MATHEMATICS OF COMPUTATION
Volume 88, Number 316, March 2019, Pages 691–716
https://doi.org/10.1090/mcom/3364
Article electronically published on June 13, 2018

TOTAL VARIATION BOUNDED FLUX LIMITERS FOR HIGH ORDER FINITE DIFFERENCE SCHEMES SOLVING ONE-DIMENSIONAL SCALAR CONSERVATION LAWS

SULIN WANG AND ZHENGFU XU

Abstract. In this paper, we focus on developing locally conservative high order finite difference methods with provable total variation stability for solving one-dimensional scalar conservation laws. We introduce a new criterion for designing high order finite difference schemes with provable total variation stability by measuring the total variation of an expanded vector. This expanded vector is created from grid values at t^{n+1} and t^n with ordering determined by upwinding information. Achievable local bounds for grid values at t^{n+1} are obtained to provide a sufficient condition for the total variation of the expanded vector not to be greater than total variation of the initial data. We apply the Flux-Corrected Transport type of bound preserving flux limiters to ensure that numerical values at t^{n+1} are within these local bounds. When compared with traditional total variation bounded high order methods, the new method does not depend on mesh-related parameters. Numerical results are produced to demonstrate: the total variation of the numerical solution is always bounded; the order of accuracy is not sacrificed. When the total variation bounded flux limiting method is applied to a third order finite difference scheme, we show that the third order of accuracy is maintained from the local truncation error point of view.

1. Introduction

Convergence results for finite difference schemes solving scalar conservation laws are very limited. For high order finite difference schemes, it is extremely challenging to provide such an analysis. The main reason is the lack of provable total variation stability and entropy stability. In this paper, we focus on developing a total variation bounded high order finite difference scheme. We first briefly review the major results and challenges in the context of a one-dimensional scalar conservation law

$$u_t + f(u)_x = 0,\ u(x,0) = u_0(x). \tag{1.1}$$

For simplicity, we ignore the boundary condition, assuming either the initial data has compact support or the problem has a periodic solution. For hyperbolic systems, the total variation is no longer a valid metric for both the exact and numerical solutions unless the system is linear and one-dimensional; see [14] and the references therein.

Received by the editor March 11, 2017, and, in revised form, May 17, 2017, and November 13, 2017.

2010 *Mathematics Subject Classification.* Primary 58J45, 65M06.

Key words and phrases. Hyperbolic conservation laws, bound preserving, flux limiters, high order scheme, total variation stability.

The authors would like to acknowledge the support of the NSF grant DMS-1316662 "High Order Maximum Principle Preserving Finite Difference Schemes for Hyperbolic Conservation Laws".

1.1. **Total variation stability.** The definition of total variation of a real-valued function $g(x)$ on one-dimensional interval $[a, b]$ is

$$\mathrm{Var}(g) = \sup_{[x_1=a<x_2<\cdots<x_{p-1}<x_p=b]} \sum_{j=1}^{p-1} |g(x_{j+1}) - g(x_j)|, \tag{1.2}$$

which equals $\int_a^b |g'(x)|dx$ when the function is differentiable. With a bounded-variation initial condition, the entropy solution to the equation (1.1) enjoys a total variation contractive property

$$\mathrm{Var}(u(\cdot, t_1)) \le \mathrm{Var}(u(\cdot, t_2)) \text{ for any } t_1 \ge t_2, \tag{1.3}$$

assuming the solution is continuous and piecewise differentiable. It would be ideal if the numerical solution shared a similar property not only for the sake of fidelity to the exact solution but also for the convergence of the numerical solution after mesh refinement. We use u_j^n for the numerical approximation of the true solution $u(x_j, t^n)$. Consistent with the definition (1.2), the total variation of a numerical solution is measured by grid point values

$$\mathrm{TV}(u^n) = \sum_j |u_{j+1}^n - u_j^n|. \tag{1.4}$$

A conservative scheme

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}) \tag{1.5}$$

is said to be total variation diminishing (TVD) if

$$\mathrm{TV}(u^{n+1}) \le \mathrm{TV}(u^n) \tag{1.6}$$

or total variation bounded (TVB) if $\mathrm{TV}(u^n)$ is uniformly bounded in time steps and mesh size. It is obvious that the TVD property is perfectly the numeric analog of the total variation contractive property (1.3) while the TVB property is a relaxed requirement. The numerical solution produced by a conservative, consistent and total variation stable finite difference scheme contains a subsequence converging to a weak solution of the initial value problem (1.1) [5, 8, 13]. In some of the earliest work by Van Leer [25, 26] and others, flux limiters and the concept of high order Riemann solvers are introduced to obtain the total variation stability and the second order of accuracy. In general, a three-point conservative scheme (1.5) can be cast into the incremental form

$$u_j^{n+1} = u_j^n + C_{j+\frac{1}{2}}^+ \Delta_{j+\frac{1}{2}} u^n - C_{j-\frac{1}{2}}^- \Delta_{j-\frac{1}{2}} u^n, \tag{1.7}$$

where $\Delta_{j+\frac{1}{2}} u^n = u_{j+1}^n - u_j^n$. For this family of schemes to be total variation stable, an explicit and sufficient condition was given by Harten [8] as

$$C_{j+\frac{1}{2}}^{\pm} \ge 0,\ C_{j+\frac{1}{2}}^+ + C_{j+\frac{1}{2}}^- \le 1 \text{ for all } j. \tag{1.8}$$

The TVD property $\mathrm{TV}(u^{n+1}) \le \mathrm{TV}(u^n)$ is induced when condition (1.8) is satisfied. Based on this criterion, various finite difference schemes were modified through flux limiting to obtain the total variation stability [1, 24]. The obvious advantage of Harten's criterion (1.8) is that it is very easy to comply with. However, it is well known that TVD schemes measured by (1.4) necessarily degenerate to first order accuracy at smooth extrema [18]. Thus those schemes are at most second order accurate when measured in L^1-norm for smooth and nonmonotone solutions.

Moreover, for two-dimensional scalar conservation law computation, Goodman and Leveque [7] (Proposition 2.2) showed that a large family of TVD schemes, measured by grid values, are at most first order accurate. The TVD stability requirement poses as a great challenge for one to design higher order schemes.

1.2. **Total variation stable high order schemes.** Efforts have been made to relax the total variation stability requirement from TVD to TVB. One sufficient condition for a scheme to be TVB is

$$\mathrm{TV}(u^{n+1}) \le (1 + M\Delta t)\mathrm{TV}(u^n), \tag{1.9}$$

which allows for the total variation to grow while remaining bounded within finite time. There are two different practices to achieve (1.9). The first is to introduce mesh-dependent limiters to reconstructed high order polynomials such that the whole scheme can still be cast into a slightly modified version of the incremental form (1.7), i.e., [20] and its application to the discontinuous Galerkin method [2]. This approach fixes the order degeneracy problem and suppresses numerical oscillation around discontinuities. However, its major drawback is that it no longer maintains the scale-invariance property of the solution (the fact that $u(\frac{x}{a}, \frac{t}{a})$ is still a solution to the differential equation when x and t are scaled by the same constant a) to the equation (1.1).

The second approach is the widely discussed essentially nonoscillatory (ENO) and weighted ENO (WENO) methods [9, 15, 21–23]. The mechanism to control the total variation of the numerical solution relies on locally adaptive, nonlinear reconstruction [21]. Although the ENO/WENO schemes produce solutions with uniformly high order accuracy and demonstrate robustness in suppressing oscillations around discontinuities, mathematically provable TVB properties are still elusive. In other words, (1.9) is not strictly followed even though extensive numerical experiments agree with it. Other concerns with this family of methods include the optimal design of the smoothness indicators and appropriate choice of a tuning parameter often used in calculating smoothness indicators. On a separate note, Tadmor and his collaborators recently found that the ENO scheme enjoys a sign property, based on which an entropy stable high order TeCNO scheme was developed [4].

1.3. **Total variation bounded schemes based on measurement of polynomials.** Sanders developed a third order TVD finite volume scheme [19], showing that controlling the total variation is compatible with high order accuracy. There, the total variation of the reconstructed polynomial (instead of grid values) is measured. Key components of this approach include delicate polynomial reconstruction and evolution of cell interface values through characteristic methods. Recently, the idea is further advanced by Zhang and Shu [30] up to sixth order finite volume schemes. This approach is limited to a one-dimensional problem due to the fact that complex steps and extra computational cost are needed for the reconstruction of high order polynomials. Meanwhile, extrema must be identified and tracked during evolution of the scheme. It is unclear how to generalize this line of approach to high order finite difference methods and multidimensional problems. To summarize, a truly high order scheme with mathematically provable total variation stability (based on discrete values) is still unknown.

1.4. **A new total variation stability criterion.** The new component of this paper is mainly to introduce a different criterion and mechanism of controlling the total variation of the numerical solution. The goal is set to be

$$\mathrm{TV}(u^{n+1}) \le \mathrm{Var}(u_0(x)) \text{ for any } n \tag{1.10}$$

by requiring that

$$\mathrm{TV}(v(u^n, u^{n+1})) \le \mathrm{Var}(u_0(x)), \tag{1.11}$$

given $\mathrm{TV}(u^n) \le \mathrm{Var}(u_0(x))$. Here, $v(u^n, u^{n+1})$ is the expanded vector made of u^n and u^{n+1}. The ordering of the new vector is determined by upwinding information of the PDE. For example, in the case of $f'(u) \ge 0$, the expanded vector is defined by

$$v(u^n, u^{n+1}) = [\ldots, u^n_{j-1}, u^{n+1}_j, u^n_j, u^{n+1}_{j+1}, u^n_{j+1}, \ldots].$$

Thus, the inequality (1.11) automatically implies (1.10).

The motivation and the idea originates from the fact that the entropy solution of the problem (1.1) does not necessarily satisfy the discrete version of TVD property (1.6). However, the entropy solution does satisfy the total variation bounded requirement (1.10) in the sense that $\mathrm{TV}(u(\cdot, t^n)) \le \mathrm{Var}(u_0(x))$ when $u(\cdot, t^n)$ is the vector of the grid values at t^n. Even though the TVD property (1.6) ensures the convergence of a subsequence of the numerical solutions to a weak solution among its many benefits, the less demanding TVB requirement (1.10) is sufficient to guarantee a convergent subsequence when meshes are refined.

Local bounds $u_{m,j}, u_{M,j}$ for the updated values u^{n+1}_j are constructed such that the inequality (1.11) is satisfied when

$$u_{m,j} \le u^{n+1}_j \le u_{M,j} \text{ for any } j. \tag{1.12}$$

The detailed process will be given in Section 4 as for obtaining these local bounds. To satisfy the inequalities (1.12), we use the bound preserving flux limiters for high order methods [27, 28] as the major tool.

The paper is organized as follows. In Section 2, we will review a high order finite difference scheme in its compressed form for the convenience of discussion. In Section 3, we will review the bound preserving flux limiters. In Section 4, we shall introduce the sufficient condition for a high order finite difference scheme to be TVB in the sense of (1.10). In Section 5, we shall give a proof that a third order finite difference scheme with the TVB flux limiters maintains third order of accuracy. We provide numerical evidences in Section 6. Finally, we conclude our discussion in Section 7.

2. Higher order finite difference scheme

Spatial disrecretization. We briefly review the approximation of the spatial operator, given the following uniform discretization of the spatial interval $[a, b]$ with the grid interval and center defined by

$$I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}],\ x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}}),\ j = 1, 2, \ldots, N. \tag{2.1}$$

Let $u_j(t)$ denote the value of the solution at center point x_j at time t. Then the conventional finite difference scheme evolves the point values of the solution in a semi-discretized conservative form:

$$\frac{d}{dt}u_j(t) + \frac{1}{\Delta x}(\hat{H}_{j+\frac{1}{2}} - \hat{H}_{j-\frac{1}{2}}) = 0, \tag{2.2}$$

where $\hat{H}_{j+\frac{1}{2}}$ is the numerical flux. Within this formulation, the finite difference methods differ in how the numerical fluxes $\hat{H}_{j+\frac{1}{2}}$ are constructed from point values $u_j(t)$. Following the pioneering first order Godunov scheme [6], many of the monotonic upstream-centered scheme for conservation laws (MUSCL) were discussed [3, 10, 18, 25] with the aim of designing high resolution schemes equipped with total variation and entropy stability. To improve the accuracy of approximation of smooth solutions and the resolution of different flow structures in fluid simulation, arbitrarily high order ENO/WENO finite difference methods were developed with adaptive construction of the numerical fluxes through high order polynomials [9, 11, 12, 16, 17, 22]. In this paper, we follow the framework by Shu and Osher [22] to construct the high order numerical fluxes. A sliding function H was introduced satisfying

$$f(u(x,t)) = \frac{1}{\Delta x}\int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} H(\xi)d\xi. \tag{2.3}$$

Thus $f(u)_x = \frac{1}{\Delta x}(H(x+\frac{\Delta x}{2}) - H(x-\frac{\Delta x}{2}))$. The rest of the construction to approximate $H(x+\frac{\Delta x}{2})$ will be same as a finite volume high order reconstruction.

Temporal integration. A fully discrete high order numerical scheme is completed by choosing a time integration method to solve the nonlinear ODE system (2.2). For the convenience of discussion, we use the third order TVD Runge-Kutta time discretization [23] as an example:

$$\begin{cases} u^{(1)} = u^n + \Delta t L(u^n), \\ u^{(2)} = u^n + \Delta t\left(\frac{1}{4}L(u^n) + \frac{1}{4}L(u^{(1)})\right), \\ u^{n+1} = u^n + \Delta t\left(\frac{1}{6}L(u^n) + \frac{4}{6}L(u^{(2)}) + \frac{1}{6}L(u^{(1)})\right). \end{cases} \tag{2.4}$$

Here, $L(u^n) := -\frac{1}{\Delta x}(\hat{H}^{(n)}_{j+\frac{1}{2}} - \hat{H}^{(n)}_{j-\frac{1}{2}})$, and u^n is the value at the nth time step. $\hat{H}^{(n)}_{j+\frac{1}{2}}$, $\hat{H}^{(1)}_{j+\frac{1}{2}}$, and $\hat{H}^{(2)}_{j+\frac{1}{2}}$ are the numerical fluxes reconstructed based on u^n, $u^{(1)}$, and $u^{(2)}$, respectively. Then we can write the scheme in compressed form,

$$u_j^{n+1} = u_j^n - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{H}^{rk}_{j-\frac{1}{2}}), \tag{2.5}$$

with $\lambda = \frac{\Delta t}{\Delta x}$ and

$$\hat{H}^{rk}_{j+\frac{1}{2}} = \frac{1}{6}\hat{H}^{(n)}_{j+\frac{1}{2}} + \frac{1}{6}\hat{H}^{(1)}_{j+\frac{1}{2}} + \frac{2}{3}\hat{H}^{(2)}_{j+\frac{1}{2}}. \tag{2.6}$$

We will explain our approach based on the compressed form (2.5). Many of the high order schemes in the form of (2.5)) do not necessarily satisfy TVB requirement (1.10). Our approach can be applied to most of the high order temporal integration as evidenced in the following sections.

3. Bounds preserving flux limiters

In this paper, the key tool to develop the TVB high order method is the bound preserving flux limiters introduced in [27, 28]. This flux-limiting method is generalized from the flux-corrected transport method [1, 29]. When the bounds are defined by $u_m = \min u_0(x), u_M = \max u_0(x)$ such as in [27, 28], the general parametrized

flux-limiting method proposed in [27] for one-dimensional scalar conservation law seeks to modify the integrated flux $\hat{H}^{rk}_{j+\frac{1}{2}}$:

$$\tilde{H}^{rk}_{j+\frac{1}{2}} = \theta_{j+\frac{1}{2}}(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j+\frac{1}{2}}) + \hat{h}_{j+\frac{1}{2}} \tag{3.1}$$

(the limiting parameter $\theta_{j+\frac{1}{2}}$ is a number between 0 and 1) such that the new high order numerical fluxes satisfy

$$u_m \le u^n_j - \lambda(\tilde{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j-\frac{1}{2}}) \le u_M. \tag{3.2}$$

Here $\hat{h}_{j+\frac{1}{2}}$ is a low order monotone flux satisfying

$$u_m \le u^n_j - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) \le u_M.$$

The numerical scheme with the modified numerical flux is still locally conservative and consistent, while retaining the designed high order of accuracy of the original scheme (2.5).

Detailed steps are given in [28] to find the limiting parameters through the coupled inequalities (3.2). For the completeness of this paper, we review those steps assuming local bounds are found as $u_{m,j}, u_{M,j}$ and preserved by the first order monotone flux

$$u_{m,j} \le u^n_j - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) \le u_{M,j}. \tag{3.3}$$

The limiting parameters $\theta_{j+\frac{1}{2}}$ are designed so that the modified numerical fluxes satisfy

$$u_{m,j} \le u^{n+1}_j = u^n_j - \lambda(\tilde{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j-\frac{1}{2}}) \le u_{M,j}. \tag{3.4}$$

For each $\theta_{j+\frac{1}{2}}$ limiting the numerical flux $\hat{H}^{rk}_{j+\frac{1}{2}}$, we are looking for parameters Λ_{-,I_j} and Λ_{+,I_j} so that when

$$\theta_{j+\frac{1}{2}} \in [0, \Lambda_{+,I_j}] \cap [0, \Lambda_{-,I_{j+1}}], \quad \text{all } j, \tag{3.5}$$

we have $u^{n+1}_j \in [u_{m,j}, u_{M,j}]$. Here, we include two parameter bounds Λ_{-,I_j} and Λ_{+,I_j} for the reason that one numerical flux affects two adjacent grid values.

We let

$$\Gamma^M_j := u_{M,j} - \left[u^n_j - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right] \ge 0,$$
$$\Gamma^m_j := u_{m,j} - \left[u^n_j - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right] \le 0.$$

To ensure $u_{m,j} \le u^{n+1}_j \le u_{M,j}$, it is sufficient to require that

$$\lambda\theta_{j-\frac{1}{2}}(\hat{H}^{rk}_{j-\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) - \lambda\theta_{j+\frac{1}{2}}(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j+\frac{1}{2}}) - \Gamma^M_j \le 0, \tag{3.6}$$

$$\lambda\theta_{j-\frac{1}{2}}(\hat{H}^{rk}_{j-\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) - \lambda\theta_{j+\frac{1}{2}}(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j+\frac{1}{2}}) - \Gamma^m_j \ge 0. \tag{3.7}$$

The discussion is case by case based on the signs of

$$F_{j\pm\frac{1}{2}} := \hat{H}^{rk}_{j\pm\frac{1}{2}} - \hat{h}_{j\pm\frac{1}{2}}.$$

1. To preserve the upper bound in (3.6),
   (a) If $F_{j-\frac{1}{2}} \le 0,\ F_{j+\frac{1}{2}} \ge 0,$

$$(\Lambda^M_{-,I_j}, \Lambda^M_{+,I_j}) = (1, 1).$$

(b) If $F_{j-\frac{1}{2}} \le 0,\ F_{j+\frac{1}{2}} < 0$,

$$(\Lambda^M_{-,I_j}, \Lambda^M_{+,I_j}) = \left(1, \min\left(1, \frac{\Gamma^M_j}{-\lambda F_{j+1/2}}\right)\right).$$

(c) If $F_{j-\frac{1}{2}} > 0,\ F_{j+\frac{1}{2}} \ge 0$,

$$(\Lambda^M_{-,I_j}, \Lambda^M_{+,I_j}) = \left(\min\left(1, \frac{\Gamma^M_j}{\lambda F_{j-1/2}}\right), 1\right).$$

(d) If $F_{j-\frac{1}{2}} > 0,\ F_{j+\frac{1}{2}} < 0$,

$$(\Lambda^M_{-,I_j}, \Lambda^M_{+,I_j}) = \left(\min\left(1, \frac{\Gamma^M_j}{\lambda F_{j-1/2} - \lambda F_{j+1/2}}\right), \min\left(1, \frac{\Gamma^M_j}{\lambda F_{j-1/2} - \lambda F_{j+1/2}}\right)\right).$$

2. Similarly, to preserve the lower bound in (3.7),

(a) If $F_{j-\frac{1}{2}} \ge 0,\ F_{j+\frac{1}{2}} \le 0$,

$$(\Lambda^m_{-,I_j}, \Lambda^m_{+,I_j}) = (1, 1).$$

(b) If $F_{j-\frac{1}{2}} \ge 0,\ F_{j+\frac{1}{2}} > 0$,

$$(\Lambda^m_{-,I_j}, \Lambda^m_{+,I_j}) = \left(1, \min\left(1, \frac{\Gamma^m_j}{-\lambda F_{j+1/2}}\right)\right).$$

(c) If $F_{j-\frac{1}{2}} < 0,\ F_{j+\frac{1}{2}} \le 0$,

$$(\Lambda^m_{-,I_j}, \Lambda^m_{+,I_j}) = \left(\min\left(1, \frac{\Gamma^m_j}{\lambda F_{j-1/2}}\right), 1\right).$$

(d) If $F_{j-\frac{1}{2}} < 0,\ F_{j+\frac{1}{2}} > 0$,

$$(\Lambda^m_{-,I_j}, \Lambda^m_{+,I_j}) = \left(\min\left(1, \frac{\Gamma^m_j}{\lambda F_{j-1/2} - \lambda F_{j+1/2}}\right), \min\left(1, \frac{\Gamma^m_j}{\lambda F_{j-1/2} - \lambda F_{j+1/2}}\right)\right).$$

Notice that the range of $\theta_{j+\frac{1}{2}}$ is required to ensure both the upper bound (3.6) and the lower bound (3.7) of numerical solutions in both cells I_j and I_{j+1}. Thus the locally defined limiting parameter is given by

$$\theta_{j+\frac{1}{2}} = \min\left\{\Lambda^M_{+,I_j}, \Lambda^M_{-,I_{j+1}}, \Lambda^m_{+,I_j}, \Lambda^m_{-,I_{j+1}}\right\}. \tag{3.8}$$

4. Local bounds and induced total variation stability

In Section 3, we provide basic steps for finding limiting parameters when local bounds are given. Such bounds have to be achievable in the sense of (3.3) in order to ensure the existence of the limiting parameters. Meanwhile, these bounds can not be too restrictive. Otherwise they will degenerate the high order accuracy. For example, in [1, 29], the bounds are set to be $u_{m,j} = \min(u^n_{j-1}, u^n_j)$, $u_{M,j} = \max(u^n_{j-1}, u^n_j)$ for the $f'(u) > 0$ case. As a result, the scheme is TVD and the accuracy is limited to at most second order.

In this section, we consider the problem (1.1) with $f'(u) > 0$. We first set the goal: numerically we would like to develop a fully discretized scheme of the type (2.5) such that

$$\mathrm{TV}(u^n) \le \mathrm{Var}(u_0) \text{ for all } n, \tag{4.1}$$

with the numerical total variation defined by (1.4). Its theoretical foundation is the total variation contractive property (1.3) of the entropy solution and the simple fact that

$$\mathrm{TV}([\ldots, g(x_{j-1}), g(x_j), g(x_{j+1}), \ldots]) \le \mathrm{Var}(g) \tag{4.2}$$

for any real-valued function $g(x)$ when $x_1 \le x_2 \le \cdots \le x_j \le \cdots \le x_{N-1} \le x_N$. The total variation bounded property (4.1) automatically implies a subsequence converging to a weak solution of the equation (1.1). Specifically, given $\mathrm{TV}(u^n) \le \mathrm{Var}(u_0)$, we would like to derive the total variation stability $\mathrm{TV}(u^{n+1}) \le \mathrm{Var}(u_0)$ when the local bounds are satisfied in (3.4) by the updated grid values. The critical steps of defining the local bounds $u_{m,j}, u_{M,j}$ and implementing the algorithm in the case of $f'(u) > 0$ include:

(1) **Calculate preliminary data by using the original high order schemes which might not be TVB in general:**

$$u_j^{rk} = u_j^n - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{H}^{rk}_{j-\frac{1}{2}}). \tag{4.3}$$

Roughly speaking, TVB stability ensures that the error from one-step approximation does not grow out of control over time. Here, u_j^{rk} provides raw data with high accuracy for a one-step approximation. This data will only be processed to produce the local bounds $u_{m,j}, u_{M,j}$ while the high accuracy information is passed along.

(2) **Combine u^{rk} with u^n to create an expanded vector with a particular ordering:**

$$v(u^n, u^{rk}) = [\ldots, u^n_{j-1}, u^{rk}_j, u^n_j, u^{rk}_{j+1}, u^n_{j+1}, \ldots].$$

Such an ordering is directly related to the assumption $f'(u) > 0$. It is obvious that $\mathrm{TV}(v(u^n, u^{rk})) \ge \mathrm{TV}(u^{rk})$. Thus it is sufficient to require that

$$\mathrm{TV}(v(u^n, u^{n+1})) \le \mathrm{Var}(u_0) \tag{4.4}$$

in order for $\mathrm{TV}(u^{n+1}) \le \mathrm{TV}(u_0(x))$. This requirement is the key to our approach and completely different from the popular Harten's sufficient condition (1.8) and its variation (1.9). In fact, the exact solution $u(x_j, t^{n+1})$ to (1.1) can be traced back to a point value $u(x_j^*, t^n)$, $x_{j-1} \le x_j^* \le x_j$ following characteristic lines assuming the solution is smooth. It is obvious that an expanded vector satisfies

$$\mathrm{TV}([\ldots, u(x_{j-1}, t^n), u(x_j^*, t^n), u(x_j, t^n), u(x_{j+1}^*, t^n), u(x_{j+1}, t^n), \ldots]) \le \mathrm{Var}(u_0)$$

even when rarefaction or shock solution forms in this region assuming $f'(u) > 0$.

(3) **Identify $u_{m,j}, u_{M,j}$.**

(a) Compute $\mathrm{TV}(v(u^n, u^{rk}))$ and if $\mathrm{TV}(v(u^n, u^{rk})) \le \mathrm{Var}(u_0)$, flux limiters are not needed. Otherwise we let

$$\mathrm{TV}_{inc} = \mathrm{TV}(v(u^n, u^{rk})) - \mathrm{Var}(u_0). \tag{4.5}$$

(b) Find the contribution of u_j^{rk} to the total variation incremental TV_{inc} in terms of

$$\mathrm{TV}_j = |u^n_{j-1} - u^{rk}_j| + |u^{rk}_j - u^n_j| - |u^n_j - u^n_{j-1}| \ge 0.$$

It is obvious, but important to notice that

$$\sum_j \mathrm{TV}_j = \mathrm{TV}(v(u^n, u^{rk})) - \mathrm{TV}(u^n) \geq \mathrm{TV}(v(u^n, u^{rk})) - \mathrm{Var}(u_0) = \mathrm{TV}_{inc}$$

since $\mathrm{TV}(u^n) \leq \mathrm{Var}(u_0)$.

(c) Calculate a proportional parameter $\eta_j = \frac{\mathrm{TV}_j}{\sum_j \mathrm{TV}_j}$ such that u_j^{rk} is modified by $\frac{1}{2}\eta_j \mathrm{TV}_{inc}$ as illustrated in Figure 1. We use u_j^* to denote the modified value. Such modified values satisfy the total variation bound

$$\mathrm{TV}(v(u^n, u^*)) = \mathrm{Var}(u_0). \tag{4.6}$$

A brief proof is given in Proposition 1. u^* is merely a high order approximation of the u_j^{rk}. It satisfies the total variation bound. However, it is not a conservative approximation.

(d) Define the local bounds

$$u_{m,j} = \min(u_{j-1}^n, u_j^*, u_j^n), \ u_{M,j} = \max(u_{j-1}^n, u_j^*, u_j^n). \tag{4.7}$$

These two local bounds are not necessarily local extrema, nor are they upper and lower bounds of the solution on an interval. They only serve as local bounds for the updated values u_j^{n+1} such that

$$\mathrm{TV}(v(u^n, u^{n+1})) \leq \mathrm{Var}(u_0)$$

when $u_j^{n+1} \in [u_{m,j}, u_{M,j}]$. A brief proof is given in Proposition 2.

(4) **Select the lower order monotone flux** $\hat{h}_{j+\frac{1}{2}} = f(u_j^n)$ and apply the flux-limiting method described in Section 3: Find the $\theta_{j+\frac{1}{2}}$'s such that $\tilde{H}_{j+\frac{1}{2}}^{rk} = \theta_{j+\frac{1}{2}}(\hat{H}_{j+\frac{1}{2}}^{rk} - \hat{h}_{j+\frac{1}{2}}) + \hat{h}_{j+\frac{1}{2}}$ satisfies

$$u_{j,m}^* \leq u_j^{n+1} = u_j^n - \lambda(\tilde{H}_{j+1/2}^{rk} - \tilde{H}_{j-1/2}^{rk}) \leq u_{j,M}^*. \tag{4.8}$$

We let $w_j = u_j^n - \lambda(\hat{h}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})$. The reason to pick the first order upwinding numerical flux is actually that $u_{j,m}^* \leq w_j \leq u_{j,M}^*$ since w_j is a convex combination of u_{j-1}^n and u_j^n. Therefore, the limiting parameters $\theta_{j+\frac{1}{2}}$'s satisfying the constraint (4.8) are guaranteed to exist.

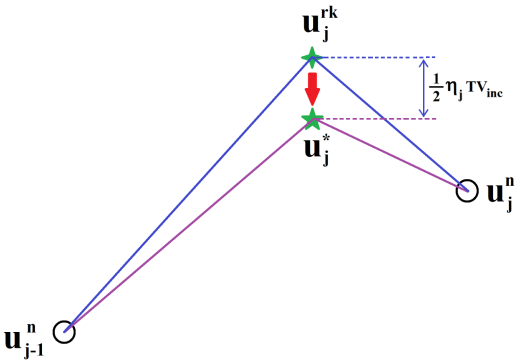


FIGURE 1. Modification of the reference value to satisfy the TVB property.

To this end, we have provided the key steps of ensuring the scheme to be TVB and detailed procedure to implement such an approach, especially for finding the local bounds. In this new scheme, the local bounds depend on global information, namely the total variation of the numerical values. However, the scheme itself is locally conservative as displayed by the local conservative formulation (3.4). As can also be observed from the above steps, this new approach does not introduce any tuning parameters. Now we prove the claim (4.6) and the claim in 3(d).

Proposition 1. *When* $\mathrm{TV}(v(u^n, u^{rk})) \geq \mathrm{Var}(u_0)$, u^* *satisfies* $\mathrm{TV}(v(u^n, u^*)) = \mathrm{Var}(u_0)$ *with* u^* *given in Figure* 1.

Proof. It is straightforward from Figure 1 that

$$\mathrm{TV}([u_{j-1}^n, u_j^*, u_j^n]) = \mathrm{TV}([u_{j-1}^n, u_j^{rk}, u_j^n]) - \eta_j \mathrm{TV}_{inc}.$$

Thus,

$$\sum_j \mathrm{TV}([u_{j-1}^n, u_j^*, u_j^n]) = \sum_j (\mathrm{TV}([u_{j-1}^n, u_j^{rk}, u_j^n]) - \eta_j \mathrm{TV}_{inc}).$$

Since $\sum_j \eta_j = 1$, the above equation can be written as

$$\mathrm{TV}(v(u^n, u^*)) = \mathrm{TV}(v(u^n, u^{rk})) - \mathrm{TV}_{inc} = \mathrm{Var}(u_0),$$

by the definition of TV_{inc} in (4.5). □

Proposition 2. *Given* $\mathrm{TV}(u^n) \leq \mathrm{Var}(u_0)$ *and local bounds defined in* (4.7)*, we have*

$$\mathrm{TV}(v(u^n, u^{n+1})) \leq \mathrm{Var}(u_0)$$

if $u_j^{n+1} \in [u_{m,j}, u_{M,j}]$ *for all* j*'s.*

Proof. Define a real-valued function $g_j(\gamma) = \mathrm{TV}([u_{j-1}^n, \gamma, u_j^n]) - |u_j^n - u_{j-1}^n|$ and $g_j(\gamma)$ is a nonnegative convex function of γ thanks to the triangular inequality. We have $g_j(u_{j-1}^n) = g_j(u_j^n) = 0$ and $g_j(u_j^*) = \mathrm{TV}([u_{j-1}^n, u_j^*, u_j^n]) - |u_j^n - u_{j-1}^n|$. If $u_j^{n+1} \in [u_{m,j}, u_{M,j}]$, it can be written as a convex combination of u_{j-1}^n, u_j^n and u_j^*. Therefore $g_j(u_j^{n+1}) \leq \max(g_j(u_{j-1}^n), g_j(u_j^*), g_j(u_j^n)) = g_j(u_j^*)$ since g_j is a nonnegative convex function. We have

$$\sum_j g_j(u_j^{n+1}) \leq \sum g_j(u_j^*),$$

which is the same as

$$\mathrm{TV}(v(u^n, u^{n+1})) - \mathrm{TV}(u^n) \leq \mathrm{Var}(u_0) - \mathrm{TV}(u^n).$$

Therefore

$$\mathrm{TV}(v(u^n, u^{n+1})) \leq \mathrm{Var}(u_0)$$

thanks to Proposition 1 and the definition of g and u_j^*. □

4.1. **The algorithm for general $f(u)$.** For general $f(u)$, we apply an indirect approach based on the Lax-Friedrich flux-splitting method. On a time interval $[0, \Delta t]$, we solve an equation

$$p_t + \frac{1}{2}(f(p) + \alpha p)_x = 0,\ p(x, 0) = u_0(x), \tag{4.9}$$

followed by solving a second equation

$$q_t + \frac{1}{2}(f(q) - \alpha q)_x = 0,\ q(x, 0) = p(x, \Delta t). \tag{4.10}$$

Theoretically, we have $q(x, \Delta t) = u(x, \Delta t)$, the solution to the original problem (1.1) with general $f(u)$. Here, $\alpha = \max_x |f'(u_0(x))|$. Numerically, equation (4.9) can be solved with the TVB high order finite difference method using the steps described above. The second equation (4.10) can also be solved with a TVB high order finite difference method with a little twist of the algorithm:

(1) Choose $\hat{h}_{j+\frac{1}{2}} = \frac{1}{2}(f(q_{j+1}) - \alpha q_{j+1})$ since the upwinding is from the right.
(2) The total variation criterion can be written as

$$\mathrm{TV}(v(q^n, q^{n+1})) \leq \mathrm{Var}(u_0) \tag{4.11}$$

with the new ordering of the expanded vector

$$v(q^n, q^{n+1}) = [\ldots, q^n_{j-1}, q^{n+1}_{j-1}, q^n_j, q^{n+1}_j, q^n_{j+1}, \ldots].$$

This provides a practical approach.

5. The local truncation error analysis

In the following, we would like to show that the third order of accuracy of both spatial and temporal discretization is maintained with the newly proposed TVB flux limiters when the solution is smooth enough. We should note that it is well known that the convergence analysis of finite difference approximation of the scalar conservation laws requires both total variation stability and entropy stability. In the absence of entropy stability, it is only realistic to perform the local truncation error analysis of the proposed approach.

As there is no rigorous error analysis for finite difference schemes with high order polynomial reconstruction by introducing the sliding function $H(x)$ and RK method for hyperbolic conservation laws (1.1), this task becomes quite challenging. To give a brief sketch of what can possibly be proven, we have to make some assumptions on the error of the original high order finite difference scheme. Notice that these assumptions have been numerically verified extensively, but have not been rigorously proven. Meanwhile, in our approach, the first step is to find the local bounds. We first assume that

$$u^{rk}_j - u(x_j, t^{n+1}) = \mathcal{O}(\Delta x^3 + \Delta t^3), \tag{5.1}$$

$$u^{*}_j - u(x_j, t^{n+1}) = \mathcal{O}(\Delta x^3 + \Delta t^3), \tag{5.2}$$

in order to prove the following theorem.

Theorem 1. *Consider solving the scalar conservation law* (1.1) *with* $f'(u) > 0$ *using a third order finite difference spatial discretization and a third order RK time discretization with the scheme written in the form of equation* (2.5)*. Assume the global error*

$$e^n_j = |u^n_j - u(x_j, t^n)| = \mathcal{O}(\Delta x^3 + \Delta t^3),\ \forall n, j,\ \textit{and}\ \mathrm{TV}(u^n) \leq \mathrm{Var}(u_0). \tag{5.3}$$

When the proposed TVB flux limiter is applied to the numerical fluxes $\hat{H}^{rk}_{j\pm\frac{1}{2}}$ *in equation* (2.5)*, and the lower order flux* $\hat{h}_{j+\frac{1}{2}}$ *in equation* (3.1) *is set to be the upwinding flux* $\hat{h}_{j+\frac{1}{2}} = f(u^n_j)$*, then*

$$|\hat{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j+\frac{1}{2}}| = \mathcal{O}(\Delta x^3 + \Delta t^3) \quad \forall j, \tag{5.4}$$

with $\lambda \max_u |f'(u)| \leq CFL < 1$*, where* $\lambda = \Delta t/\Delta x$*.*

Proof. We denote $I_{j-\frac{1}{2}} = [x_{j-1}, x_j]$. We only consider the limiters for the upper bound case since it is similar for the lower bound case. The theorem is proven via discussing four cases separately. Noticing that cases (b) and (c) are similar, so we only give the proof for the cases (a), (b), and (d).

Case (a) No limiters are introduced, i.e., $\tilde{H}^{rk}_{j+\frac{1}{2}} = \hat{H}^{rk}_{j+\frac{1}{2}}$.

Case (d) This is the case when $F_{j-\frac{1}{2}} > 0$ and $F_{j+\frac{1}{2}} < 0$. Similar to Case (a), we only need to consider the difference $|\hat{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j+\frac{1}{2}}|$ when $\frac{\Gamma^M_j}{\lambda\left(F_{j-1/2}-F_{j+1/2}\right)} < 1$.

Noticing that

$$\Gamma^M_j - \lambda\left(F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}}\right) = u_{M,j} - \left(u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{H}^{rk}_{j-\frac{1}{2}})\right) = u_{M,j} - u^{rk}_j$$

and it implies that

$$\hat{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j+\frac{1}{2}} = \frac{\lambda\left(F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}}\right) - \Gamma^M_j}{\lambda(F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}})} F_{j+\frac{1}{2}} = \frac{u^{rk}_j - u_{M,j}}{\lambda(F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}})} F_{j+\frac{1}{2}}. \tag{5.5}$$

It is sufficient to show that

$$\frac{F_{j+\frac{1}{2}}}{\lambda(F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}})}\left(u^{rk}_j - u_{M,j}\right) = \mathcal{O}(\Delta t^3 + \Delta x^3) \tag{5.6}$$

when $u_{M,j} < u^{rk}_j$. Since $u_{M,j} = \max\{u^n_{j-1}, u^*_j, u^n_j\}$, we have $u_{M,j} \geq u^{rk}_j + \mathcal{O}(\Delta t^3 + \Delta x^3)$ because of assumptions (5.1) and (5.2). Therefore we have $u_{M,j} - u^{rk}_j = \mathcal{O}(\Delta t^3 + \Delta x^3)$ and (5.6) noticing that $-\frac{1}{\lambda} < \frac{1}{\lambda} \cdot \frac{F_{j+\frac{1}{2}}}{F_{j-\frac{1}{2}} - F_{j+\frac{1}{2}}} < 0$.

Case (b) This is the case when $F_{j-\frac{1}{2}} \leq 0$ and $F_{j+\frac{1}{2}} < 0$. Similarly, we only need to consider the difference $|\hat{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j+\frac{1}{2}}|$ when $\frac{\Gamma^M_j}{-\lambda F_{j+1/2}} < 1$. Noticing that

$$\Gamma^M_j + \lambda F_{j+\frac{1}{2}} = u_{M,j} - \left(u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right)$$

and it implies that

$$\hat{H}^{rk}_{j+\frac{1}{2}} - \tilde{H}^{rk}_{j+\frac{1}{2}} = \frac{\Gamma^M_j + \lambda F_{j+\frac{1}{2}}}{\lambda} = \frac{u_{M,j} - \left(u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right)}{\lambda}. \tag{5.7}$$

Then it is sufficient to show that

$$u_{M,j} - \left(u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right) = \mathcal{O}(\Delta t^3 + \Delta x^3) \tag{5.8}$$

when $u_{M,j} - \left(u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}})\right) < 0$. The first order upwinding flux can be written as

$$\hat{h}_{j-\frac{1}{2}} = f(u^n_{j-1}) = f(u(x_{j-1}, t^n)) + \mathcal{O}(\Delta t^3 + \Delta x^3). \tag{5.9}$$

By introducing a sliding function H, the third order RK flux can be written as

$$\hat{H}^{rk}_{j+\frac{1}{2}} = \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} H(x_{j+\frac{1}{2}}, t)dt + \mathcal{O}(\Delta t^3). \tag{5.10}$$

Apply Simpson's Rule to the integral, we have

$$\hat{H}^{rk}_{j+\frac{1}{2}} = \frac{1}{6}H(x_{j+\frac{1}{2}}, t^n) + \frac{4}{6}H(x_{j+\frac{1}{2}}, t^{n+\frac{1}{2}}) + \frac{1}{6}H(x_{j+\frac{1}{2}}, t^{n+1}) + \mathcal{O}(\Delta t^3) \tag{5.11}$$

with $t^{n+\frac{1}{2}} = t^n + \frac{\Delta t}{2}$. The sliding function $H(x,t)$ can be replaced by an expansion [22]:

$$\begin{aligned} H(x_{j+\frac{1}{2}},t) &= f(u(x_{j+\frac{1}{2}},t)) - \tfrac{\Delta x^2}{24} f_{xx}(u(x_{j+\frac{1}{2}},t)) + \mathcal{O}(\Delta x^4) \\ &= -\tfrac{1}{24} f(u(x_{j-\frac{1}{2}},t)) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}},t)) - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}},t)) + \mathcal{O}(\Delta x^4). \end{aligned}$$

By substitution, the RK flux can be rewritten as

$$\begin{aligned} \hat{H}^{rk}_{j+\frac{1}{2}} = {} & \frac{1}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}},t^{n+1})) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}},t^{n+1})) - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}},t^{n+1}) \Big] \\ & + \frac{4}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}},t^{n+\frac{1}{2}})) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}},t^{n+\frac{1}{2}})) - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}},t^{n+\frac{1}{2}}) \Big] \\ & + \frac{1}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}},t^{n})) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}},t^{n})) - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}},t^{n}) \Big] \\ & + \mathcal{O}(\Delta t^3 + \Delta x^4). \end{aligned}$$

Following the characteristics, we have

$$\begin{aligned} \hat{H}^{rk}_{j+\frac{1}{2}} = {} & \frac{1}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}} - \lambda_{11}\Delta x, t^n)) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}} - \lambda_{21}\Delta x, t^n)) \\ & \qquad - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}} - \lambda_{31}\Delta x, t^n) \Big] \\ & + \frac{4}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}} - \lambda_{12}\Delta x, t^n)) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}} - \lambda_{22}\Delta x, t^n)) \\ & \qquad - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}} - \lambda_{32}\Delta x, t^n) \Big] \\ & + \frac{1}{6}\Big[-\tfrac{1}{24} f(u(x_{j-\frac{1}{2}}, t^n)) + \tfrac{13}{12} f(u(x_{j+\frac{1}{2}}, t^n)) - \tfrac{1}{24} f(u(x_{j+\frac{3}{2}}, t^n) \Big] \\ & + \mathcal{O}(\Delta t^3 + \Delta x^4), \end{aligned} \tag{5.12}$$

where the coefficients satisfy the corresponding equation:

$$\begin{aligned} &\lambda_{11} = \lambda f'(u(x_{j-\frac{1}{2}} - \lambda_{11}\Delta x, t^n)), && \lambda_{12} = \tfrac{\lambda}{2} f'(u(x_{j-\frac{1}{2}} - \lambda_{12}\Delta x, t^n)), \\ &\lambda_{21} = \lambda f'(u(x_{j+\frac{1}{2}} - \lambda_{21}\Delta x, t^n)), && \lambda_{22} = \tfrac{\lambda}{2} f'(u(x_{j+\frac{1}{2}} - \lambda_{22}\Delta x, t^n)), \\ &\lambda_{31} = \lambda f'(u(x_{j+\frac{3}{2}} - \lambda_{31}\Delta x, t^n)), && \lambda_{32} = \tfrac{\lambda}{2} f'(u(x_{j+\frac{3}{2}} - \lambda_{32}\Delta x, t^n)). \end{aligned} \tag{5.13}$$

In order to prove (5.8), we introduce some point x_p on the subinterval $I_{j-\frac{1}{2}}$ by

$$x_p := x_j - z\Delta x \tag{5.14}$$

with the parameter $z \in [0,1]$. We write $u(x,t^n)$ simply as $u(x)$, denote $u_p := u(x_p)$, $u'_p := u'(x_p)$, $u''_p := u''(x_p)$ and introduce some notations for the convenience of description:

$$\begin{aligned} &\lambda_0 := \lambda f'(u_p), \\ &\lambda_1 := \tfrac{1}{2}\lambda f''(u_p) u'_p, \\ &\lambda_2 := \tfrac{1}{2}\big[f'(u_p) u''_p + f''(u_p)(u'_p)^2 \big]. \end{aligned} \tag{5.15}$$

Now we try to find the Taylor expansions around $x = x_p$ for

$$\mathbf{LHS} := u^n_j - \lambda(\hat{H}^{rk}_{j+\frac{1}{2}} - \hat{h}_{j-\frac{1}{2}}) \tag{5.16}$$

by first finding the Taylor expansions around $x = x_p$ for the following quantities:

$$\begin{aligned} u(x) &= u_p + u_p'(x-x_p) + \tfrac{1}{2}u_p''(x-x_p)^2 + \mathcal{O}((x-x_p)^3), \\ f(u(x)) &= f(u_p) + f'(u_p)u_p'(x-x_p) + \lambda_2(x-x_p)^2 + \mathcal{O}((x-x_p)^3), \\ f'(u(x)) &= f'(u_p) + f''(u_p)u_p'(x-x_p) + \mathcal{O}((x-x_p)^2). \end{aligned} \tag{5.17}$$

Under assumption (5.3), we have

$$\begin{aligned} u_j^n &= u_p + u_p' z\Delta x + \tfrac{1}{2}u_p'' z^2\Delta x^2 + \mathcal{O}(\Delta t^3 + \Delta x^3), \\ \hat{h}_{j-\frac{1}{2}} &= f(u_p) + f'(u_p)u_p'(z-1)\Delta x + \lambda_2(z-1)^2\Delta x^2 + \mathcal{O}(\Delta t^3 + \Delta x^3). \end{aligned} \tag{5.18}$$

To obtain a full expansion of the RK flux in (5.12), for each fixed $i = 1, 2, 3$, we let

$$\begin{aligned} \lambda_{i1} &= \lambda_0 + \eta_{i1}\Delta x + \mathcal{O}(\Delta x^2), \\ \lambda_{i2} &= \tfrac{\lambda_0}{2} + \eta_{i2}\Delta x + \mathcal{O}(\Delta x^2), \end{aligned}$$

and substitute into (5.13) to determine η_{i1} and η_{i2}. We have

$$\begin{aligned} \lambda_{i1} &= \lambda_0 + 2\lambda_1\big(z + \tfrac{2i-3}{2} - \lambda_0\big)\Delta x + \mathcal{O}(\Delta x^2), \\ \lambda_{i2} &= \tfrac{\lambda_0}{2} + \lambda_1\big(z + \tfrac{2i-3}{2} - \tfrac{\lambda_0}{2}\big)\Delta x + \mathcal{O}(\Delta x^2), \end{aligned} \tag{5.19}$$

for $i = 1, 2, 3$ with the parameter z defined in (5.14). We now substitute (5.19) into (5.12) to obtain the Taylor expansion of RK flux

$$\begin{aligned} \hat{H}_{j+\frac{1}{2}}^{rk} = f(u_p) &+ f'(u_p)u_p'(z + \tfrac{1}{2} - \tfrac{\lambda_0}{2})\Delta x - \lambda_1 f'(u_p)u_p'\big(z + \tfrac{1}{2} - \tfrac{2}{3}\lambda_0\big)\Delta x^2 \\ &+ \lambda_2\Big[\big(\tfrac{1}{6} - \tfrac{1}{2}\lambda_0 + \tfrac{1}{3}\lambda_0^2\big) + (1-\lambda_0)z + z^2\Big]\Delta x^2 + \mathcal{O}(\Delta t^3 + \Delta x^3). \end{aligned} \tag{5.20}$$

Combine (5.18) and (5.20) with (5.16), we have

$$\mathbf{LHS} = u_p + u_p'\Delta x s_1 + \tfrac{u_p''}{2}\Delta x^2 s_3 + \lambda_1 u_p'\Delta x^2 s_2 + \mathcal{O}(\Delta t^3 + \Delta x^3), \tag{5.21}$$

where

$$\begin{aligned} s_1 &:= \big(-\tfrac{3}{2}\lambda_0 + \tfrac{1}{2}\lambda_0^2\big) + z, \\ s_2 &:= \big(\tfrac{5}{6} + \lambda_0 - \lambda_0^2\big) + (2\lambda_0 - 3)z, \\ s_3 &:= \big(\tfrac{5}{6}\lambda_0 + \tfrac{1}{2}\lambda_0^2 - \tfrac{1}{3}\lambda_0^3\big) + (\lambda_0^2 - 3\lambda_0)z + z^2. \end{aligned} \tag{5.22}$$

It is necessary to point out that $0 \le s_3 \le 1$ and $s_1 + \sqrt{s_3} \ge 0,\ s_1 - \sqrt{s_3} \le 0$ (see the proofs in Lemmas 1 and 2). Therefore, (5.21) can be rewritten as two different forms:

$$\mathbf{LHS} = u(x_p - \sqrt{s_3}\Delta x) + u_p'\Delta x(s_1 + \sqrt{s_3}) + \lambda_1 u_p'\Delta x^2 s_2 + \mathcal{O}(\Delta t^3 + \Delta x^3), \tag{5.23}$$

$$\mathbf{LHS} = u(x_p + \sqrt{s_3}\Delta x) + u_p'\Delta x(s_1 - \sqrt{s_3}) + \lambda_1 u_p'\Delta x^2 s_2 + \mathcal{O}(\Delta t^3 + \Delta x^3). \tag{5.24}$$

On the sufficiently refined interval $I_{j-\frac{1}{2}}$, we consider three subcases. Introduce $x_j^* \in I_{j-\frac{1}{2}}$ such that $u(x_j, t^{n+1}) = u(x_j^*)$, then the basic assumption (5.2) says

$$u_j^* = u(x_j^*) + \mathcal{O}(\Delta t^3 + \Delta x^3). \tag{5.25}$$

Let $\lambda^* := \lambda f'(u(x_j^*)) < 1$ and follow the characteristics, we have

$$x_j^* = x_j - \lambda^*\Delta x. \tag{5.26}$$

Case (I) If there is a local maximum inside the cell of $I_{j-\frac{1}{2}}$, we choose x_p so that $u(x_p)$ is the local maximum, then $u'_p = 0$. From (5.17), we obtain

$$\begin{aligned}\lambda^* &= \lambda f'(u(x_j^*)) = \lambda f'(u_p) + \lambda f''(u_p)u'_p(x_j^* - x_p) + \mathcal{O}(\Delta x^2)\\ &= \lambda_0 + \mathcal{O}(\Delta x^2).\end{aligned} \tag{5.27}$$

Then for u_j^* in (5.25), apply the Taylor expansion at $x = x_j - \lambda_0\Delta x$,

$$\begin{aligned}u_j^* &= u(x_j - \lambda^*\Delta x) + \mathcal{O}(\Delta t^3 + \Delta x^3)\\ &= u(x_j - \lambda_0\Delta x) + u'(x_j - \lambda_0\Delta x)(\lambda_0 - \lambda^*)\Delta x + \mathcal{O}(\Delta t^3 + \Delta x^3)\\ &= u(x_j - \lambda_0\Delta x) + \mathcal{O}(\Delta t^3 + \Delta x^3).\end{aligned} \tag{5.28}$$

Therefore, under assumptions (5.1)–(5.3),

$$u_{M,j} = \max\{u(x_{j-1}), u(x_j - \lambda_0\Delta x), u(x_j)\} + \mathcal{O}(\Delta t^3 + \Delta x^3). \tag{5.29}$$

Since $I_{j-\frac{1}{2}}$ is a sufficiently refined interval, u can be nondecreasing on $[x_{j-2}, x_p]$ and nonincreasing on $[x_p, x_{j+1}]$. From the equivalent forms (5.23) and (5.24), in order to prove (5.8), it is sufficient to prove

$$u(x_p - \sqrt{s_3}\Delta x) \le u_{M,j} \text{ or } u(x_p + \sqrt{s_3}\Delta x) \le u_{M,j}.$$

• If $\lambda_0 \in [0, \frac{1}{2}]$, define

$$F(\lambda_0, z) := \max\left(z - 1 + \sqrt{s_3}, \lambda_0 - z + \sqrt{s_3}\right).$$

From Lemma 4, $F(\lambda_0, z) \ge 0$, which directly implies

$$x_p - \sqrt{s_3}\Delta x \le x_{j-1} \text{ or } x_p + \sqrt{s_3}\Delta x \ge x_j - \lambda_0\Delta x.$$

Then

$$u(x_p - \sqrt{s_3}\Delta x) \le u(x_{j-1}) \text{ or } u(x_p + \sqrt{s_3}\Delta x) \le u(x_j - \lambda_0\Delta x).$$

• If $\lambda_0 \in [\frac{1}{2}, 1)$, define

$$G(\lambda_0, z) := \max\left(z - \lambda_0 + \sqrt{s_3}, -z + \sqrt{s_3}\right).$$

From Lemma 4, $G(\lambda_0, z) \ge 0$, which directly implies

$$x_p - \sqrt{s_3}\Delta x \le x_j - \lambda_0\Delta x \text{ or } x_p + \sqrt{s_3}\Delta x \ge x_j.$$

Then

$$u(x_p - \sqrt{s_3}\Delta x) \le u(x_j - \lambda_0\Delta x) \text{ or } u(x_p + \sqrt{s_3}\Delta x) \le u(x_j).$$

Case (II) If u is decreasing on the interval $I_{j-\frac{1}{2}}$, then $u'_p \le 0$ and

$$\begin{aligned}u_{M,j} &= \max\{u(x_{j-1}), u(x_j^*), u(x_j)\} + \mathcal{O}(\Delta t^3 + \Delta x^3)\\ &= u(x_{j-1}) + \mathcal{O}(\Delta t^3 + \Delta x^3).\end{aligned} \tag{5.30}$$

Noticing $s_1 + \sqrt{s_3} \ge 0$, we apply the equivalent form (5.23) and choose $x_p = x_j$, i.e., $z = 0$. With these notations, $s_3(\lambda_0) = \frac{1}{6}\lambda_0(1 + \lambda_0)(5 - 2\lambda_0)$ is strictly increasing function with the range $s_3([0, 1)) = [0, 1)$. It is easy to verify that s_3 is upper bounded away from 1 given $CFL < 1$. Thus,

$$u(x_p - \sqrt{s_3}\Delta x) \le u(x_{j-1}).$$

To prove (5.8), it is sufficient to show

$$u(x_p - \sqrt{s_3}\Delta x) + u'_p\Delta x(s_1 + \sqrt{s_3}) + u'_p\Delta x^2\tilde{s}_2 \le u(x_{j-1}) \quad \text{or} \quad u'_p = \mathcal{O}(\Delta x), \tag{5.31}$$

where $\tilde{s}_2 := \lambda_1 s_2$. We assume

$$u(x_p - \sqrt{s_3}\Delta x) + c\Delta x^2 > u(x_{j-1}), \tag{5.32}$$

where $c := |u_p'\lambda_1 s_2|$. By using the Mean Value Theorem, there exists a point $x^{\sharp 2} \in (x_{j-1}, x_p - \sqrt{s_3}\Delta x)$ s.t. $u'(x^{\sharp 2}) \le 0$ and

$$u(x_{j-1}) - u(x_p - \sqrt{s_3}\Delta x) = u'(x^{\sharp 2})(\sqrt{s_3} - 1)\Delta x.$$

Then the inequality (5.32) follows:

$$-\frac{c}{1-\sqrt{s_3}}\Delta x < u'(x^{\sharp 2}) \le 0. \tag{5.33}$$

Therefore, we get $u_p' = \mathcal{O}(\Delta x)$.

Case (III) If $u(x,t)$ is increasing in $I_{j-\frac{1}{2}}$, then $u_p' \ge 0$ and

$$\begin{aligned} u_{M,j} &= \max\{u(x_{j-1}), u(x_j^*), u(x_j)\} + \mathcal{O}(\Delta t^3 + \Delta x^3) \\ &= u(x_j) + \mathcal{O}(\Delta t^3 + \Delta x^3). \end{aligned} \tag{5.34}$$

Noticing $s_1 - \sqrt{s_3} \le 0$, we apply the equivalent form (5.24) and choose $x_p = x_{j-1}$, i.e., $z = 1$. A similar proof for (5.8) holds. To this end, the proof is completed. □

Here, we provide lemmas related to the proof of Theorem 1.

Lemma 1. *For $(\lambda_0, z) \in [0,1] \times [0,1]$, the function $s_3(\lambda_0, z)$ defined in* (5.22) *satisfies*

$$0 \le s_3(\lambda_0, z) \le 1.$$

Proof. For any fixed $\lambda_0 \in [0,1]$, $s_3(\lambda_0, z)$ can be seen as a polynomial of degree two about $z \in [0,1]$, whose axis of symmetry is given by

$$z = \tfrac{3\lambda_0 - \lambda_0^2}{2} \in [0,1].$$

So the minimum over $(\lambda_0, z) \in [0,1] \times [0,1]$ is

$$\min_{\lambda_0, z} s_3(\lambda_0, z) = s_3(\lambda_0, \tfrac{3\lambda_0 - \lambda_0^2}{2}) = \frac{1}{12}\lambda_0(1-\lambda_0)(2-\lambda_0)(5-3\lambda_0) \ge 0,$$

and the maximum over $(\lambda_0, z) \in [0,1] \times [0,1]$ is

$$\begin{aligned} \max_{\lambda_0, z} s_3(\lambda_0, z) &= \max\{s_3(\lambda_0, 0), s_3(\lambda_0, 1)\} \\ &= \max\left\{\tfrac{1}{6}\lambda_0(1+\lambda_0)(5-2\lambda_0), \tfrac{1}{6}(1-\lambda_0)(2-\lambda_0)(3-2\lambda_0)\right\} \le 1. \end{aligned}$$

□

Lemma 2. *For $(\lambda_0, z) \in [0,1] \times [0,1]$, the functions s_1 and s_3 defined in* (5.22) *satisfy $s_1 + \sqrt{s_3} \ge 0$ and $s_1 - \sqrt{s_3} \le 0$.*

Proof. Clearly, the fact

$$(s_1 + \sqrt{s_3})(s_1 - \sqrt{s_3}) = s_1^2 - s_3 = -\frac{1}{12}\lambda_0(1-\lambda_0)(2-\lambda_0)(5-3\lambda_0) \le 0$$

yields $s_1 + \sqrt{s_3} \ge 0$ and $s_1 - \sqrt{s_3} \le 0$. □

Lemma 3. *For each fixed $\lambda_0 \in [0,1]$, the function $h_1(z) := z + \sqrt{s_3}$ is increasing about $z \in [0,1]$ and $h_2(z) := -z + \sqrt{s_3}$ is decreasing about $z \in [0,1]$.*

Proof. For each fixed $\lambda_0 \in [0,1]$, notice that $s_3'(z) = 2s_1$. Differentiating $h_1(z)$ and $h_2(z)$ with respect to z and use the Lemma 2, we have

$$h_1'(z) = 1 + \frac{s_1}{\sqrt{s_3}} \geq 0, \qquad h_2'(z) = -1 + \frac{s_1}{\sqrt{s_3}} \leq 0,$$

which implies that $h_1(z)$ is increasing and $h_2(z)$ is decreasing. □

Lemma 4. *We have the following results:*

(1) *For* $(\lambda_0, z) \in [0, \frac{1}{2}] \times [0,1]$,

$$F(\lambda_0, z) := \max\big(-1 + z + \sqrt{s_3}, \lambda_0 - z + \sqrt{s_3}\big) \geq 0.$$

(2) *For* $(\lambda_0, z) \in [\frac{1}{2}, 1] \times [0,1]$,

$$G(\lambda_0, z) := \max\big(-\lambda_0 + z + \sqrt{s_3}, -z + \sqrt{s_3}\big) \geq 0.$$

Proof. (1) For each fixed $\lambda_0 \in [0, \frac{1}{2}]$, $F(\lambda_0, z)$ is a continuous function about $z \in [0,1]$ and

$$F(\lambda_0, z) = \begin{cases} h_2(z) + \lambda_0, & \text{if } \ 0 \leq z \leq \frac{\lambda_0+1}{2}, \\ h_1(z) - 1, & \text{if } \ \frac{\lambda_0+1}{2} \leq z \leq 1. \end{cases}$$

From Lemma 3, F is decreasing on $z \in [0, \frac{\lambda_0+1}{2}]$ and increasing on $z \in [\frac{\lambda_0+1}{2}, 1]$. Then F obtains its minimum at $z = \frac{\lambda_0+1}{2} \in [\frac{1}{2}, \frac{3}{4}]$,

$$F(\lambda_0, \tfrac{\lambda_0+1}{2}) = h_1(\tfrac{\lambda_0+1}{2}) - 1 = \sqrt{s_3(\lambda_0, \tfrac{\lambda_0+1}{2})} - (1 - \tfrac{\lambda_0+1}{2}).$$

It holds true since $s_3(\lambda_0, \frac{\lambda_0+1}{2}) - (1 - \frac{\lambda_0+1}{2})^2 = \frac{1}{6}\lambda_0(1-\lambda_0)(2-\lambda_0) \geq 0$.

(2) For each fixed $\lambda_0 \in [\frac{1}{2}, 1]$, $G(\lambda_0, z)$ is a continuous function about $z \in [0,1]$, and

$$G(\lambda_0, z) = \begin{cases} h_2(z), & \text{if } \ 0 \leq z \leq \frac{\lambda_0}{2}, \\ h_1(z) - \lambda_0, & \text{if } \ \frac{\lambda_0}{2} \leq z \leq 1. \end{cases}$$

From Lemma 3, G is decreasing on $z \in [0, \frac{\lambda_0}{2}]$ and increasing on $z \in [\frac{\lambda_0}{2}, 1]$. Then G obtains its minimum at $z = \frac{\lambda_0}{2} \in [\frac{1}{4}, \frac{1}{2}]$,

$$G(\lambda_0, \tfrac{\lambda_0}{2}) = h_2(\tfrac{\lambda_0}{2}) = \sqrt{s_3(\lambda_0, \tfrac{\lambda_0}{2})} - \tfrac{\lambda_0}{2}.$$

It holds true since $s_3(\lambda_0, \frac{\lambda_0}{2}) - (\frac{\lambda_0}{2})^2 = \frac{1}{6}\lambda_0(1-\lambda_0)(5-\lambda_0) \geq 0$. □

The proof provided here relies on assumptions (5.1) and (5.2). Assumption (5.1) is rather acceptable from checking the local truncation error. However, assumption (5.2) is not as obviously true and difficult to prove even though the numerical results provided in Section 6 seem to be consistent with this assumption. One alternative approach that allows for theoretical justification of assumption (5.2) is that

Remark 2. In step (4)(b) of the algorithm described in Section 4, a slight change can be made. First, we assume that S is the number of local extrema of the initial data. Among all (assume the number is P) the grid values that contribute to the incremental of total variation, we find $P - S$ ones with the smallest contribution and find u_j^* closest to u_j^{rk} while $|u_{j-1}^n - u_j^*| + |u_j^* - u_j^n| - |u_j^n - u_{j-1}^n| = 0$. The rest S values of u_j^* can be found similar to that described in (4)(b). However, it will be problematic to generalize such an alternative approach to higher dimensions.

6. Numerical simulations

In this section, we present several numerical examples for the proposed TVB flux limiters for high order finite difference schemes with high order RK time discretization. Through the tests, we would like to examine the performance of the new methods on accuracy, effect of bounded total variation around discontinuities, and other issues such as the scheme's performance when total variation bounds are not sharply estimated.

6.1. **Accuracy test.** For the accuracy test, we consider three types of reconstruction: fixed stencil reconstruction (one grid biased toward the upwinding direction, denoted as Fix3 in the tables and figures.); third order ENO reconstruction (ENO3); fifth order WENO reconstruction (WENO5). For stability, the CFL condition $\frac{\Delta t}{\Delta x}\max_u |f'(u)| \le \text{CFL}$ is applied to all tests. For WENO5 reconstruction, we choose Δt such that $\Delta t \max_u |f'(u)| = \text{CFL}(\Delta x)^{5/3}$ for the purpose of showing fifth order accuracy.

Example 5. Consider the 1D linear equation,

$$\begin{aligned} u_t + u_x &= 0, \\ u(x,0) &= u_0(x), \end{aligned} \tag{6.1}$$

on $[0, 2\pi]$ with periodic boundary conditions. We test $u_0(x) = \frac{4}{\pi}\arctan(\sin(x))$ for Fix3 and $u_0(x) = \sin(x)$ for ENO3 and WENO5 at time $T = 5$ with the same initial total variation $\text{Var}(u_0) = 4$.

For the problem (6.1), the total variation of the exact solution at any time T equals the initial total variations $\text{Var}(u_0) = 4$. We list the L^1 and L^∞ errors, orders of accuracy and the numerical total variations at $T = 5$ and CFL$= 0.9$ in Tables 1 and 2 with Fix3, ENO3 and WENO5 reconstructions. It can be observed from the tables that the orders for Fix3, ENO3, and WENO5 schemes with TVB flux limiters are consistent with the designed order of accuracy and the numerical total variations at $T = 5$ are bounded by $\text{Var}(u_0)$.

Example 6. Consider the 1D Burgers' equation

$$\begin{aligned} u_t + (\tfrac{u^2}{2})_x &= 0, \\ u(x,0) &= u_0(x), \end{aligned} \tag{6.2}$$

on $[0, 2\pi]$ with periodic boundary conditions. We test $u_0(x) = 1.1 + \frac{4}{\pi}\arctan(\sin(x))$ for Fix3 and $u_0(x) = 1.1 + \sin(x)$ (thus $f'(u) > 0$) for ENO3 and WENO5 at time $T = 0.5$ with the same initial total variations $\text{Var}(u_0) = 4$.

For these tests, the total variations of the entropy solutions at $T = 0.5$ equal the initial total variation $\text{Var}(u_0) = 4$. In Tables 3 and 4, we can observe that L^1 and L^∞ errors, orders and the numerical total variations at $T = 0.5$ for the finite difference schemes with Fix3 ($CFL = 0.9$), ENO3 and WENO5 ($CFL = 0.5$) reconstructions, respectively. The designed orders of accuracy for Fix3, ENO3 and WENO5 schemes are not reduced by the TVB flux limiters.

6.2. **Discontinuous solution.** In this part of the test, we consider the piecewise-constant initial condition

$$u_0(x) = \tfrac{1}{2}\text{sign}\left(\sin\left(\tfrac{2\pi}{3}x - \tfrac{\pi}{6}\right) + \tfrac{1}{2}\right) + \tfrac{1}{2}\text{sign}\left(\sin\left(\tfrac{2\pi}{3}x + \tfrac{\pi}{6}\right) - \tfrac{1}{2}\right) \tag{6.3}$$

at $T = 0.5$, $N = 60$, CFL$= 0.99$ and $\text{TV}(u_0) = 4$.

TABLE 1. Example 5: L^1, L^∞ errors and orders of accuracy for Fix3 reconstruction without and with TVB flux limiters.

	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
NO flux limiters	40	3.84E-02	–	1.40E-02	–	4.009968
	80	5.60E-03	2.78	2.41E-03	2.54	4.001208
	160	7.25E-04	2.95	3.29E-04	2.87	4.000114
	320	9.17E-05	2.98	4.20E-05	2.97	4.000011
	640	1.15E-05	3.00	5.27E-06	3.00	3.999998
	1280	1.44E-06	3.00	6.59E-07	3.00	4.000000
	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
TVB flux limiters	40	3.82E-02	–	1.41E-02	–	4.000000
	80	5.60E-03	2.77	2.41E-03	2.54	4.000000
	160	7.25E-04	2.95	3.29E-04	2.87	4.000000
	320	9.17E-05	2.98	4.20E-05	2.97	4.000000
	640	1.15E-05	3.00	5.27E-06	3.00	3.999997
	1280	1.44E-06	3.00	6.59E-07	3.00	4.000000

TABLE 2. Example 5: L^1, L^∞ errors and orders of accuracy for ENO3/WENO5 reconstruction with TVB flux limiters.

	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
ENO3	40	8.65E-03	–	2.29E-03	–	3.985431
	80	1.10E-03	2.98	2.87E-04	3.00	3.998527
	160	1.37E-04	3.00	3.55E-05	3.02	3.999763
	320	1.72E-05	3.00	4.41E-06	3.01	3.999966
	640	2.15E-06	3.00	5.49E-07	3.00	3.999990
	1280	2.69E-07	3.00	6.83E-08	3.01	3.999999
	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
WENO5	20	8.93E-03	–	2.41E-03	–	3.962503
	40	2.82E-04	4.99	8.39E-05	4.85	3.994505
	80	8.79E-06	5.00	2.72E-06	4.95	3.999673
	160	2.75E-07	5.00	8.33E-08	5.03	3.999904
	320	8.57E-09	5.00	2.48E-09	5.07	3.999983
	640	2.66E-10	5.01	7.19E-11	5.11	3.999992

Example 7. Consider the linear problem (6.1) with periodic boundary condition and initial condition (6.3).

The numerical results are displayed in Figures 2 and 3 for Fix3, ENO3 and WENO5 reconstructions without and with TVB flux limiters, respectively. It can be observed that without TVB limiters, the numerical results all exceed the initial total variation. When the TVB flux limiters are applied, the total variations are completely bounded by that of the initial data. In the following figures, the red/starred line is for exact solution and the blue/dotted line is for numerical solution. The incremental of total variation for the numerical results from ENO3 and

TABLE 3. Example 6: L^1, L^∞ errors and orders of accuracy for Fix3 reconstruction without and with TVB flux limiters.

	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
NO flux limiters	40	1.49E-02	–	2.86E-02	–	4.000082
	80	2.71E-03	2.46	8.81E-03	1.70	4.000017
	160	3.92E-04	2.79	1.70E-03	2.37	3.999913
	320	5.15E-05	2.93	2.50E-04	2.77	4.000002
	640	6.55E-06	2.97	3.28E-05	2.93	3.999978
	1280	8.22E-07	3.00	4.14E-06	2.99	3.999996
	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
TVB flux limiters	40	1.49E-02	–	2.86E-02	–	4.000000
	80	2.71E-03	2.46	8.81E-03	1.70	3.999966
	160	3.92E-04	2.79	1.70E-03	2.37	3.999913
	320	5.15E-05	2.93	2.50E-04	2.77	3.999999
	640	6.55E-06	2.98	3.28E-05	2.93	3.999978
	1280	8.22E-07	3.00	4.14E-06	2.99	3.999996

TABLE 4. Example 6: L^1, L^∞ errors and orders of accuracy for ENO3/WENO5 reconstruction with TVB flux limiters.

	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
ENO3	20	2.78E-02	–	1.40E-02	–	3.962806
	40	4.04E-03	2.78	3.57E-03	1.98	3.993606
	80	7.10E-04	2.51	7.83E-04	2.19	3.998529
	160	1.02E-04	2.80	1.45E-04	2.43	3.999617
	320	1.46E-05	2.81	3.36E-05	2.11	3.999944
	640	2.36E-06	2.63	7.44E-06	2.18	3.999955
	N	L^1 error	Order	L^∞ error	Order	TV(u^n)
WENO5	20	1.17E-02	–	1.00E-02	–	3.985593
	40	1.06E-03	3.47	1.15E-03	3.13	3.998238
	80	4.46E-05	4.57	7.31E-05	3.97	3.999742
	160	1.69E-06	4.72	2.55E-06	4.84	3.999830
	320	5.79E-08	4.87	1.36E-07	4.23	3.999999
	640	1.61E-09	5.17	2.78E-09	5.61	3.999964

WENO5 without TVB flux limiters is slight and almost invisible without zoomed-in view.

Example 8. Consider the nonlinear problem (6.2), initial condition (6.3).

Noticing that $f'(u)$ changes sign, we use the operator splitting version of the TVB flux limiting method described in Section 4.1. The numerical solutions at $T = 0.5$ are displayed in Figures 4 and 5 for Fix3 and WENO5 reconstructions without and with TVB flux limiters. The difference from applying the TVB limiters to the ENO3 method is minimal and almost visually undetectable. We skip its

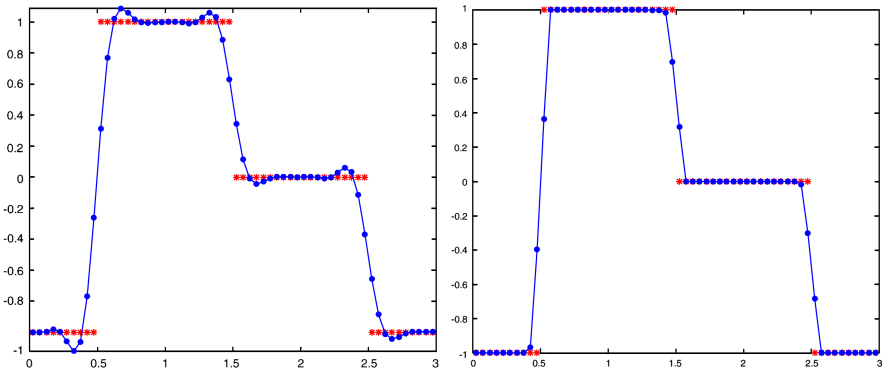


FIGURE 2. Example 7: Left: Fix3 without flux limiters $\mathrm{TV}(u_n) = 4.9542$; Right: Fix3 with TVB flux limiters $\mathrm{TV}(u_n) = 4.0000$.

representation here. Again, from these graphs, it can be observed that the numerical total variations are well controlled.

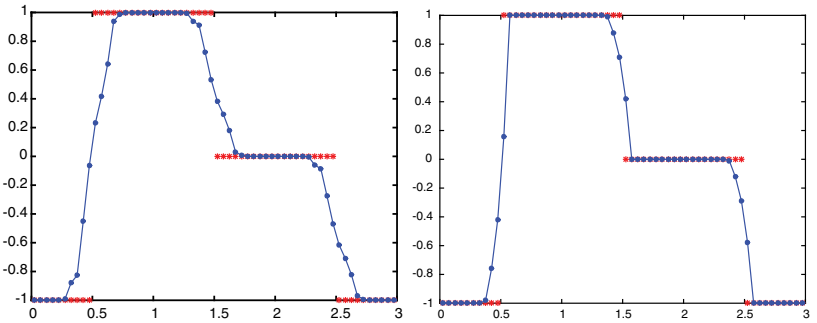
6.3. **When the estimate of the total variation is not sharp.** As is known, that the total variation strictly decays for many initial data. At a later time, the total variation of the entropy solution can be strictly less than the total variation of the initial data by a significant amount. To better understand the performance of the TVB flux limiters in this scenario, we consider such a problem below.

Example 9. The problem (6.2), with initial condition $u_0(x) = \sin(x)$ on $[0, 2\pi]$. Consider the numerical results at $T = 1.25$ and $T = 2$.

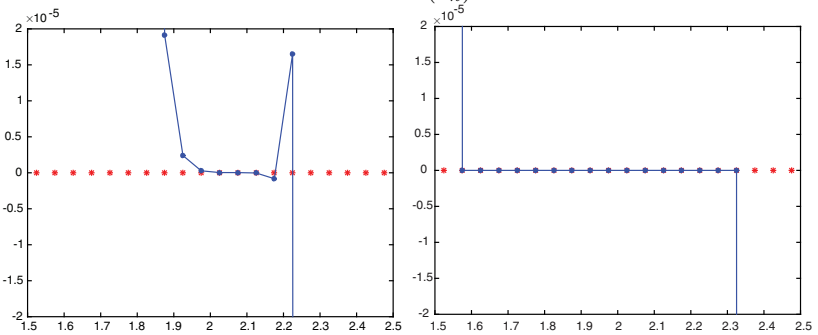
Considering $f'(u)$ changes sign in this test, we apply the splitting approach in (4.9) and (4.10) with $\alpha = 1$. We notice that the total variation $\mathrm{TV}(u) = 4$ up to the time $T = \frac{\pi}{2}$ and $\mathrm{TV}(u) < 4$ after that. We present the numerical solutions with linear reconstruction without and with the TVB flux limiters. In Figure (6a), the discontinuity is well resolved when the total variation bound is sharp. In Figures (6b,6c,6d), oscillations are observed with linear reconstruction. Meanwhile, the oscillation grows with grid refinements when TVB flux limiters are not applied. With the TVB flux limiters, the oscillation is under control by the initial total variation bound although the oscillation is not completely removed (since the total variation bound is not sharp). In this regard, ENO/WENO methods perform better since they are capable of suppressing the oscillation around discontinuities in these cases.

7. Conclusion remarks

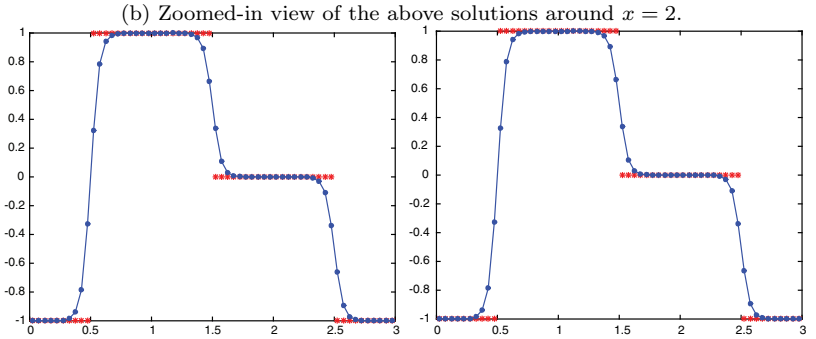
In this paper, we have designed a TVB high order finite difference scheme for scalar hyperbolic conservation laws. The novel TVB criterion based on discrete grid values can be applied to high order finite difference schemes with linear and nonlinear reconstructions. It provides a provable TVB mechanism for high order finite difference methods without sacrificing the order of accuracy. It does not introduce extra CFL constraint for the third order finite difference schemes as shown in the analysis section. The proof for schemes higher than third order is much more difficult since our analysis relies on tedious Taylor expansion. The



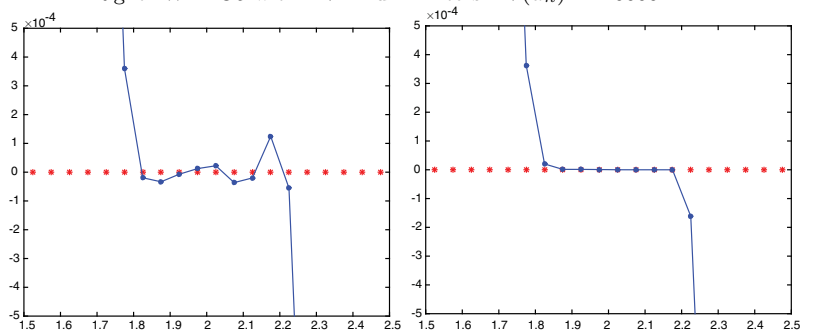
(a) Left: ENO3 without flux limiters $\mathrm{TV}(u_n) = 4.0001$; Right: ENO3 with TVB flux limiters $\mathrm{TV}(u_n) = 4.0000$



(b) Zoomed-in view of the above solutions around $x = 2$.



(c) Left: WENO5 without flux limiters $\mathrm{TV}(u_n) = 4.0014$; Right: WENO5 with TVB flux limiters $\mathrm{TV}(u_n) = 4.0000$



(d) Zoomed-in view of the above solutions around $x = 2$.

FIGURE 3. Example 7: ENO3/WENO5 without and with TVB flux limiters

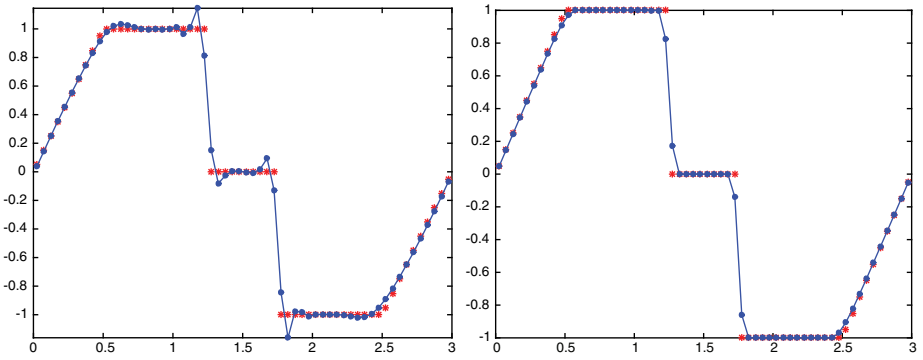


FIGURE 4. Example 8: Left: Fix3 without flux limiters $\mathrm{TV}(u_n) = 5.2751$; Right: Fix3 with TVB flux limiters $\mathrm{TV}(u_n) = 4.0000$.

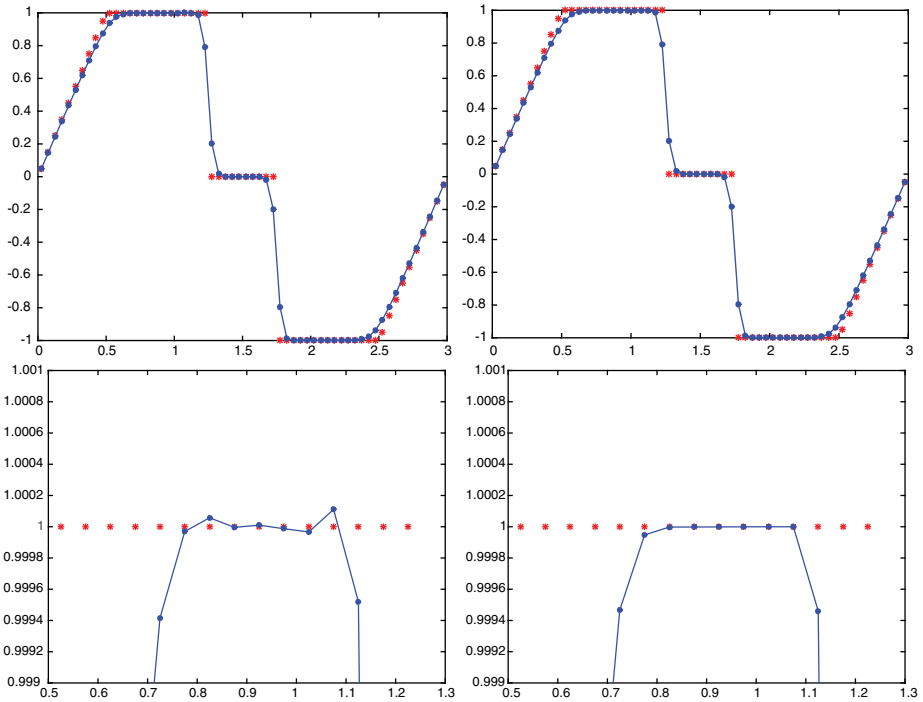
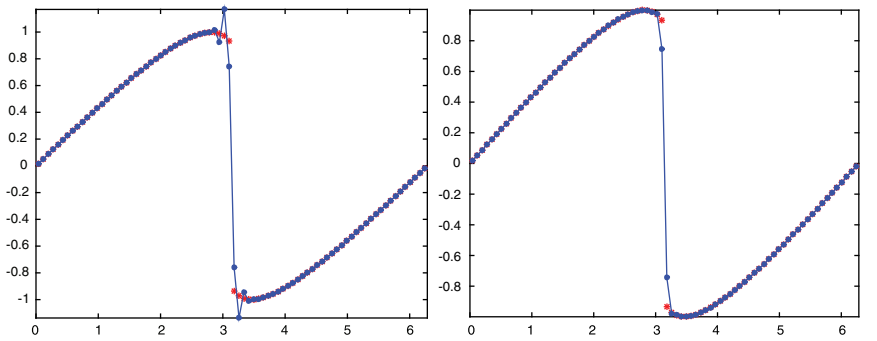
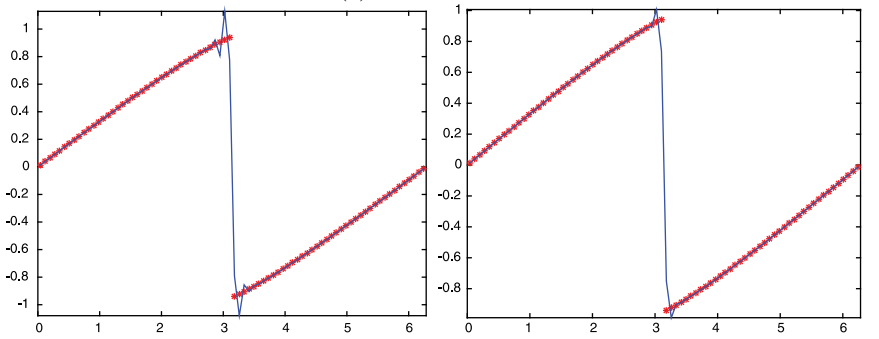


FIGURE 5. Example 8: Left: WENO5 without flux limiters $\mathrm{TV}(u_n) = 4.001$; Right: WENO5 with TVB flux limiters $\mathrm{TV}(u_n) = 4.000$.

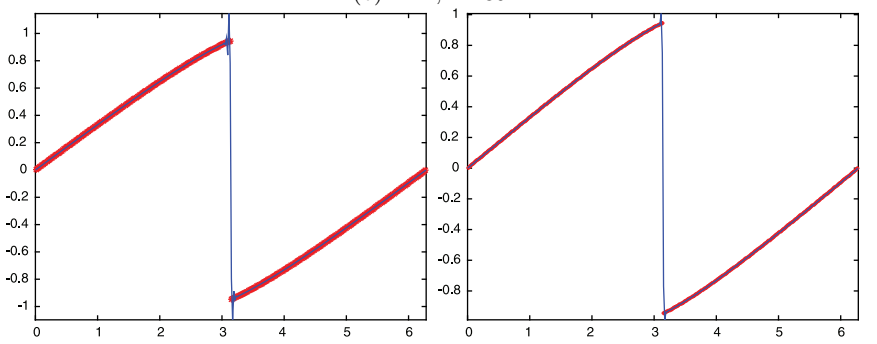
performance of the TVB flux limiters are excellent when the total variation bound is sharp. However, when total variation bound is not sharp, the classic ENO and WENO schemes performs better than the linear scheme with TVB flux limiters around discontinuities in terms of suppressing oscillations.



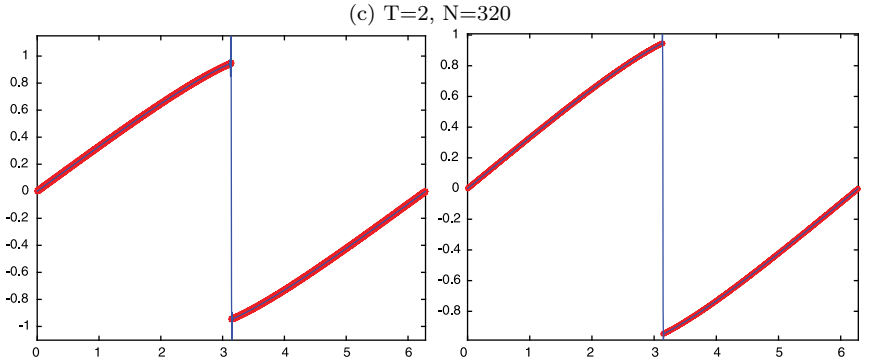
(a) T=1.25, N=80



(b) T=2, N=80



(c) T=2, N=320



(d) T=2, N=1280

FIGURE 6. Example 9: Left: Fix3 without flux limiters; Right: Fix3 with TVB flux limiters.

References

[1] J. P. Boris and D. L. Book, *Flux-corrected transport. I. SHASTA, A fluid transport algorithm that works*, Journal of computational physics **11** (1973), no. 1, 38–69.

[2] B. Cockburn and C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework*, Math. Comp. **52** (1989), no. 186, 411–435. MR983311

[3] P. Colella, *A direct Eulerian MUSCL scheme for gas dynamics*, SIAM J. Sci. Statist. Comput. **6** (1985), no. 1, 104–117. MR773284

[4] U. S. Fjordholm, S. Mishra, and E. Tadmor, *Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws*, SIAM J. Numer. Anal. **50** (2012), no. 2, 544–573. MR2914275

[5] J. Glimm, *Solutions in the large for nonlinear hyperbolic systems of equations*, Comm. Pure Appl. Math. **18** (1965), 697–715. MR0194770

[6] S. K. Godunov, *A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics* (Russian), Mat. Sb. (N.S.) **47 (89)** (1959), 271–306. MR0119433

[7] J. B. Goodman and R. J. LeVeque, *On the accuracy of stable schemes for* $2D$ *scalar conservation laws*, Math. Comp. **45** (1985), no. 171, 15–21. MR790641

[8] A. Harten, *High resolution schemes for hyperbolic conservation laws*, J. Comput. Phys. **49** (1983), no. 3, 357–393. MR701178

[9] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, *Uniformly high-order accurate essentially nonoscillatory schemes. III*, J. Comput. Phys. **71** (1987), no. 2, 231–303. MR897244

[10] A. Harten, P. D. Lax, and B. van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev. **25** (1983), no. 1, 35–61. MR693713

[11] G.-S. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes.*, Tech. report, DTIC Document, 1995.

[12] G.-S. Jiang and E. Tadmor, *Nonoscillatory central schemes for multidimensional hyperbolic conservation laws*, SIAM J. Sci. Comput. **19** (1998), no. 6, 1892–1917. MR1638064

[13] P. Lax and B. Wendroff, *Systems of conservation laws*, Comm. Pure Appl. Math. **13** (1960), 217–237. MR0120774

[14] P. D. Lax and X.-D. Liu, *Solution of two-dimensional Riemann problems of gas dynamics by positive schemes*, SIAM J. Sci. Comput. **19** (1998), no. 2, 319–340. MR1618863

[15] X.-D. Liu and S. Osher, *Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes. I*, SIAM J. Numer. Anal. **33** (1996), no. 2, 760–779. MR1388497

[16] X.-D. Liu, S. Osher, and T. Chan, *Weighted essentially non-oscillatory schemes*, J. Comput. Phys. **115** (1994), no. 1, 200–212. MR1300340

[17] X.-D. Liu and E. Tadmor, *Third order nonoscillatory central scheme for hyperbolic conservation laws*, Numer. Math. **79** (1998), no. 3, 397–425. MR1626324

[18] S. Osher and S. Chakravarthy, *High resolution schemes and the entropy condition*, SIAM J. Numer. Anal. **21** (1984), no. 5, 955–984. MR760626

[19] R. Sanders, *A third-order accurate variation nonexpansive difference scheme for single nonlinear conservation laws*, Math. Comp. **51** (1988), no. 184, 535–558. MR935073

[20] C.-W. Shu, *TVB uniformly high-order schemes for conservation laws*, Math. Comp. **49** (1987), no. 179, 105–121. MR890256

[21] C.-W. Shu, *High order weighted essentially nonoscillatory schemes for convection dominated problems*, SIAM Rev. **51** (2009), no. 1, 82–126. MR2481112

[22] C.-W. Shu and S. Osher, *Efficient implementation of essentially nonoscillatory shock-capturing schemes. II*, J. Comput. Phys. **83** (1989), no. 1, 32–78. MR1010162

[23] C.-W. Shu and S. Osher, *Efficient implementation of essentially nonoscillatory shock-capturing schemes*, J. Comput. Phys. **77** (1988), no. 2, 439–471. MR954915

[24] P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM J. Numer. Anal. **21** (1984), no. 5, 995–1011. MR760628

[25] B. Van Leer, *Towards the ultimate conservative difference scheme. II. monotonicity and conservation combined in a second-order scheme*, J. Comput. Phys. **14** (1974), no. 4, 361–370.

[26] B. van Leer, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method [J. Comput. Phys.* **32** *(1979), no. 1, 101–136]*, J. Comput. Phys. **135** (1997), no. 2, 227–248. With an introduction by Ch. Hirsch; Commemoration of the 30th anniversary {of J. Comput. Phys.}. MR1486274

[27] T. Xiong, J.-M. Qiu, and Z. Xu, *A parametrized maximum principle preserving flux limiter for finite difference RK-WENO schemes with applications in incompressible flows*, J. Comput. Phys. **252** (2013), 310–331. MR3101509

[28] Z. Xu, *Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem*, Math. Comp. **83** (2014), no. 289, 2213–2238. MR3223330

[29] S. T. Zalesak, *Fully multidimensional flux-corrected transport algorithms for fluids*, J. Comput. Phys. **31** (1979), no. 3, 335–362. MR534786

[30] X. Zhang and C.-W. Shu, *A genuinely high order total variation diminishing scheme for one-dimensional scalar conservation laws*, SIAM J. Numer. Anal. **48** (2010), no. 2, 772–795. MR2670004

Department of Mathematical Science, Michigan Tech University, Houghton, Michigan 49931

Email address: sulinw@mtu.edu

Department of Mathematical Science, Michigan Tech University, Houghton, Michigan 49931

Email address: zhengfux@mtu.edu