Selective monotonicity preservation in scalar advection

Peter N. Blossey ∗ Dale R. Durran

Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640

Abstract

An efficient method for scalar advection is developed that selectively preserves monotonicity. Monotonicity preservation is applied only where the scalar field is likely to contain discontinuities as indicated by significant grid-cell-to-grid-cell variations in a smoothness measure conceptually similar to that used in weighted essentially non-oscillatory (WENO) methods. In smooth regions, the numerical diffusion associated with monotonicity-preserving methods is avoided. The resulting method, while not globally monotonicity preserving, allows the full accuracy of the underlying advection scheme to be achieved in smooth regions. The violations of monotonicity that do occur are generally very small, as seen in the tests presented here. Strict positivity preservation may be effectively and efficiently obtained through an additional flux correction step.

The underlying advection scheme used to test this methodology is a variant of the piecewise parabolic method (PPM) that may be applied to multi-dimensional problems using density-corrected dimensional splitting and permits stable semi-Lagrangian integrations using CFL numbers larger than one. Two methods for monotonicity preservation are used here: flux correction and modification of the underlying polynomial reconstruction.

Key words: scalar advection, flux corrected transport, WENO, PPM

PACS: 47.11.-j, 47.11.Df

1 Introduction

The accurate simulation of the effectively inviscid transport of scalar fields such as temperature or the concentrations of chemical species plays an important role in the modeling of many physical systems. The inviscid transport of a conservative scalar may be described by the conservation law

\[
\frac{\partial \rho \psi}{\partial t} + \nabla \cdot (\rho u \psi) = 0
\]  

(1)

where \( \rho \), \( \psi \) and \( u \) are the density, scalar concentration and vector velocity fields, respectively. For simplicity, we ignore any source and sink terms related to chemical reactions or changes in phase that might appear in more general applications and assume those processes are integrated in a separate fractional step. Note, however, that such processes may give rise to sharp gradients in \( \psi \) that must be handled accurately in the approximation of (1).

It has long been recognized that numerical approximations to (1) often require a method that is monotonicity preserving in the vicinity of discontinuities and poorly resolved gradients, but switches to a higher-order, less diffusive formulation where the solution is smooth. One of the first techniques proposed for constructing such a scheme is flux-corrected transport (FCT) [1]. The general FCT strategy [2] is to compute approximations to the true fluxes \( \rho u \psi \) using both a low-order monotone method and a higher-order scheme. In regions where there is no possibility of generating new maxima and minima, the solution is updated by computing the divergence of the high-order fluxes. In other regions, the high-order fluxes are “corrected” to prevent the development of new maxima and minima. One systematic problem with flux correction algorithms is that they erroneously damp smooth extrema. As the phase of a wave shifts and its peak translates between a pair of grid points over an interval longer than a single time step, there should be time steps in which the maximum grid-point value of the function increases (for example, if the peak of a continuous function shifts from a position midway between two grid points to exactly coincide with one of the grid points). Flux correction algorithms do not distinguish between increases in the maximum of the grid-point values produced by the translation of a wave crest and spurious increases created by non-monotone schemes in the vicinity of poorly resolved gradients. Furthermore, in many applications, only a small percentage of the grid cells in the entire domain actually require flux correction or flux limiting, and throughout the remainder of the domain, significant computational effort is expended to apply a flux-limiting algorithm that is either unnecessary or detrimental.

Similar problems with the damping of smooth extrema are also encountered with a related class of flux-limiter and slope-limiter methods in which the magnitude of the high-order flux is “limited” to keep the solution total variation diminishing (TVD) [3–5]. Higher-order generalizations of the slope-limiter approach, in which piecewise parabolas are fitted to finite-volume averages of \( \psi \) to estimate fluxes at the cell interfaces and these parabolas are subsequently modified to ensure monotonicity preservation (the PPM method [6,7]), con-
time to suffer the same type of spurious damping.

The damping of smooth well-resolved extrema may be avoided using WENO methods [8,9]. WENO methods estimate the fluxes in (1) using polynomial approximations to $\rho \psi$ of identical order spanning different sets of grid cells and compute the actual flux as a weighted average of the fluxes given by each of these polynomial approximations. The weights are determined by the local smoothness of the polynomial approximation over each stencil such that those regions which include sharp, poorly resolved gradients receive negligible weighting. In smooth regions, the weights are computed to yield the highest possible order of accuracy.

WENO methods have been used with considerable success in a wide variety of applications, including those considered here. Nevertheless in our tests, solutions to (1) obtained using a straight-forward implementation of the 5th-order WENO approximation [9] required much more computational time than an FCT approximation of roughly similar accuracy. WENO methods also suffer from a tendency for the weights of the various stencils to revert rather slowly to the underlying high-order method as the numerical resolution of a smooth function improves [10].

Here we propose a hybrid method in which monotonicity preservation is enforced only in those regions where simple WENO-motivated smoothness criteria indicate a poorly resolved steep gradient. As will be demonstrated by several test problems in section 4, two different implementations of this approach using the PPM for the underlying advection scheme preserve smooth extrema while minimizing spurious overshoots and undershoots better than the conventional fifth-order WENO method, while requiring far less computation time. These hybrid methods can also be used in semi-Lagrangian formulations that enlarge the numerical domain of dependence to permit stable integrations at CFL numbers larger than one.

A variety of additional strategies have been proposed for detecting “troubled cells” where limiters should be applied to prevent spurious oscillations near a discontinuity while attempting to preserve smooth extrema [11]. Here we compare our WENO-motivated criteria to the troubled-cell identification methodology proposed by Zerroukat et al., [12], which is a refinement of criteria proposed earlier by Sun et al [13] and Nair et al. [14].

In the following, the underlying advection scheme and approaches for monotonicity preservation, which include flux correction and modification of the underlying polynomial reconstruction, are described in detail in section 2. The WENO-motivated criteria for selective monotonicity preservation is presented in section 3. One and two-dimensional test problems comparing the various approaches are analyzed in section 4, and section 5 contains the conclusions.

2 Formation of advection scheme

The underlying numerical method for advection is a forward-in-time method presented by Skamarock [15] that implements dimension-splitting in an effective manner and can be extended to CFL numbers larger than one.

2.1 The Eulerian algorithm in one dimension

The numerical approximation to (1) in one dimension is expressed in the conservation form

$$\frac{(\rho \phi)^{n+1}_i - (\rho \phi)^n_i}{\Delta t} + \frac{f_{i+\frac{1}{2}}(\phi^n) - f_{i-\frac{1}{2}}(\phi^n)}{\Delta x} = 0, \quad (2)$$

where $(\rho \phi)^n_i$ is an estimate of the mass of the scalar $\rho \psi$ at time $t_n \equiv n\Delta t$ integrated over the cell centered at $x_i \equiv i\Delta x$, and $f_{i+\frac{1}{2}}(\phi^n)$ is an estimate of the mass flux of the scalar $\rho \psi$ through the cell interface at $x_{i+\frac{1}{2}}$ over the period $t_n \leq t \leq t_{n+1}$. The velocities $u_{i+\frac{1}{2}}$ are defined at the cell interfaces. In the present work, the fluxes $f_{i+\frac{1}{2}}(\phi^n)$ are defined using the piecewise parabolic method (PPM) of Colella and Woodward [6]. In this approach, approximate scalar values at the cell faces are computed from the cell averages in the four surrounding cells:

$$\hat{\phi}_{i+\frac{1}{2}} = \left[7(\phi^n_i + \phi^n_{i+1}) - (\phi^n_{i-1} + \phi^n_{i+2})\right]/12 \quad (3)$$

The flux may be computed in one of two ways. In the FCT formulation, no information about the polynomial reconstruction is required (e.g. whether the reconstruction has an extrema), and the flux is given by [15]:

$$f^\text{ppm}_{i+\frac{1}{2}}(\phi^n) = (\rho u)_{i+\frac{1}{2}}[\phi^n_i + (1 - C)(\hat{\phi}_{i+\frac{1}{2}} - \phi^n_i) - \mathcal{C}(1 - C)(\hat{\phi}_{i-\frac{1}{2}} - 2\phi^n_i + \hat{\phi}_{i-\frac{1}{2}})], \quad (4)$$

for $(\rho u)_{i+\frac{1}{2}} > 0$. Here $C = (\rho u)_{i+\frac{1}{2}} \Delta t/(\rho \Delta x)$ is the Courant number. The formulas for negative $(\rho u)_{i+\frac{1}{2}}$ are given in the appendix. In (4), and also below in (7), $(\rho u)_{i+\frac{1}{2}}$ is an estimate of the total mass flux through the same cell interface over the time $t^n \leq t \leq t^{n+1}$. If density is being predicted through the integration of the continuity equation, $(\rho u)_{i+\frac{1}{2}}$ is the value used in that calculation. If the flow is non-divergent and the density is
a uniform $\rho_0$, the density can be divided out of the problem and $u_{i+1/2}$ can be estimated by interpolating or extrapolating the velocity field to time $t^{n+1/2}$.

If the underlying parabola is to be modified to preserve monotonicity, the scalar distribution in the cell upwind of $x_{i+1/2}$ is reconstructed for $(\rho u)_{i+1/2} > 0$ as:

$$\phi(\xi) = a_0 + a_1 \xi + a_2 \xi^2,$$

where $\xi = (x_{i+1/2} - x)/(x_{i+1/2} - x_{i-1/2})$ and

$$a_0 = \phi_{\xi_0}, \quad a_1 = -4\phi_{\xi_0} - 2\phi_{\xi_1} + 6\bar{\phi}, \quad a_2 = 3\phi_{\xi_0} + 3\phi_{\xi_1} - 6\bar{\phi},$$

with $\phi_{\xi_0} = \hat{\phi}_{i+1/2}$, $\phi_{\xi_1} = \hat{\phi}_{i-1/2}$, and $\bar{\phi} = \bar{\phi}_i$. Here, the coordinate $\xi$ increases (with decreasing $x$) from zero at $x_{i+1/2}$ to one at $x_{i-1/2}$, and $\bar{\phi}$ assumes values of $\phi_{\xi_0}$ at $\xi = 0$ and $\phi_{\xi_1}$ at $\xi = 1$. The cell-average scalar value between $\xi = 0$ and $\xi = 1$ is $\phi$. This definition of $\xi$ places the origin at the face where the flux will be computed. The flux at $x_{i+1/2}$ is:

$$f^\text{ppm}_{i+1/2} (\phi^n) = (\rho u)_{i+1/2} \left( \frac{1}{|C|} \int_{-1}^{1} \phi(\xi) d\xi \right) = (\rho u)_{i+1/2} \left( a_0 + a_1 |C|/2 + a_2 |C|^2/3 \right).$$

Note that (7) applies regardless of the sign of $u_{i+1/2}$.

The derivation of the polynomial reconstruction and computation of the flux presented here loosely follows that in [16]. The original method of Colella and Woodward [6] included modifications of the parabola to preserve monotonicity; a method that modifies $a_0$, $a_1$ and $a_2$ to make the reconstruction monotonic is discussed in section 2.4. Polynomial reconstructions of different order may also be employed, e.g. the piecewise cubic method of [12], and different techniques may be used to compute $\phi_{i+1/2}$, as in [16].

2.2 Extension two higher dimensions

The preceding is generalized to multiple dimensions using mass-corrected Strang splitting [17]. If the continuity equation is being integrated as part of a full dynamical model such that

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n + \Delta t \left( \frac{(\rho u)_{i-1/2,j} - (\rho u)_{i+1/2,j}}{\Delta x} + \frac{(\rho v)_{i,j-1/2} - (\rho v)_{i,j+1/2}}{\Delta y} \right)$$

the mass flux $(\rho u)$ through each interface must be saved for the scalar advection algorithm.\(^1\) The mass continuity equation is then reintegrated over each split-step using these fluxes to allow a correct diagnosis of the scalar concentration at the end of each split-step.

In two dimensions, the method is:

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n + \Delta t \left[ (\rho u)_{i-1/2,j} - (\rho u)_{i+1/2,j} \right]/\Delta x$$

$$\left( \rho \phi \right)_{i,j}^{n+1} = \left( \rho \phi \right)_{i,j}^n + \Delta t \left[ f_{i,j}^{i-1/2,j} (\phi^n) - f_{i,j}^{i+1/2,j} (\phi^n) \right]/\Delta x$$

$$\phi' = \left( \rho \phi \right)' / (\rho', \phi)$$

$$\rho_{i,j}^{n+1} = \rho_{i,j}^n + \Delta t \left[ (\rho v)_{i,j-1/2} - (\rho v)_{i,j+1/2} \right]/\Delta y$$

$$\left( \rho \phi \right)_{i,j}^{n+1} = \left( \rho \phi \right)_{i,j}^n + \Delta t \left[ g_{i,j}^{i,j-1/2} (\phi') \phi_{i,j}^{n+1} = \left( \rho \phi \right)^{n+1} \right]/\Delta y$$

Here, $v_{i,j+1/2}$ and $g_{i,j+1/2}$ are the velocity and the scalar mass flux through the “north” face of grid cell $(i, j)$. Note that even if the flow is non-divergent and the density constant throughout the domain, it is still necessary to compute $\rho_{i,j}^{n+1}$ in the first step of (9), because the velocity gradient $\partial u/\partial x$ is not generally zero. Also, the Courant number in the second split step is computed $C = (\rho v)_{i,j+1/2} \Delta t / (\rho' \Delta y)$. Strang splitting, in which the $x$-direction integration precedes the $y$-direction integration on the odd time steps and follows it on the even steps, is used to preserve second-order accuracy in time.

2.3 Flux form semi-Lagrangian advection

The individual split steps may be integrated using a semi-Lagrangian approximation to the flux divergence that extends the numerical domain of dependence beyond the adjacent upstream grid cell and allows stable computations with CFL numbers greater than unity [18,19]. The following formulae are for positive velocities, the corresponding expressions for negative velocities are given in the Appendix.

As a first step, the integer shift $s$ and fractional Courant number $\tilde{C}$ for each cell interface are chosen to satisfy

$$\tilde{C} \rho_{i-s}^{n} + \sum_{k=1}^{s} \rho_{i-k+1}^{n} = \frac{\Delta t}{\Delta x} (\rho u)_{i+1/2} \quad \text{if} \quad (\rho u)_{i+1/2} \geq 0,$$

\(^1\) If the scalar advection is integrated with a larger time step than that used to integrate the more active dynamical fields the mass fluxes would be summed over a complete cycle of fast-dynamics sub-steps.
subject to the constraint that $0 \leq \tilde{C} \leq 1$. If $\rho$ is uniform and is eliminated from (10), the right side is just the Courant number $C$, and $s$ is the integer part of $C$. Otherwise, the right side is proportional to the mass flux used in (8) to determine $\rho^{n+1}$, and $s$ is determined through trial and error by systematically increasing the number of terms in the summation, beginning at zero.

Let $\bar{u} = \tilde{C} \Delta x / \Delta t$. The scalar mass flux at $x_{i+1/2}$ is computed as

$$f_{i+1/2} = \frac{1}{\Delta t} \sum_{k=1}^{s} (\rho^n \phi^n \Delta x)_{i-k+1} + \bar{f}_{i-s+1/2},$$

where $\bar{f}_{i-s+1/2}$, the scalar flux arising from the non-integer part of the Courant number, is computed using either (4) in the FCT formulation, or (7) in the modified parabola approach, except that in both formulae $(\rho u)_{i+1/2}$ is replaced by $\rho^n \bar{u}$ and $C$ is replaced by $\tilde{C}$.

### 2.4 Approaches for Monotonicity and Positivity Preservation

The behavior of selective monotonicity preservation will be investigated in the context of two different approaches to monotonicity preservation: flux corrected transport and polynomial modification. The FCT approach modifies the flux directly, while the latter modifies the underlying polynomial reconstruction from which the flux is computed. The mechanics of these two approaches are described below.

#### 2.4.1 Flux Corrected Transport (FCT)

As in the standard FCT approach, the flux is composed of a monotonicity-preserving upwind flux

$$f_{i+1/2}^\text{up}(\phi^n) = \begin{cases} (\rho u)_{i+1/2} \phi^n_{i+1}, & \text{if } (\rho u)_{i+1/2} > 0; \\ (\rho u)_{i+1/2} \phi^n_{i+1}, & \text{otherwise}, \end{cases}$$

and a “corrected” higher-order increment such that

$$f_{i+1/2} = f_{i+1/2}^\text{up} + r_{i+1/2} f_{i+1/2}^{\text{cor}},$$

where $0 \leq r_{i+1/2} \leq 1$, and $f_{i+1/2}^{\text{cor}} = f_{i+1/2}^{\text{ppm}} - f_{i+1/2}^\text{up}$ is the difference between the flux from the higher-order PPM method and the upwind flux.

As a preliminary step to evaluating $r_{i+1/2}$, an approximate “transported and diffused” solution is evaluated using the upwind flux

$$\frac{(\rho \phi)_i}{\phi_i} = \frac{(\rho \phi)_i^n + \Delta t}{\Delta x} (f_{i+1/2}^\text{up} + f_{i+1/2}^\text{up}) - (f_{i+1/2}^\text{up} - f_{i+1/2}^\text{up})$$

the sum of the correction fluxes directed out of the cell centered at $x_i$ is computed as $f_{i+1/2}^\text{out} = \max \{f_{i+1/2}^\text{cor}, 0\} - \min \{f_{i+1/2}^\text{cor}, 0\}$, and the sum of the fluxes directed into the cell at $x_{i+1}$ is calculated as $f_{i+1}^n = \max \{f_{i+1}^\text{cor}, 0\} - \min \{f_{i+1}^\text{cor}, 0\}$.

Finally let

$$\phi_{i+1}^{\text{max}, \text{min}} = \max(\phi_{i-1}^{n}, \phi_{i-2}^{n}, \phi_{i-s+1}^{n}, \phi_{i-1}^{\text{td}}, \phi_{i}^{\text{td}}, \phi_{i+1}^{\text{td}}),$$

where $s$ is zero in Eulerian implementations, and is the integer shift defined by (10) or (A.2) in semi-Lagrangian integrations.

Assuming $f_{i+1/2}^\text{cor} > 0$, the correction factor for monotonicity preservation is defined following [2,15] as

$$r_{i+1/2}^{\text{mon}} = \max \left[ 0, \min \left( 1, \left[ \frac{((\rho \phi)_i^{\text{td}} - \tilde{\rho} \phi_{i+1}^{\text{min}}) \Delta x}{\Delta t f_{i+1}^\text{out} + \epsilon_2} - \frac{\tilde{\rho}_{i+1}^{\text{max}} (\rho \phi)_i^{\text{td}} - (\rho \phi)_i^{\text{td}} \Delta x}{\Delta t f_{i+1}^\text{cor} + \epsilon_2} \right] \right) \right]$$

where $\epsilon_2$ is a small parameter chosen to avoid division by zero and $\tilde{\rho}$ is chosen to avoid division by zero and is the density as updated in the current step. In the case of no splitting, or the final step of the Strang split method (9), $\tilde{\rho} = \rho^{n+1}$; in first fractional step of (9), $\tilde{\rho} = \rho^{n+1}$.

In the final step, the cell averages are updated to time $t^{n+1}$ using

$$(\rho \phi)_i^{n+1} = (\rho \phi)_i^{\text{td}} + \frac{\Delta t}{\Delta x} (f_{i+1/2}^{\text{mon} f_{i+1/2}^{\text{cor}}} - f_{i+1/2}^{\text{mon} f_{i+1/2}^{\text{cor}}}).$$

#### 2.4.2 Polynomial Modification (PMod)

The enforcement of monotonicity in polynomial reconstructions as a long history in slope-limiter methods [5] and in the piecewise parabolic method [6] and its successors, e.g. [18]. In the present work, a monotonicity-preserving reconstruction is ensured as follows:

(1) The estimates of the scalar value at the cell faces are required to lie within the range of values in the neighboring cells, by defining

Expressions for $f_{i+1/2}^\text{cor} < 0$ appear in the Appendix.
\[
\dot{\phi}_{i+1/2}^{\text{mon}} = \min(\phi_{i+1/2}^{\text{max}}, \max(\phi_{i+1/2}^{\text{min}}, \dot{\phi}_{i+1/2}))
\] (18)

where \(\phi_{i+1/2}^{\text{max,min}} = \max, \min(\phi_i, \phi_{i+1})\). These monotonic estimates replace the corresponding non-monotonic estimates of \(\phi\) at the cell interfaces such that \(\phi_{i+1/2}^{\text{mon}} = \phi_{i+1/2}^{\text{mon}}\) and \(\phi_{i+1/2} = \phi_{i+1/2}^{\text{mon}}\) for \(\rho u_{i+1} > 0\), which in turn modifies the coefficients of the reconstructed parabola according to (6).

(2) If the resulting parabolic has an extrema in the interior of the grid cell (the condition for which is \(\phi'(\xi) = 0\) for \(0 < \xi < 1\), or equivalently \(0 < -a_1/(2a_2) < 1\), then the polynomial reconstruction is modified as follows:

(a) If \((\bar{\phi} - \phi_{\xi_1})(\bar{\phi} - \phi_{\xi_0}) > 0\), the parabola is replaced by the piecewise constant reconstruction \(\phi(\xi) = \bar{\phi}\).

(b) Otherwise, a new parabola is constructed that moves the extrema to either \(\xi = 0\) or \(\xi = 1\). If \(|\bar{\phi} - \phi_{\xi_1}| < |\bar{\phi} - \phi_{\xi_0}|\), this is accomplished by replacing (6) with

\[
(a_0, a_1, a_2) = (-2\phi_{\xi_1} + 3\bar{\phi}, 6\phi_{\xi_1} - 6\bar{\phi}, -3\phi_{\xi_1} + 3\bar{\phi}).
\] (19)

Alternatively, if \(|\bar{\phi} - \phi_{\xi_1}| > |\bar{\phi} - \phi_{\xi_0}|\), then

\[
(a_0, a_1, a_2) = (\phi_{\xi_0}, 0, -3\phi_{\xi_0} + 3\bar{\phi}).
\] (20)

These steps ensure that the polynomial reconstruction within each grid cell is monotonic and that it does not introduce any new extrema. Except for the computation of \(\dot{\phi}_{i+1/2}^{\text{mon}}\), our approach follows that of Zerroukat et al. [16].

### 2.4.3 Positivity preservation

Regardless of the approach used for monotonicity preservation, flux correction can be employed to prevent the occurrence of negative values where even very small negative values are unacceptable, as for example, in the advective transport of chemical species. Following the approach outlined in section 2.4.1, the flux is written as the sum of an upwind component and a high-order correction as in equation (13). A correction factor sufficient to keep the scheme positive definite may be defined as

\[
r_{i+1/2} = \max\left[0, \min\left(1, \frac{(\rho\phi)'\Delta x}{\Delta tf_{\text{out}} + \epsilon_2}\right)\right].
\] (21)

Note that in contrast to (16), \(\phi_{\text{max}}^\text{min}, \phi_{\text{min}}^\text{min}\), and \(f_{\text{in}}\) are not required to evaluate \(r_{i+1/2}\), and in contrast to [20,21], \(r_{i+1/2}\) corrects only the high-order increment, not the total flux. The small parameter \(\epsilon_2\) in (16) and (21) is chosen to prevent division by zero.

### 3 Determining appropriate locations for monotonicity preservation

As noted in the introduction, smooth extrema typically experience spurious damping when monotonicity preservation is continually enforced at every point in the domain. The present approach seeks to preserve monotonicity only at those locations where it is most beneficial.

#### 3.1 Where is monotonicity preservation useful?

To appreciate where the modification of fluxes for monotonicity preservation reduces numerical error, consider a one-dimensional flow with constant wind speed on the periodic domain \(0 \leq x \leq 1\) in which the initial scalar concentration is

\[
\psi_0(x) = \begin{cases} 
0 & x \leq 1/8 \text{ or } x > 7/8 \\
\left(1 - \sin(4\pi x)/2\right) & 1/8 < x \leq 1/2 \\
1/2 & 1/2 < x \leq 7/8
\end{cases}
\] (22)

This function has discontinuities in \(\psi_0\) at \(x = 7/8\), in \(\psi_0'\) at \(x = 1/2\) and in \(\psi_0''\) at \(x = 1/8\). Also, \(\psi_0\) has a smooth extremum at \(x = 3/8\).

PPM solutions with and without monotonicity preserving flux correction (16) are compared to the exact solution in Fig. 1a, along with the third PPM solution obtained using selective flux correction, which will be described in section 3.3. The numerical solution is integrated for two time steps with a CFL number of 0.5 and then shifted back one grid cell for comparison with the initial profile. The only obvious errors in the solutions in Fig. 1a appear near the discontinuity at \(x = 7/8\).

The absolute errors, plotted in Fig. 1b, confirm that the largest errors for all methods occur in the immediate vicinity of the discontinuity in \(\psi\). Monotonicity preserving flux correction drastically reduces the error relative to the uncorrected solution in broader regions surrounding the discontinuities in \(\psi\), \(\psi'\) and \(\psi''\) where the solution is locally independent of \(x\), but such flux correction also increases the error in the immediate vicinity of the smooth maximum at \(x = 3/8\) by more than two orders of magnitude. These results suggest the PPM advection scheme can benefit from monotonicity preservation near discontinuities in the scalar field or its first two derivatives, but that it should be avoided near smooth extrema. Note that similar results are obtained when monotonicity preservation is enforced through polynomial modification, rather than flux correction.
3.2 A WENO-derived smoothness metric

Several methods have been proposed for discriminating between regions of differing smoothness in the numerical solution, including the multi-resolution analysis of polynomial interpolants [22,23], comparisons of the magnitudes of grid-point-to-grid-point jumps in the function [24,25], and weighted essentially non-oscillatory (WENO) methods [9].

WENO schemes of order $(2r - 1)$ use a linear combination of lower-order polynomial interpolants of the cell-centered fluxes, $f_i \equiv f(\phi_i)$ over $r$ adjacent stencils to yield a higher order scheme. Where the solution is very smooth, the lower order interpolants are weighted so that their linear combination will yield an order $(2r - 1)$ scheme. Near poorly resolved gradients, the weights are adjusted to reduce the contribution from those interpolants computed over stencils where the solution is not smooth. Letting $q_{i,k}(x)$ be the $k$th member of the set of $r$ interpolants for flux through the “downstream” boundary of the cell centered at $x_i$, then $\beta_{i,k}$, the standard WENO smoothness indicator for that interpolant, is defined as

$$\beta_{i,k} = 2 \sum_{m=1}^{r} \int_{x_{i-1/2}}^{x_{i+1/2}} (\Delta x)^{2m-1} \left( \frac{\partial^{m} q_{i,k}}{\partial x^{m}} \right)^{2} \, dx$$

As an example of the calculations required to evaluate one of the three $\beta$ used in a 5th-order WENO method, the most “upstream” $\beta$ is

$$\beta_{i,0} = \frac{13}{12} (f_{i-2} - 2f_{i-1} + f_{i})^2 + \frac{1}{4} (f_{i-2} - 4f_{i-1} + 3f_{i})^2.$$  

Further details are given in Jiang and Shu [9].

Being the sum of the norms of the first and second derivatives, $\beta_{i,k}$ is strongly sensitive to discontinuities in $\psi(x)$. Hill and Pullin [26] took advantage of this property by defining

$$\lambda_i = \frac{\max_{k} \beta_{i,k}}{\min_{k} \beta_{i,k} + \epsilon}$$

and switching from a tuned center-difference method to a WENO method at those $x_i$ where $\lambda_i$ exceeded a threshold value. (The small constant $\epsilon$ is set to avoid division by zero.)

3.3 Selective monotonicity preservation

We use a smoothness metric that is similar, but much simpler to compute than $\lambda_i$, and enforce monotonicity preservation where the cell-to-cell variations in this metric exceed a threshold. Motivated by (23), let $\gamma_i$ approximate

$$\gamma_i \approx \frac{1}{2} \left( (\Delta x)^2 \frac{\partial^2 \psi}{\partial x^2}(x_i) \right)^2 + \frac{1}{2} \left( 2\Delta x \frac{\partial \psi}{\partial x}(x_i) \right)^2.$$

The factors of one half and two in the preceding are strictly for computational convenience since they allow $\gamma_i$ to be efficiently calculated as

$$\gamma_i = \frac{[(\phi_{i+1} - 2\phi_i + \phi_{i-1})^2 + (\phi_{i+1} - \phi_{i-1})^2]}{2}$$

$$= (\phi_{i+1} - \phi_i)^2 + (\phi_i - \phi_{i-1})^2.$$  

$$\equiv$$
A smoothness parameter \( \lambda_{i+1/2} \) is then evaluated such that

\[
\lambda_{i+1/2} = \frac{\max_{k \in K} \gamma_k}{\min_{k \in K} \gamma_k + \epsilon},
\]

where \( K \) is the upstream weighted set of indices

\[
K = \begin{cases} 
[i - 1, i, i + 1], & \text{if } u_{i+1/2} \geq 0, \\
[i, i + 1, i + 2], & \text{otherwise},
\end{cases}
\]

and \( \epsilon \) is again a small parameter that prevents division by zero.

Monotonicity preservation, using either flux correction (section 2.4.1) or polynomial modification (section 2.4.2), is enforced only at those cell interfaces where \( \lambda_{i+1/2} \) exceeds \( \lambda_{\text{max}} \). In the FCT approach, the correction factor becomes

\[
r_{i+1/2} = \begin{cases} 
r_{\text{mon}}^{i+1/2}, & \text{if } \lambda_{i+1/2} > \lambda_{\text{max}}, \\
1, & \text{otherwise},
\end{cases}
\]

while the piecewise parabolic reconstruction used to compute the flux at \( x_{i+1/2} \) is modified to maintain monotonicity only if \( \lambda_{i+1/2} > \lambda_{\text{max}} \). In multi-dimensional problems, the selective correction of fluxes along one coordinate is based solely on the smoothness of the scalar along that coordinate.

Note that in contrast to \( \beta_{i,k} \) only one \( \gamma_i \) is computed per grid cell; (27) involves fewer algebraic operations than (24), and \( \gamma_i \) is computed using the approximate solution itself, rather than the fluxes. In addition, the stencil of grid cells required for the computation of \( \lambda_{i+1/2} \) is the same as that used to compute the PPM flux at \( x_{i+1/2} \).

After considerable experimentation we chose to set \( \lambda_{\text{max}} = 20 \), which as shown in Fig. 1c, does a good job of distinguishing regions where the function defined by (22) or its low-order derivatives are discontinuous. The role of \( \epsilon \) in the denominator of (28) can be appreciated by considering a scalar field with a uniform background \( \phi_0 \). If three consecutive background values are included in a five point stencil, the minimum value of \( \gamma_k \) will be zero, and a perturbation adjacent to the background field \( |\phi_i - \phi_0| \) exceeding \( \sqrt{\epsilon \lambda_{\text{max}}} \) will be necessary to trigger selective monotonicity preservation. We have found it effective to set \( \epsilon \) to \( 10^{-8} \) times the square of a characteristic scale for the solution \( \psi \).

Selective flux limiting, like WENO methods, comes very close to preserving strict positivity, but does allow small negatives to develop in some situations where the initial data is non-negative. To guarantee that the scheme is strictly positivity preserving, one may efficiently apply positive-definite flux correction globally as described in section 2.4.3. If selective limiting is applied within the FCT framework, enforcement of global positivity preservation is achieved by replacing (30) by

\[
r_{i+1/2} = \begin{cases} 
r_{\text{mon}}^{i+1/2}, & \text{if } \lambda_{i+1/2} > \lambda_{\text{max}}, \\
r_{\text{pos}}^{i+1/2}, & \text{otherwise},
\end{cases}
\]

When polynomial modification is used to selectively enforce monotonicity, the fluxes at each interface are corrected in a final pass using (13) and (21).

Another approach in which the local values of \( \phi_i \) are used to determine where to apply monotonicity preserving polynomial modification while attempting to preserve smooth extrema was proposed by Zerroukat et al. [12]. Building on earlier criteria proposed by Sun et al [13] and Nair et al. [14], the reconstruction of \( \phi \) at interface \( x_{i+1/2} \) is judged to be a spurious grid-scale violation of monotonicity if

\[
(\hat{\phi}_{i+1/2} - \phi_i)(\phi_{i+1} - \hat{\phi}_{i+1/2}) < 0,
\]

and at least one of the following inequalities are satisfied:

\[
\begin{align*}
(\phi_i - \phi_{i-1})(\phi_{i+2} - \phi_{i+1}) & \geq 0, \\
(\phi_i - \phi_{i-1})(\phi_{i-1} - \phi_{i-2}) & \leq 0, \\
(\phi_{i+2} - \phi_{i+1})(\phi_{i+3} - \phi_{i+2}) & \leq 0, \\
(\phi_{i+1/2} - \phi_i)(\phi_{i+1} - \phi_{i-1}) & \leq 0.
\end{align*}
\]

These inequalities recognize a discontinuity through changes in slope between consecutive (or nearby) intervals, independent of the amplitude of the changes in slope. Similar inequalities are employed to judge whether an extremum in the polynomial reconstruction within the grid cell are spurious, i.e., whether the extremum in \( \phi(\xi) \) is a subgrid-scale violation of monotonicity. In contrast, the selective monotonicity preservation method just presented recognizes discontinuities through variations in the amplitude of the cell-to-cell changes in \( \phi_i \)—although we prefer to characterize (27) as an approximation to the WENO-like smoothness metric (26) rather than simply an empirical measure of cell-to-cell variations. In the two-dimensional tests that follow, selective monotonicity preservation will be compared with a combination of the criteria (32)–(36) and further criteria that recognize subgrid-scale violations of monotonicity as in [12].
In this section, the effectiveness of selective monotonicity preservation using both flux correction and polynomial modification is evaluated by comparing it to the underlying PPM method without any monotonicity preservation, and with global monotonicity preservation. The positive-definite and semi-Lagrangian versions (CFL > 1) of the selective monotonicity preservation scheme are also evaluated.

4 Results

4.1 One-dimensional tests

In all the one-dimensional tests, $\rho = u = 1$ throughout the domain, which is periodic on the interval $0 \leq x \leq 1$. The first test considered is the long-term advection of a well-resolved sine wave. The initial profile is $\phi(x, t = 0) = \sin(2\pi x)$ and $\Delta x = 1/30$. Figure 2 displays the results in the subdomain $[0, 0.5]$ at $t = 20$ using the underlying PPM scheme without any monotonicity preservation, selective monotonicity preservation and global monotonicity preservation. Results using both the flux correction and polynomial modification approaches to selective and global monotonicity are shown. Selective monotonicity preservation gives identical results to the underlying PPM scheme in this case, since the smoothness metric $\lambda$ is less than $\lambda_{\text{max}}$ everywhere. Despite the excellent numerical resolution, global monotonicity preservation noticeably degrades the solution near the extrema and its errors in the two norm $E_2$ and the infinity norm $E_\infty$ are, respectively, five and eight times greater for global monotonic flux correction than for selective or no flux correction. Global monotonicity preservation through polynomial modification damps the extrema almost twice as much as does the FCT approach and leads to correspondingly larger $E_2$ and $E_\infty$ errors (not shown).

Although the preceding example highlights the problems generated by global monotonicity preservation near well resolved extrema, it does not provide a stringent test of the ability of the selective monotonicity preservation algorithm to distinguish between true discontinuities and extrema in coarsely resolved waves. Figure 3 compares the same methods in a case where the initial $\psi_0(x)$ is the sum of equal-amplitude $7.5\Delta x$ and $10\Delta x$ waves and the solution has translated one time around the periodic domain ($t=1$). Despite the variability in the function on scales close to the grid scale, $\lambda$ remains smaller than $\lambda_{\text{max}}$, so no corrections are applied to the fluxes in the selective scheme and it again yields results identical to the underlying PPM method. On the other hand, global monotonicity preservation damps the initial profile more vigorously, eliminating a pair of smooth extrema near $x = 0.5$ and producing roughly twice the $E_2$ and $E_\infty$ errors. The performance of each scheme on a problem with discontinuities is evaluated in Fig. 4, in which an initial top hat profile, with $\psi_0(x) = 1$ for $0.4 \leq x \leq 0.6$ and zero elsewhere, is advected five times around the periodic domain ($t=5$). When $\Delta x = 0.02$ (top row), the solutions generated by selective and global monotonicity preservation are nearly identical, and both preserve the top hat profile reasonably well. Toward the end of the integration, selective flux correction is applied at 17 of the 50 points in the numerical domain when using FCT (diamonds in Fig. 4) and 15 using polynomial modification (plus signs in Fig. 4). As expected, the uncorrected PPM scheme produces over- and undershoots. The $E_2$ and $E_\infty$ errors are similar for all methods. When the same experiment is repeated with $\Delta x$ doubled to 0.04, the poorly resolved flat top is transformed into a single peak by the weak numerical diffusion in the underlying PPM scheme (bottom row in Fig. 4). Small overshoots and significant undershoots are produced by the uncorrected scheme. Selective flux correction produces slight damping of the maximum, no perceptible undershoots, and a somewhat smaller $E_\infty$ error than the other methods. Global monotonicity correction damps the maximum much more rapidly. The fidelity of the maximum at $t=5$ obtained using selective flux correction is, however, somewhat deceptive because some overshooting occurred early in the simula-
As a final 1-D test illustrating how selective limiting alone is not able to guarantee a strictly positive solution, the initial condition was specified as
\[ \psi_0(x) = \max\left[0, \sin(6\pi x) + \sin(8\pi x)\right], \] (37)
which is the positive part of the function considered in connection with Fig. 3. (Note that again, \( \Delta x = 1/30 \).) As shown in Fig. 5, after one circuit around the periodic domain (\( t=1 \)), global monotonicity preservation keeps the solution strictly positive, but it also produces much more damping in the peaks than the unmodified PPM scheme. The peaks in the selective monotonicity preservation solution are, on the other hand, almost identical to those given by the unmodified PPM scheme. Selective flux correction also significantly reduces the magnitude of the negative undershoots generated by the unmodified PPM method, but does not eliminate them completely. These undershoots allowed by selective limiting might be deemed an acceptable tradeoff if the background concentration were nonzero. Strict positivity can nevertheless be maintained using the FCT correction presented in section 2.4.3.

When used to enforce global monotonicity preservation, polynomial modification generates considerably more damping than the flux correction approach, as evident in Figs. 2–5. Yet these two methods yield roughly similar results when they are applied selectively, although polynomial modification still produces a slight tendency toward stronger damping.

\subsection*{4.2 Two-dimensional tests}

In many problems involving scalar transport, poorly resolved steep gradients develop from initially well-resolved fields due to stretching and deformation by the velocity field. LeVeque [27] proposed a test problem to model this phenomena using a time-varying swirling flow that deforms an initially well resolved circular patch into a narrow crescent before reversing direction and returning it back to its original shape and position. Very similar tests were presented in [15,21]. One possible weakness of such tests is that the phase errors for each Fourier mode could partially or completely cancel when the velocity reverses.

Here the possibility of such cancellation is reduced by defining the reversing deformation flow as a perturbation to a steady flow in solid body rotation, so that the velocity at any point away from the origin never reverses. Our velocity field\(^3\) is defined on the unit square \( 0 \leq x,y \leq 1 \) as
\[ u(x,y,t) = u_\theta(r,t) \sin(\theta), \quad v(x,y,t) = -u_\theta(r,t) \cos(\theta), \] (38)
\[^3\text{To ensure that the discrete velocity field is divergence-free, the velocities in the tests shown here are computed from a streamfunction given in the Appendix.} \]
where
\[ r = \sqrt{(x - 0.5)^2 + (y - 0.5)^2}, \quad \theta = \tan^{-1}\left(\frac{y - 0.5}{x - 0.5}\right), \]
\[ u_\theta(r, t) = \frac{4\pi r}{T} \left[ 1 - \cos\left(\frac{2\pi t}{T}\right) \right] \left[ 1 + (4r)^6 \right] \]

The density is uniform with \( \rho = 1 \), and the initial concentration field is smooth and circularly symmetric

\[
\psi(x, y, t = 0) = \begin{cases} 
\psi_0 + \left(1 + \cos(\pi \tilde{r}) \right)^2 & \tilde{r} \leq 1 \\
\psi_0 & \tilde{r} > 1 
\end{cases}
\]  
\tag{39}

where \( \tilde{r} = 5\sqrt{(x - 0.3)^2 + (y - 0.5)^2} \). In the following, the background scalar concentration \( \psi_0 \) is zero unless otherwise specified.

The evolution of the true solution is periodic over time \( T \), as illustrated in Fig. 6. At \( t = 0 \) contour lines for the initial scalar field are circular and the tangential velocity is weaker than that for solid body rotation where \( r < 1/4 \), but where \( r > 1/4 \), it is stronger. At \( T/4 \) the scalar field has been stretched into a crescent with the highest concentrations at its tail and the perturbation deformation flow is zero, so that the total flow is instantaneously in solid body rotation. At \( T/2 \) the distribution of the scalar is again identical to \( \psi(x, y, t = 0) \) and the swirling flow is stronger than solid body rotation for \( r < 1/4 \) and weaker for \( r > 1/4 \). By \( t = 3T/4 \) the velocity is again in solid body rotation, and the scalar is distributed in a crescent with the highest concentrations at its head.

Nine numerical solutions obtained using the grid spacing \( \Delta x = \Delta y = 0.02 \) are compared at \( t = T \) in Fig. 7. The location of the correct solution, which is identical to \( \psi(x, y, t = 0) \) is shown by thick solid contours. The numerical solution never achieves significant amplitude at the lateral boundaries, so the lateral boundary conditions (zero scalar gradient at inflow, linear scalar extrapolation at outflow) have no influence on these results. The time step is chosen so that the maximum CFL number during the split steps \( \max(|u|/\Delta x, |v|/\Delta y)\Delta t \) is unity.

Undershoots of more than 10% of the initial maximum are present in the uncorrected PPM solution (Fig. 7a), although the amplitude of the initial peak is relatively well preserved. Selective monotonicity preservation (not shown) produces much smaller undershoots; the minimum value is \(-1.7 \times 10^{-4}\) using flux correction. These small undershoots are completely eliminated through the addition of positive-definite flux correction using (21), as shown in Fig. 7b,c. Selective flux correction preserves almost as much of the initial amplitude as the uncorrected PPM. On the other hand, as in the one-dimensional tests, applying monotonicity-preserving flux correction everywhere (Fig. 7d) significantly damps the peak values. Increasing the time step by a factor of four and using selective and positivity preserving flux correction with the semi-Lagrangian algorithm (10)–(11) gives a very similar result to that obtained with a maximum CFL number of unity, but the CFL\(_{\text{max}} = 4 \) solution has slightly more amplitude (Fig. 7e,f).

Two WENO solutions were also computed and plotted in Fig. 7. The unsplit solution is the 5th-order WENO method [9] integrated with the optimal third-order TVD Runge-Kutta scheme of Gottlieb and Shu [28]. Perhaps the only aspect of our implementation that is not completely standard is that only that part of the flux at \( (x_{i+1/2}, y_j) \) due to \( \phi_{i+1/2,j} \) is evaluated using WENO interpolation. The other factor in the flux, the velocity, is specified on a staggered
As evident in Fig. 7, the split WENO solution is similar to, and the unsplit WENO solution is better than, the PPM method with global monotonicity preservation, but both are inferior to the PPM with selective monotonicity preservation. The small negative undershoots in the WENO solutions could be eliminated using the same type of positivity correction applied in conjunction with selective monotonicity preservation in Figs. 7b,c.

The solution shown in Fig. 7i (hereafter ZWS PCM) is obtained by polynomial modification at cell faces flagged by the selection criteria (32)–(36). As in Zerroukhat et al [12] the underlying polynomial reconstruction is piecewise cubic (the PCM method), additional tests for subgrid-scale violation of monotonicity are applied within each cell, and the approximation is modified to ensure it is monotonic using the method described in [12,16] where violations of monotonicity are found. Positivity preservation in the approximation is guaranteed through further modifications, where necessary. In the results presented here, this scheme is implemented in a flux-based formulation with dimension-splitting as in section 2.2 rather than the unsplit semi-Lagrangian method given in [12]. The split formulation keeps the method directly comparable to the other schemes used here and is likely computationally more efficient the original unsplit method. At the resolution of $\Delta x = \Delta y = 0.02$, the ZWS scheme is similar in accuracy to the PPM solutions with selective monotonicity and global positivity preservation.

Fig. 8 shows that substantial improvement occurs when the same simulations are repeated at the higher resolution $\Delta x = \Delta y = 0.01$, although the relative performance of the PPM-based methods is unchanged. The uncorrected PPM method preserves the maximum amplitude well, but still generates strong undershoots. Global monotonicity-preserving flux correction continues to strongly damp the peak value. The combination of selective monotonicity preservation and global positivity-preserving flux correction both preserves the maximum and avoids undershoots. The lowest $E_2$ and $E_\infty$ errors are obtained using the selective, positivity-preserving scheme using flux correction with $CFL_{\max} = 4$, while the most accurate peak amplitude is found using the PCM-based ZWS scheme. At this resolution, both WENO solutions are clearly superior to the global monotonicity-preserving PPM, but still inferior to those generated using the PPM method with selective monotonicity and global positivity preservation.

When the resolution is further improved to $\Delta x = \Delta y = 0.005$, all solutions closely approximate the correct result, except for that obtained using the PPM method with global monotonicity preserving flux correction. As illustrated in Fig. 9a, global flux correction continues to produce significant errors in the shape and strength of the peak in the tracer distribution, a situation similar to that already encountered in Fig. 2. The solution obtained using selective flux correction with positivity preservation, shown in Fig. 9b, is representative.
of the other solutions. Finally, although its $E_2$ and $E_\infty$ errors are small, the uncorrected PPM solution (not shown) does generate undershoots as strong as -0.017. The tests shown in Figs. 7–9 also provide a check on the importance of errors introduced by time splitting. Although at each resolution, the unsplit

Fig. 8. As in Fig. 7 but with $\Delta x = \Delta y = 0.01$.

Fig. 9. As in Fig. 7, except that with $\Delta x = \Delta y = 0.005$ and the only solutions shown are for the PPM method using (a) global monotonicity preserving flux correction, (b) selective monotonicity preservation through polynomial modification and global positivity preserving flux correction, and (c) unsplit WENO.

WENO solution is superior to that generated by the split WENO method, it is inferior to those obtained using the time-split PPM method with selective flux correction and the time-split ZWS PCM scheme, neither of which appear to be degraded by errors introduced by time-splitting.

The locations where selective flux correction is applied in simulations of the same problem, except that a non-zero background concentration ($\psi_0 = 1$) is added, are shown in Fig. 10 for time $T/4$, a time of maximum scalar deformation. Such correction is only applied in a relatively small fraction of the total points in the domain and is focused in zones where poorly resolved gradients abut regions where the scalar concentration is approximately equal to the uniform background value. Since in this case the background concentration is non-zero, undershoots that would be eliminated by positive-definite flux correction with $\psi_0 = 0$ are revealed. Selective monotonicity preservation (here using polynomial modification) effectively minimizes these undershoots, holding them to values of order $10^{-4}$. In comparison, the undershoots generated by the uncorrected PPM method (not shown) are three orders of magnitude larger (0.122, 0.084 and 0.017 for grid spacings of 0.02, 0.01 and 0.005, respec-
A more stringent test of the treatment of discontinuities is also considered by replacing the smooth initial condition for the passive scalar (39) with the two-dimensional square wave

\[
\psi(x, y, t = 0) = \begin{cases} 
1 & \text{max}(|x - 0.3|, |y - 0.5|) \leq 0.15 \\
0 & \text{otherwise}
\end{cases}
\]

and again simulating the transport of \(\psi\) in the rotating and deforming flow (38). The approximations to \(\psi(x, y, T)\) obtained using a spatial resolution of \(\Delta x = \Delta y = 0.01\) are shown in Fig. 11. In general, the square-wave initial condition produces a tendency for some overshoots in all but the monotonicity preserving FCT solution (Fig. 38a). (Positive-definite flux correction prevents undershoots in all solutions except that generated by the WENO scheme.)

At time \(T\), the overshoots allowed by selective monotonicity preservation are negligible or small (Fig. 38b,c,e). The WENO solution (Fig. 38d) generates modest overshoots of 5.9%, and the ZWS PCM method (Fig. 38f) produces overshoots greater of 21%. The evolution of the overshoots during the simulation is, however, somewhat more complicated that might be suggested by a simple examination of the fields at \(t = T\).

Figure 12 shows the maximum value of \(\psi\) over the full spatial domain for the coarsest resolution \(\max_{x,y} \psi(x, y, t)\) with the square-wave initial condition and a rotating, deforming flow for three numerical methods.

\[
\psi(x, y, t = 0) = \begin{cases} 
1 & \text{max}(|x - 0.3|, |y - 0.5|) \leq 0.15 \\
0 & \text{otherwise}
\end{cases}
\]
generated by polynomial modification (with a maximum of 7.7%), and they persist throughout a slightly longer period (not decaying to very small values until roughly 0.7T). Despite this complicated time-evolution, the PPM methods with selective monotonicity preservation clearly do the best job of the reducing the maximum overshoot in this test—except of course, that all overshoots are eliminated using global monotonicity preservation.

A final classical two-dimensional advection test is taken from [29, p. 460], where a pointed cone and a square wave undergo one period of solid body rotation. The velocity field is as in (38) with \( u_\theta = 2\pi r \), and the initial scalar field is

\[
\psi(x, y, t = 0) = \begin{cases} 
1 & \max(|x - 0.675|, |y - 0.5|) \leq 0.125 \\
1 - \hat{r} & \hat{r} < 1 \\
0 & \text{otherwise}
\end{cases}
\]

where \( \hat{r} = [(x - 0.275)^2 + (y - 0.5)^2]^{-1/2}/0.175 \). The original formulation of the problem in [29] has been altered only to re-scale the coordinates so that \( x, y \) lie in \([0, 1]\) rather than \([-1, 1]\). Solutions computed using selective monotonicity preservation through polynomial modification with and without positive-definite flux correction, along with the result for the unsplit WENO method are shown in Fig. 13. The overall performance of each method is very good, and suggests this test is somewhat less challenging than those considered previously. These results also reinforce our previous findings that the undershoots allowed by selective monotonicity preservation (without positivity preservation) are small. In this case the magnitude of such undershoots are half that generated by the unsplit WENO method.

4.3 Convergence results

Empirically determined convergence rates for each of the methods used to simulate \( \psi \) in the rotating and deforming flow (38) with smooth initial conditions (39) may be estimated from Fig. 14, which shows the \( E_2 \) and \( E_\infty \) errors from a series of simulations in which the spatial and temporal resolutions were repeatedly doubled. Enforcing global monotonicity preservation in the PPM method leads to second-order convergence in the \( L_2 \) norm and slower-than-second-order convergence in the \( L_\infty \) norm. In contrast, third-order convergence in both the \( L_2 \) and \( L_\infty \) norms is obtained using without any flux correction and with selective monotonicity preservation implemented either through FCT or polynomial modification together with global positivity preservation. The ZWS PCM method converges at a rate slightly faster than third order once the solution is well resolved. Since Strang splitting is only second-order in time, these third-order (and higher, in the case of the ZWS PCM) convergence rates suggest that spatial truncation errors are dominating the total error in the numerical solution. Further evidence of the relatively minor impact of the time truncation error is provided by the similarity of the CFL\(_{\text{max}} = 1\) and CFL\(_{\text{max}} = 4\) errors at all but the finest resolution.

Slightly higher rates of convergence can be obtained using selective monotonicity preservation in combination with the piecewise cubic, rather than the piecewise parabolic method. As expected the benefits of higher-order are most pronounced when the solution is smooth and well resolved. Since the modest additional computation required to use selective monotonicity preservation
Table 1
Time required for the rotation and deformation test running optimized FORTRAN 95 code on a 1.3GHz Itanium 2 processor as a function of flux correction methodology, underlying numerical method, and spatial resolution. Each time is in seconds and is the average from five identical simulations.

<table>
<thead>
<tr>
<th>Monotonicity Preservation (Underlying Scheme)</th>
<th>Number of Grid Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50x50</td>
</tr>
<tr>
<td>None (PPM)</td>
<td>0.59</td>
</tr>
<tr>
<td>Global PMod (PPM)</td>
<td>0.76</td>
</tr>
<tr>
<td>Selective PMod (PPM)</td>
<td>0.90</td>
</tr>
<tr>
<td>Sel. Pos. PMod (PPM)</td>
<td>1.10</td>
</tr>
<tr>
<td>Global FCT (PPM)</td>
<td>1.14</td>
</tr>
<tr>
<td>Selective FCT (PPM)</td>
<td>1.04</td>
</tr>
<tr>
<td>Sel. Pos. FCT (PPM)</td>
<td>1.20</td>
</tr>
<tr>
<td>Mon. Pos. ZWS (PCM)</td>
<td>1.56</td>
</tr>
<tr>
<td>None (Split WENO)</td>
<td>2.79</td>
</tr>
</tbody>
</table>

As also evident in Table 1, the additional computation required to ensure that the selective monotonicity preserving schemes are strictly positive adds between 12% and 22% to the cost of the method, but still yields a relatively economical scheme. For example, selective monotonicity preservation through polynomial modification with global positivity preservation remains faster at all resolutions than the standard FCT scheme with global monotonicity preservation. It is also roughly 40% faster than the ZWS PCM method, although about half of this speed difference is due to the use of piecewise parabolas instead of piecewise cubics in our method. Recently, Zerroukat et al. [30] has suggested that parabolic splines can be more efficient and more accurate than the standard PPM approach, and the use of such splines might further reduce the computation times given in Table 1.

Finally, the split WENO method is much more costly than the other alternatives, primarily because of the three iterations required to execute the third order Runge-Kutta time integration. The unsplit WENO simulations required roughly twice the time as the split WENO method because the time step was halved in the unsplit integrations. Although the time step only needs to be reduced by a factor of $1/\sqrt{2}$ to preserve stability, even using the maximum stable time step, the unsplit WENO method would be unattractive from an efficiency standpoint. In addition, the speed of the PPM- and PCM-based methods can be dramatically increased using their semi-Lagrangian formulations. Although it is likely that the split WENO method could also be implemented in a semi-Lagrangian form, we have not pursued this since at least for this test, it would remain less accurate and much slower than the PPM method with selective monotonicity preservation enforced through polynomial modification.

5 When selective monotonicity preservation through polynomial modification with positivity preservation was tested with piecewise cubics (rather than piecewise parabolas), the method was 18% faster than the ZWS PCM for this test.
If $\phi$ and $\chi$ are the numerical approximations to two tracers, both of which satisfy (1), and at some time $t^n$ and for all $x_i$,

$$\phi^n_i = \alpha \chi^n_i + \mu, \quad (42)$$

where $\alpha$ and $\mu$ are constants, then the selective monotonicity preserving PPM method using flux correction maintains the linear correlation between $\phi$ and $\chi$ at all future time steps. This may be derived by noting that $f^{\text{mp}}(\phi) = \alpha f^{\text{mp}}(\chi) + f^{\text{mp}}(\mu)$ and $f^{\text{cor}}(\phi) = \alpha f^{\text{cor}}(\chi) + f^{\text{cor}}(\mu)$. Since $f^{\text{cor}}(\mu) = 0$, it can be seen that $r^{\text{mon}}(\phi) = r^{\text{mon}}(\chi)$. The linear correlation between $\phi$ and $\xi$ is also preserved by selective monotonicity preservation using polynomial modification.

If positivity preserving flux correction is also applied, then the correlation between $\phi$ and $\chi$ is maintained provided $\mu = 0$, but if $\mu \neq 0$, the relation $r^{\text{pos}}(\phi) = r^{\text{pos}}(\chi)$ is not guaranteed and the correlation will generally not be maintained. Linear correlations of the form (42) with $\mu \neq 0$ can nevertheless be maintained when applying positivity preserving flux correction if the same value of $r^{\text{pos}}_{i+1/2}$ is used to correct the fluxes of both chemical species at $x_{i+1/2}$ by replacing (31) with

$$r_{i+1/2} = \begin{cases} r^{\text{mon}}_{i+1/2}, & \text{if } \lambda_{i+1/2} > \lambda_{\text{max}}, \\ \min \left[ r^{\text{pos}}_{i+1/2}(\phi), r^{\text{pos}}_{i+1/2}(\chi) \right] & \text{otherwise.} \end{cases} \quad (43)$$

### 5 Conclusion

A hybrid approach for modeling scalar advection has been proposed in which monotonicity preservation is enforced solely at points where a WENO-like metric indicates the presence of discontinuities or poorly-resolved gradients. This metric is evaluated by determining the cell-to-cell variations in the sum of the squares of the normalized first and second spatial derivatives of the scalar field, and flux-correction or polynomial modification is applied only at those points where such variations exceed a threshold value. Elsewhere, including the regions surrounding smooth extrema, the high-order accuracy of the underlying numerical method is preserved.

This approach, which we call “selective monotonicity preservation,” can be used in conjunction with many different underlying advection schemes and with different methods for monotonicity preservation. Here we have demonstrated its performance using a variant of the PPM proposed by Skamarock [15], which allows efficient mass-conservative integrations in a semi-Lagrangian formulation at CFL numbers greater than unity. Both flux correction and modification of the underlying polynomial reconstruction have been tested in the selective monotonicity preservation framework and have been shown to perform similarly. Flux correction generally gave better amplitude preservation in the two-dimensional cases considered here, while polynomial modification executed slightly faster and generated somewhat smaller overshoots and undershoots.

Among the test problems considered is a non-reversing swirling shearing flow that allows an easy comparison between the numerical and exact solutions. Using the PPM scheme on this problem, empirically determined convergence rates improved from second order in the $L_2$ norm and less-than-second order in the $L_\infty$ norm when monotonicity preservation was enforced at every point, to fully third order using selective monotonicity preservation.

Although in a given time step, monotonicity is preserved in those regions where the WENO-like criteria indicates a poorly resolved sharp gradient, this method does not guarantee the global solution will be free from all overshoots and undershoots. In all the cases that we have tested, those undershoots and overshoots that are occasionally generated are very small. It is our view that some degree of over- and undershoots must be tolerated if genuinely high-order accuracy is to be maintained in the vicinity of smooth extrema.

If negative tracer concentrations must be avoided, as in the simulation of chemically reacting species, positivity-preserving flux correction may be applied at every point with only minimal additional computation. In our swirling flow tests, the computational time required to apply selective monotonicity preservation together with global positivity-preserving flux correction was comparable to that required to apply monotonicity preserving flux correction at every cell interface, and was much less than that used by standard WENO methods.

### 6 Acknowledgments

We gratefully acknowledge helpful conversations with Chris Bretherton, Mohamed Zerroukat and Bill Skamarock. This work was supported by NSF grant ATM 0506589 and ONR contract N00173-06-1-G904.

**Appendix: Expressions for negative fluxes**

*PPM flux for $u_{i+1/2} < 0$*

When $u_{i+1/2} < 0$, equation (4) becomes
\[ f^{ppm}_{i+1} = (\rho u)_{i+1/2}[\phi^n_{i+1} + (1 + C)(\dot{\phi}_{i+1/2} - \phi^n_{i+1}) + C(1 + C)(\dot{\phi}_{i+1/2} - 2\phi^n_{i+1} + \dot{\phi} \Delta z)] \]

where \( \dot{\phi} \) and \( C \) are defined below (4).

For negative velocities at \( x_{i+1/2} \), the parabolic reconstruction of the scalar distribution may be defined as in equation (5), except that the coordinate \( \xi \) increases with \( x \) in this case, as \( \xi = (x - x_{i+1/2})/(x_{i+3/2} - x_{i+1/2}) \), and that \( \phi = \phi_{i+1}, \phi_0 = \phi_{i+1/2} \) and \( \phi_{\xi} = \phi_{i+3/2} \).

Semi-Lagrangian CFL number and departure index calculation for \( u_{i+1/2} < 0 \)

Choose \( \hat{C} \) and \( s \) such that \(-1 \leq \hat{C} \leq 0 \) and

\[ \hat{C} \rho^n_{i-s+1} - \sum_{k=s}^{1} \rho^n_{i-k} = \frac{\Delta t}{\Delta x} (\rho u)_{i+1/2} \]  

(A.2)

while noting that \( s \leq 0 \) in this case. Analogous to (11), the scalar mass flux at \( x_{i+1/2} \) can then be computed as

\[ f_{i+1/2} = \frac{1}{\Delta t} \sum_{k=s}^{1} (\rho \phi^n_0 \Delta x)_{i-k} + \bar{f}_{i-s+1}. \]  

(A.3)

Flux correction for \( f^\text{cor}_{i+1/2} < 0 \)

When \( f^\text{cor}_{i+1/2} < 0 \), the renormalization factor (16) for monotonic flux correction is

\[ r_{i+1/2} = \max \left[ 0, \min \left( 1, \frac{[\rho_0 \phi^n_{\text{max}} - (\rho_0)_{i+1}] \Delta x}{\Delta f^\text{in}_i + \epsilon_2}, \frac{[\rho_0]_{i+1} - \rho_0 \phi^n_{\text{min}} \Delta x}{\Delta f^\text{out}_i + \epsilon_2} \right) \right] \]  

(A.4)

where \( \phi^n_{\text{max}}, \phi^n_{\text{min}}, f^\text{in}_i \) and \( f^\text{out}_i \) are as in section 2.1. Note that \( \epsilon_2 = 10^{-16} \) for the results shown in this paper. The renormalization factor for positivity preservation (21) in this case is

\[ r^\text{pos}_{i+1/2} = \max \left[ 0, \min \left( 1, \frac{(\rho_0)_{i+1} \Delta x}{\Delta f^\text{out}_i + \epsilon_2} \right) \right]. \]  

(A.5)

Streamfunction for two-dimensional tests

To ensure that the velocity field in the two-dimensional tests is discretely divergence-free, the velocities through cell faces are computed using finite differences from the following streamfunction

\[ \Psi(r, t) = \frac{4\pi}{T} \left\{ \frac{r^2}{2} + \cos \left( \frac{2\pi t}{T} \right) \left[ \frac{r^2}{2} + \frac{1}{96} \log(1 - 16r^2 + 256r^4) \right] - \frac{1}{48} \log(1 + 16r^2) \right. \]

\[ \left. - \frac{\sqrt{3}}{48} \arctan \left( -\frac{1 + 32r^2}{\sqrt{3}} \right) \right\} \]

evaluated at the corners of the grid cells, e.g. \((x_{i+1/2}, y_{j+1/2})\). Note that \( r \) is defined below (38). This streamfunction was integrated from the expression for \( u_0(r, t) \) using MAPLE.

References


