

# High order entropy preserving ADER-DG schemes

Elena Gaburro<sup>a,\*</sup>, Philipp Öffner<sup>b</sup>, Mario Ricchiuto<sup>a</sup>, Davide Torlo<sup>c</sup>

<sup>a</sup>*Inria, Univ. Bordeaux, CNRS, Bordeaux INP, IMB, UMR 5251, 200 Avenue de la Vieille Tour, 33405 Talence cedex, France*

<sup>b</sup>*Institut für Mathematik, Johannes Gutenberg Universität, Staudingerweg 9, 55099 Mainz, Germany*

<sup>c</sup>*SISSA Mathlab, Mathematics Area, SISSA, via Bonomea 265, 34136, Trieste, Italy*

---

## Abstract

In this paper, we develop a fully discrete entropy preserving ADER-Discontinuous Galerkin (ADER-DG) method. To obtain this desired result, we equip the space part of the method with entropy correction terms that balance the entropy production in space, inspired by the work of Abgrall. Whereas for the time-discretization we apply the relaxation approach introduced by Ketcheson that allows to modify the timestep to preserve the entropy to machine precision. Up to our knowledge, it is the first time that a provable fully discrete entropy preserving ADER-DG scheme is constructed. We verify our theoretical results with various numerical simulations.

*Keywords:* Hyperbolic Conservation Laws, ADER, Discontinuous Galerkin, Entropy Conservation/Stability, Structure Preserving, Relaxation Method

---

## 1. Introduction

In last years, the development of arbitrary high-order entropy preserving<sup>1</sup> discretizations has attracted a lot of attention in community developing numerical methods for hyperbolic PDEs. Following the seminal work by Tadmor [65, 66], several ideas have surfaced providing general strategies to have control on the discrete production of entropy. Many works were focused on finding an entropy stable semidiscretization (in space), isolating the entropy conservative flux. These techniques have been applied to many applications, *inter alia* for shallow water equations [35, 71, 50], for Euler's equation [51, 34, 13, 56, 55, 19], Navier Stokes problems [33, 72, 20, 47, 46], magneto hydro-dynamics problems [14], multiphase or multicomponent problems [18, 58, 48] and generally for hyperbolic problems [2, 4, 42, 43, 15, 10, 27, 28, 41, 38, 57], also in the Lagrangian framework [11, 22, 12]. More recently also fully discrete entropy stable or entropy conservative schemes have been introduced, e.g. [40, 54, 3, 5, 29, 44, 49, 52]. These techniques exploit quite different underlying mechanisms, e.g. artificial viscosity, limiting techniques, the summation-by-parts (SBP) framework, entropy correction terms, relaxation ansatzes or combination of those. Even if the entropy dissipative methods differ significantly in their underlying ideas, they have demonstrated excellent numerical properties in numerical simulations.

The ADER approach has proved to be an accurate numerical scheme in numerous engineering applications. It has also been shown to be an appropriate setting to design discretizations able to preserve different physical properties. First proposed in [68, 70] as an high-order extension of the classical Godunov scheme, the method has been dramatically extended and re-interpreted in different ways, among which we recall [23, 61, 25, 9, 8, 7, 26, 24, 21, 32, 16, 36]. However, up to this point no provable high-order entropy conservative/dissipative ADER method can be found in the literature. Entropy conservation/stability is however essential to ensure that our numerical approximations converge to the physical relevant weak entropy solution. In the following work, we will consider this fact and develop for the first time an entropy preserving ADER-DG method. Therefore, we adapt two general techniques to the ADER-DG

---

\*Corresponding author

*Email addresses:* elena.gaburro@inria.fr (Elena Gaburro), poeffner@uni-mainz.de (Philipp Öffner), mario.ricchiuto@inria.fr (Mario Ricchiuto), davide.torlo@sissa.it (Davide Torlo)

<sup>1</sup>We use entropy preserving to describe either entropy conservative or dissipative (stable) discretization depending on the problem which is considered.

framework. Inspired by [2, 5], we handle the space-discretization of our ADER-DG method via entropy correction terms whereas the entropy production in time is balanced through the relaxation approach proposed in [40, 53, 54].

The paper is organized as follows. In Section 2, we briefly describe hyperbolic conservation laws and the concept of entropy. Then, in Section 3, the ADER-DG method is introduced in its nowadays used explicit space-time Finite Element (FE) interpretation. In Section 4, we describe the applied entropy correction techniques and the relaxation approach used in our ADER-DG framework resulting in a provably entropy preserving method. Next, in Section 5 we verify our theoretical results via various numerical simulations. A conclusion in Section 6 finishes this paper.

## 2. Hyperbolic Conservation Laws

Various natural phenomena and engineering problems can be described by the following system of hyperbolic conservation laws

$$\partial_t \mathbf{u} + \operatorname{div} \mathbf{F}(\mathbf{u}) = 0, \quad (1)$$

where  $\mathbf{u} : \Omega \times \mathbb{R}^+ \rightarrow \Psi \subset \mathbb{R}^m$  denotes the conserved variable depending on the space coordinate  $\mathbf{x} \in \Omega \subset \mathbb{R}^D$  and time coordinate  $t \in \mathbb{R}^+$ ,  $\mathbf{F} : \Psi \rightarrow \mathbb{R}^{m \times D}$  is the flux function and  $\operatorname{div}$  is the divergence operator. As usual, system (1) is equipped with appropriate boundary and initial conditions. In the scalar case we will use the symbol  $u$  for the conservative variable. The paper focuses on the two-space dimensional problems (1), i.e.  $\Omega \subset \mathbb{R}^2$ . However note that all the ideas presented readily apply to the general case. In the simulation section, we will show [tests](#) on scalar linear advection equation, shallow water equations and gas dynamic Euler's equations. All the details on the equations are described in that section.

Due to the non-uniqueness of weak solutions of (1), the concept of entropy as an additional admissible criterion is necessary. Following the classical paper [37],  $\eta : \Psi \rightarrow \mathbb{R}$  is a convex entropy function with related entropy flux  $\mathbf{G}$  and entropy variable  $\mathbf{v} = \partial_{\mathbf{u}} \eta$  verifying  $\langle \mathbf{v}, \partial_{\mathbf{u}} \mathbf{F}_d \rangle = \partial_{\mathbf{u}} \mathbf{G}_d^T$ . Smooth solutions of (1) fulfill additionally the entropy conservation law

$$\partial_t \eta + \operatorname{div} \mathbf{G} = 0, \quad (2)$$

while admissible weak solutions are characterized by the entropy inequality

$$\partial_t \eta + \operatorname{div} \mathbf{G} \leq 0. \quad (3)$$

If  $\eta$  is strictly convex, the mapping between the entropy variables  $\mathbf{v} = \partial_{\mathbf{u}} \eta$  and the conservative variables  $\mathbf{u}$  is one-to-one. We can either work directly with  $\mathbf{v}$  or  $\mathbf{u}$  when solving (1). Moreover, it holds that the Hessian  $\partial_{\mathbf{u}\mathbf{u}} \eta$  (or Jacobian  $\partial_{\mathbf{u}} \mathbf{v}$ ) is a symmetric positive definite (SPD) matrix and we denote its inverse by  $A_0 := \partial_{\mathbf{v}} \mathbf{u}$ . Finally, our numerical approximations will be constructed in a way to solve (1) and fulfill the entropy conditions (2) or (3) in a global manner. Introducing the total entropy  $\mathcal{E}$  defined by the integral

$$\mathcal{E}(\mathbf{u}) = \int_{\Omega} \eta(\mathbf{u}) \mathrm{d}\mathbf{x}, \quad (4)$$

total entropy conservation stems immediately from (2) and reads

$$\partial_t \mathcal{E}(\mathbf{u}) + \int_{\partial\Omega} \mathbf{G} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = 0 \quad (5)$$

for smooth solutions, while (5) becomes an inequality ( $\leq$ ) in presence of shocks simply following (3). Our objective is to provide a strategy allowing to design ADER-DG schemes capable to mimic the last identity within machine precision.

## 3. High order ADER Discontinuous Galerkin schemes

### 3.1. Notation

This section introduces the basic notation used throughout the paper. As mentioned before, we follow the recent interpretation of ADER-DG as an explicit space-time FE discretization. We divide the spatial domain  $\Omega$  into cells  $\Omega_i$

(in our case triangles). The numerical solution is represented by piecewise polynomials of degree  $N \geq 0$  at time level  $t^n$  and it is defined by

$$\mathbf{u}_h^n(\mathbf{x}, t^n) = \sum_{\ell=0}^{N-1} \phi_\ell(\mathbf{x}) \hat{\mathbf{u}}_{\ell,i}^n, \quad \mathbf{x} \in \Omega_i, \quad (6)$$

where  $\phi_\ell(\mathbf{x})$  is a modal spatial basis and  $N$  denotes the total number of expansion coefficients  $\hat{\mathbf{u}}_i^n$ . The basis function  $\phi_\ell(\mathbf{x})$  is defined through a Taylor series of degree  $N$  in the cell  $\Omega_i$ . Let  $\mathbf{x}_{\mathbf{b}_i}$  be the barycenter of  $\Omega_i$  and  $h_i$  be the radius of the circumscribed circle around the triangle  $\Omega_i$ . The basis in  $\Omega_i$  is given by

$$\phi_\ell(\mathbf{x}) = \frac{(x - x_{\mathbf{b}_i})^{p_\ell} (y - y_{\mathbf{b}_i})^{q_\ell}}{p_\ell! h_i^{p_\ell} q_\ell! h_i^{q_\ell}}, \quad \ell = 0, \dots, N-1, \quad p_\ell, q_\ell \geq 0, \quad 0 \leq p_\ell + q_\ell \leq N \quad \text{with } N = \frac{1}{d!} \prod_{m=1}^d (N+m). \quad (7)$$

**Remark 3.1** (Extension to the general  $P_N P_M$  framework). *If we choose  $N = 0$ , we obtain as usual a first order FV method. Instead of increasing now the order of accuracy in one cell by setting  $N > 0$ , one can, for instance, apply a higher-order FV reconstruction up to order  $M + 1$ . Therefore, our method is included in the more general class of  $P_N P_M$  schemes proposed in [23] and further developed in [32, 31] which unifies higher-order FV and DG schemes into a common framework. Since the application of the entropy correction term inside a FV scheme is more problematic as described in [5], we restrict ourself in this paper on the ADER-DG approach only. Extensions to general  $P_N P_M$  are planned for future work.*

Different from classical Runge-Kutta discontinuous Galerkin method, the ADER-DG approach is not build on a method of **lines** (MOL) ansatz but on a local predictor-corrector approach. The advantage of the predictor-corrector procedure is that the information exchange between the elements is done only once at each time step, only during the corrector step. This results in advantages in runtime and efficiency of the scheme [6]. Therefore, we also define a space-time basis which locally approximates our predictor  $\mathbf{q}_h$  inside the element  $C_i^n = \Omega_i \times [t^n, t^{n+1}]$ . It is given by

$$\mathbf{q}_h^n(\mathbf{x}, t) = \sum_{\ell=0}^{Q-1} \theta_\ell(\mathbf{x}, t) \hat{\mathbf{q}}_\ell^n, \quad (\mathbf{x}, t) \in C_i^n, \quad (8)$$

where

$$\theta_\ell(x, y, t)|_{C_i^n} = \frac{(x - x_{\mathbf{b}_i})^{p_\ell} (y - y_{\mathbf{b}_i})^{q_\ell} (t - t^n)^{r_\ell}}{p_\ell! h_i^{p_\ell} q_\ell! h_i^{q_\ell} r_\ell! h_i^{r_\ell}}, \quad (9)$$

$$\ell = 0, \dots, Q, \quad p_\ell, q_\ell, r_\ell \geq 0, \quad 0 \leq p_\ell + q_\ell + r_\ell \leq N \quad \text{with } Q = \frac{1}{(d+1)!} \prod_{m=1}^{d+1} (N+m)$$

is our *modal space-time* basis for the polynomials of degree  $N$ . Here,  $Q$  is the number of degrees of freedom in one space-time cell  $C_i$ .

### 3.2. ADER-DG method

ADER-DG schemes consist in two steps for the update of one timestep: the local predictor and the corrector, which we describe now separately.

#### 3.2.1. Local Predictor Step

The method starts by calculating a high-order local predictor solution of the underlying PDE (1) using a classical Galerkin approach inside each element  $C_i$ . We obtain

$$\int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \theta_k \partial_t \mathbf{q}_h \, d\mathbf{x} \, dt + \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \theta_k \operatorname{div} \mathbf{F}(\mathbf{q}_h) \, d\mathbf{x} \, dt = \mathbf{0} \quad (10)$$

with  $\Omega_i^\circ = \Omega_i \setminus \partial\Omega_i$  being the interior of the spatial cell.

Applying integration by parts in time yields

$$\begin{aligned} & \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t^{n+1}) \mathbf{q}_h(\mathbf{x}, t^{n+1}) d\mathbf{x} - \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t^n) \mathbf{u}_h(\mathbf{x}, t^n) d\mathbf{x} - \\ & \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \partial_t \theta_k(\mathbf{x}, t) \mathbf{q}_h(\mathbf{x}, t) d\mathbf{x} dt + \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t) \operatorname{div} \mathbf{F}(\mathbf{q}_h(\mathbf{x}, t)) d\mathbf{x} dt = \mathbf{0}, \end{aligned} \quad (11)$$

in which now  $\mathbf{u}_h(\mathbf{x}, t^n)$  is given. Equation (11) involves only the local polynomial  $\mathbf{q}_h^n$  within a single space-time element. The space-time predictors (11) can be thus solved independently in each element. In practice, we approximate their solution by a simple fixed-point iteration procedure (discrete Picard iterations) as follows

$$\begin{aligned} & \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t^{n+1}) \mathbf{q}_h^{(k+1)}(\mathbf{x}, t^{n+1}) d\mathbf{x} - \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \partial_t \theta_k(\mathbf{x}, t) \mathbf{q}_h^{(k+1)}(\mathbf{x}, t) d\mathbf{x} dt = \\ & \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t^n) \mathbf{u}_h(\mathbf{x}, t^n) d\mathbf{x} + \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \theta_k(\mathbf{x}, t) \operatorname{div} \mathbf{F}(\mathbf{q}_h^{(k)}(\mathbf{x}, t)) d\mathbf{x} dt. \end{aligned} \quad (12)$$

For further details one can refer to [31, 30] where the same notation has been used.

### 3.2.2. Corrector Step

Once the local space-time prediction of the solution  $\mathbf{q}_h^n(\mathbf{x}, t)$  has been computed, we can write a one step update in time as follows. We start by multiplying (1) with a spatial-only test function  $\phi_k$  (cf. equation (7)), and integrate in space and time:

$$\int_{t^n}^{t^{n+1}} \int_{\Omega_i} \phi_k (\partial_t \mathbf{u} + \operatorname{div} \mathbf{F}(\mathbf{u})) d\mathbf{x} dt = \mathbf{0}. \quad (13)$$

Inserting now (6) in the time evolution variable and our local information (11) in the flux, we can apply the divergence theorem in (13), and we get our final update step of the ADER-DG method:

$$\left( \int_{\Omega_i^\circ} \phi_k \phi_l d\mathbf{x} \right) (\hat{\mathbf{u}}_\ell^{n+1} - \hat{\mathbf{u}}_\ell^n) + \int_{t^n}^{t^{n+1}} \int_{\partial\Omega_i} \phi_k \mathcal{F}(\mathbf{q}_h^-, \mathbf{q}_h^+) \cdot \mathbf{n} dS dt - \int_{t^n}^{t^{n+1}} \int_{\Omega_i^\circ} \nabla \phi_k \cdot \mathbf{F}(\mathbf{q}_h) d\mathbf{x} dt = \mathbf{0}, \quad (14)$$

where  $\mathcal{F}$  is the numerical flux, and  $\mathbf{n}$  the external normal vector at the boundaries of the cell. Since we have the local predictor information  $\mathbf{q}_h$ , we can apply high-order quadrature rules for the approximation of time-integrals in (14) and we finally obtain our update for the solution at  $t^{n+1}$ . We remark that the information exchange in-between the elements plays a role only in the corrector step. As mentioned before, the presented interpretation of ADER-DG is an explicit space-time FE method.

## 4. Fully discrete entropy preserving ADER discontinuous Galerkin schemes

As described in the previous section, the ADER-DG approach consists of two steps: a set of local predictors, and a fully explicit global corrector. The corrector step enforces conservation, and thus consistency with weak solutions of (1). However, the entropy (in)equality (2) (or (3)) might not be guaranteed. Due to the inherent space-time nature of the method, to enforce this additional constraint we need to have simultaneous control of the entropy production in space and in time. To this end we extend two techniques previously proposed in literature:

1. for the balance of the spatial discretization in our ADER-DG method (14), we apply entropy correction terms first proposed by Abgrall in [2],
2. for handling the time-integration we deal with the relaxation approach described in [40, 53].

In the following section we explain how these techniques can be adapted to (14).



#### 4.1. Entropy correction in space

The approach proposed by Abgrall in [2] is based on the following idea. We introduce a simple entropy correction term to the underlying discretization which is lead by a parameter which varies in every cell. The parameter is chosen in order to locally balance the entropy production at every degree of freedom and, simultaneously, not to violate the conservation property of the scheme. It can be interpreted as the solution of an optimization problem [5] or as an additional viscosity/anti-viscosity term. Inspired by [2], we introduce an entropy correction term with artificial diffusion into the corrector of the ADER formulation (14) using a generic element  $\psi$  from our underlying approximation space. It yields

$$\begin{aligned} & \int_{\Omega_i} \psi(\mathbf{x}) (\mathbf{u}_h^{n+1}(\mathbf{x}) - \mathbf{u}_h^n(\mathbf{x})) \, d\mathbf{x} + \int_{t^n}^{t^{n+1}} \int_{\partial\Omega_i} \psi(\mathbf{x}) \mathcal{F}_n(\mathbf{q}_h^-(\mathbf{x}, t), \mathbf{q}_h^+(\mathbf{x}, t)) \, dS \, dt - \\ & \int_{t^n}^{t^{n+1}} \int_{\Omega_i} \nabla \psi(\mathbf{x}) \cdot \mathbf{F}(\mathbf{q}_h(\mathbf{x}, t)) \, d\mathbf{x} \, dt + \int_{t^n}^{t^{n+1}} \alpha_{\Omega_i} \int_{\Omega_i} \nabla \psi(\mathbf{x}) A_0(\mathbf{q}_h) \nabla \mathbf{v}_h(\mathbf{q}_h(\mathbf{x}, t)) \, d\mathbf{x} \, dt = \mathbf{0}, \end{aligned} \quad (15)$$

where  $\mathbf{v}_h$  is the approximation of the entropy variables  $\mathbf{v} \in \mathbb{R}^m$ ,  $\mathcal{F}_n(\cdot, \cdot) := \mathcal{F}(\cdot, \cdot) \cdot \mathbf{n}$  and  $A_0 \in \mathbb{R}^{m \times m}$  is the inverse of the analytical Hessian of the entropy. Note that  $A_0$  is evaluated at each quadrature point along with the gradient terms in the integral. Since by definition each evaluation of  $A_0(x)$  is an SPD matrix, the only criterion to make sure that the sign of the bilinear form resulting from the quadrature is also strictly non negative is to have enough quadrature points wrt the polynomial degree, which is largely satisfied by usual high order quadrature formulas (see e.g. discussion in [1], Section §3.1.7).

Using the Gauss-Legendre quadrature formula for the time integration, from (15) we obtain

$$\begin{aligned} & \int_{\Omega_i} \psi(\mathbf{x}) (\mathbf{u}_h^{n+1}(\mathbf{x}) - \mathbf{u}_h^n(\mathbf{x})) \, d\mathbf{x} + \Delta t \sum_{s=0}^S \beta_s \int_{\partial\Omega_i} \psi(\mathbf{x}) \mathcal{F}_n(\mathbf{q}_h^{s,-}(\mathbf{x}), \mathbf{q}_h^{s,+}(\mathbf{x})) \, dS - \\ & \Delta t \sum_{s=0}^S \beta_s \int_{\Omega_i} \nabla \psi(\mathbf{x}) \cdot \mathbf{F}(\mathbf{q}_h^s(\mathbf{x})) \, d\mathbf{x} + \Delta t \sum_{s=0}^S \beta_s \alpha_{\Omega_i}^s \int_{\Omega_i} \nabla \psi(\mathbf{x}) A_0 \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})) \, d\mathbf{x} = \mathbf{0}, \end{aligned} \quad (16)$$

where  $\beta_s$  are the time quadrature weights. For the sake of completeness, we recall that for numerical integration we employ, on each triangle,  $(N+1)^2$  quadrature points defined by the tensor product of the one-dimensional Gauss-Jacobi formula. The points are defined on a square and one edge is collapsed into a vertex. Since all the Gauss-Legendre points are not on the boundary of the square, they are all distinct in the triangle as well, see [64]. Then, the spatial points are extruded in time again by means of the tensor product rule with Gauss-Legendre points taken on the interval  $[t^n, t^{n+1}]$ . We thus obtain a total number of  $(N+1)^{(2+1)}$  quadrature points for the space-time integration, distributed on  $N+1$  time slabs.

It is left to define all  $\alpha_{\Omega_i}^s$ . We choose them in order to have an entropy dissipative (conservative) semi-discretization update to balance with the entropy flux integral. In order to do so, we proceed by considering the following approximations. We would like to substitute  $\psi$  in the previous equation with  $\mathbf{v}_h(\mathbf{q}_h(t, \mathbf{x}))$  and compute a scalar product, in order to obtain on the left hand side an approximation of the entropy variation in time and on the right hand side some conservative and dissipative entropy flux terms. We extend equation (15) to a time dependent  $\psi$ . Then, for each entry of the vector  $\mathbf{u}_h$ , we use as  $\psi$  function the corresponding entry of  $\mathbf{v}_h(\mathbf{q}_h(t, \mathbf{x}))$  and we sum all over the components. We can see that (15) are specific cases of a more general equation where  $\psi$  is also time dependent and a vector. To obtain back (15), one should simply choose  $\psi$  constant in time and nonzero in one of the system component. This leads to

$$\begin{aligned} & \int_{\partial\Omega_i} \langle \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})), \mathcal{F}_n(\mathbf{q}_h^{s,-}(\mathbf{x}), \mathbf{q}_h^{s,+}(\mathbf{x})) \rangle \, dS - \int_{\Omega_i} \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})) : \mathbf{F}(\mathbf{q}_h^s(\mathbf{x})) \, d\mathbf{x} + \\ & \alpha_{\Omega_i}^s \int_{\Omega_i} \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x}))^T A_0 \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})) \, d\mathbf{x} \geq \int_{\partial\Omega_i} \mathcal{G}(\mathbf{v}_h(\mathbf{q}_h^{s,-}(\mathbf{x})), \mathbf{v}_h(\mathbf{q}_h^{s,+}(\mathbf{x}))) \cdot \mathbf{n} \, dS, \end{aligned} \quad (17)$$

where  $\mathcal{G}$  is a consistent numerical entropy flux. Mind that these estimates are only an approximation of what we want to achieve and not the actual proof of the entropy conservation/diffusion property, which follows in the next equations. Now, we split the numerical flux into a central part and a diffusive one as

$$\mathcal{F}_n(\mathbf{q}^-, \mathbf{q}^+) = \mathcal{F}_n^c(\mathbf{q}^-, \mathbf{q}^+) + \mathcal{F}_n^d(\mathbf{q}^-, \mathbf{q}^+). \quad (18)$$

As an example with a local Lax-Friedrichs (Rusanov) entropy type flux, we would have

$$\mathcal{F}_{\mathbf{n}}^c(\mathbf{q}^-, \mathbf{q}^+) := \frac{\mathbf{F}(\mathbf{q}^-) + \mathbf{F}(\mathbf{q}^+)}{2} \cdot \mathbf{n}, \quad \text{and} \quad \mathcal{F}_{\mathbf{n}}^d(\mathbf{q}^-, \mathbf{q}^+) := -\mathbf{a}_{LF, \mathbf{n}}(\mathbf{q}^-, \mathbf{q}^+) \frac{\mathbf{q}^+ - \mathbf{q}^-}{2}. \quad (19)$$

Here,  $\mathbf{a}_{LF, \mathbf{n}}$  denotes the dissipation matrix, i.e.,  $\mathbf{a}_{LF, \mathbf{n}} = \mathcal{I} \lambda$  with  $\lambda$  the maximum spectral radius of the flux Jacobian in the direction  $\mathbf{n}$  computed in the two arguments.

For simplicity, we introduce the following definitions of the central spatial contribution, the diffusion flux term and the entropy fix

$$\begin{aligned} \mathbf{F}_i^s &:= \int_{\partial\Omega_i} \langle \mathbf{v}_h(\mathbf{q}_{h,i}^s(\mathbf{x})), \mathcal{F}_{\mathbf{n}}^c(\mathbf{q}_h^{s,-}(\mathbf{x}), \mathbf{q}_h^{s,+}(\mathbf{x})) \rangle dS - \int_{\Omega_i^o} \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})) : \mathbf{F}(\mathbf{q}_h^s(\mathbf{x})) d\mathbf{x}, \\ \mathbf{D}_i^s &:= \int_{\partial\Omega_i} \langle \mathbf{v}_h(\mathbf{q}_{h,i}^s(\mathbf{x})), \mathcal{F}_{\mathbf{n}}^d(\mathbf{q}_h^{s,-}(\mathbf{x}), \mathbf{q}_h^{s,+}(\mathbf{x})) \rangle dS, \\ \mathbf{E}_i^s &:= \int_{\Omega_i} \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x}))^T A_0 \nabla \mathbf{v}_h(\mathbf{q}_h^s(\mathbf{x})) d\mathbf{x} \geq 0. \end{aligned}$$

With these definitions, we can finally determine  $\alpha_{\Omega_i}^s$  in such a way to ensure entropy conservation or entropy dissipation. We start focusing on the entropy dissipative case, therefore a simple choice would be to determine  $\alpha_{\Omega_i}^s$  such that the entropy contribution of the central part vanishes, i.e.,

$$\int_{\partial\Omega_i} \mathcal{G}(\mathbf{v}_h(\mathbf{q}^{s,-}(\mathbf{x})), \mathbf{v}_h(\mathbf{q}^{s,+}(\mathbf{x}))) \cdot \mathbf{n} dS = \mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s, \quad s = 1, \dots, S. \quad (20)$$

This leads to the definition of

$$\alpha_{\Omega_i}^s = \frac{\int_{\partial\Omega_i} \mathcal{G}(\mathbf{v}(\mathbf{q}^{s,-}(\mathbf{x})), \mathbf{v}(\mathbf{q}^{s,+}(\mathbf{x}))) \cdot \mathbf{n} dS - \mathbf{F}_i^s}{\mathbf{E}_i^s}. \quad (21)$$

Since the diffusive part  $\mathbf{D}_i^s$  decreases always the entropy, we obtain the following estimation

$$\begin{aligned} \int_{t^n}^{t^{n+1}} \int_{\Omega} \mathbf{v}_h(\mathbf{x}, t)^T \partial_t \mathbf{u}_h(\mathbf{x}, t) d\mathbf{x} &\approx -\Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s) \\ &= -\Delta t \sum_{s=0}^S \beta_s \left[ \sum_{\Omega_i \in \Omega} \int_{\partial\Omega_i} \mathcal{G}(\mathbf{v}(\mathbf{q}^{s,-}(\mathbf{x})), \mathbf{v}(\mathbf{q}^{s,+}(\mathbf{x}))) \cdot \mathbf{n} dS + \sum_{i \in \Omega} \mathbf{D}_i^s \right] \leq -\Delta t \sum_{s=0}^S \beta_s \int_{\partial\Omega} \mathbf{G}(\mathbf{v}(\mathbf{q}^s(\mathbf{x}))) \cdot \mathbf{n} dS. \end{aligned}$$

With this definition, we correct the contribution from the centered term of the numerical flux. In particular, we are able to split the update into an entropy conservative part ( $\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s$ ) and an entropy dissipative part ( $\mathbf{D}_i^s$ ). This splitting will be crucial when introducing the relaxation technique, as it allows to choose which part we need to balance also in time.

**Remark 4.1** (Division by zero). *The division performed in (21) is numerically dangerous as for areas where the solution is flat (i.e. uniform in space)  $\mathbf{E}_i^s$  might be 0. In order not to risk to encounter NaN, we perform that division only if  $\mathbf{E}_i^s \geq \varepsilon^s$ , otherwise we set  $\alpha_{\Omega_i}^s = 0$ . This tolerance  $\varepsilon^s$  is chosen in the following way*

$$\varepsilon^s := \bar{h}^N \max_{i \in \Omega} \mathbf{E}_i^s, \quad (22)$$

where  $N$  is the polynomial degree and  $\bar{h}$  is the average mesh size. This guarantees not to introduce extra terms where the solution is constant or almost constant.

**Remark 4.2** (Correction for first order finite volume schemes). *The special case of  $N = 0$ , i.e., the finite volume method, does not fit in our framework as the terms  $\mathbf{E}$  are all zero. In this case, we can resort to the classical characterization by Tadmor [65, 66]. On unstructured grids we introduce the jumps  $\llbracket (\cdot) \rrbracket = (\cdot)^+ - (\cdot)^-$  on a given mesh face with normal  $\mathbf{n}_f$ . The generalization of the classical Tadmor condition [56, 66]*

$$\llbracket \mathbf{v}_h \rrbracket^T \mathcal{F}_{\mathbf{n}_f}^{co} = \llbracket \boldsymbol{\psi}(u_h) \rrbracket \cdot \mathbf{n}_f \quad (23)$$

defines an entropy conservative flux  $\mathcal{F}^{co}$  consistent with the numerical entropy flux

$$\mathcal{G}_{\mathbf{n}_f} := \bar{\mathbf{v}}_h^T \mathcal{F}_{\mathbf{n}_f}^{co} - \bar{\psi}(u_h) \cdot \mathbf{n}_f, \quad (24)$$

having denoted by  $\bar{\cdot}$  the arithmetic average, and by  $\psi$  the entropy potential defined by

$$\mathbf{G} = \mathbf{v}^T \mathbf{F} - \psi. \quad (25)$$

A given numerical flux  $\mathcal{F}_{\mathbf{n}_f}$  may not verify (23). However, we can use the face jumps as a correction term and define

$$\mathcal{F}_{\mathbf{n}_f}^{co} = \mathcal{F}_{\mathbf{n}_f} - \alpha_f A_0 (\mathbf{v}^+ - \mathbf{v}^-) = \mathcal{F}_{\mathbf{n}_f} - \alpha_f A_0 \llbracket \mathbf{v} \rrbracket, \quad (26)$$

with  $A_0$  the SPD inverse of the analytical entropy Hessian in some linearized state. By imposing the entropy conservation condition (23), we compute a limiting value

$$\alpha_0 := \frac{-\llbracket \psi \rrbracket \cdot \mathbf{n}_f + \llbracket \mathbf{v} \rrbracket^T \mathcal{F}_{\mathbf{n}_f}}{\llbracket \mathbf{v} \rrbracket^T A_0 \llbracket \mathbf{v} \rrbracket} = \frac{\llbracket \mathbf{G} \rrbracket \cdot \mathbf{n}_f - \llbracket \mathbf{v}_h^T \mathbf{F} \rrbracket \cdot \mathbf{n}_f + \llbracket \mathbf{v} \rrbracket^T \mathcal{F}_{\mathbf{n}_f}}{\llbracket \mathbf{v} \rrbracket^T A_0 \llbracket \mathbf{v} \rrbracket}. \quad (27)$$

Entropy conservation is ensured for  $\alpha_f = \alpha_0$ , while semi-discrete entropy stability is obtained for  $\alpha_f > \alpha_0$ . Note that this comparison result is relatively classical, see [66]. In this work, we will focus on high order methods, and we will not discuss further this case. The authors are anyway considering this formulation to obtain an a posteriori subcell finite volume limiter, in order to deal with solutions with strong shocks. This will be object of future studies, as the coupling of the two techniques is delicate and must be handled carefully.

#### 4.2. Relaxation applied to ADER-DG

Would the time integration be exact (vis in the time continuous case), the correction introduced in the previous section would allow to satisfy *exactly* entropy conservation/stability within each element. But, unfortunately, in general this is not the case and moreover the fully discrete nature of the ADER approach requires to handle the temporal discretization simultaneously with the spatial one. To achieve exact entropy conservation thus we adapt the relaxation approach proposed in [40], and applied in different contexts [3, 39, 54, 53]. The key idea of the relaxation ansatz is to project the final numerical approximation  $\mathbf{u}_h^{n+1}$  back to the manifold where the entropy is constant (or dissipative), as in the underlying continuous setting. This is obtained multiplying the time-step  $\Delta t$  with a scalar quantity  $\gamma^n$  and using this new time-step for the update. We describe now how  $\gamma^n$  can be chosen in order to preserve (or dissipate) the total entropy (4). We set

$$\Delta \mathbf{u}_h := \mathbf{u}_h^{n+1} - \mathbf{u}_h^n, \quad (28)$$

where  $\mathbf{u}_h^{n+1}$  is obtained solving (15). Now, we have to find  $\gamma^n$  and  $\mathbf{u}_{h,\gamma^n}^{n+1} := \mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h$  such that the total entropy is controlled. We start by estimating the total entropy (4) as

$$\begin{aligned} \mathcal{E}(\mathbf{u}_{h,\gamma^n}^{n+1}) &= \int_{\Omega} \eta(\mathbf{u}_{h,\gamma^n}^{n+1}) d\mathbf{x} \approx \int_{\Omega} \eta(\mathbf{u}_h^n) d\mathbf{x} + \gamma^n \langle \eta_u(\mathbf{u}_h^n), \Delta \mathbf{u}_h \rangle d\mathbf{x} \\ &\approx \mathcal{E}(\mathbf{u}_h^n) - \gamma^n \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s) \\ &\leq \mathcal{E}(\mathbf{u}_h^n) - \gamma^n \Delta t \sum_{s=0}^S \beta_s \int_{\partial \Omega} G(\mathbf{v}^s) \cdot \mathbf{n} dS \approx \mathcal{E}(\mathbf{u}_h^n) - \int_{t^n}^{t^{n+1}} \int_{\partial \Omega} G(\mathbf{v}^s) \cdot \mathbf{n} dS. \end{aligned} \quad (29)$$

This leads us to the following (non-)linear equation:

$$\mathcal{E}(\mathbf{u}_{h,\gamma^n}^{n+1}) = \int_{\Omega} \eta(\mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h) d\mathbf{x} \stackrel{!}{=} \int_{\Omega} \eta(\mathbf{u}_h^n) d\mathbf{x} - \gamma^n \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s). \quad (30)$$

This is equivalent to finding the zero of

$$R(\gamma^n) := \int_{\Omega} \eta(\mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h) d\mathbf{x} - \int_{\Omega} \eta(\mathbf{u}_h^n) d\mathbf{x} + \gamma^n \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s). \quad (31)$$

In case we want to exactly preserve the entropy in time, we can remove the  $D_i^s$  terms from equation (30). From (30), we can finally calculate  $\gamma^n$  at every time-step and we update the new value of the approximated solution and the new time-step via

$$\begin{cases} \tilde{\mathbf{u}}_{h,\gamma^n}^{n+1} := \mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h^n, \\ t_{\gamma^n}^{n+1} := t^n + \gamma^n \Delta t. \end{cases} \quad (32)$$

Finally, the time step  $\Delta t = t^{n+1} - t^n$  is chosen according to a CFL condition

$$\Delta t = \text{CFL} \min_{\Omega_i} \frac{h_i}{|\lambda_{\max,i}|}, \quad \forall \Omega_i \in \Omega, \quad (33)$$

where the Courant-Friedrichs-Levy number has to be  $\text{CFL} < 1$  (CFL=0.9 is used in all our numerical simulations) and  $|\lambda_{\max,i}|$  is the maximum absolute value of the eigenvalues in the cell  $\Omega_i$ .

In order to have a solution of the relaxation equation (31), we need to have  $R(0) = 0$ ,  $R'(0) < 0$  and  $R''(0) > 0$  [54]. It is straightforward that  $R(0) = 0$ . It is also clear that  $R''(0) = \int_{\Omega} \Delta \mathbf{u}_h^T \eta''(\mathbf{u}_h^n) \Delta \mathbf{u}_h d\mathbf{x} > 0$  as the entropy is convex. To prove that  $R'(0) < 0$  we have to perform some steps similar to the proof in [54], where some smallness assumptions on  $\Delta t$  and  $h$  must be imposed. The proof can be found in Appendix A.

**Remark 4.3** (Division by zero (Relaxation)). *As discussed in Remark 4.1, when  $E_i^s < \varepsilon$  we set  $\alpha_{\Omega_i} = 0$ . To ensure control of the entropy production up to machine accuracy this particular case needs to be treated also in the relaxation procedure. In our approach, if  $E_i^s < \varepsilon$  we replace  $F_i^s + \alpha_{\Omega_i}^s E_i^s$  by  $\int_{\partial\Omega_i} G(\mathbf{v}^s) \cdot \mathbf{n} dS$  in (30). This guarantees that we still have*

$$\mathcal{E}(\mathbf{u}_{h,\gamma^n}^{n+1}) = \mathcal{E}(\mathbf{u}_h^n) - \sum_{s=0}^S \beta^s \left[ \int_{\partial\Omega} G(\mathbf{v}^s) \cdot \mathbf{n} dS + \sum_{\Omega_i \in \Omega} D_i^s \right], \quad (34)$$

which is otherwise guaranteed by the definition of  $\alpha_{\Omega_i}$ .

It should be pointed out that in case of quadratic entropies (30) is explicitly solvable as demonstrated in Example 4.5 below, while for more nonlinear entropies a nonlinear solver for the scalar equation (30) must be used. In our numerical simulations in Section 5, we use a simple Newton method to solve (30).

**Remark 4.4** (Newton method for (31)). *When we want to find the (non trivial) zero of (31), the Newton method or variants can help in finding the solution. For Newton method also the derivative of  $R$  is required, but it is easily computable as*

$$\frac{dR(\gamma^n)}{d\gamma^n} = \int_{\Omega} \langle \mathbf{v}(\mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h), \Delta \mathbf{u}_h \rangle d\mathbf{x} + \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (F_i^s + \alpha_{\Omega_i}^s E_i^s + D_i^s). \quad (35)$$

In that case, to compute the Jacobian, the entropy variable definition is, again, necessary. Finally, we have to stress that due to the results of [53], we can ensure that if the underlying time method is of order  $p \geq 2$ ,

- there exists always a unique  $\gamma > 0$  that solves (30). This  $\gamma$  satisfies  $\gamma = 1 + O(\Delta t^{p-1})$  under a reasonable small time-step  $\Delta t$ , and
- the relaxation solution  $\tilde{\mathbf{u}}_{h,\gamma^n}^{n+1}$  is also of order  $p$ .

The proof of these results is identical to the one in [54] where the main steps are the following. First, it is shown that  $R(1) = O(\Delta t^{p+1})$ , then, since  $R'(0) = \Theta(\Delta t^2) < 0$  because  $R''(0) > 0$ , we know that there exists a solution  $\gamma^n$  such that  $R(\gamma^n) = 0$  and it is distant an  $O(\Delta t^{p-1})$  by 1. Hence, the solution is  $\tilde{\mathbf{u}}_{h,\gamma^n}^{n+1}$ , considered at time  $t^n + \gamma^n \Delta t$ , is of order  $p$ .

Using the entropy correction terms and the relaxation approach, we obtain an entropy preserving ADER-DG method for which we will verify the properties next in the numerical experiments section, but before we finish this part with the following already mentioned example.

**Example 4.5** (Quadratic entropy). *In case of a quadratic entropy for scalar equations*

$$\eta(u) := \frac{\langle u, u \rangle}{2},$$

we can *simplify* the expression  $\mathcal{E}(\mathbf{u}_h^{n+1})$  as

$$\mathcal{E}(\mathbf{u}_h^{n+1}) = \int_{\Omega} \eta(\mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h) d\mathbf{x} = \frac{1}{2} \int_{\Omega} \langle \mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h, \mathbf{u}_h^n + \gamma^n \Delta \mathbf{u}_h \rangle \quad (36)$$

$$= \frac{1}{2} \int_{\Omega} \langle \mathbf{u}_h^n, \mathbf{u}_h^n \rangle + 2\gamma^n \langle \Delta \mathbf{u}_h, \mathbf{u}_h^n \rangle + (\gamma^n)^2 \langle \Delta \mathbf{u}_h, \Delta \mathbf{u}_h \rangle d\mathbf{x} \quad (37)$$

$$= \mathcal{E}(\mathbf{u}_h^n) + \int_{\Omega} \gamma^n \langle \Delta \mathbf{u}_h, \mathbf{u}_h^n \rangle + \frac{1}{2} (\gamma^n)^2 \langle \Delta \mathbf{u}_h, \Delta \mathbf{u}_h \rangle d\mathbf{x}, \quad (38)$$

so that (30) reduces to

$$\int_{\Omega} \gamma^n \langle \Delta \mathbf{u}_h, \mathbf{u}_h^n \rangle + \frac{1}{2} (\gamma^n)^2 \langle \Delta \mathbf{u}_h, \Delta \mathbf{u}_h \rangle d\mathbf{x} \stackrel{!}{=} -\gamma^n \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s). \quad (39)$$

The solution  $\gamma^n = 0$  is not the *desired value*, the other one is

$$\gamma^n = -2 \frac{\Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s) + \int_{\Omega} \langle \Delta \mathbf{u}_h, \mathbf{u}_h^n \rangle d\mathbf{x}}{\int_{\Omega} \langle \Delta \mathbf{u}_h, \Delta \mathbf{u}_h \rangle d\mathbf{x}}, \quad (40)$$

which is close to one *as stated above*.

## 5. Numerical results

In this section we test the proposed algorithm against many physical problems: linear advection, shallow water and Euler equations. We will focus on different aspects in the various tests. In most cases we will compare the standard ADER-DG method (*indicated in the figures legend by  $P_N$* ) with the proposed ADER-DG equipped both with the entropy fix in space and the relaxation technique (*indicated in the figures legend by the short name  $P_N$ +Rel*). We will study the convergence analysis, the entropy error in time and the value of  $\gamma^n$  in time. For each problem we will introduce the equations, the entropy variables, the entropy fluxes and the matrix  $A_0$ .

### 5.1. Linear advection and rotation

The first equation we treat is the linear advection equations in two forms. The equation we consider is

$$\partial_t u + \mathbf{a} \cdot \nabla u = 0, \quad (41)$$

with  $\mathbf{a} = (a_1, a_2)$  which can be constant or a rotation vector, i.e.,  $\mathbf{a} = (-y, x)$ . The entropy, entropy variable and entropy flux are

$$\eta(u) = u^2/2, \quad v(u) = u, \quad G(u) = \eta(u) \mathbf{a}. \quad (42)$$

In this case we have trivially that  $A_0 = 1$ .

#### 5.1.1. Test: traveling bump

The first smooth test that we consider is given by the following initial condition

$$u_0(\mathbf{x}) = \begin{cases} e^{1-\frac{1}{1-r^2}} & \text{if } r < 1, \\ 0 & \text{else,} \end{cases} \quad (43)$$

where  $r = \sqrt{x^2 + y^2}$ , the advection coefficient is  $\mathbf{a} = (1, 0)^T$ , the domain is  $\Omega = [-1.5, 1.5] \times [-1.5, 1.5]$  with periodic boundary conditions in the  $x$  direction and wall boundary conditions in the  $y$  direction. The exact solution is given by  $u_{ex}(x, t) = u_0(\mathbf{x} - \mathbf{a}t)$ .

*First*, we check that the expected order of accuracy of the relaxed schemes is conserved. In Table 1 we analyze the error at time  $t_f = 3$  with respect to the exact solution and we observe that for all elements  $P_N$  we obtain order of

Table 1: Order of convergence for the ADER-DG schemes equipped with the relaxation technique at time  $t_f = 3.0$  for the  $L_2$  norm error of the travelling bump.

$P_1 + \text{Rel}$			$P_2 + \text{Rel}$			$P_3 + \text{Rel}$		
$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$	$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$	$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$
2.37E-2	3.51E-3	-	3.31E-2	4.47E-4	-	4.30E-2	2.07E-4	-
1.69E-2	1.93E-3	1.78	2.43E-2	2.74E-4	2.77	3.12E-2	6.77E-5	3.47
1.23E-2	1.09E-3	1.82	1.75E-2	1.05E-4	2.91	2.27E-2	2.02E-5	3.79
8.91E-3	5.96E-4	1.84	1.26E-2	4.04E-5	2.94	1.63E-2	5.49E-6	3.97

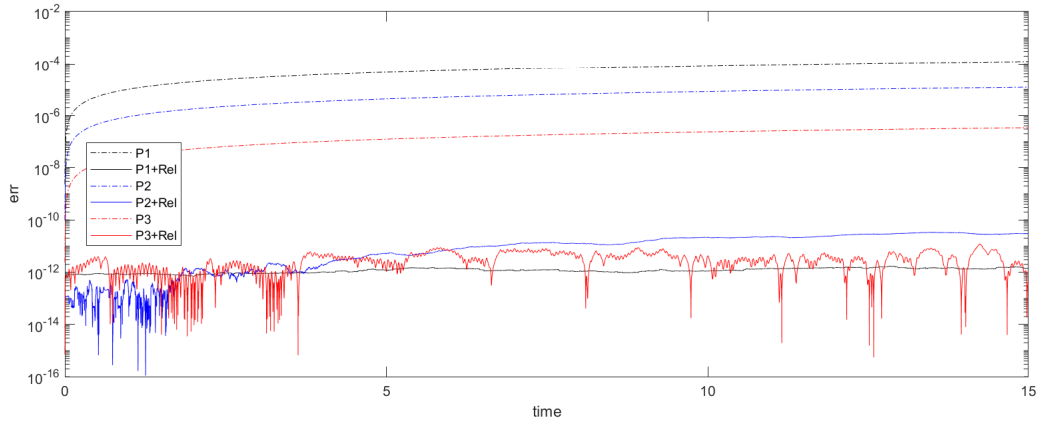


Figure 1: Total entropy error for the traveling bump test case.

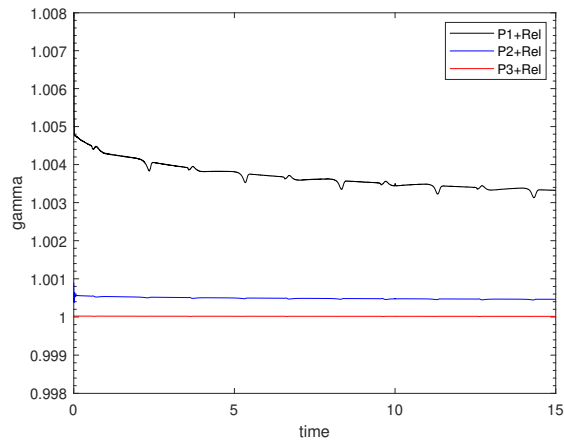


Figure 2: Values of  $\gamma^n$  for the traveling bump test case.

Table 2: Order of convergence for the ADER-DG schemes equipped with the relaxation technique at time  $t_f = 0.1$  for the  $L_2$  norm error of the rotating bump.

$P_1$ + Rel			$P_2$ + Rel			$P_3$ + Rel		
$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$	$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$	$\Delta x$	$L_2(h - h_{ex})$	$O(L_2)$
4.53E-2	2.93E-3	-	6.37E-2	8.09E-5	-	8.28E-2	4.28E-4	-
4.06E-2	2.35E-3	2.09	5.75E-2	6.16E-6	2.69	7.38E-2	2.91E-4	3.32
3.71E-2	1.97E-3	1.94	5.24E-2	4.77E-6	2.75	6.77E-2	2.12E-4	3.71
3.43E-3	1.70E-3	1.91	4.84E-2	3.71E-6	3.19	6.24E-2	1.57E-4	3.70

accuracy  $N + 1$ . Then, we compute the error between the exact total entropy and the approximated one. In Figure 1 we compare the total entropy error for simulations with different order of polynomials, with and without relaxation, but with a similar number of degrees of freedom. In particular we run the  $P_1$  scheme over a mesh of 24190 elements (72570 DOFs), the  $P_2$  scheme over a mesh of 11726 elements (70356 DOFs), and the  $P_3$  scheme over a mesh of 7104 elements (71040 DOFs). We run all the simulations for time  $t_f = 15$ . The plot always shows the error in absolute value, but we observe in the simulations that the classical ADER-DG dissipates the total entropy. Moreover, we can see in the plot that the higher the order of accuracy the lower the error is for classical ADER-DG. Finally, we observe that the relaxation methods preserve the total entropy up to machine precision also for [large values of the final time](#).

Finally, in Figure 2 we show the values of  $\gamma^n$  as a function of time for the relaxation schemes for the same meshes described above. The value is close to 1, in particular, it is much closer when the order is high.

### 5.1.2. Test: rotating bump

The second test we consider is a bump rotating around the point  $(0, 0)$ . The initial condition is

$$u_0(\mathbf{x}) = \begin{cases} e^{1-\frac{1}{1-r^2}} & \text{if } r < 1, \\ 0 & \text{else,} \end{cases} \quad (44)$$

where  $r = \sqrt{x^2 + (y - 1.5)^2}$ , the advection coefficient is  $\mathbf{a} = (-y, x)^T$  and the domain is  $\Omega = [-3, 3] \times [-3, 3]$ . Homogeneous Dirichlet boundary conditions are imposed all over the boundary.

As for the previous test, we assess the order of accuracy (as usual, without loss of generality, at a short time  $t_f = 0.1$ ) and the error of the total entropy and the values of  $\gamma^n$  (after 1 complete revolution around the origin at time  $t_f = 2\pi$ ). The order of accuracy is computed in Table 2 and, again, we observe the expected convergence order for all the polynomials degree.

In Figure 3 the entropy error with respect to time is depicted for the classical ADER-DG and the entropy corrected one. The three schemes considered in this plot employ a similar number of degrees of freedom: indeed, we run the  $P_1$  scheme over a mesh of 10410 elements (31230 DOFs), the  $P_2$  scheme over a mesh of 5192 elements (31152 DOFs), and the  $P_3$  scheme over a mesh of 3126 elements (31260 DOFs). In the classical ADER-DG the total entropy is preserved only up to the order of accuracy of the scheme and this error increases in time. On the other side, the relaxation ADER-DG obtains the preservation of the total entropy up to machine precision for all times.

In Figure 4 we see the values of  $\gamma^n$  as a function of time for the relaxation method. As expected, the values are close to 1 and, as the scheme is more accurate, we obtain values closer to 1. One must notice that for this test the values of  $\gamma$  are larger with respect to the previous test. This means that the classical ADER-DG in this test has a larger entropy error with respect to the previous one, hence, a larger correction on the timestep is needed to preserve the total entropy.

### 5.2. Shallow water equations

The shallow water equations describe the motion of water for vertically averaged variables. We define with  $h$  the water height and with  $u$  and  $v$  the vertically averaged velocity in  $x$  and  $y$  directions, respectively. We denote



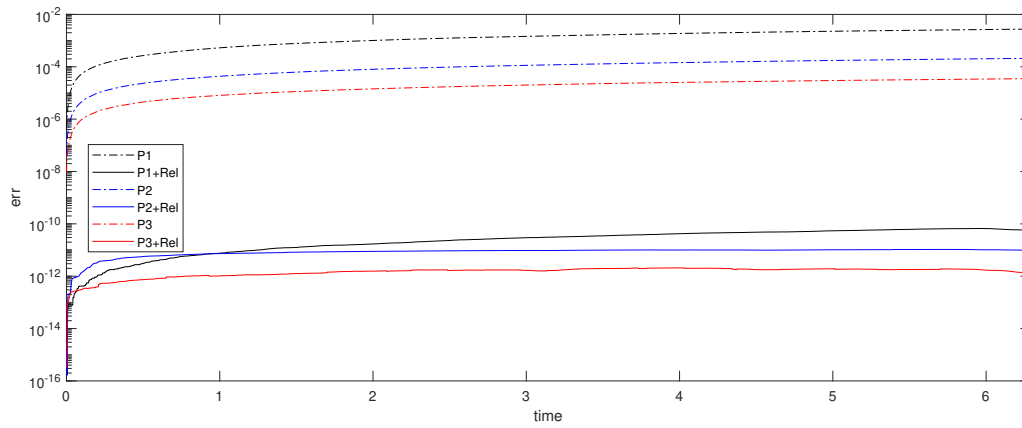


Figure 3: Total entropy error for the rotating bump test case.

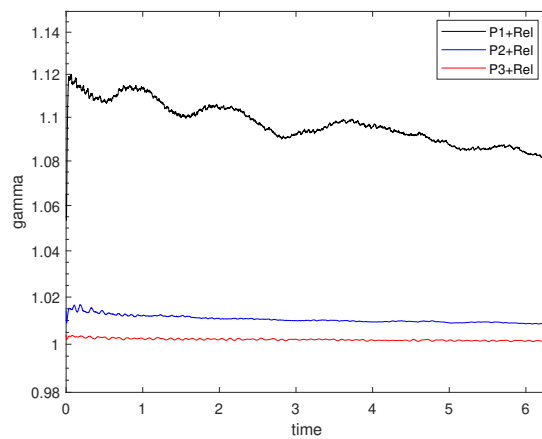


Figure 4: Values of  $\gamma^n$  for the rotating bump test case.

the discharge with  $\mathbf{q} := (hu, hv)^T$  and with  $g$  the gravitational acceleration. The shallow water equations on a flat bathymetry read

$$\begin{cases} \partial_t h + \partial_x(hu) + \partial_y(hv) = 0, \\ \partial_t(hu) + \partial_x(hu^2 + g\frac{h^2}{2}) + \partial_y(huv) = 0, \\ \partial_t(hv) + \partial_x(huv) + \partial_y(hv^2 + g\frac{h^2}{2}) = 0. \end{cases} \quad (45)$$

The associated entropy, entropy variables and entropy fluxes are

$$\eta(\mathbf{u}) = hk + \frac{gh^2}{2}, \quad \mathbf{v}(\mathbf{u}) = (gh - k, u, v)^T, \quad \mathbf{G} = \mathbf{q}(gh + k), \quad (46)$$

$$(47)$$

having introduced the kinetic energy  $k = (u^2 + v^2)/2$ . In this case we have

$$\partial_{\mathbf{u}}\mathbf{v} = \frac{1}{h} \begin{pmatrix} gh + 2k & -u & -v \\ -u & 1 & 0 \\ -v & 0 & 1 \end{pmatrix} \quad \text{and} \quad A_0 = \frac{1}{g} \begin{pmatrix} 1 & u & v \\ u & gh + u^2 & uv \\ v & uv & gh + v^2 \end{pmatrix}. \quad (48)$$

### 5.2.1. Moving vortex test

For these equations, we consider some traveling vortexes whose analytical solution is available. There are many vortexes known in literature that can serve the scope. In [59] there is a  $C^2$  vortex and in [62] there is an isentropic Euler vortex that can be adapted to shallow water which is  $C^\infty$ , but it is not compactly supported. In order to have a vortex that is both compactly supported and with enough smoothness, we will use one of the vortexes defined in [60]. In particular, we will define it on the domain  $\Omega = [0, 1]^2$  and such that it is  $C^6(\Omega)$ . It is centered in  $\mathbf{x}_c = (0.5, 0.5)^T$  with radius  $r_0 = 0.45$ . Its amplitude is  $\Delta h = 0.1$ , its frequency is  $\omega = \pi/r_0$  and its intensity is  $\Gamma = \frac{12\pi\sqrt{g\Delta h}}{r_0\sqrt{315\pi^2 - 2048}}$ . Then, we define the reference water level  $h_c = 1$  and the background velocity  $u_c = 1, v_c = 0$ . Next, we introduce the co-moving coordinates  $\mathcal{I}(\mathbf{x}, t) = \mathbf{x} - \mathbf{x}_c - (u_c t, v_c t)^T$  and we define the distance from the vertex center as  $\mathcal{R}(\mathbf{x}, t) = \|\mathcal{I}(\mathbf{x}, t)\|$ .

The analytical solution is defined as

$$\begin{pmatrix} h(x, t) \\ u(x, t) \\ v(x, t) \end{pmatrix} = \begin{cases} \begin{pmatrix} h_c + \frac{1}{g} \frac{\Gamma^2}{\omega^2} \cdot (\lambda(\omega\mathcal{R}(\mathbf{x}, t)) - \lambda(\pi)), \\ u_c + \Gamma(1 + \cos(\omega\mathcal{R}(\mathbf{x}, t)))^2 \cdot (-\mathcal{I}(\mathbf{x}, t)_y), \\ v_c + \Gamma(1 + \cos(\omega\mathcal{R}(\mathbf{x}, t)))^2 \cdot (\mathcal{I}(\mathbf{x}, t)_x), \end{pmatrix}, & \text{if } \omega\mathcal{R}(\mathbf{x}, t) \leq \pi, \\ \begin{pmatrix} h_c & u_c & v_c \end{pmatrix}^T, & \text{else,} \end{cases} \quad (49)$$

with

$$\begin{aligned} \lambda(r) = & \frac{20 \cos(r)}{3} + \frac{27 \cos(r)^2}{16} + \frac{4 \cos(r)^3}{9} + \frac{\cos(r)^4}{16} + \frac{20r \sin(r)}{3} \\ & + \frac{35r^2}{16} + \frac{27r \cos(r) \sin(r)}{8} + \frac{4r \cos(r)^2 \sin(r)}{3} + \frac{r \cos(r)^3 \sin(r)}{4}. \end{aligned}$$

We consider periodic boundary conditions and final time  $t_f = 1$ . A plot of the initial solution and of the solution at  $t_f = 1$  simulated with the relaxation ADER-DG  $P_2$  is given in Figure 5.

For this test we study the order of accuracy of the schemes for polynomials of degrees 1, 2 and 3. In Table 3 we list the errors  $L_2(h - h_{ex})$  and the experimental order of accuracy  $\mathcal{O}(L_2)$ . The order of accuracy found is the expected one, i.e., 2, 3, and 4. It is clear that the error and the accuracy of the scheme is not modified by the entropy correction and relaxation techniques.

To compare the results with the classical ADER-DG method, we consider different meshes: for  $\mathbb{P}^1$  7590 elements (22770 DOFs), for  $\mathbb{P}^2$  3726 elements (22356 DOFs) and for  $\mathbb{P}^3$  2268 elements (22680 DOFs), in order to have a comparable number of DOFs. In Figure 6 we plot the total entropy error as a function of time for these schemes. For the classical ADER-DG methods the global entropy is dissipated in time and the lower the order the larger is the

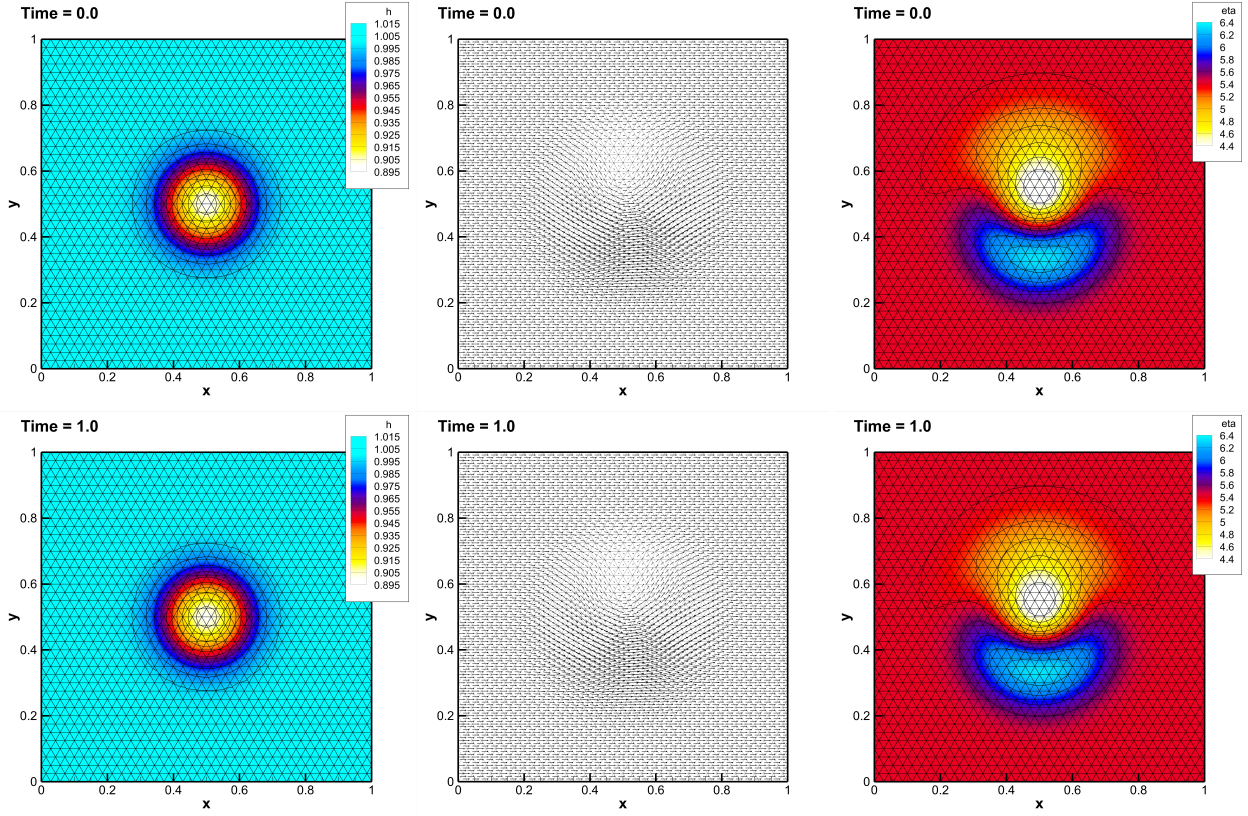


Figure 5: SW vortex test case. In the figure, we report the profile of the [water height  \$h\$](#)  (left), the velocity field (middle) and the entropy  $\eta$  (right) at time  $t = 0$  (top row) and  $t = 1$  (bottom row) obtained with a DG scheme of order 2 with relaxation on a coarse mesh made of 2840 elements.

Table 3: Order of convergence for the ADER-DG schemes equipped with the relaxation technique at time  $t_f = 1$  for the  $L_2$  norm error of the water height  $h$  in the shallow water moving vortex test case.

$\Delta x$	$P_1 + \text{Rel}$		$\Delta x$	$P_2 + \text{Rel}$		$\Delta x$	$P_3 + \text{Rel}$	
	$L_2(h - h_{ex})$	$O(L_2)$		$L_2(h - h_{ex})$	$O(L_2)$		$L_2(h - h_{ex})$	$O(L_2)$
1.64E-2	9.37E-5	-	2.36E-2	1.02E-5	-	2.98E-2	8.41E-7	-
1.27E-2	5.53E-5	2.07	1.83E-2	4.93E-6	2.89	2.36E-2	3.23E-7	4.10
1.01E-2	3.42E-5	2.04	1.43E-2	2.40E-6	2.91	1.84E-2	1.17E-7	4.07
7.88E-3	2.09E-5	2.02	1.10E-2	1.11E-6	2.93	1.43E-2	4.28E-8	4.05

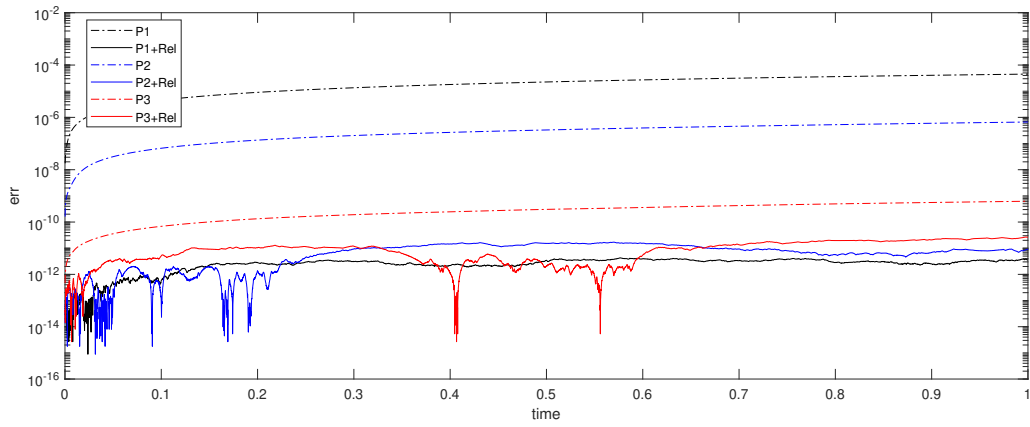


Figure 6: Global entropy error for the shallow water moving vortex test case.

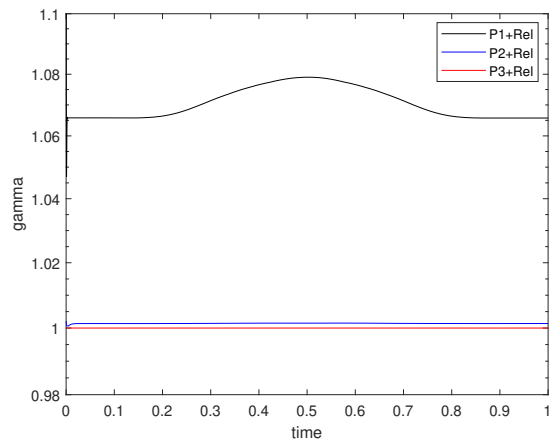


Figure 7: Values of  $\gamma^n$  for the shallow water vortex test case.

entropy error. Conversely, for the relaxation version of the schemes, the total entropy is conserved up to machine precision independently on the order of accuracy of the scheme.

In Figure 7 we can observe the values of  $\gamma^n$  in time for the three relaxation versions of the methods. We observe that for  $\mathbb{P}^1$   $\gamma^n = 1 + O(10^{-1})$ , while for higher orders the values of  $\gamma^n$  are much closer to 1, as the original scheme is already close enough to the entropy conservative solution.

Finally, we would like to stress out that one can also consider problems with bottom topography. By the application of the global flux formulation as described in [45, 17], one can adapt the entropy correction terms to obtain the desired results.

### 5.3. Euler equations of gas dynamics

The complete Euler equations of gas dynamics describe the dynamic of a gas through density  $\rho$ , momentum  $\mathbf{m} = (\rho u, \rho v)^T$  with velocity field  $(u, v)^T$ , and total energy  $E$ . Through these variables, we can define other quantities to shorten the notation of the equations:  $k = (u^2 + v^2)/2$  the specific kinetic energy,  $e := E - \rho k$  the internal energy and  $p$  the pressure linked to the conservative variables by the equation of state (EOS)

$$E = \frac{p}{\gamma - 1} + \rho k, \quad (50)$$

where  $\gamma$  is the adiabatic constant. Then, Euler's equations for a perfect gas read

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) + \partial_y(\rho v) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) + \partial_y(\rho uv) = 0, \\ \partial_t(\rho v) + \partial_x(\rho uv) + \partial_y(\rho v^2 + p) = 0, \\ \partial_t(E) + \partial_x(u(E + p)) + \partial_y(v(E + p)) = 0. \end{cases} \quad (51)$$

Denoting with  $s = p\rho^{-\gamma}$ , we consider the entropy

$$\eta(\mathbf{u}) = -\frac{\gamma + 1}{\gamma - 1} \rho s^{\frac{1}{\gamma+1}}. \quad (52)$$

The entropy for Euler's system is not unique and many other different entropies are used in literature. We have chosen one of Harten's entropies [63], but all other choices are valid. A simple change in entropy and related quantities suffices to obtain another entropy preserving method. Let us denote also the momentum as  $\mathbf{m}^T := (\rho u, \rho v)$ , then, the entropy variables and entropy fluxes are

$$\mathbf{v}(\mathbf{u}) = -(\rho p)^{-\frac{\gamma}{\gamma+1}} \begin{pmatrix} E \\ -\mathbf{m} \\ \rho \end{pmatrix}, \quad G = -\mathbf{m} \frac{\gamma + 1}{\gamma - 1} (p\rho^{-\gamma})^{\frac{1}{\gamma+1}}. \quad (53)$$

The Hessian of the entropy is

$$\partial_{\mathbf{u}} \mathbf{v} = \frac{\gamma(\gamma - 1)}{\gamma + 1} (\rho p)^{-\frac{2\gamma+1}{\gamma+1}} \begin{pmatrix} E & -\mathbf{m} & \rho \end{pmatrix} \otimes \begin{pmatrix} E \\ -\mathbf{m} \\ \rho \end{pmatrix} - (\rho p)^{-\frac{\gamma}{\gamma+1}} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (54)$$

and  $A_0$ , the inverse of  $\partial_{\mathbf{u}} \mathbf{v}$ , is

$$A_0 = \gamma \rho (\rho p)^{-\frac{1}{\gamma+1}} \begin{pmatrix} \rho & \rho u & \rho v & \rho k + \frac{p}{\gamma(\gamma-1)} \\ \rho u & E - \frac{p}{\gamma(\gamma-1)} + \frac{\rho(u^2-v^2)}{2} & \rho uv & uE \\ \rho v & \rho uv & E - \frac{p}{\gamma(\gamma-1)} - \frac{\rho(u^2-v^2)}{2} & vE \\ \rho k + \frac{p}{\gamma(\gamma-1)} & Eu & Ev & \frac{E^2}{\rho} \end{pmatrix}. \quad (55)$$

Table 4: Order of convergence for the ADER-DG schemes equipped with the relaxation technique at time  $t_f = 1$  for the  $L_2$  norm of the density variable  $\rho$  in the Shu vortex test case.

$P_1 + \text{Rel}$			$P_2 + \text{Rel}$			$P_3 + \text{Rel}$		
$\Delta x$	$L_2(\rho - \rho_{ex})$	$O(L_2)$	$\Delta x$	$L_2(\rho - \rho_{ex})$	$O(L_2)$	$\Delta x$	$L_2(\rho - \rho_{ex})$	$O(L_2)$
1.01E-1	1.48E-3	-	1.43E-1	3.10E-4	-	1.84E-1	1.81E-5	-
7.88E-2	9.02E-4	2.03	1.10E-1	1.54E-4	2.68	1.43E-1	6.18E-6	4.35
5.63E-2	4.58E-4	2.02	8.09E-2	6.65E-5	2.71	1.04E-1	1.60E-6	4.20
4.11E-2	2.43E-4	2.01	5.82E-2	2.69E-5	2.74	7.55E-2	4.37E-7	4.06

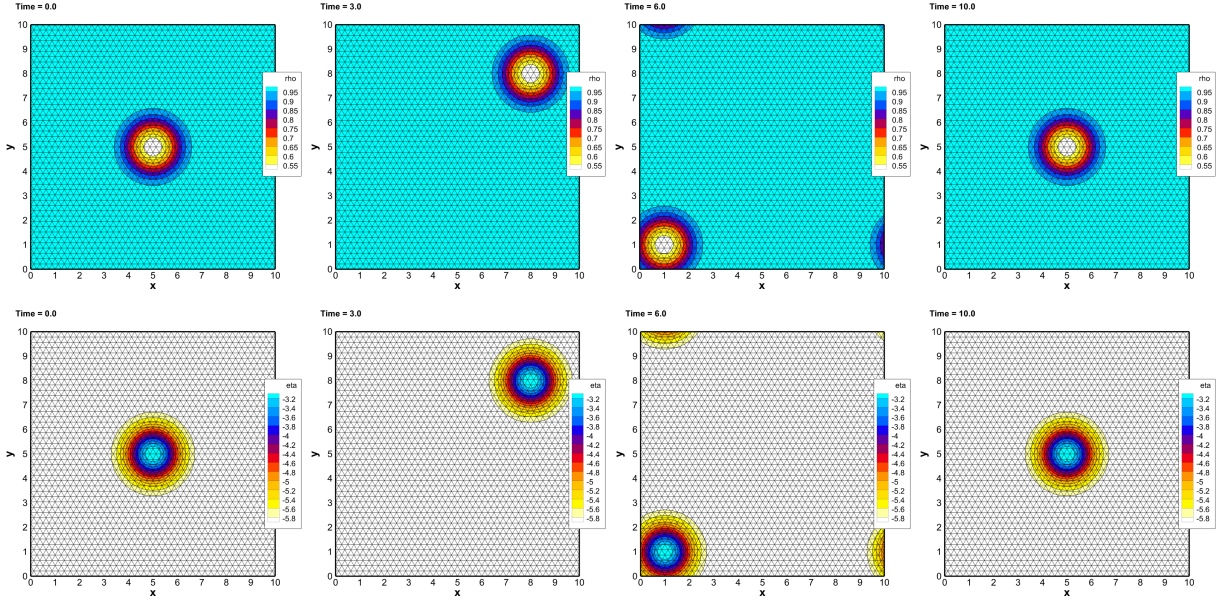


Figure 8: Shu vortex test case with periodic boundary conditions. In the figure we report the profile of the density  $\rho$  (top row) and the entropy  $\eta$  (bottom row) at time  $t = 0, 3, 6$  and  $t_f = 10$  obtained with a DG scheme of order 4 with entropy fix in space and relaxation on a very coarse mesh made of 3726 elements.

### 5.3.1. Shu vortex

We consider again a moving vortex to assess the accuracy of the scheme and to compare the entropy production of the proposed schemes. We consider a Shu vortex [62] at the beginning centered in  $(5, 5)$  with periodic boundary conditions. The simulations are run up to the final time  $t_f = 10.0$ . We plot some simulations at different times  $t = 0, 3, 6$  and  $t_f = 10$  of the variable  $\rho$  and of  $\eta(\mathbf{u})$  in Figure 8 for the DG of order 4 on a very coarse mesh. We test first of all the order of accuracy of the proposed schemes. In Table 4 we observe that all the methods behave with the expected accuracy.

Even if the error of the classical ADER-DG and the relaxation version are very close, there is a huge difference between the schemes on the total entropy. In Figure 9 we plot the total entropy error for different simulations. We considered both classical ADER-DG and relaxation schemes for different orders. The meshes are for  $\mathbb{P}_1$  of 12348 elements (37044 DOFs), for  $\mathbb{P}_2$  scheme of 6300 elements (37800 DOFs), and for  $\mathbb{P}_3$  of 3726 elements (37260 DOFs), so that the DOFs are comparable. As in the previous tests, we have that classical ADER-DG dissipates the total entropy, proportionally to the accuracy of the scheme, while the relaxation version preserves it up to machine precision.

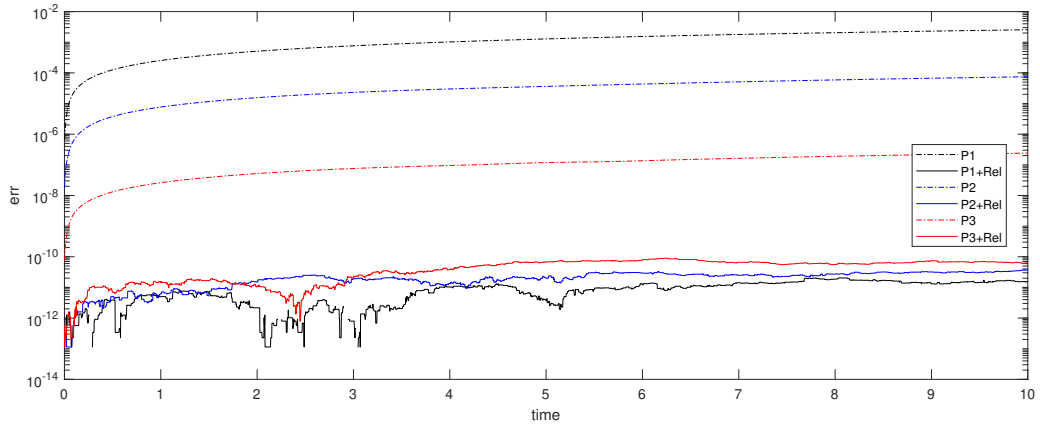


Figure 9: Entropy difference for the Shu vortex test case.

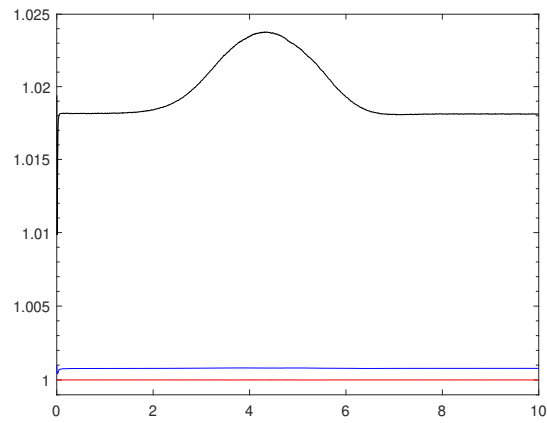


Figure 10: Values of  $\gamma^n$  for the Shu vortex test case.



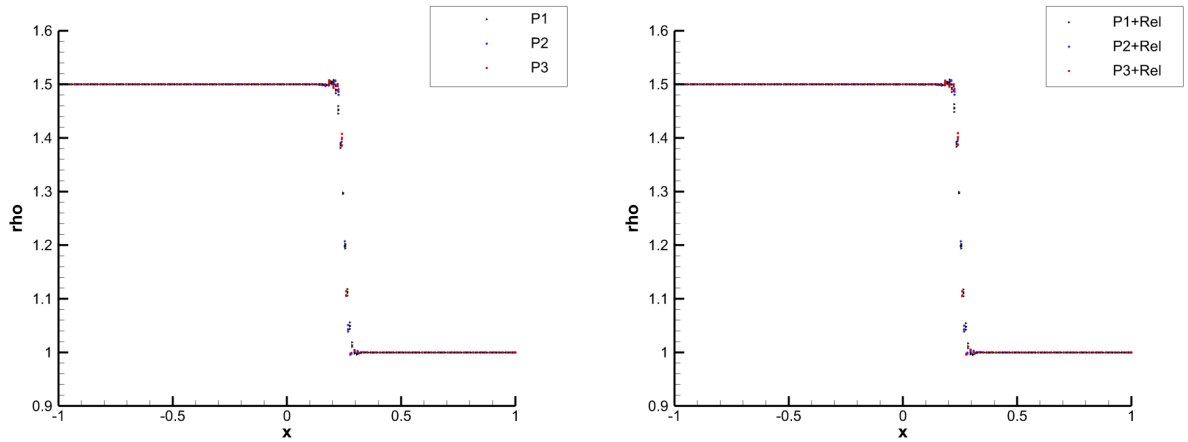


Figure 11: Simulation of contact discontinuity traveling in the domain. On the left we used ADER-DG, on the right ADER-DG with entropy correction and relaxation method.

### 5.3.2. Moving contact discontinuity

Now, we perform a simulation involving a moving contact discontinuity. We define a one dimensional Riemann problem on the rectangle  $\Omega = [-1, 1] \times [0, 1]$  with the following initial conditions for the left and right state

$$(\rho_L, u_L, v_L, p_L) = (1.5, 1, 0, 1) \text{ for } x < 0, \quad (\rho_R, u_R, v_R, p_R) = (1, 1, 0, 1) \text{ for } x \geq 0. \quad (56)$$

We have also to remark that we slightly smoothen the initial discontinuity (to avoid the need of the limiter) according to [67]

$$\mathbf{u} = \frac{1}{2} (\mathbf{u}_R + \mathbf{u}_L) + \frac{1}{2} (\mathbf{u}_R - \mathbf{u}_L) \operatorname{erf}\left(\frac{x}{2\bar{h}}\right), \quad (57)$$

where  $\bar{h}$  is the average mesh size. The solution of this Riemann problem is a contact discontinuity, hence, the entropy is preserved across it in time. We consider periodic boundary conditions on the top and bottom boundaries and inflow/outflow at the left and right boundaries, and we run our simulations up to the final time  $t_f = 0.25$ . We can compute analytically the amount of total entropy that leave the domain at the right boundary to compute the error with respect to the analytical solution.

In Figure 11 we show the final solution for the ADER-DG (left) and the relaxation ADER-DG (right). We run the  $\mathbb{P}_1$  scheme over a mesh of 5516 elements (16548 DOFs), the  $\mathbb{P}_2$  scheme over a mesh of 2860 elements (17160 DOFs), and the  $\mathbb{P}_3$  scheme over a mesh of 1744 elements (17440 DOFs). We remark that both schemes do not have a limiter suited for shocks, hence small overshoots are visible at the sides of the contact discontinuity. The difference between the two simulations is not so evident, but, again, in classical methods there is a loss of total entropy. In Figure 12 the total entropy of classical methods is lower than the exact one, and it is very similar for all the degrees of freedom. Here, we do not see huge advantages with a high order method as the solution is discontinuous. On the other side, for relaxation methods the exact value of the total entropy is preserved up to machine precision. Finally, in Figure 13 we show the values of  $\gamma^n$  in time and we can see that they are much larger than in previous tests. This is again due to the discontinuous solution that makes the accuracy of the scheme decrease, needing a larger modification to obtain the total entropy conservation.

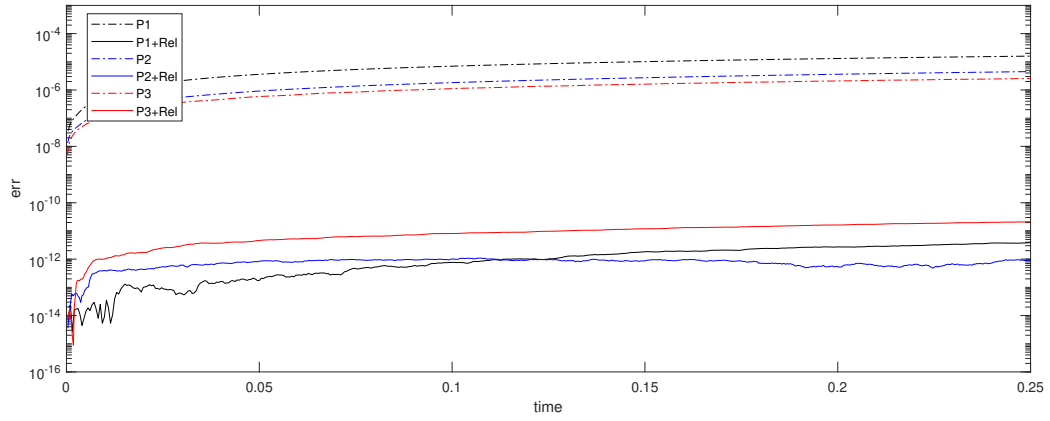


Figure 12: Entropy error for contact discontinuity test case.

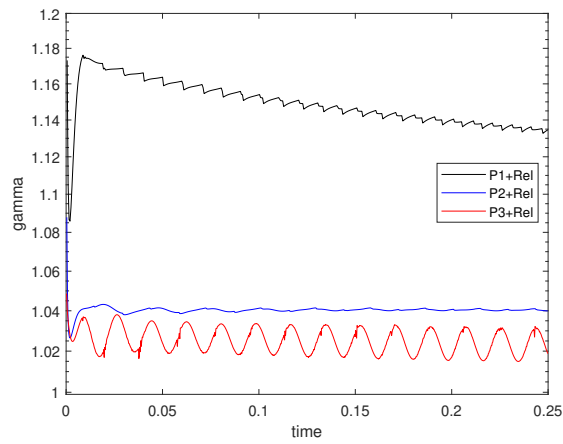


Figure 13: Values of  $\gamma^n$  in time with relaxation algorithms for different orders but similar number of DOFs.

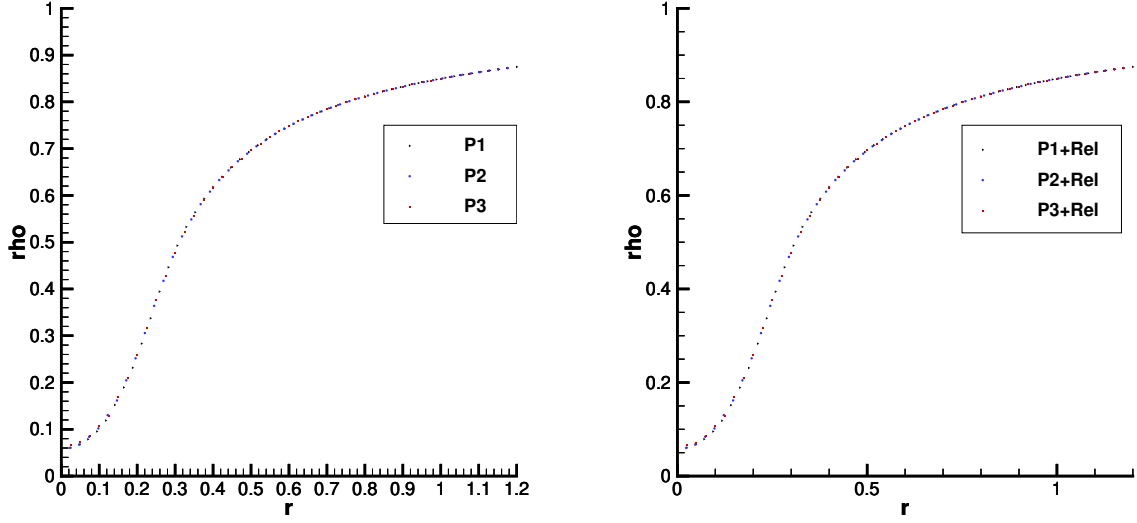


Figure 14: Simulation of the 123 problem. On the left we used ADER-DG, on the right ADER-DG with entropy correction and relaxation method.

### 5.3.3. 123 Problem

The last test that we propose is a two dimensional generalization of the 123 problem also presented in [69, Section 4.3.3]. It is defined on a domain  $\Omega = [-L, L]^2$  with  $L = 1.2$ , with initial condition

$$\begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} = \begin{pmatrix} \rho_0 \\ u \\ v \\ p_0 \end{pmatrix} = \begin{pmatrix} 1.0 \\ \frac{x}{\|x\|+1e-4} \\ v_0 \\ 0.4 \end{pmatrix}, \quad (58)$$

with  $v_0 = 2.0$ . We impose transmissive boundary conditions on all the domain.

First, we plot the profile of the density solution at the final time  $t_f = 0.15$  on the line  $y = 0$ . In Figure 14 we compare the classical ADER-DG and the relaxation ADER-DG methods. Again the different orders are run on different meshes so that the DOFs are more or less similar. We run the  $P_1$  scheme over a mesh of 10108 elements (30324 DOFs), the  $P_2$  scheme over a mesh of 5130 elements (30780 DOFs), and the  $P_3$  scheme over a mesh of 3388 elements (33880 DOFs). No visible differences are observable.

In Figure 15 we plot again the solution  $\rho$  and  $\eta(\mathbf{u})$  at the final time  $t_f = 0.15$  on the whole domain with a  $P_3$  ADER-DG scheme on a coarse mesh with 5130 elements. In this case, the boundary is affected by the propagation of the solution, hence, we do not know an analytical solution even for the total entropy. Hence, to obtain the error of the total entropy error in Figure 16 we compute numerically the integral of the entropy flux coming out at the boundary of the domain to approximate the exact value of the total entropy, i.e.,

$$\mathcal{E}(\mathbf{u}(t^n)) \approx \mathcal{E}(\mathbf{u}(t^{n-1})) - \sum_{s=0}^n \beta^s \Delta t^n \gamma^n \int_{\partial\Omega} \mathbf{G}(\mathbf{u}_h^{(s)}) \cdot \mathbf{n} \, dS. \quad (59)$$

In Figure 16 we observe again the usual behavior, the classical schemes preserve the total entropy only up to the accuracy of the method and, since the solution starts with a discontinuity in the velocity field, there is no large difference between  $P_2$  and  $P_3$ . For the relaxation ADER-DG we are able to preserve the exact total entropy up to machine precision for all the orders.

Finally, in Figure 17 we plot the  $\gamma^n$  values and we can observe that there is no large difference between  $P_2$  and  $P_3$ , as seen already for the classical methods.

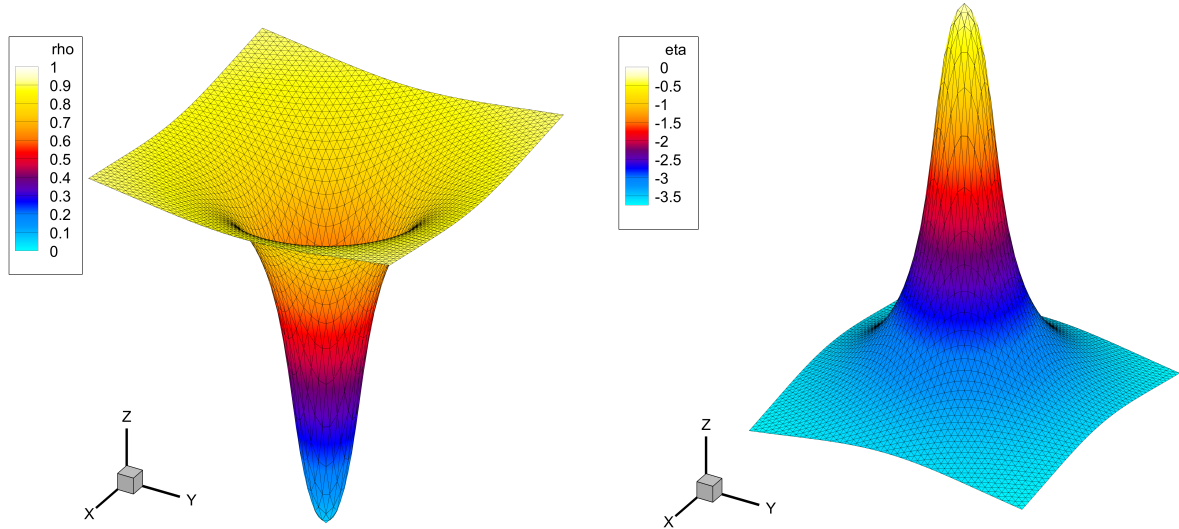


Figure 15: 123 Problem. In the figure we report the profile of the density  $\rho$  (left) and the entropy  $\eta$  (right) at time  $t = 0.15$  obtained with a DG scheme of order 3 with relaxation on a very coarse mesh made of 5130 elements.

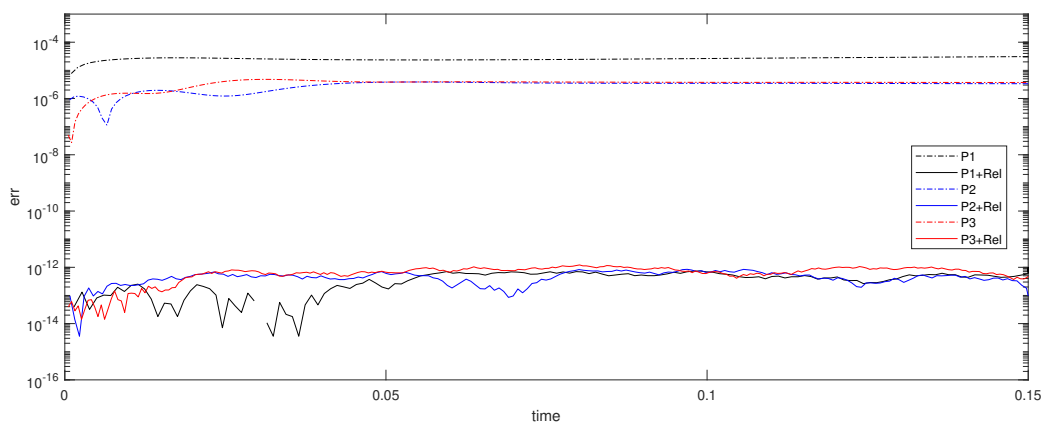


Figure 16: Entropy error for the 123 problem.

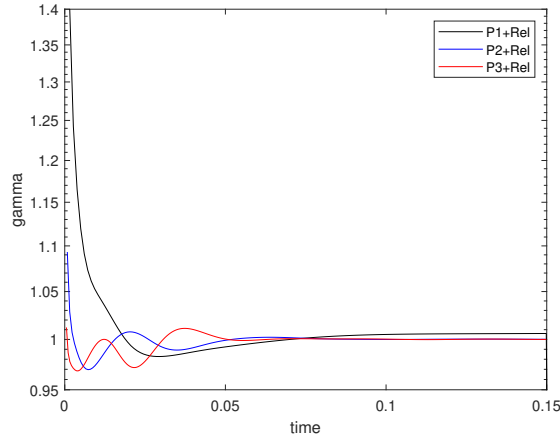


Figure 17: Values of  $\gamma^n$  in time with relaxation algorithms for different orders but similar number of DOFs.

**Remark 5.1** (Entropy stability of the corrected ADER-DG without relaxation). *It is interesting to notice that for most examples we have measured values of  $\gamma^n > 1$ . This means that the underlying ADER-DG scheme with centered flux and entropy correction is entropy stable. This means that the relaxation procedure is not required for stability, and that it must over-compensate with entropy production in time to ensure conservation up to machine precision. The only exception to this is the last test in which values of  $\gamma^n < 1$  have been observed in the initial instants of the simulation. In this case relaxation indeed is necessary to ensure at least conservation.*

## 6. Conclusions

In this work we have presented an enhancement for the ADER-DG arbitrarily high order method to provably conserve or dissipate the total entropy of the solution. The method is based on two ingredients. The first is an entropy correction term that allows to balance the spatial entropy production/destruction in each cell of the domain. Moreover, with the ADER-DG discretization is then possible to split the spatial discretization into an entropy conservative term and an entropy diffusive one. The second ingredient is the relaxation method, which exploits the split spatial discretization and it is able to guarantee a zero entropy production also during the time discretization for the entropy conservative terms. By construction the proposed scheme is able to exactly preserve the total entropy for many tests where physically this happens but, for the classical ADER-DG, this can not be guaranteed. Indeed, we have tested the method on smooth solutions and a contact discontinuity and in all cases the method has proved its properties. The tests have shown that for most cases the entropy correction with ADER-DG without relaxation provides in practice an entropy stable method. The relaxation procedure allows to correct this by over-compensating with entropy production in time allowing a control of entropy conservation within machine accuracy. For some tests, the usual behaviour is observed in which relaxation allows to reduce the entropy production in time. Overall, this shows the capabilities of the mix entropy correction/relaxation to provide full control on the entropic properties of the scheme.

The method as it is has still some limitations as we started from an unlimited version of the ADER-DG. In order to be able to deal also with shocks and more challenging tests, it will be necessary to introduce a limiter and to insert it in the spatial discretization split so that also in the relaxation process the dissipation will occur only when needed. Another extension of the method on which the authors are working on, is the  $\mathbb{P}^N\mathbb{P}^M$  method [32], to obtain entropy conservative/dissipative schemes.

## Acknowledgment

E. G. and M. R. are members of the CARDAMOM team at the Inria center of the university of Bordeaux. D. T. has been funded by a postdoctoral fellowship in Inria Bordeaux at the team CARDAMOM and by a postdoctoral

fellowship in SISSA Trieste. P. O. gratefully acknowledge the support of the Gutenberg Research College and also wants to thank M. R. for his invitation to Inria Bordeaux. E. G. gratefully acknowledges the support received from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Individual Fellowship *SuPerMan*, grant agreement No. 101025563.

## Appendix A. Proof that $R'(0) < 0$

**Theorem Appendix A.1** (Monotonicity of  $R$ ). *The function  $R$  defined in (31) is monotone decreasing if  $h, \Delta t \ll 1$  and  $\alpha_{\Omega_i}^s = O(h^2)$ .*

*Proof.* Let us introduce the forms  $a_s : (V_h)^m \times (V_h)^m \rightarrow \mathbb{R}$ , where  $V_h = \{u \in L^2(\Omega) : u|_{\Omega_i} \in C^1(\Omega_i)\}$ , as

$$\begin{aligned} a^s(\mathbf{z}, \mathbf{q}_h) &:= \sum_{\Omega_i \in \Omega} \left( - \int_{\partial\Omega_i} \langle \mathbf{z}(\mathbf{x}), \mathcal{F}_n(\mathbf{q}_h^-(\mathbf{x}), \mathbf{q}_h^+(\mathbf{x})) \rangle dS + \int_{\Omega_i} \nabla \mathbf{z}(\mathbf{x}) \cdot \mathbf{F}(\mathbf{q}_h(\mathbf{x})) d\mathbf{x} - \alpha_{\Omega_i}^s \int_{\Omega_i} \nabla \mathbf{z}(\mathbf{x}) A_0 \nabla v_h(\mathbf{q}_h(\mathbf{x})) d\mathbf{x} \right) \\ &= - \sum_{\Omega_i \in \Omega} \int_{\Omega_i} \langle \mathbf{z}(x), \nabla \cdot \mathbf{F}(\mathbf{q}_h(\mathbf{x})) \rangle d\mathbf{x} + O(h) = \sum_{\Omega_i \in \Omega} \int_{\Omega_i} \langle \mathbf{z}(x), \partial_t(\mathbf{q}_h(\mathbf{x})) \rangle d\mathbf{x} + O(h), \end{aligned} \quad (\text{A.1})$$

where the last steps used the regularity of the approximated solution  $\mathbf{q}_h$  to obtain at least that  $\alpha_{\Omega_i}^s = O(h^2)$ , see [2].  $a^s$  are linear in the first argument. Now, using (16) we notice that

$$\int_{\Omega} \langle \mathbf{z}, \Delta \mathbf{u}_h \rangle d\mathbf{x} = \Delta t \sum_{s=0}^S \beta_s a^s(\mathbf{z}, \mathbf{q}_h^s).$$

Then, with (A.1) and the definitions of  $\mathbf{E}_i^s$ ,  $\mathbf{D}_i^s$  and  $\mathbf{F}_i^s$ , we have

$$\begin{aligned} R'(0) &= \int_{\Omega} \langle \eta'(\mathbf{u}_h^n), \Delta \mathbf{u}_h \rangle d\mathbf{x} + \Delta t \sum_{s=0}^S \beta_s \sum_{i \in \Omega} (\mathbf{F}_i^s + \alpha_{\Omega_i}^s \mathbf{E}_i^s + \mathbf{D}_i^s) \\ &= \Delta t \sum_{s=0}^S \beta_s a^s(\eta'(\mathbf{u}_h^n), \mathbf{q}_h^s) - \Delta t \sum_{s=0}^S \beta_s a^s(\eta'(\mathbf{q}_h^s), \mathbf{q}_h^s) = \Delta t \sum_{s=0}^S \beta_s a^s(\eta'(\mathbf{u}_h^n) - \eta'(\mathbf{q}_h^s), \mathbf{q}_h^s) \\ &= \Delta t \sum_{s=0}^S \beta_s a^s \left( (\mathbf{u}_h^n - \mathbf{q}_h^s)^T \int_0^1 \eta''(\mathbf{u}_h^n - r(\mathbf{u}_h^n - \mathbf{q}_h^s)) dr, \mathbf{q}_h^s \right). \end{aligned}$$

We observe that  $\mathbf{q}_h^s - \mathbf{u}_h^n \approx \beta_s \Delta t \partial_t \mathbf{u}_h^n$  up to  $O(\Delta t^2)$  as well as  $\Delta \mathbf{u}_h = \mathbf{u}_h^{n+1} - \mathbf{u}_h^n \approx \Delta t \partial_t \mathbf{u}_h^n$ , where with  $\partial_t \mathbf{u}_h^n$  we mean  $\partial_t \mathbf{u}_h(x, t)|_{t=\tau^n}$ . We also approximate  $\int_0^1 \eta''(\mathbf{u}_h^n - r(\mathbf{u}_h^n - \mathbf{q}_h^s)) dr \approx \int_0^1 \eta''(\mathbf{u}_h^n + r \Delta t \partial_t \mathbf{u}_h^n) dr =: \tilde{A}$  up to  $O(\Delta t^2)$ , where  $\tilde{A}$  is a positive definite matrix. So, we can write

$$\begin{aligned} R'(0) &= -\Delta t^2 \sum_{s=0}^S \beta_s a^s \left( \beta_s \partial_t (\mathbf{u}_h^n)^T \tilde{A} + O(\Delta t^2), \mathbf{q}_h^s \right) = -\Delta t^2 \sum_{s=0}^S \beta_s^2 \sum_{\Omega_i \in \Omega} \langle \partial_t (\mathbf{u}_h^n)^T \tilde{A}, \partial_t (\mathbf{q}_h^s(\mathbf{x})) \rangle d\mathbf{x} + O(\Delta t^3) \\ &= -\Delta t^2 \sum_{s=0}^S \beta_s^2 \sum_{\Omega_i \in \Omega} \langle \partial_t (\mathbf{u}_h^n)^T \tilde{A}, \partial_t (\mathbf{u}_h^n(\mathbf{x})) \rangle d\mathbf{x} + O(\Delta t^3) < 0. \end{aligned}$$

Hence, the thesis follows.  $\square$

## References

- [1] R. Abgrall and M. Ricchiuto. High order methods for CFD. In Rene de Borst Erwin Stein and Thomas J.R. Hughes, editors, *Encyclopedia of Computational Mechanics, Second Edition*. John Wiley and Sons, 2017.
- [2] Rémi Abgrall. A general framework to construct schemes satisfying additional conservation relations. application to entropy conservative and entropy dissipative schemes. *Journal of Computational Physics*, 372:640–666, 2018.
- [3] Rémi Abgrall, Élise Le Mélédo, Philipp Öffner, and Davide Torlo. Relaxation Deferred Correction Methods and their Applications to Residual Distribution Schemes. *The SMAI Journal of computational mathematics*, 8:125–160, 2022.
- [4] Rémi Abgrall, Jan Nordström, Philipp Öffner, and Svetlana Tokareva. Analysis of the SBP-SAT stabilization for finite element methods part II: Entropy stability. *Communications on Applied Mathematics and Computation*, pages 1–23, 2020.
- [5] Rémi Abgrall, Philipp Öffner, and Hendrik Ranocha. Reinterpretation and Extension of Entropy Correction Terms for Residual Distribution and Discontinuous Galerkin Schemes: Application to Structure Preserving Discretization. *Journal of Computational Physics*, page 110955, 2022.
- [6] Dinshaw S. Balsara, Chad Meyer, Michael Dumbser, Huijing Du, and Zhiliang Xu. Efficient implementation of ADER schemes for Euler and magnetohydrodynamical flows on structured meshes – speed comparisons with Runge-Kutta methods. *J. Comput. Phys.*, 235:934–969, 2013.
- [7] Walter Boscheri, Simone Chiochetti, and Ilya Peshkov. A cell-centered implicit-explicit Lagrangian scheme for a unified model of nonlinear continuum mechanics on unstructured meshes. *Journal of Computational Physics*, 451:110852, 2022.
- [8] Walter Boscheri and Michael Dumbser. Arbitrary-Lagrangian-Eulerian discontinuous Galerkin schemes with a posteriori subcell finite volume limiting on moving unstructured meshes. *Journal of Computational Physics*, 346:449 – 479, 2017.
- [9] Walter Boscheri, Michael Dumbser, and D.S. Balsara. High-order ADER-WENO ALE schemes on unstructured triangular meshes—application of several node solvers to hydrodynamics and magnetohydrodynamics. *International Journal for Numerical Methods in Fluids*, 76(10):737–778, 2014.
- [10] Mark H Carpenter, Travis C Fisher, Eric J Nielsen, and Steven H Frankel. Entropy stable spectral collocation schemes for the Navier–Stokes equations: Discontinuous interfaces. *SIAM Journal on Scientific Computing*, 36(5):B835–B867, 2014.
- [11] G Carré, S Del Pino, Bruno Després, and Emmanuel Labourasse. A cell-centered Lagrangian hydrodynamics scheme on general unstructured meshes in arbitrary dimension. *Journal of Computational Physics*, 228(14):5160–5183, 2009.
- [12] Agnes Chan, Gérard Gallice, Raphaël Loubère, and Pierre-Henri Maire. Positivity preserving and entropy consistent approximate Riemann solvers dedicated to the high-order MOOD-based Finite Volume discretization of Lagrangian and Eulerian gas dynamics. *Computers & Fluids*, 229:105056, 2021.
- [13] Praveen Chandrashekar. Kinetic energy preserving and entropy stable finite volume schemes for compressible Euler and Navier-Stokes equations. *Commun. Comput. Phys.*, 14(5):1252–1286, 2013.
- [14] Praveen Chandrashekar and Christian Klingenberg. Entropy stable finite volume scheme for ideal compressible MHD on 2-D Cartesian meshes. *SIAM J. Numer. Anal.*, 54(2):1313–1340, 2016.
- [15] Tianheng Chen and Chi-Wang Shu. Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *Journal of Computational Physics*, 345:427–461, 2017.
- [16] Simone Chiochetti, Ilya Peshkov, Sergey Gavriluk, and Michael Dumbser. High order ADER schemes and GLM curl cleaning for a first order hyperbolic formulation of compressible flow with surface tension. *Journal of Computational Physics*, 426:109898, 2021.
- [17] Mirco Ciallella, Davide Torlo, and Mario Ricchiuto. Arbitrary high order weno finite volume scheme with flux globalization for moving equilibria preservation. *arXiv preprint arXiv:2205.13315*, 2022.
- [18] Frédéric Coquel, Claude Marmignon, Pratik Rai, and Florent Renac. An entropy stable high-order discontinuous Galerkin spectral element method for the Baer-Nunziato two-phase flow model. *Journal of Computational Physics*, 431:110135, 2021.
- [19] Jared Crean, Jason E. Hicken, David C. Del Rey Fernández, David W. Zingg, and Mark H. Carpenter. Entropy-stable summation-by-parts discretization of the Euler equations on general curved elements. *J. Comput. Phys.*, 356:410–438, 2018.
- [20] David C. Del Rey Fernández, Mark H. Carpenter, Lisandro Dalcin, Lucas Fredrich, Andrew R. Winters, Gregor J. Gassner, and Matteo Parsani. Entropy-stable  $p$ -nonconforming discretizations with the summation-by-parts property for the compressible Navier-Stokes equations. *Comput. Fluids*, 210:14, 2020. Id/No 104631.
- [21] R Dematté, VA Titarev, GI Montecinos, and EF Toro. ADER methods for hyperbolic equations with a time-reconstruction solver for the generalized Riemann problem: the scalar case. *Communications on Applied Mathematics and Computation*, 2(3):369–402, 2020.
- [22] Junming Duan and Huazhong Tang. Entropy stable adaptive moving mesh schemes for 2D and 3D special relativistic hydrodynamics. *Journal of Computational Physics*, 426:109949, 2021.
- [23] Michael Dumbser, Dinshaw S Balsara, Eleuterio F Toro, and Claus-Dieter Munz. A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes. *Journal of Computational Physics*, 227(18):8209–8253, 2008.
- [24] Michael Dumbser, Francesco Fambri, Elena Gaburro, and Anne Reinarz. On GLM curl cleaning for a first order reduction of the CCZ4 formulation of the Einstein field equations. *Journal of Computational Physics*, 404:109088, 2020.
- [25] Michael Dumbser and Claus-Dieter Munz. ADER discontinuous Galerkin schemes for aeroacoustics. *Comptes Rendus Mécanique*, 333:683–687, 2005.
- [26] Francesco Fambri, Michael Dumbser, Sven Köppel, Luciano Rezzolla, and Olindo Zanotti. Ader discontinuous Galerkin schemes for general-relativistic ideal magnetohydrodynamics. *Monthly Notices of the Royal Astronomical Society*, 477(4):4543–4564, 2018.
- [27] Travis C. Fisher and Mark H. Carpenter. High-order entropy stable finite difference schemes for nonlinear conservation laws: finite domains. *J. Comput. Phys.*, 252:518–557, 2013.
- [28] Travis C. Fisher, Mark H. Carpenter, Jan Nordström, Nail K. Yamaleev, and Charles Swanson. Discretely conservative finite-difference formulations for nonlinear conservation laws in split form: theory and boundary conditions. *J. Comput. Phys.*, 234:353–375, 2013.
- [29] Lucas Friedrich, Gero Schnücker, Andrew R. Winters, David C. Del Rey Fernández, Gregor J. Gassner, and Mark H. Carpenter. Entropy



- stable space-time discontinuous Galerkin schemes with summation-by-parts property for hyperbolic conservation laws. *J. Sci. Comput.*, 80(1):175–222, 2019.
- [30] Elena Gaburro. A unified framework for the solution of hyperbolic PDE systems using high order direct Arbitrary-Lagrangian–Eulerian schemes on moving unstructured meshes with topology change. *Archives of Computational Methods in Engineering*, 28(3):1249–1321, 2021.
- [31] Elena Gaburro, Walter Boscheri, Simone Chiochetti, Christian Klingenberg, Volker Springel, and Michael Dumbser. High order direct arbitrary-Lagrangian–Eulerian schemes on moving Voronoi meshes with topology changes. *J. Comput. Phys.*, 407:44, 2020. Id/No 109167.
- [32] Elena Gaburro and Michael Dumbser. A Posteriori Subcell Finite Volume Limiter for General PNPM Schemes: Applications from Gasdynamics to Relativistic Magnetohydrodynamics. *Journal of Scientific Computing*, 86(3):1–41, 2021.
- [33] Gregor J. Gassner, Andrew R. Winters, Florian J. Hindenlang, and David A. Kopriva. The BR1 scheme is stable for the compressible Navier-Stokes equations. *J. Sci. Comput.*, 77(1):154–200, 2018.
- [34] Gregor J Gassner, Andrew R Winters, and David A Kopriva. Split form nodal discontinuous Galerkin schemes with summation-by-parts property for the compressible Euler equations. *Journal of Computational Physics*, 327:39–66, 2016.
- [35] Gregor J Gassner, Andrew R Winters, and David A Kopriva. A well balanced and entropy conservative discontinuous Galerkin spectral element method for the shallow water equations. *Applied Mathematics and Computation*, 272:291–308, 2016.
- [36] Maria Han Veiga, Philipp Öffner, and Davide Torlo. DeC and ADER: similarities, differences and a unified framework. *J. Sci. Comput.*, 87(1):35, 2021. Id/No 2.
- [37] Amiram Harten. On the symmetric form of systems of conservation laws with entropy. *J. Comput. Phys.*, 49:151–164, 1983.
- [38] Dorian Hillebrand, Simon-Christian Klein, and Philipp Öffner. Comparison to control oscillations in high-order Finite Volume schemes via physical constraint limiters, neural networks and polynomial annihilation. *arXiv preprint arXiv:2203.00297*, 2022.
- [39] Shinho Kang and Emil M. Constantinescu. Entropy-preserving and entropy-stable relaxation IMEX and multirate time-stepping methods. *J. Sci. Comput.*, 93(1):31, 2022. Id/No 23.
- [40] David I. Ketcheson. Relaxation Runge-Kutta methods: conservation and stability for inner-product norms. *SIAM J. Numer. Anal.*, 57(6):2850–2870, 2019.
- [41] David A Kopriva, Gregor J Gassner, and Jan Nordström. On the theoretical foundation of overset grid methods for hyperbolic problems ii: Entropy bounded formulations for nonlinear conservation laws. *Journal of Computational Physics*, page 111620, 2022.
- [42] Dmitri Kuzmin. Entropy stabilization and property-preserving limiters for P1 discontinuous Galerkin discretizations of scalar hyperbolic problems. *Journal of Numerical Mathematics*, 29(4):307–322, 2021.
- [43] Dmitri Kuzmin and Manuel Quezada de Luna. Algebraic entropy fixes and convex limiting for continuous finite element discretizations of scalar hyperbolic conservation laws. *Computer Methods in Applied Mechanics and Engineering*, 372:113370, 2020.
- [44] Dmitri Kuzmin, Hennes Hajduk, and Andreas Rupp. Limiter-based entropy stabilization of semi-discrete and fully discrete schemes for nonlinear hyperbolic problems. *Comput. Methods Appl. Mech. Eng.*, 389:28, 2022. Id/No 114428.
- [45] Yogiraj Mantri, Philipp Öffner, and Mario Ricchiuto. Fully well balanced and entropy conservative global flux DGSEM for shallow water flows. *in preparation*, 2022.
- [46] Juan Manzanero, Gonzalo Rubio, David A Kopriva, Esteban Ferrer, and Eusebio Valero. An entropy–stable discontinuous Galerkin approximation for the incompressible Navier–Stokes equations with variable density and artificial compressibility. *Journal of Computational Physics*, 408:109241, 2020.
- [47] Juan Manzanero, Gonzalo Rubio, David A Kopriva, Esteban Ferrer, and Eusebio Valero. Entropy–stable discontinuous Galerkin approximation with summation–by–parts property for the incompressible Navier–Stokes/Cahn–Hilliard system. *Journal of Computational Physics*, 408:109363, 2020.
- [48] Claude Marmignon, Fabio Naddei, and Florent Renac. Energy relaxation approximation for compressible multicomponent flows in thermal nonequilibrium. *Numerische Mathematik*, 151(1):151–184, 2022.
- [49] Philipp Öffner, Jan Glaubit, and Hendrik Ranocha. Stability of correction procedure via reconstruction with summation-by-parts operators for Burgers’ equation using a polynomial chaos approach. *ESAIM, Math. Model. Numer. Anal.*, 52(6):2215–2245, 2018.
- [50] Martin Parisot. Entropy-satisfying scheme for a hierarchy of dispersive reduced models of free surface flow. *International Journal for Numerical Methods in Fluids*, 91(10):509–531, 2019.
- [51] Gabriella Puppo and Matteo Semplice. Numerical entropy and adaptivity for finite volume schemes. *Communications in Computational Physics*, 10(5):1132–1160, 2011.
- [52] Hendrik Ranocha, Jan Glaubit, Philipp Öffner, and Thomas Sonar. Stability of artificial dissipation and modal filtering for flux reconstruction schemes using summation-by-parts operators. *Appl. Numer. Math.*, 128:1–23, 2018.
- [53] Hendrik Ranocha, Lajos Lóczy, and David I. Ketcheson. General relaxation methods for initial-value problems with application to multistep schemes. *Numerische Mathematik*, 146(4):875–906, 2020.
- [54] Hendrik Ranocha, Mohammed Sayyari, Lisandro Dalcin, Matteo Parsani, and David I. Ketcheson. Relaxation Runge-Kutta methods: fully discrete explicit entropy-stable schemes for the compressible Euler and Navier-Stokes equations. *SIAM J. Sci. Comput.*, 42(2):a612–a638, 2020.
- [55] Deep Ray and Praveen Chandrashekar. Entropy stable schemes for compressible Euler equations. *Int. J. Numer. Anal. Model., Ser. B*, 4(4):335–352, 2013.
- [56] Deep Ray, Praveen Chandrashekar, Ulrik S Fjordholm, and Siddhartha Mishra. Entropy stable scheme on two-dimensional unstructured grids for Euler equations. *Communications in Computational Physics*, 19(5):1111–1140, 2016.
- [57] Florent Renac. Entropy stable DGSEM for nonlinear hyperbolic systems in nonconservative form with application to two-phase flows. *J. Comput. Phys.*, 382:1–26, 2019.
- [58] Florent Renac. Entropy stable, robust and high-order DGSEM for the compressible multicomponent Euler equations. *J. Comput. Phys.*, 445:28, 2021. Id/No 110584.
- [59] Mario Ricchiuto and Andreas Bollermann. Stabilized residual distribution for shallow water simulations. *Journal of Computational Physics*, 228(4):1071–1115, 2009.

- [60] Mario Ricchiuto and Davide Torlo. Analytical travelling vortex solutions of hyperbolic equations for validating very high order schemes. *arXiv preprint:2109.10183*, 2021.
- [61] Giulia Rossi, Michael Dumbser, and Aronne Armanini. A well-balanced path conservative SPH scheme for nonconservative hyperbolic systems with applications to shallow water and multi-phase flows. *Computers and Fluids*, 154:102–122, 2017.
- [62] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock capturing schemes. *Journal of Computational Physics*, 77:439–471, 1988.
- [63] Björn Sjögreen and HC Yee. Entropy stable method for the Euler equations revisited: central differencing via entropy splitting and SBP. *Journal of Scientific Computing*, 81(3):1359–1385, 2019.
- [64] A.H. Stroud. *Approximate Calculation of Multiple Integrals*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1971.
- [65] Eitan Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. i. *Mathematics of Computation*, 49(179):91–103, 1987.
- [66] Eitan Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12:451–512, 2003.
- [67] Maurizio Tavelli and Michael Dumbser. A high order semi-implicit discontinuous Galerkin method for the two dimensional shallow water equations on staggered unstructured meshes. *Applied Mathematics and Computation*, 234:623–644, 2014.
- [68] V. A. Titarev and E. F. Toro. ADER: Arbitrary high-order Godunov approach. *J. Sci. Comput.*, 17(1-4):609–618, 2002.
- [69] Eleuterio F Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, second edition, 1999.
- [70] Eleuterio F Toro, RC Millington, and LAM Nejad. Towards very high order Godunov schemes. In *Godunov methods. Theory and applications. International conference, Oxford, GB, October 1999*, pages 907–940. New York, NY: Kluwer Academic/ Plenum Publishers, 2001.
- [71] Niklas Wintermeyer, Andrew R Winters, Gregor J Gassner, and David A Kopriva. An entropy stable nodal discontinuous Galerkin method for the two dimensional shallow water equations on unstructured curvilinear meshes with discontinuous bathymetry. *Journal of Computational Physics*, 340:200–242, 2017.
- [72] Nail K. Yamaleev, David C. Del Rey Fernández, Jialin Lou, and Mark H. Carpenter. Entropy stable spectral collocation schemes for the 3-D Navier-Stokes equations on dynamic unstructured grids. *J. Comput. Phys.*, 399:27, 2019. Id/No 108897.