

A GENUINELY HIGH ORDER TOTAL VARIATION DIMINISHING SCHEME FOR ONE-DIMENSIONAL SCALAR CONSERVATION LAWS*

XIANGXIONG ZHANG[†] AND CHI-WANG SHU[‡]

Abstract. It is well known that finite difference or finite volume total variation diminishing (TVD) schemes solving one-dimensional scalar conservation laws degenerate to first order accuracy at smooth extrema [8], thus TVD schemes are at most second order accurate in the L^1 norm for general smooth and non-monotone solutions. However, Sanders [12] introduced a third order accurate finite volume scheme which is TVD, where the total variation is defined by measuring the variation of the reconstructed polynomials rather than the traditional way of measuring the variation of the grid values. By adopting the definition of the total variation for the numerical solutions as in [12], it is possible to design genuinely high order accurate TVD schemes. In this paper, we construct a finite volume scheme which is TVD in this sense with high order accuracy (up to sixth order) in the L^1 norm. Numerical tests for a fifth order accurate TVD scheme will be reported, which include test cases from traffic flow models.

Key words. hyperbolic conservation laws; finite volume scheme; total variation diminishing; total variation bounded; high order accuracy; conservative form

AMS subject classifications. 65M06

1. Introduction. We consider numerical solutions of one-dimensional hyperbolic scalar conservation law

$$(1.1) \quad u_t + f(u)_x = 0, \quad u(x, 0) = u_0(x),$$

where $u_0(x)$ is assumed to be a bounded variation function. The main difficulty in solving (1.1) is that the solution may contain discontinuities even if the initial condition is smooth.

Successful numerical schemes for solving (1.1) are usually total variation stable, for example the total variation diminishing (TVD) schemes [1] or the total variation bounded (TVB) schemes [14], or essentially non-oscillatory, for example the essentially non-oscillatory (ENO) schemes [2, 15] or the weighted ENO (WENO) schemes [6, 4]. ENO and WENO schemes, although uniformly high order accurate and stable in applications, do not have mathematically provable TVB properties for general solutions and do not satisfy a maximum principle. It is certainly desirable to have a TVD or TVB scheme, which shares the TVD property of the exact entropy solution of (1.1), satisfies a maximum principle, and has at least a convergent subsequence to a weak solution of (1.1) due to its compactness.

Typically, for a finite difference scheme with the numerical solution given by the grid values u_j , or a finite volume scheme with the numerical solution given by the cell averages u_j , the total variation of the numerical solution is measured by

$$(1.2) \quad TV(u) = \sum_j |u_{j+1} - u_j|$$

*Research supported by NSF grant DMS-0809086 and AFOSR grant FA9550-09-1-0126.

[†]Department of Mathematics, Brown University, Providence, RI 02912. E-mail: zhangxx@dam.brown.edu

[‡]Division of Applied Mathematics, Brown University, Providence, RI 02912. E-mail: shu@dam.brown.edu

which is the standard bounded variation semi-norm when the numerical solution is considered to be a piecewise constant function with the data u_j . A scheme is TVD if the numerical solution satisfies $TV(u^{n+1}) \leq TV(u^n)$ where u^n refers to the numerical solution at the time level t^n . A TVB scheme is one which satisfies $TV(u^n) \leq M$ for all n such that $t^n \leq T$, where the constant M does not depend on the mesh sizes but may depend on T . A sufficient condition for a scheme to be TVB is

$$TV(u^{n+1}) \leq TV(u^n) + M\Delta t, \text{ or } TV(u^{n+1}) \leq (1 + M\Delta t)TV(u^n),$$

where M is a constant and Δt is the time step.

It is well known that finite difference or finite volume TVD schemes solving (1.1), where the total variation is measured by (1.2), necessarily degenerate to first order accuracy at smooth extrema [8], thus TVD schemes are at most second order accurate in the L^1 norm for general smooth and non-monotone solutions. While the TVB schemes in [14] can overcome this accuracy degeneracy difficulty, the schemes are no longer scale-invariant (scale-invariance refers to the fact that the scheme does not change when x and t are scaled by the same factor) and involve a TVB parameter M which must be estimated and adjusted for individual problems.

In [12], Sanders introduced a third order accurate finite volume scheme which is TVD. The main idea in [12] is to define the total variation by measuring the variation of the reconstructed polynomials, rather than the traditional measurement as in (1.2). The scheme of Sanders in [12] can be summarized in the following steps.

- Start from the cell averages \bar{u}_j^0 and the cell boundary values $u_{j+\frac{1}{2}}^0$ for all j from the initial condition $u_0(x)$.
- For $n = 0, 1, \dots$, perform the following
 1. Reconstruct a piecewise quadratic polynomial solution $u^n(x)$, based on the information \bar{u}_j^n and $u_{j+\frac{1}{2}}^n$ for all j , such that $u^n(x)$ is third order accurate (degenerates to second order at isolated critical points, therefore still third order in the L^1 norm), and TVD

$$(1.3) \quad TV(u^n(x)) \leq TV(u^{n-1}(x))$$

where the total variation is measured by the standard bounded variation semi-norm of the piecewise quadratic polynomial solution $u^n(x)$. For $n = 0$, $TV(u^{n-1}(x))$ is taken as the bounded variation semi-norm of the initial condition $u_0(x)$.

2. Evolve the PDE (1.1) *exactly* for one time step Δt from the “initial condition” $u^n(x)$ at the time level t^n , and take the cell averages \bar{u}_j^{n+1} and the cell boundary values $u_{j+\frac{1}{2}}^{n+1}$ for all j from this exactly evolved solution. Then return to Step 1 above.

The crucial step in Sanders’ scheme is the reconstruction, which should be high order accurate and TVD in the sense of (1.3). The second step above, namely the exact time evolution and cell averaging, does not increase the total variation. The resulting scheme is thus TVD as long as the reconstruction is TVD. The purpose of this paper is to generalize the scheme of Sanders, mainly the step of the high order TVD reconstruction, to higher order accuracy. In order to measure the total variation of a polynomial $p(x)$ of degree k over the interval $[a, b]$, we would need to obtain the zeros of its derivative $p'(x)$ in this interval, denoted by a_1, a_2, \dots, a_{k-1} . If we also denote $a_0 = a$ and $a_k = b$, then the variation of $p(x)$ over the interval $[a, b]$ is

$\int_a^b |p'(x)| dx = \sum_{j=1}^k |p(a_j) - p(a_{j-1})|$. If we insist on working with explicit formulas for the zeros of polynomials to save cost, then we can only have the polynomial degree k of $p(x)$ up to five, hence our approach can generate schemes up to sixth order in the L^1 norm.

A major difficulty in generalizing Sanders' TVD reconstruction to higher order is the design of the nonlinear limiter, which should maintain accuracy in smooth regions while enforcing the TVD property (1.3). The original limiter of Sanders in [12] works well for the third order scheme ($k = 2$), but it seems difficult to generalize it directly to higher order. We will develop a different limiter to achieve this purpose. The finite volume schemes we develop in this paper are $(k+1)$ -th order accurate in the L^1 norm, conservative, and TVD in the sense of (1.3). Numerical tests for a fifth order accurate TVD scheme will be reported, which include test cases from traffic flow models.

This paper is divided into six sections. In Section 2 we develop a Hermite type reconstruction procedure using piecewise polynomials of degree k for functions of bounded variation. We use a quartic polynomial ($k = 4$) as an example to illustrate the procedure. The reconstruction does not increase the variation of the function being approximated, and is $(k+1)$ -order accurate in regions where the approximated function is smooth. In Section 3 we show how to evolve the approximation in time in essentially the same way as in [12]. In Section 4, we derive the conservative form of the scheme and show that the local truncation error is formally $(k+1)$ -th order accurate. In Section 5 we present some numerical examples of the fifth order scheme using quartic reconstruction polynomials. Numerical examples include those from traffic flow models. Finally, in Section 6, we give concluding remarks and remarks for future work.

2. A fifth order TVD reconstruction. For a smooth function $u(x)$ with bounded variation over an interval $J \subset \mathbf{R}$, we would like to find a piecewise polynomial $r(x)$ approximating $u(x)$ with the property that the variation of $r(x)$ does not exceed that of $u(x)$. We will use a Hermite type quartic reconstruction polynomial as an example to show how to find such an approximation. However we remark that this approximation procedure can also be applied to any reconstruction polynomials of degree up to five, including those obtained with the ENO or WENO procedure.

We will use the following notation for our mesh $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$, $\Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, which is assumed to be uniform for simplicity, and for the discretization of u over the mesh, $\bar{u}_j = \frac{1}{\Delta x} \int_{I_j} u(x) dx$, $u_{j-\frac{1}{2}} = u(x_{j-\frac{1}{2}})$. Given the cell averages and cell boundary point values of $u(x)$, there are many ways to obtain a reconstruction polynomial $p_j(x)$ over the cell I_j . We choose to use the Hermite type reconstruction of degree four; i.e., the polynomial $p_j(x)$ should satisfy

$$\frac{1}{\Delta x} \int_{I_i} p_j(x) dx = \bar{u}_i, \quad i = j-1, j, j+1; \quad \text{and} \quad p_j(x_{j\pm\frac{1}{2}}) = u(x_{j\pm\frac{1}{2}}).$$

If the reconstruction is written as $p_j(x) = a_4(x-x_j)^4 + a_3(x-x_j)^3 + a_2(x-x_j)^2 + a_1(x-x_j) + a_0$, then the coefficients can be given explicitly as

$$\begin{aligned} a_0 &= \frac{\bar{u}_{j-1} + 298\bar{u}_j + \bar{u}_{j+1} - 54(u_{j-\frac{1}{2}} + u_{j+\frac{1}{2}})}{192}, & a_1 &= \frac{\bar{u}_{j-1} - \bar{u}_{j+1} - 10(u_{j-\frac{1}{2}} - u_{j+\frac{1}{2}})}{8\Delta x}, \\ a_2 &= \frac{-(\bar{u}_{j-1} + 58\bar{u}_j + \bar{u}_{j+1}) + 30(u_{j-\frac{1}{2}} + u_{j+\frac{1}{2}})}{8\Delta x^2}, & a_3 &= \frac{\bar{u}_{j+1} - \bar{u}_{j-1} + 2(u_{j-\frac{1}{2}} - u_{j+\frac{1}{2}})}{\Delta x^3}, \\ a_4 &= \frac{5\bar{u}_{j-1} + 50\bar{u}_j + 5\bar{u}_{j+1} - 30(u_{j-\frac{1}{2}} + u_{j+\frac{1}{2}})}{12\Delta x^4}. \end{aligned}$$

It is clear that the following properties hold for $p_j(x)$:

I *Accuracy* (recall that $u(x)$ is a smooth function)

$$(2.1) \quad p_j(x) = u(x) + O(\Delta x^5), \quad \forall x \in I_j$$

II *Agreement of the cell averages*

$$(2.2) \quad \frac{1}{\Delta x} \int_{I_j} p_j(x) dx = \bar{u}_j, \quad \forall j$$

Define the piecewise polynomial $p(x) = \sum_j p_j(x) \chi_j(x)$, where χ_j is the characteristic function on I_j . Notice that for the specific Hermite polynomial $p_j(x)$ defined above, we have $p(x_{j-\frac{1}{2}}) = p_j(x_{j-\frac{1}{2}}) = p_{j-1}(x_{j-\frac{1}{2}}) = u(x_{j-\frac{1}{2}})$, that is, $p(x)$ is a continuous function. For any piecewise polynomial function $r(x) = \sum_j r_j(x) \chi_j(x)$, where $r_j(x)$ is a polynomial which may not satisfy $r_j(x_{j-\frac{1}{2}}) = u(x_{j-\frac{1}{2}})$ or $r_j(x_{j+\frac{1}{2}}) = u(x_{j+\frac{1}{2}})$, we define the variation of r on each cell by

$$Var(r_j) = \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} |r'_j(x)| dx + |r_j(x_{j-\frac{1}{2}}) - u_{j-\frac{1}{2}}| + |r_j(x_{j+\frac{1}{2}}) - u_{j+\frac{1}{2}}|,$$

and define the variation of r on the whole domain by $Var(r) = \sum_j Var(r_j)$. The standard total variation seminorm of $r(x)$ is

$$TV(r) = \sum_j \left[\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} |r'_j(x)| dx + |r_j(x_{j-\frac{1}{2}}) - r_{j-1}(x_{j-\frac{1}{2}})| \right] \leq \sum_j Var(r_j).$$

The approximation $r(x)$ is TVD if it satisfies $TV(r) = \int_J |r'(x)| dx \leq \int_J |u'(x)| dx = TV(u)$, where the integral is in the generalized sense since $r'(x)$ may contain δ -functions, because $r(x)$ is a piecewise polynomial function which may not be continuous at cell interfaces. Obviously, r is a TVD approximation if $Var(r_j) \leq \int_{I_j} |u'(x)| dx$, $\forall j$, where we assume the cell boundary values $u(x_{j-\frac{1}{2}})$ are well defined and are shared by both cells I_{j-1} and I_j . For convenience, we denote the total variation of the function u over the interval I_j by $TV(u)_j = \int_{I_j} |u'(x)| dx$.

Now the question is, given a piecewise approximation polynomial $p(x)$ and certain information of $u(x)$ including its extrema, its cell boundary values $u_{j-\frac{1}{2}}$, its cell averages \bar{u}_j and its cell total variation $TV(u)_j$ for all j , whether we can find a TVD approximation $r(x)$ by limiting $p(x)$ such that the accuracy condition (2.1) and the conservation condition (2.2) still hold.

To obtain the TVD property, it suffices to enforce

$$(2.3) \quad Var(r_j) \leq TV(u)_j, \quad \forall j.$$

We now discuss the limiting process case by case:

Case I. $p_j(x)$ is monotone in the cell I_j .

For this case, the TVD requirement (2.3) is already satisfied since $p_j(x)$ is a Hermite type polynomial which interpolates u at the cell boundaries.

Case II. $p_j(x)$ is not monotone in the cell I_j , but $u(x)$ is monotone in this cell.

In this case, the TVD property (2.3) does not hold for $p_j(x)$ because

$$Var(p_j) > |u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}}| = TV(u)_j.$$

Without loss of generality, we only discuss the situation that $u(x)$ is monotonically increasing in I_j , i.e., $u_{j-\frac{1}{2}} \leq u_{j+\frac{1}{2}}$ and $u'(x) \geq 0$, $\forall x \in I_j$. Since $p_j(x)$ is not monotone, we must have $p'_j(x) < 0$ for some $x \in I_j$. Consider $\tilde{p}_j(x) = p_j(x) - \alpha_j(x - x_j)$, $\alpha_j = \min_{x \in I_j} p'_j(x)$. We then have the following lemma.

LEMMA 2.1. *The accuracy condition (2.1) and the cell average agreement condition (2.2) still hold for $\tilde{p}_j(x)$.*

Proof. The agreement of the cell average is obvious. To prove accuracy, it suffices to show $\alpha_j = O(\Delta x^4)$.

We have $p'_j(x) - u'(x) = O(\Delta x^4)$ since $p_j(x)$ is a fourth degree Hermite approximation of $u(x)$. If $p'_j(x)$ attains its minimum at the point $x^{\min} \in I_j$, then we have $u'(x^{\min}) \geq 0$, $\alpha_j = p'_j(x^{\min}) < 0$, therefore $|\alpha_j| = -\alpha_j \leq u'(x^{\min}) - \alpha_j = O(\Delta x^4)$. \square

We now enforce the TVD requirement (2.3). For this purpose, we apply the following scaling to $\tilde{p}_j(x)$ around its cell average on the cell I_j :

$$r_j^{\theta_j}(x) = \theta_j(\tilde{p}_j(x) - \bar{u}_j) + \bar{u}_j,$$

where $\theta_j \in [0, 1]$ is a parameter to be determined by the TVD requirement (2.3).

LEMMA 2.2. *If $\text{Var}(\tilde{p}_j) > \text{TV}(u)_j$, then there exists $\theta_j \in [0, 1]$ such that $\text{Var}(r_j^{\theta_j}) = \text{TV}(u)_j$.*

Proof. Take

$$\theta_j = \min \left\{ \left| \frac{u_{j-\frac{1}{2}} - \bar{u}_j}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j} \right|, \left| \frac{u_{j+\frac{1}{2}} - \bar{u}_j}{\tilde{p}_j(x_{j+\frac{1}{2}}) - \bar{u}_j} \right| \right\}.$$

First, we need to show $\theta_j \in [0, 1]$. If $\theta_j > 1$, then

$$(2.4) \quad |u_{j-\frac{1}{2}} - \bar{u}_j| > |\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j|, \quad |u_{j+\frac{1}{2}} - \bar{u}_j| > |\tilde{p}_j(x_{j+\frac{1}{2}}) - \bar{u}_j|.$$

Since $u(x)$ and $\tilde{p}_j(x)$ are both monotonically increasing in I_j , we have

$$(2.5) \quad u_{j-\frac{1}{2}} \leq \bar{u}_j \leq u_{j+\frac{1}{2}}$$

and $\tilde{p}_j(x_{j-\frac{1}{2}}) \leq \bar{u}_j \leq \tilde{p}_j(x_{j+\frac{1}{2}})$. Therefore, (2.4) implies $u_{j-\frac{1}{2}} \leq \tilde{p}_j(x_{j-\frac{1}{2}}) \leq \bar{u}_j \leq \tilde{p}_j(x_{j+\frac{1}{2}}) \leq u_{j+\frac{1}{2}}$. We then have

$$\begin{aligned} \text{Var}(\tilde{p}_j) &= \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} |\tilde{p}'_j(x)| dx + |\tilde{p}_j(x_{j-\frac{1}{2}}) - u_{j-\frac{1}{2}}| + |\tilde{p}_j(x_{j+\frac{1}{2}}) - u_{j+\frac{1}{2}}| \\ &= (\tilde{p}_j(x_{j+\frac{1}{2}}) - \tilde{p}_j(x_{j-\frac{1}{2}})) + (\tilde{p}_j(x_{j-\frac{1}{2}}) - u_{j-\frac{1}{2}}) + (u_{j+\frac{1}{2}} - \tilde{p}_j(x_{j+\frac{1}{2}})) \\ &= u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}} = \text{TV}(u)_j. \end{aligned}$$

This contradicts $\text{Var}(\tilde{p}_j) > \text{TV}(u)_j$, therefore we have $\theta_j \in [0, 1]$.

Second, we should verify the TVD property (2.3). For convenience, we denote $r_j^{\theta_j}(x)$ by $r_j(x)$ here. By the definition of the θ_j ,

$$(2.6) \quad \begin{aligned} |r_j(x_{j-\frac{1}{2}}) - \bar{u}_j| &= \theta_j |\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j| \leq |u_{j-\frac{1}{2}} - \bar{u}_j|, \\ |r_j(x_{j+\frac{1}{2}}) - \bar{u}_j| &= \theta_j |\tilde{p}_j(x_{j+\frac{1}{2}}) - \bar{u}_j| \leq |u_{j+\frac{1}{2}} - \bar{u}_j|. \end{aligned}$$

Since $u(x)$ and $r_j(x)$ are both monotonically increasing in I_j , we have (2.5) and $r_j(x_{j-\frac{1}{2}}) \leq \bar{u}_j \leq r_j(x_{j+\frac{1}{2}})$, therefore (2.6) implies $u_{j-\frac{1}{2}} \leq r_j(x_{j-\frac{1}{2}}) \leq \bar{u}_j \leq r_j(x_{j+\frac{1}{2}}) \leq u_{j+\frac{1}{2}}$.

$r_j(x_{j+\frac{1}{2}}) \leq u_{j+\frac{1}{2}}$. We thus have

$$\begin{aligned} \text{Var}(r_j) &= \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} |r'_j(x)| dx + |r_j(x_{j-\frac{1}{2}}) - u_{j-\frac{1}{2}}| + |r_j(x_{j+\frac{1}{2}}) - u_{j+\frac{1}{2}}| \\ &= (r_j(x_{j+\frac{1}{2}}) - r_j(x_{j-\frac{1}{2}})) + (r_j(x_{j-\frac{1}{2}}) - u_{j-\frac{1}{2}}) + (u_{j+\frac{1}{2}} - r_j(x_{j+\frac{1}{2}})) \\ &= u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}} = TV(u)_j. \quad \square \end{aligned}$$

Notice that this scaling does not change the cell average. Of course, we would need to show that, with the θ_j taken in Lemma 2.2, this scaling does not destroy accuracy. We would need the following result for polynomials:

LEMMA 2.3. *If a polynomial $p(x)$ of degree k ($k \leq 4$) is monotone over an interval $I = [a, b]$, and its cell average over I is \bar{p} , then*

$$(2.7) \quad \max_{x \in I} \left| \frac{p(x) - \bar{p}}{p(b) - \bar{p}} \right| \leq C, \quad \max_{x \in I} \left| \frac{p(x) - \bar{p}}{p(a) - \bar{p}} \right| \leq C,$$

where C is a constant depending only on the degree of the polynomial k . In particular, $C(2) = 2$, $C(3) = 4$ and $C(4) = 7$.

The proof is elementary but lengthy, and is therefore deferred to the Appendix. A similar (and more general) result for quadratic polynomials can be found in the appendix of [5]. \square

Using Lemma 2.3 we can prove the accuracy property:

LEMMA 2.4. *For the θ_j chosen in Lemma 2.2, the accuracy property (2.1) holds for $r_j^{\theta_j}(x)$.*

Proof. It suffices to show that $r_j^{\theta_j}(x) - \tilde{p}_j(x) = O(\Delta x^5)$. Without loss of generality, we assume $\theta_j = \left| \frac{u_{j-\frac{1}{2}} - \bar{u}_j}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j} \right|$. In this case, since both $u(x)$ and $\tilde{p}_j(x)$ are monotonically increasing, we actually have $\theta_j = \frac{u_{j-\frac{1}{2}} - \bar{u}_j}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j}$. Then,

$$\begin{aligned} r_j^{\theta_j}(x) - \tilde{p}_j(x) &= \theta_j(\tilde{p}_j(x) - \bar{u}_j) + \bar{u}_j - \tilde{p}_j(x) = (\theta_j - 1)(\tilde{p}_j(x) - \bar{u}_j) \\ &= \left(\frac{u_{j-\frac{1}{2}} - \bar{u}_j}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j} - 1 \right) (\tilde{p}_j(x) - \bar{u}_j) = \frac{u_{j-\frac{1}{2}} - \tilde{p}_j(x_{j-\frac{1}{2}})}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j} (\tilde{p}_j(x) - \bar{u}_j) \\ &= \frac{\tilde{p}_j(x) - \bar{u}_j}{\tilde{p}_j(x_{j-\frac{1}{2}}) - \bar{u}_j} (u_{j-\frac{1}{2}} - \tilde{p}_j(x_{j-\frac{1}{2}})) = O(\Delta x^5) \end{aligned}$$

where in the last equality we have used Lemma 2.3. \square

Case III. Neither $p_j(x)$ nor $u(x)$ is monotone in I_j .

First, consider the situation that $u(x)$ has only one nontrivial extremum in I_j . Without loss of generality, we assume $u(x)$ attains its maximum at $x^{\max} \in I_j$, and we denote $u^{\max} = u(x^{\max})$.

We propose to break the polynomial $p_j(x)$ into two parts, p_j^l and p_j^r . Let \bar{u}_j^l and \bar{u}_j^r be the cell averages of u on $I_j^l = [x_{j-\frac{1}{2}}, x^{\max}]$ and $I_j^r = [x^{\max}, x_{j+\frac{1}{2}}]$ respectively, similarly let \bar{p}_j^l and \bar{p}_j^r be the cell averages of $p_j(x)$ on I_j^l and I_j^r respectively. Define

$$\begin{aligned} p_j^l(x) &= p_j(x), \quad \forall x \in I_j^l, & p_j^r(x) &= p_j(x), \quad \forall x \in I_j^r, \\ u^l(x) &= u(x), \quad \forall x \in I_j^l, & u^r(x) &= u(x), \quad \forall x \in I_j^r. \end{aligned}$$

After an adjustment of the cell averages over the two sub-cells, in each of the two intervals I_j^l and I_j^r , the situation reduces to either Case I or Case II. Here, we will only briefly discuss p_j^l , since the procedure for p_j^r is the same. Since u^{\max} is the maximum, $u^l(x)$ is monotonically increasing in I_j^l .

1. If p_j^l is monotone on I_j^l (Case I), then define

$$(2.8) \quad \tilde{p}_j^l(x) = p_j^l(x) - \bar{p}_j^l + \bar{u}_j^l, \quad (r_j^l)^{\theta_j}(x) = \theta_j(\tilde{p}_j^l(x) - \bar{u}_j^l) + \bar{u}_j^l,$$

where

$$(2.9) \quad \theta_j = \min \left\{ \left| \frac{u_{j-\frac{1}{2}} - \bar{u}_j^l}{\tilde{p}_j^l(x_{j-\frac{1}{2}}) - \bar{u}_j^l} \right|, \left| \frac{u^{\max} - \bar{u}_j^l}{\tilde{p}_j^l(x^{\max}) - \bar{u}_j^l} \right|, 1 \right\}.$$

2. If p_j^l is not monotone on I_j^l (Case II), let α be the minimum of the derivative of p_j^l over I_j^l , then $\alpha < 0$. Define

$$(2.10) \quad \tilde{p}_j^l(x) = p_j^l(x) - \bar{p}_j^l + \bar{u}_j^l - \alpha \left(x - \frac{x_{j-\frac{1}{2}} + x^{\max}}{2} \right), \quad (r_j^l)^{\theta_j}(x) = \theta_j(\tilde{p}_j^l(x) - \bar{u}_j^l) + \bar{u}_j^l,$$

where

$$(2.11) \quad \theta_j = \min \left\{ \left| \frac{u_{j-\frac{1}{2}} - \bar{u}_j^l}{\tilde{p}_j^l(x_{j-\frac{1}{2}}) - \bar{u}_j^l} \right|, \left| \frac{u^{\max} - \bar{u}_j^l}{\tilde{p}_j^l(x^{\max}) - \bar{u}_j^l} \right|, 1 \right\}.$$

We define $r_j^r(x)$ on I_j^r in a similarly way, and use $r_j(x) = r_j^l(x)\chi(I_j^l) + r_j^r(x)\chi(I_j^r)$ as our TVD approximation of $u(x)$ on I_j . By the same arguments as in Cases I and II, we can show the TVD property and agreement of cell averages for $r_j(x)$. And we can also show the accuracy for $r_j(x)$, following the same arguments as in Cases I and II.

If there are multiple nontrivial extrema of $u(x)$ inside the interval I_j , for the implementation of the limiter, we will choose arbitrarily one of the extrema of $u(x)$ (e.g. the one which gives the maximum or the minimum of $u(x)$ inside the interval I_j). Without loss of generality, we assume $u(x)$ attains its maximum at $x^{\max} \in I_j$, and $u^{\max} = u(x^{\max})$ is chosen as the extremum which will be used in the limiter. We still break the polynomial $p_j(x)$ into two parts, $p_j^l(x)$ on $I_j^l = [x_{j-\frac{1}{2}}, x^{\max}]$ and $p_j^r(x)$ on $I_j^r = [x^{\max}, x_{j+\frac{1}{2}}]$. Then we perform the following modification for $p_j^l(x)$:

1. If $(u_{j-\frac{1}{2}} - \bar{u}_j^l)(\bar{u}_j^l - u^{\max}) \geq 0$ and $p_j^l(x)$ is monotonically nondecreasing on I_j^l , then do exactly the same limiting as (2.8) and (2.9).
2. If $(u_{j-\frac{1}{2}} - \bar{u}_j^l)(\bar{u}_j^l - u^{\max}) \geq 0$ and $p_j^l(x)$ is not monotone on I_j^l , do exactly the same limiting as (2.10) and (2.11).
3. Otherwise, set $p_j^l(x) = \bar{u}_j^l$.

The limiting for $p_j^r(x)$ is similar. By the same arguments as above, we can show the TVD property and the agreement of cell averages for such modified two-piece polynomials.

REMARK 2.5. *One apparent gap in the proof above is that accuracy for Case III can only be shown to hold for the case of $u(x)$ having at most one extremum in the cell I_j . We would like to justify this restriction for sufficiently small Δx . For this purpose, we assume that the initial condition $u(x, 0)$ has only finitely many strict*

smooth extrema. A pair of adjacent extrema can only consist of one maximum M_j and one minimum m_j , if the function $u(x, 0)$ is smooth and non-constant in between, with $m_j < M_j$. Denoting $C = \min_j (M_j - m_j) > 0$. If we agree to consider accuracy only in those cells in which $|\frac{\partial}{\partial x} u(x, t)| \leq M$ for a pre-determined constant M , we will take the mesh size $\Delta x < \frac{C}{M}$. If at a later time t , $u(x, t)$ has two extrema in the same cell, then they must correspond to one such pair in the initial condition with values m_j and M_j (following characteristics, along which the solution stays constant). Clearly, by the mean value theorem, there is then a point ξ in this cell such that $|\frac{\partial}{\partial x} u(\xi, t)| \geq \frac{M_j - m_j}{\Delta x} \geq \frac{C}{\Delta x} > M$, hence we do not need to consider accuracy of the numerical approximation in this cell. In summary, the reconstruction polynomials after the modifications introduced in this section always satisfy the TVD and cell average agreement properties, however accuracy can be shown only for sufficiently small Δx .

3. A TVD finite volume scheme. Combining the high order accurate TVD reconstruction described in the previous section with the method of characteristics, we obtain a high order accurate TVD Godunov-type finite volume scheme solving the one dimensional scalar conservation law (1.1).

3.1. Time evolution. To implement the approximation of the previous section, several pieces of information must be available for each cell. Specifically, the cell averages and the left and right cell boundary point values of the function being approximated must be known. To obtain this information, we follow Sanders [12] and use a staggered spatial mesh with the method of characteristics. We only need to discuss the evolution procedure for one time step.

Let $T(u_0)(x, t)$, $t \geq 0$, denote the solution to the scalar conservation law (1.1) and $R(u_0)(x)$ denote the piecewise polynomial TVD reconstruction to $u_0(x)$ in the previous section. As in the previous section, we partition the real line into nonoverlapping intervals $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, and approximate $u_0 \in BV$ by the piecewise polynomial TVD reconstruction

$$u^0(x) = R(u_0)(x) = \sum_j r_j(x) \chi_j(x),$$

where $r_j(x)$ is either a polynomial or a (possibly discontinuous) two-piece polynomials $r_j(x) = r_j^l(x) \chi(I_j^l) + r_j^r(x) \chi(I_j^r)$. Consider a staggered partition $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$. The objective of the evolution is to determine the necessary information of $T(u^0)(x, \frac{\Delta t}{2})$ such that a TVD piecewise polynomial reconstruction at time $t = \frac{\Delta t}{2}$ can be obtained.

We assume that u^0 is given by $u^0(x) = \overline{R}(u_0)(x)$, where \overline{R} denotes a “preconditioned” version of reconstruction R . By this we mean specifically that $R(u_0)$ is modified (in a way we will discuss later) so that for all j ,

$$(3.1) \quad \max_{I_j} \left| \frac{d}{dx} u^0(x) \right| |f''(u^0(x))| \Delta t < 2.$$

Essentially, to enforce (3.1) is to push extremely large gradients of $r_j(x)$ out of the interval I_j and into the jump discontinuities at cell interfaces, see [12]. An additional condition that we assume throughout is the Courant condition; that is, for all u in the range of u_0 we assume the ratio $\lambda = \Delta t / \Delta x$ is taken so that

$$(3.2) \quad |f'(u)| \lambda < 1.$$

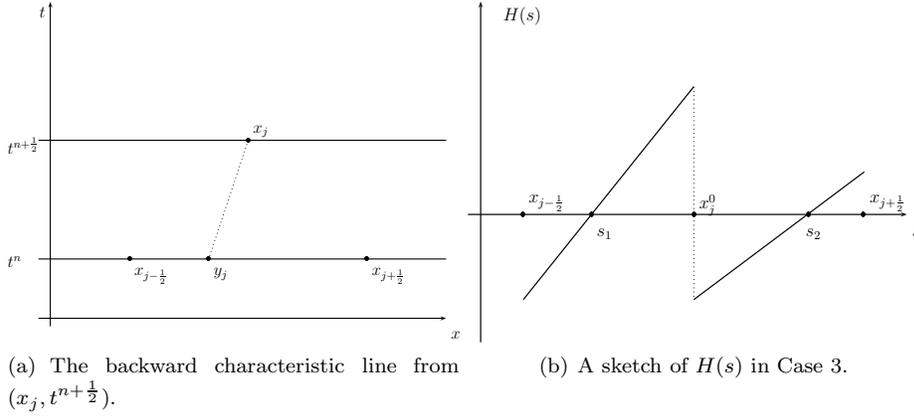


FIG. 3.1.

If we can find the backward characteristic line from each point x_j at time $t = \Delta t/2$ back to the time level $t = 0$ with its foot located at $y_j \in I_j$ (see Figure 3.1(a)) such that the exact entropy solution is constant $u_j = u^0(y_j)$ along the line, then we have already obtained the cell endpoint values of $T(u^0)(x, \frac{\Delta t}{2})$ on the staggered cell $I_{j-\frac{1}{2}}$, and we can apply the divergence theorem to (1.1) over the trapezoid defined by the points $(x_{j-1}, \Delta t/2)$, $(x_j, \Delta t/2)$, $(y_{j-1}, 0)$ and $(y_j, 0)$ to obtain the cell average of $T(u^0)(x, \frac{\Delta t}{2})$ on $I_{j-\frac{1}{2}}$ as

$$(3.3) \quad \bar{u}_{j-\frac{1}{2}}^{\frac{1}{2}} = \frac{1}{\Delta x} \left[\int_{y_{j-1}}^{y_j} u^0(x) dx - \frac{\Delta t}{2} f(u_j) + (x_j - y_j) u_j + \frac{\Delta t}{2} f(u_{j-1}) - (x_{j-1} - y_{j-1}) u_{j-1} \right].$$

Therefore, we¹ may now focus on how to obtain the backward characteristic line. With (3.1) and (3.2) we have:

LEMMA 3.1. *If the approximation function u^0 is continuous on I_j , that is, $r_j(x)$ is a polynomial, then the backward characteristic equation*

$$(3.4) \quad f'(u^0(x)) = \frac{x_j - x}{\Delta t/2}, \quad x \in I_j$$

has a unique solution $y_j \in I_j$.

Proof. Consider finding the root of the function

$$(3.5) \quad H(s) = f'(u^0(s))\Delta t + 2(s - x_j).$$

According to (3.2), we have that

$$(3.6) \quad H(x_{j-\frac{1}{2}}^+) = f'(u^0(x_{j-\frac{1}{2}}^+))\Delta t - \Delta x < 0,$$

$$(3.7) \quad H(x_{j+\frac{1}{2}}^-) = f'(u^0(x_{j+\frac{1}{2}}^-))\Delta t + \Delta x > 0.$$

Moreover, (3.1) implies that for every $s \in I_j$,

$$(3.8) \quad \frac{d}{ds} H(s) = f''(u^0(s))\Delta t \frac{d}{ds} r_j(s) + 2 > 0.$$

Therefore, $H(s)$ has a unique root $y_j \in I_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$. \square

LEMMA 3.2. *If $r_j(x) = r_j^l(x)\chi(I_j^l) + r_j^r(x)\chi(I_j^r)$ where $r_j^l(x)$ and $r_j^r(x)$ are two polynomials defined on $I_j^l = [x_{j-\frac{1}{2}}, x_j^0]$ and $I_j^r = [x_j^0, x_{j+\frac{1}{2}}]$ respectively, we can get either one or three possible candidates for the backward characteristic line from $(x_j, t^{\frac{1}{2}})$ by solving the characteristic equation (3.4).*

Proof. Again, consider finding the root of the function (3.5). We still have (3.6) and (3.7). Moreover, (3.8) holds for each piece of $s \in (x_{j-\frac{1}{2}}, x_j^0)$ and $s \in (x_j^0, x_{j+\frac{1}{2}})$, respectively. Therefore, $H(s)$ is continuous on I_j except at x_j^0 and $H(s)$ is monotonically increasing both on the left side and on the right side of x_j^0 . Thus, there are three cases depending on the signs of $H((x_j^0)^-)$ and $H((x_j^0)^+)$.

1. If $H((x_j^0)^-)H((x_j^0)^+) > 0$, there is exactly one root of $H(s)$.
2. If $H((x_j^0)^-) \leq 0$, $H((x_j^0)^+) \geq 0$, the backward characteristic line is from $(x_j^0, 0)$ to $(x_j, \Delta t/2)$ and the speed of this characteristic is $2(x_j - x_j^0)/\Delta t$. Precisely speaking there is no root of $H(s)$ if $H((x_j^0)^-) < 0$ and $H((x_j^0)^+) > 0$, but this means a rarefaction wave emanating from $(x_j^0, 0)$.
3. If $H((x_j^0)^-) \geq 0$, $H((x_j^0)^+) \leq 0$, then there are two roots of $H(s)$, one is in I_j^l and the other one is in I_j^r . Moreover, the line segment from $(x_j^0, 0)$ to $(x_j, \Delta t/2)$ is also a possible characteristic. See Figure 3.1(b). \square

From the previous lemma, we know that there is a unique backward characteristic line unless Case 3 happens. For Case 3, we can use the Lax formula to choose the correct characteristic line among the candidates, if the flux $f(u)$ is convex, see, e.g. the procedure used in [10]. However, if $f(u)$ is nonconvex, it is very difficult to single out the correct characteristic line among these candidates. We will show in next section that any choice among these candidates will maintain the desired accuracy, therefore our main concern is the TVD property of the scheme. We will choose the candidate so that the cell average obtained from (3.3) satisfies the maximum principle, i.e.

$$(3.9) \quad \min_{[y_{j-1}, y_j]} u^0(x) \leq \bar{u}_{j-\frac{1}{2}}^{\frac{1}{2}} \leq \max_{[y_{j-1}, y_j]} u^0(x).$$

Once this maximum principle is satisfied, the conditions for the analysis in Section 2 will be fulfilled and we may have a TVD reconstruction.

Notice that a choice of the correct characteristics for both cell end points y_{j-1} and y_j would produce $\bar{u}_{j-\frac{1}{2}}^{\frac{1}{2}}$ from (3.3) as the exact cell average of the entropy solution $T(u^0)(x, \Delta t/2)$ on $I_{j-\frac{1}{2}}$, therefore it would automatically satisfy the maximum principle (3.9). Thus, in Case 3 when there are more than one candidate for the characteristic line, there is at least one choice which would return a $\bar{u}_{j-\frac{1}{2}}^{\frac{1}{2}}$ satisfying (3.9). If more than one candidate satisfies this criterion, we will simply choose one of them arbitrarily. In our implementation, if there are N intervals in which Case 3 happens, then we have 3^N candidates of possible combination. We check all the 3^N candidates sequentially until we find one choice such that each $\bar{u}_{j-\frac{1}{2}}^{\frac{1}{2}}$ from (3.3) satisfies (3.9), then we will stop the search and use this choice.

After finding the backward characteristics, we can obtain the cell averages and cell boundary point values of $T(u^0)(x, \Delta t/2)$. With this information, we can apply the procedure in the previous section to obtain a piecewise Hermite type reconstruction polynomial $p(x) = \sum_j p_{j-\frac{1}{2}}(x)\chi_{j-\frac{1}{2}}(x)$ on the staggered mesh $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$. To find a TVD approximation to $T(u^0)(x, \Delta t/2)$, we still need to know the ex-

rema of $T(u^0)(x, \Delta t/2)$. In actual implementation, we use the maximum/minimum value of $u^0(x)$ on the interval $[y_{j-1}, y_j]$ to replace the maximum/minimum value of $T(u^0)(x, \Delta t/2)$, and the position of the extremum of $T(u^0)(x, \Delta t/2)$ is determined by forward characteristic lines starting from the maximum/minimum point of $u^0(x)$. For example, if x_j^{\max} is the point where the maximum value of $u^0(x)$ in $[y_{j-1}, y_j]$ is achieved, and the maximum value is u_{\max}^0 , then the forward characteristic is the line through $(x_j^{\max}, 0)$ and $(x_j^{\max} + f'(u_{\max}^0)\Delta t/2, \Delta t/2)$, i.e., we trace the maximum/minimum values by the characteristics. The two cell averages of the left and right parts can be computed in the same way as in (3.3).

REMARK 3.3. *Since the entropy solution to (1.1) satisfies the maximum principle, the maximum/minimum value of $T(u^0)(x, \Delta t/2)$ on the interval $I_{j-\frac{1}{2}}$ is less/greater than or equal to the maximum/minimum value of $u^0(x)$ on the interval $[y_{j-1}, y_j]$, which could be easily computed because u^0 is a piecewise polynomial of degree $k \leq 5$ and there are algebraic analytical formulas for the extrema.*

3.2. Precondition. We propose the following procedure as the preconditioning process to ensure (3.1). Clearly, this preconditioning does not affect accuracy in smooth cells.

1. $r_j(x)$ is a polynomial. If $r_j(x)$ fails to satisfy (3.1), i.e.,

$$\max_{x \in I_j} r_j'(x) f''(r_j(x)) > 2/\Delta t, \quad \text{or} \quad \min_{x \in I_j} r_j'(x) f''(r_j(x)) < -2/\Delta t,$$

then we use $\hat{r}_j(x) = \mu(r_j(x) - \bar{u}_j) + \bar{u}_j$ to replace $r_j(x)$, where \bar{u}_j is the cell average and

$$\mu = \min \left\{ \left| \frac{2}{\Delta t \max_{x \in I_j} f''(r_j(x)) \max_{x \in I_j} r_j'(x)} \right|, \left| \frac{2}{\Delta t \min_{x \in I_j} f''(r_j(x)) \min_{x \in I_j} r_j'(x)} \right| \right\}.$$

2. $r_j(x) = r_j^l(x)\chi(I_j^l) + r_j^r(x)\chi(I_j^r)$. We only discuss $r_j^l(x)$, the process for $r_j^r(x)$ is similar. If $r_j^l(x)$ fails to satisfy (3.1), i.e.,

$$\max_{x \in I_j^l} (r_j^l)'(x) f''(r_j(x)) > 2/\Delta t, \quad \text{or} \quad \min_{x \in I_j^l} (r_j^l)'(x) f''(r_j(x)) < -2/\Delta t,$$

then we use $\hat{r}_j^l(x) = \mu(r_j^l(x) - \bar{u}_j^l) + \bar{u}_j^l$ to replace $r_j^l(x)$, where \bar{u}_j^l is the cell average on I_j^l and

$$\mu = \min \left\{ \left| \frac{2}{\Delta t \max_{x \in I_j^l} f''(r_j^l(x)) \max_{x \in I_j^l} (r_j^l)'(x)} \right|, \left| \frac{2}{\Delta t \min_{x \in I_j^l} f''(r_j^l(x)) \min_{x \in I_j^l} (r_j^l)'(x)} \right| \right\}.$$

3.3. Algorithm flowchart. We can now formulate the algorithm flowchart for the TVD finite volume scheme. From now on we restrict ourself to the fifth order case for easy presentation, although the procedure and results are valid for all orders up to six. Given the cell averages and cell boundary values of the initial data $u_0(x)$, we use the Hermite reconstruction procedure in Section 2 to obtain a piecewise polynomial of degree four $u^0(x)$ as the numerical initial condition. We apply the limiting process in Section 2 to the numerical initial condition, still denoted by $u^0(x)$, which is then a piecewise polynomial of degree four satisfying $TV(u^0) \leq TV(u_0)$.

1. Start with piecewise polynomial of degree four $u^n(x) = \sum_j r_j(x)\chi_j(x)$ at time level n , where $r_j(x)$ is either a polynomial or two-piece polynomials on the interval I_j . Apply the preconditioning process detailed in Section 3.2 to each $r_j(x)$, still denoted as $r_j(x)$.
2. For each j , by Lemmas 3.1 and 3.2, there is either one or three solutions in I_j for the characteristic equation $f'(r_j(x)) = \frac{x_j - x}{\Delta t/2}$, $x \in I_j$. Use either explicit formulas (if $f'(r_j(x))$ is a lower degree polynomial) or an iterative method to find the solution or solutions in I_j . If there is only one solution, set it as y_j . Next, if there are N intervals in which there are three candidates for y_j , then check all the 3^N candidates of combination sequentially until we find one which returns a maximum-principle-satisfying $\bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ for all the intervals in the formula

$$(3.10) \quad \begin{aligned} \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} & \left[\int_{y_{j-1}}^{y_j} u^n(x) dx - \frac{\Delta t}{2} f(u^n(y_j)) + (x_j - y_j) u^n(y_j) \right. \\ & \left. + \frac{\Delta t}{2} f(u^n(y_{j-1})) - (x_{j-1} - y_{j-1}) u^n(y_{j-1}) \right]. \end{aligned}$$

3. Evaluate $T(u^n)(x_j, \frac{\Delta t}{2}) = u^n(y_j)$ and the average of $T(u^n)(x, \frac{\Delta t}{2})$ on $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$ by the formula (3.10). Construct the Hermite type reconstruction polynomials $p_{j-\frac{1}{2}}(x)$ on the staggered interval $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$ using the formulas in Section 2.
4. Let u_j denote $u^n(y_j)$. Evaluate the extrema of $p_{j-\frac{1}{2}}(x)$ in each staggered interval $I_{j-\frac{1}{2}}$ and the extrema of $r_j(x)$ in each I_j to calculate $Var(p_{j-\frac{1}{2}}) = \int_{x_{j-1}}^{x_j} |p'_{j-\frac{1}{2}}(x)| dx + |p_{j-\frac{1}{2}}(x_{j-1}) - u_{j-1}| + |p_{j-\frac{1}{2}}(x_j) - u_j|$ and the standard variation of $u^n(x)$ on $[y_{j-1}, y_j]$, $TV(u^n)_{[y_{j-1}, y_j]} = \int_{y_{j-1}}^{y_j} |(u^n)'(x)| dx$. If $Var(p_{j-\frac{1}{2}}) > TV(u^n)_{[y_{j-1}, y_j]}$, we perform the TVD limiting process described in Section 2 on $p_{j-\frac{1}{2}}(x)$:
 - If $u^n(x)$ is monotone on $[y_{j-1}, y_j]$, then perform the limiting procedure

$$\tilde{p}_{j-\frac{1}{2}}(x) = p_{j-\frac{1}{2}}(x) - \alpha_{j-\frac{1}{2}}(x - x_{j-\frac{1}{2}}),$$

$$\text{where } \alpha_{j-\frac{1}{2}} = \begin{cases} \min_{x \in I_{j-\frac{1}{2}}} p'_{j-\frac{1}{2}}(x), & \text{if } u_{j-1} \leq u_j \\ \max_{x \in I_{j-\frac{1}{2}}} p'_{j-\frac{1}{2}}(x), & \text{if } u_{j-1} \geq u_j \end{cases}, \text{ and}$$

$$r_{j-\frac{1}{2}}(x) = \theta_{j-\frac{1}{2}} \left(\tilde{p}_{j-\frac{1}{2}}(x) - \bar{u}_{j-\frac{1}{2}}^{n+1} \right) + \bar{u}_{j-\frac{1}{2}}^{n+1},$$

$$\text{with } \theta_{j-\frac{1}{2}} = \min \left\{ \left| \frac{u_{j-1} - \bar{u}_{j-\frac{1}{2}}^{n+1}}{\tilde{p}_{j-\frac{1}{2}}(x_{j-1}) - \bar{u}_{j-\frac{1}{2}}^{n+1}} \right|, \left| \frac{u_j - \bar{u}_{j-\frac{1}{2}}^{n+1}}{\tilde{p}_{j-\frac{1}{2}}(x_j) - \bar{u}_{j-\frac{1}{2}}^{n+1}} \right|, 1 \right\}.$$

- If $u^n(x)$ is not monotone on $[y_{j-1}, y_j]$, then choose either the maximum or the minimum of $u^n(x)$, denoted as $y_j^{ext} \in [y_{j-1}, y_j]$. Set $x_j^{ext} = y_j^{ext} + f'(u^n(y_j^{ext}))\Delta t/2$ and $u_j^{ext} = u^n(y_j^{ext})$. Divide $I_{j-\frac{1}{2}}$ into two parts $I_{j-\frac{1}{2}}^l = [x_{j-1}, x_j^{ext}]$ and $I_{j-\frac{1}{2}}^r = [x_j^{ext}, x_j]$. Calculate the averages

of $T(u^n)(x, \frac{\Delta t}{2})$ over $I_{j-\frac{1}{2}}^l$ and $I_{j-\frac{1}{2}}^r$ by

$$\begin{aligned}\bar{u}_{j-\frac{1}{2}}^l &= \frac{1}{\Delta x} \left[\int_{y_{j-1}}^{y_j^{ext}} u^n(x) dx - \frac{\Delta t}{2} f(u^n(y_j^{ext})) + (x_j^{ext} - y_j^{ext}) u^n(y_j^{ext}) \right. \\ &\quad \left. + \frac{\Delta t}{2} f(u^n(y_{j-1})) - (x_{j-1} - y_{j-1}) u^n(y_{j-1}) \right] \\ \bar{u}_{j-\frac{1}{2}}^r &= \frac{1}{\Delta x} \left[\int_{y_j^{ext}}^{y_j} u^n(x) dx - \frac{\Delta t}{2} f(u^n(y_j)) + (x_j - y_j) u^n(y_j) \right. \\ &\quad \left. + \frac{\Delta t}{2} f(u^n(y_j^{ext})) - (x_j^{ext} - y_j^{ext}) u^n(y_j^{ext}) \right].\end{aligned}$$

Set $p_{j-\frac{1}{2}}^l(x) = p_{j-\frac{1}{2}}(x)$ and $p_{j-\frac{1}{2}}^r(x) = p_{j-\frac{1}{2}}(x)$. Let $\bar{p}_{j-\frac{1}{2}}^l$ denote the average of $p_{j-\frac{1}{2}}^l(x)$ on $I_{j-\frac{1}{2}}^l$. Do the following limiting procedure to $p_{j-\frac{1}{2}}^l(x)$:

- (a) If $(u_{j-1} - \bar{u}_{j-\frac{1}{2}}^l) (\bar{u}_{j-\frac{1}{2}}^l - u_j^{ext}) < 0$, set $r_{j-\frac{1}{2}}^l(x) = \bar{u}_{j-\frac{1}{2}}^l$;
- (b) Else, if $p_{j-\frac{1}{2}}^l(x)$ is monotone and $(p_{j-\frac{1}{2}}^l(x_{j-1}) - p_{j-\frac{1}{2}}^l(x_j^{ext})) (u_{j-1} - u_j^{ext}) < 0$, then set $r_{j-\frac{1}{2}}^l(x) = \bar{u}_{j-\frac{1}{2}}^l$;
- (c) Else, if $p_{j-\frac{1}{2}}^l(x)$ is monotone and $(p_{j-\frac{1}{2}}^l(x_{j-1}) - p_{j-\frac{1}{2}}^l(x_j^{ext})) (u_{j-1} - u_j^{ext}) \geq 0$, then set $\tilde{p}_{j-\frac{1}{2}}^l(x) = p_{j-\frac{1}{2}}^l(x) - \bar{p}_{j-\frac{1}{2}}^l + \bar{u}_{j-\frac{1}{2}}^l$, and $r_{j-\frac{1}{2}}^l(x) = \theta_{j-\frac{1}{2}} (\tilde{p}_{j-\frac{1}{2}}^l(x) - \bar{u}_{j-\frac{1}{2}}^l) + \bar{u}_{j-\frac{1}{2}}^l$, where

$$\theta_{j-\frac{1}{2}} = \min \left\{ \left| \frac{u_{j-1} - \bar{u}_{j-\frac{1}{2}}^l}{\bar{p}_{j-\frac{1}{2}}^l(x_{j-1}) - \bar{u}_{j-\frac{1}{2}}^l} \right|, \left| \frac{u_j^{ext} - \bar{u}_{j-\frac{1}{2}}^l}{\tilde{p}_{j-\frac{1}{2}}^l(x_j^{ext}) - \bar{u}_{j-\frac{1}{2}}^l} \right|, 1 \right\};$$

- (d) Else, if $p_{j-\frac{1}{2}}^l(x)$ is not monotone, set

$$\tilde{p}_{j-\frac{1}{2}}^l(x) = p_{j-\frac{1}{2}}^l(x) - \bar{p}_{j-\frac{1}{2}}^l + \bar{u}_{j-\frac{1}{2}}^l - \alpha_{j-\frac{1}{2}} \left(x - \frac{x_{j-1} + x_j^{ext}}{2} \right),$$

$$\text{with } \alpha_{j-\frac{1}{2}} = \begin{cases} \min_{x \in I_{j-\frac{1}{2}}^l} (p_{j-\frac{1}{2}}^l)'(x), & \text{if } u_{j-1} \leq u_j^{ext} \\ \max_{x \in I_{j-\frac{1}{2}}^l} (p_{j-\frac{1}{2}}^l)'(x), & \text{if } u_{j-1} \geq u_j^{ext} \end{cases}, \text{ and}$$

$$r_{j-\frac{1}{2}}^l(x) = \theta_{j-\frac{1}{2}} (\tilde{p}_{j-\frac{1}{2}}^l(x) - \bar{u}_{j-\frac{1}{2}}^l) + \bar{u}_{j-\frac{1}{2}}^l,$$

$$\text{where } \theta_{j-\frac{1}{2}} = \min \left\{ \left| \frac{u_{j-1} - \bar{u}_{j-\frac{1}{2}}^l}{\bar{p}_{j-\frac{1}{2}}^l(x_{j-1}) - \bar{u}_{j-\frac{1}{2}}^l} \right|, \left| \frac{u_j^{ext} - \bar{u}_{j-\frac{1}{2}}^l}{\tilde{p}_{j-\frac{1}{2}}^l(x_j^{ext}) - \bar{u}_{j-\frac{1}{2}}^l} \right|, 1 \right\}.$$

Perform a similar modification to $p_{j-\frac{1}{2}}^r(x)$, then we obtain $r_{j-\frac{1}{2}}(x) = r_{j-\frac{1}{2}}^l(x) \chi(I_{j-\frac{1}{2}}^l) + r_{j-\frac{1}{2}}^r(x) \chi(I_{j-\frac{1}{2}}^r)$.

This finishes the evolution to $t^n + \frac{\Delta t}{2}$ on the staggered mesh $I_{j-\frac{1}{2}}$.

5. Evolve to $t^n + \Delta t$ in an analogous way back to the mesh I_j .

4. Properties of the scheme.

4.1. Conservative form. THEOREM 4.1. *The scheme given in the previous section can be written in a conservative form.*

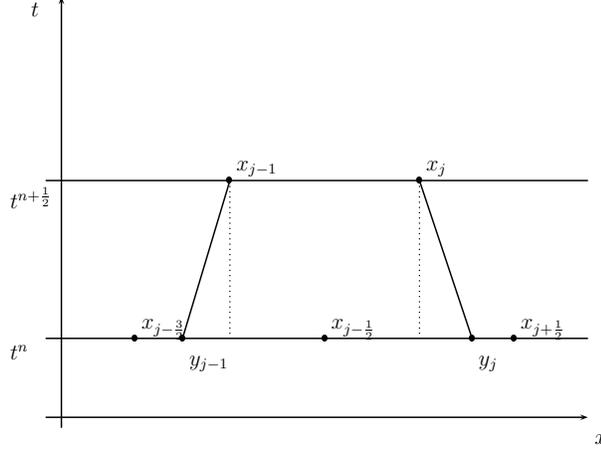


FIG. 4.1. The trapezoid region in the x - t plane.

Proof. The cell averages can be rewritten as

$$\begin{aligned}\bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}} &= \frac{1}{\Delta x} \left[\int_{y_{j-1}}^{y_j} u^n(x) dx - \frac{\Delta t}{2} f(u_j) + (x_j - y_j)u_j + \frac{\Delta t}{2} f(u_{j-1}) - (x_{j-1} - y_{j-1})u_{j-1} \right] \\ &= \bar{u}_{j-\frac{1}{2}}^n - \frac{1}{2} \frac{\Delta t}{\Delta x} (\hat{f}_j - \hat{f}_{j-1}),\end{aligned}$$

where the numerical flux is $\hat{f}_j = \frac{2}{\Delta t} \int_{y_j}^{x_j} u^n(x) dx + \left(f(u_j) - \frac{2(x_j - y_j)}{\Delta t} u_j \right)$ and $\bar{u}_{j-\frac{1}{2}}^n$ denotes the average of $u^n(x)$ on $I_{j-\frac{1}{2}}$. See Figure 4.1.

First, \hat{f}_j only depends on $u^n(x)$ over the interval I_j . Second, \hat{f}_j is consistent in the sense that $\hat{f}_j = f(\bar{u})$ if $u^n(x)$ is a constant \bar{u} . Thus, our scheme is conservative. \square

4.2. Total-variation diminishing. THEOREM 4.2. *The scheme given in the previous section is TVD: $TV(u^{n+\frac{1}{2}}(x)) \leq TV(u^n(x))$.*

Proof. As long as the cell averages satisfy the maximum principle (3.9), Lemma 2.2 in the TVD reconstruction section will hold, which ensures the TVD property of the reconstruction after limiting. The preconditioning process in Section 3.2 clear does not increase the variation. Therefore, the scheme satisfies

$$TV(u^{n+\frac{1}{2}}(x)) \leq Var(u^{n+\frac{1}{2}}(x)) \leq TV(u^n(x)). \quad \square$$

4.3. Accuracy. Except for the Case 3 in Lemma 3.2, the time evolution of our scheme is exactly the same as [12]. Therefore, following the same lines as in [12], if Case 3 never happens, then we can show that our scheme is fifth order accurate for smooth solutions away from the extrema by calculating the local truncation error defined in [12], and it will lose at most one order of accuracy near extrema. Since there is only finitely many such extrema, the L^1 order of accuracy is optimal. Here, we need to show that the local truncation error will lose at most one order of accuracy if Case 3 happens. It suffices to check the accuracy of the backward characteristic line that we choose in Case 3.

LEMMA 4.3. *Assume $f''(x)$ and $u'_0(x)$ are both bounded, then all the three candidates of the backward characteristic lines in Case 3 are fifth order accurate if the mesh is fine enough.*

Proof. Assume the three possible locations of the foot of backward characteristics are s_0 , s_1 and s_2 as shown in Figure 3.1(b) (in which s_0 is denoted as x_j^0). It suffices to show that $s_1 - s_2 = O(\Delta x^5)$ if Δx and Δt are sufficiently small.

Define $G(x) = f'(u_0(x))\Delta t + 2(x - x_j)$ and $H(x) = f'(u^0(x))\Delta t + 2(x - x_j)$. We have

$$G(x) - H(x) = f'(u_0(x))\Delta t - f'(u^0(x))\Delta t = f''(\zeta)(u_0(x) - u^0(x))\Delta t = O(\Delta x^5),$$

therefore $|H(s_0^-) - H(s_0^+)| \leq |H(s_0^-) - G(s_0)| + |G(s_0) - H(s_0^+)| = O(\Delta x^5)$, hence $H(s_0^-)H(s_0^+) < 0$ implies $|H(s_0^-)| = O(\Delta x^5)$. Notice that

$$G'(x) = f''(u_0(x))u_0'(x)\Delta t + 2,$$

hence we have $G'(x) \geq \frac{3}{2}$ for all $x \in [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ if Δt is small enough, since $f''(x)$ and $u_0'(x)$ are bounded. Now $G'(x) - H'(x) = O(\Delta x^4)$ implies that if Δx is small, $H'(x) \geq 1$ for all $x \in [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. Therefore

$$\frac{H(s_0^-) - H(s_1)}{s_0 - s_1} = H'(\xi) \geq 1, \quad \text{for some } \xi \in [s_1, s_0],$$

which implies $s_0 - s_1 \leq H(s_0^-) - H(s_1) = H(s_0^-) = O(\Delta x^5)$. Similarly $s_0 - s_2 = O(\Delta x^5)$. Therefore, the accuracy is not destroyed even if the entropic characteristic line is not necessarily chosen in Case 3. \square

We now have the following theorem on the accuracy of our scheme.

THEOREM 4.4. *Assuming the initial data $u_0(x)$ and the flux $f(u)$ are both smooth functions. Take Δt sufficiently small so that (3.1) is satisfied. The scheme is fifth order accurate away from the extrema of $u_0(x)$ and when Case 3 does not happen. Accuracy can lose at most one order near the extrema or when Case 3 happens, hence in L^1 the error is fifth order accurate.*

5. Numerical test. In this section we provide numerical examples to test our schemes.

5.1. Standard test cases. Example 1. We solve the model equation $u_t + u_x = 0$, $-1 \leq x \leq 1$, $u(x, 0) = u_0(x)$, with periodic boundary conditions.

Three initial data $u_0(x)$ are used. The first one is $u_0(x) = \sin(\pi x)$, and the second one is $u_0(x) = \sin^4(\pi x)$. We list the L^1 and L^∞ errors for the cell averages at time $t = 5$ in Table 5.1. Here and below, the mesh size $\Delta x = 2/N$. We can clearly see that the designed fifth order accuracy is achieved in both cases, at least for the L^1 error.

TABLE 5.1
 $t = 5$, $\Delta t/\Delta x = 0.95$.

N	$u_0(x) = \sin(\pi x)$				$u_0(x) = \sin^4(\pi x)$			
	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
20	1.08E-6	-	2.18E-6	-	6.25E-4	-	1.22E-3	-
40	3.34E-8	5.01	6.93E-8	4.97	2.72E-5	4.52	7.58E-5	4.01
80	1.02E-9	5.03	2.17E-9	5.00	8.46E-7	5.00	3.15E-6	4.59
160	3.14E-11	5.02	6.69E-11	5.02	2.65E-8	4.99	1.20E-7	5.13
320					8.57E-10	4.95	4.08E-9	4.47

The third initial function is $u_0(x) = \begin{cases} 1, & -1 \leq x \leq 0 \\ -1, & 0 \leq x \leq 1 \end{cases}$. The results at $t = 100$ are shown in Figure 5.2(a). We can see that numerical solution maintains a strict

maximum principle and has relatively good resolution for the discontinuity for a coarse mesh after a very long time simulation (50 time periods).

Example 2. We solve the Burgers equation with periodic boundary conditions $u_t + (u^2/2)_x = 0, -1 \leq x \leq 1, u(x, 0) = u_0(x)$. For the initial data $u_0(x) = 0.25 + 0.5 \sin(\pi x)$, the exact solution is smooth up to $t = \frac{2}{\pi}$, then it develops a moving shock which interacts with a rarefaction wave. We list the errors in Table 5.2 at $t = 0.15$. We can clearly see the designed fifth order accuracy is achieved in the L^1 norm. In Figure 5.1 we can see that the shock is captured very well at $t = \frac{2}{\pi}$ and $t = 2.0$. The errors 0.05 away from the shock (i.e., $|x - \text{shock location}| \geq 0.05$) are listed in Table 5.2 at $t = 2.0$. We can see that the designed order of accuracy is achieved or surpassed.

TABLE 5.2
 $u_0(x) = 0.25 + 0.5 \sin(\pi x), \Delta t/\Delta x = 1.2$.

N	$t = 0.15$				$t = 2$			
	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
20	1.32E-6	–	6.73E-6	–	2.19E-3	–	4.09E-2	–
40	3.26E-8	5.34	1.47E-7	5.51	3.30E-5	6.05	7.34E-4	5.80
80	6.41E-10	5.66	4.25E-9	5.12	2.68E-6	6.95	1.79E-5	5.36
160	2.30E-11	4.80	2.38E-10	4.16	8.23E-10	8.34	9.24E-8	9.24
320	7.63E-13	4.91	1.59E-11	3.90	2.70E-13	11.57	9.30E-12	13.27
640	2.46E-14	4.95	1.02E-12	3.95				

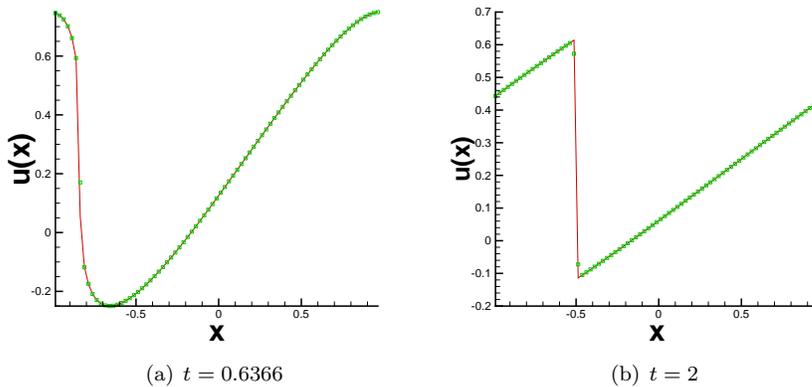


FIG. 5.1. $N = 80, \Delta t/\Delta x = 1.2$. Solid line: exact solution; Symbols: numerical solution.

Example 3. Note that we have only proved TVD for our schemes, not entropy conditions. We use nonconvex fluxes to test the convergence to the physically correct entropy solutions. The “exact” solutions are obtained from the first order Lax-Friedrichs scheme on a very fine mesh.

The first flux is the Buckley-Leverett flux $f(u) = \frac{4u^2}{4u^2 + (1-u)^2}$, with the initial data $u = 1$ in $[-\frac{1}{2}, 0]$ and $u = 0$ elsewhere. The computational result is displayed in Figure 5.2(b), which is quite satisfactory.

$$\text{The second flux is } f(u) = \begin{cases} 1, & \text{if } u < 1.6 \\ \cos(5\pi(u - 1.8)) + 2.0, & \text{if } 1.6 \leq u < 2.0 \\ -\cos(5\pi(u - 2.2)), & \text{if } 2.0 \leq u < 2.4 \\ 1, & \text{if } u \geq 2.4 \end{cases} \quad \text{with}$$

two initial conditions, which is an example used in [11]. The first initial condition is $u_0(x) = \begin{cases} 1, & \text{for } x < 0 \\ 3, & \text{for } x \geq 0 \end{cases}$ and it is shown in [11] that the numerical solutions of many high order schemes would stay stationary which is entropy-violating. Our result is shown in Figure 5.2(c) (solid line is exact solution and symbols denote numerical solution (cell averages)) which approximates the exact entropy solution very well. The other initial data that we test is $u_0(x) = \begin{cases} 3, & \text{for } -1 \leq x < 0 \\ 1, & \text{for } 0 \leq x \leq 1 \end{cases}$ with a periodic boundary condition. It is shown in [11] that convergence towards the entropy solution for this test case is slow for first order monotone schemes and may fail for many high order schemes. Our results are shown in Figure 5.2(d). There is clearly convergence with refined meshes, and the rate of convergence is faster than that of the first order schemes shown in [11].

5.2. Test cases from traffic flow models. In the subsection, we test our fifth order TVD scheme on two traffic flow problems. To describe the dynamic characteristics of traffic on a homogeneous and unidirectional highway, the Lighthill-Whitham-Richards (LWR) model is widely used. The equation for the LWR model is $\rho_t + q(\rho)_x = 0$ with suitable initial and boundary conditions. Here $\rho \in (0, \rho_{\max})$ is the density, ρ_{\max} is the maximum (jam) density, and $q(\rho) = u(\rho)\rho$ is the traffic flow on a homogeneous highway.

Example 4. The first traffic flow test example is taken from [7]. The flow-density function is given by a concave function

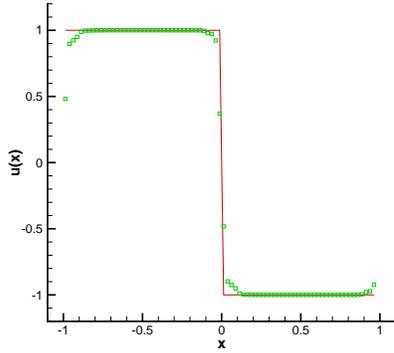
$$q(\rho) = \begin{cases} -0.4\rho^2 + 100\rho, & \rho \in [0, 50] \\ -0.1\rho^2 + 15\rho + 3500, & \rho \in [50, 100] \\ -0.024\rho^2 - 5.2\rho + 4760, & \rho \in [100, 350] \end{cases}.$$

The length of the freeway is 20 km. The entrance density is 50 veh/km. The piecewise linear initial density profile shown in Fig 5.3(a) is formed. The entrance is blocked for 10 min, after which traffic is released again from the entrance at the capacity density 75 veh/km. After 20 min, the entrance flow returns to 50 veh/h. At the exit boundary, a traffic signal is installed, with a repeated pattern of 2 min green light (zero density) followed by 1 min red light (jam density). The numerical solutions are shown in Figures 5.3(b), 5.3(c) and 5.3(d). We can observe that our TVD scheme produces very good approximations to the exact solution for this test case.

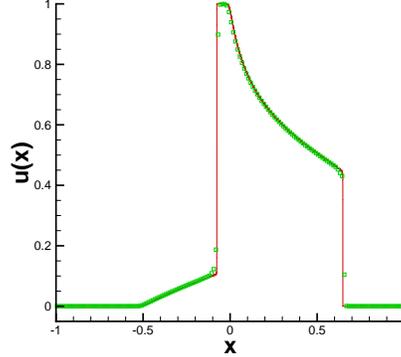
Example 5. We consider a similar problem but with a much more complicated flow-density function in [3]. The flow function $q(\rho) = \rho V_e(\rho)$ is given by

$$V_e(\rho) = \frac{\tilde{V}^2}{2V_0} \left(-1 + \sqrt{1 + \frac{4V_0^2}{\tilde{V}^2}} \right)$$

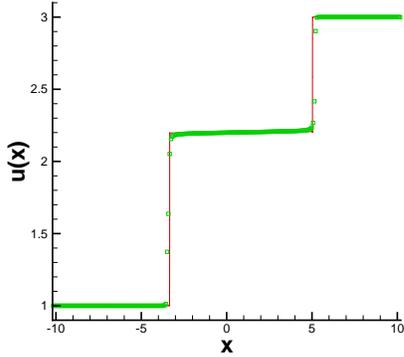
with $\tilde{V}(\rho) = \frac{1}{T_r} \left(\frac{1}{\rho} - \frac{1}{\rho_{\max}} \right) \sqrt{\frac{\alpha(\rho_{\max})}{\alpha(\rho)}}$ and $\alpha(\rho) = \alpha_0 + \Delta\alpha \left(\tanh\left(\frac{\rho - \rho_c}{\Delta\rho}\right) + 1 \right)$. Here V_0 , T_r , ρ_{\max} , α_0 , $\Delta\alpha$, ρ_c and $\Delta\rho$ are all constant parameters to be determined by fitting them to the empirical data. The physical meaning of these parameters can be found in [3]. We simply choose some typical values mentioned in [3]: $V_0 = 110$ km/h, $T_r = 1.8$ seconds, $\rho_{\max} = 160$ vehicles/km, $\alpha_0 = 0.008$, $\Delta\alpha = 0.02$, $\rho_c =$



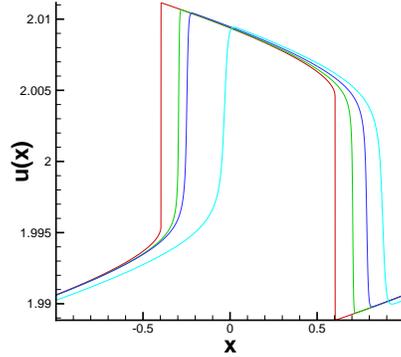
(a) $N = 40, t = 100.$



(b) $N = 160, t = 0.4, \Delta t/\Delta x = 0.3.$



(c) $N = 400, t = 2, \Delta t/\Delta x = 0.04.$

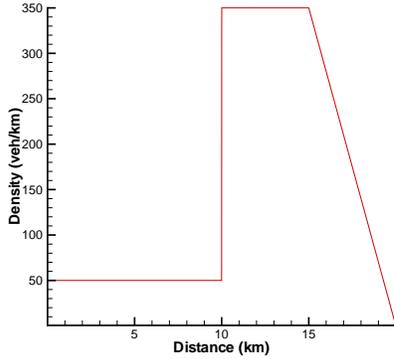


(d) $t = 2, \Delta t/\Delta x = 0.04.$ The curves, from right to left, are corresponding to $N = 800, N = 1600, N = 3200$ and the “exact” solution, respectively.

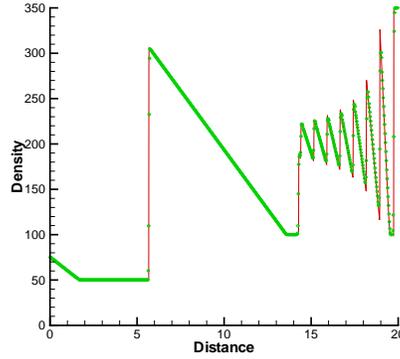
FIG. 5.2. *Example 1 and Example 3. In Fig 5.2(a), 5.2(b) and 5.2(c), solid line is exact solution and squares denote numerical solution (cell averages).*

$0.27\rho_{\max}$ and $\Delta\rho = 0.1\rho_{\max}$. With all these parameters, the flow-density function $q(\rho)$ is well-defined, the graphs of this function and its second derivative are plotted in Figure 5.4(a) and 5.4(b). It is clearly neither a globally concave nor a globally convex function. The entrance density is constant 30 veh/km. The initial condition is $\rho_0(x) = \frac{135}{2} \sin\left(\frac{\pi}{10}x\right) + \frac{145}{2}$. At the exit boundary, a traffic signal is installed, with a repeated pattern of 1 min green light ($\rho = 10$ veh/km) followed by 2 min red light ($\rho = 140$ veh/km). The numerical solutions are shown in Figure 5.4(c), with a magnified graph for the boxed region shown in Figure 5.4(d), where the solid line is the reference solution obtained by the first order Lax-Friedrichs scheme on a very fine grid ($N = 4000000$) at $t = 18$ min. We again observe very good resolution of our scheme for this nonconvex traffic flow model.

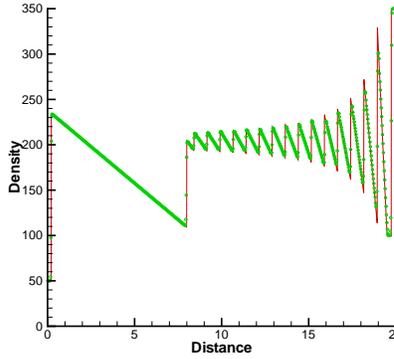
REMARK 5.1. *The advantage of our scheme is that it is high order accurate and satisfies strictly a maximum principle, therefore it has better resolution than the usual*



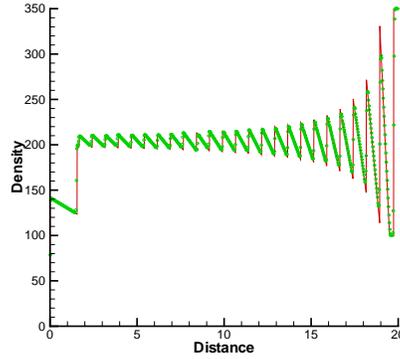
(a) The initial density.



(b) $t = 30$ min.



(c) $t = 60$ min.



(d) $t = 90$ min.

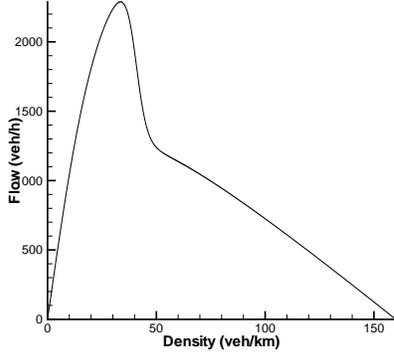
FIG. 5.3. Traffic flow Example 4: $N = 800$. Solid line: exact solution; Circles: numerical solution (cell averages).

TVD schemes when the solution contains many waves, and it does not generate any unphysical solution such as negative density. From the two examples above, we can see that all the waves (shocks and rarefactions) are well captured.

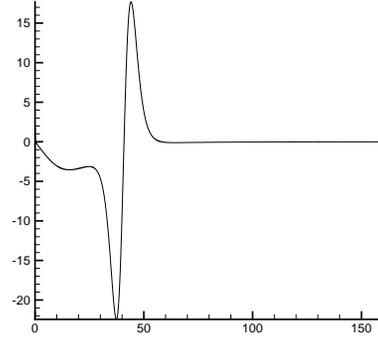
REMARK 5.2. *The boundary conditions in these two examples are all piecewise constants in time. Hence we simply use constant values on ghost cells as the numerical boundary condition for our scheme.*

5.3. A simplified scheme. In this subsection we discussed a simplified version of our finite volume scheme. We use the same method for time evolution, and a similar but simpler limiter which only enforces strict maximum principle but not TVD, without breaking a polynomial into two pieces on any interval. The algorithm satisfies all the theoretical properties of the previous scheme except for a rigorous proof of TVD. This maximum-principle-satisfying finite volume scheme is described below.

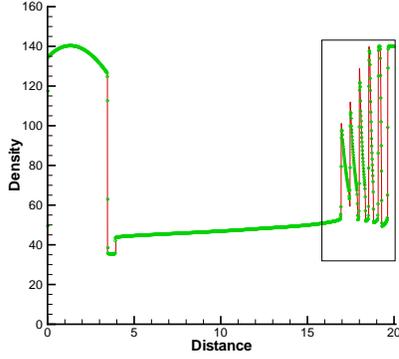
1. Start with the preconditioned version of piecewise polynomial of degree four $u^n(x) = \sum_j r_j(x)\chi_j(x)$ at time level n .



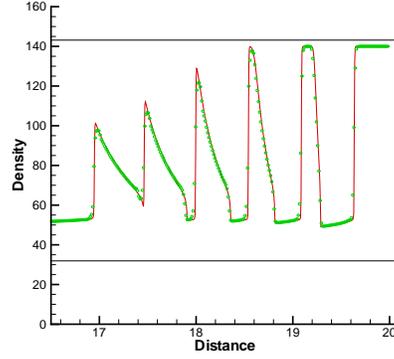
(a) the graph of $q(\rho)$.



(b) the graph of $q''(\rho)$.



(c) $t = 18$ min. The region inside the rectangle on the right is magnified in Figure 5.4(d).



(d) Magnified graph of the region inside the rectangle in Figure 5.4(c).

FIG. 5.4. Traffic flow Example 5: $N = 1600$. Solid line: exact solution; Circles: numerical solution (cell averages).

2. For each x_j , find the unique root y_j of the characteristic equation (3.4) in I_j . The uniqueness of the root is ensured by Lemma 3.1.
3. Evaluate $T(u^n)(x_j, \frac{\Delta t}{2}) = u^n(y_j)$ and the average $\bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ of $T(u^n)(x, \frac{\Delta t}{2})$ on $I_{j-\frac{1}{2}}$ by the formula (3.3). Construct the Hermite type interpolation polynomials $p_{j-\frac{1}{2}}(x)$ on the staggered interval $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$.
4. For each interval $I_{j-\frac{1}{2}}$, evaluate the maximum M_j and the minimum m_j of $p_{j-\frac{1}{2}}(x)$ on $I_{j-\frac{1}{2}}$, and the maximum \widetilde{M}_j and the minimum \widetilde{m}_j of $u^n(x)$ on $[y_{j-1}, y_j]$. Apply the following scaling $r_{j-\frac{1}{2}}(x) = \theta(p_{j-\frac{1}{2}}(x) - \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}) + \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ where θ is determined by $\theta = \min \left\{ \frac{\widetilde{M}_j - \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}}{M_j - \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}}, \frac{\widetilde{m}_j - \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}}{m_j - \bar{u}_{j-\frac{1}{2}}^{n+\frac{1}{2}}}, 1 \right\}$.
5. Apply the preconditioning process. This finishes the evolution to $t^n + \frac{\Delta t}{2}$ on the staggered mesh $I_{j-\frac{1}{2}}$.

6. Evolve to $t^n + \Delta t$ in an analogous way back to the mesh I_j .

REMARK 5.3. *We can easily prove that this scheme is conservative and fifth order accurate. Obviously it satisfies the maximum principle. Therefore we will refer to it as the maximum-principle-satisfying scheme.*

REMARK 5.4. *The maximum-principle-satisfying scheme is much easier to code than the TVD scheme. Although we cannot rigorously prove TVD or TVB of the numerical solution, we have not observed any significant difference between this scheme and the TVD scheme tested before, for all the test cases reported in this paper. We will not show these results here to save space.*

6. Concluding remarks. In this paper we have extended the work in [12] and have constructed a class of genuinely high order accurate finite volume TVD schemes for solving one dimensional scalar conservation laws. These schemes do not degenerate to lower order accuracy for solutions with smooth extrema, yet they satisfy a strict maximum principle and they can be proved to be TVD, when the total variation is measured by the bounded variation semi-norm of the reconstructed piecewise polynomials. The key ingredient of the algorithm is the TVD reconstruction, which can be efficiently implemented for order of accuracy up to six.

We have tested the fifth order scheme on a variety of examples including those from traffic flow models and those with non-convex fluxes. The solutions are high order accurate and provide good resolution to shocks and rarefaction waves.

A simplified scheme is also described, which enforces a strict maximum principle but is not rigorously TVD, however it is much simpler to implement than the TVD scheme. Numerical experiments indicate that this simplified scheme perform as nicely in all the test cases as the TVD scheme.

The advantage of the schemes constructed in this paper, compared with traditional TVD schemes, is that they do not degenerate to first order at smooth extrema, hence they give very good resolutions to solutions with complicated smooth waves. On the other hand, the advantage of the schemes constructed in this paper, compared with ENO and WENO schemes, is that they satisfy strictly the maximum principle, hence they will not generate non-physical solutions such as negative density for the traffic flows. This property is important in many applications.

In order to generalize of this scheme to two dimensions, we should abandon the requirement of TVD and insist only on the strict maximum principle, that is, along the approach of the simplified scheme in Section 5.3. Initial work along this direction has been performed in [16], after the submission of the original version of this paper. It is also possible to formally generalize the scheme to hyperbolic systems, see [13] for one possible approach. However, it would be probably more appropriate, for the system case, to enforce certain positivity preserving properties, such as positivity preserving for density and pressure for Euler equations of compressible gas dynamics, see [9] for one possible approach. These generalizations constitute ongoing research.

REFERENCES

- [1] A. Harten, *High resolution schemes for hyperbolic conservation laws*, Journal of Computational Physics, 49 (1983), pp.357-393.
- [2] A. Harten, B. Engquist, S. Osher and S. Chakravarthy, *Uniformly high order essentially non-oscillatory schemes, III*, Journal of Computational Physics, 71 (1987), 231-303.
- [3] D. Helbing, A. Hennecke, V. Shvetsov and M. Treiber, *MASTER: macroscopic traffic simulation based on a gas-kinetic, non-local traffic model*, Transportation Research Part B, 35 (2001), pp.183-211.

- [4] G.-S. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes*, Journal of Computational Physics, 126 (1996), pp.202-228.
- [5] X.-D. Liu and S. Osher, *Non-oscillatory high order accurate self similar maximum principle satisfying shock capturing schemes*, SIAM Journal on Numerical Analysis, 33 (1996), pp.760-779.
- [6] X.-D. Liu, S. Osher and T. Chan, *Weighted essentially non-oscillatory schemes*, Journal of Computational Physics, 115 (1994), 200-212.
- [7] Y. Lu, S.C. Wong, M. Zhang, C.-W. Shu and W. Chen, *Explicit construction of entropy solutions for the Lighthill-Whitham-Richards traffic flow model with a piecewise quadratic flow-density relationship*, Transportation Research Part B, 42 (2008), pp.355-372.
- [8] S. Osher and S. Chakravarthy, *High resolution schemes and the entropy condition*, SIAM Journal on Numerical Analysis, 21 (1984), pp.955-984.
- [9] B. Perthame and C.-W. Shu, *On positivity preserving finite volume schemes for Euler equations*, Numerische Mathematik, 73 (1996), pp.119-130.
- [10] J.-M. Qiu and C.-W. Shu, *Convergence of Godunov-type schemes for scalar conservation laws under large time steps*, SIAM Journal on Numerical Analysis, 46 (2008), pp.2211-2237.
- [11] J.-M. Qiu and C.-W. Shu, *Convergence of high order finite volume weighted essentially non-oscillatory scheme and discontinuous Galerkin method for nonconvex conservation laws*, SIAM Journal on Scientific Computing, 31 (2008), pp.584-607.
- [12] R. Sanders, *A third-order accurate variation nonexpansive difference scheme for single nonlinear conservation law*, Mathematics of Computation, 51 (1988), pp.535-558.
- [13] R. Sanders, *High resolution staggered mesh approach for nonlinear hyperbolic systems of conservation laws*, Journal of Computational Physics, 101 (1992), pp.314-329.
- [14] C.-W. Shu, *TVB uniformly high-order schemes for conservation laws*, Mathematics of Computation, 49 (1987), pp.105-121.
- [15] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, Journal of Computational Physics, 77 (1988), pp.439-471.
- [16] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, Journal of Computational Physics, 229 (2010), pp.3091-3120.

Appendix A. Appendix.

In this appendix we provide a proof of Lemma 2.3.

Proof: We only prove the $k = 4$ case. The proof for lower k is very similar (and simpler) and is thus omitted.

We can assume the interval is $I = [-\frac{1}{2}, \frac{1}{2}]$ by considering the rescaled variable $x' = \frac{x-(a+b)/2}{b-a}$, which leaves the ratios in (2.7) unchanged. We will still denote x' by x and, without loss of generality, we assume $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$ is increasing on $I = [-\frac{1}{2}, \frac{1}{2}]$. Notice that since $p(x)$ is monotone on $I = [-\frac{1}{2}, \frac{1}{2}]$, we only need to show that

$$(A.1) \quad \left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| \leq 7, \quad \left| \frac{p(-\frac{1}{2}) - \bar{p}}{p(\frac{1}{2}) - \bar{p}} \right| \leq 7.$$

We will only prove the first inequality in (A.1), the derivation for the second inequality being the same. We have

$$p(-\frac{1}{2}) = a_0 - \frac{1}{2}a_1 + \frac{1}{4}a_2 - \frac{1}{8}a_3 + \frac{1}{16}a_4, \quad p(\frac{1}{2}) = a_0 + \frac{1}{2}a_1 + \frac{1}{4}a_2 + \frac{1}{8}a_3 + \frac{1}{16}a_4$$

$$\bar{p} = a_0 + \frac{1}{12}a_2 + \frac{1}{80}a_4$$

Thus,

$$\left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| = \frac{p(\frac{1}{2}) - \bar{p}}{-p(-\frac{1}{2}) + \bar{p}} = \frac{\frac{1}{2}a_1 + \frac{1}{6}a_2 + \frac{1}{8}a_3 + \frac{1}{20}a_4}{\frac{1}{2}a_1 - \frac{1}{6}a_2 + \frac{1}{8}a_3 - \frac{1}{20}a_4}$$

$$= 1 + \frac{\frac{2}{3}a_2 + \frac{1}{5}a_4}{a_1 - \frac{1}{3}a_2 + \frac{1}{4}a_3 - \frac{1}{10}a_4} = 1 + \frac{2}{3} \frac{a_2 + \frac{3}{10}a_4}{a_1 - \frac{1}{3}a_2 + \frac{1}{4}a_3 - \frac{1}{10}a_4}.$$

We now discuss this in several cases.

Case 1. If $a_2 + \frac{3}{10}a_4 = 0$, then $\left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| = 1 < 7$.

Case 2. If $a_2 + \frac{3}{10}a_4 < 0$, then $\left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| = 1 + \frac{2}{3} \frac{1}{\frac{a_1 + \frac{1}{4}a_3}{a_2 + \frac{3}{10}a_4} - \frac{1}{3}}$. Since $p(x)$ is increasing,

$$\begin{aligned} & \frac{1}{2}a_1 + \frac{1}{6}a_2 + \frac{1}{8}a_3 + \frac{1}{20}a_4 = p\left(\frac{1}{2}\right) - \bar{p} \geq 0, \\ \implies & \frac{1}{2}(a_1 + \frac{1}{4}a_3) + \frac{1}{6}(a_2 + \frac{3}{10}a_4) \geq 0, \\ \implies & \frac{a_1 + \frac{1}{4}a_3}{a_2 + \frac{3}{10}a_4} \leq -\frac{1}{3}, \\ \implies & \left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| < 1. \end{aligned}$$

Case 3. If $a_2 + \frac{3}{10}a_4 > 0$, then

$$\left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| = 1 + \frac{2}{3} \frac{1}{\frac{a_1 + \frac{1}{4}a_3}{a_2 + \frac{3}{10}a_4} - \frac{1}{3}}.$$

For any $u, v \in [-\frac{1}{2}, \frac{1}{2}]$, we have

$$\begin{aligned} p'(u) &= a_1 + 2ua_2 + 3u^2a_3 + 4u^3a_4 \geq 0, \\ p'(v) &= a_1 + 2va_2 + 3v^2a_3 + 4v^3a_4 \geq 0, \\ \implies p'(u) + p'(v) &= 2a_1 + 2(u+v)a_2 + 3(u^2+v^2)a_3 + 4(u^3+v^3)a_4 \geq 0, \\ \text{(A.2)} \implies 2(a_1 + \frac{3}{2}(u^2+v^2)a_3) &+ 2(u+v)(a_2 + 2(u^2+v^2-uv)a_4) \geq 0, \end{aligned}$$

We would like to find $u, v \in [-\frac{1}{2}, \frac{1}{2}]$ satisfying

$$\begin{cases} \frac{3}{2}(u^2+v^2) &= \frac{1}{4} \\ 2(u^2+v^2-uv) &= \frac{3}{10} \end{cases}.$$

Solving this linear system, we obtain $u = -\frac{1}{2} \left(\sqrt{\frac{1}{5}} + \sqrt{\frac{2}{15}} \right)$ and $v = -\frac{1}{2} \left(\sqrt{\frac{1}{5}} - \sqrt{\frac{2}{15}} \right)$, which are apparently within $[-\frac{1}{2}, \frac{1}{2}]$. Plugging these values into (A.2), we obtain

$$\begin{aligned} & 2 \left(a_1 + \frac{1}{4}a_3 \right) - 2\sqrt{\frac{1}{5}} \left(a_2 + \frac{3}{10}a_4 \right) \geq 0, \\ \implies & \frac{a_1 + \frac{1}{4}a_3}{a_2 + \frac{3}{10}a_4} \geq \sqrt{\frac{1}{5}}, \\ \implies & \left| \frac{p(\frac{1}{2}) - \bar{p}}{p(-\frac{1}{2}) - \bar{p}} \right| \leq 1 + \frac{2}{3} \frac{1}{\sqrt{\frac{1}{5}} - \frac{1}{3}} \approx 6.85 < 7. \square \end{aligned}$$

REMARK A.1. *The result of this lemma also holds for polynomials of higher degree. We refer to Lemma 2.4. in [16] for the existence proof of the constant C . The proof in [16] does not however provide an explicit value of C .*