

A positivity-preserving high order discontinuous Galerkin scheme for convection-diffusion equations

Sashank Srinivasan^a, Jonathan Poggie^a, Xiangxiong Zhang^{b,*}

^a*School of Aeronautics and Astronautics, Purdue University,
701 W. Stadium Ave. West Lafayette, IN 47907-2045*

^b*Department of Mathematics, Purdue University,
150 N. University Street, West Lafayette, IN 47907-2067*

Abstract

For constructing high order accurate positivity-preserving schemes for convection-diffusion equations, we construct a simple positivity-preserving diffusion flux. Discontinuous Galerkin (DG) schemes with such a positivity-preserving diffusion flux are nonlinear schemes, which can be regarded as a reduction of the high order positivity-preserving DG schemes for compressible Navier-Stokes equations in [1] to scalar diffusion operators. In this paper we focus on the local DG method to discuss how to apply such a flux. A limiter on the auxiliary variable for approximating the gradient of the solution must be used so that the diffusion flux is positivity-preserving in the sense that DG schemes with this flux satisfies a weak positivity property. Together with a positivity-preserving limiter, high order DG schemes with strong stability preserving time discretizations can be rendered positivity-preserving without losing conservation or high order accuracy for convection-diffusion problems with periodic boundary conditions or a special class of Dirichlet or Neumann boundary conditions. Numerical tests on a few parabolic equations and an application to modeling electrical discharges are shown to demonstrate the performance of this scheme.

Keywords: discontinuous Galerkin finite element method, high order accuracy, convection-diffusion equations, positivity-preserving, glow discharge, drift-diffusion

1. Introduction

1.1. The positivity-preserving property in convection-diffusion equations

We consider a generic convection-diffusion equation of the form

$$u_t + f(u)_x = a(u)_{xx}, \quad u(x, 0) = u_0(x) \quad (1)$$

*Corresponding author

Email addresses: sashank.srinivasan.13@gmail.com (Sashank Srinivasan),
jpoggie@purdue.edu (Jonathan Poggie), zhan1966@purdue.edu (Xiangxiong Zhang)

or equivalently

$$u_t + f(u)_x = (b(u)u_x)_x, \quad u(x, 0) = u_0(x), \quad (2)$$

where $b(u) = a'(u) \geq 0$. The exact solution $u(x, t)$ satisfies a maximum principle or bound-preserving property, i.e., if $\max u_0(x) = M$ and $\min u_0(x) = m$, then

$$u(x, t) \in [m, M], \quad \forall x, \forall t \geq 0.$$

The particular case $m = 0$ for the lower bound, i.e., positivity-preserving, is an important desired property for numerical schemes in many applications. For instance, negative values are physically meaningless, in radionuclide transport calculations [2], chemotaxis problems [3, 4], streamer discharges simulations [5], etc. Moreover, negative values may result in ill-posedness and instability for certain nonlinear equations. In this paper, we are interested in constructing arbitrarily high order accurate positivity-preserving schemes.

1.2. Monotone schemes

A lot of first order accurate classical schemes can be shown positivity-preserving since such low order accurate schemes are usually monotone for interested equations. Consider the the scalar convection equation as an example,

$$u_t + f(u)_x = 0. \quad (3)$$

A first order monotone scheme is given by

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} \left[\widehat{f}(u_j^n, u_{j+1}^n) - \widehat{f}(u_{j-1}^n, u_j^n) \right],$$

where u_j^n denotes the numerical solution at n -th time step and j -th grid point, and $\widehat{f}(\uparrow, \downarrow)$ is a monotone numerical flux, i.e., $\widehat{f}(\cdot, \cdot)$ is non-decreasing w.r.t. its first argument and non-increasing w.r.t. its second argument. For instance, the Lax-Friedrichs flux is monotone,

$$\widehat{f}(u, v) = \frac{1}{2} [f(u) + f(v) - \alpha(v - u)], \quad \alpha = \max_u |f'(u)|. \quad (4)$$

Let $\lambda = \frac{\Delta t}{\Delta x}$ and regard the whole right hand side of the scheme as a single function $H_\lambda(u_{j-1}^n, u_j^n, u_{j+1}^n) = u_j^n - \lambda \left[\widehat{f}(u_j^n, u_{j+1}^n) - \widehat{f}(u_{j-1}^n, u_j^n) \right]$, then $H_\lambda(\uparrow, \uparrow, \uparrow)$ is a monotone function under the CFL constraint $\lambda \max_u |f'(u)| \leq 1$. The monotonicity implies the bound-preserving property: if $u_j^n \in [m, M]$, then

$$m = H_\lambda(m, m, m) \leq u_j^{n+1} \leq H_\lambda(M, M, M) = M.$$

On the other hand, the Godunov Theorem states that a linear monotone scheme is at most first order accurate for the convection equation [6].

1.3. Weak monotonicity in finite volume type schemes for convection

Even though monotonicity is not a necessary condition for numerical schemes to preserve bounds or positivity, it is a very convenient tool for constructing positivity-preserving schemes in a simple and efficient manner. To construct high order accurate bound-preserving schemes for scalar convection, we can take advantage of weak monotonicity in finite volume type schemes including discontinuous Galerkin (DG) methods, which was first used in [7, 8, 9]. For the equation (3), consider a $(k + 1)$ -th order accurate finite volume spatial discretization (or the cell average scheme in a DG method) with forward Euler time discretization on an interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$,

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \lambda \left[\hat{f} \left(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+ \right) - \hat{f} \left(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+ \right) \right], \quad (5)$$

where \hat{f} is a monotone numerical flux, and $u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+$ are approximations to $u(x_{j+\frac{1}{2}}, t^n)$ from the left and from the right respectively. Let $u_j(x)$ be the reconstruction polynomial of degree k in a finite volume scheme (or the DG polynomial of degree k in a DG method) at time step n such that \bar{u}_j^n is the cell average of $u_j(x)$ on the interval I_j and $u_{j+\frac{1}{2}}^-$ and $u_{j-\frac{1}{2}}^+$ are nodal values of $u_j(x)$ at two cell ends. Let $N = \lceil (k + 3)/2 \rceil$, i.e., N is smallest integer satisfying $2N - 3 \geq k$. We consider an N -point Legendre Gauss-Lobatto quadrature rule on the interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, which is exact for integrals of polynomials of degree up to $2N - 3$. Denote these quadrature points on I_j as

$$S_j = \{x_{j-\frac{1}{2}} = \hat{x}_j^1, \hat{x}_j^2, \dots, \hat{x}_j^{N-1}, \hat{x}_j^N = x_{j+\frac{1}{2}}\}. \quad (6)$$

Let $\hat{\omega}_\mu$ be the quadrature weights for the interval $[-\frac{1}{2}, \frac{1}{2}]$ such that $\sum_{\mu=1}^N \hat{\omega}_\mu = 1$.

Thus we have,

$$\bar{u}_j^n = \frac{1}{\Delta x} \int_{I_j} u_j(x) dx = \sum_{\mu=1}^N \hat{\omega}_\mu u_j(\hat{x}_j^\mu) = \sum_{\mu=2}^{N-1} \hat{\omega}_\mu u_j(\hat{x}_j^\mu) + \hat{\omega}_1 u_{j-\frac{1}{2}}^+ + \hat{\omega}_N u_{j+\frac{1}{2}}^-. \quad (7)$$

After plugging (7) in, we can rewrite (5) as

$$\bar{u}_j^{n+1} = \sum_{\mu=2}^{N-1} \hat{\omega}_\mu u_j(\hat{x}_j^\mu) + \hat{\omega}_1 u_{j-\frac{1}{2}}^+ + \hat{\omega}_N u_{j+\frac{1}{2}}^- - \lambda \left[\hat{f} \left(u_{j+\frac{1}{2}}^-, u_{j+\frac{1}{2}}^+ \right) - \hat{f} \left(u_{j-\frac{1}{2}}^-, u_{j-\frac{1}{2}}^+ \right) \right].$$

Under the CFL constraint $\frac{\Delta t}{\Delta x} \max_j |f'(u_{j+\frac{1}{2}}^\pm)| \leq \hat{\omega}_1 = \hat{\omega}_N = \frac{1}{N(N-1)}$, the right hand side of the scheme above is a monotonically non-decreasing function of all nodal values involved, i.e., $u_{j\pm\frac{1}{2}}^\pm$ and $u_j(\hat{x}_j^\mu)$, which is called weak monotonicity. A simple and efficient local bound-preserving limiter can be designed to control the nodal values at \hat{x}_j^μ ($\mu = 1, \dots, N$) without affecting accuracy and conservation. Together with strong stability preserving (SSP) Runge-Kutta or multistep

methods [10], which are convex combinations of several formal forward Euler steps, a high order accurate finite volume or DG scheme can be rendered bound-preserving with this limiter. Furthermore, such a result can be easily extended to multiple dimensions on cells of arbitrary shapes.

1.4. Weak monotonicity in finite volume type schemes for diffusion

To extend the bound-preserving results for convection equations to convection-diffusion problems, we consider the heat equation $u_t = u_{xx}$ as the simplest diffusion model equation. The second order accurate centered difference scheme $u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x^2}(u_{i-1}^n - 2u_i^n + u_{i+1}^n)$ is monotone under the CFL constraint $\Delta t \leq \frac{1}{2}\Delta x^2$. But conventional higher order accurate discretizations are not monotone. Unfortunately, even the weak monotonicity holds only up to second order accuracy in a conventional linear finite volume scheme for the heat equation in the sense of local truncation error analysis, see Appendix D in [1].

A non-conventional high order finite volume satisfying the weak monotonicity was constructed in [8]. However, this approach requires new implementations and it does not apply to DG methods.

For DG methods, it is still possible to construct a third order linear scheme satisfying the weak monotonicity. With special parameters, the direct DG (DDG) method, which is a generalized version of interior penalty DG method, indeed satisfies the weak monotonicity up to third order accuracy [11, 12]. However, it does not seem to be possible to extend such a result to higher order accuracy.

To construct high order schemes satisfying the weak monotonicity, one possible approach is to explore nonlinear discretizations for diffusion operators. Nonlinear discretizations and schemes for a nonlinear equation refer to those ones which are still nonlinear if the equation reduces to a linear one. Such an effort was first made in [1] for constructing arbitrarily high order accurate positivity-preserving DG schemes for compressible Navier-Stokes equations. More specifically, a positivity-preserving flux was constructed for the nonlinear Navier-Stokes diffusion operator in [1], which is a nonlinear discretization.

To construct bound-preserving or positivity-preserving high order schemes for scalar convection-diffusion problems, it seems natural to consider a nonlinear discretization satisfying the weak monotonicity, following [1]. Such a goal was achieved in [13] by rewriting a diffusion operator as a gradient flow. In this paper, for constructing positivity-preserving high order schemes, we will discuss a different nonlinear discretization by proposing a new numerical flux for diffusion operators. The method in this paper and the scheme in [13] both can be regarded as a reduction of the scheme in [1] from the Navier-Stokes system to scalar equations. On the other hand, these two approaches are very different since our approach does not require rewriting a diffusion operator as a gradient flow, which is very convenient to apply to commonly used diffusion operators such as the Laplacian.

1.5. Alternatives to construct high order positivity-preserving schemes

For constructing high order accurate positivity-preserving schemes for parabolic equations such as the heat equation, besides studying the weak monotonicity, there are a few alternatives in the literature.

The simplest centered difference gives the second order accurate five-point discrete Laplacian for two-dimensional heat equation, which can be easily shown to be positivity-preserving with forward Euler or backward Euler time discretizations. When using the fourth order accurate nine-point discrete Laplacian with backward Euler or Crank-Nicolson, the coefficient matrix in the linear system is an M-matrix thus can be shown to be positivity-preserving under certain CFL conditions. Thus a second order accurate in time and fourth order accurate in space positivity-preserving finite difference scheme can be easily constructed when the diffusion operator is the Laplacian. Obviously this approach works only for linear diffusion operators, and generalization to nonlinear problems is nontrivial.

Another simple and efficient method to enforce positivity in high order schemes is to take a convex combination of high order flux with a first order positivity-preserving one [14, 15, 16]. Even though this method applies to various high order schemes including finite difference, finite volume and DG methods and it works well in numerical tests, not only does it not seem intuitive or obvious why a linear combination with a first order scheme does not destroy the accuracy, it is also difficult to rigorously justify its accuracy in analysis.

1.6. Contributions and organization of the paper

The main contribution in this paper is a new numerical flux for nonlinear diffusion operators, which is positivity-preserving in the sense that a high order finite volume or DG scheme with such a flux satisfies a weak positivity property. As a demonstration of how to design positivity-preserving high order accurate schemes, we mainly focus on local discontinuous Galerkin (LDG) method [17, 18] as an example. Together with two simple limiters and the new diffusion flux, the LDG method can be rendered positivity-preserving for scalar convection-diffusion problems. Moreover, the high order accuracy will not be destroyed by limiters if the non-negative exact solution $u(x, t)$ satisfies the following assumption: $u_x(x, t) = 0$ whenever $u(x, t) = 0$. Such an assumption holds for any non-negative smooth solutions for periodic boundary conditions since the derivative of any smooth function becomes zero at any interior global minimum. However, the assumption may not be true at the boundary points. A simple counter example is an analytical solution $u(x, t) = e^{-t} \sin x$ for the heat equation $u_t = u_{xx}$ on the interval $[0, \pi]$ with Dirichlet or Neumann boundary conditions. The exact solution $u(x, t) = e^{-t} \sin x$ is non-negative on $[0, \pi]$ but at $x = 0$ we have $u(0, t) = 0$ and $u_x(0, t) = e^{-t}$. For such a problem, the limiters in this paper will destroy the high order accuracy in DG schemes.

The paper is organized as follows: we demonstrate the main idea and implementation details for the one dimensional case in Section 2. In Section 3, we discuss the two dimensional case. In Section 4 numerical tests will be shown.

Section 5 is an application of the positivity-preserving high order schemes to modeling electrical discharges using the Drift-Diffusion model. Section 6 consists of concluding remarks.

2. A positivity-preserving DG scheme in one dimension

In this section we first construct a diffusion flux, with which the high order DG schemes satisfy some weak positivity property for solving one-dimensional convection-diffusion problems. Then we show that together with two simple limiters the high order local DG scheme with this diffusion flux can be rendered positivity-preserving.

2.1. A positivity-preserving flux for the diffusion terms

We first consider the diffusion equation. We only consider periodic boundary conditions. For solving a generic diffusion equation of the form

$$u_t = a(u)_{xx}, \quad u(x, 0) = u_0(x), \quad (8)$$

or equivalently

$$u_t = (b(u)u_x)_x, \quad u(x, 0) = u_0(x), \quad (9)$$

where $b(u) = a'(u) \geq 0$, there are quite a few different DG formulations such as the interior penalty DG method [19, 20, 21, 22], the local DG method [17, 18], the scheme by Baumann and Oden [23], the scheme by Cheng and Shu [24], Compact DG [25], direct DG [26, 27, 28], correction procedure via reconstruction (CPR) [29, 30], Hybrid DG [31] and Embedded DG [32], etc. In this paper, we focus on the local DG (LDG) method, and the ideas presented can be adapted to other DG methods. The equation (9) is equivalent to,

$$u_t = (b^*(u)q)_x, \quad q = B(u)_x,$$

where $b^*(u) = \sqrt{b(u)}$ and $B(u) = \int^u b^*(s)ds$. On a computational domain $x \in [a, b]$ discretized as $a = x_{\frac{1}{2}} < x_{\frac{3}{2}} \cdots < x_{N_x - \frac{1}{2}} < x_{N_x + \frac{1}{2}} = b$ by a mesh T_h with N_x cells I_j of size $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, with the approximation space chosen as $V_h^k = \{v : v|_{I_j} \in P^k(I_j), \forall I_j \in T_h\}$, in the local DG method we seek $u_h, q_h \in V_h^k$ such that for test functions $\psi_h, \phi_h \in V_h^k$:

$$\int_{I_j} \frac{\partial}{\partial t} u_h \psi_h dx = \left(\widehat{b}^*(u_h) \widehat{q}_h \psi_h \right)_{j+\frac{1}{2}} - \left(\widehat{b}^*(u_h) \widehat{q}_h \psi_h \right)_{j-\frac{1}{2}} - \int_{I_j} b^*(u_h) q_h \frac{\partial}{\partial x} \psi_h dx, \quad (10)$$

$$\int_{I_j} q_h \phi_h dx = \left(\widehat{B}(u_h) \phi_h \right)_{j+\frac{1}{2}} - \left(\widehat{B}(u_h) \phi_h \right)_{j-\frac{1}{2}} - \int_{I_j} B(u_h) \frac{\partial}{\partial x} \phi_h dx.$$

The diffusion fluxes can be taken in an alternating fashion as $\widehat{B}(u) = B(u^+)$, $\widehat{q} = q^-$ and $\widehat{b}^*(u) = \frac{B(u^+) - B(u^-)}{u^+ - u^-}$.

Without loss of generality, we assume a uniform mesh, i.e., $\Delta x_j = \Delta x$. By taking the test function $\psi_h \equiv 1$ in (10), we obtain the time evolution of the cell average of u_h as

$$\frac{d}{dt} \bar{u}_h = \frac{\left(\widehat{b}^*(u)\widehat{q}\right)_{j+\frac{1}{2}} - \left(\widehat{b}^*(u)\widehat{q}\right)_{j-\frac{1}{2}}}{\Delta x}. \quad (11)$$

With the first order explicit forward Euler time stepping, (11) becomes

$$\bar{u}_h^{n+1} = \bar{u}_h^n + \frac{\Delta t}{\Delta x} \left[\left(\widehat{b}^*(u^n)\widehat{q}\right)_{j+\frac{1}{2}} - \left(\widehat{b}^*(u^n)\widehat{q}\right)_{j-\frac{1}{2}} \right]. \quad (12)$$

Higher-order explicit time discretizations will be discussed later in Section 2.5. For simplicity, the subscript h and superscript n will be omitted.

Now consider the following diffusion flux \widehat{q} :

$$\widehat{q}_{j+\frac{1}{2}} = \frac{1}{2} \left[q_{j+\frac{1}{2}}^+ + q_{j+\frac{1}{2}}^- + \beta_{j+\frac{1}{2}} \left(u_{j+\frac{1}{2}}^+ - u_{j+\frac{1}{2}}^- \right) \right], \quad (13a)$$

where the penalty parameter β is defined to achieve a weak positivity property for the scheme (12):

$$\beta_{j+\frac{1}{2}} = \begin{cases} \max \left\{ \left| \frac{q_{j+\frac{1}{2}}^+}{u_{j+\frac{1}{2}}^+} \right|, \left| \frac{q_{j+\frac{1}{2}}^-}{u_{j+\frac{1}{2}}^-} \right| \right\}, & \text{if } u_{j+\frac{1}{2}}^+ \neq 0 \text{ and } u_{j+\frac{1}{2}}^- \neq 0, \\ \left| \frac{q_{j+\frac{1}{2}}^+}{u_{j+\frac{1}{2}}^+} \right|, & \text{if } u_{j+\frac{1}{2}}^+ \neq 0 \text{ and } u_{j+\frac{1}{2}}^- = 0, \\ \left| \frac{q_{j+\frac{1}{2}}^-}{u_{j+\frac{1}{2}}^-} \right|, & \text{if } u_{j+\frac{1}{2}}^+ = 0 \text{ and } u_{j+\frac{1}{2}}^- \neq 0, \\ 0 & \text{if } u_{j+\frac{1}{2}}^+ = 0 \text{ and } u_{j+\frac{1}{2}}^- = 0. \end{cases} \quad (13b)$$

Plugging in (7), the scheme (12) with the flux (13) can be written as

$$\begin{aligned} \bar{u}_j^{n+1} &= \bar{u}_j^n + \lambda \left[\left(\widehat{b}^*\widehat{q}\right)_{j+\frac{1}{2}} - \left(\widehat{b}^*\widehat{q}\right)_{j-\frac{1}{2}} \right] \\ &= \sum_{\mu=2}^{N-1} \widehat{\omega}_\mu u_j(\widehat{x}_j^\mu) + \widehat{\omega}_1 u_{j-\frac{1}{2}}^+ + \widehat{\omega}_N u_{j+\frac{1}{2}}^- + \lambda \left[\left(\widehat{b}^*\widehat{q}\right)_{j+\frac{1}{2}} - \left(\widehat{b}^*\widehat{q}\right)_{j-\frac{1}{2}} \right] \\ &= \sum_{\mu=2}^{N-1} \widehat{\omega}_\mu u_j(\widehat{x}_j^\mu) + u_{j-\frac{1}{2}}^+ \left[\widehat{\omega}_1 - \frac{1}{2} \lambda \widehat{b}_{j-\frac{1}{2}}^* \left(\frac{q_{j-\frac{1}{2}}^+}{u_{j-\frac{1}{2}}^+} + \beta_{j-\frac{1}{2}} \right) \right] \\ &\quad + u_{j+\frac{1}{2}}^+ \left[\frac{1}{2} \lambda \widehat{b}_{j+\frac{1}{2}}^* \left(\frac{q_{j+\frac{1}{2}}^+}{u_{j+\frac{1}{2}}^+} + \beta_{j+\frac{1}{2}} \right) \right] + u_{j-\frac{1}{2}}^- \left[\frac{1}{2} \lambda \widehat{b}_{j-\frac{1}{2}}^* \left(\beta_{j-\frac{1}{2}} - \frac{q_{j-\frac{1}{2}}^-}{u_{j-\frac{1}{2}}^-} \right) \right] \\ &\quad + u_{j+\frac{1}{2}}^- \left[\widehat{\omega}_N + \frac{1}{2} \lambda \widehat{b}_{j+\frac{1}{2}}^* \left(\frac{q_{j+\frac{1}{2}}^-}{u_{j+\frac{1}{2}}^-} - \beta_{j+\frac{1}{2}} \right) \right], \end{aligned} \quad (14)$$

For the right hand side in (14) to be a positive linear combination of point values $u_{j\pm\frac{1}{2}}^\pm$ and $u_j(\hat{x}_j^\mu)$, it suffices to require $q = 0$ wherever $u = 0$ and the time step to satisfy the following constraint (notice that $\widehat{\omega}_1 = \widehat{\omega}_N$):

$$\max_j \left\{ \frac{1}{2} \widehat{b}_{j+\frac{1}{2}}^* \left(\beta_{j+\frac{1}{2}} - \frac{q_{j+\frac{1}{2}}^-}{u_{j+\frac{1}{2}}^-} \right), \frac{1}{2} \widehat{b}_{j-\frac{1}{2}}^* \left(\beta_{j-\frac{1}{2}} + \frac{q_{j-\frac{1}{2}}^+}{u_{j-\frac{1}{2}}^+} \right) \right\} \frac{\Delta t}{\Delta x} \leq \widehat{\omega}_1.$$

Notice that the Mean Value Theorem guarantees the existence of some ξ such that $\widehat{b}^*(u)_{j+\frac{1}{2}} = [B(u_{j+\frac{1}{2}}^+) - B(u_{j+\frac{1}{2}}^-)] / (u_{j+\frac{1}{2}}^+ - u_{j+\frac{1}{2}}^-) = B'(\xi) \geq 0$. Thus the condition on the time step above gives a CFL constraint for an explicit scheme:

$$\frac{\Delta t}{\Delta x} \max_j \widehat{b}_{j+\frac{1}{2}}^* \beta_{j+\frac{1}{2}} \leq \widehat{\omega}_1 = \frac{1}{N(N-1)}, \quad (15)$$

which is an acceptable time step constraint if β is not too large.

Theorem 1. *For the high order scheme (12) using the flux (13), under the CFL constraint (15), the cell average $\bar{u}_j^{n+1} \geq 0$ if the following are satisfied:*

1. *All point values at the Gauss-Lobatto quadrature points are non-negative: $u_j(\hat{x}_j^\mu) \geq 0$ for $\mu = 1, \dots, N$ and all j , which include $u_{j\pm\frac{1}{2}}^\pm$.*
2. *For all j , $q_{j+\frac{1}{2}}^- = 0$ if $u_{j+\frac{1}{2}}^- = 0$, and $q_{j+\frac{1}{2}}^+ = 0$ if $u_{j+\frac{1}{2}}^+ = 0$.*

Remark 2. *For a smooth solution $u(x, t) \geq 0$ defined on the whole real line or periodically defined on an interval, if $u(x, t) = 0$ at $x = x^*$, then $x = x^*$ is a critical point of $u(x, t)$ since zero is the global minimum value of $u(x, t)$. Thus $x = x^*$ is also a critical point of $B(u(x, t))$ since $\frac{\partial}{\partial x} B(u(x^*, t)) = B'(u(x^*, t))u_x(x^*, t) = 0$. However, for a smooth function $u(x, t) \geq 0$ defined on an interval, $u(x, t) = 0 \implies u_x(x, t) = 0$ may not be true at the two cell ends, e.g., $u(x, t) = e^{-t} \sin x$ on the interval $[0, \pi]$. Therefore the second condition in Theorem 1 is satisfied by the exact solution for periodic boundary conditions or a special class of other boundary conditions satisfying $u(x, t) = 0 \implies u_x(x, t) = 0$ at cell ends.*

Remark 3. *Theorem 1 is a positivity-preserving result, namely, (14) is only a positive linear combination rather than a convex combination. To have a bound-preserving result, then we need (14) to be a convex combination, which is nontrivial. On the other hand, for diffusion problems, the lower bound $m = 0$ is usually much more interesting or important to preserve than other lower bounds or an upper bound.*

Remark 4. *The main idea here is to use a Lax-Friedrichs type diffusion flux (13) to achieve positivity, which can be used in other DG formulations as well. For example, consider the Cheng-Shu method [24] for (8), given by*

$$\int_{I_j} u_t v dx = (\widehat{a}_x v)_{j+\frac{1}{2}} - (\widehat{a}_x v)_{j-\frac{1}{2}} + (\widehat{a} v_x)_{j-\frac{1}{2}} - (\widehat{a} v_x)_{j+\frac{1}{2}} + \int_{I_j} a(u) v_{xx} dx, \quad (16)$$

where we take the flux \widehat{a}_x as in (13), i.e., $\widehat{a}_x = \frac{[a(u)]}{[u]} (u_x^- + u_x^+ + \beta(u^+ - u^-))$, and β can be similarly defined to achieve the similar results as in Theorem 1.

Remark 5. We can also define a positivity-preserving flux for $(\widehat{b}^* \widehat{q})_{j+\frac{1}{2}}$ with which the scheme (12) is equivalent to (12) with the flux (13):

$$(\widehat{b}^* \widehat{q})_{j+\frac{1}{2}} = \frac{1}{2} \left[\widehat{b}_{j+\frac{1}{2}}^* q_{j+\frac{1}{2}}^+ + \widehat{b}_{j+\frac{1}{2}}^* q_{j+\frac{1}{2}}^- + \beta_{j+\frac{1}{2}} (u_{j+\frac{1}{2}}^+ - u_{j+\frac{1}{2}}^-) \right],$$

and

$$\beta = \begin{cases} \max \left\{ \left| \frac{\widehat{b}^* q^+}{u^+} \right|, \left| \frac{\widehat{b}^* q^-}{u^-} \right| \right\}, & \text{if } u^+ \neq 0 \text{ and } u^- \neq 0, \\ \left| \frac{\widehat{b}^* q^+}{u^+} \right|, & \text{if } u^+ \neq 0 \text{ and } u^- = 0, \\ \left| \frac{\widehat{b}^* q^-}{u^-} \right|, & \text{if } u^+ = 0 \text{ and } u^- \neq 0, \\ 0 & \text{if } u^+ = 0 \text{ and } u^- = 0. \end{cases}$$

Remark 6. For solving the heat equation $u_t = u_{xx}$, $(k+1)$ -th order accuracy can be proven for the local DG scheme using P^k basis with alternating fluxes, i.e., the scheme (10) with $\widehat{B}(u) = u^+$ and $\widehat{q} = q^-$. While for the local DG scheme with centered fluxes, i.e., $\widehat{B}(u) = \frac{1}{2}(u^+ + u^-)$ and $\widehat{q} = \frac{1}{2}(q^- + q^+)$, the order of accuracy is still $k+1$ if k is even, but only k -th order accuracy can be obtained if k is odd. See [18] for the analysis. Since the flux (13) reduces to a centered flux when $\beta = 0$, we would expect the order of accuracy cannot exceed k in general for the odd k . In the numerical tests, we observe at least k -th order accuracy for the flux (13) for odd k and $(k+1)$ -th order accuracy for even k . On the other hand, it would be highly nontrivial to rigorously prove such a claim on the order of accuracy as those for alternating and centered fluxes in [18] since DG schemes with the flux (13) is a nonlinear scheme.

2.2. Convection-diffusion problems

The local DG method for solving (2) can be written as

$$\begin{aligned} \int_{I_j} \frac{\partial}{\partial t} u_h \psi_h dx &= -(\widehat{f} \psi_h)_{j+\frac{1}{2}} + (\widehat{f} \psi_h)_{j-\frac{1}{2}} + \int_{I_j} f(u_h) \frac{\partial}{\partial x} \psi_h dx \\ &\quad + (\widehat{b}^* \widehat{q} \psi_h)_{j+\frac{1}{2}} - (\widehat{b}^* \widehat{q} \psi_h)_{j-\frac{1}{2}} - \int_{I_j} b^*(u_h) q_h \frac{\partial}{\partial x} \psi_h dx, \\ \int_{I_j} q_h \phi_h dx &= (\widehat{B}(u_h) \phi_h)_{j+\frac{1}{2}} - (\widehat{B}(u_h) \phi_h)_{j-\frac{1}{2}} - \int_{I_j} B(u_h) \frac{\partial}{\partial x} \phi_h dx. \end{aligned}$$

By taking the test function $\psi_h \equiv 1$ and using the forward Euler time discretization, we obtain

$$\begin{aligned} \bar{u}^{n+1} &= \bar{u}^n - \frac{\Delta t}{\Delta x} \left[\widehat{f}_{j+\frac{1}{2}} - \widehat{f}_{j-\frac{1}{2}} \right] + \frac{\Delta t}{\Delta x} \left[(\widehat{b}^* \widehat{q})_{j+\frac{1}{2}} - (\widehat{b}^* \widehat{q})_{j-\frac{1}{2}} \right] \\ &= \frac{1}{2} \left(\bar{u}^n - 2 \frac{\Delta t}{\Delta x} \left[\widehat{f}_{j+\frac{1}{2}} - \widehat{f}_{j-\frac{1}{2}} \right] \right) + \frac{1}{2} \left(\bar{u}^n + 2 \frac{\Delta t}{\Delta x} \left[(\widehat{b}^* \widehat{q})_{j+\frac{1}{2}} - (\widehat{b}^* \widehat{q})_{j-\frac{1}{2}} \right] \right). \end{aligned} \tag{17}$$

Thus we split the scheme (17) as an average of the convection part $\bar{u}^n - 2\lambda \left[\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} \right]$ and the diffusion part $\bar{u}^n + 2\lambda \left[(\hat{b}^* \hat{q})_{j+\frac{1}{2}} - (\hat{b}^* \hat{q})_{j-\frac{1}{2}} \right]$. With Theorem 1 and the results reviewed in Section 1, we have

Theorem 7. *If we use the fluxes (4) and (13) in the DG scheme (17), under the CFL constraint*

$$\frac{\Delta t}{\Delta x} \max_j |f'(u_{j+\frac{1}{2}}^\pm)| \leq \frac{1}{2} \hat{\omega}_1 = \frac{1}{2N(N-1)}, \quad \frac{\Delta t}{\Delta x} \max_j \hat{b}_{j+\frac{1}{2}}^* \beta_{j+\frac{1}{2}} \leq \frac{1}{2N(N-1)}. \quad (18)$$

the cell average $\bar{u}_j^{n+1} \geq 0$ if the following are satisfied:

1. All point values at the Gauss-Lobatto quadrature points are non-negative: $u_j(\hat{x}_j^\mu) \geq 0$ for $\mu = 1, \dots, N$ and all j , which include $u_{j\pm\frac{1}{2}}^\pm$.
2. For all j , $q_{j+\frac{1}{2}}^- = 0$ if $u_{j+\frac{1}{2}}^- = 0$, and $q_{j+\frac{1}{2}}^+ = 0$ if $u_{j+\frac{1}{2}}^+ = 0$.

2.3. The positivity-preserving limiter

For DG methods, the first condition in Theorem 1 and Theorem 7 can be easily enforced by the simple scaling limiter in [7]. Let $u_j(x)$ be the DG polynomial at n -th time step for the cell I_j with the cell average $\bar{u}_j^n \geq 0$, then the limiter can be defined as:

$$\tilde{u}_j(x) = \bar{u}_j^n + \theta_j (u_j(x) - \bar{u}_j^n), \theta_j = \min \left(1, \left| \frac{\bar{u}_j^n}{\bar{u}_j^n - \min_\mu u_j(\hat{x}_j^\mu)} \right| \right). \quad (19)$$

Here $\tilde{u}_j(x)$ is the modified polynomial after scaling. The limiter ensures $\tilde{u}_j(\hat{x}_j^\mu) \geq 0$ without changing the cell average. Moreover, it does not destroy the high order accuracy, see [7, 8, 1].

2.4. A limiter for the auxiliary variable q

To enforce the second condition in Theorem 1 and Theorem 7, we consider an additional modification on $q_j(x)$ in this subsection.

Whenever $|u_h| < \epsilon$ for some small positive number ϵ at cell ends, we would like to enforce $q_h(x)$ to be zero at the same location, which can be easily achieved by employing a second limiter.

In the nodal representation [33], the degree of freedoms of $q_j(x)$, a polynomial degree k , on an interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ can be represented by its $k+1$ point values $q_j(x_j^\alpha)$, where x_j^α ($\alpha = 1, \dots, k+1$) can be chosen as the $(k+1)$ -point Gauss-Lobatto quadrature points on I_j . In particular, $x_j^1 = x_{j-\frac{1}{2}}$ and $x_j^{k+1} = x_{j+\frac{1}{2}}$. Let $\tilde{q}_j(x)$ be the modified polynomial. We need to achieve $\tilde{q}_j(x_{j+\frac{1}{2}}) = 0$ if $|\tilde{u}_j(x_{j+\frac{1}{2}})| \leq \epsilon$ and $\tilde{q}_j(x_{j-\frac{1}{2}}) = 0$ if $|\tilde{u}_j(x_{j-\frac{1}{2}})| \leq \epsilon$. We also

want to minimized the distance between $\tilde{q}_j(x)$ and $q_j(x)$. Thus we can define the limiter as the following:

$$\begin{aligned} \tilde{q}_j(x_j^\alpha) &= q_j(x_j^\alpha), \quad \alpha = 2, \dots, k \\ \tilde{q}_j(x_j^1) &= \begin{cases} q_j(x_j^1), & \text{if } \left| \tilde{u}_j(x_{j-\frac{1}{2}}) \right| \geq \epsilon \\ 0, & \text{otherwise} \end{cases} \\ \tilde{q}_j(x_j^{k+1}) &= \begin{cases} q_j(x_j^{k+1}), & \text{if } \left| \tilde{u}_j(x_{j+\frac{1}{2}}) \right| \geq \epsilon \\ 0, & \text{otherwise} \end{cases}, \end{aligned} \quad (20)$$

which minimizes the distance defined as

$$\|q_j(x) - \tilde{q}_j(x)\| = \sum_{\alpha=1}^{k+1} |q_j(x_j^\alpha) - \tilde{q}_j(x_j^\alpha)|^2.$$

In a modal representation, the polynomial q_j is represented as a linear combination of basis polynomials $p^i(x)$ ($i = 0, \dots, k$):

$$q_j(x) = \sum_{i=0}^k q_j^i p^i(x).$$

Let $\tilde{q}_j(x) = \sum_{i=0}^k \tilde{q}_j^i p^i(x)$. Then the implementation of the limiter (20) is equivalent to solving a linear system of equations:

$$\begin{bmatrix} p^0(x_j^1) & p^1(x_j^1) & p^2(x_j^1) & \dots & p^k(x_j^1) \\ p^0(x_j^2) & p^1(x_j^2) & p^2(x_j^2) & \dots & p^k(x_j^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p^0(x_j^k) & p^1(x_j^k) & p^2(x_j^k) & \dots & p^k(x_j^k) \\ p^0(x_j^{k+1}) & p^1(x_j^{k+1}) & p^2(x_j^{k+1}) & \dots & p^k(x_j^{k+1}) \end{bmatrix} \begin{bmatrix} \tilde{q}_j^1 \\ \tilde{q}_j^2 \\ \vdots \\ \tilde{q}_j^k \\ \tilde{q}_j^{k+1} \end{bmatrix} = \begin{bmatrix} \tilde{q}_j(x_j^0) \\ q_j(x_j^1) \\ \vdots \\ q_j(x_j^{k-1}) \\ \tilde{q}_j(x_j^k) \end{bmatrix}.$$

The parameter β in the positivity-preserving diffusion flux (13) is essentially of the form q/u , thus the limiter (20) is necessary to remove the singularities around $u = 0$.

On the other hand, the limiter (20) does not destroy the accuracy of DG polynomials for a smooth solution with finite number of extrema.

Without loss of generality, assume the smooth function $f(x) = u(x, t^n)$ has only one interior global minimum with minimum value being zero. Let x^* be the minimum point, then $f(x^*) = 0$ and $f'(x^*) = 0$. Then $f(x)$ has no other local extrema in a small neighbourhood of x^* . For a very small number $\delta > 0$, $f(x) < \delta$ implies x is in a neighbourhood of x^* , i.e., $x \in (x^* - h, x^* + h)$ for some $h > 0$, and $f(x)$ is monotone on the two intervals $(x^* - h, x^*)$ and $(x^*, x^* + h)$ respectively. Let $f^{-1}(x)$ denote the inverse function of $f(x)$ on these two intervals $(x^* - h, x^*]$ and $[x^*, x^* + h)$ respectively. Assume $f(x)$ is smooth enough so that $f^{-1}(x)$ is lipschitz continuous with Lipschitz constant L on the two intervals respectively, where L does not depend on h or δ . Then

$f(x) - f(x^*) = f(x) < \delta$ implies that $|x - x^*| \leq L\delta = \mathcal{O}(\delta)$. By Taylor expansion of $f'(x)$, if $f(x) < \mathcal{O}(\delta)$, then $f'(x) = f'(x^*) + f''(x^*)\mathcal{O}(\delta) + \mathcal{O}(\delta^2) = \mathcal{O}(\delta)$.

Suppose $\tilde{u}_j(x)$ and $q_j(x)$ are high order accurate approximations to $u(x, t^n)$ and $\frac{\partial}{\partial x}B(u(x, t^n))$ on the interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ respectively, e.g., $\tilde{u}_j(x) - u(x, t^n) = \mathcal{O}(\Delta x^{k+1})$ and $q_j(x) - \frac{\partial}{\partial x}B(u(x, t^n)) = \mathcal{O}(\Delta x^{k+1})$ for any $x \in I_j$. Without loss of generality, consider the case $\tilde{u}_j(x_{j+\frac{1}{2}}) < \epsilon$ at a fixed point $x = x_{j+\frac{1}{2}}$, then we have $u(x_{j+\frac{1}{2}}, t^n) < \mathcal{O}(\Delta x^{k+1}) + \epsilon$. By the discussion in the previous paragraph, it implies that $\frac{\partial}{\partial x}u(x_{j+\frac{1}{2}}, t^n) = \mathcal{O}(\Delta x^{k+1}) + \mathcal{O}(\epsilon)$ for small enough Δx and ϵ . Therefore $\frac{\partial}{\partial x}B(u(x_{j+\frac{1}{2}}, t^n)) = \frac{\partial}{\partial u}B(u(x_{j+\frac{1}{2}}, t^n))\frac{\partial}{\partial x}u(x_{j+\frac{1}{2}}, t^n) = \mathcal{O}(\Delta x^{k+1}) + \mathcal{O}(\epsilon)$. Thus $q_j(x_{j+\frac{1}{2}}) = \mathcal{O}(\Delta x^{k+1}) + \mathcal{O}(\epsilon)$ and setting $\tilde{q}_j(x_{j+\frac{1}{2}}) = 0$ does not affect the accuracy up to $\mathcal{O}(\Delta x^{k+1}) + \mathcal{O}(\epsilon)$.

In our numerical tests, we set $\epsilon = 10^{-10}$ in (20).

2.5. High order time discretizations and implementation

High order time discretizations can be achieved through Strong Stability Preserving (SSP) Runge-Kutta or multi-step methods [10], which are convex combinations of several formal forward Euler time steps. If forward Euler can preserve the positivity, then so does a high order SSP time discretization due to the convex combination. For instance, a third order SSP Runge-Kutta method for an equation of the form $\frac{d}{dt}U = L(U^n)$ is given by

$$\begin{aligned} U^{(1)} &= U^n + \Delta t L(U^n), \\ U^{(2)} &= \frac{3}{4}U^n + \frac{1}{4}\left(U^{(1)} + \Delta t L(U^{(1)})\right), \\ U^{n+1} &= \frac{1}{3}U^n + \frac{2}{3}\left(U^{(2)} + \Delta t L(U^{(2)})\right). \end{aligned}$$

Given $u_j(x)$ and $q_j(x)$ at each time step or stage, the limiters can be applied as follows:

- Apply the positivity-preserving limiter (19) to $u_j(x)$ to obtain $\tilde{u}_j(x)$.
- Apply the limiter (20) to $q_j(x)$ to obtain $\tilde{q}_j(x)$. Then compute the diffusion flux (13) in the DG method.

For choosing time steps in computation, the CFL conditions (15) and (18) should not be used directly for two reasons. First, the CFL conditions (15) and (18) are only sufficient rather than necessary conditions for ensuring $\bar{u}_j^{n+1} \geq 0$. Second, for explicit schemes solving diffusion equations, linear stability requires $\Delta t = \mathcal{O}(\Delta x^2)$. The CFL condition (15) and (18) are neither necessary nor sufficient to achieve $\Delta t = \mathcal{O}(\Delta x^2)$. The linear stability CFL constraint $\Delta t = \mathcal{O}(\Delta x^2)$ must also be satisfied in implementation.

In practice, at n -th time step in a SSP Runge-Kutta method, we can start the computation with a time step size implied by linear stability, e.g., $\Delta t = \mathcal{O}(\Delta x^2)$ for solving a diffusion problem (8). If any negative cell averages emerge in any time stage before reaching t^{n+1} , restart the computation at t^n with a halved

time step, which will not be an endless loop thanks to Theorem 1 and Theorem 7. See Section 5.3 in [1] for more details.

3. The positivity-preserving DG scheme in two dimensions

In this section we discuss the positivity-preserving high order DG schemes for the diffusion problems in two dimensions. As in the previous section, we focus on the LDG scheme with the assumption of periodic boundary conditions. Convection-diffusion problems can be similarly treated as in Section 2.2. For simplicity, we focus on rectangular cells for implementation details, even though the construction of positivity-preserving diffusion fluxes can be easily extended to unstructured meshes.

3.1. The local DG scheme

Consider a generic d -dimensional nonlinear diffusion problem:

$$u_t = \nabla \cdot (b(u)\nabla u) \quad (21)$$

where $b(u)$ is a semi-positive definite matrix. There exists a semi-positive definite matrix $b^*(u)$ such that

$$b_{ij}(u) = \sum_{1 \leq l \leq d} b_{il}^*(u)b_{lj}^*(u).$$

The following is an equivalent form of the equation (21):

$$\begin{aligned} u_t &= \sum_{i=1}^d \left(\sum_{l=1}^d b_{il}^*(u)q_l \right)_{x_i}, \\ q_l &= \sum_{j=1}^d (g_{lj}(u))_{x_j}, \quad l = 1, \dots, d, \\ g_{lj}(u) &= \int^u b_{lj}^*(\tau) d\tau. \end{aligned} \quad (22)$$

For convenience, introduce $\mathbf{q} = (q_1, \dots, q_d)^t$ and the flux functions $h_i(u, \mathbf{q}) = \sum_{l=1}^d b_{il}^*(u)q_l$, $\mathbf{h}(u, \mathbf{q}) = (h_1, \dots, h_d)^t$ and $\mathbf{g}_l = (g_{l1}, \dots, g_{ld})^t$.

Let K be a polygonal cell with edges e_j ($j = 1, \dots, E$) in a mesh \mathbb{T}_h approximating a two-dimensional computational domain. Let $\psi, \phi_l, l = 1, \dots, d$ be test functions, then the LDG scheme [18] can be written as

$$\begin{aligned} \int_K u_t \psi dA &= \int_{\partial K} \widehat{\mathbf{h}} \cdot \mathbf{n}_{\partial K} \psi ds - \int_K \mathbf{h}(u, \mathbf{q}) \cdot \nabla \psi dA, \\ \int_K q_l \phi_l dA &= \int_{\partial K} \widehat{\mathbf{g}_l} \cdot \mathbf{n}_{\partial K} \phi_l ds - \int_K \mathbf{g}_l(u) \cdot \nabla \phi_l dA, \quad l = 1, \dots, d, \end{aligned} \quad (23)$$

where $\mathbf{n}_{\partial K}$ denotes the unit outward normal vector to the boundary ∂K of the element K , $\widehat{\mathbf{h}} \cdot \mathbf{n}_{\partial K}$ and $\widehat{\mathbf{g}_l} \cdot \mathbf{n}_{\partial K}$ represent the diffusion fluxes.

3.2. A positivity-preserving flux for a generic equation on unstructured meshes

Let $|e_j|$ denote the length of the edge e_j and $|K|$ be the area of the rectangle K . Then the cell average \bar{u}_K time evolution equation can be obtained by setting $\psi = 1$:

$$\frac{d}{dt}\bar{u}_K = \frac{1}{|K|} \int_{\partial K} \widehat{\mathbf{h} \cdot \mathbf{n}_{\partial K}} ds = \frac{1}{|K|} \sum_{j=1}^E \int_{e_j} \widehat{\mathbf{h} \cdot \mathbf{n}_j} ds, \quad (24)$$

where $\mathbf{n}_j = (n_{j1}, n_{j2}, \dots, n_{jd})$ denotes the unit outward normal vector on the edge e_j . Let u^{in} and u^{ext} denote the approximations to u from interior and exterior of the cell K respectively. We consider a diffusion flux $\widehat{\mathbf{h} \cdot \mathbf{n}_j}$ in the following form:

$$\begin{aligned} \widehat{\mathbf{h} \cdot \mathbf{n}_j}(u^{in}, \mathbf{q}^{in}, u^{ext}, \mathbf{q}^{ext}) &= \frac{1}{2} [(\widehat{b^*} \mathbf{q}^{in}) \cdot \mathbf{n}_j + (\widehat{b^*} \mathbf{q}^{ext}) \cdot \mathbf{n}_j + \beta_j(u^{ext} - u^{in})] \\ &= \frac{1}{2} [\mathbf{q}^{in} \cdot (\widehat{b^*} \mathbf{n}_j) + \mathbf{q}^{ext} \cdot (\widehat{b^*} \mathbf{n}_j) + \beta_j(u^{ext} - u^{in})], \end{aligned} \quad (25a)$$

where we have used the fact that $(b\mathbf{q}) \cdot \mathbf{n} = \mathbf{n}^t b \mathbf{q} = (b\mathbf{n}) \cdot \mathbf{q}$ for a symmetric matrix b and two column vectors \mathbf{q} and \mathbf{n} . The matrix $\widehat{b^*}$ is given by

$$\widehat{b^*} = \frac{g(u^{ext}) - g(u^{in})}{u^{ext} - u^{in}}, \quad (25b)$$

and the penalty parameter β_j for each edge e_j is defined as,

$$\beta_{lj} = \begin{cases} \max \left\{ \left| \frac{\mathbf{q}^{ext} \cdot (\widehat{b^*} \mathbf{n}_j)}{u^{ext}} \right|, \left| \frac{\mathbf{q}^{in} \cdot (\widehat{b^*} \mathbf{n}_j)}{u^{in}} \right| \right\}, & \text{if } u^{ext} \neq 0 \text{ and } u^{in} \neq 0, \\ \left| \frac{\mathbf{q}^{ext} \cdot (\widehat{b^*} \mathbf{n}_j)}{u^{ext}} \right|, & \text{if } u^{ext} \neq 0 \text{ and } u^{in} = 0, \\ \left| \frac{\mathbf{q}^{in} \cdot (\widehat{b^*} \mathbf{n}_j)}{u^{in}} \right|, & \text{if } u^{ext} = 0 \text{ and } u^{in} \neq 0, \\ 0 & \text{if } u^{ext} = 0 \text{ and } u^{in} = 0. \end{cases} \quad (25c)$$

The line integral along each edge e_j can be approximated by N -point Gauss-Legendre quadrature (N should be at least $k+1$ for polynomials of degree k basis functions). Let \mathbf{x}_j^α ($\alpha = 1, \dots, N$) denote the Gauss-Legendre quadrature points on the edge e_j and ω_α denote the corresponding normalized quadrature weights such that $\sum_\alpha \omega_\alpha = 1$. With first order forward Euler time discretization, the cell average scheme (24) becomes

$$\bar{u}_K^{n+1} = \bar{u}_K^n + \frac{\Delta t}{|K|} \sum_{j=1}^E \sum_{\alpha=1}^N \widehat{\mathbf{h} \cdot \mathbf{n}_j}(\mathbf{x}_j^\alpha) \omega_\alpha |e_j|. \quad (26)$$

Let $u_K(\mathbf{x})$ denote the DG polynomial approximating $u(\mathbf{x}, t^n)$ on the cell K . For convenience, we will use subscripts α, j for u and \mathbf{q} to denote the evaluation at the point \mathbf{x}_j^α , e.g., $u_{\alpha,j}^{in} = u_K(\mathbf{x}_j^\alpha)$. Assume there exists a M -point quadrature on K in which the quadrature points include all \mathbf{x}_j^α and the

smallest quadrature weight is positive. For instance, such a quadrature can be easily constructed by tensor products of Gauss-Lobatto quadrature and Gauss-Legendre quadrature on rectangular cells, see [7]. For triangular cells and more general polygons, see [8, 1, 34, 35] for how to construct such a quadrature. Let \mathbf{x}_γ ($\gamma = EN + 1, \dots, M$) denote other quadrature points, then we have

$$\bar{u}_K^n = \frac{1}{|K|} \iint_K u_K(\mathbf{x}) dA = \sum_{j=1}^E \sum_{\alpha=1}^N \omega_{\alpha,j} u_{\alpha,j}^{in} + \sum_{\gamma=EN+1}^M \omega_\gamma u_K(\mathbf{x}_\gamma),$$

where $\omega_{\alpha,j} > 0$ and $\omega_\gamma > 0$ are the corresponding normalized quadrature weights satisfying $\sum_{j=1}^E \sum_{\alpha=1}^N \omega_{\alpha,j} + \sum_{\gamma=EN+1}^M \omega_\gamma = 1$. Thus we can rewrite (26) as

$$\begin{aligned} \bar{u}_K^{n+1} &= \sum_{\gamma=NE+1}^M \omega_\gamma u_K(\mathbf{x}_\gamma) + \sum_{j=1}^E \sum_{\alpha=1}^N \omega_{\alpha,j} u_{\alpha,j}^{in} + \frac{\Delta t}{|K|} \sum_{j=1}^E \sum_{\alpha=1}^N \widehat{\mathbf{h}} \cdot \mathbf{n}_j(\mathbf{x}_j^\alpha) \omega_\alpha |e_j|, \\ &= \sum_{\gamma=NE+1}^M \omega_\gamma u_K(\mathbf{x}_\gamma) + \sum_{j=1}^E \sum_{\alpha=1}^N \omega_{\alpha,j} u_{\alpha,j}^{in} \\ &\quad + \frac{\Delta t}{|K|} \sum_{j=1}^E \sum_{\alpha=1}^N \omega_\alpha |e_j| \frac{1}{2} [\mathbf{q}_{\alpha,j}^{in} \cdot (\widehat{\mathbf{b}}_{\alpha,j}^* \mathbf{n}_j) + \mathbf{q}_{\alpha,j}^{ext} \cdot (\widehat{\mathbf{b}}_{\alpha,j}^* \mathbf{n}_j) + \beta_j (u_{\alpha,j}^{ext} - u_{\alpha,j}^{in})], \\ &= \sum_{\gamma=NE+1}^M \omega_\gamma u_K(\mathbf{x}_\gamma) + \sum_{j=1}^E \sum_{\alpha=1}^N u_{\alpha,j}^{ext} \frac{1}{2} \frac{|e_j|}{|K|} \omega_\alpha \Delta t \left(\frac{\widehat{\mathbf{b}}_{\alpha,j}^* \cdot (\widehat{\mathbf{b}}_{\alpha,j}^* \mathbf{n}_j)}{u_{\alpha,j}^{ext}} + \beta_j \right) \\ &\quad + \sum_{j=1}^E \sum_{\alpha=1}^N u_{\alpha,j}^{in} \left[\omega_{\alpha,j} + \frac{1}{2} \frac{|e_j|}{|K|} \omega_\alpha \Delta t \left(\frac{\mathbf{q}_{\alpha,j}^{in} \cdot (\widehat{\mathbf{b}}_{\alpha,j}^* \mathbf{n}_j)}{u_{\alpha,j}^{in}} - \beta_j \right) \right], \end{aligned}$$

which is a positive linear combination of point values $u_K(\mathbf{x}_\lambda)$, $u_{\alpha,j}^{in}$ and $u_{\alpha,j}^{ext}$ under the CFL constraint

$$\Delta t \frac{|e_j|}{|K|} \beta_j \leq \min_{\alpha} \frac{\omega_{\alpha,j}}{\omega_\alpha}. \quad (27)$$

Thus we have obtained a weak positivity result in two dimensions:

Theorem 8. *For the high order scheme (26) using the flux (25), under the CFL constraint (27), the cell average $\bar{u}_K^{n+1} \geq 0$ if the following are satisfied:*

1. *All point values at the M -point quadrature points (\mathbf{x}_j^α and \mathbf{x}_λ) on ∂K are non-negative: $u_{\alpha,j}^{in} = u_K(\mathbf{x}_j^\alpha) \geq 0$ and $u_K(\mathbf{x}_\lambda) \geq 0$.*
2. *For all j and α , $\mathbf{q}_{\alpha,j}^{in} = 0$ if $u_{\alpha,j}^{in} = 0$, and $\mathbf{q}_{\alpha,j}^{ext} = 0$ if $u_{\alpha,j}^{ext} = 0$.*

Remark 9. *If the equation (21) reduces to a one-dimensional problem, then the flux (25) reduces to the flux in Remark 5.*

3.3. Special cases on rectangular meshes

To see an easier construction of positivity-preserving diffusion fluxes for more commonly used diffusion operators in applications such as the Laplacian, we simplify the discussion in this subsection by considering the equation (21) with a diagonal matrix $b(u)$, i.e.,

$$u_t = \partial_{xx}[b_{11}(u)] + \partial_{yy}[b_{22}(u)],$$

where $b_{11}(u), b_{22}(u)$ are functions of u satisfying $b_{11}(u) \geq 0, b_{22}(u) \geq 0$ for any $u \geq 0$. The following is an equivalent form for such an equation:

$$\begin{aligned} u_t &= (b_{11}^*(u)q_1)_x + (b_{22}^*(u)q_2)_y, \\ q_1 &= g_{11}(u)_x, \quad q_2 = g_{22}(u)_y, \\ g_{11} &= \int^u b_{11}^*(\tau)d\tau, \quad g_{22} = \int^u b_{22}^*(\tau)d\tau, \end{aligned}$$

where $b_{11}^* = \sqrt{b_{11}}$ and $b_{22}^* = \sqrt{b_{22}}$.

Consider a rectangular cell $K = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ with $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ and $\Delta y = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$. Let x_i^α and y_j^α denote the Gauss-Legendre quadrature points on the intervals $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and $[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ respectively, then the scheme (26) can be written as:

$$\begin{aligned} \bar{u}_K^{n+1} &= \bar{u}_K^n + \frac{\Delta t}{\Delta x} \sum_{\alpha=1}^N \omega_\alpha \left[\widehat{b_{11}q_1}(x_{i+\frac{1}{2}}, y_j^\alpha) - \widehat{b_{11}q_1}(x_{i-\frac{1}{2}}, y_j^\alpha) \right] \\ &\quad + \frac{\Delta t}{\Delta y} \sum_{\alpha=1}^N \omega_\alpha \left[\widehat{b_{22}q_2}(x_i^\alpha, y_{j+\frac{1}{2}}) - \widehat{b_{22}q_2}(x_i^\alpha, y_{j-\frac{1}{2}}) \right]. \end{aligned}$$

Here we can use (25) for the numerical fluxes $\widehat{b_{11}q_1}$ and $\widehat{b_{22}q_2}$. As an alternative, we can also consider the following flux which is a straightforward extension of (13). For simplicity, we only discuss $\widehat{b_{11}q_1}$, and $\widehat{b_{22}q_2}$ can be similarly defined:

$$\widehat{b_{11}q_1} = \frac{g_{11}(u^+) - g_{11}(u^-)}{u^+ - u^-} \hat{q}_1, \quad (28a)$$

where

$$\hat{q}_1 = \frac{1}{2} (q_1^+ + q_1^- + \beta(u^+ - u^-)), \quad (28b)$$

and

$$\beta = \begin{cases} \max \left\{ \left| \frac{q_1^+}{u^+} \right|, \left| \frac{q_1^-}{u^-} \right| \right\}, & \text{if } u^+ \neq 0 \text{ and } u^- \neq 0, \\ \left| \frac{q_1^+}{u^+} \right|, & \text{if } u^+ \neq 0 \text{ and } u^- = 0, \\ \left| \frac{q_1^-}{u^-} \right|, & \text{if } u^+ = 0 \text{ and } u^- \neq 0, \\ 0 & \text{if } u^+ = 0 \text{ and } u^- = 0, \end{cases} \quad (28c)$$

where u^+ and u^- denote the approximations from the right and from the left respectively.

3.4. The positivity-preserving limiter

Let $u_K(\mathbf{x})$ denote the DG polynomial at n -th time step for the cell K with the cell average $\bar{u}_K^n \geq 0$. To enforce the first condition in Theorems 8, we can apply the simple scaling limiter in [7] to $u_K(\mathbf{x})$ at the M -point quadrature points. The limiter can be defined as:

$$\tilde{u}_K(\mathbf{x}) = \bar{u}_K^n + \theta_K (u_K(\mathbf{x}) - \bar{u}_K^n), \theta_K = \min \left(1, \left| \frac{\bar{u}_K^n}{\bar{u}_K^n - m_K} \right| \right), \quad (29)$$

where m_K is the minimum value of $u_K(\mathbf{x})$ at M -point quadrature points. See [7, 8, 1] for more details.

3.5. Limiting the auxiliary variables \mathbf{q}

To enforce the second condition in Theorems 8, we need another limiter for modifying the auxiliary variables $\mathbf{q} = (q_1, q_2)$ so that they vanish wherever $u = 0$.

We only discuss how to modify q_1 and the discussion for q_2 is the same. Let $q_1(\mathbf{x})$ denote the DG polynomial for the variable q_1 approximating u_x on the cell K . For each cell K , let \mathbf{x}_j ($j = 1, \dots, H$) denote edge quadrature (for computing numerical fluxes along ∂K) points at which $\tilde{u}_K(\mathbf{x}_j) \leq \epsilon$ where the parameter $\epsilon > 0$ is a small number, e.g., $\epsilon = 10^{-10}$. We would like to find a modified polynomial $\tilde{q}_1(\mathbf{x})$ such that $\tilde{q}_1(\mathbf{x}_j) = 0$ with the smallest distance $\|q_1(\mathbf{x}) - \tilde{q}_1(\mathbf{x})\|$ for some norm $\|\cdot\|$, which is a projection problem.

Let L be the dimension of the approximation space. For instance, if using Q^k basis functions on a rectangular mesh, where Q^k refers to the space of tensor products of 1D polynomials of degree up to k , then $L = (k + 1)^2$. If using P^k basis on any polygonal cells, $L = k(k + 1)/2$.

Assume $p^i(\mathbf{x})$ ($i = 1, \dots, L$) form a basis for Q^k polynomials on the rectangular cell K . Then any Q^k polynomial $q(\mathbf{x})$ can be written as $q(\mathbf{x}) = \sum_{i=1}^L q^i p^i(\mathbf{x})$ where q^i are coordinates of $q(\mathbf{x})$ under the basis $\{p^i(\mathbf{x})\}$. Consider the following homogeneous linear system:

$$\begin{bmatrix} p^1(\mathbf{x}_1) & p^2(\mathbf{x}_1) & \dots & p^L(\mathbf{x}_1) \\ p^1(\mathbf{x}_2) & p^2(\mathbf{x}_2) & \dots & p^L(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \vdots \\ p^1(\mathbf{x}_H) & p^2(\mathbf{x}_H) & \dots & p^L(\mathbf{x}_H) \end{bmatrix} \begin{bmatrix} q^1 \\ q^2 \\ \vdots \\ q^L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (30)$$

Let A denote the $H \times L$ coefficient matrix in the linear system (30). Let B denote a matrix whose columns form a basis to all solutions of (30), i.e., the null space of A .

Let $q_1(\mathbf{x}) = \sum_{i=1}^L q_1^i p^i(\mathbf{x})$ and $\tilde{q}_1(\mathbf{x}) = \sum_{i=1}^L \tilde{q}_1^i p^i(\mathbf{x})$. Let $\mathbf{b} = [q_1^1 \quad q_1^2 \quad \dots \quad q_1^L]^T$ and $\tilde{\mathbf{b}} = [\tilde{q}_1^1 \quad \tilde{q}_1^2 \quad \dots \quad \tilde{q}_1^L]^T$. Then we can simply define $\tilde{\mathbf{b}}$ as the Euclidean projection of \mathbf{b} to the null space of A , thus $\tilde{\mathbf{b}}$ can be computed as $\tilde{\mathbf{b}} = B(B^T B)^{-1} B^T \mathbf{b}$.

4. Numerical tests

In all our numerical tests, we set the parameter in the limiter for the auxiliary variable q as $\epsilon = 10^{-10}$.

4.1. Accuracy tests for one-dimensional equations

We first consider a diffusion equation of the form

$$u_t = \varepsilon u_{xx} \tag{31}$$

on the the domain $[0, 2\pi]$ with initial data $u_0(x) = \sin(x) + 1$ and periodic boundary conditions. The analytical solution is $u(x, t) = \exp(-\varepsilon t) \sin(x) + 1$. We consider the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$ with the positivity preserving limiters (19) and (20). As a comparison, we also consider the local DG scheme with alternating fluxes, i.e., the scheme (10) with $\hat{u} = u^-$ and $\hat{q} = q^+$. In addition to enforcing the CFL condition for achieving positivity (15), we also impose the following time step constraints in Table 1 for linear stability in a third order SSP Runge-Kutta time discretization. The errors for $\varepsilon = 0.01$ at time $t = 1$ are listed in Table 2, in which we can observe that neither the new flux (13) nor limiters (19) and (20) affect the high order accuracy.

Table 1: The CFL constraints for linear stability for DG schemes with P^k basis solving $u_t = a(u)_{xx}$ where $\mu = \frac{\Delta t}{\Delta x^2} \max a'(u)$.

P^1	P^2	P^3	P^4	P^5
$\mu \leq 0.05$	$\mu \leq 0.01$	$\mu \leq 0.0005$	$\mu \leq 0.0002$	$\mu \leq 0.0001$

Next we consider a convection diffusion equation

$$u_t + u_x = \varepsilon u_{xx} \tag{32}$$

with the same initial data, boundary conditions and computational domain. The analytical solution is given by $u(x, t) = \exp(-\varepsilon t) \sin(x - t) + 1$. The global Lax-Friedrich's flux is used for the convection term and the time step constraint $\frac{\Delta t}{\Delta x} \leq \frac{1}{2k+1}$ is also enforced for the sake of linear stability due to the convection term. Errors for $\varepsilon = 0.01$ at time $t = 1$ are listed in Table 3.

4.2. Accuracy tests for Dirichlet and Neumann boundary conditions

We test the performance of the nonlinear flux (13) and limiters for Dirichlet and Neumann boundary conditions. We first consider special Dirichlet or Neumann boundary conditions so that the exact solution satisfies $u(x, t) = 0 \implies u_x(x, t) = 0$. Consider the Dirichlet boundary condition, i.e., solving the following initial boundary value problem on the interval $[0, 2\pi]$:

$$u_t = u_{xx}, \quad u(x, 0) = \sin x + 1, \quad u(0, t) = 1, \quad u(2\pi, t) = 1,$$

Table 2: Accuracy tests for DG schemes (10) with P^k basis solving (31) with $\varepsilon = 0.01$ at time $t = 1$ using N_x cells.

	$N_x =$	LDG with alternating fluxes				DG with (13) without any limiter				DG with (13) and limiters			
		L^1		L^∞		L^1		L^∞		L^1		L^∞	
		error	order	error	order	error	order	error	order	error	order	error	order
P^1	4	6.57e-02	-	1.08e-01	-	3.18e-02	-	6.16e-02	-	6.03e-02	-	1.33e-01	-
	8	2.02e-02	1.70	3.22e-02	1.75	8.29e-03	1.94	2.03e-02	1.60	1.48e-02	2.02	8.08e-02	0.72
	16	4.59e-03	2.14	7.23e-03	2.16	2.69e-03	1.63	1.12e-02	0.85	1.71e-03	3.11	2.78e-03	4.86
	32	1.17e-03	1.97	1.84e-03	1.98	9.73e-04	1.46	8.26e-03	0.44	5.67e-04	1.59	1.00e-03	1.47
	64	2.92e-04	2.00	4.59e-04	2.00	2.01e-04	2.27	3.72e-04	4.47	2.01e-04	1.49	3.74e-04	1.42
128	7.31e-05	2.00	1.15e-04	2.00	5.50e-05	1.87	1.11e-04	1.75	5.53e-05	1.86	1.12e-04	1.74	
P^2	4	3.10e-03	-	5.01e-03	-	2.42e-03	-	3.89e-03	-	6.68e-03	-	1.76e-02	-
	8	6.59e-04	2.23	1.35e-03	1.89	7.29e-04	1.73	1.25e-03	1.64	8.13e-04	3.04	1.43e-03	3.63
	16	1.23e-04	2.42	2.17e-04	2.64	1.30e-04	2.48	2.34e-04	2.41	1.31e-04	2.64	2.36e-04	2.59
	32	1.72e-05	2.84	3.11e-05	2.81	1.68e-05	2.95	3.08e-05	2.93	1.68e-05	2.96	3.08e-05	2.94
	64	2.15e-06	3.00	3.90e-06	2.99	2.12e-06	2.98	3.87e-06	2.99	2.12e-06	2.98	3.87e-06	2.99
128	2.69e-07	3.00	4.88e-07	3.00	2.67e-07	2.99	4.85e-07	2.99	2.67e-07	2.99	4.85e-07	2.99	
P^3	4	5.47e-04	-	1.25e-03	-	3.90e-04	-	7.74e-04	-	3.90e-04	-	7.74e-04	-
	8	4.63e-05	3.56	9.31e-05	3.74	5.05e-05	2.95	9.86e-05	2.97	5.05e-05	2.95	9.86e-05	2.97
	16	3.15e-06	3.88	5.46e-06	4.09	3.68e-06	3.78	6.44e-06	3.94	3.68e-06	3.78	6.44e-06	3.94
	32	2.00e-07	3.98	3.59e-07	3.92	2.32e-07	3.99	4.87e-07	3.72	2.32e-07	3.99	4.87e-07	3.72
	64	1.25e-08	4.00	2.25e-08	4.00	1.79e-08	3.70	3.62e-08	3.75	1.79e-08	3.70	3.62e-08	3.75
128	7.80e-10	4.00	1.41e-09	4.00	1.17e-09	3.93	2.32e-09	3.96	1.17e-09	3.93	2.32e-09	3.96	
P^4	4	4.64e-05	-	8.50e-05	-	5.78e-05	-	8.70e-05	-	5.78e-05	-	8.70e-05	-
	8	1.88e-06	4.63	3.14e-06	4.76	1.97e-06	4.88	3.33e-06	4.71	1.96e-06	4.88	3.32e-06	4.71
	16	6.11e-08	4.94	1.09e-07	4.84	5.94e-08	5.04	1.11e-07	4.90	5.94e-08	5.04	1.11e-07	4.90
	32	1.89e-09	5.01	3.57e-09	4.94	1.87e-09	4.99	3.56e-09	4.96	1.87e-09	4.99	3.58e-09	4.96
	64	5.92e-11	5.00	1.12e-10	4.99	5.87e-11	4.99	1.12e-10	5.00	5.87e-11	4.99	1.12e-10	5.00
128	2.17e-12	4.77	5.01e-12	4.49	2.16e-12	4.76	5.00e-12	4.48	2.16e-12	4.76	5.00e-12	4.48	
P^5	4	3.39e-06	-	6.34e-06	-	3.79e-06	-	5.75e-06	-	8.78e-06	-	2.21e-05	-
	8	6.09e-08	5.80	1.02e-07	5.95	5.97e-08	5.98	1.07e-07	5.74	1.03e-07	6.42	3.81e-07	5.86
	16	9.72e-10	5.97	1.79e-09	5.84	9.9e-10	5.92	2.12e-09	5.66	1.16e-09	6.47	4.95e-09	6.26
	32	1.51e-11	6.01	2.87e-11	5.96	2.05e-11	5.59	4.73e-11	5.49	2.06e-11	5.81	4.73e-11	6.71
	64	6.80e-13	4.48	1.60e-12	4.16	6.45e-13	4.99	1.99e-12	4.57	6.46e-13	5.00	1.99e-12	4.57
128	2.39e-12	-1.81	4.71e-12	-1.56	2.39e-12	-1.89	4.71e-12	-1.25	2.39e-12	-1.86	4.72e-12	-1.24	

for which the exact solution is $u(x, t) = e^{-t} \sin x + 1$. We use the following straightforward boundary treatment in the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$. We set u^\pm as the Dirichlet boundary conditions at two boundary points and use extrapolation for q^\pm at two boundary points:

$$u_{\frac{1}{2}}^- = 1, \quad u_{N_x + \frac{1}{2}}^+ = 1, \quad q_{\frac{1}{2}}^- = q_{\frac{1}{2}}^+, \quad q_{N_x + \frac{1}{2}}^+ = q_{N_x + \frac{1}{2}}^-.$$

See Table 4 for the impact of such a simple boundary treatment for the DG schemes. We can observe around $(k + 1)$ -order accuracy for even k and at least k -th order accuracy for odd k for the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$. Notice that the difference for DG schemes with or without limiters does not show in three significant figures of the errors for this particular problem.

Next we consider the Neumann boundary condition, i.e., solving the following initial boundary value problem on the interval $[0, 2\pi]$:

$$u_t = u_{xx}, \quad u(x, 0) = \sin x + 1, \quad u_x(0, t) = e^{-t}, \quad u_x(2\pi, t) = e^{-t},$$

for which the exact solution is $u(x, t) = e^{-t} \sin x + 1$. We use the following straightforward boundary treatment in the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$. We set q^\pm as the Neumann boundary conditions at two boundary

Table 3: Accuracy tests for DG schemes (10) with P^k basis solving 1D convection diffusion equation (32) with $\varepsilon = 0.01$ at time $t = 1$ using Nx cells.

	$Nx =$	LDG with alternating fluxes				DG with (13) without any limiter				DG with (13) and limiters			
		L^1		L^∞		L^1		L^∞		L^1		L^∞	
		error	order	error	order	error	order	error	order	error	order	error	order
P^1	4	5.51e-02	-	1.53e-01	-	1.02e-01	-	2.16e-01	-	5.97e-02	-	2.04e-01	-
	8	1.76e-02	1.65	3.28e-02	2.22	1.69e-02	2.59	3.23e-02	2.74	1.95e-02	1.62	5.36e-02	1.93
	16	4.61e-03	1.93	8.06e-03	2.03	4.23e-03	2.00	8.02e-03	2.01	4.45e-03	2.13	1.24e-02	2.11
	32	1.16e-03	1.99	1.94e-03	2.05	1.14e-03	1.89	2.14e-03	1.90	1.14e-03	1.97	2.30e-03	2.43
	64	2.92e-04	1.99	4.72e-04	2.04	2.86e-04	1.99	4.725e-04	2.18	2.86e-04	1.99	4.96e-04	2.21
128	7.31e-05	2.00	1.16e-04	2.02	1.24e-04	1.20	9.66e-04	-1.03	7.15e-05	2.00	1.17e-04	2.08	
P^2	4	8.94e-03	-	1.74e-02	-	8.92e-03	-	1.74e-02	-	1.17e-02	-	3.18e-02	-
	8	1.14e-03	2.97	1.84e-03	3.24	1.15e-03	2.95	1.87e-03	3.21	1.19e-03	3.30	2.26e-03	3.82
	16	1.38e-04	3.05	2.45e-04	2.91	1.39e-04	3.05	2.48e-04	2.91	1.40e-04	3.09	2.48e-04	3.18
	32	1.72e-05	3.00	3.10e-05	2.98	1.74e-05	3.00	3.13e-05	2.99	1.74e-05	3.01	3.13e-05	2.99
	64	2.15e-06	3.00	3.90e-06	2.99	2.16e-06	3.01	3.91e-06	3.00	2.16e-06	3.01	3.91e-06	3.00
128	2.69e-07	3.00	4.88e-07	3.00	2.69e-07	3.01	4.87e-07	3.00	2.69e-07	3.01	4.87e-07	3.00	
P^3	4	7.45e-04	-	1.66e-03	-	7.44e-04	-	1.68e-03	-	1.85e-03	-	4.48e-03	-
	8	5.17e-05	3.85	1.02e-04	4.02	5.05e-05	3.88	1.02e-04	4.05	5.05e-05	5.20	1.02e-04	5.46
	16	3.19e-06	4.02	5.86e-06	4.12	3.12e-06	4.02	5.85e-06	4.12	3.12e-06	4.02	5.85e-06	4.12
	32	2.00e-07	4.00	3.60e-07	4.03	1.97e-07	3.99	3.60e-07	4.02	1.97e-07	3.99	3.60e-07	4.02
	64	1.25e-08	4.00	2.25e-08	4.00	1.23e-08	4.00	2.26e-08	3.99	1.23e-08	4.00	2.26e-08	3.99
128	7.80e-10	4.00	1.41e-09	4.00	7.65e-10	4.00	1.42e-09	3.99	7.65e-10	4.00	1.42e-09	3.99	
P^4	4	5.67e-05	-	1.35e-04	-	5.71e-05	-	1.37e-04	-	7.31e-05	-	1.69e-04	-
	8	1.93e-06	4.87	3.68e-06	5.20	1.95e-06	4.87	3.72e-06	5.21	1.95e-06	5.23	3.72e-06	5.51
	16	5.97e-08	5.02	1.16e-07	4.98	6.01e-08	5.02	1.17e-07	4.99	6.01e-08	5.02	1.17e-07	4.99
	32	1.89e-09	4.98	3.57e-09	5.03	1.89e-09	4.99	3.58e-09	5.03	1.89e-09	4.99	3.58e-09	5.03
	64	5.92e-11	5.00	1.12e-10	4.99	5.90e-11	5.00	1.12e-10	4.96	5.90e-11	5.00	1.12e-10	5.00
128	2.19e-12	4.76	5.09e-12	4.46	2.18e-12	4.75	5.09e-12	4.46	2.18e-12	4.75	5.09e-12	4.46	
P^5	4	3.61e-06	-	8.59e-06	-	3.60e-06	-	8.53e-06	-	1.40e-05	-	3.96e-05	-
	8	6.31e-08	5.84	1.27e-07	6.08	6.19e-08	5.86	1.27e-07	6.07	9.08e-08	7.27	3.85e-07	6.69
	16	9.60e-10	6.04	1.99e-09	6.00	9.42e-10	6.04	1.97e-09	6.00	9.93e-10	6.52	5.27e-09	6.19
	32	1.51e-11	5.99	3.10e-11	6.00	1.49e-11	5.98	3.10e-11	6.00	1.49e-11	6.06	4.29e-11	6.94
	64	6.59e-13	4.52	1.65e-12	4.24	6.55e-13	4.51	1.64e-12	4.24	6.53e-13	4.51	1.63e-12	4.72
128	2.34e-12	-1.83	4.75e-12	-1.53	2.34e-12	-1.84	4.75e-12	-1.53	2.34e-12	-1.84	4.76e-12	-1.54	

points and use extrapolation for u^\pm at two boundary points:

$$q_{\frac{1}{2}}^- = e^{-t}, \quad q_{N_x + \frac{1}{2}}^+ = -e^{-t}, \quad u_{\frac{1}{2}}^- = u_{\frac{1}{2}}^+, \quad u_{N_x + \frac{1}{2}}^+ = u_{N_x + \frac{1}{2}}^-.$$

See Table 5 for the impact of such a simple boundary treatment for the DG schemes. Notice that the difference for DG schemes with or without limiters does not show in three significant figures of the errors for this particular problem. For the P^2 basis, the order of accuracy reduces to the second order for the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$. Nonetheless we can observe at least k -th order accuracy for the DG scheme (10) using fluxes (13) and $\hat{u} = u^-$. In other words, order of accuracy of the nonlinear flux (10) may degenerate by one order lower for even k for this special Neumann boundary conditions.

Then we consider solving the heat equation with boundary conditions so that $u(x, t) = 0 \implies u_x(x, t) = 0$ is no longer true. We first consider the Dirichlet boundary condition, i.e., solving the following initial boundary value problem on the interval $[0, \pi]$:

$$u_t = u_{xx}, \quad u(x, 0) = \sin x, \quad u(0, t) = 0, \quad u(\pi, t) = 0,$$

for which the exact solution is $u(x, t) = e^{-t} \sin x$. See Table 6 for the errors. As expected, the high order accuracy is completely destroyed by limiters. We can

Table 4: Accuracy tests for DG schemes (10) with P^k basis solving 1D heat equation $u_t = u_{xx}$ with a special Dirichlet boundary condition at time $t = 1$ using N_x cells.

	$N_x =$	LDG with alternating fluxes				DG with (13) without any limiter				DG with (13) and limiters			
		L^1		L^∞		L^1		L^∞		L^1		L^∞	
		error	order	error	order	error	order	error	order	error	order	error	order
P^2	8	1.97e-03	-	1.37e-02	-	1.00e-03	-	4.36e-03	-	1.00e-03	-	4.36e-03	-
	16	1.74e-04	3.51	1.76e-03	2.96	1.01e-04	3.31	3.76e-04	3.53	1.01e-04	3.31	3.76e-04	3.53
	32	1.70e-05	3.35	2.22e-04	2.99	1.24e-05	3.03	4.63e-05	3.02	1.24e-05	3.03	4.63e-05	3.02
	64	1.83e-06	3.21	2.78e-05	3.00	1.54e-06	3.01	5.80e-06	3.00	1.54e-06	3.01	5.80e-06	3.00
P^3	8	2.48e-04	-	2.38e-03	-	4.53e-05	-	1.85e-04	-	4.53e-05	-	1.85e-04	-
	16	1.65e-05	3.91	3.13e-04	2.92	2.70e-06	4.07	1.64e-05	3.50	2.70e-06	4.07	1.64e-05	3.50
	32	1.04e-06	3.98	3.97e-05	2.98	1.66e-07	4.02	1.80e-06	3.18	1.66e-07	4.02	1.80e-06	3.18
	64	6.56e-08	3.99	4.98e-06	3.00	1.03e-08	4.01	2.15e-07	3.07	1.03e-08	4.01	2.15e-07	3.07
P^4	8	5.66e-06	-	5.78e-05	-	1.81e-06	-	1.01e-05	-	1.81e-06	-	1.01e-05	-
	16	1.13e-07	5.65	1.85e-06	4.96	4.70e-08	5.26	2.24e-07	5.49	4.70e-08	5.26	2.24e-07	5.49
	32	2.50e-09	5.50	5.83e-08	4.99	1.46e-09	5.01	7.07e-09	4.99	1.46e-09	5.01	7.07e-09	4.99
	64	6.20e-11	5.33	1.83e-09	5.00	4.57e-11	4.99	2.22e-10	4.99	4.57e-11	4.99	2.22e-10	4.99
P^5	8	4.68e-07	-	5.79e-06	-	5.89e-08	-	2.78e-07	-	5.89e-08	-	2.78e-07	-
	16	7.78e-09	5.91	1.92e-07	4.92	8.88e-10	6.05	6.14e-09	5.50	8.88e-10	6.05	6.14e-09	5.50
	32	1.24e-10	5.98	6.07e-09	4.98	1.39e-11	6.00	1.74e-10	5.14	1.39e-11	6.00	1.74e-10	5.14
	64	2.43e-12	5.67	1.90e-10	4.99	7.75e-13	4.16	5.31e-12	5.03	7.75e-13	4.16	5.31e-12	5.03

Table 5: Accuracy tests for DG schemes (10) with P^k basis solving 1D heat equation $u_t = u_{xx}$ with a special Neumann boundary conditions at time $t = 1$ using N_x cells.

	$N_x =$	LDG with alternating fluxes				DG with (13) without any limiter				DG with (13) and limiters			
		L^1		L^∞		L^1		L^∞		L^1		L^∞	
		error	order	error	order	error	order	error	order	error	order	error	order
P^2	8	6.78e-04	-	2.83e-03	-	7.19e-03	-	2.87e-02	-	7.19e-03	-	2.87e-02	-
	16	9.04e-05	2.91	3.67e-04	2.95	1.91e-03	1.91	7.65e-03	1.91	1.91e-03	1.91	7.65e-03	1.91
	32	1.78e-05	2.94	4.65e-05	2.99	4.81e-04	1.99	1.94e-03	1.98	4.81e-04	1.99	1.94e-03	1.98
	64	1.50e-06	2.97	5.81e-06	3.00	1.20e-04	2.00	4.87e-04	2.00	1.20e-04	2.00	4.87e-04	2.00
P^3	8	3.68e-05	-	1.57e-04	-	2.00e-04	-	7.99e-04	-	2.00e-04	-	7.99e-04	-
	16	2.38e-06	3.95	1.03e-05	3.93	1.08e-05	4.21	5.12e-05	3.97	1.08e-05	4.21	5.12e-05	3.97
	32	1.50e-07	3.99	6.49e-07	3.98	1.10e-06	3.30	3.22e-06	3.99	1.10e-06	3.30	3.22e-06	3.99
	64	9.36e-09	4.00	4.06e-08	4.00	1.24e-07	3.15	3.52e-07	3.19	1.24e-07	3.15	3.52e-07	3.19
P^4	8	1.26e-06	-	6.79e-06	-	3.60e-05	-	1.28e-04	-	3.60e-05	-	1.28e-04	-
	16	4.25e-08	4.89	2.23e-07	4.93	2.34e-06	3.95	8.44e-06	3.92	2.34e-06	3.95	8.44e-06	3.92
	32	1.39e-09	4.93	7.07e-09	4.98	1.48e-07	3.99	5.35e-07	3.98	1.48e-07	3.99	5.35e-07	3.98
	64	4.48e-11	4.96	2.22e-10	4.99	9.24e-09	4.00	3.35e-08	4.00	9.24e-09	4.00	3.35e-08	4.00
P^5	8	4.64e-08	-	2.42e-07	-	5.57e-07	-	2.08e-06	-	5.57e-07	-	2.08e-06	-
	16	7.51e-10	5.95	3.99e-09	5.92	7.15e-09	6.28	3.32e-08	5.97	7.15e-09	6.28	3.32e-08	5.97
	32	1.19e-11	5.98	6.34e-11	5.98	1.20e-10	5.89	5.22e-10	5.99	1.20e-10	5.89	5.22e-10	5.99
	64	8.05e-13	3.89	1.81e-12	5.13	2.73e-12	5.46	7.16e-12	6.19	2.73e-12	5.46	7.16e-12	6.19

also test the limiters on the same solution with Neumann boundary conditions, i.e., the following initial boundary value problem on the interval $[0, \pi]$:

$$u_t = u_{xx}, \quad u(x, 0) = \sin x, \quad u_x(0, t) = e^{-t}, \quad u_x(\pi, t) = -e^{-t},$$

for which the exact solution is $u(x, t) = e^{-t} \sin x$. The results of the accuracy test are similar thus not included here, i.e., the DG scheme with (13) and limiters is not high order accurate for this problem either.

4.3. Accuracy tests for a two-dimensional equation

Consider the two dimensional convection diffusion equation:

$$u_t + u_x + u_y = \varepsilon(u_{xx} + u_{yy}) \quad (33)$$

Table 6: Accuracy tests for DG schemes (10) with P^k basis solving 1D heat equation $u_t = u_{xx}$ with a generic Dirichlet boundary condition at time $t = 1$ using N_x cells.

	$N_x =$	LDG with alternating fluxes				DG with (13) and limiters			
		L^1		L^∞		L^1		L^∞	
		error	order	error	order	error	order	error	order
P^2	8	2.47e-04	-	1.76e-03	-	5.13e-04	-	2.00e-03	-
	16	2.16e-05	3.51	2.22e-04	2.99	1.29e-04	1.99	2.47e-04	3.02
	32	2.12e-06	3.35	2.78e-05	3.00	2.07e-05	2.64	3.56e-05	2.79
	64	2.29e-07	3.21	3.48e-06	3.00	4.88e-05	-1.23	9.52e-05	-1.42
P^3	8	3.05e-05	-	3.13e-04	-	5.46e-04	-	2.26e-03	-
	16	1.94e-06	3.98	3.97e-05	2.98	3.25e-04	0.75	6.38e-04	1.82
	32	1.22e-07	3.99	4.98e-06	3.00	6.25e-05	2.38	1.19e-04	2.42
	64	7.63e-09	4.00	6.23e-07	3.00	8.21e-05	-0.39	2.86e-04	-1.26
P^4	8	1.79e-07	-	1.85e-06	-	7.92e-05	-	1.43e-04	-
	16	3.53e-09	5.66	5.83e-08	4.99	3.78e-04	-2.26	6.09e-04	-2.09
	32	7.81e-11	5.50	1.83e-09	5.00	9.23e-08	11.99	7.12e-07	9.74
	64	2.31e-12	5.08	5.71e-11	5.00	1.34e-5	-7.19	3.84e-5	-5.75

on the domain $[-1, 1] \times [-1, 1]$ with initial conditions $u(x, y) = \sin(2\pi(x+y)) + 1$ and periodic boundary conditions for $\varepsilon = 0.001$. The analytical solution is given by $u(x, y, t) = \exp(-8\pi^2\varepsilon t) \sin(2\pi(x+y-2t)) + 1$. Errors at $t = 0.5$ of local DG scheme with alternating fluxes [18] and DG scheme with positivity-preserving flux and limiters are listed in Table 7.

Table 7: Accuracy tests for DG schemes (10) with Q^k basis solving the 2D convection diffusion equation (33) with $\varepsilon = 0.001$ at time $t = 0.5$.

	Grid size	LDG with alternating fluxes						DG with positivity-preserving flux and limiters					
		L^1		L^2		L^∞		L^1		L^2		L^∞	
		error	order	error	order	error	order	error	order	error	order	error	order
Q^1	6×6	6.25e-02	-	6.25e-02	-	1.81e-01	-	6.52e-01	-	7.32e-01	-	1.27	-
	11×11	1.11e-02	2.49	1.55e-02	2.36	3.01e-02	2.59	2.36e-02	4.79	3.03e-02	4.60	7.21e-02	4.14
	21×21	3.33e-03	1.74	4.33e-03	1.84	1.03e-02	1.55	3.71e-03	2.67	4.73e-03	2.68	1.23e-02	2.55
	41×41	9.45e-04	1.82	1.21e-03	1.84	2.96e-03	1.80	9.39e-04	1.98	1.21e-03	1.97	2.96e-03	2.06
Q^2	6×6	7.06e-03	-	7.06e-03	-	2.34e-02	-	9.51e-02	-	1.08e-01	-	1.74e-01	-
	11×11	1.02e-03	2.79	1.30e-03	2.81	3.20e-03	2.87	1.05e-03	6.50	1.33e-03	6.34	3.25e-03	5.75
	21×21	1.30e-04	2.97	1.66e-04	2.97	4.11e-04	2.96	1.31e-04	3.00	1.68e-04	2.99	4.41e-04	2.97
	41×41	1.63e-05	3.00	2.08e-05	2.99	5.18e-05	2.99	1.63e-05	3.00	2.09e-05	3.00	5.20e-05	2.99
Q^3	6×6	8.89e-04	-	8.89e-04	-	3.51e-03	-	8.90e-04	-	1.12e-03	-	3.50e-03	-
	11×11	5.57e-05	4.00	6.82e-05	4.12	1.83e-04	4.26	5.57e-05	4.00	6.81e-05	4.12	1.83e-04	4.26
	21×21	3.48e-06	4.00	4.33e-06	3.98	1.20e-05	3.93	3.48e-06	4.00	4.32e-06	3.98	1.20e-05	3.93
	41×41	2.18e-07	4.00	2.71e-07	4.00	7.52e-07	4.00	2.18e-07	4.00	2.71e-07	4.00	7.52e-07	3.99
Q^4	6×6	1.10e-04	-	1.10e-04	-	2.54e-04	-	1.11e-04	-	1.30e-04	-	2.54e-04	-
	11×11	3.11e-06	5.15	3.68e-06	5.14	8.58e-06	4.89	3.11e-06	5.15	3.68e-06	5.14	8.60e-06	4.89
	21×21	9.70e-08	5.00	1.14e-07	5.02	2.60e-07	5.04	9.70e-08	5.00	1.14e-07	5.02	2.61e-07	5.04
	41×41	3.03e-09	5.00	3.55e-09	5.00	8.19e-09	4.99	3.03e-09	5.00	3.55e-09	5.00	8.20e-09	4.99

4.4. One-dimensional porous medium equations

A representative test for validating the positivity-preserving property of a scheme solving nonlinear diffusion equations is the porous medium equation,

$$u_t = (u^m)_{xx} \quad , \quad m > 1. \quad (34)$$

We consider the Barenblatt analytical solution given by

$$B_m(x, t) = t^{-k} \left[\left(1 - \frac{k(m-1)}{2m} \frac{|x|^2}{t^{2k}} \right)_+ \right]^{\frac{1}{m-1}}, \quad (35)$$

where $f_+ = \max\{f, 0\}$. See also [8, 36, 13] for such a solution. To test the numerical scheme, we use the initial conditions taken as $B_m(x, 1)$ and the solution computed till time $t = 2$ on the domain $[-6, 6]$. DG schemes with the positivity-preserving flux and limiters using P^1 to P^5 basis are tested for $m \in [2, 10]$. Figure 1 consists of results for various values of m using different polynomial approximations. DG schemes without any positivity treatment will generate negative solutions for this test [36]. We use centered fluxes for $\widehat{B}(u)$, although the positivity property does not depend on this choice. High order accurate schemes without any positivity treatment can easily produce negative solutions, which may further induce instabilities.

4.5. Two-dimensional porous medium equations

We consider the two-dimensional porous medium equation given by

$$u_t = \Delta u^m$$

on the computational domain $[-2, 2] \times [-2, 2]$ for various values of m from time $t = 0$ till time $t = 0.01$. The initialization used is

$$u = \begin{cases} 1, & (x, y) \in [-0.5, 0.5] \times [-0.5, 0.5], \\ 0, & \text{otherwise.} \end{cases}$$

A centered flux for the diffusion flux $\widehat{\mathbf{g}_t \cdot \mathbf{n}}$ is used. Solutions for $m = 4$ by positivity-preserving DG schemes are plotted in Figure 2. Even though oscillations are observed in the numerical solutions, the positivity of solutions is achieved.

5. Applications to modeling electrical discharges

In this section, we present an application of the positivity preserving scheme to modeling electrical discharges, which involves solving for the species number densities. A positivity ensuring scheme is essential since a physical requirement is for the number densities to be non-negative. Discharges can be classified into various categories and our focus is on modeling the direct current glow discharge. A popular model for solving for a glow discharge is the Drift-Diffusion model

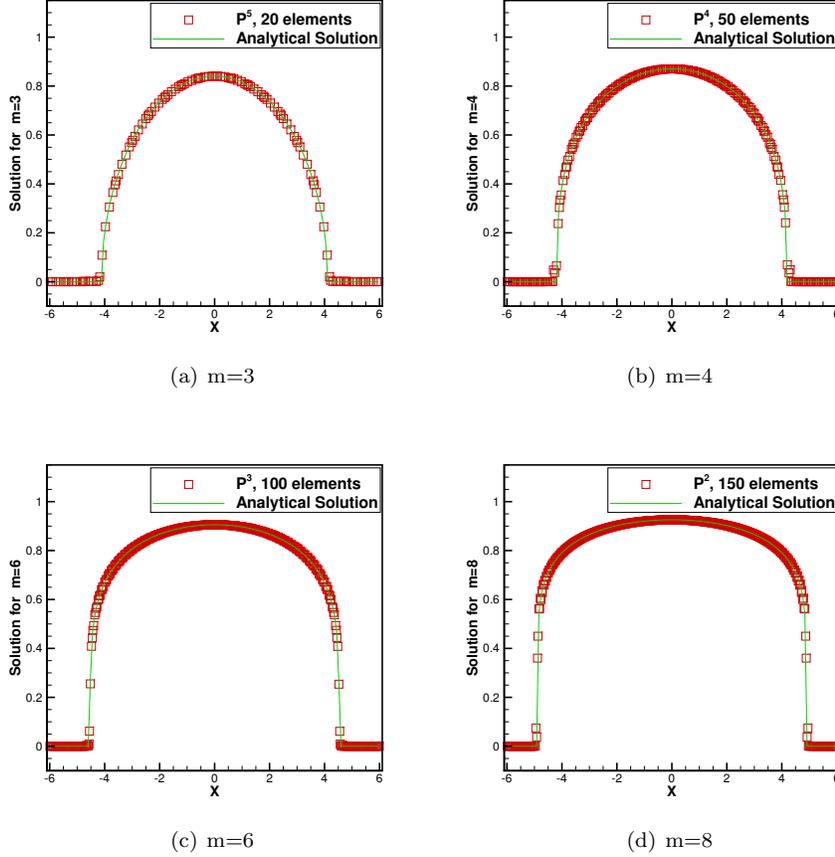


Figure 1: Porous Medium Equation solutions obtained using various polynomial approximations. The legend shows the polynomial degree and the number of elements used for discretizing the domain.

[37]. It is a traditional continuum based model that is derived from the first moment of the Boltzmann equation. It has been utilized in various works in the literature to model glow and pulsed discharges in different gases [38, 39, 40]. It is most effective in modeling glow discharges at pressures 1-50 Torr and voltage differences in the range 0.3 to 10 kV [40].

The model consists of equations that describe the spatial-temporal evolution of ion and electron number densities, solved in a coupled manner with a Poisson equation that describes the electric potential distribution in a self-consistent manner.

$$\frac{\partial n_i}{\partial t} + \nabla \cdot \mathbf{\Gamma}_i = \alpha |\mathbf{\Gamma}_e| - \beta n_i n_e, \quad (36a)$$

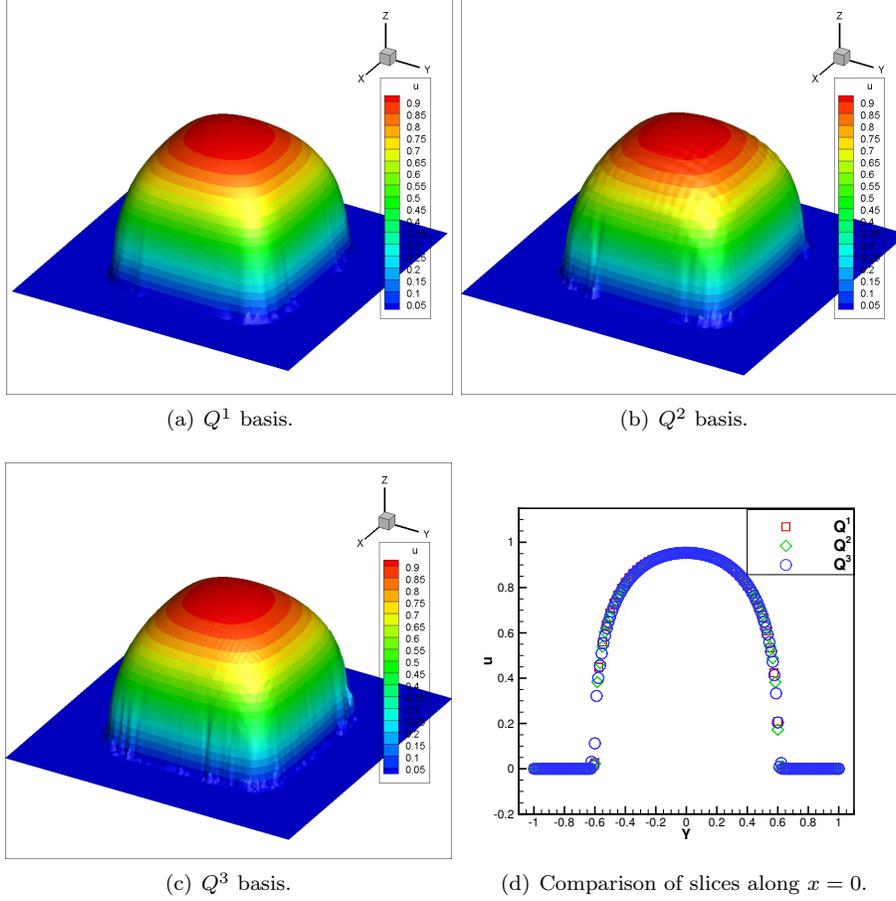


Figure 2: Positivity-preserving DG schemes with Q^k basis for solving the two-dimensional porous medium equation with $m = 4$ on a 60×60 grid.

$$\frac{\partial n_e}{\partial t} + \nabla \cdot \mathbf{\Gamma}_e = \alpha |\mathbf{\Gamma}_e| - \beta n_i n_e, \quad (36b)$$

where n_i and n_e represent the ion and electron number densities, $\mathbf{\Gamma}_i = n_i \mu_i \mathbf{E} - D_i \nabla n_i$ is the ion flux, $\mathbf{\Gamma}_e = -n_e \mu_e \mathbf{E} - D_e \nabla n_e$ is the electron flux, $\mathbf{E} = -\nabla \phi$ is the electric field, μ and D are the mobility and diffusion coefficients, e is the electric charge of the electron, ϵ_0 is the permittivity of space and ϕ is the electric potential. The spatial distribution of the electric potential ϕ is governed by Gauss' law

$$\nabla^2 \phi = -\frac{e}{\epsilon_0} (n_i - n_e) \quad (36c)$$

The two source terms represent the ionization and recombination processes,

which lead to gain and loss of species respectively. Ionization is assumed to be governed by Townsend's model, which characterizes ionization using the Townsend ionization coefficient

$$\alpha = Ape^{-\frac{B}{|\mathbf{E}|^p}} \quad (36d)$$

where A, B are coefficients dependent on the gas and p is the gas pressure. The recombination process is described using the rate coefficient β . The physics-based time step restrictions are governed by a combination of the dielectric relaxation time t_d and the Courant time step limit t_c

$$\begin{aligned} t_d &= \frac{e}{\epsilon_0(\mu_e n_e + \mu_i n_i)} \\ t_c &= \frac{\Delta x^2}{2D_e + \Delta x \mu_e E} \end{aligned} \quad (37)$$

The model in its original form (36) is more complicated than convection diffusion equations, since the right hand sides in (36a) and (36b) are Hamilton-Jacobi equations due to the fact that $\mathbf{\Gamma}_e$ contains the gradient of densities. However, a common approximation that is made to the ionization term, is to neglect the diffusion contribution to the electron flux. If this approximation is made, the set of equations can be regarded as convection-diffusion in nature, and both the ionization and recombination terms can be treated purely as source terms. This approximation has been employed in literature, for example in [40].

Though the model can be used to predict various types of discharges, our primary interest is in modeling a glow discharge. To this end, we present two physical processes: a non-ionizing transient sheath problem and a steady-state glow discharge.

5.1. High-density transient sheath

The transient sheath problem is a good prelude to glow discharge calculations, in that the plasma structure closely resembles the glow discharge structure during the initial phase. The test case we present here is that of a non-ionizing plasma in high-pressure argon gas. This particular case was originally analyzed by Poggie and Gaitonde [41] using a simplified physical model, and subsequently by Hilbun and Case [42] with the same model as employed here.

The source terms are omitted in the model used here, which is given by

$$\frac{\partial n_i}{\partial t} + \nabla \cdot \mathbf{\Gamma}_i = 0, \quad (38a)$$

$$\frac{\partial n_e}{\partial t} + \nabla \cdot \mathbf{\Gamma}_e = 0, \quad (38b)$$

solved in a coupled fashion with Gauss's law

$$\nabla^2 \phi = -\frac{e}{\epsilon_0}(n_i - n_e). \quad (38c)$$

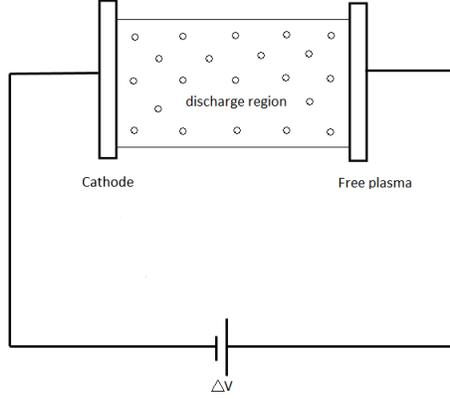


Figure 3: Sheath test case schematic representation.

A schematic of the computational domain can be regarded as given in Figure 3. The computational domain, which is the region between the cathode and the anode, is of length $200 \lambda_{De}$, where $\lambda_{De} = \sqrt{\frac{\epsilon_0 k_b T_e}{n_0 e^2}}$ is the Debye length. A constant potential difference of $\Delta V = \frac{50 k_b T_e}{e}$ is applied across the discharge gap and the potential drop does not change with time due to the absence of an external circuit. The boundary conditions imposed are tabulated in Table 8.

Table 8: Boundary conditions used for sheath calculations solved by Drift-Diffusion model.

quantity	cathode (x=0)	free plasma boundary (x=h)
n_i	$\frac{\partial n_i}{\partial x} = 0$	$n_i = n_0$
n_e	$n_e = \gamma \frac{\mu_i}{\mu_e} n_i$	$\frac{\partial n_e}{\partial x} = 0$
ϕ	$\phi = -50 k_b T_e / e$	$\phi = 0$

The secondary emission coefficient γ was taken to be 0.3 while the initial number densities were set as $n_0 = 10^{15} m^{-3}$. The ion and electron temperatures were taken to be 0.1 eV and 1 eV respectively. The transport coefficients for Argon at 100 Torr pressure were taken to be as given in Table 9.

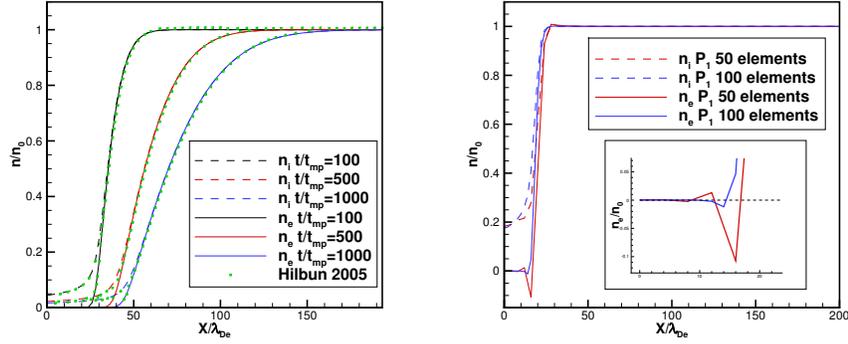
Table 9: Transport coefficients for argon at $p = 100$ Torr.

μ_i	μ_e	D_i	D_e
$10^{-3} m^2 / sV$	$0.3 m^2 / sV$	$10^{-4} m^2 / s$	$0.3 m^2 / s$

A general purpose slope limiting procedure will also be required to avoid spurious oscillations. The limiter used in this study is the minmod function

based slope limiter designed by Krivodonova in [43].

For short time computation, the solutions by the LDG scheme with alternating diffusion fluxes for the ions, central diffusive fluxes for the electrons and global Lax-Friedrichs's convective flux for both species, showed good overall agreement with calculations in [41] and [42]. Number density profiles at various instances of time are plotted in Figure 4(a), where time is denoted in a non-dimensionalized form using Maxwell's time scale for the ions $t_{mp} = \frac{\epsilon_0}{em_i\mu_i}$.



(a) LDG solutions compared with computational results by Hilbun and Case in [42]. (b) Negative electron densities at $t = 6t_{mp}$.

Figure 4: The calculations of high-density sheath by the LDG scheme with piece-wise linear polynomials basis without any positivity treatment.

However, at various instances of time, negative number densities were produced by the LDG scheme, see Figure 4(b). The degree of polynomial basis has an effect on the magnitude of negative values predicted. This shows the need for the positivity scheme to be applied. The positivity-preserving DG schemes resolve the discrepancies while maintaining overall accuracy of predictions. See Figure 5 for the results of the positivity-preserving DG scheme with quadratic polynomials.

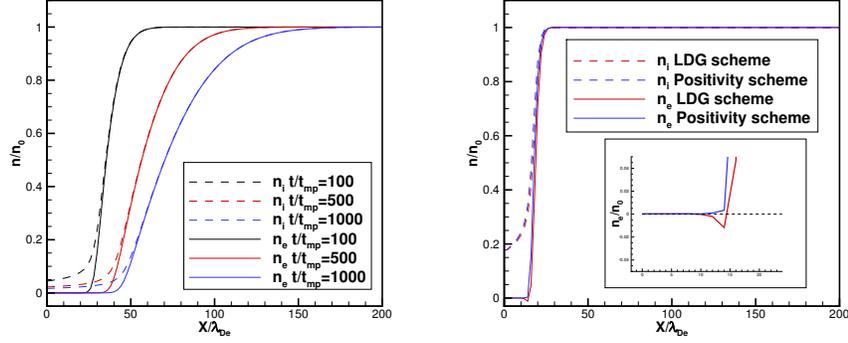
5.2. Direct current glow discharge

To model a steady state glow discharge, Ionization and recombination source terms will have to be considered, and the set of equations to be solved is given by,

$$\frac{\partial n_i}{\partial t} + \nabla \cdot \mathbf{\Gamma}_i = \alpha |\mu_e n_e \mathbf{E}| - \beta n_i n_e \quad (39a)$$

$$\frac{\partial n_e}{\partial t} + \nabla \cdot \mathbf{\Gamma}_e = \alpha |\mu_e n_e \mathbf{E}| - \beta n_i n_e \quad (39b)$$

$$\nabla^2 \phi = -\frac{e}{\epsilon_0} (n_i - n_e) \quad (39c)$$



(a) The positivity-preserving DG scheme with P^2 basis on 100 elements. (b) The positivity-preserving DG scheme with P^2 basis on 100 elements. at $t = 6t_{mp}$

Figure 5: Positivity-preserving DG scheme calculations of high-density sheath.

Note that the ionization term has been modified in order to treat it as a source term rather having to employ a Hamilton-Jacobi type formulation. The computational procedure will involve the consideration of an external circuit, which helps in restricting the calculations to remain in the glow discharge regime. A revised schematic is given in Figure 6.

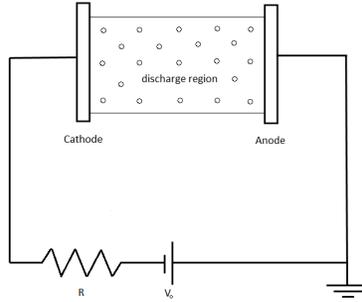


Figure 6: Glow discharge test case schematic representation.

As a consequence of the external circuit, the potential difference across the gap will vary as the discharge evolves in time. For the circuit considered, the voltage drop across the gap is given by

$$\Delta V = V_0 - iR_0, \quad (40)$$

where V_0 is the external EMF supplied, R_0 is the external resistance, i is the current in the circuit and ΔV is the voltage drop across the discharge. The

current is assumed to be the electron drift current calculated at the cathode

$$i = \int e\mu_e n_e(x=0,t) E(x=0,t) dA \quad (41)$$

Since we are currently dealing with a one-dimensional domain, the integral over the electrode area is approximated as the current density at the cathode multiplied by an effective area, taken to be $A = 1.25 \times 10^{-3} m^{-3}$. We consider a discharge in hydrogen gas at 5 Torr pressure, with constant ion and electron temperatures of 0.0258 eV and 1 eV respectively. The transport coefficients used are given in Table 10.

Table 10: Transport coefficients for the hydrogen gas calculations.

quantity	hydrogen
$\mu_i, m^2/(V \cdot s)$	$\frac{0.145}{p}$
$\mu_e, m^2/(V \cdot s)$	$\frac{44}{p}$
$D_i, m^2/s$	$\mu_i T_i$
$D_e, m^2/s$	$\mu_e T_e$
$A, 1/(m \cdot Torr)$	500
$B, V/(m \cdot Torr)$	13000

Calculation of the source terms can be done through numerical quadrature. Solutions of the LDG scheme showed negative number densities predicted, mostly in the region near the anode. This non-physicality was not as persistent as was observed for the sheath phenomena, but nevertheless has to be dealt with. Inclusion of the source terms would also affect the time step constraint to achieve positivity. The ionization term will always provide a positive contribution to the cell average. The recombination is a loss term and has to be included while deriving the CFL condition. See [44] for details of how to ensure the positivity when source term integrals in a DG scheme is computed by quadrature.

The requirement for the positivity-preserving scheme for glow discharge calculations is only at sporadic instances of time. The time step restrictions for positivity could be severely small. On the other hand, all time steps derived in this paper are only sufficient but not necessary for ensuring the positivity, thus it can be used only when it is necessary, as explained in Section 2.5. Another implementation is to employ the positivity-preserving scheme in an adaptive fashion. This can be done by using the LDG scheme at the start of each time step and depending on the need to preserve positivity, the positivity-preserving scheme can be employed. Both positivity-preserving DG schemes and such an adaptive positivity-preserving scheme yield accurate results that compare well with LDG scheme predictions. See Figure 7.

A flowchart for an adaptive positivity-preserving scheme implementation is as follows

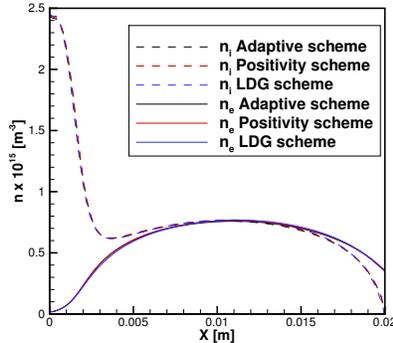


Figure 7: hydrogen discharge, 5 Torr pressure, 100 elements and P^2 basis.

- Start calculations at n -th time step of a Runge-Kutta method using
 - alternating fluxes in the LDG method [18].
 - Time step is determined by linear stability CFL criterion.
- Check for positivity of cell averages for each stage in a Runge-Kutta method,
 - If satisfied, proceed;
 - If not satisfied, go back to start of n -th time step and employ the positivity-preserving DG scheme with time steps sufficient for preserving positivity.
- First apply the slope limiter then apply the positivity-preserving limiters for each time stage.

6. Concluding remarks

In this paper we have demonstrated how to construct a positivity-preserving high order DG scheme by considering a simple positivity-preserving diffusion flux for solving nonlinear diffusion equations, with an application to modeling electrical discharges. Such a positivity-preserving diffusion flux can be regarded as one way of reducing the positivity-preserving flux for Navier-Stokes equations in [1] to scalar diffusion equations. Compared to an alternative method for gradient flow in [13], the main advantage of our approach is that the positivity-preserving flux can be easily constructed without converting a diffusion operator into a gradient flow operator, which is convenient to use for commonly used diffusion operators such as the Laplacian. On the other hand, for this simple diffusion flux to be positivity-preserving, a limiter to control the gradient of solution must be used, in addition to a positivity-preserving limiter. Numerical

tests suggest that neither this simple positivity-preserving diffusion flux nor the two limiters affect the high order accuracy in a high order DG scheme for an initial boundary value problem with a non-negative exact solution $u(x, t)$ satisfying a critical assumption $u(x, t) = 0 \implies u_x(x, t) = 0$.

Acknowledgments

The research was supported by the NSF grant DMS-1522593 and by the School of Aeronautics and Astronautics at Purdue University.

References

- [1] X. Zhang, On positivity-preserving high order discontinuous galerkin schemes for compressible navier–stokes equations, *Journal of Computational Physics* 328 (2017) 301–343.
- [2] A. Genty, C. Le Potier, Maximum and minimum principles for radionuclide transport calculations in geological radioactive waste repository: comparison between a mixed hybrid finite element method and finite volume element discretizations, *Transport in porous media* 88 (1) (2011) 65–85.
- [3] A. Blanchet, J. Dolbeault, B. Perthame, Two-dimensional keller-segel model: Optimal critical mass and qualitative properties of the solutions., *Electronic Journal of Differential Equations (EJDE)*[electronic only] 2006 (2006) Paper–No.
- [4] R. Strehl, A. Sokolov, S. Turek, Efficient, accurate and flexible finite element solvers for chemotaxis problems, *Computers & Mathematics with Applications* 64 (3) (2012) 175–189.
- [5] C. Zhuang, R. Zeng, A positivity-preserving scheme for the simulation of streamer discharges in non-attaching and attaching gases, *Communications in Computational Physics* 15 (1) (2014) 153–178.
- [6] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser Basel, 1992.
- [7] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, *Journal of Computational Physics* 229 (9) (2010) 3091–3120.
- [8] X. Zhang, Y. Liu, C.-W. Shu, Maximum-principle-satisfying high order finite volume weighted essentially nonoscillatory schemes for convection-diffusion equations, *SIAM Journal on Scientific Computing* 34 (2) (2012) A627–A658.

- [9] X. Zhang, C.-W. Shu, Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments, in: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 467, The Royal Society, 2011, pp. 2752–2776.
- [10] S. Gottlieb, D. I. Ketcheson, C.-W. Shu, *Strong stability preserving Runge-Kutta and multistep time discretizations*, World Scientific, 2011.
- [11] J. Yan, Maximum-Principle-Satisfying Direct discontinuous Galerkin method and its variation for convection diffusion equations, submitted for publication (2015).
- [12] Z. Chen, H. Huang, J. Yan, Third order maximum-principle-satisfying direct discontinuous Galerkin methods for time dependent convection diffusion equations on unstructured triangular meshes, *Journal of Computational Physics* 308 (2016) 198–217.
- [13] Z. Sun, J. A. Carrillo, C.-W. Shu, A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials, arXiv preprint.
- [14] T. Xiong, J.-M. Qiu, Z. Xu, High order maximum-principle-preserving discontinuous galerkin method for convection-diffusion equations, *SIAM Journal on Scientific Computing* 37 (2) (2015) A583–A608.
- [15] P. Yang, T. Xiong, J.-M. Qiu, Z. Xu, High order maximum principle preserving finite volume method for convection dominated problems, *Journal of Scientific Computing* 67 (2) (2016) 795–820.
- [16] L. Guo, Y. Yang, Positivity preserving high-order local discontinuous galerkin method for parabolic equations with blow-up solutions, *Journal of Computational Physics* 289 (2015) 181–195.
- [17] F. Bassi, S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations, *Journal of Computational Physics* 131 (2) (1997) 267–279.
- [18] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, *SIAM Journal on Numerical Analysis* 35 (6) (1998) 2440–2463.
- [19] J. Douglas, T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, *Computing methods in applied sciences* (1976) 207–216.
- [20] D. N. Arnold, An interior penalty finite element method with discontinuous elements, *SIAM journal on numerical analysis* 19 (4) (1982) 742–760.

- [21] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM journal on numerical analysis* 39 (5) (2002) 1749–1779.
- [22] B. Riviere, *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*, SIAM, 2008.
- [23] C. E. Baumann, J. T. Oden, A discontinuous hp finite element method for the Euler and Navier–Stokes equations, *International Journal for Numerical Methods in Fluids* 31 (1) (1999) 79–95.
- [24] Y. Cheng, C.-W. Shu, A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives, *Mathematics of Computation* 77 (262) (2008) 699–730.
- [25] A. Uranga, P.-O. Persson, M. Drela, J. Peraire, Implicit large eddy simulation of transitional flows over airfoils and wings, *Proceedings of the 19th AIAA Computational Fluid Dynamics*, AIAA 4131 (7) (2009) 67.
- [26] H. Liu, J. Yan, The direct discontinuous Galerkin (DDG) methods for diffusion problems, *SIAM Journal on Numerical Analysis* 47 (1) (2009) 675–698.
- [27] H. Liu, J. Yan, The direct discontinuous Galerkin (DDG) method for diffusion with interface corrections, *Communications in Computational Physics* 8 (3) (2010) 541.
- [28] J. Yan, A new nonsymmetric discontinuous Galerkin method for time dependent convection diffusion equations, *Journal of Scientific Computing* 54 (2-3) (2013) 663–683.
- [29] Z. Wang, H. Gao, A unifying lifting collocation penalty formulation including the discontinuous Galerkin, spectral volume/difference methods for conservation laws on mixed grids, *Journal of Computational Physics* 228 (21) (2009) 8161–8186.
- [30] T. Haga, H. Gao, Z. J. Wang, A high-order unifying discontinuous formulation for the Navier-Stokes equations on 3D mixed grids, *Mathematical Modelling of Natural Phenomena* 6 (03) (2011) 28–56.
- [31] J. Peraire, N. C. Nguyen, B. Cockburn, A Hybridizable Discontinuous Galerkin Method for the Compressible Euler and Navier-Stokes Equations, in: *Proceedings of 48th AIAA Aerospace Sciences Meeting and Exhibit*, Orlando, Florida (AIAA Paper 2010-363), 2010.
- [32] J. Peraire, N. Nguyen, B. Cockburn, An embedded discontinuous Galerkin method for the compressible Euler and Navier-Stokes equations, in: *Proceedings of the 20th AIAA Computational Fluid Dynamics Conference*, Honolulu, Hawaii, (AIAA Paper 2011-3228), 2011.

- [33] J. S. Hesthaven, T. Warburton, Nodal discontinuous Galerkin methods: algorithms, analysis, and applications, Springer Science & Business Media, 2007.
- [34] Y. Xing, X. Zhang, Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes, *Journal of Scientific Computing* 57 (1) (2013) 19–41.
- [35] Y. Lv, M. Ihme, Entropy-bounded discontinuous Galerkin scheme for Euler equations, *Journal of Computational Physics* 295 (2015) 715–739.
- [36] Y. Zhang, X. Zhang, C.-W. Shu, Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection–diffusion equations on triangular meshes, *Journal of Computational Physics* 234 (2013) 295–316.
- [37] Y. P. Raizer, J. E. Allen, Gas discharge physics, Springer-Berlin, 1997.
- [38] J. Poggie, Multi-fluid modelling of pulsed discharges for flow control applications, *International Journal of Computational Fluid Dynamics* 29 (2) (2015) 180–191.
- [39] J. Poggie, Numerical exploration of flow control with glow discharges, in: 35th AIAA Plasmadynamics and Lasers Conference, Fluid Dynamics and Co-located Conferences, AIAA Paper 2004-2658.
- [40] S. T. Surzhikov, J. S. Shang, Two-component plasma model for two-dimensional glow discharge in magnetic field, *Journal of Computational Physics* 199 (2) (2004) 437–464.
- [41] J. Poggie, D. Gaitonde, Electrode boundary conditions in magnetogasdynamic flow control, in: 40th AIAA Aerospace Sciences Meeting & Exhibit, AIAA Paper 2002-199.
- [42] W. Hilbun, B. Case, Preliminary Development of a Computational Model of a Dielectric Barrier Discharge, in: 43rd AIAA Aerospace Sciences Meeting and Exhibit, AIAA Paper 2005-1176.
- [43] L. Krivodonova, Limiters for high-order discontinuous Galerkin methods, *Journal of Computational Physics* 226 (1) (2007) 879–896.
- [44] X. Zhang, C.-W. Shu, Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms, *Journal of Computational Physics* 230 (4) (2011) 1238–1248.