ICCFD13-2026-162

Modal collocation operators with the summation-by-parts property

J. Hicken*
Corresponding author: hickej2@rpi.edu
* Rensselaer Polytechnic Institute, Troy NY, USA.

Abstract: Multidimensional summation-by-parts (SBP) operators can be used to construct high-order, conservative, and stable discretizations on complex geometries. While tensor-product SBP operators are widely used, the adoption of multidimensional SBP operators has been slow, in part because they are difficult and tedious to implement. To help address this difficulty, this work investigates an explicit formula for first-derivative SBP operators. This formula produces discretizations whose spectrum and accuracy are effectively independent of the quadrature rules used to define the solution nodes. Numerical experiments are provided that verify this behavior.

Keywords: Summation-by-parts, Collocation, Stable, High-Order.

1. Introduction

One of my research interests is the optimization of systems governed by computational fluid dynamics (CFD). Such CFD-constrained optimization problems arise in aerodynamic design, flow control, and inverse problems. These are demanding applications for CFD because they require discretizations that are efficient, robust, and geometrically flexible. High-order discretizations with the summation-by-parts (SBP) property [1] offer one way to satisfy these requirements [2, 3].

Despite being well-suited to CFD-constrained optimization, multidimensional SBP operators can be difficult to use. This difficulty stems from the non-uniqueness of SBP operators and, consequently, a scarcity of closed-form definitions that have favorable spectral properties.

The complicated implementation of multi-dimensional SBP operators likely contributes to the slow adoption of SBP operators in simplex-based codes. Unlike tensor-product operators, which are defined by a concise one-dimensional operator, multi-dimensional SBP operators require several dense matrices for each polynomial degree P. Furthermore, these matrices are rarely defined by explicit formulae; for example, the skew-symmetric operators in References [4] and [5] were defined as the minimum-norm solution of accuracy constraints, while the operators in Reference [6] were found by solving various optimization problems. If a practitioner wants to use these multi-dimensional SBP operators, they must either (i) hope a library is available in their chosen programming language, (ii) painstakingly transcribe the operators (accounting for node and face ordering), or (iii) implement the linear solves or optimizations that produce the operators from first principles. None of these options is attractive.

Thus, there is a need for a straightforward, explicit formula for multidimensional SBP operators. This formula should produce an operator with good spectral properties; ideally, the spectrum should be independent of the quadrature rule that the operator is based on. In this paper, I show that the SBP operator introduced by Chan [7] satisfies these requirements. His operator projects the collocation-based solution onto a modal basis, which is why I will refer to the operator as a modal collocation (MC) operator.

The focus of [7] was entropy-stable discontinuous Galerkin discretizations, where Chan used the MC operator in an intermediate step at the quadrature nodes to leverage two-point entropy conservative flux functions [8]. However, MC operators have received little attention as a high-order difference operator on their own. To the best of my knowledge, the only direct application of an MC operator was by Montoya and Zingg [9], who used it as a benchmark against their tensor-product modal SBP operators.

Therefore, the goal of the paper is to highlight the MC operator and characterize the spectrum of MC-based discretizations.

The following is a brief outline of the remaining paper. Section 2 begins by reviewing the definition of the MC operator, and then proceeds to the main theoretical result regarding the spectrum of MC-based semi-discretizations. To support the theory, I provide some numerical experiments in Section 3 of one- and two-dimensional discretizations of initial-boundary-value problems. Conclusions are provided in Section 4.

2. Methodology and Analysis

2.1 Modal Collocation Operators

I begin by defining the modal-collocation operator.

Definition 1 (First-derivative modal-collocation operator) *Let $X_\Omega = \{\boldsymbol{x}_i\}_{i=1}^N$ and $W_\Omega = \{w_i\}_{i=1}^N$ denote the nodes and weights, respectively, for a quadrature rule over the domain Ω that is exact for degree $2P$ polynomials. In addition, let $\mathsf{V} \in \mathbb{R}^{N\times N_P}$ be a matrix whose columns are functions from a degree P orthonormal polynomial basis evaluated at X_Ω, and let $\mathsf{V}_{x_d} \in \mathbb{R}^{N\times N_P}$ hold the derivatives of the basis with respect to the coordinate x_d also evaluated at X_Ω. Then the matrix*

$$\mathsf{D}_{x_d} = \mathsf{V}_{x_d}\mathsf{V}^T\mathsf{W},$$

is a degree P modal-collocation operator corresponding to the first derivative $\partial/\partial x_d$, $d = 1, 2, \ldots, D$, at the quadrature nodes X_Ω, where $\mathsf{W} = \mathsf{diag}(w_1, w_2, \ldots, w_N)$ is a diagonal matrix whose diagonal entries are the quadrature weights in W_Ω.

This MC operator definition uses an orthonormal basis to simplify the subsequent analysis. However, this is not necessary; indeed, Chan [7] used a more general projection operation defined as show below.

$$\mathsf{D}_{x_d} = \tilde{\mathsf{V}}_{x_d}(\tilde{\mathsf{V}}^T\mathsf{W}\tilde{\mathsf{V}})^{-1}\tilde{\mathsf{V}}^T\mathsf{W},$$

where $\tilde{\mathsf{V}}$ is a generic basis for polynomials of total degree P evaluated at the quadrature nodes. It is straightforward to show that the above definition transforms the generic basis into an orthonormal basis, so Definition 1 is a special case; see [10].

An MC operator that uses a positive quadrature rule (i.e. positive weights) is also a degree P diagonal-norm summation-by-parts operator [7]. In particular, we have the decomposition

$$\mathsf{W}\mathsf{D}_{x_d} = \mathsf{W}\mathsf{V}_{x_d}\mathsf{V}^T\mathsf{W} = \mathsf{S}_{x_d} + \frac{1}{2}\mathsf{E}_{x_d},$$

where S_{x_d} is skew-symmetric and E_{x_d} is symmetric. Furthermore, the symmetric matrix acts on the boundary and can be written [7]

$$\mathsf{E}_x = \mathsf{R}_\Gamma^T\mathsf{W}_\Gamma\mathsf{R}_\Gamma,$$

where $\mathsf{W}_\Gamma = \mathsf{diag}(w_1, w_2, \ldots, w_M)$ holds the weights of a degree $2P$ exact quadrature over the boundary $\Gamma = \partial\Omega$. In the above definition of E_x, the projection/interpolation matrix

$$\mathsf{R}_\Gamma = \mathsf{V}_\Gamma\mathsf{V}^T\mathsf{W}$$

where V_Γ holds the orthonormal basis evaluated at the boundary quadrature nodes $X_\Gamma = \{\boldsymbol{x}_j\}_{j=1}^M$.

We consider a one-dimensional discretization in the theory presented below. In the one-dimensional setting we write the MC operator as $\mathsf{D}_x = \mathsf{V}_x\mathsf{V}^T\mathsf{W}$ and it acts on the interval domain $\Omega = [x_L, x_R]$.

Furthermore, in one dimension the boundary quadrature becomes trivial — $X_\Gamma = \{x_L, x_R\}$ and $\mathsf{W}_\Gamma = \mathsf{diag}(1,1)$ — and the operator E_x simplifies to

$$\mathsf{E}_x = \mathsf{R}_R^T \mathsf{R}_R - \mathsf{R}_L^T \mathsf{R}_L,$$

where

$$\mathsf{R}_L = \mathsf{V}_L \mathsf{V}^T \mathsf{W}, \qquad \text{where} \qquad \mathsf{R}_R = \mathsf{V}_R \mathsf{V}^T \mathsf{W}.$$

The matrices $\mathsf{V}_L \in \mathbb{R}^{1\times N_P}$ and $\mathsf{V}_R \in \mathbb{R}^{1\times N_P}$ represent the orthonormal polynomial basis evaluated at x_L and x_R, respectively.

2.2 Quadrature Independence of MC Semi-discretization Spectrum

In this section, I show that an MC semi-discretization of the periodic, one-dimensional advection equation has an eigenvalue spectrum that is effectively independent of the underlying quadrature. Specifically, the non-trivial eigenvalues in the spectrum depend only on the domain/element shape and the chosen polynomial degree P. Elsewhere I show that this result extends to multidimensional, non-periodic problems [10].

The constant-coefficient linear advection equation with periodic boundary conditions is given by,

$$\begin{aligned} &\frac{\partial u}{\partial t} + \alpha \frac{\partial u}{\partial x} = 0, && \forall x \in \Omega = [x_L, x_R], t \in [0,T], \\ &u(x_L,t) = u(x_R,t), && \forall t \in [0,T], \end{aligned} \tag{1}$$

where $\alpha > 0$ is the advection speed. The standard SBP semi-discretization of (1) using one element/block and upwind simultaneous approximation terms (SATs) is given by [3, 2]

$$\frac{d\boldsymbol{u}}{dt} + \alpha \mathsf{D}_x \boldsymbol{u} + \alpha \mathsf{W}^{-1} \mathsf{R}_L^T \left(\mathsf{R}_L \boldsymbol{u} - \mathsf{R}_R \boldsymbol{u}\right) = \boldsymbol{0}, \tag{2}$$

where $\boldsymbol{u}(t) \in \mathbb{R}^N$ is the solution. The semi-discretization can be written concisely as

$$\frac{d\boldsymbol{u}}{dt} = \mathsf{A}\boldsymbol{u}, \qquad \text{where} \qquad \mathsf{A} = -\alpha \mathsf{D}_x - \alpha \mathsf{W}^{-1} \mathsf{R}_L^T \left(\mathsf{R}_L - \mathsf{R}_R\right). \tag{2}$$

I can now state and prove the main result.

Theorem 1 *Suppose the semi-discretization (2) uses a degree P modal-collocation operator D_x. If $\sigma(\mathsf{A}) = \{\lambda_i\}_{i=1}^N$ denotes the spectrum of A, then $\sigma(\mathsf{A})$ is the union of $N_Z = N - N_P$ zero eigenvalues and N_P eigenvalues that are independent of the quadrature.*

Proof: Let $\mathsf{Z} \in \mathbb{R}^{N\times(N-N_P)}$ be a matrix in the null-space of $\mathsf{V}^T\mathsf{W}$, i.e. $\mathsf{V}^T\mathsf{W}\mathsf{Z} = \mathsf{0}$. Also, without loss of generality, assume that Z is orthonormal with respect to W, so $\mathsf{Z}^T\mathsf{W}\mathsf{Z} = \mathsf{I}$. Next, decompose the solution into components corresponding to V and Z,

$$\boldsymbol{u}(t) = \mathsf{V}\boldsymbol{u}_V(t) + \mathsf{Z}\boldsymbol{u}_Z(t),$$

and substitute this decomposition into the ODE (2);

$$\frac{d}{dt}\left(\mathsf{V}\boldsymbol{u}_V + \mathsf{Z}\boldsymbol{u}_Z\right) = \mathsf{A}\left(\mathsf{V}\boldsymbol{u}_V + \mathsf{Z}\boldsymbol{u}_Z\right).$$

Next, left multiply the ODE above by $\left[\mathsf{WV}, \mathsf{WZ}\right]^T$ to obtain

$$\frac{d}{dt}\begin{bmatrix}\boldsymbol{u}_V\\ \boldsymbol{u}_Z\end{bmatrix} = \begin{bmatrix}\mathsf{V}^T\mathsf{W}\\ \mathsf{Z}^T\mathsf{W}\end{bmatrix}\mathsf{A}\begin{bmatrix}\mathsf{V} & \mathsf{Z}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_V\\ \boldsymbol{u}_Z\end{bmatrix} = \underbrace{\begin{bmatrix}\mathsf{V}^T\mathsf{WAV} & \mathsf{V}^T\mathsf{WAZ}\\ \mathsf{Z}^T\mathsf{WAV} & \mathsf{Z}^T\mathsf{WAZ}\end{bmatrix}}_{\equiv \mathsf{B}}\begin{bmatrix}\boldsymbol{u}_V\\ \boldsymbol{u}_Z\end{bmatrix}. \tag{3}$$

Note that the matrix B defined above is similar to the matrix A, since the matrix left multiplying A in Equation (3) is the inverse of the matrix right multiplying A. Consequently, A and B share the same eigenvalues; it will prove easier to characterize the eigenvalues of B.

To simplify the matrix B in the ODE (3), we first note that Z is in the null-space of A:

$$\begin{aligned}\mathsf{AZ} &= \big(-\alpha\mathsf{D}_x - \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_L + \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_R\big)\mathsf{Z}\\ &= -\alpha\mathsf{V}_x\mathsf{V}^T\mathsf{WZ} - \alpha\mathsf{V}\mathsf{V}_L^T\mathsf{V}_L\mathsf{V}^T\mathsf{WZ} + \alpha\mathsf{V}\mathsf{V}_L^T\mathsf{V}_R\mathsf{V}^T\mathsf{WZ} = \mathsf{0}\end{aligned}$$

where we used $\mathsf{V}^T\mathsf{WZ} = \mathsf{0}$. Thus, the submatrices $\mathsf{V}^T\mathsf{WAZ}$ and $\mathsf{Z}^T\mathsf{WAZ}$ in the matrix B are matrices of zeros. Next, consider the submatrix $\mathsf{Z}^T\mathsf{WAV}$ in B:

$$\begin{aligned}\mathsf{Z}^T\mathsf{WAV} &= \mathsf{Z}^T\mathsf{W}\big(-\alpha\mathsf{D}_x - \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_L + \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_R\big)\mathsf{V}\\ &= -\alpha\big(\mathsf{Z}^T\mathsf{WV}_x\mathsf{V}^T\mathsf{W} + \mathsf{Z}^T\mathsf{WV}\mathsf{V}_L^T\mathsf{V}_L\mathsf{V}^T\mathsf{W} - \mathsf{Z}^T\mathsf{WV}\mathsf{V}_L^T\mathsf{V}_R\mathsf{V}^T\mathsf{W}\big)\mathsf{V}\\ &= \mathsf{0},\end{aligned}$$

where we used $\mathsf{Z}^T\mathsf{WV} = \mathsf{0}$ and the fact that the derivatives of the orthonormal basis can be expressed in terms of the basis, so $\mathsf{Z}^T\mathsf{WV}_x = \mathsf{0}$ as well.

To summarize, we have shown that the matrix B in the ODE (3) is equivalent to

$$\mathsf{B} = \begin{bmatrix}\mathsf{V}^T\mathsf{WAV} & \mathsf{0}\\ \mathsf{0} & \mathsf{0}\end{bmatrix}.$$

Therefore, the eigenvalues of B are the union of $N_Z = N - N_P$ repeated zero eigenvalues and the eigenvalues of the submatrix $\mathsf{V}^T\mathsf{WAV}$. This submatrix simplifies because of the exactness of the quadrature. To see this, let $\mathcal{V}_i(x)$ and $\mathcal{V}_j(x)$, with $i, j \in \{1, 2, \ldots, N_P\}$, be two functions from the orthonormal basis. Then

$$\begin{aligned}\big[\mathsf{V}^T\mathsf{WAV}\big]_{ij} &= \big[\mathsf{V}^T\mathsf{W}\big(-\alpha\mathsf{D}_x - \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_L + \alpha\mathsf{W}^{-1}\mathsf{R}_L^T\mathsf{R}_R\big)\mathsf{V}\big]_{ij}\\ &= \big[-\alpha\mathsf{V}^T\mathsf{WV}_x - \alpha\mathsf{V}_L^T\mathsf{V}_L + \alpha\mathsf{V}_L^T\mathsf{V}_R\big]_{ij}, && \text{(Using } \mathsf{V}^T\mathsf{WV} = \mathsf{I}\text{)}\\ &= -\alpha\sum_{k=1}^{N}\mathcal{V}_i(x_k)w_k\frac{\partial\mathcal{V}_j}{\partial x}(x_k) - \alpha\mathcal{V}_i(x_L)\mathcal{V}_j(x_L) + \alpha\mathcal{V}_i(x_L)\mathcal{V}_j(x_R)\\ &= -\alpha\int_\Omega\mathcal{V}_i\frac{\partial\mathcal{V}_j}{\partial x}\,d\Omega + \alpha\int_\Gamma\mathcal{V}_i\hat{\mathcal{V}}_j n_x\,d\Gamma, && \text{(quadrature accuracy)}\end{aligned}$$

where the numerical flux function $\hat{\mathcal{V}}_j$ is given by

$$\hat{\mathcal{V}}_j(x) = \begin{cases}\mathcal{V}_j(x), & x = x_L,\\ \mathcal{V}_j(x), & x = x_R.\end{cases}$$

Thus, we have shown that the entries in the submatrix $\mathsf{V}^T\mathsf{WAV}$ are independent of the quadrature. Consequently, its eigenvalues are also independent of the quadrature, which completes the proof. □

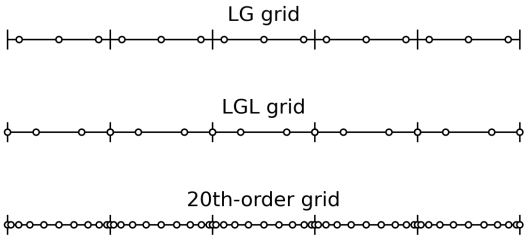

Figure 1: Example collocation nodes for $P = 2$ on a five-element grid considered in the constant-coefficient advection study.

2.3 Generalization

Constant-coefficient, periodic advection is a simple problem — does Theorem 1 generalize? For example, can we use a central-flux instead of an upwind flux in (2)? What about non-periodic boundary conditions and multi-element grids? Most importantly, what about two- and three-dimensional problems?

The result generalizes to non-periodic, constant-coefficient advection regardless of the numerical flux, number of dimensions, and the shape of the elements [10]. That is, if MC operators are used to discretize the advection problem in the usual way[1], then the spectrum of the semi-discretization has $N_Z = N - N_P$ zero eigenvalues and N_P eigenvalues that are independent of the quadrature. The result generalizes in this case because the MC semi-discretization continues to project the collocation degrees of freedom onto the polynomial basis — which removes dependence on modes in the span of Z — and the resulting residual remains in the span of the polynomial basis over each element.

For non-constant or nonlinear hyperbolic problems, the MC spectrum will depend on the quadrature. However, in this case, one can show that the non-trivial part of the spectrum is identical to that obtained using a discontinuous Galerkin semi-discretization based on the same quadrature and numerical fluxes [10].

3. Results

3.1 One-Dimensional Linear Advection Studies

I begin with the the one-dimensional constant-coefficient advection in order to verify Theorem 1. I also use this problem to demonstrate that MC-based discretizations behave as expected in terms of accuracy and stability. Thus, consider Problem (1) with $\alpha = 1$, $[x_L, x_R] = [0, 1]$, and a Gaussian-bump initial condition:

$$u(x, 0) = \exp\Big(-\tfrac{1}{2}\big((x - \tfrac{1}{2})/(2/25)\big)^2\Big), \qquad \forall x \in [0, 1].$$

I consider three different quadratures rules to demonstrate that the quadrature rule does not impact the spectrum of the semi-discretization. These quadratures are listed below for a given polynomial degree P. Figure 1 illustrates the collocation nodes for $P = 2$ on a uniform grid with five elements.

LG: This choice corresponds to a $(P + 1)$-point Legendre-Gauss (LG) quadrature rule, which is exact for polynomials of degree $Q = 2P + 1$.

LGL: This choice corresponds to a $(P + 2)$-point Legendre-Gauss-Lobatto (LGL) rule, which is also exact for polynomials of degree $Q = 2P + 1$.

20th-order: This choice corresponds to an 11-point LGL rule, which is exact for polynomials of degree $Q = 19$.

[1] That is, following a standard SBP-SAT discretization.

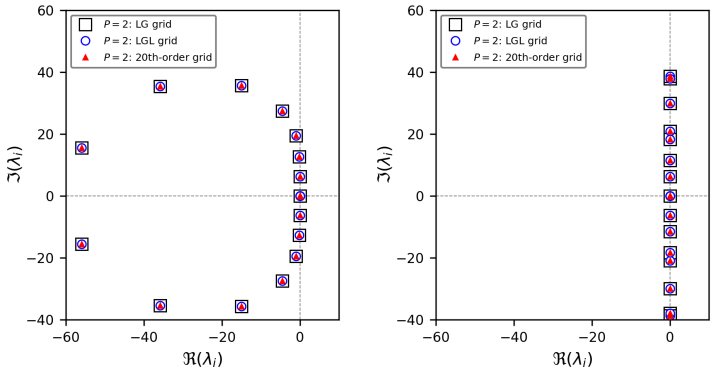

Figure 2: Spectra of $P = 2$ discretizations corresponding to operators using the LG, LGL, and 20th-order quadratures with upwind (left figure) and symmetric (right figure) SATs.

On a generic element $k \in \{1, 2, \ldots, K\}$, the SBP-SAT semi-discretization of the PDE in (1) is

$$\frac{d\boldsymbol{u}_k}{dt} = -\mathsf{D}_x \boldsymbol{u}_k - \boldsymbol{s}_k, \qquad \forall\, k \in \{1, 2, \ldots, K\}, t \in [0, T], \tag{4}$$

where $\boldsymbol{u}_k(t) \in \mathbb{R}^N$ is the solution on element k, $\mathsf{D}_x \boldsymbol{u}_k$ is the MC discretization of the spatial derivative, and $\boldsymbol{s}_k$ denotes the simultaneous approximation terms (SATs) on node k.

I consider both upwind SATs and symmetric SATs, which are given by the following expressions on element k.

$$\boldsymbol{s}_k = \mathsf{R}_L^T\big(\mathsf{R}_L \boldsymbol{u}_k - \mathsf{R}_R \boldsymbol{u}_{k-1}\big), \tag{upwind}$$

$$\boldsymbol{s}_k = \frac{1}{2}\mathsf{R}_L^T\big(\mathsf{R}_L \boldsymbol{u}_k - \mathsf{R}_R \boldsymbol{u}_{k-1}\big) - \frac{1}{2}\mathsf{R}_R^T\big(\mathsf{R}_R \boldsymbol{u}_k - \mathsf{R}_L \boldsymbol{u}_{k+1}\big). \tag{symmetric}$$

The above expressions introduce a small abuse of notation: when $k = 1$ (the first element) the index $k - 1$ should be replaced with K, and when $k = K$ (the last element) the index $k + 1$ should be replaced with 1.

When necessary, I numerically solve the system of ODEs (4) using a high-order adaptive time-marching scheme with absolute and relative tolerances set to 10^{-13} and 10^{-11}, respectively. Specifically, I use the `Vern7()` method as implemented in the Julia `DifferentialEquations.jl` package [11].

3.1.1 Spectra

Figure 2 shows spectra for the semi-discretization (4) using $P = 2$ and five elements. The left plot shows the spectra for upwind SATs and the right plot shows the spectra for symmetric SATs. The three spectra shown on each plot correspond to the three choices of quadrature points. The LG scheme has the fewest degrees of freedom and therefore the fewest eigenvalues. The LGL and 20th-order schemes have a subset of eigenvalues that match the LG scheme's eigenvalues to machine precision, consistent with Theorem 1. The remaining LGL and 20th-order eigenvalues are zero and correspond to the space orthogonal to the degree $P = 2$ basis.

3.1.2 Accuracy

Next I verify the accuracy of the semi-discretization (4). While the three quadrature rules (LG, LGL, and 20th-order) share the same non-trivial eigenvalues, their semi-discretizations will produce different errors because of projection errors; the initial condition and exact solution are non-polynomial, so the different quadrature rules produce slightly different basis coefficients when the functions are projected.

We can eliminate the differences in projection errors by using a common quadrature rule to project the initial condition and exact solution. Here we use the LG quadrature as the common rule. To clarify this approach, let V_{LG} denote the degree P polynomial evaluated at the LG nodes on one element and let W_{LG}

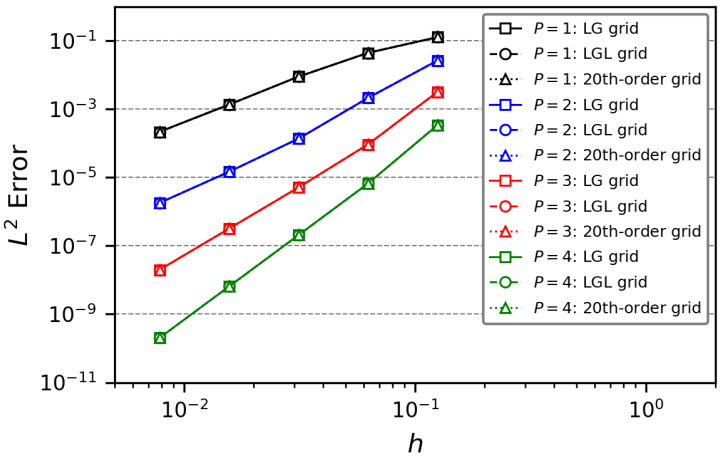

Figure 3: L^2 error in the linear advection solution at $t = 1$ for each quadrature type over a sequence of meshes and polynomial degrees P.

denote the diagonal matrix of LG weights. Then the initial condition on element k is given by

$$\boldsymbol{u}_k(0) = \mathsf{V}\mathsf{V}_{\text{LG}}^T \mathsf{W}_{\text{LG}} \boldsymbol{u}_{0,k,\text{LG}}, \qquad k \in \{1, 2, \ldots, K\},$$

where $\boldsymbol{u}_{0,k,\text{LG}}$ is the initial condition at the LG nodes of element k. A similar projection is used for the exact solution when computing the error.

Figure 3 plots the L^2 error in the numerical solution at time $t = T = 1$. The error is approximated using the global quadrature rule induced by a scheme's diagonal mass matrix and the LG-projected exact solution. The mesh spacing shown in the figure corresponds to a sequence of five grids with $K \in \{8, 16, 32, 64, 128\}$ elements. Furthermore, I considered $P \in \{1, 2, 3, 4\}$ for the polynomial exactness of the operators.

In general, all the discretizations exhibit an error convergence rate of $P+1$ based on the finest two meshes. An exception to this are the $P = 1$ schemes, which have convergence rates closer to 2.5. More relevant to the present work, all three quadrature rules produce the same error to machine precision. Note that when a common LG-projection is not applied to the initial condition and exact solution, the LGL and 20th-order quadratures produce errors that are 1.6 and 1.3 times larger, respectively, than the LG quadrature.

3.1.3 Energy Stability

I conclude the one-dimensional results with a verification of energy stability. To help make the presentation more concise, I only present the results corresponding to the $P = 4$ discretizations on five elements. The results for $P = 1$, 2, and 3 are qualitatively similar.

The left plot in Figure 4 shows the change in energy, $e(t) - e(0)$ where $e(t) = \boldsymbol{u}(t)^T \mathsf{M} \boldsymbol{u}(t)$, for the energy conservative versions of (4) that use symmetric SATs. The results confirm that the energy is indeed conserved up to machine precision in this case, regardless of the quadrature scheme used. The differences in the change in energy between the three quadrature rules can be explained by the adaptive `Vern6()` time-marching scheme.

The right plot in Figure 4 displays the change in energy versus time t when the discretization uses upwind SATs. As expected, all three quadrature schemes yield semi-discretizations with non-increasing energy.

3.2 2D Example on a Circular Domain

For the second set of results, I consider the linear advection equation on a circular domain.

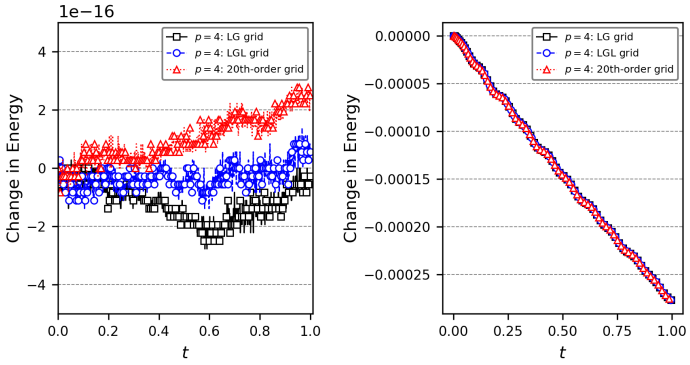

Figure 4: Change in total energy over $t \in [0, 1]$ for the energy-conservative (left figure) and energy-dissipative (right figure) $p = 4$ schemes.

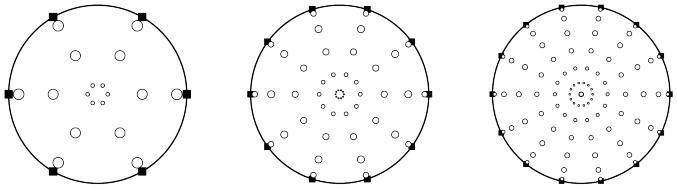

Figure 5: Nodes for degree $Q = 4$ (left), $Q = 8$ (center), and $Q = 12$ (right) quadratures on the circle. The white circles are the solution nodes and the black squares are the boundary quadrature nodes. The size of a node is proportional to the magnitude of its quadrature weight.

3.2.1 Operator Construction

I use a tensor-product grid in polar coordinates to construct degree Q quadrature rules for the unit circle. The rules have LG nodes in the radial direction and uniformly distributed nodes in the azimuthal direction. The LG rules use $Q/2 + 1$ nodes[2], so they are $Q + 1$ exact in the radial direction; the additional degree of accuracy is needed for the Jacobian of the mapping (i.e. r). There are $Q + 2$ nodes in the azimuthal direction, which is sufficient to integrate trigonometric polynomials of degree $Q + 1$. The additional degree of accuracy in the azimuthal direction is used to align the volume and boundary quadrature nodes; the boundary quadrature needs to be $Q + 1$ exact to account for the normal vector $\boldsymbol{n} = [\cos(\theta), \sin(\theta)]^T$. Figure 5 illustrates the volume and boundary nodes corresponding to $q = 4$, $q = 8$, and $q = 12$ quadrature rules.

The MC operators are constructed following Definition 1 with (normalized) Zernike polynomials [12] providing the orthonormal basis.

For comparison, I also consider SBP operators constructed using Equation (9) from [13]. I will refer to these operators as min-norm SBP operators, because their skew-symmetric matrices are equivalent to solving the accuracy conditions in a minimum-norm sense. The E_{x_d} matrices are constructed using extrapolation along the collinear nodes to produce sparse boundary operators.

3.2.2 IBVP and Discretization

I use a linear advection equation with circular streamlines to study the eigenvalues and accuracy of the MC and SBP operators for circular domains. The details of the initial-boundary-value problem are listed below.

$$\begin{aligned} \frac{\partial u}{\partial t} &= -\sum_{d=1}^{2} \frac{\partial}{\partial x_d}(\alpha_d u), && \forall \boldsymbol{x} \in \Omega, t \in [0, 2\pi], \\ u(\boldsymbol{x}, t)\alpha_n(\boldsymbol{x}) &= 0, && \forall \boldsymbol{x} \in \Gamma, t \in [0, 2\pi], \\ u(\boldsymbol{x}, 0) &= \exp(x_1 x_2), && \forall \boldsymbol{x} \in \Omega, \end{aligned} \tag{5}$$

[2] Q is always taken to be even, so $Q/2$ is an integer.

where Ω is the unit disc, $\alpha_n(\boldsymbol{x}) = \boldsymbol{\alpha}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x})$ is the normal component of the advection velocity, and the (divergence-free) advection velocity is given by

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \begin{bmatrix} \alpha_1(\boldsymbol{x}) \\ \alpha_2(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} -x_2\sqrt{x_1^2 + x_2^2} \\ x_1\sqrt{x_1^2 + x_2^2}. \end{bmatrix} = \begin{bmatrix} -\sin(\theta)r^2 \\ \cos(\theta)r^2 \end{bmatrix}.$$

The advection velocity is parallel to the disc's boundary, so the normal component $\alpha_n(\boldsymbol{x}) = 0$ and the boundary condition is satisfied.

The MC discretization of (5) is

$$\begin{aligned} \frac{d\boldsymbol{u}}{dt} &= \underbrace{-\sum_{d=1}^{2} \left(\mathsf{D}_d \, \Lambda_d - \mathsf{W}^{-1}\mathsf{E}_d \, \Lambda_d\right)\boldsymbol{u}}_{\equiv \mathsf{A}\boldsymbol{u}}, && \forall\, t \in [0, 2\pi] \\ \left[\boldsymbol{u}(0)\right]_i &= \exp\left((x_1)_i (x_2)_i\right), && \forall\, i = 1, 2, \ldots, N, \end{aligned} \tag{6}$$

where the diagonal matrix

$$\Lambda_d = \mathsf{diag}\left(\alpha_d(\boldsymbol{x}_1), \alpha_d(\boldsymbol{x}_2), \ldots, \alpha_d(\boldsymbol{x}_N)\right)$$

holds the dth component of the advection velocity evaluated at the nodes. The terms $\mathsf{W}^{-1}\mathsf{E}_d \, \Lambda_\mathsf{d}\boldsymbol{u}$ are penalty terms for the boundary condition. The SBP discretization of (5) is identical to (6) with the MC operators replaced with SBP operators.

3.2.3 Spectra

Figures 6–9 plot the eigenvalues of the matrix A arising from different basis degree P and quadrature degree Q in the semi-discretization (6). Figures 6 and 7 show the eigenvalues corresponding to $P = 6$ discretizations using MC and min-norm SBP operators, respectively. The non-trivial MC eigenvalues are invariant with respect to the three quadrature degrees $Q = 12$, $Q = 24$, and $Q = 36$. By contrast, the min-norm SBP eigenvalues change with the quadrature.

Figures 8 and 9 are analogous plots for $P = 10$ discretizations. Again, the non-trivial MC eigenvalues are invariant, whereas the SBP eigenvalues are not. In fact, the SBP discretization using the $Q = 60$ quadrature has an eigenvalue with a positive real part. Note that this instability can be eliminated by using a skew-symmetric discretization, but it is notable that the MC discretization does not require this strategy in this case.

Table 1 summarizes the maximum real parts and spectral radii of the spectra plotted in Figures 6–9. All of the discretizations produce stable matrices with the exception of the min-norm SBP discretization with $P = 10$ and $Q = 60$, as previously noted. In addition, as expected from Theorem 1, the MC discretizations have spectral radii that are independent of the quadrature degree. By contrast, the min-norm SBP discretizations have spectral radii that grow with Q.

3.2.4 Accuracy

I conclude the results for the circular-flow problem by investigating the accuracy of the MC and min-norm SBP discretizations. For this study, I solve the semi-discretization (6) using the `Vern9()` method implemented in `DifferentialEquations.jl` with absolute and relative error tolerances of 10^{-12} and 10^{-10}, respectively. Unlike the one-dimensional results, I do not use a common quadrature projection here.

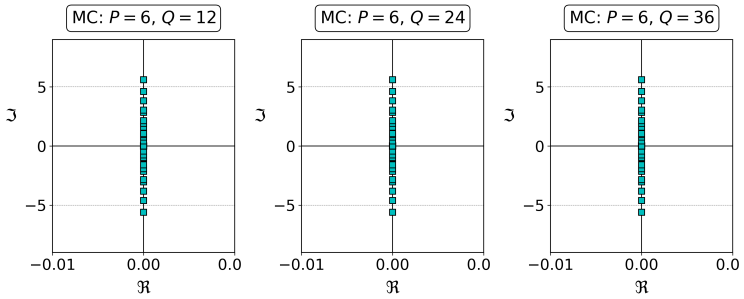

Figure 6: Eigenvalues of the MC-based semi-discretization of the circular-flow problem using $P = 6$ and various quadratures.

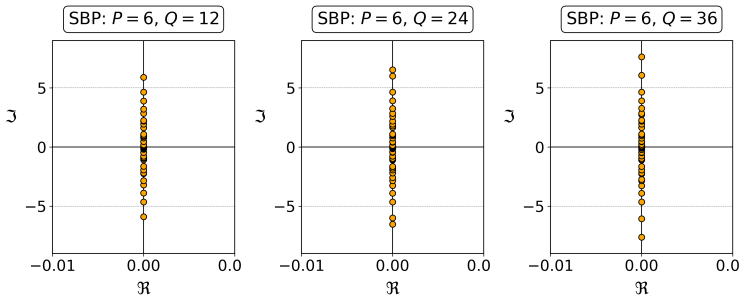

Figure 7: Eigenvalues of the min-norm-SBP-based semi-discretization of the circular-flow problem using $P = 6$ and various quadratures.

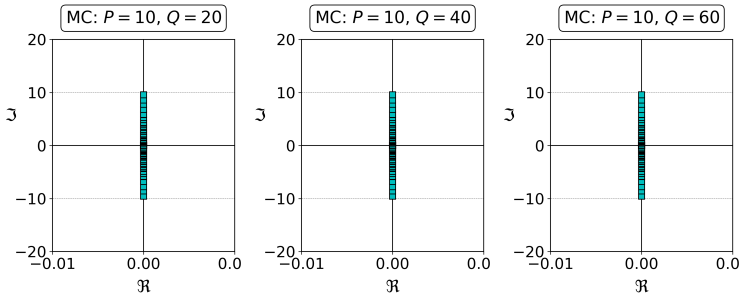

Figure 8: Eigenvalues of the MC-based semi-discretization of the circular-flow problem using $P = 10$ and various quadratures.

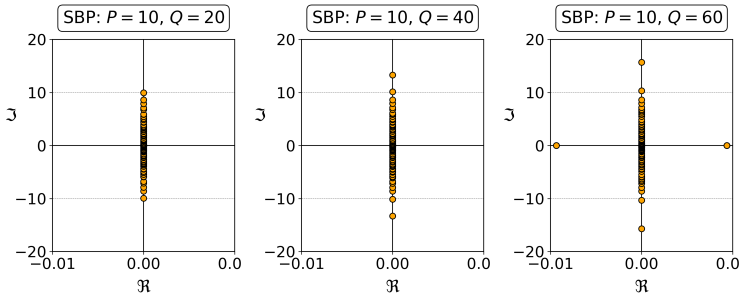

Figure 9: Eigenvalues of the min-norm-SBP-based semi-discretization of the circular-flow problem using $P = 10$ and various quadratures.

Table 1: Maximum real parts and spectral radii for the circular-flow problem using different degree P MC and SBP operators.

deg. $P = 6$

	max real part		**spectral radius**	
quad. Q	**MC**	**SBP**	**MC**	**SBP**
12	8.2678e-16	1.2479e-15	5.6	5.8987
24	1.077e-15	2.2829e-14	5.6	6.5336
36	9.5198e-16	3.5682e-14	5.6	7.6309

deg. $P = 10$

	max real part		**spectral radius**	
quad. Q	**MC**	**SBP**	**MC**	**SBP**
20	2.1778e-15	3.4132e-15	9.5652	9.9487
40	1.8609e-15	6.0315e-14	9.5652	13.278
60	2.1188e-15	0.0093463	9.5652	15.713

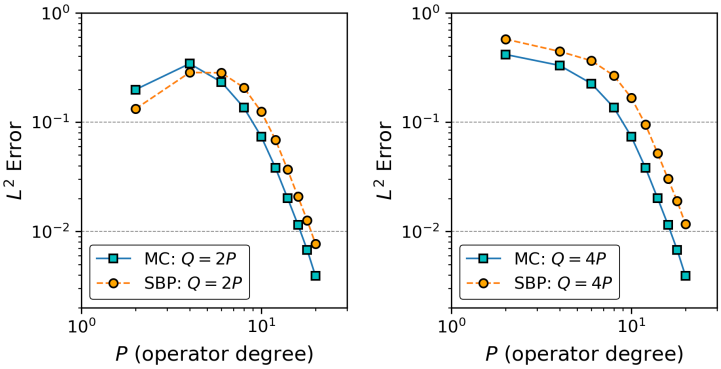

Figure 10: L^2 Error versus degree P for MC and min-norm SBP discretizations of the circular-flow problem. The left figure plots the error when the quadrature is $Q = 2P$ exact and the right figure plots the error when it is $Q = 4P$ exact.

Figure 10 shows the L^2 error versus polynomial degree for MC and min-norm SBP solutions to the circular-flow problem. The left figure plots errors based on quadratures with $Q = 2P$, and the right figure plots errors based on $Q = 4P$. Results are shown for all even degrees between $P = 2$ and $P = 20$.

A couple characteristics of the errors in Figure 10 are worth highlighting. First, the min-norm SBP errors depend on both P and Q. Second, the MC errors are independent of the quadrature Q for polynomial degrees $P > 2$. The MC discretization using $P = 2$ depends on the quadrature degree because the advection field is not a degree two polynomial.

4. Conclusions

Defining and using multidimensional SBP operators can be cumbersome, because each quadrature rule requires bespoke operators that are defined as dense matrices, in general. To address the need for a more straightforward implementation, this work investigated the modal-collocation (MC) operator proposed by Chan [7]. The MC operator is defined by an explicit formula that is straightforward to implement on arbitrary-shaped elements. Furthermore, because it projects the collocation solution onto the polynomial basis, an MC operator produces semi-discretizations whose spectral properties are essentially independent of the underlying quadrature rule.

I presented results applying the MC operator to the discretization of the constant-coefficient advection equation in one dimension, and a circular advection problem in two dimensions. The results confirmed that

the MC-based semi-discretizations are independent of the underlying quadrature accuracy for $Q \geq 2P$.

Overall, the theory and results suggest that MC operators are an attractive option for spectral-element discretizations. They seem particularly well suited to non-tensor product elements.

References

[1] H. O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In Carl de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations*, New York, 1974. Mathematics Research Center, the University of Wisconsin, Academic Press.

[2] Magnus Svärd and Jan Nordström. Review of summation-by-parts schemes for initial-boundary-value problems. *Journal of Computational Physics*, 268:17–38, 2014.

[3] David C. Del Rey Fernández, Jason E. Hicken, and David W. Zingg. Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations. *Computers & Fluids*, 95:171–196, May 2014.

[4] Jason E. Hicken, David C. Del Rey Fernández, and David W. Zingg. Multidimensional summation-by-parts operators: General theory and application to simplex elements. *SIAM Journal on Scientific Computing*, 38(4):A1935–A1958, July 2016.

[5] David C. Del Rey Fernández, Jason E. Hicken, and David W. Zingg. Simultaneous approximation terms for multi-dimensional summation-by-parts operators. *Journal of Scientific Computing*, 75:83–110, August 2017.

[6] André L. Marchildon and David W. Zingg. Optimization of multidimensional diagonal-norm summation-by-parts operators on simplices. *Journal of Computational Physics*, 411:109380, 2020.

[7] Jesse Chan. On discretely entropy conservative and entropy stable discontinuous Galerkin methods. *Journal of Computational Physics*, 362:346–374, June 2018.

[8] Eitan Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12:451–512, May 2003.

[9] Tristan Montoya and David W. Zingg. Efficient tensor-product spectral-element operators with the summation-by-parts property on curved triangles and tetrahedra. *SIAM Journal on Scientific Computing*, 46(4):A2270–A2297, 2024.

[10] Jason Hicken. A collocation scheme that is equivalent to discontinuous galerkin discretizations, 2026. https://arxiv.org/abs/2605.27327.

[11] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1), 2017.

[12] Frits von Zernike. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica*, 1(7):689–704, 1934.

[13] Jason E Hicken, Sharanjeet Kaur, and Ge Yan. Constructing stable, high-order finite-difference operators on point clouds over complex geometries. *Journal of Computational Physics*, 532:113940, 2025.