ON WAVEWISE ENTROPY INEQUALITIES FOR HIGH-RESOLUTION SCHEMES WITH SOURCE TERMS I: THE SEMI-DISCRETE CASE *

HUANAN YANG† AND NAN JIANG‡

Abstract. We extend the framework and the convergence criteria of wavewise entropy inequalities of [H. Yang, *Math. Comp.*, (1996), pp. 45-67] to a large class of semi-discrete high resolution schemes for hyperbolic conservation laws with source terms. This approach is based on an extended theory of Yang [22] on wave tracking and wave analysis and the theory of Vol'pert [21] on BV solutions. For the Cauchy problem of convex conservation laws with source terms, we use one of the criteria to prove the convergence to the entropy solution of generalized MUSCL schemes and a class of schemes using flux limiters previously discussed in 1984 by Sweby.

1. Introduction. In this paper, we extend the framework and the convergence criteria of wavewise entropy inequalities, or WEI, developed in [22] to a large class of semi-discrete high-resolution schemes for initial value problems of hyperbolic conservation laws with source terms:

$$(1.1) \qquad \begin{cases} w_t + f(w)_x = q(w), \\ w(x,0) = w_0(x), \end{cases}$$

where $f \in C^1(\mathbb{R})$, $q \in C^1(\mathbb{R})$, and $w_0 \in BV(\mathbb{R})$. Here BV stands for the subspace of L^1_{loc} consisting of functions z with bounded total variation

$$TV(z) := \sup_{h \neq 0} \int_{\mathbb{R}} \frac{|z(x+h) - z(x)|}{|h|} \, dx.$$

The homogeneous problems corresponding to (1.1) are

$$(1.2) \qquad \begin{cases} w_t + f(w)_x = 0, \\ w(x,0) = w_0(x). \end{cases}$$

To introduce the numerical scheme, let us partition the real line for the space variable into cells of equal size. The j-th cell is centered at $x_j = jh + c$, where $j = 0, \pm 1, \pm 2, \ldots$, c is a constant, and h is the space step size. For an arbitrary function v we use Δ_+ and Δ_- to denote the forward and the backward difference operators, respectively: $\Delta_\pm v_j = \pm(v_{j\pm 1} - v_j)$ or $\Delta_\pm v(x) = \pm(v(x \pm h) - v(x))$. The corresponding divided difference operators are denoted by $D_\pm = \frac{1}{h}\Delta_\pm$. Let $u_j(t)$ be the numerical approximation to the exact solution $w(x_j, t)$ or its cell average on the j-th cell $\bar{w}(x_j, t)$. We consider semi-discrete conservative schemes for (1.1) that have the form

$$(1.3) \qquad \frac{d}{dt} u_j(t) = -D_+ g_{j-\frac{1}{2}}(t) + q(u_j(t))$$

where

$$(1.4) \qquad g_{j+\frac{1}{2}}(t) = g_{j+\frac{1}{2}}(u(t)) = g(u_{j-p+1}(t), u_{j-p+2}(t), \cdots, u_j(t), \cdots, u_{j+p}(t), h).$$

---

†Department of Mathematics, Kansas State University, Manhattan, Manhattan, KS 66506-2602, USA (hyang@math.ksu.edu).

‡Department of Mathematical Sciences, University of South Dakota, Vermillion, SD 57069, USA (njiang@usd.edu).

Here g is the numerical flux which is Lipschitz continuous with respect to its first $2p$ arguments and is *consistent* with the conservation law in the sense that

$$(1.5) \qquad\qquad g(u, u, \cdots, u, h) \equiv f(u).$$

The collection of points $\{x_{j-p}, x_{j-p+1}, \cdots, x_{j+p}\}$ is said to be *the stencil of the scheme at the point* (x_j, t), and the integer $2p + 1$ is *the size of the stencil of the scheme*. We call the corresponding scheme

$$(1.6) \qquad\qquad \frac{d}{dt} u_j(t) = -D_+ g_{j-\frac{1}{2}}(t)$$

which is consistent with the problem (1.2) the *homogeneous counterpart* (HCP) of the scheme (1.3). The scheme (1.6) is said to be *self-similar* if g is independent of h. In this paper, we only consider conservative schemes with self-similar HCPs.

We denote by Q_t the solution operator of the ordinary differential equation $v'(t) = q(v(t))$. Namely, $v(t) = Q_t(v(0))$. For a numerical solution $\{u_j, j \in \mathbb{Z}\}$, let $A = \inf_j(u_j(0))$ and $B = \sup_j(u_j(0))$. Let $T > 0$ be a point within the intersection of the intervals of existence of $Q_t(A)$ and $Q_t(B)$. Finally, we set $\bar{A} = \min_{t \in [0,T]} Q_t(A)$ and $\bar{B} = \max_{t \in [0,T]} Q_t(B)$. As an easy consequence of the results of this paper, the exact solution $w(\cdot, t)$ is defined for $0 \leq t \leq T$ and its range is a subset of $[\bar{A}, \bar{B}]$. Through out this paper $Q_t, A, B, \bar{A}, \bar{B}$ and T will obey these definitions. We extend the domain of a numerical solution $\{u_j(t) : j \in \mathbb{Z}\}$ to $\mathbb{R} \times [0, T)$ by setting $u(x, t) = u_j(t)$ for $x_{j-1/2} \leq x < x_{j+1/2}$. Extended numerical solutions are represented by the same symbols as the mesh valued ones except the absence of subscripts. We use superscripts to label sequences of solutions. A scheme (1.3)-(1.5) for the Cauchy problem (1.1) *converges* if, for every initial condition w_0 in BV and for each sequence of initial data $\{u_j^k(0), j \in \mathbb{Z}\}_{k=1}^\infty$ with uniformly bounded variations that converges in $L_{loc}^1(\mathbb{R})$ to w_0, the corresponding sequence of (extended) numerical solutions $\{u^k\}$ generated by the scheme converges in $L_{loc}^1(\mathbb{R} \times [0, T))$ to the unique entropy solution w of the problem (1.1) provided that the step sizes h_k of u^k vanish as $k \to \infty$. By Helly's Theorem on the set of total variation bounded functions and Lax-Wendroff Theorem [8] on conservative schemes, a TVB conservative scheme converges provided that each limit function of the numerical solution is an entropy solution.

The analyses of convergence in the early time were focused on the numerical methods for homogeneous problems (1.2). By the end of the 1980s, the method of cell entropy inequalities (CEI) had been the dominant approach for the analysis of entropy admissibility. See, for example, [1], [4], [7], [12], [13], [15], [16] and the references therein. In the CEI approach, one tries to derive cell entropy inequalities for certain pairs of numerical entropy and entropy flux. Once this is obtained, the same arguments for Lax-Wendroff Theorem [8] ensure entropy admissibility of the limits of the numerical solutions. Unfortunately, it is too demanding, if not impossible, for a high-resolution scheme to satisfy numerical entropy inequalities at every mesh point, as required by the CEI approach. As a result, the convergence of some very popular methods, such as the MUSCL scheme in its original setting cannot be proved by this approach.

Realizing that cell entropy inequalities are not necessary for convergence, since the early 1990s, several authors have developed new approaches for convergence analysis. Lions and Souganidis [9] proved, for strictly convex homogeneous conservation laws, convergence of MUSCL schemes (which are the second-order extension of Godunov's schemes) for the resolvent equation and the implicit MUSCL schemes with

large (unbounded) CFL numbers obtained by the backward Euler time discretization of the semi-discrete MUSCL schemes. The proof is based on the theory of viscosity solutions. To our knowledge, this approach has not been extended to conservation laws with source terms.

Meanwhile, Yang [22] and [23] formed the concept of wavewise entropy inequalities (WEI) for a large class of TVD (Total Variation Diminishing) schemes. Based on this concept Yang proved several convergence criteria. In particular, for convex conservation laws, one of the criteria points out that, a wavewise entropy inequality across the area of rarefaction where $u_j \leq u_{j+1}$ for all x_j is sufficient for convergence to the entropy solution. Hence, in convergence analysis, one may safely remove the shock area from scrutiny. Further, even in the rarefaction area, a much weaker condition than CEI is sufficient for convergence. Using this criterion, Yang proved the convergence of both semi-discrete MUSCL schemes and some fully-discrete ones, and the convergence of a class of high-resolution schemes based on flux limiters, for homogeneous problems with convex flux functions (1.2).

Recently, the numerical analysis of non-homogeneous problems (1.1) has attracted much attention. This includes studying numerical methods for the approximation of (1.1), see [2], [6], [10], for example; the error bounds related to the approximation of (1.1), see [17], [20] for example. However, the analytical tools in this area remain to be CEI, and hence, suffer to the aforementioned restrictions.

In this paper, we extend the entire framework of Yang's WEI (see [22]) to non-homogeneous conservation laws provided that the numerical flux satisfies the same conditions as in the homogeneous case. In particular, we show that the aforementioned semi-discrete MUSCL schemes and the class of high-resolution schemes based on flux limiters, remain convergent in the non-homogeneous case. The WEI framework also works for fully-discrete schemes which will be discussed in a subsequent paper.

The paper is organized as follows. §2 consists of two parts. In the first part we review some properties of the discontinuities of BV weak solutions of conservation laws emphasizing entropy conditions which harbor the idea of WEI approach; and in the second part we show existence, uniqueness and total variation boundedness of the numerical solutions as infinite systems of ordinary differential equations, which ensure existence of convergent subsequences of numerical solutions whose limits are weak solutions by, again, the arguments of Lax-Wendroff (see [8]).

The main results of the paper are in section §3 where we give four WEI convergence criteria and use one of them to show convergence of the generalized MUSCL schemes and a class of high resolution schemes using flux limiters for convex conservation laws with source terms. These results are parallel to those in [22] for their HCPs. We give full proofs of the first two criteria since they are simple and reveal interesting effects of similarity transforms on the schemes with source terms. To prove the third criterion, we need to extend the extremum tracking theory of [22] to high-resolution schemes for non-homogeneous conservation laws, and we need to perform the wave separations, concentrations and splittings. The extension of the extremum tracking theory was once in doubt since, unlike their HCPs, the values of the numerical solutions with source terms may increase at local maxima and decrease at minima. Remarkably, as we are able to show, the enhancement of local extrema by the source term does not destroy the non-oscillatory property of the scheme, which makes the extension possible. We devote the entire section §4 to the, rather lengthy, extension. With the extended extremum tracking theory one could perform the wave separations, concentrations and splittings needed to complete the proof. We omit

these since they follow, almost word by word, those in [22]. We also omit the proof of the last criterion for the same reason.

In general, we only present proofs that are substantially different from the corresponding ones in [22]. Also, for better readability, we closely follow many notations in [22].

2. Preliminaries.

2.1. Review of discontinuities of weak solutions. In [22] a simple observation explains the idea of WEI approach for homogeneous conservation laws. This observation is also valid for non-homogeneous ones: Let $U(w)$ be a convex entropy function, and $F(w)$ its flux: $F' = U'f'$. In the area where the solution w is smooth, the additional conservation law $U(w)_t + F(w)_x = U'(w)q(w)$ holds, and the entropy condition is automatically satisfied. Therefore, the entropy admissibility of a weak solution is solely determined by that of its discontinuities. The following is a closer examination of this observation.

For any two distinct numbers w^- and w^+ in the domain of f, the function

$$(2.1) \qquad W(x,t) = \begin{cases} w^- & \text{if } x < st, \\ w^+ & \text{if } x > st, \end{cases}$$

is a *traveling discontinuity*, provided that

$$(2.2) \qquad s(w^+ - w^-) = f(w^+) - f(w^-)$$

holds. Clearly $W(x,t)$ is a weak solution of the homogeneous conservation law. Denote by $f[w; w^-, w^+]$ the linear function interpolating $f(w)$ at $w = w^-$ and $w = w^+$. Then W is an *admissible traveling discontinuity* if

$$(2.3) \qquad \text{sgn}(w^+ - w^-)(f[w; w^-, w^+] - f(w)) \leq 0$$

holds for all w between w^- and w^+; otherwise, it is a *traveling expansion shock*.

Generic discontinuities of BV weak solutions are inherently connected to the traveling discontinuities through Vol'pert's BV solution theory (see [21]) on which a brief discussion is in order. Let $\mu(E)$ be the Lebesgue measure of a measurable set $E \subset \mathbb{R}^n$. We use $B_r(x_0)$ to denote the open ball centered at x_0 with the radius r. Let a be an unit vector in $\mathbb{R}^n$, and $R_a(x_0)$ be the half space $\{(x - x_0) \bullet a > 0\}$ in $\mathbb{R}^n$. A point of *density (rarefaction)* for the set E is a point x for which

$$\lim_{r \to 0} \mu(E \bigcap B_r(x))/\mu(B_r(x)) = 1(0).$$

If $w(x)$ is a function defined on a set $E \subset \mathbb{R}^n$ and x_0 is not a point of rarefaction for E, then $L_E w(x_0)$ will denote the *approximate limit* of the function $w(x)$ at the point x_0 with respect to the set E. This means, by definition, $\forall \varepsilon > 0$, x_0 is a point of rarefaction of the set

$$\{x : |w(x) - L_E w(x_0)| > \varepsilon, x \in E.\}$$

DEFINITION 2.1.   Let $w(x)$ be a function defined on $\mathbb{R}^n$.

(α) A point $x_0 \in \mathbb{R}^n$ is said to be *regular* if there exists a unit vector a such that $l_a w(x_0)$ and $l_{-a} w(x_0)$ exist and are finite. Here, $l_a w(x_0) = L_{R_a(x_0)} w(x_0)$.

(β) The point x_0 is said to be a *point of jump* for $w(x)$ if it is regular and $l_a w(x_0) \neq l_{-a} w(x_0)$. The set of the jump points for $w(x)$ is denoted by $\Gamma(w)$.

(γ) If $x_0 \in \Gamma(w)$, then the value a appearing in the definition (α) is called *the normal* to $\Gamma(w)$ at the point x_0.

For simplicity, we apply the preceding concepts in the case $n = 2$ to a BV weak solution $w(x, t)$ of the conservation law, possibly with source terms. For any $(x_0, t_0) \in \Gamma(w)$, let a be the normal to $\Gamma(w)$ at the point (x_0, t_0) with positive spatial component, and let $w^+ = l_a w(x_0, t_0)$, $w^- = l_{-a} w(x_0, t_0)$. We then call W, defined by (2.1)-(2.2), *the traveling discontinuity associated with w at the jump point* (x_0, t_0).

Denote by H_n the n-dimensional Hausdorff measure. The following basic result holds.

LEMMA 2.2 (Vol'pert[21]).   *A necessary and sufficient condition for a weak solution $w \in BV$ of $w_t + f(w)_x = q(w)$ to be an entropy solution is that (2.3) holds, for H_1-almost all points in $\Gamma(w)$.*

Briefly speaking, in the WEI approach, if a sequence of total variation bounded numerical solutions approaches an entropy violating weak solutions, one may construct a sequence of numerical solutions with vanishing step size and vanishing source terms that converges to a traveling expansion shock and harbors an asymptotic traveling expansion shock, a concept that will be given in §3. Similarity transforms play the central role in the construction of such a sequence. Let $\mathcal{S}_{x_0, t_0}^{\varepsilon}$ be the similarity transform centered at a point (x_0, t_0):

$$\mathcal{S}_{x_0, t_0}^{\varepsilon}((x, t)) = (x_0 + \varepsilon x, t_0 + \varepsilon t).$$

This induces a transform $T_{x_0, t_0}^{\varepsilon}$ in the set of the functions ψ defined on a domain $\Omega \subset \mathbb{R}^n \times \mathbb{R}^+$: $T_{x_0, t_0}^{\varepsilon} \psi = \psi \circ \mathcal{S}_{x_0, t_0}^{\varepsilon} |_{\Omega}$, if $\mathcal{S}_{x_0, t_0}^{\varepsilon} \Omega \subset \Omega$, where $\phi |_{\Omega}$ denotes the restriction of ϕ to the set Ω. Define $w_{\varepsilon}(x, t)$ by

$$w_{\varepsilon}(x, t) = (T_{x_0, t_0}^{\varepsilon} w)(x, t) = w \circ \mathcal{S}_{x_0, t_0}^{\varepsilon}((x, t)) = w(x_0 + \varepsilon x, t_0 + \varepsilon t).$$

Clearly, if $w(x, t)$ is a weak solution of $w_t + f(w)_x = q(w)$, then $w_{\varepsilon}(x, t)$ is one of $w_t + f(w)_x = \varepsilon q(w)$. The following lemma (presented in [22] and still holds for non-homogeneous case) is one of the foundations of the WEI method. It shows that by successively zooming in around a jump point (x_0, t_0) of a weak solution w, one can view it locally as a traveling discontinuity.

LEMMA 2.3 (Microscope Lemma).   *Let (x_0, t_0) be a jump point of a BV weak solution w in the sense of Definition 2.1. If $\{\varepsilon_k\}_{k=1}^{\infty}$ is a sequence of positive numbers such that $\lim_{k \to \infty} \varepsilon_k = 0$, then the sequence $\{w_{\varepsilon_k}\}$ converges in L_{loc}^1 to the traveling discontinuity W associated with the jump point (x_0, t_0).*

2.2. Preliminaries of the numerical schemes. In this subsection, we discuss the existence, uniqueness and boundedness of the solutions of the schemes (1.3)-(1.5) and the convergence of the schemes to the weak solutions. Throughout the paper, we require that the numerical flux of the schemes satisfy the following assumption.

ASSUMPTION 2.4.   *The numerical fluxes $g_{j+\frac{1}{2}}(t)$, $j = 0$, ± 1, ± 2, $\cdots$, satisfy*

$$g_{j+\frac{1}{2}}(t) \geq g_{j-\frac{1}{2}}(t) \quad \text{if } u_j(t) - u_{j\pm 1}(t) \geq 0,$$

and

$$g_{j+\frac{1}{2}}(t) \leq g_{j-\frac{1}{2}}(t) \quad if \ \ u_j(t) - u_{j\pm 1}(t) \leq 0.$$

We remark that for homogeneous conservation laws, this was proposed by Tadmor [19] as a convenient TVD condition of conservative schemes.

The local existence, uniqueness and total variation boundedness of the numerical solutions for the schemes (1.3)-(1.5) under the Assumption 2.4 follow easily from Picard's iteration beginning with $u_j^{(0)}(t) = Q_{t-t_0}(u_j(t_0))$. See [11] for details. We have the following result on the "global" existence and uniqueness of the numerical solutions.

THEOREM 2.5. *Suppose the flux of the numerical scheme* (1.3)-(1.5) *satisfies Assumption* 2.4, *and the initial condition satisfies* $A \leq u_j(0) \leq B$ *for all integer* j. *Then the scheme* (1.3)-(1.5) *admits a unique solution that satisfies* $Q_t(A) \leq u_j(t) \leq Q_t(B)$ *for all* j *and for* $0 \leq t \leq T$ *provided that* $Q_t(A)$ *and* $Q_t(B)$ *are defined on the interval.*

To prove the theorem, we first establish a lemma concerning

$$S(t) := \sup_j(u_j(t)) \quad \text{and} \quad I(t) := \inf_j(u_j(t)).$$

LEMMA 2.6. *Under the Assumption* 2.4, *for any* $b > a \geq 0$, *we have*

$$S(b) \leq S(a) + \int_a^b q(S(t)) \, dt \quad and \quad I(b) \geq I(a) + \int_a^b q(I(t)) \, dt$$

provided that $S(t)$ *and* $I(t)$ *are defined on* $[a, b]$.

Proof. The proofs of the two inequalities are parallel, and we only prove the one for $S(t)$. Clearly, $S(t)$ is a Lipschitz function, which implies that

$$S(b) = S(a) + \int_a^b S'(t) \, dt.$$

It is, therefore, suffices to show that $S'(t) \leq q(S(t))$ for every $t \in (a, b]$ such that $S'(t)$ is defined. Let $b \geq t > t - \tau > a$. Fix such a t. Since u is locally TVB, there exists a sequence of integers $\{j_k\}_{k=1}^{\infty}$ such that the limits

$$v_i = \lim_{k \to \infty} u_{j_k+i}(t) \quad \text{for} \quad i = 0, \pm 1, \ldots, \pm p$$

exist with

$$v_0 = S(t) = \max_{-p \leq i \leq p} v_i.$$

We have

$$\frac{S(t) - S(t-\tau)}{\tau} \le \liminf_{k \to \infty} \frac{u_{j_k}(t) - u_{j_k}(t-\tau)}{\tau} = \liminf_{k \to \infty} \int_0^1 \frac{d}{dt} u_{j_k}(t - \theta\tau) d\theta$$

$$= \liminf_{k \to \infty} \int_0^1 \left\{ -\frac{1}{h}(g_{j_k+1/2}(t - \theta\tau) - g_{j_k-1/2}(t - \theta\tau)) + q(u_{j_k}(t - \theta\tau)) \right\} d\theta$$

$$\le \liminf_{k \to \infty} \left\{ -\frac{1}{h}(g_{j_k+1/2}(t) - g_{j_k-1/2}(t)) + q(u_{j_k}(t)) + C(h)\tau \right\}$$

$$= -\frac{1}{h}(g(v_{-p+1}, \ldots, v_0, \ldots, v_p) - g(v_{-p}, \ldots, v_{-1}, \ldots, v_{p-1})) + q(S(t)) + C(h)\tau$$

$$\le q(S(t)) + C(h)\tau.$$

In the last part of the inequality we have applied Assumption 2.4. The desired inequality is obtained by letting $\tau \to 0$. $\square$

Proof. [Proof of Theorem 2.5] Under the assumptions of the theorem, we claim that if the numerical solution $\{u_j(t), j \in \mathbb{Z}\}$ with initial condition

$$\{u_j(0), j \in \mathbb{Z}\} \subseteq [A, B]$$

exists for $0 \le t < \delta$, where δ is a positive constant, then the a-priori estimate $\{u_j(t), j \in \mathbb{Z}\} \subseteq [Q_t(A), Q_t(B)]$ holds for $0 \le t < \delta$. With this a-priori estimate, the desired result follows easily from the local existence and uniqueness and the extension procedure which are standard in the theory of ordinary differential equations (see, for example, [5]). It remains to prove the a-priori estimate which is equivalent to $S(t) \le Q_t(B)$ and $I(t) \ge Q_t(A)$. Since the proof of the two is similar, we only show the former. Let $z(t) = Q_t(B)$. If the inequality were false, then since $u_j(0) \le B$, there would be an interval (α, β) with $0 \le \alpha < \beta \le T$ such that $S(\alpha) = z(\alpha)$ and $S(t) > z(t)$ for $t \in (\alpha, \beta)$. However, Lemma 2.6 and the definition of $z(t)$ imply that, for $t \in (\alpha, \beta)$, we have

$$S(t) \le S(\alpha) + \int_\alpha^t q(S(\tau)) \, d\tau \quad \text{and} \quad z(t) = z(\alpha) + \int_\alpha^t q(z(\tau)) \, d\tau,$$

which imply that $S(t) - z(t) \le \int_\alpha^t q'(\xi(\tau))(S(\tau) - z(\tau)) \, d\tau$. Thus $S(t) - z(t) \le C^+ \int_a^t (S(\tau) - z(\tau)) \, d\tau$, where $C^+ = \max_{t \in [\alpha, \beta]} |q'(\xi(t))|$ is finite. Hence, by Gronwall's inequality, $S(t) - z(t) \le 0$ for $t \in (\alpha, \beta)$, which would be a contradiction. $\square$

The theory of ordinary differential equations and Theorem 2.5 immediately imply the following corollary.

COROLLARY 2.7. *If the numerical flux of the scheme* (1.3)-(1.5) *satisfies Assumption* 2.4, *and* $q'(u) \le C$ *for all* u, *then the numerical solution exists and is unique for* $t \ge 0$.

Let $TV_u(t)$ be the total spatial variation of $u(t)$. For conservation laws with source terms, Assumption 2.4 no longer guarantees the TVD property of the scheme (1.3)-(1.5). However, the following TVB property holds.

THEOREM 2.8. *With the conditions of Theorem* 2.5, *we have* $TV_u(t) \le TV_u(0)e^{Ct}$ *for* $t \in [0, T]$, *provided that* $\max_{u \in [\bar{A}, \bar{B}]} q'(u) \le C$ *with* $C \ge 0$.

Proof. Let $s_{j+1/2}(t) := \text{sgn}(\Delta u_j(t))$. Multiplying both sides of the forward difference of (1.3) by $s_{j+1/2}$ and then integrating, we obtain

$$|\Delta u_j(t)| = |\Delta u_j(0)| - \frac{1}{h} \int_0^t s_{j+1/2}(\tau)[\Delta g_{j+1/2}(\tau) - \Delta g_{j-1/2}(\tau)]\, d\tau$$

$$(2.4) \qquad\qquad + \int_0^t s_{j+1/2}(\tau)\Delta q(u_j(\tau))\, d\tau.$$

Next, summing the above equation over all j, using summation by parts for the second-term on the right-hand side, we obtain

$$(2.5) \qquad\qquad TV_u(t) = \phi_u(t) + \psi_u(t),$$

where $\phi_u(t)$ and $\psi_u(t)$ are defined by

$$(2.6) \quad \phi_u(t) := TV_u(0) + \frac{1}{h} \int_0^t \sum_{j=-\infty}^{\infty} (s_{j+1/2} - s_{j-1/2})(g_{j+1/2}(\tau) - g_{j-1/2}(\tau))\, d\tau,$$

$$(2.7) \qquad\qquad \psi_u(t) := \int_0^t \sum_{j=-\infty}^{\infty} s_{j+1/2}\Delta q(u_j(\tau))\, d\tau,$$

respectively. Since $\phi_u(0) = TV_u(0)$ and Assumption 2.4 implies that $\phi_u(t)$ is a nonincreasing function of t, and since $s_{j+1/2}(\tau)\Delta q(u_j(\tau)) \le C|\Delta u_j(\tau)|$, we have

$$(2.8) \qquad\qquad TV_u(t) \le TV_u(0) + C \int_0^t TV_u(\tau)\, d\tau \quad \forall\, t \ge 0.$$

Hence, $TV_u(t) \le TV_u(0)e^{Ct}$, by Gronwall's inequality. $\square$

REMARK. The functions $\phi_u(t)$ and $\psi_u(t)$ have the following invariant properties under similarity transform T_{x_0,t_0}^ε: Denote $(\hat{x}, \hat{t}) = \mathcal{S}_{x_0,t_0}^\varepsilon((x,t))$ and $\hat{u} = T_{x_0,t_0}^\varepsilon(u)$. Then, for any $t_1, t_2 \in \mathbb{R}$ with $\hat{t}_1, \hat{t}_2 \ge 0$, we have $\phi_{\hat{u}}(t_2) - \phi_{\hat{u}}(t_1) = \phi_u(\hat{t}_2) - \phi_u(\hat{t}_1)$, and $\psi_{\hat{u}}(t_2) - \psi_{\hat{u}}(t_1) = \psi_u(\hat{t}_2) - \psi_u(\hat{t}_1)$. These properties are very important in extending the WEI framework. One should notice that in the definition of ψ the source term for $\hat{u}$ is $\varepsilon q(\hat{u})$ instead of $q(\hat{u})$.

With Theorem 2.8, using Helly's Theorem on the set of total variation bounded functions and following the proof of the Lax–Wendroff Theorem [8], we obtain the following result.

THEOREM 2.9. *Suppose a scheme defined by (1.3)-(1.5) satisfies Assumption 2.4 and $\{u_j^n(t)\}_{n=1}^\infty$ are generated by the scheme. Suppose also that the step sizes $h_n \to 0$ as $n \to \infty$, and the initial conditions have uniformly bounded total variations. Then $\{u_j^n(t)\}_{n=1}^\infty$ contains a subsequence $\{u_j^{n_l}(t)\}_{l=1}^\infty$ which converges in $L_{loc}^1(\mathbb{R} \times [0,T))$ towards a weak solution of (1.1) as $l \to \infty$.*

3. WEI criteria for convergence to the entropy solution. For convenience, let Υ be the set of all sequences of numbers in $(0,1)$ with zero limit. We use bold-faced letters to represent the sequences in Υ, and use the corresponding light-faced ones with subscripts to represent the terms in such a sequence.

3.1. General TVB schemes. To obtain sufficient conditions for convergence to the entropy solution, we attack their contrapositives. Hence, we begin with a sequence of numerical solutions $\{u_j^n\}_{n=1}^{\infty}$ generated by the scheme (1.3)-(1.5) that converges to an entropy violating weak solution w. We also assume that the corresponding sequence of step-sizes $\mathbf{h} \in \Upsilon$. By Lemma 2.2, there exists a jump point (x_0, t_0) of w with an associated traveling expansion shock W. Now for any $\varepsilon \in \Upsilon$ applying similarity transforms $T_{x_0,t_0}^{\varepsilon_k}$ to u^n for each n and to w, we obtain $u_{\varepsilon_k}^n$ and w_{ε_k} respectively. The numerical solution $u_{\varepsilon_k}^n$ satisfies the same scheme for $w_t + f(w)_x = \varepsilon_k q(w)$ with step size $H_{k,n} = h_n/\varepsilon_k$. For fixed k, $u_{\varepsilon_k}^n \to w_{\varepsilon_k}$ as $n \to \infty$ in L_{loc}^1, and the step size satisfies $\lim_{n\to\infty} H_{k,n} = 0$. Applying Lemma 2.3 and using the same diagonal process as in [22], one may choose a sequence of increasing positive integers $\{n_k\}$ such that $\{u_{\varepsilon_k}^{n_k}\}_{k=1}^{\infty}$ converges in L_{loc}^1 to W and $\lim_{k\to\infty} H_{k,n_k} = 0$. For simplicity, we denote $\{u_{\varepsilon_k}^{n_k}\}_{k=1}^{\infty}$ by $\{u^k\}_{k=1}^{\infty}$ and H_{k,n_k} by h_k. Then u^k is generated by

$$(3.1) \qquad \frac{d}{dt}u_j^k(t) = -\frac{1}{h_k}[g_{j+\frac{1}{2}}(u_j^k(t)) - g_{j-\frac{1}{2}}(u_j^k(t))] + \varepsilon_k q(u_j^k(t)),$$

with $\mathbf{h}, \boldsymbol{\varepsilon} \in \Upsilon$. We call the scheme (3.1) the $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form of the scheme (1.3)-(1.5), and we have obtained our first WEI convergence criteria.

THEOREM 3.1. *A TVB scheme* (1.3)-(1.5) *for the Cauchy problem* (1.1) *converges if there exists no sequence of functions* $\{u^k(x,t)\}_{k=1}^{\infty}$ *generated by the* $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled *form* (3.1) *of the scheme with* $\mathbf{h}, \boldsymbol{\varepsilon} \in \Upsilon$ *that converges in* $L_{loc}^1(\mathbb{R} \times [0, T))$ *to a traveling expansion shock.*

3.2. Schemes with TVD HCP. Clearly, Theorem 3.1 also holds for TVB schemes approximating scalar conservation laws of several spatial dimensions. However, time dependent functions of one-spatial dimension have a property that is not shared by those of several spatial dimensions: similarity transforms $T_{x_0,t_0}^{\varepsilon}$ preserves the total variation in space. This enables us to obtain stronger and more practical convergence criteria than Theorem 3.1.

Recall that $TV_u(t)$ is the total spatial variation of u at the time t. From now on we denote $TTV_u(t_1, t_2)$ and $T\phi_u(t_1, t_2)$ as the total temporal variations of $TV_u(t)$ and $\phi_u(t)$, respectively, from t_1 to t_2. $W(x,t)$ is a traveling discontinuity defined by (2.1) with the two states w^- and w^+, where w^- and w^+ are two distinct real constants in the domain of f. Throughout the remaining part of the paper, we make a convention: the phrase "$(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form" automatically implies that $\mathbf{h}, \boldsymbol{\varepsilon} \in \Upsilon$. We call a sequence of numerical solutions $\{u^k\}_{k=1}^{\infty}$ generated by a $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form (3.1) of the scheme (1.3)-(1.5) *a TV-stable sequence of a numerical traveling discontinuity with the limit* W, if there exist $\boldsymbol{\varepsilon}' \in \Upsilon$ and positive constants C_0 and C such that
   (i) $u^k \to W$ in $L_{loc}^1(\mathbb{R} \times [0, 1))$,
   (ii) $TV_{u^k}(t) < C_0$ for all t and k, and
   (iii) $T\phi_{u^k}(0, 1) < C\varepsilon_k'$ for each k.
When W is a traveling expansion shock, we call $\{u^k\}_{k=1}^{\infty}$ *a TV-stable sequence of a numerical traveling expansion shock.*

THEOREM 3.2. *A semi-discrete scheme of the form* (1.3)-(1.5) *satisfying Assumption* 2.4 *for the Cauchy problem* (1.1) *converges if no* $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled *form* (3.1) *of the scheme is able to generate a TV-stable sequence of a numerical traveling expansion shock.*

Proof. We argue by contradiction. Assume that the convergence of a TVB scheme given by (1.3)-(1.5) fails. By Theorem 3.1 there exists a $(\mathbf{h}', \boldsymbol{\varepsilon}')$-scaled form (3.1) of the scheme which is capable of generating a sequence of functions $\{u^\nu\}_{\nu=1}^\infty$ that converge in $L^1_{loc}(\mathbb{R} \times [0,1))$ to a traveling expansion shock W of the form (2.1). Moreover, $TV_{u^\nu}(t) < C_0$ for all ν. Our goal is to find a sequence $\{\hat{u}^k\}_{k=1}^\infty$ generated by a $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form (3.1) of the scheme such that $\hat{u}^k \to W$ in L^1_{loc}, $TV_{\hat{u}}^k(t) \le C_0$ and $T\phi_{\hat{u}^k}(0,1) \le C\varepsilon'_k$. To this end, by (2.5)-(2.7) and $T\phi_{u^\nu}(t_1, t_2) = \phi_{u^\nu}(t_1) - \phi_{u^\nu}(t_2)$, we have

$$T\phi_{u^\nu}(t_1, t_2) = TV_{u^\nu}(t_1) - TV_{u^\nu}(t_2) + \varepsilon'_\nu \int_{t_1}^{t_2} \sum_{j=-\infty}^\infty s_{j+1/2} \Delta q(u_j^\nu(s)) \, ds$$

$$\le (2 + \bar{C})C_0 := C'_0,$$

where $t_1, t_2 \in [0,1]$, and $\bar{C} := \max_{u \in [\bar{A}, \bar{B}]} |q'(u)|$.

Since ϕ_{u^ν} is monotone decreasing, for any positive integers n and ν, there is an integer $m(n, \nu)$ such that $0 \le m(n, \nu) \le n - 1$ and

$$T\phi_{u^\nu}(m(n,\nu)/n, (m(n,\nu)+1)/n) \le \frac{1}{n} T\phi_{u^\nu}(0,1) \le \frac{C'_0}{n}.$$

Let $t_{n,\nu} = m(n,\nu)/n$, and $x_{n,\nu} = st_{n,\nu}$. For each k, one can first choose a sufficiently large $n = n_k$ so that $C'_0/n_k < \varepsilon'_k$. Then, since $u^\nu \to W$ in $L^1_{loc}(\mathbb{R} \times [0,T))$, one can choose a sufficiently large $\nu = \nu_k$ so that

$$\int_0^1 \int_{st-1}^{st+1} |u^{\nu_k}(x,t) - W(x,t)| \, dx \, dt < \varepsilon'_k/n_k^2,$$

and $h_k := n_k h'_{\nu_k} < \frac{1}{k}$. For simplicity we set $\hat{x}_k = x_{n_k, \nu_k}$, $\hat{t}_k = t_{n_k, \nu_k}$, and $\hat{u}^k(x,t) = T_{\hat{x}_k, \hat{t}_k}^{1/n_k} u^{\nu_k}(x,t)$. We then have

$$\int_0^1 \int_{st-n_k}^{st+n_k} |\hat{u}^k(x,t) - W(x,t)| \, dx \, dt < \varepsilon'_k,$$

since $T_{s\alpha,\alpha}^c W(x,t) = W(x,t)$ for any positive constants α and c. Therefore $\hat{u}^k \to W$ in $L^1_{loc}(\mathbb{R} \times [0,1))$. Next, since $h_k < \frac{1}{k}$, and the source term of $\hat{u}^k$ is $\varepsilon_k q(\hat{u}^k)$, where $\varepsilon_k := \frac{\varepsilon'_{\nu_k}}{n_k} \to 0$ as $k \to \infty$, $\{\hat{u}^k\}_{k=1}^\infty$ is generated by the $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form of the scheme. Moreover, $TV_{\hat{u}^k}(t) \le C_0$ since similarity transforms preserve the spatial variation. Finally, the remark following the proof of Theorem 2.8 implies that

$$T\phi_{\hat{u}^k}(0,1) = T\phi_{u^{\nu_k}}(m(n_k, \nu_k)/n_k, (m(n_k, \nu_k)+1)/n_k) \le \frac{C'_0}{n_k} < \varepsilon'_k.$$

Therefore, $\{\hat{u}^k\}_{k=1}^\infty$ is a TV-stable sequence of numerical traveling expansion shock. The Theorem is proved. □

3.3. Extremum Traceable schemes, general flux f. To connect the numerical flux with the exact flux, we make the following assumption which is stronger than Assumption 2.4. This assumption is needed to develop more practical convergence criteria.

ASSUMPTION 3.3. *The numerical fluxes $g_{j+\frac{1}{2}}(t)$, $j = 0, \pm 1, \pm 2, \cdots$, satisfy*

$$g_{j+\frac{1}{2}}(t) \ge f(u_j) \ge g_{j-\frac{1}{2}}(t) \quad \text{if } u_j(t) - u_{j\pm 1}(t) \ge 0,$$

and

$$g_{j+\frac{1}{2}}(t) \le f(u_j) \le g_{j-\frac{1}{2}}(t) \quad if \;\; u_j(t) - u_{j\pm 1}(t) \le 0.$$

In what follows we will introduce the concepts of *asymptotic traveling discontinuity* (*ATD*) and *asymptotic traveling expansion shock* (*ATES*). As in [22], we use the following notion of paths to be the boundaries of the transition areas of the discontinuities of the numerical solutions.

DEFINITION 3.4. A grid point valued function $x_{I(t)} = I(t)h + c$, $t \in [0, 1]$, is said to be an ε-*path of the first type with respect to a numerical solution* u if for $\varepsilon > 0$ the following conditions hold:

(i) There is a finite partition of $[0, 1]$: $0 = \tau_0 < \tau_1 < \cdots < \tau_n = 1$ such that $I(t)$ is a constant integer on each subinterval $(\tau_{\nu-1}, \tau_\nu)$.

(ii) For all j between $I(\tau_\nu -)$ and $I(\tau_\nu +)$,

$$u_j(\tau_\nu) = u_{I(\tau_\nu -)}(\tau_\nu) = u_{I(\tau_\nu +)}(\tau_\nu).$$

(iii) The following inequality holds:

$$\min(\operatorname{sgn}(\Delta_+ u_{I(t)}(t))\Delta_- u_{I(t)}(t), |\Delta_+ u_{I(t)}(t)|) < \varepsilon.$$

(iv) The total variation of the function $u_{I(t)}(t)$ of t for $0 \le t \le 1$ is bounded by $C\varepsilon$, for some positive constant C.

DEFINITION 3.5. A grid point valued function $x_{I(t)} = I(t)h + c$, $t \in [0, 1]$, is said to be an ε-*path of the second type with respect to a numerical solution* u if for $\varepsilon > 0$ the following conditions hold:

(i) $I(t)$ is a monotone function of t on the interval $[0, 1]$ with the property that there is a finite partition of $[0, 1]$: $0 = \tau_0 < \tau_1 < \cdots < \tau_n = 1$ such that $I(t)$ is a constant integer on each subinterval $(\tau_{\nu-1}, \tau_\nu)$. Moreover, $|I(\tau_\nu +) - I(\tau_\nu -)| = 1$ for $\nu = 1, \cdots, n-1$.

(ii) There is a constant A such that for any $t \in [0, 1]$, $|u_j(t) - A| < \varepsilon/2$ holds if x_j is in the stencil of the scheme at $(x_{I(t)}, t)$.

Along an ε-path of either type, the numerical flux and the exact flux have the following relationship.

LEMMA 3.6 (see Lemma 3.7 in [22] for the corresponding result in the homogeneous case). *Let* $\{u^k\}_{k=1}^{\infty}$ *be a sequence of functions generated by a* $(\mathbf{h}, \boldsymbol{\varepsilon})$-*scaled form* (3.1) *of the scheme* (1.3)-(1.5) *that satisfies Assumption 3.3. For each k, let* $x_{I^k(t)} = I^k(t)h_k + c$ *be an ε_k-path of either type with $\varepsilon_k \le \varepsilon$. We then have*

(3.2) $$\int_0^1 |g_{I^k(t)\pm\frac{1}{2}}(t) - f(u_{I^k(t)}^k(t))|\, dt < C\varepsilon,$$

where C depends on the Lipschitz coefficients of g and $\max_{u \in [\bar{A}, \bar{B}]} |q(u)|$.

Loosely speaking, the essential structure of a traveling wave of a numerical solution is the moving transition from a left limit to a different right limit. The transition

region is bounded either by extremum paths, i.e., ridge (trough) lines, which can be approximated by an ε-path of the first type, or by rim lines which can be approximated by an ε-path of the second type. Now we define the waves rigorously.

DEFINITION 3.7. Let $\{u^k\}_{k=1}^{\infty}$ be a sequence of functions generated by a $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form (3.1) of the scheme (1.3)-(1.5) that satisfies Assumption 3.3. We call a sequence of pairs of ε_k-paths of either type, $\{x^{(k)}(t), y^{(k)}(t)\}$, where $x^{(k)}(t) = x_{I^k(t)}^k = I^k(t)h_k + c_k$, $y^{(k)}(t) = x_{J^k(t)}^k = J^k(t)h_k + c_k$, and $0 \leq t \leq 1$, an *asymptotic traveling wave (ATW) of* $\{u^k\}$ if $x^{(k)}(t) < y^{(k)}(t)$, and if there are a linear function $x(t) = st+r$ and two distinct constants L and R such that:

(i) In the case $s = 0$, for each k, if either path of the pair $\{x^{(k)}(t), y^{(k)}(t)\}$, say $x^{(k)}(t)$, is of the second type, then $x^{(k)}(t)$ is a constant depending only on k.

(ii) Both $x^{(k)}(t)$ and $y^{(k)}(t)$ converges to $x(t)$ uniformly on the t-interval $[0, 1]$ as $k \to \infty$.

(iii) $u_{I^k(t)}^k(t)$ and $u_{J^k(t)}^k(t)$ converges to L and R respectively and uniformly on the t-interval $[0, 1]$ as $k \to \infty$.

For each k, denote by $\widehat{\Omega}^k$ the region $x_{I^k(t)} - h_k/2 < x < x_{J^k(t)} - h_k/2$, $0 \leq t \leq 1$. As in [22] we call the sequence $\{\widehat{\Omega}^k\}$ *the transition region* of the ATW, $x(t)$ *the limit path* of the ATW, L and R *the two states* of the ATW.

In order to study entropy properties of an ATW, as Osher in [12], for any convex entropy $U(w)$ and its flux $F(w)$, we adopt the numerical entropy flux $G_{j-\frac{1}{2}}(t) \overset{\text{def}}{=} F(u_j) + U'(u_j)[g_{j-\frac{1}{2}} - f(u_j)]$. Applying Lemma 3.6 for the conservation laws of the form: $U(w)_t + F(w)_x = \varepsilon_k U'(w)q(w)$ with $\boldsymbol{\varepsilon} \in \Upsilon$, then $G_{j-\frac{1}{2}}(t)$ satisfies the following.

COROLLARY 3.8. *If* $\{u^k\}_{k=1}^{\infty}$ *satisfies the conditions of Lemma* 3.6, *then*

$$\int_0^1 |G_{I^k(t)-\frac{1}{2}}(t) - F(u_{I^k(t)}^k(t))| dt < C\varepsilon.$$

Adapting Osher's proof in [12], we have the following equality for $\{u^k\}_{k=1}^{\infty}$ generated by a $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form (3.1) of the scheme (1.3)-(1.5).

$$(3.3) \qquad h_k[\frac{d}{dt}U(u_j^k(t)) + D_+ G_{j-\frac{1}{2}}(t) - \varepsilon_k U'(u_j^k(t))q(u_j^k(t))]$$

$$= \int_{u_j^k(t)}^{u_{j+1}^k(t)} U''(w)(g_{j+\frac{1}{2}} - f(w)) dw.$$

Let $\phi(x, t)$ be a smooth function with compact support in the domain $-\infty < x < \infty$, $0 < t < 1$. Set $\phi_j(t) = \phi(x_j, t)$ and define

$$(3.4) \quad \widehat{\Phi}^k \overset{\text{def}}{=} \int_0^1 \sum_{j=I^k(t)}^{J^k(t)-1} h_k[\frac{d}{dt}U(u_j^k(t)) + D_+ G_{j-\frac{1}{2}}^k(t) - \varepsilon_k U'(u_j^k(t))q(u_j^k(t))]\phi_j(t) dt.$$

We have the following important result.

LEMMA 3.9 (see Lemma 3.10 in [22] for the result of the HCP of the scheme (1.3)-(1.5)). *Suppose* $\{u^k(x, t)\}_{k=1}^{\infty}$ *satisfies the conditions of Lemma* 3.6. *Let* $\{I^k(t)h_k + $

$c_k, J^k(t)h_k + c_k\}$ *be an ATW of* $\{u^k(x,t)\}_{k=1}^{\infty}$ *with the limit path* $x(t) = st$ *and the two states* L *and* R. *We then have*

$$(3.5) \qquad \lim_{k \to \infty} \widehat{\Phi}^k = [F(R) - F(L) - s(U(R) - U(L))] \int_{x=st} \phi(x,t)dt.$$

Choosing $U(w) = w$, we have $\widehat{\Phi}^k = 0$. Hence $f(R) - f(L) = s(R - L)$, and an integration by parts give the second equality of the following Corollary.

COROLLARY 3.10. *With the conditions of Lemma 3.9, we have the following discrete Rankine-Hugoniot condition:*

$$(3.6) \qquad\qquad\qquad f(R) - f(L) = s(R - L).$$

Moreover

$$(3.7) \qquad \lim_{k \to \infty} \widehat{\Phi}^k = \int_L^R U''(w)(f[w; L, R] - f(w))dw \int_{x=st} \phi(x,t)dt.$$

DEFINITION 3.11. An ATW $\{x^k(t), y^k(t)\}_{k=1}^{\infty}$ of $\{u^k\}_{k=1}^{\infty}$ is called *an ATD of* $\{u^k\}_{k=1}^{\infty}$ if the ATW is essentially monotone in space. Namely, for each k and for $t \in [0, 1]$,

(i) if p and q are any integers such that $I^k(t) \le p < q \le J^k(t)$, then

$$-(u_q^k(t) - u_p^k(t))\mathrm{sgn}(R - L) \le \varepsilon_k,$$

and

(ii) if j is an integer such that $I^k(t) \le j \le J^k(t)$ and that $(u_{j+1}^k(t) - u_j^k(t))(R - L) < 0$, then

$$u_j^k(t), u_{j+1}^k(t) \in \mathcal{N}_{\delta_k} (\{w : \ f(w) = f[w; L, R]\})$$

for some $\boldsymbol{\delta} \in \Upsilon$, where $\mathcal{N}_\delta(S)$ denotes the δ-neighborhood of a set S. An ATD of $\{u^k\}_{k=1}^{\infty}$ is called an *ATES* of $\{u^k\}_{k=1}^{\infty}$ if the entropy condition (2.3) with $w^- = L$ and $w^+ = R$ fails. In the last case we also say that $\{u^k\}_{k=1}^{\infty}$ *harbors the ATES* $\{x^k(t), y^k(t)\}_{k=1}^{\infty}$.

Now we are ready to state the third WEI criterion for the convergence.

THEOREM 3.12. *A semi-discrete scheme of the form* (1.3)-(1.5) *satisfying Assumption 3.3 converges if no* $(\mathbf{h}, \boldsymbol{\varepsilon})$-*scaled form* (3.1) *is able to create a sequence of functions* $\{u^k(x,t)\}_{k=1}^{\infty}$ *that harbors an ATES.*

Proof. [Sketch of the proof] (for details see the proof of Theorem 3.13 in [22] for HCP of the scheme) Again, we argue by contradiction. If the convergence fails, then by Theorem 3.2 there exists a TV-stable sequence $\{u^k\}$ of a numerical traveling expansion shock generated by a $(\mathbf{h}, \boldsymbol{\varepsilon})$-scaled form (3.1) of the scheme. Since $TV_{u^k}(t) < C_0$ for all k and $t \in [0, 1]$ and since $\{u^k\}$ converges to W in $L_{loc}^1(\mathbb{R} \times [0, 1))$, in the compact domain $\Omega = \{(x,t) : (x,t) \in [st - 1, st + 1] \times [0, 1]\}$ there may exist

at most uniformly bounded number of large oscillations which asymptotically either travel away from the line $x = st$ as infinitesimally thin spikes in the graph of the numerical solutions, or move along the line. If the scheme is extremum traceable, i.e., non-oscillatory, then we can use approximate extremum paths to track these oscillations. Since the sequence $\{u^k\}$ is TV-stable, the amplitudes of these oscillations are essentially stationary, and it contains a subsequence, in which the approximate paths becomes ε-paths as described in Definitions 3.4 and 3.5. Using similarity transforms and selecting subsequences, we may push those oscillations which do not travel along the line $x = st$ out of the interested domain Ω (this effect is called wave separation). Hence, all the strong oscillations which remain in Ω travel along the line (this effect is called wave concentration). Finally it can be shown that these oscillations consists of finite number of strong ATWs which dominated the entropy estimate and the oscillations of small amplitude whose contributions to the entropy estimate are negligible, and at least one of the strong ATWs must be an ATES (this analysis is called wave splitting). The entire proof can be directly translated from that for Theorem 3.13 in [22], except the extreme tracking theory is in question now, since with the source term the numerical solution, in general, does not satisfy the Local Maximum Principal (LMP). Fortunately, a careful analysis shows that the extrema can still be traced and we devote §4 to this analysis. □

3.4. Extremum traceable schemes, convex flux f. Now we focus on the case $f''(w) \geq 0$, i.e., the convex conservation laws. We call a pair of numbers $\{L, R\}$ a *rarefying pair* if $L < R$ and $f[w; L, R] > f(w)$ when $L < w < R$. And we call a collection of data $\{v_j\}_{j=-p}^{n+p}$ a *rarefying collection with respect to the pair* $\{L, R\}$ if $L = v_0 \leq v_1 \leq \cdots \leq v_n = R$, and $L \leq v_{-1}$, and $R \geq v_{n+1}$. Let $\bar{g}_{j+1/2} := g(v_{j-p+1}, v_{j-p+2}, \cdots, v_{j+p})$, where g is the function (1.4) in its self-similar form.

Our fourth, and the last WEI criterion of convergence states that the WEI across the area of the rarefaction is sufficient for convergence.

THEOREM 3.13. *A scheme of the form* (1.3)-(1.5) *satisfying Assumption* 3.3 *converges for convex conservation laws if, for any rarefying pair* $\{L, R\}$, *there is a constant* $\delta > 0$ *such that the quadrature inequality*

$$(3.8) \qquad \sum_{j=0}^{n-1}(v_{j+1} - v_j)\bar{g}_{j+1/2} + \delta < \int_L^R f[w; L, R]dw$$

holds for all rarefying collections $\{v_j\}_{j=-p}^{n+p}$ *with respect to the pair* $\{L, R\}$.

It is startling that the theorem is exactly the same as the corresponding one for the HCP of the scheme in [22]. Indeed, the source term does not appear in the criterion. Therefore, all the concrete schemes whose convergence was established in [22] by WEI remain convergent in the presence of the source term.

For completeness, we end this section by listing these results which establish the convergence to the entropy solution of generalized MUSCL schemes and a class of schemes using flux limiters. The building blocks for both classes are called E schemes [12]. Let $g^E(\cdot, \cdot)$ be the flux of any E scheme, i.e., it is Lipschitz continuous, and satisfies

$$(3.9) \qquad \text{sgn}(w_{j+1} - w_j)(g^E(w_j, w_{j+1}) - f(w)) \leq 0$$

for all w between w_j and w_{j+1}. First, we give the result of the class of generalized MUSCL schemes

$$(3.10) \qquad \frac{du_j}{dt} = -\frac{1}{h}\Delta_+ g^E(u_{j-1} + s_{j-1}h/2, u_j - s_j h/2) + q(u_j).$$

THEOREM 3.14. *A generalized MUSCL scheme with an E scheme building block converges provided that the slope s_j satisfies: $s_j = 0$, if u_j is an extremum; $0 \leq \frac{hs_j}{\Delta_+ u_j}, \frac{hs_j}{\Delta_- u_j} \leq 2$, if u_{j-1} is a maximum or u_{j+1} is a minimum; and $0 \leq \frac{hs_j}{\Delta_+ u_j}, \frac{hs_j}{\Delta_- u_j} \leq 1$ holds if $u_j < u_{j+1}$.*

Next, we give the convergence results of a large class of high resolution schemes using flux limiters in their semi-discrete form with source terms. They have been studied by Sweby In [18]. Denote

$$(\Delta f_{j+\frac{1}{2}})^+ := f(u_{j+1}) - g^E(u_j, u_{j+1}), \quad (\Delta f_{j+\frac{1}{2}})^- := f(u_j) - g^E(u_j, u_{j+1}),$$

$$r_j^+ := (\Delta f_{j-\frac{1}{2}})^+/(\Delta f_{j+\frac{1}{2}})^+, \qquad r_j^- := (\Delta f_{j+\frac{1}{2}})^-/(\Delta f_{j-\frac{1}{2}})^-,$$

and $(Df_{j+\frac{1}{2}})^\pm := (\Delta f_{j+\frac{1}{2}})^\pm/\Delta_+ u_j$. With the Chakravarthy and Osher limiters (see [14]):

$$(3.11) \qquad \psi_c(r) = \begin{cases} 0 & r < 0, \\ r & 0 \leq r < c, \\ c & r \geq c, \end{cases}$$

and the numerical flux defined by

$$(3.12) \qquad g_{j+\frac{1}{2}} = g^E(u_j, u_{j+1}) + \frac{1}{2}\psi(r_j^+)(\Delta f_{j+\frac{1}{2}})^+ + \frac{1}{2}\psi(r_{j+1}^-)(\Delta f_{j+\frac{1}{2}})^-,$$

we have the following general result.

THEOREM 3.15. *The numerical solutions of the schemes (1.3)-(1.5) converge for the convex problem (1.1) provided that the numerical flux $g_{j+\frac{1}{2}}$ satisfies*
    (i) $g_{j+\frac{1}{2}} \geq f(u_j)$ *if* $u_j - u_{j\pm1} \geq 0$, *and* $g_{j-\frac{1}{2}} \geq f(u_j)$ *if* $u_j - u_{j\pm1} \leq 0$.
    (ii) *When $u_{j+1} > u_j$, $g_{j+\frac{1}{2}}$ is given by (3.12), where g^E is the numerical flux function of any E scheme with $\psi(r) = \psi_1(r)$ defined by (3.11).*

The result can be improved, when the building blocks of the high resolution schemes are well known monotone schemes such as the Godunov, the Engquist–Osher, or Lax-Friedrichs schemes.

THEOREM 3.16. *The numerical solutions of the schemes (1.3)-(1.5) converge for the convex problem (1.1) provided that the numerical flux $g_{j+\frac{1}{2}}$ satisfies:*
    (i) $g_{j+\frac{1}{2}} \geq f(u_j)$ *if* $u_j - u_{j\pm1} \geq 0$, *and* $g_{j-\frac{1}{2}} \geq f(u_j)$ *if* $u_j - u_{j\pm1} \leq 0$.
    (ii) *When $u_{j+1} > u_j$, $g_{j+\frac{1}{2}}$ is defined by (3.12), where $\psi(r) = \psi_c(r)$ is given by (3.11) with $1 \leq c \leq 2$, provided that $g^E(\cdot, \cdot)$ is the numerical flux of one of the following monotone schemes.*

The Godunov scheme:

$$g^{\text{God}}(u_j, u_{j+1}) = \begin{cases} \min_{u_j \leq w \leq u_{j+1}} f(w) & \text{when} \quad u_j \leq u_{j+1}, \\ \max_{u_j \geq w \geq u_{j+1}} f(w) & \text{when} \quad u_j \geq u_{j+1}. \end{cases}$$

The Engquist–Osher scheme:

$$g^{\text{EO}}(u_j, u_{j+1}) = \int_0^{u_j} \max(f'(w), 0)dw + \int_0^{u_{j+1}} \min(f'(w), 0)dw + f(0).$$

The Lax-Friedrichs scheme:

$$g^{\text{LF}}(u_j, u_{j+1}) = \frac{f(u_j) + f(u_{j+1})}{2} - \frac{a}{2}(u_{j+1} - u_j)$$

where $a \geq \max |f'(w)|$.

4. Extended extremum tracking theory. We now extend the extremum tracking theory to the non-homogeneous scheme (1.3)-(1.5), as needed in the proof of Theorem 3.12. We begin with a simple fact which comes from the Lipschitz continuity of the numerical flux function and the source term.

LEMMA 4.1. *If u is a uniformly bounded numerical solution of the scheme* (1.3)-(1.5), *then, for $h \leq 1$, there are two positive constants C_1 and C_2 such that*

$$|\frac{du_j(t)}{dt}| \leq C_1/h \quad and \quad |\frac{du_j(t'')}{dt} - \frac{du_j(t')}{dt}| \leq (C_2/h^2)|t'' - t'|.$$

Throughout this section, C_1 and C_2 are exclusively used for these two constants. The set $X = \{x_j\}_{j=-\infty}^{\infty}$ is called the *set of grid points* and $L = X \times [0, T)$ the *set of grid line*. The solution u is defined on L. A finite set of successive grid points $\{x_p, \ldots, x_q\}$ with $q \geq p$ is called the *stencil of a spatial maximum*, or simply *an M-stencil* of u at the time t, provided $u_p(t) = \cdots = u_q(t)$, $u_{p-1}(t) < u_p(t)$, and $u_{q+1}(t) < u_q(t)$. Notions of *N-stencils* for minima and *E-stencils* for general extrema are defined similarly.

The following lemma still holds for the numerical solution u generated by the non-homogeneous scheme (1.3)-(1.5).

LEMMA 4.2 (see Lemma 6.2. in [22] for the homogeneous case). *Suppose $\{x_p, \ldots, x_q\}$, $q \geq p$, is an M-stencil of u at the time t_0. Then there exists a $\delta > 0$ such that, when $|t - t_0| < \delta$, there exists at least one M-stencil $\{x_{p'}, \ldots, x_{q'}\}$ of u at the time t that is a subset of $\{x_p, \ldots, x_q\}$.*

Nothing needs to be changed when we adopt the notion of extremum paths [22] for the non-homogeneous schemes.

DEFINITION 4.3 (The extremum paths). A nonempty subset of $\{(x, t) \in L : t_1 \leq t \leq t_2\}$ denoted by M_{t_1, t_2} is called *a ridge of the numerical solution u from t_1 to t_2* if
   (i) For each $t \in [t_1, t_2]$ the set

$$P_M(t) \stackrel{\text{def}}{=} \{x_j : (x_j, t) \in M_{t_1, t_2}\} = \{x_{p(t)}, \ldots, x_{q(t)}\}$$

is non empty, and is an M-stencil of u at t.

(ii) M_{t_1,t_2} satisfies the following "connectivity" condition: for each t in $[t_1, t_2]$, there exists a neighborhood $U(t)$ of t such that if $t' \in U(t) \cap [t_1, t_2]$, then $P_M(t') \subseteq P_M(t)$.

The set $P_M(t)$ is called *the x-projection of M_{t_1,t_2} at* t. The value of u along the ridge is denoted by $V_M(t)$: $V_M(t) = u_j(t)$ for $p(t) \leq j \leq q(t)$. Hereafter, we use the notations M_{t_1,t_2}, $P_M(t)$ and $V_M(t)$ exclusively for the above notions. If for all $t \in [t_1, t_2]$ the M-stencil in (i) and (ii) is replaced by an N-stencil, then the set is called a *trough of u from t_1 to t_2*, and is denoted by N_{t_1,t_2}. The related notions $P_N(t)$ and $V_N(t)$ are defined similarly. Ridges and troughs are also called *extremum paths*. When we do not distinguish between ridges and troughs, we use E_{t_1,t_2}, $P_E(t)$ and $V_E(t)$ for either type. We add superscripts on M, N, or E to indicate several paths in one solution, sequences of paths associated with a sequence of solutions, or both. We make the convention that $E^1_{t_1,t_2} < (\leq)E^2_{t_1,t_2}$ when $\max P_{E^1}(t) < (\leq) \max P_{E^2}(t)$ for all $t \in [t_1, t_2]$. For $t'' > t'$, we say that a given E-stencil of u at t'' *can be traced back to t'* if it is the x-projection of an extremum path $E_{t',t''}$ at t''.

Two local non-oscillatory properties (LNOPs) of the scheme (1.3)-(1.5) are the foundation of the extremum tracking theory. These properties imply that no new extremum may emerge for $t > 0$. Hence one can trace any given E-stencil back to the initial time $t = 0$. In the homogeneous case, Yang [22] proved the two LNOPs by using the LMP: If x_i belongs to an $M(N)$-stencil of $u_j(t)$ at the time t, then $\frac{du_i(t)}{dt} \leq (\geq)0$. For the scheme with a source term, the LMP, in general, does not hold. Surprisingly, the LNOPs, as described in the following two lemmas, still hold, provided that the numerical flux function satisfies the same Assumption 2.4 that enforces the LMP on its HCP.

LEMMA 4.4 (LNOP 1). *With Assumption 2.4, if $\{x_p, \ldots, x_q\}$ is an M-stencil of u at some t_0, then there exists a $\delta > 0$ such that at any $t \in [t_0, t_0 + \delta)$, there is no N-stencil of u that is a subset of $\{x_p, \ldots, x_q\}$.*

Proof. Since u is continuous in t and $\{x_p, \ldots, x_q\}$ is an M-stencil of u at t_0, there exists a $\delta > 0$ such that for $t \in [t_0, t_0 + \delta)$,

$$\max(u_{p-1}(t), u_{q+1}(t)) < \min_{p \leq j \leq q} u_j(t).$$

If $q \leq p + 1$, the lemma holds trivially. Otherwise, for any j with $p < j < q$, set

$$D_j(t) \stackrel{\text{def}}{=} \min(\max_{p \leq l < j}(u_l(t) - u_j(t))_+, \max_{j < r \leq q}(u_r(t) - u_j(t))_+)$$

and

$$D(t) \stackrel{\text{def}}{=} \max_{p < j < q} D_j(t).$$

Here, $(a)_+ = \max(a, 0)$. Clearly, $D(t) \geq 0$, and $D(t) = 0$ if and only if there is no N-stencil of u at t that is a subset of $\{x_p, \ldots, x_q\}$. The function $D(t)$ is clearly Lipschitz continuous. Therefore, if

$$I := \{t : D(t) > 0\} \cap (t_0, t_0 + \delta) \neq \emptyset,$$

there are real numbers α and β satisfying $t_0 \leq \alpha < \beta \leq t_0 + \delta$ with $D(\alpha) = 0$ and $D(t) > 0$ for $\alpha < t < \beta$. Moreover, $D(t)$ is differentiable almost everywhere in the

interval (α, β) and

$$(4.1) \qquad\qquad D(t) = \int_\alpha^t D'(s)\,ds.$$

For any $t \in (\alpha, \beta)$ where $D'(t)$ is defined, there exist an i with $p < i < q$, an l with $p \le l < i$, and an r with $i < r \le q$ such that

$$D(t) = \min(u_l(t) - u_i(t), u_r(t) - u_i(t)).$$

Clearly, x_l and x_r each belongs to an M-stencil of u, and x_i belongs to an N-stencil of u at t. Hence, Assumption 2.4 implies

$$\frac{du_l(t)}{ds} \le q(u_l(t)), \quad \frac{du_r(t)}{ds} \le q(u_r(t)), \quad \text{and} \quad \frac{du_i(t)}{ds} \ge q(u_i(t)).$$

Hence, if $u_l(t) - u_i(t) = u_r(t) - u_i(t)$, then

$$(4.2) \qquad \begin{aligned} D'(t) &\le \max(u_l'(t) - u_i'(t), u_r'(t) - u_i'(t)) \\ &\le \max(q(u_l(t)) - q(u_i(t)), q(u_r(t)) - q(u_i(t))) \le \bar{C} D(t). \end{aligned}$$

Otherwise, without loss of generality, assume $u_l(t) - u_i(t) < u_r(t) - u_i(t)$. By definition, for any $s \in (\alpha, \beta)$, we have $D(s) \ge \min(u_l(s) - u_i(s), u_r(s) - u_i(s))$. Thus, $D(s) \ge u_l(s) - u_i(s)$ when $|s - t|$ is sufficiently small and the equality holds when $s = t$. This implies that

$$(4.3) \qquad\qquad D'(t) = u_l'(t) - u_i'(t) \le \bar{C}(u_l(t) - u_i(t)) = \bar{C} D(t).$$

Now from (4.2), (4.3) and (4.1), we obtain $D(t) \le \int_\alpha^t \bar{C} D(s)\,ds$ for all $t \in (\alpha, \beta)$. Applying Gronwall's inequality, we get $D(t) \equiv 0$ for all $t \in (\alpha, \beta)$. This contradicts the inequality $D(t) > 0$ for all $t \in (\alpha, \beta)$. Hence, $I = \emptyset$ as desired. □

LEMMA 4.5 (LNOP 2). *With Assumption* 2.4, *if*

$$u_{p-1}(t_0) > u_p(t_0) \ge \cdots \ge u_q(t_0) > u_{q+1}(t_0),$$

then there exists a $\delta > 0$ such that for $t \in [t_0, t_0 + \delta)$

$$(4.4) \qquad\qquad u_{p-1}(t) > u_p(t) \ge \cdots \ge u_q(t) > u_{q+1}(t).$$

Proof. Since the numerical solution is continuous in t, there exists a $\delta > 0$ such that

$$u_{p-1}(t) > \max_{p \le i \le q} u_i(t) \ge \min_{p \le i \le q} u_i(t) > u_{q+1}(t)$$

for $t \in [t_0, t_0 + \delta)$. Define

$$D(t) = \max_{p \le l < q} \big(\max_{l < r \le q} (u_r(t) - u_l(t))_+ \big).$$

Clearly, $D(t) \ge 0$, and $D(t) = 0$ if and only if (4.4) holds. It therefore suffices to show that $D(t) = 0$ for all $t \in [t_0, t_0 + \delta)$. We omit the rest of the proof, which is similar to that of Lemma 4.4. □

Without LMP, the two Monotonicity Properties (Lemmas 6.7, 6.8 in [22]) need a minor modification for the scheme (1.3)-(1.5). For M_{t_1,t_2}, define

$$R_M(t) := \inf_{\max P_M(t) < j < \infty} u_j(t).$$

Mimicking the proof of Lemma 2.6, one can easily show the following lemma.

LEMMA 4.6. *With Assumption 2.4, if M_{t_1,t_2} is a ridge of u, then for any $t_1 \leq t \leq t_2$, we have*
  (i) $V_M(t) \leq V_M(t_1) + \int_{t_1}^t q(V_M(s)) \, ds$,
  (ii) $R_M(t) \geq R_M(t_1) + \int_{t_1}^t q(R_M(s)) \, ds$.

Let $C' := \max_{u \in [\bar{A}, \bar{B}]} |q(u)|$. Immediately we have the following corollary.

COROLLARY 4.7. *With Assumption 2.4, if M_{t_1,t_2} is a ridge of u, then for any t', t'' with $t_1 \leq t' \leq t'' \leq t_2$, we have*
  (i) $V_M(t'') \leq V_M(t') + C'(t'' - t')$,
  (ii) $R_M(t'') \geq R(t') - C'(t'' - t')$.

The following two lemmas come from [22], both still hold for the numerical solution u generated by the scheme (1.3)-(1.5). The first one implies that along its route of propagation, an extremum path of u sweeps over all the grid points in between. The next lemma implies that the forward extremum path of u starting from an E-stencil at any time t_0 is unique.

LEMMA 4.8 (Sweeping Over Property). *If M_{t_1,t_2} is a ridge of u such that*

$$x_p = min[P_M(t_1) \bigcup P_M(t_2)] \quad and \quad x_q = max[P_M(t_1) \bigcup P_M(t_2)],$$

then $\bigcup_{t_1 \leq t \leq t_2} P_M(t) \supseteq \{x_j : p \leq j \leq q\}$.

LEMMA 4.9 (Order Preserving Property). *Suppose the numerical solution u satisfies Assumption 2.4 and suppose $E_{t_1,t_2}^{(1)}$ and $E_{t_1,t_2}^{(2)}$ are two extremum paths of u. If $max P_{E^{(1)}}(t_2) < min P_{E^{(2)}}(t_2)$, then $E_{t_1,t_2}^{(1)} < E_{t_1,t_2}^{(2)}$.*

We now prove the main result of this section.

LEMMA 4.10 (Backward Trace-ability Property). *With the Assumption 2.4, any M-stencil $\{x_p, \ldots, x_q\}$ at any $T_1 > 0$ can be traced back to $t = 0$.*

Proof. Let $\Im$ be the set of t for which there exists an M_{t,T_1} with $P_M(T_1) = \{x_p, x_{p+1}, \ldots, x_q\}$. We prove the lemma by showing that $\Im$ is nonempty and is both open and closed in $[0, T_1]$. It is obvious that $\Im$ is nonempty, since $T_1 \in \Im$. The proof of "open" is similar to the one in [22]. We give the proof for closeness.

"Closed": For any $\tau, \tau' \in (0, T_1)$, if $\tau > \tau'$, then $\tau' \in \Im$ implies that $\tau \in \Im$. Thus, to show the closeness it suffices to show that for any $\{\tau_\mu\}_{\mu=0}^\infty \subset \Im$ with $\tau_1 > \tau_2 > \cdots$ and $\tau_\mu \to t'$, we have $t' \in \Im$.

First, We claim that $x_{p^\mu(t)}$ and $x_{q^\mu(t)}$ are uniformly bounded: there exists a constant $X_{\max}$ such that for all μ, $x_{q^\mu(t)} < X_{\max}$ when $\tau_\mu \leq t \leq T_1$. Similarly, there exists a constant $X_{\min}$ that is a lower bound of $x_{p^\mu(t)}$.

Since $x_{q^\mu(T_1)} = x_{q(T_1)}$ for all μ, it suffices to show that there exists a constant Δt such that for any $t_0 \in (t', T_1]$, there exists a constant $X_{\max}(t_0)$ for which $x_{q^\mu(t)} \le X_{\max}(t_0)$ for all μ and $t_0 - \Delta t \le t \le t_0$ provided that $\{x_{q^\mu(t_0)}\}_{\mu=0}^\infty$ has a upper bound $x_{\sup}(t_0)$.

Indeed, applying Lemma 4.6 and argument similar to the proof of Theorem 2.5 shows that we have $R_{M^\mu}(t) \le \alpha(t) < \beta(t) \le V_{M^\mu}(t)$, for all μ and $\tau_\mu \le t \le T_1$, where $\alpha(t)$ and $\beta(t)$ are the solutions of the ordinary differential equation $v'(t) = q(v(t))$ with $\alpha(T_1) = R_{M^\mu}(T_1)$ and $\beta(T_1) = V_{M^\mu}(T_1)$ respectively. Define

$$\delta = \min_{t' \le t \le T_1} \frac{1}{4}(\beta(t) - \alpha(t)).$$

Since $\beta(t)$ is uniformly continuous on $[0, T_1]$, there is a η such that

$$|\beta(s) - \beta(r)| \le \delta,$$

provided that $s, r \in [0, T_1]$ and $|s - r| \le \eta$. Now let $\Delta t = \min(\frac{\delta h}{C_1}, \eta)$. By the definition of $R_{M^\mu}(t)$, there exists an j_0 with $x_{j_0} \ge x_{\sup}(x_0)$ for which $u_{j_0}(t) \le \alpha(t_0) + \delta$ for $t \in [t_0 - \Delta t, t_0]$. The inequality $\Delta t \le \eta$ also implies that on the same interval

$$V_{M^\mu}(t) \ge \beta(t) = \beta(t_0) + (\beta(t) - \beta(t_0)) \ge \beta(t_0) - \delta.$$

Hence

$$V_{M^\mu}(t) - u_{j_0}(t) \ge \beta(t_0) - \alpha(t_0) - 2\delta \ge 2\delta > 0.$$

Therefore, the Sweeping over Lemma 4.8 implies that for $t \in [t_0 - \Delta t, t_0]$, we have

$$q^\mu(t) \le x_{j_0}.$$

This prove the claim.

Next, we show that there exists a subsequence $\{M_{\tau_\mu, T_1}\}_{\mu=1}^\infty$ of ridges $\{M_{\tau_\mu, T_1}^\mu\}_{\mu=1}^\infty$, such that $P_M(T_1) = \{x_p, x_{p+1}, \ldots, x_q\}$ and

(4.5) $$M_{\tau_{\mu_1}, T_1} \subset M_{\tau_{\mu_2}, T_1} \quad \text{if} \quad \mu_1 < \mu_2.$$

Because of (4.5), it is not necessary to use superscripts to distinguish different ridges in this sequence.

To see this, let us assume that $\{P_i^\mu\}_{i=1}^{m_\mu}$ is the collection of M-stencils of u at τ_μ between $X_{\min}$ and $X_{\max}$. Since $x_{p^\mu(t)}$ and $x_{q^\mu(t)}$ are uniformly bounded, $\{m_\mu\}_{\mu=1}^\infty$ is a bounded sequence. Inductively, we select an M-stencil $P^\mu = P_{i_\mu}^\mu$ at τ_μ for each natural number μ as follows: We are able to select a $P^1 = P_{i_1}^1$ such that it consists of the x-projections at τ_1 of an infinite subsequence $(SQ)_1$ of $\{M_{\tau_\mu, T_1}^\mu\}_{\mu=1}^\infty$ because m_1 is bounded. In general, if $(SQ)_\mu$ and P^μ have been selected, we are able to select a $P^{\mu+1} = P_{i_{\mu+1}}^{\mu+1}$ such that it consists of the x-projections at $\tau_{\mu+1}$ of an infinite subsequence $(SQ)_{\mu+1}$ of $(SQ)_\mu$ because $\{m_\mu\}_{\mu=1}^\infty$ is a bounded sequence. Thus, P^μ is selected for every natural number μ. Now we define M_{τ_μ, T_1} to be the unique ridge of u such that $P_M(\tau_\mu) = P^\mu$ (uniqueness is guaranteed by Lemma 4.9). The relation (4.5) follows from the uniqueness.

Finally, suppose that $P_M(t) = \{x_{p(t)}, \ldots, x_{q(t)}\}$ and that

$$\widetilde{p} = \lim_{\mu \to \infty} \inf p(\tau_\mu), \quad \text{and} \quad \widetilde{q} = \lim_{\mu \to \infty} \sup q(\tau_\mu).$$

We have $\widetilde{p} \leq \widetilde{q}$. Since $p(\tau_\mu)$ and $q(\tau_\mu)$ are uniformly bounded, $q(\tau_\mu) \leq \widetilde{q}$ and $p(\tau_\mu) \geq \widetilde{p}$ for sufficiently large μ. Moreover, Lemmas 4.5 and 4.8 and the continuity of u in t imply that

$$u_{\widetilde{p}}(t') = \cdots = u_{\widetilde{q}}(t') = \lim_{\mu \to \infty} V(\tau_\mu).$$

Let $\widehat{p}$ and $\widehat{q}$ be integers such that $\widehat{p} \leq \widetilde{p} \leq \widetilde{q} \leq \widehat{q}$ with

$$u_{\widehat{p}}(t') = \cdots = u_{\widehat{q}}(t'),$$

and

$$u_{\widehat{p}-1}(t') \neq u_{\widehat{p}}(t'), \quad u_{\widehat{q}+1}(t') \neq u_{\widehat{q}}(t').$$

The collection $\{x_{\widehat{p}}, \ldots, x_{\widehat{q}}\}$ is an M-stencil of u at t', because otherwise, by Lemmas 4.4 and 4.5, $\{x_{\widehat{p}-1}, \ldots, x_{\widehat{q}+1}\}$ would not contain any M-stencil at $t = \tau_\mu$ for sufficiently large μ and the existence of the aforementioned M_{τ_μ, T_1} would be violated. Set $P_M(t') = \{x_{\widehat{p}}, \ldots, x_{\widehat{q}}\}$. Then M_{t', T_1} is well defined. Hence $t' \in \Im$. This shows that $\Im$ is closed in $[0, T_1]$. $\square$

Given a path E_{t_1, t_2}, we would like, ideally, to have a function $x_E(t) : [t_1, t_2] \to \mathbb{R}$ that satisfies the following conditions.

(1) There is a finite partition of $[t_1, t_2] : t_1 = \tau_0 < \tau_1 < \cdots < \tau_n = t_2$ such that $x_E(t)$ is a constant on each subinterval $(\tau_{\mu-1}, \tau_\mu)$ for $1 \leq \mu \leq n$.

(2) $x_E(t) \in P_E(t)$ for $t \in [t_1, t_2]$.

It is not clear that such a function $x_E(t)$ exists. However, we can construct a piecewise constant function of t that represent E_{t_1, t_2} sufficiently well. Without loss of generality, we only discuss the spatial maxima.

LEMMA 4.11. *For any $\varepsilon > 0$ and for each ridge M_{t_1, t_2} of a numerical solution u generated by the scheme (1.3)-(1.5) satisfying Assumption 2.4, there exists a piecewise constant function $x_M^\varepsilon(t) : [t_1, t_2] \to \mathbb{R}$ and a finite partition of $[t_1, t_2] : t_1 = \tau_0 < \tau_1 < \cdots < \tau_n = t_2$ such that $x_M^\varepsilon(t) \equiv c_\nu$ and*

$$(4.6) \qquad c_\nu \in P_M(\tau_{\nu-1}) \bigcap P_M(\tau_\nu),$$

on each subinterval $(\tau_{\nu-1}, \tau_\nu)$ for $1 \leq \nu \leq n$. Furthermore, at any $t \in [t_1, t_2]$, for any x_i between (and including) $x_M^\varepsilon(t)$ and $P_M(t)$, we have

$$(4.7) \qquad |V_M(t) - u_i(t)| < \varepsilon.$$

Finally, for $V_M^\varepsilon(t) \overset{\text{def}}{=} u(x_M^\varepsilon(t), t)$,

$$(4.8) \qquad TV_{t_1 \leq t \leq t_2}(V_M^\varepsilon(t)) < V_M(t_1) - V_M(t_2) + C.$$

We call the function $x_M^\varepsilon(t)$ an *ε-ridge to M_{t_1, t_2}*. Similarly, we may define an *ε-trough $x_N^\varepsilon(t)$* to an trough N_{t_1, t_2}. Both ε-ridges and ε-troughs are called *ε-E paths*.

Let $E_{t, t'}^1, E_{t, t'}^2$ be two extremum paths of u. In the following, the total spatial variation of u at s between $E_{t, t'}^1$ and $E_{t, t'}^2$ is denoted by $TV_{E^1 E^2}(s)$. We also denote $TTV_{E^1 E^2}(t, t')$ as the temporal variation of $TV_{E^1 E^2}(s)$ from t to t'.

LEMMA 4.12. *Suppose the numerical solution u of the scheme (1.3)-(1.5) satisfies Assumption 2.4. Suppose also that $E^1_{t,t'} \leq E^2_{t,t'}$ are two extremum paths of u such that either $P_{E^1}(t')$ and $P_{E^2}(t')$ are identical or they are two successive E-stencils of u at t'. Let $C = \max_{u \in [\bar{A}, \bar{B}]} |q(u)|(t' - t)$ and let*

$$u^M = \max_{\min P_{E^1}(t') \leq x \leq \max P_{E^2}(t')} u(x, t'), \; u^N = \min_{\min P_{E^1}(t') \leq x \leq \max P_{E^2}(t')} u(x, t').$$

(i) *If $\min P_{E^1}(t) \leq x_j \leq \max P_{E^2}(t)$, then*

$$u^N - TTV_{E^1 E^2}(t, t') - C \leq u_j(t) \leq u^M + TTV_{E^1 E^2}(t, t') + C.$$

(ii) *If $\min P_{E^1}(t) \leq x_p < x_q \leq \max P_{E^2}(t)$, and $u_q(t) - u_p(t)$ has the opposite sign of $V_{E^2}(t) - V_{E^1}(t)$, then*

$$|u_q(t) - u_p(t)| \leq TTV_{E^1 E^2}(t, t')/2 + C.$$

Proof. Without loss of generality, we may assume that $E^1_{t,t'}$, and $E^2_{t,t'}$ are a trough and a ridge, respectively. It follows that $u^M = V_{E^2}(t')$ and $u^N = V_{E^1}(t')$. Denote $u^M(t) := V_{E^2}(t)$ and $u^N(t) := V_{E^1}(t)$. Using Corollary 4.7, we have

$$TTV_{E^1 E^2}(t, t') \geq TV_{E^1 E^2}(t) - TV_{E^1 E^2}(t') = TV_{E^1 E^2}(t) - (u^M - u^N)$$
$$\geq u^M(t) - u_j(t) - u^M + u^N \geq -C'(t' - t) - u_j(t) + u^N,$$

and

$$TTV_{E^1 E^2}(t, t') \geq TV_{E^1 E^2}(t) - TV_{E^1 E^2}(t') = TV_{E^1 E^2}(t) - (u^M - u^N)$$
$$\geq u_j(t) - u^N(t) - u^M + u^N \geq u_j(t) - C'(t' - t) - u^M.$$

Hence, (i) holds. Next, we have

$$TTV_{E^1 E^2}(t, t') \geq TV_{E^1 E^2}(t) - TV_{E^1 E^2}(t') = TV_{E^1 E^2}(t) - (u^M - u^N)$$
$$\geq |u^M(t) - u_q(t)| + |u_q(t) - u_p(t)| + |u_p(t) - u^N(t)| - u^M + u^N$$
$$= 2(u_p(t) - u_q(t)) - 2C'(t' - t).$$

Thus, (ii) holds, and the lemma is proved. □

REMARK. Since $TV_u(t) = \sum_{j=-\infty}^{\infty} |u_{j+1}(t) - u_j(t)| \leq C_0$ implies that

$$\lim_{j \to \pm\infty} u_j(t) = C_{\pm}(t)$$

exists for some $C_{\pm}(t)$. We have $\frac{dC_{\pm}(t)}{dt} = q(C_{\pm}(t))$, by the consistency of the numerical flux g. Therefore, $-C'(t' - t) \leq C_{\pm}(t) - C_{\pm}(t') \leq C'(t' - t)$. It is easy to see that Lemma 4.12 also holds when $u(x, t')$ is monotone between $-\infty$ and $\max P_{E^2}(t')$, or when $u(x, t')$ is monotone between $\min P_{E^1}(t')$ and ∞. In the former case, $\min P_{E^1}(t)$ is replaced by $-\infty$ and $V_{E^1}(t)$ is replaced by $C_-(t)$; in the latter case, $\max P_{E^2}(t)$ is replaced by ∞ and $V_{E^2}(t)$ is replaced by $C_+(t)$. The proof is similar to that in the standard case and is omitted.

For the non-homogeneous scheme (1.3)-(1.5), Lemma 6.15 of [22] no longer holds, which plays active role in the proof of Theorem 3.12. Thus, we need a enhanced

version of Lemma 4.12 which functions the same as the Lemma 6.15 of [22]. Recall the function $\phi_u(t)$ defined by the relation (2.6) is a decreasing function of t. Hence, we have $T\phi_u(t, t') = \phi_u(t) - \phi_u(t')$. Our next goal is to replace $TTV_{E^1 E^2}(t, t')$ in Lemma 4.12 by $T\phi_u(t, t')$. For this purpose, we shall first establish the relationship between $T\phi_u(t, t')$ and $TTV_{E^1 E^2}(t, t')$. During the argument, C stands for a generic constant and $E^1_{t,t'}$, $E^2_{t,t'}$ are two arbitrary extreme paths. It is easy to see that $V_E(t)$ is Lipschitz continuous. Hence it is differentiable almost everywhere.

LEMMA 4.13. *For almost all $s \in [t, t']$, there is a $x_i = x_i(s) \in P_E(s)$, such that $\frac{dV_E(s)}{ds} = \frac{du_i(s)}{ds}$.*

Proof. For any $s \in [t, t']$, where $V_E(s)$ is differentiable, we have $P_E(s') \subseteq P_E(s)$ for s' sufficiently close to s. Hence, there is a $x_i = x_i(s) \in P_E(s)$ and a sequence $\{\tau_n\} \subseteq \mathbb{R} \setminus \{0\}$, $\tau_n \to 0$ such that $V_E(s + \tau_n) = u_i(s + \tau_n)$. Thus

$$\frac{dV_E(s)}{ds} = \lim_{\tau_n \to 0} \frac{V_E(s + \tau_n) - V_E(s)}{\tau_n} = \lim_{\tau_n \to 0} \frac{u_i(s + \tau_n) - u_i(s)}{\tau_n} = \frac{du_i(s)}{ds}.$$

□

LEMMA 4.14. *Let $M_{t,t'}$ be any ridge and $P_M(\tau) = \{x_p(\tau), \cdots, x_q(\tau)\}$. Then, with Assumption 2.4, for almost all $\tau \in [t, t']$, we have*

(4.9)
$$q(V_M(\tau)) \geq \frac{dV_M(\tau)}{d\tau}$$

$$\geq \frac{1}{2h} \sum_{j=p(\tau)}^{q(\tau)} (s_{j+\frac{1}{2}}(\tau) - s_{j-\frac{1}{2}}(\tau))(g_{j+\frac{1}{2}}(\tau) - g_{j-\frac{1}{2}}(\tau)) + q(V_M(\tau)).$$

Proof. By Assumption 2.4, the first inequality always holds. To see the second one. Let $i = i(\tau)$ be as in Lemma 4.13. Clearly, we have

$$\frac{dV_M(\tau)}{d\tau} = \frac{du_i(\tau)}{d\tau} = \max_{j \in P_M(\tau)} \frac{du_j(\tau)}{d\tau}$$

$$= \max_{j \in P_M(s)} [-\frac{1}{h}(g_{j+\frac{1}{2}}(\tau) - g_{j-\frac{1}{2}}(\tau)) + q(V_M(\tau))].$$

If $p(\tau) = q(\tau)$, then $j = i = p = q$ and $s_{j+\frac{1}{2}} - s_{j-\frac{1}{2}} = -2$. Hence

$$\frac{dV_M(\tau)}{d\tau} = \frac{1}{2h}(s_{j+\frac{1}{2}}(\tau) - s_{j-\frac{1}{2}}(\tau))(g_{j+\frac{1}{2}}(\tau) - g_{j-\frac{1}{2}}(\tau)) + q(V_M(\tau)).$$

Thus (4.9) holds.

If $p(\tau) \leq q(\tau) - 1$, then $s_{p+\frac{1}{2}} - s_{p-\frac{1}{2}} = s_{q+\frac{1}{2}} - s_{q-\frac{1}{2}} = -1$. Hence

$$\frac{dV_M(\tau)}{d\tau} = \max_{j \in P_M(\tau)} [-\frac{1}{h}(g_{j+\frac{1}{2}}(\tau) - g_{j-\frac{1}{2}}(\tau)) + q(V_M(\tau))]$$

$$\geq \frac{1}{2}[-\frac{1}{h}(g_{p+\frac{1}{2}} - g_{p-\frac{1}{2}}) - \frac{1}{h}(g_{q+\frac{1}{2}} - g_{q-\frac{1}{2}})] + q(V_M(\tau))$$

$$= \frac{1}{2h} \sum_{j=p(\tau)}^{q(\tau)} (s_{j+\frac{1}{2}}(\tau) - s_{j-\frac{1}{2}}(\tau))(g_{j+\frac{1}{2}}(\tau) - g_{j-\frac{1}{2}}(\tau)) + q(V_M(\tau)).$$

Thus (4.9) holds as well. We finished the proof of Lemma 4.14. □

The following lemma is the result of a simple observation.

LEMMA 4.15. *Without loss of generality. Let the extrema at $\tau \in [t, t']$ and between $E^1_{t,t'}$ and $E^2_{t,t'}$ are $N_1, M_1, \cdots, N_l, M_l$. We have*

$$TV_{E^1 E^2}(\tau) = V_{M_l}(\tau) - V_{N_1}(\tau) + 2 \sum_{m=1}^{l-1} (V_{M_m}(\tau) - V_{N_{m+1}}(\tau)).$$

Now, by Lemmas 4.13, 4.14 and 4.15, for almost all $\tau \in [t, t']$ we have

$$|\frac{dTV_{E^1 E^2}(\tau)}{d\tau}| \leq -\frac{1}{h} \sum_{j=-\infty}^{\infty} (s_{j+1/2} - s_{j-1/2})(g_{j+1/2} - g_{j-1/2}) + C.$$

Hence, we have

(4.10) $$TTV_{E^1 E^2}(t, t') = \int_t^{t'} |\frac{dTV_{E^1 E^2}(\tau)}{d\tau}| \, d\tau \leq T\phi_u(t, t') + C.$$

Based on the relation (4.10), we can restate Lemma 4.12 in the following important form.

LEMMA 4.16. *With the conditions and notations of Lemma 4.12,*
(i) *If $\min P_{E^1}(t) \leq x_j \leq \max P_{E^2}(t)$, then*

$$u^N - T\phi_u(t, t') - C \leq u_j(t) \leq u^M + T\phi_u(t, t') + C.$$

(ii) *If $\min P_{E^1}(t) \leq x_p < x_q \leq \max P_{E^2}(t)$, and $u_q(t) - u_p(t)$ has the opposite sign of $V_{E^2}(t) - V_{E^1}(t)$, then*

$$|u_q(t) - u_p(t)| \leq T\phi_u(t, t')/2 + C,$$

where C is a positive constant.

Lemma 4.16 and the relation (4.7), imply immediately the following result.

LEMMA 4.17. *Suppose that, in addition to the conditions and the notations of Lemma 4.12, $x_{I(t)}$ and $x_{J(t)}$ are ε-E paths to $E^1_{t,t'}$ and $E^2_{t,t'}$, respectively, such that $x_{I(t')} \in P_{E^1}(t')$ and $x_{J(t')} \in P_{E^2}(t')$.*
(i) *If $x_{I(t)} \leq x_j \leq x_{J(t)}$, then*

$$u^N - (C + \varepsilon) - T\phi_u(t, t') \leq u_j(t) \leq u^M + (C + \varepsilon) + T\phi_u(t, t').$$

(ii) *If $x_{I(t)} \leq x_p < x_q \leq x_{J(t)}$, and $u_q(t) - u_p(t)$ has the opposite sign of $V_{E^2}(t) - V_{E^1}(t)$, then*

$$|u_q(t) - u_p(t)| \leq T\phi_u(t, t')/2 + (C + 2\varepsilon).$$

REMARK.   Lemmas 4.16 and 4.17 also hold when $u(x, t')$ is monotone between $-\infty$ and $\max P_{E^2}(t')$, or when $u(x, t')$ is monotone between

$\min P_{E^1}(t')$ and ∞. This is because in both cases Lemma 4.12 and the relation (4.10) still hold.

The following result connects the notions of ε-E paths of this section, the ε-paths in the sense of Definition 3.4 and the TV-stable sequences of a numerical traveling discontinuity introduced in §3.

LEMMA 4.18. *Let* $\{u^k\}_{k=1}^\infty$ *be a TV-stable sequence of a numerical traveling discontinuity generated by a* $(\mathbf{h}, \boldsymbol{\varepsilon})$*-scaled form of the scheme* (1.3)-(1.5) *that satisfies Assumption 2.4 for each positive integer k. If* $x_E^{\varepsilon_k}(t)$ *is an* ε_k*-E path of* u^k*, then it is also an* ε_k*-path of the first type with respect to* u^k*.*

Thus, the extension of the extremum tracking theory [22] has been completed.

REFERENCES

[1] M. CRANDALL AND A. MAJDA, *Monotone difference approximations for scalar conservation laws,* Math. of Comp., 34 (1980), pp. 1–21.

[2] A. CHALABI, *Stable upwind schemes for hyperbolic conservation laws with source terms,* IMA J. of Numer. Anal., 12 (1992), pp. 217–241.

[3] A. CHALABI, *On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms,* Math. of Comp., 66:218 (1997), pp. 527–545.

[4] A. HARTEN, J. M. HYMAN AND P. D. LAX, *On finite-difference approximations and entropy conditions for shocks,* Comm. Pure and Appl. Math., 29 (1976), pp. 297–322.

[5] M. W. HIRSCH AND S. SMALE, *Differential equations, dynamical systems, and linear algebra,* Academic Press, 1974.

[6] S. JIN, *Runge-Kutta method for hyperbolic conservation laws with stiff relaxation terms,* J. Comput. Phys., 122 (1995), pp. 51–67.

[7] N. N. KUZNETSOV, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation,* USSR Comp. Math. and Math. Phys., 16 (1976), pp. 105–109.

[8] P. LAX AND B. WENDROFF, *Systems of conservation laws,* Comm. Pure Appl. Math., 13 (1960), pp. 217–237.

[9] P. LIONS AND P. SOUGANIDIS, *Convergence of MUSCL type methods for scalar conservation laws,* C. R. Acad. Sci. Paris., 311, Série I (1990), pp. 259–264.

[10] J. LEVEQUE AND H. C. YEE, *A study of numerical methods for hyperbolic conservation laws with stiff source terms,* J. Comput. Phys., 86 (1990), pp.187–210.

[11] J. NAN, PhD Dissertation, Kansas State University, (2000)

[12] S. OSHER, *Riemann solvers, the entropy condition and difference approximations,* SIAM J. Numer. Anal., 21 (1984), pp. 217–235.

[13] S. OSHER, *On convergence of generalized MUSCL schemes,* SIAM J. Numer. Anal., 22 (1985), pp. 947–961.

[14] S. OSHER AND S. CHAKRAVARTHY, *High resolution schemes and entropy condition,* SIAM J. Numer. Anal., 21 (1984), pp. 955–984.

[15] S. OSHER AND E. TADMOR, *On the convergence of difference schemes to scalar conservation laws,* Math. Comp., 50 (1988), pp. 19–51.

[16] R. SANDERS, *On convergence of monotone finite difference schemes for variable spatial differencing,* Math. Comp., 40 (1983), pp. 19–36.

[17] H. J. SCHROLL AND R. WINTHER, *Finite-difference schemes for scalar conservation laws with source terms,* IMA J. of Numer. Anal., 16 (1996), pp. 201–215.

[18] P. K. SWEBY, *High resolution schemes using flux limiters for hyperbolic conservation laws,* SIAM J. Numer. Anal., 21 (1984), pp. 995–1011.

[19] E. TADMOR, *Convenient total variation diminishing conditions for nonlinear difference schemes,* SIAM J. Numer. Anal., 25 (1988), pp. 1002–1014.

[20] T. TANG, *Convergence analysis for operator-splitting method applied to conservation laws with stiff source terms,* SIAM J. Numer. Anal., 35:5 (1998), pp. 1939–1968.

[21] A. VOL'PERT, *The space BV and quasi-linear equations,* Math. USSR Sb, 73 (1967), pp. 255–302.

[22] H.Yang, *On Wavewise Entropy Inequalities for High-Resolution Schemes I: The Semi-discrete case*, Math. Comp., 65 (1996), pp. 45–67.

[23] H. Yang, *On Wavewise Entropy Inequalities for High Resolution Schemes II: Fully discrete MUSCL Schemes with Exact Evolution In Small time*, SIAM. J. Numer. Anal., 36:1 (1999), pp. 1–31.