Journal of Computational Physics 391 (2019) 14–36

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Third order maximum-principle-satisfying DG schemes for convection-diffusion problems with anisotropic diffusivity

Hui Yu [a,*], Hailiang Liu [b]

[a] Tsinghua University, Yau Mathematical Sciences Center, Beijing, 100084, China
[b] Iowa State University, Mathematics Department, Ames, IA 50011, United States of America

ARTICLE INFO

Article history:
Received 14 December 2018
Received in revised form 10 April 2019
Accepted 13 April 2019
Available online 18 April 2019

Keywords:
Diffusion
Discontinuous Galerkin method
Maximum principle
High order accuracy

ABSTRACT

For a class of convection-diffusion equations with variable diffusivity, we construct third order accurate discontinuous Galerkin (DG) schemes on both one and two dimensional rectangular meshes. The DG method with an explicit time stepping can well be applied to nonlinear convection–diffusion equations. It is shown that under suitable time step restrictions, the scaling limiter proposed in Liu and Yu (2014) [23] when coupled with the present DG schemes preserves the solution bounds indicated by the initial data, i.e., the maximum principle, while maintaining uniform third order accuracy. These schemes can be extended to rectangular meshes in three dimension. The crucial for all model scenarios is that an effective test set can be identified to verify the desired bounds of numerical solutions. This is achieved mainly by taking advantage of the flexible form of the diffusive flux and the adaptable decomposition of weighted cell averages. Numerical results are presented to validate the numerical methods and demonstrate their effectiveness.

1. Introduction

In this paper, we construct and analyze third order maximum-principle-satisfying (MPS) discontinuous Galerkin (DG) schemes for the following problem,

$$\begin{cases} M(x)\partial_t u + \nabla \cdot f(u) = \nabla \cdot (A\nabla u), & x \in \mathbb{R}^d,\ t > 0, \\ u(0,x) = u_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.1}$$

This equation can be considered as a model for numerous physical problems. In this equation, $u = u(t,x)$ is a time-dependent unknown scalar function, $f(u)$ is the smooth vector flux, and d is the spatial dimension. We assume that the diffusion tensor $A = A(x,u)$ is a symmetric and nonnegative definite matrix, and $M(x)$ is strictly positive scalar function. In this paper, we present our results and analysis for $d = 1, 2$, extension to three dimension can be done as well.

Our concern for (1.1) arises from several typical scenarios. The first case with $f = 0$ is a model for heat conduction in a non-uniform body with $M(x) = \frac{c(x)}{\rho(x)}$, where $c(x)$ is the specific heat and $\rho(x)$ the mass density. The heat flux is proportional to the temperature difference $-A\nabla u$, known as the Fourier law of heat conduction. Here A measures the ability of the material to conduct heat, called the thermal conductivity. In the second case with $M = 1$, we have the usual convection-diffusion equation,

* Corresponding author.
E-mail addresses: huiyu@tsinghua.edu.cn (H. Yu), hliu@iastate.edu (H. Liu).

https://doi.org/10.1016/j.jcp.2019.04.028
0021-9991/

$$\partial_t u + \nabla \cdot f(u) = \nabla \cdot (A(x,u)\nabla u). \tag{1.2}$$

There are many interpretations and derivations from fluid dynamics and other application areas that motivate the convection-diffusion equation, such as Navier-Stokes equations and the porous medium equation. The third case is the Fokker-Planck equation,

$$\partial_t \rho = \nabla \cdot (\nabla \rho + \nabla V(x)\rho), \tag{1.3}$$

which can be rewritten in terms of $u = \rho e^{V(x)}$ as (1.1) with $A(x) = M(x) = e^{-V(x)}$, and $f = 0$. This setting allows for many variants.

The above three types of problems can be formulated under the model class (1.1). One important solution to (1.1) is the one that is bounded by two constants dictated by the initial data, leading to the so-called Maximum Principle (MP). In other words, if

$$c_1 = \min_x u_0, \quad c_2 = \max_x u_0,$$

then $u(t,x) \in [c_1, c_2]$ for any $x \in \mathbb{R}^d$ and $t > 0$.

From the analytical viewpoint, the maximum principle is quite general but very significant due to its physical implications. From the numerical viewpoint, it is widely recognized that maximum principle provides a valuable tool in proving solvability results (existence and uniqueness of discrete solutions), enforcing numerical stability, and deriving convergence results (a priori error estimates) for the sequence of approximate solutions. Design of a high order scheme to preserve the maximum principle is known a challenging task. Our goal is to better understand how a high order DG scheme can be constructed for (1.1) to respect the MPS property. The main difficulty in the anisotropic case with a variable weight $M(x)$ is the derivation of suitable sufficient conditions so that the weighted cell average stays in $[c_1, c_2]$ during the time evolution. Such weighted cell average is essentially used for limiting the numerical solution into $[c_1, c_2]$, without destroying accuracy.

1.1. Related work

Early discussion of the discrete maximum principle for the convection-diffusion equations includes the linear finite element solutions for parabolic equations [8], and recent developments [10,9,30,31,11], as well as [25] by the Petrov–Galerkin finite element method to solve convection dominated problems. However, they are under a different framework.

The present investigation involves the choice of numerical fluxes and monotonicity of weighted cell averages in terms of point values. In the largest sense, the origins of these ideas go all the way back to monotone schemes for hyperbolic conservation laws.

Indeed, for scalar conservation laws, i.e., (1.2) with $A = 0$, many first order classical schemes can be shown to be MPS (other names of this sort include bound-preserving, positivity preserving, or maximum-principle-preserving) since such low order accurate schemes are usually monotone. On the other hand, the Godunov Theorem states that a linear monotone scheme is at most first order accurate for the convection equation [17]. To construct high order accurate MPS schemes for scalar convection, weak monotonicity in finite volume type schemes including DG methods was first used in [35–37]. Here by weak monotonicity it means that each cell average is monotone with respect to point values in that cell, see e.g. [34]. The main idea in their work is to find sufficient conditions to preserve the desired bounds of cell averages by repeated convex combinations. A simple and efficient local MPS limiter can then be used to control the solution values at test points without affecting accuracy and conservation. Together with strong stability preserving (SSP) Runge–Kutta or multistep methods [13], which are convex combinations of several formal forward Euler steps, a high order accurate finite volume or DG scheme can be rendered MPS with the limiter.

For diffusion, a linear finite volume scheme can only be up to second order accurate in order to preserve the weak monotonicity, unless a non-conventional discretization is used in the scheme construction such as that in [37]. For DG methods, in general only second order accuracy can be obtained to feature the MPS property; see [38] for solving (1.2) on triangular meshes.

The only DG method known to satisfy the weak monotonicity up to third order accuracy is the direct DG (DDG) method introduced in [21,22]. Indeed, the special method parameters in the DDG discretization allowed us to design in [23] a third order MPS method for the linear Fokker-Planck equation (1.3). A key idea in [23] is the use of the non-logrithimic Landau formulation

$$M\partial_t u = \nabla \cdot (M\nabla u) \quad \text{with } M(x) = e^{-V(x)} \text{ and } u = \frac{\rho}{M},$$

so that the corresponding maximum principle on $\rho(t,x)$:

$$c_1 e^{-V} \le \rho(0,x) \le c_2 e^{-V} \Longrightarrow c_1 e^{-V} \le \rho(t,x) \le c_2 e^{-V} \quad \forall t > 0,$$

reduces to

$$c_1 \le u(0,x) \le c_2 \Longrightarrow c_1 \le u(t,x) \le c_2 \quad \forall t > 0.$$

With this reformulation, one can show that each weighted cell average is monotone in terms of point values under appropriate CFL conditions. The result in [23] is directly applicable to multi-dimensional diffusion on rectangular meshes. However, it gets subtle to ensure the MPS property on unstructured meshes; we refer to [2] for a third order such DDG method to solve diffusion equations on unstructured triangular meshes.

Another approach towards a positivity-preserving scheme with high order accuracy is to use the local DG (LDG) method [1,5], combined with some special positivity-preserving fluxes. Such an effort was first made in [34] for constructing high order accurate positivity-preserving DG schemes for compressible Navier–Stokes equations. Concerning further developments in this direction, we refer to [28,29] for solving convection–diffusion problems. An MPS third-order LDG method using overlapping meshes has been recently proposed in [7] for convection-diffusion equations.

One noteworthy alternative to enforce positivity in high order schemes is to take a convex combination of high order flux with a first order positivity-preserving one; the method has been applied to various high order schemes including finite difference, finite volume, and DG methods [15,32,33], while rigorous justification of accuracy for such methods seems difficult.

1.2. Present investigation

The main distinction of our present investigation from the above mentioned works is the use of weighted cell averages for both ensuring the weak monotonicity and applying the scaling limiter, with special attention on difficulty caused by the anisotropic diffusivity.

The spatial discretization explored in this work is the DDG method introduced by Liu and Yan in [21,22]. Besides the usual advantages of a DG method (see e.g. [14,26,27]), one main feature of the DDG method lies in numerical flux choices for the solution gradient, which involve second order derivatives evaluated crossing cell interfaces (see (2.2) below). With this choice, the obtained schemes are provably stable and optimally convergent as well as superconvergent for $\beta_1 \neq 0$ [18,3]. Such method has also been successfully extended to various application problems, including Fokker-Planck type equations [23,24,19,20], and the three dimensional compressible Navier–Stokes equation [6].

Built upon the work [23], we present third order DG methods for solving the initial value problem (1.1), with an application to nonlinear convection-diffusion equations of form (1.2). Let us illustrate the main ideas via a simple one-dimensional equation subject to periodic boundary condition:

$$M(x)\partial_t u = \partial_x (A(x)\partial_x u).$$

The computational cell is denoted by $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, where $x_{j+\frac{1}{2}}$'s are the grid points. Let τ and h be the time and space steps for a uniform mesh and n the index for the time stepping. The update on $\langle u_h^n \rangle_j$, the weighted cell average of the numerical approximation u_h^n, is given by

$$\langle u_h^{n+1} \rangle_j = \langle u_h^n \rangle_j - \frac{\tau}{h} \widehat{A\partial_x u_h^n}\Big|_{\partial I_j} \quad \text{where } \langle u \rangle_j := \frac{1}{h}\int_{I_j} M(x)u(x)\,dx.$$

Note that $\widehat{\partial_x u_h}$ is the approximation to $\partial_x u$ at ∂I_j, the cell interface of I_j, given by

$$\widehat{\partial_x u} = \frac{\beta_0}{h}[u] + \{\partial_x u\} + \beta_1 h[\partial_x^2 u].$$

As mentioned above, this form of the diffusive flux was originally introduced in [21,22] as part of the DDG method for solving the diffusion problem. If the flux parameters satisfy

$$\beta_0 \ge 1 \quad \text{and} \quad \frac{1}{8} \le \beta_1 \le \frac{1}{4},$$

then the procedure developed in [23] can be extended to the present setting to conclude that, under a suitable CFL condition, the simple Euler forward will keep the cell average $\bar{u}_j^n = \frac{\langle u_h^n \rangle_j}{\langle 1 \rangle_j} \in [c_1, c_2]$ in each time step, and the validity of the maximum principle when combined with a scaling limiter.

For a two dimensional problem with $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$, the DDG method on shape-regular Cartesian meshes with $\kappa^{-1} \le \frac{\Delta x}{\Delta y} \le \kappa$ can be rendered MPS if

$$\beta_0 \ge 1 + \frac{\kappa |c| L(L-1)}{2\min\{a,b\}} \quad \text{and} \quad \frac{1}{8} \le \beta_1 \le \frac{1}{4}.$$

Here L is the number of Gauss-Lobatto points used in the numerical evaluation of involved integrals. With $f(u)$ and $A = A(x,y,u)$, the MPS DDG schemes are analyzed in Section 3.1 where the parameter range and the CFL conditions are established.

The main conclusion is as follows: by applying the weighted MPS limiter introduced in [23] to the DDG scheme designed here for (1.1), with the time discretization by an SSP Runge-Kutta method (see [12]), we obtain a third order accurate scheme solving (1.1) satisfying the strict maximum principle in the sense that the numerical solution never goes out of the range $[c_1, c_2]$ as indicated by the initial data.

1.3. Organization of the paper

The rest of the paper is organized as follows. In Section 2, we design the numerical method for one dimensional problems. We first formulate the DDG scheme to solve the model heat equation, and prove the MPS property of the third order fully discretized DDG scheme, we then apply the result to show the MPS property for nonlinear convection-diffusion equations. Section 3 is organized similarly for two dimensional problems on Cartesian meshes. In Section 4, we present the MPS limiter, with which the algorithm is complete. In Section 5, numerical tests for the DDG method are reported, including examples from the heat equation, porous media equation the Buckley-Leverett equation, and two dimensional diffusion with the anisotropic diffusion. Concluding remarks are given in Section 6.

2. MPS schemes in one dimension

We first investigate the MPS property for third order DDG schemes to weighted diffusion equations, and show how to apply the scheme obtained to nonlinear convection-diffusion equations.

2.1. The diffusion equation

We begin with the heat equation of the form

$$M(x)\partial_t u = \partial_x(A(x)\partial_x u), \tag{2.1}$$

with $M(x) > 0$ and $A(x) \geq 0$ on the spatial domain Ω, subject to initial data $u_0(x)$ and the periodic boundary condition. It is known that the following maximum principle holds:

$$\text{if } c_1 \leq u_0(x) \leq c_2 \ \ \forall x \in \Omega, \ \text{then } c_1 \leq u(t,x) \leq c_2 \ \ \forall x \in \Omega, t > 0.$$

In general, the problem can be either defined on a connected compact domain with proper boundary conditions, or it can involve the whole real line with solutions vanishing at the infinity. In our numerical scheme, we will always choose the spatial domain to be a connected interval. For simplicity, the periodic boundary conditions are applied.

We partition the domain Ω by regular cells such that $\Omega = \bigcup_{j=1}^{N} I_j$ with $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. Denote $h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$ and $h = \max_j h_j$. We seek numerical solutions in the discontinuous piecewise polynomial space,

$$V_h = \{v \in L^2(\Omega) | \quad v|_{I_j} \in P^k(I_j), \quad j = 1 \cdots, N\}.$$

Here $P^k(I_j)$ is the space of k-th order polynomials on I_j. Note that the functions in V_h can be double-valued at cell interfaces. Hence notations v^- and v^+ are used for the left limit and right limit of v. The jump of these two values, $v^+ - v^-$, is denoted by $[v]$, and the average by $\{v\}$.

Throughout this paper we adopt the DDG numerical flux of the form

$$\widehat{\partial_x v} = \frac{\beta_0}{h_{j+\frac{1}{2}}}[v] + \{\partial_x v\} + \beta_1 h_{j+\frac{1}{2}}[\partial_x^2 v] \quad \text{with } h_{j+\frac{1}{2}} = \frac{h_j + h_{j+1}}{2}, \tag{2.2}$$

when crossing the cell interface $x_{j+\frac{1}{2}}$, and (β_0, β_1) are in the range to be specified so that the underlying third order scheme can weakly satisfy the maximum-principle. The parameter range was first identified in [23] for a third order DDG scheme to feature the MPS property for linear Fokker-Planck equations.

For model equation (2.1), we consider a $(k+1)$th-order DG scheme: to find $u_h \in V_h$ such that for any test function $v \in V_h$,

$$\int_{I_j} M(x)\partial_t u_h v \, dx = -\int_{I_j} A(x)\partial_x u_h \partial_x v \, dx + A\left[\widehat{\partial_x u_h} v + (u_h - \{u_h\})\partial_x v\right]\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}},$$

with the diffusive flux $\widehat{\partial_x u_h}$ as defined in (2.2). This DDG scheme with interface correction was proposed in [22] for the diffusion problem, as an improved version of that in [21]. Based on the DDG scheme in [21], we will have

$$\int_{I_j} M(x)\partial_t u_h v\,dx = -\int_{I_j} A(x)\partial_x u_h \partial_x v\,dx + \widehat{A\partial_x u_h} v\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}.$$

With the forward Euler time discretization, the weighted cell average for either of two DDG schemes evolves as

$$\langle u_h^{n+1}\rangle_j = \langle u_h^n\rangle_j + \mu h\ \widehat{A\partial_x u_h^n}\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}, \tag{2.3}$$

where $\mu = \frac{\tau}{h^2}$ is the mesh ratio. For a concise presentation, a uniform mesh is assumed. Here and in what follows, we denote the time step length as τ. Note that for periodic boundary conditions considered in this paper, we take

$$u^-_{h,\frac{1}{2}} = u^-_{h,N+\frac{1}{2}}, \quad u^+_{h,N+\frac{1}{2}} = u^+_{h,\frac{1}{2}}$$

in the DDG numerical flux formula (2.2) when $j = \frac{1}{2}, N+\frac{1}{2}$.

We also use the notation

$$\langle q(\xi)\rangle_j = \frac{1}{2}\int_{-1}^{1} M(x_j + \frac{h}{2}\xi) q(\xi) d\xi, \text{ for any } q(\xi) \text{ on } [-1,1].$$

And the cell average $\bar{u}_j$ is $\bar{u}_j = \frac{\langle u_h\rangle_j}{\langle 1\rangle_j}$. For $j = 1, \cdots, N$, we define

$$a_j = \frac{\langle \xi - \xi^2\rangle_j}{\langle 1-\xi\rangle_j}, \quad b_j = \frac{\langle \xi + \xi^2\rangle_j}{\langle 1+\xi\rangle_j}, \tag{2.4}$$

and

$$\begin{aligned}
\tilde{\omega}_j^1 &= \frac{\langle \gamma - \xi(1+\gamma) + \xi^2\rangle_j}{2(1+\gamma)},\\
\tilde{\omega}_j^2 &= \frac{\langle 1-\xi^2\rangle_j}{1-\gamma^2},\\
\tilde{\omega}_j^3 &= \frac{\langle -\gamma + \xi(1-\gamma) + \xi^2\rangle_j}{2(1-\gamma)}
\end{aligned} \tag{2.5}$$

with the weight $M(x)|_{I_j} = M(x_j + \frac{h}{2}\xi)$. We recall the following key result.

Lemma 2.1. *[23, Lemma 3.3] $\tilde{\omega}_j^i > 0$ for $i = 1, 2, 3$ if and only if*

$$\gamma \in (a_j, b_j),$$

where a_j, b_j satisfy $-1 < a_j < b_j < 1$.

Proof. Here we only show $a_j < b_j$, which means that selection of γ is always ensured. A direct calculation gives

$$b_j - a_j = 2\frac{\langle 1\rangle_j\langle \xi^2\rangle_j - \langle \xi\rangle_j^2}{\langle 1+\xi\rangle_j\langle 1-\xi\rangle_j} \geq 0,$$

where the numerator can be reformulated as

$$\begin{aligned}
&\int_{-1}^{1}\int_{-1}^{1} (\eta^2 - \xi\eta)\tilde{M}(\xi)\tilde{M}(\eta)d\xi d\eta + \int_{-1}^{1}\int_{-1}^{1} (\xi^2 - \eta\xi)\tilde{M}(\eta)\tilde{M}(\xi)d\eta d\xi\\
&= \int_{-1}^{1}\int_{-1}^{1} (\xi - \eta)^2\tilde{M}(\xi)\tilde{M}(\eta)d\xi d\eta > 0,
\end{aligned}$$

where $\tilde{M}(\xi) := M(x_j + \frac{h}{2}\xi)$ has been used. □

We thus have the following result.

Theorem 2.2. *($k = 2$) The scheme (2.3) with*

$$\beta_0 \geq 1 \quad \textit{and} \quad \frac{1}{8} \leq \beta_1 \leq \frac{1}{4} \tag{2.6}$$

is maximum-principle-satisfying, namely, $c_1 \leq \bar{u}_j^{n+1} \leq c_2$ if $u_h^n(x)$ is in $[c_1, c_2]$ on $\{S_j\}_{j=1}^N$, where

$$S_j = x_j + \frac{h}{2}\{-1, \gamma, 1\}$$

with γ satisfying

$$a_j < \gamma < b_j \quad \textit{and} \quad |\gamma| \leq 8\beta_1 - 1, \tag{2.7}$$

under the CFL condition $\mu \leq \mu_0$, where μ_0 is given in (2.12) below.

Proof. *Step 1.* Weighted integral decomposition. Define

$$p_j(\xi) = u_h\left(x_j + \frac{h}{2}\xi\right) \text{ for } \xi \in [-1, 1],$$

we see that in the case of $p_j(\xi) \in P^2[-1, 1]$, for any $\gamma \in (-1, 1)$, the unique interpolation of p_j at three points $\{-1, \gamma, 1\}$ gives the following

$$p_j(\xi) = \frac{(\xi-1)(\xi-\gamma)}{2(1+\gamma)} p_j(-1) + \frac{(\xi-1)(\xi+1)}{(\gamma-1)(\gamma+1)} p_j(\gamma) + \frac{(\xi+1)(\xi-\gamma)}{2(1-\gamma)} p_j(1). \tag{2.8}$$

This yields the following identity for the weighted average,

$$\langle u_h \rangle_j = \tilde{\omega}_j^1 p_j(-1) + \tilde{\omega}_j^2 p_j(\gamma) + \tilde{\omega}_j^3 p_j(1), \tag{2.9}$$

where $\tilde{\omega}_j^i$ given in (2.5) are ensured positive by Lemma 2.1.

Step 2. Flux representation. A direct calculation gives

$$\begin{aligned} h\,\widehat{\partial_x u_h}\Big|_{x_{j+\frac{1}{2}}} &= \alpha_3(-\gamma) p_{j+1}(-1) + \alpha_2(-\gamma) p_{j+1}(\gamma) + \alpha_1(-\gamma) p_{j+1}(1) \\ &\quad - \left[\alpha_1(\gamma) p_j(-1) + \alpha_2(\gamma) p_j(\gamma) + \alpha_3(\gamma) p_j(1)\right], \end{aligned} \tag{2.10}$$

where

$$\alpha_1(\gamma) = \frac{8\beta_1 - 1 + \gamma}{2(1+\gamma)}, \quad \alpha_2(\gamma) = 2\frac{1 - 4\beta_1}{1-\gamma^2}, \quad \alpha_3(\gamma) = \beta_0 + \frac{8\beta_1 - 3 + \gamma}{2(1-\gamma)},$$

are all positive due to (2.6) and (2.7).

Step 3. Monotonicity under some CFL condition. We now substitute (2.9) and (2.10) into (2.3) to obtain

$$\langle u_h^{n+1} \rangle_j = R_j^n(M(\cdot), \mu, h, A)$$

with

$$\begin{aligned} R_j^n(M(\cdot), \mu, h, A) &= \langle u_h^n \rangle_j + \mu\left(A h \widehat{\partial_x u_h^n}\Big|_{x_{j+\frac{1}{2}}} - A h \widehat{\partial_x u_h^n}\Big|_{x_{j-\frac{1}{2}}}\right) \\ &= \left[\tilde{\omega}_j^1 - \mu\left(\alpha_3(-\gamma) A_{j-\frac{1}{2}} + \alpha_1(\gamma) A_{j+\frac{1}{2}}\right)\right] p_j(-1) \\ &\quad + \left[\tilde{\omega}_j^2 - \mu\left(\alpha_2(-\gamma) A_{j-\frac{1}{2}} + \alpha_2(\gamma) A_{j+\frac{1}{2}}\right)\right] p_j(\gamma) \\ &\quad + \left[\tilde{\omega}_j^3 - \mu\left(\alpha_1(-\gamma) A_{j-\frac{1}{2}} + \alpha_3(\gamma) A_{j+\frac{1}{2}}\right)\right] p_j(1) \\ &\quad + \mu A_{j+\frac{1}{2}}\left[\alpha_3(-\gamma) p_{j+1}(-1) + \alpha_2(-\gamma) p_{j+1}(\gamma) + \alpha_1(-\gamma) p_{j+1}(1)\right] \\ &\quad + \mu A_{j-\frac{1}{2}}\left[\alpha_1(\gamma) p_{j-1}(-1) + \alpha_2(\gamma) p_{j-1}(\gamma) + \alpha_3(\gamma) p_{j-1}(1)\right]. \end{aligned} \tag{2.11}$$

Here it is understood that $p_0(\xi) := p_N(\xi + |\Omega|)$ and $p_{N+1}(\xi + |\Omega|) = p_1(\xi)$ for incorporating the periodic boundary conditions. Note also that $\sum_i^3 \alpha_i(\gamma) = \beta_0 = \sum_i^3 \alpha_i(-\gamma)$. Using the fact $\tilde{\omega}_j^3(\gamma) = \tilde{\omega}_j^1(-\gamma)$ and formulas for $\alpha_2(\gamma)$, $\tilde{\omega}_j^2$, we see that if μ is chosen to be smaller than μ_0 where

$$\mu_0 = \left(\max_{1\le j\le N} A(x_{j+\frac{1}{2}})\right)^{-1} \min_{1\le j\le N}\left\{\frac{\tilde{\omega}_j^1(\pm\gamma)}{\alpha_3(\mp\gamma)+\alpha_1(\pm\gamma)}, \frac{\langle 1-\xi^2\rangle_j}{4(1-4\beta_1)}\right\}, \tag{2.12}$$

(2.11) is nondecreasing in the point values $p_j(\pm1), p_j(\gamma), p_{j\pm1}(\pm1), p_{j\pm1}(\gamma)$, hence when these values are replaced with the lower and upper bounds c_1 and c_2 respectively, we have

$$c_1\sum_{i=1}^3 \tilde{\omega}_j^i \le \langle u_h^{n+1}\rangle_j \le c_2\sum_{i=1}^3 \tilde{\omega}_j^i,$$

since the terms with α_i's are canceled out. Moreover the sum of $\tilde{\omega}_j^i$ is $\langle 1\rangle_j$. Therefore

$$c_1\langle 1\rangle_j \le \langle u_h^{n+1}\rangle_j \le c_2\langle 1\rangle_j \quad\Rightarrow\quad c_1 \le \bar{u}_j^{n+1} \le c_2. \quad \square$$

Remark 2.1. For other types of boundary conditions, the boundary flux needs to be modified. A similar result to Theorem 2.2 may be established as long as the PDE problem satisfies a maximum principle.

Remark 2.2. The use of γ is essential for the success of our schemes, in particular when $M(x)$ is a function. For the special case $M(x)=1$, we have $\gamma\in(-\frac{1}{3},\frac{1}{3})$, hence $\gamma=0$, which corresponds to the usual Gauss quadrature point, is admissible. The CFL number given in (2.12) may be optimized by carefully tuning $\gamma\in(a_j,b_j)$, but not in a linear fashion. Nevertheless, it was observed in [23] that the larger $|\gamma|$ is, the better the scheme's performance for the heat equation.

2.2. Application to nonlinear convection-diffusion equation

In this section we will demonstrate how to apply our MPS DG method in section 2.1 to the nonlinear convection-diffusion equation,

$$\partial_t u + \partial_x f(u) = \partial_x(A(x,u)\partial_x u),$$

where $f(u)$ is a smooth function and diffusion coefficient $A(x,u)\ge 0$, subject to initial data $u(0,x)=u_0(x)$, and periodic boundary conditions. By applying the DG approximation, we obtain the following scheme. We seek $u_h\in V_h$ such that for any test function $v\in V_h$,

$$\begin{aligned}\int_{I_j}\partial_t u_h v\,dx &= \int_{I_j} f(u_h)\partial_x v\,dx - \hat{f}(u_h^-,u_h^+)v\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \\ &\quad - \int_{I_j} A_h\partial_x u_h\partial_x v\,dx + \{A_h\}\left[\widehat{\partial_x u_h}v + (u_h-\{u_h\})\partial_x v\right]\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}.\end{aligned} \tag{2.13}$$

For diffusion part, we adopt the DDG diffusive flux (2.2) For convection, any monotone numerical flux can be used, i.e., $\hat{f}(u,v)$ is Lipschitz continuous, nondecreasing in u and nonincreasing in v, consistent with $f(u)$ in the sense that $\hat{f}(u,u)=f(u)$. For example, the global Lax-Friedrichs flux

$$\hat{f}(u_h^-,u_h^+) = \frac{1}{2}(f(u_h^-)+f(u_h^+)-\sigma(u_h^+-u_h^-)), \quad \sigma = \max_{u\in[c_1,c_2]}|f'(u)|.$$

We consider the first order Euler forward temporal discretization of (2.13) to obtain

$$\begin{aligned}\int_{I_j}\frac{u_h^{n+1}-u_h^n}{\tau}v\,dx &= \int_{I_j} f(u_h^n)\partial_x v\,dx - \hat{f}((u_h^n)^-,(u_h^n)^+)v\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \\ &\quad - \int_{I_j} A_h^n\partial_x u_h^n\partial_x v\,dx + \{A_h^n\}\left[\widehat{\partial_x u_h^n}v + (u_h^n-\{u_h^n\})\partial_x v\right]\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}.\end{aligned} \tag{2.14}$$

By taking the test function $v=1$ on I_j and 0 elsewhere, we obtain the evolutionary update for the cell average,

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \lambda\hat{f}^n\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} + \mu h\{A_h^n\}\widehat{\partial_x u_h^n}\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}},$$

where $\lambda=\frac{\tau}{h}$ and $\mu=\frac{\tau}{h^2}$ are the mesh ratios. Assuming that $\bar{u}_j^n\in[c_1,c_2]$ for all j's, we would like to derive some sufficient conditions such that $\bar{u}_j^{n+1}\in[c_1,c_2]$ under certain restrictions on λ and μ.

For piecewise quadratic polynomials, the main result can be stated as follows.

Theorem 2.3. *($k = 2$) The scheme (2.14) with*

$$\beta_0 \geq 1 \quad \text{and} \quad \frac{1}{8} \leq \beta_1 \leq \frac{1}{4}$$

is maximum-principle-satisfying; namely, $\bar{u}_j^{n+1} \in [c_1, c_2]$ if $u_h^n(x) \in [c_1, c_2]$ on the set S_j's where

$$S_j = x_j + \frac{h}{2}\{-1, \gamma, 1\}$$

with γ satisfying

$$-\frac{1}{3} < \gamma < \frac{1}{3} \quad \text{and} \quad |\gamma| \leq 8\beta_1 - 1, \tag{2.15}$$

under the CFL condition

$$\lambda \leq \lambda_0, \quad \mu \leq \mu_0$$

for some λ_0 and μ_0 defined in (2.17) and (2.18), respectively.

Proof. We present the proof in four steps:

Step 1. Split: we split the average $\bar{u}_j^n$ into two halves so that

$$\bar{u}_j^{n+1} = \frac{1}{2}\mathcal{C}_j^n + \frac{1}{2}\mathcal{D}_j^n,$$

where the convection term is

$$\mathcal{C}_j = \bar{u}_j - 2\lambda \hat{f}\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \tag{2.16}$$

and the diffusion term is

$$\mathcal{D}_j = \bar{u}_j + 2\mu h\{A_h^n\}\widehat{\partial_x u_h}\Big|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}.$$

This split is for convenient presentation and may not lead to an optimal CFL condition.

Step 2. The integral decomposition. From (2.9) it follows

$$\bar{u}_j = \omega^1 p_j(-1) + \omega^2 p_j(\gamma) + \omega^3 p_j(1),$$

where $\omega^i = \tilde{\omega}^i$, $\langle u_h \rangle_j = \bar{u}_j$ for $M(x) \equiv 1$, and

$$\omega^1 = \frac{1+3\gamma}{6(1+\gamma)}, \quad \omega^2 = \frac{2}{3(1-\gamma^2)}, \quad \omega^3 = \frac{1-3\gamma}{6(1-\gamma)}.$$

These coefficients are positive for γ satisfying (2.15).

Step 3. The convection term.

Using the cell average decomposition and the flux formula, we rewrite (2.16) as

$$\begin{aligned}
\mathcal{C}_j &= \bar{u}_j^n - 2\lambda\left(\widehat{f}\big|_{x_{j+\frac{1}{2}}} - \widehat{f}\big|_{x_{j-\frac{1}{2}}}\right) \\
&= \omega^3 p_j(1) - 2\lambda\hat{f}(p_j(1), p_{j+1}(-1)) + \omega^2 p_j(\gamma) + \omega^1 p_j(-1) + 2\lambda\hat{f}(p_{j-1}(1), p_j(-1)) \\
&=: G(p_{j-1}(1), p_j(-1), p_j(\gamma), p_j(1), p_{j+1}(-1)).
\end{aligned}$$

For a monotone flux $\hat{f}(u, v)$ being Lipschitz continuous with Lipschitz constant $\mathcal{L}$, G is non-increasing in all the four arguments provided the following condition is met,

$$2\lambda\mathcal{L} \leq \min\{\omega^1, \omega^3\}.$$

Note that for the Lax-Friedrichs flux, $\mathcal{L} = \max_{u\in[c_1,c_2]} |f'(u)|$. Moreover, the consistency of the flux $\hat{f}(u, u) = f(u)$ implies that $G(u, u, u, u) = u$. Hence we have

$$\mathcal{C}_j \in [G(c_1, c_1, c_1, c_1, c_1), G(c_2, c_2, c_2, c_2, c_2)] = [c_1, c_2],$$

as long as the involved values are in $[c_1, c_2]$. It suffices to take

$$\lambda_0 = \frac{1}{2\mathcal{L}} \min\{\omega^1, \omega^3\} = \frac{1-3\gamma}{12\mathcal{L}(1-\gamma)}, \tag{2.17}$$

where we have used the fact that $\omega^1(\gamma) = \omega^3(-\gamma)$.

Step 4. The diffusion term. We apply the result in Theorem 2.2 to the case with $M(x) \equiv 1$ and μ replaced by 2μ to conclude that $\mathcal{D}_j \in [c_1, c_2]$ if $u_h^n(x) \in [c_1, c_2]$ on S_j and $\mu \le \mu_0$ with

$$\begin{aligned} \mu_0 &= \frac{1}{2}\left(\max_j \{A_h^n\}_{j+\frac{1}{2}}\right)^{-1} \min_{1\le j\le N}\left\{\frac{\omega^1(\pm\gamma)}{\alpha_3(\mp\gamma)+\alpha_1(\pm\gamma)}, \frac{1}{3(1-4\beta_1)}\right\} \\ &= \frac{1}{12}\left(\max_{1\le j\le N} \{A_h^n\}_{j+\frac{1}{2}}\right)^{-1} \min_{1\le j\le N}\left\{\frac{1\pm 3\gamma}{\beta_0(1\pm\gamma)+8\beta_1-2}, \frac{2}{1-4\beta_1}\right\}. \quad \square \end{aligned} \tag{2.18}$$

The above analysis can be readily carried over to (1.1) in one dimension, i.e.,

$$M(x)\partial_t u + \partial_x f(u) = \partial_x (A(x,u)\partial_x u). \tag{2.19}$$

We summarize the result in the following.

Theorem 2.4. *($k = 2$) The scheme (2.14) when applied to (2.19) with*

$$\beta_0 \ge 1 \quad \text{and} \quad \frac{1}{8} \le \beta_1 \le \frac{1}{4}$$

is maximum-principle-satisfying; namely, $\bar{u}_j^{n+1} \in [c_1, c_2]$ if $u_h^n(x) \in [c_1, c_2]$ on the set S_j's where

$$S_j = x_j + \frac{h}{2}\{-1, \gamma, 1\}$$

with γ satisfying

$$\gamma \in (a_j, b_j) \text{ as in (2.4) and } \quad |\gamma| \le 8\beta_1 - 1,$$

under the CFL condition

$$\lambda \le \lambda_0, \quad \mu \le \mu_0$$

for some λ_0 and μ_0 defined in (2.20) and (2.21), respectively.

Proof. The proof is entirely analogous to the proof of Theorem 2.3. One only needs to replace ω^i by $\tilde{\omega}^i$ for $i = 1, 2, 3$, and $\bar{u}_j$ by $\langle u_h \rangle_j$, respectively, so to obtain different CFL conditions. More precisely, (2.17) in Step 3 needs to be replaced by

$$\lambda_0 = \frac{1}{2\mathcal{L}} \min\{\tilde{\omega}^1(\gamma), \tilde{\omega}^3(\gamma)\} = \frac{1}{2\mathcal{L}} \min\{\tilde{\omega}^1(\pm\gamma)\}, \tag{2.20}$$

and (2.12) by a factor of $1/2$ gives

$$\mu_0 = \frac{1}{2}\left(\max_{1\le j\le N} A(x_{j+\frac{1}{2}})\right)^{-1} \min_{1\le j\le N}\left\{\frac{\tilde{\omega}_j^1(\pm\gamma)}{\alpha_3(\mp\gamma)+\alpha_1(\pm\gamma)}, \frac{\langle 1-\xi^2\rangle_j}{4(1-4\beta_1)}\right\}. \quad \square \tag{2.21}$$

3. MPS schemes in two dimensions

In this section, we design an MPS DG scheme to solve two dimensional problems on Cartesian meshes. Consider the model equation

$$M(x,y)\partial_t u = \nabla\cdot(A\nabla u) \text{ with } A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \quad \text{for } (x,y) \in \Omega \subset \mathbb{R}^2, \quad t > 0,$$

subject to the initial data $u_0(x, y)$ and periodic boundary conditions. Here $M(x, y) > 0$ is a given function, a, b and c are constant parameters so that A is nonnegative definite. The domain $\Omega = I \times J$ is a rectangle given by two intervals I and J in x and y direction, respectively. Let $\cup_{i=1}^{N_x} \cup_{j=1}^{N_y} K_{ij}$ be a partition of the domain Ω, with $K_{ij} = I_i \times J_j$, where

$$I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \quad J_j = [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}].$$

The finite element space is defined as

$$V_h = \{v \in L^2(\Omega), v\big|_{K_{ij}} \in Q^k(K_{ij}), i = 1, \cdots, N_x, j = 1, \cdots, N_y\}.$$

Here $Q^k(K_{ij})$ is the tensor product space of $P^k(I_i)$ and $P^k(J_j)$. Hence the DDG scheme can be formulated as follows: to find $u_h \in V_h$ such that for all $v \in V_h$,

$$\begin{aligned}\int_{K_{ij}} Mu_h^{n+1} v\,dxdy = &\int_{K_{ij}} Mu_h^n v\,dxdy - \tau \int_{K_{ij}} A\nabla u_h^n \cdot \nabla v\,dxdy \\ &+ \tau \int_{\partial K_{ij}} A\widehat{\nabla u_h^n} \cdot \nu v\,ds + \tau \int_{\partial K_{ij}} A\nabla v \cdot \nu (u_h^n - \{u_h^n\})\,ds,\end{aligned} \tag{3.1}$$

where ν is the outward unit normal to the cell boundary ∂K_{ij}, and the numerical flux

$$\widehat{\nabla u_h} = (\widehat{\partial_x u_h}, \widehat{\partial_y u_h})^\top \tag{3.2}$$

is defined as follows,

$$\begin{aligned}&\widehat{\partial_x u_h}\Big|_{(x_{i+\frac{1}{2}}, y)} = \frac{\beta_0}{\Delta x}[u_h] + \{\partial_x u_h\} + \beta_1 \Delta x[\partial_x^2 u_h],\ \widehat{\partial_y u_h}\Big|_{(x_{i+\frac{1}{2}}, y)} = \{\partial_y u_h\}, \\ &\widehat{\partial_y u_h}\Big|_{(x, y_{j+\frac{1}{2}})} = \frac{\beta_0}{\Delta y}[u_h] + \{\partial_y u_h\} + \beta_1 \Delta y[\partial_y^2 u_h],\ \widehat{\partial_x u_h}\Big|_{(x, y_{j+\frac{1}{2}})} = \{\partial_x u_h\},\end{aligned}$$

where β_0, β_1 are flux parameters to be determined to ensure the desired MPS property. Note that in $\widehat{\partial_y u_h}|_{(x_{i+\frac{1}{2}}, y)}$, jump terms do not show up since along interface $x = x_{i+\frac{1}{2}}$ and $y \in [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, there is no jump of polynomials in y direction. This argument applies to $\partial_x u_h|_{(x, y_{j+\frac{1}{2}})}$ as well. Here for a concise expression of the numerical flux, a uniform mesh has been used, with $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ and $\Delta y = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$.

To proceed, we recall some conventions similar to the one-dimensional case. The weighted cell average is defined as

$$\langle u^{n+1}\rangle_{ij} = \frac{\int_{K_{ij}} Mu_h\,dxdy}{\Delta x \Delta y} = \fint_{J_j}\fint_{I_i} Mu_h\,dxdy,$$

where $\fint$ denotes the average integral. We also define the weighted interval average in x and y, respectively,

$$\langle \phi(\xi)\rangle_i(y) = \fint_{-1}^{1} M\left(x_i + \frac{\Delta x}{2}\xi, y\right)\phi(\xi)d\xi, \quad \langle \phi(\eta)\rangle_j(x) = \fint_{-1}^{1} M\left(x, y_j + \frac{\Delta y}{2}\eta\right)\phi(\eta)d\eta.$$

The cell average

$$\overline{u}_{ij} = \frac{\int_{K_{ij}} Mu_h\,dxdy}{\int_{K_{ij}} M\,dxdy} = \frac{\langle u_h\rangle_{ij}}{\langle 1\rangle_{ij}}$$

update can be obtained from (3.1) as

$$\langle u_h^{n+1}\rangle_{ij} = \langle u^n\rangle_{ij} + \mu_x \Delta x \fint_{J_j}\left(a\widehat{\partial_x u_h^n} + c\widehat{\partial_y u_h^n}\right)dy\Big|_{\partial I_i} + \mu_y \Delta y \fint_{I_i}\left(b\widehat{\partial_y u_h^n} + c\widehat{\partial_x u_h^n}\right)dx\Big|_{\partial J_j}, \tag{3.3}$$

where $\mu_x = \frac{\tau}{(\Delta x)^2}$ and $\mu_y = \frac{\tau}{(\Delta y)^2}$. Let $\mu = \mu_x + \mu_y$ and decompose $\langle u^n\rangle_{ij}$ as

$$\langle u^n\rangle_{ij} = \frac{\mu_x}{\mu}\langle u^n\rangle_{ij} + \frac{\mu_y}{\mu}\langle u^n\rangle_{ij},$$

so that (3.3) can be rewritten as

$$\langle u^{n+1}\rangle_{ij} = \frac{\mu_x}{\mu}\fint_{J_j} H_1(y)\,dy + \frac{\mu_y}{\mu}\fint_{I_i} H_2(x)\,dx + B, \tag{3.4}$$

where

$$H_1(y) = \oint_{I_i} M(x,y)u_h^n\,dx + \mu\Delta x\, a\widehat{\partial_x u_h^n}\Big|_{\partial I_i},$$

$$H_2(x) = \oint_{J_j} M(x,y)u_h^n\,dy + \mu\Delta y\, b\widehat{\partial_y u_h^n}\Big|_{\partial J_j},$$

$$B = \frac{c\tau}{|K_{ij}|}\left[\int_{J_j}\{\partial_y u_h\}dy|_{\partial I_i} + \int_{I_i}\{\partial_x u_h\}dx|_{\partial J_j}\right].$$

Notice that B can be expressed as a combination of point values of u_h^n at four vertices of K_{ij} as

$$\begin{aligned}B = {}& \frac{c\tau}{2\Delta x\Delta y}\big(2u_h^n(x^-_{i+\frac12}, y^-_{j+\frac12}) - 2u_h^n(x^-_{i+\frac12}, y^+_{j-\frac12}) - 2u_h^n(x^+_{i-\frac12}, y^-_{j+\frac12}) + 2u_h^n(x^+_{i-\frac12}, y^+_{j-\frac12})\big)\\ &+ \frac{c\tau}{2\Delta x\Delta y}\big(u_h^n(x^+_{i+\frac12}, y^-_{j+\frac12}) + u_h^n(x^-_{i+\frac12}, y^+_{j+\frac12}) - u_h^n(x^-_{i+\frac12}, y^-_{j-\frac12}) - u_h^n(x^+_{i+\frac12}, y^+_{j-\frac12})\big)\\ &+ \frac{c\tau}{2\Delta x\Delta y}\big(u_h^n(x^+_{i-\frac12}, y^-_{j-\frac12}) + u_h^n(x^-_{i-\frac12}, y^+_{j-\frac12}) - u_h^n(x^+_{i-\frac12}, y^+_{j+\frac12}) - u_h^n(x^-_{i-\frac12}, y^-_{j+\frac12})\big).\end{aligned}$$

The two integrals in (3.4) can be approximated by the Gauss-Lobatto quadrature rule with sufficient accuracy. Let us assume that we use an L-point Gauss-Lobatto quadrature rule with $L \geq \frac{k+3}{2}$ points, which has accuracy of at least $O(h^{k+2})$. Let

$$\hat{S}_i^x = \{x_{i-\frac12} = \hat{x}_i^1 < \cdots < \hat{x}_i^\sigma < \cdots < \hat{x}_i^L = x_{i+\frac12}\},$$
$$\hat{S}_j^y = \{y_{j-\frac12} = \hat{y}_j^1 < \cdots < \hat{y}_j^\sigma < \cdots < \hat{y}_j^L = y_{j+\frac12}\}$$

denote the quadrature points on I_i and J_j, respectively, and $\hat{\omega}^\sigma$'s be the associated quadrature weights so that

$$\sum_{\sigma=1}^{L}\hat{\omega}^\sigma = 1.$$

Using the quadrature rule on the right-hand side of (3.4), we obtain the following scheme

$$\langle u_h^{n+1}\rangle_{ij} = \frac{\mu_x}{\mu}\sum_{\sigma=1}^{L}\hat{\omega}^\sigma H_1(\hat{y}_j^\sigma) + \frac{\mu_y}{\mu}\sum_{\sigma=1}^{L}\hat{\omega}^\sigma H_2(\hat{x}_i^\sigma) + B. \tag{3.5}$$

Also set

$$S_i^x = \{x_{i-\frac12}, x_i^\gamma, x_{i+\frac12}\} = x_i + \frac{\Delta x}{2}\{-1, \gamma^x, 1\},$$
$$S_j^y = \{y_{j-\frac12}, y_j^\gamma, y_{j+\frac12}\} = y_j + \frac{\Delta y}{2}\{-1, \gamma^y, 1\}$$

to denote the test sets on $[x_{i-\frac12}, x_{i+\frac12}]$ and $[y_{j-\frac12}, y_{j+\frac12}]$, respectively, with γ^x, γ^y satisfying

$$\frac{\langle\xi-\xi^2\rangle_i}{\langle 1-\xi\rangle_i}(y_j^\sigma) < \gamma^x < \frac{\langle\xi+\xi^2\rangle_i}{\langle 1+\xi\rangle_i}(y_j^\sigma), \quad |\gamma^x| \leq 8\beta_1 - 1,$$
$$\frac{\langle\eta-\eta^2\rangle_j}{\langle 1-\eta\rangle_j}(x_i^\sigma) < \gamma^y < \frac{\langle\eta+\eta^2\rangle_j}{\langle 1+\eta\rangle_j}(x_i^\sigma), \quad |\gamma^y| \leq 8\beta_1 - 1.$$

We use $\otimes$ to denote the tensor product and define

$$S_{ij} = (S_i^x \otimes \hat{S}_j^y) \cup (\hat{S}_i^x \otimes S_j^y).$$

The main result can now be stated in the following.

Theorem 3.1. *($k = 2$) Consider the two dimensional DDG scheme (3.1) on rectangular meshes. Assume the mesh is regularly shaped, i.e., $\kappa^{-1} \leq \frac{\Delta x}{\Delta y} \leq \kappa$, for some constant $\kappa > 0$, and the flux parameters (β_0, β_1) satisfy*

$$\beta_0 \geq 1 + \frac{\kappa|c|}{2\hat{\omega}^1\min\{a,b\}}. \quad \text{and} \quad \frac18 \leq \beta_1 \leq \frac14 \tag{3.6}$$

If $u_h^n(x,y) \in [c_1, c_2]$ for all $(x,y) \in S_{ij}$, then there exists $\mu_0 > 0$ such that if $\mu \le \mu_0$ the cell average $\bar{u}_{ij}^{n+1} \in [c_1, c_2]$. More precisely, we have

$$\mu_0 = \min_{i,j} \underline{\omega}_{ij} \min\left\{ \frac{\hat{\omega}^1}{\hat{\omega}^1 \max\{a,b\}\left(\beta_0 + \frac{8\beta_1 - 2}{1+\gamma}\right) + \kappa|c|}, \frac{1-\gamma^2}{4\max\{a,b\}(1-4\beta_1)} \right\}, \tag{3.7}$$

where $\underline{\omega}_{ij} > 0$ is defined in equation (A.3), and $\gamma = \max\{|\gamma^x|, |\gamma^y|\} \le 8\beta_1 - 1$.

The proof is relegated to Appendix A. We proceed to deal with nonlinear diffusion equations in the next subsection.

3.1. Application to nonlinear diffusion equations

This section is devoted to application to nonlinear diffusion equations of the form

$$\partial_t u = \nabla \cdot (A \nabla u) \text{ with } A(x,y,u) = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \quad \text{for } (x,y) \in \Omega \subset \mathbb{R}^2, \quad t > 0,$$

subject to initial data $u_0(x,y)$ and periodic boundary conditions, and $A(x,y,u)$ is nonnegative definite. This type of model arises in a wide range of applications.

Hence the DDG scheme can be formulated as follows: to find $u_h \in V_h$ such that for all $v \in V_h$,

$$\begin{aligned} \int_{K_{ij}} u_h^{n+1} v \, dxdy = & \int_{K_{ij}} u_h^n v \, dxdy - \tau \int_{K_{ij}} A_h^n \nabla u_h^n \cdot \nabla v \, dxdy \\ & + \tau \int_{\partial K_{ij}} \{A_h^n\} \widehat{\nabla u_h^n} \cdot \nu v \, ds + \tau \int_{\partial K_{ij}} \{A_h^n\} \nabla v \cdot \nu (u_h^n - \{u_h^n\}) \, ds, \end{aligned} \tag{3.8}$$

where ν is the outward unit normal to the cell boundary ∂K_{ij}, the numerical flux $\widehat{\nabla u_h}$ is defined in (3.2), and $A_h^n = A(x,y,u_h^n)$.

The cell average evolves according to

$$\bar{u}_{ij}^{n+1} = \bar{u}_{ij}^n + \mu_x \Delta x \fint_{J_j} \left(\{a_h^n\} \widehat{\partial_x u_h^n} + \{c_h^n\} \widehat{\partial_y u_h^n} \right) dy \Big|_{\partial I_i} + \mu_y \Delta y \fint_{I_i} \left(\{b_h^n\} \widehat{\partial_y u_h^n} + \{c_h^n\} \widehat{\partial_x u_h^n} \right) dx \Big|_{\partial J_j}.$$

That is

$$\bar{u}_{ij}^{n+1} = \frac{\mu_x}{\mu} \fint_{J_j} H_1(y) \, dy + \frac{\mu_y}{\mu} \fint_{I_i} H_2(x) \, dx + B, \tag{3.9}$$

where $\mu_x = \frac{\tau}{(\Delta x)^2}$, $\mu_y = \frac{\tau}{(\Delta y)^2}$ and $\mu = \mu_x + \mu_y$, with

$$\begin{aligned} H_1(y) &= \fint_{I_i} u_h^n \, dx + \mu \Delta x \, \{a_h^n\} \widehat{\partial_x u_h^n} \Big|_{\partial I_i}, \\ H_2(x) &= \fint_{J_j} u_h^n \, dy + \mu \Delta y \, \{b_h^n\} \widehat{\partial_y u_h^n} \Big|_{\partial J_j}, \\ B &= \frac{\tau}{|K_{ij}|} \left[\int_{J_j} \{c_h^n\} \{\partial_y u_h\} dy|_{\partial I_i} + \int_{I_i} \{c_h^n\} \{\partial_x u_h\} dx|_{\partial J_j} \right]. \end{aligned}$$

The main result can be stated in the following.

Theorem 3.2. *($k = 2$) Consider the two dimensional DDG scheme (3.8) on rectangular meshes. Assume the mesh is regularly shaped, i.e., $\kappa^{-1} \le \frac{\Delta x}{\Delta y} \le \kappa$, for some constant $\kappa > 0$, and the flux parameters (β_0, β_1) satisfy*

$$\beta_0 \ge 1 + \frac{2\kappa \|c\|_\infty L(L-1)}{(1-\gamma)\min\{a,b\}}, \quad \text{and} \quad \frac{1}{8} \le \beta_1 \le \frac{1}{4} \tag{3.10}$$

and $\gamma = \max\{|\gamma^x|, |\gamma^y|\} \le 8\beta_1 - 1$, *where*

$$\|c\|_\infty = \max_{(x,y)\in\Omega, u\in[c_1,c_2]} |c(x,y,u)|, \quad \min\{a,b\} = \min_{(x,y)\in\Omega, u\in[c_1,c_2]} \big\{a(x,y,u), b(x,y,u)\big\}.$$

If $u_h^n(x,y) \in [c_1,c_2]$ *for all* $(x,y) \in S_{ij}$, *then there exists* $\mu_0 > 0$ *such that if* $\mu \le \mu_0$ *the cell average* $\bar{u}_{ij}^{n+1} \in [c_1,c_2]$. *More precisely, we have*

$$\mu_0 = \min\left\{ \frac{1-3\gamma}{6\max\{a,b\}\big(\beta_0 + \frac{8\beta_1-2}{1+\gamma}\big)(1-\gamma) + 12\kappa\|c\|_\infty L(L-1)}, \frac{1}{6\max\{a,b\}(1-4\beta_1)} \right\}, \tag{3.11}$$

where

$$\max\{a,b\} = \max_{(x,y)\in\Omega, u\in[c_1,c_2]} \big\{a(x,y,u), b(x,y,u)\big\}.$$

The proof is relegated to Appendix B.

4. Scaling limiter and the MPS algorithm

4.1. Scaling limiter

The one dimensional result in Theorem 2.2 and the two dimensional result in Theorem 3.1 tell us that for the DDG scheme with forward Euler time discretization, we need to modify u_h^n such that it is in $[c_1,c_2]$ on the test set $S = S_j$ or S_{ij}. In one dimensional case, we can use the following scaling limiter

$$\tilde{u}_h(x) = \theta\left(u_h(x) - \bar{u}_j\right) + \bar{u}_j \quad \text{with } \theta = \min\left\{1, \left|\frac{\bar{u}-c_1}{\bar{u}_j - m_1}\right|, \left|\frac{c_2-\bar{u}_j}{m_2-\bar{u}_j}\right|\right\}, \tag{4.1}$$

where

$$m_1 = \min_{x\in S_j} u_h(x), \qquad m_2 = \max_{x\in S_j} u_h(x). \tag{4.2}$$

In the two dimensional case,

$$\tilde{u}_h(x,y) = \theta\left(u_h(x,y) - \bar{u}_{ij}\right) + \bar{u}_{ij} \quad \text{where } \theta = \min\left\{1, \left|\frac{\bar{u}_{ij}-c_1}{\bar{u}_{ij}-m_1}\right|, \frac{c_2-\bar{u}_{ij}}{m_2-\bar{u}_{ij}}\right|\right\}, \tag{4.3}$$

where

$$m_1 = \min_{(x,y)\in S_{ij}} u_h(x,y) \quad \text{and} \quad m_2 = \max_{(x,y)\in S_{ij}} u_h(x,y). \tag{4.4}$$

The modified polynomials are indeed in $[c_1,c_2]$ and preserve the cell average. Moreover, following [23] it can be shown that the above scaling limiters do not destroy the accuracy. We summarize this for two dimensional case only.

Lemma 4.1. *If* $\bar{u}_{ij} \in (c_1,c_2)$, *then the modified polynomial (4.3) is as accurate as* $u_h(x,y)$ *to approximate* $u(t,x,y)$ *for the same* t *in the following sense:*

$$|u_h(x,y) - \tilde{u}_h(x,y)| \le C_k \|u_h(\cdot,\cdot) - u(t,\cdot,\cdot)\|_\infty,$$

where C_k *is a constant depending on the polynomial degree* k *and the weight function* $M(x,y)$.

4.2. Algorithm

The fact that we only require u_h^n be in the desired range $[c_1,c_2]$ at certain points in $\cup_{i,j} S_{ij}$ can be used to reduce the computational cost in a great deal.

Given the weighted L^2 projection u_h^0 computed from the initial data $u_0(x,y)$, the algorithm is stated below:

(1) Initialization.
Obtain $u_h^0 \in V_h$ using the standard piecewise L^2 projection

$$\int_{I_{ij}} u_0(x,y)\phi dxdy = \int_{I_{ij}} u_h^0 \phi dxdy \quad \text{for} \quad \phi \in V_h.$$

(2) Time evolution.
For $n = 0, 1, 2, \ldots$,
(a) Check the point values of u_h^n on the test set S_{ij}. If one of them goes outside of $[c_1, c_2]$, reconstruct $\hat{u}_h^n$ using the formula (4.3) and (4.4) and set $u_h^n = \hat{u}_h^n$.
(b) Use the scheme (3.8) to compute u_h^{n+1}.
End

This algorithm is guaranteed to produce numerical solutions within the range with uniform third order accuracy for smooth exact solutions.

The algorithm with forward Euler time discretization can be extended to high order ODE solves, such as the strong stability preserving Runge-Kutta methods, since they are a convex linear combination of the forward Euler; see [35]. The desired MPS property can be ensured as long as the proper time step restrictions are respected.

5. Numerical tests

In this section, we present the results of numerical tests using our third-order maximum-principle-preserving DDG schemes. The numerical integration is computed using the Gaussian quadrature rule. Since $M(x)$ and $A(x)$ can be complex functions, sufficient number of quadrature points will guarantee the desired order of accuracy and induce small numerical errors. Hence we take 16 quadrature points in each cell through all the examples. As for the time stepping, we implement the SSP(3,3) scheme as in [12] for the strong-stability-preservation in time. The one-dimensional error is measured by the discrete norms:

$$e_p^h(t) = \left(\sum_{j=1}^{N_x} \|u_h(t,\cdot) - u(t,\cdot)\|_{L^p(I_j)}^p \right)^{\frac{1}{p}}, \text{ for } p = 1 \text{ or } 2,$$

and $u(t,x)$ is taken as the exact solution or the reference solution given by greatly refined spacial discretization. If neither of them is available, we can compute the consecutive errors between $u_h(t,x)$ and $u_{\frac{h}{2}}(t,x)$, where the subindex indicates the mesh size h and $\frac{h}{2}$. We introduce $e_\infty^h(t)$ to demonstrate the discrete MPS property at the test points in S_j:

$$e_\infty^h(t) = \max_{x \in S_j, 0 \le j \le N} \{c_1 - u_h(t,x), u_h(t,x) - c_2\}. \tag{5.1}$$

If $e_\infty^h(t) > 0$, then the discrete MPS property is violated.

5.1. One dimensional numerical tests

In our numerical tests we choose scheme parameters as

$$\beta_0 = 2,\ \beta_1 = 0.16,\ \gamma = 0.1.$$

Accuracy test

We construct a linear problem of form (1.1) to demonstrate the third order accuracy of our numerical schemes:

$$\begin{cases} M(x)\partial_t u = \partial_x\big(A(x)\partial_x u\big), & x \in [1,3],\ t > 0, \\ u(0,x) = \sin(x^2 - 1), & x \in [1,3] \end{cases} \tag{5.2}$$

with $M(x) = 4xe^{-x^2+1}$, $A(x) = \frac{e^{-x^2+1}}{x}$. The exact solution is given by

$$u(t,x) = \exp(-t)\sin(x^2 - 1 - t).$$

The boundary condition is imposed by using the exact solution. We take the final time $t = 0.1$ with different mesh sizes $N_x = 16, 32, 64$ and 128. Fig. 1 shows the logarithm of the L^2 error for u_h, denoted by circles. One can observe the third order of accuracy for u_h.

Porous medium equation

The porous medium equation

$$\partial_t u = \partial_x^2(u^m), \quad m > 1 \tag{5.3}$$

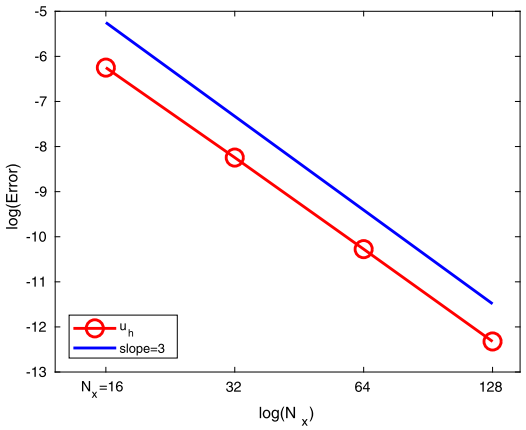


Fig. 1. The accuracy test on (5.2). The figure shows the logarithm of the L^2 error with different number of meshes for u_h, denoted by circles. One can observe the third order of accuracy for u_h.

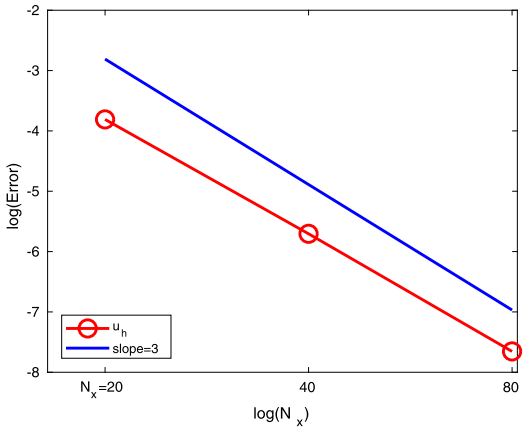


Fig. 2. The accuracy test on (5.3) with $m = 5$ at $t = 0.1$. The error e_p^h is computed by removing the two nonsmooth corners. The figure shows the logarithm of the L^2 error with different number of meshes for u, denoted by circles. One can observe the third order of accuracy for u_h.

is known to admit the Barenblatt solution of the form

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

which is compactly supported. Fig. 2 provides the accuracy test on (5.3) with $m = 5$ at $t = 0.1$. Removing the two nonsmooth corners, one can observe the third-order accuracy for the numerical solutions of u_h.

We compute the numerical solution with initial data $B_m(1,x)$ subject to zero boundary conditions for $m = 2$, up to final time $t = 3$. From the numerical results in Fig. 3(a) with the MPS limiter we see a sharp resolution of discontinuities, and keeping the solution strictly within the initial bounds everywhere for all time. Fig. 3(b) shows a zoom-in at the nonsmooth corner for $x \in [-6, -4]$ where the solution u is well simulated. In contrast, without MPS limiter, it brings in significant overshoots near the upper bound of the exact solution, as evidenced by oscillations already appearing at $t = 1.0025$ in Fig. 3(c). In addition, the scheme without the MPS limiter will blow up in a short time.

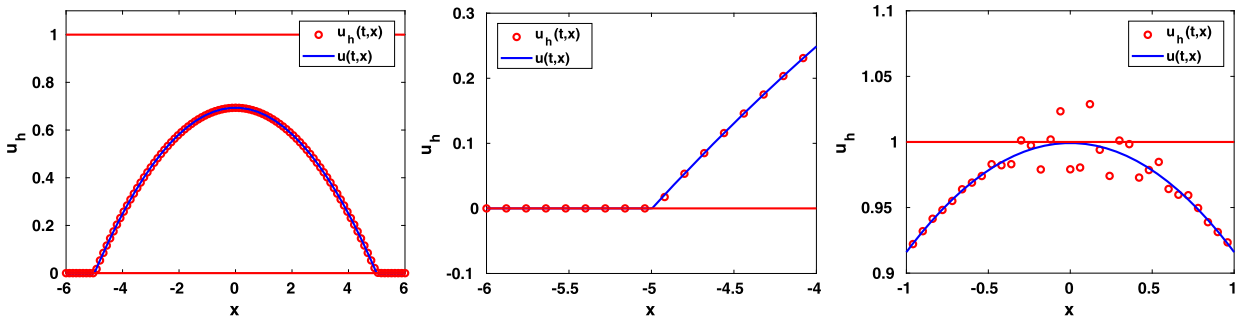
The Buckley-Leverett equation

The convection-diffusion Buckley-Leverett equation of the form

$$\partial_t u + \partial_x f(u) = \varepsilon \partial_x(\nu(u)\partial_x u), \tag{5.4}$$

is a model often used in reservoir simulations; see [16]. Here $\varepsilon > 0$ is a small parameter, f has an s-shape:

$$f(u) = \frac{u^2}{u^2 + (1-u)^2},$$



(a) $u_h(t=3,x)$. (b) Zoom-in of $u_h(t=3,x)$. (c) $u_h(t=1.0025,x)$.

Fig. 3. The numerical solution to (5.3) with $m=2$ and $N_x=200$, $\Delta t=0.0001$. $\mu_0\approx 3.66\times 10^{-2}$. Fig. 3b shows $u_h(t=3,x)$ (circles) with the MPS limiter against the exact solution (solid lines). Fig. 3c shows the numerical solution without the MPS limiter at $t=1.0025$, zoomed in $[-1,1]$. This numerical solution blows up shortly.

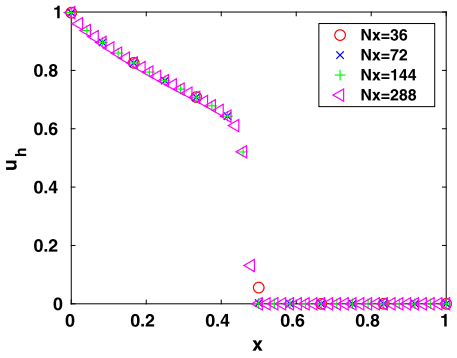


Fig. 4. The numerical solution of problem (5.4) and (5.5) for $t=0.2$ with $N_x=36,72,144,288$.

and

$$v(u)=4u(1-u)1_{0\le u\le 1}.$$

We numerically solve (5.4) with $\varepsilon=0.01$, subject to the following initial and boundary conditions

$$u(0,x)=(1-3x)1_{0\le x\le 1/3},\quad u(t,0)=1,\quad u(t,1)=0. \tag{5.5}$$

The exact solution is not available. With numerical convergence we demonstrate that our numerical scheme is capable of simulating the sharp corner of the solution that is moving in time. From the results in Fig. 4 we observe the numerical convergence when the spacial mesh is refined. Moreover, the lower bound of $u_h(t,x)$ is well preserved around the corner point $x=0.5$. We note that the numerical solution here is comparable to that obtained in [16] by the second order central scheme.

5.2. Two dimensional numerical tests

We consider the convection-diffusion equation:

$$\partial_t u+\nabla\cdot f(u)=\nabla\cdot(A\nabla u) \text{ for } (x,y)\in[-1,1]\times[-1,1] \text{ and } t>0, \tag{5.6}$$

where

$$f(u)=u\vec{v},\quad \vec{v}:=(0.01,0.01)^\top$$

and A is a symmetric, positive definite matrix. For such an equation, exact solutions can be found of the form $u(t,x,y)=a(t)\exp(-\xi^\top B(t)\xi)$ with $\xi=\boldsymbol{x}-\vec{v}t$, $\boldsymbol{x}:=(x,y)^\top$, provided $a'=-a\mathrm{tr}(AB)$ and $B'=-2B^\top AB$. Here $a(0)$ can be chosen small enough to ensure that the periodic boundary condition adopted is reasonable and accurate at a finite time. In fact, if we set $\sigma_0=0.01$ and a 2×2 matrix $\sigma(t)$ be such that

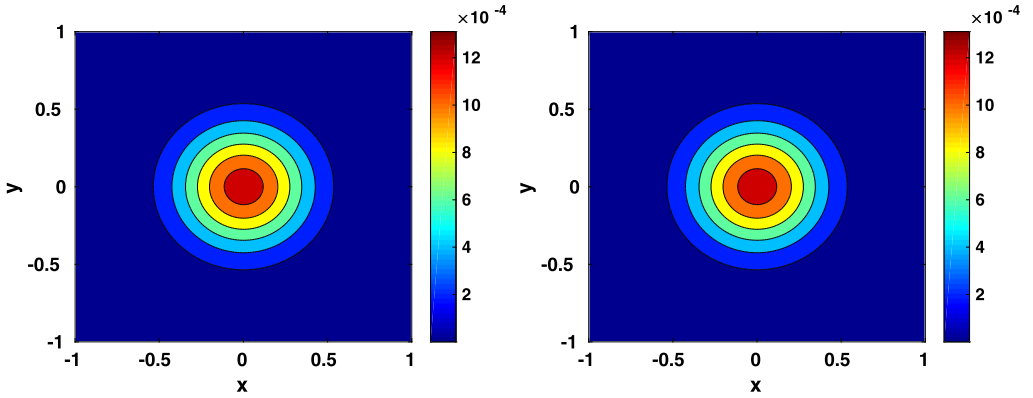
$$\sigma(t)=\sigma_0^2\mathrm{Id}+2At \quad \text{with Id being the identity matrix.}$$

Table 1
MPS limiter.

t	min(u)	max(u)
0	0	1
2×10^{-6}	0	0.96111
4×10^{-6}	0	0.92156
1×10^{-5}	0	0.81327
2×10^{-5}	0	0.68236

Table 2
No limiter: $e^h_\infty(t)$.

t	$\beta_0=2$	$\beta_0=4$
0	1.705E–005	1.705E–005
2×10^{-6}	7.023E–004	4.168E–004
4×10^{-6}	1.339E–003	6.838E–004
6×10^{-6}	1.874E–003	6.387E–004
8×10^{-6}	2.403E–003	4.414E–004



(a) $u(t,x)$. (b) $u_h(t,x)$.

Fig. 5. The contours of solutions to (5.6) with the first choice of the matrix A and $N_x=N_y=200$, $\Delta t=10^{-6}$. $\mu_0\approx4.62\times10^{-3}$. The final time $t=0.0381$. Fig. 5a shows the exact solution. Fig. 5b shows the numerical solution with the MPS limiter. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Then the function

$$u(t,x,y)=\frac{\sigma_0^2}{|\det(\sigma)|^{\frac{1}{2}}}\exp\left(-\frac{(\boldsymbol{x}-\vec{v}t)^T\sigma^{-1}(\boldsymbol{x}-\vec{v}t)}{2}\right),\quad \boldsymbol{x}:=(x,y)^\top$$

as given in [4], is a solution to equation (5.6).

In our numerical tests, we take three choices of the tensor A as

$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&2\end{pmatrix}\text{ or }\begin{pmatrix}1&1\\1&2\end{pmatrix},$$

which are usually denoted as isotropic, diagonally anisotropic and fully anisotropic diffusion tensors.

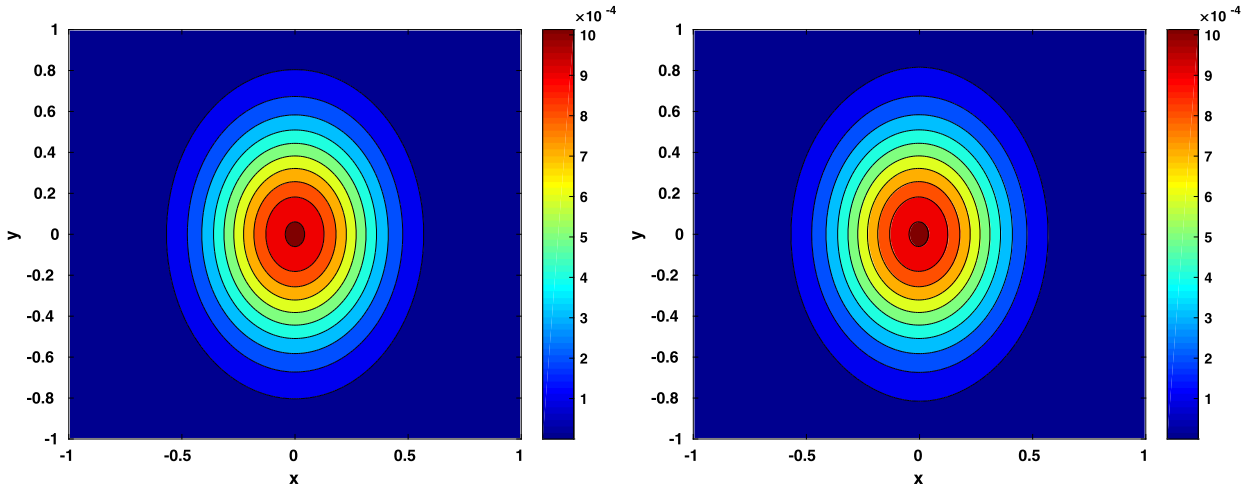
We begin to demonstrate the necessity of the MPS limiter using the first isotropic problem. Table 1 shows the minumum and maximum of the numerical solutions $u_h(t,x,y)$ using the MPS limiter over all the points in the test sets S_{ij}. They are well bounded in the interval $[0,1]$, satisfying the maximum principle. Table 2 shows the error $e^h_\infty(t)$ in (5.1) without the MPS limiter. One can observe that the numerical solutions $u_h(t,x,y)$ violates the maximum principle and the simulation will break down after some time. Larger β_0 helps to suppress the overshoot or undershoot, but cannot realize the bound preservation ideally.

In the first two cases, we take $\beta_0=2,\beta_1=0.16,\gamma=0.1$ for both variables x and y. For the fully anisotropic one, we take $\beta_0=4$. For smaller β_0, small oscillations develop in time and the scheme can become unstable. We observe a nice agreement of the numerical solution to the exact solution for all three cases of A.

For the equation (5.6) with the third choice of the matrix A, we take care of the anisotropy by using a larger $\beta_0=4$, and adpative mesh size and time steps. In the beginning, the solution is highly concentrated around the origin and requires a good resolution. Therefore, for $t\in[0,10^{-5}]$, we employ the discretization with $\Delta x=\Delta y=0.005$, $\Delta t=10^{-7}$. Afterwards, a coarser mesh with $\Delta x=\Delta y=0.01$, $\Delta t=10^{-6}$ is used. For time $t=0.01$, Fig. 7(a-b) show the contours of the exact solution and the numerical solution with the MPS limiter. Fig. 7(c) shows a slice of the exact solution (circles) and the numerical solution (crosses) for $y=0$. One can observe a good agreement of the two solutions.

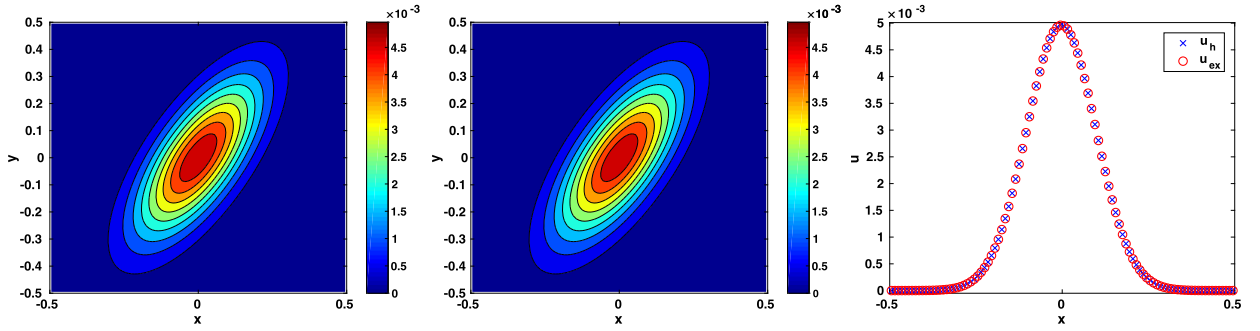
6. Concluding remarks

In this paper, we present third order accurate DDG schemes which can be proven maximum-principle-satisfying for a class of diffusion equations with variable diffusivity in terms of spatial variables and/or the unknown, in both one and



(a) $u(t,x,y)$. (b) $u_h(t,x,y)$.

Fig. 6. The contours of solutions to (5.6) with the second choice of the matrix A and $N_x = N_y = 200$, $\Delta t = 10^{-6}$. $\mu_0 \approx 4.62 \times 10^{-3}$. The final time $t = 0.03485$. Fig. 6a shows the exact solution. Fig. 6b shows the numerical solution with the MPS limiter.



(a) $u(t,x,y)$. (b) $u_h(t,x,y)$. (c) $u_h(t,x,0)$.

Fig. 7. (5.6) with the third choice of the matrix A and adaptive mesh size and time steps. $\mu_0 \approx 4.04 \times 10^{-3}$. In the beginning, the solution is highly concentrated around the origin and requires a good resolution. Therefore, for $t \in [0, 10^{-5}]$, we employ the discretization with $\Delta x = \Delta y = 0.005$, $\Delta t = 10^{-7}$. Afterwards, a coarser mesh with $\Delta x = \Delta y = 0.01$, $\Delta t = 10^{-6}$ is used. The final time $t = 0.01$. Fig. 7a shows the exact solution. Fig. 7b shows the numerical solution with the MPS limiter. Fig. 7c shows the good agreement between the exact solution (circles) and the numerical solution (crosses) for $y = 0$.

two dimensional settings. Through careful theoretical analysis and numerical tests, we show that under suitable CFL conditions, with a simple scaling limiter involving little additional computational cost, the numerical schemes satisfy the strict maximum principle while maintaining uniform third order accuracy. The methodology extends to three dimensional rectangular meshes as well. The effectiveness of the maximum-principle-satisfying DG schemes has been demonstrated through extensive numerical examples.

The presence of variable diffusivity has led to slightly more restrictive CFL conditions in order to preserve the maximum principle. In two dimensional case, the CFL condition seems a main factor for the slow numerical time evolution. It would be interesting to extend the present result to implicit schemes so to improve the computational efficiency.

Acknowledgements

Liu's research was partially supported by the National Science Foundation under Grant DMS1812666 and by NSF Grant RNMS (Ki-Net) 1107291.

Appendix A. Proof of Theorem 3.1

Upon regrouping, we decompose the right-hand side of (3.5) as

$$\langle u^{n+1}\rangle_{ij} = \frac{\mu_x}{\mu}\sum_{\sigma=2}^{L-1}\hat{\omega}^\sigma H_1(\hat{y}_j^\sigma) + \frac{\mu_y}{\mu}\sum_{\sigma=2}^{L-1}\hat{\omega}^\sigma H_2(\hat{x}_i^\sigma)$$

$$+\frac{\mu_x\hat{\omega}_1}{\mu}\big(H_1(y_{j-\frac{1}{2}})+H_1(y_{j+\frac{1}{2}})\big)+\frac{\mu_y\hat{\omega}_1}{\mu}\big(H_2(x_{i-\frac{1}{2}})+H_2(x_{i+\frac{1}{2}})\big)+B.$$

Here $\hat{\omega}_1=\hat{\omega}_L=\frac{1}{L(L-1)}$ is used. From (2.11) we see that

$$H_1(\hat{y}_j^\sigma)=R_i(M(\cdot,\hat{y}_j^\sigma),\mu,\Delta x,a),\quad H_2(\hat{x}_i^\sigma)=R_j(M(\hat{x}_i^\sigma,\cdot),\mu,\Delta y,b)$$

for $1\le\sigma\le L$. Notice that the terms in B, induced from the nontrivial c, involve only polynomials values at four vertices of K_{ij}, we proceed to regroup and combine them with $H_1(y_{j\pm\frac{1}{2}})$ and $H_2(x_{i\pm\frac{1}{2}})$, respectively, in the following way:

$$B_1=\frac{c\tau}{2\Delta x\Delta y}\big(u_h(x^-_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}})+u_h(x^+_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}})-u_h(x^-_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})-u_h(x^+_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})\big)\text{ with }H_1(y_{j-\frac{1}{2}}),$$
$$B_2=\frac{c\tau}{2\Delta x\Delta y}\big(u_h(x^-_{i+\frac{1}{2}},y^-_{j+\frac{1}{2}})-u_h(x^+_{i-\frac{1}{2}},y^-_{j+\frac{1}{2}})+u_h(x^+_{i+\frac{1}{2}},y^-_{j+\frac{1}{2}})-u_h(x^-_{i-\frac{1}{2}},y^-_{j+\frac{1}{2}})\big)\text{ with }H_1(y_{j+\frac{1}{2}}),$$
$$B_3=\frac{c\tau}{2\Delta x\Delta y}\big(u_h(x^+_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}})-u_h(x^-_{i-\frac{1}{2}},y^-_{j+\frac{1}{2}})+u_h(x^+_{i-\frac{1}{2}},y^-_{j-\frac{1}{2}})-u_h(x^+_{i-\frac{1}{2}},y^+_{j+\frac{1}{2}})\big)\text{ with }H_2(x_{i-\frac{1}{2}}),$$
$$B_4=\frac{c\tau}{2\Delta x\Delta y}\big(u_h(x^-_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})-u_h(x^-_{i+\frac{1}{2}},y^-_{j+\frac{1}{2}})+u_h(x^-_{i+\frac{1}{2}},y^-_{j-\frac{1}{2}})-u_h(x^-_{i+\frac{1}{2}},y^+_{j+\frac{1}{2}})\big)\text{ with }H_2(x_{i+\frac{1}{2}}).$$

We shall also use the following notations.

$$\tilde{\omega}_i^{x,1}(\gamma^x,y)=\frac{\langle\gamma^x-\xi(1+\gamma^x)+\xi^2\rangle_i(y)}{2(1+\gamma^x)},\qquad \tilde{\omega}_j^{y,1}(\gamma^y,x)=\frac{\langle\gamma^y-\eta(1+\gamma^y)+\eta^2\rangle_j(x)}{2(1+\gamma^y)},$$
$$\tilde{\omega}_i^{x,2}(\gamma^x,y)=\frac{\langle1-\xi^2\rangle_i(y)}{1-(\gamma^x)^2},\qquad \tilde{\omega}_j^{y,2}(\gamma^y,x)=\frac{\langle1-\eta^2\rangle_j(x)}{1-(\gamma^y)^2},$$
$$\tilde{\omega}_i^{x,3}(\gamma^x,y)=\frac{\langle-\gamma^x+\xi(1-\gamma^x)+\xi^2\rangle_i(y)}{2(1-\gamma^x)},\qquad \tilde{\omega}_j^{y,3}(\gamma^y,x)=\frac{\langle-\gamma^y+\eta(1-\gamma^y)+\eta^2\rangle_j(x)}{2(1-\gamma^y)}.$$

For the first group, we have

$$\begin{aligned}
&H_1(y_{j-\frac{1}{2}})+\frac{\mu}{\mu_x\hat{\omega}_1}B_1=R_i(M(\cdot,y_{j-\frac{1}{2}}),\mu,\Delta x,a)+\frac{\mu}{\mu_x\hat{\omega}_1}B_1\\
&=\left[\tilde{\omega}_i^{x,1}(\gamma^x,y_{j-\frac{1}{2}})-\mu a\big(\alpha_3(-\gamma^x)+\alpha_1(\gamma^x)\big)+\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right]u_h(x^+_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}})\\
&\quad+\left[\tilde{\omega}_i^{x,2}(\gamma^x,y_{j-\frac{1}{2}})-\mu a\big(\alpha_2(-\gamma^x)+\alpha_2(\gamma^x)\big)\right]u_h(x^\gamma_i,y^+_{j-\frac{1}{2}})\\
&\quad+\left[\tilde{\omega}_i^{x,3}(\gamma^x,y_{j-\frac{1}{2}})-\mu a\big(\alpha_1(-\gamma^x)+\alpha_3(\gamma^x)\big)-\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right]u_h(x^-_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})\\
&\quad+\left(\mu a\alpha_3(-\gamma^x)-\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right)u_h(x^+_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})\\
&\quad+\mu a\left[\alpha_2(-\gamma^x)u_h(x^\gamma_{i+1},y^+_{j-\frac{1}{2}})+\alpha_1(-\gamma^x)u_h(x^-_{i+\frac{3}{2}},y^+_{j-\frac{1}{2}})\right]\\
&\quad+\mu a\left[\alpha_1(\gamma^x)u_h(x^+_{i-\frac{3}{2}},y^+_{j-\frac{1}{2}})+\alpha_2(\gamma^x)u_h(x^\gamma_{i-1},y^+_{j-\frac{1}{2}})\right]\\
&\quad+\left(\mu a\alpha_3(\gamma^x)+\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right)u_h(x^-_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}}),
\end{aligned}$$

and the second group reduces to

$$\begin{aligned}
&H_1(y_{j+\frac{1}{2}})+\frac{\mu}{\mu_x\hat{\omega}_1}B_2=R_i(M(\cdot,y_{j+\frac{1}{2}}),\mu,\Delta x,a)+\frac{\mu}{\mu_x\hat{\omega}_1}B_1\\
&=\left[\tilde{\omega}_i^{x,1}(\gamma^x,y_{j+\frac{1}{2}})-\mu a\big(\alpha_3(-\gamma^x)+\alpha_1(\gamma^x)\big)-\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right]u_h(x^+_{i-\frac{1}{2}},y^-_{j+\frac{1}{2}})\\
&\quad+\left[\tilde{\omega}_i^{x,2}(\gamma^x,y_{j+\frac{1}{2}})-\mu a\big(\alpha_2(-\gamma^x)+\alpha_2(\gamma^x)\big)\right]u_h(x^\gamma_i,y^-_{j+\frac{1}{2}})\\
&\quad+\left[\tilde{\omega}_i^{x,3}(\gamma^x,y_{j+\frac{1}{2}})-\mu a\big(\alpha_1(-\gamma^x)+\alpha_3(\gamma^x)\big)+\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right]u_h(x^-_{i+\frac{1}{2}},y^-_{j+\frac{1}{2}})
\end{aligned}$$

$$
\begin{aligned}
&+\left(\mu a\alpha_3(-\gamma^x)+\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right)u_h(x^+_{i+\frac{1}{2}},y^-_{j+\frac{1}{2}})\\
&+\mu a\left[\alpha_2(-\gamma^x)u_h(x^\gamma_{i+1},y^-_{j+\frac{1}{2}})+\alpha_1(-\gamma^x)u_h(x^-_{i+\frac{3}{2}},y^-_{j+\frac{1}{2}})\right]\\
&+\mu a\left[\alpha_1(\gamma^x)u_h(x^+_{i-\frac{3}{2}},y^-_{j+\frac{1}{2}})+\alpha_2(\gamma^x)u_h(x^\gamma_{i-1},y^-_{j+\frac{1}{2}})\right]\\
&+\left(\mu a\alpha_3(\gamma^x)-\frac{c\tau\mu}{2\Delta x\Delta y\mu_x\hat{\omega}_1}\right)u_h(x^-_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}}).
\end{aligned}
$$

From the above two groups we see that all coefficients of solution values involved are nonnegative if

$$a\alpha_3(\pm\gamma^x)-\frac{\Delta x|c|}{2\Delta y\hat{\omega}_1}\geq 0, \tag{A.1a}$$

$$\mu\leq\min\left\{\frac{\tilde{\omega}_i^{x,1}(\pm\gamma^x,y_{j\pm\frac{1}{2}})}{a\big(\alpha_1(\pm\gamma^x)+\alpha_3(\mp\gamma^x)\big)+\frac{\Delta x|c|}{\Delta y\hat{\omega}_1}},\frac{\tilde{\omega}_i^{x,2}(\gamma^x,y_{j\pm\frac{1}{2}})}{2\alpha_2(\gamma^x)}\right\}. \tag{A.1b}$$

Here we used $\tilde{\omega}_i^{x,3}(\gamma^x,y)=\tilde{\omega}_i^{x,1}(-\gamma^x,y)$ and $\alpha_2(\gamma^x)=\alpha_2(-\gamma^x)$.

Similarly, all coefficients of solution values in

$$H_2(x_{i-\frac{1}{2}})+\frac{\mu}{\mu_y\hat{\omega}_1}B_3=R_j(M(x_{i-\frac{1}{2}},\cdot),\mu,\Delta y,b)+\frac{\mu}{\mu_y\hat{\omega}_1}B_3$$

and

$$H_2(x_{i+\frac{1}{2}})+\frac{\mu}{\mu_y\hat{\omega}_1}B_4=R_j(M(x_{i+\frac{1}{2}},\cdot),\mu,\Delta y,b)+\frac{\mu}{\mu_y\hat{\omega}_1}B_4$$

are also nonnegative if

$$b\alpha_3(\pm\gamma^y)-\frac{\Delta y|c|}{2\Delta x\hat{\omega}_1}\geq 0, \tag{A.2a}$$

$$\mu\leq\min\left\{\frac{\tilde{\omega}_j^{y,1}(\pm\gamma^y,x_{i\pm\frac{1}{2}})}{b\big(\alpha_1(\pm\gamma^y)+\alpha_3(\mp\gamma^y)\big)+\frac{\Delta y|c|}{\Delta x\hat{\omega}_1}},\frac{\tilde{\omega}_j^{y,2}(\gamma^y,x_{i\pm\frac{1}{2}})}{2\alpha_2(\gamma^y)}\right\}. \tag{A.2b}$$

Here we used $\tilde{\omega}_j^{y,3}(\gamma^y,x)=\tilde{\omega}_j^{y,1}(-\gamma^y,x)$ and $\alpha_2(\gamma^y)=\alpha_2(-\gamma^y)$.

Since $\tilde{\omega}_i^{x,\sigma},\tilde{\omega}_j^{y,\sigma}$ for $\sigma=1,2,3$ only depend on $M(x,y)|_{K_{ij}}$ and γ, therefore bounded from below. We set such bound as

$$\underline{\omega}_{ij}=\min_{\gamma^x,\gamma^y,x\in\hat{S}_i^x,y\in\hat{S}_j^y}\{\tilde{\omega}_i^{x,1}(\pm\gamma^x,y),\tilde{\omega}_i^{x,2}(\gamma^x,y),\tilde{\omega}_j^{y,1}(\pm\gamma^y,x),\tilde{\omega}_j^{y,2}(\gamma^y,x)\}. \tag{A.3}$$

Notice also that for $\gamma=\max\{|\gamma^x|,|\gamma^y|\}$, using $\beta_1\leq 1/4$, we have

$$\alpha_1(\pm\gamma^x)+\alpha_3(\mp\gamma^x)=\beta_0+\frac{8\beta_1-2}{1\pm\gamma^x}\leq\beta_0+\frac{8\beta_1-2}{1-\gamma},$$

which also hods when γ^x is replaced by γ^y. Hence both (A.1b) and (A.2b) are ensured by (3.7).

Observe that (A.1a) and (A.2a) are implied by

$$\beta_0+\min_{s\in[-\gamma,\gamma]}\frac{8\beta_1-3+s}{2(1-s)}\geq\frac{|c|\kappa}{2\hat{\omega}_1\min\{a,b\}}.$$

Using $\beta_1\leq 1/4$ we see that the minimum on the left hand side is -1, obtained at $s=\gamma$ when $\gamma=8\beta_1-1$, hence this relation gives the lower bound in (3.6).

Appendix B. Proof of Theorem 3.2

For simplicity of presentation, in the following we consider $A=A(x,y)$ instead of $A(x,y,u)$, for which we only need to use $\{A_h\}$ on interfaces, so that $c=c(x,y)$, $a=a(x,y)$ and $b=b(x,y)$.

Since $u_h(x, y)$ is quadratic in terms of x and y respectively, we can use the formula (2.8) to obtain

$$\begin{aligned}\dot{p}(\eta) &= \dot{\omega}^1(\eta)p(-1) + \dot{\omega}^2(\eta)(\gamma) + \dot{\omega}^3(\eta)p(1)\\ &= \frac{2\eta - 1 - \gamma}{2(1+\gamma)}p(-1) + \frac{2\eta}{(\gamma^2-1)}p(\gamma) + \frac{2\eta + 1 - \gamma}{2(1-\gamma)}p(1).\end{aligned}$$

Here η is the variable in the reference element $[-1, 1]$, and $\dot{\omega}^\sigma(\eta) = \frac{d}{d\eta}\omega^\sigma(\eta)$. Then

$$\begin{aligned}\int_{J_j} c(x, y)\{\partial_y u_h(x, y)\}dy = \{u_h(x, y^+_{j-\frac{1}{2}})\}&\int_{-1}^{1} c\left(x, y_j + \frac{\Delta y}{2}\eta\right)\dot{\omega}^1(\eta)\,d\eta\\ &+ \{u_h(x, y^\gamma_j)\}\int_{-1}^{1} c\left(x, y_j + \frac{\Delta y}{2}\eta\right)\dot{\omega}^2(\eta)\,d\eta\\ &+ \{u_h(x, y^-_{j+\frac{1}{2}})\}\int_{-1}^{1} c\left(x, y_j + \frac{\Delta y}{2}\eta\right)\dot{\omega}^3(\eta)\,d\eta.\end{aligned}$$

Here the average $\{\cdot\}$ is taken with respect to x^- and x^+. Using the quadrature rule on the right-hand side of (3.9), we obtain the following scheme

$$\bar{u}^{n+1}_{ij} = \frac{\mu_x}{\mu}\sum_{\sigma=1}^{L}\hat{\omega}^\sigma H_1(\hat{y}^\sigma_j) + \frac{\mu_y}{\mu}\sum_{\sigma=1}^{L}\hat{\omega}^\sigma H_2(\hat{x}^\sigma_i) + B. \tag{B.1}$$

The test set now reduces to

$$\begin{aligned}S^x_i &= \{x_{i-\frac{1}{2}}, x^\gamma_i, x_{i+\frac{1}{2}}\} = x_i + \frac{\Delta x}{2}\{-1, \gamma, 1\},\\ S^y_j &= \{y_{j-\frac{1}{2}}, y^\gamma_j, y_{j+\frac{1}{2}}\} = y_j + \frac{\Delta y}{2}\{-1, \gamma, 1\}\end{aligned}$$

with γ satisfying

$$|\gamma| \le \frac{1}{3} \text{ and } |\gamma| \le 8\beta_1 - 1. \tag{B.2}$$

It can be shown that there exists $L \ge \frac{2k+3}{2}$ and γ satisfying (B.2) such that $x^\gamma_i \in \hat{S}^x_i$ and $y^\gamma_j \in \hat{S}^y_j$, the corresponding quadrature weight is denoted by $\hat{\omega}^*$. Upon regrouping, the right-hand side of (B.1) may be decomposed as

$$\begin{aligned}\bar{u}^{n+1}_{ij} = &\frac{\mu_x}{\mu}\sum_{\sigma\neq 1,*,L}\hat{\omega}^\sigma H_1(\hat{y}^\sigma_j) + \frac{\mu_y}{\mu}\sum_{\sigma\neq 1,*,L}\hat{\omega}^\sigma H_2(\hat{x}^\sigma_i)\\ &+ \frac{\mu_x}{\mu}\left(\hat{\omega}^1 H_1(y_{j-\frac{1}{2}}) + \hat{\omega}^* H_1(y^\gamma_j) + \hat{\omega}^1 H_1(y_{j+\frac{1}{2}})\right)\\ &+ \frac{\mu_y}{\mu}\left(\hat{\omega}^1 H_2(x_{i-\frac{1}{2}}) + \hat{\omega}^* H_2(x^\gamma_i) + \hat{\omega}^1 H_2(x_{i+\frac{1}{2}})\right) + B.\end{aligned}$$

Here $\hat{\omega}^1 = \hat{\omega}^L$ is used. The terms in B will be regrouped correspondingly. More precisely, the first term involving integration on J_j is regrouped in terms involving solution values at $y^+_{j-\frac{1}{2}}, y^\gamma_j$ and $y^-_{j+\frac{1}{2}}$, and combined with $H_1(y_{j-\frac{1}{2}}), H_1(y^\gamma_j)$ and $H_1(y_{j+\frac{1}{2}})$, respectively. We check upon the first group only:

$$\begin{aligned}&H_1(y_{j-\frac{1}{2}}) + \frac{\mu}{\mu_x\hat{\omega}^1}B_1 = R_i(M(\cdot, y_{j-\frac{1}{2}}), \mu, \Delta x, a) + \frac{\mu}{\mu_x\hat{\omega}^1}B_1\\ &= \left[\omega^1 - \mu a\big(\alpha_3(-\gamma) + \alpha_1(\gamma)\big) - \frac{\mu\Delta x}{2\Delta y\hat{\omega}^1}\int_{-1}^{1} c\left(x_{i-\frac{1}{2}}, y_j + \frac{\Delta y}{2}\eta\right)\dot{\omega}^1(\eta)\,d\eta\right]u_h(x^+_{i-\frac{1}{2}}, y^+_{j-\frac{1}{2}})\\ &\quad + \left[\omega^2 - \mu a\big(\alpha_2(-\gamma) + \alpha_2(\gamma)\big)\right]u_h(x^\gamma_i, y^+_{j-\frac{1}{2}})\end{aligned}$$

$$
\begin{aligned}
&+\left[\omega^3-\mu a\big(\alpha_1(-\gamma)+\alpha_3(\gamma)\big)+\frac{\mu\Delta x}{2\Delta y\hat{\omega}^1}\int_{-1}^{1}c\left(x_{i+\frac{1}{2}},y_j+\frac{\Delta y}{2}\eta\right)\dot{\omega}^1(\eta)\,d\eta\right]u_h(x^-_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})\\
&+\left(\mu a\alpha_3(-\gamma)+\frac{\mu\Delta x}{2\Delta y\hat{\omega}^1}\int_{-1}^{1}c\left(x_{i+\frac{1}{2}},y_j+\frac{\Delta y}{2}\eta\right)\dot{\omega}^1(\eta)\,d\eta\right)u_h(x^+_{i+\frac{1}{2}},y^+_{j-\frac{1}{2}})\\
&+\mu a\left[\alpha_2(-\gamma)u_h(x^{\gamma}_{i+1},y^+_{j-\frac{1}{2}})+\alpha_1(-\gamma)u_h(x^-_{i+\frac{3}{2}},y^+_{j-\frac{1}{2}})\right]\\
&+\mu a\left[\alpha_1(\gamma)u_h(x^+_{i-\frac{3}{2}},y^+_{j-\frac{1}{2}})+\alpha_2(\gamma)u_h(x^{\gamma}_{i-1},y^+_{j-\frac{1}{2}})\right]\\
&+\left(\mu a\alpha_3(\gamma)-\frac{\mu\Delta x}{2\Delta y\hat{\omega}^1}\int_{-1}^{1}c\left(x_{i-\frac{1}{2}},y_j+\frac{\Delta y}{2}\eta\right)\dot{\omega}^1(\eta)\,d\eta\right)u_h(x^-_{i-\frac{1}{2}},y^+_{j-\frac{1}{2}}).
\end{aligned}
$$

From the three groups of $H_1(y_{j-\frac{1}{2}})$, $H_1(y^{\gamma}_j)$ and $H_1(y_{j+\frac{1}{2}})$, we see that all the coefficients of solutions values are nonnegative if

$$
\begin{aligned}
&a\alpha_3(\pm\gamma)-\frac{\Delta x}{2\Delta y\min\{\hat{\omega}_1,\hat{\omega}^*\}}\max_{x\in S^x_i;\sigma}\left|g^{\sigma}_j(x)\right|\geq 0,\\
&\mu\leq\min\left\{\frac{\min\{\omega^1,\omega^3\}}{a\big(\alpha_1(\pm\gamma)+\alpha_3(\mp\gamma)\big)+\frac{\Delta x}{2\Delta y\min\{\hat{\omega}_1,\hat{\omega}^*\}}\max_{x\in S^x_i;\sigma}\left|g^{\sigma}_j(x)\right|},\frac{\omega^2}{2a\alpha_2(\gamma)}\right\},
\end{aligned}
$$

where $g^{\sigma}_j(x)=\int_{-1}^{1}c\left(x,y_j+\frac{\Delta y}{2}\eta\right)\dot{\omega}^{\sigma}(\eta)\,d\eta$. The rest of terms in B when combined with $H_2(y)$ for $y\in S^y_j$ led to the following conditions

$$
\begin{aligned}
&b\alpha_3(\pm\gamma)-\frac{\Delta y}{2\Delta x\min\{\hat{\omega}_1,\hat{\omega}^*\}}\max_{y\in S^y_j;\sigma}\left|\int_{-1}^{1}c\left(x_i+\frac{\Delta x}{2}\xi,y\right)\dot{\omega}^{\sigma}(\xi)\,d\xi\right|\geq 0,\\
&\mu\leq\min\left\{\frac{\min\{\omega^1,\omega^3\}}{b\big(\alpha_1(\pm\gamma)+\alpha_3(\mp\gamma)\big)+\frac{\Delta y}{2\Delta x\min\{\hat{\omega}_1,\hat{\omega}^*\}}\max_{y\in S^y_j;\sigma}\left|g^{\sigma}_i(y)\right|},\frac{\omega^2}{2a\alpha_2(\gamma)}\right\},
\end{aligned}
$$

where $g^{\sigma}_i(y)=\int_{-1}^{1}c(x_i+\frac{\Delta x}{2}\xi,y))\dot{\omega}^{\sigma}(\xi)\,d\xi$. Note that both $|\int_{-1}^{1}c(x,y_j+\frac{\Delta y}{2}\eta)\dot{\omega}^{\sigma}(\eta)\,d\eta|$ and $|\int_{-1}^{1}c(x_i+\frac{\Delta x}{2}\xi,y))\dot{\omega}^{\sigma}(\xi)\,d\xi|$ are bounded from above by

$$
\|c\|_{\infty}\left|\int_{-1}^{1}\dot{\omega}^{\sigma}(\eta)d\eta\right|\leq\frac{4\|c\|_{\infty}}{1-\gamma}.
$$

Using the fact that $\kappa^{-1}\leq\frac{\Delta x}{\Delta y}\leq\kappa$, $\omega^3\leq\omega^1$, $\frac{1}{L(L-1)}=\hat{\omega}_1<\hat{\omega}^*$, and $\alpha_3(\pm\gamma)\geq\beta_0-1$, as well as

$$
\alpha_1(\pm\gamma)+\alpha_3(\mp\gamma)\leq\beta_0+\frac{8\beta_1-2}{1+\gamma}.
$$

We see that all the above needed constraints are implied by the more severe constraints, including (3.10) for β_0, and the CFL condition (3.11).

References

[1] F. Bassi, S. Rebay, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations, J. Comput. Phys. 131 (2) (1997) 267–279.

[2] Z. Chen, H. Huang, J. Yan, Third order maximum-principle-satisfying direct discontinuous Galerkin methods for time dependent convection diffusion equations on unstructured triangular meshes, J. Comput. Phys. 308 (2016) 198–217.

[3] W.-X. Cao, H. Liu, Z.-M. Zhang, Superconvergence of the direct discontinuous Galerkin method for convection-diffusion equations, Numer. Methods Partial Differ. Equ. 33 (2017) 290–317.

[4] Z. Chai, B. Shi, Z. Guo, A multip-relaxation-time lattice Boltzmann model for general nonlinear anisotropic convection–diffusion equations, J. Comput. Sci. 69 (2016) 355–390.

[5] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection–diffusion systems, SIAM J. Numer. Anal. 35 (6) (1998) 2440–2463.
[6] J. Cheng, X. Yang, X. Liu, T.-G. Liu, H. Luo, A direct discontinuous Galerkin method for the compressible Navier-Stokes equations on arbitrary grids, J. Comput. Phys. 327 (2016) 484–502.
[7] J. Du, Y. Yang, Maximum-principle-preserving third-order local discontinuous Galerkin method for convection-diffusion equations on overlapping meshes, J. Comput. Phys. 377 (2019) 117–141.
[8] H. Fujii, Some remarks on finite element analysis of time-dependent field problems, in: Theory and Practice in Finite Element Structural Analysis, University of Tokyo Press, Tokyo, 1973, pp. 91–106.
[9] I. Faragó, R. Horváth, Discrete maximum principle and adequate discretizations of linear parabolic problems, SIAM J. Sci. Comput. 28 (2006) 2313–2336.
[10] I. Faragó, R. Horváth, S. Korotov, Discrete maximum principle for linear parabolic problems solved on hybrid meshes, Appl. Numer. Math. 53 (2005) 249–264.
[11] I. Faragó, J. Karátson, S. Korotov, Discrete maximum principles for nonlinear parabolic PDE systems, IMA J. Numer. Anal. 32 (4) (2012) 1541–1573.
[12] S. Gottlieb, D.I. Ketcheson, C.-W. Shu, High order strong stability preserving time discretizations, J. Sci. Comput. 38 (2009) 251–289.
[13] S. Gottlieb, D.I. Ketcheson, C.-W. Shu, Strong Stability Preserving Runge–Kutta and Multistep Time Discretizations, World Scientific, 2011.
[14] J.S. Hesthaven, T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, 1st edition, Springer Publishing Company, Incorporated, 2007.
[15] Y. Jiang, Z. Xu, Parametrized maximum principle preserving limiter for finite difference WENO schemes solving convection-dominated diffusion equations, SIAM J. Sci. Comput. 35 (2013) A2524–A2553.
[16] A. Kurganov, E. Tadmor, New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations, J. Comput. Phys. 160 (1) (2000) 241–282.
[17] R.J. LeVeque, Numerical Methods for Conservation Laws, Birkhäuser, Basel, 1992.
[18] H. Liu, Optimal error estimates of the direct discontinuous Galerkin method for convection–diffusion equations, Math. Comput. 84 (2015) 2263–2295.
[19] H. Liu, Z.-M. Wang, An entropy satisfying discontinuous Galerkin method for nonlinear Fokker–Planck equations, J. Comput. Sci. 68 (2016) 1217–1240.
[20] H. Liu, Z.-M. Wang, A free energy satisfying discontinuous Galerkin method for Poisson–Nernst–Planck systems, J. Comput. Phys. 238 (2017) 413–437.
[21] H. Liu, J. Yan, The direct discontinuous Galerkin (DDG) methods for diffusion problems, SIAM J. Numer. Anal. 47 (1) (2009) 675–698.
[22] H. Liu, J. Yan, The direct discontinuous Galerkin (DDG) method for diffusion with interface corrections, Commun. Comput. Phys. 8 (3) (2010) 541–564.
[23] H. Liu, H. Yu, Maximum-principle-satisfying third order discontinuous Galerkin schemes for Fokker–Planck equations, SIAM J. Sci. Comput. 36 (5) (2014) A2296–A2325.
[24] H. Liu, H. Yu, The entropy satisfying discontinuous Galerkin method for Fokker–Planck equations, J. Sci. Comput. 62 (3) (2015) 803–830.
[25] A. Mizukami, T.J. Hughes, A Petrov-Galerkin finite element method for convection-dominated flows: an accurate upwinding technique for satisfying the maximum principle, Comput. Methods Appl. Mech. Eng. 50 (1985) 181–193.
[26] B. Riviére, Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations, Frontiers in Applied Mathematics, SIAM, 2008.
[27] C.-W. Shu, Discontinuous Galerkin methods: general approach and stability, in: Numerical Solutions of Partial Differential Equations, in: Adv. Courses Math. CRM Barcelona, Birkhauser, Basel, 2009, pp. 149–201.
[28] Z. Sun, J.A. Carrillo, C.-W. Shu, A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials, J. Comput. Phys. 382 (2018) 76–104.
[29] S. Srinivasana, J. Poggiea, X. Zhang, A positivity-preserving high order discontinuous Galerkin scheme for convection–diffusion equations, J. Comput. Phys. 366 (2018) 120–143.
[30] V. Thomee, L.B. Wahlbin, On the existence of maximum principles in parabolic finite element equations, Math. Comput. 77 (2008) 11–19.
[31] T. Vejchodský, S. Korotov, A. Hannukainen, Discrete maximum principle for parabolic problems solved by prismatic finite elements, Math. Comput. Simul. 80 (2010) 1758–1770.
[32] T. Xiong, J.-M. Qiu, Z. Xu, High order maximum-principle-preserving discontinuous Galerkin method for convection–diffusion equations, SIAM J. Sci. Comput. 37 (2) (2015) A583–A608.
[33] P. Yang, T. Xiong, J.-M. Qiu, Z. Xu, High order maximum principle preserving finite volume method for convection dominated problems, J. Sci. Comput. 67 (2) (2016) 795–820.
[34] X. Zhang, On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations, J. Comput. Phys. 328 (2017) 301–343.
[35] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, J. Comput. Phys. 229 (9) (2010) 3091–3120.
[36] X. Zhang, C.-W. Shu, Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 467, The Royal Society, 2011, pp. 2752–2776.
[37] X. Zhang, Y. Liu, C. Shu, Maximum-principle-satisfying high order finite volume weighted essentially nonoscillatory schemes for convection-diffusion equations, SIAM J. Sci. Comput. 34 (2) (2012) A627–A658.
[38] Y. Zhang, X. Zhang, C.-W. Shu, Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection-diffusion equations on triangular meshes, J. Comput. Phys. 234 (2013) 295–316.