SIAM J. SCI. COMPUT.
Vol. 41, No. 2, pp. B250–B273

HIGH-ORDER BOUND-PRESERVING DISCONTINUOUS GALERKIN METHODS FOR STIFF MULTISPECIES DETONATION*

JIE DU†, CHENG WANG‡, CHENGENG QIAN‡, AND YANG YANG§

Abstract. In this paper, we develop high-order bound-preserving discontinuous Galerkin (DG) methods for multispecies and multireaction chemical reactive flows. In this problem, density and pressure are nonnegative, and the mass fraction for the ith species, denoted as z_i, $1 \le i \le M$, should be between 0 and 1, where M is the total number of species. In [C. Wang, X. Zhang, C.-W. Shu, and J. Ning, *J. Comput. Phys.*, 231 (2012), pp. 653–665], the authors have introduced the positivity-preserving technique that guarantees the positivity of the numerical density, pressure, and the mass fraction of the first $M-1$ species. However, the extension to preserve the upper bound 1 of the mass fraction is not straightforward. There are three main difficulties. First of all, the time discretization in [C. Wang, X. Zhang, C.-W. Shu, and J. Ning, *J. Comput. Phys.*, 231 (2012), pp. 653–665] was based on Euler forward. Therefore, for problems with stiff source, the time step will be significantly limited. Secondly, the mass fraction does not satisfy a maximum principle, and most of the previous techniques cannot be applied. Thirdly, in most of the previous works for gaseous denotation, the algorithm relies on the second-order Strang splitting methods where the flux and stiff source terms can be solved separately, and the extension to high-order time discretization seems to be complicated. In this paper, we will solve all the three problems given above. The high-order time integration does not depend on the Strang splitting; i.e., we do not split the flux and the stiff source terms. Moreover, the time discretization is explicit and can handle the stiff source with large time step. Most importantly, in addition to the positivity-preserving property introduced in [C. Wang, X. Zhang, C.-W. Shu, and J. Ning, *J. Comput. Phys.*, 231 (2012), pp. 653–665], the algorithm can preserve the upper bound 1 for each species. Numerical experiments will be given to demonstrate the good performance of the bound-preserving technique and the stability of the scheme for problems with stiff source terms.

Key words. discontinuous Galerkin method, bound-preserving, mass fraction, stiff source, detonation

AMS subject classification. 65M60

DOI. 10.1137/18M122265X

1. Introduction.

In this paper, we develop high-order bound-preserving discontinuous Galerkin (DG) methods for multispecies and multireaction chemical reactive flows and investigate the following convection-reaction equation in two space dimensions

$$\rho_t + m_x + n_y = 0, \tag{1.1a}$$

$$m_t + (mu + p)_x + (nu)_y = 0, \tag{1.1b}$$

$$n_t + (mv)_x + (nv + p)_y = 0, \tag{1.1c}$$

$$E_t + ((E + p)u)_x + ((E + p)v)_y = 0, \tag{1.1d}$$

*Submitted to the journal's Computational Methods in Science and Engineering section October 24, 2018; accepted for publication (in revised form) January 28, 2019; published electronically March 26, 2019.

http://www.siam.org/journals/sisc/41-2/M122265.html

Funding: The work of first author was supported by National Natural Science Foundation of China under grant NSFC 11801302. The work of the second and third authors was supported by National Key R & D Program of China (2017YFC0804700), the National Natural Science Foundation of China under grants 11732003, Science Challenge Project (TZ2016001), and Beijing Natural Science Foundation (8182050). The work of the fourth author was supported by NSF grant DMS-1818467.

†Yau Mathematical Sciences Center, Beijing, 100084 China (jdu@tsinghua.edu.cn).

‡State Key Laboratory of Explosion Science and Technology, Beijing, 100081 China (wangcheng@bit.edu.cn, 3120170079@ bit.edu.cn).

§Michigan Technological University, Houghton, MI, 49931 (yyang7@mtu.edu).

$$(r_1)_t + (mz_1)_x + (nz_1)_y = s_1 \tag{1.1e}$$

$$\cdots$$

$$(r_{M-1})_t + (mz_{M-1})_x + (nz_{M-1})_y = s_{M-1}, \tag{1.1f}$$

where ρ, u, v, $m = \rho u$, $n = \rho v$, E, and p are the total density, velocity in x direction, velocity in y direction, momentum in x direction, momentum in y direction, the total energy, and pressure, respectively. M is the total number of chemical species. For $1 \le i \le M$, $r_i = \rho z_i$ with z_i being the mass fraction for the ith species, and $\sum_{i=1}^{M} z_i = 1$. Therefore, we have $\sum_{i=1}^{M} r_i = \rho$ and $0 \le z_i \le 1$. The equation of state is given as

$$p = (\gamma - 1)\left(E - \frac{1}{2}\rho(u^2 + v^2) - \rho z_1 q_1 - \cdots - \rho z_M q_M\right),$$

where q_i is the enthalpy of formation for the ith species, and the temperature is defined as $T = p/\rho$. The s_i given in the source term describes the chemical reactions. We consider R reactions of the form

$$\nu'_{1,r}X_1 + \nu'_{2,r}X_2 + \cdots + \nu'_{M,r}X_M \to \nu''_{1,r}X_1 + \nu''_{2,r}X_2 + \cdots + \nu''_{M,r}X_M, \quad r = 1, 2, \ldots, R,$$

where $\nu'_{i,r}$ and $\nu''_{i,r}$ are the stoichiometric coefficients of the reactants and products, respectively, of the ith species in the rth reaction. For nonequilibrium chemistry, the rate of production of the ith species can be written as

$$s_i = M_i \sum_{r=1}^{R} (\nu''_{i,r} - \nu'_{i,r}) \left[k_r(T) \prod_{j=1}^{M} \left(\frac{r_j}{M_j}\right)^{\nu'_{j,r}} \right], \quad i = 1, 2, \ldots, M,$$

where M_i is the molar mass of the ith species. $k_r(T)$, a function of the temperature T, indicates the reaction rate. In this paper, we take

$$k_r(T) = \begin{cases} B_r T^{\alpha_r}, & T > T_r, \\ 0, & T \le T_r, \end{cases}$$

where T_r is the ignition temperature for the rth reaction, and B_r and α_r are pre-exponential factor and index of temperature, respectively. Moreover, it is easy to check that $\sum_{i=1}^{M} s_i = 0$. Therefore, using the fact $\sum_{i=1}^{M} z_i = 1$, we can subtract (1.1e)–(1.1f) from (1.1a) to obtain a new equation

$$(r_M)_t + (mz_M)_x + (nz_M)_y = s_M, \tag{1.2}$$

which is similar to (1.1e)–(1.1f), and this can help us construct the bound-preserving technique.

Numerical simulations of wave propagation in gaseous detonation are essential for minimizing devastating hazards. However, the single-step model could not predict the correct ignition process of the mixture. It was argued that using detailed chemical model would reproduce results that agree with the experimental data. However, there are some challenges in the simulations of detonation wave using detailed chemical model due to complexity of chemical kinetics. Thus, designing an efficient and accurate numerical method is of practical importance. However, the construction of the numerical methods is not an easy task. There are three main difficulties. Firstly, the reaction speed of the chemical species is extremely fast, leading to stiff source

terms in the model system; see, e.g., [2, 9]. Hence, the time step would be rather small if some explicit time integration, such as Euler forward, is applied. Secondly, the exact solution of the model may contain shocks, and the direct numerical simulation may yield nonphysical numerical approximations; i.e., the density and pressure can be negative, and the mass fraction may not be between 0 and 1, especially for high-order numerical schemes, leading to ill-posedness of the problems and the numerical simulations will blow-up. Therefore, some special techniques should be constructed to make the numerical approximation to be physically relevant. Finally, due to the stiff source, direct numerical simulations on coarse meshes may yield nonphysical shock waves; see, e.g., [9] for the discussion. In this paper, we will focus on the first two problems and construct suitable high-order numerical schemes with large time steps that can preserve the physical bounds. We will extend the idea to deal with the last problem in the future. We would like to apply the DG method, as it is high-order accurate and uses piecewise polynomials as the numerical approximation and hence is easy to apply limiters.

The DG method, first introduced by Reed and Hill [14] in the framework of neutron linear transport, gained even greater popularity for good stability, high order accuracy, and flexibility on hp-adaptivity and on complex geometry. There were some previous works discussing DG methods in solving gaseous denotation; see [10, 11] as an incomplete list. However, none of them focused on the bound-preserving technique. Physically bound-preserving high-order numerical methods for conservation laws have been actively studied in the last few years. In [21], genuinely maximum-principle-preserving high-order DG schemes for scalar conservation laws and two-dimensional incompressible flows in vorticity-streamfunction formulation have been constructed. Subsequently, positivity-preserving (PP) high-order DG schemes for compressible Euler equations on rectangular meshes were given in [22], and the extension to triangular meshes was given in [24]. Later, the technique was applied to other hyperbolic systems, such as pressureless Euler equations [20], extended MHD equations [25], relativistic hydrodynamics [13], etc., and the L^1 stability was demonstrated. In [23], the authors studied the compressible Euler equations with source terms, and the idea was later extended to gaseous detonation in [18] to preserve the positivity of density, pressure, and all the mass fractions except the last one. The basic idea of the PP technique in [18] is to apply Euler forward time discretization and take the test function to be 1 in each cell to obtain an equation of the numerical cell average of the target variable, say r, and prove that the cell average, $\bar{r}$, is positive. Then we can apply a slope limiter to the numerical approximation and construct a new one:

$$\tilde{r} = \bar{r} + \theta(r - \bar{r}), \quad \theta \in [0, 1].$$

The extension to high-order time discretization is based on the strong-stability-preserving (SSP) Runge–Kutta (RK)/multistep methods [4, 15, 16], which can be written as convex combinations of Euler forwards. It is not easy to extend the idea in [18] to preserve the upper bound 1 for the mass fraction. First of all, most of the previous works that preserve two bounds (see, for example, [21, 24]), are based on the maximum-principle-preserving technique. However, the mass fraction z_i does not satisfy a maximum-principle. Recently, one of the authors studied miscible displacements in porous media and constructed a second-order DG scheme that preserves the two bounds 0 and 1 for the volumetric percentage in [5] on rectangular meshes, and the extension to triangular meshes has been given in [1]. In this paper, we follow the ideas given in [5, 1] to gaseous detonation to construct high-order DG schemes on general rectangular and triangular meshes. The basic idea is to apply the PP technique

to each r_i (or z_i) and enforce $\sum_{i=1}^{M} r_i = \rho$ (or $\sum_{i=1}^{M} z_i = 1$) by choosing consistent fluxes (see Definition 3.1). Then each z_i would be between 0 and 1. The second difficulty is the construction of high-order time integration for the stiff source term. The time discretization in the analysis in [18, 5, 1] was chosen as Euler forward method. However, in gaseous detonation, $k_r(T)$ would be a large constant, leading to an extremely stiff source s_i. Therefore, by applying the idea in [18, 5, 1], the time step will be significantly limited. One alternative is to consider backward Euler discretization and derive the PP technique. To the best knowledge of the author, the only work in this direction is given in [12], where the maximum-principle-preserving technique was investigated for hyperbolic equations. However, by using backward Euler method, the scheme is only first-order accurate in time, and the idea cannot be extended to high-order methods following [18, 5, 1] since no high-order SSP RK methods can be written as a convex combination of backward Euler methods [4]. Moreover, due to the time step restriction by the PP technique, any time integration that is the combination of Euler forward and backward Euler, such as Crank–Nicolson method, cannot be applied. Notice that the time constraint of the PP technique with Euler forward time discretization is due to the stiffness of the source. Hence, one may consult the splitting method and separate the flux and the source terms. By doing so, we can apply Euler forward time discretization for the convection term and use other suitable ODE solvers for the source term. However, the most commonly used splitting method is the second-order Strang splitting method [17], and the extension to high-order time integration is complicated. Another possible idea to construct the time integration is to apply the modified Patankar–RK scheme [7, 8]. However, the high-order schemes contain some defects as the fraction used in the trick may has zero denominator with nonzero numerator. Therefore, one has to assume the exact solution to be strictly positive. However, this may not true as one species may not appear initially and will be created during the chemical reaction. Recently, there is a new idea introduced in [6] to solve scalar hyperbolic equations with stiff source terms by using the modified exponential RK/multistep DG methods. The algorithm in [6] is not based on the splitting methods nor the Patankar–RK method. However, the idea cannot to applied to construct bound-preserving technique in the stiff multispecies detonation, since it does not preserve the total mass. Therefore, one of the necessary conditions in the bound-preserving technique, $\sum_{i=1}^{M} z_i = 1$, is not satisfied.

In this paper, we will modify the scheme introduced in [6] to preserve the total mass. Then we can apply the ideas in [5, 1] to construct the bound-preserving technique. Since the time step is not too small, it is possible to sufficiently refine the mesh to capture the correct position of the shocks. In this paper, we only discuss the bound-preserving technique on fine meshes and the numerical simulations on coarse meshes will be given in the future. Before we finish the introduction, we would like to summarize the advantages of the proposed scheme. The algorithm

1. is high-order accurate in both space and time (at least third-order accurate for multistep method);
2. is explicit and can handle stiff source term with relatively large time step;
3. is not based on the splitting technique nor the Patankar–RK methods;
4. has local mass conservation;
5. preserves the total mass;
6. preserves the bounds, such as the positivity of the density and pressure, and the two bounds 0 and 1 of the mass fraction.

The organization of this paper is as follows. In section 2, we construct the DG scheme. In section 3, we consider the flux terms only. We apply Euler forward

time integration and demonstrate the new bound-preserving technique to preserve the upper bound 1 of the mass fraction. The high-order multistep time integrations and the full algorithm will be given in section 4. The second-order RK method will be constructed in section 5. Numerical experiments will be given in section 6. We will finish in section 7 with some conclusion remarks.

2. DG methods. In this section, we will construct the DG scheme for (1.1). We rewrite (1.1) into the form of

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_{\mathbf{x}} + \mathbf{g}(\mathbf{w})_{\mathbf{y}} = \mathbf{s}(\mathbf{w}), \tag{2.1}$$

where

$$\begin{aligned}
\mathbf{w} &= (\rho, m, n, E, \rho z_1, \dots, \rho z_{M-1})^T, \\
\mathbf{f}(\mathbf{w}) &= (m, mu + p, mv, (E+p)u, mz_1, \dots, mz_{M-1})^T, \\
\mathbf{g}(\mathbf{w}) &= (n, nu, nv + p, (E+p)v, nz_1, \dots, nz_{M-1})^T, \\
\mathbf{s}(\mathbf{w}) &= (0, 0, 0, 0, s_1, \dots, s_{M-1})^T.
\end{aligned}$$

Let $\Omega_h = \{K\}$ be a quasiuniform partition of the computational domain Ω with rectangular or triangular elements. Denote h_K to be the diameter of element K, with $h = \max_K h_K$ and $|K|$ to be the area of K. We define the finite element space V_h^k as

$$V_h^k = \left\{ z : z\big|_K \in P^k(K), \forall K \in \Omega_h \right\},$$

where $P^k(K)$ denotes the set of polynomials of degree up to k in cell K.

In this paper, we also use $\mathbf{w}$ as the numerical approximations. The DG scheme is to find $\mathbf{w} \in \mathbf{V}_h = [V_h^k]^{M+3}$ such that for any test functions $\boldsymbol{\xi} \in \mathbf{V}_h$ and $K \in \Omega_h$ we have

$$\int_K \mathbf{w}_t \cdot \boldsymbol{\xi}\, d\mathbf{x} = \int_K \mathbf{F}(\mathbf{w}) \cdot \nabla \boldsymbol{\xi}\, d\mathbf{x} - \int_{\partial K} \mathbf{H}(\mathbf{w}^{int}, \mathbf{w}^{ext}, \boldsymbol{\nu}) \cdot \boldsymbol{\xi} ds + \int_K \mathbf{s}(\mathbf{w}) \cdot \boldsymbol{\xi}\, d\mathbf{x}, \tag{2.2}$$

where $\mathbf{F} = \langle \mathbf{f}, \mathbf{g} \rangle$ and $\boldsymbol{\nu}$ is the unit outer normal of ∂K in cell K. Here, $\mathbf{w}^{int}$ and $\mathbf{w}^{ext}$ are the values of $\mathbf{w}$ on the edge ∂K obtained from the interior and the exterior of K, respectively, and $\mathbf{H}(\mathbf{w}^{int}, \mathbf{w}^{ext}, \boldsymbol{\nu})$ is the numerical flux. In this paper, we consider Lax–Friedrichs flux and

$$\mathbf{H}(\mathbf{w}_1, \mathbf{w}_2, \boldsymbol{\nu}) = \frac{1}{2} \left[\mathbf{F}(\mathbf{w}_1) \cdot \boldsymbol{\nu} + \mathbf{F}(\mathbf{w}_2) \cdot \boldsymbol{\nu} - \alpha(\mathbf{w}_2 - \mathbf{w}_1) \right], \quad \alpha = \| |\langle u, v \rangle| + c \|_\infty, \tag{2.3}$$

where $c = \sqrt{\frac{\gamma p}{\rho}}$ is the sound speed.

3. Bound-preserving technique for the convection term. In this section, we take the source to be zero, i.e., $\mathbf{s}(\mathbf{w}) = \mathbf{0}$ in (2.1) or $s_1 = \cdots = s_{M-1} = 0$ in (1.1), and construct the bound-preserving technique.

3.1. Preliminaries. In this subsection, we introduce some preliminaries that to be used for the bound-preserving technique.

We first demonstrate the PP technique introduced in [18]. We use Euler forward time discretization and take $\boldsymbol{\xi} = \mathbf{1}$; then the equation satisfied by the numerical cell averages can be written as

$$\bar{\mathbf{w}}_K^{n+1} = \bar{\mathbf{w}}_K^n - \frac{\Delta t}{|K|} \int_{\partial K} \mathbf{H}(\mathbf{w}^{int}, \mathbf{w}^{ext}, \boldsymbol{\nu}) ds, \tag{3.1}$$

where $\bar{\mathbf{w}}_K^n = \frac{1}{|K|} \int_K \mathbf{w} d\mathbf{x}$ is the cell average of the numerical solution $\mathbf{w}$ in cell K at time level n and Δt is the time step size. In [18], the authors defined the convex admissible set as

$$G = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \\ E \\ r_1 \\ \cdots \\ r_{M-1} \end{pmatrix}, \rho > 0, p > 0, z_1 > 0, \ldots, z_{M-1} > 0 \right\}$$

and constructed the numerical approximations that lie in G. In this paper, we would like to define another admissible set

$$\tilde{G} = \left\{ \mathbf{w} \in G, \sum_{i=1}^{M-1} z_i < 1 \right\}.$$

The only difference between G and $\tilde{G}$ is one more condition that $z_1 + \cdots + z_{M-1} < 1$ is added in $\tilde{G}$. If we introduce a new variable $z_M = 1 - z_1 - \cdots - z_{M-1}$, then $\tilde{G}$ can be rewritten as

$$\tilde{G} = \left\{ \mathbf{w} \in G : z_M > 0, \sum_{i=1}^{M} z_i = 1 \right\}.$$

In the rest part of this paper, we will use the form of $\tilde{G}$ given above as the admissible set. Following [18], it is easy to check that $\tilde{G}$ is a convex set as p is a concave function of $\mathbf{w}$. Before we finish this subsection, we would like to demonstrate the following lemma whose proof is straightforward.

LEMMA 3.1. *Suppose* $\mathbf{w} \in \tilde{G}$*; then for any* $\tau > 0$*, we have* $\tau \mathbf{w} \in \tilde{G}$.

3.2. Rectangular meshes. Denote $\Omega = [a, b] \times [c, d]$ to be the computational domain. Let $a = x_{\frac{1}{2}} < \cdots < x_{N_x+\frac{1}{2}} = b$ and $c = y_{\frac{1}{2}} < \cdots < y_{N_y+\frac{1}{2}} = d$ be the grid points in x and y directions, respectively. Define $I_i = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ and $J_j = (y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}})$. Let $K_{ij} = I_i \times J_j$, $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, be a partition of Ω and denote $\Omega_h = \{K_{ij}\}$. For simplicity, if not otherwise stated, we always use K to denote the cell. The mesh sizes in the x and y directions are given as $\Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ and $\Delta y_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$, respectively. For simplicity, we assume uniform meshes and denote $\Delta x = \Delta x_i$ and $\Delta y = \Delta y_j$. However, this assumption is not essential.

For accuracy, we use L-point Gaussian quadratures with $L \geq k+1$ to approximate the line integrals in (3.1). More details of this requirement can be found in [3]. The Gaussian quadrature points on $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and $[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ are denoted by

$$p_i^x = \left\{ x_i^\beta : \beta = 1, \ldots, L \right\} \text{ and } p_j^y = \left\{ y_j^\beta : \beta = 1, \ldots, L \right\},$$

respectively. Also, we denote w_β as the corresponding weights on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Moreover, we use

$$\hat{p}_i^x = \left\{\hat{x}_i^\alpha : \alpha = 0, \dots, \hat{L}\right\} \text{ and } \hat{p}_j^y = \left\{\hat{y}_j^\alpha : \alpha = 0, \dots, \hat{L}\right\}$$

as the Gauss–Lobatto points on $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and $[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, respectively, with $2\hat{L} - 1 \geq k$. Also, we denote $\hat{w}_\alpha$ as the corresponding weights on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Then we can state the following theorem [18].

THEOREM 3.1. *If the DG solution $\mathbf{w}_K(x, y) \in G$ for all $(x, y) \in S$, where*

$$S = (p_i^x \otimes \hat{p}_j^y) \cup (\hat{p}_i^x \otimes p_j^y) \cup (p_i^x \otimes p_j^y), \tag{3.2}$$

then the scheme (3.1) *is PP, namely, $\bar{\mathbf{w}}_K^{n+1} \in G$ under the time step restriction*

$$\alpha(\lambda_1 + \lambda_2) \leq \hat{\omega}_1, \tag{3.3}$$

where $\lambda_1 = \frac{\Delta t}{\Delta x}$ and $\lambda_2 = \frac{\Delta t}{\Delta y}$.

Proof. In Theorem 2.1 in [18], the authors considered system (2.1) with $M = 2$. The time step restriction for the PP technique is

$$\alpha(\lambda_1 + \lambda_2) \leq a_1 \hat{\omega}_1, \quad \max_{p_i^x \otimes p_j^y} \{\Delta t \ s_1/\rho\} \leq a_2,$$

where a_1 and a_2 are two arbitrary nonnegative numbers satisfying $a_1 + a_2 = 1$. In this section, we take the source term $s_1 = 0$. Hence, we can take $a_2 = 0$, $a_1 = 1$ to obtain (3.3). For general M, since the equation satisfied by r_i, $i = 1, \dots, M-1$ are exactly the same, the time step restriction for the PP technique for r_1 also works for r_i, $i = 2, \dots, M-1$. Finally, to preserve the positivity of pressure p, we would like to define $\tilde{E} = E - \sum_{i=1}^{M-2} r_i q_i$; then it is easy to check that the scheme satisfied by $\tilde{E}$ is exactly the same as that by E and $p = (\gamma-1)(\tilde{E} - \frac{1}{2}\rho(u^2+v^2) - \rho z_{M-1} q_{M-1} - \rho z_M q_M)$. Hence we can use $\tilde{E}$ as the total energy in the proof of Theorem 2.1 in [18] and follow the same derivations. Now we finish the proof. ∎

3.3. Triangular meshes. For each triangle K we denote by ℓ_K^i $(i = 1, 2, 3)$ as the length of its three edges e_K^i $(i = 1, 2, 3)$. Assume the line integrals in (3.1) are solved by L-point Gaussian quadrature where $L \geq k + 1$. Different from the quadrature applied in the previous subsection, we consult the quadrature introduced in [24], where the quadrature points are given in the polycentric coordinates as

$$\begin{aligned} S = \Bigg\{ & \left(\frac{1}{2} + z^\beta, \left(\frac{1}{2} + \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right), \left(\frac{1}{2} - \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right)\right), \\ & \left(\left(\frac{1}{2} - \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right), \frac{1}{2} + z^\beta, \left(\frac{1}{2} + \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right)\right), \\ & \left(\left(\frac{1}{2} + \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right), \left(\frac{1}{2} - \hat{z}^\alpha\right)\left(\frac{1}{2} - z^\beta\right), \frac{1}{2} + z^\beta\right), \\ & \alpha = 0, \dots, \hat{L}, \ \beta = 1, \dots, L \Bigg\}, \end{aligned} \tag{3.4}$$

where $\hat{z}^\alpha$ $(\alpha = 0, \dots, \hat{L})$ and z^β $(\beta = 1, \dots, L)$ are the Gauss–Lobatto and Gaussian quadrature points on the reference interval $[-\frac{1}{2}, \frac{1}{2}]$, respectively. Now we can state the following theorem [18] and whose proof is basically the same as given in Theorem 3.1, and we skip it here.

THEOREM 3.2. *If the numerical solution* $\mathbf{w}_K(x,y) \in G$ *for any* $(x,y) \in S$, *then the scheme* (3.1) *is PP, i.e.,* $\bar{\mathbf{w}}_K^{n+1} \in G$ *under the time step restriction*

$$\alpha \frac{\Delta t}{|K|} \sum_{i=1}^{3} \ell_K^i \le \frac{2}{3}\hat{\omega}_1. \tag{3.5}$$

3.4. The upper bound of the mass fraction. In [18], the authors did not demonstrate how to preserve the upper bound 1 of the mass fraction, and we will demonstrate the technique in this subsection.

Instead of analyzing (1.1), we would like to study (1.1b)–(1.1f) and (1.2), which is equivalent to (1.1). Following the same discussion in this section, we will not consider the contribution from the source and take $s_1 = \cdots = s_M = 0$. Theorems 3.1 and 3.2 have the following corollary directly.

THEOREM 3.3. *Suppose the conditions of Theorems* 3.1 *and* 3.2 *are satisfied for rectangular and triangular meshes, respectively. If* $\mathbf{w}_K \in \tilde{G}$ *and* $\sum_{i=1}^{M} z_i = 1$ *for all* $(x,y) \in S$, *then* $\bar{\mathbf{w}}_{\mathbf{K}}^{\mathbf{n+1}} \in \tilde{G}$.

Proof. We only need to show that $\bar{r}_M^{n+1} = \bar{\rho}^{n+1} - \sum_{i=1}^{M-1} \bar{r}_i^{n+1} > 0$. Denote

$$\mathbf{H} = (h_\rho, h_m, h_n, h_E, h_1, \ldots, h_{M-1})^T$$

in (2.3). Then we have

$$\begin{aligned} h_\rho &= \frac{1}{2}\left[\mathbf{F}_\rho(\mathbf{w}_1)\cdot\boldsymbol{\nu} + \mathbf{F}_\rho(\mathbf{w}_2)\cdot\boldsymbol{\nu} - \alpha(\rho_2 - \rho_1)\right], \\ h_i &= \frac{1}{2}\left[\mathbf{F}_i(\mathbf{w}_1)\cdot\boldsymbol{\nu} + \mathbf{F}_i(\mathbf{w}_2)\cdot\boldsymbol{\nu} - \alpha(r_{i2} - r_{i1})\right], \quad i = 1,2,\ldots,M-1, \end{aligned}$$

where

$$\mathbf{F}_\rho = (m,n) \quad \text{and} \quad \mathbf{F}_i = (mz_i, nz_i).$$

Using the fact that $\sum_{i=1}^{M} z_i = 1$, we can subtract the DG scheme for (1.1e)–(1.1f) from that for (1.1a) to obtain the DG scheme for (1.2), and the equation satisfied by the numerical cell average $\bar{r}_M$ is

$$\bar{r}_M^{n+1} = \bar{r}_M^n - \frac{\Delta t}{|K|}\int_{\partial K} h_M(\mathbf{w}^{int}, \mathbf{w}^{ext}, \boldsymbol{\nu})ds,$$

where the numerical flux is given as

$$h_M(\mathbf{w}_1, \mathbf{w}_2, \boldsymbol{\nu}) = \frac{1}{2}\left[\mathbf{F}_M(\mathbf{w}_1)\cdot\boldsymbol{\nu} + \mathbf{F}_M(\mathbf{w}_2)\cdot\boldsymbol{\nu} - \alpha(r_{M2} - r_{M1})\right]$$

with

$$\mathbf{F}_M = (mz_M, nz_M).$$

We can observe that h_M is similar to h_i, and the only difference is that we replace i by M. Therefore, the equation satisfied by $\bar{r}_i$, $i = 1,\ldots,M$, are exactly the same. In Theorems 3.1 and 3.2, we have constructed the time step restrictions to obtain positive $\bar{r}_i^{n+1}$, $i = 1,\ldots,M-1$. Therefore, the same time step constraints also work for $\bar{r}_M^{n+1}$. Moreover, since $\sum_{i=1}^{M} \bar{r}_i^{n+1} = \bar{\rho}^{n+1}$, then $\bar{\mathbf{w}}^{n+1} \in \tilde{G}$. ∎

Remark 3.1. In the above proof, we have used the condition that $r_1+\cdots+r_M=\rho$ ($z_1+\cdots+z_M=1$) at all time levels and thus obtain an extra ghost equation satisfied by r_M, which is the same as the equations satisfied by r_i, $i=1,\ldots,M-1$. In other words, if we solve r_M by using the same ghost equation, the condition $r_1+\cdots+r_M=\rho$ should be true at the next time level. This is crucial to obtain the positivity of r_M and the upper bound of z_i, $i=1,\ldots,M$ at the next time level. Hence, the new high-order time integration which will be constructed in the next section should maintain the total mass conservation condition $r_1+\cdots+r_M=\rho$ at the next time level.

Before we finish this section, we would like to demonstrate the following definition.

DEFINITION 3.1. *We say the elements in the numerical flux* $\mathbf{H}$ *are consistent if* $h_\rho=h_i$ *if we take* $z_i=1$ *for all* $1\le i\le M-1$.

The elements in the numerical flux $\mathbf{H}$ in (2.3) are consistent, and we have used this fact to construct h_M and preserve the upper bound 1 of the mass fraction.

4. Bound-preserving technique for the full algorithm. In this section, we proceed to demonstrate the bound-preserving technique for the full algorithm. We first construct the high-order time integration and then demonstrate the full algorithm and the bound-preserving technique.

4.1. High-order time discretization. Consider the ODE

$$\mathbf{w}_t=\mathbf{F}(\mathbf{w})+\mathbf{s}(\mathbf{w}), \tag{4.1}$$

where $\mathbf{F}(\mathbf{w})$ represents the DG discretization of the convection term in this section.

We rewrite (4.1) as

$$\mathbf{w}_t+\mu\mathbf{w}=\mathbf{F}(\mathbf{w})+\mathbf{s}(\mathbf{w})+\mu\mathbf{w},$$

where $\mu\ge 0$ is a constant in each time step but may depend on n. The above equation further yields

$$(e^{\mu t}\mathbf{w})_t=e^{\mu t}(\mathbf{F}(\mathbf{w})+\mathbf{s}(\mathbf{w})+\mu\mathbf{w}).$$

We use the SSP multistep methods to discretize the above ODE to obtain the exponential multistep methods. In [6], the authors introduced second-, third-, and fourth-order schemes. For simplicity, we only discuss the second- and third-order schemes in this paper. The extension to fourth-order schemes is straightforward following the same lines. The second- and third-order schemes given in [6] are

$$\mathbf{w}^{n+1}=\frac{3}{4}e^{-\mu\Delta t}\left[\mathbf{w}^n+2\Delta t\mathbf{F}(\mathbf{w}^n)+2\Delta t(\mathbf{s}(\mathbf{w}^n)+\mu\mathbf{w}^n)\right]+\frac{1}{4}e^{-3\mu\Delta t}\mathbf{w}^{n-2}, \tag{4.2}$$

and

$$\begin{aligned}\mathbf{w}^{n+1}=&\frac{16}{27}e^{-\mu\Delta t}\left[\mathbf{w}^n+3\Delta t\mathbf{F}(\mathbf{w}^n)+3\Delta t\left(\mathbf{s}(\mathbf{w}^n)+\mu\mathbf{w}^n\right)\right]\\&+\frac{11}{27}e^{-4\mu\Delta t}\left[\mathbf{w}^{n-3}+\frac{12}{11}\Delta t\mathbf{F}(\mathbf{w}^{n-3})+\frac{12}{11}\Delta t\left(\mathbf{s}(\mathbf{w}^{n-3})+\mu\mathbf{w}^{n-3}\right)\right],\end{aligned} \tag{4.3}$$

respectively. However, it is impossible to construct the bound-preserving technique by using the time integration given above. Before we demonstrate the reasons, we would like to give the following definition

DEFINITION 4.1. *Consider a multistep method containing* $\mathbf{w}^{n+i}$, $i=-\ell,1-\ell,\ldots,0,1$ *for* (4.1). *We say the scheme is globally conservative if* $\mathbf{w}^{n+1}=\mathbf{1}$ *under the following two conditions:*

1. $\mathbf{w}^{n-\ell} = \cdots = \mathbf{w}^n = \mathbf{1}$;
2. $\mathbf{F}(\mathbf{w}) = \mathbf{s}(\mathbf{w}) = \mathbf{0}$.

It is easy to check that the globally conservative time integration implies total mass conservation, i.e., $z_1 + \cdots + z_M = 1$ ($r_1 + \cdots + r_M = \rho$) at the next time level. Also, as we will see later in Theorem 4.1, the globally conservative condition ensures us to rewrite the solution at the next time level as a convex combination of several terms and thus we only need to preserve the bounds of each term. However, (4.2) and (4.3) are not globally conservative for $\mu \neq 0$, and the necessary condition, $\sum_{i=1}^{M} z_i = 1$, in the bound-preserving technique may not be satisfied. We will modify the schemes and make them to be globally conservative. To do that, we consider Taylor's expansion of the exponential functions. For (4.2), we approximate

$$e^{-\mu\Delta t} \approx 1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2 \geq 0$$

to obtain

$$\begin{aligned}\mathbf{w}^{n+1} = {} & \frac{3}{4}\left(1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2\right)\left[\mathbf{w}^n + 2\Delta t\mathbf{F}(\mathbf{w}^n) + 2\Delta t(\mathbf{s}(\mathbf{w}^n) + \mu\mathbf{w}^n)\right] \\ & + \frac{1}{4}\left(1 - 3\mu\Delta t + \frac{9}{2}(\mu\Delta t)^2\right)\mathbf{w}^{n-2}. \end{aligned} \tag{4.4}$$

Since we expanded the exponential function to the second-order term (the error is third-order in time), then (4.4) is also second-order accurate. Now, we are ready to construct the globally conservative scheme. We simply take $\mathbf{s} = \mathbf{F} = \mathbf{0}$ and let $\mathbf{w}^{n-2} = \mathbf{w}^n = \mathbf{1}$ to obtain $\mathbf{w}^{n+1} = (1 + \frac{3}{4}(\mu\Delta t)^3)\mathbf{1}$. Therefore, the second-order globally conservative scheme is

$$\mathbf{w}^{n+1} = A_2^1\left[\mathbf{w}^n + 2\Delta t\mathbf{F}(\mathbf{w}^n) + 2\Delta t(\mathbf{s}(\mathbf{w}^n) + \mu\mathbf{w}^n)\right] + A_2^2\mathbf{w}^{n-2}, \tag{4.5}$$

where

$$A_2^1 = \frac{3}{4}\frac{1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2}{1 + \frac{3}{4}(\mu\Delta t)^3} \quad A_2^2 = \frac{1}{4}\frac{1 - 3\mu\Delta t + \frac{9}{2}(\mu\Delta t)^2}{1 + \frac{3}{4}(\mu\Delta t)^3}.$$

It is easy to check A_2^1 and A_2^2 are both positive for all Δt. The scheme is globally conservative since it is easy to check that

$$A_2^1 + 2\mu\Delta t A_2^1 + A_2^2 = A_2^1(1 + 2\mu\Delta t) + A_2^2 = 1. \tag{4.6}$$

Now, we proceed to construct the third-order globally conservative scheme. We also apply the Taylor's expansion of the exponential functions and approximate

$$e^{-\mu\Delta t} \approx 1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2 - \frac{1}{6}(\mu\Delta t)^3 + \frac{1}{24}(\mu\Delta t)^4 \geq 0.$$

Then (4.3) can be written as

$$\begin{aligned}\mathbf{w}^{n+1} = {} & \frac{16}{27}\left(1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2 - \frac{1}{6}(\mu\Delta t)^3 + \frac{1}{24}(\mu\Delta t)^4\right) \\ & \times \left[\mathbf{w}^n + 3\Delta t\mathbf{F}(\mathbf{w}^n) + 3\Delta t\left(\mathbf{s}(\mathbf{w}^n) + \mu\mathbf{w}^n\right)\right] \\ & + \frac{11}{27}\left(1 - 4\mu\Delta t + 8(\mu\Delta t)^2 - \frac{32}{3}(\mu\Delta t)^3 + \frac{32}{3}(\mu\Delta t)^4\right) \\ & \cdot \left[\mathbf{w}^{n-3} + \frac{12}{11}\Delta t\mathbf{F}(\mathbf{w}^{n-3}) + \frac{12}{11}\Delta t\left(\mathbf{s}(\mathbf{w}^{n-3}) + \mu\mathbf{w}^{n-3}\right)\right]. \end{aligned} \tag{4.7}$$

We take $\mathbf{s} = \mathbf{F} = \mathbf{0}$ and let $\mathbf{w}^{n-3} = \mathbf{w}^n = \mathbf{1}$ to obtain $\mathbf{w}^{n+1} = (1 - \frac{2}{3}(\mu\Delta t)^4 + \frac{130}{27}(\mu\Delta t)^5)\mathbf{1}$. Therefore, the third-order globally conservative scheme is

$$\begin{aligned}\mathbf{w}^{n+1} &= A_3^1\left[\mathbf{w}^n + 3\Delta t\mathbf{F}(\mathbf{w}^n) + 3\Delta t\left(\mathbf{s}(\mathbf{w}^n) + \mu\mathbf{w}^n\right)\right] \\ &\quad + A_3^2\left[\mathbf{w}^{n-3} + \frac{12}{11}\Delta t\mathbf{F}(\mathbf{w}^{n-3}) + \frac{12}{11}\Delta t\left(\mathbf{s}(\mathbf{w}^{n-3}) + \mu\mathbf{w}^{n-3}\right)\right],\end{aligned} \tag{4.8}$$

where

$$A_3^1 = \frac{16}{27}\frac{1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2 - \frac{1}{6}(\mu\Delta t)^3 + \frac{1}{24}(\mu\Delta t)^4}{1 - \frac{2}{3}(\mu\Delta t)^4 + \frac{130}{27}(\mu\Delta t)^5},$$
$$A_3^2 = \frac{11}{27}\frac{1 - 4\mu\Delta t + 8(\mu\Delta t)^2 - \frac{32}{3}(\mu\Delta t)^3 + \frac{32}{3}(\mu\Delta t)^4}{1 - \frac{2}{3}(\mu\Delta t)^4 + \frac{130}{27}(\mu\Delta t)^5}.$$

It is easy to check A_3^1 and A_3^2 are both positive for all Δt. The scheme is globally conservative and

$$A_3^1 + 3\mu\Delta t A_3^1 + A_3^2 = 1.$$

Remark 4.1. To construct the third-order scheme, we cannot expand the exponential function to the third-order term, i.e.,

$$e^{-\mu\Delta t} \approx 1 - \mu\Delta t + \frac{1}{2}(\mu\Delta t)^2 - \frac{1}{6}(\mu\Delta t)^3,$$

since the approximation above may not be positive and the PP technique will fail to work. Especially, if μ is large, the time step Δt would be extremely small to make the approximation to be positive. More details about this requirement will be discussed in the next subsection.

4.2. Full algorithm. In this subsection, we will demonstrate the bound-preserving technique for the full scheme and the construction of the physically relevant numerical approximations. We first present the following lemma.

LEMMA 4.1. *Let* $\mathbf{w} \in \tilde{G}$*; then*

$$\tilde{\mathbf{s}} = \mathbf{s}(\mathbf{w}) + \mu\mathbf{w} \in \tilde{G},$$

provided

$$\mu > \max_{0\le i\le M}\left\{-\frac{s_i}{r_i}, \frac{\sum_{j=1}^M s_j q_j}{p}, 0\right\}, \tag{4.9}$$

where $s_M = -\sum_{i=1}^{M-1} s_i$, $r_M = \rho - \sum_{i=1}^{M-1} r_i$, *and* p *is the pressure computed by using* $\mathbf{w}$.

Proof. Denote $\tilde{\mathbf{s}} = (\tilde{s}_\rho, \tilde{s}_m, \tilde{s}_n, \tilde{s}_E, \tilde{s}_1, \ldots, \tilde{s}_{M-1})^T$. It is easy to check $\tilde{s}_o = \mu o$ for $o = \rho, m, n, E$ and $\tilde{s}_i = \mu r_i + s_i$ for $i = 1, \ldots, M-1$. We further denote $\tilde{s}_M = \tilde{s}_\rho - \sum_{i=1}^{M-1}\tilde{s}_i$ and hence

$$\tilde{s}_M = \mu\rho - \sum_{i=1}^{M-1}(\mu r_i + s_i) = \mu\left(\rho - \sum_{i=1}^{M-1} r_i\right) - \sum_{i=1}^{M-1} s_i = \mu r_M + s_M.$$

Therefore, $\tilde{s}_\rho > 0$, and

$$\begin{aligned}\tilde{p} &= (\gamma-1)\left(\tilde{s}_E - \frac{1}{2}\tilde{s}_\rho(u^2+v^2) - \tilde{s}_1 q_1 - \cdots - \tilde{s}_M q_M\right)\\ &= (\gamma-1)\left(\mu E - \frac{1}{2}\mu\rho(u^2+v^2) - (\mu r_1 + s_1)q_1 - \cdots - (\mu r_M + s_M)q_M\right)\\ &= (\gamma-1)\left(\mu\left(E - \frac{1}{2}\rho(u^2+v^2) - r_1 q_1 - \cdots - r_M q_M\right) - s_1 q_1 - \cdots - s_M q_M\right)\\ &= (\gamma-1)\left(\mu p - s_1 q_1 - \cdots - s_M q_M\right) > 0,\end{aligned}$$

with $(u,v) = (\frac{\tilde{s}_m}{\tilde{s}_\rho}, \frac{\tilde{s}_n}{\tilde{s}_\rho}) = (\frac{m}{\rho}, \frac{n}{\rho})$. Moreover, we choose μ as (4.9) to obtain $\tilde{s}_i = s_i + \mu r_i > 0$ for all $1 \le i \le M$. Notice that

$$\sum_{i=1}^{M} \tilde{s}_i = \tilde{s}_\rho;$$

then

$$\sum_{i=1}^{M} \tilde{z}_i = \sum_{i=1}^{M} \frac{\tilde{s}_i}{\tilde{s}_\rho} = 1,$$

and we finish the proof. ∎

Now, we can state the main theorem for second-order scheme. For simplicity, we omit the subscript K of each $\mathbf{w}_K$.

THEOREM 4.1. *Consider the DG scheme* (2.2) *with time integration* (4.5)*, where* μ *satisfies* (4.9)*. If* $\mathbf{w}^n, \mathbf{w}^{n-2} \in \tilde{G}$ *for all* $(x,y) \in S$*, where* S *is defined in* (3.2) *and* (3.4) *for rectangular and triangular meshes, respectively. Then we have* $\bar{\mathbf{w}}^{n+1} \in \tilde{G}$ *under the condition* $\Delta t \le \frac{1}{2}\Delta\tilde{t}$*, where* $\Delta\tilde{t}$ *is the time step given in* (3.3) *and* (3.5) *for rectangular and triangular meshes, respectively.*

Proof. By Theorem 3.3, we have

$$R_1 = \overline{\mathbf{w}^n + 2\Delta t\mathbf{F}(\mathbf{w}^n)} \in \tilde{G}.$$

Moreover, by Lemmas 3.1 and 4.1, we can show that

$$R_2 = \overline{\frac{1}{\mu}(\mathbf{s}(\mathbf{w}) + \mu\mathbf{w})} \in \tilde{G}.$$

Finally, denote $R_3 = \overline{\mathbf{w}}^{n-2} \in \tilde{G}$. Notice the fact that $\tilde{G}$ is a convex set; then take cell average in (4.5) to obtain

$$\begin{aligned}\overline{\mathbf{w}}^{n+1} &= A_2^1\overline{\mathbf{w}^n + 2\Delta t\mathbf{F}(\mathbf{w}^n)} + A_2^1\overline{2\Delta t(\mathbf{s}(\mathbf{w}) + \mu\mathbf{w})} + A_2^2\overline{\mathbf{w}}^{n-2}\\ &= A_2^1 R_1 + 2\mu\Delta t A_2^1 R_2 + A_2^2 R_3 \in \tilde{G},\end{aligned}$$

where in the last step, we applied the globally conservative condition (4.6) and the fact that $A_2^1 > 0$ and $A_2^2 > 0$. ∎

Following the same analysis given above with some minor changes, we can obtain the theorem for the third-order scheme. Therefore, we only demonstrate the theorem without proof.

THEOREM 4.2. *Consider the DG scheme* (2.2) *with time integration* (4.8)*, where* μ *satisfies* (4.9)*. If* $\mathbf{w}^n, \mathbf{w}^{n-3} \in \tilde{G}$ *for all* $(x, y) \in S$*, where* S *is defined in* (3.2) *and* (3.4) *for rectangular and triangular meshes, respectively. Then we have* $\bar{\mathbf{w}}^{n+1} \in \tilde{G}$ *under the condition* $\Delta t \le \frac{1}{3}\Delta\tilde{t}$*, where* $\Delta\tilde{t}$ *is the time step given in* (3.3) *and* (3.5) *for rectangular and triangular meshes, respectively.*

Based on the above two theorems, we can construct physically relevant numerical cell averages $\bar{\mathbf{w}}$. However, the numerical approximations $\mathbf{w}$ may be out of the bounds. Hence, we need to apply suitable limiters and construct physically relevant numerical approximations. The full algorithm on each fixed element K is given as follows:

1. Set a small number $\epsilon = 10^{-13}$.
2. If $\bar{\rho} > \epsilon$, then we proceed to the next step. Otherwise, $\mathbf{w}$ is identified as the approximation to vacuum; then we take $\mathbf{w} = \bar{\mathbf{w}}$ and skip the following steps.
3. We modify the density ρ first. Compute
$$\rho_{min} = \min_{(x,y)\in S} \rho(x, y).$$
If $\rho_{min} < 0$, then take
$$\hat{\rho} = \overline{\rho} + \theta\left(\rho - \overline{\rho}\right), \qquad \hat{r}_i = \overline{r}_i + \theta\left(r_i - \overline{r}_i\right), \quad i = 1, \ldots, M-1,$$
with
$$\theta = \frac{\overline{\rho} - \varepsilon}{\overline{\rho} - \rho_{min}}.$$
Here we implicitly modify $\hat{r}_M = \overline{r}_M + \theta\left(r_M - \overline{r}_M\right)$ to keep $\sum_{i=1}^{M} \hat{r}_i = \hat{\rho}$.
4. Modify the mass fraction. For $1 \le i \le M$, define $\hat{S}_i = \{(x, y) \in S : \hat{r}_i(x, y) \le 0\}$. Take
$$\tilde{r}_i = \hat{r}_i + \theta\left(\frac{\overline{r}_i}{\overline{\rho}}\hat{\rho} - \hat{r}_i\right), 1 \le i \le M-1, \tag{4.10}$$
$$\theta = \max_{1\le i\le M} \max_{(x,y)\in\hat{S}_i} \left\{\frac{-\hat{r}_i(x, y)\bar{\rho}}{\overline{r}_i\hat{\rho}(x, y) - \hat{r}_i(x, y)\bar{\rho}}, 0\right\}.$$
5. Modify the pressure. Denote $\tilde{\mathbf{w}} = (\hat{\rho}, m, n, E, \tilde{r}_1, \ldots, \tilde{r}_{M-1})^T$. For each $\mathbf{x} \in S$, if $\tilde{\mathbf{w}}(\mathbf{x}) \in \tilde{G}$, then take $\theta_{\mathbf{x}} = 1$. Otherwise, take
$$\theta_{\mathbf{x}} = \frac{p(\overline{\mathbf{w}})}{p(\overline{\mathbf{w}}) - p(\tilde{\mathbf{w}}(\mathbf{x}))}.$$
Then, we use
$$\mathbf{w}^{new} = \overline{\mathbf{w}} + \theta(\tilde{\mathbf{w}} - \overline{\mathbf{w}}), \quad \theta = \min_{\mathbf{x}\in S} \theta_{\mathbf{x}},$$
as the new DG approximation. The proof for $p(\mathbf{w}^{new}) \ge 0$ can be found in [18].

Remark 4.2. In step 3, we can simply take $\hat{r}_i = r_i$, $i = 1, \ldots, M-1$ and implicitly modify $\hat{r}_M = \hat{\rho} - \sum_{i=1}^{M-1} r_i$. Therefore, one may not need to apply the limiter to r_i, $i = 1, \ldots, M-1$. In the numerical experiments, we only compute $\hat{\rho}$ and keep r_i unchanged for $1 \le i \le M-1$.

Remark 4.3. In step 4, it is easy to check that $\tilde{r}_i(x, y) \ge 0$ for all $(x, y) \in S$, $i = 1 \ldots, M-1$. If we further define
$$\tilde{r}_M = \hat{r}_M + \theta\left(\frac{\overline{r}_M}{\overline{\rho}}\hat{\rho} - \hat{r}_M\right),$$

then $\tilde{r}_M(x,y) \geq 0$ for all $(x,y) \in S$. Since $\sum_{i=1}^M \hat{r}_i = \hat{\rho}$, then

$$\sum_{i=1}^M \tilde{r}_i = \sum_{i=1}^M \hat{r}_i + \theta \left(\sum_{i=1}^M \frac{\bar{r}_i}{\bar{\rho}} \hat{\rho} - \sum_{i=1}^M \hat{r}_i \right) = \hat{\rho} + \theta \left(\frac{\bar{\rho}}{\bar{\rho}} \hat{\rho} - \hat{\rho} \right) = \hat{\rho}.$$

Therefore, we have $0 \leq \tilde{r}_i(x,y) \leq \hat{\rho}(x,y), \forall (x,y) \in S, i = 1, \ldots, M$, which further indicates that the mass fraction of each species is within the range [0,1].

5. Second-order globally conservative RK method. In this section, we proceed to construct a second-order globally conservative RK method, and the third-order one will be discussed in the future. For the RK method, time steps can change in different time levels. Hence, for practical problems in which the wave speed changes quickly or even widely, the RK method can be an alternative to the multistep method.

Following the analysis in section 4.2, (4.1) yields

$$(e^{\mu t}\mathbf{w})_t = e^{\mu t}(\mathbf{F}(\mathbf{w}) + \mathbf{s}(\mathbf{w}) + \mu \mathbf{w}).$$

The second-order RK scheme given in [6] is

$$\mathbf{w}^{(1)} = e^{-\mu \Delta t} \left(\mathbf{w}^n + \Delta t \mathbf{F}(\mathbf{w}^n) + \Delta t (\mathbf{s}(\mathbf{w}^n) + \mu \mathbf{w}^n) \right), \tag{5.1}$$

$$\mathbf{w}^{n+1} = \frac{1}{2} e^{-\mu \Delta t} \mathbf{w}^n + \frac{1}{2} \left[\mathbf{w}^{(1)} + \Delta t \mathbf{F}(\mathbf{w}^{(1)}) + \Delta t (\mathbf{s}(\mathbf{w}^{(1)}) + \mu \mathbf{w}^{(1)}) \right]. \tag{5.2}$$

Similar to the multistep method, the above scheme is not globally conservative for $\mu \neq 0$. We consider Taylor's expansion of the exponential functions and approximate

$$e^{-\mu \Delta t} \approx 1 - \mu \Delta t + \frac{1}{2} (\mu \Delta t)^2 \geq 0.$$

Following the same idea in section 4.2, we take $\mathbf{s} = \mathbf{F} = \mathbf{0}$ and let $\mathbf{w}^n = \mathbf{1}$ in (5.1) to obtain $\mathbf{w}^{(1)} = (1 - \frac{1}{2}(\mu \Delta t)^2 + \frac{1}{2}(\mu \Delta t)^3)\mathbf{1}$, then take $\mathbf{w}^n = \mathbf{w}^{(1)} = \mathbf{1}$ in (5.2) to obtain $\mathbf{w}^{n+1} = (1 + \frac{1}{4}(\mu \Delta t)^2)\mathbf{1}$. Therefore, the second-order globally conservative scheme is

$$\mathbf{w}^{(1)} = B_1^1 \left[\mathbf{w}^n + \Delta t \mathbf{F}(\mathbf{w}^n) + \Delta t (\mathbf{s}(\mathbf{w}^n) + \mu \mathbf{w}^n) \right], \tag{5.3}$$

$$\mathbf{w}^{n+1} = B_2^1 \mathbf{w}^n + B_2^2 \left[\mathbf{w}^{(1)} + \Delta t \mathbf{F}(\mathbf{w}^{(1)}) + \Delta t (\mathbf{s}(\mathbf{w}^{(1)}) + \mu \mathbf{w}^{(1)}) \right], \tag{5.4}$$

where

$$B_1^1 = \frac{1 - \mu \Delta t + \frac{1}{2}(\mu \Delta t)^2}{1 - \frac{1}{2}(\mu \Delta t)^2 + \frac{1}{2}(\mu \Delta t)^3}, \quad B_2^1 = \frac{1}{2} \frac{1 - \mu \Delta t + \frac{1}{2}(\mu \Delta t)^2}{1 + \frac{1}{4}(\mu \Delta t)^2}, \quad B_2^2 = \frac{1}{2} \frac{1}{1 + \frac{1}{4}(\mu \Delta t)^2}.$$

It is easy to check B_1^1, B_2^1, and B_2^2 are positive for all Δt, and the scheme is globally conservative:

$$(1 + \mu \Delta t) B_1^1 = B_2^1 + (1 + \mu \Delta t) B_2^2 = 1. \tag{5.5}$$

Then following the analyses in Theorem 4.1, we can easily obtain the following one.

THEOREM 5.1. *Consider the DG scheme* (2.2) *with time integration* (5.3) *and* (5.4)*, where μ satisfies* (4.9) *for $\mathbf{w} = \mathbf{w}^n$. If $\mathbf{w}^n \in \tilde{G}$ for all $(x,y) \in S$, where S is defined in* (3.2) *and* (3.4) *for rectangular and triangular meshes, respectively. Then we have $\bar{\mathbf{w}}^{(1)} \in \tilde{G}$ under the condition $\Delta t \leq \Delta \tilde{t}$, where $\Delta \tilde{t}$ is the time step given in* (3.3) *and* (3.5) *for rectangular and triangular meshes, respectively. In addition to the above conditions, if μ satisfies* (4.9) *for $\mathbf{w} = \mathbf{w}^{(1)}$ and $\mathbf{w}^{(1)} \in \tilde{G}$ for all $(x,y) \in S$, then we have $\bar{\mathbf{w}}^{n+1} \in \tilde{G}$ under the condition $\Delta t \leq \Delta \tilde{t}$.*

Remark 5.1. In the above theorem, μ satisfies (4.9) for $\mathbf{w} = \mathbf{w}^{(1)}$, hence it is not easy to find out μ in (5.3). In practice, we choose two different μ in (5.3) and (5.4), say μ^n and $\mu^{(1)}$, satisfying (4.9) for $\mathbf{w} = \mathbf{w}^n$ and $\mathbf{w} = \mathbf{w}^{(1)}$, respectively. Numerical experiments in section 6 demonstrate that the scheme is also second-order accurate in time.

Based on the above theorem, we can construct physically relevant numerical cell averages $\bar{\mathbf{w}}^{(1)}$ and $\bar{\mathbf{w}}^{n+1}$. However, we still need the bound-preserving limiter discussed in section 4.2 to modify the numerical approximations $\mathbf{w}^{(1)}$ and $\mathbf{w}^{n+1}$.

Finally, different from the multistep method, it is not easy to observe the accuracy for (5.3)–(5.4) as we introduced second-order errors in the denominators in $B's$. Before we state the accuracy of the scheme, we would like to demonstrate the following lemma, whose proof follows from direct computation, hence we omit it here.

LEMMA 5.1.

$$B_2^1 + B_2^2 B_1^1 (1+\mu\Delta t)^2 = \frac{1 - \frac{1}{4}(\mu\Delta t)^2 + \mathcal{O}(\Delta t^3)}{1 - \frac{1}{4}(\mu\Delta t)^2 + \mathcal{O}(\Delta t^3)} = 1 + \mathcal{O}(\Delta t^3),$$

$$B_2^2 \Delta t[(1+\mu\Delta t)B_1^1 + 1] = \Delta t \frac{1 + \mathcal{O}(\Delta t^2)}{1 + \mathcal{O}(\Delta t^2)} = \Delta t + \mathcal{O}(\Delta t^3),$$

$$B_2^2 = \frac{1}{2} + \mathcal{O}(\Delta t^2).$$

We will prove that the scheme is indeed second-order accurate in the following theorem.

THEOREM 5.2. *Consider the ordinary differential system* $\mathbf{w}_t = \mathbf{L}(\mathbf{w})$*, with* $\mathbf{L}(\mathbf{u}) = \mathbf{F}(\mathbf{u}) + \mathbf{s}(\mathbf{u})$*. The globally conservative RK scheme* (5.3)–(5.4) *is second-order accurate.*

Proof. We can rewrite (5.3)–(5.4) as

$$\mathbf{w}^{(1)} = B_1^1 \left[\mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \mu\Delta t \mathbf{w}^n\right], \tag{5.6}$$

$$\mathbf{w}^{n+1} = B_2^1 \mathbf{w}^n + B_2^2 \left[\mathbf{w}^{(1)} + \Delta t \mathbf{L}(\mathbf{w}^{(1)}) + \mu\Delta t \mathbf{w}^{(1)}\right]. \tag{5.7}$$

By using (5.5), we have

$$\mathbf{w}^{(1)} = \mathbf{w}^n + (1 + \mathcal{O}(\Delta t))\Delta t \mathbf{L}(\mathbf{w}^n) = \mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \mathcal{O}(\Delta t^2). \tag{5.8}$$

Substitute (5.6) into (5.7) to obtain

$$\begin{aligned}
\mathbf{w}^{n+1} &= B_2^1 \mathbf{w}^n + B_2^2 (1+\mu\Delta t)\mathbf{w}^{(1)} + B_2^2 \Delta t \mathbf{L}(\mathbf{w}^{(1)}) \\
&= [B_2^1 + B_2^2 B_1^1 (1+\mu\Delta t)^2]\mathbf{w}^n + B_2^2 \Delta t[(1+\mu\Delta t)B_1^1 + 1]\mathbf{L}(\mathbf{w}^n) \\
&\quad + B_2^2 \Delta t(\mathbf{L}(\mathbf{w}^{(1)}) - \mathbf{L}(\mathbf{w}^n)) \\
&= \mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \frac{1}{2}\Delta t(\mathbf{L}(\mathbf{w}^{(1)}) - \mathbf{L}(\mathbf{w}^n)) + \mathcal{O}(\Delta t^3) \\
&= \mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \frac{1}{2}\Delta t(\mathbf{L}(\mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \mathcal{O}(\Delta t^2)) - \mathbf{L}(\mathbf{w}^n)) + \mathcal{O}(\Delta t^3) \\
&= \mathbf{w}^n + \Delta t \mathbf{L}(\mathbf{w}^n) + \frac{1}{2}\Delta t^2 \nabla \mathbf{L}(\mathbf{w}^n)\mathbf{L}(\mathbf{w}^n) + \mathcal{O}(\Delta t^3),
\end{aligned}$$

where in the third step we use Lemma 5.1 and (5.5), and the fourth step requires (5.8). Substitute $\mathbf{w}^n$ and $\mathbf{w}^{n+1}$ with the exact solutions, we obtain that the local truncation error is $\mathcal{O}(\Delta t^2)$. ∎

Remark 5.2. The construction of third-order globally conservative RK scheme is highly nontrivial. We will discuss this in the future.

6. Numerical examples. In this section, we will use numerical experiments to demonstrate the effect of the bound-preserving DG method. We refine the meshes to match the positions of the shocks with those given in [19]. Therefore, we only plot the numerical approximations in this section. Moreover, the numerical results obtained by using RK method and multistep method are similar. If not otherwise stated, the figures in this section are obtained by using DG method with piecewise P^1 polynomials and second-order time integration given in (4.5).

6.1. Test of the ODE solver.

Example 6.1 (accuracy test for the ODE solver). We first test the stability and accuracy of the ODE solver, and study the following problem,

$$u'(t) = -cu^7, \qquad u(0) = u_0,$$

where c is a parameter that we can adjust. The problem becomes stiff as c increases. The exact solution is

$$u(t) = u_0(6ctu_0^6 + 1)^{-1/6}.$$

We take the final time to be $t = 0.5$ and denote the total number of time steps as N_t.

We first take $u_0 = 0.1$ with $c = 10000$. Numerical results for all ODE solvers proposed in this paper are listed in Table 6.1. The initial condition is well prepared, and we can observe optimal convergence rates. Next, we take $u_0 = 1$; the results are given in Table 6.2. For this problem, the initial condition is not well prepared, and we can only observe optimal convergence rate for all time integrations if the problem is not stiff, e.g., $c = 1$. However, if the problem is stiff, e.g., $c = 100$ or 100,000, we can hardly observe the expected accuracy.

TABLE 6.1
Accuracy test for ODE solvers with $u_0 = 0.1$ with $c = 10{,}000$.

N_t	L^∞ norm	Order	L^∞ norm	Order	L^∞ norm	Order
	2nd order RK		2nd order multistep		3rd order multistep	
2	1.30E-08	–	5.24E-08	–	3.20E-09	–
4	3.25E-09	2.00	1.17E-08	2.16	2.13E-10	3.90
8	8.10E-10	2.00	3.62E-09	1.70	3.52E-11	2.60
16	2.02E-10	2.00	1.01E-09	1.84	5.21E-12	2.76
32	5.06E-11	2.00	2.65E-10	1.93	6.97E-13	2.90
64	1.26E-11	2.00	6.80E-11	1.96	9.00E-14	2.95

6.2. One space dimension.

Example 6.2 (accuracy test for one-dimensional (1D) system). In this example, we consider periodic boundary condition and take $u = 1$ and $p = 0$ in the exact solution. We choose $M = 2$ and the source term is given as $s_1 = -cr_1^7$. Hence, we need to solve the following system

$$\begin{cases} \rho_t + \rho_x = 0, \\ (r_1)_t + (r_1)_x = -c(r_1)^7, \end{cases} \quad x \in [0, 2\pi].$$

TABLE 6.2
Accuracy test for ODE solvers with $u_0 = 1$.

	c=1		c=100		c=10,000	
N_t	L^∞ norm	Order	L^∞ norm	Order	L^∞ norm	Order
	2nd order RK					
20	7.07E-05	–	7.96E-04	–	1.93E-03	–
40	1.74E-05	2.02	4.59E-04	0.79	8.92E-04	1.11
80	4.28E-06	2.02	2.33E-04	0.98	4.24E-04	1.07
160	1.07E-06	2.00	8.65E-05	1.43	1.89E-04	1.17
320	2.68E-07	2.00	2.45E-05	1.82	7.13E-05	1.40
	2nd order multistep					
20	3.07E-04	–	2.92E-03	–	2.84E-04	–
40	8.47E-05	1.86	1.25E-03	1.22	6.02E-05	2.24
80	2.22E-05	1.94	4.99E-04	1.33	1.56E-04	−1.37
160	5.71E-06	1.96	1.90E-04	1.39	1.51E-04	0.05
320	1.45E-06	1.97	6.71E-05	1.50	9.35E-05	0.69
	3rd order multistep					
20	6.68E-05	–	3.90E-02	–	7.06E-02	–
40	1.02E-05	2.71	6.96E-03	2.49	5.82E-02	0.28
80	1.41E-06	2.86	3.19E-04	4.45	4.23E-02	0.46
160	1.87E-07	2.91	1.19E-04	1.43	2.62E-02	0.69
320	2.42E-08	2.95	3.76E-05	1.66	1.23E-02	1.10

The initial conditions are given as $r_1(x,0) = 0.1\big(1 + sin(x)\big)$ and $\rho(x,0) = 0.1\big(2 + sin(x) + cos(x)\big)$. The parameter c can be used to adjust the stiffness of the equation. For this problem, the total density ρ should be nonnegative and the mass fraction r_1/ρ should be between 0 and 1.

We apply DG method with piecewise P^1 (P^2) polynomials coupled with second-order multistep and RK (third-order multistep) time discretizations with and without bound-preserving limiter. The final time is taken as $t = 0.5$. Both stiff and nonstiff cases are calculated. The errors are listed in Table 6.3, and the last column shows the percentage of cells that have been modified by the limiter. We can observe optimal orders of accuracy with and without limiter.

Example 6.3 (a 1D detonation wave with 3 species and 1 reaction). In this case, we solve a reacting model with three species and one reaction,

$$2H_2 + O_2 \to 2H_2O.$$

The parameters are taken as $T_1 = 2.0, B_1 = 500, \alpha_1 = 1, q_1 = 1000, q_2 = 0, q_3 = 0$, $M_1 = 2, M_2 = 32, M_3 = 18$. The computational domain is $[0, 50]$, and the initial condition is given as piecewise constants

$$(\rho, u, p, z_1, z_2, z_3)(x,0) = \begin{cases} (2.0, 10.0, 40.0, 0.325, 0.0, 0.625), & x \leqslant 2.5, \\ (1.0, 0.0, 1.0, 0.4, 0.6, 0.0), & x > 2.5. \end{cases}$$

We take the final time to be $t = 3$. This is a simple one-step chemical model for hydrogen-oxygen mixtures. The fuel rich hydrogen-oxygen mixture is on the right-hand side. And the mixture is totally burnt on the left-hand side.

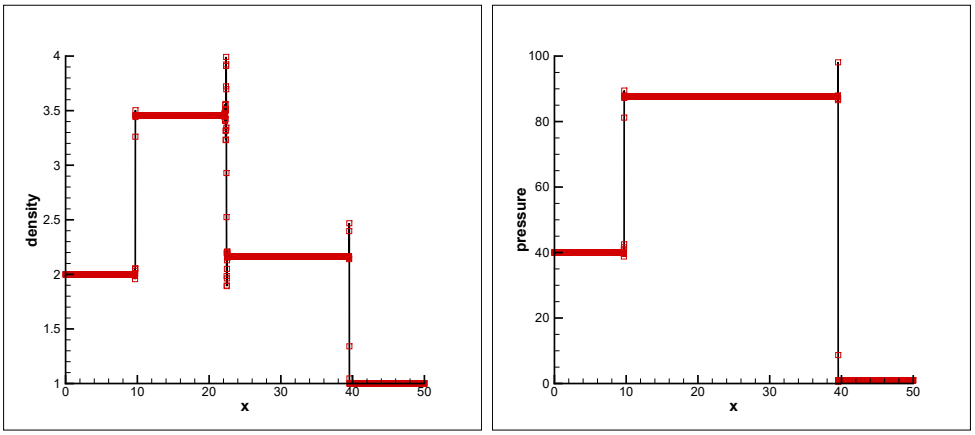
To resolve the thin reaction zone, we take $\Delta x = 0.01$ and CFL $= 0.05$. The profiles of density, pressure, and mass fraction for each species are shown in Figure 6.1. From the figure, we can see that the shocks are captured well and the shock positions are

TABLE 6.3
Accuracy test for the one dimensional problem.

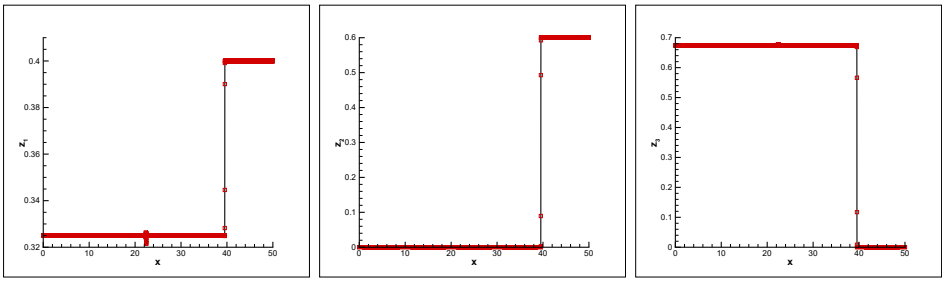
	Without limiter				With limiter				
N	L^∞ norm	Order	L^2 norm	Order	L^∞ norm	Order	L^2 norm	Order	Percentage
	2nd order multistep method, cfl=0.1, c=100								
10	1.19E-02	–	9.86E-03	–	1.33E-02	–	1.03E-02	–	20.00%
20	3.16E-03	1.91	2.55E-03	1.95	3.78E-03	1.81	2.63E-03	1.97	20.00%
40	8.04E-04	1.98	6.41E-04	1.99	9.83E-04	1.94	6.66E-04	1.98	12.50%
80	2.02E-04	1.99	1.60E-04	2.00	2.78E-04	1.82	1.65E-04	2.02	10.00%
160	5.07E-05	2.00	4.01E-05	2.00	7.05E-05	1.98	4.10E-05	2.01	9.38%
	2nd order multistep method, cfl=0.1, c=10,000								
10	1.21E-02	–	9.85E-03	–	1.34E-02	–	1.01E-02	–	20.00%
20	3.20E-03	1.91	2.56E-03	1.94	3.82E-03	1.80	2.58E-03	1.97	20.00%
40	8.13E-04	1.97	6.55E-04	1.97	9.93E-04	1.94	6.58E-04	1.97	12.50%
80	2.05E-04	1.99	1.65E-04	1.98	2.81E-04	1.82	1.63E-04	2.01	10.00%
160	5.15E-05	1.99	4.16E-05	1.99	7.13E-05	1.98	4.05E-05	2.01	9.37%
	2nd order RK method, cfl=0.1, c=100								
10	1.18E-02	–	9.89E-03	–	1.59E-02	–	1.11E-02	–	50.00 %
20	3.15E-03	1.91	2.54E-03	1.96	4.38E-03	1.86	2.77E-03	2.00	40.00 %
40	8.04E-04	1.97	6.40E-04	1.99	1.13E-03	1.95	6.89E-04	2.00	25.00 %
80	2.02E-04	1.99	1.60E-04	2.00	3.17E-04	1.82	1.68E-04	2.03	20.00 %
160	5.07E-05	2.00	4.01E-05	2.00	8.10E-05	1.97	4.15E-05	2.02	18.75 %
	2nd order RK method, cfl=0.1, c=10,000								
10	1.18E-02	–	9.62E-03	–	1.65E-02	–	1.10E-02	–	50.00 %
20	3.15E-03	1.90	2.49E-03	1.95	4.37E-03	1.91	2.69E-03	2.04	40.00 %
40	8.03E-04	1.97	6.31E-04	1.98	1.12E-03	1.95	6.58E-04	2.03	22.50 %
80	2.02E-04	1.99	1.58E-04	1.99	3.17E-04	1.83	1.62E-04	2.02	13.75 %
160	5.07E-05	2.00	3.96E-05	2.00	8.10E-05	1.97	4.04E-05	2.01	10.63 %
	3rd order multistep method, cfl=0.05, c=100								
10	1.78E-04	–	3.08E-04	–	2.20E-04	–	3.35E-04	–	10.00%
20	2.13E-05	3.06	3.77E-05	3.03	2.13E-05	3.37	3.80E-05	3.14	15.00%
40	3.08E-06	2.79	5.38E-06	2.81	3.08E-06	2.79	5.38E-06	2.82	12.50%
80	3.77E-07	3.03	6.61E-07	3.02	3.77E-07	3.03	6.61E-07	3.02	10.00%
160	4.73E-08	2.99	8.31E-08	2.99	4.73E-08	2.99	8.31E-08	2.99	8.75%
	3rd order multistep method, cfl=0.05, c=10,000								
10	3.36E-04	–	4.78E-04	–	3.36E-04	–	4.89E-04	–	10.00%
20	4.35E-05	2.95	5.46E-05	3.13	4.35E-05	2.95	5.47E-05	3.16	15.00%
40	5.86E-06	2.89	6.83E-06	3.00	5.86E-06	2.89	6.83E-06	3.00	12.50%
80	7.00E-07	3.06	8.41E-07	3.02	7.00E-07	3.06	8.41E-07	3.02	10.00%
160	8.63E-08	3.02	1.05E-07	3.00	8.63E-07	3.02	1.05E-07	3.00	8.75%

correct. Moreover, the density and pressure are positive, and all mass fractions are in the interval $[0, 1]$. Since we only implemented the bound-preserving limiter, there are some oscillations in the density and mass fractions.

In addition, we also compared the results obtained using PP technique in [18] with the one achieved by using bound-preserving technique demonstrated in this paper. Figure 6.2 shows the profile of mass fraction of H_2O (the last species that not solved explicitly in the DG scheme) with different techniques at time t=0.055. We can observe some negative values in the left panel, where we used the PP technique in [18]. As expected, the bound-preserving technique given in this paper yields positive value of the mass fraction in the right panel.

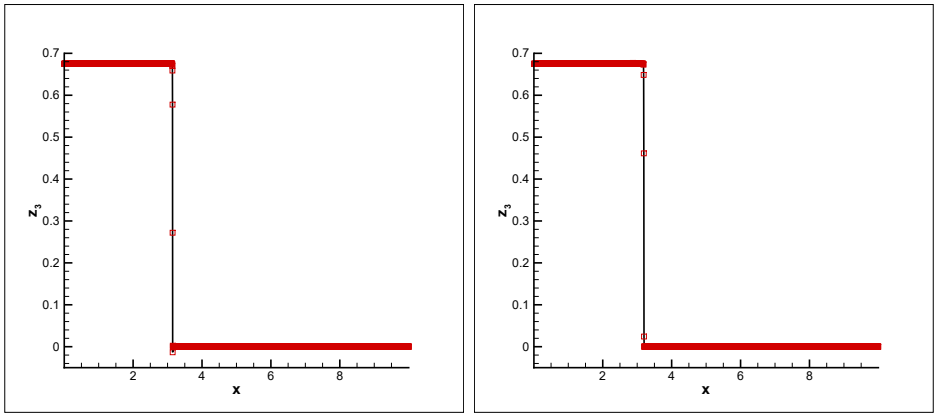


(a) density (b) pressure



(c) mass fraction of z_1 (d) mass fraction of z_2 (e) mass fraction of z_3

FIG. 6.1. *Numerical solutions of Example* 6.3 *at* $t = 3$.



(a) using PP technique in [18] (b) using bound-preserving technique

FIG. 6.2. *Numerical solutions of Example* 6.3 *at* $t = 0.055$.

We also investigate the influence of the denominators of A_2^1 and A_2^2 in the time integral scheme (4.5). We take the denominators to be 1, the numerical scheme is not stable until we reduce the CFL to 0.001. This is because the lack of mass conservation will lead to the "add mass" effect, especially for μ is large. Therefore, a sufficiently small Δt is necessary to suppress this effect.

Moreover, we applied the second-order RK method and the observations are similar.

Example 6.4 (a 1D detonation wave with 5 species and 2 reaction). In this example, we would resolve a two-step chemical reaction model with 4 species for hydrogen-oxygen-nitrogen mixture.

$$H_2 + O_2 \to 2OH,\ 2OH + H_2 \to 2H_2O.$$

Here nitrogen is considered as a catalyst. The parameters are $T_1 = 2.0, T_2 = 10, B_1 = B_2 = 10^6, \alpha_1 = \alpha_2 = 0, q_1 = q_2 = 0, q_3 = -20, q_4 = -100, q_5 = 0, M_1 = 2, M_2 = 32, M_3 = 17, M_4 = 18, M_5 = 28$. The initial data are as follows:

$$(\rho, u, p, z_1, z_2, z_3, z_4, z_5)(x, 0) = \begin{cases} (2.0, 10.0, 40.0, 0.0, 0.0, 0.17, 0.63, 0.2), & x \leqslant 2.5, \\ (1.0, 0.0, 1.0, 0.08, 0.72, 0.0, 0.0, 0.2), & x > 2.5. \end{cases}$$

The computational domain is $[0, 10]$, and final time is $t = 0.5$.

In this example, we take $\Delta x = 0.005$ and CFL $= 0.01$. Figure 6.3 shows the numerical density, pressure, and mass fractions. All shock waves are captured accurately, as well as the mass fraction of the intermediate component OH. Also, Figure 6.3 shows that all bounds are preserved.

6.3. Two space dimensions.

Example 6.5 (accuracy test for two-dimensional (2D) system). From now on, we consider the 2D problem. In this example, we consider periodic boundary condition and take $u = v = 1$ and $p = 0$ in the exact solution. We choose $M = 2$ and the source is given as $s_1 = -cr_1^7$. Hence, we need to solve the following system

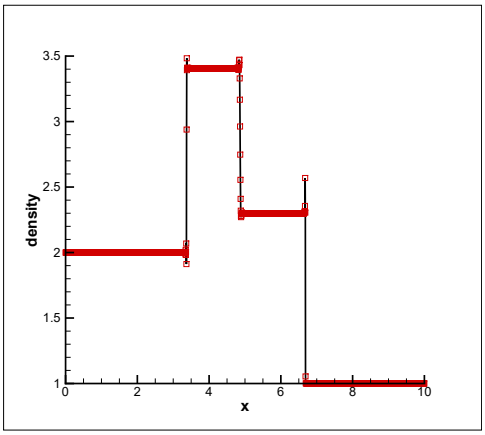
$$\begin{cases} \rho_t + \rho_x + \rho_y = 0, \\ (r_1)_t + (r_1)_x + (r_1)_y = -c(r_1)^7, \end{cases} \quad (x, y) \in [0, 2\pi]^2.$$

The initial conditions are given as $\rho(x, y, 0) = 0.1\big(2 + sin(x + y) + cos(x + y)\big)$ and $r_1(x, y, 0) = 0.1\big(1 + sin(x + y)\big)$, respectively. For this problem, the total density ρ should be nonnegative and the mass fraction r_1/ρ should be between 0 and 1.

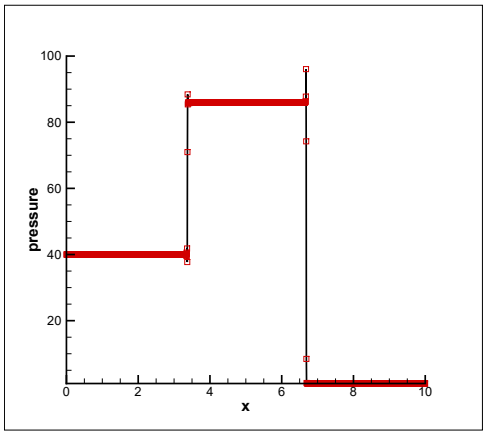
We use piecewise $P^1(P^2)$ polynomials coupled with second-order (third-order) time discretizations and take the final time to be $t = 0.5$. Numerical errors for different time discretizations with different c are given in the left column of Table 6.4. From the left column of Table 6.4 we can again observe the expected high order of accuracy of our scheme. We further add the limiter to preserve the lower bound of ρ and the two bounds of r_1/ρ and show the results in the right part of the error table. The percentage of cells that have been modified by the limiter is listed in the last column. By comparing the results with and without limiter, we can see that the limiter dose not harm the original high order of accuracy.

Example 6.6 (a 2D detonation wave with 4 species and 1 reaction). In this example, we test a 2D reacting model with four species and one reaction. A prototype reaction for this model is

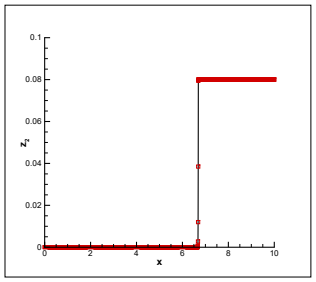
$$CH_4 + 2O_2 \to CO_2 + 2H_2O.$$



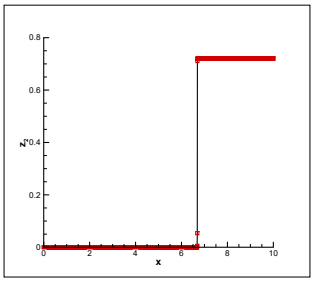
(a) density



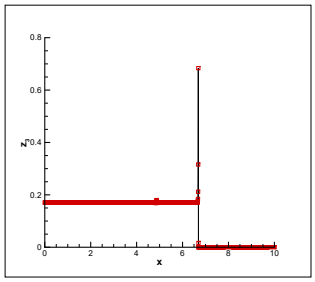
(b) pressure



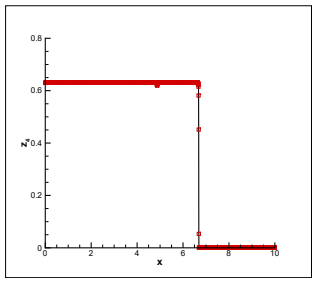
(c) mass fraction of z_1



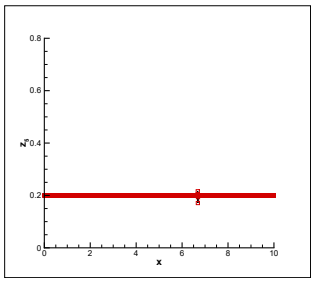
(d) mass fraction of z_2



(e) mass fraction of z_3



(f) mass fraction of z_3



(g) mass fraction of z_3

FIG. 6.3. *Numerical solutions of Example* 6.4 *at* $t = 0.5$.

The parameters are $T_1 = 2$, $B_1 = 10^6$, $\alpha_1 = 0$, $q_1 = 200$, $q_2 = 0$, $q_3 = 0$, $q_4 = 0$, $M_1 = 16$, $M_2 = 32$, $M_3 = 44$, $M_4 = 18$. The initial values consist of totally burnt gas inside of a circle with radius 10 and totally unburnt gas everywhere outside this circle. The set up is as follows

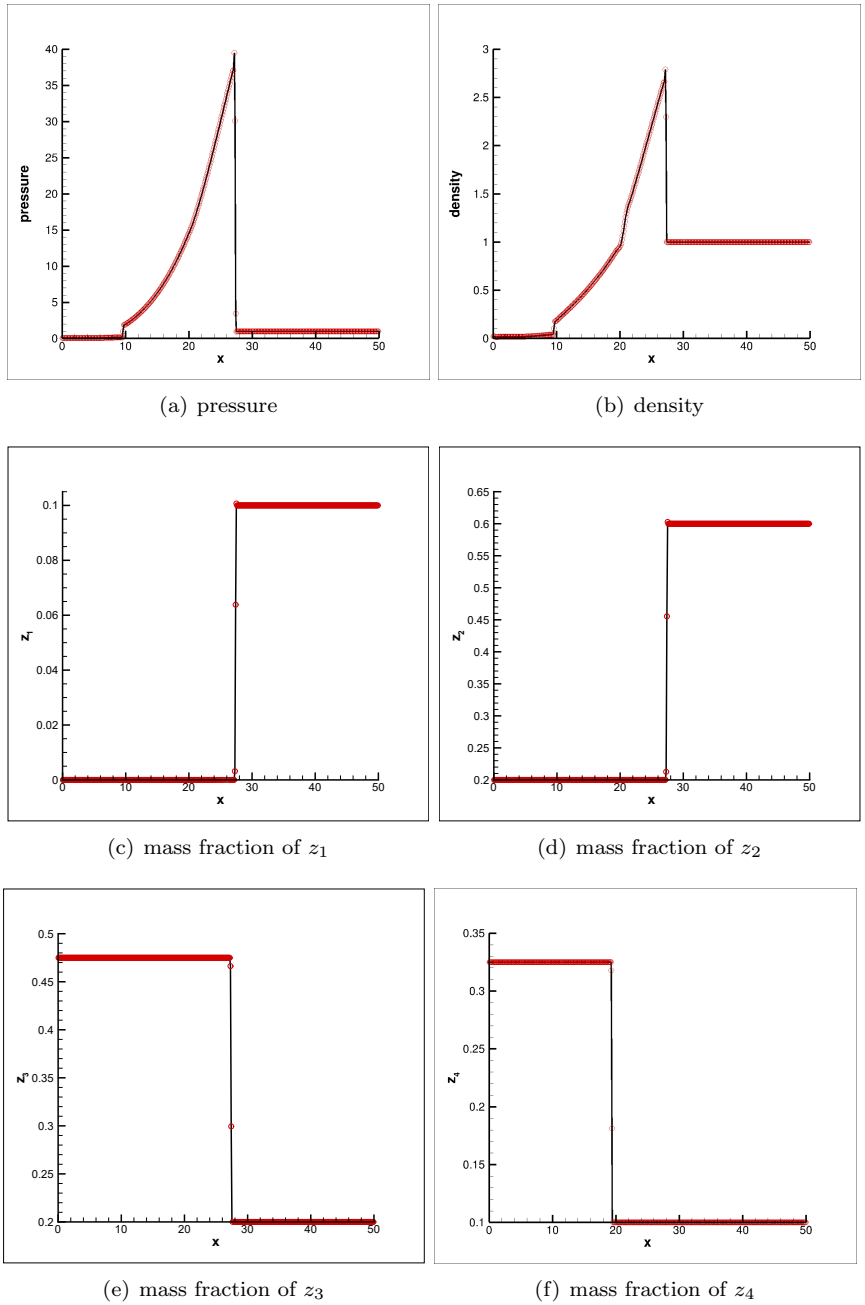
$$(\rho, u, v, p, z_1, z_2, z_3, z_4)(x, y, 0) = \begin{cases} (2, 10x/r, 10y/r, 40, 0, 0.2, 0.475, 0.325), & r \leqslant 10, \\ (1, 0, 0, 1, 0.1, 0.6, 0.2, 0.1), & r > 10. \end{cases}$$

The computational domain is $[0, 50] \times [0, 50]$.

TABLE 6.4
Accuracy test for the two dimensional system.

	Without limiter				With limiter				
N	L^2 norm	Order	L^∞ norm	Order	L^2 norm	Order	L^∞ norm	Order	Percentage
	2nd order RK, cfl=0.1, c=100								
10	5.23E-03	–	1.28E-02	–	5.99E-03	–	1.71E-02	–	28.00%
20	1.31E-03	2.00	3.76E-03	1.76	1.49E-03	2.01	5.25E-03	1.70	13.75%
40	3.26E-04	2.00	1.01E-03	1.90	3.63E-04	2.04	1.42E-03	1.89	5.88%
80	8.14E-05	2.00	2.60E-04	1.95	8.85E-05	2.04	3.74E-04	1.92	2.48%
160	2.03E-05	2.00	6.60E-05	1.98	2.16E-05	2.03	1.00E-04	1.90	1.05%
	2nd order RK, cfl=0.1, c=10,000								
10	5.12E-03	–	1.33E-02	–	5.91E-03	–	1.74E-02	–	30.00%
20	1.29E-03	1.99	3.88E-03	1.77	1.46E-03	2.02	5.10E-03	1.78	14.00%
40	3.22E-04	2.00	1.03E-03	1.91	3.49E-04	2.07	1.44E-03	1.82	5.31%
80	8.04E-05	2.00	2.66E-04	1.96	8.43E-05	2.05	3.80E-04	1.92	1.80%
160	2.01E-05	2.00	6.73E-05	1.98	2.08E-05	2.02	1.02E-04	1.90	0.63%
	2nd order multistep, cfl=0.1, c=100								
10	5.13E-03	–	1.30E-02	–	5.78E-03	–	1.70E-02	–	34.00%
20	1.31E-03	1.97	3.77E-03	1.79	1.48E-03	1.97	5.22E-03	1.70	21.50%
40	3.27E-04	2.00	1.01E-03	1.90	3.61E-04	2.03	1.41E-03	1.89	8.81%
80	8.18E-05	2.00	2.60E-04	1.95	8.83E-05	2.03	3.73E-04	1.92	3.84%
160	2.05E-05	2.00	6.61E-05	1.98	2.16E-05	2.03	9.96E-05	1.90	1.59%
	2nd order multistep, cfl=0.1, c=10,000								
10	5.01E-03	–	1.36E-02	–	5.65E-03	–	1.65E-02	–	37.00%
20	1.29E-03	1.96	3.92E-03	1.79	1.44E-03	1.97	5.07E-03	1.71	22.00%
40	3.23E-04	1.99	1.05E-03	1.91	3.47E-04	2.05	1.45E-03	1.80	7.88%
80	8.11E-05	2.00	2.70E-04	1.95	8.46E-05	2.04	3.85E-04	1.92	2.53%
160	2.03E-05	2.00	6.97E-05	1.96	2.10E-05	2.01	1.03E-04	1.90	0.89%
	3rd order multistep, cfl=0.03, c=100								
10	6.42E-04	–	3.45E-03	–	1.26E-03	–	5.58E-03	–	14.00%
20	8.07E-05	2.99	4.33E-04	2.99	8.47E-05	3.90	4.96E-04	3.49	3.00%
40	1.01E-05	3.00	5.41E-05	3.00	1.01E-05	3.07	5.45E-05	3.19	0.56%
80	1.26E-06	3.00	6.75E-06	3.00	1.26E-06	3.00	6.75E-06	3.01	0.11%
160	1.58E-07	3.00	8.44E-07	3.00	1.58E-07	3.00	8.44E-07	3.00	0.04%
	3rd order multistep, cfl=0.03, c=10,000								
10	7.22E-04	–	4.19E-03	–	1.27E-03	–	6.09E-03	–	15.00%
20	9.34E-05	2.95	6.11E-04	2.78	9.55E-05	3.74	6.11E-04	3.32	1.25%
40	1.17E-05	2.99	7.65E-05	3.00	1.18E-05	3.02	7.65E-05	3.00	0.38%
80	1.47E-06	3.00	9.65E-06	2.99	1.47E-06	3.00	9.65E-06	2.99	0.11%
160	1.84E-07	3.00	1.20E-06	3.00	1.84E-07	3.00	1.20E-06	3.00	0.04%

This is a radially symmetric problem, and the detonation front is circular. We take $N_x = N_y = 600$ and CFL $= 0.01$. We test both the second-order RK method and the second-order multistep method with piecewise P^1 polynomials. Figure 6.4 shows the one dimensional cuts of pressure, density, and mass fractions along the line $x = y$ at $t = 2$ by using RK method. We can see that our scheme preserve the positivity of the density and pressure, and the two bounds 0 and 1 of each mass fraction. Also, we can see that our schemes can capture the detonations well. The results obtained by using second-order multistep method is exactly the same, so we skip them here.



(a) pressure (b) density

(c) mass fraction of z_1 (d) mass fraction of z_2

(e) mass fraction of z_3 (f) mass fraction of z_4

FIG. 6.4. *Numerical solutions of Example* 6.6 *along the line* $x = y$ *at* $t = 2$. *The* 2*nd order RK method with piecewise* P^1 *polynomials.*

7. Conclusion. In this paper, we have introduced the high-order conservative bound-preserving DG methods for stiff multispecies detonation. A new explicit time integration has been constructed. Numerical experiments demonstrated the good performance of the scheme.

REFERENCES

[1] N. Chuenjarern, Z. Xu, and Y. Yang, *High-order bound-preserving discontinuous Galerkin methods for compressible miscible displacements in porous media on triangular meshes*, J. Comput. Phys., 378 (2019), pp. 110–128.

[2] J. F. Clarke, S. Karni, J. J. Quirk, P. L. Roe, L. G. Simmonds, and E. F. Toro, *Numerical computation of two-dimensional unsteady detonation waves in high energy solids*, J. Comput. Phys., 106 (1993), pp. 215–233.

[3] B. Cockburn, S. Hou, and C.-W. Shu, *The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws* IV*: The multidimensional case*, Math. Comput., 54 (1990), pp. 545–581.

[4] S. Gottlieb, C.-W. Shu, and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Rev., 43 (2001), pp. 89–112.

[5] H. Guo and Y. Yang, *Bound-preserving discontinuous Galerkin method for compressible miscible displacement problem in porous media*, SIAM J. Sci. Comput., 39 (2017), pp. A1969–A1990.

[6] J. Huang and C.-W. Shu, *Bound-preserving modified exponential Runge–Kutta discontinuous Galerkin methods for scalar hyperbolic equations with stiff source terms,* J. Comput. Phys., 361 (2018), pp. 111–135.

[7] S. Kopecz and A. Meister, *On order conditions for modified Patankar–Runge-Kutta schemes*, Appl. Numer. Math., 123 (2018), pp. 159–179.

[8] S. Kopecz and A. Meister, *Unconditionally positive and conservative third order modified Patankar–Runge-Kutta discretizations of production-destruction systems*, BIT, 58 (2018), pp. 691–728.

[9] R. J. LeVeque and H. C. Yee, *A study of numerical methods for hyperbolic conservation laws with stiff source terms,* J. Comput. Phys., 86 (1990), pp. 187–210.

[10] Y. Lv and M. Ihme, *Discontinuous Galerkin method for multicomponent chemically reacting flows and combustion*, J. Comput. Phys., 270 (2014), pp. 105–137.

[11] Y. Lv and M. Ihme, *High-order discontinuous Galerkin method for applications to multicomponent and chemically reacting flows*, Acta Mechanica Sinica, 33 (2017), pp. 486–499.

[12] T. Qin and C.-W. Shu, *Implicit positivity-preserving high order discontinuous Galerkin methods for conservation laws,* SIAM J. Sci. Comput., 40 (2018), pp. A81–A107.

[13] T. Qin, C.-W. Shu, and Y. Yang, *Bound-preserving discontinuous Galerkin methods for relativistic hydrodynamics*, J. Comput. Phys., 315 (2016), pp. 323–347.

[14] W. H. Reed and T. R. Hill, *Triangular mesh methods for the Neutron transport equation*, Los Alamos Scientific Laboratory Report LA-UR-73-479, Los Alamos, NM, 1973.

[15] C.-W. Shu, *Total-variation-diminishing time discretizations*, SIAM J. Statist. Sci. Comput., 9 (1988), pp. 1073–1084.

[16] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, J. Comput. Phys., 77 (1988), pp. 439–471.

[17] G. Strang, *On the construction and comparison of difference schemes,* SIAM J. Numer. Anal., 5 (1968), pp. 506–517.

[18] C. Wang, X. Zhang, C.-W. Shu, and J. Ning, *Robust high order discontinuous Galerkin schemes for two-dimensional gaseous detonations,* J. Comput. Phys., 231 (2012), pp. 653–665

[19] W. Wang, C.-W. Shu, H. C. Yee, D. V. Kotov, and B. Sj ogreen, *High order finite difference methods with subcell resolution for stiff multispecies detonation capturing*, Comm. Comput. Phys., 17 (2015), pp. 317–336.

[20] Y. Yang, D. Wei and, C.-W. Shu, *Discontinuous Galerkin method for Krause's consensus models and pressureless Euler equations*, J. Comput. Phys., 252 (2013), pp. 109–127.

[21] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, J. Comput. Phys., 229 (2010), pp. 3091–3120.

[22] X. Zhang and C.-W. Shu, *On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, J. Comput. Phys., 229 (2010), pp. 8918–8934.

[23] X. Zhang and C.-W. Shu, *Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms*, J. Comput. Phys., 230 (2011), pp. 1238–1248.

[24] X. Zhang, Y. Xia and 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), pp. 29–32.

[25] X. Zhao, Y. Yang, and C. Seyler, *A positivity-preserving semi-implicit discontinuous Galerkin scheme for solving extended magnetohydrodynamics equations*, J. Comput. Phys., 278 (2014), pp. 400–415.