# High-resolution FEM-TVD schemes based on a fully multidimensional flux limiter

D. Kuzmin* and S. Turek

*Institute of Applied Mathematics (LS III), University of Dortmund*
*Vogelpothsweg 87, D-44227, Dortmund, Germany*

**Abstract**

A new approach to the derivation of local extremum diminishing finite element schemes is presented. The monotonicity of the Galerkin discretization is enforced by adding discrete diffusion so as to eliminate all negative off-diagonal matrix entries. The resulting low-order operator of upwind type acts as a preconditioner within an outer defect correction loop. A generalization of TVD concepts is employed to design solution-dependent antidiffusive fluxes which are inserted into the defect vector to preclude excessive smearing of solution profiles by numerical diffusion. Standard TVD limiters can be applied edge-by-edge using a special reconstruction of local three-point stencils. As a fully multidimensional alternative to this technique, a new limiting strategy is introduced. A node-oriented flux limiter is constructed so as to control the ratio of upstream and downstream edge contributions which are associated with the positive and negative off-diagonal coefficients of the high-order transport operator, respectively. The proposed algorithm can be readily incorporated into existing flow solvers as a 'black-box' postprocessing tool for the matrix assembly routine. Its performance is illustrated by a number of numerical examples for scalar convection problems and incompressible flows in two and three dimensions.

**Key Words:** convection-dominated problems; high-resolution schemes; flux limiters; finite elements; unstructured grids

# 1 Introduction

An adequate treatment of unstable convective terms is a matter of utmost importance for the majority of CFD applications. Solutions produced by standard discretization techniques are typically corrupted by nonphysical oscillations and/or excessive numerical diffusion. Traditionally, these problems have been dealt with by means of a nonlinear shock-capturing viscosity. Modern high-resolution schemes are based on flux/slope limiters which switch between linear high- and low-order discretizations adaptively depending on the smoothness of the solution. The foundations of this methodology were laid by Boris and Book [5] who introduced the pioneering concept of flux-corrected-transport (FCT). A genuinely multidimensional generalization of the original FCT algorithm was proposed by Zalesak [47] and carried over to finite elements by Löhner *et al.* [32],[33].

*Correspondence to: kuzmin@math.uni-dortmund.de

In a series of recent publications [21],[22],[23],[24] we extended the classical FEM-FCT formulation to implicit time-stepping and developed a new approach to the design of positivity-preserving schemes on arbitrary meshes. In this paper, we employ similar tools to investigate and promote another important class of high-resolution schemes which was established by Harten [15],[16]. His total variation diminishing (TVD) methods and extensions thereof rest on a firm mathematical basis and have enjoyed an increasing popularity over the past two decades. However, both the theory and the algorithms are essentially one-dimensional and no rigorous generalization to the multidimensional case is available to date. As a matter of fact, most of the CFD codes utilizing a (quasi-) TVD discretization of convective terms resort to directional splitting on Cartesian grids.

In light of the above, TVD limiters have hardly been used in the finite element context. In scarce publications on that subject, TVD-like artificial viscosities were designed using an *ad hoc* reconstruction of local one-dimensional stencils associated with edges of the finite element mesh [1],[34],[35]. Even though the numerical results were found to be quite promising, such schemes cannot be guaranteed to be positivity-preserving and may fail to suppress the nonphysical oscillations in some cases. Moreover, a piecewise-linear approximation on a simplex mesh is a prerequisite for the derivation of the underlying edge-based data structure [38]. Therefore, this approach is not suitable for multilinear and higher order finite elements. Currently, there is a pronounced trend towards the use of discontinuous Galerkin methods [7],[8] and fluctuation splitting / residual distribution schemes [6],[9] which are applicable to unstructured meshes and share some characteristics of both finite element and finite volume techniques.

The algebraic approach pursued in this paper consists in modifying the discrete transport operators so as to render the discretization local extremum diminishing. To this end, the oscillatory high-order operator is transformed into a 'monotone' low-order one by a conservative elimination of negative off-diagonal entries [22]. To recover the high accuracy of the original discretization in regions where the solution is sufficiently smooth, excessive numerical diffusion is removed by adding a limited amount of compensating antidiffusion. The diffusive and antidiffusive terms are decomposed into a sum of skew-symmetric internodal fluxes which are associated with edges of the sparsity graph for the global matrix. In general, two neighboring nodes may engage in a bilateral mass exchange if their basis functions have overlapping supports. A conservative flux decomposition is also feasible for the discretized convective terms [23]. A complete transition to an edge-based data structure may reduce the overhead incurred by indirect addressing and offer considerable savings in terms of both CPU time and memory requirements [31],[34]. At the same time, it is not mandatory, unlike in the FEM-TVD method of Lyra *et al.* [35].

The formation of wiggles is precluded by a flux limiter which controls the slope ratio in the direction of each edge or the interplay of upstream and downstream edge contributions to each node. In either case, the space discretization is proved to be local extremum diminishing. The first method (slope limiting) represents a generalization of the standard approach based on stencil reconstruction. The second one (flux limiting) is conceptually different and utilizes some features of the multidimensional FCT formulation. It turns out to be more robust and exhibits much better convergence rates to the steady state, at least in the case of certain nonconforming finite elements which lend themselves to the numerical treatment of the incompressible Navier-Stokes equations. Moreover, no *ad hoc* recovery of solution values at fictitious 'dummy nodes' is necessary, so that the limiter is completely independent of the mesh topology. All the required information is inferred from the coefficients and the sparsity pattern of the discrete operators. In fact, the algebraic nature of the proposed algorithm makes it a generic CFD tool which is applicable to finite element, finite volume and finite difference discretizations alike.

# 2 One-dimensional TVD schemes

Let us briefly review the classical TVD methodology in order to introduce some basic concepts and identify the main difficulties which hinder a truly multidimensional generalization. A detailed presentation of this material can be found in [17],[29],[44]. It was shown by Lax [28] that any physically admissible solution to a scalar conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \tag{1}$$

where $f$ is a flux function, is characterized by a nonincreasing total variation defined as

$$TV(u) = \int_{-\infty}^{\infty} \left| \frac{\partial u}{\partial x} \right| \,\mathrm{d}x. \tag{2}$$

In fact, the exact solution is constant along the characteristic lines given by $\frac{\mathrm{d}x}{\mathrm{d}t} = a(u)$, where $a(u) = \frac{\mathrm{d}f(u)}{\mathrm{d}u}$ is the propagation speed. In light of the above, it is natural to require that the total variation of the numerical solution be nonincreasing as well, i.e.

$$TV(u^{n+1}) \leq TV(u^n), \qquad \text{where} \quad TV(u^n) = \sum_i |u^n_{i+1} - u^n_i|. \tag{3}$$

Here and below $u_i$ stand for the values of the approximate solution at the mesh nodes $x_i$ and the superscripts refer to the time level at which it is evaluated. The neighboring grid points $x_i$ and $x_{i\pm1}$ exchange the conserved quantities via numerical fluxes $f_{i\pm1/2}$ which are supposed to be consistent with the underlying continuous flux $f(u)$ [29].

Harten [15] proved that a conservative semi-discrete difference scheme

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + \frac{f_{i+1/2} - f_{i-1/2}}{\Delta x} = 0 \tag{4}$$

is *total variation diminishing* provided that it can be rewritten in the form

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = c_{i-1/2}(u_{i-1} - u_i) + c_{i+1/2}(u_{i+1} - u_i) \tag{5}$$

with (possibly nonlinear) nonnegative coefficients $c_{i-1/2} \geq 0$ and $c_{i+1/2} \geq 0$. Furthermore, an additional CFL-like condition must be reckoned with unless the fully implicit backward Euler method is employed for the time discretization [15],[22].

According to the Godunov theorem [13], linear TVD schemes are doomed to be at most first-order accurate. To circumvent this serious limitation, the numerical fluxes $f_{i\pm1/2}$ can be constructed in a nonlinear way by blending high- and low-order approximations

$$f_{i\pm1/2} = f^L_{i\pm1/2} + \Phi_{i\pm1/2}[f^H_{i\pm1/2} - f^L_{i\pm1/2}]. \tag{6}$$

Here $\Phi_{i\pm 1/2}$ is an adaptive correction factor which is referred to as the *flux limiter* and designed so as to satisfy Harten's TVD conditions (5). As a rule of thumb, it should be equal to zero in the vicinity of steep gradients and approach (or even exceed) unity in regions where the solution is sufficiently smooth. This corresponds to adding a proper amount of nonlinear *antidiffusion* to the low-order flux approximation $f^L_{i\pm 1/2}$ so as to improve the accuracy without generating spurious wiggles and violating the TVD property.

To elucidate the ins and outs of this hybrid method in a rather simple setting, let us consider the linear convection equation

$$\frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} = 0 \tag{7}$$

with a constant velocity $v > 0$. In the course of space discretization, the underlying flux function $f(u) = vu$ is typically approximated by

$$f^L_{i+1/2} = vu_i \qquad \text{upwind difference method} \tag{8}$$

or

$$f^H_{i+1/2} = v\frac{u_{i+1} + u_i}{2} \qquad \text{central difference method.} \tag{9}$$

The incoming fluxes $f^L_{i-1/2}$ and $f^H_{i-1/2}$ are defined simply by shifting the index $i$. It can be easily verified that the upwind-biased flux approximation yields a first-order accurate (linear) TVD scheme with the coefficients $c_{i-1/2} = v/\Delta x,\ c_{i+1/2} = 0$. To recover the second-order accuracy of the central difference discretization whenever possible, we insert the above approximations into (6) to obtain the nonlinear TVD flux

$$f_{i+1/2} = vu_i + \frac{v}{2}\Phi_{i+1/2}(u_{i+1} - u_i). \tag{10}$$

To determine the magnitude of admissible antidiffusive correction, the flux limiter $\Phi$ should examine the local smoothness of the solution. A suitable sensor is provided by the ratio of consecutive gradients evaluated at the grid point located upwind. In the case of a positive velocity, we have $\Phi_{i+1/2} = \Phi(r_i)$. The involved slope ratio

$$r_i = \frac{u_i - u_{i-1}}{u_{i+1} - u_i} \tag{11}$$

is obviously negative at a local extremum (see Figure 1), relatively small for smooth data and large if the numerical solution tends to change abruptly.

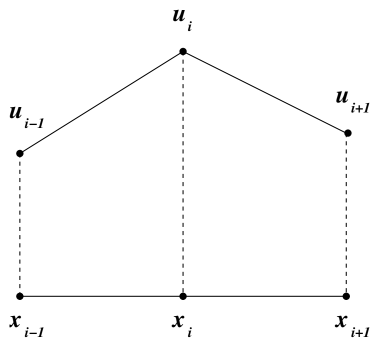


Figure 1. Three-point stencil in one dimension.

As shown by Jameson [18], it is worthwhile to interpret and implement the flux limiter $\Phi$ in terms of a special limited average operator $\mathcal{L}$ being a function of two variables such that $\Phi(r) = \mathcal{L}(1, r)$. This operator is supposed to possess the following properties:

P1. $\mathcal{L}(a, b) = \mathcal{L}(b, a)$.

P2. $\mathcal{L}(ca, cb) = c\mathcal{L}(a, b)$.

P3. $\mathcal{L}(a, a) = a$.

P4. $\mathcal{L}(a, b) = 0 \quad \text{if } ab \leq 0$.

In particular, the last property ensures that $\Phi(r) = 0$ if $r \leq 0$. Thus, the accuracy of a TVD discretization inevitably degrades to first order at local extrema. Another important implication is the symmetry of the flux limiter

$$\Phi(r) = \mathcal{L}(1, r) = r\mathcal{L}(1/r, 1) = r\Phi(1/r) \tag{12}$$

which follows from the properties (P1) and (P2). By virtue of this identity, the antidiffusive flux received by the grid point $i$ from its right neighbor is given by

$$\Phi(r_i)(u_{i+1} - u_i) = \mathcal{L}(u_{i+1} - u_i, u_i - u_{i-1}) = \Phi(1/r_i)(u_i - u_{i-1}) \tag{13}$$

and has the same effect as a diffusive flux from the left neighbor if $\Phi(1/r_i) > 0$. Note that working with $\mathcal{L}$ rather than $\Phi$ prevents division by zero in the denominator of $r_i$.

Making use of the definition (11), one can represent the semi-discrete scheme for the convection equation in the desired form (5) with the coefficients

$$c_{i-1/2} = \frac{v}{2\Delta x}\left[2 + \frac{\Phi(r_i)}{r_i} - \Phi(r_{i-1})\right], \qquad c_{i+1/2} = 0. \tag{14}$$

To meet the requirements of Harten's theorem, the expression in the brackets must be nonnegative. A variety of limiters designed to enforce this condition have been proposed in the literature. Some of the most popular two-parameter ones are as follows

$$\begin{aligned}
&\text{minmod:} && \mathcal{L}(a, b) = \mathcal{S}(a, b) \cdot \min\{|a|, |b|\},\\
&\text{Van Leer:} && \mathcal{L}(a, b) = \mathcal{S}(a, b) \cdot \frac{2|a||b|}{|a|+|b|},\\
&\text{MC:} && \mathcal{L}(a, b) = \mathcal{S}(a, b) \cdot \min\left\{\frac{|a+b|}{2}, 2|a|, 2|b|\right\},\\
&\text{superbee:} && \mathcal{L}(a, b) = \mathcal{S}(a, b) \cdot \max\{\min\{2|a|, |b|\}, \min\{|a|, 2|b|\}\}.
\end{aligned}$$

Here $\mathcal{S}(a, b)$ stands for the synchronized sign function

$$\mathcal{S}(a, b) = \frac{\text{sign}(a) + \text{sign}(b)}{2} = \begin{cases} 1 & \text{if } a > 0 \ \wedge \ b > 0, \\ -1 & \text{if } a < 0 \ \wedge \ b < 0, \\ 0 & \text{otherwise.} \end{cases}$$

The associated one-parameter limiters $\Phi$ yield correction factors lying in the range $[0, 2]$. Note that the integer values $0, 1, 2$ result in the standard upwind, central and downwind approximation, respectively.

Nowadays, TVD schemes of this kind are widely used in CFD software in conjunction with finite difference discretizations on structured grids and a number of promising high-resolution finite volume schemes is available for unstructured ones (see e.g. [4]). At the same time, a finite element practitioner considering the implementation of a TVD-like method in a general purpose code is faced by (at least) the following challenges:

- how to cast a finite element scheme in conservation form?
- how to generalize Harten's TVD criteria to multidimensions?
- how to perform upwinding in the finite element context?
- how to deal with a variable mesh size and/or velocity field?
- how to discretize nonlinear source and/or sink terms?
- how to define the parameter $r_i$ on unstructured grids?
- how to treat nonlinear problems in an implicit fashion?

These problems proved to be a rather hard nut to crack, so that the development of FEM-TVD methods has eventually ended up in a deadlock. In what follows, we will give answers to the questions listed above and propose a new methodology for the design of TVD-type discretizations on unstructured meshes. Some peculiarities of a finite element implementation are discussed in the next two sections, although they are not essential for the *algebraic flux correction* paradigm to be presented afterwards.

# 3 Finite element discretization

As a model problem, we consider the multidimensional counterpart of equation (7) which represents a mass conservation law for a scalar quantity $u$

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) = 0 \qquad \text{in } \Omega, \tag{15}$$

where the velocity field $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ is assumed to be known analytically or computed numerically from a momentum equation solved in a parallel way. The initial data are given by $u(\mathbf{x}, 0) = u_0(\mathbf{x})$ and boundary conditions are to be prescribed only at the inlet $\Gamma_{\text{in}} = \{\mathbf{x} \in \Gamma : \mathbf{v} \cdot \mathbf{n} < 0\}$, where $\mathbf{n}$ denotes the unit outward normal to the boundary $\Gamma$.

The weak form of the continuity equation is derived by integrating the weighted residual over the domain $\Omega$ and setting the result equal to zero

$$\int_\Omega w \left[\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u)\right] \mathrm{d}\mathbf{x} = 0, \qquad \forall w. \tag{16}$$

If the boundary conditions are specified in terms of fluxes rather than actual values of $u$ at the inlet, it is worthwhile to integrate the convective term by parts and substitute the incoming fluxes into the resulting surface integral.

A common practice in finite element methods for conservation laws is to interpolate the convective fluxes in the same way as the numerical solution

$$u_h = \sum_j u_j \varphi_j, \qquad (\mathbf{v}u)_h = \sum_j (\mathbf{v}_j u_j)\,\varphi_j, \tag{17}$$

where $\varphi_i$ denote the basis functions spanning the finite-dimensional subspace. This kind of approximation was called the *group finite element formulation* by Fletcher [12] who found it to provide a very efficient treatment of nonlinear convective terms and even lead to a small gain of accuracy for the 2D Burgers equation discretized on a uniform grid.

The substitution of (17) into (16) yields the following semi-discrete problem

$$\sum_j \left[\int_\Omega \varphi_i \varphi_j \,\mathrm{d}\mathbf{x}\right] \frac{\mathrm{d}u_j}{\mathrm{d}t} + \sum_j \left[\int_\Omega \varphi_i \mathbf{v}_j \cdot \nabla \varphi_j \,\mathrm{d}\mathbf{x}\right] u_j = 0. \tag{18}$$

This is a system of ordinary differential equations for the nodal values of the approximate solution which can be written compactly in matrix form

$$M_C \frac{\mathrm{d}u}{\mathrm{d}t} = Ku, \tag{19}$$

where $M_C = \{m_{ij}\}$ denotes the consistent mass matrix and $K = \{k_{ij}\}$ stands for the discrete transport operator. The matrix entries are given by

$$m_{ij} = \int_\Omega \varphi_i \varphi_j \,\mathrm{d}\mathbf{x}, \qquad k_{ij} = -\mathbf{v}_j \cdot \mathbf{c}_{ij}, \qquad \mathbf{c}_{ij} = \int_\Omega \varphi_i \nabla \varphi_j \,\mathrm{d}\mathbf{x}. \tag{20}$$

For fixed meshes, the coefficients $m_{ij}$ and $\mathbf{c}_{ij}$ remain unchanged throughout the simulation and, consequently, need to be evaluated just once during the initialization step. This enables us to update the matrix $K$ in a very efficient way by computing its entries $k_{ij}$ from formula (20) without resorting to costly numerical integration. The auxiliary coefficients $\mathbf{c}_{ij}$ correspond to the discretized space derivatives and have zero row sums, i. e. $\sum_j \mathbf{c}_{ij} = 0$ as long as the sum of basis functions $\varphi_j$ is equal to one at every point.

# 4 Conservative flux decomposition

It is common knowledge that the Galerkin FEM is globally conservative [14]. Indeed, summing equations (18) over $i$ and taking into account that the basis functions sum to unity, one recovers the integral form of the conservation law. Therefore, the total mass of $u$ in $\Omega$ may only change due to boundary fluxes. At the same time, the finite element discretization of convective terms does not admit a natural decomposition into a sum of numerical fluxes from one node into another. Since most high-resolution schemes operate with such fluxes, their extension to finite elements proved to be a difficult task. Peraire *et al.* [38] demonstrated that a conservative flux decomposition is feasible for $P_1$ finite elements (piecewise-linear approximation on a simplex mesh). A similar technique was proposed by Barth [3],[4] who investigated the relationship between finite element and finite volume discretizations. The transition to an edge-based data structure reportedly offers a number of significant advantages as compared to the conventional element-based formulation. Moreover, it paves the way for a straightforward extension of many popular high-resolution schemes (including TVD) to unstructured meshes [31],[34],[36].

In [23] we developed a flux decomposition technique which is applicable to general finite element approximations on arbitrary meshes including quadrilateral and hexahedral ones. Integration by parts in the weak formulation (16) and the fact that the coefficients $\mathbf{c}_{ij}$ have zero row sums make it possible to decompose the contribution of convective terms to interior nodes into a sum of skew-symmetric internodal fluxes [23]

$$(Ku)_i = \sum_{j\neq i} g_{ij}, \qquad \text{where} \quad g_{ij} = (\mathbf{v}_j \cdot \mathbf{c}_{ji})u_j - (\mathbf{v}_i \cdot \mathbf{c}_{ij})u_i. \tag{21}$$

A promising approach to the derivation of nonoscillatory finite element methods consists in replacing the centered flux $g_{ij}$ by another consistent numerical flux [34]. In spite of the advantages offered by such an edge-oriented solution strategy, it is often desirable to use an already existing finite element code based on conventional data structures. Due to the fact that the Galerkin method is conservative, it suffices to guarantee that all subsequent matrix manipulations to be performed at the discrete level do not violate this property. To this end, we introduce the concept of *discrete diffusion operators* which are defined as symmetric matrices having zero row and column sums [22]

$$D = \{d_{ij}\} \quad \text{such that} \quad \sum_i d_{ij} = \sum_j d_{ij} = 0, \quad d_{ij} = d_{ji}. \tag{22}$$

Some typical examples are the discrete Laplacian, the streamline-diffusion operator and the artificial diffusion introduced by row-sum mass lumping.

A discrete diffusion operator $D$ applied to the vector $u$ yields

$$(Du)_i = \sum_j d_{ij}u_j = \sum_{j\neq i} d_{ij}(u_j - u_i) \tag{23}$$

due to the zero row sum property. Hence, the contribution of diffusive terms to node $i$ can be decomposed into a sum of numerical fluxes:

$$(Du)_i = \sum_{j\neq i} f_{ij}, \qquad \text{where} \quad f_{ij} = d_{ij}(u_j - u_i). \tag{24}$$

The so-defined flux $f_{ij}$ from node $j$ into node $i$ is proportional to the difference between the nodal values, so it leads to a steepening or flattening of solution profiles depending on the sign of the diffusion coefficient $d_{ij}$. Furthermore, the symmetry of the matrix $D$ implies that $f_{ji} = -f_{ij}$ so that there is no net loss or gain of mass.

The skew-symmetric diffusive fluxes can be associated with edges of the graph representing the sparsity pattern of the global stiffness matrix. For linear finite elements their number equals the number of actual mesh edges, whereas multilinear and high-order FEM approximations allow for interactions of all nodes sharing the same element. As we are about to see, artificial diffusion operators satisfying the above conditions constitute a very useful tool for the design of multidimensional high-resolution schemes.

# 5 Design criteria

In order to prevent the formation of spurious undershoots and overshoots in the vicinity of steep gradients, we need a suitable generalization of Harten's theorem to multidimensions. Assume that the semi-discretized transport equation can be cast in the generic form

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \sum_j l_{ij} u_j, \qquad \text{where} \quad l_{ii} = -\sum_{j \neq i} l_{ij}. \tag{25}$$

In particular, this is feasible for our nodal ODE system (19) if $M_C$ is replaced by its diagonal counterpart $M_L$ resulting from row-sum mass lumping

$$l_{ij} = \frac{k_{ij}}{m_i}, \qquad \text{where} \qquad m_i = \sum_j m_{ij}, \quad M_L = \mathrm{diag}\{m_i\}$$

and the velocity field $\mathbf{v}$ is discretely divergence-free in the sense that the lumped-mass $L_2$-projection for the recovery of continuous nodal gradients yields the approximation $(\nabla \cdot \mathbf{v})_i = \frac{1}{m_i} \sum_j \mathbf{v}_j \cdot \mathbf{c}_{ij} = -\sum_j l_{ij} = 0$.

As long as the coefficients of the numerical scheme have zero row sums, the right-hand side of (25) can be represented in terms of the off-diagonal ones

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \sum_{j \neq i} l_{ij}(u_j - u_i). \tag{26}$$

It was shown by Jameson [18], [19], [20] that negative coefficients in the above expression are the 'villains' responsible for the nonphysical oscillations produced by standard high-order methods. Indeed, if $l_{ij} \geq 0,\ \forall j \neq i$ then the spatial discretization proves stable in the $L_\infty$-norm due to the fact that

- maxima do not increase: $u_i = \max_j u_j \quad \Rightarrow \quad u_j - u_i \leq 0 \quad \Rightarrow \quad \frac{\mathrm{d}u_i}{\mathrm{d}t} \leq 0,$
- minima do not decrease: $u_i = \min_j u_j \quad \Rightarrow \quad u_j - u_i \geq 0 \quad \Rightarrow \quad \frac{\mathrm{d}u_i}{\mathrm{d}t} \geq 0.$

As a rule, the coefficient matrices are sparse, so that $l_{ij} = 0$ unless $i$ and $j$ are adjacent nodes. Arguing as above, one can show that a *local* maximum cannot increase, and a *local* minimum cannot decrease. Therefore, semi-discrete schemes of this type are *local extremum diminishing*. For three-point finite difference methods, the LED constraint reduces to Harten's TVD conditions. If homogeneous Dirichlet boundary conditions are prescribed at both endpoints, the total variation of the solution can be expressed as [18]

$$TV(u_h) = 2\left(\sum u_{\max} - \sum u_{\min}\right) \tag{27}$$

and is obviously nonincreasing as long as the local maxima and minima do not grow. Therefore, one-dimensional LED schemes are necessarily total variation diminishing. At the same time, Jameson's criterion is applicable in multidimensions and the positivity of matrix coefficients is easy to verify for arbitrary discretizations on unstructured meshes.

After the time discretization, an additional condition may need to be imposed in order to make sure that the solution values remain nonnegative if this should be the case for physical reasons. In general, a fully discrete scheme is *positivity-preserving* if it can be represented in the form

$$Au^{n+1} = Bu^n, \tag{28}$$

where $B = \{b_{ij}\}$ has no negative entries and $A = \{a_{ij}\}$ is a so-called *M-matrix* defined as a nonsingular discrete operator such that $a_{ij} \leq 0$ for $j \neq i$ and all the coefficients of its inverse are nonnegative. These properties imply that the positivity of the old solution $u^n$ carries over to $u^{n+1} = A^{-1}Bu^n$.

As a useful byproduct, our algebraic positivity criterion yields a readily computable upper bound for admissible values of the time step $\Delta t = t^{n+1} - t^n$. In particular, the fully discretized local extremum diminishing scheme

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \theta \sum_{j \neq i} l_{ij}^{n+1}(u_j^{n+1} - u_i^{n+1}) + (1-\theta) \sum_{j \neq i} l_{ij}^n (u_j^n - u_i^n) \tag{29}$$

is unconditionally positivity-preserving for $\theta = 1$ (backward Euler method) and subject to the following CFL-like condition otherwise [22],[23]

$$1 + \Delta t(1-\theta) \min_i l_{ii}^n \geq 0 \qquad \text{for} \quad 0 \leq \theta < 1. \tag{30}$$

Note that this estimate is based solely on the magnitude of the diagonal coefficients $l_{ii}^n$, which makes it a handy tool for adaptive time step control.

# 6 Required modifications

The basic idea for the construction of a high-resolution scheme at the algebraic level is rather simple. It can be traced back to the concepts of flux-corrected transport [5],[47]. Roughly speaking, the governing equation is discretized in space by an arbitrary linear high-order method (e.g. central differences or Galerkin FEM) and the resulting matrices are modified *a posteriori* so as to satisfy the imposed constraints. The flow chart of required algebraic manipulations is sketched in Figure 2.

First, we perform mass lumping and transform the high-order operator $K$ into its nonoscillatory low-order counterpart $L$ by adding a discrete diffusion operator $D$ designed so as to get rid of all negative off-diagonal coefficients. In the next step, excessive artificial diffusion is removed. This is accomplished by applying a limited amount of compensating antidiffusion $F$ which depends on the local solution behavior and improves the accuracy in smooth regions. Both diffusive and antidiffusive terms admit a conservative flux decomposition so that the proposed algebraic modifications do not affect the total mass. It is worth mentioning that the final operator $K^*$ does have some negative off-diagonal coefficients. Nevertheless, the resulting discretization proves local extremum diminishing if (for a given solution $u$) there exists a matrix $L^*$ such that all off-diagonal entries $l_{ij}^*$ are nonnegative and $L^*u = K^*u$. In the remainder of this paper we will dwell on the design of discrete diffusion/antidiffusion operators and introduce multidimensional flux limiters of TVD type which guarantee the existence of $L^*$ without constructing it explicitly.

1. Linear high-order scheme (e.g. Galerkin FEM)

$$M_C \frac{\mathrm{d}u}{\mathrm{d}t} = Ku, \quad \text{such that} \quad \exists\, j \neq i: \; k_{ij} < 0$$

2. Linear low-order scheme $L = K + D$

$$M_L \frac{\mathrm{d}u}{\mathrm{d}t} = Lu, \quad \text{such that} \quad l_{ij} \geq 0, \; \forall j \neq i$$

3. Nonlinear high-resolution scheme $K^* = L + F$

$$M_L \frac{\mathrm{d}u}{\mathrm{d}t} = K^* u, \quad \text{such that} \quad \exists\, j \neq i: \; k^*_{ij} < 0$$

Equivalent LED representation $L^* u = K^* u$

$$M_L \frac{\mathrm{d}u}{\mathrm{d}t} = L^* u, \quad \text{such that} \quad l^*_{ij} \geq 0, \; \forall j \neq i$$

Figure 2. Roadmap of matrix manipulations.

# 7 Discrete upwinding

For finite difference and finite volume discretizations, the first-order accurate upwind method yields an operator $L$ which corresponds to the least diffusive linear LED scheme. Unfortunately, it has been largely unclear how to construct such an optimal low-order discretization in the finite element framework. Streamline-diffusion methods like SUPG are stable but not monotonicity-preserving, whereas other upwind-biased finite element schemes resort to a finite volume approximation of convective terms [1],[2],[46]. At the same time, the LED constraint can be enforced by elimination of negative off-diagonal coefficients from the discrete transport operator. Interestingly enough, this algebraic approach to the design of 'monotone' low-order methods reduces to standard upwinding for the one-dimensional convection equation [21],[22].

As a starting point, we consider a linear high-order discretization, e. g. our semi-discrete problem (19) for the Galerkin method. After mass lumping, each nodal value $u_i$ satisfies an ordinary differential equation of the form

$$m_i \frac{\mathrm{d}u_i}{\mathrm{d}t} = \sum_{j \neq i} k_{ij}(u_j - u_i) + \delta_i u_i, \qquad \text{where} \quad \delta_i = \sum_j k_{ij}. \tag{31}$$

The first term in the right-hand side is associated with the 'incompressible' part of the discrete transport operator $K$ since $\delta_i u_i$ is an approximation of $-u\nabla \cdot \mathbf{v}$ (see above) which vanishes for divergence-free velocity fields and is responsible for a physical growth of local extrema otherwise. For the concomitant low-order scheme to be local extremum diminishing, all off-diagonal coefficients of the linear operator $L = K + D$ must be nonnegative.

Hence, the optimal diffusion coefficients are given by [22],[23]

$$d_{ij} = d_{ji} = \max\{0, -k_{ij}, -k_{ji}\}, \qquad d_{ii} = -\sum_{j\neq i} d_{ij}. \tag{32}$$

By construction, $D = \{d_{ij}\}$ is a discrete diffusion operator. It follows that the difference between the resulting scheme and the original one can be represented as a sum of skew-symmetric diffusive fluxes $f_{ij}^d = d_{ij}(u_j - u_i)$ between adjacent nodes whose basis functions have overlapping supports. Recall that this is sufficient to guarantee mass conservation at the algebraic level. The above manipulations lead to the desired semi-discrete scheme of low order (see the roadmap in Figure 2) which reads

$$M_L \frac{\mathrm{d}u}{\mathrm{d}t} = Lu, \quad \text{such that} \quad l_{ij} \geq 0, \ \forall j \neq i. \tag{33}$$

In practice, the elimination of negative off-diagonal entries is performed edge-by-edge without assembling the global matrix $D$. After the initialization $L := K$, we examine each pair of nonzero off-diagonal coefficients $l_{ij}$ and $l_{ji}$. If the smaller one is negative, it is set equal to zero and three other entries are modified so as to restore row/column sums:

$$\begin{aligned} l_{ii} &:= l_{ii} - d_{ij}, \qquad & l_{ij} &:= l_{ij} + d_{ij}, \\ l_{ji} &:= l_{ji} + d_{ij}, \qquad & l_{jj} &:= l_{jj} - d_{ij}. \end{aligned} \tag{34}$$

Let us orient the edges of the sparsity graph so that $l_{ji} \geq l_{ij} = \max\{0, k_{ij}\}$ for the edge $\overrightarrow{ij}$. This orientation convention implies that node $i$ is located 'upwind' and corresponds to the row number of the eliminated negative entry (if any). Furthermore, the nodes can be renumbered so as to transform $L$ into an upper or lower triangular matrix and design very efficient solvers/smoothers/preconditioners for the resulting linear system.

The 'postprocessing' technique described in this section will be referred to as *discrete upwinding*. Note that the LED constraint is imposed only on the incompressible part of the transport operator. The 'reactive' term $\delta_i u_i$ is not affected by artificial diffusion since $\sum_j l_{ij} = \sum_j (k_{ij} + d_{ij}) = \sum_j k_{ij}$ due to the zero row sum property of $D$. If the governing equation contains sources and sinks, they may need to be linearized as proposed by Patankar [37] and explained in [23],[24]. Furthermore, physical diffusion can be built into the matrix either before or after discrete upwinding. In the former case, it is automatically detected and the amount of artificial diffusion is reduced accordingly. In our experience, TVD limiters should be applied to the convective operator alone. Therefore, it is advisable to incorporate the contribution of physical diffusion into $L$ rather than $K$.

# 8 Generalized TVD formulation

The modified Galerkin discretization derived by means of row-sum mass lumping and discrete upwinding is optimal in the sense that it incorporates just as much artificial diffusion as is necessary to preclude the birth and growth of spurious wiggles regardless of the local smoothness. Nevertheless, this LED scheme is linear and therefore at most first-order accurate [13]. To circumvent the restrictive Godunov theorem, a nonoscillatory high-resolution scheme must be nonlinear even for a linear partial differential equation. It is worth mentioning that the majority of CFD applications are governed by nonlinear conservation laws to begin with, so that the computational overhead needed to remove excessive artificial diffusion is less significant than in the linear case.

In order to offset the loss of accuracy in smooth regions, we modify the discrete transport operator again by adding a limited amount of nonlinear antidiffusion:

$$K^*(u) = L + F(u) = K + D + F(u), \tag{35}$$

where both $D$ and $F(u)$ possess the properties of discrete diffusion operators.

In a practical implementation, the contribution of nonlinear antidiffusive terms to the right-hand side of the final semi-discrete scheme (cf. the flow chart in Figure 2)

$$M_L \frac{\mathrm{d}u}{\mathrm{d}t} = K^* u \qquad \longleftarrow \qquad (Fu)_i = \sum_{j \neq i} f_{ij}^a \tag{36}$$

is assembled edge-by-edge from the skew-symmetric internodal fluxes

$$f_{ij}^a := \min\{\Phi(r_i) d_{ij}, l_{ji}\}(u_i - u_j), \qquad f_{ji}^a := -f_{ij}^a. \tag{37}$$

The antidiffusive flux $f_{ij}^a$ from node $j$ into its **upwind** (in the sense of our orientation convention $l_{ji} \geq l_{ij}$) neighbor $i$ depends on the diffusion coefficient $d_{ij}$ for discrete upwinding and on the entry $l_{ji} = \max\{k_{ji}, k_{ji} - k_{ij}\}$ of the low-order transport operator. Furthermore, $\Phi$ is a standard one-parameter limiter applied to a suitable smoothness indicator $r_i$ (to be specified below). By definition, the downwind node $j$ receives the flux $f_{ji}^a$ of the same magnitude but with the opposite sign so that mass conservation is guaranteed.

Let us derive a *sufficient* condition for scheme (36) to be local extremum diminishing. If $\Phi(r_i) = 0$ or $d_{ij} = 0$, the antidiffusive flux $f_{ij}^a$ vanishes and does not pose any hazard. Therefore, we restrict ourselves to the nontrivial case $f_{ij}^a \neq 0$ which implies that both $\Phi(r_i)$ and $d_{ij}$ are strictly positive. Our objective is to prove the existence of a LED operator $L^*$ which is equivalent to $K^*$ for the given solution $u$ (see the last box in Figure 2). Clearly, the sensor $r_i$ cannot be chosen arbitrarily. The symmetry property (12) of the limiter $\Phi$ makes it possible to represent the antidiffusive flux in the form

$$f_{ij}^a = \Phi(r_i) a_{ij}(u_i - u_j) = \Phi(1/r_i) a_{ij} \Delta u_{ij}, \tag{38}$$

where the *antidiffusion coefficient* $a_{ij}$ and the *upwind difference* $\Delta u_{ij}$ are defined as follows

$$a_{ij} := \min\{d_{ij}, l_{ji}/\Phi(r_i)\}, \qquad \Delta u_{ij} := r_i(u_i - u_j). \tag{39}$$

For the numerical solution to be nonoscillatory, the antidiffusive fluxes must behave as diffusive ones, cf. equation (13). The assumption $d_{ij} > 0$ implies that $k_{ij} < 0$ and $l_{ij} = 0$ for the edge $\overrightarrow{ij}$ which links an upwind node $i$ and a downwind node $j$. Therefore, its contributions to the modified convective term $K^*u$ in (36) can be written as

$$k_{ij}^*(u_j - u_i) = f_{ij}^a, \qquad k_{ji}^*(u_i - u_j) = l_{ji}(u_i - u_j) - f_{ij}^a. \tag{40}$$

The increment to node $j$ is obviously of diffusive nature and satisfies the LED criterion, since the coefficient $k_{ji}^* = l_{ji} - \Phi(r_i)a_{ij}$ is nonnegative by construction (see the definition of $a_{ij}$). Furthermore, it follows from relation (38) that the negative off-diagonal entry $k_{ij}^* = -\Phi(r_i)a_{ij}$ of the nonlinear transport operator $K^*$ is acceptable provided that $\Delta u_{ij}$ admits the following representation

$$\Delta u_{ij} = \sum_{k \neq i} \alpha_{ik}(u_k - u_i), \qquad \text{where} \quad \alpha_{ik} \geq 0, \quad \forall k \neq i. \tag{41}$$

In other words, the limited antidiffusive flux $f_{ij}^a$ from node $j$ into node $i$ should be interpreted as a sum of diffusive fluxes contributed by other neighbors. It remains to devise a multidimensional smoothness indicator $r_i$ and check if the corresponding upwind difference $\Delta u_{ij}$ satisfies the above condition.

# 9 Slope-limiter FEM-TVD algorithm

For classical finite difference TVD schemes, the quantity $r_i$ is given by the slope ratio (11) at the upwind node so that $\Delta u_{ij} = u_k - u_i$, where $k = i-1$ is the second neighbor of node $i$. However, this natural definition of $r_i$ is no longer possible in multidimensions, whereby each node interacts with more than two neighbors. A geometric approach commonly employed in the literature is to reconstruct a local one-dimensional stencil by insertion of equidistant *dummy nodes* on the continuation of each mesh edge [1],[18],[34],[35]. The difference $\Delta u_{ij}$ is defined as before using the interpolated or extrapolated solution value at the dummy node $k$ adjacent to the upwind node $i$. The construction of a three-point stencil for an unstructured triangular mesh is illustrated in Figure 3.

The numerical behavior of various techniques for the recovery of $u_k$ was analyzed in detail by Lyra [34]. His comparative study covers the following algorithms:

1. Interpolation/extrapolation using the adjacent triangle $T_i$ containing node $i$.

2. Interpolation using the actual triangle $T_k$ containing the dummy node $k$.

3. Extrapolation using a least squares reconstruction for the gradient at node $i$.

Numerical experiments revealed that solutions depend strongly on the employed strategy. The first option was proved to provide the LED property but failed to produce nonoscillatory results for some aerodynamic applications. The second procedure based on the actual triangle was favored due to the enhanced robustness as compared to the use of the adjacent triangle. However, the resulting discretization is no longer local extremum diminishing and neither is the gradient reconstruction method which corresponds to

$$\Delta u_{ij} = (\mathbf{x}_i - \mathbf{x}_j) \cdot \nabla_h u_i, \qquad \text{where} \qquad \nabla_h u_i = \frac{1}{m_i} \sum_{k \neq i} \mathbf{c}_{ik}(u_k - u_i). \tag{42}$$

Here $\nabla_h u_i$ stands for a continuous approximation to the solution gradient at the upwind node $i$ recovered by means of a consistent $L_2$-projection. The involved coefficients $\mathbf{c}_{ik}$ are

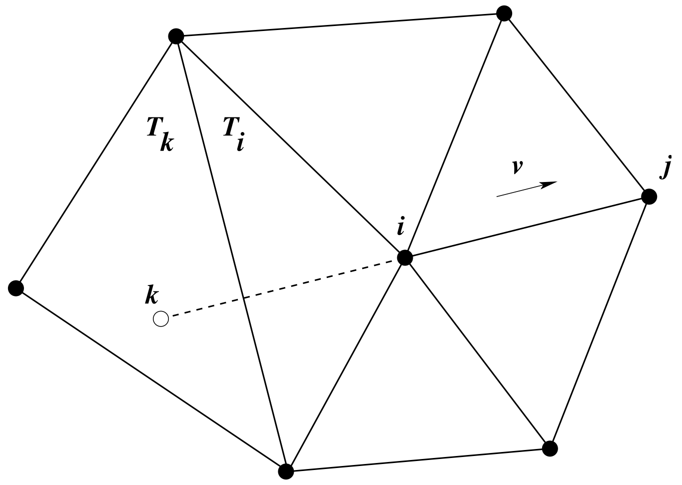


Figure 3. Three-point stencil in two dimensions.

defined in (20) for the standard Galerkin method. This approach is relatively simple to implement and more efficient than the linear interpolation techniques. However, Lyra [34] reported its performance to be quite poor and emphasized the need for the development of a more robust algorithm for the reconstruction of nodal gradients.

Let us explain why the above choice of the upwind difference may prove unsatisfactory. If any of the scalar products $\mathbf{c}_{ik} \cdot (\mathbf{x}_i - \mathbf{x}_j)$ is negative, then the formula for $\Delta u_{ij}$ is not of the form (41), so the numerical scheme may fail to satisfy the LED criterion. To rectify this, one can employ a monotone projection operator constructed by resorting to discrete upwinding. Note that $\mathbf{c}_{ki} = -\mathbf{c}_{ik}$ for internal nodes, so that the elimination of negative off-diagonal coefficients leads to the following LED-type reconstruction procedure [25]

$$\Delta u_{ij} = \frac{2}{m_i} \sum_{k \neq i} \max\{0, \mathbf{c}_{ik} \cdot (\mathbf{x}_i - \mathbf{x}_j)\}(u_k - u_i). \tag{43}$$

For uniform meshes in one dimension, this kind of extrapolation corresponds to using the upwind gradient and yields $\Delta u_{ij} = u_k - u_i$, where $k$ is the upwind neighbor of $i$.

By virtue of (39), the upwind difference $\Delta u_{ij}$ can be converted into the smoothness sensor $r_i$ which provides an estimate of the gradient jump along the edge $\overrightarrow{ij}$. Hence, this geometric approach to the design of nonlinear LED schemes will be referred to as the *slope-limiter* FEM-TVD algorithm. It can be equipped with any standard limiter $\Phi$. At the same time, the numerical results are rather sensitive to the alignment of the three-point stencil and to the algorithm employed to recover the solution values at the dummy node. Moreover, such methods are computationally expensive and may experience severe convergence problems for steady-state applications. This shortcoming was also noticed by Lyra [34] who explained it by the lack of background dissipation and indicated that the convergence rates can be improved to some extent by 'freezing' the antidiffusive terms as the solution approaches the steady state. The main reason for the insufficient robustness seems to be the unidirectional nature of stencil reconstruction and the independent limiting of antidiffusive fluxes associated with the same upwind node.

## 10 Flux-limiter FEM-TVD algorithm

Let us abandon the stencil reconstruction technique and design the smoothness indicator $r_i$ in a different way. The one-dimensional convection equation (7) discretized in space by the lumped-mass Galerkin FEM or by the central difference method can be written in the form (5) where the two coefficients are given by

$$c_{i-1/2} = \frac{v}{2\Delta x} > 0, \qquad c_{i+1/2} = -\frac{v}{2\Delta x} < 0. \tag{44}$$

Remarkably, the ratio of upwind and downwind contributions to node $i$

$$r_i = \frac{c_{i-1/2}(u_{i-1} - u_i)}{c_{i+1/2}(u_{i+1} - u_i)} \tag{45}$$

reduces to the slope ratio $r_i$ defined in (11) as long as the velocity $v$ is constant. Moreover, this interpretation leads to a conceptually different limiting strategy which guarantees the TVD property for variable velocity fields and carries over to multidimensions. As we are about to see, the new algorithm is akin to that proposed by Zalesak [47] in the framework of flux-corrected transport methods, so we adopt his notation to reflect this relationship.

In the multidimensional case, the incompressible part of the original convective term $Ku$ can be decomposed into a sum of edge contributions with negative coefficients and a sum of those with positive coefficients

$$P_i = \sum_{j\neq i} \min\{0, k_{ij}\}(u_j - u_i), \quad Q_i = \sum_{j\neq i} \max\{0, k_{ij}\}(u_j - u_i) \tag{46}$$

which are due to mass transfer from the downstream and upstream directions, respectively. The sum $P_i$ is composed from the *raw antidiffusive fluxes* which offset the error incurred by elimination of negative matrix entries in the course of discrete upwinding. They are responsible for the formation of spurious wiggles and must be securely limited. At the same time, the constituents of the sum $Q_i$ are harmless since they resemble diffusive fluxes and do satisfy the LED criterion. Thus, it is natural to require that the net antidiffusive flux into node $i$ be a limited average of the original increments $P_i$ and $Q_i$.

Due to the property (P4) of TVD limiters, it is worthwhile to distinguish between the positive and negative edge contributions to both sums

$$P_i = P_i^+ + P_i^-, \qquad P_i^\pm = \sum_{j\neq i} \min\{0, k_{ij}\} \begin{matrix}\min\\ \max\end{matrix} \{0, u_j - u_i\}, \tag{47}$$

$$Q_i = Q_i^+ + Q_i^-, \qquad Q_i^\pm = \sum_{j\neq i} \max\{0, k_{ij}\} \begin{matrix}\max\\ \min\end{matrix} \{0, u_j - u_i\} \tag{48}$$

and limit the positive and negative antidiffusive fluxes separately. To this end, we pick a standard limiter $\Phi$ and compute the *nodal correction factors*

$$R_i^\pm = \Phi(Q_i^\pm / P_i^\pm) \tag{49}$$

which determine the percentage of $P_i^\pm$ that can be retained without violating the LED constraint for row $i$ of the modified transport operator $K^*$. Clearly, $R_i^\pm$ does not need to be evaluated if the raw antidiffusion $P_i^\pm$ vanishes.

For each edge $\overrightarrow{ij}$ of the sparsity graph, the antidiffusive flux $f_{ij}^a$ from its downwind node $j$ into the upwind node $i$ is constructed as follows:

$$f_{ij}^a := \begin{cases} \min\{R_i^+ d_{ij}, l_{ji}\}(u_i - u_j) & \text{if } u_i \geq u_j, \\ \min\{R_i^- d_{ij}, l_{ji}\}(u_i - u_j) & \text{if } u_i < u_j, \end{cases} \qquad f_{ji}^a := -f_{ij}^a. \tag{50}$$

Importantly, the same correction factor $R_i^\pm$ is applied to all positive/negative antidiffusive fluxes which represent the interactions of node $i$ with its neighbors located downstream in the sense of our orientation convention.

The node-oriented limiting strategy makes it possible to control the combined effect of antidiffusive fluxes *acting in concert* rather than merely the variation of the solution along each edge. Moreover, the new limiter extracts all information from the original matrix $K$ and does not need the coordinates of nodes or other geometric details. The equivalence of (37) and (50) reveals that the underlying smoothness sensor $r_i$ is implicitly defined by

$$r_i = \begin{cases} Q_i^+/P_i^+ & \text{if } u_i \geq u_j, \\ Q_i^-/P_i^- & \text{if } u_i < u_j. \end{cases} \tag{51}$$

It is easy to verify that the upwind difference $\Delta u_{ij} = r_i(u_i - u_j)$ satisfies condition (41) since all coefficients in the sum of upwind contributions $Q_i^\pm$ are nonnegative and

$$\Delta u_{ij} = \alpha_{ij} Q_i^\pm, \qquad \text{where} \quad \alpha_{ij} = \genfrac{}{}{0pt}{}{\max}{\min}\{0, u_i - u_j\}/P_i^\pm \geq 0. \tag{52}$$

Thus, our nonlinear semi-discrete scheme (36) proves local extremum diminishing if the antidiffusive fluxes are computed from (50). The new algorithm will be called the *flux-limiter* FEM-TVD method to distinguish it from the one described in the preceding section. In one dimension, both generalizations reduce to their finite difference prototype.

# 11 Iterative defect correction

The algorithm presented so far can be classified as a *method of lines* in which the spatial derivatives are already discretized and a system of coupled ordinary differential equations for the time-dependent nodal values $u_i$ is to be solved. In principle, the discretization of system (36) in time can be performed using any numerical method for initial value problems. First- or second-order accuracy is sufficient for our purposes, so we can use the standard $\theta$-scheme. Furthermore, we restrict ourselves to implicit time-stepping methods ($0 < \theta \leq 1$) because the implementation of the fully explicit one is straightforward. As a result, we end up with a nonlinear algebraic system of the form

$$M_L \frac{u^{n+1} - u^n}{\Delta t} = \theta K^*(u^{n+1})u^{n+1} + (1-\theta)K^*(u^n)u^n \tag{53}$$

which must be solved iteratively. Recall that the time step $\Delta t$ is subject to a CFL-like positivity condition unless $\theta = 1$.

Successive approximations to the end-of-step solution $u^{n+1}$ can be computed e. g. by the fixed-point defect correction scheme [46]

$$u^{(m+1)} = u^{(m)} + A^{-1} r^{(m)}, \qquad m = 0, 1, 2, \dots \tag{54}$$

where $r^{(m)}$ denotes the residual vector for the $m$-th cycle and $A$ is a 'preconditioner' which should be easy to invert. The iteration process continues until the norm of the defect or that of the relative changes becomes small enough.

In a practical implementation, the 'inversion' of $A$ is also performed by a suitable iterative method for solving the linear subproblem

$$A \Delta u^{(m)} = r^{(m)}, \qquad m = 0, 1, 2, \dots \tag{55}$$

After a certain number of inner iterations, the resulting solution increment $\Delta u^{(m)}$ is applied to the last iterate, whereby $u^n$ provides a reasonable initial guess

$$u^{(m+1)} = u^{(m)} + \Delta u^{(m)}, \qquad u^{(0)} = u^n. \tag{56}$$

Incidentally, the auxiliary problem (55) does not have to be solved very accurately at each outer iteration. A moderate improvement of the residual (1-2 digits) is sufficient to obtain a good overall accuracy. By construction, the low-order evolution operator

$$A = M_L - \theta \Delta t L, \qquad L = K + D \tag{57}$$

for the underlying linear LED scheme (33) enjoys the M-matrix property and constitutes an excellent preconditioner. Furthermore, the diagonal dominance of $A$ can be enhanced by means of an implicit underrelaxation [11]. In fact, iterative defect correction preconditioned by the monotone upwind operator is frequently used to enhance the robustness of CFD solvers. This practice is to be recommended even in the linear case, since an iterative method may fail to converge if applied directly to the ill-conditioned matrix originating from a high-order discretization of the troublesome convective terms.

The defect vector and the constant right-hand side are given by

$$r^{(m)} = b^n - [A - \theta \Delta t F(u^{(m)})] u^{(m)}, \tag{58}$$

$$b^n = M_L u^n + (1 - \theta) \Delta t [L + F(u^n)] u^n. \tag{59}$$

Note that both expressions consist of a low-order contribution augmented by limited antidiffusion of the form (36). The antidiffusive fluxes $f_{ij}^a$ are evaluated edge-by-edge at the corresponding time level and inserted into the global vectors. If they are omitted, we recover the nonoscillatory linear scheme (33) which is overly diffusive. The task of the flux limiter is to determine how much artificial diffusion can be safely removed without violating the LED criterion. It is worth mentioning that our algebraic FEM-TVD algorithm is directly applicable to steady-state problems as well as to time-dependent equations written as stationary boundary value problems in the space-time domain.

## 12 Summary of the algorithm

The proposed multidimensional generalization of TVD schemes can be implemented on arbitrary grids using either the conventional or the edge-based data structure. As a matter of fact, the algorithm is applicable to finite differences, finite elements and finite volumes alike, since the origin of the discrete transport operator $K$ is immaterial. The required modifications are limited to the matrix assembly routine which is to be called repeatedly in the outer defect correction loop (54) for the iterative treatment of nonlinearities. The sequence of 'postprocessing' steps to be performed can be summarized as follows:

*In a loop over edges:*

1. Retrieve the entries $k_{ij}$ and $k_{ji}$ of the high-order transport operator.
2. Determine the artificial diffusion coefficient $d_{ij}$ from equation (32).
3. Update the four entries of the preconditioner $A$ as required by (34).
4. Adopt the edge orientation $\overrightarrow{ij}$ such that node $i$ is located upwind.
5. Store the order of nodes as well as $d_{ij}$ and $l_{ji}$ for future reference.

*In a loop over nodes:*

6. Calculate the ratio of upstream/downstream contributions $Q_i^{\pm}$ and $P_i^{\pm}$.
7. Apply a TVD limiter $\Phi$ to obtain the nodal correction factors $R_i^{\pm}$.

*In a loop over edges:*

8. Compute the diffusive flux $f_{ij}^d = d_{ij}(u_j - u_i)$ due to discrete upwinding.
9. Check the sign of $u_i - u_j$ and evaluate the flux $f_{ij}^a$ from formula (50).
10. Insert the corrected (anti-)diffusive flux $f_{ij} = f_{ij}^d + f_{ij}^a$ into the defect.

The numerical examples that follow illustrate the performance of this algorithm. The slope-limiter version can be coded in a similar way using the upwind difference (43) to determine the slope ratio $r_i = \Delta u_{ij}/(u_i - u_j)$ and the corresponding correction factors $\Phi(r_i)$. Note that no loop over nodes is needed in this case. Indeed, the recovery of $\Delta u_{ij}$ via stencil reconstruction is performed independently for each edge. As an alarming consequence, the contributions of other edges are not taken into account, so that the total antidiffusive flux cannot be properly controlled. However, on regular grids the numerical results produced by a slope-limiter FEM-TVD method are typically quite good [25].

# 13 Numerical examples

## 13.1 Solid body rotation

Rotation of solid bodies with discontinuities and small scale features is frequently used as a challenging test problem for transport algorithms. In the first example, we consider the benchmark configuration proposed by LeVeque [30]. It is intended to examine the ability of a numerical method to reproduce both discontinuous and smooth profiles. To this end, a slotted cylinder, a cone and a smooth hump are exposed to the nonuniform velocity field $\mathbf{v} = (0.5 - y, x - 0.5)$ and undergo a counterclockwise rotation about the center of the square domain $\Omega = (0,1) \times (0,1)$. Each of these bodies lies within a circle of radius $r_0 = 0.15$ centered at a point with Cartesian coordinates $(x_0, y_0)$.

The exact solution to the linear convection equation after each full revolution matches the initial data depicted in Figure 4 (left). Let us introduce the normalized distance function $r(x,y) = \frac{1}{r_0}\sqrt{(x-x_0)^2 + (y-y_0)^2}$. It follows that $u(x,y,0) = 0$ for $r(x,y) > 1$. Elsewhere, the reference shape of the three bodies is given by

$$\text{Cylinder:} \quad (x_0, y_0) = (0.5, 0.75), \quad u(x,y,0) = \begin{cases} 1, & \text{if } |x - x_0| \geq 0.025 \vee y \geq 0.85, \\ 0, & \text{otherwise.} \end{cases}$$

$$\text{Cone:} \quad (x_0, y_0) = (0.5, 0.25), \quad u(x,y,0) = 1 - r(x,y).$$

$$\text{Hump:} \quad (x_0, y_0) = (0.25, 0.5), \quad u(x,y,0) = 0.25[1 + \cos(\pi \min\{r(x,y), 1\})].$$

The numerical solution at $t = 2\pi$ produced by the FEM-TVD scheme with the Crank-Nicolson time-stepping and the superbee flux limiter is shown in Figure 4 (right). It was computed on a uniform mesh of $128 \times 128$ bilinear elements using the time step $\Delta t = 10^{-3}$.

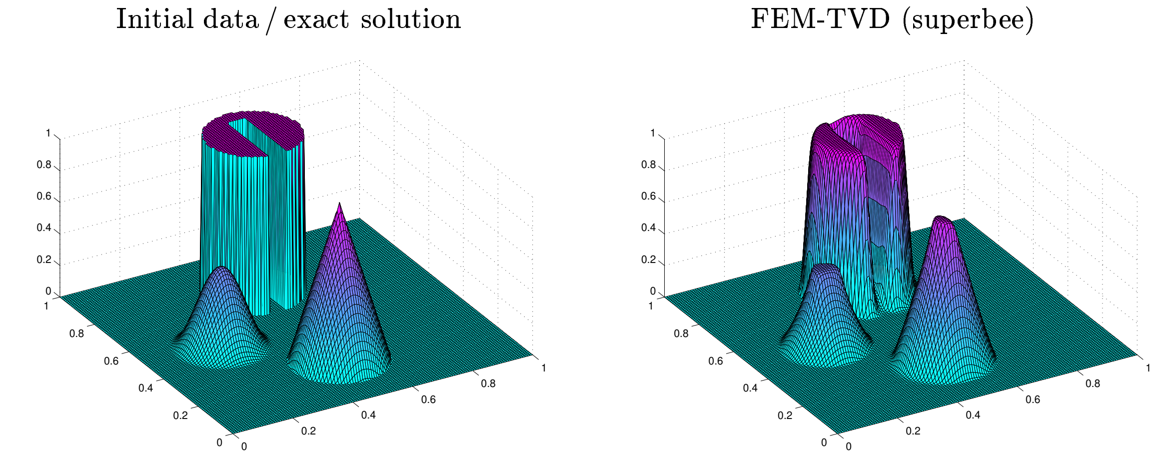


Figure 4. Solid body rotation, $128 \times 128$ bilinear elements, $t = 2\pi$.

No spurious wiggles are observed and the resolution of discontinuities is far superior to that achievable with discrete upwind or a similar low-order method. Even the narrow bridge of the slotted cylinder is largely preserved. However, the irrecoverable error induced by mass lumping leads to some loss of accuracy which manifests itself in a noticeable erosion of the ridges. Furthermore, the employed superbee limiter is known to be slightly underdiffusive (see the next example), which results in an artificial steepening of solution gradients and a pronounced peak flattening for the cone and hump.

## 13.2 Rotation of a Gaussian hill

The second test case proposed by Lapin [27] makes it possible to evaluate the magnitude of artificial diffusion due to the discretization in space and time. This can be accomplished by applying certain statistical tools to the convection-diffusion equation

$$\frac{\partial u}{\partial t} + \mathbf{v} \cdot \nabla u = \epsilon \Delta u \quad \text{in } \Omega = (-1,1) \times (-1,1), \tag{60}$$

where $\mathbf{v} = (-y, x)$ is the velocity field and $\epsilon = 10^{-3}$ is the physical diffusion coefficient.

The initial condition to be imposed is given by $u(x, y, 0) = \delta(x_0, y_0)$, where $\delta$ stands for the Dirac delta function. Clearly, it is impossible to initialize the solution by a singular function in a practical implementation. Instead, it is reasonable to concentrate the whole mass at a single node. The integral of a discrete function over the domain $\Omega$ can be computed as the sum of nodal values multiplied by the entries of the lumped mass matrix: $\int_\Omega u_h \, d\mathbf{x} = \int_\Omega \sum_i u_i \varphi_i \, d\mathbf{x} = \sum_i m_i u_i$. The total mass of a delta function equals unity. Hence, one should find node $i$ closest to the peak location $(x_0, y_0)$ and set $u_i^0 = 1/m_i$, $u_j^0 = 0,\ j \neq i$. Alternatively, one can start with the exact solution at a time $t_0 > 0$.

In the rotating Lagrangian reference frame, the convective term vanishes and the resulting diffusion problem can be solved analytically. It can be readily verified that the exact solution of (60) is a Gaussian hill defined by the normal distribution function

$$u(x, y, t) = \frac{1}{4\pi\epsilon t} e^{-\frac{r^2}{4\epsilon t}}, \qquad r^2 = (x - \hat{x})^2 + (y - \hat{y})^2,$$

where $\hat{x}$ and $\hat{y}$ denote the time-dependent peak coordinates

$$\hat{x}(t) = x_0 \cos t - y_0 \sin t, \qquad \hat{y}(t) = -x_0 \sin t + y_0 \cos t.$$

The actual peak coordinates for a numerical approximation may be quite different. They can be calculated as the mathematical expectation of the center of mass under the probability distribution with density $u_h$ given by the finite element solution

$$\hat{x}_h(t) = \int_\Omega x u_h(x, y, t) \, d\mathbf{x}, \qquad \hat{y}_h(t) = \int_\Omega y u_h(x, y, t) \, d\mathbf{x}.$$

The quality of approximation can be assessed by considering the standard deviation

$$\sigma_h^2(t) = \int_\Omega r_h^2 u_h(x, y, t) \, d\mathbf{x}, \qquad r_h^2 = (x - \hat{x}_h)^2 + (y - \hat{y}_h)^2,$$

which quantifies the rate of smearing caused by both physical and the numerical diffusion. Due to all sorts of discretization errors, $\sigma_h^2$ may differ considerably from the exact value $\sigma^2 = 4\epsilon t$. This discrepancy represented by the relative variance error

$$\Delta\sigma_{\text{rel}} = \frac{\sigma_h^2 - \sigma^2}{\sigma^2} = \frac{\sigma_h^2}{4\epsilon t} - 1$$

serves as an excellent indicator of numerical diffusion inherent to the discretization scheme.

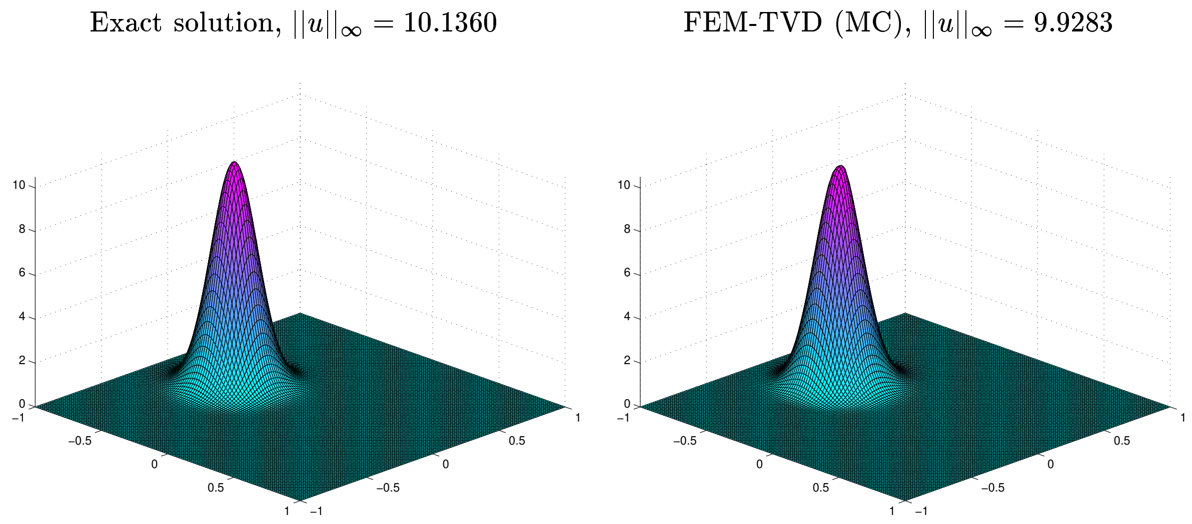


Figure 5. Rotation of a Gaussian hill, $128 \times 128$ bilinear elements, $t = 2.5\,\pi$.

Let us start with the analytical solution corresponding to $x_0 = 0$, $y_0 = 0.5$ and $t_0 = 0.5\,\pi$. Figure 5 depicts the exact and numerical solution after one full revolution of the Gaussian hill. The mesh size and time step are the same as in the previous example. The solution produced by the FEM-TVD method with the Crank-Nicolson time-stepping and the MC flux limiter proves to be very accurate, although some 'peak clipping' does occur. The global maximum drops to 9.9283 as compared to 10.1360 for the exact solution.

As the Gaussian hill moves around the origin, it is being gradually smeared by diffusion. The error estimator $\Delta\sigma_{\rm rel}$ enables us to compare the performance of standard TVD limiters and to investigate the influence of the time discretization. If the first-order accurate backward Euler method is employed, the temporal part of the relative variance error plays an important role at large time steps and decreases linearly as the time step

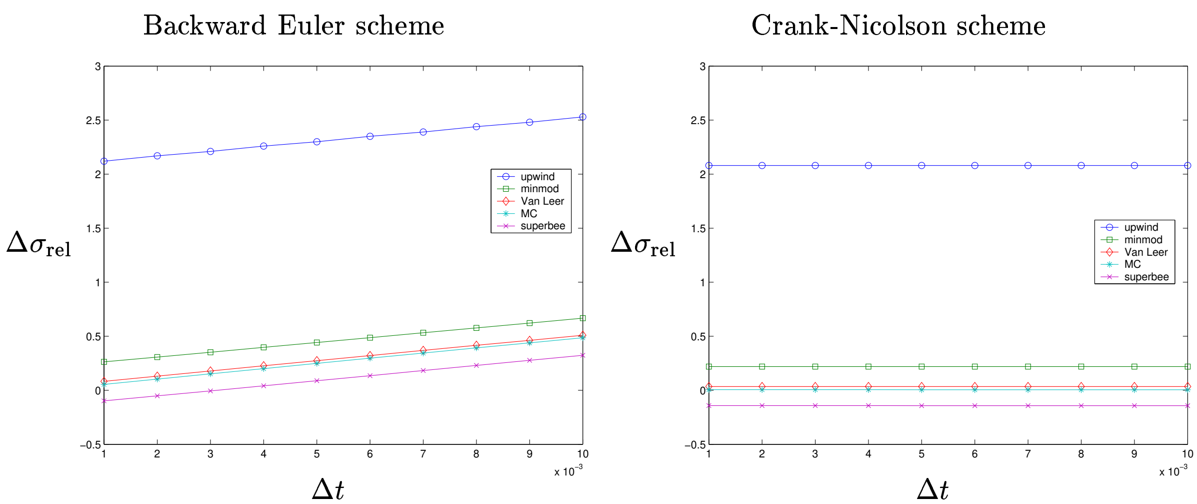


Figure 6. Gaussian hill: relative variance error vs. the time step.

is refined (see Figure 6, left). For the second-order accurate Crank-Nicolson scheme, the temporal discretization error is negligibly small. This is why the time step does not affect the values of $\Delta\sigma_{\rm rel}$ displayed in Figure 6 (right). In this case, the accuracy is determined by the space discretization and the choice of the flux limiter is decisive.

As expected, by far the most diffusive solutions are produced by the discrete upwind scheme, whereas the nonlinear FEM-TVD correction leads to a dramatic improvement. At the same time, a comparison of standard TVD limiters with one another reveals that the relative variance error may differ appreciably from case to case. The most diffusive limiter is minmod followed by Van Leer. The MC limiter used to obtain the numerical solution shown in Figure 5 outperforms both of them. Moreover, it is more suitable for the treatment of smooth profiles than Roe's superbee limiter. The latter turns out to be underdiffusive, so that $\Delta\sigma_{\rm rel}$ is negative if the spatial discretization error dominates. It is important to be aware of this fact when using the superbee limiter in CFD simulations.

## 13.3 Steady-state convection-diffusion

The FEM-TVD algorithm for the design of discrete transport operators can be applied to stationary problems directly or in conjunction with a pseudo-time-stepping technique. In the latter case, the steady-state solution is obtained by marching into the stationary limit of the associated time-dependent problem. The temporal accuracy is immaterial in this case, since the time step is merely an artificial parameter which determines the convergence rates. Hence, it is desirable to choose time steps as large as possible, so as to reduce the computational cost. The restrictive CFL condition prevents explicit schemes from operating with large time steps and makes them too inefficient for our purposes. This drawback can be rectified to some extent by resorting to local time-stepping but it is obvious that steady-state problems call for an implicit treatment.

In light of the above, the fully implicit backward Euler method, which was found to be quite diffusive for transient problems, constitutes an excellent iterative solver for steady or creeping flows. Let us investigate the numerical behavior of the backward Euler TVD scheme for the singularly perturbed convection-diffusion equation

$$\mathbf{v}\cdot\nabla u-\epsilon\Delta u=0 \qquad \text{in } \Omega=(0,1)\times(0,1),$$

where $\mathbf{v}=(\cos 10^{\circ},\sin 10^{\circ})$ and $\epsilon=10^{-3}$. The concomitant boundary conditions read

$$\frac{\partial u}{\partial y}(x,1)=0, \qquad u(x,0)=u(1,y)=0, \qquad u(0,y)=\begin{cases}1, & y\geq 0.5,\\ 0, & y<0.5.\end{cases}$$

The solution to this elliptic problem is characterized by the presence of a sharp front next to the line $x=1$. The boundary layer develops because the solution of the reduced problem ($\epsilon=0$) does not satisfy the homogeneous Dirichlet boundary condition.

A reasonable initial approximation for the pseudo-time-stepping loop is given by

$$u(x,y,0)=\begin{cases}1-x, & y\geq 0.5,\\ 0, & y<0.5.\end{cases}$$

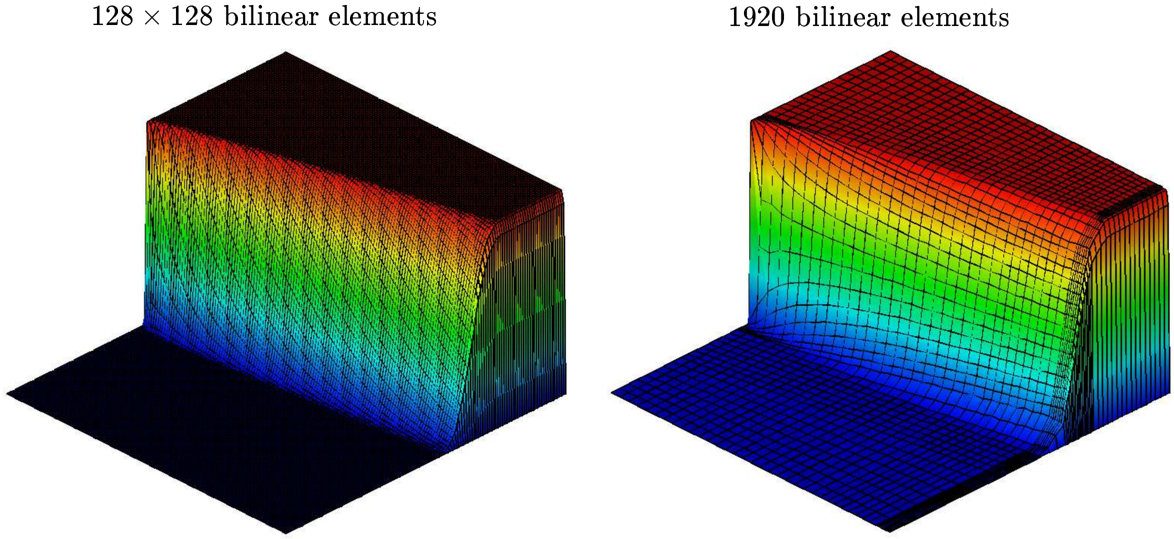


Figure 7. Convection-diffusion, $\epsilon = 10^{-3}$. FEM-TVD (superbee).

It is worthwhile to start with the discrete upwind scheme and use the converged low-order solution as initial data for the time-dependent FEM-TVD algorithm. This 'educated guess' should be close enough to the steady-state limit. Hence, the computational overhead due to the assembly and limiting of antidiffusive fluxes will be insignificant.

The numerical solutions depicted in Figure 7 demonstrate that the FEM-TVD method combined with backward Euler time discretization is capable of producing nonoscillatory solutions with a sharp resolution of steep fronts and boundary layers. The left diagram was computed as before on the Cartesian grid of $128 \times 128$ bilinear elements, while the computational mesh for the right one consists of just 1920 elements and is refined in regions where the solution gradients are large. Due to the adaptive mesh refinement and a special alignment of the grid lines, the accuracy is comparable with that achieved on the uniform mesh at a much higher computational cost. This example indicates that our multidimensional FEM-TVD algorithm can be successfully applied on adaptive meshes and benefit from the unconditional stability of implicit time-stepping.

## 13.4 Standing vortex

The generality of the fully discrete approach has enabled us to integrate the FEM-TVD algorithm into the incompressible flow solver FEATFLOW based on the nonconforming Rannacher-Turek finite elements (a discontinuous rotated bilinear approximation of the velocity components on quadrilateral meshes) [39],[45]. The multidimensional flux limiter was built into the matrix assembly routine for the convective terms and no other modifications of the code were necessary. The incompressibility constraint for the Navier-Stokes equations was imposed in the framework of the Multilevel Pressure Schur Complement formulation which unites coupled solution techniques and discrete projection methods which decouple the velocity and pressure using suitable operator-splitting tools [46].

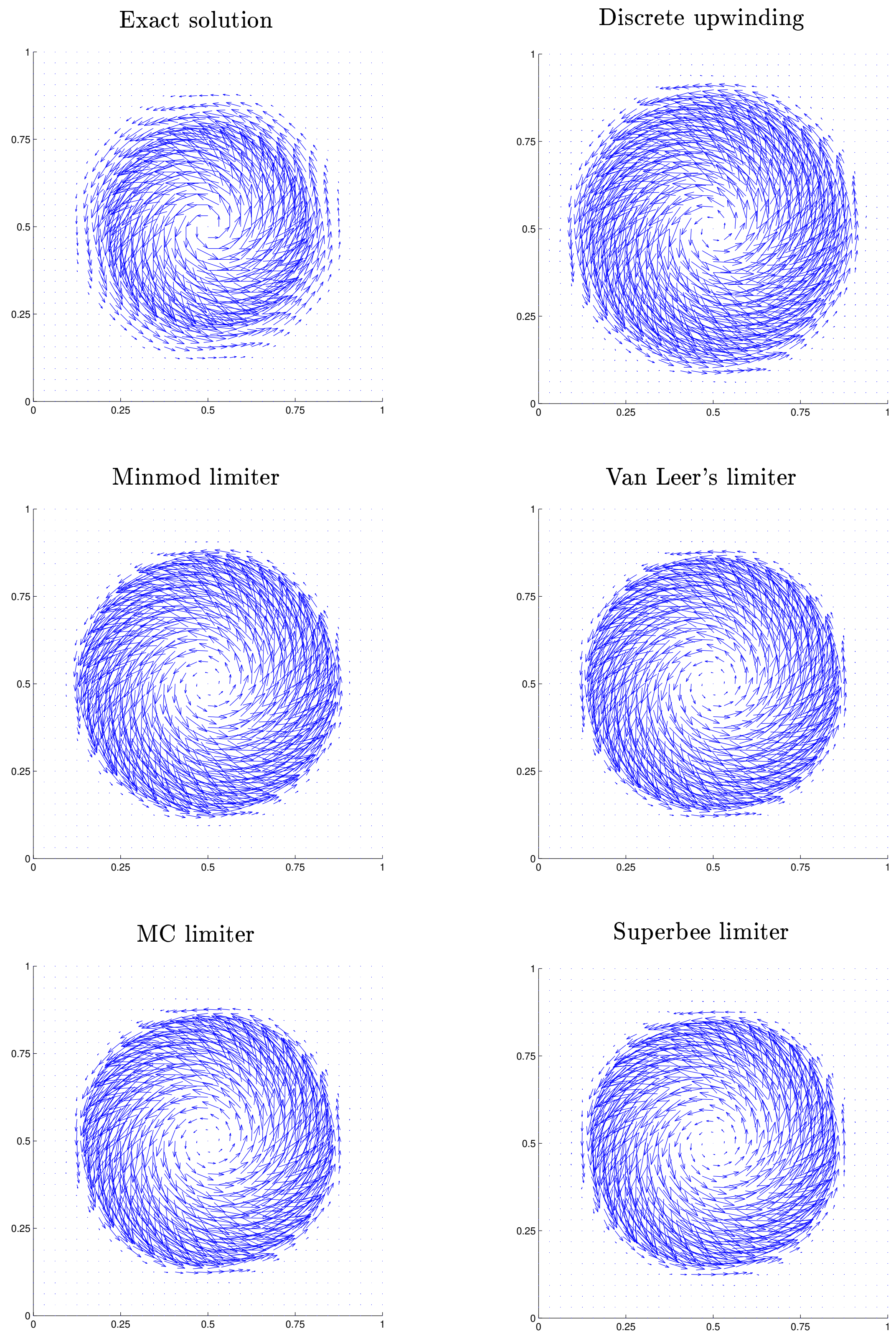


Figure 8. Standing vortex problem. FEM-TVD solution at $t = 3$.

Let us first apply a nonstationary projection solver to the well-known *standing vortex* problem in order to verify the dissipative properties of TVD limiters. The incompressible Navier-Stokes equations for an inviscid flow ($Re = \infty$) are solved in a unit square

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = 0, \quad \nabla \cdot \mathbf{u} = 0, \qquad \text{in } \Omega = (0,1)^2. \tag{61}$$

The initial condition is an axisymmetric vortex which also represents the exact steady-state solution. In polar coordinates, the velocity $\mathbf{u}$ can be decomposed into the radial component $u_r$ and the angular component $u_\theta$ which are initialized by

$$u_r = 0, \qquad u_\theta = \begin{cases} 5r, & r < 0.2, \\ 2 - 5r, & 0.2 \leq r \leq 0.4, \\ 0, & r > 0.4, \end{cases}$$

where $r = \sqrt{(x-0.5)^2 + (y-0.5)^2}$ denotes the distance from the center.

The objective is to test the ability of the discretization scheme to reproduce the original vortex. The numerical results produced by the Galerkin method equipped with discrete upwinding and four standard TVD limiters are compared to the exact solution in Figure 8. They were obtained at $t = 3$ using a mesh of $64 \times 64$ quadrilateral elements and backward Euler time-stepping. The artificial viscosity introduced by discrete upwinding is seen to degrade the accuracy of the solution appreciably, while FEM-TVD injects enough antidiffusion to alleviate the smearing of the vortex. The differences between the performance of the flux limiters under consideration are marginal but a closer look reveals that their diffusivities compare as discussed above for the Gaussian hill problem.

## 13.5 Flow around a cylinder

The second incompressible flow problem to be dealt with is the established benchmark *Flow around a cylinder* developed for the priority program "Flow simulation on high-performance computers" under the auspices of DFG, the German Research Association [40]. This project was intended to facilitate the evaluation of various numerical algorithms for the incompressible Navier-Stokes equations in the laminar flow regime. A quantitative comparison of simulation results is possible on the basis of relevant flow characteristics such as drag and lift coefficients, for which reliable reference values are available. Moreover, the efficiency of solution techniques can be assessed in an objective manner.

Let us consider the steady incompressible flow around a cylinder with a circular cross-section. An in-depth description of the geometrical details and boundary conditions for the 2D/3D case can be found in references [40],[46] which contain all relevant information regarding this benchmark configuration. The flow at $Re = 20$ is actually dominated by diffusion and could be simulated by the standard Galerkin method without any extra stabilization (as far as the discretization is concerned; the iterative solver may require using a stabilized preconditioner). Our goal is to investigate the behavior of FEM-TVD schemes for such low Reynolds number flows, for which their use is hardly optimal.

In particular, it is instructive to study the interplay of finite element discretizations for the convective and diffusive terms. As already mentioned above, discrete upwinding can be performed for the cumulative transport operator or just for the convective part. In the case of the nonconforming $\tilde{Q}_1$-elements, the discrete Laplacian operator originating from the Galerkin approximation of viscous terms is a positive-definite matrix but some of its off-diagonal coefficients are negative. Our numerical experiments indicate that it is worthwhile to leave it unchanged and apply the TVD postprocessing to the discretized convective term. In the case of linear or bilinear elements, physical diffusion can be taken into account in the formula (32) but the auxiliary quantities $P_i^\pm$ and $Q_i^\pm$ for the flux limiter should still be evaluated using the coefficients of the convective operator.

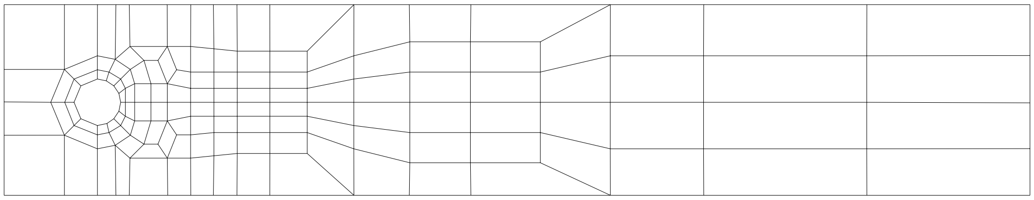

Figure 9. Coarse mesh for the DFG benchmark "Flow around a cylinder".

To generate the hierarchical data structures for the geometric multigrid algorithm which constitutes the core of our incompressible flow solver FEATFLOW, we introduce a sequence of successively refined quadrilateral meshes. The elements of the coarse mesh shown in Figure 9 are subdivided into four subelements at each refinement level, and the 2D mesh is extended into the third dimension for a 3D simulation. The two-dimensional results produced by the FEM-TVD algorithm embedded in a discrete projection method with pseudo-time-stepping are presented in Table 1. The computational mesh for the multigrid level NLEV contains NMT midpoints and NEL elements. For the employed $\tilde{Q}_1/Q_0$ finite element pair (rotated bilinear velocity, piecewise constant pressure), NMT represents the number of unknowns for each velocity component, while NEL equals the number of degrees of freedom associated with the pressure.

| NLEV | NMT | NEL | $C_D$ | $C_L$ |
|---|---|---|---|---|
| 3 | 4264 | 2080 | 5.6502 | $0.5233 \cdot 10^{-2}$ |
| 4 | 16848 | 8320 | 5.5852 | $0.8080 \cdot 10^{-2}$ |
| 5 | 66976 | 33280 | 5.5760 | $0.9946 \cdot 10^{-2}$ |
| 6 | 267072 | 133120 | 5.5707 | $0.1042 \cdot 10^{-1}$ |
| 7 | 1066624 | 532480 | 5.5692 | $0.0963 \cdot 10^{-1}$ |

Table 1. FEM-TVD (MC), projection solver.

The drag and lift coefficients listed in Table 1 exhibit a monotone convergence behavior but fall short of the reference values $C_D \approx 5.5795$, $C_L \approx 0.01061$ on fine meshes. This discrepancy can be attributed to the inexact evaluation of surface integrals. Furthermore, the accuracy of approximation for such derived quantities cannot be properly controlled in a flow solver developed primarily for genuinely time-dependent problems. Fortunately, the FEM-TVD algorithm is readily applicable in the stationary case, so that the use of pseudo-time-stepping is not mandatory. Taking advantage of this fact, we integrated TVD limiters into another FEATFLOW module which lends itself to the solution of the stationary

Navier-Stokes equations. It is based on the local Multilevel Pressure Schur Complement approach with adaptive patching, whereby small subproblems are solved exactly within an outer block-Gauss-Seidel/Jacobi iteration [41],[46]. This strongly coupled solution technique is very robust and far superior to projection schemes at low Reynolds numbers.

| NLEV | $C_D$ | $C_L$ | NL | MG | CPU |
|---|---|---|---|---|---|
| 3 | 5.8084 | $0.1733 \cdot 10^{-2}$ | 22 | 33 | 38 |
| 4 | 5.6514 | $0.7171 \cdot 10^{-2}$ | 10 | 19 | 87 |
| 5 | 5.6003 | $0.9734 \cdot 10^{-2}$ | 8 | 15 | 291 |
| 6 | 5.5854 | $0.1040 \cdot 10^{-1}$ | 7 | 13 | 1074 |
| 7 | 5.5811 | $0.1058 \cdot 10^{-1}$ | 5 | 9 | 3144 |

Table 2. FEM-TVD (MC), coupled solver.

| NLEV | $C_D$ | $C_L$ | NL | MG | CPU |
|---|---|---|---|---|---|
| 3 | 5.6699 | $0.5694 \cdot 10^{-2}$ | 9 | 26 | 29 |
| 4 | 5.6004 | $0.9700 \cdot 10^{-2}$ | 8 | 22 | 98 |
| 5 | 5.5841 | $0.1048 \cdot 10^{-1}$ | 7 | 19 | 364 |
| 6 | 5.5806 | $0.1060 \cdot 10^{-1}$ | 6 | 16 | 1280 |
| 7 | 5.5798 | $0.1061 \cdot 10^{-1}$ | 5 | 13 | 4390 |

Table 3. Samarski's upwind, coupled solver.

Tables 2 and 3 demonstrate that the drag and lift coefficients produced by the coupled solver are in a good agreement with the reference values and with those obtained using Samarski's upwind method based on a classical artificial viscosity which depends on the local Reynolds number [46]. It is worth mentioning that this finite-volume-like discretization of convective terms, which has traditionally been used in FEATFLOW, involves a free parameter which must be determined by trial and error. Clearly, the 'optimal' value is hard to find from *a priori* considerations. In addition, Samarski's hybrid method is only suitable for intermediate and low Reynolds numbers, since it becomes increasingly diffusive and degenerates into the standard upwind scheme in the limit of inviscid flow. At the same time, the nonlinear FEM-TVD discretization remains remarkably accurate for arbitrarily large Reynolds numbers (see the previous example), whereby the tradeoff between accuracy and stability is managed automatically by the flux limiter.

Note that the number of nonlinear iterations NL and linear multigrid steps MG reduces as the mesh is refined. Moreover, the improvement of convergence rates is faster than that for the hybrid upwind method which is more efficient than FEM-TVD on coarse meshes but less efficient on fine ones, as can be seen from the presented CPU times. In our experience, the differences are even more pronounced if Newton's method is employed. On the other hand, its advantages in comparison with the straightforward defect correction seem to fade at high Reynolds numbers, since the nonlinearity inherent to the discretization procedure plays an increasingly important role. In a nutshell, the design of robust and efficient iterative solvers for nonlinear high-resolution schemes with flux limiters is a nontrivial task, so there is a lot of room for further research.

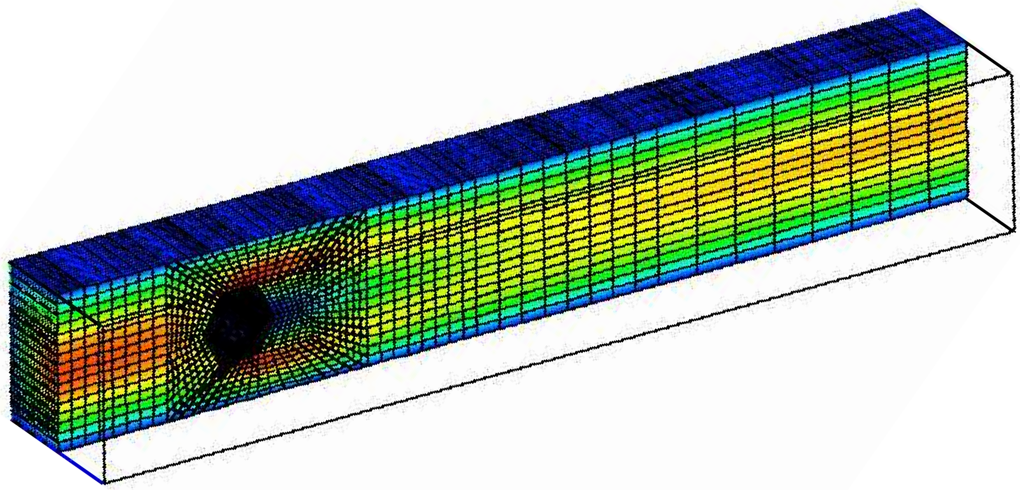

Figure 10. Horizontal velocity. FEM-TVD (MC), 3D projection solver.

Last but not least, it is worth mentioning that our 'black-box' postprocessing routine for the convective operator proved to be applicable in 3D without any modifications. The three-dimensional simulation results produced by the FEM-TVD algorithm combined with a projection solver from the FEATFLOW package are presented in Figure 10. The underlying mesh consists of 49152 hexahedral elements and gives rise to 151808 unknowns for each velocity component. In Table 4, the drag and lift coefficients for three different meshes are compared with the reference values published in [40] and with those obtained previously using finite volume upwinding (UPW), Samarski's hybrid scheme (SAM) and streamline diffusion (SD) stabilization [46]. The FEM-TVD method was found to perform very well, whereas the results produced by its competitors are rather sensitive to the value of the involved user-defined constant which is not known *a priori*.

| NLEV | UPW-1st | SAM-1.0 | SD-0.25 | SD-0.5 | TVD |
|---|---|---|---|---|---|
| 3 | 6.08/ 1.01 | 5.72/ 0.28 | 5.78/-0.44 | 5.98/-0.52 | 6.40/ 0.35 |
| 4 | 6.32/ 1.20 | 6.07/ 0.62 | 6.13/ 0.26 | 6.26/ 0.18 | 6.24/ 0.64 |
| 5 | 6.30/ 1.20 | 6.14/ 0.83 | 6.17/ 0.70 | 6.23/ 0.64 | 6.19/ 0.80 |
| Ref | 6.18/ 0.85 | 6.18/ 0.85 | 6.18/ 0.85 | 6.18/ 0.85 | 6.18/ 0.85 |

Table 4. FEM-TVD vs. upwind and streamline diffusion.

# 14 Conclusions

A generalization of one-dimensional TVD schemes to finite element discretizations on unstructured meshes was presented. The discrete transport operator was modified by elimination of negative off-diagonal coefficients followed by insertion of compensating anti-diffusion. The LED principle was utilized to generalize the concept of upwinding and to design a family of multidimensional TVD limiters controlling the local solution gradients or the ratio of diffusive/antidiffusive fluxes. The proposed algorithm renders a centered discretization of convective terms local extremum diminishing and positivity-preserving. It is characterized by a remarkable flexibility and the ease of implementation. Furthermore, it can be readily integrated into existing CFD software as a modular extension to the matrix assembly routine. The potential of the new methodology was demonstrated by application to scalar convection and incompressible flow problems. It can be extended to hyperbolic systems of conservation laws in the framework of an approximate Riemann solver using a suitable transformation to the characteristic variables (see [1],[10],[42],[43]).

**Acknowledgments**

The authors would like to thank Dr. R. Schmachtel for the implementation of FEM-TVD in the coupled Navier-Stokes solver and providing the data presented in Tables 2–3.

# References

[1] P. Arminjon and A. Dervieux, Construction of TVD-like artificial viscosities on 2-dimensional arbitrary FEM grids. *INRIA Research Report* **1111** (1989).

[2] K. Baba and M. Tabata, On a conservative upwind finite element scheme for convective diffusion equations. *RAIRO Numerical Analysis* **15** (1981) 3–25.

[3] T. J. Barth, Numerical aspects of computing viscous high Reynolds number flows on unstructured meshes. Technical report 91-0721, *AIAA paper*, 1991.

[4] T. J. Barth, Aspects of unstructured grids and finite volume solvers for the Euler and Navier-Stokes equations. In *von Karman Institute for Fluid Dynamics Lecture Series* Notes 1994-05, Brussels, 1994.

[5] J. P. Boris and D. L. Book, Flux-corrected transport. I. SHASTA, A fluid transport algorithm that works. *J. Comput. Phys.* **11** (1973) 38–69.

[6] J.-C. Carette, H. Deconinck, H. Paillère and P.L. Roe, Multidimensional upwinding: Its relation to finite elements. *Int. J. Numer. Methods Fluids* **20** (1995) Nr. 8-9, 935-955.

[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] B. Cockburn, G. E. Karniadakis and C.-W. Shu, The development of discontinuous Galerkin methods. In: *Discontinuous Galerkin methods. Theory, computation and applications,* Lect. Notes Comput. Sci. Eng. **11**, Springer, 2000, 3-50.

[9] H. Deconinck, H. Paillère, R. Struijs and P.L. Roe, Multidimensional upwind schemes based on fluctuation-splitting for systems of conservation laws. *Comput. Mech.* **11** (1993) Nr. 5-6, 323-340.

[10] J. Donea, V. Selmin and L. Quartapelle, Recent developments of the Taylor-Galerkin method for the numerical solution of hyperbolic problems. *Numerical methods for fluid dynamics III*, Oxford, 171-185 (1988).

[11] J.H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*. Springer, 1996.

[12] C.A.J. Fletcher, The group finite element formulation. *Comput. Methods Appl. Mech. Engrg.* **37** (1983) 225-243.

[13] S.K. Godunov, Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. *Mat. Sbornik* **47** (1959) 271-306.

[14] P. Hansbo, Aspects of conservation in finite element flow computations. *Comput. Methods Appl. Mech. Engrg.* **117** (1994) 423-437.

[15] A. Harten, High resolution schemes for hyperbolic conservation laws, *J. Comput. Phys.* **49** (1983) 357–393.

[16] A. Harten, On a class of high resolution total-variation-stable finite-difference-schemes. *SIAM J. Numer. Anal* 21 (1984) 1-23.

[17] C. Hirsch, *Numerical Computation of Internal and External Flows. Vol. II: Computational Methods for Inviscid and Viscous Flows.* John Wiley & Sons, Chichester, 1990.

[18] A. Jameson, Analysis and design of numerical schemes for gas dynamics 1. Artificial diffusion, upwind biasing, limiters and their effect on accuracy and multigrid convergence, *International Journal of Computational Fluid Dynamics* **4** (1995) 171-218.

[19] A. Jameson, Computational algorithms for aerodynamic analysis and design. *Appl. Numer. Math.* **13** (1993) 383-422.

[20] A. Jameson, Positive schemes and shock modelling for compressible flows. *Int. J. Numer. Meth. Fluids* **20** (1995) 743–776.

[21] D. Kuzmin, Positive finite element schemes based on the flux-corrected transport procedure, In: *Computational Fluid and Solid Mechanics*, Elsevier, 887-888 (2001).

[22] D. Kuzmin and S. Turek, Flux correction tools for finite elements. *J. Comput. Phys.* **175** (2002) 525-558.

[23] D. Kuzmin, M. Möller and S. Turek, Multidimensional FEM-FCT schemes for arbitrary time-stepping. *Int. J. Numer. Meth. Fluids* **42** (2003) 265-295.

[24] D. Kuzmin, M. Möller and S. Turek, Implicit flux-corrected transport algorithm for finite element simulation of the compressible Euler equations. To appear in: M. Krizek *et al.* (eds), *Finite Element Methods: Fifty Years of Conjugate Gradients*, Proc. Conf., Univ. of Jyväskylä, 2002, Springer, Berlin.

[25] D. Kuzmin and S. Turek, Finite element discretization tools for gas-liquid flows. In: M. Sommerfeld (ed.), *Bubbly Flows: Analysis, Modelling and Calculation*, Springer, 2004, 191-201.

[26] D. Kuzmin, M. Möller and S. Turek, High-resolution FEM-FCT schemes for multidimensional conservation laws. Technical report **231**, University of Dortmund, 2003. Submitted to *Comput. Methods Appl. Mech. Engrg.*

[27] A. Lapin, University of Stuttgart. Private communication.

[28] P. D. Lax, *Systems of Conservation Laws and Mathematical Theory of Shock Waves.* SIAM Publications, Philadelphia, 1973.

[29] R. J. LeVeque, *Numerical Methods for Conservation Laws.* Birkhäuser, 1992.

[30] R. J. LeVeque, High-resolution conservative algorithms for advection in incompressible flow. *Siam J. Numer. Anal.* **33** (1996) 627–665.

[31] R. Löhner, *Applied CFD Techniques: An Introduction Based on Finite Element Methods.* Wiley, 2001.

[32] R. Löhner, K. Morgan, J. Peraire and M. Vahdati, Finite element flux-corrected transport (FEM-FCT) for the Euler and Navier-Stokes equations. *Int. J. Numer. Meth. Fluids* **7** (1987) 1093–1109.

[33] R. Löhner, K. Morgan, M. Vahdati, J. P. Boris and D. L. Book, FEM-FCT: combining unstructured grids with high resolution. *Commun. Appl. Numer. Methods* **4** (1988) 717–729.

[34] P. R. M. Lyra, *Unstructured Grid Adaptive Algorithms for Fluid Dynamics and Heat Conduction.* PhD thesis, University of Wales, Swansea, 1994.

[35] P. R. M. Lyra, K. Morgan, J. Peraire and J. Peiro, TVD algorithms for the solution of the compressible Euler equations on unstructured meshes. *Int. J. Numer. Meth. Fluids* **19** (1994) 827–847.

[36] K. Morgan and J. Peraire, Unstructured grid finite element methods for fluid mechanics. *Reports on Progress in Physics*, **61** (1998), no. 6, 569-638.

[37] S. V. Patankar, *Numerical Heat Transfer and Fluid Flow.* McGraw-Hill, New York, 1980.

[38] J. Peraire, M. Vahdati, J. Peiro and K. Morgan, The construction and behaviour of some unstructured grid algorithms for compressible flows. *Numerical Methods for Fluid Dynamics* **IV**, Oxford University Press, 221-239 (1993).

[39] R. Rannacher and S. Turek, A simple nonconforming quadrilateral Stokes element. *Numer. Meth. PDEs* **8** (1992), no. 2, 97-111.

[40] M. Schäfer and S. Turek (with support of F. Durst, E. Krause, R. Rannacher), Benchmark computations of laminar flow around cylinder, In E.H. Hirschel (ed.), *Flow Simulation with High-Performance Computers II*, Vol. 52 von *Notes on Numerical Fluid Mechanics*, Vieweg, 1996, 547–566.

[41] R. Schmachtel, *Robuste lineare und nichtlineare Lösungsverfahren für die inkompressiblen Navier-Stokes-Gleichungen.* PhD thesis, University of Dortmund, 2003.

[42] V. Selmin, Finite element solution of hyperbolic equations. I. One-dimensional case. *INRIA Research Report* **655** (1987).

[43] V. Selmin, Finite element solution of hyperbolic equations. II. Two-dimensional case. *INRIA Research Report* **708** (1987).

[44] P. K. Sweby, High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.* **21** (1984), 995–1011.

[45] S. Turek *et al.*, *FEATFLOW: finite element software for the incompressible Navier-Stokes equations.* User manual, University of Dortmund, 2000. Available at the URL *http://www.featflow.de.*

[46] S. Turek, *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach*, LNCSE 6, Springer, 1999.

[47] S. T. Zalesak, Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.* **31** (1979) 335–362.