Mathematical Models and Methods in Applied Sciences
© World Scientific Publishing Company

Error estimates for forward Euler shock capturing finite element approximations of the one dimensional Burgers’ equation

Erik Burman
Department of Mathematics,
University College London, Gower Street, London,
UK–WC1E 6BT,
e.burman@ucl.ac.uk

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We propose an error analysis for a shock capturing finite element method for the Burgers’ equation using the duality theory due to Tadmor. The estimates use a one sided Lipschitz stability ($Lip^+$-stability) estimate on the discrete solution and are obtained in a weak norm, but thanks to a total variation a priori bound on the discrete solution and an interpolation inequality, error estimates in $L^p$-norms ($1 \leq p < \infty$) are deduced. Both first order artificial viscosity and a nonlinear shock capturing term that formally is of second order are considered. For the discretization in time we use the forward Euler method.

In the numerical section we verify the convergence order of the nonlinear scheme using the forward Euler method and a second order strong stability preserving Runge-Kutta method. We also study the $Lip^+$-stability property numerically and give some examples of when it holds strictly and when it is violated.

Keywords: burgers’ equation; shock capturing; stabilized finite element method; error estimates.

AMS Subject Classification: 76M10, 65M12, 65M15, 65M60

1. Introduction

There exists a vast litterature on the design and convergence of numerical methods for nonlinear scalar conservation laws dating back to the seminal work of Krushkov. Error estimates are typically obtained using entropy stability and the so-called variable doubling technique. Asymptotic results have also been obtained using entropy stability and compensated compactness. Different numerical methods have been considered. For work on finite difference methods we refer to Refs. 22, 23, 12, 13, 26, 11, 14, 24, finite volume methods to Refs. 9, 10, 4, finite element methods to Refs. 20, 21, 6, 28 and finally spectral methods to Refs. 25, 31, 5, 16. For an introduction to these techniques we refer to the review article by Cockburn or the one by Tadmor.

The aim of this work is to analyse a shock capturing finite element method
proposed by the author\textsuperscript{2} using the framework developed by Nessyahu and Tadmor\textsuperscript{26} which in its turn draws from the stability analysis of Tadmor.\textsuperscript{30} This appears not to have been proposed earlier, which is curious since similar duality techniques were proposed later for a posteriori error estimation for finite element methods using duality.\textsuperscript{23,19} The key argument of the analysis is to use a duality argument to get continuous dependence on initial data for the adjoint perturbation equation of the Burgers’ equation in the $\text{Lip}$-norm. This leads to error estimates in the $\text{Lip'}$-norm, i.e. the norm associated to the dual of the space of Lipschitz-continuous functions. The dual stability estimate crucially relies on the stability properties of the numerical method, in particular that the discrete solution satisfies the Oleinik E-condition,\textsuperscript{29} which corresponds to a one sided Lipschitz condition. To prove error estimates the finite element residual must also be a priori bounded by initial data which typically requires a bound in the BV-norm. Estimates in general $L^p$-norms may then be recovered using interpolation between the $\text{Lip'}$-norm and the BV-norm.

We will discuss how a nonlinear shock capturing finite element method can be designed so that the resulting method allows for both a posteriori and a priori error estimates derived using the framework proposed by Nessyahu and Tadmor. Our analysis uses the following three main ingredients,

\begin{itemize}
  \item one sided Lipschitz stability estimates for the finite element methods,
  \item a priori stability estimates on a linearized dual problem derived by Tadmor,\textsuperscript{30}
  \item Galerkin orthogonality and approximability.
\end{itemize}

For the case of the nonlinear viscosity proposed in Ref. 2 it is not possible to impose Oleinik’s E-condition at every time step, so some growth of the discrete positive gradient is inevitable. Here we prove a BV bound for the fully discrete case using nonlinear viscosity and then suggest a slightly modified form of the viscosity, which allows us to prove $\text{Lip}^+$-stability. The key idea is to smooth the viscosity in a neighbourhood of the maximum gradient. Some numerical results using the nonlinear scheme showing its sharp resolution of non-smooth solutions and (global) high order convergence for smooth solutions are given. We also investigate numerically what conditions are necessary on the time-step in order for the discrete solution computed using the nonlinear viscosity to be $\text{Lip}^+$-stable. Below we will let $C$ denote a generic constant that may change at each appearance, but which is always independent of $h$. We will also use the notation $a \lesssim b$ for $a \leq Cb$. 
2. The Burgers’ equation

Consider the model case of the Burgers’ equation on the space-time domain \( Q := \mathbb{R} \times I \), where \( I = (0, T) \).

\[
\partial_t u + \partial_x \frac{u^2}{2} = 0 \quad \text{in} \quad Q
\]

(2.1)

\[ u(x,0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}. \]

Assume that \( u_0 \) has compact support and bounded variation. It is well known that the equation (2.1) admits a unique entropy solution that satisfies Oleinik’s E-condition:

\[
\frac{u(x+h,t) - u(x,t)}{h} < \frac{E}{t} \quad \text{for some} \quad E > 0 \quad \text{and} \quad \forall x \in \mathbb{R}, \forall h > 0.
\]

(2.2)

Indeed the classical entropy condition and (2.2) are known to be equivalent. The entropy solution is known to satisfy a maximum principle on the form:

\[
\text{ess sup}_{(x,t) \in Q} |u(x,t)| \leq \text{ess sup}_{x \in \mathbb{R}} |u_0(x)|.
\]

(2.3)

Oleinik’s E-condition is also equivalent to the satisfaction of a one sided Lipschitz condition of the solution, more precisely the following a priori estimate is known to hold:\(^30\)

\[
\|u(\cdot,t)\|_{Lip^+} \leq \frac{1}{\|u_0\|_{Lip^+} + t}
\]

where

\[
\|u(\cdot,t)\|_{Lip^+} := \text{ess sup}_{x \neq y} \left( \frac{u(x,t) - u(y,t)}{x-y} \right)^+.\]

We also recall the Lip-seminorm and the associated dual seminorm,

\[
\|u\|_{Lip} := \text{ess sup}_{x \neq y} \left| \frac{u(x) - u(y)}{x-y} \right|,
\]

\[
\|u\|_{Lip'} := \sup_{v \in Lip, \|v\|_{Lip} = 1} (u - \bar{u},v), \quad \bar{u} := \int_{\mathbb{R}} u \, dx.
\]

The associated function spaces will be denoted by Lip and Lip’. We will also use the norm of bounded variation that we define as

\[
\|u\|_{BV} := \|u\|_{L^1(\mathbb{R})} + BV(u),
\]

where

\[
BV(u) := \sup \left\{ \int_{\mathbb{R}} u \partial_x \phi \, dx : \phi \in C^1_c(\mathbb{R}), \|\phi\|_{L^\infty(\mathbb{R})} = 1 \right\}.
\]

Recall that for functions in \( W^{1,1}(\mathbb{R}) \) there holds \( \|u\|_{W^{1,1}(\mathbb{R})} = \|u\|_{BV} \).
3. Artificial viscosity finite element method

We denote the computational nodes by \( x_i := ih, \) \( i \in \mathbb{Z}, \) defining the elements \( K_i := [x_i, x_{i+1}], \) \( i \in \mathbb{Z}, \) and the standard, piecewise linear and continuous, nodal basis functions \( \{v_i\}_{i=-\infty}^{\infty}, \) such that \( v_i(x_j) = \delta_{ij}, \) with \( \delta_{ij} \) the Kronecker delta. The finite element space is given by

\[
V_h := \left\{ \sum_{i \in \mathbb{Z}} u_i v_i, \text{ where } u_i \in \mathbb{R} \right\}.
\]

We define the standard \( L^2 \) inner product on \( X \subset \mathbb{R} \) by

\[
(v_h, w_h)_X := \int_X v_h w_h \, dx.
\]

The discrete form corresponding to mass-lumping reads

\[
(v_h, w_h)_h := \sum_{i=-\infty}^{\infty} v_h(x_i) w_h(x_i) h.
\]

The associated norms are defined by \( \|v\|_X := (v, v)^{\frac{1}{2}}_X, \) for all \( v \in L^2(X), \) if \( X \) coincides with \( \mathbb{R} \) the subscript is dropped, and \( \|v_h\|_h := (v_h, v_h)^{\frac{1}{2}}_h \) for all \( v_h \in V_h. \)

Note that, by norm equivalence on discrete spaces, for all \( v_h \in V_h \)

\[
\|v_h\|_h \lesssim \|v_h\| \lesssim \|v_h\|_h.
\]

Using the above notation the artificial viscosity finite element space semi-discretization of (2.1) reads, given \( u_0 \in BV(\mathbb{R}) \cap Lip^+(\mathbb{R}), \) with compact support, find \( u_h(t) \in V_h \) such that \( u_h(0) = \pi_{BV} u_0, \) where \( \pi_{BV} \) is a special interpolation operator to be defined, and

\[
(\partial_t u_h, v_h)_h + \left( \partial_x u_h^2 h^{-1}, v_h \right) + \left( \hat{\nu} \partial_x u_h, \partial_x v_h \right) = 0, \quad \text{for all } v_h \in V_h \text{ and } t > 0, \quad (3.1)
\]

where we propose two different forms of \( \hat{\nu}: \)

1. first order artificial viscosity:

\[
\hat{\nu}(u_h) := h\frac{1}{2} \|u_h\|_{L^\infty(\mathbb{R})}; \quad (3.2)
\]

2. nonlinear weakly consistent artificial viscosity:

\[
\hat{\nu}(u_h) = \nu_0(u_h) \quad (3.3)
\]

with

\[
\nu_0(u_h)|_{K_i} := h\frac{1}{2} \|u_h\|_{L^\infty(\mathbb{R})} \max_{k \in \{i, i+1\}} |\phi_k| \quad (3.4)
\]

where

\[
\phi_k := \frac{[\partial_x u_h]_{x_k}}{2 \|[\partial_x u_h]\|_{x_k}}, \quad \text{(when } \|[\partial_x u_h]\|_{x_k} \neq 0, \text{ otherwise } \phi_k := 0).}
\]
where we have introduced the jump of the gradient over node $x_k$

$$[\partial_x u_h]_{x_k} := \partial_x u_h|_{K_k} - \partial_x u_h|_{K_{k-1}},$$

and the average of $|\partial_x u_h|$, $|\partial_x u_h|_{x_k} := \frac{1}{2}(|\partial_x u_h|_{K_k} + |\partial_x u_h|_{K_{k-1}})$, where $|\partial_x u_h|_{K_k} := |\partial_x u_h|_{x_k}$.

In this paper we only consider time discretization using the forward Euler method. However, since we show that the forward Euler scheme is total variation diminishing for either of the two artificial viscosities proposed above, we know that strong stability preserving Runge-Kutta methods will inherit this property and we will investigate the performance of a second order Runge-Kutta scheme in the numerical section. To define the fully discrete scheme we introduce the discrete time levels $0 = t_0 < t_1 < \ldots < t_N = T$, with time-step $k_n = t_n - t_{n-1}$, $n = 1, \ldots, N$. We denote the time intervals $I_j := (t_j, t_{j+1})$ and the space time slabs $Q_j := \mathbb{R} \times I_j$.

Using the above notation the artificial viscosity finite element space discretization and explicit Euler discretization in time of (2.1) reads, given $u_0 \in \text{Lip}^+(\mathbb{R})$ find $u_n^h \in V_h$ such that $u_h(0) = \pi_{BV} u_0$ and

$$k_n^{-1}(u_n^h - u_{n-1}^h), v_h) + \left(\frac{(u_{n-1}^h)^2}{2}, v_h\right) + \nu(u_{n-1}^h)\partial_x u_{n-1}^h, \partial_x v_h = 0,$$

for all $v_h \in V_h$ and $n = 1, \ldots, N$. \hfill (3.5)

For simplicity we will below assume that the time step is constant for all $n$. We end this section by defining $\pi_{BV}$ and proving some key properties of this interpolant.

**Definition 3.1.** (BV-stable interpolant, $\pi_{BV}$) Let

$$\pi_{BV} u := \sum_{i \in \mathbb{Z}} \bar{u}_i v_i \text{ where } \bar{u}_i = h^{-1} \int_{x_i - \frac{h}{2}}^{x_i + \frac{h}{2}} u \, dx.$$

The properties of the interpolant are collected in the following Lemma.

**Lemma 3.1.** Assume that $u \in BV$. Then there holds

$$\|\pi_{BV} u\|_{BV} \leq \|u\|_{BV} \text{ and } \|u - \pi_{BV} u\|_{L^1(\mathbb{R})} \lesssim h \|u\|_{BV}, \hfill (3.6)$$

$$(u - \pi_{BV} u, v) \lesssim h^2 \|u\|_{BV} \|v\|_{\text{Lip}}, \forall v \in \text{Lip and } \|\pi_{BV} u\|_{\text{Lip}^+} \leq \|u\|_{\text{Lip}^+}. \hfill (3.7)$$

**Proof.** For the proof of (3.6) see Ref. 1. The first inequality of (3.7) follows by first observing that since nodal quadrature integrates piecewise linear functions exactly and $\pi_{BV} u(x_i) = \bar{u}_i$ there holds for all elements $K$

$$\int_{K} (\pi_0 u - \pi_{BV} u) \, dx = 0,$$
where \( \pi_0 u \) is defined by \( \pi_0 u \big|_{(x_i - \frac{h}{2}, x_i + \frac{h}{2})} = \bar{u}_i \). Therefore

\[
(u - \pi_{BV} u, v) = (u - \pi_0 u, v - \pi_0 v) + \sum_i (\pi_0 u - \pi_{BV} u, v - v(x_i))_{K_i},
\]

\[
\lesssim h (\|u - \pi_0 u\|_{L^1(\mathbb{R})} + \|u - \pi_{BV} u\|_{L^1(\mathbb{R})}) \|v\|_{L^{ip}}.
\]

(3.8)

and the result follows from the \( L^1 \)-estimate (3.6). For the second inequality of (3.7) we see that

\[
\|\pi_{BV} u\|_{L^{ip}} = \sup_{x \in \mathbb{R}} h^{\frac{1}{2}} \left( \int_{x - h/2}^{x + h/2} u \, dx - \int_{x - h/2}^{x + h/2} \pi_{BV} u \, dx \right)
\]

\[
= \sup_{i} h^{-1} \int_{x_i - h/2}^{x_i + h/2} \frac{u(x + h) - u(x)}{h} \, dx 
\]

\[
\leq \sup_{i} h^{-1} \int_{x_i - h/2}^{x_i + h/2} \|u\|_{L^{ip}} \, dx = \|u\|_{L^{ip}}. \]

\[\square\]

4. Application of the theory of Nessyahu-Tadmor

We follow the abstract framework proposed in Section 4.2 of Ref. 32. We first recall the adjoint equation associated to the perturbation equation associated to (2.1) and (3.1)

\[
-\partial_t \varphi - a(u, u_h) \partial_x \varphi = 0 \quad \text{in} \quad Q
\]

\[
\varphi(\cdot, T) = \psi \quad \text{in} \quad \mathbb{R}
\]

(4.1)

and the following stability estimate.\(^{30}\)

**Proposition 4.1.** Consider the linear transport equation (4.1) with Lipschitz continuous final time data. We assume that for \( n \geq 0 \) the discrete solution \( u_h^n \)

1. satisfies \( U_n \lesssim U_0 \), with \( U_n := \|u_h^n\|_{L^\infty(\mathbb{R})} \);
2. satisfies the one sided Lipschitz condition

\[
D_n := \sup_{x \in \mathbb{R}} \|\partial_x u_h^n\| = \|u_h^n\|_{L^{ip}} \leq m(t_n), \quad m(t_n) \in L^1[0, t_n].
\]

(4.2)

Then for \( t > 0 \) there exists a unique Lipschitz continuous solution \( \varphi(x, t) \) such that the following estimate holds

\[
\|\varphi(\cdot, t)\|_{L^{ip}} \leq \|\psi(\cdot)\|_{L^{ip}} \cdot e^{M(t)}, \quad M(t) \equiv \int_0^t m(\tau) \, d\tau, \quad t > 0.
\]

Observe that since the \( Lip \)-seminorm is equivalent to the \( W^{1, \infty} \)-seminorm in one dimension, it follows from Proposition 4.1 and equation (4.1) that

\[
\|\partial_t \varphi(\cdot, t)\|_{L^\infty(\mathbb{R})} \lesssim U_0 \|\psi(\cdot)\|_{L^{ip}} \cdot e^{M(t)}, \quad t > 0.
\]

(4.3)

We can then prove the result,
Proposition 4.2. Let $u$ be the entropy solution of (2.1) and $u_h$ the solution of (3.1) with $0 \leq \nu_h \leq Ch U_n$. Assume that $u_h$ satisfies assumption 2. of Proposition 4.1 and in addition

\[ \|u_h(\cdot, t_n)\|_{BV} \lesssim \|u_h(\cdot, 0)\|_{BV}, \quad n \geq 1, \] (4.4)

where the constant $C$ is independent of $n$ and $h$.

Then there holds

\[ \|(u_h - u)(\cdot, T)\|_{L^p} \lesssim \eta(u_h) \] (4.5)

with

\[
\eta(u_h) := \|\pi_{BV} u_0 - u_0\|_{L^p} + h\|\partial_t u_h + \frac{1}{2} \partial_x (u_h)^2\|_{L^1(Q)} \\
+ \|\partial_x u_h\|_{L^1(Q)} + h^2\|\partial_x \partial_t u_h\|_{L^1(Q)} \\
+ k(\|\partial_t u_h + \frac{1}{2} \partial_x (u_h)^2\|_{L^1(Q)} + \|\partial_x u_h\|_{L^1(Q)}) \lesssim h + k, (4.6)
\]

where

\[
u_h(x, t)|_{I_j} : = u^j_h(x), \quad \partial_t u_h|_{I_j} : = k^{-1}(u^{j+1}_h(x) - u^j_h(x)).
\]

The error in $L^p$-norm satisfies the bound

\[ \|(u_h - u)(\cdot, T)\|_{L^p(R)} \lesssim \eta(u_h)^{\frac{1}{2p}} \lesssim (h + k)^{\frac{1}{2p}}. \] (4.7)

Proof. This result follows using the techniques from Theorem 4.1 and Corollary 4.1 of Ref. 32 adapted to the finite element framework. We give full details below for completeness.

First observe that by the definition of (4.1) there holds with $e = u_h - u$,

\[
(e(\cdot, T), \psi) = (u_h(\cdot, 0), \varphi(\cdot, 0)) + \sum_{i=0}^{N-1} (u^{i+1}_h - u^i_h, \varphi(\cdot, t_{i+1})) \\
- (u(\cdot, T), \psi) + (u, \partial_t \varphi)_Q - \left( \left( \frac{u^2_h}{2} - \frac{u^2}{2} \right), \partial_x \varphi \right)_Q \\
= (e(\cdot, 0), \varphi(\cdot, 0)) + (\partial_t u_h + \partial_x \left( \frac{u^2_h}{2} \right), \varphi)_Q + \sum_{i=0}^{N-1} (\partial_t u_h, \varphi(\cdot, t_{i+1}) - \varphi)_Q.
\]

Applying Galerkin orthogonality with the nodal interpolant of $\varphi(x, t)$, defined to be constant in time over each time interval $I_j$ as follows,

\[ I_h \varphi|_{I_j} := I_h \varphi(\cdot, t_j) \]
we have

\[
(e(\cdot), \psi) = (e(\cdot, 0), \varphi(\cdot, 0)) + (\partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right), \varphi - \mathcal{I}_h \varphi)_{Q} \\
- (\hat{\nu}(u_h) \partial_x u_h, \partial_x \mathcal{I}_h \varphi)_{Q} + \int_{Q} (\partial_t u_h \mathcal{I}_h \varphi - (\mathcal{I}_h (\partial_t u_h) \mathcal{I}_h \varphi)) \, dx \\
+ \sum_{i=0}^{N-1} (\partial_t u_h, \varphi(\cdot, t_{i+1}) - \varphi)_{Q_i}.
\]

For the last term in the right hand side observe that

\[
\sum_{i=0}^{N-1} (\partial_t u_h, \varphi(\cdot, t_{i+1}) - \varphi)_{Q_i} = \sum_{i=0}^{N-1} (\partial_t u_h, \int_{t}^{t_{i+1}} \partial_s \varphi(\cdot, s) \, ds)_{Q_i},
\]

\[
\leq k \|\partial_t u_h\|_{L^1(Q)} \|\partial_t \varphi\|_{L^{\infty}(Q)}
\]

using a similar approach in time and standard interpolation in space we also have the following bound on the interpolation error

\[
\|\varphi - \mathcal{I}_h \varphi\|_{L^{\infty}(Q)} = \sup_j \|\varphi - \mathcal{I}_h \varphi\|_{L^{\infty}(Q_j)}
\]

\[
\lesssim \sup_j \|\varphi - \mathcal{I}_h \mathcal{I}_h \varphi\|_{L^{\infty}(Q_j)} + \sup \|\varphi(\cdot, t_j) - \mathcal{I}_h \varphi\|_{L^\infty(\mathbb{R})}
\]

\[
\lesssim k \|\partial_t \varphi\|_{L^{\infty}(Q)} + h \sup_{t \in [0, T]} \|\varphi(\cdot, t)\|_{L^{\text{lip}}}.
\]

Using these upper bounds, Hölders inequality and \(\|\partial_x \mathcal{I}_h \varphi\|_{L^{\infty}(\mathbb{R})} \leq \|\varphi\|_{L^{\text{lip}}}\) we obtain

\[
(e(\cdot), \psi) \lesssim \left( \|\pi_{BV} u - u_0\|_{L^{p'}} + h \|\partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right)\|_{L^1(Q)} \right)
\]

\[
+ \|\hat{\nu}(u_h) \partial_x u_h\|_{L^1(Q)} + h^2 \|\partial_x \partial_t u_h\|_{L^1(Q)} \sup_{t \in [0, T]} \|\varphi(\cdot, t)\|_{L^{\text{lip}}}
\]

\[
+ k \left( \|\partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right)\|_{L^1(Q)} + \|\partial_t u_h\|_{L^1(Q)}\right) \|\partial_t \varphi\|_{L^{\infty}(Q)}.
\]

Note that a uniform upper bound on \(a(u, u_h)\) holds since \(U_n \leq BV(u_h^n)\). It then follows from Proposition 4.1 that by taking the supremum over all \(\psi\) such that \(\|\psi\|_{L^{\text{lip}}} = 1\)

\[
\|e(\cdot, T)\|_{L^{p'}} \lesssim e^{M(T)} \left[ \|\pi_{BV} u_0 - u_0\|_{L^{p'}} + h \|\partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right)\|_{L^1(Q)} \right]
\]

\[
+ \|\hat{\nu}(u_h) \partial_x u_h\|_{L^1(Q)} + h^2 \|\partial_x \partial_t u_h\|_{L^1(Q)}
\]

\[
+ k \left( \|\partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right)\|_{L^1(Q)} + \|\partial_t u_h\|_{L^1(Q)}\right).
\]
The error estimate (4.7) now follows from (4.5), (4.6) and the a priori control (4.4) where
\[ - (1 - \pi) \frac{1}{T} \int_0^T \| \partial_t u_h(x_i) \|^2_{L^1(\Omega)} \, dt \leq \left( \frac{1}{2} \sum_i \| \partial_t u_h(x_i) \|_{L^1(\Omega)}^2 \right) \leq U_0 \| \partial_x u_h \|_{L^1(\Omega)} \leq U_0 \| u_0 \|_{BV} \]
and therefore \( \| \partial_t u_h \|_{L^1(\Omega)} \leq U_0 T \| u_0 \|_{BV} \) and
\[ \| \partial_t u_h + \partial_x \left( \frac{u_h^2}{2} \right) \|_{L^1(Q)} \leq \left( \| \partial_t u_h \|_{L^1(Q)} + U_0 \| \partial_x u_h \|_{L^1(Q)} \right) \leq \| u_0 \|_{BV}. \]

Note that using an inverse inequality and the upper bound on \( \hat{\nu}(u_h) \) there holds
\[ h \| \partial_x \partial_t u_h \|_{L^1(Q)} + h^{-1} \| \hat{\nu} \partial_x u_h \|_{L^1(Q)} \leq \| \partial_t u_h \|_{L^1(Q)} + U_0 \| \partial_x u_h \|_{L^1(Q)} \leq T \| u_0 \|_{BV} \]
and (4.6) follows.

The \( L^p \)-error estimate finally is a consequence of a Gagliardo-Nirenberg inequality valid in one space dimension, \(^{27}\)
\[ \| \partial_x^j u \|_{L^p(\mathbb{R})} \leq \| \partial_x^m \|_{L^p(\mathbb{R})} \| u \|_{L^q(\mathbb{R})} \]
where
\[ \frac{1}{p} = j + \left( \frac{1}{r} - m \right) \alpha + \frac{1 - \alpha}{q} . \]
To apply (4.9) let \( E \) be a function with compact support such that \( \partial_x E = e - \bar{e} \), where \( e = \pi_{BV} u - u_h \). Then observe that
\[ \| E \|_{L^1(\mathbb{R})} = \sup_{\| v \|_{Lip}=1} \| E, \partial_x v \| = \sup_{\| v \|_{Lip}=1} \| e - \bar{e}, v \| = \| e \|_{Lip'} . \]
It follows that taking \( j = r = q = s = 1 \) and \( m = 2 \) in (4.9) we get, with \( 1 - \alpha = (2p)^{-1} \)
\[ \| e \|_{L^p(\mathbb{R})} = \| \partial_x E \|_{L^p(\Omega)} \leq \| \partial_x E \|_{L^1(\mathbb{R})} \| E \|_{L^1(\mathbb{R})}^{1-\alpha} \| e \|_{L^1(\mathbb{R})} \| e \|_{Lip'}^{2p} . \]
The estimate follows in the standard fashion by applying the above bounds and the corresponding \( L^p \)-interpolation error estimate\(^{1}\) for \( \pi_{BV} \) in the right hand side of
\[ \| u - u_h \|_{L^p(\mathbb{R})} \leq \| e \|_{L^p(\mathbb{R})} + \| \pi_{BV} u - u \|_{L^p(\mathbb{R})} . \]
The error estimate (4.7) now follows from (4.5), (4.6) and the a priori control (4.4) and the fact that \( u \) and \( u_h \) both are bounded in \( BV \). \( \square \)

It is possible to use interpolation to derive error estimates in stronger norms as well.\(^{20}\)
Remark 4.1. We observe that since the following a priori estimate is known for the exact solution

\[ \|u(\cdot, t)\|_{Lip^+} \leq \frac{1}{\|u_0\|_{Lip^+}^{-1} + t} \]

and assuming that a similar form will be satisfied for the discrete solution

\[ \|u_h(\cdot, t)\|_{Lip^+} \leq \frac{1}{\|u_0\|_{Lip^+}^{-1} + \sigma t} \quad 0 < \sigma \leq 1 \]

then the perturbation growth in time will be bounded by the factor

\[ e^{M(t)} = e^{\sigma^{-1} \log_e (1 + \|u_0\|_{Lip^+} \sigma t)} = (1 + \|u_0\|_{Lip^+} \sigma t)^{1/\sigma}. \]

Remark 4.2. It may be noted that assumption 2 of Proposition 4.1, implies that \( u_h \), with compact support, is bounded in the \( BV \)-norm since

\[ 0 = \int_{\mathbb{R}} \partial_x u^n_h \, dx = \int_{\mathbb{R}} (\partial_x u^n_h)_+ \, dx + \int_{\mathbb{R}} (\partial_x u^n_h)_- \, dx \]

and hence

\[ \int_{\mathbb{R}} (\partial_x u_h)_+ \, dx = - \int_{\mathbb{R}} (\partial_x u_h)_- \, dx. \]

Then

\[ \int_{\mathbb{R}} |\partial_x u^n_h| \, dx = 2 \int_{\mathbb{R}} (\partial_x u^n_h)_+ \, dx \leq 2 \text{meas}(\text{supp}(\partial_x u^n_h)) \sup_{x \in \mathbb{R}} \partial_x u^n_h(x) \]

\[ \leq 2 \text{meas}(\text{supp}(\partial_x u^n_h)) m(t). \]

However the constant in the above estimate depends on the measure of the support of \( u^n_h \) that can be difficult to quantify for artificial viscosity methods. In the following we will give independent proofs for the two properties.

5. Stability estimates for the forward Euler shock capturing scheme

It follows from the above analysis that we only need to prove that the numerical scheme satisfies the assumptions of Propositions 4.1 and 4.2. This amount to proving the \( Lip^+ \)-stability and the \( BV \)-stability of the discrete solution. The former result is the stronger one and as we shall see below, we can prove the latter result for cases where we do not manage to prove the former. One may prove that if the solution of (3.5) is \( BV \)-stable, then the sequence of finite element approximations converges to the unique entropy solution of Burgers’ equation.\(^2\)

We will first consider the first order scheme obtained using the viscosity (3.2) and prove that the solution is \( Lip^+ \)-stable and \( BV \)-stable. Then we will consider the scheme using the nonlinear viscosity (3.3). In this case we will first prove that the solution is \( BV \)-stable, extending the results of Ref. 2 to the fully discrete case. For the \( Lip^+ \)-stability on the other hand it is easy to show that the maximum derivative
can grow from one time-step to the next. Drawing from the analysis of the space semi-discretized case, where a modification of (3.4) was introduced ensuring the Lip$^+$-stability, we then propose a modified form of the artificial viscosity that is regularized locally in the neighbourhood of local maxima in the first derivative. This is a natural extension of the argument to the fully discrete case and for this perturbed shock-capturing term we prove the Lip$^+$-stability. In all cases we will evaluate the integrals of (3.5) and analyse the method as a finite difference scheme. In particular we will use the two forms given in the following Lemma. To avoid overloading the notation we sometimes drop the superscript $n - 1$ in the derivations below.

**Lemma 5.1.** Let $u_h$ be the solution of (3.5). Then the following relations hold

1. **Difference scheme for $u^n_h := u^n_h(x_i)$:**

$$u^n_i = u^{n-1}_i - h^{-1}k(\hat{\nu}(u^{n-1})|_{K_{i-1}} + \hat{u}_{i-} \partial_x u^{n-1}_h|_{K_{i-1}} + h^{-1}k(\hat{\nu}(u^{n-1})|_{K_{i}} - \hat{u}_{i+} \partial_x u^{n-1}_h|_{K_{i}}) \quad (5.1)$$

where

$$\hat{u}_{i-} := \int_{K_{i-1}} u^{n-1}_h v_i \, dx, \quad \hat{u}_{i+} := \int_{K_{i}} u^{n-1}_h v_i \, dx;$$

2. **Difference scheme for $\partial_x u^n_h|_{K_i}$:**

$$\partial_x u^n_h|_{K_i} = \partial_x u^{n-1}_h|_{K_i} + kh^{-2}(T_1(u^{n-1}_h) + T_2(u^{n-1}_h)), \quad (5.2)$$

where

$$T_1(u_h) := \frac{1}{6}h^2(\partial_x u_h|_{K_{i-1}})^2 - \frac{2}{3}h^2(\partial_x u_h|_{K_i})^2 - \frac{1}{6}h^2(\partial_x u_h|_{K_{i+1}})^2 - \frac{1}{2}hu_h(x_i)(\partial_x u_h|_{K_i} - \partial_x u_h|_{K_{i-1}}) - \frac{1}{2}hu_h(x_{i+1})(\partial_x u_h|_{K_{i+1}} - \partial_x u_h|_{K_i})$$

$$T_2(u_h) := (\hat{\nu}(u_h) \partial_x u_h)|_{K_{i-1}} - 2(\hat{\nu}(u_h) \partial_x u_h)|_{K_i} + (\hat{\nu}(u_h) \partial_x u_h)|_{K_{i+1}}.$$

**Proof.** Let $v_i$ denote the nodal basis function associated to node $x_i$. By testing (3.5) with $v_i$ we obtain by inspection

$$(k^{-1}(u^{n}_h - u^{n-1}_h), v_i)_h = k^{-1}(u^{n}_i - u^{n-1}_i)$$

$$\int_{\mathbb{R}} u^{n-1}_h \partial_x u^{n-1}_h v_i \, dx = \hat{u}_{i-} \partial_x u^{n-1}_h|_{K_{i-1}} + \hat{u}_{i+} \partial_x u^{n-1}_h|_{K_i} \quad (5.3)$$

where

$$\hat{u}_{i-} := \int_{K_{i-1}} u^{n-1}_h v_i \, dx, \quad \hat{u}_{i+} := \int_{K_{i}} u^{n-1}_h v_i \, dx.$$

For future reference we introduce the notation

$$u^n_i := u^n_h(x_i), \quad \hat{u}_i := \hat{u}_{i-} + \hat{u}_{i+}. \quad (5.4)$$
Applying (5.3) and (5.5) in the formulation (3.5) we may write,
\[ \int_{\mathbb{R}} \dot{\nu}(u_h^{n-1}) \partial_x u_h^{n-1} \partial_x v_i \, dx = (\dot{\nu}(u_h^{n-1})) \partial_x u_h^{n-1} |_{K_{i-1}} - (\dot{\nu}(u_h^{n-1})) \partial_x u_h^{n-1} |_{K_i}. \] (5.5)

Applying (5.3) and (5.5) in the formulation (3.5) we may write,
\[ u^n_i = u_i - h^{-1}k(\dot{\nu}(u_h)|_{K_{i-1}} + \hat{\nu}_i) \partial_x u_h|_{K_{i-1}} + h^{-1}k(\dot{\nu}(u_h)|_{K_i} - \hat{\nu}_i) \partial_x u_h|_{K_i}. \] (5.6)

To prove the second relation we may test (3.5) with \( h^{-1}(v_{i+1} - v_i) \) to obtain
\[ \partial_x u_h|_{K_i} = \partial_x u_h|_{K_i} - \frac{1}{h^2} \int_{x_{i-1}}^{x_{i+2}} u_h \partial_x u_h (v_{i+1} - v_i) \, dx \]
\[ - \frac{1}{h^2} \int_{x_{i-1}}^{x_{i+2}} \dot{\nu} \partial_x u_h \partial_x (v_{i+1} - v_i) \, dx = \partial_x u_h|_{K_i} + \frac{1}{h^2} (T_1(u_h) + T_2(u_h)). \]

Decomposing the integrals \( T_1 \) and \( T_2 \) on the contributions from \( v_i \) and \( v_{i+1} \) we have after integration
\[ T_1(u_h) = - \frac{1}{6} h^2 (\partial_x u_h|_{K_{i-1}})^2 - \frac{2}{3} h^2 (\partial_x u_h|_{K_i})^2 - \frac{1}{6} h^2 (\partial_x u_h|_{K_{i+1}})^2 \]
\[ - \frac{1}{2} hu_h(x_i)(\partial_x u_h|_{K_i} - \partial_x u_h|_{K_{i-1}}) - \frac{1}{2} hu_h(x_{i+1})(\partial_x u_h|_{K_{i+1}} - \partial_x u_h|_{K_i}) \]
and
\[ T_2 = - \int_{x_{i-1}}^{x_{i+2}} \dot{\nu} \partial_x u_h \partial_x (v_{i+1} - v_i) \, dx \]
\[ = (\dot{\nu}(u_h) \partial_x u_h)|_{K_{i-1}} - 2(\dot{\nu}(u_h) \partial_x u_h)|_{K_i} + (\dot{\nu}(u_h) \partial_x u_h)|_{K_{i+1}}. \]

To prove that the schemes are monotonicity preserving and total variation diminishing we will use Harten’s positivity criterion that we here recall on a form suitable for our purpose.

**Theorem 5.1.** *(Harten’s positivity criterion)* If the scheme (3.5) may be written on the form
\[ u^n_i(x_i) = u_h^{n-1}(x_i) - C_{i-1} h \partial_x u_h^{n-1}|_{K_{i-1}} + D_i h \partial_x u_h^{n-1}|_{K_i}, \] (5.7)
with the coefficients \( C_i \) and \( D_i \) satisfying
\[ C_{i-1} \geq 0, \quad D_i \geq 0 \text{ and } C_i + D_i \leq 1 \] (5.8)
then the scheme is total variation diminishing and satisfies \( \max_i u^n_i \leq \max_i u^0_i \).

**Proof.** For a proof we refer to Refs. 17, 18. \( \square \)
5.1. First order schemes

We will first consider the artificial viscosity method obtained using the definition (3.2) in (3.5).

Proposition 5.1. Let \( u_h \) be the solution of (3.1) using the linear viscosity (3.2). Then, if the CFL condition \( k \leq C_{oh}/U_{n-1} \), with \( C = \frac{1}{2} \) is satisfied, the following bounds hold for the discrete solution \( u_h^n \), \( n \geq 0 \),

\[
U_n \leq U_0, \quad BV(u_h^n) \leq BV(u_0)
\]

and

\[
\|u_h^n\|_{Lip^+} \leq \left[ \|u_0\|_{Lip^+} + \frac{1}{4} t_n \right].
\]

Proof. First we will apply Theorem 5.1, to prove equation (5.9). Compare equation (5.1) with (5.7) and identify

\[
C_{i-1} := h^{-2} k \left( \nu(u_h^{n-1})|_{K_{i-1}} + \hat{u}_{i-1} \right) \quad \text{and} \quad D_i := h^{-2} k \left( \nu(u_h^{n-1})|_{K_i} - \hat{u}_{i+1} \right).
\]

It follows from the definition of the viscosity, (3.2), that under the CFL condition \( k \leq \frac{1}{2} h/U_{n-1} \) there holds for all \( j \),

\[
0 \leq h^{-2} k \left( \frac{1}{2} hU_{n-1} + \hat{u}_{i-1} \right) = C_{i-1} \leq h^{-1} kU_{n-1} \leq \frac{1}{2},
\]

\[
0 \leq h^{-2} k \left( \frac{1}{2} hU_{n-1} - \hat{u}_{i+1} \right) = D_i \leq h^{-1} kU_{n-1} \leq \frac{1}{2}
\]

and we conclude that all inequalities of (5.8) are satisfied, proving (5.9). Turning to (5.10) this we consider equation (5.2) and note that since \( \nu(u_h) \) is constant on \( \mathbb{R} \) we have

\[
(T_1(u_h) + T_2(u_h)) = (\nu(u_h) + \frac{1}{2} h u_h(x_i)) (\partial_x u_h|_{K_{i-1}} - \partial_x u_h|_{K_i})
\]

\[
+ (\nu(u_h) - \frac{1}{2} h u_h(x_{i+1})) (\partial_x u_h|_{K_{i+1}} - \partial_x u_h|_{K_i})
\]

\[
- \frac{1}{6} h^2 (\partial_x u_h|_{K_{i-1}})^2 - \frac{2}{3} h^2 (\partial_x u_h|_{K_i})^2 - \frac{1}{6} h^2 (\partial_x u_h|_{K_{i+1}})^2
\]

It follows that (5.2) may be written

\[
\partial_x u_h^n|_{K_i} = \left( 1 - c_i - d_i \right) \partial_x u_h^{n-1}|_{K_i} + c_i \partial_x u_h^{n-1}|_{K_{i-1}} + d_i \partial_x u_h^{n-1}|_{K_{i+1}}
\]

\[
- \frac{1}{6} k (\partial_x u_h^{n-1}|_{K_{i-1}})^2 - \frac{2}{3} k (\partial_x u_h^{n-1}|_{K_i})^2 - \frac{1}{6} k (\partial_x u_h^{n-1}|_{K_{i+1}})^2
\]

with (by the definition of \( \nu \)) \( 0 \leq c_i = kh^{-2}(\nu(u_h) + \frac{1}{2} h u_h(x_i)) \) and \( 0 \leq d_i = kh^{-2}(\nu(u_h) - \frac{1}{2} h u_h(x_i)) \). Under the CFL-condition we get

\[
c_i + d_i = 2kh^{-2}\nu(u_h) \leq \frac{1}{2}.
\]
As a consequence
\[ D_n \leq (1 - \frac{k}{3}D_{n-1})D_{n-1}. \]

We now show that using induction on this expression one arrives at the bound
\[ D_n \leq \frac{D_0}{1 + D_0k/3}. \]

First observe that since \(1 - x \leq 1/(1 + x)\) there holds
\[ D_1 \leq (1 - \frac{k}{3}D_0)D_0 \leq \frac{D_0}{1 + D_0k/3}. \]

Then assume that \(D_n \leq D_0/(1 + D_0nk/3)\) and observe that
\[
D_{n+1} \leq (1 - \frac{k}{3}D_n)D_n \leq \frac{D_n}{1 + D_nk/3} \leq \frac{D_0}{(1 + D_0k/3)(1 + D_0nk/3)} \leq \frac{D_0}{(1 + D_0(n+1)k/3 + D_0^n k^2/9)} \leq \frac{D_0}{(1 + D_0(n+1)k/3)}. 
\]

This bound and the \(Lip^+\) stability of \(\pi_{BV}\), (3.7) proves (5.10).

5.2. The nonlinear shock-capturing method

In the nonlinear case \(Lip^+\)-stability was proved in the space semi-discretized case, provided the nonlinear viscosity was modified close to local maxima of the gradient.\(^3\)

The perturbed viscosity was defined by
\[
\hat{\nu}(u_h^{n-1})|_{K_i} := \nu_0(u_h^{n-1})|_{K_i} + \frac{1}{2}\xi(u_h)(\nu_0(u_h^{n-1})|_{K_{i-1}} + \nu_0(u_h^{n-1})|_{K_{i+1}}). \quad (5.11)
\]

where \(\xi\) denotes a correction factor, taking the value 1 in cells where the gradient takes a local maximum and zero elsewhere. It is not straightforward to make these ideas carry over to the fully discrete case. The reason for this is that the corrective factor \(\xi\) acts instantaneously to counter any growth in the positive gradient. In the fully discrete case the factor can only change at each time level so the \(Lip^+\)-stability can be violated at time level \(n\) in cells where the gradient is large, but not a local maximum, at time level \(n-1\). This problem arises where the nonlinear viscosity switches off abruptly and can be cured by smoothing the viscosity locally as we shall see below.

In this section we will first prove that the approximation obtained with the scheme using the unperturbed nonlinear viscosity satisfies the BV-estimate necessary for convergence and then we will extend the ideas of perturbing the viscosity close to local maxima to the fully discrete case and prove that we can obtain \(Lip^+\)-stability.
Proposition 5.2. (BV-bound for the nonlinear shock-capturing) Let $u_h^n$ be the solution of (3.1) computed under the CFL-condition $k \leq \frac{1}{4} hU_{n-1}^{-1}$ and using the nonlinear viscosity (3.3). Then the following bounds hold:

$$U_n \leq U_0, \quad BV(u_h^n) \leq BV(u_0). \tag{5.12}$$

Proof. To prove that the scheme is TVD we will use Harten’s positivity criterion. Starting from (5.1), assume that $\hat{u}_i \leq 0$ and add and subtract $\hat{u}_i - \partial_x u_h|_{K_i}$ in the right hand side of equation (5.1) to obtain

$$u_i^n = u_i - h^{-1}k((\hat{\nu}(u_h)|_{K_{i-1}} + \hat{u}_i - \partial_x u_h|_{K_i})$$

$$+ h^{-1}k((\hat{\nu}(u_h)|_{K_i} - \hat{u}_i + \partial_x u_h|_{K_i})$$

$$= u_i - h^{-1}k((\hat{\nu}(u_h)|_{K_{i-1}} - \hat{u}_i - \partial_x u_h|_{K_i})$$

$$+ h^{-1}k((\hat{\nu}(u_h)|_{K_i} - \hat{u}_i - \partial_x u_h|_{K_i}).$$

Using now that

$$\hat{u}_i - [\partial_x u_h]_{x_i} = 2\hat{u}_i - \phi_i(u_h)\{[\partial_x u_h]_{x_i} = \hat{u}_i - \phi_i(u_h)\{[\partial_x u_h|_{K_i} + [\partial_x u_h|_{K_{i-1}}\}$$

we may write, with the notation $s_i := \text{sign}(\partial_x u_h|_{K_i})$

$$u_i^n = u_i - h^{-1}k((\hat{\nu}(u_h)|_{K_{i-1}} - \hat{u}_i - \phi_i(u_h)s_i)\partial_x u_h|_{K_{i-1}}$$

$$+ h^{-1}k((\hat{\nu}(u_h)|_{K_i} - \hat{u}_i - \phi_i(u_h)s_i)\partial_x u_h|_{K_i}.$$  

We then identify the coefficients

$$C_{i-1} = h^{-2}k((\hat{\nu}(u_h)|_{K_{i-1}} - \hat{u}_i - \phi_i s_{i-1})$$

and

$$D_i = h^{-2}k((\hat{\nu}(u_h)|_{K_i} - \hat{u}_i - \phi_i s_i).$$

By the definition of $\hat{\nu}(u_h)$ we see that

$$(\hat{\nu}(u_h)|_{K_{i-1}} - \hat{u}_i - \phi_i(u_h)s_i) \geq \frac{1}{2}(U_{n-1} - \|u_h\|_{L^\infty(K_{i-1})}) \max_{k \in \{j-1,j\}} |\phi_k| h \geq 0 \tag{5.13}$$

and

$$(\hat{\nu}(u_h)|_{K_i} + \hat{u}_i - \phi_i(u_h)s_i) \geq \frac{1}{2}(U_{n-1} - \|u_h\|_{L^\infty(K_{i-1})}) \max_{k \in \{j,j+1\}} |\phi_k| h \geq 0 \tag{5.14}$$

It then follows that $C_{i-1} \geq 0$ and, since we assumed that $\hat{u}_i < 0$, $D_i \geq 0$. To see that the condition $C_i + D_i \leq 1$ may be satisfied we write

$$C_i + D_i \leq h^{-2}k(2\hat{\nu}(u_h)|_{K_i} + 3h\|u_h\|_{L^\infty(K_{i-1} \cup K_i)}) \leq 4kh^{-1}U_{n-1}. \tag{5.15}$$

Hence the desired bound holds under the CFL-condition

$$k \leq \frac{1}{4} hU_{n-1}^{-1}.$$  

The case when $\hat{u}_i \geq 0$ is similar, but this time $\hat{u}_i + \partial_x u_h|_{K_{i-1}}$ is added and subtracted instead.
The difficulty in proving the $\text{Lip}^+$-stability for the nonlinear shock-capturing stems from the strong variations in the viscosity coefficient that may appear if the solution abruptly changes from rough to smooth. These fluctuations may translate into oscillations of the gradient that violates the $\text{Lip}^+$-condition. The nonlinear indicator function $\xi$ that prohibited growth of gradient does not work in the fully discrete case due to its instantaneous effect. Only under a strengthened CFL does the correction eliminate spurious gradient oscillations. This effect will be explained in the analysis below and illustrated in the numerical section. To obtain a fully discrete, $\text{Lip}^+$-stable, scheme we propose a method that uses a regularized version of the nonlinear viscosity. As a substitute for the instantaneous action of the nonlinear perturbation $\xi(u)$ we categorise the elements in one part where the gradient is so big that time discretization can result in a violation of the $\text{Lip}^+$ stability over one time step and another part where a moderate growth in the gradient can be permitted. The solution in the part where a violation is possible is then computed using a regularized viscosity coefficient that is a natural generalization of $\xi$ to the fully discrete case.

Recall that $D_n := \sup_i \partial_x u^n_h|_{K_i}$ and fix $0 < \delta < 1$. Then define

$$K_{\xi} := \{ K_i : \partial_x u^n_h|_{K_i} / D_n > \delta \}$$

and $K_{\xi} := \{ K_i \notin K_{\xi} = 1 \}$. We propose the following generalization of the definition of the viscosity $\hat{\nu}$.

**Definition 5.1.** Let $\hat{\nu}(u_h)|_{K_i}$ be defined depending on the solution $u_h$ in the neighborhood of element $K_i$ according to:

- $K_i \in K_{\xi}=0$. Then $\hat{\nu}(u_h)|_{K_i} = \nu_0(u_h)|_{K_i}$
- $K_i \in K_{\xi}=1$. For each connected subset of elements $K \in K_{\xi}=1$ let $K_1, \ldots, K_M$, $M \geq 1$, denote the elements in $K$. Let $K_0$ and $K_{M+1}$ denote the left and right neighbours of the interval $\cup_{i=1}^M K_i$. Define the locally regularized viscosity by $\nu_S(u_h)|_{K_i}$, $i = 1, \ldots, M$ such that

$$-\nu_S|_{K_{i-1}} + 2\nu_S|_{K_i} - \nu_S|_{K_{i+1}} = 2\nu_0|_{K_i}$$

(5.17)

with the boundary conditions $\nu_S|_{K_0} = \nu_0|_{K_0}$, $\nu_S|_{K_{M+1}} = \nu_0|_{K_{M+1}}$. Then define:

$$\hat{\nu}(u_h)|_{K_i} := \min \left\{ 3/2 U_n-1 h, \nu_S|_{K_i} \right\}.$$

First observe that the condition in point two leads to an implicit definition of the viscosity, a linear system has to be solved for every connected interval in $K_{\xi}=1$. Since the left hand side of (5.17) is an M-matrix and the right hand side is positive, there holds $\hat{\nu} \geq 0$. Also note that the method is still formally of second order away from local extrema. However, point 2 of Definition 5.1 above also shows that in certain situations the first order artificial viscosity can extend into the smooth part of the domain, possibly making the order of convergence degenerate. The actual
performance of this method will depend on the choice of $\delta$ and of the distribution of the elements in $K_{\xi=1}$.

**Remark 5.1.** To see the relation to the semi-discretized case, consider an isolated element $K_i$ in $K_{\xi=1}$, and such that $\hat{\nu}(u^h_{n-1})|_{K_i} < \frac{3}{2}U_{n-1}h$. In this case there holds

$$\hat{\nu}_{K_i} = \nu_0|_{K_i} + \frac{1}{2}(\nu_0|_{K_{i-1}} + \nu_0|_{K_{i+1}}). \quad (5.18)$$

We see that (5.18) coincides with the definition of (5.11), with $\xi|_{K_i} = 1$.

**Proposition 5.3.** Let $u_h$ be the solution of (3.5) computed under the CFL-condition $k \leq \frac{1}{6}hU_{n-1}^{-1}$ and with $\hat{\nu}$ from Definition 5.1 above. Then there holds

$$U_n \leq U_0, \quad BV(u^h_n) \leq BV(u_0). \quad (5.19)$$

**Proof.** The proof is equivalent to that of Proposition 5.2. Using that this time $\nu_0|_{K_i} \leq \hat{\nu}|_{K_i} \leq \frac{3}{2}U_{n-1}h$, showing that $\hat{\nu}$ is sufficiently large for (5.13) and (5.14) to hold. Condition (5.15) then holds under the strengthened CFL-condition. \(\square\)

**Theorem 5.2.** Let $u_h^n$ be the solution of (3.5) with $\hat{\nu}$ given in Definition 5.1 and computed under the CFL-condition

$$k_n \leq Co hU_{n-1}^{-1}, \quad Co = \frac{1}{4}(1 - \delta)$$

where $\delta$ is the parameter introduced in (5.16). Then there holds

$$\|u^h_n\|_{Lip^*} \leq \frac{1}{\|u_0\|_{Lip^*}} + \frac{2\epsilon}{\delta}.$$ 

**Proof.** Consider an element $K_i$, we must consider the two possibilities $K_i \in K_{\xi=0}$ and $K_i \in K_{\xi=1}$ separately. We evaluate the right hand side of (5.2) in each case. Let $s_i := \text{sign} (\partial_x u^h_{n-1}|_{K_i})$.

- $K_i \in K_{\xi=0}$. First consider the case $s_i < 0$. Then at least one of $u_h(x_i) \geq 0$ and $u_h(x_{i+1}) \leq 0$ must be true. For simplicity assume that the latter case holds. Starting from (5.2) we may write,

$$\partial_x u^h_n|_{K_i} = -\frac{1}{6}k(\partial_x u_h|_{K_{i-1}})^2 + \frac{2}{3}k(\partial_x u_h|_{K_i})^2 - \frac{1}{6}k(\partial_x u_h|_{K_{i+1}})^2 + \left(1 - \frac{1}{2}kh^{-1}u_h(x_i) + \frac{1}{2}kh^{-1}u_h(x_{i+1}) - 2kh^{-2}\hat{\nu}|_{K_i}\right)\partial_x u_h|_{K_i}$$

$$+ kh^{-2}\left(\hat{\nu}|_{K_{i-1}} + \frac{1}{2}hu_h(x_i)\right)\partial_x u_h|_{K_{i-1}}$$

$$+ kh^{-2}\left(\hat{\nu}|_{K_{i+1}} - \frac{1}{2}hu_h(x_{i+1})\right)\partial_x u_h|_{K_{i+1}}.$$
We now use that
\[
\frac{1}{2} h u_h(x_i) \partial x u_h|_{K_i=1} = -\frac{1}{2} h u_h(x_i) \phi_i s_{i-1} \partial x u_h|_{K_i=1}
+ \frac{1}{2} h u_h(x_i)(1 - s_i \phi_i) \partial x u_h|_{K_i} \quad (5.20)
\]
to obtain
\[
\partial x u_h^n|_{K_i} = -\frac{1}{6} k(\partial x u_h|_{K_i-1})^2 - \frac{2}{3} k(\partial x u_h|_{K_i})^2 - \frac{1}{6} k(\partial x u_h|_{K_{i+1}})^2
+ \left(1 - \frac{1}{2} k h^{-1}(u_h(x_i)(2 - s_i \phi_i) + u_h(x_{i+1})) - 2 k h^{-2} \hat{\nu}|_{K_i}\right) \partial x u_h^{n-1}|_{K_i}
\]
\[
+ kh^{-2} \left(\hat{\nu}|_{K_{i-1}} - \frac{1}{2} h u_h(x_i) \phi_i s_{i-1}\right) \partial x u_h^{n-1}|_{K_{i-1}}
\]
\[
+ kh^{-2} \left(\hat{\nu}|_{K_{i+1}} - \frac{1}{2} h u_h(x_{i+1})\right) \partial x u_h^{n-1}|_{K_{i+1}}.
\]

It follows from the definition of \( \hat{\nu} \) and \( Co \) that \( b_i \geq 1 - \frac{5}{2} Co \geq 0 \), \( 0 \leq c_i \leq 2 Co \) and \( 0 \leq d_i \leq 2 Co \). We assume the worst case \( \partial x u_h^{n-1}|_{K_{i-1}} \geq 0 \) and \( \partial x u_h^{n-1}|_{K_{i+1}} \geq 0 \) to obtain the upper bound
\[
\partial x u_h^n|_{K_i} \leq -\frac{1}{6} k(\partial x u_h^{n-1}|_{K_{i-1}})^2 - \frac{1}{6} k(\partial x u_h^{n-1}|_{K_{i+1}})^2
+ 2 Co(\partial x u_h^{n-1}|_{K_{i-1}} + \partial x u_h^{n-1}|_{K_{i+1}}) \leq D_{n-1} - \frac{1}{3} k D_{n-1}^2,
\]
where we have used the inequality (??).
Now we turn to the case \( s_i \geq 0 \). First assume that \( \hat{\nu}|_{K_{i-1}} < \frac{1}{2} u_{n-1} h \) and \( \hat{\nu}|_{K_{i+1}} < \frac{1}{2} u_{n-1} h \). We may then use the relation
\[
\partial x u_h|_{K_{i-1}} - \partial x u_h|_{K_i} = -\phi_i s_{i-1}(\partial x u_h|_{K_{i-1}} - \partial x u_h|_{K_i}) - \phi_i (s_{i-1} + s_i) \partial x u_h|_{K_i} \quad (5.21)
\]
and similarly for $\partial_x u_h|_{K_{i+1}}$, in (5.2) to obtain

$$
kh^{-2}(T_1 + T_2) = -\frac{1}{6} k(\partial_x u_h|_{K_{i-1}})^2 - \frac{2}{3} k(\partial_x u_h|_{K_i})^2 - \frac{1}{6} k(\partial_x u_h|_{K_{i+1}})^2 \\
+ kh^{-2}(\hat{\nu}(u_h)|_{K_{i-1}} - \frac{1}{2} u_h(x_i)h\phi_{i-1})(\partial_x u_h|_{K_{i-1}} - \partial_x u_h|_{K_i}) \\
+ kh^{-2}(\hat{\nu}(u_h)|_{K_{i+1}} - \frac{1}{2} u_h(x_{i+1})h\phi_{i+1})(\partial_x u_h|_{K_{i+1}} - \partial_x u_h|_{K_i}) \\
+ kh^{-2}(\hat{\nu}(u_h)|_{K_{i-1}} - 2\hat{\nu}(u_h)|_{K_i} + \hat{\nu}(u_h)|_{K_{i+1}})\partial_x u_h|_{K_i} \\
- \frac{1}{2} u_h(x_i)kh^{-1}(s_{i-1} + s_i)\phi_i\partial_x u_h|_{K_i} \\
- \frac{1}{2} u_h(x_{i+1})kh^{-1}(s_{i+1} + s_i)\phi_{i+1}\partial_x u_h|_{K_i}.
$$

(5.22)

If either $\hat{\nu}(u_h)|_{K_{i-1}}$ or $\hat{\nu}(u_h)|_{K_{i+1}}$ is larger than $\frac{1}{2} hU_{n-1}$ then it is not necessary to introduce the $\phi$ factor in the corresponding factor $c_i$ or $d_i$. By the assumption on $\hat{\nu}$ and under the CFL condition we have $0 \leq c_i \leq 2Co$ and $0 \leq d_i \leq 2Co$, with the maximum values taken if $\hat{\nu}|_{K_{i-1}}$ or $\hat{\nu}|_{K_{i+1}}$ takes the maximum value $\frac{3}{2}U_{n-1}h$.

We proceed to bound the last three terms of (5.22). Under the CFL-condition and assuming $\hat{\nu}|_{K_{i-1}} \leq \frac{1}{2} hU_{n-1}$ and $\hat{\nu}|_{K_{i+1}} \leq \frac{1}{2} hU_{n-1}$ we obtain the bound

$$
kh^{-2}(\hat{\nu}(u_h)|_{K_{i-1}} - 2\hat{\nu}(u_h)|_{K_i} + \hat{\nu}(u_h)|_{K_{i+1}})\partial_x u_h|_{K_i} \\
- \frac{1}{2} u_h(x_i)kh^{-1}(s_{i-1} + s_i)\phi_i\partial_x u_h|_{K_i} \\
- \frac{1}{2} u_h(x_{i+1})kh^{-1}(s_{i+1} + s_i)\phi_{i+1}\partial_x u_h|_{K_i} \leq 3Co \partial_x u_h|_{K_i}.
$$

It is easy to see that the same inequality holds if one or both of $\hat{\nu}(u_h)|_{K_{i-1}}$ and $\hat{\nu}(u_h)|_{K_{i+1}}$ are larger than $\frac{1}{2} hU_{n-1}$, since then the corresponding terms including the $\phi$ factors are omitted. It follows that

$$
\partial_x u_h^n|_{K_i} \leq -\frac{1}{6} k(\partial_x u_h^{n-1}|_{K_{i-1}})^2 - \frac{2}{3} k(\partial_x u_h^{n-1}|_{K_i})^2 - \frac{1}{6} k(\partial_x u_h^{n-1}|_{K_{i+1}})^2 \\
+ (1 - c_i - d_i + 3Co)\partial_x u_h^{n-1}|_{K_i} + c_i\partial_x u_h^{n-1}|_{K_{i-1}} + d_i\partial_x u_h^{n-1}|_{K_{i+1}} \\
\leq \partial_x u_h^{n-1}|_{K_i} + 4Co D_{n-1} - \frac{1}{3} k D_{n-1}^2
$$

where we once again used the inequality (??). Using now the definition of $Co$ and of $K_{\xi=0}$ we have

$$
\partial_x u_h^n|_{K_i} \leq \delta D_{n-1} + (1 - \delta) D_{n-1} - \frac{1}{3} k D_{n-1}^2 \leq D_{n-1} - \frac{1}{3} k D_{n-1}^2.
$$

• $K_i \in K_{\xi=1}$. Observe that in this case by definition $\partial_x u_h|_{K_i} > 0$. First assume that $\hat{\nu}(u_h)|_{K_i} < \frac{1}{2} hU_{n-1}$ and $0 \leq u_h(x_i) \leq u_h(x_{i+1})$ then we consider again
equation (5.2), and use (5.20) in the second term of $T_1$ to obtain

$$
kh^{-2}(T_1 + T_2) = -\frac{k}{6} \frac{1}{2} (\partial_x u_h |_{K_{i-1}})^2 - \frac{2}{3} k (\partial_x u_h |_{K_i})^2 - \frac{1}{6} k (\partial_x u_h |_{K_{i+1}})^2
+ \underbrace{kh^{-2} (\tilde{\nu}(u_h) |_{K_{i-1}} + \frac{1}{2} u_h(x_i)) (\partial_x u_h |_{K_{i-1}} - \partial_x u_h |_{K_i})}_{\kappa_i}
+ \underbrace{kh^{-2} (\tilde{\nu}(u_h) |_{K_{i-1}} - \frac{1}{2} u_h(x_{i+1}) h \phi_{i+1} |_{s_{i+1}}) (\partial_x u_h |_{K_{i+1}} - \partial_x u_h |_{K_i})}_{\tilde{\nu}}
+ \underbrace{kh^{-2} (\tilde{\nu}(u_h) |_{K_{i-1}} - 2 \tilde{\nu}(u_h) |_{K_i} + \tilde{\nu}(u_h) |_{K_{i+1}}) \partial_x u_h |_{K_i}}_{\partial_x u_h |_{K_{i+1}}}
- \frac{1}{2} u_h(x_{i+1}) kh^{-1} (s_{i+1} + s_i) \phi_{i+1} \partial_x u_h |_{K_i},
$$

(5.23)

This time $0 \leq c_i \leq 2 Co$ by the sign of $u_h(x_i)$ and $0 \leq d_i \leq 2 Co$ and $c_i + d_i \leq 4 Co < 1$ as in the previous case. Note that by equation (5.17) we have

$$
\tilde{\nu}(u_h) |_{K_{i-1}} - 2 \tilde{\nu}(u_h) |_{K_i} + \tilde{\nu}(u_h) |_{K_{i+1}} - \frac{1}{2} u_h(x_i) kh(s_{i+1} + s_i) \phi_{i+1}
\leq \nu S |_{K_{i-1}} - 2 \nu S |_{K_i} + \nu S |_{K_{i+1}} + u_h(x_i) kh |_{\phi_{i+1}}
= -2 \nu v_0 + u_h(x_i) kh |_{\phi_{i+1}} \leq 0.
$$

Therefore we may this time conclude that

$$
\partial_x u_h |_{K_i} \leq (1 - c_i - d_i) \partial_x u_h |_{K_{i-1}} + c_i \partial_x u_h |_{K_{i-1}} + d_i \partial_x u_h |_{K_{i+1}}
- \frac{1}{6} k (\partial_x u_h |_{K_{i-1}})^2 - \frac{2}{3} k (\partial_x u_h |_{K_i})^2 - \frac{1}{6} k (\partial_x u_h |_{K_{i+1}})^2
\leq D_{n-1} - \left( \frac{2}{3} \nu^2 + \frac{1}{3} \right) k D_{n-1}^2.
$$

The case $u_h(x_i) \leq u_h(x_{i+1}) \leq 0$ is equivalent, but this time the $c_i$ factor will have the $\phi_i$ perturbation and $d_i$ will be positive using only the sign of $u_h(x_{i+1})$.

Finally we need to consider the configuration where there is a sonic point in $K_i$, i.e. $u_h(x_i) \leq 0$ and $u_h(x_{i+1}) \geq 0$. Assume that $\partial_x u_h |_{K_{i-1}} < 0$ then there is a local minimum in $x_i$ and therefore $\tilde{\nu} |_{K_{i-1}} \geq \frac{1}{2} U_{n-1} h$. This means that $c_i$ is positive and the result follows as before. If on the other hand $\partial_x u_h |_{K_{i-1}} > 0$ we use that

$$
kh^{-2} (\tilde{\nu}(u_h) |_{K_{i-1}} + \frac{1}{2} u_h(x_i)) (\partial_x u_h |_{K_{i-1}} - \partial_x u_h |_{K_i})
\leq kh^{-2} (\tilde{\nu}(u_h) |_{K_{i-1}} + \frac{1}{2} u_h(x_{i+1})) (\partial_x u_h |_{K_{i-1}} - \partial_x u_h |_{K_i})
+ \frac{1}{2} k (\partial_x u_h |_{K_i})^2 - \frac{1}{2} k \partial_x u_h |_{K_i} \partial_x u_h |_{K_{i-1}}.
$$

In the first line of the right hand side we see that we have reverted back to the situation of a positive convective velocity so this part can be treated as before.
The first term in the last line can be absorbed by the second term of $T_1$ and the second term is negative so it may be dropped. We conclude that

$$\partial_x u^n_h|_{K_i} \leq (1 - \frac{1}{3}kD_{n-1})D_{n-1}.$$  

Now consider the case $\tilde{\nu}|_{K_i} = \frac{2}{3}U_{n-1}h$, then we must consider three different cases depending on the neighbouring elements.

1. $K_{i-1}, K_{i+1} \in K_\ell = 1$. Then observe that by the definition of $(5.17)$ there holds $\frac{2}{3}U_{n-1}h \leq \tilde{\nu}|_{K_{i+1}} \leq \frac{2}{3}U_{n-1}h$. We start from (5.2) and write

$$\begin{align*}
kh^{-2}(T_1 + T_2) &= -\frac{1}{6}k(\partial_x u_h|_{K_{i+1}})^2 - \frac{2}{3}k(\partial_x u_h|_{K_i})^2 - \frac{1}{6}k(\partial_x u_h|_{K_{i-1}})^2 \\
&\quad + kh^{-2}(\tilde{\nu}(u_h)|_{K_{i+1}} + \frac{1}{2}u_h(x_i)h)(\partial_x u_h|_{K_{i-1}} - \partial_x u_h|_{K_i}) \\
&\quad + kh^{-2}(\tilde{\nu}(u_h)|_{K_{i+1}} - \frac{1}{2}u_h(x_{i+1})h)(\partial_x u_h|_{K_{i+1}} - \partial_x u_h|_{K_i}) \\
&\quad + \bigg(\frac{1}{2}u_h(x_{i+1})h(s_{i+1} + s_i)\phi_{i+1}\bigg) \\
&\quad - U_{n-1}h + |u_h(x_i)||\phi_{i+1}| \leq 0.
\end{align*}$$

As before $0 \leq c_i, d_i \leq 2Co, c_i + d_i < 1$. Thanks to the cut off at $3/2U_{n-1}h$ we see that $\tilde{\nu}(u_h)|_{K_{i+1}} - \frac{1}{2}u_h(x_{i+1})h(s_{i+1} + s_i)\phi_{i+1} \leq 0$.

2. $K_{i-1} \in K_\ell = 1, K_{i+1} \in K_\ell = 0$. In this case note that $\frac{2}{3}U_{n-1}h \leq \tilde{\nu}|_{K_{i-1}} \leq \frac{2}{3}U_{n-1}h$. Then considering equation (5.23) we observe once again $0 \leq c_i \leq 2Co$ (using the bounds on $\tilde{\nu}(u_h)|_{K_{i-1}}$, $0 \leq d_i \leq 2Co$ (using the definition of $\tilde{\nu}(u_h)|_{K_{i+1}}$, $c_i + d_i < 1$ and we may conclude by observing that by the bounds on the viscosity $\tilde{\nu}$ in the elements we have

$$\begin{align*}
\tilde{\nu}(u_h)|_{K_{i-1}} - 2\tilde{\nu}(u_h)|_{K_i} + \tilde{\nu}(u_h)|_{K_{i+1}} - \frac{1}{2}u_h(x_i)h(s_{i+1} + s_i)\phi_{i+1} \\
&\leq -U_{n-1}h + |u_h(x_i)||\phi_{i+1}| \leq 0. \ (5.24)
\end{align*}$$

(3) $K_{i-1} \in K_\ell = 0, K_{i+1} \in K_\ell = 1$. We proceed as in point 2), but with the $\phi_i$ contribution in the factor $c_i$. Then $0 \leq c_i \leq Co$ and since $\frac{2}{3}U_{n-1}h \leq \tilde{\nu}|_{K_{i+1}} \leq \frac{2}{3}U_{n-1}h$ we also have $0 \leq d_i \leq 2Co$ and under the condition on $Co, c_i + d_i < 1$. Finally we observe as in 2) that $\tilde{\nu}(u_h)|_{K_{i-1}} - 2\tilde{\nu}(u_h)|_{K_i} + \tilde{\nu}(u_h)|_{K_{i+1}} \leq -U_{n-1}h$ and therefore the equivalent of (5.24) holds.

Using the same arguments as in the proof of (5.10) we have

$$D_n \leq \left(1 - \frac{k}{3}D_{n-1}\right)D_{n-1}$$

in all three cases.

The result once again follows by induction over $D_n$. \[\square\]
**Remark 5.2.** It should be noted that several aspects of the above regularization could be modified in order to minimize \( \hat{\nu} \). For instance the right hand side (5.17) could be modified, making it zero if the last term in (5.23) is negative. The present form was chosen in order to convey the idea with all the desireable results, without having to consider too many special cases.

**Remark 5.3.** The order of the method depends on the interaction between the space and time discretization. To obtain a first order scheme we may use a fixed \( \delta \) and take \( k_n = \frac{1}{4} U_{n-1} (1 - \delta) \). Since \( \delta \) does not go to zero with \( h \), the artificial viscosity may remain \( O(h) \) in a subset of \( \mathbb{R} \) that does not go to zero when reducing the mesh size. A higher order scheme is obtained by choosing \( \delta = (1 - h^s), s > 0 \). Then the measure of the set \( K_{\xi=1} \) will go to zero and the CFL condition will be strengthened to \( k_n = \frac{1}{4} U_{n-1} h^{1+s} \), making the time discretization error higher order as well.

6. Numerical examples

In this section we will illustrate the practical implications of the above analysis. First we consider two model problems, one with smooth initial data and one with rough data and verify that the nonlinear shock capturing method has global second order convergence for the smooth solution and first order convergence for the rough solution. We then turn to the question of the \( \text{Lip}^+ \)-stability of the methods and give some examples of violations of \( \text{Lip}^+ \)-stability and also show how to improve the behavior. In the study below we will consider three different methods:

1. the forward Euler method with the artificial viscosity (3.3) ("method I" below);
2. the forward Euler method with the artificial viscosity (3.3) and the local correction (5.11) acting only in cells with local maxima of the gradient ("method II" below);
3. an SSP second order Runge-Kutta method, i.e. Heun’s method ("method III" below).

6.1. Convergence studies

We first consider a problem with smooth initial data

\[
 u_0 = \frac{1}{2} (\cos(\pi x) + 1)
\]

on the interval \((-1, 1)\). We compute the solution at \( T = 0.5 \), before shock formation and solve for the exact solution on a mesh with 6400 mesh points using fixed point iteration. The initial data and the final solutions are given in Figure 1. First we consider time discretization using method I and the following relation between the time and space discretization parameters, \( k = h^2 \). In Table 1 errors in several different norms are presented on four consecutive meshes. Instead of reporting the
Error estimates for SC-FEM for Burgers’ equation

Fig. 1. Left smooth initial condition; right solution at $T = 0.5$.

Table 1. smooth solution, forward Euler in time, $k = h^2$

<table>
<thead>
<tr>
<th>N</th>
<th>$|u - u_h|_{L^1}\text{(R)}$</th>
<th>$|u - u_h|_{L^2}\text{(R)}$</th>
<th>$|u - u_h|_{-1}$ ($\sim H^{-1}\text{(R)}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>$2.5 \cdot 10^{-3}$</td>
<td>$3.6 \cdot 10^{-3}$</td>
<td>$3.0 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>200</td>
<td>$6.7 \cdot 10^{-4}$</td>
<td>$1.0 \cdot 10^{-3}$</td>
<td>$7.0 \cdot 10^{-5}$</td>
</tr>
<tr>
<td>400</td>
<td>$1.8 \cdot 10^{-4}$</td>
<td>$3.0 \cdot 10^{-4}$</td>
<td>$1.7 \cdot 10^{-5}$</td>
</tr>
<tr>
<td>800</td>
<td>$4.6 \cdot 10^{-5}$</td>
<td>$8.9 \cdot 10^{-5}$</td>
<td>$4.2 \cdot 10^{-6}$</td>
</tr>
</tbody>
</table>

$\text{Lip}^+$-norm, we give the the errors in the following weak norm related to $H^{-1}$,

$$\|u\|_{-1}^2 := \|\nabla \bar{u}\|^2 + \|\bar{u}\|^2,$$

where $\bar{u}$ solves $-\partial_{xx} \bar{u} + \bar{u} = u$, $\bar{u}(-1) = \bar{u}(1) = 0$.

Experimental convergence rates are given in parenthesis. We observe second order convergence in the $L_1$-norm and in the $\|\cdot\|_{-1}$ norm. The convergence in the $L_2$-norm is slightly below second order, but should be compared with the $O(h^{3/2})$ convergence that is expected for stabilized finite element methods.

Now we consider a problem with non-smooth solution. The initial data and final time exact solution is given in Figure 2. We compute the solution at $T = 0.5$ when the shock has formed. The exact solution is computed using the method of characteristics on a mesh with 12800 elements. We present tables with the same errors as in the previous cases in Table 2. We observe first order convergence for the $L^1$-error and the $H^{-1}$-norm error and 1/2-order convergence in the $L^2$-norm.

We then consider the same computations using method III, under the CFL-condition $k = \frac{1}{4} h$. The corresponding results are presented in Tables 3 and 4, the conclusions are similar as in the previous case. Not in any case were any violations of the maximum principle observed.

6.2. Investigation of $\text{Lip}^+$ stability

We now consider the initial data given by the left plot of figure 2 and solve until $T = 1$. First we consider the forward Euler methods under the CFL-condition $k = \frac{1}{4} h$, which is the limit value of Proposition (5.12). We study the relative violation of the $\text{Lip}^+$-stability by reporting the time evolution of $\lambda_n = D_n/\sup_x \partial_x u(x, t_n)$, which should be one for a strictly $\text{Lip}^+$ stable method. The results for method I
to assume that regularization in one cell would be enough, indeed are observed. Since the chosen CFL-condition corresponds to \( \delta \) of Figure 3, we give the same curves for method II. Similar results as for method I visible that the \( \text{Lip}_h \) are presented in the left plot of Figure 3. The three curves correspond to varying \( h \). From top to bottom, \( h = 2400^{-1} \), \( h = 1200^{-1} \) and \( h = 400^{-1} \). It is clearly visible that the \( \text{Lip}^2 \)-stability deteriorates under mesh refinement. In the right plot of Figure 3, we give the same curves for method II. Similar results as for method I are observed. Since the chosen CFL-condition corresponds to \( \delta = 0 \) it is optimistic to assume that regularization in one cell would be enough, indeed \( K_{\xi=1}^\circ \) contains

\[
\begin{array}{c|c|c|c}
 N & \|u - u_h\|_{L^1(\mathbb{R})} & \|u - u_h\|_{L^2(\mathbb{R})} & \|u - u_h\|_{-1} \sim H^{-1}(\mathbb{R}) \\
100 & 0.036 & 0.071 & 6.4 \cdot 10^{-3} \\
200 & 0.018 (1.0) & 0.049 (0.5) & 3.2 \cdot 10^{-3} (1.0) \\
400 & 9.4 \cdot 10^{-3} (0.9) & 0.034 (0.5) & 1.6 \cdot 10^{-3} (1.0) \\
800 & 4.7 \cdot 10^{-3} (1.0) & 0.023 (0.6) & 7.9 \cdot 10^{-4} (1.0) \\
\end{array}
\]

Table 2. nonsmooth solution, forward Euler in time, \( k = h^2 \)

\[
\begin{array}{c|c|c|c}
 N & \|u - u_h\|_{L^1(\mathbb{R})} & \|u - u_h\|_{L^2(\mathbb{R})} & \|u - u_h\|_{-1} \sim H^{-1}(\mathbb{R}) \\
100 & 2.6 \cdot 10^{-3} & 3.7 \cdot 10^{-3} & 9.7 \cdot 10^{-4} \\
200 & 6.9 \cdot 10^{-4} (1.9) & 1.0 \cdot 10^{-3} (1.9) & 2.3 \cdot 10^{-4} (2.0) \\
400 & 1.8 \cdot 10^{-4} (1.9) & 3.0 \cdot 10^{-4} (1.7) & 5.6 \cdot 10^{-5} (2.0) \\
800 & 4.7 \cdot 10^{-5} (1.9) & 8.9 \cdot 10^{-5} (1.8) & 1.4 \cdot 10^{-5} (2.0) \\
\end{array}
\]

Table 3. smooth solution, 2nd order Runge-Kutta in time, \( k = \frac{1}{4} h \)

are presented in the left plot of Figure 3. The three curves correspond to varying \( h \). From top to bottom, \( h = 2400^{-1} \), \( h = 1200^{-1} \) and \( h = 400^{-1} \). It is clearly visible that the \( \text{Lip}^2 \)-stability deteriorates under mesh refinement. In the right plot of Figure 3, we give the same curves for method II. Similar results as for method I are observed. Since the chosen CFL-condition corresponds to \( \delta = 0 \) it is optimistic to assume that regularization in one cell would be enough, indeed \( K_{\xi=1}^\circ \) contains

\[
\begin{array}{c|c|c|c}
 N & \|u - u_h\|_{L^1(\mathbb{R})} & \|u - u_h\|_{L^2(\mathbb{R})} & \|u - u_h\|_{-1} \sim H^{-1}(\mathbb{R}) \\
100 & 3.6 \cdot 10^{-2} & 7.2 \cdot 10^{-2} & 1.6 \cdot 10^{-2} \\
200 & 1.8 \cdot 10^{-2} (1.0) & 4.9 \cdot 10^{-2} (0.6) & 8.0 \cdot 10^{-3} (1.0) \\
400 & 9.2 \cdot 10^{-3} (1.0) & 3.3 \cdot 10^{-2} (0.6) & 4.0 \cdot 10^{-3} (1.0) \\
800 & 4.6 \cdot 10^{-3} (1.0) & 2.3 \cdot 10^{-2} (0.5) & 2.0 \cdot 10^{-3} (1.0) \\
\end{array}
\]

Table 4. nonsmooth solution, 2nd order Runge-Kutta in time, \( k = \frac{1}{4} h \)
all elements with positive gradient. Therefore in the left plot of Figure 4 we report
the same quantity computed with \( h = 2400^{-1} \) and the strengthened CFL-condition
\( k = \frac{1}{4} h^{1.25} \), corresponding to \( \delta = 1 - h^{\frac{1}{4}} \), both for methods I and II. We see that the
strengthened CFL-condition reduces the perturbations of the gradient in both cases.
However only for method II is the solution \( Lip^+ \)-stable. Numerical experiences not
reported here showed that under the CFL-condition \( k = h^{2} \) used in the convergence
study in the previous section also method I was completely \( Lip^+ \)-stable.

We then considered the same computation using method III, \( k = \frac{1}{4} h \). The time
evolution of \( \lambda_n \) is reported in Figure 4, right plot. This time the upper curve is
\( h = 400^{-1} \) (dotted), the middle curve \( h = 1200^{-1} \) and the bottom curve \( h = 2400^{-1} \),
so we see that in this case the \( Lip^+ \) stability improves under mesh refinement.

To illustrate the instability qualitatively we present the upper section of the
rarefaction wave at \( T = 1 \), computed under the hyperbolic CFL-condition \( k = \frac{1}{4} h \)
in Figure 5 using method I (left plot) and method III (right plot). The spurious
oscillations created by the singularity at the crest of the rarefaction are clearly seen
in the left plot, but not present in the right. Even more striking are the plots in
Figure 6, showing the gradient in the rarefaction wave for the same computation,
using method I (left plot) and method III (right plot).

7. Conclusion

We have studied some shock capturing finite element methods using the framework
for error estimation introduced by Nessyahu and Tadmor. We proved that a finite
element method using standard first order artificial viscosity and lumped mass for
the time derivative satisfies the necessary stability conditions for the error estimates.
We then showed that nonlinear shock capturing leads to BV-stable approximations
when discretized using the forward Euler method in time and therefore also for
Fig. 4. Left: Time evolution of the relative $\text{Lip}^+$ violation, $\lambda_n$, $h = 2400^{-1}$, $k = \frac{1}{4}h^{1.25}$. Upper curve: method I; lower curve: method II. Right: Time evolution of the relative $\text{Lip}^+$ violation, $\lambda_n$, $k = \frac{1}{4}h$, method III. Upper curve: $h = 400^{-1}$ (dotted); middle curve $h = 1200^{-1}$; lower curve $h = 2400^{-1}$.

Fig. 5. Zoom of the solution in the rarefaction for $k = \frac{1}{4}h$, $h = 2400^{-1}$. Left: method I. Right: method III.

strong stability preserving Runge-Kutta methods. To prove a priori stability estimates that are sufficient for the error analysis to hold also in the case of nonlinear shock-capturing we introduced a modified viscosity coefficient, regularized closed to local maxima of the gradient. Some of the unproven conjectures were verified numerically, in particular the global high order of the nonlinear scheme applied to smooth solutions.
The numerical investigations showed that the SSP Runge-Kutta method has excellent $\text{Lip}^+$-stability properties without any additional modification of the nonlinear viscosity (3.4) and that the $\text{Lip}^+$ stability of the forward Euler method was substantially improved provided the CFL-condition was strengthened.

References


