21st AIAA Computational Fluid Dynamics Conference, June 25, 2013, San Diego, CA

High-Order Compact-Stencil Summation-By-Parts Operators for the Compressible Navier-Stokes Equations

D. C. Del Rey Fernández* and D. W. Zingg†

University of Toronto Institute for Aerospace Studies, Toronto, Ontario, M3H 5T6, Canada

A general framework is presented for deriving minimum-stencil high-order summation-by-parts finite-difference operators for the second derivative with variable coefficients, for orders of accuracy $3^{\mathbf{rd}}$ through $6^{\mathbf{th}}$. These operators can be used to construct time-stable numerical schemes with simultaneous approximation terms to weakly impose boundary conditions. The derivation of these operators leads to various free parameters which can be used for optimization of the operator about criteria such as spectral radius and truncation error. The operators are $2p$ accurate on the interior, where the interior stencil has $2p+1$ nodes, but p accurate at the boundaries. Nonetheless, for purely parabolic problems, they can be shown to be $p+2$ globally accurate. However, for the Navier-Stokes equations, the continuity equation renders the method $p+1$ accurate. We present a novel means of circumventing this degradation in accuracy by using $p+2$ globally accurate operators for the continuity equation, and we prove that the new discretization remains amenable to the energy method, a necessary condition to prove time stability. Numerical tests on the one-dimensional linear convection-diffusion equation and the one- and three-dimensional Navier-Stokes equations using the method of manufactured solutions are used for verification and characterization studies. We show that for the Navier-Stokes equations using a $p+2$ globally accurate first derivative for the continuity equation substantially increases the accuracy benefits of the minimum-stencil operator.

I. Introduction

Despite Moore's law and the advent of massively parallel computers, solution of industrially relevant PDEs remains a time-intensive effort. As an example, the direct-numerical simulation of turbulent flow over a commercial aircraft would require approximately 10^{18} flops of computational power,[1] and thus computational efficiency remains one of the primary concerns of CFD practitioners. In the early 1970s, Kreiss and Oliger[2] and Swartz and Wendroff[3] 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 Δx is the mesh spacing. Thus, for a given accuracy, HO methods require coarser mesh spacing relative to lower-order methods. HO methods have been shown to be more computationally efficient than lower-order methods; some examples are the linear advection equation[4] and the compressible Navier-Stokes (NS) equations.[5,6] In the present paper, we examine discretizations based on HO summation-by-parts (SBP) finite-difference operators,[7–13] with simultaneous approximation terms (SATs) for boundary and interface treatment,[14–23] which have been successfully applied to various problems including the linear advection diffusion equation,[17] electromagnetic wave propagation[24] and the compressible Euler and NS equations.[18,20,23,25–27]

HO finite-difference methods trade mathematical complexity for computational efficiency, where the main

*PhD Candidate

†Professor and Director, Tier 1 Canada Research Chair in Computational Aerodynamics and Environmentally Friendly Aircraft Design, J. Armand Bombardier Foundation Chair in Aerospace Flight, Associate Fellow AIAA.

difficulty in their implementation arises from the boundary treatment. SBP operators provide a systematic means of deriving HO finite-difference operators with suitably accurate HO boundary treatments that are time stable.[12] In conjunction with SATs to weakly impose boundary and interface conditions, SBP operators naturally give rise 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, i.e. there is no need for halo nodes. In curvilinear coordinates, time stability can only be proven for diagonal-norm SBP operators;[28] thus we limit ourselves to those operators. The advantages of diagonal-norm SBP operators come at the cost of reduced accuracy; the operators have interior schemes that are $2p$ accurate with boundary accuracy that is p accurate and therefore the global order of accuracy is $p+1$.[29] However, Mattsson and Nordström[12] have shown that for parabolic PDEs, utilization of SBP operators for the second derivative garners an additional degree of accuracy, so the method is formally $p+2$ order accurate for minimum-stencil operators (interior scheme with the same number of nodes as the minimum-stencil first derivative) and $p+1$ accurate for wide-stencil operators (application of the first derivative twice). Furthermore, Hicken and Zingg have shown that if the formulation is dual consistent,[30] then functionals converge with the order of accuracy of the interior scheme.[31] Our research deals with aerodynamic optimization; we therefore need a robust and efficient algorithm to solve the compressible NS equations on complex geometries; the SBP-SAT method has proven to be robust and efficient in both direct computations of the NS equations and in the context of RANS, see Hicken and Zingg,[32] Osusky and Zingg,[33] and Osusky et al.[34]

Minimum-stencil operators have several numerical advantages over wide-stencil operators,[12] wide-stencil referring to the application of the first derivative twice: they have lower global error and are more dissipative of high wavenumber modes. Moreover, they have a smaller bandwidth and thus require less computational resources, particularly if one is interested in adjoint-based optimization for which the Jacobian is constructed. Finally, although one can use the wide-stencil operators, doing so with SBP operators results in the loss of one order of accuracy in the context of parabolic PDEs.

Beyond a certain order of accuracy, SBP operators contain free parameters that can be used for optimizing their behaviour. General solutions to SBP operators for the first derivative have been well studied and have been optimized by other authors (see Diener et al.[10]). Less well understood are SBP operators for the second derivative with constant or variable coefficients. Although particular instances for the constant coefficient case[11,12] and, recently, the variable coefficient case[13] have been presented, general solutions have not been derived. Besides being of academic interest, general solutions to SBP operators for the second derivative provide the ability to optimize the operators, thereby maximizing their efficiency. Therefore, the primary goal of this paper is to present a framework for deriving general solutions to SBP operators for the second derivative. We then show how to optimize these general solutions to produce operators with optimal characteristics.

For parabolic problems, SBP operators for the second derivative have the advantage of gaining an additional order of accuracy. However, our interest is in the solution of the compressible NS equations and we lose this additional order since the NS equations are incompletely parabolic by virtue of the continuity equation. The second goal of this paper is to present a novel means of circumventing this problem by using a first derivative for the continuity equation that is one order of accuracy higher than used for the remaining equations. In order to be a feasible solution, we must show two things: first, that we do not lose the SBP property and second, that we gain in terms of accuracy.

II. Spatial Discretization

Here we introduce SBP finite-difference operators for the first derivative and the second derivative with constant or variable coefficients. SBP operators for the first derivative were derived by Kreiss and Scherer,[7] refined by Strand[8] and applied by various authors (see Mattsson et al.,[12] Hicken and Zingg,[32] Diener et al.[10] and, Dias and Zingg[25]). SBP operators are finite-difference schemes, using centred differences on the interior of the domain, that do not include boundary conditions; these must be applied by some other means, in our case using SATs, see Section III. Our interest is in the solution of the compressible NS equations over complex geometries for the purpose of aerodynamic shape optimization. Structured meshes are used to

capture these geometries, so the NS equations are transformed to curvilinear coordinates. As a result, the discussion of SBP operators is limited to those with a diagonal norm, as these are provably time stable in curvilinear coordinates.[28] Given that the solution of the NS equations is the main interest, we require an SBP approximation to $\partial_x(\beta\partial_x Q)$ that is conservative, has the SBP property and is compatible with the first derivative, such that it can be used to prove time stability for the linearized NS equations.[12] To make explicit the means of deriving the proposed operators, we present detailed derivations of the required derivatives that are 4^{th}-order in the interior of the domain as examples.

A. Notation and definitions

We follow the conventions laid out by Hicken and Zingg.[31] SBP difference operators are generically defined on a uniformly spaced grid of N points on the domain $[0,1]$ and thus the grid spacing is $h = 1/(N-1)$. Typically the domain we are interested in is not $[0,1]$, but we assume that for the set of problems of interest a sufficiently differentiable invertible transformation exists from the domain of interest to $[0,1]$.

Capital letters with a script type are used to denote functions on a specified domain, and so $\mathcal{U}(x) \in C^p[a,b]$ denotes that the function, $\mathcal{U}(x)$, is a p times differentiable function on $[a,b]$. Roman letters in bold font denote the restriction $\mathcal{U}(x)$ onto an N-point grid of corresponding continuous functions, as an example $\mathbf{u} \in \mathbb{R}^N$ means $\mathbf{u} = [\mathcal{U}(x_1),\mathcal{U}(x_2),..,\mathcal{U}(x_N)]^T$. In discussing the imposition of boundary conditions using SATs, we refer to the unit vectors $\mathbf{e}_L, \mathbf{e}_R \in \mathbb{R}^N$, which are

$$\mathbf{e}_L = [1,0,...,0]^T, \mathbf{e}_R = [0,..,0,1]^T.$$

The operators for the first and second derivatives have different orders of accuracy on the interior, at the boundary, and globally. In order to differentiate between operators and the various orders of accuracy we follow the convention that we append a superscript to operators for the various orders of accuracy and a subscript to denote which derivative we are approximating. For example, $D_{i,e}^{(a,b,c)}$, denotes the operator for the i^{th} derivative with interior order of accuracy of a, boundary closure accuracy of b and results in a solution with global order of accuracy c, while the additional subscript e is to differentiate amongst various versions of the operator. In some cases we will not be interested in one or several of the orders of accuracy and will insert colons; as an example; $D_3^{(2,:,:)}$ denotes an approximation to the third derivative which is second-order on the interior where we specify neither the accuracy of the operator at the boundary nor the global order of accuracy. For the second derivative with variable coefficients we will make explicit the dependence on the variable coefficients, β, by denoting these operators as $D_2^{(a,b,c)}(B)$, where B is a diagonal matrix with the variable coefficients along its diagonal.

B. First derivative

The SBP property of the first derivative is mimetic of $\int_a^b \mathcal{Q}\partial_x\mathcal{Q}dx$, which leads to the following definition:

SBP Diagonal Norm First Derivative The matrix $D \in \mathbb{R}^{N\times N}$ is an SBP operator for the first derivative if it approximates the first derivative and is of the form $D = H^{-1}\Theta$, where $H \in \mathbb{R}^{N\times N}$, is a positive-definite diagonal matrix, called the norm, and $\Theta \in \mathbb{R}^{N\times N}$ has the property $\Theta + \Theta^T = \text{diag}(-1,0,...,0,1) = E$.

To understand how the SBP property is a discrete analogue of integration by parts we examine the continuous case. First we define the inner product of two real valued-functions, $\mathcal{U},\mathcal{V} \in [0,1]$ by $(\mathcal{U},\mathcal{V}) = \int_0^1 \mathcal{U}\mathcal{V}\text{d}x$, and hence the norm $\sqrt{(\mathcal{U},\mathcal{U})} = \|\mathcal{U}\|_2 = \sqrt{\int_0^1 \mathcal{U}^2\text{d}x}$.

Now consider the linear convection equation,

$$\frac{\partial\mathcal{Q}}{\partial t} = -\frac{\partial\mathcal{Q}}{\partial x}. \tag{1}$$

The energy method involves multiplying (1) by $\mathcal{Q}$ and integrating in space, i.e. taking the inner product of the PDE with respect to the solution $\mathcal{Q}$. We get

$$\frac{\text{d}\|\mathcal{Q}\|_2^2}{\text{d}t} = -\left.\mathcal{Q}^2\right|_0^1, \tag{2}$$

and we can see that stability of the equations depends solely on the boundary values of $\mathcal{Q}$.

Ignoring the boundary conditions, the semi-discrete form of the linear convection equation is

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = -D\mathbf{q} = -H^{-1}\Theta\mathbf{q}. \tag{3}$$

We similarly define a discrete inner product and norm as $(\mathbf{u},\mathbf{v})_H = \mathbf{u}^T H\mathbf{v}$ and $\|\mathbf{u}\|_H = \sqrt{\mathbf{u}^T H\mathbf{u}}$. Taking the discrete inner product of (3), i.e. multiplying through by $\mathbf{q}^T H$, and adding the transpose, based on the properties of H and Θ, gives

$$\frac{\mathrm{d}\|\mathbf{q}\|_H^2}{\mathrm{d}t} = -\left\{[\mathbf{q}(N)]^2 - [\mathbf{q}(1)]^2\right\}. \tag{4}$$

As in the continuous case, (4) only depends on boundary values, and we can see that the SBP property is mimetic, in the discrete case, of integration by parts and thus allows for a discrete energy method.

The SBP property also automatically ensures that the first derivative is conservative. To see this we consider the continuous case:

$$\int_0^1 \frac{\partial \mathcal{U}}{\partial x}\mathrm{d}x = \mathcal{U}|_0^1. \tag{5}$$

The discrete norm matrix H represents a high-order quadrature of order $2p$,[35] such that

$$\int_0^1 \mathcal{V}\mathcal{U}\mathrm{d}x \approx \mathbf{v}^T H\mathbf{u},$$

and the discrete analogue of (5) is

$$\mathbf{1}^T HD\mathbf{u} = \mathbf{1}^T\Theta\mathbf{u}.$$

However, the SBP property $\Theta + \Theta^T = E$ gives us that $\Theta = E - \Theta^T$, and we have that

$$\mathbf{1}^T\Theta\mathbf{u} = \mathbf{1}^T\left(E - \Theta^T\right)\mathbf{u}.$$

On the other hand we know that for consistency the discrete derivative of a constant vector must be zero, therefore we have $D\mathbf{1} = 0$, which implies that $\Theta\mathbf{1} = 0$, and we conclude that $\mathbf{1}^T\Theta^T = 0$, and we have that

$$\int_0^1 \frac{\partial \mathcal{U}}{\partial x}\mathrm{d}x \approx \mathbf{1}^T HD\mathbf{u} = \mathbf{1}^T E\mathbf{u} = \mathbf{u}(N) - \mathbf{u}(1),$$

and the first-derivative operator is discretely conservative.

SBP operators for the first derivative have been well studied.[8,10,15,17,19,22,36,37] Here we present a brief account of how to derive them since we construct the SBP operator for the second derivative with variable coefficients by applying the first derivative twice and adding a corrective term that makes the operator have an interior stencil that spans $2p+1$ nodes. We are interested in operators approximating the first derivative that on the interior are centred-difference approximations, satisfy the SBP property, and have the minimum number of nodes for a given order of accuracy, i.e. for interior order of accuracy $2p$ the stencil has $2p+1$ nodes. The interior stencil is known, so what remains is to derive the stencils for boundary nodes.

To summarize what is known about $D = H^{-1}\Theta$:

1. D is $2p$ accurate at interior points and p accurate at $2p$ points at the left and right boundaries;[8]
2. $\Theta^T + \Theta = \mathrm{diag}(-1,0,0,...,0,0,1)$, implying that Θ is nearly skew-symmetric with $\Theta(1,1) = -\frac{1}{2}$, $\Theta(N,N) = \frac{1}{2}$, and the remaining diagonal entries are 0;
3. H is diagonal, positive definite and $H(j,j) = H(N-(j-1), N-(j-1))$ for $j \in [1,2p]$ with the remaining diagonal entries equal to h.

We use $D_1^{(4,2,3)}$ as an example, i.e. $p = 2$. As it is 4^{th} order in the interior, it has a 5-point interior stencil, and 4 boundary stencils need to be derived. The global order of accuracy of the solution is $p + 1 = 3$. The interior stencil is given as $(\frac{1}{12}, -\frac{2}{3}, 0, \frac{2}{3}, -\frac{1}{12},)$, giving

$$D = H^{-1}\Theta =$$

$$\frac{1}{h}\begin{bmatrix} \frac{1}{H_{11}} & & & & & & \\ & \frac{1}{H_{22}} & & & & & \\ & & \frac{1}{H_{33}} & & & & \\ & & & \frac{1}{H_{44}} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \end{bmatrix} \times$$

$$\begin{bmatrix} -\frac{1}{2} & \theta_{12} & \theta_{13} & \theta_{14} & & & & \\ -\theta_{12} & 0 & \theta_{23} & \theta_{24} & & & & \\ -\theta_{13} & -\theta_{23} & 0 & \theta_{34} & -\frac{1}{12} & & & \\ -\theta_{14} & -\theta_{24} & -\theta_{34} & 0 & \frac{2}{3} & -\frac{1}{12} & & \\ 0 & 0 & \frac{1}{12} & -\frac{2}{3} & 0 & \frac{2}{3} & -\frac{1}{12} & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

where h is the grid spacing. Note that the above form is based upon the minimum boundary stencil width, as derived by Strand,[8] to satisfy order $2p$ accuracy on the interior and order p accuracy at $2p$ points. The $D_1^{(4,2,3)}$ operator gives the following approximations at the boundary nodes:

$$\text{Node 1: } \frac{1}{H_{11}\times h}\left(-\tfrac{1}{2}q_j + \theta_{12}q_{j+1} + \theta_{13}q_{j+2} + \theta_{14}q_{j+3}\right),$$

$$\text{Node 2: } \frac{1}{H_{22}\times h}\left(-\theta_{12}q_{j-1} + \theta_{23}q_{j+1} + \theta_{24}q_{j+2}\right)$$

$$\vdots$$

The above equations must satisfy the accuracy criteria, in this case $p = 2$, i.e. second-order accuracy. Inserting Taylor series expansions of the q_{j+k} gives a system of equations; as an example the first node gives

$$-\tfrac{1}{2} + \theta_{12} + \theta_{13} + \theta_{14} = 0, \text{ (consistency)}$$

$$\theta_{12} + 2\,\theta_{13} + 3\,\theta_{14} = H_{11}, \text{ (first-order accuracy)}$$

$$\tfrac{1}{2}\,\theta_{12} + 2\,\theta_{13} + \tfrac{9}{2}\,\theta_{14} = 0, \text{ (second-order accuracy).}$$

There are $\frac{2p(2p-1)}{2} + 2p = 10$ coefficients, but there are $2p(p+1) = 12$ equations, so the system appears overdetermined. However, this does not turn out to be the case since not all of the equations are linearly independent. In fact, the following occurs: for $D_1^{(2,1,2)}$ and $D_1^{(4,2,3)}$, the operators are unique, while for $D_1^{(6,3,4)}$ and $D_1^{(8,4,5)}$, the operators have 1 and 3 free parameters, respectively, that can be used to optimize their behaviour. For interior orders beyond 8, it is not possible to derive an operator that has a positive definite H, and hence a discrete norm, with only $2p$ boundary nodes that are p order accurate (see Strand[8]). In order to surpass these limitations one could increase the number of boundary nodes (see Diener et al.[10] and Sjogreen and Yee[38]).

C. Second derivative with variable coefficients

The discrete SBP operator for the second derivative with variable coefficients results in a large system of nonlinear equations that must be solved for the boundary closures. As the order of the operator increases so too does the complexity of the system of nonlinear equations that needs to be solved and the number of free parameters that need to be specified once a solution is found.

First we define the SBP operators for the second derivative with constant or variable coefficients and give a generic structure to construct them. We define the discrete SBP operator for the second derivative with constant coefficients, $B = \text{diag}(1, \ldots, 1)$, or variable coefficients as follows:

SBP Second Derivative The matrix $D_2^{(2p,p,\hat{p})}(B) \in \mathbb{R}^{N\times N}$ is an SBP operator for the second derivative, with global order of accuracy $\hat{p}$, if it approximates the second derivative and is of the form, $D_2^{(2p,p,\hat{p})}(B) = H^{-1}\{-M + EBD_b\}$, where H is a diagonal positive-definite matrix called the norm, $E = \text{diag}(-1, 0, ..., 0, 1)$, $B = \text{diag}(\beta_0, \ldots, \beta_N)$, $D_b^{(:,\geq p+1,:)}$ is an approximation to the first derivative at the boundaries, $M = \left(D_1^{(2p,p,p+1)}\right)^T HBD_1^{(2p,p,p+1)} + R$, M and R are positive-semi-definite (PSD) and symmetric, and B is PSD.

In order to show that the proposed SBP-SAT discretization is time-stable for the linearized NS equations, the first derivative used for the inviscid portion of the equations must be the same as that used in the construction of the second derivative (the formulation is said to be consistent, see Mattsson and Nordström[11]). Before proceeding let us make clear how the given definition leads to an SBP operator and show the relationship between the minimum-stencil operator and the application of the first derivative twice, wide-stencil operator. As before we start with the continuous case. Consider the variable-coefficient diffusion equation on the interval $[0, 1]$:

$$\frac{\partial \mathcal{Q}}{\partial t} = \frac{\partial}{\partial x}\left(\beta \frac{\partial \mathcal{Q}}{\partial x}\right). \tag{6}$$

Applying the energy method to (6) gives

$$\frac{\mathrm{d}||\mathcal{Q}||^2}{\mathrm{d}t} = 2\left(\beta Q \frac{\partial \mathcal{Q}}{\partial x}\right)\Big|_0^1 - 2\int_0^1 \beta\left(\frac{\partial \mathcal{Q}}{\partial x}\right)^2 \mathrm{d}x. \tag{7}$$

Ignoring boundary conditions, the semi-discrete equations are

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = H^{-1}\{-M + EBD_b\}\,\mathbf{q}.$$

Multiplying by $\mathbf{q}^T H$ and adding the transpose of the product gives

$$\begin{aligned} \tfrac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{d}t} &= 2\mathbf{q}^T EBD_b\mathbf{q} - 2\mathbf{q}^T M\mathbf{q} \\ &= 2\mathbf{q}^T EBD_b\mathbf{q} - 2(D\mathbf{q})^T HBD\mathbf{q} - 2\mathbf{q}^T R\mathbf{q}. \end{aligned} \tag{8}$$

Let us first consider the application of the first derivative twice, in this case $R = 0$ and $D_b = D$, and we have

$$\frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{d}t} = 2\mathbf{q}^T EBD\mathbf{q} - 2(D\mathbf{q})^T HBD\mathbf{q}.$$

Since the discrete norm represents a $2p$-order quadrature we see that $(D\mathbf{q})^T HBD\mathbf{q} \approx \int_0^1 \frac{\partial \mathcal{U}}{\partial x} b \frac{\partial \mathcal{U}}{\partial x}\mathrm{d}x$ and $\mathbf{q}^T EBD\mathbf{q} \approx b\frac{\partial \mathcal{U}}{\partial x}\big|_0^1$. The difference in the minimum-stencil operator is the addition of the term $\mathbf{q}^T R\mathbf{q}$. From (8) we see that for time-stability R must be PSD, and from an accuracy standpoint it must add an error that is no larger than the global order of accuracy. In order to show that the proposed SBP-SAT discretization is time-stable for the linearized NS equations, H in the above definition must be the same norm as used with the first derivative. Thus the formulation is said to be compatible with the first derivative.[12]

To construct these operators, we need R. For the operator with $2p$ accuracy on the interior, we posit the general form:

$$R_p = \frac{1}{h}\sum_{i=p+1}^{2p} \alpha_i (\tilde{D}_{i,p}^{(2,1,:)})^T C_i^p B \tilde{D}_{i,p}^{(2,1,:)},$$

where h is the mesh spacing. The $\tilde{D}_{i,p}^{(2,1,:)}$ operators have an interior stencil with $2p+1$ nodes, while the boundary stencils have $3p$ nodes. The interior stencil is a second-order centred-difference approximation to the i^{th} derivative while the boundary stencils are first-order accurate. The tilde notation denotes an undivided difference approximation. Constructed thus, the operator is guaranteed to be PSD, as long as the C_i^p, which are diagonal matrices of the form $C_i^p = \text{diag}\left(c_{11}^p, \ldots, c_{2p2p}^p, 1, \ldots, 1, c_{2p2p}^p, \ldots, c_{11}^p\right)$, where the superscript p is to differentiate amongst the various C_i, are PSD.

The interior stencil that is compatible with the proposed construction of the second derivative is given as

$$D_{2,\text{int}}^{(2p,:,:)} = -\left\{ (D^{(2p,:,:)})^T B D^{(2p,:,:)} + \frac{1}{h^2}\sum_{p+1}^{2p} \alpha_i (\tilde{D}_{i,p}^{(2,:,:)})^T B \tilde{D}_{i,p}^{(2,:,:)} \right\}.$$

The global order of accuracy, $\hat{p}$, will depend on the particular problem being solved and can be in the range $\hat{p} \in [p+1, p+2]$. We explicitly give the general form of the SBP operators for $p \in [2,3,4]$ as

$$D_2^{(2,1,:)}(B) = H^{-1}\left\{ -\left(D_1^{(2,1,2)}\right)^T H B D_1^{(2,1,2)} - \tfrac{1}{4h}\left(\tilde{D}_2^{(2,1,:)}\right)^T C_2^1 B \tilde{D}_{2,1}^{(2,1,:)} + \tfrac{1}{h} E B \tilde{D}_{1,1}^{(:,2,:)} \right\},$$

$$\begin{aligned} D_2^{(4,2,:)}(B) = H^{-1}\Bigg\{ &-\left(D_1^{(4,2,3)}\right)^T H B D_1^{(4,2,3)} - \tfrac{1}{18h}\left(\tilde{D}_{3,2}^{(2,1,:)}\right)^T C_3^2 B \tilde{D}_{3,2}^{(2,1,:)} \\ &- \tfrac{1}{48h}\left(\tilde{D}_{4,2}^{(2,1,:)}\right)^T C_4^2 B \tilde{D}_{4,2}^{(2,1,:)} + \tfrac{1}{h} E B \tilde{D}_1^{(:,3,:)} \Bigg\}, \end{aligned}$$

$$\begin{aligned} D_2^{(6,3,:)}(B) = H^{-1}\Bigg\{ &-\left(D_1^{(6,3,4)}\right)^T H B D_1^{(6,3,4)} - \tfrac{1}{80h}\left(\tilde{D}_{4,3}^{(2,1,:)}\right)^T C_4^3 B \tilde{D}_{4,3}^{(2,1,:)} \\ &- \tfrac{1}{100h}\left(\tilde{D}_{5,3}^{(2,1,:)}\right)^T C_5^3 B \tilde{D}_{5,3}^{(2,1,:)} - \tfrac{1}{720h}\left(\tilde{D}_{6,3}^{(2,1,:)}\right)^T C_6^3 B \tilde{D}_{6,3}^{(2,1,:)} + \tfrac{1}{h} E B \tilde{D}_1^{(:,\geq 4,:)} \Bigg\}, \end{aligned} \tag{9}$$

$$\begin{aligned} D_2^{(8,4,:)}(B) = H^{-1}\Bigg\{ &-\left(D_1^{(8,4,5)}\right)^T H B D_1^{(8,4,5)} - \tfrac{1}{350h}\left(\tilde{D}_{5,4}^{(2,1,:)}\right)^T C_5^4 B \tilde{D}_{5,4}^{(2,1,:)} \\ &- \tfrac{1}{252h}\left(\tilde{D}_{6,4}^{(2,1,:)}\right)^T C_6^4 B \tilde{D}_{6,4}^{(2,1,:)} - \tfrac{1}{980h}\left(\tilde{D}_{7,4}^{(2,1,:)}\right)^T C_7^4 B \tilde{D}_{7,4}^{(2,1,:)} \\ &- \tfrac{1}{11200h}\left(\tilde{D}_{8,4}^{(2,1,:)}\right)^T C_8^4 B \tilde{D}_{8,4}^{(2,1,:)} + \tfrac{1}{h} E B \tilde{D}_1^{(:,\geq 5,:)} \Bigg\}. \end{aligned}$$

The above formulation leads to multiple solutions with many free parameters for which values must be chosen. In this paper we explore two alternatives:

- Wide-boundary-node operator: use all of the free parameters to optimize the operators about some criteria, and thus there are $3p$ boundary nodes that are p accurate, with the widest boundary stencil containing $4p$ nodes;
- Narrow-boundary-node operator: set free parameters such that the operator has $2p$ boundary nodes that are p accurate, with the widest boundary stencil containing $3p$ nodes, as in the case of the first derivative, and use the remaining free parameters to optimize the operator about some criteria.

D. Navier-Stokes equations with $p+2$ treatment of the continuity equation

In contrast to parabolic PDEs, such as the linear convection-diffusion (LCD) equation – see Section V.A, the minimum- and wide-stencil approximations to the second derivative result in nearly identical convergence

plots for the NS equations, see Section V.B, suggesting that the error from the continuity equation dominates. We propose using a $p+2$ globally accurate first-derivative SBP operator to discretize the first derivative in the continuity equation to mitigate this reduced accuracy. However, in order that the resultant discretization can be provably time stable for linear PDEs, we must show that the discretization is still amenable to the energy method. To prove that the energy method can still be applied, it is sufficient to examine a hyperbolic system of equations such as:

$$\frac{\partial \mathcal{Q}}{\partial t} = -A\frac{\partial \mathcal{Q}}{\partial x}, \text{ for } x \in [0,1], \tag{10}$$

where A is a constant symmetric matrix, i.e. $A^T = A$ (note that this condition is not a restriction since the energy method is only applicable to PDEs that can be symmetrized). Applying the energy method gives

$$\frac{\mathrm{d}||\mathcal{Q}||^2}{\mathrm{dt}} = -\mathcal{Q}^T A \mathcal{Q}\big|_0^1.$$

Let us examine what happens when we treat all the equations with the same first derivative operator. The semi-discrete form is given as

$$\frac{\mathrm{d}\mathbf{q}}{dt} = -\hat{A}D\mathbf{q},$$

where $D = D_1^{(2p,p,p+1)} \otimes I_5$, $\hat{A} = I_N \otimes A$, I_n are $n \times n$ identity matrices unless otherwise noted, $\otimes$ is the Kronecker product, and N is the number of nodes in the one-dimensional grid under consideration. Applying the energy method, i.e. multiplying (10) by $\mathbf{q}^T H$ and adding the transpose of this product, gives

$$\frac{\mathrm{d}\mathbf{q}^T}{\mathrm{dt}} H\mathbf{q} + \mathbf{q}^T H \frac{\mathrm{d}\mathbf{q}}{\mathrm{dt}} = \mathbf{q}^T \left\{ H\hat{A}D + D^T \hat{A}^T H^T \right\} \mathbf{q}, \tag{11}$$

where $H = H^{(2p,p,:)} \otimes I_5$, $H = H^T$, and $D = H^{-1}\Theta \otimes I_5$. We will make use of the following rules for Kronecker products:

mixed-product rule: $(A \otimes B)(C \otimes D) = AC \otimes BD$,

bilinearity rule: $A \otimes (B + C) = A \otimes B + A \otimes C.$,

transpose rule: $(A \otimes B)^T = A^T \otimes B^T$

For the LHS of (11) we note that the terms $\frac{\mathrm{d}\mathbf{q}^T}{\mathrm{dt}} H\mathbf{q}$, and $\mathbf{q}^T H \frac{\mathrm{d}\mathbf{q}}{\mathrm{dt}}$ are scalars and the transpose of a scalar is equal to the scalar thus

$$\frac{\mathrm{d}\mathbf{q}^T}{\mathrm{dt}} H\mathbf{q} + \mathbf{q}^T H \frac{\mathrm{d}\mathbf{q}}{\mathrm{dt}} = \frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{dt}},$$

where $||\mathbf{q}||_H^2 = \mathbf{q}^T H \mathbf{q}$ is the discrete inner product. Now by the mixed-product rule

$$H\hat{A}D = (H \otimes I_5)(I_N \otimes A)(H^{-1}\Theta \otimes I_5) = HI_N H^{-1}\Theta \otimes I_5 A I_5 = \Theta \otimes A,$$

and similarly

$$D^T \hat{A}^T H^T = (\Theta^T H^T \otimes I_5)(I_N \otimes A^T)(H^T \otimes I_5) = \Theta^T H I_N H \otimes I_5 A^T I_5 = \Theta^T \otimes A.$$

Therefore, (11) reduces to

$\frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{dt}} = \mathbf{q}^T \left\{ \Theta \otimes A + \Theta^T \otimes A \right\} \mathbf{q}$, which by the bilinearity rule becomes $= \left(\Theta + \Theta^T\right) \otimes A$, and by the properties of Θ reduces to $\mathrm{diag}(-1, 0, \ldots, 0, 1) \otimes A$

$$\frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{dt}} = -\left(\mathbf{q}^T(N) A \mathbf{q}(N) - \mathbf{q}^T(1) A \mathbf{q}(1)\right),$$

and the above is the discrete analogue of the continuous case. The question is what happens when we treat one of the equations with a higher-order derivative.

We define

$$D = D_1^{(2(p+1),p+1,p+2)} \otimes \text{diag}(1,0,0,0,0) + D_1^{(2p,p,p+1)} \otimes \text{diag}(0,1,1,1,1) = D_1^{(p+2)} \otimes \tilde{I}_1 + D_1^{(p+1)} \otimes \tilde{I}_4$$

and

$$H = H^{(2(p+1),p+1,p+2)} \otimes \text{diag}(1,0,0,0,0) + H^{(2p,p,p+1)} \otimes \text{diag}(0,1,1,1,1,1) = H^{(p+2)} \otimes \tilde{I}_1 + H^{(p+1)} \otimes \tilde{I}_4.$$

Multiplying the new semi-discrete equation by $\mathbf{q}^T H$ and adding the transpose of that product gives

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(\mathbf{q}^T H \mathbf{q}\right) = \frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{dt}} = -\mathbf{q}^T \left[H\hat{A}D + D^T \hat{A} H\right] \mathbf{q}.$$

Now

$$H\hat{A}D =$$

$$\left[H^{(p+2)} \otimes \tilde{I}_1 + H^{(p+1)} \otimes \tilde{I}_4\right] [I_N \otimes A] \times \left[\left(H^{(p+2)}\right)^{-1} \otimes \tilde{I}_1 + \left(H^{(p+1)}\right)^{-1} \otimes \tilde{I}_4\right] \times \left[\Theta^{(p+2)} \otimes \tilde{I}_1 + \Theta^{(p+1)} \otimes \tilde{I}_4\right].$$

We are concerned with the products

$$\left[H^{(p+2)} \otimes \tilde{I}_1\right] [I_N \otimes A] \left[\left(H^{(p+1)}\right)^{-1} \otimes \tilde{I}_4\right] \text{ and } \left[H^{(p+1)} \otimes \tilde{I}_4\right] [I_N \otimes A] \left[\left(H^{(p+2)}\right)^{-1} \otimes \tilde{I}_1\right].$$

Using the Kronecker product rule we have

$$\left[H^{(p+2)} \otimes I_1\right] [I_N \otimes A] \left[\left(H^{(p+1)}\right)^{-1} \otimes I_4\right] = H^{p+2} I_N \left(H^{p+1}\right)^{-1} \otimes I_1 A I_4 = H^{(p+2)} \left(H^{(p+1)}\right)^{-1} \otimes 0 = 0$$

and

$$\left[H^{(p+1)} \otimes I_4\right] [I_N \otimes A] \left[\left(H^{(p+2)}\right)^{-1} \otimes I_1\right] = H^{p+1} I_N \left(H^{p+2}\right)^{-1} \otimes I_4 A I_1 = H^{(p+1)} \left(H^{(p+2)}\right)^{-1} \otimes 0 = 0.$$

Therefore we have

$$H\hat{A}D = \Theta^{p+2} \otimes \tilde{I}_1 A + \Theta^{p+1} \otimes \tilde{I}_4 A,$$

and similarly, using the transpose rule

$$D^T \hat{A} H = (H\hat{A}D)^T = \left(\Theta^{p+2}\right)^T \otimes I_1 A + \left(\Theta^{p+1}\right)^T \otimes I_4 A.$$

Finally, using the bilinearity rule, we find that

$$\frac{\mathrm{d}||\mathbf{q}||_H^2}{\mathrm{dt}} = -\mathbf{q}^T \left[\left(\Theta^{p+2} + \left(\Theta^{p+2}\right)^T\right) \otimes \tilde{I}_1 A + \left(\Theta^{p+1} + \left(\Theta^{p+1}\right)^T\right) \otimes \tilde{I}_4 A\right] \mathbf{q} =$$

$$-\mathbf{q}^T \text{diag}(-1,0,\ldots,0,1) \otimes A\mathbf{q} =$$

$$-\left(\mathbf{q}^T(N) A \mathbf{q}(N) - \mathbf{q}^T(1) A \mathbf{q}(1)\right),$$

as desired. Therefore, we have shown that we can use a $p+2$ globally accurate first derivative for the continuity equations while retaining the ability to apply the energy method.

III. Boundary conditions

With SBP operators it is typical to use SATs to enforce boundary conditions weakly. To explain the SAT concept, consider the LCD equation with variable coefficients (see Mattsson[22] for the case with $b = 1$):

$$\frac{\partial \mathcal{Q}}{\partial t} = -a\frac{\partial \mathcal{Q}}{\partial x} + \epsilon \frac{\partial}{\partial x}\left(b\frac{\partial \mathcal{Q}}{\partial x}\right), 0 \leq x \leq 1, t \geq 0$$

$$\mathcal{Q}(0,t) + \alpha b \frac{\partial \mathcal{Q}(0,t)}{\partial x} = 0, \frac{\partial \mathcal{Q}(1,t)}{\partial x} = 0,$$

$$\mathcal{Q}(x,0) = f(x),$$

$$\frac{-2\epsilon}{a} \leq \alpha \leq 0,$$

and its semi-discrete analogue with SATs applied,

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{dt}} = -aD\mathbf{q} + \epsilon D_2(b)\mathbf{q} + SAT.$$

SATs impose the boundary conditions as penalty terms; rather than imposing the boundary term exactly, they do so within the discretization error. This method has been found to be preferable to strong enforcement of the boundary conditions. In their comparison of weakly and strongly enforced Dirichlet boundary conditions for the solution of the advection-diffusion equation and the incompressible NS equations, Bazilevs and Hughes[39] found that weakly imposed boundary conditions resulted in faster convergence to steady state. Similarly, Eliasson et al.[40] found that weakly enforced boundary conditions for the NS equations resulted in faster convergence to steady state, suggesting that the reason for this was an improved eigenspectrum for the semi-discrete equations.

The SAT term has the following form:[22]

$$\begin{aligned}\mathrm{SAT_{L/R}} = \quad & -\tfrac{H^{-1}}{2}\left(\tau_L \mathbf{e}_L^T\left[q - \mathcal{Q}(0)\mathbf{e}_L + \alpha b D_b q - b\tfrac{\partial \mathcal{Q}(0)}{\partial x}\mathbf{e}_L\right]\right), \\ & -\tfrac{H^{-1}}{2}\left(\tau_R \mathbf{e}_R^T\left[bD_b q - b\tfrac{\partial \mathcal{Q}(1)}{\partial x}\mathbf{e}_R\right]\right).\end{aligned}$$

where $\tau_L = \frac{-2\epsilon}{\alpha}$, and $\tau_R = 2\epsilon$ and L/R refers to the left and right boundaries respectively.

For the compressible NS equations we rely on the work in Svärd and Nordström,[20] and Nordström et al.[23] In the current work we deal with two boundary types, characteristic boundaries, and block interface boundaries. For the inviscid portion of the NS equations both types of boundary conditions are treated identically and the SATs take the form:[23]

$$SAT^{\mathrm{I}}_{(\xi_i,\mathrm{L})} = -\tfrac{H^{-1}}{J}A^{+}_{\xi_i}\left(\mathbf{q}(\xi_i, L)_{\mathrm{computed}} - \mathbf{q}(\xi_i, \mathrm{L})_{\mathrm{prescribed}}\right),$$

$$SAT^{\mathrm{I}}_{(\xi_i,\mathrm{R})} = \tfrac{H^{-1}}{J}A^{-}_{\xi_i}\left(\mathbf{q}(\xi_i, R)_{\mathrm{computed}} - \mathbf{q}(\xi_i, \mathrm{R})_{\mathrm{prescribed}}\right),$$

where

$$A^{\pm}_{\xi_i} = \frac{A_{\xi_i} \pm |A|_{\xi_i}}{2}, \; A_{\xi_i} \in \left[\frac{\partial \hat{E}}{\partial \hat{Q}}, \frac{\partial \hat{F}}{\partial \hat{Q}}, \frac{\partial \hat{G}}{\partial \hat{Q}}\right], \text{ and } \xi_i \in [\xi, \eta, \zeta].$$

The subscripts $(\xi_i, \mathrm{L/R})$ refer to the SAT boundary term for the left and right boundaries in direction ξ_i, the superscript I indicates that the SAT term is for the inviscid portion of the NS equations, and $\mathbf{q}_{\mathrm{computed}}$ is the computed solution and $\mathbf{q}_{\mathrm{prescribed}}$ is the known boundary condition. In addition, the matrix $|A|_{\xi_i} = X^T_{\xi_i}\Lambda_{\xi_i}X_{\xi_i}$, where X_{ξ_i} is the right eigenmatrix of A_{ξ_i}in the ξ_i direction, and Λ_{ξ_i} contains the eigenvalues along its diagonal. The matrix A_{ξ_i} is constructed using Roe averaging.

The SAT terms for the viscous portion of the NS equations take the following form:

$$SAT^{\mathrm{v}}_{\xi_i} = \frac{\sigma^{\mathrm{v}}_{\mathrm{L/R}}}{\mathrm{Re}} \left\{ \begin{array}{l} B\left(D_{b,\xi_i}\hat{\mathbf{q}}(\xi_i,\mathrm{L/R})_{\mathrm{computed}} - \frac{\partial}{\partial\xi_i}\hat{\mathbf{q}}(\xi_i,\mathrm{L/R})_{\mathrm{prescribed}}\right) \\ \\ +\Psi\frac{\lambda}{J}\left(\mathbf{q}(\xi_i,\mathrm{L/R})_{\mathrm{computed}} - \mathbf{q}(\xi_i,\mathrm{L/R})_{\mathrm{prescribed}}\right) \end{array} \right\}$$

where the superscript v denotes the SAT for the viscous portion of the NS equations , $\sigma^{\mathrm{v}}_L = 1$ and $\sigma^{\mathrm{v}}_R = -1$, while λ depends on the eigenvalues of a matrix which is constructed from the linearized and symmetrized NS equations and the order of the discretization,[23] while J is the Jacobian of the curvilinear transformation. Ψ is 0 for characteristic boundaries and 1 for block interfaces.

IV. Optimization of Operators

The free-parameters in the SBP operators for the second derivative with variable coefficients offer the opportunity to optimize the operators in some way. For optimization we must choose some metrics(s) and a context. For example we can choose the metrics of truncation error or spectral radius and the context could be the LCD equation, which is what we will do in this paper. What is unclear is whether or not the optimum found for a particular metric/context scenario translates into equivalent behaviour for the same metric but a different context, say for example the NS equations.

We use the steady LCD equation with the method of manufactured solutions (MMS) to optimize the operators. The PDE we solve is

$$-\frac{\partial\mathcal{U}}{\partial x} + \frac{\partial}{\partial x}\left(\mathcal{B}\frac{\partial\mathcal{U}}{\partial x}\right) + g = 0,$$

where g is a source term we add so that the assumed solution satisfies the PDE, and $\mathcal{U}$ and $\mathcal{B}$ are defined as:

$\mathcal{U} = \sin(x)$, and $\mathcal{B} = 1 + \epsilon\cos(x)$, where $\epsilon = 0.9$ and $x \in [\pi/3, 2\pi/3]$.

We will examine 4 operators for each order of accuracy, which will be referred to using the following:

- D: application of the first derivative twice
- L: Wide boundary node stencil operator from Mattsson,[13] available up to 6^{th} order accuracy, used as a benchmark for comparison
- W: Optimized wide boundary node stencil operator,
- M: Optimized minimum boundary node stencil operator.

For orders of accuracy beyond 2, all operators have free parameters and we optimize the operators to construct specific instances. All free parameters reside in the boundary stencil and we therefore use as an objective function the mean square boundary truncation error (T_B):

$$T_B = \sqrt{\frac{1}{N_B}\sum_{i=1}^{N_B} = T^2_{B,i}},$$

where N_B is the number of boundary nodes and $T^2_{B,i}$ is the sum of the squares of the coefficients of the leading truncation error term associated with node i. We numerically minimize the objective function using Maple's Minimize optimization subroutine. We note that increasing orders of accuracy result in increasing numbers of solutions and free parameters for the various operators.

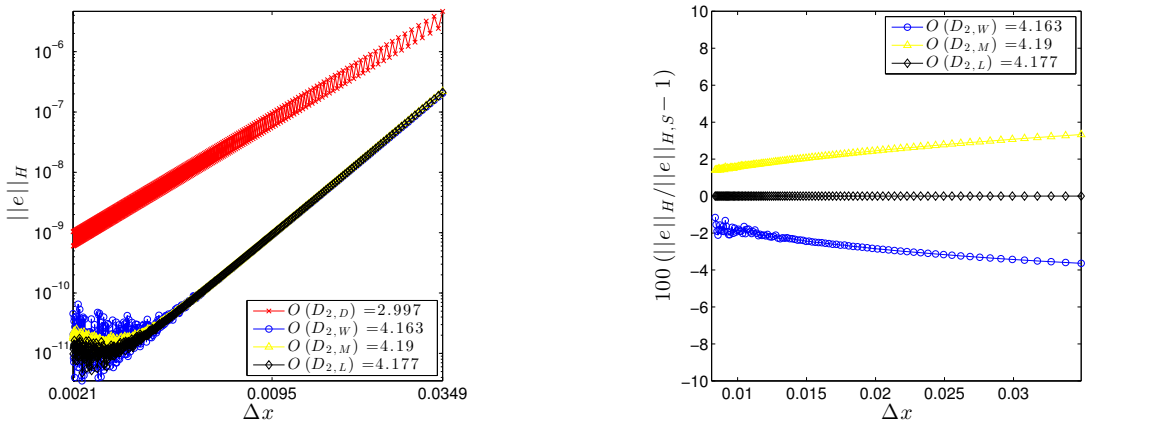
Given that we are using MMS we have access to the analytical solution and can therefore examine the error of the computed solution directly. We define the solution error as

$$\|\mathbf{e}\|_H = \sqrt{\mathbf{e}^T H \mathbf{e}},$$

where, $\mathbf{e} = (\mathbf{u} - \mathbf{u}_a)$, and $\mathbf{u}_a$ is the restriction of the analytical solution onto the grid.

Table 1. Convergence rates of $D_2(B)^{(4,:,:)}$ operators

Operator	Order
D	2.997
W	4.163
M	4.19
L	4.177



(a) $\|\mathbf{e}\|_H$

(b) Percent difference of $\|\mathbf{e}\|_H$ relative to the L operator

Figure 1. Performance of $D_2(B)^{(4,:,:)}$ operator for the LCD equation

Table 2. Convergence rates of $D_2(B)^{(6,:,:)}$ operators

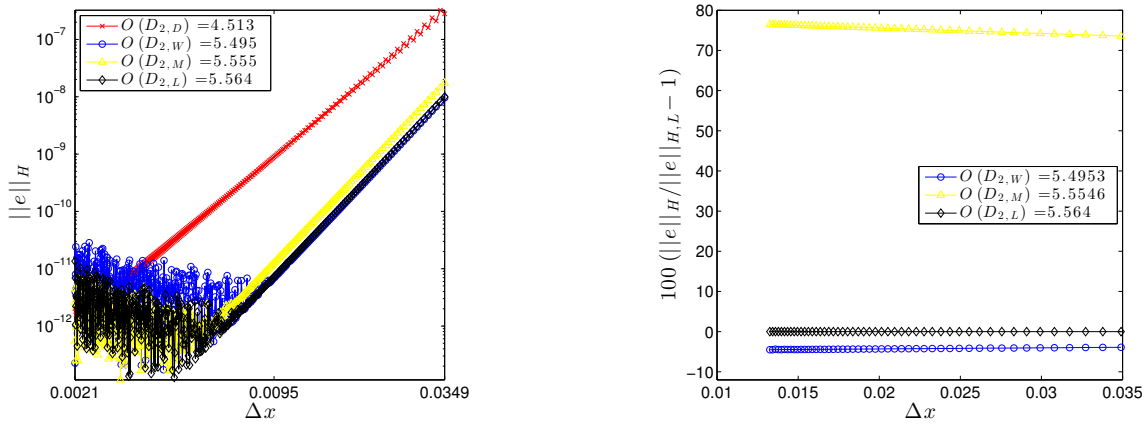
Operator	Order
D	4.513
W	5.495
M	5.555
L	5.564

V. Results

A. Linear convection-diffusion equation

In this section we solve the LCD equation subject to MMS (see Section IV). Table 1 gives the convergence rates of the various operators, while Figure 1 shows $\|\mathbf{e}\|_H$ and the percentage difference of $\|\mathbf{e}\|_H$ relative to Mattsson's operator (L). As expected, the minimum-stencil operators have a higher rate of convergence and a lower global error than the first-derivative twice operator (D). Moreover, we can see that the proposed optimized wide-boundary-stencil operator (W) is slightly more accurate than Mattsson's (L) operator, while the optimized minimum-boundary-stencil operator (M) is slightly less accurate.

For the $D_2^{(6,3,:)(B)}$ operator the results are similar to that of the $D_2(B)^{(4,:,:)}$ operator with the W operator outperforming both the M and L operators; see Table 2 and Figure 2. However, the behaviour of the M operator is a fair bit worse in this case, which is in contrast to what we observe for the three-dimensional NS equations, see Section V.B.



(a) $\|\mathbf{e}\|_H$

(b) Percent difference of $\|\mathbf{e}\|_H$ relative to the L operator

Figure 2. Performance of $D_2(B)^{(6,:,:)}$ operator for the LCD equation

B. Navier-Stokes equations

In this section we compare and contrast the various operators in the context of the three-dimensional NS equations using MMS, where we specify ρ, u, v, and p and add an appropriate source term such that the assumed solution satisfies the modified PDE. The assumed solution is constructed using sine and cosine functions; as an example ρ has form

$$\rho = r_0 + r_x \sin(a_{rx}x) + r_y \cos(a_{ry}y) + r_z \cos(a_{rz}z) + r_{xy} \cos(a_{rxy}xy) + r_{xz} \sin(a_{rxz}xz) + r_{yz} \cos(a_{ryz}yz)$$

where the terms r_j, and a_{rj} are specified coefficients. For this case we need to use artificial dissipation compatible with the SBP scheme such that we retain an energy estimate.[9,10]

Table 3 gives the convergence rates for the $D_2(B)^{(4,:,:)}$ operators and we can see that the convergence rates are very similar. However, examining Figure 3 we can see a significant difference between the various operators, with the wide-boundary-stencil operator (W) outperforming the other operators and the minimum-boundary-node stencil operator (M) performing better than Mattson's operator (L), in contrast to the LCD case, and the application of the first derivative twice (D).

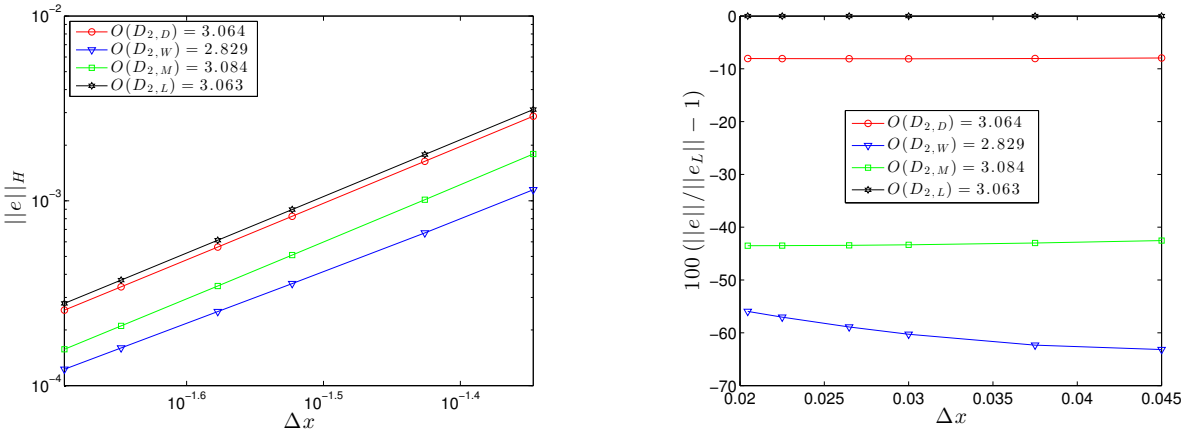
Table 4 gives the convergence rates for the $D_2(B)^{(6,:,:)}$ and again the operators have similar convergence rates. Examining Figure 4 we can see a significant difference between the various operators, with the M operator outperforming the L, and D operators. In this case we were unable to get converged solutions with the W operator for the same solver inputs, our hypothesis is that the additional degrees of freedom in D_b used to optimize the operator, which lead to large coefficients in the stencil at the first and last node, may have lead to some type of degradation, such as a large spectral radius, which did not manifest itself in the context of the LCD equation. Finally, Figure 5 compares the convergence of the $D_2(B)^{(4,:,:)}$ and $D_2(B)^{(6,:,:)}$ operators, and we see that not only do the proposed operators, M and W, have smaller error than the L operator but that the $p = 3$ operators have much smaller error than the $p = 2$ operators.

Table 3. Convergence rates of $D_2(B)^{(4,:,:)}$ operators for the three-dimensional NS equations using MMS with a Reynolds number of unity

Operator	$O(e)$
D	3.064
W	2.829
M	3.084
L	3.063

Table 4. Convergence rates of $D_2(B)^{(6,:,:)}$ operators for the three-dimensional NS equations using MMS with a Reynolds number of unity

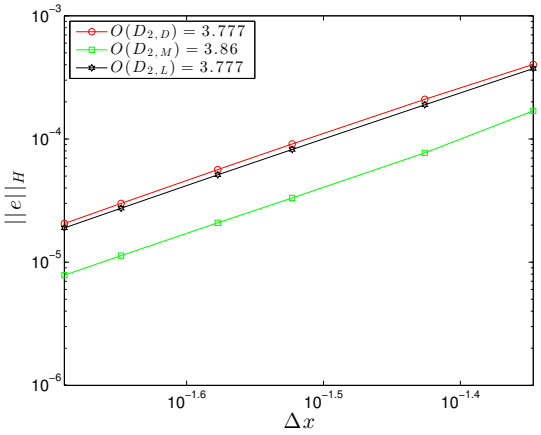
Operator	$O(e)$
D	3.777
M	3.86
L	3.777



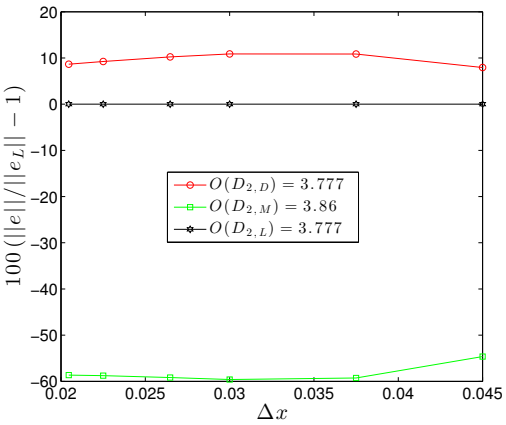
(a) $\|\mathbf{e}\|_H$

(b) Percent difference of $\|\mathbf{e}\|_H$ relative to the L operator

Figure 3. Performance of $D_2(B)^{(4,:,:)}$ operator for the 3D NS equations using MMS with a Reynolds number of unity



(a) $\|\mathbf{e}\|_H$



(b) Percent difference of $\|\mathbf{e}\|_H$ relative to the L operator

Figure 4. Performance of $D_2(B)^{(6,:,:)}$ operator for the 3D NS equations using MMS with a Reynolds number of unity

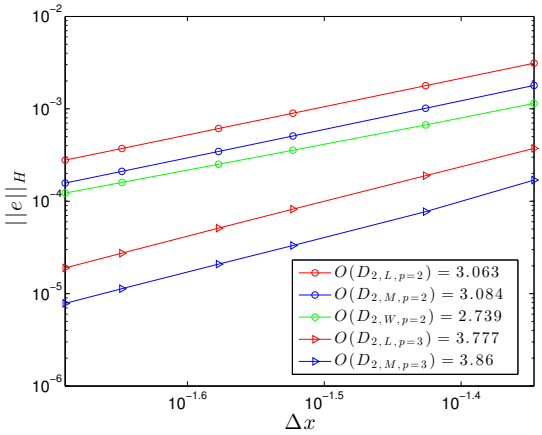


Figure 5. Performance of minimum-stencil operators $D_2(B)^{(4,:,:)}$ and $D_2(B)^{(6,:,:)}$ for the 3D NS equations using MMS with a Reynolds number of unity

VI. Higher-order $(p+2)$ treatment of the continuity equation

In the previous section we saw that although the various operators differ in their global error their rates of convergence were identical (for $p+1$ treatment of the continuity equation). In this section we examine the effects of increasing, by one order, the accuracy of the operator used to compute the first derivative in the continuity equation. In Section D we proved that we can do so while retaining an energy estimate: To keep things simple we focus our attention on the $D_2^{(4,:,:)}$ operators and solve the one-dimensional NS equations subject to an MMS solution of the form:

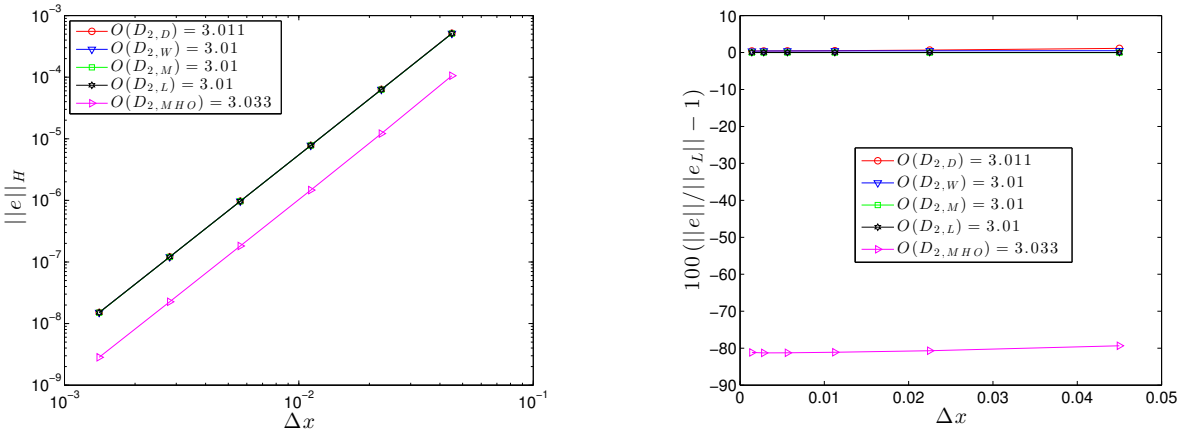
$$\rho = 2 + \cos(x),\ u = 2 + \cos(x),\ p = 2 + \cos(x)$$

The tests cases are run without adding numerical dissipation in order to examine the derivative operators in isolation.

Table 5 gives the convergence rates for the various $D_2(B)^{(4,:,:)}$ operators. The operator MHO refers to the optimized minimum-boundary-stencil operator with the continuity equation treated with a first derivative that is $p+2$ globally accurate. All operators except the MHO operator have practically identical rates of convergence, as we saw for the three-dimensional NS equations. The MHO operator does not show the expected $p+2$ convergence rate. However, as shown in Figure 6, the MHO operator does produce a significantly smaller error than the other operators, almost an order of magnitude smaller.

Table 5. Convergence rates of $D_2(B)^{(4,:,:)}$ operators one-dimensional NS equations using MMS with a Reynolds number of unity

Operator	$O(e)$
D	3.011
W	3.010
M	3.010
L	3.010
MHO	3.033



(a) $\|\mathbf{e}\|_H$

(b) Percent difference of $\|\mathbf{e}\|_H$ relative to the L operator

Figure 6. Performance of $D_2(B)^{(4,:,:)}$ operator for the 1D NS equations using MMS, no dissipation, and a Reynolds number of unity

VII. Conclusions and Future Work

The focus of this paper has been the derivation of HO minimum-stencil SBP operators for the second derivative with variable coefficients. In combination with SATs to weakly enforce boundary conditions, HO SBP operators provide an efficient means of solving PDEs discretized on structured grids. We have given a complete picture as to how to derive the various relevant operators necessary for PDEs containing first and second derivatives with variable coefficients.

The second derivative can be constructed by application of the first derivative twice or by derivation of a minimum-stencil SBP operator. The benefits of the minimum-stencil form are smaller bandwidth and better accuracy and damping characteristics. We presented our derivation for the minimum-stencil SBP approximation to the second derivative with variable coefficients. We have been able to derive the general solution to the $D_2^{(2p,p,:)}(B)$ for $p \in [1,2,3]$ operators, and specific instances of the $D_2^{(8,4,:)}(B)$ operator (not presented here) and propose a systematic means of optimizing the various SBP operators for the free variables that the operators possess.

We optimized the operators in terms of a metric on the coefficients of the leading truncation terms of the boundary stencils. We examined the behaviour of the resultant operators, using MMS, in the context of the LCD equation and the three-dimensional NS equations. We observed that the error characteristics of the operators in the context of the LCD equation did not transfer exactly to the NS equations. In particular, using our $D_2(B)^{(6,:,:)}$ wide-boundary-stencil operator, we where unable to get converged solutions for the NS equations while the operator performed well for the LCD equation. This leads us to two conclusions, first, we must look at other metrics besides the coefficients of the leading truncation term, and second, that we must examine and accordingly adjust these operators within the context of interest.

For parabolic problems, using SBP operators for the second derivative results in a global order of accuracy of $p+2$, for the minimum-stencil operator, and $p+1$, for the application of the first derivative twice, and we show that the minimum-stencil operator not only has better convergence but smaller error. However, the compressible NS equations are incompletely parabolic and we find that the minimum-stencil operator has the same error characteristics as the application of the first derivative twice. In order to retain the desirable accuracy characteristics of the minimum-stencil operator, we propose using a $p+2$ accurate first derivative for the continuity equation, while proving that the discretization is still amenable to the energy method, a prerequisite for proving time stability. Our initial results show significant improvements in accuracy.

Future work will concentrate on application of these operators to practical aerodynamic problems in order to ascertain their efficiency in that context.

References

[1]Joslin, R. D., "Discussion of DNS: Past, Present and Future." Tech. rep., Nasa Langely Research Center, 1997.

[2]Kreiss, H.-O. and Oliger, J., "Comparison of accurate methods for the integration of hyperbolic equations," *Tellus*, Vol. 24, No. 3, 1972, pp. 199–215.

[3]Swartz, B. and Wendroff, B., "The relative efficiency of finite difference and finite element methods: I Hyperbolic problems and splines," *SIAM Journal on Numerical Analysis*, Vol. 11, No. 5, 1974, pp. 979–993.

[4]Zingg, D. W., "Comparison of high-accuracy finite-difference methods for linear wave propagation," *SIAM Journal on Scientific Computing*, Vol. 22, No. 2, 2000, pp. 476–502.

[5]Zingg, D. W., De Rango, S., Nemec, M., and Pulliam, T. H., "Comparison of several spatial discretizations for the Navier-Stokes equations," *Journal of Computational Physics*, Vol. 160, No. 2, 2000, pp. 683–704.

[6]De Rango, S. and Zingg, D. W., "Higher-order spatial discretization for turbulent aerodynamic computations," *AIAA Journal*, Vol. 39, No. 7, 2001, pp. 1296–1304.

[7]Kreiss, H.-O. and Scherer, G., *Mathematical aspects of finite elements in partial differential equations*, chap. Finite element and finite difference methods for hyperbolic partial differential equations, Academic Press, New York/London, 1974.

[8]Strand, B., "Summation by parts for finite difference approximations for d/dx," *Journal of Computational Physics*, Vol. 110, No. 1, 1994, pp. 47–67.

[9]Mattsson, K., Svärd, M., and Nordström, J., "Stable and accurate artificial dissipation," *Journal of Scientific Computing*, Vol. 21, No. 1, 2004, pp. 57–79.

[10]Diener, P., D., E. N., Schnetter, E., and Tiglio, M., "Optimized high-order derivative and dissipation operators satisfying Summation by parts, and applications in three-dimensional mulit-block evolutions," *Journal of Scientific Computing*, Vol. 32, No. 1, 2007, pp. 109–145.

[11]Mattsson, K. and Nordström, J., "Summation by parts operators for finite difference approximations of second derivatives," *Journal of Computational Physics*, Vol. 199, 2004, pp. 503–540.

[12]Mattsson, K., Svard, M., and Shoeybi, M., "Stable and accurate schemes for the compressible Navier–Stokes equations," *Journal of Computational Physics*, Vol. 227, No. 4, 2008, pp. 2293–2316.

[13]Mattsson, K., "Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients," *Journal of Scientific Computing*, Vol. 51, No. 3, 2011, pp. 650–682.

[14]Funaro, D. and Gottlieb, D., "A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations," *Mathematics of Computation*, Vol. 51, No. 184, 1988, pp. 599–613.

[15]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.

[16]Hesthaven, J. S., "A stable penalty method for the compressible Navier-Stokes equations: III multidimensional domain decomposition schemes," *SIAM Journal on Scientific Computing*, Vol. 20, No. 1, 1988.

[17]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.

[18]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.

[19]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.

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

[21]Svärd, M. and Nordström, J., "A stable high-order finite difference scheme for the compressible Navier–Stokes equations: No-slip wall boundary conditions," *Journal of Computational Physics*, Vol. 227, No. 10, 2008, pp. 4805–4824.

[22]Mattsson, K., "Boundary Procedures for Summation-by-Parts Operators," *Journal of Scientific Computing*, Vol. 1, No. 133-153, 2003.

[23]Nordström, J., Gong, J., van der Weide, E., and Svärd, M., "A stable and conservative high order multi-block method for the compressible Navier-Stokes equations," *Journal of Computational Physics*, Vol. 228, No. 24, 2009, pp. 9020–9035.

[24]Nordström, J. and Gustafsson, R., "High order finite difference approximations of electromagnetic wave propagation close to material discontinuities," *Journal of Scientific Computing*, Vol. 18, No. 2, 2003, pp. 215–234.

[25]Dias, S. C. and Zingg, D. W., "A high-order parallel Newton-Krylov flow solver for the Euler equations," *AIAA paper 2009-3657*, 2009.

[26]Del Rey Fernández, D. C. and Zingg, D. W., "High-order compact-stencil summation-by-parts operators for the second derivative with variable coefficients," *ICCFD7-2803*, 2012.

[27]Boom, P. D. and Zingg, D. W., "Time-accurate flow simulations using an efficient Newton-Krylov-Schur approach with high-order temporal and spatial discretization," *AIAA paper 2013-0383*, 2013.

[28]Svärd, M., "On coordinate transformations for summation-by-parts operators," *Journal of Scientific Computing*, Vol. 20, No. 1, 2004, pp. 29–42.

[29]Gustafsson, B., "The convergence rate for difference approximations to general mixed initial boundary value problems," *SIAM Journal on Numerical Analysis*, Vol. 18, No. 2, 1981, pp. 179–190.

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

[31]Hicken, J. E. and Zingg, D. W., "Superconvergent functional estimates from summation-by-parts finite-difference discretizations." *SIAM Journal on Scientific Computing*, Vol. 33, No. 2, 2011, pp. 893–922.

[32]Hicken, J. E. and Zingg, D. W., "A parallel Newton-Krylov solver for the Euler equations discretized using simultaneous approximation terms," *AIAA Journal*, Vol. 46, No. 11, 2008, pp. 2773–2786.

[33]Osusky, M. and Zingg, D. W., "A Parallel Newton-Krylov-Schur flow solver for the Reynolds-averaged Navier-Stokes equations," *AIAA Paper 2012-0442*, 2012.

[34]Osusky, M., Boom, P. D., Del Rey Fernández, D. C., and Zingg, D. W., "An efficient Newton-Krylov-Shur parallel solution algorithm for the steady and unsteady Navier-Stokes equations," *ICCFD7*, 2012.

[35]Hicken, J. E. and Zingg, D. W., "Summation-by-parts operators and high-order quadrature," *Journal of Computational and Applied Mathematics*, Vol. 237, 2013, pp. 111–125.

[36]Olsson, P., "Summation by parts, projections, and stability. I," *Mathematics of Computation*, Vol. 64, No. 211, 1995, pp. 1035–1065.

[37]Gerritsen, M. and Olsson, P., "Designing an efficient solution strategy for fluid flows: 1. A stable high order finite difference scheme and sharp shock resolution for the Euler equations," *Journal of Computational Physics*, Vol. 129, No. 2, 1996, pp. 245–262.

[38]Sjogreen, B. and Yee, H. C., "On Tenth Order Central Spatial Schemes," *Fifth International Symposium on Turbulence and Shear Flow Phenomena*, 2007.

[39]Bazilevs, Y. and Hughes, T. J. R., "Weak imposition of Dirichlet boundary conditions in fluid mechanics," *Computers & Fluids*, Vol. 36, No. 1, 2007, pp. 12–26.

[40]Eliasson, P., Eriksson, S., and Nordström, J., "The influence of weak and strong solid wall boundary conditions on the convergence to steady-state of the Navier-Stokes equations," *AIAA Paper 2009-3551*, 2009.