DIFFERENCE SCHEMES STABILIZED BY DISCRETE MOLLIFICATION FOR DEGENERATE PARABOLIC EQUATIONS IN TWO SPACE DIMENSIONS

CARLOS D. ACOSTA[A] AND RAIMUND BÜRGER[B]

ABSTRACT. The discrete mollification method is a convolution-based filtering procedure for the regularization of ill-posed problems. This method is applied here to stabilize explicit schemes, which were first analyzed by Karlsen & Risebro [*M2AN Math. Model. Numer. Anal.* **35** (2001), 239–269], for the solution of initial value problems of strongly degenerate parabolic PDEs in two space dimensions. Two new schemes are proposed, which are based on direction-wise and two-dimensional discrete mollification of the second partial derivatives forming the Laplacian of the diffusion function, respectively. The mollified schemes permit to use substantially larger time steps than the original (basic) scheme. It is proven that both schemes converge to the unique entropy solution of the initial value problem. Numerical examples demonstrate that the mollified schemes are competitive in efficiency, and in many cases significantly more efficient, than the basic scheme.

1. INTRODUCTION

1.1. **Scope.** We study explicit finite difference schemes for the initial value problem

$$u_t + f(u)_x + g(u)_y = \Delta A(u), \quad (x, y) \in \mathbb{R}^2, \quad t \in (0, T], \tag{1.1}$$

$$u(x, y, 0) = u_0(x, y), \quad (x, y) \in \mathbb{R}^2, \tag{1.2}$$

where we assume that

$$A(u) = \int_0^u a(s)\,\mathrm{d}s, \quad a(u) \geq 0. \tag{1.3}$$

We allow that $a(u) = 0$ on u-intervals of positive length, so (1.1) is, in general, a strongly degenerate parabolic equation. Its solutions are in general discontinuous, and need to be defined as entropy solutions. It is well known that certain monotone, and therefore first-order, finite difference schemes converge to the entropy solution of (1.1), (1.2) (Evje & Karlsen, 2000; Karlsen & Risebro, 2001). The term $\Delta A(u)$ is usually discretized in a standard way by direction-wise summation of second finite differences of $A(u)$. If Δx and Δt denote the meshwidth of the underlying Cartesian spatial mesh and the time step, respectively, and $\lambda := \Delta t/\Delta x$, $\mu := \Delta t/\Delta x^2$, then for an explicit scheme a CFL stability condition of the type

$$\alpha\lambda\big(\|f'\|_\infty + \|g'\|_\infty\big) + \beta\mu\|a\|_\infty \leq 1 \tag{1.4}$$

must be satisfied, where the coefficients $\alpha, \beta > 0$ depend on the precise numerical scheme and the choice of the underlying convergence theory.

We herein study numerical schemes in which $\Delta A(u)$ is discretized by means of discrete mollification. Roughly speaking, the mollification-based discretization of this term consists in either taking convex combinations of second differences of $A(u)$ in each direction, taken with respect to multiples of Δx, or in directly approximating $\Delta A(u)$ via convolution of discrete numerical values of u with a kernel in a neighborhood of the meshpoint of interest, and subtracting the value of u at that point. (The first of these alternatives corresponds to a direction-wise implementation of the mollified schemes for one-dimensional problems by

Date: March 4, 2011.

[A]Universidad Nacional de Colombia, Department of Mathematics and Statistics, Manizales, Colombia. E-Mail: cdacostam@unal.edu.co.

[B]CI[2]MA and Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Casilla 160-C, Concepción, Chile. E-Mail: rburger@ing-mat.udec.cl.

Acosta et al. (2011).) Both procedures, addressed herein as "Scheme 1" and "Scheme 2", respectively, give rise to a CFL condition of the type

$$\alpha\lambda\big(\|f'\|_\infty + \|g'\|_\infty\big) + \varepsilon_\eta \beta\mu \|a\|_\infty \leq 1, \tag{1.5}$$

where $\varepsilon_\eta > 0$ is a small parameter associated with the width $2\eta+1$ of the stencil of mollification. Condition (1.5) is more advantageous than (1.4) for $\varepsilon_\eta < 1$, a condition satisfied in most circumstances, since for a given value of Δx it permits to employ a larger time step for the corresponding mollified version than with the basic version of a numerical scheme.

The present paper serves two purposes. Firstly, we prove that both Schemes 1 and 2 converge to the unique entropy solution of the initial value problem (1.1), (1.2). In doing to we will appeal to the solution concept utilized by Karlsen and Risebro (2001), which is based on requiring that $\nabla A(u) \in L^2$ rather than $\nabla A(u) \in L^\infty$ (as in Acosta et al., 2011). This solution concept permits to prove convergence without restrictions on the value of η or on the initial spatial total variation of $\nabla A(u_0)$. Such restrictions were found necessary for the one-dimensional case treated in Acosta et al. (2011), so the present treatment generalizes these results, apart, of course, from the extension to two space dimensions. The proof that the limit u of numerical solutions generated by Scheme 1 or 2 indeed satisfies $\nabla A(u) \in L^2$ leads to fairly involved calculations, but we show that the extension of the corresponding analysis by Karlsen & Risebro (2001) to mollified schemes is straightforward.

Secondly, we present a number of numerical experiments with Schemes 1 and 2 applied to non-degenerate and degenerate problems. Our special interest is the issue of efficiency, that is the reduction of numerical error per CPU time. Schemes 1 and 2 permit to use substantially larger time steps than the basic scheme, but of course one numerical evaluation of $\Delta A(u)$ requires the evaluation of stencils of $4\eta+1$ or even $(2\eta+1)^2$ points (with Schemes 1 and 2, respectively) instead of five points with the basic scheme. Nevertheless, our numerical experiments show that Schemes 1 and 2 in many situations are still more efficient than the basic scheme, and otherwise only slightly less efficient than the basic scheme. We expect that the performance of mollified schemes can still be improved by an optimal choice of the mollification weights, which we will, however, not pursue herein.

1.2. **Motivation and related work.** Equations of the type (1.1) include a large number of known equations such as the heat equation, one-point-degenerate porous-medium-type equations (where $f = g \equiv 0$, $A(u) = u^m$), the two-point degenerate reservoir flow equation (in one space dimension defined by $f(u) = u^2/(u^2 + (1-u)^2)$, $A(u) = u(1-u)$), and strongly degenerate parabolic equations, which appear in models of traffic flow (Rouvre & Gagneux, 1999; Bürger & Karlsen, 2003), sedimentation-consolidation processes (Berres et al., 2003), and aggregation (Betancourt et al., 2011). Furthermore, (1.1) explicitly includes the case $A \equiv 0$, that is, a first-order, nonlinear scalar conservation law. Let us point out, however, that for this case the mollified schemes presented herein reduce to known methods (see e.g. Karlsen & Risebro, 2001), since the mollification device affects the discretization of the parabolic part of (1.1) only. For a short introduction to the well-posedness analysis of strongly degenerate parabolic equations and an up-to-date list of references we refer e.g. to the introductory parts of Holden et al. (2010).

Concerning numerical schemes, we mention that monotone schemes for first-order conservation laws (corresponding to $A \equiv 0$) were introduced by Harten et al. (1976) and Crandall & Majda (1980). It is well known that these schemes convergence to an entropy solution, which remains valid for the application to strongly degenerate parabolic equations. This was first exploited by Evje & Karlsen (2000). Related analyses include implicit monotone schemes for degenerate parabolic equations (Evje & Karlsen, 1999), problems with boundary conditions (Bürger et al., 2006), multidimensional degenerate parabolic equations (Karlsen & Risebro, 2001), equations with discontinuous coefficients (Karlsen et al., 2002, 2003; Bürger et al., 2005), and problems of parameter identification (Coronel et al., 2003) (this list is far from being complete). The disadvantage of monotone schemes is their well-known generic limitation to first-order accuracy.

Discrete mollification is a versatile convolution-based filtering procedure for the regularization of ill-posed problems and the stabilization of explicit schemes for the solution of PDEs. This technique was introduced by Diego A. Murio and collaborators in a series of papers (cf., e.g., Murio, 1993, 2002; Mejía & Murio 1995, 1996; Murio et al., 2001). Acosta & Mejía (2008, 2009) introduced the mollification method as a stabilizer for numerical schemes for strictly parabolic convection-diffusion equations and nonlinear scalar conservation laws. Acosta & Mejía (2010) show that a particular discrete approximation of the second derivative of a smooth function, based on discrete mollification, stabilizes operator splitting methods (Karlsen & Risebro, 1997) for the numerical solution of convection-diffusion problems. Acosta et al. (2011) applied the method of Acosta & Mejía (2010) to strongly degenerate parabolic equations.

1.3. **Outline of the paper.** The remainder of the paper is organized as follows. Section 2 provides some preliminaries, including precise assumptions on the functions A, f and g and the initial datum u_0, a definition of an entropy solution, and L^1_{loc} and L^2_{loc} compactness criteria. In Section 3 we recall some basic facts from one-dimensional discrete mollification and introduce the two-dimensional discrete mollification operator, the basic montone numerical scheme, and its two mollified versions, Scheme 1 and 2. The respective CFL conditions are discussed. Section 4, which is at the core of this paper, is devoted to the convergence analysis, which is split into a series of lemmas. In Lemmas 4.1 and 4.2 we prove that Schemes 1 and 2 are monotone under appropriate CFL conditions, and in Lemma 4.3 we invoke standard arguments to deduce that both schemes are L^∞- and L^1-stable, and total variation diminishing (TVD). In Lemma 4.4 we prove that the numerical solutions generated by both schemes are L^1 Hölder continuous in time. The proof of this lemma is based on the "interpolation lemma" by Kružkov (1969), and includes results related to summation by parts for the mollified schemes that are useful in several instances. Then we prove in Lemmas 4.5 and 4.7 the discrete analogue of $\nabla A(u) \in L^2(\Pi_T)$ for Scheme 1 and 2, respectively. To this end we need to strengthen the CFL condition previously imposed to ensure monotonicity of the schemes. (Lemma 4.6 cites a result from Karlsen & Risebro (2001).) Lemmas 4.8 and 4.9 state discrete entropy inequalities satisfied by Scheme 1 and 2, respectively. The main convergence result is stated in Theorem 4.1. Its proof is based on the previous lemmas, and includes a proof of L^2 continuity in time of the discrete analogue of $A(u)$. Finally, the convergence proof is concluded by appealing to the L^1_{loc} and L^2_{loc} compactness criteria. In Section 5 we present numerical examples for five different cases, and in Section 6 we collect some conclusions.

2. Preliminaries

2.1. **Assumptions.** Concerning the functions A, f and g we assume that

$$A \in \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}), \text{ and } A(\cdot) \text{ is nondecreasing with } A(0) = 0, \tag{2.1}$$

$$f, g : \mathbb{R} \to \mathbb{R}, \quad f, g, f', g' \in \mathrm{Lip}(\mathbb{R}, \mathbb{R}), \tag{2.2}$$

where $f' \equiv \mathrm{d}f/\mathrm{d}u$, $g' \equiv \mathrm{d}g/\mathrm{d}u$. Moreover, the initial datum u_0 is assumed to satisfy

$$u_0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2). \tag{2.3}$$

2.2. **Definition and uniqueness of an entropy solution.**

Definition 2.1. *A measurable function $u = u(x, y, t)$ is said to be an* entropy solution *of* (1.1), (1.2) *if the following conditions are satisfied:*

(1) $u \in L^1(\Pi_T) \cap L^\infty(\Pi_T) \cap C(0, T; L^1(\mathbb{R}^2))$.
(2) *The following entropy inequality holds for all $k \in \mathbb{R}$ and all non-negative test functions $\varphi \in C_0^\infty(\Pi_T)$:*

$$\begin{aligned} \iint_{\Pi_T} \Big\{ &|u-k|\varphi_t + \mathrm{sgn}(u-k)\big((f(u)-f(k))\varphi_x + (g(u)-g(k))\varphi_y\big) \\ &+ \big|A(u)-A(k)\big|\Delta\varphi \Big\} \,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y \geq 0. \end{aligned} \tag{2.4}$$

(3) $A(u) \in L^2(0, T; H^1(\mathbb{R}^2))$.
(4) *The initial condition* (1.2) *is satisfied in the following sense:*

$$\operatorname*{ess\,lim}_{t \downarrow 0} \int_{\mathbb{R}^2} \big|u(x, y, t) - u_0(x, y)\big| \,\mathrm{d}x\,\mathrm{d}y = 0.$$

Satbility of entropy solutions with respect to initial data, and therefore uniqueness, follows from the analysis of a more general equation by Karlsen & Risebro (2003). We may state the following theorem.

Theorem 2.1 (L^1 stability of entropy solutions). *Assume that* (2.1) *and* (2.2) *hold, and that* $u, v \in L^\infty(0,T;BV(\mathbb{R}^2))$ *are entropy solutions of* (1.1), (1.2) *with respective initial data* u_0 *and* v_0, *which are both assumed to satisfy* (2.3). *Then* $\|u(\cdot,t) - v(\cdot,t)\|_{L^1(\mathbb{R}^2)} \leq \|u_0 - v_0\|_{L^1(\mathbb{R}^2)}$ *for almost all* $t \in (0,T)$.

2.3. **Compactness criteria.** To state the following L^1_{loc} compactness criterion we recall that a modulus of continuity is a nondecreasing function $\nu : \mathbb{R}^0_+ \to \mathbb{R}^0_+$ with $\nu(0) = 0$. The following two lemmas are stated in Karlsen & Risebro (2001).

Lemma 2.1 (L^1_{loc} compactness lemma). *Assume that* $\{z_h\}_{h>0}$ *is a sequence of functions defined on* $\mathbb{R}^d \times (0,T)$ *which satisfy the following:*

(1) *There exists a constant* $C_1 > 0$, *which is independent of* h, *such that* $\|z_h(\cdot,t)\|_{L^1(\mathbb{R}^d)} \leq C_1$ *and* $\|z_h(\cdot,t)\|_{L^\infty(\mathbb{R}^d)} \leq C_1$ *for all* $t \in (0,T)$.

(2) *There exists a spatial modulus of continuity* ν *which is independent of* h *such that*
$$\big\|z_h(\cdot + y, t) - z_h(\cdot, t)\big\|_{L^1(\mathbb{R}^d)} \leq \nu\big(|y|; z_h\big) \quad \text{as } y \to 0, \text{ for all } t \in (0,T).$$

(3) *There exists a temporal modulus of continuity* ω *which is independent of* h *such that*
$$\big\|z_h(\cdot, t+\tau) - z_h(\cdot, t)\big\|_{L^1(\mathbb{R}^d)} \leq \omega(\tau; z_h) \quad \text{for all } t \in (0,T), \text{ whenever } \tau \in (0,T).$$

Then $\{z_h\}_{h>0}$ *is compact in the strong topology of* $L^1_{\mathrm{loc}}(\mathbb{R}^d \times (0,T))$. *Moreover, any limit point of* $\{z_h\}_{h>0}$ *belongs to* $L^1(\mathbb{R}^d \times (0,T)) \cap L^\infty(\mathbb{R}^d \times (0,T)) \cap C(0,T;L^1(\mathbb{R}^d))$.

Lemma 2.2 (L^2_{loc} compactness lemma). *Assume that* $\{z_h\}_{h>0}$ *is a sequence of functions defined on* $\mathbb{R}^d \times (0,T)$ *for which there exist constants* $C_1, C_2, C_3 > 0$, *which may depend on* T, *but not on* h, *such that*
$$\begin{aligned} \|z_h\|_{L^2(\mathbb{R}^d \times (0,T))} &\leq C_1, \\ \big\|z_h(\cdot + y, \cdot) - z_h(\cdot, \cdot)\big\|_{L^2(\mathbb{R}^d \times (0,T))} &\leq C_2\big(|y| + h\big) \quad \text{for all } y \text{ as } h \downarrow 0, \\ \big\|z_h(\cdot, \cdot + \tau) - z_h(\cdot, \cdot)\big\|_{L^2(\mathbb{R}^d \times (0,T-\tau))} &\leq C_3\sqrt{|\tau| + h} \quad \text{for all } \tau > 0 \text{ as } h \downarrow 0. \end{aligned}$$

Then $\{z_h\}_{h>0}$ *is compact in the strong topology of* $L^2_{\mathrm{loc}}(\mathbb{R}^d \times (0,T))$. *Moreover, any limit point of* $\{z_h\}_{h>0}$ *belongs to* $L^2(0,T;H^1(\mathbb{R}^d))$.

3. Discrete mollification and numerical schemes

3.1. **One-dimensional discrete mollification.** The mollification method is based on replacing the discrete function $y = \{y_j\}_{j\in\mathbb{Z}}$, which can, for example, consist of evaluations or cell averages of a real function $y = y(x)$ given at equidistant grid points $x_j = x_0 + j\Delta x$, $\Delta x > 0$, $j \in \mathbb{Z}$, by its mollified version $J_\eta y$, where J_η is the so-called *mollification operator* defined by
$$[J_\eta y]_j := \sum_{i=-\eta}^{\eta} w_i y_{j+i}, \quad j \in \mathbb{Z},$$
where $\eta \in \mathbb{N}$ is the support parameter (indicating the width of the mollification stencil) and the so-called weights w_i satisfy
$$w_i = w_{-i}, \quad 0 \leq w_i \leq w_{i-1}, \quad i = 1, \dots, \eta; \quad w_{-\eta} + w_{-\eta+1} + \dots + w_\eta = 1. \tag{3.1}$$
The weights w_i are obtained by numerical integration of the truncated Gaussian kernel
$$\kappa_{p\delta}(t) := \begin{cases} A_p \delta^{-1} \exp(-t^2/\delta^2) & \text{for } |t| \leq p\delta, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{where } A_p := \left(\int_{-p}^{p} \exp(-s^2)\,\mathrm{d}s \right)^{-1},$$
and δ and p are positive parameters. This kernel satisfies $\kappa_{p\delta} \geq 0$, $\kappa_{p\delta} \in C^\infty(-p\delta, p\delta)$, $\kappa_{p\delta} = 0$ outside $[-p\delta, p\delta]$, and $\int_{\mathbb{R}} \kappa_{p\delta} = 1$. Then we define $\xi_{j-1/2} := (j - 1/2)\Delta x$ for $j \in \mathbb{Z}$ and compute the weights by
$$w_i := \int_{\xi_{i-1/2}}^{\xi_{i+1/2}} \kappa_{p\delta}(-s)\,\mathrm{d}s, \quad i = -\eta, \dots, \eta.$$

η	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
0	1								
1	0.84272	0.07864							
2	0.60387	0.19262	5.4438e-3						
3	0.45556	0.23772	3.3291e-2	1.2099e-3					
4	0.36266	0.24003	6.9440e-2	8.7275e-3	4.7268e-4				
5	0.30028	0.22625	9.6723e-2	2.3430e-2	3.2095e-3	2.4798e-4			
6	0.25585	0.20831	0.11241	4.0192e-2	9.5154e-3	1.4905e-3	1.5434e-4		
7	0.22270	0.19058	0.11942	5.4793e-2	1.8403e-2	4.5234e-3	8.1342e-4	1.0697e-4	
8	0.19708	0.17444	0.12097	6.5725e-2	2.7973e-2	9.3255e-3	2.4348e-3	4.9782e-4	7.9691e-5
⋮									
12	0.13476	0.12729	0.10727	8.0645e-2	5.4093e-2	3.2370e-2	1.7282e-2	8.2314e-3	3.4977e-3
η	$i=9$	$i=10$	$i=11$	$i=12$					
12	1.3260e-3	4.4843e-4	1.3529e-4	3.6414e-5					

TABLE 1. Discrete mollification weights w_i.

Usually $p = 3$ is taken and δ, whose role is to determine the shape of the kernel's Gaussian bell, is considered as regularization parameter, and it is estimated by means of methods like Generalized Cross Validation (GCV) (Mejía & Murio, 1996; Murio, 2002). In any case, in this work the main relationship between δ and η is given by $\delta = (\eta + 1/2)\Delta x/p$. This choice generates weights $w_{-\eta}, \dots, w_\eta$, that are independent of Δx. For $p = 3$ these weights are same as those used in previous work (Acosta & Mejía, 2008, 2009; Acosta et al., 2011), and are listed in Table 1.

We conclude this section with some approximation and stability results.

Lemma 3.1. *The discrete mollification operator can be written in the forms*

$$[J_\eta y]_j = y_j + (\psi_j - \psi_{j-1}) = y_j - \sum_{i=1}^{\eta} \rho_i \Delta y_{j-i+1/2} + \sum_{i=1}^{\eta} \rho_i \Delta y_{j+i-1/2},$$

where we define

$$\psi_j := \sum_{k=1}^{\eta} \rho_k \left(y_{j+k} - y_{j-k+1}\right) = \sum_{k=-\eta+1}^{\eta-1} Q_{-k} \Delta y_{j+k+1/2},$$

$$\rho_k := \sum_{i=k}^{\eta} w_i, \quad k = -\eta, \dots, \eta; \quad Q_{-k} = Q_k := \sum_{i=k+1}^{\eta} \rho_i, \quad k = 0, \dots, \eta - 1.$$

We assume that g is a sufficiently smooth real function, set $y_j = g(x_j)$, and employ the Taylor expansion

$$y_{j+i} = y_j + (i\Delta x)\, g'(x_j) + \frac{1}{2}(i\Delta x)^2 g''(x_j) + \frac{1}{6}\,(i\Delta x)^3\, g'''(x_j) + \frac{1}{24}\,(i\Delta x)^4\, g^{(4)}(\xi_{j,i}),$$

where $\xi_{j,i}$ is a real number between x_j and x_{j+i}. Then, defining

$$C_\eta := \left(\sum_{i=-\eta}^{\eta} i^2 w_i\right)^{-1} = \left(2\sum_{i=1}^{\eta} i^2 w_i\right)^{-1}, \tag{3.2}$$

we can write

$$[J_\eta y]_j = \sum_{i=-\eta}^{\eta} w_i y_{j+i} = y_j + \frac{\Delta x^2}{2C_\eta} g''(x_j) + \frac{\Delta x^4}{24} \sum_{i=-\eta}^{\eta} i^4 w_i g^{(4)}(\xi_{j,i}).$$

Theorem 3.1. *Let $g \in C^4(\mathbb{R})$ with $g^{(4)}$ bounded on $\mathbb{R}$, and set $y_j = g(x_j)$. If the data $\{y_j^\varepsilon\}_{j\in\mathbb{Z}}$ satisfy $|y_j^\varepsilon - y_j| \le \varepsilon$ for all $j \in \mathbb{Z}$, then $|[J_\eta y^\varepsilon]_j - [J_\eta y]_j| \le \varepsilon$ for all $j \in \mathbb{Z}$. Additionally, for each compact set $K = [a,b]$ there exists a constant $C = C(K)$ such that*

$$\left| [J_\eta y]_j - g(x_j) - \frac{\Delta x^2}{2C_\eta} g''(x_j) \right| \le C\Delta x^4 \quad \text{for all } j \in \mathbb{Z}. \tag{3.3}$$

Moreover, the following inequalities hold for all $j \in \mathbb{Z}$, where C is a different constant in each inequality:

$$\left|[J_\eta y]_j - g(x_j)\right| \le C\Delta x^2, \qquad \left|\Delta_0 [J_\eta y]_j - \Delta x g'(x_j)\right| \le C\Delta x^3,$$
$$\left|\Delta_+ [J_\eta y]_j - \Delta x g'(x_j)\right| \le C\Delta x^2, \qquad \left|\Delta_-\Delta_+ [J_\eta y]_j - \Delta x^2 g''(x_j)\right| \le C\Delta x^4.$$

3.2. **Two-dimensional discrete mollification.** In this section we consider a uniform bidimensional grid for the (x,y)-plane of the form $(x_i, y_j) = (x_0 + i\Delta x, y_0 + j\Delta y)$ with $\Delta x, \Delta y > 0$. Now, consider a discrete function G defined on the grid by $G(x_i, y_j) = G_{ij}$, again G can be the result of evaluations or cell averages of another function defined on $\mathbb{R}^2$. For each η we define the two-dimensional discrete mollification of G as

$$\left[J_\eta^2 G\right]_{ij} = \sum_{k=-\eta}^{\eta}\sum_{l=-\eta}^{\eta} w_k w_l G_{i+l,j+k},$$

where we use the standard weights w_k of the one-dimensional discrete mollification.

We begin by observing that

$$\left[J_\eta^2 G\right]_{ij} = \sum_{k=-\eta}^{\eta} w_k \left(\sum_{l=-\eta}^{\eta} w_l G_{i+l,j+k} \right) = \sum_{k=-\eta}^{\eta} w_k \left[J_\eta^x G\right]_{i,j+k} = J_\eta^y \left[J_\eta^x G\right]_{ij},$$

where the operators J_η^x and J_η^y are the usual one-dimensional discrete mollification working with respect to the variables x and y respectly. Additionally, if $G_{ij} = G(x_i, y_j)$ is the evaluation of a sufficiently smooth function, from (3.3) we have

$$\begin{aligned}
\left[J_\eta^2 G\right]_{ij} &= \sum_{k=-\eta}^{\eta} w_k \left[J_\eta^x G\right]_{i,j+k} = \sum_{k=-\eta}^{\eta} w_k \left(G_{i,j+k} + \frac{\Delta x^2}{2C_\eta} G_{xx}(x_i, y_{j+k}) + \mathcal{O}(\Delta x^4) \right) \\
&= \sum_{k=-\eta}^{\eta} w_k G_{i,j+k} + \frac{\Delta x^2}{2C_\eta} \sum_{k=-\eta}^{\eta} w_k G_{xx}(x_i, y_{j+k}) + \mathcal{O}(\Delta x^4) \\
&= G_{ij} + \frac{\Delta y^2}{2C_\eta} G_{yy}(x_i, y_j) + \mathcal{O}(\Delta y^4) + \frac{\Delta x^2}{2C_\eta} \left(G_{xx}(x_i, y_j) + \frac{\Delta y^2}{2C_\eta} G_{xxyy}(x_i, y_j) + \mathcal{O}(\Delta y^4) \right) + \mathcal{O}(\Delta x^4) \\
&= G_{ij} + \frac{\Delta x^2}{2C_\eta} G_{xx}(x_i, y_j) + \frac{\Delta y^2}{2C_\eta} G_{yy}(x_i, y_j) + \left[\mathcal{O}(\Delta y^4) + \mathcal{O}(\Delta x^4) + \mathcal{O}(\Delta x^2\Delta y^2) + \mathcal{O}(\Delta x^2 \Delta y^4) \right].
\end{aligned}$$

In the case $\Delta x = \Delta y$, we have

$$\left[J_\eta^2 G\right]_{ij} = G_{ij} + \frac{\Delta x^2}{2C_\eta}(G_{xx} + G_{yy})(x_i, y_j) + \mathcal{O}(\Delta x^4) = G_{ij} + \frac{\Delta x^2}{2C_\eta}(\Delta G)(x_i, y_j) + \mathcal{O}(\Delta x^4). \tag{3.4}$$

For future use we also note the easily verifiable identity

$$[J_\eta^2 G]_{ij} - G_{ij} = \sum_{k=-\eta}^{\eta} w_k \sum_{l=1}^{\eta} w_l (G_{i+k,j+l} - 2G_{i+k,j} + G_{i+k,j-l}) + \sum_{k=1}^{\eta} w_k (G_{i+k,j} - 2G_{ij} + G_{i-k,j}). \tag{3.5}$$

3.3. **Basic and mollified schemes.** For the basic scheme and its mollified versions we assume a Cartesian grid with $\Delta x = \Delta y$, and set $x_i = i\Delta x$, $y_j = j\Delta x$, $x_{i+1/2} = (i+1/2)\Delta x$, $y_{j+1/2} = (j+1/2)\Delta x$. The initial condition (1.2) is discretized by cell averaging, i.e.,

$$u_{ij}^0 = \frac{1}{\Delta x^2} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} u_0(x,y)\,\mathrm{d}y\,\mathrm{d}x \quad \text{for all } i,j \in \mathbb{Z}.$$

η	C_η	ε^1_η	$\tilde\varepsilon^2_\eta$	ε^2_η
1	6.3581	1.0000	0.9214	1.8427
2	2.3321	0.9238	0.7409	1.4817
3	1.2097	0.7130	0.5189	1.0379
4	0.8280	0.5277	0.3595	0.7191
5	0.5672	0.3969	0.2580	0.5160
6	0.4116	0.3063	0.1923	0.3847
7	0.3118	0.2424	0.1482	0.2964
8	0.2442	0.1961	0.1173	0.2347
⋮				
12	0.1142	0.0988	0.1121	0.0560

TABLE 2. Stability parameters ε^1_η for Scheme 1 (cf. (3.9)) and $\tilde\varepsilon^2_\eta$ and ε^2_η for Scheme 2 (cf. (3.10) and (3.13)) for the weights w_i given in Table 1.

The basic scheme is given by

$$u^{n+1}_{ij} = u^n_{ij} - \lambda\Delta^+_x F\bigl(u^n_{i-1,j}, u^n_{ij}\bigr) - \lambda\Delta^+_y G\bigl(u^n_{i,j-1}, u^n_{ij}\bigr) + \mu\bigl(\Delta^2_x A\bigl(u^n_{ij}\bigr) + \Delta^2_y A\bigl(u^n_{ij}\bigr)\bigr). \tag{3.6}$$

Here, the numerical flux by Engquist & Osher (1981) is given by $F(u,v) = f^+(u) + f^-(v)$, where

$$f^+(u) := f(0) + \int_0^u \max\bigl\{f'(s), 0\bigr\}\,\mathrm{d}s, \quad f^-(u) := \int_0^u \min\bigl\{f'(s), 0\bigr\}\,\mathrm{d}s;$$

the function G is defined in the same way with f replaced by g.

The first mollified version of the basic scheme (3.6), denoted Scheme 1, is based on discretizing the terms $A(u)_{xx}$ and $A(u)_{yy}$ separately by applications of the one-dimensional discrete mollification operator J_η in x- and y-direction, respectively. Denoting the respective versions of J_η by J^x_η and J^y_η, we obtain the following numerical scheme:

$$\begin{aligned} u^{n+1}_{ij} &= u^n_{ij} - \lambda\Delta^+_x F\bigl(u^n_{i-1,j}, u^n_{ij}\bigr) - \lambda\Delta^+_y G\bigl(u^n_{i,j-1}, u^n_{ij}\bigr) \\ &\quad + 2\mu C_\eta\Bigl(\bigl[J^x_\eta A(u^n)\bigr]_{ij} + \bigl[J^y_\eta A(u^n)\bigr]_{ij} - 2A\bigl(u^n_{ij}\bigr)\Bigr). \end{aligned} \tag{3.7}$$

Alternatively, we may discretize $\Delta A(u)$ in terms of the two-dimensional discrete mollification operator J^2_η introduced in Section 3.2. The resulting mollified scheme, Scheme 2, has the form

$$u^{n+1}_{ij} = u^n_{ij} - \lambda\Delta^+_x F\bigl(u^n_{i-1,j}, u^n_{ij}\bigr) - \lambda\Delta^+_y G\bigl(u^n_{i,j-1}, u^n_{ij}\bigr) + 2\mu C_\eta\Bigl(\bigl[J^2_\eta A(u^n)\bigr]_{ij} - A\bigl(u^n_{ij}\bigr)\Bigr). \tag{3.8}$$

3.4. **CFL conditions and stabilization.** We briefly summarize the CFL conditions that will appear in our analysis. The basic scheme (3.6) results to be monotone under the CFL condition

$$\lambda\bigl(\|f'\|_\infty + \|g'\|_\infty\bigr) + 4\mu\|a\|_\infty \leq 1.$$

In the next section we will prove that the same property holds under the CFL condition

$$\lambda\bigl(\|f'\|_\infty + \|g'\|_\infty\bigr) + 4\varepsilon^1_\eta\mu\|a\|_\infty \leq 1, \quad \varepsilon^1_\eta := (1-w_0)\,C_\eta \tag{3.9}$$

for Scheme 1 and under the CFL condition

$$\lambda\bigl(\|f'\|_\infty + \|g'\|_\infty\bigr) + 4\tilde\varepsilon^2_\eta\mu\|a\|_\infty \leq 1, \quad \tilde\varepsilon^2_\eta := \frac{1-w_0^2}{2}C_\eta \tag{3.10}$$

for Scheme 2. Table 2 shows the values of ε^1_η and $\tilde\varepsilon^2_\eta$ for the weights given in Table 1 and selected values of η. We see that ε^1_η and $\tilde\varepsilon^2_\eta$ decrease with η, and that $\varepsilon^1_\eta < 1$ and $\tilde\varepsilon^2_\eta < 1$ for $\eta \geq 2$, so Schemes 1 and 2 are monotone for larger values of Δt for a given meshwidth Δx than is the basic scheme.

For the convergence analysis, however, we will need to impose a more restrictive CFL condition to ensure that the limit of approximate solutions satisfies item (3) of Definition 2.1. For the basic scheme, this condition, which we refer to as "strengthened CFL condition", is given by

$$8\lambda\big(\|f'\|_\infty+\|g'\|_\infty\big)+8\mu\|a\|_\infty\leq 1-\varepsilon \tag{3.11}$$

for a number $\varepsilon\in(0,1)$, see Karlsen & Risebro (2001). The corresponding strengthened CFL condition for Scheme 1 is given by

$$8\lambda\big(\|f'\|_\infty+\|g'\|_\infty\big)+8\varepsilon^1_\eta\mu\|a\|_\infty\leq 1-\varepsilon,\quad 0<\varepsilon<1, \tag{3.12}$$

while that for Scheme 2 is given by

$$8\lambda\big(\|f'\|_\infty+\|g'\|_\infty\big)+8\varepsilon^2_\eta\mu\|a\|_\infty\leq 1-\varepsilon,\quad 0<\varepsilon<1,\quad \varepsilon^2_\eta:=2\tilde{\varepsilon}^2_\eta. \tag{3.13}$$

From Table 2 we infer that for the weights of Table 1, only for $\eta\geq 4$ we have $\varepsilon^2_\eta<1$, and Scheme 2 thus has a favorable CFL condition compared with that of the basic scheme. However, for large values of η the coefficients ε^1_η and ε^2_η are nearly equal.

4. Convergence analysis

Lemma 4.1. *Scheme 1 given by* (3.7) *is monotone under the CFL condition* (3.9).

Proof. The proof is similar to the one of Lemma 3.2 by Acosta et al. (2011). We denote by u^n and v^n the respective data $\{u^n_{ij}\}_{i,j\in\mathbb{Z}}$ and $\{v^n_{ij}\}_{i,j\in\mathbb{Z}}$, and assume that $u^n_{ij}=v^n_{ij}$ for $i,j\in\mathbb{Z}$ with the exception of $i=k$, $j=l$, for which we assume that $u^n_{kl}\leq v^n_{kl}$. We rewrite Scheme 1, (3.7), as $u^{n+1}_{ij}=\mathcal{S}_{ij}(u^n)$, where $\mathcal{S}_{ij}(u^n)$ denotes the right-hand side of (3.7). Analogously, we define $v^{n+1}_{ij}=\mathcal{S}_{ij}(v^n)$. Clearly, $\mathcal{S}_{ij}(u^n)-\mathcal{S}_{ij}(v^n)=0$ if $k<i-\eta$ or $k>i+\eta$ and $l<j-\eta$ or $l>j+\eta$. Similarly, in the remaining cases combinations of arguments of the proof of Lemma 3.2 by Acosta et al. (2011) will be sufficient to establish that $\mathcal{S}_{ij}(u^n)-\mathcal{S}_{ij}(v^n)\leq 0$ provided that $i\neq k$ or $j\neq l$. It remains to deal with the case $i=k$, $j=l$. We then have

$$\begin{aligned}\mathcal{S}_{ij}(u^n)-\mathcal{S}_{ij}(v^n)&=u^n_{ij}-v^n_{ij}-\lambda\big(F\big(u^n_{ij},u^n_{i+1,j}\big)-F\big(v^n_{ij},u^n_{i+1,j}\big)\big)+\lambda\big(F\big(u^n_{i-1,j},u^n_{ij}\big)-F\big(u^n_{i-1,j},v^n_{ij}\big)\big)\\&\quad-\lambda\big(G\big(u^n_{ij},u^n_{i,j+1}\big)-G\big(v^n_{ij},u^n_{i,j+1}\big)\big)+\lambda\big(G\big(u^n_{i,j-1},u^n_{ij}\big)-G\big(u^n_{i,j-1},v^n_{ij}\big)\big)\\&\quad+2\mu C_\eta\Big(\big[J^x_\eta\big(A(u^n)-A(v^n)\big)\big]_{ij}+\big[J^y_\eta\big(A(u^n)-A(v^n)\big)\big]_{ij}-2\big(A\big(u^n_{ij}\big)-A\big(v^n_{ij}\big)\big)\Big).\end{aligned}$$

Considering that

$$\big[J^x_\eta\big(A(u^n)-A(v^n)\big)\big]_{ij}=\sum_{\nu=-\eta}^{\eta}w_\nu\big(A\big(u^n_{i+\nu,j}\big)-A\big(v^n_{i+\nu,j}\big)\big)=w_0\big(A\big(u^n_{ij}\big)-A\big(v^n_{ij}\big)\big)$$

(and analogously $[J^y_\eta(A(u^n)-A(v^n))]_{ij}=w_0(A(u^n_{ij})-A(v^n_{ij})))$ and that we have

$$\begin{aligned}F\big(u^n_{ij},u^n_{i+1,j}\big)-F\big(v^n_{ij},u^n_{i+1,j}\big)&=f^+\big(u^n_{ij}\big)-f^+\big(v^n_{ij}\big),\\F\big(u^n_{i-1,j},u^n_{ij}\big)-F\big(u^n_{i-1,j},v^n_{ij}\big)&=f^-\big(u^n_{ij}\big)-f^-\big(v^n_{ij}\big),\\G\big(u^n_{ij},u^n_{i,j+1}\big)-G\big(v^n_{ij},u^n_{i,j+1}\big)&=g^+\big(u^n_{ij}\big)-g^+\big(v^n_{ij}\big),\\G\big(u^n_{i,j-1},u^n_{ij}\big)-G\big(u^n_{i,j-1},v^n_{ij}\big)&=g^-\big(u^n_{ij}\big)-g^-\big(v^n_{ij}\big)\end{aligned} \tag{4.1}$$

by the definition of F and G, we obtain for $i=k$ and $j=l$

$$\begin{aligned}\mathcal{S}_{ij}(u^n)-\mathcal{S}_{ij}(v^n)&=u^n_{ij}-v^n_{ij}-\lambda\big(f^+\big(u^n_{ij}\big)-f^+\big(v^n_{ij}\big)\big)+\lambda\big(f^-\big(u^n_{ij}\big)-f^-\big(v^n_{ij}\big)\big)\\&\quad-\lambda\big(g^+\big(u^n_{ij}\big)-g^+\big(v^n_{ij}\big)\big)+\lambda\big(g^-\big(u^n_{ij}\big)-g^-\big(v^n_{ij}\big)\big)+4\mu C_\eta(w_0-1)\big(A\big(u^n_{ij}\big)-A\big(v^n_{ij}\big)\big)\\&=\int_{u^n_{ij}}^{v^n_{ij}}\Big\{-1+\lambda\max\big\{f'(s),0\big\}-\lambda\min\big\{f'(s),0\big\}+\lambda\max\big\{g'(s),0\big\}\\&\qquad-\lambda\min\big\{g'(s),0\big\}+4\mu C_\eta(1-w_0)a(s)\Big\}\,\mathrm{d}s\end{aligned}$$

$$\leq -\int_{u_{ij}^n}^{v_{ij}^n} \Big\{1 - \lambda\big|f'(s),0\big| - \lambda\big|g'(s),0\big| - 4\mu C_\eta(1-w_0)a(s)\Big\}\,\mathrm{d}s. \tag{4.2}$$

Under the CFL condition (3.9) the integrand in (4.2) is non-negative, so $\mathcal{S}_{ij}(u^n) \leq \mathcal{S}_{ij}(v^n)$ and Scheme 1, (3.7), is indeed montone. □

Lemma 4.2. *Scheme 2 defined by* (3.8) *is monotone under the CFL condition* (3.10).

Proof. The proof is similar to that of Lemma 4.1. We define u^n and v^n as in the proof of Lemma 4.1, and denote by $\tilde{\mathcal{S}}_{ij}(u^n)$ the right-hand side of (3.8), so that $u_{ij}^{n+1} = \tilde{\mathcal{S}}_{ij}(u^n)$ and $v_{ij}^{n+1} = \tilde{\mathcal{S}}_{ij}(v^n)$. By arguments similar to those of the proof of Lemma 4.1 one can straightforwardly show that $\tilde{\mathcal{S}}_{ij}(u^n) - \tilde{\mathcal{S}}_{ij}(v^n) \leq 0$ if $i \neq k$ or $j \neq l$. For $i=k$ and $j=l$ we obtain

$$\begin{aligned}\tilde{\mathcal{S}}_{ij}(u^n) - \tilde{\mathcal{S}}_{ij}(v^n) &= u_{ij}^n - v_{ij}^n - \lambda\big(F(u_{ij}^n, u_{i+1,j}^n) - F(v_{ij}^n, u_{i+1,j}^n)\big) + \lambda\big(F(u_{i-1,j}^n, u_{ij}^n) - F(u_{i-1,j}^n, v_{ij}^n)\big)\\ &\quad - \lambda\big(G(u_{ij}^n, u_{i,j+1}^n) - G(v_{ij}^n, u_{i,j+1}^n)\big) + \lambda\big(G(u_{i,j-1}^n, u_{ij}^n) - G(u_{i,j-1}^n, v_{ij}^n)\big)\\ &\quad + 2\mu C_\eta\Big(\big[J_\eta^2\big(A(u^n) - A(v^n)\big)\big]_{ij} - \big(A(u_{ij}^n) - A(v_{ij}^n)\big)\Big).\end{aligned}$$

Now, taking into account that

$$\big[J_\eta^2\big(A(u^n) - A(v^n)\big)\big]_{ij} = \sum_{\nu=-\eta}^{\eta}\sum_{\kappa=-\eta}^{\eta} w_\nu w_\kappa \big(A(u_{i+\nu,j+\kappa}^n) - A(v_{i+\nu,j+\kappa}^n)\big) = w_0^2\big(A(u_{ij}^n) - A(v_{ij}^n)\big)$$

and using again (4.1), we obtain for $i=k$ and $j=l$

$$\begin{aligned}\tilde{\mathcal{S}}_{ij}(u^n) - \tilde{\mathcal{S}}_{ij}(v^n) &= u_{ij}^n - v_{ij}^n - \lambda\big(f^+(u_{ij}^n) - f^+(v_{ij}^n)\big) + \lambda\big(f^-(u_{ij}^n) - f^-(v_{ij}^n)\big)\\ &\quad - \lambda\big(g^+(u_{ij}^n) - g^+(v_{ij}^n)\big) + \lambda\big(g^-(u_{ij}^n) - g^-(v_{ij}^n)\big) + 2\mu C_\eta\left(w_0^2 - 1\right)\left(A(u_{ij}^n) - A(v_{ij}^n)\right).\end{aligned}$$

The remainder of the proof is now analogous to that of Lemma 4.1. □

In what follows, we denote by u^Δ the piecewise constant function that satisfies

$$u^\Delta(x,y,t) = u_{ij}^n \quad \text{for } x_{i-1/2} \leq x < x_{i+1/2},\ y_{j-1/2} \leq y < y_{j+1/2},\ t_n \leq t < t_{n+1}. \tag{4.3}$$

Moreover, we define

$$F_{i-1/2,j}^n := F\big(u_{i-1,j}^n, u_{ij}^n\big), \quad G_{i,j-1/2}^n := G\big(u_{i,j-1}^n, u_{ij}^n\big), \quad A^n := A(u^n), \quad A_{ij}^n := A\big(u_{ij}^n\big).$$

A sum over "i,j" is understood as simultaneous summation over $i \in \mathbb{Z}$, $j \in \mathbb{Z}$, and a sum over "i,j,n" denotes summation over $i \in \mathbb{Z}$, $j \in \mathbb{Z}$ and $n = 0,\dots,N-1$. Furthermore, C denotes a constant whose meaning may change from line to line, but whose value is always independent of $\Delta = (\Delta x, \Delta t)$.

Lemma 4.3. *Under the respective CFL conditions* (3.9) *and* (3.10) *the approximate solutions generated by Schemes 1 and 2 satisfy the uniform L^∞ and L^1 bounds*

$$\|u^n\|_\infty \leq \|u_0\|_\infty \quad \text{for } n = 1,\dots,N, \tag{4.4}$$

$$\|u^n\|_1 \leq \|u_1\|_1 \quad \text{for } n = 1,\dots,N, \tag{4.5}$$

and the uniform total variation diminishing (TVD) property holds:

$$\sum_{i,j}\Big(\big|u_{i+1,j}^{n+1} - u_{ij}^{n+1}\big| + \big|u_{i,j+1}^{n+1} - u_{ij}^{n+1}\big|\Big) \leq \sum_{i,j}\Big(\big|u_{i+1,j}^{n} - u_{ij}^{n}\big| + \big|u_{i,j+1}^{n} - u_{ij}^{n}\big|\Big), \quad n = 0,\dots,N-1. \tag{4.6}$$

Proof. Inequality (4.4) follows from monotonicity by a standard argument if we take into account that if $w_{ij}^n = \|u_0\|_\infty$ for all $i,j \in \mathbb{Z}$ or $w_{ij}^n = -\|u_0\|_\infty$ for all $i,j \in \mathbb{Z}$ and $w_{ij}^{n+1} = \mathcal{S}_{ij}(w^n)$ or $w_{ij}^{n+1} = \tilde{\mathcal{S}}_{ij}(w^n)$, then $w_{ij}^{n+1} = \|u_0\|_\infty$ for all $i,j \in \mathbb{Z}$ or $w_{ij}^{n+1} = -\|u_0\|_\infty$ for all $i,j \in \mathbb{Z}$. Inequalities (4.5) and (4.6) are standard properties of monotone schemes, and are established by an application of the Crandall-Tartar lemma (see Karlsen & Risebro, 2001). □

Lemma 4.4. *For both Schemes 1 and 2 there exists a constant C, which is independent of Δ, such that*

$$\left\|u^{\Delta}(\cdot,t_1)-u^{\Delta}(\cdot,t_2)\right\|_{L^1(\mathbb{R}^2)} \leq C\sqrt{|t_1-t_2|}. \tag{4.7}$$

Proof. The proof is similar to that of Karlsen & Risebro (2001), and proceeds by appealing to the Kružkov interpolation lemma (Kružkov, 1969; see Lemma 2.4 by Karlsen & Risebro, 2001). The proof proceeds by selecting a test function $\varphi=\varphi(x,y)$, multiplying (3.7) (or (3.8)) by $\Delta x^2\varphi_{ij}=\varphi(x_i,y_j)$, summing the result over $i,j\in\mathbb{Z}$, and applying summation by parts. Taking into account that for Scheme 1, we have that

$$\begin{aligned}
&C_\eta\sum_{i,j}\Big(\big[J_\eta^x A^n\big]_{ij}+\big[J_\eta^y A^n\big]_{ij}-2A_{ij}^n\Big)\varphi_{ij}\\
&=C_\eta\sum_{i,j}\left(\sum_{k=1}^{\eta}w_k\Big(\big(A_{i+k,j}^n-2A_{ij}^n+A_{i-k,j}^n\big)+\big(A_{i,j+k}^n-2A_{ij}^n+A_{i,j-k}^n\big)\Big)\right)\varphi_{ij}\\
&=C_\eta\sum_{k=1}^{\eta}w_k\sum_{i,j}\Big(\big(A_{i+k,j}^n-A_{ij}^n\big)\big(\varphi_{ij}-\varphi_{i+k,j}\big)+\big(A_{i,j+k}^n-A_{ij}^n\big)\big(\varphi_{ij}-\varphi_{i,j+k}\big)\Big)\\
&=-C_\eta\sum_{k=1}^{\eta}w_k\sum_{i,j}\sum_{p,q=0}^{k-1}\Big(\big(\Delta_x^+A_{i+p,j}^n\big)\big(\Delta_x^+\varphi_{i+q,j}\big)+\big(\Delta_y^+A_{i,j+p}^n\big)\big(\Delta_y^+\varphi_{i,j+q}\big)\Big)\\
&=-C_\eta\sum_{k=1}^{\eta}k^2w_k\sum_{i,j}\Big(\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+\varphi_{ij}\big)+\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+\varphi_{ij}\big)\Big)\\
&=-\frac{1}{2}\sum_{i,j}\Big(\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+\varphi_{ij}\big)+\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+\varphi_{ij}\big)\Big),
\end{aligned}\tag{4.8}$$

we deduce that

$$2\mu C_\eta\Delta x^2\left|\sum_{i,j}\Big(\big[J_\eta^x A^n\big]_{ij}+\big[J_\eta^y A^n\big]_{ij}-2A_{ij}^n\Big)\varphi_{ij}\right|\leq\Delta t\Delta x\|\nabla\varphi\|_\infty\sum_{i,j}\Big(\big|\Delta_x^+A_{ij}^n\big|+\big|\Delta_y^+A_{ij}^n\big|\Big).$$

In a similar way, in light of (3.5) we obtain for Scheme 2, by a summation by parts and using (3.2),

$$\begin{aligned}
&C_\eta\sum_{i,j}\Big(\big[J_\eta^2A^n\big]_{ij}-A_{ij}^n\Big)\varphi_{ij}\\
&=-C_\eta\left(\sum_{k=-\eta}^{\eta}w_k\sum_{l=1}^{\eta}w_l\sum_{i,j}\big(A_{i+k,j+l}^n-A_{i+k,j}^n\big)(\varphi_{i+k,j}-\varphi_{ij})+\sum_{k=1}^{\eta}w_k\sum_{i,j}\big(A_{i+k,j}^n-A_{ij}^n\big)(\varphi_{i+k,j}-\varphi_{ij})\right)\\
&=-C_\eta\left(\sum_{k=-\eta}^{\eta}w_k\sum_{l=1}^{\eta}w_l\sum_{i,j}\sum_{p,q=0}^{l-1}\big(\Delta_y^+A_{i+k,j+p}^n\big)\big(\Delta_y^+\varphi_{i,j+q}\big)+\sum_{k=1}^{\eta}w_k\sum_{i,j}\sum_{r,s=0}^{k-1}\big(\Delta_x^+A_{i+r,j}^n\big)\big(\Delta_x^+\varphi_{i+s,j}\big)\right)\\
&=-C_\eta\left(\sum_{k=-\eta}^{\eta}w_k\sum_{l=1}^{\eta}l^2w_l\sum_{i,j}\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+\varphi_{ij}\big)+\sum_{k=1}^{\eta}k^2w_k\sum_{i,j\in\mathbb{Z}}\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+\varphi_{ij}\big)\right)\\
&=-\frac{1}{2}\sum_{i,j}\Big(\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+\varphi_{ij}\big)+\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+\varphi_{ij}\big)\Big),
\end{aligned}\tag{4.9}$$

which implies that

$$2\mu\Delta x^2C_\eta\left|\sum_{i,j}\Big(\big[J_\eta^2A^n\big]_{ij}-A_{ij}^n\Big)\varphi_{ij}\right|\leq\Delta t\Delta x\|\nabla\varphi\|_\infty\sum_{i,j}\Big(\big|\Delta_x^+A_{ij}^n\big|+\big|\Delta_y^+A_{ij}^n\big|\Big).$$

Consequently, for both schemes there exists a constant C, which is independent of Δ, such that

$$\sum_{i,j}\left|u_{ij}^{n+1}-u_{ij}^{n}\right|\varphi_{ij}\Delta x^2 \le C\Delta t\|\nabla\varphi\|_{L^\infty(\mathbb{R}^2)}\sum_{i,j}\Big(\Big(\left|F_{i-1/2,j}^n\right|+\left|G_{i,j-1/2}^n\right|\Big)\Delta x^2+\Big(\left|\Delta_x^+A_{ij}^n\right|+\left|\Delta_y^+A_{ij}^n\right|\Big)\Delta x\Big).$$

The proof of (4.7) can now be concluded exactly as in Karlsen & Risebro (2001) by an application of the Kružkov interpolation lemma, and appealing to the uniform L^1 and BV bounds (4.5) and (4.6). ☐

Lemma 4.5. *Assume that the CFL condition* (3.12) *is satisfied. Then for Scheme 1 there exists a constant C, which is independent of Δ, such that*

$$\sum_{i,j,n}\left[\left(\frac{\Delta_x^+A_{ij}^n}{\Delta x}\right)^2+\left(\frac{\Delta_y^+A_{ij}^n}{\Delta x}\right)^2\right]\Delta x^2\Delta t\le C. \tag{4.10}$$

For the proof of Lemma 4.4, we need the following result by Karlsen & Risebro (2001).

Lemma 4.6. *The following inequality holds:*

$$S:=\Delta x\Delta t\sum_{i,j,n}\big(\Delta_x^+F_{i-1/2,j}^n+\Delta_y^+G_{i,j-1/2}^n\big)u_{ij}^n\ge\frac{\Delta t\Delta x}{M_{f',g'}}\tilde{S}, \tag{4.11}$$

where $M_{f',g'}:=2\max\{\|f'\|_\infty,\|g'\|_\infty\}$ *and*

$$\begin{aligned}\tilde{S}:=\sum_{i,j,n}\Big(&\big(f^+(u_{ij}^n)-f^+(u_{i-1,j}^n)\big)^2+\big(f^-(u_{i+1,j}^n)-f^-(u_{ij}^n)\big)^2\\&+\big(g^+(u_{ij}^n)-g^+(u_{i,j-1}^n)\big)^2+\big(g^-(u_{i,j+1}^n)-g^-(u_{ij}^n)\big)^2\Big).\end{aligned}$$

The proof of Lemma 4.6 follows as a special case from the analysis of Section 4 of Karlsen & Risebro (2001). For sake of completeness we provide the proof in the Appendix.

Proof of Lemma 4.5. Noting that the one-dimensional discrete mollification operator J_η^x satisfies

$$[J_\eta^xA^n]_{ij}-A_{ij}^n=\sum_{k=-\eta}^{\eta}w_kA_{i+k,j}^n-A_{ij}^n=\sum_{k=1}^{\eta}w_k\big(A_{i+k,j}^n-2A_{ij}^n+A_{i-k,j}^n\big)$$

(with an analogous identity for J_η^y), we can rewrite the marching formula (3.7) as

$$\begin{aligned}&-2\mu C_\eta\sum_{k=1}^{\eta}w_k\Big(\big(A_{i+k,j}^n-2A_{ij}^n+A_{i-k,j}^n\big)+\big(A_{i,j+k}^n-2A_{ij}^n+A_{i,j-k}^n\big)\Big)\\&=u_{ij}^{n+1}-u_{ij}^n-\lambda\big(\Delta_x^+F_{i-1/2,j}^n+\Delta_y^+G_{i,j-1/2}^n\big).\end{aligned} \tag{4.12}$$

Next, we multiply (4.12) by $\Delta x^2u_{ij}^n$, sum the result over $i,j\in\mathbb{Z}$ and $n=0,\dots,N-1$, use that

$$\big(u_{ij}^{n+1}-u_{ij}^n\big)u_{ij}^n=\frac{1}{2}\Big(\big(u_{ij}^{n+1}\big)^2-\big(u_{ij}^n\big)^2-\big(u_{ij}^{n+1}-u_{ij}^n\big)^2\Big),$$

apply summation by parts and take into account (4.9) (with φ_{ij} replaced by u_{ij}^n) to obtain

$$\begin{aligned}&\sum_{i,j,n}\Big(\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+u_{ij}^n\big)+\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+u_{ij}^n\big)\Big)\\&=-\frac{\Delta x^2}{2}\sum_{i,j,n}\Big(\big(u_{ij}^{n+1}\big)^2-\big(u_{ij}^n\big)^2\Big)+\frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1}-u_{ij}^n\big)^2-S\\&=\frac{\Delta x^2}{2}\sum_{i,j\in\mathbb{Z}}\big(u_{ij}^0\big)^2+\frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1}-u_{ij}^n\big)^2-S.\end{aligned}$$

We define $a^*:=\|a\|_\infty$. Then we can write

$$\big(\Delta_x^+A_{ij}^n\big)\big(\Delta_x^+u_{ij}^n\big)\ge\frac{1}{a^*}\big(\Delta_x^+A_{ij}^n\big)^2,\quad\big(\Delta_y^+A_{ij}^n\big)\big(\Delta_y^+u_{ij}^n\big)\ge\frac{1}{a^*}\big(\Delta_y^+A_{ij}^n\big)^2. \tag{4.13}$$

Consequently, taking into account that $\mu\Delta x^2 = \Delta t$ and Lemma 4.6, we obtain the inequality

$$\frac{\Delta t}{a^*}\sum_{i,j,n}\Big(\big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big) + \frac{\Delta t\Delta x}{M_{f',g'}}\tilde{S} \leq \frac{\Delta x^2}{2}\sum_{i,j\in\mathbb{Z}}\big(u_{ij}^0\big)^2 + \frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1} - u_{ij}^n\big)^2. \tag{4.14}$$

On the other hand, noting that $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain from (3.7)

$$\begin{aligned}
&\frac{1}{2}\big(u_{ij}^{n+1} - u_{ij}^n\big)^2 - 2\lambda^2\big(\Delta_x^+ F_{i-1/2,j}^n + \Delta_y^+ G_{i,j-1/2}^n\big)^2 \\
&\leq 4\mu^2 C_\eta^2\left(\sum_{k=-\eta}^{\eta} w_k A_{i+k,j}^n - A_{ij}^n + \sum_{k=-\eta}^{\eta} w_k A_{i,j+k}^n - A_{ij}^n\right)^2 \\
&\leq 8\mu^2 C_\eta^2\left[\left(\sum_{k=-\eta}^{\eta} w_k A_{i+k,j}^n - A_{ij}^n\right)^2 + \left(\sum_{k=-\eta}^{\eta} w_k A_{i,j+k}^n - A_{ij}^n\right)^2\right] \\
&= 8\mu^2 C_\eta^2\left[\left(\sum_{k=1}^{\eta} w_k\Big[\big(A_{i+k,j}^n - A_{ij}^n\big) + \big(A_{i-k,j}^n - A_{ij}^n\big)\Big]\right)^2 + \left(\sum_{k=-\eta}^{\eta} w_k\Big[\big(A_{i,j+k}^n - A_{ij}^n\big) + \big(A_{i,j-k}^n - A_{ij}^n\big)\Big]\right)^2\right].
\end{aligned}$$

Applying the Cauchy-Schwarz inequality and noting that $w_1 + \cdots + w_\eta = (1-w_0)/2$, we get

$$\begin{aligned}
&\frac{1}{2}\big(u_{ij}^{n+1} - u_{ij}^n\big)^2 - 2\lambda^2\big(\Delta_x^+ F_{i-1/2,j}^n + \Delta_y^+ G_{i,j-1/2}^n\big)^2 \\
&\quad\leq 8\mu^2 C_\eta^2(1-w_0)\sum_{k=1}^{\eta} w_k\Big(\big(A_{i+k,j}^n - A_{ij}^n\big)^2 + \big(A_{i-k,j}^n - A_{ij}^n\big)^2 + \big(A_{i,j+k}^n - A_{ij}^n\big)^2 + \big(A_{i,j-k}^n - A_{ij}^n\big)^2\Big).
\end{aligned} \tag{4.15}$$

Next, we note that

$$\begin{aligned}
&16\mu^2\Delta x^2 C_\eta^2(1-w_0)\sum_{k=1}^{\eta} w_k\sum_{i,j,n}\Big(\big(A_{i+k,j}^n - A_{ij}^n\big)^2 + \big(A_{i,j+k}^n - A_{ij}^n\big)^2\Big) \\
&= 16\mu^2\Delta x^2 C_\eta^2(1-w_0)\sum_{k=1}^{\eta} w_k\sum_{i,j,n}\sum_{p,q=0}^{k-1}\Big(\big(\Delta_x^+ A_{i+p,j}^n\big)\big(\Delta_x^+ A_{i+q,j}^n\big) + \big(\Delta_y^+ A_{i,j+p}^n\big)\big(\Delta_y^+ A_{i,j+q}^n\big)\Big) \\
&= 16\mu^2\Delta x^2 C_\eta^2(1-w_0)\sum_{k=1}^{\eta} k^2 w_k\sum_{i,j,n}\Big(\big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big) \\
&= 8\mu^2\Delta x^2 C_\eta(1-w_0)\sum_{i,j,n}\Big(\big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big).
\end{aligned} \tag{4.16}$$

Thus, we obtain the following inequality by multiplying inequality (4.15) by Δx^2 and summing the result over $i,j\in\mathbb{Z}$ and $n = 0,\dots,N-1$:

$$\begin{aligned}
\frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1} - u_{ij}^n\big)^2 &\leq 2\Delta t^2\sum_{i,j,n}\big(\Delta_x^+ F_{i-1/2,j}^n + \Delta_y^+ G_{i,j-1/2}^n\big)^2 \\
&\qquad + 8\mu^2\Delta x^2 C_\eta(1-w_0)\sum_{i,j,n}\Big(\big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big).
\end{aligned}$$

Taking into account that by (A.1),

$$\sum_{i,j,n}\big(\Delta_x^+ F_{i-1/2,j}^n + \Delta_y^+ G_{i,j-1/2}^n\big)^2 \leq 4\tilde{S},$$

we arrive at

$$\frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1} - u_{ij}^n\big)^2 \leq 8\Delta t^2\tilde{S} + 8\mu^2\Delta x^2 C_\eta(1-w_0)\sum_{i,j,n}\Big(\big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big). \tag{4.17}$$

Adding (4.14) and (4.17) we obtain the inequality

$$\left(\frac{\Delta t}{a^*}-8\mu^2\Delta x^2 C_\eta(1-w_0)\right)\sum_{i,j,n}\Big(\big(\Delta_x^+A_{ij}^n\big)^2+\big(\Delta_y^+A_{ij}^n\big)^2\Big)+\left(\frac{\Delta t\Delta x}{M_{f',g'}}-8\Delta t^2\right)\tilde S\le\frac{\Delta x^2}{2}\sum_{i,j}\big(u_{ij}^0\big)^2:=\tilde C. \tag{4.18}$$

Now, if the strengthened CFL condition (3.12) is imposed instead of (3.9), we get

$$8\mu^2\Delta x^2 C_\eta(1-w_0)=8\mu C_\eta(1-w_0)\Delta t\le\frac{\Delta t(1-\varepsilon)}{a^*},\quad 8\Delta t\le\frac{(1-\varepsilon)\Delta x}{\|f'\|_\infty+\|g'\|_\infty}\le\frac{(1-\varepsilon)\Delta x}{M_{f',g'}}, \tag{4.19}$$

and therefore

$$\frac{\Delta t}{a^*}-8\mu\Delta x^2 C_\eta(1-w_0)\ge\frac{\varepsilon\Delta t}{a^*},\quad \frac{\Delta t\Delta x}{M_{f',g'}}-8\Delta t^2=\Delta t\left(\frac{\Delta x}{M_{f',g'}}-8\Delta t\right)\ge\frac{\varepsilon\Delta x\Delta t}{M_{f',g'}}>0. \tag{4.20}$$

Clearly, under the CFL condition (3.12) the coefficient of $\tilde S$ on the left-hand side of (4.18) is positive. Thus, we get from (4.18) the inequality

$$\frac{\varepsilon}{a^*}\Delta t\Delta x^2\sum_{i,j,n}\left[\left(\frac{\Delta_x^+A_{ij}^n}{\Delta x}\right)^2+\left(\frac{\Delta_y^+A_{ij}^n}{\Delta x}\right)^2\right]\le\tilde C,$$

from which we deduce that (4.10) holds with $C=\tilde C a^*/\varepsilon$. □

Lemma 4.7. *Assume that the CFL condition* (3.13) *is satisfied. Then inequality* (4.10) *is also valid for Scheme 2.*

Proof. The proof is similar to that of Lemma 4.4. In fact, we obtain by multiplying (3.8) by $\Delta x^2 u_{ij}^n$, summing the result over $i,j\in\mathbb{Z}$ and $n=0,\dots,N-1$, applying summation by parts (where we bear in mind (4.9) with φ_{ij} replaced by u_{ij}^n), using (4.13) and applying Lemma 4.6, that (4.14) is also valid for Scheme 2. On the other hand, a straightforward computation yields

$$\begin{aligned}\Big(2\mu C_\eta\Big(\big[J_\eta^2A^n\big]_{ij}-A_{ij}^n\Big)\Big)^2&=4\mu^2C_\eta^2\left(\sum_{k,l=-\eta}^{\eta}w_kw_l\big(A_{i+l,j+k}^n-A_{ij}^n\big)\right)^2\\&\le4\mu^2C_\eta^2\left(\sum_{k,l=-\eta}^{\eta}w_kw_l-w_0^2\right)\sum_{k,l=-\eta}^{\eta}w_kw_l\big(A_{i+l,j+k}^n-A_{ij}^n\big)^2\\&\le8\mu^2C_\eta^2(1-w_0^2)\sum_{k,l=-\eta}^{\eta}w_kw_l\Big(\big(A_{i+l,j+k}^n-A_{i+l,j}^n\big)^2+\big(A_{i+l,j}^n-A_{ij}^n\big)^2\Big),\end{aligned}$$

from which we deduce (by arguments similar to those used in (4.16)) that

$$\begin{aligned}&\sum_{i,j}\Big(2\mu C_\eta\Big(\big[J_\eta^2A^n\big]_{ij}-A_{ij}^n\Big)\Big)^2\\&\le8\mu^2C_\eta^2(1-w_0^2)\left(\sum_{k,l=-\eta}^{\eta}w_lw_k\sum_{i,j}\big(A_{i+l,j+k}^n-A_{i+l,j}^n\big)^2+\sum_{l=-\eta}^{\eta}w_l\sum_{i,j}\big(A_{i+l,j}-A_{ij}^n\big)^2\right)\\&=8\mu^2C_\eta^2(1-w_0^2)\left(2\sum_{l=-\eta}^{\eta}w_l\sum_{k=1}^{\eta}k^2w_k\sum_{i,j\in\mathbb{Z}}\big(\Delta_y^+A_{ij}^n\big)^2+2\sum_{l=1}^{\eta}l^2w_l\sum_{i,j}\big(\Delta_x^+A_{ij}^n\big)^2\right)\\&=8\mu^2C_\eta(1-w_0^2)\sum_{i,j}\Big(\big(\Delta_x^+A_{ij}^n\big)^2+\big(\Delta_y^+A_{ij}^n\big)^2\Big).\end{aligned}$$

Taking into account this inequality, we obtain from the inequality

$$\frac12\big(u_{ij}^{n+1}-u_{ij}^n\big)^2\le2\lambda^2\big(\Delta_x^+F_{i-1/2,j}^n+\Delta_y^+G_{i,j-1/2}^n\big)^2+\Big(2\mu C_\eta\Big(\big[J_\eta^2A^n\big]_{ij}-A_{ij}^n\Big)\Big)^2, \tag{4.21}$$

by multiplying (4.21) by Δx^2, summing the result over $i,j\in\mathbb{Z}$ and $n=0,\dots,N-1$ and dealing with the first term on the right-hand side exactly as in the proof of Lemma 4.4, the following analogue of (4.17):

$$\frac{\Delta x^2}{2}\sum_{i,j,n}\big(u_{ij}^{n+1}-u_{ij}^n\big)^2 \le 8\Delta^2\tilde{S}+8\mu^2C_\eta(1-w_0^2)\sum_{i,j,n}\Big(\big(\Delta_x^+A_{ij}^n\big)^2+\big(\Delta_y^+A_{ij}^n\big)^2\Big).$$

Adding (4.14) to this inequality we obtain

$$\left(\frac{\Delta t}{a^*}-8\mu^2\Delta x^2C_\eta(1-w_0^2)\right)\sum_{i,j,n}\Big(\big(\Delta_x^+A_{ij}^n\big)^2+\big(\Delta_y^+A_{ij}^n\big)^2\Big)+\left(\frac{\Delta t\Delta x}{M_{f',g'}}-8\Delta t^2\right)\tilde{S}\le\frac{\Delta x^2}{2}\sum_{i,j}\big(u_{ij}^0\big)^2=:\hat{C}.$$

Now let impose instead of (3.10) the strengthened condition (3.13). Then we have

$$\frac{\Delta t}{a^*}-8\mu\Delta x^2C_\eta(1-w_0^2)\ge\frac{\varepsilon\Delta t}{a^*},$$

while the second inequalities in (4.19) and (4.20) are valid as before. Consequently, (4.10) holds as in the proof of Lemma 4.4. □

Lemma 4.8. *We recall the standard notation $a\wedge b:=\min\{a,b\}$ and $a\vee b:=\max\{a,b\}$, and define the numerical entropy fluxes*

$$Q_f(k,u,v):=F(u\vee k,v\vee k)-F(u\wedge k,v\wedge k),\quad Q_g(k,u,v):=G(u\vee k,v\vee k)-G(u\wedge k,v\wedge k).$$

Then Scheme 1 satisfies the following cell entropy inequality:

$$\begin{aligned}\forall k\in\mathbb{R}:\quad \big|u_{ij}^{n+1}-k\big| &\le \big|u_{ij}^n-k\big|-\lambda\Delta_x^-Q_f\big(k,u_{ij}^n,u_{i+1,j}^n\big)-\lambda\Delta_y^-Q_g\big(k,u_{ij}^n,u_{i,j+1}^n\big)\\ &\quad+2\mu C_\eta\sum_{l=-\eta}^{\eta}w_l\Big(\big|A_{i+l,j}^n-A(k)\big|-\big|A_{ij}^n-A(k)\big|+\big|A_{i,j+l}^n-A(k)\big|-\big|A_{ij}^n-A(k)\big|\Big).\end{aligned}\tag{4.22}$$

Proof. Replacing every occurrence of u_{ij}^n in the definition of $\mathcal{S}_{ij}(u^n)$ (i.e., in the right-hand side of (3.7)) by $u_{ij}^n\vee k$, we obtain the following identity, where $u^n\vee k=\{u_{ij}^n\vee k\}_{i,j\in\mathbb{Z}}$:

$$\begin{aligned}\mathcal{S}_{ij}(u^n\vee k)&=u_{ij}^n\vee k-\lambda\Delta_x^-F\big(u_{ij}^n\vee k,u_{i+1,j}^n\vee k\big)-\lambda\Delta_y^-G\big(u_{ij}^n\vee k,u_{i,j+1}^n\vee k\big)\\ &\quad+2\mu C_\eta\left(\sum_{l=-\eta}^{\eta}w_l\big(A\big(u_{i+l,j}^n\vee k\big)-A\big(u_{ij}^n\vee k\big)\big)+\sum_{l=-\eta}^{\eta}w_l\big(A\big(u_{i,j+l}^n\vee k\big)-A\big(u_{ij}^n\vee k\big)\big)\right)\end{aligned}$$

The same identity holds if every "$\vee$" is replaced by "$\wedge$", and we define $u^n\wedge k=\{u_{ij}^n\wedge k\}_{i,j\in\mathbb{Z}}$. Subtracting the second version from the first, we get

$$\begin{aligned}\mathcal{S}_{ij}(u^n\vee k)-S_{ij}(u^n\wedge k)&=u_{ij}^n\vee k-u_{ij}^n\wedge k-\lambda\Delta_x^-Q_f\big(k,u_{ij}^n,u_{i+1,j}^n\big)-\lambda\Delta_y^-Q_g\big(k,u_{ij}^n,u_{i,j+1}^n\big)\\ &\quad+2\mu C_\eta\Bigg(\sum_{l=-\eta}^{\eta}w_l\Big(\big(A\big(u_{i+l,j}^n\vee k\big)-A\big(u_{i+l,j}^n\wedge k\big)\big)-\big(A\big(u_{ij}^n\vee k\big)-A\big(u_{ij}^n\wedge k\big)\big)\Big)\\ &\quad+\sum_{l=-\eta}^{\eta}w_l\Big(\big(A\big(u_{i,j+l}^n\vee k\big)-A\big(u_{i,j+l}^n\wedge k\big)\big)-\big(A\big(u_{ij}^n\vee k\big)-A\big(u_{ij}^n\wedge k\big)\big)\Big)\Bigg)\end{aligned}\tag{4.23}$$

Since A is nondecreasing, we can rewrite the right-hand side of (4.23) as that of (4.22). On the other hand, due to the monotonicity of the scheme we have that

$$\mathcal{S}_{ij}(u^n\vee k)-\mathcal{S}_{ij}(u^n\wedge k)\ge\mathcal{S}_{ij}(u^n)\vee k-\mathcal{S}_{ij}(u^n)\wedge k=\big|u_{ij}^n-k\big|.\tag{4.24}$$

Combining (4.23) and (4.24) we arrive at the desired entropy inequality (4.22). □

Lemma 4.9. *Scheme 2 satisfies the following cell entropy inequality:*

$$\begin{aligned}\forall k \in \mathbb{R}: \quad \big|u_{ij}^{n+1} - k\big| \le \big|u_{ij}^n - k\big| &- \lambda \Delta_x^- Q_f\big(k, u_{ij}^n, u_{i+1,j}^n\big) - \lambda \Delta_y^- Q_g\big(k, u_{ij}^n, u_{i,j+1}^n\big) \\ &+ 2\mu C_\eta \left(\sum_{p,q=-\eta}^{\eta} w_p w_q \big|A_{i+p,j+q}^n - A(k)\big| - \big|A_{ij}^n - A(k)\big| \right).\end{aligned} \tag{4.25}$$

Proof. The proof is similar to that of Lemma 4.8 if we take into account that

$$\begin{aligned}&\tilde{\mathcal{S}}_{ij}(u^n \vee k) - \tilde{\mathcal{S}}_{ij}(u^n \wedge k) \\ &= \big|u_{ij}^n - k\big| - \lambda \Delta_x^- Q_f\big(k, u_{ij}^n, u_{i+1,j}^n\big) - \lambda \Delta_y^- Q_g\big(k, u_{ij}^n, u_{i,j+1}^n\big) \\ &\quad + 2\mu C_\eta \left(\sum_{p,q=-\eta}^{\eta} w_p w_q \Big(A\big(u_{i+p,j+q}^n \vee k\big) - A\big(u_{i+p,j+q}^n \wedge k\big)\Big) - \Big(A\big(u_{ij}^n \vee k\big) - A\big(u_{ij}^n \wedge k\big)\Big) \right).\end{aligned}$$

□

Theorem 4.1. *Assume that the integrated diffusion coefficient A, the functions f and g, and the initial datum u_0 satisfy the respective assumptions* (2.1), (2.2), *and* (2.3). *Assume that for the cases of Scheme 1 and 2, Δt and Δx are always chosen such that the respective strengthened CFL condition* (3.12) *or* (3.13) *is satisfied. Then the piecewise constant approximate solutions* (4.3) *generated by Scheme 1 or 2 converge to the unique entropy solution of* (1.1), (1.2) *as $\Delta = (\Delta x, \Delta t) \downarrow 0$.*

Proof. From Lemma 4.3 we deduce that there exists a constant C_1, which does not depend on Δ, such that

$$\big\|u^\Delta(\cdot,t)\big\|_{L^1(\mathbb{R}^2)} \le C, \quad \big\|u^\Delta(\cdot,t)\big\|_{L^\infty(\mathbb{R}^2)} \le C, \quad \big|u^\Delta(\cdot,t)\big|_{BV(\mathbb{R}^2)} \le C. \tag{4.26}$$

In view of the uniform L^1 and BV estimates in (4.26) and (4.7) in Lemma 4.4 we may appeal to Lemma 2.1 to conclude that $\{u^\Delta\}_{\Delta>0}$ is compact in $L^1_{\mathrm{loc}}(\Pi_T)$, and any limit point u (obtained as $\Delta \downarrow 0$) will satisfy items (1) and (4) in Definition 2.1.

Next, we prove that u satisfies (2) in Definition 2.1, i.e., the entropy inequality (2.4). This can be done by a standard Lax-Wendroff-type argument; namely, we choose a nonnegative test function $\phi \in C^\infty(\Pi_T)$ with compact support on $\mathbb{R}^2 \times [0,T)$, multiply the discrete entropy inequalities for Schemes 1 and 2, (4.22) and (4.25), respectively, by $\Delta x^2 \phi_{ij}^n$, where $\phi_{ij}^n = \phi(x_i, y_i, t^n)$, sum the result over $i,j \in \mathbb{Z}$ and $n = 0, \dots, N-1$, apply summation by parts, and let $\Delta \downarrow 0$. Details will be omitted here (see, e.g., Acosta et al., 2011; Evje & Karlsen, 2000; Karlsen & Risebro, 2001), however, we will explicitly demonstrate how the summation by parts procedure works for the mollification-based discretizations of $\Delta A(u)$ for Schemes 1 and 2, respectively.

As a slight extension of the calculus of Acosta et al. (2011) we here obtain

$$\begin{aligned}&\Delta x^2 \Delta t \sum_{i,j,n} \frac{2C_\eta}{\Delta x^2} \left(\sum_{l=-\eta}^{\eta} w_l \Big(\big|A_{i+l,j}^n - A(k)\big| - \big|A_{ij}^n - A(k)\big| + \big|A_{i,j+l}^n - A(k)\big| - \big|A_{ij}^n - A(k)\big| \Big) \right) \phi_{ij}^n \\ &= \Delta x^2 \Delta t \sum_{i,j,n} \big|A_{ij}^n - A(k)\big| \frac{2C_\eta}{\Delta x^2} \left(\sum_{l=-\eta}^{\eta} w_l \phi_{i-l,j}^n - \phi_{ij}^n + \sum_{l=-\eta}^{\eta} w_l \phi_{i,j-l}^n - \phi_{ij}^n \right) \\ &= \Delta x^2 \Delta t \sum_{i,j,n} \big|A_{ij}^n - A(k)\big| \big(\phi_{xx}(x_i, y_j, t^n) + \phi_{yy}(x_i, y_j, t^n) \big) + \mathcal{O}(\Delta x^2),\end{aligned} \tag{4.27}$$

while the analogous result for Scheme 2 is

$$
\begin{aligned}
&\Delta x^2 \Delta t \sum_{n=0}^{N-1} \frac{2C_\eta}{\Delta x^2} \sum_{i,j} \phi_{ij}^n \sum_{p,q=-\eta}^{\eta} w_p w_q \Big(\big|A_{i+p,j+q}^n - A(k)\big| - \big|A_{ij}^n - A(k)\big| \Big) \\
&= \Delta x^2 \Delta t \sum_{n=0}^{N-1} \frac{2C_\eta}{\Delta x^2} \sum_{p,q=-\eta}^{\eta} w_p w_q \sum_{i,j} \Big(\big|A_{i+p,j+q}^n - A(k)\big| - \big|A_{ij}^n - A(k)\big| \Big) \phi_{ij}^n \\
&= \Delta x^2 \Delta t \sum_{n=0}^{N-1} \frac{2C_\eta}{\Delta x^2} \sum_{p,q=-\eta}^{\eta} w_p w_q \sum_{i,j} \big|A_{ij}^n - A(k)\big| \big(\phi_{i-p,j-q}^n - \phi_{ij}^n \big) \\
&= \Delta x^2 \Delta t \sum_{i,j,n} \big|A_{ij}^n - A(k)\big| \left(\frac{2C_\eta}{\Delta x^2} \sum_{p,q=-\eta}^{\eta} w_p w_q \big(\phi_{i+p,j+q}^n - \phi_{ij}^n \big) \right) \\
&= \Delta x^2 \Delta t \sum_{i,j,n} \big|A_{ij}^n - A(k)\big| \frac{2C_\eta}{\Delta x^2} \Big(\big[J_\eta^2 \phi(\cdot,\cdot,t^n) \big]_{ij} - \phi(x_i, y_j, t^n) \Big) \\
&= \Delta x^2 \Delta t \sum_{i,j,n} \big|A_{ij}^n - A(k)\big| \big(\phi_{xx}(x_i, y_j, t^n) + \phi_{yy}(x_i, y_j, t^n) \big) + \mathcal{O}(\Delta x^2).
\end{aligned}
\tag{4.28}
$$

In both calculations, (4.27) and (4.28), we use that $w_p = w_{-p}$ for $p = 0, \dots, \eta$, and that ϕ is smooth, so (3.4) holds for $G = \phi$. Clearly, both expressions duely converge to $\iint_{\Pi_T} |A(u) - A(k)| \Delta\phi \, \mathrm{d}t \, \mathrm{d}x \, \mathrm{d}y$ as $\Delta \downarrow 0$.

It remains to prove that a limit u of $\{u^\Delta\}_{\Delta>0}$ satisfies item (3) of Definition 2.1. This will be done by deriving a weak BV estimate (see Champier et al., 1993; Eymard et al., 1998a, 1998b; Afif & Amaziane, 2002). To this end, we may directly follow the analysis by Karlsen & Risebro (2001) to deduce from the weak space estimate (4.10) that there exists a constant $C > 0$, which is independent of Δ, such that

$$
\big\| A\big(u^\Delta(\cdot + y, \cdot)\big) - A\big(u^\Delta(\cdot,\cdot)\big) \big\|_{L^2(\Pi_T)} \leq C\big(|y| + h\big).
$$

The weak space estimate and the difference schemes themeselves can be employed to prove that $A(u^\Delta)$ is also L^2 continuous in time. To this end, let $n(t) \in \mathbb{N}_0$ be chosen such that $t \in [t_{n(t)}, t_{n(t)+1})$. Then

$$
\iint_{\Pi_{T-\tau}} \Big(A\big(u^\Delta(x, t+\tau)\big) - A\big(u^\Delta(x,t)\big) \Big)^2 \mathrm{d}t \, \mathrm{d}x \leq \|a\|_\infty \int_0^{T-\tau} B(t) \, \mathrm{d}t,
$$

where

$$
B(t) := \Delta x^2 \sum_{i,j} \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) \big(u_{ij}^{n(t+\tau)} - u_{ij}^{n(t)} \big) = \sum_{n=n(t)}^{n(t)+n(\tau)-1} Q_n(t),
$$

where we define

$$
Q_n(t) := \Delta x^2 \sum_{i,j} \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) \big(u_{ij}^{n+1} - u_{ij}^n \big).
$$

Consider now Scheme 1. Then $Q_n(t) = \tilde{Q}_n(t) + Q_n^1(t)$, where

$$
\begin{aligned}
\tilde{Q}_n(t) &= -\Delta x \Delta t \sum_{i,j} \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) \big(\Delta_x^+ F_{i-1/2,j}^n + \Delta_y^+ G_{i,j-1/2}^n \big), \\
Q_n^1(t) &= 2\Delta t C_\eta \sum_{i,j} \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) \big([J_\eta^x A^n]_{ij} + [J_\eta^y A^n]_{ij} - 2A_{ij}^n \big).
\end{aligned}
$$

By summation by parts we obtain

$$
\tilde{Q}_n(t) = \Delta x \Delta t \sum_{i,j} \Big(\Delta_x^+ \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) F_{i-1/2,j}^n + \Delta_y^+ \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)} \big) G_{i,j-1/2}^n \Big).
$$

Consequently,

$$\begin{aligned}\big|\tilde{Q}_n(t)\big| &\le \Delta x^2 \Delta t \sum_{i,j}\Big(\big(F^n_{i-1/2,j}\big)^2 + \big(G^n_{i,j-1/2}\big)^2\Big)\\ &\quad + \frac{\Delta t}{2}\sum_{i,j}\Big(\big(\Delta_x^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_x^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t)}\big)^2\Big),\end{aligned}$$

so there exists a constant C_1, which is independent of Δ, such that

$$\big|\tilde{Q}_n(t)\big| \le C_1\Delta t + \frac{\Delta t}{2}\sum_{i,j}\Big(\big(\Delta_x^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_x^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t)}\big)^2\Big).$$

On the other hand, following (4.8) with φ_{ij} replaced by $A_{ij}^{n(t)}$ and $A_{ij}^{n(t+\tau)}$ we get

$$\big|Q_n^1(t)\big| = \Delta t\left|\sum_{i,j}\Big(\big(\Delta_x^+ A_{ij}^n\big)\big(\Delta_x^+ A_{ij}^{n(t+\tau)} - \Delta_x^+ A_{ij}^{n(t)}\big) + \big(\Delta_y^+ A_{ij}^n\big)\big(\Delta_y^+ A_{ij}^{n(t+\tau)} - \Delta_y^+ A_{ij}^{n(t)}\big)\Big)\right|,$$

which implies that

$$\big|Q_n^1(t)\big| \le \frac{\Delta t}{2}\sum_{i,j}\Big(\big(\Delta_x^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_x^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t+\tau)}\big)^2 + 2\big(\Delta_x^+ A_{ij}^n\big)^2 + 2\big(\Delta_y^+ A_{ij}^n\big)^2\Big) \tag{4.29}$$

and therefore

$$\begin{aligned}\big|\tilde{Q}_n(t)\big| + \big|Q_n^1(t)\big| &\le C_1\Delta t + \Delta t\sum_{i,j}\Big(\big(\Delta_x^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_x^+ A_{ij}^{n(t+\tau)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t)}\big)^2 + \big(\Delta_y^+ A_{ij}^{n(t+\tau)}\big)^2\\ &\quad + \big(\Delta_x^+ A_{ij}^n\big)^2 + \big(\Delta_y^+ A_{ij}^n\big)^2\Big).\end{aligned} \tag{4.30}$$

Consequently,

$$\int_0^{T-\tau} B(t)\,\mathrm{d}t \le (T-\tau)C_1 n(\tau)\Delta t + \int_0^{T-\tau}\big(B_1(t) + \cdots + B_6(t)\big)\,\mathrm{d}t,$$

where we define

$$B_1(t) := \sum_{n=n(t)}^{n(t)+n(\tau)-1}\Delta t\sum_{i,j}\big(\Delta_x^+ A_{ij}^{n(t)}\big)^2,$$

and similarly $B_2(t),\dots,B_6(t)$ according to the summands in (4.30). Then, in view of (4.10), we get

$$\begin{aligned}\int_0^{T-\tau} B_1(t)\,\mathrm{d}t &= \int_0^{T-\tau}\left(\Delta t\sum_{i,j}\sum_{n=n(t)}^{n(t)+n(\tau)-1}\big(\Delta_x^+ A_{ij}^{n(t)}\big)^2\right)\mathrm{d}t = \sum_{m=0}^{N-n(\tau)}\Delta t^2\sum_{i,j}\sum_{n=n(t_m)}^{n(t_m)+n(\tau)-1}\big(\Delta_x^+ A_{ij}^{n(t_m)}\big)^2\\ &\le n(\tau)\Delta t\left(\Delta t\Delta x^2\sum_{i,j,n}\left(\frac{\Delta_x^+ A_{ij}^n}{\Delta x}\right)^2\right) \le C(\tau + \Delta t),\end{aligned} \tag{4.31}$$

where C does not depend on Δt. Analogous inequalities hold for B_2, B_3 and B_4. On the other hand,

$$\int_0^{T-\tau} B_5(t)\,\mathrm{d}t \le \sum_{m=0}^{N-n(\tau)}\Delta t^2\sum_{n=m}^{m+n(\tau)-1}\sum_{i,j}\big(\Delta_x^+ A_{ij}^n\big)^2 \le \Delta t\sum_{k=0}^{n(\tau)-1}\left(\Delta t\Delta x^2\sum_{i,j,n}\left(\frac{\Delta_x^+ A_{ij}^n}{\Delta x}\right)^2\right) \le C(\tau+\Delta t). \tag{4.32}$$

An analogous bound holds for the corresponding integral over B_6. In view of the bounds (4.31) and (4.32), we deduce that

$$\int_0^{T-\tau} B(t)\,\mathrm{d}t \leq C(\Delta t + \tau),$$

and therefore the following bound holds for approximate solutions u^Δ generated by Scheme 1:

$$\big\| A\big(u^\Delta(\cdot,\cdot+\tau)\big) - A\big(u^\Delta(\cdot,\cdot)\big)\big\|_{L^2(\mathbb{R}^2\times(0,T-\tau))} \leq C\sqrt{\Delta t + \tau}. \tag{4.33}$$

To deal with Scheme 2, note that in this case $Q_n(t) = \tilde{Q}_n(t) + Q_n^2(t)$, where $\tilde{Q}_n(t)$ is defined above, and

$$Q_n^2(t) := 2\Delta t C_\eta \sum_{i,j} \big(A_{ij}^{n(t+\tau)} - A_{ij}^{n(t)}\big)\Big(\big[J_\eta^2 A^n\big]_{ij} - A_{ij}^n\Big).$$

Following (4.9) with φ_{ij} replaced by $A_{ij}^{n(t)}$ and $A_{ij}^{n(t+\tau)}$ we obtain that (4.29) remains valid with $Q_n^1(t)$ replaced by $Q_n^2(t)$, so following exactly the same steps as for Scheme 1 we may deduce that (4.33) also holds for Scheme 2. In view of Lemma 2.2 we conclude that for both schemes, $A(u^\Delta)$ converges strongly in $L^2_{\mathrm{loc}}(\mathbb{R}^2\times(0,T))$ to a limit $\bar{A}$ as $\Delta \downarrow 0$, with $\bar{A} \in L^2(0,T;H^1(\mathbb{R}))$. In light of the strong convergence $u^\Delta \to u$ a.e., we conclude that $\bar{A} = A(u)$, so the limit function u also satisfies (3) of Definition 2.1. □

5. Numerical examples

In this section we include a well selected collection of numerical examples. The purpose here is to evaluate the performance of Scheme 1, (3.7), and Scheme 2, (3.8), versus the basic scheme (3.6). We employ the weights given in Table 2, and select the values $\eta = 3$, $\eta = 8$ and $\eta = 12$ for comparison. Examples 1 and 2 are two-dimensional linear convection-diffusion problems with known exact solution and are used for accuracy and order comparisons, these examples were taken from Cecchi & Pirozzi (2005). Next we present two nonlinear examples, Examples 3 and 4, which were suggested by Karlsen & Risebro (1997) as test for multidimensional nonlinear convection-diffusion problems. Finally, Example 5 for a strongly degenerate parabolic equation was taken from Holden et al. (2000).

In all examples, the time step Δt is chosen for each of the basic scheme and Schemes 1 and 2 such that the respective strengthened CFL condition, (3.11), (3.12) or (3.13), is satisfied with $1-\varepsilon = 0.98$. The plots and tables of errors and illustrations of numerical solutions are taken at the respective final time $t = T$. A discussion of the numerical results will be provided in Section 6.

5.1. **Examples 1 and 2: accuracy and order.** In Examples 1 and 2 we consider the two dimensional linear convection-diffusion equation

$$u_t + u_x + u_y = \varepsilon_0\left(u_{xx} + u_{yy}\right). \tag{5.1}$$

In Example 1 we consider this equation for $(x,y) \in [0,2\pi]^2$ and $0 < t \leq T = 1$ along with the initial datum

$$u(x,y,0) = u_0(x,y) = \sin x \sin y.$$

This problem has the exact solution

$$u(x,y,t) = \exp\left(-2\varepsilon_0 t\right)\sin\left(x-t\right)\sin\left(y-t\right).$$

Here, we use the parameter value $\varepsilon_0 = 1/2$. The results are summarized in Figure 1 (a), while Table 3 provides details of the error history for $\eta = 8$.

In Example 2 we consider $(x,y) \in [0,3]^2$ for $0 < t \leq T = 1$ and the initial datum

$$u(x,y,0) = u_0(x,y) = \exp\left(-\frac{(x-1)^2 + (y-1)^2}{4\varepsilon_0}\right). \tag{5.2}$$

The exact solution of the initial value problem (5.1), (5.2) is given by

$$u(x,y,t) = \frac{1}{1+t}\exp\biggl(-\frac{\left(x-(1+t)\right)^2 + \left(y-(1+t)\right)^2}{4\varepsilon_0\left(1+t\right)}\biggr).$$

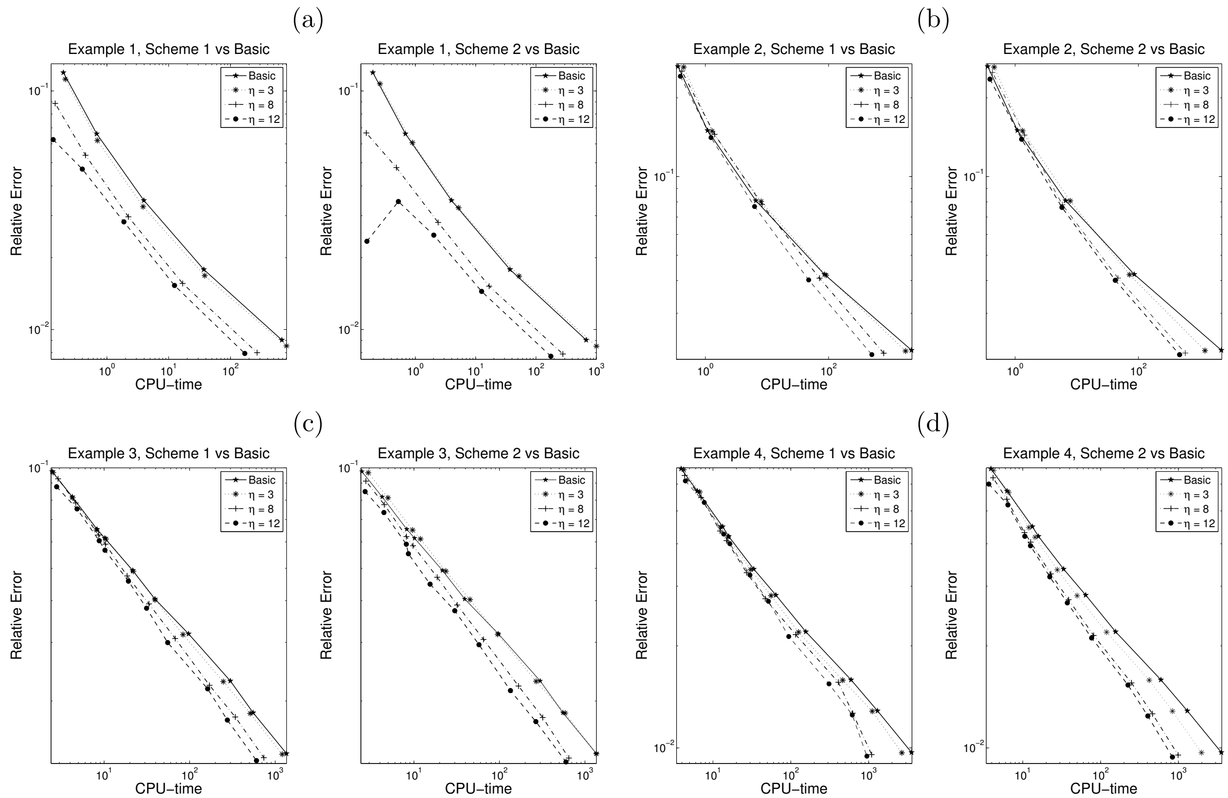

FIGURE 1. Examples 1 (a), 2 (b), 3 (c) and 4 (d): CPU time versus L^1 error, corresponding to the numerical solution at final time $t = T$ in each example.

	Scheme 1, $\eta = 8$			Scheme 2, $\eta = 8$			Basic scheme		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
$\pi/16$	0.0886	–	0.145	0.0666	–	0.153	0.1194	–	0.197
$\pi/32$	0.0538	0.7208	0.451	0.0477	0.4822	0.485	0.0661	0.8523	0.684
$\pi/64$	0.0297	0.8568	2.224	0.0281	0.7643	2.387	0.0348	0.9274	3.963
$\pi/128$	0.0156	0.9281	16.83	0.0152	0.8861	16.73	0.0178	0.9625	37.09
$\pi/256$	0.0080	0.9637	270.1	0.0079	0.9443	279.0	0.0090	0.9813	679.6

TABLE 3. Example 1: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]).

We here employ the parameter value $\varepsilon_0 = 1/16$. See Figure 1 (b) for a summary of the numerical results, and Table 4 for details on the error history for $\eta = 12$.

5.2. **Examples 3 and 4: nonlinear, non-degenerate convection-diffusion problems.** In Example 3 we focus on the nonlinear problem

$$u_t + \left(u + (u-0.25)^3\right)_x - \left(u + u^2\right)_y = \varepsilon_0 \left(u_{xx} + u_{yy}\right), \quad (x,y) \in [-2,4]^2, \quad 0 < t \le T = 0.5,$$

$$u_0(x,y) = \begin{cases} 1 & \text{for } (x-0.25)^2 + (y-0.25)^2 < 5, \\ 0 & \text{otherwise.} \end{cases}$$

	Scheme 1, $\eta = 12$			Scheme 2, $\eta = 12$			Basic scheme		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/16	0.2418	–	0.416	0.2358	–	0.428	0.2636	–	0.359
1/32	0.1409	0.7789	1.411	0.1388	0.7640	1.419	0.1500	0.8130	1.088
1/64	0.0768	0.8757	8.305	0.0762	0.8656	5.845	0.0809	0.8916	6.610
1/128	0.0403	0.9319	71.07	0.0401	0.9255	46.65	0.0423	0.9366	85.65
1/256	0.0209	0.9486	784.3	0.0208	0.9448	575.8	0.0217	0.9633	2217

TABLE 4. Example 2: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]).

	Scheme 1, $\eta = 3$			Scheme 2, $\eta = 3$			Basic scheme		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/20	0.0968	–	2.537	0.0967	–	2.925	0.0975	–	2.464
1/24	0.0813	0.9556	4.283	0.0813	0.9514	4.975	0.0818	0.9605	4.253
1/30	0.0651	0.9989	8.119	0.0651	0.9962	9.640	0.0655	0.9993	8.263
1/32	0.0612	0.9589	10.52	0.0612	0.9568	11.92	0.0616	0.9552	10.17
1/40	0.0490	0.9984	21.99	0.0490	0.9956	23.62	0.0492	1.0000	21.57
1/48	0.0402	1.0795	39.35	0.0403	1.0773	45.24	0.0404	1.0819	39.31
1/60	0.0316	1.0754	83.72	0.0317	1.0733	95.39	0.0318	1.0757	97.21
1/80	0.0229	1.1282	247.4	0.0229	1.1267	268.3	0.0230	1.1290	299.2
1/96	0.0184	1.2045	510.5	0.0184	1.2035	585.1	0.0184	1.2058	552.4
1/120	0.0139	1.2605	1209	0.0139	1.2596	1358	0.0139	1.2611	1355

TABLE 5. Example 3: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]).

	Scheme 1, $\eta = 8$			Scheme 2, $\eta = 8$			Scheme 1, $\eta = 12$			Scheme 2, $\eta = 12$		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/20	0.0927	–	2.892	0.0913	–	2.745	0.0878	–	2.786	0.0849	–	2.687
1/24	0.0785	0.9113	4.649	0.0776	0.8932	4.510	0.0753	0.8399	4.820	0.0735	0.7941	4.455
1/30	0.0629	0.9916	8.632	0.0622	0.9865	8.174	0.0605	0.9845	8.762	0.0590	0.9792	8.154
1/32	0.0591	0.9789	10.36	0.0584	0.9818	9.766	0.0567	1.0098	10.23	0.0553	1.0233	8.611
1/40	0.0474	0.9832	18.59	0.0470	0.9744	18.68	0.0457	0.9579	19.24	0.0448	0.9400	15.43
1/48	0.0391	1.0587	33.41	0.0388	1.0491	32.03	0.0380	1.0247	31.32	0.0373	1.0054	30.02
1/60	0.0308	1.0714	67.75	0.0306	1.0661	65.07	0.0299	1.0620	55.36	0.0295	1.0525	57.65
1/80	0.0223	1.1225	170.8	0.0222	1.1171	167.1	0.0218	1.1099	162.3	0.0215	1.1001	134.5
1/96	0.0179	1.1975	341.5	0.0179	1.1923	320.1	0.0175	1.1837	276.3	0.0174	1.1740	266.2
1/120	0.0135	1.2594	736.5	0.0135	1.2553	647.0	0.0133	1.2539	605.2	0.0131	1.2469	599.3

TABLE 6. Example 3: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]) (continued). The corresponding results for the basic scheme are given in Table 5.

Here and in Examples 4 and 5, the reference solution is computed by the basic scheme (3.6) using $\Delta x = 1/480$. A sample numerical solution of the problem, together with the reference solution, is shown in Figure 2 (a). The results concerning convergence are summarized in Figure 1 (c). Moreover, for the purpose of comparison we select this example for a record of detailed information on the convergence history for all three parameters, $\eta = 3$, $\eta = 8$ and $\eta = 12$. See Tables 5 and 6.

In Example 4 we consider the Buckley-Leverett-type problem

$$u_t + f(u)_x + g(u)_y = \varepsilon_0 (u_{xx} + u_{yy}), \quad (x, y) \in [-3, 3]^2, \quad 0 < t \leq T = 1,$$

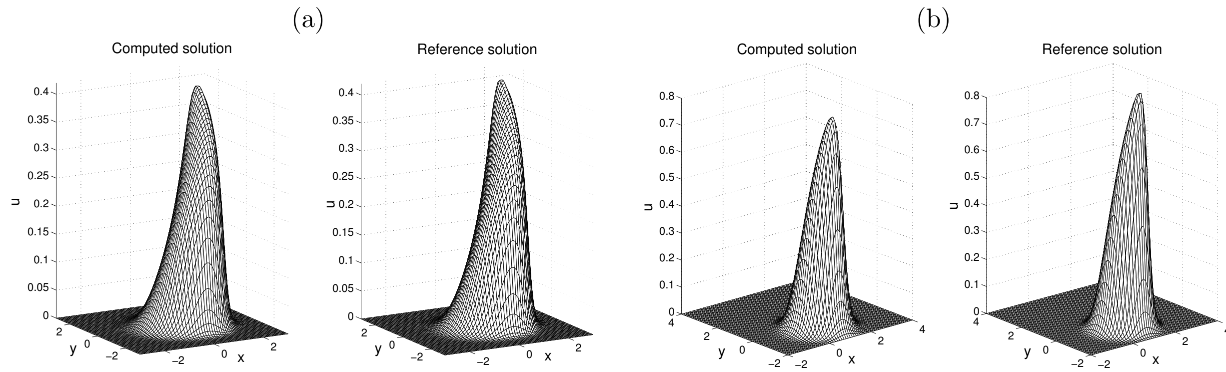

FIGURE 2. Example 3 (a): (left) numerical solution computed by Scheme 1, (3.7), with $\eta = 8$ and $\Delta x = 1/20$, (right) reference solution; Example 4 (b): (left) numerical solution computed by Scheme 2, (3.8), with $\eta = 8$ and $\Delta x = 1/40$, (right) reference solution. In both examples, the reference solution is computed by the basic scheme with $\Delta x = 1/480$.

	Scheme 1, $\eta = 8$			Scheme 2, $\eta = 8$			Basic scheme		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/20	0.0636	–	4.317	0.0626	–	4.155	0.0666	–	3.905
1/24	0.0548	0.8223	7.031	0.0541	0.8025	6.244	0.0572	0.8297	6.284
1/30	0.0436	1.0181	12.32	0.0432	1.0124	10.51	0.0450	1.0833	13.30
1/32	0.0409	1.0204	14.95	0.0405	1.0055	12.64	0.0421	1.0211	15.92
1/40	0.0329	0.9730	27.13	0.0326	0.9671	22.97	0.0337	0.9970	33.46
1/48	0.0275	0.9707	47.08	0.0274	0.9628	38.85	0.0282	0.9683	64.39
1/60	0.0216	1.0918	115.2	0.0215	1.0872	81.40	0.0220	1.1143	155.8
1/80	0.0156	1.1251	408.2	0.0156	1.1199	249.3	0.0159	1.1357	595.1
1/96	0.0127	1.1545	624.4	0.0126	1.1507	462.9	0.0129	1.1568	1301
1/120	0.0096	1.2597	1093	0.0095	1.2564	990.8	0.0097	1.2656	3555

TABLE 7. Example 4: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]).

where we choose the functions

$$g(u) = \frac{u^2}{u^2 + (1-u)^2}, \quad f(u) = g(u)\bigl(1 - 5(1-u)^2\bigr),$$

the parameter $\varepsilon_0 = 0.1$, and the initial datum

$$u(x,y,0) = u_0(x,y) = \begin{cases} 1 & \text{for } (x-0.25)^2 + (y-0.25)^2 < 5, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 2 (b) shows a sample numerical solution of the problem, together with the reference solution. The convergence history is summarized in Figure 1 (d), and details for $\eta = 8$ are provided in Table 7.

5.3. **Example 5: strongly degenerate parabolic problem.** Consider the strongly degenerate parabolic Burgers-like equation

$$u_t + (u^2)_x + (u^2)_y = \Delta A(u), \quad (x,y) \in [-1.5, 1.5]^2, \quad 0 < t \le T = 0.5,$$

	Scheme 1, $\eta = 3$			Scheme 2, $\eta = 3$			Basic scheme		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/20	0.0958	–	0.774	0.0986	–	0.864	0.0901	–	0.720
1/24	0.0750	1.3421	1.041	0.0772	1.3376	1.200	0.0718	1.2466	0.998
1/30	0.0673	0.4831	1.632	0.0695	0.4745	1.851	0.0630	0.5850	1.562
1/32	0.0602	1.7301	1.885	0.0622	1.7089	2.118	0.0564	1.7264	1.781
1/40	0.0505	0.7912	3.117	0.0520	0.8061	3.567	0.0475	0.7670	2.988
1/48	0.0418	1.0299	4.927	0.0430	1.0333	5.752	0.0393	1.0359	5.055
1/60	0.0341	0.9107	9.245	0.0352	0.9033	10.86	0.0319	0.9423	11.92
1/80	0.0251	1.0645	21.55	0.0259	1.0691	26.16	0.0234	1.0684	31.71
1/96	0.0203	1.1833	40.31	0.0209	1.1789	55.94	0.0189	1.1737	65.75
1/120	0.0162	0.9982	89.49	0.0167	0.9963	124.2	0.0150	1.0387	163.2

TABLE 8. Example 5: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]).

	Scheme 1, $\eta = 8$			Scheme 2, $\eta = 8$			Scheme 1, $\eta = 12$			Scheme 2, $\eta = 12$		
Δx	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]	L^1-re	L^1-cr	CPU [s]
1/20	0.1210	–	0.685	0.1343	–	0.677	0.1479	–	0.694	0.1723	–	0.671
1/24	0.0948	1.3376	0.924	0.1052	1.3383	0.918	0.1170	1.2824	0.957	0.1376	1.2357	0.905
1/30	0.0858	0.4505	1.411	0.0939	0.5109	1.428	0.1048	0.4940	1.494	0.1204	0.5953	1.404
1/32	0.0772	1.6220	1.594	0.0846	1.6115	1.602	0.0948	1.5501	1.707	0.1092	1.5233	1.570
1/40	0.0647	0.7972	2.607	0.0702	0.8378	2.602	0.0786	0.8399	2.765	0.0892	0.9036	2.461
1/48	0.0538	1.0078	4.231	0.0583	1.0182	3.945	0.0654	1.0086	4.295	0.0740	1.0235	3.561
1/60	0.0445	0.8505	7.231	0.0481	0.8595	7.275	0.0542	0.8410	7.668	0.0607	0.8879	6.287
1/80	0.0330	1.0375	15.85	0.0357	1.0410	14.53	0.0405	1.0157	16.64	0.0451	1.0316	13.12
1/96	0.0268	1.1433	28.16	0.0290	1.1383	24.77	0.0330	1.1200	28.73	0.0368	1.1158	21.98
1/120	0.0218	0.9213	63.72	0.0236	0.9225	51.33	0.0271	0.8823	57.34	0.0302	0.8967	43.98

TABLE 9. Example 5: L^1 relative errors (L^1-re), convergence rates (L^1-cr) and CPU times (CPU [s]) (continued). The corresponding results for the basic scheme are given in Table 8.

where $A(u)$ is defined by (1.3) with

$$a(u) = \begin{cases} 0.1 & \text{for } |u| \geq 0.25, \\ 0 & \text{otherwise,} \end{cases}$$

with zero boundary conditions. The initial condition is given by

$$u_0(x,y) = \begin{cases} -1 & \text{for } (x-0.5)^2 + (y-0.5)^2 < 0.16, \\ 1 & \text{for } (x+0.4)^2 + (y+0.4)^2 < 0.16, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 3 (a) shows a sample numerical solution. Figures 3 (b) and (c) present cuts of the numerical solutions generated for $\Delta x = 1/60$ and different values of η, and for $\eta = 3$ and different values of Δx, respectively. Tables 8 and 9 provide a detailed error history for $\eta = 3$, $\eta = 8$ and $\eta = 15$. Figure 4 summarizes this information graphically.

6. Conclusions

The analysis of Section 4 shows that Schemes 1 and 2 converge to the unique entropy solution of the initial value problem (1.1), (1.2). The corresponding proofs only rely on the generic properties (3.1) of the mollification weights, and do not depend on the particular way in which these are generated. Furthermore,

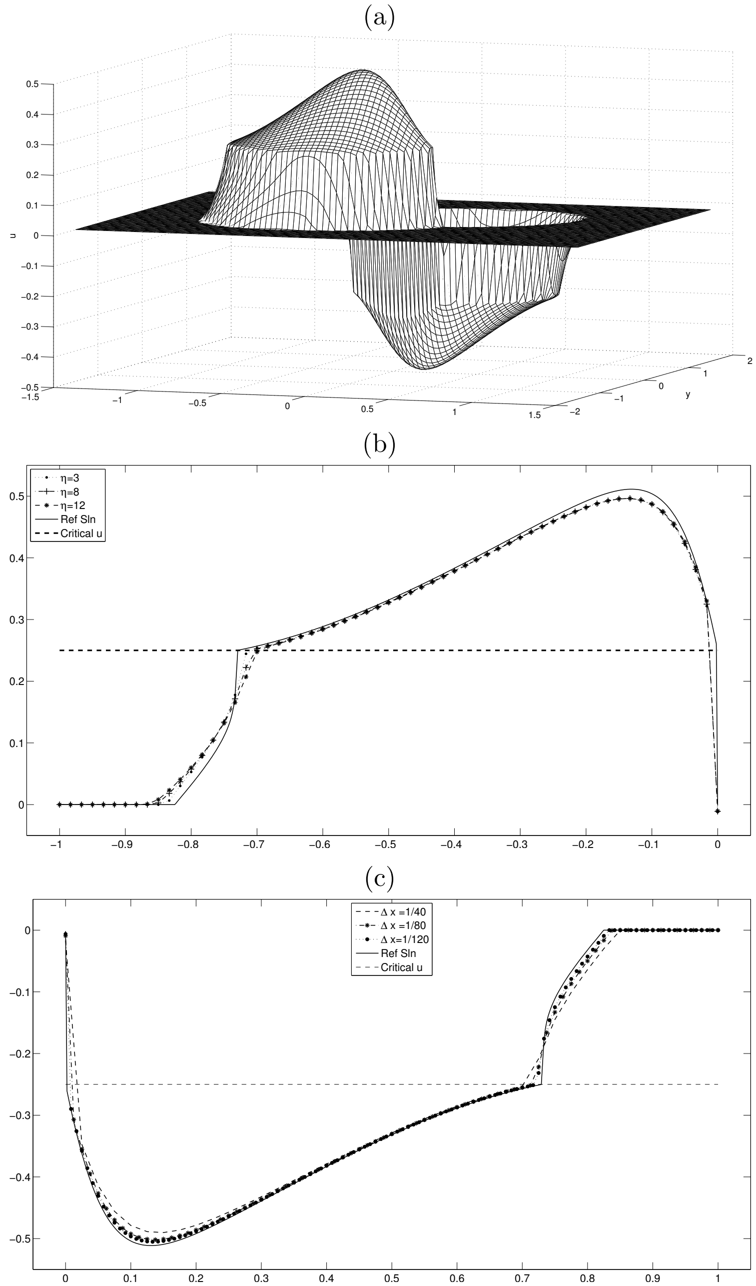

FIGURE 3. Example 5: numerical solution computed with Scheme 1 (a) with $\eta = 3$ and $\Delta x = 1/60$, (b) enlarged view of the diagonal of the numerical solution computed with Scheme 1 with $\eta = 3, 8, 12$ and $\Delta x = 1/60$, (c) enlarged view of the diagonal of the numerical solution computed with Scheme 1 with $\eta = 3$ and $\Delta x = 1/40, 1/80, 1/120$.

it is clear hat simpler versions of the present proofs will cover the one-dimensional case; on the other hand, the analysis could be extended to a general number of space dimensions and convective fluxes that depend on position, see Karlsen & Risebro (2001). Concerning the one-dimensional case, let us emphasize that our solution concept differs from that of Acosta et al. (2011). The one-dimensional analogue of item (3) of

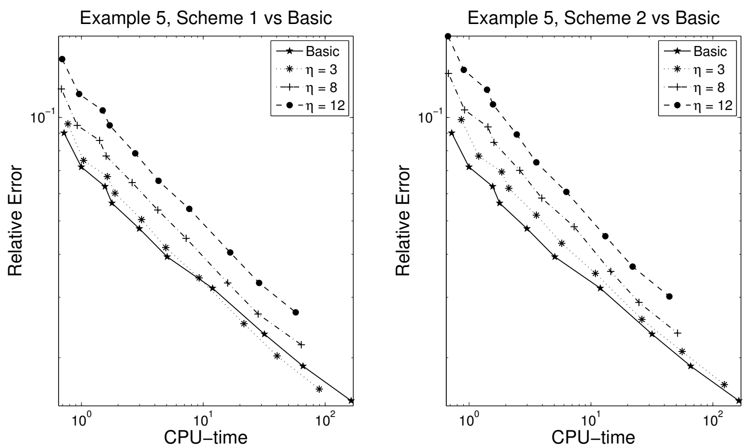

FIGURE 4. Example 5: CPU time versus approximate L^1 error.

our Defintion 2.1 implies the requirement $A(u)_x \in L^2(\Pi_T)$, while the requirement in Acosta et al. (2011), $A(u) \in C^{1,1/2}(\Pi_T)$, where $\Pi_T = \mathbb{R} \times (0,T)$, amounts to establishing a uniform L^∞ bound on $A(u^\Delta)_x$. In Lemma 3.5 of Acosta et al. (2011) this bound is established, but appears to be feasible only under restrictions on η; for example, for the weights of Table 1, one needs to impose $\eta \leq 5$. On the other hand, in Acosta et al. (2011) we seek solutions in $BV(\Pi_T)$, which makes it necessary to establish an L^1 Lipschitz continuity in time bound (Lemma 3.3 of Acosta et al., 2011); the proof of this bound depends, in turn, on a uniform bound of the spatial total variation of $A(u_0)_x$. None of the two restrictions appears in the present analysis. Having said this, it should be pointed out that the final CFL condition imposed by Acosta et al. (2011) is analogous to the milder condition (3.9) (obviously, with the term $\|g'\|_\infty$ not present in 1-d and a coefficient 2 instead of 4), while the present analysis relies on strengthened CFL conditions.

Concerning the numerical results, let us for a moment consider the CPU times only. We observe in general that for $\eta = 3$, Schemes 1 and 2 generate little savings in CPU time (compared with the basic scheme), and in some instances Scheme 2 or even both mollified schemes are slower (always considering the same value of Δx), as can be seen in Tables 4, 5 or 8. This is in agreement with the entries for $\eta = 3$ in Table 2. For Scheme 2, $\varepsilon_3^2 = 1.0379 > 1$ means that Δt even has to be slightly smaller than for the basic scheme. Combined with the fact that even for that small value of η, Scheme 2 is based on a stencil of 49 points for one evaluation of $A(u)$, it seems surprising that at least for small values of Δx, this scheme can still compete with the basic scheme. Indeed, all tables indicate that the CPU times for Scheme 2 with $\eta = 3$ are close to, and in many cases for small Δx, even smaller than those of the basic scheme. For $\eta = 8$ and $\eta = 12$, the CPU times for Schemes 1 and 2 are substantially smaller than those of the basic scheme, with the exception of large values of Δx. This trend is evident in Tables 3, 4, 6, 7 and 9. Moreover, the CPU times for $\eta = 12$ become consistently smaller than those for $\eta = 8$ as Δx increases. For example, the two last lines of Tables 6 and Table 9 for $\eta = 12$, corresponding to $\Delta x = 1/96$ and $\Delta x = 1/120$ in the respective cases of Examples 3 and 5, indicate a factor of acceleration (compared with the basic scheme) ranging between 2.00 and 3.71. This is remarkable since although the maximal factor by which Δt can be increased is $1/\varepsilon_{12}^1 = 10.12$ and $1/\varepsilon_{12}^2 = 8.92$ for Scheme 1 and 2, respectively, the evaluation of $\Delta A(u)$ for each of these schemes is based on a stencil of $4\eta + 1 = 49$ and $(2\eta+1)^2 = 625$ points, respectively.

Let us now relate the CPU times to the observed exact or approximate relative L^1 errors. First of all, the observed L^1 convergence rates are consistent with the first order of accuracy of all numerical schemes considered herein. Moreover, for Examples 1 to 4 the errors for the mollified versions are, for a given value of Δx, slightly larger, but in general very close to those of the basic scheme. For Example 5, which has a discontinuous exact solution, the error produced by the mollified versions can become twice as large as that of the basic scheme, as can be seen, for example, in the case of Scheme 2 with $\eta = 12$ and $\Delta x = 120$ for

Example 5 (see Table 9). Of course, Scheme 2 generates the solution in only 27% of the CPU time required by the basic scheme. Thus, we need to assess which scheme is most efficient, i.e., reduces the error below a given treshold in shortest time. To this end we have plotted the error histories of all runs in L^1 error versus CPU time diagrams, see Figures 1 for Examples 1 to 4 and Figure 4 for Example 5, respectively. In each diagram the graph that is lowest, at least for small errors and large CPU times (corresponding to small values of Δx) corresponds to the most efficient method. In this sense, we conclude from the plots of Figure 1 that for these examples, the schemes for $\eta = 8$ or $\eta = 12$ are always more efficient than the basic scheme or Scheme 1 or 2 for $\eta = 3$. On the other hand, Figure 4 indicates that for Example 5, these schemes (with the exception of Scheme 1 for $\eta = 3$ and small Δx) are always slightly less efficient than the basic scheme. This is most likely related to the non-smooth nature of the solution.

To put this observation into the proper perspective, let us recall that we are using a given set of weights to be consistent with previous work (see Section 3.1). Although this should be a very interesting question for further research, we are not investigating the problem of how to construct the family of weights optimally for a given PDE. In other words, through a change in the weights the performance of Schemes 1 and 2 for degenerate problems (such as our Example 5) can possibly be improved, and these schemes could eventually turn out to be more efficient than the basic scheme. Having said this, there are applications in which the use of a larger time step through a mollified scheme is an asset even if the scheme is not more efficient. For example, the solution of certain parameter identification problems for strongly degenerate parabolic equations by the adjoint equation method (see, e.g., Coronel et al., 2003) requires the storage of the full numerical solution of the direct problem as a function of time and space. Evidently, the storage space requirements would be decreased by the factor the time step can be increased.

Finally, we mention that for simplicity it is assumed that the parameter η and the weights employed are the same on all portions of the computational domain, and for all times. It would be very interesing to make the effort to generate additional gains in accuracy, and savings in CPU time, by choosing these parameters adaptively in response to local information on the smoothness of the approximate solution.

Appendix

Proof of Lemma 4.6. We recall that

$$\begin{aligned} \Delta_x^+ F_{i-1/2,j}^n &= f^-\big(u_{i+1,j}^n\big) - f^-\big(u_{ij}^n\big) + f^+\big(u_{ij}^n\big) - f^+\big(u_{i-1,j}^n\big), \\ \Delta_y^+ G_{i,j-1/2}^n &= g^-\big(u_{i,j+1}^n\big) - g^-\big(u_{ij}^n\big) + g^+\big(u_{ij}^n\big) - g^+\big(u_{i,j-1}^n\big). \end{aligned} \tag{A.1}$$

If we define the functions $\mathcal{F}^\pm(u) := \int_0^u s f^\pm(s)\,\mathrm{d}s$ and $\mathcal{G}^\pm(u) := \int_0^u s g^\pm(s)\,\mathrm{d}s$, then for all $a, b \in \mathbb{R}$

$$\mathcal{F}^\pm(b) - \mathcal{F}^\pm(a) = b\big(f^\pm(b) - f^\pm(a)\big) - \int_a^b \big(f^\pm(b) - f^\pm(s)\big)\,\mathrm{d}s,$$

with the same identity holding for $\mathcal{G}^\pm$ and $g^\pm$. We then obtain

$$\begin{aligned} \big(f^-\big(u_{i+1,j}^n\big) - f^-\big(u_{ij}^n\big)\big)u_{ij}^n &= \mathcal{F}^-\big(u_{i+1,j}^n\big) - \mathcal{F}^-\big(u_{ij}^n\big) - \int_{u_{i+1,j}^n}^{u_{ij}^n} \big(f^-(s) - f^-\big(u_{i+1,j}^n\big)\big)\,\mathrm{d}s, \\ \big(f^+\big(u_{ij}^n\big) - f^+\big(u_{i-1,j}^n\big)\big)u_{ij}^n &= \mathcal{F}^+\big(u_{ij}^n\big) - \mathcal{F}^+\big(u_{i-1,j}^n\big) - \int_{u_{i-1,j}^n}^{u_{ij}^n} \big(f^+(s) - f^+\big(u_{i-1,j}^n\big)\big)\,\mathrm{d}s. \end{aligned} \tag{A.3}$$

Therefore, considering that the sums over $i, j \in \mathbb{Z}$ of the differences of evaluations of $\mathcal{F}^\pm$ and $\mathcal{G}^\pm$ in (A.3) and related identities involving $\mathcal{G}^\pm$ and $g^\pm$ are zero, we obtain

$$\begin{aligned} S = \Delta x \Delta t \sum_{i,j,n} \Bigg\{ &\int_{u_{i-1,j}^n}^{u_{ij}^n} \big(f^+(s) - f^+\big(u_{i-1,j}^n\big)\big)\,\mathrm{d}s - \int_{u_{i+1,j}^n}^{u_{ij}^n} \big(f^-(s) - f^-\big(u_{i+1,j}^n\big)\big)\,\mathrm{d}s \\ &+ \int_{u_{i,j-1}^n}^{u_{ij}^n} \big(g^+(s) - g^+\big(u_{i,j-1}^n\big)\big)\,\mathrm{d}s - \int_{u_{i,j+1}^n}^{u_{ij}^n} \big(g^-(s) - g^-\big(u_{i,j+1}^n\big)\big)\,\mathrm{d}s \Bigg\}. \end{aligned} \tag{A.3}$$

As a consequence of the identity

$$\left|\int_a^b \big(h(\xi)-h(a)\big)\,\mathrm{d}\xi\right| \geq \frac{(h(b)-h(a))^2}{2L_h} \quad \text{for all } a,b\in\mathbb{R}$$

valid for all monotone Lipschitz continuous functions with Lipschitz constant L_h (see Karlsen & Risebro, 2001), we obtain

$$\begin{aligned}\int_{u_{i-1,j}^n}^{u_{ij}^n} \big(f^+(s)-f^+\big(u_{i-1,j}^n\big)\big)\,\mathrm{d}s &\geq \frac{1}{2\|f'\|}\big(f^+\big(u_{ij}^n\big)-f^+\big(u_{i-1,j}^n\big)\big)^2,\\ -\int_{u_{i+1,j}^n}^{u_{ij}^n} \big(f^-(s)-f^-\big(u_{i+1,j}^n\big)\big)\,\mathrm{d}s &\geq \frac{1}{2\|f'\|}\big(f^-\big(u_{i+1,j}^n\big)-f^-\big(u_{ij}^n\big)\big)^2,\end{aligned}$$

and analogous inequalities for the two remaining integrals in (A.3). This concludes the proof of (4.11).

Acknowledgements

CDA and RB acknowledge support by BASAL project CMM, Universidad de Chile and Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción. In addition, RB is supported by Fondecyt project 1090456 and CDA by Universidad Nacional de Colombia, DIMA Project 20201005601.

References

Acosta, C.D., Bürger, R. & Mejía, C.E. (2011) Monotone difference schemes stabilized by discrete mollification for strongly degenerate parabolic equations. *Numer. Meth. Partial Diff. Eqns.*, to appear.

Acosta, C.D. & Mejía, C.E. (2008) Stabilization of explicit methods for convection-diffusion problems by discrete mollification. *Comput. Math. Applic.*, **55**, 363–380.

Acosta, C.D. & Mejía, C.E. (2009) Approximate solution of hyperbolic conservation laws by discrete mollification. *Appl. Numer. Math.*, **59**, 2256–2265.

Acosta, C.D. & Mejía, C.E. (2010) A mollification-based operator splitting method for convection-diffusion equations. *Comput. Math. Applic.*, **59**, 1397–1408.

Afif, M. & Amaziane, B. (2002) Convergence of finite volume schemes for a degenerate convection-diffusion equation arising in flow in porous media. *Comput. Methods Appl. Mech. Engrg.*, **191**, 5265–5286.

Berres, S., Bürger, R., Karlsen, K.H. & Tory, E.M. (2003) Strongly degenerate parabolic-hyperbolic systems modeling polydisperse sedimentation with compression. *SIAM J. Appl. Math.*, **64**, 41–80.

Betancourt, F., Bürger, R. & Karlsen, K.H. (2011) A strongly degenerate parabolic aggregation equation. *Commun. Math. Sci.*, to appear.

Bürger, R., Coronel, A. & Sepúlveda, M. (2006) A semi-implicit monotone difference scheme for an initial-boundary value problem of a strongly degenerate parabolic equation modelling sedimentation-consolidation processes. *Math. Comp.*, **75**, 91–112.

Bürger, R. & Karlsen, K.H. (2003) On a diffusively corrected kinematic-wave traffic model with changing road surface conditions. *Math. Models Methods Appl. Sci.* **13**, 1767–1799.

Bürger, R., Karlsen, K.H. & Towers, J.D. (2005) A mathematical model of continuous sedimentation of flocculated suspensions in clarifier-thickener units. *SIAM J. Appl. Math.*, **65**, 882–940.

Cecchi, M. & Pirozzi, M. (2005) High order finite difference numerical methods for time-dependent convection–dominated problems. *Appl. Numer. Math.*, **55**, 334–356.

Champier, S., Gallouët, T. & Herbin, R. (1993) Convergence of an upstream finite volume scheme for a nonlinear hyperbolic equation on a triangular mesh. *Numer. Math.*, **66**, 139–157.

Coronel, A., James, F. & Sepúlveda, M. (2003) Numerical identification of parameters for a model of sedimentation processes. *Inverse Problems*, **19**, 951–972.

Crandall, M.G. & Majda, A. (1980) Monotone difference approximations for scalar conservation laws. *Math. Comp.*, **34**, 1-21.

Engquist, B. & Osher, S. (1981) One-sided difference approximations for nonlinear conservation laws. *Math. Comp.*, **36**, 321–351.

Evje, S. & Karlsen, K.H. (1999) Degenerate convection-diffusion equations and implicit monotone difference schemes. In: Freistühler, H. & Warnecke, G. (eds.), Hyperbolic Problems: Theory, Numerics, Applications, Birkhäuser-Verlag, Basel, Switzerland, 285–294.

Evje, S. & Karlsen, K.H. (2000) Monotone difference approximations of BV solutions to degenerate convection-diffusion equations. *SIAM J. Numer. Anal.*, **37**, 1838–1860.

Eymard, R., Gallouët, T., Ghilani, M. & Herbin, R. (1998a) Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes. *IMA J. Numer. Anal.*, **18**, 563–594.

Eymard, R., Gallouët, T., Hilhorst, D. & Naït Slimane, Y. (1998b) Finite volumes and nonlinear diffusion equations. *RAIRO Modél. Math. Anal. Numér.*, **32**, 747–761.

Harten, A., Hyman, J.M. & Lax, P.D. (1976) On finite difference approximations and entropy conditions for shocks. *Comm. Pure Appl. Math.*, **29**, 297–322.

Holden, H., Karlsen, K.H. & Lie, K.-A. (2000) Operator splitting methods for degenerate convection-diffusion equations II: numerical examples with emphasis on reservoir simulation and sedimentation. *Comput. Geosci.*, **4**, 287–322.

Holden, H., Karlsen, K.H., Lie, K.-A. & Risebro, N.H. (2010) Splitting Methods for Partial Differential Equations with Rough Solutions. European Mathematical Society, Zurich, Switzerland.

Karlsen, K.H. & Risebro, N.H. (1997) An operator splitting method for nonlinear convection-diffusion equations. *Numer. Math.*, **77**, 365–382.

Karlsen, K.H. & Risebro, N.H. (2001) Convergence of finite difference schemes for viscous and inviscid conservation laws with rough coefficients. *M2AN Math. Model. Numer. Anal.*, **35**, 239–269.

Karlsen, K.H. & Risebro, N.H. (2003) On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients. *Discrete Contin. Dyn. Syst.*, **9**, 1081-1104.

Karlsen, K.H., Risebro, N.H. & Towers, J.D. (2002) Upwind difference approximations for degenerate parabolic convection-diffusion equations with a discontinuous coefficient. *IMA J. Numer. Anal.*, **22**, 623–664.

Karlsen, K.H., Risebro, N.H. & Towers, J.D. (2003) L^1 stability for entropy solutions of nonlinear degenerate parabolic convection-diffusion equations with discontinuous coefficients. *Skr. K. Nor. Vid. Selsk.*, **3**, 49 pp.

Kružkov, S.N. (1969) Results on the nature of the continuity of solutions of parabolic equations, and certain applications thereof. *Mat. Zametki*, **6**, 97–108 (in Russian).

Mejía, C.E. & Murio, D.A. (1995) Numerical identification of diffusivity coefficient and initial condition by discrete mollification. *Comput. Math. Applic.*, **30**, 35–50.

Mejía, C.E. & Murio, D.A. (1996) Numerical solution of generalized IHCP by discrete mollification. *Comput. Math. Applic.*, **32**, 33–50.

Murio, D.A. (1993) The Mollification Method and the Numerical Solution of Ill-Posed Problems. John Wiley, New York.

Murio, D.A. (2002) Mollification and space marching. In: Woodbury, K. (Ed.), Inverse Engineering Handbook, CRC Press, Boca Raton, FL, USA.

Murio, D.A., Mejía, C.E. & Zhan, S. (2001) Some applications of the mollification method, in: Lassonde, M. (Ed.), Approximation, Optimization and Mathematical Economics, Physica-Verlag.

Rouvre, E. & Gagneux, G. (1999) Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées. *C. R. Acad. Sci. Paris Sér. I* **329**, 599–602.