High-order limiting methods using maximum principle bounds derived from the Boltzmann equation I: Euler equations

Tarik Dzanic[a,*], Luigi Martinelli[b]

[a]Center for Applied Scientific Computing, Lawrence Livermore National Lab, Livermore, CA 94550, USA
[b]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

Keywords:
Euler equations
Boltzmann equation
High-order
Limiting
Positivity-preserving
Discontinuous Galerkin

ABSTRACT

The use of limiting methods for high-order numerical approximations of hyperbolic conservation laws generally requires defining an admissible region/bounds for the solution. In this work, we present a novel approach for computing solution bounds and limiting for the Euler equations through the kinetic representation provided by the Boltzmann equation, which allows for extending limiters designed for linear advection directly to the Euler equations. Given an arbitrary set of solution values to compute bounds over (e.g., numerical stencil) and a desired linear advection limiter, the proposed approach yields an analytic expression for the admissible region of particle distribution function values, which may be numerically integrated to yield a set of bounds for the density, momentum, and total energy. These solution bounds are shown to preserve positivity of density/pressure/internal energy and, when paired with a limiting technique, can robustly resolve strong discontinuities while recovering high-order accuracy in smooth regions without any ad hoc corrections (e.g., relaxing the bounds). This approach is demonstrated in the context of an explicit unstructured high-order discontinuous Galerkin/flux reconstruction scheme for a variety of difficult problems in gas dynamics, including cases with extreme shocks and shock-vortex interactions. Furthermore, this work presents a foundation for limiting techniques for more complex macroscopic governing equations that can be derived from an underlying kinetic representation for which admissible solution bounds are not well-understood.

1. Introduction

The understanding of the dynamics of inviscid fluid flow governed by the compressible Euler equations plays a critical role in predicting the behavior of fluids in high-speed regimes. These gas dynamics equations are not only fundamental as a governing model for the study of compressible flow phenomena such as shock waves but also serve as a stepping-stone to more complex simulations of viscous compressible flows. Numerical methods for simulating transport-dominated physics have widely relied on the Euler equations as the prototypical model problem, largely due to the challenges posed by the strongly nonlinear interactions between shock waves, discontinuities, and vortical features in the flow. These challenges have driven the development of sophisticated numerical techniques to yield robust and high-fidelity resolution of these complex flow features.

Various numerical approaches have been utilized for gas dynamics simulations, with the brunt of applications focusing on second-order finite volume and finite difference schemes. Although well-established in the literature and broadly robust for many practical uses, their relatively high numerical dissipation and low accuracy make them prohibitively expensive for high-fidelity (i.e., scale-resolving) simulations of complex unsteady flows. As a result, a particular class of numerical schemes which have more recently grown in popularity are high-order methods, which broadly encompass finite difference, finite volume, and finite element methods that can achieve arbitrarily high-order accuracy. Due to this increased accuracy paired with lower numerical dissipation, these methods can generally offer better resolution of small-scale flow features and transport-dominated flow physics at a lower relative computational cost than low-order schemes [1]. However, the high-order approximations which underpin these methods are primarily designed to yield high accuracy in smooth regions of the flow. Around discontinuities or severe gradients, the use of high-order schemes can cause spurious oscillations in the numerical solution, often referred to as Gibbs phenomena. These spurious oscillations result in a numerically ill-behaved solution and, in many cases, the failure of the numerical scheme altogether.

*Corresponding author
dzanic1@llnl.gov (T. Dzanic)
ORCID(s): 0000-0003-3791-1134 (T. Dzanic); 0000-0001-6319-7646 (L. Martinelli)

To extend the use of high-order method to problems that exhibit discontinuities and steep gradients, it is often necessary to augment the scheme with an additional numerical stabilization approach to increase its robustness in the vicinity of such features. One such stabilization approach is through *a priori* or *a posteriori* limiting on the discrete solution, which relies on using a secondary numerical approximation that is more robust (but generally less accurate). The solutions as predicted by the high-order scheme and the secondary low-order scheme are then combined in a manner which ideally ensures sufficient numerical stabilization around discontinuities without sacrificing the accuracy of the high-order scheme in smooth regions of the flow. Various forms of limiting have been applied to hyperbolic conservation laws, typically falling into variants of the flux-corrected transport method of Boris and Book [2] (e.g., [3–8]) or through limiting/modification of the underlying solution/operator approximations (e.g., [9–20]). We remark here that these examples only cover a small subset of the numerical approaches in the literature and do not by any means encompass all the possible limiting techniques used over the decades.

One of the primary difficulties in designing performant limiting methods for high-order schemes stems not necessarily from the development of the limiting schemes themselves but instead from the difficulty of defining exactly what a "well-behaved" numerical solution is for nonlinear systems of equations. Limiting between the low-order and high-order solution is typically dictated by some ad hoc heuristic (e.g., indicator functions) or through a set of bounds (i.e., admissible regions/sets) for the solution. While heuristic approaches can be effective for some applications, they often rely on empirical problem-dependent parameters which can misidentify regions that require limiting and degrade the accuracy and robustness of the approach. On the other hand, limiting based on a set of admissible bounds for the solution is typically a highly-robust alternative, often providing guarantees that the limited solution will abide by certain physical constraints [8, 21] (although typically at a larger computational expense). However, it is unclear what the optimal bounds are to enforce on the solution, tying in to the problem of defining what a well-behaved numerical solution is. Bounds on the numerical properties of the solution (e.g., total variation diminishing [9], local extremum diminishing methods [11], etc.) have shown to be robust for many applications, but these bounds are often quite restrictive and unnecessarily degrade the accuracy of the high-order scheme in smooth regions without ad hoc corrections, whereas more advanced stencil modification techniques (e.g., ENO/WENO/TENO-type schemes and limiters [12, 22]) can retain this accuracy but are typically prohibitively expensive and algorithmically complex for arbitrary mixed unstructured meshes. Bounds based on attributes of the underlying governing equations/systems have shown promise, relying on either physical constraints (e.g., positivity of density/pressure [8, 13]) or additional mathematical properties (e.g., minimum entropy principle [17, 18], invariant region [4, 23]). Nevertheless, they still exhibit some difficulties, such as sometimes also reducing to first-order accuracy around smooth extrema (requiring ad hoc corrections such as relaxing the bounds), requiring nonlinear limiting procedures which may be computationally expensive, and, most importantly, relying on mathematical properties which may not easily extend to systems which are of more practical interest (e.g., Navier–Stokes equations, multi-physics applications, etc.).

The methods to be presented in this work are motivated by two primary observations. The first observation is that numerical approximations of the nonlinear Boltzmann equation, a general kinetic description of molecular gas dynamics which underpins the behavior of fluids governed by the Euler and Navier–Stokes equations, are significantly more robust than approximations of the associated continuum conservation laws. In fact, high-order schemes for the Boltzmann equation have been effectively used to directly resolve shock structures without any numerical shock capturing approaches (or with just simple limiting techniques) while accurately resolving smooth features in the flow [24–27]. The second observation is that linear advection, which governs particle transport in the Boltzmann equation, is a significantly easier governing equation to design high-order limiting schemes for. Stabilization methods for high-order approximations of the advection equation have been developed for decades, and due to the linear scalar nature of the equation, simple yet effective bounds can be formed for the solution (e.g., local maximum principle). As such, there exist a variety of high-order limiting schemes for the advection equation that one can choose from, with varying degrees of robustness and accuracy.

With these observations in mind, the goal of this work is to introduce a novel approach to computing and enforcing bounds for limiting schemes used in high-order approximations of the Euler equations. The basis of this work is the connection between the Euler equations and the underlying kinetic representation provided by the Boltzmann equation, which essentially allows for one to extend limiting techniques designed for linear advection to the Euler equations. The proposed approach relies on the existence of an exact representation of the particle distribution function of the Boltzmann equation in the inviscid limit, for which the desired linear advection limiting technique can analytically be applied to yield a set of admissible bounds for the *conserved variables* (i.e., density, momentum, energy) in the Euler equations. We demonstrate this approach in the context of an explicit unstructured high-order discontinuous

Galerkin/flux reconstruction scheme [28, 29] utilizing an associated local maximum principle preserving limiter [30] for the underlying advection equation — however, this approach is broadly applicable in the sense of computing bounds given a set of discrete solution values (e.g., numerical stencil, local neighborhood, etc.), which can be extended to other bounds-preserving limiting methods such as flux-corrected transport and/or slope-limiting techniques. This approach is used to simulate a variety of difficult problems in gas dynamics, including cases with extreme shocks, near-vacuum conditions, and small-scale smooth flow features, showcasing its robustness around strong discontinuities and accuracy in smooth regions of the flow. Furthermore, this work presents a foundation for limiting techniques for more complex macroscopic governing equations that can be derived from an underlying kinetic representation (e.g., Navier–Stokes equations, magnetohydrodynamics, radiation-hydrodynamics, etc.) for which admissible solution bounds are not necessarily well-understood.

The remainder of this manuscript is organized as follows. In Section 2, some preliminaries on the Euler and Boltzmann equations as well as limiting methods for high-order discontinuous Galerkin schemes are presented. The proposed approach is then introduced in Section 3, followed by numerical analysis in Section 4 and implementation details in Section 5. The results of a variety of numerical experiments are then presented in Section 6 followed by conclusions in Section 7.

2. Preliminaries

2.1. Euler equations

The system of interest in this work is represented by the compressible Euler equations for gas dynamics in d dimensions without source terms, written in conservation form as

$$\partial_t \mathbf{w}(\mathbf{x}, t) + \boldsymbol{\nabla} \cdot \mathbf{F}(\mathbf{w}) = 0, \tag{1}$$

where the solution $\mathbf{w}$ and flux $\mathbf{F}(\mathbf{w})$ are defined as

$$\mathbf{w} = \begin{bmatrix} \rho \\ \mathbf{m} \\ E \end{bmatrix} \quad \text{and} \quad \mathbf{F}(\mathbf{w}) = \begin{bmatrix} \mathbf{m}^T \\ \mathbf{m} \otimes \mathbf{U} + P\mathbf{I} \\ (E + P)\mathbf{U}^T \end{bmatrix}. \tag{2}$$

Here, ρ is the density, $\mathbf{m}$ is the momentum, and E is the total energy. The symbol $\mathbf{I}$ denotes the identity matrix in $\mathbb{R}^{d\times d}$, $\mathbf{U} = \mathbf{m}/\rho$ denotes the velocity, and P denotes the pressure, computed through the ideal gas assumption as

$$P = (\gamma - 1)\rho e = (\gamma - 1)\left(E - \frac{1}{2}\mathbf{m}\cdot\mathbf{m}/\rho\right), \tag{3}$$

where γ is the specific heat ratio and e is the specific internal energy. It is sometimes convenient to represent the solution in terms of the primitive variables $\mathbf{q} = [\rho, \mathbf{U}, P]^T$. The solution of the Euler equations is endowed with a convex invariant set corresponding to the positivity of density ($\rho \geq 0$) and positivity of pressure/internal energy ($P, \rho e \geq 0$).

2.2. Boltzmann equation

The Euler equations can be recovered as the limiting behavior of the Boltzmann equation, given as

$$\partial_t f(\mathbf{x}, \mathbf{u}, t) + \mathbf{u}\cdot\nabla f = \frac{1}{Kn} C(f, f'), \tag{4}$$

where $f(\mathbf{x}, \mathbf{u}, t) \in \mathbb{R}^d$ is a scalar particle distribution function, Kn is the Knudsen number, and $C(f, f')$ is the collision operator which models particle interactions [31]. This distribution function is defined over a higher-dimensional space, which includes not only the spatial position $\mathbf{x}$ but also the microscopic velocity $\mathbf{u}$, and represents the probability density of a particle existing at some location $\mathbf{x}$ with some velocity $\mathbf{u}$. For brevity, we neglect the internal energy domain here, but the effects of internal energy are later presented. The conserved flow variables $\mathbf{w}(\mathbf{x}, t)$ can be recovered through the moments of the distribution function as

$$\mathbf{w}(\mathbf{x}, t) = [\rho, \mathbf{m}, E]^T = \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{u}, t)\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u}, \tag{5}$$

where $\boldsymbol{\psi}(\mathbf{u}) := [1, \mathbf{u}, (\mathbf{u}\cdot\mathbf{u})/2]^T$ is the vector of collision invariants. We utilize the notation that the lowercase symbol $\mathbf{u}$ refers to the microscopic velocity and the uppercase symbol $\mathbf{U}$ refers to the macroscopic velocity.

In the limit of zero Knudsen number (i.e., zero viscosity), the distribution function $f(\mathbf{x}, \mathbf{u}, t)$ tends towards a unique equilibrium distribution function $g(\mathbf{x}, \mathbf{u}, t)$ which minimizes (across the velocity domain) the entropy $H(z)$, i.e.,

$$g = \arg\min_z H(z), \qquad H(z) = \int_{\mathbb{R}^d} z \log(z)\, \mathrm{d}\mathbf{u}. \tag{6}$$

and possesses identical primary moments, i.e.,

$$\int_{\mathbb{R}^d} g(\mathbf{x}, \mathbf{u}, t)\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u} = \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{u}, t)\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u}. \tag{7}$$

From the H-theorem, the equilibrium distribution function (for a monatomic particle) is a Maxwell–Boltzmann distribution of the form

$$g(\mathbf{x}, \mathbf{u}, t) = \frac{\rho(\mathbf{x}, t)}{[2\pi\theta(\mathbf{x}, t)]^{d/2}} \exp\left[-\frac{\|\mathbf{u} - \mathbf{U}(\mathbf{x}, t)\|_2^2}{2\theta(\mathbf{x}, t)}\right], \tag{8}$$

where $\theta = P/\rho$ is a scaled temperature and $\|\cdot\|_2^2$ denotes the squared norm along the spatial dimension. We sometimes use the notation $g(\mathbf{w})$ to denote the equilibrium distribution corresponding to the macroscopic state $\mathbf{w}$. The effects of particle collision drive the distribution function f to this equilibrium state g. In particular, the collision operator is known to conserve primary moments, i.e.,

$$\int_{\mathbb{R}^d} C(f, f')\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u} = \mathbf{0}, \tag{9}$$

and is identically zero at equilibrium, i.e.,

$$C(g, g') = 0. \tag{10}$$

2.3. Discontinuous Galerkin methods

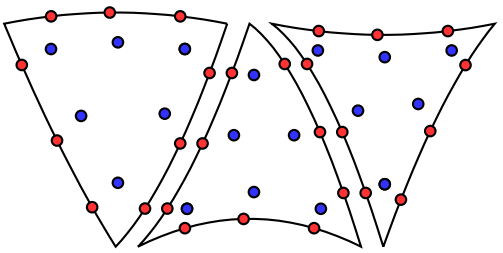


Figure 1: Schematic of an unstructured discontinuous Galerkin approximation with volume (blue) and surface (red) quadrature nodes.

While the proposed approach to be presented is somewhat agnostic to the numerical method of choice, this work particularly focuses on nodal discontinuous Galerkin (DG)-type [28] approximations to Eq. (1), including strong-form representations such as flux reconstruction [29] schemes and collocated discontinuous spectral element methods. In these approaches, the domain Ω is partitioned into N elements Ω_k such that $\Omega = \bigcup_N \Omega_k$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. With a slight abuse of notation, we use $\mathbf{w}(\mathbf{x})$ to denote the discrete solution within each element Ω_k, which is approximated by a set of N_s nodal basis functions as

$$\mathbf{w}(\mathbf{x}) = \sum_{i=1}^{N_s} \mathbf{w}_i \phi_i(\mathbf{x}) \subset V_h, \tag{11}$$

where $\phi_i(\mathbf{x})$ are the nodal basis functions, $\mathbf{w}_i$ are their associated coefficients, and V_h is the piece-wise polynomial space spanned by the basis functions. The order of the approximation is denoted as $\mathbb{P}_p$, where p is the maximal order of $\mathbf{u}(\mathbf{x})$.

By integrating Eq. (1) with respect to a test function $\mathbf{w}_h(\mathbf{x}) \subset V_h$, the DG formulation is obtained as

$$\sum_{k=1}^{N} \left\{ \int_{\Omega_k} \partial_t \mathbf{w} \cdot \mathbf{w}_h \, \mathrm{d}V + \int_{\partial\Omega_k} \overline{\mathbf{F}}(\mathbf{w}^-, \mathbf{w}^+, \mathbf{n}) \cdot \mathbf{w}_h \, \mathrm{d}\mathbf{s} - \int_{\Omega_k} \mathbf{F}(\mathbf{w}) \cdot \nabla \mathbf{w}_h \, \mathrm{d}V \right\} = 0. \tag{12}$$

These integrals are typically performed by quadrature over volume and surface quadrature nodes, shown by the blue and red nodes in Fig. 1, respectively. Similar strong form formulations can be obtained through a correction function-type approach [29], which use corresponding solution and flux nodes akin to volume and surface quadrature nodes. Along each element interface $\partial\Omega_k$, two solutions exist, the solution within the element of interest (denoted by the $\mathbf{w}^-$) and the solution of its face-adjacent neighbor (denoted by $\mathbf{w}^+$). The flux function here is replaced by a numerical interface flux $\overline{\mathbf{F}}(\mathbf{w}^-, \mathbf{w}^+, \mathbf{n})$, which depends on these two solutions as well as the outward facing normal vector $\mathbf{n}$. The numerical flux is generally computed using exact [32] or approximate Riemann solvers (e.g., Rusanov [33]).

While the solution of high-order DG schemes generally does not abide by certain key properties of the underlying hyperbolic systems (e.g., maximum principles, invariant regions, etc.), the scheme possesses a useful feature in that the evolution of the element-wise mean of the solution $\overline{\mathbf{w}}$, defined as

$$\overline{\mathbf{w}} = \frac{\int_{\Omega_k} \mathbf{w}(\mathbf{x}) \, \mathrm{d}\mathbf{x}}{\int_{\Omega_k} \mathrm{d}\mathbf{x}}, \tag{13}$$

is dependent on only the interface flux, i.e.,

$$\partial_t \overline{\mathbf{w}} = - \int_{\partial\Omega_k} \overline{\mathbf{F}}(\mathbf{w}^-, \mathbf{w}^+, \mathbf{n}) \, \mathrm{d}\mathbf{s}, \tag{14}$$

which allows it to be represented as a convex combination of temporal updates of first-order Godunov schemes [34]. Under some relatively minor assumptions on the numerical scheme (e.g., Courant-–Friedrichs—Lewy condition, appropriate Riemann solver, explicit strong stability preserving time integration, appropriate quadrature, etc.), this ensures that the element-wise mean preserves convex invariants of the system [30].

2.4. Maximum principle preserving limiting

This property of the element-wise mean of high-order DG schemes allows it to be used in many numerical stabilization approaches, some examples of which are referenced as follows [13, 17, 18, 26, 30, 35]. In the simplest example, the element-wise mean can be used to enforce a maximum principle on the advection equation, given as

$$\partial_t u(\mathbf{x}, t) + \mathbf{c} \cdot \nabla u = 0, \tag{15}$$

where $\mathbf{c}$ is the velocity. It is known that the solution of scalar hyperbolic conservation laws abides by a maximum principle in the sense that the extrema are bounded by the initial data $u(\mathbf{x}, 0)$. In a discrete setting, this can be formulated as a discrete maximum principle as

$$m \leq u(\mathbf{x}_i, t) \leq M \;\forall\; i \in V, t \geq 0, \tag{16}$$

where m and M are the lower and upper bounds of the solution, respectively, and V is some set of nodes over which the maximum principle should be enforced. In general, V corresponds to the discrete locations which the solution needs to be evaluated (e.g., volume/surface quadrature nodes for DG schemes, solution/flux nodes for flux reconstruction schemes, etc.).

One can apply a global maximum principle, where the bounds are computed over the domain as

$$m = \min_{\mathbf{x} \in \Omega} u(\mathbf{x}, 0), \quad M = \max_{\mathbf{x} \in \Omega} u(\mathbf{x}, 0), \tag{17}$$

which results in a single set of bounds for the entire solution. The discrete solution within each element at some time step t^n can be limited in an *a priori* (or *a posteriori*) way using the approach of Zhang and Shu [30] as

$$\widetilde{u}^n(\mathbf{x}) = \overline{u}^n + \alpha \left(u^n(\mathbf{x}) - \overline{u}^n \right), \tag{18}$$

where $\widetilde{u}^n(\mathbf{x})$ is the limited solution, $\overline{u}^n$ is the element-wise mean, and α is the limiting factor. It can be shown that if α is computed as

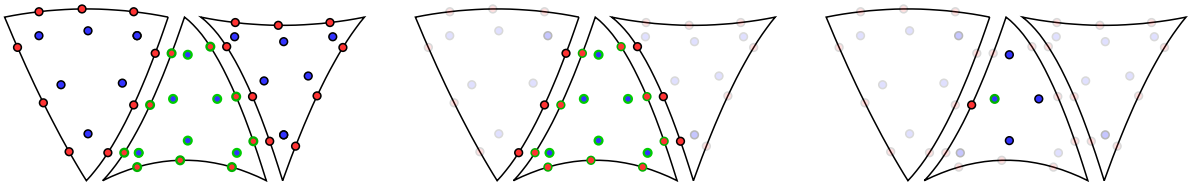
$$\alpha = \min\left[1, \left|\frac{M-\overline{u}}{\max\limits_{i\in V} u(\mathbf{x}_i)-\overline{u}}\right|, \left|\frac{m-\overline{u}}{\min\limits_{i\in V} u(\mathbf{x}_i)-\overline{u}}\right|\right], \tag{19}$$

the limited solution is guaranteed to preserve the maximum principle, i.e., $m \leq \widetilde{u}(\mathbf{x}_i) \leq M \;\forall\, i \in V$.

In practice, this is often not restrictive enough since local extrema can exist in the solution which may not be adequately stabilized by global bounds. It is generally more advisable to enforce a discrete *local* maximum principle, for which the bounds are computed at each time step over some local stencil S for each element Ω_k as

$$m_k^{n+1} = \min_{i\in S} u(\mathbf{x}_i, t^n), \quad M_k^{n+1} = \max_{i\in S} u(\mathbf{x}_i, t^n). \tag{20}$$

The local maximum preserving limiting approach requires the user to choose this local stencil S for some corresponding nodal set V to apply the bounds over. Various approaches for this exist, some examples of which are shown in Fig. 2. A common approach is to apply bounds uniformly across an element, with each volume/surface quadrature node having the same bounds, which may be computed over a local domain of influence such as the face/Voronoi neighbors for that element (Fig. 2a) or the numerical stencil for that element (Fig. 2b). More localized/restrictive approaches exist as well, where unique bounds can be computed over a local subset of nodes (e.g., a first-order stencil) within an element and enforced for each single node (Fig. 2c). These different approaches broadly have their own advantages and disadvantages, with less restrictive bounds (e.g., face neighbors) often resulting in less numerical dissipation around smooth regions of the solution and more restrictive bounds (e.g., nodal neighbors) often resulting in better behaved solutions around discontinuities (particularly for very high-order approximations) and vice versa. The numerical intricacies of these approaches as well as their unique drawbacks are described in more detail in Anderson et al. [36], although the examples shown here and in the referenced work for local maximum principle limiters of DG approximations of linear transport are by no means exhaustive.



(a) Face neighbors (b) Numerical stencil (c) Nodal neighbors

Figure 2: Various examples of choices for the stencil to compute bounds over (S, shown by solid red/blue nodes) as well as the nodes to enforce bounds over (V, shown by nodes with green outlines), including face/Voronoi neighbors (left), numerical stencil (middle), and nodal neighbors (right).

3. Methodology

The overarching goal of this work is to use insights from high-order limiting schemes for the Boltzmann equation to guide the development of limiting schemes for the Euler equations. Essentially, this work aims to find a novel approach to computing and enforcing bounds on the solution of Euler equations through its connection with an underlying kinetic transport equation, which allows for the use of limiting techniques developed for linear advection. Due to the scalar linear nature of the advection equation, these limiting techniques have become significantly more established over the decades and are simpler and more robust than methods required for nonlinear hyperbolic systems of equations, with many options of varying degrees of algorithmic simplicity, robustness, and accuracy. Most importantly, through this approach, one instead recovers local bounds for the *conserved variables* of the solution, i.e.,

$$\rho(\mathbf{x}_i) \in [\rho_{\min}, \rho_{\max}], \tag{21a}$$

$$\mathbf{m}_j(\mathbf{x}_i) \in [\mathbf{m}_{j,\min}, \mathbf{m}_{j,\max}] \quad \text{for} \quad j \in \{1, \dots, d\}, \tag{21b}$$

$$E(\mathbf{x}_i) \in [E_{\min}, E_{\max}], \tag{21c}$$

for some arbitrary point $\mathbf{x}_i$, which allows for the use of simple scalar limiting techniques and differs from many limiting approaches which enforce bounds on linear/nonlinear functionals of the solution (e.g., density, pressure, entropy, etc.).

3.1. Boltzmann representation of the Euler equations

The Euler equations can be represented in terms of the zero Knudsen number limit of the Boltzmann equation. Consider an *explicit* approximation of the Boltzmann equation at some time step t^n (denoted by the superscript n), written as

$$\partial_t f^n(\mathbf{x}, \mathbf{u}) = -\mathbf{u}\cdot\nabla f^n + \frac{1}{Kn} C(f^n, f^{n'}). \tag{22}$$

For zero Knudsen number, the distribution function is in equilibrium, such that we obtain

$$f^n(\mathbf{x}, \mathbf{u}) = g^n(\mathbf{x}, \mathbf{u}) = g(\mathbf{w}^n(\mathbf{x})), \tag{23}$$

i.e., the distribution function is the Maxwell–Boltzmann distribution in Eq. (8) for the corresponding macroscopic solution state $\mathbf{w}$. Furthermore, from Eq. (10), the collision term in the equilibrium limit *for a distribution function in equilibrium* is zero, i.e.,

$$\lim_{Kn\to 0} \frac{1}{Kn} C(g, g') = 0. \tag{24}$$

Therefore, a single *explicit* temporal integration step for the Boltzmann equation with a solution initially at equilibrium ($f^n = g^n$) yields the collisionless Boltzmann equation, written as

$$\partial_t f^n(\mathbf{x}, \mathbf{u}) = -\mathbf{u}\cdot\nabla f^n, \tag{25}$$

which can be integrated across the velocity domain to yield the Euler equations as

$$\int_{\mathbb{R}^d} \partial_t f^n(\mathbf{x}, \mathbf{u}) \boldsymbol{\psi}(\mathbf{u}) \, \mathrm{d}\mathbf{u} = \partial_t \mathbf{w}^n(\mathbf{x}) \quad \text{and} \quad \int_{\mathbb{R}^d} -\mathbf{u}\cdot\nabla f^n(\mathbf{x}, \mathbf{u}) \boldsymbol{\psi}(\mathbf{u}) \, \mathrm{d}\mathbf{u} = -\boldsymbol{\nabla}\cdot\mathbf{F}\left(\mathbf{w}^n(\mathbf{x})\right). \tag{26}$$

Note that this observation is only valid for a single explicit step for an equilibrium state, as any deviations from equilibrium (driven by the transport term) are then immediately reversed by the (infinitely large) collision term at the next time step.

3.2. Maximum principle bounds for the Boltzmann equation

The collisionless Boltzmann equation can be considered as a high-dimensional linear transport equation, and for any arbitrary point in the velocity domain $\mathbf{u} = \mathbf{u}_0$, $f_0 = f(\mathbf{u}_0)$, this simply yields a corresponding linear advection equation,

$$\partial_t f_0(\mathbf{x}, t) + \mathbf{u}_0\cdot\nabla f_0 = 0. \tag{27}$$

As such, we can apply standard stabilization techniques developed for high-order approximations of linear advection such as a local maximum principle preserving limiter. For example, one can apply element-wise forward-in-time maximum principle bounds as presented in Section 2.4, which, for an arbitrary element Ω_k, enforces bounds on the nodes $j \in V$ computed over the stencil $i \in S$ as

$$\min_{i\in S} f_0^n(\mathbf{x}_i) \le f_0^{n+1}(\mathbf{x}_j) \le \max_{i\in S} f_0^n(\mathbf{x}_i) \quad \forall\, j \in V. \tag{28}$$

We present this in terms a forward Euler approximation, but one can straightforwardly generalize this to higher-order time stepping schemes. Since $\mathbf{u}_0$ is arbitrary, this local maximum principle can be applied uniformly across the velocity domain, i.e.,

$$f_{\min}^{n+1}(\mathbf{u}) \le f^{n+1}(\mathbf{x}_j, \mathbf{u}) \le f_{\max}^{n+1}(\mathbf{u}) \quad \forall\, j \in V\ ,\ \forall\, \mathbf{u} \in \mathbb{R}^d. \tag{29}$$

where

$$f_{\min}^{n+1}(\mathbf{u}) = \min_{i\in S} f^n(\mathbf{x}_i, \mathbf{u}) \quad \text{and} \quad f_{\max}^{n+1}(\mathbf{u}) = \max_{i\in S} f^n(\mathbf{x}_i, \mathbf{u}). \tag{30}$$

We note here that limiting methods of this form are not necessarily just constrained to collisionless Boltzmann approximations, and various techniques that include the effects of collision can also be used (e.g., operator splitting, retaining only positivity bounds [26, 37], etc.). However, this is not necessary in the proposed approach.

3.3. From kinetic to macroscopic

Since discrete velocity approximations of the Boltzmann equation are effectively identical to the advection equation, the goal of this approach is to extend limiting techniques (such as the aforementioned maximum principle preserving limiter) from the advection equation to the Euler equations through the Boltzmann equation. The connection between the local maximum principle bounds in Eq. (29) and the Euler equations can be made through the Maxwell–Boltzmann distribution in Eq. (8), which analytically defines the distribution function as

$$f^n(\mathbf{x}_i, \mathbf{u}) = g(\mathbf{w}^n(\mathbf{x}_i)). \tag{31}$$

An example of this is shown in the schematic in Fig. 3, which showcases an example Euler solution within a DG element (which aims to local maximum principle bounds on the discrete numerical stencil) and the respective distribution function for the discrete solution nodes across that stencil.

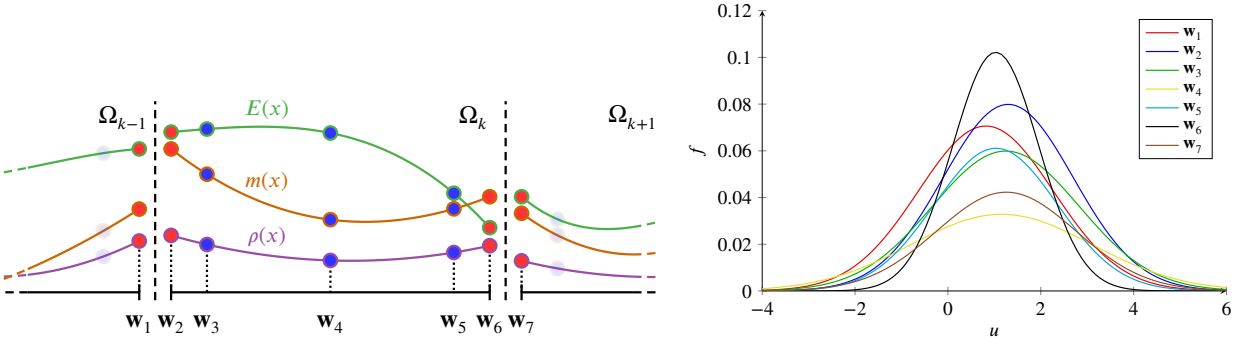


Figure 3: Example schematic for a one-dimensional Euler solution within an element Ω_k, including volume (solid blue circles) and surface (solid red circles) quadrature nodes for the local discrete stencil (denoted by $\{\mathbf{w}_1, \ldots, \mathbf{w}_7\} \in S$). Corresponding equilibrium distribution function (in the velocity domain) for each solution node shown on the right.

Given the analytic definition of the distribution function for each solution node in the stencil S, the bound for the distribution function at the next time step can be computed as

$$f^{n+1}_{\min}(\mathbf{x}_i, \mathbf{u}) = \min_{i \in S} g(\mathbf{w}^n(\mathbf{x}_i) \quad \text{and} \quad f^{n+1}_{\max}(\mathbf{x}_i, \mathbf{u}) = \max_{i \in S} g(\mathbf{w}^n(\mathbf{x}_i)), \tag{32}$$

which are piecewise defined over the velocity domain and are effectively the lower and upper bounds of the corresponding distribution functions of each solution node in the stencil S. A schematic of these distribution function bounds for the corresponding solution and stencil in Fig. 3 is shown in Fig. 4. This essentially yields an expression for the region of "admissible" distribution function values over the velocity domain, bounded from below and above by $f_{\min}$ and $f_{\max}$, respectively.

Once the admissible distribution function values are defined, it is possible to then recover the corresponding admissible macroscopic states. In particular, we are interested in finding the minimum and maximum values of each component of the macroscopic solution $[\rho, \mathbf{m}, E]^T$ that can be recovered from a distribution function f which resides in the admissible region $f_{\min} \leq f \leq f_{\max}$. For the density and energy, where the respective collision invariants are strictly non-negative (i.e., a non-negative measure), we can simply integrate the moments of the minimum/maximum distribution function to yield the minimum/maximum density and energy bounds, i.e.,

$$\rho^{n+1}_{\min} = \int_{\mathbb{R}^d} f^{n+1}_{\min} \, \mathrm{d}\mathbf{u} \quad \text{and} \quad \rho^{n+1}_{\max} = \int_{\mathbb{R}^d} f^{n+1}_{\max} \, \mathrm{d}\mathbf{u}, \tag{33}$$

$$E^{n+1}_{\min} = \int_{\mathbb{R}^d} f^{n+1}_{\min} \left(\frac{1}{2}\mathbf{u} \cdot \mathbf{u}\right) \mathrm{d}\mathbf{u} \quad \text{and} \quad E^{n+1}_{\max} = \int_{\mathbb{R}^d} f^{n+1}_{\max} \left(\frac{1}{2}\mathbf{u} \cdot \mathbf{u}\right) \mathrm{d}\mathbf{u} \, \mathrm{d}\mathbf{u}. \tag{34}$$

For momentum, where the collision invariants are not strictly non-negative, the process is slightly more involved, requiring splitting the integration over the velocity domain as

$$\mathbf{m}^{n+1}_{i,\min} = \int_{\substack{\mathbb{R}^d \\ \mathbf{u}_i < 0}} f^{n+1}_{\max} \, \mathbf{u}_i \, \mathrm{d}\mathbf{u} + \int_{\substack{\mathbb{R}^d \\ \mathbf{u}_i \geq 0}} f^{n+1}_{\min} \, \mathbf{u}_i \, \mathrm{d}\mathbf{u} \tag{35}$$

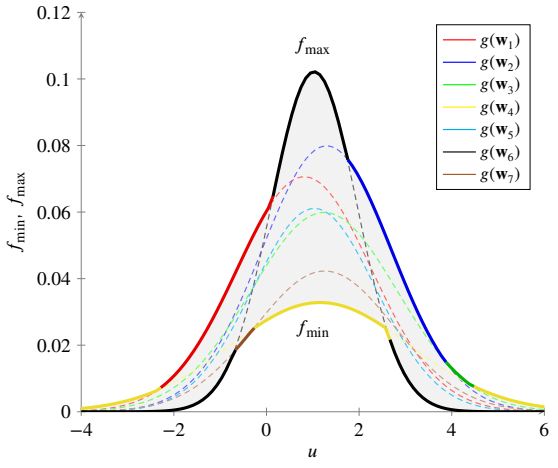


Figure 4: Example schematic for minimum and maximum bounds for the distribution function ($f_{\min}$ and $f_{\max}$, respectively) for an element Ω_k as computed by the local discrete stencil of the macroscopic solution $\{\mathbf{w}_1, \dots, \mathbf{w}_7\} \in S$ from Fig. 3. Distribution function portions highlighted by the corresponding equilibrium distribution functions. Shaded region denotes the admissible set of the distribution function values.

and

$$\mathbf{m}_{i,\max}^{n+1} = \int_{\substack{\mathbb{R}^d \\ \mathbf{u}_i<0}} f_{\min}^{n+1}\, \mathbf{u}_i \, \mathrm{d}\mathbf{u} + \int_{\substack{\mathbb{R}^d \\ \mathbf{u}_i\geq 0}} f_{\max}^{n+1}\, \mathbf{u}_i \, \mathrm{d}\mathbf{u}, \tag{36}$$

for each momentum component i. This yields a set of bounds for the conserved variables at the next time step which can be used for the limiting method of choice, which we denote by $\mathbf{w}_{i,\min}$ and $\mathbf{w}_{i,\max}$, respectively, for each solution component $1 \leq i \leq d+2$. The technique for approximating these integrals is later presented in Section 3.7.

3.4. Limiting

As the bounds for the proposed approach are given in terms of the conserved variables instead of functionals thereof (e.g., pressure, entropy, etc.), the limiting can be performed by straightforwardly treating the individual solution components as scalar local maximum principle limiting tasks. In particular, the limiting method in this work is the squeeze limiter of Zhang and Shu [30] presented in Section 2.4, which uniformly contracts the high-order solution to the element-wise mean. For the proposed approach, the limiting factor α for each solution component i can be computed as

$$\alpha_i = \min\left[1, \left|\frac{\mathbf{w}_{i,\max} - \overline{\mathbf{w}}_i}{\max\limits_{j\in V} \mathbf{w}_i(\mathbf{x}_j) - \overline{\mathbf{w}}_i}\right|, \left|\frac{\mathbf{w}_{i,\min} - \overline{\mathbf{w}}_i}{\min\limits_{j\in V} \mathbf{w}_i(\mathbf{x}_j) - \overline{\mathbf{w}}_i}\right|\right]. \tag{37}$$

We set the limiting factor explicitly as $\alpha_i = 0$ if $|\min\limits_{j\in V} \mathbf{w}_i(\mathbf{x}_j) - \overline{\mathbf{w}}_i| < \epsilon$ or $|\max\limits_{j\in V} \mathbf{w}_i(\mathbf{x}_j) - \overline{\mathbf{w}}_i| < \epsilon$, where $\epsilon = 10^{-13}$. This limiting factor is computed separately for each element in the domain.

As the bounds are decoupled by solution component, the limiting can be performed independently (where each solution component is limited by its respective limiting factor) or coupled (where each solution component is limited by the minimum of all limiting factors). For this work, we use the latter approach, such that the limited solution $\widetilde{\mathbf{w}}(\mathbf{x})$ is obtained as

$$\widetilde{\mathbf{w}}(\mathbf{x}) = \overline{\mathbf{w}} + \alpha_{\min}(\mathbf{w}(\mathbf{x}) - \overline{\mathbf{w}}), \tag{38}$$

where

$$\alpha_{\min} = \min_i \alpha_i, \tag{39}$$

although one may obtain a marginally less dissipative approach by using the former. It will later be shown in Section 4 that a limiting factor $0 \leq \alpha_{\min} \leq 1$ is guaranteed to exist for which the limited solution discretely satisfies the imposed bounds, i.e.,

$$\min_{j\in V} \widetilde{\mathbf{w}}_i(\mathbf{x}_j) \geq \mathbf{w}_{i,\min} \quad \text{and} \quad \max_{j\in V} \widetilde{\mathbf{w}}_i(\mathbf{x}_j) \leq \mathbf{w}_{i,\max} \quad \text{for} \quad 1 \leq i \leq d+2.$$

3.5. Relaxing the bounds

While strictly enforcing a local maximum principle on the high-order scheme is often enough to yield a well-behaved solution around discontinuities (although in some cases more restrictive conditions are needed, see Anderson et al. [36] for example), it is often too restrictive of a condition in smooth regions of the flow and will often degrade the accuracy of the scheme around smooth extrema, sometimes referred to as the *peak clipping* effect. This effect is seen in various limiting approaches for high-order approximations of scalar transport and gas dynamics [38], and one remedy for this effect is to marginally relax the bounds to recover high-order accuracy in smooth regions of the flow.

However, as will be shown in the numerical experiments, *the proposed approach does not appear to require relaxing the bounds to recover high-order accuracy around smooth extrema whilst simultaneously enforcing sufficient stabilization to ensure convergence around discontinuities.* Nevertheless, it can be beneficial to relax the bounds, both to allow for a slight numerical tolerance in terms of limiting and to compensate for any minor numerical integration errors in computing the bounds from the distribution function. As such, we relax the bounds in a relative manner as

$$f_{\min}^{n+1}(\mathbf{x}_i, \mathbf{u}) = \min_{i \in S} (1 - r_{\text{fac}}) g(\mathbf{w}^n(\mathbf{x}_i) \quad \text{and} \quad f_{\max}^{n+1}(\mathbf{x}_i, \mathbf{u}) = \max_{i \in S} (1 + r_{\text{fac}}) g(\mathbf{w}^n(\mathbf{x}_i)), \tag{40}$$

where the relaxation factor is set as $r_{\text{fac}} = 10^{-3}$, which is identical to relaxing $f_{\min}^{n+1}$ and $f_{\max}^{n+1}$ themselves. Similar relaxation approaches are taken in other limiting methods such as in the works of Guermond et al. [21]. The effects of enforcing bounds with and without relaxation for smooth and discontinuous solutions will later be shown in the numerical results, but, unless otherwise stated, the relaxed bounds are used in this work.

3.6. Modeling internal energy

On its own, the Maxwell–Boltzmann distribution models a monatomic particle which has d translation degrees of freedom. As such, this would constrain the specific heat ratio to $\gamma = 1 + 2/d$. To recover any arbitrary value of $\gamma \in [1, 1 + 2/d]$ and effectively model polyatomic molecules, it is necessary to include the effects of internal degrees of freedom in the distribution function. This is typically achieved through modeling an additional dimension for the internal energy domain, $\zeta \in [0, \infty)$, with the macroscopic state then computed as

$$\mathbf{w}(\mathbf{x}, t) = \int_0^\infty \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{u}, \zeta, t)\, \boldsymbol{\psi}(\mathbf{u}, \zeta)\, \mathrm{d}\mathbf{u}\, \mathrm{d}\zeta, \tag{41}$$

where the updated collision invariants are defined as $\boldsymbol{\psi}(\mathbf{u}, \zeta) := [1, \mathbf{u}, (\mathbf{u} \cdot \mathbf{u})/2 + \zeta]^T$ (see Baranger et al. [39], for example). However, naively implementing an internal energy model would add another dimension for evaluating/integrating the distribution function, which can drastically increase the computational cost. Since we are only interested in the effects of the internal energy on the macroscopic state, we can instead take a reduced distribution approach [39] by analytically integrating over the internal energy domain to compute the macroscopic state as

$$\mathbf{w}(\mathbf{x}, t) = \int_{\mathbb{R}^d} f_T(\mathbf{x}, \mathbf{u}, \zeta, t)\, [1, \mathbf{u}, (\mathbf{u} \cdot \mathbf{u})/2]^T\, \mathrm{d}\mathbf{u} + \int_{\mathbb{R}^d} f_I(\mathbf{x}, \mathbf{u}, \zeta, t)\, [0, 0, 1]^T\, \mathrm{d}\mathbf{u}, \tag{42}$$

where f_T and f_I represent the translation (monatomic) and internal energy components of the distribution function. This is further simplified through the relation that for an equilibrium distribution function,

$$g_I = \frac{\theta \delta}{2} g_T, \tag{43}$$

where δ is the number of internal degrees of freedom and θ is the scaled temperature of the macroscopic state. By setting $\delta = \frac{2}{\gamma - 1} - d$, one can recover the desired specific heat ratio γ. From a computational perspective, computing the bounds for the distribution function for a particle with internal energy effects (as opposed to the monatomic case) simply requires the addition of the $\frac{\theta \delta}{2}$ term in the collision invariant for the total energy.

3.7. Integrating the distribution function

Computing the bounds requires integrating the distribution functions $f_{\min}$ and $f_{\max}$. Although it is possible in theory to perform this integration analytically, it is not tractable to do it in any computationally feasible manner outside of simple one-dimensional cases. As such, this integration must be performed numerically via quadrature. The quadrature techniques for this approach broadly follow the techniques used in numerical approximations of the Boltzmann

equation, where similar integration over the velocity domain is necessary to recover the moments of the distribution function. However, since the distribution functions to be integrated are no longer the Maxwell–Boltzmann distribution (but a piecewise combination thereof), certain techniques which rely on its properties do not directly apply (e.g., Gauss–Hermite quadrature).

First, the infinite velocity domain can be truncated to a reduced velocity domain $\Omega_u \subset \mathbb{R}^d$. Since the Maxwell–Boltzmann distribution is a Gaussian distribution centered around $\mathbf{U}$ with a standard deviation of $\sqrt{\theta}$, one can quantify the maximum error of truncating the domain to k standard deviations around the mean (i.e., $\mathbf{U} \pm k\sqrt{\theta}$). For example, $k = 4$ and $k = 8$ bound the relative error of

$$\int_{\mathbb{R}^d} g \, \mathrm{d}\mathbf{u} - \int_{\Omega_u} g \, \mathrm{d}\mathbf{u}$$

by approximately $6 \cdot 10^{-5}$ and $1 \cdot 10^{-15}$, respectively. We use a similar approach to truncate the velocity domain, where the largest extent over the stencil (along each direction) is used, i.e.,

$$\Omega_u = \left[\min_{i \in S} \left(\mathbf{U}(\mathbf{x_i}) - k\sqrt{\theta(\mathbf{x}_i)} \right), \max_{i \in S} \left(\mathbf{U}(\mathbf{x_i}) - k\sqrt{\theta(\mathbf{x}_i)} \right) \right] \tag{44}$$

We assume that k is chosen appropriately to achieve the desired tolerance, which should be related to the relaxation factor used (i.e., k should ostensibly be chosen such that the relative error is $\lesssim r_{\text{fac}}$). For $r_{\text{fac}} = 10^{-3}$ in this work, this implies that $k \geq 3.5$, whereas without relaxation, one should integrate the bounds to machine precision (i.e, $k = 8$). We note here however that exact integration (even over the truncated domain) isn't actually necessary, and any numerical integration errors simply result in minor under/overprediction in the limiting factors without affecting the positivity-preserving properties of the approach.

With the truncated domain defined, one can introduce a quadrature operation as

$$\sum_{i=1}^{N_u} M_i f(\mathbf{u}_i)\psi(\mathbf{u}_i) \approx \int_{\mathbb{R}^d} f(\mathbf{u})\psi(\mathbf{u}) \, \mathrm{d}\mathbf{u}, \tag{45}$$

where M_i is a quadrature weight and $\mathbf{u}_i \in \Omega_u$ correspond to a set of N_u quadrature nodes in the velocity domain. A wide variety of numerical integration techniques can be used here. In this work, we particularly rely on trapezoidal rule integration over an equispaced Cartesian mesh which is commonly employed for discrete velocity approximations of the Boltzman equation. Higher-order (albeit less robust) quadrature techniques can also be used, although the piecewise nature of the underlying distribution function tends to mitigate benefits of many higher-order techniques which have an underlying assumption of the smoothness of the integrand.

3.8. Bounding density and pressure

As will be shown in Section 4, the proposed bounds guarantee that the density and pressure/internal energy will be positive for any $\mathbf{w} \in [\mathbf{w}_{\min}, \mathbf{w}_{\max}]$. However, it is often desirable to enforce some minimum threshold on the density/pressure since just strictly enforcing positivity can lead to numerical difficulties for very small values of density/pressure (e.g., ill-behaved Riemann solver, unfeasible time step restrictions). As such, we enforce a floor value of 10^{-12} on the density and pressure. For density, this entails simply replacing the minimum density bound $\rho_{\min}$ with $\max(\rho_{\min}, 10^{-12})$. For pressure, this process is slightly more involved since the proposed bounds are given in terms of the conserved variables. If the discrete limited pressure, computed as

$$P_{\min} = \min_{i \in V} P\left(\widetilde{\mathbf{w}}(\mathbf{x}_i)\right), \tag{46}$$

was below the threshold, we apply another limiting step as

$$\widetilde{\widetilde{\mathbf{w}}}(\mathbf{x}) = \overline{\mathbf{w}} + \alpha_P(\widetilde{\mathbf{w}}(\mathbf{x}) - \overline{\mathbf{w}}), \tag{47}$$

with the limiting factor computed as

$$\alpha_P = \left| \frac{10^{-12} - P(\overline{\mathbf{w}})}{P_{\min} - P(\overline{\mathbf{w}})} \right|, \tag{48}$$

The quasiconcavity of the pressure functional ensures that $\min_{i \in V} P\left(\widetilde{\widetilde{\mathbf{w}}}(\mathbf{x}_i)\right) \geq 10^{-12}$.

3.9. Overview

With the details of the proposed approach presented, we summarize here an overview of the method applied to high-order approximations of the Euler equations. This overview is shown in terms of a single temporal update $t^n \to t^{n+1}$ for an arbitrary element Ω_k, but this can directly be generalized to every element and each substage of higher-order (SSP) time integration schemes. The necessary code for the implementation of this approach is provided in the electronic supplementary material.

1. Collect the macroscopic solution $\mathbf{w}^n(\mathbf{x}_i)$ at each node in the chosen stencil S.
2. Evaluate the distribution function bounds $f_{\min}^{n+1}$ and $f_{\max}^{n+1}$ (per Eq. (32) and Eq. (40)) and the collision invariants $\boldsymbol{\psi} = [1, \mathbf{u}, \mathbf{u}\cdot\mathbf{u}/2 + \zeta]$ at each discrete velocity node $\mathbf{u}_i$. Perform a quadrature summation over the discrete velocity nodes ($1 \le i \le N_u$) to recover the macroscopic bounds per Eqs. (33) to (36).
3. Apply a minimum threshold on the density bound.
4. Advance solution in time $\mathbf{w}^n(\mathbf{x}) \to \mathbf{w}^{n+1}(\mathbf{x})$.
5. Compute the limited solution per Eq. (38).
6. Compute the minimum pressure of the limited solution per Eq. (46).
7. If minimum pressure is below the minimum threshold, apply a second limiting step per Eq. (47).
8. Replace the high-order DG solution $\mathbf{w}^{n+1}(\mathbf{x})$ with the limited solution $\widetilde{\mathbf{w}}^{n+1}(\mathbf{x})$) (or $\widetilde{\widetilde{\mathbf{w}}}^{n+1}(\mathbf{x})$).

4. Analysis

With this formulation, we now move on to state and prove the numerical properties of the proposed approach. The subsequent proofs rely on some (or all) of the following assumptions.

(i) Let there be a stencil S over which the bounds are computed for an arbitrary element Ω_k, and let G be the set of solutions with positive density $\rho > 0$ and positive pressure/internal energy $P, \rho e > 0$. The initial solution $\mathbf{w}_i \in G \; \forall \; i \in S$ at some time t^n.

(ii) The distribution functions are integrated via a discrete integration operator $\mathbf{M}$ consisting of N_u positive entries corresponding to integration over a subdomain $\Omega_u \subset \mathbb{R}^d$ for which we denote its extent as $u_{\max} = \max_{\mathbf{u}\in\Omega_u} \|\mathbf{u}\|_2$ which is assumed to be at least two thermal/absolute velocity units (i.e., $k \ge 2$). It is assumed that the integration operator can integrate moments of an arbitrary distribution function f to some desired numerical tolerance, i.e.,

$$\int_{\mathbb{R}^d} f(\mathbf{u})\psi(\mathbf{u})\,\mathrm{d}\mathbf{u} \approx \sum_{i=1}^{N_u} M_i f(\mathbf{u}_i)\psi(\mathbf{u}_i). \tag{49}$$

We denote this operation in shorthand as $\mathbf{M}\cdot f \otimes \psi(\mathbf{u})$.

(iii) The nodes of the stencil over which to compute bounds (denoted by S) and the nodes over which the limiting is performed (denoted by $V \subset S$) are chosen appropriately such that:

(a) There exists a set of positive quadrature weights c_i such that

$$\overline{\mathbf{w}} = \frac{\int_{\Omega_k} \mathbf{w}(\mathbf{x})\,\mathrm{d}\mathbf{x}}{\int_{\Omega_k} \mathrm{d}\mathbf{x}} = \frac{\sum_{i\in V} c_i \mathbf{w}(\mathbf{x}_i)}{\sum_{i\in V} c_i}.$$

(b) There exists a subset of nodes $V' \subset V$ (e.g., surface quadrature nodes) with a corresponding set of positive quadrature weights c_i' such that

$$\int_{\partial\Omega_k} \mathbf{w}(\mathbf{x})\,\mathrm{d}\mathbf{s} \approx \sum_{i\in V'} c_i' \mathbf{w}(\mathbf{x}_i).$$

(iv) The interface numerical fluxes are computed using a Rusanov-type approach [33] as

$$\overline{\mathbf{F}}(\mathbf{w}^+, \mathbf{w}^-, \mathbf{n}) = \frac{1}{2}\left[\mathbf{F}(\mathbf{w}^-) + \mathbf{F}(\mathbf{w}^+)\right] \cdot \mathbf{n} - \frac{1}{2}\lambda_{\max}(\mathbf{w}^+, \mathbf{w}^-)\left[\mathbf{w}^+ - \mathbf{w}^-\right],$$

where $\lambda_{\max}$ is an upper bound on the local maximum wavespeed of the system.

Theorem 4.1 (Constant state preservation)**.** *The proposed bounds preserve constant states in the sense that if* $\mathbf{w}(\mathbf{x}_i, t^n) = \mathbf{w}_0 \ \forall\, i \in S$ *for some constant state* $\mathbf{w}_0$*, then* $\mathbf{w}_{\min}^{n+1} = \mathbf{w}_{\max}^{n+1} = \mathbf{w}_0$.

PROOF OF THEOREM 4.1. The proof follows directly from the definition of $f_{\min}^{n+1}$ and $f_{\max}^{n+1}$ in Eq. (32), which yields

$$f_{\min}^{n+1} = f_{\max}^{n+1} = g(\mathbf{w}_0). \tag{50}$$

By integrating the moments, one yields

$$\mathbf{w}_{\min}^{n+1} = \mathbf{w}_{\max}^{n+1} = \mathbf{M} \cdot g(\mathbf{w}_0) \otimes \boldsymbol{\psi}(\mathbf{u}) \approx \int_{\mathbb{R}^d} g(\mathbf{w}_0)\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u} = \mathbf{w}_0. \tag{51}$$

Theorem 4.2 (Positivity preservation)**.** *The proposed bounds preserve positivity of density and pressure/internal energy for all* $\mathbf{w} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$.

PROOF OF THEOREM 4.2. To show positivity of density, it suffices to show that $\rho_{\min}^{n+1} > 0$, which can be expressed as the condition

$$\mathbf{M} \cdot f_{\min}^{n+1} > 0. \tag{52}$$

From Eq. (8), the distribution function $g(\mathbf{w})$ is strictly positive if $\mathbf{w} \in G$. Therefore, $f_{\min}^{n+1}(\mathbf{u}), f_{\max}^{n+1}(\mathbf{u}) > 0$ for all $\mathbf{u} \in \mathbb{R}^d$, which means that $\rho_{\min}^{n+1} > 0$.

To show positivity of pressure/internal energy for all $\mathbf{w} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$, we use the relation that any $\mathbf{w}_0 \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$ can be expressed as

$$\mathbf{w}_0 = \mathbf{M} \cdot f_0^{n+1} \otimes \boldsymbol{\psi}(\mathbf{u}) \tag{53}$$

for some distribution function $f_{\min}^{n+1}(\mathbf{u}) \le f_0^{n+1}(\mathbf{u}) \le f_{\max}^{n+1}(\mathbf{u})$, where $f_0^{n+1}(\mathbf{u}) > 0$. The internal energy can be calculated as

$$\rho e = E - \frac{1}{2}\rho \mathbf{U}\cdot\mathbf{U} = \mathbf{M} \cdot f_0^{n+1}\left(\frac{1}{2}\mathbf{u}\cdot\mathbf{u} - \frac{1}{2}\mathbf{u}\cdot\mathbf{U}\right) = \mathbf{M} \cdot \frac{1}{2} f_0^{n+1}\mathbf{c}\cdot\mathbf{c}, \tag{54}$$

where $\mathbf{c} = \mathbf{u} - \mathbf{U}$ is the peculiar velocity. From the positivity of the distribution function and measure $\mathbf{c}\cdot\mathbf{c}$, it follows that $\rho e > 0$ and, by extension, $P = (\gamma - 1)\rho e > 0$.

Theorem 4.3 (Riemann-averaged states)**.** *For any arbitrary* $\mathbf{w}_i, \mathbf{w}_j \in G$ *and corresponding bounds* $[f_{\min}, f_{\max}]$ *and* $[\mathbf{w}_{\min}, \mathbf{w}_{\max}]$*, the Riemann-averaged state*

$$\overline{\mathbf{w}}_{ij} = \frac{1}{2}\left(\mathbf{w}_i + \mathbf{w}_j\right) - \frac{1}{2\lambda_{\max}}\left(\mathbf{F}(\mathbf{w}_j) - \mathbf{F}(\mathbf{w}_i)\right) \cdot \mathbf{n}_{ij} \tag{55}$$

preserves the bounds in the sense that $\overline{\mathbf{w}}_{ij} \in [\mathbf{w}_{\min}, \mathbf{w}_{\max}]$ *for any arbitrary normal vector* $\mathbf{n}_{ij}$ *if* $\lambda_{\max}$ *is an upper bound on the local maximum wavespeed of the system.*

PROOF OF THEOREM 4.3. By using the relation that

$$\mathbf{F}(\mathbf{w}) = \int_{\mathbb{R}^d} \mathbf{u} g(\mathbf{w})\, \boldsymbol{\psi}(\mathbf{u})\, \mathrm{d}\mathbf{u}, \tag{56}$$

the Riemann-averaged state can be expressed in terms of a distribution function $\overline{g}_{ij}$ as

$$\overline{\mathbf{w}}_{ij} = \mathbf{M} \cdot \overline{g}_{ij} \otimes \boldsymbol{\psi}(\mathbf{u}) = \mathbf{M} \cdot \left[\frac{1}{2}\left(g_i + g_j\right) - \frac{1}{2\lambda_{\max}}\left(g_j - g_i\right)\mathbf{u}\cdot\mathbf{n}_{ij}\right] \otimes \boldsymbol{\psi}(\mathbf{u}). \tag{57}$$

Let $\lambda_{\max} = u_{\max} \geq \max\left(|\mathbf{U}_i \cdot \mathbf{n}_{ij}| + \sqrt{\gamma\theta_i}, |\mathbf{U}_j \cdot \mathbf{n}_{ij}| + \sqrt{\gamma\theta_j}\right)$. The Riemann-averaged distribution function can be expressed as

$$\overline{g}_{ij} = g_i\left(\frac{1}{2} + \frac{1}{2}\frac{\mathbf{u}\cdot\mathbf{n}}{\lambda_{\max}}\right) + g_j\left(\frac{1}{2} - \frac{1}{2}\frac{\mathbf{u}\cdot\mathbf{n}}{\lambda_{\max}}\right). \tag{58}$$

Since $\lambda_{\max} \geq |\mathbf{u}\cdot\mathbf{n}_{ij}|$, $\overline{g}_{ij}$ is a convex combination of the states g_i and g_j. Therefore,

$$f_{\min} = \min\left(g_i, g_j\right) \leq \overline{g}_{ij} \leq \max\left(g_i, g_j\right) = f_{\max}, \tag{59}$$

from which it directly follows that $\overline{\mathbf{w}}_{ij} \in [\mathbf{w}_{\min}, \mathbf{w}_{\max}]$.

Remark. *The proof of Theorem 4.3 is based on an upper bound on the local maximum particle speed $u_{\max}$ (on the truncated domain Ω_u) instead of the local maximum macroscopic wavespeed (e.g., the Davis [40] wavespeed estimate $\lambda_{\max} = |\mathbf{U}\cdot\mathbf{n}| + \sqrt{\gamma\theta}$). At high Mach numbers, where the thermal velocity is small compared to the flow velocity, these estimates do not drastically differ, but at low Mach numbers, the local maximum particle speed may be several times larger than the maximum macroscopic wavespeed — a factor of $k/\sqrt{\gamma}$ — which can affect the maximum admissible time step required to satisfy the bounds. However, we observed in empirical studies that setting $\lambda_{\max}$ as the local maximum macroscopic wavespeed still always ensured that the Riemann-averaged state satisfied the bounds, but we do not have a rigorous mathematical proof for this behavior.*

Theorem 4.4 (First-order scheme)**.** *For any arbitrary $\mathbf{w}_0^n, \mathbf{w}_1^n, \mathbf{w}_2^n \in G$ and corresponding bounds $[\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$, the first-order Lax–Friedrichs scheme,*

$$\mathbf{w}_1^{n+1} = \mathbf{w}_1^n - \frac{\Delta t}{\Delta x}\left(\overline{\mathbf{F}}(\mathbf{w}_1^n, \mathbf{w}_2^n) - \overline{\mathbf{F}}(\mathbf{w}_0^n, \mathbf{w}_1^n)\right), \tag{60}$$

where

$$\overline{\mathbf{F}}(\mathbf{w}_a, \mathbf{w}_b) = \frac{1}{2}\left[\mathbf{F}(\mathbf{w}_a) + \mathbf{F}(\mathbf{w}_b)\right] - \frac{1}{2}\lambda_{\max}(\mathbf{w}_a, \mathbf{w}_b)\left[\mathbf{w}_b - \mathbf{w}_a\right] \tag{61}$$

preserves bounds in the sense that $\mathbf{w}_1^{n+1} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$ under the standard Courant-–Friedrichs—Lewy (CFL) condition $\lambda_{\max}\frac{\Delta t}{\Delta x} \leq 1$.

Proof of Theorem 4.4. The first-order Lax–Friedrichs scheme can be written in terms of the Riemann-averaged states as

$$\mathbf{w}_1^{n+1} = \left(1 - \lambda_{\max}^{0,1}\frac{\Delta t}{\Delta x} - \lambda_{\max}^{1,2}\frac{\Delta t}{\Delta x}\right)\mathbf{w}_1^n + \lambda_{\max}^{0,1}\frac{\Delta t}{\Delta x}\overline{\mathbf{w}}_{01} + \lambda_{\max}^{1,2}\frac{\Delta t}{\Delta x}\overline{\mathbf{w}}_{12}, \tag{62}$$

where $\lambda_{\max}^{0,1} = \lambda_{\max}(\mathbf{w}_0, \mathbf{w}_1)$, $\lambda_{\max}^{1,2} = \lambda_{\max}(\mathbf{w}_1, \mathbf{w}_2)$, and $\mathbf{n}_{01} = \mathbf{n}_{12} = 1$. If Δt is sufficiently small such that $\lambda_{\max}\frac{\Delta t}{\Delta x} \leq 1$, where $\lambda_{\max} = \max(\lambda_{\max}^{0,1}, \lambda_{\max}^{1,2})$, then $\mathbf{w}_1^{n+1}$ is a convex combination of states in $[\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$. As a result, $\mathbf{w}_1^{n+1} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$.

Theorem 4.5 (Mean bounds preservation)**.** *For some arbitrary element Ω_k and its associated mean $\overline{\mathbf{w}}^n = \overline{\mathbf{w}}(t^n)$, there exists a sufficiently small $\Delta t > 0$ such that $\overline{\mathbf{w}}(t^n + \Delta t) = \overline{\mathbf{w}}^{n+1} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$.*

Proof of Theorem 4.5. The element-wise mean can be computed as

$$\overline{\mathbf{w}}^n = \frac{\sum_{i\in V} c_i\left[\mathbf{M}\cdot g(\mathbf{w}_i^n) \otimes \boldsymbol{\psi}(\mathbf{u})\right]}{\sum_{i\in V} c_i}, \tag{63}$$

which, by linearity of the quadrature operators, can be expressed as

$$\overline{\mathbf{w}}^n = \mathbf{M}\cdot\left[\frac{\sum_{i\in V} c_i g(\mathbf{w}_i^n)}{\sum_{i\in V} c_i}\right] \otimes \boldsymbol{\psi}(\mathbf{u}), \tag{64}$$

Since, by definition, $g(\mathbf{w}_i^n) \in [f_{\min}^{n+1}, f_{\max}^{n+1}]\ \forall\, \mathbf{u} \in \mathbb{R}^d\ \forall\, i \in V$ and $[f_{\min}^{n+1}, f_{\max}^{n+1}]$ is a convex set, the convex combination $\sum_{i\in V} c_i g(\mathbf{w}_i^n)/\sum_{i\in V} c_i \in [f_{\min}^{n+1}, f_{\max}^{n+1}]\ \forall\, \mathbf{u} \in \mathbb{R}^d\ \forall\, i \in V$. Therefore, $\overline{\mathbf{w}}^n \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$.

It now remains to be shown that $\overline{\mathbf{w}}^{n+1} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$. The temporal update of the element-wise mean can be represented as

$$\overline{\mathbf{w}}^{n+1} = \overline{\mathbf{w}}^{n} - \Delta t \sum_{i \in V'} c_i' \overline{\mathbf{F}}(\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+}, \mathbf{n}_i) = \frac{\sum_{i \in V} c_i \mathbf{w}_i^n}{\sum_{i \in V} c_i} - \Delta t \sum_{i \in V'} c_i' \overline{\mathbf{F}}(\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+}, \mathbf{n}_i), \tag{65}$$

where $\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$ for all $i \in V'$. We use the shorthand $\mathbf{w}_i$ to denote $\mathbf{w}(\mathbf{x}_i)$. $\overline{\mathbf{w}}^{n+1}$ can be equivalently expressed as a convex combination as

$$\overline{\mathbf{w}}^{n+1} = \frac{\sum_{i \in V \setminus V'} c_i \mathbf{w}_i^n + \sum_{i \in V'} c_i \mathbf{w}_i^*}{\sum_{i \in V} c_i}, \tag{66}$$

where

$$\mathbf{w}_i^* = \mathbf{w}_i^{n,-} - \Delta t \frac{c_i'}{c_i} \overline{\mathbf{F}}(\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+}, \mathbf{n}_i). \tag{67}$$

Thus, it is sufficient to show that the intermediate states $\mathbf{w}_i^* \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$ for all $i \in V'$. These states can be represented in terms of the Riemann-averaged states as

$$\mathbf{w}_i^* = [1 - C]\, \mathbf{w}_i^{n,-} + C \overline{\mathbf{w}}_{i^- i^+}, \tag{68}$$

where

$$C = \Delta t \lambda_{\max}(\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+}) \frac{c_i'}{c_i}. \tag{69}$$

It can be seen that $\mathbf{w}_i^* \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$ for any $0 \leq C \leq 1$. Note here that c_i'/c_i is $\mathcal{O}(1/h)$, where h is some local mesh scale, as it is a surface quadrature weight normalized by a volume quadrature weight. Therefore, the above condition is met under a CFL-like constraint $\Delta t \lambda_{\max}(\mathbf{w}_i^{n,-}, \mathbf{w}_i^{n,+}) \frac{c_i'}{c_i} \leq 1$, which then guarantees that $\overline{\mathbf{w}}^{n+1} \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$.

Corollary 4.5.1 (Existence of solution)**.** *If* $\overline{\mathbf{w}}(t^{n+1}) \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]$*, there exists a limiting factor* $\alpha \in [0, 1]$ *such that* $\widetilde{\mathbf{w}}(\mathbf{x}_i, t^{n+1}) \in [\mathbf{w}_{\min}^{n+1}, \mathbf{w}_{\max}^{n+1}]\ \forall\ i \in V$.

5. Implementation

The proposed limiting approach was implemented using the flux reconstruction (FR) scheme of Huynh [29], a strong-form generalization of the nodal DG method, within PyFR [41], a high-order unstructured FR solver that can efficiently target massively-parallel CPU and GPU computing architectures. To showcase generalizability with respect to solution node sets, we utilize open solution/flux node sets (i.e., where the interface flux nodes are not a subset of the interior solution nodes), namely the Gauss–Legendre/Gauss–Legendre solution/flux nodes for line/quadrilateral elements and Williams–Shunn/Gauss–Legendre solution/flux nodes for triangular elements [42]. Temporal integration was performed using a three-stage, third-order strong stability preserving Runge–Kutta scheme (SSP-RK3 [43]), with limiting performed after each substage of the scheme. Common interface numerical fluxes were computed using the HLLC Riemann solver of Toro et al. [44]. We note here that while the numerical analysis in this work only presents some theoretical guarantees for Rusanov-type Riemann solvers, it was found empirically in the numerical experiments that many other positivity-preserving entropy-stable Riemann solvers always preserved the bounds on the element-wise mean (under the standard assumptions), such that less dissipative approaches like HLLC were found to be more performant.

The bounds were computed over the numerical stencil S, consisting of the interior solution and flux nodes as well as the exterior (face-adjacent) flux nodes, and limiting was performed over the element stencil V, consisting of just the interior solution and flux nodes, as shown in Fig. 2b. The distribution functions were integrated over the velocity domain using uniform trapezoidal rule quadrature with the domain extent set to four thermal/absolute velocity units (i.e., $k = 4$). Similar accuracy was observed with more advanced quadrature techniques such as Romberg integration. For the one- and two-dimensional experiments to be presented, the velocity domain resolution was fixed separately.

For one-dimensional problems, we purposely use a large highly-resolved velocity domain ($k = 8$, $N_u = 2048$) which can integrate the bounds to essentially near machine precision levels so that we can effectively compare the accuracy of the approach using strict and relaxed bounds without introducing any implicit relaxation in the bounds due to numerical under-integration of the distribution functions. For two-dimensional problems, we use a more appropriate resolution of $k = 4$, $N_u = 32^2$ for every problem, although we note that similar results could be obtained with a lower resolution of $N_u = 16^2$. In theory, one can use more efficient adaptive quadrature techniques to yield a more performant algorithm since the piecewise nature of the distribution function gives *a priori* information about the required local resolution, but this is not explored in this work as its primary purpose is to evaluate the accuracy of the proposed approach.

6. Results

6.1. Sod shock tube

N	$\mathbb{P}_2$	Rate	$\mathbb{P}_3$	Rate	$\mathbb{P}_4$	Rate	$\mathbb{P}_5$	Rate
25	1.258×10^{-2}	-	9.788×10^{-3}	-	1.222×10^{-2}	-	1.196×10^{-2}	-
50	4.513×10^{-3}	1.479	4.392×10^{-3}	1.156	4.621×10^{-3}	1.402	4.921×10^{-3}	1.282
100	2.424×10^{-3}	0.897	2.178×10^{-3}	1.012	2.443×10^{-3}	0.920	2.503×10^{-3}	0.975
200	1.259×10^{-3}	0.945	1.181×10^{-3}	0.884	1.359×10^{-3}	0.846	1.576×10^{-3}	0.667
400	6.227×10^{-4}	1.016	5.747×10^{-4}	1.039	7.750×10^{-4}	0.810	8.608×10^{-4}	0.873
800	3.287×10^{-4}	0.922	2.886×10^{-4}	0.994	4.877×10^{-4}	0.668	4.786×10^{-4}	0.847
1600	1.731×10^{-4}	0.925	1.469×10^{-4}	0.974	2.961×10^{-4}	0.720	2.911×10^{-4}	0.717
3200	9.132×10^{-5}	0.923	7.815×10^{-5}	0.911	1.819×10^{-4}	0.703	1.677×10^{-4}	0.795
		0.987		0.989		0.834		0.851

Table 1: Convergence of the L^1 norm of the density error for the Sod shock tube problem at $t = 0.2$ computed using varying approximation orders and mesh resolution *with relaxed bounds*. Average rate of convergence shown on bottom.

N	$\mathbb{P}_2$	Rate	$\mathbb{P}_3$	Rate	$\mathbb{P}_4$	Rate	$\mathbb{P}_5$	Rate
25	1.273×10^{-2}	-	1.030×10^{-2}	-	1.231×10^{-2}	-	1.309×10^{-2}	-
50	4.602×10^{-3}	1.468	4.879×10^{-3}	1.078	5.212×10^{-3}	1.240	5.602×10^{-3}	1.224
100	2.496×10^{-3}	0.882	2.475×10^{-3}	0.979	2.902×10^{-3}	0.845	3.104×10^{-3}	0.852
200	1.314×10^{-3}	0.926	1.403×10^{-3}	0.819	1.715×10^{-3}	0.759	1.979×10^{-3}	0.650
400	6.586×10^{-4}	0.996	7.438×10^{-4}	0.916	9.820×10^{-4}	0.804	1.130×10^{-3}	0.808
800	3.529×10^{-4}	0.900	4.100×10^{-4}	0.859	5.689×10^{-4}	0.788	6.749×10^{-4}	0.744
1600	1.887×10^{-4}	0.903	2.349×10^{-4}	0.804	3.292×10^{-4}	0.789	4.132×10^{-4}	0.708
3200	1.001×10^{-4}	0.914	1.424×10^{-4}	0.722	1.936×10^{-4}	0.766	2.593×10^{-4}	0.673
		0.969		0.879		0.830		0.784

Table 2: Convergence of the L^1 norm of the density error for the Sod shock tube problem at $t = 0.2$ computed using varying approximation orders and mesh resolution *without relaxed bounds*. Average rate of convergence shown on bottom.

An initial evaluation of the shock capturing and discontinuity resolving capabilities of the proposed approach was performed with the Sod shock tube [45]. This canonical gas dynamics test case consists of the three main features of the Riemann problem: shock waves, rarefaction waves, and contact discontinuities. The domain was set as $\Omega = [0, 1]$ with Dirichlet boundary conditions, and the initial conditions were set as

$$\mathbf{q}(x, 0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 0.5, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{where} \quad \mathbf{q}_l = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 0.125 \\ 0 \\ 0.1 \end{bmatrix}. \tag{70}$$

The problem was solved using varying approximation orders and mesh resolution, and the convergence of the density error with respect to the exact solution was computed, calculated as the point-mean L^1 norm on the solution nodes. The errors and rates are shown in Table 1 as computed with the *relaxed* bounds. It can be seen that the optimal rate of unity was nearly achieved at all approximation orders. A marginally higher convergence rate was observed at lower approximation orders ($\mathbb{P}_2$ and $\mathbb{P}_3$) than at higher approximation orders ($\mathbb{P}_4$ and $\mathbb{P}_5$), which is consistent with

observations that element-wise local maximum principle limiters are better behaved at lower approximation orders [36]. The experiments were also repeated with *strict* bounds, the results of which are tabulated in Table 2. It can be seen that very similar error magnitudes were obtained with and without relaxing the bounds, with the relaxed bounds even achieving marginally higher convergence rates.

6.2. Smooth convecting density pulse

N	$\mathbb{P}_2$	Rate	$\mathbb{P}_3$	Rate	$\mathbb{P}_4$	Rate	$\mathbb{P}_5$	Rate
10	1.663×10^{-1}	-	3.622×10^{-2}	-	5.097×10^{-3}	-	6.811×10^{-4}	-
15	6.048×10^{-2}	2.495	4.934×10^{-3}	4.917	4.187×10^{-4}	6.164	4.657×10^{-5}	6.617
20	2.605×10^{-2}	2.927	1.346×10^{-3}	4.515	8.759×10^{-5}	5.439	4.800×10^{-6}	7.898
25	1.014×10^{-2}	4.227	4.821×10^{-4}	4.601	2.517×10^{-5}	5.588	1.958×10^{-6}	4.020
30	5.177×10^{-3}	3.689	1.925×10^{-4}	5.036	8.523×10^{-6}	5.940	6.527×10^{-7}	6.025
35	2.551×10^{-3}	4.590	1.008×10^{-4}	4.196	4.250×10^{-6}	4.514	2.559×10^{-7}	6.075
40	1.435×10^{-3}	4.311	5.225×10^{-5}	4.920	2.331×10^{-6}	4.499	1.021×10^{-7}	6.877
		3.469		4.696		5.570		6.248

Table 3: Convergence of the L^∞ norm of the density error for the convecting density pulse problem at $t = 1$ computed using varying approximation orders and mesh resolution *with relaxed bounds*. Average rate of convergence shown on bottom.

N	$\mathbb{P}_2$	Rate	$\mathbb{P}_3$	Rate	$\mathbb{P}_4$	Rate	$\mathbb{P}_5$	Rate
10	1.663×10^{-1}	-	3.622×10^{-2}	-	5.098×10^{-3}	-	6.812×10^{-4}	-
15	6.048×10^{-2}	2.495	4.934×10^{-3}	4.917	4.187×10^{-4}	6.164	4.650×10^{-5}	6.621
20	2.605×10^{-2}	2.927	1.346×10^{-3}	4.515	8.773×10^{-5}	5.433	4.752×10^{-6}	7.928
25	1.014×10^{-2}	4.227	4.821×10^{-4}	4.602	2.503×10^{-5}	5.620	1.958×10^{-6}	3.974
30	5.177×10^{-3}	3.689	1.926×10^{-4}	5.033	8.401×10^{-6}	5.988	5.215×10^{-7}	7.255
35	2.551×10^{-3}	4.590	1.007×10^{-4}	4.202	4.155×10^{-6}	4.568	2.559×10^{-7}	4.620
40	1.435×10^{-3}	4.311	5.231×10^{-5}	4.908	2.183×10^{-6}	4.821	1.021×10^{-7}	6.877
		3.469		4.695		5.605		6.290

Table 4: Convergence of the L^∞ norm of the density error for the convecting density pulse problem at $t = 1$ computed using varying approximation orders and mesh resolution *without relaxed bounds*. Average rate of convergence shown on bottom.

To verify that the proposed approach retains the high-order accuracy of the underlying DG scheme even at smooth extrema, a smooth convecting density pulse was simulated. The problem was evaluated on the periodic domain $\Omega = [-0.5, 0.5]$, and the initial conditions are given as

$$\rho = 1 + \exp(-\sigma x^2) \tag{71a}$$

$$u = 1, \tag{71b}$$

$$P = 1, \tag{71c}$$

where $\sigma = 500$ is the strength of the Gaussian wave. This problem consists of a density perturbation convecting through the domain at constant velocity. To evaluate the behavior of the limiting approach at smooth extrema, the L^∞ norm of the density error was calculated at the solution nodes after one flow-through of the domain ($t = 1$). The error and convergence rates at varying approximation orders and mesh resolution are shown in Table 3 and Table 4 for the limiting with and without relaxing the bounds, respectively. It can be seen that the approach recovers the expected high-order (i.e., approximately $p + 1$) convergence rates for all approximation orders. Most importantly, it can also be seen that strictly enforcing the bounds does not degrade the accuracy of the scheme at smooth extrema, indicating that the proposed approach is not overly restrictive in smooth regions of the flow. This is in contrast to many limiting techniques which require ad hoc corrections to recover high-order accuracy at smooth extrema.

6.3. Double expansion wave

The positivity-preserving properties of the proposed approach at near-vacuum conditions were then evaluated through the double expansion wave problem. This problem, known as the *123 problem* (or Test 2) in Toro [32],

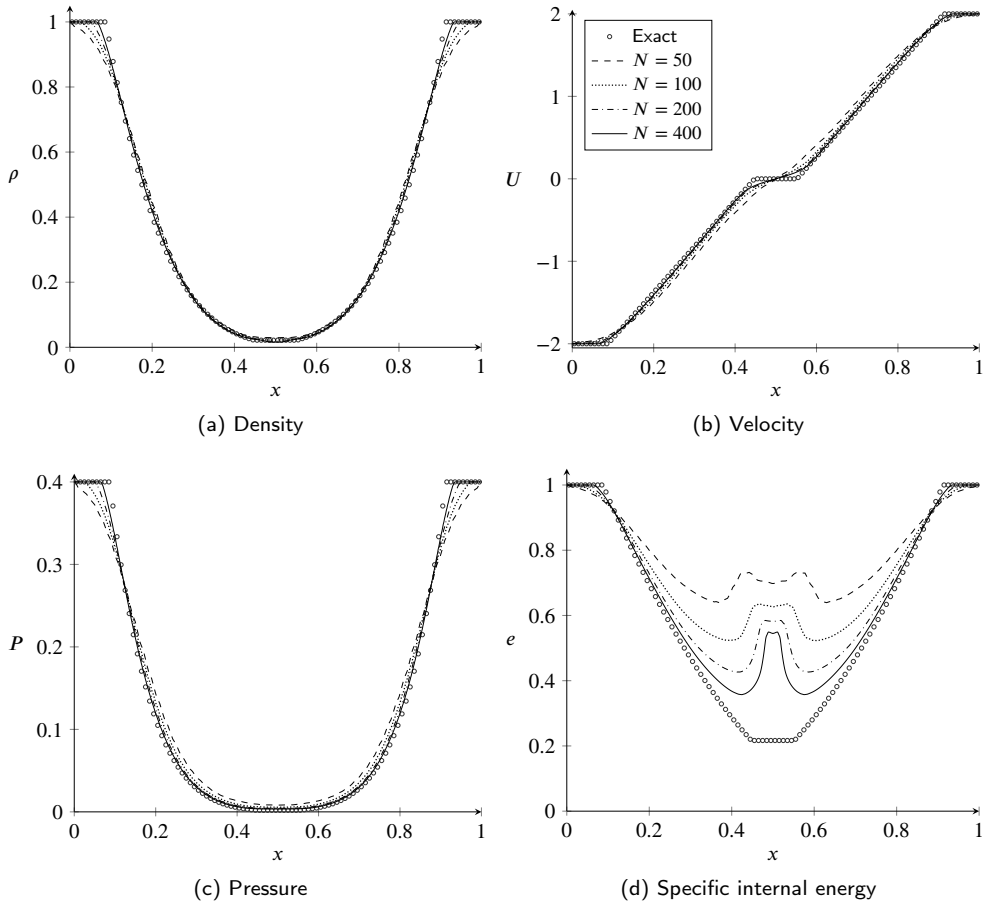


Figure 5: Density (top left), velocity (top right), pressure (bottom left), and specific internal energy (bottom right) profiles for the double expansion wave problem at $t = 0.15$ computed using a $\mathbb{P}_2$ approximation with varying mesh resolution.

was solved on the domain $\Omega = [0, 1]$ with Dirichlet boundary conditions, and the initial conditions were set as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_L, & \text{if } x \leq 0.5, \\ \mathbf{q}_R, & \text{else,} \end{cases} \qquad \text{where} \quad \mathbf{q}_L = \begin{bmatrix} 1 \\ -2 \\ 0.4 \end{bmatrix}, \quad \mathbf{q}_R = \begin{bmatrix} 1 \\ 2 \\ 0.4 \end{bmatrix}. \tag{72}$$

This Riemann problem develops into two outrunning expansion waves, causing a low density and pressure region in the center. The resulting symmetric wave structure and the near-vacuum center region are notably difficult to resolve, with even robust first-order schemes often performing poorly or failing altogether.

The problem was solved using a $\mathbb{P}_2$ approximation with varying mesh resolution between $N = 50$ and $N = 400$ elements. The profiles of density, velocity, pressure, and specific internal energy at $t = 0.15$ are shown in Fig. 5. The predicted solution profiles showed clear qualitative convergence to the exact solution with increasing resolution, even in the near-vacuum center region. The specific internal energy profile showed a distinct spike in the center, a very commonly encountered numerical artifact in this problem indicative of spurious physical entropy generation in the center [44]. However, as expected, this spike decreased with increasing resolution, also converging toward the exact solution.

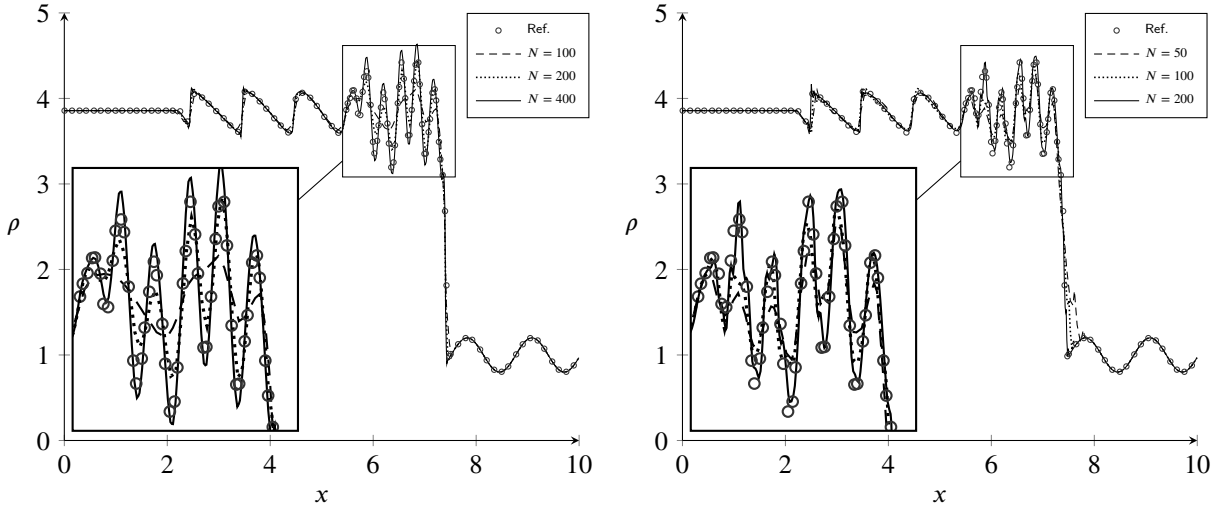
6.4. Shu–Osher problem

The Shu–Osher problem [46] was then simulated, which consists of more complex flow behavior including shock waves and oscillatory small-scale features. The domain was set as $\Omega = [0, 10]$ with Dirichlet boundary conditions,

and the initial conditions were set as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 1, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{where} \quad \mathbf{q}_l = \begin{bmatrix} 3.857143 \\ 2.629369 \\ 10.333333 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1 + 0.2\sin 5x \\ 0 \\ 1 \end{bmatrix}. \tag{73}$$

The problem consists of a shock wave interacting with a sinusoidally-perturbed density field which induces instabilities in the flow field that may be erroneously damped by overly dissipative shock capturing schemes.



(a) $\mathbb{P}_1$ (b) $\mathbb{P}_3$

Figure 6: Density profile for the Shu–Osher problem at $t = 1.8$ computed using a $\mathbb{P}_1$ (left) and $\mathbb{P}_3$ (right) approximation with varying mesh resolution.

To explore the efficacy of the proposed approach at both low and high orders, the problem was simulated with both a $\mathbb{P}_1$ and $\mathbb{P}_3$ approximation. A reference solution was computed using a highly-resolved exact Godunov-type solver [34]. The density profiles at $t = 1.8$ as computed by a $\mathbb{P}_1$ and $\mathbb{P}_3$ approximation are shown in Fig. 6 in comparison to the reference simulation. It can be seen that the solution was well-behaved around the shock at both approximation orders and various mesh resolutions, with minimal spurious oscillations. Furthermore, good resolution of the oscillatory density field was obtained, with clear qualitative convergence to the exact solution with increasing resolution. As expected, the $\mathbb{P}_1$ was better behaved, with fewer spurious oscillations. However, the $\mathbb{P}_3$ approximation showed better resolution of the oscillatory flow features at lower mesh resolution, indicating that the high-order accuracy of the underlying scheme was sufficiently retained in smooth regions.

6.5. Leblanc shock tube

As a final one-dimensional numerical experiment, the Leblanc shock tube was simulated to validate the approach for extreme shocks and flow conditions. This notoriously difficult test case consists of pressure ratios of $\mathcal{O}(10^9)$ and is extremely challenging for numerical schemes to accurately and robustly resolve [47]. The problem was solved on the domain $\Omega = [0, 9]$ with Dirichlet boundary conditions, and the initial conditions were set as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 3, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{where} \quad \mathbf{q}_l = \begin{bmatrix} 1 \\ 0 \\ (\gamma-1)10^{-1} \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 10^{-3} \\ 0 \\ (\gamma-1)10^{-10} \end{bmatrix}. \tag{74}$$

This problem consists of a right-running strong shock and contact discontinuity as well as a left-running expansion wave. Due to the large density jump at the contact discontinuity, which can be smeared by numerical dissipation, it is common for numerical approximations to overpredict the specific internal energy at the contact, which, in turn, overpredicts the shock speed. As such, fine mesh resolution is often required to accurately predict the shock location [47].

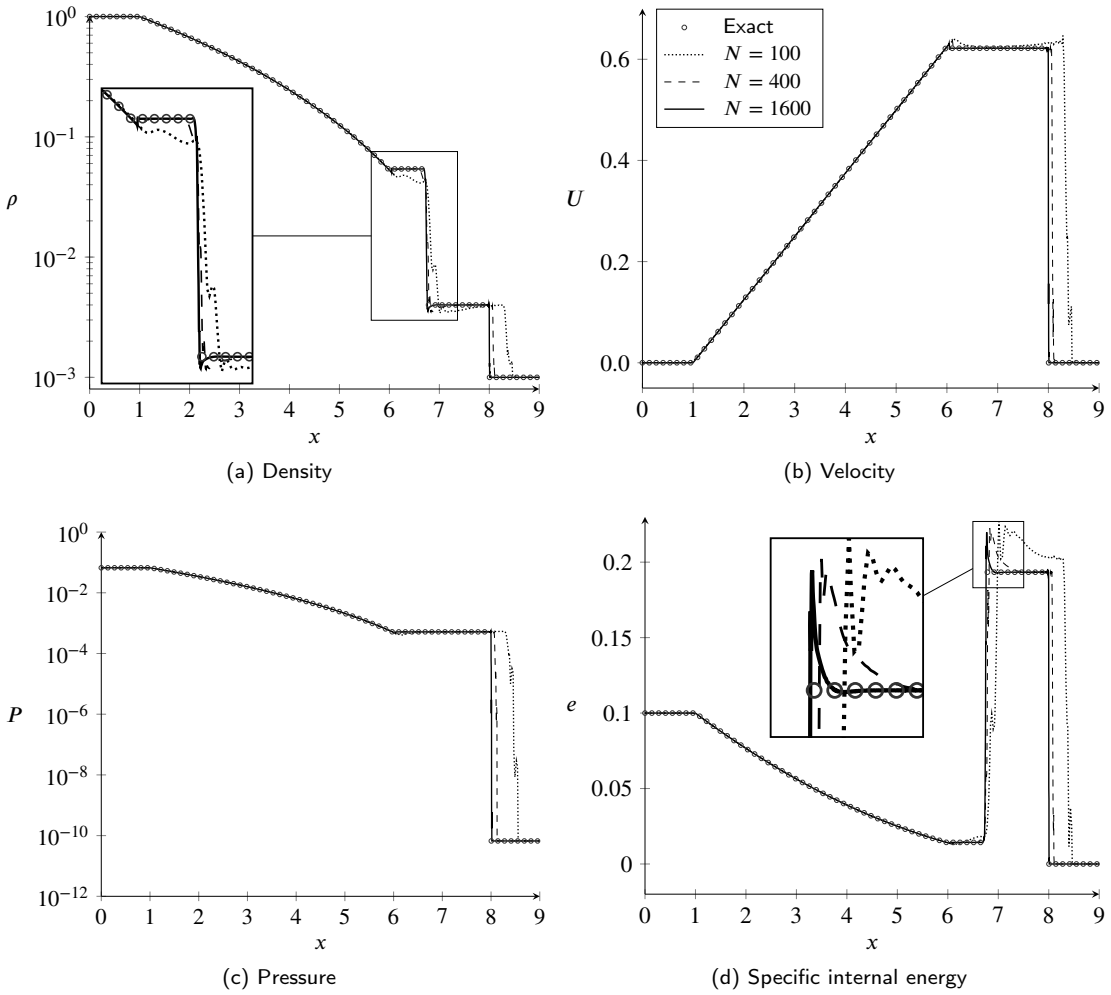


Figure 7: Density (top left), velocity (top right), pressure (bottom left), and specific internal energy (bottom right) profiles for the Leblanc shock tube problem at $t = 6$ computed using a $\mathbb{P}_3$ approximation with varying mesh resolution.

The problem was solved using a $\mathbb{P}_3$ approximation with various mesh resolutions ranging from $N = 100$ to $N = 1600$. The predicted density, velocity, pressure, and specific internal energy fields at $t = 6$ are shown in Fig. 7. At all resolutions, the proposed limiting approach effectively stabilized the high-order solution around very strong discontinuities. At $N = 100$, the shock speed was noticeably overpredicted, and some minor oscillations were observed around the shock and contact. However, at $N = 400$, a notably accurate solution was obtained despite the relatively coarse resolution, with only a minor overprediction of the shock speed and specific internal energy at the contact. At $N = 1600$, excellent agreement with the exact solution was obtained.

Since the problem has an exact solution, the convergence behavior of the proposed approach was evaluated for cases with extreme discontinuities. The error of various solution quantities was computed, calculated as the point-mean L^1 norm on the solution nodes normalized by the domain size, including the conserved variables (ρ, m, E) and primitive variables (u, P) as well as non-quasiconcave functionals (e). Furthermore, the physical entropy $s = P\rho^{-\gamma}$ was also computed. The error and convergence rates for these quantities are shown in Table 5 for the approach *with relaxed bounds*. It can be seen that the optimal rate of unity was achieved for all quantities. Furthermore, the error and convergence rates for the approach *without relaxed bounds* are shown in Table 6. Similarly to the Sod shock tube, marginally better error and convergence rates were observed with the relaxed bounds in comparison to the strict bounds.

N	ρ	Rate	m	Rate	E	Rate	U	Rate
50	3.925×10^{-3}	-	1.685×10^{-3}	-	6.329×10^{-4}	-	4.179×10^{-2}	-
100	2.215×10^{-3}	0.825	9.243×10^{-4}	0.866	3.595×10^{-4}	0.816	2.996×10^{-2}	0.480
200	1.122×10^{-3}	0.982	5.062×10^{-4}	0.869	1.880×10^{-4}	0.935	1.322×10^{-2}	1.180
400	5.471×10^{-4}	1.036	2.269×10^{-4}	1.157	8.839×10^{-5}	1.088	7.492×10^{-3}	0.819
800	2.788×10^{-4}	0.972	1.257×10^{-4}	0.853	4.683×10^{-5}	0.916	3.456×10^{-3}	1.116
1600	1.441×10^{-4}	0.953	6.126×10^{-5}	1.036	2.354×10^{-5}	0.992	1.896×10^{-3}	0.866
3200	7.047×10^{-5}	1.032	3.164×10^{-5}	0.953	1.180×10^{-5}	0.996	8.490×10^{-4}	1.159
N	P	Rate	e	Rate	s	Rate		
50	1.755×10^{-4}	-	2.009×10^{-2}	-	6.250×10^{-1}	-		
100	1.020×10^{-4}	0.783	1.464×10^{-2}	0.456	4.595×10^{-1}	0.444		
200	4.714×10^{-5}	1.113	6.995×10^{-3}	1.066	2.204×10^{-1}	1.060		
400	2.514×10^{-5}	0.907	3.754×10^{-3}	0.898	1.188×10^{-1}	0.891		
800	1.186×10^{-5}	1.084	1.876×10^{-3}	1.001	5.928×10^{-2}	1.003		
1600	6.309×10^{-6}	0.911	1.021×10^{-3}	0.878	3.200×10^{-2}	0.889		
3200	3.025×10^{-6}	1.060	4.759×10^{-4}	1.101	1.493×10^{-2}	1.100		

Table 5: Convergence of the L^1 error of the density, momentum, total energy, velocity, pressure, specific internal energy, and physical entropy for the Leblanc shock tube problem at $t = 6$ computed using a $\mathbb{P}_3$ approximation *with relaxed bounds*.

N	ρ	Rate	m	Rate	E	Rate	U	Rate
50	3.892×10^{-3}	-	1.679×10^{-3}	-	6.303×10^{-4}	-	4.179×10^{-2}	-
100	2.198×10^{-3}	0.824	9.214×10^{-4}	0.866	3.584×10^{-4}	0.815	3.000×10^{-2}	0.478
200	1.140×10^{-3}	0.947	5.131×10^{-4}	0.845	1.916×10^{-4}	0.903	1.324×10^{-2}	1.180
400	6.058×10^{-4}	0.912	2.478×10^{-4}	1.050	9.849×10^{-5}	0.960	7.546×10^{-3}	0.811
800	3.276×10^{-4}	0.887	1.423×10^{-4}	0.801	5.498×10^{-5}	0.841	3.490×10^{-3}	1.112
1600	1.844×10^{-4}	0.829	7.491×10^{-5}	0.925	3.030×10^{-5}	0.860	1.938×10^{-3}	0.849
3200	9.658×10^{-5}	0.933	4.054×10^{-5}	0.886	1.616×10^{-5}	0.907	8.740×10^{-4}	1.149
N	P	Rate	e	Rate	s	Rate		
50	1.733×10^{-4}	-	2.007×10^{-2}	-	6.245×10^{-1}	-		
100	1.012×10^{-4}	0.776	1.463×10^{-2}	0.456	4.584×10^{-1}	0.446		
200	4.963×10^{-5}	1.028	6.987×10^{-3}	1.066	2.197×10^{-1}	1.061		
400	3.144×10^{-5}	0.659	3.781×10^{-3}	0.886	1.194×10^{-1}	0.880		
800	1.724×10^{-5}	0.867	1.872×10^{-3}	1.014	5.877×10^{-2}	1.022		
1600	1.077×10^{-5}	0.678	1.015×10^{-3}	0.883	3.154×10^{-2}	0.898		
3200	5.858×10^{-6}	0.879	4.792×10^{-4}	1.083	1.490×10^{-2}	1.082		

Table 6: Convergence of the L^1 error of the density, momentum, total energy, velocity, pressure, specific internal energy, and physical entropy for the Leblanc shock tube problem at $t = 6$ computed using a $\mathbb{P}_3$ approximation *without relaxed bounds*.

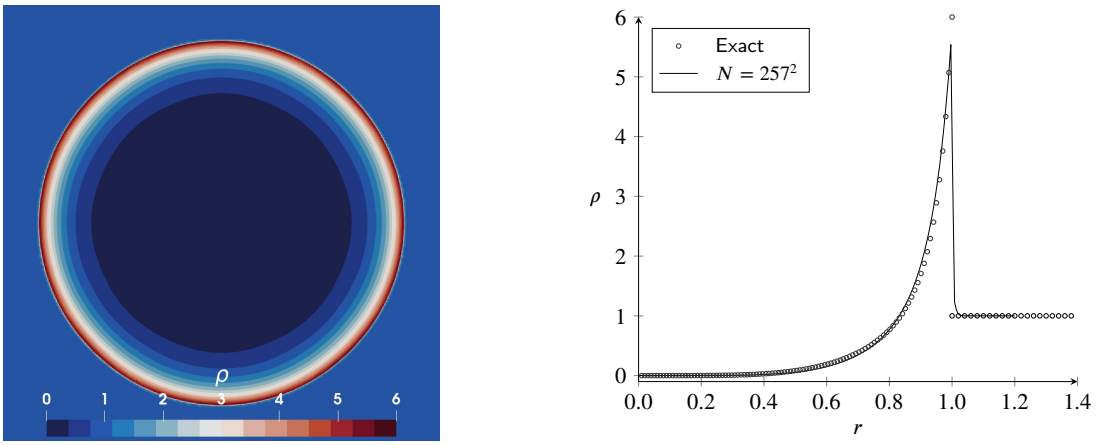
6.6. Sedov blast wave

The extension to two-dimensional problems was first performed with the Sedov blast wave problem [48], which consists of an energy point-source in an ambient gas that drives an outward-running radial explosion. The domain was set as $\Omega = [-1.2, 1.2]^2$ with Dirichlet boundary conditions, and the initial conditions were set to consist of an ambient gas with a constant density $\rho_0 = 1$ and velocity $\mathbf{U} = \mathbf{0}$. We take an identical problem setup to Maire [49] by setting the pressure to an ambient value of $P_a = 10^{-6}$ everywhere except in the element centered at $x = y = 0$, where an overpressure value is set in terms of the internal energy e_0 and cell volume V_0 as

$$P_0 = 4(\gamma - 1)\rho_0 e_0 / V_0. \tag{75}$$

Here, $V_0 = \Delta x \Delta y = 2.4^2/N$ is the element volume and $\epsilon_0 = 0.244816$ is the initial specific internal energy. The factor of 4 in the initial pressure is used to account for the fact that the entire domain is being simulated instead of one quadrant.

The problem was solved using a $\mathbb{P}_2$ approximation with $N = 257^2$ quadrilateral elements. The contours of density as well as the azimuthally-averaged density profile at $t = 1$ are shown in Fig. 8. We observe the canonical blast wave



(a) Contours (b) Profiles

Figure 8: Contours of density (left) and azimuthally-averaged density profile for the two-dimensional Sedov blast problem at $t = 1$ computed using a $\mathbb{P}_2$ approximation and $N = 257^2$ elements.

structure in the solution, with a low density center region and a sharp outward-running shock front. Furthermore, good agreement with the exact solution was obtained, with sharp resolution of the shock and minimal spurious oscillations.

6.7. Double Mach reflection

The double mach reflection problem of Woodward and Colella [50] was subsequently used to evaluate the proposed approach in the context of strong shocks with small-scale flow features. The problem consists of a Mach 10 shock impinging on a 30 degree ramp, which results in multiple strong shock interactions as well as the onset of small-scale Kelvin–Helmholtz instabilities. The problem was solved on the domain $\Omega = [0, 4] \times [0, 1]$, and the initial conditions were set as

$$\mathbf{q}(\mathbf{x}, 0) = \begin{cases} \mathbf{q}_l, & \text{if } x < 1/6 + \tan(30°)y, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{where} \quad \mathbf{q}_l = \begin{bmatrix} 8 \\ 7.14471 \\ -4.125 \\ 116.5 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1.4 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Dirichlet boundary conditions were primarily used, with the post-shock state $\mathbf{q}_l$ at the left boundary and the pre-shock state $\mathbf{q}_r$ for the right boundary. At the bottom boundary, the post-shock state $\mathbf{q}_l$ was set for $x < 1/6$ and a slip adiabatic wall boundary condition was applied for $x \geqslant \frac{1}{6}$. For the top boundary, the exact solution was set, given as

$$\mathbf{q}(\mathbf{x}, t)|_{y=1} = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 1/6 + \tan(30°)y + \frac{10}{\cos(30°)}t, \\ \mathbf{q}_r, & \text{else.} \end{cases}$$

The problem was computed using a $\mathbb{P}_3$ approximation with $N = 960 \times 240$ quadrilateral elements. The contours of density at $t = 0.2$ are shown in Fig. 9. Excellent resolution of the strong shocks in the flow was observed, with effectively sub-element resolution of discontinuities without noticeable spurious oscillations. Furthermore, the onset of the Kelvin–Helmholtz instabilities rolling up off the contact line was well-resolved by the approach, indicating that the limiter does not erroneously dissipate small-scale flow features.

6.8. Richtmyer–Meshkov instability

To further evaluate the ability of the approach in resolving small-scale vortical flow features in the vicinity of discontinuities, the Richtmyer–Meshkov instability was simulated. This instability, predicted analytically by Richtmyer [51] and shown experimentally by Meshkov [52], stems from an impulse acting upon a contact discontinuity, generally as a result of a propagating shock wave. The problem was solved on the domain $\Omega = [0, 0.5] \times [0, 4]$, and the initial

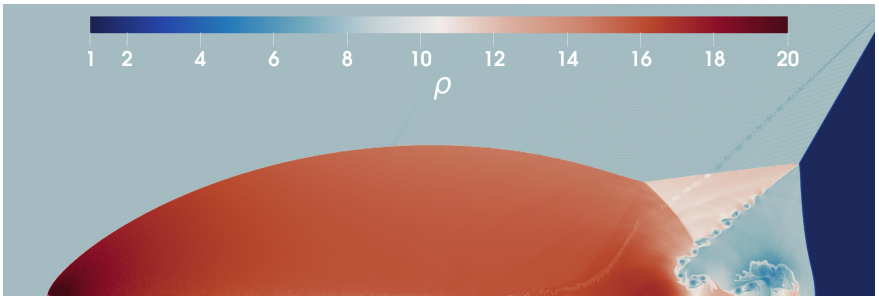
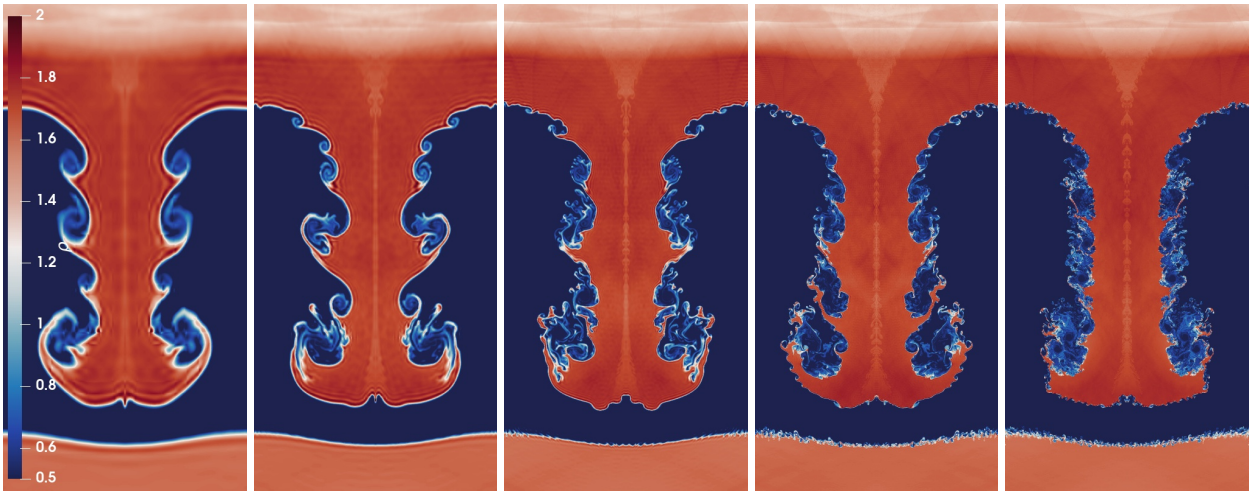


Figure 9: Contours of density for the Double Mach reflection problem at $t = 0.2$ on the subregion $0 \leq x \leq 3$ computed using a $\mathbb{P}_3$ approximation with $N = 960 \times 240$ elements.



(a) $N = 10 \times 80$ (b) $N = 20 \times 160$ (c) $N = 40 \times 320$ (d) $N = 80 \times 640$ (e) $N = 160 \times 1280$

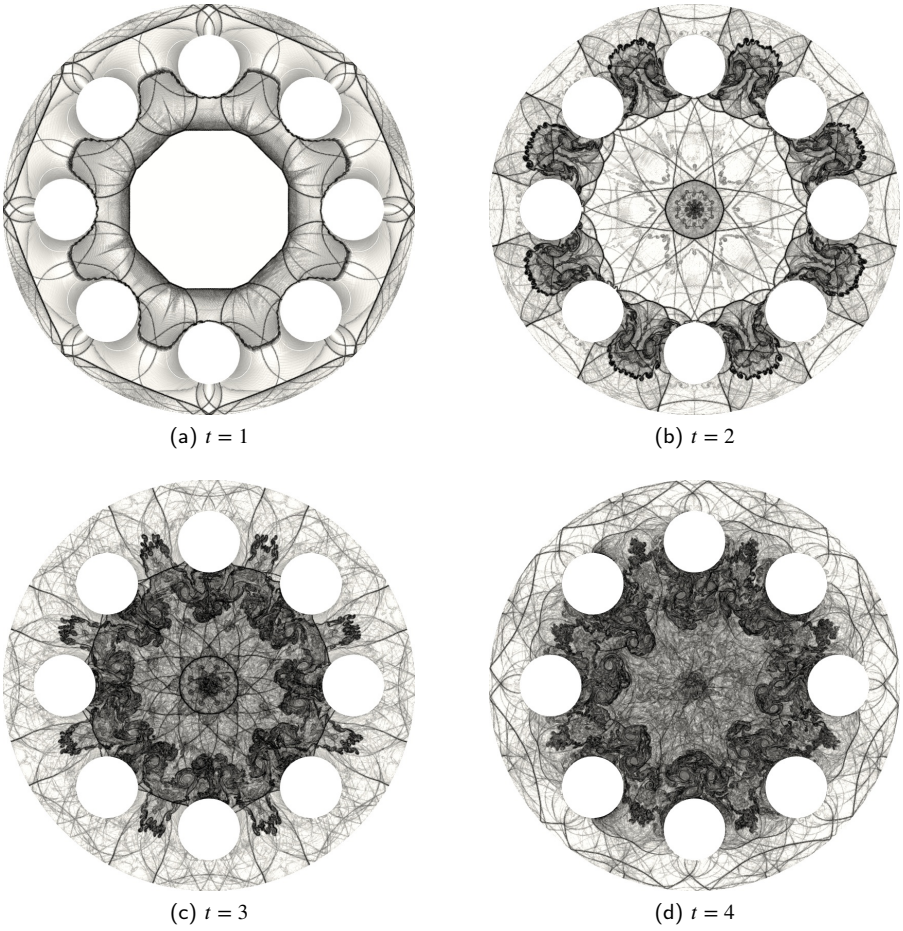
Figure 10: Contours of density for the Richtmyer–Meshkov instability problem at $t = 2$ on the subregion $2 \leq y \leq 4$ computed using a $\mathbb{P}_4$ approximation with varying mesh resolution. Contours are reflected about the y-axis for visualization.

conditions were set as

$$\mathbf{q}(\mathbf{x}, 0) = \begin{cases} \mathbf{q}_t, & \text{if } y \geq 2 - \frac{1}{2}\cos(2\pi x), \\ \mathbf{q}_m, & \text{if } \frac{1}{2} \leq y < 2 - \frac{1}{2}\cos(2\pi x), \\ \mathbf{q}_b, & \text{else,} \end{cases} \quad \text{where} \quad \mathbf{q}_b = \begin{bmatrix} 0.25 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_m = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_t = \begin{bmatrix} 8/3 \\ 0 \\ 0 \\ 4.5 \end{bmatrix}.$$

These conditions yield an Atwood number of 0.6. Slip adiabatic wall boundary conditions were enforced along the left and right boundaries (mimicking symmetry) and Dirichlet boundary conditions were enforced along the top and bottom boundaries.

The problem was computed using a $\mathbb{P}_4$ approximation with progressively finer mesh resolution, starting with $N = 10 \times 80$ quadrilateral elements and increasing to $N = 160 \times 1280$ quadrilateral elements. The density contours at $t = 2$ as computed on the various meshes are shown in Fig. 10. Good resolution of the discontinuities in the flow was observed even with the higher-order approximation, and the rollup of the vortices was effectively captured by the approach with increasingly smaller-scale vortical structures appearing with higher mesh resolution. These results indicate that the low numerical dissipation and scale-resolving capabilities of the underlying high-order DG scheme are not noticeably degraded with the proposed limiter.



(a) $t = 1$ (b) $t = 2$

(c) $t = 3$ (d) $t = 4$

Figure 11: Schlieren-type representation of the density gradient norm for the obstructed radial explosion problem at varying times computed using a $\mathbb{P}_2$ approximation with 9.68×10^5 unstructured/curved triangular elements.

6.9. Obstructed radial explosion

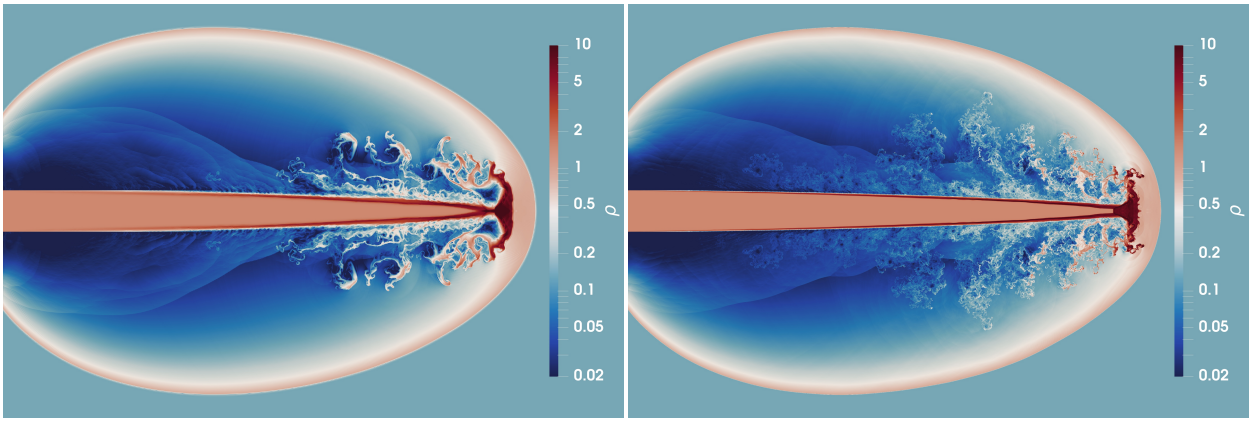
The explosion problem of Nazarov and Larcher [53] was then used to showcase the extension of the approach to unstructured/curved grids. The problem consists of a circular domain of radius 2, with eight circular cylinders of radius 0.3 placed azimuthally equispaced such that their centers were located 1.4 units from the origin. An initial overpressure region in the center drives an outward-running shock and contact discontinuity, which interact with the curved surfaces to form complex shock interactions. The initial conditions were set similarly to a radially-symmetric Sod shock tube problem, with an inner gas region of radius $\sqrt{0.4}$ (centered at the origin) set to $\mathbf{q} = [1, 0, 0, 1]$ and the remaining domain set to $\mathbf{q} = [0.125, 0, 0, 0.1]$. Adiabatic slip wall boundary conditions were applied to all surfaces.

An unstructured mesh was generated with approximately 9.68×10^5 second-order triangular elements. The problem was then simulated on this mesh using a $\mathbb{P}_2$ approximation. A Schlieren-type representation of the density gradient norm is shown in Fig. 11 at several time instances. The interactions of the shocks with the contact line, which form Richtmyer–Meshkov instabilities and complex vortical flow, were well captured by the approach. Good resolution of the discontinuities in the flow as well as the small-scale flow features was observed, showcasing the applicability of the approach on unstructured meshes.

6.10. Mach 800 astrophysical jet

As a final validation study and an evaluation of the approach for very extreme flow conditions, the problem of high-speed astrophysical jets was considered. The test case was introduced by Balsara [54] and consists of a Mach 800 jet expanding into an ambient gas. An identical problem setup to Dzanic and Witherden [17] was used. The half-

domain was set as $\Omega = [0, 0.5] \times [0, 1.5]$, with adiabatic slip wall (symmetry) boundary conditions along the y-axis and Dirichlet boundary conditions on the outer boundaries. The initial conditions were set to $\mathbf{q} = [0.1\gamma, 0, 0, 1]$. At the $y = 0$ boundary, the inlet region was defined on $x \leqslant 0.05$, with the solution set as $\mathbf{q} = [\gamma, 0, 800, 1]^T$, while the remaining boundary region was set to a Dirichlet boundary conditions.



(a) $N = 200 \times 600$ (b) $N = 800 \times 2400$

Figure 12: Contours of density for the Mach 800 astrophysical jet problem at $t = 0.002$ computed using a $\mathbb{P}_3$ approximation with $N = 200 \times 600$ (left) and $N = 800 \times 2400$ (right) elements. Contours are reflected about the y-axis for visualization, and view is reoriented such that the $+y$ direction is shown left-to-right.

The problem was solved using a $\mathbb{P}_3$ approximation on a coarse and fine mesh, consisting of $N = 200 \times 600$ and $N = 800 \times 2400$ quadrateral elements, respectively. The contours of density at $t = 0.002$ as computed on these meshes are shown in Fig. 12. The strong leading shock wave was robustly resolved without noticeable spurious oscillations, and small-scale flow features around the jet showed excellent resolution. With increasing mesh resolution, smaller-scale flow features were observed around the jet. Given the identical case setup, the results can be compared to the entropy filtering approach in Dzanic and Witherden [17]. It can be seen that noticeably better resolution of the small-scale flow features was obtained with the proposed approach with similar resolution of the shock structure.

7. Conclusions

In this work, we introduced a novel limiting approach for high-order approximations of the Euler equations. Motivated by the observation that high-order approximations of the Boltzmann equation remain robust around discontinuities and that linear particle transport is notably easier to design limiters for, the proposed approach connects the evolution of the macroscopic solution of the Euler equations to the underlying kinetic representation provided by the Boltzmann equation, which allows for extending limiters designed for linear advection directly to the Euler equations. The technique takes an arbitrary stencil of local solution values and, by applying the corresponding local maximum principle to the analytic expression of linear transport of the Boltzmann particle distribution function in the inviscid limit, yields an expression for the admissible region of particle distribution function values, which may be numerically integrated to yield a set of bounds for the density, momentum, and total energy. Several key properties of these bounds are shown, including preservation of the positivity of density/pressure/internal energy. This approach was applied to an explicit unstructured high-order discontinuous Galerkin/flux reconstruction scheme for a variety of problems in gas dynamics such as extreme shocks, smooth flows, and shock-vortex interactions. It was observed that the approach could robustly resolve strong discontinuities with essentially optimal convergence rates while recovering high-order accuracy in smooth regions without any ad hoc corrections.

This work presents potential for various extensions and improvements which are in the process of being developed. The approach is not strictly restricted to the limiting method used in this work as it primarily is used to compute solution bounds given a set of discrete solution values. Therefore, it may directly be applied to flux-corrected transport-type techniques which rely on enforcing bounds over a set of Riemann-averaged states [21] since it was shown that the Riemann-averaged states preserve these bounds. Furthermore, the kinetic representation provided by the Boltzmann

equation underpins many complex macroscopic governing equations, some for which admissible solution bounds are not necessarily well-understood. As such, the extension of this method to these systems, namely the Navier–Stokes equations and multi-physics systems such as magnetohydrodynamics, can result in more physically consistent limiting techniques. Additionally, algorithmic improvements can be made to increase performance, such as applying adaptive quadrature-based integration methods for the distribution functions.

Acknowledgements

This work was partially performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE–AC52–07NA27344 and the LLNL-LDRD Program under Project tracking No. 24–ERD–050. Release number LLNL–JRNL–867251–DRAFT.

References

[1] B.C. Vermeire, F.D. Witherden, and P.E. Vincent. On the utility of GPU accelerated high-order methods for unsteady flow simulations: A comparison with industry-standard tools. *Journal of Computational Physics*, 334:497–521, April 2017. doi: 10.1016/j.jcp.2016.12.049.

[2] Jay P. Boris and David L. Book. Flux-corrected transport. *Journal of Computational Physics*, 135(2):172–186, August 1997. doi: 10.1006/jcph.1997.5700.

[3] Steven T Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of Computational Physics*, 31(3):335–362, June 1979. doi: 10.1016/0021-9991(79)90051-2.

[4] Will Pazner. Sparse invariant domain preserving discontinuous Galerkin methods with subcell convex limiting. *Computer Methods in Applied Mechanics and Engineering*, 382:113876, August 2021. doi: 10.1016/j.cma.2021.113876.

[5] Yimin Lin, Jesse Chan, and Ignacio Tomas. A positivity preserving strategy for entropy stable discontinuous Galerkin discretizations of the compressible Euler and Navier-Stokes equations. *Journal of Computational Physics*, 475:111850, February 2023. doi: 10.1016/j.jcp.2022.111850.

[6] Andrés M. Rueda-Ramírez, Will Pazner, and Gregor J. Gassner. Subcell limiting strategies for discontinuous Galerkin spectral element methods. *Computers & Fluids*, 247:105627, October 2022. doi: 10.1016/j.compfluid.2022.105627.

[7] Ahmad Peyvan, Khemraj Shukla, Jesse Chan, and George Karniadakis. High-order methods for hypersonic flows with strong shocks and real chemistry. *Journal of Computational Physics*, 490:112310, October 2023. doi: 10.1016/j.jcp.2023.112310.

[8] Dmitri Kuzmin. Monolithic convex limiting for continuous finite element discretizations of hyperbolic conservation laws. *Computer Methods in Applied Mechanics and Engineering*, 361:112804, April 2020. doi: 10.1016/j.cma.2019.112804.

[9] Ami Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49(3):357–393, March 1983. doi: 10.1016/0021-9991(83)90136-5.

[10] Timothy Barth and Dennis Jespersen. The design and application of upwind schemes on unstructured meshes. In *27th Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, January 1989. doi: 10.2514/6.1989-366.

[11] Antony Jameson. Positive schemes and shock modelling for compressible flows. *International Journal for Numerical Methods in Fluids*, 20 (8–9):743–776, April 1995. doi: 10.1002/fld.1650200805.

[12] Chi-Wang Shu. *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, page 325–432. Springer Berlin Heidelberg, 1998. ISBN 9783540498049. doi: 10.1007/bfb0096355.

[13] Xiangxiong Zhang and Chi-Wang Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *Journal of Computational Physics*, 229(23):8917–8934, November 2011. doi: 10.1016/j.jcp.2010.08.016.

[14] Hongqiang Zhu, Yue Cheng, and Jianxian Qiu. A comparison of the performance of limiters for Runge-Kutta discontinuous Galerkin methods. *Advances in Applied Mathematics and Mechanics*, 5(3):365–390, 2013. doi: 10.1017/S2070073300002733.

[15] Di Zhang, Chunbo Jiang, Dongfang Liang, and Liang Cheng. A review on TVD schemes and a refined flux-limiter for steady-state calculations. *Journal of Computational Physics*, 302:114–154, December 2015. doi: 10.1016/j.jcp.2015.08.042.

[16] Michael Dumbser and Raphaël Loubère. A simple robust and accurate a posteriori sub-cell finite volume limiter for the discontinuous Galerkin method on unstructured meshes. *Journal of Computational Physics*, 319:163–199, August 2016. doi: 10.1016/j.jcp.2016.05.002.

[17] T. Dzanic and F.D. Witherden. Positivity-preserving entropy-based adaptive filtering for discontinuous spectral element methods. *Journal of Computational Physics*, 468:111501, November 2022. doi: 10.1016/j.jcp.2022.111501.

[18] Eric J. Ching, Ryan F. Johnson, Sarah Burrows, Jacklyn P. Higgs, and Andrew D. Kercher. Positivity-preserving and entropy-bounded discontinuous Galerkin method for the chemically reacting, compressible Navier-Stokes equations. In *AIAA SCITECH 2023 Forum*. American Institute of Aeronautics and Astronautics, January 2023. doi: 10.2514/6.2023-0664.

[19] Aditya K. Pandare, Jacob Waltz, Weizhao Li, Hong Luo, and Jozsef Bakosi. On the design of stable, consistent, and conservative high-order methods for multi-material hydrodynamics. *Journal of Computational Physics*, 490:112313, October 2023. doi: 10.1016/j.jcp.2023.112313.

[20] Fan Zhang and Jian Cheng. Analysis on physical-constraint-preserving high-order discontinuous Galerkin method for solving Kapila's five-equation model. *Journal of Computational Physics*, 492:112417, November 2023. doi: 10.1016/j.jcp.2023.112417.

[21] Jean-Luc Guermond, Murtazo Nazarov, Bojan Popov, and Ignacio Tomas. Second-order invariant domain preserving approximation of the Euler equations using convex limiting. *SIAM Journal on Scientific Computing*, 40(5):A3211–A3239, January 2018. doi: 10.1137/17m1149961.

[22] Lin Fu, Xiangyu Y. Hu, and Nikolaus A. Adams. A family of high-order targeted ENO schemes for compressible-fluid simulations. *Journal of Computational Physics*, 305:333–359, January 2016. doi: 10.1016/j.jcp.2015.10.037.

[23] Jean-Luc Guermond and Bojan Popov. Invariant domains and first-order continuous finite element approximation for hyperbolic systems. *SIAM Journal on Numerical Analysis*, 54(4):2466–2489, January 2016. doi: 10.1137/16m1074291.

[24] Dominik Wilde, Andreas Krämer, Dirk Reith, and Holger Foysi. High-order semi-Lagrangian kinetic scheme for compressible turbulence. *Physical Review E*, 104(2), August 2021. doi: 10.1103/physreve.104.025301.

[25] Zhaoli Guo, Ruijie Wang, and Kun Xu. Discrete unified gas kinetic scheme for all Knudsen number flows. II. Thermal compressible case. *Physical Review E*, 91(3), March 2015. doi: 10.1103/physreve.91.033313.

[26] T. Dzanic, F.D. Witherden, and L. Martinelli. A positivity-preserving and conservative high-order flux reconstruction method for the polyatomic Boltzmann–BGK equation. *Journal of Computational Physics*, 486:112146, August 2023. doi: 10.1016/j.jcp.2023.112146.

[27] Tarik Dzanic and Luigi Martinelli. Direct molecular gas dynamics simulations of re-entry vehicles via the Boltzmann equation. In *AIAA SCITECH 2024 Forum*. American Institute of Aeronautics and Astronautics, January 2024. doi: 10.2514/6.2024-2726.

[28] Jan S. Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods*. Springer New York, 2008. doi: 10.1007/978-0-387-72067-8.

[29] H. T. Huynh. A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods. In *18th AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics, June 2007. doi: 10.2514/6.2007-4079.

[30] Xiangxiong Zhang and Chi-Wang Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, May 2010. doi: 10.1016/j.jcp.2009.12.030.

[31] Carlo Cercignani. *The Boltzmann Equation and Its Applications*. Springer New York, 1988. doi: 10.1007/978-1-4612-1039-9.

[32] Eleuterio F. Toro. The Riemann problem for the Euler equations. In *Riemann Solvers and Numerical Methods for Fluid Dynamics*, chapter 4, pages 115–157. Springer Berlin Heidelberg, 1997. doi: 10.1007/978-3-662-03490-3_4.

[33] V.V Rusanov. The calculation of the interaction of non-stationary shock waves and obstacles. *USSR Computational Mathematics and Mathematical Physics*, 1(2):304–320, January 1962. doi: 10.1016/0041-5553(62)90062-9.

[34] Sergei Konstantinovich Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[35] Xiangxiong Zhang. On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations. *Journal of Computational Physics*, 328:301–343, January 2017. doi: 10.1016/j.jcp.2016.10.002.

[36] R. Anderson, V. Dobrev, Tz. Kolev, D. Kuzmin, M. Quezada de Luna, R. Rieben, and V. Tomov. High-order local maximum principle preserving (MPP) discontinuous Galerkin finite element method for the transport equation. *Journal of Computational Physics*, 334:102–124, April 2017. doi: 10.1016/j.jcp.2016.12.031.

[37] Shashank Jaiswal. An entropy stable scheme for the non-linear Boltzmann equation. *Journal of Computational Physics*, 463:111289, August 2022. doi: 10.1016/j.jcp.2022.111289.

[38] Dmitri Kuzmin and Friedhelm Schieweck. A parameter-free smoothness indicator for high-resolution finite element schemes. *Open Mathematics*, 11(8), January 2013. doi: 10.2478/s11533-013-0254-4.

[39] C. Baranger, Y. Dauvois, G. Marois, J. Mathé, J. Mathiaud, and L. Mieussens. A BGK model for high temperature rarefied gas flows. *European Journal of Mechanics - B/Fluids*, 80:1–12, March 2020. doi: 10.1016/j.euromechflu.2019.11.006.

[40] S. F. Davis. Simplified second-order Godunov-type methods. *SIAM Journal on Scientific and Statistical Computing*, 9(3):445–473, May 1988. doi: 10.1137/0909030.

[41] F.D. Witherden, A.M. Farrington, and P.E. Vincent. PyFR: An open source framework for solving advection–diffusion type problems on streaming architectures using the flux reconstruction approach. *Computer Physics Communications*, 185(11):3028–3040, November 2014. doi: 10.1016/j.cpc.2014.07.011.

[42] D.M. Williams, L. Shunn, and A. Jameson. Symmetric quadrature rules for simplexes based on sphere close packed lattice arrangements. *Journal of Computational and Applied Mathematics*, 266:18–38, August 2014. doi: 10.1016/j.cam.2014.01.007.

[43] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43(1): 89–112, January 2001. doi: 10.1137/s003614450036757x.

[44] E. F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock Waves*, 4(1):25–34, July 1994. doi: 10.1007/bf01414629.

[45] Gary A Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27(1):1–31, April 1978. doi: 10.1016/0021-9991(78)90023-2.

[46] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, August 1988. doi: 10.1016/0021-9991(88)90177-5.

[47] Wei Liu, Juan Cheng, and Chi-Wang Shu. High order conservative Lagrangian schemes with Lax–Wendroff type time discretization for the compressible Euler equations. *Journal of Computational Physics*, 228(23):8872–8891, December 2009. doi: 10.1016/j.jcp.2009.09.001.

[48] L I Sedov. *Similarity and dimensional methods in mechanics, tenth edition*. CRC Press, 10 edition, September 1993.

[49] Pierre-Henri Maire. A high-order cell-centered Lagrangian scheme for two-dimensional compressible fluid flows on unstructured meshes. *Journal of Computational Physics*, 228(7):2391–2425, April 2009. doi: 10.1016/j.jcp.2008.12.007.

[50] Paul Woodward and Phillip Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of Computational Physics*, 54(1):115–173, April 1984. doi: 10.1016/0021-9991(84)90142-6.

[51] Robert D. Richtmyer. Taylor instability in shock acceleration of compressible fluids. *Communications on Pure and Applied Mathematics*, 13 (2):297–319, May 1960. doi: 10.1002/cpa.3160130207.

[52] E. E. Meshkov. Instability of the interface of two gases accelerated by a shock wave. *Fluid Dynamics*, 4(5):101–104, 1972. doi: 10.1007/bf01015969.

[53] Murtazo Nazarov and Aurélien Larcher. Numerical investigation of a viscous regularization of the Euler equations by entropy viscosity. *Computer Methods in Applied Mechanics and Engineering*, 317:128–152, April 2017. doi: 10.1016/j.cma.2016.12.010.

[54] Dinshaw S. Balsara. Self-adjusting, positivity preserving high order schemes for hydrodynamics and magnetohydrodynamics. *Journal of Computational Physics*, 231(22):7504–7517, September 2012. doi: 10.1016/j.jcp.2012.01.032.