Journal of Computational and Applied Mathematics 263 (2014) 299–311

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A posteriori error estimation for the Lax–Wendroff finite difference scheme

J.B. Collins [a], Don Estep [b,*], Simon Tavener [a]

[a] Department of Mathematics, Colorado State University, Fort Collins, CO 80523, United States

[b] Department of Statistics, Colorado State University, Fort Collins, CO 80523, United States

HIGHLIGHTS

- Formulation of the Lax–Wendroff difference scheme for conservation laws as a finite element method.
- Goal oriented a posteriori error estimate for the Lax–Wendroff finite difference scheme.
- Investigation of accuracy of the computational error estimate.

ARTICLE INFO

Article history:
Received 11 April 2013
Received in revised form 3 November 2013

Keywords:
Burgers equation
Conservation law
Finite difference scheme
A posteriori error estimate
Dual problem

ABSTRACT

In many application domains, the preferred approaches to the numerical solution of hyperbolic partial differential equations such as conservation laws are formulated as finite difference schemes. While finite difference schemes are amenable to physical interpretation, one disadvantage of finite difference formulations is that it is relatively difficult to derive the so-called goal oriented a posteriori error estimates. A posteriori error estimates provide a computational approach to numerically compute accurate estimates in the error in specified quantities computed from a numerical solution. Widely used for finite element approximations, a posteriori error estimates yield substantial benefits in terms of quantifying reliability of numerical simulations and efficient adaptive error control.

The chief difficulties in formulating a posteriori error estimates for finite difference schemes is introducing a variational formulation – and the associated adjoint problem – and a systematic definition of residual errors. In this paper, we approach this problem by first deriving an equivalency between a finite element method and the Lax–Wendroff finite volume method. We then obtain an adjoint based error representation formula for solutions obtained with this method. Results from linear and nonlinear viscous conservation laws are given.

1. Introduction

In this paper, we derive a computable, goal-specific a posteriori error estimate for the Lax–Wendroff finite difference scheme for the viscous nonlinear conservation law in one dimension,

$$\begin{cases} u_t + f(u)_x = \epsilon u_{xx}, & x \in \mathbb{S}^1,\ 0 < t \le T, \\ u(x,0) = u_0(x), & x \in \mathbb{S}^1, \end{cases} \tag{1}$$

* Corresponding author. Tel.: +1 9704916722.
E-mail addresses: jbcolli2@gmail.com (J.B. Collins), estep@stat.colostate.edu, don.estep@gmail.com (D. Estep), tavener@math.colostate.edu (S. Tavener).

0377-0427/$ – see front matter
http://dx.doi.org/10.1016/j.cam.2013.12.035

where $\epsilon > 0, f : \mathbb{R} \to \mathbb{R}$ is smooth, and $\mathbb{S}^1$ is the one dimensional unit sphere, i.e. we assume periodic boundary conditions. We also apply the estimate to an example with $\epsilon = 0$, in which case, we also assume f is convex. Periodic boundary conditions greatly simplifies the presentation since boundary conditions can introduce serious complications for hyperbolic and convection-dominated problems. Generally, a posteriori error estimates can be extended to include the effects of error in boundary conditions and pursuing such analysis for hyperbolic equations is an interesting problem.

In contrast to a priori convergence and accuracy analysis, a posteriori error estimate yields an accurate estimate of the error in information $\mathcal{Q}(u)$ computed from a particular numerical solution U. The ingredients of the a posteriori error analysis include variational analysis, adjoint operators, and computable residuals. Computable accurate error estimates are an important component of reliability, uncertainty quantification, and adaptive error control. Adjoint-based a posteriori error estimation has been developed and implemented widely over the past few decades within the finite element community [1–4]. Much of the work in a posteriori error estimation has been directed towards elliptic and parabolic problems, however there is some recent research targeting conservation laws. Barth and Larson, [5–7], considered error estimation for the discontinuous Galerkin method and certain Godunov methods. Other work, [8–11] has addressed adaptivity and the necessary error estimation for various conservation laws. All of the studies for conservation laws assume the approximate solution is obtained by a finite element method e.g., discontinuous Galerkin. This method is well-suited for a posteriori error estimation, but it is also relatively new.

The first methods developed for hyperbolic problems were finite difference methods. These included methods such as Lax–Wendroff [12], Godunov [13], MacCormack [14], upwind [15], and many others[16,17]. See [18,19] for a review of some of the early finite difference schemes for hyperbolic problems. These methods were developed to deal with hyperbolic problems, in particular, to capture discontinuities effectively. They were also developed to have low computational cost, being explicit methods. Therefore, many large scale codes implement these methods [20–22]. Therefore, it is useful to obtain a posteriori error estimates for solutions obtained by these finite difference methods.

In this paper, we derive an a posteriori error estimate for the Lax–Wendroff scheme. The main ideas of the analysis can be used to derive estimates for other finite difference schemes, though the specific details would depend on the particular scheme in question. To derive the estimate, we first rewrite the Lax–Wendroff method as a "nodally equivalent" finite element method. We then perform an adjoint based error analysis for this finite element method, which can then be interpreted as an estimate for the original difference scheme. The error estimate can be partitioned into a sum of contributions, each corresponding to specific approximations made in the discretization. This quantification of various contributions to the error is essential to obtain an accurate estimate and it is also useful for adaptivity, as discussed in the conclusion. Since this scheme is explicit, we use the work of [23], where error estimation was performed for explicit time stepping schemes for ordinary differential equations.

The structure of the paper is organized as follows. We recall the derivation of the Lax–Wendroff scheme in Section 2. In Section 3, we formulate a finite element method that is equivalent to the Lax–Wendroff scheme. We present the a posteriori error estimate in Section 4. Numerical results for the linear advection and Burgers equations are presented in Section 5.

2. A review of the Lax–Wendroff finite difference scheme

The Lax–Wendroff scheme is an explicit second order difference scheme. As with other simple difference schemes, simplicity of implementation is an attractive feature. However, the price of higher order approximation is that the Lax–Wendroff scheme is dispersive, which limits usefulness for problems with shocks. Nonetheless, it is still an extremely popular method that is embedded in many legacy codes.

We partition the temporal domain by the nodes, $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$, and define $k_n = t_n - t_{n-1}$, while the spatial domain is partitioned by the nodes, $x_M = x_0 < x_1 < \cdots < x_{M-1} < x_M = x_0$, with the uniform spatial step $h = x_i - x_{i-1}$. The Lax–Wendroff scheme is originally derived for a pure convection problem, that is (1) with $\epsilon = 0$, based on a truncated Taylor series expansion. Assuming $u(x, t)$ is a smooth solution of (1) with $\epsilon = 0$, we consider the approximation of the solution generated by truncating the Taylor series in time:

$$u(x, t + k_n) \approx u(x, t) + k_n u_t(x, t) + \frac{k_n^2}{2} u_{tt}(x, t). \tag{2}$$

Then, using (1), we replace all temporal derivatives with spatial derivatives,

$$u_t = f(u)_x, \qquad u_{tt} = f(u)_{xt} = f(u)_{tx} = (f'(u)u_t)_x = (f'(u)f(u)_x)_x.$$

Approximating the spatial derivatives with centered differences, and using subscripts and superscripts to denote the finite difference approximation, we obtain the update formula for the Lax–Wendroff method,

$$u_i^n = u_i^{n-1} - \frac{k_n}{2h}(f_{i+1}^{n-1} - f_{i-1}^{n-1}) - \frac{k_n^2}{2h^2}\left[f_{i-1/2}'^{n-1}(f_i^{n-1} - f_{i-1}^{n-1}) + f_{i+1/2}'^{n-1}(f_i^{n-1} - f_{i+1}^{n-1})\right], \tag{3}$$

for $n = 1, \ldots, N$ and $i = 0, \ldots, M-1$, and

$$u_i^n = u(x_i, t_n), \qquad f_i^n = f(u(x_i, t_n)), \qquad f_{i+1/2}'^n = f'(u(x_{i+1/2}, t_n)).$$

It is also common to approximate the propagation speed evaluated at the midpoint by, $f_{i+1/2}'^n \approx f'\left(\frac{u_i^n + u_{i+1}^n}{2}\right)$.

To include the viscous term, we use a simple second order finite difference approximation of the second derivative. This retains the order of the method without adding excessive computational expense. The final Lax–Wendroff update formula for Eq. (1) is given by,

$$u_i^n = u_i^{n-1} - \frac{k_n}{2h}(f_{i+1}^{n-1} - f_{i-1}^{n-1}) - \frac{k_n^2}{2h^2}\left[f'^{n-1}_{i-1/2}(f_i^{n-1} - f_{i-1}^{n-1}) + f'^{n-1}_{i+1/2}(f_i^{n-1} - f_{i+1}^{n-1})\right] + \frac{\epsilon k_n}{h^2}[u_{i+1}^{n-1} - 2u_i^{n-1} + u_{i-1}^{n-1}].$$

3. Derivation of an equivalent finite element method

The Lax–Wendroff approximation is defined on a finite set of points on the spatio-temporal domain. In order to perform adjoint-based a posteriori error analysis, we employ a variational formulation of the differential equation over the entire spatio-temporal continuum. It is natural to apply this formulation to a finite element function, which is defined at every point in space and time, but less obvious how to apply it to a finite difference scheme directly. We approach this issue by constructing a finite element method that agrees exactly with the Lax–Wendroff scheme at the nodes, which we call nodal equivalence. The idea of nodal equivalence was introduced in [23] for this purpose.

To obtain such an equivalent method, we examine the various approximations made by the Lax–Wendroff method, and incorporate them into the variational formulation of a finite element approximation. We then employ particular quadratures to evaluate the integrals in the variational formulation, and this yields the nodally equivalent finite element approximation.

3.1. Finite element discretization

We begin by stating the weak form of the viscous conservation law. Find $u \in H^1([0,T]; H^1(\mathbb{S}^1))$ such that,

$$\begin{cases} \displaystyle\int_0^T \langle u_t, v\rangle\, dt = -\int_0^T \left\langle \frac{d}{dx} f(u), v \right\rangle + \epsilon \left\langle \frac{d}{dx} u, \frac{d}{dx} v \right\rangle dt, \quad \forall\, v \in L^2([0,T]; H^1(\mathbb{S}^1)), \\ u(x,0) = u_0(x), \quad x \in \mathbb{S}^1, \end{cases} \tag{4}$$

where $\langle \cdot, \cdot \rangle$ is the $L^2(\mathbb{S}^1)$ inner product and $\epsilon > 0$.

The finite element method we use is similar to the standard continuous Galerkin method. Let,

$$V_h^p = \{v(x) \in H^1(\mathbb{S}^1) : v(x)|_{I_i} \in \mathcal{P}^p(I_i),\ i = 1, \ldots, M\} \tag{5}$$

where $\mathcal{P}^p$ is the space of polynomial functions of degree $\leq p$ and,

$$V_n^{p,q} = \left\{ g(x,t) : g(x,t) = \sum_{j=0}^{q} t^j v_j(x),\ v_j(x) \in V_h^p,\ (x,t) \in S_n \right\}, \tag{6}$$

where $S_n = \mathbb{S}^1 \times [t_{n-1}, t_n]$.

We choose the approximation spaces to be second order, i.e. let $p = q = 1$. We begin by defining the finite element method with exact evaluation of integrals as: Find $U(x,t) \in V_n^{1,1}$ such that $U(x,0) = u_0(x)$ and,

$$\begin{cases} \displaystyle\int_{t_{n-1}}^{t_n} \langle U_t, v\rangle\, dt = -\int_{t_{n-1}}^{t_n} \left[\sum_{j=1}^{M} \left\langle \frac{d}{dx} S_j f(P_n U), v \right\rangle_{I_j} \right. \\ \qquad \left. + \dfrac{k_n}{2} \left\langle f'(P_n U) \dfrac{d}{dx} S_j f(P_n U), \dfrac{d}{dx} v \right\rangle_{I_j} + \epsilon \left\langle \dfrac{d}{dx} P_n U, \dfrac{d}{dx} v \right\rangle_{I_j} \right] dt \quad \forall\, v \in V_n^{1,0}, \\ U(x, t_{n-1}^-) = U(x, t_{n-1}^+), \end{cases} \tag{7}$$

for $n = 1, \ldots, N$. For simplicity, we assume that the spatial grid remains constant for all time. It is possible to have spatial grids that change at each time step by introducing a projection into the continuity condition, but it makes establishing the equivalence with an analogous Lax–Wendroff scheme significantly more complicated.

We now introduce four changes to the ideal finite element method (7) in order to obtain nodal equivalence to the Lax–Wendroff scheme. These are:

1. a spatial projection operator S_j,
2. a temporal mapping P_n,
3. a "correction" term,
4. special quadrature formulas.

The spatial approximation operator $S_j : H^1(I_j) \to \mathcal{P}^1(I_j)$ is a projection onto $\mathcal{P}^1(I_j)$. The effect is to replace spatial derivatives with finite differences. The temporal approximation operator $P_n : H^1([t_{n-1}, t_n]; H^1(\mathbb{S}^1)) \to H^1([t_{n-1}, t_n]; H^1(\mathbb{S}^1))$ is defined by,

$$P_n U(x, t) = U(x, t_{n-1}).$$

This gives an explicit time integration scheme. Next, we introduce the nominal correction term,

$$\frac{k_n}{2}\left\langle f'(P_n U)\frac{d}{dx}S_j f(P_n U), \frac{d}{dx}v\right\rangle_{I_j}.$$

An approximate version of this term is added to the method to correct the loss of order that results from using the crude explicit time extrapolation operator in order to recover second order convergence in time [24]. An analogous term arises in the truncation error analysis of the Lax–Wendroff scheme. The last step in constructing the finite element method is to use certain quadratures to evaluate some of the integrals. In particular, we define the trapezoid and midpoint quadratures by the discrete inner products $\langle\cdot,\cdot\rangle_T$ and $\langle\cdot,\cdot\rangle_M$ respectively. With these changes, the finite element method is written as: Find $U(x, t) \in V_n^{1,1}$ such that,

$$\begin{cases} \displaystyle\int_{t_{n-1}}^{t_n} \langle U_t, v\rangle_T \, dt = -\int_{t_{n-1}}^{t_n} \sum_{j=1}^{M} \left\langle \frac{d}{dx}S_j f(P_n U), v\right\rangle_{I_j} \\ \quad \displaystyle + \frac{k_n}{2}\left\langle f'(P_n U)\frac{d}{dx}S_j f(P_n U), \frac{d}{dx}v\right\rangle_{I_j,M} + \epsilon\left\langle \frac{d}{dx}P_n U, \frac{d}{dx}v\right\rangle_{I_j}, \quad \forall\, v \in V_n^{1,0}, \\ U(x, t_{n-1}^-) = U(x, t_{n-1}^+), \end{cases} \tag{8}$$

for $n = 1, \ldots, N$. This finite element method produces an approximate solution that is nodally equivalent to a Lax–Wendroff approximation, as shown in the following theorem.

Theorem 1 (*Nodal Equivalence*). *If $U(x, t)$ is a solution of (8) and $\{u_i^n\}$ is the Lax–Wendroff finite difference approximation, then*

$$U(x_i, t_n) = u_i^n, \quad i = 0, \ldots, M-1, n = 1, \ldots, N.$$

Proof. In order to show that the values $U(x_i, t_n)$ satisfy the update formula (3) for time t_n to t_{n+1} any node x_i, we choose the test function in (8) to have support $I_i \cup I_{i+1} \times [t_{n-1}, t_n]$ i.e.,

$$v_i^n(x, t) = \begin{cases} \dfrac{x - x_{i-1}}{h} & x \in I_i, t \in [t_{n-1}, t_n] \\ \dfrac{x_i - x}{h} & x \in I_{i+1}, t \in [t_{n-1}, t_n], \end{cases} \tag{9}$$

and examine and evaluate each term in (8) separately.

(i) First, the temporal derivative term:

$$\begin{aligned} \int_0^T \langle U_t, v_i^n\rangle_T \, dt &= \int_{t_{n-1}}^{t_n} \langle U_t, v_i^n\rangle_T \, dt \\ &= \int_{t_{n-1}}^{t_n} \frac{h}{2}(U_t(x_i, t) + U_t(x_i, t)) \, dt \\ &= h(U_i^n - U_i^{n-1}), \end{aligned} \tag{10}$$

where for notational simplicity, we define $U_i^n = U(x_i, t_n)$.

(ii) The flux term: for this and the correction term we use the fact that $\frac{d}{dx}S_j f(P_n U) = \frac{1}{h}(f(U_j^{n-1}) - f(U_{j-1}^{n-1}))$.

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \sum_{j=1}^{M} \left\langle \frac{d}{dx}S_j f(P_n U), v_i^n\right\rangle_{I_j} dt &= \int_{t_{n-1}}^{t_n} \sum_{j=i}^{i+1} \left\langle \frac{d}{dx}S_j f(P_n U), v_i^n\right\rangle_{I_j} dt \\ &= \int_{t_{n-1}}^{t_n} \frac{1}{2}(f(U_i^{n-1}) - f(U_{i-1}^{n-1})) + \frac{1}{2}(f(U_{i+1}^{n-1}) - f(U_i^{n-1})) \, dt \\ &= \frac{k_n}{2}(f(U_{i+1}^{n-1}) - f(U_{i-1}^{n-1})). \end{aligned} \tag{11}$$

(iii) The correction term:

$$\begin{aligned}
&\frac{k_n}{2}\int_{t_{n-1}}^{t_n}\sum_{j=i}^{i+1}\left\langle f'(P_nU)\frac{d}{dx}S_jf(P_nU),\,\frac{d}{dx}v_i^n\right\rangle_{I_j,M}dt\\
&\quad=\frac{k_n}{2h^2}\left[\int_{t_{n-1}}^{t_n}(f(U_i^{n-1})-f(U_{i-1}^{n-1}))\langle f'(P_nU),\,1\rangle_{I_i,M}-(f(U_{i+1}^{n-1})-f(U_i^{n-1}))\langle f'(P_nU),\,1\rangle_{I_{i+1},M}\right]\\
&\quad=\frac{k_n^2}{2h}\left[f'(U_{i-1/2}^{n-1})(f(U_i^{n-1})-f(U_{i-1}^{n-1}))+f'(U_{i+1/2}^{n-1})(f(U_i^{n-1})-f(U_{i+1}^{n-1}))\right],
\end{aligned}\tag{12}$$

where $U_{i+1/2}^{n-1}=\frac{1}{2}(U_i^{n-1}+U_{i+1}^{n-1})$.

(iv) The viscosity term:

$$\begin{aligned}
\epsilon\int_{t_{n-1}}^{t_n}\sum_{j=i}^{i+1}\left\langle\frac{d}{dx}P_nU,\,\frac{d}{dx}v_i^n\right\rangle&=\frac{\epsilon}{h}\int_{t_{n-1}}^{t_n}(U_i^{n-1}-U_{i-1}^{n-1})-(U_{i+1}^{n-1}-U_i^{n-1})\,dt\\
&=\frac{\epsilon k_n}{h}(-U_{i+1}^{n-1}+2U_i^{n-1}-U_{i-1}^{n-1}).
\end{aligned}\tag{13}$$

Using (10)–(13) in (8) and dividing by h, we see that U_i^n satisfies the Lax–Wendroff update formula and therefore that $U_i^n=u_i^n$. □

Given this nodal equivalence, existence and convergence properties of the finite element method can therefore be obtained through the known properties of the Lax–Wendroff method.

4. An error representation formula

We estimate the error in a quantity of interest $\mathcal{Q}(U)$ specified as a linear functional of the solution. Common quantities of interest include the average value, the value at a point, and various moments of the solution. We can compute a general quantity of interest as

$$\mathcal{Q}(U)=\int_0^T\langle U,\,\psi\rangle\,dt+\langle U(x,T),\,\psi_T\rangle,\tag{14}$$

for a specified function ψ and vector ψ_T. This general form includes a component evaluated over the entire spatio-temporal domain and another evaluated over space at the final time. Defining the error to be $e=u-U$, the error estimate is based on an exact error representation formula for the quantity of interest $\mathcal{Q}(e)$. The representation distinguishes various contributions to the error.

The error representation is given as a sum of weighted residuals, where each residual represents a specific discretization effect while the weighting functions are determined by the solution to an adjoint equation to (1). The abstract adjoint equation is given by: Find $\varphi\in H^1([0,T];H^1(\mathbb{S}^1))$ such that,

$$\begin{cases}\displaystyle\int_0^T\left\langle v,\,-\varphi_t-A^*\frac{d}{dx}\varphi\right\rangle+\epsilon\left\langle\frac{d}{dx}v,\,\frac{d}{dx}\varphi\right\rangle=\langle v,\,\psi\rangle,\quad\forall\,v\in L^2([0,T];H^1(\mathbb{S}^1))\\ \varphi(x,T)=\psi_T,\end{cases}\tag{15}$$

where $A=\int_0^1 f'(su+(1-s)U)\,ds$. Note that the functions which define the quantity of interest, ψ and ψ_T, are used as data in the adjoint equation, and that the adjoint problem is solved backwards in time.

Nominally, the concept of an adjoint problem applies to a linear problem and there are several ways to define an adjoint to a nonlinear problem. In the case of a viscous conservation law, we adopt the standard technique of writing the nonlinear flux in "linear" form using the integral mean value theorem, and defining an adjoint with respect to this form. However, (15) cannot be solved numerically in practice because it requires knowledge of the true solution u. The common approach is to use the replacement

$$A\approx\int_0^1 f'(sU+(1-s)U)\,ds=f'(U)$$

to obtain an approximation of the adjoint problem that can be solved numerically. This substitution works well in most situations, but does become problematic in the case of a discontinuous solution, since it can lead to a significant error in the adjoint convection coefficient.

The a posteriori error analysis using the theoretical adjoint problem (15) yields a so-called "error representation formula".

Theorem 2. *Let $U(x,t)$ be a solution of (8). Then,*

$$\int_0^T \langle e,\psi\rangle\,dt + \langle e(T),\psi_T\rangle = \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\sum_{j=1}^{M}\Big(\underbrace{R_j^n(U,\varphi-\pi_h\varphi)}_{\text{Spatial Discretization}} + \underbrace{R_j^n(U,\pi_h\varphi-\pi_k\pi_h\varphi)}_{\text{Temporal Discretization}} + \underbrace{SE1_j^n + SE2_j^n + TE1_j^n + TE2_j^n + TE3_j^n}_{\text{Explicit terms}} + \underbrace{Q1_j^n + Q2_j^n}_{\text{Quadrature}}\Big)\,dt + \underbrace{\langle e(x,0),\varphi(x,0)\rangle}_{\text{Initial error}},$$

where the residual of the finite element method over the spatio-temporal "box" $I_j\times[t_{n-1},t_n]$ is defined by,

$$R_j^n(U,\varphi) = \left\langle -U_t - \frac{d}{dx}S_jf(P_nU),\varphi\right\rangle_{I_j} - \left\langle \frac{k_n}{2}f'(P_nU)\frac{d}{dx}S_jf(P_nU) + \epsilon\frac{d}{dx}P_nU,\frac{d}{dx}\varphi\right\rangle_{I_j},$$

and π_k and π_h are projections onto the temporal and spatial test spaces for the finite element method respectively. The explicit terms are,

- *Spatial Explicit Terms*

$$SE1_j^n = \left\langle \frac{d}{dx}\big(S_jf(U) - f(U)\big),\varphi\right\rangle_{I_j}$$

$$SE2_j^n = \frac{k_n}{2}\left\langle f'(U)\frac{d}{dx}S_jf(U),\frac{d}{dx}\varphi\right\rangle_{I_j}.$$

- *Temporal Explicit Terms*

$$TE1_j^n = \left\langle \frac{d}{dx}S_j\big(f(P_nU) - f(U)\big),\varphi\right\rangle_{I_j}$$

$$TE2_j^n = \frac{k_n}{2}\left\langle f'(P_nU)\frac{d}{dx}S_jf(P_nU) - f'(U)\frac{d}{dx}S_jf(U),\frac{d}{dx}\varphi\right\rangle_{I_j}$$

$$TE3_j^n = \epsilon\left\langle \frac{d}{dx}\big(P_nU - U\big),\frac{d}{dx}\varphi\right\rangle_{I_j}.$$

The quadrature terms are,

$$Q1_j^n = \langle U_t,\pi_k\pi_h\varphi\rangle_{I_j,T} - \langle U_t,\pi_k\pi_h\varphi\rangle_{I_j}$$

$$Q2_j^n = \frac{k_n}{2}\left\langle f'(P_nU)\frac{d}{dx}S_jf(P_nU),\frac{d}{dx}\pi_k\pi_h\varphi\right\rangle_{I_j,M} - \frac{k_n}{2}\left\langle f'(P_nU)\frac{d}{dx}S_jf(P_nU),\frac{d}{dx}\pi_k\pi_h\varphi\right\rangle_{I_j}.$$

The first two terms on the right hand side of (16) (spatial and temporal discretization) quantify the effect of the approximation of the true solution space by a finite dimensional finite element space. The remaining terms in (16) quantify the effects of various approximations to the differential operator. The first two terms are considered the principle "discretization" expressions. They are affected by the so-called Galerkin orthogonality, or cancellation of local discretization errors that results from the Galerkin formulation. This is reflected in the adjoint weights, which have the form of differences between the adjoint solution and a projection into the finite element space. This is known as "Galerkin orthogonality". The other terms in the error representation do not have these projections.

Proof. We define two nonlinear forms that represent the residual for (1) and (7).

$$\mathcal{N}(u,v) := \int_0^T \langle u_t + f(u)_x, v\rangle + \epsilon\langle u_x,v_x\rangle\,dt$$

$$\mathcal{N}_M(U,v) := \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\sum_{j=1}^{M}\left\langle U_t + \frac{d}{dx}S_jf(P_nU),v\right\rangle_{I_j} + \left\langle \epsilon\frac{d}{dx}P_nU - \frac{k_n}{2}f'(P_nU)\frac{d}{dx}S_jf(P_nU),v_x\right\rangle_{I_j}dt.$$

We now have the following relation between the residual R_j^n and the above forms,

$$\begin{aligned}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\sum_{j=1}^{M}R_j^n(U,v)\,dt &= \mathcal{N}(u,v) - \mathcal{N}_M(U,v)\\ &= \mathcal{N}(u,v) - \mathcal{N}(U,v) + \mathcal{N}(U,v) - \mathcal{N}_M(U,v)\\ &= \mathcal{N}'(su + (1-s)U,v;e) + \mathcal{N}(U,v) - \mathcal{N}_M(U,v),\end{aligned}$$

where $\mathcal{N}'$ is the Fréchet derivative and,

$$\mathcal{N}'(su + (1-s)U, v; e) = \int_0^T \langle e_t + (Ae)_x, v\rangle + \epsilon\langle e_x, v_x\rangle \, dt.$$

This gives the following relation between the computable residual and the error,

$$\mathcal{N}_M(U, v) - \mathcal{N}(U, v) + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\sum_{j=1}^{M} R_j^n(U, v) = \int_0^T \langle e_t + (Ae)_x, v\rangle + \epsilon\langle e_x, v_x\rangle \, dt. \tag{16}$$

Using the adjoint problem (16) and integration by parts, we have,

$$\begin{aligned}\int_0^T \langle e, \psi\rangle \, dt &= \int_0^T \langle e, -\varphi_t - A^*\varphi_x\rangle + \epsilon\langle e_x, \varphi_x\rangle \, dt \\ &= \int_0^T \langle e_t + (Ae)_x, \varphi\rangle + \epsilon\langle e_x, \varphi_x\rangle \, dt + \langle e(x,0), \varphi(x,0)\rangle - \langle e(x,T), \psi_T\rangle.\end{aligned}$$

Rearranging and using (16),

$$\int_0^T \langle e, \psi\rangle \, dt + \langle e(x,T), \psi_T\rangle = \mathcal{N}_M(U, \varphi) - \mathcal{N}(U, \varphi) + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\sum_{j=1}^{M} R_j^n(U, \varphi)\, dt + \langle e(x,0), \varphi(x,0)\rangle.$$

The first two terms on the right hand side are equal to the explicit terms in (16). Galerkin orthogonality and strategic addition of zero are used to obtain the rest of the terms in the formula. □

5. Numerical results

We consider two example problems, a linear advection problem, and the nonlinear Burgers equation. The numerical results are designed to demonstrate the accuracy of the error estimate. We define the effectivity ratio to be,

$$\mathcal{E} = \frac{\text{Error Estimate}}{\text{Exact Error}}.$$

To evaluate this, we construct a problem that has a specified exact solution or we compute a highly accurate numerical reference solution that can be used to approximate the exact error. An accurate estimator has an effectivity ratio close to 1. We plot $|1 - \mathcal{E}|$ to show that this quantity is small. We also use the estimate to investigate the relative contributions of the various components that comprise the total error.

It is well-known that the time step must be constrained to ensure stability when using explicit methods to solve time-dependent problems. For hyperbolic problems, the constraint is typically of the form $k_n = O(h)$, whereas for parabolic problems, it is typically $k_n = O(h^2)$. For $\epsilon > 0$, we do not want to simply choose the stricter parabolic bound on the time step, but use von Neumann analysis to determine a condition that depends on the size of ϵ, i.e., the degree of hyperbolicity or parabolicity of the problem.

If $\epsilon = 0$, that is the problem is purely hyperbolic, we use the typical CFL condition,

$$k_n \le cah,$$

where $0 < c < 1$ is the CFL number and $a = \max_{t\in[t_{n-1},t_n]}\{f'(u)\}$. For $\epsilon > 0$, we use von Neumann analysis to derive a constraint that approaches the hyperbolic condition as $\epsilon \to 0$ and approaches the parabolic condition as ϵ increases.

$$k_n \le \min\left\{\frac{ch^2}{2\epsilon}, \left|c^2\left(\sqrt{\frac{\epsilon^2}{a^4} + \frac{h^2}{a^2}} - \epsilon\right)\right|\right\}.$$

This is not a strict theoretical bound, but it is adequate to maintain stability.

The numerical solution of the adjoint problem involves several steps. First, we recall the linearization approximation, where we replace the exact linearization operator by $f'(U) \approx A$. We must then compute an approximate adjoint solution $\Phi \approx \varphi$ to (15). This is used in the error representation formula (16) to make the formula computable. Also since we project the adjoint solution onto the finite element space in (16), we must compute Φ with a higher order method or finer mesh. We use the third order in space and second order in time continuous Galerkin finite element method to solve the approximate adjoint problem, which allows direct evaluation of the Galerkin orthogonality weight $\Phi - \pi_k\pi_h\Phi$.

These choices for solving the adjoint problem work well for PDEs with smooth solutions, but can be less effective if the solution is discontinuous, as we illustrate in Section 5.3. In the case of pure conservation laws, the definition and numerical solution of an appropriate adjoint problem remains an interesting research problem.

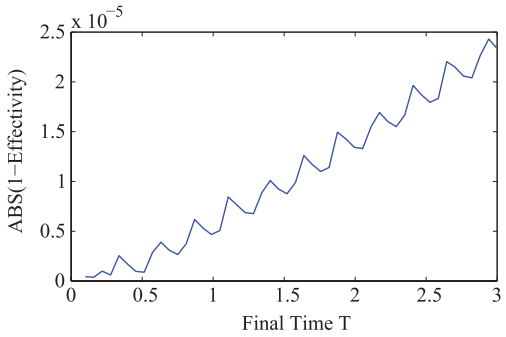

Fig. 1. Error in the effectivity ratio for the linear advection problem. The initial data is defined as $u_0 = \sin(\pi x)$ with coefficients $a = 1, \epsilon = 0.01$. The quantity of interest is a point value at the final time.

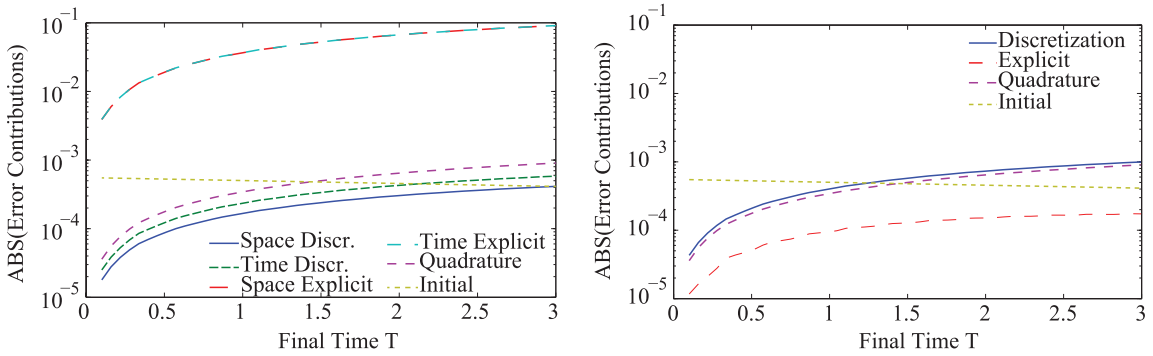

Fig. 2. Error contributions for the linear advection problem. The initial data is defined as $u_0 = \sin(\pi x)$ with coefficients $a = 1, \epsilon = 0.01$. The quantity of interest is a point value at the final time.

5.1. Smooth linear advection

We begin by considering the linear advection problem, that is,

$$\begin{cases} u_t + a u_x = \epsilon u_{xx}, & x \in \mathbb{S}^1, 0 < t \leq T, \\ u(x, 0) = u_0(x), \end{cases} \tag{17}$$

where $a \in \mathbb{R}$ denotes the velocity of propagation. We choose $a = 1$. When $\epsilon = 0$ the solution is simply direct transport of the initial condition with constant speed and direction. The number of spatial steps is $N = 32$, and the CFL number is $c = 0.95$.

5.1.1. Smooth initial conditions

We first consider a smooth initial condition, $u_0(x) = \sin(\pi x)$. We choose $\epsilon = 0.01$ to obtain a convection dominated problem. The quantity of interest is a point value at a final time, and we estimate the error for multiple final times T. Fig. 1 plots the error in the effectivity ratio $|1 - \mathcal{E}|$, against the final time, while Fig. 2 plots the various error contributions against the final time. The error estimator is extremely accurate, at least four digits, for this problem. The error in the effectivity increases monotonically as we solve to greater final time, but not rapidly and is still roughly as accurate.

While derivation of the a posteriori estimate presented above nominally does not hold for $\epsilon = 0$, this case is often the primary focus in practice. We examine the estimate's behavior in this case. The results are comparable to the viscous case presented above. The effectivity ratio is of the same accuracy, though it remains constant instead of growing with time, and the error contributions are very similar.

We also consider the rate of convergence of the various error contributions. Fig. 3 shows how each contribution converges. Of particular interest are the explicit contributions. Both the spatial and temporal explicit contributions converge to first order, however, due to cancellation, their sum converges to third order. This illustrates a property of the Lax–Wendroff method, in which first order spatial and temporal approximations are combined to form a higher order method. We also note that the total error estimate converges to second order, as expected for the Lax–Wendroff scheme.

5.1.2. Effect of explicit time integration

Examining the individual error contributions, we see that the explicit error dominates for both the inviscid and viscous linear advection problems. However, a good deal of cancellation occurs between the spatial and temporal explicit errors, and so when the total explicit error is considered, it is comparable to the other error contributions.

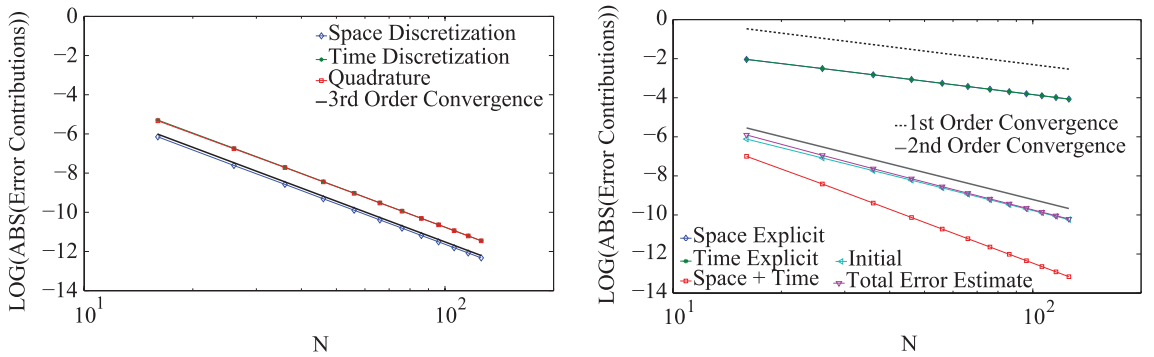

Fig. 3. Error contributions of the error estimate versus number of spatial steps. Linear advection with initial data $u_0(x) = \sin(\pi x)$, $a = 1$ and $\epsilon = 0$. Quantity of interest is a point value at final time.

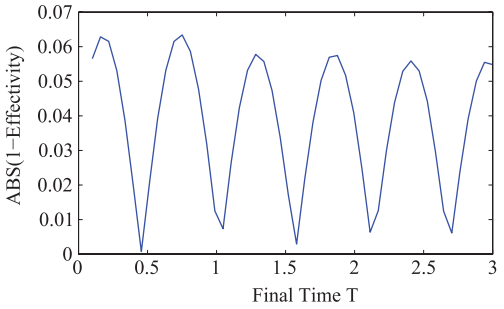

Fig. 4. Error in the effectivity ratio for the linear advection problem. The initial data is a top hat function with coefficients $a = 1, \epsilon = 0$. The quantity of interest is a point value at the final time.

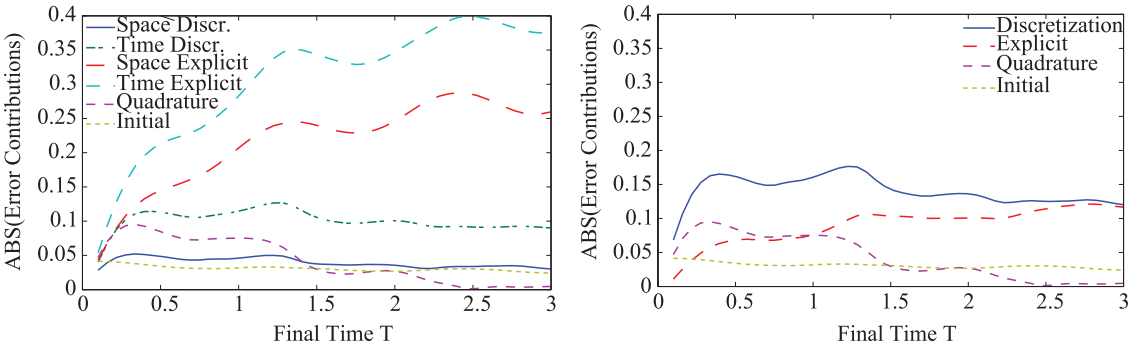

Fig. 5. Error contributions for the linear advection problem. The initial data is a top hat function with coefficients $a = 1, \epsilon = 0$. The quantity of interest is a point value at the final time.

5.1.3. *Discontinuous linear advection*

We next consider linear advection with a discontinuous initial condition, and no diffusion. We test the error estimator with a quantity of interest at the points of discontinuity at the final time. Fig. 4 shows the error in the effectivity and the exact error. Fig. 5 shows the various error contributions. We see that error in the effectivity is considerably higher than for smooth initial conditions. However, the error is still estimated to roughly one digit, and so is relatively accurate.

5.2. *Viscous Burgers equation*

The second example is Burgers equation,

$$\begin{cases} u_t + \left(\frac{1}{2}u^2\right)_x = \epsilon u_{xx}, & x \in \mathbb{S}^1, 0 < t \leq T, \\ u(x, 0) = u_0(x). \end{cases} \tag{18}$$

We begin by demonstrating the effectiveness of the error estimator on a smooth, diffusion dominated problem. We use initial data $u_0(x) = \sin(\pi x)$, and viscosity coefficient $\epsilon = .15$. The exact solution is plotted for various times in Fig. 6. The

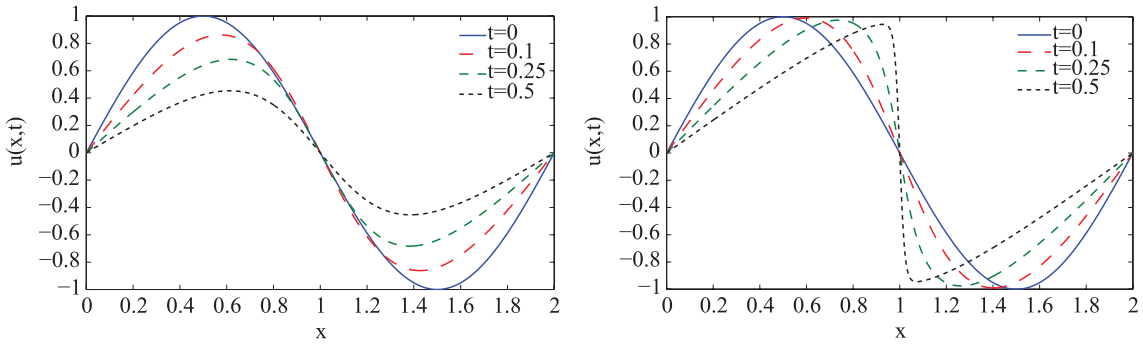

Fig. 6. Exact solution of the viscous Burgers equation with initial data $u_0(x) = \sin(\pi x)$ and viscosity coefficient (a) $\epsilon = 0.15$ and (b) $\epsilon = 0.01$.

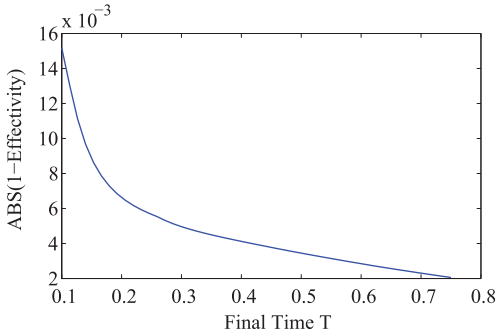

Fig. 7. Error in the effectivity ratio. The Burgers equation with initial data $u_0(x) = \sin(\pi x)$, $\epsilon = 0.15$. Quantity of interest is a point value at the final time.

quantity of interest is the point value at $x = 1$ at the final time. The number of spatial steps and CFL number are the same as for the linear example above. Fig. 7 plots the error in the effectivity ratio, and Fig. 8 plots the various error contributions. We present the signed errors in this case to show how the various contributions can cancel and combine to determine the total error. The error contributions have all roughly the same magnitude and there is no significant cancellation.

Of particular interest is the temporal explicit error contribution. We see here that it dominates and there is no significant cancellation between it and the spatial explicit contributions. For this example, the error could be reduced by using an implicit method, essentially eliminating the temporal explicit contribution.

Next we decrease the viscosity coefficient and consider the accuracy of the error estimator as the problem becomes convection dominated on a fixed mesh. As $\epsilon \to 0$, the solution forms a sharper and sharper internal layer. The exact solution for $\epsilon = 0.01$ is shown in Fig. 6. This sharp gradient makes it difficult to estimate the error, due to the fact that the solution fails to be resolved at some point since the fixed mesh becomes too coarse. This increases the linearization error which leads to error in the adjoint solution. The adjoint solution at time $t = 0$ is plotted in Fig. 9. Note that the error is highly sensitive near where the shock is forming. Fig. 11 shows the error in the effectivity ratio as ϵ varies. We clearly see a reduction in the accuracy of the error estimator as $\epsilon \to 0$, due to the error in the adjoint solution.

However, we see that the estimator is accurate for $\epsilon > 0.02$, therefore we examine the contributions to the error as ϵ decreases to this value. Fig. 10 shows the error contributions as ϵ decreases. From this plot we can obtain some interesting information about the error. We see that the temporal discretization error becomes significant and temporal explicit error contribution dominates as $\epsilon \to 0$.

5.3. Inviscid Burgers equation

We next consider the inviscid Burgers equation, this is (18) with $\epsilon = 0$. A well known property of this problem is that the solution can become discontinuous in finite time, even with smooth initial data.

Since the Lax–Wendroff is a conservative method, $\frac{d}{dt}\int_{\mathbb{S}^1} u(x,t)\,dx = 0$ should hold in a discrete sense. If the initial condition can be represented by the approximation space, then for a conservative method the error in a quantity of interest defined by $\psi = 1, \psi_T = 0$ should be zero up to machine precision. Using this quantity of interest, we show that not only is the analogous finite element method nodally equivalent to Lax–Wendroff, it also preserves the discrete conservative property. We use piecewise linear continuous initial data, shown in Fig. 12, which can be represented exactly in the finite element space. Table 1 shows the estimated error in the quantity of interest. This demonstrates the accuracy of the error estimator. The approximate solution obtained from the Lax–Wendroff method is also shown in Fig. 12. Note that while considerable oscillations occur around the point of discontinuity, they are generated in a particular way so that the quantity is conserved.

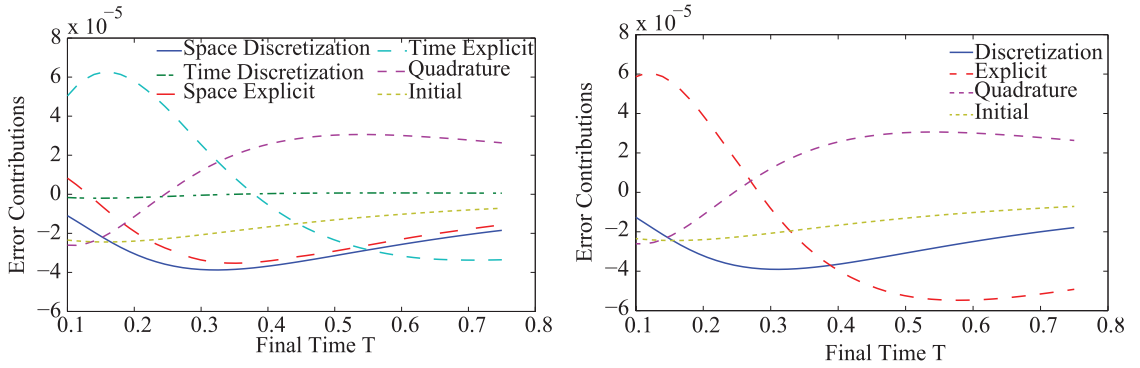

Fig. 8. Error contributions. The Burgers equation with initial data $u_0(x) = \sin(\pi x)$, $\epsilon = 0.15$. Quantity of interest is a point value at the final time.

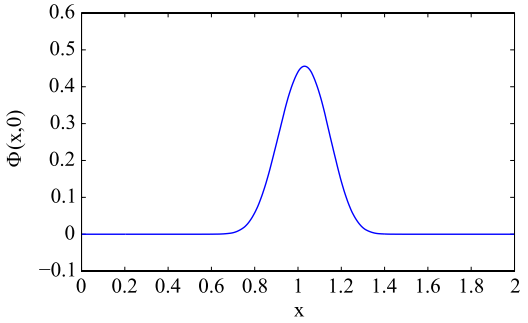

Fig. 9. Solution of the adjoint problem at final time for viscous Burgers equations with $\epsilon = 0.1$. Quantity of interest is a point value at the final time.

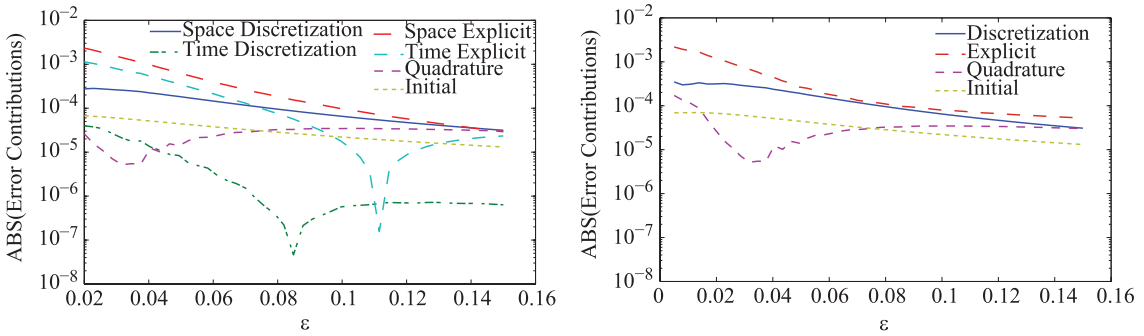

Fig. 10. Error contributions. The Burgers equation with initial data $u_0(x) = \sin(\pi x)$ as $\epsilon \to 0$. Quantity of interest is a point value at the final time.

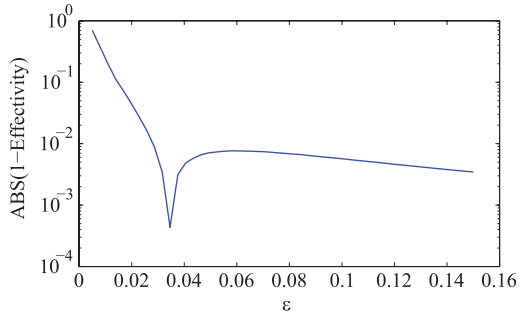

Fig. 11. Error contributions as a function of ϵ. The Burgers equation with initial data $u_0(x) = \sin(\pi x)$. Quantity of interest is a point value at the final time.

Finally we consider a problem in which a shock forms from smooth initial conditions $u_0(x) = \sin(\pi x)$. In this problem a stationary shock forms at $x = 0$ at time $t = \frac{1}{\pi}$. As above, we choose the quantity of interest to be a point value at the final

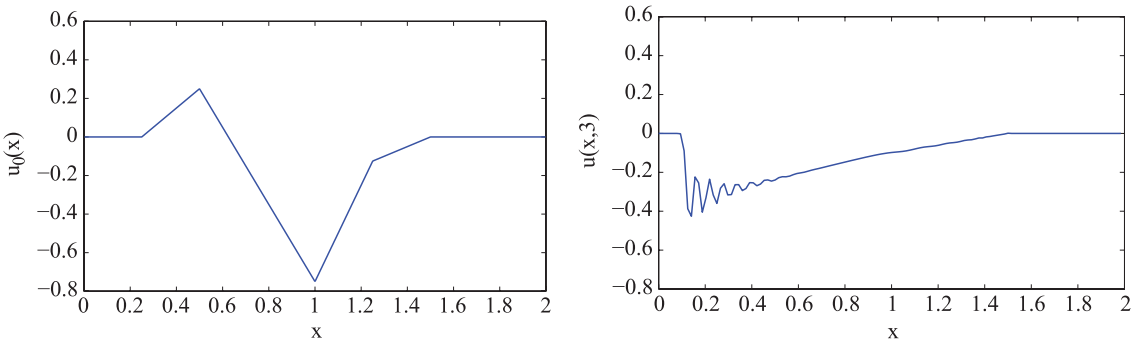

Fig. 12. (Left) Piecewise linear initial conditions used to test the conservative property of Lax–Wendroff. This is chosen so that it can be exactly represented by the finite element space. (Right) Lax–Wendroff approximation at final time $T = 3$.

Table 1
Test of conservativeness of Lax–Wendroff.

N	Error estimate
8	2.259e−16
16	4.906e−16
32	2.708e−16
64	1.398e−15
128	1.794e−15

time, in particular at $x = 0$, and examine the effectivity ratio for various final times. Fig. 13 shows the error in the effectivity ratio as a function of the final time. The time when the shock is formed is marked on the plot.

We see that after the formation of the shock, the error estimate is no longer reliably accurate. A significant factor appears to be "linearization error" in the formulation of the computational adjoint problem. We see this by considering the linearized coefficient $A(x)$ in the adjoint problem at the point of discontinuity. By taking the limit from the left and right sides of the shock, we see,

$$\lim_{x\to 1^-} A(x,t) = \int_0^1 f'(su_L + (1-s)U(0,t))\,ds,$$
$$\lim_{x\to 1^+} A(x,t) = \int_0^1 f'(su_R + (1-s)U(0,t))\,ds,$$

where u_L and u_R are the limiting values of the exact solution at the left and right of the shock. The problem being that the exact solution is discontinuous, and the approximate solution is piecewise smooth. Therefore, when we approximate A with $f'(U)$, we introduce an error of $O(|u_L - u_R|)$. So for strong shocks, like the one in this problem, the linearization error is likely to be large and this could make the adjoint solution highly inaccurate, thus destroying the accuracy of the error estimator. This provides motivation to consider some of the alternate ways to define adjoints to nonlinear operators [25].

6. Conclusions

In this work, we derive an a posteriori error estimator for the Lax–Wendroff method. To do this, we derive a finite element method that is nodally equivalent to the Lax–Wendroff method, and perform error estimation on the finite element approximation.

We show that the error estimator works well under expected conditions, that is diffusion dominated nonlinear problems and smooth linear problems. We also see that the error estimator works for linear problems with discontinuous solutions, though not as well as for smooth data. However, for nonlinear problems, such as Burgers equation, we see that the estimator fails when shocks begin to form. This may be due to the linearization error in the formulation of the adjoint problem, which provides motivation for further study of an appropriate choice of adjoint problem for problems with shocks.

A possibility for future work in this area is adaptivity. The error estimator can be used for classical mesh refinement adaptivity. However, the splitting of the error into various contributions allows for a type of adaptivity where the method itself changes on certain portions of the domain to reduce the error. This was illustrated in [23] for ODEs. By modifying the approximating operators S_j and P_n and the quadrature rules on different intervals, a different method can be used on each interval. If these methods are chosen to reduce the dominating error contributions, i.e. a higher order quadrature rule when the quadrature contribution is too large, then the total error can be reduced for less computation cost than refining the grid. This is especially true for hyperbolic problems, as it is not possible to refine the temporal and spatial components of the grid independently due to stability constraints.

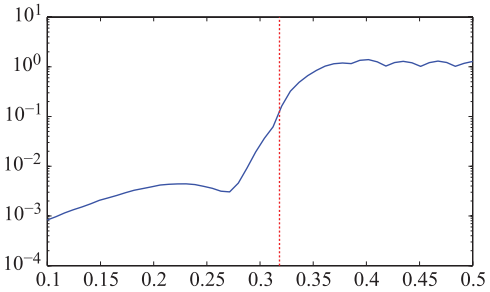

Fig. 13. Effectivity error for the inviscid Burgers equation with initial condition $u_0(x) = \sin(\pi x)$. The dotted line denotes time when shock forms in solution.

Acknowledgments

J. Collin's work is supported in part by the Lawrence Livermore National Laboratory (B590495). D. Estep's work is supported in part by the Defense Threat Reduction Agency (HDTRA1-09-1-0036), Department of Energy (DE-FG02-04ER25620, DE-FG02-05ER25699, DE-FC02-07ER54909, DE-SC0001724, DE-SC0005304, INL00120133, DE0000000SC9279), Idaho National Laboratory (00069249, 00115474), Lawrence Livermore National Laboratory (B573139, B584647, B590495), National Science Foundation (DMS-0107832, DMS-0715135, DGE-0221595003, MSPA-CSE-0434354, ECCS-0700559, DMS-1065046, DMS-1016268, DMS-FRG-1065046, DMS-1228206), National Institutes of Health (#R01GM096192). S. Tavener's work is supported in part by the Department of Energy (DE-FG02-04ER25620, INL00120133), Idaho National Laboratory (00069249, 00115474), Lawrence Livermore National Laboratory (B590495), and the National Science Foundation (DMS-1228206).

References

[1] M. Ainsworth, J. Oden, A posteriori error estimation in finite element analysis, Comput. Methods Appl. Mech. Eng. 142 (1) (1997) 1–88.
[2] R. Becker, R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, Acta Numer. 10 (1) (2001) 1–102.
[3] D. Estep, A posteriori error bounds and global error control for approximation of ordinary differential equations, SIAM J. Numer. Anal. (1995) 1–48.
[4] M. Giles, E. Süli, Adjoint methods for pdes: a posteriori error analysis and postprocessing by duality, Acta Numer. 11 (1) (2002) 145–236.
[5] T. Barth, M. Larson, A posteriori error estimates for higher order godunov finite volume methods on unstructured meshes, in: Finite Volumes for Complex Applications III, London.
[6] T. Barth, Space–time error representation and estimation in navier-stokes calculations, Complex Effects Large Eddy Simul. (2007) 29–48.
[7] M. Larson, T. Barth, A posteriori error estimation for discontinuous Galerkin approximations of hyperbolic systems, Lecture Notes Comput. Sci. Eng. 11 (1999) 363–368.
[8] R. Hartmann, P. Houston, Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws, SIAM J. Sci. Comput. 24 (3) (2003) 979–1004.
[9] R. Hartmann, Multitarget error estimation and adaptivity in aerodynamic flow simulations, SIAM J. Sci. Comput. 31 (1) (2008) 708–731.
[10] R. Hartmann, J. Held, T. Leicht, Adjoint-based error estimation and adaptive mesh refinement for the rans and $k^2 - \omega$ turbulence model equations, J. Comput. Phys. 230 (11) (2011) 4268–4284.
[11] L. Wang, D. Mavriplis, Adjoint-based hp adaptive discontinuous Galerkin methods for the 2d compressible euler equations, J. Comput. Phys. 228 (20) (2009) 7643–7661.
[12] P. Lax, B. Wendroff, Systems of conservation laws, Comm. Pure Appl. Math. 13 (1960) 217–237.
[13] S. Godunov, A. Zabrodin, G. Prokopov, A computational scheme for two-dimensional non stationary problems of gas dynamics and calculation of the flow from a shock wave approaching a stationary state, USSR Comput. Math. Math. Phys. 1 (4) (1962) 1187–1219.
[14] R. MacCormack, Numerical solution of the interaction of a shock wave with a laminar boundary layer, in: Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics, Springer, 1971, pp. 151–163.
[15] R. Richtmyer, K. Morton, Differential Methods for Initial Value Problems, Wiley, New York.
[16] A. Harten, G. Zwas, Self-adjusting hybrid schemes for shock computations, J. Comput. Phys. 9 (3) (1972) 568–583.
[17] V. Rusanov, The calculation of the interaction of non-stationary shock waves and obstacles, USSR Comput. Math. Math. Phys. 1 (2) (1962) 304–320.
[18] R. Richtmyer, A survey of difference methods for non-steady fluid dynamics, no. 63, National Center for Atmospheric Research, 1963.
[19] G. Sod, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, J. Comput. Phys. 27 (1) (1978) 1–31.
[20] F. Brunel, J. Leboeuf, T. Tajima, J. Dawson, M. Makino, T. Kamimura, Magnetohydrodynamic particle code: Lax–Wendroff algorithm with finer grid interpolations, J. Comput. Phys. 43 (2) (1981) 268–288.
[21] R. Cameron, L. Gizon, K. Daiffallah, Slim: a code for the simulation of wave propagation through an inhomogeneous, magnetised solar atmosphere, Astronom. Nachr. 328 (3–4) (2007) 313–318.
[22] J. Desombre, D. Morichon, M. Mory, Rans $v^2 - f$ simulation of a swash event: detailed flow structure, Coastal Eng. 71 (2013) 1–12.
[23] J. Collins, D. Estep, S. Tavener, A posteriori error estimates for explicit time integration methods, BIT Numerical Mathematics.
[24] J. Trangenstein, Numerical Solution of Hyperbolic Partial Differential Equations, Cambridge University Press, 2009.
[25] G.I. Marchuk, V.I. Agoshkov, V.P. Shutyaev, Adjoint Equations and Perturbation Algorithms in Nonlinear Problems, CRC Press, Boca Raton, FL, 1996.