

# Genuinely multi-dimensional stationarity preserving global flux Finite Volume formulation for nonlinear hyperbolic PDEs

Wasilij Barsukow<sup>a</sup>, Mirco Ciallella<sup>b</sup>, Mario Ricchiuto<sup>c</sup>, Davide Torlo<sup>d</sup>

<sup>a</sup>*Institut de Mathématiques de Bordeaux, Université de Bordeaux, CNRS UMR 5251, Talence, France*

<sup>b</sup>*Laboratoire Jacques-Louis Lions, Université Paris Cité, CNRS UMR 7598, Paris, France*

<sup>c</sup>*Centre Inria de l'Université de Bordeaux, CNRS UMR 5251, Talence, France*

<sup>d</sup>*Dipartimento di Matematica, Università di Roma La Sapienza, Rome, Italy*

---

## Abstract

Classical Finite Volume methods for multi-dimensional problems include stabilization (e.g. via a Riemann solver), that is derived by considering several one-dimensional problems in different directions. Such methods therefore ignore a possibly existing balance of contributions coming from different directions, such as the one characterizing multi-dimensional stationary states. Instead being preserved, they are usually diffused away by such methods. Stationarity preserving methods use a better suited stabilization term that vanishes at the stationary state, allowing the method to preserve it. This work presents a general approach to stationarity preserving Finite Volume methods for nonlinear conservation/balance laws. It is based on a multi-dimensional extension of the global flux approach. The new methods are shown to significantly outperform existing ones even if the latter are of higher order of accuracy and even on non-stationary solutions.

*Keywords:* Stationarity preservation, Global flux, Finite Volume, Multi-dimensional well-balancing, Hyperbolic equations

---

## 1. Introduction

This paper focuses on the numerical solution of nonlinear hyperbolic systems of conservation laws in two dimensions:

$$\partial_t q + \partial_x f + \partial_y g = 0, \quad (1)$$

where  $q$ ,  $f$  and  $g$  are the vectors of conservative variables and fluxes. Numerical methods for hyperbolic partial differential equations (PDEs) need numerical diffusion to achieve entropy stability and in order to deal with solutions characterized by strong gradients. The majority of numerical methods for multi-dimensional problems, though, are developed following a dimension-by-dimension approach, meaning that the numerical diffusion is usually derived in a one-dimensional framework and that the diffusion term associated to an edge (or a face, in 3D) usually involves only two states. Standard numerical methods with one-dimensional Riemann solvers typically introduce a diffusion term of the type

$$\partial_t q + \partial_x f + \partial_y g = \Delta x \partial_x (\nu_x \partial_x q) + \Delta y \partial_y (\nu_y \partial_y q), \quad (2)$$

where  $\Delta x$  and  $\Delta y$  provide the size of the discretization, and  $\nu_x$  and  $\nu_y$  represent the diffusion coefficients, which are often chosen proportional to the spectral radius of the flux Jacobian. This one-dimensional approach does not take into account possible multi-dimensional features of the numerical solution, such as the stationary states characterized by a balance of contributions coming from different directions [8]. For equation (1), stationary states are governed by

$$\partial_x f + \partial_y g = 0. \quad (3)$$

For classical methods, with the two-dimensional diffusion term designed following one-dimensional approaches, the solution will be completely diffused instead of being kept stationary [9, 11]. In contrast to Equation (3), the discrete stationary states are characterized by the much more restrictive conditions  $\partial_x f = 0$  and  $\partial_y g = 0$ . States where  $\partial_x f \neq 0$  is balanced by  $-\partial_y g$  is not a stationary state of the numerical method.

This can be prevented by choosing more sophisticated diffusion operators [40, 45, 28, 39]. Such methods are called stationarity preserving [8]. More recent developments of ad-hoc diffusion operators were developed for geostrophic equilibria in the linear and nonlinear case [4, 5]. A connection to these equilibria in the context of low Mach number limit of the Euler equations, which is related to the long-time limit of linear acoustics, was provided in [30, 31] through the preservation of discrete divergence with ad-hoc functional spaces. Early examples of stationarity preserving methods for nonlinear conservation laws can be found in [7]. So far, however, no general theory for the agnostic detection of stationary states of nonlinear multi-dimensional hyperbolic partial differential equations is available.

The method presented in this work is based on the global flux [27, 16, 18] approach initially introduced for hyperbolic balance laws in one dimension,

$$\partial_t q + \partial_x f = s, \quad (4)$$

with the original goal of developing well-balanced methods [3, 12, 17], and the treatment of source terms present in the mathematical model. The global flux has already been successfully applied to different contexts and numerical methods [19, 20, 23, 35, 36, 32] to preserve one dimensional equilibria.

The idea of the global flux consists in rewriting the source term as a flux  $R$ :

$$R := \int^x s dx. \quad (5)$$

In this framework, equation (4) can be recast as

$$\partial_t q + \partial_x (f - R) = 0, \quad (6)$$

where discrete steady states satisfy the relation

$$\partial_x (f - R) = 0 \Leftrightarrow f - R \equiv \mathcal{F}_0, \quad (7)$$

with  $\mathcal{F} = f - R$  the so-called global flux and  $\mathcal{F}_0 = \mathcal{F}(x_0)$  for a given  $x_0$  in the domain. In the same spirit, a similar approach that integrates the Coriolis term into an apparent bathymetry term was also developed in [14].

The concept of well-balancing is a particular case of the preservation of general stationary solutions. The overarching idea is to design numerical schemes in which the artificial diffusion vanishes at relevant equilibria. The development of well-balanced schemes in one-dimensional problems has reached high levels of maturity in the last decades, but the multi-dimensional extensions are often tackled with trivial dimension-by-dimension approaches [19, 21, 38], which only allows the preservation of 1D-like equilibria. In [11], some of the authors presented a way of achieving multi-dimensional stationarity preservation through global flux in the context of a 2D linear problem, by considering that

$$\begin{aligned}\partial_t q + \partial_x f + \partial_y g &= \partial_t q + \partial_x \partial_y \left( \int^y f \, dy \right) + \partial_y g \\ &= \partial_t q + \partial_x f + \partial_y \partial_x \left( \int^x g \, dx \right) = 0.\end{aligned}\tag{8}$$

By combining the two derivatives, the two-dimensional formulation of the conservation law (1) can be written as

$$\partial_t q + \partial_x f + \partial_y g = \partial_t q + \partial_x \partial_y (F + G) = 0,\tag{9}$$

by defining

$$F := \int^y f \, dy, \quad G := \int^x g \, dx.\tag{10}$$

The new divergence operator  $\partial_x \partial_y (F + G)$  now is easy to preserve at the discrete level. Thanks to this formulation, it becomes also straightforward to consider multi-dimensional balance laws with source terms,

$$\partial_t q + \partial_x f + \partial_y g = s.\tag{11}$$

In this case, the source flux  $R$  can be defined as

$$R := \int^x \int^y s \, dx \, dy,\tag{12}$$

and directly included in the global flux. Setting

$$\mathcal{F} = F + G - R,\tag{13}$$

the multi-dimensional global flux form of the problem now becomes again

$$\partial_t q + \partial_{xy} \mathcal{F} = 0.\tag{14}$$

In this work, we present how this idea can be used to design first-order finite volume methods preserving multi-dimensional steady states, not known a priori, **for general nonlinear hyperbolic PDEs**. Moreover, a thorough analysis of the method is presented for linear problems showing a link with other stationarity preserving methods, as well as a discrete energy estimate. The approach proposed here naturally leads to the introduction of nonlinear **genuinely multi-dimensional fluxes at cell corners**, which have been shown to provide

fundamental enhancements in the numerical solutions, and enjoy many theoretical properties ([26, 10]). However, differently from previous works, the formulation proposed here naturally leads to corner fluxes, without any hypotheses on the type of quadrature. This approach is a starting point for the development of new families of stationarity preserving high-order methods based on high-degree polynomial reconstruction [22, 23], or discontinuous Galerkin methods [35, 46].

The paper is organized as follows. In section 2, we present the examples of PDEs that will be considered when assessing the performance of the method experimentally. In section 3, we recall the global flux method in a one-dimensional framework for hyperbolic balance laws. In section 4, we present the extension of the global flux method to two-dimensional nonlinear hyperbolic PDEs. Here, we discuss the finite volume formulation, the stabilization technique, boundary conditions, the treatment of source terms, as well as stability and consistency of the method for a linear model. In section 5, we present the standard finite volume method with piecewise constant and piecewise linear reconstructions used for comparison with the global flux method. Several numerical experiments are presented in section 6 to show the performance of the method. Finally, we draw some conclusions in section 7.

## 2. Mathematical models

The numerical method presented in this work is rather general and, in order to show its potential, we exemplify it on several mathematical models described by both linear and nonlinear hyperbolic systems. In particular, herein we will focus on three systems: linear acoustics, Euler equations for gas dynamics and the shallow water equations. For all of them, we focus on two-dimensional problems.

### 2.1. Linear acoustic system

The system of linear acoustic is a simple model that directly embeds non-trivial divergence-free steady states. It can be written in the following 2D and vectorial forms as:

$$\begin{cases} \partial_t u + \partial_x p &= 0, \\ \partial_t v + \partial_y p &= 0, \\ \partial_t p + \partial_x u + \partial_y v &= 0, \end{cases} \quad \begin{cases} \partial_t \mathbf{v} + \nabla p &= 0, \\ \partial_t p + \nabla \cdot \mathbf{v} &= 0, \end{cases} \quad (15)$$

where  $p$  is the pressure and  $\mathbf{v} = (u, v)$  is the velocity. The system can also be written in the compact form (1) with

$$q = \begin{bmatrix} u \\ v \\ p \end{bmatrix}, \quad f = \begin{bmatrix} p \\ 0 \\ u \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ p \\ v \end{bmatrix}. \quad (16)$$

The steady states of this system are given by

$$\partial_t q \equiv 0 \quad \Leftrightarrow \quad \nabla \cdot \mathbf{v} \equiv 0 \quad \text{and} \quad p \equiv p_0 = \text{const.} \quad (17)$$

## 2.2. Euler equations

The Euler equations are a simplification of the full Navier-Stokes system that do not include viscosity effects. Their use is widespread for the simulation of compressible gas dynamics. The system can be written in vectorial form as:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \mathbf{v}) &= 0, \\ \partial_t (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} + p \mathbf{I}) &= 0, \\ \partial_t (\rho E) + \nabla \cdot (\rho H \mathbf{v}) &= 0, \end{cases} \quad (18)$$

having denoted by  $\rho$  the density, by  $\mathbf{v}$  the velocity field, by  $E = e + \|\mathbf{v}\|^2/2$  the specific total energy, being  $e$  the specific internal energy and  $\mathbf{I}$  is the identity matrix. Finally, the total specific enthalpy is  $H = h + \|\mathbf{v}\|^2/2$ , with  $h = e + p/\rho$  the specific enthalpy. To close the system, we use the classical perfect gas equation of state  $p = (\gamma - 1)\rho e$  with  $\gamma$  the constant ratio of specific heats ( $\gamma = 1.4$  for air).

The nonlinear system of Euler equations can also be recast in the compact form (1) with

$$q = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{bmatrix}, \quad f = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho Hu \end{bmatrix}, \quad g = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho Hv \end{bmatrix}. \quad (19)$$

Steady states of the Euler equations are more complex but, after some manipulations, the smooth steady states can be characterized by the following relations:

$$\nabla \cdot (\rho \mathbf{v}) = 0, \quad (\rho \mathbf{v} \cdot \nabla) \mathbf{v} + \nabla p = 0, \quad \mathbf{v} \cdot \nabla H = 0. \quad (20)$$

## 2.3. Shallow water system

The Saint-Venant or shallow water equations describe the dynamics of hydrostatic free surface waves influenced by gravity. This model is valid under the hypothesis of very large wavelengths, or very shallow depths, and is applied in various engineering fields, including river and estuarine hydrodynamics, urban flood management, and tsunami risk assessment. In particular, when working with large scale problems, this simplified model becomes crucial to speed-up the computational time.

The system can be written in vectorial form as:

$$\begin{cases} \partial_t h + \nabla \cdot (h \mathbf{v}) &= 0, \\ \partial_t (h \mathbf{v}) + \nabla \cdot (h \mathbf{v} \otimes \mathbf{v} + \frac{1}{2} g h^2 \mathbf{I}) &= -g h \nabla b, \end{cases} \quad (21)$$

where  $h$  is the water height,  $\mathbf{v}$  the velocity field,  $b$  is the bathymetry and  $g$  is the gravity constant. The system can also be written in the classical compact notation (11) with

$$q = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \quad f = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2} g h^2 \\ huv \end{bmatrix}, \quad g = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2} g h^2 \end{bmatrix}, \quad s = \begin{bmatrix} 0 \\ -gh \partial_x b \\ -gh \partial_y b \end{bmatrix}. \quad (22)$$

This system admits a large variety of equilibria depending on the interaction between the flux and the source. The most studied equilibria in the context of well-balanced methods

are the so-called “lake at rest” states, which are characterized by a constant free surface level  $\eta := h + b \equiv \eta_0$  and a zero velocity  $\mathbf{v} \equiv 0$ . However, in the presence of a non-flat bathymetry, the system can also admit non-trivial equilibria, which are characterized by a non-zero velocity and a non-flat free surface level:

$$\nabla \cdot (h\mathbf{v}) = 0 \quad (\mathbf{v} \cdot \nabla)\mathbf{v} + g\nabla(h + b) = 0 \quad (23)$$

Several works have been devoted to the study of these equilibria in one dimension or in a quasi-1D framework [37, 35, 21]. In this work, we are interested in truly multi-dimensional well-balanced schemes that are capable of preserving all these equilibria at the discrete level.

### 3. Global flux for 1D balance laws

In this section, we recall the main principle of the global flux method and its initial usage in a 1D framework. Consider a general nonlinear balance law (4) and define a global flux  $\mathcal{F}$  as

$$\mathcal{F} = f - \int^x s dx, \quad (24)$$

such that (4) can now be written in a quasi-conservative form as

$$\partial_t q + \partial_x \mathcal{F} = 0. \quad (25)$$

Steady states given by  $\partial_t q = 0$  are equivalently characterized by the condition  $\partial_x \mathcal{F} = 0 \Leftrightarrow \mathcal{F} \equiv \mathcal{F}_0 \in \mathbb{R}$ .

In a finite volume framework, the computational domain  $\Omega$  is split into  $N$  cells and the equation (4) is integrated over each cell  $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ :

$$\frac{d}{dt} \bar{q}_i + \frac{\widehat{\mathcal{F}}_{i+\frac{1}{2}} - \widehat{\mathcal{F}}_{i-\frac{1}{2}}}{\Delta x} = 0, \quad (26)$$

where the cell average is defined as

$$\bar{q}_i := \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q dx. \quad (27)$$

The numerical global flux  $\widehat{\mathcal{F}}_{i+\frac{1}{2}}$  is considered to be a function of the two values of the global flux  $\mathcal{F}_{i+\frac{1}{2}}^L$  and  $\mathcal{F}_{i+\frac{1}{2}}^R$  reconstructed at both sides of interface  $x_{i+\frac{1}{2}}$ . For piecewise constant reconstructions of the global flux one simply has  $\mathcal{F}_{i+\frac{1}{2}}^L = \mathcal{F}_i$  and  $\mathcal{F}_{i+\frac{1}{2}}^R = \mathcal{F}_{i+1}$ .

**Remark 1** (Numerical global flux). *It is important to underline that structure preservation can only be achieved if the interface global flux  $\widehat{\mathcal{F}}_{i+\frac{1}{2}}$  depends only on global fluxes  $\{\mathcal{F}_j\}_{j \in \mathbb{Z}}$  in the cells, and not on the values  $\{q_j\}_{j \in \mathbb{Z}}$  of the conservative variables. This is due to the fact that, at equilibria, only global fluxes are constant while conservative variables may vary.*

In our previous work [23], we employed the following upwind flux:

$$\widehat{\mathcal{F}}_{i+\frac{1}{2}} = \mathbb{1}^+ \mathcal{F}_{i+\frac{1}{2}}^L + \mathbb{1}^- \mathcal{F}_{i+\frac{1}{2}}^R \quad \text{where} \quad \mathbb{1}^\pm := L^{-1} \frac{\mathbb{1} \pm \text{sign } \Lambda}{2} L, \quad (28)$$

where  $L$  is the matrix of left eigenvectors of the flux Jacobian  $\partial_q f$ , and  $\text{sign } \Lambda$  is the diagonal matrix of the sign of the eigenvalues of the flux Jacobian, evaluated using any (average) state at the interface<sup>1</sup>.

For the development of the global flux method, a consistent approximation of  $R := \int^x s \, dx$  is necessary to define  $\mathcal{F} = f - R$ . This integral can be computed in a recursive manner, starting from the beginning of the domain, by integrating the source in each element. To simplify the description of the method, we will assume that  $q$  and  $s$  are constant in each cell, therefore the source integral can be computed as

$$R_i := \underbrace{\int_{R_{i-1}}^{x_{i-1}} s \, dx}_{R_{i-1}} + \int_{x_{i-1}}^{x_{i-\frac{1}{2}}} s \, dx + \int_{x_{i-\frac{1}{2}}}^{x_i} s \, dx = R_{i-1} + \frac{\Delta x}{2} \bar{s}_{i-1} + \frac{\Delta x}{2} \bar{s}_i. \quad (29)$$

Hence, the global flux will now depend on both the conservative flux and the source term

$$\mathcal{F}_i = f(\bar{q}_i) - R_i = f(\bar{q}_i) - R_{i-1} - \frac{\Delta x}{2} (\bar{s}_{i-1} + \bar{s}_i). \quad (30)$$

Similarly, the recursive procedure gives us the following values for  $\mathcal{F}_{i-1}$  and  $\mathcal{F}_{i+1}$ :

$$\mathcal{F}_{i-1} = f(\bar{q}_{i-1}) - R_{i-1}, \quad (31)$$

$$\mathcal{F}_{i+1} = f(\bar{q}_{i+1}) - R_{i-1} - \Delta x \left( \frac{1}{2} \bar{s}_{i-1} + \bar{s}_i + \frac{1}{2} \bar{s}_{i+1} \right). \quad (32)$$

It can be noticed that, when considering a simple numerical flux

$$\widehat{\mathcal{F}}_{i+\frac{1}{2}}(\mathcal{F}_i, \mathcal{F}_{i+1}) = \mathcal{F}_i \quad (33)$$

equation (26) can be recast as

$$\frac{d}{dt} \bar{q}_i = - \frac{\mathcal{F}_i - \mathcal{F}_{i-1}}{\Delta x} = - \frac{f(\bar{q}_i) - f(\bar{q}_{i-1})}{\Delta x} + \frac{\bar{s}_i + \bar{s}_{i-1}}{2}, \quad (34)$$

which shows already a difference with respect to the classical finite volume method, where the source term would be treated in a centered way.

When the upwind numerical flux (28) is used, one has (having temporarily made the dependence of  $\mathbb{1}^\pm$  on the interface explicit)

$$\frac{d}{dt} \bar{q}_i = - \frac{\mathcal{F}_i - \mathcal{F}_{i-1}}{\Delta x} \quad (35)$$

$$\begin{aligned} &= - \frac{\mathbb{1}_{i+\frac{1}{2}}^- f(\bar{q}_{i+1}) + (\mathbb{1}_{i+\frac{1}{2}}^+ + \mathbb{1}_{i-\frac{1}{2}}^-) f(\bar{q}_i) - \mathbb{1}_{i-\frac{1}{2}}^+ f(\bar{q}_{i-1})}{\Delta x} \\ &\quad + (\mathbb{1}_{i+\frac{1}{2}}^+ + \mathbb{1}_{i+\frac{1}{2}}^- - \mathbb{1}_{i-\frac{1}{2}}^-) \frac{\bar{s}_i + \bar{s}_{i-1}}{2} + \mathbb{1}_{i+\frac{1}{2}}^- \frac{\bar{s}_i + \bar{s}_{i-1}}{2} + \underbrace{(\mathbb{1}_{i+\frac{1}{2}}^+ + \mathbb{1}_{i+\frac{1}{2}}^-)}_{=1} - \underbrace{(\mathbb{1}_{i-\frac{1}{2}}^+ + \mathbb{1}_{i-\frac{1}{2}}^-)}_{=1} R_{i-1}, \end{aligned} \quad (36)$$

---

<sup>1</sup>In principle, therefore, one should write  $\mathbb{1}_{i+\frac{1}{2}}^\pm$  to make clear that they differ from interface to interface; we do not make this depends explicit to ensure readability

where the contribution from  $R_{i-1}$  cancels out even if  $\mathbb{1}^\pm$  depend on the interface, since they nevertheless add up to  $\mathbb{1}$  on each of them.

**Remark 2** (Compactness of global fluxes). *It can be noticed that, although the global flux in (30) is defined globally with  $R_{i-1}$  that depends on previous values, the time residual (34) shows that the stencil is actually compact due to the cancellation of these terms. A full analogy with compact residual distribution methods on a dual cell is presented later in section 4.5.*

## 4. Global flux for multi-dimensional hyperbolic PDEs

### 4.1. Numerical method

When dealing with multi-dimensional conservation laws, non-trivial equilibria arise also in absence of a source term in the equation. For steady states  $\partial_t q = 0$ , it is no longer just  $\partial_x f = 0$  that follows, but instead the divergence  $\partial_x f + \partial_y g = 0$ , which in general might have many solutions.

In this section, we will show how to extend the global flux method to multi-dimensional conservation and balance laws. We start by rewriting (1) as (9) using the definitions in (10) to obtain

$$\partial_t q + \partial_{xy} \mathcal{F} = 0. \quad (37)$$

with the global flux  $\mathcal{F} := F + G$ . Then, integration of (14) over the cell  $C_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$  yields

$$\Delta x \Delta y \frac{d}{dt} \bar{q}_{i,j} + \mathcal{F}(t, x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}) - \mathcal{F}(t, x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}}) - \mathcal{F}(t, x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}}) + \mathcal{F}(t, x_{i-\frac{1}{2}}, y_{j-\frac{1}{2}}) = 0, \quad (38)$$

where the cell average is defined as

$$\bar{q}_{i,j} := \frac{1}{\Delta x \Delta y} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} q \, dx dy.$$

We therefore are led to introduce the numerical **corner fluxes**  $\widehat{\mathcal{F}}_{i\pm\frac{1}{2},j\pm\frac{1}{2}}$  that then allow to write the evolution equation as

$$\frac{d}{dt} \bar{q}_{i,j} + \frac{\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} - \widehat{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}} - \widehat{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}} + \widehat{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}}}{\Delta x \Delta y} = 0. \quad (39)$$

Recall that the global flux  $\mathcal{F}$  is obtained by integrating the physical fluxes and the source, see equations (10), (12), and (13). In practice, the integrated fluxes  $F$  and  $G$  are computed by performing integrals along, respectively, the  $y$  and  $x$  directions in a 1D fashion. In particular, the value of  $F_i$  in the barycenter of a given cell  $i$  can be computed recursively starting from the beginning of the domain, similarly to section 3, as

$$F_{i,j} = \int_{y_{j-1}}^{y_{j-1}} f \, dy + \int_{y_{j-1}}^{y_j} f \, dy = F_{i,j-1} + \frac{\Delta y}{2} (f_{i,j-1} + f_{i,j}), \quad \forall i, \quad (40)$$

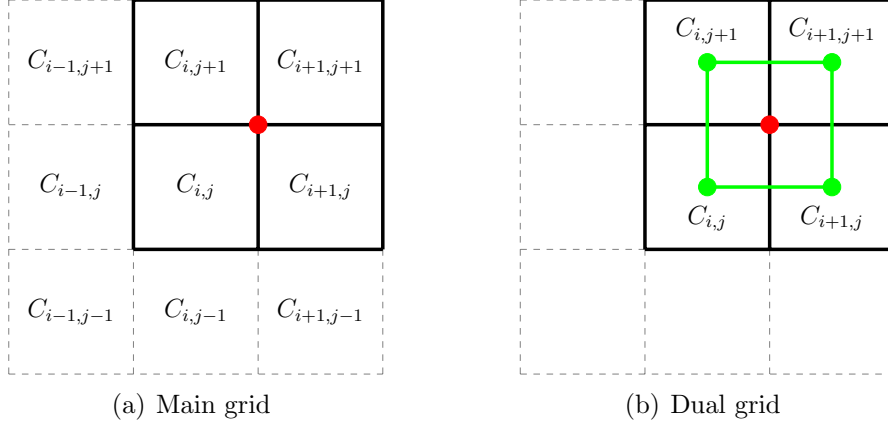


Figure 1: Cell labeling for the 2D grid.

where, for a first order method, trapezoidal rule is accurate enough. Similarly,  $G_i$  can be computed as

$$G_{i,j} = \int_{x_{i-1}}^{x_i} g \, dx + \int_{x_{i-1}}^{x_i} g \, dx = G_{i-1,j} + \frac{\Delta x}{2}(g_{i-1,j} + g_{i,j}), \quad \forall j. \quad (41)$$

When dealing with hyperbolic PDEs with source terms, as for the shallow water equations in section 2.3, the integral of the source term  $R := \int^x \int^y s \, dx dy$  can also be embedded into the global flux. Similarly,  $R$  can be recursively defined as

$$R_{i,j} = \int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} s \, dy dx + \int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} s \, dy dx = R_{i-1,j-1} + \frac{\Delta x \Delta y}{4} (s_{i-1,j-1} + s_{i-1,j} + s_{i,j-1} + s_{i,j}), \quad (42)$$

for every  $i, j$ . More details about the treatment of source terms for the shallow water equations 2.3, are given in section 4.6.

#### 4.2. Numerical corner fluxes

The conservative formulation obtained using global fluxes (39) requires the definition of numerical corner fluxes to update cell averages. This definition is better achieved by considering the evolution of all the cells neighbouring a given node. To this end it is preferable to recast (39) as

$$\frac{d}{dt} \bar{q}_{i,j} + \frac{\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)}}{\Delta x \Delta y} = 0. \quad (43)$$

where we have added the superscript  $(i,j)$  to account for the fact that the signs used in the formula involve the flux balance for  $\bar{q}_{i,j}$ . They thus are interpreted as depending on the orientation of the corner normal for the four cells  $(i+\ell, j+r)$  for  $\ell, r \in \{0, 1\}$  with respect to the corner  $(i+\frac{1}{2}, j+\frac{1}{2})$ , defined by

$$\mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)} := \begin{pmatrix} (-1)^{\ell+1} \\ (-1)^{r+1} \end{pmatrix}, \text{ i.e.,} \quad (44)$$

$$\mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} := \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} := \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} := \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} := \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (45)$$

Thus  $\mathbf{n}_p^c$  is a normal at corner  $p$  pointing into cell  $c$ . Corner normals and corner fluxes alongside a modified concept of conservation associated to corners rather than edges are used e.g. in [10] for general polygonal grids, where they are shown to be crucial for structure preservation. There, corner normals are defined as the average of the two edge-normals adjacent to the node, scaled with the respective edge lengths. We believe that the definition above is sufficient in the context of Cartesian meshes.

We further define the scalar  $n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)}$  as

$$n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)} = n(\mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)}) = (-1)^{\ell+1}(-1)^{r+1} = (-1)^{\ell+r}, \quad (46)$$

where  $n(\mathbf{n}) := \mathbf{n}_x \mathbf{n}_y$  is the product of the two components.

The reinterpreted definition of the corner fluxes above requires a different setting compared to the classical one, and appears in the context of genuinely multi-dimensional Riemann solvers using more than two states as input (see e.g. [6, 25, 26] and references therein).

Next, we aim at defining the notion of the numerical global flux in the multi-dimensional context as generally as possible. Then, we elucidate the conditions imposed on the functional form of the numerical flux by consistency, conservation, and preservation of steady states. We define the numerical corner fluxes at the corner  $(i + \frac{1}{2}, j + \frac{1}{2})$  with respect to the four cells  $(i + \ell, j + r)$  with  $\ell, r \in \{0, 1\}$  as

$$\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)} = \widehat{\mathcal{F}}(\mathcal{F}_{i,j}, \mathcal{F}_{i,j+1}, \mathcal{F}_{i+1,j}, \mathcal{F}_{i+1,j+1}; \bar{q}_{i,j}, \bar{q}_{i,j+1}, \bar{q}_{i+1,j}, \bar{q}_{i+1,j+1} | \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)}).$$

We can now formulate local consistency as

$$\widehat{\mathcal{F}}(\mathcal{F}, \mathcal{F}, \mathcal{F}, \mathcal{F}; q, q, q, q | \mathbf{n}) = \mathcal{F} n(\mathbf{n}). \quad (47)$$

A stronger property is the steady state preservation requirement which can be expressed as

$$\widehat{\mathcal{F}}(\mathcal{F}, \mathcal{F}, \mathcal{F}, \mathcal{F}; \bar{q}_{i,j}, \bar{q}_{i,j+1}, \bar{q}_{i+1,j}, \bar{q}_{i+1,j+1} | \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}) = \mathcal{F} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}. \quad (48)$$

**Remark 3** (Steady state subspace). *Following [11], it will be shown below that steady state preservation may be actually already proven if*

$$\mathcal{F}_{i,j}^* = \mathcal{F} + \alpha_i + \beta_j \quad (49)$$

for any two data distributions  $\alpha$  and  $\beta$ , such that  $\alpha$  only depends on  $i$  and  $\beta$  only on  $j$ . In this case, a conservation property more general of (48) reads

$$\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}(\mathcal{F}_{i,j}^*, \mathcal{F}_{i,j+1}^*, \mathcal{F}_{i+1,j}^*, \mathcal{F}_{i+1,j+1}^*; \bar{q}_{i,j}, \bar{q}_{i,j+1}, \bar{q}_{i+1,j}, \bar{q}_{i+1,j+1} | \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}) = \mathcal{F} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}. \quad (50)$$

*This condition will be used also in the consistency analysis of Section 4.5.*

Conservation cannot be expressed by face in this framework, as in standard finite volume methods. It is instead formulated at corners as follows:

$$\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} + \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} + \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} = 0. \quad (51)$$

Having established the necessary conditions that the numerical global fluxes have to satisfy, we next propose a particular choice. As in [26], we define numerical global fluxes as the sum of a consistent central flux plus a diffusion term  $\mathcal{D}$ . In the present paper, corner fluxes are obtained extending to quadrilaterals the multi-dimensional Osher-Solomon flux proposed in [26], and combining it with the recent work [11] on global flux compatible SUPG stabilization. We define the numerical corner flux around the corner  $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$ , which involves the cells  $C_{i,j}$ ,  $C_{i+1,j}$ ,  $C_{i,j+1}$  and  $C_{i+1,j+1}$  (see figure 1). The same principle is applied to the other corners. We set

$$\begin{aligned}\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} &= \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)}, & \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} &= \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)}, \\ \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} &= \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)}, & \widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} &= \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)},\end{aligned}\tag{52}$$

where

$$\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{4}(\mathcal{F}_{i,j} + \mathcal{F}_{i+1,j} + \mathcal{F}_{i,j+1} + \mathcal{F}_{i+1,j+1})$$

is the average of the global fluxes at the corner.

To define the numerical dissipation we consider corner dual cells, as depicted on the right on figure 1. The conservation condition (51) requires that

$$\mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} + \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} = 0.$$

To define the corner dissipation, we cannot proceed as in [26], since this would break the stationarity preserving property. Instead, we take inspiration from the streamline upwind stabilization (SUPG), studied in the global flux context in [11]. To this end, on the dual cell  $\tilde{C}_{i+\frac{1}{2},j+\frac{1}{2}}$  we compute SUPG stabilizing terms

$$\begin{aligned}\mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)} &:= \mathcal{D}(\widetilde{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}, \bar{q}_{i+\frac{1}{2},j+\frac{1}{2}} | \mathbf{n}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+r)}) \\ &= \alpha \Delta \int_{\tilde{C}_{i+\frac{1}{2},j+\frac{1}{2}}} \left( \frac{1}{\Delta x} J^x \partial_\xi \varphi_{\ell,r} + \frac{1}{\Delta y} J^y \partial_\eta \varphi_{\ell,r} \right) \partial_{\xi\eta} \widetilde{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} d\xi d\eta,\end{aligned}\tag{53}$$

where  $\widetilde{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}$  is a bi-linear  $\mathbb{Q}^1$  reconstruction of the global flux on the dual cell from the four adjacent values  $\mathcal{F}_{i,j}$ ,  $\mathcal{F}_{i+1,j}$ ,  $\mathcal{F}_{i,j+1}$ ,  $\mathcal{F}_{i+1,j+1}$ . Here,  $J^x$  and  $J^y$  are the Jacobians of the fluxes  $f$  and  $g$  computed in the average value  $\bar{q}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{q_{i,j} + q_{i+1,j} + q_{i,j+1} + q_{i+1,j+1}}{4}$ .  $\Delta = \frac{\sqrt{\Delta x^2 + \Delta y^2}}{\sqrt{2}}$  is the characteristic mesh size, and  $\alpha = 1/\lambda_m$  with  $\lambda_m$  the maximal spectral radius of the flux Jacobians computed with the average state of the four reconstructed values at the corner. The  $\varphi_{\ell,r}$  for  $\ell, r \in \{0, 1\}$  in the above definition are the standard bi-linear finite element basis functions on the quadrilateral  $\tilde{C}$  defined by

$$\varphi_{\ell,r}(\xi, \eta) = \frac{1}{4}(1 + (-1)^{\ell+1}\xi)(1 + (-1)^{r+1}\eta),\tag{54}$$

i.e.,

$$\begin{aligned}\varphi_{0,0}(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 - \eta), & \varphi_{1,0}(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 - \eta) \\ \varphi_{0,1}(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 + \eta), & \varphi_{1,1}(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 + \eta)\end{aligned}\tag{55}$$

on the reference element  $\xi, \eta \in [-1, 1]$ . With this, we can explicitly evaluate the streamline upwind dissipation terms as

$$\mathcal{D}(\widetilde{\mathcal{F}}, \bar{q}|\mathbf{n}) = \frac{\alpha\Delta}{4} \left( \frac{\mathbf{n}_x}{\Delta x} J^x + \frac{\mathbf{n}_y}{\Delta y} J^y \right) \Phi(\widetilde{\mathcal{F}}), \quad (56)$$

i.e.,

$$\begin{aligned} \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} &= \frac{\alpha\Delta}{4} \left( -\frac{J^x}{\Delta x} + \frac{J^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j+\frac{1}{2}}, & \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} &= \frac{\alpha\Delta}{4} \left( +\frac{J^x}{\Delta x} + \frac{J^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j+\frac{1}{2}}, \\ \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} &= \frac{\alpha\Delta}{4} \left( -\frac{J^x}{\Delta x} - \frac{J^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j+\frac{1}{2}}, & \mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} &= \frac{\alpha\Delta}{4} \left( +\frac{J^x}{\Delta x} - \frac{J^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j+\frac{1}{2}}, \end{aligned} \quad (57)$$

with

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}} := \Phi(\widetilde{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}) := \mathcal{F}_{i+1,j+1} - \mathcal{F}_{i,j+1} - \mathcal{F}_{i+1,j} + \mathcal{F}_{i,j} = \int_{\tilde{C}_{i+\frac{1}{2},j+\frac{1}{2}}} \partial_{xy} \widetilde{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} dx dy. \quad (58)$$

The next sections are devoted to the analysis of some properties of the scheme obtained with the above definitions, as well as some enhancements.

#### 4.3. Compactness of the method

Taking into account only the central flux, without the diffusion, one obtains

$$\frac{d}{dt} \bar{q}_{i,j} = - \frac{\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \widehat{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)}}{\Delta x \Delta y} \quad (59a)$$

$$= - \frac{\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}} n_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}} n_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}} n_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)}}{\Delta x \Delta y} \quad (59b)$$

$$= - \frac{\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} - \overline{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}} - \overline{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}} + \overline{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}}}{\Delta x \Delta y}. \quad (59c)$$

Define

$$\overline{F}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{4} (F_{i,j} + F_{i+1,j} + F_{i,j+1} + F_{i+1,j+1}) \quad (60)$$

$$\overline{G}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{4} (G_{i,j} + G_{i+1,j} + G_{i,j+1} + G_{i+1,j+1}) \quad (61)$$

such that  $\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} = \overline{F}_{i+\frac{1}{2},j+\frac{1}{2}} + \overline{G}_{i+\frac{1}{2},j+\frac{1}{2}}$ , and use the recursions (40)–(41) to obtain

$$\overline{F}_{i+\frac{1}{2},j+\frac{1}{2}} - \overline{F}_{i+\frac{1}{2},j-\frac{1}{2}} = \frac{\Delta y}{8} (f_{i+1,j+1} + 2f_{i+1,j} + f_{i+1,j-1}) + \frac{\Delta y}{8} (f_{i,j+1} + 2f_{i,j} + f_{i,j-1}). \quad (62)$$

We find an analogous formula for  $\bar{F}_{i-\frac{1}{2},j+\frac{1}{2}} - \bar{F}_{i-\frac{1}{2},j-\frac{1}{2}}$ , such that in the end

$$\frac{d}{dt}\bar{q}_{i,j} = -\frac{1}{2\Delta x}\langle f_{i+1,\cdot} - f_{i-1,\cdot} \rangle_j - \frac{1}{2\Delta y}\langle g_{\cdot,j+1} - g_{\cdot,j-1} \rangle_i \quad (63)$$

with the average

$$\langle a_{i,\cdot} \rangle_j := \frac{1}{4}(a_{i,j+1} + 2a_{i,j} + a_{i,j-1}). \quad (64)$$

One observes that all the global fluxes drop out and the central part of the method is local.

Next, we turn to the numerical stabilization. Defining

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^F = F_{i+1,j+1} - F_{i,j+1} - F_{i+1,j} + F_{i,j} \quad (65)$$

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^G = G_{i+1,j+1} - G_{i,j+1} - G_{i+1,j} + G_{i,j} \quad (66)$$

such that  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}} = \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^F + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^G$ , and using again the recursions (40)–(41), one obtains

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^F = \frac{\Delta y}{2}(f_{i+1,j+1} + f_{i+1,j} - f_{i,j+1} - f_{i,j}). \quad (67)$$

An analogous formula is valid for  $\Phi_{i-\frac{1}{2},j+\frac{1}{2}}^F$ , and also

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^G = \frac{\Delta x}{2}(g_{i+1,j+1} + g_{i,j+1} - g_{i+1,j} - g_{i,j}). \quad (68)$$

One observes again that all the global fluxes drop out.

The Jacobians involved in the update of  $q_{i,j}$  are evaluated at the four corners, such that the method with only the numerical stabilization reads

$$\begin{aligned} \frac{d}{dt}\bar{q}_{i,j} &= -\frac{\mathcal{D}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \mathcal{D}_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \mathcal{D}_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \mathcal{D}_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)}}{\Delta x \Delta y} \\ &= -\frac{1}{\Delta x \Delta y} \frac{\alpha \Delta}{4} \left[ \left( -\frac{J_{i+\frac{1}{2},j+\frac{1}{2}}^x}{\Delta x} - \frac{J_{i+\frac{1}{2},j+\frac{1}{2}}^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j+\frac{1}{2}} + \left( \frac{J_{i-\frac{1}{2},j+\frac{1}{2}}^x}{\Delta x} - \frac{J_{i-\frac{1}{2},j+\frac{1}{2}}^y}{\Delta y} \right) \Phi_{i-\frac{1}{2},j+\frac{1}{2}} \right. \\ &\quad \left. + \left( -\frac{J_{i+\frac{1}{2},j-\frac{1}{2}}^x}{\Delta x} + \frac{J_{i+\frac{1}{2},j-\frac{1}{2}}^y}{\Delta y} \right) \Phi_{i+\frac{1}{2},j-\frac{1}{2}} + \left( \frac{J_{i-\frac{1}{2},j-\frac{1}{2}}^x}{\Delta x} + \frac{J_{i-\frac{1}{2},j-\frac{1}{2}}^y}{\Delta y} \right) \Phi_{i-\frac{1}{2},j-\frac{1}{2}} \right]. \end{aligned} \quad (69)$$

To give the spirit of the method, assume for the moment, however, that the Jacobians

are evaluated on same state. Then, the numerical stabilization becomes

$$\frac{d}{dt}\bar{q}_{i,j} = -\frac{1}{\Delta x \Delta y} \frac{\alpha \Delta}{4} \left[ \frac{J^x}{\Delta x} \left( -\Phi_{i+\frac{1}{2},j+\frac{1}{2}} + \Phi_{i-\frac{1}{2},j+\frac{1}{2}} - \Phi_{i+\frac{1}{2},j-\frac{1}{2}} + \Phi_{i-\frac{1}{2},j-\frac{1}{2}} \right) \right. \quad (70)$$

$$\left. + \frac{J^y}{\Delta y} \left( -\Phi_{i+\frac{1}{2},j+\frac{1}{2}} - \Phi_{i-\frac{1}{2},j+\frac{1}{2}} + \Phi_{i+\frac{1}{2},j-\frac{1}{2}} + \Phi_{i-\frac{1}{2},j-\frac{1}{2}} \right) \right] \\ = \frac{\alpha \Delta}{2} \left[ \frac{J^x}{\Delta x^2} \langle f_{i+1,\cdot} - 2f_{i,\cdot} + f_{i-1,\cdot} \rangle_j + \frac{J^x}{4\Delta x \Delta y} (g_{i+1,j+1} - g_{i-1,j+1} - g_{i+1,j-1} + g_{i-1,j-1}) \right. \\ \left. + \frac{J^y}{4\Delta x \Delta y} (f_{i+1,j+1} - f_{i-1,j+1} - f_{i+1,j-1} + f_{i-1,j-1}) + \frac{J^y}{\Delta y^2} \langle g_{\cdot,j+1} - 2g_{\cdot,j} + g_{\cdot,j-1} \rangle_i \right]. \quad (71)$$

Here, the average  $\langle \cdot \rangle$  introduced in (64) has been used again.

Finally, the method can be expressed in classical flux form

$$\frac{d}{dt}q_{i,j} + \frac{\hat{f}_{i+\frac{1}{2},j} - \hat{f}_{i-\frac{1}{2},j}}{\Delta x} + \frac{\hat{g}_{i,j+\frac{1}{2}} - \hat{g}_{i,j-\frac{1}{2}}}{\Delta y} = 0. \quad (72)$$

This demonstrates that additionally to the notion (51), the method is also conservative in the classical sense. The numerical flux through the edge  $(i + \frac{1}{2}, j)$  reads

$$\hat{f}_{i+\frac{1}{2},j} = \frac{1}{2} \langle f_{i+1,\cdot} + f_{i,\cdot} \rangle_j - \frac{\alpha \Delta}{2} \frac{J_{i+\frac{1}{2},j+\frac{1}{2}}^x \Phi_{i+\frac{1}{2},j+\frac{1}{2}} + J_{i+\frac{1}{2},j-\frac{1}{2}}^x \Phi_{i+\frac{1}{2},j-\frac{1}{2}}}{2\Delta x \Delta y}. \quad (73)$$

In a quasi-1D situation, i.e. when nothing depends on  $j$  and when  $g = 0$ , the flux is

$$\hat{f}_{i+\frac{1}{2}} = \frac{1}{2} (f_{i+1} + f_i) - \frac{\alpha}{2} J_{i+\frac{1}{2}}^x (f_{i+1} - f_i). \quad (74)$$

With this, next we discuss the interplay between the numerical stabilization and stationarity preservation.

#### 4.4. Analysis of the method for the linear acoustic system

In this section, we focus on the analysis of the new numerical method by analysing the numerical diffusion that allows to achieve stationarity preservation, and obtaining an energy estimate.

##### 4.4.1. Numerical diffusion and stationarity preservation

We start by considering the linear acoustic system, but similar results can be shown for nonlinear problems. A classical dimensionally split finite volume scheme with a local Lax-Friedrichs numerical flux provides the following discretization of the linear acoustic system:

$$\begin{aligned} \frac{d}{dt}p + D_x u + D_y v &= \frac{\lambda_m \Delta x}{2} D_{xx} p + \frac{\lambda_m \Delta y}{2} D_{yy} p, \\ \frac{d}{dt}u + D_x p &= \frac{\lambda_m \Delta x}{2} D_{xx} u, \\ \frac{d}{dt}v + D_y p &= \frac{\lambda_m \Delta y}{2} D_{yy} v, \end{aligned} \quad (75)$$

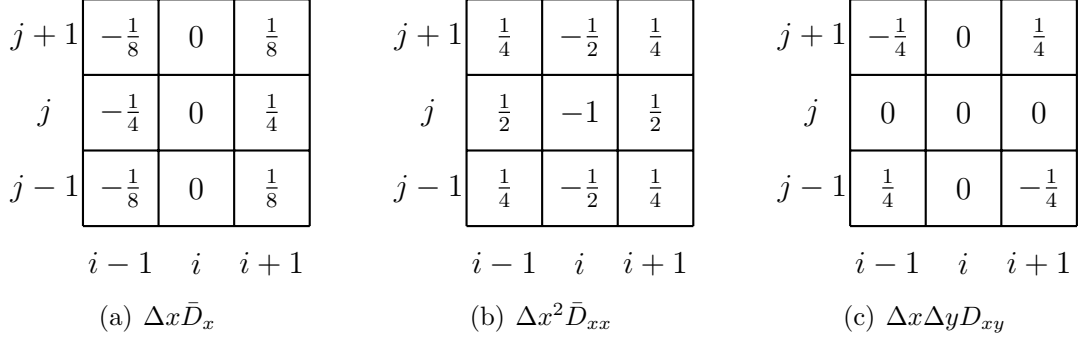


Figure 2: Finite difference-like stencils for global flux differential operators.

where  $D$  represent the discrete derivative operators given by

$$(D_x q)_{i,j} = \frac{q_{i+1,j} - q_{i-1,j}}{2\Delta x}, \quad (D_{xx} q)_{i,j} = \frac{q_{i+1,j} - 2q_{i,j} + q_{i-1,j}}{\Delta x^2}, \quad (76)$$

and similarly for  $D_y$  and  $D_{yy}$ . As can be noticed, the stencil used in this discretization is a simple 5-points stencil.

Contrary to this, the stencils involved in the new first order global flux method, equipped with SUPG corner fluxes as described above, includes the cell itself and its eight neighbors (9-points stencil):

$$\begin{aligned} \frac{d}{dt} p + \bar{D}_x u + \bar{D}_y v &= \frac{\alpha \Delta}{2} (\bar{D}_{xx} p + \bar{D}_{yy} p), \\ \frac{d}{dt} u + \bar{D}_x p &= \frac{\alpha \Delta}{2} (\bar{D}_{xx} u + D_{xy} v), \\ \frac{d}{dt} v + \bar{D}_y p &= \frac{\alpha \Delta}{2} (D_{xy} u + \bar{D}_{yy} v), \end{aligned} \quad (77)$$

where the new finite difference operators are given by the stencils in figure 2.  $\bar{D}_x$  and  $\bar{D}_{xx}$  are standard discrete first and second order derivatives in  $x$ , but including a particular averaging in  $y$  direction, introduced in (64) as  $\langle a_{i,\cdot} \rangle_j := \frac{1}{4}(a_{i,j+1} + 2a_{i,j} + a_{i,j-1})$ . These operators have first appeared in [40] and then in virtually all subsequent works on stationarity and vorticity preservation for linear acoustics on Cartesian grids, e.g. in [40, 45, 28, 39, 34, 8].

In [11] these finite difference operators appeared naturally as Kronecker products of uni-direction operators:  $D_{xy} = D_x \otimes D_y I_y$ ,  $\bar{D}_{xx} = D_{xx} \otimes D_y I_y$  etc., with matrices  $I_y$  responsible for the integration and  $D_y I_y$  being the particular averaging matrix corresponding to  $\langle \cdot \rangle$ .

One observes that the diffusion operators for the velocity no longer depend on second derivatives of individual components, but instead on the gradient of the divergence operator. Although this characteristic of the scheme is more readily visible for a simplified model like the linear acoustic system (77), similar considerations can be drawn also for more complex nonlinear systems, as is obvious from (71).

**Remark 4** (Numerical diffusion for nonlinear problems). *For the shallow water equations, the first order global flux method with SUPG corner fluxes leads to the following discrete*

evolution equations:

$$\begin{aligned}
\frac{d}{dt}h + \bar{D}_x f_h + \bar{D}_y g_h &= \frac{\alpha\Delta}{2}(\bar{D}_{xx}f_{hu} + D_{xy}g_{hu}) + \frac{\alpha\Delta}{2}(D_{xy}f_{hv} + \bar{D}_{yy}g_{hv}), \\
\frac{d}{dt}hu + \bar{D}_x f_{hu} + \bar{D}_y g_{hu} &= \frac{\alpha\Delta}{2}(g\bar{h} - \bar{u}^2)(\bar{D}_{xx}f_h + D_{xy}g_h) - \frac{\alpha\Delta}{2}\bar{u}\bar{v}(D_{xy}f_h + \bar{D}_{yy}g_h) + \\
&\quad + \alpha\Delta\bar{u}(\bar{D}_{xx}f_{hu} + D_{xy}g_{hu}) + \frac{\alpha\Delta}{2}\bar{v}(D_{xy}f_{hu} + \bar{D}_{yy}g_{hu}) + \frac{\alpha\Delta}{2}\bar{u}(D_{xy}f_{hv} + \bar{D}_{yy}g_{hv}), \quad (78) \\
\frac{d}{dt}hv + \bar{D}_x f_{hv} + \bar{D}_y g_{hv} &= \frac{\alpha\Delta}{2}(g\bar{h} - \bar{v}^2)(D_{xy}f_h + \bar{D}_{yy}g_h) - \frac{\alpha\Delta}{2}\bar{u}\bar{v}(\bar{D}_{xx}f_h + D_{xy}g_h) + \\
&\quad + \alpha\Delta\bar{v}(D_{xy}f_{hv} + \bar{D}_{yy}g_{hv}) + \frac{\alpha\Delta}{2}\bar{v}(\bar{D}_{xx}f_{hu} + D_{xy}g_{hu}) + \frac{\alpha\Delta}{2}\bar{u}(\bar{D}_{xx}f_{hv} + D_{xy}g_{hv}),
\end{aligned}$$

where we considered constant Jacobians defined in an average state  $\bar{q} = (\bar{h}, \bar{h}\bar{u}, \bar{h}\bar{v})$  to regroup the terms in a compact form. For notational convenience, we have introduced the terms  $f_{hu}$  and  $g_{hu}$  to denote the fluxes for the momentum equation in  $hu$ , and similarly for the other equations. Again, we obtain diffusion terms that depends on the gradient of the divergence operator, which is essential for stationarity preservation.

#### 4.4.2. Semi-discrete energy stability

In this section, we focus on the semi-discrete energy estimates for the linear acoustic system. In particular, in the continuous setting, it can be easily proven that the conserved energy of the system is  $\mathcal{E} = \frac{u^2+v^2}{2} + \frac{p^2}{2}$ , by multiplying (15) by  $q^T$ , summing the three equations and integrating over the whole domain  $\Omega$ :

$$\begin{aligned}
\int_{\Omega} [q^T \partial_t q + q^T \mathbf{J} \nabla q] d\mathbf{x} &= \int_{\Omega} [u \partial_t u + u \partial_x p + v \partial_t v + v \partial_y p + p \partial_t p + p \partial_x u + p \partial_y v] d\mathbf{x} \\
&= \frac{d}{dt} \int_{\Omega} \left[ \frac{u^2 + v^2}{2} + \frac{p^2}{2} \right] d\mathbf{x} + \int_{\partial\Omega} p \mathbf{v} \cdot \mathbf{n} dS
\end{aligned} \quad (79)$$

where the second term is zero for periodic boundary conditions. In our discrete framework, we would like to prove that

$$\frac{d}{dt} \int_{\Omega} \left[ \frac{u^2 + v^2}{2} + \frac{p^2}{2} \right] dx \leq 0.$$

To do that, we will write the new differential operators introduced above in a tensor product form to split the contributions from the two dimensions, for more details see [11],

$$\begin{aligned}
\bar{D}_x &= (D_+ M_-) \otimes (M_+ M_-), & \bar{D}_{xx} &= (D_+ D_-) \otimes (M_+ M_-), \\
\bar{D}_{yy} &= (M_+ M_-) \otimes (D_+ D_-), & D_{xy} &= (D_+ M_-) \otimes (D_+ M_-)
\end{aligned}$$

where the derivative,  $D$ , and average,  $M$ , operators are defined as

$$D_+ = \begin{bmatrix} -1 & 1 & 0 & \dots \\ 0 & -1 & 1 & \dots \\ & \ddots & \ddots & \\ \dots & 0 & -1 & 1 \\ 1 & \dots & 0 & -1 \end{bmatrix}, \quad M_+ = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \dots \\ 0 & \frac{1}{2} & \frac{1}{2} & \dots \\ & \ddots & \ddots & \\ \dots & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \dots & 0 & \frac{1}{2} \end{bmatrix}, \quad (80)$$

with periodic boundary conditions, and by

$$D_- = -D_+^T =: D \quad \text{and} \quad M_- = M_+^T =: M.$$

**Proposition 5** (Semi-discrete energy inequality). *The following semi-discrete energy inequality holds,*

$$\frac{d}{dt} \sum_{i,j} \mathcal{E}_{i,j} \leq 0. \quad (81)$$

*Proof.* The central part of the method preserves the energy since, e.g.  $u^T \bar{D}_x p + p^T \bar{D}_x u = u^T \bar{D}_x p + u^T \bar{D}_x^T p = 0$  up to boundary terms. The evolution of the energy of the system is thus entirely given by the numerical stabilization as follows:

$$\frac{2}{\alpha \Delta} \frac{d}{dt} \sum_{i,j} \mathcal{E}_{i,j} = p^T \bar{D}_{xx} p + p^T \bar{D}_{yy} p + u^T \bar{D}_{xx} u + u^T D_{xy} v + v^T D_{xy} u + v^T \bar{D}_{yy} v, \quad (82)$$

where the terms on the right-hand side can be recast as

$$\begin{aligned} p^T \bar{D}_{xx} p &= p^T (D_+ D_-) \otimes (M_+ M_-) p = -\|(D \otimes M)p\|^2 \leq 0, \\ p^T \bar{D}_{yy} p &= p^T (M_+ M_-) \otimes (D_+ D_-) p = -\|(M \otimes D)p\|^2 \leq 0, \\ u^T \bar{D}_{xx} u &= u^T (D_+ D_-) \otimes (M_+ M_-) u = -[(D \otimes M)u]^T [(D \otimes M)u], \\ u^T D_{xy} v &= u^T (D_+ M_+) \otimes (M_- D_-) v = -[(D \otimes M)u]^T [(M \otimes D)v], \\ v^T D_{xy} u &= v^T (M_+ D_+) \otimes (D_- M_-) u = -[(M \otimes D)v]^T [(D \otimes M)u], \\ v^T \bar{D}_{yy} v &= v^T (M_+ M_-) \otimes (D_+ D_-) v = -[(M \otimes D)v]^T [(M \otimes D)v], \end{aligned}$$

where the mixed operator was manipulated thanks to  $MD = DM$ . Hence, the semi-discrete energy is found to decrease:

$$\begin{aligned} \frac{2}{\alpha \Delta} \frac{d}{dt} \sum_{i,j} \mathcal{E}_{i,j} &\leq -[(D \otimes M)u]^T [(D \otimes M)u + (M \otimes D)v] - [(M \otimes D)v]^T [(D \otimes M)u + (M \otimes D)v] \\ &= -\|(D \otimes M)u + (M \otimes D)v\|^2 \leq 0. \end{aligned}$$

This is a discrete version of  $\int_{\Omega} \mathbf{v} \cdot \nabla (\nabla \cdot \mathbf{v}) \, d\mathbf{x} = - \int_{\Omega} (\nabla \cdot \mathbf{v})^2 \, d\mathbf{x} + \text{boundary terms}.$

□

#### 4.5. Analogy with Residual Distribution and discrete steady states

The recent work of [26] has provided a general analysis of the relations between multi-dimensional finite volume methods with point fluxes and residual distribution schemes. Earlier, it has been shown in [2] that residual distribution schemes can be reformulated in terms of a global flux finite volume method. This section elaborates on these aspects for the global flux finite volume approach proposed here. This allows us to give more details on the discrete steady states of the method in a more general setting.

Following the last reference we start from the conservation condition (51) at each corner. Consider, instead of (52), the following ansatz for the numerical global flux, given by the trace of the cell (global) flux plus a fluctuation:

$$\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = \mathcal{F}_{i,j} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)}, \quad \ell, m \in \{0, 1\}. \quad (83)$$

All the properties of the numerical flux can be translated into requirements on the fluctuations  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)}$ . The most interesting ones are related to conservation and stationarity preservation. Corner conservation is written by using the above ansatz in (51), which leads to the requirement

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j)} + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+1,j+1)} + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j+1)} = - \sum_{\ell,m \in \{0,1\}} \mathcal{F}_{i+\ell,j+m} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = -\Phi_{i+\frac{1}{2},j+\frac{1}{2}},$$

where  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}}$  is the global flux integral on the corner dual cell as defined in (58). By virtue of (83), defining a corner flux is thus equivalent to defining a residual distribution scheme satisfying

$$\sum_{\ell,m \in \{0,1\}} \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = -\Phi_{i+\frac{1}{2},j+\frac{1}{2}}. \quad (84)$$

This analogy goes much further, and it is in fact a full equivalence. In particular, we can prove the following facts.

**Proposition 6** (Equivalence with RD). *Consider the multi-dimensional global flux finite volume method (43), with numerical fluxes written in terms of fluctuations (58). Then,*

1. *the multidimensional finite volume global flux method (43) with piecewise constant data is equivalent to the Residual Distribution scheme*

$$\Delta x \Delta y \frac{d}{dt} \bar{q}_{i,j} + \Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \Phi_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \Phi_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \Phi_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)} = 0, \quad (85)$$

*with fluctuations  $\Phi_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^{(i,j)}$  verifying the conservation condition (84) at each corner  $(i \pm \frac{1}{2}, j \pm \frac{1}{2})$ ;*

2. *the finite volume global flux method (43) with average flux  $\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)}$ ,  $\forall \ell, m \in \{0, 1\}$  is equivalent to the residual distribution scheme with fluctuations*

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = (\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathcal{F}_{i+\ell,j+m}) n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)}$$

3. *the finite volume method including the numerical dissipation in (52), defined by the streamline upwind terms (57), is equivalent to the residual distribution scheme defined by*

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = (\overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathcal{F}_{i+\ell,j+m}) n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} - \frac{1}{4} \delta_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} \Phi_{i+\frac{1}{2},j+\frac{1}{2}}, \quad \ell, m \in \{0, 1\}$$

*with  $\delta_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = \alpha \Delta \left( (-1)^\ell \frac{J^x}{\Delta x} + (-1)^m \frac{J^y}{\Delta y} \right)$  according to (57);*

4. both the centered and the stabilized method are steady state preserving with respect to global fluxes of the type (49). In particular, they both admit discrete steady solutions verifying

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}} = 0 \quad \forall i, j ;$$

5. both the centered and the stabilized method are formally second order accurate at steady state for smooth enough solutions.

*Proof.* The first fact is a consequence of the identity  $n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + n_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + n_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + n_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)} = 0$ , so when replacing (83) in (43) we obtain immediately (85): Using (85) or (43) to represent the scheme is absolutely equivalent. The second and third properties are obtained by simply subtracting from the average flux, and from (52) the cell contributions to obtain the corresponding fluctuations.

Concerning the discrete kernel property, consider first the average flux without extra dissipation, and compute explicitly the sum of the corner contributions:

$$\begin{aligned} & \overline{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}} n_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i+\frac{1}{2},j-\frac{1}{2}} n_{i+\frac{1}{2},j-\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i-\frac{1}{2},j+\frac{1}{2}} n_{i-\frac{1}{2},j+\frac{1}{2}}^{(i,j)} + \overline{\mathcal{F}}_{i-\frac{1}{2},j-\frac{1}{2}} n_{i-\frac{1}{2},j-\frac{1}{2}}^{(i,j)} \\ &= \frac{\mathcal{F}_{i+1,j+1} - \mathcal{F}_{i+1,j-1} - \mathcal{F}_{i-1,j+1} + \mathcal{F}_{i-1,j-1}}{4} \end{aligned} \quad (86)$$

$$= \frac{1}{4} \left( \Phi_{i+\frac{1}{2},j+\frac{1}{2}} + \Phi_{i-\frac{1}{2},j+\frac{1}{2}} + \Phi_{i+\frac{1}{2},j-\frac{1}{2}} + \Phi_{i-\frac{1}{2},j-\frac{1}{2}} \right). \quad (87)$$

This shows that  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}} = 0$  is in the kernel of the average flux method. The same is true for the stabilized scheme for which we can write after assembly around cell  $i, j$

$$\begin{aligned} & \sum_{\ell, m \in \{-1, 1\}} \left[ \overline{\mathcal{F}}_{i+\frac{\ell}{2},j+\frac{m}{2}} n_{i+\frac{\ell}{2},j+\frac{m}{2}}^{(i,j)} - \frac{1}{4} \delta_{i+\frac{\ell}{2},j+\frac{m}{2}}^{(i,j)} \Phi_{i+\frac{\ell}{2},j+\frac{m}{2}} \right] \\ &= \sum_{\ell, m \in \{-1, 1\}} \frac{1}{4} \left[ \left( I - \delta_{i+\frac{\ell}{2},j+\frac{m}{2}}^{(i,j)} \right) \Phi_{i+\frac{\ell}{2},j+\frac{m}{2}} \right] \end{aligned} \quad (88)$$

So to prove point 4., we just check that  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}} = 0$  is also a consequence of the steady state preservation condition of Remark 3 in (50). For  $\mathcal{F}_{i,j} = \mathcal{F} + \alpha_i + \beta_j$  as in (49), in each dual cell

$$\Phi_{i+\frac{1}{2},j+\frac{1}{2}} = \mathcal{F} + \alpha_{i+1} + \beta_{j+1} - \mathcal{F} - \alpha_i - \beta_{j+1} - \mathcal{F} - \alpha_{i+1} - \beta_j + \mathcal{F} + \alpha_i + \beta_j = 0$$

and thus  $\widehat{\mathcal{F}}_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)} = \mathcal{F}$ . Finally, the second-order consistency at steady state is a consequence of results for residual distribution schemes with bounded distribution coefficients on linear and bi-linear elements (see [24, 42, 1] and references therein). Second-order accuracy at steady state thus follows from the boundedness of the coefficients  $1/4$ , and  $(I - \delta_{i+\frac{1}{2},j+\frac{1}{2}}^{(i+\ell,j+m)})/4$  which appear in the equivalent forms of the scheme (87) and (88).  $\square$

The switch from conservative finite volume to residual distribution methods contains some nuances on which we would like to comment. In particular, the proof uses two different

writings of the average flux scheme. This may be confusing as to which is the proper form to use. The confusion is originated from the fact that for this simple case, due to cancellation, the same global assembly can be obtained from several different local contributions, not all fitting into the same conservation framework. For example, the proof shows that the scheme

$$\Delta x \Delta y \frac{d}{dt} \bar{q}_{i,j} + \frac{\Phi_{i+\frac{1}{2},j+\frac{1}{2}}}{4} + \frac{\Phi_{i+\frac{1}{2},j-\frac{1}{2}}}{4} + \frac{\Phi_{i-\frac{1}{2},j+\frac{1}{2}}}{4} + \frac{\Phi_{i-\frac{1}{2},j-\frac{1}{2}}}{4} = 0 \quad (89)$$

is equivalent to the average flux scheme. In light of (85), this may lead to the conclusion, that the definition  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}}^{(i,j)} = \Phi_{i+\frac{1}{2},j+\frac{1}{2}}/4$  is a viable one; which is, however, wrong. Such a definition is not acceptable here, as it does not satisfy the local conservation constraint (84), which has a minus on the right hand side. This change of sign is related to the difference between internal and exterior oriented normals. Scheme (89) can also be shown to be conservative in the classical cell-vertex residual distribution framework, with the appropriate conservation constraint. On Cartesian meshes, this mismatch cancels out and one ends up with the same discretization after assembly. However, on general meshes the change in sign must be carefully accounted for (see e.g. [26]), and the appropriate local conservation and consistency conditions respected. In particular, the central RD (89) can indeed be written in a flux form, but the numerical flux has a much more involved expression than the simple average flux. The interested reader can refer to the last reference, and to [2] for more details.

#### 4.6. Source modification to preserve solutions at rest

This subsection has the goal of providing a direct way to embed bathymetry source terms typical of the shallow water system with bottom topography. For simplicity, this part revolves around this PDE system, but the same approach can be also applied for other cases. In particular, the goal is to achieve stationarity preservation for motionless equilibria, i.e. lake at rest preservation, coming from the balance between the hydrodynamic pressure and bottom topography, which is present in both 1D and 2D configurations:

$$\begin{aligned} \text{1D: } \begin{cases} h(x) + b(x) \equiv \eta_0, \\ u(x) \equiv 0, \end{cases} \quad \text{2D: } \begin{cases} h(x, y) + b(x, y) \equiv \eta_0, \\ u(x, y) = v(x, y) \equiv 0. \end{cases} \end{aligned} \quad (90)$$

As also shown in section 3, in a 1D global flux framework ([23]) it has been proposed to integrate the source terms within the flux derivative, thus obtaining a quasi-conservative formulation of the PDE:

$$\begin{cases} \partial_t h + \partial_x q_x &= 0, \\ \partial_t q_x + \partial_x \left( \frac{q_x^2}{h} + g \frac{h^2}{2} \right) &= -gh \partial_x b, \end{cases} \implies \begin{cases} \partial_t h + \partial_x q_x &= 0, \\ \partial_t q_x + \partial_x \left( \frac{q_x^2}{h} + g \frac{h^2}{2} + \int^x gh \partial_\xi b d\xi \right) &= 0. \end{cases} \quad (91)$$

However, contrary to classical source terms, the bathymetry term shall be treated differently given the presence of its derivative. To achieve consistency and well-balancedness for the high order method, in [23] the integral of the bathymetry source is considered to jump at each interface. In the current low order framework the approach amounts to

$$R_i^x = \int^{x_i} gh \partial_\xi b d\xi = \int^{x_{i-1}} gh \partial_\xi b d\xi + \int_{x_{i-1}}^{x_i} gh \partial_\xi b d\xi = R_{i-1}^x + g \frac{h_i + h_{i-1}}{2} (b_i - b_{i-1}), \quad (92)$$

where the second integral has been computed using a consistent approximation of  $\partial_\xi b$ , while for the  $h$  a simple trapezoidal rule has been used.

The same approach can also be developed for 2D systems, including the the source term containing  $\partial_x$  in the  $x$ -flux, and the one containing  $\partial_y$  in the  $y$ -flux. Starting from equation (21), we can integrate the source terms present in the momentum equation as follows:

$$\begin{cases} \partial_t h + \partial_x(hu) + \partial_y(hv) = 0 \\ \partial_t(hu) + \partial_x\left(hu^2 + g\frac{h^2}{2} + \int^x gh\partial_\xi b \, d\xi\right) + \partial_y(huv) = 0, \\ \partial_t(hv) + \partial_x(huv) + \partial_y\left(hv^2 + g\frac{h^2}{2} + \int^y gh\partial_\eta b \, d\eta\right) = 0. \end{cases} \quad (93)$$

**Proposition 7** (Lake at rest preservation). *The 2D global flux scheme of the system (93) with the source term quadrature provided in equation (92) is exactly well-balanced for the lake at rest solution (90).*

*Proof.* To prove that the family of equilibria (90) is exactly preserved when  $u = v \equiv 0$  and  $\eta = h + b \equiv \eta_0$ , Since the mixed terms depend only on the velocity and thus vanish, the two momentum equations can be treated separately for the  $x$  and  $y$  contribution. In particular we want to show that given a zero velocity and constant free surface elevation,  $f_{i,j} + R_{i,j}^x = f_{i-1,j} + R_{i-1,j}^x$ ,  $\forall j$ . Without loss of generality, we will show this result only for the  $x$  direction.

By substitution of the relevant quantities, we obtain

$$f_i + R_i^x - f_{i-1} - R_{i-1}^x = g\frac{h_i^2 - h_{i-1}^2}{2} + g\frac{h_i + h_{i-1}}{2}(b_i - b_{i-1}) \quad (94)$$

$$= g\frac{h_i + h_{i-1}}{2}(\eta_i - \eta_{i-1}) = 0, \quad \forall j, \quad (95)$$

where the last equality holds when  $\eta_i = \eta_{i-1} \equiv \eta_0$ , with  $\eta_i = h_i + b_i$ .  $\square$

#### 4.7. Compatible boundary conditions

Boundary conditions play an essential role in the practical application of numerical methods. Let us for example consider the usual ghost cell approach. In a classical dimension-by-dimension finite volume method, homogeneous Neumann boundary conditions on the state variables can be simply enforced by copying the state. This is somewhat consistent with the internal treatment based on one dimensional Riemann fluxes using two states. However, this approach cannot be steady state preserving since it is not based on the global flux. One should consider corner fluxes on the boundaries, and choose the ghost states consistently with the equilibrium condition. Here, we construct compatible Neumann boundaries based on the discrete constraint  $\Phi_{\text{corner}} = 0$ , which we have shown to be the relevant multi-dimensional characterization of our discrete steady states. Compatible transmissive conditions require this relation to be verified by the ghost cells at each boundary corner.

Let us consider the right boundary domain (see figure 3). We can impose  $\mathcal{F}_{N+1,j-1} := \mathcal{F}_{N,j-1}$  for all  $j$  on the global fluxes, instead of computing the state variables as for classical homogeneous Neumann conditions. Then, we can use directly the global flux at the boundary corners. This will lead to the following relation

$$\Phi_{N+\frac{1}{2},j-\frac{1}{2}} = \mathcal{F}_{N+1,j} - \mathcal{F}_{N,j} - \mathcal{F}_{N+1,j-1} + \mathcal{F}_{N,j-1} = 0, \quad (96)$$

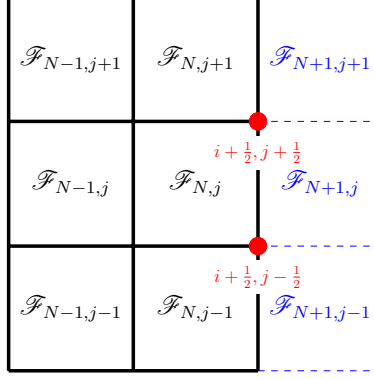


Figure 3: Ghost cell labelling: ghost cells are highlighted in blue.

which gives a steady state solution.

On corner ghosts, similarly, one has to impose  $\mathcal{F}_{N_x+1,N_y+1} := \mathcal{F}_{N_x,N_y+1}$ , if also on the top side transmissive conditions must be imposed, this will results in

$$\mathcal{F}_{N_x+1,N_y+1} = \mathcal{F}_{N_x,N_y+1} = \mathcal{F}_{N_x+1,N_y} = \mathcal{F}_{N_x,N_y}.$$

Similar ideas can be used also for other boundary types, but in our experience transmissive conditions are the most critical to correctly maintain the internal structure, since they involve no external data, which provide some link to the correct solution for other boundary conditions.

When seeking to preserve steady states, it is crucial to ensure that the number of equations imposed—either by boundary conditions or by the steady state equations—does not exceed the number of unknowns, or that these equations are mutually compatible.

The steady state conditions enforced by  $\Phi_{i+\frac{1}{2},j+\frac{1}{2}}$  for  $i = 0, \dots, N_x$  and  $j = 0, \dots, N_y$  introduce  $N_{eq}(N_x + 1)(N_y + 1)$  linearly independent constraints on  $N_{eq}N_xN_y$  unknowns  $\bar{q}_{i,j}$  (for  $i = 1, \dots, N_x$ ,  $j = 1, \dots, N_y$ ). To satisfy these extra constraints,  $N_{eq}(2N_x + 2N_y + 4)$  ghost cell values are introduced. This leaves  $N_{eq}(N_x + N_y + 3)$  equations that can be specified at the boundaries, typically through Dirichlet conditions. Homogeneous Neumann conditions, as previously discussed, are compatible with the internal constraints and do not introduce additional equations. Therefore, the  $N_{eq}(N_x + N_y + 3)$  remaining degrees of freedom correspond to at most two sides where Dirichlet boundary conditions can be imposed, and if not consecutive sides, they should not have all corners included.

In summary, when seeking equilibria, the boundary conditions must be compatible with the internal constraints. It is only possible, and necessary, to impose complete Dirichlet conditions on (at most) two boundaries. In this respect, this work is changing the perspective on this issue. For many years, schemes similar to the one obtained here have been considered as flawed due to the existence of the steady states characterized by proposition 6. This is due to the fact that spurious oscillating modes may also satisfy the condition  $\Phi_{i+1/2,j+1/2} = 0 \forall i, j$  (see e.g. [42, 1] and references therein). However, this is only true if one considers the problem locally, which is a wrong way to define multidimensional steady states as they must include boundary conditions. If these are imposed in a compatible manner, spurious modes can be controlled. This work, as well as the work discussed in [11], contributes to rectifying this notion.

## 5. Standard finite volume scheme used for comparison

In this section, we present the standard finite volume (FV) scheme used for comparison with the novel global flux (GF) scheme in the numerical experiments presented in section 6. The classical finite volume formulation for the 2D nonlinear hyperbolic problem (11) can be written by integrating it in the cell  $C_{i,j}$ :

$$\frac{d}{dt} \bar{q}_{i,j} + \frac{\hat{f}_{i+\frac{1}{2},j} - \hat{f}_{i-\frac{1}{2},j}}{\Delta x} + \frac{\hat{g}_{i,j+\frac{1}{2}} - \hat{g}_{i,j-\frac{1}{2}}}{\Delta y} = \bar{s}_{i,j}, \quad (97)$$

where the numerical flux  $\hat{f}_{i+\frac{1}{2},j}$  is computed through the local Lax-Friedrichs (or Rusanov) flux:

$$\hat{f}_{i+\frac{1}{2},j} = \frac{1}{2} \left( f_{i+\frac{1}{2},j}^L + f_{i+\frac{1}{2},j}^R \right) - \frac{\lambda_m}{2} \left( q_{i+\frac{1}{2},j}^R - q_{i+\frac{1}{2},j}^L \right), \quad (98)$$

where

$$f_{i+\frac{1}{2},j}^L = f(q_{i+\frac{1}{2},j}^L), \quad f_{i+\frac{1}{2},j}^R = f(q_{i+\frac{1}{2},j}^R), \quad (99)$$

and similarly for the others. The source term is computed by integrating the source term over the cell  $C_{i,j}$ :

$$\bar{s}_{i,j} = \frac{1}{\Delta x \Delta y} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} s dx dy. \quad (100)$$

In this work, we are going to compare the first order global flux scheme against the standard finite volume with both piece-wise constant (first order accurate) and piece-wise linear (second order accurate) reconstructions. Since the solution update can be performed in a dimension-by-dimension way, we can focus only on the  $x$ -direction for simplicity. For piece-wise constant reconstruction, the left and right states at interface  $i + \frac{1}{2}$  simply are

$$q_{i+\frac{1}{2},j}^L = \bar{q}_{i,j}, \quad q_{i+\frac{1}{2},j}^R = \bar{q}_{i+1,j}, \quad \forall j. \quad (101)$$

While in the second order case, the left and right states are computed through a piece-wise linear reconstruction of the solution as

$$\tilde{q}(x, y) = \bar{q}_{i,j} + (x - x_i) (\partial_x q)_{i,j} + (y - y_j) (\partial_y q)_{i,j}, \quad x, y \in C_{i,j}. \quad (102)$$

Hence, the left and right states at interface  $i + \frac{1}{2}$  are

$$q_{i+\frac{1}{2},j}^L = \bar{q}_{i,j} + \frac{\Delta x}{2} (\partial_x q)_{i,j}, \quad q_{i+\frac{1}{2},j}^R = \bar{q}_{i+1,j} - \frac{\Delta x}{2} (\partial_x q)_{i+1,j}, \quad \forall j. \quad (103)$$

Here, the slopes  $(\partial_x q)_{i,j}$  are evaluated using the generalized minmod limiter [41]:

$$(\partial_x q)_{i,j} = \text{minmod} \left( \vartheta \frac{\bar{q}_{i+1,j} - \bar{q}_{i,j}}{\Delta x}, \frac{\bar{q}_{i+1,j} - \bar{q}_{i-1,j}}{2\Delta x}, \vartheta \frac{\bar{q}_{i,j} - \bar{q}_{i-1,j}}{\Delta x} \right), \quad (104)$$

where  $\vartheta$  is used to control the amount of dissipation. In particular, the larger  $\vartheta$  is, the sharper and more oscillatory the reconstruction will be. For the simulations presented in section 6, when not specified, we set the parameter  $\vartheta = 1.3$ . The same approach has been used along the  $y$  direction.

The classical minmod function is defined as

$$\text{minmod}(a, b, c) = \begin{cases} \min(a, b, c), & \text{if } a, b, c > 0, \\ \max(a, b, c), & \text{if } a, b, c < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (105)$$

## 6. Numerical experiments

In this section, the goal is to show the performance of the new global flux scheme (GF) when compared to the classical finite volume first order (FV-1) and second order (FV-2) approaches presented in section 5. Several test cases are presented to study the impact of the method on both linear and nonlinear hyperbolic problems: linear acoustics, shallow water and the Euler equations. The numerical experiments are performed taking the gravity  $g = 9.812$ , for the shallow water system, and the ratio of specific heats  $\gamma = 1.4$  for the Euler system. All convergence analyses are performed on a set of nested quadrilateral meshes with  $N_x = N_y = 20, 40, 80, 160, 320$ . Time integration is performed through classical explicit Euler and second order Runge-Kutta methods.

### 6.1. Linear acoustic system

#### 6.1.1. Stationary vortex

The first test case considered here concerns the simulation of the linear acoustic system. The initial condition is a compactly supported vortex centered in  $(x_0, y_0) = (0.5, 0.5)$  defined on the square  $[0, 1] \times [0, 1]$  with periodic boundary conditions, which is given by

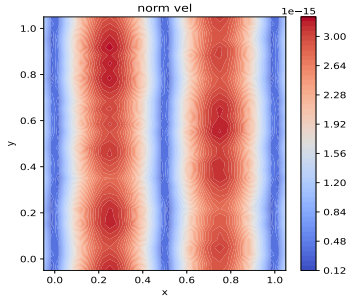
$$\begin{aligned} p(x, y) &= 1, \\ u(x, y) &= (y - y_0)f(\rho(x, y)), \\ v(x, y) &= -(x - x_0)f(\rho(x, y)), \end{aligned}$$

with  $\rho(x, y) = \frac{\sqrt{(x-x_0)^2 + (y-y_0)^2}}{r_0}$ , where  $r_0 = 0.45$ ,  $f(\rho) = \gamma(1 + \cos(\pi\rho))^2$  and  $\gamma = \frac{12\pi\sqrt{0.981}}{r_0\sqrt{315\pi^2 - 2048}}$ . This vortex is taken from the work [44], where its derivation is described. This initial condition is a steady state of the acoustic system.

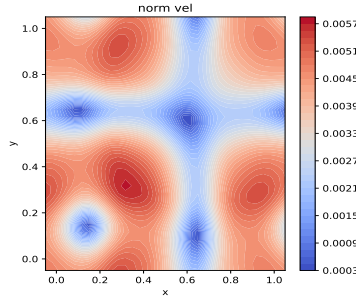
In table 1, the errors computed with the  $L_2$  norm and convergence rates are shown. As can be noticed, the GF method outperforms the standard FV-1 and FV-2 methods in terms of discretization errors. Although the GF is in principle first order accurate due to the piece-wise constant reconstruction, a superconvergence behavior is experienced for stationary solutions (compare Proposition 6). Hence, the GF method is not only able to preserve the vortex structure, but does so at second order accuracy. In particular, this is evident when increasing the final time of the simulation, as shown in figure 4. Classical methods like FV-1 and FV-2 are not able to preserve the vortex structure for long times, due to their numerical dissipation that spoils the final solution.

Table 1: Linear acoustic system: vortex ( $t_f = 1$ ).  $L_2$  error and order of accuracy  $\tilde{n}$  for FV-1, FV-2 and GF.

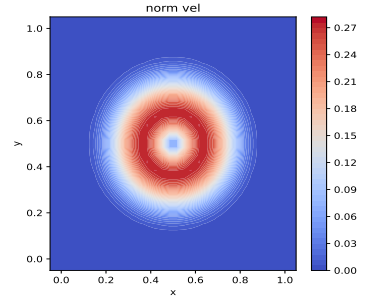
| $N_x, N_y$ | $p$      |             | $u$      |             | $v$      |             |
|------------|----------|-------------|----------|-------------|----------|-------------|
|            | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ |
| FV-1       |          |             |          |             |          |             |
| 20         | 3.51E-05 | —           | 6.51E-02 | —           | 6.51E-02 | —           |
| 40         | 4.58E-05 | -0.38       | 5.42E-02 | 0.26        | 5.42E-02 | 0.26        |
| 80         | 2.71E-05 | 0.75        | 3.97E-02 | 0.44        | 3.97E-02 | 0.44        |
| 160        | 1.06E-05 | 1.35        | 2.54E-02 | 0.64        | 2.54E-02 | 0.64        |
| 320        | 3.34E-06 | 1.66        | 1.47E-02 | 0.79        | 1.47E-02 | 0.79        |
| FV-2       |          |             |          |             |          |             |
| 20         | 4.31E-04 | —           | 2.58E-02 | —           | 2.58E-02 | —           |
| 40         | 2.54E-04 | 0.76        | 6.12E-03 | 2.07        | 6.12E-03 | 2.07        |
| 80         | 6.37E-05 | 1.99        | 1.61E-03 | 1.93        | 1.61E-03 | 1.93        |
| 160        | 1.29E-05 | 2.30        | 4.46E-04 | 1.84        | 4.46E-04 | 1.84        |
| 320        | 2.73E-06 | 2.24        | 1.25E-04 | 1.83        | 1.25E-04 | 1.83        |
| GF         |          |             |          |             |          |             |
| 20         | 4.72E-05 | —           | 3.95E-04 | —           | 3.95E-04 | —           |
| 40         | 4.53E-05 | 0.05        | 9.17E-05 | 2.10        | 9.17E-05 | 2.10        |
| 80         | 1.95E-05 | 1.21        | 2.26E-05 | 2.01        | 2.26E-05 | 2.01        |
| 160        | 6.42E-06 | 1.60        | 5.58E-06 | 2.01        | 5.58E-06 | 2.01        |
| 320        | 1.85E-06 | 1.79        | 1.38E-06 | 2.01        | 1.38E-06 | 2.01        |



(a) FV-1



(b) FV-2



(c) GF

Figure 4: Linear acoustic system: vortex. Isocontours of the velocity norm obtained with FV-1, FV-2 and GF after a long time integration ( $t_f = 200$ ).

## 6.2. Euler equations

### 6.2.1. Isentropic vortex

In this section, we test the proposed method on a smooth isentropic vortex [29]. The initial condition is given in terms of primitive variables and it consists in superposition of a homogeneous background flow and a perturbation:

$$(\rho, u, v, p) = (1 + \delta\rho, u_0 + \delta u, v_0 + \delta v, 1 + \delta p).$$

The test case is set up in a  $[0, 10] \times [0, 10]$  domain with periodic boundary conditions and vortex radius  $r = \sqrt{(x-5)^2 + (y-5)^2}$ . The vortex strength is  $\epsilon = 5$ , and the entropy perturbation is assumed to be zero. Given these hypotheses, the perturbations on velocity and temperature can be written as

$$\begin{bmatrix} \delta u \\ \delta v \end{bmatrix} = \frac{\epsilon}{2\pi} \exp\left(\frac{1-r^2}{2}\right) \begin{bmatrix} -(y-5) \\ (x-5) \end{bmatrix}, \quad \delta T = -\frac{(\gamma-1)\epsilon^2}{8\gamma\pi^2} \exp(1-r^2).$$

It follows that the perturbations on density and pressure read

$$\delta\rho = (1 + \delta T)^{\frac{1}{\gamma-1}} - 1, \quad \delta p = (1 + \delta T)^{\frac{\gamma}{\gamma-1}} - 1.$$

This test case is a stationary solution of the Euler equations.

In table 2, the convergence analysis for the isentropic vortex is presented by comparing the FV-1, FV-2 and GF methods by running the simulation of a static vortex, i.e.  $u_0 = v_0 = 0$ , until a final time  $t_f = 1$ . As observed above, the GF shows superconvergent behavior with order 2. In terms of discretization errors, it outperforms not only the classical piecewise constant finite volume method, but also the second-order approach equipped with a linear reconstruction. Even after very long simulations times (see figure 5) the new GF method is able to maintain the vortex, while the first order FV-1 dissipates everything away, and FV-2 significantly distorts the vortex structure and still diffuses it more than GF. Observe that the nonlinearity of the equations makes this test significantly more challenging than the corresponding test for linear acoustics, where in particular advection is not present.

In table 3, the convergence analysis for a moving isentropic vortex is presented with  $u_0 = v_0 = 1$  and a final time  $t_f = 10$ . Here we can directly observe that the GF method is indeed first order accurate, as expected, since the reconstruction is piecewise constant. No superconvergence is observed in this case, as the solution is not stationary. The results show an relevant improvement of the GF method over the FV-1 in both discretization errors and convergence rates, while the FV-2 method is, in this case, the best since it is able to achieve second order accuracy. In figure 6, the solution at the final time is shown for the three methods.

### 6.2.2. Perturbation of the isentropic vortex

In this section, we present a test case for the Euler equations that consists in a perturbation of the isentropic vortex presented in the previous section. The initial conditions for the three schemes FV-1, FV-2 and GF are taken as the final results  $q_{\text{eq}}$  of the respective simulations run until final time  $t_f = 50$  with a  $80 \times 80$  mesh.

Table 2: Euler equations: isentropic vortex with  $u_0 = v_0 = 0$  ( $t_f = 1$ ).  $L_2$  error and order of accuracy  $\tilde{n}$  for FV-1, FV-2 and GF methods.

| $N_x, N_y$ | $\rho$   |             | $\rho u$ |             | $\rho v$ |             | $\rho E$ |             |
|------------|----------|-------------|----------|-------------|----------|-------------|----------|-------------|
|            | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ |
| FV-1       |          |             |          |             |          |             |          |             |
| 20         | 3.58E-01 | –           | 6.77E-01 | –           | 6.77E-01 | –           | 1.16E+00 | –           |
| 40         | 2.47E-01 | 0.53        | 4.40E-01 | 0.62        | 4.40E-01 | 0.62        | 8.29E-01 | 0.48        |
| 80         | 1.49E-01 | 0.72        | 2.59E-01 | 0.76        | 2.59E-01 | 0.76        | 5.15E-01 | 0.68        |
| 160        | 8.33E-02 | 0.84        | 1.43E-01 | 0.85        | 1.43E-01 | 0.85        | 2.91E-01 | 0.82        |
| 320        | 4.42E-02 | 0.91        | 7.56E-02 | 0.91        | 7.56E-02 | 0.91        | 1.56E-01 | 0.90        |
| FV-2       |          |             |          |             |          |             |          |             |
| 20         | 1.06E-01 | –           | 2.05E-01 | –           | 2.00E-01 | –           | 4.32E-01 | –           |
| 40         | 3.62E-02 | 1.55        | 6.74E-02 | 1.60        | 6.71E-02 | 1.57        | 1.20E-01 | 1.85        |
| 80         | 1.07E-02 | 1.76        | 1.93E-02 | 1.80        | 1.95E-02 | 1.78        | 2.91E-02 | 2.04        |
| 160        | 2.39E-03 | 2.16        | 5.58E-03 | 1.78        | 5.61E-03 | 1.79        | 7.04E-03 | 2.04        |
| 320        | 5.12E-04 | 2.22        | 1.39E-03 | 2.00        | 1.39E-03 | 2.01        | 1.56E-03 | 2.17        |
| GF         |          |             |          |             |          |             |          |             |
| 20         | 1.52E-02 | –           | 3.67E-02 | –           | 3.67E-02 | –           | 4.59E-02 | –           |
| 40         | 5.95E-03 | 1.35        | 1.15E-02 | 1.67        | 1.15E-02 | 1.67        | 1.54E-02 | 1.57        |
| 80         | 1.76E-03 | 1.76        | 3.06E-03 | 1.90        | 3.06E-03 | 1.90        | 4.35E-03 | 1.82        |
| 160        | 4.69E-04 | 1.90        | 7.87E-04 | 1.96        | 7.87E-04 | 1.96        | 1.16E-03 | 1.90        |
| 320        | 1.21E-04 | 1.95        | 2.00E-04 | 1.97        | 2.00E-04 | 1.97        | 3.02E-04 | 1.94        |

Table 3: Euler equations: isentropic vortex with  $u_0 = v_0 = 1$  ( $t_f = 10$ ).  $L_2$  error and order of accuracy  $\tilde{n}$  for FV-1, FV-2 and GF methods.

| $N_x, N_y$ | $\rho$   |             | $\rho u$ |             | $\rho v$ |             | $\rho E$ |             |
|------------|----------|-------------|----------|-------------|----------|-------------|----------|-------------|
|            | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ |
| FV-1       |          |             |          |             |          |             |          |             |
| 20         | 6.50E-01 | –           | 1.54E+00 | –           | 1.54E+00 | –           | 3.12E+00 | –           |
| 40         | 6.21E-01 | 0.06        | 1.46E+00 | 0.07        | 1.46E+00 | 0.07        | 3.01E+00 | 0.05        |
| 80         | 5.82E-01 | 0.09        | 1.31E+00 | 0.15        | 1.31E+00 | 0.15        | 2.83E+00 | 0.09        |
| 160        | 5.13E-01 | 0.18        | 1.06E+00 | 0.30        | 1.07E+00 | 0.29        | 2.48E+00 | 0.19        |
| 320        | 4.01E-01 | 0.35        | 7.58E-01 | 0.49        | 7.63E-01 | 0.49        | 1.92E+00 | 0.36        |
| FV-2       |          |             |          |             |          |             |          |             |
| 20         | 5.29E-01 | –           | 1.07E+00 | –           | 1.12E+00 | –           | 2.48E+00 | –           |
| 40         | 2.45E-01 | 1.10        | 4.42E-01 | 1.28        | 4.84E-01 | 1.21        | 1.17E+00 | 1.08        |
| 80         | 6.55E-02 | 1.90        | 1.26E-01 | 1.80        | 1.37E-01 | 1.81        | 2.82E-01 | 2.04        |
| 160        | 1.85E-02 | 1.82        | 3.22E-02 | 1.97        | 3.42E-02 | 2.00        | 6.03E-02 | 2.22        |
| 320        | 3.38E-03 | 2.45        | 7.49E-03 | 2.10        | 8.12E-03 | 2.07        | 1.28E-02 | 2.23        |
| GF         |          |             |          |             |          |             |          |             |
| 20         | 5.46E-01 | –           | 1.30E+00 | –           | 1.13E+00 | –           | 2.58E+00 | –           |
| 40         | 4.44E-01 | 0.30        | 1.02E+00 | 0.34        | 7.71E-01 | 0.55        | 2.10E+00 | 0.29        |
| 80         | 3.18E-01 | 0.48        | 6.96E-01 | 0.56        | 4.92E-01 | 0.64        | 1.53E+00 | 0.45        |
| 160        | 1.99E-01 | 0.67        | 4.12E-01 | 0.75        | 2.96E-01 | 0.73        | 9.73E-01 | 0.65        |
| 320        | 1.12E-01 | 0.82        | 2.24E-01 | 0.87        | 1.68E-01 | 0.81        | 5.55E-01 | 0.80        |

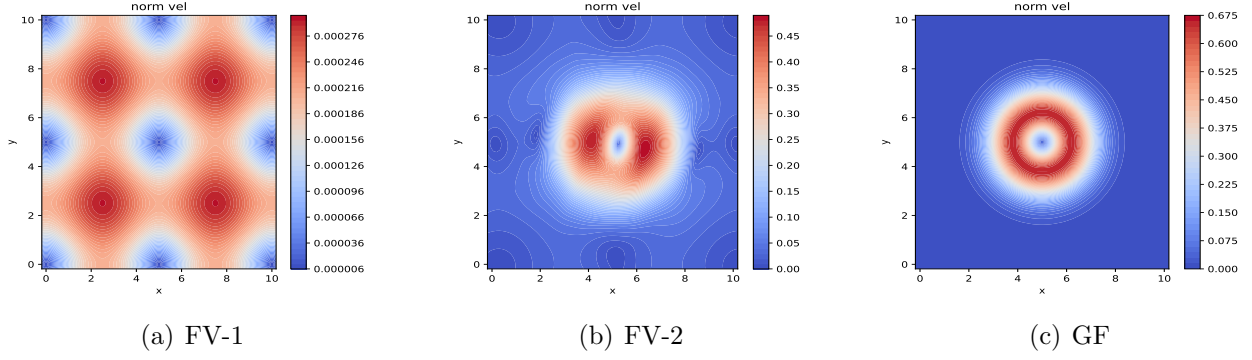


Figure 5: Euler equations: isentropic vortex with  $u_0 = v_0 = 0$ . Isocontours of the velocity norm obtained with FV-1, FV-2 and GF after a long time integration ( $t_f = 200$ ).

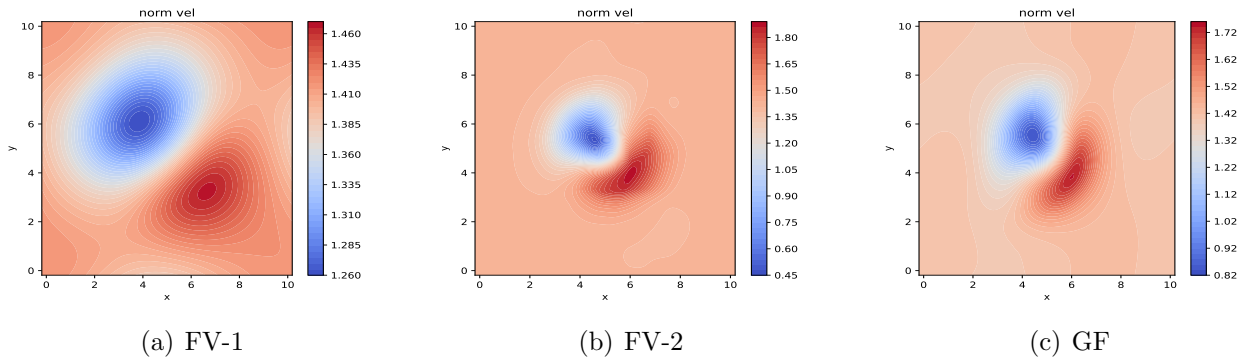


Figure 6: Euler equations: isentropic vortex with  $u_0 = v_0 = 1$ . Isocontours of the velocity norm obtained with FV-1, FV-2 and GF at  $t_f = 10$ .

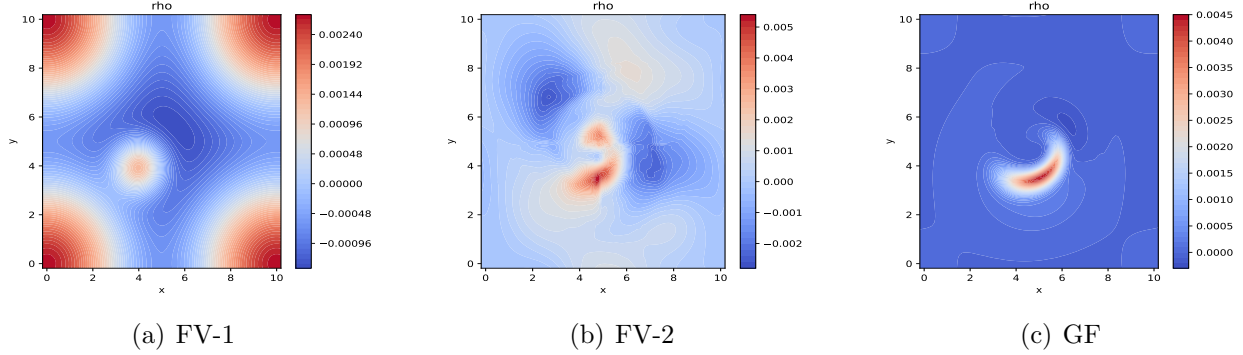


Figure 7: Euler equations: perturbation of the isentropic vortex. Isocontours of the  $\rho - \rho_{\text{eq}}$  norm obtained with FV-1, FV-2 and GF at final time  $t_f = 2$  with a  $80 \times 80$  mesh.

Then, we add to the initial conditions a density perturbation  $\delta\rho$  centered in  $(4, 4)$  of the form:

$$\delta\rho = Ae^{-\frac{(x-4)^2+(y-4)^2}{\sigma^2}}$$

where  $A = 5 \cdot 10^{-3}$  and  $\sigma = 0.8$ . The simulation is run until a final time  $t = 2$  to compare the effect of the numerical viscosity on the evolution of the perturbation.

In figure 7, we show the density contour plot at the final time for the three methods. The GF method is able to capture the perturbation sharply, while the FV-1 and FV-2 methods have discretization errors too large to capture it properly. By looking at the isocontours scales, it is clear that the perturbation is completely dissipated for the FV-1 method, while for the FV-2 method the perturbation is still visible but with a much larger error compared to the expected solution.

### 6.2.3. Sod's circular problem

Here, we test the robustness of the global flux method on the Euler equations for the Sod circular problem. The problem is a two-dimensional extension of the classical shock tube problem. The simulation is performed on a domain  $[-1, 1] \times [-1, 1]$  and the initial condition is given by

$$Q(\mathbf{x}, 0) = \begin{cases} Q_i & \text{if } r \leq R, \\ Q_e & \text{if } r > R, \end{cases}$$

with  $r = \sqrt{x^2 + y^2}$ . The circle of radius  $R = 0.5$  is centered in the origin and separates the inner state  $Q_i$  from the outer state  $Q_e$ , where  $Q = (\rho, u, v, p)$ . The initial conditions are given by  $Q_i = (1, 0, 0, 1)$  and  $Q_e = (0.125, 0, 0, 0.1)$ . For a reference solution of this problem, we refer to [13].

To have a smoother initial condition, the two states are connected by a smooth transition region given by an erfc function defined as

$$\zeta(r) = \frac{1}{2} \text{erfc} \left( \frac{r - R}{\delta} \right),$$

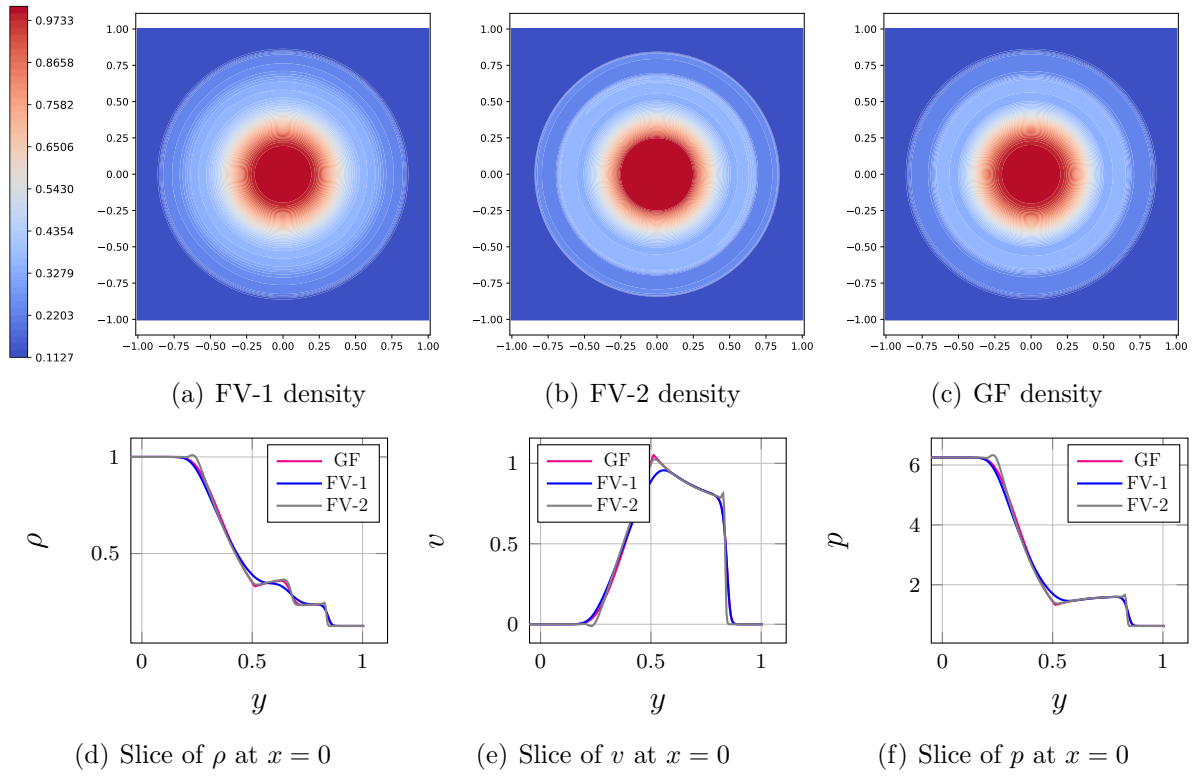


Figure 8: Euler equations: Sod's circular problem. Numerical results obtained on a  $400 \times 400$  mesh with FV-1, FV-2 and GF methods run until a final time  $t_f = 0.2$ .

where  $\delta = 0.01$ . Therefore, we can define the smoothed initial condition as

$$Q(r, 0) = \zeta(r)Q_i + (1 - \zeta(r))Q_e.$$

The simulation is run until final time  $t_f = 0.2$  before the shock waves reach the boundaries.

In figure 8, we present the numerical results obtained on a  $400 \times 400$  mesh with the three methods FV-1, FV-2 and GF. For all situations, we show the density contour plot at the final time, along with a slice of the density and vertical velocity on the  $x = 0$  axis. It can be noticed that, among all three simulations, the GF performs much better than the standard first order scheme and it is clearly comparable to a second order one. It is able to sharply capture the three waves, which are smoothed out by the classical FV-1. GF also avoids oscillations at the beginning of the rarefaction, while the FV-2 shows small oscillations. On the foot of the rarefaction, the GF show sharper results, while it is a little more diffusive on the contact discontinuity with respect to the FV-2. On the shock, the GF does not oscillate, while the FV-2 shows minimal oscillations and is a little more sharply representing the discontinuity.

This test case not only allows us to show the robustness of the method to deal with unsteady shock propagation. It also provides interesting insights into its low dissipation even though the method has not been designed to have any particular properties on unsteady solutions.

#### 6.2.4. Kelvin-Helmholtz instability

The last test case we present for the Euler system is a smooth Kelvin-Helmholtz instability, introduced in [33] to study the ability of a numerical scheme to cope with low Mach number flow and to assess qualitatively the numerical diffusion of the method. There is a large body of work available in the literature concerning the shortcomings of classical Finite Volume methods in the subsonic regime (see e.g. [9]). The effect of stabilizing diffusion becomes bigger as the Mach number decreases, making it necessary to use highly resolved grids in order to capture the features of the flow. Numerical methods that are not low Mach number compliant typically also stabilize Kelvin-Helmholtz setups in an artificial way.

The simulations are performed in the domain  $[0, 2] \times [-1/2, 1/2]$  until a final time  $t_f = 80$ . The initial condition is given by the following primitive variables:

$$\rho = \gamma + \mathcal{H}(y)r, \quad u = M\mathcal{H}(y), \quad v = \delta M \sin(2\pi x), \quad p = 1,$$

where the Mach number parameter is  $M = 10^{-2}$  and we use  $r = 10^{-3}$  and  $\delta = 0.1$ . The function  $\mathcal{H}(y)$  is defined as,

$$\mathcal{H}(y) = \begin{cases} -\sin\left(\frac{\pi}{\omega}\left(y + \frac{1}{4}\right)\right), & \text{if } -\frac{1}{4} - \frac{\omega}{2} \leq y < -\frac{1}{4} + \frac{\omega}{2}, \\ -1, & \text{if } -\frac{1}{4} + \frac{\omega}{2} \leq y < \frac{1}{4} - \frac{\omega}{2}, \\ \sin\left(\frac{\pi}{\omega}\left(y - \frac{1}{4}\right)\right), & \text{if } \frac{1}{4} - \frac{\omega}{2} \leq y < \frac{1}{4} + \frac{\omega}{2}, \\ 1 & \text{else,} \end{cases}$$

where  $\omega = 1/16$ . Observe that the shear flow is smooth such that for short times, there exists a solution to which numerical methods converge upon mesh refinement ([33]).

In figure 9, we present the numerical results obtained with FV-1, FV-2 and GF for the Kelvin-Helmholtz instability arising from the aforementioned initial conditions. The

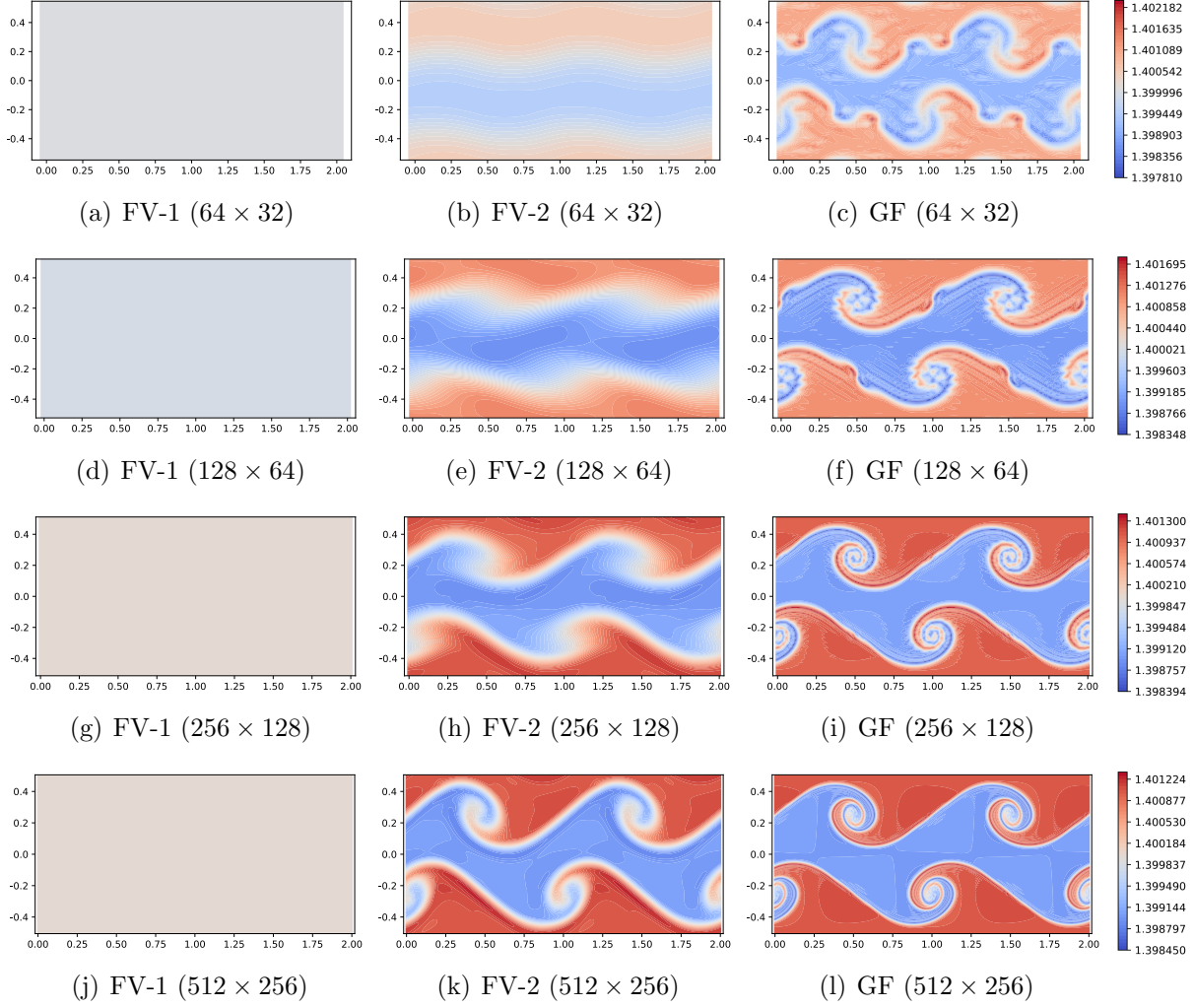


Figure 9: Euler equations: Kelvin-Helmholtz instability. Density isocontours are presented for a set of nested meshes to compare FV-1 (top), FV-2 (middle) and GF (bottom).

simulations are performed on a set of 4 nested grids from a  $64 \times 32$ , the coarsest, to  $512 \times 256$ , the finest.

The FV-1 scheme is not able to capture any of the features arising from the instability. No vortices form, since FV-1 is not low Mach number compliant.

Much improved results are obtained using the FV-2 with a linear reconstruction of the conservative variables. Here, the higher order of accuracy helps to overcome excessive diffusion at this Mach number and for this simulation time. However, the structures still appear diffused and would need even more resolution for the vortex details to be captured.

Very differently from these methods, the GF method is able to capture all details of the flow very accurately. Already on the coarsest mesh, the fluid structures start to appear and develop. Here, some spurious vortices are visible, which are a known artefact of virtually any numerical method (see e.g. [15]). When increasing the resolution, the fluid features converge to the expected solution found in other references [33]. Comparison to the results obtained with low Mach compliant methods studied in [33] shows that the GF method is at least as good.

It has been suggested in [8] that numerical methods for the Euler equations whose linearization (= method for linear acoustics) is stationarity preserving, are low Mach number compliant. A nonlinear stationarity preserving method naturally has this property, and some experimental examples of this behaviour can also be found in [7]. Thus, even though we set out to improve the performance of the numerical method at stationary state, here we observe that this property is beneficial even for solutions far away from it.

### 6.3. Shallow water system

#### 6.3.1. Potential flow

The first test case implemented for the shallow water equations is an equilibrium (see [43]) characterized by a known exact solution, for which it is possible to perform a convergence analysis. The initial condition is a potential flow defined on the square  $[-1, 1] \times [-1, 1]$  with Dirichlet boundary conditions, which is given by

$$\begin{aligned} h(x, y) &= (x - x_0)(y - y_0) + C, \\ u(x, y) &= (x - x_0), \\ v(x, y) &= -(y - y_0) \end{aligned}$$

where  $C = 3/2$  and  $(x_0, y_0) = (0, 0)$ . The 2D equilibrium is achieved thanks to a special bathymetry given by

$$b(x, y) = \frac{1}{g} \left( 30 - \frac{x^2 + y^2}{2} \right) - xy - C.$$

The solution of this potential flow is shown in figure 10, and the convergence rates computed at final time  $t_f = 1$  are presented in table 4, demonstrating the improvement brought about by the global flux formulation. Again, since the setup is stationary, superconvergence is observed. Moreover, the new method is even able to outperform FV-2.

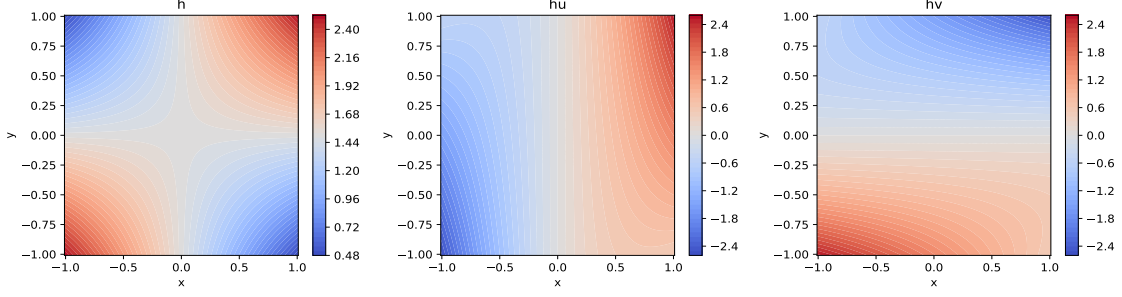


Figure 10: Shallow water system: potential flow. Reference solution of the conservative variables.

Table 4: Shallow water system: potential flow ( $t_f = 1$ ).  $L_2$  error and order of accuracy  $\tilde{n}$  for FV-1, FV-2 and GF.

| $N_x, N_y$ | $h$      |             | $hu$     |             | $hv$     |             |
|------------|----------|-------------|----------|-------------|----------|-------------|
|            | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ |
| FV-1       |          |             |          |             |          |             |
| 20         | 1.54E-02 | —           | 1.57E-01 | —           | 1.71E-01 | —           |
| 40         | 8.25E-03 | 0.89        | 1.08E-01 | 0.53        | 1.11E-01 | 0.62        |
| 80         | 4.30E-03 | 0.94        | 6.60E-02 | 0.71        | 6.43E-02 | 0.78        |
| 160        | 2.18E-03 | 0.97        | 3.68E-02 | 0.84        | 3.47E-02 | 0.88        |
| 320        | 1.10E-03 | 0.99        | 1.95E-02 | 0.91        | 1.80E-02 | 0.94        |
| FV-2       |          |             |          |             |          |             |
| 20         | 2.49E-04 | —           | 1.06E-03 | —           | 1.47E-03 | —           |
| 40         | 5.26E-05 | 2.24        | 2.61E-04 | 2.02        | 3.25E-04 | 2.17        |
| 80         | 1.09E-05 | 2.27        | 7.11E-05 | 1.87        | 8.17E-05 | 1.99        |
| 160        | 2.24E-06 | 2.28        | 1.86E-05 | 1.93        | 2.06E-05 | 1.98        |
| 320        | 4.81E-07 | 2.21        | 4.69E-06 | 1.98        | 5.17E-06 | 1.99        |
| GF         |          |             |          |             |          |             |
| 20         | 1.15E-04 | —           | 4.29E-04 | —           | 1.11E-03 | —           |
| 40         | 2.69E-05 | 2.09        | 1.01E-04 | 2.08        | 2.39E-04 | 2.21        |
| 80         | 6.49E-06 | 2.05        | 2.46E-05 | 2.03        | 5.50E-05 | 2.11        |
| 160        | 1.59E-06 | 2.02        | 6.08E-06 | 2.01        | 1.32E-05 | 2.05        |
| 320        | 3.95E-07 | 2.01        | 1.51E-06 | 2.00        | 3.24E-06 | 2.02        |

Table 5: Shallow water system: lake at rest ( $t_f = 0.1$ ).  $L_2$  error and order of accuracy  $\tilde{n}$  for FV-1 and FV-2 schemes with the novel GF.

| $N_x, N_y$ | $h$      |             | $hu$     |             | $hv$     |             |
|------------|----------|-------------|----------|-------------|----------|-------------|
|            | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ | $L_2$    | $\tilde{n}$ |
| FV-1       |          |             |          |             |          |             |
| 20         | 7.13E-03 | –           | 4.67E-02 | –           | 3.61E-02 | –           |
| 40         | 2.79E-03 | 1.35        | 2.42E-02 | 0.95        | 2.04E-02 | 0.82        |
| 80         | 1.20E-03 | 1.22        | 1.21E-02 | 1.00        | 1.10E-02 | 0.89        |
| 160        | 5.46E-04 | 1.12        | 5.99E-03 | 1.00        | 5.71E-03 | 0.94        |
| 320        | 2.60E-04 | 1.06        | 2.98E-03 | 1.00        | 2.91E-03 | 0.97        |
| FV-2       |          |             |          |             |          |             |
| 20         | 2.92E-03 | –           | 2.17E-03 | –           | 2.26E-03 | –           |
| 40         | 6.76E-04 | 2.10        | 3.65E-04 | 2.57        | 3.94E-04 | 2.51        |
| 80         | 1.60E-04 | 2.07        | 7.58E-05 | 2.26        | 8.10E-05 | 2.28        |
| 160        | 3.87E-05 | 2.04        | 1.70E-05 | 2.15        | 1.78E-05 | 2.18        |
| 320        | 9.50E-06 | 2.02        | 4.02E-06 | 2.08        | 4.13E-06 | 2.11        |
| GF         |          |             |          |             |          |             |
| 20         | 4.50E-16 | –           | 6.24E-15 | –           | 6.36E-15 | –           |
| 40         | 9.71E-16 | –           | 1.36E-14 | –           | 1.31E-14 | –           |
| 80         | 1.58E-15 | –           | 3.03E-14 | –           | 3.29E-14 | –           |
| 160        | 3.38E-15 | –           | 8.62E-14 | –           | 8.67E-14 | –           |
| 320        | 6.47E-15 | –           | 2.28E-13 | –           | 2.29E-13 | –           |

### 6.3.2. Lake at rest

In this section, we test the well-balanced property, proven in section 4.6, of the global flux method for lake at rest solutions of the shallow water system. The problem is set in a rectangular domain  $[0, 1] \times [0, 1]$  with periodic boundary conditions. The initial and exact solution is given by

$$h(x, y) = 1 - b(x, y), \quad u(x, y) = v(x, y) \equiv 0,$$

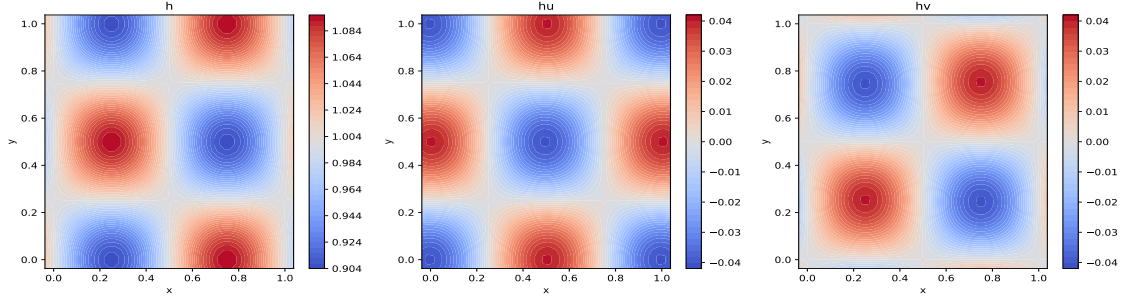
where the bathymetry is defined as

$$b(x, y) = 0.1 \sin(2\pi x) \cos(2\pi y).$$

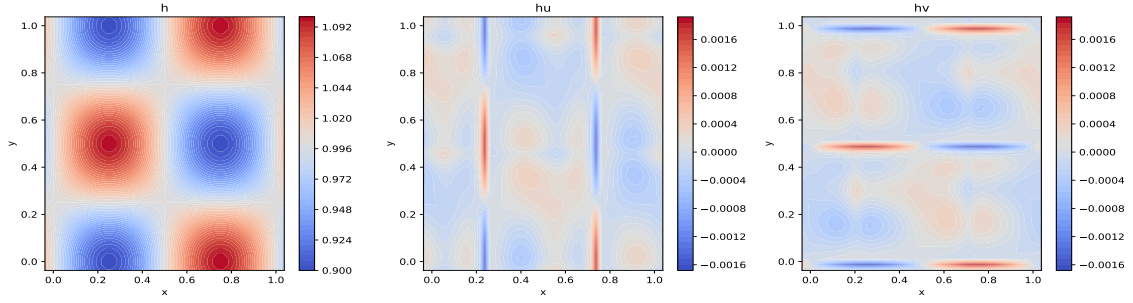
In table 5, a convergence study is presented at final time  $t_f = 0.1$ . As expected, thanks to the well-balanced property of the global flux method, the GF is able to achieve machine precision errors. The standard FV-1 and FV-2 methods show only the classical first and second order convergence slopes. In figure 11, we present the comparison between the well-balanced GF method and the non-well-balanced FV-1 and FV-2 methods.

### 6.3.3. 2D supercritical equilibria

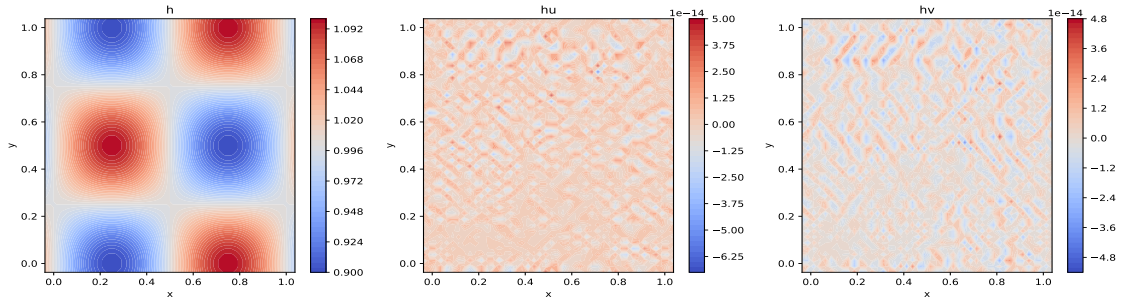
We consider two fully multi-dimensional steady states of the shallow water system, characterized by constant momentum in supercritical regimes. However, contrary to the one-dimensional version of such equilibria [23, 21], no exact solution is known for the simulations presented in this section. For this reason, the numerical results obtained through the three schemes will be compared qualitatively. The problems are simulated on a rectangular domain



(a) FV-1



(b) FV-2



(c) GF

Figure 11: Shallow water system: lake at rest. Numerical results for the lake at rest solution on the coarse mesh  $40 \times 40$  at final time  $t_f = 0.1$  obtained with FV-1, FV-2 and GF.

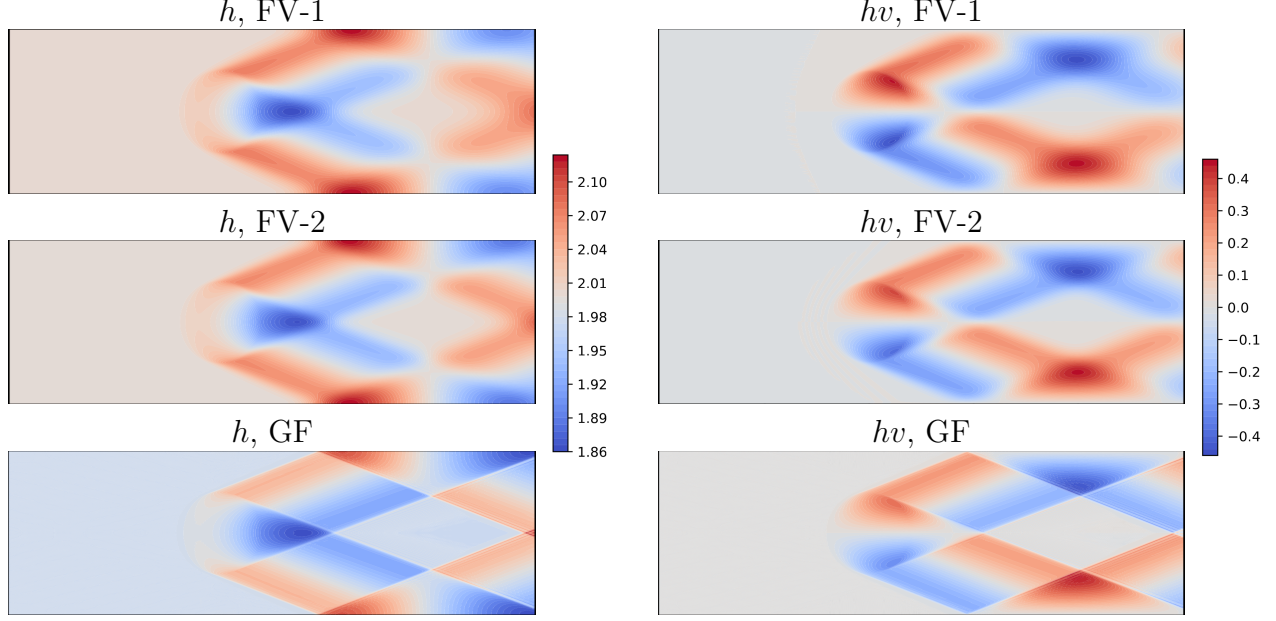


Figure 12: Shallow water system: 2D supercritical equilibria. Numerical results obtained with FV-1, FV-2 and GF to steady state for  $N_x = N_y = 450$ .

$[0, 25] \times [0, 8]$ , and are made fully multi-dimensional by employing a 2D bathymetry that is a function of both  $x$  and  $y$ , given by

$$b(x, y) = \begin{cases} \frac{1}{5} \left( 1 - \left( \frac{r(x, y)}{2} \right)^2 \right), & \text{where } r(x, y) < 2 \\ 0, & \text{elsewhere} \end{cases}$$

with  $r(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$  and  $(x_0, y_0) = (10, 4)$ . The initial conditions of the first problem are given by

$$h(x, y, 0) = 2 - b(x, y), \quad q_x(x, y, 0) = 24, \quad q_y(x, y, 0) = 0.$$

Inlet boundary conditions (equal to the initial conditions) are imposed on the left boundary of the domain, and outlet (homogeneous Neumann) on the right. Top and bottom of the domain are periodic boundaries. In figure 12, we present the numerical solutions for the conservative variables when the numerical steady state is reached (time residual close to machine precision). All simulations are performed on a mesh of  $450 \times 450$  elements. GF is able to capture and resolve sharply the many shocks appearing behind the bathymetry bump. Although the mesh resolution for this case is quite fine, FV-1 still presents a highly diffused result. An improvement is experienced when using the linear reconstruction for FV-2, although all the waves still appear as smooth transitions. However, the results of the classical methods still remain significantly inferior to those of GF, which captures all waves sharply in much fewer cells.

As a second problem, and in order to test the robustness of the method, we also consider a crooked supercritical equilibrium with the same bathymetry but different initial conditions:

$$h(x, y, 0) = 2 - b(x, y), \quad q_x(x, y, 0) = 24, \quad q_y(x, y, 0) = 4\pi.$$

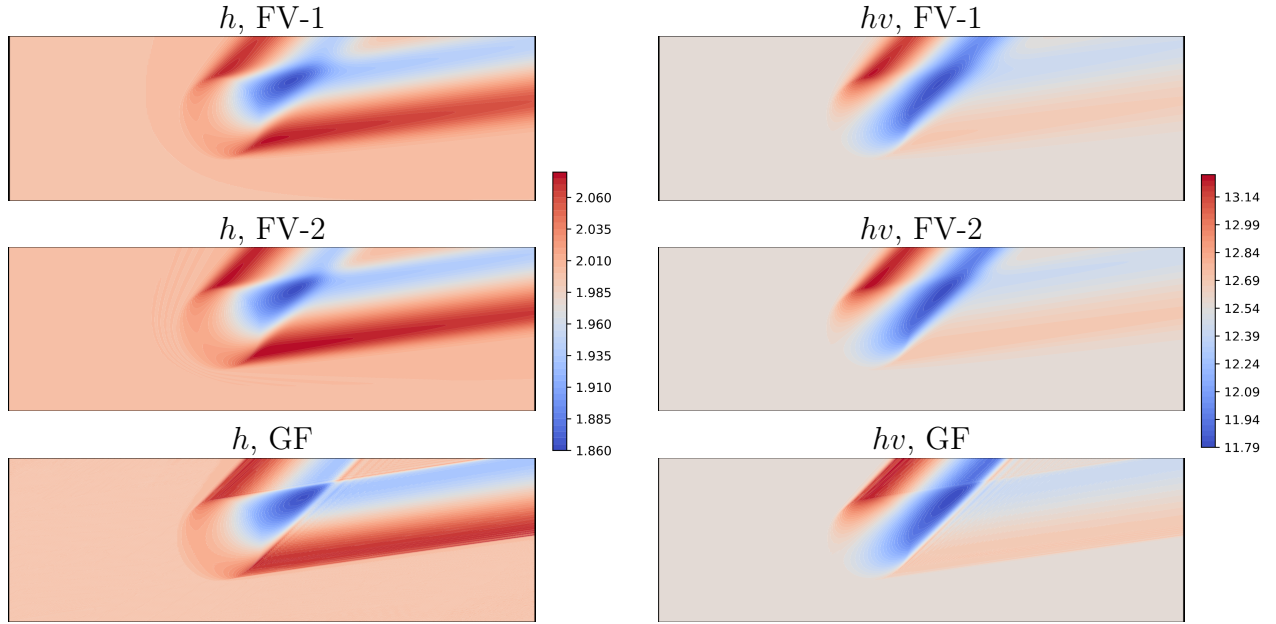


Figure 13: Shallow water system: 2D crooked supercritical equilibria. Numerical results obtained with FV-1, FV-2 and GF to steady state for  $N_x = N_y = 450$ .

In this case, left and bottom boundaries are inlet boundaries, while right and top are outlets. It should be noticed, that this test case is even more challenging than the one shown before since no part of the fluid is aligned with the background Cartesian mesh. In figure 13, the results obtained for the conservative variables are presented, where the same conclusion about the quality of the result of the GF method can be drawn as for the previous test. All the physical features of the equilibria are well captured, while they are significantly more diffused by the FV-1 and FV-2 methods.

## 7. Conclusions and perspectives

In this work, we have presented a new way to derive finite volume methods for nonlinear multi-dimensional hyperbolic systems, which is based on the global flux approach (9), introduced in [11]. It is a general way to obtain stationarity preserving schemes for nonlinear problems. Besides its generality, the method is also able to achieve superconvergence on steady problems. Despite a focus on stationary states during its design, we observe highly accurate solutions for unsteady multi-dimensional problems, outperforming standard first and even second-order finite volume methods.

This work opens the way to several future developments. In particular, the extension of the finite volume formulation to high order methods by using high-degree polynomial reconstruction techniques like WENO [29] is a natural next step, following the work on the 1D global flux WENO approach introduced in [23]. Moreover, the first order finite volume method can also be seen as the starting point to develop a new family of multi-dimensional high order discontinuous Galerkin methods based on the global flux formulation. More investigations will also be dedicated to the extension of the method to deal with mathematical models characterized by curl-free solutions, like the Maxwell equations. Extending,

for instance, the observation that stationarity preserving methods are also low Mach number compliant, theoretical work will include further analysis of the method in unsteady situations.

## References

- [1] R. Abgrall and M. Ricchiuto. High order methods for CFD. In Rene de Borst Erwin Stein and T. J.R. Hughes, editors, *Encyclopedia of Computational Mechanics, Second Edition*. John Wiley and Sons, 2017.
- [2] R. Abgrall and M. Ricchiuto. *Hyperbolic Balance Laws: Residual Distribution, Local and Global Fluxes*, pages 177–222. Springer Nature Singapore, Singapore, 2022.
- [3] E. Audusse, François Bouchut, Marie-Odile Bristeau, Rupert Klein, and Benoit Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal on Scientific Computing*, 25(6):2050–2065, 2004.
- [4] E. Audusse, Minh Hieu Do, Pascal Omnes, and Yohan Penel. Analysis of modified Godunov type schemes for the two-dimensional linear wave equation with Coriolis source term on cartesian meshes. *Journal of Computational Physics*, 373:91–129, 2018.
- [5] E. Audusse, V. Dubos, N. Gaveau, and Y. Penel. Energy-stable and linearly well-balanced numerical schemes for the nonlinear shallow water equations with the coriolis force. *SIAM Journal on Scientific Computing*, 47(1):A1–A23, 2025.
- [6] D.S. Balsara. Multidimensional HLLE Riemann solver: Application to Euler and magnetohydrodynamic flows. *J. Comput. Phys.*, 229:1970–1993, 2010.
- [7] W. Barsukow. *Low Mach number finite volume methods for the acoustic and Euler equations*. Doctoral thesis, University of Wuerzburg, 2018.
- [8] W. Barsukow. Stationarity preserving schemes for multi-dimensional linear systems. *Mathematics of Computation*, 88(318):1621–1645, 2019.
- [9] W. Barsukow. Truly multi-dimensional all-speed schemes for the euler equations on cartesian grids. *Journal of Computational Physics*, 435:110216, 2021.
- [10] W. Barsukow, Raphael Loubère, and Pierre-Henri Maire. A node-conservative vorticity preserving finite volume method for linear acoustics on unstructured grids. *Math. Comp.*, 94:2299–2343, 2025.
- [11] W. Barsukow, M. Ricchiuto, and D. Torlo. Structure preserving nodal continuous finite elements via global flux quadrature. *Numerical Methods for Partial Differential Equations*, 41(1):e23167, 2025.
- [12] C. Berthon and C. Chalons. A fully well-balanced, positive and entropy-satisfying Godunov-type method for the shallow-water equations. *Mathematics of Computation*, 85(299):1281–1307, 2016.

- [13] W. Boscheri and M. Dumbser. Arbitrary-Lagrangian-Eulerian one-step WENO finite volume schemes on unstructured triangular meshes. *Communications in Computational Physics*, 14(5):1174–1206, 2013.
- [14] F. Bouchut, J. Le Sommer, and V. Zeitlin. Frontal geostrophic adjustment and nonlinear wave phenomena in one-dimensional rotating shallow water. Part 2. high-resolution numerical simulations. *Journal of Fluid Mechanics*, 514:35–63, 2004.
- [15] David L Brown. Performance of under-resolved two-dimensional incompressible flow simulations. *Journal of Computational Physics*, 122(1):165–183, 1995.
- [16] V. Caselles, R. Donat, and G. Haro. Flux-gradient and source-term balancing for certain high resolution shock-capturing schemes. *Computers & Fluids*, 38(1):16–36, 2009.
- [17] M.J. Castro and C. Parés. Well-balanced high-order finite volume methods for systems of balance laws. *Journal of Scientific Computing*, 82(2):48, 2020.
- [18] Y. Cheng, A. Chertock, M. Herty, A. Kurganov, and T. Wu. A new approach for designing moving-water equilibria preserving schemes for the shallow water equations. *Journal of Scientific Computing*, 80(1):538–554, 2019.
- [19] A. Chertock, S. Cui, A. Kurganov, Ş.N. Özcan, and E. Tadmor. Well-balanced schemes for the Euler equations with gravitation: Conservative formulation using global fluxes. *Journal of Computational Physics*, 358:36–52, 2018.
- [20] A. Chertock, A. Kurganov, Xin Liu, Yongle Liu, and T. Wu. Well-balancing via flux globalization: Applications to shallow water equations with wet/dry fronts. *Journal of Scientific Computing*, 90(1):1–21, 2022.
- [21] M. Ciallella, L. Micalizzi, V. Michel-Dansac, P. Öffner, and D. Torlo. A high-order, fully well-balanced, unconditionally positivity-preserving finite volume framework for flood simulations. *GEM-International Journal on Geomathematics*, 16(1):1–33, 2025.
- [22] M. Ciallella, L. Micalizzi, P. Öffner, and D. Torlo. An arbitrary high order and positivity preserving method for the shallow water equations. *Computers & Fluids*, 247:105630, 2022.
- [23] M. Ciallella, D. Torlo, and M. Ricchiuto. Arbitrary high order WENO finite volume scheme with flux globalization for moving equilibria preservation. *Journal of Scientific Computing*, 96(2):53, 2023.
- [24] H. Deconinck and M. Ricchiuto. Residual distribution schemes: Foundations and analysis. In *Encyclopedia of Computational Mechanics Second Edition*, pages 1–53. John Wiley & Sons, Ltd, 2017.
- [25] M. Dumbser D.S. Balsara and R. Abgrall. Multidimensional HLLC Riemann Solver for Unstructured Meshes - With Application to Euler and MHD Flows. *J. Comput. Phys.*, 261:172–208, 2014.

- [26] E. Gaburro, M. Ricchiuto, and M. Dumbser. On general and complete multidimensional riemann solvers for nonlinear systems of hyperbolic conservation laws, 2025. arXiv eprint, 2506.00207, math.na.
- [27] L.L. Gascón and J.M. Corberán. Construction of second-order TVD schemes for nonhomogeneous hyperbolic conservation laws. *Journal of Computational Physics*, 172(1):261–297, 2001.
- [28] R. Jeltsch and M. Torrilhon. On curl-preserving finite volume discretizations for shallow water equations. *BIT Numerical Mathematics*, 46:35–53, 2006.
- [29] G.-S. Jiang and C.-C. Wu. A high-order WENO finite difference scheme for the equations of ideal magnetohydrodynamics. *Journal of Computational Physics*, 150(2):561–594, 1999.
- [30] J. Jung and V. Perrier. Steady low mach number flows: identification of the spurious mode and filtering method. *J. Comput. Phys.*, 468:111462, 2022.
- [31] J. Jung and V. Perrier. Behavior of the discontinuous galerkin method for compressible flows at low mach number on triangles and tetrahedrons. *SIAM J. Sci. Comput.*, 46(1):A452–A482, 2024.
- [32] M. Kazolea, C. ParésMadroñal, and M. Ricchiuto. Approximate well-balanced WENO finite difference schemes using a global-flux quadrature method with multi-step ode integrator weights. *arXiv preprint arXiv:2501.06155*, 2025.
- [33] Giovanni L., R. Andrassy, W. Barsukow, J. Higl, P.V.F. Edelmann, and F.K. Röpke. Performance of high-order godunov-type methods in simulations of astrophysical low mach number flows. *Astronomy & Astrophysics*, 686:A34, 2024.
- [34] TB Lung and PL Roe. Toward a reduction of mesh imprinting. *International Journal for Numerical Methods in Fluids*, 76(7):450–470, 2014.
- [35] Y. Mantri, P. Öffner, and M. Ricchiuto. Fully well-balanced entropy controlled discontinuous galerkin spectral element method for shallow water flows: global flux quadrature and cell entropy correction. *Journal of Computational Physics*, 498:112673, 2024.
- [36] L. Micalizzi, M. Ricchiuto, and R. Abgrall. Novel well-balanced continuous interior penalty stabilizations. *Journal of Scientific Computing*, 100(1):14, 2024.
- [37] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography or manning friction. *Journal of Computational Physics*, 335:115–154, 2017.
- [38] V. Michel-Dansac, C. Berthon, Stéphane Clain, and Françoise Foucher. A two-dimensional high-order well-balanced scheme for the shallow water equations with topography and manning friction. *Computers & Fluids*, 230:105152, 2021.

- [39] S. Mishra and E. Tadmor. Constraint preserving schemes using potential-based fluxes I. multidimensional transport equations. *Communications in Computational Physics*, 9(3):688–710, 2011.
- [40] Keith William Morton and Philip L Roe. Vorticity-preserving lax–wendroff-type schemes for the system wave equation. *SIAM Journal on Scientific Computing*, 23(1):170–192, 2001.
- [41] H. Nessyahu and E. Tadmor. Non-oscillatory central differencing for hyperbolic conservation laws. *Journal of computational physics*, 87(2):408–463, 1990.
- [42] M. Ricchiuto. *Contributions to the development of residual discretizations for hyperbolic conservation laws with application to shallow water flows*. Habilitation à diriger des recherches, Université Sciences et Technologies - Bordeaux I, December 2011. [pdf].
- [43] M. Ricchiuto and A. Bollermann. Stabilized residual distribution for shallow water simulations. *Journal of Computational Physics*, 228(4):1071–1115, 2009.
- [44] M. Ricchiuto and D. Torlo. Analytical travelling vortex solutions of hyperbolic equations for validating very high order schemes. *arXiv preprint arXiv:2109.10183*, 2021.
- [45] D. Sidilkover. Factorizable schemes for the equations of fluid flow. *Applied numerical mathematics*, 41(3):423–436, 2002.
- [46] J. Zhang, Y. Xia, and Y. Xu. Well-balanced discontinuous galerkin method with flux globalization for rotating shallow water equations. *Journal of Computational Physics*, page 114094, 2025.