Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection-diffusion equations on triangular meshes\textsuperscript{1}

Yifan Zhang\textsuperscript{2}, Xiangxiong Zhang\textsuperscript{3} and Chi-Wang Shu\textsuperscript{4}

Abstract

We propose second order accurate discontinuous Galerkin (DG) schemes which satisfy a strict maximum principle for general nonlinear convection-diffusion equations on unstructured triangular meshes. Motivated by genuinely high order maximum-principle-satisfying DG schemes for hyperbolic conservation laws [14, 26], we prove that under suitable time step restriction for forward Euler time stepping, for general nonlinear convection-diffusion equations, the same scaling limiter coupled with second order DG methods preserves the physical bounds indicated by the initial condition while maintaining uniform second order accuracy. Similar to the purely convection cases, the limiters are mass conservative and easy to implement. Strong stability preserving (SSP) high order time discretizations will keep the maximum principle. Following the idea in [30], we extend the schemes to two-dimensional convection-diffusion equations on triangular meshes. There are no geometric constraints on the mesh such as angle acuteness. Numerical results including incompressible Navier-Stokes equations are presented to validate and demonstrate the effectiveness of the numerical methods.

Keywords: discontinuous Galerkin method; maximum principle; positivity preserving; convection-diffusion equations; incompressible Navier-Stokes equations; degenerate parabolic equations; triangular meshes

\textsuperscript{1}Research supported by DOE grant DE-FG02-08ER25863 and NSF grant DMS-1112700.

\textsuperscript{2}Division of Applied Mathematics, Brown University, Providence, RI 02912. E-mail: yifan.zhang@brown.edu

\textsuperscript{3}Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139. E-mail: zhangxx@mit.edu

\textsuperscript{4}Division of Applied Mathematics, Brown University, Providence, RI 02912. E-mail: shu@dam.brown.edu
1 Introduction

1.1 Motivation

In this paper we are interested in constructing maximum-principle-satisfying discontinuous Galerkin (DG) schemes for nonlinear convection-diffusion equations:

\[ u_t + f(u)_x = a(u)_{xx}, \quad u(x, 0) = u_0(x), \]  

or equivalently,

\[ u_t + f(u)_x = (b(u)u_x)_x, \quad u(x, 0) = u_0(x), \]

where \( b(u) = a'(u) \geq 0 \), and their two-dimensional versions. The exact solution of (1.1) or (1.2) satisfies a strict maximum principle, i.e., if \( \max u_0(x) = M, \min u_0(x) = m \), then

\[ u(x, t) \in [m, M], \quad \forall x \in \mathbb{R}, \quad \forall t \geq 0. \]  

In particular, for the case \( m = 0 \), it is the positivity-preserving property, which is preferred for numerical schemes in many applications. For physical quantities like densities and probability distributions, negative values are non-physical in the numerical simulations, for instance, radionuclide transport calculations [9], and may lead to ill-posedness and instability for certain nonlinear equations such as chemotaxis problems [18].

The discrete maximum principle (DMP) for numerical schemes solving elliptic type equations were studied since 1960s. Even though DMP is more difficult to achieve for parabolic type equations, maximum-principle-satisfying finite element methods were also well studied in a similar fashion. The earliest work about DMP for parabolic equations is [8] in which Fujii investigated DMP for the heat equation using piecewise linear finite elements. For the most recent results on nonlinear parabolic equations, see [7] and the references therein. The main techniques to achieve DMP for parabolic equations in the literature appear to be mainly algebraic, for instance, the lumped mass method. Such approaches tend to pose certain restrictions on the spatial mesh such as angle acuteness [19] and also the rather unnatural time step restriction that the time step must be larger than a threshold, in order to use the larger numerical viscosity for implicit time discretization with larger time steps.
On the other hand, to have a maximum-principle-satisfying scheme for a convection-diffusion problem (1.1) which has a significant or dominate convection part, one would rather use an explicit time discretization and would first need to construct such schemes for the purely convection case. It turns out that one can construct arbitrarily high order maximum-principle-satisfying DG schemes solving scalar conservation laws on unstructured triangular meshes [26, 30]. The most interesting observations in [26, 30] consist of two parts. First, if an intuitive reasonable sufficient condition is satisfied, i.e., the values of the DG polynomials at the current time level and at certain quadrature points are in the desired range \([m, M]\), then the cell averages of the numerical solutions at the next time step will also be in this range under suitable time step restrictions with an explicit Euler forward time stepping. Second, a simple scaling limiter can enforce the sufficient condition of keeping the DG polynomials at these quadrature points in the desired bounds, once we know their cell averages are already in these bounds, without destroying accuracy and conservation. High order accuracy in time can be achieved by strong stability preserving Runge-Kutta or multi-step methods, which are convex combinations of Euler forward steps. Thus DG methods with this simple scaling limiter will satisfy the maximum principle for scalar conservation laws.

It would be appealing if one can extend the results [26, 30] to convection-diffusion equations in a straightforward way since the method is very easy to implement without any constraint on the spatial mesh. Unfortunately, it is quite difficult to do so, at least for arbitrarily high order schemes. We will explain in more detail about this difficulty in the next subsection. In this paper, we will discuss the generalization of this technique for piecewise linear DG methods in solving general nonlinear convection-diffusion problems.

1.2 The methodology

A maximum-principle-satisfying DG scheme for (1.1) can be constructed in two steps. First we consider the purely convection case:

\[
  u_t + f(u)_x = 0, \quad u(x, 0) = u_0(x).
\]
Motivated by [14], a general framework to construct genuinely high order maximum principle satisfying DG scheme for (1.4) was proposed in [26, 30]. It was the first time that genuinely high order maximum-principle-satisfying finite volume and DG schemes for multidimensional nonlinear scalar conservation laws on arbitrary triangular meshes became available. This methodology has been successfully applied to various problems [22, 27, 20, 29, 23, 24, 15, 17, 3] for positivity-preserving of physically relevant properties such as density, pressure and water height, where bound preserving is a desirable property for applications. For a survey of the recent developments of such schemes and applications, see [28].

To illustrate the ideas regarding maximum-principle-satisfying DG schemes for (1.4) in [26], we consider the forward Euler time discretization. On the mesh \( x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{\frac{N-1}{2}} < x_{\frac{N+1}{2}} \), the equation satisfied by the cell averages of the numerical solutions obtained for the DG scheme is:

\[
\bar{u}_j^{n+1} = \bar{u}_j^n - \Delta t \frac{h}{r} \left[ \hat{f} \left( u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+ \right) - \hat{f} \left( u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+ \right) \right] \tag{1.5}
\]

where \( h = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}} \) is the mesh size, assumed to be uniform for simplicity, \( n \) refers to the time level and \( j \) denotes the spatial cell, and \( \bar{u}_j \) is the numerical approximation to the cell average obtained by the DG scheme. \( u_{j+\frac{1}{2}}^\pm \) are the values of the numerical solution at the point \( x_{j+\frac{1}{2}} \) evaluated from the cell \( I_{j+1} \) and \( I_j \) respectively at time level \( n \). \( \hat{f}(\cdot, \cdot) \) is a monotone flux, which is an increasing function of the first argument and a decreasing function of the second argument, for example, the global Lax-Friedrichs flux.

The key ingredients in constructing a high order maximum-principle-satisfying DG scheme for (1.4) in [26] are:

- The cell average at the next time step by forward Euler time stepping, i.e., (1.5) is a monotone function with respect to certain point values (Gauss-Lobatto points for the one-dimensional case), under a suitable CFL condition.

- A simple limiter modifies the DG polynomial \( p_j(x) \) on cell \( I_j \) into \( \tilde{p}_j(x) \) such that \( \tilde{p}_j(x) \in [m, M] \) at these special points without changing its cell average. Moreover, it
can be proven that the modified polynomial $\tilde{p}_j(x)$ is also a high order approximation just as $p_j(x)$. Thus we have $\tilde{u}^{n+1} \in [m, M]$ if all the degree of freedoms at time level $n$ in the right hand side of (1.5) are replaced by those of modified polynomials $\tilde{p}_j(x)$.

- Replace the forward Euler by the strong stability preserving (SSP) high order time discretizations which are convex combinations of forward Euler thus will keep the bounds. So it suffices to prove maximum principle for (1.5).

Therefore we reach the conclusion that the DG scheme with the simple scaling limiter is high order accurate and will satisfy the maximum principle in the sense that the cell averages and point values at certain quadrature points will not go out of the range $[m, M]$.

To extend the idea above to convection-diffusion equations (1.1) or (1.2), it suffices to consider the purely diffusion equations with $f(u) = 0$. However, it appears very difficult to seek a direct generalization for arbitrary high order DG schemes. The main difficulty lies in the first step above, namely to write the formula for obtaining the cell average at the next time step by forward Euler time stepping, in terms of certain point values and to ensure monotonicity with respect to these point values, which is a key ingredient in [26]. To achieve arbitrary high order approximation, a non-conventional maximum-principle-satisfying finite volume scheme for convection-diffusion equations was developed in [25]. The main idea is to define a quantity called double cell average over the interval $I_j$, by utilizing the cell averaging operator twice on the original solution. Unlike conventional finite volume schemes whose numerical solutions are cell averages, in [25], the high order approximations are reconstructed from those double cell averages. It turns out that the formula for obtaining those double cell averages at the next time step by forward Euler time stepping can be written as monotone functions with respect to point values at certain quadrature points, just like in the purely convective cases. It is then straightforward to apply the methodology in [26] including the scaling limiter to ensure a strict maximum principle for the numerical solutions while keeping high order accuracy. Even though this finite volume method suggests a possible alternative when looking for high order maximum-principle-satisfying numerical schemes for
convection-diffusion equations, it is an unconventional finite volume method requiring new implementations. It is also not obvious how to generalize the technique to finite volume methods on arbitrary unstructured meshes, or to DG methods.

The main content of this paper is to extend the framework in [26, 30] to second order accurate $P^1$-DG schemes solving one- and two-dimensional convection-diffusion equations on arbitrary triangular meshes. The second order accurate maximum-principle-satisfying DG scheme designed in this paper share all the good properties as those in [26, 30] including mass conservation and easiness of implementation. We also discuss applications of this scheme to nonlinear degenerate parabolic equations and incompressible Navier-Stokes equations.

We remark that most schemes satisfying DMP for parabolic equations in the literature used implicit time discretizations. But the small time step $\Delta t = O(h^2)$ turns out to be necessary even for implicit time discretizations to have the contraction property, see [6]. We use explicit time discretizations which are advantageous for solving nonlinear equations, especially for convection dominated problems. Moreover, our method is straightforward to implement. The limiter is a local operator on each cell. In practice, for each time step in a multistep method or each time stage in a Runge-Kutta method, one only needs to add the limiter to the DG scheme as a pre-processing step. See Section 3.3 for an easy implementation.

1.3 Organization of the paper

This paper is organized as follows: we first discuss maximum-principle-satisfying second order DG schemes in one space dimension in Section 2. In Section 3, we show an extension to two space dimensions on arbitrary triangular meshes. In Sections 4 and 5, numerical tests for the maximum-principle-satisfying DG schemes in one and two dimensions will be shown respectively. In Section 6, we discuss the application of this maximum-principle-satisfying DG scheme to two-dimensional incompressible Navier-Stokes equations in the vorticity stream-function formulation. Concluding remarks are given in Section 7.
2 Maximum-principle-satisfying DG schemes in one dimension

For simplicity, we will assume periodic boundary condition from now on. All the results regarding maximum principle can be easily incorporated into any initial-boundary value problem.

2.1 Preliminaries

We first review briefly how to establish the monotonicity of (1.5).

The equation satisfied by the cell averages of a DG method with forward Euler time discretization for a one-dimensional scalar conservation laws (1.4) can be written as (1.5). Any monotone numerical fluxes can be used, for example the global Lax-Friedrichs flux

\[ \tilde{f}(u, v) = \frac{1}{2} (f(u) + f(v) - a(v - u)), \quad a = \max_u |f'(u)|. \]  

(2.1)

The important property we will use for the monotone flux \( \tilde{f}(u, v) \) is that it is an increasing function (a non-decreasing function) of \( u \) and a decreasing function of \( v \), hence the first order scheme

\[ H_\lambda(u, v, w) = v - \lambda(\tilde{f}(v, w) - \tilde{f}(u, v)) \]  

(2.2)

is a monotonically increasing function of all three arguments \( u, v \) and \( w \) provided a CFL condition is satisfied. For the Lax-Friedrichs flux (2.1) the CFL condition is

\[ \lambda a \leq 1, \]  

(2.3)

where \( \lambda = \frac{\Delta t}{h} \).

Let the piecewise polynomials \( p_j(x) \) of degree \( k \) denote the DG polynomials at time level \( n \) and \( \bar{u}_j \) denote the cell averages of \( p_j(x) \). \( u_{j+1/2}^{\pm} \) are the values of the numerical solution at the point \( x_{j+1/2} \) evaluated from the cell \( I_{j+1} \) and \( I_j \) respectively at time level \( n \). We consider an \( N \)-point Legendre Gauss-Lobatto quadrature rule on the interval \( I_j = [x_{j-1/2}, x_{j+1/2}] \), which is exact for the integral of polynomials of degree up to \( 2N - 3 \). We denote these quadrature
2.2 Maximum-principle-satisfying DG schemes for convection-diffusion

There exist several different types of DG formulations to treat the second order viscous
terms. In this paper, for the one dimensional case, we focus on three different formulations:
the DG formulation of Cheng and Shu for higher order derivatives [2], the interior penalty
DG (IPDG) scheme [21, 1, 16] and the local DG (LDG) scheme [5]. We start with the
Cheng-Shu DG scheme to take care of the second order derivatives in [2]. For (1.1), their scheme can be described as: For a computational mesh of the domain \([a, b]\): \(a = x_\frac{1}{2} < x_\frac{3}{2} < \cdots < x_{N-\frac{1}{2}} < x_{N+\frac{1}{2}} = b\), where the mesh size \(h = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}\) is assumed to be uniform here only for simplicity, define the approximation space \(V_h^k = \{ v : v|_{I_j} \in P_k(I_j), \forall j \}\), where \(P_k(I_j)\) denotes the set of all polynomials up to degree \(k\) defined on the cell \(I_j\). For any test function \(v \in V_h^k\), the DG formulation is to find \(u \in V_h^k\) such that,

\[
\int_{I_j} u_t v dx - \int_{I_j} f(u)v_x dx - \int_{I_j} a(u)v_{xx} dx + \hat{f}_{j+\frac{1}{2}} v^+_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} v^-_{j-\frac{1}{2}} - \hat{a}_{x_j+\frac{1}{2}} v^+_{j+\frac{1}{2}} + \hat{a}_{x_j-\frac{1}{2}} v^-_{j-\frac{1}{2}} = 0, \tag{2.7}
\]

where the convection flux \(\hat{f}_{j+\frac{1}{2}} = \hat{f}(u_{j+\frac{1}{2}}, u_{j+\frac{1}{2}})\) is chosen as described in the previous subsection, and the diffusion fluxes are chosen in an alternating way with additional penalty on the second order derivative term:

\[
\hat{a}_x(u) = \frac{[a(u)]}{[u]} \left( (u)^-_x + \frac{\alpha}{h} [u] \right), \quad \hat{a}(u) = a(u^+). \tag{2.8}
\]

Here \(\alpha\) is a sufficiently large positive coefficient, and \([u] = u^+ - u^-\) denotes the jump across the cell boundaries. It is proved in [2] that, by this treatment of the diffusion fluxes (2.8), the DG scheme (2.7) is stable with a sub-optimal error estimate. Numerical experiments indicate optimal \((k+1)\)-th order convergence in \(L^2\).

By taking the test function \(v = 1\), we obtain the evolutionary equation for the cell averages of the numerical scheme:

\[
\frac{d}{dt} \bar{u}_j = -\frac{1}{h} \left( (\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}) - \hat{a}_{xj+\frac{1}{2}} - \hat{a}_{xj-\frac{1}{2}} \right). \tag{2.9}
\]

We consider the first order forward Euler time discretization for (2.9); higher order time discretization will be discussed later. Our first result is:

**Theorem 2.1.** For the \(P^1\)-DG scheme (2.7) with forward Euler time stepping, the maximum principle holds for the cell averages \(\bar{u}_j^{n+1} \in [m, M]\) for all \(j\), if \(u_{j+\frac{1}{2}}^\pm \in [m, M]\) for all \(j\) and the following CFL conditions hold

\[
\alpha \geq 1, \quad \lambda \max_u |f'(u)| \leq \frac{1}{4}, \quad \mu(\alpha + 1) \max_u (a'(u)) \leq \frac{1}{4} \tag{2.10}
\]
where $\mu = \frac{1}{h} = \frac{\Delta t}{\Delta x}$.

Proof. Applying forward Euler time stepping on (2.9), we obtain the formula for the cell averages of the numerical solution at the next time step

$$\bar{u}^{n+1} = \bar{u}^n - \lambda \left( \hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} \right) + \lambda \left( (\hat{a}_x)_{j+\frac{1}{2}} - (\hat{a}_x)_{j-\frac{1}{2}} \right).$$  \hfill (2.11)

To utilize the result in the purely convection case [26], we divide the cell average $\bar{u}_j^n$ into two halves:

$$\bar{u}_j^{n+1} = \left( \frac{1}{2} \bar{u}_j^n - \lambda \left( \hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} \right) \right) + \left( \frac{1}{2} \bar{u}_j^n + \lambda ((\hat{a}_x)_{j+\frac{1}{2}} - (\hat{a}_x)_{j-\frac{1}{2}}) \right).$$  \hfill (2.12)

This decomposition may not lead to the optimal CFL restriction for the proof but is adopted for easy presentation.

The diffusion term:

$$D_j = \frac{1}{2} \bar{u}_j^n + \lambda ((\hat{a}_x)_{j+\frac{1}{2}} - (\hat{a}_x)_{j-\frac{1}{2}}).$$  \hfill (2.13)

By the linearity of the approximation space with $k = 1$:

$$\bar{u}_j = \frac{u^-_{j+\frac{1}{2}} + u^+_{j-\frac{1}{2}}}{2}, \quad (u_x)_{j+\frac{1}{2}} = \frac{u^-_{j+\frac{1}{2}} - u^+_{j-\frac{1}{2}}}{h}$$

$$(\hat{a}_x)_{j+\frac{1}{2}} = \xi_{j+\frac{1}{2}} \left( \frac{\alpha u^+_{j+\frac{1}{2}} - (\alpha - 1) u^-_{j+\frac{1}{2}} - u^+_{j-\frac{1}{2}}}{h} \right), \quad \xi_{j+\frac{1}{2}} = \frac{[a]_{j+\frac{1}{2}}}{[u]_{j+\frac{1}{2}}}.$$

By the mean value theorem, there exists $\xi_0$ such that $\xi_{j+\frac{1}{2}} = \frac{[a]_{j+\frac{1}{2}}}{[u]_{j+\frac{1}{2}}} = a'(\xi_0) \geq 0$. Rewriting (2.13), omitting the superscript denoting time level $n$, we have

$$D_j = \frac{1}{4} u^-_{j+\frac{1}{2}} + \frac{1}{4} u^+_{j-\frac{1}{2}} + \mu \left( \xi_{j+\frac{1}{2}} \alpha u^+_{j+\frac{1}{2}} - \xi_{j+\frac{1}{2}} (\alpha - 1) u^-_{j+\frac{1}{2}} - (\xi_{j+\frac{1}{2}} + \alpha \xi_{j-\frac{1}{2}}) u^+_{j-\frac{1}{2}} \right)$$

$$+ \xi_{j-\frac{1}{2}} (\alpha - 1) u^-_{j-\frac{1}{2}} + \xi_{j-\frac{1}{2}} u^+_{j-\frac{1}{2}} \right)$$

$$= \mu \xi_{j+\frac{1}{2}} \alpha u^+_{j+\frac{1}{2}} + \left( \frac{1}{4} - \mu \xi_{j+\frac{1}{2}} (\alpha - 1) \right) u^-_{j+\frac{1}{2}} + \left( \frac{1}{4} - \mu (\xi_{j+\frac{1}{2}} + \alpha \xi_{j-\frac{1}{2}}) \right) u^+_{j-\frac{1}{2}}$$

$$+ \mu (\alpha - 1) \xi_{j-\frac{1}{2}} u^-_{j-\frac{1}{2}} + \mu \xi_{j-\frac{1}{2}} u^+_{j-\frac{1}{2}}.$$

The convection term:

$$C_j = \frac{1}{2} \bar{u}_j^n - \lambda \left( \hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} \right).$$  \hfill (2.14)
According to the discussion in the preliminaries, in the $P^1$-DG case, we can choose the 2-point Gauss-Lobatto quadrature with the weights $\hat{\omega}_1 = \hat{\omega}_2 = \frac{1}{2}$. Then we have

$$C_j = \frac{1}{4} H_1 + \frac{1}{4} H_2$$

(2.15)

where

$$H_1 = u^-_{j + \frac{1}{2}} - 4\lambda \left( \hat{f}(u^-_{j + \frac{1}{2}}, u^+_{j + \frac{1}{2}}) - \hat{f}(u^+_{j - \frac{1}{2}}, u^-_{j + \frac{1}{2}}) \right) = H_{4\lambda} \left( u^+_{j - \frac{1}{2}}, u^-_{j + \frac{1}{2}}, u^+_{j + \frac{1}{2}} \right),$$

$$H_2 = u^+_{j - \frac{1}{2}} - 4\lambda \left( \hat{f}(u^+_{j - \frac{1}{2}}, u^-_{j + \frac{1}{2}}) - \hat{f}(u^-_{j - \frac{1}{2}}, u^+_{j - \frac{1}{2}}) \right) = H_{4\lambda} \left( u^-_{j - \frac{1}{2}}, u^+_{j - \frac{1}{2}}, u^-_{j + \frac{1}{2}} \right).$$

Combining the results for the convection and diffusion terms together, under the CFL conditions (2.10), $\bar{u}^{n+1}_j$ becomes a monotonically increasing function of point values ranging in $[m, M]$, thus it also belongs to $[m, M]$. $\square$

**Remark 2.2.** We have only proved the result for the first order forward Euler time stepping. The result also holds for high order SSP Runge-Kutta or multistep methods [10] as they are convex combinations of forward Euler steps. For instance, the second order SSP Runge-Kutta method (with the CFL coefficient $c = 1$) is

$$u^{(1)} = u^n + \Delta t F(u^n)$$

$$u^{n+1} = \frac{1}{2} u^n + \frac{1}{2} (u^{(1)} + \Delta t F(u^{(1)}))$$

(2.16)

where $F(u)$ is the spatial operator, the CFL coefficient $c$ for a SSP time discretization refers to the fact that, if we assume the forward Euler time discretization for solving the equation $u_t = F(u)$ is stable in a norm or a semi-norm under a time step restriction $\Delta t \leq \Delta t_0$, then the high order SSP time discretization is also stable in the same norm or semi-norm under the time step restriction $\Delta t \leq c\Delta t_0$.

**Remark 2.3.** This $P^1$-DG result can also be extended to other penalty type DG formulations for second order derivatives, for example the interior penalty Galerkin methods [21, 1, 16], with slightly different CFL conditions and stability coefficient restrictions. If we consider the
interior penalty DG formulation for (1.2): for any test function \( v \in V^k_h \), the IPDG scheme is to find \( u \in V^k_h \), such that

\[
\int_{\Omega} u_t v dx = \sum_{j=1}^{N} \left( \int_{I_j} f(u)v_x dx - \hat{f}_{j+\frac{1}{2}} v^-_{j+\frac{1}{2}} + \hat{f}_{j-\frac{1}{2}} v^+_{j-\frac{1}{2}} \right) - \sum_{j=1}^{N} \left( \int_{I_j} b(u)u_x v_x + \{b v_x\}_{j+\frac{1}{2}} [v]_{j+\frac{1}{2}} \pm \{b v_x\}_{j+\frac{1}{2}} [u]_{j+\frac{1}{2}} + \frac{\alpha}{h} [u]_{j+\frac{1}{2}} [v]_{j+\frac{1}{2}} \right)
\]

where \( \{b v_x\} = \frac{1}{2}((b v_x)^+ + (b v_x)^-) \), and \( \alpha \) is the penalty coefficient. The plus and minus signs above in front of the \( \{b v_x\}_{j+\frac{1}{2}} [u]_{j+\frac{1}{2}} \) term refer to symmetric [21, 1] and nonsymmetric penalty methods [16] respectively. The diffusion term (2.13) then becomes:

\[
D_j = \frac{1}{2} \pi_j + \lambda \left( \{b u_x\}_{j+\frac{1}{2}} - \{b u_x\}_{j-\frac{1}{2}} + \frac{\alpha}{h} [u]_{j+\frac{1}{2}} - \frac{\alpha}{h} [u]_{j-\frac{1}{2}} \right) \quad (2.17)
\]

which can be written out as (omitting the superscript denoting the time level \( n \)):

\[
D_j = \frac{1}{4} u^-_{j+\frac{1}{2}} + \frac{1}{4} u^+_{j-\frac{1}{2}} + \mu \left( \frac{1}{2} b^+_{j+\frac{1}{2}} u^-_{j+\frac{1}{2}} + (\alpha - \frac{1}{2} b^+_{j+\frac{1}{2}}) u^+_{j+\frac{1}{2}} \right)
\]

\[
- (\alpha + \frac{1}{2} b^-_{j+\frac{1}{2}} - \frac{1}{2} b^+_{j+\frac{1}{2}}) u^-_{j+\frac{1}{2}} - (\alpha + \frac{1}{2} b^-_{j+\frac{1}{2}} - \frac{1}{2} b^+_{j+\frac{1}{2}}) u^+_{j+\frac{1}{2}}
\]

\[
+ (\alpha - \frac{1}{2} b^-_{j+\frac{1}{2}}) u^-_{j-\frac{1}{2}} + \frac{1}{2} b^-_{j-\frac{1}{2}} u^+_{j-\frac{1}{2}}
\]

here \( b^+_{j+\frac{1}{2}} = b(u^+_{j+\frac{1}{2}}) \geq 0 \). Collecting terms and using simple algebra, we can easily show that, if \( u^\pm_{j+\frac{1}{2}} \in [m, M] \) for all \( j \), then under the CFL condition specified in Table 2.1, the conclusion of Theorem (2.1) holds.

**Remark 2.4.** This \( P^1 \)-DG result can also be extended to the LDG method for second order derivatives. To obtain an LDG formulation for (1.2), we first rewrite it as:

\[
u_t + f(u)_x = (b^*(u)q)_x, \quad q - B(u)_x = 0 \quad (2.18)\]

where \( b^*(u) = \sqrt{b(u)} \), \( B(u) = \int^u b^*(s)ds \). We then seek \( u, q \in V^k_h \), such that for test functions \( v, p \in V^k_h \):

\[
\int_{I_j} u_t v dx - \int_{I_j} (f(u) - b^*(u)q)v_x dx + (\hat{f} - \hat{b}^*\hat{q}) j+\frac{1}{2} v^-_{j+\frac{1}{2}} - (\hat{f} - \hat{b}^*\hat{q}) j-\frac{1}{2} v^+_{j-\frac{1}{2}} = 0 \quad (2.19)
\]
\[ \int_{I_1} q\, dx + \int_{I_1} B(u)p_x\, dx - \hat{B}_{j+\frac{1}{2}}p^{-}_{j+\frac{1}{2}} + \hat{B}_{j-\frac{1}{2}}p^{+}_{j-\frac{1}{2}} = 0 \]  \quad (2.20)

where the diffusion fluxes are:

\[ \hat{b}^* = \frac{B(u^*) - B(u^-)}{u^* - u^-}, \quad \hat{q} = q^-, \quad \hat{B} = B(u^*). \]

Similarly as before, the diffusion term for (2.19) is:

\[ D_j = \frac{1}{2} \pi_j + \lambda (\hat{b}^*_j \hat{q}_{j+\frac{1}{2}} - \hat{b}^*_j \hat{q}_{j-\frac{1}{2}}), \quad (2.21) \]

where the \( \hat{q} \) is based on \( q \) which is solved locally from (2.20). If we replace the volume integral by a quadrature rule with sufficient accuracy (see [4]), we have

\[ q_j(x) = \frac{(B^+_{j+\frac{1}{2}} - B^+_{j-\frac{1}{2}})}{h} + 6\frac{(B^+_{j+\frac{1}{2}} - B^+_{j+\frac{3}{2}})}{h}(x - x_j) \]

Plugging it into (2.21), omitting the superscript \( n \), we obtain:

\[ D_j = \frac{1}{4} \, u^{-}_{j+\frac{1}{2}} + \frac{1}{4} \, u^{+}_{j+\frac{1}{2}} + \mu \left[ \xi_{j+\frac{1}{2}} (4B^+_{j+\frac{1}{2}} - 3B^-_{j+\frac{1}{2}} - B^+_{j+\frac{3}{2}}) 
- \xi_{j-\frac{1}{2}} (4B^+_{j-\frac{1}{2}} - 3B^-_{j-\frac{1}{2}} - B^+_{j-\frac{3}{2}}) \right]. \]

Here \( \xi = \frac{|b^*|}{|u^*|} \geq 0 \). Collecting terms, we have the diffusion term:

\[ D_j = 4\mu \xi_{j+\frac{1}{2}} B^+_{j+\frac{1}{2}} + \left( \frac{1}{4} \, u^{-}_{j+\frac{1}{2}} - 3\mu \xi_{j+\frac{1}{2}} B^-_{j+\frac{1}{2}} \right) + \left( \frac{1}{4} \, u^{+}_{j+\frac{1}{2}} - \mu (\xi_{j+\frac{1}{2}} + 4\xi_{j-\frac{1}{2}}) B^+_{j-\frac{1}{2}} \right) 
+ 3\mu \xi_{j-\frac{1}{2}} B^-_{j-\frac{1}{2}} + \mu \xi_{j-\frac{1}{2}} B^+_{j-\frac{3}{2}}. \]

Without loss of generality, assume \( m_0 \) is a number smaller than all \( u \) and \( B(u) = \int_{m_0}^u b^*(s)\, ds \), \( b^* \geq 0 \). By the mean value theorem, we have \( B(u) = b^*(\bar{u})(u - m_0) \), with \( \bar{u} \in [m_0, u] \). Thus, by a simple algebra, we have

\[ D_j = 4\mu \xi_{j+\frac{1}{2}} b^*(\bar{u}_{1}) u^{+}_{j+\frac{1}{2}} + \left( \frac{1}{4} - 3\mu \xi_{j+\frac{1}{2}} b^*(\bar{u}_{2}) \right) u^{-}_{j+\frac{1}{2}} + \left( \frac{1}{4} - \mu (\xi_{j+\frac{1}{2}} + 4\xi_{j-\frac{1}{2}}) b^*(\bar{u}_{3}) \right) u^{+}_{j-\frac{1}{2}} 
+ 3\mu \xi_{j-\frac{1}{2}} b^*(\bar{u}_{4}) u^{-}_{j-\frac{1}{2}} + \mu \xi_{j-\frac{1}{2}} b^*(\bar{u}_{5}) u^{+}_{j-\frac{3}{2}}. \]

Now, if \( u^{\pm}_{j+\frac{1}{2}} \in [m, M] \) for all \( j \), then under the CFL condition shown specifically in Table 2.1, the conclusion for Theorem (2.1) still holds.
Table 2.1: The CFL conditions for (1.1) or (1.2) and restrictions for penalty coefficients in the one dimensional case.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>convection CFL</th>
<th>diffusion CFL</th>
<th>penalty coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheng-Shu DG</td>
<td>( \lambda \max</td>
<td>f'(u)</td>
<td>\leq \frac{1}{4} )</td>
</tr>
<tr>
<td>IPDG</td>
<td>( \lambda \max</td>
<td>f'(u)</td>
<td>\leq \frac{1}{4} )</td>
</tr>
<tr>
<td>LDG</td>
<td>( \lambda \max</td>
<td>f'(u)</td>
<td>\leq \frac{1}{4} )</td>
</tr>
</tbody>
</table>

Remark 2.5. From Table 2.1, we note that the CFL conditions for maximum-principle-preserving DG schemes are comparable with or just slightly more restrictive than the standard time step restrictions for linear stability of DG methods for diffusion terms.

2.3 Implementation

To enforce the sufficient conditions in Theorem 2.1, i.e., \( u_{j+\frac{1}{2}} \) at time level \( n \), we can use the same limiter as in [26]. Given approximation polynomials \( p_j(x) \) evolved from a P1-DG scheme, with its cell averages \( \overline{u}_j^n \in [m, M] \), we would like to modify \( p_j(x) \) into \( \tilde{p}_j(x) \) by

\[
\tilde{p}_j(x) = \theta(p_j(x) - \overline{u}_j^n) + \overline{u}_j^n, \quad \theta = \min \left\{ 1, \frac{M - \overline{u}_j^n}{M_j - \overline{u}_j^n}, \frac{m - \overline{u}_j^n}{m_j - \overline{u}_j^n} \right\},
\]

(2.22)

where

\[
M_j = \max\{p_j(x_{j+\frac{1}{2}}), p_j(x_{j-\frac{1}{2}})\}, \quad m_j = \min\{p_j(x_{j+\frac{1}{2}}), p_j(x_{j-\frac{1}{2}})\}.
\]

It is clear that \( \tilde{p}_j(x_{j+\frac{1}{2}}) \in [m, M] \) and \( \tilde{p}_j(x) \) has the same cell average \( \overline{u}_j^n \). Moreover, we have \( ||\tilde{p}_j - p_j||_\infty = O(h^2) \) if the exact solution is smooth, see [26] for the proof.

For forward Euler time discretization, at time level \( n \), assume the P1-DG polynomial in cell \( I_j \) is \( p_j(x) \), and the cell average of \( p_j(x) \) is \( \overline{u}_j^n \in [m, M] \), then

- In each cell, monitor the point values of \( p_j(x) \) at the two end points \( x_{j+\frac{1}{2}} \) by computing \( \theta \) in (2.22).
- If \( \theta = 1 \), leave the DG solution \( p_j(x) \) unchanged.
- If \( \theta < 1 \), use \( \tilde{p}_j(x) \) instead of \( p_j(x) \) in the DG scheme with forward Euler in time under the CFL condition in Table 2.1.
For SSP high order time discretizations, we need to use the limiter in each stage for a Runge-Kutta method or in each step for a multistep method. Notice that the CFL conditions in Table 2.1 are sufficient but not necessary to achieve maximum principle. A more efficient implementation would be enforcing the stringent CFL conditions in Table 2.1 only when a preliminary calculation to next time step with a normal time step violates the maximum principle. See Section 3.3 for more details regarding this implementation.

3 Maximum-principle-satisfying DG schemes in two dimensions

In this section, we extend our results to P1-DG schemes on triangular meshes solving the initial value problem for two-dimensional nonlinear convection-diffusion equations in a general form:

$$\frac{\partial u}{\partial t} + \nabla \cdot F(u) = \nabla \cdot (A \nabla u), \quad F(u) = \langle f(u), g(u) \rangle,$$

where $A = A(u, x)$ is a $2 \times 2$ symmetric semi-positive-definite matrix. Let $[m, M]$ denote the desired range indicated by the initial condition. We will assume periodic boundary condition again for simplicity. All the results regarding maximum principle can be easily incorporated into any initial-boundary value problem.

3.1 Preliminaries

In [30], Zhang, Xia and Shu explained in details how to construct maximum-principle-satisfying high order DG schemes for hyperbolic conservation laws on triangular meshes

$$u_t + \nabla \cdot F(u) = 0, \quad F(u) = \langle f(u), g(u) \rangle.$$  

We review the results briefly. On a triangulation $T_h$, for each triangle $K$, we denote by $|K|$ the area of the triangle and $l^i_K$ ($i = 1, 2, 3$) the length of the corresponding edges $e^i_K$ ($i = 1, 2, 3$), with outward unit normal vectors being $\vec{\nu}_i$ ($i = 1, 2, 3$). We also denote the neighboring triangle along $e^i_K$ as $K_i$. Any one-dimensional monotone flux can be used, for
example the global Lax-Friedrichs flux

\[ \hat{f}(u, v, \mathbf{\tilde{v}}) = \frac{1}{2} (F(u) \cdot \mathbf{\tilde{v}} + F(v) \cdot \mathbf{\tilde{v}} - a(v - u)), \quad a = \max_u |F'(u) \cdot \mathbf{\tilde{v}}|. \] (3.3)

The first order Lax-Friedrichs scheme for (3.2) on triangular meshes is defined as:

\[ u_{n+1}^K = u_n^K - \frac{\Delta t}{|K|} \sum_{i=1}^{3} \hat{f}(u_n^K, u_n^{i,K}, \mathbf{\tilde{v}}_i)^l_i \]

By the same token, the scheme satisfied by the cell averages of DG schemes for (3.2) can be written as:

\[ \bar{u}_{n+1}^K = \bar{u}_n^K - \frac{\Delta t}{|K|} \sum_{i=1}^{3} \int_{e_i^K} \hat{f}(u_{n^i}^{in}, u_{n^i}^{out}, \mathbf{\tilde{v}}_i) ds \]

where \( \bar{u}_K \) is the cell average of the numerical solution over the triangle \( K \), \( u_{n^i}^{in} \) and \( u_{n^i}^{out} \) are the values of the numerical solution on the edge \( e_i^K \) evaluated from the DG polynomials defined in \( K \) and in \( K_i \) respectively. Given the approximation space \( P^k(K) \), with sufficient accuracy, the line integrals along the edges can be approximated by the \( (k+1) \)-point Gauss quadrature rules:

\[ \bar{u}_{n+1}^K = \bar{u}_n^K - \frac{\Delta t}{|K|} \sum_{i=1}^{3} \sum_{\beta=1}^{k+1} \hat{f}(u_{n^i}^{i,\beta}, u_{n^i}^{i,\beta}, \mathbf{\tilde{v}}_i) \omega_{\beta} l_i^K \] (3.4)

where \( \omega_{\beta} \) denotes the quadrature weights for the \( (k+1) \)-point Gauss quadrature rule on \([-\frac{1}{2}, \frac{1}{2}]\), and \( u_{n^i}^{i,\beta} \) denotes the value of the DG polynomial defined on \( K \) evaluated at the \( \beta \)-th Gauss quadrature point on the edge \( e_i^i \).

In [30], the authors proved that the right hand side of (3.4) can be rewritten as a monotonically increasing function with respect to certain point values of the DG polynomials, thus the maximum principle can be easily achieved by the methodology for the one-dimensional case. These points, denoted by \( S_k^K \) on each triangle \( K \), contain the Gauss quadrature points on the three edges. For DG method with polynomials of arbitrary degree \( k \), \( S_k^K \) can be constructed by taking the Dubiner transformation of the tensor product of Gauss-Lobatto and Gauss quadratures, which is then summarized (see [30] for details). In particular, for the \( P^1 \)-DG method under consideration in this paper, the set \( S_k^K \) is defined as the six points
shown in Figure 3.1 where the symbols denote the 2-point Gaussian quadrature points for each edge.

Let \( u_K(x) \) denote the DG polynomial on the cell \( K \). For the \( P^1 \)-DG method, if the point values of \( u_K(x) \) on the three vertices are in the range \([m, M]\) for each \( K \) at time level \( n \), then the point values of \( u_K(x) \) on any points inside \( K \) including on the stencil \( S^1_K \) are in the range \([m, M]\) due to the linearity, thus following the result in [30] the solutions of (3.4) satisfy \( \mathbf{u}_K^{n+1} \in [m, M] \) under the following CFL condition:

\[
a \frac{\Delta t}{|K|} \sum_{i=1}^{3} l^i_K \leq \frac{1}{3}.
\] (3.5)

Therefore, to seek a sufficient condition to ensure the maximum principle for the convection-diffusion equations, it suffices to focus only on the diffusion part. Straightforward extension to convection-diffusion equations can then be obtained using similar cell average splitting approach as in the one-dimensional case. Thus we will first consider the two-dimensional diffusion equation, i.e., (3.1) with \( \mathbf{F} \equiv 0 \).

### 3.2 Monotonicity of the scheme for the diffusion part

We consider in this section only the Cheng-Shu DG scheme in [2] for (3.1) with \( \mathbf{F} \equiv 0 \). For a given triangulation \( T_h \), define the approximation space as: \( V_h^k = \{ v : v|_K \in P^k(K) \} \). The DG scheme is to seek \( u \in V_h^k \), such that for any test function \( \varphi \in V_h^k \), the following equality
holds:

\[
\frac{\partial}{\partial t} \int\int_K u \varphi dx = \int\int_{\partial K} (A \nabla u \cdot \vec{v}) \varphi ds - \int\int_{\partial K} \left( \tilde{u} A \nabla \varphi \right) \cdot \vec{v} ds + \int\int_K u \nabla \cdot (A \nabla \varphi) dx. \tag{3.6}
\]

The numerical fluxes on each edge \( e_i^K \) are:

\[
(A \nabla u \cdot \vec{v}) = A(u^-) \nabla u^- \cdot \vec{v} + \frac{\alpha_A}{\lambda_A} (u^\text{out} - u^\text{in}), \quad \tilde{u} A = u^+ A(u^+), \tag{3.7}
\]

where the maximum can be taken either locally on each edge or globally over the whole domain, \( \vec{v} \) is the outward unit normal vector on the edges, \( \|A\| \) is the matrix 2-norm or the spectral radius (maximum eigenvalue in absolute value) for the symmetric matrix \( A \), \( \alpha \) is again a sufficiently large positive constant to ensure stability. \( u^\pm \) denotes the numerical solution on the edges, evaluated from \( K \) or \( K_i \). The '±' for each edge \( e_i^K \) is determined by the inner product of \( \vec{v}_i \) and a predetermined constant vector \( \vec{v}_0 \) which is not parallel to any edge in the mesh: for each edge \( e_i^K \) in the cell \( K \),

\[
\begin{align*}
    u^- &= u_K, \quad u^+ = u_{K_i} \quad \text{if} \quad \vec{v}_0 \cdot \vec{v}_i < 0; \\
    u^+ &= u_K, \quad u^- = u_{K_i} \quad \text{if} \quad \vec{v}_0 \cdot \vec{v}_i > 0.
\end{align*}
\]

Taking \( \varphi = 1 \) on \( K \) and \( \varphi = 0 \) everywhere else,

\[
\frac{\partial}{\partial t} \tilde{u}_K = \frac{1}{|K|} \int_{\partial K} (A \nabla u \cdot \vec{v}) ds.
\]

Therefore, the scheme satisfied by the cell averages of the DG method with forward Euler on a triangular mesh is

\[
\tilde{u}_K^{n+1} = \tilde{u}_K^n + \frac{\Delta t}{|K|} \int_{\partial K} (A \nabla u^n \cdot \vec{v}) ds.
\]

The integral can be approximated by a 2-point Gauss quadrature with sufficient accuracy for P1-DG. We use \( x_{i,\beta} (\beta = 1, 2) \) to denote the local Gauss quadrature points for each edge \( e_i^K \). Let \( \omega_{\beta} (\beta = 1, 2) \) denote the corresponding quadrature weights for the 2-point Gauss
quadrature on the interval \([-\frac{1}{2}, \frac{1}{2}]\). The scheme then becomes (omitting the superscripts for time level \(n\) on the right hand side):

\[
\bar{u}_K^{n+1} = \frac{\Delta t}{|K|} \sum_{i=1}^{\ell_i} \sum_{\beta=1}^{2} \omega_{\beta} (\mathbf{A} \nabla u \cdot \mathbf{v}_i)_{\mathbf{x}=\mathbf{x}_{i,\beta}} + \frac{\Delta t}{|K|} \sum_{i=1}^{\ell_i} \sum_{\beta=1}^{2} \omega_{\beta} (\mathbf{A} \nabla u \cdot \mathbf{\bar{v}}_i)_{\mathbf{x}=\mathbf{x}_{i,\beta}}.
\]

(3.8)

Now the idea is to rewrite the right hand side above as a convex combinations of certain point values. To find an explicit expression for the flux \((\mathbf{A} \nabla u \cdot \mathbf{v}_i)\), we consider two different cases.

- **Case A:** \(u^- = u^{out}, u^+ = u^{in}\). For this case, we have \(u_{i,\beta} = u_{K_i}(\mathbf{x}_{i,\beta})\). The subscript or superscript \((i, \beta, K_i)\) for \(u\) and \(\nabla u\) will denote the value evaluated at \(u_{K_i}(\mathbf{x}_{i,\beta})\). Then we have

\[
(A \nabla u \cdot \mathbf{v}_i)_{\mathbf{x}=\mathbf{x}_{i,\beta}} = A_{i,\beta}^K \nabla u_{K_i}^{i,\beta} \cdot \mathbf{v}_i + \alpha A_{i,\beta}^K (u_{K_i}^{i,\beta} - u_{K_i}^i).
\]

(3.9)

Let \(\mathbf{v}_{i,\beta} = A_{i,\beta}^K \mathbf{v}_i\), then \(\mathbf{v}_{i,\beta} \cdot \mathbf{v}_i = A_{i,\beta}^K \mathbf{v}_i \cdot \mathbf{v}_i \geq 0\) because \(A\) is semi-positive definite. Namely, \(\mathbf{v}_{i,\beta}\) is a direction pointing into \(K_i\) or along the edge \(e_{i,K_i}\). Since \(A\) is symmetric, we have a directional derivative

\[
A_{i,\beta}^K \nabla u_{K_i}^{i,\beta} \cdot \mathbf{v}_i = \nabla u_{K_i}^{i,\beta} \cdot (A_{i,\beta}^K \mathbf{v}_i) = \frac{\partial u_{K_i}(\mathbf{x})}{\partial \mathbf{v}_{i,\beta}}_{\mathbf{x}=\mathbf{x}_{i,\beta}}.
\]

For each point \(\mathbf{x}_{i,\beta}\) on the edge \(e_{i,K_i}\), consider the straight line described by the parametric equation \(\mathbf{r}(t) = t\mathbf{v}_{i,\beta} + \mathbf{x}_{i,\beta}, t \in \mathbb{R}\), which intersects with the other two edges of \(K_i\) at one point denoted by \(\mathbf{x}_{i,\beta}^*\). See Figure 3.2(a) for an illustration.

By the linearity of \(u_{K_i}(\mathbf{x})\), the directional derivative can be written as:

\[
\frac{\partial u_{K_i}(\mathbf{x})}{\partial \mathbf{v}_{i,\beta}}_{\mathbf{x}=\mathbf{x}_{i,\beta}} = \frac{u_{K_i}(\mathbf{x}_{i,\beta}^*) - u_{K_i}^{i,\beta}}{||\mathbf{x}_{i,\beta} - \mathbf{x}_{i,\beta}^*||}.
\]

(3.10)

Plugging (3.10) into (3.9), we obtain:

\[
(A \nabla u \cdot \mathbf{v}_i)_{\mathbf{x}=\mathbf{x}_{i,\beta}} = \frac{1}{||\mathbf{x}_{i,\beta} - \mathbf{x}_{i,\beta}^*||} u_{K_i}(\mathbf{x}_{i,\beta}^*) + \left(\frac{\alpha A_{i,\beta}^K}{l_i^K} - \frac{1}{||\mathbf{x}_{i,\beta} - \mathbf{x}_{i,\beta}^*||}\right) u_{K_i}^{i,\beta} - \frac{\alpha A_{i,\beta}^K}{l_i^K} u_{K_i}^i.
\]

(3.11)
• **Case B:** \( u^+ = u^{\text{out}}, \ u^- = u^{\text{in}} \). For this case, we have \( u_{i,\beta}^- = u_K(x_{i,\beta}) \). Then

\[
(A \nabla u \cdot \vec{v}_i) \bigg|_{x = x_{i,\beta}} = A_{i,\beta}^{i,\beta} \nabla u_K^{i,\beta} \cdot \vec{v}_i + \frac{\alpha A}{l_K} (u_{i,\beta}^K - u_{i,\beta}^{i,\beta})
\]

Let \( \vec{v}_{i,\beta} = -A_{i,\beta}^{i,\beta} \vec{v}_i \), then \( \vec{v}_{i,\beta} \cdot \vec{v}_i = -A_{i,\beta}^{i,\beta} \vec{v}_i \cdot \vec{v}_i \leq 0 \). Namely, \( \vec{v}_{i,\beta} \) is a direction pointing into \( K \) or along the edge \( e_i^K \). The directional derivative is

\[
A_{i,\beta}^{i,\beta} \nabla u_K^{i,\beta} \cdot \vec{v}_i = \nabla u_K^{i,\beta} \cdot (-\vec{v}_{i,\beta}) = -\frac{\partial u_K(x)}{\partial \vec{v}_{i,\beta}} \bigg|_{x = x_{i,\beta}}.
\]

For each point \( x_{i,\beta} \) on the edge \( e_i^K \), consider the straight line described by the parametric equation

\[
r(t) = t \vec{v}_{i,\beta} + x_{i,\beta}, \quad t \in \mathbb{R},
\]

which intersects with the other two edges of \( K \) at one point denoted by \( x_{i,\beta}^* \). See Figure 3.2 (b) for an illustration. Assume \( x_{i,\beta}^* \) lies on the edge \( e_{j}^K \) (\( j \neq i \)). Since \( u_K(x) \) is a linear function, \( u_K(x_{i,\beta}^*) \) can be represented as a linear combination of point values at the two Gauss quadrature points \( u_{j,\beta}^{(i,\beta),1}, u_{j,\beta}^{(i,\beta),2} \) on the edge \( e_{j}^K \) where \( j \) depends on \( i \) and \( \beta \):

\[
u_K(x_{i,\beta}^*) = c_{i,\beta}^1 u_{j,\beta}^{(i,\beta),1} + c_{i,\beta}^2 u_{j,\beta}^{(i,\beta),2}
\]

(3.12)

with \( c_{i,\beta}^1, c_{i,\beta}^2 \in [-\sqrt{3}/2, \sqrt{3}/2] \) and \( c_{i,\beta}^1 + c_{i,\beta}^2 = 1 \).

![Figure 3.2: Two cases for the numerical flux on triangular meshes](image)

By the linearity of \( u_K(x) \), the directional derivative can be written as:

\[
\frac{\partial u_K(x)}{\partial \vec{v}_{i,\beta}} \bigg|_{x = x_{i,\beta}} = \frac{u_K(x_{i,\beta}^*) - u_{i,\beta}^{i,\beta}}{||x_{i,\beta} - x_{i,\beta}^*||}.
\]

20
Thus we obtain:

\[
(A \nabla u \cdot \mathbf{v})_{x=x_i,\beta} = -\frac{1}{||x_{i,\beta} - x^*_{i,\beta}||} u_K(x^*_{i,\beta}) - \left( \frac{\alpha_A}{l_i^K} - \frac{1}{||x_{i,\beta} - x^*_{i,\beta}||} \right) u^i_{i,\beta} + \frac{\alpha_A}{l_i^K} u_{i,\beta}. \\
(3.13)
\]

By the exactness of the quadrature in Figure 3.1 for the linear polynomial \(u_K(x)\) on the triangle \(K\) (see [30] for how to construct this quadrature), we can decompose the cell average \(\bar{u}_K^n\) as a convex combination of the point values on the stencil \(S^1_K\),

\[
\bar{u}_K^n = \frac{1}{|K|} \int_K u_K(x) \, dx = \sum_{i=1}^3 \sum_{\beta=1}^2 \frac{1}{3} \omega_{\beta} u^i_{i,\beta}. \\
(3.14)
\]

Let \(h_{i,\beta}\) denote \(||x_{i,\beta} - x^*_{i,\beta}||\). Combining (3.11), (3.12), (3.13) and (3.14), we can rewrite (3.8) as:

\[
\bar{u}_K^{n+1} = \bar{u}_K^n + \frac{\Delta t}{|K|} \sum_{i=1}^3 l_i^K \sum_{\beta=1}^2 \omega_{\beta} (A \nabla u \cdot \mathbf{v})_{x=x_i,\beta} \\
= \frac{1}{2} \bar{u}_K^n + \frac{1}{2} \bar{u}_K^n + \frac{\Delta t}{|K|} \sum_{i=1}^3 l_i^K \sum_{\beta=1}^2 \omega_{\beta} (A \nabla u \cdot \mathbf{v})_{x=x_i,\beta} \\
= \sum_{i \in \text{Case A}} \sum_{\beta=1}^2 \omega_{\beta} \left[ \left( \frac{1}{6} - \Delta t \frac{\alpha_A}{|K|} \right) u^i_{i,\beta} + \frac{\Delta t}{|K|} \frac{l_i^K}{h_{i,\beta}} u_{K_i}(x^*_{i,\beta}) + \frac{\Delta t}{|K|} \left( \frac{\alpha_A}{h_{i,\beta}} - \frac{l_i^K}{h_{i,\beta}} \right) u_{i,\beta} \right] \\
+ \sum_{i \in \text{Case B}} \sum_{\beta=1}^2 \omega_{\beta} \left[ \left( \frac{1}{6} - \Delta t \frac{\alpha_A}{|K|} + \frac{\Delta t}{|K|} \frac{l_i^K}{h_{i,\beta}} \right) u^i_{i,\beta} + \Delta t \frac{\alpha_A}{|K|} u_{i,\beta} \right] \\
+ \sum_{i \in \text{Case B}} \sum_{\beta=1}^2 \omega_{\beta} \left( -\Delta t \frac{l_i^K}{|K|} u_K(x^*_{i,\beta}) \right) + \frac{1}{2} \bar{u}_K^n. \\
(3.15)
\]

Notice that we can rewrite the cell average as

\[
\bar{u}_K^n = \frac{1}{3} \sum_{i=1}^3 \bar{u}_K^n = \frac{1}{3} \left( \sum_{i \in \text{Case A}} \bar{u}_K^n + \sum_{i \in \text{Case B}} \bar{u}_K^n \right) = \frac{1}{3} \left( \sum_{i \in \text{Case A}} \bar{u}_K^n + \sum_{i \in \text{Case B}} \sum_{\beta=1}^2 \omega_{\beta} u^i_{i,\beta} \right),
\]

so the third term in (3.15) can be further expanded:

\[
\sum_{i \in \text{Case B}} \sum_{\beta=1}^2 \omega_{\beta} \left( -\Delta t \frac{l_i^K}{|K|} u_K(x^*_{i,\beta}) \right) + \frac{1}{2} \bar{u}_K^n.
\]

21
we have,}

\[
\tilde{u}_K^{n+1} = \sum_{i \in \text{Case}A} \sum_{\beta = 1}^{2} \omega_\beta \left[ \left( \frac{1}{6} - \Delta t \frac{\alpha_A}{|K|} \right) u_{i,\beta} + \Delta t \frac{l_K}{|K|h_{i,\beta}} u_{K_i}(x_{i,\beta}) + \Delta t \left( \alpha_A - \frac{l_K}{h_{i,\beta}} \right) u_{K_i} \right] \\
+ \sum_{i \in \text{Case}B} \sum_{\beta = 1}^{2} \omega_\beta \left[ \left( \frac{1}{6} - \Delta t \frac{\alpha_A}{|K|} + \Delta t \frac{l_K}{|K|h_{i,\beta}} \right) u_{i,\beta} + \Delta t \frac{\alpha_A}{|K|} u_{i,\beta} \right] + \frac{1}{6} \sum_{i \in \text{Case}A} \tilde{u}_K^n \\
+ \sum_{i \in \text{Case}B} \sum_{\beta = 1}^{2} \omega_\beta \left[ \frac{1}{18} \sum_{k \neq j} \sum_{\gamma = 1}^{2} \omega_\gamma u_{k,\gamma}^{i,\beta} + \left( \frac{1}{36} - \Delta t \frac{l_K}{|K|h_{i,\beta}} c_{1,i,\beta} \right) u_{j,1}^{i,\beta} + \left( \frac{1}{36} - \Delta t \frac{l_K}{|K|h_{i,\beta}} c_{2,i,\beta} \right) u_{j,2}^{i,\beta} \right].
\]

(3.16)

To have a convex combination in (3.16), it suffices to require the following terms to be non-negative, which will imply a time step restriction,

\[
\frac{1}{6} - \Delta t \frac{\alpha_A}{|K|}, \quad \alpha_A = \frac{l_K}{h_{i,\beta}}, \quad \frac{1}{36} - \Delta t \frac{l_K}{|K|h_{i,\beta}} c_{1,i,\beta}.
\]

(3.17)

Next, we derive an explicit CFL condition for a regular triangulation, namely a triangulation with the minimum angle bounded uniformly away from zero. Recall that for each triangle \( K \), \( x_{i,\beta} (\beta = 1, 2) \) are the two Gaussian quadrature points on the edge \( e^i_K \). By simple geometry relationship, we have,

\[
h_{i,\beta} \geq \tilde{C} \sin(\tilde{\theta}) l_K^i, \quad \frac{l_K^i}{h_{i,\beta}} \leq \frac{1}{\tilde{C} \sin(\tilde{\theta})}
\]

(3.18)
Here $\tilde{C} = \frac{D(x_{i,\beta})}{l_K}$, with $D(x_{i,\beta})$ denotes the distance from the point $x_{i,\beta}$ to vertices of the edge $e_K$. Notice that there is a uniform lower bound enforcing $\tilde{C} \geq \frac{3-\sqrt{3}}{6}$, since $x_{i,\beta}$ are 2-point Gauss quadrature points on $e^i_K$. $\tilde{C}$ here denotes the angle between the edge $e^i_K$ and $e^j_K$ or $e^j_{K,i}$ $(j \neq i)$ on which the point $x^*_{i,\beta}$ lies on. Thus, we let $\theta_{K,i,j}$, $i, j \in \{1, 2, 3\}, i \neq j$ denote the angles between edges $e^i_K$ and $e^j_K$. Notice that $c_{i,\beta} \in [-\frac{\sqrt{3}-1}{2}, \frac{\sqrt{3}+1}{2}]$, it suffices to have the following conditions for the terms in (3.17) to be non-negative,

$$\frac{1}{\tilde{C} \min_{K,i,j} (\sin(\theta^i_{K,j}))} \leq \frac{1}{6}, \quad \Delta t \frac{1}{\tilde{C} \min_{K,i,j} (\sin(\theta^i_{K,j}))} \leq \frac{\sqrt{3}-1}{36},$$

which can be simplified as

$$\alpha \geq \frac{(3 + \sqrt{3})}{\min_{K,i,j} (\sin(\theta^i_{K,j}))}, \quad \Delta t \frac{\alpha}{|K|} \leq \frac{\sqrt{3}-1}{36}. \quad (3.19)$$

### 3.3 A second order maximum-principle-satisfying DG scheme

For a given triangulation $T$, define the approximation space as: $V^k_h = \{v : v|_K \in P^k(K)\}$.

The DG scheme for a convection-diffusion problem (3.1) is to seek $u \in V^k_h$, such that for any test function $\varphi \in V^k_h$, the following equality holds:

$$\frac{\partial}{\partial t} \int_K u \varphi dx = \int_K F \cdot \nabla \varphi dx - \int_{\partial K} \hat{F} \cdot \hat{\nu} \varphi ds$$

$$+ \int_{\partial K} (A \nabla u \cdot \nu) ds - \int_{\partial K} \left( u \tilde{A} \nabla \varphi \right) \cdot \tilde{\nu} ds + \int_K u \nabla \cdot (A \nabla \varphi) dx,$$

with the numerical fluxes:

$$F \cdot \nu = \frac{1}{2} \left( F(u^*) \cdot \nu + F(u^{out}) \cdot \nu - a(u^{out} - u^*) \right), \quad a = \max_{u,\tilde{u}} |F(u) \cdot \tilde{\nu}|,$$

$$(A \nabla u \cdot \nu) = A(u^-) \nabla u^- \cdot \nu + \alpha \frac{l_K}{|K|} (u^{out} - u^*), \quad u^- = u + A(u^*).$$

For forward Euler time stepping, take the test function as the indicator function on the cell $K$, and replace the line integral with 2-point Gauss quadrature, we get the cell average scheme for P1-DG method:

$$\bar{u}^{n+1}_K = \bar{u}^n_K - \frac{\Delta t}{|K|} \sum_{i=1}^3 l^i_K \sum_{\beta=1}^2 \omega_{\beta} (F \cdot \nu)_{i,\beta} + \frac{\Delta t}{|K|} \sum_{i=1}^3 l^i_K \sum_{\beta=1}^2 \omega_{\beta} (A \nabla \nu \cdot \tilde{\nu} + \frac{\Delta t}{l_K} (A \nabla u \cdot \nu)) \bigg|_{x=x_{i,\beta}}. \quad (3.21)$$
We can decompose (3.21) as $\tau^{n+1}_K = C_K + D_K$ with

$$C_K = \frac{1}{2} \tau^n_K - \frac{\Delta t}{|K|} \sum_{i=1}^{3} l_i^2 \sum_{\beta=1}^{2} \omega_{\beta} (\hat{F} \cdot \nu)_{i,\beta}$$  \hspace{1cm} (3.22)$$

$$D_K = \frac{1}{2} \tau^n_K + \frac{\Delta t}{|K|} \sum_{i=1}^{3} l_i^2 \sum_{\beta=1}^{2} \omega_{\beta} (A \nabla u \cdot \nu_i)_{|x=x_{i,\beta}}$$  \hspace{1cm} (3.23)$$

By the results in [30], the convection part (3.22) is a monotonically increasing function of point values of DG polynomials under the CFL condition $a \Delta t |K| \leq \frac{1}{6}$. By the discussion in the previous subsection, the diffusion part (3.23) is a convex combination of point values of DG polynomials under the constraints $\alpha_A \geq (3 + \sqrt{3}) \frac{1}{\min_{K,i,j} (\sin(\theta_{ij}^K))}$, $\Delta t \frac{\alpha_A}{|K|} \leq \frac{\sqrt{3} - 1}{72}$, with one half smaller time step than (3.19) since the cell average $\tau^n_K$ is halved in (3.23).

Notice that at time level $n$, for the linear DG polynomials $u_K(x)$, if $u_K(x) \in [m, M]$ at three vertices of the triangle $K$, then $u_K(x) \in [m, M]$ for any $x \in K$. Thus, we have,

**Theorem 3.1.** Consider the $P^1$-DG scheme (3.20), if the DG polynomial at time level $n$ on each triangle $K$ satisfies $u_K(x) \in [m, M]$ on all three vertices of the triangle $K$, then the cell average at next time step satisfies the maximum principle, i.e., $\tau^{n+1}_K \in [m, M]$, under the CFL conditions:

$$a \frac{\Delta t}{|K|} \sum_{i=1}^{3} l_i^2 \leq \frac{1}{6}, \quad \Delta t \frac{\alpha_A}{|K|} \leq \frac{\sqrt{3} - 1}{72}, \quad \alpha_A \geq (3 + \sqrt{3}) \frac{1}{\min_{K,i,j} (\sin(\theta_{ij}^K))},$$ \hspace{1cm} (3.24)$$

where $\theta_{K}^{i,j}$ is defined as in Section 3.2, and $\alpha_A = \lambda_A \alpha$, $\lambda_A = \max_{u \in [m, M], x \in T_n} \| A(u, x) \|$.

**Remark 3.2.** The CFL conditions (3.24) is only a sufficient condition rather than a necessary condition for maximum principle to hold. Larger time steps are possible by trying different expansions in (3.15).

Given linear polynomials $p_K(x)$ as the DG polynomial defined on triangle $K$ with $\tau^n_K$ as its cell average over $K$ at time level $n$, if $\tau^n_K \in [m, M]$, the following limiter ensures $\tilde{p}_K(x) \in [m, M]$ for any $x \in K$:

$$\tilde{p}_K(x) = \theta (p_K(x) - \tau^n_K) + \tau^n_K, \quad \theta = \min \left\{1, \left\{ \frac{M - \tau^n_K}{M_K - \tau^n_K}, \frac{m - \tau^n_K}{m_K - \tau^n_K} \right\} \right\}$$ \hspace{1cm} (3.25)$$
where $M_k$ and $m_k$ are the maximum and minimum of $p_K(x)$ on the triangle $K$. Notice that the limiter (3.25) does not change the cell average. Moreover, we have $||\tilde{p}_K - p_K||_\infty = \mathcal{O}(h^2)$ if the exact solution is smooth, see [26] for the proof.

An efficient implementation of our maximum-principle-satisfying $P^1$-DG method (3.20) is:

1. Find the desired bound $[m, M]$ by evaluating the minimum and maximum of the initial condition. Given the triangular mesh, find the mesh constant $\min_{K,i,j} \left( \sin(\theta_{i,j}^K) \right)$ and a direction $\vec{v}_0$ which is not parallel to any edge in the mesh. Calculate $\lambda_A = \max_{u \in [m,M], x \in T_h} ||A(u, x)||$. Take the penalty constant as $\alpha = (3 + \sqrt{3}) \frac{1}{\lambda_A \min_{K,i,j} \left( \sin(\theta_{i,j}^K) \right)}$ as suggested by (3.24).

2. Choose a discretization of the initial data such that the cell averages of the numerical initial data are in $[m, M]$. For instance, the standard $L^2$-projection.

3. Use the second order SSP time discretizations, for instance, the second order Runge-Kutta method (2.16).

4. For each forward Euler time stepping in (2.16), use the limiter (3.25) to pre-process the DG polynomials.

5. Consider the DG scheme with limiter as an operator. Given the numerical solution $u^n$ at time level $n$ as an input, let $\tilde{u}^{n+1}$ denote the output with a normal time step,

- if the cell averages of $\tilde{u}^{n+1}$ violates the maximum principle, then go back to time level $n$ and evolve $u^n$ with a smaller time step suggested by (3.24).
- otherwise use $\tilde{u}^{n+1}$ as the solution at the next time level, i.e., set $u^{n+1} = \tilde{u}^{n+1}$ and proceed.

This algorithm is guaranteed to produce numerical solutions within the range $[m, M]$ with uniform second order accuracy for smooth exact solutions.
4 One dimensional numerical tests

In this section, we give numerical tests for the Cheng-Shu DG formulation [2]. The same set of numerical experiments have also been carried out using the IPDG and LDG schemes. The results are similar, thus we will not list them here to save space.

4.1 Accuracy test

We first consider the linear equation

\[ u_t + u_x = \varepsilon u_{xx} \]  \hspace{1cm} (4.1)

with initial data

\[ u_0(x) = \sin(x) \]

on \([0, 2\pi]\) and periodic boundary conditions. The exact solution is: \(u(x, t) = \exp(-\varepsilon t) \sin(x - t)\). We take \(\varepsilon = 0.001\) and the final time is \(t = 1\). The time step is taken as suggested in (2.10). From Table 4.1, we observe the expected second order accuracy for the maximum-principle-satisfying DG scheme.

<table>
<thead>
<tr>
<th>No. of Elements</th>
<th>DG without Limiter</th>
<th>DG with Limiter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(L^1) error</td>
<td>order</td>
</tr>
<tr>
<td>20</td>
<td>1.83E-02</td>
<td>–</td>
</tr>
<tr>
<td>40</td>
<td>4.55E-03</td>
<td>2.01</td>
</tr>
<tr>
<td>80</td>
<td>1.13E-03</td>
<td>2.00</td>
</tr>
<tr>
<td>160</td>
<td>2.82E-04</td>
<td>2.01</td>
</tr>
<tr>
<td>320</td>
<td>7.03E-05</td>
<td>2.01</td>
</tr>
<tr>
<td>640</td>
<td>1.75E-05</td>
<td>2.01</td>
</tr>
</tbody>
</table>

4.2 Porous medium equations

A typical example of degenerate parabolic equations is the porous medium equation. In one space dimension, the porous medium equation can be written as follows:

\[ u_t = (u^m)_{xx}, \quad m > 1. \]  \hspace{1cm} (4.2)
We test the proposed scheme using the Barenblatt solution:

\[ B_m(x, t) = t^{-k}[\left(1 - \frac{k(m - 1)|x|^2}{2m t^{2k}}\right)]^{m-1} \]

(4.3)

which is an exact solution to (4.2) with compact support. The initial condition is taken to be \( B_m(x, 1) \), and the numerical solution is computed to \( t = 2 \), with zero boundary conditions.

We compare the numerical solutions obtained from the maximum-principle-satisfying DG scheme to the analytical solutions, see Figure 4.1 (a). The maximum-principle-satisfying DG scheme resolves the discontinuities in the solutions quite well, and keeps the solution strictly within the initial bounds everywhere for all time. If the DG schemes are applied without the limiter, there are significant undershoots near the foot of the numerical solution. See Figure 4.1 (b) and more clearly in the zoomed version Figure 4.1 (c).

### 4.3 The Buckley-Leverett equation

Now we consider the convection-diffusion Buckley-Leverett equation, which is a model often used in reservoir simulations and a standard numerical test, see [11].

\[ u_t + f(u)_x = \varepsilon(\nu(u)u_x)_x, \]

(4.4)

The numerical solution is computed at \( t = 0.2 \) with \( \varepsilon = 0.01 \) and boundary condition is \( u(0, t) = 1 \). The \( \nu(u) \) and the initial condition are given as:

\[ \nu(u) = \begin{cases} 4u(1-u) & 0 \leq u \leq 1 \\ 0 & \text{otherwise} \end{cases}, \quad u(x, 0) = \begin{cases} 1 - 3x & 0 \leq x \leq \frac{1}{3} \\ 0 & \frac{1}{3} \leq x \leq 1 \end{cases} \]

(4.5)

\( f(u) \) has an s-shape:

\[ f(u) = \frac{u^2}{u^2 + (1-u)^2} \]

As shown in Figure 4.2, the maximum-principle-satisfying DG scheme provides good approximation and at the same time eliminates all the negative values.

### 5 Two-dimensional numerical tests

In this section, we list numerical results obtained from the Cheng-Shu DG formulation on triangular meshes. We first give examples of the meshes that we use in the computation, see
Figure 4.1: The numerical results for one-dimensional porous medium equation.
(a) Numerical solutions computed by DG with and without limiter.

(b) Zoomed in results near the corner of the numerical solutions.

Figure 4.2: The numerical results for one-dimensional Buckley-Leverett equation.
Figure 5.1. As we have shown in Section 3, our maximum-principle-satisfying DG schemes put no additional constraints on the shape of the elements. To be more specific, the schemes would preserve the physical bounds strictly even if there are obtuse elements in the triangulation. All the meshes in this section are unstructured, generated by EashMesh [13]. If not specified, the unstructured meshes used in the computation are refined independently, which means for each refinement we generate a new mesh with half of the previous mesh size.

Figure 5.1: An illustration of meshes used in Section 5.1.

5.1 Accuracy test

Similar as in the one-dimensional case, we test the accuracy of the maximum-principle-satisfying DG scheme by solving:

\[ u_t + \nabla \cdot u = \varepsilon \Delta u, \quad u_0(x) = \sin(2\pi(x + y)) \]  

(5.1)
on \( (x, y) \in [0, 1] \times [0, 1] \) with periodic boundary conditions. The exact solution is \( u(x, y, t) = \exp(-8\pi^2\varepsilon t)\sin(2\pi(x + y - 2t)) \). We take \( \varepsilon = 0.0001 \) and the final time is \( t = 0.1 \). The time step is taken as suggested in (3.24). Again, we observe the expected second order convergence rate on meshes like that in Figure 5.1 (a), see Table 5.1. We also compute the numerical solutions on the obtuse mesh shown in Figure 5.1 (b), where the mesh size \( h = 0.025 \),
and the largest angle among all the elements is about \( \frac{2\pi}{3} \). From the DG scheme without limiter, we obtain the maximum and minimum of the numerical solutions to be 1.0958 and -1.0783 respectively; while by the DG scheme with limiter, the maximum and minimum of the numerical solutions are exactly 1.0000 and -1.0000, the same as indicated by the initial conditions. If we refine the unstructured obtuse mesh like Figure 5.1 (b) in a structured way, namely every triangle is divided into four similar subtriangles for each refinement, we still observe uniform second order accuracy for the DG scheme with or without limiter, see Table 5.2, as expected from the conclusions in Section 3.

<table>
<thead>
<tr>
<th>Mesh size</th>
<th>DG without Limiter</th>
<th>DG with Limiter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( L^1 ) error</td>
<td>order</td>
</tr>
<tr>
<td>0.1</td>
<td>1.71E-02</td>
<td>-</td>
</tr>
<tr>
<td>0.05</td>
<td>4.27E-03</td>
<td>2.00</td>
</tr>
<tr>
<td>0.025</td>
<td>1.05E-03</td>
<td>2.02</td>
</tr>
<tr>
<td>0.0125</td>
<td>2.56E-04</td>
<td>2.04</td>
</tr>
<tr>
<td>0.00625</td>
<td>6.12E-05</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td>( L^1 ) error</td>
<td>order</td>
</tr>
<tr>
<td></td>
<td>6.00E-03</td>
<td>2.11</td>
</tr>
<tr>
<td></td>
<td>1.33E-03</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>2.85E-04</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>6.19E-05</td>
<td>2.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mesh size</th>
<th>DG without Limiter</th>
<th>DG with Limiter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( L^1 ) error</td>
<td>order</td>
</tr>
<tr>
<td>0.025</td>
<td>4.21E-03</td>
<td>2.24</td>
</tr>
<tr>
<td>0.0125</td>
<td>8.92E-04</td>
<td>2.22</td>
</tr>
<tr>
<td>0.00625</td>
<td>1.92E-04</td>
<td>2.22</td>
</tr>
<tr>
<td>0.003125</td>
<td>4.21E-05</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td>( L^1 ) error</td>
<td>order</td>
</tr>
<tr>
<td></td>
<td>7.29E-03</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>1.51E-03</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>2.85E-04</td>
<td>2.41</td>
</tr>
</tbody>
</table>

5.2 Porous medium equations

To test the proposed scheme with the two-dimensional porous medium equation:

\[ u_t = \Delta(u^2) \quad (5.2) \]

with a periodic boundary condition and the initial condition

\[ u_0(x, y) = \begin{cases} 
  1, & \text{if } (x, y) \in [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}] \\
  0, & \text{otherwise on } (x, y) \in [-1, 1] \times [-1, 1],
\end{cases} \quad (5.3) \]
we compute our numerical solution up to time $t = 0.005$. For this degenerate parabolic equation, if we solve with the DG scheme only, the non-physical negative values will grow as time steps accumulate, while the maximum-principle-satisfying DG scheme will always keep the numerical solution nonnegative, thus also ensuring the well-posedness of the equation (5.2) and the stability of the numerical scheme. In Figure 5.2 (b), we compare our numerical results with the non-conventional finite volume scheme proposed in [25]. We also compare our maximum-principle-satisfying DG scheme with the unlimited DG scheme, see Figure 5.3. Here we only computed our numerical solutions up to a relatively shorter time, since the negative values in the unlimited DG scheme will grow as time accumulates, and eventually lead to blowing up of the numerical solution. Nevertheless, our limited DG solution can be used for long time. The minimum of our numerical solutions on different meshes is reported in Table 5.2, which turns out to be identically zero.

![Graph and Diagrams](image)

Figure 5.2: The figure on the left and the dotted curve on the right are the numerical solutions at $T = 0.005$ of the maximum-principle-satisfying DG scheme on a $h = 0.0125$ mesh. The figure on the right is the cut of the numerical solution computed by maximum-principle-satisfying DG compared with the reference solution computed by the maximum-principle-satisfying finite volume WENO scheme [25] on a $128 \times 128$ mesh.
Table 5.3: The minimum of the numerical solutions at $t = 0.005$.

<table>
<thead>
<tr>
<th>Mesh size</th>
<th>0.05</th>
<th>0.025</th>
<th>0.0125</th>
<th>0.00625</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG with limiter</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 5.3: The numerical solutions computed by DG with and without limiter at $t = 0.0005$. The mesh size $h = 0.05$

6 Applications to incompressible Navier-Stokes equations

6.1 Preliminaries

To extend our maximum-principle-satisfying DG scheme to incompressible Navier-Stokes equations, we consider the following vorticity and stream function formulation:

$$w_t + (uw)_x + (vw)_y = \frac{1}{Re} \Delta w$$

(6.1)

$$\Delta \phi = w, \quad \langle u, v \rangle = \langle -\phi_y, \phi_x \rangle$$

(6.2)

$$w(x, y, 0) = w_0(x, y), \quad \langle u, v \rangle \cdot \vec{v} = given \ on \ \partial \Omega.$$ 

The exact solution of (6.1)-(6.2) satisfies the maximum principle, i.e.,

$$w(x, y, t) \in [m, M], \quad \forall (x, y) \in \Omega, \quad \forall t \geq 0$$
where \( m = \min w_0(x, y) \), \( M = \max w_0(x, y) \). In [12], Liu and Shu developed a high order DG method for solving (6.1). The flowchart of the scheme is:

- Solve (6.2) by a standard continuous finite element Poisson solver for the stream function \( \phi \).
- Take \( \mathbf{u} = \langle u, v \rangle = -\langle \phi_y, \phi_x \rangle \), notice here \( \langle u, v \rangle \cdot \mathbf{\nu}_i \) is continuous across the edges \( e_i \) since \( \phi \) is continuous.
- Finally solve (6.1) by a DG scheme defined as follows: for a given triangulation \( T_h \) of the domain \( \Omega \), define the approximation spaces as:

\[
V^k_h = \{ v : v|_K \in P^k(K), \forall K \in T_h \}
\]

For given \( \mathbf{u} = \langle u, v \rangle \), which is obtained from the previous step, and test function \( \varphi \in V^k_h \), the DG scheme is to find \( w \in V^k_h \) such that

\[
\int_K (\partial_t w) \varphi dx - \int_K w \mathbf{u} \cdot \nabla \varphi dx + \sum_{e \in \partial K} \int_e \mathbf{u} \cdot \mathbf{\nu} \hat{w} \varphi ds = \frac{1}{Re} \left( \sum_{e \in \partial K} \int_e (\tilde{\nabla} w \cdot \mathbf{\nu}) \varphi ds - \sum_{e \in \partial K} \int_e \tilde{\omega}(\nabla \varphi \cdot \mathbf{\nu}) ds + \int_K w \Delta \varphi dx \right)
\]  

(6.3)

Since \( \mathbf{u} \cdot \mathbf{\nu} \) is continuous across the element edges, \( \mathbf{u} \cdot \mathbf{\nu} \hat{w} \) can be defined as the Lax-Friedrichs upwind biased flux, See [30]:

\[
\mathbf{u} \cdot \mathbf{\nu} \hat{w} = \frac{1}{2} [\mathbf{u} \cdot \mathbf{\nu} (w^{in} + w^{out}) - a(w^{in} - w^{out})],
\]

(6.4)

where \( a \) is taken to be the maximum of \( |\mathbf{u} \cdot \mathbf{\nu}| \), either globally or locally. The diffusion fluxes \( \tilde{\nabla} w \cdot \mathbf{\nu}, \tilde{\omega} \) are chosen as described in Section 3, (3.7):

\[
\tilde{\nabla} w \cdot \mathbf{\nu} = \nabla w^- + \frac{\alpha}{h_K} (w^{out} - w^{in}), \quad \tilde{\omega} = w^+ \quad (6.5)
\]

We use the same triangular mesh as in Section 3, and both the Poisson solver and the DG method (6.3) use \( P^1 \) elements.
In [30], it is proved that for incompressible Euler equations on triangular meshes, the high order DG scheme, which is the same as in (6.3) omitting the diffusion terms, coupled with a linear scaling limiter, satisfies the maximum principle under the CFL condition:

\[
a \frac{\Delta t}{|K|} \sum_{i=1}^{3} l_i K \leq \frac{2}{3} \hat{w}_1,
\]

where \( \hat{w}_1 \) is the one-dimensional Gauss-Lobatto quadrature weight for two cell ends. In particular, for the P1 case, we have the following CFL condition for purely convection case:

\[
a \frac{\Delta t}{|K|} \sum_{i=1}^{3} l_i K \leq \frac{1}{3}
\]

(6.7)

Following the discussion in Section 3 and the result in [30], Theorem 3.1 also holds for (6.1).

### 6.2 Accuracy test

Firstly, we solve (6.1) with the initial condition: \( w(x, y, 0) = -2 \sin(x) \sin(y) \) on \([0, 2\pi] \times [0, 2\pi]\) and periodic boundary conditions. The exact solution is available: \( w(x, y, t) = -2 \sin(x) \sin(y) \exp(-2t/Re) \). In the computation, the Reynolds number is \( Re = 100 \) and the final time is \( T = 0.1 \). As shown in Table 6.1, the maximum-principle-satisfying DG scheme is second order accurate.

<table>
<thead>
<tr>
<th>Mesh size</th>
<th>( L^1 ) error</th>
<th>order</th>
<th>( L^\infty ) error</th>
<th>order</th>
<th>( L^1 ) error</th>
<th>order</th>
<th>( L^\infty ) error</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>3.36E-02</td>
<td>–</td>
<td>2.22E-01</td>
<td>–</td>
<td>3.85E-02</td>
<td>–</td>
<td>2.50E-01</td>
<td>–</td>
</tr>
<tr>
<td>0.4</td>
<td>8.71E-03</td>
<td>1.94</td>
<td>5.84E-02</td>
<td>1.93</td>
<td>8.73E-03</td>
<td>2.14</td>
<td>5.83E-02</td>
<td>2.10</td>
</tr>
<tr>
<td>0.2</td>
<td>2.18E-03</td>
<td>2.00</td>
<td>1.57E-02</td>
<td>1.90</td>
<td>2.18E-03</td>
<td>2.00</td>
<td>1.57E-02</td>
<td>1.90</td>
</tr>
<tr>
<td>0.1</td>
<td>5.78E-04</td>
<td>1.92</td>
<td>3.66E-03</td>
<td>2.10</td>
<td>5.78E-04</td>
<td>1.92</td>
<td>3.72E-03</td>
<td>2.07</td>
</tr>
<tr>
<td>0.05</td>
<td>1.50E-04</td>
<td>1.95</td>
<td>1.00E-03</td>
<td>1.87</td>
<td>1.50E-04</td>
<td>1.92</td>
<td>1.00E-03</td>
<td>1.90</td>
</tr>
</tbody>
</table>

### 6.3 The vortex patch problem

We test our maximum-principle-satisfying DG scheme by solving Navier-Stokes equations with the initial condition:

\[
w_0(x, y, 0) = \begin{cases} 
-1, & \text{if } (x, y) \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right] \times \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right] \\
1, & \text{if } (x, y) \in \left[-\frac{3\pi}{4}, \frac{\pi}{4}\right] \times \left[-\frac{3\pi}{4}, \frac{\pi}{4}\right] \\
0, & \text{if otherwise on } [0, 2\pi] \times [0, 2\pi]
\end{cases}
\]

(6.8)
and periodic boundary conditions. Reynolds number is specified to be \( Re = 100 \). We compare our numerical scheme with the unlimited DG scheme at \( t = 0.1 \) (see Table 6.2) and \( t = 1 \) (see Figure 6.1).

<table>
<thead>
<tr>
<th>Mesh size</th>
<th>DG without Limiter</th>
<th>DG with Limiter</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>Min: -1.002485131209995, Max: 1.008209167432207</td>
<td>Min: -0.996833462986856, Max: 1.000000000000000</td>
</tr>
<tr>
<td>0.4</td>
<td>Min: -1.012248342818721, Max: 1.029509746350882</td>
<td>Min: -1.000000000000000, Max: 1.000000000000000</td>
</tr>
<tr>
<td>0.2</td>
<td>Min: -1.063529424090108, Max: 1.048046525326357</td>
<td>Min: -1.000000000000003, Max: 1.000000000000016</td>
</tr>
<tr>
<td>0.1</td>
<td>Min: -1.028127448686370, Max: 1.037932424206786</td>
<td>Min: -1.0000000000000000, Max: 1.0000000000000000</td>
</tr>
</tbody>
</table>

7 Concluding remarks

In this paper, we present second order accurate maximum-principle-satisfying discontinuous Galerkin schemes for convection-diffusion equations. Through careful theoretical analysis and extensive numerical tests, we show that under suitable CFL conditions, with a simple scaling limiter involving little additional computational cost, the numerical schemes satisfy the strict maximum principle while maintaining uniform second order accuracy. The methodology works on both structured and unstructured triangular meshes, and also for two dimensional incompressible Navier-Stokes equations in the vorticity / streamfunction formulation. An important application is the positivity-preserving property for nonlinear degenerate parabolic equations, which may become ill-posed for negative values. The effectiveness of the maximum-principle-satisfying DG schemes has been demonstrated through extensive numerical examples including two-dimensional porous medium equations and the vortex patch problems. For maximum principle preserving, the proposed schemes require only slightly more restrictive time step restrictions than that required for linear stability, with explicit SSP time discretizations. Generalizations to implicit time marching schemes constitute our ongoing work. It would also be interesting to generalize the result to higher order discontinuous Galerkin schemes, although this appears to be difficult and there might be a need to modify the numerical fluxes used in the DG formulation.
Figure 6.1: The contour of vorticity at time $t = 1$.
References


[9] A. Genty and C. Le Potier, *Maximum and minimum principles for radionuclide transport calculations in geological radioactive waste repository: Comparison between a mixed*


