Hgher accuracy methods for 1D advection

Wenjun Zhao

January 25, 2019

1 Overview

This document is a discription for numerical solvers for 1D advection equation

$$s_t + [u(x)f(s)]_x = 0,$$

where u is a spatially dependent velocity. We discuss a conservative solver with finite volume method, which requires accurate estimation for flux. To get higher order estimations for flux as an average over time and space, we consider the advection for linear/quadratic polynomials as approximations for s in each cell.

The temporal integrator is described in section 2, and in section 3 linear and quadratic reconstructions are discussed, which should have second/third order for constant advection, and second order for non-constant advection. These two sections follow from the 2D BDS (Bell-Dawson-Shubin) paper [1]. To tackle the case with nonsmooth solution and reduce the oscillations introduced by numerical, in section 4 some limiters are discussed: van Leer and BDS limiter for linear profile [1]; For quadratic reconstruction, we introduce: monotone limiting on BDS limiter [1], and the original Piecewise Parabolic Method (PPM) [2], and extreme-preserving PPM [3].

We have some numerical results after implementing these methods with periodic boundary condition. In section 5, we consider the equation with constant advection. Results for advecting a sharp Gaussian peak, a semi-circle, and a square wave function will be shown and analysed quantitively by L^1 relative error, estimated convergence rate and extremum values. In section 6, for non-constant advection problem, the convergence rate is verified by method of manufactured solution.

2 Temporal integrator

2.1 Constant advection

Spatially update with finite difference gives:

$$s_j^{n+1} = s_j^n - u\frac{\Delta t}{\Delta x}[s_{j+1/2}^{n+1/2} - s_{j-1/2}^{n+1/2}],$$

where the notation s_j^{n+1} is the averaged of j-th cell, after $(n+1)$ time steps. Without loss of generality, assume the flow is moving from the left to the right ($u > 0$), then the flux is averaged in the left cell over one time step:

$$s_{j+1/2}^{n+1/2} = \frac{1}{u\Delta t}\int_{x_{j+1/2}-u\Delta t}^{x_{j+1/2}} \left[s_{xx,j}(x - x_j)^2 + s_{x,j}(x - x_j) + s_j\right] dx,$$

where the quadratic function in the integral is the reconstructed quadratic approximation in the j-th cell. It reduces to linear reconstruction when $s_{xx,j}$ is set to 0.

If $u < 0$, we will take the integral over $x_{j+1/2}$ to $x_{j+1/2} - u\Delta t$, which is the flux coming for the cell at right side. For a fixed CFL number $\sigma = u\Delta t/\Delta x$, this estimation is third order accurate with quadratic reconstruction and second order with linear reconstruction.

2.2 Spatially dependent advection

With spatially-dependent advection, the conservation law gives:

$$s_j^{n+1} = s_j^n - \frac{\Delta t}{\Delta x}[u_{j+1/2}s_{j+1/2}^{n+1/2} - u_{j-1/2}s_{j-1/2}^{n+1/2}].$$

We can rewrite the equation as:

$$s_t + us_x + su_x = 0,$$

so the influence of second term in the average flux has to be considered. First, consider u as a constant, and do the same thing as in constant advection problems, denote a predictor step as

$$(s_{j+1/2}^{n+1/2})^p = \frac{1}{u_{j+1/2}\Delta t}\int_{x_{j+1/2}-u_{j+1/2}\Delta t}^{x_{j+1/2}} \left[s_{xx,j}(x-x_j)^2 + s_{x,j}(x-x_j) + \bar{s}_j\right] dx,$$

The exact formula for this integral gives

$$(s_{j+1/2}^{n+1/2})^p = \bar{s}_j + \frac{\Delta x - u_{j+1/2}\Delta t}{2}s_{x,j} + s_{xx,j}\left[\frac{1}{4}(\Delta x)^2 - \frac{1}{2}\Delta x u_{j+1/2}\Delta t + \frac{1}{3}(u_{j+1/2}\Delta t)^2\right].$$

then consider u_x as a constant approximated by finite difference from the upwinding direction, and take the average over time (therefore with coefficient of $\Delta t/2$) gives:

$$s_{j+1/2}^{n+1/2} = (s_{j+1/2}^{n+1/2})^p - \frac{\Delta t}{2}\frac{(u_{j+1/2} - u_{j-1/2})}{\Delta x}(s_{j+1/2}^{n+1/2})^p.$$

With a spatially dependent u, this estimation is second order accurate for linear/quadratic construction.

3 Piecewise reconstruction

3.1 Linear reconstruction

First, we construct a linear representation of s at time t^n in the form of

$$p_j^l(x) = s_{x,j}(x - x_j) + \hat{s},$$

where x_j denotes the cell center of cell j. And the estimates for the edges of cell is given by

$$s_{j+1/2} = \frac{1}{12}\left[7(s_j + s_{j+1}) - (s_{j+2} + s_{j-1})\right], \tag{1}$$

which is third order accurate for smooth functions, proposed in [2] by constructing a parabola on $[x_{j-1}, x_{j+1}]$ and calculate the mean of slope in the interval.

Then, the parameter for linear term in the reconstruction is given by centered finite difference between the two corners:

$$s_{x,j} = \frac{s_{j+1/2} - s_{j-1/2}}{\Delta x} = \frac{-s_{j+2} + 8s_{j+1} - 8s_{j-1} + s_{j-2}}{12\Delta x}. \tag{2}$$

3.2 Quadratic reconstruction

Here we construct a new quadratic representation of s at time t^n in the form of

$$p_j^q(x) = s_{xx,i}(x - x_j)^2 + s_{x,j}(x - x_j) + \bar{s},$$

where x_j denotes the cell center of cell j, and we no longer use the notation $\hat{s}$, as the constant term will no longer be equal to s_j. The first order coefficient is same as linear reconstruction. As proposed in [1], eq 25a, the coefficient for quadratic term is given by half of the second order derivative at cell center i, approximately:

$$s_{xx,j} = \frac{1}{2}\frac{(-s_{j-2} + 12s_{j-1} - 22s_j + 12s_{j+1} - s_{j+2})}{8(\Delta x)^2},$$

which is exact for polynomials up to order 5, and the coefficient for second order term is hence given by the derivatives by 2, and the constant $\bar{s}$ could be calculated by matching the average over cell j:

$$\bar{s} = s_j - \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} s_{xx,j}(x - x_j)^2 dx = s_j - \frac{1}{12} s_{xx,j}(\Delta x)^2.$$

4 Limiters

The purpose of limiting process is to guaratee that the numerical method does not introduce new extrema at the edges and cause large amplitude oscillations near non-smoothness. Here we introduce some limiters on linear and quadratic terms.

4.1 Limiters for linear reconstruction

4.1.1 van Leer limiter

The idea of van Leer limiter is that, the absolute value of the coefficient can not exceed twice the value for the finite difference evaluated at left/right side only. In this report, as our linear coefficient is given by

$$s_{x,j} = \frac{1}{12\Delta x}[-s_{j+2} + 8s_{j+1} - 8s_{j-1} + s_{j-2}]$$

For example, if the derivative for both sides are positive, then:

$$s^l_{x,j} = \min\left\{\frac{-s_{j+2} + 8s_{j+1} - 8s_{j-1} + s_{j-2}}{12\Delta x}, \frac{-s_{j+2} + 8s_{j+1} - 7s_j}{6\Delta x}, \frac{-8s_{j-1} + s_{j-2} + 7s_j}{6\Delta x}\right\}. \tag{3}$$

If the finite differences have different signs, then assign it to zero.

4.1.2 BDS limiter

Here we use the BDS limiter as mentioned in [1] for the 1D case, which says that if a local extreme happens, the slope should be adjusted such that the L^1 norm between it and the original interpolation is minimized, subject to the constraint that the interpolated face values lie in the rage of adjacent cell averages, and the averaged value in the cell is equal to the original one:

$$\min_{s^l_{x,i}} |s^l_{x,i} - s_{x,i}|, \quad s.t. s_i - s^l_{x,i}\frac{\Delta x}{2} \in [\min(s_{i-1}, s_i), \max(s_{i-1}, s_i)], s_i + s^l_{x,i}\frac{\Delta x}{2} \in [\min(s_{i+1}, s_i), \max(s_{i+1}, s_i)].$$

In the code, this is done in an iterative way, which is:

- For each cell, calculate the number of face values that fall outside of the range of adjacent cells, i.e.

  $$s_i - s^l_{x,i}\frac{\Delta x}{2} \notin [\min(s_{i-1}, s_i), \max(s_{i-1}, s_i)], s_i + s^l_{x,i}\frac{\Delta x}{2} \notin [\min(s_{i+1}, s_i), \max(s_{i+1}, s_i)].$$

  This number can only be equal to 0 or 1 or 2.

- If the value is equal to 1, then just adjust that value such that the change in the absolute value of slope is minimized. For example, if $s_i - s_{x,i}\Delta x/2 < \min(s_{i-1}, s_i)$, and the unlimited slope is positive, then we just adjust the slope such that $s_i - s_{x,i}\Delta x/2 = \min(s_{i-1}, s_i)$.

- If the value is equal to 2, we adjust the slope with the same principle of minimizing L^1 norm. How to adjust the slope depends on which face value is farther from the constraint range. For example:

  If $|s_i + s_{x,i}\Delta x/2 - \max(s_{i+1}, s_i)| > |s_i - s_{x,i}\Delta x/2 - \min(s_{i-1}, s_i)|$, then we have to adjust according to the constraint on the right side, otherwise it is not enough to let the interpolation on the other side also satisfy the constraint.

4.2 Limiters for quadratic reconstruction

4.2.1 Monotone limiting

This limiter is also discussed in [1]: to guarantee the monotonicity, we do the limiting by constrain that the extrema of quadratic polynomial does not occur in the cell:

$$(p_i^q)_x = 2s_{xx,i}(x - x_i) + s_{x,i} \neq 0 \quad \forall x \in [x_{i-1/2}, x_{i+1/2}] \implies |s_{xx,i}| \leq \frac{|s_{x,i}|}{\Delta x}.$$

Therefore, it gives that if $|s_{xx,i}| > \frac{|s_{x,i}|}{\Delta x}$, then set $s_{xx,i} = sgn(s_{xx,i})\frac{|s_{x,i}|}{\Delta x}$. In the code, it is implemented in the way that: we start from the unlimited linear interpolation and apply this quadratic constraint, if new extreme is created, then start from the limited linear interpolation and constrain the quadratic terms again. If it still creates new extreme, then set it to piecewise constant.

4.2.2 PPM limiter (with extremum preservation)

This limiter is designed for preserving smooth extreme values, as the monotone one overshoot sometimes, because it is more restrictive than monotonicity preserving. In [3], the authors proposed another PPM limiter.

1. Interpolating face values: As discussed in the previous section, the interpolation of face values gives

$$s_{j+1/2} = \frac{1}{12}\left[7(s_j + s_{j+1}) - (s_{j+2} + s_{j-1})\right],$$

we first constrain it by van Leer limiter as mentioned in [3], then limit this value by using a nonlinear combination of approximations to the second derivative. If it falls outside of the constraint range, then we impose the following constraint:

$$\begin{aligned}(D^2s)_{j+1/2} &= \frac{3}{h^2}(s_j - 2s_{j+1/2} + s_{j+1}) \\ (D^2s)_{j+1/2,L} &= \frac{1}{h^2}(s_{j-1} - 2s_j + s_{j+1}) \\ (D^2s)_{j+1/2,R} &= \frac{1}{h^2}(s_j - 2s_{j+1} + s_{j+2})\end{aligned}$$

If the signs of them are all the same, we define

$$(D^2s)^l_{j+1/2} = sgn((D^2s)_{j+1/2})\min(C|(D^2s)_{j+1/2,L}|, C|(D^2s)_{j+1/2,R}|, |(D^2s)_{j+1/2}|).$$

Here $C > 1$ is a constant independent on the discretization grid size. In [3], the authors mentioned that numerical experiments have shown that the solution is not sensitive to the value C in the range of $[1.25, 5]$. If the signs are not the same, then limit it to zero.

So,

$$s_{j+1/2} = (s_j + s_{j+1})/2 - h^2(D^2s)^l_{j+1/2}/8.$$

In [3], the denominator is 3, which I think is a typo...? As this formula is exact for quadratic polynomials with denominator 8.

2. Constructing the parabolic interpolant

First, set $s_{j,+} = s_{j+1/2}$, and $s_{j,-} = s_{j-1/2}$.

(a) Case when extremum occurs:

If

$$(s_{j,+} - s_j)(s_{j,-} - s_j) \geq 0 \text{ or} (s_{j+1} - s_j)(s_{j-1} - s_j) \geq 0,$$

then it is a local extremum, then we approximate the second order derivative as

$$
\begin{aligned}
(D^2 s)_j &= \frac{-2}{h^2}(6s_j - 3(s_{j,+} + s_{j,-})) \\
(D^2 s)_{j,C} &= \frac{1}{h^2}(s_{j-1} - 2s_j + s_{j+1}) \\
(D^2 s)_{j,L} &= \frac{1}{h^2}(s_{j-2} - 2s_{j-1} + s_j) \\
(D^2 s)_{j,R} &= \frac{1}{h^2}(s_j - 2s_{j+1} + s_{j+2})
\end{aligned}
$$

Again, if their signs are all the same, then

$$
(D^2 s)_j^l = sgn((D^2 s)_{j+1/2}) \min(C|(D^2 s)_{j,L}|, C|(D^2 s)_{j,R}|, C|(D^2 s)_{j,C}|, |(D^2 s)_j|).
$$

Then we can use it as quadratic coefficient after dividing it by 2.

If (2a) does not hold, then proceed to the original PPM limiter, which is:

(b) If $(s_{j,+} - s_j)(s_{j,-} - s_j) > 0$ and (2a) does not hold, then set it piecewisely constant: $s_{j,+} = s_{j,-} = s_j$.

(c) Otherwise, if one of $|s_{j,\pm} - s_j| \geq 2|s_{j,\mp} - s_j|$, then for that choice of $\pm$ we set

$$
s_{j,\pm} = s_j - 2(s_{j,\mp} - s_j).
$$

The second and third items are the original PPM method [2], and the full extremum-preserving PPM will be refered as PPM2 later. To make the parameterization consistent with BDS, we make the translation between $s_{j,\pm}$ and $s_{xx,j}$, $s_{x,j}$ and $s_{0,j}$, which is exact for quadratic polynomials:

$$
s_{xx,j} = \frac{3(s_{j,+} - s_{j,-}) - 6s_j}{(\Delta x)^2}, \quad s_{x,j} = \frac{s_{j,+} - s_{j,-}}{\Delta x}, \quad s_{0,j} = s_j - \frac{s_{xx,j}}{12}(\Delta x)^2.
$$

Theoretically, the BDS limiter is providing a hard constraint on the interpolated face values, while the van Leer limiter, of PPM limiters only constrain (in a soft way, for example, by a loose bound with constant C in PPM) on the amplitude of derivative. So one could expect that BDS will overshoot more than van Leer in linear interpolation. For quadratic reconstruction, the extremum-preserving PPM breaks the monotonicity constraint, which is too strong and often not necessary, so it overshoots the least. This agrees with numerical experiments with constant advection in this report, and also with the experiments in BDS paper [1] in 2D cases and spatially dependent advection problems. However, it has its advantage for more difficult problems when PPM/PPM2 (Table 13 in [1]) and van Leer (see the square wave case with large CFL in this report) could fail.

5 Numerical examples with constant advection

We test the algorithm for a system with over an interval $[0, 1]$ with $N = 32, 64, 128, 256, 512$ equi-spaced points and periodic boundary condition. The initial conditions are given by:

- Gaussian: $s(x, 0) = \exp(-256(x - \frac{1}{2})^2)$–smooth.
- Semi-circle: $s(x, 0) = \left(\max\left(\frac{1}{16} - (x - \frac{1}{2})^2\right), 0\right)^{\frac{1}{2}}$–continuous but non-smooth.
- Square wave: $s(x, 0) = \mathcal{X}_{|x-\frac{1}{2}|\leq\frac{1}{4}}$–discontinuous.

The advection equation is given with a constant coefficient:

$$
s_t(x, t) + us_x(x, t) = 0, \quad x \in [0, 1], s(0, t) = s(1, t),
$$

where $u = 1$, so we can have the analytical solution $s_{exact}(x, t) = s(mod(x - t, 1), 0)$. The errors are quantified by the relative norm at time $t = 10$:

$$
\epsilon_h(t) = \frac{||s_h - \bar{s}_{exact}||_1}{||\bar{s}_{exact}||_1},
$$

where the $\bar{s}_{exact}$ is the cell-average of analytical solution.
We will estimate the convergence rate empirically with the following formula:

$$ratio(h) = \log(\frac{\epsilon_{2h}}{\epsilon_h})/\log(2).$$

The methods we will be testing is:

- Linear reconstruction: no limiting/van Leer limiter/BDS limiter [1] ;
- Quadratic reconstruction: no limiting/BDS+monotone limiter [1] /PPM [2] /PPM2(extremum-preserving) [3]. The constant in constraint is chosen as the same value in [2]: $C = 1.25$.

For a comparison, with same initial condition/advection, this ratio for these problems in [3] with PPM is respectively approaching 3 for Gaussian problem, 1.2 for semi-circle and 0.8 for square wave problem, comparable with the results below.

When initializing the numerical solver, we set the initial value to be a 4-th order approximation of average over each cell, as mentioned in [3]:

$$s_j^0 = s(j\Delta x, 0) + \frac{1}{24}(s((j-1)\Delta x, 0) - 2s(j\Delta x, 0) + s((j+1)\Delta x, 0)),$$

so it does not influence the result of empirical order estimation, as theoretically all of these methods are second/third order accurate for smooth problems .

5.1 Gaussian initial condition

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	error	1.2948	0.8763	0.3759	0.1045	0.0266
	ratio	–	0.5633	1.2209	1.8464	1.9734
linear, VL	error	0.7656	0.3998	0.1766	0.0897	0.0302
	ratio	–	0.9373	1.1790	0.9774	1.5698
linear, BDS	error	0.8758	0.4785	0.1991	0.0969	0.0306
	ratio	–	0.8720	1.2647	1.0392	1.6625
quadratic, unlimit	error	0.5518	0.1670	0.0224	0.0025	0.0003
	ratio	–	1.7247	2.9006	3.1652	3.0686
quadratic, BDS+monotone	error	0.8448	0.3896	0.1321	0.0328	0.0076
	ratio	–	1.1166	1.5605	2.0116	2.1140
quadratic, PPM	error	0.6278	0.2221	0.0655	0.0106	0.0019
	ratio	–	1.4988	1.7616	2.6322	2.4736
quadratic, PPM2	error	0.5591	0.1637	0.0344	0.0038	0.0005
	ratio	–	1.7722	2.2523	3.1897	2.8279

Table 1: Relative L^1 error and empirical order at $t = 10$ with CFL 0.2 for constant advection and Gaussian peak initial condition with different reconstruction.

- This is a smooth problem, and the relative L^1 error and its ratio (unlimited) indicates second order for linear reconstruction, third order for quadratic reconstruction as expected.
- In this case, the limiter only overshoots, so the linear reconstruction reduces to a combination of piecewise constant/linear, hence has a ratio between 1 and 2 with limiter. In the same sense, with limiting, quadratic reconstruction should have a ratio between 2 and 3 theoretically, and it is consistent with numerical experiment.

The maximum of numerical solution occurs at the peak, and it tells us how much the flux limiter is overshooting by how far it is from 1. From the figure of solutions (peak) and table for maximum, in linear reconstruction, BDS limiter overshoots more than van Leer limiter. In quadratic reconstruction, the extremum-preserving PPM performs very well, less overshoots than the original PPM. PPM is better than BDS with monotone limiting on its quadratic term.

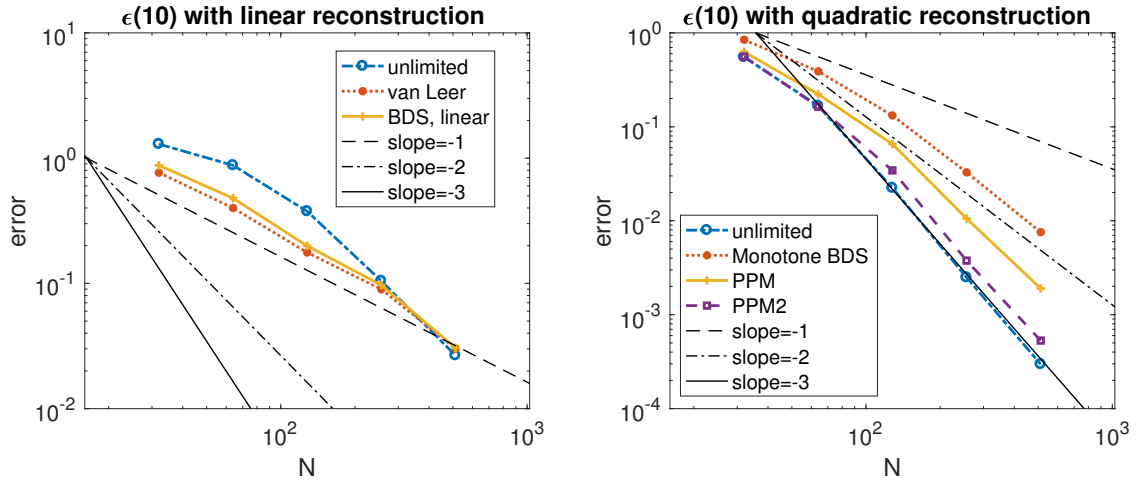

Figure 1: With linear/quadratic reconstruction, relative L^1 error for Gaussian peak constant advection with CFL 0.2 at $t = 10$ as a function of number of cells, plot in logarithm scale.

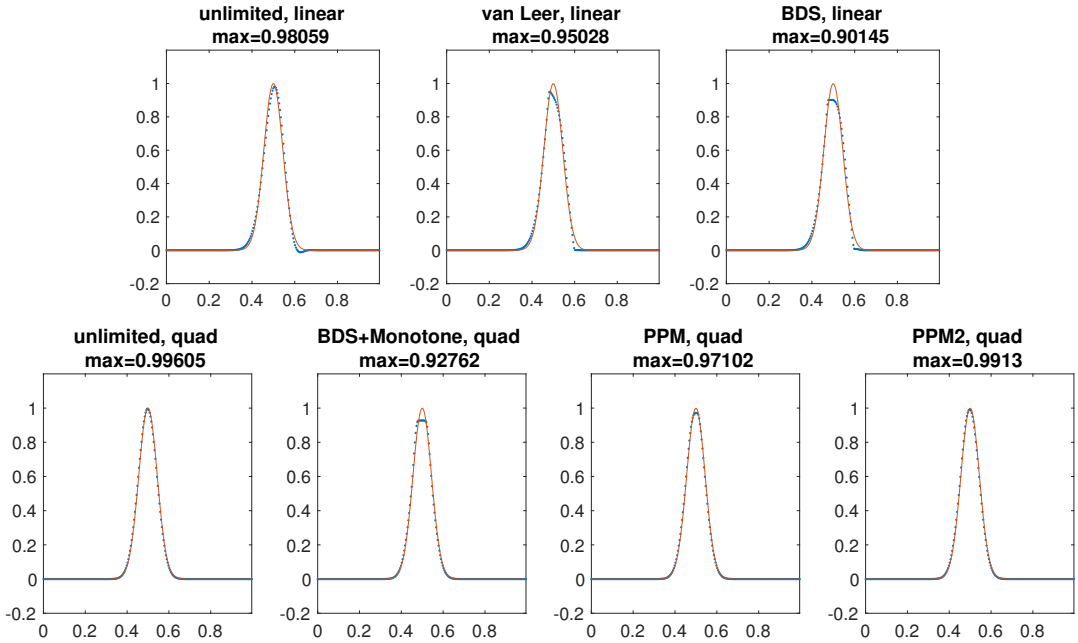

Figure 2: Numerical solution (blue) vs truth solution (red) for advection of Gaussian peak with $N = 256$ at $t = 10$.

	$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	0.4946	0.7012	0.8908	0.9806	0.9979
linear, VL	0.4461	0.6644	0.8610	0.9503	0.9791
linear, BDS	0.3610	0.5666	0.7778	0.9015	0.9580
quadratic, unlimit	0.6285	0.8599	0.9730	0.9961	0.9994
quadratic, BDS+monotone	0.3740	0.6132	0.8207	0.9276	0.9732
quadratic, PPM	0.4935	0.7610	0.9079	0.9710	0.9913
quadratic,PPM2	0.5379	0.8278	0.9582	0.9913	0.9978

Table 2: Maximum at $t = 10$ with CFL 0.2 for constant advection and Gaussian peak initial condition with different reconstruction. Theoretically, its value should be equal to 1.

5.2 Semi-circle initial condition

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	error	0.2687	0.1584	0.0886	0.0476	0.0252
	ratio	–	0.7625	0.8383	0.8956	0.9181
linear, VL	error	0.1601	0.0925	0.0537	0.0313	0.0165
	ratio	–	0.7913	0.7844	0.7786	0.9251
linear, BDS	error	0.1578	0.0933	0.0546	0.0322	0.0172
	ratio	–	0.7582	0.7733	0.7606	0.9026
quadratic, unlimit	error	0.0944	0.0447	0.0201	0.0089	0.0040
	ratio	–	1.0779	1.1555	1.1690	1.1606
quadratic, BDS+monotone	error	0.0920	0.0539	0.0230	0.0100	0.0044
	ratio	–	0.7716	1.2264	1.1974	1.1750
quadratic, PPM	error	0.0817	0.0493	0.0227	0.0105	0.0049
	ratio	–	0.7295	1.1195	1.1074	1.0985
quadratic, PPM2	error	0.0820	0.0492	0.0227	0.0105	0.0049
	ratio	–	0.7364	1.1174	1.1175	1.0943

Table 3: Relative L^1 error and empirical order at $t = 10$ with CFL 0.2 for constant advection and semi-circle initial condition with different reconstruction.

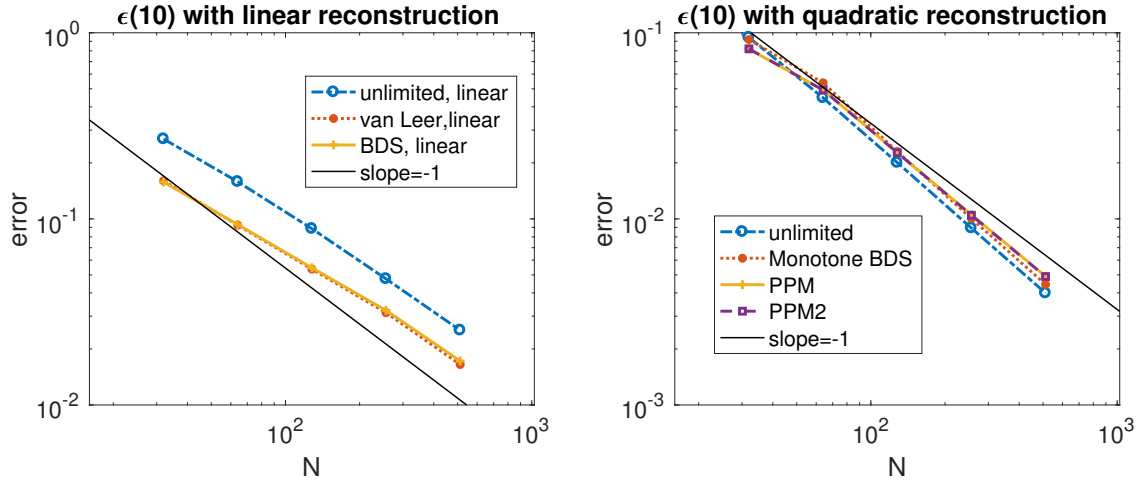

Figure 3: Relative L^1 error for semi-circle constant advection with CFL 0.2 at $t = 10$ as a function of number of cells, plot in logarithm scale.

The ratio of error is slightly less than 1 for linear reconstruction, and slightly more than 1 for quadratic reconstruction. The solution is continuous, but has discontinuous first order derivative at two points, so both reconstruction will behave like around first order accurate, as the first order derivative can not be estimated well only by finite difference between adjacent cells. However, it can do well for linear/quadratic estimation when it is far from the two discontinuities, so it makes sense to have an accuracy order above one with quadratic reconstruction.

In this problem, there is a smooth maximum, and are infinitely many minimums. With no limiter, both linear and quadratic reconstruction has oscillations, as its max/min has larger absolute value than truth. All of the limiters succeed to keep the values in the possible range. Basically, the conclusion is same as Gaussian test: BDS overshoots more than van Leer and PPMs, at both the smooth maximum and the minimums its values are farther from the truth.

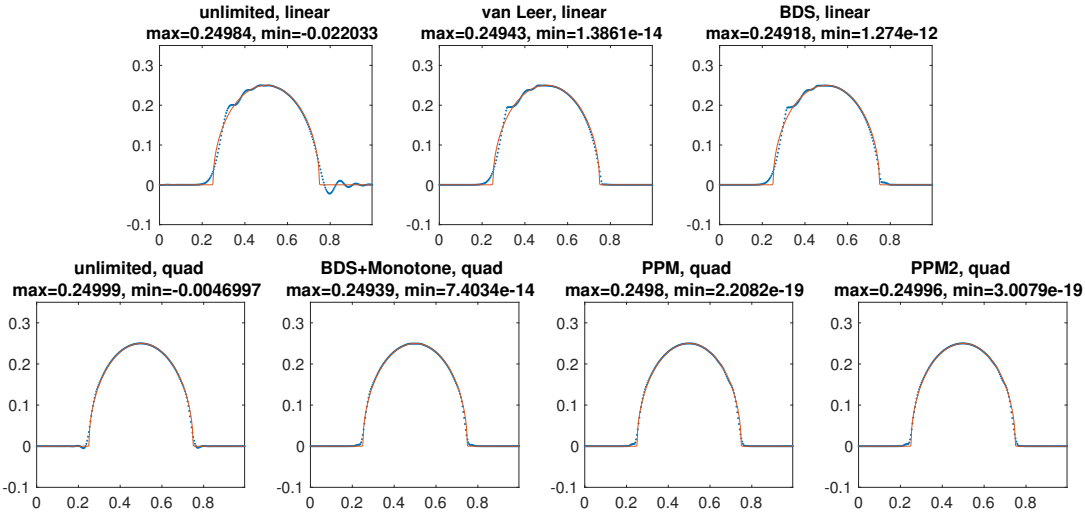

Figure 4: Numerical solution for semi-circle constant advection with CFL 0.2, $N = 256$ at $t = 10$.

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	max	0.2698	0.2649	0.2541	0.2498	0.2500
	min	-0.0351	-0.0315	-0.0264	-0.0220	-0.0180
linear, VL	max	0.2431	0.2433	0.2483	0.2494	0.2499
	min	0.0010	0.0000	0.0000	0.0000	0.0000
linear, BDS	max	0.2309	0.2427	0.2477	0.2492	0.2497
	min	0.0032	0.0001	0.0000	0.0000	0.0000
quadratic, unlimit	max	0.2540	0.2502	0.2500	0.2500	0.2500
	min	-0.0105	-0.0085	-0.0063	-0.0047	-0.0035
quadratic, BDS+monotone	max	0.2338	0.2449	0.2483	0.2494	0.2498
	min	0.0023	0.0000	0.0000	0.0000	0.0000
quadratic, PPM	max	0.2409	0.2476	0.2494	0.2498	0.2499
	min	0.0004	0.0000	0.0000	0.0000	0.0000
quadratic, PPM2	max	0.2409	0.2482	0.2494	0.2500	0.2500
	min	0.0002	0.0000	0.0000	0.0000	0.0000

Table 4: Maximum/minimum at $t = 10$ with CFL 0.2 for constant advection and semi-circle initial condition with different reconstruction. The maximum of analytical solution is 0.25, and minimum is 0.

5.3 Square wave initial condition

Small CFL number: 0.2

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	error	0.4237	0.2899	0.2019	0.1355	0.0902
	ratio	–	0.5473	0.5221	0.5749	0.5870
linear, VL	error	0.2489	0.1597	0.1021	0.0652	0.0415
	ratio	–	0.6405	0.6443	0.6486	0.6515
linear, BDS	error	0.2535	0.1635	0.1052	0.0676	0.0434
	ratio	–	0.6332	0.6361	0.6383	0.6384
quadratic, unlimit	error	0.1925	0.1203	0.0703	0.0406	0.0235
	ratio	–	0.6779	0.7750	0.7931	0.7906
quadratic, BDS monotone	error	0.1887	0.1110	0.0652	0.0384	0.0227
	ratio	–	0.7649	0.7679	0.7632	0.7604
quadratic, PPM	error	0.1775	0.1043	0.0618	0.0373	0.0233
	ratio	–	0.7679	0.7542	0.7275	0.6772
quadratic, PPM2	error	0.1775	0.1043	0.0618	0.0373	0.0233
	ratio	–	0.7667	0.7557	0.7275	0.6768

Table 5: Relative L^1 error and empirical order at $t = 10$ with CFL 0.2 for constant advection and Gaussian peak initial condition with different reconstruction.

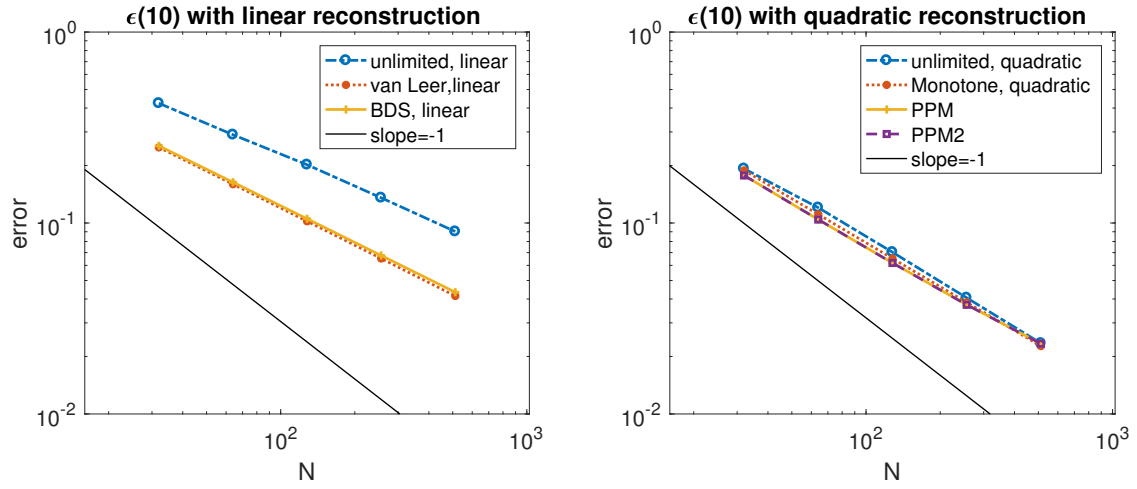

Figure 5: Relative L^1 error for square wave constant advection with CFL 0.2 at $t = 10$ as a function of number of cells, plot in logarithm scale.

This test problem has two discontinuities, and each point is either a global maximum or global minimum. This is more difficult than the semi-circle one, as there is discontinuities in itself (which can not be resolved well because finite difference in interpolation introduce some numerical 'correlation' between adjacent cells in this case), so the accuracy order will be lower for both linear and quadratic reconstructions.

In this case, when the resolution is high enough, with all of the limiters the maximum and minimum are very accurate, while there can be large oscillations near discontinuities with no limiter (see the figure of solutions). Still, BDS overshoots more, underestimates maximum and overestimates minimum, and PPM2 does slightly better than PPM (which is not obvious from figures, but from max/min values and L^1 error).

This numerical test is done with small CFL number 0.2, which is not good for this discontinuous problem, as the more time steps it takes, the more 'correlation' there will be between adjacent cells, so we see that a lot of points are needed to resolve the discontinuities. In the following, results on the same problem, but with CFL 0.9 will be considered, and it provideds important conclusions.

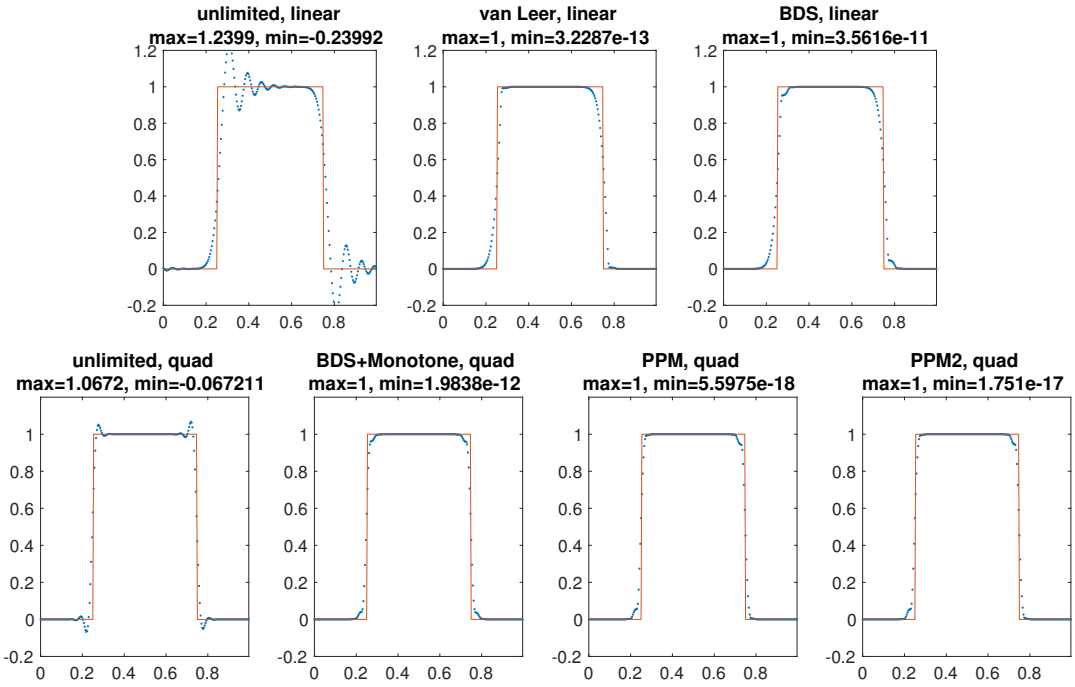

Figure 6: Numerical solution for square wave constant advection with CFL 0.2, $N = 256$ at $t = 10$.

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	max	1.2069	1.2201	1.2297	1.2399	1.2471
	min	0.9887	0.9998	1.0000	1.0000	1.0000
linear, VL	max	0.9887	0.9998	1.0000	1.0000	1.0000
	min	0.0113	0.0002	0.0000	0.0000	0.0000
linear, BDS	max	0.9769	0.9990	1.0000	1.0000	1.0000
	min	0.0231	0.0010	0.0000	0.0000	0.0000
quadratic, unlimit	max	1.0698	1.0719	1.0692	1.0672	1.0656
	min	-0.0698	-0.0719	-0.0692	-0.0672	-0.0656
quadratic, BDS+monotone	max	0.9784	0.9993	1.0000	1.0000	1.0000
	min	0.0216	0.0007	0.0000	0.0000	0.0000
quadratic, PPM	max	0.9951	1.0000	1.0000	1.0000	1.0000
	min	0.0049	0.0000	0.0000	0.0000	0.0000
quadratic, PPM2	max	0.9962	1.0000	1.0000	1.0000	1.0000
	min	0.0038	0.0000	0.0000	0.0000	0.0000

Table 6: Maximum/minimum at $t = 10$ with CFL 0.2 for constant advection and square wave initial condition with different reconstruction.

Large CFL number: 0.9

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	error	0.1943	0.1407	0.0965	0.0655	0.0423
	ratio	–	0.4665	0.5441	0.5577	0.6300
linear, VL	error	0.1305	0.0856	0.0565	0.0386	0.0252
	ratio	–	0.6094	0.5977	0.5490	0.6187
linear, BDS	error	0.1322	0.0849	0.0549	0.0359	0.0221
	ratio	–	0.6391	0.6279	0.6142	0.7022
quadratic, unlimit	error	0.1424	0.0788	0.0455	0.0258	0.0148
	ratio	–	0.8531	0.7943	0.8156	0.8068
quadratic, BDS+monotone	error	0.1261	0.0694	0.0405	0.0240	0.0138
	ratio	–	0.8622	0.7749	0.7554	0.7989
quadratic, PPM	error	0.1229	0.0666	0.0388	0.0230	0.0133
	ratio	–	0.8852	0.7781	0.7560	0.7842
quadratic, PPM2	error	0.1229	0.0666	0.0388	0.0230	0.0133
	ratio	–	0.8849	0.7781	0.7560	0.7842

Table 7: Relative L^1 error and empirical order at $t = 10$ with CFL 0.9 for constant advection and Gaussian peak initial condition with different reconstruction.

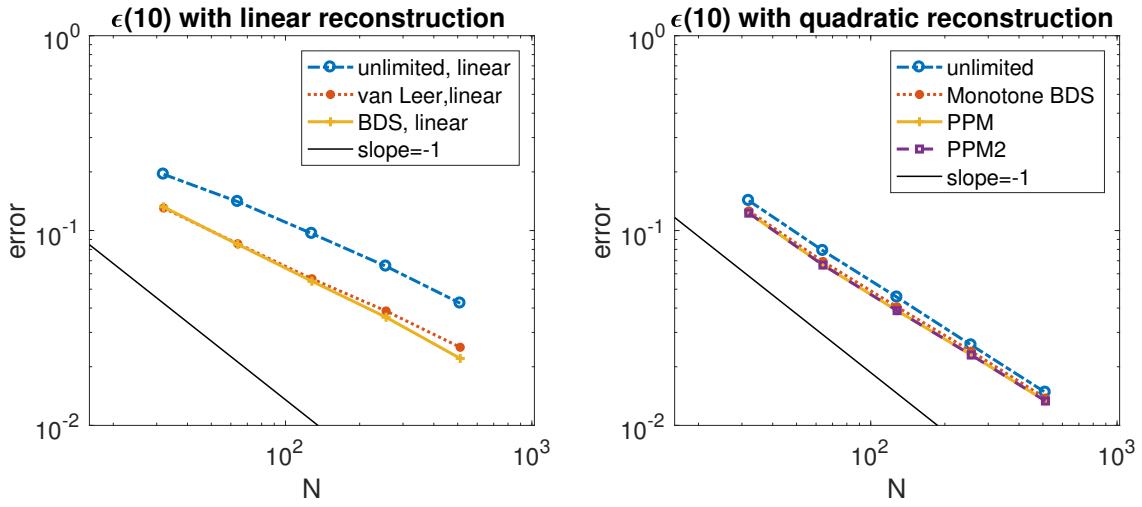

Figure 7: Relative L^1 error for square wave constant advection with CFL 0.9 at $t = 10$ as a function of number of cells, plot in logarithm scale.

As mentioned in the results with CFL number equal to 0.2, with a larger CFL number, the L^1 error is reduced because less cells are needed to resolve the discontinuities, even for the unlimited methods. And a new observation is that, in this case, the van Leer limiter is not enough for limiting, as oscillations are still seen in solutions near discontinuities. Although BDS overshoots, it always succeed to keep extemum values in the possible range, which might be important for more complicated problems, as in the 2D cases when initial data is discontinuous, and velocity spacially dependent (Table 13, [1], where PPM/PPM2 are insufficient in limiting).

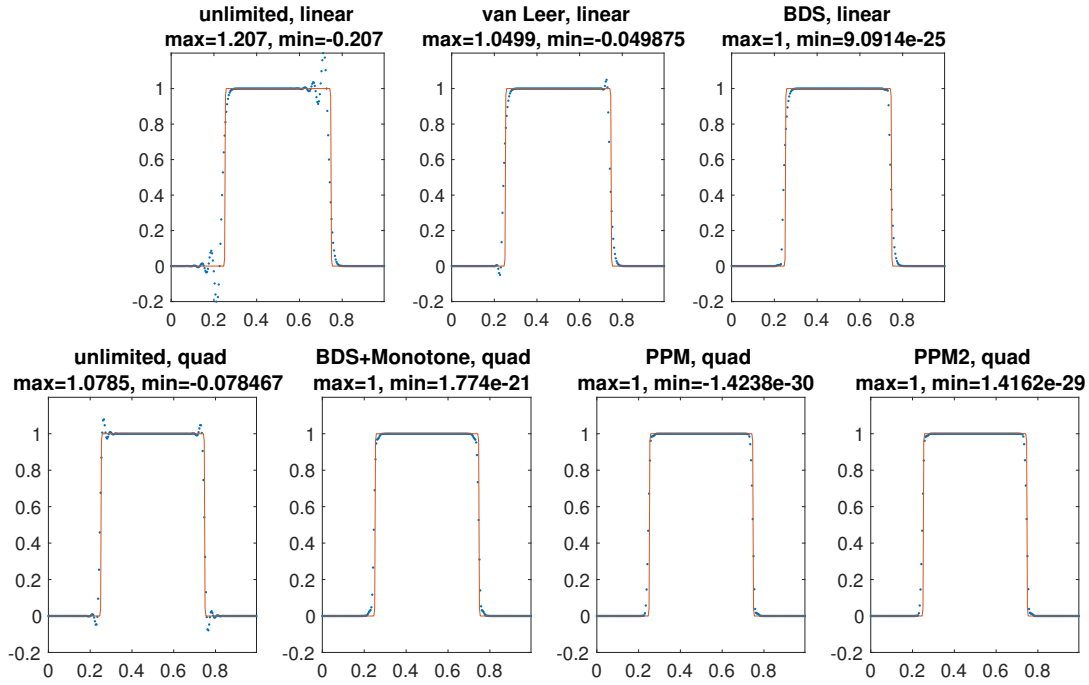

Figure 8: Numerical solution for square wave constant advection with CFL 0.9, $N = 256$ at $t = 10$.

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	max	1.1400	1.1691	1.1877	1.2070	1.2228
	min	-0.1400	-0.1691	-0.1877	-0.2070	-0.2228
linear, VL	max	1.0085	1.0138	1.0253	1.0499	1.0778
	min	-0.0085	-0.0138	-0.0253	-0.0499	-0.0778
linear, BDS	max	0.9995	1.0000	1.0000	1.0000	1.0000
	min	0.0005	0.0000	0.0000	0.0000	0.0000
quadratic, unlimit	max	1.0893	1.0860	1.0828	1.0785	1.0789
	min	-0.0893	-0.0860	-0.0828	-0.0785	-0.0789
quadratic, limited	max	0.9992	1.0000	1.0000	1.0000	1.0000
	min	0.0008	0.0000	0.0000	0.0000	0.0000
quadratic, PPM	max	1.0000	1.0000	1.0000	1.0000	1.0000
	min	0.0000	0.0000	0.0000	-0.0000	0.0000
quadratic, PPM2	max	1.0000	1.0000	1.0000	1.0000	1.0000
	min	0.0000	0.0000	0.0000	-0.0000	0.0000

Table 8: Maximum/minimum at $t = 10$ with CFL 0.2 for constant advection and square wave initial condition with different reconstruction.

6 A numerical example with spatially-dependent advection

In this section, we examine the accuracy for spatially-dependent advection problem. As no problem with known analytical solution is provided, here we use method of manufactured solution, with:

$$s_t(x,t) + [(\sin(2\pi x) + 2)s(x,t)]_x = 0, \quad s(x,0) = \cos(2\pi x), \quad x \in [0,1], \quad s(0,t) = s(1,t).$$

If we prescribe the truth solution as $s_{exact} = \cos(2\pi(x+t))$, the accuracy can be tested by solving the equation with a source term:

$$s_t(x,t) + [(\sin(2\pi x) + 2)s(x,t)]_x = (s_{exact})_t(x,t) + [(\sin(2\pi x) + 2)s_{exact}(x,t)]_x = f(x,t).$$

where the function $f(x,t)$ can be calculated explicitly as $2\pi\cos(2\pi(2x+t)) - 6\pi\sin(2\pi(x+t))$. The influence of source term in implementation is in two parts:

- In the predictor step of face values, the equation should go half of the time step with the forcing:

$$s^p_{i+1/2} = \frac{1}{u_{i+1/2}\Delta t}\int_{x_{i+1/2}-u_{i+1/2}\Delta t}^{x_{i+1/2}} \left[s_{xx,i}(x-x_i)^2 + s_{x,i}(x-x_i) + s_i\right] dx + f_i\Delta t/2.$$

- In the temporal integrator, the equation should go one full time step forward:

$$s^{n+1}_j = s^n_j - \frac{\Delta t}{\Delta x}[u_{j+1/2}s_{j+1/2} - u_{j-1/2}s_{j-1/2}] + f_j\Delta t.$$

The method is initialized analytically with cell average at time 0, and we compare the exact solution as cell average analytically in the same way, so they are exact. The error and convergence ratio are shown in the table below, with $t = 10$, and the maximum value for CFL number is $\max_x(u(x))\Delta t/\Delta x = 3\Delta t/\Delta x = 0.6$. The table 9 and logarithm plot of relative global error 10 shows a second order convergence for both linear and quadratic reconstruction (slightly, quadratic one is better, and limiting only does overshoot and no help in this problem), and the local error converge to a smooth function in space (see figure 9) , which indicates it has reached its asymptotic limit.

		$N = 32$	$N = 64$	$N = 128$	$N = 256$	$N = 512$
linear, unlimit	error	0.0055	0.0015	0.0004	0.0001	0.0000
	ratio	–	1.9659	1.9881	1.9977	1.9995
linear, BDS	error	0.0056	0.0015	0.0004	0.0001	0.0000
	ratio	–	1.9843	1.9854	1.9960	1.9984
quadratic, unlimit	error	0.0053	0.0014	0.0003	0.0001	0.0000
	ratio	–	1.9622	1.9938	1.9992	1.9999
quadratic, BDS+monotone	error	0.0054	0.0014	0.0003	0.0001	0.0000
	ratio	–	1.9966	2.0027	1.9994	2.0002

Table 9: Relative L^1 error and empirical order at $t = 10$ with maximal CFL 0.6 for non-constant advection with manufactured solution and square wave initial condition.

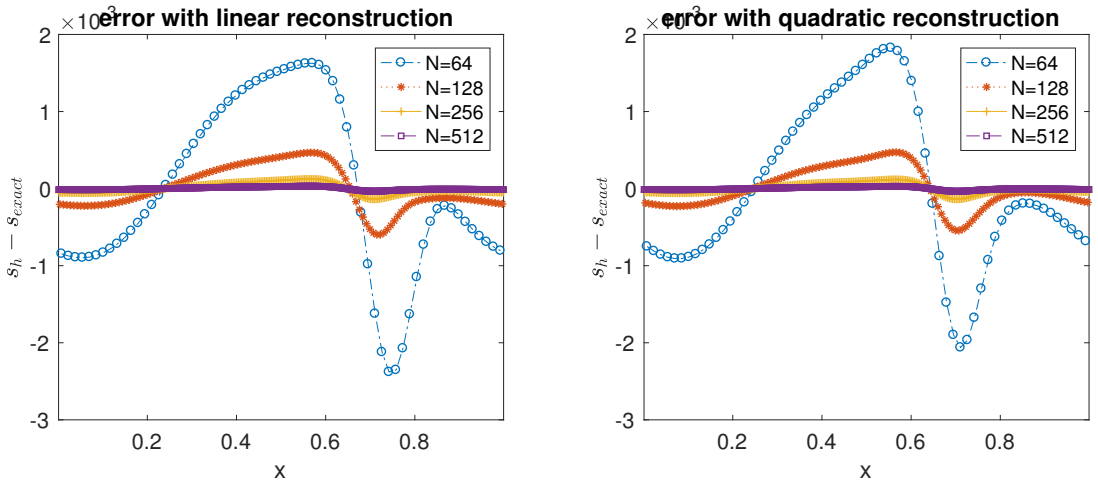

Figure 9: Local error for linear/quadratic reconstruction as function of space without limiting.

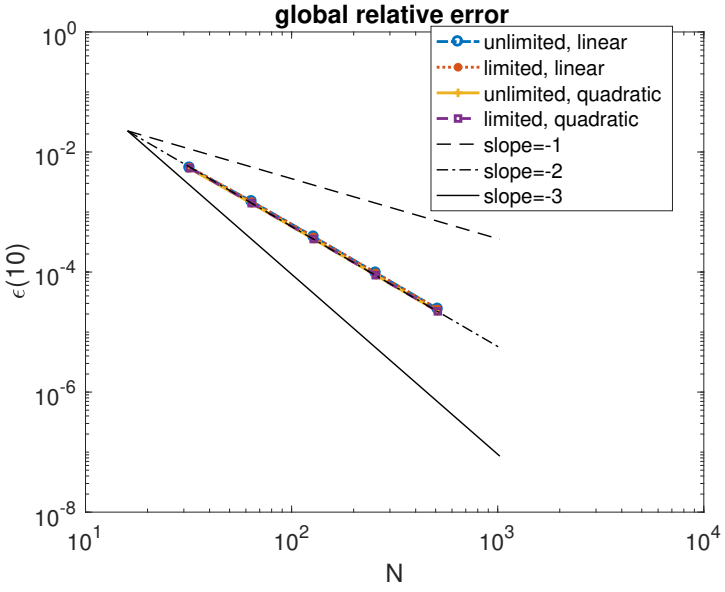

Figure 10: Logarithm plot for global relative error at $t = 10$ as function of number of cells.

References

[1] *An Unsplit, Higher-order Godunov Method Using Quadratic Reconstruction for Advection in Two Dimensions*, Sandra May, Andrew Nonaka, Ann Almgren and John Bell.

[2] *The Piecewise Parabolic Method (PPM) for Gas-dynamical Simulations*, Phillip Colella.

[3] *A limiter for PPM that preserves accuracy at smooth extrema*, Phillip Colella, Michael D. Sekora.