A POSTERIORI ERROR ESTIMATES FOR NUMERICAL APPROXIMATIONS TO SCALAR CONSERVATION LAWS: SCHEMES SATISFYING STRONG AND WEAK ENTROPY INEQUALITIES

LAURENT GOSSE[1] AND CHARALAMBOS MAKRIDAKIS[1 2]

[1] *Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion, Crete,*

[2] *Department of Mathematics University of Crete, 71409 Heraklion, Crete, GREECE*

Abstract. Explicit a posteriori error estimates are proposed for schemes approximating the entropy solution of the scalar conservation law. A key property for deriving these estimates is the discrete entropy. We propose estimates valid for monotone finite volume schemes and for second order one dimensional finite difference schemes.

1. Introduction. We consider numerical (finite difference and finite volume) schemes approximating the entropy solution of the conservation law

$$\partial_t u + \mathrm{div} F(u) = 0, \quad x \in \mathbb{R}^d, t > 0, \tag{1}$$

with initial data $u(x,0) = u_0(x) \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, [9]. Our goal is to provide explicit a posteriori error estimates for schemes satisfying certain cell entropy inequalities. These estimates is a first step towards designing adaptive algorithms based on appropriate mesh refinement strategy for the efficient approximation of (1). The main tool for establishing such estimates is an appropriate (and simple) use of Kruzkov's estimates in the framework of [1] and [7]. For previous work on a posteriori error estimates in L^1 cf. [3], see also [11]. A posteriori error estimates were previously considered in [6] for finite element schemes approximating the $\varepsilon-$viscous one dimensional systems of conservation laws in a weighted space-time L^2 norm. Estimates in a dual Lip_+ norm for (1) in the case $d = 1$ were proved in [13].

We shall see that a basic property of a numerical scheme that can lead to the a posteriori error estimate is the discrete entropy. We first consider schemes satisfying "strong" entropy inequalities. Therefore in section 2 we consider monotone finite volume schemes in unstructured meshes. These schemes satisfy an a posteriori error estimate in L^1, Theorem 1. This bound is of order $O(h^{1/2})$ if the approximate solution is bounded in BV. A local version of the same result can also be obtained. A sketch of the (simple) proof of Theorem 1 is presented in Section 4. In section 3 we consider schemes that satisfy "weak" entropy inequalities. Typical examples of such schemes are the second order schemes considered in [15] and the central schemes of [12]. It turns out that for such schemes one can derive similar a posteriori error estimates, Theorem 2. Of course we would like to know whether these estimates are meaningful at least in the case where we have more information available. Theorem 3 shows that, in the case of one dimensional convex flux, one can recover convergence rates known before [14]. We emphasize that a posteriori estimates, such as these presented in this note, provide the following useful information: If the (computable) upper bound of the error is less than a given tolerance τ we know that the error measured in the corresponding norm is at most τ even if we cannot prove the convergence of the scheme.

IACM-FORTH Technical Report 98.4 (1998)

2. Monotone Finite Volume Schemes. We begin by describing the finite volume schemes. Let $\mathcal{T}_h$ be a decomposition of $\mathbb{R}^d$ into non-overlapping, nonempty and open polyhedra such that $\bigcup_{K\in\mathcal{T}_h}\bar{K}=\mathbb{R}^d$. The set of faces of K is denoted by ∂K and, on each face e on K, $\nu_{e,K}\in\mathbb{R}^d$ represents the outward unit normal to the face e. Γ_h will denote the set of all faces of the decomposition $\mathcal{T}_h$. Given a face e of K, K_e denotes the unique polyhedron that shares the face e with K. The volume of K is denoted by $|K|$ and the $(d-1)$-measure of e by $|e|$. Let h_K be the diameter of the polyhedron K and let $h=\sup_{K\in\mathcal{T}_h}h_K<1$. We assume that our decomposition is regular, i.e., if ρ_K is the diameter of the largest ball B, $B\subset K$, $h_K\le\gamma\rho_K$, $\quad K\in\mathcal{T}_h$, with a constant γ independent of h.

Monotone schemes are based on consistent discrete fluxes g^K,

$$g^K(u,v)=-g^{K_e}(v,u),\quad g^K(u,u)=F(u)\cdot\nu_{e,K},\quad \frac{\partial g^K}{\partial u}\ge 0\ ,\frac{\partial g^K}{\partial v}\le 0 \tag{2}$$

Let δ be the time step and $t^n=n\delta$. Then we seek a piecewise constant function $(u_{h,\delta},u_{h,\delta}|_{K\times[t^n,t^{n+1})}=u_K^n)$, such that

$$u_K^{n+1}=u_K^n-\frac{\delta}{|K|}\sum_{e\in\partial K}|e|g^K(u_K^n,u_{K_e}^n), \tag{3}$$

where $u_K^0=\frac{1}{|K|}\int_K u_0(x)\,dx$. Then one can show that the scheme (2) satisfies the entropy inequalities:

$$|u_K^{n+1}-\xi|+\frac{\delta}{|K|}\sum_{e\in\partial K}|e|D_\xi^K(u_K^n,u_{K_e}^n)\le|u_K^n-\xi|\ ,\quad \xi\in\mathbb{R}, \tag{4}$$

under, of course, appropriate CFL conditions. Here $D_\xi^K(u,v)=g^K(u\vee\xi,v\vee\xi)-g^K(u\wedge\xi,v\wedge\xi)$, where for $a,b\in\mathbb{R}$ $a\vee b=\max\{a,b\}$ and $a\wedge b=\min\{a,b\}$. We call (4) "strong" entropy inequality since the entropy flux is consistent, $D_\xi^K(u,u)=\operatorname{sgn}(u-\xi)(F(u)-F(\xi))\cdot\nu_{e,K}\quad u\in\mathbb{R}$. It is well known that these schemes are L^1 contractive and also that satisfy a weak dissipation estimate, [2], [16]. In section 4 we sketch the proof of:

Theorem 1. *Let $u_{h,\delta}$ be the solution of the finite volume scheme* (3)*. Then the following estimate holds*

$$\sup_{t\le t^n}\|u(t)-u_{h,\delta}(t)\|_{L^1(\mathbb{R}^d)}\le\|u_{h,\delta}(0)-u_0\|_{L^1(\mathbb{R}^d)}+C\sqrt{TV(u_0)M}\Big\{\Big(\sum_{m=0}^{n}\delta\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}h_K^d|u_K^m-u_{K_e}^m|\Big)^{1/2}$$
$$+\sqrt{t^n}\max_{m\le n+1}\Big(\sum_{K\in\mathcal{T}_h}|K|\,|u_K^m-u_K^{m-1}|\Big)^{1/2}\Big\}\,.$$

This result provides en explicit computable error bound for the error. This bound is of order $O(h^{1/2})$ if the approximate solution is bounded in BV. Using the weak dissipation estimate of [2], [16] a local version of Theorem 1 implies the asymptotic rate of convergence $h^{1/4}$ proved in [2], [16] under the CFL condition needed for the validity of (4).

In addition a local estimate of $\|u(t)-u_{h,\delta}(t)\|_{L^1(K)}$ can be obtained, cf. (11) below. In this case the error terms will be the same as above with sums in appropriate extended cones of dependence of the form $B_t=B(x_0,R+M(T-t)+\Delta_h)$, where $\Delta_h=O(h^{1/4})$. In an independent work Kröner and Ohlberger [8] have obtained a posteriori error bounds for these schemes similar to our Theorem 1 for fluxes $F=F(x,t,u)$. The estimate of the error in [8] is given in the *space-time* L^1 norm under the additional assumption that the mesh is quasi-uniform.

3. Second Order Finite Difference Schemes. We restrict now to one dimensional finite difference schemes. Let $K_j = (x_{j+\frac{1}{2}}, x_{j-\frac{1}{2}})$, $j \in \mathbb{Z}$, and x_j denotes the middle point of K_j. Then $|K_j| = h_j \leq h$. We consider piecewise constant numerical approximations $(v_h, v_h|_{K_j \times [t^n, t^{n+1})} = v_j^n)$ generated by schemes of the form

$$v_j^{n+1} = v_j^n - \frac{\delta}{h_j}\,(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n)\,, \tag{5}$$

with appropriate initial data v_h^0. Where for a given $2p$ consistent discrete flux g,

$$g_{j+\frac{1}{2}}^n = g(v_{j-p+1}^n, \dots, v_{j+p}^n)\,, \quad g(v, \dots v) = f(v) \;\; v \in \mathbb{R}. \tag{6}$$

We consider second order schemes (5) in their viscous form, cf. [15], [12], for which one introduces a *modified* numerical flux $\tilde{f}_j^n = f(v_j^n) + \ell_j^n$:

$$v_j^{n+1} = v_j^n - \frac{\delta}{2h_j}\,(\tilde{f}_{j+1}^n - \tilde{f}_{j-1}^n)) + \frac{\delta}{2h_j}\,(Q_{j+\frac{1}{2}}^n\,(v_{j+1}^n - v_j^n) - Q_{j-\frac{1}{2}}^n\,(v_j^n - v_{j-1}^n))\,. \tag{7}$$

Our estimates are based on the following *weak* cell entropy inequality, [5], see also [15], [12].

Lemma 1. *If the CFL condition* $|\Delta \tilde{f}_{j+\frac{1}{2}}^n| \leq Q_{j+\frac{1}{2}}^n \leq \frac{h_j}{2\delta}$ *is met, then for* $U_\xi(v) := |v - \xi|$, $m_{j+\frac{1}{2}}^n = Q_{j+\frac{1}{2}}^n - \Delta \tilde{f}_{j+\frac{1}{2}}^n$, $p_{j-\frac{1}{2}}^n = Q_{j-\frac{1}{2}}^n + \Delta \tilde{f}_{j-\frac{1}{2}}^n$, $\Delta \tilde{f}_{j+\frac{1}{2}}^n := \frac{\tilde{f}_{j+1}^n - \tilde{f}_j^n}{v_{j+1}^n - v_j^n}$, *and all* $\xi \in \mathbb{R}$, *there holds*

$$U_\xi(v_j^{n+1}) - U_\xi(v_j^n) + \frac{\delta}{2h_j}\big[(m_{j+\frac{1}{2}}^n(U_\xi(v_{j+1}^n) - U_\xi(v_j^n)) + p_{j-\frac{1}{2}}^n(U_\xi(v_j^n) - U_\xi(v_{j-1}^n))\big] \leq 0\,.$$

Lemma 1 is the main tool in establishing

Theorem 2. [5] *Under the assumptions of Lemma 1 the following error bound holds*

$$\begin{aligned}
\sup_{t \leq t^n} \|v_h(t) - u(t)\|_{L^1} \leq \|v_h(0) - u_0\|_{L^1} + C\sqrt{TV(u_0)M}\Big\{\Big(\sum_{m=0}^{n} \delta \sum_{j\in\mathbb{Z}} h_j|v_j^m - v_{j-1}^m|\Big)^{1/2} \\
+\sqrt{t^n} \max_{m \leq n+1} \Big(\sum_{j\in\mathbb{Z}} h_j|v_j^m - v_j^{m-1}|\Big)^{1/2}\Big\} + C\sum_{m=0}^{n} \delta \sum_{j\in\mathbb{Z}} (\omega_{j-\frac{1}{2}}^m + \nu_{j-\frac{1}{2}}^m)\,,
\end{aligned} \tag{8}$$

where for $A(v_{j-1}^m, v_j^m\,;\theta) = 2\big((1-\theta)f(v_j^m) + \theta f(v_{j-1}^m) - f((1-\theta)v_j^m + \theta v_{j-1}^m)\big)$ $\theta \in [0,1]$,

$$\omega_{j-\frac{1}{2}}^m = \omega(v_{j-1}^m, v_j^m) = \max\,\big\{0\,,\,\big(\chi_{(v_j^m - v_{j-1}^m)^+} - \chi_{(v_{j-1}^m - v_j^m)^+}\big)\max_{\theta\in[0,1]} A(v_{j-1}^m, v_j^m\,;\theta)\big\}\,.$$

and $\nu_{j-\frac{1}{2}}^m = |\ell_j^m - \ell_{j-1}^m|$.

Note that the constant C in Theorem 2 depends only on the constant of the estimate in Lemma 2 below and can be estimated, [5]. The error bound (8) is similar to this of Theorem 1 with the exception of the presence of the $(\omega_{j-\frac{1}{2}}^m + \nu_{j-\frac{1}{2}}^m)$ term. If $F = f$ is smooth then it is bounded by $C(f)(v_{j-1}^m - v_j^m)^2$ which is (presumably) small in the areas of the smoothness of u. Note also that Oleinik's entropy condition implies that if u_-, u_+ are the values of u at a shock then $\omega(u_-, u_+) = 0$. In the case where f is convex this term is of order one provided that the scheme is Lip_+ stable:

Theorem 3. [5] *Assume in addition that f is convex, then the estimate of Theorem 2 is valid with $\omega^m_{j-\frac{1}{2}}$ bounded by*

$$\omega^m_{j-\frac{1}{2}} \leq \max_{s\in[v^m_{j-1},v^m_j]} f''(s).\max\{0, v^m_j - v^m_{j-1}\}.|v^m_j - v^m_{j-1}|\,.$$

Since, by the CFL condition the difference in time terms can be transformed in to space terms, and since these schemes are TVD we recover the error bound of [14] for the L^1 norm provided that the discrete Lip_+ seminorm of v_h norm grows at most as $h^{-1/2}$ and the sum of $\nu^m_{j-\frac{1}{2}}$ is of the order $h^{1/2}$. Note that if we consider schemes of the form (7) with $\tilde{f}^n_j = f(v^n_j)$ the estimate (8) holds with $\nu^m_{j-\frac{1}{2}} = 0$ and the last term is $O(h)$.

4. Proof of Theorem 1. We present the basic steps of the Proof of Theorem 1 which is is based on the following approximation lemma, [1], [7]. For a proof, cf. [7].

Lemma 2. *Let $u_h, u \in L^\infty_{loc}((0,\infty), L^1_{loc}(\mathbb{R}^d))$ be right continuous in t, with values in $L^1_{loc}(\mathbb{R}^d)$. Assume that u is the entropy solution of the conservation law (1.1), with initial value $u^0 \in BV(\mathbb{R}^d)$. Let Ψ be a nonnegative test function, $\Psi \in C^\infty_c((0,\infty)\times\mathbb{R}^d)$, and assume that u_h satisfies*

$$\begin{aligned}
(9)\qquad & -\iint_{(0,\infty)\times\mathbb{R}^d} \big(|u_h - k|\partial_t\Psi + \mathrm{sgn}(u_h - k)[F(u_h) - F(k)]\cdot\nabla_x\Psi\big)dtdx \\
& \leq \iint_{(0,\infty)\times\mathbb{R}^d} \Big(\beta_O B_O(\Psi) + \alpha_G|\partial_t\Psi| + \sum_j \beta^j_H B^j_H(\frac{\partial\Psi}{\partial x_j})\Big)dxdt, \quad \textit{for all } k\in\mathbb{R},
\end{aligned}$$

where $F = (F_1,\dots,F_d)$ and α_G, β_O, β^j_H, are nonnegative k-independent functions in $L^1_{loc}([0,\infty)\times\mathbb{R}^d)$ and $\alpha_G, \in L^\infty_{loc}([0,\infty), L^1_{loc}(\mathbb{R}^d))$. Let $\Delta > 0$, and $\mathcal{S}_h = \{S\}$ be an element decomposition of $[0,\infty)\times\mathbb{R}^d$, into elements S, such that $\mathrm{diam}(S_t) \leq \Delta$ in the case where B_O or B^j_H, is not identically zero where $S_t = \{x : (t,x)\in S\}$.

If in addition, for all $(t,x)\in S$, $1\leq i,j\leq d$,

$$(10)\qquad |B_O(\Psi)(t,x)| \leq C\sup_{x'\in S_t}|\Psi(t,x')|\,,\quad |B^j_H(\frac{\partial\Psi}{\partial x_j})(t,x)| \leq C\sup_{x'\in S_t}|\frac{\partial\Psi}{\partial x_j}(t,x')|\,,$$

where C is a uniform constant independent of Ψ and the element decomposition $\mathcal{S}_h$, then the following estimate holds: for any $T\geq 0$, $x_0\in\mathbb{R}^d$, $R>0$, $\rho>0$ with $M = Lip(f)$, we have:

(11)

$$\begin{aligned}
&\int_{|x-x_0|<R} |u_h(T,x) - u(T,x)|dx \leq \int_{B_0} |u_h(0,x) - u(0,x)|dx + C(M+1)\Delta TV(u^0) \\
&+ C(1 + \frac{T(1+M)}{\Delta})\sup_{0\leq t\leq T+\rho}\int_{B_t}\alpha_G(t,x)dx + C\iint_{0\leq t\leq T\ x\in B^\Delta_t}\Big(\beta_O(t,x) + \frac{1}{\Delta}\sum_{j=1}^d \beta^j_H(t,x)\Big)dxdt\,,
\end{aligned}$$

where $B_t = B(x_0, R + M(T-t) + \Delta)$, $B^\Delta_t = B(x_0, R + M(T-t) + 2\Delta)$.

Let Ψ a nonnegative test function with compact support and set $V^n_K := |u^n_K - \xi|$ and $V^n_{F,K} := [F(u^n_K) - F(\xi)]\mathrm{sgn}(u^n_K - \xi)$. To apply Lemma 2 we have to estimate

$$E := -\sum_{n=0}^\infty \int_{I_n}\sum_{K\in\mathcal{T}_h}\int_K [V^n_K\Psi_t + V^n_{F,K}\cdot\nabla_x\Psi]dx\,dt,$$

where $I_n = [n\delta, (n+1)\delta)$. We denote by $\bar{\Psi}^e = \int_e \Psi\, ds$, $\bar{\Psi}^K = \int_K \Psi\, dt$ and $\bar{\Psi}_t^K = \int_K \Psi_t\, dt$. Note that since $\sum_{e\in\partial K} |e| D_\xi^K(u_K^n, u_K^n) = 0$, multiplying the discrete entropy inequality (4) by $\frac{1}{\delta}\int_{I_n} \bar{\Psi}^K dt$ we get

$$\sum_{n=0}^{\infty}\sum_{K\in\mathcal{T}_h} |u_K^{n+1} - \xi| \big(\frac{1}{\delta}\int_{I_n} \bar{\Psi}^K dt\big) + \sum_{n=0}^{\infty}\sum_{e\in\Gamma_h} \Big\{ \frac{|e|}{|K|}[D_\xi^K(u_K^n, u_{K_e}^n) - D_\xi^K(u_K^n, u_K^n)]\big(\int_{I_n} \bar{\Psi}^K dt\big)$$

$$+ \frac{|e|}{|K_e|}[D_\xi^{K_e}(u_{K_e}^n, u_K^n) - D_\xi^{K_e}(u_{K_e}^n, u_{K_e}^n)]\big(\int_{I_n} \bar{\Psi}^{K_e} dt\big)\Big\} \le \sum_{n=0}^{\infty}\sum_{K\in\mathcal{T}_h} |u_K^n - \xi| \big(\frac{1}{\delta}\int_{I_n} \bar{\Psi}^K dt\big).$$

Therefore, it turns out that:

$$E \le \sum_{n=0}^{\infty}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K} \{L_g |u_K^n - u_{K_e}^n|\} \int_{I_n} \Big|\bar{\Psi}^e - \frac{|e|}{|K|}\bar{\Psi}^K\Big| dt$$
$$+ \sum_{n=0}^{\infty}\sum_{K\in\mathcal{T}_h} |u_K^{n+1} - u_K^n| \int_K \Big|\Psi(t^{n+1}, x) - \frac{1}{\delta}\int_{I_n} \Psi(t,x)dt\Big| dx$$

where L_g denotes the Lipschitz constant of the flux g. To adjust to the notation of Lemma 2 let $\mathcal{S}_h = \{S_K^n\}$, $S_K^n = I_n \times K$, $K \in \mathcal{T}_h$ be a partition of $[0, +\infty) \times \mathbb{R}^d$. Then, for any $t > 0$, $(S_K^n)_t = K$. If

$$B_H(\nabla_x \Psi)\Big|_{S^K}(x,t) = \frac{1}{|K|}\Big|\Big(|e|\Psi(x,t) - \bar{\Psi}^e(t)\Big)\Big|,$$

then (10) is satisfied since,

$$\frac{1}{|K|}\Big|\Big(|e|\Psi(x,t) - \bar{\Psi}^e(t)\Big)\Big| \le \frac{1}{|K|} C h_K |e| \sup_{x'\in K} |\nabla\Psi(x',t)| \le C \sup_{x'\in K} |\nabla\Psi(x',t)|\,.$$

Therefore $u_{h,\delta}$ satisfies (9) with $\beta_H\big|_{S_K^n} = \sum_{e\in\partial K} L_g |u_K^n - u_{K_e}^n|$ and $\alpha_G\big|_{S_K^n} = |u_K^{n+1} - u_K^n|$. Hence the desired result follows if $\rho = \delta$ by minimizing over Δ the error bound (11).

Acknowledgments. This work is partially supported by European Union TMR grant No. ERBFMRXCT960033 (HCL).

References

[1] F. Bouchut and B. Perthame, *Kruzhkov's estimates for scalar conservation laws revisited*, Trans. Amer. Math. Soc. **350** (1998), 2847-2870.

[2] B. Cockburn, F. Coquel and P. LeFloch, *An error estimate for finite volume methods for conservation laws*, Math. Comp. **64** (1994), 77-103.

[3] B. Cockburn and H. Gau, *A posteriori error estimates for general numerical methods for scalar conservation laws*, Mat. Apl. Comput. **14** (1995), 37-47.

[4] B. Cockburn and P.-A. Gremaud, *A priori error estimates for numerical methods for scalar conservation laws. Part I: The general approach*, Math. Comp. **65** (1996), 533-573.

[5] L. Gosse and Ch. Makridakis, *Two a posteriori error estimates for one-dimensional scalar conservation laws*, (To appear).

[6] C. Johnson and A. Szepessy, *Adaptive finite element methods for conservation laws. Part I: The general approach*, Comm. Pure Appl. Math. **48** (1995), 199-234.

[7] M.A. Katsoulakis, G. Kossioris and Ch. Makridakis, *Convergence and error estimates of relaxation schemes for multidimensional conservation laws*, Preprint 97-3, Dept. of Math., University of Crete, (1997). To appear in Comm. Part. Diff. Eq..

[8] D. Kröner and M. Ohlberger, *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multidimensions*, Preprint no. 2-1998, Math. Fakultat, Univ. Freiburg, (1998).

[9] S.N. Kruzhkov, *First order quasilinear equations with several independent variables*, Math. USSR Sbornik **10** (1970), 217-243.

[10] N.N. Kuznetzov, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation*, USSR Comp. Math. and Math. Phys. **16** (1976), 105-119.

[11] B. J. Lucier, *A moving mesh numerical method for hyperbolic conservation laws*, Math. Comp. **40** (1983), 91-106.

[12] H. Nessyahu and E. Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comp. Phys. **87** (1990), 408-463.

[13] H. Nessyahu and E. Tadmor, *The convergence rate of approximate solutions for nonlinear scalar conservation laws*, SIAM J. Numer. Anal. **29** (1992), 1505-1519.

[14] H. Nessyahu, E. Tadmor and T. Tassa, *The convergence rate of Godunov type schemes*, SIAM J. Numer. Anal. **31** (1994), 1-16.

[15] S. Osher and E. Tadmor, *On the convergence of difference approximations to scalar conservation laws*, Math. Comp. **50** (1988), 19-51.

[16] J.-P. Vila, *Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws*, Math. Mod. Numer. Anal. **28** (1994), 267-295.