A HYBRID FINITE DIFFERENCE–FINITE VOLUME APPROACH TO SOLVE FIRST-ORDER HYPERBOLIC CONSERVATION LAWS WITH SUPERIOR ACCURACY

XIANYI ZENG

Abstract. A hybrid finite difference–finite volume (FD-FV) approach for discretization in space is proposed to solve first-order hyperbolic conservation laws. Unlike any conventional finite difference method (FDM) or finite volume method (FVM), this approach uses both cell-averaged values and nodal values as degrees of freedom (DOF). Consequently it is inherently conservative like FVM and easy to extend to high-order accuracy in space like FDM. The proposed FD-FV approach works for arbitrary flux functions, whether convex or non-convex; and it does not require any exact or approximate Riemann solver hence it is also computationally economical. Method of lines is adopted for time integration in present work; in particular, explicit Runge-Kutta methods are employed. It is theoretically proven and numerically confirmed that in general, the proposed FD-FV methods possess superior accuracy than conventional FDM or FVM. Linear stability is studied for general FD-FV schemes – both space-accurate and time-stable FD-FV schemes of up to fifth-order accuracy in both space and time are presented. Numerical examples show that as long as the solutions are smooth, the proposed FD-FV methods are more efficient than conventional FVM of the same order, at least when explicit time-integrators are applied.

Key words. first-order hyperbolic conservation law, high-order accuracy, hybrid finite difference–finite volume approach, linear stability, superior spatial accuracy

AMS subject classifications. 65M06, 65M08, 65M12, 65M20, 35L65

1. Introduction. Hyperbolic conservation laws (HCL) have wide applications in areas like computational fluid dynamics (CFD), computational aeroacoustics (CAA), and magnetohydrodynamics (MHD). Since analytical solutions to most HCLs are difficult to construct, effective numerical methods become indispensable tools for both theoretical analysis and applications. Three desired properties of practical numerical schemes for such problems are: (1) conservation, which ensures correct shock speed; (2) high-order accuracy when the solution is smooth; and (3) stability. In addition, versatility is also preferred, which means that the method should work for general flux function that meets minimum mathematical requirements, yet still provides correct solutions. Most popular numerical techniques, like finite difference methods (FDM), finite volume methods (FVM) and finite element methods (FEM), have their own strength and weakness in resolving these issues. This work, on the other hand, makes an attempt towards constructing versatile numerical schemes such that all these goals can be naturally achieved.

Finite volume methods, arguably the most popular technique in practice for HCLs, is naturally conservative. Usually, these methods are of Godunov type, which requires an exact or approximate Riemann solver. These Riemann solvers often highly depend on the specific flux function, and can be expensive to compute, especially when the flux function is non-convex. In addition, since these methods typically operate on cell-averaged values, high-order accuracy in space is difficult to achieve. Most practical FVMs are only second-order accurate, for example, the monotone upstream-centered schemes for conservation laws (MUSCL) and flux-corrected transport (FCT) and the Kolgan’s method. This is partially due to the fact that cell-averaged value can be treated as a second-order approximation to

*Department of Aeronautics and Astronautics, Stanford University, Stanford, California, 94305-4035 (rhodusz@stanford.edu).
the nodal value at cell centroid \[21, 20\]. However, to construct higher than second-order FVMs, one must be very careful to treat the cell-averaging nature of discrete data properly, which leads to further complication. Some examples are the piecewise parabolic method (PPM) [9], the k-exact method [3, 27] and essentially non-oscillatory (ENO) [17] or weighted essentially non-oscillatory (WENO) [25] schemes.

On the contrary, finite difference method, being numerical methods based on nodal values and conventional Taylor series expansions, are easy to be designed to high-order accuracy. In addition, since most discrete operators of FDM are linear, von Neumann analysis can be applied for preliminary study of the linear stability when designing a scheme. However, these methods often lack inherent conservation and additional measures must be taken [15, 6, 36] to ensure the correct shock speed.

Finally, the third type of methods, FEMs, are not originally designed for advection dominated problems, like most HCLs of concern. To account for instability caused by advection, two approaches are often employed to enhance or modify the original Galerkin type methods. The first approach introduces additional terms [18, 31] to stabilize the whole system, without modifying the Galerkin framework. These methods are conservative, as long as constant function is included in the weighting functional space [2]. The other approach, represented by the discontinuous Galerkin methods [8], requires numerical fluxes like FVMs at element interfaces, thus they depend on the particular physical flux function.

In this work, a hybrid approach is proposed for spatial discretization, such that it combines the advantages of both FDM and FVM in the sense that the method is inherently conservative like FVMs as well as easy to extend to high-order accuracy like FDMs. In addition, the new approach is suitable for general flux functions, convex or not, and it requires no numerical flux functions. To this end, both cell-averaged values and nodal values at cell faces are counted as degrees of freedom in the proposed FD-FV approach. Method of lines is adopted for time integration; in particular, explicit Runge-Kutta time stepping methods are employed in this work although it can be extended to implicit methods like conventional FDMs and FVMs. For discretizing the cell-averaged values, fluxes at cell faces are computed as exact flux functions evaluated at nodal values. To discretize the nodal values in space, polynomial reconstruction taking into account of both adjacent nodal values and cell-averaged values is performed. This reconstruction leads to a linear discrete differential operator (DDO) at cell faces, which enables the von Neumann stability analysis to the proposed FD-FV schemes.

The present paper focuses on 1D case. It is proven theoretically and verified numerically that the proposed FD-FV approach generally has superior accuracy property compared to conventional finite difference methods providing the same order of polynomial reconstruction. This property is demonstrated by numerical example to show that it also carry to 2D case on Cartesian grids. Thus the proposed methods are potentially more computational efficient than conventional numerical methods to solve HCLs.

Some related works in literature are briefly discussed in the following. These methods, like the present one, make use of DOFs with different mathematical meanings as dependent variables. The constrained interpolation profile method (CIP) [42] and the interpolated differential operator method (IDO) [1, 30] have spatial derivatives and nodal values as DOFs. They employs Hermite interpolation for local data reconstruct. However, these methods are more suitable for elliptic equations and parabolic equations in which case the solutions are more regular than HCLs. The
enhanced versions to handle conservation laws, naming CIP-CSL [41] (CSL stands for conservative semi-Lagrangian) and IDO-CF [19] (CF stands for conservation form), are able to handle certain hyperbolic equations like the Euler equations and incompressible Navier-Stokes equations in conservation form. These enhanced versions are closely related to the present approach, in the sense that they include cell-averaged values into the DOFs and discard all the spatial derivatives. However, compared to CIP-CSL and IDO-CF, the current approach is more general and it is suitable for all flux functions. In addition, to the best knowledge of the author, provable superior accuracy in space is never shown before for alike numerical schemes.

To this effect, the remainder of the paper is organized as follows. Section 2 presents the general framework of the proposed approach for one-dimensional equations as well as a simple extension to two-dimensional case. The convergence property in space is studied in Section 3. This section also shows that in most cases, the FD-FV methods have superior accuracy property than conventional FDMs providing the same order polynomial reconstruction. Next, Section 4 studies the linear stability property of both the semi-discretized and the fully discretized forms of the proposed methods; in the latter case explicit Runge-Kutta time-integrators are employed for simplicity. Practical FD-FV schemes that are both high-order accurate and linearly stable are presented in Section 5. Numerical examples that confirm the theoretical results are offered in Section 6, where a number of benchmark problems are solved to show both the accuracy and the robustness of the new approach. Finally, Section 7 concludes this paper.

2. A hybrid finite-difference and finite-volume (FD-FV) approach. Here and throughout the remainder of the paper, the following notations are used:

- The normal fonts \((x, y, t, u, \text{ etc.})\) designate scalar variables; and the bolded fonts \((x, w, f, \text{ etc.})\) designate vector-valued variables.
- Unless otherwise noted, the subscripts \(x, y\) mean partial derivatives in space, and the subscript \(t\) indicates time derivative.
- \((x, t)\) designates the exact value of a variable \(†\) evaluated at the location \(x\) (in Cartesian coordinates) and time \(t\).
- The overline \(\overline{\cdot}\) indicates the cell-averaged value of a variable \(†\) over a control volume \(C\); whereas in absence of the overline, nodal value is indicated.

2.1. General FD-FV formulation for one-dimensional hyperbolic conservation laws. Consider the system of first-order hyperbolic conservation laws (hyperbolic conservation laws or HCL for short throughout the remainder of this paper) in one space dimension:

\[ w_t + f(w)_x = 0 \]  \hspace{1cm} (2.1)

where \(w \in \mathbb{R}^d\) is a real-valued vector and \(f : \mathbb{R}^d \to \mathbb{R}^d\) is a differentiable function of \(w\). Denote the Jacobian matrix of the flux function as \(J(w) = \partial f / \partial w\) and it is supposed to be known analytically.

Consider the fixed computational domain \(\Omega = [a, b]\) and a uniform mesh with \(N\) control volumes: \(\Omega = \cup_{i=1}^{N} \mathcal{C}_i, \mathcal{C}_i = [a + (i - 1)h, a + ih],\) where \(h = (b - a)/N\) is the cell size. The cell centers are denoted by \(x_i: x_i = a + (i - 1/2)h,\) \(i = 1, 2, \cdots, N;\) and the cell faces by \(x_{i+1/2}: x_{i+1/2} = a + ih,\) \(i = 0, 1, \cdots, N.\)

The proposed FD-FV methods seek the numerical approximations to both the nodal values at cell faces, \(w_{i+1/2}^n \approx w(x_{i+1/2}, t^n);\) and approximations to the cell-averaged
values:
\[ \bar{w}_i^n \approx \frac{1}{|C_i|} \int_{C_i} w(x, t^n) dx \]  
(2.2)

Here \(|C_i| = x_{i+1/2} - x_{i-1/2} = h\) is the size of the control volume \(C_i\).

Method of lines is adopted in present work, which means Eq. (2.1) is first discretized in space to obtain an ODE system for the time-dependent variables:

\[ \bar{w}_i(t) \approx \frac{1}{|C_i|} \int_{C_i} w(x, t) dx \]  
(2.3)

The resulting ODE system is then solved by any preferred ODE solver. In this paper, the explicit Runge-Kutta methods are employed. Here and throughout the remainder of the paper, \(\bar{w}_i^n, \bar{w}_i^{n+1/2}\) and \(\hat{w}_i, \hat{w}_i^{n+1/2}\) designate approximated cell-averaged and nodal solutions to the fully discretized system and semi-discretized system, respectively. In the latter case, the dependence in \(t\) is omitted for convenience.

To derive the semi-discretized formula for cell-averaged values, the governing equation is integrated over each cell \(C_i\) to obtain:

\[ \frac{d}{dt} \left( \frac{1}{|C_i|} \int_{C_i} w(x, t) dx \right) + \frac{1}{|C_i|} \left[ f \left( w(x_{i+1/2}, t) \right) - f \left( w(x_{i-1/2}, t) \right) \right] = 0 \]

which suggests the following semi-discretized scheme for \(\bar{w}_i\):

\[ \frac{d\bar{w}_i}{dt} + \frac{1}{|C_i|} \left[ f \left( \bar{w}_{i+1/2} \right) - f \left( \bar{w}_{i-1/2} \right) \right] = 0 \]  
(2.4)

Note that Eq. (2.4) is exact in the sense that there is no truncation error associated with this semi-discretization formula. On the other hand, to obtain the semi-discretized formula for nodal values, Eq. (2.1) is written in non-conservative form at a cell face \(x_{i+1/2}\):

\[ w_i(x_{i+1/2}, t) + J \left( w(x_{i+1/2}, t) \right) w_x(x_{i+1/2}, t) = 0 \]

The following numerical scheme is proposed here for solving \(w_{i+1/2}\):

\[ \frac{dw_{i+1/2}}{dt} + J \left( \dot{w} \right) [D_x \dot{w}]_{i+1/2} = 0, \quad \dot{w} = w_{i+1/2} \]  
(2.5)

\(\dot{w}\) does not have to be chosen as \(w_{i+1/2}\), but it is the simplest and the most natural choice and is adopted in this paper. This choice is also shown to simplify a number of issues when nonlinear problem is concerned (see Section 5). \(D_x\) is the discrete differential operator that characterizes the FD-FV scheme, which takes into account both nearby nodal values and cell-averaged values.

In this paper, linear \(D_x\) is considered so that von Neumann stability analysis applies. It can be expressed in general form as follows:

\[ [D_x w]_{i+1/2} = \frac{1}{h} \sum_{l=-q}^{q} \alpha_l \bar{w}_{i+l} + \frac{1}{h} \sum_{l=-q}^{q} \beta_l \bar{w}_{i+1/2+l} \]  
(2.6)

Here \(\alpha_l\) and \(\beta_l\) are constants, and \(q\) indicates the stencil of the numerical scheme. Note that Eq. (2.6) is constructed to a desired order of accuracy based on Taylor
In the 1D case, Eq. (2.12) is not exact, because edge-averaged fluxes are replaced by flux

\[ w(x_{i+1/2}, t) = w(x_{i+1/2}, t) + \sum_{m=1}^{\infty} \frac{(lh)^m}{m!} \partial^m_x w(x_{i+1/2}, t) \]  

(2.7)

and for cell-averaged values, the following expansion is employed:

\[ \bar{w}_{i+1}(t) = \frac{1}{h} \int_{x_{i+1/2} + (l-1)h}^{x_{i+1/2} + lh} w(x, t) dx \]

\[ = \frac{1}{h} \int_{x_{i+1/2} + (l-1)h}^{x_{i+1/2} + lh} w(x_{i+1/2}, t) + \sum_{m=1}^{\infty} \frac{(x - x_{i+1/2})^m}{m!} \partial^m_x w(x_{i+1/2}, t) \]  

(2.8)

In theory, Eqs. (2.7, 2.8) enable one to construct \( D_x \) to arbitrary order of accuracy in space.

**Remark:** In the special case \( \alpha_l = 0, \forall l \), Eq. (2.6) is the conventional finite difference operators. Thus in this case, the proposed FD-FV methods reduce to classical FDM, with cell-averaged values being numerical artifacts.

### 2.2. A simple extension to two dimensions on Cartesian meshes.

The methodology can be easily extended to solve 2D hyperbolic conservation laws in the case of Cartesian meshes. The governing equation is:

\[ w_t + f(w)_x + g(w)_y = 0 \]  

(2.9)

Let \( J = \partial f / \partial w \) and \( K = \partial g / \partial w \) be the Jacobian matrices for the flux functions in \( x \)- and \( y \)-directions, and they are supposed to be known analytically.

Consider a Cartesian mesh with cells \( C_{i,j} = [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}] \), where the cell center of \( C_{i,j} \) is \((x_i, y_j) = ((x_{i-1/2} + x_{i+1/2})/2, (y_{j-1/2} + y_{j+1/2})/2) \). In the semi-discretized form, the cell-averaged dependent variables are naturally defined as:

\[ \bar{w}_{i,j}(t) \approx \frac{1}{|C_{i,j}|} \int_{C_{i,j}} w(x, t) dx \]  

(2.10)

To reuse the DDO proposed in the 1D case, the nodal values at the following locations are chosen in this work:

\[ w_{i,j+1/2}(t) \approx w(x_i, y_{j+1/2}, t), \quad w_{i+1/2,j}(t) \approx w(x_{i+1/2}, y_j, t) \]  

(2.11)

The locations of the dependent variables are illustrated in Figure 2.1.

The cell-averaged variables are advanced in time as follows:

\[ \frac{d\bar{w}_{i,j}}{dt} + \frac{f(\bar{w}_{i+1/2,j}) - f(\bar{w}_{i-1/2,j})}{h_x} + \frac{g(\bar{w}_{i,j+1/2}) - g(\bar{w}_{i,j-1/2})}{h_y} = 0 \]  

(2.12)

Here \( h_x \) and \( h_y \) are edge lengths in \( x \)- and \( y \)-directions of the control volumes. Unlike the 1D case, Eq. (2.12) is not exact, because edge-averaged fluxes are replaced by flux
functions evaluated at edge centers, which introduces second-order error in space. Higher-order fluxes can be obtained, for example, using nearby nodal values on the same axis. However, Eq. (2.12) is adopted for numerical examples in this paper, for the reason that it is sufficient to show the superior accuracy property (to be presented in Section 3) of the proposed FD-FV schemes in 2D case. FD-FV schemes with higher-order fluxes in multiple dimensions will be addressed in future work.

The semi-discretized formula for nodal values are given by the following:

\[
\frac{d}{dt}w_{i+1/2,j} + J(w_{i+1/2,j})[D_xw]_{i+1/2,j} + K(w_{i+1/2,j})[D_yw]_{i+1/2,j} = 0
\] (2.13)

\[
\frac{d}{dt}w_{i,j+1} + J(w_{i,j+1})[D_xw]_{i,j+1} + K(w_{i,j+1})[D_yw]_{i,j+1} = 0
\] (2.14)

Here \(D_x\) and \(D_y\) are the 1D discrete differential operators developed in previous section. On the other hand, \(\tilde{D}_x\) and \(\tilde{D}_y\) are classical discrete operators in FDMs that only uses nodal values. For example, using first-order one-sided operators, one has

\[
[\tilde{D}_y^\text{Forward}w]_{i+1/2,j} = \frac{1}{h_y} (w_{i+1/2,j+1} - w_{i+1/2,j})
\]

or \([\tilde{D}_y^\text{Backward}w]_{i+1/2,j} = \frac{1}{h_y} (w_{i+1/2,j} - w_{i+1/2,j-1})\)

\[
[\tilde{D}_x^\text{Forward}w]_{i,j+1/2} = \frac{1}{h_x} (w_{i+1,j+1/2} - w_{i,j+1/2})
\]

or \([\tilde{D}_x^\text{Backward}w]_{i,j+1/2} = \frac{1}{h_x} (w_{i,j+1/2} - w_{i-1,j+1/2})\)

Finally, before going into analysis of the accuracy and stability properties of the proposed method, a few remarks should be addressed:

**Remark 1:** This section provides the general framework of the FD-FV schemes. To construct stable and accurate FD-FV methods to solve hyperbolic problems, adequate upwinding property (or more precisely, linear stability) must be considered. These are provided in Section 5 where the practical FD-FV schemes are presented.
Remark 2: The 2D FD-FV scheme framework presented is a simple one. Nevertheless, it is sufficient to illustrate the superior accuracy property (see Section 3) in 2D case by this simple framework as shown in Section 6.

Remark 3: In the 2D case, locations of dependent variables other than the one in Figure 2.1 can be explored. For example, one can choose \((x_{i+1/2}, y_{j+1/2})\) to be the points where nodal values are evaluated. A discussion of the advantages and disadvantages of each of these configurations is outside the scope of this paper, and will be addressed in future work.

Remark 4: Some related works [41, 19] introduce edge-averaged values along control volume edges as dependent variables. However, since the flux function and edge-averaging operator typically don’t commute, one still needs to: (1) approximate edge-averaged fluxes with edge-averaged conservative variables; and (2) develop governing equations for the edge-averaged variables. To avoid such complications, edge-averaged variables are not considered in this work.

3. Accuracy analysis. Consider applying the scheme given by Eqs. (2.4–2.5) to solve 1D scalar advection equation with constant propagation velocity:

\[ u_t + cu_x = 0 \]  \hspace{1cm} (3.1)

The flux function is given by \( f(u) = cu \), and the Jacobian is \( J(u) = c \). The accuracy and stability properties of the proposed FD-FV scheme are studied by applying Eqs. (2.4–2.5) to solve Eq. (3.1) with simple wave solution.

To this effect, the discrete operator \( \mathcal{D}_x \) that characterizes the FD-FV scheme, which is given by Eq. (2.6), is divided into two operators \( \mathcal{A} \) and \( \mathcal{N} \) as follows:

\[
\begin{align*}
\mathcal{A}w_{i+1/2} &= \frac{1}{h} \sum_{l=-q+1}^{q} \alpha_l \mathcal{W}_{i+l}, \\
\mathcal{N}w_{i+1/2} &= \frac{1}{h} \sum_{l=-q}^{q} \beta_l \mathcal{W}_{i+1/2+l}
\end{align*}
\]  \hspace{1cm} (3.2)

And it follows:

\[
[\mathcal{D}_x w]_{i+1/2} = [\mathcal{A}w]_{i+1/2} + [\mathcal{N}w]_{i+1/2}
\]

Note that no subscript \( x \) is introduced to \( \mathcal{A} \) or \( \mathcal{N} \), because they do not necessarily constitute a consistent discrete differential operator that approximates first-order spatial derivative at \( x_{i+1/2} \). And it is shown later that, this is the key to the superior accuracy property of the proposed FD-FV schemes.

3.1. Application of the proposed schemes to 1D advection equation with simple wave solutions. Consider the simple wave solution to Eq. (3.1) with wave number \( k \):

\[ u(x, t) = e^{ik(x-ct)} \]  \hspace{1cm} (3.3)

In the following, the superscript \( * \) designates the exact values corresponding to (3.3). To avoid confusion, \( i \) is reserved for imaginary unit in the remainder of this section, and \( j \) will be used as a generic subscript that indicates a particular control volume.

It is also supposed that the mesh is uniform with cell size \( h \), and \( x_j = jh, x_{j+1/2} = (j + 1/2)h \). Then \( \mathcal{W}_j(t) \) and \( u_{j+1/2}^*(t) \), the exact cell-averaged value and nodal value at time \( t \), are given by:

\[
\begin{align*}
\mathcal{W}_j^* &= \frac{1}{h} \int_{(j-1/2)h}^{(j+1/2)h} e^{ik(x-ct)} \, dx = \frac{1}{i\theta} e^{-ickt} e^{ij\theta} \left( e^{i\theta/2} - e^{-i\theta/2} \right) \hspace{1cm} (3.4) \\
u_{j+1/2}^* &= e^{-ickt} e^{i(j+1/2)\theta}
\end{align*}
\]  \hspace{1cm} (3.5)
Here \( \theta = kh \). Define the following quantities for convenience:

\[
A^\star(t) = N^\star(t) = e^{-ickt}, \quad R_j(\theta) = \frac{1}{\theta} e^{i\theta} \left(e^{i\theta/2} - e^{-i\theta/2}\right), \quad S_{j+1/2}(\theta) = e^{i(j+1/2)\theta}
\]

Then the exact solutions (3.4–3.5) are:

\[
\begin{align*}
\pi^\star_j(t) &= A^\star(t) R_j(\theta), \quad u^\star_j(t) = N^\star(t) S_{j+1/2}(\theta) \\
\end{align*}
\]

(3.6)

Next apply the FD-FV scheme (2.4–2.5) to solve Eq. (3.1). It follows that the ODE systems obtained from discretization in space for dependent variables \( \pi_j \) and \( u_{j+1/2} \) are:

\[
\begin{align*}
\frac{d\pi_j}{dt} + \frac{c}{\theta} (u_{j+1/2} - u_{j-1/2}) &= 0 \\
\frac{du_{j+1/2}}{dt} + c [A u]_{j+1/2} + c [N u]_{j+1/2} &= 0
\end{align*}
\]

(3.7)

(3.8)

Here \( A \) and \( N \) are given by Eqs. (3.2).

To find the solutions to the ODE systems (3.7–3.8) with exact initial data:

\[
\begin{align*}
\pi_j(0) &= R_j(\theta), \quad u_{j+1/2}(0) = S_{j+1/2}(\theta)
\end{align*}
\]

(3.9)

It is supposed first and verified afterwards that the solutions are given by:

\[
\begin{align*}
\pi_j &= A(t) R_j(\theta), \quad u_{j+1/2} = N(t) S_{j+1/2}(\theta)
\end{align*}
\]

(3.10)

with \( A(0) = N(0) = 1 \). Note that both \( A(t) \) and \( N(t) \) also depend on \( \theta \). However, to simplify the notations and also to be consistent with the vocabulary “ODE” that follows, the dependence on \( \theta \) is omitted but should be kept in mind. Then it follows that \( A(\cdot) \) and \( N(\cdot) \), which are approximations to \( A^\star(\cdot) \) and \( N^\star(\cdot) \), satisfy the ODEs:

\[
\begin{align*}
A'(t) + i c k N(t) &= 0 \\
N'(t) + c k a(\theta) A(t) + c k b(\theta) N(t) &= 0
\end{align*}
\]

(3.11)

Here \( a(\cdot) \) and \( b(\cdot) \) are functions of \( \theta \):

\[
a(\theta) = \frac{1 - e^{-i\theta}}{i \theta^2} \sum_{l=-q+1}^{q} e^{i\theta \alpha_l}, \quad b(\theta) = \frac{1}{\theta} \sum_{l=-q}^{q} e^{i\theta \beta_l}
\]

(3.12)

The ODEs (3.10–3.11) are obtained by inserting Eqs. (3.9) to Eqs. (3.7–3.8) and applying the following equalities:

\[
\begin{align*}
i \theta R_j(\theta) &= S_{j+1/2}(\theta) - S_{j-1/2}(\theta), \quad S_{j+1/2+1}(\theta) = S_{j+1/2}(\theta) e^{i\theta} \\
R_{j+1}(\theta)/S_{j+1/2}(\theta) &= \frac{1}{i \theta} e^{i\theta} (1 - e^{-i\theta})
\end{align*}
\]

The ODE system (3.10–3.11) is the key for accuracy and stability analysis. Its analytical solution is given by:

\[
\begin{bmatrix}
A(t) \\
N(t)
\end{bmatrix} = e^{-c k C(\theta) t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C(\theta) = \begin{bmatrix} 0 & i \\ a(\theta) & b(\theta) \end{bmatrix}
\]

(3.13)
The eigenvalue decomposition of the coefficient matrix $C(\theta)$ is:

$$C(\theta) = \begin{bmatrix} i & i \\ \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \frac{\lambda_2}{i(\lambda_2-\lambda_1)} & -\frac{1}{\lambda_2-\lambda_1} \\ \frac{1}{\lambda_2-\lambda_1} & \frac{\lambda_1}{i(\lambda_2-\lambda_1)} \end{bmatrix}$$  \tag{3.14}

where the eigenvalues are functions of $\theta$:

$$\lambda_1(\theta) = \frac{1}{2} \left( b(\theta) + \sqrt{b(\theta)^2 + 4ia(\theta)} \right), \quad \lambda_2(\theta) = \frac{1}{2} \left( b(\theta) - \sqrt{b(\theta)^2 + 4ia(\theta)} \right)$$  \tag{3.15}

Consequently, the solutions are expressed analytically as:

$$A(t) = \frac{\lambda_2 - i}{\lambda_2 - \lambda_1} e^{-ck\lambda_1 t} - \frac{\lambda_1 - i}{\lambda_2 - \lambda_1} e^{-ck\lambda_2 t}$$  \tag{3.16}

$$N(t) = \frac{\lambda_1(\lambda_2 - i)}{i(\lambda_2 - \lambda_1)} e^{-ck\lambda_1 t} - \frac{\lambda_2(\lambda_1 - i)}{i(\lambda_2 - \lambda_1)} e^{-ck\lambda_2 t}$$  \tag{3.17}

Note that the dependence on $\theta$ in $\lambda_{1,2}$ is omitted here and throughout the remainder of this section for convenience.

Since $R_j(\theta)$ and $S_{j+1/2}(\theta)$ are known, it is easy to verify that Eqs. (3.16–3.17) lead to the exact solutions to the semi-discretized problem (3.7–3.8) through Eqs. (3.9).

Comparing with Eqs. (3.4–3.5), the order of spatial accuracy of the FD-FV scheme, or the expected convergence rate in space, is defined as follows:

**Definition 3.1.** The FD-FV scheme defined by the DDO $D_x$ is $p$th-order ($p \geq 1$) accurate in space, if:

$$\frac{A(t)}{A^*(t)} = 1 + c_a(t)\theta^p + O(\theta^{p+1}), \quad \frac{N(t)}{N^*(t)} = 1 + c_n(t)\theta^p + O(\theta^{p+1})$$  \tag{3.18}

Here $c_a$ and $c_n$ are functions of time that are independent of $\theta$, and at least one of them is not zero.

Comparing the exponential terms in Eqs. (3.16–3.17) and noticing that $A^*(t) = N^*(t) = e^{-ickt}$, one of the two eigenvalues $\lambda_{1,2}$ is a numerical approximation to the imaginary unit $i$, whereas the other one is spurious. Without loss of generality, $\lambda_1$ is chosen to be close to $i$ and $\lambda_2$ is supposed to be spurious.

The designed order of accuracy of the DDO $D_x$ is defined as follows:

**Definition 3.2.** The DDO $D_x$ given by Eq. (2.6) is $p$th-order ($p \geq 1$) accurate in space, if:

$$[D_x v^*]_{j+1/2} = \partial_x v^*_{j+1/2} + c_p \partial_x^{p+1} v^*_{j+1/2} h^p + O(h^{p+1})$$  \tag{3.19}

for any sufficiently smooth function $v^*(x,t)$. Here $c_p$ is a non-zero constant that is independent of the function $v^*$ and cell size $h$.

In the case of conventional FDMs (or $A \equiv 0$ as explained before), the numerical scheme usually has the same order of accuracy as $D_x$. However, it is to be shown next that this is not the case for most FD-FV methods proposed in this paper.
3.2. Accuracy analysis of \( D_x \) and the corresponding FD-FV scheme. In order to study the accuracy of the FD-FV scheme that is determined by the operator \( D_x \), the relationship between the coefficients \( \alpha_l \), \( \beta_l \) and the order of \( D_x \) is first established by the following proposition:

**Proposition 3.3.** The DDO \( D_x \) given by Eq. (2.6) is \( p \)th-order (\( p \geq 1 \)) accurate with leading error coefficient \( c_p \neq 0 \) if and only if the constants \( \alpha_l \), \( -q + 1 \leq l \leq q \) and \( \beta_l \), \( -q \leq l \leq q \) satisfy:

\[
\begin{align*}
\sum_{l=-q+1}^{q} \alpha_l + \sum_{l=-q}^{q} \beta_l &= 0 \quad (3.20a) \\
\sum_{l=-q+1}^{q} \frac{2l-1}{2} \alpha_l + \sum_{l=-q}^{q} l \beta_l &= 1 \quad (3.20b) \\
\sum_{l=-q+1}^{q} \frac{\ell^{m+1} - (l-1)^{m+1}}{(m+1)!} \alpha_l + \sum_{l=-q}^{q} \frac{l^m}{m!} \beta_l &= 0, \quad m = 2, \cdots, p \quad (3.20c) \\
\sum_{l=-q+1}^{q} \frac{\ell^{p+2} - (l-1)^{p+2}}{(p+1)!} \alpha_l + \sum_{l=-q}^{q} \frac{\ell^{p+1}}{(p+1)!} \beta_l &= c_p \quad (3.20d)
\end{align*}
\]

The proof is straightforward: the Taylor series expansions (2.7–2.8) of cell-averaged variables and nodal variables are plugged into Eq. (2.6), then Eqs. (3.20) are obtained according to Eq. (3.19).

Defining the following quantities:

\[
\begin{align*}
a_m &= \frac{1}{(m+1)!} \sum_{l=-q+1}^{q} (l^{m+1} - (l-1)^{m+1}) \alpha_l, \quad b_m &= \frac{1}{m!} \sum_{l=-q}^{q} l^m \beta_l, \quad m = 0, 1, \cdots
\end{align*}
\]

then Eqs. (3.20) are simplified to:

\[
a_0 + b_0 = 0, \quad a_1 + b_1 = 1, \quad a_m + b_m = 0, \quad m = 2, \cdots, p \quad a_{p+1} + b_{p+1} = c_p \quad (3.21)
\]

In addition, from Eqs. (3.12), one obtains by Taylor series expansions of \( e^{i\theta} \) the following expressions of \( a(\theta) \) and \( b(\theta) \) in terms of \( a_m \) and \( b_m \):

\[
\begin{align*}
a(\theta) &= \frac{a_0}{\theta} + \sum_{m=0}^{\infty} \frac{\ell^{m+1} a_{m+1} \theta^m}{\ell^{m+1}} \quad (3.22) \\
b(\theta) &= \frac{b_0}{\theta} + \sum_{m=0}^{\infty} \frac{\ell^{m+1} b_{m+1} \theta^m}{\ell^{m+1}} \quad (3.23)
\end{align*}
\]

Suppose \( D_x \) is designed to be \( p \)th-order accurate in space, so that Eqs. (3.21) hold, it follows:

\[
a(\theta) + b(\theta) = i + r(\theta)\theta^p, \quad r(\theta) = \sum_{m=0}^{\infty} \frac{\ell^{m+p+1} (a_{m+p+1} + b_{m+p+1}) \theta^m}{\ell^{m+p+1}}
\]

To compute the two eigenvalues of \( C(\theta) \) using Eqs. (3.15), the term in the square root operation is first evaluated:

\[
b^2 + 4ia = b^2 + 4i(i - b + r\theta^p) = (2i - b)^2 + 4ir\theta^p = \left(2i - b + \frac{2ir\theta^p}{2i - b}\right)^2 + \frac{4r^2\theta^{2p}}{(2i - b)^2}
\]
Note that it is assumed here that $b(\theta) \neq 2i$, which happens in the special case $b_0 = 0$, $b_1 = 2$, $b_m = 0$, $m \geq 2$ according to Eq. (3.23). The last term in the parentheses has a crucial impact on the accuracy of the FD-FV scheme, and it is expressed as follows by applying Eq. (3.23) and the leading terms of $r(\theta)$:

$$\frac{2i\theta b}{2i - b} = \frac{2i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} + O(\theta^{p+2-s_\beta})$$

(3.24)

To study the FD-FV scheme’s spatial order of accuracy, the index $s_\beta$ is defined:

**Definition 3.4.** Given the operator $N$, its index $s_\beta$ is defined by:

$$s_\beta = \begin{cases} 
0 & b_0 \neq 0 \\
1 & b_0 = 0, b_1 \neq 2 \\
s \geq 2 & b_0 = 0, b_1 = 2, b_2 = \cdots = b_{s-1} = 0, b_s \neq 0
\end{cases}$$

(3.25)

Consequently, the leading term of the denominator of Eq. (3.24) is of order $O(\theta^{s_\beta})$. Hence one establishes:

$$\frac{2i\theta b}{2i - b} = \frac{2i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} + O(\theta^{p+2-s_\beta})$$

(3.26)

Here $\chi_{s_\beta = 1}$ takes the value 1 if $s_\beta = 1$ and 0 otherwise. Thus the square root in Eqs. (3.15) is:

$$\sqrt{b^2 + 4ia} = 2i - b + \frac{2i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} + O(\theta^{p+2-s_\beta}) + O(\theta^{1-s_\beta})$$

(3.27)

The last term accounts for the contribution of $\frac{1i^{2\theta b^2}}{2i - b} = O(\theta^{2p+2-2s_\beta})$ in the expansion of $b^2 + 4ia$. In particular, suppose $2i - b + \frac{2i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} = O(\theta^{1-s_\beta})$, one has:

$$\min(\hat{p} + \hat{q}, 2\hat{p}) = 2p + 2 - 2s_\beta$$

Note that the particular choice of the square root from the two available values in Eq. (3.27) is due to the consideration that $\lambda_1$ should approximate $i$, see the discussion following Definition 3.1. To this end, by applying Eq. (3.27), one obtains from Eqs. (3.15):

$$\lambda_1 = i + \frac{i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} + O(\theta^{\min(p+2-s_\beta, \hat{p})})$$

(3.28)

$$\lambda_2 = \frac{b_0}{\theta} - i + \sum_{m=0}^{\infty} i^{m+1} b_{m+1} \theta_m - \frac{i^{p+2-s_\beta} \theta_{p+1-s_\beta}}{2 \chi_{s_\beta = 1} - b_{s_\beta}} \theta_{p+1-s_\beta} + O(\theta^{\min(p+2-s_\beta, \hat{p})})$$

(3.29)

Note that $\lambda_1$ at best approximates $i$ to $(p + 1 - s_\beta)$th-order accuracy in space. Thus to maximize the accuracy of $\lambda_1$, $s_\beta$ should be zero, that is, $b_0 \neq 0$. This leads to the major result in this paper:

**Theorem 3.5.** If $s_\beta = 0$, or equivalently $b_0 = \sum_{i=-q}^{q} \beta_i \neq 0$, a pth-order ($p \geq 1$) accurate operator $D_x$ leads to a $(p + 1)$th-order accurate FD-FV scheme provided also...
that \( cb_0 > 0 \).

Note that the Definitions [3.1][3.2] are used here. They do not guarantee convergence of the underlying FD-FV schemes unless stability is also addressed, which is provided later.

**Proof.** Given \( s_\beta = 0 \), there is \( \chi(s_\beta=1) = 0 \) and consequently:

\[
2i - b + \frac{2i^{p+2-s_\beta}c_p}{2\chi(s_\beta=1) - b_{s_\beta}} \theta^{p+1-s_\beta}
\]

\[
= 2i - \left( \frac{b_0}{\theta} + \sum_{m=0}^{\infty} \frac{i^{m+1}b_{m+1}\theta^m}{\theta} \right) - \frac{2i^{p+2}c_p}{b_0} \theta^{p+1} = O(\theta^{-1})
\]

Thus \( \hat{q} = -1 \), and \( \hat{p} \) satisfies \( \min(\hat{p} - 1, 2\hat{p}) = 2p + 2 \), or equivalently \( \hat{p} = 2p + 3 \). Then Eqs. (3.28)(3.29) reduce to:

\[
\lambda_1 = i - \frac{i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2}), \quad \lambda_2 = \frac{b_0}{\theta} + (b_1 - 1)i + O(\theta)
\]

(3.30)

Noticing that:

\[
\lambda_2 - \lambda_1 = -\sqrt{b^2 + 4ia} = -2i + b + \frac{2i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2}) = \frac{b_0}{\theta} + O(1)
\]

and following Eqs. (3.16)(3.17) one has:

\[
\frac{A(t)}{A^*(t)} = \left( 1 + \frac{\lambda_1 - i}{\lambda_2 - \lambda_1} \right) e^{-ck(\lambda_1-i)t} - \frac{\lambda_1 - i}{\lambda_2 - \lambda_1} e^{-ck(\lambda_2-i)t}
\]

(3.31)

\[
\frac{N(t)}{N^*(t)} = \left( 1 + \frac{\lambda_2(\lambda_1 - i)}{i(\lambda_2 - \lambda_1)} \right) e^{-ck(\lambda_2-i)t} - \frac{\lambda_2(\lambda_1 - i)}{i(\lambda_2 - \lambda_1)} e^{-ck(\lambda_2-i)t}
\]

(3.32)

The leading terms of every single entity are evaluated:

\[
\frac{\lambda_1 - i}{\lambda_2 - \lambda_1} = -\left( \frac{i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2}) \right) = -\frac{i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2})
\]

\[
\frac{\lambda_2(\lambda_1 - i)}{i(\lambda_2 - \lambda_1)} = \left( \frac{b_0}{i\theta} + O(1) \right) \left( -\frac{i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2}) \right) = -\frac{i^{p+2}c_p}{b_0} \theta^{p+1} + O(\theta^{p+2})
\]

\[
e^{-ck(\lambda_1-i)t} = e^{-ck[t-(i^{p+2}c_p/b_0)\theta^{p+1}+O(\theta^{p+2})]} = 1 + \frac{i^{p+2}c_pkt}{b_0} \theta^{p+1} + O(\theta^{p+2})
\]

\[
e^{-ck(\lambda_2-i)t} = e^{-ck[t/b_0+(b_1-2)i+O(\theta)]} = e^{\frac{-cb_0}{2}t} e^{-ick(b_1-2)t} (1 + O(\theta))
\]

Since \( h > 0 \), one has \( e^{-cb_0 t/h} < 1 \) due to the assumption that \( cb_0 > 0 \). Hence \( e^{-ck(\lambda_2-i)t} = O(1) \) and it goes to zero exponentially as \( h \to 0 \). Thus the spurious modes associated with \( \lambda_2 \) can be safely removed from Eqs. (3.31)(3.32) for polynomial order-of-accuracy analysis, and they are expressed as \( \epsilon(\theta) \) in the following formula.
Finally, it is obtained:

$$A(t) = \left( 1 - \frac{i^{p+2}c_p \theta^{p+2}}{b_0} + O(\theta^{p+3}) \right) \left( 1 + \frac{i^{p+2}c_p k t \theta^{p+1}}{b_0} + O(\theta^{p+2}) \right) + \epsilon(\theta)$$

$$N(t) = \left( 1 - \frac{i^{p+2}c_p \theta^{p+1}}{b_0} + O(\theta^{p+2}) \right) \left( 1 + \frac{i^{p+2}c_p k t \theta^{p+1}}{b_0} + O(\theta^{p+2}) \right) + \epsilon(\theta)$$

which completes the proof that the corresponding FD-FV scheme is $(p + 1)$th-order accurate in space according to Definition 3.1.

Remark 1: The key to the superior accuracy property demonstrated by this theorem is $s_\beta = 0$ or $b_0 = \sum_{l=-q}^{q} \beta_l \neq 0$. This explains why it can not be expected for conventional finite difference schemes: FDMs correspond to $D_x$ that satisfies $A \equiv 0$, in which case $a_0 = \sum_{l=-q}^{q} \alpha_l = 0$. Thus the consistency condition (the first equation in Eqs. (3.21)) enforces $b_0 = 0$, hence $s_\beta > 0$ and the order-of-accuracy of the scheme is at best the designed order-of-accuracy of the discrete differential operator.

Remark 2: In the case $s_\beta = 0$, the spurious modes associated with $A_2$ has a factor $e^{-cb_0 t/h}$, which can either decay or grow exponentially as the mesh is refined, depending on the sign of $cb_0$. Thus it is crucial to design $D_x$ such that $b_0$ has the same sign as $c$, the propagation velocity. This suggests that for conservation law systems, the characteristic form of the governing equation should be employed, which requires the system to be hyperbolic. Note that the condition $cb_0 > 0$ can be regarded as a mathematical extension (or explanation) of the “upwind” concept for solving advection dominated problems.

Remark 3: The motivation to use FD-FV schemes that has $s_\beta = 0$ is not solely due to the proven superior accuracy property in space. A potential danger of designing FD-FV scheme with $s_\beta \neq 0$ such that $a_0 = b_0 = 0$ is: in this case, Eqs. (2.4-2.5) allow the “constant” solutions in the form $\overline{w}_j = \overline{w}_a$, $\forall j$, $\overline{w}_{j+1/2} = \overline{w}_b$, $\forall j$ where $\overline{w}_a$ and $\overline{w}_b$ are arbitrary constants. Since $a_0 = b_0 = 0$, there is no mechanism in the numerical scheme that enforces $\overline{w}_a = \overline{w}_b$. Thus in the case $s_\beta \neq 0$, the FD-FV scheme allows unphysical constant solutions.

Throughout the remainder of this paper, it is always assumed that $b_0 \neq 0$, hence Theorem 3.5 applies. Several such $D_x$ applied at $x_{j+1/2}$ with up to fourth-order accuracy in space are listed in Table 3.1. In the table, “F-biased” and “B-biased” stands for forward-biased and backward-biased respectively.

4. Von Neumann Stability analysis. This section studies the linear stability of the proposed FD-FV schemes in the case $b_0 \neq 0$. To this end, the stability of the solutions to the ODE systems (3.7-3.8) due to discretization in space is first studied in Section 4.1. Next, time-integrators are considered in Section 4.2 where asymptotic behavior when $h \to 0$ is studied. This asymptotic analysis only requires the knowledge of $b_0$ of the spatial operator $D_x$. Finally, sample fully discretized FD-FV schemes are studied in Section 4.3.
4.1. Stability analysis of the semi-discretized form in the case $b_0 \neq 0$. Suppose $c > 0$, then the “forward” and “backward” directions in Table 3.1 correspond to the “downwind” and “upwind” directions respectively.

Looking at Eqs. (3.16)–(3.17), there are five sources of numerical errors:

- **Dispersion**: the computed wave number is $\tilde{\omega} = Re(k\lambda_1/i)$; or more conveniently one may study $\tilde{\theta} = Re(\lambda_1 \theta/i)$. $\tilde{\theta}$ is related to $\theta$ by the first equation in Eqs. (3.30). Their difference represents the error in propagation speed of the simple wave.

- **Dissipation**: the exact solution does not damp or amplify in time; but the numerical solution may do. This error is measured by $\varepsilon = Im(\lambda_1 \theta/i)$. Given $c > 0$, a negative $\varepsilon$ is desired since in this case the magnitude of the numerical solution does not blow up as time increases.

- **Stationary dispersion**: this phase error is associated with the arguments of time-independent quantities in front of $e^{-ck(\lambda_i-i)t}$ in Eqs. (3.16)–(3.17). The corresponding errors in the cell-averaged values and nodal values are:

\[
\varphi_a = \arg \left( \frac{\lambda_2 - i}{\lambda_2 - \lambda_1} \right), \quad \varphi_n = \arg \left( \frac{\lambda_1(\lambda_2 - i)}{i(\lambda_2 - \lambda_1)} \right)
\]

Note that the effect of these terms in the solution is independent of time, but only depends on the mesh size.

- **Stationary dissipation**: this error is associated with the magnitude of the same quantities above. And they are defined by:

\[
\varphi_a = \left| \frac{\lambda_2 - i}{\lambda_2 - \lambda_1} \right|, \quad \varphi_n = \left| \frac{\lambda_1(\lambda_2 - i)}{i(\lambda_2 - \lambda_1)} \right|
\]

Again, these quantities only affect the solution via mesh sizes, no matter how long the ODEs are integrated.

- **Noise**: unlike most FDMs, the proposed FD-FV methods have a spurious mode in the solution and it is called “noise” in this paper. This quantity is measured by the magnitude of the second terms in Eqs. (3.16)–(3.17):

\[
\nu = \left| \frac{\lambda_2(\lambda_1 - i)}{i(\lambda_2 - \lambda_1)} e^{-ck(\lambda_2-i)t} \right|
\]
Fig. 4.1. Dispersion ($\tilde{\theta}$, left) and dissipation ($\varepsilon$, right) properties of selected $\mathcal{D}_x$

These quantities are more difficult to evaluate since they are functions of both wave number and time. However, as described before, $\lambda_2$ is a function of space only and the leading term is $b_0/\theta$; thus providing $c b_0 > 0$, as required by Theorem 3.5 the exponential terms in $\nu$ and $\varphi$ are expected to decay exponentially with increasing $t$ and deceasing $\theta$. In the plots provided later for selected $\mathcal{D}_x$ in Table 3.1 the time is fixed to $t = h/c$, which corresponds to one time step size computed with Courant number 1.0.

Next, these errors are plotted against $\theta$ for the operators listed in Table 3.1. Note, however, only the $\mathcal{D}_x$ with $b_0 > 0$ are chosen due to Theorem 3.5 and the assumption that $c > 0$. In particular, the chosen operators are the following: $\mathcal{D}_{x,up}^1$ (1st-order backward), $\mathcal{D}_{x,up}^2$ (2nd-order backward), $\mathcal{D}_{x,up-biased}^3$ (3rd-order B-biased), $\mathcal{D}_{x,up}^4$ (4th-order B-biased) and $\mathcal{D}_{x,up}^4$ (4th-order backward).

The dispersion relations between $\tilde{\theta}$ and $\vartheta$ are provided in the left of Figures 4.1 the dispersion is plotted in the right of Figures 4.1. The four stationary quantities $\tilde{\varphi}_a$, $\tilde{\varphi}_n$ and $\varphi_n$ are plotted as functions of $\theta$ in Figures 4.2. The noises $\nu$ and $\varphi$ at fixed time $t = h/c$ are plotted as functions of $\theta$ in Figures 4.3.

Note that the $\theta$-abscissa is chosen to be in the range $[0, \pi]$, in which case the largest value corresponds to one cell (or two DOFs) per wavelength. The left figure in Figures 4.1 illustrates: high-order $\mathcal{D}_x$ is more accurate than low-order ones; and given the order $p$, the upwind-biased version is more accurate than the upwind version. However, looking at the $\epsilon-\theta$ plot in the right part of Figures 4.1, it is clear that the fully upwind 4th-order accurate $\mathcal{D}_{x,up}^4$ is unstable since $\epsilon > 0$ when $\theta \to \pi$; whereas the 4th-order upwind-biased operator $\mathcal{D}_{x,up-biased}^4$ has no such stability problem. This observation is against common intuition that “upwind” is closely related to positive numerical viscosity and hence it tends to stabilize the numerical scheme. Thus, one must pay special attention not to equalize “upwind” property with linear stability when designing FD-FV schemes, especially when high-order accuracy in space is concerned. Figures 4.2 indicate that the stationary errors do not blow up for the range of $\theta$ of interests and for all $\mathcal{D}_x$ considered. Figures 4.3 show that the noises due to the spurious mode are bounded and close to zero in the range of $\theta$ for all $\mathcal{D}_x$ whereas it is again observed that $\mathcal{D}_{x,up}^4$, the 4th-order fully upwind operator leads to the largest noise among all six operators.

4.2. Asymptotic stability analysis of fully-discretized problem as $h \to 0$.

The ODE system (3.7–3.8) must be supplemented with a time-integrator in order to
construct a practical numerical solver. In this section, several explicit Runge-Kutta methods up to fifth-order accurate in time are chosen to illustrate the linear stability property of selected DDO $D_x$.

The building block of these time-integrators is the forward Euler (FE) method, which is provided below for the general formula (2.4–2.5) given fixed time step size $\Delta t = t^{n+1} - t^n$:

$$
\begin{align*}
\bar{w}_{j}^{n+1} &= w_j^n - \frac{\Delta t}{h} \left[ f(w_{j+1/2}^n) - f(w_{j-1/2}^n) \right] \\
 w_{j+1/2}^{n+1} &= w_{j+1/2}^n - \Delta t J(w_{j+1/2}^n) [D_x w]_{j+1/2}^n
\end{align*}
$$

Fig. 4.2. Stationary quantities: (1) upper-left, dispersion for cell-averaged values ($\bar{\theta}_n$); (2) upper-right, dispersion for nodal values ($\bar{\phi}_a$); (3) lower-left, dissipation for cell-averaged values ($\bar{\nu}_n$); (4) lower-right, dissipation for nodal values ($\bar{\nu}_n$)

Fig. 4.3. Noises associated with spurious mode $\lambda_2$: (1) left, for cell-averaged values ($\bar{\nu}$); (2) right, for nodal values ($\nu$)
Simplify the formula by defining the operator $\mathcal{K}$ as follows:

$$W^{n+1} = W^n + \Delta t K_1, \quad K_1 = \mathcal{K}(W^n)$$

Here $W$ is the collection of variables $w_j$ and $w_{j+1/2}$. Then the considered Runge-Kutta methods are summarized in Table 4.1. Note that the particular Runge-Kutta methods are chosen for the purpose of matching the expected order of accuracy of the FD-FV schemes. The schemes in Table 4.1 are chosen for simplicity in presentation due to the simple coefficients. Optimized schemes, for example, towards stability preserving or low storage considerations, are not explored here.

Before combining with $D_x$ in Table 3.1, asymptotic stability property of each time-integrator in Table 4.1 as $h \to 0$, or equivalently $\theta \to 0$, is studied.

To this end, first consider applying FE method to solve Eqs. (3.10–3.11), which leads to:

$$A^{n+1} = A^n - i\lambda N^n, \quad N^{n+1} = N^n - \lambda (a(\theta)A^n + b(\theta)N^n)$$

Here $\lambda = c\Delta t/h$ is the signed Courant number. Employing the coefficient matrix $C(\theta)$ given in Eqs. (3.13), the equations above become:

$$\begin{bmatrix} A^{n+1} \\ N^{n+1} \end{bmatrix} = (I - \lambda \theta C(\theta)) \begin{bmatrix} A^n \\ N^n \end{bmatrix}$$

Providing eigenvalue decomposition of $C(\theta)$ by Eq. (3.14) and initial condition $A^0 = N^0 = 1$, one has:

$$\begin{bmatrix} A^n \\ N^n \end{bmatrix} = \begin{bmatrix} i & i \\ \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} (1 - \lambda \theta \lambda_1)^n & 0 \\ 0 & (1 - \lambda \theta \lambda_2)^n \end{bmatrix} \begin{bmatrix} \frac{\lambda_2}{\lambda(\lambda_2-\lambda_1)} - \frac{1}{\lambda_2-\lambda_1} \\ \frac{\lambda_1}{\lambda(\lambda_2-\lambda_1)} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

| RK2 | $K_1 = \mathcal{K}(W^n)$  
$K_2 = \mathcal{K}(W^n + \Delta t K_1)$  
$W^{n+1} = W^n + \frac{1}{2} \Delta t (K_1 + K_2)$ |
| RK3 | $K_1 = \mathcal{K}(W^n)$  
$K_2 = \mathcal{K}(W^n + \frac{1}{2} \Delta t K_1)$  
$K_3 = \mathcal{K}(W^n + \frac{1}{2} \Delta t K_2)$  
$W^{n+1} = W^n + \frac{1}{2} \Delta t (K_1 + K_2)$ |
| RK4 | $K_1 = \mathcal{K}(W^n)$  
$K_2 = \mathcal{K}(W^n + \frac{1}{2} \Delta t K_1)$  
$K_3 = \mathcal{K}(W^n + \frac{1}{3} \Delta t K_2)$  
$K_4 = \mathcal{K}(W^n + \frac{1}{3} \Delta t K_3)$  
$W^{n+1} = W^n + \frac{1}{3} \Delta t (K_1 + 2K_2 + 2K_3 + K_4)$ |
| RK5 | $K_1 = \mathcal{K}(W^n)$  
$K_2 = \mathcal{K}(W^n + \Delta t K_1)$  
$K_3 = \mathcal{K}(W^n + \frac{1}{2} \Delta t K_1)$  
$K_4 = \mathcal{K}(W^n + \frac{1}{3} \Delta t K_2)$  
$K_5 = \mathcal{K}(W^n + \frac{1}{3} \Delta t K_3)$  
$W^{n+1} = W^n + \frac{1}{4} \Delta t (7K_1 + 9K_2 + 32K_3 + 12K_4 + 32K_5 + 32K_6)$ |

Table 4.1

Explicit Runge-Kutta methods with up to fifth-order accuracy.
Thus a necessary condition for the method to be stable is:

\[ |1 - \lambda \theta \lambda_{1,2}| \leq 1 \]  

(4.6)

Recalling Eqs. (3.30), in the limit \( \theta \to 0 \), a necessary condition of the Courant number for achieving linearly stability is given by:

\[ |\lambda| \leq \frac{2}{|b_0|} \]

Similar analysis can be conducted for other methods in Table 4.1 and the corresponding bounds are reported in Table 4.2.

Though not mathematically proven, it turns out that for the combinations of \( D_x \) and time-integrators considered in this paper, Table 4.2 provides convenient estimates of the largest allowable Courant number for the given FD-FV scheme to be stable.

### 4.3. Stability property of various FD-FV schemes

The spatial operators studied in Section 4.1 are combined with proper time-integrators provided in Section 4.2 to give full FD-FV schemes. The chosen combinations are provided in the second and third columns in Table 4.3. Note that \( D_x^{4,up} \) is not included, since it is shown to be unstable on the semi-discretized level in Section 4.1. Also note that the time-integrator is chosen to be one-order higher than the designed order of the pairing \( D_x \). This is due to the superior accuracy property proved in Theorem 3.5.

The solution to the fully discretized scheme providing simple wave initial condition can be expressed as:

\[
\begin{bmatrix}
A^n \\
N^n
\end{bmatrix} = P(\lambda \theta C(\theta)) \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\]

see for example Eq. (4.5) in the case of forward Euler time-integrator. Here \( P(\cdot) \) is a polynomial that is solely determined by the time-integrator. The FD-FV scheme is stable if and only if:

\[ |P(\lambda \theta \lambda_{1,2}(\theta))| \leq 1 \]  

(4.7)

Thus to illustrate the stability property of each FD-FV methods listed in Table 4.3 the contours of maximum value \( \max(|P(\lambda \theta \lambda_{1,2}(\theta))|, |P(\lambda \theta \lambda_{1,2}(\theta))|) \) is plotted on the...
These issues are addressed in following sub-sections.

Fig. 4.4. Contours of \( \max\{P(\lambda \theta \lambda_1(\theta))\} \) on \( \theta - \lambda \) plane: (1) upper-left, \( D_2^{1,up-RK2} \); (2) upper-right, \( D_2^{3,up-RK3} \); (3) middle-left, \( D_2^{3,up-biased-RK4} \); (4) middle-left, \( D_2^{3,up-RK4} \); (5) bottom, \( D_2^{4,up-biased-RK5} \)

\( \theta - \lambda \) plane, see Figures 4.4 In the figures, as \( \theta \to 0 \), the contours with value 1 intersect the \( \lambda \)-axis at the values \( \lambda_{0} \), which are provided in Section 4.2 by asymptotic analysis. These values are reported in the fifth column in Table 4.3. The last column provides the maximum allowable Courant number \( \lambda_{max} \) for linear stability of the FD-FV scheme. It is clear by comparing the last two columns that, \( \lambda_{0} \) is a good estimate to \( \lambda_{max} \), and the computation of the former value is much simpler than the latter.

5. Accurate and stable FD-FV schemes in practice. The FD-FV scheme provided by Eqs. (2.4) needs to be polished for practical use. To this end, at least two issues must be addressed:

- First, the stability analysis in previous section indicates that the operator \( D_x \) must be designed to use data in a certain stencil which takes into account of the direction of characteristic velocity. This is not reflected in Eqs. (2.4).
- Second, boundary conditions must be treated properly.

These issues are addressed in following sub-sections.

5.1. Applying DDO to characteristics. Due to stability consideration, \( D_x \) should be applied to characteristics rather than components of \( w \), thus Eq. (2.5)
should be modified, with the aid of the eigenvalue decomposition of $\hat{J} = J(\hat{w})$:

$$\hat{J} = R \hat{\Lambda} L$$

Here $\hat{\Lambda}$ is a real diagonal matrix with diagonal elements $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_d$. $\hat{R} = [\hat{r}_1 \hat{r}_2 \cdots \hat{r}_d]$ and $\hat{L} = \hat{R}^{-1} = [\hat{l}_1 \hat{l}_2 \cdots \hat{l}_d]^t$, where $\hat{r}_i, \hat{l}_i \in \mathbb{R}^d$ are right- and left-eigenvectors associated with $\hat{\lambda}_i, i = 1, 2, \cdots, d$. Consequently Eq. (2.5) can be modified as:

$$\hat{L} \frac{d\hat{w}_{i+1/2}}{dt} + \hat{\Lambda} \left[ D_x (\hat{L} \hat{w}) \right]_{i+1/2} = 0 \quad (5.1)$$

Note that this is a numerical approximation to the linearized equation of Eq. (2.1):

$$\omega_t + \hat{J} \omega_x = 0 \quad (5.2)$$

With the choice $\hat{w} = \omega_{i+1/2}$ (see Section 2), the equation above is exact; in this case, $\hat{J}(\omega_{i+1/2})$ is denoted by $\hat{J}_{i+1/2}$ for short. If $\hat{w}$ is selected differently for a $p$th-order accurate FD-FV scheme, $\hat{J}$ must be constructed carefully such that $\hat{J} = \hat{J}_{i+1/2} + O(h^p)$; usually, this means that $\hat{w}$ should be chosen such that $\hat{w} = \omega_{i+1/2} + O(h^p)$ given sufficiently smooth flux function $f(\cdot)$. This issue does not arise with the choice in this paper.

Eq. (5.1) leads locally to a system of decoupled advection equations:

$$\frac{d\omega^m_{i+1/2}}{dt} + \lambda_m \left[ D_x \omega^m \right]_{i+1/2} = 0, \quad \omega^m = l_m \omega \quad (5.3)$$

Then the chosen $D_x$ in Section 3 are applied to compute approximations to space derivatives of each $\omega^m$. For example, if $D^4_{x,up-biased}$ is chosen, the scheme described by Eq. (5.3) is as follows:

$$\left[ D_x \omega^m \right]_{i+1/2} = \begin{cases} 
[D^4_{x,B-biased} \omega^m]_{i+1/2} & \lambda_m > 0 \\
[D^4_{x,F-biased} \omega^m]_{i+1/2} & \lambda_m < 0 
\end{cases}$$

Here the meaning of the notations in the superscripts of $D_x$ are given in Table 3.1.

### 5.2. Boundary condition treatments.

Implementation of Dirichlet type and Neumann type boundary conditions are considered here for $p$th-order accurate FD-FV scheme and for forward Euler time-stepping method. For simplicity, one-dimensional scalar problem is supposed:

$$\omega_t + f(\omega)_x = 0 \quad (5.4)$$

In addition, the chosen DDO is denoted by $D_x^{-1}$, which is designed to be $(p-1)$th-order accurate. Without loss of generality, suppose the boundary condition at $x_{1/2} = 0$, which is the left end point of the computational domain, is given by:

$$\omega(0,t) = g_d(t), \quad \text{Dirichlet BC} \quad (5.5)$$

or $\omega_x(0,t) = g_n(t), \quad \text{Neumann BC} \quad (5.6)$

First consider the Dirichlet boundary condition (5.5). To enforce this boundary condition at $t^n$ such that the solution at next time step $t^{n+1}$ can be obtained, the following steps are taken:
1. Set $\omega_{1/2}^n = g_d(t^n)$.
2. If for some $i > 0$, $[D_x^{p-1}\omega^n]_{i+1/2}$ requires inputs $\omega_{j+1/2}^n$, $j < 0$ that are not available, a different $D_x^{p-2,i+1/2}$ with: (1) a biased stencil that do not require any $\omega_{j+1/2}^n$, $j < 0$; (2) has $b_0 \neq 0$ and is designed to be $(p-1)$th-order accurate in space, is applied at $x_{i+1/2}$ instead of $D_x^{p-1}$.

For example, if $D_x^{1,\text{Backward}}$, $D_x^2\text{Backward}$ or $D_x^4, B\text{-biased}$ are applied around $x = 0$ in the FD-FV scheme, only the first step is required since none of these operators requires information more than one cell away. On the other hand, if $D_x^3, B\text{-biased}$ or $D_x^4, F\text{-biased}$ is chosen instead, the operator at $x_{i+1/2}$ should be modified to $D_x^4, B\text{-biased}$ and $D_x^4, F\text{-biased}$ respectively.

Next, consider the Neumann type boundary condition (5.6). It is enforced at $t^n$ as follows:

1. Replace $\omega_0^n$ with the value such that $[D_x^{p-1}\text{Forward}\omega^n]_{1/2} = g_n(t^n)$ is satisfied. Here $D_x^{p-1,\text{Forward}}$ is the DDO that is designed to be $(p-1)$th-order accurate, with $b_0 \neq 0$ and only use variables in the forward direction.
2. If for some $i > 0$, $[D_x^{p-1}\omega^n]_{i+1/2}$ requires non-existent variables, the treatment is the same as in the Dirichlet boundary condition case.

The first step can be illustrated by considering, for example, the 4-th order accurate FD-FV scheme ($p = 4$) with operator $D_x^{5, \text{Backward}}$. Then at $t^n$, the value $\omega_{1/2}^n$ is computed, such that:

$$\frac{\omega_2^n - 8\omega_3^n + 17\omega_4^n - 10\omega_5^n}{2h} = [D_x^{3,\text{Forward}}\omega^n]_{1/2} = g_n(t^n)$$

or equivalently

$$\omega_{1/2}^n = \frac{1}{10} \left( \omega_2^n - 8\omega_3^n + 17\omega_4^n - 2h g_n(t^n) \right)$$

**Remark:** In practice, to solve first-order equations like the HCL considered in this paper, the spatial order of accuracy of the numerical scheme can be one-order lower at the boundary without affecting the overall order of accuracy in space [15]. Thus the previous implementations of boundary conditions can be simplified by employing one-order lower $D_x$ at and/or near the boundary.

6. **Numerical examples.** The proposed FD-FV schemes in Table 4.3 are tested by solving various 1D problems. The superior accuracy property is demonstrated by solving 1D advection equations and 1D Euler equations with smooth solutions. Transport problems with discontinuous solutions are also tested, in which case linear stability analysis does not apply. As expect, Gibbs phenomenon appears near the discontinuity; however, the magnitude of the oscillations seems to be independent of the mesh size. This is in contrary to conventional FVMs, in which case solutions blow up quickly as mesh is refined, unless further stabilization measure (like slope limiting or flux limiting) is taken. The last example in 1D shows the versatility of the proposed FD-FV methods, by solving a scalar nonlinear HCL with non-convex flux function. Finally, the first-order operator $D_x^{1,\text{Up}}$ is extended to solve a 2D Euler problem using the framework described in Section 2. In this 2D example, the superior accuracy property proven in Theorem 5.3 is shown to be carried to the 2D case; and performance comparison with MUSCL solves show that the proposed second-order FD-FV scheme is more efficient than conventional second-order FVMs.
For all the tests, the initial condition of cell-averaged value must be computed carefully. One choice is to use the exact cell-averaged data obtained from the initial condition; an alternative method is to use quadrature points to compute high-order accurate cell-averaged values. In particular, if a $p$th-order FD-FV scheme is tested, the initial cell-averaged data is computed to be $p$th-order accurate approximations to the exact value. The former method is adopted for all the tests with discontinuous initial conditions; whereas for all tests with smooth initial conditions the second approach is employed.

6.1. One-dimensional advection equations. The first 1D advection problem has smooth solution and periodic boundary condition:

$$u_t + 2u_x = 0, \quad (x, t) \in [-1, 1] \times [0, 1]$$

$$u(x, 0) = 1 + \frac{1}{2} \sin(\pi x), \quad x \in [-1, 1]$$

$$u(-1, t) = u(1, t), \quad t \in [0, 1]$$

(6.1)

The exact solution at $t = 1$ is used to compute the $L_1$ errors of both cell-averaged variables and nodal variables. Four uniform meshes with number of cells ranging from 20 to 160 are used to study the convergence behavior. The $L_1$ errors are reported in logarithmic scales in Figure 6.1.
The second problem has smooth solution and Dirichlet boundary condition:

\[
\begin{aligned}
    u_t + u_x &= 0, \quad (x, t) \in [-0.5, 0.5] \times [0, 0.5] \\
    u(x, 0) &= \begin{cases} 
        1 + \frac{1}{2}x^3 \sin(2\pi x) & x \in [-0.5, 0] \\
        1 & x \in [0, 0.5] 
    \end{cases} \\
    u(-0.5, t) &= 1 + \frac{1}{2}(t + \frac{1}{2})^3 \sin(2\pi(t + \frac{1}{2})), \quad t \in [0, 0.5] \\
    u(0.5, t) &= 1.0, \quad t \in [0, 0.5] 
\end{aligned}
\] (6.2)

Four uniform meshes with number of cells ranging from 20 to 160 are used to study the convergence behavior. The \(L_1\) errors are reported in logarithmic scales in Figure 6.2.

All the simulations use Courant numbers just below the \(\lambda_{max}\) in Table 4.3. It is clear that all the FD-FV schemes are one-order higher than the designed order of \(D_x\), thus confirm the superior accuracy property given by Theorem 3.5.

Remark: In order to solve Eqs. (6.2), no special treatment is taken to apply the Dirichlet boundary condition at intermediate Runge-Kutta stages. As pointed out by Mark H. Carpenter et al. [7], this is not appropriate for high-order Runge-Kutta schemes. However, there seems to be no fix of this issue in literature for general explicit Runge-Kutta methods. Thus in this test, the fix to RK4 in their work is not applied. Nevertheless, the obtained \(L_1\) convergence behavior is still as expected and no degrading of convergence rates is observed. This may be partially due to the linearity of the problem.
Next, 1D advection equation with discontinuous initial condition is solved:

\[ u_t + u_x = 0, \quad (x,t) \in [-1.5, 1.5] \times [0, 1] \]

\[ u(x,0) = \begin{cases} 
1 & x \in [-1.5, -1] \cup [0, 1.5] \\
2 & x \in [-1, 0] 
\end{cases} \]

(6.3)

Each of the five FD-FV schemes is tested with three uniform meshes with 30, 60 and 120 cells respectively. These solutions are plotted in Figures 6.3. For comparison, solutions obtained from a second-order FVM using Roe solver, but without any limiting strategy, are also presented.

Since the differential operators \( D_x \) are designed based on Taylor series expansions, which are not valid when discontinuity appears, the Gibbs phenomenon is expected near the discontinuities. These are clearly observed in Figures 6.3. However, it is interesting to see that the magnitude of the spurious oscillations are independent of the mesh sizes; whereas in the case of second-order FVM without limiting, the magnitude increase slightly when mesh is refined.
6.2. One-dimensional Euler equations. The first 1D nonlinear problem is chosen to be the 1D Euler equations, in which case the flux function is convex. First, 1D Euler equations with smooth solutions and periodic boundary condition is tested:

\[
\begin{align*}
& w_t + f(w)_x = 0, \quad (x,t) \in [-1, 1] \times [0, 0.3] \\
& u(x,0) = 2 + \frac{1}{2} \sin(\pi x) \\
& \rho(x,0) = 1 + \frac{1}{2} \sin(\pi x) \\
& p(x,0) = 1 + \frac{1}{2} \sin(\pi x) \\
& w(-1,t) = w(1,t), \quad t \in [0, 1]
\end{align*}
\] (6.4)

Here the conservative vector and flux function are:

\[
\begin{align*}
& w = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}, \\
& f(w) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (E + p)u \end{pmatrix}
\end{align*}
\]

And the system is closed by ideal gas law with specific heat ratio \( \gamma = 1.4 \):

\[
E = \frac{p}{\gamma - 1} + \frac{1}{2} \rho u^2
\]

To compute the errors, a reference solution is computed by the fifth-order accurate FD-FV scheme \( D^{4,\text{up-biased}}_{x} - \text{RK5} \) on a uniform mesh with 2560 cells. Each of the five schemes is tested using four uniform meshes with number of cells ranging from 20 to 160. The \( L_1 \) errors in density are reported in logarithmic scales in Figures 6.4.

The \( L_1 \) errors in velocity and pressure show similar convergence behavior and are
Solved on (\ref{eq:6.4}) Next the classical Sod’s shock tube problem \cite{34} is tested: the 1D Euler equations are

\[
\begin{align*}
\frac{\partial 
\begin{pmatrix} 
\rho 
\end{pmatrix}}{\partial t} + \frac{\partial 
\begin{pmatrix} 
\rho u 
\end{pmatrix}}{\partial x} = 0,
\frac{\partial 
\begin{pmatrix} 
\rho u 
\end{pmatrix}}{\partial t} + \frac{\partial 
\begin{pmatrix} 
\rho u^2 + p 
\end{pmatrix}}{\partial x} = 0,
\frac{\partial 
\begin{pmatrix} 
\rho e 
\end{pmatrix}}{\partial t} + \frac{\partial 
\begin{pmatrix} 
\rho e u 
\end{pmatrix}}{\partial x} = 0,
\end{align*}
\]

with initial conditions:

\[
\begin{align*}
\rho(x,0) &= \begin{cases} 1 & x < 0 \\ 0.125 & x \geq 0 \end{cases}, \\
\rho(x,0) &= \begin{cases} 0.125 & x < 0 \\ 1 & x \geq 0 \end{cases},
\end{align*}
\]

\[
\begin{align*}
u(x,0) &= \begin{cases} 0 & x < 0 \\ 0.0 & x \geq 0 \end{cases},
\end{align*}
\]

\[
\begin{align*}
p(x,0) &= \begin{cases} 1 & x < 0 \\ 0.1 & x \geq 0 \end{cases}.
\end{align*}
\]

The superior accuracy property is confirmed for both cell-averaged values and nodal values.

Next the classical Sod’s shock tube problem \cite{34} is tested: the 1D Euler equations are solved on \((x, t) \in [-2, 2] \times [0, 0.8]\) with initial conditions:

\[
\begin{align*}
\rho(x,0) &= 1.0, \\
u(x,0) &= 0.0, \quad x \in [-2, 0], \\
p(x,0) &= 1.0.
\end{align*}
\]

The superior accuracy property is confirmed for both cell-averaged values and nodal values.

Table 6.1

| FD-FV scheme | Cells | \(L_1\) errors | Rates | \(L_1\) errors | Rates |
|--------------|------|----------------|-------|----------------|-------|
| \(D_{x,up-RK2}^1\) | 40 | 6.61E - 3 | 4.72E - 3 |  |  
| | 80 | 1.80E - 3 | 1.33E - 3 | 1.83 |  
| | 160 | 4.70E - 4 | 3.50E - 4 | 1.92 |  
| | 320 | 1.20E - 4 | 8.98E - 5 | 1.96 |  
| \(D_{x,up-RK3}^2\) | 40 | 4.05E - 4 | 3.54E - 4 |  |  
| | 80 | 5.18E - 5 | 4.54E - 5 | 2.96 |  
| | 160 | 5.88E - 6 | 5.11E - 6 | 3.15 |  
| | 320 | 8.21E - 7 | 7.22E - 7 | 2.82 |  
| \(D_{x,up-biased-RK4}^3\) | 40 | 7.66E - 5 | 6.61E - 5 |  |  
| | 80 | 5.20E - 6 | 4.46E - 6 | 3.89 |  
| | 160 | 3.36E - 7 | 2.87E - 7 | 3.96 |  
| | 320 | 2.13E - 8 | 1.82E - 8 | 3.98 |  
| \(D_{x,up-RK4}^4\) | 40 | 4.39E - 5 | 3.63E - 5 |  |  
| | 80 | 2.82E - 6 | 2.43E - 6 | 3.90 |  
| | 160 | 1.79E - 7 | 1.54E - 7 | 3.98 |  
| | 320 | 1.12E - 8 | 9.70E - 8 | 3.99 |  
| \(D_{x,up-biased-RK5}^5\) | 40 | 8.12E - 6 | 6.63E - 6 |  |  
| | 80 | 2.94E - 7 | 2.69E - 7 | 4.62 |  
| | 160 | 9.78E - 9 | 9.09E - 9 | 4.89 |  
| | 320 | 3.15E - 10 | 2.93E - 10 | 4.95 |  

Note that in this case, the velocity and pressure of cell-averaged values are computed by following equations:

\[
u = u(w) = \frac{\bar{w}}{\rho}, \quad p = p(w) = \left(\gamma - 1\right) \left(\frac{\bar{E}}{\rho} - \frac{1}{2} \left(\frac{\bar{w}}{\rho}\right)^2\right)
\]

It is again observed that: (1) Gibbs phenomenon appears near the discontinuities; (2) the magnitudes of the oscillations around both contact discontinuity and the shock are independent of the mesh sizes. For comparison, solutions obtained using the same
The problem is solved until $T = 1.8$. In this case, not all the five FD-FV schemes succeed in producing solutions. In particular, density profiles obtained by $D_x^{1,up}$–RK2, $D_x^{2,up}$–RK3, $D_x^{3,up}$–biased–RK4, $D_x^{4,up}$–biased–RK5, $D_x^{4,up}$–biased–RK5 succeed in producing solutions. In particular, density profiles obtained by $D_x^{1,up}$–RK2, $D_x^{2,up}$–RK3, $D_x^{3,up}$–biased–RK4, $D_x^{4,up}$–biased–RK5, $D_x^{4,up}$–biased–RK5 succeed in producing solutions.

 meshes by a second-order FVM using Roe flux and without any limiting strategy are also provided in Figure 6.3. Note that the unlimited FVM performs poorly near the shock: no data is obtained at all on the finest mesh. Hence it is observed, and conjectured that the proposed FD-FV methods have some inherent stabilization mechanism that is not possessed by conventional FVM methods.

It must be emphasized, however, the author do not mean by this observation, that the present FD-FV schemes are competitive to modern FVMs, especially when discontinuity is expected. The reason is that nonlinear stability analysis of FVMs are very mature, and many techniques are developed over decades to ensure certain stability property (like TVD, TVB, or discrete maximum principle) for finite volume schemes. Nevertheless, it worths to study the nonlinear stability property of the proposed FD-FV schemes to enhance the robustness of the methods; this is an important direction of future work on FD-FV schemes.

The last test with 1D Euler equations is a more challenging one, the Shu-Osher problem [33]. This problem solves flow obtained by a moving Mach 3 shock wave interacting with a sinusoidal density profile. The computational domain is $x \in [-5, 5]$, with initial conditions:

$$
\begin{align*}
\rho(x, 0) &= 3.857143 \\
u(x, 0) &= 2.629369 \\
p(x, 0) &= 10.33333 \\
\rho(x, 0) &= 1 + \frac{1}{5} \sin(5x) \\
\rho(x, 0) &= 0.0 \\
p(x, 0) &= 1.0 \\
x \in [-5, -4] \quad , \\
x \in [-4, 5]
\end{align*}
$$

The problem is solved until $T = 1.8$. In this case, not all the five FD-FV schemes succeed in producing solutions. In particular, density profiles obtained by $D_x^{1,up}$–RK2,

| FD-FV scheme | Cells | $L_1$ errors | Rates | $L_1$ errors | Rates |
|--------------|-------|--------------|-------|--------------|-------|
| $D_x^{1,up}$–RK2 | 40    | $7.54E-3$ | 1.91 | $5.98E-3$ | 1.90 |
| | 80    | $2.01E-3$ | 1.96 | $1.61E-3$ | 1.90 |
| | 160   | $5.17E-4$ | 1.98 | $4.14E-4$ | 1.96 |
| | 320   | $1.31E-4$ | 1.98 | $1.05E-4$ | 1.98 |
| $D_x^{2,up}$–RK3 | 40    | $4.62E-4$ | 2.96 | $4.10E-4$ | 2.97 |
| | 80    | $5.92E-5$ | 3.15 | $5.23E-5$ | 3.17 |
| | 160   | $6.68E-6$ | 2.84 | $5.83E-6$ | 2.82 |
| | 320   | $9.33E-7$ | 3.99 | $8.26E-7$ | 3.99 |
| $D_x^{3,up}$–biased–RK4 | 40    | $7.85E-5$ | 3.96 | $6.49E-5$ | 3.93 |
| | 80    | $5.29E-6$ | 3.99 | $4.26E-6$ | 3.97 |
| | 160   | $3.39E-7$ | 3.99 | $2.72E-7$ | 3.99 |
| | 320   | $2.13E-8$ | 3.99 | $1.71E-8$ | 3.99 |
| $D_x^{4,up}$–biased–RK5 | 40    | $4.32E-5$ | 3.94 | $3.27E-5$ | 3.88 |
| | 80    | $2.82E-6$ | 3.98 | $2.21E-6$ | 3.97 |
| | 160   | $1.78E-7$ | 3.99 | $1.41E-7$ | 3.99 |
| | 320   | $1.12E-8$ | 3.99 | $8.89E-8$ | 3.99 |
| $D_x^{4,up}$–biased–RK5 | 40    | $8.40E-6$ | 4.76 | $6.51E-6$ | 4.62 |
| | 80    | $3.09E-7$ | 4.92 | $2.64E-7$ | 4.86 |
| | 160   | $1.02E-8$ | 4.97 | $9.13E-9$ | 4.86 |
| | 320   | $3.27E-10$ | 4.97 | $2.94E-10$ | 4.96 |

Table 6.2: $L_1$ error in pressure of various FD-FV schemes applied to solve (6.4)
$D^2_{x,up}$-RK3 and $D^3_{x,up-biased}$-RK4 with a reduced Courant number are presented in Figures 6.6 and 6.7. For comparison, the density profile obtained by the MUSCL scheme with RK2 and van Albada limiter is also plotted. In the same figure, the number of cells for FVM is twice as large as number of cells for FD-FV schemes, so that the numbers of DOF are the same. The reference solution is computed using MUSCL-RK2 with 10000 uniform cells.
Comparing the two second-order schemes (MUSCL–RK2 and $D^{1,ap}_{x}$–RK2), given the same number of DOF: (1) the FD-FV scheme provides much more accurate magnitude of the solution, indicating the proposed scheme is much less dissipative than FVM with the same order of accuracy; (2) on the other hand, the phase error obtained from the second-order FD-FV scheme is larger than MUSCL–RK2. This phase error is greatly reduced if higher-order FD-FV methods are used.

6.3. One-dimensional scalar conservation law with non-convex flux.
The last 1D test case concerns non-convex flux [17], in which case approximates Riemann solver in FVMs is much harder to design. The governing equation is given by the scalar HCL:

$$u_t + f(u)_x = 0, \quad f(u) = \frac{1}{4}(u^2 - 1)(u^2 - 4) \quad (6.6)$$

The problem is solved on $(x, t) \in [-2, 2] \times [0, 0.04]$ with initial condition:

$$u(x, 0) = -3, \quad x \in [-2, 0]; \quad u(x, 0) = 3, \quad x \in [0, 2]$$

The exact solution at $T = 0.04$ is given by:

$$u(x, T) = \begin{cases} 
-3 & x \leq -19.5T \\
g(x/T) & -19.5T \leq x < 0 \\
-g(-x/T) & 0 < x \leq 19.5T \\
3 & 19.5T \leq x 
\end{cases}$$

Here $g(\cdot)$ satisfies $u = f'(g(u))$. All the five FD-FV schemes are tested. The first series of results are obtained using 80 uniform cells, and the results are plotted with the exact solution in Figures 6.8.

Note that the magnitudes of the oscillations are not symmetric about $x = 0$, this is due to the fact that $x = 0$ coincides with a grid point, and the initial condition at this point is set to be equal to 3, i.e. the right status. If odd number of uniform cells is used, for example 81 cells, the asymmetry disappears, as shown in Figures 6.9.

It is also interesting to observe that, using odd number of cells, the spurious oscillations near the weak discontinuities almost disappear.

6.4. Two-dimensional Euler equations. Finally, a 2D problem, the isentropic vortex advection problem [32], is tested to illustrate the superior accuracy.
The computation is performed on \((x, y, t) = [-5, 5] \times [-5, 5] \times [0, 10]\), with periodic boundary condition on all four edges. The initial condition is an isentropic vortex
given by:

\[
\begin{align*}
    u(x, y, 0) &= 1 - \frac{\epsilon y}{2\pi} \exp\left(\frac{1}{2}(1 - x^2 - y^2)\right) \\
    v(x, y, 0) &= 1 + \frac{\epsilon x}{2\pi} \exp\left(\frac{1}{2}(1 - x^2 - y^2)\right) \\
    \rho(x, y, 0) &= \left(1 - \frac{(\gamma - 1)c^2}{8\gamma\pi^2} \exp\left(1 - x^2 - y^2\right)\right)^{\frac{1}{\gamma}} \\
    p(x, y, 0) &= \left(1 - \frac{(\gamma - 1)c^2}{8\gamma\pi^2} \exp\left(1 - x^2 - y^2\right)\right)^{\frac{1}{\gamma}}
\end{align*}
\]

This initial condition leads to a vortex with uniform entropy \(p/\rho^\gamma = 1.0\). The constant \(\epsilon\) is chosen to be \(\epsilon = 5\), and the exact solution at one period \(T = 10\) equals the initial condition. Four uniform meshes are employed to test the convergence behavior with grid sizes ranging from \(20 \times 20\) to \(160 \times 160\). The computed \(L_1\) errors are reported in Table 6.3.

It is clear from the table that the asymptotic convergence rates in all components are second-order. Note that the coarsest mesh corresponds to using only 8 cells per dimension to resolve the vortex; hence the computed \(L_1\) errors are large. The computational costs of this FD-FV scheme is compared to MUSCL–RK2, which has the same order of accuracy in both space and time. To this end, Table 6.4 presents the comparison results obtained by choosing the same mesh size for both schemes, as
well as the comparison results using roughly the same number of DOFs. Here all the tests are performed using C++ programs and g++ compiler under 64bit CentOS 5.8 system with single processor (Intel(R) Xeon (TM) CPU 3.00GHz).

It can be concluded from the table that: (1) given the same mesh size, although the number of DOFs of $D^1_{x,up}$–RK2 is two times larger than MUSCL–RK2, the CPU times are almost the same; and the FD-FV scheme gives much better results than FVM; (2) given the same number of DOFs, $D^1_{x,up}$–RK2 provides less accurate solutions than MUSCL–RK2, but the former approaches asymptotic convergence faster than the latter as the meshes are refined – the difference between the $L_1$ errors is less significant when meshes are refined; (3) given the same number of DOFs, $D^1_{x,up}$–RK2 is much faster than MUSCL–RK2.

**Remark:** This comparative study only considers explicit time-integrators, in which case the cost of FVMs is mainly due to numerical flux evaluations. However, implicit methods are often employed in practice to allow larger time steps. When implicit time-integrators are used, a major additional computational cost is from the nonlinear solver, for example, by Newton methods. In this case, the connectivity plays an important role in determining the efficiency of the algorithm. Thus the statements made here may or may not applies to implicit methods, which will be investigated in future work.

### Table 6.4

| Scheme | Mesh size | $N_{dof}$ | $E_\rho$ | $N_{iter}$ | $T_{cpu}$ (sec.) |
|--------|-----------|-----------|----------|-----------|-----------------|
| FD-FV  | 20 × 20   | 4,960     | 1.87E + 0| 200       | 20              |
| FVM    | 20 × 20   | 1,600     | 2.23E + 0| 200       | 18              |
| FVM    | 35 × 35   | 4,900     | 1.26E + 0| 350       | 92              |
| FD-FV  | 40 × 40   | 19,520    | 8.12E - 1| 400       | 158             |
| FVM    | 40 × 40   | 6,400     | 1.02E + 0| 400       | 139             |
| FVM    | 70 × 70   | 19,600    | 3.76E - 1| 700       | 731             |
| FD-FV  | 80 × 80   | 77,440    | 2.41E - 1| 800       | 1,252           |
| FVM    | 80 × 80   | 25,600    | 2.97E - 1| 800       | 1,101           |
| FVM    | 140 × 140 | 78,400    | 1.25E - 1| 1,400     | 5,852           |
| FD-FV  | 160 × 160 | 308,380   | 6.26E - 2| 1,600     | 9,945           |
| FVM    | 160 × 160 | 102,400   | 1.07E - 1| 1,600     | 8,730           |
| FVM    | 278 × 278 | 309,136   | 4.73E - 2| 2,780     | 45,938          |

Performance comparison between second-order FD-FV ($D^1_{x,up}$–RK2) and second-order FVM (MUSCL–RK2 with van Albada limiter): $N_{dof}$ – number of DOFs; $E_\rho$ – $L_1$ error in density for cell-averaged data; $N_{iter}$ – number of iterations; $T_{cpu}$ – computational time in seconds

### 7. Concluding remarks.

A hybrid finite difference–finite volume approach is proposed for solving general first-order hyperbolic conservation laws. It is a departure from conventional finite difference schemes and finite volume schemes in that both cell-averaged values and nodal values are counted as dependent variables and evolved in time. By distinguishing between cell-averaged DOFs and nodal DOFs, the proposed FD-FV schemes are: (1) natural conservative like FVMs, but work for general flux functions and do not require any exact or approximated Riemann solver; (2) easy to extend to high-order accuracy in space like FDMs. The present work focuses on discretization in space, and integration in time is achieved by method of lines; in
particular, explicit Runge-Kutta methods are adopted for simplicity. It is theoretically proven and numerically illustrated that the proposed FD-FV methods have superior spatial accuracy property than conventional FDMs and FVMs. This means that the achieved order of accuracy in space is one-order higher than the designed order of accuracy of the discrete differential operator applied in the FD-FV scheme. When shock occurs in the solutions, second-order FVM fail if no other stabilization mechanism is applied, especially when the mesh is sufficiently refined. However, it is observed for the proposed FD-FV schemes, that the magnitudes of the oscillations near both contact discontinuities and shocks are independent of the mesh size. Hence it is conjectured that the FD-FV schemes have some inherently stabilization mechanism that worths further investigation.

There are remaining issues, however, to be addressed in future work. For example, nonlinear stability of the methods is to be investigated, and the present schemes need to be enhanced to avoid any spurious oscillations near discontinuities. In addition, since the conservation law alone does not prevent non-physical solutions, entropy inequality should be considered in designing practical FD-FV schemes.

**Acknowledgments.** The author thanks support from Farhat Research Group at Stanford University, and greatly appreciates Charbel Farhat and George C. Papaioannou (both from Stanford University) for reviewing the manuscript and their insightful suggestions. The author also gratefully thanks Guglielmo Scovazzi from Duke University for fruitful discussions.

**REFERENCES**

[1] T. Aoki, Interpolated differential operator (ido) scheme for solving partial differential equations, Comput. Phys. Commun., 102 (1997), pp. 132–146.

[2] T. J. Barth, Aspects of unstructured grids and finite-volume solvers for the euler and navier-stokes equations, in Special Course on Unstructured Grid Methods for Advection Dominated Flows, T. J. Barth and H. Deconinck, eds., AGARD, May 1992. AGARD Report 787.

[3] T. J. Barth, Recent developments in high order k-exact reconstruction on unstructured meshes, in AIAA 31st Aerospace Sciences Meeting & Exhibit, January 1993, Reno, Nevada.

[4] J. P. Boris and D. L. Book, Flux-corrected transport. i. shasta, a fluid transport algorithm that works, J. Comput. Phys., 11 (1973), pp. 38–69.

[5] A. Bressan, Hyperbolic conservation laws, in Mathematics of Complexity and Dynamical Systems, R. A. Meyers, ed., Springer New York, 2011, pp. 729–739. Subject: Mathematics and Statistics.

[6] M. H. Carpenter, D. Gottlieb, and S. Abarbanel, Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes, J. Comput. Phys., 111 (1994), pp. 220–236.

[7] M. H. Carpenter, D. Gottlieb, S. Abarbanel, and W.-S. Don, The theoretical accuracy of runge-kutta time discretizations for initial boundary value problems: a study of the boundary error, SIAM J. Sci. Comput., 16 (1995), pp. 1241–1252.

[8] B. Cockburn and C.-W. Shu, The runge-kutta discontinuous galerkin method of conservation laws v: multidimensional systems, J. Comput. Phys., 141 (1998), pp. 199–224.

[9] P. Colella and P. R. Woodward, The piecewise parabolic method (ppm) for gas-dynamical simulations, J. Comput. Phys., 54 (1984), pp. 174–201.

[10] C. M. Dafermos, Hyperbolic Conservation Laws in Continuum Physics, vol. 325 of Grundlehren der mathematischen Wissenschaften, Springer, 3 ed., 2010.

[11] B. Einfeldt, On godunov-type methods for gas dynamics, SIAM J. Numer. Anal., 25 (1988), pp. 294–318.

[12] S. K. Godunov, A difference scheme for numerical computation of discontinuous solution of hyperbolic equation, Matematicheskii Sbornik, 47 (1959), pp. 271–306. In Russian.

[13] S. Gottlieb and C.-W. Shu, Total variational diminishing runge-kutta schemes, Math. Comput., 67 (1998), pp. 73–85.
[14] S. Gottlieb, C.-W. Shu, and E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev., 43 (2001), pp. 89–112.
[15] B. Gustafsson, High Order Difference Methods for Time Dependent PDE, vol. 38 of Springer Series in Computational Mathematics, Springer, 1 ed., 2008.
[16] E. Hairer, S. P. Nørsett, and G. Wanner, Solving Ordinary Differential Equations I: Nonstiff Problems, vol. 8 of Springer Series in Computational Mathematics, Springer, 2 ed., 1993.
[17] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, Uniformly high order accurate essentially non-oscillatory schemes, iii, J. Comput. Phys., 71 (1987), pp. 231–303.
[18] T. J. R. Hughes and T. E. Tzelepi, Finite element methods for first-order hyperbolic systems with particular emphasis on the compressible euler equations, 45 (1984), pp. 217–284.
[19] Y. Imai, T. Aoki, and K. Takizawa, Conservative form of interpolated differential operator scheme for compressible and incompressible fluid dynamics, J. Comput. Phys., 227 (2008), pp. 2253–2283.
[20] A. Jameson, Positive schemes and shock modelling for compressible flows, Int. J. Numer. Meth. FL., 20 (1995), pp. 743–776.
[21] A. Jameson, W. Schmidt, and E. Turkel, Numerical solutions of the euler equations by finite volume methods with runge-kutta time stepping schemes, in AIAA 14th Fluid and Plasma Dynamics Conference, June 1981. Palo Alto, California.
[22] C. A. Kennedy, M. H. Carpenter, and R. M. Lewis, Low-storage, explicit runge-kutta schemes for the compressible navier-stokes equations, Appl. Numer. Math., 35 (2000), pp. 177–219.
[23] D. I. Ketcheson, Highly efficient strong stability-preserving runge-kutta methods with low-storage implementations, SIAM J. Sci. Comput., 30 (2008), pp. 2113–2136.
[24] V. F. Kolgan, Application of the principle of minimizing the derivative to the construction of finite-difference schemes for computing discontinuous solutions of gas dynamics, J. Comput. Phys., 230 (2011), pp. 2384–2390. Short Note, Originally published in Uchenie Zapiski TsAGI (TsAGI Research Notes), 3 (6), 6876, 1972.
[25] X.-D. Liu, S. Osher, and T. Chan, Weighted essentially non-oscillatory schemes, J. Comput. Phys., 115 (1994), pp. 200–212.
[26] H. A. Luther, Further explicit fifth-order runge-kutta formulas, SIAM Rev., 8 (1966), pp. 374–389.
[27] C. Michalak and C. Ollivier-Gooch, Accuracy preserving limiter for the high-order accurate solution of the euler equations, J. Comput. Phys., 228 (2012), pp. 8693–9711.
[28] P. L. Roe, Approximate riemann solvers, parameter vectors, and difference schemes, J. Comput. Phys., 43 (1981), pp. 357–372.
[29] S. J. Ruuth, Global optimization of explicit strong-stability-preserving runge-kutta methods, Math. Comput., 75 (2006), pp. 183–207.
[30] K. Sakurai, T. Aoki, W.-H. Lee, and K. Kato, Poisson equation solver with fourth-order accuracy by using interpolated differential operator scheme, Comput. Math. Appl., 43 (2002), pp. 621–630.
[31] F. Shakib, T. J. R. Hughes, and Z. Johan, A new finite element formulation for computational fluid dynamics: X. the compressible euler and navier-stokes equations, 89 (1991), pp. 141–219.
[32] C.-W. Shu, High-order finite difference and finite volume weno schemes and discontinuous galerkin methods for cfd, Int. J. Comput. Fluid D., 17 (2003), pp. 107–118.
[33] C.-W. Shu and S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii, J. Comput. Phys., 83 (1989), pp. 32–78.
[34] G. A. Sod, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, J. Comput. Phys., 27 (1978), pp. 1–31. Review.
[35] R. J. Spiteri and S. J. Ruuth, A new class of optimal high-order strong-stability-preserving time discretization methods, SIAM J. Numer. Anal., 40 (2002), pp. 469–491.
[36] M. Svärd, M. H. Carpenter, and J. Nordström, A stable high-order finite difference scheme for the compressible navierstokes equations, far-field boundary conditions, J. Comput. Phys., 225 (2007), pp. 1020–1038.
[37] E. F. Toro, M. Spruce, and W. Speares, Restoration of the contact surface in the hll-riemann solver, Shock Waves, 4 (1994), pp. 25–34.
[38] G. D. van Albada, B. van Leer, and J. W. W. Roberts, A comparative study of computational methods in cosmic gas dynamics, Astron. Astrophys., 108 (1982), pp. 76–84.
[39] B. van Leer, Towards the ultimate conservative difference scheme v. a second-order sequel to godunov's method, J. Comput. Phys., 32 (1979), pp. 101–136.
[40] J. H. Williamson, Low-storage runge-kutta schemes, J. Comput. Phys., 35 (1980), pp. 48–56.
[41] F. Xiao, A. Ikebata, and T. Hasegawa, Numerical simulations of free-interface fluids by a multi-integrated moment method, Comput. Struct., 83 (2005), pp. 409–423.
[42] Y. Yabe, F. Xiao, and T. Utsumi, The constrained interpolation profile method for multiphase analysis, J. Comput. Phys., 169 (2001), pp. 556–593.