CORRECTED OPERATOR SPLITTING FOR NONLINEAR PARABOLIC EQUATIONS

Kenneth Hvistendahl Karlsen, Nils Henrik Risebro

Abstract. We present a corrected operator splitting (COS) technique for nonlinear parabolic equations of convection-diffusion type. The main feature of this method is the ability to correctly resolve steep fronts for large time steps, as opposed to standard operator splitting (OS) which fails do so. COS is based on solving a conservation law for modeling convection, a heat equation for modeling diffusion, and finally a certain "residual" conservation law for necessary correction. The residual equation, which is ignored in OS, has an anti-diffusive effect whose purpose is to counter-balance some of the diffusion introduced by the heat equation. It is shown that COS generates a compact sequence of approximate solutions which converges to the solution of the problem. The method of Dafermos constitutes an important part of our solution strategy. Finally, some numerical examples are presented.

0. Introduction. In this paper we introduce a novel operator splitting method for constructing approximate solutions to nonlinear parabolic convection-diffusion problems of the form

$$u_t + f(u)_x = \varepsilon\nu(u)_{xx}, \qquad u(x,0) = u_0(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, T], \tag{1}$$

where $u_0(x)$, $\nu(u)$, and $f(u)$ are given, sufficiently smooth functions, and $\varepsilon > 0$ is a small scaling parameter. Partial differential equations from mathematical physics sometimes appear in the non-conservative form

$$u_t + f(u)_x = \varepsilon \left(d(u)u_x\right)_x, \tag{2}$$

where we can assume that $d(u)$ is a strictly positive function, so that (2) is parabolic and admits classical solutions. The mixed hyperbolic/parabolic case ($d(u) \geq 0$) will be considered elsewhere. In the parabolic context we can obviously write (2) in conservative form (1), so that any solution strategy presented for (1) applies equally well to (2). Consequently, we choose to work with (1) in this paper. Existence and uniqueness of a classical solution to (1) is well known, see for example [18,21]. Furthermore, the notion of a classical solution coincides with the notion of a weak solution for parabolic equations such as (1), see [21].

Equations such as (1) arise in a variety of applications, ranging from models of turbulence [4], via traffic flow [20] and financial modeling [3], to two phase flow in porous media [22]. Equation (1) can also be viewed as a model problem for a system of convection-diffusion equations, such as three phase flow in porous media [26], or the Navier-Stokes equations. Of particular importance is the case where convection dominates diffusion, i.e., ε is small compared with other scales in (1). This is often the case in models of two phase flow in oil reservoirs. Accurate numerical simulation of such models are consequently often complicated by both unphysical oscillations and numerical diffusion. The operator splitting approach presented here is particularly well suited to the case where $\varepsilon \ll 1$.

If $\varepsilon \ll 1$, then (1) is "almost hyperbolic", and it is natural to exploit this when constructing numerical methods. A widely used strategy is viscous operator splitting (OS henceforth), that is, splitting (1) into a hyperbolic conservation law and a parabolic heat equation, each of which is solved by some proper numerical scheme. This approach, or at least certain variations on this approach, has indeed been taken by several authors, we only mention Beale and Majda [2], Douglas and Russell [9], [24], [12], Espedal and Ewing [10], [11], Dawson [8], and more recently Karlsen and Risebro [14]. In [24], a characteristic element method is used to solve the hyperbolic part of (1). In [9], error estimates are obtained for a linear version of (1), and in [14] it is shown that that the viscous splitting method converges to the solution of (1) in the case of linear diffusion, sufficiently smooth flux functions, and any (discontinuous) initial function of bounded variation.

However, numerical experiments [14] suggest that OS can be severely diffusive near steep shock fronts, at least when the time step is large. Let us elaborate on this feature by studying an application of OS to Burgers' equation [4], i.e., $f(u) = \frac{1}{2}u^2$ and $\nu(u) = u$, with Riemann initial data $u_0(x) = \chi_{[0,\infty\rangle}(x)$. The true solution is a single (self-sharpening) shock front moving with positive velocity. In particular, the size of the shock layer is $\mathcal{O}(\varepsilon)$ (see e.g. [25]), which contrasts the the well-known $\mathcal{O}(\sqrt{\varepsilon})$ - layers seen in linear equations.

Key words and phrases. nonlinear parabolic equations, operator splitting, front tracking, initial value problem.
The first author has been supported by VISTA, a research cooperation between the Norwegian Academy of Science and Letters and Den norske stats oljeselskap a.s. (Statoil).

Let $\mathcal{S}^f(t)$ denote the entropy satisfying solution operator associated with the nonlinear conservation law

$$v_t + f(v)_x = 0, \tag{3}$$

and let $\mathcal{H}(t)$ denote the solution operator associated with the (linear) heat equation

$$w_t = \varepsilon w_{xx}. \tag{4}$$

Then the operator splitting (OS) approximation takes the form

$$u(x, n\Delta t) \approx \left[\mathcal{H}(\Delta t) \circ \mathcal{S}^f(\Delta t)\right]^n u_0(x). \tag{5}$$

Let us calculate the first step in (5) for Burgers' equation. The entropy weak solution to the convex conservation law (3) is $v(x, \Delta t) = \chi_{[\Delta t/2, \infty)}(x)$. Using $v(x, \Delta t)$ as (discontinuous) initial data for the heat equation (4), we obtain the following explicit formula for the OS approximation

$$u(x, \Delta t) \approx \left[\mathcal{H}(\Delta t)\mathcal{S}^f(\Delta t)\right] u_0(x) \equiv \int_{\Delta t/2}^{\infty} \frac{1}{\sqrt{4\pi\varepsilon\Delta t}} \exp\left[\frac{-(x-y)^2}{4\varepsilon\Delta t}\right] dy. \tag{6}$$

It is not difficult to deduce from this expression that the shock layer has size $\mathcal{O}(\sqrt{\varepsilon\Delta t})$. Consequently, we do not expect that the layer is properly resolved unless a small time step ($\Delta t = \mathcal{O}(\varepsilon)$) is used, a claim that is in fact supported by numerical evidence [14,15].

An interesting observation is the following. Let $f_c(u)$ denote the upper concave envelope of $f(u) = \frac{1}{2}u^2$ in the interval $[0, 1]$. Applying OS to the linear equation $u_t + f_c(u)_x = \varepsilon u_{xx}$ still yields the solution (6). In fact, applying OS to $u_t + g(u)_x = \varepsilon u_{xx}$, for any convex flux function $g(u)$ that lies below or equals $f_c(u)$, will give the approximation (6). The OS solution of Burgers' equation does not take into account the convex shape of the flux function, that is, the self-sharpening nature of the front. However, the part of the flux function that is neglected can be identified as a residual flux term of the form $f_{\mathrm{res}} \equiv f - f_c$.

Now the idea is to take a third correction step to reduce the superfluous diffusion introduced by the heat equation (4), i.e., instead of (5) we use an approximation formula of the form

$$u(x, n\Delta t) \approx \left[\mathcal{C}(\tau) \circ \mathcal{H}(\Delta t) \circ \mathcal{S}^f(\Delta t)\right]^n u_0(x), \qquad \tau > 0, \tag{7}$$

where $\mathcal{C}(\tau)$ is the solution operator associated with the "residual" conservation law $v_t + f_{\mathrm{res}}(v)_x = 0$ at time τ. Due to the special form of f_{res}, convex with $f_{\mathrm{res}}(0) = f_{\mathrm{res}}(1) = 0$, we see that $\mathcal{C}(\tau)$ possesses the desired anti-diffusive (sharpening) property when applied to (6). Of course, we should not take τ too large, typically not larger than Δt, because the diffusive front (6) then will be sharpened into a discontinuity. When choosing τ we should have in mind the that the OS layer and the true layer have sizes $\mathcal{O}(\sqrt{\varepsilon\Delta t})$ and $\mathcal{O}(\varepsilon)$, respectively. In addition, we should take into account the fact that "particles" upon action of $\mathcal{C}(\tau)$ move a distance not exceeding $\tau\|(f_{\mathrm{res}})'\|_\infty$ (finite speed of propagation).

As we have seen, for a single Riemann problem it is possible to derive *à priori* the explicit expression for the residual flux term f_{res}. This was first observed by Espedal and Ewing [EspEw] (see also Dahle [7]) who suggested a splitting method based on the linear conservation law $v_t + f_c(v)_x = 0$ and the nonlinear diffusion equation $w_t + f_{\mathrm{res}}(w)_x = \varepsilon w_{xx}$, instead of (3) and (4). This two-step method, which can be viewed as an alternative to our three-step method (7), has the advantage of giving the correct size of the shock layer and making it possible extend the characteristic methods [9,24] to nonlinear problems without severe time step restrictions.

Of course, an *à priori* construction of the residual flux f_{res} is not possible for general problems. The main purpose of the present paper is to demonstrate that it is possible to dynamically construct a residual flux term $f_{\mathrm{res}}(x, \cdot)$ for general problems when using front tracking, as defined by Dafermos [6] (see also [13]), to solve the nonlinear conservation law (3). Consequently, the three-step corrected operator splitting approach (7) makes sense in general. Our construction relies heavily upon the fact that Dafermos' method is based on solving Riemann problems. We also prove that (7) converges to the solution of (1) as various discretization parameters tend to zero.

The rest of this paper is organized as follows: In section 1 we present some useful information about parabolic equations and the front tracking method. In section 2 we explain in detail the semi-discrete corrected operator splitting scheme (7) and the construction of the residual flux term. In section 3 we obtain compactness of the sequence of approximate solutions generated by the corrected splitting scheme. In section 4 we present an application of a fully discrete scheme in which (4) is solved by the Galerkin finite element method.

1. Preliminaries. We shall always assume that $f(u)$ and $\nu(u)$ are continuously differentiable ($f, \nu \in C^1$), and that $u_0(x)$ is a function of bounded variation ($u_0 \in BV$). Under these assumptions it is well known that there exists a unique classical solution to (1), with the initial data assumed in the weak sense, i.e.,

$$\int \phi(x)\,(u(x,t) - u_0(x)) \to 0, \qquad \text{as } t \to 0+, \qquad \text{for all } \phi \in C_c.$$

At all points of continuity, the data is assumed in the usual pointwise sense. Moreover, the solution $u(x,t)$ is bounded and all the derivatives appearing in the equation are continuous. We call $u(x,t)$ a weak solution if

$$\iint_0^T (u\phi_t + f(u)\phi_x + \varepsilon\nu(u)\phi_{xx})\, dt\, dx + \int u_0(x)\phi(x,0)\, dx = 0, \qquad \text{for all } \phi \in C_0^\infty(\mathbb{R} \times \langle 0, T]) . \tag{8}$$

For parabolic equations of the form (1), it is known that the notion of weak and classical solutions coincide. Consequently, in order to show that the limit of a converging sequence is the unique classical solution to (1), it is sufficient to demonstrate that the limit satisfies (8). We refer to [18,21] for a survey of the mathematical theory of nonlinear parabolic equations such as (1) and (2).

For later use, let us collect some of the properties that the solution of (1) possesses.

Lemma 1.1 (Parabolic equations). *For $u_0 \in BV$ and $f, \nu \in C^1$, let $u(x,t)$ be the unique classical solution to*

$$u_t + f(u)_x = \varepsilon\nu(u)_{xx}, \qquad u(x,0) = u_0(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, T] .$$

Then the following two à priori bounds hold

$$\|u(\cdot,t)\|_\infty \le \|u_0\|_\infty, \qquad TV\,(u(\cdot,t)) \le TV\,(u_0)\,, \qquad t \in \langle 0, T] . \tag{9}$$

For $g \in C^1$ and $v_0 \in BV$, let $v(x,t)$ be the unique classical solution to

$$v_t + g(v)_x = \varepsilon\nu(u)_{xx}, \qquad v(x,0) = v_0(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, T] .$$

Then the following stability (comparison) result holds

$$\|u(\cdot,t) - v(\cdot,t)\|_1 \le \|u_0 - v_0\|_1 + t\,\|f - g\|_{Lip} \min\left(TV\,(u_0)\,, TV\,(v_0)\right), \qquad t \in \langle 0, T] . \tag{10}$$

Proof. The first property is well known and follows from the maximum principle. Let us therefore concentrate on proving the two remaining claims. The difference $e(x,t) = u(x,t) - v(x,t)$ satisfies the error equation

$$e_t + (a(x,t)e(x,t))_x - \varepsilon\,(b(x,t)e)_{xx} = \mathcal{T}(x,t), \tag{11}$$

where the coefficients $a(x,t)$ and $b(x,t)$ are given by

$$a(x,t) = \int_0^1 f'\,(\xi v(x,t) + (1-\xi)u(x,t))\; d\xi,$$
$$b(x,t) = \int_0^1 \nu'\,(\xi v(x,t) + (1-\xi)u(x,t))\; d\xi.$$

Here we assume that both e and e_x tend to zero as $|x| \to \infty$, and let $\mathcal{T}(x,t)$ denote the truncation error, i.e.,

$$\mathcal{T}(x,t) = v_t + f\,(v)_x - \varepsilon\nu(v)_{xx} .$$

Let $\psi(x,t)$ be the solution of the backward problem,

$$\psi_t + a(x,t)\psi_x + \varepsilon b(x,t)\psi_{xx} = 0, \qquad \psi(x,T) = \phi(x), \qquad (x,t) \in \mathbb{R} \times [0, T\rangle\,, \tag{12}$$

where $\phi(x)$ is smooth and $|\phi|$ tends to zero as $|x| \to \infty$. It is well known that the (classical) solution satisfies the maximum principle

$$\|\psi(\cdot,t)\|_\infty \le \|\phi\|_\infty, \qquad t < T. \tag{13}$$

By integrating the error equation (11) against ψ over $\mathbb{R}\times\langle 0,T\rangle$, and noting that (12) is just the adjoint problem of the error equation, we obtain

$$\int e(x,T)\phi(x)\,dx = \int e(x,0)\psi(x,0)\,dx + \iint_0^T \mathcal{T}(x,t)\psi(x,t)\,dt\,dx. \tag{14}$$

Let us first show the stability with respect to the initial function, that is, let $g\equiv f$. In this case $\mathcal{T}=0$, and by choosing $\phi=\operatorname{sign}(e)$, we obtain

$$\|u(\cdot,T)-v(\cdot,T)\|_1 \le \|u_0-v_0\|_1. \tag{15}$$

An important consequence of this stability result is that solutions of (1) have total variation that is bounded by the initial variation. Recall that for a function $h=h(x)$, the total variation can be defined as

$$TV(h) = \limsup_{\delta\to 0}\frac{1}{\delta}\int |h(x)-h(x-\delta)|\,dx.$$

Thanks to translation invariance and estimate (15), we readily calculate that

$$\begin{aligned}
&TV(u(\cdot,T))\\
&\quad= \limsup_{\delta\to 0}\frac{1}{\delta}\int |u(x,T)-u(x-\delta,T)|\;dx\\
&\quad\le \limsup_{\delta\to 0}\frac{1}{\delta}\int |u_0(x)-u_0(x-\delta)|\;dx\\
&\quad= TV(u_0).
\end{aligned}$$

We now use this estimate and (13) to establish stability with respect to the flux function,

$$\begin{aligned}
&\left|\iint_0^T \mathcal{T}(x,t)\psi(x,t)\,dt dx\right|\\
&\quad= \left|\iint_0^T (f(v)-g(v))_x\,\psi(x,t)\,dt\,dx\right|\\
&\quad\le \iint_0^T \|f'-g'\|_\infty |v_x(x,t)||\psi(x,t)|\,dt\,dx \le T\,\|f'-g'\|_\infty\, TV(v_0)\,\|\phi\|_\infty.
\end{aligned}$$

Choosing $\phi=\operatorname{sign}(e)$ in (14) and using symmetry, we derive the desired result (10). □

Since the front tracking method will be important, we give a brief description. Let $f(u)$ be a Lipschitz continuous, piecewise linear function with breakpoints located at $\{u_i\}$. Consider the nonlinear conservation law

$$u_t + f(u)_x = 0, \qquad u(x,0)=u_0(x), \qquad (x,t)\in\mathbb{R}\times\langle 0,T], \tag{16}$$

where we assume that u_0 is piecewise constant, taking a finite number of values, say, $\{u_{0,i}\}\subset\{u_i\}$. Consider first the Riemann problem with $u_L=u_k$ and $u_R=u_j$ for some k and j. Let $f_c(u;u_L,u_R)$ be given by

$$f_c(u\,;u_L,u_R) = \begin{cases} \text{the lower convex envelope of } f \text{ between } u_L \text{ and } u_R, \text{ if } u_L<u_R,\\ \text{the upper concave envelope of } f \text{ between } u_R \text{ and } u_L, \text{ if } u_L>u_R. \end{cases}$$

Since f is piecewise linear, then so is f_c. Let $\{\bar u_i\}$, $i=1,\dots,M$, be such that

$$\bar u_0 = u_L, \qquad \bar u_M = u_R, \qquad \{\bar u_0,\dots,\bar u_M\}\subseteq\{u_k,\dots,u_l\},$$

and such that f_c is linear on each interval $[\bar u_i,\bar u_{i+1}]$, $i=0,\dots,M-1$. The solution of the Riemann problem with left state u_L and right state u_R is then given by

$$u(x,t) = \begin{cases} u_L, & \text{for } x<\bar s_0 t,\\ \bar u_i, & \bar s_i t\le x\le \bar s_{i+1}t, \qquad i=0,\dots,M-2,\\ u_R, & x>\bar s_{M-1}t, \end{cases} \tag{17}$$

where

$$\bar{s}_i = \frac{\bar{f}_{i+1} - \bar{f}_i}{\bar{u}_{i+1} - \bar{u}_i}, \qquad i = 0, \ldots, M-1,$$

and $\bar{f}_i = f_c(\bar{u}_i; u_L, u_R)$. The solution of the more general problem (16) is constructed as follows. Observe that each jump in the initial data u_0 defines a Riemann problem. The solution of these problems leads to a series of discontinuities propagating in the (x,t) plane. By "gluing" together the solutions of the Riemann problems we have the global solution until, at some point, two or more of these discontinuities interact, and we have what is called a shock collision. When two or more neighboring discontinuities collide, they define a new Riemann problem with left and right states given by the values immediately to the left and to the right of the collision. This Riemann problem is then solved, and we have the solution until the next shock collision occurs. This collision is of course handled in the same way, similarly for subsequent collisions. The front tracking method for constructing the exact solution to (16) may briefly be summarized as follows:

(1) Solve the Riemann problems defined by the piecewise constant initial data.
(2) Keep track of shock collisions and solve Riemann problems arising at the collision points.

For various implementation aspects of the front tracking method we refer to [23] and [19].

Remark. The front tracking method for general conservation laws (arbitrary f and u_0) consists in replacing f with a piecewise linear approximation and u_0 with a piecewise constant approximation, and then to solve the resulting perturbed problem exactly according to the procedure described above. For a more detailed treatment of the front tracking method we refer to [13].

Next, let us introduce a dynamic grid $\{z_j^n\}$ on which the approximate solutions will be defined. Let the grid cells be of the form $z_j^n = [x_j^n, x_{j+1}^n\rangle$, and introduce the projection operator $\pi = \pi(\{z_j^n\})$ as

$$\pi g(x) = \frac{1}{|z_j^n|}\int_{z_j^n} g(\tilde{x})\, d\tilde{x}, \qquad \text{for } x \in z_j^n. \tag{18}$$

Here π is to be considered as an operator from the space of functions of bounded variation to functions that are constant on each grid cell z_j^n. The grid cells z_j^n can be of varying size, but the grid is assumed to be regular in the sense that

$$\Delta x_{\min} \le |z_j^n| \le \Delta x_{\max} \equiv \Delta x, \qquad \frac{\Delta x_{\max}}{\Delta x_{\min}} \le \text{Const.} \tag{19}$$

This means that we can adjust the grid (see section 4) to follow the dynamics of the solution in order to enable optimal resolution. However, the adjustment must be done so that the mesh regularity condition (19) is not violated. The projection operator will be used to generate piecewise constant data for the front tracking method. We shall later need the following three properties

$$\|\pi g\|_\infty \le \|g\|_\infty, \qquad TV(\pi g) \le TV(g), \qquad \|\pi g - g\|_1 \le TV(g)\,\Delta x, \tag{20}$$

which are easily seen to hold for any function $g(x)$ of bounded variation.

Finally, let $f_\delta(u)$ denote a piecewise linear and continuous approximation to $f(u)$, chosen so that

$$\|f_\delta\|_{Lip} \le \|f\|_{Lip}, \qquad \|f - f_\delta\|_\infty = \mathcal{O}(\delta^2), \qquad \|f - f_\delta\|_{Lip} = \mathcal{O}(\delta), \qquad \text{as } \delta \to 0. \tag{21}$$

2. The semi-discrete method. In this section we describe the semi-discrete corrected operator splitting method (COS henceforth). The COS strategy is first presented within an abstract framework, and then we give an explicit realization of the strategy using the residual flux function briefly introduced in the first section. Fix $T > 0$ and an integer $N \ge 1$, and choose Δt such that $N\Delta t = T$. We demand that the timestep Δt and the space discretization Δx are related as follows

$$\frac{\Delta x^2}{\Delta t} \to 0, \qquad \frac{\Delta x}{\Delta t} \le \text{Const.}, \qquad \text{as } \Delta x, \Delta t \to 0. \tag{22}$$

Note that in contrast to finite difference methods, (22) allows for large time steps. Let now u^n denote the piecewise constant approximate solution to (1) at time $t = n\Delta t$, for some fixed $n = 0, \ldots, N-1$. For notational convenience we have suppressed the dependency on Δx, Δt, and δ in u^n. Next, we describe how to inductively construct the piecewise constant function u^{n+1} from u^n.

Step 1 (convection step): Let $v(x,t)$ be the entropy weak solution (in the sense of Kruzkov [17]) of the hyperbolic conservation law

$$v_t + f_\delta(v)_x = 0, \qquad v(x,0) = u^n(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, \Delta t] . \tag{23}$$

Recall that $v(x,t)$ coincides with the solution generated by the front tracking method. Let $\mathcal{S}^{f_\delta}(t)$ denote the solution operator associated with the conservation law $v_t + f_\delta(v)_x = 0$ at time t. We then define an intermediate solution

$$u^{n+1/3} = \mathcal{S}^{f_\delta}(\Delta t) u^n.$$

Step 2 (diffusion step): Let $w(x,t)$ be the solution of the nonlinear heat equation

$$w_t = \varepsilon \nu(w)_{xx}, \qquad w(x,0) = u^{n+1/3}(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, \Delta t] . \tag{24}$$

Let $\mathcal{H}^\nu(t)$ denote the solution operator associated with $w_t = \varepsilon\nu(w)_{xx}$ at time t, and $\mathcal{H}^\nu_{\Delta x}(t) = \pi \circ \mathcal{H}^\nu(t)$ its projection onto the grid $\{z_j^n\}$. We define the second intermediate solution by

$$u^{n+2/3} = \mathcal{H}^\nu_{\Delta x}(\Delta t) u^{n+1/3}.$$

Step 3 (correction step): Let τ be a positive parameter (referred to as the correction time) chosen so that $\tau = c\Delta t$ for some constant $c > 0$. For each fixed $\tau > 0$, let $\mathcal{C}(\tau) : BV \to BV$ be a given operator referred to as the correction operator. We assume that $\mathcal{C}(\tau)$ satisfies the following four regularity estimates:

$$\left\| \mathcal{C}(\tau) u^{2/3} \right\|_\infty \le M_1, \tag{C1}$$

$$TV\left(\mathcal{C}(\tau) u^{2/3} \right) \le M_2, \tag{C2}$$

$$\left\| \mathcal{C}(\tau) u^{2/3} - u^{2/3} \right\|_1 \le M_3 \Delta t, \tag{C3}$$

$$\left| \int \left(\mathcal{C}(\tau) u^{2/3}(x) - u^{2/3}(x) \right) \phi(x) \right| \le M_4 \left(\Delta t\right)^{1+\gamma}, \qquad \text{for all } \phi \in C_0^\infty, \tag{C4}$$

where $\gamma > 0$ is a strictly positive number, and $M_1, \dots, M_4$ are constants independent of Δx, Δt, δ, and τ. The four conditions (C1)–(C4) are imposed in order to ensure convergence to the solution of (1). The idea is that $\mathcal{C}(\tau)$ should have an anti-diffusive effect (the amount depending on the size of τ) to counter-balance some of the diffusion introduced in step 2, so that correct balance between nonlinear convection and diffusion is achieved even for large Δt. With the current notation in hand, we finally define the corrected operator splitting solution by

$$u^{n+1} = \mathcal{C}(\tau) u^{n+2/3}.$$

The COS solution $\{u^n\}_{n=0}^N$ is constructed inductively by applying the above three-step procedure to construct u^{n+1} from u^n,

$$u^{n+1} = \left[\mathcal{C}(\tau) \circ \mathcal{H}^\nu_{\Delta x}(\Delta t) \circ \mathcal{S}^{f_\delta}(\Delta t) \right] u^n, \qquad \tau > 0, \tag{25}$$

where the induction is initiated by setting $u^0 = \pi u_0$. The rest of this section is devoted to suggesting an explicit construction of the operator $\mathcal{C}(\tau)$. The essential part of this construction is the residual flux term f_{res}. The motivation for considering the residual flux term is found in the introduction.

Observe that the front tracking solution is a step function whose discontinuities always are entropy satisfying shocks. Consequently, it is possible to construct a residual flux term with respect to each shock in this solution. Then, to obtain a global correction operator, we should connect these local terms properly. Below we propose a "connection strategy" that is computer efficient and easy to implement. But it is not difficult construct other methods for connecting the local terms that will yield (slightly) different COS schemes than the one presented here. However, when deriving realizations of the correction operator we must have in mind the conditions (C1)–(C4), which imply that the resulting COS schemes converge to the exact solution of (1).

Suppose that the function $u^{n+1/3}(x)$ is piecewise constant with its discontinuities located at the points $\{y_i\}$. Let $\tilde{u}_{i+1}$ denote the value of $u^{n+1/3}$ in the interval $[y_i, y_{i+1}\rangle$, and let $f^i_{\text{res}}(u)$ denote the ith local residual flux term constructed with respect to the left and right shock values $\tilde{u}_i$ and $\tilde{u}_{i+1}$ located at $x = y_i$. More precisely, we define $f^i_{\text{res}}(u)$ in terms of the formula

$$f^i_{\text{res}}(u) = \begin{cases} f_\delta(u) - f_{\delta,c}\left(u; \tilde{u}_i, \tilde{u}_{i+1}\right), & \text{when } u \in [\tilde{u}_i, \tilde{u}_{i+1}] \text{ for some } i, \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

After ignoring the terms that are zero, we obtain a (finite) sequence of non-zero local residual terms $f^i_{\rm res}(u)$, $i = 1,\dots,N_n$, ordered with respect to increasing (location) y_i - values. Let $\bar{x}_i$, $i = 0,\dots,N_{n+1}$, be spatial positions (degrees of freedom) chosen such that y_i is located somewhere in the interval $\langle \bar{x}_i, \bar{x}_{i+1}\rangle$. Then we define the global residual flux term (suppressing the δ and n - dependency) by

$$f_{\rm res}(x,u) = \begin{cases} f^i_{\rm res}(u), & \text{when } x \in [\bar{x}_i, \bar{x}_{i+1}\rangle \text{ for some } i = 0,\dots,N_n, \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

The function $f_{\rm res}(x,u)$ may be discontinuous as a function of x for fixed u, but observe that $f_{\rm res}\left(x, u^{n+1/3}(x)\right) = 0$ for all x. Furthermore, for fixed x we see that $f_{\rm res}(x,u)$ is a piecewise linear function of u.

Let $v_i(x,t)$ denote the entropy weak solution (in the sense of Bardos *et al.* [1]) to the nonlinear conservation law

$$v_t + f^i_{\rm res}(v)_x = 0, \qquad (x,t) \in \langle \bar{x}_i, \bar{x}_{i+1}\rangle \times \langle 0,\tau], \tag{28}$$

with initial data and boundary data (whenever necessary) imposed as follows

$$\begin{aligned} v(x,0) &= u^{n+2/3}(x), && x \in \langle \bar{x}_i, \bar{x}_{i+1}\rangle, \\ v(\bar{x}_i,t) &= u^{n+2/3}(\bar{x}_i+), && t \in \langle 0,\tau], \\ v(\bar{x}_{i+1},t) &= u^{n+2/3}(\bar{x}_{i+1}-), && t \in \langle 0,\tau]. \end{aligned} \tag{29}$$

Observe that $v_i(x,t)$ coincides with the solution generated by the front tracking method. Introduce the globally defined function

$$v(x,t) = \sum_{i=0}^{N_n} v_i(x,t)\chi_{[\bar{x}_i,\bar{x}_{i+1}\rangle}(x), \qquad (x,t) \in \mathbb{R} \times \langle 0,\tau], \tag{30}$$

and let $\mathcal{C}(\tau)$ denote the correction operator that takes $u^{n+2/3}$ to the function $v(x,\tau)$, i.e.,

$$\mathcal{C}(\tau)u^{n+2/3}(x) = v(x,\tau), \qquad \tau > 0.$$

Consequently, our correction operator is realized by solving initial–boundary value problems for nonlinear conservation laws with carefully chosen flux functions.

Remark. According to (27), a local residual term $f^i_{\rm res}$ is constructed with respect to each discontinuity in the piecewise constant front tracking solution. In particular, a large number of these local flux terms will be so small that they have no significant influence on the solution. Consequently, when doing numerical calculations a local term $f^i_{\rm res}$ is not taken into account if $|\tilde{u}_{i+1} - \tilde{u}_i| \le c_{\rm tr}$, where $c_{\rm tr}$ is some small (problem dependent) threshold parameter. This means that computational effort in terms of correction is only spent in the regions where significant shock fronts are located.

From a computational point of view, we ought to be more specific about the choice of the interval $\langle \bar{x}_i, \bar{x}_{i+1}\rangle$, which is used explicitly in (28) and (29). Let therefore $\hat{x}_i$, $i = 1,\dots,N_n$, be the midpoint of each interval where $u^{n+1/3}(x)$ is constant, i.e., $\hat{x}_i = \frac{1}{2}(y_{i-1} + y_i)$, and define $\hat{x}_0 = y_0 - 1$, $\hat{x}_{N_n+1} = y_{N_n} + 1$. We now let the ith interval $\langle \bar{x}_i, \bar{x}_{i+1}\rangle$ be given by

$$\langle \bar{x}_i, \bar{x}_{i+1}\rangle \equiv \langle \hat{x}_i - \Delta, \hat{x}_{i+1} + \Delta\rangle,$$

where $\Delta \ge 0$ is a small parameter that can depend on i. When choosing Δ, one should have in mind that the correction effect on $\langle \bar{x}_i, \bar{x}_{i+1}\rangle$ is "maximized" when

$$f^i_{\rm res}\left(u^{n+2/3}(\bar{x}_i+)\right) \approx 0, \qquad f^i_{\rm res}\left(u^{n+2/3}(\bar{x}_{i+1}-)\right) \approx 0.$$

In the computational study presented in section 4, we have typically taken the parameter Δ to be of the same size as the polygonal approximation parameter δ. This strategy seems to work well for the numerical examples presented here, but still, as mentioned above, other strategies for choosing $\{\bar{x}_i\}$ should be explored in the future. Furthermore, one should also investigate alternatives to (30) for connecting the local residual flux terms.

We now show that the conditions (C1)–(C4) ensuring convergence are satisfied. By construction we know that that the operator $\mathcal{C}(t)$ does not introduce new minima or maxima and thus (C1) holds. Let $v^E_i(x,t)$ denote the entropy weak solution to the conservation law

$$v_t + f^i_{\rm res}(v)_x = 0, \qquad (x,t) \in \mathbb{R} \times \langle 0,\tau],$$

with extended initial data given by

$$v(x,0)=\begin{cases} u^{n+2/3}(x_i+), & \text{for } x\le \bar{x}_i,\\ u^{n+2/3}(x), & \text{for } x\in\langle \bar{x}_i,\bar{x}_{i+1}\rangle,\\ u^{n+2/3}(x_{i+1}-), & \text{for } x\ge \bar{x}_{i+1}.\end{cases}$$

Then observe that $v_i^E\big|_{\langle \bar{x}_i,\bar{x}_{i+1}\rangle}=v_i$ and that $TV\left(v_i^E(\cdot,t)\right)\le TV\left(v_0^E\right)$, see Lemma 3.2. This gives $TV\left(v_i(\cdot,t)\right)\le TV\left(v_i(\cdot,0)\right)$, from which it obviously follows that

$$TV\left(\mathcal{C}(\tau)u^{n+2/3}\right)\le TV\left(u^{n+2/3}\right).$$

Hence, (C2) holds. Since the shocks in $v_i(x,t)$ propagate at finite speed and since the variation is finite (C2), we obtain

$$\begin{aligned}
&\left|\int\left(\mathcal{C}(\tau)u^{n+2/3}(x)-u^{n+2/3}(x)\right)\,dx\right|\\
&\quad\le\sum_i\int_{\bar{x}_i}^{\bar{x}_{i+1}}\left|v_i(x,\tau)-v_i(x,0)\right|\,dx\\
&\quad\le \text{Const.}\cdot\sum_i TV\left(u^{n+2/3}\big|_{\langle \bar{x}_i,\bar{x}_{i+1}\rangle}\right)\tau\\
&\quad\le \text{Const.}\cdot TV\left(u^{n+2/3}\right)\tau\le M_3\Delta t,
\end{aligned}\tag{31}$$

where we have assumed that $\tau=c\Delta t$ for some non-zero constant c. Thus, (C3) holds. By assuming a slightly stronger relation, namely, $\tau=c\,(\Delta t)^{1+\gamma}$ as $\Delta t\to 0$, where $\gamma>0$ is an arbitrary small number, (C4) follows from (C3). This condition has no practical consequences from a computational point of view, but note that in "worst case" scenarios the operator $\mathcal{H}^{\nu}_{\Delta x}(\Delta t)$ can destroy significant structures of $u^{n+1/3}$ (e.g. monotonicity properties and local extrema), so that $\mathcal{C}(\tau)$ possesses no correction effect when applied to $u^{n+2/3}$. Consequently, the condition $\tau=c\,(\Delta t)^{1+\gamma}$ becomes necessary in order to ensure convergence to the true solution.

However, in more typical applications where $\mathcal{H}^{\nu}_{\Delta x}(\Delta t)$ conserves the structures of $u^{n+1/3}$, this estimate can certainly be improved upon. To this end, we will therefore apply the correction operator only if the following two conditions are fulfilled. First, we assume that

$$\left(f^i_{\text{res}}\right)'\left(u^{n+2/3}(\bar{x}_i)\right)\ge 0,\qquad \left(f^i_{\text{res}}\right)'\left(u^{n+2/3}(\bar{x}_{i+1})\right)\le 0,\qquad i=0,\dots,N_n.\tag{32}$$

Note that (32) is necessary for the conservation law (28) to actually possess a correction effect on $u^{n+2/3}$. Consequently, on intervals $\langle \bar{x}_i,\bar{x}_{i+1}\rangle$ where (32) is violated, the correction operator operator is not applied since it would not have the "right" correction effect there. Secondly, we assume, the relation

$$\left|u^{n+2/3}(\bar{x}_i)-u^{n+1/3}(\bar{x}_i)\right|\le C\sqrt{\Delta t},\qquad i=0,\dots,N_{n+1},\tag{33}$$

where C is some finite constant independent of the discretization parameters and the position $\bar{x}_i$, but dependent on the initial function and the flux function. This condition is, however, merely a technical assumption associated with the subsequent convergence analysis. The assumptions (32) and (33) imply that

$$\left|v\left(\bar{x}_i,t\right)-\tilde{u}_i\right|=\mathcal{O}\left(\sqrt{\Delta t}\right),\qquad i=0,\dots,N_{n+1},\tag{34}$$

where $v(x,t)$ is given by (30) and $t\in\langle 0,\tau]$. Now let $v_i(x,t)$ denote the solution of (28) and (29), and let

$$v_i^h(x,t)=(\omega_h * v_i(\cdot,t))(x,t),\qquad (x,t)\in\langle \bar{x}_i+h,\bar{x}_{i+1}-h\rangle\times\langle 0,\tau],$$

where $\omega_h(x)$ is a standard C_0^∞ - mollifier with smoothing radius h. Then we have that the smooth function $v_i^h(x,t)$ satisfies the equation

$$\left[v_i^h\right]_t+\left[g_i^h\right]_x=0,\qquad g_i^h(x,t)=\left(\omega_h * g_i(v_i(\cdot,t))\right)(x,t),\qquad g_i(v)=f^i_{\text{res}}(v).$$

Integrating $\left[v_i^h\right]_t + \left[g_i^h\right]_x = 0$ against a test function $\phi \in C_0^\infty$ over $\langle \bar{x}_i + h, \bar{x}_{i+1} - h\rangle \times \langle 0, \tau\rangle$, and using the fact that $f_{\mathrm{res}}\left(x, u^{n+1/3}(x)\right) = 0$ for all x, we can calculate as follows

$$
\begin{aligned}
&\left|\int_{\bar{x}_i}^{\bar{x}_{i+1}} \left(v_i(x,\tau) - v_i(x,0)\right)\phi(x)\right| dx \\
&\quad = \lim_{h\to 0}\left|\int_{\bar{x}_i+h}^{\bar{x}_{i+1}-h} \left(v_i^h(x,\tau) - v_i^h(x,0)\right)\phi(x)\right| dx \\
&\quad = \lim_{h\to 0}\left|\int_0^\tau \int_{\bar{x}_i+h}^{\bar{x}_{i+1}-h} \left[g_i^h(x,t)\right]_x \phi(x)\,dx\,dt\right| \\
&\quad \le \lim_{h\to 0}\left|\int_0^\tau \int_{\bar{x}_i+h}^{\bar{x}_{i+1}-h} g_i^h(x,t)\left[\phi(x)\right]_x\,dx\,dt\right| + \lim_{h\to 0}\left|BT_i^h\right| \\
&\quad \le \|\phi_x\|_\infty \int_0^\tau \int_{\bar{x}_i}^{\bar{x}_{i+1}} \left|g_i(v(x,t)) - g_i\left(u^{n+1/3}(x)\right)\right|\,dx\,dt + \lim_{h\to 0}\left|BT_i^h\right| \\
&\quad \le \|\phi_x\|_\infty \int_0^\tau \int_{\bar{x}_i}^{\bar{x}_{i+1}} \left|g_i\left(v(x,t)\right) - g_i\left(u^{n+2/3}(x)\right)\right|\,dx\,dt \\
&\qquad + \|\phi_x\|_\infty \int_0^\tau \int_{\bar{x}_i}^{\bar{x}_{i+1}} \left|g_i\left(u^{n+2/3}(x)\right) - g_i\left(u^{n+1/3}(x)\right)\right|\,dx\,dt + \lim_{h\to 0}\left|BT_i^h\right| \\
&\quad \le \mathrm{Const.}\cdot\|\phi_x\|_\infty \int_0^\tau \int_{\bar{x}_i}^{\bar{x}_{i+1}} \left|v_i(x,t) - u^{n+2/3}(x)\right|\,dx\,dt \\
&\qquad + \mathrm{Const.}\cdot\|\phi_x\|_\infty \int_0^\tau \int_{\bar{x}_i}^{\bar{x}_{i+1}} \left|u^{n+2/3}(x) - u^{n+1/3}(x)\right|\,dx\,dt + \lim_{h\to 0}\left|BT_i^h\right|.
\end{aligned}
$$

Here BT_i^h denotes the boundary terms arising from integration by parts, i.e.,

$$
BT_i^h = \int_0^\tau \left(g_i^h(\bar{x}_{i+1} - h, t)\phi(\bar{x}_{i+1} - h) - g_i^h(\bar{x}_i + h, t)\phi(\bar{x}_i + h)\right)\,dt.
$$

We estimate the boundary terms $BT_i \equiv \lim_{h\to 0} BT_i^h$ as follows

$$
\begin{aligned}
|BT_i| &= \left|\int_0^\tau \left(g_i\left(v_i(\bar{x}_{i+1},t)\right)\phi(\bar{x}_{i+1}) - g_i\left(v_i(\bar{x}_i,t)\right)\phi(\bar{x}_i)\right)\,dt\right| \\
&\le \|\phi\|_\infty \int_0^\tau \left|g_i\left(v_i(\bar{x}_{i+1},t)\right)\right|\,dt + \|\phi\|_\infty \int_0^\tau \left|g_i\left(v_i(\bar{x}_i,t)\right)\right|\,dt \\
&\equiv I_i\left(\bar{x}_{i+1}\right) + I_i\left(\bar{x}_i\right).
\end{aligned}
$$

Having (34) and $g_i\left(\tilde{u}_i\right) = 0$ in mind, we can estimate $I_i\left(\bar{x}_i\right)$ (and similarly for $I_i\left(\bar{x}_{i+1}\right)$) as

$$
\begin{aligned}
I_i\left(\bar{x}_i\right) &= \|\phi\|_\infty \int_0^\tau \left|g_i\left(v_i(\bar{x}_i,t)\right) - g_i\left(\tilde{u}_i\right)\right|\,dt \\
&\le \mathrm{Const.}\cdot\|\phi\|_\infty \int_0^\tau \left|v_i(\bar{x}_i,t) - \tilde{u}_i\right|\,dt \le \mathrm{Const.}\cdot\|\phi\|_\infty\,\tau\sqrt{\Delta t},
\end{aligned}
$$

so that $|BT_i| \le \mathrm{Const.}\cdot\|\phi\|_\infty\,\tau\Delta x$. Consequently, by summing over all $i = 0, \dots, N_n$, we get

$$
\begin{aligned}
&\left|\int \left(\mathcal{C}(\tau)u^{n+2/3}(x) - u^{n+2/3}(x)\right)\phi(x)\,dx\right| \\
&\le \mathrm{Const.}\cdot\|\phi_x\|_\infty \left(\int_0^\tau \int \left|v(x,t) - u^{n+2/3}(x)\right| + \left|u^{n+2/3}(x) - u^{n+1/3}(x)\right|\,dx\,dt\right) + \mathrm{Const.}\cdot\|\phi\|_\infty\,\tau\sqrt{\Delta t} \\
&= \mathcal{O}\left(\left(\Delta t\right)^{3/2}\right),
\end{aligned}
$$

where we also have used Lemma 3.4, the second part of (22), and that $\tau = c\Delta t$. Thus we have obtained that the correction operator, whenever applied, satisfies (C4) with $\gamma = 1/2$ when $\tau = c\Delta t$.

Remark. Note that the correction time τ is a parameter which has to be chosen properly in order to decrease the temporal splitting error. Observe therefore that it is possible to define the COS algorithm alternatively as follows: We let step 1 remain as before, whereas steps 2 and 3 are replaced by a new step $2'$. The new step consists of solving the parabolic problem

$$w_t + [f_{\mathrm{res}}(x,w) - \varepsilon\nu(w)_x]_x = 0, \qquad w(x,0) = u^{n+1/3}(x), \qquad (x,t) \in \mathbb{R} \times \langle 0, \Delta t] \tag{35}$$

in the weak sense, thereby yielding a splitting formula of the form

$$u^{n+1} = \left[\mathcal{P}^{f_{\mathrm{res}}}(\Delta t) \circ \mathcal{S}^{f_\delta}(\Delta t)\right] u^n.$$

Here $\mathcal{P}^{f_{\mathrm{res}}}(t)$ is the solution operator associated with the parabolic equation in (35). This alternative way of viewing COS is important for practical applications of the method, since the undetermined parameter τ now is eliminated. By construction, equation (35) contains the necessary information for ensuring correct balance between convection and diffusion. In fact, when the solution of (1) is simply a moving shock front, it is possible to show that (35) will generate shock layers of the correct size $\mathcal{O}(\varepsilon)$. Consequently, one should have this equation in mind when choosing the correction time τ.

Furthermore, note that the function $\mathcal{C}(\Delta t) \circ \mathcal{H}^\nu_{\Delta x}(\Delta t)u^{n+1/3}$ can be viewed as an approximation to the solution of (35) at time $t = \Delta t$. In practical applications one might introduce local time stepping in order to circumvent the problem of determining τ. Choose $\tilde{\Delta t}$ and $\tilde{n}$ such that $\tilde{n}\tilde{\Delta t} = \Delta t$ and use the approximation

$$\left[\mathcal{C}\left(\tilde{\Delta t}\right) \circ \mathcal{H}^\nu_{\Delta x}\left(\tilde{\Delta t}\right)\right]^{\tilde{n}} u^{n+1/3}(x) \tag{36}$$

instead of (35), or alternatively the Strang-type splitting

$$\left[\mathcal{H}^\nu_{\Delta x}\left(\frac{\tilde{\Delta t}}{2}\right) \circ \mathcal{C}\left(\tilde{\Delta t}\right) \circ \mathcal{H}^\nu_{\Delta x}\left(\frac{\tilde{\Delta t}}{2}\right)\right]^{\tilde{n}} u^{n+1/3}.$$

The reason for doing this is to better capture the nonlinearity inherent in (35), i.e., to obtain the correct balance between nonlinear convection and diffusion.

Finally, let us mention that an implementation based on the alternative COS algorithm (step 1 and $2'$) is presented and thoroughly tested in the companion paper [15]. Here, the solution of the parabolic equation (35) is approximated by a Petrov-Galerkin type finite element method. The main observation is that a substantial decrease in the temporal splitting error is obtained by solving (35) instead of merely the heat equation.

3. Convergence analysis. In this section we justify the term "approximate solution" by showing that a sequence of COS approximations is compact in L^{loc}_1, and that the limit of a convergent subsequence is a solution to (1). Convergence of (25) is obtained for any correction operator $\mathcal{C}(\tau)$ that satisfies the four regularity conditions (C1)–(C4). The compactness argument is standard in the context of conservation laws, and consists in establishing *à priori* bounds on the amplitude and the derivatives of the approximate solutions independent of the discretization parameters Δx, Δt, and δ.

We need to consider functions defined in the interval $\langle 0, T]$, and not merely on the time-strips $t = n\Delta t$. To accomplish this, define the sequence $\{u_\eta\}_{\eta>0}$ by

$$u_\eta(x,t) = \begin{cases} \mathcal{S}^{f_\delta}\left(2\left(t - n\Delta t\right)\right) u^n(x), & t \in \left\langle n\Delta t, \left(n+\frac{1}{2}\right)\Delta t\right\rangle, \\ \mathcal{H}^\nu\left(2\left(t - \left(n+\frac{1}{2}\right)\Delta t\right)\right) u^{n+1/3}(x), & t \in \left[\left(n+\frac{1}{2}\right)\Delta t, (n+1)\Delta t\right\rangle, \\ \mathcal{C}\left(\tau\right) u^{n+2/3}(x), & t = (n+1)\Delta t, \end{cases} \tag{37}$$

where $\eta = (\Delta x, \Delta t, \delta)$, $u^{n+1/3} = \mathcal{S}^{f_\delta}(\Delta t)\, u^n$, and $u^{n+2/3} = \mathcal{H}^\nu_{\Delta x}(\Delta t)\, u^{n+1/3}$. This method of extending $\{u^n\}_{n=0}^N$ to a function defined for all $t > 0$ was first used by Crandall and Majda [5], see also [14].

Lemma 3.1. *The following maximum principle holds*

$$\left\|u_\eta(\cdot,t)\right\|_\infty \le M_1, \qquad t \in \langle 0, T]\,. \tag{38}$$

Proof. By the construction based on solution of Riemann problems, we know that that the operator $\mathcal{S}^{f_\delta}(t)$ do not introduce new minima or maxima, and neither does the projection operator π nor the diffusion operator $\mathcal{H}^\nu(t)$. According to assumption (C1), $\left\|\mathcal{C}(\tau)u^{n+2/3}\right\|_\infty \le M_1$. Thus, Lemma 3.1 follows by induction on n. □

Lemma 3.2. *We have the following bound on the total variation*

$$TV\left(u_\eta(\cdot,t)\right) \le M_2, \qquad t \in \langle 0,T]. \tag{39}$$

Proof. Again by construction (see [13]) we have that $TV\left(\mathcal{S}^{f_\delta}(t)u^n\right) \le TV\left(u^n\right)$. From general theory of parabolic equations we know that the same is true for the operator $\mathcal{H}^\nu(t)$. According to (20), $TV\left(\pi g\right) \le TV\left(g\right)$. From assumption (C2) we have that $TV\left(\mathcal{C}(\tau)\right) \le M_2$. The lemma now follows by induction on n. ☐

Lemma 3.3. *Let there be finite constants C_1 and C_2 such that the function $u : \mathbb{R} \times \langle 0,T] \to \mathbb{R}$ satisfies*

$$\begin{aligned} \|u(\cdot,t)\|_\infty &\le C_1, && \text{for all } t \in \langle 0,T], \\ TV\left(u(\cdot,t)\right) &\le C_2, && \text{for all } t \in \langle 0,T]. \end{aligned}$$

Assume that $u(x,t)$ is weakly Lipschitz continuous in the time variable in the sense that

$$\left|\int \phi(x)\big(u(x,t_2) - u(x,t_1)\big)\,dx\right| \le Const. \cdot (t_2 - t_1)\left(\|\phi\|_\infty + \|\phi_x\|_\infty\right), \tag{40}$$

for all $\phi \in C_0^1$ and $0 < t_1 \le t_2 \le T$. Then there is a constant C, depending in particular on C_1 and C_2, such that the following interpolation result is valid

$$\|u(x,t_2) - u(x,t_1)\|_1 \le C\sqrt{t_2 - t_1}, \qquad 0 < t_1 \le t_2 \le T. \tag{41}$$

Proof. Let $\omega_h(x)$ be a standard C_0^∞-mollifier with smoothing radius h. Let $d(x) = u(x,t_2) - u(x,t_1)$, and define $\beta(x) = \operatorname{sign}(d(x))$ for $|x| \le r - h$ and $\beta(x) = 0$ for $|x| > r - h$, where $r > h$. Moreover, define $\beta^h = \omega_h * \beta$, and note that $\beta^h \in C^\infty$ with support in $[-r,r]$. By choosing $\phi = \beta^h$ in (40) and recalling several elementary properties of the mollifier function (see e.g. [17, Lemma 1]), it follows that

$$\begin{aligned} &\left|\int_{-r}^{r} |u(x,t_2) - u(x,t_1)|\,dx\right| \\ &\quad\le \left|\int_{-r}^{r} \big||d(x)| - \beta^h(x)d(x)\big|\,dx\right| + \left|\int_{-r}^{r} \left|\beta^h(x)d(x)\right|\,dx\right| \\ &\quad\le \tilde{C}_1 h + \tilde{C}_2\left(t_2 - t_1\right)/h. \end{aligned}$$

Here we mention that $\tilde{C}_1$ depends on C_2 and $\tilde{C}_2$ depends on C_1. Choosing $h = \sqrt{t_2 - t_1}$ and letting $r \to \infty$, we obtain (41). ☐

Lemma 3.4. *There is finite constant M, independent of η, such that*

$$\left\|u_\eta\left(\cdot,t_2\right) - u_\eta\left(\cdot,t_1\right)\right\|_1 \le M\sqrt{t_2 - t_1}, \qquad 0 < t_1 \le t_2 \le T. \tag{42}$$

Proof. According to the previous lemma it is enough, thanks to the *à priori* bounds (38) and (39), to establish weak Lipschitz continuity in time of the splitting solutions. Without loss of generality, assume that $t_1 = k\Delta t$ and $t_2 = l\Delta t$ for some integers k and l with $k \le l$. Integrating the differential equation for $w(x,t) = \mathcal{H}^\nu(t)u^{n+1/3}(x)$ against $\phi(x)$ over $\mathbb{R} \times \langle t, t+\Delta t\rangle$, gives

$$\begin{aligned} &\left|\int \phi(x)\big(\mathcal{H}^\nu(t+\Delta t)u^{n+1/3}(x) - \mathcal{H}^\nu(t)u^{n+1/3}(x)\big)\,dx\right| \\ &\quad= \left|\int_t^{t+\Delta t}\int \phi(x)\varepsilon\nu(w(x,\xi))_{xx}\,dx\,d\xi\right| \\ &\quad= \left|\int_t^{t+\Delta t}\int \phi(x)_x\varepsilon\nu(w(x,\xi))_x\,dx\,d\xi\right| \\ &\quad\le \varepsilon\|\nu'\|_\infty\|\phi_x\|_\infty \int_t^{t+\Delta t}\int |w(x,\xi)_x|\,dx\,d\xi \\ &\quad\le \text{Const.} \cdot \|\phi_x\|_\infty TV\left(u^{n+1/3}\right)\Delta t. \end{aligned}$$

Due to finite speed of propagation, we have a stronger estimate for the solution operator $\mathcal{S}^{f_\delta}(t)$, namely

$$\left|\int \left(\mathcal{S}^{f_\delta}(t+\Delta t)u^n(x) - \mathcal{S}^{f_\delta}(t)u^n(x)\right)\,dx\right| \leq \text{Const.} \cdot TV\left(u^n\right)\Delta t.$$

By assumption (C3), a similar estimate holds for the operator $\mathcal{C}(\tau)$. Using the recently obtained continuity estimates, the last part of (20), and Lemma 3.4, we readily compute that

$$\begin{aligned}
&\left|\int \phi(x)\big(u_\eta(x,t_2) - u_\eta(x,t_1)\big)\,dx\right| \\
&\quad \leq \sum_{n=k}^{l-1}\left|\int \phi(x)\big(u_\eta(x,(n+1)\Delta t) - u_\eta(x,n\Delta t)\big)\,dx\right| \\
&\quad \leq \sum_{n=k}^{l-1} \text{Const.} \cdot \Delta t\left(||\phi||_\infty + ||\phi_x||_\infty + ||\phi||_\infty\frac{\Delta x}{\Delta t} + ||\phi||_\infty\frac{\tau}{\Delta t}\right) \\
&\quad \leq \text{Const.} \cdot (t_2 - t_1)\left(||\phi||_\infty + ||\phi_x||_\infty\right),
\end{aligned}$$

where we also have taken into account the second part of (22) and that $\tau = c\Delta t$. The proof is now closed by appealing to Lemma 3.3. $\square$

Lemma 3.5. *Let $\{\eta\}$ be any sequence tending to zero. Then there exists a subsequence $\{\eta_j\}$ and a function u such that the corresponding subsequence $\{u_{\eta_j}\}$ converges to u in $C\left([0,T];L_1^{\text{loc}}(\mathbb{R})\right)$.*

Proof. Let us fix $t \in \langle 0,T]$, and let $\langle -r,r\rangle$ be a bounded open set in $\mathbb{R}$. From Lemma 3.3 and 3.4, we know that $\{u_\eta(\cdot,t)\}$ is bounded in $L_1(\langle -r,r\rangle) \cap BV(\langle -r,r\rangle)$. Using the compactness of the imbedding

$$L_1(\langle -r,r\rangle) \cap BV(\langle -r,r\rangle) \to L_1(\langle -r,r\rangle),$$

we know that there must exist a subsequence $\{u_{\eta_j}(\cdot,t)\}$ converging strongly in $L_1(\langle -r,r\rangle)$. Thus, from standard diagonal arguments one deduces the existence of a further subsequence, still denoted by $\{u_{\eta_j}(\cdot,t)\}$, and a function $u(\cdot,t) \in L_1^{\text{loc}}$ such that

$$u_{\eta_j}(\cdot,t) \to u(\cdot,t) \text{ in } L_1^{\text{loc}}.$$

Now let $\{t_m\}$ be a dense countable sequence in $[0,T]$. Applying the previous argument to each t_m and doing another diagonalization, we find a subsequence, also denoted by $\{u_{\eta_j}(\cdot,t)\}$, and a function u such that

$$u_{\eta_j}(\cdot,t_m) \to u(\cdot,t_m) \text{ in } L_1^{\text{loc}} \text{ for all } m.$$

Using continuity in time of u_{η_j}, i.e., Lemma 3.4, it is follows that $\{u_{\eta_j}(\cdot,t)\}$ is a Cauchy sequence in L_1^{loc} for all $t \in [0,T]$. Thus $u_{\eta_j}(\cdot,t) \to u(\cdot,t)$ in L_1^{loc} for all these t-values. A closer inspection will show that this convergence is, in fact, uniform in t for $t \in [0,T]$. This concludes the proof of the lemma. $\square$

Next, we justify the term "approximate solution" by showing:

Theorem 3.6 (Convergence). *Suppose that $u_0(x)$ is of bounded variation and that $f(u)$ and $\nu(u)$ are (locally) continuously differentiable. Let $\{\eta\}$ be any sequence of real numbers tending to zero. Then the corresponding sequence of COS solutions, $\{u_\eta(x,t)\}$, converges to the solution of the initial value problem*

$$u_t + f(u)_x = \varepsilon\nu(u)_{xx}, \qquad u(x,0) = u_0(x), \qquad (x,t) \in \mathbb{R} \times \langle 0,T].$$

Proof. From Lemma 3.5 we know that there exists a subsequence, for notational convenience denoted by $\{u_\eta\}$, which converges to a function u. We will continue along the lines of [14], by showing that the limit function u is a weak solution in the sense (8). For $\phi \in C_0^\infty(\mathbb{R} \times \langle 0,T])$, define the functional

$$\mathcal{L}(u,f,\phi) = \iint_0^T \left(u\phi_t + f_\delta(u)\phi_x + \varepsilon\nu(u)\phi_{xx}\right)\,dt\,dx + \int u_0(x)\phi(x,0)\,dx.$$

We shall prove that $\mathcal{L}_\phi(u) = 0$ for all proper test functions ϕ. Let $v_n(t) = \mathcal{S}^{f_\delta}(t)u^n$, $t \in \langle 0, \Delta t]$, and define a new test function $\tilde{\phi}$ by $\tilde{\phi}(x,t) = \phi(x, t/2)$. Then the following equality holds

$$
\begin{aligned}
&\iint_{n\Delta t}^{(n+\frac{1}{2})\Delta t} \left(\frac{1}{2} u_\eta \phi_t + f_\delta\left(u_\eta\right)\phi_x\right) dt\,dx \\
&= \frac{1}{2}\iint_0^{\Delta t} \left(v_n(x,\xi)\tilde{\phi}\left(x, \xi + 2n\Delta t\right)_\xi + f_\delta(v_n(x,\xi))\tilde{\phi}\left(x,\xi+2n\Delta t\right)_x\right) d\xi\,dx \\
&= \frac{1}{2}\int u_\eta\left(x, \left(n+\frac{1}{2}\right)\Delta t\right)\phi\left(x, \left(n+\frac{1}{2}\right)\Delta t\right) dx \\
&\quad - \frac{1}{2}\int u_\eta\left(x, n\Delta t\right)\phi\left(x, n\Delta t\right) dx,
\end{aligned}
\tag{44}
$$

where we have used the substitution $\xi = 2(t - n\Delta t)$. Similarly

$$
\begin{aligned}
&\iint_{(n+\frac{1}{2})\Delta t}^{(n+1)\Delta t} \left(\frac{1}{2} u_\eta \phi_t + \nu(u_\eta)\phi_{xx}\right) dt\,dx \\
&= \frac{1}{2}\int u_\eta\left(x, (n+1)\Delta t-\right)\phi\left(x, (n+1)\Delta t\right) dx \\
&\quad - \frac{1}{2}\int u_\eta\left(x, \left(n+\frac{1}{2}\right)\Delta t\right)\phi\left(x, \left(n+\frac{1}{2}\right)\Delta t\right) dx.
\end{aligned}
\tag{45}
$$

Adding (44) and (45), and summing over $n = 0, \dots, N-1$, we obtain

$$
|\mathcal{L}(u_\eta, f_\delta, \phi)| \leq \sum_{i=1}^{4} |E_i|,
$$

where

$$
\begin{aligned}
E_1 &= \sum_{n=0}^{N-1}\int\left(\int_{n\Delta t}^{(n+1)\Delta t} f_\delta\left(u_\eta\right)\phi_x\,dt - 2\int_{n\Delta t}^{(n+\frac{1}{2})\Delta t} f_\delta\left(u_\eta\right)\phi_x\,dt\right) dx, \\
E_2 &= \varepsilon\sum_{n=0}^{N-1}\int\left(\int_{n\Delta t}^{(n+1)\Delta t} \nu(u_\eta)\phi_{xx}\,dt - 2\int_{(n+\frac{1}{2})\Delta t}^{(n+1)\Delta t} \nu(u_\eta)\phi_{xx}\,dt\right) dx, \\
E_3 &= \sum_{n=0}^{N-1}\int\left(\pi u_-^{n+2/3}(x) - u_-^{n+2/3}(x)\right)\phi\left(x, (n+1)\Delta t\right) dx, \\
E_4 &= \sum_{n=0}^{N-1}\int\left(\mathcal{C}(\tau)u^{n+2/3}(x) - u^{n+2/3}(x)\right)\phi\left(x, (n+1)\Delta t\right) dx,
\end{aligned}
$$

and $u_-^{n+2/3}(x)$ denotes $\mathcal{H}^\nu\left(\Delta t\right)u^{n+1/3}(x)$. Let us first consider E_2, which we rewrite as $E_2 = E_{2,1} + E_{2,2}$, where

$$
\begin{aligned}
E_{2,1} &= \varepsilon\sum_{n=0}^{N-1}\int\left(\int_{n\Delta t}^{(n+1)\Delta t}\phi_{xx}(x,t)\,dt - 2\int_{(n+\frac{1}{2})\Delta t}^{(n+1)\Delta t}\phi_{xx}(x,t)\,dt\right)\nu(u_\eta(x, n\Delta t))\,dx \\
E_{2,2} &= \varepsilon\sum_{n=0}^{N-1}\int\left(\int_{n\Delta t}^{(n+1)\Delta t}\left(\nu(u_\eta(x,t)) - \nu(u_\eta(x,n\Delta t))\right)\phi_{xx}(x,t)\,dt\right. \\
&\qquad\left. - 2\int_{(n+\frac{1}{2})\Delta t}^{(n+1)\Delta t}\left(\nu(u_\eta(x,t)) - \nu(u_\eta(x,n\Delta t))\right)\phi_{xx}(x,t)\,dt\right) dx.
\end{aligned}
$$

Since $\phi \in C_0^\infty(\mathbb{R}\times\langle 0, T])$, we may write $\phi_{xx}(x,t) = \phi_{xx}(x, n\Delta t) + \mathcal{O}(t - n\Delta t)$ for $t \geq n\Delta t$. With the aid of this, it is easy to see that $|E_{2,1}| = \mathcal{O}\left(\Delta t\right)$. The L^1-continuity in time (42) implies that

$$
|E_{2,2}| \leq \text{Const.}\cdot\varepsilon\|\nu'\|_\infty\sum_{n=0}^{N-1}\int_{n\Delta t}^{(n+1)\Delta t}\int|u_\eta(x,t) - u_\eta(x,n\Delta t)|\,dx\,dt = \mathcal{O}\left(\sqrt{\Delta t}\right).
$$

Consequently, $|E_2| = \mathcal{O}(\sqrt{\Delta t})$. Similarly, one can deduce that $|E_1| = \mathcal{O}\left(\sqrt{\Delta t}\right)$.

Next, the error due to the projection operator can be bounded as follows

$$
\begin{aligned}
|E_3| = &\left| \sum_{n=0}^{N-1} \sum_j \int_{z_j^n} \left(\pi u_-^{n+2/3}(x) - u_-^{n+2/3}(x)\right) \phi\left(x_j^n, (n+1)\Delta t\right)\, dx \right.\\
&\left. + \sum_{n=0}^{N-1} \sum_j \int_{z_j^n} \left(\pi u_-^{n+2/3}(x) - u_-^{n+2/3}(x)\right) \left(\phi\left(x, (n+1)\Delta t\right) - \phi\left(x_j^n, (n+1)\Delta t\right)\right) dx \right| \\
&= \left| \sum_{n=0}^{N-1} \sum_j \int_{z_j^n} \left(\pi u_-^{n+2/3}(x) - u_-^{n+2/3}(x)\right) \left(\phi\left(x, (n+1)\Delta t\right) - \phi\left(x_j^n, (n+1)\Delta t\right)\right) dx \right| \\
&\le \|\phi_x\|_\infty \Delta x \sum_{n=0}^{N-1} \sum_j \int_{z_j^n} \left|\pi u_-^{n+2/3}(x) - u_-^{n+2/3}(x)\right| dx \\
&\le TV\left(u_-^{n+2/3}\right) \|\phi_x\|_\infty N (\Delta x)^2 = \mathcal{O}\left(\Delta x\right),
\end{aligned}
$$

where we have used (18), (20), Lemma 3.3, and the first part of (22).

It remains to estimate E_4. Using assumption (C4), we obtain that $|E_4| = \mathcal{O}\left(\sqrt{\Delta t} + (\Delta t)^\gamma\right)$. Thus we have arrived at

$$
|\mathcal{L}(u_\eta, f_\delta, \phi)| = \mathcal{O}\left(\Delta x + \sqrt{\Delta t} + (\Delta t)^\gamma\right).
$$

Having the second part of (21) in mind, we easily calculate

$$
\begin{aligned}
|\mathcal{L}(u_\eta, f, \phi)| &= |\mathcal{L}(u_\eta, f_\delta, \phi)| + |\mathcal{L}(u_\eta, f, \phi) - \mathcal{L}(u_\eta, f_\delta, \phi)| \\
&= \mathcal{O}\left(\Delta x + \sqrt{\Delta t} + (\Delta t)^\gamma\right) + \left| \int \int_0^T \left(f\left(u_\eta\right) - f_\delta\left(u_\eta\right) \phi_x\right) dt\, dx \right| \\
&= \mathcal{O}\left(\Delta x + \sqrt{\Delta t} + (\Delta t)^\gamma\right) + \text{Const.} \cdot \|f - f_\delta\|_\infty \\
&= \mathcal{O}\left(\Delta x + \sqrt{\Delta t} + (\Delta t)^\gamma + \delta^2\right).
\end{aligned}
$$

In view of Lebesgue's dominated convergence theorem, we conclude that

$$
\mathcal{L}(u, f, \phi) = \lim_{\eta \to 0} \mathcal{L}\left(u_\eta, f, \phi\right) = 0,
$$

where $u = \lim_{\eta \to 0} u_\eta$. Since the solution of (1) is unique, the whole sequence $\{u_\eta\}$ converges. This concludes the proof of the theorem. $\square$

4. A discrete method. In this section consider an application of the corrected operator splitting algorithm in the case of a linear diffusion coefficient. To obtain a fully discrete version of

$$
u^{n+1} = \left[\mathcal{C}(\tau) \circ \mathcal{H}_{\Delta x}^{\nu}(\Delta t) \circ \mathcal{S}^{f_\delta}(\Delta t)\right] u^n, \qquad n = 0, \dots, N-1, \qquad N\Delta t = T,
$$

we must choose a proper numerical scheme for integrating the linear heat equation (4). From a computationally point of view we ought to impose some kind of boundary conditions on the parabolic equation (1). For ease of presentation, consider therefore the parabolic equation

$$
u_t + f(u)_x = \varepsilon u_{xx}, \qquad (x,t) \in \langle a, b \rangle \times \langle 0, T], \tag{46}
$$

with initial and boundary data imposed as follows

$$
\begin{aligned}
u(x,0) &= u_0(x), && x \in \langle a, b\rangle, \\
u(a,t) &= u_a, && t \in \langle 0, T], \\
u(b,t) &= u_b, && t \in \langle 0, T],
\end{aligned} \tag{47}
$$

where $\langle a, b\rangle \subset \mathbb{R}$, and u_a and u_b are constants. We assume that the initial and boundary data are such that the convection solution (23) remains consistent with the boundary data for all time $t \in \langle 0, T]$.

We use a finite element method for the solution of the linear heat equation (35), with "elements" determined by the discontinuities in the front tracking solution $\mathcal{S}^{f_\delta}(t)v_0(x)$. In order to ensure convergence of our method, we add nodes whenever the spacing between two discontinuities becomes larger than Δx. Let $\mathcal{H}^\nu_{\Delta x}(t)$ denote the operator which takes an initial function

$$w_0(x) = \sum_{j=1}^{M} \xi_j \varphi_j(x)$$

to the (projected) approximate solution of (35) obtained by the element method using basis functions $\varphi_j(x)$, $j = 1, \dots, M$. We assume that these basis functions are associated with Δx such that $M \to \infty$ as $\Delta x \to 0$. The approximate solution is then written as

$$\mathcal{H}^\nu_{\Delta x}(t) w_0(x) = \pi \left(\sum_{j=1}^{M} \xi_j(t) \varphi(x) \right),$$

where $\xi_j(t)$ is the solution of the following system of ordinary differential equations

$$\sum_{j=1}^{M} \dot{\xi}_j(t) \left(\varphi_i, \varphi_j\right) + \varepsilon \sum_{j=1}^{M} \xi_j(t) l\left(\varphi_i, \varphi_j\right) = 0, \qquad i = 1, \dots, M. \tag{48}$$

Here $l(\cdot,\cdot)$ denotes the usual bilinear form associated with the right-hand side of (35). For a description of finite element methods for problems such as (35), see [16].

Now let $u_\eta(x,t)$ denote the fully discrete analog of (37). By mimicking the proofs in section 3, it is not difficult to prove the following lemma.

Lemma 4.1. *We have that the fully discrete $u_\eta(x,t)$ satisfies the following three à priori estimates*

$$\|u_\eta(\cdot,t)\|_\infty \le K, \qquad TV\left(u_\eta(\cdot,t)\right) \le K, \qquad \|u_\eta(\cdot,t_2) - u_\eta(x,t_1)\|_1 \le Kh(|t_2 - t_1|),$$

where K is some number independent of the discretization parameters η, and $h(t)$ is a uniformly continuous function with $h(0) = 0$.

Consequently, we obtain compactness of the sequence $\{u_\eta\}$. Furthermore, it is not difficult to demonstrate that the limit of a converging subsequence is a weak solution to (46), and then that the following convergence theorem is valid.

Theorem 4.2. *Assume that $u_0(x)$ is of bounded variation and that $f(u)$ is continuously differentiable. Then $u(x,t) = \lim_{\eta \to 0} u_\eta(x,t)$ is the solution of the initial-boundary value problem (46) and (47).*

We close this paper by presenting some numerical experiments with the fully discrete COS method applied to the Burgers equation [4] and the Buckley-Leverett equation [22]. To clearly demonstrate the effect of the correction operator, COS is compared with the standard splitting OS. Its implementation coincides with the one obtained by setting the correction time to zero in COS.

In the computations presented below we have set the distance between the interpolation points in the flux function to $\delta = 0.05$. The spatial domain is discretized using 64 nodes, and we are consequently using one time step to reach final computing time $t = T$. The diffusion coefficient ε is kept fixed at 0.005 and the correction threshold parameter c_{tr} at 0.1. The integration of (48) is done by Euler's backward method. Furthermore, the finite element method uses piecewise linear basis functions of the type

$$\varphi_j(x) = \begin{cases} 0, & \text{if } x \le x^n_{j-1}, \\ 1, & \text{if } x = x^n_j, \\ 0, & \text{if } x \ge x^n_{j+1}, \end{cases}$$

where the numbers $\{x^n_j\}$ are the grid nodes that coincide with the discontinuities in the front tracking solution.

According to (25) the COS solution at final computing time $t = T$ is piecewise constant. In applications one should replace this solution by a proper piecewise linear function so that second order accuracy in space is obtained. However, for clarity of presentation, we have chosen to visualize the OS solutions as piecewise linear functions and the COS solutions as step functions. For comparison, we have generated "exact" solutions using OS with very fine discretization parameters.

Example 1 (The Burgers equation). We first consider the Burgers equation

$$u_t + \left(\frac{1}{2}u^2\right)_x = \varepsilon u_{xx}, \qquad (x,t) \in \langle 0,1\rangle \times \langle 0,T],$$

with initial and boundary data; $u(x,0) = \chi_{\langle 0,1\rangle - [0.2,0.6]}(x)$ and $u|_{x=0} = u|_{x=1} = 0$. We compute solutions up to time $T = 0.3$ using one time step ($\Delta t = 0.3$). In Figure 4.1 we show the results of COS using correction times $\tau = 0.2 \cdot \Delta t$ (middle plot) and $\tau = 0.37 \cdot \Delta t$ (right plot). The residual flux term generated by the COS algorithm is shown in the leftmost plot. The true solution is non-monotone and contains a strong shock front located around 0.75. Our *only* interest is to see if COS can resolve the shock layer using one time step. For the problems under consideration we know that the size of the shock layer should be $\mathcal{O}(\varepsilon)$ [25]. We see that the layer produced by OS is (several) orders of magnitude too wide. A slight improvement is seen for COS with correction time $\tau = 0.2 \cdot \Delta t$. However, by increasing the correction parameter to $\tau = 0.37 \cdot \Delta t$, COS obtains the correct size of the shock layer. To sum up, for this example we see that the correction operator has the promised properties; that is, it manages to correct most of the errors introduced by the heat equation, at least when then correction time is properly chosen.

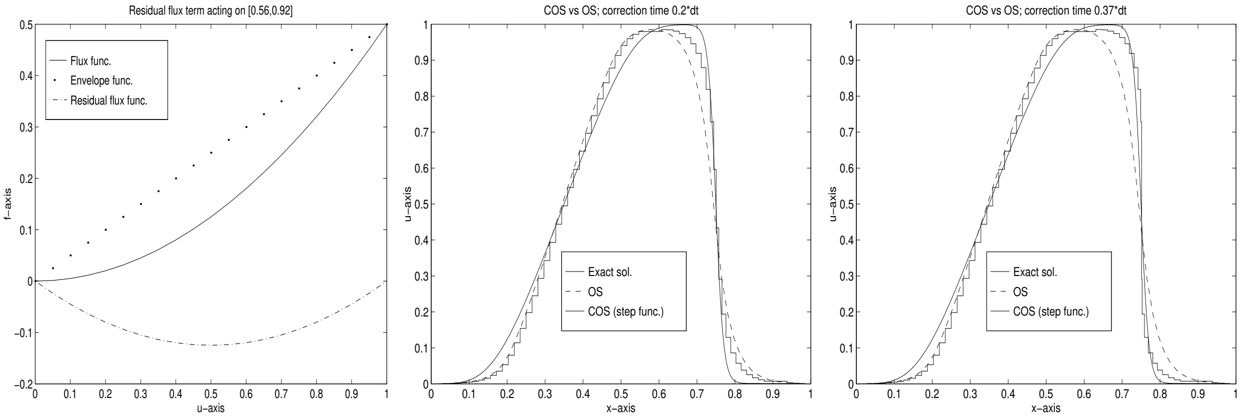

Figure 4.1. Example 1. The computations are done with 64 grid cells, 20 linear interpolation points, and 1 time step. Left: Residual flux function. Middle: COS with correction time 0.06. Right: COS with correction time 0.111. For a proper choice of the correction time we clearly see (right) that the error introduced by the heat equation is corrected so that (almost) the right size of the shock layer is obtained.

Example 2 (The Buckley-Leverett equation). Next we consider the two phase flow equation

$$u_t + \left(\frac{u^2}{u^2 + (1-u)^2}\right)_x = \varepsilon u_{xx}, \qquad (x,t) \in \langle 0,1.5\rangle \times \langle 0,T], \tag{49}$$

with initial data

$$u_0(x) = \begin{cases} 1 - 3x, & \text{for } 0 < x < \frac{1}{3}, \\ 0, & \text{for } \frac{1}{3} \le x < 1.5, \end{cases}$$

and boundary data; $u|_{x=0} = 1$ and $u|_{x=1.5} = 0$. We now compute solutions up to time $T = 0.6$ using one time step ($\Delta t = 0.6$). The true solution still contains a strong shock front. The results are presented in Figure 4.2. Also this time we see that OS is far too diffusive in the shock front region. Furthermore, COS manages to resolve the shock front when the correction time is $\tau = 0.18 \cdot \Delta t$, see the rightmost picture.

Finally, let us consider (49) with non-monotone data; $u(x,0) = \chi_{\langle 0,1.5\rangle - [0.2,0.6]}(x)$ and $u|_{x=0} = u|_{x=1.5} = 0$, which results in the generation of two residual flux terms, see Figure 4.3 (left and middle). Solutions are computed up to final time $T = 0.4$ in one step ($\Delta t = 0.4$) and they are depicted in Figure 4.3 (right plot). Here we have only shown COS with an "optimal" correction time; $\tau = 0.2 \cdot \Delta t$, in which case the desired correction effect in the shock regions is as evident as in the previous computations. Observe that there is a slight loss of amplitude that comes from the diffusion step and that this loss is a consequence of the large time step. This type of error cannot be corrected by adjusting the parameter τ. Instead a smaller time step, or alternatively the splitting formula (36), should be employed. The "local time stepping" formula (36) is designed so that significant loss of amplitude cannot occur during the diffusion step.

5. Concluding remarks. The standard two-step viscous splitting method has a tendency to be too diffusive when applied to convection dominated parabolic equations. In the present paper we have shown that it is

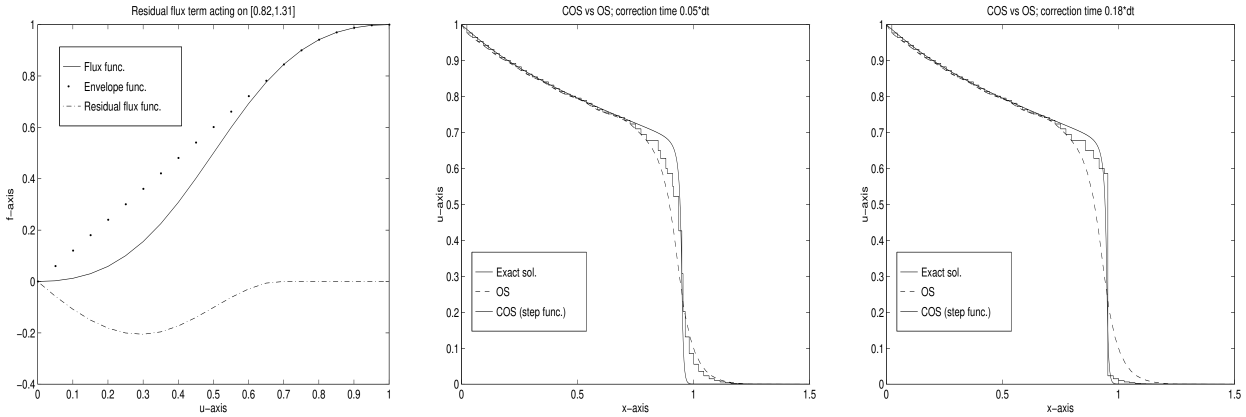

Figure 4.2. Example 2 (monotone data). The computations are done with 64 grid cells, 20 linear interpolation points, and 1 time step. Left: Residual flux function. Middle: COS with correction time 0.03. Right: COS with correction time 0.108.

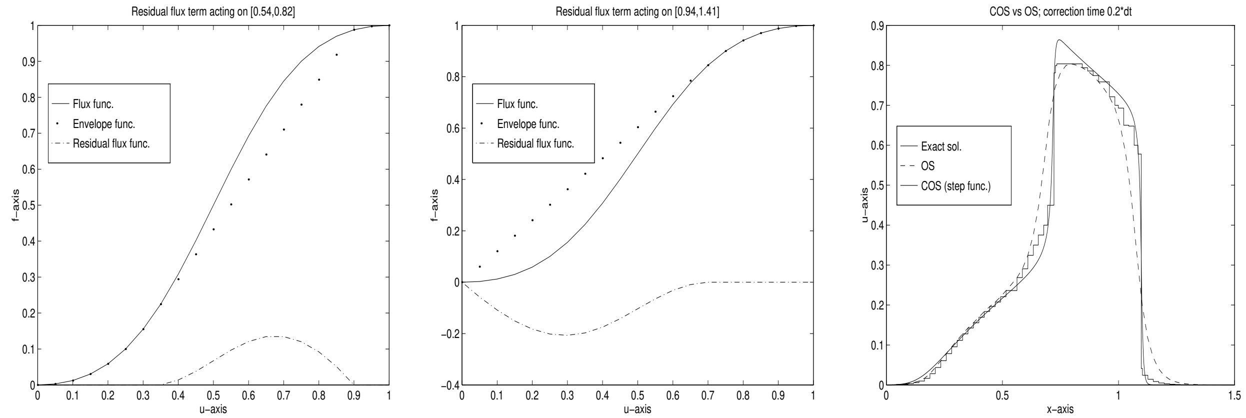

Figure 4.3. Example 2 (non-monotone data). The computations are done with 64 grid cells, 20 linear interpolation points, and 1 time step. Left and Middle: Residual flux function. Right: COS with correction time 0.08. Also in this non-monotone case (with several anti-diffusive flux terms) we see that COS resolves the shock front notably better than OS. Note that there is a loss of amplitude due to a large time step (see the text for a partial discussion).

possible to construct a residual flux term which can be employed in a third step to correct smearing errors introduced in the diffusion step. Alternatively, as pointed out in section 2, this residual term can also be included in the diffusion step, yielding a more complicated equation modeling diffusion. However, this equation contains the necessary information to ensure the correct balance between convection and diffusion, see [15] for numerical examples verifying this claim. The numerical examples given in the present paper indicate that the correction (anti-diffusive) effect is significant in the shock layer regions when the residual flux term is used in an explicit correction operator (25). The front tracking method plays a central role in the construction of the residual flux term.

Finally, we would like to mention that we plan to extend this approach to degenerate (mixed hyperbolic/parabolic) equations [27] and equations with explicit spatially dependent coefficients, so that our techniques can be applied to the equations arising in the modeling of two phase flow in a porous medium [22].

Acknowledgment. The authors thank S. Evje, M. S. Espedal, H. K. Dahle, and K-A. Lie for interesting discussions.

References

1. C. Bardos, A. Y. LeRoux, J. C. Nedelec, *First order quasilinear equations with boundary conditions*, Comm. Partial Diff. Eqs. **4(9)** (1979), 1017–1034.
2. J. T. Beale, A. Majda, *Rates of convergence for viscous splitting of the Navier-Stokes equations*, Math. Comp. **37(156)** (1981), 243–259.
3. F. Black, M. Scholes, *The pricing of options and corporate liabilities*, J. Pol. Econ. **81** (1973), 637–659.
4. J. M. Burgers, *The Nonlinear Diffusion Equation*, Reidel, Dordrecht, 1974.
5. M. Crandall, A. Majda, *The method of fractional steps for conservation laws*, Numer. Math. **34** (1980), 285–314.
6. C. M. Dafermos, *Polygonal approximations of solutions of the initial value problem for a conservation law*, J. Math. Anal. Appl. **38** (1972), 33–41.
7. H. K. Dahle, *Adaptive characteristic operator splitting techniques for convection-dominated diffusion problems in one and two space dimensions*, PhD. Thesis, Dept. of Math., University of Bergen, Norway (1988).

8. C. N. Dawson, *Godunov-mixed methods for advective flow problems in one space dimension*, SIAM J. Num. Anal. **28(5)** (1991), 1282–1309.
9. J. Douglas, T. F. Russell, *Numerical methods for convection dominated diffusion problems based on combining the method of characteristiscs with finite element or finite difference procedures*, SIAM J. Numer. Anal. **19** (1982), 871–885.
10. M. Espedal, R. E. Ewing, *Characteristic Petrov–Galerkin subdomain methods for two-phase immiscible flow*, Comput. Methods. Appl. Mech. Engrg. **64** (1987), 113–135.
11. R. E. Ewing, *Operator splitting and Eulerian-Lagrangian localized adjoint methods for multiphase flow*, The mathematics of finite elements and Applications VII MAFELAP (J. Whiteman, eds.), Academic Press, San Diego, 1990, pp. 215–232.
12. R. E. Ewing, T. F. Russell, *Efficient time stepping methods for miscible displacement problems in porous media*, SIAM J. Numer. Anal. **19** (1982), 1–67.
13. H. Holden, L. Holden, R. Høegh-Krohn, *A numerical method for first order nonlinear scalar conservation laws in one–dimension*, Comput. Math. Applic. **15** (1988), 595–602.
14. K. Hvistendahl Karlsen, N. H. Risebro, *An operator splitting method for nonlinear convection-diffusion equations*, To appear in Numer. Math.
15. K. Hvistendahl Karlsen, K. Brusdal, H. K. Dahle, S. Evje, K-A. Lie, *The corrected operator splitting approach applied to a nonlinear advection-diffusion problem*, Preprint, University of Bergen (1997).
16. C. Johnson, *Numerisk lösning av partiella differentialekvationer med finita element metoden*, Studentlitteratur, Lund, 1980.
17. S. N. Kružkov, *First order quasilinear equations in several independent variables*, Mat. Sbornik **10** (1970), 217–243.
18. S. N. Kružkov, O. Oleinik, *Quasi-linear second order equations with several indendent variables*, Uspehi. Mat. Nauk **16** (1961), 115–155.
19. J. O. Langseth, *On an implementation of a front tracking method for hyperbolic conservation laws*, Advances in Engineering Software **26** (1996).
20. M. J. Lighthill, G.B. Whitham, *On kinematic waves. II. Theory of traffic flow on long crowded roads*, Proc. Roy. Soc. **229A** (1955), 317–345.
21. O. Oleinik, *Discontinuous solutions of non-linear differential equations*, Amer. Math. Soc. Transl. **26** (1963), 95–172.
22. D.W. Peaceman, *Fundamentals of Numerical Reservoir Simulation*, Elsevier, Amsterdam, 1977.
23. N. H. Risebro, A. Tveito, *Front tracking applied to a nonstrictly hyperbolic system of conservation laws*, SIAM J. Sci. Comput. **12** (1991), 1401–1419.
24. T. F. Russell, *Galerkin time stepping along characteristics for Burgers' equation*, Scientific computing (R. Stepleman, eds.), IMACS, North-Holland, 1983, pp. 183–192.
25. J. S. Scroggs, *Shock-layer bounds for a singularly perturbed equation*, Quarterly of Applied Mathematics **LII(3)** (1995), 423–431.
26. J. A. Trangenstein, *Mathematical structure of the black-oil model for petroleum reservoir simulation*, SIAM J. Appl. Math. **49** (1989), 749–783.
27. A. I. Vol'pert and S. I. Hudjaev, *Cauchy's problem for degenerate second order quasilinear parabolic equations*, Math. USSR Sbornik **7** (1969), 365–387.

Department of Mathematics
University of Bergen
Johs. Brunsgt. 12
N–5008 Bergen, Norway
(Hvistendahl Karlsen)
E-mail address: kennethk@mi.uib.no

Department of Mathematics
University of Oslo
P.O. Box 1053, Blindern
N–0316 Oslo, Norway
(Risebro)
E-mail address: nilshr@math.uio.no