

# Numerical Scheme for a Scalar Conservation Law with Memory

**Małgorzata Peszyńska**

*Department of Mathematics, Oregon State University, Corvallis, Oregon 97331*

*Received 13 January 2013; accepted 13 May 2013*

*Published online in Wiley Online Library (wileyonlinelibrary.com).*

*DOI 10.1002/num.21806*

We consider approximation of solutions to conservation laws with memory represented by a Volterra term with a smooth decreasing but possibly unbounded kernel. The numerical scheme combines Godunov method with a treatment of the integral term following from product integration rules. We prove stability for both linear and nonlinear flux functions and demonstrate the expected order of convergence using numerical experiments. The problem is motivated by modeling advective transport in heterogeneous media with subscale diffusion. © 2013 Wiley Periodicals, Inc. *Numer Methods Partial Differential Eq* 000: 000–000, 2013

*Keywords:* Godunov scheme; conservation laws; memory terms; Volterra integral terms

## I. INTRODUCTION

In computational modeling of flow and transport in heterogeneous domains, one frequently has to work with a practical modeling scale which is much larger than that of some coupled diffusion processes. This occurs, in particular, in highly heterogeneous porous media where the advection dominates any other transport mechanisms in the highly permeable part of the domain of flow. This advective transport is coupled to the local diffusion phenomena in the adjacent, much less permeable domains, in which advection is negligible. To avoid modeling at all the spatial scales involved, one can use an upscaled model in which appropriate memory terms account for the subscale diffusion [1–3].

Transport models with memory are used in industrial applications, for example, to describe gas transport in coal seams involving adsorption and diffusion in meso- and micropores [4–9], a process accompanying methane recovery and carbon sequestration in coalbeds. Another example is that of transport models in subsurface and hyporheic zones [10–12]. Memory terms appear in other applications such as viscoelasticity, heat conduction in materials with memory, and electromagnetism in dispersive media. We review various model applications for transport with memory in Section II.

*Correspondence to:* Małgorzata Peszyńska, Department of Mathematics, Oregon State University, Corvallis, OR 97331 (e-mail: mpezs@math.oregonstate.edu)

Contract grant sponsor: National Science Foundation; contract grant numbers: DMS 1115827 (“Hybrid modeling in porous media”), DMS 0511190 (“Model Adaptivity for porous media”), and Department of Energy grant 98089 (“Modeling, analysis, and simulation of preferential flow”).

© 2013 Wiley Periodicals, Inc.

## 2 PESZYŃSKA

In this article, we develop and analyze numerical discretization of the prototype model

$$u_t + \beta * u_t + f(u)_x = 0, \quad x \in \mathbb{R}, t \in (0, T), \quad (1.1)$$

in which the unknown  $u$  is, for example, the concentration of some species being advected in the medium,  $f$  is a linear or strictly increasing nonlinear convex flux function, and the Volterra convolution integral

$$(\beta * u_t)(x, t) := \int_0^t \beta(t-s)u_t(x, s)ds, \quad (1.2)$$

has kernel  $\beta$  whose properties are crucial for analysis and efficient approximation of (1.1).

Throughout the article, we assume that  $\beta$  is nonnegative, nonincreasing, and smooth,

$$\beta(t) \geq 0, \quad \forall t > 0, \quad (1.3)$$

$$\beta'(t) \leq 0, \quad \forall t > 0, \quad (1.4)$$

$$\beta \in L^1_{loc}(\mathbb{R}_+). \quad (1.5)$$

It is important to distinguish two cases depending on behavior of  $\beta$  at  $t = 0$

$$\beta \in \mathcal{B}_\infty := \{\beta \in C^0(0, \infty) : \lim_{t \rightarrow 0} \beta(t) = +\infty\}, \quad (1.6)$$

$$\beta \in \mathcal{B}_b := \{\beta \in C^0([0, \infty)) : \beta(0) = b\}. \quad (1.7)$$

Typical examples of kernels from these two classes are, respectively,  $\beta(t) = t^{-1/2}$ , and  $\beta(t) = \exp(-t)$ . For notation convenience, we also consider  $\beta \in \mathcal{B}_0$ , that is,  $\beta \equiv 0$ .

Many of the technical results we develop for (1.1) can be easily extended, to the fractional conservation law

$$\beta * u_t + f(u)_x = 0, \quad x \in \mathbb{R}, t \in (0, T), \quad (1.8)$$

in which  $\beta \in \mathcal{B}_\infty$ . In particular, if  $\beta(t) = \frac{1}{\Gamma(r)}t^{r-1}$ ,  $r \in (0, 1)$ , then one can identify  $\beta * u_t$  as the fractional derivative of  $u$  of order  $1 - r$  [13]. The model (1.8) is itself important in applications, see Section II.

It is therefore natural to combine (1.1) and (1.8) and consider the initial value problem

$$\alpha u_t + \beta * u_t + f(u)_x = 0, \quad x \in \mathbb{R}, t \in (0, T) \quad (1.9a)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R} \quad (1.9b)$$

with  $\alpha = 1$  or  $\alpha = 0$ , respectively. The case with an arbitrary  $\alpha > 0$  can be considered via rescaling.

When  $\alpha = 1$  and  $\beta \in \mathcal{B}_0$ , the Eq. (1.9a) is a scalar conservation law whose analysis and approximation encounters well-known difficulties. The approximation of scalar conservation laws is well-understood [14–17]. In contrast, we are not aware of any known results for numerical approximation to (1.9) if  $\beta \neq 0$ . In this case, the solutions may have an interesting structure due to the presence of the integral terms which give rise to certain smoothing effects, for example, prevent or mollify development of singularities. In particular, as recently reviewed in Ref. [18] with  $\alpha = 1$ ,  $\beta \in \mathcal{B}_\infty$ , and linear  $f$ , the solution to Riemann problem, that is, with  $u_0 = 1 - H(x)$ ,

where  $H$  is the Heaviside function, is continuous for  $t > 0$ ; see also Refs. [19–22]. On the other hand, as shown in Ref. [18], there is no such smoothing involved if  $f$  is nonlinear or  $\beta \in \mathcal{B}_b$ . However, if  $\alpha = 0$ ,  $\beta \in \mathcal{B}_\infty$ , the solution is continuous [22].

The purpose of this article is to define and analyze a numerical approximation scheme for (1.9). We consider the well-known upwind scheme which arises naturally for the case  $\alpha = 1$ ,  $\beta \equiv 0$ , and extend it to (1.9) by an appropriate treatment of the convolution term  $\beta * u_t$ , following the idea from [23, 24] where a similar term appeared in a parabolic equation with memory. The scheme and the notation are introduced in Section III.

In Section IV, we consider the case of linear flux  $f$  and use von-Neumann stability analysis to show that an explicit scheme proposed in Section III is stable as long as a condition similar to CFL condition [14] holds. Specifically for  $\alpha = 1$ , the time step  $k$  must satisfy  $k \leq C(\beta, f)h$ , where  $h$  is the spatial step, and where  $C(\beta, f)$  grows with the increasing singularity of  $\beta$  and decreases with increasing Lipschitz constant for  $f$ . The case  $\alpha = 0$  is handled similarly but the stability is guaranteed only if  $\beta \in \mathcal{B}_\infty$  and under a severe restriction on  $k$ ; in particular, for  $r = \frac{1}{2}$  we find that  $k = O(h^2)$  is required. Thus, in this case, an implicit scheme which is shown to be unconditionally stable, may be preferable. For nonlinear flux  $f$ ,  $\beta \in \mathcal{B}_\infty$ , and under additional assumptions, we show in Section V that the explicit scheme is  $TV$ -stable.

In Section VI, we present numerical examples. On one hand, we confirm our theoretical convergence results. On the other hand, we provide simulations to illustrate the smoothness of the solutions depending on the singularity of  $\beta$ .

In this article, we consider only first-order schemes in one spatial dimension for scalar problems. The extensions to higher-order schemes as well as to realistic multicomponent systems and multidimensional application examples are forthcoming. This article is a first step toward the understanding of interplay between smoothing effects due to memory effects and nonlinear advection, and the numerical stability and instabilities arising from these, respectively.

Finally, the analysis of the discrete scheme for (1.9) in this article is completely different from that in Refs. [23, 24] which relied on energy estimates. It is also distinct from the analysis of schemes for viscoelasticity and other problems with memory terms  $\beta * Au$  with nonsingular and weakly singular kernels  $\beta$  [25–27] which strongly relied on energy estimates for the second-order differential operator  $A$ .

## II. APPLICATIONS OF CONSERVATION LAWS WITH MEMORY

The model (1.9) is a simplification of a more general equation motivated by the experimental results in [1, 28]

$$u_t + \beta * u_t + vu_x - Du_{xx} + \beta_1 * u_{xxt} + \beta_2 * u_{xt} = 0, \tag{2.1}$$

which we obtained in Ref. [2]. Here,  $D > 0$ ,  $v$  are the macroscopic diffusion and advection coefficients, and the kernels  $\beta, \beta_1, \beta_2$  are derived from the solutions to microscopic advection-diffusion cell problems. In Ref. [3] the first set of simulations confirmed good agreement between (2.1) and an underlying model of transport with memory in highly heterogeneous media comprised of highly permeable medium interlaced with low permeability zones. The model (2.1) differs from the usual homogenized models in that it does not assume scale separation and allows a continuum of contrasts between the two zones. The scale separation constant plays a role in the relative importance and magnitude of the kernels  $\beta_1, \beta_2$  with respect to  $\beta$ , with the latter known as the primary diffusion kernel. The kernel  $\beta$  can be computed analytically or numerically for simple

geometries of microblocks, for example, is given by weakly-singular Fourier–Bessel series (see Ref. [23]), whereas its qualitative nature is well-described by  $\beta(t) \sim t^{-1/2}$  [23, 29]. For high contrast between the zones and large flow rates,  $D \approx 0$ ,  $\beta_1 \approx \beta_2 \approx 0$ . More studies and rigorous analyses are underway extending [3] and this article.

A model similar to (2.1) with  $\beta_1 \equiv \beta_2 \equiv 0$

$$u_t + \beta * u_t + vu_x - Du_{xx} = 0, \quad (2.2)$$

was used in modeling hyporheic zones [10–12] where the contaminants moving quickly in a stream or river interact with the storage in the sediment and banks surrounding the stream. Here,  $\beta(t) \sim t^r$ ,  $r = -0.28$  was found by a match with experiments.

The use of memory terms as models of subscale diffusion dates back to double-porosity models such as the Warren–Root model [29–31], which has been popular in modeling fractured oil reservoirs. More recently, such models are used routinely to describe gas transport involving adsorption and diffusion of methane and carbon dioxide in coalbeds [4–9]. In particular, consider

$$u_t + a_t + vu_x = 0, \quad (2.3)$$

where  $u$  is the concentration of species being transported through coal and  $a$  is the amount adsorbed, and  $v$  is the transport velocity. Because coals have multiporous structure, it is known that an equilibrium model  $a = g(u)$  is not adequate. [Here,  $g(\cdot)$  is an increasing smooth function called isotherm [32].] Instead one should use the full-microblock diffusion-adsorption model in which  $a_t = \beta * g(u)_t$ ; here  $\beta(t) \sim t^{-1/2}$ . Via a change of variable, this model shares structural similarity with (1.9).

As a simplification, Ref. [5] uses the quasi-static (kinetic) model in which

$$a_t + \tau(a - g(u)) = 0. \quad (2.4)$$

It is not difficult to calculate that with  $\beta(t) = \tau \exp(-\tau t)$

$$a_t(x, t) = \beta * g(u(x, t))_t + \beta(t)(a(x, 0) - g(u(x, 0))), \quad (2.5)$$

thus making (2.3)–(2.5) similar to the nonlinear model (1.9) with a bounded kernel.

A more general multicomponent model of methane and/or carbon dioxide adsorption in coal formulated for the needs of industry in Refs. [4–9] should account for dependence of  $v = v(u)$ , an extension to a system, as well as compressibility effects. See Ref. [32] for an exposition on models of methane evolution and a multicomponent nonequilibrium example. A comprehensive model will require approximate Riemann solvers and an equation of state, but the analysis is outside the present scope.

The model (1.9) we study in this article is thus clearly a simplification of each of (2.1), (2.2), and (2.3). It includes, however, the memory terms, that is, those elements of each of (2.1), (2.2), and (2.3) which set them apart from standard models. It is the goal of this article to address and analyze its numerical discretization, so that eventually more comprehensive models can be treated.

### III. DISCRETIZATION SCHEMES FOR (1.9)

We follow the usual nomenclature as in Refs. [14, 15, 17, 33, 34]. To avoid unnecessary technicalities associated with the treatment of boundaries, we consider the solution to the IVP

(1.9) only for initial data with compact support and assume that the spatial domain of interest is covered by a uniform spatial grid with parameter  $h$ . Thus, we consider the spatial grid at  $x_j = jh, j = -\infty \dots \infty$ . We consider only uniform time-stepping with a parameter  $k$ , that is, with time steps given by  $t_n = nk, n = 0, 1, \dots$ . We define and analyze approximations  $u_j^n \approx u(x_j, t_n)$ . Let  $W^n := (w_j^n)_{j=-\infty}^\infty, n = 0, 1, \dots$  denote a grid function. The grid norms are defined as

$$\|W^n\|_1 := h \sum_j |w_j^n|, \tag{3.1}$$

$$\|W^n\|_2 := \sqrt{h} \sqrt{\sum_j |w_j^n|^2}. \tag{3.2}$$

We denote  $\lambda := \frac{k}{h}$  as usual and keep it fixed.

**A. Discretization with No Memory Terms**

We first recall basic notation and facts for a linear scalar conservation law without memory, that is,  $\alpha = 1, \beta \equiv 0$ , and  $f(u) = vu$  [14, 34]. Let  $v > 0$ . Recall the explicit upwind scheme for (1.9) which defines an approximation  $u_j^n \approx u(x_j, t_n)$  explicitly as

$$u_j^n = u_j^{n-1} - v\lambda(u_j^{n-1} - u_{j-1}^{n-1}) := u_j^{n-1} - D_{j,-1}^{n-1}u, \quad n = 1, 2, \dots \tag{3.3}$$

We assume  $(u_j^0)_j$  is an appropriate projection of the initial data  $u_0$ .

This scheme has a local truncation error of first-order in  $h$  and  $k$ , and is first-order convergent provided the stability condition (Courant-Friedrichs-Levy (CFL) condition)  $v\lambda \leq 1$  holds. This well-known fact follows directly, for example, from von-Neumann stability analysis for (3.3), in which the ansatz  $u_j^n := \rho^n e^{ij\theta}$  in (3.3), after some algebraic manipulations, yields for the magnification factor in (3.3)  $\rho = \rho_{\text{upwind}} = 1 - \gamma_{\text{upwind}} := 1 - v\lambda(1 - e^{-i\theta})$ . Assuming  $v\lambda \leq 1$  we have  $|\rho_{\text{upwind}}| = |1 - \gamma_{\text{upwind}}| \leq 1$ . This ensures that the numerical scheme uses only the values from the domain of dependence of (1.8) which does not include, for  $v > 0$ , the point  $x_{j+1}^{n-1}$ . For that reason, the centered scheme

$$u_j^n = u_j^{n-1} - \frac{v}{2}\lambda(u_{j+1}^{n-1} - u_{j-1}^{n-1}) := u_j^{n-1} - D_{j,0}^{n-1}u \tag{3.4}$$

is unstable, with the corresponding von-Neumann analysis delivering

$$\rho_{\text{center}} = 1 - \gamma_{\text{center}} := 1 - v\lambda i \sin(\theta), \tag{3.5}$$

and the magnification factor clearly satisfying  $|\rho_{\text{center}}| \geq 1$ .

We also recall that the upwind and centered schemes are unconditionally stable if we use an implicit in time treatment of the spatial derivative terms, that is, if we replace  $D_{j,p}^{n-1}$  in (3.3) or (3.4) by  $D_{j,p}^n$ . This follows from

$$|(1 + \rho)^{-1}| \leq 1, \quad \forall \rho, \tag{3.6}$$

and, thus, for any  $\lambda$ , it holds for  $\rho = \rho_{\text{upwind}}, \rho = \rho_{\text{center}}$ . However, as implicit schemes require the solution of a linear system at each time step, they are rarely used for purely advective transport.

**B. Discretization of Memory Terms**

Now we consider  $\beta \neq 0$  and discuss the convolution term  $\beta * u_t$  in (1.9a). We set

$$b_m := \int_{t_m}^{t_{m+1}} \beta(s) ds. \tag{3.7}$$

Since  $t_m = mk$ , we can write

$$b_{n-m} = \int_{t_{n-m}}^{t_{n-m+1}} \beta(s) ds = \int_{t_n-t_m}^{t_n-t_{m-1}} \beta(s) ds = \int_{t_{m-1}}^{t_m} \beta(t_n - s) ds.$$

Since  $\beta$  is integrable, the following approximation is well defined even if  $\beta \in \mathcal{B}_\infty$ :

$$\begin{aligned} (\beta * u_t)|_{x,t_n} &= \sum_{m=1}^n \int_{t_{m-1}}^{t_m} \beta(t_n - s) u_t(x, s) ds \\ &\approx \sum_{m=1}^n \frac{1}{k} (u_j^m - u_j^{m-1}) \int_{t_{m-1}}^{t_m} \beta(t_n - s) ds = \sum_{m=1}^n \frac{1}{k} (u_j^m - u_j^{m-1}) b_{n-m}. \end{aligned} \tag{3.8}$$

This discretization based on product integration rules was used in Refs. [23, 24] for a parabolic equation with memory for which we proved appropriate convergence results for both bounded and weakly singular kernels. Other approaches were used in Refs. [25–27] for bounded kernels and Volterra terms involving derivatives other than  $u_t$ .

We can easily show that (3.8) is first-order consistent.

**Lemma 3.1.** *Let (1.5) hold. Then the local truncation error in (3.8) is of order  $O(k)$ .*

**Proof.** Recall that the consistency error is the residual remaining after we plug the exact solution of a problem  $u(x, t)$  evaluated at the appropriate discrete points  $(x_j, t_m)$  into the scheme. The calculations use Taylor series expansions of  $u$  in terms of its time derivatives and the consistency order is derived formally assuming  $u$  is smooth.

Consider

$$\frac{1}{k} \sum_{m=1}^n (u(x, t_m) - u(x, t_{m-1})) b_{n-m} = \frac{1}{k} \int_{t_{m-1}}^{t_m} \sum_{m=1}^n (u(x, t_m) - u(x, t_{m-1})) \beta(t_n - s) ds. \tag{3.9}$$

Now expand using Taylor series

$$\begin{aligned} u(x, t_m) &= u(x, s) + (t_m - s)u_t(x, s) + \frac{1}{2}(t_m - s)^2 u_{tt}(x, s) + \dots \\ u(x, t_{m-1}) &= u(x, s) + (t_{m-1} - s)u_t(x, s) + \frac{1}{2}(t_{m-1} - s)^2 u_{tt}(x, s) + \dots \end{aligned}$$

The difference of these two equations can be used to estimate the contribution to the truncation error from approximation of the integral term, that is, the difference of the term on the left side of (3.8) and its discretization in (3.9)

$$\begin{aligned} & |(\beta * u_t)|_{x,t_n} - \frac{1}{k} \sum_{m=1}^n (u(x, t_m) - u(x, t_{m-1})) b_{n-m} | \\ & \leq \frac{1}{k} \int_{t_{m-1}}^{t_m} \sum_{m=1}^n \beta(t_n - s) (C_u k^2) = C_u k \|\beta\|_{L_1(0,t_n)}. \end{aligned}$$

where we have used (1.3), (1.5), and  $C_u := \frac{1}{2} \max_{x \in \mathbb{R}, t \in (0,t_n)} |u_{tt}(x, t)|$ . ■

**C. Memory-Upwind Scheme for Linear Flux**

With (3.8), we formulate a general memory-upwind scheme for (1.9a) in the form

$$\alpha \frac{1}{k} (u_j^n - u_j^{n-1}) + \sum_{m=1}^n \frac{1}{k} (u_j^m - u_j^{m-1}) b_{n-m} + D_{j,-1}^{n*} u = 0, \tag{3.10}$$

in which the approximation of the convolution term (3.8) is combined with the upwind approximation  $D_{j,-1}^{n*} u$  to  $(vu_x)|_{(x_j,t_n)}$  as in (3.3). Here, one can use either the explicit  $n^* = n - 1$  or the implicit-in time  $n^* = n$  treatment.

Lemma 3.1 combined with standard truncation error analysis for the nonmemory terms in (3.10) such as, for example, in Ref. [14], lead to the following

**Corollary 3.2.** *The order of consistency of (3.10) is that of  $O(h + k)$ .*

This consistency result combined with appropriate stability results to be proved in Section IV under CFL-like conditions imply convergence of the scheme (3.10). Extension to nonlinear  $f$  is discussed in Section V. The order of convergence is numerically confirmed for smooth solutions, whereas a suboptimal order is obtained for nonsmooth solutions; see Section VI.

**IV. STABILITY OF MEMORY-UPWIND SCHEME FOR LINEAR FLUX**

We prove below that for  $\alpha = 1$ , the stability of the explicit scheme (3.10) is conditional and based on a CFL-like condition which enforces  $k = O(h)$  and thus leads, by Corollary 3.2, to optimal linear convergence for smooth solutions. For  $\alpha = 0$ , the stability condition depends on  $\beta$ ; in particular, for the fractional conservation law with  $\beta(t) \sim t^{-1/2}$ , we require  $k = O(h^2)$  and the linear convergence order  $O(k + h)$  is suboptimal.

An implicit scheme is unconditionally stable for either  $\alpha = 0$  or  $\alpha = 1$ , and the convergence order is linear.

In the stability proof, we use the following consequence of Rouché’s theorem [35].

**Lemma 4.1.** *Let  $n \geq 1$  and  $\rho$  be the root of the polynomial*

$$\rho^n + a_{n-1} \rho^{n-1} + \dots + a_1 \rho + a_0 = 0. \tag{4.1}$$

*Then*

$$|\rho| \leq \max \left\{ 1, \sum_{m=0}^{n-1} |a_m| \right\}. \tag{4.2}$$

**Theorem 4.2.** *Let  $\beta$  satisfy (1.3)–(1.5). The following schemes are stable under the additional conditions as below.*

i. For  $\alpha = 1$  and  $\beta \in \mathcal{B}_\infty \cup \mathcal{B}_b$  the explicit upwind scheme

$$(u_j^n - u_j^{n-1}) + \sum_{m=1}^n (u_j^m - u_j^{m-1})b_{n-m} + \lambda v(u_j^{n-1} - u_{j-1}^{n-1}) = 0. \tag{4.3}$$

is stable provided  $v\lambda \leq 1$ .

ii. For  $\alpha = 0$  and  $\beta \in \mathcal{B}_\infty$ , the explicit upwind scheme

$$\sum_{m=1}^n (u_j^m - u_j^{m-1})b_{n-m} + \lambda v(u_j^{n-1} - u_{j-1}^{n-1}) = 0. \tag{4.4}$$

is stable provided  $v\lambda \leq b_0 - b_1$ .

iii The implicit upwind scheme

$$\alpha(u_j^n - u_j^{n-1}) + \sum_{m=1}^n (u_j^m - u_j^{m-1})b_{n-m} + \lambda v(u_j^n - u_{j-1}^n) = 0, \tag{4.5}$$

is unconditionally stable for  $\alpha = 1$  with  $\beta \in \mathcal{B}_\infty \cup \mathcal{B}_b$ , or when  $\alpha = 0$  with  $\beta \notin \mathcal{B}_0$ .

**Proof.** First rewrite the terms in (3.8)

$$\begin{aligned} (\beta * u_t)|_{x,t_n} &\approx \frac{1}{k} [(u_j^1 - u_j^0)b_{n-1} + (u_j^2 - u_j^1)b_{n-2} + \dots + (u_j^{n-1} - u_j^{n-2})b_1 + (u_j^n - u_j^{n-1})b_0] \\ &= \frac{1}{k} [b_0 u_j^n + (b_1 - b_0)u_j^{n-1} + (b_2 - b_1)u_j^{n-2} + \dots + (b_{n-1} - b_{n-2})u_j^1 - b_{n-1}u_j^0] \\ &= \frac{1}{k} \left[ b_0 u_j^n + (b_1 - b_0)u_j^{n-1} + \sum_{m=1}^{n-2} u_j^m (b_{n-m} - b_{n-m-1}) - u_j^0 b_{n-1} \right]. \end{aligned} \tag{4.6}$$

To prove (i)–(ii), we apply (4.6) in (3.10), that is, in (4.3) and (4.4). We get

$$(\alpha + b_0)(u_j^n - u_j^{n-1}) + \left[ b_1 u_j^{n-1} + \sum_{m=1}^{n-2} u_j^m (b_{n-m} - b_{n-m-1}) - u_j^0 b_{n-1} \right] + v\lambda(u_j^{n-1} - u_{j-1}^{n-1}) = 0. \tag{4.7}$$

It follows next, with the ansatz  $u_j^n := \rho^n e^{ij\theta}$ , that the magnification factor  $\rho$  satisfies

$$(\alpha + b_0)(\rho^n - \rho^{n-1}) + \left[ b_1 \rho^{n-1} + \sum_{m=1}^{n-2} \rho^m (b_{n-m} - b_{n-m-1}) - b_{n-1} \right] + v\lambda \rho^{n-1} (1 - e^{-i\theta}) = 0. \tag{4.8}$$

Rearranging terms we get

$$(\alpha + b_0)\rho^n + (b_1 - \alpha - b_0)\rho^{n-1} + v\lambda \rho^{n-1} (1 - e^{-i\theta}) + \sum_{m=1}^{n-2} \rho^m (b_{n-m} - b_{n-m-1}) - b_{n-1} = 0. \tag{4.9}$$

Equivalently,  $\rho$  is the root of the (monic) polynomial (4.1) where, with  $\gamma_{\text{upwind}}$  as in (3.3), we have

$$a_{n-1} = \frac{1}{\alpha + b_0} (b_1 - \alpha - b_0 + \gamma_{\text{upwind}}), \tag{4.10}$$

$$a_m := \frac{1}{\alpha + b_0} (b_{n-m} - b_{n-m-1}), \quad 1 \leq m \leq n - 2, \tag{4.11}$$

$$a_0 := \frac{-b_{n-1}}{\alpha + b_0}. \tag{4.12}$$

By Lemma 4.1, to bound  $\rho \leq 1$ , and therefore to conclude the stability proof, it remains to show

$$\sum_{m=0}^{n-1} |a_m| \leq 1. \tag{4.13}$$

To this aim, by (1.3) and (1.4) we have  $b_m \geq 0$  and  $b_m - b_{m+1} \geq 0$ . Thus, for any  $0 \leq m \leq n - 1$  we have

$$\begin{aligned} \sum_{m=0}^{n-1} |a_m| &= \sum_{m=0}^{n-2} |a_m| + |a_{n-1}| = \frac{1}{\alpha + b_0} [|-b_{n-1}| + |b_{n-1} - b_{n-2}| + \dots + |b_2 - b_1|] + |a_{n-1}| \\ &= \frac{1}{\alpha + b_0} [b_{n-1} + b_{n-2} - b_{n-1} + \dots + b_1 - b_2] + |a_{n-1}| \\ &= \frac{1}{\alpha + b_0} (b_1 + |\alpha + b_0 - b_1 - \gamma_{\text{upwind}}|). \end{aligned} \tag{4.14}$$

Now (4.13) is equivalent to demonstrating

$$|\alpha + b_0 - b_1 - \gamma_{\text{upwind}}| \leq \alpha + b_0 - b_1. \tag{4.15}$$

Because  $\alpha + b_0 - b_1 \geq b_0 - b_1 \geq 0$ , (4.15) can only be true if  $\gamma_{\text{upwind}}$  has a nontrivial real part satisfying an additional condition to be determined.

First, we note that the assumptions in (i) and (ii) imply for both  $\alpha = 0$  and  $\alpha = 1$

$$v\lambda \leq \alpha + b_0 - b_1. \tag{4.16}$$

For  $\alpha = 1$ , we see that  $v\lambda \leq 1$  is sufficient for  $|1 - \gamma_{\text{upwind}}| \leq 1$ . Now (4.16) suffices for  $\frac{v\lambda}{1+b_0-b_1} \leq 1$  and  $|1 - \frac{\gamma_{\text{upwind}}}{1+b_0-b_1}| \leq 1$ , thus also for (4.15), and (i) is proved.

For  $\alpha = 0$ , similarly, (4.15) is equivalent to  $|1 - \frac{\gamma_{\text{upwind}}}{b_0-b_1}| \leq 1$  as long as  $b_0 - b_1 > 0$  which holds if  $\beta \notin \mathcal{B}_0$  and in particular if  $\beta \in \mathcal{B}_\infty$ . Thus  $\frac{v\lambda}{b_0-b_1} \leq 1$  leads to (4.13) and (ii) is proved. Superficially, it appears that  $\beta \in \mathcal{B}_b$  suffices for (ii); however, the stability condition cannot be satisfied; see Remark 1.

To prove (iii), we proceed similarly as before and find that the magnification factor  $\rho$  is the root of (4.1) with the coefficients  $a_m$  given for the implicit scheme as

$$a_{n-1} = \frac{1}{\alpha + b_0 + \gamma_{\text{upwind}}} (b_1 - \alpha - b_0), \tag{4.17}$$

$$a_m := \frac{1}{\alpha + b_0 + \gamma_{\text{upwind}}} (b_{n-m} - b_{n-m-1}), \quad 1 \leq m \leq n - 2, \tag{4.18}$$

$$a_0 := \frac{-b_{n-1}}{\alpha + b_0 + \gamma_{\text{upwind}}}. \tag{4.19}$$

To show a statement analogous to (4.13), we telescope the terms and see that one needs to show  $0 < \alpha + b_0 \leq |\alpha + b_0 + \gamma_{\text{upwind}}|$ . But this follows directly and unconditionally from (3.6) for the upwind scheme (and for centered scheme) as long as  $\alpha + b_0 > 0$  which is guaranteed for each of the two cases in hypotheses of (iii). ■

**Remark 1.** *To interpret the stability condition in Theorem 4.2 (ii), for  $\alpha = 0$ , consider first  $\beta \in \mathcal{B}_\infty$  and in particular;  $\beta(t) = t^r$ ,  $r \in (-1, 0)$ . Then*

$$b_0 - b_1 = \int_0^k \beta(t)dt - \int_k^{2k} \beta(t)dt = \frac{2 - 2^{r+1}}{r + 1} k^{r+1}. \tag{4.20}$$

Therefore, in order to satisfy

$$v \frac{k}{h} = v\lambda \leq b_0 - b_1, \tag{4.21}$$

we need  $g_r(k) := k^{r+1} \leq C_r \frac{h}{v}$  with some constant  $C_r > 0$  dependent on  $r$ . This can be done because  $\lim_{k \rightarrow 0} g_r(k) = 0$ . In particular, for  $r = -1/2$ , we see that  $\sqrt{k} \leq C_r \frac{h}{v}$  or, equivalently, that  $k$  needs to scale with  $O(h^2)$ . This condition, albeit restrictive, can be satisfied in practical computations.

We note that the condition  $k = O(h^2)$  derived for the fractional conservation law with  $r = 1/2$  is reminiscent of stability conditions for explicit finite difference discretizations of a diffusion equation. The latter can be considered a “square” of the former and thus the stability condition is not surprising.

**Remark 2.** *Now consider  $\beta \in \mathcal{B}_b$ , for example,  $\beta(t) = \exp(-t)$ . Here, we find that for (4.21) to hold, one needs  $g(k) = \frac{k}{1 - 2e^{-k} + e^{-2k}} \leq \frac{h}{v}$  to hold. Now  $g(k)$  is unbounded at 0 thus (4.21) cannot hold with  $h, k \rightarrow 0$ . This illustrates why the condition  $\beta \in \mathcal{B}_\infty$  is necessary in Theorem 4.2, (ii).*

## V. STABILITY FOR NONLINEAR FLUX

Consider (1.9a) with nonlinear  $f(\cdot)$ . For conservation laws with no memory, that is,  $\beta \equiv 0$ , various discretizations have been formulated. In convergent schemes, approximations to the flux term  $f(u)_x$  must be consistent and conservative. In this article, we use first-order Godunov’s method which is well-known to possess such properties [14–16] and which exhibits first-order convergence for smooth solutions.

For simplicity in what follows we assume  $f(\cdot)$  is increasing and convex, for example, as in Burger’s equation  $f(u) = \frac{u^2}{2}$ . For such  $f(\cdot)$ , the Godunov method is the same as the nonlinear version of the explicit upwind method.

For nonlinear scalar conservation law with memory, we extend easily (3.10) to the Godunov-memory scheme

$$\alpha(u_j^n - u_j^{n-1}) + \sum_{m=1}^n (u_j^m - u_j^{m-1})b_{n-m} + \lambda(f(u_j^{n-1}) - f(u_{j-1}^{n-1})) = 0. \tag{5.1}$$

The stability of (5.1) is assessed differently than for the linear  $f(\cdot)$  in Theorem 4.2, and the analysis involves the total variation of a solution

$$TV_T(U) := \sum_{n=1}^N kTV(U^n) + \|U^n - U^{n-1}\|_1,$$

whose spatial part  $TV(U^n) := \sum_j |u_j^n - u_{j-1}^n|$ , and the temporal part  $\|U^n - U^{n-1}\|_1$  are bounded independently.

The boundedness of  $TV_T(U)$  allows to take advantage of the compactness properties of (sets of) solutions with bounded total variation and with compactly supported initial data to prove convergence of numerical schemes, see, for example, [14], [Theorem 15.2]. Below, we prove this stability property for  $\alpha = 1$  and weakly singular kernels and demonstrate other results applicable to more general circumstances. The proof is achieved in a sequence of auxiliary results.

**A. Estimate Variation in Space**

When no memory terms are present, one can prove  $l_1$ -contractiveness of the scheme which implies that the total variation is diminishing which in turn implies the TV-stability, see [15].

When  $\beta \neq 0$ , the contracting type bounds do not hold. In particular,  $TV(U^n) \leq TV(U^{n-1})$  may not be true because  $TV(U^n)$  depends on the entire sequence

$$TV(U^m), m = 0, 1, \dots = n - 1.$$

We are, however, able to show that  $TV(U^n) \leq TV(U^0)$ , thus is uniformly bounded.

**Lemma 5.1.** *Given some  $W^0$ , consider  $W^n = (w_j^n)_j, n \geq 1$  satisfying*

$$\alpha(w_j^n - w_j^{n-1}) + \sum_{m=1}^n (w_j^m - w_j^{m-1})b_{n-m} + \lambda v_{j,n} w_j^{n-1} - \lambda v_{j-1,n} w_{j-1}^{n-1} = 0, \tag{5.2}$$

where  $(v_{j,n})_j$  are given.

Assume (i)  $b_m \geq 0, \forall m$ , (ii)  $0 \leq \lambda v_{j,n} \leq \alpha + b_0 - b_1, \forall j, n$ , (iii)  $b_m - b_{m-1} \geq 0, \forall m$ . Then  $W^n$  satisfies

$$\sum_j |w_j^n| \leq C_0 := \sum_j |w_j^0|. \tag{5.3}$$

**Proof.** We first demonstrate the assertion for  $v_{j,n}$  constant in  $n$  and in the end we comment on the general case.

Proof of (5.3) is by induction. For  $n = 1$  (5.2) reads

$$\alpha(w_j^1 - w_j^0) + (w_j^1 - w_j^0)b_0 + \lambda(v_j w_j^0 - v_{j-1} w_{j-1}^0) = 0.$$

By (i), (ii) we can estimate

$$(\alpha + b_0)|w_j^1| \leq |w_j^0|(\alpha + b_0 - \lambda v_j) + \lambda v_{j-1}|w_{j-1}^0|$$

Summing over  $j$  and multiplying by  $h$  we get

$$\begin{aligned} (\alpha + b_0) \sum_j |w_j^1| &\leq \sum_j |w_j^0|(\alpha + b_0 - \lambda v_j) + \sum_j \lambda v_{j-1}|w_{j-1}^0|. \\ &= \sum_j |w_j^0|(\alpha + b_0) - \sum_j |w_j^0|\lambda v_j + \sum_j \lambda v_{j-1}|w_{j-1}^0|. \end{aligned}$$

We can cancel the (equal) sums involving  $v_j$  and  $v_{j-1}$  in the last identity because  $w_j^0$  is zero outside a finite set of indices. This means we proved (5.3) for  $n = 1$ .

Now, assuming (5.3) holds for  $m = 1, 2, \dots, n - 1$ , we demonstrate it for  $m = n$ . Rewrite (5.2) taking advantage of (3.8)

$$\begin{aligned} (\alpha + b_0)w_j^n &= \alpha w_j^{n-1} + (b_0 - b_1)w_j^{n-1} + (b_1 - b_2)w_j^{n-2} + \dots + (b_{n-2} - b_{n-1})w_j^1 + b_{n-1}w_j^0 \\ &\quad - [\lambda v_j w_j^{n-1} - \lambda v_{j-1} w_{j-1}^{n-1}]. \end{aligned} \tag{5.4}$$

To estimate, we group the first and last terms on the right hand side, exploit (i), (ii), (iii), and sum over  $j$

$$\begin{aligned} (\alpha + b_0) \sum_j |w_j^n| &\leq \sum_j |w_j^{n-1}|(\alpha + b_0 - b_1 - \lambda v_j) + \sum_j \lambda v_{j-1}|w_{j-1}^{n-1}| \\ &\quad + (b_1 - b_2) \sum_j |w_j^{n-2}| + \dots + (b_{n-2} - b_{n-1}) \sum_j |w_j^1| + b_{n-1} \sum_j |w_j^0| \end{aligned} \tag{5.5}$$

Similarly as above, we collapse together the first two sums. Also, by induction assumption,  $\sum_j |w_j^m| \leq C_0$  for each  $m$ . Collecting terms we get

$$(\alpha + b_0) \sum_j |w_j^n| \leq C_0(\alpha + b_0 - b_1 + b_1 - b_2 + \dots b_{n-2} - b_{n-1} + b_{n-1}) = C_0(\alpha + b_0),$$

which ends the proof for  $v_{j,n}$  constant in  $n$ . To complete the proof for general  $v_{j,n}$  we notice that the steps above do not depend on its variability between time steps. ■

Now we use Lemma 5.1 to prove that  $TV(U^n)$  is bounded uniformly for all  $n$ . We follow calculations and steps similar to those in the proof of Ref. [[14], Chapter 15].

**Lemma 5.2.** *Let  $f$  be increasing and strictly convex. Also, let  $\beta$  satisfy (1.3)–(1.5) and  $\beta \in \mathcal{B}_\infty \cup \mathcal{B}_b \setminus \mathcal{B}_0$ . Let  $U^n, \tilde{U}^n$  be defined by (5.1), each with separate unital data  $U^0, \tilde{U}^0$ , respectively. Let  $M > 0$  be given depending on  $U^0, \tilde{U}^0$ . Assume also that the time step  $k$  satisfies*

$$0 < \lambda f'(\theta) \leq 1, \forall \theta : |\theta| \leq M. \tag{5.6}$$

Then

$$\sum_j |u_j^n - \tilde{u}_{j-1}^n| \leq \sum_j |u_j^0 - \tilde{u}_{j-1}^0|. \tag{5.7}$$

**Proof.** We provide only a sketch as the proof is similar to those in Ref. [[14], Example 15.6]. First, we note that the assumptions on  $\beta$  imply (i),(iii) in Lemma 5.1. To estimate the difference  $U^n - \tilde{U}^n$ , we subtract the two difference equations defining these grid functions. Now consider the terms  $f(u_j^{n-1}) - f(\tilde{u}_j^{n-1})$  that appear in the result and apply mean value theorem yielding  $f(u_j^{n-1}) - f(\tilde{u}_j^{n-1}) = v_j(u_j^{n-1} - \tilde{u}_j^{n-1})$ , where  $v_j = f'(\theta_j)$  for some  $\theta_j$  between  $u_j^{n-1}$  and  $\tilde{u}_j^{n-1}$ . To see that  $\theta_j$  is bounded, we can proceed by induction considering its value at  $n = 1$  depending on  $u_j^0$  and  $\tilde{u}_j^0$ . By (5.6), the assumption (ii) in Lemma 5.1 holds. Thus, Lemma 5.1 can be applied to  $W = U^n - \tilde{U}^n$  which completes the proof. ■

Now let  $U^n$  be the approximation defined by (5.1) corresponding to the original initial value problem (1.9a) and (1.9b), with initial data  $U^0$  obtained from  $u_0$ . Let  $\tilde{U}^n$  be a grid function obtained from  $U^n$  by shifting indices, that is, by  $\tilde{u}_j^n = u_{j-1}^n$ . Now, we use Lemma 5.2 to get a bound on  $TV(U^n)$ .

**Corollary 5.3.** *Assume hypotheses of Lemma 5.2. Then*

$$TV(U^n) \leq TV(U^0). \tag{5.8}$$

**B. Estimate Variation in Time**

Now, we obtain an  $O(k)$  bound for  $\| U^n - U^{n-1} \|_1$ ; this is necessary to show that the total variation is uniformly bounded.

For  $\beta \equiv 0$ , a simple proof via Gronwall’s-type inequality is provided in Ref. [[14], Lemma 15.1].

When  $\beta \not\equiv 0$ , this simple strategy does not work because the presence of memory terms causes a tremendous accumulation of error. To show this need not occur, we exploit the special structure arising from the Volterra integral with weakly singular kernels. We use delicate discrete inequalities that were proved in, for example, [36, 37]. Specifically, we take advantage of the result [[37], Theorem 1] which was formulated for the needs of convergence analysis of discretized weakly singular Volterra integral equation; see also Ref. [[38], Theorem 6.1]. We extract the relevant part of the proof of this result in the following Lemma.

**Lemma 5.4** (Part of Theorem 1 Ref. [37]). *Let  $(e_n)_{n=1}^N$  be a nonnegative sequence of numbers for which it holds*

$$e_n \leq A + \sum_{m=1}^{n-1} B_{nm} e_m, \quad 1 \leq m \leq N, \tag{5.9}$$

where  $A > 0$  and  $B_{nm} := Ck(t_n - t_m)^r$  for some  $-1 < r < 0$  and  $C > 0$ . Then

$$e_n \leq A \sum_{m=0}^{n-1} \frac{R^m}{(m!)^{1+r}} \leq A\mathcal{H}, \tag{5.10}$$

where  $R > 0, \mathcal{H} > 0$  are constants independent of  $k$  but dependent on  $r$  and  $T = Nk$ .

This Lemma will be used in the stability proof below for  $\beta(t) = t^r$ .

We assume below as is customarily done that  $f(\cdot)$  is globally Lipschitz continuous. This is not satisfied, for example, for Burgers' flux  $f(u) = \frac{u^2}{2}$  if its arguments live outside a compact set. However, one can modify easily the line of reasoning in Ref. [[14], Section 15] to see that (5.8) implies that  $U^n$  is bounded and thus a hypothesis of global Lipschitz constant for  $f$  acting on arguments from  $U^n$  is justified.

**Proposition 5.5.** *Assume that the flux function  $f(\cdot)$  has a Lipschitz constant  $L_f$  and that  $\beta(t) = t^r$  for some  $-1 < r < 0$ . Assume also  $\alpha = 1$ . Then*

$$\| U^n - U^{n-1} \|_1 \leq C_T k, \tag{5.11}$$

where  $C_T > 0$  is independent of  $k$  but dependent on  $f$ .

**Proof.** To apply the construction apparent in Lemma 5.4, we first consider  $n = 1$  in (5.1) where we have

$$(\alpha + b_0)(u_j^1 - u_j^0) = -\lambda(f(u_j^0) - f(u_{j-1}^0)).$$

Next, we proceed similarly as in one step of Ref. [[14], Lemma 15.1]. Since  $f$  is Lipschitz continuous, we have

$$(\alpha + b_0)|u_j^1 - u_j^0| \leq \lambda L_f |u_j^0 - u_{j-1}^0| = \frac{k}{h} L_f |u_j^0 - u_{j-1}^0|.$$

Sum over  $j$ , multiply by  $h$ , and divide by  $\alpha + b_0$  to get

$$\| U^1 - U^0 \|_1 = h \sum_j |u_j^1 - u_j^0| \leq \frac{kL_f}{\alpha + b_0} |u_j^0 - u_{j-1}^0| = \frac{kL_f}{\alpha + b_0} TV(U^0),$$

and we set  $e_1 := \| U^1 - U^0 \|_1, A := kL_f TV(U^0)$ .

Now for  $n > 1$  we rewrite (5.1) to get

$$(\alpha + b_0)(u_j^n - u_j^{n-1}) = -\lambda(f(u_j^{n-1}) - f(u_{j-1}^{n-1})) - \sum_{m=1}^{n-1} b_{n-m}(u_j^m - u_j^{m-1}).$$

Now we sum over  $j$ , multiply by  $h$ , and estimate

$$(\alpha + b_0)h \sum_j |u_j^n - u_j^{n-1}| \leq \lambda h L_f \sum_j |u_j^{n-1} - u_{j-1}^{n-1}| + \sum_{m=1}^{n-1} b_{n-m} h \sum_j |u_j^m - u_j^{m-1}|.$$

By (5.8), we get

$$(\alpha + b_0) \| U^n - U^{n-1} \|_1 \leq kL_f TV(U^0) + \sum_{m=1}^{n-1} b_{n-m} \| U^m - U^{m-1} \|_1.$$

To apply Lemma 5.4, we need now to check if we can bound  $b_{n-m}$  by  $B_{n,m} = Ck(t_n - t_m)^r$ . This follows directly by estimating  $b_{n-m} = \int_{t_{n-m}}^{t_{n-m+1}} t^r dt$  from above by the left rectangular rule  $b_{n-m} \leq k(t_{n-m})^r = k^{1+r}(n - m)^r$  and we are done. ■

At the current time, we are unable to provide a proof for an arbitrary  $\beta \in \mathcal{B}_\infty$  or  $\alpha = 0$  so that Proposition 5.5 can be extended. However, our numerical experiments (not reported) indicate that  $\| U^n - U^{n-1} \|_1$  is of order  $k$  for any kernel and for  $\alpha = 0$ .

**C. Summary of Nonlinear Stability Results**

Taking advantage of Corollary 5.3 and Proposition 5.5, we see that the total variation of the numerical approximation to (1.9) is bounded, at least for the case of weakly singular kernels of type  $t^r$ , and  $\alpha = 1$ .

**Theorem 5.6.** *Let the assumptions of Corollary 5.3 and Proposition 5.5 hold. Then the solution  $U^n$  to the numerical scheme (5.1) satisfies*

$$\sum_{n=1}^N kTV(U^n) + \| U^n - U^{n-1} \|_1 \leq C_{TVS}. \tag{5.12}$$

The bound (5.12) combined with the knowledge of compact support of  $U^n$  together mean that the scheme (5.1) is TV-stable. Although we do not provide a formal proof concerning the support of  $U^n$ , we remark that it can be carried out mimicking any such proof for  $\beta \equiv 0$ .

**VI. NUMERICAL RESULTS**

In this section, we confirm the first-order convergence rate expected for the discretization schemes, also for cases not covered by our theory. We also illustrate smoothing effects associated with memory terms.

We provide results only for explicit schemes. As we have shown and as is easy to believe, the implicit schemes are unconditionally stable. For problems without memory, they are rarely used because they require significantly more computational work, except when the benefits of unconditional stability counter those of cost per time step. In addition, it is well-known that the implicit schemes introduce errors due to numerical diffusion much larger than those in explicit schemes; see, for example, Ref. [[17], Section 2.3]. In our experience, for example, with an implicit scheme for Burgers' equation with no memory, the fronts are substantially smeared out even for a relatively fine grid. Because one of the objectives in this article is to illustrate the delicate interplay between nonlinearity and smoothing effects due to memory terms, we choose to use only explicit schemes. Otherwise, the use of implicit schemes would make it hard to recognize the source of the smoothing between numerical diffusion and subscale diffusion encapsulated by memory terms.

**A. Set-Up of Experiments**

The results presented below for explicit schemes vary depending on the smoothness of the solutions which, in turn, are directly related to the choice of the initial data  $u_0$ , flux function  $f$ , and memory kernel  $\beta$ . We consider several representative choices.

Linear flux examples use  $f(u) = u$  and the nonlinear flux is the Burger’s flux  $f(u) = \frac{u^2}{2}$ . For the latter, as long as  $\max_x(|u_0(x)|) = 1$ , we have conveniently that the global Lipschitz constant  $V = \max |f'(u_0)| = 1$  for both choices of  $f$ .

For the memory kernels  $\beta$ , we use one of three choices of  $\beta_i, i = 0, 1, 2$  with  $\beta_1(t) = e^{-t} \in \mathcal{B}_b$  or  $\beta_2(t) = t^{-1/2} \in \mathcal{B}_\infty$ . For comparison purposes, we also consider  $\beta_0(t) = 0 \in \mathcal{B}_0$ .

We consider initial data  $u_0 = g_k, k = 1, \dots, 4$ , with  $g_1(x) = \sin(x)$ ,  $g_2(x) = \max(0, 1 - |x|)$ ,  $g_3(x) = 1 - H(x)$ ,  $g_4(x) = H(x)$ , where  $H(x)$  is the Heaviside function. Although  $g_1 \in C^\infty(\mathbb{R})$  we have  $g_2 \in C^0(\mathbb{R})$  and is piecewise  $C^1$ -smooth, whereas  $g_3, g_4$  are only piecewise  $C^0$ -smooth. The initial data is in fact truncated to have compact support. When computing the grid norm, we only consider  $x_j \in (a, b)$  with  $(a, b) = (0, 2\pi)$  for  $g_1$ , and  $(a, b) = (-1, 3)$  for all other choices. We set  $L = b - a$  to denote the length of the simulation window.

In all experiments, we ensure that the stability condition is satisfied. In fact, if  $\alpha = 1$ , we use  $k = 0.99h/V$ , that is, the time step close to but not equal to the “magic” time step  $k = h/V$ . If  $\alpha = 0$  and  $\beta = \beta_2$ , we use  $k = \frac{1}{2}(\frac{h}{V})^2$ .

If the true solution  $u(x, t)$  is known, we compare it to  $u_h$  constructed from  $(u^n_j)_{j,n}$ . In fact, we report only the error at  $t_N = T$  and consider various error norms of  $u(x_j, t_n) - u^n_j$ . We denote the error by  $E_p(\alpha, g, \beta; h)$  for  $p = 1, 2$  grid norms. For additional interest, we report on the  $E_\infty(\alpha, g, \beta; h)$  error in  $L_\infty$  grid norm for which no theory is given but which for smooth enough  $u$  behaves similarly to the errors measured in  $L_1$  and  $L_2$  grid norms. The order of convergence is calculated by comparison of errors from subsequent time steps. The error is shown to be of linear order if the solutions are smooth enough. In the convergence plots, we include the plots of  $h$  versus  $h$  labeled “linear” and of  $\sqrt{h}$  labeled “root”.

If the true solution  $u(x, t)$  is not known, we use instead the solution  $u_{h_{\min}}$  computed at a very fine grid with  $h_{\min}$  as given below. Ideally,  $h/h_{\min}$  should be at least an order of magnitude. However, in some examples below for  $\alpha = 0$ , we cannot afford a very small  $h_{\min}$  due to the stability constraints. Therefore, we report also on experiments in which  $h/h_{\min} \approx 3$ .

**B. Convergence Rate for Linear Flux  $f$  with Known  $u$**

Here, we let the analytical solution  $u(x, t) = \sin(x - t)$  be given so that the initial data is  $u_0 = g_1$ . We compute the right-hand side function in an extension to (1.9a)

$$\alpha u_t + \beta * u_t + \nu u_x = F(x, t). \tag{6.1}$$

In the implementation of (4.3), we include appropriately the term  $F^n_j = F(x_j, t^n)$  to account for the right hand side.

Because the solution  $u$  is  $C^\infty$  smooth, a very good behavior of convergence is expected. To confirm, consider first the conservation law with memory that is  $\alpha = 1$ . For  $\beta = \beta_1$ , we can compute the right hand side  $F(x, t) = u_t + \beta * u_t + u_x$  with

$$\beta * u_t = -\frac{1}{2} (\sin(x)(\sin(t) - \cos(t) + e^{-t}) + \cos(x)(\cos(t) + \sin(t) - e^{-t})).$$

For the weakly singular  $\beta = \beta_2 \in \mathcal{B}_\infty$ , the convolution integral in  $\beta * u_t$  is not elementary; we compute  $F(x, t)$  numerically with Gauss–Kronrod’s quadrature in MATLAB®’s function `quadgk`.

TABLE I. Convergence results for linear flux,  $u_0 = g_1$ , and  $\alpha = 1$  (top), and  $\alpha = 0$  (bottom).

$L/h$	$E_p(1, g_1, \beta_1; h)$			Order	$E_p(1, g_1, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
100	0.03685	0.03548	0.04032		0.02882	0.02378	0.03538	
200	0.01342	0.01219	0.01837	1.1021	0.01860	0.01804	0.02055	0.9865
300	0.00863	0.00819	0.01241	1.0908	0.01243	0.01209	0.01379	0.9938
400	0.00632	0.00618	0.00937	1.0837	0.00933	0.00910	0.01038	0.9972
500	0.00496	0.00496	0.00753	1.0789	0.00746	0.00729	0.00832	0.9995

$L/h$	$E_p(0, g_1, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
100	0.06349	0.02943	0.01765	
200	0.03318	0.01533	0.00917	0.9362
300	0.02245	0.01036	0.00619	0.9635
400	0.01696	0.00782	0.00467	0.9743
500	0.01363	0.00628	0.00376	0.9800

When  $\alpha = 0$ , we compute  $F$  similarly; we only use weakly singular  $\beta = \beta_2$  for which the scheme is stable.

For all the cases reported here, we confirm linear order of convergence of the scheme (4.3) for  $h$  small enough, see Table I and Figure 1 for evidence. As  $h$  decreases, the order becomes close to 1.

The results reported for this example are entirely unsurprising but serve as a good reference for the more interesting cases to follow.

**C. Convergence Rate for Linear Flux with Unknown  $u$**

Now, we use  $F = 0$ ,  $L = 4$ , and initial data  $g_2$  or  $g_3$ . Because the true solution is not known, we approximate it with  $u_{h_{\min}}$  with  $h_{\min} = L/5000$  except when  $\alpha = 0$  in which case  $h_{\min} = L/500$ .

When computing the error at  $t = T$ , an additional consideration is necessary. Because  $k$  is chosen automatically for each  $h$ , and as only uniform time stepping is used, we have to interpolate between the time steps to find  $u_h|_{t=T}$  for some choices of  $h$ . Thus, the error behavior can be somewhat poor for large  $h$ . This is easily understood once we realize the deficiency of interpolation in time, for example, for the travelling wave solution. For reference, the error computed this way for  $\beta_0$  and  $g_2$  has a ragged behavior until  $h \leq L/1000$ , whereas the use of a known analytical solution rewards us with a perfect linear order of convergence.

Consider now the results in Figure 2 for piecewise smooth initial data  $g_2$  with the corresponding numerical evidence in Table II. These confirm that the order of convergence is linear for  $E_1$  and  $E_2$ . In fact, when  $\alpha = 1$ , the average order in  $E_1$  is 1.13 and 1.14 for  $\beta_1$  and  $\beta_2$ , respectively. When  $\alpha = 0$ , the average order is 1.0762. We note that in the latter case we can afford only a small number of experiments and relatively large  $h$  due to the suboptimal stability requirements on  $k$ .

Equally interesting is the linearly convergent behavior of the error in  $L_\infty$  grid norm for the weakly singular kernel  $\beta_2$ . This appears to suggest that the true solution (which we do not know) is smoother than the initial data. This does not occur for bounded kernel  $\beta_1$  for which  $E_\infty(\cdot; h)$  appears to behave like  $O(\sqrt{h})$ .

Our next set of results is for piecewise constant initial data  $g_3$ ; see Figure 3 and Table III. With the bounded kernel  $\beta_1$ , the solution is not smooth enough and we cannot expect linear order of

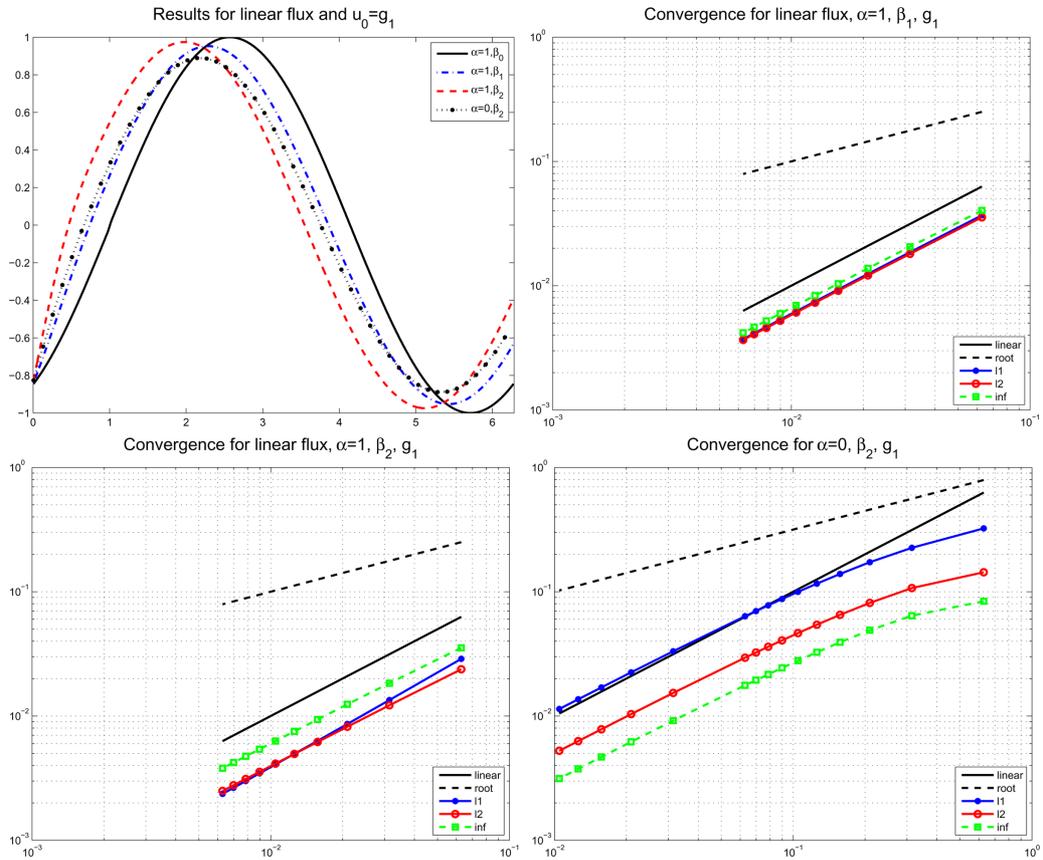


FIG. 1. Results for linear flux and  $u_0 = g_1$ . Top left: plot of numerical solutions at  $T = 1$ , computed with  $L/h = 500$ . Other figures: plots of  $E_1(\cdot, h)$ ,  $E_2(\cdot, h)$ ,  $E_\infty(\cdot, h)$  for different data. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

convergence in  $E_2$ , and the error  $E_\infty$  may not even converge to zero. In fact, we observe sublinear order of convergence for small  $h$  in  $E_1(\cdot, h)$  and  $E_2(\cdot, h)$ . However, when  $\beta_2$  is used, we observe linear order of convergence in all  $E_1, E_2, E_\infty$ , suggesting that  $u$  is indeed quite smooth; the latter is also suggested by the plots of solutions.

**D. Nonlinear Flux Examples**

Here, we use  $F = 0$ ,  $L = 4$ , and  $h_{\min} = L/5000$  if  $\alpha = 1$ , and  $h_{\min} = L/500$  if  $\alpha = 0$ . When computing the error, we use  $u_{h_{\min}}$  instead of  $u$  which is unknown.

It is well-known [14] that for nonlinear flux  $f$  and no memory effects  $\beta \equiv 0$ , even for smooth initial data, the solution  $u$  to (1.9) develops discontinuities in finite time. Therefore, optimal linear convergence is generally not expected around the singularities, however, subtle differences appear depending on the initial data; see convergence plots in Figure 4 and plots of corresponding solutions included in Figures 5 and Figures 6.

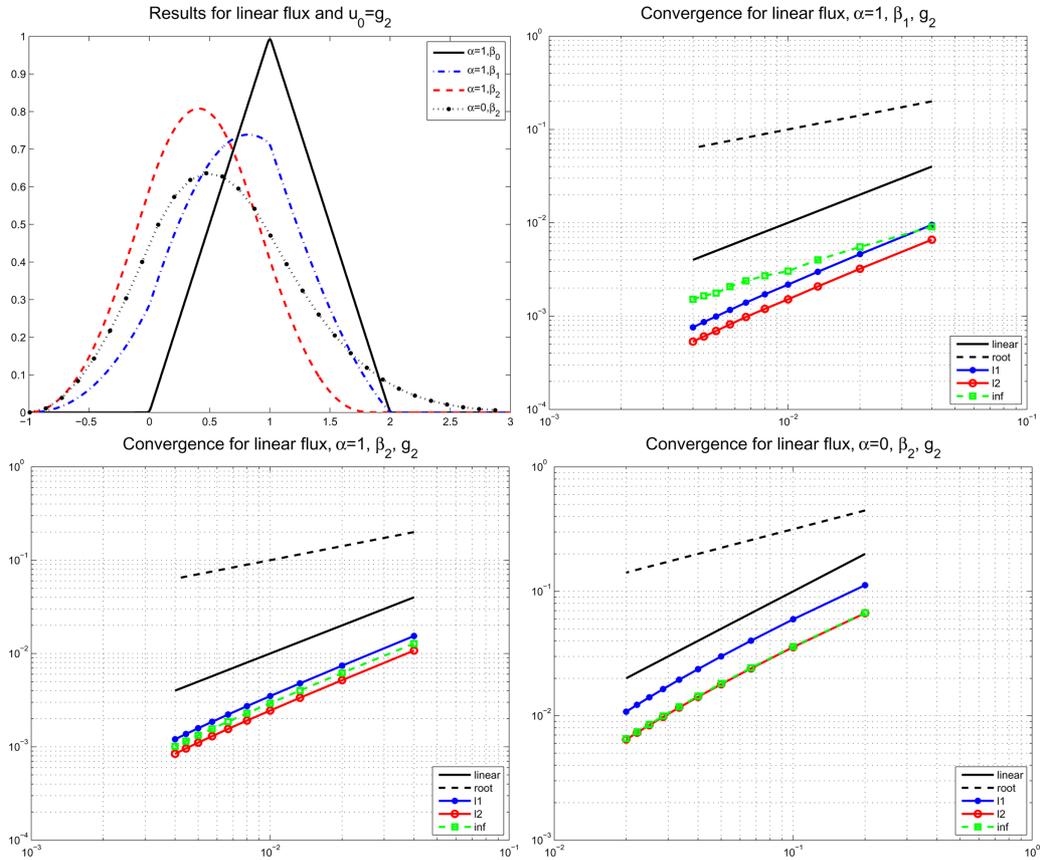


FIG. 2. Results for linear flux and  $u_0 = g_2$ . Top left: plot of numerical solutions at  $T = 1$ , computed with  $L/h = 500$  for  $\alpha = 1$  and  $L/h = 100$  for  $\alpha = 0$ . Other figures: plots of  $E_1(\cdot, h)$ ,  $E_2(\cdot, h)$ ,  $E_\infty(\cdot, h)$  for different data. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

If memory effects are present, with  $g_3$  as initial data and nonlinear flux, some singularities likely persist according to Ref. [18]. From theory, it is not clear what smoothing and what convergence rates can be expected. However, our numerical experiments and convergence plots offer a hint to what could be expected.

The plots in Figure 5 and 6 show the solutions to (5.1) and convergence plots for piecewise smooth initial data  $g_2$ , and piecewise constant data  $g_3$ , respectively. See also numerical evidence in Tables IV and V, respectively. We discuss these results below.

The case of piecewise smooth  $g_2$  appears to be qualitatively similar to that of  $g_1$ , even if bounded memory kernel  $\beta_1$  is used, and we observe optimal linear convergence properties of the scheme (5.1).

For piecewise constant initial data  $g_3$ , one is led to believe that the solutions to (5.1) for  $\alpha = 1, \beta_1$  or  $\alpha = 0, \beta_2$  may not be smooth enough for optimal convergence. This is confirmed for profiles evolving from  $g_3$ , where sublinear convergence is observed for  $E_2$  and poor behavior of  $E_\infty$  is observed. However,  $E_1$  appears to have at least linear convergence, even for bounded kernels.

TABLE II. Convergence results for linear flux,  $u_0 = g_2$ , and  $\alpha = 1$  (top), and  $\alpha = 0$  (bottom).

$L/h$	$E_p(1, g_2, \beta_1; h)$			Order	$E_p(1, g_2, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
100	0.00949	0.00658	0.00905		0.01538	0.01074	0.01277	
200	0.00461	0.00321	0.00550	1.0417	0.00740	0.00518	0.00619	1.0553
300	0.00298	0.00208	0.00400	1.0729	0.00479	0.00335	0.00401	1.0742
400	0.00217	0.00151	0.00303	1.1042	0.00349	0.00244	0.00293	1.0962
500	0.00171	0.00119	0.00270	1.0746	0.00273	0.00191	0.00229	1.1071

$L/h$	$E_p(0, g_2, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
20	0.11196	0.06686	0.06725	
40	0.05957	0.03562	0.03595	0.9104
60	0.04014	0.02400	0.02430	0.9737
80	0.02998	0.01792	0.01816	1.0146
100	0.02371	0.01417	0.01437	1.0512

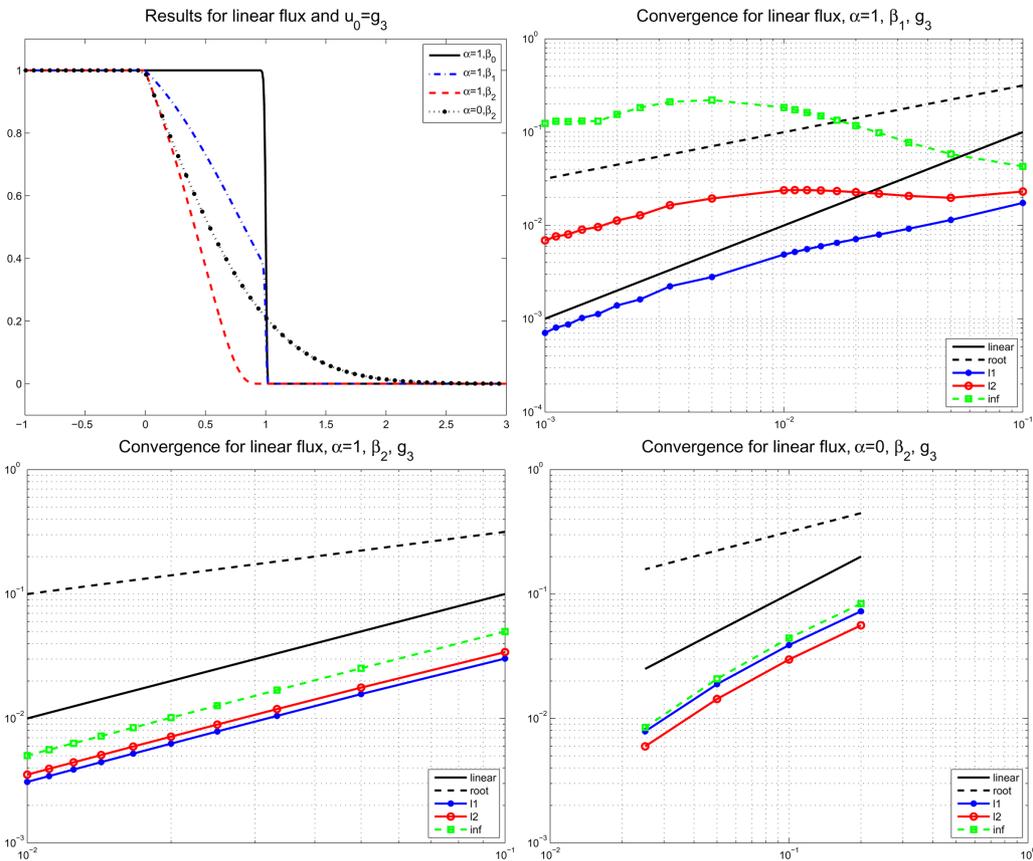


FIG. 3. Results for linear flux and  $u_0 = g_3$ . Top left: plot of numerical solutions at  $T = 1$ , computed with  $L/h = 800$  for  $\alpha = 1$  and  $L/h = 160$  for  $\alpha = 0$ . Other figures: plots of  $E_1(\cdot, h)$ ,  $E_2(\cdot, h)$ ,  $E_\infty(\cdot, h)$  for different data. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

TABLE III. Convergence results for linear flux,  $u_0 = g_3$ , and  $\alpha = 1$  (top),  $\alpha = 0$  (bottom).

$L/h$	$E_p(1, g_3, \beta_1; h)$			Order	$E_p(1, g_3, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
100	0.01125	0.01999	0.07223		0.012396	0.01399	0.01994	
200	0.00717	0.02049	0.11238	0.6482	0.006085	0.00692	0.00982	1.0264
400	0.00374	0.01509	0.12437	0.9396	0.002902	0.00332	0.00472	1.0681
800	0.00188	0.00835	0.07808	0.9941	0.001320	0.00151	0.00216	1.1358

$L/h$	$E_p(0, g_3, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
20	0.07265	0.05601	0.08359	
40	0.03889	0.02973	0.04422	0.9014
80	0.01884	0.01432	0.02077	1.0454
160	0.00786	0.00596	0.00847	1.2603

E. Discussion

Our theoretical results show stability in both  $L_2$  and  $L_1$  grid norms for linear  $f(\cdot)$ . For nonlinear flux and weakly singular kernels with  $\alpha = 1$ , we proved stability in  $L_1$  grid norm. In all cases covered by our analysis, we confirmed that the schemes are stable and produce linearly convergent solutions. In particular, this covers the cases of  $\alpha = 1, f(u) = u, \beta \in \mathcal{B}_\infty \cup \mathcal{B}_b$ , and  $\alpha = 0, f(u) = u, \beta \in \mathcal{B}_\infty$ , and  $\alpha = 1, f(u) = \frac{u^2}{2}, \beta \in \mathcal{B}_\infty$ . In addition, convergence appears to be of linear order also in cases not covered by our theoretical results, that is  $\alpha = 1, f(u) = \frac{u^2}{2}, \beta \in \mathcal{B}_b$ , and  $\alpha = 0, f(u) = \frac{u^2}{2}, \beta \in \mathcal{B}_\infty$ .

The simulation results and convergence plots analyzed together suggest that there is sufficient smoothing involved when memory terms are present. The smoothing associated with  $\beta_2$  appears to counteract, to a certain degree, the front sharpening caused by nonlinearity of  $f$ . On the other hand, the smoothing associated with bounded memory kernel  $\beta_1$  does not seem significantly inferior to that for weakly singular  $\beta_2$  for nonlinear flux or discontinuous initial data. At the same time, the effect of memory terms with bounded kernels is weak for linear flux. Finally, there is

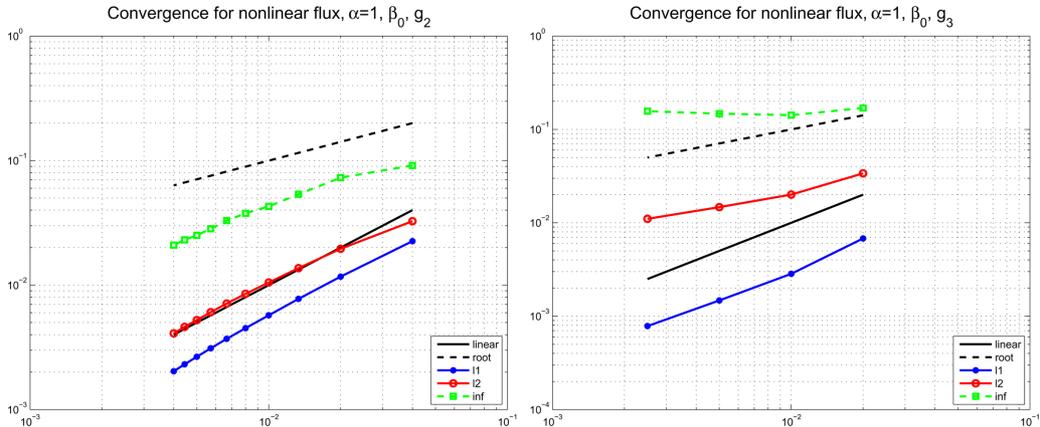


FIG. 4. Convergence results in the absence of memory effects ( $\beta = \beta_0$ ) for nonlinear flux and  $u_0 = g_2$  or  $u_0 = g_3$ . [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

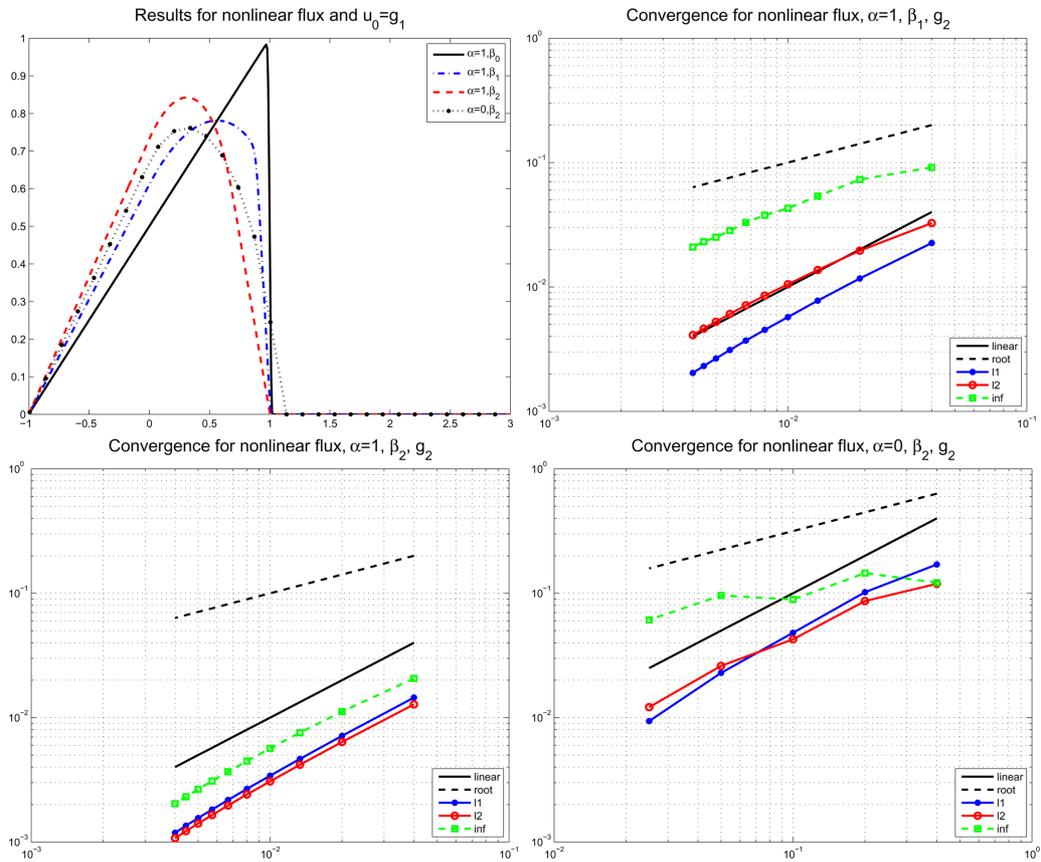


FIG. 5. Results for nonlinear flux and  $u_0 = g_2$ . Top left: plot of numerical solutions at  $T = 1$ , computed with  $L/h = 500$  for  $\alpha = 1$  and  $L/h = 160$  for  $\alpha = 0$ . Other figures: plots of  $E_1(\cdot, h)$ ,  $E_2(\cdot, h)$ ,  $E_\infty(\cdot, h)$  for different data. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

not much qualitative difference in the smoothness of solutions obtained for  $\alpha = 1$  and  $\alpha = 0$  when weakly singular memory kernels are involved, but the effect for  $\alpha = 0$  seems quantitatively stronger.

To support these remarks, we provide an additional set of simulation results for initial data  $g_4(x) = H(x)$ . See Figure 7 where we show plots of solutions for both linear and nonlinear flux. (Because the corresponding convergence results are very similar to those for  $g_2$ , we do not show nor discuss these). It is interesting, for, for example,  $\alpha = 1, \beta = \beta_1$ , to compare the solutions for the linear and nonlinear flux, as their evolution is qualitatively different. This difference parallels the one for  $\alpha = 1, \beta = \beta_0$ . In contrast, the solutions obtained with  $\beta = \beta_2$  are very similar to one another, that is, their smoothness is comparable whether or not the flux is linear.

## VII. SUMMARY

In this article, we considered scalar conservation laws with rate-dependent memory terms in the form of Volterra convolution integrals. We have indicated several applications where such partial differential equations (PDEs) are important.

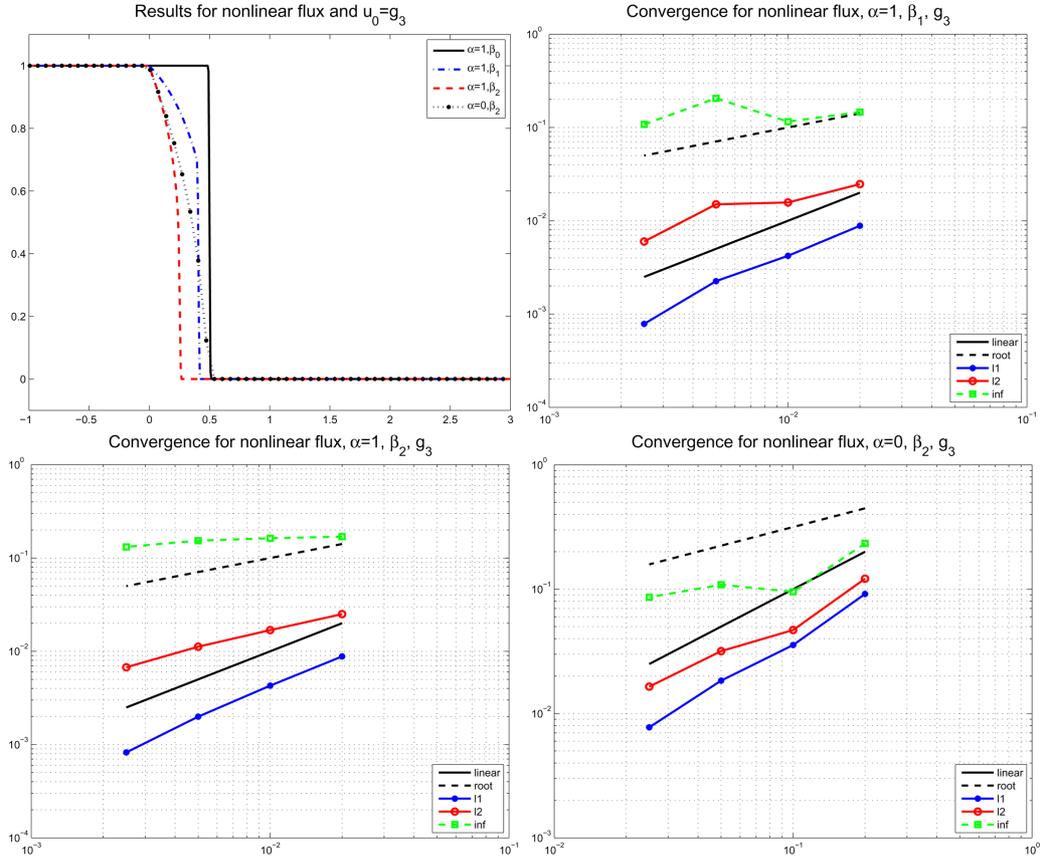


FIG. 6. Results for nonlinear flux and  $u_0 = g_3$ . Top left: plot of numerical solutions at  $T = 1$ , computed with  $L/h = 800$  for  $\alpha = 1$  and  $L/h = 160$  for  $\alpha = 0$ . Other figures: plots of  $E_1(\cdot, h)$ ,  $E_2(\cdot, h)$ ,  $E_\infty(\cdot, h)$  for different data. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

TABLE IV. Convergence results for nonlinear flux,  $u_0 = g_2$ , and  $\alpha = 1$  (top),  $\alpha = 0$  (bottom).

$L/h$	$E_p(1, g_2, \beta_1; h)$			Order	$E_p(1, g_2, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$		$p = 1$	$p = 2$	$p = \infty$	
100	0.02254	0.03261	0.09127		0.01449	0.01276	0.02065	
200	0.01169	0.01955	0.07294	0.9468	0.00713	0.00637	0.01119	1.0214
300	0.00775	0.01368	0.05369	1.0119	0.00465	0.00418	0.00756	1.0539
400	0.00572	0.01049	0.04288	1.0555	0.00341	0.00307	0.00565	1.0784
500	0.00451	0.00851	0.03770	1.0653	0.00267	0.00241	0.00447	1.0965
$L/h$	$E_p(0, g_2, \beta_2; h)$			Order				Order
	$p = 1$	$p = 2$	$p = \infty$		$p = 1$	$p = 2$	$p = \infty$	
20	0.10190	0.08651	0.14509					
40	0.04813	0.04259	0.08915					1.0820
80	0.02290	0.02605	0.09596					1.0716
160	0.00937	0.01213	0.06106					1.2879

TABLE V. Convergence results for nonlinear flux,  $u_0 = g_3$ , and  $\alpha = 1$  (top),  $\alpha = 0$  (bottom).

$L/h$	$E_p(1, g_3, \beta_1; h)$			Order	$E_p(1, g_3, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
100	0.01734	0.05182	0.24487		0.02010	0.05579	0.25935	
200	0.00882	0.02473	0.14634	0.9750	0.00881	0.02502	0.16954	1.1896
400	0.00420	0.01569	0.11548	1.0683	0.00427	0.01687	0.16302	1.0429
800	0.00225	0.01499	0.20534	0.9025	0.00199	0.01119	0.15355	1.1031

$L/h$	$E_p(0, g_3, \beta_2; h)$			Order
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$
20	0.09147	0.12129	0.23352	
40	0.03551	0.04696	0.09521	1.3648
80	0.01837	0.03174	0.10863	0.9507
160	0.00773	0.01646	0.08609	1.2481

To our knowledge, this article is the first in which a numerical scheme for such PDEs has been considered. We formulated and demonstrated convergence of simple Godunov-like schemes combined with a consistent approximation to the memory terms allowing for weak singularity of the convolution memory kernels. The conditions necessary for stability of the explicit schemes for (1.1) are comparable to, or milder than, those for conservation laws without memory. However, the conditions for fractional conservation laws, that is, for (1.8), require strict time step constraints, and might in practice call for implicit rather than explicit schemes.

Using these numerical schemes, we provided illustration of the smoothing effects that the memory terms introduce. The smoothing effects are quite dramatic if the kernel is weakly singular or for linear flux, and are less pronounced for nonlinear flux or bounded kernels. These results are consistent with the available analysis of such PDEs. In addition, our numerical experiments offered intuition concerning the nature and smoothness of solutions in cases not covered by the theory.

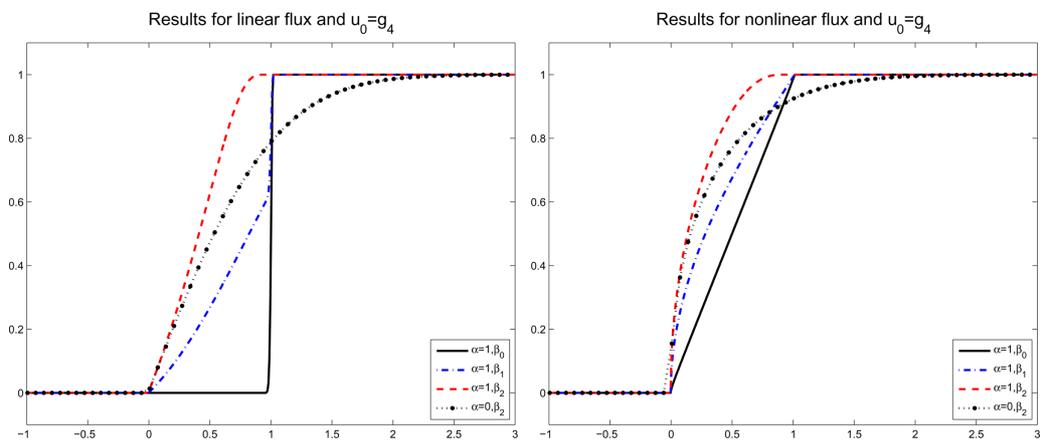


FIG. 7. Simulation results for  $u_0 = g_4$  and linear (left) and nonlinear flux (right). They were computed with  $L/h = 800$  for  $\alpha = 1$  and  $L/h = 160$  for  $\alpha = 0$ . [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

More work extending that presented here is underway. In particular, we are considering more applications-related terms and couplings, and rate-independent memory terms, as well as seek ways to improve the efficiency of the schemes. More analysis is needed to discern the fine details of smoothness of solutions to problems with memory and to analyze higher order convergent schemes.

We thank the referees for detailed comments which helped to improve this paper.

## References

1. B. Zinn, L. C. Meigs, C. F. Harvey, R. Haggerty, W. J. Peplinski, and C. F. von Schwerin, Experimental visualization of solute transport and mass transfer processes in two-dimensional conductivity fields with connected regions of high conductivity, *Environ Sci Technol* 38 (2004), 3916–3926.
2. M. Peszyńska and R. E. Showalter, Multiscale elliptic-parabolic systems for flow and transport, *Electron J Diff Equ* (2007), 30.
3. S.-Y. Yi, M. Peszyńska, and R. Showalter, Numerical upscaled model of transport with non-separated scales, *Proceedings of CMWR XVIII in Barcelona, June 21–24, 2010*, available online at <http://congress.cimne.com/CMWR2010/Proceedings>, 2010. paper 188.
4. E. Ruckenstein, A. S. Vaidyanathan, and G. R. Youngquist, Sorption by solids with bidisperse pore structures, *Chem Eng Sci* 26 (1971), 1305–1318.
5. G. R. King, T. Ertekin, and F. C. Schwerer, Numerical simulation of the transient behavior of coal-seam degasification wells, *SPE Formation Eval* 2 (1986), 165–183.
6. C. R. Clarkson and R. M. Bustin, The effect of pore structure and gas pressure upon the transport properties of coal: a laboratory and modeling study. 1, isotherms and pore volume distributions, *Fuel* 78 (1999), 1333–1344.
7. J. Q. Shi and S. Durucan, A bidisperse pore diffusion model for methane displacement desorption in coal by CO<sub>2</sub> injection, *Fuel* 82 (2003), 1219–1229.
8. A. Busch, Y. Gensterblum, B. M. Krooss, and R. Littke, Methane and carbon dioxide adsorption-diffusion experiments on coal: upscaling and modeling, *Int J Coal Geol* 60 (2004), 151–168.
9. J.-Q. Shi, S. Mazumder, K.-H. Wolf, and S. Durucan, Competitive methane desorption by supercritical CO<sub>2</sub>; injection in coal, *Trans Porous Media* 75 (2008), 35–54.
10. R. Haggerty and S. M. Gorelick, Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour Res* 31 (1995), 2383–2400.
11. M. N. Gooseff, S. M. Wondzell, R. Haggerty, and J. Anderson, Comparing transient storage modeling and residence time distribution (rtd) analysis in geomorphically varied reaches in the Lookout Creek basin, Oregon, USA, *Adv Water Resour* 26 (2003), 925–937.
12. R. Haggerty, S. M. Wondzell, and M. A. Johnson, Power-law residence time distribution in the hyporheic zone of a 2nd-order mountain stream, *Geophys Res Lett* 29 (2002), 18–1–18–4.
13. B. Cockburn, G. Gripenberg, and S.-O. Londen, On convergence to entropy solutions of a single conservation law, *J Diff Equ* 128 (1996), 206–251.
14. R. J. LeVeque, *Numerical methods for conservation laws*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1990.
15. R. J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.
16. J. A. Trangenstein, *Numerical solution of hyperbolic partial differential equations*, Cambridge University Press, Cambridge, 2009.
17. D. Kröner, *Numerical schemes for conservation laws*, Wiley-Teubner Series Advances in Numerical Mathematics, John Wiley & Sons Ltd., Chichester, 1997.

18. G. Gripenberg, Nonsmoothing in a single conservation law with memory, *Electron J Diff Equ* 9 (2001), 1–8.
19. J. A. Nohel, A nonlinear conservation law with memory, In *Volterra and functional-differential equations* (Blacksburg, Va., 1981), Vol. 81 of *Lecture Notes in Pure and Appl. Math.*, Dekker, New York, 1982, pp. 91–123.
20. J. Prüss, Positivity and regularity of hyperbolic Volterra equations in Banach spaces, *Math Ann* 279 (1987), 317–344.
21. R. Malek-Madani and J. A. Nohel, Formation of singularities for a conservation law with memory, *SIAM J Math Anal* 16 (1985), 530–540.
22. G. Gripenberg, P. Clément, and S.-O. Londen, Smoothness in fractional evolution equations and conservation laws, *Ann Scuola Norm Sup Pisa Cl Sci* (4) 29 (2000), 231–251.
23. M. Peszyńska, Analysis of an integro-differential equation arising from modelling of flows with fading memory through fissured media, *J Partial Diff Equ* 8 (1995), 159–173.
24. M. Peszyńska, Finite element approximation of diffusion equations with convolution terms, *Math Comp* 65 (1996), 1019–1037.
25. W. McLean and V. Thomée, Asymptotic behaviour of numerical solutions of an evolution equation with memory, *Asymptot Anal* 14 (1997), 257–276.
26. W. McLean and V. Thomée, Numerical solution of an evolution equation with a positive-type memory term, *J Austral Math Soc Ser B* 35 (1993), 23–70.
27. V. Thomée and L. B. Wahlbin, Long-time numerical solution of a parabolic equation with memory, *Math Comp* 62 (1994), 477–496.
28. R. Haggerty, S. A. McKenna, and L. C. Meigs, On the late-time behavior of tracer test breakthrough curves, *Water Resour Res* 36 (2000), 3467–3479.
29. U. Hornung and R. E. Showalter, PDE-models with hysteresis on the boundary, In *Models of hysteresis* (Trento, 1991), Vol. 286 of *Pitman Res. Notes Math. Ser.*, Longman Sci. Tech., Harlow, 1993, pp. 30–38.
30. J. E. Warren and P. J. Root, The behavior of naturally fractured reservoirs, *Soc Petro Eng Jour* 3 (1963), 245–255.
31. T. Arbogast, J. Douglas Jr., and U. Hornung, Derivation of the double porosity model of single phase flow via homogenization theory, *SIAM J Math Anal* 21 (1990), 823–836.
32. M. Peszynska, Methane in subsurface: Mathematical modeling and computational challenges, C. Dawson and M Gerritsen, editors, In *IMA Vol. 156: Computational Challenges in the Geosciences*, Springer, 2013.
33. J. C. Strikwerda, *Finite difference schemes and partial differential equations* 2nd Ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2004.
34. R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007. Steady-state and time-dependent problems.
35. M. J. Ablowitz and A. S. Fokas, *Complex variables: introduction and applications*, Cambridge Texts in Applied Mathematics 2nd Ed., Cambridge University Press, Cambridge, 2003.
36. J. Norbury and A. M. Stuart, Volterra integral equations and a new Gronwall inequality. I. The linear case, *Proc Roy Soc Edinburgh Sect A* 106 (1987), 361–373.
37. L. Tao and H. Yong, A generalization of discrete Gronwall inequality and its application to weakly singular Volterra integral equation of the second kind, *J Math Anal Appl* 282 (2003), 56–62.
38. J. Dixon and S. McKee, Weakly singular discrete Gronwall inequalities, *Z Angew Math Mech* 66 (1986), 535–544.