



Stable and conservative boundary treatment for difference methods, with application to cut-cell discretizations

Nek Sharan^{*}, Peter T. Brady[†] and Daniel Livescu[‡]
Los Alamos National Laboratory, Los Alamos, NM, 87544

Provably time-stable finite-difference schemes that apply boundary conditions strongly (or exactly) are presented for hyperbolic systems. The proof of stability is constructed using the energy method. Sufficient conditions for stability and conservation are derived for scalar hyperbolic equation and coupled system of hyperbolic equations. Boundary stencils and norms that satisfy the sufficient conditions are derived for the centered second- and fourth-order interior stencils. A framework to further derive higher-order stencils is provided. The discretization uses non-square derivative operator to allow energy and conservation statements in terms of solution values at grid points excluding the boundary point where physical boundary condition is applied. The approach for strong boundary conditions on uniform grid is then applied to cut-cell grid configurations to derive cut-cell boundary stencils. The derived stencils do not have a small-cell problem and can be easily implemented in two- and three-dimensions following a dimensionally split discretization. Various linear and non-linear numerical tests that verify the accuracy and stability of the method are presented.

I. Introduction

Fluid-flow simulations for practical applications, involving wall boundaries or finite-domain inflow/outflow boundaries, require stable boundary treatment to allow long-time calculations typical of turbulent flows. High-order centered finite-difference schemes are commonly used for high-fidelity turbulent flow simulations and wave propagation problems because of their non-dissipative property, ease of implementation, and computational efficiency. However, the non-dissipative character of centered schemes also makes them prone to numerical instabilities when the boundary stencil, in combination with the interior stencil, is unstable.

Various numerical stability definitions exist that bound the solution of an initial-boundary value problem in terms of constants independent of grid spacing and initial/boundary data [1]. The classical definition allows non-physical solution growth in time, even though the solution may converge on successive grid refinements [2], which can be detrimental to long-time calculations. In this study, the boundary stencils are, therefore, derived to satisfy the time stability (also called strict or energy stability) definition, which provides a uniform bound of solution in time, preventing non-physical growth in time.

Commonly used time-stable boundary treatments include the weak imposition of boundary conditions (BCs) with simultaneous-approximation-term (SAT)[2] and the projection method[3, 4]. The SAT approach imposes BCs using a penalty term, whereas the projection method uses a projection matrix to incorporate BCs into the system of ODEs solved for the discrete solution. The extent to which the boundary point may satisfy the BC with SAT approach depends on the value of the penalty parameter. A higher value may better satisfy the BC, however, it makes the ODE system stiffer, with adverse implications on time stepping. In case of non-homogeneous BCs, the projection method also does not satisfy the BC exactly because the projected ODE system imposes the time-derivative of BC, and the time-integration of the

^{*}CCS-2, Los Alamos National Laboratory. AIAA Member.

[†]CCS-2, Los Alamos National Laboratory. AIAA Member.

[‡]CCS-2, Los Alamos National Laboratory. AIAA Associate Fellow.

ODE system may not be exact. Therefore, we seek a time-stable method with strong BC enforcement, which satisfies the BC exactly, that may be better suited for accurate near-wall turbulence statistics in direct numerical simulations of wall-bounded flows.

Weak boundary conditions for fluid-flow simulations, solving the Euler or the Navier-Stokes equations, require target values for all conservative variables, which are often not available. Other commonly used approaches, e.g., characteristic boundary conditions [5–7], circumvent that requirement by imposing the known boundary values strongly, which determine the incoming characteristics, and, then, computing the remaining unknowns from differential equations cast in terms of the incoming characteristic variables. While numerical stability analyses of schemes using the weak implementation are widely available [8–12], a similar analysis for strong implementations is hindered by the challenge of constructing a closed system of (differential) equations that incorporates strong boundary conditions and a lack of methodology to examine stability of such systems [13]. This work proposes a framework for stability analysis with strong BCs, facilitating derivation of boundary stencils that are time-stable with strong enforcement of BCs.

The developed framework is then employed to derive closed-form finite-difference cut-cell boundary stencils. Cut-cell methods [14–16] are widely used for fluid-flow simulations over non-trivial solid geometries and moving bodies. The popularity of the cut-cell approach stems from its advantages in grid generation and its computational efficiency. It highly simplifies grid generation on complex domains by considering a Cartesian fluid grid with solid body, intersecting the domain, simply cut out from the fluid domain, as shown schematically in figure 1. Though extensively used for aerodynamic calculations, e.g., [17, 18], the use of cut-cell methods for high-fidelity turbulent flow simulations is limited by several shortcomings in the existing approaches.

Most existing cut-cell approaches use a finite-volume discretization with fluxes approximated using upwind or essentially non-oscillatory reconstructions [19, 20]. These reconstructions help stabilize simulations, especially at cut-cell boundaries, however, they introduce numerical dissipation that may unfavorably influence turbulence/mixing statistics [21, 22]. Cut cells generated in the fluid domain, after extracting the intersecting body, can become arbitrarily small. The small cells require unreasonably small time steps for numerical stability in explicit time integration and they ill-condition the system in case of implicit time integration. This issue, commonly referred to as the small-cell problem, is solved by cell mixing/merging/linking approaches that tend to lower the accuracy of the overall scheme and introduce additional complexity in form of index changes and book keeping. This study aims at resolving the above-mentioned issues by considering a conservative finite-difference discretization that, by construction, does not have a small-cell problem and is provably time-stable [23]. The analytical proof of stability used in this study provides stencils that are different from the cut-cell boundary stencils of [24] derived using optimization procedures.

The paper is organized as follows. Sections II.A and II.B provide the proof of time-stability for a finite-difference discretization applying strong BCs to solve a hyperbolic scalar equation and a hyperbolic system of equations, respectively, on uniform grids. The stencils resulting from the proofs are described in section II.C. Numerical results from the application of the uniform-grid boundary stencils are discussed in section II.D. Section III describes the extension of the uniform-grid framework to derive cut-cell boundary stencils. Numerical results from the cut-cell stencils are provided in section III.B. Section IV discusses the results and the conclusions of the study.

II. Time-stability of strong boundary conditions on uniform grids

In this section, the conditions on boundary stencils for time-stable enforcement of strong BCs to solve a hyperbolic scalar equation and hyperbolic system of equations discretized on uniform grids are derived. Boundary stencils that satisfy those conditions are then applied to various problems to verify the accuracy and stability of the method.

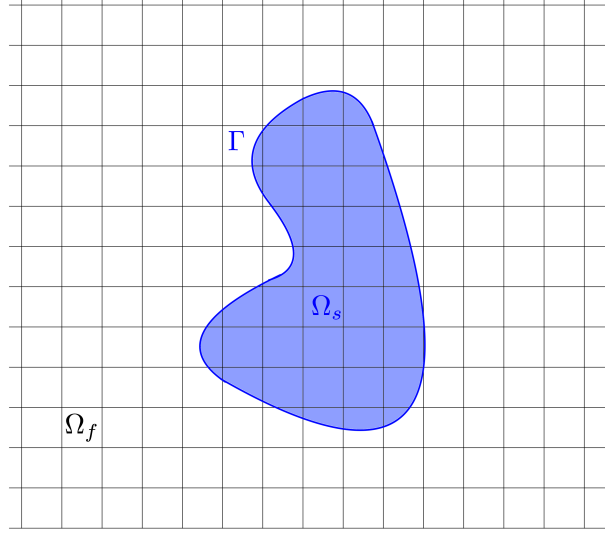


Fig. 1 Schematic of a cut-cell grid. The solid body intersecting the fluid domain (Ω_f) is denoted by Ω_s . The cut-cell boundary is denoted by Γ .

A. Scalar hyperbolic problem

Consider the scalar hyperbolic equation

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = 0, \quad x_0 \leq x \leq x_n, \quad t \geq 0, \quad (1)$$

with the initial and the boundary conditions given by

$$U(x, 0) = f(x), \quad U(x_0, t) = g(t). \quad (2)$$

A semi-discretization of (1)-(2) with strong boundary conditions and $n + 1$ equidistant grid points in the domain can be written as

$$\frac{d\tilde{\mathbf{u}}}{dt} = -D\mathbf{u}, \quad (3)$$

where $\mathbf{u}(t) = [u_0(t) \quad \dots \quad u_n(t)]^T$, with $u_0(t) \equiv g(t)$, is the discrete solution vector. $\tilde{\mathbf{u}}(t) = [u_1(t) \quad \dots \quad u_n(t)]^T$ is the solution vector without the first element, corresponding to the grid point where boundary condition is imposed. D , a matrix of size $n \times (n + 1)$, denotes the derivative operator. The entries of D will be denoted by d_{ij} , where $1 \leq i \leq n$ and $0 \leq j \leq n$. Its non-square structure prevents computation at the first point, essentially, turning it into a flux point.

Define a scalar product and norm for discrete real-valued vector functions $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ by

$$(\mathbf{v}, \mathbf{w})_H = \mathbf{v}^T H \mathbf{w} = \sum_{i,j=1}^{\kappa} h_{ij} v_i w_j \Delta x + \sum_{i=\kappa+1}^{n-\kappa} v_i w_i \Delta x + \sum_{i,j=n-\kappa+1}^n h_{ij} v_i w_j \Delta x, \quad (4)$$

$$\|\mathbf{v}\|_H = \sqrt{(\mathbf{v}, \mathbf{v})_H}, \quad (5)$$

where Δx denotes the grid spacing, κ represents the depth of boundary stencil, and h_{ij} are the entries of a symmetric positive-definite (norm) matrix H .

Multiplying (3) by $\tilde{\mathbf{u}}^T H$, where H is a norm matrix of size $n \times n$, and using the chain rule yields

$$\frac{d}{dt} \|\tilde{\mathbf{u}}\|_H^2 = -\tilde{\mathbf{u}}^T H D \mathbf{u} - (D \mathbf{u})^T H \tilde{\mathbf{u}}. \quad (6)$$

Time-stability of (3) can be ensured by showing

$$\frac{d}{dt} \|\tilde{\mathbf{u}}\|_H^2 \leq K |g|^2, \quad (7)$$

where K must be a constant independent of f , g , Δx , and time step Δt .

To simplify the analysis, the non-square operator $Q = HD$ can be decomposed such that

$$\tilde{\mathbf{u}}^T H D \mathbf{u} = \tilde{\mathbf{u}}^T Q \mathbf{u} = \tilde{\mathbf{u}}^T \tilde{Q} \tilde{\mathbf{u}} + \tilde{\mathbf{u}}^T \mathbf{q}_0 g, \quad (8)$$

where \tilde{Q} is a square ($n \times n$) matrix containing all columns of Q except the first and vector \mathbf{q}_0 is the the first column of Q . Substituting (8) in r.h.s. of (6) provides the time-stability condition:

$$-\tilde{\mathbf{u}}^T H D \mathbf{u} - (D \mathbf{u})^T H \tilde{\mathbf{u}} = -\tilde{\mathbf{u}}^T (\tilde{Q} + \tilde{Q}^T) \tilde{\mathbf{u}} - 2\tilde{\mathbf{u}}^T \mathbf{q}_0 g \leq K |g|^2. \quad (9)$$

The quantity $S = \int_{x_0}^{x_n} U dx$, governed by (1), should depend only on boundary fluxes, *i.e.*,

$$\frac{dS}{dt} = g(t) - U(x_n, t). \quad (10)$$

A discrete statement for (10) is given by

$$\frac{dS}{dt} \approx \sum_{i=1}^n \left(\frac{d}{dt} H \tilde{\mathbf{u}} \right)_i = - \sum_{i=1}^n (H D \mathbf{u})_i = g(t) - u_n(t), \quad (11)$$

where $(\mathbf{v})_i$ denotes the i -th component of vector \mathbf{v} and the entries of H constitute a quadrature for the domain $x_0 \leq x \leq x_n$. In terms of the operators defined in (8), condition (11) translates to

$$\sum_{i=1}^n (\mathbf{q}_0)_i = -1, \quad \sum_{i=1}^n q_{ij} = \begin{cases} 1 & j = n \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

where q_{ij} denotes the element of matrix \tilde{Q} at i -th row and j -th column for $1 \leq i, j \leq n$.

We seek derivative approximations, D , and norm matrices, H , that satisfy the time-stability condition (9) and the discrete conservation statement (11) for various orders of accuracy. The derivation proceeds by assuming an extent of non-zero elements in vector \mathbf{q}_0 , denoted by β , *i.e.*, let $\mathbf{q}_0 = \begin{bmatrix} q_{10} & \cdots & q_{\beta 0} & 0 & \cdots & 0 \end{bmatrix}^T$. $\beta > 0$ represents the depth of boundary stencils that use the physical boundary point, where strong boundary condition is applied, for derivative approximation. A non-zero (row) entry in \mathbf{q}_0 requires a corresponding non-zero diagonal entry in \tilde{Q} to satisfy (9), as shown by the following.

The time-stability condition (9) is satisfied if, for $1 \leq i, j \leq n$ and $\beta > 0$,

$$q_{ij} \begin{cases} = -q_{ji} & \text{if } i \neq j, \\ > 0 & \text{if } i = j \leq \beta, \\ \geq 0 & \text{if } i = j > \beta. \end{cases} \quad (13)$$

The conservation statement (12) is concurrently satisfied if the latter two conditions in (13), for the diagonal entries of \tilde{Q} , are replaced by stricter conditions, given by

$$q_{ij} = \begin{cases} -q_{ji} & \text{if } i \neq j, \\ -\frac{1}{2}q_{i0} > 0 & \text{if } i = j \leq \beta, \\ 0 & \text{if } \beta < i = j < n, \\ \frac{1}{2} & \text{if } i = j = n, \end{cases} \quad (14)$$

and $\sum_{i=1}^{\beta} q_{i0} = -1$. A proof for (13) and (14) is provided in Appendix A.

B. System of hyperbolic equations

In this section, the conditions for time stability of a semi-discretization for a system of one-dimensional hyperbolic equations with strong boundary conditions are discussed. The hyperbolic system coupled at the boundaries, discussed in [2, 25], that provides a severe test of stability for numerical schemes is considered. The system, on domain $0 \leq x \leq 1$ and $t \geq 0$, is given by

$$\frac{\partial \mathbf{U}^I}{\partial t} + \Lambda^I \frac{\partial \mathbf{U}^I}{\partial x} = 0, \quad (15)$$

$$\frac{\partial \mathbf{U}^{II}}{\partial t} + \Lambda^{II} \frac{\partial \mathbf{U}^{II}}{\partial x} = 0, \quad (16)$$

where

$$\mathbf{U}^I = \left[U^1(x, t) \quad \cdots \quad U^k(x, t) \right]^T, \quad \Lambda^I = \text{diag}(\lambda_1, \dots, \lambda_k), \quad \text{for } \lambda_1 > \lambda_2 > \cdots > \lambda_k > 0 \quad (17)$$

describe a system of right-moving waves and

$$\mathbf{U}^{II} = \left[U^{k+1}(x, t) \quad \cdots \quad U^r(x, t) \right]^T, \quad \Lambda^{II} = \text{diag}(\lambda_{k+1}, \dots, \lambda_r), \quad \text{for } 0 > \lambda_{k+1} > \lambda_{k+2} > \cdots > \lambda_r, \quad (18)$$

a system of left-moving waves. The system (15)-(16) is well-posed for arbitrary initial conditions with continuous derivative and boundary conditions given by

$$\mathbf{U}^I(0, t) = L\mathbf{U}^{II}(0, t) + \mathbf{g}^I(t), \quad (19)$$

$$\mathbf{U}^{II}(1, t) = R\mathbf{U}^I(1, t) + \mathbf{g}^{II}(t), \quad (20)$$

where L and R are constant matrices of size $k \times (r - k)$ and $(r - k) \times k$, respectively, and \mathbf{g}^I and \mathbf{g}^{II} are vectors of size k and $r - k$, respectively. The system (15)-(20) has a non-growing solution in time if \mathbf{g}^I and \mathbf{g}^{II} are zero and

$$\|L\| \|R\| \leq 1. \quad (21)$$

The matrix norm for real matrices is defined by $\|L\|^2 = \rho(L^T L)$, where $\rho(\cdot)$ denotes the spectral radius.

A semi-discretization of (15)-(20) using strong boundary conditions can be written as

$$\frac{d\mathbf{w}}{dt} = -\mathcal{D}\mathbf{w}, \quad (22)$$

where $\mathbf{w}(t) = \begin{bmatrix} \tilde{\mathbf{u}}^I(t) & \tilde{\mathbf{u}}^{II}(t) \end{bmatrix}^T$ with $\tilde{\mathbf{u}}^I(t) = \begin{bmatrix} \tilde{\mathbf{u}}^1(t) & \dots & \tilde{\mathbf{u}}^k(t) \end{bmatrix}$ and $\tilde{\mathbf{u}}^{II}(t) = \begin{bmatrix} \tilde{\mathbf{u}}^{k+1}(t) & \dots & \tilde{\mathbf{u}}^r(t) \end{bmatrix}$. The unknowns for each equation in the system are given, assuming a discretization with $n+1$ grid points, as described in section II.A, by $\tilde{\mathbf{u}}^\phi(t) = \begin{bmatrix} u_1^\phi(t) & \dots & u_n^\phi(t) \end{bmatrix}^T$ for $1 \leq \phi \leq k$ and by $\tilde{\mathbf{u}}^\phi(t) = \begin{bmatrix} u_0^\phi(t) & \dots & u_{n-1}^\phi(t) \end{bmatrix}^T$ for $k+1 \leq \phi \leq r$, where $\tilde{\mathbf{u}}^\phi(t)$ is the solution vector without the element corresponding to the grid point where boundary condition is applied. Therefore, the solution vectors for the first k equations do not contain the element corresponding to the first point and the rest do not contain the element corresponding to the last point.

The derivative operator, \mathcal{D} , is then given by

$$\mathcal{D} = \Lambda \mathcal{H}^{-1} \mathcal{Q}, \quad (23)$$

where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$,

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_{11} & 0 \\ 0 & \mathcal{H}_{22} \end{bmatrix}, \quad \text{and} \quad \mathcal{Q} = \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21} & \mathcal{Q}_{22} \end{bmatrix}. \quad (24)$$

The submatrices

$$\mathcal{H}_{11} = I_k \otimes H, \quad \mathcal{H}_{22} = I_{r-k} \otimes H^\#, \quad (25)$$

$$\mathcal{Q}_{11} = I_k \otimes \tilde{\mathcal{Q}}, \quad \mathcal{Q}_{12} = L \otimes \mathcal{Q}_0, \quad \mathcal{Q}_{21} = -R \otimes \mathcal{Q}_0^\#, \quad \mathcal{Q}_{22} = -I_{r-k} \otimes \tilde{\mathcal{Q}}^\#, \quad (26)$$

where I_m denotes an identity matrix of size $m \times m$, \otimes denotes the Kronecker product and the superscript $\#$ denotes matrices/vectors rotated by 180° , for example,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^\# = \begin{bmatrix} d & c \\ b & a \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a \\ b \end{bmatrix}^\# = \begin{bmatrix} b \\ a \end{bmatrix}. \quad (27)$$

\mathcal{Q}_0 is a $n \times n$ matrix with \mathbf{q}_0 as the first column and remaining columns zero. The vector \mathbf{q}_0 and matrices H and $\tilde{\mathcal{Q}}$ are as described in section II.A.

Let the discrete energy be defined as (e.g., [2, 25])

$$E(t) = \sum_{\phi=1}^k \frac{\|R\|}{\lambda_\phi} (\tilde{\mathbf{u}}^\phi)^T H \tilde{\mathbf{u}}^\phi + \sum_{\phi=k+1}^r \frac{\|L\|}{|\lambda_\phi|} (\tilde{\mathbf{u}}^\phi)^T H^\# \tilde{\mathbf{u}}^\phi, \quad (28)$$

which provides the time stability condition:

$$\frac{dE}{dt} = \sum_{\phi=1}^k \frac{\|R\|}{\lambda_\phi} \frac{d}{dt} (\tilde{\mathbf{u}}^\phi)^T H \tilde{\mathbf{u}}^\phi + \sum_{\phi=k+1}^r \frac{\|L\|}{|\lambda_\phi|} \frac{d}{dt} (\tilde{\mathbf{u}}^\phi)^T H^\# \tilde{\mathbf{u}}^\phi \leq 0. \quad (29)$$

The choice of $\mathbf{g}^I = 0$ and $\mathbf{g}^{II} = 0$ in (19)-(20) results in the time-stability requirement of non-positive dE/dt in (29). Non-zero \mathbf{g}^I and \mathbf{g}^{II} will require a bound in terms of $\|\mathbf{g}^I\|$ and $\|\mathbf{g}^{II}\|$. However, for stability analysis, it suffices to assume that both \mathbf{g}^I and \mathbf{g}^{II} vanish, without loss of generality [2].

$$q_{ij} \begin{cases} = -q_{ji} & \text{if } i \neq j, \\ > 0 & \text{if } i = j \leq \beta, \\ = 0 & \text{if } \beta < i = j < n, \\ = \frac{1}{2} & \text{if } i = j = n, \end{cases} \quad (35)$$

and $\sum_{j=0}^{\beta} \sum_{i=1}^{\kappa} q_{ij} = -1$, such that the following conservation statement holds,

$$\frac{dS}{dt} \approx \sum_{i=1}^n \left(\frac{d}{dt} H \tilde{\mathbf{u}} \right)_i = - \sum_{i=1}^n (H D \mathbf{u})_i = g(t) - u_n(t) + O(\Delta x), \quad (36)$$

in place of (11). As obvious, in the limit of $\Delta x \rightarrow 0$, the statement (36) tends to (11).

3. Higher-order schemes

Using the approach described in the previous section, several stencils for 3 – 6 – 3 and 4 – 8 – 4 schemes were found that satisfy (13) and (36), and are time-stable for the coupled system in section II.D.2. However, a detailed discussion of these schemes is beyond the scope of this article and will be a subject of future publication.

D. Numerical results from uniform-grid simulations

In this section, numerical results from application of boundary stencils derived in the previous section are discussed. The derived stencils are used at the boundary where physical boundary condition is applied. At the outflow boundary, where no BC is applied, diagonal-norm SBP stencils derived in [26] are used. In all cases, time integration is performed using the classical fourth-order Runge-Kutta (RK4) method. For convergence studies, the time step is taken small enough such that the temporal errors are insignificant compared to the spatial truncation errors.

1. 1-D scalar advection equation

Consider the scalar hyperbolic equation (1), on a spatial domain $0 \leq x \leq 1$, with initial and boundary conditions given by

$$u(x, 0) = \sin 2\pi x, \quad u(0, t) = g(t) = \sin 2\pi(-t). \quad (37)$$

The exact solution to the problem is $u(x, t) = \sin 2\pi(x - t)$. A semi-discretization to the problem, using strong BCs, the notation of (3), and the decomposition described in (8), is given by

$$\frac{d\tilde{\mathbf{u}}}{dt} = -D\mathbf{u} = -H^{-1}\tilde{Q}\tilde{\mathbf{u}} - H^{-1}\mathbf{q}_0g. \quad (38)$$

For a bounded boundary data $g(t)$, the stability of the semi-discretization depends on the properties of the matrix $M = -H^{-1}\tilde{Q}$, referred to as the system matrix. Figures 2 and 3 show the eigenvalue spectrum of the system matrix using various grid points for the 1 – 2 – 1 and the 2 – 4 – 2 scheme, respectively. All eigenvalues lie in strict left half of the complex plane and, therefore, the discretization is time-stable.

Table 1 shows the L_2 - and L_∞ -norm of the solution error, denoted by ε , and the respective convergence rates from the two schemes. As expected, the 1 – 2 – 1 scheme converges with second-order accuracy and the 2 – 4 – 2 scheme converges with third-order accuracy.

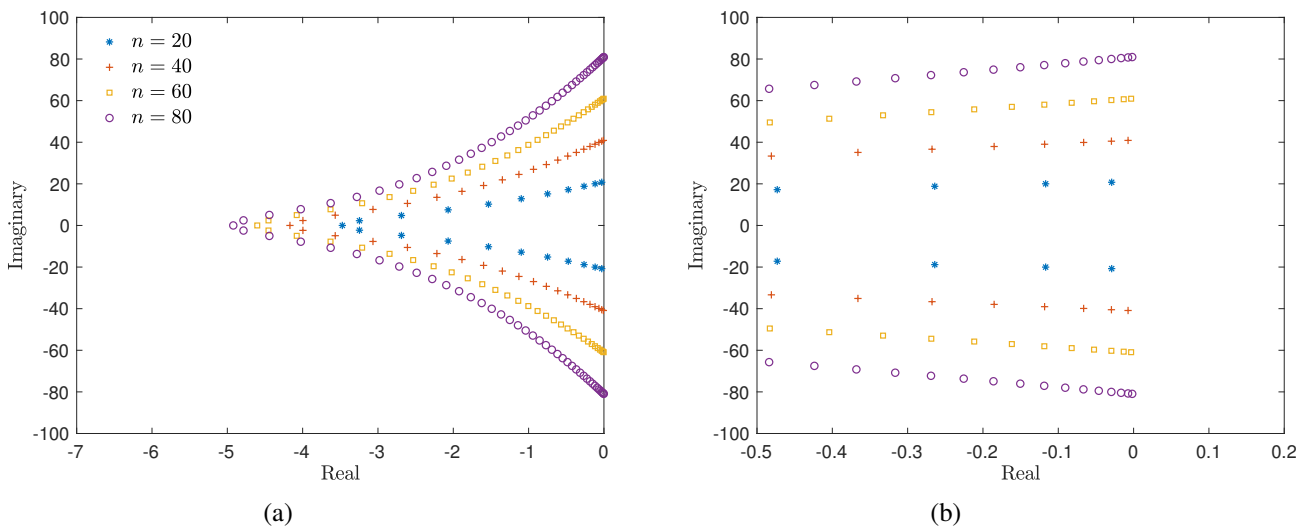


Fig. 2 Eigenvalue spectrum of the system matrix to solve (1) with initial and boundary condition given by (37) using 1 – 2 – 1 scheme for various number of grid points. (a) All eigenvalues, (b) Magnified view near the imaginary axis. Legend is the same for both plots.

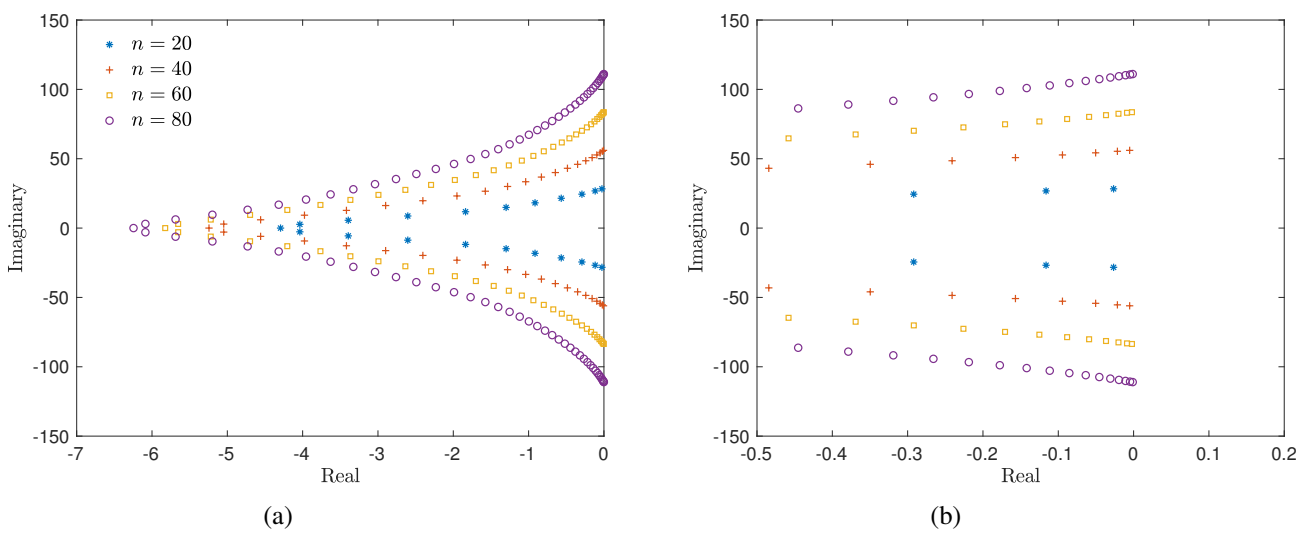


Fig. 3 Eigenvalue spectrum of the system matrix to solve (1) with initial and boundary condition given by (37) using 2 – 4 – 2 scheme for various number of grid points. (a) All eigenvalues, (b) Magnified view near the imaginary axis. Legend is the same for both plots.

n	1 – 2 – 1				2 – 4 – 2			
	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate
20	-1.442427		-1.234263		-1.828907		-1.541334	
40	-2.044080	1.999	-1.834978	1.996	-2.789357	3.215	-2.335029	2.637
80	-2.644558	1.995	-2.435158	1.994	-3.729319	3.298	-3.204515	2.888
160	-3.245543	1.996	-3.039630	2.008	-4.653197	3.110	-4.099137	2.972
320	-3.846993	1.998	-3.646874	2.017	-5.567189	3.046	-5.000487	2.994
640	-4.448730	1.999	-4.250385	2.005	-6.475805	3.027	-5.903084	2.998

Table 1 L_2 - and L_∞ -norm of the error and convergence rates for the 1 – 2 – 1 and 2 – 4 – 2 scheme. Error calculations performed at $t_f = 1.0$.

2. 1-D coupled hyperbolic system

Consider the hyperbolic system, on domain $0 \leq x \leq 1$ and $t \geq 0$,

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = 0, \quad (39)$$

$$\frac{\partial V}{\partial t} - \frac{\partial V}{\partial x} = 0. \quad (40)$$

$$\text{Initial conditions : } U(x, 0) = \sin 2\pi x, \quad V(x, 0) = -\sin 2\pi x. \quad (41)$$

$$\text{Boundary conditions : } U(0, t) = V(0, t), \quad V(1, t) = U(1, t). \quad (42)$$

This system provides a severe test of numerical stability because it is neutrally stable, *i.e.*, the energy, $\int_0^1 [U(x, t)^2 + V(x, t)^2] dx$, remains constant with time.

Let $\mathbf{u}(t) = [u_0(t) \ \cdots \ u_n(t)]^T$ and $\mathbf{v}(t) = [v_0(t) \ \cdots \ v_n(t)]^T$ denote the grid function, assuming a spatial discretization of the above system with $n + 1$ grid points. A semi-discretization of (39)-(42) with strong boundary conditions is given by

$$\frac{d\mathbf{w}}{dt} = -\mathcal{D}\mathbf{w}, \quad (43)$$

where $\mathbf{w}(t) = [\tilde{\mathbf{u}}(t) \ \tilde{\mathbf{v}}(t)]^T$ with $\tilde{\mathbf{u}}(t) = [u_1(t) \ \cdots \ u_n(t)]^T$ and $\tilde{\mathbf{v}}(t) = [v_0(t) \ \cdots \ v_{n-1}(t)]^T$. The derivative operator, \mathcal{D} , is given by

$$\mathcal{D} = \begin{bmatrix} H & 0 \\ 0 & H^\# \end{bmatrix}^{-1} \begin{bmatrix} \tilde{Q} & Q_0 \\ -Q_0^\# & -\tilde{Q}^\# \end{bmatrix} = \mathcal{H}^{-1}Q, \quad (44)$$

where \tilde{Q} and Q_0 are as described in (8) and (26), respectively, and the superscript $\#$ denotes matrices/vectors rotated by 180° , as described in (27).

To the best of our knowledge, there are no third- or higher-order finite-difference stencils that can solve the problem (39)-(42) with strong BCs in a time-stable manner, *i.e.*, stably for long times. Figure 4(a) shows the eigenvalue spectrum of the system matrix used to solve (39)-(42) with strong BCs and first-derivative stencils from some popular references [26, 33, 34]. Few eigenvalues in each case lie in right half of the complex plane, *i.e.*, the maximum real part of the eigenvalues is positive. Therefore, each scheme will exhibit unphysical solution growth in time, as shown in figure 4(b).

Figures 5 and 6 show the the eigenvalue spectrum of the system matrix, given by $-\mathcal{D}$ in (43), using various number

of grid points for the 1 – 2 – 1 and the 2 – 4 – 2 scheme, respectively. All eigenvalues lie in strict left half of the complex plane indicating a time-stable discretization. Table 2 shows the L_2 - and L_∞ -norm of the solution error, denoted by ϵ , and the respective convergence rates from the two schemes. As expected, the 1 – 2 – 1 scheme converges with second-order accuracy and the 2 – 4 – 2 scheme converges with third-order accuracy.

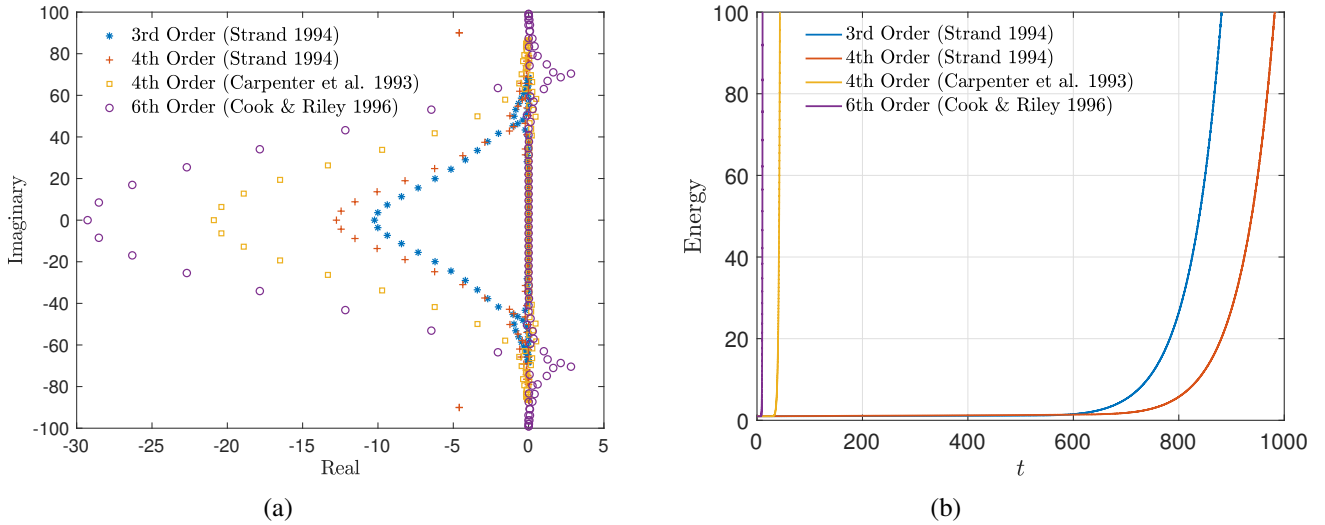


Fig. 4 (a) Eigenvalue spectrum of the system matrix and (b) solution energy from solving the coupled hyperbolic system (39)-(42) using various spatial schemes with strong boundary conditions and 51 grid points.

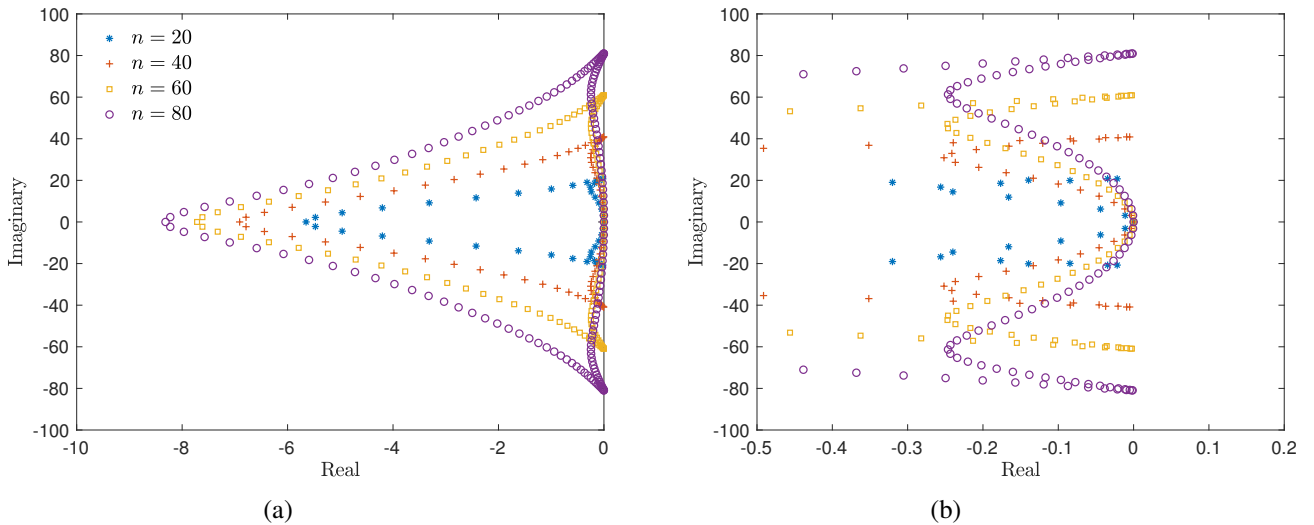


Fig. 5 Spectrum of the system matrix to solve (39)-(42) using 1 – 2 – 1 scheme for various number of grid points. (a) All eigenvalues, (b) Magnified view near the imaginary axis. Legend is the same for both plots.

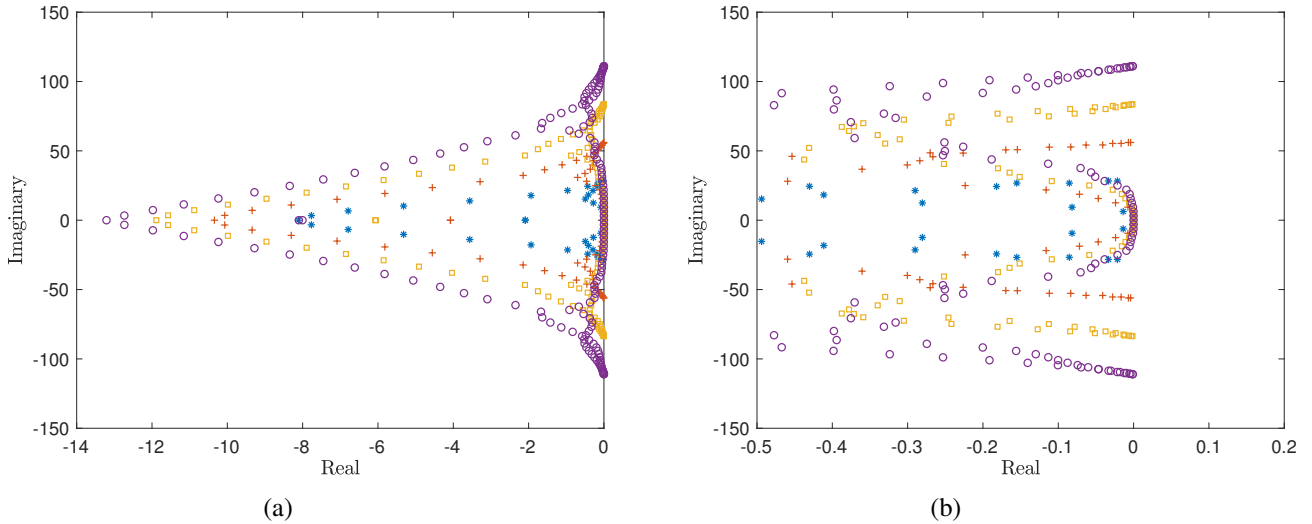


Fig. 6 Spectrum of the system matrix to solve (39)-(42) using 2-4-2 scheme for various number of grid points. (a) All eigenvalues, (b) Magnified view near the imaginary axis. Legend is the same for both plots.

n	1-2-1				2-4-2			
	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate
20	-1.217223		-1.225890		-1.676188		-1.508359	
40	-1.803716	1.948	-1.770808	1.810	-2.643277	3.215	-2.351858	2.802
80	-2.398761	1.977	-2.353810	1.937	-3.582599	3.120	-3.206750	2.840
160	-2.997715	1.990	-2.955241	1.998	-4.505004	3.064	-4.099017	2.964
320	-3.598344	1.995	-3.555882	1.995	-5.417936	3.035	-5.000116	2.993
640	-4.199721	1.998	-4.157098	1.997	-6.325949	3.016	-5.902821	2.999

Table 2 L_2 - and L_∞ -norm of the error and convergence rate with the 1-2-1 and 2-4-2 scheme. Error calculations performed at $t_f = 1.0$.

3. 2-D variable-coefficient advection equation

Consider the scalar problem

$$\frac{\partial \phi}{\partial t} + u \frac{\partial \phi}{\partial x} + v \frac{\partial \phi}{\partial y} = 0, \quad 0 \leq x, y \leq L \quad t \geq 0, \quad (45)$$

$$u(x, y) = \frac{\partial r}{\partial x}, \quad v(x, y) = \frac{\partial r}{\partial y}, \quad (46)$$

$$r(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2}, \quad (47)$$

where $L = \sqrt{2}$, $x_0 = -0.25$ and $y_0 = -0.25$. The initial and boundary conditions are given by

$$\phi(x, y, 0) = \sin 2\pi r, \quad (48)$$

and

$$\phi(0, y, t) = \sin 2\pi (r(0, y) - t), \quad \phi(x, 0, t) = \sin 2\pi (r(x, 0) - t), \quad (49)$$

respectively. The exact solution to the problem is $\phi(x, y, t) = \sin 2\pi (r - t)$.

Figures 7(a) and (b) show the L_∞ -error from a long-time simulation using 1-2-1 and 2-4-2 scheme, respectively. A low and a high value of CFL numbers are used with various number of grid points to show that the derived boundary stencils yield accurate results at reasonable time steps. The error remains constant with time indicating a time-stable behavior. As expected, errors from the 2-4-2 scheme are smaller than that from the 1-2-1 scheme. Increasing the CFL number from 0.3 to 0.8 does not have a visible influence on error profile indicating that the spatial truncation error dominates in these runs.

Table 3 shows the errors and convergence rates from the two schemes used to solve (46)-(49). As desired, the 1-2-1 scheme converges with second-order accuracy and the 2-4-2 scheme converges with third-order accuracy.

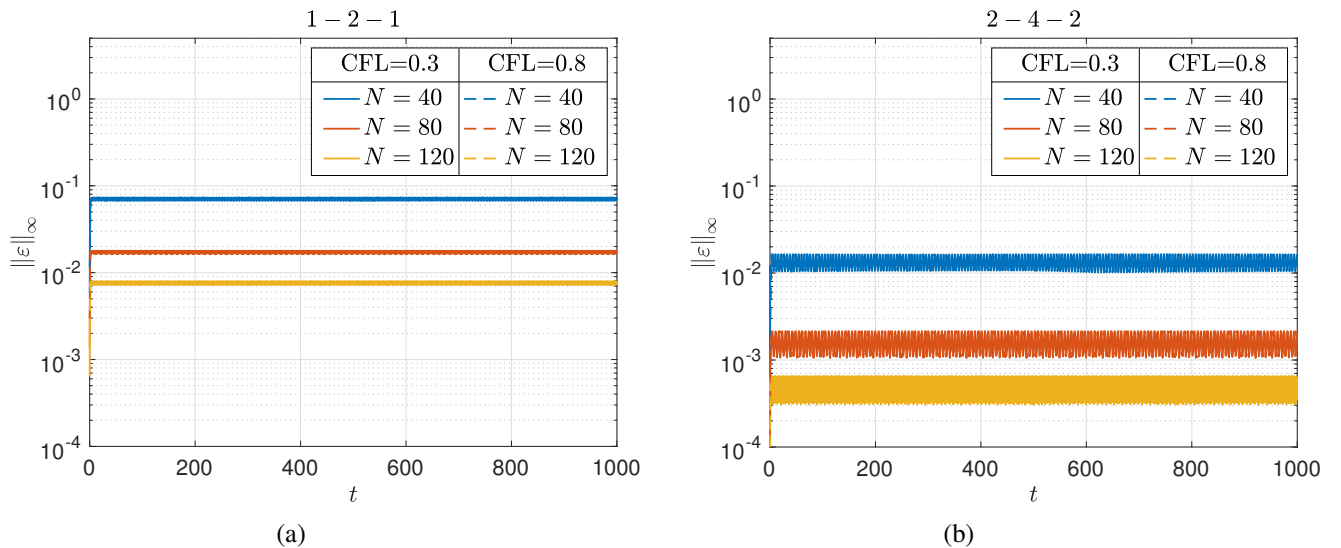


Fig. 7 L_∞ -error from a long-time simulation of (46)-(49) using two CFL numbers, $N \times N$ grid points, and (a) 1-2-1 and (b) 2-4-2 scheme.

N	1 - 2 - 1				2 - 4 - 2			
	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate
30	-1.404196		-1.037912		-2.118175		-1.439252	
60	-2.018962	1.993	-1.615432	1.872	-3.092760	3.160	-2.445756	3.263
120	-2.626948	1.995	-2.207732	1.944	-4.035679	3.095	-3.435688	3.249
240	-3.232256	1.999	-2.801850	1.962	-4.954626	3.034	-4.343758	2.998

Table 3 L_2 - and L_∞ -norm of the error and convergence rate with the 1 - 2 - 1 and 2 - 4 - 2 scheme used to solve (46)-(49) on a $N \times N$ grid. Error calculations performed at $t_f = 1.0$.

III. Time-stability of strong boundary conditions on cut-cell grids

On a cut-cell grid, boundary point(s) generated by the embedded boundary may not be uniformly spaced with respect to other grid points, as seen in figure 1. A one-dimensional analogue of such a configuration, for solving the scalar hyperbolic problem (1)-(2), is shown in figure 8. $\alpha = 1$ corresponds to a uniform grid. For cut-cell configurations, assuming $x_0 = 0$ for $\alpha = 1$, the boundary coordinate becomes

$$x_0 = (1 - \alpha) \Delta x = x_1 - \alpha \Delta x, \quad (50)$$

where $x_1 = \Delta x$ is fixed.

Following the same approach as section II.A, a semi-discretization for the cut-cell configuration can be written as (3), where strong BC $u_0(t) \equiv U(x_0, t) \equiv g(t)$ is applied. The time-stability condition and the discrete conservation statement are, similarly, given by (9) and (11), respectively. The conditions on the entries of $Q = HD$ to satisfy (9) and (11), given by (13)-(14), remain the same. The cut-cell boundary location, x_0 , influences the truncation error calculation (affecting the boundary stencils in D) and the quadrature for conserved quantity calculation (affecting the entries of H near cut-cell boundary). However, constraints on the entries of matrix Q for stability and conservation remain unchanged. For the system of hyperbolic equations (section II.B), the constraints (30) for time-stability hold.

The task, as in the uniform-grid case, is to find a symmetric positive-definite matrix, H , and a derivative operator, D , for various orders of accuracy that satisfy the time-stability and conservation constraints.

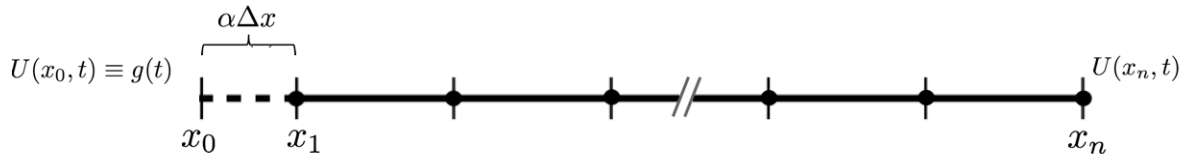


Fig. 8 One-dimensional grid with variable first grid point location, analogous to a cut-cell boundary point. Solid dots denote grid points where the governing equation is solved. $0 \leq \alpha \leq 1$.

A. Stencils for various orders of accuracy

As mentioned in section II.C, the present study focuses on diagonal-norm stencils. For a conservative finite-difference scheme, the entries of the diagonal norm matrix, H , represent cell sizes for a corresponding flux-based discretization given at a point, to solve (1), by (e.g., [35])

$$\frac{du_i}{dt} = -\frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{h_{ii}\Delta x}, \quad (51)$$

integration is performed using the classical fourth-order Runge-Kutta (RK4) method. For convergence studies, the time step is taken small enough such that the temporal errors are insignificant compared to the spatial truncation errors.

1. 1-D scalar advection equation

Consider the scalar hyperbolic equation (1), on a spatial domain $x_0 \leq x \leq 1$, where x_0 is given by (50), with initial and boundary conditions given by

$$u(x, 0) = \sin 2\pi x, \quad u(x_0, t) = g(t) = \sin 2\pi (x_0 - t). \quad (53)$$

The exact solution to the problem is $u(x, t) = \sin 2\pi (x - t)$.

Figures 10(a) and (b) show the L_∞ -error from a long-time simulation using $1-2-1$ and $2-4-2$ scheme, respectively. A low and a high value of CFL numbers are used with various values of α to show that the derived boundary stencils do not have the small-cell problem. The error remains constant with time indicating a time-stable behavior. As expected, errors from the $2-4-2$ scheme are smaller than that from the $1-2-1$ scheme. Increasing the CFL number from 0.3 to 0.8 does not have a visible influence on error profile indicating that the spatial truncation error dominates in these runs.

Figures 11(a) and (b) show the L_∞ -error and the convergence rates of the two schemes for various values of α . As expected, the $1-2-1$ scheme converges with second-order accuracy and the $2-4-2$ scheme converges with third-order accuracy. The error profiles in figures 10 and 11 suggest that the prefactor in leading-order truncation error term does not vary much with α and, therefore, the error magnitude is similar for various values of α .

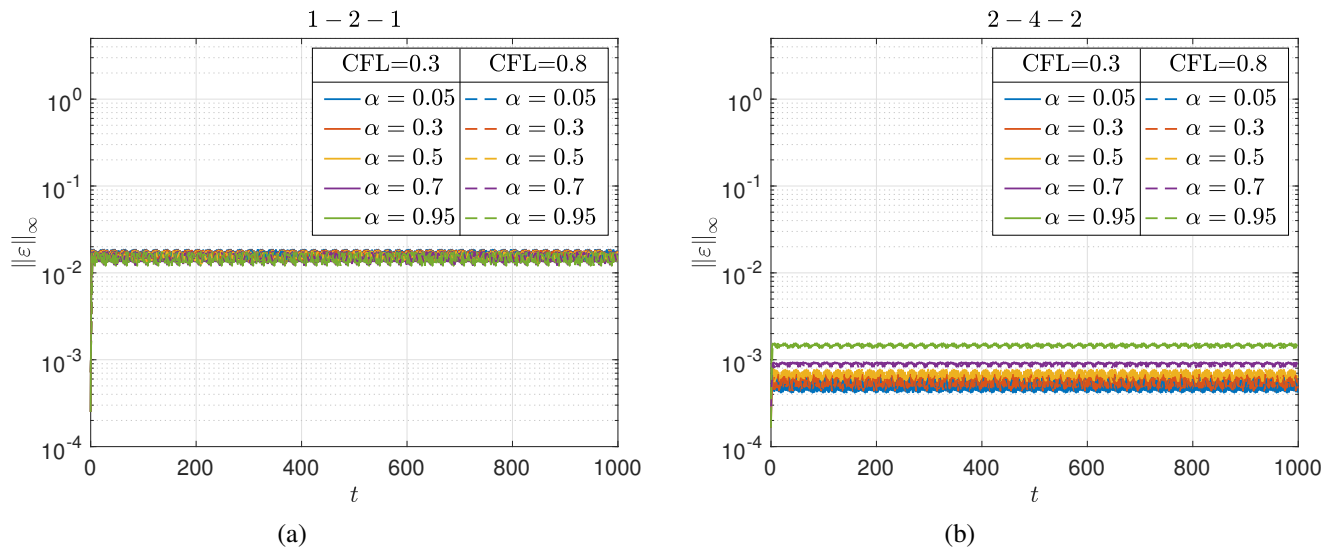


Fig. 10 L_∞ -error from a long-time simulation of (1) using x_0 given by (50), initial and boundary condition given by 53, various values of α , and (a) $1-2-1$ and (b) $2-4-2$ scheme at two CFL numbers.

2. 2-D variable-coefficient advection equation

Consider a cut-cell grid as shown in figure 12(a) and the scalar problem of section II.D.3 with a different domain extent, given by

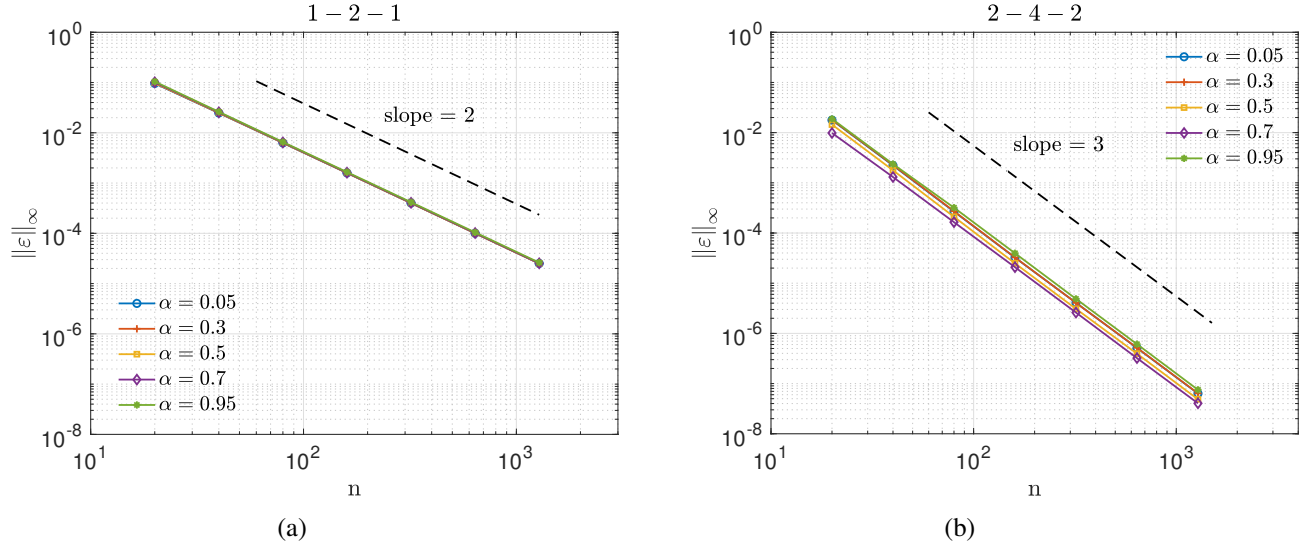


Fig. 11 L_∞ -norm of the solution error and convergence rate with the (a) 1 – 2 – 1 and (b) 2 – 4 – 2 scheme used to solve (1) with x_0 given by (50) and a domain with n grid points. Error calculations performed at $t_f = 1.0$. Dashed black line show the expected order-of-accuracy.

$$\frac{\partial \phi}{\partial t} + u \frac{\partial \phi}{\partial x} + v \frac{\partial \phi}{\partial y} = 0, \quad -L \leq x, y \leq L \quad t \geq 0, \quad (54)$$

where u and v are expressed as (46), and $L = 1$, $x_0 = 0$ and $y_0 = 0$ is assumed. The initial condition is given by (48) and the boundary condition at the cut-cell boundary $r = \sqrt{(x - x_0)^2 + (y - y_0)^2} = r_0$ is

$$\phi(x, y, t) = \sin 2\pi(r_0 - t) \quad \text{at} \quad \sqrt{(x - x_0)^2 + (y - y_0)^2} = r_0. \quad (55)$$

The exact solution to the problem is $\phi(x, y, t) = \sin 2\pi(r - t)$. The grid points with $r < r_0$ are blanked out, *i.e.*, the governing equation is not solved there. Figure 12(b) shows a surface plot of the initial condition on the cut-cell grid.

Figures 13(a) and (b) show the L_∞ -error from a long-time simulation of (54) using the 1 – 2 – 1 and the 2 – 4 – 2 scheme, respectively. A low and a high value of CFL numbers are used with various number of grid points to show that the small-cell problem does not arise with the derived boundary stencils. The error remains constant with time indicating a time-stable behavior. As expected, errors from the 2 – 4 – 2 scheme are smaller than that from the 1 – 2 – 1 scheme. Increasing the CFL number from 0.3 to 0.8 does not have a visible influence on error profile indicating that the spatial truncation error dominates in these runs.

Table 4 shows the errors and convergence rates from the two schemes. As desired, the 1 – 2 – 1 scheme converges with second-order accuracy and the 2 – 4 – 2 scheme converges with third-order accuracy.

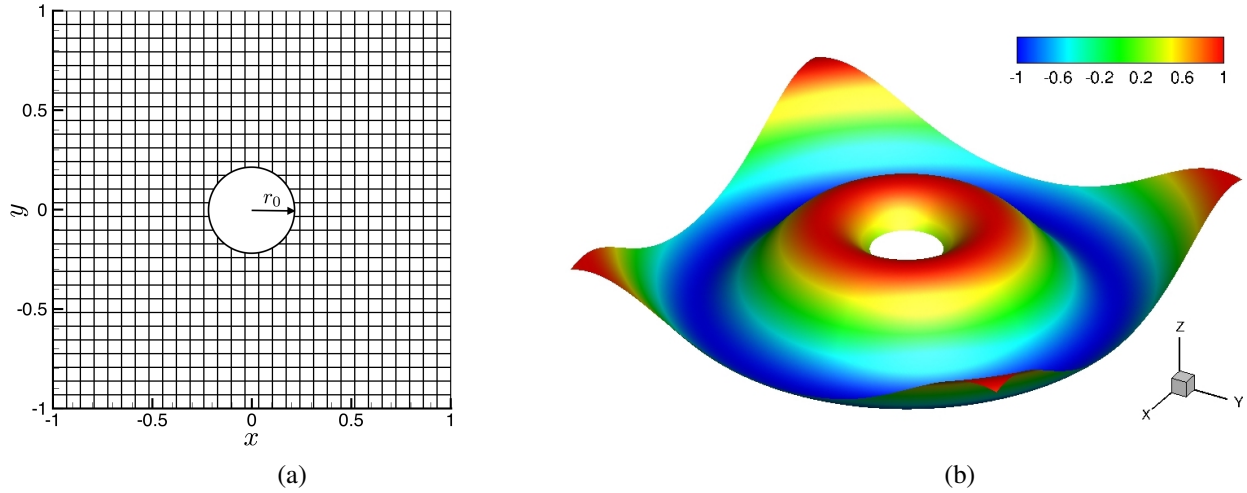


Fig. 12 (a) An example cut-cell grid for the scalar problem (54) and (b) a surface plot of the initial condition on the cut-cell domain shown in subfigure (a).

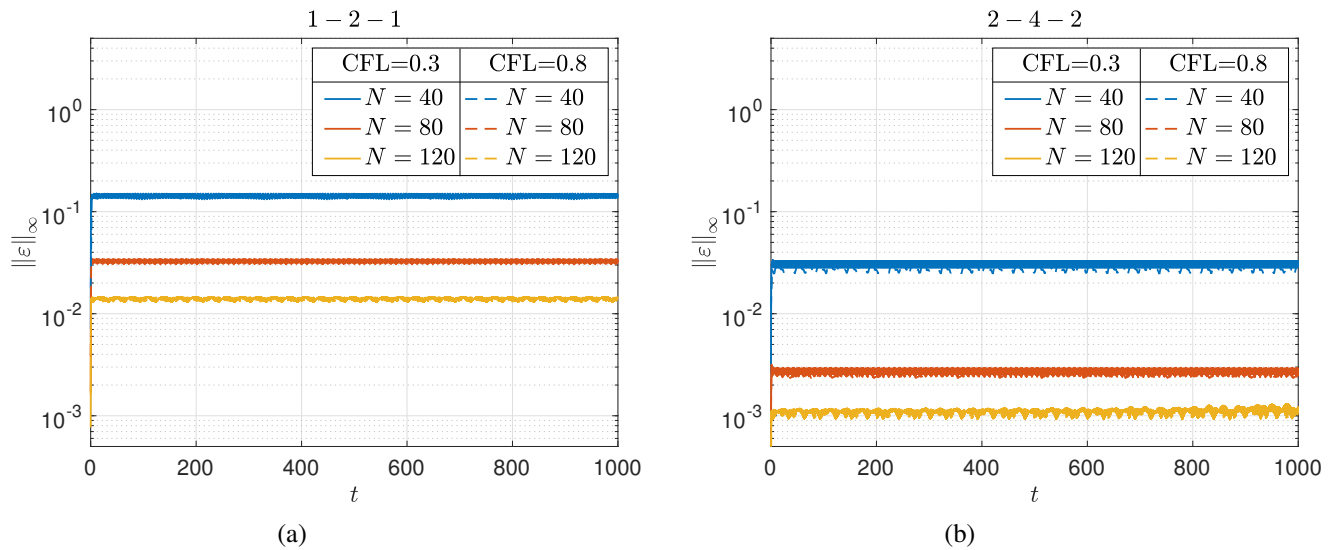


Fig. 13 L_∞ -error from a long-time simulation of (54) using two CFL numbers and $N \times N$ grid points with (a) 1 - 2 - 1 and (b) 2 - 4 - 2 scheme.

N	1 - 2 - 1				2 - 4 - 2			
	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate	$\log_{10} \ \varepsilon\ _2$	Rate	$\log_{10} \ \varepsilon\ _\infty$	Rate
30	-0.989345		-0.643610		-1.666727		-1.196508	
60	-1.662429	2.182	-1.273019	2.041	-2.639233	3.153	-2.191315	3.225
120	-2.355574	2.275	-1.921113	2.127	-3.573114	3.065	-3.125201	3.102
240	-3.019971	2.194	-2.553477	2.088	-4.482657	3.021	-4.025420	2.990

Table 4 L_2 - and L_∞ -norm of the error and convergence rate with the 1 - 2 - 1 and 2 - 4 - 2 schemes. Error calculations performed at $t_f = 1.0$.

3. 2-D Euler equations

In this section, we discuss the extension of the cut-cell approach to solve the Euler equations. The conservative form of the two-dimensional Euler equations is given by

$$\frac{\partial Q}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0, \quad (56)$$

$$Q = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ E \end{bmatrix}, \quad F = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{bmatrix}, \quad G = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{bmatrix}, \quad (57)$$

$$E = \frac{p}{\gamma - 1} + \rho \left(\frac{u^2 + v^2}{2} \right), \quad (58)$$

where u, v are the Cartesian velocity components, ρ denotes the density, p the pressure and E is the total energy. γ denotes the ratio of specific heats.

Assume a cut-cell boundary as shown in figure 14. A semi-discretization for the grid point shown in red, denoted by subscript ij , using strong BCs is given by

$$\frac{d\mathbf{q}_{ij}}{dt} = -S \left(\left[S^{-1} (D_x^{\text{out}} \mathbf{f})_{ij} \right]_+ + \left[S^{-1} (D_x^{\text{in}} \mathbf{f})_{ij} \right]_- \right) - (D_y \mathbf{g})_{ij}, \quad (59)$$

where columns of S are the right eigenvectors of the Jacobian matrix $A = \partial F / \partial x = S \Lambda S^{-1}$. The matrices S and Λ can be found in [11]. D_x^{in} and D_x^{out} are the inflow and outflow boundary stencils, respectively. As mentioned before, diagonal-norm SBP stencils are used as outflow stencil and the boundary stencils derived in section III.A are used as inflow stencil. The characteristic decomposition in semi-discretization (59) is performed only in x -direction because the grid point, shown in red in figure 14, needs to use the boundary stencil only in x -direction. In y -direction interior stencil is used. In the case where a grid point has to use the boundary stencil in y -direction, a similar decomposition ought to be performed in y -direction.

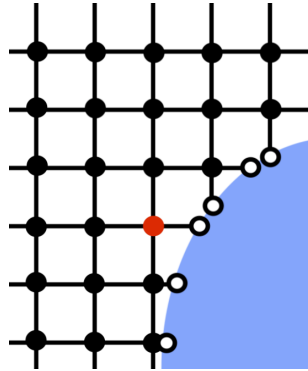


Fig. 14 Two-dimensional schematic of grid points near a cut-cell boundary.

To examine the performance of the cut-cell method, we solve the two-dimensional Euler equations for the propagation of a compressible isentropic vortex on a domain with cut boundaries, as shown in figure 15. Boundaries in x -direction are assumed to be periodic, which allows the vortex to loop through the domain multiple times, assessing long-time

stability of the method. Characteristic boundary conditions are applied in y -direction using the exact solution for strong enforcement.

The exact solution to the problem is given by

$$\begin{aligned} \rho &= \left(1 - \frac{\varpi^2(\gamma - 1)}{8\pi^2 c_0^2} e^{1-\varphi^2 r^2}\right)^{\frac{1}{\gamma-1}}, & u &= u_0 - \frac{\varpi}{2\pi} \varphi(y - y_0 - v_0 t) e^{\frac{1-\varphi^2 r^2}{2}}, \\ v &= v_0 + \frac{\varpi}{2\pi} \varphi(x - x_0 - u_0 t) e^{\frac{1-\varphi^2 r^2}{2}}, & E &= \frac{P}{\gamma - 1} + \frac{1}{2} \rho(u^2 + v^2), \\ p &= \rho^\gamma, & r^2 &= (x - x_0 - u_0 t)^2 + (y - y_0 - v_0 t)^2, \end{aligned} \quad (60)$$

where (x_0, y_0) denotes the initial position of the vortex, (u_0, v_0) denotes the vortex convective velocity, φ is a scaling factor and ϖ denotes the non-dimensional circulation. Unless otherwise stated, we use $v_0 = 0$, $\gamma = 1.4$, $\varphi = 11$ and $\varpi = 1$. All quantities in (60) are non-dimensional, obtained from the density scale $= \rho_0^*$, velocity scale $u_0^* = \frac{c_0^*}{\sqrt{\gamma}}$, unit length scale and pressure scale $= \rho_0^* u_0^{*2}$, where $*$ denotes the dimensional quantities. The non-dimensional ambient speed of sound is $c_0 = \sqrt{\gamma}$.

Figures 13(a) and (b) show the L_∞ -error from a long-time simulation using the 1-2-1 and the 2-4-2 scheme, respectively. A subsonic ($u_0 = 1.0$) and a supersonic ($u_0 = 2.0$) convective velocity is used to examine the robustness of boundary implementation for cases where all characteristics leave/enter the domain as well as cases where some characteristics enter, while the others leave the domain. The error remains constant with time indicating a time-stable behavior. As expected, errors from the 2-4-2 scheme are smaller than that from the 1-2-1 scheme. The calculations were performed at a CFL number of 0.5, showing for this non-linear problem that the method does not have a small-cell problem.

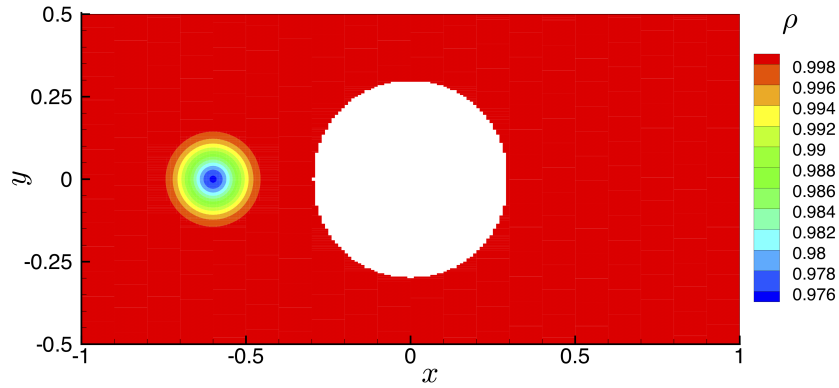


Fig. 15 Cut-cell domain showing the initial density contour for isentropic convecting vortex.

IV. Conclusions

A framework to examine the stability of finite-difference methods with strong (exact) boundary conditions is developed and used to derive provably time-stable boundary stencils for hyperbolic (inviscid) systems on uniform grid. The stencils allow stable long-time simulation of systems that previously required additional stabilization measures or a weak implementation of boundary conditions. The framework is then used to derive time-stable boundary stencils for cut-cell grids. The derived stencils do not have the small-cell problem, commonly encountered with cut-cell methods.

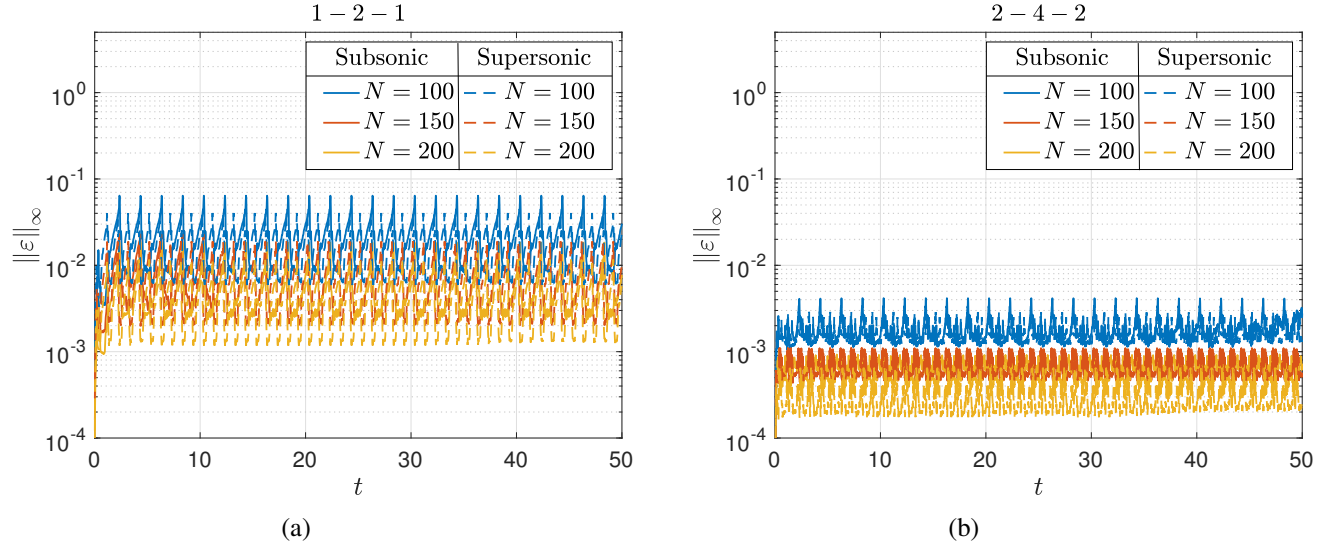


Fig. 16 L_∞ -error from a long-time simulation of the Euler equations for the convecting vortex problem at subsonic and supersonic convective velocity using $N \times N$ grid points with (a) 1-2-1 and (b) 2-4-2 scheme.

For simulations in higher dimensions, a dimensionally split approach is used, which highly simplifies the implementation of the method and is computationally efficient. Several linear and non-linear inviscid tests confirm the stability and robustness of the approach.

Acknowledgements

This work was supported by the US Department of Energy through the Los Alamos National Laboratory. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001). Discussions with Dr. Harsha Nagarajan on using the global optimization solver Alpine is gratefully acknowledged.

Appendix

A. Proof of conditions (13) and (14)

Matrix \tilde{Q} with entries satisfying $q_{ij} = -q_{ji}$, for $i \neq j$, yields

$$\frac{\tilde{Q} + \tilde{Q}^T}{2} = \text{diag} (q_{11}, \dots, q_{\beta\beta}, \dots, q_{nn}), \quad (61)$$

whose substitution in (9), with $\mathbf{q}_0 = [q_{10} \ \dots \ q_{\beta 0} \ 0 \ \dots \ 0]^T$, provides

$$-\tilde{\mathbf{u}}^T (\tilde{Q} + \tilde{Q}^T) \tilde{\mathbf{u}} - 2\tilde{\mathbf{u}}^T \mathbf{q}_0 g = - \sum_{i=1}^n 2q_{ii} u_i^2 - \sum_{i=1}^{\beta} 2q_{i0} u_i g \quad (62)$$

$$= \sum_{i=1}^{\beta} \left[-2q_{ii} \left(u_i + \frac{q_{i0}}{2q_{ii}} g \right)^2 + \frac{q_{i0}^2}{2q_{ii}} g^2 \right] - \sum_{i=\beta+1}^n 2q_{ii} u_i^2 \leq K_1 g^2, \quad (63)$$

where the last inequality holds for $q_{ii} > 0$ if $1 \leq i \leq \beta$ and $q_{ii} \geq 0$ if $\beta < i \leq n$ (the conditions in (13)), and $K_1 = \sum_{i=1}^{\beta} \frac{q_{i0}^2}{2q_{ii}}$. This proves the first statement of the theorem.

Substituting the values of (14) in (63), and using $\sum_{i=1}^{\beta} q_{i0} = -1$, provides

$$-\tilde{\mathbf{u}}^T \left(\tilde{Q} + \tilde{Q}^T \right) \tilde{\mathbf{u}} - 2\tilde{\mathbf{u}}^T \mathbf{q}_0 g \leq g^2, \quad (64)$$

which ensures time-stability. The following shows that (14) also satisfies the conservation statement (12).

The rows of a derivative approximation, D , sum to zero and, hence, the rows of HD also sum to zero. It provides, using the relation between HD and the elements of \mathbf{q}_0 and \tilde{Q} defined in (8),

$$\sum_{j=0}^n q_{ij} = q_{i0} + q_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n q_{ij} = 0 \quad \forall 1 \leq i \leq n, \quad (65)$$

where $q_{i0} = 0$ if $i > \beta$. Using $q_{ij} = -q_{ji}$ for $i \neq j$ yields

$$\sum_{\substack{j=1 \\ j \neq i}}^n q_{ij} = - \sum_{\substack{j=1 \\ j \neq i}}^n q_{ji} \quad \forall 1 \leq i \leq n. \quad (66)$$

Adding $-q_{ii}$ to both sides of (66) and using (65) provides

$$- \sum_{j=1}^n q_{ji} = \sum_{j=1}^n q_{ij} - 2q_{ii} = -q_{i0} - 2q_{ii} \quad \forall 1 \leq i \leq n. \quad (67)$$

To satisfy (12), then, $q_{ii} = -\frac{1}{2}q_{i0}$ if $1 \leq i < n$ and $q_{ii} = \frac{1}{2} - \frac{1}{2}q_{i0}$ if $i = n$. But, since $q_{i0} = 0$ if $i > \beta$, $q_{ii} = 0$ if $\beta < i < n$ and $q_{ii} = \frac{1}{2}$ if $i = n$, which are the parameter values in (14). This completes the proof.

B. Proof of condition (30)

The individual terms in summations of (29), that denote the contribution from each equation of the system, are given by

$$\frac{d}{dt} (\tilde{\mathbf{u}}^\phi)^T H \tilde{\mathbf{u}}^\phi = \frac{d}{dt} \|\tilde{\mathbf{u}}^\phi\|_H^2 = -\lambda_\phi (\tilde{\mathbf{u}}^\phi)^T \left(\tilde{Q} + \tilde{Q}^T \right) \tilde{\mathbf{u}}^\phi - 2\lambda_\phi (\tilde{\mathbf{u}}^\phi)^T \mathbf{q}_0 \left(L \tilde{\mathbf{u}}_0^{II} \right)_\phi, \quad 1 \leq \phi \leq k, \quad (68)$$

$$\frac{d}{dt} (\tilde{\mathbf{u}}^\phi)^T H^\# \tilde{\mathbf{u}}^\phi = \frac{d}{dt} \|\tilde{\mathbf{u}}^\phi\|_{H^\#}^2 = -\lambda_\phi (\tilde{\mathbf{u}}^\phi)^T \left(\tilde{Q}^\# + \left(\tilde{Q}^\# \right)^T \right) \tilde{\mathbf{u}}^\phi - 2\lambda_\phi (\tilde{\mathbf{u}}^\phi)^T \mathbf{q}_0^\# \left(R \tilde{\mathbf{u}}_n^I \right)_\phi, \quad k+1 \leq \phi \leq r, \quad (69)$$

where $\tilde{\mathbf{u}}_0^{II} = \left[u_0^{k+1}(t) \quad u_0^{k+2}(t) \quad \dots \quad u_0^r(t) \right]^T$ and $\tilde{\mathbf{u}}_n^I = \left[u_n^1(t) \quad u_n^2(t) \quad \dots \quad u_n^k(t) \right]^T$. Assuming $q_{ij} = -q_{ji}$, for $i \neq j$ in matrix \tilde{Q} , the contribution to (29) from the first term in r.h.s. of (68) and (69) can be calculated from, respectively,

$$\sum_{\phi=1}^k (\tilde{\mathbf{u}}^\phi)^T \left(\tilde{Q} + \tilde{Q}^T \right) \tilde{\mathbf{u}}^\phi = 2 \sum_{i=1}^n q_{ii} \sum_{\phi=1}^k \left(u_i^\phi \right)^2 = 2 \sum_{i=1}^n q_{ii} \|\tilde{\mathbf{u}}_i^I\|^2, \quad (70)$$

$$\sum_{\phi=k+1}^r (\tilde{\mathbf{u}}^\phi)^T \left(\tilde{Q}^\# + \left(\tilde{Q}^\# \right)^T \right) \tilde{\mathbf{u}}^\phi = -2 \sum_{i=1}^n q_{ii} \sum_{\phi=k+1}^r \left(u_{n-i}^\phi \right)^2 = -2 \sum_{i=1}^n q_{ii} \|\tilde{\mathbf{u}}_{n-i}^{II}\|^2, \quad (71)$$

where $\|\tilde{\mathbf{u}}_i^I\|^2 = \sum_{\phi=1}^k (u_i^\phi)^2$ and $\|\tilde{\mathbf{u}}_{n-i}^{II}\|^2 = \sum_{\phi=k+1}^r (u_{n-i}^\phi)^2$. Further, assuming $\mathbf{q}_0 = [q_{10} \ \dots \ q_{\beta 0} \ 0 \ \dots \ 0]^T$, the contribution to (29) from the second term in r.h.s. of (68) and (69) can be estimated from, respectively,

$$\sum_{\phi=1}^k (\tilde{\mathbf{u}}^\phi)^T \mathbf{q}_0 (L\tilde{\mathbf{u}}_0^{II})_\phi = \sum_{i=1}^{\beta} q_{i0} \sum_{\phi=1}^k u_i^\phi (L\tilde{\mathbf{u}}_0^{II})_\phi, \quad (72)$$

$$\sum_{\phi=k+1}^r (\tilde{\mathbf{u}}^\phi)^T \mathbf{q}_0^\# (R\tilde{\mathbf{u}}_n^I)_\phi = - \sum_{i=1}^{\beta} q_{i0} \sum_{\phi=k+1}^r u_{n-i}^\phi (R\tilde{\mathbf{u}}_n^I)_\phi. \quad (73)$$

Using

$$\sum_{\phi=1}^k u_i^\phi (L\tilde{\mathbf{u}}_0^{II})_\phi \leq \|\tilde{\mathbf{u}}_i^I\| \|L\| \|\tilde{\mathbf{u}}_0^{II}\| \quad \text{and} \quad \sum_{\phi=k+1}^r u_{n-i}^\phi (R\tilde{\mathbf{u}}_n^I)_\phi \leq \|\tilde{\mathbf{u}}_{n-i}^{II}\| \|R\| \|\tilde{\mathbf{u}}_n^I\| \quad (74)$$

in (72) and (73), respectively, and, in turn, using (68)-(69) with (70)-(73) in (29), assuming $q_{ii} \geq 0$ for $\beta < i < n$, it can be shown

$$\frac{dE}{dt} \leq \left\{ \sum_{i=1}^{\beta} (-2q_{ii} \|R\| \|\tilde{\mathbf{u}}_i^I\|^2 + 2|q_{i0}| \|L\| \|R\| \|\tilde{\mathbf{u}}_i^I\| \|\tilde{\mathbf{u}}_0^{II}\|) - 2q_{nn} \|L\| \|\tilde{\mathbf{u}}_0^{II}\|^2 \right\} \quad (75)$$

$$+ \left\{ \sum_{i=1}^{\beta} (-2q_{ii} \|L\| \|\tilde{\mathbf{u}}_{n-i}^{II}\|^2 + 2|q_{i0}| \|L\| \|R\| \|\tilde{\mathbf{u}}_n^I\| \|\tilde{\mathbf{u}}_{n-i}^{II}\|) - 2q_{nn} \|R\| \|\tilde{\mathbf{u}}_n^I\|^2 \right\}. \quad (76)$$

The time-stability condition (29) is satisfied if both curly brackets in (76) are non-positive. Introducing $\sum_{i=1}^{\beta} a_i = 1$, where $a_i > 0$, the last term in the curly brackets can be written as

$$2q_{nn} \|L\| \|\tilde{\mathbf{u}}_0^{II}\|^2 = 2 \sum_{i=1}^{\beta} a_i q_{nn} \|L\| \|\tilde{\mathbf{u}}_0^{II}\|^2 \quad \text{and} \quad 2q_{nn} \|R\| \|\tilde{\mathbf{u}}_n^I\|^2 = 2 \sum_{i=1}^{\beta} a_i q_{nn} \|R\| \|\tilde{\mathbf{u}}_n^I\|^2. \quad (77)$$

Substituting (77) in (76), it can be shown that $dE/dt \leq 0$ if

$$q_{ii} \geq \frac{q_{i0}^2}{4q_{nn}a_i} \|L\| \|R\|, \quad 1 \leq i \leq \beta. \quad (78)$$

This proves the first statement of the theorem.

Appendix A showed that (14) with $\sum_{i=1}^{\beta} q_{i0} = -1$, where $q_{i0} \leq 0$, satisfies the discrete conservation statement (12) for the scalar advection equation. As mentioned before, the discrete conservation statement for the system (15)-(16) is same as that for the scalar advection equation. Therefore, a stencil satisfying (14) provides a conservative scheme for the system (15)-(16). It remains to be shown that the stencil also satisfies the stability condition (29).

Using (14) with $\sum_{i=1}^{\beta} q_{i0} = -1$ and $a_i = -q_{i0}$ in (78) yields the condition

$$1 \geq \|L\| \|R\|, \quad (79)$$

which is satisfied from (21). This completes the proof.

C. 2 – 4 – 2 scheme for uniform grid

$$\begin{array}{llll}
h_{11} = 1.117853598033634 & h_{22} = 1.734954607723689 & h_{33} = 0.493492831348563 & h_{44} = 1.153698962894113 \\
d_{10} = -0.558055563977424 & d_{20} = -0.177806646597481 & d_{30} = 0.197577181565075 & d_{40} = 0.053103321910167 \\
d_{11} = 0.206193447640676 & d_{21} = -0.148032843241780 & d_{31} = -0.349146497048670 & d_{41} = 0.031031686127352 \\
d_{12} = 0.229753040942520 & d_{22} = 0.010938409310223 & d_{32} = -0.469159274307636 & d_{42} = -0.272872172147738 \\
d_{13} = 0.154135831102631 & d_{23} = 0.133448297494816 & d_{33} = 0.026584989564182 & d_{43} = -0.326375382961636 \\
d_{14} = -0.032026755708402 & d_{24} = 0.181452783034222 & d_{34} = 0.763007924163851 & d_{44} = 0.009492491845307 \\
d_{15} = 0 & d_{25} = 0 & d_{35} = -0.168864323936802 & d_{45} = 0.577851491687484 \\
d_{16} = 0 & d_{26} = 0 & d_{36} = 0 & d_{46} = -0.072231436460936
\end{array}$$

D. 2 - 4 - 2 scheme for cut-cell grid

$$\begin{array}{ll}
h_{11} = \frac{18\alpha^5 + 117\alpha^4 + 270\alpha^3 + 270\alpha^2 + 117\alpha + 17}{24(9\alpha^2 + 9\alpha + 2)} & h_{22} = \frac{-42\alpha^5 - 177\alpha^4 - 78\alpha^3 + 294\alpha^2 + 275\alpha + 59}{24(9\alpha^2 + 9\alpha + 2)} \\
h_{33} = \frac{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 186\alpha^2 + 191\alpha + 43}{24(9\alpha^2 + 9\alpha + 2)} & h_{44} = \frac{-6\alpha^5 - 15\alpha^4 - 6\alpha^3 + 222\alpha^2 + 221\alpha + 49}{24(9\alpha^2 + 9\alpha + 2)} \\
d_{10} = -\frac{36(\alpha + 1)^2(2\alpha + 1)}{18\alpha^5 + 117\alpha^4 + 270\alpha^3 + 270\alpha^2 + 117\alpha + 17} & d_{11} = \frac{6(\alpha + 1)(3\alpha + 2)}{18\alpha^5 + 117\alpha^4 + 270\alpha^3 + 270\alpha^2 + 117\alpha + 17} \\
d_{12} = \frac{-18\alpha^5 - 45\alpha^4 + 126\alpha^3 + 378\alpha^2 + 279\alpha + 59}{36\alpha^5 + 234\alpha^4 + 540\alpha^3 + 540\alpha^2 + 234\alpha + 34} & d_{13} = -\frac{2(3\alpha + 1)(3\alpha + 2)}{18\alpha^5 + 117\alpha^4 + 270\alpha^3 + 270\alpha^2 + 117\alpha + 17} \\
d_{14} = \frac{3(6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 6\alpha^2 - 5\alpha - 1)}{36\alpha^5 + 234\alpha^4 + 540\alpha^3 + 540\alpha^2 + 234\alpha + 34} & d_{15} = 0 \quad d_{16} = 0 \\
d_{20} = -\frac{36\alpha^2(2\alpha + 1)}{42\alpha^5 + 177\alpha^4 + 78\alpha^3 - 294\alpha^2 - 275\alpha - 59} & d_{21} = \frac{-18\alpha^5 - 45\alpha^4 + 126\alpha^3 + 378\alpha^2 + 279\alpha + 59}{2(42\alpha^5 + 177\alpha^4 + 78\alpha^3 - 294\alpha^2 - 275\alpha - 59)} \\
d_{22} = -\frac{6\alpha(3\alpha + 1)}{42\alpha^5 + 177\alpha^4 + 78\alpha^3 - 294\alpha^2 - 275\alpha - 59} & d_{23} = \frac{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 258\alpha^2 + 263\alpha + 59}{-84\alpha^5 - 354\alpha^4 - 156\alpha^3 + 588\alpha^2 + 550\alpha + 118} \\
d_{24} = \frac{2\alpha(12\alpha^4 + 30\alpha^3 + 12\alpha^2 - 3\alpha - 1)}{42\alpha^5 + 177\alpha^4 + 78\alpha^3 - 294\alpha^2 - 275\alpha - 59} & d_{25} = 0 \quad d_{26} = 0 \\
d_{30} = 0 & d_{31} = \frac{2(3\alpha + 1)(3\alpha + 2)}{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 186\alpha^2 + 191\alpha + 43} \\
d_{32} = -\frac{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 258\alpha^2 + 263\alpha + 59}{60\alpha^5 + 150\alpha^4 + 60\alpha^3 + 372\alpha^2 + 382\alpha + 86} & d_{33} = 0 \\
d_{34} = \frac{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 258\alpha^2 + 263\alpha + 59}{60\alpha^5 + 150\alpha^4 + 60\alpha^3 + 372\alpha^2 + 382\alpha + 86} & d_{35} = -\frac{2(3\alpha + 1)(3\alpha + 2)}{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 186\alpha^2 + 191\alpha + 43} \\
d_{36} = 0 & d_{40} = 0 \\
d_{41} = \frac{3(6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 6\alpha^2 - 5\alpha - 1)}{2(6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 222\alpha^2 - 221\alpha - 49)} & \\
d_{42} = -\frac{2\alpha(12\alpha^4 + 30\alpha^3 + 12\alpha^2 - 3\alpha - 1)}{6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 222\alpha^2 - 221\alpha - 49} & d_{43} = \frac{30\alpha^5 + 75\alpha^4 + 30\alpha^3 + 258\alpha^2 + 263\alpha + 59}{2(6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 222\alpha^2 - 221\alpha - 49)}
\end{array}$$

$$d_{44} = 0$$

$$d_{45} = -\frac{16(3\alpha + 1)(3\alpha + 2)}{6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 222\alpha^2 - 221\alpha - 49}$$

$$d_{46} = \frac{2(3\alpha + 1)(3\alpha + 2)}{6\alpha^5 + 15\alpha^4 + 6\alpha^3 - 222\alpha^2 - 221\alpha - 49}$$

References

- [1] Gustafsson, B., *High order difference methods for time dependent PDE*, Vol. 38, Springer Science & Business Media, 2007.
- [2] Carpenter, M. H., Gottlieb, D., and Abarbanel, S., “Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes,” *Journal of Computational Physics*, Vol. 111, No. 2, 1994, pp. 220–236.
- [3] Olsson, P., “Summation by parts, projections, and stability. I,” *Mathematics of Computation*, Vol. 64, No. 211, 1995, pp. 1035–1065.
- [4] Olsson, P., “Summation by parts, projections, and stability. II,” *Mathematics of Computation*, Vol. 64, No. 212, 1995, pp. 1473–1493.
- [5] Thompson, K. W., “Time dependent boundary conditions for hyperbolic systems,” *Journal of computational physics*, Vol. 68, No. 1, 1987, pp. 1–24.
- [6] Poinot, T. J., and Lele, S., “Boundary conditions for direct simulations of compressible viscous flows,” *Journal of computational physics*, Vol. 101, No. 1, 1992, pp. 104–129.
- [7] Lodato, G., Domingo, P., and Vervisch, L., “Three-dimensional boundary conditions for direct and large-eddy simulation of compressible viscous flows,” *Journal of Computational Physics*, Vol. 227, No. 10, 2008, pp. 5105–5143.
- [8] Svård, M., Carpenter, M. H., and Nordström, J., “A stable high-order finite difference scheme for the compressible Navier–Stokes equations, far-field boundary conditions,” *Journal of Computational Physics*, Vol. 225, No. 1, 2007, pp. 1020–1038.
- [9] Svård, 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.
- [10] Nordström, J., Gong, J., Van der Weide, E., and Svård, 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.
- [11] Sharan, N., “Time-stable high-order finite difference methods for overset grids,” Ph.D. thesis, University of Illinois at Urbana-Champaign, 2016.
- [12] Sharan, N., Pantano, C., and Bodony, D. J., “Time-stable overset grid method for hyperbolic problems using summation-by-parts operators,” *Journal of Computational Physics*, Vol. 361, 2018, pp. 199–230.
- [13] Brady, P. T., and Livescu, D., “High-order, stable, and conservative boundary schemes for central and compact finite differences,” *Computers & Fluids*, Vol. 183, 2019, pp. 84–101.
- [14] Clarke, D. K., Hassan, H., and Salas, M., “Euler calculations for multielement airfoils using Cartesian grids,” *AIAA journal*, Vol. 24, No. 3, 1986, pp. 353–358.
- [15] Ingram, D. M., Causon, D. M., and Mingham, C. G., “Developments in Cartesian cut cell methods,” *Mathematics and Computers in Simulation*, Vol. 61, No. 3-6, 2003, pp. 561–572.
- [16] Berger, M., and Aftosmis, M., “Progress towards a Cartesian cut-cell method for viscous compressible flow,” *50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, 2012, p. 1301.

- [17] Aftosmis, M. J., “Solution adaptive Cartesian grid methods for aerodynamic flows with complex geometries,” *VKI Lecture Series*, Vol. 2, 1997, p. 1997.
- [18] Pattinson, J., Malan, A., and Meyer, J., “A cut-cell non-conforming Cartesian mesh method for compressible and incompressible flow,” *International Journal for Numerical Methods in Engineering*, Vol. 72, No. 11, 2007, pp. 1332–1354.
- [19] Schneiders, L., Hartmann, D., Meinke, M., and Schröder, W., “An accurate moving boundary formulation in cut-cell methods,” *Journal of Computational Physics*, Vol. 235, 2013, pp. 786–809.
- [20] Muralidharan, B., and Menon, S., “A high-order adaptive Cartesian cut-cell method for simulation of compressible viscous flow over immersed bodies,” *Journal of Computational Physics*, Vol. 321, 2016, pp. 342–368.
- [21] Mohammadi-Aragh, M., Klingbeil, K., Brüggemann, N., Eden, C., and Burchard, H., “The impact of advection schemes on restratification due to lateral shear and baroclinic instabilities,” *Ocean Modelling*, Vol. 94, 2015, pp. 112–127.
- [22] Sharan, N., Matheou, G., and Dimotakis, P. E., “Mixing, scalar boundedness, and numerical dissipation in large-eddy simulations,” *Journal of Computational Physics*, Vol. 369, 2018, pp. 148–172.
- [23] Sharan, N., Brady, P., and Livescu, D., “High-order energy-stable boundary treatment for finite-difference cut-cell method,” *Bulletin of the American Physical Society*, 2019.
- [24] Brady, P. T., and Livescu, D., “Stable, High-Order and Conservative Cut-Cell Methods,” *AIAA Scitech 2019 Forum*, 2019, p. 1991.
- [25] Abarbanel, S. S., Chertock, A. E., and Yefet, A., “Strict stability of high-order compact implicit finite-difference schemes: the role of boundary conditions for hyperbolic PDEs, II,” *Journal of Computational Physics*, Vol. 160, No. 1, 2000, pp. 67–87.
- [26] Strand, B., “Summation by Parts for Finite Difference Approximations for d/dx ,” *Journal of Computational Physics*, Vol. 110, No. 1, 1994, pp. 47–67.
- [27] Svärd, M., “On coordinate transformations for summation-by-parts operators,” *Journal of Scientific Computing*, Vol. 20, No. 1, 2004, pp. 29–42.
- [28] Gustafsson, B., “The convergence rate for difference approximations to mixed initial boundary value problems,” *Mathematics of Computation*, Vol. 29, No. 130, 1975, pp. 396–406.
- [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] Inc., W. R., “Mathematica, Version 12.0,” , ??? URL <https://www.wolfram.com/mathematica>.
- [31] Nagarajan, H., Lu, M., Wang, S., Bent, R., and Sundar, K., “An adaptive, multivariate partitioning algorithm for global optimization of nonconvex programs,” *Journal of Global Optimization*, 2019. doi:10.1007/s10898-018-00734-1.
- [32] Abarbanel, S. S., Chertock, A. E., and Yefet, A., “Strict stability of high-order compact implicit finite-difference schemes: the role of boundary conditions for hyperbolic PDEs, II,” *Journal of Computational Physics*, Vol. 160, No. 1, 2000, pp. 67–87.
- [33] Carpenter, M. H., Gottlieb, D., and Abarbanel, S., “The stability of numerical boundary treatments for compact high-order finite-difference schemes,” *Journal of Computational Physics*, Vol. 108, No. 2, 1993, pp. 272–295.
- [34] Cook, A. W., and Riley, J. J., “Direct numerical simulation of a turbulent reactive plume on a parallel computer,” *Journal of Computational Physics*, Vol. 129, No. 2, 1996, pp. 263–283.
- [35] Perot, J. B., “Discrete conservation properties of unstructured mesh schemes,” *Annual review of fluid mechanics*, Vol. 43, 2011, pp. 299–318.