A New Semi-Discrete Finite-Volume Active Flux Method for Hyperbolic Conservation Laws

Rémi Abgrall, Alina Chertock, Alexander Kurganov, and Lorenzo Micalizzi§

Abstract

In this work, we introduce a new active flux (AF) method for hyperbolic systems of conservation laws. Following an AF approach recently proposed in [R. ABGRALL, Commun. Appl. Math. Comput., 5 (2023), pp. 370–402], we consider two different formulations of the studied system (the original conservative formulation and a primitive one containing nonconservative products), and discretize them on overlapping staggered meshes using two different numerical schemes. The novelty of our method is twofold. First, we introduce an original paradigm making use of overlapping finitevolume (FV) meshes over which cell averages of conservative and primitive variables are evolved using semi-discrete FV methods: The nonconservative system is discretized by a path-conservative central-upwind scheme and its solution is used to evaluate very simple numerical fluxes for the discretization of the original conservative system. Second, to ensure the nonlinear stability of the resulting AF method, we design a post-processing, which also guarantees a conservative coupling between the two sets of variables. We test the proposed semi-discrete FV AF method on a number of benchmarks for the one- and two-dimensional Euler equations of gas dynamics.

Key words: Active flux methods; overlapping staggered meshes; path-conservative central-upwind schemes; conservative post-processing. states.

AMS subject classification: 76M12, 65M08, 76N99, 35L65, 35L67.

1 Introduction

This paper focuses on the development of numerical methods for hyperbolic systems of conservation laws, which in the two-dimensional (2-D) case read as

$$U_t + F(U)_x + G(U)_y = 0.$$
(1.1)

Here, *x* and *y* are the spatial variables, *t* is time, $U \in \mathbb{R}^M$ is the vector of conserved variables, $F, G \in \mathbb{R}^M$ are the fluxes, whose Jacobians, $\frac{\partial F}{\partial U}, \frac{\partial G}{\partial U} \in \mathbb{R}^{M \times M}$, are assumed to be real-diagonalizable, and $M \in \mathbb{N}$ with $M \ge 1$ is the number of equations in the system.

^{*}Institute for Mathematics & Computational Science, University of Zurich, 8057 Zurich, Switzerland; remi.abgrall@math.uzh.ch

[†]Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA; chertock@math.ncsu.edu

[‡]Department of Mathematics and Shenzhen International Center for Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China; alexander@sustech.edu.cn

[§]Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA; lmicali@ncsu.edu

It is well-known that solutions of nonlinear hyperbolic conservation laws may become nonsmooth even when the initial and boundary data are infinitely smooth. Therefore, the solution of (1.1) has to be defined in a weak (integral) sense, and hence conservative finite-volume (FV) methods seem to be one of the natural choices to be considered. In these methods, the computational domain is covered by FV cells and the numerical solution is realized in terms of its cell averages, which are evolved in time using an integral form of (1.1). For a variety of existing FV methods, we refer the reader to [21,24,29,47] and references therein.

In this work, we are interested in active flux (AF) methods, which were first introduced in [18, 19] as an extension of the scheme proposed in [48] for linear hyperbolic partial differential equations (PDEs). The key idea of AF methods is to evolve several pieces of information on the computed solution, namely, both cell averages and point values at cell interfaces. This gives additional degrees of freedom, which can be used not only to enhance the accuracy of the resulting scheme, but also to successfully hybridize conservative and nonconservative numerical methods. As demonstrated in [4], nonconservative methods cannot be used for accurately solving hyperbolic systems of conservation laws since nonconservative numerical schemes typically converge to non-entropy (non-physical) weak solutions. However, in the AF framework, highly accurate nonconservative methods can be used to evolve the cell interface point values, while the convergence toward the physically relevant solutions is guaranteed by a conservative evolution of the cell averages. In the past decade, a variety of AF methods have been developed; see, e.g., [1-3, 5, 6, 13, 40, 52].

A similar idea of obtaining additional degrees of freedom by evolving several pieces of information was used in the methods on overlapping cells. These methods include both FV [31,33,34,50] and discontinuous Galerkin [32, 50, 51] ones. In the FV methods on overlapping cells, hierarchical reconstruction limiters are used to achieve high-order non-oscillatory approximations of the computed solution, whose cell averages on overlapping cells are evolved in time.

In this paper, we incorporate the overlapping FV approach into the AF framework to develop a novel second-order semi-discrete FV AF method. To this end, we first consider the following nonconservative system, which is equivalent to (1.1) for smooth solutions:

$$\boldsymbol{V}_t + \widetilde{\boldsymbol{F}}(\boldsymbol{V})_x + \widetilde{\boldsymbol{G}}(\boldsymbol{V})_y = \boldsymbol{B}(\boldsymbol{V})\boldsymbol{V}_x + \boldsymbol{C}(\boldsymbol{V})\boldsymbol{V}_y, \qquad (1.2)$$

where $V \in \mathbb{R}^M$ is the vector of primitive variables, $\widetilde{F}, \widetilde{G} : \mathbb{R}^M \to \mathbb{R}^M$, and $B, C \in \mathbb{R}^{M \times M}$. To cite an example, one may consider the Euler equations of gas dynamics, which read as (1.1) with

$$\boldsymbol{U} = (\rho, \rho u, \rho v, E)^{\mathsf{T}}, \quad \boldsymbol{F}(\boldsymbol{U}) = (\rho u, \rho u^2 + p, \rho u v, u(E+p))^{\mathsf{T}}, \quad \boldsymbol{G}(\boldsymbol{U}) = (\rho u, \rho u v, \rho v^2 + p, v(E+p))^{\mathsf{T}}, \quad (1.3)$$

where ρ is the density, *u* and *v* are the *x*- and *y*-velocities, *E* is the total energy, and *p* is the pressure. We use the classical closure, obtained with the help of the equation of state of ideal fluids,

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho(u^2 + v^2), \qquad (1.4)$$

in which γ is the specific heat ratio. The system (1.1), (1.3)–(1.4) can be rewritten in the nonconservative form in many different ways, for instance, using the primitive variables ρ , u, v, and p, for which the corresponding nonconservative system reads as (1.2) with

$$\boldsymbol{V} = (\rho, u, v, p)^{\mathsf{T}}, \quad \widetilde{\boldsymbol{F}}(\boldsymbol{V}) = \left(\rho u, \frac{u^2}{2}, 0, pu\right)^{\mathsf{T}}, \quad \widetilde{\boldsymbol{G}}(\boldsymbol{V}) = \left(\rho u, 0, \frac{v^2}{2}, pv\right)^{\mathsf{T}},$$

$$B^{x}(V) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\rho} \\ 0 & 0 & -u & 0 \\ 0 & -(\gamma - 1)p & 0 & 0 \end{pmatrix}, \quad B^{y}(V) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -v & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\rho} \\ 0 & 0 & -(\gamma - 1)p & 0 \end{pmatrix}.$$

We discretize systems (1.1) and (1.2) on overlapping meshes using a semi-discrete approach, and then solve the resulting systems of ordinary differential equations (ODEs) simultaneously. The numerical fluxes for the conservative system (1.1) are taken simply as F(U(V)) and G(U(V)), while the discretization of the nonconservative system (1.2) is more involved as its solutions cannot be understood in the sense of distributions. In [16], a concept of Borel measure solutions of nonconservative systems was introduced (see also [27, 28]) and later utilized to develop path-conservative numerical methods; see, e.g., [8–10, 37, 38] and references therein. Here, we discretize (1.2) using a modified version of the Riemann-problem-solver-free path-conservative central-upwind (PCCU) scheme from [10]. The modification is intended to reduce the amount of numerical dissipation present in the original PCCU scheme and is carried out by replacing the central-upwind (CU) numerical flux from [26], which was used in [10], with a less dissipative CU numerical flux from [25].

Since the conservative numerical fluxes do not employ any limiting procedures, one can expect the computed U to be oscillatory. At At the same time, variables V, computed in a non-oscillatory manner by the PCCU scheme, may converge to a non-physical weak solution. We therefore introduce a conservative post-processing, which couples the evolution of the two sets of variables. The resulting numerical solution is (essentially) oscillation-free and the scheme converges to the physically relevant solution of (1.1). We test the proposed FV AF method on a number of benchmarks for one- and two-dimensional Euler equations of gas dynamics.

The rest of the paper is organized as follows. In §2, we introduce the new one-dimensional (1-D) FV AF method, and then extend it to the 2-D case in §3. In §4, we report the results of several numerical experiments for both 1-D and 2-D Euler equations of gas dynamics. Finally, in §5, we provide concluding remarks and discuss future perspectives.

2 One-Dimensional Semi-Discrete FV AF Method

In this section, we introduce the new semi-discrete FV AF method for the 1-D version of (1.1):

$$\boldsymbol{U}_t + \boldsymbol{F}(\boldsymbol{U})_x = \boldsymbol{0}, \tag{2.1}$$

whose nonconservative formulation reads as

$$\boldsymbol{V}_t + \boldsymbol{F}(\boldsymbol{V})_x = \boldsymbol{B}(\boldsymbol{V})\boldsymbol{V}_x. \tag{2.2}$$

We consider overlapping FV meshes consisting of uniform cells $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, j = 1, ..., N and $I_{j+\frac{1}{2}} = [x_j, x_{j+1}]$, j = 0, ..., N with $x_{j+1} = x_{j+\frac{1}{2}} + \Delta x/2 = x_j + \Delta x$ for all *j*. As in all FV methods, the computed quantities are cell averages of U and V, which are obtained on the above two grids, namely,

$$\overline{U}_j :\approx \int_{I_j} U(x,t) \, \mathrm{d}x \quad \text{and} \quad \overline{V}_{j+\frac{1}{2}} :\approx \int_{I_{j+\frac{1}{2}}} V(x,t) \, \mathrm{d}x$$

Note that both \overline{U}_j and $\overline{V}_{j+\frac{1}{2}}$, like many other indexed quantities below, depend on time, but we omit this dependence for the sake of notation brevity.

The semi-discretization of the conservative system (2.1) is obtained by integrating (2.1) in space over the cells I_i , which results in

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{U}_{j} = -\frac{1}{\Delta x} \Big[\mathcal{F}_{j+\frac{1}{2}} - \mathcal{F}_{j-\frac{1}{2}} \Big], \qquad (2.3)$$

in which we take the following simple numerical fluxes:

$$\mathcal{F}_{j+\frac{1}{2}} = F(U_{j+\frac{1}{2}}), \quad U_{j+\frac{1}{2}} := U(\overline{V}_{j+\frac{1}{2}}),$$
(2.4)

where U(V) denotes the transformation from primitive to conserved variables. Note that for secondorder methods, cell averages and point values formally differ by $O((\Delta x)^2)$, which makes the transformation used in (2.4) straightforward, while higher order extensions would require a suitable higher order reconstruction of point values.

The semi-discretization of the nonconservative system (2.2) is obtained through the modified PCCU scheme and reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathbf{V}}_{j+\frac{1}{2}} = -\frac{1}{\Delta x} \Big[\widetilde{\mathcal{F}}_{j+1} - \widetilde{\mathcal{F}}_j - \mathbf{B}_{j+\frac{1}{2}} - \frac{a_j^+}{a_j^+ - a_j^-} \mathbf{B}_{\Psi,j} + \frac{a_{j+1}^-}{a_{j+1}^+ - a_{j+1}^-} \mathbf{B}_{\Psi,j+1} \Big],$$
(2.5)

where $\widetilde{\mathcal{F}}_{j}$ are the CU numerical fluxes from [25] given by

$$\widetilde{\mathcal{F}}_{j} = \frac{a_{j}^{+}\widetilde{F}(V_{j}^{-}) - a_{j}^{-}\widetilde{F}(V_{j}^{+})}{a_{j}^{+} - a_{j}^{-}} + \frac{a_{j}^{+}a_{j}^{-}}{a_{j}^{+} - a_{j}^{-}} \Big(V_{j}^{+} - V_{j}^{-} - \delta V_{j}\Big).$$
(2.6)

Here, V_j^{\pm} are one-sided point values of V obtained using a piecewise linear reconstruction applied to the local characteristic variables Γ of(1.2). To this end, we follow [36] and introduce

$$\Gamma_{j-\frac{1}{2}} = Q_{j+\frac{1}{2}}^{-1} \overline{V}_{j-\frac{1}{2}}, \quad \Gamma_{j+\frac{1}{2}} = Q_{j+\frac{1}{2}}^{-1} \overline{V}_{j+\frac{1}{2}}, \quad \Gamma_{j+\frac{3}{2}} = Q_{j+\frac{1}{2}}^{-1} \overline{V}_{j+\frac{3}{2}},$$

where $Q_{j+\frac{1}{2}}$ is a matrix such that $Q_{j+\frac{1}{2}}^{-1} \mathcal{A}_{j+\frac{1}{2}} Q_{j+\frac{1}{2}}$ is a diagonal matrix and $\mathcal{A}_{j+\frac{1}{2}} = \mathcal{A}(\overline{V}_{j+\frac{1}{2}}) := \frac{\partial \widetilde{F}}{\partial V}(\overline{V}_{j+\frac{1}{2}}) - B(\overline{V}_{j+\frac{1}{2}})$. Equipped with $\Gamma_{j-\frac{1}{2}}, \Gamma_{j+\frac{1}{2}}$, and $\Gamma_{j+\frac{3}{2}}$, we apply a generalized minmod limiter (see, e.g., [30,45]) to evaluate

$$(\Gamma_x)_{j+\frac{1}{2}} = \operatorname{minmod}\left(\theta \, \frac{\Gamma_{j+\frac{1}{2}} - \Gamma_{j-\frac{1}{2}}}{\Delta x}, \frac{\Gamma_{j+\frac{3}{2}} - \Gamma_{j-\frac{1}{2}}}{2\Delta x}, \theta \, \frac{\Gamma_{j+\frac{3}{2}} - \Gamma_{j+\frac{1}{2}}}{\Delta x}\right), \quad \theta \in [1, 2],$$

where the minmod function, defined as

$$\min(c_1, c_2, \ldots) = \begin{cases} \min_i c_i & \text{if } c_i > 0 \ \forall i, \\ \max_i c_i & \text{if } c_i < 0 \ \forall i, \\ 0 & \text{otherwise,} \end{cases}$$

is applied in a component-wise manner. We then obtain

$$\Gamma_{j}^{+} = \Gamma_{j+\frac{1}{2}} - \frac{\Delta x}{2} (\Gamma_{x})_{j+\frac{1}{2}}, \quad \Gamma_{j+1}^{-} = \Gamma_{j+\frac{1}{2}} + \frac{\Delta x}{2} (\Gamma_{x})_{j+\frac{1}{2}},$$

and hence the corresponding point values of V are

$$V_j^+ = Q_{j+\frac{1}{2}}\Gamma_j^+, \quad V_{j+1}^- = Q_{j+\frac{1}{2}}\Gamma_{j+1}^-.$$

In (2.5) and (2.6), a_j^{\pm} are the one-sided local speeds of propagation, which can be estimated by

$$a_{j}^{-} = \min \left\{ \lambda_{1}(V_{j}^{-}), \lambda_{1}(V_{j}^{+}), 0 \right\}, \quad a_{j}^{+} = \max \left\{ \lambda_{M}(V_{j}^{-}), \lambda_{M}(V_{j}^{+}), 0 \right\}$$

where $\lambda_1(V) \leq \ldots \leq \lambda_M(V)$ are the eigenvalues of $\mathcal{A}(V)$. The term δV_j in (2.6) represents "built-in" anti-diffusion and is given by (see [25]),

$$\delta V_j = \operatorname{minmod} \left(V_j^* - V_j^-, V_j^+ - V_j^* \right), \quad V_j^* := \frac{a_j^+ V_j^+ - a_j^- V_j^- - \widetilde{F}(V_j^+) + \widetilde{F}(V_j^-)}{a_j^+ - a_j^-}$$

Finally, the terms $B_{j+\frac{1}{2}}$ and $B_{\Psi,j}$ in (2.5) are given by (see [10] for details)

$$B_{j+\frac{1}{2}} = B(\overline{V}_{j+\frac{1}{2}})(V_{j+1}^{-} - V_{j}^{+}), \quad B_{\Psi,j} = \frac{1}{2} \left[B(V_{j}^{-}) + B(V_{j}^{+}) \right] (V_{j}^{+} - V_{j}^{-}).$$

Remark 2.1 We would like to stress that the proposed FV AF method is not tied to the PCCU scheme and, in principle, one can numerically solve the nonconservative system (2.2) using an alternative second-order stable numerical method instead.

The ODE systems (2.3) and (2.5) should be numerically solved by a stable and sufficiently accurate ODE solver. However, the solution obtained upon completion of a time step evolution will have two significant drawbacks. First, the solution realized by $\{\overline{U}_j(t + \Delta t)\}$ will likely be oscillatory since no limiting procedures are employed in the computation of numerical fluxes in (2.4). Second, the solution realized by $\{\overline{V}_{j+\frac{1}{2}}(t + \Delta t)\}$ will not necessarily be conservative, that is, $\sum_j U(\overline{V}_{j+\frac{1}{2}}(t + \Delta t))$ may not be equal to $\sum_j U(\overline{V}_{j+\frac{1}{2}}(t))$. Therefore, to ensure that the resulting numerical solution is (essentially) oscillation-free and converges to the physically relevant solution of (2.1), we propose the conservative post-processing procedure presented in the next subsection.

2.1 Post-Processing

Let us assume that the solution was evolved from time *t* to $t + \Delta t$ with the help of an ODE solver and denote the obtained solutions by $\{\overline{U}_{j}^{*}\}$ and $\{\overline{V}_{j+\frac{1}{2}}^{*}\}$. Our goal is to modify these values through a suitable post-processing procedure to obtain non-oscillatory sets of $\{\overline{U}_{j}(t + \Delta t)\}$ and $\{\overline{V}_{j+\frac{1}{2}}(t + \Delta t)\}$.

The proposed post-processing can be presented algorithmically through the following four steps. **Step 1.** Compute the conserved variables at $x = x_{j+\frac{1}{2}}$ using the transformation from the primitive ones:

$$\boldsymbol{U}_{j+\frac{1}{2}}^* \coloneqq \boldsymbol{U}(\overline{\boldsymbol{V}}_{j+\frac{1}{2}}^*).$$

Step 2. Perform the piecewise linear reconstruction for U using the slopes computed by the minmod limiter

$$(U_x)_j^* = 2 \operatorname{minmod}\left(\frac{U_j^* - U_{j-\frac{1}{2}}^*}{\Delta x}, \frac{U_{j+\frac{1}{2}}^* - U_j^*}{\Delta x}\right),$$

which results in

$$\boldsymbol{U}_{j+\frac{1}{2}}^{*,-} := \overline{\boldsymbol{U}}_{j}^{*} + \frac{\Delta x}{2} (\boldsymbol{U}_{x})_{j}^{*}, \quad \boldsymbol{U}_{j+\frac{1}{2}}^{*,+} := \overline{\boldsymbol{U}}_{j+1}^{*} - \frac{\Delta x}{2} (\boldsymbol{U}_{x})_{j+1}^{*}.$$
(2.7)

Step 3. Set

$$\boldsymbol{U}_{j+\frac{1}{2}}^{**} := \frac{1}{2} \Big(\boldsymbol{U}_{j+\frac{1}{2}}^{*,-} + \boldsymbol{U}_{j+\frac{1}{2}}^{*,+} \Big), \tag{2.8}$$

and recompute the primitive variables at $x = x_{j+\frac{1}{2}}$ using the transformation from the conservative ones:

$$\overline{\boldsymbol{V}}_{j+\frac{1}{2}}(t+\Delta t) = \boldsymbol{V}(\boldsymbol{U}_{j+\frac{1}{2}}^{**}).$$

Step 4. Correct the conserved variables by setting

$$\overline{U}_{j}(t+\Delta t) = \frac{1}{2} \Big(U_{j-\frac{1}{2}}^{**} + U_{j+\frac{1}{2}}^{**} \Big).$$
(2.9)

It is important to stress that the post-processing is conservative as shown in the next proposition.

Proposition 2.1 The post-processed cell averages of the conserved variables satisfy

$$\sum_{j} \overline{U}_{j}(t + \Delta t) = \sum_{j} \overline{U}_{j}^{*}.$$

Proof. The proof consists of simple direct computations (assuming no contributions from the boundary terms):

$$\sum_{j} \overline{U}_{j}(t + \Delta t) \stackrel{(2.9)}{=} \frac{1}{2} \sum_{j} U_{j-\frac{1}{2}}^{**} + \frac{1}{2} \sum_{j} U_{j+\frac{1}{2}}^{**} = \sum_{j} U_{j+\frac{1}{2}}^{**} \stackrel{(2.8)}{=} \frac{1}{2} \sum_{j} U_{j+\frac{1}{2}}^{*,-} + \frac{1}{2} \sum_{j} U_{j+\frac{1}{2}}^{*,+}$$

$$\stackrel{(2.7)}{=} \frac{1}{2} \sum_{j} \left[\overline{U}_{j}^{*} + \frac{\Delta x}{2} (U_{x})_{j}^{*} \right] + \frac{1}{2} \sum_{j} \left[\overline{U}_{j+1}^{*} - \frac{\Delta x}{2} (U_{x})_{j+1}^{*} \right] = \sum_{j} \overline{U}_{j}^{*}.$$

3 Two-Dimensional Semi-Discrete FV AF Method

We now extend the proposed semi-discrete FV AF method to the 2-D case, in which the conservative and nonconservative (primitive) formulations of the governing equations are given by (1.1) and (1.2), respectively.

We consider overlapping Cartesian meshes (see the sketch in Figure 3.1) consisting of uniform cells $I_{j,k} := [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}] \times [y_{k-\frac{1}{2}}, y_{k+\frac{1}{2}}], j = 1, ..., N_x, k = 1, ..., N_y, I_{j+\frac{1}{2},k}^x = [x_j, x_{j+1}] \times [y_{k-\frac{1}{2}}, y_{k+\frac{1}{2}}], j = 0, ..., N_x, k = 1, ..., N_y$, and $I_{j,k+\frac{1}{2}}^y = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}] \times [y_k, y_{k+1}], j = 1, ..., N_x, k = 0, ..., N_y$ with $x_{j+1} = x_{j+\frac{1}{2}} + \Delta x/2 = x_j + \Delta x$ for all j and $y_{k+1} = y_{k+\frac{1}{2}} + \Delta y/2 = y_k + \Delta y$ for all k. The computed quantities are the following cell averages obtained on the above three grids:

$$\overline{U}_{j,k} :\approx \int_{I_{j,k}} U(x,t) \, \mathrm{d}x \mathrm{d}y, \quad \overline{V}_{j+\frac{1}{2},k}^x :\approx \int_{I_{j+\frac{1}{2},k}^x} V(x,t) \, \mathrm{d}x \mathrm{d}y, \quad \overline{V}_{j,k+\frac{1}{2}}^y :\approx \int_{I_{j,k+\frac{1}{2}}^y} V(x,t) \, \mathrm{d}x \mathrm{d}y.$$



Figure 3.1: Sketch of the overlapping cells: $I_{j,k}$ (black-bordered); $I_{j-\frac{1}{2},k}^x$ and $I_{j+\frac{1}{2},k}^X$ (red-filled); $I_{j,k-\frac{1}{2}}^y$ and $I_{j,k+\frac{1}{2}}^y$ (blue-filled).

As in the 1-D case, the semi-discretization of the conservative system (1.1) is obtained by integrating (1.1) in space over the cells $I_{j,k}$, which leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{U}_{j,k} = -\frac{1}{\Delta x} \Big[\mathcal{F}_{j+\frac{1}{2},k} - \mathcal{F}_{j-\frac{1}{2},k} \Big] - \frac{1}{\Delta y} \Big[\mathcal{G}_{j,k+\frac{1}{2}} - \mathcal{G}_{j,k-\frac{1}{2}} \Big],$$

with the following second-order numerical fluxes:

$$\mathcal{F}_{j+\frac{1}{2},k} = F(U_{j+\frac{1}{2},k}), \ U_{j+\frac{1}{2},k} := U(\overline{V}_{j+\frac{1}{2},k}^x) \text{ and } \mathcal{G}_{j,k+\frac{1}{2}} = G(U_{j,k+\frac{1}{2}}), \ U_{j,k+\frac{1}{2}} := U(\overline{V}_{j,k+\frac{1}{2}}^y).$$

Two *independent* semi-discretizations of the nonconservative system (1.2) are obtained through the modified 2-D PCCU scheme applied on the $I_{j+\frac{1}{2},k}^x$ and $I_{j,k+\frac{1}{2}}^y$ meshes, respectively, as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t} \overline{\mathbf{V}}_{j+\frac{1}{2},k}^{x} = -\frac{1}{\Delta x} \Big[\widetilde{\mathbf{\mathcal{F}}}_{j+1,k}^{x} - \widetilde{\mathbf{\mathcal{F}}}_{j,k}^{x} - \mathbf{B}_{j+\frac{1}{2},k}^{x} - \frac{a_{j,k}^{x,+}}{a_{j,k}^{x,+} - a_{j,k}^{x,-}} \mathbf{B}_{\Psi,j,k}^{x} + \frac{a_{j+1,k}^{x,-}}{a_{j+1,k}^{x,+} - a_{j+1,k}^{x,-}} \mathbf{B}_{\Psi,j+1,k}^{x} \Big]
- \frac{1}{\Delta y} \Big[\widetilde{\mathbf{\mathcal{G}}}_{j+\frac{1}{2},k+\frac{1}{2}}^{x} - \widetilde{\mathbf{\mathcal{G}}}_{j+\frac{1}{2},k-\frac{1}{2}}^{x} - \mathbf{C}_{j+\frac{1}{2},k}^{x} \\ - \frac{b_{j+\frac{1}{2},k-\frac{1}{2}}^{x,+} - b_{j+\frac{1}{2},k-\frac{1}{2}}^{x}}{b_{j+\frac{1}{2},k-\frac{1}{2}}^{x,-} - b_{j+\frac{1}{2},k-\frac{1}{2}}^{x}} \mathbf{C}_{\Psi,j+\frac{1}{2},k+\frac{1}{2}}^{x,-} - \frac{b_{j+\frac{1}{2},k+\frac{1}{2}}^{x,+} - b_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-} - b_{j+\frac{1}{2},k+\frac{1}{2}}^{x}} \Big] \\ \frac{\mathrm{d}}{\mathrm{d}t} \overline{\mathbf{V}}_{j,k+\frac{1}{2}}^{y} = -\frac{1}{\Delta x} \Big[\widetilde{\mathbf{\mathcal{F}}}_{j+\frac{1}{2},k+\frac{1}{2}}^{y} - \widetilde{\mathbf{\mathcal{F}}}_{j-\frac{1}{2},k+\frac{1}{2}}^{y} - \mathbf{B}_{j,k+\frac{1}{2}}^{y} \\ - \frac{a_{j+\frac{1}{2},k+\frac{1}{2}}^{y,+} - \widetilde{\mathbf{\mathcal{F}}}_{j-\frac{1}{2},k+\frac{1}{2}}^{y} - \mathbf{B}_{j,k+\frac{1}{2}}^{y} - \mathbf{B}_{j,k+\frac{1}{2}}^{y} - a_{j+\frac{1}{2},k+\frac{1}{2}}^{y,-} - a_{j+\frac{1}{2},k+\frac{1}{2}}^{y,-} B_{\Psi,j+\frac{1}{2},k+\frac{1}{2}}^{y,-} \Big] \\ - \frac{1}{\Delta y} \Big[\widetilde{\mathbf{\mathcal{G}}}_{j,k+1}^{y} - \widetilde{\mathbf{\mathcal{G}}}_{j,k}^{y} - \mathbf{C}_{j,k+\frac{1}{2}}^{y} - \frac{b_{j,k+\frac{1}{2}}^{y,+} - b_{j,k}^{y,-}}{b_{j,k}^{y,+} - b_{j,k}^{y,-}} \mathbf{C}_{\Psi,j,k}^{y} + \frac{b_{j,k+1}^{y,-} - b_{j,k+1}^{y,-} - b_{j,k+1}^{y,-}}{b_{j,k+1}^{y,+} - b_{j,k+1}^{y,-}} \mathbf{C}_{\Psi,j,k+1}^{y} \Big]$$

$$(3.2)$$

Once again, we stress that the right-hand sides (RHSs) of (3.1) and (3.2) are computed independently and thus their computations can be performed in parallel.

The terms on the RHS of (3.1) are computed in a dimension-by-dimension manner and given by

$$\begin{split} \widetilde{\mathcal{F}}_{j,k}^{x} &= \frac{a_{j,k}^{x,+} F(\mathbf{V}_{j,k}^{x,-}) - a_{j,k}^{x,-} F(\mathbf{V}_{j,k}^{x,+})}{a_{j,k}^{x,+} - a_{j,k}^{x,-}} + \frac{a_{j,k}^{x,+} a_{j,k}^{x,-}}{a_{j,k}^{x,+} - a_{j,k}^{x,-}} \Big(\mathbf{V}_{j,k}^{x,+} - \mathbf{V}_{j,k}^{x,-} - \delta \mathbf{V}_{j,k}^{x} \Big), \\ \delta \mathbf{V}_{j,k}^{x} &= \min \left(\mathbf{V}_{j,k}^{x,+} - \mathbf{V}_{j,k}^{x,-}, \mathbf{V}_{j,k}^{x,+} - \mathbf{V}_{j,k}^{x,+} \right), \quad \mathbf{V}_{j,k}^{x,+} &= \frac{a_{j,k}^{x,+} \mathbf{V}_{j,k}^{x,-} - \widetilde{F}(\mathbf{V}_{j,k}^{x,+}) + \widetilde{F}(\mathbf{V}_{j,k}^{x,-})}{a_{j,k}^{x,+} - a_{j,k}^{x,-}} \Big), \\ \widetilde{\mathbf{G}}_{j+\frac{1}{2},k+\frac{1}{2}}^{x} &= \frac{b_{j+\frac{1}{2},k+\frac{1}{2}}^{x,+\frac{1}{2}} \widetilde{G}(\mathbf{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-} - \mathbf{O}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-} - \mathbf{O}_{j+\frac{1}{2},k+\frac{1}{2}}^{$$

where $V_{j,k}^{x,\pm}$ and $V_{j+\frac{1}{2},k+\frac{1}{2}}^{x,\pm}$ are reconstructed point values obtained from the cell averages $\{\overline{V}_{j,k}^x\}$. As in the 1-D case, we compute these values by exploiting the local characteristic decomposition. In the *x*direction, we have

$$V_{j,k}^{x,+} = Q_{j+\frac{1}{2},k}^x \Gamma_{j,k}^{x,+}, \quad V_{j+1,k}^{x,-} = Q_{j+\frac{1}{2},k}^x \Gamma_{j+1,k}^{x,-},$$

where

$$\Gamma_{j,k}^{x,+} = \Gamma_{j+\frac{1}{2},k}^{x} - \frac{\Delta x}{2} (\Gamma_{x})_{j+\frac{1}{2},k}^{x}, \quad \Gamma_{j+1,k}^{x,-} = \Gamma_{j+\frac{1}{2},k}^{x} + \frac{\Delta x}{2} (\Gamma_{x})_{j+\frac{1}{2},k}^{x}$$

with the slopes computed using the generalized minmod limiter:

$$(\Gamma_{x})_{j+\frac{1}{2},k}^{x} = \operatorname{minmod}\left(\theta \frac{\Gamma_{j+\frac{1}{2},k}^{x} - \Gamma_{j-\frac{1}{2},k}^{x}}{\Delta x}, \frac{\Gamma_{j+\frac{3}{2},k}^{x} - \Gamma_{j-\frac{1}{2},k}^{x}}{2\Delta x}, \theta \frac{\Gamma_{j+\frac{3}{2},k}^{x} - \Gamma_{j+\frac{1}{2},k}^{x}}{\Delta x}\right), \quad \theta \in [0,1].$$
(3.3)

In (3.3), one has

$$\Gamma_{j-\frac{1}{2},k}^{x} = (\mathcal{Q}_{j+\frac{1}{2},k}^{x})^{-1} \overline{V}_{j-\frac{1}{2},k}^{x}, \quad \Gamma_{j+\frac{1}{2},k}^{x} = (\mathcal{Q}_{j+\frac{1}{2},k}^{x})^{-1} \overline{V}_{j+\frac{1}{2},k}^{x}, \quad \Gamma_{j+\frac{3}{2},k}^{x} = (\mathcal{Q}_{j+\frac{1}{2},k}^{x})^{-1} \overline{V}_{j+\frac{3}{2},k}^{x},$$

where the matrix $Q_{j+\frac{1}{2},k}^x$ is such that $(Q_{j+\frac{1}{2},k}^x)^{-1} \mathcal{A}_{j+\frac{1}{2},k}^x Q_{j+\frac{1}{2},k}^x$ is diagonal with $\mathcal{A}_{j+\frac{1}{2},k}^x = \mathcal{A}(\overline{V}_{j+\frac{1}{2},k}^x) :=$ $\frac{\partial \widetilde{F}}{\partial V}(\overline{V}_{j+\frac{1}{2},k}^{x}) - B(\overline{V}_{j+\frac{1}{2},k}^{x}).$ Similarly, in the *y*-direction, we have

$$\boldsymbol{V}_{j+\frac{1}{2},k-\frac{1}{2}}^{x,+} = P_{j+\frac{1}{2},k}^{x} \boldsymbol{\Gamma}_{j+\frac{1}{2},k-\frac{1}{2}}^{x,+}, \quad \boldsymbol{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-} = P_{j+\frac{1}{2},k}^{x} \boldsymbol{\Gamma}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-},$$

where

$$\Gamma_{j+\frac{1}{2},k-\frac{1}{2}}^{x,+} = \Gamma_{j+\frac{1}{2},k}^{x} - \frac{\Delta y}{2} (\Gamma_{y})_{j+\frac{1}{2},k}^{x}, \quad \Gamma_{j+\frac{1}{2},k}^{x,-} = \Gamma_{j+\frac{1}{2},k}^{x} + \frac{\Delta y}{2} (\Gamma_{y})_{j+\frac{1}{2},k}^{x}$$

with the slope given by

$$(\Gamma_{y})_{j+\frac{1}{2},k}^{x} = \operatorname{minmod}\left(\theta \frac{\Gamma_{j+\frac{1}{2},k}^{x} - \Gamma_{j+\frac{1}{2},k-1}^{x}}{\Delta y}, \frac{\Gamma_{j+\frac{1}{2},k+1}^{x} - \Gamma_{j+\frac{1}{2},k-1}^{x}}{2\Delta y}, \theta \frac{\Gamma_{j+\frac{1}{2},k+1}^{x} - \Gamma_{j+\frac{1}{2},k}^{x}}{\Delta y}\right), \quad \theta \in [0,1].$$
(3.4)

In (3.4), one has

$$\Gamma_{j+\frac{1}{2},k-1}^{x} = (P_{j+\frac{1}{2},k}^{x})^{-1}\overline{V}_{j+\frac{1}{2},k-1}^{x}, \quad \Gamma_{j+\frac{1}{2},k}^{x} = (P_{j+\frac{1}{2},k}^{x})^{-1}\overline{V}_{j+\frac{1}{2},k}^{x}, \quad \Gamma_{j+\frac{1}{2},k+1}^{x} = (P_{j+\frac{1}{2},k}^{x})^{-1}\overline{V}_{j+\frac{1}{2},k+1}^{x},$$

where the matrix $P_{j+\frac{1}{2},k}^x$ is such that $(P_{j+\frac{1}{2},k}^x)^{-1}\mathcal{B}_{j+\frac{1}{2},k}^x P_{j+\frac{1}{2},k}^x$ is diagonal with $\mathcal{B}_{j+\frac{1}{2},k}^x = \mathcal{B}(\overline{V}_{j+\frac{1}{2},k}^x) := \frac{\partial \widetilde{G}}{\partial V}(\overline{V}_{j+\frac{1}{2},k}^x) - C(\overline{V}_{j+\frac{1}{2},k}^x)$.

The one-sided local speeds of propagation in (3.1) are estimated by

$$a_{j,k}^{x,-} = \min\left\{\lambda_1(\mathbf{V}_{j,k}^{x,-}), \lambda_1(\mathbf{V}_{j,k}^{x,+}), 0\right\}, \qquad a_{j,k}^{x,+} = \max\left\{\lambda_M(\mathbf{V}_{j,k}^{x,-}), \lambda_M(\mathbf{V}_{j,k}^{x,+}), 0\right\}, \\ b_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-} = \min\left\{\mu_1(\mathbf{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-}), \mu_1(\mathbf{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,+}), 0\right\}, \qquad b_{j+\frac{1}{2},k+\frac{1}{2}}^{x,+} = \max\left\{\mu_M(\mathbf{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,-}), \mu_M(\mathbf{V}_{j+\frac{1}{2},k+\frac{1}{2}}^{x,+}), 0\right\},$$

where $\lambda_1(V) \leq \ldots \leq \lambda_M(V)$ are the eigenvalues of $\mathcal{A}(V)$ and $\mu_1(V) \leq \ldots \leq \mu_M(V)$ are the eigenvalues of $\mathcal{B}(V)$.

The terms on the RHS of (3.2), related to the update of $\overline{V}_{j,k+\frac{1}{2}}^{y}$, can be analogously computed through the same formulae used to evaluate the RHS of (3.1), but with the first index shifted by $-\frac{1}{2}$, the second index shifted by $+\frac{1}{2}$, and the superscripts "x" replaced by "y".

Remark 3.1 We would like to emphasize that, as in the 1-D case, the proposed 2-D FV AF method is not tied to the PCCU scheme and the nonconservative system (1.2) can be numerically solved by an alternative second-order stable numerical method.

Finally, the post-processing is carried out in a "dimension-by-dimension" manner by alternating sweeps in the *x*-direction, in which we modify the cell averages of \overline{U} and \overline{V}^x , and in the *y*-direction, in which we modify the cell averages of \overline{U} and \overline{V}^y . We use the same notation as in the 1-D case: $\{\overline{U}_{j,k}^*\}$, $\{\overline{V}_{j+\frac{1}{2},k}^{x,*}\}$, and $\{\overline{V}_{j,k+\frac{1}{2}}^{y,*}\}$ denote the solution values evolved from time *t* to $t + \Delta t$ by an ODE solver, and our goal is to post-process them to obtain $\{\overline{U}_{j,k}(t + \Delta t)\}, \{\overline{V}_{j+\frac{1}{2},k}^x(t + \Delta t)\}$ and $\{\overline{V}_{j,k+\frac{1}{2}}^y(t + \Delta t)\}$.

The proposed post-processing can be presented algorithmically through the following eight steps, with Steps 1–4 corresponding to the post-processing in the x-direction and Steps 5–8 corresponding to the post-processing in the y-direction.

Step 1. Compute the conserved variables at $(x_{j+\frac{1}{2}}, y_k)$ using the transformation from the primitive ones:

$$\boldsymbol{U}_{j+\frac{1}{2},k}^{*} = \boldsymbol{U}(\overline{\boldsymbol{V}}_{j+\frac{1}{2},k}^{x,*}).$$

Step 2. Perform the piecewise linear reconstruction for U in the x-direction using the slopes computed by the minmod limiter

$$(\boldsymbol{U}_{x})_{j,k}^{*} = 2 \operatorname{minmod}\left(\frac{\overline{\boldsymbol{U}}_{j,k}^{*} - \boldsymbol{U}_{j-\frac{1}{2},k}^{*}}{\Delta x}, \frac{\boldsymbol{U}_{j+\frac{1}{2},k}^{*} - \overline{\boldsymbol{U}}_{j,k}^{*}}{\Delta x}\right),$$

which results in

$$\boldsymbol{U}_{j+\frac{1}{2},k}^{*,-} := \overline{\boldsymbol{U}}_{j,k}^{*} + \frac{\Delta x}{2} (\boldsymbol{U}_{x})_{j,k}^{*}, \quad \boldsymbol{U}_{j+\frac{1}{2},k}^{*,+} := \overline{\boldsymbol{U}}_{j+1,k}^{*} - \frac{\Delta x}{2} (\boldsymbol{U}_{x})_{j+1,k}^{*}.$$

Step 3. Set

$$\boldsymbol{U}_{j+\frac{1}{2},k}^{**} := \frac{1}{2} \Big(\boldsymbol{U}_{j+\frac{1}{2},k}^{*,-} + \boldsymbol{U}_{j+\frac{1}{2},k}^{*,+} \Big),$$

and recompute the primitive variables at $(x_{j+\frac{1}{2}}, y_k)$ using the transformation from the conservative ones:

$$\overline{\boldsymbol{V}}_{j+\frac{1}{2},k}^{x}(t+\Delta t) = \boldsymbol{V}(\boldsymbol{U}_{j+\frac{1}{2},k}^{**}).$$

Step 4. Correct the conserved variables by setting

$$\widetilde{\boldsymbol{U}}_{j,k} := \frac{1}{2} \Big(\boldsymbol{U}_{j-\frac{1}{2},k}^{**} + \boldsymbol{U}_{j+\frac{1}{2},k}^{**} \Big).$$

Step 5. Compute the conserved variables at $(x_j, y_{k+\frac{1}{2}})$ using the transformation from the primitive ones:

$$\boldsymbol{U}^*_{j,k+\frac{1}{2}} = \boldsymbol{U}(\overline{\boldsymbol{V}}^{y,*}_{j,k+\frac{1}{2}}).$$

Step 6. Perform the piecewise linear reconstruction for U in the y-direction using the slopes computed by the minmod limiter

$$(\boldsymbol{U}_{y})_{j,k}^{*} = 2 \operatorname{minmod}\left(\frac{\widetilde{\boldsymbol{U}}_{j,k} - \boldsymbol{U}_{j,k-\frac{1}{2}}^{*}}{\Delta y}, \frac{\boldsymbol{U}_{j,k+\frac{1}{2}}^{*} - \widetilde{\boldsymbol{U}}_{j,k}}{\Delta y}\right),$$

which results in

$$\boldsymbol{U}_{j,k+\frac{1}{2}}^{*,-} := \widetilde{\boldsymbol{U}}_{j,k} + \frac{\Delta y}{2} (\boldsymbol{U}_{y})_{j,k}^{*}, \quad \boldsymbol{U}_{j,k-\frac{1}{2}}^{*,+} := \widetilde{\boldsymbol{U}}_{j,k+1}^{*} - \frac{\Delta y}{2} (\boldsymbol{U}_{y})_{j,k+1}^{*}.$$

Step 7. Set

$$\boldsymbol{U}_{j,k+\frac{1}{2}}^{**} := \frac{1}{2} \Big(\boldsymbol{U}_{j,k+\frac{1}{2}}^{*,-} + \boldsymbol{U}_{j,k+\frac{1}{2}}^{*,+} \Big),$$

and recompute the primitive variables at $(x_j, y_{k+\frac{1}{2}})$ using the transformation from the conservative ones:

$$\overline{\boldsymbol{V}}_{j,k+\frac{1}{2}}^{y}(t+\Delta t) = \boldsymbol{V}(\boldsymbol{U}_{j,k+\frac{1}{2}}^{**}).$$

Step 8. Correct the conserved variables by setting

$$\overline{\boldsymbol{U}}_{j,k}(t+\Delta t) = \frac{1}{2} \Big(\boldsymbol{U}_{j,k-\frac{1}{2}}^{**} + \boldsymbol{U}_{j,k+\frac{1}{2}}^{**} \Big).$$

4 Numerical Examples

In this section, we present the numerical results obtained for the Euler equations of gas dynamics. In the 1-D case, these equations can be written in either the conservative form (2.1) with

$$\boldsymbol{U} = (\rho, \rho \boldsymbol{u}, E)^{\mathsf{T}}, \quad \boldsymbol{F}(\boldsymbol{U}) = (\rho \boldsymbol{u}, \rho \boldsymbol{u}^2 + p, \boldsymbol{u}(E+p))^{\mathsf{T}},$$

and the equation of state $E = \frac{p}{\gamma-1} + \frac{1}{2}\rho u^2$, or the nonconservative form (2.2) with

$$\boldsymbol{V} = (\rho, u, p)^{\mathsf{T}}, \quad \widetilde{\boldsymbol{F}}(\boldsymbol{V}) = \left(\rho u, \frac{u^2}{2}, pu\right)^{\mathsf{T}}, \quad \boldsymbol{B}(\boldsymbol{V}) = \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & -\frac{1}{\rho}\\ 0 & -(\gamma - 1)p & 0 \end{pmatrix}.$$

In all of the numerical examples below, the time discretization is performed using the three-stage third-order strong stability-preserving Runge-Kutta method (see, e.g., [22, 23]) with the time-step selected adaptively with the CFL number set to be 0.25. The minmod parameter has been taken $\theta = 1.3$ in all of the examples except for Example 5, in which we have used $\theta = 1.1$. In all of the examples, we set $\gamma = 1.4$.

Example 1 (accuracy test for unsteady isentropic vortex). In the first example taken from [7,35,42], we numerically verify the accuracy of the proposed FV AF method.

We consider a smooth unsteady vortex on the computational domain $[-10, 10] \times [-10, 10]$, endowed with periodic boundary conditions. The initial data are

$$\begin{aligned} \rho(x, y, 0) &= \left(1 - \frac{(\gamma - 1)\kappa^2}{2\gamma}\right)^{\frac{1}{\gamma - 1}}, \quad p(x, y, 0) = \rho^{\gamma}(x, y, 0), \\ u(x, y, 0) &= 1 - \kappa y, \quad v(x, y, 0) = 1 + \kappa x, \quad \kappa = \frac{5}{2\pi}e^{\frac{1 - \kappa^2 - \gamma^2}{2}}. \end{aligned}$$

The exact solution is given by U(x, y, t) = U(x - t, y - t, 0), modulo the periodic boundary conditions.

We compute the solution until the final time t = 0.1 on a sequence of uniform $N \times N$ meshes with N = 100, 200, 400, 800, and 1600. The obtained results are reported in Table 4.1, in which we demonstrate that the expected order of convergence is achieved in the *U*-solution (we show the ρ -, ρu -, and *E*-components), V^x -solution (we show the *v*-component), and V^y -solution (we show the *p*-component). The same order is observed in all other components not shown here for the sake of brevity.

Ν	ρ -error	rate	<i>ρu</i> -error	rate	<i>E</i> -error	rate	v-error	rate	<i>p</i> -error	rate
100	1.29e-2	_	2.88e-2	_	7.09e-2	_	6.18e-2	_	3.48e-2	_
200	3.46e-3	1.90	6.99e-3	2.05	1.78e-2	2.00	1.59e-2	1.96	8.11e-3	2.10
400	9.72e-4	1.83	1.77e-3	1.98	4.56e-3	1.96	4.13e-3	1.94	1.93e-3	2.07
800	2.50e-4	1.96	4.32e-4	2.03	1.11e-3	2.04	1.04e-3	1.99	4.59e-4	2.07
1600	5.93e-5	2.08	1.02e-4	2.09	2.48e-4	2.16	2.56e-4	2.02	1.07e-4	2.11

Table 4.1: Example 1: L^1 -errors and corresponding convergence rates for the ρ -, ρu -, and Ecomponents of the U-solution, v-component of the V^x -solution, and p-component of the V^y solution.

It should be observed that in this example, the solution is smooth and thus it can, in principle, be computed without the post-processing. We have verified that in this case, the second order of convergence is also obtained (the results are omitted for the sake of brevity).

Example 2 (Sod shock-tube problem). In the second example, we consider the Sod shock-tube problem from [44]. The Riemann initial data,

$$\boldsymbol{V}(x,0) = \begin{cases} (1,0,1)^{\top}, & x < 0.5, \\ (0.125,0,0.1)^{\top}, & \text{otherwise.} \end{cases}$$

are prescribed in the computational domain [0, 1] with the free boundary conditions.

We compute the solution until the final time t = 0.2 on a uniform mesh with N = 200 and plot it in Figure 4.1 together with the exact solution, obtained using the library NUMERICA [46], and the solution computed by the FV AF method but without post-processing. As one can see, the FV AF solution is oscillation-free and its rather sharply resolved discontinuities are located at correct locations thanks to the conservative post-processing. On the contrary, the FV AF method without the post-processing produces very large spurious spikes in the U-solution (note that these spikes are located exactly where the initial condition was discontinuous) and a non-oscillatory V-solution, which apparently captures a wrong weak solution as expected (see [4]).



Figure 4.1: Example 2: U-solutions (upper row) and V-solutions (lower row) computed by the FV AF method and the FV AF method without the post-processing.

Example 3 (double rarefaction problem). In this problem taken from [47], we consider a Riemann problem, whose solution contains two rarefaction waves, which expand and form a near-vacuum-area in the middle of the computational domain [0, 1]. The initial conditions,

$$\boldsymbol{V}(x,0) = \begin{cases} (1,-2,0.4)^{\mathsf{T}}, & x < 0.5, \\ (1,2,0.4)^{\mathsf{T}}, & \text{otherwise,} \end{cases}$$

are supplemented with the free boundary conditions.

We compute the solution until the final time t = 0.15 on a uniform mesh with N = 200 and plot the obtained results together with exact solution, once again generated using the library NUMERICA [46] in Figure 4.2. As one can see, the computed solution is oscillation-free and preserve the positivity of both the density and pressure, which is not the case for many existing schemes.



Figure 4.2: Example 3: U-solutions (upper row) and V-solutions (lower row).

Example 4 (shock-turbulence interaction problem). In this example taken from [43], a shock interacts with a turbulent flow characterized by high-frequency oscillations.

The initial data,

$$\boldsymbol{V}(x,0) = \begin{cases} (3.857143, 2.629369, 10.333333)^{\mathsf{T}}, & x < -4, \\ (1+0.2\sin(5x), 0, 1)^{\mathsf{T}}, & \text{otherwise}, \end{cases}$$

are prescribed in the computational domain [-5, 5] with the inflow boundary conditions at the left boundary and free boundary conditions at the right one.

We compute the solution until the final time t = 1.8 on a uniform mesh with N = 1000 and report the obtained results in Figure 4.3. Since in this example no exact solution is available, we compute the reference solution (shown in Figures 4.3 and 4.4) by a second-order semi-discrete CU scheme from [25] on a much finer uniform mesh consisting of 200000 cells. As one can see, the solution structures are correctly resolved.

We also compare the performance of the FV AF method with a recently proposed low-dissipation central-upwind (LDCU) scheme from [12] (implemented on the same mesh, using the same spatial reconstruction of the characteristic variables of the conserved system and time discretization). The obtained results are shown in Figure 4.4, where one can see that the FV AF method better resolves smooth, oscillating parts of the computed solution.



Figure 4.3: Example 4: U-solutions (upper row) and V-solutions (lower row).



Figure 4.4: Example 4: Densities computed by the FV AF method and LDCU scheme (left) and zoom at $x \in [0.5, 2.5]$ (right).

Example 5 (Woodward-Colella problem). In this problem, which was introduced in [49], the initial conditions,

$$V(x,0) = \begin{cases} (1,0,10^3)^{\top}, & x < 0.1, \\ (1,0,10^2)^{\top}, & x > 0.9, \\ (1,0,10^{-2})^{\top}, & \text{otherwise} \end{cases}$$

are prescribed in the computational domain [0, 1] with solid wall boundary conditions.

In Figure 4.5, we plot the solution computed at the final time t = 0.038 on a uniform mesh with N = 400 along with the reference solution generated by a second-order semi-discrete CU scheme from



Figure 4.5: Example 5: U-solutions (upper row) and V-solutions (lower row).

Example 6 (explosion problem). This problem taken from [47], is a multidimensional extension of the Sod shock-tube problem considered in Example 2.

We take the following initial conditions:

$$\boldsymbol{V}(x, y, 0) = \begin{cases} (1, 0, 0, 1)^{\mathsf{T}}, & \sqrt{x^2 + y^2} < 0.4, \\ (0.125, 0, 0, 0.1)^{\mathsf{T}}, & \text{otherwise}, \end{cases}$$

which are prescribed in the computational domain $[-1, 1] \times [-1, 1]$ with the free boundary conditions.

We compute the solution until the final time t = 0.25 on a uniform mesh with $N_x = N_y = 400$. In Figure 4.6, we report three-dimensional plots of the density- and energy-components of the *U*-solution and the density- and pressure-components of the V^x -solution (we do not show the V^y -solution as it is practically the same as the V^x -one). The 1-D slices of ρ , $\sqrt{(\rho u)^2 + (\rho v)^2}$, and *E* from the *U*-solution and of ρ , $\sqrt{u^2 + v^2}$, and *p* from the V^x -solution along the line y = x are shown in Figure 4.7 (the reference solution plotted there has been obtained by a second-order semi-discrete FV scheme with reconstruction of characteristic variables and exact Riemann solver numerical flux on a much finer uniform mesh consisting of 3000×3000 cells). As one can see, neither spurious oscillations nor anomalous features are visible, and the solution is sharply captured. One can also observe that the radial symmetry is preserved.



Figure 4.6: Example 6: Density- and energy-components of the U-solution (upper row) and densityand pressure-components of the V^x -solution (lower row).

Example 7 (shock-vortex interaction). In the last 2-D example, which was studied in, e.g., [17,20,39], we consider a moving vortex with Mach number $M_v := 0.9$ interacting with a stationary shock with Mach number $M_s := 1.5$. Such an interaction gives rise to complex flow tructures, making this numerical test challenging for higher-order numerical schemes.

A schematic of the initial setup is shown in Figure 4.8. Specifically, the computational domain $[0, 2] \times [0, 1]$ is divided into two main subdomains by the vertical line x = 0.5. The vortex is initially located within a circular area centered in $(x_c, y_c) := (0.25, 0.5)$ and occupies regions I and II: the first one is a circle of radius a = 0.075 and the latter one is a concentrical annulus with inner radius a and outer radius b = 0.175. The states in regions III and IV correspond to a stationary shock. In particular, given the left constant state from region III,

$$\rho_{\rm III} = 1, \quad u_{\rm III} = \sqrt{\gamma} M_s, \quad v_{\rm III} = 0, \quad p_{\rm III} = 1,$$

the right state, $(\rho_{IV}, u_{IV}, v_{IV}, p_{IV})$, can be easily computed through the Rankine-Hugoniot conditions (see [47, Section 3.1.3]), leading to

$$\rho_{\rm IV} = \frac{(\gamma+1)M_s^2}{(\gamma-1)M_s^2+2}\rho_{\rm III}, \quad u_{\rm IV} = \frac{(\gamma-1)M_s^2+2}{(\gamma+1)M_s^2}u_{\rm III} \quad v_{\rm IV} = 0, \quad p_{\rm IV} = \frac{2\gamma M_s^2 - (\gamma-1)}{\gamma+1}p_{\rm III}.$$



Figure 4.7: Example 6: 1-D slices along y = x for the *U*-solution (upper row) and *V*^{*x*}-solution (lower row).



Figure 4.8: Example 7: Sketch of the simulation set-up.

The velocity profile in vortex regions I and II is given, in terms of radial coordinates (r, ϑ) with respect to the center (x_c, y_c) :

$$u(r, \vartheta) = u_{\text{III}} - v_{\vartheta} \sin \vartheta, \quad v(r, \vartheta) = v_{\text{III}} + v_{\vartheta} \cos \vartheta,$$

where

$$v_{\vartheta} := \begin{cases} v_m \frac{r}{a}, & r \le a, \\ v_m \frac{a}{a^2 - b^2} \left(r - \frac{b^2}{r} \right), & a < r < b \\ 0, & r \ge b, \end{cases}$$

with $v_m := M_v \sqrt{\gamma}$ being the maximal angular velocity.

Density and pressure profiles inside the vortex are obtained imposing a balance between centripetal force and pressure gradients [39], resulting in

$$p = p_{\mathrm{III}} \left(\frac{T}{T_{\mathrm{III}}} \right)^{\frac{\gamma}{\gamma-1}}, \quad \rho = \rho_{\mathrm{III}} \left(\frac{T}{T_{\mathrm{III}}} \right)^{\frac{1}{\gamma-1}}$$

where $T_{\text{III}} = \frac{p_{\text{III}}}{\rho_{\text{III}}R}$ is the constant temperature associated with the state of region III, with R = 287 J/kg-K being the specific gas constant of the fluid, and T(r) is the temperature within the vortex:

$$T(r) = \begin{cases} A + \frac{\gamma - 1}{R\gamma} \frac{v_m^2}{a^2} \frac{r^2}{2}, & r \le a, \\ B + \frac{\gamma - 1}{R\gamma} v_m^2 \frac{a^2}{(a^2 - b^2)^2} \left(\frac{r^2}{2} - 2b^2 \ln r - b^4 \frac{r^{-2}}{2}\right), & a < r < b, \\ T_{\text{III}}, & r \ge b, \end{cases}$$

where

$$\begin{split} B &= T_{\mathrm{III}} - \frac{\gamma - 1}{R\gamma} v_m^2 \frac{a^2}{(a^2 - b^2)^2} \left(\frac{b^2}{2} - 2b^2 \ln b - b^4 \frac{b^{-2}}{2} \right), \\ A &= B + \frac{\gamma - 1}{R\gamma} v_m^2 \frac{a^2}{(a^2 - b^2)^2} \left(\frac{a^2}{2} - 2b^2 \ln a - b^4 \frac{a^{-2}}{2} \right) - \frac{\gamma - 1}{R\gamma} \frac{v_m^2}{2}. \end{split}$$

Figure 4.9 shows a Schlieren image of the numerical results displaying the magnitude of the density gradient field, $\|\nabla \rho\|_2$, computed on a uniform mesh with $N_x = 1200$ and $N_y = 601$ at time t = 0.7. In this figure, we have used the following shading function:

$$\exp\left(-\frac{K||\nabla\rho||_2}{\max||\nabla\rho||_2}\right),\,$$

where the numerical density derivatives are computed using central differencing. The computed results agree well with those obtained in the literature.

5 Conclusions

In this paper, we have introduced a new active flux (AF) method for hyperbolic systems of conservation laws, for which we consider two different formulations of the studied system (the original conservative formulation and a primitive one containing nonconservative products), and discretize them on overlapping staggered meshes using two different numerical methods. Both the conservative and primitive variables are evolved in time using second-order semi-discrete finite-volume (FV) methods. The non-conservative system is discretized by a path-conservative central-upwind scheme and its solution is used



Figure 4.9: Example 7: Schlieren image of the density gradient.

to evaluate very simple numerical fluxes for the original conservative system. The nonlinear stability of the resulting FV AF method is enforced with the help of a post-processing, which also guarantees a conservative coupling between the two sets of variables. The performance of the proposed method has been demonstrated on a number of one- and two-dimensional benchmarks.

We would like to emphasize that the main difficulty in designing higher-order extensions of the proposed FV AF method is related to developing a high-order post-processing. Such post-processing should be based on a uniformly accurate reconstruction/interpolation. One can use, for example, CWENO reconstructions/interpolations (see [14, 15, 41] and references therein), and we plan to explore this possibility in our future work.

We would also like to point out that having two sets of data for the discrete solution, which we evolve at each evolution step, has been used to design a smoothness indicator based on the difference between these two solutions; see [11]. This smoothness indicator can be used to develop different adaptation strategies, which may substantially enhance the resolution achieved by the FV AF method.

Acknowledgment

The work of A. Chertock was supported in part by NSF grant DMS-2208438. The work of A. Kurganov was supported in part by NSFC grants 12171226 and W2431004. The work of L. Micalizzi was supported in part by the LeRoy B. Martin, Jr. Distinguished Professorship Foundation.

In memoriam

This paper is dedicated to the memory of Prof. Arturo Hidalgo López (*July 03rd 1966 - †August 26th 2024) of the Universidad Politecnica de Madrid, organizer of HONOM 2019 and active participant in many other editions of HONOM. Our thoughts and wishes go to his wife Lourdes and his sister María Jesús, whom he left behind.

References

[1] R. ABGRALL, A combination of residual distribution and the active flux formulations or a new class

of schemes that can combine several writings of the same hyperbolic problem: application to the *1D Euler equations*, Commun. Appl. Math. Comput., 5 (2023), pp. 370–402.

- [2] R. ABGRALL AND W. BARSUKOW, Extensions of active flux to arbitrary order of accuracy, ESAIM Math. Model. Numer. Anal., 57 (2023), pp. 991–1027.
- [3] R. ABGRALL, W. BARSUKOW, AND C. KLINGENBERG, A semi-discrete active flux method for the Euler equations on Cartesian grids, J. Sci. Comput., 102 (2025). Paper No. 36.
- [4] R. ABGRALL AND S. KARNI, A comment on the computation of non-conservative products, J. Comput. Phys., 229 (2010), pp. 2759–2763.
- [5] W. BARSUKOW, The active flux scheme for nonlinear problems, J. Sci. Comput., 86 (2021). Paper No. 3.
- [6] W. BARSUKOW, J. HOHM, C. KLINGENBERG, AND P. L. ROE, The active flux scheme on Cartesian grids and its low Mach number limit, J. Sci. Comput., 81 (2019), pp. 594–622.
- [7] W. BOSCHERI AND M. DUMBSER, Arbitrary-Lagrangian-Eulerian one-step WENO finite volume schemes on unstructured triangular meshes, Commun. Comput. Phys., 14 (2013), pp. 1174–1206.
- [8] M. J. CASTRO, P. G. LEFLOCH, M. L. MUÑOZ-RUIZ, AND C. PARÉS, Why many theories of shock waves are necessary: convergence error in formally path-consistent schemes, J. Comput. Phys., 227 (2008), pp. 8107–8129.
- [9] M. J. CASTRO, T. MORALES DE LUNA, AND C. PARÉS, Well-balanced schemes and path-conservative numerical methods, in Handbook of numerical methods for hyperbolic problems, vol. 18 of Handb. Numer. Anal., Elsevier/North-Holland, Amsterdam, 2017, pp. 131–175.
- [10] M. J. CASTRO DÍAZ, A. KURGANOV, AND T. MORALES DE LUNA, Path-conservative central-upwind schemes for nonconservative hyperbolic systems, ESAIM Math. Model. Numer. Anal., 53 (2019), pp. 959–985.
- [11] A. CHERTOCK, A. KURGANOV, AND L. MICALIZZI, New smoothness indicaactive flux framework. within an Submitted: preprint is available tor at https://chertock.wordpress.ncsu.edu/files/2025/05/CKM25.pdf.
- [12] S. CHU, A. KURGANOV, AND R. XIN, New low-dissipation central-upwind schemes. Part II, J. Sci. Comput., 96 (2023). Paper No. 56.
- [13] E. CHUDZIK, C. HELZEL, AND D. KERKMANN, *The Cartesian grid active flux method: linear stability and bound preserving limiting*, Appl. Math. Comput., 393 (2021). Paper No. 125501.
- [14] I. CRAVERO, G. PUPPO, M. SEMPLICE, AND G. VISCONTI, CWENO: uniformly accurate reconstructions for balance laws, Math. Comp., 87 (2018), pp. 1689–1719.
- [15] I. CRAVERO, M. SEMPLICE, AND G. VISCONTI, Optimal definition of the nonlinear weights in multidimensional central WENOZ reconstructions, SIAM J. Numer. Anal., 57 (2019), pp. 2328–2358.
- [16] G. DAL MASO, P. G. LEFLOCH, AND F. MURAT, Definition and weak stability of nonconservative products, J. Math. Pures Appl. (9), 74 (1995), pp. 483–548.

- [17] M. DUMBSER, O. ZANOTTI, R. LOUBÈRE, AND S. DIOT, A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws, J. Comput. Phys., 278 (2014), pp. 47–75.
- [18] T. EYMANN AND P. ROE, *Active flux schemes for systems*, 2011. In: 20th AIAA Computational Fluid Dynamics Conference, Paper No. 3840.
- [19] T. A. EYMANN AND P. L. ROE, *Multidimensional active flux schemes*, 2013. In: 21st AIAA Computational Fluid Dynamics Conference, Paper No. 2940.
- [20] M. GEISENHOFER, F. KUMMER, AND B. MÜLLER, A discontinuous Galerkin immersed boundary solver for compressible flows: adaptive local time stepping for artificial viscosity-based shock-capturing on cut cells, Internat. J. Numer. Methods Fluids, 91 (2019), pp. 448–472.
- [21] E. GODLEWSKI AND P.-A. RAVIART, Numerical approximation of hyperbolic systems of conservation laws, vol. 118 of Applied Mathematical Sciences, Springer-Verlag, New York, 2021. Second edition.
- [22] S. GOTTLIEB, D. KETCHESON, AND C.-W. SHU, Strong stability preserving Runge-Kutta and multistep time discretizations, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011.
- [23] S. GOTTLIEB, C.-W. SHU, AND E. TADMOR, Strong stability-preserving high-order time discretization methods, SIAM Rev., 43 (2001), pp. 89–112.
- [24] J. S. HESTHAVEN, Numerical methods for conservation laws, vol. 18 of Computational Science & Engineering, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2018. From analysis to algorithms.
- [25] A. KURGANOV AND C.-T. LIN, On the reduction of numerical dissipation in central-upwind schemes, Commun. Comput. Phys., 2 (2007), pp. 141–163.
- [26] A. KURGANOV, S. NOELLE, AND G. PETROVA, Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton-Jacobi equations, SIAM J. Sci. Comput., 23 (2001), pp. 707–740.
- [27] P. G. LEFLOCH, *Hyperbolic systems of conservation laws*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2002. The theory of classical and nonclassical shock waves.
- [28] —, Graph solutions of nonlinear hyperbolic systems, J. Hyperbolic Differ. Equ., 1 (2004), pp. 643–689.
- [29] R. J. LEVEQUE, *Finite volume methods for hyperbolic problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.
- [30] K.-A. LIE AND S. NOELLE, On the artificial compression method for second-order nonoscillatory central difference schemes for systems of conservation laws, SIAM J. Sci. Comput., 24 (2003), pp. 1157–1174.
- [31] Y. LIU, Central schemes on overlapping cells, J. Comput. Phys., 209 (2005), pp. 82–104.

- [32] Y. LIU, C.-W. SHU, E. TADMOR, AND M. ZHANG, Central discontinuous Galerkin methods on overlapping cells with a nonoscillatory hierarchical reconstruction, SIAM J. Numer. Anal., 45 (2007), pp. 2442–2467 (electronic).
- [33] —, *Non-oscillatory hierarchical reconstruction for central and finite volume schemes*, Commun. Comput. Phys., 2 (2007), pp. 933–963.
- [34] Y. LIU, C.-W. SHU, AND Z. XU, Hierarchical reconstruction with up to second degree remainder for solving nonlinear conservation laws, Nonlinearity, 22 (2009), pp. 2799–2812.
- [35] L. MICALIZZI, D. TORLO, AND W. BOSCHERI, *Efficient iterative arbitrary high-order methods: an adaptive bridge between low and high order*, Communications on Applied Mathematics and Computation, (2023), pp. 1–38.
- [36] L. MICALIZZI AND E. F. TORO, Impact of numerical fluxes on high order semidiscrete WENO-DeC finite volume schemes, Sib. Electron. Math. Rep., 21 (2024), pp. B232–B287.
- [37] C. PARÉS, Numerical methods for nonconservative hyperbolic systems: a theoretical framework, SIAM J. Numer. Anal., 44 (2006), pp. 300–321.
- [38] —, *Path-conservative numerical methods for nonconservative hyperbolic systems*, in Numerical methods for balance laws, vol. 24 of Quad. Mat., Dept. Math., Seconda Univ. Napoli, Caserta, 2009, pp. 67–121.
- [39] A. RAULT, G. CHIAVASSA, AND R. DONAT, *Shock-vortex interactions at high Mach numbers*, J. Sci. Comput., 19 (2003), pp. 347–371. Special issue in honor of the sixtieth birthday of Stanley Osher.
- [40] P. ROE, Designing CFD methods for bandwidth—a physical approach, Comput. & Fluids, 214 (2021). Paper No. 104774.
- [41] M. SEMPLICE AND G. VISCONTI, Efficient implementation of adaptive order reconstructions, J. Sci. Comput., 83 (2020), pp. Paper No. 6, 27 pp.
- [42] C.-W. SHU, Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws, in Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997), vol. 1697 of Lecture Notes in Math., Springer, Berlin, 1998, pp. 325–432.
- [43] C.-W. SHU AND S. OSHER, Efficient implementation of essentially nonoscillatory shock-capturing schemes. II, J. Comput. Phys., 83 (1989), pp. 32–78.
- [44] 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.
- [45] P. K. SWEBY, High resolution schemes using flux limiters for hyperbolic conservation laws, SIAM J. Numer. Anal., 21 (1984), pp. 995–1011.
- [46] E. F. TORO, NUMERICA: A Library of Source Codes for Teaching, Research and Applications, Numeritek Ltd., Cheshire, UK, www.numeritek.com, 1999.
- [47] E. F. TORO, *Riemann solvers and numerical methods for fluid dynamics: A practical introduction*, Springer-Verlag, Berlin, Heidelberg, third ed., 2009.

- [48] B. VAN LEER, Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection, J. Comput. Phys., 23 (1977), pp. 276–299.
- [49] P. WOODWARD AND P. COLELLA, *The numerical solution of two-dimensional fluid flow with strong shocks*, J. Comput. Phys., 54 (1984), pp. 115–173.
- [50] Z. XU, Y. LIU, H. DU, G. LIN, AND C.-W. SHU, Point-wise hierarchical reconstruction for discontinuous Galerkin and finite volume methods for solving conservation laws, J. Comput. Phys., 230 (2011), pp. 6843–6865.
- [51] Z. XU, Y. LIU, AND C.-W. SHU, Hierarchical reconstruction for discontinuous Galerkin methods on unstructured grids with a WENO-type linear reconstruction and partial neighboring cells, J. Comput. Phys., 228 (2009), pp. 2194–2212.
- [52] X. ZENG, A high-order hybrid finite difference-finite volume approach with application to inviscid compressible flow problems: a preliminary study, Comput. & Fluids, 98 (2014), pp. 91–110.