

SHARP L^1 A POSTERIORI ERROR ANALYSIS FOR NONLINEAR CONVECTION DIFFUSION PROBLEMS

ZHIMING CHEN AND GUANGHUA JI

ABSTRACT. We derive sharp $L^\infty(L^1)$ a posteriori error estimates for initial boundary value problems of nonlinear convection diffusion equations of the form

$$\frac{\partial u}{\partial t} + \operatorname{div} f(u) - \Delta A(u) = g,$$

which displays both parabolic and hyperbolic behavior in a way that depends on the solution itself. The problem is discretized implicitly in time via the method of characteristics and in space via continuous piecewise linear finite elements. The analysis is based on the Kruřkov “doubling of variables” device and the recently introduced “boundary layer sequence” technique to derive the entropy error inequality on bounded domains. The derived a posteriori error estimators have the correct convergence order in the region where the solution is smooth and recover the standard a posteriori error estimators known for parabolic equations with strong diffusions.

1. INTRODUCTION

A posteriori error estimates are computable quantities in terms of the discrete solution and data that measure the actual discrete errors without the knowledge of exact solutions. The adaptive finite element method based on a posteriori error estimates initiated in [3] provides a systematic way to refine or coarsen the mesh according to the local a posteriori error estimators on the elements. There are considerable efforts in the literature devoted to the development of a posteriori error analysis and efficient adaptive algorithms for various linear and nonlinear parabolic partial differential equations (see e.g. [14, 16, 7, 5, 8] and the reference therein).

Let Ω is a bounded domain in \mathbf{R}^d ($d = 1, 2, 3$) with Lipschitz boundary and $T > 0$. In this paper, we consider the sharp a posteriori error analysis for the following nonlinear convection diffusion equation

$$(1.1) \quad \frac{\partial u}{\partial t} + \operatorname{div} f(u) - \Delta A(u) = g \quad \text{in } Q$$

with the initial and boundary conditions

$$(1.2) \quad u|_{t=0} = u_0, \quad u|_{\partial\Omega \times (0, T)} = 0.$$

2000 *Mathematics Subject Classification.* 65N15, 65N30, 65N50.

This author was supported in part by China NSF under grant 10025102 and by China MOST under grant G1999032802. Part of the work was done when the first author was participating in the 2003 Programme Computational Challenges in Partial Differential Equations at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK.

Here $u = u(x, t) \in \mathbf{R}$, with $(x, t) \in Q = \Omega \times (0, T)$. We assume that the function $f : \mathbf{R} \rightarrow \mathbf{R}^d$ is locally Lipschitz continuous, the function $A : \mathbf{R} \rightarrow \mathbf{R}$ is nondecreasing and locally Lipschitz continuous, $g \in L^\infty(Q)$ and $u_0 \in L^\infty(\Omega)$.

Problems of the type (1.1) model a wide variety of physical phenomena including porous media flow, flow of glaciers and sedimentation processes [19]. Our motivation comes from the simulation of flow transport through unsaturated porous media which is governed by the so-called Richards equation [2, 17]

$$(1.3) \quad \frac{\partial S}{\partial t} - \frac{\partial K(S)}{\partial z} - \operatorname{div}(K(S)\nabla p) = 0,$$

where S is the volumetric water content, p is the pressure head, and $K(S)$ is the relative permeability. One of the widely used nonlinear constitutive relations for $S = S(p)$ and $K = K(S)$ in the engineering literature, the so called van Genuchten-Mualem formula, reads as follows

$$S(p) = (1 + |\alpha p|^n)^{-m}, \quad K(S) = S^{1/2}(1 - (1 - S^{1/m})^m)^2$$

where $m = 1 - 1/n$, $\alpha > 0$, $n > 1$ are shape constants which vary for different types of porous media. For (1.3), the existence of weak solutions is considered in [2] and the uniqueness of weak solutions is proved in [24] based on Kruřkov “doubling of variables” technique. Entropy solutions for (1.1) are studied in [4] and [22]. The mathematical techniques developed in [22] play an important role in the analysis in this paper.

Our discretization of (1.1) is based on combining continuous piecewise linear finite elements in space with the characteristic finite difference in time. The method of characteristics originally proposed in [13, 26] is widely used to solve convection diffusion problems in finite element community (cf. e.g. [17, 1, 16, 8, 18]). Given U_h^{n-1} as the finite element approximation of the solution at time t^{n-1} , let τ_n and $V_0^n \subset H_0^1(\Omega)$ be the time step and the conforming linear finite element space at the n th time step, then our discrete scheme reads as following: find $U_h^n \in V_0^n$ such that

$$\left\langle \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n}, v \right\rangle + \langle \nabla A(U_h^n), \nabla v \rangle = \langle \bar{g}^n, v \rangle \quad \forall v \in V_0^n,$$

where $\bar{g}^n = \tau_n^{-1} \int_{t^{n-1}}^{t^n} g(x, t) dt$, $\bar{U}_h^{n-1}(x) = U_h^{n-1}(\tilde{X}(t^{n-1}))$, and the approximate characteristics $\tilde{X}(t)$ is defined by

$$d\tilde{X}/dt = f'(U_h^{n-1}(\tilde{X}(t))), \quad \tilde{X}(t^n) = x.$$

In the linear case when $f(u) = \mathbf{v}u$, $A(u) = \epsilon u$ for some small constant $\epsilon > 0$, $L^2(L^2)$ a posteriori error estimate is proved in [16] based on the duality argument.

The well-known Kruřkov “doubling of variables” technique originally appeared in [21] plays a decisive role in the error estimation (both a posteriori and a priori) for numerical schemes solving the Cauchy problems of nonlinear conservation laws (see e.g. [10, 11, 12, 20] and the reference therein). It is also used recently in [23] for the implicit vortex centered finite volume discretization of the Cauchy problems of (1.1). The common feature of these studies is that the derived local a posteriori error estimators is of the order \sqrt{h} in the region where the solution is smooth, where h is the local mesh size. This degeneration of the order of the local estimators may cause over-refinements for the solution of (1.1) in the region where the diffusion is dominant. We also refer to [15] for a different approach of a posteriori error estimation for nonlinear conservation laws.

The basic assumption in this paper is that the diffusion is positive

$$A'(s) > 0, \quad \forall s \in \mathbf{R}.$$

This assumption includes the Richards equation (1.3) and the viscosity regularization of degenerate parabolic equations, for example, the regularized continuous casting problem which is considered in [8]. The novelty of our analysis with respect to the analysis in [11, 12, 20, 23] lies in the treatment of boundary conditions for the a posteriori error analysis using the Kruřkov technique. This is achieved by using the recently introduced ‘‘boundary layer sequence’’ technique in [22]. The nice consequence of the analysis in this paper is that our a posteriori error estimates are able to recover the standard sharp a posteriori error estimators in the literature derived for parabolic problem with diffusion coefficients bounded uniformly away from zero (see Remark 5.8). Further remarks about the differences of the a posteriori error estimates in this paper from those in [11, 12, 20, 23] can be found in Remark 5.10.

The rest of the paper is organized as follows. In section 2 we set the notation and briefly recall the definition of entropy solutions for (1.1). In section 3 we introduce the discrete problem. In section 4 we derive the important entropy error inequality by using boundary layer sequence technique. In section 5 we derive the a posteriori error estimates and present several remarks. In section 6 we show a numerical example for linear convection diffusion problems.

2. SETTING

Let Ω be a bounded domain in \mathbf{R}^d with Lipschitz boundary. Defining \mathcal{B} as the set of all possible Lipschitz coverings of $\partial\Omega$ in the sense that $\partial\Omega \subset \cup_{B \in \mathcal{B}} B$, and, in some local coordinates $x = (x', x_d)$, there exists a Lipschitz function $x_d = \rho(x')$ such that $B \cap \partial\Omega = \{x \in B : x_d = \rho(x')\}$, $B \cap \Omega = \{x \in B : x_d < \rho(x')\}$. Given $T > 0$, let $Q = \Omega \times (0, T)$. We start by stating the hypotheses concerning the data.

(H1) $f : \mathbf{R} \rightarrow \mathbf{R}^d$ is locally Lipschitz continuous, $f(0) = 0$; f' is uniformly Lipschitz continuous in \mathbf{R} .

(H2) $A : \mathbf{R} \rightarrow \mathbf{R}$ is locally Lipschitz continuous, $A(0) = 0$; $A'(s) > 0$ for any $s \in \mathbf{R}$, and $A' \circ A^{-1}$ is uniformly Holder continuous in \mathbf{R} , i.e.

$$|(A' \circ A^{-1})(s) - (A' \circ A^{-1})(r)| \leq C|s - r|^\gamma \quad \forall s, r \in \mathbf{R},$$

for some constant $0 < \gamma \leq 1$ and some constant $C > 0$.

(H3) $g \in L^\infty(Q)$, $u_0 \in L^\infty(\Omega)$.

We recall the following definition of entropy solutions to the problem (1.1)-(1.2) in [22].

Definition 2.1. A function $u \in L^\infty(Q)$ is an entropy solution of the problem (1.1)-(1.2) if:

(i) (*regularity*) we have

$$A(u) \in L^2(0, T; H_0^1(\Omega)),$$

and, for every $B \in \mathcal{B}$, and any non-negative $\psi \in C_c^\infty(B)$ we have (here $x = (x', x_d)$, $x_d = \rho(x')$ on $B \cap \partial\Omega$)

$$(2.1) \quad (-|u|\psi, \operatorname{sgn}(u)(\nabla A(u) - f(u))\psi) \in \mathcal{DM}(Q)_2,$$

where $DM(Q)_2$ is the set of divergence-measure vector fields in Q defined by

$$DM(Q)_2 = \{(w, v) \in L^2(Q) \times L^2(Q)^d : \exists C > 0 \text{ such that} \\ \left| \int_Q (w\varphi_t + v\nabla\varphi) dx dt \right| \leq C \|\varphi\|_{L^\infty(Q)} \quad \forall \varphi \in C_c^\infty(Q)\};$$

(ii) (*entropy condition in the interior of Q*) u is an entropy solution of the equation with test functions zero on the boundary, i.e.,

$$(2.2) \quad - \int_Q |u - k| \partial_t \phi - \int_Q \operatorname{sgn}(u - k) [f(u) - f(k) - \nabla A(u)] \cdot \nabla \phi \\ \leq \int_Q \operatorname{sgn}(u - k) g \phi,$$

for any $\phi \in H_0^1(Q)$, $\phi \geq 0$, for any $k \in \mathbf{R}$;

(iii) (*initial condition*) the initial condition is assumed to be the limit in L^1 sense,

$$(2.3) \quad \operatorname{esslim}_{t \rightarrow 0^+} \int_\Omega |u(x, t) - u_0(x)| dx = 0.$$

Note that since $A'(s) > 0$ for any $s \in \mathbf{R}$ by (H2), the entropy boundary condition in [22] is satisfied trivially [22, Remark 1.2], and thus we have not included it in above definition of entropy solutions for (1.1)-(1.2). The main implication of the regularity property (2.1) lies in that it provides a proper meaning of the normal trace of the vector $(-|u|\psi, \operatorname{sgn}(u)(\nabla A(u) - f(u))\psi)$ on the boundary. Since our analysis does not involve properties of divergence-measure vector fields, we refer the interested readers to [22] and the reference therein for further discussion on divergence-measure vector fields. It is proved in [22] that (1.1)-(1.2) has a unique entropy solution u in the sense of Definition 2.1. Another definition of entropy solutions for (1.1)-(1.2) can be found in [4].

By taking $k > \operatorname{esssup}_Q u(x, t)$ and $k < \operatorname{essinf}_Q u(x, t)$ in (2.2), it is easy to see that an entropy solution is also a weak solution of the same problem in the following sense.

Definition 2.2. A function u is a weak solution of the problem (1.1)- (1.2) if

$$\partial_t u \in L^2(0, T; H^{-1}(\Omega)), \quad f(u) \in L^2(0, T; L^2(\Omega)), \quad A(u) \in L^2(0, T; H_0^1(\Omega)),$$

and

$$(2.4) \quad \int_0^T \langle \partial_t u, \varphi \rangle dt + \int_Q (-f(u) + \nabla A(u)) \cdot \nabla \varphi dx dt = \int_Q g \varphi dx dt$$

for any $\varphi \in L^2(0, T; H_0^1(\Omega))$ such that $\phi(\cdot, 0) = \phi(\cdot, T) = 0$. The initial condition is assumed to be the limit in L^1 sense as in (2.3). Here $\langle \cdot, \cdot \rangle$ stands for either the inner product in $L^2(\Omega)$ or the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

The existence of weak solutions follows directly from the existence of entropy solutions. Since A is strictly increasing we know that the weak solution u of (2.4) is also unique. By a straightforward application of the weak maximum principle we have $u \in L^\infty(Q)$ (cf. e.g. [22, Lemma 2.4]). Then standard theory for parabolic equations implies that $\partial_t u \in L^2(Q)$. Consequently, by (1.1), $\Delta A(u) \in L^2(Q)$. We summarize these results in the following theorem.

Theorem 2.3. *Let the hypotheses (H1)-(H3) be satisfied. Then there exists a unique weak solution u to (1.1)-(1.2). Moreover, the following regularity results are valid*

$$u \in L^\infty(Q), \quad \partial_t u \in L^2(Q), \quad \Delta A(u) \in L^2(Q).$$

We remark that if the domain Ω has a smooth boundary or Ω is convex, then we can obtain the regularity $u \in L^2(0, T; H^2(\Omega))$ by the standard regularity theory for elliptic equations.

3. DISCRETIZATION

We now introduce the fully discrete problem, which combines continuous piecewise linear finite elements in space with the characteristic finite difference in time. In fact, we use the method of characteristics to discretize the convection [13, 26, 17, 1, 16, 18]. We denote by τ_n the n -th time step and set

$$t^n := \sum_{i=1}^n \tau_i, \quad \varphi^n(\cdot) := \varphi(\cdot, t^n)$$

for any function φ continuous in $(t^{n-1}, t^n]$. Let N be the total number of time steps, that is $t^N \geq T$. If

$$(3.1) \quad dX(t)/dt = f'(u(X(t), t)), \quad X(t^n) = x$$

defines the forward characteristics in (t^{n-1}, t^n) , then $U(t) = u(X(t), t)$ satisfies

$$(3.2) \quad dU(t)/dt = \partial_t u + f'(u) \cdot \nabla u = \partial_t u + \operatorname{div} f(u).$$

The characteristic finite difference method is based on writing

$$\bar{x}^{n-1} = X(t^{n-1}), \quad \bar{u}^{n-1}(x) = u(\bar{x}^{n-1}, t^{n-1}),$$

for $n \geq 1$ and discretizing (3.1) by means of backward differences as follows:

$$\frac{dU^n}{dt} \approx \frac{U^n - U^{n-1}}{\tau_n} \implies \partial_t u^n + \operatorname{div} f(u^n) \approx \frac{u^n - \bar{u}^{n-1}}{\tau_n}.$$

Therefore, the discretization in time of (1.1) reads

$$(3.3) \quad \frac{u^n - \bar{u}^{n-1}}{\tau_n} - \Delta A(u^n) = \bar{g}^n \quad \text{in } \Omega,$$

where $\bar{g}^n(x) = \tau_n^{-1} \int_{t^{n-1}}^{t^n} g(x, t) dt$. Since \bar{u}^{n-1} is well defined only for $\bar{x}^{n-1} \in \bar{\Omega}$, one has to properly extend \bar{u}^{n-1} beyond the inflow boundary according to the boundary condition imposed.

Let \mathcal{M}^n be a regular triangulation of Ω into simplexes. The mesh \mathcal{M}^n is obtained by refinement/coarsening of \mathcal{M}^{n-1} , and thus \mathcal{M}^n and \mathcal{M}^{n-1} are compatible. Two meshes are compatible if one is the local refinement by bisection of the other. For any $K \in \mathcal{M}^n$, h_K stands for its diameter. We also denote \mathcal{B}^n the collection of boundaries e of \mathcal{M}^n in Ω ; h_e stands for the size of $e \in \mathcal{B}^n$.

Let V^n indicate the usual space of continuous piecewise linear finite elements over \mathcal{M}^n and $V_0^n = V^n \cap H_0^1(\Omega)$. Let $U_h^0 \in V^n$ be some discretization of the initial function u_0 so that $\|u_0 - U_h^0\|_{L^1(\Omega)}$ can be arbitrarily small when refining the initial mesh \mathcal{M}_0 .

Discrete problem. Given $U_h^{n-1} \in V^{n-1}$, then \mathcal{M}^{n-1} and τ_{n-1} are modified to get \mathcal{M}^n and τ_n and thereafter $U_h^n \in V_0^n$ computed according to

$$(3.4) \quad \left\langle \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n}, v \right\rangle + \langle \nabla A(U_h^n), \nabla v \rangle = \langle \bar{g}^n, v \rangle \quad \forall v \in V_0^n,$$

where $\bar{U}_h^{n-1} = U_h^{n-1}(\bar{x}^{n-1})$, $\bar{x}^{n-1} = \tilde{X}(t^{n-1})$, and the approximate characteristics $\tilde{X}(t)$ is defined by

$$(3.5) \quad d\tilde{X}/dt = f'(U_h^{n-1}(\tilde{X}(t))), \quad \tilde{X}(t^n) = x.$$

The characteristics $\tilde{X}(t)$ satisfying (3.5) may not be computed exactly and it can be calculated by multistep Euler method or Runge-Kutta method as suggested in [1] or [16]. If the time-step size τ_n is small enough (depending on the boundedness of U_h^{n-1}), then due to (H1) it can be easily proved that the approximate characteristics do not cross each other (cf. e.g. [18]). In this paper, we will not elaborate on this issue and simply assume this to be the case and still denote by $\tilde{X}(t)$ this approximate characteristics. Further details on the application of the method of characteristics to nonlinear convection-diffusion problem can be found in [17, 18].

We also remark that since $A(\cdot)$ is strictly monotone, (3.4) can be solved by nonlinear SOR [8] if appropriate mass lumping is used for computing $\langle U_h^n, v \rangle$ for $v \in V^n$, and the nonlinear relation $A(U_h^n)$ is enforced nodewise, i.e. to replace $\langle \nabla A(U_h^n), \nabla v \rangle$ in (3.4) by $\langle \nabla I^n A(U_h^n), \nabla v \rangle$, where $I^n : C(\Omega) \rightarrow V^n$ is the standard finite element Lagrange interpolant. The a posteriori error analysis below can be easily extended to cover these situations by including appropriate error indicators for quadrature error. To avoid inessential complications, we will not consider the extensions in this paper.

We conclude this section with some notation. Let the jump of $\nabla A(U_h^n)$ across some $e \in \mathcal{B}^n$ be

$$[\nabla A(U_h^n)]_e := (\nabla A(U_h^n)|_{K_1} - \nabla A(U_h^n)|_{K_2}) \cdot \nu_e$$

with the convention the unit normal vector ν_e to e points from K_2 to K_1 and so that the jump $[\nabla A(U_h^n)]_e$ is well-defined. Let U_h denote the piecewise linear extension of $\{U_h^n\}$, that is $U_h(\cdot, 0) = U_h^0(\cdot)$, and for all $t^{n-1} < t \leq t^n$,

$$U_h(\cdot) = \frac{t^n - t}{\tau_n} U_h^{n-1}(\cdot) + \frac{t - t^{n-1}}{\tau_n} U_h^n(\cdot).$$

Finally, we introduce the mesh-dependent norms

$$\|h_n \varphi\|_{L^2(\Omega)} = \left(\sum_{K \in \mathcal{M}^n} \|h_K \varphi\|_{L^2(K)}^2 \right)^{\frac{1}{2}}, \quad \|h_n^{1/2} \varphi\|_{L^2(\Omega)} = \left(\sum_{e \in \mathcal{B}^n} h_e \|\varphi\|_{L^2(e)}^2 \right)^{\frac{1}{2}}.$$

4. ENTROPY ERROR INEQUALITY

We start by introducing some notation. For any $\varepsilon > 0$, let

$$H_\varepsilon(z) = \text{sgn}(z) \min(1, |z|/\varepsilon)$$

be the regularization of the sign function $\text{sgn}(z)$. For any $k \in \mathbf{R}$, define the entropy pair $(U_\varepsilon, F_\varepsilon)$ by

$$U_\varepsilon(z, k) = \int_k^z H_\varepsilon(A(r) - A(k)) dr, \quad F_\varepsilon(z, k) = \int_k^z H_\varepsilon(A(r) - A(k)) f'(r) dr.$$

The following result is well-known (cf. e.g. [4, 22]).

Lemma 4.1. *For any $\phi \in L^2(0, T; H_0^1(\Omega))$ such that $\phi(\cdot, 0) = \phi(\cdot, T) = 0$, and any $k \in \mathbf{R}$, we have*

$$(4.1) \quad - \int_Q U_\varepsilon(u, k) \partial_t \phi - \int_Q F_\varepsilon(u, k) \cdot \nabla \phi + \int_Q H_\varepsilon(A(u) - A(k)) \nabla A(u) \cdot \nabla \phi \\ + \int_Q H'_\varepsilon(A(u) - A(k)) |\nabla A(u)|^2 \phi = \int_Q g H_\varepsilon(A(u) - A(k)) \phi.$$

By letting $\varepsilon \rightarrow 0$ in (4.1) one obtains the entropy condition in the interior of Q (2.2). In this paper, however, we will not use this limit interior entropy condition.

Let $(H^1(\Omega))'$ be the dual space of $H^1(\Omega)$, we define the discrete residual $\mathcal{R} \in L^2(0, T; (H^1(\Omega))')$ through the following relation, for any $\varphi \in H^1(\Omega)$,

$$(4.2) \quad \langle \partial_t U_h, \varphi \rangle - \langle f(U_h), \nabla \varphi \rangle + \langle \nabla A(U_h), \nabla \varphi \rangle = \langle g, \varphi \rangle - \langle \mathcal{R}, \varphi \rangle.$$

Then similar to Lemma 4.1, we have the following result.

Lemma 4.2. *For any $\phi \in L^2(0, T; H_0^1(\Omega))$ such that $\phi(\cdot, 0) = \phi(\cdot, T) = 0$, and any $k' \in \mathbf{R}$, we have*

$$(4.3) \quad - \int_Q U_\varepsilon(U_h, k') \partial_t \phi - \int_Q F_\varepsilon(U_h, k') \cdot \nabla \phi + \int_Q H_\varepsilon(A(U_h) - A(k')) \nabla A(u) \cdot \nabla \phi \\ + \int_Q H'_\varepsilon(A(U_h) - A(k')) |\nabla A(U_h)|^2 \phi \\ = \int_Q g H_\varepsilon(A(U_h) - A(k')) \phi - \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(k')) \phi \rangle.$$

Proof. For the sake of completeness, we sketch the proof here. We take $\varphi = H_\varepsilon(A(U_h) - A(k')) \phi$ in (4.2), integrate in time over $(0, T)$, and rewrite each term as follows. First, by integration by parts, we get

$$\int_0^T \langle \partial_t U_h, H_\varepsilon(A(U_h) - A(k')) \phi \rangle = \int_Q \partial_t U_\varepsilon(U_h, k') \phi = - \int_Q U_\varepsilon(U_h, k') \partial_t \phi.$$

Next, let $\psi_\varepsilon(z, k') = \partial_z [H_\varepsilon(A(z) - A(k'))] = H'_\varepsilon(A(z) - A(k')) A'(z)$, then it is easy to see that

$$F_\varepsilon(z, k') = \int_{k'}^z H_\varepsilon(A(r) - A(k')) f'(r) dr \\ = f(z) H_\varepsilon(A(z) - A(k')) - \int_{k'}^z f(r) \psi_\varepsilon(r, k') dr.$$

Thus, by doing integration by parts, we have

$$- \int_0^T \langle f(U_h), \nabla \varphi \rangle \\ = - \int_Q (\psi_\varepsilon(U_h, k') \nabla U_h \cdot f(U_h) \phi + H_\varepsilon(A(U_h) - A(k')) f(U_h) \cdot \nabla \phi) \\ = - \int_Q F_\varepsilon(U_h, k') \cdot \nabla \phi - \int_Q \operatorname{div} \left(\phi \int_{k'}^{U_h} f(r) \psi_\varepsilon(r, k') dr \right) \\ = - \int_Q F_\varepsilon(U_h, k') \cdot \nabla \phi.$$

The rest of the proof is straightforward and we omit the details. \square

Now we are going to apply the Kruřkov “doubling of variables” technique and will always write $u = u(y, s)$, $U_h = U_h(x, t)$, unless otherwise stated. If necessary, in the following we will write $Q_{(x,t)}$ or $Q_{(y,s)}$ to stress the domain of integration with respect to (x, t) or (y, s) respectively, although $Q \times Q$ will mainly denote the domain of integration with respect to four variables. The following entropy error identity extends similar result in [4].

Lemma 4.3. *Let $\phi = \phi(x, t; y, s)$ be non-negative function such that*

$$\begin{aligned} (x, t) &\mapsto \phi(x, t; y, s) \in C_c^\infty(Q) \text{ for every } (y, s) \in Q, \\ (y, s) &\mapsto \phi(x, t; y, s) \in C_c^\infty(Q) \text{ for every } (x, t) \in Q. \end{aligned}$$

Then we have

$$\begin{aligned} (4.4) \quad & - \int_{Q \times Q} U_\varepsilon(u, U_h) (\partial_t \phi + \partial_s \phi) - \int_{Q \times Q} F_\varepsilon(u, U_h) (\nabla_x \phi + \nabla_y \phi) \\ & + \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \phi + \nabla_y \phi) \\ & + \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot (\nabla_x \phi + \nabla_y \phi) \\ & + \int_{Q \times Q} H'_\varepsilon(A(u) - A(U_h)) |\nabla_x A(U_h) - \nabla_y A(u)|^2 \phi \\ = & - \int_{Q \times Q} \partial_t [U_\varepsilon(U_h, u) - U_\varepsilon(u, U_h)] \phi \\ & - \int_{Q \times Q} \nabla_x [F_\varepsilon(U_h, u) - F_\varepsilon(u, U_h)] \phi \\ & - \int_{Q_{(y,s)}} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \phi \rangle dt. \end{aligned}$$

Proof. Recall that we write $u = u(y, s)$ and so we can take $k = U_h(x, t)$ in (4.1). Similarly, we can take $k' = u(y, s)$ in (4.3). The lemma follows from the following two identities which can be easily proved by integration by parts

$$\begin{aligned} & \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nabla_x \phi \\ & = \int_{Q \times Q} H'_\varepsilon(A(u) - A(U_h)) \nabla_x A(U_h) \cdot \nabla_y A(u) \phi \\ & \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot \nabla_y \phi \\ & = \int_{Q \times Q} H'_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot \nabla_y A(u) \phi. \end{aligned}$$

\square

The next objective is to remove the restriction that the test functions in the entropy error identity (4.4) must have vanishing trace. This is achieved by using the technique of boundary layer sequence introduced in [22]. The properties of the boundary layer sequence are summarized in the following lemma. For a proof, we refer to [22].

Lemma 4.4. For any $\delta > 0$, let ζ_δ be the solution of the elliptic problem

$$-\delta^2 \Delta \zeta_\delta + \zeta_\delta = 1 \quad \text{in } \Omega, \quad \zeta_\delta = 0 \quad \text{on } \partial\Omega.$$

Then we have

$$\lim_{\delta \rightarrow 0} \zeta_\delta = 1 \quad \text{a.e. in } \Omega; \quad 0 \leq \zeta_\delta \leq 1, \quad -\Delta \zeta_\delta \geq 0 \quad \text{in } \Omega.$$

Moreover, for any $v \in L^2(0, T; H^1(\Omega))$, and for any $\xi \in H^1(Q) \cap C(Q)$ such that $\xi(\cdot, 0) = \xi(\cdot, T) = 0$,

$$\lim_{\delta \rightarrow 0} \int_Q (v \cdot \nabla \zeta_\delta) \xi dx = - \int_\Sigma (v \cdot \nu) \xi,$$

where $\Sigma = \partial\Omega \times (0, T)$.

Now we specify the choice of the test function ϕ in the entropy error identity (4.4), which is similar to that used in [22].

Definition 4.5. Let

$$(4.5) \quad \phi(x, t, y, s) = \zeta_\delta(x) \zeta_\eta(y) \xi(x, t, y, s) \theta(t),$$

where $\theta \in C_c^\infty(0, T)$ such that $\theta \geq 0$, and ξ is defined as follows. Let $\{\varphi_j\}_{0 \leq j \leq J}$ be a partition of unity subordinate to open sets B_0, B_1, \dots, B_J such that $\Omega \subset \cup_{j=0}^J B_j$, $B_0 \subset \subset \Omega$ and $\partial\Omega \subset \cup_{j=1}^J B_j$. Let $\hat{\varphi}_j \in C_c^\infty(\mathbf{R}^d)$, $0 \leq \hat{\varphi}_j \leq 1$, such that $\text{supp}(\hat{\varphi}_j) \subset B_j$ and $\hat{\varphi}_j(x) = 1$ on the support of φ_j so that $\varphi_j(x) \hat{\varphi}_j(x) = \varphi_j(x)$. We use φ_j as a function of y and $\hat{\varphi}_j$ as a function of x , and denote $\hat{\varphi}_j(x) \varphi_j(y) = \psi_j(x, y)$. Define

$$(4.6) \quad \xi(x, t, y, s) = \sum_{j=0}^J \omega_l(t-s) \omega_m(x' - y') \omega_n(x_d - y_d) \psi_j(x, y),$$

where ω_l, ω_n are sequences of symmetric mollifiers in \mathbf{R} , ω_m is a sequence of symmetric mollifier in \mathbf{R}^{d-1} , and for $j = 1, 2, \dots, J$, $x = (x', x_d), y = (y', y_d)$ are local coordinates induced by $\psi_j(x, y)$ in B_j , that is, $B_j \cap \partial\Omega = \{x \in B_j : x_d = \rho_j(x')\}$, $B \cap \Omega = \{x \in B_j : x_d < \rho_j(x')\}$ for some Lipschitz continuous function $\rho_j : \mathbf{R}^{d-1} \rightarrow \mathbf{R}$.

The following theorem is the main result of this section.

Theorem 4.6. *Let θ and ξ be defined in Definition 4.5. Then we have the following entropy error inequality*

$$\begin{aligned}
(4.7) \quad & - \int_{Q \times Q} U_\varepsilon(u, U_h) \xi \theta_t - \int_{Q \times Q} K_\varepsilon(u, U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \theta \\
& + \int_{Q \times Q} H'_\varepsilon(A(u) - A(U_h)) |\nabla_x A(U_h) - \nabla_y A(u)|^2 \xi \theta \\
\leq & - \int_{Q \times Q} \partial_t [U_\varepsilon(U_h, u) - U_\varepsilon(u, U_h)] \xi \theta \\
& - \int_{Q \times Q} \nabla_x [F_\varepsilon(u, U_h) - F_\varepsilon(u, U_h)] \xi \theta \\
& - \int_{Q(y, s)} \int_{\Sigma(x, t)} \left(F_\varepsilon(u, U_h) - H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \right) \cdot \nu_x \xi \theta \\
& - \int_{Q(x, t)} \int_{\Sigma(y, s)} \left(F_\varepsilon(u, U_h) - H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \right) \cdot \nu_y \xi \theta \\
& - \int_{Q(y, s)} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \xi \theta \rangle dt,
\end{aligned}$$

where $K_\varepsilon(u, U_h) = F_\varepsilon(u, U_h) - H_\varepsilon(A(u) - A(U_h)) (\nabla_y A(u) - \nabla_x A(U_h))$, $\Sigma = \partial\Omega \times (0, T)$, and $\Sigma_{(x, t)}$ or $\Sigma_{(y, s)}$ are the domain of integration of Σ with respect to (x, t) or (y, s) respectively.

The proof of the theorem depends on the following lemmas.

Lemma 4.7. *We have*

$$(4.8) \quad \lim_{\delta, \eta \rightarrow 0} - \int_{Q \times Q} U_\varepsilon(u, U_h) (\partial_t \phi + \partial_s \phi) = - \int_{Q \times Q} U_\varepsilon(u, U_h) \xi \theta_t,$$

and

$$\begin{aligned}
\lim_{\delta, \eta \rightarrow 0} - \int_{Q \times Q} F_\varepsilon(u, U_h) (\nabla_x \phi + \nabla_y \phi) &= - \int_{Q \times Q} F_\varepsilon(u, U_h) (\nabla_x \xi + \nabla_y \xi) \theta \\
&+ \int_{Q(y, s)} \int_{\Sigma(x, t)} F_\varepsilon(u, U_h) \cdot \nu_x \xi \theta \\
&+ \int_{Q(x, t)} \int_{\Sigma(y, s)} F_\varepsilon(u, U_h) \cdot \nu_y \xi \theta
\end{aligned}$$

Proof. By the definition of ϕ in (4.5) and ξ in (4.6), we know that $\partial_t \phi + \partial_s \phi = \zeta_\delta \zeta_\eta \xi \theta_t$. Thus

$$\int_{Q \times Q} U_\varepsilon(u, U_h) (\partial_t \phi + \partial_s \phi) = \int_{Q \times Q} U_\varepsilon(u, U_h) \zeta_\delta \zeta_\eta \xi \theta_t.$$

Then (4.8) follows by letting $\delta, \eta \rightarrow 0$ in above equality and using Lebesgue dominated convergence theorem.

Next, we note that

$$(4.9) \quad \nabla_x \phi + \nabla_y \phi = \zeta_\delta \zeta_\eta (\nabla_x \xi + \nabla_y \xi) \theta + (\zeta_\eta \nabla_x \zeta_\delta + \zeta_\delta \nabla_y \zeta_\eta) \xi \theta.$$

Thus

$$\begin{aligned} \int_{Q \times Q} F_\varepsilon(u, U_h)(\nabla_x \phi + \nabla_y \phi) &= \int_{Q \times Q} F_\varepsilon(u, U_h) \zeta_\delta \zeta_\eta (\nabla_x \xi + \nabla_y \xi) \theta \\ &\quad + \int_{Q \times Q} F_\varepsilon(u, U_h) (\zeta_\eta \nabla_x \zeta_\delta + \zeta_\delta \nabla_y \zeta_\eta) \xi \theta. \end{aligned}$$

Now we let $\delta, \eta \rightarrow 0$. The first term can be treated by using Lebesgue dominated convergence theorem and the second term can be treated by Lemma 4.4 because of $F_\varepsilon(u, U_h) \in L^2(0, T; H^1(\Omega))$. This proves (4.9). \square

Lemma 4.8. *We have*

$$\begin{aligned} &\liminf_{\delta, \eta \rightarrow 0} \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \phi + \nabla_y \phi) \\ &\geq \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \xi + \nabla_y \xi) \theta \\ &\quad - \int_{Q(y, s)} \int_{\Sigma(x, t)} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nu_x \xi \theta. \end{aligned}$$

Proof. By (4.9) we have

$$\begin{aligned} (4.10) \quad &\int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \phi + \nabla_y \phi) \\ &= \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \xi + \nabla_y \xi) \zeta_\delta \zeta_\eta \theta \\ &\quad + \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nabla_x \zeta_\delta \zeta_\eta \xi \theta \\ &\quad + \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta. \end{aligned}$$

By Lebesgue dominated convergence theorem, we get

$$\begin{aligned} &\lim_{\delta, \eta \rightarrow 0} \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \xi + \nabla_y \xi) \zeta_\delta \zeta_\eta \theta \\ &= \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot (\nabla_x \xi + \nabla_y \xi) \theta. \end{aligned}$$

By Lemma 4.4 and Lebesgue dominated convergence theorem, we have

$$\begin{aligned} &\lim_{\delta, \eta \rightarrow 0} \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nabla_x \zeta_\delta \zeta_\eta \xi \theta \\ &= - \int_{Q(y, s)} \int_{\Sigma(x, t)} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nu_x \xi \theta. \end{aligned}$$

To deal with the third term on the right-hand of (4.10), we define

$$\Phi_\varepsilon(z) = \int_0^z H_\varepsilon(r) dr, \quad \forall z \in \mathbf{R}.$$

It is easy to see that $\Phi_\varepsilon(z) = \Phi_\varepsilon(-z)$ and $\Phi_\varepsilon(z_1 + z_2) \leq \Phi_\varepsilon(z_1) + \Phi_\varepsilon(z_2)$. Thus

$$\Psi_\varepsilon(u, U_h) = \Phi_\varepsilon(A(u) - A(U_h)) + \Phi_\varepsilon(A(u)) - \Phi_\varepsilon(A(U_h)) \geq 0 \quad \text{a.e. in } Q \times Q.$$

Note that

$$\begin{aligned}
& \int_{Q \times Q} H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta \\
= & \int_{Q \times Q} \nabla_y \Phi_\varepsilon(A(u) - A(U_h)) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta \\
= & \int_{Q \times Q} \nabla_y [\Phi_\varepsilon(A(u) - A(U_h)) - \Phi_\varepsilon(A(U_h))] \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta \\
= & \int_{Q \times Q} \nabla_y \Psi_\varepsilon(u, U_h) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta - \int_{Q \times Q} \nabla_y \Phi_\varepsilon(A(u)) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta \\
=: & L_1 + L_2.
\end{aligned}$$

Since $\Psi_\varepsilon(u, U_h) \geq 0$, $-\Delta_y \zeta_\eta \geq 0$, by integrating by parts we get

$$\begin{aligned}
L_1 &= - \int_{Q \times Q} \Psi_\varepsilon(u, U_h) \Delta_y \zeta_\eta \zeta_\delta \xi \theta - \int_{Q \times Q} \Psi_\varepsilon(u, U_h) \nabla_y \zeta_\eta \cdot \nabla_y \xi \zeta_\delta \theta \\
&\geq - \int_{Q \times Q} \Psi_\varepsilon(u, U_h) \nabla_y \zeta_\eta \cdot \nabla_y \xi \zeta_\delta \theta.
\end{aligned}$$

Since $\Psi_\varepsilon(u, U_h)|_{\Sigma_{(y,s)}} = 0$, we deduce by using Lemma 4.4 that

$$\liminf_{\delta, \eta \rightarrow 0} L_1 \geq 0.$$

Notice that, by using Lebesgue dominated convergence theorem, we have

$$\begin{aligned}
\lim_{\delta \rightarrow 0} L_2 &= - \int_{Q \times Q} H_\varepsilon(A(u)) \nabla_y A(u) \cdot \nabla_y \zeta_\eta \xi \theta \\
&= \int_{Q \times Q} H_\varepsilon(A(u)) \nabla_y A(u) \cdot \nabla_y [(1 - \zeta_\eta) \xi \theta] \\
&\quad - \int_{Q \times Q} H_\varepsilon(A(u)) \nabla_y A(u) \cdot \nabla_y (\xi \theta) (1 - \zeta_\eta) \\
=: & L_{21} + L_{22}.
\end{aligned}$$

Since $\zeta_\eta \rightarrow 1$ a.e. in Ω , we obtain

$$\lim_{\eta \rightarrow 0} L_{22} = 0.$$

To deal with L_{21} , we denote by

$$\Theta(y, s) = (1 - \zeta_\eta(y)) \int_{Q_{(x,t)}} \xi(x, t, y, s) \theta(t) dx dt.$$

Since $\theta \in C_c^\infty(0, T)$, for sufficiently large l , we may assume that $\Theta(\cdot, 0) = \Theta(\cdot, T) = 0$. Thus we can take $\varphi = H_\varepsilon(A(u)) \Theta$ as the test function in (2.4). Using similar argument leading to (4.1), we can show that

$$\begin{aligned}
(4.11) \quad & - \int_Q U_\varepsilon(u, 0) \partial_s \Theta - \int_Q F_\varepsilon(u, 0) \cdot \nabla_y \Theta + \int_Q H_\varepsilon(A(u)) \nabla_y A(u) \cdot \nabla_y \Theta \\
& + \int_Q H'_\varepsilon(A(u)) |\nabla_y A(u)|^2 \Theta = \int_Q g H_\varepsilon(A(u)) \Theta.
\end{aligned}$$

Therefore

$$\begin{aligned} L_{21} &= \int_Q g H_\varepsilon(A(u)) \Theta + \int_Q U_\varepsilon(u, 0) \partial_s \Theta - \int_Q \operatorname{div}_y F_\varepsilon(u, 0) \Theta \\ &\quad - \int_Q H'_\varepsilon(A(u)) |\nabla_y A(u)|^2 \Theta, \end{aligned}$$

which tends to zero as $\eta \rightarrow 0$ by using the fact that $\Theta, \partial_s \Theta \rightarrow 0$ as $\eta \rightarrow 0$. This proves that $\lim_{\eta \rightarrow 0} \lim_{\delta \rightarrow 0} L_2 = 0$. Similarly, we can show that $\lim_{\delta \rightarrow 0} \lim_{\eta \rightarrow 0} L_2 = 0$. This completes the proof. \square

Now we are ready to prove Theorem 4.6.

Proof of Theorem 4.6. The proof lies in taking the test function ϕ in the entropy error identity (4.4) as in Definition 4.5, and then take the limit $\delta, \eta \rightarrow 0$. By Lemmas 4.7 and 4.8 and Lebesgue dominated convergence theorem, we only remain to consider the limit of the following quantity

$$\begin{aligned} L_3 &= \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot (\nabla_x \phi + \nabla_y \phi) \\ &\quad + \int_{Q(y,s)} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \phi \rangle dt. \end{aligned}$$

From (4.9) and (4.2), it is easy to check that

$$\begin{aligned} L_3 &= \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \zeta_\delta \zeta_\eta \theta \\ &\quad + \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot \nabla_y \zeta_\eta \zeta_\delta \xi \theta \\ &\quad + \int_{Q(y,s)} \int_0^T \langle g - \partial_t U_h, H_\varepsilon(A(U_h) - A(u)) \zeta_\delta \zeta_\eta \xi \theta \rangle dt \\ &\quad + \int_{Q(y,s)} \int_0^T \langle f(U_h), \nabla_x (H_\varepsilon(A(U_h) - A(u)) \xi \theta) \zeta_\delta \zeta_\eta \rangle dt \\ &\quad + \int_{Q(y,s)} \int_0^T \langle f(U_h), \nabla_x \zeta_\delta \zeta_\eta H_\varepsilon(A(U_h) - A(u)) \xi \theta \rangle dt \\ &\quad - \int_{Q(y,s)} \int_0^T \langle \nabla_x A(U_h), \nabla_x (H_\varepsilon(A(U_h) - A(u)) \xi \theta) \zeta_\delta \zeta_\eta \rangle dt \\ &=: L_{31} + \dots + L_{36}. \end{aligned}$$

By Lebesgue dominated convergence theorem, we have

$$\lim_{\delta, \eta \rightarrow 0} L_{31} = \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \theta,$$

and

$$\lim_{\delta, \eta \rightarrow 0} (L_{33} + L_{34} + L_{36}) = \int_{Q(y,s)} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \xi \theta \rangle dt.$$

By Lemma 4.4, we get

$$\lim_{\delta, \eta \rightarrow 0} L_{32} = - \int_{Q(x,t)} \int_{\Sigma(y,s)} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot \nu_y \xi \theta.$$

Moreover, since $f(U_h) = 0$ on $\Sigma_{(x,t)}$, we have

$$\lim_{\delta, \eta \rightarrow 0} L_{35} = 0.$$

Therefore we have

$$\begin{aligned} \lim_{\delta, \eta \rightarrow 0} L_3 &= \int_{Q \times Q} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \theta \\ &\quad - \int_{Q_{(x,t)}} \int_{\Sigma_{(y,s)}} H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h) \cdot \nu_y \xi \theta \\ &\quad + \int_{Q_{(y,s)}} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \xi \theta \rangle dt. \end{aligned}$$

This completes the proof. \square

5. A POSTERIORI ERROR ANALYSIS

We start by prove the following elementary estimate which extends the result in [12, Corollary 6.4].

Lemma 5.1. *For any $\varepsilon > 0$ and $z \in \mathbf{R}$, define*

$$\nu(\varepsilon, z) = \min\{A'(s) : |A(s) - A(z)| \leq \varepsilon\}.$$

Then for any $k \in \mathbf{R}$ and $z \in \mathbf{R}$, we have

$$\begin{aligned} |\partial_z [U_\varepsilon(z, k) - U_\varepsilon(k, z)]| &\leq \frac{\varepsilon^\gamma}{\nu(\varepsilon, z)} \mathcal{K}_1, \\ |\partial_z [F_\varepsilon(z, k) - F_\varepsilon(k, z)]| &\leq \frac{\varepsilon^\gamma}{\nu(\varepsilon, z)} \mathcal{K}_2, \end{aligned}$$

where $0 < \gamma \leq 1$ is the Hölder exponent of $A' \circ A^{-1}$ in (H2), $\mathcal{K}_1 = \mathbf{H}(A' \circ A^{-1})$, $\mathcal{K}_2 = \mathcal{K}_1 \|f'\|_{L^\infty(\mathbf{R})} + \mathbf{L}(f')$ with $\mathbf{H}(A' \circ A^{-1})$ and $\mathbf{L}(f')$ being the Hölder constant of $A' \circ A^{-1}$ and the Lipschitz constant of f' respectively.

Proof. We only prove the estimate for F_ε . The estimate for U_ε is similar. By definition,

$$\begin{aligned} &\partial_z [F_\varepsilon(z, k) - F_\varepsilon(k, z)] \\ &= \partial_z \int_k^z \left[H_\varepsilon(A(r) - A(k)) + H_\varepsilon(A(r) - A(z)) \right] f'(r) dr \\ &= H_\varepsilon(A(z) - A(k)) f'(z) - \int_k^z H'_\varepsilon(A(r) - A(z)) A'(z) f'(r) dr \\ &= \int_k^z H'_\varepsilon(A(r) - A(z)) A'(r) f'(z) dr - \int_k^z H'_\varepsilon(A(r) - A(z)) A'(z) f'(r) dr \\ &= \int_k^z H'_\varepsilon(A(r) - A(z)) A'(r) (f'(z) - f'(r)) dr \\ &\quad + \int_k^z H'_\varepsilon(A(r) - A(z)) (A'(r) - A'(z)) f'(r) dr. \end{aligned}$$

Without loss of generality, we may assume $z > k$. Then since $H'_\varepsilon(A(r) - A(z))$ vanishes outside the set $\{A'(s) : |A(s) - A(z)| \leq \varepsilon\}$, we know that

$$\begin{aligned} & \left| \int_k^z H'_\varepsilon(A(r) - A(z)) A'(r) (f'(z) - f'(r)) dr \right| \\ & \leq \frac{1}{\varepsilon} \int_{A^{-1}(A(z)-\varepsilon)}^z A'(r) |f'(r) - f'(z)| dr \\ & \leq \frac{1}{\varepsilon} \mathbf{L}(f') |z - A^{-1}(A(z) - \varepsilon)| \int_{A^{-1}(A(z)-\varepsilon)}^z A'(r) dr \\ & \leq \mathbf{L}(f') \frac{\varepsilon}{\nu(\varepsilon, z)}, \end{aligned}$$

where we have used the fact that $z = A^{-1}(A(z))$. Moreover,

$$\begin{aligned} & \left| \int_k^z H'_\varepsilon(A(r) - A(z)) (A'(r) - A'(z)) f'(r) dr \right| \\ & \leq \mathbf{H}(A' \circ A^{-1}) \varepsilon^{\gamma-1} \int_{A^{-1}(A(z)-\varepsilon)}^z |f'(r)| dr \\ & \leq \mathbf{H}(A' \circ A^{-1}) \varepsilon^{\gamma-1} \|f'\|_{L^\infty(\mathbf{R})} |z - A^{-1}(A(z) - \varepsilon)| \\ & \leq \mathbf{H}(A' \circ A^{-1}) \|f'\|_{L^\infty(\mathbf{R})} \frac{\varepsilon^\gamma}{\nu(\varepsilon, z)}. \end{aligned}$$

This completes the proof. \square

The next step is to let the parameters in the mollifier functions $l, m, n \rightarrow \infty$ in the entropy error inequality (4.7) to complete the Kruřkov “doubling of variables” technique.

Lemma 5.2. *We have*

$$(5.1) \quad \lim_{l, m, n \rightarrow \infty} \int_{Q \times Q} U_\varepsilon(u, U_h) \xi \theta_t = \int_Q U_\varepsilon(u, U_h) \theta_t,$$

$$(5.2) \quad \lim_{l, m, n \rightarrow \infty} \int_{Q \times Q} K_\varepsilon(u, U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \theta = 0,$$

$$(5.3) \quad \begin{aligned} \lim_{l, m, n \rightarrow \infty} \int_{Q \times Q} H'_\varepsilon(A(u) - A(U_h)) |\nabla_x A(U_h) - \nabla_y A(u)|^2 \xi \theta \\ = \int_Q H'_\varepsilon(A(u) - A(U_h)) |\nabla A(U_h) - \nabla A(u)|^2 \theta. \end{aligned}$$

Proof. From the definition of $\xi(x, t, y, s)$ in (4.6) and the property of mollifier functions, we have

$$\lim_{l, m, n \rightarrow \infty} \int_{Q \times Q} U_\varepsilon(u, U_h) \xi \theta_t = \sum_{j=0}^J \int_Q U_\varepsilon(u, U_h) \psi_j(x, x) \theta_t = \int_Q U_\varepsilon(u, U_h) \theta_t,$$

where we have used the fact that

$$(5.4) \quad \sum_{j=0}^J \psi_j(x, x) = \sum_{j=0}^J \varphi_j(x) \hat{\varphi}_j(x) = \sum_{j=0}^J \varphi_j(x) = 1.$$

This proves (5.1). Similarly, we can show (5.3). To see (5.2), we note that

$$\nabla_x \xi + \nabla_y \xi = \sum_{j=0}^J \omega_l(t-s) \omega_m(x' - y') \omega_n(x_d - y_d) (\nabla_x \psi_j + \nabla_y \psi_j).$$

Thus

$$\lim_{l,m,n \rightarrow \infty} \int_{Q \times Q} K_\varepsilon(u, U_h) \cdot (\nabla_x \xi + \nabla_y \xi) \theta = \sum_{j=0}^J \int_Q K_\varepsilon(u, U_h) \nabla_x \psi_j(x, x) \theta,$$

which vanishes due to (5.4). \square

Lemma 5.3. *We have*

$$(5.5) \quad \overline{\lim}_{l,m,n \rightarrow \infty} \left| \int_{Q \times Q} \partial_t (U_\varepsilon(U_h, u) - U_\varepsilon(u, U_h)) \xi \theta \right| \leq \mathcal{K}_1 \varepsilon^\gamma \int_Q \frac{1}{\nu(\varepsilon, U_h)} |\partial_t U_h| \theta,$$

$$(5.6) \quad \overline{\lim}_{l,m,n \rightarrow \infty} \left| \int_{Q \times Q} \nabla_x (F_\varepsilon(U_h, u) - F_\varepsilon(u, U_h)) \xi \theta \right| \leq \mathcal{K}_2 \varepsilon^\gamma \int_Q \frac{1}{\nu(\varepsilon, U_h)} |\nabla_x U_h| \theta.$$

Proof. By Lemma 5.1, we have

$$\begin{aligned} & \left| \int_{Q \times Q} \partial_t (U_\varepsilon(U_h, u) - U_\varepsilon(u, U_h)) \xi \theta \right| \\ & \leq \mathcal{K}_1 \varepsilon^\gamma \int_{Q(x,t)} \frac{1}{\nu(\varepsilon, U_h)} |\partial_t U_h| \theta \cdot \int_{Q(y,s)} \xi dy ds. \end{aligned}$$

From the definition of ξ in (4.6) and (5.4), it is easy to see that

$$\lim_{l,m,n \rightarrow \infty} \int_{Q(y,s)} \xi(x, t, y, s) = 1.$$

This proves the estimate (5.5). Similarly we can show (5.6). \square

Lemma 5.4. *We have*

$$(5.7) \quad \lim_{n \rightarrow \infty} \lim_{l,m \rightarrow \infty} \int_{Q(y,s)} \int_{\Sigma(x,t)} (F_\varepsilon(u, U_h) - H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u)) \cdot \nu_x \xi \theta = 0,$$

$$(5.8) \quad \lim_{n \rightarrow \infty} \lim_{l,m \rightarrow \infty} \int_{Q(x,t)} \int_{\Sigma(y,s)} (F_\varepsilon(u, U_h) - H_\varepsilon(A(U_h) - A(u)) \nabla_x A(U_h)) \cdot \nu_y \xi \theta = 0.$$

Proof. We modify the idea in [22, §3] to show (5.7). Since $U_h = 0, A(U_h) = 0$ on $\Sigma(x,t)$, defining $N_j(y') = (-\nabla \rho_j(y'), 1)$ (see Definition 4.5 for the notation) we get

$$\begin{aligned} & \lim_{l,m \rightarrow \infty} \int_{Q(y,s)} \int_{\Sigma(x,t)} (F_\varepsilon(u, U_h) - H_\varepsilon(A(u) - A(U_h)) \nabla_y A(u)) \cdot \nu_x \xi \theta \\ & = \sum_{j=1}^J \int_{Q(y,s)} (F_\varepsilon(u, 0) - H_\varepsilon(A(u)) \nabla_y A(u)) \cdot N_j(y') \omega_n(\rho_j(y') - y_d) \psi_j \theta \\ & = : L_4, \end{aligned}$$

where $\psi_j(y) = \hat{\varphi}_j(y', \rho_j(y'))\varphi_j(y', y_d)$. Let $w_n(y) = 2 \int_{y_d - \rho_j(y')}^0 \omega_n(s) ds$, then $\nabla w_n = -2\omega_n(\rho_j(y') - y_d)N_j(y')$. Thus

$$\begin{aligned} L_4 &= -\frac{1}{2} \sum_{j=1}^J \int_{Q_{(y,s)}} (F_\varepsilon(u, 0) - H_\varepsilon(A(u))\nabla_y A(u)) \cdot \nabla w_n \psi_j \theta \\ &= \frac{1}{2} \sum_{j=1}^J \int_{Q_{(y,s)}} (F_\varepsilon(u, 0) - H_\varepsilon(A(u))\nabla_y A(u)) \cdot \nabla [(1 - w_n)\psi_j \theta] \\ &\quad - \frac{1}{2} \sum_{j=1}^J \int_{Q_{(y,s)}} (F_\varepsilon(u, 0) - H_\varepsilon(A(u))\nabla_y A(u)) \cdot \nabla (\psi_j \theta)(1 - w_n) \\ &=: L_{41} + L_{42}. \end{aligned}$$

Notice that $w_n \rightarrow 1$ a.e. in Ω , by Lebesgue dominated convergence theorem, we have $\lim_{n \rightarrow \infty} L_{42} = 0$. Moreover, by (4.11) and using the argument in dealing with the limit L_{21} in the proof of Lemma 4.8, we can show that $\lim_{n \rightarrow \infty} L_{41} = 0$. Therefore, L_4 tends to 0 as $n \rightarrow \infty$. This proves (5.7).

The proof of (5.8) is simpler. Since U_h is a finite element function, the trace of $F_\varepsilon(U_h, 0) + F_\varepsilon(A(U_h)) \cdot \nabla_x A(U_h)$ on $\Sigma_{(x,t)}$ is well-defined and is equal to 0. One can easily prove the integral in (5.8) converges to zero as $l, m, n \rightarrow \infty$. \square

Lemma 5.5. *Let θ be defined in Definition 4.5, then we have*

$$\begin{aligned} (5.9) \quad & - \int_Q U_\varepsilon(u, U_h)\theta_t + \int_Q H'_\varepsilon(A(u) - A(U_h))|\nabla(A(U_h) - A(u))|^2 \theta \\ & \leq \mathcal{K}\varepsilon^\gamma \int_Q \frac{1}{\nu(\varepsilon, U_h)} (|\partial_t U_h| + |\nabla_x U_h|)\theta - \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u))\theta \rangle dt. \end{aligned}$$

where $\mathcal{K} = \max(\mathcal{K}_1, \mathcal{K}_2)$.

Proof. We let first $l, m \rightarrow \infty$ then $n \rightarrow \infty$ in the entropy error inequality (4.7). By Lemmas 5.2-5.4, we are remained to consider the limit of

$$\begin{aligned} L_5 : &= - \int_{Q_{(y,s)}} \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u))\xi \theta \rangle dt \\ &= - \int_{Q \times Q} (g - \partial_t U_h) H_\varepsilon(A(U_h) - A(u)) \xi \theta \\ &\quad - \int_{Q \times Q} (f(U_h) - \nabla_x A(U_h)) \cdot \nabla_x (H_\varepsilon(A(U_h) - A(u))) \xi \theta \\ &\quad - \int_{Q \times Q} (f(U_h) - \nabla_x A(U_h)) \cdot \nabla_x \xi H_\varepsilon(A(U_h) - A(u)) \theta. \end{aligned}$$

Notice that, by integration by parts, we have

$$\begin{aligned} & \int_{Q \times Q} (f(U_h) - \nabla_x A(U_h)) \cdot \nabla_y \xi H_\varepsilon(A(U_h) - A(u)) \theta \\ &= - \int_{Q \times Q} (f(U_h) - \nabla_x A(U_h)) \cdot \nabla_y (H_\varepsilon(A(U_h) - A(u))) \xi \theta \\ &\quad + \int_{Q_{(x,t)}} \int_{\Sigma_{(y,s)}} H_\varepsilon(A(U_h)) (f(U_h) - \nabla_x A(U_h)) \cdot \nu_y \xi \theta. \end{aligned}$$

Thus

$$\begin{aligned}
L_5 &= - \int_{Q \times Q} (g - \partial_t U_h) H_\varepsilon(A(U_h) - A(u)) \xi \theta \\
&\quad - \int_{Q \times Q} H'_\varepsilon(A(U_h) - A(u)) (f(U_h) - \nabla_x A(U_h)) \cdot (\nabla_x A(U_h) - \nabla_y A(u)) \xi \theta \\
&\quad - \int_{Q \times Q} (f(U_h) - \nabla_x A(U_h)) \cdot (\nabla_x \xi + \nabla_y \xi) H_\varepsilon(A(U_h) - A(u)) \theta \\
&\quad + \int_{Q(x,t)} \int_{\Sigma(y,s)} H_\varepsilon(A(U_h)) (f(U_h) - \nabla_x A(U_h)) \cdot \nu_y \xi \theta \\
&=: L_{51} + \cdots + L_{54}.
\end{aligned}$$

Similar to the proof of (5.1) in Lemma 5.2, it is easy to see that

$$\lim_{l,m,n \rightarrow \infty} (L_{51} + L_{52}) = - \int_0^T \langle \mathcal{R}, H_\varepsilon(A(U_h) - A(u)) \theta \rangle dt.$$

Similar to the proof of (5.2) in Lemma 5.2, we know that $\lim_{l,m,n \rightarrow \infty} L_{53} = 0$. Finally, since $H_\varepsilon(A(U_h)) = 0$ on $\Sigma(x,t)$, we can easily prove that $L_{54} \rightarrow 0$ as $l, m, n \rightarrow \infty$. This completes the proof. \square

To proceed, we introduce the interior residual

$$R^n := \bar{g}^n - \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n} + \Delta A(U_h^n) \quad \text{on any } K \in \mathcal{M}^n,$$

where we recall that $\bar{g}^n = \tau_n^{-1} \int_{t^{n-1}}^{t^n} g(x, t) dt$.

The following theorem is the main result of this paper.

Theorem 5.6. *Let the assumptions (H1)-(H3) be satisfied. For $n \geq 1$, let $\varepsilon_n = (\sum_{i=1}^3 \mathcal{E}_i^n)^{\frac{2}{1+\gamma}}$, where γ is the Hölder exponent of $A' \circ A^{-1}$, and $\mathcal{E}_1^n, \mathcal{E}_2^n, \mathcal{E}_3^n$ are the error indicators defined below. Denote $Q_n = \Omega \times (t^{n-1}, t^n)$, and define*

$$(5.10) \quad \Lambda_n = \max \left(1, \int_{Q_n} \frac{1}{\nu(\varepsilon_n, U_h)} (|\partial_t U_h| + |\nabla U_h|) + \int_{\Omega} \frac{1}{\nu(\varepsilon_n, U_h^n)} \right),$$

where for any $z \in \mathbf{R}$, $\nu(\varepsilon_n, z) = \min\{A'(s) : |A(s) - A(z)| \leq \varepsilon_n\}$. Then there exists a constant C depending only on the minimum angles of the meshes \mathcal{M}^n , $n = 1, \dots, m$, such that the following a posteriori error estimate is valid

$$\|u^m - U_h^m\|_{L^1(\Omega)} \leq \mathcal{E}_0 + \sum_{n=1}^m (\mathcal{E}_4^n + \mathcal{E}_5^n) + C \sum_{n=1}^m \Lambda_n^{\frac{1}{1+\gamma}} \left(\sum_{i=1}^3 \mathcal{E}_i^n \right)^{\frac{2\gamma}{1+\gamma}},$$

where the error indicators $\mathcal{E}_0, \mathcal{E}_i^n, i = 1, \dots, 5$, are defined by

$$\begin{aligned}
\mathcal{E}_0 &= \|u_0 - U_h^0\|_{L^1(\Omega)} && \text{initial error} \\
\mathcal{E}_1^n &= \tau_n^{1/2} \|h_n^{1/2} [\nabla A(U_h^n)]\|_{L^2(\Omega)} && \text{jump residual} \\
\mathcal{E}_2^n &= \tau_n^{1/2} \|h_n R^n\|_{L^2(\Omega)} && \text{interior residual} \\
\mathcal{E}_3^n &= \tau_n^{1/2} \|\nabla(A(U_h^n) - A(U_h^{n-1}))\|_{L^2(\Omega)} && \text{time residual} \\
\mathcal{E}_4^n &= \int_{t^{n-1}}^{t^n} \left\| \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n} - (\partial_t U_h + \operatorname{div} f(U_h)) \right\|_{L^1(\Omega)} dt && \text{characteristics} \\
\mathcal{E}_5^n &= \int_{t^{n-1}}^{t^n} \|g - \bar{g}^n\|_{L^1(\Omega)} dt && \text{source.}
\end{aligned}$$

Proof. In the proof we will make use of the Clément interpolation operator $\Pi^n : H_0^1(\Omega) \rightarrow V_0^n$, which satisfies the following local approximation properties [9], for any $\varphi \in H_0^1(\Omega)$,

$$(5.11) \quad \|\varphi - \Pi^n \varphi\|_{L^2(K)} + h_K \|\nabla(\varphi - \Pi^n \varphi)\|_{L^2(K)} \leq C^* h_K \|\nabla \varphi\|_{L^2(N(K))},$$

$$(5.12) \quad \|\varphi - \Pi^n \varphi\|_{L^2(e)} \leq C^* h_e^{1/2} \|\nabla \varphi\|_{L^2(N(e))},$$

where $N(A)$ is the union of all elements in \mathcal{M}^n surrounding the sets $A = K \in \mathcal{M}^n$ or $A = e \in \mathcal{B}^n$. The constant C^* depends only on the minimum angle of the mesh \mathcal{M}^n .

Denote $\zeta = H_\varepsilon(A(U_h) - A(u))\theta$. Then by (4.2) and (3.4), we know that, for $t \in (t^{n-1}, t^n]$,

$$\begin{aligned}
\langle \mathcal{R}, \zeta \rangle &= \langle g - \bar{g}^n, \zeta \rangle + \langle R^n - \Delta A(U_h^n), \zeta - \Pi^n \zeta \rangle - \langle \nabla A(U_h^n), \nabla(\zeta - \Pi^n \zeta) \rangle \\
&\quad + \left\langle \frac{U_h^{n-1} - \bar{U}_h^{n-1}}{\tau_n} - \operatorname{div} f(U_h), \zeta \right\rangle - \langle \nabla(A(U_h) - A(U_h^n)), \nabla \zeta \rangle,
\end{aligned}$$

where $\Delta A(U_h^n)$ is understood in elementwise sense. Thus, after integrating by parts, we get

$$\begin{aligned}
& - \int_0^T \langle \mathcal{R}, H_\varepsilon(A(u) - A(U_h))\theta \rangle dt \\
= & - \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \langle g - \bar{g}^n, \zeta \rangle - \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \langle R^n, \zeta - \Pi^n \zeta \rangle \\
& + \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \sum_{e \in \mathcal{B}^n} \int_e [\nabla A(U_h^n)]_e (\zeta - \Pi^n \zeta) \\
& - \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \left\langle \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n} - (\partial_t U_h + \operatorname{div} f(U_h)), \zeta \right\rangle \\
& + \sum_{n=1}^N \int_{t^{n-1}}^{t^n} \langle \nabla(A(U_h) - A(U_h^n)), \nabla \zeta \rangle dt \\
= &: \text{I} + \dots + \text{V}.
\end{aligned}$$

Now we assume $0 \leq \theta \leq 1$ and θ is supported in (t^{n-1}, t^n) . Since $H_\varepsilon(A(u) - A(U_h)) \leq 1$, we have

$$\text{I} + \text{IV} \leq \mathcal{E}_4^n + \mathcal{E}_5^n$$

By (5.11)-(5.12), we have

$$\text{II} + \text{III} \leq C \int_{t^{n-1}}^{t^n} \left(\|h_n R^n\|_{L^2(\Omega)}^2 + \|h_n^{1/2} [\nabla A(U_h^n)]\|_{L^2(\Omega)}^2 \right)^{1/2} \|\nabla \zeta\|_{L^2(\Omega)}.$$

Recall that $\nabla \zeta = H'_\varepsilon(A(U_h) - A(u)) \nabla(A(U_h) - A(u)) \theta$. By Young's inequality,

$$\text{II} + \text{III} \leq \frac{1}{4} \int_{Q_n} H'_\varepsilon(A(u) - A(U_h)) |\nabla(A(u) - A(U_h))|^2 \theta + C\varepsilon^{-1} (\mathcal{E}_1^n + \mathcal{E}_2^n)^2.$$

Similarly, we have

$$\text{V} \leq \frac{1}{4} \int_{Q_n} H'_\varepsilon(A(u) - A(U_h)) |\nabla(A(u) - A(U_h))|^2 \theta + C\varepsilon^{-1} (\mathcal{E}_3^n)^2.$$

Substitute these estimates into (5.9) we arrive at, for any $\theta \in C_c^\infty(t^{n-1}, t^n)$ satisfying $0 \leq \theta \leq 1$,

$$(5.13) \quad - \int_{Q_n} U_\varepsilon(u, U_h) \theta_t \leq \kappa \varepsilon^\gamma \int_{Q_n} \frac{1}{\nu(\varepsilon, U_h)} (|\partial_t U_h| + |\nabla_x U_h|) + (\mathcal{E}_4^n + \mathcal{E}_5^n) + C\varepsilon^{-1} \left(\sum_{i=1}^3 \mathcal{E}_i^n \right)^2.$$

The following argument to choose θ is standard (see e.g. [19]). For any $t^{n-1} < t_1 < t_2 < t^n$, take α sufficiently small such that $t_1 - \alpha > t^{n-1}$, $t_2 + \alpha < t^n$, and define

$$\theta(t) = \int_{t-t_2}^{t-t_1} \omega_\alpha(s) ds,$$

where ω_α is the symmetric mollifier in \mathbf{R} . Then it is clear that $0 \leq \theta \leq 1$ and $\theta_t = \omega_\alpha(t - t_1) - \omega_\alpha(t - t_2)$. Thus

$$\begin{aligned} - \int_{Q_n} U_\varepsilon(u, U_h) \theta_t &= \int_{Q_n} U_\varepsilon(u, U_h) \omega_\alpha(t - t_2) - \int_{Q_n} U_\varepsilon(u, U_h) \omega_\alpha(t - t_1) \\ &\rightarrow \int_{\Omega} U_\varepsilon(u, U_h)(t_2) dx - \int_{\Omega} U_\varepsilon(u, U_h)(t_1) dx, \quad \text{as } \alpha \rightarrow 0. \end{aligned}$$

From the definition of the entropy function $U_\varepsilon(z, k) = \int_k^z H_\varepsilon(A(r) - A(k)) dr$, it is easy to prove that

$$|z - k| - \frac{\varepsilon}{\nu(\varepsilon, k)} \leq U_\varepsilon(z, k) \leq |z - k|,$$

where $\nu(\varepsilon, k) = \min\{A'(s) : |A(s) - A(k)| \leq \varepsilon\}$ is defined the same as that in Lemma 5.1. Thus

$$\begin{aligned} & - \lim_{\alpha \rightarrow 0} \int_{Q_n} U_\varepsilon(u, U_h) \theta_t \\ & \geq \| (u - U_h)(t_2) \|_{L^1(\Omega)} - \| (u - U_h)(t_1) \|_{L^1(\Omega)} - \varepsilon \int_{\Omega} \frac{1}{\nu(\varepsilon, U_h(x, t_2))} dx \\ & \rightarrow \| u^n - U_h^n \|_{L^1(\Omega)} - \| u^{n-1} - U_h^{n-1} \|_{L^1(\Omega)} - \varepsilon \| \nu(\varepsilon, U_h^n)^{-1} \|_{L^1(\Omega)}, \end{aligned}$$

as $t_2 \rightarrow t^n$ and $t_1 \rightarrow t^{n-1}$. Therefore we deduce from (5.13) that

$$\begin{aligned} \|u^n - U_h^n\|_{L^1(\Omega)} &\leq \varepsilon^\gamma \left(\mathcal{K} \int_{Q_n} \frac{1}{\nu(\varepsilon, U_h)} (|\partial_t U_h| + |\nabla_x U_h|) + \int_{\Omega} \frac{1}{\nu(\varepsilon, U_h^n)} \right) \\ &\quad + \|u^{n-1} - U_h^{n-1}\|_{L^1(\Omega)} + \mathcal{E}_4^n + \mathcal{E}_5^n + C\varepsilon^{-1} \left(\sum_{i=1}^3 \mathcal{E}_i^n \right)^2. \end{aligned}$$

Now let $\varepsilon_n = (\sum_{i=1}^3 \mathcal{E}_i^n)^{\frac{2}{1+\gamma}}$ and take $\varepsilon = \varepsilon_n / \Lambda_n^{\frac{1}{1+\gamma}}$, where Λ_n is defined in (5.10). Since $\varepsilon \leq \varepsilon_n$, we have $\nu(\varepsilon, U_h) \geq \nu(\varepsilon_n, U_h)$, $\nu(\varepsilon, U_h^n) \geq \nu(\varepsilon_n, U_h^n)$, and consequently,

$$\int_{Q_n} \frac{1}{\nu(\varepsilon, U_h)} (|\partial_t U_h| + |\nabla_x U_h|) + \int_{\Omega} \frac{1}{\nu(\varepsilon, U_h^n)} \leq \Lambda_n.$$

Thus

$$\|u^n - U_h^n\|_{L^1(\Omega)} \leq \|u^{n-1} - U_h^{n-1}\|_{L^1(\Omega)} + \mathcal{E}_4^n + \mathcal{E}_5^n + C\Lambda_n^{\frac{1}{1+\gamma}} \left(\sum_{i=1}^3 \mathcal{E}_i^n \right)^{\frac{2\gamma}{1+\gamma}}.$$

This completes the proof upon summing n from 1 to m , $m \geq 1$. \square

To conclude this section, we give several remarks about the a posteriori error estimate derived in this section.

Remark 5.7. In practical computations, the error indicator \mathcal{E}_0 for the initial error can be easily reduced by refining the initial mesh, and the source error indicator \mathcal{E}_5 can be controlled by reducing time step sizes. The characteristic error indicator \mathcal{E}_4 can be reduced by reducing the time step size if the approximate characteristics $\tilde{X}(t)$ in (3.5) is solved by convergent multistep Euler method or high order Runge-Kutta method.

Remark 5.8. In the case of strong diffusion $A'(s) \geq \beta > 0$ for any $s \in \mathbf{R}$ and A' is uniformly Lipschitz continuous, then the Hölder exponent $\gamma = 1$ in (H2) and Λ_n is bounded by $\beta^{-1} \|U_h\|_{BV(Q_n)}$ which is expected to be bounded in practical computations. The a posteriori error estimator in Theorem 5.6 then recovers the standard a posteriori error estimator derived in the literature for parabolic problems [25, 7]. In particular, the space error indicators $\mathcal{E}_1^n, \mathcal{E}_2^n$, which control the adaptation of finite element meshes at each time step, are sharp in the sense that a local lower bound for the error can be established by extending the argument in [7, Theorem 2.2] for linear parabolic equations.

Remark 5.9. In the case of small constant viscosity $A' = \epsilon$, then the Hölder exponent $\gamma = 1$ in (H2), and $\Lambda_n = C\epsilon^{-1}$. The estimators derived in Theorem 5.6 are closely related to the estimators in [16], in which $L^2(L^2)$ a posteriori error estimates are derived based on the duality argument for the linear convection dominated equation

$$(5.14) \quad \frac{\partial u}{\partial t} + \operatorname{div}(\mathbf{v}u) - \epsilon \Delta u = g \quad \text{in } Q,$$

where $\mathbf{v} \in C(\bar{Q})^2$ such that $\operatorname{div} \mathbf{v} = 0$. For the linear problem (5.14), we remark that one can derive an $L^\infty(L^1)$ a posteriori error estimate of the same form as in Theorem 5.6 without using the Kružkov “doubling of variables” technique. We now describe briefly this simple argument. The weak formulation of (5.14) is

$$(5.15) \quad \langle \partial_t u, \varphi \rangle - \langle \mathbf{v}u, \nabla \varphi \rangle + \epsilon \langle \nabla u, \nabla \varphi \rangle = \langle g, \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega).$$

The discrete problem is the same as in (3.4) and we define the discrete residual $\mathcal{R} \in L^2(0, T; H^{-1}(\Omega))$ similar to (4.2), for any $\varphi \in H_0^1(\Omega)$,

$$(5.16) \quad \langle \partial_t U_h, \varphi \rangle - \langle \mathbf{v} U_h, \nabla \varphi \rangle + \epsilon \langle \nabla U_h, \nabla \varphi \rangle = \langle g, \varphi \rangle - \langle \mathcal{R}, \varphi \rangle.$$

Subtracting (5.15) from (5.16) we get the following error equation, for any $\varphi \in H_0^1(\Omega)$,

$$\langle \partial_t(u - U_h), \varphi \rangle - \langle \mathbf{v}(u - U_h), \nabla \varphi \rangle + \epsilon \langle \nabla(u - U_h), \nabla \varphi \rangle = \langle \mathcal{R}, \varphi \rangle.$$

The a posteriori error estimate can be readily derived by taking $\varphi = H_\delta(u - U_h)$, where $H_\delta(s) = s/\sqrt{s^2 + \delta^2}$ is a regularization of $\text{sgn}(s)$, using the following Galerkin orthogonality for $t \in (t^{n-1}, t^n]$

$$\begin{aligned} \langle \mathcal{R}, \varphi \rangle &= \langle g - \bar{g}^n, \varphi \rangle + \left\langle \frac{U_h^n - \bar{U}_h^{n-1}}{\tau_n} - (\partial_t U_h + \mathbf{v} \nabla U_h), \varphi \right\rangle \\ &\quad + \langle R^n, \varphi - \Pi^n \varphi \rangle - \epsilon \langle \nabla U_h^n, \nabla(\varphi - \Pi^n \varphi) \rangle - \epsilon \langle \nabla(U_h - U_h^n), \nabla \varphi \rangle, \end{aligned}$$

and exploiting the standard argument in the a posteriori error analysis. We remark, however, that this simple argument can not be extended to deal with the nonlinear problem considered in this paper.

Remark 5.10. The a posteriori error analysis in this paper is different from the a posteriori error analysis for nonlinear conservation laws in [11, 12, 20] or nonlinear degenerate parabolic equations in [23] in the following aspects. Firstly, only Cauchy problems are considered in [11, 12, 20, 23]. The difficulty to include boundary condition is essential. In this paper, we have used the recently introduced technique of “boundary layer sequence” in [22] to overcome the difficulty. We also remark that the use of the technique of “boundary layer sequence” allows us to extend the analysis in the paper to treat other types of boundary conditions. We will report the progress in this respect in future studies. Secondly, the nature of the estimators are different: our estimators emphasize the diffusion effect of the problem which requires the assumption $A'(s) > 0$ for any $s \in \mathbf{R}$; the estimates in [23] are valid for any nonlinear function A such that $A'(s) \geq 0$. Consequently, the estimators in [23] do not have the right order in the region when the solution is smooth. Finally, the methods of analysis are different. Recall that there are several parameters introduced in the analysis

- The regularizing parameter ϵ in $H_\epsilon(z)$;
- The boundary layer sequence parameters δ, η and the mollifier parameters l, m, n .

The analysis for Cauchy problems in [11, 12, 20, 23] is based on letting $\epsilon \rightarrow 0$ and taking finite mollifier parameters l, m, n . Note that there are no boundary layer sequence parameters δ, η for the analysis for Cauchy problems. The analysis in this paper is based on letting $\delta, \eta \rightarrow 0$ and $l, m, n \rightarrow \infty$ but taking a finite ϵ . We are not able to use the same technique as that in [11, 12, 20, 23] by choosing finite mollifier parameters l, m, n to treat the problem with boundary conditions.

6. A NUMERICAL EXAMPLE

In this section we report a numerical example for the following linear convection diffusion equation

$$\frac{\partial u}{\partial t} + \operatorname{div}(\mathbf{v}u) - \epsilon \Delta u = g \quad \text{in } Q.$$

The computation makes use of the adaptive finite element toolbox ALBERT [27] and is based on the a posteriori error estimate derived in this paper. Further numerical experiments including the application to nonlinear convection diffusion equations will be reported in a forthcoming paper.

By Theorem 5.6 or the argument in Remark 5.9, we know that

$$\begin{aligned} \max_{1 \leq m \leq N} \|u^m - U_h^m\|_{L^1(\Omega)} &\leq \mathcal{E}_0 + \sum_{n=1}^N (\mathcal{E}_4^n + \mathcal{E}_5^n) + C \left(\sum_{n=1}^N \tau_n \eta_{\text{time}}^n \right)^{1/2} \\ &\quad + C \left(\sum_{n=1}^N \tau_n \eta_{\text{space}}^n \right)^{1/2}, \end{aligned}$$

where the time error indicator η_{time}^n and space error indicator η_{space}^n are given by

$$\eta_{\text{time}}^n = \epsilon \|\nabla(U_h^n - U_h^{n-1})\|_{L^2(\Omega)}^2, \quad \eta_{\text{space}}^n = \sum_{K \in \mathcal{M}^n} \eta_K^n$$

with the local error indicator η_K^n defined as

$$\eta_K^n = \epsilon^{-1} \|h_K R^n\|_{L^2(K)}^2 + \epsilon \|h_c^{1/2} [\nabla U_h^n]\|_{L^2(\partial K)}^2.$$

Let the time and space tolerances TOL_{time} and $\text{TOL}_{\text{space}}$ be given. At each time step $n \geq 1$, the time step size τ_n is determined through the requirements

$$\eta_{\text{time}}^n \leq \frac{\text{TOL}_{\text{time}}^2}{4T}, \quad \frac{1}{\tau_n} (\mathcal{E}_4^n + \mathcal{E}_5^n) \leq \frac{\text{TOL}_{\text{time}}}{2T}.$$

The set of elements marked for refinements $\mathcal{M}_{\text{refine}}^n$ and the set of elements marked for coarsening $\mathcal{M}_{\text{coarse}}^n$ are determined by the relations

$$\sum_{K \in \mathcal{M}_{\text{refine}}^n} \eta_K^n \geq \theta_r^2 \sum_{K \in \mathcal{M}^n} \eta_K^n, \quad \sum_{K \in \mathcal{M}_{\text{coarse}}^n} \eta_K^n \leq \theta_c^2 \sum_{K \in \mathcal{M}^n} \eta_K^n.$$

The iteration for the mesh adaptation at each time step n is terminated whenever $\eta_{\text{space}}^n \leq \text{TOL}_{\text{space}}^2/T$ is satisfied.

We consider the so-called rotating cylinder problem from [16, Example 6.3]. Let $\Omega = (0, 1)^2$, $T = 0.5$, $g = 0$, $\mathbf{v} = -2\pi(2x_2 - 1, 1 - 2x_1)^T$ and

$$u_0 = \begin{cases} 1, & \text{for } s \leq 1/4; \\ 0, & \text{otherwise,} \end{cases}$$

where $s^2 = (2x_1 - 1/2)^2 + (2x_2 - 1)^2$.

In the computations we take $\text{TOL}_{\text{time}} = \text{TOL}_{\text{space}}$ and $\theta_r = 0.5$, $\theta_c = 0.1$. The initial mesh \mathcal{M}_0 at time $t = 0$ is so chosen that $\mathcal{E}_0 = \|u_0 - U_h^0\|_{L^1(\Omega)} \leq \text{TOL}_{\text{initial}}$. In our computations, we take $\text{TOL}_{\text{initial}} \ll \text{TOL}_{\text{space}}$ so that the initial errors are negligible. Table 6.1 shows the initial errors and the numbers of nodes of the corresponding initial meshes for different choices of ϵ , $\text{TOL}_{\text{initial}}$, and the corresponding $\text{TOL} = \text{TOL}_{\text{space}} + \text{TOL}_{\text{time}}$.

TABLE 6.1. The initial errors and the numbers of nodes of the initial meshes for different choices of ϵ , TOL and TOL_{initial}.

	TOL	TOL _{initial}	\mathcal{E}_0	number of nodes
$\epsilon=1e-5$	6.0	2.0e-4	1.47e-4	8415
	4.0	1.0e-4	9.58e-5	13407
	2.0	5.0e-5	4.96e-5	24751
	1.0	2.5e-5	1.77e-5	68051
$\epsilon=1e-3$	2.0	2.0e-3	1.51e-3	835
	1.0	1.0e-3	5.73e-4	2247
	0.5	5.0e-4	3.91e-4	3287
	0.25	2.5e-4	1.47e-4	8415
	0.125	1.25e-4	9.58e-5	13407

Tables 6.2 and 6.3 show the total number of nodes $M = \sum_{n=1}^N M_n$, where M_n is the number of nodes of \mathcal{M}_n , the total estimated error η and the convergence rate α for $\epsilon = 10^{-3}$ and $\epsilon = 10^{-5}$ respectively. For two different TOL _{i} , let η_i and M_i be the corresponding total estimated error and total number of nodes, the convergence rate α is computed by

$$\alpha = \frac{\log(\eta_1/\eta_2)}{\log(M_1/M_2)}.$$

TABLE 6.2. The total number of nodes M , the total estimated error η and the convergence rate α when $\epsilon = 10^{-3}$.

TOL	M	η	α
2.0	20540	1.2044	
1.0	113238	0.6392	-0.3711
0.5	728872	0.3267	-0.3605
0.25	4434790	0.1637	-0.3827
0.125	25173197	0.0814	-0.4024

TABLE 6.3. The total number of nodes M , the total estimated error η and the convergence rate α when $\epsilon = 10^{-5}$.

TOL	M	η	α
6.0	102111	3.7485	
4.0	267457	2.6576	-0.3572
2.0	2046604	1.2973	-0.3524
1.0	14878263	0.6426	-0.3541

We observe from Tables 6.2 and 6.3 that the a posteriori error estimator η is roughly proportional to $M^{-1/3}$, i.e. $\eta \approx CM^{-1/3}$ for some constant $C > 0$. This indicates that

$$(6.1) \quad \max_{1 \leq n \leq N} \|u - U^n\|_{L^1(\Omega)} \leq CM^{-1/3}.$$

We remark that because of the singular nature of the solutions, the numerical scheme (3.4)-(3.5) with uniform refinements both in space and time will not produce the convergence rate (6.1) in terms of the error reduction.

Figure 6.1 shows the meshes and the surface plots of the solutions at time $t = 0.251278$ and $t = 0.500878$ when $\epsilon = 10^{-5}$. We observe that the meshes “follow” the positions of the cylinder. For this problem, the “leakage” of the numerical solutions is observed in [16] in the following sense: the mesh is coarser in the regions of the cylinder closest to and farthest from the center of rotation. We do not observe, however, this phenomenon in our computation as indicated in Figure 6.1. This may be explained by the difference of the error indicators used in two papers.

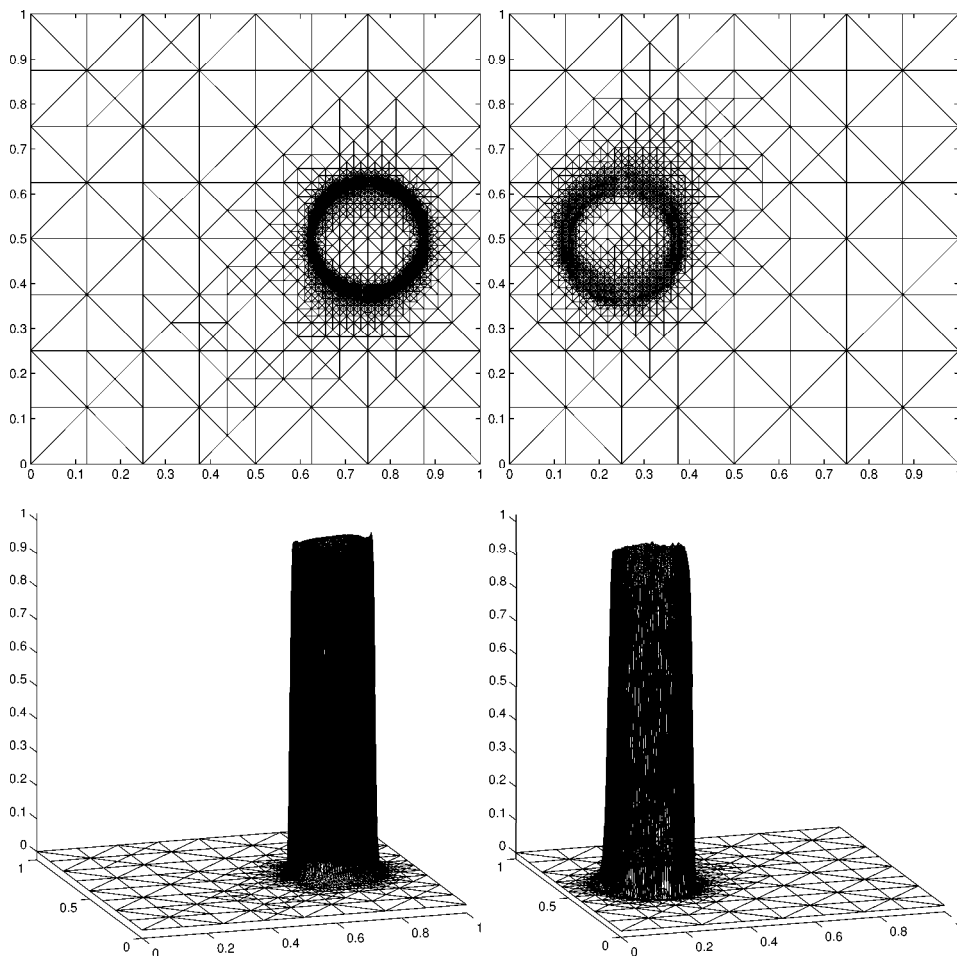


FIGURE 6.1. The meshes (top) and the surface plots (bottom) of the solutions at time $t = 0.251278$ and $t = 0.500878$ when $\epsilon = 10^{-5}$. The number of nodes are 3133 ($t = 0.251278$) and 2143 ($t = 0.500878$) respectively.

REFERENCES

- [1] T. Arbogast and M.F. Wheeler, *A characteristic-mixed finite element method for advection-dominated transport problems*, SIAM J. Numer. Anal. 32 (1995), 404-424.
- [2] H.W. Alt and S. Luckhaus, *Quasilinear elliptic-parabolic differential equations*, Math. Z. 183 (1983), 311-341.
- [3] I. Babuška and C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. 15 (1978), 736-754.
- [4] J. Carrillo, *Entropy solutions for nonlinear degenerate problems*, Arch. Rational Mech. Anal. 147 (1999), 269-361.
- [5] Z. Chen and S. Dai, *Adaptive Galerkin methods with error control for a dynamical Ginzburg-Landau model in superconductivity*, SIAM J. Numer. Anal. 38 (2001), 1961-1985.
- [6] Z. Chen and S. Dai, *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients*. SIAM J. Sci. Comput. 24 (2002), 443-462.
- [7] Z. Chen and F. Jia, *An adaptive finite element method with reliable and efficient error control for linear elliptic problems*, Math. Comp. (to appear).
- [8] Z. Chen, R.H. Nochetto and A. Schmidt, *A characteristic Galerkin method with adaptive error control for continuous casting problem*, Comput. Methods Appl. Mech. Engrg. 189 (2000), 249-276.
- [9] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numer. 9 (1975), 77-84.
- [10] B. Cockburn, *A simple introduction to error estimation for nonlinear hyperbolic conservation laws*, Department of Mathematics, University of Minnesota, 2003. <http://www.math.umn.edu/~cockburn/LectureNotes.html>
- [11] B. Cockburn, F. Coquel and P.G. Lefloch, *An error estimate for finite volume methods for multidimensional conservation laws*, Math. Comp. 63 (1994), 77-103.
- [12] B. Cockburn and P.-A. Gremaud, *Error estimates for finite element methods for scalar conservation laws*, SIAM J. Numer. Anal. 33 (1996), 522-554.
- [13] J. Douglas, Jr. and T.F. Russell, *Numerical methods for convection-dominated diffusion problem based on combining the method of characteristic with finite element or finite difference procedures*, SIAM J. Numer. Anal. 19 (1982), pp. 871-885.
- [14] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal. 28 (1991), 43-77.
- [15] C. Johnson and A. Szepessy, *Adaptive finite element methods for conservation laws based on a posteriori error estimates*, Comm. Pure Appl. Math. XLVIII (1995), 199-234.
- [16] P. Houston and E. Süli, *Adaptive Lagrange-Galerkin methods for unsteady convection-diffusion problems*, Math. Comp. 70 (2000), 77-106.
- [17] K. Huang, R. Zhang and M.T. van Genuchten, *An Eulerian-Lagrangian approach with an adaptively corrected method of characteristics to simulate variably saturated water flow*, Water Resource Research 30 (1994), 499-507.
- [18] J. Kačur, *Solution of degenerate convection-diffusion problems by the method of characteristics*, SIAM J. Numer. Anal. 39 (2001), 858-879.
- [19] K.H. Karlsen and N.H. Risebro, *On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients*, Discrete Contin. Dynam. Systems - Series A 9 (2003), 1081-1104.
- [20] D. Kröner and M. Ohlberger, *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multi-dimensions*, Math. Comp. 69 (2000), 25-39.
- [21] N.N. Kružkov, *First order quasi-linear equations in several independent variables*, Math. USSR Sbornik 10 (1970), 217-243
- [22] C. Mascia, A. Porretta and A. Terracina, *Nonhomogeneous Dirichlet problems for degenerate parabolic-hyperbolic equations*, Arch. Rational Mech. Anal. 163 (2002), 87-124.
- [23] M. Ohlberger, *A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations*, Math. Model. Numer. Anal. 35 (2001), 355-387.
- [24] F. Otto, *L^1 -contraction and uniqueness for quasilinear elliptic-parabolic equations*, J. Diff. Equations 131 (1996), 20-38.
- [25] M. Picasso, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg. 167 (1998), 223-237.

- [26] O. Pironneau, *On the transport-diffusion algorithm and its application to the Navier-Stokes equations*, Numer. Math. 38 (1982), pp. 309-332.
- [27] A. Schmidt and K.G. Siebert, *ALBERT: An adaptive hierarchical finite element toolbox*, IAM, University of Freiburg, 2000. <http://www.mathematik.uni-freiburg.de/IAM/Research/projectsdz/albert>.

LSEC, INSTITUTE OF COMPUTATIONAL MATHEMATICS, ACADEMY OF MATHEMATICS AND SYSTEM SCIENCES, CHINESE ACADEMY OF SCIENCES, BEIJING 100080, CHINA.
E-mail address: zmchen@lsec.cc.ac.cn

INSTITUTE OF COMPUTATIONAL MATHEMATICS, ACADEMY OF MATHEMATICS AND SYSTEM SCIENCES, CHINESE ACADEMY OF SCIENCES, BEIJING 100080, CHINA
E-mail address: ghji@lsec.cc.ac.cn