High order conservative positivity-preserving discontinuous Galerkin method for stationary hyperbolic equations*

Ziyao Xu[†] and Chi-Wang Shu[‡]

Abstract

This is a follow-up work of Yuan et al. (SIAM J Sci Comput 38:A2987-A3019, 2016) and Ling et al. (J Sci Comput 77: 1801-1831, 2018) that further investigates the positivity-preserving discontinuous Galerkin (DG) methods for stationary hyperbolic equations. In 2016, Yuan et al. proposed a high order positivity-preserving DG method for stationary hyperbolic equations with constant coefficients, but the scheme has to be used in combination with a non-conservative rotational limiter introduced in case of negative cell averages. Ling et al. (2018) improved the results in one dimensional space by rigorously proving the positivity of cell averages of the unmodulated DG scheme, which allows the conservative scaling limiter in Zhang et al. (J Comput Phys 229:8918-8934, 2010) to be used to maintain positivity without affecting accuracy, but extension to two space dimensions requires an augmentation of the DG space and works only in the second order case. Considering that the aforementioned works only address stationary hyperbolic equations with constant coefficients and higher than second order conservative methods are still unavailable in two and three space dimensions, we propose high order conservative positivity-preserving DG methods for variable coefficient and nonlinear stationary hyperbolic equations in one dimension and constant coefficient stationary hyperbolic equations in two and three dimensions, via a suitable quadrature in the DG framework. We show the good performance of the algorithms by ample numerical experiments, including their applications in time-dependent problems.

Key Words: high order accuracy, positivity-preserving, conservative schemes, discontinuous Galerkin methods, stationary hyperbolic equations

1 Introduction

In this paper, we are interested in numerical methods for stationary hyperbolic equations. In the one dimensional space, we consider the variable coefficient and nonlinear stationary hyperbolic equations

$$(a(x)u)_x + \lambda u = f(x), \quad x \in \Omega = [0,1], \tag{1.1}$$

*Research supported by AFOSR grant FA9550-20-1-0055 and NSF grant DMS-2010107.

[†]Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. E-mail: ziyao_xu1@brown.edu

[‡]Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. E-mail: chi-wang_shu@brown.edu

where $a(x)$ does not change sign and, without loss of generality $a(x) > 0$, and

$$(a(u)u)_x + \lambda u = f(x), \quad x \in \Omega = [0, 1], \tag{1.2}$$

where $a(u)$ does not change sign and, without loss of generality $a(u) > 0$. Here $\lambda \geq 0$ is a constant. In two and three dimensional spaces, we consider the constant coefficient stationary hyperbolic equations

$$au_x + bu_y + \lambda u = f(x, y), \quad (x, y) \in \Omega = [0, 1]^2, \tag{1.3}$$

and

$$au_x + bu_y + cu_z + \lambda u = f(x, y, z), \quad (x, y, z) \in \Omega = [0, 1]^3, \tag{1.4}$$

respectively, where $\lambda \geq 0$ is a constant and, without loss of generality, we assume $a, b, c > 0$.

The stationary hyperbolic equations (1.1)-(1.4) have wide applications in steady-state transport problems. Moreover, the equations form the building block of the linear radiative transfer equation (RTE), which is an integro-differential equation that describes the distribution of radiative intensity in a medium, based on the discrete-ordinate method (DOM) [9, 11] and iterative procedure on the source terms, see [19, 13] for more details.

The discontinuous Galerkin (DG) method is one of the most popular numerical methods to solve hyperbolic equations, for its advantages in obtaining high order accuracy, flexibility for complex geometry and easiness to be parallelized. In 1970, Reed et al. [14] proposed the first DG scheme to solve the linear steady-state RTE for neutron transport problems. It was later developed into Runge-Kutta discontinuous Galerkin (RKDG) methods by Cockburn et al. in a series of papers [7, 6, 5, 4, 8] to solve time-dependent hyperbolic equations such as the Burgers equation, Euler equations, and shallow water equations, etc. In this paper, we will adopt the classic DG method to solve the stationary hyperbolic equations.

For stationary hyperbolic equations, it is well-known that their physical solutions satisfy the positivity-preserving property, i.e. the solutions are nonnegative, provided the corresponding boundary conditions and source terms are nonnegative. When designing numerical methods, one naturally wants to maintain the positivity-preserving property on the numerical solution, since negative values are not only physically unacceptable, but also may cause severe robustness issues in the simulations, especially when coupled with other physical systems.

There have been intensive studies on positivity-preserving DG methods. In 2010, the genuinely maximum-principle-satisfying DG method was proposed by Zhang et al. in [21] for time-dependent scalar hyperbolic equations. The method is called positivity-preserving when the lower bound in the maximum-principle is zero, which is the case in our problems. The general framework of the positivity-preserving method is composed of two parts. The first part is to obtain the solution at the next time step with nonnegative cell averages from the original, unlimited DG scheme, probably under certain step-size conditions. Once the cell averages of solution are guaranteed nonnegative, the scaling limiter in [21], which maintains the high order accuracy and mass conservation, is applied to modify the solution such that the entire solution becomes nonnegative. Based on this simple but powerful framework, positivity-preserving and maximum-principle-satisfying DG methods for time-dependent problems have been rapidly developed later, e.g. for the Euler equations [22, 23], Navier-Stokes equations [24, 12], shallow water equations[17, 16], convection-diffusion equations [25, 18], and compressible miscible displacements [10], among others.

In 2016, Yuan et al. [19] proposed a high order positivity-preserving DG method for constant coefficient stationary hyperbolic equations. Taking the one dimensional case as an example, their algorithm is as follows: Firstly, they proved a fundamental result that the numerical solution $u(x)$ solved from the unmodulated DG method satisfies $\max\{\bar{u}_K, u_K(x_c)\} \geq 0$ on every cell K of the mesh, where $\bar{u}_K$ is the cell average on K, and x_c is the right end point (the downwind point) of K. They then modify the solution $u_K(x)$ on cell K based on the principle that, if $\bar{u}_K \geq 0$, the conservative scaling limiter [21]

$$\tilde{u}_K(x) = \theta\left(u_K(x) - \bar{u}_K\right) + \bar{u}_K, \text{where } \theta = \min\{\frac{\bar{u}_K}{\bar{u}_K - \min_K u_K(x)}, 1\} \tag{1.5}$$

is applied, otherwise a non-conservative rotational limiter [19] centered at x_c is used. Their algorithm can maintain positivity without affecting high order accuracy, however, since the cell average $\bar{u}_K$ can be changed by the rotation, the algorithm is not conservative in general, which is also true when the algorithm is extended to two-dimensional rectangular [19] or triangular [20] meshes. In 2018, Ling et al. [13] improved the result by rigorously proving that the solution of the unmodulated DG method in one dimension actually satisfies $\bar{u}_K \geq 0$ for all K. Therefore the scaling limiter (1.5) can always be used, which yields a high order conservative positivity-preserving DG method. In their work, a special test function ξ that recovers cell averages $\bar{u}_K$ from the left hand side of the DG scheme was proved to be nonnegative, which implies $\bar{u}_K \geq 0$ since the source term and boundary terms on the right hand side of the DG scheme are both nonnegative, see more details in [13]. Unfortunately, direct extension to two dimensions fails due to the fact that such test function ξ is no longer nonnegative over the cell in rectangular meshes, even for second order DG method with P^1 or Q^1 spaces. Instead, the authors obtained a second order positivity-preserving conservative scheme on rectangular meshes by augmenting the P^1 finite element space, but the extension of this approach to higher space dimensions or to higher order schemes was not carried out in [13] and is highly nontrivial.

In this paper, we further investigate high order conservative positivity-preserving DG method for stationary hyperbolic equations. We put our effort on proving the positivity of cell averages of the scheme so that the conservative scaling limiter (1.5) can be applied directly to maintain high order accuracy and positivity. The main difficulty is that the unmodulated DG method fails in positivity-preserving for cell averages in all the equations we consider in this paper, which will be illustrated by concrete examples in later sections. To resolve this difficulty, we modify the original DG method by adopting appropriate quadrature rules to replace the exact integrals in the schemes, which is a common practice in the implementation of DG schemes, not only because the exact integral is often difficult to obtain, but also for the purpose of achieving specific properties, e.g. maximum-principle-satisfying [21] or entropy stability [3]. The quadrature rules adopted in the schemes are easy to implement and can be directly extended to high dimensions. More importantly, we will show that the cell averages of the DG schemes with such quadrature rules are positive, by proving the positivity of the test function that recovers the cell average from the left hand side of the schemes. The rest of the paper is organized as follows. In Section 2, we propose the conservative positivity-preserving method in the one dimensional space by introducing the desired quadrature rules in the DG formulation, which do not evaluate the integrals in the DG scheme exactly. We give an example to explain why such quadratures are necessary, and rigorously prove the positivity-preserving property of our method. In Section 3, we propose the positivity-preserving DG methods for two and three space dimensions, based on direct extensions from the 1D algorithm. We detail the implementation of the positivity-preserving scaling limiter (1.5) and summarize the complete positivity-preserving algorithm in Section 4. The good performance of the schemes are demonstrated by ample numerical experiments in Section 5. Due to the inaccurate quadrature, the order of convergence is suboptimal in two and three space dimensions, but we observe optimal convergence in all one dimensional tests. Finally, we end in Section 6 with concluding remarks.

2 Numerical algorithm in one space dimension

In this section, we construct high order conservative positivity-preserving DG methods for stationary hyperbolic equations (1.1) and (1.2) in the one dimensional space. The schemes can be arbitrarily high order for the case of (1.1) with $\lambda = 0$, but for the other cases we are only able to prove the positivity-preserving property for P^1 and P^2 (second and third order) DG schemes.

2.1 Notations

We take the partition $0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = 1$ on $\Omega = [0,1]$, and denote the j-th cell by $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, with the cell size $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$ and the cell center $x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$ for $j = 1, 2, \ldots, N$.

The finite element space of P^k-DG scheme is defined as

$$V_h^k = \{v \in L^2([0,1]) : v|_{I_j} \in P^k(I_j), j = 1, 2, \ldots, N\}, \tag{2.1}$$

where $P^k(I)$ is the polynomial space of order no greater than k on I. For $v \in V_h^k$, we define the cell average $\bar{v}_j = \frac{1}{\Delta x_j}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} v(x)dx$ on I_j. Moreover, we denote by $v^-_{j+\frac{1}{2}}$ and $v^+_{j+\frac{1}{2}}$ the left and right limits of v at $x_{j+\frac{1}{2}}$, respectively, i.e. $v^\pm_{j+\frac{1}{2}} = v(x_{j+\frac{1}{2}} \pm 0)$.

For the purpose of positivity-preserving, we adopt the Gauss-Legendre quadrature rule of k points to evaluate volume integrals in the P^k-DG scheme, and denote this quadrature by $⨍_{I_j} v(x)dx = \Delta x_j \sum_{\alpha=1}^{k} \hat{\omega}_\alpha v(\hat{x}_\alpha)$, where $\{\hat{x}_\alpha, \alpha = 1, \ldots, k\}$ are the quadrature points on I_j and $\{\hat{\omega}_\alpha, \alpha = 1, \ldots, k\}$ are the quadrature weights satisfying $\sum_{\alpha=1}^{k} \hat{\omega}_\alpha = 1$.

2.2 Variable coefficient stationary hyperbolic equation in one space dimension

Consider the variable coefficient stationary hyperbolic equation (1.1) with $f(x) \geq 0$ in Ω. As mentioned before, without loss of generality we assume $a(x) > 0$ and the corresponding boundary condition $u(0) = u_0 \geq 0$. The case $a(x) < 0$ with the boundary condition $u(1) = u_0 \geq 0$ can be obtained by the change of variable $x' = 1 - x$.

Firstly, we give an example to show that the original DG scheme with exact integrals may produce negative cell averages, even when the upwind boundary condition and the source term are both positive. The original P^k-DG scheme of the equation (1.1) is to seek $u \in V_h^k$, s.t. $\forall w \in V_h^k$,

$$-\int_{I_j} (a(x)uw_x - \lambda uw)\, dx + a(x_{j+\frac{1}{2}})u^-_{j+\frac{1}{2}}w^-_{j+\frac{1}{2}} = a(x_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}}w^+_{j-\frac{1}{2}} + \int_{I_j} fwdx, \tag{2.2}$$

for $j = 1, 2, \ldots, N$, where we let $u^-_{\frac{1}{2}} = u_0$. We adopt the P^1-DG scheme and take $a(x) = 1 + x, \lambda = 0$ and $u_0 > 0$. It is easy to check that $\xi(x) = \frac{6+5\Delta x_1}{6+8\Delta x_1+2\Delta x_1^2} - \frac{3x}{\Delta x_1(3+\Delta x_1)}$ is the unique function in $P^1(I_1)$ such that $-\int_{I_1} a(x)v\xi_x dx + a(x_{\frac{3}{2}})v^-_{\frac{3}{2}}\xi^-_{\frac{3}{2}} = \bar{v}_1$ for all $v \in V_h^1$, and $\xi(x_{\frac{3}{2}}) = -\frac{\Delta x_1}{2(3+4\Delta x_1+\Delta x_1^2)} < 0$. By taking the test function $w = \xi$ (where we extend $w = 0$ outside I_1) in the scheme, we can construct $f(x) \geq 0$ that takes large values around $x_{\frac{3}{2}}$ such that $\bar{u}_1 = a(0)u_0\xi(x_{\frac{1}{2}}) + \int_{I_1} f\xi dx < 0$. One can check that if we adopt P^2, P^3, P^4, P^5-DG schemes and take $a(x) = 1 + x^2, a(x) = 1 + x^3, a(x) = 1 + x^4, a(x) = 1 + x^5$, respectively, negative cell averages may also appear following the same lines, see the details in Appendix B.

However, we are going to show that the positivity-preserving property can be achieved simply by replacing the exact integrals in the scheme by the Gauss-Legendre quadratures of k points. The positivity-preserving P^k-DG scheme of (1.1) is to seek $u \in V_h^k$, s.t. $\forall w \in V_h^k$,

$$-\oint_{I_j} (a(x)uw_x - \lambda uw)\, dx + a(x_{j+\frac{1}{2}})u^-_{j+\frac{1}{2}}w^-_{j+\frac{1}{2}} = a(x_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}}w^+_{j-\frac{1}{2}} + \oint_{I_j} fwdx, \tag{2.3}$$

for $j = 1, 2, \ldots, N$.

Cockburn et al. have proved in [4] that a sufficient condition for the quadrature in P^k-DG scheme to attain optimal convergence is to have algebraic degree of accuracy $2k$. Though this condition is not satisfied by the quadrature in (2.3), we observe optimal order of convergence in all one dimensional tests.

Based on the framework of [21], we only need to put our effort on proving the positivity of cell averages of the scheme (2.3), then the scaling limiter (1.5) can be used to achieve positivity of the entire solution without losing mass conservation and accuracy. Same as in [13], it suffices to prove the positivity of the test function $\xi \in V_h^k$ that recovers the cell average of the solution from the left hand side of the scheme (2.3).

We assume that $a(x) \in C^k(I_j), j = 1, 2, \ldots, N$, in the P^k-DG scheme to make sense of some norms to be used. We first consider the case $\lambda = 0$ and give the main result as follows.

Lemma 2.1. *Define $\xi(x) = \frac{1}{\Delta x_j}\int_x^{x_{j+\frac{1}{2}}} \mathcal{L}[\frac{1}{a(t)}]dt$ for $x \in I_j$, where $\mathcal{L}[\cdot]$ is the Lagrange interpolation operator at the Gauss-Legendre points $\{\hat{x}_\alpha\}_{\alpha=1}^k$, then ξ is the unique function in $P^k(I_j)$ that satisfies*

$$-\oint_{I_j} a(x)v\xi_x dx + a(x_{j+\frac{1}{2}})v^-_{j+\frac{1}{2}}\xi^-_{j+\frac{1}{2}} = \bar{v}_j, \quad \forall v \in P^k(I_j). \tag{2.4}$$

Moreover, for $k = 1$, $\xi \geq 0$ on I_j; for $k \geq 2$, $\xi \geq 0$ on I_j if the mesh size satisfies

$$\Delta x_j \leq \left(\frac{(2k)!}{k!||a(x)||_{L^\infty(I_j)}||\frac{d^k}{dx^k}\left(\frac{1}{a(x)}\right)||_{L^\infty(I_j)}} \right)^{\frac{1}{k}}. \tag{2.5}$$

Proof. By definition, $\xi \in P^k(I_j)$, $\xi_x(x) = -\frac{1}{\Delta x_j}\mathcal{L}[\frac{1}{a(t)}](x)$, and $\xi^-_{j+\frac{1}{2}} = 0$. Therefore, it follows from direct computation that, $\forall v \in P^k(I_j)$,

$$
\begin{aligned}
&-\int_{I_j} a(x) v \xi_x dx + a(x_{j+\frac{1}{2}}) v^-_{j+\frac{1}{2}} \xi^-_{j+\frac{1}{2}} \\
=& \sum_{\alpha=1}^{k} \hat{\omega}_\alpha a(\hat{x}_\alpha) v(\hat{x}_\alpha) \mathcal{L}[\frac{1}{a(t)}](\hat{x}_\alpha) + 0 \\
=& \sum_{\alpha=1}^{k} \hat{\omega}_\alpha v(\hat{x}_\alpha) = \bar{v}_j,
\end{aligned}
$$

where the last equality holds because the k-point Gauss-Legendre quadrature is exact for integrals of polynomials of order at most k.

As to the uniqueness, we consider the corresponding homogeneous linear problem: Find $\eta \in P^k(I_j)$, s.t.

$$
-\int_{I_j} a(x) v \eta_x dx + a(x_{j+\frac{1}{2}}) v^-_{j+\frac{1}{2}} \eta^-_{j+\frac{1}{2}} = 0, \quad \forall v \in P^k(I_j).
$$

If we take v as the $k+1$ Lagrange basis at $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_k, x_{j+\frac{1}{2}}$, the above linear problem is converted to the system of linear equations

$$
\begin{cases}
\eta_x(\hat{x}_\alpha) = 0, \quad \alpha = 1, 2, \ldots, k \\
\eta(x_{j+\frac{1}{2}}) = 0.
\end{cases}
$$

Since $\eta_x \in P^{k-1}(I_j)$, we have $\eta_x \equiv 0$ from the uniqueness of Lagrange interpolation, which implies $\eta \equiv 0$ since $\eta(x_{j+\frac{1}{2}}) = 0$. Therefore, the function satisfying (2.4) is unique in $P^k(I_j)$.

To show the positivity of ξ, it suffices to prove its integrand $\mathcal{L}[\frac{1}{a(x)}] \geq 0$ on I_j. When $k = 1$, this is clear because the Lagrange interpolant $\mathcal{L}[\frac{1}{a(x)}] = \frac{1}{a(\hat{x}_1)}$ is a constant. When $k \geq 2$, we need the error formula[2] of the Lagrange polynomial for $g(x) \in C^k(I_j)$ interpolating at $\hat{x}_1, \ldots, \hat{x}_k$,

$$
g(x) - \mathcal{L}[g](x) = \frac{g^{(k)}(\zeta(x))}{k!}(x - \hat{x}_1)(x - \hat{x}_2) \cdots (x - \hat{x}_k),
$$

where $\zeta(x) \in I_j$ is generally unknown. Moreover, let us recall that the standard k-th order Legendre polynomial satisfies $|P_k(r)| \leq 1$ for $r \in [-1, 1]$, and has the explicit formula

$$
P_k(r) = \frac{(2k)!}{2^k (k!)^2}(r - \hat{r}_1)(r - \hat{r}_2) \cdots (r - \hat{r}_k),
$$

where $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_k$ are the roots of the k-th order Legendre polynomial. The properties of the Legendre polynomials imply $|\frac{1}{k!}(x - \hat{x}_1)(x - \hat{x}_2) \cdots (x - \hat{x}_k)| = \frac{(\Delta x_j)^k k!}{(2k)!} \left| P_k \left(\frac{x - \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})}{\Delta x_j / 2} \right) \right| \leq \frac{(\Delta x_j)^k k!}{(2k)!}$. Therefore, we have the lower bound estimates for $\mathcal{L}[\frac{1}{a(x)}]$ on I_j as follows,

$$\begin{aligned}\mathcal{L}[\frac{1}{a(t)}](x) &= \frac{1}{a(x)} - \frac{d^k}{dx^k}\left(\frac{1}{a(x)}\right)\Big|_{x=\zeta} \cdot \frac{1}{k!}(x-\hat{x}_1)(x-\hat{x}_2)\cdots(x-\hat{x}_k) \\ &\geq \frac{1}{||a(x)||_{L^\infty(I_j)}} - \frac{(\Delta x_j)^k k!}{(2k)!}||\frac{d^k}{dx^k}\left(\frac{1}{a(x)}\right)||_{L^\infty(I_j)} \\ &\geq 0, \quad \forall x \in I_j,\end{aligned}$$

under the condition $\Delta x_j \leq \left(\frac{(2k)!}{k!||a(x)||_{L^\infty(I_j)}||\frac{d^k}{dx^k}\left(\frac{1}{a(x)}\right)||_{L^\infty(I_j)}}\right)^{\frac{1}{k}}$ on the mesh size. □

Remark 2.1. *The condition* (2.5) *is drawn from the requirement that the Lagrange interpolation* $\mathcal{L}[\frac{1}{a(x)}]$ *being nonnegative on* I_j*. Since we have assumed the smoothness of* $a(x)$*, which implies* $\frac{1}{a(x)}$ *is smooth and lower bounded away from zero, the mesh size condition should not be severe. Indeed, since we merely need the integration* $\int_x^{x_{j+\frac{1}{2}}} \mathcal{L}[\frac{1}{a(t)}]dt \geq 0, x \in I_j$*, to guarantee the positivity of* ξ*, the actual condition needed on the mesh size may be even more relaxed.*

Based on the lemma above, if we assume the inflow condition $u^-_{j-\frac{1}{2}} \geq 0$, we can immediately obtain the positivity of $\bar{u}_j$ by taking the test function $w = \xi$ (extend $\xi = 0$ outside I_j) in the scheme (2.3) and using the fact that the source term f and coefficient $a(x)$ are positive. We can therefore obtain the result for the positivity-preserving property of the scheme (2.3) with $\lambda = 0$ as follows.

Theorem 2.2. *For the variable coefficient stationary hyperbolic equation* (1.1) *with* $\lambda = 0$*, if the source term and inflow conditions from upstream cells (including the inflow condition on the first cell) are positive, then the cell averages of the scheme* (2.3) *are positive, under the mesh size condition in Lemma 2.1.*

We then consider the case $\lambda > 0$ and give the main result as follows.

Lemma 2.3. *Define the functions*

$$\xi_1(x) = \frac{2(x_{j+\frac{1}{2}} - x)}{\Delta x_j(2a(\hat{x}_1) + \lambda\Delta x_j)}, \quad x \in I_j,$$

and

$$\xi_2(x) = \frac{6(x_{j+\frac{1}{2}} - x)\left(\tilde{\lambda}(x - x_{j-\frac{1}{2}}) + a(\hat{x}_1) + a(\hat{x}_2)\right)}{\Delta x_j\left(12a(\hat{x}_1)a(\hat{x}_2) + 3\Delta x_j\lambda(a(\hat{x}_1) + a(\hat{x}_2)) + \Delta x_j^2\lambda^2\right)}, \quad x \in I_j,$$

where $\tilde{\lambda} = \lambda + \frac{\sqrt{3}(a(\hat{x}_1) - a(\hat{x}_2))}{\Delta x_j}$*, for* P^1*-DG and* P^2*-DG schemes, respectively, then* ξ_1 *and* ξ_2 *are the unique functions in* $P^k(I_j)$ *that satisfies*

$$-\fint_{I_j}(a(x)v\xi_x - \lambda v\xi)\,dx + a(x_{j+\frac{1}{2}})v^-_{j+\frac{1}{2}}\xi^-_{j+\frac{1}{2}} = \bar{v}_j, \quad \forall v \in P^k(I_j), \tag{2.6}$$

for $k = 1$ and $k = 2$, respectively.

Moreover, $\xi_1 \geq 0$ on I_j; $\xi_2 \geq 0$ on I_j if $\lambda \geq p_j(a)$, or otherwise $\Delta x_j \leq \frac{2\min_{x\in I_j} a(x)}{p_j(a)-\lambda}$, where $p_j(\cdot)$ is the one-sided Lipschitz seminorm [1] defined as

$$p_j(v) = \sup_{x,y\in I_j, x\neq y} \left(\frac{v(x)-v(y)}{x-y} \right)_+, \text{ where } z_+ = \max(0, z).$$

Proof. It is easy to check by solving the linear equation/system that $\xi_1(x)$ and $\xi_2(x)$ are the unique solutions of the linear problem (2.6) for $k = 1$ and $k = 2$, respectively.

It is also clear that $\xi_1(x) \geq 0$ on I_j, since $a(x), \lambda > 0$ by assumption.

As for $k = 2$, the positivity of $\xi_2(x)$ is always the same to its factor $\tilde{\lambda}(x - x_{j-\frac{1}{2}}) + a(\hat{x}_1) + a(\hat{x}_2)$. Note that $\tilde{\lambda} = \lambda - \frac{a(\hat{x}_1)-a(\hat{x}_2)}{\hat{x}_1-\hat{x}_2} \geq \lambda - p_j(a)$, thereby $\tilde{\lambda}(x - x_{j-\frac{1}{2}}) + a(\hat{x}_1) + a(\hat{x}_2) \geq a(\hat{x}_1) + a(\hat{x}_2) \geq 0$ if $\lambda \geq p_j(a)$, or $\tilde{\lambda}(x - x_{j-\frac{1}{2}}) + a(\hat{x}_1) + a(\hat{x}_2) \geq a(\hat{x}_1) + a(\hat{x}_2) - (p_j(a) - \lambda)\Delta x_j \geq 0$ if $\lambda < p_j(a)$. Both cases indicate that $\xi_2(x) \geq 0$ on I_j. □

Following the same arguments as before, we can immediately get the positivity of $\bar{u}_j$ if we assume the positivity of the inflow condition and the source term. We can therefore obtain the result for the positivity-preserving property of the scheme (2.3) with $\lambda > 0$ (in fact it also applies to the case of $\lambda = 0$) as follows.

Theorem 2.4. *For the variable coefficient stationary hyperbolic equation* (1.1) *with $\lambda > 0$, if the source term and the inflow conditions from upstream cells (including the inflow condition on the first cell) are positive, then the cell averages of the scheme* (2.3) *are positive, under the conditions in Lemma 2.3.*

Remark 2.2. *We are only able to prove the positivity-preserving property for P^k-DG methods with $k = 1$ and $k = 2$ here. For the cases $k \geq 3$, the positivity of test function ξ satisfying* (2.6) *is too complicated to be analyzed generally. However, we have investigated these cases for some special $a(x)$ and the results are promising, which are shown in Appendix B.*

2.3 Nonlinear stationary hyperbolic equation in one space dimension

Consider the nonlinear stationary hyperbolic equation (1.2) with $f \geq 0$ in Ω. We assume $a(u) \geq c > 0$, $\frac{d(a(u)u)}{du} > 0$ for all u, and the boundary condition $u(0) \geq 0$.

Formally, we still have the same positivity-preserving results as in the variable coefficient case if we adopt the scheme: seek $u \in V_h^k$, s.t. $\forall w \in V_h^k$,

$$-\int_{I_j} (a(u)uw_x - \lambda uw)\, dx + a(u^-_{j+\frac{1}{2}})u^-_{j+\frac{1}{2}}w^-_{j+\frac{1}{2}} = a(u^-_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}}w^+_{j-\frac{1}{2}} + \int_{I_j} fwdx, \tag{2.7}$$

for $j = 1, 2, \ldots, N$, since $a(u)$ in the scheme can be regarded as $a(u(x))$ in the variable coefficient case. However, because $u(x)$ is unknown, the mesh size conditions established before for positivity-preserving is unavailable for $k \geq 2$. To resolve this difficulty, we give a P^2-DG scheme which is positivity-preserving on arbitrary meshes: seek $u \in V_h^2$, s.t. $\forall w \in V_h^2$,

$$-\int_{I_j} (a(u)uw_x - \lambda uw)\, dx + a(u^-_{j+\frac{1}{2}})u^-_{j+\frac{1}{2}}w^-_{j+\frac{1}{2}} = a(u^-_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}}w^+_{j-\frac{1}{2}} + \int_{I_j} fw dx, \tag{2.8}$$

for $j = 1, 2, \ldots, N$, where $\int_{I_j}$ denotes the Simpson's quadrature rule.

We give the main result for the P^2-DG scheme (2.8) as follows.

Lemma 2.5. *Let $u(x)$ be the solution of the scheme* (2.8) *and define the function*

$$\xi(x) = \frac{6(x_{j+\frac{1}{2}} - x)\left(\tilde{\lambda}(x - x_{j-\frac{1}{2}}) + a(u(\hat{x}_1)) + a(u(\hat{x}_2))\right)}{\Delta x_j \left(12a(u(\hat{x}_1))a(u(\hat{x}_2)) + 3\Delta x_j \lambda (a(u(\hat{x}_1)) + a(u(\hat{x}_2))) + \Delta x_j^2 \lambda^2\right)}, \quad x \in I_j, \tag{2.9}$$

where $\tilde{\lambda} = \lambda + \frac{\sqrt{3}(a(u(\hat{x}_1)) - a(u(\hat{x}_2)))}{\Delta x_j}$, then $\xi \in P^2(I_j)$ satisfies

$$-\int_{I_j} (a(u)v\xi_x - \lambda v\xi)\, dx + a(u^-_{j+\frac{1}{2}})v^-_{j+\frac{1}{2}}\xi^-_{j+\frac{1}{2}} = \bar{v}_j, \quad \forall v \in P^2(I_j). \tag{2.10}$$

Moreover, $\xi \geq 0$ at the points $\{x_{j-\frac{1}{2}}, x_j, x_{j+\frac{1}{2}}\}$.

Proof. It can be verified by direct computations similar to the proofs before. □

Following the same arguments as in the variable coefficient case, we immediately get the positivity of $\bar{u}_j$, if we assume the positivity of inflow condition and source term. Though the expression of ξ in (2.9) contains the unknown solution u, it is not a problem since we actually do not use ξ in the implementation of the positivity-preserving algorithm. We can therefore obtain the result for the positivity-preserving property of the schemes (2.7) for $k = 1$ and (2.8) for $k = 2$ as follows.

Theorem 2.6. *For the nonlinear stationary hyperbolic equation* (1.2)*, if the source term and inflow conditions from upstream cells (including the inflow condition on the first cell) are positive, then the cell averages of the schemes* (2.7) *for $k = 1$ and* (2.8) *for $k = 2$ are positive on arbitrary meshes.*

3 Numerical algorithm in two and three space dimensions

In this section, we construct high order conservative positivity-preserving DG schemes for constant coefficient stationary hyperbolic equations (1.3) and (1.4) in two and three dimensions, respectively. The schemes are direct extensions from the algorithm in one space dimension. We are only able to give rigorous proofs of positivity-preserving for limited cases but numerical computation shows strong evidence that the schemes are positivity-preserving for Q^k-DG for arbitrary k in two dimensions, and for odd $k = 1, 3, 5, 7, \dots$ in three dimensions.

3.1 Notations

We take the partition $0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N_x+\frac{1}{2}} = 1$, $0 = y_{\frac{1}{2}} < y_{\frac{3}{2}} < \cdots < y_{N_y+\frac{1}{2}} = 1$, and $0 = z_{\frac{1}{2}} < z_{\frac{3}{2}} < \cdots < z_{N_z+\frac{1}{2}} = 1$ in the x, y and z directions, respectively, and define the mesh sizes $\Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, i = 1, \dots, N_x$, $\Delta y_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}, j = 1, \dots, N_y$, and $\Delta z_l = z_{l+\frac{1}{2}} - z_{l-\frac{1}{2}}, l = 1, \dots, N_z$, with cell centers $x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}), i = 1, \dots, N_x$, $y_j = \frac{1}{2}(y_{j-\frac{1}{2}} + y_{j+\frac{1}{2}}), j = 1, \dots, N_y$, and $z_l = \frac{1}{2}(z_{l-\frac{1}{2}} + z_{l+\frac{1}{2}}), l = 1, \dots, N_z$. Moreover, we denote by $K_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}], i = 1, \dots, N_x, j = 1, \dots, N_y$ the cells in the two dimensional domain $\Omega = [0,1]^2$, and $K_{i,j,l} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}] \times [z_{l-\frac{1}{2}}, z_{l+\frac{1}{2}}], i = 1, \dots, N_x, j = 1, \dots, N_y, l = 1, \dots, N_z$ the cells in the three dimensional domain $\Omega = [0,1]^3$.

The finite element spaces of the Q^k-DG scheme are defined as

$$V_h^k = \{v \in L^2([0,1]^2) : v|_{K_{i,j}} \in Q^k(K_{i,j}), i = 1, \dots, N_x, j = 1, \dots, N_y\}, \tag{3.1}$$

and

$$V_h^k = \{v \in L^2([0,1]^3) : v|_{K_{i,j,l}} \in Q^k(K_{i,j,l}), i = 1, \dots, N_x, j = 1, \dots, N_y, l = 1, \dots, N_z\}, \tag{3.2}$$

in two and three dimensional domains, respectively, where $Q^k(K)$ is the tensor product polynomial space of order no greater than k on the cell K. For $v \in V_h^k$, we denote the cell average by $\bar{v}_{i,j}$ on $K_{i,j}$, and $\bar{v}_{i,j,l}$ on $K_{i,j,l}$. In two space dimensions, we define the left/right and lower/upper limits of v on the vertical and horizontal cell interfaces by $v(x^{\pm}_{i+\frac{1}{2}}, y) = v(x_{i+\frac{1}{2}} \pm 0, y)$ and $v(x, y^{\pm}_{j+\frac{1}{2}}) = v(x, y_{j+\frac{1}{2}} \pm 0)$, respectively. In three space dimensions, the limits on cell interfaces are defined similarly.

We let $\{\hat{r}_\alpha, \hat{\omega}_\alpha\}_{\alpha=1}^k$ and $\{\tilde{r}_\alpha, \tilde{\omega}_\alpha\}_{\alpha=1}^{k+1}$ be the Gauss-Legendre quadrature rules with k and $k+1$ quadrature points on $[-1,1]$, respectively. As in the previous section, we use the notation $⨍$ to denote the approximate integration via the k-point Gauss-Legendre quadrature. If not otherwise stated, the usual integral notation $\int$ stands for the exact integral, which can be evaluated by the $k+1$ point Gauss-Legendre quadrature in the Q^k-DG scheme for the constant coefficient problems. Finally, we denote by $\{\ell_i(x), i = 1, \dots, k\}$ the Lagrange interpolation basis at $\{\hat{r}_\alpha\}_{\alpha=1}^k$ with $\ell_i(\hat{r}_\alpha) = \delta_{i,\alpha}$, and by $\ell_i'(x)$ the derivative of $\ell_i(x)$.

3.2 Constant coefficient stationary hyperbolic equation in two space dimensions

Consider the constant coefficients stationary hyperbolic equation (1.3) with $f(x,y) \geq 0$ in Ω. As mentioned before, without loss of generality, we may assume $a, b > 0$, because the other cases can be obtained by the change of variables $x' = 1 - x$ and/or $y' = 1 - y$. The corresponding boundary conditions are given by $u(0,y) = g_1(y), u(x,0) = g_2(x)$, where $g_1, g_2 \geq 0$.

Firstly, we would like to remark that the original DG methods are not positivity-preserving for the cell averages in general, even for the P^1-DG or Q^1-DG schemes. One can refer to the counterexamples constructed in [13].

The positivity-preserving Q^k-DG scheme of (1.3) is to seek $u \in V_h^k$ s.t. $\forall w \in V_h^k$,

$$
\begin{aligned}
&-⨍_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} ⨍_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} (auw_x + buw_y - \lambda uw)\, dxdy + \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} au(x_{i+\frac{1}{2}}^-, y) w(x_{i+\frac{1}{2}}^-, y) dy + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} bu(x, y_{j+\frac{1}{2}}^-) w(x, y_{j+\frac{1}{2}}^-) dx \\
&= \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} au(x_{i-\frac{1}{2}}^-, y) w(x_{i-\frac{1}{2}}^+, y) dy + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} bu(x, y_{j-\frac{1}{2}}^-) w(x, y_{j-\frac{1}{2}}^+) dx + ⨍_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} ⨍_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} fw dxdy,
\end{aligned}
\tag{3.3}
$$

for $i = 1, \ldots, N_x, j = 1, \ldots, N_y$. If $x_{i-\frac{1}{2}} = 0$, we let $u(x_{i-\frac{1}{2}}^-, y) = g_1(y)$, similarly if $y_{j-\frac{1}{2}} = 0$, we let $u(x, y_{j-\frac{1}{2}}^-) = g_2(x)$. The quadrature adopted in (3.3) does not satisfy the condition for optimal convergence established in [4], which results in sub-optimal convergence as we will show in the numerical tests.

Without loss of generality, we only consider scheme (3.3) on the reference cell $K = [-1,1] \times [-1,1]$, as any cell $K_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ can be transferred to K by changing of coordinates which only rescales a, b, λ, f without altering their signs. We give the main result as follows.

Lemma 3.1. *Define $\xi(x,y;a,b,\lambda) = (1-x)(1-y)\eta(x,y;a,b,\lambda)$ for $(x,y) \in [-1,1]^2$, where $\eta(x,y;a,b,\lambda) = \sum_{i,j=1}^k \eta_{ij}(a,b,\lambda)\ell_i(x)\ell_j(y)$, and $\{\eta_{ij}(a,b,\lambda)\}_{i,j=1}^k$ is the solution of the linear system*

$$
\begin{aligned}
\sum_{i,j=1}^k &\left(a\left((1-\hat{r}_\alpha)(1-\hat{r}_\beta)\ell_i'(\hat{r}_\alpha)\delta_{\beta,j} - (1-\hat{r}_\beta)\delta_{\alpha,i}\delta_{\beta,j}\right) + b\left((1-\hat{r}_\alpha)(1-\hat{r}_\beta)\ell_j'(\hat{r}_\beta)\delta_{\alpha,i} - (1-\hat{r}_\alpha)\delta_{\alpha,i}\delta_{\beta,j}\right)\right. \\
&\left. -\lambda(1-\hat{r}_\alpha)(1-\hat{r}_\beta)\delta_{\alpha,i}\delta_{\beta,j}\right)\eta_{ij} = -\frac{1}{4}, \quad \alpha, \beta = 1, 2, \ldots, k,
\end{aligned}
\tag{3.4}
$$

then $\xi(x,y;a,b,\lambda) \in Q^k([-1,1]^2)$ satisfies

$$
-⨍_{-1}^{1} ⨍_{-1}^{1} (av\xi_x + bv\xi_y - \lambda v\xi)\, dxdy + \int_{-1}^{1} av(1,y)\xi(1,y)dy + \int_{-1}^{1} bv(x,1)\xi(x,1)dx = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} vdxdy, \tag{3.5}
$$

for any $v \in Q^k([-1,1]^2)$.

Moreover, for $k = 1, 2$, we have $\xi(x, y; a, b, \lambda) \geq 0$ on $[-1, 1]^2$; for $k = 3$, we can show $\xi(\hat{r}_\alpha, \hat{r}_\beta; a, b, 0) \geq 0$, $\alpha, \beta = 1, 2, 3$ and $\xi(-1, \tilde{r}_\alpha; a, b, 0), \xi(\tilde{r}_\alpha, -1; a, b, 0) \geq 0$, $\alpha = 1, 2, 3, 4$.

Proof. By definition of $\xi(x, y)$, we can compute that $\xi_x(x, y) = (1 - x)(1 - y) \sum_{i,j=1}^k \eta_{ij} \ell_i'(x)\ell_j(y) - (1 - y) \sum_{i,j=1}^k \eta_{ij} \ell_i(x)\ell_j(y)$ and $\xi_y(x, y) = (1-x)(1-y) \sum_{i,j=1}^k \eta_{ij} \ell_i(x)\ell_j'(y) - (1-x) \sum_{i,j=1}^k \eta_{ij} \ell_i(x)\ell_j(y)$, thereby it can be checked that $\{\eta_{ij}\}_{i,j=1}^k$ is the solution of the linear system (3.4) if and only if ξ satisfies

$$a\xi_x(\hat{r}_\alpha, \hat{r}_\beta) + b\xi_y(\hat{r}_\alpha, \hat{r}_\beta) - \lambda\xi(\hat{r}_\alpha, \hat{r}_\beta) = -\frac{1}{4}, \quad \alpha, \beta = 1, 2, \ldots, k.$$

Moreover, we have $\xi(1, y) = \xi(x, 1) = 0$ from the definition. Therefore, it follows from direct computation that

$$\begin{aligned} &-\fint_{-1}^1 \fint_{-1}^1 (av\xi_x + bv\xi_y - \lambda v\xi)\, dxdy + \int_{-1}^1 av(1, y)\xi(1, y)dy + \int_{-1}^1 bv(x, 1)\xi(x, 1)dx \\ =& -4 \sum_{\alpha,\beta=1}^k \hat{\omega}_\alpha \hat{\omega}_\beta v(\hat{r}_\alpha, \hat{r}_\beta) \left(a\xi_x(\hat{r}_\alpha, \hat{r}_\beta) + b\xi_y(\hat{r}_\alpha, \hat{r}_\beta) - \lambda\xi(\hat{r}_\alpha, \hat{r}_\beta)\right) + 0 + 0 \\ =& \sum_{\alpha,\beta=1}^k \hat{\omega}_\alpha \hat{\omega}_\beta v(\hat{r}_\alpha, \hat{r}_\beta) = \frac{1}{4} \int_{-1}^1 \int_{-1}^1 v dxdy, \qquad \forall v \in Q^k([-1, 1]^2), \end{aligned}$$

where the last equality follows from the fact that the tensor product of k-point Gauss-Legendre quadrature is accurate for $v \in Q^k([-1, 1]^2)$.

It remains to show the positivity of ξ, or equivalently η.

When $k = 1$, by solving the linear equation (3.4), we have $\eta(x, y; a, b, \lambda) = \frac{1}{4(a+b+\lambda)} > 0$.

When $k = 2$, by solving the linear system (3.4), we have $\eta(x, y; a, b, \lambda) = C^{-1}(6a^3 + 15a^2b + 15ab^2 + 6b^3 + 9a^2\lambda + 17ab\lambda + 9b^2\lambda + 5a\lambda^2 + 5b\lambda^2 + \lambda^3 + 3a^2bx + 9ab^2x + 6b^3x + 3a^2\lambda x + 9ab\lambda x + 9b^2\lambda x + 3a\lambda^2x + 5b\lambda^2x + \lambda^3x + 6a^3y + 9a^2by + 3ab^2y + 9a^2\lambda y + 9ab\lambda y + 3b^2\lambda y + 5a\lambda^2y + 3b\lambda^2y + \lambda^3y + 9a^2bxy + 9ab^2xy + 3a^2\lambda xy + 9ab\lambda xy + 3b^2\lambda xy + 3a\lambda^2xy + 3b\lambda^2xy + \lambda^3xy)$, where $C = \frac{16}{9}(3a^2 + 3ab + 3b^2 + 3a\lambda + 3b\lambda + \lambda^2)(3a^2 + 6ab + 3b^2 + 3a\lambda + 3b\lambda + \lambda^2) > 0$. Since $\eta \in Q^1([-1, 1]^2)$ and $\eta(-1, -1) = C^{-1}(12a^2b + 12ab^2 + 8ab\lambda) > 0$, $\eta(-1, 1) = C^{-1}(12a^3 + 12a^2b + 12a^2\lambda + 8ab\lambda + 4a\lambda^2) > 0$, $\eta(1, -1) = C^{-1}(12ab^2 + 12b^3 + 8ab\lambda + 12b^2\lambda + 4b\lambda^2) > 0$, $\eta(1, 1) = C^{-1}(12a^3 + 36a^2b + 36ab^2 + 12b^3 + 24a^2\lambda + 44ab\lambda + 24b^2\lambda + 16a\lambda^2 + 16b\lambda^2 + 4\lambda^3) > 0$, we have $\eta(x, y; a, b, \lambda) > 0$ for $(x, y) \in [-1, 1]^2$.

Now we consider the case $k = 3$ with $\lambda = 0$. Firstly, we note that from the definition, $\xi(x, y; a, b, \lambda) = C\xi(x, y; Ca, Cb, C\lambda)$ and $\eta(x, y; a, b, \lambda) = C\eta(x, y; Ca, Cb, C\lambda)$, $\forall C > 0$. Therefore it suffices to investigate the case $a = 1, b > 0$ since $\xi(x, y; a, b, 0) = \frac{1}{a}\xi(x, y; 1, \frac{b}{a}, 0)$. By solving the linear system (3.4), we get

$\eta_{ij}(1,b,0) = \frac{P_{ij}(b)}{Q(b)}, i,j = 1,2,3$, where $P_{ij}(b)$ and $Q(b)$ are polynomials defined as:

$$P_{11}(b) = 2(5(5-\sqrt{15}) + 5(17-4\sqrt{15})b + (195-31\sqrt{15})b^2 + (240-38\sqrt{15})b^3 + (195-31\sqrt{15})b^4 + 5(17-4\sqrt{15})b^5 + 5(-5+\sqrt{15})b^6)$$

$$P_{12}(b) = 20 + (95+3\sqrt{15})b + 180b^2 + 14(15-\sqrt{15})b^3 + (195-29\sqrt{15})b^4 + 25(5-\sqrt{15})b^5 + 10(5-\sqrt{15})b^6$$

$$P_{13}(b) = 2(5(5+\sqrt{15}) + 5(8+\sqrt{15})b + (45+\sqrt{15})b^2 + 30b^3 + (45-\sqrt{15})b^4 + 5(8-\sqrt{15})b^5 + 5(5-\sqrt{15})b^6)$$

$$P_{21}(b) = 10(5-\sqrt{15}) + 25(5-\sqrt{15})b + (195-29\sqrt{15})b^2 + 14(15-\sqrt{15})b^3 + 180b^4 + (95+3\sqrt{15})b^5 + 20b^6$$

$$P_{22}(b) = 20 + 95b + 198b^2 + 249b^3 + 198b^4 + 95b^5 + 20b^6$$

$$P_{23}(b) = 10(5+\sqrt{15}) + 25(5+\sqrt{15})b + (195+29\sqrt{15})b^2 + 14(15+\sqrt{15})b^3 + 180b^4 + (95-3\sqrt{15})b^5 + 20b^6$$

$$P_{31}(b) = 2(5(5-\sqrt{15}) + 5(8-\sqrt{15})b + (45-\sqrt{15})b^2 + 30b^3 + (45+\sqrt{15})b^4 + 5(8+\sqrt{15})b^5 + 5(5+\sqrt{15})b^6)$$

$$P_{32}(b) = 20 + (95-3\sqrt{15})b + 180b^2 + 14(15+\sqrt{15})b^3 + (195+29\sqrt{15})b^4 + 25(5+\sqrt{15})b^5 + 10(5+\sqrt{15})b^6$$

$$P_{33}(b) = 2(5(5+\sqrt{15}) + 5(17+4\sqrt{15})b + (195+31\sqrt{15})b^2 + (240+38\sqrt{15})b^3 + (195+31\sqrt{15})b^4 + 5(17+4\sqrt{15})b^5 + 5(5+\sqrt{15})b^6)$$

$$Q(b) = 16(1+b)(5 + 15b + 27b^2 + 31b^3 + 27b^4 + 15b^5 + 5b^6)$$

One can observe that all coefficients in the above polynomials are positive. Therefore, we have $\eta(\hat{r}_\alpha, \hat{r}_\beta; a, b, 0) = \frac{1}{a}\eta(\hat{r}_\alpha, \hat{r}_\beta; 1, \frac{b}{a}, 0) = \frac{1}{a}\frac{P_{\alpha,\beta}(b/a)}{Q(b/a)} > 0$, for $\alpha, \beta = 1,2,3$. Further more, since $\eta(x,y;1,b,0) = \sum_{i,j=1}^{3} \eta_{ij}(1,b,0)\ell_i(x)\ell_j(y) = \frac{\sum_{i,j=1}^{3} P_{ij}(b)\ell_i(x)\ell_j(y)}{Q(b)}$, the values of η at the quadrature points $\{(-1,\tilde{r}_\alpha), \alpha = 1,2,3,4\}$ and $\{(\tilde{r}_\alpha, -1), \alpha = 1,2,3,4\}$ are also rational functions of b. By direct computation, one can check that the coefficients of these rational functions are all positive, which implies the positivity of $\eta(x,y;a,b,0)$ at these points. We omit the details of computation since it is straightforward but lengthy. □

Remark 3.1. *By the Cramer's rule, we always have $\eta_{ij}(1,b,0) = \frac{P_{ij}(b)}{Q(b)}$, where $P_{ij}(b)$ and $Q(b)$ are polynomials, $i,j = 1,2,\ldots,k$, for general k. However, Mathematica is unable to afford the symbolic calculation for $k > 3$. We sample some values of b and solve the corresponding values of $P_{i,j}(b)$ and $Q(b)$ numerically. By interpolation, we recover the expressions of $P_{i,j}(b)$ and $Q(b)$, and find that all coefficients of them are nonnegative for $k = 4$. Unfortunately, even numerical computation are difficult for the case $k \geq 5$.*

Based on the lemma above, if we assume the positivity of the inflow conditions $u(x^-_{i-\frac{1}{2}}, y)$ and $u(x, y^-_{j-\frac{1}{2}})$, we can prove the positivity of $\bar{u}_{i,j}$ by taking the test function $w = \xi$ (extend $\xi = 0$ outside $K_{i,j}$) in the scheme (3.3) and using the fact that the source term f and coefficients a, b are positive. We can therefore obtain the result for the positivity-preserving property of the scheme (3.3) as follows.

Theorem 3.2. *For the constant coefficient stationary hyperbolic equation* (1.3)*, if the source term and inflow* *conditions from upstream cells (including the inflow conditions on inflow boundary cells) are positive, then the cell averages of the scheme* (3.3) *are positive for the* Q^1, Q^2*-DG schemes with* $\lambda \geq 0$*, and* Q^3*-DG scheme with* $\lambda = 0$*.*

Though we are not able to give rigorous proofs for the positivity-preserving property of the scheme (3.3) with $k = 3, \lambda > 0$ or $k > 3, \lambda \geq 0$ due to the difficulty of symbolically solving the large linear system (3.4), we can still investigate these cases numerically.

For any given values of a, b, λ, we can always solve for $\{\eta_{ij}\}_{i,j=1}^{k}$ numerically from the linear system (3.4) to obtain the values of η at the quadrature points $\{(\hat{r}_\alpha, \hat{r}_\beta)\}_{\alpha,\beta=1}^{k}, \{(-1, \tilde{r}_\alpha)\}_{\alpha=1}^{k+1}$ and $\{(\tilde{r}_\alpha, -1)\}_{\alpha=1}^{k+1}$ used on the right hand side of (3.3). The scheme is positivity-preserving if η is positive at all these quadrature points. Moreover, we can take advantage of the relationship $\eta(x, y; a, b, \lambda) = C\eta(x, y; Ca, Cb, C\lambda)$, $\forall C > 0$, to reduce the computation. If $\lambda \geq \max\{a, b\}$, we use $\eta(x, y; a, b, \lambda) = \frac{1}{\lambda}\eta(x, y; \frac{a}{\lambda}, \frac{b}{\lambda}, 1)$; otherwise we assume $a \geq \max\{b, \lambda\}$ without loss of generality and use $\eta(x, y; a, b, \lambda) = \frac{1}{a}\eta(x, y; 1, \frac{b}{a}, \frac{\lambda}{a})$. Therefore, we only need to numerically investigate the positivity of η in the two cases $0 \leq a, b \leq 1, \lambda = 1$ and $a = 1, 0 \leq b, \lambda \leq 1$.

We define

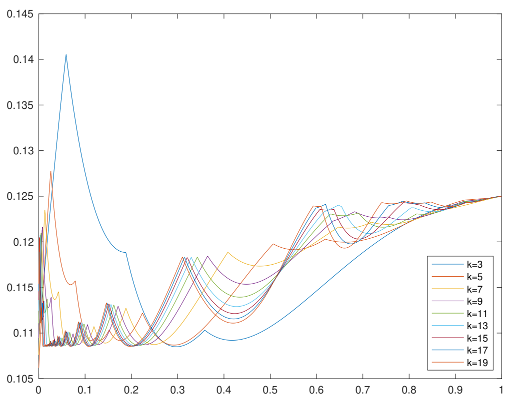
$$\eta_1(k) = \min_{0\leq a,b\leq 1} \min_{1\leq\alpha\leq k+1} \{\eta(-1, \tilde{r}_\alpha; a, b, 1), \eta(\tilde{r}_\alpha, -1; a, b, 1)\},$$

$$\eta_2(k) = \min_{0\leq b,\lambda\leq 1} \min_{1\leq\alpha\leq k+1} \{\eta(-1, \tilde{r}_\alpha; 1, b, \lambda), \eta(\tilde{r}_\alpha, -1; 1, b, \lambda)\},$$

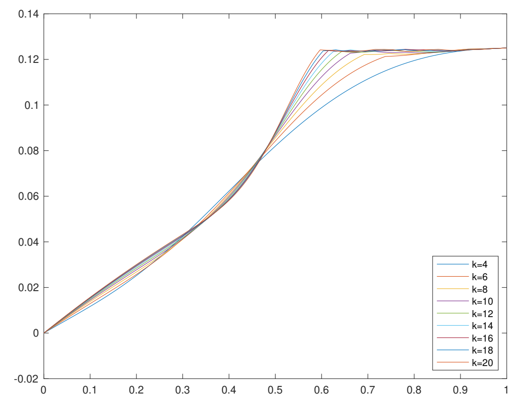
$$\eta_3(k) = \min_{0\leq a,b\leq 1} \min_{1\leq\alpha,\beta\leq k} \eta(\hat{r}_\alpha, \hat{r}_\beta; a, b, 1),$$

$$\eta_4(k) = \min_{0\leq b,\lambda\leq 1} \min_{1\leq\alpha,\beta\leq k} \eta(\hat{r}_\alpha, \hat{r}_\beta; 1, b, \lambda),$$

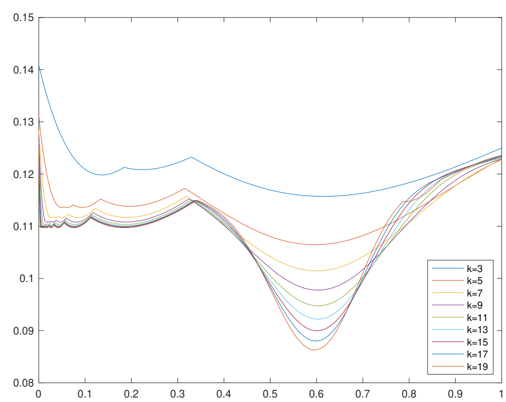
and equally space 1000×1000 points of (a, b) or (b, λ) on $[0, 1] \times [0, 1]$ to approximate $\min_{0\leq a,b\leq 1}$ and $\min_{0\leq b,\lambda\leq 1}$, and give the approximate values $\tilde{\eta}_i(k), i = 1, 2, 3, 4$ in Table 1 and Table 2 for odd and even k, respectively. From the tables, we can observe that the minimum value of η at the quadrature points is zero (machine epsilon) on boundaries when k is even, and strictly positive in all other cases. Moreover, we visualize a particular case $\lambda = 0$, and plot $h_1^k(b) = \min_{1\leq\alpha\leq k+1}\{\eta(-1, \tilde{r}_\alpha; 1, b, 0), \eta(\tilde{r}_\alpha, -1; 1, b, 0)\}$, $h_2^k(b) = \min_{1\leq\alpha,\beta\leq k} \eta(\hat{r}_\alpha, \hat{r}_\beta; 1, b, 0)$ for $b \in [0, 1]$ in the Figure 1, from which we can observe the same pattern as shown in the tables.



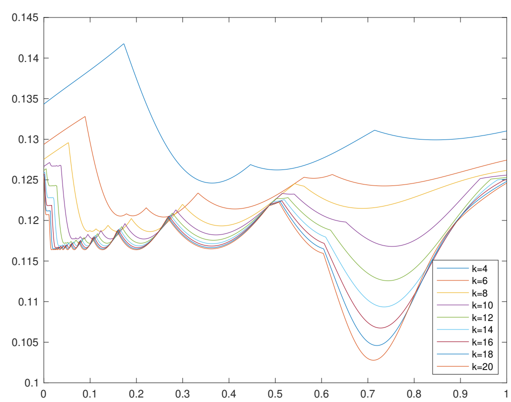
(a) $h_1^k(b)$, k is odd



(b) $h_1^k(b)$, k is even



(c) $h_2^k(b)$, k is odd



(d) $h_2^k(b)$, k is even

Figure 1: $h_1^k(b)$ and $h_2^k(b)$ for different k, 1000 points equally spaced on $[0, 1]$

k	3	5	7	9	11	13	15	17	19
$\tilde{\eta}_1$	4.75E-02	4.59E-02	4.65E-02	4.73E-02	4.80E-02	4.86E-02	4.91E-02	4.93E-02	4.92E-02
$\tilde{\eta}_2$	4.75E-02	4.59E-02	4.65E-02	4.73E-02	4.80E-02	4.86E-02	4.91E-02	4.93E-02	4.92E-02
$\tilde{\eta}_3$	5.67E-02	5.17E-02	5.01E-02	4.93E-02	4.90E-02	4.88E-02	4.86E-02	4.85E-02	4.85E-02
$\tilde{\eta}_4$	5.67E-02	5.17E-02	5.01E-02	4.93E-02	4.90E-02	4.88E-02	4.86E-02	4.85E-02	4.85E-02

Table 1: $\tilde{\eta}_i(k), i = 1, 2, 3, 4$ with odd k

k	4	6	8	10	12	14	16	18	20
$\tilde{\eta}_1$	-1.11E-15	-1.78E-15	-2.66E-15	-4.44E-15	-5.33E-15	-3.02E-14	-2.84E-14	-5.68E-14	-3.20E-14
$\tilde{\eta}_2$	-2.22E-16	-2.78E-16	-3.89E-16	-2.36E-16	-4.72E-16	-1.05E-15	-7.77E-16	-1.16E-15	-7.22E-16
$\tilde{\eta}_3$	5.98E-02	5.64E-02	5.51E-02	5.44E-02	5.40E-02	5.37E-02	5.33E-02	5.29E-02	5.27E-02
$\tilde{\eta}_4$	5.98E-02	5.64E-02	5.51E-02	5.44E-02	5.40E-02	5.37E-02	5.33E-02	5.29E-02	5.27E-02

Table 2: $\tilde{\eta}_i(k), i = 1, 2, 3, 4$ with even k

3.3 Constant coefficient stationary hyperbolic equation in three space dimensions

Consider the constant coefficient stationary hyperbolic equation (1.4) with $f(x, y, z) \geq 0$ in Ω. Without loss of generality, we assume $a, b, c > 0$. The corresponding boundary conditions are given by $u(0, y, z) = g_1(y, z), u(x, 0, z) = g_2(x, z)$ and $u(x, y, 0) = g_3(x, y)$, where $g_1, g_2, g_3 \geq 0$.

The positivity-preserving Q^k-DG scheme of (1.4) is to seek $u \in V_h^k$, where k is odd, s.t. $\forall w \in V_h^k$

$$
\begin{aligned}
&-\fint_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \fint_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \fint_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} (auw_x + buw_y + cuw_z - \lambda uw)\, dxdydz + \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \int_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} au(x_{i+\frac{1}{2}}^-, y, z) w(x_{i+\frac{1}{2}}^-, y, z) dydz \\
&+ \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} bu(x, y_{j+\frac{1}{2}}^-, z) w(x, y_{j+\frac{1}{2}}^-, z) dxdz + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} cu(x, y, z_{l+\frac{1}{2}}^-) w(x, y, z_{l+\frac{1}{2}}^-) dxdy \\
&= \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \int_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} au(x_{i-\frac{1}{2}}^-, y, z) w(x_{i-\frac{1}{2}}^+, y, z) dydz + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} bu(x, y_{j-\frac{1}{2}}^-, z) w(x, y_{j-\frac{1}{2}}^+, z) dxdz \\
&+ \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} cu(x, y, z_{l-\frac{1}{2}}^-) w(x, y, z_{l-\frac{1}{2}}^+) dxdy + \fint_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \fint_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \fint_{z_{l-\frac{1}{2}}}^{z_{l+\frac{1}{2}}} fw dxdydz,
\end{aligned}
\tag{3.6}
$$

for $i = 1, \dots, N_x, j = 1, \dots, N_y, l = 1, \dots, N_z$. If $x_{i-\frac{1}{2}} = 0$, we let $u(x^-_{i-\frac{1}{2}}, y, z) = g_1(y, z)$, similarly, if $y_{j-\frac{1}{2}} = 0$ or $z_{l-\frac{1}{2}} = 0$, we let $u(x, y^-_{j-\frac{1}{2}}, z) = g_2(x, z)$ or $u(x, y, z^-_{l-\frac{1}{2}}) = g_3(x, y)$, respectively. The sub-optimal convergence is observed in numerical experiments due to the inaccurate quadrature rule adopted in the scheme.

Without loss of generality, we only consider the scheme (3.6) on the reference cell $K = [-1, 1]^3$, as any cell $K_{i,j,l}$ can be transferred to K by changing of coordinates with only rescales a, b, c, λ, f without altering their signs. We give the main results as follows.

Lemma 3.3. *Define $\xi(x, y, z; a, b, c, \lambda) = (1-x)(1-y)(1-z)\eta(x, y, z; a, b, c, \lambda)$, where $\eta(x, y, z; a, b, c, \lambda) = \sum_{i,j,l=1}^k \eta_{ijl}(a, b, c, \lambda)\ell_i(x)\ell_j(y)\ell_l(z)$, and $\{\eta_{ijl}(a, b, c, \lambda)\}_{i,j,l=1}^k$ is the solution of the linear system*

$$
\begin{aligned}
\sum_{i,j,l=1}^{k} &\left(a\left((1-\hat{r}_\alpha)(1-\hat{r}_\beta)(1-\hat{r}_\gamma)\ell_i'(x_\alpha)\delta_{\beta,j}\delta_{\gamma,l} - (1-\hat{r}_\beta)(1-\hat{r}_\gamma)\delta_{\alpha,i}\delta_{\beta,j}\delta_{\gamma,l}\right)\right. \\
&+ b\left((1-\hat{r}_\alpha)(1-\hat{r}_\beta)(1-\hat{r}_\gamma)\ell_j'(x_\beta)\delta_{\alpha,i}\delta_{\gamma,l} - (1-\hat{r}_\alpha)(1-\hat{r}_\gamma)\delta_{\alpha,i}\delta_{\beta,j}\delta_{\gamma,l}\right) \\
&+ c\left((1-\hat{r}_\alpha)(1-\hat{r}_\beta)(1-\hat{r}_\gamma)\ell_l'(x_\gamma)\delta_{\alpha,i}\delta_{\beta,j} - (1-\hat{r}_\alpha)(1-\hat{r}_\beta)\delta_{\alpha,i}\delta_{\beta,j}\delta_{\gamma,l}\right) \\
&\left.-\lambda(1-\hat{r}_\alpha)(1-\hat{r}_\beta)(1-\hat{r}_\gamma)\delta_{\alpha,i}\delta_{\beta,j}\delta_{\gamma,l}\right)\eta_{ijl} \\
&= -\frac{1}{8}, \quad \alpha, \beta, \gamma = 1, 2, \dots, k,
\end{aligned}
\tag{3.7}
$$

then $\xi(x, y, z; a, b, c, \lambda) \in Q^k([-1, 1]^3)$ satisfies

$$
\begin{aligned}
&-\int_{-1}^1\int_{-1}^1\int_{-1}^1 (av\xi_x + bv\xi_y + cv\xi_z - \lambda v\xi)\, dxdydz + \int_{-1}^1\int_{-1}^1 av(1, y, z)\xi(1, y, z)dydz + \int_{-1}^1\int_{-1}^1 bv(x, 1, z)\xi(x, 1, z)dxdz \\
&+ \int_{-1}^1\int_{-1}^1 cv(x, y, 1)\xi(x, y, 1)dxdy = \frac{1}{8}\int_{-1}^1\int_{-1}^1\int_{-1}^1 vdxdy,
\end{aligned}
\tag{3.8}
$$

for any $v \in Q^k([-1, 1]^3)$.

Moreover, for $k = 1$, we have $\xi(x, y, z; a, b, c, \lambda) \geq 0$ for $(x, y) \in [-1, 1]^3$.

Proof. By definition of $\xi(x, y, z)$, we can compute that $\xi_x(x, y, z) = (1-x)(1-y)(1-z)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i'(x)\ell_j(y)\ell_l(z) - (1-y)(1-z)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i(x)\ell_j(y)\ell_l(z)$, $\xi_y(x, y, z) = (1-x)(1-y)(1-z)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i(x)\ell_j'(y)\ell_l(z) - (1-x)(1-z)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i(x)\ell_j(y)\ell_l(z)$, and $\xi_z(x, y, z) = (1-x)(1-y)(1-z)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i(x)\ell_j(y)\ell_l'(z) - (1-x)(1-y)\sum_{i,j,l=1}^k \eta_{ijl}\ell_i(x)\ell_j(y)\ell_l(z)$, thereby it is easy to check that $\{\eta_{ijl}\}_{i,j,l=1}^k$ is the solution of the linear system (3.7) if and only if ξ satisfies

$$
a\xi_x(\hat{r}_\alpha, \hat{r}_\beta, \hat{r}_\gamma) + b\xi_y(\hat{r}_\alpha, \hat{r}_\beta, \hat{r}_\gamma) + c\xi_z(\hat{r}_\alpha, \hat{r}_\beta, \hat{r}_\gamma) - \lambda\xi(\hat{r}_\alpha, \hat{r}_\beta, \hat{r}_\gamma) = -\frac{1}{8}, \quad \alpha, \beta, \gamma = 1, 2, \dots, k.
$$

Moreover, $\xi(1,y,z) = \xi(x,1,z) = \xi(x,y,1) = 0$ from the definition. Therefore, it follows from direct computation that (3.8) holds. When $k=1$, we can solve ξ from (3.7) to obtain $\xi(x,y,z) = \frac{1}{8(a+b+c+\lambda)}(1-x)(1-y)(1-z) \geq 0$ in $[-1,1]^3$. $\square$

Based on the above lemma, we can obtain the result for the positivity-preserving property of the scheme (3.6) as follows.

Theorem 3.4. *For the constant coefficient stationary hyperbolic equation* (1.4)*, if the source term and inflow conditions from upstream cells (including the inflow conditions on inflow boundary cells) are positive, then the cell averages of the scheme* (3.6) *are positive for the Q^1-DG scheme.*

We are of course not satisfied with only Q^1-DG positivity-preserving scheme, which has first order convergence rate by numerical experiments. Similar to the two dimensional case, we numerically investigate the positivity of $\eta(x,y,z)$ at the quadrature points used on the right hand side of (3.6) for larger k. It suffices to consider two cases: $0 \leq a,b,c \leq 1, \lambda = 1$ and $a = 1, 0 \leq b,c,\lambda \leq 1$ because of the property $\xi(x,y,z;a,b,c,\lambda) = C\xi(x,y,z;Ca,Cb,Cc,C\lambda), \forall C > 0$ and the symmetry in x,y,z directions.

We define

$$\eta_1(k) = \min_{0\leq a,b,c\leq 1} \min_{1\leq\alpha,\beta\leq k+1} \{\eta(-1,\hat{r}_\alpha,\hat{r}_\beta;a,b,c,1), \eta(\hat{r}_\alpha,-1,\hat{r}_\beta;a,b,c,1), \eta(\hat{r}_\alpha,\hat{r}_\beta,-1;a,b,c,1)\},$$

$$\eta_2(k) = \min_{0\leq b,c,\lambda\leq 1} \min_{1\leq\alpha,\beta\leq k+1} \{\eta(-1,\hat{r}_\alpha,\hat{r}_\beta;1,b,c,\lambda), \eta(\hat{r}_\alpha,-1,\hat{r}_\beta;1,b,c,\lambda), \eta(\hat{r}_\alpha,\hat{r}_\beta,-1;1,b,c,\lambda)\},$$

$$\eta_3(k) = \min_{0\leq a,b,c\leq 1} \min_{1\leq\alpha,\beta,\gamma\leq k} \eta(\hat{r}_\alpha,\hat{r}_\beta,\hat{r}_\gamma;a,b,c,1),$$

$$\eta_4(k) = \min_{0\leq b,c,\lambda\leq 1} \min_{1\leq\alpha,\beta,\gamma\leq k} \eta(\hat{r}_\alpha,\hat{r}_\beta,\hat{r}_\gamma;1,b,c,\lambda)$$

and equally space $100\times100\times100$ points for $k=2,3,4$, $30\times30\times30$ points for $k=5,6,\ldots,10$, of (a,b,c) or (b,c,λ) on $[0,1]^3$ to approximate $\min_{0\leq a,b,c\leq 1}$ and $\min_{0\leq b,c,\lambda\leq 1}$. We give the approximate values $\tilde{\eta}_i(k), i = 1,2,3,4$ in Table 3. From the table, we can observe that the minimum value of η at quadrature points is negative on boundaries when k is even, and strictly positive in all other cases, which suggest that we should use odd k for the purpose of positivity-preserving.

4 Implementation of the algorithms

In this section, we summarize the results obtained in the previous sections and illustrate the implementation of the positivity-preserving algorithms.

k	2	3	4	5	6	7	8	9	10
$\tilde{\eta}_1$	-4.44E-16	1.04E-02	-4.00E-15	9.75E-03	-1.60E-14	1.00E-02	-1.70E-03	1.04E-02	-6.05E-03
$\tilde{\eta}_2$	-3.97E-06	1.04E-02	-1.63E-03	9.75E-03	-6.15E-03	1.00E-02	-1.25E-02	1.04E-02	-2.01E-02
$\tilde{\eta}_3$	2.61E-02	1.39E-02	1.61E-02	1.20E-02	1.43E-02	1.14E-02	1.38E-02	1.18E-02	1.36E-02
$\tilde{\eta}_4$	2.61E-02	1.39E-02	1.61E-02	1.20E-02	1.43E-02	1.14E-02	1.38E-02	1.18E-02	1.36E-02

Table 3: $\tilde{\eta}_i(k), i = 1, 2, 3, 4$

Firstly, we introduce a robust version of the positivity-preserving limiter (1.5) used in practice. We set a small threshold $\epsilon > 0$, e.g. $\epsilon = 10^{-14}$, and denote by S the set of points where we want to preserve the positivity of function values. The set S must include the quadrature points used on the inflow boundaries in the schemes for the purpose of positivity-preserving. To be more precise, S must include the point $x_{i+\frac{1}{2}}$ on I_i in one space dimension, the points $\{(x_{i+\frac{1}{2}}, \tilde{y}_\alpha)\}_{\alpha=1}^{k+1}$, $\{(\tilde{x}_\alpha, y_{j+\frac{1}{2}})\}_{\alpha=1}^{k+1}$ on $K_{i,j}$ in two space dimensions, and the points $\{(x_{i+\frac{1}{2}}, \tilde{y}_\alpha, \tilde{z}_\beta)\}_{\alpha,\beta=1}^{k+1}$, $\{(\tilde{x}_\alpha, y_{j+\frac{1}{2}}, \tilde{z}_\beta)\}_{\alpha,\beta=1}^{k+1}$, $\{(\tilde{x}_\alpha, \tilde{y}_\beta, z_{l+\frac{1}{2}})\}_{\alpha,\beta=1}^{k+1}$ on $K_{i,j,l}$ in three space dimensions, where $\tilde{x}_\alpha = x_i + \frac{1}{2}\Delta x_i \tilde{r}_\alpha$, $\tilde{y}_\alpha = y_j + \frac{1}{2}\Delta y_j \tilde{r}_\alpha$, $\tilde{z}_\alpha = z_l + \frac{1}{2}\Delta z_l \tilde{r}_\alpha$, $\alpha = 1, 2, \ldots, k+1$ are the $(k+1)$-point Gauss-Legendre quadrature points in different directions. On a cell K with the cell average $\bar{u}_K \geq 0$, if $\bar{u}_K \leq \epsilon$, we take the modified solution $\tilde{u}_K(\mathbf{x}) \equiv \bar{u}_K$, otherwise, we take the modified solution as

$$\tilde{u}_K(\mathbf{x}) = \theta\,(u_K(\mathbf{x}) - \bar{u}_K) + \bar{u}_K, \text{where } \theta = \min\{\frac{\bar{u}_K - \epsilon}{\bar{u}_K - \min_{\mathbf{x}\in S} u_K(\mathbf{x})}, 1\}, \tag{4.1}$$

where $\mathbf{x}$ denotes the coordinates in one, two or three space dimensions.

In one dimensional space, we compute the solution u_i on cell I_i based on the solution $\tilde{u}_{i-1}$ with $\tilde{u}_{i-1}(x) \geq 0, x \in S$. Once u_i is obtained from the scheme with $\bar{u}_i \geq 0$, we apply the above limiter to obtain a modified solution $\tilde{u}_i$, which will be used in the computation on the next cell.

Similarly, in two dimensional space, we compute the solution $u_{i,j}$ on cell $K_{i,j}$ based on the solution $\tilde{u}_{i-1,j}, \tilde{u}_{i,j-1}$ with $\tilde{u}_{i-1,j}(x,y), \tilde{u}_{i,j-1}(x,y) \geq 0, (x,y) \in S$. Once $u_{i,j}$ is obtained, we apply the above limiter to obtain the modified solution $\tilde{u}_{i,j}$, which will be used in later computations. In three dimensional space, we compute the solution $u_{i,j,l}$ on cell $K_{i,j,l}$ based on the solution $\tilde{u}_{i-1,j,l}, \tilde{u}_{i,j-1,l}, \tilde{u}_{i,j,l-1}$ with $\tilde{u}_{i-1,j,l}(x,y,z), \tilde{u}_{i,j-1,l}(x,y,z), \tilde{u}_{i,j,l-1}(x,y,z) \geq 0, (x,y,z) \in S$. Once $u_{i,j,l}$ is obtained, we apply the above limiter to obtain the modified solution $\tilde{u}_{i,j,l}$ and use it in the later computations.

We would like to remark that, under certain mesh size conditions, the positivity of the solution at the interfaces $x^-_{j+\frac{1}{2}}, j = 1, 2, \dots, N$ in one dimensional space is automatically maintained even without the positivity-preserving limiter, i.e. $u^-_{j+\frac{1}{2}} \geq 0, j = 1, 2, \dots, N$ provided $f, u_0 \geq 0$. This fact allows us to apply the positivity-preserving limiter simultaneously for all cells after the DG solution has been obtained for all cells. The detailed theorem and its proof are given in Appendix A.

5 Numerical examples

In this section, we perform numerical experiments to show the good performance of the positivity-preserving methods established in the previous sections. Many of the examples are taken from [19, 13, 20]. We take the set S in the positivity-preserving limiter of the Section 4 as the union of the necessary points introduced therein and 100 equally spaced points on 1D cells, or 50×50 equally spaced points on 2D cells, or $20 \times 20 \times 20$ equally spaced points on 3D cells. If not otherwise stated, we use uniform meshes with mesh sizes satisfying the conditions of positivity-preserving established in the previous sections.

Example 5.1. *We solve the equation* (1.1) *with* $a(x) = \frac{1}{2+\sin(4\pi x)}, \lambda = 0$ *and* $f(x) = x^2$ *on the domain* $\Omega = [0, 1]$. *The boundary condition is given by* $u(0) = 0$ *and the exact solution is* $u(x) = \frac{2}{3}x^3 + \frac{1}{3}\sin(4\pi x)x^3$. *We compute the solution based on the positivity-preserving scheme* (2.3) *and give the errors, order of convergence, and data about positivity in Tables 4 and 5 for the cases without and with the limiter, respectively. From the tables, we can see that the orders of convergence are optimal, and the negative values of the solution of the scheme without limiter are eliminated by the positivity-preserving limiter.*

Example 5.2. *We solve the equation* (1.1) *with* $a(x) = 1, \lambda = 6000$ *and* $f(x) = \lambda\left(\frac{1}{9}\cos^4(x) + \epsilon\right) - \frac{4}{9}\cos^3(x)\sin(x)$ *on the domain* $\Omega = [0, \pi]$. *We take* $\epsilon = 10^{-14}$ *such that the source term is nonnegative. The boundary condition is given by* $u(0) = \frac{1}{9} + \epsilon$ *and the exact solution is* $u(x) = \frac{1}{9}\cos^4(x) + \epsilon$. *This example has been tested in [13] with a rigorously proved high order conservative positivity-preserving method. However, since the inaccurate integral is adopted in our scheme, the results of our algorithm will be different. We collect the numerical errors, orders of convergence, and data about positivity in Tables 6 and 7 for the schemes* (2.3) *without and with the limiter, respectively, from which we can observe the optimal convergence, and the negative values of the solution being eliminated by the positivity-preserving limiter.*

Example 5.3. *We solve the equation* (1.1) *with* $a(x) = 1 + x, \lambda = 10000$ *and* $f(x) = (\lambda + 1)\left(\frac{1}{9}\cos^4(x) + \epsilon\right) - (1 + x)\left(\frac{4}{9}\cos^3(x)\sin(x)\right)$ *on the domain* $\Omega = [0, 2\pi]$. *We take* $\epsilon = 2 \times 10^{-14}$ *such that the source term is*

k	N	L^1 error	order	L^∞ error	order	$\min u_h$
1	20	1.78E-03	-	2.89E-02	-	-8.71E-06
	40	4.41E-04	2.01	7.27E-03	1.99	-5.96E-07
	80	1.10E-04	2.00	1.83E-03	1.99	-3.81E-08
	160	2.75E-05	2.00	4.57E-04	2.00	-2.39E-09
	320	6.88E-06	2.00	1.14E-04	2.00	-1.50E-10
2	20	8.34E-05	-	1.53E-03	-	-3.46E-06
	40	1.06E-05	2.98	2.10E-04	2.86	-4.17E-07
	80	1.32E-06	3.01	2.91E-05	2.85	-5.24E-08
	160	1.64E-07	3.00	3.81E-06	2.93	-6.63E-09
	320	2.05E-08	3.00	4.86E-07	2.97	-8.38E-10
3	20	4.42E-06	-	1.15E-04	-	-9.64E-07
	40	2.76E-07	4.00	7.31E-06	3.98	-7.63E-08
	80	1.72E-08	4.01	4.57E-07	4.00	-5.03E-09
	160	1.07E-09	4.00	2.86E-08	4.00	-3.18E-10
	320	6.70E-11	4.00	1.79E-09	4.00	-2.00E-11
4	20	1.36E-07	-	3.41E-06	-	-6.96E-08
	40	4.23E-09	5.01	1.09E-07	4.97	-1.14E-09
	80	1.34E-10	4.98	3.39E-09	5.00	-1.80E-11
	160	4.17E-12	5.00	1.06E-10	5.00	-2.81E-13
	320	1.30E-13	5.00	3.32E-12	5.00	-4.40E-15

Table 4: Results of Example 5.1 without limiter

k	N	L^1 error	order	L^∞ error	order	Limited cells (%)
1	20	1.78E-03	-	2.89E-02	-	5.00
	40	4.41E-04	2.01	7.27E-03	1.99	2.50
	80	1.10E-04	2.00	1.83E-03	1.99	1.25
	160	2.75E-05	2.00	4.57E-04	2.00	0.63
	320	6.88E-06	2.00	1.14E-04	2.00	0.31
2	20	8.41E-05	-	1.52E-03	-	5.00
	40	1.07E-05	2.97	2.10E-04	2.85	2.50
	80	1.34E-06	3.00	2.92E-05	2.85	1.25
	160	1.67E-07	3.00	3.82E-06	2.93	0.63
	320	2.09E-08	3.00	4.88E-07	2.97	0.31
3	20	5.45E-06	-	1.13E-04	-	5.00
	40	3.84E-07	3.83	7.17E-06	3.98	2.50
	80	2.52E-08	3.93	4.46E-07	4.01	1.25
	160	1.61E-09	3.97	2.79E-08	4.00	0.63
	320	1.02E-10	3.99	1.74E-09	4.00	0.31
4	20	2.35E-07	-	3.50E-06	-	5.00
	40	5.81E-09	5.33	1.10E-07	4.99	2.50
	80	1.54E-10	5.23	3.42E-09	5.01	1.25
	160	4.44E-12	5.12	1.07E-10	5.00	0.63
	320	1.47E-13	4.92	3.34E-12	5.00	0.31

Table 5: Results of Example 5.1 with limiter

k	N	L^1 error	order	L^∞ error	order	$\min u_h$
1	20	2.14E-03	-	2.64E-03	-	-1.33E-03
	40	4.66E-04	2.20	6.76E-04	1.96	-2.90E-04
	80	7.94E-05	2.55	1.70E-04	2.00	-4.34E-05
	160	1.15E-05	2.78	4.21E-05	2.01	-1.41E-06
	320	2.41E-06	2.26	1.04E-05	2.02	-1.20E-10
	640	5.94E-07	2.02	2.51E-06	2.05	-1.66E-11
2	20	3.40E-04	-	4.59E-04	-	-2.68E-04
	40	6.79E-05	2.32	1.04E-04	2.15	-4.43E-05
	80	1.00E-05	2.76	1.86E-05	2.48	-8.12E-08
	160	1.25E-06	3.00	2.12E-06	3.13	-3.03E-07
	320	9.42E-08	3.72	1.55E-07	3.78	-8.65E-09
	640	6.07E-09	3.96	9.80E-09	3.98	-1.52E-10

Table 6: Results of Example 5.2 without limiter

k	N	L^1 error	order	L^∞ error	order	Limited cells (%)
1	20	1.20E-03	-	2.64E-03	-	10.00
	40	2.97E-04	2.02	6.76E-04	1.96	5.00
	80	6.59E-05	2.17	1.70E-04	2.00	2.50
	160	1.14E-05	2.53	4.21E-05	2.01	1.25
	320	2.41E-06	2.24	1.04E-05	2.02	0.31
	640	5.94E-07	2.02	2.51E-06	2.05	0.16
2	20	2.41E-04	-	4.59E-04	-	15.00
	40	5.56E-05	2.11	1.04E-04	2.15	7.50
	80	9.98E-06	2.48	1.86E-05	2.48	2.50
	160	1.24E-06	3.01	2.12E-06	3.13	1.88
	320	9.41E-08	3.72	1.55E-07	3.78	0.94
	640	6.07E-09	3.95	9.80E-09	3.98	0.31

Table 7: Results of Example 5.2 with limiter

nonnegative. The boundary condition is given by $u(0) = \frac{1}{9} + \epsilon$ *and the exact solution is* $u(x) = \frac{1}{9}\cos^4(x) + \epsilon$. *We compute the solution using the scheme* (2.3) *and give the numerical errors, orders of convergence, and data about positivity in Tables 8 and 9 for the case without and with the limiter, respectively. From the tables, we can see that the orders of convergence are optimal, and that negative values appear without limiter and the positivity is maintained under the modification of the limiter.*

k	N	L^1 error	order	L^∞ error	order	$\min u_h$
1	20	8.35E-03	-	1.07E-02	-	-3.03E-03
	40	1.76E-03	2.25	2.83E-03	1.93	-5.01E-04
	80	4.10E-04	2.10	7.02E-04	2.01	-1.10E-04
	160	9.38E-05	2.13	1.73E-04	2.02	-1.60E-05
	320	2.15E-05	2.13	4.28E-05	2.02	-4.49E-07
	640	5.19E-06	2.05	1.05E-05	2.02	-3.21E-10
2	20	1.62E-03	-	1.26E-03	-	-4.51E-04
	40	3.73E-04	2.11	2.89E-04	2.12	-1.03E-04
	80	7.96E-05	2.23	7.02E-05	2.04	-1.62E-05
	160	1.29E-05	2.62	1.32E-05	2.41	-1.19E-06
	320	1.35E-06	3.26	1.85E-06	2.83	-1.45E-07
	640	1.03E-07	3.71	1.66E-07	3.48	-3.78E-09

Table 8: Results of Example 5.3 without limiter

Example 5.4. *We solve the equation* (1.2) *with* $a(u) = u^2 + 0.01, \lambda = 5$ *and* $f(x) = -8\sin(x)\cos^7(x)\left(3(\cos^8(x)+\epsilon)^2 + 0.01\right) + \lambda\left(\cos^8(x)+\epsilon\right)$ *on the domain* $\Omega = [0, \pi]$. *We take* $\epsilon = 10^{-14}$ *such that the source term is nonnegative. The boundary condition is given by* $u(0) = 1 + \epsilon$ *and the exact solution is* $u(x) = \cos^8(x) + \epsilon$. *We give the errors, orders of convergence, and data about positivity in Tables 10 and 11 for the scheme* (2.7) *with* $k = 1$ *and scheme* (2.8) *with* $k = 2$ *in the case without and with the limiter, respectively, with the same conclusion about accuracy and positivity-preserving as before.*

Example 5.5. *We solve the equation* (1.3) *with* $a = 0.7, b = 0.3, \lambda = 1.0$ *and* $f = 0$ *on the domain* $\Omega = [0, 1] \times [0, 1]$. *The boundary conditions are given by* $u(x, 0) = 0$ *for* $0 \le x \le 1$ *and* $u(0, y) = \sin^6(\pi y)$ *for*

k	N	L^1 error	order	L^∞ error	order	Limited cells (%)
1	20	6.44E-03	-	1.07E-02	-	10.00
	40	1.53E-03	2.08	2.83E-03	1.93	5.00
	80	3.75E-04	2.03	7.02E-04	2.01	3.75
	160	9.12E-05	2.04	1.73E-04	2.02	1.88
	320	2.15E-05	2.09	4.28E-05	2.02	0.94
	640	5.19E-06	2.05	1.05E-05	2.02	0.31
2	20	1.44E-03	-	1.37E-03	-	20.00
	40	3.36E-04	2.10	3.25E-04	2.07	10.00
	80	7.62E-05	2.14	7.37E-05	2.14	6.25
	160	1.29E-05	2.56	1.32E-05	2.49	3.13
	320	1.35E-06	3.26	1.85E-06	2.83	1.56
	640	1.03E-07	3.71	1.66E-07	3.48	0.78

Table 9: Results of Example 5.3 with limiter

k	N	L^1 error	order	L^∞ error	order	$\min u_h$
1	20	1.20E-01	-	1.28E-01	-	-1.17E-01
	40	8.11E-03	3.89	1.86E-02	2.78	-2.97E-03
	80	1.55E-03	2.39	4.39E-03	2.08	-2.03E-07
	160	3.87E-04	2.01	1.09E-03	2.02	-1.04E-13
	320	9.66E-05	2.00	2.71E-04	2.00	8.96E-15
	640	2.41E-05	2.00	6.76E-05	2.00	9.86E-15
2	20	8.56E-02	-	1.76E-01	-	-7.20E-02
	40	1.16E-02	2.88	4.47E-02	1.98	-1.44E-02
	80	9.28E-04	3.65	4.78E-03	3.22	-9.60E-05
	160	8.20E-05	3.50	4.31E-04	3.47	-4.44E-09
	320	8.27E-06	3.31	4.41E-05	3.29	-7.82E-14
	640	9.25E-07	3.16	4.89E-06	3.17	9.98E-15

Table 10: Results of Example 5.4 without limiter

k	N	L^1 error	order	L^∞ error	order	Limited cells (%)
1	20	3.85E-02	-	1.28E-01	-	15.00
	40	7.02E-03	2.46	1.86E-02	2.78	12.50
	80	1.55E-03	2.18	4.39E-03	2.08	3.75
	160	3.87E-04	2.01	1.09E-03	2.02	1.25
	320	9.66E-05	2.00	2.71E-04	2.00	0.00
	640	2.41E-05	2.00	6.76E-05	2.00	0.00
2	20	1.89E-02	-	7.65E-02	-	50.00
	40	6.35E-03	1.58	4.26E-02	0.85	37.50
	80	9.15E-04	2.80	4.78E-03	3.15	18.75
	160	8.20E-05	3.48	4.31E-04	3.47	6.25
	320	8.27E-06	3.31	4.41E-05	3.29	1.25
	640	9.25E-07	3.16	4.89E-06	3.17	0.00

Table 11: Results of Example 5.4 with limiter

$0 \le y \le 1$. *It is easy to check that the exact solution of the problem is*

$$u(x,y) = \begin{cases} 0, & y < \frac{b}{a}x \\ \sin^6(\pi(y - \frac{b}{a}x))e^{-\frac{\lambda}{a}x} & y \ge \frac{b}{a}x \end{cases}$$

We compute the solution based on the scheme (3.3) *with* $k = 1, 2, 3, 4, 5$. *The errors, orders of convergence and data about positivity are given in Tables 12 and 13 for the cases without and with positivity-preserving limiter, respectively, from which the sub-optimal convergence can be observed. Moreover, we plot the results of the scheme with the limiter for* $k = 1, 2, 3, 4$ *on the* 40×40 *mesh in Figure 2, in which we put white dots on those cells where negative values appear before the limiting process.*

Example 5.6. *We solve the equation* (1.3) *with* $a = 0.6, b = 0.4, \lambda = 0$ *and* $f = 0$ *on the domain* $\Omega = [0,1]^2$. *The boundary condition is given by* $u(x,0) = 1$ *for* $0 < x \le 1$ *and* $u(0,y) = 0$ *for* $0 \le y \le 1$. *The exact solution of the problem is*

$$u(x,y) = \begin{cases} 1, & y < \frac{b}{a}x \\ 0, & y \ge \frac{b}{a}x \end{cases}$$

This problem can be interpreted as a two-dimensional radiative transfer model in transparent medium, see [13]. We plot the contours of the numerical solution solved from the scheme (3.3) *with positivity-preserving limiter for* $k = 1, 2, 3, 4$ *on* 40×40 *rectangular mesh in Figure 3, where white dots are drawn on the cells with negative values appearing before the limiting process. Moreover, we cut the profile of the solutions along the line* $x = 0.5$, *and compare them with the exact solution and the numerical solution solved without limiter in Figure 4, from which we can clearly see that the scheme without limiter produces negative values while the positivity of the solution is maintained with the limiter.*

Example 5.7. *We solve the equation* (1.3) *with* $a = 0.6, b = 0.4, \lambda = 1$ *and* $f = 0$ *on the domain* $\Omega = [0,1]^2$. *The boundary condition is given by* $u(x,0) = 1$ *for* $0 < x \le 1$ *and* $u(0,y) = 0$ *for* $0 \le y \le 1$. *The exact solution of the problem is*

$$u(x,y) = \begin{cases} e^{-\frac{\lambda}{b}y}, & y < \frac{b}{a}x \\ 0, & y \ge \frac{b}{a}x \end{cases}$$

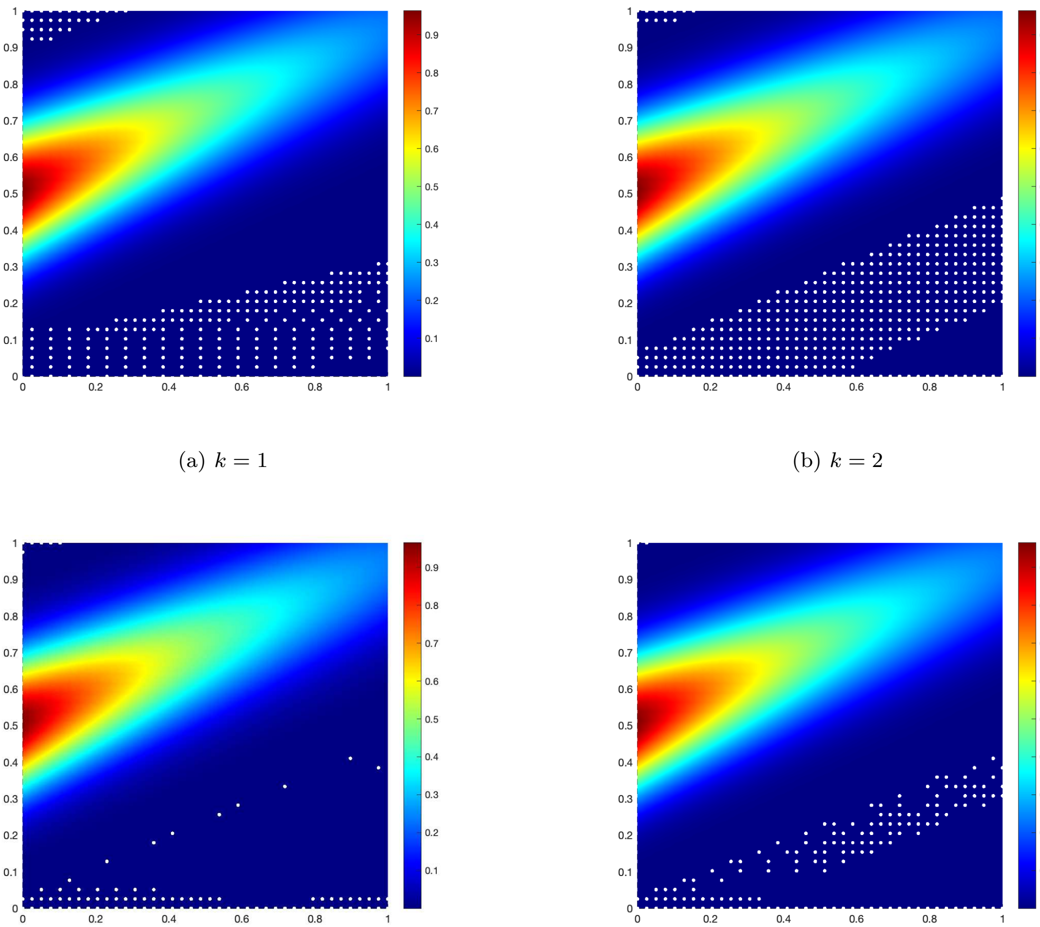
The problem can be viewed as a two-dimensional radiative transfer model in purely absorbing medium, see [13]. We plot the contour of the numerical solution solved from the scheme (3.3) *with positivity-preserving*

k	$N_x \times N_y$	L^1 error	order	L^∞ error	order	$\min u_h$
1	10×10	1.87E-02	-	2.96E-01	-	-1.01E-01
	20×20	6.04E-03	1.63	1.08E-01	1.46	-6.43E-03
	40×40	2.45E-03	1.30	4.80E-02	1.16	-2.42E-04
	80×80	1.14E-03	1.11	2.36E-02	1.02	-4.50E-06
	160×160	5.56E-04	1.03	1.18E-02	1.00	-7.34E-08
	320×320	2.77E-04	1.01	5.89E-03	0.99	-1.16E-09
2	10×10	1.70E-03	-	3.97E-02	-	-8.78E-03
	20×20	3.77E-04	2.17	1.27E-02	1.65	-2.37E-03
	40×40	9.11E-05	2.05	3.48E-03	1.86	-1.70E-04
	80×80	2.27E-05	2.01	9.17E-04	1.93	-2.53E-06
	160×160	5.68E-06	2.00	2.35E-04	1.97	-1.14E-07
	320×320	1.42E-06	2.00	5.95E-05	1.98	-2.89E-09
3	10×10	1.45E-04	-	4.56E-03	-	-6.59E-04
	20×20	1.49E-05	3.29	4.76E-04	3.26	-4.35E-05
	40×40	1.70E-06	3.13	6.13E-05	2.96	-7.75E-07
	80×80	2.06E-07	3.04	7.78E-06	2.98	-1.22E-08
	160×160	2.56E-08	3.01	9.80E-07	2.99	-1.92E-10
	320×320	3.19E-09	3.00	1.23E-07	2.99	-3.01E-12
4	10×10	1.23E-05	-	4.28E-04	-	-7.15E-05
	20×20	6.95E-07	4.14	3.63E-05	3.56	-2.49E-06
	40×40	4.13E-08	4.07	2.55E-06	3.83	-4.55E-08
	80×80	2.53E-09	4.03	1.68E-07	3.93	-7.49E-10
	160×160	1.57E-10	4.01	1.07E-08	3.97	-1.19E-11
	320×320	9.77E-12	4.00	6.77E-10	3.99	-1.87E-13

Table 12: Results of Example 5.5 without limiter

k	$N_x \times N_y$	L^1 error	order	L^∞ error	order	Limited cells (%)
1	10×10	1.98E-02	-	3.17E-01	-	36.00
	20×20	6.13E-03	1.69	1.08E-01	1.56	21.25
	40×40	2.45E-03	1.32	4.80E-02	1.16	14.69
	80×80	1.14E-03	1.11	2.36E-02	1.02	5.27
	160×160	5.56E-04	1.03	1.18E-02	1.00	1.21
	320×320	2.77E-04	1.01	5.89E-03	0.99	0.23
2	10×10	2.34E-03	-	3.92E-02	-	49.00
	20×20	3.91E-04	2.58	1.27E-02	1.63	37.25
	40×40	9.12E-05	2.10	3.48E-03	1.86	25.50
	80×80	2.27E-05	2.01	9.17E-04	1.93	13.34
	160×160	5.68E-06	2.00	2.35E-04	1.97	6.25
	320×320	1.42E-06	2.00	5.95E-05	1.98	3.43
3	10×10	2.69E-04	-	5.18E-03	-	27.00
	20×20	1.80E-05	3.90	5.01E-04	3.37	13.25
	40×40	1.72E-06	3.39	6.13E-05	3.03	5.81
	80×80	2.06E-07	3.06	7.78E-06	2.98	3.97
	160×160	2.56E-08	3.01	9.80E-07	2.99	2.95
	320×320	3.19E-09	3.00	1.23E-07	2.99	2.45
4	10×10	3.29E-05	-	8.94E-04	-	29.00
	20×20	1.01E-06	5.03	3.81E-05	4.55	14.00
	40×40	4.37E-08	4.53	2.55E-06	3.90	9.31
	80×80	2.55E-09	4.10	1.68E-07	3.93	4.20
	160×160	1.57E-10	4.02	1.07E-08	3.97	2.28
	320×320	9.77E-12	4.01	6.77E-10	3.99	1.67

Table 13: Results of Example 5.5 with limiter



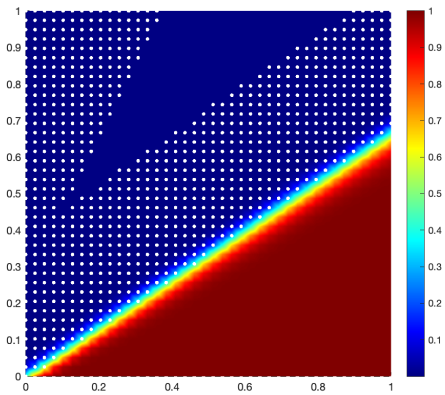
(a) $k = 1$

(b) $k = 2$

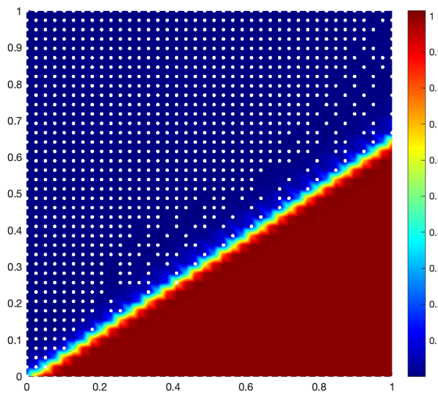
(c) $k = 3$

(d) $k = 4$

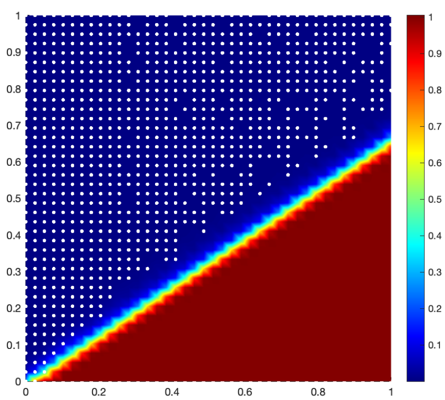
Figure 2: Solutions of Example 5.5 with limiter



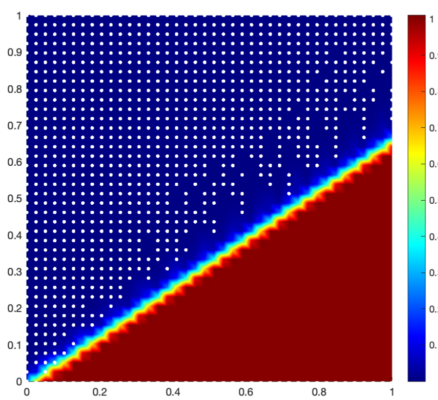
(a) $k = 1$



(b) $k = 2$

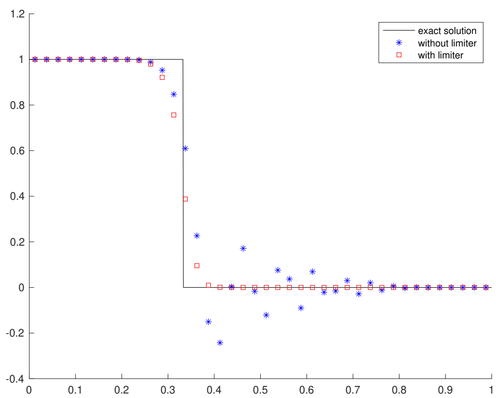


(c) $k = 3$

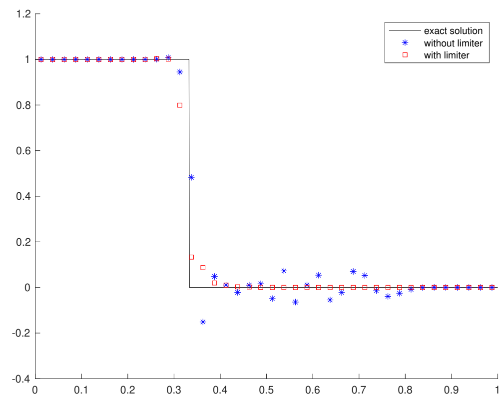


(d) $k = 4$

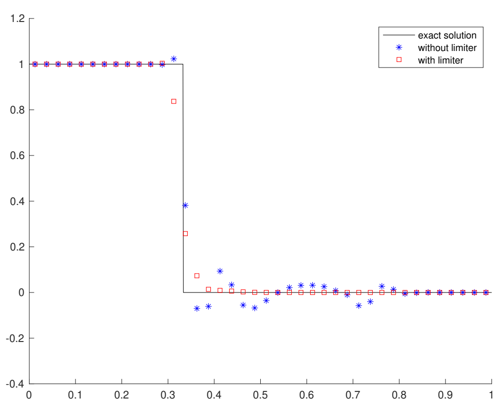
Figure 3: Solutions of Example 5.6 with limiter



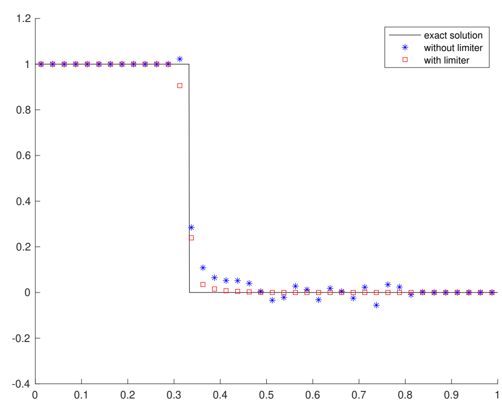
(a) $k = 1$



(b) $k = 2$



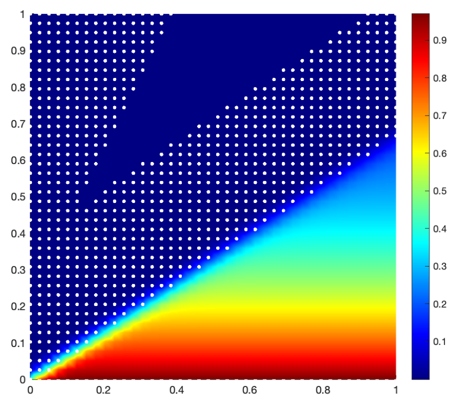
(c) $k = 3$



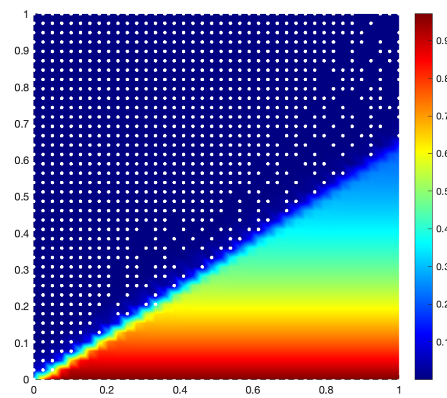
(d) $k = 4$

Figure 4: Solutions of Example 5.6 cut along $x = 0.5$

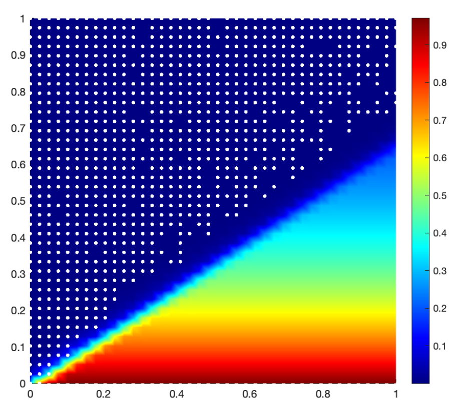
limiter for $k = 1, 2, 3, 4$ on 40×40 rectangular mesh in Figure 5, where white dots are drawn on the cells with negative values appearing before the limiting process. Moreover, we cut the profile of the solution along the line $x = 0.5$, and compare them with the exact solutions and the numerical solutions solved without limiter in Figure 6, from which we can see the positivity of solution is attained under the positivity-preserving limiter.



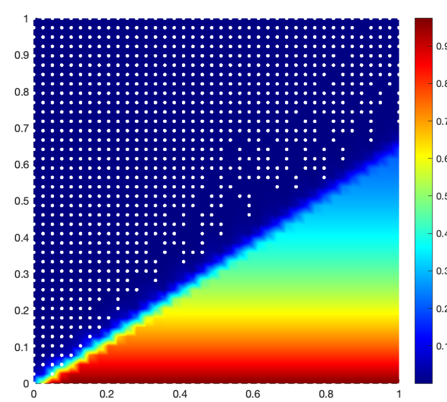
(a) $k = 1$



(b) $k = 2$



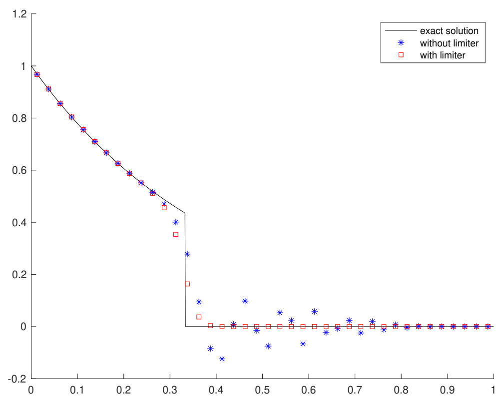
(c) $k = 3$



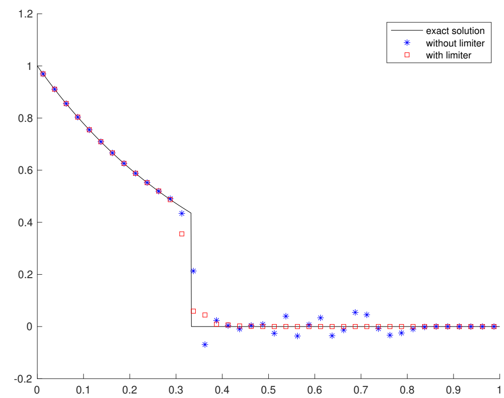
(d) $k = 4$

Figure 5: Solutions of Example 5.7 with limiter

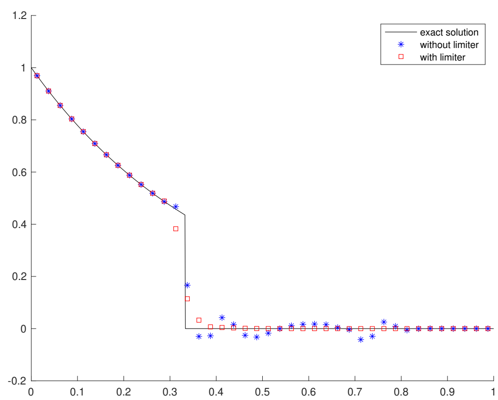
Example 5.8. *We consider the time-dependent linear problem $u_t + u_x = 0$ on the domain $\Omega = [0, 2]$ with*



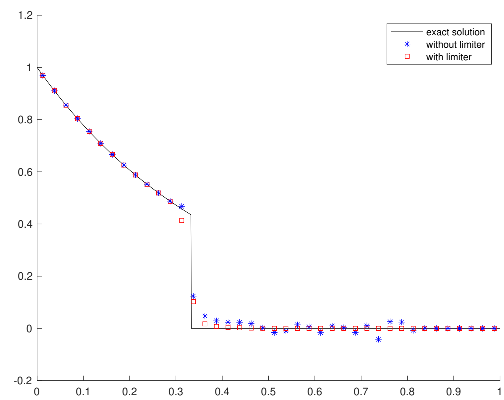
(a) $k = 1$



(b) $k = 2$



(c) $k = 3$



(d) $k = 4$

Figure 6: Solutions of Example 5.7 cut along $x = 0.5$

boundary condition $u(0) = 0$ *and discontinuous initial condition*

$$u_0(x) = \begin{cases} 1, & x \in [\frac{1}{4}, \frac{3}{4}] \\ 0, & \textit{otherwise.} \end{cases}$$

The solution of the problem is $u(x,t) = u_0(x-t)$. *We use the space-time DG approach that treats the time as an extra dimension, and solve the problem based on the scheme* (1.3). *The mesh is* 80×40 *on the space-time domain* $\Omega \times [0,T]$. *We plot the numerical solutions at* $t = 1$ *and compare it with the exact solution and the solution solved without limiter in Figure 7. From the figures, we can see the solutions have negative values without the positivity-preserving limiter, while the positivity is maintained after the limiting process.*

Example 5.9. *We consider the time-dependent linear problem* $u_t + au_x + bu_y + \lambda u = f$ *with* $a = 0.7, b = 0.3, \lambda = 0.5$ *on the domain* $\Omega = [0,1]^2$. *The initial condition is*

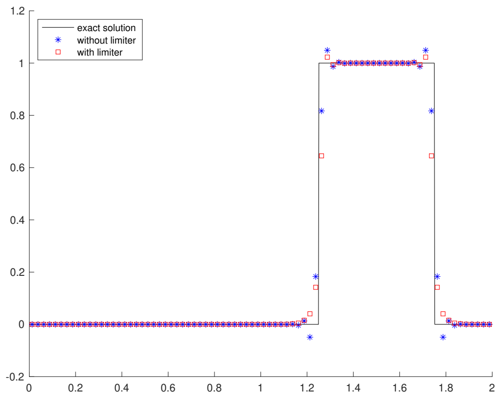
$$u_0(x,y) = \begin{cases} 0, & y < \frac{b}{a}x, \\ \sin^6(\pi(y - \frac{b}{a}x))e^{-2\lambda x}, & y \geq \frac{b}{a}x. \end{cases}$$

The exact solution of the problem is $u(x,y,t) = u_0(x - at, y - bt)e^{-\lambda t}$. *The boundary conditions are given according to the exact solution on the inflow boundaries. We use the space-time DG approach that treats the time as an extra dimension, and solve the problem based on the scheme* (1.4) *on the space-time domain* $\Omega \times T$ *with* $T = 0.5$. *The errors, orders of convergence and data about positivity on the whole space-time domain* $\Omega \times [0,T]$ *are given in Tables 14 and 15 for the cases without and with limiter, respectively, from which we can observe sub-optimal convergence and the positivity of solution being maintained by the positivity-preserving limiter.*

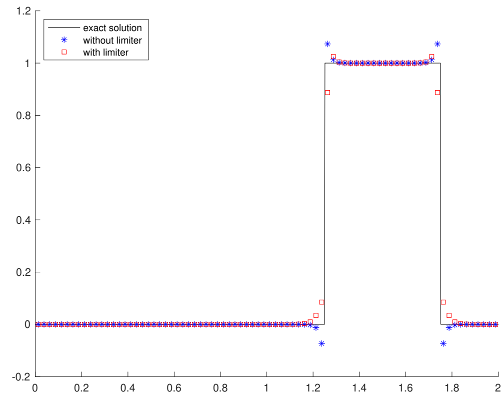
6 Concluding remarks

In this paper, we have constructed the high order conservative positivity-preserving DG method for stationary hyperbolic equations, via suitable quadrature rules in the DG framework.

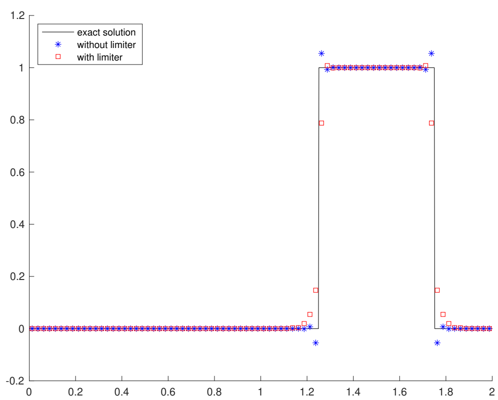
In one space dimension, we propose the conservative positivity-preserving scheme with arbitrary high order for the variable coefficient equation (1.1) with $\lambda = 0$, and second and third orders for the variable coefficient equation (1.1) with $\lambda > 0$ and nonlinear equation (1.2) with $\lambda \geq 0$, which is a vast extension of the previous works in [13, 19] since only constant coefficient equations were addressed therein.



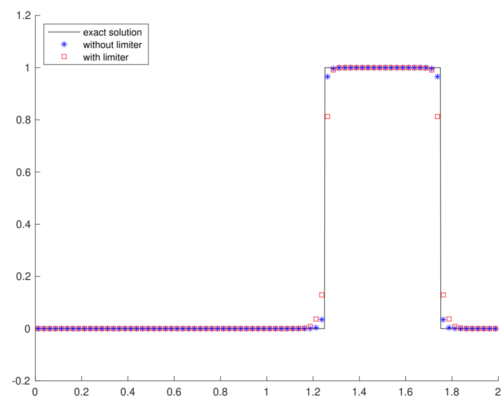
(a) $k = 1$



(b) $k = 2$



(c) $k = 3$



(d) $k = 4$

Figure 7: Solutions of Example 5.8 at $T = 1$

k	$N_x \times N_y \times N_t$	L^1 error	order	L^∞ error	order	$\min u_h$
1	$10 \times 10 \times 5$	1.17E-02	-	8.74E-01	-	-4.52E-01
	$20 \times 20 \times 10$	4.80E-03	1.29	3.61E-01	1.27	-8.95E-02
	$40 \times 40 \times 20$	2.27E-03	1.08	1.87E-01	0.95	-7.47E-03
	$80 \times 80 \times 40$	1.12E-03	1.02	9.51E-02	0.98	-2.51E-04
	$160 \times 160 \times 80$	5.55E-04	1.01	4.80E-02	0.99	-2.10E-05
	$320 \times 320 \times 160$	2.77E-04	1.00	2.41E-02	0.99	-3.49E-06
3	$10 \times 10 \times 5$	1.19E-04	-	1.55E-02	-	-2.63E-03
	$20 \times 20 \times 10$	1.29E-05	3.20	1.82E-03	3.09	-1.40E-04
	$40 \times 40 \times 20$	1.52E-06	3.09	2.27E-04	3.00	-2.63E-06
	$80 \times 80 \times 40$	1.86E-07	3.03	2.89E-05	2.98	-4.65E-08
	$160 \times 160 \times 80$	2.31E-08	3.01	3.66E-06	2.98	-7.58E-10
	$320 \times 320 \times 160$	2.89E-09	3.00	4.61E-07	2.99	-3.68E-11
5	$10 \times 10 \times 5$	8.73E-07	-	1.43E-04	-	-5.46E-05
	$20 \times 20 \times 10$	2.47E-08	5.14	3.71E-06	5.27	-5.46E-07
	$40 \times 40 \times 20$	7.35E-10	5.07	1.14E-07	5.02	-1.11E-08
	$80 \times 80 \times 40$	2.25E-11	5.03	3.63E-09	4.98	-2.02E-10
	$160 \times 160 \times 80$	7.01E-13	5.01	1.15E-10	4.98	-3.37E-12
	$320 \times 320 \times 160$	2.20E-14	4.99	3.62E-12	4.99	-5.44E-14

Table 14: Results of Example 5.9 without limiter

k	$N_x \times N_y$	L^1 error	order	L^∞ error	order	Limited cells (%)
1	$10 \times 10 \times 5$	1.28E-02	-	7.43E-01	-	78.20
	$20 \times 20 \times 10$	5.11E-03	1.33	4.81E-01	0.63	58.20
	$40 \times 40 \times 20$	2.29E-03	1.16	1.87E-01	1.36	40.61
	$80 \times 80 \times 40$	1.12E-03	1.03	9.51E-02	0.98	29.66
	$160 \times 160 \times 80$	5.55E-04	1.01	4.80E-02	0.99	21.89
	$320 \times 320 \times 160$	2.77E-04	1.00	2.41E-02	0.99	17.11
3	$10 \times 10 \times 5$	8.84E-04	-	1.02E-01	-	48.40
	$20 \times 20 \times 10$	5.36E-05	4.04	1.85E-02	2.46	34.55
	$40 \times 40 \times 20$	2.18E-06	4.62	8.90E-04	4.38	25.93
	$80 \times 80 \times 40$	1.92E-07	3.51	2.89E-05	4.94	22.40
	$160 \times 160 \times 80$	2.32E-08	3.05	3.66E-06	2.98	19.09
	$320 \times 320 \times 160$	2.89E-09	3.01	4.61E-07	2.99	16.19
5	$10 \times 10 \times 5$	3.35E-04	-	9.43E-02	-	42.40
	$20 \times 20 \times 10$	3.02E-05	3.47	3.03E-02	1.64	33.68
	$40 \times 40 \times 20$	1.01E-06	4.90	1.55E-03	4.29	27.34
	$80 \times 80 \times 40$	1.28E-08	6.30	3.83E-05	5.34	22.33
	$160 \times 160 \times 80$	1.24E-10	6.69	6.96E-07	5.78	18.93
	$320 \times 320 \times 160$	1.40E-12	6.47	1.15E-08	5.92	16.36

Table 15: Results of Example 5.9 with limiter

We also propose the conservative positivity-preserving scheme for constant coefficient equations with arbitrary high order in two space dimensions, and arbitrary odd order in three space dimensions, which improves the existing results in [13, 19] that are either non-conservative with high order accuracy or conservative with second order accuracy. We only give rigorous proofs for limited cases but the results of numerical experiments in Section 3 for general cases are very promising.

Finally, we would like to mention that, even though we have not discussed it in this paper, one important application of the positivity-preserving schemes for stationary hyperbolic equations is to radiative transfer equations. One can refer to [13, 19, 20] for details.

References

[1] Y. Brenier and S. Osher, The discrete one-sided Lipschitz condition for convex scalar conservation laws, SIAM Journal on Numerical Analysis, 25, 1988, 8-23.

[2] R.L. Burden, J.D. Faires and A.M. Burden, Numerical Analysis, Cengage Learning, 2015.

[3] T. Chen and C.-W. Shu, Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws, Journal of Computational Physics, 345, 2017, 427-461.

[4] B. Cockburn, S. Hou and C.-W. Shu, The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: the multidimensional case, Mathematics of Computation, 54, 1990, 545-581.

[5] B. Cockburn, S.-Y. Lin and C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems, Journal of Computational Physics, 84, 1989, 90-113.

[6] B. Cockburn and C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework, Mathematics of computation, 52, 1989, 411-435.

[7] B. Cockburn and C.-W. Shu, The Runge-Kutta local projection-discontinuous-Galerkin finite element method for scalar conservation laws, ESAIM: Mathematical Modelling and Numerical Analysis, 25, 1991, 337-361.

[8] B. Cockburn and C.-W. Shu, The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems, Journal of Computational Physics, 141, 1998, 199-224.

[9] W.A. Fiveland, Discrete-ordinates solutions of the radiative transport equation for rectangular enclosures, 1984.

[10] H. Guo and Y. Yang, Bound-preserving discontinuous Galerkin method for compressible miscible displacement in porous media, SIAM Journal on Scientific Computing, 39, 2017, A1969-A1990.

[11] K.D. Lathrop and B.G. Carlson, Discrete ordinates angular quadrature of the neutron transport equation, Los Alamos Scientific Laboratory, 1964.

[12] M. Li, Y. Cheng, J. Shen and X. Zhang, A bound-preserving high order scheme for variable density incompressible Navier-Stokes equations, Journal of Computational Physics, 425, 2021, 109906.

[13] D. Ling, J. Cheng and C.-W. Shu, Conservative high order positivity-preserving discontinuous Galerkin methods for linear hyperbolic and radiative transfer equations, Journal of Scientific Computing, 77, 2018, 1801-1831.

[14] W.H. Reed and T.R. Hill, Triangular mesh methods for the neutron transport equation, Los Alamos Scientific Laboratory, 1973.

[15] T. Qin and C.-W. Shu, Implicit positivity-preserving high-order discontinuous Galerkin methods for conservation laws, SIAM Journal on Scientific Computing, 40, 2018, A81-A107.

[16] Y. Xing and X. Zhang, Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes, Journal of Scientific Computing, 57, 2013, 19-41.

[17] Y. Xing, X. Zhang and C.-W. Shu, Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations. Advances in Water Resources, 33, 2010, 1476-1493.

[18] T. Xiong, J.-M. Qiu and Z. Xu, High order maximum-principle-preserving discontinuous Galerkin method for convection-diffusion equations, SIAM Journal on Scientific Computing, 37, 2015, A583-A608.

[19] D. Yuan, J. Cheng and C.-W. Shu, High order positivity-preserving discontinuous Galerkin methods for radiative transfer equations, SIAM Journal on Scientific Computing, 38, 2016, A2987-A3019.

[20] M. Zhang, J. Cheng and J. Qiu, High order positivity-preserving discontinuous Galerkin schemes for radiative transfer equations on triangular meshes, Journal of Computational Physics, 397, 2019, 108811.

[21] X. Zhang and C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, Journal of Computational Physics, 229, 2010, 3091-3120.

[22] X. Zhang and C.-W. Shu, On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes, Journal of Computational Physics, 229, 2010, 8918-8934.

[23] X. Zhang and C.-W. Shu, Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms, Journal of Computational Physics, 230, 2011, 1238-1248.

[24] X. Zhang, On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations, Journal of Computational Physics, 328, 2017, 301-343.

[25] Y. Zhang, X. Zhang and 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.

Appendices

A The positivity of solution at downwind points in one space dimension

In this appendix, we prove that the solutions of schemes proposed in Section 2 are nonnegative at downwind points under certain mesh size conditions, provided the positivity of the boundary condition and source term.

Theorem A.1. *For the problem* (1.1) *with* $a(x) > 0$ *and* $f, u(0) \geq 0$*, the solution of the scheme* (2.3) *satisfies* $u^-_{j+\frac{1}{2}} \geq 0, j = 1, 2, \ldots, N$*, for* $k = 1, 2, 3, \ldots$ *if* $\lambda = 0$*, and for* $k = 1, 2$ *if* $\lambda \Delta x_j \leq 2 \min_{x \in I_j} a(x), j = 1, 2, \ldots, N$.

Proof. If $\lambda = 0$, we take the test function $w = 1$ in the scheme (2.3) to yield the equations

$$a(x_{j+\frac{1}{2}})u^-_{j+\frac{1}{2}} = a(x_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}} + \int_{I_j} f dx, j = 1, 2, \ldots, N,$$

satisfied by the solution. Since $a(x) > 0, f(x) \geq 0$ on Ω, and $u^-_{\frac{1}{2}} = u(0) \geq 0$, by induction, we have $u^-_{j+\frac{1}{2}} \geq 0, j = 1, 2, \ldots, N$.

If $\lambda > 0$ and $k = 1$, it is easy to check that the test function $\xi(x) = \frac{1}{a(x_{j+\frac{1}{2}})} - \frac{2\lambda(x_{j+\frac{1}{2}}-x)}{a(x_{j+\frac{1}{2}})(2a(x_j)+\lambda\Delta x_j)} \in P^1(I_j)$ satisfies

$$-\int_{I_j} (a(x)v\xi_x - \lambda v\xi)\, dx + a(x_{j+\frac{1}{2}})v^-_{j+\frac{1}{2}}\xi^-_{j+\frac{1}{2}} = v^-_{j+\frac{1}{2}}, \tag{A.1}$$

for all $v \in P^1(I_j)$. Moreover, $\xi(x_{j+\frac{1}{2}}) = \frac{1}{a(x_{j+\frac{1}{2}})} > 0$ and $\xi(x_{j-\frac{1}{2}}) = \frac{2a(x_j)-\lambda\Delta x_j}{a(x_{j+\frac{1}{2}})(2a(x_j)+\lambda\Delta x_j)} \geq 0$ if $\lambda\Delta x_j \leq 2\min_{x\in I_j} a(x)$, which implies that $\xi(x) \geq 0$ on I_j. Therefore, by taking the test function $w = \xi$ (extends to zero outside I_j) in the scheme (2.3), we have

$$u^-_{j+\frac{1}{2}} = a(x_{j-\frac{1}{2}})u^-_{j-\frac{1}{2}}\xi^+_{j-\frac{1}{2}} + \int_{I_j} f\xi dx, \tag{A.2}$$

which implies $u^-_{j+\frac{1}{2}} \geq 0$ if $u^-_{j-\frac{1}{2}} \geq 0$. Since $u^-_{\frac{1}{2}} = u(0) \geq 0$, by induction, we have $u^-_{j+\frac{1}{2}} \geq 0$ for $j = 1, 2, \ldots, N$.

If $\lambda > 0$ and $k = 2$, one can check that $\xi(x) = \xi_1 L_1(x) + \xi_2 L_2(x) + \xi_3 L_3(x)$ satisfies the equation (A.1) for all $v \in P^2(I_j)$, where $L_1(x), L_2(x), L_3(x)$ are the Lagrange basis at $\{\hat{x}_1, \hat{x}_2, x_{j+\frac{1}{2}}\}$ with $L_1(\hat{x}_1) = 1, L_2(\hat{x}_2) = 1, L_3(x_{j+\frac{1}{2}}) = 1$, and

$$\begin{aligned}
\xi_1 &= \frac{2\sqrt{3}a(\hat{x}_1)(2\sqrt{3}a(\hat{x}_2) - \lambda\Delta x_j)}{a(x_{j+\frac{1}{2}})\left(12a(\hat{x}_1)a(\hat{x}_2) + 3a(\hat{x}_1)\lambda\Delta x_j + 3a(\hat{x}_2)\lambda\Delta x_j + \lambda^2\Delta x_j^2\right)},\\
\xi_2 &= \frac{2\sqrt{3}a(\hat{x}_2)(2\sqrt{3}a(\hat{x}_1) + \lambda\Delta x_j)}{a(x_{j+\frac{1}{2}})\left(12a(\hat{x}_1)a(\hat{x}_2) + 3a(\hat{x}_1)\lambda\Delta x_j + 3a(\hat{x}_2)\lambda\Delta x_j + \lambda^2\Delta x_j^2\right)},\\
\xi_3 &= \frac{1}{a(x_{j+\frac{1}{2}})}.
\end{aligned}$$

Moreover, if $\lambda\Delta x_j \leq 2\min_{x\in I_j} a(x)$, we have $\xi(\hat{x}_1) = \xi_1 \geq 0, \xi(\hat{x}_2) = \xi_2 \geq 0$ and

$$\xi(x_{j-\frac{1}{2}}) = \frac{12a(\hat{x}_1)a(\hat{x}_2) - 3a(\hat{x}_1)\lambda\Delta x_j - 3a(\hat{x}_2)\lambda\Delta x_j + \lambda^2\Delta x_j^2}{a(x_{j+\frac{1}{2}})\left(12a(\hat{x}_1)a(\hat{x}_2) + 3a(\hat{x}_1)\lambda\Delta x_j + 3a(\hat{x}_2)\lambda\Delta x_j + \lambda^2\Delta x_j^2\right)} \geq 0.$$

Therefore, follow the same lines as in the case $k = 1$, we obtain $u^-_{j+\frac{1}{2}} \geq 0$ for $j = 1, 2, \ldots, N$. □

Almost the same arguments can be used to prove a similar theorem for the scheme (2.7) with $k = 1$ and scheme (2.8) with $k = 2$, except that the positivity of ξ at the midpoint need to be checked due to the quadrature rules adopted on the right hand side the schemes. The theorem is stated as follows and the proof is omitted.

Theorem A.2. *For the problem* (1.2) *with* $a(u) \geq c > 0$, *and* $f, u(0), \lambda \geq 0$, *the solutions of the scheme* (2.7) *with* $k = 1$ *and scheme* (2.8) *with* $k = 2$ *satisfy* $u^-_{j+\frac{1}{2}} \geq 0, j = 1, 2, \ldots, N$ *if* $\lambda\Delta x_j \leq 2c, j = 1, 2, \ldots, N$.

Remark A.1. *For the time-dependent linear problem* $u_t+(a(x)u)_x=0$*, the backward Euler time discretization approach yields the stationary equations* $(a(x)u^n)_x+\Delta t^{-1}u^n=\Delta t^{-1}u^{n-1}, n=1,2,\ldots,\frac{T}{\Delta t}$*. Therefore, the backward Euler discontinuous Galerkin scheme* (2.3) *is positivity-preserving under the CFL condition* $\min_{x\in I_j} a(x)\frac{\Delta t}{\Delta x_j}\geq\frac{1}{2},\forall j$*, if the positivity-preserving limiter is not applied until the computation of* u^n *is completed at the time level* n*. This result can be viewed as an extension of the theoretical result of positivity-preserving backward Euler discontinuous Galerkin method for* $u_t+u_x=0$ *analyzed in [15].*

B Investigation of the schemes (2.2) and (2.3) for some special $a(x)$

The unmodulated P^k-DG schemes (2.2) for the equation (1.1) could result in negative cell averages in the solution for some special $a(x)$. For instance, one can take $a(x)=1+x, a(x)=1+x^2, a(x)=1+x^3, a(x)=1+x^4, a(x)=1+x^5$ in the unmodulated P^1,P^2,P^3,P^4,P^5-DG schemes, respectively, for some particular λ.

More precisely, for the test function $\xi\in P^k([0,h])$, s.t.

$$-\int_0^h (a(x)v\xi_x-\lambda v\xi)\,dx+a(h)v(h)\xi(h)=\frac{1}{h}\int_0^h vdx, \quad \forall v\in P^k([0,h]),$$

where $a(x)=1+x^k$, $k=1,2,3,4,5$, $\xi(h)$ is strictly negative for sufficiently small h, though $\lim_{h\to 0}\xi(hx)=1-x\geq 0$ for $x\in[0,1]$.

One can check that, if $\lambda=0$:

- For $k=1$, $\lim_{h\to 0}\frac{\xi(h)}{h}=-\frac{1}{6}$.
- For $k=2$, $\lim_{h\to 0}\frac{\xi(h)}{h^2}=-\frac{1}{30}$.
- For $k=3$, $\lim_{h\to 0}\frac{\xi(h)}{h^3}=-\frac{1}{140}$.
- For $k=4$, $\lim_{h\to 0}\frac{\xi(h)}{h^4}=-\frac{1}{630}$.
- For $k=5$, $\lim_{h\to 0}\frac{\xi(h)}{h^5}=-\frac{1}{2772}$.

One can also check that, if $\lambda=\frac{1}{2}$:

- For $k=1$, $\lim_{h\to 0}\frac{\xi(h)}{h}=-\frac{1}{12}$.
- For $k=2$, $\lim_{h\to 0}\frac{\xi(h)}{h^2}=-\frac{7}{240}$.

- For $k=3$, $\lim_{h\to 0}\frac{\xi(h)}{h^3} = -\frac{47}{6720}$.
- For $k=4$, $\lim_{h\to 0}\frac{\xi(h)}{h^4} = -\frac{383}{241920}$.
- For $k=5$, $\lim_{h\to 0}\frac{\xi(h)}{h^5} = -\frac{349}{967680}$.

Therefore, we can construct proper source term $f(x) \geq 0$ with large values around $x = h$, such that the average of the solution on the cell $[0, h]$ is negative.

However, using the positivity-preserving scheme defined in (2.3), the above problems are resolved. One can check that, for the test function $\xi \in P^k([0,h])$, s.t.

$$-\int_0^h (a(x)v\xi_x - \lambda v\xi)\, dx + a(h)v(h)\xi(h) = \frac{1}{h}\int_0^h v dx, \quad \forall v \in P^k([0,h]),$$

where $a(x) = 1 + x^k$, $k = 1, 2, 3, 4, 5$, we still have $\lim_{h\to 0}\xi(hx) = 1 - x$ but now $\xi(h) = 0$ in all those cases.