TVD SCHEMES FOR TWO-DIMENSIONAL SCALAR CONSERVATION LAWS

JOHN D. TOWERS

Abstract. A finite difference scheme for a two-dimensional scalar conservation law cannot have conservation form, a fixed and finite stencil, formal second order accuracy, and be total variation decreasing (TVD) - this is a well known theorem of Goodman and Leveque. This paper investigates the possibility of relaxing the restriction on the stencil in order to construct a scheme that is conservative, has second order accuracy, and is TVD. A nonlocal limiter algorithm is introduced, which operates on either flux corrections (for flux-limiter schemes) or slopes (for slope limiter schemes) in order to enforce the TVD property. The main results of the paper are compatibility of the non-local limiter algorithm with formal second order accuracy and convergence of the resulting approximations (modulo extraction of a subsequence) to a weak solution of the conservation law.

1. Introduction

This paper investigates shock-capturing finite difference algorithms for computing approximate solutions to the two-dimensional Cauchy problem

$$u_t + f(u)_x + g(u)_y = 0, \quad u(x,y,0) = u_0(x,y) \tag{1.1}$$

in $\mathbf{R}^2 \times [0,\infty)$. The initial data is assumed to be locally integrable, of bounded variation, and constant for sufficiently large x and y; there are constants B and u_∞ such that $u_0(x,y) = u_\infty$ for $|(x,y)| \geq B$.

Even when the initial data is smooth, solutions of 1.1 develop discontinuities, and so weak solutions are sought. A weak solution to the Cauchy problem 1.1 is a bounded measurable function $u : \mathbf{R}^2 \times [0,\infty) \to \mathbf{R}^1$ satisfying

$$\int_0^\infty \int_{\mathbf{R}^2} \Big(u\psi_t + f(u)\psi_x + g(u)\psi_y \Big) dxdydt + \int_{\mathbf{R}^2} u_0(x,y)\psi(x,y,0)dxdy = 0 \tag{1.2}$$

for every smooth test function $\psi(x,y,t)$ with compact support in $\mathbf{R}^2 \times [0,\infty)$.

The concept of total variation (TV) plays a central role in both the theory of scalar conservation laws and in computing approximate solutions. For a locally integrable function $u : \mathbf{R}^1 \to \mathbf{R}^1$, its total variation $TV(u)$ is

$$TV(u) = TV^x(u) = \sup_{|h|>0} \int_{\mathbf{R}^1} \frac{|u(x+h) - u(x)|}{|h|} dx,$$

1991 *Mathematics Subject Classification.* Primary 35L65, 65M06, 65M12.
Key words and phrases. conservation laws, finite difference approximations, TVD, SOR.
2dtvd01.tex Jan 04 , 2001 jtowers@cts.com.

and for $u : \mathbf{R}^2 \to \mathbf{R}^1$

(1.3)

$$\begin{aligned} TV(u) &= TV^x(u) + TV^y(u) \\ &= \sup_{|h|>0} \int_{\mathbf{R}^2} \frac{|u(x+h,y) - u(x,y)|}{|h|} dxdy + \sup_{|h|>0} \int_{\mathbf{R}^2} \frac{|u(x,y+h) - u(x,y)|}{|h|} dxdy. \end{aligned}$$

A numerical scheme which generates approximations $U^n = u^\Delta(\cdot,\cdot,t^n)$ that are nonincreasing in total variation from one time level to the next, $TV(U^{n+1}) \le TV(U^n)$, is described as total variation decreasing (TVD). Entropy solutions of the Cauchy problem 1.1 have decreasing total variation, and so TVD schemes produce approximations that mimic an important property of the desired solution. From a more practical point of view, the TVD property rules out spurious overshoots and oscillations that often appear near steep gradients with the traditional higher order schemes such as the Lax-Wendroff algorithm. Finally, the TVD property, when accompanied by a uniform bound on the approximations, and an easily obtained time-continuity estimate, guarantee that the sequence of approximations are convergent (modulo extraction of a subsequence) as the mesh size shrinks. If the scheme is in conservation form, the Lax-Wendroff [20] theorem guarantees that the limit of the subsequence is a weak solution. If in addition, the numerical approximations satisfy a set of discrete entropy inequalities, then the whole sequence converges to the physically correct entropy solution.

Monotone schemes, such as the Lax-Friedrichs scheme, Godunov's scheme, and the Engquist-Osher scheme are TVD, and additionally satisfy a large set of discrete entropy inequalities. This makes it possible to prove that monotone schemes converge to the physically unique entropy solution of the Cauchy problem 1.1 [5], [6], [15], [19]. This convergence result holds quite generally, and includes multidimensional schemes with variable (both spatial and temporal) meshes [24], [27]. Monotone schemes are well-behaved from the point of view that they do not generate spurious oscillations near steep gradients. The upwind algorithms such as the Engquist-Osher scheme and Godunov's scheme resolve shocks without smearing, usually spreading the shock over only one or two mesh points. Nevertheless, monotone schemes are at best first order accurate [15] and thus of limited usefulness for multidimensional problems.

Harten discovered a simple criterion (see lemma 2.1) for one-dimensional schemes for scalar conservation laws to be TVD, and showed that it was possible to construct schemes of higher than first order accuracy (at least away from extrema) that were also TVD. Schemes that have formal second order accuracy, except possibly at extrema are said to have second order resolution (SOR). Schemes which are both TVD and SOR are referred to as TVD-SOR. Harten's TVD criterion initiated an enormous amount of activity constructing and analyzing TVD schemes during the eighties. The resulting algorithms, and their extensions, were applied successfully not only to one-dimensional scalar problems, but also to systems via field by field decomposition, and also to multidimensional problems using dimension-by-dimension TVD schemes. TVD schemes degenerate to first order accuracy at nonsonic extrema [25]. In order to overcome this problem, total variation bounded (TVB) schemes were developed by Shu [28], and essentially nonoscillatory (ENO) schemes were developed by Harten, Osher, Engquist, and Chakravarthy [14], [16]. To reduce the complexity required in solving Riemann problems, Nessyahu and Tadmor [26] developed high order central schemes which are TVD-SOR and converge to the unique entropy solution for the case of a convex flux. Osher and Tadmor presented a fairly complete theory of explicit TVD-SOR algorithms, including conditions guaranteeing convergence to the physically correct entropy condition, assuming a concave or convex flux [25].

A weaker condition than TVD which also can be used to guarantee convergence when the flux is convex, at least in the one-dimensional setting, is the so-called one-sided Lipschitz condition (OSLC). Brenier and Osher [2] and Goodman and Leveque [11] have constructed second order schemes satisfying the OSLC condition. It is not known whether this approach could be used for multidimensional schemes. If so, a result of Hoff [17] seems to imply that it would have to be restricted to conservation laws having an isotropic flux, meaning that $(f(u), g(u)) = f(u)\mathbf{a}$, where f is convex, and $\mathbf{a}$ is a unit vector. The isotropic situation is of some interest since the equations of inviscid gas dynamics are isotropic.

Goodman and Leveque showed that for two-dimensional conservation laws the TVD property is incompatible with second (or higher) order accuracy. More precisely, they showed that it is not possible for a 2D scheme to have the following four properties:

- conservation form,
- formal second order accuracy in smooth, noncritical regions,
- TVD,
- a fixed local stencil.

The present paper considers the consequences of relaxing the fourth condition. Specifically, the stencil is allowed to expand to the width of the computational domain, a situation that will be described herein as nonlocal. Nonlocal stencils are normally associated with implicit schemes, where the stencil necessarily becomes nonlocal because of the systems of equations that must be solved. In the present context, the stencil becomes nonlocal due to the regularization process required to keep the scheme simultaneously second order accurate and TVD.

Coquel and LeFloch [3], [4] made an important contribution to the theory of finite difference algorithms for conservation laws by showing that it is possible to construct higher than first order multidimensional schemes that are convergent without requiring a variation bound. They used entropy estimates to obtain a weak estimate of the discrete space derivatives. These estimates are weaker than a bound on the total variation but sufficient to apply the theory of measure valued solutions of DiPerna [8].

The two-dimensional schemes described in this paper will be constructed within the fractional step framework. (See eg., reference [6].) Fractional step methods are widely used not only for difference schemes, but also with front tracking methods [18], [22]. Much of their appeal comes from the fact that they provide a natural way to build a multidimensional algorithm from a simpler one-dimensional algorithm. This approach also simplifies the analysis in the following sections. The spatial domain $\mathbf{R}^2$ is discretized into uniform cells $I_j \times I_k = [x_j - \Delta x/2, x_j + \Delta x/2) \times [y_k - \Delta y/2, y_k + \Delta y/2)$ centered at the points $(x_j, y_k) = (j\Delta x, k\Delta y)$ for $j, k \in \mathbf{Z}$. Similarly, the time domain $\mathbf{R}^+$ is divided into time strips $I^n = [t^n, t^{n+1})$, where $t^n = n\Delta t$ for $n \in \mathbf{Z}^+$. Let $\chi^n_{j,k}$ be the characteristic function for the box $I_j \times I_k \times I^n$, and let $\chi_j(x)$, $\chi_k(y)$, and $\chi^n(t)$ be the characteristic functions for the intervals I_j, I_k and I^n. For each mesh size $\Delta = \Delta t$, the finite difference scheme generates a piecewise constant approximation

$$u^{\Delta}(x,y,t) = \sum_{n\geq 0} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \chi_{j,k}^{n}(x,t) U_{j,k}^{n}. \tag{1.4}$$

Cell averages are used to discretize the initial data:

$$u^{\Delta}(x,y,0) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \chi_{j,k}(x) U_{j,k}^{0}, \qquad U_{j,k}^{0} = \frac{1}{\Delta x \Delta y} \int_{I_j} \int_{I_k} u_0(x,y) dx dy,$$

where $\chi_{j,k}(x,y)$ is the characteristic function for the rectangle $I_j \times I_k$.

Let $\lambda^x = \Delta t/\Delta x$, $\lambda^y = \Delta t/\Delta y$. The fractional step algorithms considered here will be based on one-dimensional operators F consistent with $u_t + f(u)_x = 0$:

$$F(U;\lambda^x)_{j,k} = U_{j,k} - \lambda^x \Delta_-^x \bar{f}(U)_{j,k}, \tag{1.5}$$

and G consistent with $u_t + g(u)_y = 0$:

$$G(U;\lambda^y)_{j,k} = U_{j,k} - \lambda^y \Delta_-^y \bar{g}(U)_{j,k}. \tag{1.6}$$

Here Δ_-^x and Δ_-^y are backward difference operators which operate on the j and k indices respectively; Δ_+^x and Δ_+^y will be used to denote the corresponding forward differences. The numerical fluxes $\bar{f}$ and $\bar{g}$ are consistent with f and g.

The Godunov method of fractional steps is the following two step procedure:

$$\begin{aligned} U^{n+\frac{1}{2}} &= F(U^n;\lambda^x), \\ U^{n+1} &= G(U^{n+\frac{1}{2}};\lambda^y). \end{aligned} \tag{1.7}$$

The Strang method of fractional steps is

$$\begin{aligned} U^{n+\frac{1}{4}} &= F(U^n;\lambda^x/2), \\ U^{n+\frac{3}{4}} &= G(U^{n+\frac{1}{4}};\lambda^y), \\ U^{n+1} &= F(U^{n+\frac{3}{4}};\lambda^x/2). \end{aligned} \tag{1.8}$$

The order of the x and y operators can be obviously be interchanged with either method. In the case where the operators F and G are monotone, the resulting fractional step schemes are also monotone, and so all of the standard monotone theory as in eg, [5], [6] applies.

If $F(\cdot;\lambda^x)_{\cdot,k}$ $(G(\cdot;\lambda^y)_{j,\cdot})$ is a second order accurate algorithm for the 1D conservation law $u_t + f(u)_x = 0$ $(v_t + g(v)_x = 0)$, then Strang's fractional step method is a second order scheme for 1.1, while Godunov's fractional step method results in second order spatial accuracy and first order accuracy in time. It is clear that either of the two fractional step methods will be TVD in the fully 2D sense as long as each of the fractional time steps is.

By way of notation, lower case variables such as u, v, w will generally denote functions defined on all of $\mathbf{R}^2 \times [0,\infty)$, $\mathbf{R}^1 \times [0,\infty)$, $\mathbf{R}^2$, or $\mathbf{R}^1$, while uppercase will be reserved for step functions defined on the discretized versions of these spaces. Given a step function U on the discretized version of $\mathbf{R}^2 \times [0,\infty)$ with the mesh refinement parameter Δ, a function u^{Δ} defined on all of $\mathbf{R}^2 \times [0,\infty)$ is constructed in the obvious way using equation 1.4. Passage from mesh functions on the discretized versions of the other spaces $\mathbf{R}^1 \times [0,\infty)$, $\mathbf{R}^2$, or $\mathbf{R}^1$ is accomplished in an analogous way. With this convention, the various norms and functionals (L^1, L^∞, TV) apply to mesh functions, eg., for U defined on a discretized $\mathbf{R}^1$,

$$||U||_1 = ||u^\Delta||_1 = \Delta x \sum_j |U_j|.$$

For a step function on $\mathbf{R}^1$, $u^\Delta(x) = \sum_j \chi_j(x) U_j$, the total variation is just the sum of the jumps:

$$TV(U) = TV(u^\Delta) = \sum_j |U_{j+1} - U_j|.$$

For a step function on $\mathbf{R}^2$, $u^\Delta(x,y) = \sum_j \sum_k \chi_{j,k}(x,k) U_{j,k}$, the total variation is

$$\begin{aligned} TV(U) &= TV(u^\Delta) = TV^x(U) + TV^y(U) \\ &= \Delta y \sum_k \sum_j |U_{j+1,k} - U_{j,k}| + \Delta x \sum_j \sum_k |U_{j,k+1} - U_{j,k}|. \end{aligned} \tag{1.9}$$

The remainder of the paper is organized as follows. Section 2 derives a system of inequalities that are sufficient for a flux-limiter one-dimensional sweep of a fractional step method to be TVD, while section 3 derives the analogous system of inequalities for slope limiter schemes. The basic form of the system of inequalities is the same in both cases. Section 4 describes the nonlocal limiter algorithm, which is a method of solving these inequalities. This section contains a proof that the algorithm actually does produce a solution to the system of inequalities, and that the solution is optimal from one point of view. Section 5 investigates the compatibility of the limiter with second order accuracy, the main result being that away from zeros of u_x and u_y, the nonlocal limiter leaves the flux corrections or slopes unchanged once the mesh size becomes small enough. Section 6 proves that, modulo extraction of a subsequence, the scheme generates approximations which converge to a weak solution to the conservation law as the mesh size shrinks. Section 7 contains some concluding remarks.

2. TVD conditions for flux-limiter schemes

This section derives conditions guaranteeing that each of the one-dimensional operators of a fractional step method is TVD in the fully 2-D sense. Each of the one-dimensional operators will correspond to a flux-limited high resolution scheme for a one-dimensional conservation law. The SOR property is achieved by adding second order corrections to a monotone flux. For the sake of concreteness, the Engquist-Osher (EO) [9] flux will be used for the monotone flux. The one-dimensional limiter is based in ideas found in [29]. The EO scheme for a one-dimensional scalar conservation law $u_t + f(u)_x = 0$ is

$$V_j = U_j - \lambda^x \Delta_-^x f^{eo}_{j+\frac{1}{2}}. \tag{2.1}$$

Here $U_j = U_j^n$, $V_j = U_j^{n+1}$, i.e., the time superscipts have been suppressed to simplify the notation; this convention will be observed wherever no confusion is possible. The EO numerical flux is defined by

$$f^{eo}_{j+\frac{1}{2}} = \frac{1}{2}(f(U_j) + f(U_{j+1})) - \frac{1}{2}\int_{U_j}^{U_{j+1}} |f'(w)| dw. \tag{2.2}$$

The resulting three-point scheme can also be written in incremental form:

$$V_j = U_j + C^+_{j+\frac{1}{2}}\Delta_+ U_j - C^-_{j-\frac{1}{2}}\Delta_- U_j, \tag{2.3}$$

where

$$C^-_{j+\frac{1}{2}} = \frac{\lambda^x}{\Delta_+ U_j}\int_{U_j}^{U_{j+1}} \chi_+(w) f'(w)dw, \quad C^+_{j+\frac{1}{2}} = \frac{-\lambda^x}{\Delta_+ U_j}\int_{U_j}^{U_{j+1}} \chi_-(w) f'(w)dw. \tag{2.4}$$

Here $\chi_-(w)$ $(\chi_-(w))$ is the characteristic function for the set where $f'(w) < 0$ $(f'(w) > 0)$. The incremental coefficients $C^-_{j+\frac{1}{2}}$ and $C^+_{j-\frac{1}{2}}$ are nonnegative, and it is not hard to check that with the CFL condition

$$\lambda^x |f'(w)| \leq 1, \quad \forall w \in [\min_j(U_j), \max_j(U_j)], \tag{2.5}$$

the following inequality holds:

$$C^-_{j+\frac{1}{2}} + C^+_{j+\frac{1}{2}} \leq 1. \tag{2.6}$$

By the following of lemma of Harten [12], the one-dimensional scheme 2.1, is TVD.

Lemma 2.1 (Harten)**.** *Suppose that a one-dimensional scalar scheme can be written in the following incremental form:*

$$U^{n+1}_j = U^n_j + C^+_{j+\frac{1}{2}}\Delta_+ U^n_j - C^-_{j-\frac{1}{2}}\Delta_- U^n_j, \tag{2.7}$$

where $C^+_{j+\frac{1}{2}} \geq 0$, $C^-_{j-\frac{1}{2}} \geq 0$, and $C^+_{j+\frac{1}{2}} + C^-_{j+\frac{1}{2}} \leq 1$. Then the scheme is TVD, $TV(U^{n+1}) \leq TV(U^n)$.

Harten, Hyman, and Lax [15] showed that monotone schemes, including the EO algorithm just described, are only first order accurate. Reference [15] gives an explicit formula for the truncation error $TE(\Delta t)$ of a monotone scheme. Applying that formula to the EO scheme gives

$$TE(\Delta t) = -(\Delta t)^2(\beta(u,\lambda^x)u_x)_x + O(\Delta t^3), \quad \beta(u,\lambda^x) = \frac{1}{2\lambda^x}|f'(u)|(1-\lambda^x|f'(u)|). \tag{2.8}$$

A second order version of EO can be constructed by following the procedure in Sweby's paper [29]. Let

$$R^+_j = -\frac{1}{2}C^+_{j-\frac{1}{2}}(1 - C^+_{j-\frac{1}{2}})\Delta_- U_j, \quad R^-_j = \frac{1}{2}C^-_{j+\frac{1}{2}}(1 - C^-_{j+\frac{1}{2}})\Delta_+ U_j, \tag{2.9}$$

and define

$$\hat{R}^+_j = M(R^+_j, R^+_{j+1}), \quad \hat{R}^-_j = M(R^-_j, R^-_{j-1}), \tag{2.10}$$

where M is the two-argument minmod function defined by

$$M(a,b) = \frac{1}{2}(\sigma(a) + \sigma(b))min(|a|,|b|).$$

Here $\sigma(a) = a/|a|$ for $a \neq 0$, and zero otherwise. A new four-point numerical flux is constructed using the correction terms just defined:

$$f^{sor}_{j+\frac{1}{2}} = f^{sor}(U_{j+2}, U_{j+1}, U_j, U_{j-1}) = f^{eo}_{j+\frac{1}{2}} - \frac{1}{\lambda^x}\hat{R}^+_{j+1} + \frac{1}{\lambda^x}\hat{R}^-_j. \tag{2.11}$$

This scheme can be written in incremental form as:

$$V_j = U_j + \tilde{C}^+_{j+\frac{1}{2}}\Delta_+ U_j - \tilde{C}^-_{j-\frac{1}{2}}\Delta_- U_j, \tag{2.12}$$

where

$$\tilde{C}^{+}_{j+\frac{1}{2}} = C^{+}_{j+\frac{1}{2}} + \frac{\Delta_{+}\hat{R}^{+}_{j}}{\Delta_{+}U_{j}}, \quad \tilde{C}^{-}_{j-\frac{1}{2}} = C^{-}_{j-\frac{1}{2}} + \frac{\Delta_{-}\hat{R}^{-}_{j}}{\Delta_{-}U_{j}}. \tag{2.13}$$

It is readily verified (see eg., [29]) that the modified incremental coefficients satisfy

$$\tilde{C}^{\pm}_{j+\frac{1}{2}} \in [0,1], \quad \tilde{C}^{+}_{j+\frac{1}{2}} + \tilde{C}^{-}_{j+\frac{1}{2}} \leq 1$$

with the reduced CFL condition

$$\lambda^{x}|f'(w)| \leq 2/3, \quad \forall w \in [\min_{j}(U_j), \max_{j}(U_j)], \tag{2.14}$$

and so, by lemma 2.1 again, the one-dimensional scheme

$$V_j = F^{sor}(U;\lambda^x)_j = U_j - \lambda^x \Delta^x_{-} f^{sor}_{j+\frac{1}{2}}$$

is TVD. The resulting 1-D algorithm has formal second order accuracy at any point where the solution is smooth and noncritical, i.e., the scheme is SOR. This follows by comparing the correction terms 2.9, 2.10 with the trunction error 2.8 of the first order version of the scheme.

By further reducing the CFL condition to

$$\lambda^{x}|f'(w)| \leq 1/3, \quad \forall w \in [\min_{j}(U_j), \max_{j}(U_j)], \tag{2.15}$$

the incremental coefficients also satisfy

$$\tilde{C}^{+}_{j+\frac{1}{2}} + \tilde{C}^{-}_{j-\frac{1}{2}} \leq 1$$

which guarantees that V_j lies in the convex hull of $\{U_{j-1}, U_j, U_{j+1}\}$, and provides an L^∞ bound on the approximations.

Returning to the two-dimensional setting, the scheme just constructed is viewed as a single x-step of a fractional step algorithm. Let

$$V_{j,k} = U_{j,k} - \alpha^x \lambda^x \Delta^x_{-} f^{sor}_{j+\frac{1}{2},k} = F^{sor}(U; \alpha^x \lambda^x)_{j,k}, \tag{2.16}$$

where α^x depends on the fractional step method; $\alpha^x = 1$ for the Godunov splitting, and $\alpha^x = 1/2$ for the Strang splitting. The numerical flux is given by

$$f^{sor}_{j+\frac{1}{2},k} = f^{eo}_{j+\frac{1}{2},k} + \hat{d}_{j+\frac{1}{2},k}, \quad f^{eo}_{j+\frac{1}{2},k} = f^{eo}(U_{j+1,k}, U_{j,k}). \tag{2.17}$$

The modified correction terms $\hat{d}_{j+\frac{1}{2},k}$ are yet to be specified. Let

$$d_{j+\frac{1}{2},k} = -\frac{1}{\alpha^x \lambda^x}\hat{R}^{+}_{j+1,k} + \frac{1}{\alpha^x \lambda^x}\hat{R}^{-}_{j,k}. \tag{2.18}$$

It will be assumed that $d_{j+\frac{1}{2},k}$ has been replaced by zero if $d_{j+\frac{1}{2},k}d_{j-\frac{1}{2},k} < 0$ or $d_{j+\frac{1}{2},k}d_{j+\frac{3}{2},k} < 0$. Away from sonic zeros of u_x, this leaves $d_{j+\frac{1}{2},k}$ unchanged, and so the SOR property is not affected by this modification. Doing this ensures that

$$d_{j+\frac{1}{2},k}d_{j-\frac{1}{2},k} \geq 0, \tag{2.19}$$

and simplifies the nonlocal limiter algorithm of section 4

The first additional condition to be placed on the modified correction term $\hat{d}_{j+\frac{1}{2},k}$ is

$$\hat{d}_{j+\frac{1}{2},k}/d_{j+\frac{1}{2},k} \in [0,1]. \tag{2.20}$$

Condition 2.20 guarantees that $TV^x(V) \leq TV^x(U)$. The objective now is to find additional conditions on $\hat{d}_{j+\frac{1}{2},k}$ so that also $TV^y(V) \leq TV^y(U)$. Applying the approach in [27] to the present situation, the y differences after an x-step can be written

(2.21)
$$\begin{aligned}\Delta_+^y V_{j,k} = \Big(1 - A_{j,k+\frac{1}{2}} - B_{j,k+\frac{1}{2}} - \frac{\alpha^x \lambda^x \Delta_+^x \hat{d}_{j-\frac{1}{2},k+1}}{\Delta_+^y U_{j,k}} + \frac{\alpha^x \lambda^x \Delta_+^x \hat{d}_{j-\frac{1}{2},k}}{\Delta_+^y U_{j,k}}\Big) \Delta_+^y U_{j,k} \\ + A_{j-1,k+\frac{1}{2}} \Delta_+^y U_{j-1,k} + B_{j-1,k+\frac{1}{2}} \Delta_+^y U_{j-1,k},\end{aligned}$$

where

$$A_{j,k+\frac{1}{2}} = \alpha^x \lambda^x \Big(\frac{f^{eo}(U_{j+1,k+1}, U_{j,k+1}) - f^{eo}(U_{j+1,k+1}, U_{j,k})}{U_{j,k+1} - U_{j,k}} \Big), \tag{2.22}$$

$$B_{j,k+\frac{1}{2}} = \alpha^x \lambda^x \Big(\frac{f^{eo}(U_{j,k}, U_{j-1,k}) - f^{eo}(U_{j,k+1}, U_{j-1,k})}{U_{j,k+1} - U_{j,k}} \Big). \tag{2.23}$$

The coefficients $A_{j,k+\frac{1}{2}}$ and $B_{j,k+\frac{1}{2}}$ are nonnegative, thanks to the monotonicity of the EO numerical flux. If the additional CFL-type condition

$$A_{j,k+\frac{1}{2}} + B_{j,k+\frac{1}{2}} \le 1 - \rho, \quad \rho \in (0,1), \tag{2.24}$$

is imposed, along with the requirement that

$$\Big| \frac{\alpha^x \lambda^x \Delta_+^x \hat{d}_{j-\frac{1}{2},k+1}}{\Delta_+^y U_{j,k}} \Big| \le \rho/2, \quad \Big| \frac{\alpha^x \lambda^x \Delta_+^x \hat{d}_{j-\frac{1}{2},k}}{\Delta_+^y U_{j,k}} \Big| \le \rho/2, \tag{2.25}$$

then taking absolute values and summing 2.21 over j gives

$$\sum_j |V_{j+1,k} - V_{j,k}| \le \sum_j |U_{j,k+1} - U_{j,k}| - \lambda^x \alpha^x \sum_j \sigma^y_{j,k+\frac{1}{2}} \Delta_+^y \Delta_+^x \hat{d}_{j-\frac{1}{2},k}. \tag{2.26}$$

Here $\sigma^y_{j,k+\frac{1}{2}} = \sigma(\Delta_+^y U_{j,k})$. Multiplying 2.26 by Δx and summing over k yields

$$\begin{aligned} TV^y(V) &\le TV^y(U) - \lambda^x \alpha^x \sum_k \Delta x \sum_j \sigma^y_{j,k+\frac{1}{2}} \Delta_+^y \Delta_+^x \hat{d}_{j-\frac{1}{2},k} \\ &= TV^y(U) + \lambda^x \alpha^x \sum_j \Delta x \sum_k \Delta_+^x \hat{d}_{j-\frac{1}{2},k} (\sigma^y_{j,k+\frac{1}{2}} - \sigma^y_{j,k-\frac{1}{2}}). \end{aligned} \tag{2.27}$$

Thus the additional condition

$$\sigma^y_{j,k+\frac{1}{2}} \ne \sigma^y_{j,k-\frac{1}{2}} \Rightarrow \Delta_+^x \hat{d}_{j-\frac{1}{2},k} = 0, \tag{2.28}$$

along with 2.20, 2.25 will give the desired variation bound $TV^y(V) \le TV^y(U)$.

It is clear that satisfaction of the following set of inequalities is sufficient for the conditions 2.20, 2.25, 2.28 to hold:

$$|\hat{d}_{j+\frac{1}{2},k} - \hat{d}_{j-\frac{1}{2},k}| \le \frac{\rho}{2\alpha^x \lambda^x} |M(\Delta_-^y U_{j,k}, \Delta_+^y U_{j,k})|, \tag{2.29}$$

$$0 \le \hat{d}_{j+\frac{1}{2},k} / d_{j+\frac{1}{2},k} \le 1. \tag{2.30}$$

The CFL-type condition 2.24 reduces to the following inequality for the EO flux:

$$\frac{\alpha^x \lambda^x}{\Delta_+^y U_{j,k}} \int_{U_{j,k}}^{U_{j,k+1}} |f'(w)| dw \le 1 - \rho. \tag{2.31}$$

This condition will hold with $\rho = 2/3$ assuming that the CFL condition $\alpha^x \lambda^x ||f'||_\infty \le 1/3$ is satisfied, which will be the case under the conditions of theorem 6.1.

Similarly, for a 1-dimensional y-step $W = G(U;\alpha^y\lambda^y)$ the corresponding set of inequalities to be satisfied by the modified flux corrections $\hat{e}_{j,k-\frac{1}{2}}$ is

$$|\hat{e}_{j,k+\frac{1}{2}} - \hat{e}_{j,k-\frac{1}{2}}| \leq \frac{\rho}{2\alpha^y\lambda^y}|M(\Delta^x_- U_{j,k}, \Delta^x_+ U_{j,k})| \tag{2.32}$$

$$0 \leq \hat{e}_{j,k+\frac{1}{2}}/e_{j,k+\frac{1}{2}} \leq 1. \tag{2.33}$$

In smooth noncritical regions, the left sides of 2.29 and 2.32 are $O(\Delta^2)$ while the right sides are $O(\Delta)$, and so it is plausible that there exist nontrivial solutions which allow for formal second accuracy. Sections 4 and 5 are devoted to showing that this is indeed the case. The results of this section are summarized in the following lemma.

Lemma 2.2. *If the x-correction terms $\hat{d}_{j+\frac{1}{2},k}$, satisfy 2.29, 2.30, and the y-correction terms $\hat{e}_{j,k+\frac{1}{2}}$ satisfy 2.32, 2.33, then the flux-limiter Strang fractional step method*

$$U^{n+1} = \hat{F}^{sor}(\cdot;\lambda^x/2) \circ \hat{G}^{sor}(\cdot;\lambda^y) \circ \hat{F}^{sor}(U^n;\lambda^x/2) \tag{2.34}$$

and the flux-limiter Godunov fractional step method

$$U^{n+1} = \hat{G}^{sor}(\cdot;\lambda^y) \circ \hat{F}^{sor}(U^n;\lambda^x) \tag{2.35}$$

are TVD in the fully two-dimensional sense. Here

$$\hat{F}^{sor}(U;\alpha^x\lambda^x)_{j,k} = U_{j,k} - \alpha^x\lambda^x\Delta^x_-(f^{eo}_{j+\frac{1}{2},k} + \hat{d}_{j+\frac{1}{2},k}), \tag{2.36}$$

$$\hat{G}^{sor}(U;\alpha^y\lambda^y)_{j,k} = U_{j,k} - \alpha^y\lambda^y\Delta^y_-(g^{eo}_{j,k+\frac{1}{2}} + \hat{e}_{j,k+\frac{1}{2}}). \tag{2.37}$$

3. TVD conditions for slope-limiter schemes

This section discusses conditions assuring that each of the one-dimensional operators of a fractional step method is TVD, assuming that they are slope-limiter high resolution schemes. Algorithms of this type were originally proposed by VanLeer [32], who referred to them as Monotone Upwind Schemes for Conservation Laws (MUSCL). See reference [23] for a proof of convergence to the entropy solution in the one-dimensional situation. The one-dimensional algorithms are of the form

$$V_j = U_j - \lambda^x\Delta^x_- f^{sor}_{j+\frac{1}{2}} \tag{3.1}$$

where

$$f^{sor}_{j+\frac{1}{2}} = \frac{1}{\Delta t}\int_0^{\Delta t} f(S^x(t)L^xU(x_{j+\frac{1}{2}}))dt + O(\Delta^2). \tag{3.2}$$

Here $S^x(t)$ denotes the exact solution operator of the one-dimensional conservation law $u_t + f(u)_x = 0$. The function L^xU is a piecewise linear reconstruction of the step function

$$u^\Delta(x) = \sum_j \chi_j(x)U_j.$$

The piecewise linear reconstruction has the form

$$L^xU(x) = L^xu^\Delta(x) = \sum_j \chi_j(x)\Big(U_j + \frac{x-x_j}{\Delta x}S_j\Big) \tag{3.3}$$

where the so-called slopes S_j are intended to approximate $\Delta_+ U_j$ and $\Delta_- U_j$, but are not yet fully specified. It is assumed that the one-dimensional scheme 3.1 can be written abstractly in the form

$$V = \mathcal{A} \circ E^x(\Delta t) \circ L^x U. \tag{3.4}$$

Here, $E^x(\Delta t)$ is an approximate evolution operator, $\mathcal{A}$ is the averaging operator,

$$\mathcal{A}w(x) = \sum_j \chi_j(x) \frac{1}{|I_j|} \int_{I_j} w(x)dx, \tag{3.5}$$

and L^x is a reconstruction operator taking piecewise constant functions into piecewise linear functions. The scheme 3.1, 3.4 has formal second order accuracy (see eg., [16]) if, for a smooth solution $w(x,t)$

$$W_j(t^n + \Delta t) = W_j(t^n) - \lambda \Delta_- f^{sor}_{j+\frac{1}{2}} + O(\Delta^3), \quad W_j(t) = \frac{1}{\Delta x} \int_{I_j} w(x,t)dx. \tag{3.6}$$

If

$$S_j = \Delta_+ W_j(t^n) + O(\Delta^2), \quad S_{j+1} = \Delta_+ W_j(t^n) + O(\Delta^2) \tag{3.7}$$

and if all of the $O(\Delta^2)$ quantities are smooth, then 3.6 holds, and the scheme is formally second order accurate.

The averaging operator $\mathcal{A}$ is TVD. Thus, assuming that the approximate evolution operator $E^x(\Delta t) = E^x$ is TVD, the one-dimensional scheme 3.4 will be TVD if

$$TV(L^x U) \le TV(U). \tag{3.8}$$

It is not hard to check that 3.8 is satisfied (see eg., Leveque's textbook [21]) with the choice

$$S_j = M(\Delta_- U_j, \Delta_+ U_j), \tag{3.9}$$

and that if $\hat{S}_j$ is another set of slopes satisfying $\hat{S}_j / S_j \in [0,1]$, then 3.8 continues to hold.

Returning to the two-dimensional setting, the single x-step

$$V_{j,k} = U_{j,k} - \alpha^x \lambda^x \Delta^x_- f^{sor}_{j+\frac{1}{2},k} = F^{sor}(U; \alpha^x \lambda^x)_{j,k} \tag{3.10}$$

with the modified x-slopes satisfying

$$\hat{S}_{j,k} / S_{j,k} \in [0,1], \quad S_{j,k} = M(\Delta^x_+ U_{j,k}, \Delta^x_- U_{j,k}) \tag{3.11}$$

will be TVD in the two-dimensional sense, $TV(V) \le TV(U)$, if $TV^y(L^x U) \le TV^y(U)$. Here the L^x refers to piecewise linear reconstruction in the x direction, i.e.,

$$L^x U(x,y) = L^x u^\Delta(x,y) = \sum_k \sum_j \chi_k(y) \chi_j(x) \Big(U_{j,k} + \frac{x - x_j}{\Delta x} \hat{S}_{j,k} \Big). \tag{3.12}$$

The y-component of $TV(L^xU)$ is computed as follows:

(3.13)
$$\begin{aligned} TV^y(L^xU) &= \sum_k \sum_j \int_{I_j} dx |L^xU(x, y_{k+1}) - L^xU(x, y_k)| \\ &= \sum_k \sum_j \int_{I_j} dx \Big| \Big(U_{j,k+1} + \frac{x - x_j}{\Delta x} \hat{S}_{j,k+1}\Big) - \Big(U_{j,k} + \frac{x - x_j}{\Delta x} \hat{S}_{j,k}\Big) \Big| \\ &= \sum_k \sum_j \int_{I_j} dx \Big| 1 - \frac{x - x_j}{\Delta x} \frac{\hat{S}_{j,k+1} - \hat{S}_{j,k}}{\Delta_+^y U_{j,k}} \Big| |\Delta_+^y U_{j,k}|. \end{aligned}$$

If the following inequality holds for all $x \in I_j$:

$$1 - \frac{x - x_j}{\Delta x} \frac{\hat{S}_{j,k+1} - \hat{S}_{j,k}}{\Delta_+^y U_{j,k}} \geq 0, \tag{3.14}$$

then

$$\begin{aligned} &\int_{I_j} dx \Big| 1 - \frac{x - x_j}{\Delta x} \frac{\hat{S}_{j,k+1} - \hat{S}_{j,k}}{\Delta_+^y U_{j,k}} \Big| |\Delta_+^y U_{j,k}| \\ &= \int_{I_j} dx \Big(1 - \frac{x - x_j}{\Delta x} \frac{\hat{S}_{j,k+1} - \hat{S}_{j,k}}{\Delta_+^y U_{j,k}} \Big) |\Delta_+^y U_{j,k}| \\ &= \Delta x |\Delta_+^y U_{j,k}| + (\hat{S}_{j,k+1} - \hat{S}_{j,k}) \sigma(\Delta_+^y U_{j,k}) \int_{I_j} dx \frac{x - x_j}{\Delta x} \\ &= \Delta x |\Delta_+^y U_{j,k}|, \end{aligned} \tag{3.15}$$

which gives $TV^y(L^xU) \leq TV^y(U)$, since $\int_{I_j} dx(x - x_j) = 0$. It is readily verified that a sufficient condition for 3.14 to be satisfied is

$$|\hat{S}_{j,k+1} - \hat{S}_{j,k}| \leq 2|\Delta_+^y U_{j,k}|. \tag{3.16}$$

The following lemma is often useful when constructing TVD schemes. An operator possessing property 3.17 is said to be monotone or order preserving, and property 3.18 is referred to as L^1-contractiveness.

Lemma 3.1 (Crandall-Tartar lemma [7]). *Let Ω be a measure space and $C \subseteq L^1(\Omega)$ have the property that $f \vee g \in C$ whenever $f, g \in C$. Let $T : C \to L^1(\Omega)$ preserve the integral:*

$$\int_\Omega Tf = \int_\Omega f, \quad \forall f \in C.$$

Then the following properties are equivalent:

$$f \leq g \ \text{ a.e.} \Rightarrow Tf \leq Tg \ \text{ a.e.}, \quad \forall f, g \in C, \tag{3.17}$$

$$\int_\Omega |Tf - Tg| \leq \int_\Omega |f - g|, \quad \forall f, g \in C. \tag{3.18}$$

Lemma 3.2. *The one-dimensional slope-limiter x-step F^{sor} appearing in 3.10 is TVD in the fully two-dimensional sense, $TV(V) \leq TV(U)$, if the underlying one-dimensional approximate evolution operator $E^x(\Delta t)$ is TVD and L^1-contractive and the modified slopes $\hat{S}_{j,k}$ satisfy 3.11 and 3.16.*

Proof. Since $\hat{S}_{j,k}/S_{j,k} \in [0,1]$, $TV^x(V) \leq TV^x(U)$. To show that $TV^y(V) \leq TV^y(U)$, take one operator at a time. That $TV^y(L^x U) \leq TV^y(U)$ was just demonstrated. From the assumption that the approximate evolution operator is L^1-contractive, it follows that $TV^y(E^x(\Delta t) \circ L^x U) \leq TV^y(L^x U)$. Specifically, L^1 contractiveness yields

(3.19)

$$\begin{aligned} TV^y(E^x(\Delta t) \circ L^x U) &= \sum_k \Delta x \sum_j |(E^x(\Delta t) \circ L^x U)_{j,k+1} - (E^x(\Delta t) \circ L^x U)_{j,k}| \\ &= \sum_k ||(E^x(\Delta t) \circ L^x U)_{\cdot,k+1} - (E^x(\Delta t) \circ L^x U)_{\cdot,k}||_1 \\ &\leq \sum_k ||(L^x U)_{\cdot,k+1} - (L^x U)_{\cdot,k}||_1 = TV^y(L^x U). \end{aligned}$$

Finally, the averaging operator $\mathcal{A}$ is monotone and conservative, thus also L^1-contractive by the Crandall-Tartar Lemma. A calculation similar to the one above gives the estimate $TV^y(\mathcal{A} \circ E^x(\Delta t) \circ L^x U) \leq TV^y(E^x(\Delta t) \circ L^x U)$. $\square$

Lemma 3.2 places fairly strong restrictions (TVD and L^1-contractive) on the approximate evolution operator $E^x(\Delta t)$. Obviously, the exact solution operator satisfies the requirements, but this is often not a practical approach. For nonlinear problems, a usable formula for the resulting numerical flux is not readily available for other than piecewise constant data, the resulting flux in that case being Godunov's. In practice, various approximations to the exact solution operator are used, and the TVD property is often one of the design goals, but the L^1-contractiveness property is not always guaranteed. One example of an operator which has both properties is the transport-collapse operator, which was shown to be a generalization of the Engquist-Osher scheme by Brenier [1]. The transport-collapse operator is somewhat easier to work with analytically than the exact solution operator, making it possible to derive the resulting numerical flux f^{sor} in a form that is usable for computations [31].

It will simplify the nonlocal limiter algorithm described in section 4 if $S_{j,k}S_{j,k+1} \geq 0$. For that reason, it will be assumed in what follows that $S_{j,k}$ has been set equal to zero if either $S_{j,k}S_{j,k+1} < 0$ or $S_{j,k}S_{j,k-1} < 0$. This only affects the slopes at zeros of u_x, and so the SOR property of the scheme is not affected.

For a one-dimensional y time step $W = G(U; \alpha^y \lambda^y)$, with slopes

$$T_{j,k} = M(\Delta_+^y U_{j,k}, \Delta_-^y U_{j,k}) \tag{3.20}$$

set of inequalities to be satisfied is

$$0 \leq \hat{T}_{j,k}/T_{j,k} \leq 1, \quad |\hat{T}_{j+1,k} - \hat{T}_{j,k}| \leq 2|\Delta_+^x U_{j,k}|. \tag{3.21}$$

In generic smooth regions where neither u_x nor u_y vanishes,

$$\Delta_+^y S_{j,k} = O(\Delta^2), \quad \Delta_+^x T_{j,k} = O(\Delta^2),$$

while

$$\Delta_+^y U_{j,k} = O(\Delta), \quad \Delta_+^x U_{j,k} = O(\Delta),$$

so it seems reasonable to expect that for small enough Δ there are nontrivial solutions to the systems of equations 3.11, 3.16 for the modified slopes $\hat{S}_{j,k}$ and the analogous system 3.21 for the modified slopes $\hat{T}_{j,k}$. This will be proven in the next two sections. The following lemma summarizes the results of this section.

Lemma 3.3. *If the slope-limiter SOR fluxes f^{sor} and g^{sor} are constructed as in equation 3.2 with the x-slopes $\hat{S}_{j+\frac{1}{2},k}$, satisfying 3.11, 3.16, and the y-slopes $\hat{T}_{j,k+\frac{1}{2}}$ satisfying 3.20, 3.21, then the slope-limiter Strang fractional step method*

$$U^{n+1} = \hat{F}^{sor}(\cdot;\lambda^x/2) \circ \hat{G}^{sor}(\cdot;\lambda^y) \circ \hat{F}^{sor}(U^n;\lambda^x/2) \tag{3.22}$$

and the slope-limiter Godunov fractional step method

$$U^{n+1} = \hat{G}^{sor}(\cdot;\lambda^y) \circ \hat{F}^{sor}(U^n;\lambda^x) \tag{3.23}$$

are TVD in the fully two-dimensional sense. Here

$$\hat{F}^{sor}(U;\alpha^x\lambda^x)_{j,k} = U_{j,k} - \alpha^x\lambda^x\Delta_-^x f^{sor}_{j+\frac{1}{2},k}, \tag{3.24}$$

$$\hat{G}^{sor}(U;\alpha^y\lambda^y)_{j,k} = U_{j,k} - \alpha^y\lambda^y\Delta_-^y g^{sor}_{j,k+\frac{1}{2}}. \tag{3.25}$$

4. The nonlocal limiter algorithm

This section presents an algorithm for solving systems of inequalities of the form

$$|\hat{z}_{i+1} - \hat{z}_i| \le \theta_i, \tag{4.1}$$

$$0 \le \hat{z}_i/z_i \le 1. \tag{4.2}$$

The unknowns are $\hat{z}_i$, and the data z_i, θ_i satisfy

$$\theta_i \ge 0, \quad z_i z_{i+1} \ge 0. \tag{4.3}$$

The z_i are assumed to vanish for sufficiently large values of the index i. Specifically, there are indices i_*, i^* such that

$$i \le i_* \Rightarrow z_i = 0, \quad i \ge i^* \Rightarrow z_i = 0. \tag{4.4}$$

Systems of inequalities of this form were derived in section 2 and section 3. In section 2, each x-step of the fractional step algorithm produced a system of this form with

$$z_i = d_{i-\frac{1}{2},k}, \quad \theta_i = \frac{\rho}{2\alpha^x\lambda^x}|M(\Delta_-^y U_{i,k}, \Delta_+^y U_{i,k})|.$$

Each y-step produced a similar system. In section 3, for each fixed j, there was a system of this form with

$$z_i = S_{j,i}, \quad \theta_i = 2|\Delta_+^y U_{j,i}|.$$

For both the flux-limiter schemes and the slope-limiter schemes, the z_i were modified specifically so that $z_i z_{i+1} \ge 0$ would hold. This condition simplifies the algorithm presented below. The existence of i^* and i_* results from the assumption that the solution is constant for large $|(x,y)|$.

The nonlocal limiter algorithm.

Forward sweep: For $i = i_*$ increasing to $i = i^* - 1$: If $|z_{i+1}| > |z_i|$, then $z_{i+1} \leftarrow z_i + \sigma(z_{i+1} - z_i)\min(|z_{i+1} - z_i|, \theta_i)$.

Backward sweep: For $i = i^*$ decreasing to $i = i_* + 1$: If $|z_{i-1}| > |z_i|$, then $z_{i-1} \leftarrow z_i + \sigma(z_{i-1} - z_i)\min(|z_{i-1} - z_i|, \theta_{i-1})$.

Here the left arrow $\leftarrow$ is the replacement operator. The algorithm can be written compactly as

$$\hat{Z} = \Phi(Z, \Theta) = \Phi^{-}(\Phi^{+}(Z, \Theta), \Theta),$$

where Φ^{+} and Φ^{-} denote the forward and backward sweeps, and

$$\hat{Z} = \{\hat{z}_i\}, \quad Z = \{z_i\}, \quad \Theta = \{\theta_i\}.$$

The operation of the limiter algorithm is best understood by first considering the case where all of the z_i are nonnegative. The forward sweep visits each point z_i in the order of increasing i. If $z_{i-1} \geq z_i$, nothing happens to z_i on the forward sweep. If $z_{i-1} < z_i$, the constraint $|z_i - z_{i-1}| \leq \theta_{i-1}$ is checked. Nothing happens to z_i if the constraint is satisfied, but if it is violated, then z_i is moved toward z_{i-1} (decreased) by exactly enough to satisfy the constraint. The points z_{i_*} and z_{i^*} are clamped at zero, so they never change. On the backward sweep, each point z_i is visited, this time in the order of decreasing i. Nothing happens to z_i if $z_i \leq z_{i+1}$. Otherwise the constraint $|z_i - z_{i+1}| \leq \theta_i$ is checked, and if there is a violation, then z_i is decreased just enough to satisfy the constraint. The algorithm behaves in an analogous way when all of the z_i are nonpositive. Any contiguous pair of intervals where the z_i have opposite signs are separated by a z_i equaling zero, which never changes in value. Thus the algorithm operates independently on intervals where the z_i do not change sign. This allows the algorithm to be understood in terms of how it operates on data of one sign.

The assumption $z_i z_{i+1} \geq 0$ simplifies the limiter algorithm and its analysis, but is not necessary. To handle the more general situation, the only modification required to the algorithm is that when one of the z_i is moved in order to satisfy a constraint, it only moves as far as the origin, i.e., no sign changes are allowed to occur. In the interest of not obscuring the main idea, this more general situation is not pursued further.

Lemma 4.1. *The output of the nonlocal limiter algorithm solves the system of inequalities 4.1, 4.2. The solution is optimal in the sense that if $\{\tilde{z}_i\}$ is another solution, then $|\tilde{z}_i| \leq |\hat{z}_i|$ for all i. The solution is independent of the ordering of the sweeps, i.e., the following symmetry property holds*

$$\hat{Z} = \Phi^{-}(\Phi^{+}(Z, \Theta), \Theta) = \Phi^{+}(\Phi^{-}(Z, \Theta), \Theta). \tag{4.5}$$

Proof. Let $\bar{z}_{i+1} = \Phi^{+}(Z, \Theta)_{i+1}$, and $\bar{Z} = \{\bar{z}_i\}$. It is easy to check by induction on i that

$$0 \leq \bar{z}_{i+1}/z_{i+1} \leq 1. \tag{4.6}$$

Similarly, with the notation $\hat{z}_{i-1} = \Phi(\bar{Z}, \Theta)_{i-1}$,

$$0 \leq \hat{z}_{i-1}/\bar{z}_{i-1} \leq 1. \tag{4.7}$$

Shifting indices, and multiplying 4.6 by 4.7 gives

$$0 \leq \hat{z}_{i-1}/z_{i-1} \leq 1. \tag{4.8}$$

From the definitions of the forward and backward sweeps, it is clear that

$$|\bar{z}_{i+1}| \geq |\bar{z}_i| \Rightarrow |\bar{z}_{i+1} - \bar{z}_i| \leq \theta_i, \tag{4.9}$$

$$|\hat{z}_{i-1}| \geq |\hat{z}_i| \Rightarrow |\bar{z}_{i-1} - \hat{z}_i| \leq \theta_{i-1}. \tag{4.10}$$

Now, suppose that $|\hat{z}_{i+1}| \geq |\hat{z}_i|$. It follows from the definition of the backward sweep that $\hat{z}_i = \bar{z}_i$. Then since $|\hat{z}_{i+1}| \leq |\bar{z}_{i+1}|$,

$$|\bar{z}_{i+1}| \geq |\bar{z}_i|. \tag{4.11}$$

By 4.9, $|\bar{z}_{i+1} - \bar{z}_i| \leq \theta_i$. and since $\hat{z}_{i+1}$ lies between $\hat{z}_i$ and $\bar{z}_{i+1}$,

$$|\hat{z}_{i+1} - \hat{z}_i| \leq \theta_i. \tag{4.12}$$

The proof that $\hat{z}_i$ solves the inequalities is completed by combining 4.8, 4.10, and 4.12.

The proof that $\{\hat{z}_i\}$ is optimal begins by showing that it is impossible to increase $|\bar{z}_i|$ without violating one of the constraints 4.1, 4.2. This is immediate for $i = i_*$ since $z_{i_*} = 0$. Proceeding inductively, suppose that this is true for $i = i_*, i_* + 1, \ldots, i_* + m$. Consider $\bar{z}_{i_*+m+1}$. If $\bar{z}_{i_*+m+1} = z_{i_*+m+1}$, then it is impossible to increase $|\bar{z}_{i_*+m+1}|$ without violating inequality 4.2. If $\bar{z}_{i_*+m+1} \neq z_{i_*+m+1}$, then z_{i_*+m+1} was modified by the forward sweep of the limiter and

$$|\bar{z}_{i_*+m+1} - \bar{z}_{i_*+m}| = \theta_{i_*+m}, \quad |\bar{z}_{i_*+m+1}| \geq |\bar{z}_{i_*+m}|.$$

Any increase in $|\bar{z}_{i_*+m+1}|$ will require the same increase in $|\bar{z}_{i_*+m}|$ to keep the constraint 4.1 satisfied. But this would violate the induction hypothesis, and so it is impossible to increase $|\bar{z}_{i_*+m+1}|$. The same type of argument, applied to the backward sweep of the limiter proves that it is impossible to increase $|\hat{z}_i|$ without violating a constraint, which completes the proof of optimality.

The symmetry property 4.5 results from the fact that the preceding argument works equally well on the reversed version of the algorithm. In particular, the reversed version also produces an optimal solution. The proof is completed with the observation that the optimal solution is unique. □

Viewing the computational domain as an array with x varying along rows, and y varying along columns, the limiter Φ operates either one row at a time, or one column at a time. For an x-step of the slope limiter scheme, Φ operates after all of the slopes have been computed, and operates on one column at a time, i.e. perpendicular to the differencing direction. For an x step of the flux limiter scheme, Φ is applied one row (parallel to the differencing direction) at a time, and can be applied as soon as the flux corrections $d_{j+\frac{1}{2},k}$ have been computed for a given row.

5. Compatibility of the nonlocal limiter with second order accuracy

It remains to check that applying the nonlocal limiter Φ does not destroy the formal second order accuracy of the one-dimensional operators. The main result of this section is that for a sufficiently fine mesh size, the second order terms in either type of scheme are unchanged by the nonlocal limiter, away from zeros of u_x and u_y. Let $B_r(\xi,\eta) = \{(x,y) : |x-\xi| < r, |y-\eta| < r\}$.

Lemma 5.1. *Let $u : \mathbf{R}^2 \to \mathbf{R}^1$ be C^2 in a neighborhood of the point (ξ,η) where*

$$u_x(\xi,\eta) \neq 0, \quad u_y(\xi,\eta) \neq 0. \tag{5.1}$$

Assume that $u(x,y) = u_\infty$ for $|(x,y)|$ sufficiently large, so that the limiter Φ is well-defined on the x-slopes $S^\Delta_{j,k} = M(\Delta^x_+ u(x_j,y_k), \Delta^x_- u(x_j,y_k))$, and the y-slopes $T^\Delta_{j,k} = M(\Delta^y_+ u(x_j,y_k), \Delta^y_- u(x_j,y_k))$. Let

$$\hat{S}^\Delta_{j,k} = \Phi(\{S^\Delta_{j,i}\}, \{2|\Delta^y_+ u(x_j,y_i)|\})_k, \quad \hat{T}^\Delta_{j,k} = \Phi(\{T^\Delta_{i,k}\}, \{2|\Delta^x_+ u(x_i,y_k)|\})_j$$

where Φ is understood to operate on $S^{\Delta}_{j,k}$ with j fixed and k as the variable, and operate on $T^{\Delta}_{j,k}$ with k fixed and j as the variable . Then there is a mesh size $\Delta_0(\xi,\eta) > 0$ and a $\delta(\xi,\eta) > 0$ such that for $\Delta \le \Delta_0$,

$$(x_j, y_k) \in B_\delta(\xi,\eta) \Rightarrow \hat{S}^{\Delta}_{j,k} = S^{\Delta}_{j,k},\ \hat{T}^{\Delta}_{j,k} = T^{\Delta}_{j,k}.$$

Proof. Choose $\delta > 0$ and $\epsilon > 0$ so that $u \in C^2(B_{3\delta}(\xi,\eta))$ and for $(x,y) \in B_{3\delta}(\xi,\eta)$

$$|u_x(x,y)| > \epsilon, \quad |u_y(x,y)| > \epsilon.$$

Choose $\Delta_0 > 0$ such that the following conditions hold for $\Delta < \Delta_0$ and $(x_j, y_k) \in B_{2\delta}(\xi,\eta)$:

$$|S^{\Delta}(x_j, y_{k+1}) - S^{\Delta}(x_j, y_k)| \le 2|u(x_j, y_{k+1}) - u(x_j, y_k)|, \tag{5.2}$$

$$\sigma(u(x_j, y_{k+1}) - u(x_j, y_k)) = \sigma(u_y(\xi,\eta)), \tag{5.3}$$

$$|S^{\Delta}_{j,k}| \le \epsilon\delta/2, \tag{5.4}$$

$$\sigma(S^{\Delta}_{j,k}) = \sigma(u_x(\xi,\eta)). \tag{5.5}$$

That it is possible to choose Δ_0 so that 5.2 is satisfied is a result of the fact that the left side of 5.2 is uniformly $O(\Delta^2)$ on $\overline{B_{2\delta}(\xi,\eta)}$ while the right side is $O(\Delta)$ with the coefficient of the Δ term bounded away from zero (because of 5.3). That 5.3, 5.4, and 5.5 hold for $\Delta \le \Delta_0$ is a consequence of the mean value theorem. Without loss of generality, assume that $u_x(x,y) > 0$ and $S^{\Delta}_{j,k} > 0$ in the region of interest. Let $y^- = \eta - \delta$, $y^+ = \eta + \delta$. The immediate objective is to prove that for $\Delta \le \Delta_0$,

$$\bar{S}^{\Delta}_{j,k} = S^{\Delta}_{j,k}, \quad \forall (x_j, y_k) \in (\xi - 2\delta, \xi + 2\delta) \times (y^-, \eta + 2\delta), \tag{5.6}$$

where $\bar{S}^{\Delta}_{j,k}$ is the output of the forward sweep of the nonlocal limiter algorithm. By way of contradiction, suppose that 5.6 fails, and choose $x_J \in (\xi - 2\delta, \xi + 2\delta)$ and $\Delta_1 \le \Delta_0$ such that

$$\bar{S}^{\Delta_1}_{J,K} \ne S^{\Delta_1}_{J,K} \tag{5.7}$$

for some $y_K \in (y^-, \eta + 2\delta)$. Because of assumption 5.2, it must be that the forward pass of the limiter has modified all of the slopes $S^{\Delta}_{J,k}$ with $y_k \in [y^- - \delta, y_K)$, and so for $y^- - \delta \le y_k < y_K$,

$$|\bar{S}^{\Delta_1}_{J,k+1} - \bar{S}^{\Delta_1}_{J,k}| = 2|\Delta^y_+ U^{\Delta_1}_{J,k}|. \tag{5.8}$$

Since all of the slopes in the area of interest are assumed positive and have been modified by the forward pass of the limiter,

$$0 \le \bar{S}^{\Delta_1}_{J,k} \le \bar{S}^{\Delta_1}_{J,k+1} \tag{5.9}$$

for $y^- - \delta \le y_k < y_K$. Let $P = \max\{p \in \mathbf{Z}^+ : p \le \delta/\Delta_1\}$. Combining 5.8 and 5.9 yields

$$\bar{S}^{\Delta_1}_{J,K-p} - \bar{S}^{\Delta_1}_{J,K-p-1} = 2|\Delta^y_+ u(x_J, y_{K-p-1})| \tag{5.10}$$

for $0 \le p \le P - 1$. Summing 5.10 over p results in

$$\bar{S}^{\Delta_1}_{J,K} - \bar{S}^{\Delta_1}_{J,K-P} = 2\sum_{p=0}^{P-1} |\Delta^y_+ u(x_J, y_{K-p})| \ge 2P\epsilon\Delta_1 \ge \delta\epsilon. \tag{5.11}$$

On the other hand, it follows from 5.4 and 5.9 that

$$\bar{S}^{\Delta_1}_{J,K} - \bar{S}^{\Delta_1}_{J,K-P} \le \bar{S}^{\Delta_1}_{J,K} \le \delta\epsilon/2,$$

which gives the desired contradiction.

A symmetric argument applied to the backward sweep of the nonlocal limiter algorithm, operating on $\{\bar{S}^{\Delta}_{j,k}\}$, proves that for some $0 < \tilde{\Delta}_0 \le \Delta_0$,

$$\hat{S}^{\Delta}_{j,k} = S^{\Delta}_{j,k}, \quad \forall (x_j, y_k) \in (\xi - 2\delta, \xi + 2\delta) \times (y^-, y^+),$$

for $\Delta \le \tilde{\Delta}_0$. Replacing Δ_0 by $\tilde{\Delta}_0$ completes the argument for the x-slopes. Repeating the argument for the y-slopes, and replacing Δ_0 as necessary completes the proof. □

Lemma 5.2. *Let $u : \mathbf{R}^2 \to \mathbf{R}^1$ be C^2 in a neighborhood of the point (ξ, η) where 5.1 holds. Assume that $u(x, y) = u_\infty$ for $|(x, y)|$ sufficiently large, so that the limiter Φ is well-defined on the x-corrections $d^{\Delta}_{j+\frac{1}{2},k}$, and the similarly defined y-corrections $e^{\Delta}_{j,k+\frac{1}{2}}$. Let*

$$\hat{d}^{\Delta}_{j-\frac{1}{2},k} = \Phi\Big(\{d^{\Delta}_{i-\frac{1}{2},k}\}, \{\frac{\rho}{2\alpha^x\lambda^x}|M(\Delta^y_- u(x_i, y_k), \Delta^y_+ u(x_i, y_k))|\}\Big)_j,$$

$$\hat{e}^{\Delta}_{j,k-\frac{1}{2}} = \Phi\Big(\{e^{\Delta}_{j,i-\frac{1}{2}}\}, \{\frac{\rho}{2\alpha^y\lambda^y}|M(\Delta^x_- u(x_j, y_i), \Delta^x_+ u(x_j, y_i))|\}\Big)_k,$$

where Φ is understood to operate on $d_{j+\frac{1}{2},k}$ with k fixed and j as the variable, and operate on $e_{j,k+\frac{1}{2}}$ with j fixed and k as the variable . Then there is a mesh size $\Delta_0(\xi, \eta) > 0$ and a $\delta(\xi, \eta) > 0$ such that for $\Delta \le \Delta_0$,

$$(x_j, y_k) \in B_\delta(\xi, \eta) \Rightarrow \hat{d}^{\Delta}_{j+\frac{1}{2},k} = d^{\Delta}_{j+\frac{1}{2},k},\ \hat{e}^{\Delta}_{j,k+\frac{1}{2}} = e^{\Delta}_{j,k+\frac{1}{2}}.$$

Proof. Take the x-corrections $d_{j+\frac{1}{2},k}$. As in the proof of lemma 5.1, choose $\Delta_0 > 0$ such that the following conditions hold for $\Delta < \Delta_0$ and $(x_j, y_k) \in B_{2\delta}(\xi, \eta)$:

(5.12)
$$|d^{\Delta}_{j+\frac{1}{2},k} - d^{\Delta}_{j-\frac{1}{2},k}| \le \frac{\rho}{2\alpha^x\lambda^x}|M((u(x_j, y_{k+1}) - u(x_j, y_k)), (u(x_j, y_{k-1}) - u(x_j, y_k)))|,$$

$$\sigma(u(x_j, y_{k+1}) - u(x_j, y_k)) = \sigma(u_y(\xi, \eta)), \tag{5.13}$$

$$|d^{\Delta}_{j+\frac{1}{2},k}| \le \epsilon\delta/4, \tag{5.14}$$

$$\sigma(d^{\Delta}_{j+\frac{1}{2},k}) = \sigma(u_x(\xi, \eta)). \tag{5.15}$$

It is possible to choose Δ_0 to satisfy conditions 5.12, 5.13, and 5.14 for essentially the same reasons as the corresponding conditions 5.2, 5.3, and 5.4 in the proof of lemma 5.1. That it is possible to satisfy 5.15 can be verified by checking formulas 2.9 and 2.18, which show that the x-corrections $d_{j+\frac{1}{2},k}$ have the same sign as $\Delta^x_+ u(x_j, y_k)$. The rest of the proof is basically the same as the proof of lemma 5.1. □

In order to guarantee that the z_i vanish for sufficiently large i, it has been assumed that $u(x, y) = u_\infty$ for $|(x, y)|$ sufficiently large. This assumption is not always realistic. For more general situations, it may be necessary to simply set the slopes (for slope limiter schemes) or flux corrections (for flux limiter schemes) equal to zero at the boundary of the computational domain. The proof of theorem 5.1 shows that the action of the limiter is very localized once the mesh is small enough, so that simply setting the z_i equal to zero at the boundaries will not destroy the SOR property at interior points where $u_x \ne 0$, $u_y \ne 0$.

6. Convergence to a weak solution

This section provides a proof that the sequence of computed approximations is convergent (modulo extraction of a subsequence), and that the limit is a weak solution of the cauchy problem 1.2.

Theorem 6.1. *The Strang and Godunov fractional step methods 2.34 and 2.35 built from flux-limiter one-dimensional x-steps of the form 2.36 and y-steps of the form 2.37 as in lemma 2.2, with the modified flux corrections computed using the nonlocal limiter Φ, are TVD in the fully two-dimensional sense. The Strang method has formal second order accuracy, and the Godunov method has formal second order spatial accuracy, at any point where neither u_x or u_y vanishes. Assume that the initial data u_0 is bounded, has compact support, and lies in $BV(\mathbf{R}^2)$. Let the mesh size $\Delta \to 0$ with λ^x and λ^y fixed and satsifying the CFL condition*

$$(6.1)\quad \alpha^x\lambda^x|f'(w)| \le 1/3, \quad \alpha^y\lambda^y|g'(w)| \le 1/3, \quad \forall w \in \Big[\inf_{x,y} u_0(x,y), \sup_{x,y} u_0(x,y)\Big].$$

Let u^Δ be the sequence of approximations generated by the flux-limiter scheme, with either the Strang splitting or the Godunov splitting. There is a weak solution u and a subsequence u^{Δ_i} such that $u^{\Delta_i} \to u$ in L^1_{loc} and boundedly a.e.

Proof. The TVD property is a result of the way that the scheme was constructed, and the properties of the limiter Φ. Lemma 5.1 establishes the accuracy claim. To prove convergence, take the first x-step $V_{j,k} = U_{j,k} - \alpha^x\lambda^x\Delta_-^x f^{sor}_{j+\frac{1}{2},k}$ where $\alpha^x = 1$ for the Godunov fractional step method or $\alpha^x = 1/2$ for the Strang method. This step can be written in incremental form

$$V_{j,k} = U_{j,k} + \tilde{C}^+_{j+\frac{1}{2},k}\Delta_+U_{j,k} - \tilde{C}^-_{j-\frac{1}{2},k}\Delta_-U_{j,k}.$$

Thanks to the CFL condition the incremental coefficients satisfy

$$\tilde{C}^+_{j+\frac{1}{2},k} \in [0,1], \quad \tilde{C}^-_{j-\frac{1}{2},k} \in [0,1], \quad \tilde{C}^+_{j+\frac{1}{2},k} + \tilde{C}^-_{j-\frac{1}{2},k} \le 1,$$

from which it follows that $V_{j,k}$ lies in the convex hull of $U_{j-1,k}$, $U_{j,k}$, and $U_{j+1,k}$. Thus $V_{j,k} \in [\underline{u}_0, \overline{u}_0]$, where $\underline{u}_0 = \inf u_0(x,y)$, $\overline{u}_0 = \sup u_0(x,y)$. Thus the CFL condition remains satisfied. By a similar argument applied to the y-step, $W = \tilde{F}^y(V;\alpha^y\lambda^y)$, $W_{j,k} \in [\underline{u}_0, \overline{u}_0]$. Continuing this way by induction establishes both the uniform L^∞ bound

$$(6.2)\qquad ||u^\Delta(\cdot,\cdot,t^n)||_\infty \le ||u_0||_\infty$$

and the fact that the CFL condition for the Godunov splitting

$$\lambda^x \sup_{j,k}|f'(U^n_{j,k})| \le 1/3, \quad \lambda^y \sup_{j,k}|g'(U^{n+1/2}_{j,k})| \le 1/3.$$

or the CFL condition for the Strang splitting

$$\frac{\lambda^x}{2}\sup_{j,k}|f'(U^n_{j,k})| \le 1/3, \quad \lambda^y \sup_{j,k}|g'(U^{n+1/4}_{j,k})| \le 1/3, \quad \frac{\lambda^x}{2}\sup_{j,k}|f'(U^{n+1/2}_{j,k})| \le 1/3$$

is satisfied at each time step. This makes it possible to take $\rho = 2/3$ in the inequalities 2.29, 2.32. Take $B > 0$ such that $u_0(x,y) = 0$ for (x,y) outside of $[-B,B]\times[-B,B]$. Due to the fact that the solution after each fractional timestep lies within the convex hull of the three closest mesh points at the lower time level, the support of the generated approximations remains bounded (at each time level), independent of the mesh size:

$$|x| > B + t^n/\lambda^x \Rightarrow u^\Delta(x,y,t^n) = 0, \quad |y| > B + t^n/\lambda^y \Rightarrow u^\Delta(x,y,t^n) = 0. \tag{6.3}$$

This yields the following L^1 bound, which holds independent of the mesh size Δ for $t \le T$:

$$||u^\Delta(\cdot,\cdot,t)||_1 \le (B + T/\lambda^x)(B + T/\lambda^y)||u_0||_\infty. \tag{6.4}$$

For the time continuity estimate, take the Godunov fractional step method; the Strang splitting is similar. Let L_f be a Lipschitz constant for the EO flux f^{eo}, eg., $||f'||_\infty$. For the x-step

$$\begin{aligned}\Delta x\Delta y \sum_{j,k} |U^{n+1/2}_{j,k} - U^n_{j,k}| &= \Delta x\Delta y \sum_{j,k} |\lambda^x \Delta^x_- f^{eo}_{j+\frac12,k} + \lambda^x \Delta^x_- \hat{d}_{j+\frac12,k}| \\ &\le \Delta x\Delta y \sum_{j,k} \Big(\lambda^x L_f(|\Delta^x_+ U^n_{j,k}| + |\Delta^x_- U^n_{j,k}|) + \lambda^x |\Delta^x_- \hat{d}_{j+\frac12,k}|\Big).\end{aligned} \tag{6.5}$$

The flux corrections satisfy

$$\sum_{j,k} |\Delta^x_- \hat{d}_{j+\frac12,k}| \le 2\sum_{j,k} |\hat{d}_{j+\frac12,k}| \le 2\sum_{j,k} |d_{j+\frac12,k}| \le 2L_f \sum_{j,k} |\Delta^x_+ U^n_{j,k}|.$$

Plugging this into 6.5 and using the TVD property yields

$$\Delta x\Delta y \sum_{j,k} |U^{n+1/2}_{j,k} - U^n_{j,k}| \le K^x \Delta x TV(U^0), \tag{6.6}$$

where $K^x = 4\lambda^x L_f$. The corresponding estimate for the y-step is, with $K^y = 4\lambda^y L_g$,

$$\Delta x\Delta y \sum_{j,k} |U^{n+1}_{j,k} - U^{n+1/2}_{j,k}| \le K^y \Delta y TV(U^0). \tag{6.7}$$

Combining 6.6 and 6.7,

$$\Delta x\Delta y \sum_{j,k} |U^{n+1}_{j,k} - U^n_{j,k}| \le (K^x/\lambda^x + K^y/\lambda^y) TV(u_0)\Delta t, \tag{6.8}$$

from which it follows that

$$||u(\cdot,\cdot,t+\tau) - u(\cdot,\cdot,t)||_1 \le K(\tau + \Delta t) \tag{6.9}$$

where K is independent of the mesh size Δ and the time t. Combining the variation bound resulting from the fact that the scheme is TVD with the time continuity estimate 6.9, the L^∞ bound 6.2, and the L^1 bound 6.4, the existence of a convergent subsequence is guaranteed, by standard arguments as in eg., [5] or [27].

It remains to verify that the limit is a weak solution. In the interest of simplicity, take the Godunov splitting; the argument is similar but more complicated for the Strang fractional step method. Composing the two half steps gives

$$\begin{aligned}U^{n+1}_{j,k} = U^n_{j,k} &- \lambda^x \Delta^x_+ (f^{eo}(U^n_{j,k}, U^n_{j-1,k}) - \hat{d}^n_{j-\frac12,k}) \\ &- \lambda^y \Delta^y_+ (g^{eo}(U^n_{j,k}, U^n_{j,k-1}) - \hat{e}^{n+1/2}_{j,k-\frac12}) \\ &- \lambda^y \Delta^y_+ (g^{eo}(U^{n-1/2}_{j,k}, U^{n+1/2}_{j,k-1}) - g^{eo}(U^n_{j,k}, U^n_{j,k-1})).\end{aligned} \tag{6.10}$$

Let ψ be a C^1 test function with compact support contained in $\mathbf{R}^2 \times [0,\infty)$. Proceeding as in the proof of the Lax-Wendroff theorem (see eg., [5], [6], [20], or [21]) gives, in addition to the usual terms that would appear in the case of a two-dimensional monotone scheme, terms of the form

$$\Delta t \Delta x \Delta y \sum_{n\geq 0} \sum_{j,k} \lambda^x \hat{d}^n_{j-\frac{1}{2},k} (\Delta^x_+ \psi^n_{j,k}/\Delta x), \tag{6.11}$$

$$\Delta t \Delta x \Delta y \sum_{n\geq 0} \sum_{j,k} \lambda^y \hat{e}^{n+1/2}_{j,k-\frac{1}{2}} (\Delta^y_+ \psi^n_{j,y}/\Delta y), \tag{6.12}$$

(6.13)
$$\Delta t \Delta x \Delta y \sum_{n\geq 0} \sum_{j,k} \lambda^y (g^{eo}(U^{n+1/2}_{j,k}, U^{n+1/2}_{j,k-1}) - g^{eo}(U^n_{j,k}, U^n_{j,k-1}))(\Delta^y_+ \psi^n_{j,k}/\Delta y).$$

For the first term,

$$|\hat{d}^n_{j-\frac{1}{2},k}| \leq |d^n_{j-\frac{1}{2},k}| \leq \Big(\frac{1}{2}|\Delta^x_+ U^n_{j,k}| + \frac{1}{2}|\Delta^x_- U^n_{j,k}|\Big)/\lambda^x,$$

and so the sum 6.11 tends to zero as $\Delta \to 0$, thanks to the uniform bound on $TV^x(U^n)$. A similar observation guarantees that the sum 6.12 tends to zero, this time using the bound for $TV^y(U^{n+1/2})$. Using the Lipschitz continuity of the flux g^{eo} (with Lipschitz constant L_g) the last term is bounded by

(6.14)
$$\Delta t \Delta x \Delta y L_g \sum_{n\geq 0} \sum_{j,k} \lambda^y (|\Delta^y_- U^{n+1/2}_{j,k}| + |\Delta^y_- U^n_{j,k}| + 2|U^{n+1/2}_{j,k} - U^n_{j,k}|)|\Delta^y_+ \psi^n_{j,k}/\Delta y|.$$

The sums involving $|\Delta^y_- U^{n+1/2}_{j,k}|$ and $|\Delta^y_- U^n_{j,k}|$ tend to zero because of the uniform bounds on $TV^y(U^n)$ and $TV^y(U^{n+1/2})$. It only remains to show that

$$\Delta t \Delta x \Delta y L_g \sum_{n\geq 0} \sum_{j,k} \lambda^y |U^{n+1/2}_{j,k} - U^n_{j,k}||\Delta^y_+ \psi^n_{j,k}/\Delta y| \to 0. \tag{6.15}$$

But this is immediate from inequality 6.6, from which it follows that the sum appearing in 6.15 converges to zero. □

Theorem 6.2. *The Strang and Godunov fractional step methods 3.22, 3.23 built from one-dimensional slope-limiter steps of the form 3.24, 3.25, with slopes $\hat{S}_{j,k}$ satisfying 3.11, 3.16 and $\hat{T}_{j,k}$ satisfying 3.20, 3.21 is TVD in the fully two-dimensional sense. The Strang splitting method has formal second order accuracy, and the Godunov splitting method has formal second order spatial accuracy, at any point where neither u_x nor u_y vanishes. Assume that the approximate evolution operators $E^x(\Delta t)$ and $E^y(\Delta t)$ are L^1-contractive, and reduce to monotone schemes with Lipschitz continuous numerical fluxes when applied to piecewise constant data (the trivial reconstruction). Assume that the initial data u_0 is bounded, has compact support, and lies in $BV(\mathbf{R}^2)$. Let the mesh size $\Delta \to 0$ with λ^x and λ^y fixed. Let u^Δ be the sequence of approximations generated the slope-limiter scheme. There is a weak solution u and a subsequence u^{Δ_i} such that $u^{\Delta_i} \to u$ in L^1_{loc} and boundedly a.e.*

Proof. Only the convergence claim requires proof. For the required L^∞ bound, take an x-step

$$V = \mathcal{A} \circ E^x \circ L^x U.$$

Because the minmod limiter is used in computing $L^x U$, $||L^x U||_\infty \leq ||U||_\infty$. With the assumption that E^x is L^1-contractive, it must also be monotone, thanks to the

Crandall-Tartar lemma, and so $||E^x \circ L^x U||_\infty \leq ||L^x U||_\infty$. Finally, the averaging operator creates no new extrema, and so $||\mathcal{A} \circ E^x \circ L^x U||_\infty \leq ||E^x \circ L^x U||_\infty$, and so $||V||_\infty \leq ||U||_\infty$. The same argument applies to a y-step,

$$W = \mathcal{A} \circ E^y \circ L^y V,$$

thus establishing the L^∞ bound, which holds uniformly in $\Delta > 0$. Having established the L^∞ bound, the approximate solutions u^Δ are uniformly bounded in L^1_{loc}. The bound on the total variation is immediate, due to the manner in which the scheme was constructed. For the time-continuity estimate, consider the Godunov splitting. An application of the triangle inequality yields

$$||U^{n+1} - U^n||_1 \leq ||U^{n+1} - U^{n+1/2}||_1 + ||U^{n+1/2} - U^n||_1. \tag{6.16}$$

Using the fact that $U^{n+1/2} = \mathcal{A} \circ E^x \circ L^x U^n$, and the fact that $\mathcal{A} \circ E^x$ is L^1-contractive,

$$\begin{aligned} ||U^{n+1/2} - U^n||_1 &= ||\mathcal{A} \circ E^x \circ L^x U^n - U^n||_1 \\ &\leq ||\mathcal{A} \circ E^x \circ L^x U^n - \mathcal{A} \circ E^x U^n||_1 + ||\mathcal{A} \circ E^x U^n - U^n||_1 \\ &\leq ||L^x U^n - U^n||_1 + ||\mathcal{A} \circ E^x U^n - U^n||_1. \end{aligned} \tag{6.17}$$

By the definition of the reconstruction operator,

$$||L^x U^n - U^n||_1 \leq \Delta x \Delta y \sum_{j,k} \frac{1}{4} |\hat{S}_{j,k}| \leq \frac{1}{4} \Delta x TV^x(U^n) = O(\Delta t), \tag{6.18}$$

and since $\mathcal{A} \circ E^x$ is a monotone scheme with a Lipschitz continuous flux (with Lipschitz constant L_f)

$$||\mathcal{A} \circ E^x U^n - U^n||_1 \leq 2 L_f \lambda^x \Delta x TV^x(U^n) = O(\Delta t),$$

and thus $||U^{n+1/2} - U^n||_1 = O(\Delta t)$. By a similar argument, $||U^{n+1} - U^{n+1/2}||_1 = O(\Delta t)$, and so by 6.16, $||U^{n+1} - U^n||_1 = O(\Delta t)$. The time continuity estimate is established in a similar way for the Strang splitting. Convergence of a subsequence now follows as in the proof of theorem 6.1.

To show that the limit is a weak solution,

$$\begin{aligned} ||U^{n+1/2} - \mathcal{A} \circ E^x U^n||_1 &= ||\mathcal{A} \circ E^x \circ L^x U^n - \mathcal{A} \circ E^x U^n||_1 \\ &\leq ||L^x U^n - U^n||_1 = O(\Delta t), \end{aligned} \tag{6.19}$$

and thus

$$U^{n+1/2} = \mathcal{A} \circ E^x U^n + O(\Delta t). \tag{6.20}$$

Similarly,

$$U^{n+1} = \mathcal{A} \circ E^y U^{n+1/2} + O(\Delta t). \tag{6.21}$$

Combining 6.20 and 6.21 yields

$$U^{n+1} = (\mathcal{A} \circ E^y) \circ (\mathcal{A} \circ E^x) U^n + O(\Delta t). \tag{6.22}$$

Since $\mathcal{A} \circ E^x$ and $\mathcal{A} \circ E^y$ are standard one-dimensional monotone schemes, 6.22 says that aside from terms that are $O(\Delta t)$, the scheme is a a fractional step method whose one-dimensional operators are standard monotone schemes. The Lax-Wendroff theorem is well-known to hold for fractional step methods built from monotone one-dimensional schemes (see eg., [6]). Also, the slope limiter scheme is in conservation form. It follows from 6.22 that the terms arising in the Lax-Wendroff proof that are not present for the purely monotone version are of the form

$$\Delta t \Delta x \Delta y \sum_{n \geq 0} \sum_{j,k} (\Delta_+^x \psi_{j,k}^n / \Delta x) a_{j,k}^n, \quad \Delta t \Delta x \Delta y \sum_{n \geq 0} \sum_{j,k} (\Delta_+^y \psi_{j,k}^n / \Delta y) b_{j,k}^n \tag{6.23}$$

where

$$\Delta x \Delta y \sum_{j,k} |a_{j,k}^n| = O(\Delta t), \quad \Delta x \Delta y \sum_{j,k} |b_{j,k}^n| = O(\Delta t).$$

Thus the sums appearing in 6.23 tend to zero as $\Delta \to 0$, which completes the proof. □

7. Conclusion

This paper has presented a nonlocal limiter that allows for conservation form schemes for two-dimensional conservation laws to be simultaneously TVD and have higher than first order accuracy away from zeros of u_x and u_y. An important issue not addressed is uniqueness and the satisfaction of entropy conditions, and is a potential topic of further investigation.

In numerical tests of these algorithms, the solutions are well-behaved, as expected, with no overshoots or ripples near steep gradients. The amount of diffusion is significantly less than that of a first order scheme, but more than that of a dimension-by-dimension TVD-SOR scheme. To understand this, consider an x-step of both SOR schemes. On an x-step, the dimension-by-dimension TVD scheme degrades to first order accuracy whenever it encounters a zero of u_x, while the fully TVD version degrades to first order at every zero of of u_y in addition to those of u_x. With this in mind, one potential area for further investigation is to develop a TVB version, in order to reduce the action of the limiter to those areas where it is truly needed to control oscillations.

It may be that the nonlocal limiter can be applied to variable coefficient advection equations of the form $u_t + a(x,y)u_x + b(x,y)u_y$, which arise in incompressible flow. Dimension-by-dimension TVD schemes are not always as well behaved on these problems as on nonlinear conservation laws. Reference [30] shows some examples exhibiting signficant overshoot (figures 9.12.6, 9.12.8) using dimension-by-dimension TVD schemes applied to a rotating flow problem.

Acknowledgments

I thank Stanley Osher who suggested this problem as a thesis topic, and generously shared his time, ideas, and enthusiasm while acting as my thesis advisor. I am also grateful to Richard Sanders, who was generous with his help and encouragement.

References

1. Y. Brenier, *Averaged multivalued solutions for scalar conservation laws* , SIAM J. Numer. Anal., 21 (1984), pp. 1013–1037.
2. Y. Brenier, S. Osher, *The discrete one-sided Lipschitz condition for scalar conservation laws* , SIAM J. Numer. Anal., 25 (1988), pp. 8–23.
3. F. Coquel, P. LeFloch, *Convergence of finite difference schemes for conservation laws in several space dimensions: the corrected antidiffusive flux approach*, Math. Comp, 57 (1991), pp. 169–210.
4. F. Coquel, P. LeFloch, *Convergence of finite difference schemes for conservation laws in several space dimensions: a general theory*, SIAM J. Numer. Anal., 30 (1993), pp. 675–700.

5. M. G. Crandall, A. Majda, *Monotone difference approximations for scalar conservation laws*, Math. Comp., 34 (1980), pp. 1–21.
6. M. G. Crandall, A. Majda, *The method of fractional steps for conservation laws*, Numer. Math., 34 (1980), pp. 285–314.
7. M. G. Crandall, L. Tartar, *Some relations between nonexpansive and order preserving mappings*, Proc. Amer. Math. Soc., 78 (1980), pp. 385–390.
8. R.J. DiPerna, *Convergence of approximate solutions to conservation laws*, Arch. Rational Mech. Anal., 82 (1983), pp. 27–70.
9. B. Engquist, S. Osher, *One sided difference approximations for nonlinear conservation laws*, Math. Comp., 36 (1980), pp. 45–75.
10. J. Goodman, R. LeVeque, *On the accuracy of stable schemes for 2D scalar conservation laws*, Math. Comp. 45, pp. 15-21.
11. J. Goodman, R. LeVeque, *A geometric approach to high resolution TVD schemes*, SIAM J. Numer. Anal. 25, pp. 268–284.
12. A. Harten, *High resolution schemes for conservation laws*, J. Comp. Phys., 49, pp. 357–393.
13. A. Harten, *On a class of high resolution total-variation-stable finite difference schemes*, SIAM J. Numer. Anal., 21 (1984), pp. 1–23.
14. A. Harten, B. Engquist, S. Osher, S. Chakravarthy, *Uniformly high order accurate essentially nonoscillatory schemes, III*, J. Comp. Phys., 71 (1987), p.231.
15. A. Harten, J.M. Hyman, P.D. Lax, *On finite difference approximations and entropy conditions for shocks*, Comm. Pure Appl. Math., 29 (1976), pp. 297–322.
16. A. Harten, S. Osher, *Uniformly high-order accurate non-oscillatory schemes I.*, SIAM J. Num. Anal., 24 (1987), pp. 279–309.
17. D. Hoff, *The sharp form of Oleinik's entropy condition in several space variable*, Transactions of the American Mathematical Society, 276 (1983), pp. 707-714.
18. H. Holden, N.H. Risebro, *A method of fractional steps for scalar conservation laws without the CFL condition*, Math. Comp., 60 (1993), pp. 221-232.
19. N. N. Kuznetsov, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasilinear equation*, USSR Comp. Math. Phys., 16, 6 (1976), pp. 105–119.
20. P.D. Lax, B. Wendroff, *Systems of conservation laws*, Comm. Pure Appl. Math., 23 (1960), pp. 217-237.
21. R. J. LeVeque, *Numerical methods for conservation laws*, Birkhauser Verlag, Boston, MA, 1992.
22. K.A. Lie, V. Haugse, K.H. Karlsen, *Dimensional splitting with front tracking and adaptive grid refinement*, Preprint, Available at the URL http://www.math.ntnu.no/conservation/.
23. S. Osher, *Convergence of generalized MUSCL schemes*, SIAM J. Numer. Anal., 22 (1985), pp. 217-235.
24. S. Osher, R. Sanders, *Numerical approximations to nonlinear conservation laws with locally varying time and space grids*, Math. Comp., 41 (1983), pp. 321-336.
25. S. Osher, E. Tadmor, *On the convergence of difference approximations to scalar conservation laws*, Math. Comp., 50 (1988), pp. 19-51.
26. H. Nessyahu, E.Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comp. Phys. 87 (1990), pp. 408–463.
27. R. Sanders, *On convergence of monotone difference schemes with variable spatial differencing*, Math. Comp., 40 (1983), pp. 91–106.
28. C. Shu, *TVB uniformly high order schemes for conservation laws*, Math. Comp., 49 (1987), pp. 105–121.
29. P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM J. Numer. Anal. 21 (1984), pp. 995-1011.
30. J.W. Thomas, *Numerical partial differential equations*, Spring Verlag, New York, NY, 1999.
31. J.D.Towers, *High Resolution finite difference algorithms for conservation laws*, Ph.D. Dissertation, U.C.L.A., 1987.
32. B. Van Leer, *Towards the ultimate conservative difference scheme, II*, J. Comp. Phys. ,14 (1974), pp. 361–370.

jtowers@cts.com