

# TVB Uniformly High-Order Schemes for Conservation Laws

By Chi-Wang Shu

**Abstract.** In the computation of conservation laws  $u_t + f(u)_x = 0$ , TVD (total-variation-diminishing) schemes have been very successful. But there is a severe disadvantage of all TVD schemes: They must degenerate locally to first-order accuracy at nonsonic critical points. In this paper we describe a procedure to obtain TVB (total-variation-bounded) schemes which are of uniformly high-order accuracy in space including at critical points. Together with a TVD high-order time discretization (discussed in a separate paper), we may have globally high-order in space and time TVB schemes. Numerical examples are provided to illustrate these schemes.

**1. Introduction.** In this paper we consider the numerical solutions to the hyperbolic conservation law

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

$$(1.1b) \quad u(x, 0) = u_0(x).$$

Here  $u = (u_1, \dots, u_s)^T$ , and the Jacobian matrix  $A(u) = \partial f / \partial u$  has  $s$  real eigenvalues

$$\lambda_1(u) \leq \lambda_2(u) \leq \dots \leq \lambda_s(u)$$

and a complete set of eigenvectors.

On a computational grid  $x_j = j\Delta x$ ,  $t_n = n\Delta t$ , we use  $u_j^n$  to denote the computed approximation to the exact solution  $u(x_j, t_n)$  of (1.1).

We restrict our theoretical discussions to the scalar case  $s = 1$ . Extensions to systems can be done following the lines of, e.g., [7], [8], although a convergence theory for general nonlinear systems is still unknown at present.

We only consider pure initial value problems, i.e., we assume  $u_0(x)$  in (1.1b) to be either periodic or to have a compact support. Initial-boundary value problems will be discussed in a separate paper [11].

The main difficulty in solving (1.1) is that the solution may develop discontinuities (shocks, contact discontinuities, etc.), even if the initial condition  $u_0(x)$  in (1.1b) is smooth. For this reason, weak solutions have to be considered and the standard numerical analysis is very hard to apply. In recent years, TVD (total-variation-diminishing) schemes have been constructed and proven to be very successful in solving (1.1); see, e.g., [1], [2], [6], [7], and the references listed there. The total variation of a discrete scalar solution is defined by

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

---

Received July 7, 1986.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65M05, 35L65, 65M10.

©1987 American Mathematical Society  
0025-5718/87 \$1.00 + \$.25 per page

and if

$$(1.3) \quad \text{TV}(u^{n+1}) \leq \text{TV}(u^n)$$

we say the scheme is TVD; while if

$$(1.4) \quad \text{TV}(u^n) \leq B$$

for some fixed  $B > 0$  depending only on  $\text{TV}(u^0)$  and all possible  $n$  and  $\Delta t$  such that  $n\Delta t \leq T$ , we say that the scheme is TVB in  $0 \leq t \leq T$ . Clearly, TVD implies TVB.

One major advantage of TVB schemes is that there is a convergent (in  $L_1^{\text{local}}$ ) subsequence as  $\Delta x \rightarrow 0$  to a weak solution of (1.1). If an additional entropy condition which implies uniqueness of weak solution to (1.1) is satisfied, then the scheme is convergent. See, e.g., [2].

One important disadvantage of all TVD schemes is that they are at most first-order accurate at nonsonic critical points (i.e., points at which  $\partial u(x, t)/\partial x = 0 \neq f'(u(x, t))$ ). This restricts the accuracy of all TVD schemes to be at most first order in the  $L_\infty$ -norm and at most second order in the  $L_1$ -norm for general problems, no matter how accurate they are in monotone smooth regions. See, e.g., [6].

To overcome this difficulty, recently Harten, Engquist, Osher, and Chakravarthy [4], [5], [3] constructed ENO (essentially nonoscillatory) schemes which are of globally high-order accuracy in smooth regions and extremely stable according to their extended analysis and numerical experiments, although a complete convergence theory (e.g., TV boundedness) for such schemes is still unavailable. In this paper we use another approach and present a simple modification of some existing TVD schemes such that the resulting schemes can be proven to be TVB and of globally high-order accuracy in space. The construction procedure is simple, and the schemes can be applied to steady state calculations (e.g., implemented with the TVD Runge-Kutta type time discretizations in [9]) as well as to time-dependent problems (e.g., implemented with the TVD high-order time discretizations in [10]). But since these methods use a fixed stencil, they may smear discontinuities and pollute more than the ENO schemes, which use an adaptive stencil to obtain information from smooth regions.

The formal “order of accuracy” in this paper is in the sense of local truncation errors, i.e., if the local truncation error  $\text{TE} = O(\Delta x^{r+1})$ , we say the scheme is (formally)  $r$ th order accurate (see, e.g., [1]).

We will use the standard notation

$$(1.5) \quad \Delta_+ u_j = u_{j+1} - u_j; \quad \Delta_- u_j = u_j - u_{j-1}.$$

A semidiscrete (method of lines) scheme to (1.1) of conservation form is a system of ODE's,

$$(1.6) \quad \frac{\partial}{\partial t} u_j = -\frac{1}{\Delta x} (\hat{f}_{j+1/2} - \hat{f}_{j-1/2}),$$

where the numerical flux  $\hat{f}_{j+1/2}$  is defined by

$$(1.7) \quad \hat{f}_{j+1/2} = \hat{f}(u_{j-k+1}, \dots, u_{j+k}),$$

which is Lipschitz continuous in all its arguments, and satisfies the consistency condition

$$(1.8) \quad \hat{f}(u, \dots, u) = f(u).$$

We also consider Euler forward time discretization of (1.6),

$$(1.9) \quad u_j^{n+1} = u_j^n - \lambda (\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n).$$

Here  $\lambda = \Delta t / \Delta x$  is called the CFL number.

If a consistent numerical flux  $\hat{f}_{j+1/2}$  satisfies

$$(1.10) \quad \text{sign}(u_{j+1} - u_j)(\hat{f}_{j+1/2} - f(u)) \leq 0$$

for all  $u$  between  $u_j$  and  $u_{j+1}$ , it is said to be an E-flux. See, e.g., [7].

E-schemes (schemes (1.6) and (1.9) with fluxes  $\hat{f}_{j+1/2}$  as E-fluxes) are TVD and entropy-condition-satisfying, so they are convergent. But E-schemes can be at most first-order accurate. In many parts of this paper we use E-fluxes as our building blocks. We will use  $h_{j+1/2}$  to denote an E-flux, and keep  $\hat{f}_{j+1/2}$  as a notation for a general flux. The following are a few useful 3-point E-fluxes.

#### Engquist-Osher

$$(1.11) \quad h_{j+1/2}^{\text{EO}} = \int_0^{u_{j+1}} \min(f'(s), 0) ds + \int_0^{u_j} \max(f'(s), 0) ds + f(0);$$

#### Godunov

$$(1.12) \quad h_{j+1/2}^{\text{G}} = \begin{cases} \min_{u_j \leq u \leq u_{j+1}} f(u), & \text{if } u_j \leq u_{j+1}, \\ \max_{u_j \geq u \geq u_{j+1}} f(u), & \text{if } u_j > u_{j+1}; \end{cases}$$

#### Roe (with entropy fix)

$$(1.13) \quad h_{j+1/2}^{\text{Rf}} = \frac{1}{2} \left[ f(u_{j+1}) + f(u_j) - \left| \frac{\Delta_+ f(u_j)}{\Delta_+ u_j} \right| \Delta_+ u_j \right]$$

unless

$$u_j < \bar{u} < u_{j+1};$$

then take any Lipschitz function so that

$$h_{j+1/2}^{\text{Rf}} \leq f(\bar{u}).$$

Here  $\bar{u}$  is the unique sonic point,  $f'(\bar{u}) = 0$ , and  $f''(u) \geq 0$  for all  $u$ .

#### Lax-Friedrichs

$$(1.14) \quad h_{j+1/2}^{\text{LF}} = \frac{1}{2} [f(u_{j+1}) + f(u_j) - \alpha \Delta_+ u_j]$$

where  $\alpha$  is a fixed positive number satisfying

$$\alpha \geq \max_{\min_j u_j \leq u \leq \max_j u_j} |f'(u)|.$$

The format of this paper is as follows: In Section 2, a TVB modification procedure is applied to the TVD schemes in [7], yielding TVB semidiscrete schemes and their Euler forward-in-time version of uniformly high-order spatial accuracy; in Section 3, the schemes in Section 2 are equipped with the TVD high-order time discretizations in [10], giving globally high-order in space and time TVB schemes; in Section 4, we illustrate the generality of the TVB modification procedure by applying it to another kind of second-order TVD schemes, making them uniformly second-order; in the appendix, several numerical results are included to illustrate the schemes in Section 3.

**2. TVB Semidiscrete Schemes of Uniformly High-Order Accuracy.** We shall consider semidiscrete (method of lines) schemes which are of conservation form

$$(2.1) \quad u_t = -\frac{1}{\Delta x} (\hat{f}_{j+1/2} - \hat{f}_{j-1/2})$$

and their first-order Euler forward time discretization

$$(2.2) \quad u_j^{n+1} = u_j^n - \lambda (\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n).$$

Starting with a three-point E-flux  $h_{j+1/2}$  (1.10), define

$$(2.3a) \quad df_{j+1/2}^- = h_{j+1/2} - f(u_j),$$

$$(2.3b) \quad df_{j+1/2}^+ = f(u_{j+1}) - h_{j+1/2}$$

(upwind-downwind decomposition).

In [7], Osher and Chakravarthy defined the following flux:

$$(2.4) \quad \begin{aligned} \hat{f}_{j+1/2} = h_{j+1/2} + \sum_{k=-m+1}^{m-1} \left[ \mu_k^m + (-1)^k \beta \binom{2m-2}{k+m-1} \right] (df_{j+k+1/2}^-)^{(k)} \\ + \sum_{k=-m+1}^{m-1} \left[ \mu_k^m - (-1)^k \beta \binom{2m-2}{k+m-1} \right] (df_{j+k+1/2}^+)^{(k)}. \end{aligned}$$

Here  $m \geq 2$  is an integer,  $0 < \beta < (m \binom{2m}{m})^{-1}$ ,

$$\binom{A}{B} = \frac{A!}{B!(A-B)!}$$

is the usual binomial coefficient, and

$$(2.5a) \quad \nu_{m-1}^m = (-1)^{m-1} \left( m \binom{2m}{m} \right)^{-1} \quad \text{for } m \geq 2,$$

$$(2.5b) \quad \nu_0^m = \frac{1}{2},$$

$$(2.5c) \quad \nu_k^m = -\nu_{-k}^m, \quad k = 1, \dots, m-1,$$

$$(2.5d) \quad \nu_k^{m+1} = \nu_k^m + (-1)^k \left( (m+1) \binom{2m+2}{m+1} \right)^{-1} \frac{k}{m} \binom{2m}{m-k} \\ \text{for } k = 1, 2, \dots, m-1,$$

$$(2.5e) \quad \mu_k^m = \nu_k^m \quad \text{for } k = \pm 1, \pm 2, \dots, \pm(m-1),$$

$$(2.5f) \quad \mu_0^m = -\frac{1}{2}.$$

Also,

$$(2.6a) \quad (df_{j+1/2}^-)^{(k)} = m(df_{j+1/2}^-, bdf_{j-k+1/2}^-, bdf_{j-k+3/2}^-) \\ \text{for all } k \text{ with } 0 \neq k \neq 1,$$

$$(2.6b) \quad (df_{j+1/2}^-)^{(0)} = m(df_{j+1/2}^-, bdf_{j+3/2}^-),$$

$$(2.6c) \quad (df_{j+1/2}^-)^{(1)} = m(df_{j+1/2}^-, bdf_{j-1/2}^-),$$

$$(2.6d) \quad (df_{j+1/2}^+)^{(k)} = m(df_{j+1/2}^+, bdf_{j-k+1/2}^+, bdf_{j-k-1/2}^+) \\ \text{for all } k \text{ with } 0 \neq k \neq -1,$$

$$(2.6e) \quad (df_{j+1/2}^+)^{(0)} = m(df_{j+1/2}^+, bdf_{j-1/2}^+),$$

$$(2.6f) \quad (df_{j+1/2}^+)^{(-1)} = m(df_{j+1/2}^+, bdf_{j+3/2}^+).$$

The parameter  $b$  is in the range

$$(2.7) \quad 0 < b \leq \left( \sum_{j=2}^m \frac{1}{2j-1} \right)^{-1} \left( 1 + 2\beta \binom{2m-2}{m-1} \right)$$

and the minmod function  $m$  is defined by

$$(2.8) \quad m(a_1, \dots, a_l) = \begin{cases} s \cdot \min(|a_1|, \dots, |a_l|), & \text{if } \text{sign}(a_1) = \dots = \text{sign}(a_l) = s, \\ 0, & \text{otherwise.} \end{cases}$$

They proved the following

- THEOREM 2.1** (Osher and Chakravarthy). (a) *The scheme (2.1), (2.3), (2.7) is TVD;*  
 (b) *For  $m \leq 8$ ,  $b$  can be chosen to be greater than one;*  
 (c) *The unlimited version (i.e., with all the superscripts on  $df^+$ ,  $df^-$  removed) of the scheme is  $(2m-1)$ -order accurate;*  
 (d) *The scheme (2.2)–(2.7) is TVD under the CFL restriction*

$$(2.9) \quad \lambda \left( \frac{df_{j+1/2}^+ - df_{j+1/2}^-}{\Delta + u_j} \right) \leq \left\{ 1 + b \left[ \frac{1}{2} \sum_{j=2}^m \left( \frac{1}{2j-1} \right) + \frac{1}{2} - \sum_{j=2}^m \frac{1}{2j(2j-1)} - \beta \binom{2m-2}{m-1} - \beta \binom{2m-2}{m-2} \right] + \sum_{j=2}^m \frac{1}{2j(2j-1)} + \beta \binom{2m-2}{m-2} \right\}^{-1}.$$

The CFL number (2.9) is not too small. For example, with  $m = 3$  (5th order) the CFL number is 0.44.

They also have a class of  $\alpha$ -schemes having similar properties.

Because these schemes are TVD, they can be only first order at nonsonic critical points. Hence, no matter how accurate they are at other points, when applied to a general problem with critical points in its solutions, these schemes are only first-order accurate in the  $L_\infty$ -norm and second-order accurate in the  $L_1$ -norm. (In steady state calculations, numerical evidence seems to indicate an increase of accuracy by one order; see [7].) We should also point out that if the E-flux  $h_{j+1/2}$  is not smooth enough, the accuracy at sonic points (where  $f'(u(x, t)) = 0$ ) is also restricted: If Godunov or Roe fluxes (which are just Lipschitz continuous) are used, we only have first order at sonic points; if Engquist-Osher flux (which is  $C^1$ ) is used, we have second order at sonic points; and if Lax-Friedrichs flux (which is  $C^\infty$ ) is used, we may have the full high order of the scheme at sonic points.

We begin to modify the schemes (2.2)–(2.7), trying to get schemes which

- (a) are TVB;
- (b) have the same high order of accuracy as the unlimited version in smooth regions, including at critical and sonic points.

For a positive number  $M \geq 0$  and  $\Delta x$ , we define the function  $mc(M, \Delta x)$  as

$$(2.10) \quad mc(M, \Delta x)(a_1, a_2, \dots, a_l) \\ = m(a_1, a_2 + M\Delta x^2 \text{sign}(a_1), \dots, a_l + M\Delta x^2 \text{sign}(a_1)),$$

where the “minmod” function  $m$  is defined as in (2.8).

We modify (2.6) by changing all the functions  $m$  there to  $mc(M, \Delta x)$ :

$$(2.11a) \quad (df_{j+1/2}^-)^{(k)} = mc(M, \Delta x)(df_{j+1/2}^-, bdf_{j-k+1/2}^-, bdf_{j-k+3/2}^-) \\ \text{for all } k \text{ with } 0 \neq k \neq 1,$$

$$(2.11b) \quad (df_{j+1/2}^-)^{(0)} = mc(M, \Delta x)(df_{j+1/2}^-, bdf_{j+3/2}^-),$$

etc.

This simple change leads to a scheme satisfying both (a) and (b).

**THEOREM 2.2.** (a) *For any  $M > 0$ , the scheme (2.2)–(2.5), (2.11), (2.7) is TVB in  $0 \leq t \leq T$  with a fixed  $T > 0$ , under the same CFL restriction (2.9);*

(b) *If the E-flux  $h_{j+1/2}$  has  $2m - 1$  Lipschitz continuous partial derivatives about its two arguments, then for  $m \leq 8$ , for any given  $A > 0$ , there exists an  $M > 0$  such that the scheme (2.2), (2.4), (2.5), (2.11), (2.7) is  $(2m - 1)$ -order accurate in space in any region where  $|u|, |u_x|, |u_{xx}| \leq A$ .*

*Proof.* (a) If we denote the flux (2.4) with limiters (2.6) by  $\hat{f}_{j+1/2}$  and the flux (2.4) with limiters (2.11) by  $\hat{f}_{j+1/2}^M$ , then, from the definition of  $mc(M, \Delta x)$  in (2.10), we easily see

$$(2.12) \quad \hat{f}_{j+1/2}^M = \hat{f}_{j+1/2} + \hat{c}_{j+1/2}$$

with

$$(2.13) \quad |\hat{c}_{j+1/2}| \leq \sum_{k=-m+1}^{m-1} \left| \mu_k^m + (-1)^k \beta \left( \frac{2m-2}{k+m-1} \right) \right| M\Delta x^2 \\ + \sum_{k=-m+1}^{m-1} \left| \nu_k^m - (-1)^k \beta \left( \frac{2m-2}{k+m-1} \right) \right| M\Delta x^2 \\ = 2M\Delta x^2 \sum_{k=-m+1}^{m-1} |\mu_k^m| = B\Delta x^2,$$

where  $B$  is a fixed number.

In the second to the last equality above we used the fact

$$(2.14) \quad \beta \left( \frac{2m-2}{k+m-1} \right) \leq |\mu_k^m| = |\nu_k^m|$$

which was proved in [7].

Let

$$(2.15) \quad \hat{u}_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2})$$

be the solution of (2.4), (2.6); then the solution of (2.4), (2.11) is

$$(2.16) \quad u_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2}^M - \hat{f}_{j-1/2}^M) \\ = u_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}) - \lambda(\hat{c}_{j+1/2} - \hat{c}_{j-1/2}) \\ = \hat{u}_j^{n+1} - \lambda(\hat{c}_{j+1/2} - \hat{c}_{j-1/2}).$$

So, since (2.4), (2.6) is TVD, we have

$$(2.17) \quad \begin{aligned} \text{TV}(u^{n+1}) &\leq \text{TV}(\dot{u}^{n+1}) + 2\lambda \sum_j |\hat{c}_{j+1/2} - \hat{c}_{j-1/2}| \\ &\leq \text{TV}(u^n) + 4\lambda N B \Delta x^2 = \text{TV}(u^n) + \bar{B} \Delta t, \end{aligned}$$

where  $\bar{B} = 4BL$ ,  $L$  is the length of the  $x$ -support of  $u$  at time  $T$  (or the length of one period in the periodic case), and  $N$  is the total number of intervals.

Now (2.17) trivially implies

$$(2.18) \quad \text{TV}(u^n) \leq \text{TV}(u^0) + \bar{B}T$$

for all  $n$  and  $\Delta t$  such that  $n\Delta t \leq T$ , i.e., the scheme (2.4), (2.11) is TVB in  $0 \leq t \leq T$ .

(b) By consistency  $f(u_j) = h(u_j, u_j)$ , so

$$(2.19) \quad \begin{aligned} df_{j+1/2}^- &= h(u_j, u_{j+1}) - h(u_j, u_j) \\ &= h_2(u_j, u_j) \Delta_+ u_j + h_{22}(u_j, \xi) \frac{(\Delta_+ u_j)^2}{2} = c \Delta x + d \Delta x^2, \end{aligned}$$

where  $c, d$  depend on  $|u|$ ,  $|u_x|$ ,  $|u_{xx}|$  and the derivatives of  $h$ .

Assume  $\Delta x$  is small. If  $c \neq 0$ , then locally  $df_{j+1/2}^-$  and  $df_{j \pm k+1/2}^-$  are of the same sign, and by continuity

$$(2.20) \quad \lim_{\Delta x \rightarrow 0} \frac{df_{j \pm k+1/2}^-}{df_{j+1/2}^-} = 1 < b.$$

In this case, the two functions  $mc(M, \Delta x)$  and  $m$  are the same, and they both pick the first argument  $df_{j+1/2}^-$ .

If  $c = 0$ , there exists an  $M$  (depending on  $A$ ) such that

$$|df_{j \pm k+1/2}^-| \leq M \Delta x^2 / (1 + b)$$

for  $k \leq m - 1$ , hence the  $mc(M, \Delta x)$  function again picks the first argument  $df_{j+1/2}^-$  (unlike the  $m$  function which may pick either argument in this case, which is the source of local degeneracy of accuracy at critical or sonic points). So for any  $A > 0$ , we can always choose  $M$  such that in any region where  $|u|$ ,  $|u_x|$ ,  $|u_{xx}| \leq A$ , the  $mc(M, \Delta x)$  function picks only the first argument, hence the scheme is identical to the unlimited version in such regions, thus is  $(2m - 1)$ -order accurate.

Theorem 2.2 is now proven.  $\square$

**Remark 2.1.** The total variation estimates (2.17), (2.18) are very crude. We can see that  $\hat{c}_{j+1/2} \neq 0$  only at points near which  $df_{j+1/2}^+$  or  $df_{j+1/2}^-$  changes sign (or approaches zero), i.e., near critical points or (in the case of upwind E-fluxes) sonic points, or points of discontinuity. Since usually there are only a fixed number of such points (independent of  $\Delta x$ ), the increment of  $\text{TV}(u^n)$  is actually  $O(\Delta x^2)$  instead of  $O(\Delta x)$  for each time step. So the total variation bound (2.18) can be improved in most cases to  $\text{TV}(u^n) \leq \text{TV}(u^0) + O(\Delta x)$  for  $0 \leq n\Delta t \leq T$ , hence the computation of steady state solutions (where we need a large  $t$ ) is also possible.

**Remark 2.2.** The requirement that the E-flux  $h_{j+1/2}$  must have  $2m - 1$  Lipschitz continuous partial derivatives is essential. Without this, we will get local degeneracy of accuracy at sonic points. This fact is demonstrated by numerical evidence (see the

appendix). For the purpose of smoothness, the Lax-Friedrichs flux (1.14) is the best, because it is in the class  $C^\infty$ . If we do not like the severe dissipativity of Lax-Friedrichs flux, we may also use Roe flux (1.13) with the absolute value function  $|\bar{a}_{j+1/2}|$  (which is not smooth at 0) replaced by a smooth function  $g(\bar{a}_{j+1/2})$  such that  $g(s) = |s|$  if  $|s| \geq \varepsilon$  for some small positive  $\varepsilon$ . Although the first-order Lax-Friedrichs scheme is very dissipative, the higher-order TVB schemes based on Lax-Friedrichs flux seem to perform quite well according to our (limited) numerical experiments.

**3. TVB Schemes of Globally High-Order Accuracy in Space and Time.** The TVB schemes in Section 2 can be equipped with the TVD high-order time discretizations in [10] to give schemes with globally high-order accuracy in space and time.

Let

$$(3.1) \quad L^{(1)}(u)_j = -\lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}).$$

Then (2.2) becomes

$$(3.2) \quad u_j^{n+1} = u_j^n + L^{(1)}(u)_j,$$

which is  $r$ th order in space:

$$(3.3) \quad L^{(1)}(u) = \Delta t(-f(u))_x + O(\Delta x^{r+1})$$

and TVB under some CFL condition

$$(3.4) \quad \lambda \leq \lambda_0$$

(see (2.9)).

If we apply the same scheme to  $u_t = (f(u))_x$  instead of (1.1), we get

$$(3.5) \quad u_j^{n+1} = u_j^n + L^{(-1)}(u)_j,$$

$$(3.6) \quad L^{(-1)}(u) = \Delta t(f(u))_x + O(\Delta x^{r+1}).$$

Then, it is proven in [10] that the scheme

$$(3.7) \quad u_j^{n+1} = \sum_{k=0}^m \left[ \alpha_k u_j^{n-k} + \beta_k (\text{sign}(\beta_k) L^{(\text{sign}(\beta_k))}(u^{n-k})_j) \right]$$

is TVB if

$$(3.8) \quad \alpha_k \geq 0, \quad k = 0, 1, \dots, m,$$

and

$$(3.9) \quad \sum_{k=0}^m \alpha_k = 1,$$

under the CFL condition

$$(3.10) \quad \lambda \leq \lambda_0 \cdot \min_k \left( \frac{\alpha_k}{|\beta_k|} \right),$$

and  $s$ -order accurate in time and space ( $1 \leq s \leq r$ ) if (3.9) and

$$(3.11a) \quad \beta_0 - \sum_{k=1}^m (k\alpha_k - \beta_k) = 1,$$

$$(3.11b) \quad (-1)^l \sum_{k=1}^m k^{l-1} (k\alpha_k - l\beta_k) = 1, \quad l = 2, 3, \dots, s,$$

are satisfied.



In [10], it is also proven that for any  $r$ , the coefficients  $\alpha_k$  and  $\beta_k$  can be found to satisfy (3.8)–(3.11).

Generalizations to equations with forcing terms and to multi-dimensional problems are also given in [10].

Hence, the TVB modification procedure in Section 2, plus the above-mentioned TVD high-order time discretizations in [10], gives globally high-order (up to 15th order) in space and time TVB schemes.

Numerical examples for applying these schemes are given in the appendix of this paper.

**4. TVB Modification Procedure Applied to Other TVD Schemes.** The TVB modification procedure discussed in Section 2 is also applicable to change certain other TVD schemes to TVB schemes with globally high order of accuracy. We use an example to illustrate this. In [1] Harten used a “modified flux” approach to get a second-order TVD scheme from the first-order TVD Roe scheme, by comparing the flux with that of the second-order Lax-Wendroff scheme. His numerical flux is given by

$$(4.1) \quad \hat{f}_{j+1/2} = \frac{1}{2} \left[ f(u_j) + f(u_{j+1}) + g_j + g_{j+1} - |\bar{a}_{j+1/2} + \bar{\gamma}_{j+1/2}| \Delta_+ u_j \right],$$

where

$$(4.2) \quad \bar{a}_{j+1/2} = \frac{f(u_{j+1}) - f(u_j)}{\Delta_+ u_j},$$

$$(4.3) \quad g_j = m(g_{j-1/2}, g_{j+1/2}),$$

with

$$(4.4) \quad g_{j+1/2} = \sigma(\bar{a}_{j+1/2}) \Delta_+ u_j,$$

and the function  $\sigma$  is defined by

$$(4.5) \quad \sigma(a) = \frac{1}{2} |a| (1 - \lambda |a|).$$

Also,

$$(4.6) \quad \gamma_{j+1/2} = \frac{g_{j+1/2} - g_{j-1/2}}{\Delta_+ u_j}.$$

In [1] it was shown that  $\hat{f}_{j+1/2}$  defined by (4.1)–(4.6) is identical to the second-order Lax-Wendroff flux

$$(4.7) \quad \hat{f}_{j+1/2}^{\text{LW}} = \frac{1}{2} \left[ f(u_j) + f(u_{j+1}) - \lambda (\bar{a}_{j+1/2})^2 \Delta_+ u_j \right]$$

up to  $O(\Delta x^2)$ , hence the scheme is second-order at the points where the coefficient of  $\Delta x^2$  term is smooth. Actually, the scheme is second-order except at critical points and sonic points. The scheme is TVD because it is just the Roe scheme applied to the “modified flux”  $f_j + g_j$ . In [1] it is also proved that the new CFL condition is still

$$(4.8) \quad \lambda \max_j |\bar{a}_{j+1/2}| \leq 1.$$

In order to apply the TVB modification procedure, we first change the definition  $g_j$  in (4.3) a little. Notice that the only requirement for  $g_j$  is

$$(4.9) \quad g_j = \Delta x \sigma(a_j)(u_x)_j + O(\Delta x^2).$$

We simply define

$$(4.10) \quad g_j^b = \frac{1}{2}m(g_{j+1/2}, bg_{j-1/2}) + \frac{1}{2}m(g_{j-1/2}, bg_{j+1/2})$$

with  $g_{j+1/2}$  defined by (4.4), (4.5), and  $1 < b \leq 3$ .

Then we can make the TVB modification:

$$(4.11) \quad g_j^{bM} = \frac{1}{2}mc(M, \Delta x)(g_{j+1/2}, bg_{j-1/2}) + \frac{1}{2}mc(M, \Delta x)(g_{j-1/2}, bg_{j+1/2}).$$

We have the following

**THEOREM 4.1.** (a) Scheme (2.2), (4.1)–(4.6), with  $g_j$  in (4.3) replaced by  $g_j^b$  in (4.10), is TVD under the CFL condition (4.8), and second-order accurate except at critical and sonic points;

(b) Scheme (2.2), (4.1)–(4.6), with  $g_j$  in (4.3) replaced by  $g_j^{bM}$  in (4.11), is TVB and second-order accurate except at sonic points, again under CFL condition (4.8).

*Proof.* (a) Since  $b > 1$ , in smooth regions away from critical and sonic points, we have, when  $\Delta x$  is small,

$$g_j^b = \frac{1}{2}g_{j+1/2} + \frac{1}{2}g_{j-1/2} = \Delta x \sigma(a_j)(u_x)_j + O(\Delta x^2),$$

hence (4.9) is satisfied. Also, it is easily seen that  $g_j^b \cdot g_{j+1}^b \geq 0$ , so

$$(4.12) \quad \begin{aligned} |g_{j+1}^b - g_j^b| &\leq \max(|g_{j+1}^b|, |g_j^b|) \\ &\leq \max\left(\frac{1+b}{2}|g_{j+1/2}|, \frac{1+b}{2}|g_{j-1/2}|\right) = \frac{1+b}{2}|g_{j+1/2}|. \end{aligned}$$

Hence,

$$(4.13) \quad |\bar{\gamma}_{j+1/2}^b| = \left| \frac{g_{j+1}^b - g_j^b}{\Delta_+ u_j} \right| \leq \frac{1+b}{2} \frac{|g_{j+1/2}|}{|\Delta_+ u_j|} = \frac{1+b}{2} \sigma(\bar{a}_{j+1/2}).$$

Moreover, if  $b \leq 3$  and  $|v_{j+1/2}| \equiv \lambda|\bar{a}_{j+1/2}| \leq 1$ , then

$$\begin{aligned} \lambda|\bar{a}_{j+1/2} + \bar{\gamma}_{j+1/2}^b| &\leq \lambda|\bar{a}_{j+1/2}| + \lambda|\bar{\gamma}_{j+1/2}^b| \\ &\leq |v_{j+1/2}| + \frac{1+b}{2}\lambda\frac{1}{2}|\bar{a}_{j+1/2}||1 - \lambda|\bar{a}_{j+1/2}|| \\ &\leq |v_{j+1/2}| + (1 - |v_{j+1/2}|) = 1, \end{aligned}$$

i.e., part (a) is proven.

(b) Following the lines of the proof of Theorem 2.2. Except here the absolute value function in the definition of  $\hat{f}_{j+1/2}$  is not smooth at 0, so we have a local degeneracy to first-order accuracy at sonic points.  $\square$

Harten's modified flux technique can of course be applied to any first-order E-flux. In order to get TVB full second-order, we again try to use it to the smooth Lax-Friedrichs flux. Define

$$(4.14) \quad \hat{f}_{j+1/2} = \frac{1}{2} \left[ f(u_j) + f(u_{j+1}) + g_j + g_{j+1} - \frac{1}{\lambda} \Delta_+ u_j \right].$$

Here we use  $\alpha = 1/\lambda$  in (1.14).

By comparing (4.14) with the Lax-Wendroff second-order flux (4.7), we see that we must have

$$(4.15) \quad \frac{1}{2}(g_j + g_{j+1}) = \frac{1}{2} \left( \frac{1}{\lambda} - \lambda a^2 \right) \Delta_+ u_j + O(\Delta x^2)$$

in order to get second-order accuracy.

Let

$$(4.16) \quad \sigma(a) = \frac{1}{2} \left( \frac{1}{\lambda} - \lambda a^2 \right)$$

and  $g_{j+1/2}$ ,  $g_j$ ,  $g_j^b$ ,  $g_j^{bM}$  be defined by (4.4), (4.3), (4.10) and (4.11), respectively; we have the following

**THEOREM 4.2.** (a) *Scheme (2.2), (4.14), (4.16), (4.2)–(4.4) is TVD and second-order accurate, except at critical points, under the CFL condition (4.8); Scheme (2.2), (4.14), (4.16), (4.2), (4.3), (4.10) is TVD and second-order accurate, except at critical points, under the CFL condition*

$$(4.17) \quad \lambda \max_j |\bar{a}_{j+1/2}| \leq \frac{4}{b+1} - 1.$$

(b) *Scheme (2.2), (4.14), (4.16), (4.2), (4.3), (4.11) is TVB in  $0 \leq t \leq T$ , and second-order accurate in smooth regions, under the CFL condition (4.17).*

*Proof.* (a) Following the lines of proofs in [1] and Theorem 4.1, we have, for (4.3),  $\bar{\gamma}_{j+1/2} \leq \sigma(\bar{a}_{j+1/2})$ , so we need

$$\begin{aligned} \lambda |\bar{a}_{j+1/2} + \bar{\gamma}_{j+1/2}| &\leq \lambda |\bar{a}_{j+1/2}| + \lambda |\bar{\gamma}_{j+1/2}| \\ &\leq \lambda |\bar{a}_{j+1/2}| + \frac{1}{2} \left[ 1 - (\lambda |\bar{a}_{j+1/2}|)^2 \right] \leq 1, \end{aligned}$$

which is guaranteed if  $\lambda |\bar{a}_{j+1/2}| \leq 1$ , because

$$\max_{0 \leq s \leq 1} \left( s + \frac{1}{2}(1 - s^2) \right) = 1.$$

For (4.10), we have  $\bar{\gamma}_{j+1/2} \leq (1+b)\sigma(\bar{a}_{j+1/2})/2$ , so for CFL condition we need

$$\lambda |\bar{a}_{j+1/2} + \bar{\gamma}_{j+1/2}| \leq \lambda |\bar{a}_{j+1/2}| + \frac{1+b}{4} \left[ 1 - (\lambda |\bar{a}_{j+1/2}|)^2 \right] \leq 1,$$

which is guaranteed if  $\lambda |\bar{a}_{j+1/2}| \leq 4/(b+1) - 1$ , because

$$s + \frac{1+b}{4}(1 - s^2) \leq 1 \quad \text{and} \quad 0 \leq s \leq 1$$

if and only if

$$0 \leq s \leq \frac{4}{b+1} - 1.$$

(b) Following the lines of the proof of Theorem 2.2.  $\square$

**Remark 4.1.** The CFL condition (4.17) is  $4/(1+b) - 1 < 1$  for  $b > 1$ . (For example, if  $b = 1.5$ , the CFL number is  $\frac{3}{5}$ .) The decrease of CFL number is compensated by the second-order accuracy at sonic points.

**Remark 4.2.** The TVB modification procedure should also be applicable to other TVD schemes involving the “minmod” function  $m$ .

**Acknowledgment.** The author would like to thank Professor Stanley Osher for his valuable help and suggestions.

**Appendix. Numerical Results.** The numbers in this appendix are often written in exponential forms; for example,  $4.2(-3)$  means  $4.2 \times 10^{-3}$ .

All the tables are collected at the end.

*Example 1.* TVD and TVB schemes (3.7) (third-order version, i.e.,  $m = 2$  in (2.4), and  $\alpha_i = \frac{16}{27}, 0, 0, \frac{11}{27}$ ;  $\beta_i = \frac{16}{9}, 0, 0, \frac{44}{27}$ ; CFL number  $= \frac{1}{3}\lambda_0$  in (3.7); see [10] for details) are used to solve Riemann problems of Burgers' equation

$$(A.1) \quad u_t + \left( \frac{u^2}{2} \right)_x = 0, \quad u(x, 0) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0. \end{cases}$$

Case (i):  $u_L = 2$ ,  $u_R = -1$ . The exact solution is a moving shock:

$$u(x, t) = \begin{cases} 2, & x < \frac{1}{2}t, \\ -1, & x > \frac{1}{2}t. \end{cases}$$

We use  $\Delta x = 1/80$ ,  $t = 1.0$  and print out the numerical shock transitions for six different schemes. EO and LF refer to the building blocks (2.3), (1.11), (1.14); the linear scheme is the scheme with all the limiters in (2.4) removed; we take the largest permissible  $b$  in (2.7) and the largest permissible CFL number (2.9), (3.10).  $M$  in (2.11) is taken to be 50 (we use  $M = 50$  for most of the computations); 3-3 means the scheme is third-order in time and space.

3-3-LF-TVD:

2.0000, 1.9989, 1.4344\*, -0.6141, -1.0000

3-3-EO-TVD:

2.0000, 2.0000, 1.6567\*, -0.8375, -1.0000

3-3-LF-TVB:

2.0000, 2.0004, 1.9991, 1.4342\*, -0.6142, -1.0001, -1.0000, -1.0000

3-3-EO-TVB:

2.0000, 1.9997, 2.0008, 1.6577\*, -0.8378, -1.0013, -0.9999, -1.0000

3-3-LF-Linear:

2.0364, 1.8316, 2.4171, 1.5368\*, -0.7970, -1.2230, -0.9861, -0.9933

3-3-EO-Linear:

2.0579, 1.8415, 2.2696, 1.7280\*, -1.0718, -0.9782, -1.0163, -0.9975

(The starred positions are the positions of the shock.)

*Several Observations:* (1) The transitions of 3-3-LF and 3-3-EO schemes are both very sharp: about 3-point transitions;

(2) There is no substantial difference between LF and EO building blocks;

(3) Oscillations in TVB schemes are small ( $< 0.1M\Delta x^2$ );

(4) Oscillations in linear schemes are large (of order  $O(1)$ ), as expected.

Case (ii):  $u_L = \frac{4}{3}$ ,  $u_R = 2$ . The exact solution is a rarefaction wave:

$$(A.2) \quad u(x, t) = \begin{cases} \frac{4}{3} & \text{if } x < \frac{4}{3}t, \\ x/t & \text{if } \frac{4}{3}t \leq x \leq 2t, \\ 2 & \text{if } x > 2t. \end{cases}$$

There is also an entropy-violating expansion shock solution:

$$(A.3) \quad u(x, t) = \begin{cases} \frac{4}{3} & \text{if } x < \frac{5}{3}t, \\ 2 & \text{if } x > \frac{5}{3}t. \end{cases}$$

This time we use  $\Delta x = 1/80$ ,  $t = 0.4$  and two schemes: 1-3-EO-TVD (Euler forward in time) and 3-3-EO-TVD (since there is no sonic point, any upwind scheme is the same in this case). The numerical output (which we omit to print out) shows that the 1-3-EO-TVD scheme converges to the wrong solution (A.3), while 3-3-EO-TVD scheme converges to the correct solution (A.2). This indicates that Euler forward in time is dangerous for entropy condition, while 3-3 schemes are better.

(We have also used our 3-3 schemes for many other test problems, including some problems with nonconvex  $f(u)$ . No entropy-violation solution was observed.)

*Example 2.* TVD and TVB globally high-order schemes (3.7) are used to solve a linear periodic problem with smooth initial condition (the scheme is nonlinear!):

$$(A.4) \quad u_t + u_x = 0, \quad u(x, 0) = \sin \pi x.$$

The solution is computed in  $-1 \leq x < 1$  up to  $t = 2$  (after one period in time).

We use a third-order (same as in Example 1) and a fifth-order scheme ( $m = 3$  in (2.4);

$$\alpha_i = \frac{7}{20}, \frac{3}{10}, \frac{4}{15}, 0, \frac{7}{120}, \frac{1}{40};$$

$$\beta_i = \frac{291201}{108000}, \frac{-198401}{86400}, \frac{88063}{43200}, 0, \frac{-17969}{43200}, \frac{73061}{432000}$$

in (3.7));  $M = 50$  in (2.11), the largest permissible  $b$  and  $\lambda$  in (2.7), (2.9), (3.10).

The errors and numerical orders are listed in Table 1.

We observe that the results in Table 1 agree with the theory pretty well: The TVD schemes are first-order in  $L_\infty$ -norm and second-order in  $L_1$ -norm, while the TVB schemes have full high-order of accuracy both in  $L_1$  and in  $L_\infty$ .

*Example 3.* The third-order schemes in Example 2 are used to solve the same linear equation with a discontinuous initial condition:

$$(A.5) \quad u(x, 0) = \sin \pi x + 0.5 \text{sign}(\sin \pi x).$$

We are interested in the behavior of the schemes near discontinuities and in smooth regions. We use  $\Delta x = \frac{1}{80}$  and compute the solution up to  $t = 2$  (after one period in time).

From the printout (which we omit) we observe that the schemes behave similarly as in Example 2 in smooth regions. For example, at  $x = 0.5$  (which is a critical point), the errors for TVD and TVB schemes are  $2.6(-3)$  and  $1.8(-5)$ , respectively. These numbers are almost identical to the corresponding errors  $[2.6(-3)$  and  $2.2(-5)]$  in the smooth problem of Example 2 with the same  $\Delta x$ .

Near the discontinuity  $x = 0$ , we print out the errors at ten points for each scheme:

3-3-TVD:

$$-3.9(-2), \quad 5.8(-3), \quad 8.3(-2), \quad 1.9(-1), \quad 3.2(-1), \quad *, \quad -3.8(-1), \quad -2.5(-1), \\ -1.4(-1), \quad -5.5(-2), \quad 7.4(-4)$$

3-3-TVB:

$$-3.9(-2), \quad 5.4(-3), \quad 8.2(-2), \quad 1.9(-1), \quad 3.2(-1), \quad *, \quad -3.8(-1), \quad -2.5(-1), \\ -1.4(-1), \quad -5.4(-2), \quad 9.6(-4)$$

where the star  $*$  is the position of the discontinuity  $x = 0$ .

We can see that the TVD and TVB schemes are almost identical near the discontinuity. There is no extra oscillation for the TVB scheme in this case.

*Example 4.* Same schemes as in Example 3 are used to solve the nonlinear Burgers' equation with periodic initial conditions:

$$(A.6) \quad u_t + \left( \frac{u^2}{2} \right)_x = 0, \quad -1 \leq x \leq 1, \quad u(x, 0) = \alpha + \frac{1}{2} \sin \pi x.$$

The exact solution is smooth up to  $t = 2/\pi$ , then it develops a moving (if  $\alpha \neq 0$ ) shock which also acts with the rarefaction waves. We can get the exact solution by Newton iteration. For details, see [5].

*Case (1):*  $\alpha = 1$ . There is no sonic point in this case. Hence we may use either EO or LF building blocks. We use the same parameters  $b$ ,  $\lambda$ ,  $M$  as in Example 2.

The errors of the numerical solutions at  $t = 0.3$  are listed in Table 2. Since the exact solution is still smooth, the numerical results are very similar to those of Example 2: TVD are first-order in  $L_\infty$  and second-order in  $L_1$ , while TVB are third-order globally. There is no substantial difference between the LF and EO building blocks.

At  $t = 2/\pi$ , the shock begins to form. We use  $\Delta x = 1/80$  and print out the errors at 10 points near the shock for both schemes:

3-3-LF-TVD:

$-1.5(-3)$ ,  $2.0(-3)$ ,  $1.8(-2)$ ,  $1.8(-2)$ ,  $-6.5(-2)^*$ ,  $9.8(-2)$ ,  $4.8(-2)$ ,  $-7.0(-3)$ ,  
 $-1.2(-2)$ ,  $-5.3(-2)$

3-3-LF-TVD:

$-1.5(-3)$ ,  $2.2(-3)$ ,  $1.9(-2)$ ,  $1.8(-2)$ ,  $-6.6(-2)^*$ ,  $9.8(-2)$ ,  $4.8(-2)$ ,  $-7.0(-3)$ ,  
 $-1.2(-2)$ ,  $-5.3(-3)$

where the star  $*$  is the position of the shock.

We can see that there is essentially no difference between the TVD and TVB schemes in this case.

In smooth regions the numerical solutions are very good. We compute the  $L_1$  and  $L_\infty$  errors in the smooth region 0.1 away from the shock (i.e.,  $|x - \text{shock}| \geq 0.1$ ) and list them in Table 3. From the table we can see that the error is of the same size as in the smooth case  $t = 0.3$ .

At  $t = 1.1$ , the reaction between the shock and the rarefaction waves is over. The solution becomes monotone between shocks. In this case the TVD and TVB schemes yield almost the same results. For smooth region errors see Table 3. For the behaviors near the shock, we again print out the errors at 10 points for each scheme:

3-3-LF-TVD:

$-3.5(-4)$ ,  $1.0(-3)$ ,  $1.8(-3)$ ,  $-2.1(-3)$ ,  $-6.0(-2)$ ,  $*$ ,  $6.9(-2)$ ,  $6.7(-5)$ ,  
 $-1.3(-3)$ ,  $-6.0(-4)$ ,  $-5.2(-4)$

3-3-LF-TVB:

$-8.6(-4)$ ,  $-9.9(-4)$ ,  $4.1(-3)$ ,  $3.8(-3)$ ,  $-5.4(-2)$ ,  $*$ ,  $6.3(-2)$ ,  $-4.7(-3)$ ,  
 $-3.0(-3)$ ,  $-2.6(-4)$ ,  $3.2(-4)$

where the star  $*$  is again the position of the shock.

As before, we do not observe any essential difference between the TVD and TVB schemes near the shock.

*Case (ii):*  $\alpha = 0.25$ . There exist sonic points in this case, so LF is preferred to EO. For errors at  $t = 0.3$  (smooth solution) see Table 2. We observe that 3-3-EO (both TVD and TVB) becomes first-order in  $L_\infty$  and second-order in  $L_1$ , while 3-3-LF-TVB is still globally third-order. The errors in smooth regions at  $t = 2/\pi$  and at  $t = 1.1$  are again listed in Table 3. The shock transitions are similar to those of Case (i) above, so we omit printing them.

*Example 5.* The same 3-3 scheme as in the previous examples (2-dim. version, see [10]) is used to solve a 2-dim. linear problem with smooth periodic initial condition:

$$(A.7) \quad u_t + u_x + u_y = 0, \quad -1 \leq x, y \leq 1, \quad u(x, y, 0) = \sin \pi x \sin \pi y.$$

We compute the solution at  $t = 2$  (after one period in time). For CFL number we just use the 1-dim. restriction (3.10).

For numerical errors see Table 4.

We observe that the errors (TVB case) are roughly twice those of the corresponding errors in the 1-dim. case (Table 1). This is expected by theory.

*Example 6.* The same scheme as in Example 5 is used to solve the same linear equation with periodic discontinuous initial condition:

$$(A.8) \quad u(x, y, 0) = \begin{cases} 1 & \text{if } (x, y) \in S, \\ 0 & \text{if } (x, y) \notin S, \end{cases}$$

where  $S = \{(x, y): |x - y| < 1/\sqrt{2}, |x + y| < 1/\sqrt{2}\}$  is a unit square centered at the origin and rotated by  $\pi/4$  (see [3]). We use  $\Delta x = \frac{1}{10}$  (which is not too small) and again the 1-dim. CFL restriction (3.10). For the TVB scheme we use  $M = 50$  in (2.11).

The numerical values at  $t = 2$  (after one period in time), at  $t = 8$  (after four periods in time) and at  $t = 16$  (after eight periods in time) are computed. We print out the values at  $x \leq 0$ ,  $y = 0$  and  $y = -0.4$ . The exact solutions are 0 to the left of the star \* and is 1 to the right.

$t = 2, y = 0$ ; TVD: .008, .014, .068, \*.24, .48, .72, .90, .99, 1.00, 1.00, 1.00  
 $t = 2, y = 0$ ; TVB: -.082, -.041, .10, \*.33, .59, .83, .98, 1.05, 1.06, 1.04, 1.04  
 $t = 2, y = -0.4$ ; TVD: .011, .009, .010, .079, .12, .10, .32, \*.53, .73, .88, .90  
 $t = 2, y = -0.4$ ; TVB: -.016, -.026, -.036, -.024, .038, .16, .35, \*.56, .77, .92, .98  
 $t = 8, y = 0$ ; TVD: .056, .058, .084, \*.19, .36, .55, .70, .78, .81, .81, .76  
 $t = 8, y = 0$ ; TVB: -.025, .018, .12, \*.28, .46, .65, .82, .96, 1.07, 1.14, 1.16  
 $t = 8, y = -0.4$ ; TVD: .041, .055, .081, .090, .12, .24, .35, \*.49, .61, .68, .70  
 $t = 8, y = -0.4$ ; TVB: -.041, -.033, .006, .077, .18, .30, .44, \*.58, .70, .78, .82  
 $t = 16, y = 0$ ; TVD: .13, .13, .13, \*.18, .28, .40, .51, .58, .60, .60, .58  
 $t = 16, y = 0$ ; TVB: .012, .038, .12, \*.24, .39, .56, .73, .89, 1.01, 1.09, 1.12  
 $t = 16, y = -0.4$ ; TVD: .11, .12, .12, .13, .19, .20, .22, \*.34, .41, .48, .51  
 $t = 16, y = -0.4$ ; TVB: -.019, -.011, .031, .10, .20, .31, .43, \*.55, .64, .71, .73

*Observations:* (1) With the 1-dim. CFL number the TVD and TVB schemes seem to be still stable. This indicates that the CFL restriction in [10] (half of the CFL restriction of 1-dim.) may be too conservative.

(2) The TVB scheme performs better than the TVD. The overshoots and undershoots are not growing drastically when  $t$  is large.

TABLE 1 (Example 2)

$t = 2$ ; E: type of error;  $r$ : numerical order

E	3-3 Scheme					5-5 Scheme			
	$\Delta x$	TVD	$r$	TVB	$r$	TVD	$r$	TVB	$r$
$L_\infty$	1/10	8.1(-2)		1.6(-2)		1.8(-1)		1.1(-3)	
	1/20	2.5(-2)	1.67	2.0(-3)	2.96	6.4(-2)	1.48	2.6(-5)	5.33
	1/40	8.2(-3)	1.63	2.5(-4)	3.01				
	1/80	2.7(-3)	1.62	2.4(-5)	3.40				
$L_1$	1/10	2.3(-2)		1.0(-2)		9.4(-2)		6.8(-4)	
	1/20	5.0(-3)	2.24	1.3(-3)	2.98	2.4(-2)	1.95	1.6(-5)	5.38
	1/40	1.0(-3)	2.29	1.6(-4)	3.02				
	1/80	2.0(-4)	2.33	1.5(-5)	3.45				

TABLE 2 (Example 4)  
 $t = 0.3$ ; E: type of error;  $r$ : numerical order

$\alpha$		EO building block					LF building block				
	E	$\Delta x$	TVD	$r$	TVB	$r$	TVD	$r$	TVB	$r$	
1.00	$L_\infty$	1/10	1.6(-2)		1.3(-2)		1.6(-2)		1.7(-2)		
		1/20	6.7(-3)	1.26	2.6(-3)	2.36	6.7(-3)	1.21	2.6(-3)	2.67	
		1/40	2.2(-3)	1.57	3.6(-4)	2.83	2.2(-3)	1.58	3.7(-4)	2.95	
		1/80	7.4(-4)	1.60	5.0(-5)	2.86	7.4(-4)	1.60	4.8(-5)	2.95	
	$L_1$	1/10	4.4(-3)		2.5(-3)		4.5(-3)		2.9(-3)		
		1/20	1.0(-3)	2.14	3.8(-4)	2.75	1.2(-3)	1.86	4.0(-4)	2.84	
		1/40	1.9(-4)	2.38	4.9(-5)	2.95	2.4(-4)	2.36	5.1(-5)	2.98	
		1/80	3.7(-5)	2.38	6.8(-6)	2.86	4.4(-5)	2.47	7.4(-6)	2.79	
0.25	$L_\infty$	1/10	1.1(-2)		1.5(-2)		1.2(-2)		1.2(-2)		
		1/20	3.7(-3)	1.63	5.5(-3)	1.45	3.9(-3)	1.61	1.9(-3)	2.62	
		1/40	1.6(-3)	1.18	2.3(-3)	1.29	1.7(-3)	1.16	2.4(-4)	3.00	
		1/80	8.2(-4)	1.00	1.3(-3)	0.84	5.9(-4)	1.57	2.9(-5)	3.05	
	$L_1$	1/10	3.0(-3)		2.5(-3)		3.7(-3)		2.1(-3)		
		1/20	8.4(-4)	1.82	5.9(-4)	2.17	8.9(-4)	2.07	2.8(-4)	2.91	
		1/40	1.9(-4)	2.15	1.3(-4)	2.17	1.7(-4)	2.40	3.3(-5)	3.07	
		1/80	4.1(-5)	2.21	2.7(-5)	2.27	3.1(-5)	2.45	4.1(-6)	3.01	

TABLE 3 (Example 4)

Errors in smooth region  $|x - \text{shock}| \geq 0.1$ ;  $\Delta x = 1/80$ ; E: error type

$\alpha$		EO building block				LF building block			
		$t = \pi/2$		$t = 1.1$		$t = \pi/2$		$t = 1.1$	
	E	TVD	TVB	TVD	TVB	TVD	TVB	TVD	TVB
1.0	$L_\infty$	9.5(-4)	1.1(-4)	2.8(-5)	3.3(-5)	9.6(-4)	1.5(-4)	2.9(-4)	4.0(-5)
	$L_1$	5.9(-5)	9.5(-6)	1.1(-5)	1.1(-5)	7.2(-4)	1.2(-5)	1.0(-5)	1.0(-5)
0.25	$L_\infty$	1.1(-3)	7.8(-4)	1.2(-3)	9.3(-4)	7.6(-4)	8.5(-5)	1.2(-5)	8.8(-6)
	$L_1$	5.7(-5)	1.7(-5)	2.2(-5)	2.1(-5)	4.7(-5)	6.1(-6)	1.3(-6)	1.2(-6)

TABLE 4 (Example 5)

$L_\infty$ :  $L_\infty$ -error;  $L_1$ :  $L_1$ -error;  $r$ : numerical order

3-3 TVD					3-3TVB			
$\Delta x$	$L_\infty$	$r$	$L_1$	$r$	$L_\infty$	$r$	$L_1$	$r$
1/10	1.6(-1)		2.2(-2)		3.2(-2)		1.3(-2)	
1/20	5.0(-2)	1.63	4.8(-3)	2.18	4.1(-3)	2.95	1.7(-3)	2.96
1/40	1.7(-2)	1.61	1.1(-3)	2.15	5.2(-4)	2.98	2.1(-4)	2.98

Institute for Mathematics and Its Applications  
 University of Minnesota  
 Minneapolis, Minnesota 55455

1. A. HARTEN, "High resolution schemes for hyperbolic conservation laws," *J. Comput. Phys.*, v. 49, 1983, pp. 357-393.
2. A. HARTEN, "On a class of high resolution total-variation-stable finite difference schemes," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 1-23.
3. A. HARTEN, *Preliminary Results on the Extension of ENO Schemes to Two-Dimensional Problems*, Proc. Internat. Conf. on Hyperbolic Problems, Saint-Etienne, January 1986.



4. A. HARTEN & S. OSHER, "Uniformly high-order accurate nonoscillatory schemes, I," *SIAM J. Numer. Anal.*, v. 24, 1987, pp. 279–309.
5. A. HARTEN, B. ENGQUIST, S. OSHER & S. CHAKRAVARTHY, "Uniformly high order accurate essentially non-oscillatory schemes," *J. Comput. Phys.* (To appear.)
6. S. OSHER & S. CHAKRAVARTHY, "High resolution schemes and the entropy condition," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 955–984.
7. S. OSHER & S. CHAKRAVARTHY, *Very High Order Accurate TVD Schemes*, ICASE Report #84-44, 1984, IMA Volumes in Mathematics and its Applications, vol. 2, Springer-Verlag, 1986, pp. 229–274.
8. P. ROE, "Approximate Riemann solvers, parameter vectors, and difference schemes," *J. Comput. Phys.*, v. 43, 1981, pp. 357–372.
9. C. SHU, "TVD time discretization I—Steady state calculations." (Preprint.)
10. C. SHU, "TVD time discretization II—Time dependent problems." (Preprint.)
11. C. SHU, "TVB boundary treatment for numerical solutions of conservation laws," *Math. Comp.*, v. 49, 1987, pp. 123–134.