Local subcell monolithic DG/FV convex property preserving scheme on unstructured grids and entropy consideration

François Vilar

Local subcell monolithic DG/FV convex property preserving scheme on unstructured grids and entropy consideration

François Vilar[a]

[a]*IMAG, Univ Montpellier, CNRS, Montpellier, France*

Abstract

This article aims at presenting a new local subcell monolithic Discontinuous-Galerkin/Finite-Volume (DG/FV) convex property preserving scheme solving system of conservation laws on 2D unstructured grids. This is known that DG method needs some sort of nonlinear limiting to avoid spurious oscillations or nonlinear instabilities which may lead to the crash of the code. The main idea motivating the present work is to improve the robustness of DG schemes, while preserving as much as possible its high accuracy and very precise subcell resolution. To do so, a convex blending of high-order DG and first-order FV scheme will be locally performed, at the subcell scale, where it is needed. To this end, by means of the theory developed in [58, 59], we first recall that it is possible to rewrite DG scheme as a subcell FV scheme on a subgrid provided with some specific numerical fluxes referred to as DG reconstructed fluxes. Then, the monolithic DG/FV method will be defined as following: each face of each subcell will be assigned with two fluxes, a 1st-order FV one and a high-order reconstructed one, that in the end will be blended in a convex way. The goal is then to determine, through analysis, optimal blending coefficients to achieve the desire properties. Numerical results on various type problems will be presented to assess the very good performance of the design method.

A particular emphasis will be put on entropy consideration. By means of this subcell monolithic framework, we will attempt to address the following questions: is this possible through this monolithic framework to ensure any entropy stability? what do we mean by entropy stability? What is the cost of such constraints? Is this absolutely needed while aiming for high-order accuracy?

Keywords: Structure-preserving scheme, subcell monolithic scheme, entropy stability, arbitrary high-order, DG subcell FV formulation, positivity-preserving scheme, hyperbolic conservation laws

Email address: `francois.vilar@umontpellier.fr` (François Vilar)

1. Introduction

This paper is concerned with solving system of conservation laws, and it is well known that hyperbolic partial differential equations frequently lead to discontinuous weak solutions within a finite time frame, posing significant challenges for numerical simulations. These challenges revolve around handling discontinuities, ensuring accuracy, and maintaining solutions within an admissibility set, such as guaranteeing positive density and internal energy in gas dynamics. Addressing these constraints simultaneously is particularly difficult because they often conflict with one another, demanding sophisticated approaches in the design of numerical methods for hyperbolic problems. A large number of numerical scheme have been developed these past fifty years to achieve such goal, and one which has particularly stood out is the Discontinuous Galerkin (DG) method. This scheme, initially introduced by Reed and Hill for neutron transport [46], has become one of the most widely used numerical schemes, particularly in computational fluid dynamics. Significant advancements by Cockburn and Shu in a series of seminal papers, see for instance [11] and the references within, have propelled DG methods forward. These methods theoretically allow for achieving any arbitrary order of accuracy while maintaining a compact stencil, and they exhibit desirable properties such as L_2 stability and hp-adaptivity. The DG scheme is renowned for its high accuracy and precise subcell resolution, even demonstrating superconvergence in some cases. However, robustness is a critical concern alongside accuracy. High-order DG schemes are known to produce spurious oscillations in the presence of discontinuities and possibly non-physical solutions (e.g., negative density or pressure in gas dynamics), which may lead to nonlinear instability or code crashes. Therefore, stabilization techniques are essential. This fundamental issue has been extensively tackled in the past. There is thus a vast literature on that topic, among which [2, 4, 33, 64, 29, 35, 41]. For a lot more detailed description of the state of the art limiters, we refer to [58] and the references within.

These past fifteen years, great progresses have been made in the direction of stabilizing and improving the robustness of high-order DG. And to do so, two main properties were under study: convex set preserving also referred to as Invariant Domain Preserving (IDP) and high-order entropy stability. In the former, the goal is to ensure that the numerical solution remains, at all time, in a convex admissible set. This property particularly permits to guarantee global maximum principle for Scalar Conservation Laws (SCL) and for instance positivity of the density and pressure in the Euler system case. The articles on this subject have flourished in recent years. Although not exhaustive, it worth mentioning [66, 67] and [5] where some polynomial limiters have been developed to ensured such property, as well as [63] where a new framework, refereed to as geometric quasi-linearization, has been introduced to also develop bound-preserving numerical methods. Another wide family of schemes also concerned with this convex property preserving issue is the one gathering Flux-Corrected Transport (FCT) techniques, Algebraic Flux Corrections (AFC) and convex limiting schemes, see for instance [3, 65, 20, 43, 19, 21, 36, 24, 38]. Most of these aforementioned methods share a similar philosophy, meaning blending high-order and low-order fluxes, operators or schemes to ensure the preservation of convex properties or more generally to be IDP.

Now, regarding high-order entropy stable schemes, a new class of numerical methods has recently extensively grown in popularity, see for instance [16, 17, 6, 18, 8, 7, 15]. Those schemes, generally referred to as entropy conservative/stable DG Spectral Element Method (DGSEM), were initially developed in the context of finite difference Summation-By-Parts (SBP) operators and Simultaneous Approximation Term (SAT) by T. Fisher and M. Carpenter in their seminal paper [16]. In the context of DG, they rely on the use of particular quadrature rules inducing a mass lumping type

diagonalization of the mass matrix and specific collocation of the flux in order to exhibit discrete SBP properties. Then, a substitution of the flux collocation values by a combination of entropy conservative numerical fluxes ensure an entropy conservation or stability while preserving the high-order accuracy. While some SBP operators and entropy stable DGSEM scheme have been successfully extended to simplex meshes, see for instance [26, 8], this family of numerical methods are generally restrained to one-dimension in space or tensor-product multi-dimensional grids.

A third direction which has been particularly embraced this past decade and has shown some of the most promising developments is subcell techniques. Here, the idea is to subdivide the bad cells, and to adopt a special procedure with the hope of curing the negative aspects of the original scheme. Some examples of this strategy can be found in [28, 53, 13]. For example, in [28], the authors use a convex combination between high-order DG schemes and first-order Finite-Volumes (FV) on a subgrid, allowing them to retain the very high accurate resolution of DG in smooth areas and ensuring the scheme's robustness in the presence of shocks. Similarly, in [53, 13], after having detected the troubled zones, cells are then subdivided into subcells and a robust first-order FV scheme, or alternatively other robust scheme (second-order TVD FV scheme, WENO scheme, ...), is performed on the subgrid in troubled cells. Let us emphasize that these subcell techniques offer several advantages. They preserve the high accuracy of DG schemes in smooth regions by applying corrections only where necessary. This local modification approach ensures that the majority of the grid remains unaffected by the stabilization process, maintaining computational efficiency and accuracy. Let us however emphasize that if a correction is used at the subgrid level, all the subcells contained in a bad cell are generally impacted. Since one of the main advantage of high-order scheme is to be able to use coarse grids while still being very precise, one can see that there is a waste of information here, as well as unnecessary computational effort made. This is particularly the case in the vicinity of discontinuities since the polynomials are globally modified. This problem was addressed in the one-dimensional case in [58] and in the two-dimensional unstructured case in [59]. This new technique relies on the reformulation of DG schemes as a FV-like scheme defined on a subgrid, through the definition of particular fluxes referred to as reconstructed fluxes. Then, after computing a DG candidate solution and check if this solution is admissible, one returns if needed to the previous time step and correct locally, at the subcell scale, the numerical solution. In the subcells where the solution was detected as bad, one substitutes the DG reconstructed flux on the subcell boundaries by a robust first-order numerical flux. And for subcells detected as admissible, one keeps the high-order reconstructed flux which allows to retain conservativity as well as the very high accurate resolution of DG schemes. Consequently, only the solution inside troubled subcells and their first neighbors will have to be recomputed. Elsewhere, the solution remains unchanged. This correction procedure is then extremely local, and has proved in different contexts its high capability to ensure a stable and robust behavior while maintaining the very high accuracy of DG schemes, see [58, 22, 59, 23].

Finally, in recent years the interest in combining these three family of schemes and techniques, namely convex limiting methods, high-order entropy stable schemes and subcell techniques has grown tremendously. Have therefore emerged new methods, as [37, 45, 50, 49, 48, 42], which combine, at the subcell scale, high-order and low-order schemes to ensure different properties on the numerical solution, while trying to preserve accuracy. For instance in [48], the authors develop a subcell monolithic scheme blending high-order DGSEM based on Gauss-Lobatto quadrature points and a first-order FV scheme. Different conditions on the blending coefficients, corresponding to different properties as positivity or local maximum principle for the concern of spurious oscillations, are

presented. Similarly, in [42], a monolithic Gauss-Lobatto DGSEM and FV scheme is also presented but along with a particular blending procedure, based of the resolution of a continuous Knapsack optimization problem, enabling the preservation of the high-order accuracy of the scheme while ensuring a cell entropy inequality. Let us underline that the numerical scheme presented in the present article falls also in this category. Indeed, the aim of this paper is to introduced a new monolithic scheme in which DG and first-order FV methods will be blended, locally at the subcell scale, to ensure any convex property as well as different entropy stabilities, while trying to preserve as much as possible the high accuracy of DG schemes. Let us emphasize that the monolithic schemes presented in [48, 42], because being based on Gauss-Lobatto solution representation and flux collocation, are limited to one-dimension in space, or by tensor-product extension to multi-dimensions on Cartesian grids. And because the theory developed in [59], namely reformulating DG scheme into a FV-like one, is very general in the sens that it can be extended to any dimension and any type of grids and cell subdivision, we aim here at presenting a monolithic scheme applicable to generic polytopal meshes. Let us emphasize however that, if the whole theory and scheme are indeed developed on any type of grid, numerical results are only shown on triangular meshes. The practical implementation of those schemes on generic polygonal grids is still an ongoing project.

Now, to present this local subcell monolithic DG/FV scheme, the remainder of this paper is organized as follows: we recall in Section 2 how unstructured grid DG scheme can be reinterpreted as a subgrid FV-like scheme, through the definition of particular fluxes that we referred to as reconstructed fluxes. While this part mainly relies on Section 2 of our previous article, [59], this theoretical analysis will be taken further here as the case where the number of subcells does not fit the dimension of the functional space will be addressed. Then, the local subcell monolithic DG/FV scheme will be introduced in Section 3. Practically, each face of each subcell will be assigned two fluxes, one reconstructed flux giving the equivalency with a high-order DG scheme and one first-order FV numerical flux. These two fluxes will then be blended in a convex manner through a blending coefficient between zero and one. The goal is now to determine, through analysis, the optimal coefficients to reach the desired properties while trying to maintain as much as possible the high accuracy of the scheme. Following this, we present in Section 4 different definitions for the blending coefficients to reach different types of entropy stability, meaning from a discrete subcell entropy inequality for any entropy to a semi-discrete cell entropy inequality for a given entropy. Only the latter one will proved to allow high-order accuracy preservation. Numerical results and a preliminary conclusion on those entropy stabilities will be given at the end of this section. Finally, section 5 is devoted to the definition of the blending coefficients ensuring different maximum principles, and by this make the monolithic scheme IDP. These theoretical parts as well as the different numerical results will be presented for both SCL and the Euler compressible gas dynamics system.

2. DG scheme reformulation

This section is devoted to the demonstration of the equivalency between DG schemes and a FV-like method on a subgrid provided the definition of particular fluxes. This theoretical part mainly relies on Section 2 of our previous article, [59]. Consequently, only the essential ingredients of such reformulation will be recalled at this time. Let us yet note that this theoretical analysis will be taken further here than in [59], as the case where the number of subcells does not fit the dimension of the functional space will be addressed. To remain as simple as possible, two-dimensional SCL will be considered in this section. The system extension is perfectly straightforward. Let then $u = u(\boldsymbol{x}, t)$, for $\boldsymbol{x} \in \omega \subset \mathbb{R}^2$ and $t \in [0, T]$, be the solution of the following problem

$$\begin{cases} \partial_t\, u(\boldsymbol{x}, t) + \nabla_x \cdot \boldsymbol{F}\left(u(\boldsymbol{x}, t)\right) = 0, & (\boldsymbol{x}, t) \in \omega \times [0, T], \quad (1a) \\ u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), & \boldsymbol{x} \in \omega, \quad (1b) \end{cases}$$

where u_0 is the initial data and $\boldsymbol{F}(u)$ the flux function. For the subsequent discretization, let us introduce the following notation. $\{\omega_c\}_c$ would be a generic partition of the domain ω into non-overlapping cells, with $|\omega_c|$ being the size of ω_c. We also partition the time domain in intermediate times $(t^n)_n$ with $\Delta t^n = t^{n+1} - t^n$ the n^{th} time step. In the DG frame, the numerical solution is considered piecewise polynomial over the domain, and hence developed on each cell onto $\mathbb{P}^k(\omega_c)$, the set of polynomials of degree up to k defined on cell ω_c. This space approximation theoretically leads to a $(k+1)^{th}$ space order accurate scheme. Let $u_h^c = \sum_{m=1}^{N_k} u_m^c(t)\, \sigma_m^c(\boldsymbol{x})$ be the restriction of u_h, the piecewise polynomial approximation of the solution u, over the cell ω_c, where the u_m^c are the N_k successive components of u_h^c over the polynomial basis $\{\sigma_m^c\}_m$. We recall that in the two-dimensional case $N_k = \frac{(k+1)(k+2)}{2}$. The coefficients u_m^c are the solution moments to be computed. To this end, by means of the weak formulation of equation (1a) on ω_c, restricting the solution functional space and the test function space to $\mathbb{P}^k(\omega_c)$ and then substituting the solution u by its approximated polynomial counterpart u_h^c, one gets

$$\int_{\omega_c} \frac{\partial\, u_h^c}{\partial t}\, \psi \,\mathrm{d}V = \int_{\omega_c} \boldsymbol{F}(u_h^c) \cdot \nabla_x \psi \,\mathrm{d}V - \int_{\partial\omega_c} \psi\, \mathcal{F}_n \,\mathrm{d}S, \qquad \forall\, \psi \in \mathbb{P}^k(\omega_c). \tag{2}$$

The DG numerical solution u_h^c is then the unique polynomial function defined in $\mathbb{P}^k(\omega_c)$ satisfying equation (2) for all function $\psi \in \mathbb{P}^k(\omega_c)$. In (2), the numerical flux function $\mathcal{F}_n$, in addition to ensure the scheme conservation, is the cornerstone of any FV or DG scheme regarding fundamental considerations as stability, positivity and entropy among others. In this context, this numerical flux is defined as a function of the two states on the left and right of each interface, $\mathcal{F}_n = \mathcal{F}\left(u^-, u^+, \boldsymbol{n}\right)$, with $u^- = \lim\limits_{\epsilon \to 0^+} u_h^c(\boldsymbol{x}_i - \epsilon\, \boldsymbol{n},\, t)$ and $u^+ = \lim\limits_{\epsilon \to 0^+} u_h^v(\boldsymbol{x}_i + \epsilon\, \boldsymbol{n},\, t)$, where ω_v is a face neighboring cell of ω_c, while $\boldsymbol{x}_i$ and $\boldsymbol{n}$ respectively stand for a point and the unit outward normal of the separating interface. From now on, we refer to the set containing the face neighboring cells of ω_c as $\mathcal{V}_c$. The numerical flux function is generally obtained through the resolution of an exact or approximated Riemann problem. In this context of SCL, we make use of the following well-known general numerical flux definition

$$\mathcal{F}(u^-, u^+, \boldsymbol{n}) = \frac{\left(\boldsymbol{F}(u^-) + \boldsymbol{F}(u^+)\right)}{2} \cdot \boldsymbol{n} - \frac{\gamma(u^-, u^+, \boldsymbol{n})}{2}\,(u^+ - u^-). \tag{3}$$

Under condition $\gamma(u^-, u^+, \boldsymbol{n}) \geq \max_{w \in I(u^-, u^+)} \left(|\boldsymbol{F}'(w) \cdot \boldsymbol{n}|\right)$, where $I(a, b) = [\min(a, b), \max(a, b)]$, the numerical flux (3) is nothing but an E-Flux, [44, 54]. A FV scheme relying on such a numerical

flux will be positivity-preserving and ensure a discrete entropy inequality for any entropy. Let us emphasize that for systems, the whole theory and scheme development that will follow can be easily extended to classical numerical fluxes, as HLL and HLL-C.

Now, taking in (2) the test function ψ among the polynomial basis functions leads to the following linear system allowing the calculation of the solution moments u_m^c

$$\sum_{m=1}^{N_k} \frac{\mathrm{d}\, u_m^c}{\mathrm{d}t} \int_{\omega_c} \sigma_m^c\, \sigma_p^c \,\mathrm{d}V = \int_{\omega_c} \boldsymbol{F}(u_h^c) \cdot \nabla_x \sigma_p^c \,\mathrm{d}V - \int_{\partial\omega_c} \sigma_p^c\, \mathcal{F}_n \,\mathrm{d}S, \qquad \forall\, p \in [\![1, N_k]\!]. \tag{4}$$

Terms $\int_{\omega_c} \boldsymbol{F}(u_h^c) \cdot \nabla_x \sigma_p^c \,\mathrm{d}V$ and $\int_{\partial\omega_c} \sigma_p^c\ \mathcal{F}_n \,\mathrm{d}S$ are respectively referred to as volume and surface integrals.

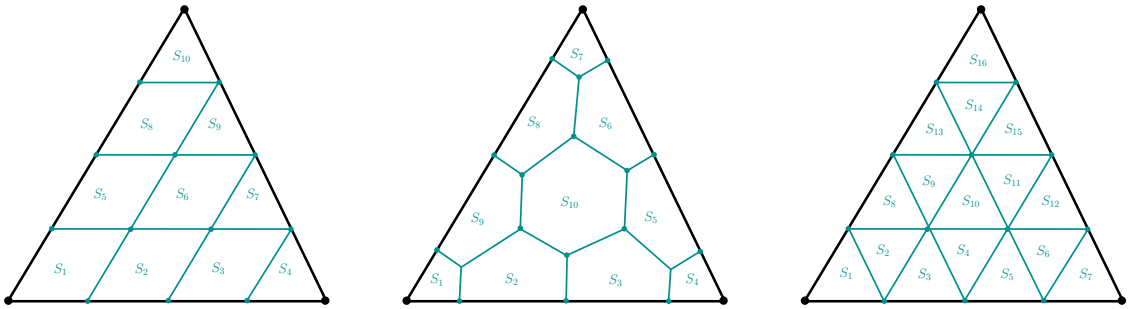
Remark 2.1. *Let us emphasize that these volume and surface integrals have to be computed in practice. Considering complex SCL with non-polynomial flux or non-linear systems as the Euler compressible gas dynamics one with non-convex equation of state for instance, exact integration may be difficult nay impossible. Generally, people either use quadrature rules, as originally introduced in [12], or a collocation of the flux, as it is done in nodal DG [25] or in DGSEM [17]. As demonstrated in [10], quadrature rules exact for polynomial of degree respectively $2\,k$ for volume integrals and $2\,k+1$ for surface ones have to be used to reach the excepted accuracy. In this article, such approach is chosen. In the remainder, $\oint$ will refer to quadrature approximated integration, while $\int$ holds for exact integration. Obviously, for polynomials of degree up to $2\,k$ and $2\,k+1$ respectively for volume and surface integrals, $\int$ and $\oint$ are indeed equivalent.*

Remark 2.2. *In [30], G.-S. Jiang and C.-W. Shu proved that DG schemes solving SCL do ensure a cell entropy inequality, for the square entropy $\eta(u) = \frac{1}{2}\,u^2$. Similarly in [27], this square stability analysis has been extended by S. Hou and X.-D. Liu to symmetric systems. However, those demonstrations rely on exact calculation of integrals, and thus does not applied if one uses quadrature rules. They are also limited to the square entropy.*

In (4), we identify $\int_{\omega_c} \sigma_m^c\, \sigma_p^c \,\mathrm{d}V = (M_c)_{mp}$ as the generic coefficient of the symmetric mass matrix $M_c \in \mathcal{M}_{N_k}$. The scheme (4) can then be reformulated in a compact matrix-vector form as

$$M_c \frac{\mathrm{d}\, U_c}{\mathrm{d}t} = \Phi_c, \tag{5}$$

with $(U_c)_m = u_m^c$ the solution vector filled with the polynomial moments, and where the so-called DG residuals $\Phi_c \in \mathbb{R}^{N_k}$ is defined as $(\Phi_c)_m = \oint_{\omega_c} \boldsymbol{F}(u_h^c) \cdot \nabla_x \sigma_m^c \,\mathrm{d}V - \oint_{\partial\omega_c} \sigma_m^c\ \mathcal{F}_n \,\mathrm{d}S$. Now, aiming at reformulating DG scheme (5) as a subgrid FV-like scheme, let us subdivide the mesh cells into subcells, similarly to what we did in [58, 59]. Let us emphasize that here we can relax the constraint requiring the number of subcells to match the dimension of the functional space. Hence, the subdivision can be chosen very freely. If N_k stands for the number of degrees of freedom in a given cell, let define N_s as the number of subcells in the cell. The only constraint we impose here is to have enough subcells not to be under-resolved, hence we impose $N_s \geq N_k$. In Figure 1, we present some easily generalizable subdivisions for triangle cells. Let us mention that in the first two, Figures 1(a) and 1(b), $N_s = N_k$. This is no more the case in the last one, Figure 1(c), as the number of subcells exceed by a lot the dimension of $\mathbb{P}^3$.



(a) Quad/tri: $N_s = N_k$ (b) Voronoi-type: $N_s = N_k$ (c) Tri: $N_s > N_k$

Figure 1: Examples of easily generalizable subdivisions for a triangular cell and a $\mathbb{P}^3$ DG scheme ($N_k = 10$)

Remark 2.3. *Let us emphasize that, while only triangular grids are considered for numerical applications, see Sections 4.4 and 5.3, the following demonstration as well as the local subcell monolithic DG/FV scheme presented in the remainder are not limited to this case. Any grid made of generic polygonal cells can be considered. Furthermore, apart from the coding aspects, the present analysis and monolithic scheme can also be directly extended to 3D geometries.*

That being said, let us consider a cell ω_c and its subdivision into N_s subcells S_m^c, for $m \in [\![1, N_s]\!]$. Then, we define the numerical solution subcell mean values, also referred to as submean values, as $\overline{u}_m^c = \frac{1}{|S_m^c|} \int_{S_m^c} u_h^c \, \mathrm{d}V$. To express DG scheme as a subgrid FV-like method, we want find the so-called reconstructed fluxes $\widehat{F_{mp}}$ such that

$$\frac{\mathrm{d}\,\overline{u}_m^c}{\mathrm{d}t} = -\frac{1}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \widehat{F_{mp}}. \tag{6}$$

In equation (6), $\mathcal{V}_m^c$ denotes the set of face neighboring subcells of S_m^c, while l_{mp} stands for the length of the interface f_{mp} between subcells S_m^c and S_p^v. Let us highlight that $S_p^v \in \mathcal{V}_m^c$ can either be inside cell ω_c or in one of its neighbors $\omega_v \in \mathcal{V}_c$. As in [59], we impose on the boundary of cell ω_c, so for $S_p^v \not\subset \omega_c$, that the reconstructed flux is nothing but the DG numerical flux, *i.e.* $l_{mp}\, \widehat{F_{mp}} = \oint_{f_{mp}} \mathcal{F}\,(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp})\, \mathrm{d}S$. As details of the proof have been given in [59], let us simply recall the final formula to compute the subcell interior faces reconstructed fluxes

$$\widehat{F_c} = -A_c^{\mathrm{t}}\, \mathcal{L}_c^{-1} \left(D_c\, P_c\, M_c^{-1}\, \Phi_c + B_c \right). \tag{7}$$

In (7), if N_f^c denotes the number of subcells' faces inside ω_c, meaning not belonging to $\partial\omega_c$, the vector $\widehat{F_c} \in \mathbb{R}^{N_f^c}$ then contains all the interior faces reconstructed fluxes weighted by the face length, *i.e.* $l_{mp}\, \widehat{F_{mp}}$. Matrix $A_c \in \mathcal{M}_{N_s \times N_f^c}$ stands for the adjacency matrix, $\mathcal{L}_c^{-1} \in \mathcal{M}_{N_s}$ the generalized inverse of the graph Laplacian matrix of the subdivision, $D_c = \mathrm{diag}(|S_1^c|, \ldots, |S_{N_k}^c|) \in \mathcal{M}_{N_s}$ the subcells volume matrix, $P_c \in \mathcal{M}_{N_s \times N_k}$ the projection matrix such that $(P_c)_{mp} = \frac{1}{|S_m^c|} \int_{S_m^c} \sigma_p^c \, \mathrm{d}V$ and $(B_c)_m = \oint_{\partial S_m^c \cap \partial\omega_c} \mathcal{F}_n \, \mathrm{d}S$ the cell boundary contribution. Definition of all these matrices can be found in [59]. Let us just recall that matrices A_c and $\mathcal{L}_c^{-1}$ only depends on the cell subdivision connectivity, while D_c, P_c and M_c depends on the chosen basis function and their values on the subcells. All those matrices can be computed initially, once and for all. The only time dependent quantities in (7) are Φ_c, the DG residual which is computed and available in any DG code, and term

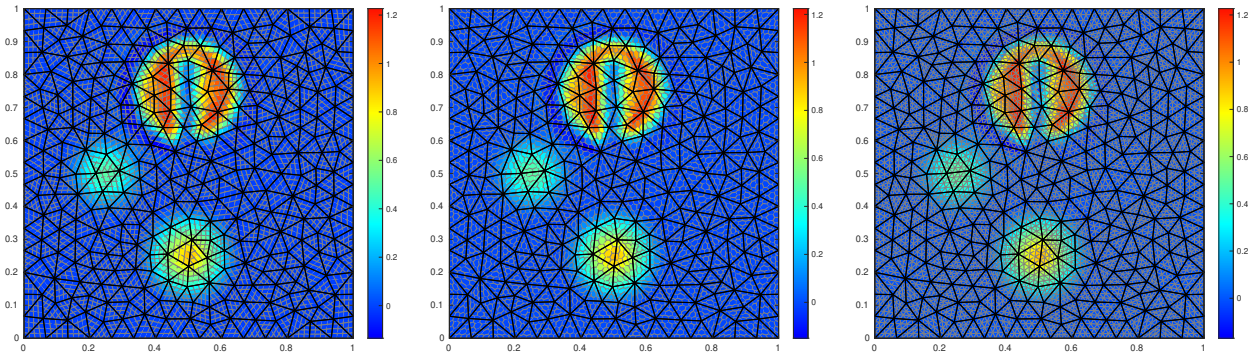
B_c which is required to close the linear system to solve and ensure that (7) is indeed the unique solution. Let us mention that different cell subdivisions will lead to different reconstructed flux values, but will still be equivalent to the same unique DG numerical solution, see Figure 2.

Remark 2.4. *To go from the polynomial representation of the solution to its submean values, we make use of the projection matrix P_c as $\overline{U}_c = P_c U_c$, where $\overline{U}_c \in \mathbb{R}^{N_s}$ is the vector containing all the subcell mean values in cell ω_c. Now, working with the piecewise constant representation of the numerical solution on the subcells through equation* (6)*, one still needs the polynomial representation of the solution in the computation of the DG residual. Then, to go from the submean values $\overline{u}_m^c$ to the polynomial moments u_m^c, we make use of the following least square procedure*

$$U_c = (P_c^t P_c)^{-1} P_c^t \overline{U}_c. \tag{8}$$

In the light of (8)*, we consider a subdivision to be admissible if matrix $P_c^t P_c$ is indeed invertible, which has been the case for all subdivisions we have studied. Let us emphasize that in the case where $N_s = N_k$, relation* (8) *reduces to $U_c = P_c^{-1} \overline{U}_c$.*

To make sure that, regardless the type of cell subdivision, the FV-like scheme (6) provided with the reconstructed fluxes definition (7) does indeed produce the DG numerical solution defined in equation (2), let us run the classical solid body rotation test case taken from [40]. To this end, we then consider (1a) with a divergence-free velocity field corresponding to a rigid rotation, defined by $\boldsymbol{F}(u, \boldsymbol{x}) = (\frac{1}{2} - y,\, x - \frac{1}{2})^{\mathrm{t}}\, u$. We apply this solid body rotation to an initial datum which includes both a plotted disk, a cone and a smooth hump. We run this test as a FV-like scheme (6) associated with definition (7), where the DG residual has been computed as for a $\mathbb{P}^3$ DG scheme. The three types of subdivision displayed in Figure 1 have been considered, see Figure 2.



(a) Quad/tri subdivision (b) Voronoi-type subdivision (c) Triangular subdivision

Figure 2: $\mathbb{P}^3$reconstructed flux FV schemes on 576 cells: subcells mean values

In the light of Figures 2 and 3, the three computations, involving three different types of cell subdivision, do produce the same numerical solution, which is nothing but the one that a $\mathbb{P}^3$ DG code would have produced. Figures 2 and 3, although demonstrating how accurate DG schemes are, also highlight the need of further limitation or correction to ensure an admissible behavior. Indeed, while the unique entropic weak solution is supposed to remained bounded by the minimum and maximum of the initial datum u_0 (the so-called maximum principle), the numerical solutions in Figure 3 clearly violate this principle. In the context of systems, as the Euler compressible gas

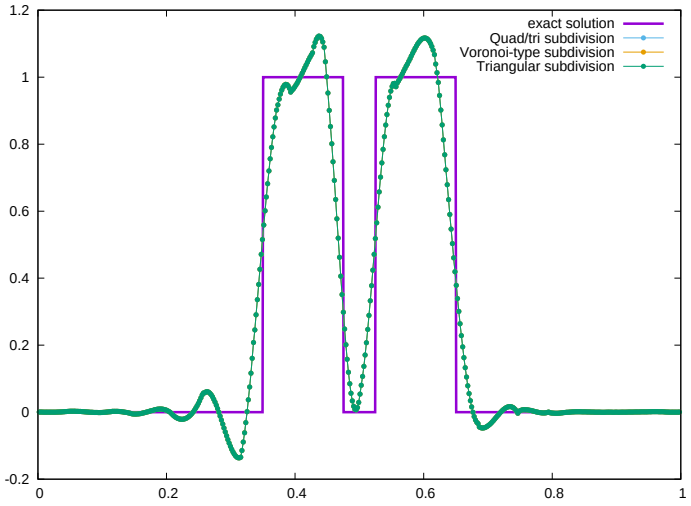


Figure 3: $\mathbb{P}^3$ reconstructed flux FV-type scheme on 576 cells: polynomial solution values on line $y = 0.75$

dynamics one for instance, this maximum principle translates into a positivity principle, where some quantities have to remain positive, as the density and internal energy in the aforementioned Euler case. The non-preservation of the positivity of the numerical solution is absolutely critical, as it can lead to the crash of the simulation code. In both Figures 2 and 3, one can furthermore clearly see the well-known Gibbs phenomenon, for which the approximation of a discontinuity through a high-order scheme will produce spurious oscillations. Those phenomena, along with the capture of non-entropic weak solutions, require some stabilization or correction techniques. This paper aims at presenting a local subcell monolithic DG/FV scheme, where DG scheme will be blended at the subcell scale with a first-order FV scheme, in order to combine the best of the two worlds, namely accuracy and robustness.

3. Local subcell monolithic DG/FV scheme

3.1. Blended fluxes and intermediate Riemann states

The previous reformulations of DG scheme into subcell FV-like scheme through the definition of reconstructed fluxes enable us to construct our local subcell monolithic DG/FV scheme. To this end, each face f_{mp} of each subcell S_m^c will be assigned two fluxes, one reconstructed flux $\widehat{F_{mp}}$ giving the equivalency with a high-order DG scheme and one first-order FV numerical flux $\mathcal{F}_{mp}^{\mathrm{FV}} = \mathcal{F}\left(\overline{u}_m^c, \overline{u}_p^v, \boldsymbol{n}_{mp}\right)$, where $\boldsymbol{n}_{mp}$ is outward unit normal of face f_{mp}. Then, these two fluxes will be blended in a convex manner through a blending coefficient $\theta_{mp} \in [0, 1]$ as in following

$$\widetilde{F_{mp}} = \mathcal{F}_{mp}^{\mathrm{FV}} + \theta_{mp} \underbrace{\left(\widehat{F_{mp}} - \mathcal{F}_{mp}^{\mathrm{FV}}\right)}_{\Delta F_{mp}}. \tag{9}$$

A blending coefficient set to 0 will lead to a first-order FV numerical flux, while a coefficient set at 1 will induce a high-order DG reconstructed flux. The local subcell monolithic DG/FV then writes as follows

$$\frac{\mathrm{d}\,\overline{u}_m^c}{\mathrm{d}t} = -\frac{1}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \widetilde{F_{mp}}. \tag{10}$$

The goal is now to determine, through analysis, the optimal coefficients to reach the desired properties while trying to maintain as much as possible the high accuracy of the scheme. To do so, we rewrite the monolithic scheme (10) as a Godunov-like scheme. But first, as only the semi-discrete version of the analysis and the monolithic scheme have been presented, we make use of SSP Runge-Kutta (RK) time integration method [51] to achieve high-accuracy in time. In the light of the fact that these multistage time integration methods write as convex combinations of first-order forward Euler scheme, the monolithic DG/FV scheme will be presented for the simple case of this latter time numerical scheme, for sake of simplicity. The semi-discrete scheme (10) provided with first-order forward Euler time integration writes

$$\overline{u}_m^{c,n+1} = \overline{u}_m^{c,n} - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \widetilde{F_{mp}}, \tag{11}$$

where all the quantities involved in the definition of the blended flux $\widetilde{F_{mp}}$ are taken at time level n (at the previous Runge-Kutta stage in a RK time integration). Defining $\gamma_{mp} = \gamma\left(\overline{u}_m^{c,n}, \overline{u}_p^{v,n}, \boldsymbol{n}_{mp}\right)$ and recalling that $\sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \boldsymbol{n}_{mp} = \boldsymbol{0}$, let us now rewrite $\overline{u}_m^{c,n+1}$ as a convex combination of quantities defined at the previous time step

$$\begin{aligned}
\overline{u}_m^{c,n+1} &= \overline{u}_m^{c,n} - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \widetilde{F_{mp}} \pm \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\, \overline{u}_m^{c,n} + \frac{\Delta t}{|S_m^c|} \boldsymbol{F}(\overline{u}_m^{c,n}) \cdot \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \boldsymbol{n}_{mp}, \\
&= \Big(1 - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\Big)\, \overline{u}_m^{c,n} + \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp} \Big(\overline{u}_m^{c,n} - \frac{\widetilde{F_{mp}} - \boldsymbol{F}(\overline{u}_m^{c,n}) \cdot \boldsymbol{n}_{mp}}{\gamma_{mp}}\Big).
\end{aligned}$$

Then, defining the left blended Riemann intermediate state $\widetilde{u_{mp}}^- = \overline{u}_m^{c,n} - \frac{\widetilde{F_{mp}} - \boldsymbol{F}(\overline{u}_m^{c,n}) \cdot \boldsymbol{n}_{mp}}{\gamma_{mp}}$, the previous expression can finally be recast into the following convex form

$$\overline{u}_m^{c,n+1} = \Big(1 - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\Big)\, \overline{u}_m^{c,n} + \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\, \widetilde{u_{mp}}^-. \tag{12}$$

Consequently, $\overline{u}_m^{c,n+1}$ does indeed write as a convex combination of previous time step quantities under the standard CFL condition in this subcell context

$$\Delta t \leq \frac{|S_m^c|}{\displaystyle\sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}}. \tag{13}$$

By mean of the numerical flux definition (3), the left blended Riemann intermediate state $\widetilde{u_{mp}}^-$ can be rewritten into the following form

$$\widetilde{u_{mp}}^- = \frac{\overline{u}_m^{c,n} + \overline{u}_p^{v,n}}{2} - \frac{\big(\boldsymbol{F}(\overline{u}_p^{v,n}) - \boldsymbol{F}(\overline{u}_m^{c,n})\big) \cdot \boldsymbol{n}_{mp}}{2\,\gamma_{mp}} - \theta_{mp} \frac{\Delta F_{mp}}{\gamma_{mp}} = u_{mp}^{*,\mathrm{FV}} - \theta_{mp} \frac{\Delta F_{mp}}{\gamma_{mp}}, \tag{14}$$

where $u_{mp}^{*,\mathrm{FV}}$ is nothing but the first-order FV Riemann intermediate state. It is straightforward to prove that, since $\gamma_{mp} \geq \max_{w \in I(\overline{u}_m^{c,n}, \overline{u}_p^{v,n})} \big(|\boldsymbol{F}'(w) \cdot \boldsymbol{n}_{mp}|\big)$, we have then $u_{mp}^{*,\mathrm{FV}} \in I(\overline{u}_m^{c,n}, \overline{u}_p^{v,n})$, see Appendix A.1.

For sake of clarity, let us specify conservativity relation. Obviously, we have that $\boldsymbol{n}_{pm} = -\boldsymbol{n}_{mp}$ as well as $\mathcal{F}^{\mathrm{FV}}_{pm} = -\mathcal{F}^{\mathrm{FV}}_{mp}$, $\widehat{F_{pm}} = -\widehat{F_{mp}}$ and $\widetilde{F_{pm}} = -\widetilde{F_{mp}}$, while $\theta_{pm} = \theta_{mp}$ and $u^{*,\mathrm{FV}}_{pm} = u^{*,\mathrm{FV}}_{mp}$. In the light of these relations, it is clear the right blended Riemann intermediate state $\widetilde{u_{mp}}^{+} := \widetilde{u_{pm}}^{-}$ hence writes $\widetilde{u_{mp}}^{+} = u^{*,\mathrm{FV}}_{mp} + \theta_{mp}\frac{\Delta F_{mp}}{\gamma_{mp}}$. It then appears that, where in first-order FV scheme we have only one Riemann intermediate state, here we have two, $\widetilde{u_{mp}}^{\pm} = u^{*,\mathrm{FV}}_{mp} \pm \theta_{mp}\frac{\Delta F_{mp}}{\gamma_{mp}}$, which both rely on the admissible first-order one $u^{*,\mathrm{FV}}_{mp}$. Consequently, introducing G, a convex admissible set where the solution has to remain in, if the numerical initial solution does lie in G, then it is always possible to find blending coefficients θ_{mp} to ensure that $\overline{u}^{c,n}_{m}$ remains in G during the whole calculation.

Remark 3.1. *To ensure the approximated solution to be in G at the initial time, the initialization has to be carried out by computing the subcell mean values and then use the projection matrix P_c to recover the cell polynomial representation, and not by a L_2 projection or an interpolation as it is generally done in DG schemes.*

As long as the first-order FV scheme used should achieve the desired properties, one can find blending coefficients for the high-order local subcell monolithic scheme to do as well. The different conditions on the blending coefficients will be formulated as inequalities. Thus, to combine several properties, one just have to take the minimum of the corresponding conditions.

Let us enlighten that it has been previously observed, [58, 59], that a stiff transition from first-order to a fully high-order scheme would produce more oscillatory solutions. It will be the case in 2D if a subcell will be assigned with first-order FV fluxes on some faces (corresponding to $\theta_{mp} = 0$) as well as fully high-order reconstructed fluxes (corresponding to $\theta_{mp} = 1$) on some other faces. In order to avoid such strong variation in fluxes accuracy, we will make use blending coefficients smoothers.

3.2. Blending coefficients smoothening

In our previous work [59], *a posteriori* blending of high-order reconstructed fluxes with first-order FV fluxes have been performed with arbitrary blending coefficients. And, in the context of non-linear problems, to avoid to yield too strong transition from high to low orders, a wider blending stencil with increasing coefficients (increasing according to the sequence $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$) was used. Here, in this *a priori* monolithic framework, we make use of a simple procedure to avoid very stiff order transition. Let us mention that we only make use of this coefficient smoother for solving non-linear problems, as it is no needed in the linear case. Furthermore, two different versions will be used, as a more constraining one will help in the context of non-convex fluxes SCL.

3.2.1. Coefficients smoothening n°1

Each subcell S^c_m will be given a blending coefficient θ^c_m defined as the average of its faces blending coefficients, as in follows

$$\theta^c_m = \frac{1}{\#\mathcal{V}^c_m} \sum_{S^v_p \in \mathcal{V}^c_m} \theta_{mp}. \tag{15}$$

Then, each subcell's face blending coefficient might be potentially reduced to the average of subcells' blending coefficients of every subcells sharing a node with face f_{mp}, as

$$\widetilde{\theta}_{mp} = \min\left(\theta_{mp}, \frac{1}{\#\mathcal{V}_{mp}} \sum_{S^v_q \in \mathcal{V}_{mp}} \theta^v_q\right).$$

Here, $\mathcal{V}_{mp}$ is the set containing all the subcells that share at least one node with face f_{mp}.

3.2.2. Coefficients smoothening n°2

This second smoother is nothing but the first one where the different averaged values are substituting by minimum values. Consequently, each subcell S_m^c will be given a blending coefficient θ_m^c defined as the minimum of its faces blending coefficients, as in follows

$$\theta_m^c = \min_{S_p^v \in \mathcal{V}_m^c} \theta_{mp}.$$

Then, each subcell's face blending coefficient might be potentially reduced by taking

$$\widetilde{\theta}_{mp} = \min\left(\theta_{mp}, \min_{S_q^v \in \mathcal{V}_{mp}} \theta_q^v\right).$$

Those two smoothening techniques have proved to improve the numerical results while preserving all the different properties presented in the next sections. While in Section 5 we focus on imposing positivity and local maximum principles to control spurious oscillations, which will prove to produce the best results, in the next section we first address the different questions regarding entropy stability.

4. Entropy stabilities

This section is devoted to entropy stability. By means of this local subcell monolithic DG/FV framework, we will attempt to address the following questions: Is it possible to find θ_{mp} the blending coefficients ensuring entropy stability? What do we mean by entropy stability? What is the cost of such constraints? Is this absolutely needed while aiming for high-order accuracy? To this end, we first introduce the definition of blending coefficients ensuring different type of entropy stabilities, while discussing the cost of such properties. Numerical results are then presented to confirm the developed theory, and to help us answer the stated queries.

For sake of simplicity, entropy stability will be addressed here in the simple case of SCL. Nonetheless, the extension to systems is perfectly straightforward. In the remainder, let then $\eta(u)$ be a strictly convex entropy, while $\boldsymbol{\phi}(u)$ be the associated entropy flux. $v(u) = \eta'(u)$ refers to the entropy variable, $\boldsymbol{\psi}(u) = v(u)\,\boldsymbol{F}(u) - \boldsymbol{\phi}(u)$ and $\boldsymbol{\Psi}(v) = \boldsymbol{\psi}\big(u(v)\big)$ to the entropy potential flux. Thanks to the entropy convexity, the mapping between u and v is indeed a diffeomorphism.

4.1. Discrete subcell entropy stability for any entropy $\Longrightarrow$ First-order

First, let us enforce, at the discrete level and subcell scale, an entropy inequality for any given entropy. To this end, we seek a blending coefficient θ_{mp} to make the blended flux $\widetilde{F_{mp}}$ an E-flux. Such flux does yield the desired property, as recalled in Appendix A.2. Let introduce $\widetilde{\gamma_{mp}}$ such that

$$\widetilde{F_{mp}} = \frac{\boldsymbol{F}\big(\overline{u}_m^{c,n}\big) + \boldsymbol{F}\big(\overline{u}_p^{v,n}\big)}{2} \boldsymbol{\cdot}\, \boldsymbol{n}_{mp} - \underbrace{\big(\gamma_{mp} - 2\,\theta_{mp}\,\Delta F_{mp}\big)}_{\widetilde{\gamma_{mp}}} \frac{\overline{u}_p^{v,n} - \overline{u}_m^{c,n}}{2}.$$

Then, to ensure a discrete entropy inequality, at the subcell level, for any entropy, it is sufficient to take θ_{mp} such that $\widetilde{\gamma_{mp}} \geq \max_{w \in I(\overline{u}_m^{c,n}, \overline{u}_p^{v,n})} \big(|\boldsymbol{F}'(w) \boldsymbol{\cdot}\, \boldsymbol{n}_{mp}|\big)$. This conditions leads to the following condition: if $\Delta F_{mp} \,.\, (\overline{u}_p^{v,n} - \overline{u}_m^{c,n}) > 0$ then

$$\theta_{mp} \leq \min\left(1, \frac{\big(\gamma_{mp} - \gamma_{\max}\big)\,\big(\overline{u}_p^{v,n} - \overline{u}_m^{c,n}\big)}{2\,\Delta F_{mp}}\right), \tag{16}$$

where $\gamma_{\max} := \max\limits_{w \in \mathrm{I}(\overline{u}_m^{c,n}, \overline{u}_p^{v,n})} \Big(|\boldsymbol{F}'(w) \boldsymbol{\cdot}\, \boldsymbol{n}_{mp}|\Big)$.

Remark 4.1. *While this particular choice does ensure the desired entropy property, it does also, as expected, lead to a* ***first-order*** *accurate scheme, as displayed in Table 1 and Figure 5. As stated in [56], an E-scheme is indeed first-order accurate, and in light of condition* (16)*, one can see for instance that* θ_{mp} *will be partially set to zero in the simple case of linear advection or if one uses global Lax-Friedrichs numerical flux.*

4.2. Semi-discrete subcell entropy stability for one given entropy $\Longrightarrow$ *Second-order*

Because condition (16) would lead to a first-order scheme, we may relaxed our expectations for sake of accuracy. Instead of a discrete entropy stability, for any entropy, one may aim for a semi-discrete entropy inequality, for a given entropy - entropy flux pair (η, $\boldsymbol{\phi}$). To this end, we make use of the following Tadmor two-point entropy conservation/dissipation condition, see [55, 56]

$$\widetilde{F_{mp}}\left(v\big(\overline{u}_p^{v,n}\big) - v\big(\overline{u}_m^{c,n}\big)\right) \leq \left(\boldsymbol{\psi}\big(\overline{u}_p^{v,n}\big) - \boldsymbol{\psi}\big(\overline{u}_m^{c,n}\big)\right) \boldsymbol{.}\, \boldsymbol{n}_{mp}.$$

As the first-order FV flux does ensure such inequality, see Appendix A.3, a sufficient condition for the blended flux to do as well is: if $\Delta F_{mp} \,.\, \left(v\big(\overline{u}_p^{v,n}\big) - v\big(\overline{u}_m^{c,n}\big)\right) > 0$ then

$$\theta_{mp} \leq \min\left(1,\ \frac{\left(\dfrac{\boldsymbol{\psi}\big(\overline{u}_p^{v,n}\big) - \boldsymbol{\psi}\big(\overline{u}_m^{c,n}\big)}{v\big(\overline{u}_p^{v,n}\big) - v\big(\overline{u}_m^{c,n}\big)}\right) \boldsymbol{.}\, \boldsymbol{n}_{mp} - \mathcal{F}_{mp}^{\mathrm{FV}}}{\Delta F_{mp}}\right) \tag{17}$$

Remark 4.2. *As one would expect, [56], condition* (17) *does decrease the accuracy to* ***second-order****, see Figure 5. Practically, it has been observed that this condition even further reduces the accuracy, as illustrated in Table 1.*

4.3. Semi-discrete cell entropy stability for one given entropy $\Longrightarrow$ *High-order*

To get a second-order accurate scheme, we had no choice but to relax a bit the aimed entropy stability, by getting a semi-discrete entropy stable scheme, at the subcell level, for one given pair (η, $\boldsymbol{\phi}$). In order to preserve the order of accuracy of this monolithic DG/FV scheme, we go from a subcell semi-discrete entropy stability to a cell based one, again for a given entropy. To do so, let us first introduce $\{\varphi_m^c\}_m$, a particular set of $\mathbb{P}^k$ basis functions. Let us emphasize that in the case where $N_s > N_k$, those functions form a spanning sets. Those functions, previously introduced in [58, 59] and that we refer to as sub-resolution basis functions. Those functions, can be seen as the L_2 projection of the subcell indicator functions $\mathbb{1}_{S_m^c}(\boldsymbol{x})$ onto $\mathbb{P}^k(\omega_c)$. They are defined such that $\forall\, \psi \in \mathbb{P}^k(\omega_c)$ and $\forall\, m = 1, \ldots, N_s$

$$\int_{\omega_c} \varphi_m\, \psi \,\mathrm{d}V = \int_{S_m^c} \psi \,\mathrm{d}V. \tag{18}$$

We note $\underline{v}_m^c$ the corresponding moments such that $v_h^c = \sum_{m=1}^{N_s} \underline{v}_m^c\, \varphi_m^c$. Now, let us express the time variation of a given entropy η over cell ω_c

$$\Delta\eta_c := \frac{\mathrm{d}}{\mathrm{d}t} \oint_{\omega_c} \eta\big(u_h^c\big) \,\mathrm{d}V = \oint_{\omega_c} v\big(u_h^c\big)\, \partial_t u_h^c \,\mathrm{d}V.$$

Referring by v_h^c the L_2 projection of the entropy variable $v\big(u_h^c\big)$ onto $\mathbb{P}^k(\omega_c)$, and by means of (18), the entropy variation can be put into the following simple expression

$$\Delta\eta_c = \int_{\omega_c} v_h^c \, \partial_t u_h^c \, \mathrm{d}V = \sum_{m=1}^{N_s} \underline{v}_m^c \int_{\omega_c} \varphi_m^c \, \partial_t u_h^c \, \mathrm{d}V = \sum_{m=1}^{N_s} \underline{v}_m^c \int_{S_m^c} \partial_t u_h^c \, \mathrm{d}V.$$

In the light of (8) and remark 2.4, the entropy time variation leads to the following compact form

$$\Delta\eta_c = \sum_{m=1}^{N_s} |S_m^c| \, \underline{v}_m^c \, \frac{\mathrm{d}\,\overline{u}_m^c}{\mathrm{d}t}. \tag{19}$$

In the case where $N_s > N_k$, this last equality has to be understood in a least square sens. The use of the semi-discrete scheme (10) provides us with the following expression

$$\Delta\eta_c = -\sum_{m=1}^{N_s} \underline{v}_m^c \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp} \, \widetilde{F_{mp}}. \tag{20}$$

Now, let $\mathfrak{f}_c$ be the set containing the subcell's faces of any subcell in ω_c, while $\breve{\mathfrak{f}}_c$ would be the set of subcell's faces inside ω_c, meaning not belonging to $\partial\omega_c$. By means of previous definition, $\#\,\breve{\mathfrak{f}}_c = N_f^c$, and if $f_{mp} \in \mathfrak{f}_c \setminus \breve{\mathfrak{f}}_c := \hat{\mathfrak{f}}_c$ that means $f_{mp} \subset \partial\omega_c$. Manipulating the two sums in (20) and recalling that the reconstructed fluxes are set to the DG numerical flux on the cell boundary, we are able to recast the cell entropy time evolution as

$$\begin{aligned}\Delta\eta_c \;=\; & \sum_{f_{mp} \in \breve{\mathfrak{f}}_c} l_{mp} \, \big(\underline{v}_p^c - \underline{v}_m^c\big) \, \widetilde{F_{mp}} \; - \sum_{f_{mp} \in \hat{\mathfrak{f}}_c} (1-\theta_{mp}) \, l_{mp} \, \underline{v}_m^c \, \mathcal{F}_{mp}^{\mathrm{FV}} \\ & - \sum_{f_{mp} \in \hat{\mathfrak{f}}_c} \theta_{mp} \, \underline{v}_m^c \oint_{f_{mp}} \mathcal{F}\left(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp}\right) \mathrm{d}S.\end{aligned}$$

Adding and retrieving $\sum_{f_{mp} \in \hat{\mathfrak{f}}_c} \theta_{mp} \oint_{f_{mp}} v\big(u_h^c\big) \, \mathcal{F}\left(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp}\right) \mathrm{d}S$ to the previous relation, the entropy variation can be separated into two terms, *i.e.* $\Delta\eta_c = \mathrm{A} + \mathrm{B}$, where

$$\mathrm{A} = \sum_{f_{mp} \in \breve{\mathfrak{f}}_c} l_{mp} \, \big(\underline{v}_p^c - \underline{v}_m^c\big) \, \widetilde{F_{mp}} + \sum_{f_{mp} \in \hat{\mathfrak{f}}_c} \theta_{mp} \oint_{f_{mp}} \Big(v\big(u_h^c\big) - \underline{v}_m^c\Big) \, \mathcal{F}\left(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp}\right) \mathrm{d}S \tag{21}$$

and

$$\mathrm{B} = -\sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp} \left((1-\theta_{mp}) \, \underline{v}_m^c \, \mathcal{F}_{mp}^{\mathrm{FV}} + \frac{\theta_{mp}}{l_{mp}} \oint_{f_{mp}} v\big(u_h^c\big) \, \mathcal{F}\left(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp}\right) \mathrm{d}S \right). \tag{22}$$

A sufficient to ensure a correct cell entropy inequality would then be to yield that

$$\mathrm{A} \leq \sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp} \left((1-\theta_{mp}) \, \boldsymbol{\Psi}\big(\underline{v}_m^c\big) + \frac{\theta_{mp}}{l_{mp}} \oint_{f_{mp}} \boldsymbol{\psi}\big(u_h^c\big) \, \mathrm{d}S \right) \boldsymbol{.}\, \boldsymbol{n}_{mp}. \tag{23}$$

Indeed, this sufficient condition would ensure the following inequality

$$\Delta\eta_c \leq -\sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp} \left((1-\theta_{mp}) \, \Big(\underline{v}_m^c \, \mathcal{F}_{mp}^{\mathrm{FV}} - \boldsymbol{\Psi}\big(\underline{v}_m^c\big) \boldsymbol{.}\, \boldsymbol{n}_{mp}\Big) + \frac{\theta_{mp}}{l_{mp}} \oint_{f_{mp}} \Big(v\big(u_h^c\big) \, \mathcal{F}_n - \boldsymbol{\psi}\big(u_h^c\big) \boldsymbol{.}\, \boldsymbol{n}_{mp}\Big) \, \mathrm{d}S \right). \tag{24}$$

The right hand side contains two contributions, a low-order one and a high-order one. Regarding the latter, if the numerical flux used in DG for the calculation of the reconstructed fluxes ensures the two-point Tadmor condition

$$\mathcal{F}\left(u_L, u_R, \boldsymbol{n}\right)\,\big(v(u_R) - v(u_L)\big) \leq \big(\boldsymbol{\psi}(u_R) - \boldsymbol{\psi}(u_L)\big)\,\boldsymbol{.}\,\boldsymbol{n}, \tag{25}$$

then the high-order part induces

$$-\frac{1}{l_{mp}} \oint_{f_{mp}} \Big(v\big(u_h^c\big)\,\mathcal{F}_n - \boldsymbol{\psi}\big(u_h^c\big)\,\boldsymbol{.}\,\boldsymbol{n}_{mp}\Big)\,\mathrm{d}S \leq -\frac{1}{l_{mp}} \oint_{f_{mp}} \phi^*\big(u_h^c,\, u_h^v,\, \boldsymbol{n}_{mp}\big)\,\mathrm{d}S := -\widehat{\phi_{mp}},$$

where the consistent numerical entropy flux ϕ^* is defined as

$$\phi^*(u_L, u_R, \boldsymbol{n}) = \frac{\big(v(u_L) + v(u_R)\big)}{2}\,\mathcal{F}(u_L, u_R, \boldsymbol{n}) - \frac{\big(\boldsymbol{\psi}(u_L) + \boldsymbol{\psi}(u_R)\big)}{2}\,\boldsymbol{.}\,\boldsymbol{n}. \tag{26}$$

Similarly, by means of an *appropriate* FV flux $\mathcal{F}_{mp}^{\mathrm{FV}}$, the low-order contribution in (24) produces

$$-\Big(\underline{v}_m^c\,\mathcal{F}_{mp}^{\mathrm{FV}} - \boldsymbol{\Psi}\big(\underline{v}_m^c\big)\,\boldsymbol{.}\,\boldsymbol{n}_{mp}\Big) \leq -\phi^*\Big(u\big(\underline{v}_m^c\big),\, u\big(\underline{v}_p^v\big),\, \boldsymbol{n}_{mp}\Big) := -\phi_{mp}^{\mathrm{FV}}. \tag{27}$$

Combining the low and high contributions, the sufficient condition (23) ensures that

$$\frac{\mathrm{d}}{\mathrm{d}t} \oint_{\omega_c} \eta\big(u_h^c\big)\,\mathrm{d}V \leq -\sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp}\,\Big((1-\theta_{mp})\,\phi_{mp}^{\mathrm{FV}} + \theta_{mp}\,\widehat{\phi_{mp}}\Big) := -\sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp}\,\widetilde{\phi_{mp}}, \tag{28}$$

which guarantees the entropy stability over cell ω_c, for a given entropy η.

Remark 4.3. *We previously said that by means of appropriate FV fluxes $\mathcal{F}_{mp}^{FV}$, relation* (27) *is ensured. To this end, the first-order FV fluxes, previously defined as $\mathcal{F}_{mp}^{FV} = \mathcal{F}\left(\overline{u}_m^c, \overline{u}_p^v, \boldsymbol{n}_{mp}\right)$, have to be modified for this entropy consideration as follows*

$$\mathcal{F}_{mp}^{FV} = \mathcal{F}\Big(u\big(\underline{v}_m^c\big),\, u\big(\underline{v}_p^v\big),\, \boldsymbol{n}_{mp}\Big). \tag{29}$$

While this definition, along with condition (23)*, guarantees entropy stability, global maximum and positivity preserving principles, introduced in the next Section 5, may not be assured anymore.*

Lastly, to ensure the sufficient condition (27), let us show that is possible to reformulate it as a continuous Knapsack problem, similarly to [42]. In this aforementioned paper, Y. Lin and J. Chan have introduced a way to ensure a semi-discrete cell entropy inequality for monolithic SEMDG schemes, by ultimately solving a continuous Knapsack optimization problem. This technique has been applied to systems of conservation laws in one-dimensional or tensor-product multi-dimensional settings. Investigations are currently carried out in the extension to simplex elements, but based on connecting quadrature points to build the SBP low-operator and not on any subcell representation. The main difference between the monolithic framework used in [48, 42] and the one presented here resides in the fact that the former are for now mostly limited to tensor-product Cartesian grids, as the summation-by-parts property of the scheme derives from specific quadrature points based solution approximation and flux collocation. The framework presented here is by construction multi-dimensional and can be theoretically applied to any type of grids with great flexibility in the

choice of cell sub-partition. Let us emphasize nonetheless that in [48, 42], thanks to the Gauss-Lobatto representation, the solution point-values can be simultaneously considered as subcell mean values. This characteristic, while mainly limiting the scheme to one-dimensional or tensor-product multi-dimensional geometries, makes the entropy analysis way simpler, as only one set of data is involved. Here, two sets of data are required in the entropy development, see definition (19), namely the solution subcell mean values $\overline{u}_m^c$ and the entropy solution sub-resolution moments $\underline{v}_m^c$. This is the reason why, as said in Remark 4.3, the first-order FV numerical fluxes require to be modified if one aims at this cell entropy stability.

Let us finally formulate the sufficient condition (27) as the following continuous Knapsack problem

$$\mathbf{C}_c \,.\, \mathbf{\Theta}_c \leq \mathrm{D}_c, \tag{30}$$

where the vector $\mathbf{\Theta}_c = \Big(\theta_1, \ldots, \theta_{\#\hat{\mathfrak{f}}_c}, \theta_{\#\hat{\mathfrak{f}}_c+1}, \ldots, \theta_{\#\mathfrak{f}_c}\Big)^{\mathrm{t}}$ contains all the boundary subcells' faces and interior subcells' faces, where D_c the right hand side is defined as

$$\mathrm{D}_c \;=\; \sum_{f_{mp} \in \hat{\mathfrak{f}}_c} l_{mp}\, \mathbf{\Psi}\big(\underline{v}_m^c\big) \,.\, \boldsymbol{n}_{mp} \;-\; \sum_{f_{mp} \in \check{\mathfrak{f}}_c} l_{mp}\, \big(\underline{v}_p^c - \underline{v}_m^c\big)\, \mathcal{F}_{mp}^{\mathrm{FV}}, \tag{31}$$

while vector $\mathbf{C}_c$ writes as, with $f_i = f_{mp}$

$$\mathrm{C}_i^c \;=\; \begin{cases} \displaystyle\oint_{f_{mp}} \Big(\big(v(u_h^c) - \underline{v}_m^c\big)\, \mathcal{F}_n - \big(\boldsymbol{\psi}(u_h^c) - \mathbf{\Psi}(\underline{v}_m^c)\big) \,.\, \boldsymbol{n}_{mp} \Big)\, \mathrm{d}S, & \forall i = 1, \ldots, \#\hat{\mathfrak{f}}_c, \\[2ex] l_{mp}\, \big(\underline{v}_p^c - \underline{v}_m^c\big)\, \Delta F_{mp}, & \forall i = \#\hat{\mathfrak{f}}_c + 1, \ldots, \#\mathfrak{f}_c. \end{cases}$$

Let us first state that (30) is indeed solvable as $\mathrm{D}_c \geq 0$, hence in the worse case $\mathbf{\Theta}_c$ can be set to zero. The positivity of D_c is easily verifiable as, through (29) and (26), it directly follows that $-\big(\underline{v}_p^c - \underline{v}_m^c\big)\, \mathcal{F}_{mp}^{\mathrm{FV}} \leq -\big(\mathbf{\Psi}(\underline{v}_m^c) - \mathbf{\Psi}(\underline{v}_p^v)\big) \,.\, \boldsymbol{n}_{mp}$, and D_c can be recast into

$$\mathrm{D}_c \geq \sum_{m=1}^{N_s} \left(\sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \boldsymbol{n}_{mp} \right) \,.\, \mathbf{\Psi}(\underline{v}_m^c) = 0.$$

Following the steps of [42], we efficiently solved (30) through a Greedy algorithm (see [42] for a detailed algorithm) by finding all the θ_{mp} under the constraint $0 \leq \theta_{mp} \leq \theta_{mp}^e \leq 1$, where the θ_{mp}^e can be any additional constraint on the blending coefficient, while maximizing $\sum_{f_{mp} \in \mathfrak{f}_c} \theta_{mp}$.

Remark 4.4. *While conditions* (16) *and* (17) *would respectively reduced the accuracy to first and second order, condition* (30) *does allow the preservation of the high-order accuracy of the scheme, see for instance Table 1 and Figure 5. To make sure of it, let us raise that substituting u_h by a smooth function u leads to $\mathbf{C}_c \,.\, \mathbf{1} - \mathrm{D}_c = |\omega_c|\, \mathcal{O}(h_c^{k+1})$, where h_c is the diameter of cell ω_c. Indeed, after some simple manipulations, it is possible to write that*

$$\mathbf{C}_c \,.\, \mathbf{1} - \mathrm{D}_c = E_{\partial\omega_c}\big(\boldsymbol{\phi}(u) \,.\, \boldsymbol{n}\big) + E_{\omega_c}\big(\boldsymbol{F}(u) \,.\, \nabla_x v_h^c\big) - E_{\partial\omega_c}\big(v_h^c\, \boldsymbol{F}(u) \,.\, \boldsymbol{n}\big) + \int_{\omega_c} \big(v(u) - v_h^c\big)\, \nabla_x \,.\, \boldsymbol{F}(u)\, \mathrm{d}V,$$

where $E_\Omega(f) = \oint_\Omega f \, \mathrm{d}\Omega - \int_\Omega f \, \mathrm{d}\Omega$. *Then, using similar arguments as in [10], and since we make use of quadrature rules respectively exact for polynomials up to degree* $2k$ *over* ω_c *and* $2k+1$ *over* $\partial\omega_c$, *we obtain the desired result. Then, by means of the greedy algorithm, see [42], it is possible to state that* $(\theta_{mp} - 1) \, \Delta F_{mp} = \frac{|\omega_c|}{l_{mp}} \, \mathcal{O}(h_c^{k+1}) = \mathcal{O}(h_c^{k+2})$. *In the end, as the blended flux can be re-expressed as* $\widetilde{F_{mp}} = \widehat{F_{mp}} + (\theta_{mp} - 1) \, \Delta F_{mp} = \widehat{F_{mp}} + \mathcal{O}(h_c^{k+2})$, *the monolithic DG/FV scheme reduces to pure DG, up to* $\mathcal{O}(h_c^{k+2})$, *in this smooth solution context.*

4.4. Numerical results: entropy stabilities

To end this section concerned with entropy stabilities and to confirm the previous stated results regarding the different entropy stability and their respective cost in accuracy, we run some numerical tests on some classical problems. Let us emphasize that a lot more problems and test cases will be considered in the section devoted to maximum principles, Section 5.3, as the following one is simply devoted to the questions regarding entropy announced in the introduction of this part. In the following, if not stated otherwise, the subcell mean values will be displayed.

4.4.1. 1D linear advection case

First, let us consider the very simple case of 1D linear advection $\partial_t u + a \partial_x u = 0$, where the velocity is set to $a = 1$. We start from a smooth initial condition $u_0(x) = \sin(2\pi x)$ and assume periodic boundary conditions. We assess the scheme accuracy after one full period. In Figure 4, numerical solutions obtained by means of the $\mathbb{P}^8$ monolithic DG/FV scheme based on the three different blending coefficients, conditions (16), (17) and (30), ensuring the three different types of entropy stability, are plotted.

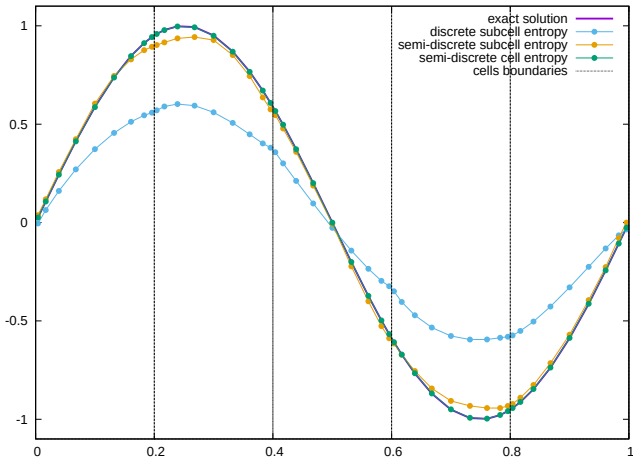


Figure 4: $\mathbb{P}^8$-DG/FV solutions on 5 cells with $\eta(u) = \frac{1}{2}u^2$

As expected, one can see how the first condition does reduce the accuracy to first-order, the second to second-order while the third one is the only one able to preserve the high accuracy, as with only 5 cells the numerical solution is extremely close to the exact one. The rates of convergence gathered in Table 1 further confirm this result. Let us emphasize that, as previously proved, in this smooth solution context the local subcell monolithic DG/FV scheme with the third entropy stability condition reduces to a pure DG scheme, up to machine precision.

	Entropy stability n°1		Entropy stability n°2		Entropy stability n°3	
h	E_{L_2}	q_{L_2}	E_{L_2}	q_{L_2}	E_{L_2}	q_{L_2}
1/1	6.97E-1	0.21	5.76E-1	1.35	1.32E-2	4.64
1/2	6.01E-1	0.48	2.26E-1	1.43	5.27E-4	6.36
1/4	4.29E-1	1.54	8.40E-2	1.41	6.41E-6	5.98
1/8	2.64E-1	0.92	3.16E-2	1.22	1.02E-7	5.89
1/16	1.48E-1	-	1.36E-2	-	1.72E-9	-

Table 1: Convergence rates for the linear advection case for $\mathbb{P}^5$-DG/FV monolithic scheme

We follow with the classical case of the linear advection of a composite signal, introduced in [31]. This signal is composed by the succession of a Gaussian, rectangular, triangular and parabolic signals.

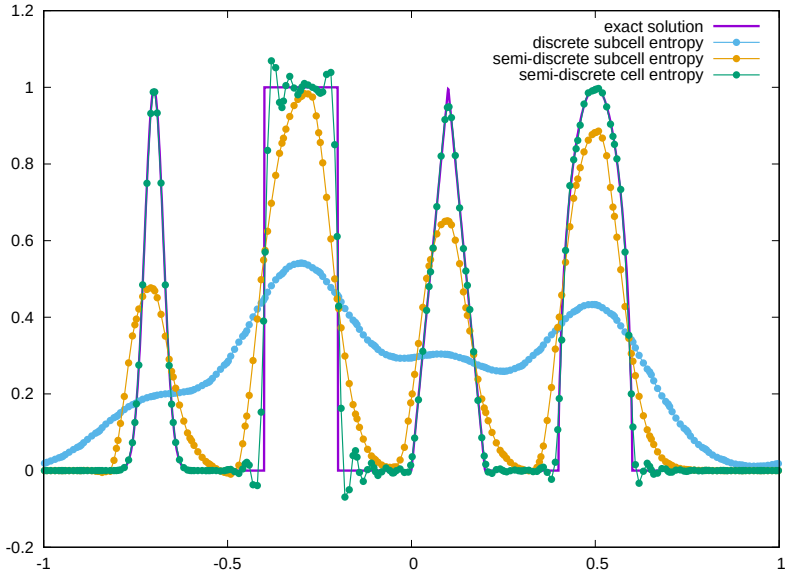
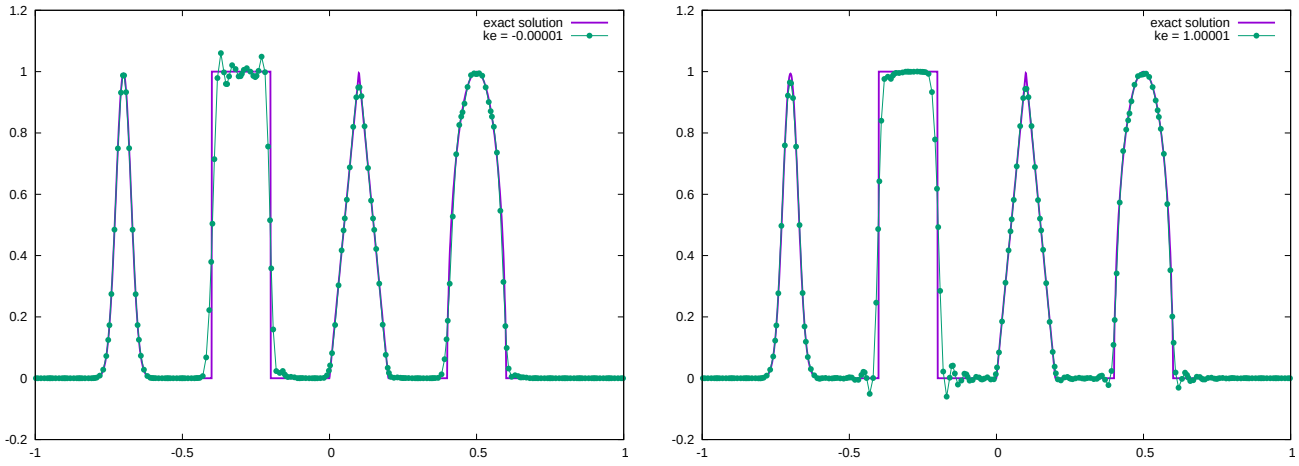


Figure 5: $\mathbb{P}^5$-DG/FV solutions on 40 cells and $\eta(u) = \frac{1}{2}u^2$

Figure 5, in which monolithic $\mathbb{P}^5$-DG/FV solutions ensuring the three types of entropy stability are displayed, confirms furthermore our previous conclusion on accuracy and entropy. However, let us note that the numerical solution ensuring the semi-discrete cell entropy inequality, for $\eta(u) = \frac{1}{2}\,u^2$, is very close to what a pure DG scheme would produce and thus exhibits the same pathologies. To be able to see a more significant impact of this entropy inequality enforcement, let us use different entropies $\eta(u) = |u - k_e|^{1+\epsilon} \backslash (1+\epsilon)$. Those can be seen as a smoothed version of the Kruzkov's entropies. The coefficient $\epsilon > 0$ is set here at $\epsilon = 0.25$, while different values of k_e will be used. In Figure 6, we make use of two different ones, respectively $k_e = -0.00001$ and $k_e = 1.00001$. One can clearly see how the numerical solution has now been impacted by the semi-discrete cell entropy stability, and how this condition put the emphasis around $u = k_e$ making in the first case the numerical solution almost positive while in the second case roughly less than one.

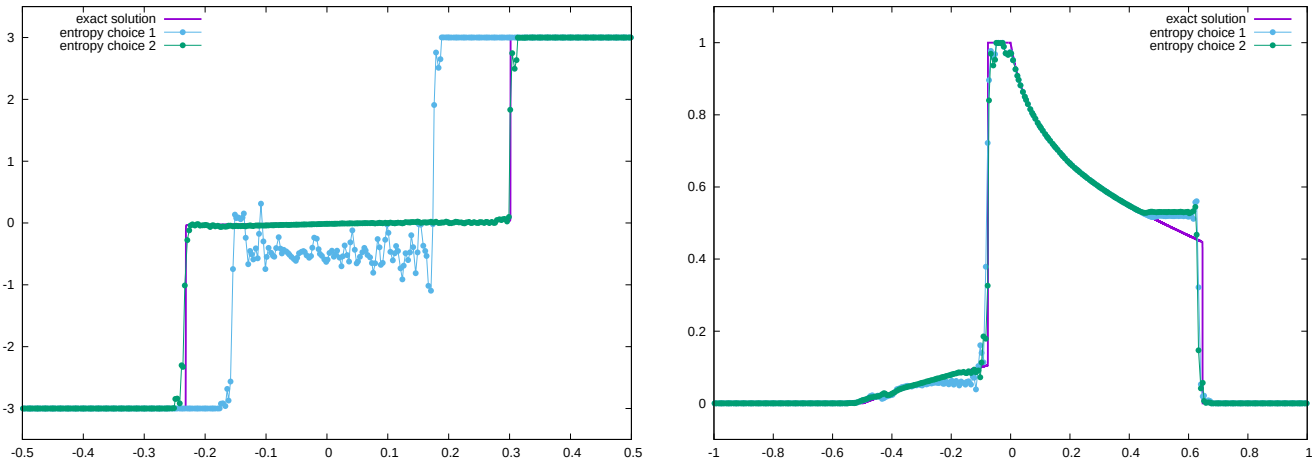


(a) $k_e = -0.00001$ (b) $k_e = 1.00001$

Figure 6: $\mathbb{P}^5$-DG/FV submean values on 40 cells: $\epsilon = 0.25$ and $\eta(u) = |u - k_e|^{1+\epsilon} \backslash (1+\epsilon)$

4.4.2. 1D Buckley non-convex case

Now, we address the challenging 1D Buckley problem. The Buckley equation is defined as $\partial_t u + \partial_x F(u) = 0$, where the non-convex flux function writes $F(u) = \frac{4\,u^2}{4\,u^2+(1-u)^2}$. As said in Remark 2.1, since the flux function is now a complex rational function, it is not practical to analytically integrate the volume integrals. And due to that, entropy stability proved in [30] does not hold anymore. Furthermore, approximated integration or collocation of the flux may also produce some aliasing effects, see [58] for some examples. Here, we want to observe the benefit of entropy stability and check if a semi-discrete cell entropy inequality for one given entropy is practically enough to capture the correct unique entropic weak solution. To this end, two different test cases will be addressed. The first one has been introduced by T. Chen and C.-W. Shu in [8]. This Riemann problem, consisting in an initial discontinuity located at $x = 0$ and taking -3 and 3 as left and right values, admits an entropic solution containing two shock waves connected by a flat rarefaction that is close to 0. We run this test case on a domain $[-0.5, 0.5]$ and end at time $t = 1$. As said in the aforementioned paper, in this case the choice of the entropy function is critical. To corroborate this statement, we make use of the same two entropy functions as in [8], meaning the energy one $\eta(u) = \frac{1}{2}\,u^2$ and $\eta(u) = \int \arctan(20\,u)\,\mathrm{d}u$ a mollified version of the Kruzkov's entropy $|u|$. To solely observe the effect of entropy stability, and not other choices of blending coefficients ensuring other properties, see Section 5, we additionally use here the DG maximum principle preserving limitation of X. Zhang and C.-W. Shu [66] to make sure the computation goes through. This additional limitation will only be used here as in the next section maximum principle and positivity will be ensured directly by appropriate choice of blending coefficients. In Figure 7(a), $\mathbb{P}^3$-DG/FV monolithic scheme is used on 80 cells, ensuring the high accuracy preserving semi-discrete cell entropy stability, with the two different entropies discussed before. In Figure 7(a), it is clear how the scheme using the energy entropy has failed to capture the entropic weak solution, while using another entropy, putting the stress around 0 because being an approximation of $|u|$, has solved this issue. This is perfectly consistent to the results obtained in [8]. Now, making use of the same two entropies, we run a second test case in which we start from the initial solution $u^0(x) = 1$ if $x \in [-\frac{1}{2}, 0]$ and $u^0(x) = 0$ elsewhere. The results obtained again by means of $\mathbb{P}^3$-DG/FV monolithic scheme is used on 80 cells are shown



(a) Test case n°1 (b) Test case n°2

Figure 7: $\mathbb{P}^3$-DG/FV submean values on 80 cells: $\eta_1(u) = \frac{1}{2}u^2$ and $\eta_2(u) = \int \operatorname{atan}(20u)\,\mathrm{d}u$

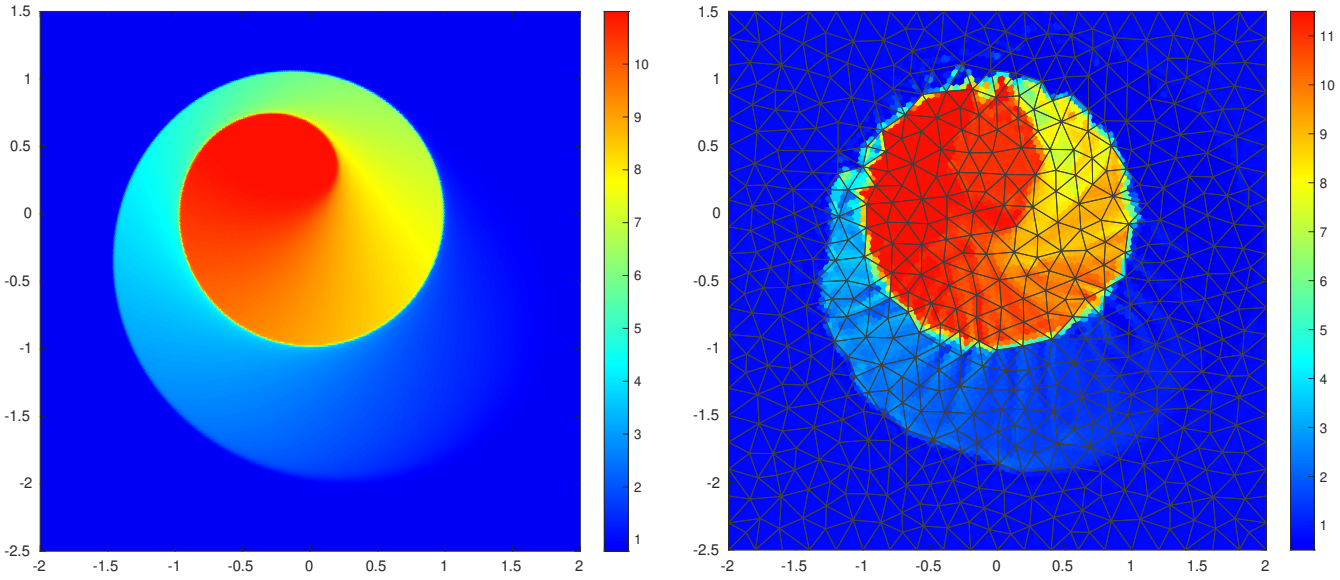
in Figure 7(b). One can clearly see how none of these two choices of entropy has enabled the scheme to capture the entropic solution. A proper entropy has to be designed in some retro-engineering process to fit not only the PDE considered but also the test case to hope to capture the correct solution. And this may be even not feasible if the solution presents very complex structures. Let us emphasize that similar results would have been obtained ensuring a semi-discrete subcell entropy inequality, meaning by means of condition (17), as this entropy stability is ensured only for one chosen entropy. Only the first condition (16) enforcing discrete entropy stability for any entropy will succeed in capturing the unique weak solution in both cases. But, as recalled, it will reduce the accuracy to first-order.

4.4.3. 2D KPP non-convex case

We now turn our attention to the 2D KPP problem proposed by Kurganov, Petrova, Popov (KPP) in [34]. For this particular problem, the flux function is given by $\boldsymbol{F}(u) = (\sin(u), \cos(u))^{\mathrm{t}}$. Considering the computational domain $[-2, 2] \times [-2.5, 1.5]$, the initial condition reads as follows

$$u^0(x) = \begin{cases} \frac{7\pi}{2} & \text{if } x < \frac{1}{2}, \\ \frac{\pi}{4} & \text{if } x > \frac{1}{2}. \end{cases}$$

This test is very challenging to many high-order schemes as the solution has a two-dimensional composite wave structure. Thus, generally, to be able to capture such rotation composite structure, very fine grids are used. Here, we compare a referential solution, obtained through a first-order FV scheme on a very fine grid made of 209184 triangular cells, with the one obtained with the $\mathbb{P}^3$-DG/FV monolithic scheme with the accuracy preserving semi-discrete cell entropy stability for $\eta(u) = \frac{1}{2}u^2$, on a coarse mesh made of 1054 cells, see Figure 8. Because that it has been previously observed in [59] that in this particular case, voronoi-type cell subdivision 1(b) produces slightly better results, this cell sub-partition is then used here. One can notice how this entropy stability condition does not guarantee the capture of the correct weak entropy solution. And refining the mesh would not cure this issue.



(a) 1^{th}-order FV on 209184 cells (b) $\mathbb{P}^3$-DG/FV on 1054 cells

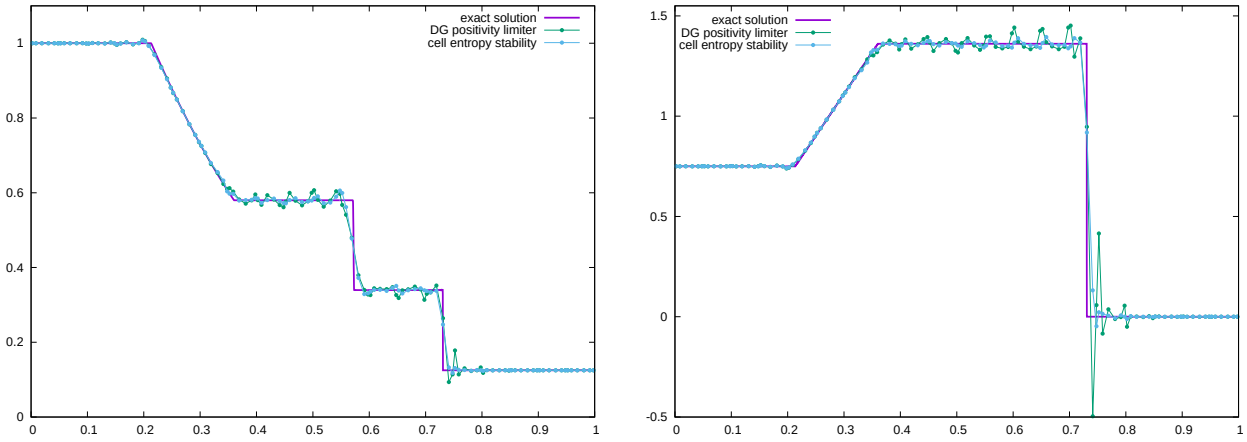
Figure 8: $\mathbb{P}^3$-DG/FV entropic scheme: non-entropic solution

4.4.4. 1D modified Sod shock tube problem

Finally, to end this section regarding entropy stability, we now consider the 1D version of the Euler compressible gas system of equations (33). A classical test case where numerical schemes may capture non-physical weak solution is the modified Sod shock tube problem, [57]. This problem is a modified version of the popular Sod's test [52]; the solution consists of a right shock wave, a right traveling contact wave and a left sonic rarefaction wave. This test is very useful in assessing the entropy satisfaction property of numerical methods, as some of them may present a shock at the sonic point in the rarefaction. In Figure 9 both solutions obtained by means of $\mathbb{P}^5$ pure DG scheme with the X. Zhang and C.-W. Shu positivity-preserving limiter [66] and our $\mathbb{P}^5$-DG/FV monolithic scheme with the accuracy preserving semi-discrete cell entropy stability. The chosen entropy in the Euler case is $\eta(\mathsf{U}) = -\rho \log(p/\rho^\gamma)$. In both cases, Rusanov type of numerical flux has been used. Firstly, it is important to note that while pure DG without positivity limiter crashes and thus fails to produce a solution, monolithic scheme with cell entropy stability does run without any need of positivity limiter. Secondly, one can see how the cell entropy stability reduces the amplitude of spurious oscillations. However, let us emphasize that both schemes successfully capture the correct entropic solution, as none of them present the non-physical shock in the rarefaction wave. Similar results can be obtained with the use of global Lax-Friedrichs, HLL and HLL-C numerical fluxes.

4.5. Conclusion on entropy stabilities?

Let us recall the questions we were asking ourselves at the beginning of this section. First, is it possible to find θ_{mp}, the blending coefficients, ensuring entropy stability? The answer is obviously yes. But what do we mean by entropy stability? We have shown how to ensure three types of entropy stabilities, meaning a fully-discrete subcell entropy inequality for any entropy, a semi-discrete subcell entropy one for a given entropy and a semi-discrete cell entropy one for a given entropy. The follow-up question is then what are the costs of such constraints? Well, if one wants to ensure en-



(a) Density (b) Velocity

Figure 9: $\mathbb{P}^5$ pure DG with positivity limiter and DG/FV monolithic scheme with cell entropy stability on 20 cells

tropy stability for any entropy, the scheme has no choice but to reduce to a first-order one. Relaxing this objective and aiming for an entropy stability for only one entropy, we may achieve second or even arbitrary high-order of accuracy. However, to indeed preserve the high-order accuracy of DG scheme, the cell entropy inequality requires to modify the definition of the subcell's faces FV fluxes in the monolithic DG/FV scheme, see Remark 4.3, which may invalidate other properties as positivity for instance. Finally, regarding the last question, is entropy stability absolutely needed? In the design of first-order scheme or any robust scheme forming the safe basis of a MOOD-type [9, 13] or monolithic scheme, it totally is. A FV scheme with the Roe solver without entropy fix would fail for instance to capture the correct solution in the modified Sod test case 4.4.4. That being said, if the goal is to go to very high order of accuracy, one has no choice but to relax its expectations and aim for an entropy stability for only one entropy. And then, it may always be possible to design a test case, considering a particular scalar conservation law with complex fluxes, that will trick high-order entropy conservative/stable schemes and make them to fail to capture the unique entropic solution. From our experience, while high-order entropy stability does slightly introduce numerical diffusion and consequently reduces spurious oscillations, it is generally not enough to capture the correct entropic solution in complicated cases. An additional shock capturing technique is added, which brings further diffusion and hence does the trick. Now, considering systems of equations, as the Euler compressible gas one, generally no high-order entropy stability is needed as a pure DG scheme with entropic stable numerical fluxes would be enough to capture the entropic solution, see Figure 9. It may not be the case considering complex equations of state, with non-concave pressure for instance, but in classical settings we are not aware of any test case where an high-order DG scheme, based on entropic numerical fluxes as Rusanov or HLL, does capture a non-entropic solution.

While entropy stability may not be absolutely required, preserving global bounds of some variables generally is, as negative density in the compressible gas case would lead to a crash of the code. The next section aims at defining the blending coefficients to guarantee the numerical solution to remain in a convex admissible set, as well as imposing a local maximum principle to reduce the apparition of spurious oscillations. Those conditions will prove to introduce enough numerical diffusion for the scheme to capture the correct entropic solution as well as to produce the best results.

5. Global and local maximum principles

This section is devoted to the definition of blending coefficients to ensure different maximum principles. It is important to emphasize that we do distinguish physical maximum principles, that the unique entropic solution should ensure, from numerical maximum principles used in the following to avoid, as much as possible, the apparition of non-physical oscillations.

5.1. Physical maximum principles

In here, we present the minimum requirements on the numerical solution to ensure the simulation code to be robust.

5.1.1. Scalar conservation laws and global maximum principle

Considering SCL, equation (1), if the initial datum yields $u_0 \in [\alpha,\, \beta]$, then the unique entropic weak solution u ensures $u(.,t) \in [\alpha,\, \beta]$ for any time t. To guarantee that the numerical solution submean values do ensure such property, it is sufficient, by convexity of relation (12), that the blended Riemann intermediate state $\widetilde{u_{mp}}^{\pm}$ also remain in $[\alpha,\, \beta]$. Since $u_{mp}^{*,\mathrm{FV}}$ does, a sufficient condition is to take θ_{mp} such that

$$\theta_{mp} \leq \min\Big(1,\, \Big|\frac{\gamma_{mp}}{\Delta F_{mp}}\Big|\, \min\big(\beta - u_{mp}^{*,\mathrm{FV}},\, u_{mp}^{*,\mathrm{FV}} - \alpha\big)\Big). \tag{32}$$

5.1.2. Euler system and positivity preservation

Let us consider the Euler compressible gas dynamics system

$$\begin{cases} \partial_t\, \mathsf{U}(\boldsymbol{x},t) + \nabla_x \boldsymbol{.}\, \mathsf{F}\big(\mathsf{U}(\boldsymbol{x},t)\big) = 0, & (\boldsymbol{x},t) \in \omega \times [0,T], \qquad (33a) \\ \mathsf{U}(\boldsymbol{x},0) = \mathsf{U}_0(\boldsymbol{x}), & \boldsymbol{x} \in \omega, \qquad (33b) \end{cases}$$

with $\mathsf{U} = (\rho,\, \boldsymbol{q},\, E)^{\mathrm{t}}$ and $\mathsf{F}\big(\mathsf{U}\big) := \big(\boldsymbol{F}^{\rho},\, \boldsymbol{F}^{\boldsymbol{q}},\, \boldsymbol{F}^{E}\big)^{\mathrm{t}} = (\boldsymbol{q},\, \boldsymbol{v}\otimes\boldsymbol{q} + p\, I_d,\, (E+p)\,\boldsymbol{v})^{\mathrm{t}}$. The conserved variables ρ, $\boldsymbol{q} = \rho\,\boldsymbol{v}$ and E then respectively stand for the density, momentum and total energy, while $\boldsymbol{v}$ characterizes the fluid velocity. The thermodynamic closure is given by the equation of state $p = p(\rho,\, \varepsilon)$ where $\varepsilon = E - \frac{1}{2}\rho\|\boldsymbol{v}\|^2$ denotes the internal energy. In this paper, we make use of a gamma gas law, *i.e.* $p = (\gamma-1)\,\varepsilon$, where γ is the polytropic index of the gas. Although the whole theory presented here has been introduced in the simple case of scalar conservation laws, the extension to the system case is perfectly straightforward.

Now, defining the admissible convex set $G = \big\{\mathsf{U} = (\rho,\, \boldsymbol{q},\, E)^{\mathrm{t}},\quad \rho > 0,\, p > 0\big\}$, we want to ensure that any subcell mean value $\overline{u}_m^c$ remains in G at all time. Following similar steps than in [36, 48], we first ensure the positivity of the density and then of the internal energy. Firstly, we introduce a first temporary blending coefficient $\theta_{mp}^{(1)}$ such that

$$\theta_{mp}^{(1)} \leq \min\Big(1,\, \Big|\frac{\gamma_{mp}}{\Delta F_{mp}^{\rho}}\Big|\, \rho_{mp}^{*,\mathrm{FV}}\big)\Big). \tag{34}$$

Then, defining quantities A_{mp}, B_{mp} and M_{mp} as in the following

$$\begin{cases} A_{mp} = \dfrac{1}{(\gamma_{mp})^2}\, \left(\dfrac{1}{2}\, \big\|\Delta\boldsymbol{F}_{mp}^{\boldsymbol{q}}\big\|^2 - \theta_{mp}^{(1)}\, \Delta F_{mp}^{\rho}\, \Delta F_{mp}^{E}\right), \\ B_{mp} = \dfrac{1}{\gamma_{mp}}\, \left(\boldsymbol{q}_{mp}^{*,\mathrm{FV}} \boldsymbol{.}\, \Delta\boldsymbol{F}_{mp}^{\boldsymbol{q}} - \rho_{mp}^{*,\mathrm{FV}}\, \Delta F_{mp}^{E} - \theta_{mp}^{(1)}\, E_{mp}^{*,\mathrm{FV}}\, \Delta F_{mp}^{\rho}\right), \\ M_{mp} = \rho_{mp}^{*,\mathrm{FV}}\, E_{mp}^{*,\mathrm{FV}} - \frac{1}{2}\, \big\|\boldsymbol{q}_{mp}^{*,\mathrm{FV}}\big\|^2, \end{cases} \tag{35}$$

where $M_{mp} > 0$ in the case of a positivity-preserving FV scheme, we introduce a second temporary blending coefficient as

$$\theta_{mp}^{(2)} \leq \min\left(1,\ \frac{M_{mp}}{\left|B_{mp}\right| + \max\left(0, A_{mp}\right)}\right). \tag{36}$$

Finally, the blending coefficients will be defined as the product of these two, *i.e.* $\theta_{mp} = \theta_{mp}^{(1)}\,\theta_{mp}^{(2)}$.

Remark 5.1. *The previous formula have been given for a numerical flux of form* (3)*, which includes global Lax-Friedrichs and Rusanov fluxes, producing a single FV intermediate state* $\mathsf{U}_{mp}^{*,FV}$*. Everything can be easily extended to other types of numerical fluxes, as HLL or HLL-C fluxes for instance. In those latter cases, for sake of simplicity, we can set* γ_{mp}*, present in the previous formula, as* $\gamma_{mp} = \max(|S_{mp}^L|, |S_{mp}^R|)$*, where* S_{mp}^L *and* S_{mp}^R *are the smallest and highest velocities of the two acoustic waves, see [1] for a proper definition of those velocities to ensure a positivity-preserving behavior. Doing so, the first-order FV numerical flux will then produce two FV Riemann intermediate states* $\mathsf{U}_{mp}^{*,\pm}$*, left and right, being defined as a convex combination of the HLL or HLL-C intermediate states and the initial left and right states* $\overline{\mathsf{U}}_m^c$ *and* $\overline{\mathsf{U}}_p^v$*.*

Remark 5.2. *It is essential to emphasize that those maximum or positivity principles impose the subcell mean values to remain in G, the admissible set. However, nothing is said on the values of the solution polynomial reconstruction, required in the computation of the DG residuals to define the reconstructed fluxes. It is thus perfectly possible that the polynomial solution in a cell yields a non-admissible value, at a cell interface for instance, which will lead to non-admissible reconstructed fluxes (even possibly NaN values). Obviously, this situation is automatically treated in this monolithic framework as, if the reconstructed flux* $\widehat{F_{mp}}$ *presents a pathological value, as NaN for instance, the blending coefficient will be set to zero. In the end, this means that if the uncorrected DG solution is nowhere to be saved inside the cell and the DG code would have then crashed, the monolithic scheme will then reduced to a first-order FV scheme applied on each subcell contained in the pathological cell, see for instance the results obtained for the Mach 20 hypersonic flow test case 5.3.10.*

5.2. Local maximum principles

Now, to reduce as much as possible the apparition of spurious oscillations in the approximation of discontinuous solution, we choose to impose a local maximum principle, at the subcell level.

5.2.1. Scalar conservation laws and local maximum principle

For SCL, thanks to their hyperbolic nature, the solution at a point should remained bounded by the minimum and maximum values of the solution at a previous time, taken in a large enough domain including the point under consideration. To ensure a low oscillatory behavior of the numerical solution, we will mimic such property at the discrete level. To do so, we will impose the submean value $\overline{u}_m^{c,n+1}$ on subcell S_m^c to be bounded by the submean values at the previous time step (or the previous RK step in the general case) in a given subcells set, as

$$\alpha_m^c := \min_{S_q^w \in \mathcal{N}(S_m^c)} \left(\overline{u}_q^{w,n}\right) \ \leq\ \overline{u}_m^{c,n+1} \ \leq\ \max_{S_q^w \in \mathcal{N}(S_m^c)} \left(\overline{u}_q^{w,n}\right) := \beta_m^c, \tag{37}$$

where $\mathcal{N}(S_m^c)$ is some set of S_m^c neighboring subcells, including subcell S_m^c, yet to be defined. The wider the set $\mathcal{N}(S_m^c)$ is, the softer this local maximum principle will be. Reversely, a smaller set would lead to a larger first-order FV contribution in flux blending. In this work, similarly to the one used in the non-linear numerical results section of [59], $\mathcal{N}(S_m^c)$ will be constituted by subcell

S_m^c, as well as all its face and node neighboring subcells S_q^v, either they belong to the same cell or not. By introducing $\mathcal{P}_m^c$ the set of vertices of subcell S_m^c, as well as $\mathcal{V}_p$ the set of subcells that share $\boldsymbol{x}_p$ as a vertex, *i.e.* $S_p^v \in \mathcal{V}_q \implies \boldsymbol{x}_q \in \mathcal{P}_p^v$, this definition of $\mathcal{N}(S_m^c)$ can be rewritten as $\mathcal{N}(S_m^c) = \bigcup_{\boldsymbol{x}_p \in \mathcal{P}_m^c} \mathcal{V}_p$.

Now, to guarantee condition (37) for any subcell, in the light of relation (12) it is sufficient to ensure that $\widetilde{u_{mp}}^- \in [\alpha_m^c, \beta_m^c]$ along with $\widetilde{u_{mp}}^+ \in [\alpha_p^v, \beta_p^v]$. Since $u_{mp}^{*,\mathrm{FV}}$ insures both conditions, it is sufficient to take θ_{mp} such that

$$\theta_{mp} \leq \min\left(1, \left|\frac{\gamma_{mp}}{\Delta F_{mp}}\right| \left\{ \begin{array}{ll} \min\left(\beta_p^v - u_{mp}^{*,\mathrm{FV}},\, u_{mp}^{*,\mathrm{FV}} - \alpha_m^c\right) & \text{if } \Delta F_{mp} > 0 \\ \min\left(\beta_m^c - u_{mp}^{*,\mathrm{FV}},\, u_{mp}^{*,\mathrm{FV}} - \alpha_p^v\right) & \text{if } \Delta F_{mp} < 0 \end{array} \right.\right). \tag{38}$$

Remark 5.3. *Let us enlighten that the local maximum principle* (37) *relies on subcell mean values. And because this constraint is not concerned with the whole polynomial set of values, it is very well-known that one has to relax it to preserve scheme accuracy in the presence of smooth extrema.*

Smooth extrema relaxation. In order to preserve high-accuracy in the vicinity of smooth extrema, we make use of a subcell level version of the smooth detector we introduced in our previous article, [59]. This latter is closed related to the smoothness indicator for finite elements, [39]. The basic idea of this detector is the following: the numerical solution is supposed to exhibit a smooth extrema if at least the linearized version of the numerical solution spatial derivatives present a monotonous profile. To this end, let us introduce the following subcell linear reconstructions

$$\left\{ \begin{array}{l} v_x^m(\boldsymbol{x}) = \overline{\partial_x u_h^c}^{\,m} + \overline{\nabla_x\left(\partial_x u_h^c\right)}^{\,m} \cdot (\boldsymbol{x} - \boldsymbol{x}_m^c), \qquad (39a) \\ v_y^m(\boldsymbol{x}) = \overline{\partial_y u_h^c}^{\,m} + \overline{\nabla_x\left(\partial_y u_h^c\right)}^{\,m} \cdot (\boldsymbol{x} - \boldsymbol{x}_m^c). \qquad (39b) \end{array} \right.$$

In (39), $\boldsymbol{x}_m^c$ denotes the centroid of subcell S_m^c, while $\overline{\partial_{x\backslash y} u_h^c}^{\,m}$ and $\overline{\nabla_x\left(\partial_{x\backslash y} u_h^c\right)}^{\,m}$ are nothing but the averaged values on S_m^c of the successive partial derivatives of u_h^c. In practice, this smoothness indicator works as a vertex-based limiter on $v_{x\backslash y}^m(\boldsymbol{x})$. Due to their linearity, functions $v_{x\backslash y}^m(\boldsymbol{x})$ attain their extrema at the vertices $\boldsymbol{x}_q \in \mathcal{P}_m^c$. Then, we consider that the exact weak solution underlying the numerical solution u_h presents a smooth profile in subcell S_m^c if, for any vertex $\boldsymbol{x}_q \in \mathcal{P}_m^c$, the linearized spatial derivative functions ensure the following constraint

$$v_{x,\,q}^{\min} \leq v_x^m(\boldsymbol{x_q}) \leq v_{x,\,q}^{\max} \qquad \text{and} \qquad v_{y,\,q}^{\min} \leq v_y^m(\boldsymbol{x_q}) \leq v_{y,\,q}^{\max}, \tag{40}$$

where $v_{x\backslash y,\,q}^{\min}$ and $v_{x\backslash y,\,q}^{\max}$ are respectively defined as

$$v_{x\backslash y,\,q}^{\min} = \min_{v \in \mathcal{V}_q} v_{x\backslash y}^m(\boldsymbol{x}_q) \qquad \text{and} \qquad v_{x\backslash y,\,q}^{\max} = \max_{v \in \mathcal{V}_q} v_{x\backslash y}^m(\boldsymbol{x}_q). \tag{41}$$

Practically, if for any vertex $\boldsymbol{x}_q \in \mathcal{P}_m^c$, conditions (40) are ensured, we then consider that the numerical solution presents a smooth profile on subcell S_m^c. Finally, if both the solution is considered smooth in both subcells S_m^c and S_p^v, the blended coefficient constraint through the local maximum condition (37) is relaxed. This procedure allows in practice the preservation of smooth extrema along with the order of accuracy for smooth problems, see Section 5.3. Let us emphasize that this smooth extrema is not needed for second-order approximation. Furthermore, considering third-order local subcell monolithic DG/FV scheme, the smoothness detection has to be performed at the cell level, instead of the subcell level, as in this case the second derivatives $\nabla_x\left(\partial_{x\backslash y} u_h^{c,\,n}\right)$ are constant over the cell.

5.2.2. Euler system and local maximum principle

Regarding the local maximum principle previously introduced, the natural system counterpart would be to apply the previous criteria to the Riemann invariants. However, in the non-linear system case, those quantities are not easy to get nor to manipulate. We could have used a linearized version of the Riemann invariants, as in [61] for instance, but for sake of simplicity we naively apply the local maximum principle to one of the conserved variable. This local maximum principle is then relaxed, by means of the same smooth extrema detector previously introduced, but this time based on the chosen conservative variable. In the numerical results Section 5.3, we choose to either work with the density or the total energy, as these physical quantities would be sensitive to any type of wave. Also, as theoretically there is no local maximum for the conserved variables, we add the FV Riemann intermediate state to the local bounds, as

$$\alpha_m^c := \min_{S_q^w \in \mathcal{N}(S_m^c)} \left(\overline{v}_q^{w,n},\, v_{mq}^{*,\mathrm{FV}} \right) \;\leq\; \overline{v}_m^{c,n+1} \;\leq\; \max_{S_q^w \in \mathcal{N}(S_m^c)} \left(\overline{v}_q^{w,n},\, v_{mq}^{*,\mathrm{FV}} \right) := \beta_m^c, \tag{42}$$

where $v \in \left\{ \rho,\, q^x,\, q^y,\, E \right\}$. The blending coefficient θ_{mp} is then defined as previously (38). Similarly to the SCL case, this local maximum principle has to be relaxed to preserve accuracy in the presence of smooth extrema.

5.3. Numerical results: global and local maximum principles

Similarly to the numerical results section devoted to entropy stability, Section 4.4, we make use here of several widely addressed and challenging test cases to demonstrate the performance and robustness of this local subcell monolithic DG/FV scheme ensuring Global Maximum Principle and positivity in the system case (GMP), as well as a relaxed Local Maximum Principle (LMP) to reduce the apparition of non-physical oscillations. In all following test cases, if not stated otherwise, the simple case of global Lax-Friedrichs numerical flux will be used for both the DG scheme and subcell first-order FV one. Regarding the cell decomposition into subcells, as it has been observed in [59], this does not have a major influence on the quality of the numerical results, especially in the non-linear case. Consequently, the simple case of the quad/tri subdivision, Figure 1(a), will be used if not specified differently. Last, for the 2D non-linear problems, we make use of the blending coefficients smoothening procedures, see Section 3.2. The less diffusive one, introduced in Section 3.1, will be used if not stated otherwise.

5.3.1. 1D linear advection case

As for the entropy stability section, we first consider the very simple case of 1D linear advection of composite signal. In Figure 10, the numerical solution obtained by means of the $\mathbb{P}^6$ monolithic DG/FV scheme ensuring GMP and a relaxed-LMP one, on a coarse grid made of only 40 cells, is displayed. One can see in Figure 10 how the scheme behaves, producing an extremely accurate solution while ensuring the preservation of the global maximum principle and a very low oscillatory profile. We can also observe how the smooth extrema relaxation has permitted to accurately capture the smooth part of the solution.

5.3.2. 1D Buckley non-convex case

Now, by means of the same two test cases used previously in the context of 1D non-convex Buckley SCL, we display in Figure 11 the numerical solutions obtained through the GMP and relaxed-LMP $\mathbb{P}^6$ monolithic DG/FV using only 40 cells. While in Section 4.4 we have shown that entropy stability for a given entropy is generally not enough to capture the unique entropic solution, in Figure 11 one

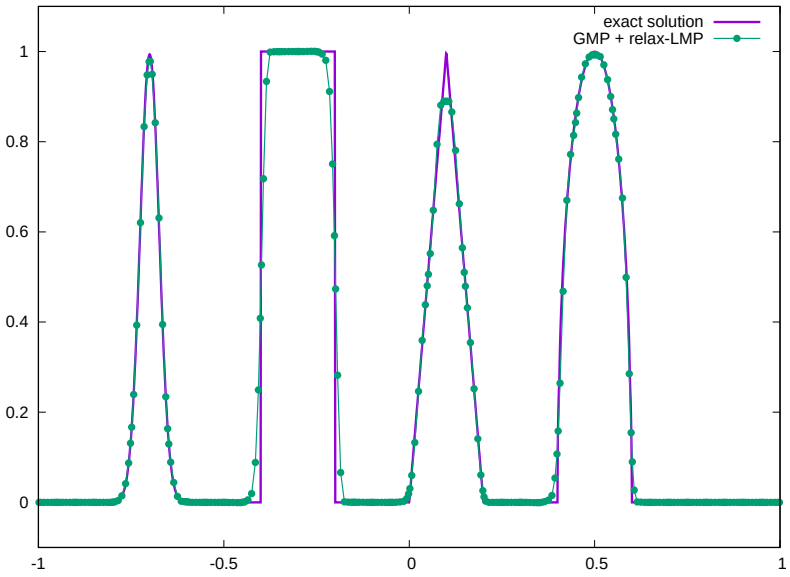
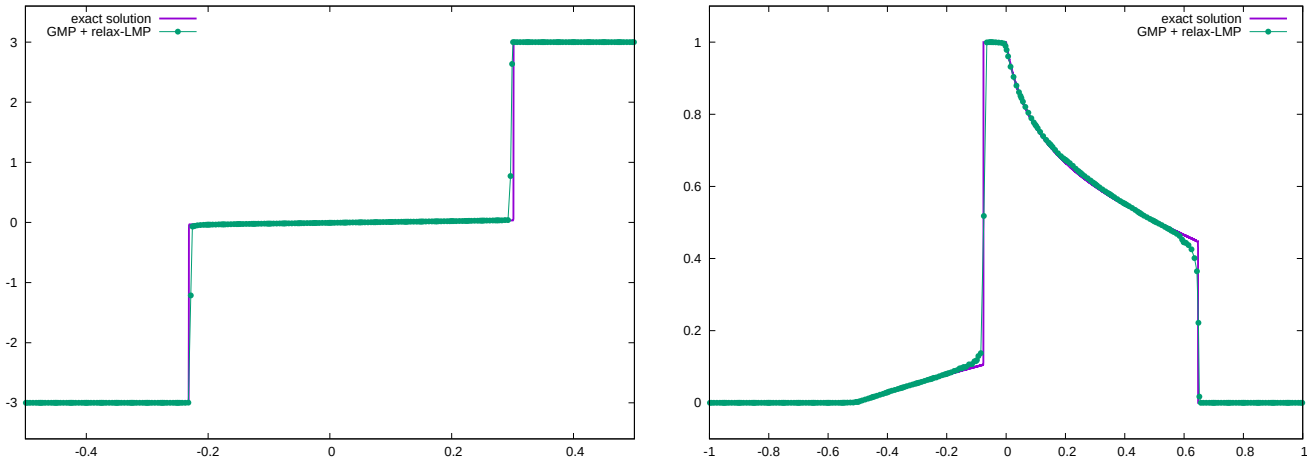


Figure 10: $\mathbb{P}^6$-DG/FV with GMP and relaxed-LMP on 40 cells



(a) Test case n°1 (b) Test case n°2

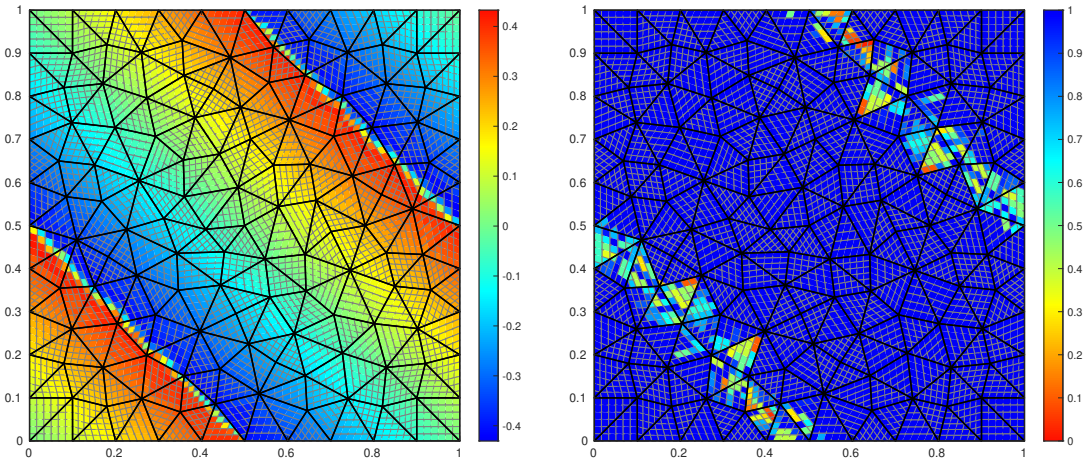
Figure 11: $\mathbb{P}^6$-DG/FV with GMP and relaxed-LMP on 40 cells

can see how the two maximum principles imposed here allow the very accurate and robust resolution of the problem under consideration, as even in this very coarse grid context the numerical solutions are extremely close to the exact entropic solutions.

5.3.3. 2D non-linear Burgers problem case

Now, to illustrated how the monolithic fluxes blending operates, let us make use of the Burgers equation, defined through (1a) and flux function $\boldsymbol{F}(u) = \frac{1}{2}\left(u^2,\, u^2\right)^{\mathrm{t}}$, with the smooth initial solution $u_0(\boldsymbol{x}) = \sin(2\pi\,(x+y))$. The domain is chosen as the unit square $[0,\,1]^2$ with periodic boundary condition. Through time, the exact solution will exhibit two stationary shocks along the lines

defined by $\big(\boldsymbol{x} \in [0,\,1]^2,\; x+y=0.5\big)$ and $\big(\boldsymbol{x} \in [0,\,1]^2,\; x+y=1.5\big)$. We run this test case until $t=0.5$ with a sixth-order monolithic DG/FV scheme ensuring GMP and a relaxed-LMP, on a very coarse unstructured grid made of 242 cells. In Figure 12(a), we display the subcells' mean values while in Figure 12(b) the subcell blending coefficients, see definition (15), are shown.



(a) Solution submean values

(b) Blending coefficients

Figure 12: $\mathbb{P}^5$-DG/FV scheme with GMP and relaxed-LMP on 242 cells

First, Figure 12(b) illustrates very well how the monolithic DG/FV scheme works and is able to accurately capture the discontinuities, as only the subcells in a small vicinity of the shocks will be computed trough a convex blending of high-order DG reconstructed fluxes and low-order FV fluxes. Elsewhere the blending coefficients are automatically set to one, which tells us that only the high-order reconstructed fluxes, which gives the equivalency with a pure DG scheme, are used.

5.3.4. 2D KPP non-convex case

Now, we consider once more the non-convex flux case of KPP SCL. A similar set up is used here, but whilst in Figure 8(b) cell entropy stability was enforced, in Figure 13 we make use of GLM and relaxed-LMP, combined with the blending coefficient smoother introduced in Section 3.2. While we have seen that the high-order entropy stable monolithic scheme fails to capture the entropic solution, see Figure 8(b), here the two maximum principles, GMP and relaxed-LMP, allows the correct approximation of the unique entropic solution, even in the difficult context of high-order schemes and coarse grids.

5.3.5. 1D modified Sod shock tube problem

Once more, we make use of the modified Sod shock tube problem. In Section 4.4.4, we have seen how even a pure DG scheme with a positivity-preserving limiter was able to capture the correct solution and that entropy stability increases robustness and reduces spurious oscillations. Now, in a similar configuration, instead of entropy stability, we display the solution obtained through the $\mathbb{P}^5$ monolithic DG/FV scheme ensuring positivity and relaxed-LMP, see Figure 14. Compared to Figure 9, one can observed on Figure 14 how those two principles allow us to obtain excellent results, as the numerical solution is non-oscillatory and extremely close to the exact one, even in this extreme coarse grid and very high order context. Let us however emphasize that this local subcell monolithic

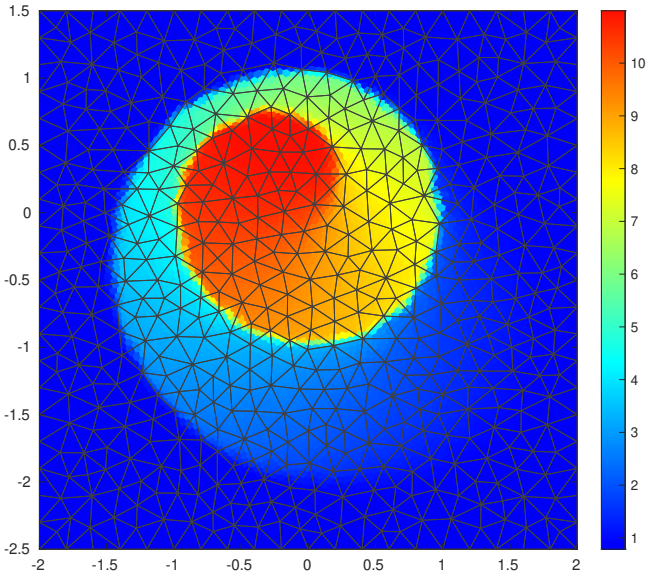
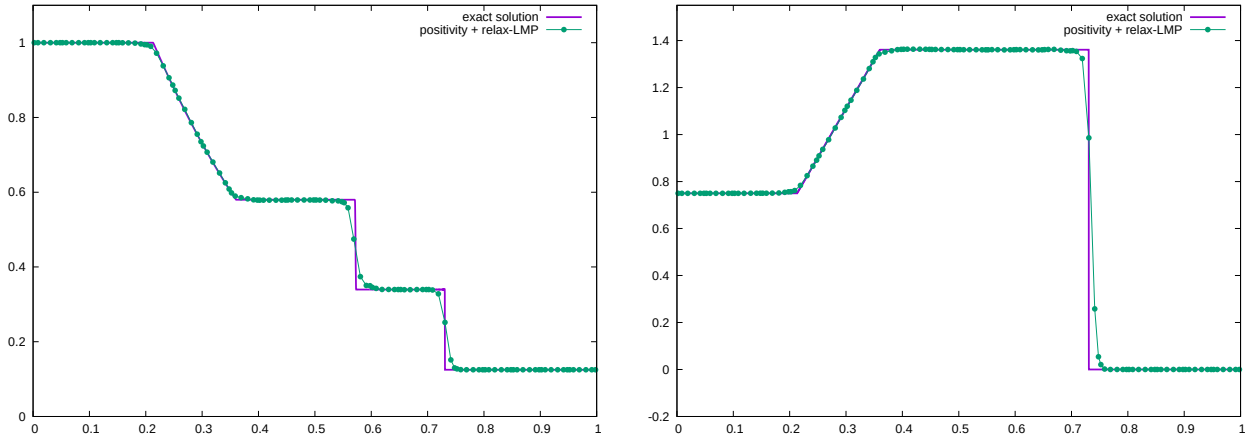


Figure 13: $\mathbb{P}^3$-DG/FV scheme with GMP and relaxed-LMP on 1054 cells



(a) Density (b) Velocity

Figure 14: $\mathbb{P}^5$-DG/FV scheme with positivity and relaxed-LMP on 20 cells

DG/FV scheme is obviously not limited to the case of very high-order of accuracy on coarse grids. It also preforms very well at second or third order, see Figure 15. In the light of Figure 15, one can see how the third-order monolithic scheme on 100 cells, with positivity and relaxed-LMP conditions, does produce a numerical solution very close to the exact one.

5.3.6. 1D smooth isentropic solution

To test the accuracy of the local subcell monolithic DG/FV scheme in the case of system, we make use of a smooth test case initially introduced in [60]. This example has been derived in the isentropic case, for the perfect gas equation of state with the polytropic index $\gamma = 3$. In this special situation, the characteristic curves of the Euler equations become straight lines and the governing equations

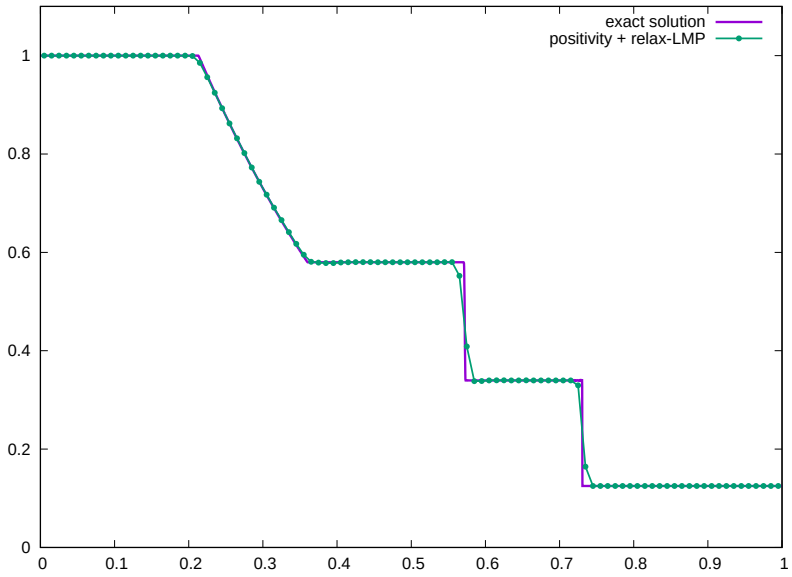


Figure 15: $\mathbb{P}^2$-DG/FV scheme with positivity and relaxed-LMP on 100 cells: cells' mean values

reduce to two Burgers equations. It is then simple to solve analytically this problem. Here, similarly to [58], we modify the initial data to yield a more challenging example, as

$$\rho^0(x) = 1 + 0.9999999\sin(\pi x), \quad u^0(x) = 0, \quad p^0(x) = \rho^0(x)^\gamma, \quad x \in [-1, 1],$$

provided with periodic conditions. This means that initially $\rho^0(-\frac{1}{2}) = 1.E-7$ and $p^0(-\frac{1}{2}) = 1.E-21$. The density and pressure being so close to zero, any numerical scheme not ensuring a positivity preservation would fail. This is the case of unlimited DG schemes. In Figure 16, numerical solutions obtained by means of the $\mathbb{P}^4$ monolithic DG/FV scheme with positivity and relaxed-LMP are depicted at time $t = 0.1$, using 20 cells. Figure 16 demonstrates how robust the monolithic scheme is as no loss of positivity is possible due to the theory, and how accurate the scheme is, the numerical solution being extremely close to the exact one even with only 20 cells. In Table 2, we gather the global errors and rates of convergence related to the 5th order scheme, along with the global minimum and the average, in space over the whole domain and in time covering the whole calculation, of the blending coefficients. The results confirm the expected fifth-order rate of convergence, even though the solution has been locally corrected.

	L_1		L_2		θ_{mp}	
h	$E^h_{L_1}$	$q^h_{L_1}$	$E^h_{L_2}$	$q^h_{L_2}$	min. θ_{mp}	aver. θ_{mp}
$\frac{1}{10}$	9.07E-4	5.86	1.23E-3	5.90	2.00E-1	0.981
$\frac{1}{20}$	1.56E-5	4.03	2.05E-5	3.83	1.92E-1	0.997
$\frac{1}{40}$	9.53E-7	4.89	1.44E-6	4.85	5.65E-4	0.999
$\frac{1}{80}$	3.21E-8	4.80	5.00E-8	4.87	3.48E-5	0.999
$\frac{1}{160}$	1.15E-9	-	1.71E-9	-	1.00	1.00

Table 2: Convergence rates computed on the pressure for the $\mathbb{P}^4$-DG/FV scheme with positivity and relaxed-LMP

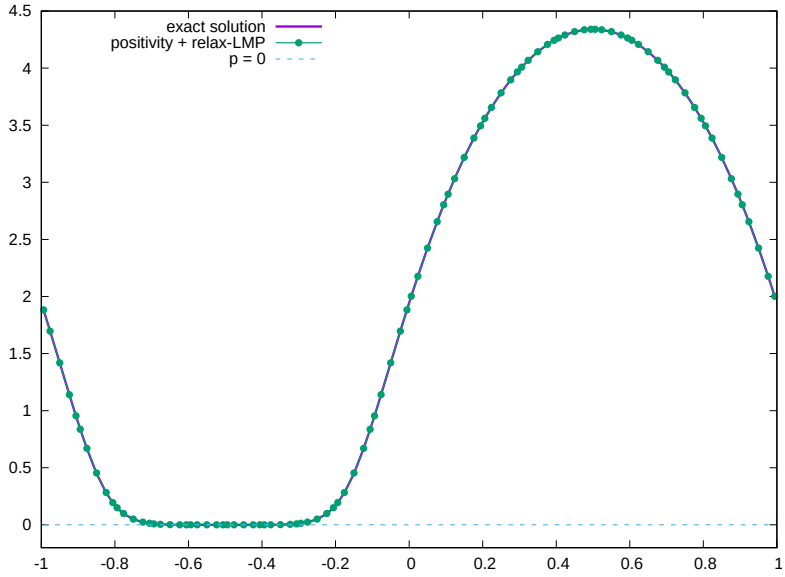
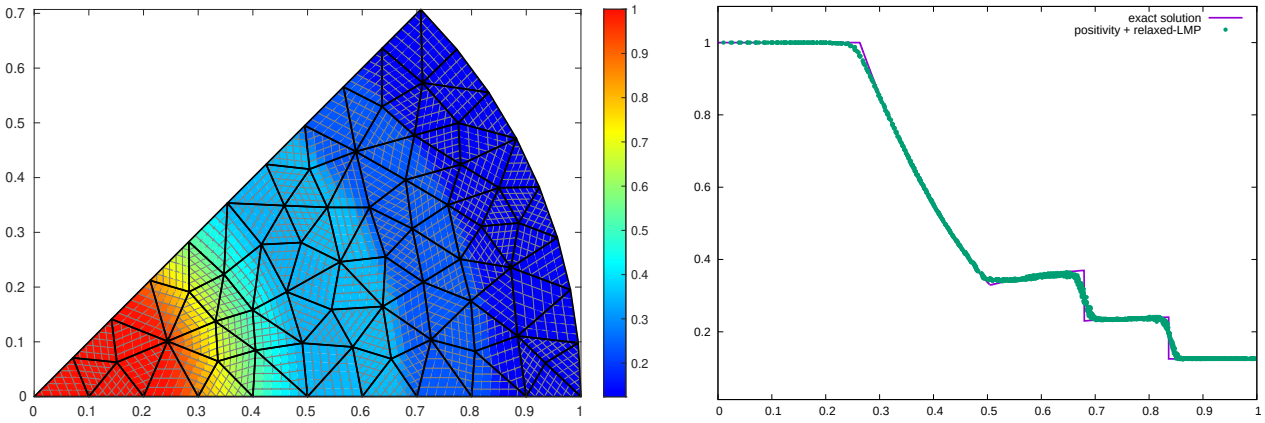


Figure 16: $\mathbb{P}^4$-DG/FV scheme with positivity and relaxed-LMP on 20 cells: pressure

We can also notice in the light of Table 2 that refining the mesh, the amount of first-order FV required to stabilize the scheme decreases more and more, as the minimum and averaged blending coefficients tend to one. One could have expected such conclusion since, in this smooth solution context, pure DG is expected to converge to the exact solution. Hence no blending should be required in the end.

5.3.7. 2D Sod shock tube problem

To close this numerical application section and assess once again the high capability of the local subcell monolithic DG/FV method presented here, the 2D Euler compressible gas dynamics system (33) case will be now addressed. First, we consider the extension of the classical Sod shock tube [52] to the case of the cylindrical geometry. This problem consists of a cylindrical shock tube of unity radius. The interface is located at $r = 0.5$. At the initial time, the states on the left and on the right sides of the interface are constant. The left state is a high pressure fluid characterized by $(\rho_0^L, p_0^L, \boldsymbol{v}_0^L) = (1, 1, \mathbf{0})$, the right state is a low pressure fluid defined by $(\rho_0^R, p_0^R, \boldsymbol{v}_0^R) = (0.125, 0.1, \mathbf{0})$. The gamma gas law is defined by $\gamma = \frac{7}{5}$. The computational domain is defined in polar coordinates by $(r, \theta) \in [0, 1] \times [0, \frac{\pi}{4}]$. We prescribe symmetry boundary conditions at the boundaries $\theta = 0$ and $\theta = \frac{\pi}{4}$, and an outflow condition at $r = 1$. The exact solution consists of three circular waves, a shock followed by a contact discontinuity and rarefaction wave. The aim of this test case is then to assess the local subcell monolithic DG/FV scheme accuracy while ensuring a non-oscillatory behavior, and its ability to preserve the radial symmetry. In Figure 17, the $\mathbb{P}^5$ monolithic DG/FV scheme with positivity and relaxed-LMP has been used on a very coarse anisotropic mesh made of only 110 triangular cells. In the light of Figure 17(a), one can see how the radial wave structure has been accurately capture, even in this extremely coarse mesh case, and how the three types of waves, meaning expansion wave, contact discontinuity and shock wave, travel and go through the large cells. Figure 17(b), where the subcells' mean values versus the subcell centroid radii $\sqrt{x^2 + y^2}$ are displayed, confirms this statement as the different points for a given radius do coincide.

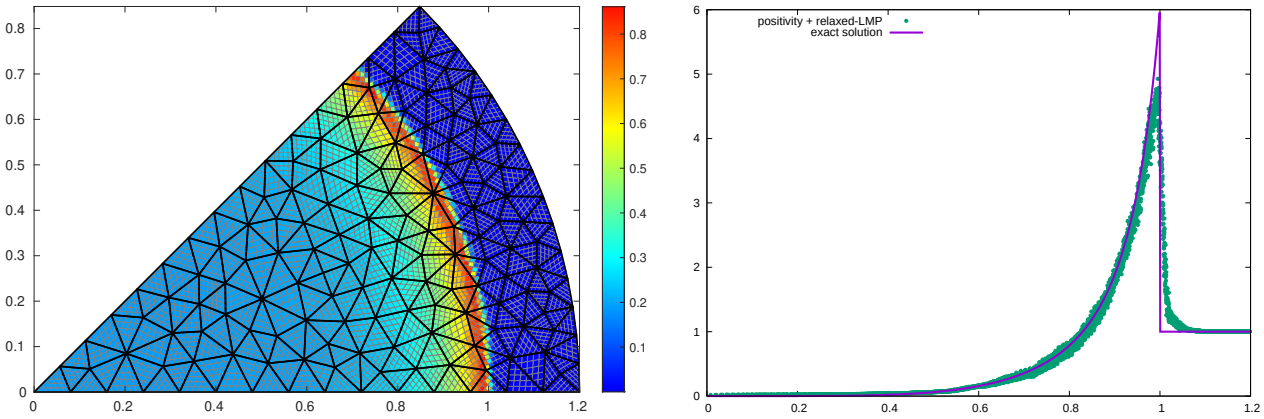


(a) Density map (b) Density profile

Figure 17: $\mathbb{P}^5$-DG/FV scheme with positivity and relaxed-LMP on a 110 cells mesh

5.3.8. 2D Sedov point blast problem

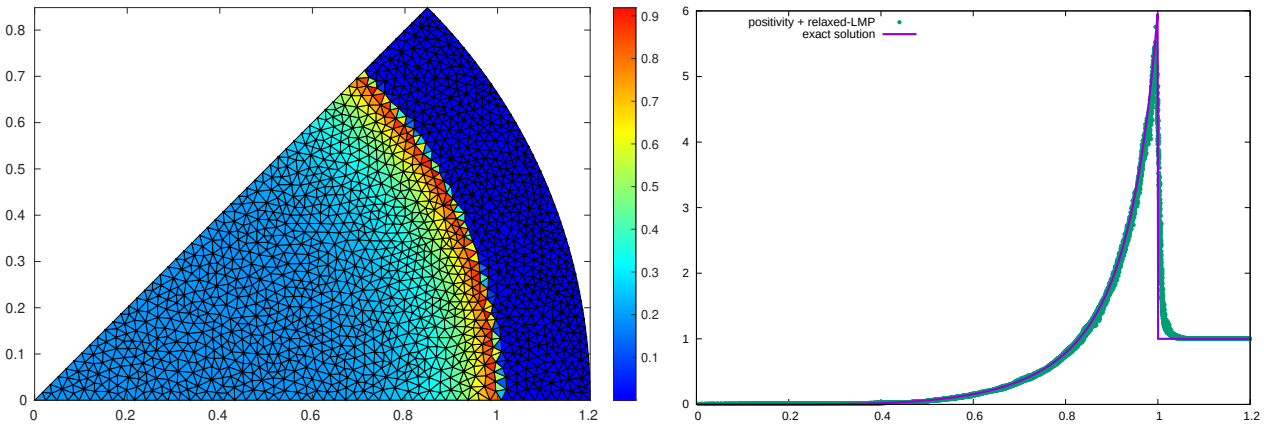
We consider the Sedov problem for a point-blast in a uniform medium. An exact solution based on self-similarity arguments is available, see for instance [32]. The initial conditions are characterized by $(\rho_0, p_0, \boldsymbol{v}_0) = (1, 10^{-14}, \mathbf{0})$, and the polytropic index is equal to $\frac{7}{5}$. We set an initial delta-function energy source at the origin prescribing the pressure in a control volume, yet to be defined, containing the origin as follows, $p_{or} = (\gamma - 1)\,\frac{\varepsilon_0}{v_{or}}$, where v_{or} denotes the measure of the chosen control volume and ε_0 the total amount of release energy. By choosing $\varepsilon_0 = 0.244816$, as suggested in [32], the solution consists of a diverging infinite strength shock wave whose front is located at radius $r = 1$ at $t = 1$, with a peak density reaching 6. The computational domain is defined in polar coordinates by $(r, \theta) \in [0,\, 1.2] \times [0, \frac{\pi}{4}]$. Similarly to the polar Sod shock tube problem, we prescribe symmetry boundary conditions at the boundaries $\theta = 0$ and $\theta = \frac{\pi}{4}$, and an outflow condition at $r = 1.2$. Regarding the control volume in which the delta-function energy will be dropped off, generally the cell containing the origin is considered. Here, similarly to [59] and to make this test case even more challenging, we choose to restrict the energy source only to the one subcell containing the origin. This means that initially, in one grid element the pressure in one subcell will be set to p_{or}, while in the remainder of the cell the pressure will be 10^{-14}. Let us further emphasize that generally in this test case, because one cannot simulate vacuum, the initial pressure is set to 10^{-6} over the domain, except at the origin. Here, to make it once again more challenging, we set the initial pressure to 10^{-14}. We run this modified Sedov point blast problem with the $\mathbb{P}^5$ monolithic DG/FV scheme, with positivity and relaxed-LMP conditions, on a very coarse grid made of 271 cells. In this particular case, the amount of total energy contained in the subcell located at the origin reaches 1947.5, while in the rest of the cell as well as in the remainder of the domain the total energy is set to 2.5E-14. Any scheme lacking positivity-preserving property or a rigorous stabilization technique would fail solving this test problem. In Figure 18, one can see how the circular aspect of the shock has been accurately captured by the scheme and the shock wave front is correctly located. This latter further goes inside and through different cells, enlightening the very robust and precise subcell resolution of this local subcell monolithic DG/FV method. The numerical solution produced remains quite close to the one-dimensional self-similar exact solution, see Figure 18(b).



(a) Energy map (b) Density profile

Figure 18: $\mathbb{P}^5$-DG/FV scheme with positivity and relaxed-LMP on a 271 cells mesh

Once again, this local subcell monolithic DG/FV scheme performs also very well for lower order methods, as depicted in Figure 19 where a finer grid made of 2894 cells has been used. Indeed, the $\mathbb{P}^2$ monolithic scheme solution is very close to the one-dimensional analytical solution. In Figure 19(a), only the cell total energy means values are represented, and not the submean values as we generally do, for a better readability of the results in this finer grid context.



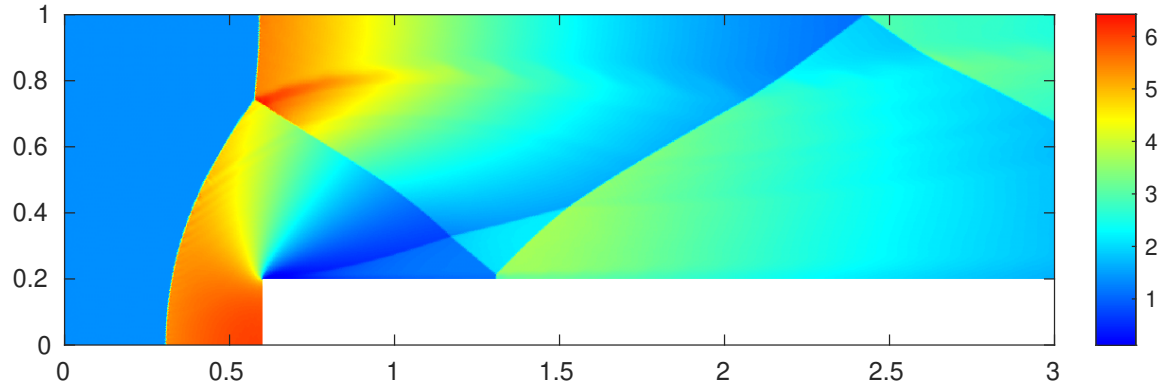
(a) Cells' mean total energy map (b) Subcells' mean density versus radii

Figure 19: $\mathbb{P}^2$-DG/FV scheme with positivity and relaxed-LMP on 2894 cells

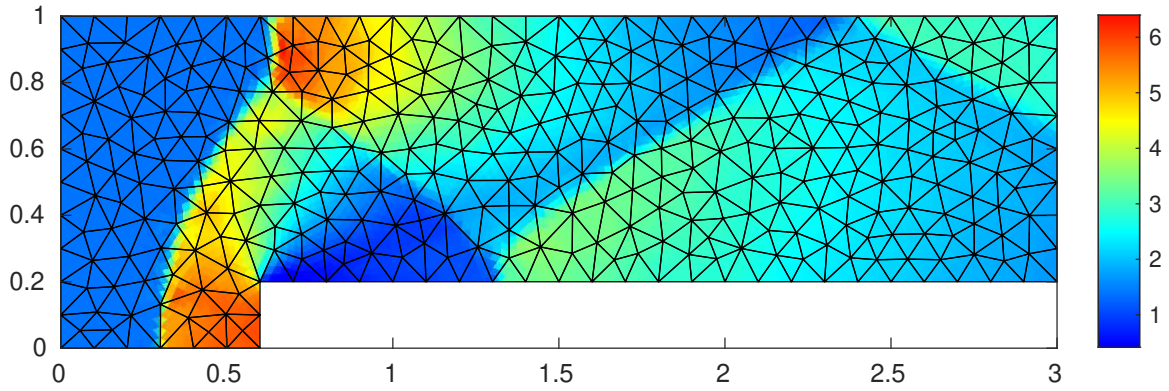
5.3.9. 2D forward-facing step problem

We now consider the forward facing step problem, which has been initially introduced by A. Emery in [14], and further studied by P. Woodward and P. Colella in [62]. This challenging test case consists in a Mach 3 flow in a 3 units in length and 1 unit in width wind tunnel. Initially, the tunnel is filled with a gamma gas law with $\gamma = \frac{7}{5}$, which everywhere has density $\rho_0 = 1.4$, pressure $p_0 = 1$ and velocity $\boldsymbol{v}_0 = (3,\, 0)^{\mathrm{t}}$. The 0.2 high step being located at $x = 0.6$, the computational domain is then $([0,\, 3] \times [0,\, 1]) \setminus ([0.6,\, 3] \times [0.2,\, 1])$. Gas with this density, pressure and velocity is continually

fed in from the left-hand boundary. Let us emphasize that unlike as it is generally done, we did not refine the mesh near the corner, see Figure 20(b) for instance, nor modify in any way our monolithic DG/FV scheme.



(a) $\mathbb{P}^1$ on 84108 cells



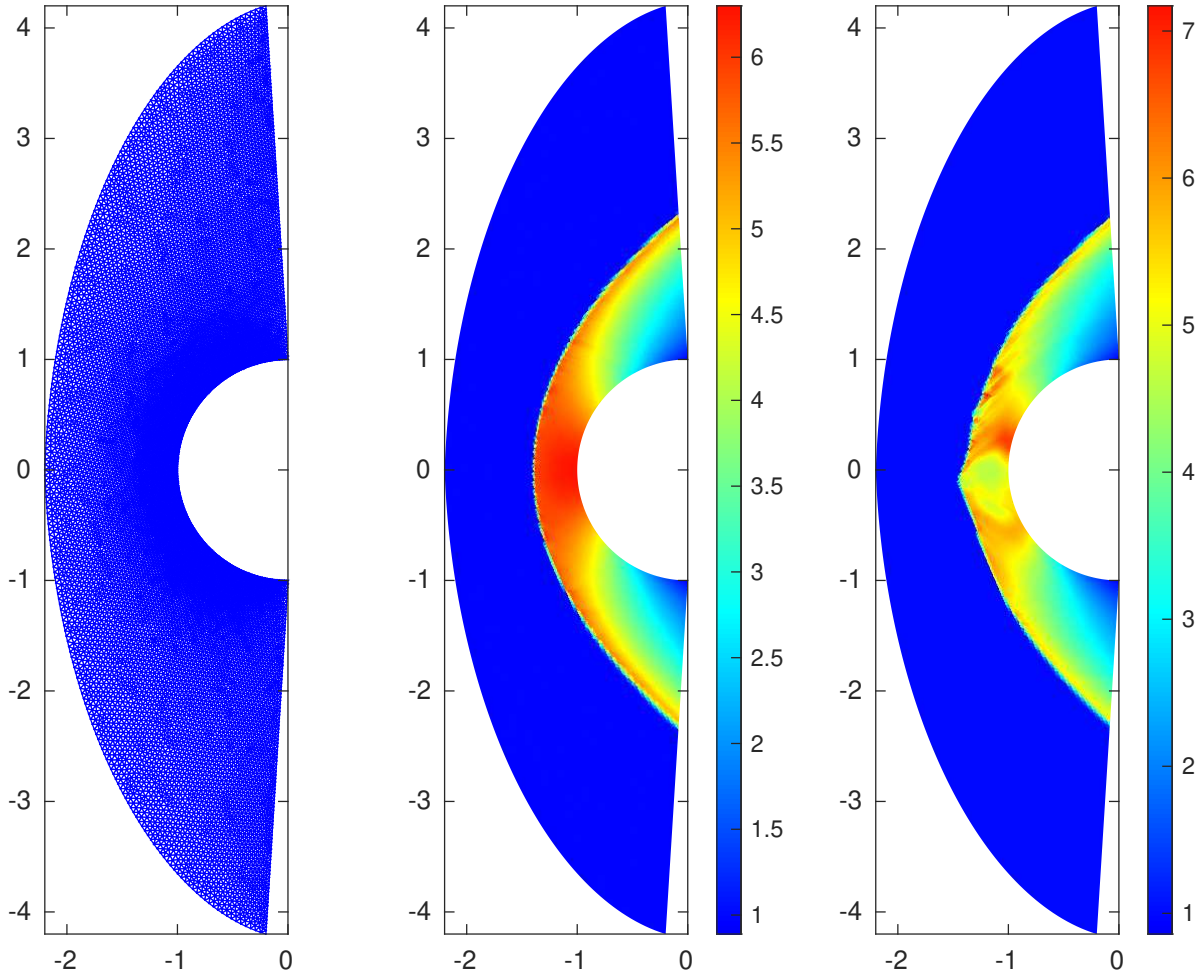
(b) $\mathbb{P}^3$ on 680 cells

Figure 20: Local subcell monolithic DG/FV scheme: subcells' density mean values

In Figure 20, we compare the numerical solutions obtained by means of our local subcell monolithic DG/FV scheme, ensuring positivity and a relaxed-LMP, respectively with a $\mathbb{P}^1$ on a fine grid made of 84108 cells and $\mathbb{P}^3$ on coarse grid made of 680 cells. Firstly, let us note that the $\mathbb{P}^1$ monolithic DG/FV scheme has produced a quite satisfactory solution close to the expected one and has been able to capture the Kelvin-Helmholtz instabilities in the top of the channel. Secondly, let us point out that due to the complexity of the flow, with multiple shocks waves and walls interacting, the benefit of high-order schemes over low order schemes may be limited. However, as depicted in Figure 20(b), high-order schemes on coarse grids may be a good solution to obtain the main features of the solution under investigation. Indeed, despite the coarseness of the mesh used in Figure 20(b), the shocks and the rarefaction fan created around the corner are quite well resolved, while ensuring a low oscillatory behavior. Now, to capture the finer structures of the solution, a finer mesh has to be used, keeping in mind that high-order monolithic DG/FV scheme will always outclass the lower-order ones, but being obviously more computationally costly.

5.3.10. 2D hypersonic flow over half cylinder problem

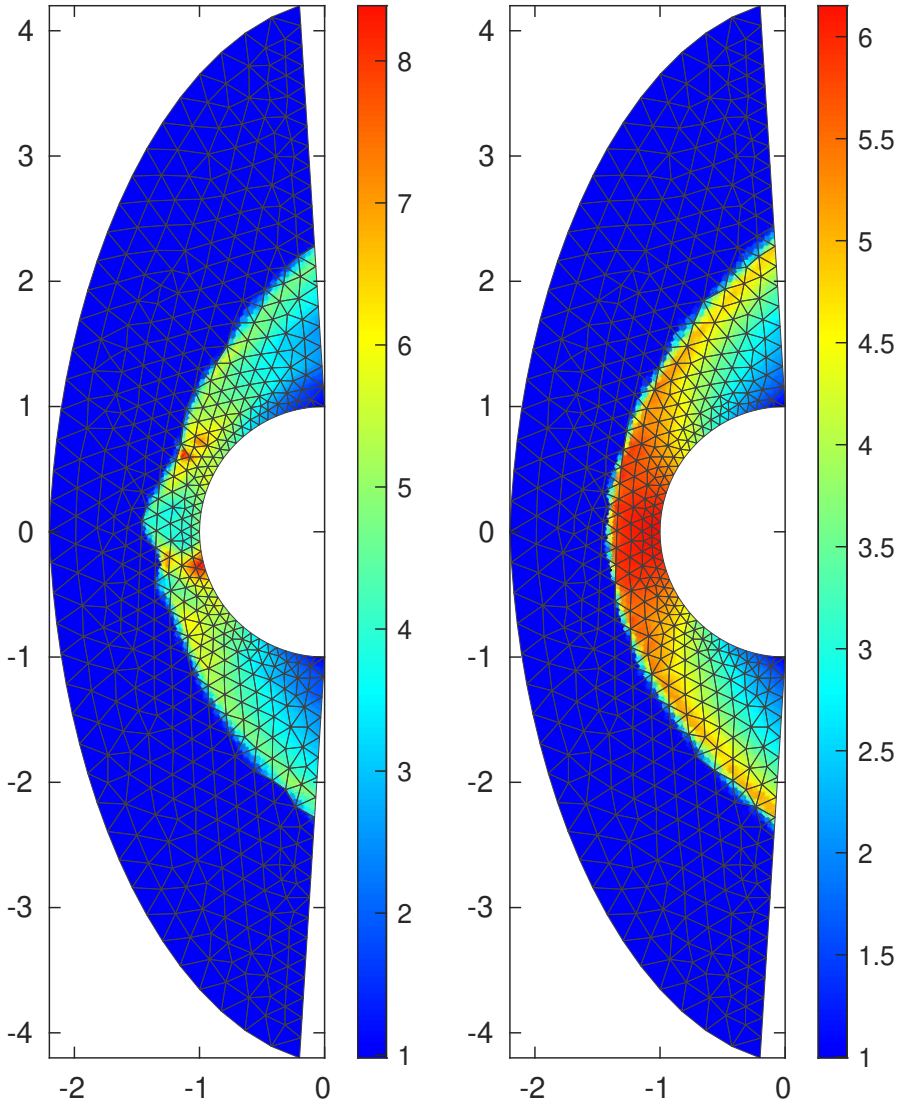
The hypersonic flow over a half-cylinder test case is a well-documented test case used to challenge numerical methods. Specifically, some schemes may develop the infamous carbuncle phenomenon, even using classical FV scheme. Instead of producing a smooth bow shock profile upstream of the half-cylinder, the carbuncle issue manifests as a pair of oblique shocks ahead of the stagnation region, compromising the overall flow predictions around the cylinder. Following the approach in [47], we simulate an inviscid flow at Mach $M_a = 20$ around a half-cylinder blunt body subjected to an incoming hypersonic flow characterized by $(\rho_i,\, p_i,\, \boldsymbol{v}_i) = \left(1,\, 1,\, (M_a\sqrt{\gamma},\, 0)^{\mathrm{t}}\right)$ with $\gamma = \frac{7}{5}$. The steady-state resulting flow is simulated using an explicit time-marching procedure, ending at time $t = 2.5$. The computational domain is sufficiently large, containing half of a cylinder centered at the origin with a radius $r = 1$ and a left incoming hypersonic flow. At the cylinder surface, a wall-type boundary condition is applied, while the bottom and upper boundary conditions are free outflow and an inflow condition is applied at the left boundary. First, to exhibit the so-called carbuncle effect, we display in Figure 21 the numerical solutions obtained using a first-order FV scheme, based respectively on HLL and HLL-C numerical fluxes, on a fine grid made of 25266 cells, see Figure 21(a).



(a) Grid used (b) HLL numerical flux (c) HLL-C numerical flux

Figure 21: 1st-order FV scheme on a grid made of 25266 cells with HLL and HLL-C numerical fluxes

As expected, the use of HLL-C numerical flux does trigger the carbuncle effect, see Figure 21(c), while the use of HLL numerical flux does not, see Figure 21(b). Now, in addition to assess the robustness of our monolithic scheme in this Mach 20 hypersonic flow context, we want to assess how the carbuncle effect translates going to higher orders of accuracy. To do so, we run our $\mathbb{P}^2$ monolithic DG/FV scheme, ensuring positivity and a relaxed-LMP, on a coarse mesh made of 1044 cells, and display the solutions in Figure 22.

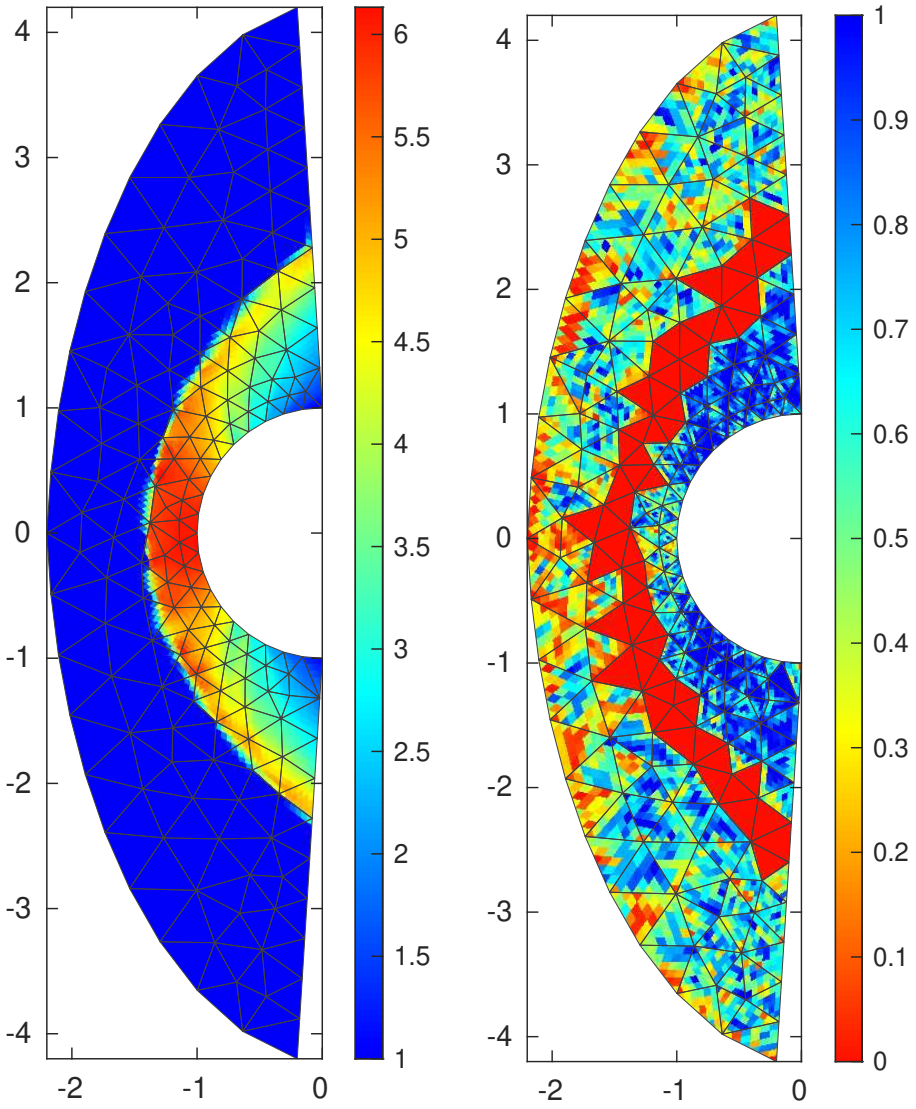


(a) HLL-C/HLL-C (b) HLL-C/HLL

Figure 22: $\mathbb{P}^2$-DG/FV scheme with positivity and relaxed-LMP on 1044 cells: density

Let us first emphasize that it is absolutely not required to use the same numerical flux to compute the DG residual in (7) and for the first-order FV flues in (9). Thus, to see the repercussion of such choices on the potential trigger of the carbuncle effect, we compare in Figure 22 the solutions obtained by means of our $\mathbb{P}^2$ monolithic DG/FV scheme respectively using HLL-C numerical flux for both DG (hence the reconstructed fluxes) and the first-order FV fluxes in the first case, Figure 22(a),

and HLL-C for DG and HLL for the FV fluxes in the second case, Figure 22(b). In Figure 22(a), one can see that even going to high order of accuracy, the use of HLL-C numerical flux does trigger the carbuncle effect. However, if in the monolithic scheme, the first-order numerical method forming the safe base of this blended scheme uses HLL numerical flux, the approximated solution will not present any carbuncle effect. This highlights the fact that, with appropriate choices for the DG and FV numerical fluxes, it should be possible to ensure some properties on the high-order parts of solution and others on the low-order parts of the solution. Finally, to test once more the high capability and robustness of this monolithic scheme going to very high-order of accuracy and very coarse meshes, even in the simulation of this complex high Mach hypersonic flow, we show in Figure 23 the solution obtained by means of the $\mathbb{P}^5$ monolithic DG/FV scheme, ensuring positivity and a relaxed-LMP, on an extremely coarse mesh made of only 292 cells.



(a) HLL-C/HLL: density (b) HLL-C/HLL: blending coeff.

Figure 23: $\mathbb{P}^5$-DG/FV scheme with positivity and relaxed-LMP on 292 cells

In the light of Figure 23(a), one can see once more how robust and accurate the monolithic scheme is, taking into account the extreme coarseness of the grid used, as the scheme has been able to capture the bow structure of the shock. Furthermore, the blending coefficients displayed on Figure 23(b) allow us to illustrate what has been said in Remark 5.2. Indeed, one can see that, in cells containing the bow shock, every subcells are computed through a first-order FV scheme, as the θ_m^c are equal to zero everywhere in the cell. This shows that the reconstructed fluxes may have developed non-admissible values, revealing that DG scheme has totally failed in those cells. This generally comes from the computation of square root of negative values in the evaluation of the DG numerical fluxes. But as long as the first-order scheme is robust, the monolithic and its associated simulation code could not crash.

6. Conclusion

This article is concerned with the construction of a new type of monolithic scheme, based on general unstructured grids, blending locally at the subcell level DG scheme and first-order FV scheme. This subcell blending procedure relies on the expression of DG methods as a finite volume scheme on a subgrid. By means of this theoretical part, we combine in a convex manner, at the subcell level, the so-called reconstructed fluxes and first-order FV numerical fluxes, through a blending coefficient θ_{mp}. The next step is then to determine those blending coefficients to achieve all the desired properties. In the first part of this paper, we have first focused our attention of the issue of entropy stability. Different types of entropy stabilities were hence introduced, along with the corresponding blending coefficients. In particular, a semi-discrete cell entropy stability, for a given entropy, has been proposed which allows the preservation of the high-order accuracy of the local subcell monolithic DG/FV scheme. However, the numerical results showed that this entropy criterion might be, in some complex case, too relaxed to capture the unique entropic weak solution. Furthermore, the cost of such entropy condition is very high as the first-order FV numerical fluxes have to be specifically modified, which may lead to the loss of other desired properties as positivity for instance. In a second part, we focus on the imposition of different maximum principles, a global one to ensure the numerical solution to remain in a convex admissible set, and a local one to address the issues of spurious oscillations. A wide number of test cases on different problems have been used to depict the very good performance and robustness of the presented monolithic scheme using those principles.

In a very near future, we expect to apply those monolithic schemes to the problematic of coastal flows simulation and their coupling with a moving object. We also intend to extend this local subcell monolithic DG/FV framework to multi-dimensional approximated Riemann solvers, and then to moving grid configurations, both in ALE and Lagrangian formalisms. Finally, we plan as well to extend the present scheme to the three-dimensional case, as the theory developed should remain exactly the same.

Appendices

A. Some properties on FV schemes and E-fluxes

This appendix aims at recalling some properties yield by a FV scheme

$$\overline{u}_m^{c,\,n+1} = \overline{u}_m^{c,\,n} - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \mathcal{F}\left(\overline{u}_m^{c,\,n}, \overline{u}_p^{v,\,n}, \boldsymbol{n}_{mp}\right), \tag{A.1}$$

relying on a general numerical flux (3), in the simple case of SCL. All the following properties can be easily extended to systems and other types of fluxes as HLL for instance.

A.1. Discrete maximum principle

First, let us show that following discrete maximum principle holds

$$\min\left(\overline{u}_m^{c,\,n}, \min_{S_p^v \in \mathcal{V}_m^c} \overline{u}_p^{v,\,n}\right) \leq \overline{u}_m^{c,\,n+1} \leq \max\left(\overline{u}_m^{c,\,n}, \max_{S_p^v \in \mathcal{V}_m^c} \overline{u}_p^{v,\,n}\right). \tag{A.2}$$

Following the steps presented in Section 3, scheme (A.1) can be recast into the following Godunov-type form

$$\overline{u}_m^{c,n+1} = \left(1 - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\right) \overline{u}_m^{c,n} + \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\, u_{mp}^{*,\,\mathrm{FV}}, \tag{A.3}$$

where $u_{mp}^{*,\,\mathrm{FV}} := u^*(\overline{u}_m^{c,n}, \overline{u}_p^{v,n}, \boldsymbol{n}_{mp})$, the FV Riemann intermediate state, writes as

$$u^*(u_L, u_R, \boldsymbol{n}) = \frac{u_L + u_R}{2} - \frac{\left(\boldsymbol{F}(u_R) - \boldsymbol{F}(u_L)\right) \boldsymbol{\cdot}\, \boldsymbol{n}}{2\,\gamma(u_L, u_R, \boldsymbol{n})}. \tag{A.4}$$

A sufficient condition to ensure (A.2) is then to show that $u^*(u_L, u_R, \boldsymbol{n})$ lies in $I(u_L, u_R)$. This condition is evident, as

$$u^*(u_L, u_R, \boldsymbol{n}) = u_L \left(\frac{1}{2} + \frac{\beta}{2\,\gamma}\right) + u_R \left(\frac{1}{2} - \frac{\beta}{2\,\gamma}\right),$$

with $\gamma := \gamma(u_L, u_R, \boldsymbol{n})$ and $\beta = \frac{(\boldsymbol{F}(u_R) - \boldsymbol{F}(u_L)) \boldsymbol{\cdot}\, \boldsymbol{n}}{u_R - u_L}$. Since $\gamma \geq \max_{w \,\in\, I(u_L, u_R)}(|\boldsymbol{F}'(w) \boldsymbol{\cdot}\, \boldsymbol{n}|)$, it directly follows that $|\beta| \leq \gamma$. The intermediate state u^* hence writes as a convex combination of u_L and u_R.

A.2. Discrete entropy stability

Now, let us recall that the discrete scheme (A.1) is entropy stable for any entropy, as it implies the following inequality

$$\eta\big(\overline{u}_m^{c,\,n+1}\big) \leq \eta\big(\overline{u}_m^{c,\,n}\big) - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \phi^*\left(\overline{u}_m^{c,\,n}, \overline{u}_p^{v,\,n}, \boldsymbol{n}_{mp}\right), \tag{A.5}$$

with function ϕ^*, a consistent numerical entropy flux, such that $\phi^*(u, u, \boldsymbol{n}) = \boldsymbol{\phi}(u) \boldsymbol{\cdot}\, \boldsymbol{n}$. To be coherent with the numerical flux $\mathcal{F}$ under consideration (3), we define the numerical entropy flux as

$$\phi^*(u_L, u_R, \boldsymbol{n}) = \frac{\left(\boldsymbol{\phi}(u_L) + \boldsymbol{\phi}(u_R)\right) \boldsymbol{\cdot}\, \boldsymbol{n}}{2} - \frac{\gamma(u_L, u_R, \boldsymbol{n})}{2} \big(\eta(u_R) - \eta(u_L)\big). \tag{A.6}$$

To demonstrate (A.5), let us first see that, by convexity of the entropy, the convex relation (A.3) leads to

$$\eta\left(\overline{u}_m^{c,n+1}\right) \leq \eta\left(\overline{u}_m^{c,n}\right) - \frac{\Delta t}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} l_{mp}\, \gamma_{mp}\, \left(\eta\left(\overline{u}_m^{c,n}\right) - \eta\left(u_{mp}^{*,\mathrm{FV}}\right)\right). \tag{A.7}$$

Thus, it directly follows that a sufficient condition to obtain (A.5) is

$$-\gamma_{mp}\left(\eta\big(\overline{u}_m^{c,n}\big) - \eta\big(u_{mp}^{*,\mathrm{FV}}\big)\right) \leq -\Big(\phi^*\big(\overline{u}_m^{c,n}, \overline{u}_p^{v,n}, \boldsymbol{n}_{mp}\big) - \boldsymbol{\phi}\big(\overline{u}_m^{c,n}\big) \cdot \boldsymbol{n}_{mp}\Big).$$

By means of the numerical entropy flux (A.6), this sufficient condition can be reformulated as follows

$$\eta\big(u^*(u_L, u_R, \boldsymbol{n})\big) \leq \frac{\eta(u_L) + \eta(u_R)}{2} - \frac{\big(\boldsymbol{\phi}(u_R) - \boldsymbol{\phi}(u_L)\big) \cdot \boldsymbol{n}}{2\,\gamma(u_L, u_R, \boldsymbol{n})}. \tag{A.8}$$

To ensure that condition (A.8) is indeed guaranteed, let us introduce the following Riemann problem

$$\begin{cases} \partial_t\, u(\boldsymbol{x}, t) + \nabla_x \cdot \boldsymbol{F}\left(u(\boldsymbol{x}, t)\right) = 0, & (\boldsymbol{x}, t) \in \mathbb{R}^d \times \mathbb{R}^+, \qquad \text{(A.9a)} \\ u(\boldsymbol{x}, 0) = \begin{cases} u_L & \text{if } (\boldsymbol{x} \cdot \boldsymbol{n}) < 0, \\ u_R & \text{if } (\boldsymbol{x} \cdot \boldsymbol{n}) > 0. \end{cases} & \qquad \text{(A.9b)} \end{cases}$$

Introducing rotated space variables $\xi_n = (\boldsymbol{x} \cdot \boldsymbol{n})$ and $\xi_\tau = (\boldsymbol{x} \cdot \boldsymbol{\tau})$, with $\boldsymbol{n}$ a given unit normal and $\boldsymbol{\tau} = \boldsymbol{n}^\perp$, and due to the invariance by rotation of SCL, equation (A.9a) rewrites as

$$\partial_t\, u + \partial_{\xi_n}\left(\boldsymbol{F}(u) \cdot \boldsymbol{n}\right) + \partial_{\xi_\tau}\left(\boldsymbol{F}(u) \cdot \boldsymbol{\tau}\right) = 0.$$

Finally, because u_0 does only depends on ξ_n, the solution u does as well, and the PDE reduces to the following 1D problem

$$\partial_t\, u + \partial_{\xi_n} F_n(u) = 0, \tag{A.10}$$

where $F_n(u) = (\boldsymbol{F}(u) \cdot \boldsymbol{n})$. Now, let us show that $u^*(u_L, u_R, \boldsymbol{n})$, defined in (A.4), is nothing but the average value of $\mathcal{W}\big(\frac{\xi_n}{t};\, u_L,\, u_R\big)$, the unique entropic weak solution of the considered Riemann problem, as long as $\gamma := \gamma(u_L, u_R, \boldsymbol{n}) \geq \max_{w \in I(u_L, u_R)}(|\boldsymbol{F}'(w) \cdot \boldsymbol{n}|)$. In this case, the waves produced by the initial discontinuity will remain left and right bounded by respectively $\xi_n = \pm\,\gamma\, t$, as depicted by Figure A.24. First, integrating equation (A.10) onto the time-space box $[-\gamma\, t, \gamma\, t] \times [0, \Delta t]$, displayed in Figure A.24, one gets

$$\frac{1}{2\,\gamma\,\Delta t} \int_{-\gamma\, t}^{\gamma\, t} \mathcal{W}\Big(\frac{\xi_n}{\Delta t};\, u_L,\, u_R\Big)\, \mathrm{d}\,\xi_n = \frac{u_L + u_R}{2} - \frac{\big(\boldsymbol{F}(u_R) - \boldsymbol{F}(u_L)\big) \cdot \boldsymbol{n}}{2\,\gamma} := u^*. \tag{A.11}$$

Because $\mathcal{W}\big(\frac{\xi_n}{t};\, u_L,\, u_R\big)$, the unique entropic weak solution, lies in $I(u_L, u_R)$, it furthers demonstrates that $u^* \in I(u_L, u_R)$ as well. Now, let us recall that the unique solution $u(\boldsymbol{x}, t) := \mathcal{W}\big(\frac{\xi_n}{t};\, u_L,\, u_R\big)$ ensures, in a weak sens, the following entropic inequalities

$$\partial_t\, \eta(u) + \partial_{\xi_n} \phi_n(u) \leq 0, \tag{A.12}$$

for any couple entropy - entropy flux $(\eta, \boldsymbol{\phi})$, where $\phi_n(u) = (\boldsymbol{\phi}(u) \cdot \boldsymbol{n})$. Similarly as before, integrating (A.12) onto $[-\gamma\, t, \gamma\, t] \times [0, \Delta t]$, it follows that

$$\frac{1}{2\,\gamma\,\Delta t} \int_{-\gamma\, t}^{\gamma\, t} \eta\left(\mathcal{W}\Big(\frac{\xi_n}{\Delta t};\, u_L,\, u_R\Big)\right)\, \mathrm{d}\,\xi_n \leq \frac{\eta(u_L) + \eta(u_R)}{2} - \frac{\big(\boldsymbol{\phi}(u_R) - \boldsymbol{\phi}(u_L)\big) \cdot \boldsymbol{n}}{2\,\gamma}. \tag{A.13}$$

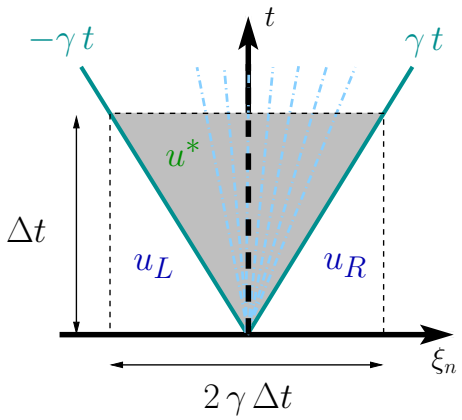


Figure A.24: Riemann fan

Finally, due to entropy convexity and using Jensen's inequality

$$\eta\,(u^*) := \eta\left(\frac{1}{2\,\gamma\,\Delta t}\int_{-\gamma\,t}^{\gamma\,t}\mathcal{W}\Big(\frac{\xi_n}{\Delta t};\,u_L,\,u_R\Big)\,\mathrm{d}\,\xi_n\right) \leq \frac{1}{2\,\gamma\,\Delta t}\int_{-\gamma\,t}^{\gamma\,t}\eta\left(\mathcal{W}\Big(\frac{\xi_n}{\Delta t};\,u_L,\,u_R\Big)\right)\,\mathrm{d}\,\xi_n,$$

relation (A.13) reduces to the desired condition (A.8).

A.3. Two-point Tadmor relation

Here, we show that the use of a numerical flux (3) with $\gamma(u_L, u_R, \boldsymbol{n}) \geq \max_{w\,\in\,I(u_L,u_R)}(|\boldsymbol{F}'(w)\,\boldsymbol{.}\,\boldsymbol{n}|)$, for which we have just displayed how it ensures a discrete entropy inequality for any entropy, also guarantees the Tadmor inequality (25). Such inequality ensures the semi-discrete FV scheme to be entropy stable for a given entropy, [55, 56]. To this end, we will exhibit how a numerical flux (3) can be put into the following form, with $D \geq 0$

$$\mathcal{F}\,(u_L, u_R, \boldsymbol{n}) = \frac{\boldsymbol{\Psi}(v_R) - \boldsymbol{\Psi}(v_L)}{v_R - v_L}\,\boldsymbol{.}\,\boldsymbol{n} - \frac{D}{2}\,(v_R - v_L), \tag{A.14}$$

where $v_{L/R} = v(u_{L/R})$. If the entropy viscosity dissipation coefficient $D = 0$, the semi-discrete FV scheme will be entropy conservative, while being entropy dissipative for $D > 0$. The combination of definitions (3) and (A.14) states that the entropy dissipation coefficient is given by

$$\begin{aligned} D &= \gamma\,\left(\frac{u_R - u_L}{v_R - v_L}\right) + \frac{2}{(v_R - v_L)^2}\left(\boldsymbol{\psi}_R - \boldsymbol{\psi}_L - \frac{(\boldsymbol{F}_L + \boldsymbol{F}_R)}{2}(v_R - v_L)\right)\boldsymbol{.}\,\boldsymbol{n}, \\ &= \gamma\,\left(\frac{u_R - u_L}{v_R - v_L}\right) + \frac{2}{(v_R - v_L)^2}\left(\int_{v_L}^{v_R}\boldsymbol{\Psi}'(v)\,\mathrm{d}\,v - \frac{(\boldsymbol{F}_R + \boldsymbol{F}_L)}{2}\int_{u_L}^{u_R} v'(u)\,\mathrm{d}\,u\right)\boldsymbol{.}\,\boldsymbol{n}, \end{aligned}$$

where $\gamma := \gamma(u_L, u_R, \boldsymbol{n})$, while $\boldsymbol{\psi}_{L/R} = \boldsymbol{\psi}(u_{L/R}) = \boldsymbol{\Psi}(v_{L/R})$ and $\boldsymbol{F}_{L/R} = \boldsymbol{F}(u_{L/R})$. By definition of the entropy potential flux, we have that $\boldsymbol{\Psi}'(v(u)) = \boldsymbol{F}(v(u))$. By a change of variable in the first

integral, it follows that

$$
\begin{aligned}
D &= \gamma \left(\frac{u_R - u_L}{v_R - v_L}\right) + \frac{2}{(v_R - v_L)^2} \left(\int_{u_L}^{u_R} \boldsymbol{F}(u)\, v'(u)\, \mathrm{d}\, u - \frac{(\boldsymbol{F}_R + \boldsymbol{F}_L)}{2} \int_{u_L}^{u_R} v'(u)\, \mathrm{d}\, u \right) \boldsymbol{.}\, \boldsymbol{n}, \\
&= \gamma \left(\frac{u_R - u_L}{v_R - v_L}\right) - \frac{1}{(v_R - v_L)^2} \int_{u_L}^{u_R} \boldsymbol{n} \boldsymbol{.} \left(\boldsymbol{F}_R + \boldsymbol{F}_L - 2\,\boldsymbol{F}(u)\right) v'(u)\, \mathrm{d}\, u, \\
&= \left(\frac{u_R - u_L}{v_R - v_L}\right)^2 \frac{1}{u_R - u_L} \int_{u_L}^{u_R} \underbrace{\left(\gamma - \frac{\left(\boldsymbol{F}_R + \boldsymbol{F}_L - 2\,\boldsymbol{F}(u)\right) \boldsymbol{.}\, \boldsymbol{n}}{u_R - u_L}\right)}_{\Gamma(u)} v'(u)\, \mathrm{d}\, u.
\end{aligned}
$$

Finally, thanks to the entropy convexity, $v'(u) > 0$, if $\forall\, u \in I(u_L, u_R), \Gamma(u) \geq 0$ then $D \geq 0$. This is actually the case because

$$
\begin{aligned}
\Gamma(u) &\geq \gamma - \left| \frac{\left(\boldsymbol{F}_R + \boldsymbol{F}_L - 2\,\boldsymbol{F}(u)\right) \boldsymbol{.}\, \boldsymbol{n}}{u_R - u_L} \right| \geq \gamma - \frac{\left|\left(\boldsymbol{F}_R - \boldsymbol{F}(u)\right) \boldsymbol{.}\, \boldsymbol{n}\right| + \left|\left(\boldsymbol{F}_L - \boldsymbol{F}(u)\right) \boldsymbol{.}\, \boldsymbol{n}\right|}{|u_R - u_L|}, \\
&= \gamma - \left(\left| \frac{u - u_L}{u_R - u_L} \right| \left| \frac{\left(\boldsymbol{F}(u) - \boldsymbol{F}_L\right) \boldsymbol{.}\, \boldsymbol{n}}{u - u_L} \right| + \left| \frac{u_R - u}{u_R - u_L} \right| \left| \frac{\left(\boldsymbol{F}_R - \boldsymbol{F}(u)\right) \boldsymbol{.}\, \boldsymbol{n}}{u_R - u} \right| \right), \\
&\geq \gamma - \left(\frac{|u - u_L| + |u_R - u|}{|u_R - u_L|} \right) \max_{w \in I(u_L, u_R)} (|\boldsymbol{F}'(w) \boldsymbol{.}\, \boldsymbol{n}|), \\
&= \gamma - \max_{w \in I(u_L, u_R)} (|\boldsymbol{F}'(w) \boldsymbol{.}\, \boldsymbol{n}|).
\end{aligned}
$$

Under the condition that $\gamma \geq \max_{w \in I(u_L, u_R)}(|\boldsymbol{F}'(w) \boldsymbol{.}\, \boldsymbol{n}|)$, the numerical flux (3) can indeed be expressed as in (A.14) with an entropy dissipation coefficient $D \geq 0$.

References

[1] P. Batten, N. Clarke, C. Lambert, and Causon. On the choice of wavespeeds for the HLLC Riemann solver. *SIAM J. Sci. Comput.*, 18:1553–1570, 1997.

[2] R. Biswas, K. Devine, and J.E. Flaherty. Parallel adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14:255–284, 1994.

[3] J.-P. Boris and D.-L. Book. Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works. *J. Comp. Phys.*, 11(1):38–69, 1973.

[4] A. Burbeau, P. Sagaut, and C.-H. Bruneau. A problem-independent limiter for high-order Runge Kutta discontinuous Galerkin methods. *J. Comp. Phys.*, 169:111–150, 2001.

[5] V. Carlier and F. Renac. Invariant domain preserving high-order spectral discontinuous approximations of hyperbolic systems. *SIAM J. Sci. Comput.*, 45(3):A1385–A1412, 2023.

[6] M. H. Carpenter, T. Fisher, E. Nielsen, and S. Frankel. Entropy Stable Spectral Collocation Schemes for the Navier–Stokes Equations: Discontinuous Interfaces. *SIAM J. Sci. Comput.*, 36:B835–B867, 2014.

[7] J. Chan. On discretely entropy conservative and entropy stable discontinuous galerkin methods. *J. Comp. Phys.*, 362:346–374, 2018.

[8] T. Chen and C.-W. Shu. Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *J. Comp. Phys.*, 345:427–461, 2017.

[9] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for hyperbolic systems: Multi-dimensional Optimal Order Detection (MOOD). *J. Comp. Phys.*, 230:4028–4050, 2011.

[10] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: The multidimensional case. *Math. Comp.*, 54:545–581, 1990.

[11] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta Discontinuous Galerkin Method for Conservation Laws V: Multidimensional Systems. *J. Comp. Phys.*, 141:199–224, 1998.

[12] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework. *Math. Comp.*, 52:411–435, 1989.

[13] M. Dumbser and R. Loubère. A simple robust and accurate a posteriorisub-cell finite volume limiter for the discontinuousGalerkin method on unstructured meshes. *J. Comp. Phys.*, 319:163–199, 2016.

[14] A. Emery. An evaluation of several differencing methods for inviscid flow problems. *J. Comp. Phys.*, 2(3):306–331, 1968.

[15] F. Renac. Entropy stable, robust and high-order dgsem for the compressible multicomponent euler equations. *J. Comp. Phys.*, 445:110584, 2021.

[16] T. C. Fisher and M. H. Carpenter. High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains. *J. Comp. Phys.*, 252:518–557, 2013.

[17] G. Gassner. A Skew-Symmetric Discontinuous Galerkin Spectral Element Discretization and Its Relation to SBP-SAT Finite Difference Methods. *SIAM J. Sci. Comput.*, 35:1233–1253, 2013.

[18] G. J. Gassner, A. R.Winters, and D. A. Kopriva. Split form nodal discontinuous Galerkin schemes with summation-by-parts property for the compressible Euler equations. *J. Comp. Phys.*, 327:39–66, 2016.

[19] J.-L. Guermond, M. Nazarov, B. Popov, and I. Tomas. Second-order invariant domain preserving approximation of the euler equations using convex limiting. *SIAM Journal on Scientific Computing*, 40(5):A3211–A3239, 2018.

[20] J.-L. Guermond and B. Popov. Invariant domains and first-order continuous finite element approximation for hyperbolic systems. *SIAM J. Numer. Anal.*, 54(4):2466–2489, 2016.

[21] J.-L. Guermond, B. Popov, and I. Tomas. Invariant domain preserving discretization-independent schemes and convex limiting for hyperbolic systems. *Comput. Methods Appl. Mech. and Engrg.*, 347:143–175, 2019.

[22] A. Haidar, F. Marche, and F. Vilar. A posteriori finite-volume local subcell correction of high-order discontinuous Galerkin schemes for the nonlinear shallow-water equations. *J. Comp. Phys.*, 452:110902, 2022.

[23] A. Haidar, F. Marche, and F. Vilar. Free-boundary problems for wave structure interactions in shallow-water: Dg-ale description and local subcell correction. *J. Sci. Comput.*, 98(2), 2024.

[24] H. Hajduk. Monolithic convex limiting in discontinuous galerkin discretizations of hyperbolic conservation laws. *Computers and Mathematics with Applications*, 87:120–138, 2021.

[25] J.S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer Publishing Company, Incorporated, 2007.

[26] J.E. Hicken, D.C. Del Rey Fernández, and D.W. Zingg. Multidimensional summation-by-parts operators: General theory and application to simplex elements. *SIAM J. Sci. Comput.*, 38(4):A1935–A1958, 2016.

[27] S. Hou and X.-D. Liu. Solutions of Multi-dimensional Hyperbolic Systems of Conservation Laws by Square Entropy Condition Satisfying Discontinuous Galerkin Method. *J. Sci. Comput.*, 31:127–151, 2007.

[28] A. Huerta, E. Casoni, and J. Peraire. A simple shock-capturing technique for high-order discontinuous Galerkin methods. *Int. J. Numer. Meth. Fluids*, 69:1614–1632, 2012.

[29] J. S. Park and S.-H. Yoon and C. Kim. Multi-dimensional limiting process for hyperbolic conservation laws on unstructured grids. *J. Comp. Phys.*, 229:788–812, 2010.

[30] G.-S. Jiang and C.-W. Shu. On cell entropy inequality for discontinuous galerkin method for a scalar hyperbolic equation. *Mathematics of Computation*, 62:531–538, 1994.

[31] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted eno schemes. *J. Comp. Phys.*, 126:202–228, 1996.

[32] J.R. Kamm and F.X. Timmes. On efficient generation of numerically robust Sedov solutions. Technical Report LA-UR-07-2849, Los Alamos National Laboratory, 2007.

[33] L. Krivodonova. Limiters for high-order discontinuous Galerkin methods. *J. Comp. Phys.*, 226:879–896, 2007.

[34] A. Kurganov, G. Petrova, and B. Popov. Adaptive semi-discrete central-upwind schemes for non convex hyperbolic conservation laws. *SIAM J. Sci. Comput.*, 29:2381–2401, 2007.

[35] D. Kuzmin. A vertex-based hierarchical slope limiter for p-adaptative discontinuous Galerkin methods. *J. Comp. Appl. Math.*, 233:3077–3085, 2009.

[36] D. Kuzmin. Monolithic convex limiting for continuous finite element discretizations of hyperbolic conservation laws. *Comput. Methods Appl. Mech. and Engrg.*, 361:112804, 2020.

[37] D. Kuzmin and M.Q. de Luna. Subcell flux limiting for high-order bernstein finite element discretizations of scalar hyperbolic conservation laws. *J. Comp. Phys.*, 411:109411, 2020.

[38] D. Kuzmin, M.Q. de Luna, D.I. Ketcheson, and J. Grüll. Bound-preserving flux limiting for high-order explicit Runge Kutta time discretizations of hyperbolic conservation laws. *J. Sci. Comput.*, 91, 2022.

[39] D. Kuzmin and F. Schieweck. A parameter-free smoothness indicator for high-resolution finite element schemes. *Centr. Eur. J. Math.*, 11:1478–1488, 2013.

[40] R. J. LeVeque. High-resolution conservative algorithms for advection in compressible flow. *SIAM J. Numer. Anal.*, 33:627–665, 1996.

[41] L. Li and Q. Zhang. A new vertex-based limiting approach for nodal discontinuous Galerkin methods on arbitrary unstructured meshes. *Computers and Fluids*, 159:316–326, 2017.

[42] Y. Lin and J. Chan. High order entropy stable discontinuous galerkin spectral element methods through subcell limiting. *J. Comp. Phys.*, 498:112677, 2024.

[43] C. Lohmann, D. Kuzmin, J.N. Shadid, and S. Mabuza. Flux-corrected transport algorithms for continuous Galerkin methods based on high order Bernstein finite elements. *J. Comp. Phys.*, 344:151–186, 2017.

[44] S. Osher. Riemann solvers, the entropy condition and difference approximations. *SIAM J. Numer. Anal.*, 21:217–235, 1984.

[45] W. Pazner. Sparse invariant domain preserving discontinuous galerkin methods with subcell convex limiting. *Comput. Methods Appl. Mech. and Engrg.*, 382:113876, 2021.

[46] W. H. Reed and T. R. Hill. Triangular Mesh Methods for the Neutron Transport Equation. Technical Report LA-UR-73-479, Los Alamos National Laboratory, 1973.

[47] A.V. Rodionov. Artificial viscosity Godunov-type schemes to cure the carbuncle phenomenon. *J. Comp. Phys.*, 345:308–329, 2017.

[48] A. Rueda-Ramìrez, B. Bolm, D. Kuzmin, and G. Gassner. Monolithic convex limiting for legendre-gauss-lobatto discontinuous galerkin spectral-element methods. *Commun. Appl. Math. Comput.*, 2024.

[49] A.M. Rueda-Ramìrez, W. Pazner, and G.J. Gassner. Subcell limiting strategies for discontinuous galerkin spectral element methods. *Computers and Fluids*, 247:105627, 2022.

[50] S. Hennemann and A.M. Rueda-Ramìrez and F.J. Hindenlang and G.J. Gassner. A provably entropy stable subcell shock capturing approach for high order split form dg for the compressible euler equations. *J. Comp. Phys.*, 426:109935, 2021.

[51] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comp. Phys.*, 77:439–471, 1988.

[52] G. A. Sod. A survey of several finite difference methods for systems of non-linear hyperbolic conservation laws. *J. Comp. Phys.*, 27:1–31, 1978.

[53] M. Sonntag and C. D. Munz. Shock capturing for discontinuous Galerkin methods using finite volume subcells. In *Finite Volumes for Complex Applications VII*, pages 945–953. Springer, 2014.

[54] E. Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Mathematics of Computation*, 168(43):369–381, 1984.

[55] E. Tadmor. The numerical viscosity of entropy stable schemes for systemes of conservation laws. i. *Mathematics of Computation*, 179(49):91–103, 1987.

[56] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12:451–512, 2003.

[57] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics*. Springer-Verlag, 1999.

[58] F. Vilar. A posteriori correction of high-order discontinuous Galerkin scheme through subcell finite volume formulation and flux reconstruction. *J. Comp. Phys.*, 387:245–279, 2018.

[59] F. Vilar and R. Abgrall. A Posteriori Local Subcell Correction of High-Order Discontinuous Galerkin Scheme for Conservation Laws on Two-Dimensional Unstructured Grids. *SIAM J. Sci. Comput.*, 46(2):A851–A883, 2024.

[60] F. Vilar, P.-H. Maire, and R. Abgrall. Cell-centered discontinuous Galerkin discretizations for two-dimensional scalar conservation laws on unstructured grids and for one-dimensional Lagrangian hydrodynamics. *Computers and Fluids*, 46(1):498–604, 2011.

[61] F. Vilar, P.-H. Maire, and R. Abgrall. A discontinuous Galerkin discretization for solving the two-dimensional gas dynamics equations written under total Lagrangian formulation on general unstructured grids. *J. Comp. Phys.*, 276:188–234, 2014.

[62] P. Woodward and P. Collela. The numerical-simulation of two-dimensional fluid-flow with strong shocks. *J. Comp. Phys.*, 54(1):115–173, 1984.

[63] K. Wu and C.-W. Shu. Geometric quasilinearization framework for analysis and design of bound-preserving schemes. *SIAM Review*, 65(4):1031–1073, 2023.

[64] M. Yang and Z.J. Wang. A parameter-free generalized moment limiter for high-order methods on unstrucured grids. *Adv. Appl. Math. Mech.*, 4:451–480, 2009.

[65] S.T. Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comp. Phys.*, 31(3):335–362, 1979.

[66] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comp. Phys.*, 229:3091–3120, 2010.

[67] X. Zhang, Y. Xia, and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high order discontinuous galerkin schemes for conservation laws on triangular meshes. *J. Sci. Comput.*, 50:29–62, 2012.