Shock Capturing Methods in High-Order Flux Reconstruction: Low-Order Graph Viscosity and Convex Limiting Approaches

W. Trojak*, T. Dzanic†, and F. D. Witherden‡
Department of Ocean Engineering, Texas A&M University, College Station, TX 77843

In this work, a method for constructing schemes is presented that enables shock-capturing in high-order nodal discontinuous spectral element methods that is devoid of tunable parameters. A high-order flux reconstruction scheme and a low-order summation-by-parts scheme are introduced and combined via a convex limiting approach to increase the fidelity of the solution without sacrificing physicality. The computational overhead of the limiting procedure is minimised by applying an entropy viscosity approach in conjunction with an entropy residual shock sensor. Numerical results for the Euler equations demonstrate that the method is able to accurately resolve discontinuous solutions even in challenging cases with near-vacuum conditions and large magnitude shock waves.

I. Introduction

Shock capturing has been an important aspect of computational fluid dynamics (CFD) as the presence of shocks is typical in many flows of interest. Early computational approaches, such as Harlow's particle-in-cell method [1], often used straightforward central difference type approaches primarily due to resource limitations. However, as available computational resources increased, more complex approaches could be considered. Subsequent methods still made use of small stencils and were primarily first or second-order, but improved upon central difference type approaches with methods such as the work of Rusanov [2] and, later, methods inspired by Godunov's work [3], such as that of van Leer [4] and Harten et al. [5]. These methods can be broadly classified as approaches that aim to resolve discontinuities through imposing physical assumptions on the discretised governing equations. Around this time, Jameson et al. [6] developed a scheme which took a somewhat different approach by applying second and fourth-order artificial dissipation to the solution in order remove oscillations and improve shock resolution.

In regions away from discontinuities, these schemes are generally limited by their first-order accuracy which brings about high grid requirements, and in response, high-order methods were developed. One early method was the total variation diminishing (TVD) scheme of Harten [7] that aimed to detect oscillations and modify the flux function to remove them. Although not a high-order method, the techniques were applied in a high-order framework to produce the essentially non-oscillatory scheme (ENO) [8] and, later, the weighted ENO (WENO) scheme [9]. These methods have been highly successful in approximating solutions with discontinuities such as shock waves and phase interfaces in multi-physics simulations.

Another class of high-order methods are spectral element methods (SEM) which have become increasingly popular as they combine the efficiency of spectral methods with the geometric flexibility of finite volume schemes. The high-order accuracy in this setting is achieved through an element-wide stencil which can have the effect of bringing information from outside of the physical domain of dependence into the solution. This, coupled with the issues of Gibbs phenomenon [10, 11], means that discontinuities pose challenges when they occur in high-order simulations, resulting in the same behaviour that ENO and WENO methods sought to confront. The focus of this work is on a particular SEM called flux reconstruction (FR) [12, 13], a scheme closely related to discontinuous Galerkin (DG) methods [14, 15].

There are several shock capturing schemes that have been developed for DG methods, most of which are applicable to FR. These schemes can again be classified as either physics based or diffusion based. The method of Persson and Peraire [16] , where the modal energy decay rate is used to activate and scale an additional diffusive term, is typical of many AV approaches that may be applied in conjunction with SEM to facilitate shock capturing. The benefit of this method is its simplicity and, in some cases, its relative effectiveness. Another related example is that of Barter and Darmofal [17]. These methods can excel at shock capturing in many applications, but for optimal performance to be achieved, the multiple control parameters typically require case-specific optimisation. A mathematically connected set

*Post Doctoral Research Scholar, AIAA Member
†PhD Candidate, AIAA Student Member
‡Assistant Professor, AIAA Member

of methods are the filtering approaches introduced by Tadmor [18] for spectral and pseudo-spectral schemes. There are multiple adaptations and modifications (see Tadmor [19], Ma [20], Maday et al. [21]); however, the approach may be generalised as aiming to remove oscillations through some application of filtering, be that through a direct filtering operation or indirectly through the connection between filtering and hyper-viscosity. These methods can be effective, but they, like artificial viscosity approaches, also require some degree of parameter optimisation and are less coupled to the physical characteristics of the system being solved.

A branch of methods that offer more general applicability are sub-cell and graph viscosity methods. Research into sub-cell methods within DG frameworks have shown their utility in shock capturing; one such high-fidelity method is that of Dumbser and Loubère [22] which utilises a sub-cell WENO approach in troubled elements. This method has a low degree of parameterisation, but comes with a high degree of complexity. In a series of papers [23–25], the concept of graph-viscosity was introduced. This may be thought of as a natural generalisation of the finite volume sub-cell method to a high-order stencil in a finite element framework, and the abstract nature of this approach makes it amenable to many numerical methods. The approach relies on building a graph of dependence for each solution point and looks to apply an artificial viscosity term for each edge in the graph. To completely remove parameterisation, the artificial viscosity terms are calculated by posing each edge as a Riemann problem. Guermond et al. [25] were able to prove many advantageous properties of the scheme, chief among which is invariant domain preservation, a property which will be discussed in detail later in this text. However, achieving these properties comes with the consequence of excessive diffusion that increases with scheme order. It is this deficiency that is motivation for the present work, which aims to explore the application of this method to FR, adaptations to reduce the overly diffuse nature of the method, and the use of convex limiting techniques to improve solution quality.

Throughout this work, the approximate solution to hyperbolic equations of m conservation laws in d-dimensions is explored, written as

$$\partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) = 0, \quad \text{for} \quad (x,t) \in \mathbb{R} \times \mathbb{R}_+, \quad \text{and} \quad \mathbf{u}(x, t=0) = \mathbf{u}_0(x), \tag{1}$$

for a solution $\mathbf{u} \in \mathbb{R}^m$ and flux function $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^{m \times d}$.

This paper is structured with preliminaries on Riemann problems and invariant sets and domains in Section II, an introduction to the flux reconstruction method in Section III, the proposed shock capturing scheme in Section VI, numerical test cases and results in Section V, and conclusions in Section VI.

II. Admissible Solutions and Invariant Domains

In the approximation of solutions to hyperbolic conservation laws, several properties of the Riemann problem can be useful in proving the generalised stability properties of numerical schemes. In these preliminaries, some properties of the Riemann problem as well as invariant sets and invariant domain preserving methods will be presented. The theory behind these properties is closely related to the works of Lax [26], Glimm [27], Chueh et al. [28], Hoff [29], and Guermond and Popov [23].

The Riemann problem can be defined as

$$\partial_t \mathbf{u} + \partial_x (\mathbf{F} \cdot \mathbf{n}) = 0, \quad (x,t) \in \mathbb{R} \times \mathbb{R}_+, \tag{2a}$$

$$\mathbf{u}(x, t=0) = \begin{cases} \mathbf{u}_L, & \text{if } x < 0, \\ \mathbf{u}_R, & \text{if } x > 0, \end{cases} \tag{2b}$$

where $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{F} \in \mathbb{R}^{m \times d}$, and $\mathbf{n}$ is some normal $\mathbf{n} \in \partial B^d$ where ∂B^d is the boundary of a d-dimension closed unit ball B^d. There is an admissible set $\mathcal{A} \subset \mathbb{R}^m$ such that for $(\mathbf{u}_L, \mathbf{u}_R) \in \mathcal{A} \times \mathcal{A}$, there is a unique self-similiar solution to Eq. (2) given by $\mathbf{u}(x,t) = \mathbf{v}(x/t, \mathbf{n}, \mathbf{u}_L, \mathbf{u}_R) \in \mathcal{A}$ for $|\mathbf{u}_R - \mathbf{u}_L| < \delta$. This is a key result of Lax [26], which shows that, first, solutions exist provided that $\mathbf{u}_L$ and $\mathbf{u}_R$ are sufficiently close, and, second, these solutions are self-similar with respect to the parameter x/t.

If the Riemann problem characterised by $\mathbf{u}_L$, $\mathbf{u}_R$, and $\mathbf{n}$ has a maximum absolute eigenvalue $\lambda_{\max}$, then

$$\mathbf{u}(x,t) = \begin{cases} \mathbf{u}_L, & \text{if } x < -t\lambda_{\max}, \\ \mathbf{u}_R, & \text{if } x > t\lambda_{\max}, \end{cases} \tag{3}$$

i.e. there is region defined by a finite maximum wave speed of the system outside of which the solution takes the value of the initial condition [30]. As a result, the following lemma can be posed.

Lemma II.1. *Let* $(\mathbf{u}_L, \mathbf{u}_R) \in \mathcal{A} \times \mathcal{A}$ *and* $\mathbf{n} \in \partial B^d$*, and let* $\overline{\mathbf{v}}(t) = \int_{-0.5}^{0.5} \mathbf{v}(x/t, \mathbf{n}, \mathbf{u}_L, \mathbf{u}_R)\, \mathrm{d}x$ *define the average solution for the system in Eq.* (2)*. For a sufficiently small* t *such that* $t\lambda_{\max} < 0.5$*, the following relation holds.*

$$\overline{\mathbf{v}}(t) = \frac{1}{2}(\mathbf{u}_L + \mathbf{u}_R) - t(\mathbf{F}_R - \mathbf{F}_L) \cdot \mathbf{n}. \tag{4}$$

For the system defined in Eq. (1) and an entropy-flux pair $(\sigma, \boldsymbol{\Sigma})$, the entropy condition on the Riemann solution is taken to be

$$\partial_t \sigma\big(\mathbf{v}(\mathbf{n}, \mathbf{u}_L, \mathbf{u}_R)\big) + \partial_x \boldsymbol{\Sigma}\big(\mathbf{v}(\mathbf{n}, \mathbf{u}_L, \mathbf{u}_R)\big) \leqslant 0. \tag{5}$$

If the solution to Riemann problem satisfies this condition, the following corollary can be established.

Corollary II.1.1. *For an entropy-flux pair* $(\sigma, \boldsymbol{\Sigma})$ *that satisfies the entropy condition in Eq.* (5)*, the average solution in Eq.* (4) *for the Riemann problem in Eq.* (2) *satisfies the following inequality provided the CFL condition* $t\lambda_{\max} < 0.5$*.*

$$\sigma(\overline{\mathbf{v}}(t)) \leqslant \frac{1}{2}\big(\sigma(\mathbf{u}_R) + \sigma(\mathbf{u}_L)\big) - t(\boldsymbol{\Sigma}(\mathbf{u}_R) - \boldsymbol{\Sigma}(\mathbf{u}_L)) \cdot \mathbf{n}, \tag{6}$$

Through this inequality, an approximate solution to the Riemann problem can be shown to be entropy stable given a maximum eigenvalue and effective time. Using this, the concept of an invariant set and domain can be introduced which may be subsequently used to prove absolute stability of numerical schemes for hyperbolic conservation laws.

Definition II.1 (Invariant Set)**.** *A set* $\mathcal{B} \subset \mathcal{A} \subset \mathbb{R}$ *is called an invariant set of Eq.* (1) *if for any* $(\mathbf{u}_L, \mathbf{u}_R) \in \mathcal{B} \times \mathcal{B}$*,* $\mathbf{n} \in \partial B^d$*, and* $t > 0$*, the entropy stable solution over the fan for Eq.* (2) *is also in* $\mathcal{B}$*, i.e.* $\overline{\mathbf{v}}(\mathbf{n}, \mathbf{u}_L, \mathbf{u}_R) \in \mathcal{B}$ *for* $0 < t < \frac{1}{2}\lambda_{\max}$*.*

The work of Hoff [29] proves that for genuinely non-linear hyperbolic equations, the invariant set $\mathcal{B}$ is convex; this is a powerful property in relation to the analysis of stability. Through the definition of the invariant set, the invariant domain can then be defined.

Definition II.2 (Invariant domain)**.** *For a set of states* $\mathbf{U} = \{\mathbf{u}_{i_1}, \mathbf{u}_{i_2}, \dots\} \subset \mathcal{B}$ *and an operator* $\mathcal{R} : \mathbf{U} \to \mathbf{u}$*,* $\mathcal{B}$ *is said to be an invariant domain for* $\mathcal{R}$ *if* $\mathcal{R}(\mathbf{U}) = \mathbf{u} \in \mathcal{B}$*.*

This definition may be interpreted in terms of a temporal update for some numerical scheme $\mathcal{R}$. If the temporal update of the solution at some point can be written as a function of the states $\mathbf{U}$ and $\mathbf{U} \subset \mathcal{B}$, then the solution at the next time step is guaranteed to be in $\mathcal{B}$ if $\mathcal{R}$ is an invariant domain preserving method. In several works on shock capturing, positivity preservation has been considered [31], and the differences between positivity preservation and invariant domain preservation are shown in the following example.

Example II.1. *The Euler equations can written in the form of Eq.* (1) *as*

$$\mathbf{u} = \begin{bmatrix} \rho \\ \rho\mathbf{V} \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho\mathbf{V} \\ \rho\mathbf{V} \otimes \mathbf{V} + p\mathbf{I} \\ (E + p)\mathbf{V} \end{bmatrix}, \tag{7}$$

where ρ *is the density,* $\mathbf{V}$ *is the velocity vector,* E *is the total energy,* $p = (\gamma - 1)\left(E - \frac{1}{2}\rho\|\mathbf{V}\|_2^2\right)$ *is the pressure, and* γ *is the ratio of specific heats. The typical approach taken for proving positivity preservation is to show that for a sufficiently small* Δt*, the strictly positive functionals* ρ *and* p *remain positive, i.e. for some scheme* $\mathcal{P}(\mathbf{U})$*, if*

$$\mathcal{P}(\mathbf{U}) = \mathbf{u}, \quad \textit{then} \quad p(\mathbf{u}) > 0 \quad \textit{and} \quad \rho(\mathbf{u}) > 0. \tag{8}$$

This can be augmented, as in the case of Hu et al. [31], by stating that if

$$\mathcal{P}(\mathbf{U}) = \mathbf{u}, \quad \textit{then} \quad p(\mathbf{u}) > p_{\min} \quad \textit{and} \quad \rho(\mathbf{u}) > \rho_{\min}, \tag{9}$$

where $p_{\min}$ *and* $\rho_{\min}$ *are the minimum density and pressure from the initial conditions. However, this does not account for the entropy, and instead, it is more physical to expect the solution to be in the invariant set*

$$\mathcal{B} = \{\mathbf{u} \in \mathcal{A} \mid \rho \geqslant 0, p(\mathbf{u}) \geqslant 0, \sigma(\mathbf{u}) \geqslant \sigma_0\} \tag{10}$$

for some entropy functional, σ*. This is a different statement to just considering the positivity of density and pressure and has a considerable impact on the physicality of the solutions.*

Furthermore, if considering the approximations given in Eq. (4)*, then given Corollary II.1.1, entropy stable solutions to the Riemann problem must lie within a small convex subset of* $\mathcal{B}$*, the size of which is dependent on* $\lambda_{\max}$*. This will form the basis the flux limiting approach taken in this work.*

A. Invariant Domain Preserving Schemes

In the work of Guermond and Popov [23], the divergence of the flux at a point i was formed through a summation over the points in the set $\mathcal{I}(i)$. This set, shown in Fig. 1a, can be thought of as the numerical domain of influence for a given point. The individual contribution from each of the points $\mathbf{x}_j \in \mathcal{I}(i)$ to the total divergence of the flux at $\mathbf{x}_i$ can be represented as the product of the flux $\mathbf{F}_j$ and the differentiation coefficients $\mathbf{c}_{ij}$.

$$\nabla \cdot \mathbf{F}_i \approx \sum_{j \in \mathcal{I}(i)} \mathbf{F}_j \cdot \mathbf{c}_{ij}. \tag{11}$$

The advantage of considering a numerical scheme in this manner is that the abstraction allows for some properties to proven for a broad array of methods that can be expressed in this framework. An example of the formation of the $\mathbf{c}_{ij}$ coefficients is shown for a finite difference scheme.

Example II.2. *Consider a fourth order central difference scheme in one dimension with uniform grid spacing h, a single conserved variable u, and the associated flux function $f(u)$. Let $u_i = u(x = x_i)$ and $f_i = f(u_i)$ where $x_i = ih$. Then, the gradient of the flux at x_i can be expressed as*

$$\frac{\mathrm{d}f_i}{\mathrm{d}x} \approx \frac{1}{12h}\left(f_{i-2} - 8f_{i-1} + 8f_{i+1} - f_{i+2}\right).$$

Transforming this scheme into the notation of Eq. (11)*, the numerical domain of influence is given by $\mathcal{I}(i) = \{i-2, i-1, i, i+1, i+2\}$, and the entries of $\mathbf{c}_{ij}$ take on values*

$$\mathbf{c}_{i,i-2} = \frac{1}{12h}, \quad \mathbf{c}_{i,i-1} = -\frac{8}{12h}, \quad \mathbf{c}_{i,i+1} = \frac{8}{12h}, \quad \mathbf{c}_{i,i+2} = -\frac{1}{12h}, \quad \text{and} \quad \mathbf{c}_{i,i} = 0.$$

Hence, the scheme may be written as

$$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=x_i} \approx \sum_{j \in \mathcal{I}(i)} f_j \mathbf{c}_{ij}.$$

This abstract presentation of a numerical method was then stabilised by including a graph viscosity term such that the resulting scheme was invariant domain preserving under strong-stability preserving temporal integration. A scheme augmented with a graph viscosity may be expressed as

$$\nabla \cdot \mathbf{F}_i = \sum_{j \in \mathcal{I}(i)} \mathbf{F}_j \cdot \mathbf{c}_{ij} - \sum_{j \in \mathcal{I}(i)} d_{ij}(\mathbf{u}_j - \mathbf{u}_i), \tag{12}$$

subject to the following conditions on the differentiation coefficients $\mathbf{c}$ and graph viscosity coefficients d,

$$d_{ii} = -\sum_{j \in \mathcal{I}(i) \setminus i} d_{ij} \quad \therefore \quad \sum_{j \in \mathcal{I}(i)} d_{ij} = 0, \tag{13}$$

and

$$\sum_{j \in \mathcal{I}(i)} \mathbf{c}_{ij} = 0. \tag{14}$$

Utilising Eq. (14), an equivalent form of Eq. (12) may be written as

$$\nabla \cdot \mathbf{F}_i \approx \sum_{j \in \mathcal{I}(i)} (\mathbf{F}_j + \mathbf{F}_i) \cdot \mathbf{c}_{ij} - \sum_{j \in \mathcal{I}(i)} d_{ij}(\mathbf{u}_j - \mathbf{u}_i) \tag{15}$$

which makes the structure of the Riemann problem clearer. The term d_{ij} is then set as

$$d_{ij} = \max\left\{\lambda_{\max}(\mathbf{n}_{ij}, \mathbf{u}_i, \mathbf{u}_j)\|\mathbf{c}_{ij}\|_2, \lambda_{\max}(\mathbf{n}_{ji}, \mathbf{u}_j, \mathbf{u}_i)\|\mathbf{c}_{ji}\|_2\right\}. \tag{16}$$

With this formulation of the graph viscosity, several properties of the scheme can be established [23, 25], particularly the property that the scheme preserves every convex invariant of the system, leading to a strong form of stability. For a detailed presentation and proof of the properties of invariant domain preserving schemes, the reader is invited to read the references.

III. Flux Reconstruction

To present the FR methodology for approximating the solution of hyperbolic PDEs, we consider a generic PDE in one-dimension, written as

$$\partial_t u + \partial_x f = 0. \tag{17}$$

The FR method makes use of a reference domain $\hat{x} \in \hat{\Omega} = [-1, 1]$ coupled to a transformation $T_k : \hat{\Omega} \rightarrow \Omega_k$ where Ω_k is a sub-domain. A nodal approximation is formed such that the reference domain shape functions $\hat{l}_i(\hat{x})$ are the Lagrange interpolating polynomials for the set of $p + 1$ unique nodes $\{\hat{x}_0, ..., \hat{x}_p\}$. With these shape functions and their corresponding nodes, the interpolation operators I_l and I_r are defined such that $I_l v = v_l = v(-1)$ and $I_r v = v_r = v(1)$ where $v \in \mathbb{P}_p$. The approximation of the solution u at $x_j \in \Omega_k$ for Eq. (17) is then given by

$$u(x_j) = u_j \approx \sum_{i=0}^{p} u(\hat{x}_i) l_i(T_k^{-1} x_j), \tag{18}$$

and the FR approximation of the spatial derivative of the flux f at $x_j \in \Omega_k$ is given by

$$\frac{\partial f(x_j)}{\partial x} \approx \left(\left. \frac{\mathrm{d}T_k}{\mathrm{d}\hat{x}} \right|_{\hat{x}=T_k^{-1}x_j} \right)^{-1} \left[\sum_{i=0}^{p} f(u_i) \frac{\mathrm{d}l_i}{\mathrm{d}\hat{x}} + (f_l^I - I_l f) \frac{\mathrm{d}h_l}{\mathrm{d}\hat{x}} + (f_r^I - I_r f) \frac{\mathrm{d}h_r}{\mathrm{d}\hat{x}} \right]_{\hat{x}=T_k^{-1}x_j} \tag{19}$$

The spatial Jacobian, simplified for the 1D case here, is just

$$\mathbf{J}(\mathbf{x}_i) = \left. \frac{\mathrm{d}T_k}{\mathrm{d}\hat{x}} \right|_{\hat{x}=T_k^{-1}x_j} = \left[\partial_\xi x \right].$$

The FR method as presented can be readily extended to tensor product elements, as demonstrated by Huynh [12], and may also be extended to simplex elements and prisms [32]. With this approximation of the spatial components of the governing equations, ODE methods may be used to integrate the solution in time in a method of lines fashion. Although there has been some success in using FR in combination with implicit time stepping methods [33], we will restrict ourselves to explicit methods in this work.

A. Temporal Integration

For the shock capturing method to be presented, several properties of the scheme such as invariant domain preservation are reliant on the strong stability of the temporal integration, e.g. forward Euler. To maintain these properties and achieve a high temporal order of accuracy, strong-stability-preserving explicit Runge–Kutta (SSP-ERK) schemes are employed [34]. This family of methods preserves the strong stability of the forward Euler method and allows for the straightforward extension of the properties derived for the forward Euler scheme to a more practical class of explicit RK schemes. In this work, the 3rd order SSP-ERK scheme is used, whose coefficients may be written in a Butcher tableau as

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array} = \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1/2 & 1/4 & 1/4 & 0 \\ \hline & 1/6 & 1/6 & 1/3 \end{array} \tag{20}$$

For convenience, we also include the CFL limits [35] in Table 1 for the FR scheme with DG correction functions and upwinded interfaces as calculated by the method of Vincent et al. [36] and Ketcheson and Ahmadia [37].

Table 1 Upwind FR-DG SSP-ERK3 CFL limits.

p	$\mathrm{CFL}_{\max}$	p	$\mathrm{CFL}_{\max}$
1	0.40959	5	0.06611
2	0.20975	6	0.05102
3	0.13010	7	0.04072
4	0.08968	8	0.03336

IV. Invariant Domain Preserving Shock Capturing

In this section, the invariant domain preserving shock capturing method is presented in terms of a high-order flux reconstruction scheme and a low-order summation-by-parts scheme followed by a convex limiting procedure between the two.

A. High-order Scheme

Utilising the description of FR in Section III, the high-order method can be cast in the abstract form of Eq. (12) to make it compatible with the shock capturing method. The set of solution points in the reference domain, $\hat{\mathbf{x}}_i \forall i \in \mathcal{I}(\mathcal{K})$, is not explicitly constrained to be the Gauss–Legendre–Lobatto points to offer a more general description. The FR approximation of the gradient of the flux in the reference domain may then be written as

$$\nabla \cdot \mathbf{F}_i = \sum_{j \in \mathcal{I}(\mathcal{K})} \mathbf{F}_j \cdot \mathbf{c}_{ij}^{\mathcal{K}} + \sum_{j \in \mathcal{I}(\mathcal{K}^E)} \mathbf{F}_j \cdot \mathbf{c}_{ij}^{\delta} - \sum_{j \in \mathcal{I}(\mathcal{K}^I)} \mathbf{F}_j \cdot \mathbf{c}_{ij}^{\delta} - \sum_{j \in \mathcal{I}(\mathcal{K}^E)} \alpha u_j \mathbf{e} \cdot \mathbf{c}_{ij}^{\delta} + \sum_{j \in \mathcal{I}(\mathcal{K}^I)} \alpha u_j \mathbf{e} \cdot \mathbf{c}_{ij}^{\delta} \quad \forall i \in \mathcal{I}(\mathcal{K}). \tag{21}$$

Here, $\mathbf{e}$ is a vector of ones and α is a wave speed term defined such that a common interface flux as typically calculated by an approximate Riemann solver is recovered on the interface flux points. The differentiation coefficients are then defined as

$$\mathbf{c}_{ij}^{\mathcal{K}} = \nabla \phi_j(\mathbf{x}_i), \tag{22a}$$

$$\mathbf{c}_{ij}^{\delta} = \frac{1}{2} \begin{bmatrix} n_{1j} \partial_{x_1} h_j(x_i) \\ n_{2j} \partial_{x_2} h_j(x_i) \\ \vdots \\ n_{2j} \partial_{x_d} h_j(x_i) \end{bmatrix}, \tag{22b}$$

where ϕ_j is the jth basis function and h_j is the correction function for (flux) point j. By utilising the following simplification

$$\mathbf{c}_{ij} = \begin{cases} \mathbf{c}_{ij}^{\mathcal{K}}, & \text{if } j \in \mathcal{I}(\mathcal{K}) \\ -\mathbf{c}_{ij}^{\delta}, & \text{if } j \in \mathcal{I}(\mathcal{K}^I) \\ \mathbf{c}_{ij}^{\delta}, & \text{if } j \in \mathcal{I}(\mathcal{K}^E), \end{cases} \tag{23}$$

for the sets shown in Fig. 1a, a compact form of Eq. (21) can be written as

$$\nabla \cdot \mathbf{F}_i = \sum_{j \in \mathcal{I}(i)} \mathbf{F}_j \cdot \mathbf{c}_{ij} - \sum_{j \in \mathcal{I}(\mathcal{K}^E)} \alpha u_j \mathbf{e} \cdot \mathbf{c}_{ij}^{\delta} + \sum_{j \in \mathcal{I}(\mathcal{K}^I)} \alpha u_j \mathbf{e} \cdot \mathbf{c}_{ij}^{\delta} \quad \forall i \in \mathcal{I}(\mathcal{K}). \tag{24}$$

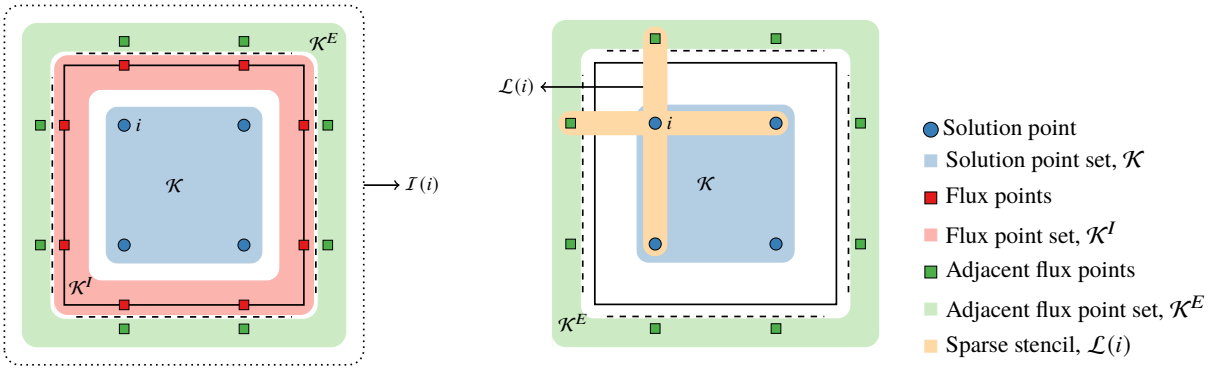
Assuming a Lagrange basis for FR, i.e. $l_i(\hat{\mathbf{x}}_j) = \delta_{ij}$,

$$\hat{\mathbf{c}}_{ij}^{\mathcal{K}} = \hat{\nabla} l_j(\hat{\mathbf{x}}_i),$$

which implies

$$\mathbf{c}_{ij}^{\mathcal{K}} = \mathbf{J}_i^{-1} \hat{\mathbf{c}}_{ij}^{\mathcal{K}}.$$

Therefore, in a typical FR computational setting, the terms $\hat{\mathbf{c}}_{ij}^{\mathcal{K}}$ can be taken directly from the terms in the flux divergence matrix.



(a) Full graph viscosity. (b) Sparse graph viscosity.

Fig. 1 Example of point sets for $p = 1$ FR on a quadrilateral.

Remark. *As can be seen from Fig. 1a, the number of points in the stencil $\mathcal{I}(i)$ grows with order with a varying rate depending on the geometry and basis used. Consequently, the dissipation increases with the spatial degree which coincides with our understanding of polynomial interpolation and the growth of Runge phenomena. As a result, the full graph viscosity approach becomes a remarkably dissipative and computationally expensive method of forming a low-order approximation even at moderate orders.*

B. Low-order Scheme

To combat the excessive dissipation and complexity of the graph viscosity method for forming a low-order approximation, Pazner [38] proposed a low-order scheme for DG such that only the immediately adjacent points were required, significantly reducing the extent of the stencil. To define a low-order FR method, we utilise the summation-by-parts (SBP) framework. This framework allows for simple test criteria for conservation, convergence, and linear stability [39] while also allowing for the abstraction of the low-order scheme and the use of existing FR solvers with only the redefinition of the operator matrices. The SBP framework may be summarised in the following definition.

Definition IV.1 (Summation-by-parts)**.** *The set of operators* $\mathbf{M}, \mathbf{D}, \mathbf{L}$*, and* $\mathbf{B}$ *is said to be summation-by-parts operators if*

$$\langle \hat{\mathbf{v}}, \hat{\mathbf{u}} \rangle_M \equiv \hat{\mathbf{v}}^T \mathbf{M} \hat{\mathbf{u}} \approx \int_{\hat{\Omega}} \hat{v}\hat{u} \mathrm{d}x, \tag{25a}$$

$$\mathbf{D}\hat{\mathbf{u}} \approx \left. \frac{\partial \hat{u}}{\partial x} \right|_{x=\hat{\mathbf{x}}}, \tag{25b}$$

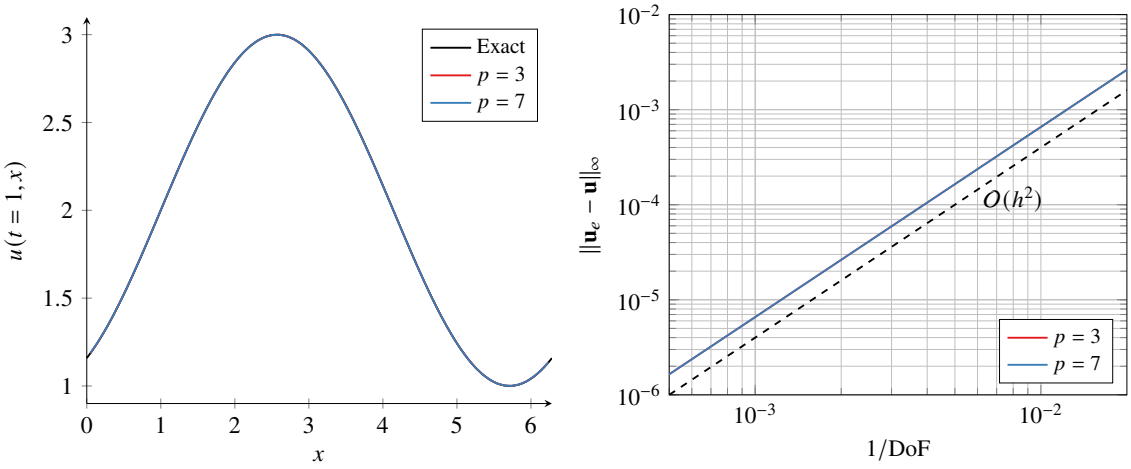
$$\hat{\mathbf{u}}^T \mathbf{M}\mathbf{D}\hat{\mathbf{v}} + \hat{\mathbf{u}}^T \mathbf{D}^T \mathbf{M}\hat{\mathbf{v}} \approx \int_{\hat{\Omega}} u\partial_x v \mathrm{d}x + \int_{\hat{\Omega}} v\partial_x u \mathrm{d}x = uv\big|_{\partial\hat{\Omega}}, \tag{25c}$$

$$\mathbf{M}\mathbf{D} + \mathbf{D}^T \mathbf{M} = \mathbf{L}^T \mathbf{B} \mathbf{L}, \tag{25d}$$

where $\mathbf{B} = \mathrm{diag}(-1, 1)$ *and* $\mathbf{L}$ *is the interpolation to* $\partial\Omega$*, i.e.* $\mathbf{L}\hat{\mathbf{u}} \approx [\hat{\mathbf{u}}_L, \hat{\mathbf{u}}_R]^T$.

This definition makes it clear that summation-by-parts is a discrete analogue to integration-by-parts. Furthermore, a scheme that yields SBP operators is useful as the rich literature of SBP can be utilised to prove properties such as stability. We now present a low-order method which will be used throughout this work. The mass matrix and differentiation matrix are given by

$$\overline{\mathbf{M}} = \mathrm{diag}(\mathbf{w}), \quad \overline{\mathbf{D}} = \frac{\overline{\mathbf{M}}^{-1}}{2} \begin{bmatrix} -1 & 1 & 0 & & \\ -1 & 0 & 1 & \ddots & \\ 0 & \ddots & \ddots & \ddots & 0 \\ & \ddots & -1 & 0 & 1 \\ & & 0 & -1 & 1 \end{bmatrix}, \tag{26}$$



(a) DoF = 64, $t = 1$

(b) L_∞ norm of error.

Fig. 2 Low-order FR-SBP applied to linear advection $u_0 = 2 + \sin x$.

where $\mathbf{w}$ are the quadrature weights of the solution points and over-lining is used to differentiate these from their high-order counterparts. The projection and correction matrices are given by

$$\overline{\mathbf{L}} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}, \quad (\overline{\mathbf{C}})^T = \overline{\mathbf{M}}^{-1} \begin{bmatrix} -1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}. \tag{27}$$

Theorem IV.1 (Low-order FR-SBP)**.** *For a one dimensional affine reference element $\hat{\mathcal{K}}$ on the domain $\hat{\Omega}$, if the spatial derivative is approximated using the FR method as*

$$-\partial_x \hat{\mathbf{f}} \approx -\mathbf{D}\mathbf{f} - \mathbf{C}(\mathbf{f}^I - \mathbf{L}\mathbf{f}). \tag{28}$$

then the definitions of the operators in Eqs. (26) *and* (27) *construct an FR-SBP scheme that is stable and conservative and therefore consistent.*

Proof. By substitution, it can be shown that these definitions satisfy the SBP property of Eq. (25). From Ranocha et al. [39], it is sufficient to show that $\mathbf{1}^T \overline{\mathbf{M}}\overline{\mathbf{C}} = \mathbf{1}^T \overline{\mathbf{L}}^T \mathbf{B}$ for conservation and $\overline{\mathbf{C}} = \overline{\mathbf{M}}^{-1} \overline{\mathbf{L}}^T \mathbf{B}$ for linear stability. Straightforward substitution shows that both of these conditions are met. □

In Fig. 2, the FR-SBP scheme was applied to the linear advection equation using SSP-ERK3 temporal integration, a constant $\Delta t = 1 \times 10^{-4}$, and centred differencing at element interfaces. A second-order rate of convergence was observed, although it is not strictly necessary for the FR-SBP scheme to be low-order as it is used in conjunction with a graph viscosity. The advantage of this method is that the size of the stencil can be modified to incorporate various schemes from subcell methods to high-order FR with full graph viscosity, and the SBP approach would permit straightforward verification of its numerical properties.

This low-order scheme is linearly stable and conservative, but to guarantee stability for non-linear hyperbolic problems, it is used in conjunction with a graph viscosity. Due to the increased sparsity of the low-order scheme in comparison to the full high-order scheme, this approach is termed sparse graph viscosity. It can be shown that in this particular case, the method reduces to a subcell approach that uses a Riemann solver at the subcell interfaces, but again, this method attempts to abstract the approach to a wider class of reduced order approximations.

Using the graph viscosity presented in Sections II.A and IV.A, the new operators of Theorem IV.1 lead to the altered definitions of the sparse space function matrices as

$$\hat{\tilde{\mathbf{c}}}^{\mathcal{K}} = \frac{\overline{\mathbf{M}}^{-1}}{2} \begin{bmatrix} -1 & 1 & & \\ -1 & 0 & 1 & \\ & \ddots & \ddots & \ddots \end{bmatrix}, \quad \hat{\tilde{\mathbf{c}}}^{\mathcal{K}^I} = \frac{\overline{\mathbf{M}}^{-1}}{2} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad \hat{\tilde{\mathbf{c}}}^{\mathcal{K}^E} = \frac{\overline{\mathbf{M}}^{-1}}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

which can be combined to give

$$\hat{\overline{\mathbf{c}}} = \frac{\overline{\mathbf{M}}^{-1}}{2}\begin{bmatrix} -1 & 0 & 1 & 0 & & & \\ 0 & -1 & 0 & 1 & \ddots & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & -1 & 0 & 1 & 0 \\ & & & 0 & -1 & 0 & 1 \end{bmatrix} \tag{29}$$

where the first and last columns are the left and right exterior flux points, respectively. The hats on the terms indicate that these are in the reference domain.

The low-order temporal update equation with graph viscosity can then be written as

$$\mathbf{u}_i^{L,n+1} = \mathbf{u}_i^n - \Delta t \sum_{j\in\mathcal{L}(i)} \mathbf{F}_j \cdot \overline{\mathbf{c}}_{ij} - \overline{d}_{ij}(\mathbf{u}_j - \mathbf{u}_i), \tag{30}$$

where the superscript L on the LHS indicates the updated value is from the low-order method. We have also introduced the new set, $\mathcal{L}(i)$, which is the modified stencil for this low-order component–this can be understood diagrammatically from Fig. 1b. The low-order graph viscosity coefficients $\overline{d}_{ij}$ are defined with $\overline{\mathbf{n}}_{ij} = \overline{\mathbf{c}}_{ij}/|\overline{\mathbf{c}}_{ij}|$, which simplifies to

$$\overline{d}_{ij} = \max\left(\lambda_{\max}(\overline{\mathbf{n}}_{ij}, \mathbf{u}_i, \mathbf{u}_j)|\overline{\mathbf{c}}_{ij}|,\ \lambda_{\max}(\overline{\mathbf{n}}_{ji}, \mathbf{u}_i, \mathbf{u}_j)|\overline{\mathbf{c}}_{ji}|\right), \tag{31}$$

where the final term accounts for non-linearity in the spatial Jacobian. A key observation is that the definitions of $\overline{\mathbf{c}}_{ij}$ and $\overline{d}_{ij}$ remain consistent with the conditions presented in Section II.A. As a result, the following properties can be shown.

Theorem IV.2 (Local Invariance)**.** *For some sub-domain Ω_k, let $\mathbf{u}_i^{L,n} = \mathbf{u}_{i,k}^{L,n} \in \mathcal{B}\ \forall\ i \in \{1, ..., p+1\}$ for the set $\mathcal{B}$ defined in Definition II.1. For the scheme defined by Eq.* (30) *and d_{ij} defined by Eq.* (31)*, there exists a strictly positive Δt such that $\mathbf{u}_i^{n+1} \in \mathcal{B}$.*

Proof. This proof follows closely that of Guermond et al. [25] Thm. 3.6. Utilising $\sum_{j\in\mathcal{L}(i)} \mathbf{c}_{ij} = 0$ and $\sum_{j\in\mathcal{L}(i)} d_{ij} = 0$, Eq. (30) can be rewritten as

$$\mathbf{u}_i^{L,n+1} = \mathbf{u}_i^{L,n} - \Delta t \sum_{j\in\mathcal{L}(i)} (\mathbf{F}_j^{L,n} - \mathbf{F}_i^{L,n}) \cdot \overline{\mathbf{c}}_{ij} - \overline{d}_{ij}(\mathbf{u}_j^{L,n} - \mathbf{u}_i^{L,n}), \tag{32}$$

which may be rearranged to

$$\mathbf{u}_i^{L,n+1} = \mathbf{u}_i^{L,n} - \Delta t \sum_{j\in\mathcal{L}(i)\setminus\{i\}} 2\overline{d}_{ij}\mathbf{u}_i^{L,n} + (\mathbf{F}_j^{L,n} - \mathbf{F}_i^{L,n}) \cdot \overline{\mathbf{c}}_{ij} - \overline{d}_{ij}(\mathbf{u}_j^{L,n} + \mathbf{u}_i^{L,n}). \tag{33}$$

The auxiliary states are then defined as

$$\tilde{\mathbf{u}}_{ij} = \frac{1}{2}(\mathbf{u}_j + \mathbf{u}_i) - \frac{1}{2\overline{d}_{ij}}(\mathbf{F}_j - \mathbf{F}_i) \cdot \overline{\mathbf{c}}_{ij}, \tag{34}$$

which, from the definition of $\overline{d}_{ij}$ in Eq. (31) and the preliminaries on invariant sets, means that $\tilde{\mathbf{u}}_{ij} \in \mathcal{B}$. Applying this definition to Eq. (33) gives

$$\mathbf{u}_i^{L,n+1} = \left[1 - 2\Delta t \sum_{j\in\mathcal{L}(i)\setminus\{i\}} \overline{d}_{ij}\right]\mathbf{u}_i^{L,n} - 2\Delta t \sum_{j\in\mathcal{L}(i)\setminus\{i\}} \overline{d}_{ij}\tilde{\mathbf{u}}_{ij}^n. \tag{35}$$

Given that $\mathbf{u}_i^{L,n}, \tilde{\mathbf{u}}_{ij}^n \in \mathcal{B}$ and $d_{ij} > 0\ \forall j \in \mathcal{L}(i) \setminus \{i\}$, then for a sufficiently small Δt, $\mathbf{u}_i^{L,n+1}$ is a convex combination of states in the invariant set. As the invariant set is convex, $\mathbf{u}_i^{L,n+1} \in \mathcal{B}$. □

As was discussed in the preliminaries, a scheme with this property is imbued with an absolute form of stability that guarantees the solution is physical given the conditions imposed on Δt in the following corollary.

Corollary IV.2.1 (CFL Limit)**.** *Given a solution $\mathbf{u}^n \in \mathcal{B}$, a sufficient condition for the local stability of the numerical scheme defined in Eq.* (30) *is that*

$$\Delta t < -\frac{1}{2\overline{d}_{ii}},$$

where $\overline{d}_{ij} \geqslant 0$ if $j \neq i$ and $\overline{d}_{ii} = -\sum_{j\in\mathcal{L}(i)\setminus i} \overline{d}_{ij}$.

The proof of this property is substantially similar to that presented by Guermond et al. [25] and is based on requiring $\mathbf{u}_i^{n+1}$ to be a convex combination of states in the invariant set. Although this forms a sufficient condition, it is not a necessary condition for the update to remain in the invariant set as discussed by Dzanic et al. [40].

Remark. *Consideration of the local invariance property and the CFL limit leads us to a global invariance property, i.e. if a global Δt is sufficiently small such that for all points the CFL limit is met, then the scheme will define a global invariant domain. Furthermore, by extension of the invariant domain preserving property, the scheme is entropy stable as shown by Guermond et al. [25].*

C. Convex Limiting

As the SBP methodology provides a robust manner in which to obtain a low-order approximation of the solution in the setting of a high-order spectral element scheme, it can be easily paired with the high-order scheme. The solution predicted by the low-order method is guaranteed to remain in the invariant set while the prediction by the high-order method can be more accurate but may not remain in the set. To alleviate this problem and achieve solutions which converge rapidly onto the true solution, a limiting technique is applied. We apply a similar approach as Guermond et al. [25], where a convex minimisation problem is solved to give the least graph-viscosity necessary. The adaption we provide is similar to the method of flux-corrected transport [41] where the limiting is performed only locally and not across the complete graph used to form the artificial viscosity. The motivation behind this is that the full approach has very high memory requirements and results in a significant increase in run time. This simplification is justified as it still provides a solution within the invariant domain, and due to the limited graph used in the low-order method, the computational cost is greatly reduced.

The temporal updated of the solution from the high-order and low-order methods may be written as

$$\mathbf{u}_i^{H,n+1} = \mathbf{u}_i^n - \Delta t \sum_{j\in\mathcal{I}(i)} \mathbf{F}_j \cdot \mathbf{c}_{ij},$$
$$\mathbf{u}_i^{L,n+1} = \mathbf{u}_i^n - \Delta t \sum_{j\in\mathcal{L}(i)} \mathbf{F}_j \cdot \overline{\mathbf{c}}_{ij} - \overline{d}_{ij}(\mathbf{u}_j - \mathbf{u}_i)$$

where the superscripts H and L indicate the high-order and low-order updates, respectively. A convex combination of these solutions can be written as

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^{L,n+1} + \alpha \mathbf{P}_i, \quad \text{for} \quad \alpha \in [0,1], \tag{37}$$

where

$$\mathbf{P}_i = \Delta t \left[\sum_{j\in\mathcal{L}(i)} \left(\mathbf{F}_j \cdot \overline{\mathbf{c}}_{ij} - \overline{d}_{ij}(\mathbf{u}_j - \mathbf{u}_i) \right) - \sum_{j\in\mathcal{I}(i)} \mathbf{F}_j \cdot \mathbf{c}_{ij} \right]. \tag{38}$$

The limiting coefficient α is then found such that

$$\rho_i^{n+1} \geqslant \rho_{\min,i}, \quad \text{and} \quad \rho_i^{n+1} \leqslant \rho_{\max,i}, \tag{39}$$

where

$$\rho_{\min,i} = \min_{j\in\mathcal{L}(i)\setminus i} \left[\rho(\mathbf{u}_i), \rho(\tilde{\mathbf{u}}_{ij}) \right], \tag{40a}$$
$$\rho_{\max,i} = \max_{j\in\mathcal{L}(i)\setminus i} \left[\rho(\mathbf{u}_i), \rho(\tilde{\mathbf{u}}_{ij}) \right]. \tag{40b}$$

These conditions impose bounds on the density based on the auxiliary states, but to define the invariant set, we also impose the bound on entropy such that

$$\psi(\mathbf{u}_i^{n+1}, \phi_{\min,i}) \geqslant 0. \tag{41}$$

where

$$\psi(\mathbf{u}, \phi_{\min}) = \rho^{\gamma+1}(\phi - \phi_{\min}), \tag{42a}$$

$$\phi = e\rho^{1-\gamma} \tag{42b}$$

$$\phi_{\min,i} = \min_{j \in \mathcal{L}(i)\setminus i} \left[\phi(\mathbf{u}_i), \phi(\tilde{\mathbf{u}}_{ij})\right]. \tag{42c}$$

These conditions are sufficient to define a convex invariant set with each functional itself being convex. Due to the convexity of these functions, if $\alpha = 1$ does not satisfy all three constraints, then a single root that satisfies these conditions is guaranteed to exist for $\alpha \in [0, 1)$. To find this root, we use a modification of the method presented by Maier and Kronbichler [42] where instead of using a quadratic-Newton type approach, we use a bisection method to avoid the numerous divisions that can be computationally expensive. After the final bisection step, the precision of the root is increased by performing a linear interpolation step followed by a check to ensure that the bounds are being enforced. Due to the proximity of the bounds by this point, this approximation was found to be highly accurate and reduced the noise in the solution. Furthermore, these improvements were seldom rejected for violating the limiting conditions.

Algorithm 1: Bisection Convex Limiting$(\mathbf{u}^{n+1,L}, \mathbf{P}, \rho_{\min}, \rho_{\max}, \phi_{\min}, n_{\max})$.

Result: α

$\alpha_l := 0$

$$\alpha_{rm} := \begin{cases} 1 & \text{if} \quad \rho(\mathbf{u}^{n+1,L} + \mathbf{P}) \geqslant \rho_{\min} \\ |\rho_{\min} - \rho(\mathbf{u}^{n+1,L})|/|\rho(\mathbf{P})| & \text{otherwise} \end{cases}$$

$$\alpha_{rM} := \begin{cases} 1 & \text{if} \quad \rho(\mathbf{u}^{n+1,L} + \mathbf{P}) \leqslant \rho_{\max} \\ |\rho_{\max} - \rho(\mathbf{u}^{n+1,L})|/|\rho(\mathbf{P})| & \text{otherwise} \end{cases}$$

$\alpha_r := \min\left[\alpha_{rm}, \alpha_{rM}\right]$

$\Psi_r := \psi(\mathbf{u}^{n+1,L} + \alpha_r \mathbf{P}, \phi_{\min})$

if $\Psi_r \geqslant 0$ **then return** $\alpha := \alpha_r$

for $n = 1$ **to** $n_{\max}$ **do**

  $\alpha_c := 0.5(\alpha_r + \alpha_l)$

  $\Psi_c := \psi(\mathbf{u}^{n+1,L} + \alpha_c \mathbf{P}, \phi_{\min})$

  $\alpha_r := \min\left[\alpha_r - 1, (1 - \alpha_c)\text{sign}(\Psi_c)\right] + 1$

  $\alpha_l := \max\left[\alpha_l, \alpha_c \text{sign}(\Psi_c)\right]$

end

$\Psi_r := \psi(\mathbf{u}^{n+1,L} + \alpha_r \mathbf{P}, \phi_{\min})$

$\Psi_l := \psi(\mathbf{u}^{n+1,L} + \alpha_l \mathbf{P}, \phi_{\min})$

$\alpha := (\alpha_l|\Psi_l| + \alpha_r|\Psi_r|)/(|\Psi_l| + |\Psi_l|)$

$\Psi := \psi(\mathbf{u}^{n+1,L} + \alpha \mathbf{P}, \phi_{\min})$

if $\Psi \leqslant 0$ **then** $\alpha := \alpha_l$

return α

The details of the bisection method used are presented in Algorithm 1. The computational efficiency can be improved when implemented by making use of conditional move statements and min/max as well as setting $n_{\max}$ a priori such that the loop can be fully unrolled. As $\alpha \in [0, 1]$, $n_{\max}$ will give α with precision of $2^{n_{\max}}$ before the final linear interpolation step. $n_{\max} = 7$ was found to be sufficient. Furthermore, there are several powers involved in the calculation of the entropy and convex entropy functions ϕ and ψ. These can be completely avoided by making use of the standard exponent-log approach. The advantage of this approach is that both of these operations can be vectorised, unlike powers, and are less prone to the well-known issues in error control when computing powers which can lead to excessive run times. An example of this alternative form in one dimension is given by

$$\psi = \rho^{\gamma+1}(\phi - \phi_{\min}) = \left(\rho E - \frac{1}{2}m^2\right) - \phi_{\min} \exp\left((\gamma + 1)\log\rho\right). \tag{43}$$

As is highlighted in Eqs. (37) and (38), the limiting is performed locally by only varying a single parameter α. This is in contrast to the approach of Guermond et al. [25] where the contribution to the update at point i from each point in $\mathcal{I}(i)$ was limited. This has a large computational overhead as not only does the actual limiting have to be performed, but also a large amount of data for each point has to be stored. The approach used here uses the points in the stencil $\mathcal{S}(i)$ and $\overline{d}_{ij}$ to form the extrema of the convex functional, but this can be done locally while only storing $\overline{d}_{ij}$ in registers. Pazner [38] seems to indicate that such an approach could lead to the method being excessively dissipative. However, as we will show later, this was not our experience, with a larger factor being which convex functionals are used in the limiting process.

As discussed by Guermond et al. [24, 25] and first investigated by Khobalatte and Perthame [43], when applying strict bounds on specific entropy, it is postulated that one cannot recover a second-order or higher rate of convergence in the L^2 and L^∞ norms. Instead, the entropy bound has to be relaxed, such as by applying some Laplacian smoothing to the field of entropy bounds. In this work, this step is not performed, but some other works have had success with this relaxation technique.

D. Entropy Viscosity and Shock Sensor

To further improve upon the shock capturing approach, two additional methods are used. The first is the entropy viscosity method of Guermond et al. [44], where an entropy residual is used to define a viscosity. The philosophy of this approach is that through applying a physically proportionate amount of viscosity, the high-order solution can be stabilised without causing excessive dissipation of the solution. For the method proposed here, as the complete scheme is not reliant on this as the only means of producing a physical solution, exhaustive tuning of the parameters is not necessary as long as they are sufficiently small.

The entropy viscosity method relies on the calculation of the R_σ, the point-wise entropy residual, as

$$\partial_t \sigma + \nabla \cdot \mathbf{\Sigma} = R_\sigma, \tag{44}$$

where $(\sigma, \mathbf{\Sigma})$ is an entropy-flux pair. From Lax [45], it is known that $R_\sigma = 0$ almost everywhere, except at shocks where it takes a value less than zero. For the Euler equations, a suitable pair is $\sigma = \rho \log(p\rho^{-\gamma})/(\gamma - 1)$ and $\mathbf{\Sigma} = \sigma \mathbf{v}$. A set of viscosities can then be defined as

$$\mu_{E,\mathcal{K}} = c_E h_\mathcal{K}^2 \|\rho\|_{\infty,\mathcal{K}} \|R_\sigma\|_{\infty,\mathcal{K}}, \tag{45a}$$

$$\mu_{\max,\mathcal{K}} = c_{\max} h_\mathcal{K} \|\rho\|_{\infty,\mathcal{K}} \left\| \|\mathbf{v}\|_2 + \sqrt{\gamma T} \right\|_{\infty,\mathcal{K}}, \tag{45b}$$

where

$$T = \frac{(\gamma - 1)}{\rho} \left(E - \frac{1}{2} \|\mathbf{V}\|_2^2 \right), \tag{46}$$

From this, a physical viscosity and thermal conductivity can be defined as

$$\mu_{\sigma,\mathcal{K}} = \min \left[\mu_{E,\mathcal{K}}, \mu_{\max,\mathcal{K}} \right], \quad \kappa_{\sigma,\mathcal{K}} = \frac{P_r}{\gamma - 1} \mu_{\sigma,\mathcal{K}}. \tag{47}$$

The final equation defines the thermal conductivity needed for the viscous terms in the Navier–Stokes equations with $P_r = 0.71$ being the Prandtl number. In the original work [44], a value of 1 was used with the argument that as this is an artificial term, the exact value is not of great importance. The terms c_E and $c_{\max}$ are free parameters, and the suggested values for spectral element methods, $c_E \in [0.1, 1]$ and $c_{\max} = 10^{-2}/p$, are used.

The entropy viscosity is used in conjunction with the viscous flux of the Navier–Stokes equations which, in three dimensions, takes the form

$$\mathbf{F}^v = \begin{bmatrix} 0 & 0 & 0 \\ \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \\ \mathbf{V}_1\tau_{xx} + \mathbf{V}_2\tau_{xy} + \mathbf{V}_3\tau_{xz} - \mathbf{q}_1 & \mathbf{V}_1\tau_{yx} + \mathbf{V}_2\tau_{yy} + \mathbf{V}_3\tau_{yz} - \mathbf{q}_2 & V_1\tau_{zx} + \mathbf{V}_2\tau_{zy} + \mathbf{V}_3\tau_{zz} - \mathbf{q}_3 \end{bmatrix}, \tag{48}$$

for

$$\boldsymbol{\tau} = \mu_{\sigma,\mathcal{K}} \left(\nabla \mathbf{V} + [\nabla \mathbf{V}]^T - \frac{2}{3} [\nabla \cdot \mathbf{V}] \mathbf{I} \right) \quad \text{and} \quad \mathbf{q} = -\kappa_{\sigma,\mathcal{K}} \nabla T. \tag{49}$$

The bulk viscosity is assumed to be zero, which may not be a valid assumption for high Mach number flows but is deemed acceptable as the entropy viscosity is an artificial term.

To evaluate R_σ, an additional equation is added to the PDE system such that an approximation to $\nabla \cdot \mathbf{\Sigma}$ may be obtained. However, as the residual cannot be integrated in time, the first sub-step and the current sub-step at each ERK stage can be used to form a temporal derivative for R_σ. For the initial sub-step in a simulation, there is insufficient data to form a temporal derivative, so the entropy viscosity is set to zero for this step. In comparison to the original method of Guermond et al. [44], the mass residual term from Eq. (45a) was neglected as it was deemed insignificant for the cases encountered, and in the occasional instance where it did make an impact, convex limiting remedied the problem.

As the entropy viscosity calculations requires the calculation of the entropy residual, which can be a good indicator of whether an element contains a shock, the residual was employed as a shock sensor to reduce the reliance on the convex limiting approach. For an element $\mathcal{K}$, a straight-forward sensor was constructed as

$$T_{\mathcal{K}} = \begin{cases} 1 & \text{if } \|R_\sigma\|_{L^2,\mathcal{K}} > c_S h, \\ 0 & \text{else,} \end{cases} \tag{50}$$

where $T_{\mathcal{K}} = 1$ indicates that the element $\mathcal{K}$ is troubled. Although there can be sensitivity to the parameter c_S, we found that for all the cases considered, $c_S \in [1, 10]$ was sufficient with $c_S = 1$ taken for all cases.

E. Maximum Eigenvalue Approximation

In Theorem IV.2, the proof of local invariance relied on the auxiliary state being in the invariant set which, through Corollary II.1.1 and Definition II.1, means that the stability of the method is dependent on $\lambda_{\max}$. If this is not a strict upper bound on the maximum absolute eigenvalues, then stability cannot be guaranteed. Alternatively, if an overly-conservative bound is used, then the maximum stable time step will be diminished. We present some methods for the calculation of $\lambda_{\max}$, particularly with respect to the Euler equations.

The canonical estimate is that of Davis [46], defined in its second form as

$$\lambda_{\max} = \max\left(|u_l| + a_l, |u_r| + a_r\right). \tag{51}$$

This estimate is a strict bound when the solution to the Riemann problem is a double rarefaction, but, in general, it is insufficient to bound $\lambda_{\max}$. A further canonical choice is that of Einfeldt [47] where the Roe-average states are used to give

$$\lambda_{\max} = \max\left(|\tilde{u} - \tilde{a}|, |\tilde{u} + \tilde{a}|\right), \tag{52}$$

for

$$\tilde{u} = \frac{u_l\sqrt{\rho_l} + u_r\sqrt{\rho_r}}{\sqrt{\rho_l} + \sqrt{\rho_r}}, \quad \tilde{H} = \frac{H_l\sqrt{\rho_l} + H_r\sqrt{\rho_r}}{\sqrt{\rho_l} + \sqrt{\rho_r}}, \quad \text{and} \quad \tilde{a} = \sqrt{(\gamma - 1)\left(\tilde{H} - \frac{1}{2}\tilde{u}^2\right)} \tag{53}$$

where $H = (E + p)/\rho$ is the enthalpy. This approach inherits many deficiencies from the associated Roe [48] methods with inaccurate estimates for rarefactions and near-vacuum conditions posing significant challenges.

Three direct methods to form strict upper bounds were recently presented by Toro et al. [49] where the estimates are formed through differing state interpolations. A full description of these methods is beyond the scope of this paper; however, it is worth stating that care must be taken in the implementation to handle problems stemming from floating-point precision.

Lastly, Guermond and Popov [50] introduced a fast, indirect method for finding an estimate of the upper bound to great accuracy. In the course of finding this bound, a direct method was also introduced which is used as a starting point for the calculation. However, Toro et al. [49] found that the direct schemes they introduced were more accurate in the majority of tests compared to the direct scheme of Guermond and Popov [50], an observation shared in this work.

V. Numerical Experiments

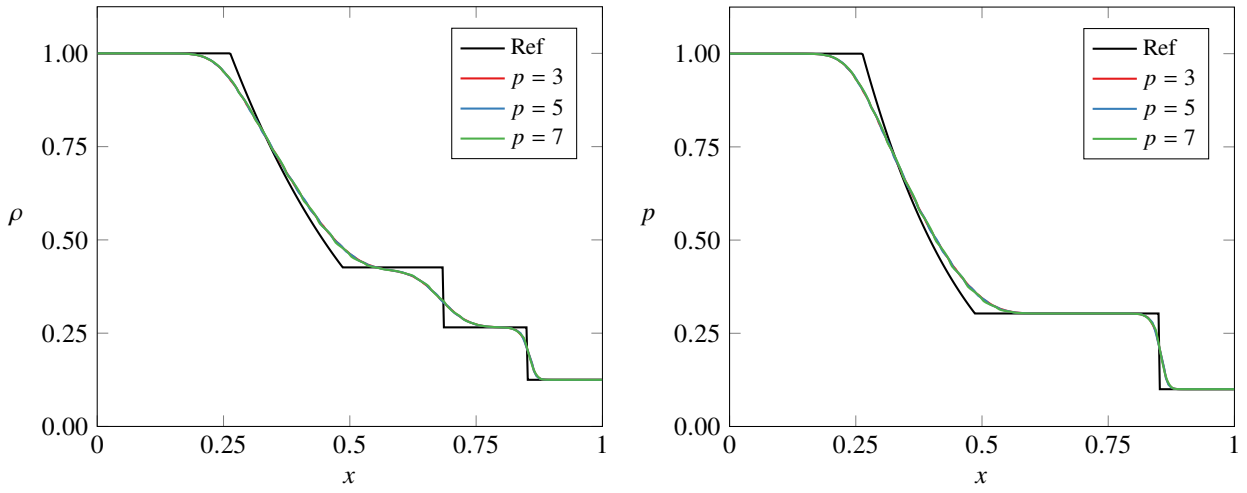
The shock capturing methodology described in Section IV was applied to a series of one dimensional test cases for the Euler equations. For the high-order scheme, FR with DG correction functions was used with the solution points placed using a Gauss–Legendre quadrature rule, and temporal integration was performed using the SSP-ERK3 scheme. The low-order scheme was employed corresponding to Eq. (29) where $\lambda_{\max}$ was set exactly using an indirect approach. For the shock sensor and entropy viscosity, the parameters were set as $c_E = 5 \times 10^{-3}$ and $c_{\max} = c_E/p$, and the entropy residual switch was set to h, unless otherwise stated.

A. Sod Shock Tube

The first test case was the Sod shock tube [51], solved on the domain $x \in \Omega = [0, 1]$ with the initial conditions

$$\mathbf{w}_l = \begin{bmatrix} \rho_l = 1 \\ u_l = 0 \\ p_l = 1 \end{bmatrix}, \quad \mathbf{w}_r = \begin{bmatrix} \rho_r = 0.125 \\ u_r = 0 \\ p_r = 0.1 \end{bmatrix}, \quad \mathbf{w} = \begin{cases} \mathbf{w}_l, & x \leqslant 0.5 \\ \mathbf{w}_r, & x > 0.5 \end{cases},$$

where $\mathbf{w}$ denotes the vector of primitive variables. For this case, an analytic reference solution may be calculated. A constant time step of $\Delta t = 5 \times 10^{-4}$ was used for the temporal integration from $t = 0$ to $t = 0.2$.



(a) Density (b) Pressure

Fig. 3 Sod shock tube at $t = 0.2$ for DoF ≈ 256.

Fig. 3 shows the results of the low-order scheme with artificial viscosity using the indirect exact $\lambda_{\max}$ [50]. The effect of the dissipation is clear and small oscillations are clearly visible in the solution. In the full graph viscosity method, it is common to see jagged artefacts due to the AV [38]. Furthermore, Fig. 3 shows that the dissipation does not adversely affect the solution as the order is increased, with the solution looking approximately the same across orders. This is consistent with the number of points in the low-order stencil being constant with order.

The results of applying the convex limiting procedure of Section IV.C are presented in Fig. 4. It is evident that a significant improvement to the results is observed. The rarefaction in this problem makes clear a property of invariant-preserving methods, namely that these methods do not necessarily produce monotonic solutions. However, that does not mean that this solution is aphysical. The small oscillations in the solution can be explained by considering the solution in the limit of vanishing viscosity of the Navier–Stokes equations. As the viscosity tends to zero, the formation of the bump becomes visible prior to the limit being reached.

Fig. 5 shows the solution when the convex limiting approach is combined with the shock sensor and entropy viscosity. In comparison to the purely limited approach, the oscillation downstream of the rarefaction was slightly reduced, and the upstream front of the rarefaction was better predicted.

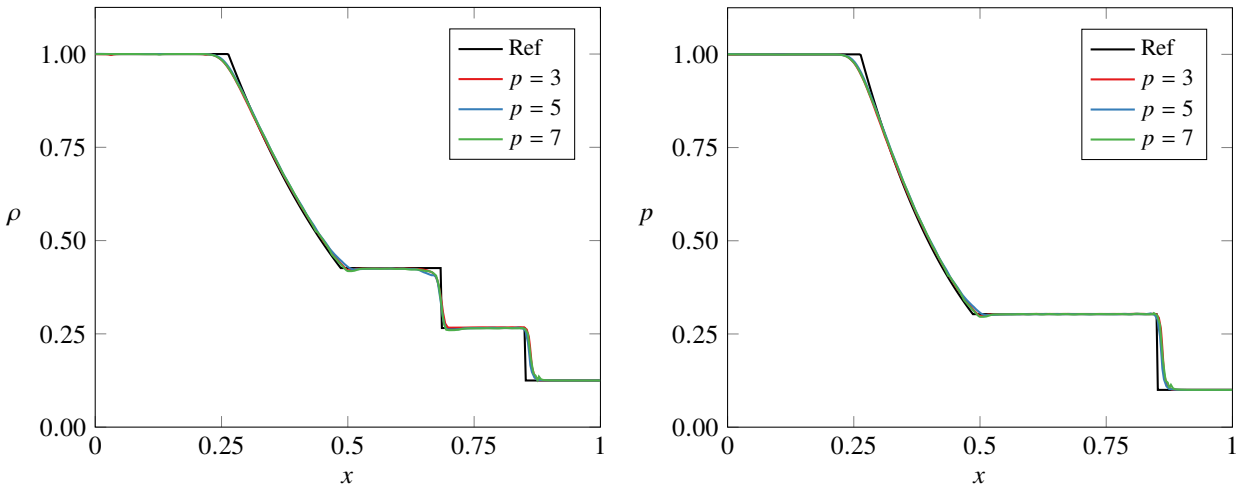
As the Sod shock tube permits an analytic solution, the error can be calculated. The L_1 and L_2 norms of the error were defined as

$$\|e\|_1 = \int_\Omega |\rho - \rho_{\text{exact}}| \, \mathrm{d}x, \tag{54}$$

$$\|e\|_2 = \sqrt{\int_\Omega (\rho - \rho_{\text{exact}})^2 \, \mathrm{d}x}, \tag{55}$$

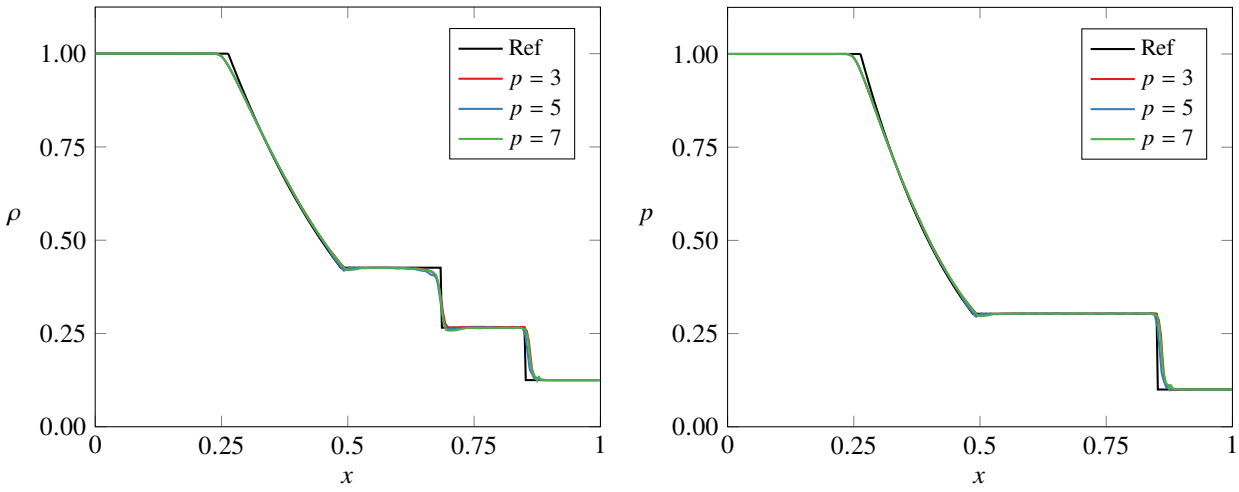
respectively.

The results for $p = 7$, displayed in Table 2, show that the scheme is first-order accurate in the L_1 norm. From the discussion in Section IV.C, Guermond et al. [25], and [43], with strict bounds upon specific entropy it should be possible



(a) Density (b) Pressure

Fig. 4 Sod shock tube at $t = 0.2$ for DoF ≈ 256 with convex limiting.



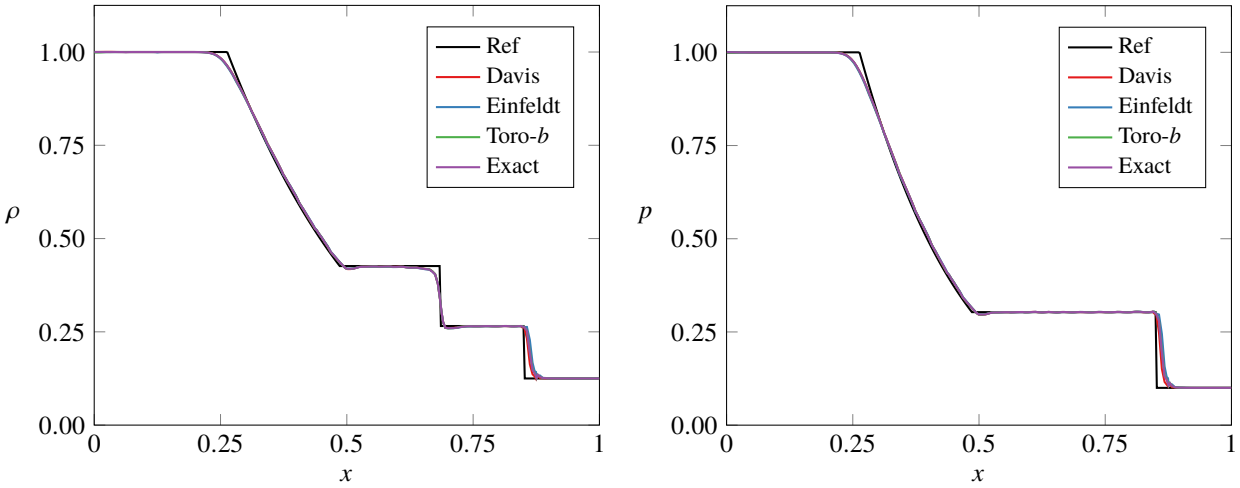
(a) Density (b) Pressure

Fig. 5 Sod shock tube at $t = 0.2$ for DoF ≈ 256 with convex limiting, shock detection, and entropy viscosity.

to attain second-order convergence in the L_1 norm but not in other norms. It is not clear whether the low-order accuracy is due to the simplified limiting procedure or if the over-prediction of the right-hand shock speed is dominating the error, polluting the convergence rate calculation. As a result of the over-prediction of the right-hand shock speed, the convergence rate in the L_2 norm is low and the L_∞ (not shown) norm is approximately constant. Further investigation is required to more fully understand the rates of convergence for this approach.

Table 2 Sod shock tube error for $p = 7$ with various DoF.

DoF	$\|e\|_1$	$\|e\|_2$
128	9.84×10^{-3}	2.05×10^{-2}
256	6.73×10^{-3}	1.95×10^{-2}
512	4.13×10^{-3}	1.59×10^{-2}
1024	3.14×10^{-3}	1.52×10^{-2}
2048	2.45×10^{-3}	1.45×10^{-2}
RoC	0.52	0.14



(a) Density

(b) Pressure

Fig. 6 Sod shock tube at $t = 0.2$ for DoF ≈ 256, $p = 7$ with convex limiting for various $\lambda_{\max}$ estimates.

As was detailed in Section IV.E, there are several direct approaches to calculate an approximate maximum wavespeed. To understand the effect of the various approximations on the accuracy of the method, the Sod shock tube was used and the results are presented in Fig. 6 with the error presented in Table 3. For this case, the various approximations did not have a significant impact upon the result, with the approach of Davis showing less error than the more complex approaches. Furthermore, the data show that the Toro-*a* [49] approximation results in a lower error than the exact wavespeed. This indicates that relaxing the bound while maintaining stability can lead to less overall dissipation. However, as is evident from the systematic $\lambda_{\max}$ over-estimates obtained by the Einfeldt method, too much additional dissipation will lead to a degradation in the solution error. The primary observation is that the methods of Toro et al. [49] are comparable to the exact $\lambda_{\max}$ without requiring Newton iterations to calculate.

B. Shu–Osher Shock Tube

Shu and Osher [52] introduced a more challenging case where the flow left of the initial discontinuity has some velocity and the density field right of the discontinuity is oscillatory. This leads to a stronger shock and the generation of high-frequency waves within the solution which can be challenging for highly dissipative schemes to resolve. The

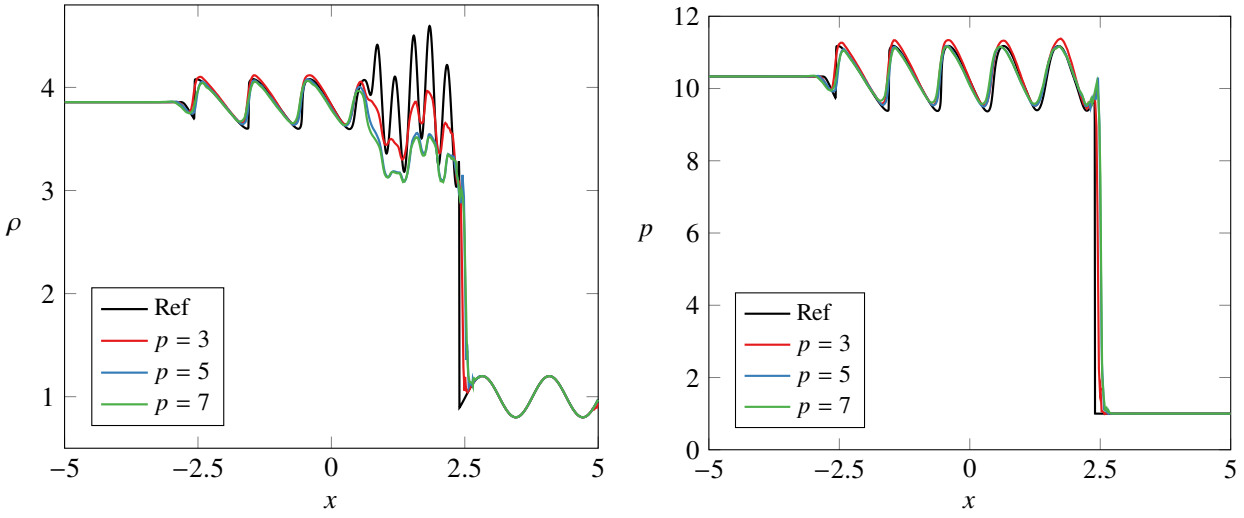
Table 3 Error for the Sod shock tube with DoF ≈ 256, $p = 7$ **for various** $\lambda_{\max}$ **estimates.**

Method	$\|e\|_1$	$\|e\|_2$
Davis	5.7235×10^{-3}	1.5697×10^{-2}
Einfeldt	5.7235×10^{-3}	1.8917×10^{-2}
Toro-a	6.1801×10^{-3}	1.6931×10^{-2}
Toro-b	6.2532×10^{-3}	1.7341×10^{-2}
Toro-c	6.2889×10^{-3}	1.7219×10^{-2}
Exact	6.2065×10^{-3}	1.7189×10^{-2}

initial condition is given by

$$\mathbf{w}_l = \begin{bmatrix} \rho_l = 3.857143 \\ u_l = 2.629369 \\ p_l = 10.333333 \end{bmatrix}, \quad \mathbf{w}_r = \begin{bmatrix} \rho_r = 1 + 0.2 \sin 5x \\ u_r = 0 \\ p_r = 1 \end{bmatrix}, \quad \mathbf{w} = \begin{cases} \mathbf{w}_l, & x \leqslant -4, \\ \mathbf{w}_r, & x > -4, \end{cases}$$

for $x \in \Omega = [-5, 5]$.

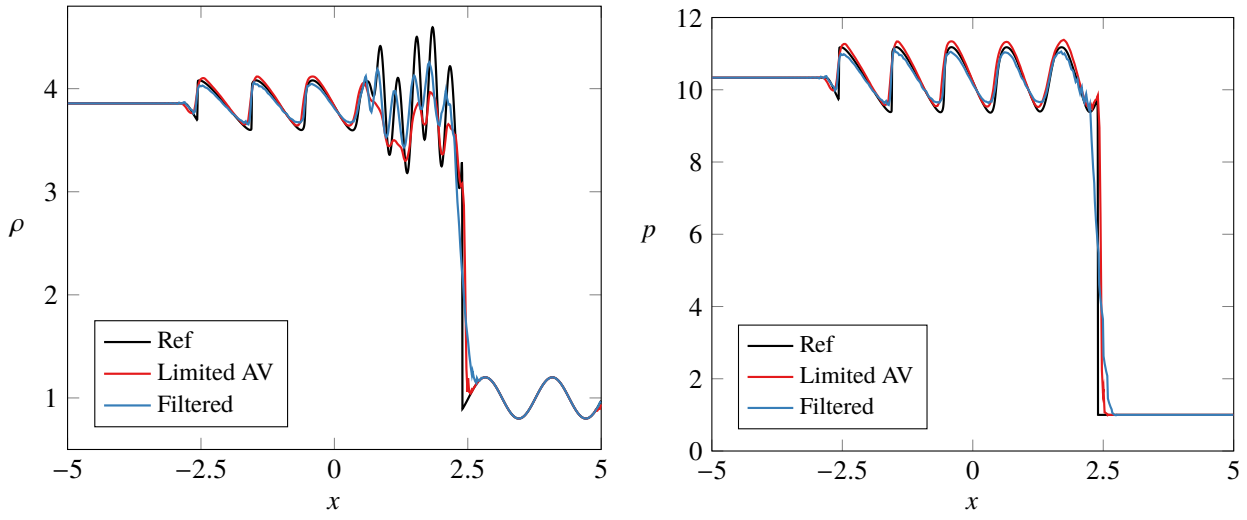


(a) Density (b) Pressure

Fig. 7 Shu–Osher shocktube at $t = 1.8$ **for** DoF ≈ 512 **with convex limiting, shock detection, and entropy viscosity.**

Fig. 7 shows the results of the convex limiting approach paired with the shock detection and entropy viscosity methods for a fixed $\Delta t = 2.5 \times 10^{-4}$. As the order was increased, the diffusion also increased, with the $p = 3$ result performing well for the given resolution but higher orders significantly damping out the oscillations upstream of the shock. This effect can be attributed to the entropy viscosity which can be excessively applied when large gradients in entropy are observed due to the increased level of Gibbs phenomenon at higher orders. Therefore, although the entropy viscosity procedure can reduce the need for convex limiting, it can bring about increasingly dissipated solutions.

The Shu–Osher test case is also useful for comparison due to its high frequency content that can be mistakenly damped out by many methods. Fig. 8 presents a comparison between the presented method and a filtering approach for DoF ≈ 500 at $p = 3$. The filtering approach is a modal filter calculated based on a p^{-2} decay in the energy of the density modes. Lower dissipation in the oscillatory region upwind of the primary shock is observed compared to the AV approach. However, there is evidence of Gibbs phenomenon, most notably in the pressure field in Fig. 8b. Although the filtering approach is relatively successful here, there are no guarantees as to the physicality of the solution.



(a) Density

(b) Pressure

Fig. 8 Shu–Osher shocktube at $t = 1.8$ for DoF ≈ 500, $p = 3$, comparing the AV approach to filtering.

C. Woodward–Colella Shock Tube

The problem of Woodward and Colella [53], in the form presented by Toro [54], is challenging for many schemes due to the large magnitude of the shocks present. The initial conditions are given by

$$\mathbf{w}_l = \begin{bmatrix} \rho_l = 1 \\ u_l = 0 \\ p_l = 10^3 \end{bmatrix}, \quad \mathbf{w}_r = \begin{bmatrix} \rho_r = 1 \\ u_r = 0 \\ p_r = 10^{-2} \end{bmatrix}, \quad \mathbf{w} = \begin{cases} \mathbf{w}_l, & x \leqslant 0.5, \\ \mathbf{w}_r, & x > 0.5, \end{cases}$$

for $\mathbf{x} \in \Omega = [0, 1]$. The results are presented in Fig. 9 for the convex limiting approach with shock detection and entropy viscosity. It is evident that, with the much larger relative magnitude of the shocks in this case, the entropy viscosity is introducing some spurious oscillations in a region of low density that is immediately upwind of a discontinuity. Zeroing the viscosity terms, as displayed in Fig. 10, shows the effect that these terms were having. Primarily, the oscillations are no longer observed. However, without the stabilisation of the high-resolution component of the solution, more reliance is put on the low-resolution component. A side effect of this is an over-prediction of the shock speeds due to the upwinding used. It was stated by Pazner [38] that similar limiting approaches which used only point-wise data were overly dissipative, and although this is not directly evident here, we posit that a benefit of a larger limiting stencil would result in more accurate predictions of the shock speeds.

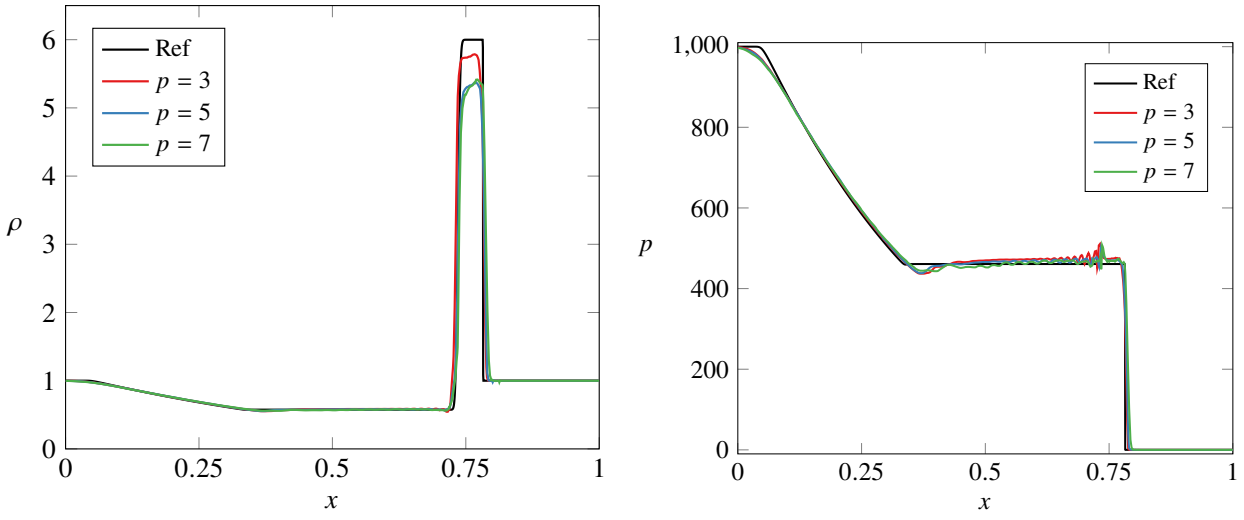
D. Near-Vacuum Shock Tube

A concern of shock capturing methods is their behaviour as the density field approaches a vacuum as small oscillations can cause the density to attain a negative value, leading the simulation to fail. To validate the convex limiting approach for these scenarios, we introduce the following test case, defined by the initial conditions

$$\mathbf{w}_l = \begin{bmatrix} \rho_l = 4 + (1 - \epsilon)\cos(\pi k x) \\ u_l = \sqrt{\frac{8\gamma}{3}} \\ p_l = 2 \end{bmatrix}, \quad \mathbf{w}_r = \begin{bmatrix} \rho_r = 1 - (1 - \epsilon)\cos(\pi k x) \\ u_r = 0 \\ p_r = 1 \end{bmatrix}, \quad \mathbf{w} = \begin{cases} \mathbf{w}_l, & x \leqslant 0.5, \\ \mathbf{w}_r, & x > 0.5. \end{cases}$$

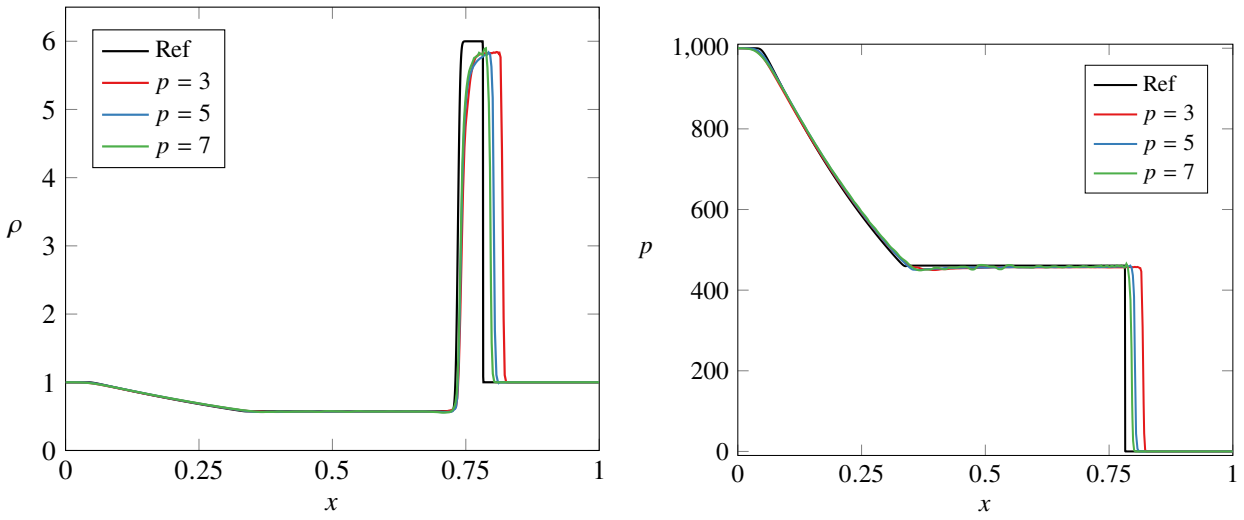
which is solved on the domain $\Omega = [0, 1]$ until $t = 0.07$ with $\epsilon = 10^{-2}$ and $k = 12$. This case is preferred over the classic 123 problem [54] where traditional methods are adequately able to produce a solution. A reference solution was generated for this case using an exact Godunov-type solver [54] with DoF $= 3 \times 10^4$.

Fig. 11 presents results for several orders with convex limiting and shock detection applied. Similar to the Woodward–Colella case, it was found that due to the large magnitude of the shocks particularly close to a near-vacuum



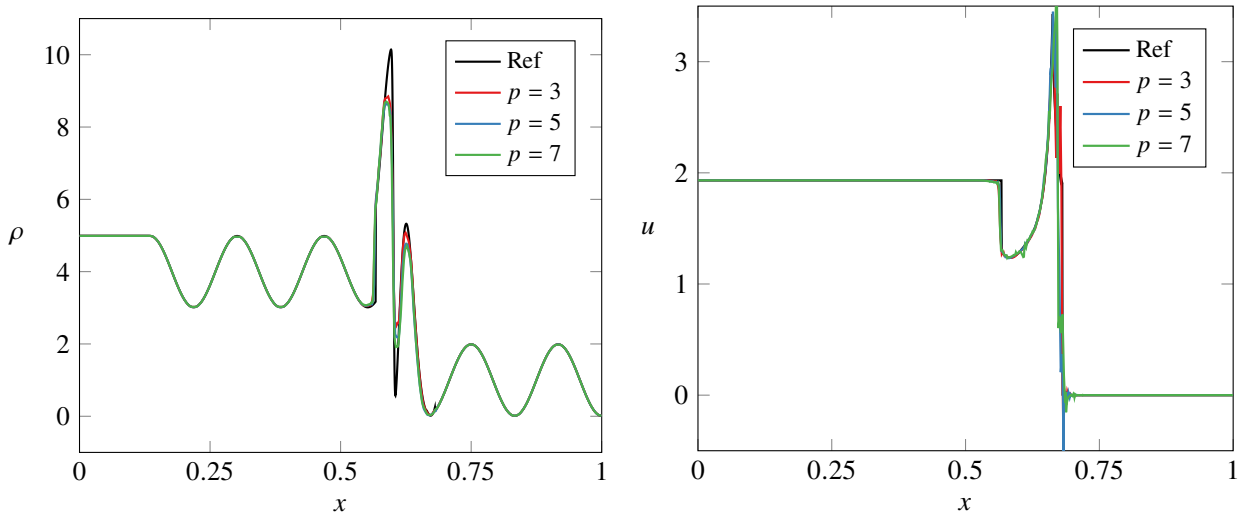
(a) Density (b) Pressure

Fig. 9 Woodward–Colella shocktube at $t = 0.012$ for DoF ≈ 512 with convex limiting, shock detection, and entropy viscosity.



(a) Density (b) Pressure

Fig. 10 Woodward–Colella shocktube at $t = 0.012$ for DoF ≈ 512 with convex limiting and shock detection.



(a) Density (b) Velocity

Fig. 11 Near vacuum shocktube at $t = 0.07$ for DoF ≈ 512 with convex limiting.

state, the entropy viscosity terms had a tendency to cause oscillations which led to negative density. The solution with convex limiting and shock detection performs well although slight oscillations, primarily in the velocity field, are noticeable in regions of near-vacuum in the vicinity of the shocks. The low density in these regions locally reduces the artificial viscosity which can be seen when considering that $(\rho u)_j - (\rho u)_i \sim \mathcal{O}(\rho)$ but $\lambda_{\max} \sim \mathcal{O}(\rho^{-1/2})$. Hence, when approaching a vacuum, oscillations that may be physical in terms on the invariant set may become more appreciable.

E. Computation Profiling

To give an estimate of the additional computational expense of this shock capturing approach, the one-dimensional implementation was profiled in a single thread workload with vectorisation enabled and MKL providing the linear algebra library. The case used was the Sod shock tube at $p = 3$ with 2048 DoF with convex limiting and the entropy residual shock detector. To reduce computational time, after the high and low-order flux divergences are calculated, the shock detection is performed and used to control the calculation of the sparse graph viscosity and the limiting. If a shock is not detected, then these components of the calculation are skipped.

Table 4 Sparse graph viscosity with convex limiting, shock sensor, and entropy viscosity. Runtime=$4.907s$**, CPI=**0.398**.**

Stage	Runtime fraction	exp/log fraction
High-order divergence	38.1%	18.3%
Low-order divergence	24.4%	20.2%
Entropy viscosity	11.3%	26.8%
Shock senor	1.1%	-
Sparse AV	3.0%	9%
Time update/limiter	18.3%	12.6%

The results of the profiling are set out in Tables 4 and 5. Although the implementation is not optimal, it does give some indication of the benefit of applying the shock sensor. When not using the shock sensor, the runtime is significantly longer with the majority of the additional computational time occurring in the limiting routine. The AV calculation fraction also increases, and in both cases, this was due to the large number of `exp` and `log` operations carried out. Although these operations are faster and more readily permit vectorisation than the `pow` operation, there is still a

Table 5 Sparse graph viscosity with convex limiting and *without* the shock sensor or entropy viscosity. Runtime=14.335*s*, **CPI=**0.472**.**

Stage	Runtime fraction	exp/log fraction
High-order divergence	10.5%	-
Low-order divergence	6.5%	-
Sparse AV	16.7%	38.4%
Time update/limiter	65.0%	55.3%

considerable computational overhead associated with the additional exponential and log calls required for the entropy divergence calculation which is highlighted with the use of the shock sensor and entropy viscosity. Our intention is to proceed to implement this in multiple dimensions in conjunction with GPU accelerators, and we included this brief account of profiling data to highlight the importance of shock sensing in conjunction with the other techniques applied here.

VI. Conclusion

An approach to enable high-order SEM to handle shocks has been presented wherein a low-order invariant domain preserving scheme is combined with a high-order flux reconstruction scheme via a convex limiting procedure. Summation-by-parts was utilised as a framework for finding conservative and linearly stable methods in a straightforward manner, and a non-parametric graph viscosity term was used to augment these schemes to produce a non-linearly stable, invariant domain preserving scheme. With this formulation of the low-order scheme, the stencil size is greatly reduced in comparison to the original graph viscosity approach, reducing the computational complexity of the convex limiting procedure. To further reduce the cost of the convex limiting approach, an entropy viscosity and entropy residual shock sensor have been applied. Numerical tests on the 1D Euler equations show that this approach achieves significantly increased accuracy over standard low-order methods while enforcing physical bounds in challenging cases with large discontinuities and near-vacuum conditions.

References

[1] Harlow, F. H., Dickman, D. O., Harris, D. E., and Martin, R. E., "Two-dimensional hydrodynamic calculations: LA-2301," Tech. rep., Los Alamos Scientific Laboratory, Los Alamos, 1959.

[2] Rusanov, V., "The calculation of the interaction of non-stationary shock waves and obstacles," *USSR Computational Mathematics and Mathematical Physics*, Vol. 1, No. 2, 1962, pp. 304–320. https://doi.org/10.1016/0041-5553(62)90062-9.

[3] Godunov, S. K., "A Difference Mehtod for Numerical Calculation of Discontinuous Solutions of the Equations of Hydrodynamics," *Mat. Sb.*, Vol. 47(89), No. 3, 1959, pp. 271–306.

[4] van Leer, B., "Flux-vector splitting for the Euler equations," *Eighth International Conference on Numerical Methods in Fluid Dynamics*, Springer Berlin Heidelberg, 1982, pp. 507–512. https://doi.org/10.1007/3-540-11948-5_66.

[5] Harten, A., Lax, P. D., and van Leer, B., "On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws," *SIAM Review*, Vol. 25, No. 1, 1983, pp. 35–61. https://doi.org/10.1137/1025002.

[6] Jameson, A., Schmidt, W., and Turkel, E., "Numerical solution of the Euler equations by finite volume methods using Runge Kutta time stepping schemes," *14th Fluid and Plasma Dynamics Conference*, American Institute of Aeronautics and Astronautics, 1981. https://doi.org/10.2514/6.1981-1259.

[7] Harten, A., "High resolution schemes for hyperbolic conservation laws," *Journal of Computational Physics*, Vol. 49, No. 3, 1983, pp. 357–393. https://doi.org/10.1016/0021-9991(83)90136-5.

[8] Harten, A., Engquist, B., Osher, S., and Chakravarthy, S. R., "Uniformly high order accurate essentially non-oscillatory schemes, III," *Journal of Computational Physics*, Vol. 71, No. 2, 1987, pp. 231–303. https://doi.org/10.1016/0021-9991(87)90031-3.

[9] Liu, X.-D., Osher, S., and Chan, T., "Weighted Essentially Non-oscillatory Schemes," *Journal of Computational Physics*, Vol. 115, No. 1, 1994, pp. 200–212. https://doi.org/10.1006/jcph.1994.1187.

[10] Wilbraham, H., "On a certain periodic function," *The Cambridge and Dublin Mathematical Journal*, Vol. 3, 1848, pp. 198–201.

[11] Lax, P. D., "Gibbs Phenomena," *Journal of Scientific Computing*, Vol. 28, No. 2-3, 2006, pp. 445–449. https://doi.org/10.1007/s10915-006-9075-y.

[12] Huynh, H. T., "A Flux Reconstruction Approach to High-Order Schemes Including Discontinuous Galerkin Methods," *18th AIAA Computational Fluid Dynamics Conference*, American Institute of Aeronautics and Astronautics, 2007. https://doi.org/10.2514/6.2007-4079.

[13] Vincent, P. E., Castonguay, P., and Jameson, A., "A New Class of High-Order Energy Stable Flux Reconstruction Schemes," *Journal of Scientific Computing*, Vol. 47, No. 1, 2010, pp. 50–72. https://doi.org/10.1007/s10915-010-9420-z.

[14] Reed, W. H., and Hill, T. R., "Triangular mesh methods for the neutron transport equation," , No. LA-UR-73-479, 1973, pp. 1–12.

[15] Cockburn, B., Karniadakis, G. E., and Shu, C.-W., "The Development of Discontinuous Galerkin Methods," *Lecture Notes in Computational Science and Engineering*, Springer Berlin Heidelberg, 2000, pp. 3–50. https://doi.org/10.1007/978-3-642-59721-3_1.

[16] Persson, P.-O., and Peraire, J., "Sub-Cell Shock Capturing for Discontinuous Galerkin Methods," *44th AIAA Aerospace Sciences Meeting and Exhibit*, American Institute of Aeronautics and Astronautics, 2006. https://doi.org/10.2514/6.2006-112.

[17] Barter, G., and Darmofal, D., "Shock Capturing with Higher-Order, PDE-Based Artificial Viscosity," *18th AIAA Computational Fluid Dynamics Conference*, American Institute of Aeronautics and Astronautics, 2007. https://doi.org/10.2514/6.2007-3823.

[18] Tadmor, E., "Shock capturing by the spectral viscosity method," *Computer Methods in Applied Mechanics and Engineering*, Vol. 80, No. 1-3, 1990, pp. 197–208. https://doi.org/10.1016/0045-7825(90)90023-f.

[19] Tadmor, E., "Super Viscosity And Spectral Approximations Of Nonlinear Conservation Laws," 1993.

[20] Ma, H., "Chebyshev–Legendre Super Spectral Viscosity Method for Nonlinear Conservation Laws," *SIAM Journal on Numerical Analysis*, Vol. 35, No. 3, 1998, pp. 893–908. https://doi.org/10.1137/s0036142995293912.

[21] Maday, Y., Kaber, S. M. O., and Tadmor, E., "Legendre Pseudospectral Viscosity Method for Nonlinear Conservation Laws," *SIAM Journal on Numerical Analysis*, Vol. 30, No. 2, 1993, pp. 321–342. https://doi.org/10.1137/0730016.

[22] Dumbser, M., and Loubère, R., "A simple robust and accurate a posteriori sub-cell finite volume limiter for the discontinuous Galerkin method on unstructured meshes," *Journal of Computational Physics*, Vol. 319, 2016, pp. 163–199. https://doi.org/10.1016/j.jcp.2016.05.002.

[23] Guermond, J.-L., and Popov, B., "Invariant Domains and First-Order Continuous Finite Element Approximation for Hyperbolic Systems," *SIAM Journal on Numerical Analysis*, Vol. 54, No. 4, 2016, pp. 2466–2489. https://doi.org/10.1137/16m1074291.

[24] Guermond, J.-L., Nazarov, M., Popov, B., and Tomas, I., "Second-Order Invariant Domain Preserving Approximation of the Euler Equations Using Convex Limiting," *SIAM Journal on Scientific Computing*, Vol. 40, No. 5, 2018, pp. A3211–A3239. https://doi.org/10.1137/17m1149961.

[25] Guermond, J.-L., Popov, B., and Tomas, I., "Invariant domain preserving discretization-independent schemes and convex limiting for hyperbolic systems," *Computer Methods in Applied Mechanics and Engineering*, Vol. 347, 2019, pp. 143–175. https://doi.org/10.1016/j.cma.2018.11.036.

[26] Lax, P. D., "Hyperbolic systems of conservation laws II," *Communications on Pure and Applied Mathematics*, Vol. 10, No. 4, 1957, pp. 537–566. https://doi.org/10.1002/cpa.3160100406.

[27] Glimm, J., "Solutions in the large for nonlinear hyperbolic systems of equations," *Communications on Pure and Applied Mathematics*, Vol. 18, No. 4, 1965, pp. 697–715. https://doi.org/10.1002/cpa.3160180408.

[28] Chueh, K. N., Conley, C. C., and Smoller, J. A., "Positively Invariant Regions for Systems of Nonlinear Diffusion Equations," *Indiana University Mathematics Journal*, Vol. 26, No. 2, 1977, pp. 373–392. URL http://www.jstor.org/stable/24891350.

[29] Hoff, D., "A finite difference scheme for a system of two conservation laws with artificial viscosity," *Mathematics of Computation*, Vol. 33, No. 148, 1979, pp. 1171–1171. https://doi.org/10.1090/s0025-5718-1979-0537964-9.

[30] Lax, P. D., "Weak solutions of nonlinear hyperbolic equations and their numerical computation," *Communications on Pure and Applied Mathematics*, Vol. 7, No. 1, 1954, pp. 159–193. https://doi.org/10.1002/cpa.3160070112.

[31] Hu, X. Y., Adams, N. A., and Shu, C.-W., "Positivity-preserving method for high-order conservative schemes solving compressible Euler equations," *Journal of Computational Physics*, Vol. 242, 2013, pp. 169–180. https://doi.org/10.1016/j.jcp.2013.01.024.

[32] Castonguay, P., Vincent, P. E., and Jameson, A., "A New Class of High-Order Energy Stable Flux Reconstruction Schemes for Triangular Elements," *Journal of Scientific Computing*, Vol. 51, No. 1, 2011, pp. 224–256. https://doi.org/10.1007/s10915-011-9505-3.

[33] Wang, L., and Yu, M., "An implicit high-order preconditioned flux reconstruction method for low-Mach-number flow simulation with dynamic meshes," *International Journal for Numerical Methods in Fluids*, Vol. 91, No. 7, 2019, pp. 348–366. https://doi.org/10.1002/fld.4759.

[34] Gottlieb, S., Shu, C.-W., and Tadmor, E., "Strong Stability-Preserving High-Order Time Discretization Methods," *SIAM Review*, Vol. 43, No. 1, 2001, pp. 89–112. https://doi.org/10.1137/s003614450036757x.

[35] Courant, R., Friedrichs, K., and Lewy, H., "On the Partial Difference Equations of Mathematical Physics," *IBM Journal of Research and Development*, Vol. 11, No. 2, 1967, pp. 215–234. https://doi.org/10.1147/rd.112.0215.

[36] Vincent, P., Castonguay, P., and Jameson, A., "Insights from von Neumann analysis of high-order flux reconstruction schemes," *Journal of Computational Physics*, Vol. 230, No. 22, 2011, pp. 8134–8154. https://doi.org/10.1016/j.jcp.2011.07.013.

[37] Ketcheson, D., and Ahmadia, A., "Optimal stability polynomials for numerical integration of initial value problems," *Communications in Applied Mathematics and Computational Science*, Vol. 7, No. 2, 2012, pp. 247–271. https://doi.org/10.2140/camcos.2012.7.247.

[38] Pazner, W., "Sparse invariant domain preserving discontinuous Galerkin methods with subcell convex limiting," , 2020.

[39] Ranocha, H., Öffner, P., and Sonar, T., "Summation-by-Parts and Correction Procedure via Reconstruction," *Lecture Notes in Computational Science and Engineering*, Springer International Publishing, 2017, pp. 627–637. https://doi.org/10.1007/978-3-319-65870-4_45.

[40] Dzanic, T., Trojak, W., and Witherden, F. D., "A Riemann Difference Scheme for Shock Capturing in Discontinuous Finite Element Methods," , 2020.

[41] Boris, J. P., and Book, D. L., "Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works," *Journal of Computational Physics*, Vol. 11, No. 1, 1973, pp. 38–69. https://doi.org/10.1016/0021-9991(73)90147-2.

[42] Maier, M., and Kronbichler, M., "Massively parallel 3D computation of the compressible Euler equations with an invariant-domain preserving second-order finite-element scheme," , 2020.

[43] Khobalatte, B., and Perthame, B., "Maximum principle on the entropy and second-order kinetic schemes," *Mathematics of Computation*, Vol. 62, No. 205, 1994, pp. 119–119. https://doi.org/10.1090/s0025-5718-1994-1208223-4.

[44] Guermond, J.-L., Pasquetti, R., and Popov, B., "Entropy viscosity method for nonlinear conservation laws," *Journal of Computational Physics*, Vol. 230, No. 11, 2011, pp. 4248–4267. https://doi.org/10.1016/j.jcp.2010.11.043.

[45] Lax, P. D., "Shock Waves and Entropy," *Contributions to Nonlinear Functional Analysis*, Elsevier, 1971, pp. 603–634. https://doi.org/10.1016/b978-0-12-775850-3.50018-2, URL https://doi.org/10.1016/b978-0-12-775850-3.50018-2.

[46] Davis, S. F., "Simplified Second-Order Godunov-Type Methods," *SIAM Journal on Scientific and Statistical Computing*, Vol. 9, No. 3, 1988, pp. 445–473. https://doi.org/10.1137/0909030.

[47] Einfeldt, B., "On Godunov-Type Methods for Gas Dynamics," *SIAM Journal on Numerical Analysis*, Vol. 25, No. 2, 1988, pp. 294–318. https://doi.org/10.1137/0725021.

[48] Roe, P., "Approximate Riemann solvers, parameter vectors, and difference schemes," *Journal of Computational Physics*, Vol. 43, No. 2, 1981, pp. 357–372. https://doi.org/10.1016/0021-9991(81)90128-5.

[49] Toro, E., Müller, L., and Siviglia, A., "Bounds for Wave Speeds in the Riemann Problem: Direct Theoretical Estimates," *Computers & Fluids*, Vol. 209, 2020, p. 104640. https://doi.org/10.1016/j.compfluid.2020.104640.

[50] Guermond, J.-L., and Popov, B., "Fast estimation from above of the maximum wave speed in the Riemann problem for the Euler equations," *Journal of Computational Physics*, Vol. 321, 2016, pp. 908–926. https://doi.org/10.1016/j.jcp.2016.05.054.

[51] Sod, G. A., "A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws," *Journal of Computational Physics*, Vol. 27, No. 1, 1978, pp. 1–31. https://doi.org/10.1016/0021-9991(78)90023-2.

[52] Shu, C.-W., and Osher, S., "Efficient implementation of essentially non-oscillatory shock-capturing schemes," *Journal of Computational Physics*, Vol. 77, No. 2, 1988, pp. 439–471. https://doi.org/10.1016/0021-9991(88)90177-5.

[53] Woodward, P., and Colella, P., "The numerical simulation of two-dimensional fluid flow with strong shocks," *Journal of Computational Physics*, Vol. 54, No. 1, 1984, pp. 115–173. https://doi.org/10.1016/0021-9991(84)90142-6.

[54] Toro, E. F., *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer Berlin Heidelberg, 2009. https://doi.org/10.1007/b79761.