Numerical Investigation of Tensor-Product Summation-by-Parts Discretization Strategies and Operators

David C. Del Rey Fernández* Pieter D. Boom† Mehrdad Shademan‡

David W. Zingg§

University of Toronto Institute for Aerospace Studies, Center for Research in Sustainable Aviation,

4925 Dufferin St., Toronto, Ontario, M3H 5T6, Canada

This paper presents a numerical investigation of the tradeoffs between various discretization approaches and operators, based on diagonal-norm summation-by-parts (SBP) operators, using the two-dimensional linear convection equation and simultaneous approximation terms (SATs) for the weak imposition of boundary conditions and interface coupling. In particular, it focuses on operators which include boundary nodes. Of the operators considered, the hybrid-Gauss-trapezoidal-Lobatto SBP operators are the most efficient. Little difference in efficiency is observed between the divergence and skew-symmetric forms, making the latter preferred given its provable stability on curved meshes. The traditional finite-difference refinement strategy is the most efficient, and the discontinuous element approach the least. The continuous element refinement strategy has comparable efficiency to the traditional approach when not exhibiting lower convergence rates. This motivates a hybrid approach whereby discontinuous elements are constructed from continuous sub-elements. This hybrid approach is found to inherit the higher convergence rates of the traditional and discontinuous approaches, and higher efficiency relative to the discontinuous approach.

I. Introduction

In this paper, we are interested in spatial discretizations of partial differential equations (PDEs) having the summation-by-parts (SBP) property.[1–5] SBP methods are attractive as they are mimetic of integration by parts, and when combined with simultaneous approximation terms (SATs)[6–9] for the weak imposition of boundary conditions and inter-element coupling, lead to consistent, conservative, and provably linearly and nonlinearly stable[10–13] discretizations of PDEs.

There is increasing recognition that numerous discretization methodologies can satisfy the SBP property.[3–5,14] Indeed, it has been shown that SBP operators can be developed for structured and unstructured meshes,[15,16] and can be implemented with discontinuous or continuous coupling[15] and using the traditional finite-difference or element approaches.[17] Moreover, given a distribution of nodes in computational space, novel sets of SBP operators can be constructed with preferential computational properties, such as reduced spectral radii or increased sparsity.

The generalized SBP framework[3] that we employ gives great flexibility not only in what operators to use but also how to implement the discretization. However, while this freedom is advantageous, it is necessary to determine which attributes of particular SBP-SAT combinations are useful. The purpose of this paper is therefore to perform an initial investigation of the utility of various combinations of the available SBP-SAT discretization approaches and operators. In particular, we are interested in the class of diagonal-norm

*Postdoctoral Fellow and AIAA member (dcdelrey@gmail.com)

†Postdoctoral Fellow and AIAA member (pieter.boom@mail.utoronto.ca)

‡Postdoctoral Fellow and AIAA member (mehrdad.shademan@utoronto.ca)

§University of Toronto Distinguished Professor of Computational Aerodynamics and Sustainable Aviation, Director, Center for Sustainable Aviation, and Associate Fellow AIAA (dwz@oddjob.utias.utoronto.ca)

SBP operators that include boundary nodes and in comparing the divergence form of the PDE with the skew-symmetric form.

The paper is organized as follows: in Section II, the notation used in the paper is set forth. Next, in Section III, the two classes of SBP operators, the two means of coupling subdomains in a domain decomposition strategy, and the two approaches to mesh refinement are reviewed. In Sections IV and V, the divergence and skew-symmetric continuous and semi-discrete forms of the linear convection equation in curvilinear coordinates are delineated. In Section VI, we characterize families of SBP operators and the various means of implementing them in terms of a number of metrics. Finally, conclusions are discussed in Section VII.

II. Notation

The notation used is adapted from Ref. 15. In the interest of simplicity, the presentation is restricted to two-dimensional operators; extension to three dimensions is straightforward.

Vectors are denoted with small bold letters, for example, $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_N]^{\mathrm{T}}$, while matrices are presented using capital letters with sans-serif font, for example, M. Capital letters with a script type are used to denote continuous functions on a specified domain. As an example, $\mathcal{U}(\xi,\eta,t) \in L^2\left([\xi_1,\xi_{\mathrm{N}}]\times[\eta_1,\eta_N]\times[0,T]\right)$ denotes a square integrable function on the domain $(\xi,\eta)\in[\xi_1,\xi_{\mathrm{N}}]\times[\eta_1,\eta_N]$. Lower case bold font is used to denote the restriction of such functions onto a grid; for example, the restriction of $\mathcal{U}$ onto a grid of N nodes is given by

$$\boldsymbol{u} = \left[\mathcal{U}\left(\xi_1,\eta_1\right), \ldots, \mathcal{U}\left(\xi_{\mathrm{N}},\eta_N\right)\right]^{\mathrm{T}},$$

Throughout this paper, the restriction of monomials is represented by $\boldsymbol{\xi}^k = \left[\xi_1^k, \ldots, \xi_N^k\right]^{\mathrm{T}}$, with the convention that $\boldsymbol{\xi}^k = \boldsymbol{0}$ if $k < 0$.

We discuss the degree of SBP operators, which is the degree of monomial for which they are exact. We reserve the word order to refer to the order of the solution error and discuss the accuracy of the operators and the discretizations exclusively in terms of degree.

III. Implementation of tensor-product SBP operators

Tensor-product SBP operators can be used to discretize a problem in a number of ways on structured and block-structured meshes. The first class of SBP operators, denoted traditional operators, can be applied to nodal distributions with essentially arbitrary numbers of nodes while retaining their point-wise error characteristics; this is achieved by increasing or decreasing the number of times the interior-point operator is applied. This implies that there is no restriction, beyond some minimum, on the number of nodes within each subdomain. For example, let us consider the simplest classical first-order SBP operator on a uniform nodal distribution. This operator is given as

$$\mathsf{D}_\xi = \frac{1}{h}\begin{bmatrix} -1 & 1 & & & \\ -\frac{1}{2} & 0 & \frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{2} & 0 & \frac{1}{2} \\ & & & -1 & 1 \end{bmatrix}, \quad \mathsf{H}_\xi = h\begin{bmatrix} \frac{1}{2} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & \frac{1}{2} \end{bmatrix},$$

$$\mathsf{Q}_\xi = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & & & \\ -\frac{1}{2} & 0 & \frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{2} & 0 & \frac{1}{2} \\ & & & -\frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

where h is the mesh spacing. SBP operators require a mechanism to introduce the effects of boundary conditions; in this paper we use SATs to enforce boundary conditions weakly.

The second class of SBP operators, denoted element-type, has interior degrees of freedom (DOFs) in each element. As an example of an element-type SBP operator, consider the following operator constructed on the Legendre-Gauss-Lobatto nodes $\boldsymbol{\xi} = [-1, 0, 1]^T$:

$$\mathsf{D}_\xi = \begin{bmatrix} -3 & \frac{5}{4}+\frac{5}{4}\sqrt{5} & \frac{5}{4}-\frac{5}{4}\sqrt{5} & \frac{1}{2} \\ -\frac{1}{4}-\frac{1}{4}\sqrt{5} & 0 & \frac{1}{2}\sqrt{5} & \frac{1}{4}-\frac{1}{4}\sqrt{5} \\ \frac{1}{4}\sqrt{5}-\frac{1}{4} & -\frac{1}{2}\sqrt{5} & 0 & \frac{1}{4}\sqrt{5}+\frac{1}{4} \\ -\frac{1}{2} & -\frac{5}{4}+\frac{5}{4}\sqrt{5} & -\frac{5}{4}-\frac{5}{4}\sqrt{5} & 3 \end{bmatrix}, \mathsf{H}_\xi = \begin{bmatrix} \frac{1}{6} & & & \\ & \frac{5}{6} & & \\ & & \frac{5}{6} & \\ & & & \frac{1}{6} \end{bmatrix},$$

$$\mathsf{Q}_\xi = \begin{bmatrix} -\frac{1}{2} & \frac{5}{24}+\frac{5\sqrt{5}}{24} & \frac{5}{24}-\frac{5\sqrt{5}}{24} & \frac{1}{12} \\ -\frac{5}{24}-\frac{5\sqrt{5}}{24} & 0 & \frac{5\sqrt{5}}{12} & \frac{5}{24}-\frac{5\sqrt{5}}{24} \\ -\frac{5}{24}+\frac{5\sqrt{5}}{24} & -\frac{5\sqrt{5}}{12} & 0 & \frac{5}{24}+\frac{5\sqrt{5}}{24} \\ -\frac{1}{12} & -\frac{5}{24}+\frac{5\sqrt{5}}{24} & -\frac{5}{24}-\frac{5\sqrt{5}}{24} & \frac{1}{2} \end{bmatrix}.$$

Such SBP operators can be used to discretize PDEs using the discontinuous approach. In such an approach, the domain is divided into K elements or blocks, and the SBP operator for the domain is constructed as

$$\mathsf{D}_\xi = \begin{bmatrix} (\mathsf{D}_\xi)_1 & & & \\ & (\mathsf{D}_\xi)_2 & & \\ & & \ddots & \\ & & & (\mathsf{D}_\xi)_K \end{bmatrix},$$

where the notation $()_i$ is used to reference the fact that this operator is for the i^{th} element or block. In this approach, the derivative operator for the domain does not have coupling between elements or blocks; the solution is multi-valued at interfaces and a mechanism to introduce coupling is necessary. This coupling can be achieved using SATs.

Alternatively, if the SBP operator contains boundary nodes, then the continuous approach can be used.[15] The global norm matrix, H_ξ, and stiffness matrix, Q_ξ, are assembled as follows:

$$\mathsf{H}_\xi = \sum_{i=1}^{K} \mathsf{T}_i \left(\mathsf{H}_\xi\right)_i \mathsf{T}_i^{\mathrm{T}},$$

and

$$\mathsf{Q}_\xi = \sum_{i=1}^{K} \mathsf{T}_i \left(\mathsf{Q}_\xi\right)_i \mathsf{T}_i^{\mathrm{T}},$$

where T_i is of size $[K(N-1)+1] \times N$ and

$$\mathsf{T}_i((i-1)(N-1)+1 : i(N-1)+1, :) = \mathrm{diag}(1, \ldots, 1).$$

For the four-node Legendre-Gauss-Lobatto example, if we have two elements, T_1 and T_2 are given by

$$\mathsf{T}_1^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathsf{T}_2^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

For operators that can only be constructed as elements, mesh refinement is carried out by increasing the number of elements — we denote this the element approach to mesh refinement. Alternatively, for traditional SBP operators, mesh refinement can be carried out using either the element approach or the traditional finite-difference approach, where the number of nodes within a block is increased.[17,18]

IV. The two-dimensional linear convection equation

We consider the two-dimensional linear convection equation on a unit square

$$\frac{\partial \mathcal{U}}{\partial t} + \beta_x \frac{\partial \mathcal{U}}{\partial x} + \beta_y \frac{\partial \mathcal{U}}{\partial y} = 0, \quad \forall (x,y) \in [0,1] \times [0,1], \quad t \geq 0, \tag{1}$$

where β_x and β_y are constants and here are assumed greater than zero. The initial and boundary conditions are

$$\begin{aligned} &\mathcal{U}(x,y,0) = \mathcal{F}(x,y), \\ &\mathcal{U}(x,0,t) = \mathcal{B}_x(x,t), \quad \forall (x,y) \in [0,1] \times [0,1], \quad t \geq 0, \\ &\mathcal{U}(0,y,t) = \mathcal{B}_y(y,t). \end{aligned}$$

Tensor-product SBP operators are applied in a rectilinear computational space. First, the domain Ω is partitioned into K nonoverlapping elements Ω_i. On each element, the PDE (1) is mapped from physical coordinates (x,y) to computational coordinates (ξ,η), resulting in

$$\frac{\partial \left(\mathcal{J}^{-1}\mathcal{U}\right)}{\partial t} + \frac{\partial \left(\lambda_\xi \mathcal{U}\right)}{\partial \xi} + \frac{\partial \left(\lambda_\eta \mathcal{U}\right)}{\partial \eta} = 0, \tag{2}$$

where

$$\lambda_\xi = \beta_x \frac{\partial y}{\partial \eta} - \beta_y \frac{\partial x}{\partial \eta}, \quad \lambda_\eta = -\beta_x \frac{\partial y}{\partial \xi} + \beta_y \frac{\partial x}{\partial \xi}, \quad \mathcal{J}^{-1} = \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi}.$$

The divergence form of the PDE (2) can be recast into skew-symmetric form, given as

$$\frac{\partial \left(\mathcal{J}^{-1}\mathcal{U}\right)}{\partial t} + \frac{1}{2}\frac{\partial \left(\lambda_\xi \mathcal{U}\right)}{\partial \xi} + \frac{1}{2}\frac{\partial \left(\lambda_\eta \mathcal{U}\right)}{\partial \eta} + \frac{\lambda_\xi}{2}\frac{\partial \mathcal{U}}{\partial \xi} + \frac{\lambda_\eta}{2}\frac{\partial \mathcal{U}}{\partial \eta} = 0. \tag{3}$$

SBP-SAT semi-discretizations of skew-symmetric forms of variable coefficient problems can be directly proven to be stable using the energy method.[16, 19–21] From a practical point of view, it has been observed that such discretizations offer increased robustness for discretizations of nonlinear PDEs.[22]

V. Spatial discretization

Before introducing the discretizations of (2) and (3), we define generalized SBP operators for the first derivative, applicable to general nodal distributions.[3]

Definition 1. Summation-by-parts operator for the first derivative: *A matrix operator,* $\mathsf{D}_\xi \in \mathbb{R}^{N\times N}$*, is an SBP operator approximating the derivative* $\frac{\partial}{\partial \xi}$*, on the nodal distribution* $\boldsymbol{\xi} \in [\xi_0, \xi_1]$*, of order and degree* p *if*

1. $\mathsf{D}_\xi \boldsymbol{\xi}^k = \mathsf{H}_\xi^{-1}\mathsf{Q}_\xi \boldsymbol{\xi}^k = \mathsf{H}_\xi^{-1}\left(\mathsf{S}_\xi + \frac{1}{2}\mathsf{E}_\xi\right)\boldsymbol{\xi}^k = k\boldsymbol{\xi}^{k-1}, \quad k = 0,1,\ldots,p;$
2. H_ξ*, denoted the norm matrix, is symmetric positive definite;*
3. $\mathsf{E}_\xi = \mathsf{E}_\xi^{\mathrm{T}}$, $\mathsf{S}_\xi = -\mathsf{S}_\xi^{\mathrm{T}}$*, therefore,* $\mathsf{Q}_\xi + \mathsf{Q}_\xi^{\mathrm{T}} = \mathsf{E}_\xi$*; and*
4. $\left(\boldsymbol{\xi}^i\right)^{\mathrm{T}} \mathsf{E}_\xi \boldsymbol{\xi}^j = \xi_1^{i+j} - \xi_0^{i+j}, \quad i,\ j = 0,1,\ldots,r,\ r \geq p.$

The definition of E_ξ given above allows for nodal distributions that do not contain boundary nodes. For the purpose of imposing boundary conditions using SATs, it is convenient to further decompose E_ξ as[3]

$$\mathsf{E}_\xi = \boldsymbol{t}_{\xi_1}\boldsymbol{t}_{\xi_1}^{\mathrm{T}} - \boldsymbol{t}_{\xi_0}\boldsymbol{t}_{\xi_0}^{\mathrm{T}},$$

where

$$\boldsymbol{t}_{\xi_0}^{\mathrm{T}}\boldsymbol{\xi}^k = \xi_0^k, \quad \boldsymbol{t}_{\xi_1}\boldsymbol{\xi}^k = \xi_1^k, \quad k = 0,1,\ldots,r.$$

In this paper, we concentrate on tensor-product SBP operators that include boundary nodes and take

$$\boldsymbol{t}_{\xi_0} = [1,0,0,\ldots,0]^{\mathrm{T}}, \quad \boldsymbol{t}_{\xi_1} = [0,0,\ldots,1]^{\mathrm{T}}.$$

The SBP-SAT discretization of the divergence form (2) is presented in terms of two blocks with a shared interface of constant ξ. The discretization of the divergence form of the PDE in the left (L) element is given as

$$\begin{aligned}\frac{\mathrm{d}\mathsf{J}_L^{-1}\boldsymbol{u}_L}{\mathrm{d}t} &+ \left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L + \left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\mathsf{\Lambda}_{\eta_L}\boldsymbol{u}_L = \\ &\frac{1}{2}\left(\mathsf{H}_{\xi_L}^{-1}\boldsymbol{t}_{L,\xi_1}\otimes\Sigma_{\mathrm{L}}\right)\left[\left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L - \left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R\right],\end{aligned}$$

where Σ_{L} is a diagonal matrix of SAT coefficients, and $\mathsf{I}_{L,\xi}$ and I_η are identity matrices. The variable coefficients λ_ξ and λ_η and the metric Jacobian, for example in the left element, are constructed as follows

$$\begin{aligned}\mathsf{\Lambda}_{\xi_L} &= \mathrm{diag}\left[\beta_x\left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{y}_L - \beta_y\left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{x}_L\right], \\ \mathsf{\Lambda}_{\eta_L} &= \mathrm{diag}\left[-\beta_x\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\boldsymbol{y}_L + \beta_y\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\boldsymbol{x}_L\right], \\ \mathsf{J}_L^{-1} &= \mathrm{diag}\left[\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\boldsymbol{x}_L\odot\left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{y}_L - \left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{x}_L\odot\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\boldsymbol{y}_L,\right],\end{aligned}$$

where $\odot$ denotes the Hadamard product. The vectors $\boldsymbol{x}_L$ and $\boldsymbol{y}_L$ contain the x and y locations of the nodes. Likewise, the discretization in the right (R) element is given as

$$\begin{aligned}\frac{\mathrm{d}\mathsf{J}_R^{-1}\boldsymbol{u}_R}{\mathrm{d}t} &+ \left(\mathsf{D}_{\xi_R}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R + \frac{1}{2}\left(\mathsf{I}_{R,\xi}\otimes\mathsf{D}_\eta\right)\mathsf{\Lambda}_{\eta_R}\boldsymbol{u}_R = \\ &\frac{1}{2}\left(\mathsf{H}_{\xi_R}^{-1}\boldsymbol{t}_{R,\xi_0}\otimes\Sigma_{\mathrm{R}}\right)\left[\left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R - \left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L\right].\end{aligned}$$

Here it has been assumed that in the η direction, the two elements are conforming and have the same nodal distribution.

There are two common choices for the coefficients in the SAT matrices Σ_{L} and Σ_{R}. The first can be thought of as using an upwind flux function and has the form

$$\Sigma_{\mathrm{L}} = \frac{-\tilde{\mathsf{\Lambda}}_{L,\xi} + |\tilde{\mathsf{\Lambda}}_{L,\xi}|}{|\tilde{\mathsf{\Lambda}}_{L,\xi}|}, \quad \Sigma_{\mathrm{R}} = \frac{-\tilde{\mathsf{\Lambda}}_{R,\xi} - |\tilde{\mathsf{\Lambda}}_{R,\xi}|}{|\tilde{\mathsf{\Lambda}}_{R,\xi}|},$$

where $\tilde{\mathsf{\Lambda}}_{L,\xi} = \left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\left(\mathbf{1}_{L,\xi}\otimes\mathsf{I}_\eta\right)$, $\tilde{\mathsf{\Lambda}}_{R,\xi} = \left(\boldsymbol{t}_{R,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\left(\mathbf{1}_{R,\xi}\otimes\mathsf{I}_\eta\right)$, and $\mathbf{1}$ denotes a column vector of ones. The second is typically referred to as a symmetric SAT and has the form

$$\Sigma_{\mathrm{L}} = \mathsf{I}_\eta, \quad \Sigma_{\mathrm{R}} = -\mathsf{I}_\eta.$$

Symmetric SATs provide no dissipation and sometimes lead to under-convergence when using the element approach.[4,15,23] Therefore, this article exclusively uses upwind SATs.

The discretization of the skew-symmetric form (3) in the left element is given as

$$\begin{aligned}\frac{\mathrm{d}\mathsf{J}_L^{-1}\boldsymbol{u}_L}{\mathrm{d}t} &+ \frac{1}{2}\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L + \frac{1}{2}\left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\mathsf{\Lambda}_{\eta_L}\boldsymbol{u}_L \\ &+ \frac{1}{2}\mathsf{\Lambda}_{\xi_L}\left(\mathsf{D}_{\xi_L}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_L + \frac{1}{2}\mathsf{\Lambda}_{\eta_L}\left(\mathsf{I}_{L,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{u}_L = \\ &\quad \frac{1}{4}\left(\mathsf{H}_{\xi_L}^{-1}\boldsymbol{t}_{L,\xi_1}\otimes\Sigma_{\mathrm{L}}\right)\left[\left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L - \left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R\right] \\ &\quad + \frac{1}{4}\mathsf{\Lambda}_{\xi_L}\left(\mathsf{H}_{\xi_L}^{-1}\boldsymbol{t}_{L,\xi_1}\otimes\Sigma_{\mathrm{L}}\right)\left[\left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_L - \left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_R\right],\end{aligned}$$

and in the right element we have

$$\begin{aligned}\frac{\mathrm{d}\mathsf{J}_R^{-1}\boldsymbol{u}_R}{\mathrm{d}t} &+ \frac{1}{2}\left(\mathsf{D}_{\xi_R}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R + \frac{1}{2}\left(\mathsf{I}_{R,\xi}\otimes\mathsf{D}_\eta\right)\mathsf{\Lambda}_{\eta_R}\boldsymbol{u}_R \\ &+ \frac{1}{2}\mathsf{\Lambda}_{\xi_R}\left(\mathsf{D}_{\xi_R}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_R + \frac{1}{2}\mathsf{\Lambda}_{\eta_R}\left(\mathsf{I}_{R,\xi}\otimes\mathsf{D}_\eta\right)\boldsymbol{u}_R = \\ &\quad \frac{1}{4}\left(\mathsf{H}_{\xi_R}^{-1}\boldsymbol{t}_{R,\xi_0}\otimes\Sigma_{\mathrm{R}}\right)\left[\left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_R}\boldsymbol{u}_R - \left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\mathsf{\Lambda}_{\xi_L}\boldsymbol{u}_L\right] \\ &\quad + \frac{1}{4}\mathsf{\Lambda}_{\xi_R}\left(\mathsf{H}_{\xi_R}^{-1}\boldsymbol{t}_{R,\xi_0}\otimes\Sigma_{\mathrm{R}}\right)\left[\left(\boldsymbol{t}_{R,\xi_0}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_R - \left(\boldsymbol{t}_{L,\xi_1}^{\mathrm{T}}\otimes\mathsf{I}_\eta\right)\boldsymbol{u}_L\right].\end{aligned}$$

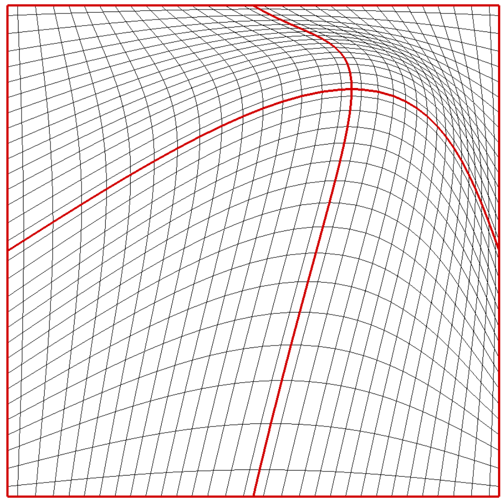


Figure 1. Example grid: Four-element curvilinear mesh using HGTL$_{17}(8,4)$ nodal distribution in each element.

VI. Results and discussion

The results presented here are for the two-dimensional linear convection equation described in equation (1) with $\beta_x = \beta_y = 1$. The initial and boundary conditions are

$$\begin{aligned}
\mathcal{F}(x,y) &= \sin(2\pi x) + \sin(2\pi y), \\
\mathcal{B}_x(x,t) &= \sin(2\pi(x-t)) + \sin(-2\pi t), \quad \forall (x,y) \in [0,1]\times[0,1], \quad t \geq 0, \\
\mathcal{B}_y(y,t) &= \sin(-2\pi t) + \sin(2\pi(y-t)),
\end{aligned}$$

and the domain is square with $x, y \subset [0,1]$. The computational domain is Cartesian, with each block or element discretized with the nodal distribution associated with the applied operator. Curved meshes are obtained by applying a perturbation to the computational mesh described above (See example in Figure 1):

$$\begin{aligned}
x &= \xi + \frac{1}{5}\sin(\pi\xi)\sin(\pi\eta), \\
y &= \eta + \frac{1}{5}\exp(1-\eta)\sin(\pi\xi)\sin(\pi\eta).
\end{aligned}$$

The solution is integrated from $t = 0$ to $t = 3$ using the standard fourth-order Runge-Kutta time-marching method and a time step such that

$$\rho\Delta t = 0.5,$$

where ρ is the spectral radius of the system matrix; with this choice, the error from the time-marching method was found to be negligible. Convergence rates on curved meshes show no influence of the temporal error for all operators considered. The accuracy of each simulation is computed relative to the exact solution of the continuous PDE and integrated using the global norm of the discretization, H, assembled from the local element norm matrices scaled by the appropriate transformation Jacobian determinant on each element

$$||\boldsymbol{u} - \mathcal{U}(\boldsymbol{x},\boldsymbol{y})||_{\mathsf{H}} = \sqrt{(\boldsymbol{u} - \mathcal{U}(\boldsymbol{x},\boldsymbol{y}))^{\mathrm{T}}\mathsf{H}(\boldsymbol{u} - \mathcal{U}(\boldsymbol{x},\boldsymbol{y}))}.$$

The basic properties of the operators investigated in this paper are summarized in Table 1 along with the acronyms used henceforth. When applying the element refinement strategies, the CSBP and HGTL elements were constructed with one more than the minimum number of DOFs in each direction. Extensive convergence studies were performed to verify the implementation of these operators.

The following conventions are used in reference to the results:

- The degree of the interior-point operator and the global degree of an operator are given by appending an ordered pair after the acronym; for example, HGTL$(\tilde{p}, p)$ refers to the hybrid-Gauss-trapezoidal-Lobatto SBP operator with an interior-point operator of degree $\tilde{p}$ and global degree p. If the operator

Operator	Abbreviation	Operator Degree	Norm Degree	Minimum # DOFs	# non-zeros per DOF
Classical FD-SBP	$\text{CSBP}_N(2,1)$	1	1	2	2
	$\text{CSBP}_N(4,2)$	2	3	8	$\frac{36+4(N-8)}{N} \leq 4.5$
	$\text{CSBP}_N(6,3)$	3	5	12	$\frac{81+6(N-12)}{N} \leq 6.75$
	$\text{CSBP}_N(8,4)$	4	7	16	$\frac{144+8(N-16)}{N} \leq 9$
Hybrid-Gauss-Trapezoidal-Lobatto	$\text{HGTL}_N(4,2)$	2	3	8	$\frac{36+4(N-8)}{N} \leq 4.5$
	$\text{HGTL}_N(6,3)$	3	5	12	$\frac{81+6(N-12)}{N} \leq 6.75$
	$\text{HGTL}_N(8,4)$	4	7	16	$\frac{144+8(N-16)}{N} \leq 9$
Diagonal Norm Lobatto	$\text{LGL}_N(N-1)$	$N-1$	$2N-3$	N	N

Table 1. Operators investigated: summary of their associated abbreviations and general properties, where N denotes the number of degrees of freedom in the operator.

is of uniform pointwise degree, only a single value is given; for example, LGL(6) refers to an LGL operator of degree 6.

- A numerical subscript is used to denote the number of DOFs in the operator, N; for example, $\text{CSBP}_5(2,1)$ refers to a classical FD-SBP operator of global degree 1 with 5 DOFs.
- Following the operator description, a suffix is appended denoting the form of the continuous PDE discretized: -d for the divergence form and -s for the skew-symmetric form.
- The refinement strategy employed is denoted with a second suffix: /t for the traditional finite-difference (FD) refinement, /c for the continuous element refinement, /d for discontinuous element refinement, and /h/K for the hybrid implementation to be described below, where K is the number of continuous sub-elements in each coordinate direction.

A. Relative efficiency

Figure 2 compares the efficiency of discretizations using CSBP, HGTL, and LGL operators on the curved meshes. Similar results are obtained for discretizations of the divergence and skew-symmetric forms. While the skew-symmetric form is slightly more accurate, it is also slightly more expensive per DOF for this test case. To minimize the data shown in the figures, only the results for the skew-symmetric form are shown.

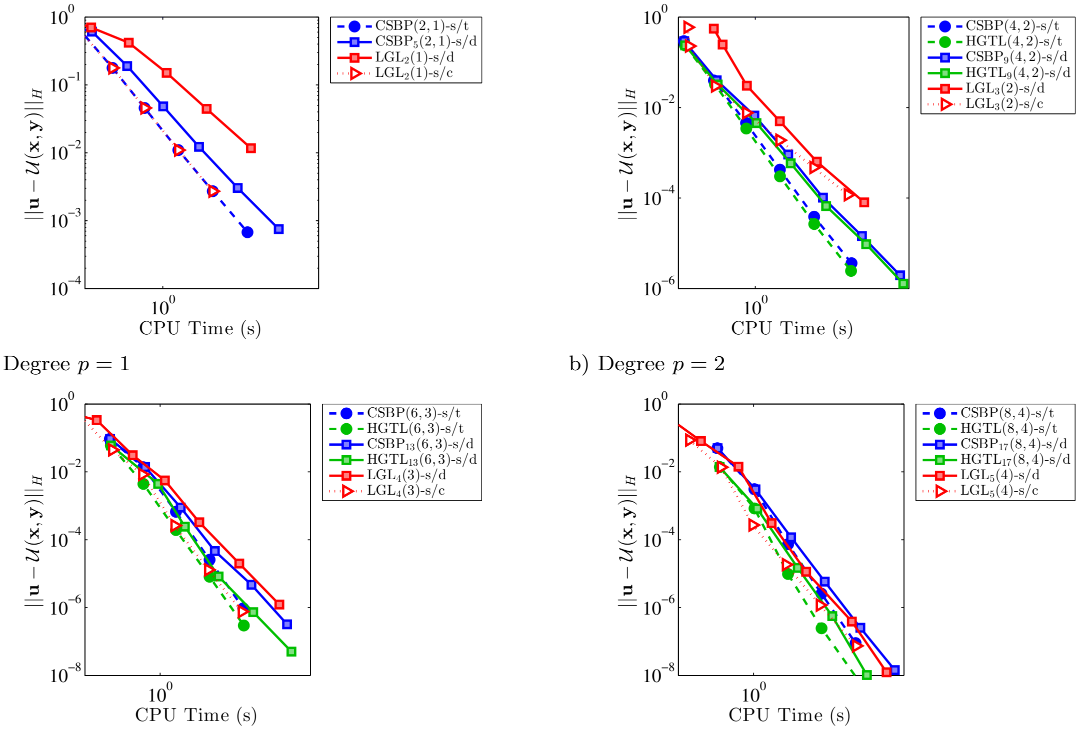
For a given degree of accuracy and refinement strategy, discretizations using HGTL operators tend to be the most efficient, followed by CSBP, and lastly LGL. The one exception is the degree four discontinuous element CSBP and LGL discretizations, which have comparable efficiency. In terms of refinement strategy, discretizations using the traditional FD approach tend to be the most efficient, and those using the discontinuous element approach the least efficient. For odd degrees of accuracy, discretizations using the continuous element approach have comparable efficiency to the traditional FD approach; however, this does not hold for even degrees of accuracy. This is expected and discussed further below.

There are a number of factors that influence these results, especially when using explicit time-marching, which include: relative accuracy, spectral radius, and the time required to evaluate the semi-discrete right-hand side (RHS). These factors are discussed further in the sections below.

1. *Accuracy*

Figure 3 shows the relative accuracy of the discretizations as a function of average mesh spacing

$$\Delta x_{\text{avg}} = \frac{1}{\sqrt{K}(N-1)}, \tag{4}$$



a) Degree $p = 1$

b) Degree $p = 2$

c) Degree $p = 3$

d) Degree $p = 4$

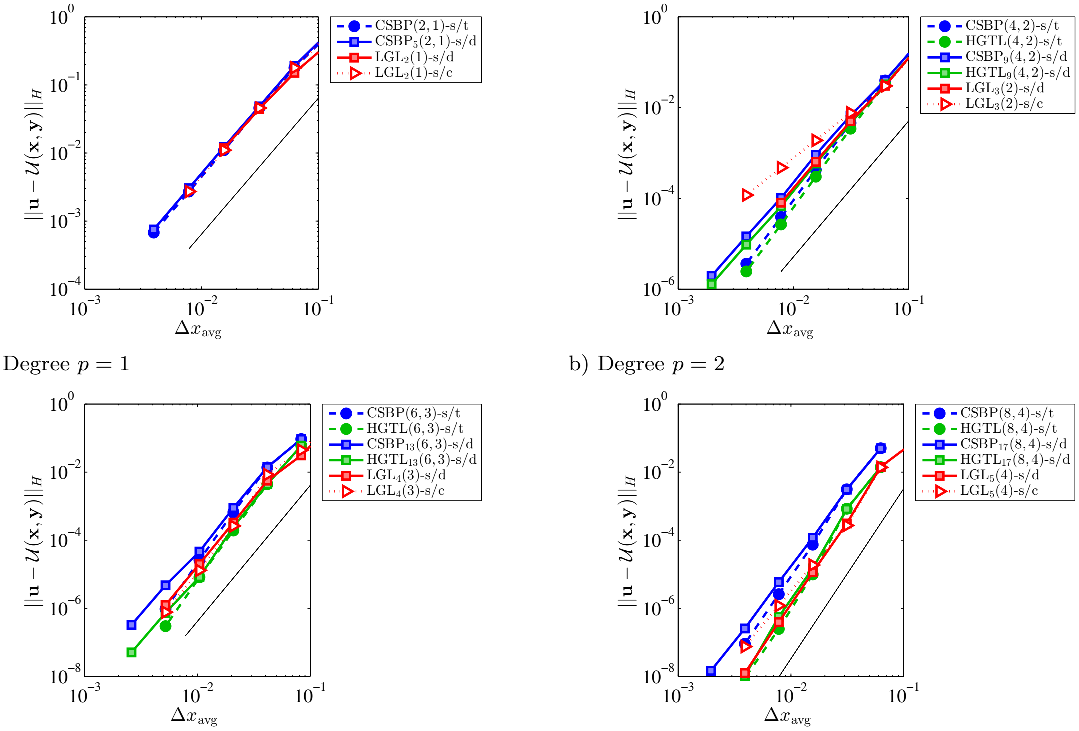
Figure 2. Relative efficiency: Comparison of discretizations on curved meshes using CSBP, HGTL, and LGL operators with a degree of accuracy of 1 through 4.

where K is the total number of elements, and N is the number of nodes per element. The discretizations all exhibit a convergence rate one order greater than their formal degree of accuracy, except for continuous element discretizations with even degrees of accuracy. These exhibit a convergence rate equal to their formal degree of accuracy. This can be seen relative to the reference slopes in the plots, purposely drawn one order higher than the formal degree of accuracy. The lower convergence rate of the continuous approach has been observed in the literature.[4,15,23]

Discretizations using the traditional FD refinement strategy tend to be the most accurate, but are limited to the use of HGTL and CSBP operators. For a given degree of accuracy, HGTL and CSBP operators share the same repeating interior-point operator and require the same minimum number of degrees of freedom. However, HGTL operators exploit a fixed number of non-uniform nodal locations at the boundaries to reduce the truncation error coefficient. HGTL discretizations are therefore consistently more accurate than CSBP discretizations. These operators are similar in form to operators proposed by Mattsson et al[24] and we anticipate that the latter should share the same preferential properties observed here. For discontinuous element discretizations, the use of HGTL operators remains the most accurate option. For degrees of accuracy one and two, CSBP operators tend to be more accurate than LGL operators, but this trend reverses for degrees of accuracy three and four. Finally, continuous element LGL discretizations with odd degrees of accuracy exhibit comparable accuracy to the traditional FD HGTL discretizations. For even degrees of accuracy, the lower rate of convergence quickly renders them inaccurate relative to other discretizations with the same degree of accuracy.

2. Spectral radii

For explicit time-marching, the spectral radius dictates the maximum stable time step size. In time-stability limited simulations, this has a direct impact on relative efficiency. The present simulations include the influence of spectral radius by selecting a time step which is a constant factor of the maximum stable time step. The spectral radius of a discretization strategy within a family of grids tends toward an asymptotic value, when scaled by the average mesh spacing. Figure 4 shows the scaled spectral radii of the various



a) Degree $p = 1$

b) Degree $p = 2$

c) Degree $p = 3$

d) Degree $p = 4$

Figure 3. Relative accuracy: Comparison of discretizations on curved meshes using CSBP, HGTL, and LGL operators with a degree of accuracy of 1 through 4.

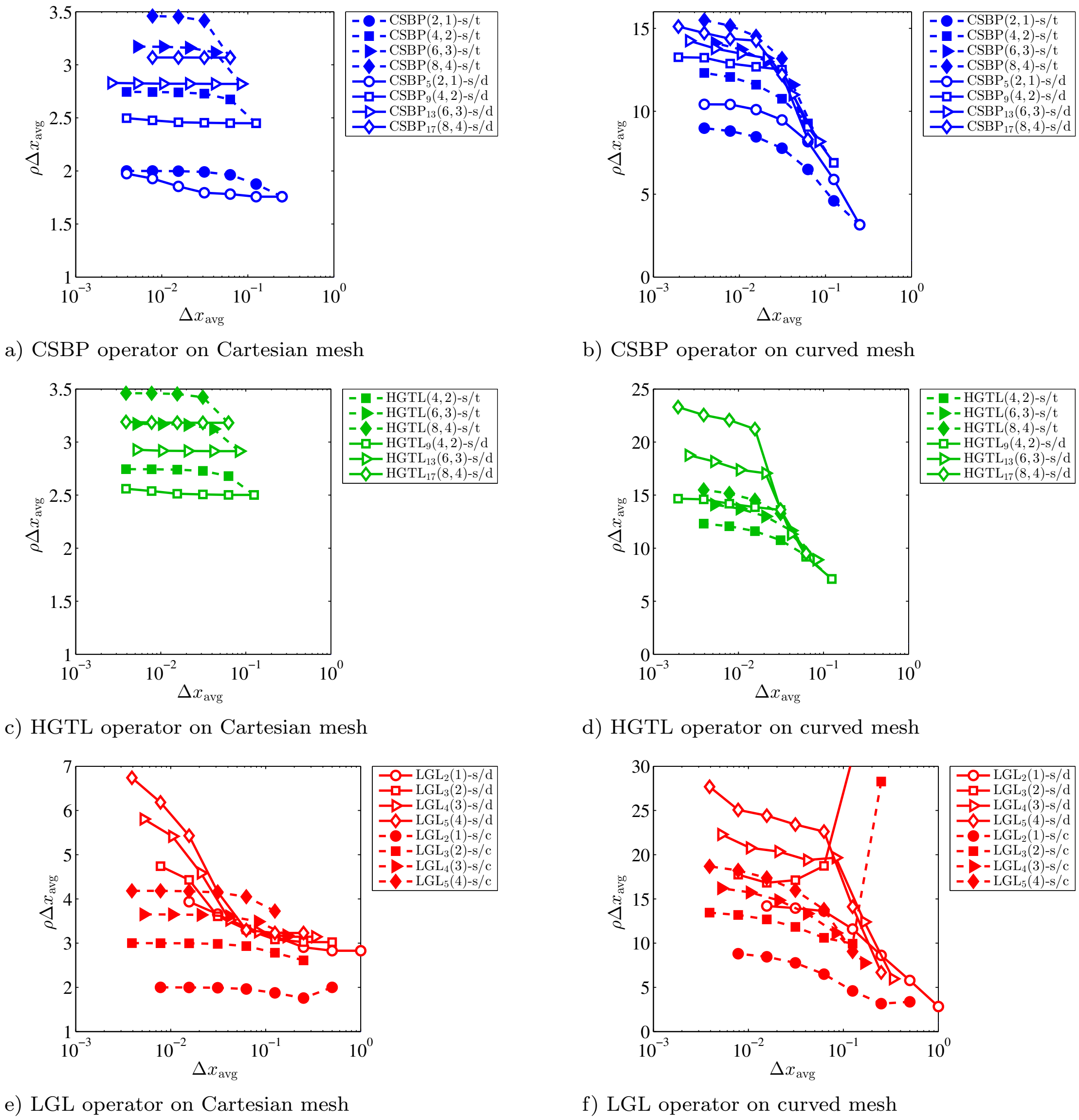
discretization strategies considered in this paper for both families of Cartesian and curved meshes described above. Results obtained for the divergence and skew-symmetric forms of the PDE are very similar, therefore only the results for the skew-symmetric form are shown.

For the Cartesian family of grids, the scaled spectral radii of CSBP and HGTL discretizations of the same degree of accuracy are very similar. The spectral radii of the discontinuous element CSBP and HGTL discretizations are consistently lower than those using a traditional FD refinement strategy. This reduces the number of time steps required, increasing the potential efficiency of the discontinuous element discretizations. Moving to the family of curved meshes, the spectral radii increase significantly. However, the spectral radii of the discontinuous element discretizations increase more than those using the traditional FD approach. This can especially be seen with the HGTL discretizations, where now the discontinuous element approach consistently yields higher spectral radii than the traditional FD approach. It requires, therefore, more time steps for stability, reducing its potential efficiency.

The LGL operators considered in this paper have a diagonal norm; therefore, no matrix inversion is required for the continuous element approach and explicit time-marching is a reasonable option for time integration. This justifies the evaluation of spectral radii for continuous element discretizations. A comparison of the spectral radii for LGL discretizations using the continuous and discontinuous element refinement strategies is also shown in Figure 4. The spectral radii of the continuous element discretizations asymptote quickly with grid resolution, and tend toward significantly lower values than the discontinuous element approach. Thus, when not affected by lower convergence rates, the continuous element approach may be an efficient alternative. Section B also evaluates a hybrid approach, where discontinuous elements are formed by a number of continuous sub-elements in order to combine the advantages of both approaches.

3. *RHS evaluation time*

For explicit time-marching, the time required to evaluate the semi-discrete RHS plays an equally important role in the relative efficiency of a simulation. This is influenced by the number of neighbors required to compute the point-difference operator at each node, the number of redundant DOFs introduced by element



a) CSBP operator on Cartesian mesh

b) CSBP operator on curved mesh

c) HGTL operator on Cartesian mesh

d) HGTL operator on curved mesh

e) LGL operator on Cartesian mesh

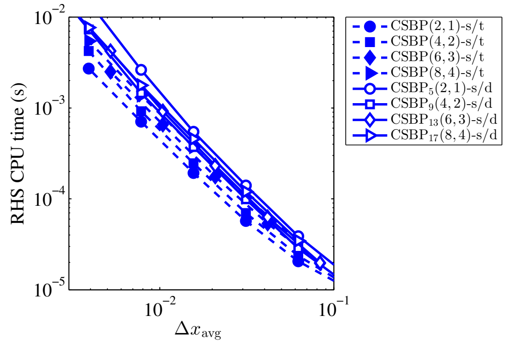
f) LGL operator on curved mesh

Figure 4. Spectral radius: Comparison of discretizations on both Cartesian (left) and curved (right) meshes for CSBP (top), HGTL (middle), and LGL (bottom) operators.

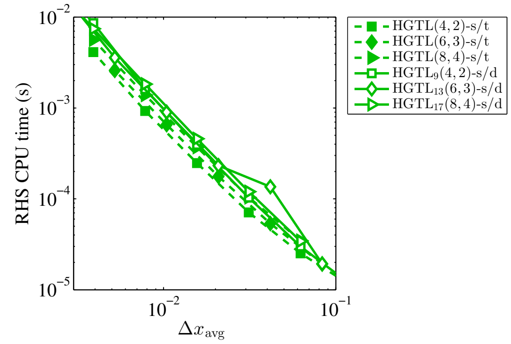
interfaces, and the relative cost of evaluating element interfaces.

Figure 5 shows the CPU time required to compute the semi-discrete RHS using the various approaches. For traditional FD and continuous element discretizations, the CPU time increases with the degree of accuracy of the method. The reason for this is that higher-order methods require more neighbors to compute the point-difference operator at each node in the mesh.

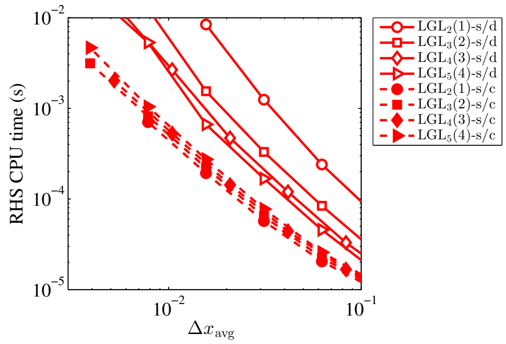
The discontinuous element discretizations consistently require more CPU time for the same average mesh spacing than the other refinement strategies. This results from the fact that the discontinuous interfaces add computational expense. The figure also shows that CPU time does not always increase monotonically with degree of accuracy. This can be seen especially for LGL discretizations, where the opposite is true. While the higher-degree operators require more neighbors to compute their point operators, the lower-degree methods require additional discontinuous interfaces that do not contribute to increasing the resolution of the simulation.



a) CSBP operator



b) HGTL operator



c) LGL operator

Figure 5. RHS evaluation time: Comparison of discretizations on curved meshes.

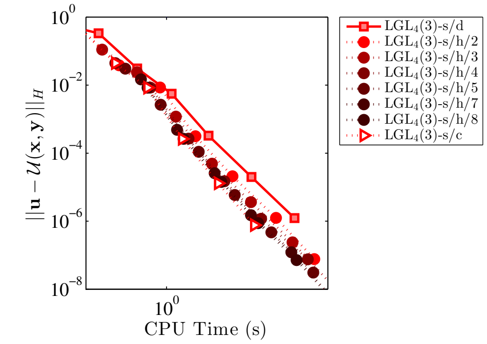
B. Hybrid approach

The hybrid element approach attempts to combine some of the advantages unique to both discontinuous and continuous element approaches. It applies a discontinuous refinement strategy to elements formed by a fixed number of continuous sub-elements in each coordinate direction. Fixing the number of sub-elements in each direction was chosen for simplicity; the method can be equally applied to elements with varying numbers of sub-elements.

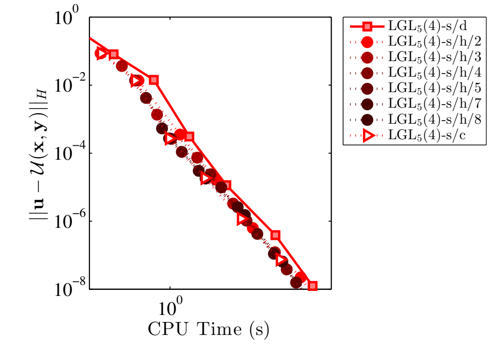
Figure 6 shows the relative efficiency, scaled spectral radius, and convergence rate of degree three and four LGL discretizations. This includes results from the continuous, discontinuous and hybrid element approaches. As before, the continuous element discretizations are more efficient than those using the discontinuous element approach. There is a significant increase in efficiency between the full discontinuous element results and the hybrid element approach with only two continuous sub-elements. The efficiency then slowly approaches the continuous approach with increasing number of continuous sub-elements.

The bottom plots show the convergence rate of the LGL discretizations. As discussed above, the even order continuous LGL discretizations exhibit a convergence rate one order lower than the discontinuous approach. At some point the lower convergence rate will render continuous element discretizations less efficient than the discontinuous element approach. Nevertheless, in the present simulations the continuous element discretizations remain more efficient. The advantage of the hybrid element discretizations is that the additional order of convergence exhibited by the discontinuous approach is recovered, along with reduced error.

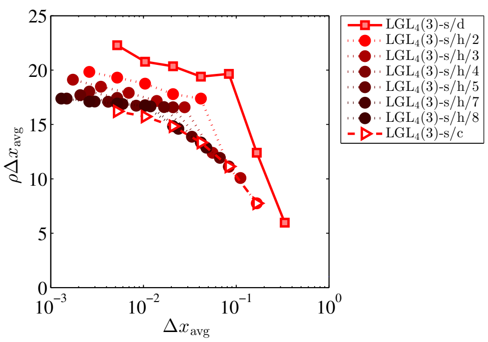
The spectral radius of the hybrid element discretizations is also reduced compared to those using the discontinuous element approach. Again, using only two continuous sub-elements has a significant impact on spectral radius. In terms of evaluation time of the semi-discrete RHS, using continuous sub-elements reduces the number of discontinuous interfaces. This reduces the number of redundant degrees of freedom required, and reduces the additional expense of computing discontinuous interfaces. Both factors contribute to the higher relative efficiency of the hybrid approach.



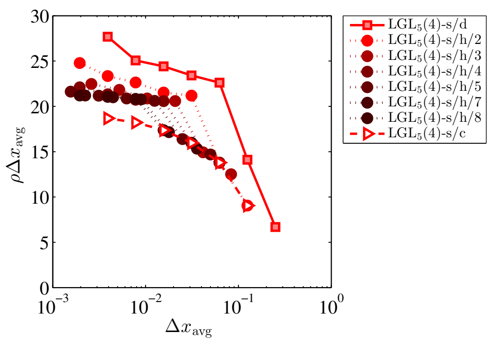
a) Efficiency degree $p = 3$



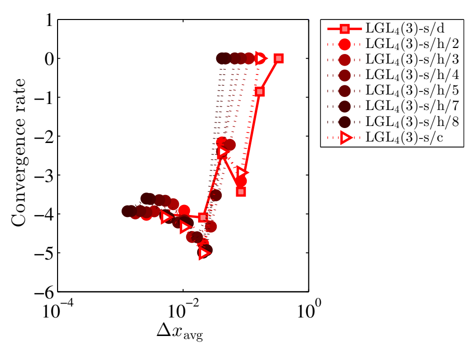
b) Efficiency degree $p = 4$



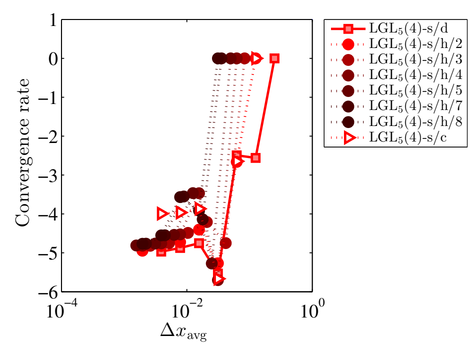
c) Spectral radius degree $p = 3$



d) Spectral raidus degree $p = 4$



e) Convergence rate degree $p = 3$



f) Convergence rate degree $p = 4$

Figure 6. Comparison of continuous, discontinuous and hybrid element approaches on curved meshes for LGL operators.

VII. Conclusions

This paper presents an initial investigation of the tradeoffs between various SBP-SAT discretization approaches and operators. The focus is on diagonal-norm SBP operators that include boundary nodes. The results are obtained using the two-dimensional linear convection equation on both Cartesian and curved meshes. The main conclusions of the study are:

- Little difference in efficiency is observed between the divergence and skew-symmetric forms. This motivates continued evaluation of discretizations of skew-symmetric forms, given the ability to prove stability on curved meshes for linear PDEs.
- HGTL operators resulted in discretizations that were consistently the most efficient for a given degree of accuracy and refinement strategy. This results from the use of non-uniform nodal locations at the boundaries, which reduce the truncation error coefficient. Future work should include application of these operators to nonlinear problems to determine whether they retain their preferential efficiency properties.

- The traditional finite-difference and continuous element refinement strategies lead to the most efficient simulations when the continuous element approach does not exhibit lower convergence rates. These simulations do not consider meshes with high stretching, where a discontinuous approach may hold an advantage; future work will consider the influence of mesh smoothness and stretching.
- The hybrid element approach leads to discretizations that inherit many of the benefits of the continuous and discontinuous element approaches. Moreover, this represents a potential methodology for mesh adaptation. This will be considered in future work.

References

[1]Del Rey Fernández, D. C., Hicken, J. E., and Zingg, D. W., "Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations," *Computers & Fluids*, Vol. 95, No. 22, 2014, pp. 171–196.

[2]Svärd, M. and Nordström, J., "Review of summation-by-parts schemes for initial-boundary-value-problems," *Journal of Computational Physics*, Vol. 268, No. 1, 2014, pp. 17–38.

[3]Del Rey Fernández, D. C., Boom, P. D., and Zingg, D. W., "A generalized framework for nodal first derivative summation-by-parts operators," *Journal of Computational Physics*, Vol. 266, No. 1, 2014, pp. 214–239.

[4]Gassner, G. J., "A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods," *SIAM Journal on Scientific Computing*, Vol. 35, No. 3, 2013, pp. A1233–A1253.

[5]Carpenter, M. H. and Gottlieb, D., "Spectral methods on arbitrary grids," *Journal of Computational Physics*, Vol. 129, No. 1, 1996, pp. 74–86.

[6]Carpenter, M. H., Gottlieb, D., and Abarbanel, S., "Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes," *Journal of Computational Physics*, Vol. 111, No. 2, 1994, pp. 220–236.

[7]Carpenter, M. H., Nordström, J., and Gottlieb, D., "A stable and conservative interface treatment of arbitrary spatial accuracy," *Journal of Computational Physics*, Vol. 148, No. 2, 1999, pp. 341–365.

[8]Nordström, J. and Carpenter, M. H., "Boundary and interface conditions for high-order finite-difference methods applied to the Euler and Navier-Stokes equations," *Journal of Computational Physics*, Vol. 148, No. 2, 1999, pp. 621–645.

[9]Nordström, J. and Carpenter, M. H., "High-order finite-difference methods, multidimensional linear problems, and curvilinear coordinates," *Journal of Computational Physics*, Vol. 173, No. 1, 2001, pp. 149–174.

[10]Parsani, M., Carpenter, M. H., and Nielsen, E. J., "Entropy stable wall boundary conditions for the three-dimensional compressible Navier-Stokes equations," *Journal of Computational Physics*, Vol. 292, No. C, 2015, pp. 88–113.

[11]Carpenter, M. H., Fisher, T. C., Nielsen, E. J., and Frankel, S. H., "Entropy Stable Spectral Collocation Schemes for the Navier–Stokes Equations: Discontinuous Interfaces," *SIAM Journal on Scientific Computing*, Vol. 36, No. 5, 2014, pp. B835–B867.

[12]Fisher, T. C. and Carpenter, M. H., "High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains," *Journal of Computational Physics*, Vol. 252, No. 1, 2013, pp. 518–557.

[13]Fisher, T. C., *High-order L2 stable multi-domain finite difference method for compressible flows*, Ph.D. thesis, Purdue University, 2012.

[14]Ranocha, H., Öffner, P., and Sonar, T., "Summation-by-parts operators for correction procedure via reconstruction," *Journal of Computational Physics*, Vol. 311, No. C, 2016, pp. 299–328.

[15]Hicken, J. E., Del Rey Fernández, D. C., and Zingg, D. W., "Multi-dimensional Summation-By-Parts Operators: General Theory and Application to Simplex Elements," *SIAM Journal on Scientific Computing*, Vol. 38, No. 4, 2016, pp. A1935–A1958.

[16]Del Rey Fernández, D. C., Hicken, J. E., and Zingg, D. W., "Simultaneous Approximation Terms for Multidimensional Summation-by-parts Operators," *Submitted to Journal of Scientific Computing (see arXiv:1605.03214v2 [math.NA])*, 2016.

[17]Hicken, J. E., Del Rey Fernández, D. C., and Zingg, D. W., "Opportunities for efficient high-order methods based on the summation-by-parts property," *22nd AIAA Computational Fluid Dynamics Conference*, No. AIAA–2015–3198, Dallas, Texas, June 2015.

[18]Fernández, D. C. D. R., Boom, P. D., and Zingg, D. W., "Corner-corrected summation-by-parts operators for the first derivative with increased order of accuracy," *Accepted to the Journal of Computational Physics*, 2016.

[19]Olsson, P. and Oliger, J., "Energy and maximum norm estimates for nonlinear conservation laws," Tech. Rep. 94–01, The Research Institute of Advanced Computer Science, 1994.

[20]Nordström, J., "Conservative Finite Difference Formulations, Variable Coefficients, Energy Estimates and Artificial Dissipation," *Journal of Scientific Computing*, Vol. 29, No. 3, 2006, pp. 375–404.

[21]Kopriva, D. A. and Gassner, G. J., "An energy stable discontinuous Galerkin spectral element discretization for variable coefficient advection problems," *SIAM Journal on Scientific Computing*, Vol. 4, No. 36, 2014, pp. A2076–A2099.

[22]Gassner, G. J. and Andrew R. Winters, D. A. K., "Split Form Nodal Discontinuous Galerkin Schemes with Summation-By-Parts Property for the Compressible Euler Equations," *Journal of Computational Physics*, Vol. 327, 2016, pp. 39–66.

[23]Hesthaven, J. S. and Warburton, T., *Nodal Discontinuous Galerkin Methods Algorithms, Analysis, and Applications*, Springer, 2008.

[24]Mattsson, K., Almquist, M., and Carpenter, M. H., "Optimal diagonal-norm SBP operators," *Journal of Computational Physics*, Vol. 264, No. 1, 2014, pp. 91–111.