

No. 589

July 2018

Gradient-based limiting and stabilization of
continuous Galerkin methods

D. Kuzmin

ISSN: 2190-1767

Gradient-based limiting and stabilization of continuous Galerkin methods

Dmitri Kuzmin

Abstract

In this paper, we stabilize and limit continuous Galerkin discretizations of a linear transport equation using an algebraic approach to derivation of artificial diffusion operators. Building on recent advances in the analysis and design of edge-based algebraic flux correction schemes for singularly perturbed convection-diffusion problems, we derive algebraic stabilization operators that generate nonlinear high-order stabilization in smooth regions and enforce discrete maximum principles everywhere. The correction factors for antidiffusive element or edge contributions are defined in terms of nodal gradients that vanish at local extrema. The proposed limiting strategy is linearity-preserving and provides Lipschitz continuity of constrained terms. Numerical examples are presented for two-dimensional test problems.

Keywords: *hyperbolic conservation laws, finite element methods, discrete maximum principles, algebraic flux correction, linearity preservation*

1.1 Introduction

Bound-preserving discretizations of hyperbolic conservation laws and convection-dominated transport problems use limiting techniques to enforce discrete maximum principles. Recent years have witnessed an increased interest of the finite element community in algebraic flux correction (AFC) schemes [9] based on various generalizations of flux-corrected transport (FCT) algorithms and total variation diminishing (TVD) methods. A major breakthrough in the theoretical analysis of AFC for continuous finite elements was achieved by Barrenechea et al. [3, 4] whose recent work has provided a set of design principles for derivation of limiters that lead

TU Dortmund University, Institute of Applied Mathematics (LS III), Vogelpothsweg 87, 44227 Dortmund, Germany

to well-posed nonlinear problems in the context of stationary convection-diffusion equations. Limiting techniques for continuous Galerkin discretizations of hyperbolic problems were proposed in [2, 7, 12]. As shown in [7, 12], the use of the standard Galerkin method as the AFC target for hyperbolic conservation laws may give rise to bounded ripples and nonphysical weak solutions. In fact, the Galerkin discretization may even produce singular matrices on criss-cross (Union Jack) meshes [13]. The use of limiters restricts the range of possible solution values but does not rule out spurious oscillations within this range. In this paper, we design artificial diffusion operators that introduce high-order stabilization in smooth regions and enforce preservation of local bounds in the vicinity of steep fronts. The element or edge contributions to the residual of the nonlinear system are constrained using limiters defined in terms of nodal gradients rather than nodal correction factors. This approach leads to a limiting procedure that satisfies all essential design criteria.

1.2 Artificial diffusion operators

To make the presentation self-contained, we begin with an outline of the basic AFC methodology [9] for C^0 finite element discretizations of the hyperbolic equation

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) = 0 \quad \text{in } \Omega \quad (1.1)$$

to be solved in a bounded domain Ω with a Lipschitz boundary Γ . The velocity field \mathbf{v} is assumed to be known. At the inlet $\Gamma_{\text{in}} = \{\mathbf{x} \in \Gamma : \mathbf{v} \cdot \mathbf{n} < 0\}$, we impose the boundary condition $u = 0$ in a weak sense by using the variational formulation

$$\int_{\Omega} w \frac{\partial u}{\partial t} d\mathbf{x} + \int_{\Gamma} w u \max\{0, \mathbf{v} \cdot \mathbf{n}\} ds - \int_{\Omega} \nabla w \cdot (\mathbf{v}u) d\mathbf{x} = 0, \quad (1.2)$$

where \mathbf{n} is the unit outward normal and $w \in H^1(\Omega)$ is an admissible test function. The numerical solution $u_h = \sum_{j=1}^N u_j \varphi_j$ is defined in terms of continuous piecewise-linear or multilinear Lagrange basis functions φ_j associated with vertices \mathbf{x}_j of a mesh (alias *triangulation*) \mathcal{T}_h . The standard Galerkin discretization leads to

$$M_C \frac{du}{dt} + Au = 0, \quad (1.3)$$

where $M_C = \{m_{ij}\}$ is the consistent mass matrix, $A = \{a_{ij}\}$ is the discrete transport operator, and $u = \{u_i\}$ is the vector of time-dependent nodal values. Introducing the lumped mass matrix $M_L = \{\delta_{ij} \sum_j m_{ij}\}$ and a symmetric artificial diffusion operator $D = \{d_{ij}\}$, we construct the low-order approximation

$$M_L \frac{du}{dt} + (A - D)u = 0 \quad (1.4)$$

which is provably bound-preserving if $\sum_j d_{ij} = 0$ and $d_{ij} \geq \max\{a_{ij}, 0, a_{ji}\}$ for all $j \neq i$ [4, 9]. The original Galerkin discretization (1.3) can be written as

$$M_L \frac{du}{dt} + (A - D)u = f \left(u, \frac{du}{dt} \right), \quad f_i = \sum_m f_i^m, \quad (1.5)$$

where $f = \{f_i\}$ is the antidiffusive part that requires limiting. In edge-based AFC schemes, f_i^m is the contribution of edge m to node i , and there exists a neighbor node $j \neq i$ such that $f_j^m = -f_i^m$ [4, 7, 9]. In element-based versions, f_i^m is the contribution of element m to node i and $\sum_i f_i^m = 0$ by definition [10, 12]. In the 1D case, the decompositions of f into edge and element contributions are equivalent.

In the process of limiting, each component f_i^m is multiplied by a solution-dependent correction factor $\alpha^m \in [0, 1]$. This modification leads to the nonlinear system

$$M_L \frac{du}{dt} + (A - D)u = \bar{f} \left(u, \frac{du}{dt} \right), \quad \bar{f}_i = \sum_m \alpha^m f_i^m. \quad (1.6)$$

We define f_i^m and α^m in the next section. The discretization in time can be performed using a strong stability preserving (SSP) Runge-Kutta method [6]. Note that only the backward Euler method is SSP without any restrictions on the time step.

1.3 Limiting of antidiffusive terms

First and foremost, the definition of correction factors α^m should guarantee that the limited antidiffusive term \bar{f}_i be *local extremum diminishing* (LED), i.e.,

$$u_i^{\max} := \max_{j \in \mathcal{N}_i} u_j = u_i \quad \Rightarrow \quad \bar{f}_i \leq 0, \quad (1.7)$$

$$u_i^{\min} := \min_{j \in \mathcal{N}_i} u_j = u_i \quad \Rightarrow \quad \bar{f}_i \geq 0, \quad (1.8)$$

where $\mathcal{N}_i = \{j \in \{1, \dots, N\} : m_{ij} \neq 0\}$ is the computational stencil of node i .

Obviously, the LED property (1.7),(1.8) holds for α^m satisfying (cf. [3, 5])

$$\alpha^m \leq \alpha_i := \min \left\{ 1, \frac{\gamma_i \min\{u_i^{\max} - u_i, u_i - u_i^{\min}\}}{\max\{u_i^{\max} - u_i, u_i - u_i^{\min}\} + \varepsilon h} \right\} \quad \forall i \in \mathcal{N}^m, \quad (1.9)$$

where \mathcal{N}^m is the set of nodes belonging to the element or edge, $\gamma_i > 0$ is a parameter to be defined in Section 1.3.2, h is the mesh size, and ε is a small positive constant.

Theoretical and numerical studies of AFC schemes indicate that the use of linearity-preserving limiters is an essential prerequisite for achieving optimal accuracy on general meshes [3, 5, 9, 10]. The bound α_i in formula (1.9) is linearity preserving if $\alpha_i = 1$ whenever u_h is linear on the patch $\bar{\Omega}_i = \{K \in \mathcal{T}_h : \mathbf{x}_i \in K\}$ of elements containing an internal node $\mathbf{x}_i \in \Omega$. According to the analysis of Barrenechea et al.

[3, 5], the nodal correction factor α_i defined by (1.9) possesses this property if

$$\gamma_i \geq \gamma_i^{\min} := \frac{\max_{\mathbf{x}_j \in \partial \Omega_i} |\mathbf{x}_i - \mathbf{x}_j|}{\text{dist}\{\mathbf{x}_i, \partial \Omega_i^{\text{conv}}\}}, \quad (1.10)$$

where $|\cdot|$ is the Euclidean norm and Ω_i^{conv} is the convex hull of points $\mathbf{x}_j \in \bar{\Omega}_i$.

Implicit time integration requires iterative solution of nonlinear systems and only converged solutions are guaranteed to be bound preserving. Therefore, convergence behavior of iterative solvers should also be taken into account. It is essential to guarantee that each product $\alpha^m f_i^m$ be a Lipschitz-continuous function of nodal values. This property is used in proofs of existence and uniqueness for the nonlinear discrete problem [3, 4] and secures convergence of fixed-point iterations based on deferred correction methods (see [1], Proposition 4.3). Faster convergence can be achieved, e.g., using Anderson acceleration [9] or differentiable LED limiters [2].

1.3.1 Nonlinear high-order stabilization

The straightforward choice $\alpha^m = \min_{i \in \mathcal{N}^m} \alpha_i$ of the correction factor α^m for f_i^m corresponds to using the oscillatory Galerkin scheme (1.3) as the limiting target. In this section, we construct a stabilized AFC scheme using a definition of α^m in terms of limited nodal gradients \mathbf{g}_i^* such that $\mathbf{g}_i^* = 0$ if $u_i = u_i^{\max}$ or $u_i = u_i^{\min}$ at an internal node $\mathbf{x}_i \in \Omega$. Additionally, the gradient recovery procedure should be exact for linear functions. In Section 1.3.2, we use nodal correction factors α_i of the form (1.9) to correct a linearity-preserving gradient reconstruction \mathbf{g}_i and obtain a Lipschitz-continuous approximation $\mathbf{g}_i^* = \alpha_i \mathbf{g}_i$ satisfying the above requirements.

An element-based AFC scheme with a stabilized high-order target is defined by

$$\alpha^m = \left(\frac{\min\{\|\nabla u_h^m\|_{C(K^m)}, p \min_{i \in \mathcal{N}^m} |\mathbf{g}_i^*|\}}{\|\nabla u_h^m\|_{C(K^m)} + \varepsilon} \right)^q, \quad (1.11)$$

where $\|\cdot\|_{C(K)}$ is the maximum norm. The parameters $p \geq 1$ and $q \in \mathbb{N}$ act as steepeners that make the limiter less diffusive. By default, we use $p = q = 2$.

If a local extremum is attained at node i for any $i \in \mathcal{N}^m$, then $|\mathbf{g}_i^*| = 0$ and, therefore, $\alpha^m = 0$ in accordance with the LED criterion. If u_h is linear on $\bar{\Omega}_i$ and the parameter γ_i is defined by (1.16), then $\alpha_i = 1$ and, therefore, $\mathbf{g}_i^* = \mathbf{g}_i = \nabla u_h^m$. In general, our formula (1.11) will produce $\alpha^m = 1$ if the magnitude of ∇u_h^m does not exceed that of the smallest nodal gradient by more than a factor of p . Lipschitz continuity of $\alpha^m f_i^m$ can be shown following Lohmann's [11] proofs for edge-based tensor limiters.

An edge-based counterpart of our gradient-based limiter (1.11) is defined by

$$\alpha^m = \left(\frac{\min\{p|\mathbf{g}_i^* \cdot (\mathbf{x}_i - \mathbf{x}_j)|, |u_i - u_j|, p|\mathbf{g}_j^* \cdot (\mathbf{x}_i - \mathbf{x}_j)|\}}{|u_i - u_j| + \varepsilon} \right)^q. \quad (1.12)$$

A proof of Lipschitz continuity for $q \in \mathbb{N}$ follows from Lohmann's analysis [11].

If the gradient is nonsmooth, our method may produce $\alpha^m < 1$ even in the case when $\mathbf{g}_i^* = \mathbf{g}_i$ for all i . In contrast to limiters designed to recover the standard Galerkin discretization whenever it satisfies the LED constraints, our definition of α^m generates nonlinear high-order dissipation in elements free of local extrema. On a uniform mesh of 1D linear elements both (1.11) and (1.12) lead to a *symmetric limited positive* (SLIP) scheme [8] that switches between second- and fourth-order dissipation. In predictor-corrector algorithms of FCT type, high-order dissipation can also be generated by adding nonlinear entropy viscosity [7] or linear gradient-based stabilization [10, 12]. However, the use of such artificial diffusion operators in iterative AFC schemes inhibits convergence due to the lack of Lipschitz continuity.

1.3.2 Recovery of nodal gradients

If u_h is linear on $\bar{\Omega}_i$, then $\mathbf{g}_i = \nabla u_h(\mathbf{x}_i)$ holds for any weighted average \mathbf{g}_i of the one-sided element gradients $\nabla u_h|_K(\mathbf{x}_i)$, $K \in \bar{\Omega}_i$. For example, the global lumped-mass L^2 projection yields the averaged nodal gradient [9]

$$\mathbf{g}_i = \frac{1}{m_i} \sum_{j \in \mathcal{N}_i} \mathbf{c}_{ij} u_j = \frac{1}{m_i} \sum_{j \in \mathcal{N}_i \setminus \{i\}} \mathbf{c}_{ij} (u_j - u_i), \quad (1.13)$$

$$m_i = \int_{\Omega} \phi_i d\mathbf{x}, \quad \mathbf{c}_{ij} = \int_{\Omega} \phi_i \nabla \phi_j d\mathbf{x}. \quad (1.14)$$

However, the so-defined \mathbf{g}_i does not necessarily vanish if a local maximum or minimum is attained at $\mathbf{x}_i \in \Omega$. To rectify this, we consider the limited gradient

$$\mathbf{g}_i^* = \alpha_i \mathbf{g}_i = \frac{1}{m_i} \sum_{j \in \mathcal{N}_i \setminus \{i\}} \mathbf{c}_{ij} \alpha_i (u_j - u_i), \quad (1.15)$$

where α_i is the nodal correction factor defined by (1.9). The limited gradient reconstruction \mathbf{g}_i^* does vanish at local extrema. Lipschitz continuity of $\alpha_i(u_j - u_i)$ can be shown using Lemma 6 in [4]. Linearity preservation is guaranteed under condition (1.10) since $\alpha_i = 1$ if u_h is linear on $\bar{\Omega}_i$. The use of the sharp bound $\gamma_i := \gamma_i^{\min}$ defined by (1.10) requires calculation of the distance to the convex hull and leads to rather diffusive *minmod* limiters like the one proposed in [5]. To simplify the formula for γ_i and make the LED constraints less restrictive, we define γ_i as follows.

The anisotropy of a mesh element $K \in \mathcal{T}_h$ can be characterized by the ratio of the local mesh size $h_K = \text{diam}(K)$ and the diameter ρ_K of the largest ball that fits into K . A family of triangulations $\{\mathcal{T}_h\}$ is called regular if there exists a constant $\sigma > 0$ such that $\frac{h_K}{\rho_K} \leq \sigma$ for all $K \in \mathcal{T}_h$ and all h . For triangular elements, ρ_K is the diameter of the inscribed circle. Since $h_K \geq \max_{\mathbf{x}_j \in K} |\mathbf{x}_i - \mathbf{x}_j|$ and $\rho_K \leq \text{dist}\{\mathbf{x}_i, \partial K \cap \partial \Omega_i^{\text{conv}}\}$ for

all $K \in \bar{\Omega}_i$, condition (1.10) holds for parameters $\gamma_i = \gamma_i(s)$, $s \geq 1$ defined by

$$\gamma_i = s \frac{\max_{K \in \bar{\Omega}_i} h_K}{\min_{K \in \bar{\Omega}_i} \rho_K}. \quad (1.16)$$

A reasonable default setting for iterative AFC schemes is $s = 2$. The limiter becomes less diffusive as s is increased but the use of $\gamma_i \gg \gamma_i^{\min}$ may cause convergence problems when it comes to iterative solution of nonlinear systems.

1.3.3 Recovery of nodal time derivatives

By (1.3) and (1.6), the term to be limited is given by $f(u, \dot{u}) = (M_L - M_C)\dot{u} - Du$. Using the Neumann series approximation [7] to M_C^{-1} , we obtain

$$\dot{u} = M_L^{-1}(I + (M_L - M_C)M_L^{-1})(\bar{f}(u, 0) - (A - D)u). \quad (1.17)$$

This definition of \dot{u} makes it possible to determine $\bar{f}(0, \dot{u})$ without recalculating $\bar{f}(u, 0)$. Moreover, the correct steady state behavior is preserved for $\dot{u} = 0$.

In our AFC scheme for time-dependent problems, we limit $f_i^m(u, 0)$ using α^m defined in Section 1.3.1. To provide continuity, the contribution of $f_i^m(0, \dot{u})$ is limited using a correction factor $\dot{\alpha}^m$ such that $\dot{\alpha}^m = 0$ if $\alpha^m = 0$ (see below). In the element-based version of (1.6), the limited antidiffusive components are given by

$$\bar{f}_i^m(u, \dot{u}) = \sum_{j \in \mathcal{N}^m \setminus \{i\}} [m_{ij}^m \dot{\alpha}^m (\dot{u}_i - \dot{u}_j) + d_{ij}^m \alpha^m (u_i - u_j)], \quad i \in \mathcal{N}^m. \quad (1.18)$$

The coefficients m_{ij}^m and d_{ij}^m represent the contribution of $K^m \in \mathcal{T}_h$ to the global matrix entries m_{ij} and d_{ij} , respectively. Algebraic residual correction schemes based on such decompositions into element contributions can be found in [10, 12].

The evolutionary part $f_i^m(0, \dot{u})$ is constrained using correction factors of the form

$$\dot{\alpha}^m = \min \left\{ 1, \beta^m \frac{\alpha^m \|\nabla u_h^m\|_{C(K^m)}}{\|\nabla \dot{u}_h^m\|_{C(K^m)} + \varepsilon} \right\} \quad (1.19)$$

such that $(\dot{\alpha}^m \|\nabla \dot{u}_h^m\|_{C(K^m)}) / (\alpha^m \|\nabla u_h^m\|_{C(K^m)}) \leq \beta^m$, where $\beta^m > 0$ should have units of the reciprocal second s^{-1} . In this work, we use $\beta^m = \|\mathbf{v}\|_{C(K^m)} / (2h_{K^m})$.

In the edge-based version of (1.6), we limit $f_i^m = f_{ij}$ and $f_j^m = -f_{ij}$ as follows:

$$\bar{f}_i^m(u, \dot{u}) = m_{ij} \dot{\alpha}^m (\dot{u}_i - \dot{u}_j) + d_{ij} \alpha^m (u_i - u_j), \quad \{i, j\} =: \mathcal{N}^m, \quad (1.20)$$

$$\dot{\alpha}^m = \min \left\{ 1, \beta^m \frac{\alpha^m |u_i - u_j|}{|\dot{u}_i - \dot{u}_j| + \varepsilon} \right\}. \quad (1.21)$$

In pseudo-time-stepping schemes for steady-state computations, we use $\dot{\alpha}^m = 0$.

1.4 Numerical examples

In Figures 1 and 2, we present the AFC results for the time-dependent solid body rotation benchmark and a stationary circular convection test. For the formulation of the corresponding (initial-)boundary value problems, we refer the reader to [9, 10]. In this numerical study, we use the element-based version of (1.6). The stationary problem is solved using implicit pseudo-time-stepping and a fixed-point iteration method [9]. The employed parameter settings and discretization parameters are summarized in the captions. The constrained Galerkin solutions satisfy local maximum principles if $\alpha^m = \alpha^m(\mathbf{g}_h^*)$ is defined by (1.11). To assess the amount of intrinsic high-order stabilization, we also present the numerical solutions obtained using the target scheme ($\alpha^m = \alpha^m(\mathbf{g}_h)$ and $\dot{\alpha}^m := \alpha^m$ in the unsteady case). The L^1 convergence rates for the circular convection test without the discontinuous portion of the inflow profile are shown in Table 1. The AFC scheme based on $\alpha^m = \alpha^m(\mathbf{g}_h^*)$ exhibits second-order superconvergence on uniform meshes. The convergence rates on perturbed meshes are comparable to those for $\alpha^m = \alpha^m(\mathbf{g}_h)$ and higher than the optimal order 1.5 for continuous P_1 finite element discretizations of (1.1).

h	uniform, $\alpha^m(\mathbf{g}_h^*)$	EOC	perturbed, $\alpha^m(\mathbf{g}_h^*)$	EOC	perturbed, $\alpha^m(\mathbf{g}_h)$	EOC
1/32	0.185E-01		0.150E-01		0.141E-01	
1/64	0.511E-02	1.85	0.473E-02	1.67	0.446E-02	1.66
1/128	0.117E-02	2.14	0.155E-02	1.61	0.133E-02	1.75
1/256	0.256E-03	2.19	0.499E-03	1.64	0.438E-03	1.60

Table 1.1 Circular convection: L^1 convergence history for triangular meshes consisting of $2/h^2$ cells, P_1 approximation, AFC scheme based on (1.11), $s = 2$, smooth inflow profile.

Acknowledgments: This research was supported by the German Research Association (DFG) under grant KU 1530/23-1. The author would like to thank Christoph Lohmann (TU Dortmund University) for helpful discussions and suggestions.

References

1. R. Abgrall, High order schemes for hyperbolic problems using globally continuous approximation and avoiding mass matrices. *J. Sci. Comput.* **73** (2017) 461–494.
2. S. Badia and J. Bonilla, Monotonicity-preserving finite element schemes based on differentiable nonlinear stabilization. *Computer Methods Appl. Mech. Engrg.* **313** (2017) 133–158.
3. G. Barrenechea, V. John, P. Knobloch, and R. Rankin, A unified analysis of algebraic flux correction schemes for convection-diffusion equations. SeMA (2018).
<https://doi.org/10.1007/s40324-018-0160-6>
4. G. Barrenechea, V. John, and P. Knobloch, Analysis of algebraic flux correction schemes. *SIAM J. Numer. Anal.* **54** (2016) 2427–2451.
5. G. Barrenechea, V. John, and P. Knobloch, A linearity preserving algebraic flux correction scheme satisfying the discrete maximum principle on general meshes. *Mathematical Models and Methods in Applied Sciences (M3AS)* **27** (2017) 525–548.

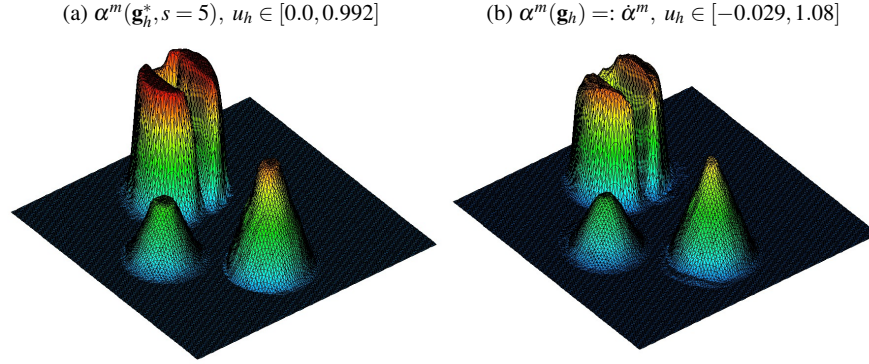


Fig. 1.1 Solid body rotation: uniform triangular mesh, 2×128^2 P_1 elements, AFC scheme based on (1.11), explicit second-order SSP-RK time-stepping, $\Delta t = 10^{-3}$, solutions at $T = 2\pi$.

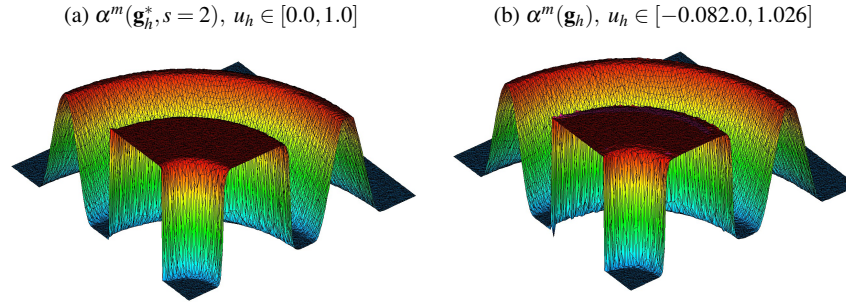


Fig. 1.2 Circular convection: perturbed triangular mesh, 2×128^2 P_1 elements, AFC scheme based on (1.11), backward Euler pseudo-time-stepping, steady-state solutions.

6. S. Gottlieb, C.-W. Shu, and E. Tadmor, Strong stability-preserving high-order time discretization methods. *SIAM Review* **43** (2001) 89–112.
7. J.-L. Guermond, M. Nazarov, B. Popov, and Y. Yang, A second-order maximum principle preserving Lagrange finite element technique for nonlinear scalar conservation equations. *SIAM J. Numer. Anal.* **52** (2014) 2163–2182.
8. A. Jameson, Positive schemes and shock modelling for compressible flows. *Int. J. Numer. Meth. Fluids* **20** (1995) 743–776.
9. D. Kuzmin, Algebraic flux correction I. Scalar conservation laws. In: D. Kuzmin, R. Löhner and S. Turek (eds.) *Flux-Corrected Transport: Principles, Algorithms, and Applications*. Springer, 2nd edition: 145–192 (2012).
10. D. Kuzmin, S. Basting, and J.N. Shadid, Linearity-preserving monotone local projection stabilization schemes for continuous finite elements. *Computer Methods Appl. Mech. Engrg.* **322** (2017) 23–41.
11. C. Lohmann, Algebraic flux correction schemes preserving the eigenvalue range of symmetric tensor fields. *Ergebnisse Inst. Angew. Math.* **584**. TU Dortmund University, 2018.
12. C. Lohmann, D. Kuzmin, J.N. Shadid, and S. Mabuza, Flux-corrected transport algorithms for continuous Galerkin methods based on high order Bernstein finite elements. *J. Comput. Phys.* **344** (2017) 151–186.
13. B. Neta and R.T. Williams, Stability and phase speed for various finite element formulations of the advection equation. *Computers & Fluids* **14** (1986) 393–410.