Lawrence Berkeley National Laboratory
Recent Work

Title
A single-stage flux-corrected transport algorithm for high-order finite-volume methods

Permalink
https://escholarship.org/uc/item/5hx8m952

Journal
Communications in Applied Mathematics and Computational Science, 12(1)

ISSN
1559-3940

Authors
Chaplin, C
Colella, P

Publication Date
2017

DOI
10.2140/camcos.2017.12.1

Peer reviewed
A SINGLE STAGE FLUX-CORRECTED TRANSPORT ALGORITHM FOR HIGH-ORDER FINITE-VOLUME METHODS

CHRISTOPHER CHAPLIN AND PHILLIP COLELLA

ABSTRACT. We present a new limiter method for solving the advection equation using a high-order, finite-volume discretization. The limiter is based on the flux-corrected transport algorithm. We modify the classical algorithm by introducing a new computation for solution bounds at smooth extrema, as well as improving the pre-constraint on the high-order fluxes. We compute the high-order fluxes via a method of lines approach with fourth order Runge-Kutta as the time integrator. For computing low-order fluxes, we select the corner transport upwind method due to its improved stability over donor-cell upwind. Several spatial differencing schemes are investigated for the high-order flux computation, including centered difference and upwind schemes. We show that the upwind schemes perform well on account of the dissipation of high wavenumber components. The new limiter method retains high-order accuracy for smooth solutions and accurately captures fronts in discontinuous solutions. Further, we need only apply the limiter once per complete time step.

1. Introduction

We wish to solve hyperbolic conservation laws of the form

\[ \frac{\partial U}{\partial t} + \nabla \cdot (\tilde{\mathbf{F}}(U)) = 0 \]  

(1)

where \( U \) represents a vector of conserved values and \( \tilde{\mathbf{F}}(U) \), \( \tilde{\mathbf{F}} = (F^1 \ldots F^D) \) are corresponding fluxes. The discrete solution of these equations at a given time \( t^{n+1} \) and spatial location \( i \) is given by:

\[ \langle U \rangle^{n+1}_i = \langle U \rangle^n_i - \frac{\Delta t}{\Delta x} \sum_d \left[ (\tilde{F}^d)^{n+\frac{1}{2}}_{i+\frac{1}{2}e_d} - (\tilde{F}^d)^{n+\frac{1}{2}}_{i-\frac{1}{2}e_d} \right] 

(2) \]
where \( \langle U \rangle^n \) and \( (F^d)^{n+\frac{1}{2}} \) approximate the average of \( U \) over a rectangular Cartesian control volumes indexed by \( i \) at time step \( n \) and the average of the fluxes in time over the respective faces, respectively.

Methods for accurately computing these fluxes \( (F^d)^{n+\frac{1}{2}} \) to obtain high-order accuracy for smooth solutions are well-understood. However, in the presence of discontinuities or underresolved gradients, such methods generate oscillatory errors, and limiters are typically introduced to selectively introduce dissipation that will damp these errors. For high-order methods based on the method of lines, the typical approach to limiters has been to apply a limiter independently at every flux evaluation in a multi-stage high-order time integration method, e.g. Runge-Kutta methods. However, such an approach is problematic, particularly for limiters that are intended to preserve high-order accuracy at smooth extrema. Such limiters typically make use of second- or higher-order spatial derivative information at extrema to determine whether the solution is smooth enough to not require limiting. This leads to the use of a larger stencil (for the limiter) at every stage of the time integration than that of the simpler high-order spatial discretization of the fluxes without limiting, thus increasing communication costs in a domain-decomposed parallel algorithm. In addition, only the linear combination of Runge-Kutta stage computations is guaranteed to produce a high-order accurate solution; the individual stage fluxes and solutions are typically lower-order accurate in time. In adaptive mesh refinement methods with refinement in time, this feature complicates the design of the time interpolation step for computing ghost cell values at refinement boundaries [4]. Finally, many of the stage-by-stage limiting schemes are geometric in nature, such that the effective fully limited method corresponds to donor-cell spatial differencing for the fluxes [7, 10]. This typically has a more restrictive CFL time step stability condition than the high-order methods.

The starting point for addressing these problems is to use a version of the classic Flux-Corrected Transport (FCT) algorithms [1, 9, 13]. These methods introduce dissipation through a nonlinear hybridization of a high-order flux with a dissipative, low-order flux. The FCT algorithm is particularly advantageous because it has a straightforward multi-dimensional expression [13] and has been used extensively with finite-volume discretizations. We will look at a family of high-order methods based on centered- and one-point-upwinded linear finite volume interpolation, following the ideas in [4], combined with the method of lines with fourth-order Runge-Kutta time discretization as the method for time discretization. All four intermediate stages of this high-order method will be computed without any limiting applied, in order to produce a high-order flux to go into the FCT hybridization. The low-order method will be the corner-transport upwind (CTU) method [2, 11]. The method for computing the hybridization coefficient will be based on the approach used in [4, 5] for
interpolation-based limiting, modified to fit into the formalism in [13]. This combination of methods addresses the issues alluded to above. The CFL stability condition for the CTU method is independent of the direction of propagation of the waves relative to the coordinate directions and of the dimensionality of the problem, and is typically much less restrictive than the CFL condition for donor-cell / method-of-lines. The limiter is applied only once – at the end of the time step – thus minimizing the impact on the communications cost.

For this study, we restrict our attention to the equation of scalar advection equation. This allows us to explore design decisions in a simple setting, but one that is still relatively unforgiving, and important for real applications (transport of scalars in the atmosphere, Vlasov equations in phase space). The extension to systems will be discussed in future work. One question we will not address here is that of positivity preservation. The reason is that, in more than one dimension for any linear first-order method that is not donor cell, it is possible to construct a discretely divergence-free advection velocity and non-negative initial data such that the solution becomes negative after one time step. Our preferred approach is to post-process the solution at the end of each time step by redistributing negative increments to nearby cells in a way that would lead to an overall non-negative solution, while still conserving [8, 12].

1.1. Advection Equation. We will consider the linear advection equation in the following form.

\[
\frac{\partial q}{\partial t} + \nabla \cdot (q\vec{u}) = 0 \tag{3}
\]

\[
\nabla \cdot \vec{u} = 0 \tag{4}
\]

on a \( D \)-dimensional square domain \( \Omega = [0, 1]^D \). In this case \( \vec{u} \) is an advective velocity and \( q \) is a scalar field. The partial differential equation above can also be written as:

\[
\frac{dq}{dt} = 0 \quad \frac{d\vec{x}}{dt} = \vec{u} \tag{5}
\]

Provided that an initial condition is specified \( (q_0 = q(\vec{x}(t_0), t_0)) \) this system of ordinary differential equations yields a unique solution for any \( q(\vec{x}(t), t) \) and \( \vec{x}(t) \). The solution arrived at by integrating the equations is that \( q \) is constant along characteristic curves defined by \( \vec{x}(t) \). Even though there is a simple solution to this equation, the
analysis is still quite useful since there is no diffusion or entropy condition built-in to the equation: any numerical errors introduced are propagated through the domain.

1.2. Finite-Volume Discretization. Our solution approach is to use a finite-volume method to discretize the physical domain into a union of control volumes $V_i$ (Eq. 6)

$$V_i = [ih, (i + e)h], \ i \in \mathbb{Z}^D, \ e = (1, 1, ..., 1)$$

where $h$ is the grid spacing ($\Delta x$) and $i$ is a $D$-dimensional index denoting location. Values of the conserved scalar quantity $q$ are stored as cell-averages $\langle q \rangle$ over each cell $V_i$ (Eq. 7). The fluxes $F^d = qu^d$ are stored as averages $\langle F^d \rangle_{i \pm \frac{1}{2}e^d}$ over the surface faces $A^\pm_d$ of each cell (Eq. 8).

$$\langle q \rangle_i(t) = \frac{1}{h^D} \int_{V_i} q(x, t)dx$$

$$\langle F^d \rangle_{i \pm \frac{1}{2}e^d}(t) = \frac{1}{h^{D-1}} \int_{A^\pm_d} F^d(x, t) \, dx$$

Applying the finite-volume discretization (Eq. 6) to Eq. 3 yields a semi-discrete system of ordinary differential equations (ODE) in time:

$$\frac{d\langle q \rangle_i}{dt} = -\frac{1}{h^D} \int_{V_i} (\nabla \cdot (\vec{F})) dx$$

$$= -\frac{1}{h} \sum_d [\langle F^d \rangle_{i + \frac{1}{2}e^d} - \langle F^d \rangle_{i - \frac{1}{2}e^d}]$$

where Eq. 10 is the result of applying the divergence theorem to Eq. 9. The integration of the above system with respect to time from $t^n$ to $t^{n+1}$ produces the solution given below:
The resulting challenge is to accurately compute $\langle F_d \rangle_{i+\frac{1}{2}e^d}$. It is important to note that no approximations have been made at this point: Eq. 11 - Eq. 12 are exact relationships. However, to obtain a full discrete approximation, we need quadrature rules for the surface fluxes in Eq. 10 and for the time-averaged fluxes in Eq. 12. The quadrature rules for computing these fluxes are defined following ideas from [4]. In that work, the high-order quadratures were computed using a method of lines approach. The surface fluxes were computed using a high-order centered difference method and the temporal integration was computed using the classic fourth-order Runge-Kutta (RK4) method. We retained the use of RK4 in this study and investigated several high-order methods for computing the surface fluxes.

1.3. Hybridization. Returning to the flux description in Eq. 12, we may now define the hybridization

$$\langle F_d \rangle _{i+\frac{1}{2}e^d} = (\eta_{i+\frac{1}{2}e^d}) \langle F_d^H \rangle _{i+\frac{1}{2}e^d} + (1 - \eta_{i+\frac{1}{2}e^d}) \langle F_d^L \rangle _{i+\frac{1}{2}e^d}$$

(13)

where the subscripts $H$ and $L$ refer to the high-order and low-order fluxes, and $\eta_{i+\frac{1}{2}e^d}$ is the hybridization coefficient.

In the following sections of the paper we will describe the design choices and procedures for computing the high-order flux, the low-order flux, and the hybridization coefficient.

2. High-Order Flux Computation

We compute the high-order fluxes using the method of lines. Two schemes must be chosen: a scheme for integrating the solution in time and a scheme for computing the spatial derivatives. High-order accuracy requires that both schemes be high-order accurate.
2.1. High-Order Temporal Integration Scheme. We use the RK4 scheme to advance the solution forward in time. Returning to the system of ODE (Eq. 9):

\[
\frac{d\langle q \rangle}{dt} = -D \cdot \langle \vec{F} \rangle(t)
\]

\[
= -\frac{1}{h} \sum_d \langle F^d \rangle_{i+\frac{1}{2}e^d} - \langle F^d \rangle_{i-\frac{1}{2}e^d}
\]

we want to integrate the system from \( t^n \) to \( t^{n+1} \). RK4 is a fourth-order integration scheme that consists of computing a linear combination of stage update variables \( k_s \). The updates are defined below:

\[
\langle q \rangle^0 = \langle q \rangle(t^n) \quad k_1 = -D \cdot \langle \vec{F}(\langle q \rangle^0) \rangle \Delta t
\]

\[
\langle q \rangle^1 = \langle q \rangle^0 + \frac{k_1}{2} \quad k_2 = -D \cdot \langle \vec{F}(\langle q \rangle^1) \rangle \Delta t
\]

\[
\langle q \rangle^2 = \langle q \rangle^0 + \frac{k_2}{2} \quad k_3 = -D \cdot \langle \vec{F}(\langle q \rangle^2) \rangle \Delta t
\]

\[
\langle q \rangle^3 = \langle q \rangle^0 + k_3 \quad k_4 = -D \cdot \langle \vec{F}(\langle q \rangle^3) \rangle \Delta t
\]

Each update variable \( k_s \) requires computing stage fluxes \( \langle F^d \rangle_s^{i+\frac{1}{2}e^d} = \langle q u^d \rangle_s^{i+\frac{1}{2}e^d} \). The stage fluxes are functions of the stage values \( \langle q \rangle_s^i \) and \( \langle u^d \rangle_s^i \) alone and the procedure for computing the fluxes will be described in the next section.

To perform the RK4 integration:

\[
\langle q \rangle(t^n + \Delta t) = \langle q \rangle(t^n) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)
\]

Returning to the conservation notation, this RK4 integration can be described by

\[
\langle q \rangle_{i}^{n+1} = \langle q \rangle_{i}^{n} - \frac{\Delta t}{h} \sum_{d=1}^{P} \left[ \langle F^d_H \rangle_{i+\frac{1}{2}e^d} - \langle F^d_H \rangle_{i-\frac{1}{2}e^d} \right]
\]

\[
\langle F^d_H \rangle_{i+\frac{1}{2}e^d} = \frac{1}{6} \left[ \langle F^d_H \rangle_{i+\frac{1}{2}e^d}^{(0)} + 2\langle F^d_H \rangle_{i+\frac{1}{2}e^d}^{(1)} + 2\langle F^d_H \rangle_{i+\frac{1}{2}e^d}^{(2)} + \langle F^d_H \rangle_{i+\frac{1}{2}e^d}^{(3)} \right]
\]
2.2. High-Order Spatial Difference Schemes. We use high-order finite-difference methods to approximate the surface fluxes associated with the spatial derivatives. The fluxes \( \langle q u^d \rangle_{i \pm \frac{1}{2} e^d} \) for the spatial derivatives are functions only of the cell-averaged \( \langle q \rangle_i \) and \( \langle u^d \rangle_i \) at any time. Several methods were explored in this study for computing \( \langle q \rangle_{i \pm \frac{1}{2} e^d} \), including high-order centered difference schemes and upwind schemes. The advantage of the upwind methods is that they have greater diffusion especially in regimes where the phase error begins to rise (Figure 1). The upwind methods only require a small additional computation and the stability between similar order centered difference and upwind methods is almost identical (Table 1).

The interpolation formulae corresponding to the spatial differencing schemes used are presented below. For compactness, the following notation will be used:

\[
\langle q \rangle^n_{i + \frac{1}{2} e^d} = \sum_{s=-S}^{S} a_s \langle q \rangle^n_{i + s e^d}
\]

where \( S \) is the width of the stencil and \( a_s \) are the coefficients. The odd ordered methods use the full range of coefficients, whereas the even ordered methods have no coefficient at \( s = -S \).

Fourth Order Centered Difference (\( S = 2 \))

\[
\{a_s : s = -S + 1, \ldots, S\} = \frac{1}{12} \{-1, 7, 7, -1\}
\]
Fifth Order Upwind \((S = 2)\)
\[
\{a_s : s = -S, \ldots, S\} = \frac{1}{60}\{2, -13, 47, 27, -3\} \tag{24}
\]

Sixth Order Centered Difference \((S = 3)\)
\[
\{a_s : s = -S + 1, \ldots, S\} = \frac{1}{60}\{1, -8, 37, 37, -8, 1\} \tag{25}
\]

Seventh Order Upwind \((S = 3)\)
\[
\{a_s : s = -S, \ldots, S\} = \frac{1}{420}\{-3, 25, -101, 319, 214, -38, 4\} \tag{26}
\]

Ninth Order Upwind \((S = 4)\)
\[
\{a_s : s = -S, \ldots, S\} = \frac{1}{2520}\{4, -41, 199, -641, 1879, 1375, -305, 55, -5\} \tag{27}
\]

2.3. **Product Rule.** To complete the flux computation, we must compute the average of the product of the scalar variable and the velocity \((\langle qu^d \rangle_{i+\frac{1}{2}e^d})\). The 2D product rules for second, fourth, and sixth-order accuracy are:

\[
\langle qu^d \rangle_{i+\frac{1}{2}e^d} = \langle q \rangle_{i+\frac{1}{2}e^d}\langle u^d \rangle_{i+\frac{1}{2}e^d} + O(h^2) \tag{28}
\]

\[
\langle qu^d \rangle_{i+\frac{1}{2}e^d} = \langle q \rangle_{i+\frac{1}{2}e^d}\langle u^d \rangle_{i+\frac{1}{2}e^d} + \frac{h^2}{12} \sum_{d' \neq d} \frac{\partial q}{\partial x_{d'}} \frac{\partial u^d}{\partial x_{d'}} + O(h^4) \tag{29}
\]

\[
\langle qu^d \rangle_{i+\frac{1}{2}e^d} = \langle q \rangle_{i+\frac{1}{2}e^d}\langle u^d \rangle_{i+\frac{1}{2}e^d} + \frac{h^2}{12} \sum_{d' \neq d} \left( \frac{\partial q}{\partial x_{d'}} \frac{\partial u^d}{\partial x_{d'}} \right) + \frac{h^4}{1440} \sum_{d' \neq d} \left( 3 \frac{\partial^3 q}{\partial x_{d'}^3} \frac{\partial u^d}{\partial x_{d'}} + 3 \frac{\partial^3 u^d}{\partial x_{d'}^3} \frac{\partial q}{\partial x_{d'}} + 2 \frac{\partial^2 u^d}{\partial x_{d'}^2} \frac{\partial^2 q}{\partial x_{d'}^2} \right) + O(h^6) \tag{30}
\]

The possible sources of error in the product formulae above are computing the averages \(\langle q \rangle_{i+\frac{1}{2}e^d}, \langle u^d \rangle_{i+\frac{1}{2}e^d}\), and computing the partial derivative sums. We have already discussed several methods and their accuracy for computing \(\langle q \rangle_{i+\frac{1}{2}e^d}\). The velocity fields are analytic for advection, so \(\langle u^d \rangle_{i+\frac{1}{2}e^d}\) introduces no error. The derivative terms in the summations above were computed exclusively using centered difference approximations of appropriate accuracy. For example, the derivatives in the fourth-order accurate product formula were computed using a second-order centered difference. The derivatives in the sixth-order formula were computed to fourth-order (for the term multiplied by \(h^2\)) and to second-order (for the term multiplied by \(h^4\)).
Table 1. Stability of Method for Varying Spatial Difference Operators

<table>
<thead>
<tr>
<th>Method</th>
<th>Stability Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th Center</td>
<td>$\sigma \lesssim 2.06/D$</td>
</tr>
<tr>
<td>5th Upwind</td>
<td>$\sigma \lesssim 1.73/D$</td>
</tr>
<tr>
<td>6th Center</td>
<td>$\sigma \lesssim 1.78/D$</td>
</tr>
<tr>
<td>7th Upwind</td>
<td>$\sigma \lesssim 1.69/D$</td>
</tr>
<tr>
<td>9th Upwind</td>
<td>$\sigma \lesssim 1.60/D$</td>
</tr>
</tbody>
</table>

2.4. Stability. We compute the stability for each high-order scheme to determine the allowable time step size following the procedure in [3]. Stability for the method of lines requires the eigenvalues of the right hand side to lie within the stability region of the time integrator. These eigenvalues are computed by diagonalizing the the semi-discrete system (Eq. 14). For advection the eigenvalues are defined as the product of the velocity and the spatial derivative operator (Eq. 31-Eq. 32). The particular eigenvalues for each spatial differencing scheme will be presented later.

\[
\frac{d\langle q \rangle}{dt} = \lambda \langle q \rangle
\]  
\[
\lambda \langle q \rangle = -\vec{u} \frac{\partial}{\partial x} \langle q \rangle
\]

These eigenvalues must lie within the stability region of the time integrator. The stability region for RK4 is well known, and can be described by its characteristic polynomial

\[
P(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24}
\]

where $z = \Delta t \lambda$. Stability for this problem requires that $|P(z)| \leq 1$. The resulting stability constraints for each spatial differencing scheme are presented in Table 1 where $\sigma = |u| \Delta t / h$.

Along with stability, the phase error and dissipation were computed (Figure 1). The dissipation was defined as $(1 - |g|)$ where:
\[ |g| = \sqrt{(\text{Re}(g)^2 + \text{Im}(g)^2)} \]  \hspace{1cm} (34)

\[ \text{Re}(g) = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}\right) - \frac{y^2}{2} \left(1 + x + \frac{x^2}{2}\right) + \frac{y^4}{24} \]  \hspace{1cm} (35)

\[ \text{Im}(g) = y \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6}\right) - \frac{y^3}{6} (1 + x) \]  \hspace{1cm} (36)

and \( \lambda = x + iy \). The normalized phase error, \( |1 - \alpha| \), is defined using the following computation

\[ \alpha = \frac{\alpha(\beta)}{|u|} = -\frac{1}{\sigma \beta} \frac{\text{Im}(g)}{\text{Re}(g)} \]  \hspace{1cm} (37)

2.4.1. Spatial Differencing Eigenvalues. The eigenvalues for each of the different high-order spatial differencing schemes are presented below. In each of the eigenvalue descriptions, \( \beta_d \) may range from \(-\pi\) to \(\pi\), and is defined as \(2\pi k_d h\) with \(k_d = 0, \pm 1, \pm 2, ..., \pm N/2\).

**Fourth Order Centered Difference:**

\[ \lambda_4 = \frac{i}{12h} \sum_{d=1}^{D} u^d \left[16 \sin(\beta_d) - 2 \sin(2\beta_d)\right] \]  \hspace{1cm} (38)

**Fifth Order Upwind:**

\[ \lambda_5 = \frac{1}{60h} \sum_{d=1}^{D} u^d \left[-2 \cos(3\beta_d) + 12 \cos(2\beta_d) - 30 \cos(\beta_d) + 20\right] + \]

\[ i \left(2 \sin(3\beta_d) - 18 \sin(2\beta_d) + 90 \sin(\beta_d)\right) \]  \hspace{1cm} (39)

**Sixth Order Centered Difference:**

\[ \lambda_6 = \frac{i}{60h} \sum_{d=1}^{D} u^d \left[2 \sin(3\beta_d) - 18 \sin(2\beta_d) + 90 \sin(\beta_d)\right] \]  \hspace{1cm} (40)
SINGLE STAGE FLUX-CORRECTED TRANSPORT ALGORITHM

Seventh Order Upwind:

\[ \lambda_7 = \frac{1}{420h} \sum_{d=1}^{D} u^d \left[ (3 \cos(4\beta_d) - 24 \cos(3\beta_d) + 84 \cos(2\beta_d) - 168 \cos(\beta_d) + 105) + \right. \\
\left. i \left( -3 \sin(4\beta_d) + 32 \sin(3\beta_d) - 168 \sin(2\beta_d) + 672 \sin(\beta_d) \right) \right] \] (41)

Ninth Order Upwind:

\[ \lambda_9 = \frac{1}{2520h} \sum_{d=1}^{D} u^d \left[ (-4 \cos(5\beta_d) + 40 \cos(4\beta_d) - 180 \cos(3\beta_d) + \\
480 \cos(2\beta_d) - 840 \cos(\beta_d) + 504) + \right. \\
\left. i \left( 4 \sin(5\beta_d) - 50 \sin(4\beta_d) + 300 \sin(3\beta_d) - \\
1200 \sin(2\beta_d) + 4200 \sin(\beta_d) \right) \right] \] (42)

3. Low-Order Flux Computation

The low order fluxes are computed using the CTU method \[2, 11\]. CTU is a first-order time advancement scheme. The method is desirable over the simpler donor-cell upwind method because its stability is independent of dimensionality. However, this increased stability comes with a price. Instead of a single flux being defined by a single upwind value, the CTU flux is dependent upon a set of upwinded values. These values are determined by tracing the characteristic paths from the nodes that define the flux surface. This process involves an increasing number of Riemann solves as the dimensionality of the problem increases. In the 1D case, CTU is identical to donor-cell upwind.

4. Computing the Hybridization Coefficient

We compute the hybridization coefficient \( \eta \) using a modified multidimensional flux-corrected transport (FCT) algorithm. Note that the time superscript notation \((n)\) for fluxes is dropped from here out, but it is implied. Our algorithm is based upon the method described first in \[13\]. Here is the generic FCT procedure:

(1) Compute the high-order fluxes \( \langle F^d_{H} \rangle_{i \pm \frac{1}{2}^e} \) over the cell volume \( V_i \).
(2) Compute the low-order fluxes $\langle F^d_L \rangle_{i \pm \frac{1}{2} e^d}$ and the corresponding low-order update $\langle q \rangle_{i}^{td}$.

$$\langle q \rangle_{i}^{td} = \langle q \rangle_{i}^{n} - \frac{\Delta t}{h} \sum_d \left[ (\langle F^d_L \rangle_{i+\frac{1}{2}e^d} - \langle F^d_L \rangle_{i-\frac{1}{2}e^d}) \right]$$  \hspace{1cm} (43)

(3) Compute the antidiffusive fluxes $\langle A^d \rangle_{i \pm \frac{1}{2} e^d}$

$$\langle A^d \rangle_{i \pm \frac{1}{2} e^d} = \langle F^d_H \rangle_{i \pm \frac{1}{2} e^d} - \langle F^d_L \rangle_{i \pm \frac{1}{2} e^d}$$ \hspace{1cm} (44)

(4) Limit the antidiffusive fluxes

$$\langle A^d_{\eta} \rangle_{i \pm \frac{1}{2} e^d} = \eta^d_{i \pm \frac{1}{2} e^d} \langle A^d \rangle_{i \pm \frac{1}{2} e^d}$$ \hspace{1cm} (45)

\[0 \leq \eta^d_{i \pm \frac{1}{2} e^d} \leq 1\]

(5) Update the solution with the limited antidiffusive fluxes

$$\langle q \rangle_{i}^{n+1} = \langle q \rangle_{i}^{n} - \frac{\Delta t}{h} \sum_d \left[ \langle A^d_{\eta} \rangle_{i+\frac{1}{2}e^d} - \langle A^d_{\eta} \rangle_{i-\frac{1}{2}e^d} \right]$$  \hspace{1cm} (46)

4.1. Limiting the antidiffusive flux. The primary challenge in the above formulation is computing the hybridization coefficients $(\eta_{i \pm \frac{1}{2} e^d})$. Following the procedure in [13], we compute the coefficients in the following manner:

Pre-constrain the high-order fluxes $\langle F^d_H \rangle_{i \pm \frac{1}{2} e^d}$. This is a pre-limiting step that seeks to keep the high-order fluxes from generating new or steeper extrema.

Compute the sum $P^\pm_i$ of all the antidiffusive fluxes into and out of the cell and a measure of the diffusion $(Q^\pm_i)$. 

\[
P_i^+ = \sum_{d} \left[ \max \left( \langle A^d \rangle_{i+\frac{1}{2}e^d}, 0 \right) - \min \left( \langle A^d \rangle_{i-\frac{1}{2}e^d}, 0 \right) \right]
\]
(47)

\[
Q_i^+ = \left( (q_{\text{max}})_i - \langle q \rangle_i \right) \frac{h}{\Delta t}
\]
(48)

\[
P_i^- = \sum_{d} \left[ \max \left( \langle A^d \rangle_{i-\frac{1}{2}e^d}, 0 \right) - \min \left( \langle A^d \rangle_{i+\frac{1}{2}e^d}, 0 \right) \right]
\]
(49)

\[
Q_i^- = \left( \langle q \rangle_i - (q_{\text{min}})_i \right) \frac{h}{\Delta t}
\]
(50)

Compute the least upper bounds \( R_i^\pm \).

\[
R_i^+ = \begin{cases} 
\min \left( 1.0, \frac{Q_i^+}{P_i^+} \right) & \text{if } P_i^+ > 0.0 \\
0.0 & \text{otherwise}
\end{cases}
\]
(51)

\[
R_i^- = \begin{cases} 
\min \left( 1.0, \frac{Q_i^-}{P_i^-} \right) & \text{if } P_i^- > 0.0 \\
0.0 & \text{otherwise}
\end{cases}
\]
(52)

Select the hybridization coefficient with the most restrictive upper bound.

\[
\eta_{i+\frac{1}{2}e^d} = \begin{cases} 
\min \left( R_{i+\frac{1}{2}e^d}^+, R_i^+ \right) & \text{if } \langle A^d \rangle_{i+\frac{1}{2}e^d} > 0.0 \\
\min \left( R_i^+, R_{i+\frac{1}{2}e^d}^- \right) & \text{if } \langle A^d \rangle_{i+\frac{1}{2}e^d} \leq 0.0
\end{cases}
\]
(53)

In the above description the user is provided with two design choices: pre-constraint for the high-order flux and method of computing the solution bounds \((q_{\text{max}})_i\) and \((q_{\text{min}})_i\).

### 4.2. Computing the Solution Bounds.

Compute initial estimates of the solution bounds, \((q_{\text{max}})_i\) and \((q_{\text{min}})_i\). First, compute the bounded solutions in a rectangular stencil \((B_i)\) that is \([2s_i + 1]^D\) cells in size, where \(s_i\) is the stencil size. In this study, \(s_i\) was allowed to vary depending on the local velocity field in the following manner:
\[ s_i = \begin{cases} 2 & \text{if } \sigma \max_d |(u^d)_i| \geq 0.5 \\ 1 & \text{otherwise} \end{cases} \]  

(54)

where \( \sigma = \max(|\bar{u}|)\Delta t/h \) is the global CFL number.

The stencil was allowed to vary to ensure good performance over the spectrum of possible CFL numbers. It was found that at low CFL numbers, \( s_i = 2 \) lead to excess diffusion near discontinuities. Alternatively at high CFL numbers, \( s_i = 1 \) lead to excess diffusion. Allowing the stencil size to vary mitigated these issues.

After the stencil is determined, four bounds are computed: (1) max based on \( \langle q \rangle^n \), (2) min based on \( \langle q \rangle^n \), (3) max based on \( \langle q \rangle^{td} \), and (4) min based on \( \langle q \rangle^{td} \).

\[
(q_{\max})^n_i = \max(B_i(\langle q \rangle^n)) \tag{55}
\]
\[
(q_{\min})^n_i = \min(B_i(\langle q \rangle^n)) \tag{56}
\]
\[
(q_{\max})^{td}_i = \max(B_i(\langle q \rangle^{td})) \tag{57}
\]
\[
(q_{\min})^{td}_i = \min(B_i(\langle q \rangle^{td})) \tag{58}
\]

Then select the upper and lower bound of the two estimates.

\[
(q_{\max})_i = \max((q_{\max})^n_i, (q_{\max})^{td}_i) \tag{59}
\]
\[
(q_{\min})_i = \min((q_{\min})^n_i, (q_{\min})^{td}_i) \tag{60}
\]

4.2.1. Accurate Solution Bounds at Smooth Extremum. For the vast majority of cells within the domain, the previous bound computation is sufficiently accurate. However, computing bounds at extremum is more complicated. Ideally the bounds need to keep the solution monotonic and positive, but the bounds should also not “clip” the solution. There are few different methods for avoiding clipping, and we use a geometric construction that is only applied at smoothly varying extrema. It is based on the ideas in [5].

The smooth extremum criteria in 1D is:
\[(ext_d)_i = \min[ (dq)_i \cdot (dq)_{i+e^d} \cdot (dq)_{i-e^d} \cdot (dq)_{i+2e^d} ] \leq 0.0 \quad \& \& \]  
\[1.25 \cdot (dqtot)_i < (tv)_i \]  

where the following definitions are used:

\[(dq)_i = \langle q \rangle^td_i - \langle q \rangle^td_{i-e^d} \]  
\[(dqtot)_i = |\langle q \rangle^td_{i+2e^d} - \langle q \rangle^td_{i-2e^d}| \]  
\[(tv)_i = |(dq)_{i+2e^d}| + |(dq)_{i+e^d}| + |(dq)_i| + |(dq)_{i-e^d}| \]  

This criteria has two parts. First, check for a sign change in the first derivative. The sign change will indicate either an extremum or a discontinuity in the solution. Second, ensure that the solution locally is not a perturbation of a discontinuity.

For a smooth multidimensional extremum either \((ext_d)_i\) must be true in all dimensions or it must be true for some \(d\) and the solution must remain constant along the dimensions in which \((ext_d)_i\) is not true. We use the following criterion to determine if the solution is constant:

\[ \max(|(q^d_{\text{max}})_i^td_i^d - \langle q \rangle^td_i^d|, |(q^d_{\text{min}})_i^td_i^d - \langle q \rangle^td_i^d|) \leq 10^{-14} \]  

where the following definitions hold:

\[(q^d_{\text{max}})_i^td_i^d = \max (\langle q \rangle^td_i^d, \langle q \rangle^td_{i-e^d}, \langle q \rangle^td_{i+e^d}) \]  
\[(q^d_{\text{min}})_i^td_i^d = \min (\langle q \rangle^td_i^d, \langle q \rangle^td_{i-e^d}, \langle q \rangle^td_{i+e^d}) \]  

Once we have determined that the solution at \(V_i\) is at a smooth extremum, we compute new values of \((q_{\text{max}})_i\) and \((q_{\text{min}})_i\). The first step is to construct a parabolic function from the local values of \(\langle q \rangle\) centered at \(x_i\):

\[q^d(x) = \left( \frac{d^2q_i^n}{2} \right) x^2 + \left( \langle q \rangle_{i+e^d} - \langle q \rangle_{i-e^d} \right) x + \langle q \rangle_i \]  

where
\[(d2q)_i^n = \langle q \rangle_{i+e^d}^n + \langle q \rangle_{i-e^d}^n - 2\langle q \rangle_i^n\]  \hspace{1cm} (69)

The location of the vertex \((x_c)\) is given by the ratio \(-b/2a\), where \(a\) and \(b\) are the quadratic and linear coefficients (Eq. 68):

\[x_c = \frac{-\langle q \rangle_{i+e^d}^n - \langle q \rangle_{i-e^d}^n}{2(d2q)_i^n}\]  \hspace{1cm} (70)

and \(-0.5 \leq x_c \leq 0.5\). Then, evaluate the quadratic at the vertex to find the extremum value as well as deconvolve to get an estimate of the point value:

\[q_{ext}^d = \left(\frac{(d2q)_i^n}{2}\right)x_c^2 + \left(\frac{(\langle q \rangle_{i+e^d}^n - \langle q \rangle_{i-e^d}^n)}{2}\right)x_c + \langle q \rangle_i - \frac{(d2q)_i^n}{24}\]  \hspace{1cm} (71)

Select the largest \((q_{ext}^d)_i\) or smallest \((q_{ext}^d)_i\) depending on the sign of the second derivative:

\[q_{ext}^d = \begin{cases} 
\max_d((q_{ext}^d)_i, (q_{max})_i) & \text{if } \text{sgn}((d2q)_i^n) \leq 0.0 \\
\min_d((q_{ext}^d)_i, (q_{min})_i) & \text{otherwise}
\end{cases}\]  \hspace{1cm} (72)

Finally, compute the appropriate extremum bound by augmenting the solution value at the previous time by a scaled difference between the extremum value and the solution value:

\[q_{max}^i = \begin{cases} 
q_i^n + \max(0.0, 2.0 |(q_{ext})_i - q_i^n|) & \text{if } \text{sgn}((d2q)_i^n) \leq 0.0 \\
(q_{max})_i & \text{otherwise}
\end{cases}\]  \hspace{1cm} (73)

\[q_{min}^i = \begin{cases} 
q_i^n + \min(0.0, 2.0 |(q_{ext})_i - q_i^n|) & \text{if } \text{sgn}((d2q)_i^n) > 0.0 \\
(q_{min})_i & \text{otherwise}
\end{cases}\]  \hspace{1cm} (74)

**4.2.2. Updating \(R_t^\pm\) at Extrema.** We flag the extrema at which the Laplacian is changing sign. We compute the \(D\) dimensional approximation to the Laplacian over a 3-point stencil:
\[
\Delta q_i = \sum_{d=1}^{D} \frac{\partial^2 q_i}{\partial x_d^2} \approx \sum_{d=1}^{D} \frac{(d2q)^n_i}{h^2}
\] (75)

If \( \Delta q \) changes sign anywhere in the 3-point vicinity of \( i \), then we flag that cell \( i \). We then update the least upper bound multiplier at the flagged cells:

\[
R^\pm_i = 0 \text{ if } i \text{ flagged}
\] (76)

### 4.3. Pre-constraining the High-Order Flux

We pre-constrain the high-order flux by modifying the high-order fluxes where they would otherwise accentuate or produce a new extremum. In practice, the value of the antidiffusive flux is edited instead of the high-order flux directly. Following [13], we set the antidiffusive flux to zero in these regions.

The baseline condition for applying the pre-constraint:

\[
\langle A^d_{i+\frac{1}{2}e^d}(\langle q\rangle_{i+e^d}^{td} - \langle q\rangle_i^{td}) \leq 0.0
\] (77)

However, this condition alone is not sufficient. This condition will occasionally be satisfied at smooth areas in the solution. We add the following conditions to make sure we only apply this condition away from smooth areas:

\[
\min \left[ (d2q)^n_{i+e^d} \cdot (d2q)_i^n \cdot (d2q)^n_{i-e^d}, (d2q)^n_{i+e^d} \cdot (d2q)^n_{i+2e^d} \right] < 0.0
\] (78)

\[
|\langle A^d_{i+\frac{1}{2}e^d} \rangle| \leq \frac{|(u^d)^{i+\frac{1}{2}e^d}| h}{2} (1 - \sigma_{i+\frac{1}{2}e^d}) \frac{|(d2q)_i + (d2q)^n_{i+e^d}|}{2}
\] (79)

where \( \sigma_{i+\frac{1}{2}e^d} = |(u^d)^{i+\frac{1}{2}e^d}| \Delta t/h \).

The first constraint above attempts to detect a discontinuity in the solution. Discontinuities are natural places for high-order fluxes and consequently antidiffusive fluxes to generate a new extremum. However, there are smooth multidimensional solutions in which the second derivative naturally changes sign. The second constraint seeks to preclude this case. The term on the right hand side of the inequality (Eq. 79) is the d-directional dissipation term, scaled by the cell size, in the modified equation analysis (Eq. 80) of CTU applied to the advection equation.
\[
\frac{\partial q}{\partial t} + \sum_d^D \left( u^d \frac{\partial q}{\partial x_d} \right) = \sum_d^D \left( \frac{u^d h}{2} (1 - \sigma_d) \frac{\partial^2 q}{\partial x_d^2} \right) + O(h^2) \quad (80)
\]

We are interested in comparing the magnitude of this dissipative term to the antidiffusive flux. While it is true that the second derivative takes on a zero value right at a discontinuity, the second derivative takes on large values immediately surrounding this point. This means that the dissipative term is large in the neighborhood of discontinuities. If the magnitude of the dissipative term is large relative to the the antidiffusive term, then we assume we are near a discontinuity and allow the pre-constraint.

5. Results

Results in one and two dimensions are presented. A total of four initial conditions were investigated. Of the four, one initial condition was smooth and the others contained a discontinuity. For the two dimensional tests, we used two different velocity fields (Eq. 81-Eq. 82): constant diagonal and solid body rotation.

\[
u = [1, 1] \quad (81)
\]
\[
u = 2\pi[y - 0.5, 0.5 - x] \quad (82)
\]

The center for the constant velocity initial condition was in the middle of the domain, whereas it was offset by 0.25 of the grid height for the solid body rotation examples

\[
x^\text{const}_c = (0.5, 0.5)
\]
\[
x^\text{solid}_c = (0.5, 0.75)
\]

5.1. Initial Conditions. The smooth initial condition was constructed as a power of cosines.

\[
q_i(t_0) = \begin{cases} 
\cos^8 \left( \frac{\pi}{2} \cdot \frac{R}{R_0} \right) & \text{if } R \leq R_0 \\
0 & \text{otherwise} \end{cases} \quad (83)
\]
\[ x_i \in [0, 1] \]
\[ R = \sqrt{(x_i - x_c)^2} \]

For this initial condition, \( R_0 = 15 \) was chosen.

Three different discontinuous initial conditions were investigated. The first was a square and is described as:

\[
q_i(t_0) = \begin{cases} 
1 & \text{if for each } D: |x_i^D - x_c^D| \leq 0.15 \\
0 & \text{otherwise}
\end{cases} \quad (84)
\]

The next is a semi-ellipse:

\[
q_i(t_0) = \begin{cases} 
\sqrt{1.0 - \left(\frac{R}{R_0}\right)^2} & \text{if } R \leq R_0 \\
0 & \text{otherwise}
\end{cases} \quad (85)
\]

In this case \( R_0 = 0.25 \).

The last test case is the classic slotted cylinder in two dimensions [13]. For this paper we used \( N = 256 \) to represent the cylinder. This is roughly twice as resolved as the original.

5.2. One-Dimensional Tests. The first requirement for the limiter method is that it reduces to the high-order scheme away from discontinuities. All of the high-order schemes achieved similar errors for smooth solutions (Figure 2). The rate of convergence for each method was 4.0, which is the expected rate due to the usage of RK4 for the time integration scheme. This implies that the limiter is not being activated in smooth regions, and the first requirement is met.

The second requirement is that the limiter method accurately represent discontinuities. To determine if this requirement is met, we analyzed the performance of the high-order schemes with and without the limiter. The first test was a comparison between the high-order schemes without the limiter (Figure 3). In the presence of a discontinuity, the centered difference solutions produced more pronounced oscillations than the upwind solutions. This result is consistent with the amplitude and phase error analysis presented earlier.
The more important comparisons are how these high-order schemes perform with the limiter on. For the square initial condition, there is very little difference in the solution. The only noticeable difference is some additional diffusion on the tailing edge of the centered difference solution. However, for the semi-ellipse, the centered difference schemes produce oscillations even in the limited solution near the discontinuity. The oscillations are bounded by FCT; however, they are clearly undesirable. The upwind schemes also produce some oscillations, but again the magnitude of the oscillations is smaller. This is motivation for selecting the upwind methods moving forward.
5.3. Two Dimensional Tests. The smooth test errors in two dimensions are reported with the limiter for both velocity fields (Figure 6). As in the one dimensional case, the high-order solution accuracy requirement is met. Each method achieves fourth order accuracy, at a minimum. Interestingly, for the solid body rotation solution, the error reduction is greater than fourth order. Since the spatial differencing error is sixth order accurate, this suggests that the time discretization error is smaller than expected even for large values of $\sigma$. Also the ninth order scheme is still convergent right at its theoretical stability limit ($\sigma \sim 0.8$).
The limiter also performs quite well at representing discontinuities in two dimensions. Various solution plots for discontinuous initial conditions are presented (Figure 7 - Figure 9). All of the two dimensional plots were generated using the ninth order scheme in space, running right at the method’s theoretical stability limit (\(\sigma = 0.8\) in 2D). The square solution under constant velocity has few, if any, ripples and is nicely bounded (Figure 7). There is some distortion of the corners, particularly at the top-left and bottom-right. The square solution under solid body rotation looks similar to the constant velocity solution and the corner issue is mitigated.

The semi-ellipse solution is likewise well resolved (Figure 8). As with the one dimensional case there are some dispersive errors on the leading edge, but they are small. The solutions under both velocity fields are accurately bounded. The semi-ellipse solution under solid body rotation was centered at \(\mathbf{x}_c^{\text{solid}} = (1.0, 1.5)\) to keep the edge...
of the condition away from the domain boundary. The domain was also expanded to $x_i \in [0, 2]$.

The final test was the slotted cylinder (Figure 9). The limiter method keeps the solution bounded and resolves the fronts quite nicely. At lower grid resolutions the slot can fill in and the bounds may not be enforced. But as the grid is refined, both of these issues are resolved.

6. Conclusions

We presented a new flux limiter based upon FCT that retains high-order accuracy for smooth solutions and captures fronts well. Our algorithm presented here uses
CTU for low-order fluxes, upwind schemes for high-order fluxes, and RK4 for time integration. Our additions to the previous FCT method included a new computation for the extrema and an expanded pre-constraint on the high-order fluxes. This new pre-constraint is more restrictive than the original one, and seeks to only constrain the fluxes near discontinuities in the solution.

Extensions for this work are applying the limiter to systems of hyperbolic conservation laws, possibly developing new high-order upwind methods with corner-coupling, and further improving the pre-constraint on the high-order fluxes. Applying the limiter to hyperbolic systems is the most straightforward extension of this work. It would be relatively simple and informative to apply this limiter in a gas dynamics solver. High-order, corner-coupled upwind methods for use with general multi-stage time
integrators could remove the dimensional dependence of the stability. However, no upwind methods of this nature currently exists. The pre-constraint on the high-order fluxes is another area where additional study could pay off. In this work we found that the pre-constraint affects a delicate balance between effectively representing discontinuities and retaining high-order accuracy in smooth areas. It was relatively simple to get one or the other. Ensuring both required testing many versions of the pre-constraint. Different versions of the pre-constraint have been proposed [6], but they did not produce good results with the algorithm presented here.

Other versions of FCT [9] have implemented high-order centered difference schemes with hyper-diffusive fluxes. These schemes, like the high-order upwind schemes, seek to add diffusion to the large wave-number solution components. These centered difference schemes were not examined in this study, but they do offer a possible alternative to the upwind ones used here.

References


