Journal of Computational Physics 406 (2020) 109219

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

High-order bound-preserving finite difference methods for miscible displacements in porous media ☆

Hui Guo [a], Xinyuan Liu [a], Yang Yang [b,*]

[a] College of Science, China University of Petroleum, Qingdao 266580, China
[b] Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931, United States of America

ARTICLE INFO

Article history:
Received 1 September 2019
Received in revised form 13 December 2019
Accepted 22 December 2019
Available online 31 December 2019

Keywords:
Miscible displacements
Bound-preserving
High-order
Finite difference method
Multi-component fluid
Flux limiter

ABSTRACT

In this paper, we develop high-order bound-preserving (BP) finite difference (FD) methods for the coupled system of compressible miscible displacements. We consider the problem with multi-component fluid mixture and the (volumetric) concentration of the jth component, c_j, should be between 0 and 1. It is well known that c_j does not satisfy a maximum-principle. Hence most of the existing BP techniques cannot be applied directly. The main idea in this paper is to construct the positivity-preserving techniques to all $c_j's$ and enforce $\sum_j c_j = 1$ simultaneously to obtain physically relevant approximations. By doing so, we have to treat the time derivative of the pressure dp/dt as a source in the concentration equation and choose suitable "consistent" numerical fluxes in the pressure and concentration equations. Recently, the high-order BP discontinuous Galerkin (DG) methods for miscible displacements were introduced in [4]. However, the BP technique for DG methods is not straightforward extendable to high-order FD schemes. There are two main difficulties. Firstly, it is not easy to determine the time step size in the BP technique. In finite difference schemes, we need to choose suitable time step size first and then apply the flux limiter to the numerical fluxes. Subsequently, we can compute the source term in the concentration equation, leading to a new time step constraint that may not be satisfied by the time step size applied in the flux limiter. Therefore, it would be very difficult to determine how large the time step is. Secondly, the general treatment for the diffusion term, e.g. centered difference, in miscible displacements may require a stencil whose size is larger than that for the convection term. It would be better to construct a new spatial discretization for the diffusion term such that a smaller stencil can be used. In this paper, we will solve both problems. We first construct a special discretization of the convection term, which yields the desired approximations of the source. Then we can find out the time step size that suitable for the BP technique and apply the flux limiters. Moreover, we will also construct a special algorithm for the diffusion term whose stencil is the same as that used for the convection term. Numerical experiments will be given to demonstrate the high-order accuracy and good performance of the numerical technique.

☆ The first author was supported by National Natural Science Foundation of China Grant 11571367 and the Fundamental Research Funds for the Central Universities 18CX05003A. The last author was supported by NSF grant DMS-1818467.

* Corresponding author.
E-mail addresses: sdugh@163.com (H. Guo), s18090001@s.upc.edu.cn (X. Liu), yyang7@mtu.edu (Y. Yang).

https://doi.org/10.1016/j.jcp.2019.109219
0021-9991/

1. Introduction

In this paper, we are interested in constructing high-order bound-preserving (BP) finite difference (FD) schemes for compressible miscible displacements in porous media. We consider the fluid mixture with N components and the governing equations over the computational domain $\Omega = [0,1]\times[0,1]$ read

$$d(\mathbf{c})\frac{\partial p}{\partial t} + \nabla\cdot\mathbf{u} = d(\mathbf{c})\frac{\partial p}{\partial t} - \nabla\cdot\left(\frac{\kappa(x,y)}{\mu(\mathbf{c})}\nabla p\right) = q,\ \ (x,y)\in\Omega,\ \ 0<t\le T, \tag{1.1}$$

$$\phi\frac{\partial c_j}{\partial t} + \nabla\cdot(\mathbf{u}c_j) - \nabla\cdot(\mathbf{D}\nabla c_j) = \tilde{c}_j q - \phi c_j z_j p_t,\ \ (x,y)\in\Omega,\ \ 0<t\le T,\ \ j=1,\cdots,N-1. \tag{1.2}$$

Here the dependent variables are the pressure in fluid mixture denoted by p, the Darcy velocity of the mixture (volume flowing across a unit across-section per unit time) denoted by $\mathbf{u} = (u,v)^T$ and the concentration of interested species measured in amount of species per unit volume denoted by $\mathbf{c} = (c_1,\cdots,c_N)^T$, with c_j being the concentration of the jth component. ϕ and κ are the porosity and permeability of the rock, respectively. μ refers to the concentration-dependent viscosity. q is the external volumetric flow rate, and $\tilde{c}_j$ is the concentration of the fluid in the external flow. $\tilde{c}_j$ must be specified at points where injection ($q>0$) takes place, and is assumed to be equal to c_j at production points ($q<0$). The diffusion coefficient $\mathbf{D}$ is symmetric and arises from two aspects: molecular diffusion and dispersion. Its form is

$$\mathbf{D} = \phi(x,y)(d_{\text{mol}}\mathbf{I} + d_{\text{long}}|\mathbf{u}|\mathbf{E} + d_{\text{tran}}|\mathbf{u}|\mathbf{E}^{\perp}), \tag{1.3}$$

where $\mathbf{E}$, a 2×2 matrix, represents the orthogonal projection along the velocity vector given as

$$\mathbf{E} = (e_{ij}(\mathbf{u})) = \left(\frac{u_i u_j}{|\mathbf{u}|^2}\right),\quad \mathbf{u} = (u_1,u_2),$$

and $\mathbf{E}^{\perp} = \mathbf{I} - \mathbf{E}$ is the orthogonal complement. The diffusion coefficient d_{long} measures the dispersion in the direction of the flow and d_{tran} shows that transverse to the flow. To ensure the stability of the scheme, $\mathbf{D}$ is assumed to be strictly positive definite in almost all of the previous works. If the flow vectors are essentially parallel to the x-axis then the dispersion term $\nabla(\mathbf{D}\nabla c)$ could be replaced by the sum [20,10]

$$d_{\text{long}}c_{xx} + d_{\text{tran}}c_{yy}.$$

Further, if we consider molecular diffusion only then $\mathbf{D} = \phi(x,y)d_{\text{mol}}\mathbf{I}$ [33,9]. In this paper, we consider the above two simplified cases and assume $\mathbf{D}$ to be a diagonal matrix. Moreover, the pressure is uniquely determined up to a constant, thus we assume $\int_\Omega p\ dxdy = 0$ at $t=0$. However, this assumption is not essential. Other coefficients can be stated as follows:

$$c_N = 1 - \sum_{j=1}^{N-1} c_j,\quad d(\mathbf{c}) = \phi\sum_{j=1}^{N} z_j c_j,$$

where z_j is the compressibility factor of the jth component of the fluid mixture. With the identity given above, we can sum (1.2) over j and subtract which from (1.1) to obtain a ghost equation satisfied by c_N as

$$\phi\frac{\partial c_N}{\partial t} + \nabla\cdot(\mathbf{u}c_N) - \nabla\cdot(\mathbf{D}\nabla c_N) = \tilde{c}_N q - \phi c_N z_N p_t,\ \ (x,y)\in\Omega,\ \ 0<t\le T. \tag{1.4}$$

We can see that c_N satisfies the same equation as those for $c_j's$, $1\le j\le N-1$. Therefore, the positivity-preserving technique for (1.2) should also work for (1.4). In this paper, we consider periodic boundary condition for simplicity. Moreover, the initial solutions are given as

$$c_j(x,y,0) = c_{j_0}(x,y),\quad p(x,y,0) = p_0(x,y),\quad (x,y)\in\Omega.$$

The miscible displacements in porous media were first presented in [7,8], where mixed finite element methods were applied. Later, the compressible problem was studied in [9] and the optimal order estimates in L^2-norm and almost optimal order estimates in L^∞-norm were given in [3]. Subsequently, many new numerical methods were introduced, such as the finite difference method [34–36], characteristic finite element method [19], splitting positive definite mixed element method [29] and H1-Galerkin mixed method [2]. Besides the above, in [25], an accurate and efficient simulator was developed for problems with wells. Later, the authors introduced an Eulerian-Lagrangian localized adjoint method to solve the transport equation for concentration, while a mixed finite element method to solve the pressure equation [24]. Moreover, the discontinuous Galerkin (DG) methods were also introduced in [5,6,30,31,14,32,33]. However, none of the works given above discussed the BP techniques. In many numerical simulations, the approximations of c_j can be placed out of the interval $[0,1]$. Especially for problems with large gradients, the value of $d(\mathbf{c})$ might be negative, leading to ill-posedness of

the problem, and the numerical approximations will blow up. Therefore, the BP technique is crucial in constructing reliable numerical approximations.

Recently, two of the authors in this paper applied DG methods for compressible miscible displacements in porous media and constructed the second-order BP technique in [13] and the extension to high-order schemes was also discussed in [4]. The ideas introduced in [13,4] are not straightforward extendable to FD schemes. Before we demonstrate the difficulties, we would like to review the BP technique for the DG methods. For simplicity, we only consider Euler forward time discretization. Given the numerical approximations at time level n, say c^n and p^n, we need to construct physically relevant numerical approximations at time level $n+1$, namely c^{n+1} and p^{n+1}. The algorithm can be demonstrated as follows:

1. Solve p_t from (1.1) and use which as the source in (1.2).
2. Discretize (1.1) and (1.2) with Euler forward time discretization. Sum up the scheme for (1.2) over j then subtract which from that for (1.1) to obtain the numerical scheme for (1.4). We want the scheme for (1.4) to be basically the same as that for (1.2).
3. Choose the "consistent" numerical fluxes (see Definition 2.1) to the convection term and write

$$\widehat{uc} = \hat{u}\hat{c} - \alpha[c], \tag{1.5}$$

   where α is the penalty parameter to be chosen by the BP technique, and $[c]$ is the jump of c across the cell interfaces.
4. Investigate the time step requirements in the BP technique [37,40,38,39] for the convection term and source term.
5. Extend the second-order BP technique [41] to the diffusion part.
6. Apply the flux limiters [15,27,28] to the numerical fluxes in the convection and diffusion terms. High and low-order numerical approximations of c shall be used in the flux limiter while only high-order approximations of u are considered.

The flow chart of the algorithm can be summarized as follows

$$\left.\begin{array}{c}\{c_H^n, p^n\} \to u, p_t \to \Delta t \\ c_H^n \to c_L^n\end{array}\right\} \to \{c_H^{n+1}, p^{n+1}\},$$

where the subscripts H and L represent the high-order and low-order numerical approximations in the flux limiters, respectively. Now, we are ready to demonstrate why the algorithm given above does not work for FD methods. The main reason is that for FD method it is very difficult to construct high-order numerical approximations if we write the numerical fluxes for the convection term as (1.5). Actually, $\widehat{uc}$ is obtained via the reconstruction procedure based on the numerical approximations of uc at the grid points in the stencil. Though, we can construct $\hat{u}$ in (1.1), it is not useful for $\widehat{uc}$ if we have applied the WENO algorithm which requires nonlinear weights in the reconstruction procedure. Secondly, in the flux limiters, we need high-order and low-order numerical fluxes, which are based on stencils with different sizes. Hence, the numerical fluxes may not be "consistent" with $\hat{u}$. Though, we can modify the numerical fluxes $\hat{u}$ to make it to be "consistent" with $\widehat{uc}$, the modification may yield a new numerical approximations of p_t, leading to a new time step constraint which may not be satisfied by the time step chosen in the flux limiter used for the convection terms. The paradox can be summarized as follows:

$$\left.\begin{array}{c}\{c_H^n, p^n\} \to u, p_t \to \Delta t \\ c_H^n \to c_L^n\end{array}\right\} \to \widehat{uc} \to \hat{u} \to u^{new}, p_t^{new} \to \Delta t^{new}$$

Numerical experiments demonstrated that Δt can be greater than Δt^{new}. Therefore, it is very difficult to estimate the time step Δt in the flux limiters. The second minor issue (compared with the time step paradox) is the discretization of the diffusion term. Most of the previous BP techniques based on FD methods work for the diffusion term in the form of $\Delta A(c)$, see e.g. [17]. However, with centered difference, the stencil for the diffusion term in miscible displacements could be larger than that for the convection term. To solve these two problems, we would like to introduced the following algorithms:

1. Use the same stencil to reconstruct both the high-order and low-order numerical fluxes for the convection term. To construct the high-order numerical fluxes for the convection term, we first choose a stencil I. Then apply the (WENO) reconstruction procedure to compute $\widehat{uc}^H$ based on the numerical approximations of uc within the stencil, see [22,23,16,18,1] for more details about the reconstruction procedure. To compute the low-order numerical fluxes $\widehat{uc}^L$, we would like to use the same stencil I, but change the values of c to be the first order approximations. Then we can calculate $\hat{u}$ based on $\widehat{uc}$ such that they are consistent, no matter whether we use high-order or low-order numerical fluxes. Therefore, high-order and low-order numerical schemes will yield the same $\hat{u}$, leading to the same approximation of p_t. The algorithm can be summarized as follows:

$$\{p^n, c^n\} \to u \to \widehat{uc}^H \to \left\{\begin{array}{c}\widehat{uc}^L \\ \hat{u} \to p_t\end{array}\right\} \to \Delta t \to \{p^{n+1}, c^{n+1}\}.$$

2. Apply the algorithm based on Taylor's expansion introduced in [23] to construct the numerical fluxes for the diffusion term. The most commonly used spatial discretization of the diffusion term is the centered difference. However, the stencil can be larger than that for the convection term. With the idea of Taylor's expansion [23], the stencil size is the same as that for the convection term, and the algorithm keeps the high-order accuracy. Different from the reconstruction procedure for the convection term, we do not need to use the same stencil for the high-order and low-order numerical fluxes, since the diffusion term in the concentration equation does not appear in the pressure equation.

The rest of the paper is organized as follows: we first discuss the FD scheme in two space dimensions in Section 2. In Section 3, we demonstrate the high-order BP technique in one space dimension. The extension to two space dimensions will be given in section 4. In Section 5, some numerical experiments will be provided. We will end in Section 6 with concluding remarks. The reconstruction procedure will be given in Appendix A.

2. The finite difference scheme

In this section, we will construct the FD scheme for compressible miscible displacements in porous media. We consider rectangular meshes and define the grid points as

$$0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N_x-\frac{1}{2}} < x_{N_x+\frac{1}{2}} = 1, \tag{2.1}$$

$$0 = y_{\frac{1}{2}} < y_{\frac{3}{2}} < \cdots < y_{N_y-\frac{1}{2}} < y_{N_y+\frac{1}{2}} = 1. \tag{2.2}$$

Moreover, we denote

$$I_{i,k} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{k-\frac{1}{2}}, y_{k+\frac{1}{2}}], \tag{2.3}$$

as the rectangular cells. The grid centers and grid size are given as

$$x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}), \quad \Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, \quad i = 1, \cdots, N_x, \tag{2.4}$$

$$y_k = \frac{1}{2}(y_{k-\frac{1}{2}} + y_{k+\frac{1}{2}}), \quad \Delta y_k = y_{k+\frac{1}{2}} - y_{k-\frac{1}{2}}, \quad k = 1, \cdots, N_y. \tag{2.5}$$

In this paper, we consider uniform partition only, and denote $\Delta x = \Delta x_i$ and $\Delta y = \Delta y_k$. This assumption is essential in the reconstruction procedure [22] of the high-order numerical fluxes.

To construct the FD scheme, we first rewrite the system (1.1)-(1.2) into the following form

$$d(\mathbf{c})p_t + \nabla \cdot \mathbf{u} = q, \tag{2.6}$$

$$a(\mathbf{c})\mathbf{u} = -\nabla p, \tag{2.7}$$

$$(\phi c_j)_t + \nabla \cdot (\mathbf{u}c_j) - \nabla \cdot (\mathbf{D}(\mathbf{u})\nabla c_j) = \tilde{c}_j q - \phi c_j z_j p_t, \quad j = 1, 2, \cdots, N-1, \tag{2.8}$$

where $a(\mathbf{c}) = \dfrac{\mu(\mathbf{c})}{\kappa}$.

For simplicity, we denote $(\cdot)_{i,k}$ as the numerical approximation at (x_i, y_k). Then the semi-discretized FD scheme can be written as

$$d(\mathbf{c}_{i,k})\frac{d}{dt}p_{i,k} = -\frac{1}{\Delta x}(\hat{u}_{i+\frac{1}{2},k} - \hat{u}_{i-\frac{1}{2},k}) - \frac{1}{\Delta y}(\hat{v}_{i,k+\frac{1}{2}} - \hat{v}_{i,k-\frac{1}{2}}) + q_{i,k}, \tag{2.9}$$

$$a(\mathbf{c}_{i,k})u_{i,k} = -\frac{1}{\Delta x}(\hat{p}_{i+\frac{1}{2},k} - \hat{p}_{i-\frac{1}{2},k}), \tag{2.10}$$

$$a(\mathbf{c}_{i,k})v_{i,k} = -\frac{1}{\Delta y}(\hat{p}_{i,k+\frac{1}{2}} - \hat{p}_{i,k-\frac{1}{2}}), \tag{2.11}$$

$$\begin{aligned}\frac{d}{dt}(r_j)_{i,k} = &-\frac{1}{\Delta x}(\widehat{uc}_{j\,i+\frac{1}{2},k} - \widehat{uc}_{j\,i-\frac{1}{2},k}) - \frac{1}{\Delta y}(\widehat{vc}_{j\,i,k+\frac{1}{2}} - \widehat{vc}_{j\,i,k-\frac{1}{2}}) \\ &+ \frac{1}{\Delta x}(\hat{H}_{i+\frac{1}{2},k} - \hat{H}_{i-\frac{1}{2},k}) + \frac{1}{\Delta y}(\hat{G}_{i,k+\frac{1}{2}} - \hat{G}_{i,k-\frac{1}{2}}) + (\tilde{c}_j)_{i,k}q_{i,k} - (r_j)_{i,k}z_j\frac{d}{dt}p_{i,k},\end{aligned} \tag{2.12}$$

where $r_j = \phi c_j$. Here, $\hat{u}_{i-\frac{1}{2},k}$ is the numerical fluxes at $(x_{i-\frac{1}{2}}, y_k)$ and is used for the spatial derivative along x-axis. Likewise for the other numerical fluxes.

Next, we proceed to demonstrate the reconstruction procedure for the numerical flux $\widehat{uc}_j$. We choose a stencil $I = \{x_{i-r}, \cdots, x_{i+1+s}\}$, where $r + s + 1 = m$ is the order of accuracy of the scheme. In general, we want $r = s$. We apply the flux splitting [22] to the scheme and write

$$\widehat{uc}_{j_{i+\frac{1}{2},k}} = \frac{1}{2}\left(\widehat{uc}^{+}_{j_{i+\frac{1}{2},k}} + \widehat{uc}^{-}_{j_{i+\frac{1}{2},k}}\right), \tag{2.13}$$

where

$$\widehat{uc}^{+}_{j_{i+\frac{1}{2},k}} = f^{+}((uc_j)_{i-r,k} + \alpha(c_j)_{i-r,k}, \cdots, (uc_j)_{i+s,k} + \alpha(c_j)_{i+s,k}),$$

and

$$\widehat{uc}^{-}_{j_{i+\frac{1}{2},k}} = f^{-}((uc_j)_{i+1-r,k} - \alpha(c_j)_{i+1-r,k}, \cdots, (uc_j)_{i+1+s,k} - \alpha(c_j)_{i+1+s,k}),$$

with parameter α being determined by the bound-preserving technique. The reconstruction function $f^{\pm}$ has been well-developed in [22]. We will demonstrate the reconstruction procedure in Appendix A and only demonstrate the following useful property:

$$f^{+}(a_{-r}, \cdots, a_s) = \sum_{\ell=-r}^{s} \omega_\ell^{+} a_\ell, \tag{2.14}$$

$$f^{-}(a_{1-r}, \cdots, a_{1+s}) = \sum_{\ell=1-r}^{s+1} \omega_\ell^{-} a_\ell, \tag{2.15}$$

where the weights $\omega_\ell^{\pm}$ can be obtained by the reconstruction procedure introduced in [22]. We can use either linear weights or nonlinear weights with WENO reconstruction. Moreover, it is easy to verify that

$$\sum_{\ell=-r}^{s} \omega_\ell^{+} = \sum_{\ell=1-r}^{s+1} \omega_\ell^{-} = 1. \tag{2.16}$$

For the diffusion part, we drop the subindex j for simplicity and consider the construction of $\hat{H}_{i+\frac{1}{2},k}$ only. The numerical flux $\hat{G}_{i,k+\frac{1}{2}}$ can be obtained with some minor changes. We assume $\mathbf{D} = (D_{ij})$, $i, j = 1, 2$ and apply the Taylor's expansion introduced in [23] to write

$$\hat{H}_{i+\frac{1}{2},k} = D_{11}c_x|_{x_{i+\frac{1}{2}},y_k} - \frac{\Delta x^2}{24}\frac{\partial^2}{\partial x^2}(D_{11}c_x)|_{x_{i+\frac{1}{2}},y_k} + \frac{7\Delta x^4}{5760}\frac{\partial^4}{\partial x^4}(D_{11}c_x)|_{x_{i+\frac{1}{2}},y_k}. \tag{2.17}$$

To evaluate $\hat{H}$, we would like to use the point values of c_j and u within the stencil I. We construct the polynomials interpolation of D_{11} and c_j, denoted as p_{11} and p_c, respectively. Then $p^x = p_{11}(p_c)_x$ is a high-order approximation of $D_{11}c_x$. We can read the point value of p^x as well as its derivatives to obtain the approximations of $D_{11}c_x$ and its derivatives. For example, consider a fifth-order scheme ($m = 5$), we need p_c to be a 6th order approximation of c_j, then a stencil with 6 points should be used.

Before we construct the numerical fluxes $\hat{u}$ and $\hat{v}$, we would like to demonstrate the following definition.

Definition 2.1. We say the flux $\widehat{uc}_j$ is consistent with $\hat{u}$ if $\widehat{uc}_j = \hat{u}$ by taking $c_j = 1$ in Ω.

Following the analysis in [13,4], we want the numerical fluxes $\widehat{uc}_j$ and $\hat{u}$ to be consistent, likewise for $\widehat{vc}_j$ and $\hat{v}$. Therefore, we take $\hat{u}_{i+\frac{1}{2},k}$ and $\hat{v}_{i,k+\frac{1}{2}}$ to be

$$\hat{u}_{i+\frac{1}{2},k} = \widehat{u1}_{i+\frac{1}{2},k}, \quad \hat{v}_{i,k+\frac{1}{2}} = \widehat{v1}_{i,k+\frac{1}{2}}.$$

Finally, we can discuss the numerical flux $\hat{p}_{i-\frac{1}{2},k}$. We can simply follow the algorithm in computing $\hat{u}_{i-\frac{1}{2},k}$ but the weights are symmetric about the axis $x = x_{i+\frac{1}{2}}$. The construction of $\hat{p}_{i,k-\frac{1}{2}}$ is similar.

Remark 2.1. The "consistent" numerical flux is used to guarantee $\sum_{j=1}^{N} c_j = 1$ at each time step (See Theorem 3.6). Thanks to the positivity-preserving techniques, we have $c_j \geq 0$ $j = 1, \cdots, N$, which further yields $0 \leq c_j \leq 1$.

Remark 2.2. Another way to discretize the diffusion term is to rewrite the diffusion part $\nabla(\mathbf{D}(\mathbf{u})\nabla c)$ into a system

$$\begin{aligned} \nabla \cdot (\mathbf{D}(\mathbf{u})\nabla c) &= \nabla \cdot s, \\ \mathbf{D}(\mathbf{u})\mathbf{q} &= \mathbf{s}, \\ \nabla c &= \mathbf{q}. \end{aligned}$$

We compare the sizes of the stencils used in the two algorithms and take $m = 5$ as an example. In (2.17), to obtain fifth-order accuracy, we use six points, and the stencil is exactly the same as that in the reconstruction procedure for the convection term. The method in this remark requires more than six points. Although the approximations of the diffusion term obtained by these two methods can be both fifth order accurate, numerical experiments demonstrate that (2.17) is better.

Before constructing the BP FD scheme, we would like to demonstrate the following key points that are quite different from most of the previous works.

1. Treat p_t in (2.8) as a source and apply the positivity-preserving techniques.
2. Apply flux limiters to the high-order scheme by combining the first-order and high-order fluxes. The stencils used for the two numerical fluxes are the same.
3. Choose a special "consistent" flux pair for (2.6) and (2.8). In Section 3.1, we will suitably choose the numerical fluxes and prove $c_j < 1$.
4. Apply Taylor's expansion to construct the numerical fluxes for the diffusion term.

3. BP technique for problems in one space dimension

In this section, we proceed to study the BP technique. We discuss the problems in one space dimension first, and the extension to the two dimensional problems will be given in the next section.

Though the FD scheme provides an approximation with high-order accuracy, the numerical solution may not be within the physical bounds when solving (2.8). To construct the BP technique, we consider the one-dimensional version of (2.9)-(2.12):

$$d(\mathbf{c}_i)\frac{d}{dt}p_i = -\frac{1}{\Delta x}(\hat{u}_{i+\frac{1}{2}} - \hat{u}_{i-\frac{1}{2}}) + q_i, \tag{3.1}$$

$$a(\mathbf{c}_i)u_i = -\frac{1}{\Delta x}(\hat{p}_{i+\frac{1}{2}} - \hat{p}_{i-\frac{1}{2}}), \tag{3.2}$$

$$\frac{d}{dt}(r_j)_i = -\frac{1}{\Delta x}(\widehat{uc}_{j\,i+\frac{1}{2}} - \widehat{uc}_{j\,i-\frac{1}{2}}) + \frac{1}{\Delta x}(\hat{H}_{i+\frac{1}{2}} - \hat{H}_{i-\frac{1}{2}}) + (\tilde{c}_j)_i q_i - (r_j)_i z_j \frac{d}{dt}p_i, \tag{3.3}$$

where $i = 1, \cdots, M$ with M being the total number of gridpoints. We use a parameterized positivity-preserving (PP) flux limiter to modify the high-order fluxes $\widehat{uc}_j$ and $\hat{H}$ in (3.3) towards a first-order fluxes, denoted as $\widehat{uc}_j{}^L$ and $\hat{h}$, by taking a linear combination of them. We will prove that the first-order scheme preserves the positivity of the numerical approximations. Following the analysis in [13,4], by using the consistent numerical flux pair, the numerical solution c_j can be bounded between 0 and 1.

3.1. First-order scheme

In this subsection, we apply Euler forward time discretization and construct first-order BP technique. We use o^n to represent the numerical approximation of o at time level n. The first-order fully-discretized scheme of (3.3) can be written as

$$(c_j)_i^{n+1} = F_i^c + F_i^d + F_i^s, \tag{3.4}$$

where

$$F_i^c = \frac{1}{3}(c_j)_i^n - \frac{\lambda}{\phi_i}(\widehat{uc}_{j\,i+\frac{1}{2}} - \widehat{uc}_{j\,i-\frac{1}{2}}), \tag{3.5}$$

$$F_i^d = \frac{1}{3}(c_j)_i^n + \frac{\lambda}{\phi_i}(\hat{h}_{i+\frac{1}{2}} - \hat{h}_{i-\frac{1}{2}}), \tag{3.6}$$

$$F_i^s = \frac{1}{3}(c_j)_i^n + \Delta t(\frac{(\tilde{c}_j)_i^n q_i^n}{\phi_i} - (c_j)_i^n z_j p_{t_i}^n), \tag{3.7}$$

with $\lambda = \frac{\Delta t}{\Delta x}$ being the ratio of the time and space mesh sizes. For simplicity, if we consider the numerical approximations at time level n, the superscript n will be dropped in the rest of this section. The following lemma is useful in the construction of first-order numerical fluxes for the convection term.

Lemma 3.1. *Let $\hat{f}^{\pm}$ be given in (2.14) and (2.15), then we have*

$$f^+(u_{i-r}c + \alpha c, \cdots, u_{i+s}c + \alpha c) = f^+(u_{i-r}, \cdots, u_{i+s})c + \alpha c,$$

and

$$f^-(u_{i+1-r}c-\alpha c,\cdots,u_{i+1+s}c-\alpha c)=f^-(u_{i+1-r},\cdots,u_{i+1+s})c-\alpha c.$$

Proof. We only proof for f^+ and the case for f^- is similar. From (2.14), we have,

$$\begin{aligned}f^+(u_{i-r}c+\alpha c,\cdots,u_{i+s}c+\alpha c)&=\sum_{\ell=-r}^{s}\omega_\ell^+(u_{i+\ell}c+\alpha c)\\&=\sum_{\ell=-r}^{s}\omega_\ell^+u_{i+\ell}c+\sum_{\ell=-r}^{s}\omega_\ell^+\alpha c\\&=\sum_{\ell=-r}^{s}\omega_\ell^+u_{i+\ell}c+\alpha c\\&=f^+(u_{i-r,k},\cdots,u_{i+s,k})c+\alpha c,\end{aligned}$$

where in the third step, we applied (2.16). □

To obtain a consistent flux pair, we need to construct a special low-order numerical flux $\widehat{uc}_j{}^L$ such that $\widehat{u1}^L=\hat{u}$. In (3.5), we also apply the flux splitting [22] to the numerical fluxes. Following the construction in (2.13) and Lemma 3.1, we choose

$$\widehat{uc}_j{}^L_{i+\frac{1}{2}}=\frac{1}{2}\left((\hat{f}^+_{i+\frac{1}{2}}+\alpha)c_{j_i}+(\hat{f}^-_{i+\frac{1}{2}}-\alpha)c_{j_{i+1}}\right),\tag{3.8}$$

where

$$\hat{f}^+_{i+\frac{1}{2}}=f^+(u_{i-r},\cdots,u_{i+s}),\quad \hat{f}^-_{i+\frac{1}{2}}=f^-(u_{i-r+1},\cdots,u_{i+s+1}).$$

In (3.6), the numerical flux of the diffusion term is given as

$$\hat{h}_{i+\frac{1}{2}}=\frac{1}{2}(D(u)_i+D(u)_{i+1})\frac{(c_j)_{i+1}-(c_j)_i}{\Delta x}.\tag{3.9}$$

We can prove that if Δt is sufficiently small, then F_i^c, F_i^d and F_i^s are all positive, and the results are given in the following three lemmas. For simplicity of presentation, if the denominator in a fraction is zero, then the value of the fraction is defined as ∞. For simplicity, if not otherwise stated, we will drop the subscript j for the jth component and use c, z, $\tilde{c}$ for c_j, z_j, $\tilde{c}_j$, respectively. We prove $F_i^c>0$ first.

Lemma 3.2. *Suppose $c_i>0$ for all i, then $F_i^c>0$ under the conditions*

$$\lambda\le\min_i\frac{2\phi_i}{3(\hat{f}^+_{i+\frac{1}{2}}-\hat{f}^-_{i-\frac{1}{2}}+2\alpha)},\tag{3.10}$$

and

$$\alpha\ge\max_i\{-\hat{f}^+_{i+\frac{1}{2}},\hat{f}^-_{i+\frac{1}{2}},0\}.\tag{3.11}$$

Proof. It is easy to check that

$$\begin{aligned}F_i^c&=\frac{1}{3}c_i-\frac{\lambda}{\phi_i}(\widehat{uc}^L_{i+\frac{1}{2}}-\widehat{uc}^L_{i-\frac{1}{2}})\\&=\frac{1}{3}c_i-\frac{\lambda}{2\phi_i}\left((\hat{f}^+_{i+\frac{1}{2}}+\alpha)c_i+(\hat{f}^-_{i+\frac{1}{2}}-\alpha)c_{i+1}-(\hat{f}^+_{i-\frac{1}{2}}+\alpha)c_{i-1}-(\hat{f}^-_{i-\frac{1}{2}}-\alpha)c_i\right)\\&=\left(\frac{1}{3}-\frac{\lambda}{2\phi_i}(\hat{f}^+_{i+\frac{1}{2}}-\hat{f}^-_{i-\frac{1}{2}}+2\alpha)\right)c_i+\frac{\lambda}{2\phi_i}(\alpha-\hat{f}^-_{i+\frac{1}{2}})c_{i+1}+\frac{\lambda}{2\phi_i}(\alpha+\hat{f}^+_{i-\frac{1}{2}})c_{i-1}.\end{aligned}$$

Clearly, under the conditions (3.10) and (3.11), all the coefficients are positive. Hence $F_i^c>0$. □

Now we proceed to prove $F_i^d>0$ and the result is given in the following lemma.

Lemma 3.3. *Suppose $c_i > 0$ for all i, then $F_i^d > 0$ under the conditions*

$$\Lambda = \frac{\Delta t}{\Delta x^2} \le \frac{\phi_i}{6D_{max}}, \quad D_{max} = \max_i D(u)_i. \tag{3.12}$$

Proof. Following the same analysis for the convection term, we write

$$\begin{aligned} F_i^d &= \frac{1}{3}c_i + \frac{\lambda}{\phi_i}(\hat{h}_{i+\frac{1}{2}} - \hat{h}_{i-\frac{1}{2}}) \\ &= \frac{1}{3}c_i + \frac{\Lambda}{\phi_i}\left(\frac{D(u)_i + D(u)_{i+1}}{2}(c_{i+1} - c_i) - \frac{D(u)_{i-1} + D(u)_i}{2}(c_i - c_{i-1})\right) \\ &= \frac{1}{3}c_i - \frac{\Lambda}{2\phi_i}(D(u)_{i+1} + 2D(u)_i + D(u)_{i-1})c_i \\ &+ \frac{\Lambda}{2\phi_i}(D(u)_i + D(u)_{i+1})c_{i+1} + \frac{\Lambda}{2\phi_i}(D(u)_{i-1} + D(u)_i)c_{i-1}. \end{aligned}$$

Since $D(u)_i \ge 0, i = 1, \cdots, M$, we have $F_i^d > 0$ under the condition (3.12). □

Next, we proceed to analyze F_i^s.

Lemma 3.4. *Suppose $c_i > 0$ for all i, then $F_i^s > 0$ under the condition*

$$\Delta t \le \min\{\frac{1}{6zp_{max}}, \frac{\phi_i}{6\max(-q_i, 0)}\}, \tag{3.13}$$

where

$$p_{max} = \max_i(p_{ti}, 0). \tag{3.14}$$

Proof. We can write the source term as

$$F_i^s = \frac{1}{3}c_i + \Delta t\left(\frac{\tilde{c}_i q_i}{\phi_i} - c_i z p_{ti}\right) = S^1 + S^2,$$

where

$$S^1 = (\frac{1}{6} - \Delta t z p_{ti})c_i, \quad S^2 = \frac{1}{6}c_i + \Delta t\frac{\tilde{c}_i q_i}{\phi_i}.$$

Clearly, we have $S^1 \ge 0$ if (3.13) is satisfied. Moreover, if $q_i \ge 0$, then $S^2 > 0$. Hence we only need to consider the case with $q_i < 0$. When $q_i < 0$, we have $\tilde{c}_i = c_i$. Then under the condition (3.13)

$$S^2 = (\frac{1}{6} + \Delta t\frac{q_i}{\phi_i})c_i > 0.$$

We finish the proof. □

Base on the above three lemmas, we can state the following theorem.

Theorem 3.5. *Consider the FD scheme (3.1)-(3.3) with Euler forward time discretization. Suppose $c_i > 0$ for all i, then $c_i^{n+1} > 0$ under the conditions (3.10), (3.11), (3.12) and (3.13).*

Proof. By (3.4), we only need F_i^c, F_i^d, and F_i^s to be positive, which follows directly from Lemmas 3.2-3.4, respectively. □

Now we have the fact that $(c_j)^{n+1} > 0$ for $j = 1, \cdots, N-1$ from Theorem 3.5. However, we still need to show $(c_j)^{n+1} < 1$, and the result is given below.

Theorem 3.6. *Suppose the conditions in Theorem 3.5 are all satisfied. We assume $c_j \ge 0$, $j = 1, \cdots, N$, with $c_N = 1 - \sum_{j=1}^{n-1} c_j$ and the flux of u and the flux of uc_j are consistent. Then $0 \le c_j^{n+1} \le 1$, under the condition*

$$\Delta t \le \frac{1}{6z_{max}p_{max}}, \tag{3.15}$$

where p_{max} is given in (3.14) and $z_{max} = \max_{1\le j\le N} z_j$.

Proof. Since the fluxes $\hat{u}$ and $\widehat{uc}_j^{\,L}$ are consistent, then

$$\widehat{uc}_N + \sum_{j=1}^{N-1} \widehat{uc}_j = \widehat{u1} = \hat{u},$$

where the flux $\widehat{uc}$ can be either the high-order or the low-order one. Add up the first-order scheme of (3.3) for $j = 1, \cdots, N-1$ and subtract the result from the first-order scheme of (3.1) to obtain a scheme satisfied by c_N

$$\frac{d}{dt}(r_N)_i + \frac{1}{\Delta x}(\widehat{uc_N}^L_{i+\frac{1}{2}} - \widehat{uc_N}^L_{i-\frac{1}{2}}) - \frac{1}{\Delta x}(\hat{h}_{i+\frac{1}{2}} - \hat{h}_{i-\frac{1}{2}}) = (\tilde{c}_N)_i q_i - (r_N)_i z_N \frac{d}{dt} p_i, \tag{3.16}$$

where $\tilde{c}_N = 1 - \sum_{j=1}^{N-1} \tilde{c}_j$ and $r_N = c_N \phi$. We can find that (3.16) is exactly (3.3) with c_j, $\tilde{c}_j$, and z_j replaced by c_N, $\tilde{c}_N$, and z_N, respectively. Therefore, following the same analysis in Theorem 3.5, with Euler forward time integration we can obtain $c_N^{n+1} > 0$ under the conditions in this theorem, which further implies $c_j^{n+1} < 1$. □

3.2. Bound-preserving technique for high-order schemes

In this subsection, we will apply the parameterized PP flux limiter to construct high-order BP technique by taking a linear combination of high-order and first-order numerical fluxes. We also consider Euler forward time discretization and drop the subscript j if not otherwise stated.

We write (3.3) as

$$c_i^{n+1} = c_i - \lambda(\hat{F}_{i+\frac{1}{2}} - \hat{F}_{i-\frac{1}{2}}) + \Delta t S_i, \tag{3.17}$$

where

$$\hat{F}_{i+\frac{1}{2}} = \frac{1}{\phi_i}(\widehat{uc}_{i+\frac{1}{2}} - \hat{H}_{i+\frac{1}{2}}), \quad S_i = \frac{\tilde{c}_i q_i}{\phi_i} - c_i z p_{t_i}.$$

We apply the flux limiter to obtain positive numerical approximation c_i^{n+1}. To do so, we replace the numerical flux $\hat{F}_{i+\frac{1}{2}}$ in (3.17) by the modified one

$$\tilde{F}_{i+\frac{1}{2}} = \theta_{i+\frac{1}{2}}(\hat{F}_{i+\frac{1}{2}} - \hat{f}_{i+\frac{1}{2}}) + \hat{f}_{i+\frac{1}{2}}, \tag{3.18}$$

where $\hat{f}_{i+\frac{1}{2}} = \frac{1}{\phi_i}(\widehat{uc}^L_{i+\frac{1}{2}} - \hat{h}_{i+\frac{1}{2}})$ is the first-order flux discussed in Section 3.1. We choose the parameters $\theta_{i+\frac{1}{2}} \in [0, 1]$ such that

$$c_i - \lambda(\tilde{F}_{i+\frac{1}{2}} - \tilde{F}_{i-\frac{1}{2}}) + \Delta t S_i \geq 0,$$

which can be rewritten as

$$\lambda\theta_{i-\frac{1}{2}}(\hat{F}_{i-\frac{1}{2}} - \hat{f}_{i-\frac{1}{2}}) - \lambda\theta_{i+\frac{1}{2}}(\hat{F}_{i+\frac{1}{2}} - \hat{f}_{i+\frac{1}{2}}) - \Gamma_i^m \geq 0, \tag{3.19}$$

where

$$\Gamma_i^m = -c_i + \lambda(\hat{f}_{i+\frac{1}{2}} - \hat{f}_{i-\frac{1}{2}}) - \Delta t S_i. \tag{3.20}$$

Theorem 3.5 guarantees that $\Gamma_i^m \leq 0$. For simplicity, we denote $F_{i\pm\frac{1}{2}} = \hat{F}_{i\pm\frac{1}{2}} - \hat{f}_{i\pm\frac{1}{2}}$. We consider the ith node and look for a locally defined pair of numbers $(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i})$ such that (3.19) is satisfied for

$$\theta_{i-\frac{1}{2}} \in [0, \Lambda^m_{-\frac{1}{2},I_i}], \quad \theta_{i+\frac{1}{2}} \in [0, \Lambda^m_{+\frac{1}{2},I_i}]. \tag{3.21}$$

Following [28],

1. If $F_{i-\frac{1}{2}} \geq 0$ and $F_{i+\frac{1}{2}} \leq 0$, let
$$(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i}) = (1, 1).$$
2. If $F_{i-\frac{1}{2}} \geq 0$ and $F_{i+\frac{1}{2}} > 0$, let
$$(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i}) = (1, \min(1, \frac{\Gamma_i^m}{-\lambda F_{i+\frac{1}{2}} - \epsilon})).$$

3. If $F_{i-\frac{1}{2}} < 0$ and $F_{i+\frac{1}{2}} \le 0$, let

$$(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i}) = (\min(1, \frac{\Gamma^m_i}{\lambda F_{i-\frac{1}{2}} - \epsilon}), 1).$$

4. If $F_{i-\frac{1}{2}} < 0$ and $F_{i+\frac{1}{2}} > 0$,
   (a) When (3.19) holds with $(\theta_{i-\frac{1}{2}}, \theta_{i+\frac{1}{2}}) = (1, 1)$, let

$$(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i}) = (1, 1);$$

   (b) Otherwise, let

$$(\Lambda^m_{-\frac{1}{2},I_i}, \Lambda^m_{+\frac{1}{2},I_i}) = (\frac{\Gamma^m_i}{\lambda F_{i-\frac{1}{2}} - \lambda F_{i+\frac{1}{2}} - \epsilon}, \frac{\Gamma^m_i}{\lambda F_{i-\frac{1}{2}} - \lambda F_{i+\frac{1}{2}} - \epsilon}).$$

In the algorithm given above, ϵ is chosen to be a very small positive number to avoid the denominator being 0. For example, $\epsilon = 10^{-13}$. Then the locally defined limiting parameter is given as

$$\theta_{i+\frac{1}{2}} = \min(\Lambda^m_{+\frac{1}{2},I_i}, \Lambda^m_{-\frac{1}{2},I_{i+1}}), \quad i = 0, \cdots, M. \tag{3.22}$$

Finally, for each $j = 1, \cdots, N$, we can find the parameter θ^j following the algorithm given above. Since $c_j's$ satisfy the same equation, the parameter θ applied in the scheme (3.18) should be

$$\tilde{\theta}_{i+\frac{1}{2}} = \min_{1 \le j \le N} \theta^j_{i+\frac{1}{2}}.$$

Remark 3.1. In the algorithm above, we choose $\tilde{\theta}$ to be the smallest one among $j = 1, \cdots, N$, then we can obtain positive c_j for all $j = 1, \cdots, N$. Then the condition $\sum_j c_j = 1$ yields $0 \le c_j \le 1$.

Remark 3.2. It is not a good idea to use a stencil with 2 points to construct the low-order numerical flux $\widehat{uc_j}^L$ for the BP technique, since such a numerical flux cannot be consistent with $\hat{u}$. Later, we will apply the flux limiter to the numerical flux and take $\widehat{uc_j} = \theta \widehat{uc_j}^H + (1-\theta)\widehat{uc_j}^L$. If $\widehat{uc_j}^H$ and $\widehat{uc_j}^L$ are not consistent with $\hat{u}$, we must choose a suitable θ (which depends on Δt) such that $\widehat{uc_j}$ given above is consistent with $\hat{u}$, which further yields the approximation of p_t from (2.9), leading to a new constraint for the time step Δt (see Lemma 3.4). Numerical experiments (See Example 5.2) demonstrate that the new time step size requirement can be smaller than that in the flux limiter. We call this to be the time step paradox. However, if we choose $\widehat{uc_j}^H$ and $\widehat{uc_j}^L$ to be consistent with $\hat{u}$, then after the flux limiter, $\widehat{uc_j}$ is also consistent with $\hat{u}$, no matter what the time step is. Hence, we can compute p_t from (2.9) and construct the time step restriction first, and then apply the flux limiter.

3.3. High-order time discretization

All the previous analyses are based on Euler forward time discretization. For high-order ones, we use the third-order strong-stability-preserving (SSP) Runge-Kutta time discretization to solve the ODE system $\mathbf{u_t} = \mathbf{L}(\mathbf{u})$:

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

Another choice is the third-order SSP multi-step method:

$$\mathbf{u}^{n+1} = \frac{16}{27}(\mathbf{u}^n + 3\Delta t \mathbf{L}(\mathbf{u}^n, t^n)) + \frac{11}{27}(\mathbf{u}^{n-3} + \frac{12}{11}\Delta t \mathbf{L}(\mathbf{u}^{n-3}, t^{n-3})).$$

More details can be found in [11,12,21]. Since the SSP time discretization is a convex combination of Euler forward time discretization, we can apply the flux limiter designed in Section 3.2 to each stage/step. The numerical solution obtained from the full scheme is also physically relevant.

The above idea is to apply the flux limiter to each stage of the Runge-Kutta method. In [26], the authors provided another idea and applied the flux limiter in the final stage of the Runge-Kutta method. However, this algorithm will also lead to the time step paradox. In fact, to perform the Runge-Kutta method, we need to choose the time step first. After we finish all the stages, we can calculate the value of p_t, leading to a new time step size constraint in the bound-preserving technique.

4. The bound-preserving technique in two space dimensions

In this section, we briefly discuss the BP technique for miscible displacements in two space dimensions. The algorithm would be basically the same as that discussed in Section 3. Therefore, we only demonstrate the results and skip the proofs. For simplicity, we only discuss Euler forward time discretization, and drop the superscript n and the subscript j if possible.

With Euler forward time discretization, (2.12) can be written as

$$\begin{aligned} r_{i,k}^{n+1} = r_{i,k} &- \lambda_x(\widehat{uc}_{i+\frac{1}{2},k} - \widehat{uc}_{i-\frac{1}{2},k}) - \lambda_y(\widehat{vc}_{i,k+\frac{1}{2}} - \widehat{vc}_{i,k-\frac{1}{2}}) \\ &+ \lambda_x(\hat{H}_{i+\frac{1}{2},k} - \hat{H}_{i-\frac{1}{2},k}) + \lambda_y(\hat{G}_{i,k+\frac{1}{2}} - \hat{G}_{i,k-\frac{1}{2}}) + \Delta t(\tilde{c}_{i,k}q_{i,k} - r_{i,k}zp_{ti,k}), \end{aligned} \tag{4.1}$$

where $\lambda_x = \frac{\Delta t}{\Delta x}$, $\lambda_y = \frac{\Delta t}{\Delta y}$. The high-order numerical fluxes have been constructed in Section 2. Now, we proceed to construct the first-order numerical fluxes. For the convection part, we choose

$$\widehat{uc}^L_{i+\frac{1}{2},k} = \frac{1}{2}\left((\hat{f}^+_{i+\frac{1}{2},k} + \alpha_x)c_{i,k} + (\hat{f}^-_{i+\frac{1}{2},k} - \alpha_x)c_{i+1,k}\right), \tag{4.2}$$

$$\widehat{vc}^L_{i,k+\frac{1}{2}} = \frac{1}{2}\left((\hat{f}^+_{i,k+\frac{1}{2}} + \alpha_y)c_{i,k} + (\hat{f}^-_{i,k+\frac{1}{2}} - \alpha_y)c_{i,k+1}\right), \tag{4.3}$$

where

$$\hat{f}^+_{i+\frac{1}{2},k} = f^+(u_{i-r,k}, \cdots, u_{i+s,k}), \quad \hat{f}^-_{i+\frac{1}{2},k} = f^-(u_{i-r+1,k}, \cdots, u_{i+s+1,k}).$$

$\hat{f}^+_{i,k+\frac{1}{2}}$ and $\hat{f}^-_{i,k+\frac{1}{2}}$ can be defined analogously. The parameters α_x and α_y are to be chosen by the BP technique.

The fluxes for the diffusion part are

$$\hat{h}_{i+\frac{1}{2},k} = (\bar{D}^{11}_{i+\frac{1}{2},k})\frac{c_{i+1,k} - c_{i,k}}{\Delta x}, \quad \hat{g}_{i,k+\frac{1}{2}} = (\bar{D}^{22}_{i,k+\frac{1}{2}})\frac{c_{i,k+1} - c_{i,k}}{\Delta y},$$

where $\bar{D}^{11}_{i+\frac{1}{2},k} = \frac{1}{2}(D_{11}(\mathbf{u})_{i,k} + D_{11}(\mathbf{u})_{i+1,k})$. Likewise for $\bar{D}^{22}_{i,k+\frac{1}{2}}$. Thus we can rewrite (4.1) as

$$c^{n+1}_{i,k} = F^c_{i,k} + F^d_{i,k} + F^s_{i,k}, \tag{4.4}$$

where

$$F^c_{i,k} = \frac{1}{3}c_{i,k} - \frac{\lambda_x}{\phi_{i,k}}(\widehat{uc}_{i+\frac{1}{2},k} - \widehat{uc}_{i-\frac{1}{2},k}) - \frac{\lambda_y}{\phi_{i,k}}(\widehat{vc}_{i,k+\frac{1}{2}} - \widehat{vc}_{i,k-\frac{1}{2}}), \tag{4.5}$$

$$F^d_{i,k} = \frac{1}{3}c_{i,k} + \frac{\lambda_x}{\phi_{i,k}}(\hat{h}_{i+\frac{1}{2},k} - \hat{h}_{i-\frac{1}{2},k}) + \frac{\lambda_y}{\phi_{i,k}}(\hat{g}_{i,k+\frac{1}{2}} - \hat{g}_{i,k-\frac{1}{2}}), \tag{4.6}$$

$$F^s_{i,k} = \frac{1}{3}c_{i,k} + \Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{ti,k}). \tag{4.7}$$

Following the same analyses given in Section 3.1 along x and y directions, we can obtain the following results regarding $F^c_{i,k}$, $F^c_{i,k}$ and $F^s_{i,k,}$.

Lemma 4.1. *Suppose $c_{i,k} > 0$ for all i, k, then $F^c_{i,k} > 0$ under the conditions*

$$\lambda_x = \frac{\Delta t}{\Delta x} \le \frac{\phi_{i,k}}{3(\hat{f}^+_{i+\frac{1}{2},k} - \hat{f}^-_{i-\frac{1}{2},k} + 2\alpha_x)}, \quad \lambda_y = \frac{\Delta t}{\Delta y} \le \frac{\phi_{i,k}}{3(\hat{f}^+_{i,k+\frac{1}{2}} - \hat{f}^-_{i,k-\frac{1}{2}} + 2\alpha_y)}, \tag{4.8}$$

and

$$\alpha_x \ge \max_{i,k}\{-\hat{f}^+_{i+\frac{1}{2},k}, \hat{f}^-_{i+\frac{1}{2},k}, 0\}, \quad \alpha_y \ge \max_{i,k}\{-\hat{f}^+_{i,k+\frac{1}{2}}, \hat{f}^-_{i,k+\frac{1}{2}}, 0\}. \tag{4.9}$$

Lemma 4.2. *Suppose $c_{i,k} > 0$ for all i, k, then $F^d_{i,k} > 0$ under the conditions*

$$\frac{\Delta t}{\Delta x^2} \le \frac{\phi_{i,k}}{12D^M_{11}}, \quad \frac{\Delta t}{\Delta y^2} \le \frac{\phi_{i,k}}{12D^M_{22}}, \tag{4.10}$$

where

$$D^M_{\ell\ell} = \max_{i,k}(D_{\ell\ell}(\mathbf{u})_{i,k}), \quad \ell = 1, 2.$$

Lemma 4.3. *Suppose $c_{i,k} > 0$ for all i, k, then $F^s_{i,k} > 0$ under the conditions*

$$\Delta t \le \frac{1}{6zp_{max}}, \quad \Delta t \le \frac{\phi_{i,k}}{6\max(-q_{i,k}, 0)}, \tag{4.11}$$

where

$$p_{max} = \max_{i,k}(p_{t\,i,k}, 0). \tag{4.12}$$

Based on the above three lemmas, we can state the following one.

Theorem 4.4. *Suppose $c_{i,k} > 0$ for all i, k, then $c^{n+1}_{i,k} > 0$ under the conditions (4.9), (4.8), (4.10), and (4.11).*

Following the proof of Theorem 3.6 with minor changes, we can obtain the following theorem.

Theorem 4.5. *Suppose the conditions in Theorem 4.4 are satisfied. We assume $c_j \ge 0$ for all $j = 1, \cdots, N$ and the flux of $(\hat{u}, \hat{v})$ and the flux of $(\widehat{uc}_j, \widehat{vc}_j)$ are consistent. Then $0 \le c^{n+1}_j \le 1$, under an extra condition*

$$\Delta t \le \frac{1}{6z_{max}p_{max}}. \tag{4.13}$$

Now, we have constructed first-order BP technique. To obtain high-order BP property, we need to apply the flux limiter and combine the first-order and high-order numerical fluxes. Similar to the BP technique in one space dimension, we are looking for the modified numerical fluxes

$$\tilde{F}^x_{i+\frac{1}{2},k} = \theta_{i+\frac{1}{2},k}(\hat{F}^x_{i+\frac{1}{2},k} - \hat{f}^x_{i+\frac{1}{2},k}) + \hat{f}^x_{i+\frac{1}{2},k}, \quad \tilde{F}^y_{i,k+\frac{1}{2}} = \theta_{i,k+\frac{1}{2}}(\hat{F}^y_{i,k+\frac{1}{2}} - \hat{f}^y_{i,k+\frac{1}{2}}) + \hat{f}^y_{i,k+\frac{1}{2}}, \tag{4.14}$$

where

$$\begin{aligned}
&\hat{F}^x_{i+\frac{1}{2},k} = \frac{1}{\phi_{i,k}}(\widehat{uc}_{i+\frac{1}{2},k} - \hat{H}_{i+\frac{1}{2},k}), \quad \hat{f}^x_{i+\frac{1}{2},k} = \frac{1}{\phi_{i,k}}(\widehat{uc}^L_{i+\frac{1}{2},k} - \hat{h}_{i+\frac{1}{2},k}),\\
&\hat{F}^y_{i,k+\frac{1}{2}} = \frac{1}{\phi_{i,k}}(\widehat{vc}_{i,k+\frac{1}{2}} - \hat{G}_{i,k+\frac{1}{2}}), \quad \hat{f}^y_{i,k+\frac{1}{2}} = \frac{1}{\phi_{i,k}}(\widehat{vc}^L_{i,k+\frac{1}{2}} - \hat{g}_{i,k+\frac{1}{2}}).
\end{aligned} \tag{4.15}$$

We want to choose $\theta_{i+\frac{1}{2},k}$ and $\theta_{i,k+\frac{1}{2}}$ such that

$$c_{i,k} - \lambda_x(\tilde{F}^x_{i+\frac{1}{2},k} - \tilde{F}^x_{i-\frac{1}{2},k}) - \lambda_y(\tilde{F}^y_{i,k+\frac{1}{2}} - \tilde{F}^y_{i,k-\frac{1}{2}}) + \Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{t\,i,k}) \ge 0. \tag{4.16}$$

Following the analysis in Section 3, we can divide (4.16) into two parts

$$\frac{1}{2}c_{i,k} - \lambda_x(\tilde{F}^x_{i+\frac{1}{2},k} - \tilde{F}^x_{i-\frac{1}{2},k}) + \frac{1}{2}\Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{t\,i,k}) \ge 0, \tag{4.17}$$

$$\frac{1}{2}c_{i,k} - \lambda_y(\tilde{F}^y_{i,k+\frac{1}{2}} - \tilde{F}^y_{i,k-\frac{1}{2}}) + \frac{1}{2}\Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{t\,i,k}) \ge 0. \tag{4.18}$$

which can be rewritten as

$$\lambda_x\theta_{i-\frac{1}{2},k}(\hat{F}^x_{i-\frac{1}{2},k} - \hat{f}^x_{i-\frac{1}{2},k}) - \lambda_x\theta_{i+\frac{1}{2},k}(\hat{F}^x_{i+\frac{1}{2},k} - \hat{f}^x_{i+\frac{1}{2},k}) - \Gamma^{m_x}_{i,k} \ge 0, \tag{4.19}$$

$$\lambda_y\theta_{i,k-\frac{1}{2}}(\hat{F}^y_{i,k-\frac{1}{2}} - \hat{f}^y_{i,k-\frac{1}{2}}) - \lambda_y\theta_{i,k+\frac{1}{2}}(\hat{F}^y_{i,k+\frac{1}{2}} - \hat{f}^y_{i,k+\frac{1}{2}}) - \Gamma^{m_y}_{i,k} \ge 0, \tag{4.20}$$

where

$$\Gamma^{m_x}_{i,k} = -\frac{1}{2}c_{i,k} + \lambda_x(\hat{f}^x_{i+\frac{1}{2},k} - \hat{f}^x_{i-\frac{1}{2},k}) - \frac{1}{2}\Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{t\,i,k}) \le 0, \tag{4.21}$$

$$\Gamma^{m_y}_{i,k} = -\frac{1}{2}c_{i,k} + \lambda_y(\hat{f}^y_{i,k+\frac{1}{2}} - \hat{f}^y_{i,k-\frac{1}{2}}) - \frac{1}{2}\Delta t(\frac{\tilde{c}_{i,k}q_{i,k}}{\phi_{i,k}} - c_{i,k}zp_{t\,i,k}) \le 0. \tag{4.22}$$

For brevity, we define

$$F^x_{i\pm\frac{1}{2},k} = \hat{F}^x_{i\pm\frac{1}{2},k} - \hat{f}^x_{i\pm\frac{1}{2},k}, \quad F^y_{i,k\pm\frac{1}{2}} = \hat{F}^y_{i,k\pm\frac{1}{2}} - \hat{f}^y_{i,k\pm\frac{1}{2}}. \tag{4.23}$$

Table 1
Example 5.1: Accuracy test for the fifth-order FD scheme with and without BP technique.

M	Without limiter			With limiter		
	L^∞ error	Order	Minimum value	L^∞ error	Order	Minimum value
40	8.07e−04	–	−7.04e−04	9.00e−04	–	4.92e−05
80	2.66e−05	4.92	−2.29e−05	4.23e−05	4.41	1.66e−07
160	8.43e−07	4.98	−7.43e−07	1.51e−06	4.81	1.81e−07
320	2.65e−08	4.99	−1.82e−08	4.93e−08	4.94	4.68e−10
640	8.28e−10	5.00	−5.75e−10	1.67e−09	4.88	6.08e−10

Following the same procedure in Section 3, we consider the node (i,k) and can obtain the locally defined pairs of numbers $(\Lambda^m_{i-\frac{1}{2},k}, \Lambda^m_{i+\frac{1}{2},k})$ and $(\Lambda^m_{i,k-\frac{1}{2}}, \Lambda^m_{i,k+\frac{1}{2}})$ along the x and y directions, respectively. Then the locally defined limiting parameters are given as

$$\theta_{i+\frac{1}{2},k} = \min(\Lambda^m_{i-\frac{1}{2},k}, \Lambda^m_{i+\frac{1}{2},k}), \quad i = 1, \cdots, N_x, \tag{4.24}$$

$$\theta_{i,k+\frac{1}{2}} = \min(\Lambda^m_{i,k-\frac{1}{2}}, \Lambda^m_{i,k+\frac{1}{2}}), \quad k = 1, \cdots, N_y. \tag{4.25}$$

For each $j = 1, 2, \cdots, N$ we can find θ^j following the algorithm given above. Similar to the BP technique in Section 3, we choose

$$\tilde{\theta}_{i+\frac{1}{2},k} = \min_{1\le j\le N} \theta^j_{i+\frac{1}{2},k}, \quad \tilde{\theta}_{i,k+\frac{1}{2}} = \min_{1\le j\le N} \theta^j_{i,k+\frac{1}{2}}$$

to be the parameters used in the flux limiter (4.14).

5. Numerical experiments

In this section, we provide numerical experiments to test the accuracy and stability of the high-order BP FD schemes. In all the examples, if not otherwise stated, we consider fluid mixtures with two components and use linear weights in the reconstruction procedure. Moreover, we use third-order SSP Runge-Kutta discretization in time and fifth-order finite difference scheme ($r = s = 2$) in space.

5.1. One dimensional case

In this subsection, we solve (1.1)-(1.2) in one space dimension on the computational domain $[0, 2\pi]$. In the first example, we test the accuracy of the BP FD scheme.

Example 5.1. We choose the initial condition as

$$c(x,0) = \sin^4(x), \quad p(x,0) = -x,$$

and the source variables are taken as

$$q = \gamma, \quad \tilde{c} = 0.$$

Other parameters are chosen as

$$\phi(x) = \mu(c) = k(x) = z_1 = z_2 = 1, \quad D(u) = 0.$$

We choose suitable Dirichlet conditions for the pressure equation and periodic boundary condition for the concentration equation such that the exact solutions are

$$c(x,t) = e^{-\gamma t}(\sin^4(x-t)), \quad p(x,t) = \gamma t - x,$$

where $\gamma = 10^{-5}$. In the numerical simulation, we take $\Delta t = 0.6\Delta x^2$ and compute up to $T = 1$. We calculate the L^∞-norm of the error between the numerical and exact solutions of c, and the results are given in Table 1. From the table, we can observe fifth-order accuracy of the FD scheme with and without the BP limiter. In Addition, negative values emerged in the case of without BP limiter, and the parameterized PP flux limiters remedied the negative values in a conservative way. Therefore, the BP technique does not degenerate the accuracy while it works.

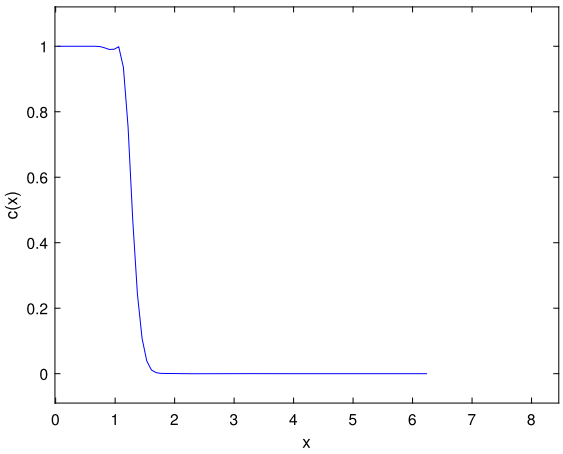


Fig. 1. Example 5.2: Numerical approximation of c at $t = 1$.

Example 5.2. We choose the initial condition as

$$c(x,0) = \begin{cases} 1, x < 1, \\ 0, x > 1, \end{cases} \qquad p(x,0) = \begin{cases} 5, x < 1, \\ 0, x > 1. \end{cases}$$

Other parameters are taken as

$$q(x,t) = 0, z_1 = 0.1, k(x) = \mu(x) = z_2 = 1, \phi(x) = 1, D(u) = 0.$$

We compute up to $T = 1$ with $M = 80$ and $\Delta t = 0.1\Delta x^2$ to reduce the time error. We solve the problem with the BP technique and the numerical approximation of c is given in Fig. 1. We can observe that the numerical approximation is between 0 and 1. Moreover, to test the effectiveness of the BP technique, we also solve the problem without the BP limiter and the numerical approximation blows up at $t \approx 0.09$. We have been demonstrated that the reason for the blow-up phenomenon is the ill-posedness of the system in [13]. This example demonstrates the necessity of the BP technique in solving compressible miscible displacements in porous media.

To demonstrate the time step paradox we use the two algorithms given in Remark 3.2 and obtain the maximum time step for them. When we choose the "consistent" algorithm, the maximum time step is $\Delta t = 0.17\Delta x^2$. Then we apply this time step to the "inconsistent" one, the numerical approximations of c will be out of the bound and the BP technique fails to work under this time step. Actually, numerical tests show that the maximum time step of the "inconsistent" algorithm is $\Delta t = 0.12\Delta x^2$. Thus, the special treatment of the numerical flux in the convection term is useful.

Example 5.3. We choose the initial condition as

$$c(x,0) = 1, \quad p(x,0) = 1.$$

and the source variables are taken as

$$q(x,t) = \begin{cases} 50, 3 < x < 3.2, \\ 0, \quad \text{otherwise}, \end{cases} \qquad \tilde{c}(x,t) = 0.$$

Other parameters are chosen as

$$\phi(x) = \mu(c) = k(x) = 1, \quad z_1 = 0.1, \quad z_2 = 1, \quad D(u) = |u|.$$

We compute up to $T = 0.1$ with $M = 40$ and $\Delta t = 0.01\Delta x^2$. We solve the problem with two algorithms for the diffusion term and the results are given in Fig. 2. We use blue curve to represent the numerical result obtained by the first algorithm (Applying Taylor's expansion to the diffusion term) and use red curve to represent the numerical result obtained by the second algorithm (The algorithm in Remark 2.2). From Fig. 2, we can observe that the numerical approximation obtained by the second algorithm is oscillatory, leading to larger pollution region, while the first one does not yield significant oscillations.

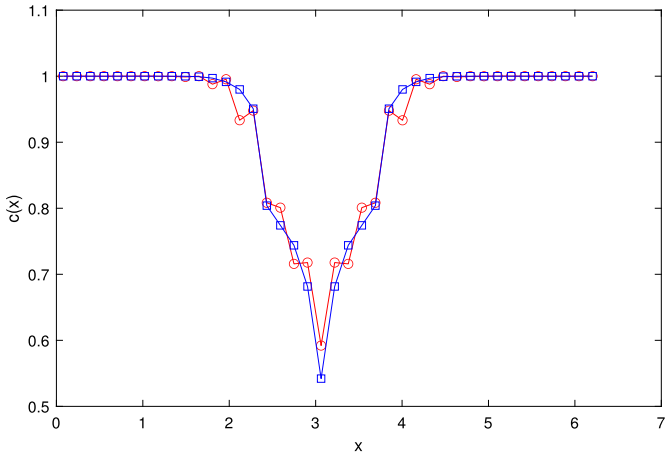
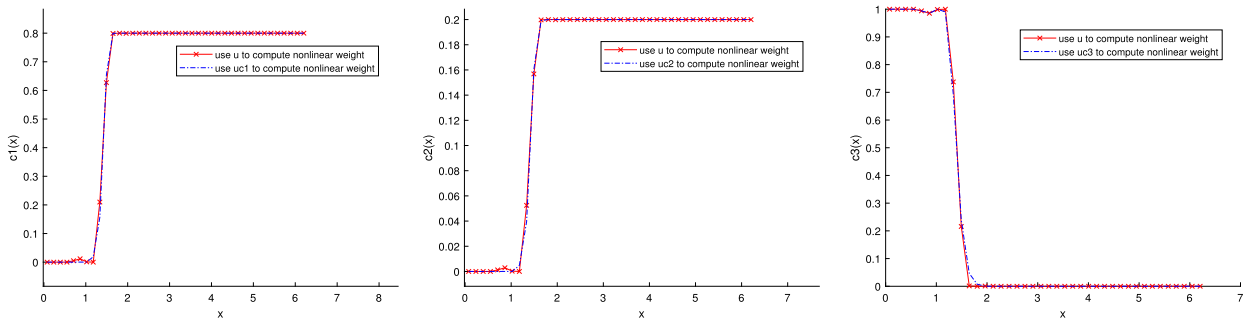


Fig. 2. Example 5.3: Numerical approximation of c at $t = 0.1$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)



(a) c_1 at T=0.1 s (b) c_2 at T=0.1 s (c) c_3 at T=0.1 s

Fig. 3. Example 5.4: Concentrations of c_1, c_2 and c_3.

Example 5.4. In this example, we consider a fluid mixture with three components. We choose the initial condition as

$$c_1(x,0) = \begin{cases} 0, & x < 1, \\ 0.8, & x > 1, \end{cases} \qquad c_2(x,0) = \begin{cases} 0, & x < 1, \\ 0.2, & x > 1, \end{cases}$$

$$c_3(x,0) = \begin{cases} 1, & x < 1, \\ 0, & x > 1, \end{cases} \qquad p(x,0) = \begin{cases} 5, & x < 1, \\ 0, & x > 1, \end{cases}$$

and the source variables are taken as

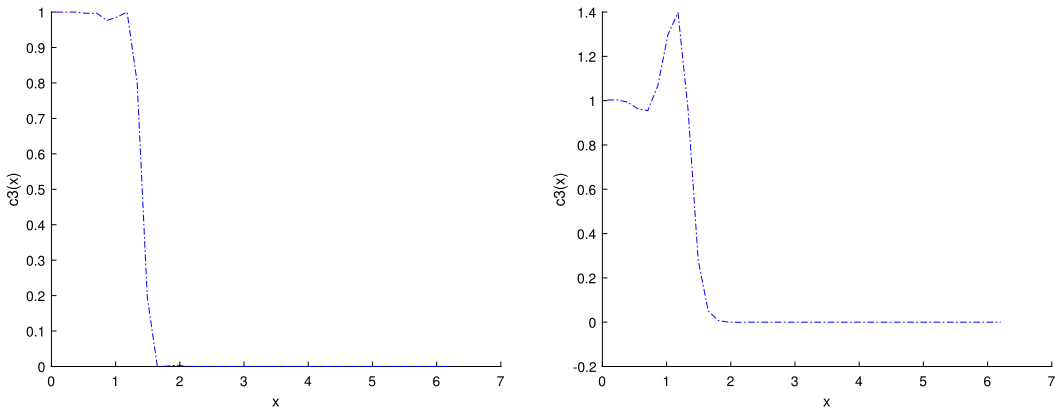
$$q = 0, \quad \tilde{c} = 0.$$

Other parameters are chosen as

$$\phi(x) = \mu(c) = k(x) = z_1 = z_2 = z_3 = 1, \quad D(u) = \gamma.$$

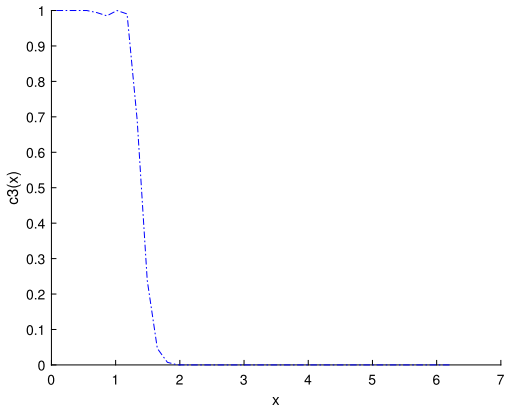
where $\gamma = 10^{-5}$.

We apply the WENO algorithm and compute the numerical approximations up to $T = 0.1$ with $M = 40$ and $\Delta t = 0.1\Delta x^2$. Due to the "consistency" requirements of the numerical fluxes, we want the nonlinear weights to be the same in the reconstruction procedure for $\hat{u}$ and $\widehat{uc_j}$, $j = 1, 2, 3$. In the numerical simulations, we use uc_1, uc_2, uc_3, and u to obtain the nonlinear weights in the WENO reconstruction procedure, respectively. The numerical results are given in Fig. 3. We can observe that the BP technique together with the WENO algorithm can suppress oscillations significantly near the discontinuity if the weights are based on uc_1 and uc_2. However, if we use u and uc_3 to calculate the nonlinear weights, we can still observe slight oscillations near the shocks. To test whether the oscillations are suppressed by WENO or the BP technique, we carry out the following four experiments: (1) BP without WENO, (2) WENO without BP, (3) BP and WENO,

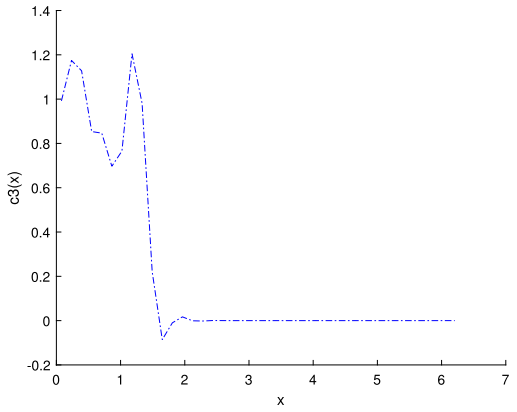


(a) BP without WENO.

(b) WENO without BP.



(c) Both BP and WENO.



(d) Neither BP nor WENO.

Fig. 4. Example 5.4: Numerical approximation of c_3 at $t = 0.1$.

(4) neither BP nor WENO. In all the experiments, the nonlinear weights in the WENO reconstruction procedure are based on the point values of uc_3, and the results are given in Fig. 4. We can observe that the WENO procedure itself cannot effectively suppress oscillations. We anticipate the oscillations are due to the discontinuity in the source term p_t. To verify your anticipation, we choose a continuous $p(x,t)$ as

$$p(x,t) = e^{-t}\sin(x),$$

and repeat the four experiments given above. The numerical results are given in Fig. 5. We can see that the WENO procedure can effectively suppress the oscillations. However, if we use WENO procedure only, the numerical approximations of c can be 1.003, which is greater than 1, see Fig. 6. Therefore, though p is continuous, we still need the BP technique to make the numerical approximations to be physically relevant.

5.2. Two dimensional case

In this subsection, we solve (1.1)-(1.2) in two space dimensions. The computational domain is set to be $\Omega = [0, 2\pi] \times [0, 2\pi]$. We first construct an analytical solutions and test the accuracy of the BP FD schemes

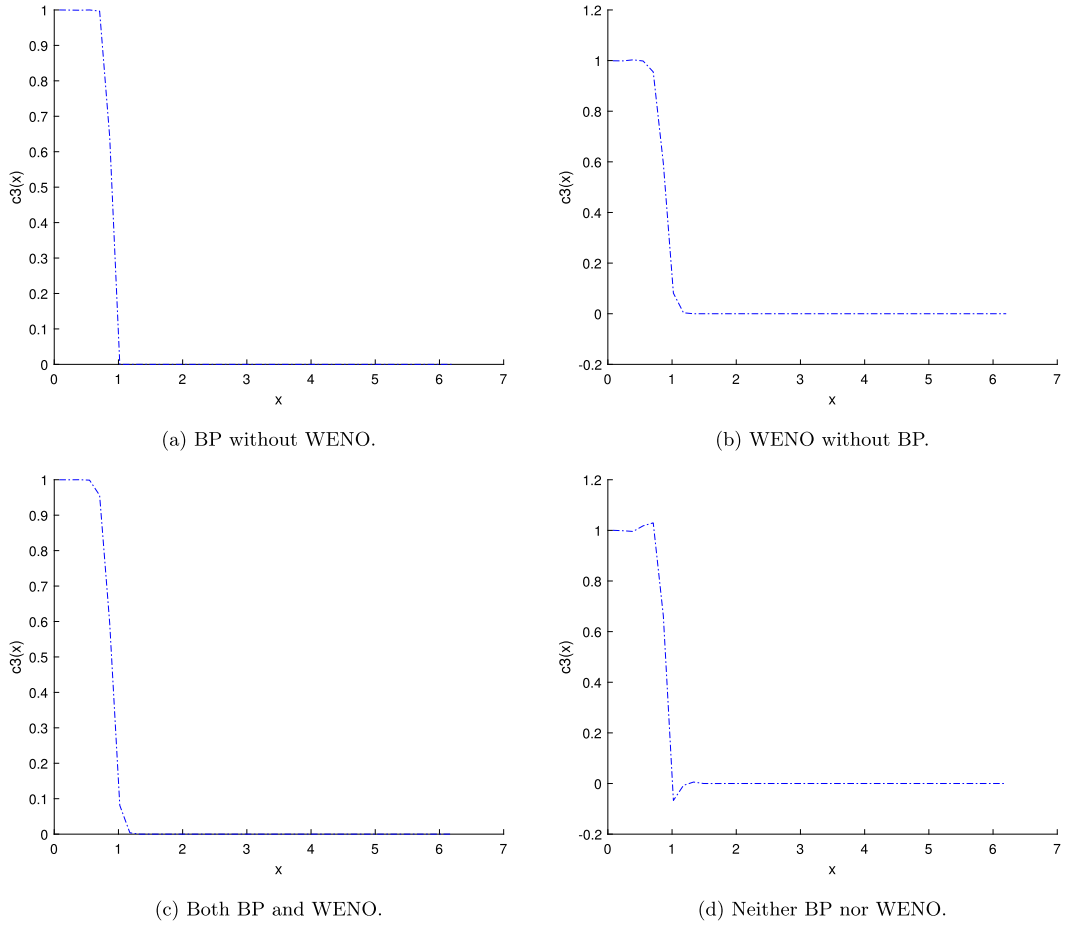
Example 5.5. We choose the initial condition as

$$c(x, y, 0) = \sin^4(x + y), \quad p(x, y, 0) = -x - y.$$

and the source variables are taken as

$$q = \gamma, \quad \tilde{c} = 0.$$

Other parameters are chosen to be



(a) BP without WENO.

(b) WENO without BP.

(c) Both BP and WENO.

(d) Neither BP nor WENO.

Fig. 5. Example 5.4: Numerical approximation of c_3 at $t = 0.1$.

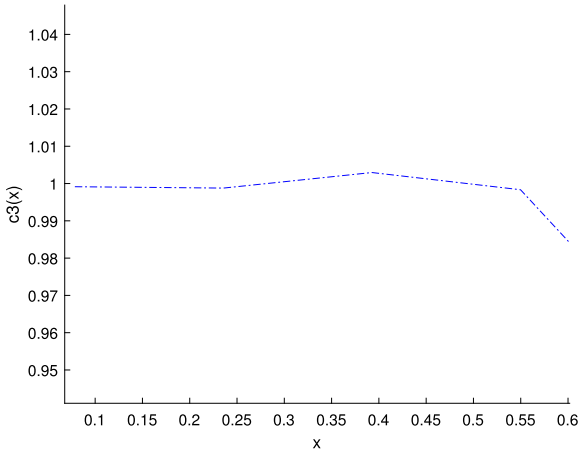


Fig. 6. Example 5.4: Numerical approximation of c_3 at $t = 0.1$ near the discontinuity.

$$\phi(x, y) = \mu(c) = k(x, y) = z_1 = z_2 = 1, \quad \mathbf{D}(\mathbf{u}) = 0.$$

We choose suitable Dirichlet boundary conditions for the pressure equation and periodic boundary condition for the concentration equation such that the exact solutions are

$$c(x, y, t) = e^{-\gamma t}(\sin^4(x + y - 2t)), \quad p(x, y, t) = \gamma t - x - y,$$

Table 2
Example 5.5: Accuracy test for the fifth-order FD scheme with and without BP technique.

M	Without limiter			With limiter		
	L^∞ error	Order	Minimum value	L^∞ error	Order	Minimum value
20	4.30e−03	–	−3.50e−03	4.30e−03	–	2.04e−06
40	1.61e−04	4.74	−1.25e−04	1.61e−04	4.74	1.95e−13
80	5.34e−06	4.91	−3.10e−06	7.64e−06	4.40	6.52e−11
160	1.68e−07	4.99	−1.48e−07	2.25e−07	5.09	2.00e−13
320	5.29e−09	4.99	−4.48e−09	8.17e−09	4.78	2.95e−10

where $\gamma = 10^{-5}$. In the numerical simulation, we choose $\Delta t = 0.1 \min\{\Delta x^2, \Delta y^2\}$ and compute up to $T = 0.1$. We calculate the L^∞-norm of the error between the numerical and exact solutions of c, and the results are given in Table 2. We can observe fifth-order accuracy of the FD scheme with and without the BP technique. From the table, negative values emerged in the case of without BP limiter, and, again, the parameterized PP flux limiters remedied the negative values in a conservative way. Therefore, the BP technique does not kill the accuracy when it works.

Example 5.6. We investigate the displacement of 3-phase porous media flow in the five-spot arrangement of injection and production wells. The computational domain is a square region taken as quarter-of-a-five-spot pattern. The three phases are light oil c_1 (with low viscosity and high compressibility), heavy oil c_2 (with high viscosity and low compressibility) and water c_3 (with medium viscosity and medium compressibility).

The initial concentrations of oil (water) are

$$c_{1,0}(x, y) = \begin{cases} 1, & x \le \frac{\pi}{2}, y \le \frac{\pi}{2}, \\ 0, & \text{otherwise}, \end{cases} \qquad c_{2,0}(x, y) = \begin{cases} 0, & x \le \frac{\pi}{2}, y \le \frac{\pi}{2}, \\ 1, & \text{otherwise}, \end{cases} \qquad c_{3,0}(x, y) = 0.$$

Therefore, the lower-left part of the region is light oil enrichment area while the other part is heavy oil enrichment area. Moreover, no water exists initially and the initial pressure is taken as 0 in the whole computational domain. To simulate the random perturbation of porosity and permeability around their average value, we choose the porosity and permeability as

$$\phi(x, y) = 0.5 + 0.05 \sin(5x) \sin(5y) \quad \text{and} \quad k(x, y) = 1.0 + 0.1 \cos(5x) \cos(5y),$$

respectively. Other parameters are taken as

$$\mu(c_1, c_2, c_3) = 0.4c_1 + 2.0c_2 + 1.0c_3, \quad z_1 = 1.2, \quad z_2 = 0.8, \quad z_3 = 1.0, \quad \mathbf{D} = diag(\gamma, \gamma).$$

where $\gamma = 0.01$.

The injection well is located in lower-left corner with $q = \frac{1}{\Delta x \Delta y}$; and production well is located in upper-right corner with $q = -\frac{1}{\Delta x \Delta y}$.

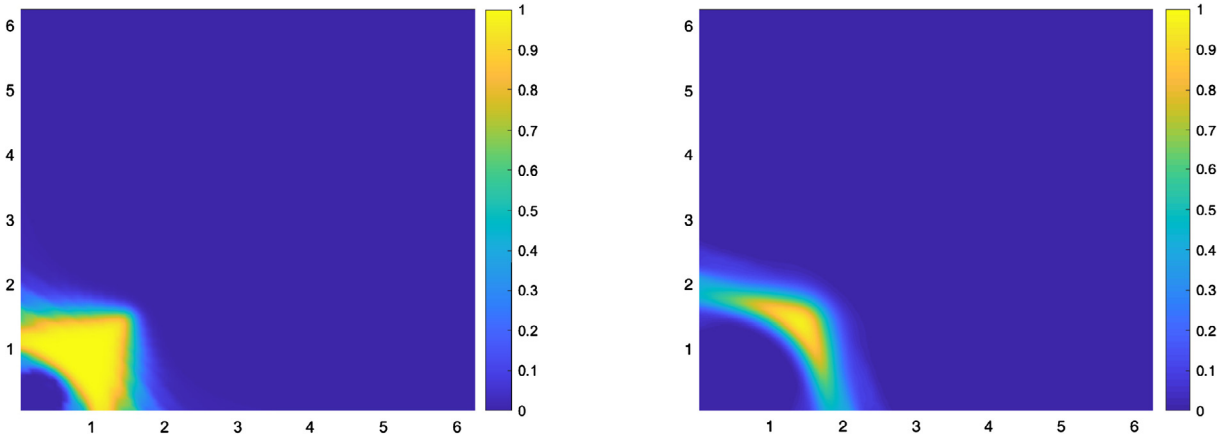
This example is used for petroleum production simulations. We compute the components c_1 and c_2 at time $T = 0.2, 0.8$ with $N_x = N_y = 80$ and $\Delta t = 0.1 \min\{\Delta x^2, \Delta y^2\}$. The distributions of c_1, c_2 and $c_1 + c_2$ at different time are shown in Figs. 7a-7f, respectively. From the figure we can see that c_1, c_2 and $c_1 + c_2$ are all between 0 and 1.

Example 5.7. To show the significance of the bound-preserving technique in real petroleum production simulations, we choose the exact parameters in Example 5.6, expect $\mathbf{D} = \mathbf{0}$ in order to avoid any dissipation to the scheme which is resulted from the diffusion term.

This example is used for petroleum production simulations when diffusion effect is negligible. We compute the components c_1 and c_2 at time $T = 0.2, 0.8$ with $N_x = N_y = 80$ and $\Delta t = 0.1 \min\{\Delta x^2, \Delta y^2\}$. The distributions of c_1, c_2 and c_3 with BP technique at different time along diagonal $y = x$ are shown in Figs. 8a-8f, respectively. From the figure we can see that c_1, c_2 and c_3 are all between 0 and 1. We also solve the problem without BP technique, the distributions of c_1, c_2 and c_3 along diagonal at different time are shown in Figs. 9a-9f, from which we can observe strong oscillations and physically irrelevant values. This implies the necessity of the BP technique.

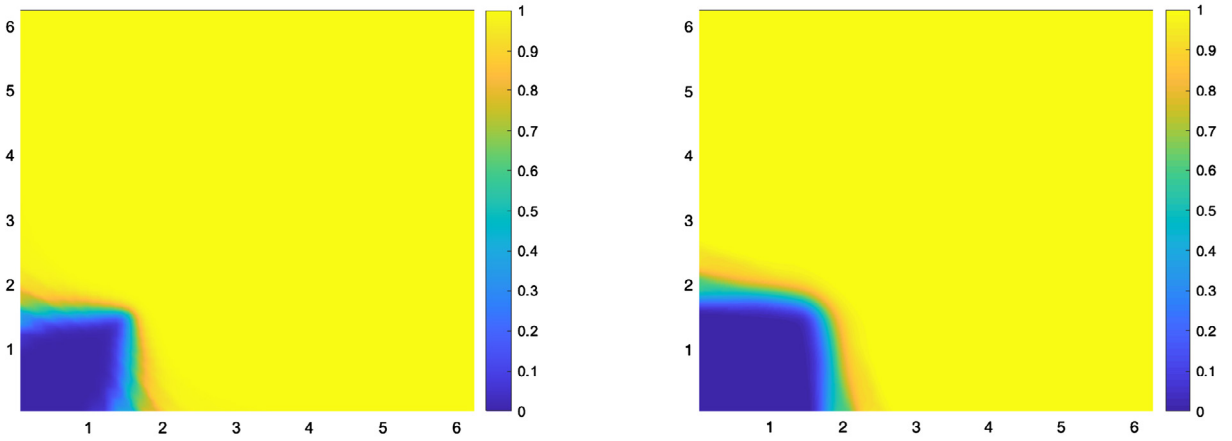
6. Concluding remarks

In this paper, we constructed high-order BP FD methods for compressible miscible displacements in porous media on rectangular meshes. We have applied the technique to the problem with multi-component fluid mixtures. Numerical simulations have shown the accuracy and necessity of the BP technique.



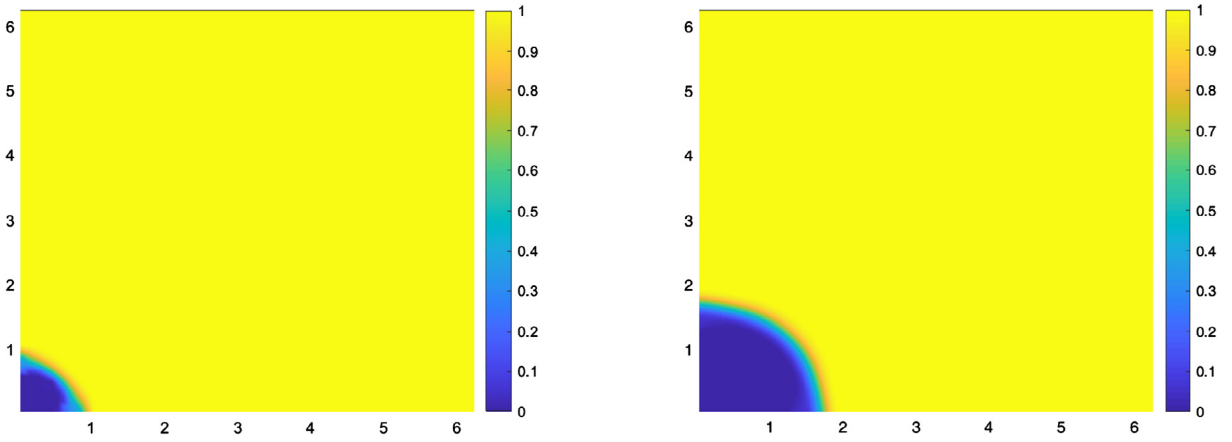
(a) c_1 at T=0.2 s

(b) c_1 at T=0.8 s



(c) c_2 at T=0.2 s

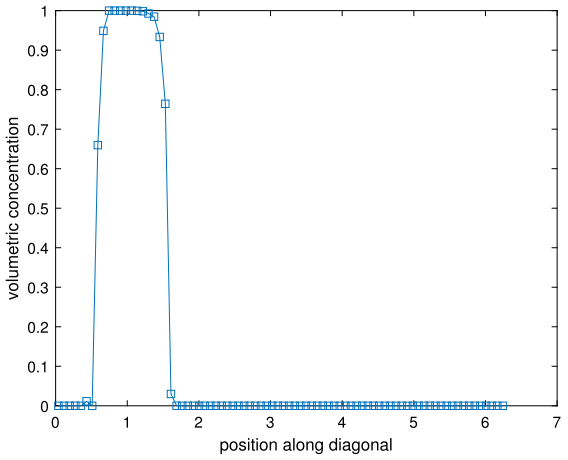
(d) c_2 at T=0.8 s



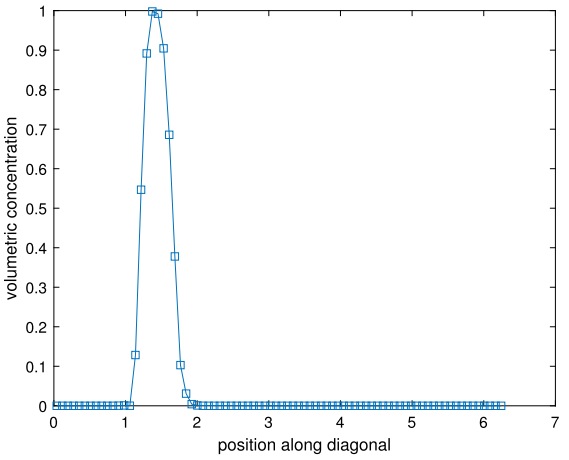
(e) $c_1 + c_2$ at T=0.2 s

(f) $c_1 + c_2$ at T=0.8 s

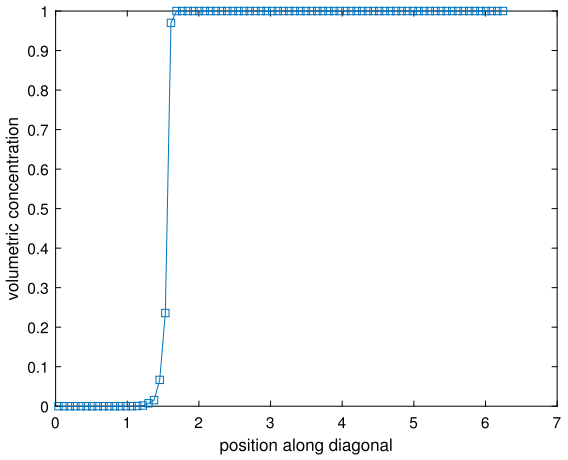
Fig. 7. Example 5.6: Concentrations of c_1, c_2 and $c_1 + c_2$.



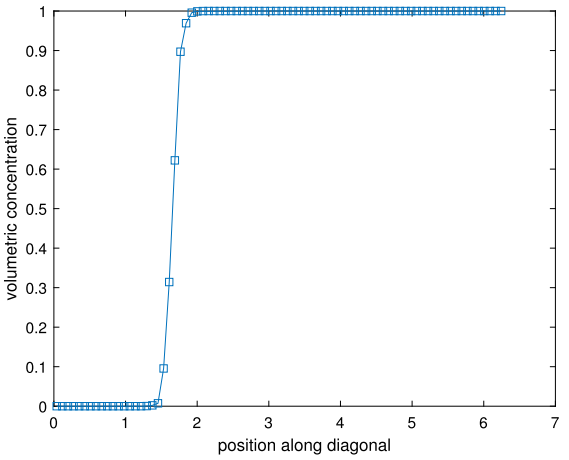
(a) c_1 at T=0.2 s



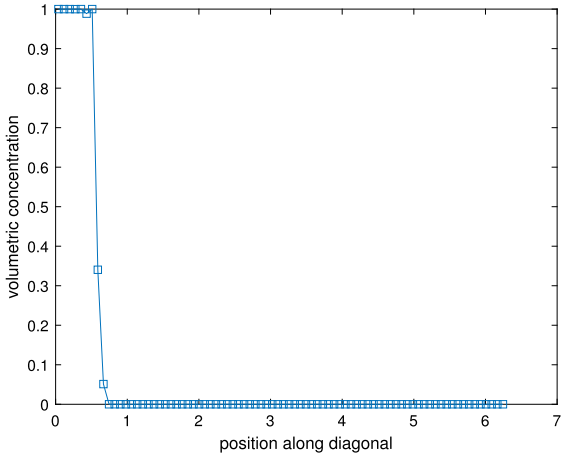
(b) c_1 at T=0.8 s



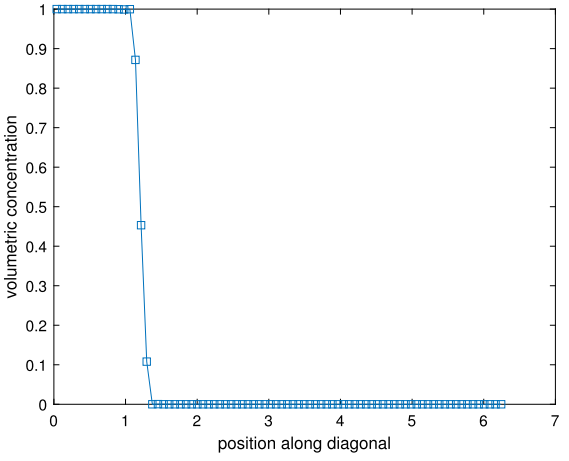
(c) c_2 at T=0.2 s



(d) c_2 at T=0.8 s

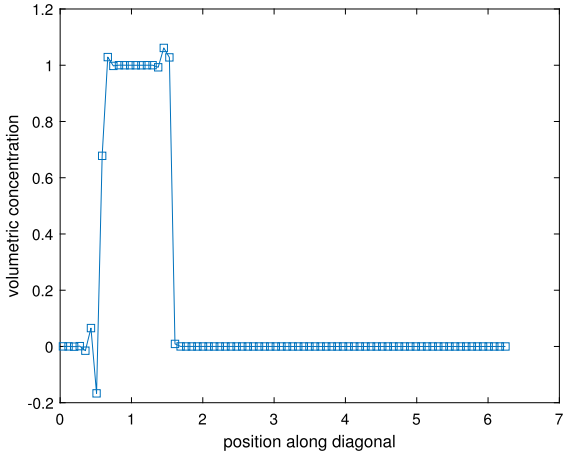


(e) c_3 at T=0.2 s

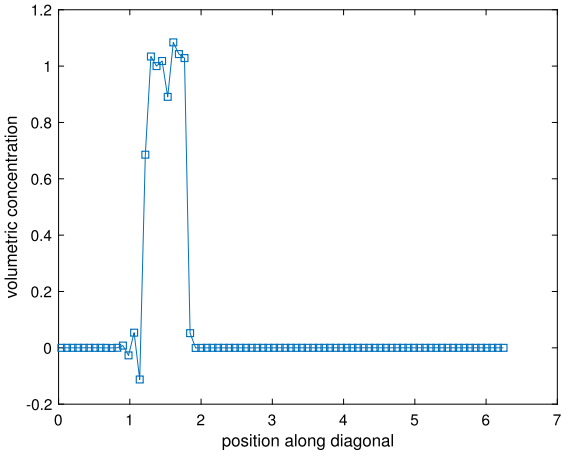


(f) c_3 at T=0.8 s

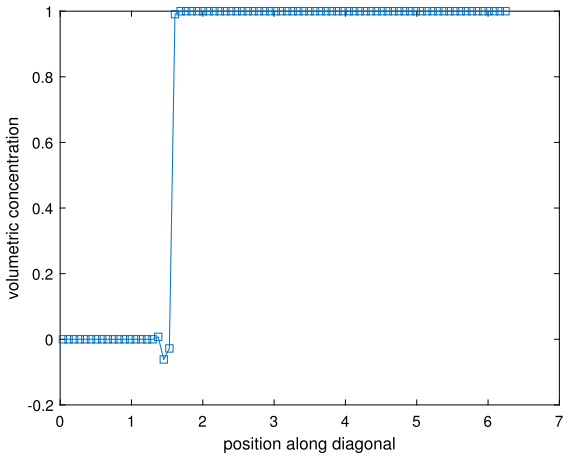
Fig. 8. Example 5.7: Concentrations of c_1, c_2 and c_3.



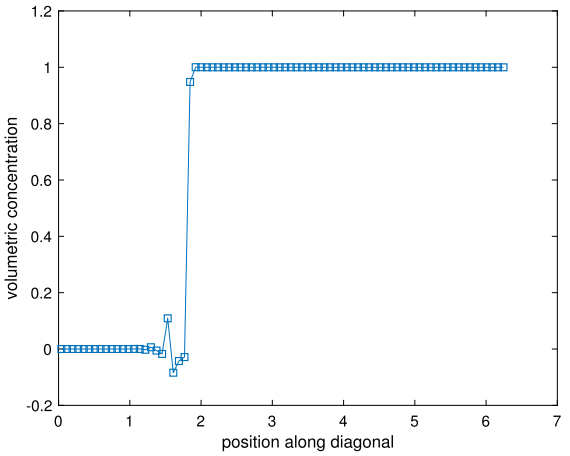
(a) c_1 at T=0.2 s



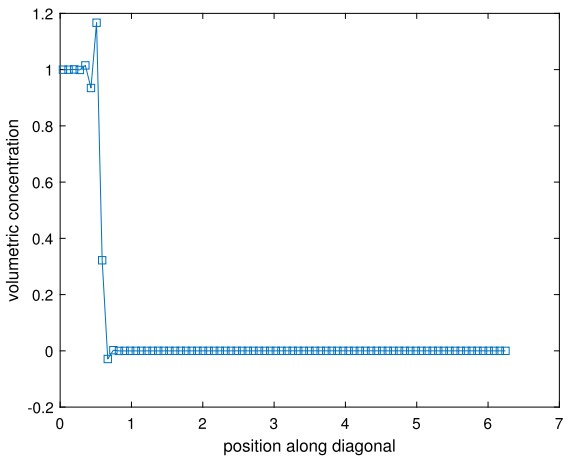
(b) c_1 at T=0.8 s



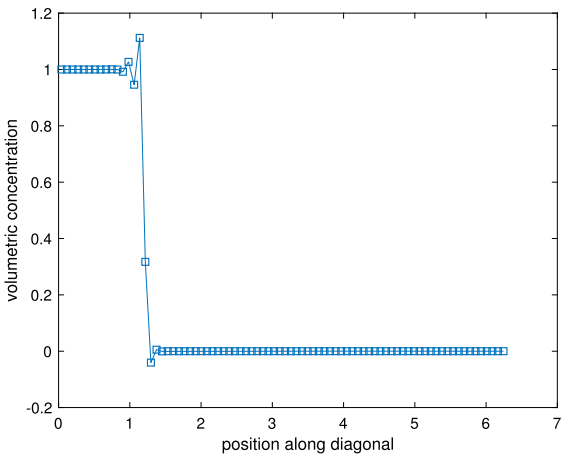
(c) c_2 at T=0.2 s



(d) c_2 at T=0.8 s



(e) c_3 at T=0.2 s



(f) c_3 at T=0.8 s

Fig. 9. Example 5.7: Concentrations of c_1, c_2 and c_3.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

In Section 2, we used reconstruction function $f^{\pm}$ to obtain the numerical flux $\widehat{uc}_j$. In this appendix, we demonstrate how to construct $f^{\pm}$. Let m, an odd integer in general, be the desired accuracy of the scheme, we take $\ell = (m+1)/2$. We only demonstrate how to obtain $f^{+}(v_{i-\ell+1}, \cdots, v_{i+\ell-1})$ at $x = x_{i+\frac{1}{2}}$ and the case for f^{-} is mirror symmetric with respect to $x_{i+\frac{1}{2}}$. Therefore, in the rest of the appendix, we will skip the superscript $+$, and use f for f^{+}. To obtain the reconstruction function f, we choose ℓ candidate stencils:

$$I_r = \{x_{i-r}, \cdots, x_{i-r+\ell-1}\}, \quad r = 0, \cdots, \ell-1, \tag{A.1}$$

which produce ℓ different values of f via the reconstruction procedure:

$$f^{(r)} = \sum_{s=0}^{\ell-1} c_{r,s} v_{i-r+s}, \quad r = 0, \cdots, \ell-1, \tag{A.2}$$

where the constants $c_{r,s}$ can be obtained in [22] and satisfy

$$\sum_{s=0}^{\ell-1} c_{r,s} = 1, \tag{A.3}$$

for all r. For example, in the numerical experiments we took $m = 5$, then $\ell = 2$ and we can choose

$$\begin{aligned} f^{(0)} &= \frac{1}{3} v_i + \frac{5}{6} v_{i+1} - \frac{1}{6} v_{i+2}, \\ f^{(1)} &= -\frac{1}{6} v_{i-1} + \frac{5}{6} v_i + \frac{1}{3} v_{i+1}, \\ f^{(2)} &= \frac{1}{3} v_{i-2} + -\frac{7}{6} v_{i-1} + \frac{11}{6} v_i. \end{aligned}$$

The WENO reconstruction would take a convex combination of all $f^{(r)}$ defined in (A.2) as a new approximation:

$$f = \sum_{r=0}^{\ell-1} \omega_r f^{(r)}. \tag{A.4}$$

The weights ω_r are the key to the success of WENO. We need

$$\omega_r \geq 0, \quad \sum_{r=0}^{\ell-1} \omega_r = 1 \tag{A.5}$$

for stability and consistency.

Before calculating the nonlinear weights ω_r, we need linear weights d_r, which satisfies the same properties as ω_r. In this paper, we choose

$$d_0 = \frac{3}{10}, \quad d_1 = \frac{3}{5}, \quad d_2 = \frac{1}{10}. \tag{A.6}$$

Then we can calculate the nonlinear weights as

$$\omega_r = \frac{\alpha_r}{\sum_{s=0}^{\ell-1} \alpha_s}, \quad r = 0, \cdots, \ell-1, \tag{A.7}$$

with

$$\alpha_r = \frac{d_r}{(\epsilon + \beta_r)^2}. \tag{A.8}$$

Here ϵ is chosen to be a very small positive number to avoid the denominator being 0. In our numerical experiments, we take $\epsilon = 10^{-6}$. β_r are the so-called "smoothness indicators" of I_r given by

$$\beta_0 = \frac{13}{12}(v_i - 2v_{i+1} + v_{i+2})^2 + \frac{1}{4}(3v_i - 4v_{i+1} + v_{i+2})^2,$$
$$\beta_1 = \frac{13}{12}(v_{i-1} - 2v_i + v_{i+1})^2 + \frac{1}{4}(v_{i-1} - v_{i+1})^2,$$
$$\beta_2 = \frac{13}{12}(v_{i-2} - 2v_{i-1} + v_i)^2 + \frac{1}{4}(v_{i-2} - 4v_{i-1} + 3v_i)^2.$$

References

[1] D.S. Balsara, C.-W. Shu, Monotonicity preserving weighted essentially non-oscillatory schemes with increasingly high order of accuracy, J. Comput. Phys. 160 (2000) 405–452.
[2] H.-Z. Chen, H. Wang, An optimal-order error estimate on an H^1-Galerkin mixed method for a nonlinear parabolic equation in porous medium flow, Numer. Methods Partial Differ. Equ. 26 (2010) 188–205.
[3] S.-H. Chou, Q. Li, Mixed finite element methods for compressible miscible displacement in porous media, Math. Comput. 57 (1991) 507–527.
[4] N. Chuenjarern, Z. Xu, Y. Yang, High-order bound-preserving discontinuous Galerkin methods for compressible miscible displacements in porous media on triangular meshes, J. Comput. Phys. 378 (2019) 110–128.
[5] M. Cui, A combined mixed and discontinuous Galerkin method for compressible miscible displacement problem in porous media, J. Comput. Appl. Math. 198 (2007) 19–34.
[6] M. Cui, Analysis of a semidiscrete discontinuous Galerkin scheme for compressible miscible displacement problem, J. Comput. Appl. Math. 214 (2008) 617–636.
[7] J. Douglas Jr., R.E. Ewing, M.F. Wheeler, A time-discretization procedure for a mixed finite element approximation of miscible displacement in porous media, RAIRO. Anal. Numér. 17 (1983) 249–256.
[8] J. Douglas Jr., R.E. Ewing, M.F. Wheeler, The approximation of the pressure by a mixed method in the simulation of miscible displacement, RAIRO. Anal. Numér. 17 (1983) 17–33.
[9] J. Douglas Jr., J. Roberts, Numerical methods for a model for compressible miscible displacement in porous media, Math. Comput. 41 (1983) 441–459.
[10] A.O. Garder Jr., D.W. Peaceman, A.L. Pozzi Jr., Numerical calculation of multidimensional miscible displacement by the method of characteristics, Soc. Pet. Eng. J. 4 (1964) 683.
[11] S. Gottlieb, D. Ketcheson, C.-W. Shu, High order strong stability preserving time discretizations, J. Sci. Comput. 38 (2009) 251–289.
[12] S. Gottlieb, C.-W. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. 43 (1) (2001) 89–112.
[13] H. Guo, Y. Yang, Bound-preserving discontinuous Galerkin method for compressible miscible displacement in porous media, SIAM J. Sci. Comput. 39 (2017), A1969–A1990.
[14] H. Guo, Q. Zhang, Error analysis of the semi-discrete local discontinuous Galerkin method for compressible miscible displacement problem in porous media, Appl. Math. Comput. 259 (2015) 88–105.
[15] L. Guo, Y. Yang, Positivity preserving high-order local discontinuous Galerkin method for parabolic equations with blow-up solutions, J. Comput. Phys. 289 (2015) 181–195.
[16] G. Jiang, C.-W. Shu, Efficient implementation of weighted ENO schemes, J. Comput. Phys. 126 (1996) 202–228.
[17] Y. Jiang, Z. Xu, Parametrized maximum principle preserving limiter for finite difference WENO schemes solving convection-dominated diffusion equations, SIAM J. Sci. Comput. 35 (6) (2013) A2524–A2553.
[18] X.-D. Liu, S. Osher, T. Chan, Weighted essentially nonoscillatory schemes, J. Comput. Phys. 115 (1994) 200–212.
[19] N. Ma, D. Yang, T. Lu, L^2-norm error bounds of characteristics collocation method for compressible miscible displacement in porous media, Int. J. Numer. Anal. Model. 2 (2005) 28–42.
[20] D.W. Peaceman, H.H. Rachford Jr., Numerical calculation of multidimensional miscible displacement, Soc. Pet. Eng. J. 2 (1962) 471.
[21] C.-W. Shu, Total-variation-diminishing time discretizations, SIAM J. Sci. Stat. Comput. 9 (1988) 1073–1084.
[22] C.-W. Shu, Essentially Non-Oscillatory and Weighted Essentially Non-Oscillatory Schemes for Hyperbolic Conservation Laws, Technical Report, 1997.
[23] C.-W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, J. Comput. Phys. 77 (1988) 439–471.
[24] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons, G. Qin, An approximation to miscible fluid flows in porous media with point sources and sinks by an Eulerian-Lagrangian localized adjoint method and mixed finite element methods, SIAM J. Sci. Comput. 22 (2000) 561–581.
[25] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons, G. Qin, An accurate approximation to compressible flow in porous media with wells, in: Numerical Treatment of Multiphase Flows in Porous Media, in: Lecture Notes in Physics, vol. 552, 2000, pp. 324–332.
[26] T. Xiong, J.-M. Qiu, Z. Xu, A parametrized maximum principle preserving flux limiter for finite difference RK-WENO schemes with applications in incompressible flows, J. Comput. Phys. 252 (2013) 310–331.
[27] T. Xiong, J.-M. Qiu, Z. Xu, High order maximum-principle-preserving discontinuous Galerkin method for convection-diffusion equations, SIAM J. Sci. Comput. 37 (2015) A583–A608.
[28] Z. Xu, Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem, Math. Comput. 83 (2014) 310–331.
[29] D. Yang, A splitting positive definite mixed element method for miscible displacement of compressible flow in porous media, Numer. Methods Partial Differ. Equ. 17 (2001) 229–249.
[30] J. Yang, Y. Chen, A priori error estimates of a combined mixed finite element and discontinuous Galerkin method for compressible miscible displacement with molecular diffusion and dispersion, J. Comput. Math. 28 (2010) 1005–1022.
[31] J. Yang, A posteriori error of a discontinuous Galerkin scheme for compressible miscible displacement problems with molecular diffusion and dispersion, Int. J. Numer. Methods Fluids 65 (2011) 781–797.
[32] J. Yang, Y. Chen, A priori error analysis of a discontinuous Galerkin approximation for a kind of compressible miscible displacement problems, Sci. China Math. 53 (2010) 2679–2696.
[33] F. Yu, H. Guo, N. Chuenjarern, Y. Yang, Conservative local discontinuous Galerkin method for compressible miscible displacements in porous media, J. Sci. Comput. 73 (2017) 1249–1275.
[34] Y. Yuan, The characteristic finite difference fractional steps methods for compressible two-phase displacement problem, Sci. China Ser. A 42 (1999) 48–57.
[35] Y. Yuan, The upwind finite difference fractional steps methods for two-phase compressible flow in porous media, Numer. Methods Partial Differ. Equ. 19 (2003) 67–88.
[36] Y. Yuan, The modified upwind finite difference fractional steps method for compressible two-phase displacement problem, Acta Math. Appl. Sin. 20 (2004) 381–396.
[37] X. Zhang, C.-W. Shu, On maximum-principle-satisfying high order schemes for scalar conservation laws, J. Comput. Phys. 229 (2010) 3091–3120.

[38] X. Zhang, C.-W. Shu, On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes, J. Comput. Phys. 229 (2010) 8918–8934.
[39] X. Zhang, C.-W. Shu, Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms, J. Comput. Phys. 230 (2011) 1238–1248.
[40] X. Zhang, Y. Xia, C.-W. Shu, Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes, J. Sci. Comput. 50 (2012) 29–32.
[41] 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.