Frame-invariant directional vector limiters for discontinuous Galerkin methods

H. Hajduk, D. Kuzmin, V. Aizinger

Frame-invariant directional vector limiters for discontinuous Galerkin methods

Hennes Hajduk[a], Dmitri Kuzmin[a], Vadym Aizinger[b,c]

[a]*Institute of Applied Mathematics (LS III), TU Dortmund University,*
Vogelpothsweg 87, D-44227 Dortmund, Germany
[b]*Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research,*
Am Handelshafen 12, D-27570 Bremerhaven, Germany
[c]*Applied Mathematics 1, University Erlangen-Nuremberg,*
Cauerstr. 11, D-91058 Erlangen, Germany

Abstract

Second and higher order numerical approximations of conservation laws for vector fields call for the use of limiting techniques based on generalized monotonicity criteria. In this paper, we introduce a family of directional vertex-based slope limiters for tensor-valued gradients of formally second-order accurate piecewise-linear discontinuous Galerkin (DG) discretizations. The proposed methodology enforces local maximum principles for scalar products corresponding to projections of a vector field onto the unit vectors of a frame-invariant orthogonal basis. In particular, we consider anisotropic limiters based on singular value decompositions and the Gram-Schmidt orthogonalization procedure. The proposed extension to hyperbolic systems features a sequential limiting strategy and a global invariant domain fix. The pros and cons of different approaches to vector limiting are illustrated by the results of numerical studies for the two-dimensional shallow water equations and for the Euler equations of gas dynamics.

Keywords: Hyperbolic conservation laws, discontinuous Galerkin methods, vector limiters, objectivity, shallow water equations, Euler equations

Email addresses: `hennes.hajduk@math.tu-dortmund.de` (Hennes Hajduk), `kuzmin@math.uni-dortmund.de` (Dmitri Kuzmin), `vadym.aizinger@awi.de` (Vadym Aizinger)

1. Introduction

The most robust finite volume and discontinuous Galerkin (DG) methods for transport problems with steep gradients use discretizations that guarantee boundedness and/or monotonicity preservation. In fact, no DG approximation with polynomial basis functions of degree one or higher can guarantee the absence of spurious oscillations in the vicinity of discontinuities without using some kind of flux or slope limiting. The importance of using numerical schemes that strictly adhere to maximum principles varies strongly in different applications and may be critical if the species transport is coupled to other physical processes, e.g., chemical reactions. A variety of limiters for DG schemes can be found in the literature [5, 6, 7, 19, 33, 40]. In our previous work, we developed vertex-based limiters for high-order finite elements [1, 22], anisotropic transport problems [3], and hyperbolic systems [9].

A typical limiting technique constrains a scalar quantity (density, total energy, volume fraction etc.) or a set of scalar quantities of interest to satisfy problem-dependent discrete maximum principles. Less common are limiting procedures specifically designed for vectors such as the velocity field in equations of fluid dynamics or the magnetic field in models of ideal or resistive magnetohydrodynamics (MHD). The reasons for this state of affairs are twofold. First, numerical solutions are usually less sensitive to small fluctuations in the velocity field than to violations of maximum principles for concentrations. Hence, the use of linear stabilization techniques may be sufficient. Second, appropriate maximum principles are far more difficult to formulate and enforce if the variable to be constrained is a vector field rather than a scalar quantity. The 'naive' approach based on separate componentwise limiting may violate the principle of objectivity and destroy rotational or planar symmetries (if any), as reported in [28]. On the other hand, limiters designed to constrain scalar functions of the vector field using a common correction factor for all components are too diffusive and incapable of delivering optimal convergence rates for smooth data [9].

The desired properties of any limiter for high-order terms include

- reliable detection and elimination of nonphysical oscillations;

- preservation of optimal approximation order in smooth regions;

- preservation of cell averages in the process of slope limiting.

In addition to these fundamental requirements, a well-designed vector limiter should be objective (i.e., frame-invariant) and possess good symmetry preservation properties. Applications that call for the use of such limiters include many problems of shock hydrodynamics and ideal MHD, as well as interface problems and wetting/drying processes in geophysical flows.

Vector limiting techniques based on the above design criteria were recently proposed in [18, 28, 29, 30, 31, 32, 39]. Luttwak and Falcovitz [28] were the first to emphasize the importance of objectivity in this context and propose a generalized monotonicity criterion for vectors. Their vector image polygon (VIP) limiter constrains a high-order approximation of the vector field to lie in the convex hull of samples corresponding to a physically admissible but inaccurate low-order approximation. A conceptually simpler and less diffusive version of this limiting strategy involves the use of bounding boxes instead of convex hulls [30]. More straightforward extensions of scalar limiting procedures perform componentwise limiting in frame-invariant (and problem-dependent) reference frames. For example, a vector limiter of this kind may be configured to constrain scalar products with the density gradient or the principal axes of a strain rate tensor [31, 32, 39].

This paper addresses the design of vector limiters for property-preserving DG methods and is organized as follows. In Section 2, we introduce some notation that we need to formulate the general limiting problem for a piecewise-linear vector field. In Section 3, we cast the scalar vertex-based limiter, as presented in [22], into a form suitable for directional vector limiting. In Sections 4 and 5, we dwell on the aspects of frame invariance and explore different approaches to definition of objective limiting directions. Section 6 presents a sequential limiting strategy for systems of conserved quantities. Numerical studies for the 2D shallow water equations and for the compressible Euler equations are performed in Section 7. The outcomes of this work are summarized and conclusions are drawn in Section 8.

2. Notation and preliminaries

Let $\mathbf{v} : \Omega \mapsto \mathbb{R}^d$, $d \in \{2, 3\}$ be a vector field defined in a bounded domain $\Omega \subset \mathbb{R}^d$ with a Lipschitz boundary $\partial\Omega$. Given a conforming mesh

$$\mathcal{T}_h = \left\{ K^1, \ldots, K^{N_{\mathrm{ele}}} \ : \ \bar{\Omega} = \bigcup_{e=1}^{N_{\mathrm{ele}}} K^e \right\},$$

3

we define the piecewise-linear DG space

$$\mathbb{V}_h^1 = \{v_h \in L^2(\Omega) : v_h|_K \in \mathbb{P}_1(K) \quad \forall K \in \mathcal{T}_h\} \qquad (1)$$

and its piecewise-constant counterpart

$$\mathbb{V}_h^0 = \{v_h \in L^2(\Omega) : v_h|_K \in \mathbb{P}_0(K) \quad \forall K \in \mathcal{T}_h\}. \qquad (2)$$

The vertices of an element $K^e \in \mathcal{T}^e$ will be denoted by $\mathbf{x}_1^e, \ldots, \mathbf{x}_{N_{\text{dof}}^e}^e$. The linear Taylor polynomial of $\mathbf{v}_h \in (\mathbb{V}_h^1)^d$ restricted to K^e is given by

$$\mathbf{v}_h^e(\mathbf{x}) = \mathbf{v}_0^e + \mathbf{G}^e(\mathbf{x} - \mathbf{x}_0^e), \qquad \mathbf{x} \in K^e, \qquad (3)$$

where

$$\mathbf{x}_0^e = \frac{1}{|K^e|} \int_{K^e} \mathbf{x} \, d\mathbf{x}$$

is the center of mass, $\mathbf{G}^e = \nabla \mathbf{v}_h^e \in \mathbb{R}^{d \times d}$ is the tensor-valued constant gradient of the linear *shape function* $\mathbf{v}_h^e$ and

$$\mathbf{v}_0^e = \frac{1}{|K^e|} \int_{K^e} \mathbf{v}_h^e \, d\mathbf{x}$$

is the cell average which coincides with the value $\mathbf{v}_h^e(\mathbf{x}_0^e)$ at the center of mass. The piecewise-constant component of the vector field $\mathbf{v}_h \in (\mathbb{V}_h^1)^d$ can be extracted using the L^2 projection operator $P_h^0 : \mathbb{V}_h^1 \to \mathbb{V}_h^0$ defined by

$$P_h^0 \mathbf{v}_h^e(\mathbf{x}) = \mathbf{v}_0^e \qquad \forall \mathbf{x} \in K^e \in \mathcal{T}_h. \qquad (4)$$

The use of slope limiters in DG and second-order finite volume methods for scalar conserved quantities is intended to guarantee that the solution values at certain control points (e.g., vertices, edge/face barycenters or quadrature points) are bounded by the local maxima and minima of the piecewise-constant component, i.e., by cell averages which are assumed to satisfy all relevant maximum principles. In extensions to vector fields, the objective is to blend the shape functions $\mathbf{v}_h^e$ and $P_h^0 \mathbf{v}_h^e \equiv \mathbf{v}_0^e$ in a way that prevents violations of generalized maximum principles without introducing inordinately large amounts of numerical diffusion and losing superlinear accuracy in smooth regions [22]. To that end, the gradient $\mathbf{G}^e$ is replaced by a limited gradient $\bar{\mathbf{G}}^e$, and the linear shape function (3) is overwritten by

$$\bar{\mathbf{v}}_h^e(\mathbf{x}) = \mathbf{v}_0^e + \bar{\mathbf{G}}^e(\mathbf{x} - \mathbf{x}_0^e), \qquad \mathbf{x} \in K^e \qquad (5)$$

4

at the end of each time step or stage of a strong stability preserving (SSP) Runge-Kutta scheme [13]. The definition of $\bar{\mathbf{G}}^e$ should guarantee boundedness in terms of cell averages and preservation of accuracy in smooth regions. Additionally, a physics-compatible limiting procedure should be objective, i.e., independent of the reference frame. In the next sections, we discuss the desired properties of vector limiters in greater detail and present some practical approaches to calculating $\bar{\mathbf{G}}^e$ in constrained DG-$\mathbb{P}_1$ schemes.

3. Directional vector limiting

A directional limiter for a vector field $\mathbf{v}_h \in \mathbb{V}_h^1$ may be designed to constrain scalar products of the form $\mathbf{q}^e \cdot \mathbf{v}_i^e$, where $\mathbf{q}^e$ is a given limiting direction and $\mathbf{v}_i^e$ is the value of $\mathbf{v}_h$ at the vertex $\mathbf{x}_i^e$. Slope limiting techniques of this kind were proposed by Maire et al. [31, 32] in the context of cell-centered finite volume approximations. Zeng and Scovazzi [39] used physics-aware directional limiting to constrain continuous finite element approximations in flux-corrected remap algorithms for vector fields. In this section, we present a vertex-based scalar product limiter based on the same design principles as the algorithms developed in [31, 32, 39]. New approaches to the choice of frame-invariant limiting directions are proposed in Section 5.

Introducing a diagonal matrix $\mathbf{A}^e = \text{diag}\{\alpha_1^e, \ldots, \alpha_d^e\}$ of directional correction factors α_j^e associated with the columns $\mathbf{q}_j^e$ of a given orthogonal matrix $\mathbf{Q}^e = [\mathbf{q}_1^e, \ldots, \mathbf{q}_d^e]$, we define the limited gradient as follows:

$$\bar{\mathbf{G}}^e = \mathbf{L}^e \mathbf{G}^e, \qquad \mathbf{L}^e = \mathbf{Q}^e \mathbf{A}^e (\mathbf{Q}^e)^T. \tag{6}$$

Let $\bar{\mathbf{v}}_i^e$ denote the value of the constrained linear shape function $\bar{\mathbf{v}}_h^e$ at the vertex $\mathbf{x}_i^e$, $i = 1, \ldots, N_{\text{dof}}^e$. Substituting (6) into (5) and exploiting the fact that the columns of $\mathbf{Q}^e$ form an orthonormal basis, we find that

$$\mathbf{q}_j^e \cdot (\bar{\mathbf{v}}_i^e - \mathbf{v}_0^e) = \mathbf{q}_j^e \cdot \bar{\mathbf{G}}^e (\mathbf{x}_i^e - \mathbf{x}_0^e) = \alpha_j^e \, \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e). \tag{7}$$

For each limiting direction $\mathbf{q}_j^e$, the scalar product $\mathbf{q}_j^e \cdot \mathbf{v}_h^e$ is a linear polynomial which attains its maxima and minima at the vertices of K^e. Let $\mathcal{E}_i^e$ denote the integer set containing the numbers of elements to which the vertex $\mathbf{x}_i^e$ belongs. A simple extension of the scalar vertex-based limiter [22] to vector fields is based on the directional local maximum principle

$$\min_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot \mathbf{v}_0^m \le \mathbf{q}_j^e \cdot \bar{\mathbf{v}}_i^e \le \max_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot \mathbf{v}_0^m, \qquad i = 1, \ldots, N_{\text{dof}}^e. \tag{8}$$

5

In view of (7), the so-defined inequality constraints are equivalent to

$$\min_{m\in\mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{v}_0^m - \mathbf{v}_0^e) \leq \alpha_j^e \, \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e) \leq \max_{m\in\mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{v}_0^m - \mathbf{v}_0^e) \qquad (9)$$

and can be satisfied for all $i = 1, \ldots, N_{\mathrm{dof}}^e$ using the correction factor

$$\alpha_j^e = \min_{1\leq i \leq N_{\mathrm{dof}}^e} \begin{cases} \min\left\{1, \dfrac{\max\limits_{m\in\mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{v}_0^m - \mathbf{v}_0^e)}{\mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e)}\right\} & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e) > 0, \\[3mm] 1 & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e) = 0, \\[3mm] \min\left\{1, \dfrac{\min\limits_{m\in\mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{v}_0^m - \mathbf{v}_0^e)}{\mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e)}\right\} & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e) < 0. \end{cases} \qquad (10)$$

Using this formula to determine the diagonal entries α_j^e of $\mathbf{A}^e$, we obtain a limiting operator $\mathbf{L}^e = \mathbf{Q}^e \mathbf{A}^e (\mathbf{Q}^e)^T$ such that conditions (9) hold for all scalar products $\alpha_j^e \, \mathbf{q}_j^e \cdot (\mathbf{v}_i^e - \mathbf{v}_0^e)$ representing projections of $\mathbf{L}^e \mathbf{G}^e (\mathbf{x}_i^e - \mathbf{x}_0^e)$ onto the orthonormal basis vectors $\mathbf{q}_j^e$ of a rotated reference frame.

4. Objectivity requirements

A highly desirable property of a limiter for vector fields is objectivity [18, 28, 31, 32, 39]. That is, the results should be invariant under translations and/or rotations defined by affine linear transformations of the form

$$\hat{\mathbf{x}} = \mathbf{R}\mathbf{x} + \mathbf{c}, \qquad (11)$$

where $\mathbf{R}$ is an orthogonal second-order tensor, and $\mathbf{c}$ is a constant vector.

To derive a transformation rule for the limited gradient tensor defined by (6), we consider a generic mesh element $K^e \in \mathcal{T}_h$ and omit the superscript e for brevity. The vector quantities to be limited are given by

$$\mathbf{v}_i - \mathbf{v}_0 = \mathbf{G}(\mathbf{x}_i - \mathbf{x}_0) \qquad (12)$$

in the original reference frame associated with $\mathbf{x}$ and by

$$\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_0 = \mathbf{R}\mathbf{G}(\mathbf{x}_i - \mathbf{x}_0) = \mathbf{R}\mathbf{G}\mathbf{R}^T(\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_0) \qquad (13)$$

in the reference frame associated with $\hat{\mathbf{x}}$. The limiting operators

$$\mathbf{L} = \mathbf{Q}\mathbf{A}\mathbf{Q}^T, \qquad \hat{\mathbf{L}} = \hat{\mathbf{Q}}\hat{\mathbf{A}}\hat{\mathbf{Q}}^T \qquad (14)$$

6

produce the same results in the original and rotated frame if

$$\mathbf{RL}(\mathbf{v}_i - \mathbf{v}_0) = \hat{\mathbf{L}}(\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_0). \tag{15}$$

Invoking (12)-(13), we find that this transformation rule is equivalent to

$$\mathbf{RLGR}^T(\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_0) = \hat{\mathbf{L}}\mathbf{RGR}^T(\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_0) \tag{16}$$

These considerations reveal that directional limiting is frame-invariant if the matrices of limiting directions and correction factors satisfy

$$\hat{\mathbf{Q}}\hat{\mathbf{A}}\hat{\mathbf{Q}}^T = \mathbf{RQA}(\mathbf{RQ})^T. \tag{17}$$

The use of $\hat{\mathbf{Q}} = \mathbf{I} = \mathbf{Q}$, where $\mathbf{I}$ is the unit tensor, corresponds to scalar limiting for individual components of the vector field in the current reference frame. Limiting techniques of this kind are not objective because the diagonal entries of the matrices $\hat{\mathbf{A}}$ and $\mathbf{A}$ depend on the choice of the limiting directions. We illustrate this fact by a numerical example in Section 7.1.

5. Objective limiting frames

To satisfy condition (17), the columns of $\mathbf{Q}$ should form a frame-invariant orthonormal basis. Such a basis can be constructed, e.g., using the eigenvectors of the symmetric tensor $\frac{1}{2}(\mathbf{G} + \mathbf{G}^T)$, as proposed in [31, 32, 39]. In this work, we construct objective reference frames for directional limiters using singular value decompositions and the Gram-Schmidt orthogonalization.

The singular value decomposition (SVD) of the gradient $\mathbf{G}$ is given by

$$\mathbf{G} = \mathbf{USW}^T, \tag{18}$$

where $\mathbf{S}$ is the diagonal matrix of singular values, $\mathbf{U}$ is an orthogonal matrix of left singular vectors and $\mathbf{W}$ is an orthogonal matrix of right singular vectors. The corresponding polar decomposition is given by $\mathbf{G} = \mathbf{RP}$, where $\mathbf{R} = \mathbf{UW}^T$ is is orthogonal and $\mathbf{P} = \mathbf{WSW}^T$ is symmetric. The matrix $\mathbf{S}$ of a singular value decomposition is unique. If all singular values are nondegenerate (distinct) and nonzero, then the matrices $\mathbf{U}$ and $\mathbf{W}$ are unique up to multiplication of each row/column by -1.

The directional vector limiter using $\mathbf{Q} = \mathbf{U}$ is objective if the SVD is unique (which is not the case, e.g., for $\mathbf{G} = \mathbf{I}$). In the degenerate case, the result may depend on the numerical algorithm for calculating the SVD.

7

Applying the limiting operator (14) to $\mathbf{G} = \mathbf{QSW}^T$, we obtain

$$\bar{\mathbf{G}} = (\mathbf{QAQ}^T)(\mathbf{QSW}^T) = \mathbf{QASW}^T. \qquad (19)$$

Hence, this directional limiting strategy corresponds to multiplication of the singular values s_j and scalar products $s_j \mathbf{w}_j \cdot (\mathbf{x}_i - \mathbf{x}_0)$ by the correction factors α_j stored in diagonal entries of $\mathbf{A}$. According to the physical interpretation of polar decompositions in terms of rotations and stretches, the gradient-based SVD limiter adjusts stretches without changing the orientation.

Clearly, the limiting directions defined as left singular vectors of an evolving gradient $\mathbf{G}$ must be updated after each time step or Runge-Kutta stage. To avoid the high overhead cost associated with gradient-based directional limiting, the matrix $\mathbf{Q}$ may be defined using an orthogonal basis depending on the orientation of element K. The use of mesh-dependent limiting directions does not violate the principle of objectivity because the mesh is attached to the computational domain and undergoes the same transformations (translations and rotations) under changes of coordinate systems.

On tensor product meshes, the normals to the edges/faces of K form a natural orthogonal basis for directional limiting. In many applications, the normal components of a vector field determine the magnitude of convective fluxes and divergence errors. The use of normals as limiting directions would provide an additional mechanism for direct control of these quantities. However, the number of normals for a general mesh cell exceeds the number of space dimensions. In this case, the only way to enforce maximum principles of the form (8) for each normal is to use the same correction factor $\alpha^e = \min_j \alpha_j^e$ for all components of the gradient, as in synchronized limiters for systems of conservation laws [24, 27]. This limiting strategy is likely to produce poor results in situations when the solution exhibits small variations in some limiting directions and large variations in other directions.

Information about the shape and orientation of a general mesh cell is also encoded in the Jacobian $\mathbf{J}(\mathbf{x})$ of the mapping $F_K : \hat{K} \mapsto K$ from a reference element $\hat{K}$ on which the basis functions are commonly defined in finite element codes. The left singular values of $\mathbf{J}_0 := \mathbf{J}(\mathbf{x}_0)$ can be used to define the orthogonal matrix $\mathbf{Q}$ for directional limiting. In contrast to the gradient-based SVD approach, this matrix needs to be calculated just once for each element provided that the mesh does not change in time.

If K is a simplex in $\mathbb{R}^d$, the Jacobian $\mathbf{J} \equiv \mathbf{J}_0$ is constant on K. For instance, the reference triangle $\hat{K} = \text{conv}\{(0,0), (1,0), (0,1)\}$ is mapped onto

a general element $K = \mathrm{conv}\,\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ of a triangular mesh using

$$F_K(\hat{\mathbf{x}}) = \begin{pmatrix} \mathbf{x}_2 - \mathbf{x}_1 & \mathbf{x}_3 - \mathbf{x}_1 \end{pmatrix} \hat{\mathbf{x}} + \mathbf{x}_1. \tag{20}$$

The Jacobian-based SVD limiter is objective if the singular values of $\mathbf{J}_0$ are distinct (which is not the case, e.g., on a uniform rectangular mesh). In the case of coinciding singular values, the outcome depends on the numerical algorithm for calculating the SVD and may be sensitive to small perturbations of the mesh. However, the limiting directions are not affected by changes of the solution in contrast to the gradient-based SVD limiter.

As a cheaper alternative which guarantees uniqueness, we consider the possibility of generating the matrix $\mathbf{Q}$ using the Gram-Schmidt algorithm to extract an orthonormal basis of limiting directions $\mathbf{q}_1, \ldots, \mathbf{q}_d$ from the linearly independent columns of the nonsingular Jacobian matrix $\mathbf{J}_0$. The so-defined directional limiter is unconditionally frame-invariant, but the limiting directions depend on the numbering of the vertices of K.

In Section 7, we verify the claims regarding objectivity properties and use the above limiting techniques to constrain velocity fields in 2D.

6. Sequential limiting for systems

In many applications of practical interest, mathematical models are based on systems of conservation laws, and the set of quantities to be limited include vector fields of form $\mathbf{v} = \varrho \mathbf{u}$. For example, the momentum density is defined as the product of a velocity field $\mathbf{u}$ and a scalar variable ϱ (e.g., density in the Euler and Navier-Stokes equations of fluid mechanics or water height in shallow water models of geophysical flows). Extensions of vertex-based limiters to finite element discretizations of such problems can be found in [4, 9, 15, 27]. In addition to local maximum principles for conserved variables, the nature of the problem at hand may require preservation of global bounds for some derived quantities such as pressure, internal energy, or entropy [15].

As an alternative to directional limiting for a momentum-like vector field $(\varrho \mathbf{u})_h$, the sequential limiting procedure proposed in [9] may be used to constrain the velocity-like derived quantity $\mathbf{u}_h = (\varrho \mathbf{u})_h / \varrho_h$ as follows:

1. Calculate the correction factor α_ϱ^e for the gradient of the density-like variable ϱ_h using the scalar version of the vertex-based limiter.
2. Limit ϱ_h^e and update the gradient of $(\varrho \mathbf{u})_h^e$ using the product rule

$$\mathbf{G}^e := \nabla(\varrho \mathbf{u})_h^e - \mathbf{u}_0^e \otimes \nabla \bar{\varrho}_h^e, \tag{21}$$

9

where $\mathbf{u}_0^e = \frac{(\varrho\mathbf{u})_0^e}{\varrho_0^e}$ is a low-order approximation to $\mathbf{u}_h^e$ and $\nabla\bar{\varrho}_h^e = \alpha_\varrho^e \nabla\varrho_h^e$ is the gradient of the limited density field $\bar{\varrho}_h^e$.

3. Given an orthogonal matrix $\mathbf{Q}^e$ of limiting directions, calculate a diagonal matrix $\mathbf{A}^e$ of directional correction factors α_j^e defined by

$$
\alpha_j^e = \min_{1 \leq i \leq N_{\mathrm{dof}}^e} \begin{cases} \min\left\{1, \dfrac{\max\limits_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{u}_0^m - \mathbf{u}_0^e)}{\mathbf{q}_j^e \cdot (\mathbf{u}_i^e - \mathbf{u}_0^e)}\right\} & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{u}_i^e - \mathbf{u}_0^e) > 0, \\[4mm] 1 & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{u}_i^e - \mathbf{u}_0^e) = 0, \\[4mm] \min\left\{1, \dfrac{\min\limits_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot (\mathbf{u}_0^m - \mathbf{u}_0^e)}{\mathbf{q}_j^e \cdot (\mathbf{u}_i^e - \mathbf{u}_0^e)}\right\} & \text{if} \quad \mathbf{q}_j^e \cdot (\mathbf{u}_i^e - \mathbf{u}_0^e) < 0. \end{cases}
\tag{22}
$$

4. Replace $\mathbf{G}^e$ by $\bar{\mathbf{G}}^e = \mathbf{Q}^e \mathbf{A}^e (\mathbf{Q}^e)^T \mathbf{G}^e$, i.e., apply the limiting operator $\mathbf{L}^e = \mathbf{Q}^e \mathbf{A}^e (\mathbf{Q}^e)^T$ to produce the shape function

$$
\overline{(\varrho\mathbf{u})}_h^e = \bar{\varrho}_h^e \mathbf{u}_0^e + \mathbf{L}^e((\varrho\mathbf{u})_h^e - \bar{\varrho}_h^e \mathbf{u}_0^e)
\tag{23}
$$

satisfying the directional local maximum principles

$$
\bar{\varrho}_i^e \min_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot \mathbf{u}_0^m \leq \mathbf{q}_j^e \cdot \overline{(\varrho\mathbf{u})}_i^e \leq \bar{\varrho}_i^e \max_{m \in \mathcal{E}_i^e} \mathbf{q}_j^e \cdot \mathbf{u}_0^m
\tag{24}
$$

for all $i = 1, \dots, N_{\mathrm{dof}}^e$ and all limiting directions $\mathbf{q}_j^e$, $j = 1, \dots, d$.

For a detailed presentation of the underlying design philosophy and additional limiting tools for invariant domain preserving finite element discretizations of hyperbolic systems, we refer the interested reader to [9, 15]. A closely related sequential limiter for symmetry- and essentially-bound-preserving flux-corrected remapping of momentum can be found in [36].

7. Numerical examples

In this section, we illustrate the implications of objectivity by a numerical example and apply directional limiting techniques to hyperbolic systems. The vector limiters considered in our numerical study are abbreviated by

- DL-FIX($\mathbf{I}$): componentwise limiting in the Cartesian (x, y) plane;

- DL-FIX($\mathbf{R}$): limiting in the plane defined by a rotation matrix $\mathbf{R}$;

- DL-SVD($\mathbf{G}$): gradient-based SVD limiter;

- DL-SVD($\mathbf{J}$): Jacobian-based SVD limiter;

- DL-GS($\mathbf{J}$): Jacobian-based Gram-Schmidt limiter.

When it comes to solving the shallow water equations in Section 7.2 and the Euler equations in Section 7.3, we adopt the sequential limiting strategy outlined in Section 6. An additional fix is proposed to enforce global bounds associated with invariant domains of hyperbolic systems.

7.1. Illustration of rotational invariance

We begin with a numerical example designed to verify the rotational invariance of vector limiters and quantify changes (if any) caused by rotations of the computational domain. To that end, we define the vector field

$$\mathbf{w}_h(x,y) = (|x| + |y|) \begin{cases} (1,0) & \text{if } x, y < 0, \\ (0,-1) & \text{otherwise} \end{cases} \qquad (25)$$

on the triangulation shown in the left panel of Fig. 1. For testing purposes, we rotate the domain and the vector field using the orthogonal matrix

$$\mathbf{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \theta = \frac{\pi}{6}.$$

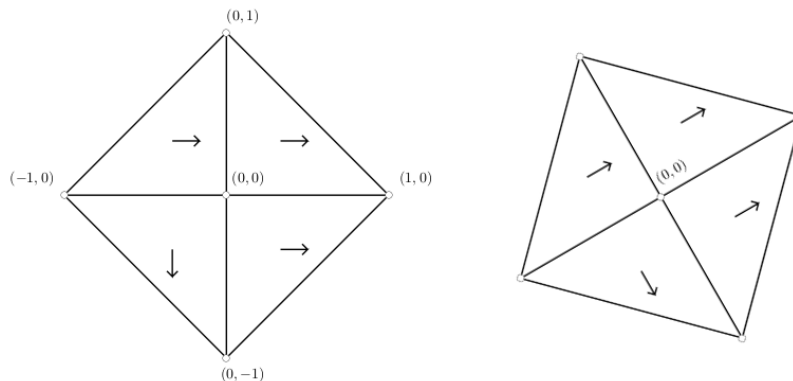The result of this transformation is shown in the right panel of Fig. 1.

Figure 1: Visualization of the piecewise-linear vector field $\mathbf{w}_h$ defined by (25) on a mesh of 4 triangular elements in the original (left: $\theta = 0$) and rotated (right: $\theta = \frac{\pi}{6}$) reference frame. The arrows show the magnitude and orientation of the centroid vectors $\mathbf{w}_0^e$.

The application of the limiter DL-FIX($\mathbf{I}$) to $\mathbf{w}_h$ reveals the following behavior: On the original grid, the correction factors for the center vertex are equal to 1 since the components of $\mathbf{w}_h(0,0) = (0,0)$ are bounded above and below by the x- and y-components of the centroid vectors. The rotation of the domain about the origin produces a vector field $\mathbf{R}\mathbf{w}_h$ which has the same value at $(0,0)$ and the same components in the rotated reference frame defined by the matrix $\mathbf{R}$. However, the x-components of all centroid vectors become strictly positive after this rotation. As a consequence, the DL-FIX($\mathbf{I}$) limiter is activated to enforce a positive lower bound on the rotated mesh, and a different result is obtained if the limited vector field is rotated back using the inverse mapping of the domain to the original configuration.

In Table 1, we present the L^∞ norms of the difference between the limited piecewise-linear approximations $\overline{\mathbf{w}}_h$ and $\mathbf{R}^T(\overline{\mathbf{R}\mathbf{w}_h})$. For reasons explained above, there is a significant discrepancy between the results calculated using the componentwise DL-FIX($\mathbf{I}$) limiter on the two grids. It is easy to verify that all other directional vector limiters satisfy the objectivity criterion (17). Therefore, the corresponding L^∞ errors are zero to machine precision.

scheme	$\|\overline{\mathbf{w}}_h - \mathbf{R}^T(\overline{\mathbf{R}\mathbf{w}_h})\|_{L^\infty}$
DL-FIX($\mathbf{I}$)	0.33
DL-FIX($\mathbf{R}$)	0.43E–15
DL-SVD($\mathbf{G}$)	0.20E–15
DL-SVD($\mathbf{J}$)	0.17E–15
DL-GS($\mathbf{J}$)	0.43E–15

Table 1: L^∞ norms of the difference between the results of directional limiting on the original ($\theta = 0$) and rotated ($\theta = \frac{\pi}{6}$) mesh of 4 triangular elements.

If the actual limiting is performed using a scaled Taylor basis (as in [22, 23]), any dependence of the scaling factors on the orientation of mesh elements may result in a lack of objectivity. For the directional limiter to be objective, the normalization factors should be defined using the transformed coordinates of the vertices rather than the Cartesian coordinates of the xy reference frame. We also remark that the results for DL-SVD may be sensitive to practical implementation of singular value decompositions.

Now that the use of objective limiters for projections of generic vector fields is motivated by an admittedly simple theoretical example, we proceed

to numerical studies of the proposed approaches in the context of sequential limiting for nonlinear hyperbolic PDE systems.

7.2. Case study: shallow water equations

The two-dimensional shallow water equations (SWE) are given by

$$\frac{\partial H}{\partial t} + \nabla \cdot (H\mathbf{u}) = 0, \qquad (26)$$

$$\frac{\partial (Hu)}{\partial t} + \nabla \cdot (Hu\mathbf{u}) + \frac{g}{2}\frac{\partial H^2}{\partial x} + g H \frac{\partial b}{\partial x} + \tau_{bf} Hu - f_c Hv = 0, \qquad (27)$$

$$\frac{\partial (Hv)}{\partial t} + \nabla \cdot (Hv\mathbf{u}) + \frac{g}{2}\frac{\partial H^2}{\partial y} + g H \frac{\partial b}{\partial y} + \tau_{bf} Hv + f_c Hu = 0, \qquad (28)$$

where $\mathbf{u} = [u, v]^T$ is the depth-averaged velocity, and $H = \xi - b$ is the total water height, that is, the difference between the free surface elevation ξ and the bed topography b, also called bathymetry. The source terms proportional to τ_{bf} and f_c are due to bottom friction and Coriolis forces, respectively.

The nonlinear system can be written in the generic divergence form

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}, \nabla b), \qquad (29)$$

where

$$\mathbf{U} = \begin{bmatrix} H \\ H\mathbf{u} \end{bmatrix}, \qquad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} H\mathbf{u} \\ H\mathbf{u} \otimes \mathbf{u} + \frac{gH^2}{2}\mathbf{I} \end{bmatrix},$$

$$\mathbf{S}(\mathbf{U}, \nabla b) = \begin{bmatrix} 0 \\ f_c vH - \tau_{bf} uH - gH\frac{\partial b}{\partial x} \\ -f_c uH - \tau_{bf} vH - gH\frac{\partial b}{\partial y} \end{bmatrix}.$$

For any element $K^- \in \mathcal{T}_h$ and any test function $\mathbf{w}_h \in (\mathbb{V}_h)^3$, the element-local semi-discrete variational form of system (29) is given by [2, 16]

$$\int_{K^-} \mathbf{w}_h \cdot \partial_t \mathbf{U}_h \, \mathrm{d}\mathbf{x} \;-\; \int_{K^-} \nabla \mathbf{w}_h : \mathbf{F}(\mathbf{U}_h) \, \mathrm{d}\mathbf{x}$$

$$+ \int_{\partial K^-} \mathbf{w}_h \cdot \widehat{\mathbf{F}}(\mathbf{U}_h^-, \mathbf{U}_h^+; \boldsymbol{\nu}_{K^-}) \, \mathrm{d}s = \int_{K^-} \mathbf{w}_h \cdot \mathbf{S}(\mathbf{U}_h, \nabla b_h) \, \mathrm{d}\mathbf{x}, \qquad (30)$$

where $\boldsymbol{\nu}_{K^-}$ is the unit outward normal and $\widehat{\mathbf{F}}(\mathbf{U}_h^-, \mathbf{U}_h^+; \boldsymbol{\nu}_{K^-})$ is a numerical flux defined in terms of the one-sided limits $\mathbf{U}_h^{\pm}$ (see [16] for details).

13

The current implementation of the DG method for the SWE system is based on the MATLAB / GNU Octave toolbox FESTUNG [12, 20, 34]. In the numerical study below, we use the Lax-Friedrichs flux with Roe-Pike averaging and a globally continuous piecewise linear approximation b_h to the bathymetry. Time integration is performed using an explicit second-order SSP Runge-Kutta scheme (Heun's method). A full description of the employed numerical methods can be found in [17].

7.2.1. Unidirectional dam break

Dam break problems are commonly used to illustrate the necessity of slope limiting in numerical methods for shallow water models and perform numerical studies of limiting techniques. The solution of the one-dimensional Riemann problem with a discontinuous initial height H and initial velocity $\mathbf{u} = \mathbf{0}$ is a rarefaction wave traveling in one direction and a shock propagating in the opposite one. The model parameters of the 1D shallow water equations are given by $g = 9.81$, $b \equiv \mathrm{const} = -1$, $\tau_{bf} = f_c = 0$. The analytical solution of this 1D test problem is known and can be found in [8].

In this experiment, we calculate pseudo-one-dimensional numerical solutions by solving the 2D shallow water problem (26)–(28) in the domain $\Omega = (0, 1000) \times (0, 1)$. The initial condition for H is given by

$$H(x, y, 0) = \begin{cases} 10 & \text{if } x \leq 500, \\ 1 & \text{otherwise.} \end{cases}$$

The triangular computational mesh is generated from a uniform Cartesian grid with spacing $\Delta x = \Delta y = 1$ by subdividing each square element into two right triangles. We consider the time interval $(0, 20)$ and use the constant time step $\Delta t = 0.001$ in this test. In Figure 2, we show the results produced by different directional vector limiters at the final time $T = 20$. It can be seen that the snapshots are in good agreement with the analytical solutions for the free surface elevation ξ and the horizontal velocity u. The lack of spurious oscillations in the velocity distribution is due to the fact that we use the sequential form of the vector limiter (as described in Section 6) to enforce maximum principles for the velocity rather than momentum.
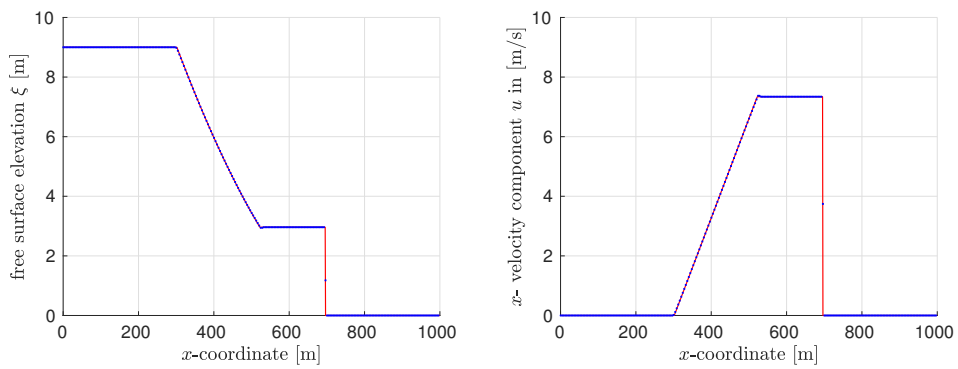
The DG method is capable of resolving the discontinuous initial condition for H exactly provided that the discontinuity occurs on the boundary between two elements. The employed mesh satisfies this requirement since it is fitted to the interface $x = 500$. To check if a given vector limiting

14
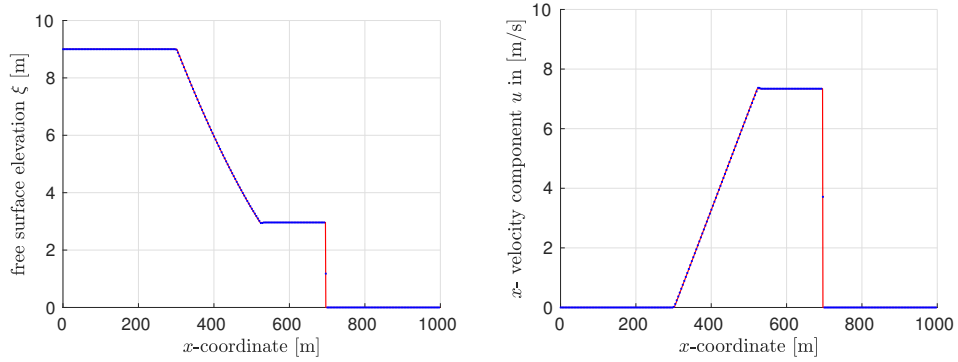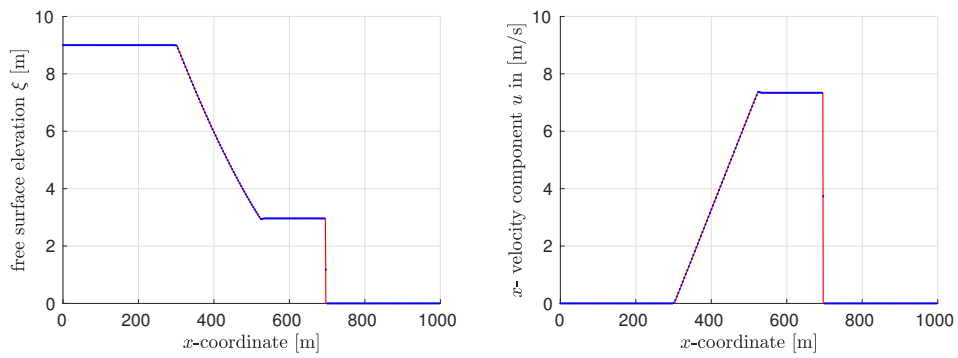
Figure 2: Analytical solution (red solid line) of the one-dimensional dam break problem and numerical approximations (blue dashed lines) obtained with different limiting schemes at time $T = 20$ $[s]$.

scheme	e_H	$e_{H\mathbf{u}}$	$e_{\mathbf{u}}$
DL-FIX($\mathbf{I}$)	4.79E–02	1.01E+00	3.34E–01
DL-FIX($\mathbf{R}$)	2.71E–09	2.11E–08	4.68E–09
DL-SVD($\mathbf{G}$)	1.03E–05	4.08E–04	1.07E–04
DL-SVD($\mathbf{J}$)	3.07E–08	2.80E–07	6.42E–08
DL-GS($\mathbf{J}$)	2.71E–09	2.11E–08	4.68E–09

Table 2: Objectivity errors caused by a rotation of the reference frame.

technique is objective, we rotate the domain by an angle θ, define the initial condition for H using the corresponding rotated coordinates $\hat{\mathbf{x}} = \mathbf{R}\mathbf{x}$, run the DG code on the rotated mesh, and transform back to the original ($\theta = 0$) reference frame using the matrix $\mathbf{R}^T$. The resulting numerical solution is denoted by the superscript rot. The difference between the slope-limited DG-$\mathbb{P}_1$ approximations $\bar{\mathbf{U}}_h$ and $\bar{\mathbf{U}}_h^{rot}$ is measured using the error metrics

$$
\begin{aligned}
e_H &:= \|\bar{H}_h - \bar{H}_h^{rot}\|_{L^\infty(\Omega)}, \\
e_{H\mathbf{u}} &:= \|\overline{(H\mathbf{u})}_h - \overline{(H\mathbf{u})}_h^{rot}\|_{L^\infty(\Omega)^2}, \\
e_{\mathbf{u}} &:= \|\bar{\mathbf{u}}_h - \bar{\mathbf{u}}_h^{rot}\|_{L^\infty(\Omega)^2}.
\end{aligned}
$$

The errors listed in Table 2 summarize the results of the above objectivity test for the rotation angle $\theta = \frac{\pi}{6}$. The objectivity criterion implies that any differences between the numerical solutions calculated on the original and rotated mesh should be due to round-off errors. As mentioned previously, the componentwise DL-FIX($\mathbf{I}$) limiter violates the objectivity requirements formulated in Section 4. The first line of Table 2 confirms that rotations of the mesh may give rise to significant nonphysical changes of numerical solutions if the limiting directions are not rotated by the same angle. The objectivity errors of limiters that employ mesh-dependent limiting directions are of the order of machine precision. The DL-SVD($\mathbf{G}$) version produces errors that are small compared to DL-FIX($\mathbf{I}$) and large compared to other limiters. As simulation continues, objectivity errors increase at a rate depending on the choice of the limiting strategy and its practical implementation.

The higher sensitivity of DL-SVD solutions to changes of the reference frame could be explained by the possible nonuniqueness of the rotation matrix $\mathbf{Q}$. The singular value decomposition of a general matrix is nonunique and its computation is ill-posed if the singular values are (almost) equal or

one of them approaches zero. For DL-SVD($\mathbf{G}$) this situation occurs wherever the gradient is singular or is a scaled rotation matrix itself. The limiting directions of DL-SVD($\mathbf{J}$) are nonunique when the rows of the Jacobian matrix $\mathbf{J}$ have comparable Euclidean norms, i.e., on triangles with two edges of approximately the same length. Note that all elements of the mesh employed in our numerical study are right triangles with equal legs. The DL-GS($\mathbf{J}$) version is more robust and produces the same results as DL-FIX($\mathbf{R}$) in this example because the legs of right triangles are parallel to the axes of the rotated coordinate system. In light of the above, we recommend the use of DL-FIX($\mathbf{R}$) for uniform meshes and DL-GS($\mathbf{J}$) for general meshes.

7.2.2. Radially symmetric dam break

As a real 2D test case, we consider a radially symmetric version of the above example. The parameters of the two-dimensional shallow water model are given by $g = 9.81$, $b \equiv \text{const} = -0.1$, $\tau_{bf} = f_c = 0$. The computational domain $\Omega = (-10, 10)^2$ is discretized using a uniform mesh of triangular elements with spacing $\Delta x = \Delta y = 0.1$. The initial height is given by

$$H(x, y, 0) = \begin{cases} 1 & \text{if } x^2 + y^2 \leq 16, \\ 0.1 & \text{otherwise.} \end{cases}$$

The time interval for this test is $(0, 1)$ and the time step is $\Delta t = 0.004$.

The pattern of wave propagation in each radial direction is supposed to reproduce the qualitative behavior of the solution to the 1D dam break problem. Additionally, a well-designed vector limiter should preserve the radial symmetry of the initial data. The numerical solution presented in Fig. 3 demonstrates the ability of the objective directional limiter DL-SVD($\mathbf{G}$) to capture shocks and rarefaction waves while preserving radial symmetry. The frame-dependent DL-FIX($\mathbf{I}$) limiter produces similar results (not shown here). Furthermore, the differences between the approximations produced by DL-SVD($\mathbf{G}$), DL-SVD($\mathbf{J}$) and DL-GS($\mathbf{J}$) are rather small as can be seen from Fig. 4–5. Due to the high computational cost of DL-SVD($\mathbf{G}$), the Jacobian-based methods are preferable. As in the previous example, the use of sequential limiting provides crisp resolution of steep velocity gradients without generating undershoots and overshoots that may occur if the density and momentum fields are limited separately.
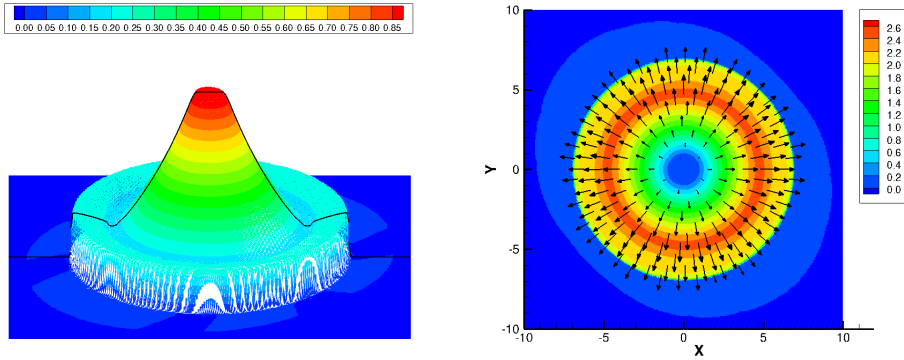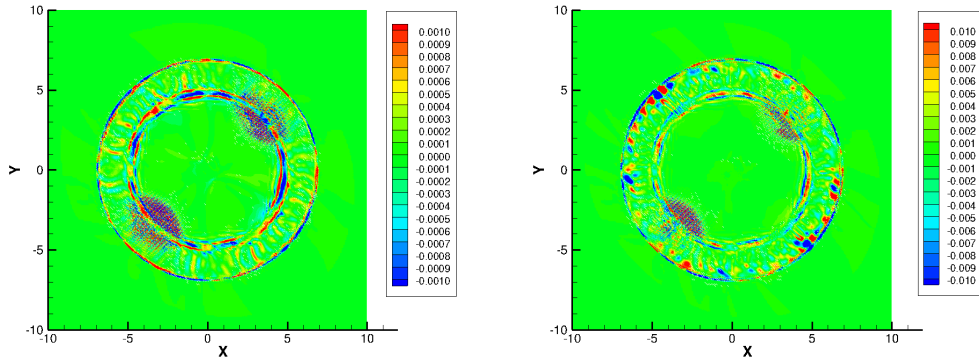
Figure 3: DL-SVD($\mathbf{G}$)

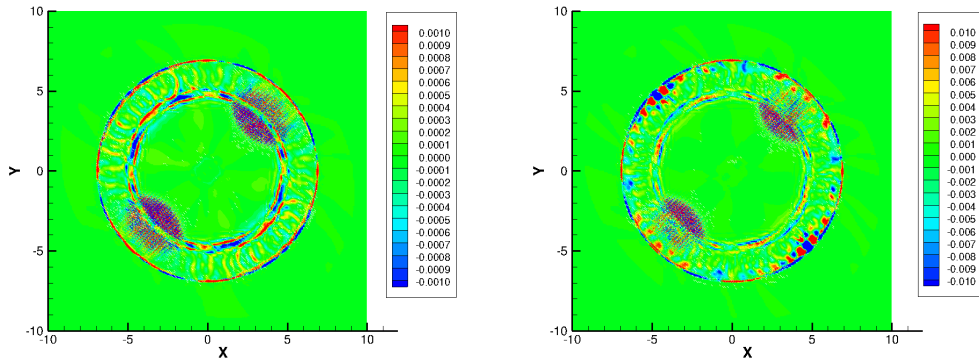Figure 4: Difference plot DL-SVD($\mathbf{J}$) vs. DL-SVD($\mathbf{G}$)

Figure 5: Difference plot DL-GS($\mathbf{J}$) vs. DL-SVD($\mathbf{G}$)

7.2.3. Vorticity advection in 2D

For a better quantitative comparison of the new limiters, we perform a grid convergence study for a smooth vortex traveling at the constant speed

18

$M = \frac{1}{2}$ in the direction specified by the angle $\alpha = \frac{\pi}{6}$. A detailed description and the smooth analytical solution of this 2D test problem can be found in Section 4.3.1 of [11]. We consider the homogeneous shallow water system corresponding to (26)–(28) with $g = 1$, $b \equiv \text{const} = -1$, $\tau_{bf} = f_c = 0$. The computational domain for this test is $\Omega = (-50, 50) \times (-50, 50)$. The Dirichlet boundary conditions for all variables are inferred from the known analytical solution. Using successive refinement in space and time, we compute the $L^2(\Omega)$ errors for each unknown at the final time $T = 100$. The Jacobian-based limiters deliver at least second-order convergence rates for each quantity of interest. Hence, these limiters perform well not only in the proximity of shocks but also for problems with smooth solutions. The inferior convergence behavior of DL-SVD($\mathbf{G}$) could be improved by using more stable numerical implementations of the singular value decomposition.

7.3. Case study: Euler equations

The Euler equations represent a hyperbolic system of conservation laws

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \mathbf{u}) = 0, \tag{31}$$

$$\frac{\partial (\varrho \mathbf{u})}{\partial t} + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) = 0, \tag{32}$$

$$\frac{\partial (\varrho E)}{\partial t} + \nabla \cdot (\varrho E \mathbf{u} + p\mathbf{u}) = 0, \tag{33}$$

where ϱ is the density, $\mathbf{u}$ is the velocity, and E is the specific total energy.

The pressure p of an ideal gas with the heat capacity ratio γ is given by

$$p = (\gamma - 1)\left(\varrho E - \frac{|\varrho \mathbf{v}|^2}{2\varrho}\right). \tag{34}$$

The generic divergence form of the nonlinear system (31)-(33) reads

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}) = \mathbf{0}, \tag{35}$$

where

$$\mathbf{U} = \begin{bmatrix} \varrho \\ \varrho \mathbf{u} \\ \varrho E \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} \varrho \mathbf{u} \\ \varrho \mathbf{u} \otimes \mathbf{u} + p\mathbf{I} \\ \varrho E \mathbf{u} + p\mathbf{u} \end{bmatrix}. \tag{36}$$

19

h	e_H	EOC$_H$	e_{uH}	EOC$_{uH}$	e_{vH}	EOC$_{vH}$	e_u	EOC$_u$	e_v	EOC$_v$
10.000	1.25E−01	–	1.12E+00	–	1.11E+00	–	1.12E+00	–	1.11E+00	–
5.000	1.04E−01	0.27	7.64E−01	0.55	7.65E−01	0.53	7.69E−01	0.55	7.72E−01	0.53
2.500	4.06E−02	1.36	2.92E−01	1.39	2.84E−01	1.43	2.94E−01	1.39	2.87E−01	1.43
1.250	9.63E−03	2.08	1.09E−01	1.42	1.07E−01	1.41	1.10E−01	1.42	1.08E−01	1.41

Table 3: Vortex advection with DL-SVD($\mathbf{G}$): errors and EOC for each unknown.

h	e_H	EOC$_H$	e_{uH}	EOC$_{uH}$	e_{vH}	EOC$_{vH}$	e_u	EOC$_u$	e_v	EOC$_v$
10.000	1.25E−01	–	1.12E+00	–	1.11E+00	–	1.12E+00	–	1.12E+00	–
5.000	9.83E−02	0.34	6.67E−01	0.74	6.75E−01	0.72	6.72E−01	0.74	6.82E−01	0.71
2.500	2.37E−02	2.05	1.54E−01	2.12	1.57E−01	2.11	1.55E−01	2.12	1.59E−01	2.10
1.250	2.75E−03	3.11	2.44E−02	2.66	2.48E−02	2.66	2.45E−02	2.66	2.50E−02	2.67

Table 4: Vortex advection with DL-SVD($\mathbf{J}$): errors and EOC for each unknown.

h	e_H	EOC$_H$	e_{uH}	EOC$_{uH}$	e_{vH}	EOC$_{vH}$	e_u	EOC$_u$	e_v	EOC$_v$
10.000	1.25E−01	–	1.11E+00	–	1.09E+00	–	1.11E+00	–	1.10E+00	–
5.000	9.86E−02	0.34	6.90E−01	0.68	6.88E−01	0.67	6.94E−01	0.68	6.94E−01	0.66
2.500	2.36E−02	2.06	1.52E−01	2.19	1.46E−01	2.23	1.53E−01	2.19	1.48E−01	2.23
1.250	2.78E−03	3.09	2.45E−02	2.63	2.34E−02	2.64	2.47E−02	2.63	2.36E−02	2.65

Table 5: Vortex advection with DL-GS($\mathbf{J}$): errors and EOC for each unknown.

In the below numerical study, this system is solved using a slope-limited DG-$\mathbb{P}_1$ space discretization based on the HLL numerical flux (see [9, 23, 24] for details). Time integration is performed using the explicit third-order SSP Runge-Kutta method. At the end of each RK stage, the gradients of the conserved variables are adjusted using the DL-FIX($\mathbf{R}$) and DL-SVD($\mathbf{G}$) versions of the sequential limiting strategy presented in Section 6. Singular value decompositions are calculated using the LINPACK subroutine `dsvdc` [26]. Since all computations are performed on uniform rectangular meshes, we use $\mathbf{R} = \mathbf{I}$ assuming that the axes of the Cartesian coordinate system are already aligned with the orthogonal limiting directions for DL-FIX($\mathbf{R}$).

The use of slope limiters based on local maximum principles for scalar variables and frame-invariant projections of vector fields is typically sufficient to detect and eliminate spurious oscillations. However, additional fixes are generally required to guarantee preservation of global bounds corresponding to the invariant domains of the given hyperbolic system. According to Guermond et al. [15], the invariant set of the Euler system is defined by

$$\mathcal{A} = \{\mathbf{U} \mid \varrho > 0, \ e > 0, \ s \geq s^{\min}\},$$

where $e = E - \frac{1}{2}|\mathbf{u}|^2$ is the specific internal energy and $s^{\min}$ is a lower bound for the specific entropy $s(\varrho, e) = s_0 + \log\left(e^{\frac{1}{\gamma-1}}\varrho^{-1}\right)$. The equation of state (34) can be written as $p = (\gamma-1)\varrho e$. Hence, nonnegativity of internal energy is equivalent to nonnegativity of the pressure for a polytropic ideal gas.

Suppose that the cell averages are physically admissible, i.e., $\mathbf{U}_0^e \in \mathcal{A}$. Since the density ϱ_h is constrained using the scalar vertex-based limiter, the density cannot become negative. The bounds-preserving low-order targets for the momentum and total energy are defined by [9]

$$\overline{(\varrho\mathbf{u})}_h^{e,L} = \mathbf{u}_0^e(\bar{\varrho}_h^e - \varrho_0^e), \qquad \overline{(\varrho E)}_h^{e,L} = E_0^e(\bar{\varrho}_h^e - \varrho_0^e). \tag{37}$$

This definition implies that the internal energy $E_0^e - \frac{1}{2}|\mathbf{u}_0|^2$ remains constant and nonnegative on K^e. We also remark that the upper bounds $\varrho_i^{e,\max}$ for the vertex-based density limiter can be defined so that the specific entropy remains bounded below by $s^{\min}$ on K^e. Hence, the shape functions

$$\mathbf{U}_h^{e,L}(\mathbf{x}) = \mathbf{U}_0^e\frac{\bar{\varrho}_h^e(\mathbf{x})}{\varrho_0^e} = \mathbf{U}_0^e + \frac{\mathbf{U}_0^e\nabla\bar{\varrho}_h^e \cdot (\mathbf{x} - \mathbf{x}_0^e)}{\varrho_0^e}, \qquad \mathbf{x} \in K^e \tag{38}$$

stay in the invariant set $\mathcal{A}$. The gradients of the limited high-order targets

$$\mathbf{U}_h^{e,H}(\mathbf{x}) = \mathbf{U}_h^{e,L}(\mathbf{x}) + \bar{\mathbf{G}}^e(\mathbf{x} - \mathbf{x}_0^e), \qquad \mathbf{x} \in K^e \tag{39}$$

21

are determined by the choice of the correction factors for the momentum components and total energy. If a nonphysical state $\mathbf{U}_i^{e,H} = \mathbf{U}_h^{e,H}(\mathbf{x}_i) \notin \mathcal{A}$ is encountered at any vertex $\mathbf{x}_i^e$ of element K^e, the invariant domain preservation property can be easily enforced by setting $\overline{\mathbf{G}}^e := 0$ and accepting the gradient of $\mathbf{U}_h^{e,L}$ as the final product of sequential limiting. This simple fix traces its origins to 'failsafe' flux-corrected transport (FCT) algorithms [25, 37] in which negative pressures are eliminated by reverting to positivity-preserving low-order approximations in offending cells.

In summary, the proposed sequential limiting strategy for the Euler equations involves the following correction steps:

1. Limit the density ϱ_h^e using the scalar vertex-based limiter (and entropy-adjusted upper bounds to enforce $s \geq s^{\min}$ if desired).
2. Limit the momentum $(\varrho\mathbf{u})_h^e$ using the sequential vector limiter presented in Section 6 and a suitably defined rotation matrix $\mathbf{Q}^e$.
3. Limit the total energy $(\varrho E)_h^e$ using the sequential scalar limiter (cf. [9]) and local bounds defined in terms of E_0^e.
4. Overwrite the limited gradients of $(\varrho\mathbf{u})_h^e$ and $(\varrho E)_h^e$ by $\mathbf{u}_0^e \otimes \nabla\bar{\varrho}_h^e$ and $E_0^e \nabla\bar{\varrho}_h^e$, respectively, if $\overline{\mathbf{U}}_i^e \notin \mathcal{A}$ for any $i = 1, \ldots, N_{\mathrm{dof}}^e$.

The last step is a security check intended to prevent violations of global bounds defined by the invariant set $\mathcal{A}$. The corresponding *a posteriori* corrections need to be performed in just a small number of troubled cells (if any) where the use of low-order compatible targets is appropriate for safety reasons. In all other cells, the sequential limiter for the conserved variables will produce physically admissible and essentially nonoscillatory solutions.
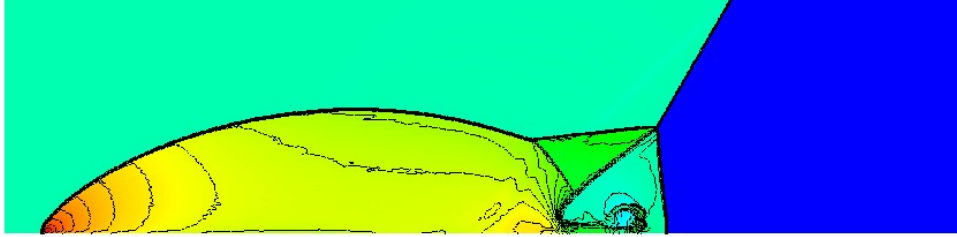
7.3.1. Double Mach reflection

The double Mach reflection problem [38] is a standard benchmark for the two-dimensional Euler equations. The computational domain for this test is the rectangle $\Omega = (0,4) \times (0,1)$. The flow pattern features a propagating Mach 10 shock in air ($\gamma = 1.4$) which initially makes a $60°$ angle with a reflecting wall. The following pre-shock and post-shock values of the flow variables are used to define the initial and boundary conditions

$$\begin{bmatrix} \varrho_L \\ u_L \\ v_L \\ p_L \end{bmatrix} = \begin{bmatrix} 8.0 \\ 8.25\cos(30°) \\ -8.25\sin(30°) \\ 116.5 \end{bmatrix}, \qquad \begin{bmatrix} \varrho_R \\ u_R \\ v_R \\ p_R \end{bmatrix} = \begin{bmatrix} 1.4 \\ 0.0 \\ 0.0 \\ 1.0 \end{bmatrix}. \tag{40}$$

Initially, the post-shock values (subscript L) are prescribed in the subdomain $\Omega_L = \{(x, y) \mid x < 1/6 + y/\sqrt{3}\}$ and the pre-shock values (subscript R) in $\Omega_R = \Omega \backslash \Omega_L$. The reflecting wall corresponds to $1/6 \leq x \leq 4$ and $y = 0$. No boundary conditions are required along the line $x = 4$. On the rest of the boundary, the post-shock conditions are assigned for $x < 1/6 + (1 + 20t)/\sqrt{3}$ and the pre-shock conditions elsewhere. The so-defined values along the top boundary describe the exact motion of the initial Mach 10 shock.
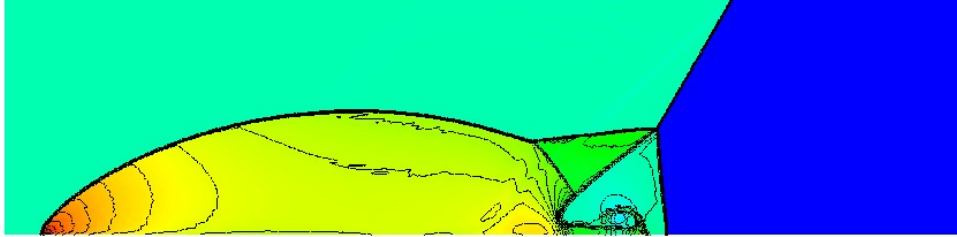
(a) DL-FIX($\mathbf{I}$)

(b) DL-SVD($\mathbf{G}$)

Figure 6: Double Mach reflection: ϱ at $t = 0.2$. DG-HLL-$\mathbb{P}_1$ discretization, sequential limiting + pressure fix, $h = \frac{1}{256}$, $\Delta t = 2.5 \cdot 10^{-5}$.

In Figs 6 and 7, we present snapshots of the density distribution at $t = 0.2$ calculated using the DL-FIX($\mathbf{I}$) and DL-SVD($\mathbf{G}$) versions of our limiting strategy. Positivity preservation is enforced using the invariant domain fix based on the use of linearized gradients $\mathbf{u}_0^e \otimes \nabla \bar{\varrho}_h^e$ and $E_0^e \nabla \bar{\varrho}_h^e$ in cells where negative pressures are found after sequential limiting of the conserved quantities. In this simulation, the number of such cells is 0 for many time steps and typically as small as 1-2 at time steps that do require a pressure fix. The numerical solutions produced by DL-FIX($\mathbf{I}$) and DL-SVD($\mathbf{G}$) look similar, so the higher cost of the latter version does not pay off in this example.
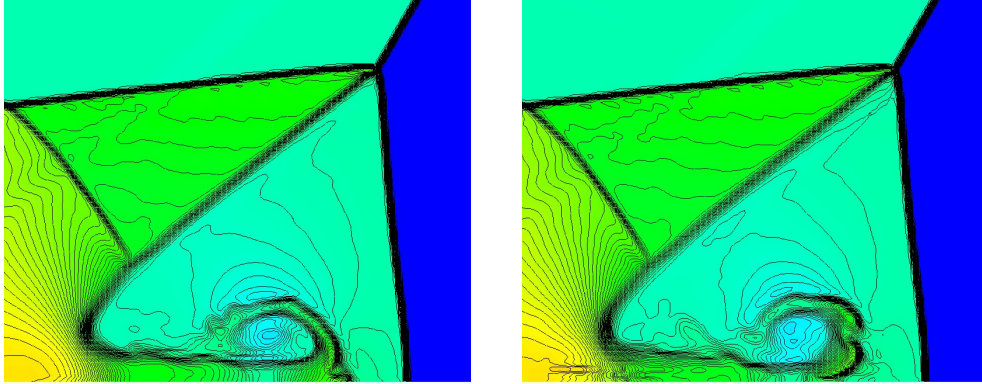
23

(a) DL-FIX($\mathbf{I}$)	(b) DL-SVD($\mathbf{G}$)

Figure 7: Double Mach reflection: zoom of ϱ at $t = 0.2$. DG-HLL-$\mathbb{P}_1$ discretization, sequential limiting + pressure fix, $h = \frac{1}{256}$, $\Delta t = 2.5 \cdot 10^{-5}$.

7.3.2. Smooth vortex

In the last numerical example, we consider the isentropic vortex problem [10, 35] which admits a smooth analytical solution. The adiabatic constant for the ideal gas equation of state (34) is $\gamma = 1.4$. Periodic boundary conditions are prescribed on the boundary of $\Omega = (-5, 5)^2$. The initial condition and exact solution at the final time $t = 10.0$ are given by [10, 35]

$$\varrho_0 = T_0^{\frac{1}{\gamma - 1}}, \qquad p_0 = \varrho_0 T_0, \qquad T_0 = 1 - \frac{(\gamma - 1)\beta^2}{8\gamma\pi^2} e^{1-r^2}, \qquad (41)$$

$$\mathbf{u}_0 = (1, 1) + \frac{\beta}{2\pi} e^{0.5(1-r^2)}(-y, x), \qquad (42)$$

where $\beta = 5.0$ is the vortex strength and $r = \sqrt{x^2 + y^2}$. We remark that β should be β^2 in the initial condition for the temperature in reference [10].

The density snapshots displayed in Fig. 8 were calculated using the HLL discontinuous Galerkin scheme equipped with the sequential DL-FIX($\mathbf{I}$) limiter. No visible changes in the symmetric shape of the density profile are observed as the center of the smooth vortex travels along the line $y = x$ with the mean velocity $(1, 1)$. The DL-SVD($\mathbf{G}$) results and the unconstrained $\mathbb{P}_1$ solutions look the same and, therefore, are not shown here. The L^1 convergence history presented in Table 6 reveals that the experimental order of convergence (EOC) of the piecewise-constant HLL scheme is very low. The $\mathbb{P}_1$ version exhibits second-order convergence in the absence of limiting.

24

The results of a grid convergence study for the DL-FIX($\mathbf{I}$) and DL-SVD($\mathbf{G}$) limiters are summarized in Table 7. The optional pressure fix is activated in both versions but never needs to be invoked in this test. In contrast to the numerical study presented in [9], both limiting techniques under investigation deliver at least second-order convergence rates without using any smoothness indicators. The DL-FIX($\mathbf{I}$) version produces the unconstrained DG-$\mathbb{P}_1$ solution on the finest meshes. The convergence of DL-SVD($\mathbf{G}$) is not quite as fast but second-order accuracy is achieved on fine meshes. The slightly better performance of DL-FIX($\mathbf{I}$) in terms of absolute L^1 errors can be explained by higher sensitivity of DL-SVD($\mathbf{G}$) to round-off errors and the choice of tolerances in numerical algorithms.

The ability of the new sequential limiters to preserve second-order accuracy of the DG-$\mathbb{P}_1$ scheme in this challenging test can be attributed to the fact that they enforce local maximum principles for the conserved quantities and are linearity-preserving w.r.t. these quantities. The bounds that guarantee preservation of invariant domains are global and do not affect the rates of
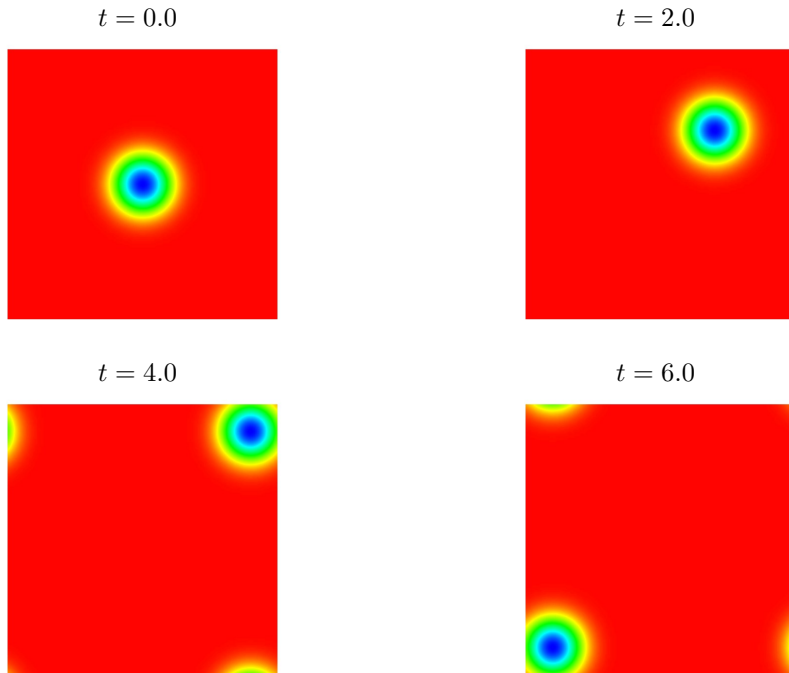
$t = 0.0$ $t = 2.0$

$t = 4.0$ $t = 6.0$

Figure 8: Smooth vortex: snapshots of the evolving density. DG-HLL-$\mathbb{P}_1$ discretization, DL-FIX($\mathbf{I}$) sequential limiting + pressure fix, $h = \frac{10}{128}$, $\Delta t = 5 \cdot 10^{-3}$.

convergence to smooth solutions. The imposition of local bounds on the specific kinetic and total energy (as proposed in [9]) leads to far more restrictive constraints and tends to degrade the rates of convergence if no smoothness indicators are used to avoid unnecessary limiting in smooth cells.

h	$\mathbb{P}_0$	EOC	$\mathbb{P}_1$	EOC
10/32	0.256E+01		0.215E+00	
10/64	0.227E+01	0.17	0.445E–01	2.27
10/128	0.180E+01	0.33	0.894E–02	2.32
10/256	0.125E+01	0.53	0.199E–02	2.17

Table 6: Smooth vortex: $E_1^\varrho = \|\varrho - \varrho_h\|_{L^1(\Omega)}$ for $\mathbb{P}_0$ and $\mathbb{P}_1$ without limiting.

h	DL-FIX($\mathbf{I}$)	EOC	DL-SVD($\mathbf{G}$)	EOC
10/32	0.343E+00		0.467E+00	
10/64	0.516E–01	2.73	0.134E+00	1.80
10/128	0.895E–02	2.53	0.313E–01	2.10
10/256	0.199E–02	2.17	0.767E–02	2.03

Table 7: Smooth vortex: $E_1^\varrho = \|\varrho - \varrho_h\|_{L^1(\Omega)}$ for $\mathbb{P}_1$ with sequential limiting.

8. Conclusions

The proposed approaches to directional limiting of vector fields are based on straightforward extensions of scalar limiting techniques. The presented numerical study indicates that a good frame-invariant vector limiter does not need to be very sophisticated. In fact, the use of scalar limiters for individual components of the vector field may be the best limiting strategy for rectangular domains and uniform meshes. Indeed, componentwise limiting in a global Cartesian reference frame does not require any additional transformations of variables and is objective if the orientation of this reference frame is defined by the mesh or by the computational domain. The use of local reference frames depending to the shape and orientation of mesh elements may be a good strategy for unstructured and anisotropic meshes. The use of limiting directions based on singular value decompositions of tensor-valued solution gradients is computationally intensive, and the associated cost is unlikely to pay off unless the solution is highly anisotropic. Regardless of the choice of

26

orthogonal limiting directions, no problems with preservation of radial symmetry were observed in numerical experiments. The multidimensional nature of our finite element schemes and vertex-based limiters seems to cure these potential problems just as well as the use of generalized maximum principles based on convex hull / VIP criteria or bounding boxes [28, 29, 30].

The proposed methodology provides a general framework for directional vector limiting and great flexibility in the choice of frame-invariant limiting directions. It is readily portable to continuous finite element approximations (see [9] for a presentation of the vertex-based limiter in a format suitable for this purpose) and flux-corrected remapping of vector quantities in arbitrary Lagrangian-Eulerian (ALE) methods [36, 39] for multimaterial flows.

The outcomes of this work also include a better understanding of the interplay between local and global constraints for different quantities of interest in high-resolution finite element schemes for systems of conservation laws, such the shallow water equations and the Euler equations considered in this work. In our experience, any attempt to impose **local** bounds on nonlinear functions of the conserved variables (pressure, internal or kinetic energy, entropy) inhibits the ability of the limiting procedure to deliver optimal convergence rates for smooth data. On the other hand, mere preservation of invariant domains does not guarantee the absence of undershoots/overshoots (or even convergence to correct weak solutions [14]). Therefore, a well-designed limiter for hyperbolic systems should employ both local bounds based on certain smoothness criteria and global bounds based on the knowledge of invariant domains of the exact solution or other physical considerations. If the use of numerical fluxes defined in terms of limited P_1 shape functions results in a violation of global bounds for cell averages, the invariant domain preservation property should be enforced using FCT-like flux limiters (cf. [33, 37]).

Acknowledgments

References

[1] V. Aizinger, A geometry independent slope limiter for the discontinuous Galerkin method. In: *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, **115** (2011) 207–217.

[2] V. Aizinger and C. Dawson, A discontinuous Galerkin method for two-dimensional flow and transport in shallow water. *Advances in Water Resources* **25** (2002) 67–84.

[3] V. Aizinger, A. Kosík, D. Kuzmin and B. Reuter, Anisotropic slope limiting for discontinuous Galerkin methods. *Int. J. Numer. Methods Fluids* **84** (2017) 543-565.

[4] P. Azerad, J.-L. Guermond and B. Popov, Well-balanced second-order approximation of the shallow water equation with continuous finite elements. *SIAM J. Numer. Anal.* **55** (2017) 3203–3224.

[5] S. Badia, J. Bonilla and A. Hierro, Differentiable monotonicity-preserving schemes for discontinuous Galerkin methods on arbitrary meshes. *Computer Methods Appl. Mech. Engrg.* **320** (2017) 582–605.

[6] T. Barth and D.C. Jespersen, The design and application of upwind schemes on unstructured meshes. *AIAA Paper*, 89-0366, 1989.

[7] B. Cockburn and C.-W. Shu, The Runge-Kutta discontinuous Galerkin method for conservation laws V. Multidimensional Systems. *J. Comput. Phys.* **141** (1998) 199–224.

[8] O. Delestre, C. Lucas, P.-A. Ksinant, F. Darboux, C. Laguerre, T.-N. Vo, F. James, S. Cordier and others, SWASHES: a compilation of shallow water analytic solutions for hydraulic and environmental studies. *Int. J. Numerical Methods in Fluids* **72** (2013) 269–300.

[9] V. Dobrev, Tz. Kolev, D. Kuzmin, R. Rieben and V. Tomov, Sequential limiting in continuous and discontinuous Galerkin methods for the Euler equations. *J. Comput. Phys.* **356** (2018) 372–390.

[10] M. Dumbser, O. Zanotti, R. Loubère and S. Diot, A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *J. Comput. Phys.* **278** (2014) 47–75.

[11] U. S. Fjordholm, *Structure preserving finite volume methods for the shallow water equations.* M.Sc. thesis, University of Oslo, 2009.

[12] F. Frank, B. Reuter, V. Aizinger and P. Knabner, FESTUNG: A MATLAB / GNU Octave toolbox for the discontinuous Galerkin method, Part I: Diffusion operator. *Computers & Mathematics with Applications* **70** (2015) 11–46.

[13] S. Gottlieb, C.-W. Shu and E. Tadmor, Strong stability-preserving high-order time discretization methods. *SIAM Review* **43** (2001) 89–112.

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

[15] J.-L. Guermond, M. Nazarov, B. Popov and I. Tomas, Second-order invariant domain preserving approximation of the Euler equations using convex limiting. `arXiv: 1710.00417v1 [math.NA] 1 Oct 2017`

[16] H. Hajduk, *Numerical Investigation of Direct Bathymetry Reconstruction Based on a Modified Shallow-Water Model.* M.Sc. thesis, University Erlangen-Nuremberg, 2017.

[17] H. Hajduk, B. R. Hodges, V. Aizinger, B. Reuter, Locally Filtered Transport for computational efficiency in multi-component advection-reaction models. In: *Environmental Modelling & Software*, **102** (2018) 185–198.

[18] Ph. Hoch and E. Labourasse, A frame invariant and maximum principle enforcing second-order extension for cell-centered ALE schemes based on Local Convex Hull Preservation. *Int. J. Numer. Methods Fluids* **76** (2014) 1043-1063.

[19] H. Hoteit, Ph. Ackerer, R. Mosé, J. Erhel and B. Philippe, New two-dimensional slope limiters for discontinuous Galerkin methods on arbitrary meshes. *Int. J. Numer. Meth. Engrg.* **61** (2004) 2566–2593.

[20] A. Jaust, B. Reuter, V. Aizinger, J. Schütz and P. Knabner, FESTUNG: A MATLAB / GNU Octave toolbox for the discontinuous Galerkin method, Part III: Hybridized discontinuous Galerkin (HDG) formulation. *Computers & Mathematics with Applications*, 2018 (accepted).

[21] L. Krivodonova, Limiters for high-order discontinuous Galerkin methods. *J. Comput. Phys.* **226** (2007) 879–896.

[22] D. Kuzmin, A vertex-based hierarchical slope limiter for p-adaptive discontinuous Galerkin methods. *J. Comput. Appl. Math.* **233** (2010) 3077–3085.

[23] D. Kuzmin, Hierarchical slope limiting in explicit and implicit discontinuous Galerkin methods. *J. Comput. Phys.* **257** (2014) 1140–1162.

[24] D. Kuzmin and C. Lohmann, Synchronized slope limiting in discontinuous Galerkin methods for the equations of gas dynamics. *Ergebnisber. Inst. Angew. Math.* **541**, TU Dortmund, 2016.

[25] D. Kuzmin, M. Möller, J.N. Shadid and M. Shashkov, Failsafe flux limiting and constrained data projections for equations of gas dynamics. *J. Comput. Phys.* **229** (2010) 8766–8779.

[26] LINPACK, `http://www.netlib.org/linpack/`.

[27] C. Lohmann and D. Kuzmin, Synchronized flux limiting for gas dynamics variables. *J. Comput. Phys.* **326** (2016) 973–990.

[28] G. Luttwak and J. Falcovitz, Slope limiting for vectors: A novel vector limiting algorithm. *Int. J. Numer. Methods Fluids* **65** (2011) 1365–1375.

[29] G. Luttwak and J. Falcovitz, VIP (Vector Image Polygon) multi-dimensional slope limiters for scalar variables. *Computers & Fluids* **83** (2013) 90-97.

[30] G. Luttwak, On the extension of monotonicity to multi-dimensional flows. Manuscript, `http://dy123d.com/gabilut_files/papers/monotonicity_in_multidim.pdf`

[31] P.-H. Maire, A high-order one-step sub-cell force-based discretization for cell-centered Lagrangian hydrodynamics on polygonal grids. *Computers & Fluids* **46** (2011) 341–347.

[32] P.-H. Maire, R. Loubère and P. Váchal, Staggered Lagrangian discretization based on cell-centered Riemann solver and associated hydrodynamics scheme. *Commun. Comput. Phys.* **10** (2011) 940–978.

[33] S.A. Moe, J.A. Rossmanith and D.C. Seal, Positivity-preserving discontinuous Galerkin methods with Lax-Wendroff time discretizations. *Journal of Scientific Computing* **71** (2017) 44-70.

[34] B. Reuter, V. Aizinger, M. Wieland, F. Frank and P. Knabner, FESTUNG: A MATLAB / GNU Octave toolbox for the discontinuous Galerkin method. Part II: Advection operator and slope limiting. *Computers and Mathematics with Applications* **72** (2016) 1896–1925.

[35] C.-W. Shu, Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. NASA/CR-97-206253, *ICASE Report* No. 97-65, 1997.

[36] J. Velechovsky, M. Kucharik, R. Liska, M. Shashkov and P. Vachal, Symmetry- and essentially-bound-preserving flux-corrected remapping of momentum in staggered ALE hydrodynamics. *J. Comput. Phys.* (2013) **255** (2013) 590–611.

[37] P.R. Woodward and P. Colella, The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.* **54** (1984) 115–173.

[38] S.T. Zalesak, The design of Flux-Corrected Transport (FCT) algorithms for structured grids. In: D. Kuzmin, R. Löhner, S. Turek (eds), *Flux-Corrected Transport: Principles, Algorithms, and Applications.* Springer, 2nd edition, 2012, pp. 23-66.

[39] X. Zeng and G. Scovazzi, A frame-invariant vector limiter for flux corrected nodal remap in arbitrary Lagrangian-Eulerian flow computations. *J. Comput. Phys.* **270** (2014) 753-783.

[40] X. Zhang and C-W. Shu, Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. A* **467** (2011) 2752–2776.