High order positivity-preserving discontinuous Galerkin methods for steady state or implicit time discretization of hyperbolic equations

Chi-Wang Shu

Division of Applied Mathematics

Brown University

Joint work with Tong Qin, and with Juan Cheng, Dan Ling and Daming Yuan

Outline

- Introduction
- Framework for explicit time stepping
- Implicit time stepping: general boundary conditions
- Implicit time stepping: inflow boundary conditions
- Conclusions and future work

Introduction

We are interested in numerically solving hyperbolic conservation laws

$$u_t + \nabla \cdot \mathbf{F}(u) = 0, \qquad u(\mathbf{x}, 0) = u_0(\mathbf{x}) \tag{1}$$

or other related hyperbolic or convection dominated equations. In particular, we are interested in the bound-preserving properties of high order numerical schemes.

We assume the exact solution of the PDE (1) has a convex invariant region G:

- If $u(\cdot, 0) \in G$, then $u(\cdot, t) \in G$ for all $t > 0$.

For a convex region G, if $u_1, \cdots, u_m \in G$, $\alpha_i \geq 0$, $\sum_{i=1}^m \alpha_i = 1$, then $u = \sum_{i=1}^m \alpha_i u_i \in G$. We will heavily use this property when building our high order bound-preserving schemes.

Several examples:

- If (1) is a scalar conservation law, an important property of the entropy solution (which may be discontinuous) is that it satisfies a strict maximum principle: If

$$M = \max_{\mathbf{x}} u_0(\mathbf{x}), \quad m = \min_{\mathbf{x}} u_0(\mathbf{x}), \tag{2}$$

  then $u(\mathbf{x}, t) \in [m, M]$ for any $\mathbf{x}$ and t.

  Therefore, $G = [m, M]$ is an invariant region. It is clearly convex.

- For the compressible Euler equations:

$$u_t + f(u)_x = 0$$

with

$$u = \begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix}, \qquad f(u) = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ v(E+p) \end{pmatrix},$$

where $E = e + \frac{1}{2}\rho v^2$. The internal energy e is related to density and pressure through the equation of states (EOS). For the ideal gas, we have $e = \frac{p}{\gamma - 1}$ with $\gamma = 1.4$ for air.

In this case, we can verify that the set

$$G = \{u : \rho \geq 0, e \geq 0\} \tag{3}$$

is invariant. It is also easy to check that G is convex (for this we need to check that the internal energy e is a concave function of the conservative variable u, then Jensen's inequality implies the convexity of G).

For many EOS, e.g. that for the ideal gas, the region G defined in (3) is equivalent to

$$G = \{u : \rho \geq 0, p \geq 0\}.$$

- Consider the relativistic hydrodynamics

$$u_t + f(u)_x = 0$$

with

$$u = \begin{pmatrix} D \\ m \\ E \end{pmatrix}, f(u) = \begin{pmatrix} Dv \\ mv + p \\ m \end{pmatrix}$$

where p, D, m and E are the thermal pressure, mass density, momentum and energy, respectively. v is the velocity. Moreover, units are normalized such that the speed of light is $c = 1$.

If we denote ρ to be the proper rest-mass density, then the conservative variable u can be written as

$$\begin{aligned} D &= \gamma\rho, \\ m &= Dh\gamma v, \\ E &= Dh\gamma - p, \end{aligned}$$

where $\gamma = (1 - v^2)^{-1/2}$ is the Lorentz factor and h is the specific enthalpy. To close the system, we specify an equation of state $h = h(p, \rho)$. For ideal gas

$$\rho h = \rho + p\Gamma/(\Gamma - 1)$$

with Γ being the specific heat ratio, such that $1 < \Gamma \leq 2$

It can be shown that the density D and pressure p are positive, and the velocity satisfies $v^2 \leq 1$, if they are initially in these cases. Therefore,

$$G = \{u : D > 0, E > 0, p > 0, v^2 \leq 1\}$$

is an invariant region. It is convex and can be represented as

$$G = \{\mathbf{w} : D > 0, E > \sqrt{D^2 + m^2}\}.$$

Wu and Tang, JCP 2015 and Qin, Shu and Yang, JCP 2016.

Framework for explicit time stepping

It is of course desirable to have the invariant region G also to be an invariant region for the numerical solution. That is, we wish that, if the initial condition $u(\cdot, 0) \in G$ then $u(\cdot, t) \in G$ for later time $t > 0$. This time u stands for the numerical solution.

We first review a framework to achieve this under explicit time stepping. The framework consists of the following ingredients:

- We start with a first order scheme which can be proved to be bound preserving under a certain CFL condition

$$\lambda = \frac{\Delta t}{\Delta x} \leq \lambda_0.$$

For example, for the scalar conservation laws

$$u_t + f(u)_x + g(u)_y = 0,$$

the first order monotone schemes satisfy maximum principles; for Euler equations of compressible gas dynamics, many first order schemes, including Godunov, Lax-Friedrichs, kinetic and HLLC schemes are proved to be positivity-preserving for density and internal energy; for relativistic hydrodynamics, the first order Lax-Friedrichs scheme is bound-preserving for the invariant region G mentioned above (Wu and Tang, JCP 2015 and Qin, Shu and Yang, JCP 2016).

We emphasize that sometimes it is already non-trivial to find first order schemes which are bound-preserving, e.g. for MHD equations. Since our high order bound-preserving schemes discussed later are built upon first order bound-preserving schemes, the very first task when one would like to solve a new PDE is to find a first order bound-preserving scheme.

- We will call a high order finite volume or DG scheme bound-preserving, if we have

$$m \leq u^{n+1}(x) \leq M, \qquad \forall x$$

provided

$$m \leq u^{n}(x) \leq M, \qquad \forall x.$$

Here $u^n(x)$ refers to the piecewise polynomial numerical solution, either reconstructed in a finite volume scheme or evolved for a DG scheme.

A suitable modification to evaluate the bounds only at certain quadrature points will be given later to facilitate easy implementation.

The flowchart for designing a high order finite volume or DG bound-preserving scheme (we use maximum principle in scalar case below as an example) is as follows:

1. Start with $u^n(x)$ which is high order accurate

$$|u(x,t^n) - u^n(x)| \leq C\Delta x^p$$

and satisfies

$$m \leq u^n(x) \leq M, \qquad \forall x$$

therefore of course we also have

$$m \leq \bar{u}_j^n \leq M, \qquad \forall j.$$

2. Evolve for one time step to get

$$m \leq \bar{u}_j^{n+1} \leq M, \qquad \forall j. \tag{4}$$

   This is usually the most difficult step, since it must be proved for the unmodulated scheme.

3. Given (4) above, use a simple scaling limiter to obtain $u^{n+1}(x)$ (reconstruction or evolution) which
   - satisfies the maximum principle

$$m \leq u^{n+1}(x) \leq M, \qquad \forall x;$$

   - is high order accurate

$$|u(x, t^{n+1}) - u^{n+1}(x)| \leq C\Delta x^p.$$

In Zhang and Shu, JCP 2010a, a procedure is designed to obtain

$$m \leq \bar{u}_j^{n+1} \leq M, \qquad \forall j$$

with simple Euler forward or SSP Runge-Kutta or multi-step methods using the unmodulated finite volume or DG scheme:

The evolution of the cell average for a higher order finite volume or DG scheme satisfies

$$\begin{aligned} \bar{u}_j^{n+1} &= G(\bar{u}_j^n, u^-_{j-\frac{1}{2}}, u^+_{j-\frac{1}{2}}, u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}}) \\ &= \bar{u}_j^n - \lambda[h(u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}}) - h(u^-_{j-\frac{1}{2}}, u^+_{j-\frac{1}{2}})], \end{aligned}$$

where

$$G(\uparrow, \uparrow, \downarrow, \downarrow, \uparrow)$$

therefore there is no maximum principle.

The polynomial $p_j(x)$ (either reconstructed in a finite volume method or evolved in a DG method) is of degree k, defined on I_j such that $\bar{u}_j^n$ is its cell average on I_j, $u^+_{j-\frac{1}{2}} = p_j(x_{j-\frac{1}{2}})$ and $u^-_{j+\frac{1}{2}} = p_j(x_{j+\frac{1}{2}})$.

We take a Legendre Gauss-Lobatto quadrature rule which is exact for polynomials of degree k, then

$$\bar{u}_j^n = \sum_{\ell=0}^{m} \omega_\ell p_j(y_\ell)$$

with $y_0 = x_{j-\frac{1}{2}}$, $y_m = x_{j+\frac{1}{2}}$. The scheme for the cell average is then rewritten as

$$
\begin{aligned}
\bar{u}_j^{n+1} &= \omega_m \left[u^-_{j+\frac{1}{2}} - \frac{\lambda}{\omega_m} \left(h(u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}}) - h(u^+_{j-\frac{1}{2}}, u^-_{j+\frac{1}{2}}) \right) \right] \\
&\quad + \omega_0 \left[u^+_{j-\frac{1}{2}} - \frac{\lambda}{\omega_0} \left(h(u^+_{j-\frac{1}{2}}, u^-_{j+\frac{1}{2}}) - h(u^-_{j-\frac{1}{2}}, u^+_{j-\frac{1}{2}}) \right) \right] \\
&\quad + \sum_{\ell=1}^{m-1} \omega_\ell p_j(y_\ell) \\
&= \omega_m H_{\lambda/\omega_m}(u^+_{j-\frac{1}{2}}, u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}}) + \omega_0 H_{\lambda/\omega_0}(u^-_{j-\frac{1}{2}}, u^+_{j-\frac{1}{2}}, u^-_{j+\frac{1}{2}}) \\
&\quad + \sum_{\ell=1}^{m-1} \omega_\ell p_j(y_\ell).
\end{aligned}
$$

where $H_\lambda(a, b, c)$ is the first order monotone operator which satisfies the maximum principle when $\lambda \leq \lambda_0$.

Therefore, if

$$m \leq p_j(y_\ell) \leq M$$

at all Legendre Gauss-Lobatto quadrature points and a reduced CFL condition

$$\lambda/\omega_m = \lambda/\omega_0 \leq \lambda_0$$

is satisfied, then

$$m \leq \bar{u}_j^{n+1} \leq M.$$

Now, given

$$m \leq \bar{u}_j^{n+1} \leq M, \qquad \forall j$$

for the unmodulated finite volume or DG scheme, again in Zhang and Shu, JCP 2010a, a procedure is designed to obtain a bound-preserving $u^{n+1}(x)$ with a very simple scaling limiter, which only requires the evaluation of unmodulated $u^{n+1}(x)$ at certain pre-determined quadrature points and does not destroy accuracy:

We replace $p_j(x)$ by the limited polynomial $\tilde{p}_j(x)$ defined by

$$\tilde{p}_j(x) = \theta_j(p_j(x) - \bar{u}_j^n) + \bar{u}_j^n$$

where

$$\theta_j = \min\left\{\left|\frac{M - \bar{u}_j^n}{M_j - \bar{u}_j^n}\right|, \left|\frac{m - \bar{u}_j^n}{m_j - \bar{u}_j^n}\right|, 1\right\},$$

with

$$M_j = \max_{x \in S_j} p_j(x), \qquad m_j = \min_{x \in S_j} p_j(x)$$

where S_j is the set of Legendre Gauss-Lobatto quadrature points of cell I_j. Clearly, this limiter is just a simple scaling of the original polynomial around its average.

The following lemma, guaranteeing the maintenance of accuracy of this simple limiter, is proved in Zhang and Shu, JCP 2010a:

Lemma: Assume $\bar{u}_j^n \in [m, M]$ and $p_j(x)$ is an $O(\Delta x^p)$ approximation, then $\tilde{p}_j(x)$ is also an $O(\Delta x^p)$ approximation.

We have thus obtained a high order accurate scheme satisfying the following maximum principle: If

$$m \leq u^n(x) \leq M, \qquad \forall x \in S_j,$$

then

$$m \leq u^{n+1}(x) \leq M, \qquad \forall x \in S_j.$$

Recall that S_j is the set of Legendre Gauss-Lobatto quadrature points of cell I_j.

Clearly, the framework easily generalizes to 2D (or higher dimensions) in structured or unstructured meshes, and to high order time accuracy through explicit SSP (also called TVD) Runge-Kutta or multi-step time discretizations.

Implicit time stepping: general boundary conditions

Many first order implicit schemes (e.g. for scalar equations, first order monotone schemes with backward Euler time discretization) can be proved to be unconditionally bound-preserving. Unfortunately, this does not easily generalize to higher order schemes.

In (Qin and Shu, SISC 2018), we have developed a general framework to obtain high order bound-preserving finite volume or DG schemes with backward Euler time discretization, which consists of the following ingredients:

- We first establish the following property:

  Property: If the numerical solution $u^n(x)$ is within the desired bounds at the Gauss-Lobatto quadrature points, then the cell averages $\bar{u}_j^{n+1}$ of the numerical solution for the unmodulated high order finite volume or DG scheme is within the desired bounds under the following lower bound of the CFL condition

$$\lambda \geq \lambda_0. \tag{5}$$

  This property is rigorously proved for the linear equations and demonstrated numerically for nonlinear equations including Euler equations of compressible gas dynamics.

- Given that the cell averages $\bar{u}_j^{n+1}$ of the numerical solution for the unmodulated high order finite volume or DG scheme is within the desired bounds, we may use the same scaling limiter as described above to obtain a high order accurate solution $u^{n+1}(x)$ (without changing the cell averages) within the desired bounds at the Gauss-Lobatto quadrature points.

Unlike in the explicit case, here the lower bound in (5) actually improves (becomes smaller) with higher order of accuracy:

Table 1: Values of λ_0 for polynomial degrees $k = 1, \ldots, 5$ in the finite volume or DG scheme, and for Legendre-Gauss-Lobatto (LGL) and Legendre-Gauss quadrature rules respectively.

	LGL rule		LG rule	
k	N_q	λ_0	N_q	r^k
1	3	0.333	2	0.333
2	4	0.262	3	0.344
3	5	0.177	4	0.177
4	6	0.177	5	0.212
5	7	0.121	6	0.121

Example 1. Steady-state solution to Burgers' problems.

Table 2: Error table for the approximation of the steady state solution to the Burgers' equation without the positivity preserving limiter.

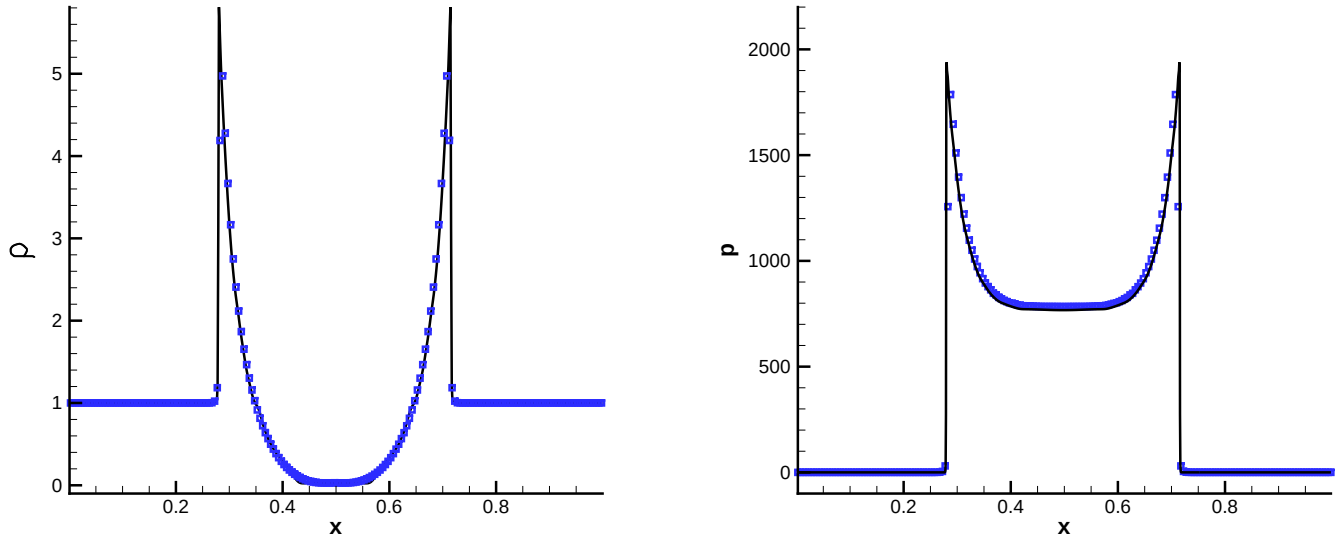
k	N	L^2 error	order	L^∞ error	order	$\min u_h$
2	20	2.915 E-6	–	5.085 E-6	–	-8.073 E-6
	40	3.508 E-7	3.05	6.358 E-7	3.00	-1.009 E-6
	80	4.293 E-8	3.03	7.948 E-8	3.00	-1.261 E-7
	160	5.304 E-9	3.02	9.934 E-9	3.00	-1.577 E-8
3	20	2.336 E-9	–	1.648 E-9	–	-3.855 E-10
	40	1.445 E-10	4.02	1.030 E-10	4.00	-1.204 E-11
	80	9.010 E-11	4.00	6.525 E-11	3.98	-3.763 E-13
	160	6.031 E-12	3.90	4.978 E-12	3.71	-1.175 E-14
4	20	3.403 E-10	–	6.312 E-10	–	-1.497 E-9
	40	1.010 E-11	5.07	1.970 E-11	5.00	-4.678 E-11
	80	3.116 E-13	5.02	5.398 E-13	5.19	-1.462 E-12

Table 3: Error table for the approximation of the steady state solution to the Burgers' equation with the positivity preserving limiter.

k	N	L^2 error	order	L^∞ error	order	$\min u_h$
2	20	3.207 E-6	–	4.408 E-6	–	1.000 E-13
	40	3.696 E-7	3.12	6.010 E-7	2.87	1.000 E-13
	80	4.435 E-8	3.06	8.171 E-8	2.88	1.000 E-13
	160	5.414 E-9	3.03	1.083 E-8	2.92	1.000 E-13
3	20	2.338 E-9	–	1.648 E-9	–	1.000 E-13
	40	1.445 E-10	4.02	1.030 E-10	4.00	1.000 E-13
	80	9.010 E-11	4.00	6.511 E-11	3.98	1.000 E-13
	160	6.031 E-12	3.90	4.987 E-12	3.71	1.051 E-14
4	20	4.792 E-10	–	9.061 E-10	–	1.000 E-13
	40	1.253 E-11	5.26	3.102 E-11	4.87	1.000 E-13
	80	3.653 E-13	5.10	1.086 E-12	4.84	1.000 E-13

Example 2. Sedov point-blast wave.

Initially the gas is steady with uniform density one in the whole domain. The pressure is set to be $p = 10^{-9}$, except in the central cell, where the pressure is as high as $p = 10^4$. Then a blast-wave starts to propagate from the central cell with a shock front. In Figure 1 we present the numerical approximation and the exact solution for the pressure and density respectively. The positivity-preserving limiter not only helps keep the low-density region positive, but also add robustness to the nonlinear solver. Without it, the code breaks down due to the failure of convergence of the nonlinear solver.



(a) density (b) Pressure

Figure 1: Sedov point-blast wave. $T = 0.003$, with $N = 200$ and $CFL = 2$. With the positivity-preserving limiter on. Solid line is the exact solution and blue squares are the numerical approximations.

Implicit time stepping: inflow boundary conditions

We consider the following linear steady hyperbolic equation,

$$\alpha u_x + \gamma u = f, \tag{6}$$

where α and $\gamma \geq 0$ are constants and $f \geq 0$ is the source term. Without loss of generality we assume $\alpha > 0$, in this case the boundary condition is given at the left boundary for the existence and uniqueness of the solution. Notice that the time dependent case can be treated either by space-time DG (adding one dimension), or by backward Euler time discretization which reduces to the same form as (6).

The earliest DG scheme in Reed and Hill, Los Alamos Technical Report 1973 was designed for solving the neutron transport equation, which is essentially (6). Given the boundary condition, we can obtain the polynomial solution u in the first cell. Then, given the "boundary condition" for the second cell (provided by the numerical solution in the first cell), we can obtain the polynomial solution u in the second cell, etc.

In Yuan, Cheng and Shu, SISC 2016, we have proved that the unmodulated DG solution $u(x)$ is positive at one point inside the cell I_j, which is a convex combination of the cell average and the downwind cell boundary value of $u(x)$, provided that the inflow boundary condition and the source term are both positive.

Even though the exact location of this point is in general not known, it does not prevent the design of a positivity-preserving limiter, which is a combination of the scaling limiter mentioned above when the cell average is positive, and a rotational limiter designed in Yuan, Cheng and Shu, SISC 2016 when the cell average is negative. This combined limiter can be proved to maintain the original high order accuracy, but it may affect conservation when the original cell averages are altered through the rotational limiter (in the sense that the cell average from the unmodulated DG solver is altered in the limiting process). A Lax-Wendroff type theorem is proved in Yuan, Cheng and Shu, SISC 2016 to guarantee convergence to weak solutions (with correct shock speeds).

In Ling, Cheng and Shu, JSC to appear, we prove that, for the one-dimensional case (6), the cell average of the unmodulated DG solution $u(x)$ is positive, provided that the inflow boundary condition and the source term are both positive. Even though this property is not valid for P^k or Q^k DG solutions in 2D, we have found augmented DG spaces (i.e. by adding additional basis function(s) into the DG space) so that the cell average of the unmodulated DG solution $u(x)$ is positive. Explicit form of this augmented DG space for the P^1 case in rectangular meshes is given in Ling, Cheng and Shu, JSC to appear. With a positive cell average, we can then use just the scaling limiter and avoid the rotational limiter, hence obtaining a positivity-preserving conservative high order DG scheme.

Conclusions and future work

- We have designed bound-preserving high order finite volume or DG schemes under a lower bound of CFL conditions, for general boundary conditions of conservation laws. Rigorous analysis is performed only for linear PDEs but numerical experiments are performed for nonlinear equations including systems.
- We have designed bound-preserving high order DG schemes for inflow boundary conditions. The schemes work for any polynomial degree k in 1D and for augmented DG spaces in 2D.

- Future work includes:
  - Analysis for more general cases;
  - Generalization to convection-diffusion equations;
  - Design of augmented DG spaces for higher order of accuracy in 2D for both rectangular and triangular meshes.

References:

[1] T. Qin and C.-W. Shu, *Implicit positivity-preserving high order discontinuous Galerkin methods for conservation laws*, SIAM Journal on Scientific Computing, v40 (2018), pp.A81-A107.

[2] D. Yuan, J. Cheng and C.-W. Shu, *High order positivity-preserving discontinuous Galerkin methods for radiative transfer equations*, SIAM Journal on Scientific Computing, v38 (2016), pp.A2987-A3019.

[3] D. Ling, J. Cheng and C.-W. Shu, *Conservative high order positivity-preserving discontinuous Galerkin methods for linear hyperbolic and radiative transfer equations*, Journal of Scientific Computing, to appear.

THANK YOU!