative and dissipation operators satisfying summation by parts, and applications in three-dimensional multi-block evolutions, Journal of Scientific Computing 32 (2007) 109–145.

[8] T.C. Fisher, M.H. Carpenter, N.K. Yamaleev, S.H. Frankel, Boundary closures for fourth-order energy stable weighted essentially non-oscillatory finite-difference schemes, Journal of Computational Physics 230 (2011) 3727–3752.

[9] M.S. Floater, K. Hormann, Barycentric rational interpolation with no poles and high rates of approximation, Numerische Mathematik 107 (2007) 315–331.

[10] D. Funaro, D. Gottlieb, A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations, Mathematics of Computation 51 (1988) 599–613.

[11] G.J. Gassner, A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods, SIAM Journal on Scientific Computing 35 (2013) A1233–A1253.

[12] A. Gil, J. Segura, N.M. Temme, Numerical Methods for Special Functions, Society for Industrial and Applied Mathematics, 2007.

[13] J. Gong, J. Nordström, Interface procedures for finite difference approximations of the advection-diffusion equation, Journal of Computational and Applied Mathematics 236 (2011) 602–620.

[14] B. Gustafsson, The convergence rate for difference approximations to mixed initial boundary value problems, Mathematics of Computation 29 (1975) 396–406.

[15] B. Gustafsson, High Order Difference Methods for Time Dependent PDE, Springer, 2008.

[16] B. Gustafsson, H.O. Kreiss, J. Oliger, Time Dependent Problems and Difference Methods, Willey-Interscience, 1996.

[17] F. Ham, K. Mattsson, G. Iaccarino, Accurate and stable finite volume operators for unstructured flow solvers, Center for Turbulence Research Annual Briefes (2006).

[18] J.S. Hesthaven, A stable penalty method for the compressible Navier-Stokes equations: III. multidimensional domain decomposition schemes, SIAM Journal on Scientific Computing 20 (1988) 62–93.

[19] J.S. Hesthaven, A stable penalty method for the compressible Navier-Stokes equations: II. one-dimensional domain decomposition schemes, SIAM Journal on Scientific Computing 18 (1997) 658–685.

[20] J.S. Hesthaven, D. Gottlieb, A stable penalty method for the compressible Navier-Stokes equations: I. open boundary conditions, SIAM Journal on Scientific Computing 17 (1996) 579–612.

[21] J.E. Hicken, D.W. Zingg, The role of dual consistency in functional accuracy: error estimation and superconvergence, AIAA paper 2011-3070 (2011).

[22] J.E. Hicken, D.W. Zingg, Superconvergent functional estimates from summation-by-parts finite-difference discretizations, SIAM Journal on Scientific Computing 33 (2011) 893–922.

[23] J.E. Hicken, D.W. Zingg, Summation-by-parts operators and high-order quadrature, Journal of Computational and Applied Mathematics 237 (2013) 111–125.

[24] J.E. Hicken, D.W. Zingg, Dual consistency and functional accuracy: A finite-difference perspective, Journal of Computational Physics 256 (2014) 161–182.

[25] A. Kitson, R.I. McLachlan, N. Robidoux, Skew-adjoint finite difference methods on nouniform grids, New Zealand Journal of Mathematics 32 (2003) 139–159.

[26] G. Klein, Applications of Linear Barycentric Rational Interpolation, Ph.D. thesis, University of Fribourg, 2012.

[27] G. Klein, J.P. Berrut, Linear barycentric rational quadrature, BIT Numerische Mathematik 52 (2012).

[28] H.O. Kreiss, J. Lorenz, Initial-Boundary Value Problems and the Navier-Stokes Equations, Volume 47 of *Classics in Applied Mathematics*, SIAM, 2004.

[29] H.O. Kreiss, J. Oliger, Comparison of accurate methods for the integration of hyperbolic equations, Tellus 24 (1972) 199–215.

[30] H.O. Kreiss, G. Scherer, Finite element and finite difference methods for hyperbolic partial differential equations, in: Mathematical aspects of finite elements in partial differential equations, Academic Press, New York/London, 1974, pp. 195–212.

[31] D.W. Levy, K.R. Laflin, E.N. Tinoco, J.C. Vassberg, M. Mani, B. Rider, C. Rumsey, R.A. Wahls, J.H. Morrison, O.P. Brodersen, S. Crippa, D.J. Mavriplis, M. Murayama, Summary of data from the fifth AIAA CFD drag prediction workshop, AIAA paper 2013-0046 (2013).

[32] K. Mattsson, Boundary procedures for summation-by-parts operators, Journal of Scientific Computing 18 (2003) 133–153.

[33] K. Mattsson, Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients, Journal of Scientific Computing 51 (2012) 650–682.

[34] K. Mattsson, M. Almquist, A solution to the stability issues with block norm summation by parts operators, Journal of Computational Physics 15 (2013) 418–442.

[35] K. Mattsson, J. Nordström, Summation by parts operators for finite difference approximations of second derivatives, Journal of Computational Physics 199 (2004) 503–540.

[36] K. Mattsson, M. Svärd, J. Nordström, Stable and accurate artificial dissipation, Journal of Scientific Computing 21 (2004) 57–79.

[37] K. Mattsson, M. Svärd, M. Shoeybi, Stable and accurate schemes for the compressible Navier-Stokes equations, Journal of Computational Physics 227 (2008) 2293–2316.

[38] J. Nordström, M. Björck, Finite volume approximations and strict stability for hyperbolic problems, Applied Numerical Mathematics 38 (2001) 237–255.

[39] J. Nordström, M.H. Carpenter, Boundary and interface conditions for high-order finite-difference methods applied to the Euler and Navier-Stokes equations, Journal of Computational Physics 148 (1999) 621–645.

[40] J. Nordström, K. Forsberg, C. Adamsson, P. Eliasson, Finite volume methods, unstructured meshes and strict stability for hyperbolic problems, Applied Numerical Mathematics 45 (2003) 453–473.

[41] J. Nordström, J. Gong, E. van der Weide, M. Svärd, A stable and conservative high order multi-block method for the compressible Navier-Stokes equations, Journal of Computational Physics 228 (2009) 9020–9035.

[42] H. O'Hara, F.J. Smith, Error estimation in the Clenshaw-Curtis quadrature formula, The Computer Journal 11 (1968) 213–219.

[43] A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics, Volume 37 of *Texts in Applied Mathematics*, 2 ed., Springer, 2007.

[44] A. Reichert, M.T. Heath, D.J. Bodony, Energy stable numerical method for hyperbolic partial differential equations using overlapping domain decomposition, Journal of Computational Physics 231 (2012) 5243–5265.

[45] B. Strand, Summation by parts for finite difference approximations for d/dx, Journal of Computational Physics 110 (1994) 47–67.

[46] M. Svärd, On coordinate transformations for summation-by-parts operators, Journal of Scientific Computing 20 (2004) 29–42.

[47] M. Svärd, M.H. Carpenter, J. Nordström, A stable high-order finite difference scheme for the compressible Navier-Stokes equations, far-field boundary conditions, Journal of Computational Physics 225 (2007) 1020–1038.

[48] M. Svärd, J. Nordström, Stability of finite volume approximations for the Laplacian operator on quadrilateral and triangular grids, Applied Numerical Mathematics 51 (2004) 101–125.

[49] M. Svärd, J. Nordström, A stable high-order finite difference scheme for the compressible Navier-Stokes equations: No-slip wall boundary conditions, Journal of Computational Physics 227 (2008) 4805–4824.

[50] B. Swartz, B. Wendroff, The relative efficiency of finite difference and finite element methods. I: Hyperbolic problems and splines, SIAM Journal on Numerical Analysis 11 (1974) 979–993.

[51] L.N. Trefethen, Is Gauss quadrature better than Clenshaw-Curtis?, SIAM Review 50 (2008) 67–87.

[52] N.K. Yamaleev, M.H. Carpenter, A systematic methodology for constructing high-order energy stable WENO schemes, Journal of Computational Physics 228 (2009) 4248–4272.

[53] N.K. Yamaleev, M.H. Carpenter, Third-order energy stable WENO scheme, Journal of Computational Physics 228 (2009) 3025–3047.