Bound-preserving modified exponential Runge-Kutta discontinuous Galerkin methods for scalar hyperbolic equations with stiff source terms

Juntao Huang[1] and Chi-Wang Shu[2]

Abstract

In this paper, we develop bound-preserving modified exponential Runge-Kutta (RK) discontinuous Galerkin (DG) schemes to solve scalar hyperbolic equations with stiff source terms by extending the idea in Zhang and Shu [43]. Exponential strong stability preserving (SSP) high order time discretizations are constructed and then modified to overcome the stiffness and preserve the bound of the numerical solutions. It is also straightforward to extend the method to two dimensions on rectangular and triangular meshes. Even though we only discuss the bound-preserving limiter for DG schemes, it can also be applied to high order finite volume schemes, such as weighted essentially non-oscillatory (WENO) finite volume schemes as well.

Keywords: Discontinuous Galerkin method; finite volume scheme; scalar hyperbolic equations; stiff source; bound-preserving scheme; high order accuracy; exponential Runge-Kutta method

[1]Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua University, Beijing 100084, China. E-mail: huangjt13@mails.tsinghua.edu.cn

[2]Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. E-mail: shu@dam.brown.edu. Research supported by ARO grant W911NF-15-1-0226 and NSF grants DMS-1418750 and DMS-1719410.

1 Introduction

We consider the scalar hyperbolic equations with a stiff source term in the multi-dimensional case:

$$u_t + \nabla \cdot F(u) = \frac{1}{\varepsilon} s(u), \quad u(x,0) = u_0(x), \tag{1.1}$$

with $\varepsilon > 0$ being a constant. We assume that the source term $s = s(u)$ is dissipative in the sense that

$$s'(u) \leq 0, \quad s(0) = 0. \tag{1.2}$$

Then the entropy solution to (1.1) satisfies the bound-preserving properties (see e.g. [23]):

i) $\|u(\cdot,t)\|_{L^\infty} \leq \|u_0\|_{L^\infty}$ for $t > 0$;

ii) If $u_0(x) \geq 0$ for any x, then $u(x,t) \geq 0$ for any x and $t > 0$; Also, if $u_0(x) \leq 0$ for any x, then $u(x,t) \leq 0$ for any x and $t > 0$.

In this paper, our target is to design schemes that numerically preserve these properties independent of ε.

We remark that the hypothesis (1.2) on the source term is realistic. Indeed, there are many physical models which fall into this category, e.g., the model of combustion [2], Euler equations of gas dynamics with heat transfer [19], and shallow water equations with friction force of the bottom [5]. The motivation of this study is to construct numerical schemes for the simple scalar model (1.1). In the future, we expect to generalize the idea in this work to hyperbolic systems with stiff source terms which have similar dissipation properties.

Moreover, without the assumption (1.2), the L^∞-norm decreasing property for the solution can no longer be expected to hold [3]. For scalar equations with more general nonlinear source terms, e.g. the source term with multiple equilibrium points in [24], we cannot expect the monotonicity for the L^∞-norm in general. However, the invariant region, between two equilibrium points of the source term in [24], could be expected to be preserved. In [9], Donat et al. studied the conditions that preserve the invariant region for the implicit-explicit

(IMEX) RK schemes. Nevertheless, it is quite frustrating that even the first-order semi-implicit scheme could only preserve the invariant region under the conditions dependent on the stiffness parameter ε. This issue is beyond the scope of this paper.

Bound-preserving high order numerical methods for conservation laws have been actively studied in the last few years. Zhang and Shu constructed uniformly high order accurate schemes satisfying a strict maximum principle for scalar conservation laws [43]. By splitting the cell average into a convex combination of values at quadrature points and then rewriting the scheme for the cell average into a convex combination of formally monotone schemes, the numerical solutions satisfy a maximum principle with the aid of a simple scaling limiter first introduced in [27]. The framework was then successfully applied to compressible Euler equations [44, 46], shallow water equations [37, 36], relativistic hydrodynamics [28], convection-diffusion equations [47], Navier-Stokes equations [42] and so on. There exists another popular approach for preserving physical bounds of the numerical solutions by Xu [39] where the parameterized maximum principle preserving flux limiter was proposed. This approach works well numerically but could be shown to keep accuracy only up to third-order. This technique was also applied to various problems for preserving the bound of physical quantities (see e.g. [38, 26, 35, 6]). For details, we refer readers to the review papers [45, 40].

The numerical approximation of hyperbolic equations with stiff source terms has been intensively studied since the pioneering work by LeVeque and Yee [24]. The main difficulty is the incorrect propagation speed of the discontinuities which may happen with insufficient spatial or temporal resolution. Various techniques have been proposed to overcome this difficulty (see e.g. [1, 10, 34]). Error estimates to the schemes for hyperbolic equations with stiff source terms have been derived (see e.g. [33, 30, 3]).

There are rare works on bound-preserving schemes for hyperbolic equations with stiff sources. Zhao et al. developed a positivity-preserving semi-implicit DG scheme for the extended magnetohydrodynamics [48]. In their work, the time splitting integration is applied and thus the scheme is at most first-order accurate in time. Recently, Chertock et al.

proposed a class of second-order semi-implicit time integration methods with steady-state and sign-preserving property for systems of ODEs with stiff terms [4], and successfully applied it to the shallow water equations with stiff friction term [5]. More recently, we constructed a class of second-order positivity-preserving implicit-explicit (IMEX) RK methods for the system of ordinary differential equations arising from the semi-discretization of the Kerr-Debye model [18]. We remark that all the schemes mentioned above are designed for specific models and are only up to second-order accuracy in time. It is highly nontrivial to design a time integrator which has high-order accuracy and allows the time step size much larger than the stiffness parameter ε.

In this work, we first construct high-order exponential strong stability preserving (SSP) RK and multi-step methods. Since the exponential functions decay to zero too fast, the scheme only produces solutions with essentially zero value given inconsistent initial values in the unresolved region. However, in some cases (e.g. source terms of polynomial type, see numerical examples in section 5), the exact solution is of order ε instead of being essentially zero. For capturing this kind of solutions more accurately, we modify the exponential scheme by replacing the exponential function with a polynomial without destroying the accuracy, hence obtaining high-order modified exponential RK methods. The bound-preserving property and the weak asymptotic-preserving property are shown for these methods. Combining these time integration methods with the semi-discrete DG schemes, we successfully obtain the bound-preserving DG schemes.

The paper is organized as follows. In section 2, we construct exponential RK and multi-step methods and then propose modified exponential RK methods by replacing the exponential functions with polynomials. For these methods, we show the bound-preserving property and the weak asymptotic-preserving property. In section 3, we combine the time discretization with the DG method and discuss the bound-preserving property by using the scaling limiter. Numerical examples for non-stiff and stiff ODEs are conducted in section 4. We validate our method for solving the scalar hyperbolic equations with non-stiff and stiff source

terms in 1D and 2D cases in section 5. Some concluding remarks are given in section 6.

2 Time discretization

In this section, we focus on time discretization and consider initial value problem of ordinary differential equation (ODE) for the scalar function $u = u(t)$:

$$u_t = f(u) + \frac{1}{\varepsilon}s(u), \quad u(0) = u_0, \tag{2.1}$$

with constant $\varepsilon > 0$. Here $f(u)$ denotes the non-stiff part and $\frac{1}{\varepsilon}s(u)$ the stiff part, which correspond to the convection term and the source term in semi-discrete schemes for (1.1), respectively. We make the following assumption on (2.1):

Assumption 2.1. 1. The Euler forward scheme for the non-stiff part satisfies the maximum principle: There exists $\Delta t_E > 0$, such that for any $0 < \Delta t \le \Delta t_E$, if $0 \le u \le M$, then

$$0 \le u + \Delta t f(u) \le M;$$

if $-M \le u \le 0$, then

$$-M \le u + \Delta t f(u) \le 0.$$

2. The stiff source term satisfies the dissipative property (1.2): $s'(u) \le 0$ and $s(0) = 0$.

Our goal is to design a scheme for (2.1) which enjoys two properties:

i) $|u(t)| \le |u(0)|$ for $t > 0$;

ii) If $u(0) \ge 0$, then $u(t) \ge 0$ for $t > 0$; Also, if $u(0) \le 0$, then $u(t) \le 0$ for $t > 0$.

We follow the idea of exponential Runge-Kutta methods (see e.g. [16]) and rewrite (2.1) as follows:

$$u_t = f(u) + \frac{1}{\varepsilon}(s(u) + \mu u) - \frac{\mu}{\varepsilon}u, \tag{2.2}$$

with μ a constant to be determined. Then the exponential form is obtained:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{\frac{\mu}{\varepsilon}t}u\right) = e^{\frac{\mu}{\varepsilon}t}\left(f(u) + \frac{1}{\varepsilon}(s(u) + \mu u)\right). \tag{2.3}$$

In the following, we will use explicit Runge-Kutta methods and multi-step methods to discretize (2.3).

2.1 Euler forward

We start with the Euler forward time discretization on (2.3) and obtain

$$u^{n+1} = e^{-\frac{\mu}{\varepsilon}\Delta t}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right). \tag{2.4}$$

The bound-preserving property for (2.4) is easily shown below:

Proposition 2.1. *For the exponential Euler forward scheme* (2.4)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*, and* $\mu \ge \sup_{0\le u\le M}(-s'(u))$*, then*

$$0 \le u^{n+1} \le e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \le M. \tag{2.5}$$

The conclusion is the same for the negative value of u^n*.*

Proof. On Assumption 2.1, we have

$$0 \le u^n + \Delta t f(u^n) \le M,$$

with $0 \le \Delta t \le \Delta t_E$. Set $G(u) := s(u) + \mu u$. Since $\mu \ge \sup_{0\le u\le M}(-s'(u))$, $G(u)$ is non-decreasing for $0 \le u \le M$. Then $G(u^n) \ge G(0) = s(0) = 0$ and $G(u^n) \le G(M) = s(M) + \mu M \le \mu M$ (Here we use $s'(u) \le 0$ and thus $s(M) \le s(0) = 0$). Finally, we arrive at the estimate for u^{n+1}:

$$0 \le u^{n+1} \le e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \le M,$$

where we have used the inequality $e^x \ge 1 + x$. The argument is the same for $-M \le u^n \le 0$ with $M > 0$. □

Notice that the exponential function $e^{-\frac{\mu}{\varepsilon}\Delta t}$ in (2.4) decays to zero too fast. If we take the initial value $u_0 = O(1)$ and $\Delta t \gg \varepsilon$, then $\frac{\mu}{\varepsilon}\Delta t \gg 1$ and thus numerically only solution with the value being essentially zero is obtained. However, in some cases (e.g. source terms

of polynomial type, see numerical examples in section 5), the exact solution is of order ε instead of being essentially zero. For capturing this kind of solutions more accurately, we modify the scheme (2.4) by replacing the exponential function with a polynomial without destroying the accuracy:

$$u^{n+1} = \frac{1}{1+\frac{\mu}{\varepsilon}\Delta t}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right). \tag{2.6}$$

It is easy to see that this scheme also enjoys the bound-preserving property with the same assumption in Proposition 2.1.

Recall that a classical asymptotic-preserving (AP) scheme requires that, the numerical scheme projects any initial data into the local equilibrium [20], with an accuracy of $O(\varepsilon)$ in one time step $\Delta t \gg \varepsilon$, i.e.,

$$s(u^n) = O(\varepsilon), \quad n \geq 1, \tag{2.7}$$

for any initial value u_0. Obviously, the exponential Euler forward method (2.4) is AP in this classical sense. However, the modified Euler forward scheme (2.6) does not satisfy this property. Instead, it is AP in the weak sense [20].

Proposition 2.2 (weak AP of the modified Euler forward scheme)**.** *On Assumption 2.1, and further assuming that $s(u) \neq 0$ for $u \neq 0$, the modified exponential Euler forward scheme* (2.6) *is AP in the weak sense: for any $\varepsilon > 0$ and any initial value u_0, and $\Delta t \gg \varepsilon$, there exists an integer $N_\varepsilon \geq 1$ (independent of Δt), such that*

$$s(u^n) = O(\varepsilon), \quad n \geq N_\varepsilon. \tag{2.8}$$

Proof. Without loss of generality, we assume that $u_0 \geq 0$. We provide a proof by contradiction. Suppose that there exists $\Delta t \gg \varepsilon$, such that $\forall N \geq 1$, $\exists n > N$, $s(u^n) = O(\varepsilon)$ does not hold. Since u^n is non-increasing w.r.t. n, we deduce that $\exists N \geq 1$ such that $u^n \gg \varepsilon$ for any $n \geq N$.

$$u^{n+1} = \frac{1}{1+\frac{\mu}{\varepsilon}\Delta t}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right),$$

$$
\begin{aligned}
&\leq u^n \left(1 + \frac{\frac{\Delta t}{\varepsilon}\frac{s(u^n)}{u^n}}{1+\frac{\Delta t}{\varepsilon}\mu(u^n)}\right),\\
&\leq u^n \left(1 + \frac{\frac{s(u^n)}{u^n}}{1+\mu(u^n)}\right),
\end{aligned}
$$

Since $\varepsilon \ll u^n \leq u^0$ for any $n \geq N$ and $s(u) \neq 0$ for $u \neq 0$, we have $\frac{\frac{s(u^n)}{u^n}}{1+\mu(u^n)} \leq -r$ with $r = r(\varepsilon, u^0) > 0$. Thus $u^{n+1} \leq (1-r)u^n$. By iteration, we have $u^{N+k} \leq (1-r)^k u^N$ for any $k \geq 1$, which makes a contradiction if we take k sufficiently large. □

2.2 Runge-Kutta method

2.2.1 Second-order Runge-Kutta method

Now let us discretize (2.3) with the second-order SSP RK method [14] and obtain

$$u^{(1)} = e^{-\frac{\mu}{\varepsilon}\Delta t}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right), \tag{2.9a}$$

$$u^{n+1} = \frac{1}{2}e^{-\frac{\mu}{\varepsilon}\Delta t}u^n + \frac{1}{2}\left((u^{(1)} + \Delta t f(u^{(1)})) + \frac{\Delta t}{\varepsilon}(s(u^{(1)}) + \mu u^{(1)})\right). \tag{2.9b}$$

Proposition 2.3. *For the second-order exponential RK scheme* (2.9)*, if* $0 \leq u^n \leq M$*, and* $0 < \Delta t \leq \Delta t_E$*,* $\mu \geq \sup_{0\leq u\leq M}(-s'(u))$*, then*

$$0 \leq u^{(1)} \leq e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \leq M, \tag{2.10}$$

$$0 \leq u^{n+1} \leq e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2)M \leq M. \tag{2.11}$$

The conclusion is the same for the negative value of u^n*.*

Proof. The first stage is Euler forward. By Proposition 2.1, we deduce that,

$$0 \leq u^{(1)} \leq e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M.$$

For the second stage, since $\mu \geq \sup_{0\leq u\leq M}(-s'(u))$, $G(u) := s(u) + \mu u$ is non-decreasing for $0 \leq u \leq M$. As $0 \leq u^{(1)} \leq e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \leq M$, we have $G(u^{(1)}) \geq G(0) = 0$ and

$$G(u^{(1)}) \leq G(e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M) \leq \mu e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M. \tag{2.12}$$

Now it is easy to see that $u^{n+1} \geq 0$. The upper bound of u^{n+1} is:

$$\begin{aligned} u^{n+1} \leq & \frac{1}{2} e^{-\frac{\mu}{\varepsilon}\Delta t} M + \frac{1}{2} e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M + \frac{\Delta t}{2\varepsilon}\mu e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \\ = & e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2)M \leq M. \end{aligned}$$

The argument is the same for the negative value of u^n. □

Remark 2.4. Note that in the proof of Propositions 2.1 and 2.3, the upper bound of u^{n+1} is controlled by using the inequalities $e^x \geq 1 + x$ and $e^x \geq 1 + x + \frac{x^2}{2}$, respectively, which are exactly the first several terms of the Taylor series of e^x at $x = 0$. Actually, in the estimate of the upper bound, f and s are both dropped and do not make any contribution. Hence, the problem is essentially equivalent to considering only the case that $f = s = 0$. In this case, the ODE (2.1) is reduced to $\frac{\mathrm{d}u}{\mathrm{d}t} = 0$, which is rewritten as $\frac{\mathrm{d}}{\mathrm{d}t}(e^{\frac{\mu}{\varepsilon}t}u) = \frac{\mu}{\varepsilon}e^{\frac{\mu}{\varepsilon}t}u$ in the exponential form. If we set $v = e^{\frac{\mu}{\varepsilon}t}u$ and $\lambda = \frac{\mu}{\varepsilon}$, then the simplified case is nothing but the test equation $\frac{\mathrm{d}v}{\mathrm{d}t} = \lambda v$. When the explicit RK methods are applied, we will obtain $v^{n+1} = \phi(\lambda\Delta t)v^n$ with ϕ being the stability function, which is equivalent to $u^{n+1} = e^{-\lambda\Delta t}\phi(\lambda\Delta t)u^n$. It is well-known that, for an s-stage explicit RK method with order s, the stability function $\phi(z) = 1 + z + \cdots + \frac{z^s}{s!}$. Hence, it is of no surprise that the upper bound of u^{n+1} is always controlled by the Taylor expansion inequality of e^x at $x = 0$. We will rediscover this fact for the third-order exponential Runge-Kutta method later.

Replacing exponential functions by polynomials in (2.9) results in the following second-order modified RK scheme:

$$u^{(1)} = \frac{1}{1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right), \tag{2.13a}$$

$$u^{n+1} = \frac{1}{2(1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2)}u^n + \frac{1}{2}\left((u^{(1)} + \Delta t f(u^{(1)})) + \frac{\Delta t}{\varepsilon}(s(u^{(1)}) + \mu u^{(1)})\right), \tag{2.13b}$$

which has the following bound-preserving property:

Proposition 2.5. *For the second-order modified exponential RK scheme* (2.13)*, if* $0 \leq u^n \leq M$*, and* $0 < \Delta t \leq \Delta t_E$*,* $\mu \geq \sup_{0 \leq u \leq M}(-s'(u))$*, then*

$$0 \leq u^{(1)} \leq \frac{1 + \frac{\mu}{\varepsilon}\Delta t}{1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2} M \leq M, \tag{2.14}$$

$$0 \le u^{n+1} \le M. \tag{2.15}$$

The conclusion is similar for the negative value of u^n.

The proof is similar to that of Proposition 2.1. The only difference is that here in the estimate of u^{n+1}, we should use the sharper upper bound $\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}M$ for $u^{(1)}$. This fact will make a difference in the bound-preserving limiters for the DG and finite volume methods. We will turn back to this issue in section 3.

Proposition 2.6 (weak AP of the modified second-order exponential RK scheme)**.** *On Assumption 2.1, and further assuming that $s(u) \neq 0$ for $u \neq 0$, the second-order modified exponential Euler forward scheme is AP in the weak sense: for any $\varepsilon > 0$ and any initial value u^0, and $\Delta t \gg \varepsilon$, there exists an integer $N_\varepsilon \geq 1$ (independent of Δt), such that*

$$s(u^n) = O(\varepsilon), \quad n \geq N_\varepsilon. \tag{2.16}$$

Proof. Without loss of generality, we assume that $u^0 \geq 0$. We provide a proof by contradiction. Suppose that there exists $\Delta t \gg \varepsilon$, such that $\forall N \geq 1$, $\exists n > N$, $s(u^n) = O(\varepsilon)$ does not hold. Since u^n is non-increasing w.r.t. n, we deduce that $\exists N \geq 1$ such that $u^n \gg \varepsilon$ for any $n \geq N$. With $z = \frac{\mu}{\varepsilon}\Delta t$ and $x = \frac{\Delta t}{\varepsilon}$, we have

$$\begin{aligned}
u^{n+1} \leq & \frac{1}{2(1+z+\frac{1}{2}z^2)}u^n + \frac{1}{2}(1+z)u^{(1)} \\
\leq & \frac{1}{2(1+z+\frac{1}{2}z^2)}u^n + \frac{1}{2}(1+z)\frac{1}{1+z+\frac{1}{2}z^2}(u^n + \frac{\Delta t}{\varepsilon}(s(u^n)+\mu u^n)) \\
= & u^n(1 + \frac{1+z}{2(1+z+\frac{1}{2}z^2)}\frac{\Delta t}{\varepsilon}\frac{s(u^n)}{u^n}) \\
= & u^n(1 + \frac{x(1+\mu(u^n)x)}{2(1+\mu(u^n)x+\frac{1}{2}(\mu(u^n)x)^2)}\frac{s(u^n)}{u^n}).
\end{aligned}$$

With $\mu > 0$, the expression $\frac{x(1+\mu(u^n)x)}{2(1+\mu(u^n)x+\frac{1}{2}(\mu(u^n)x)^2)}$, as a function of x, is non-decreasing for $0 \leq x < \infty$. Thus for $x \gg 1$, we have

$$\frac{1+\mu(u^n)}{2(1+\mu(u^n)+\frac{1}{2}(\mu(u^n))^2)} < \frac{x(1+\mu(u^n)x)}{2(1+\mu(u^n)x+\frac{1}{2}(\mu(u^n)x)^2)} < \frac{1}{\mu(u^n)},$$

and thus

$$u^{n+1} < u^n(1 + \frac{1+\mu(u^n)}{2(1+\mu(u^n)+\frac{1}{2}(\mu(u^n))^2)}\frac{s(u^n)}{u^n}).$$

Since $\varepsilon \ll u^n \le u^0$ for any $n \ge N$ and $s(u) \neq 0$ for $u \neq 0$, we have $\frac{1+\mu(u^n)}{2(1+\mu(u^n)+\frac{1}{2}(\mu(u^n))^2)}\frac{s(u^n)}{u^n} \le -r$ with $r = r(\varepsilon, u^0) > 0$. Immediately we obtain $u^{n+1} \le (1-r)u^n$. By iteration $u^{N+k} \le (1-r)^k u^N$ for $k \ge 1$, which makes a contradiction if we take k sufficiently large. □

2.2.2 Third-order Runge-Kutta method

We discretize (2.3) with the classical third-order SSP RK method [31]:

$$u^{(1)} = e^{-\frac{\mu}{\varepsilon}\Delta t}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right), \tag{2.17a}$$

$$u^{(2)} = \frac{3}{4}e^{-\frac{1}{2}\frac{\mu}{\varepsilon}\Delta t}u^n + \frac{1}{4}e^{\frac{1}{2}\frac{\mu}{\varepsilon}\Delta t}\left((u^{(1)} + \Delta t f(u^{(1)})) + \frac{\Delta t}{\varepsilon}(s(u^{(1)}) + \mu u^{(1)})\right), \tag{2.17b}$$

$$u^{n+1} = \frac{1}{3}e^{-\frac{\mu}{\varepsilon}\Delta t}u^n + \frac{2}{3}e^{-\frac{1}{2}\frac{\mu}{\varepsilon}\Delta t}\left((u^{(2)} + \Delta t f(u^{(2)})) + \frac{\Delta t}{\varepsilon}(s(u^{(2)}) + \mu u^{(2)})\right). \tag{2.17c}$$

The bound-preserving property is given as follows. The proof is omitted since it is similar to that for the second order scheme.

Proposition 2.7. *For the third-order exponential RK scheme* (2.17)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*,* $\mu \ge \sup_{0\le u\le \frac{3}{e}M}(-s'(u))$*, then*

$$0 \le u^{(1)} \le e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t)M \le M, \tag{2.18}$$

$$0 \le u^{(2)} \le e^{-\frac{\mu}{2\varepsilon}\Delta t}(1 + \frac{\mu}{2\varepsilon}\Delta t + \frac{1}{4}(\frac{\mu}{\varepsilon}\Delta t)^2)M \le \frac{3}{e}M, \tag{2.19}$$

$$0 \le u^{n+1} \le e^{-\frac{\mu}{\varepsilon}\Delta t}(1 + \frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2 + \frac{1}{6}(\frac{\mu}{\varepsilon}\Delta t)^3)M \le M. \tag{2.20}$$

The conclusion is similar for the negative value of u^n*.*

Remark 2.8. Note that we could only prove that $0 \le u^{(2)} \le e^{-\frac{\mu}{2\varepsilon}\Delta t}(1 + \frac{\mu}{2\varepsilon}\Delta t + \frac{1}{4}(\frac{\mu}{\varepsilon}\Delta t)^2)M$. The upper bound of $u^{(2)}$ is not less than M necessarily, but could be controlled by $\frac{3}{e}M$.

By replacing exponential functions by polynomials in (2.17), we have the following scheme:

$$u^{(1)} = \frac{1}{1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3}\left((u^n + \Delta t f(u^n)) + \frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right), \tag{2.21a}$$

$$u^{(2)} = \frac{3}{4(1+\frac{1}{2}z+\frac{1}{8}z^2+\frac{1}{48}z^3)}u^n + \frac{1+z+\frac{1}{2}z^2+\frac{1}{6}z^3}{4(1+\frac{1}{2}z+\frac{1}{8}z^2+\frac{1}{48}z^3)}\left((u^{(1)}+\Delta t f(u^{(1)}))+\frac{\Delta t}{\varepsilon}(s(u^{(1)})+\mu u^{(1)})\right), \tag{2.21b}$$

$$u^{n+1} = \frac{1}{3(1+z+\frac{1}{2}z^2+\frac{1}{6}z^3)}u^n + \frac{2(1+\frac{1}{2}z+\frac{1}{8}z^2+\frac{1}{48}z^3)}{3(1+z+\frac{1}{2}z^2+\frac{1}{6}z^3)}\left((u^{(2)}+\Delta t f(u^{(2)}))+\frac{\Delta t}{\varepsilon}(s(u^{(2)})+\mu u^{(2)})\right), \tag{2.21c}$$

with $z = \frac{\mu}{\varepsilon}\Delta t$ and $\mu = \mu(u^n) = \sup_{0\le u\le 1.13652u^n}(-s'(u))$.

Proposition 2.9. *For the third-order polynomial RK scheme* (2.21)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*,* $\mu \ge \sup_{0\le u\le 1.13652M}(-s'(u))$*, then*

$$0 \le u^{(1)} \le \frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2+\frac{1}{6}(\frac{\mu}{\varepsilon}\Delta t)^3}M \le M, \tag{2.22}$$

$$0 \le u^{(2)} \le \frac{1+\frac{1}{2}\frac{\mu}{\varepsilon}\Delta t+\frac{1}{4}(\frac{\mu}{\varepsilon}\Delta t)^2}{1+\frac{1}{2}\frac{\mu}{\varepsilon}\Delta t+\frac{1}{8}(\frac{\mu}{\varepsilon}\Delta t)^2+\frac{1}{48}(\frac{\mu}{\varepsilon}\Delta t)^3}M \le 1.13652M, \tag{2.23}$$

$$0 \le u^{n+1} \le M. \tag{2.24}$$

The conclusion is similar for the negative value of u^n*.*

The proof to the weak AP property of the third-order modified exponential RK method (2.21) is similar as before and thus omitted here.

2.2.3 Fourth-order Runge-Kutta method

There exists no explicit four-stage fourth-order SSP Runge-Kutta method with non-negative coefficients [22, 13]. However, fourth order SSP methods with more than four stages do exist. Here, we take two fourth-order SSP Runge-Kutta methods as examples. The first one is the optimal five-stage fourth-order Runge-Kutta method, which is developed in [22, 32] and proved to be optimal in [29]. The second one is the ten-stage fourth-order method by Ketcheson in [21], which has excellent SSP coefficient, low storage and simple rational coefficients.

With the optimal five-stage fourth-order Runge-Kutta method, the time discretization

for (2.3) is presented in the Shu-Osher form [31]:

$$u^{(1)} = e^{-\frac{\mu}{\varepsilon}c_1\Delta t}\left(\alpha_{1,0}u^n + \beta_{1,0}\Delta t f(u^n) + \beta_{1,0}\frac{\Delta t}{\varepsilon}(s(u^n)+\mu u^n)\right), \tag{2.25a}$$

$$u^{(2)} = e^{-\frac{\mu}{\varepsilon}c_2\Delta t}\alpha_{2,0}u^n + e^{-\frac{\mu}{\varepsilon}(c_2-c_1)\Delta t}\left(\alpha_{2,1}u^{(1)} + \beta_{2,1}\Delta t f(u^{(1)}) + \beta_{2,1}\frac{\Delta t}{\varepsilon}(s(u^{(1)})+\mu u^{(1)})\right), \tag{2.25b}$$

$$u^{(3)} = e^{-\frac{\mu}{\varepsilon}c_3\Delta t}\alpha_{3,0}u^n + e^{-\frac{\mu}{\varepsilon}(c_3-c_2)\Delta t}\left(\alpha_{3,2}u^{(2)} + \beta_{3,2}\Delta t f(u^{(2)}) + \beta_{3,2}\frac{\Delta t}{\varepsilon}(s(u^{(2)})+\mu u^{(2)})\right), \tag{2.25c}$$

$$u^{(4)} = e^{-\frac{\mu}{\varepsilon}c_4\Delta t}\alpha_{4,0}u^n + e^{-\frac{\mu}{\varepsilon}(c_4-c_3)\Delta t}\left(\alpha_{4,3}u^{(3)} + \beta_{4,3}\Delta t f(u^{(3)}) + \beta_{4,3}\frac{\Delta t}{\varepsilon}(s(u^{(3)})+\mu u^{(3)})\right), \tag{2.25d}$$

$$\begin{aligned} u^{n+1} =& e^{-\frac{\mu}{\varepsilon}(c_5-c_2)\Delta t}\alpha_{5,2}u^{(2)} + e^{-\frac{\mu}{\varepsilon}(c_5-c_3)\Delta t}\left(\alpha_{5,3}u^{(3)} + \beta_{5,3}\Delta t f(u^{(3)}) + \beta_{5,3}\frac{\Delta t}{\varepsilon}(s(u^{(3)})+\mu u^{(3)})\right) \\ &+ e^{-\frac{\mu}{\varepsilon}(c_5-c_4)\Delta t}\left(\alpha_{5,4}u^{(4)} + \beta_{5,4}\Delta t f(u^{(4)}) + \beta_{5,4}\frac{\Delta t}{\varepsilon}(s(u^{(4)})+\mu u^{(4)})\right), \end{aligned} \tag{2.25e}$$

with

$$\begin{aligned} \alpha_{1,0} =& 1, \\ \alpha_{2,0} =& 0.444370493651235, \quad \alpha_{2,1} = 0.555629506348765, \\ \alpha_{3,0} =& 0.620101851488403, \quad \alpha_{3,2} = 0.379898148511597, \\ \alpha_{4,0} =& 0.178079954393132, \quad \alpha_{4,3} = 0.821920045606868, \\ \alpha_{5,2} =& 0.517231671970585, \quad \alpha_{5,3} = 0.096059710526147, \quad \alpha_{5,4} = 0.386708617503269, \end{aligned}$$

and

$$\begin{aligned} \beta_{1,0} =& 0.391752226571890, \quad \beta_{2,1} = 0.368410593050371, \quad \beta_{3,2} = 0.251891774271694, \\ \beta_{4,3} =& 0.544974750228521, \quad \beta_{5,3} = 0.063692468666290, \quad \beta_{5,4} = 0.226007483236906, \end{aligned}$$

and

$$\begin{aligned} c_1 =& 0.391752226571890, \quad c_2 = 0.586079689311540, \quad c_3 = 0.474542363121400, \\ c_4 =& 0.935010630967653, \quad c_5 = 1. \end{aligned}$$

The bound-preserving property is easily shown in the same approach:

Proposition 2.10. *For the five-stage fourth-order exponential RK scheme* (2.25)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*,* $\mu \ge \sup_{0\le u\le 1.27332M}(-s'(u))$*, then*

$$0 \le u^{n+1} \le e^{-z}(1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \gamma z^5)M \le M,$$

with $z = \frac{\mu}{\varepsilon}\Delta t$ *and* $\gamma \approx 0.004477718303076 < \frac{1}{120}$*. The conclusion is similar for the negative value of* u^n*.*

Now we replace exponential functions by polynomials and construct modified exponential RK methods. Set the polynomial of degree 5 as

$$p^{(5)}(z) := 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \gamma z^5, \tag{2.26}$$

where the parameter γ is the same with that in Proposition 2.10. Then we replace $e^{-c_i z}$ by $1/p^{(5)}(c_i z)$ for $i = 1, 2, 3, 4$ and $e^{-(c_m - c_n)z} = e^{c_n z}/e^{c_m z}$ by $p^{(5)}(c_n z)/p^{(5)}(c_m z)$ for $(m, n) = (2, 1), (3, 2), (4, 3), (5, 2), (5, 3), (5, 4)$ in (2.25) and deduce the schemes:

$$u^{(1)} = \frac{1}{p^{(5)}(c_1 z)}\left(\alpha_{1,0}u^n + \beta_{1,0}\Delta t f(u^n) + \beta_{1,0}\frac{\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right), \tag{2.27a}$$

$$u^{(2)} = \frac{1}{p^{(5)}(c_2 z)}\alpha_{2,0}u^n + \frac{p^{(5)}(c_1 z)}{p^{(5)}(c_2 z)}\left(\alpha_{2,1}u^{(1)} + \beta_{2,1}\Delta t f(u^{(1)}) + \beta_{2,1}\frac{\Delta t}{\varepsilon}(s(u^{(1)}) + \mu u^{(1)})\right), \tag{2.27b}$$

$$u^{(3)} = \frac{1}{p^{(5)}(c_3 z)}\alpha_{3,0}u^n + \frac{p^{(5)}(c_2 z)}{p^{(5)}(c_3 z)}\left(\alpha_{3,2}u^{(2)} + \beta_{3,2}\Delta t f(u^{(2)}) + \beta_{3,2}\frac{\Delta t}{\varepsilon}(s(u^{(2)}) + \mu u^{(2)})\right), \tag{2.27c}$$

$$u^{(4)} = \frac{1}{p^{(5)}(c_4 z)}\alpha_{4,0}u^n + \frac{p^{(5)}(c_3 z)}{p^{(5)}(c_4 z)}\left(\alpha_{4,3}u^{(3)} + \beta_{4,3}\Delta t f(u^{(3)}) + \beta_{4,3}\frac{\Delta t}{\varepsilon}(s(u^{(3)}) + \mu u^{(3)})\right), \tag{2.27d}$$

$$\begin{aligned} u^{n+1} =& \frac{p^{(5)}(c_2 z)}{p^{(5)}(c_5 z)}\alpha_{5,2}u^{(2)} + \frac{p^{(5)}(c_3 z)}{p^{(5)}(c_5 z)}\left(\alpha_{5,3}u^{(3)} + \beta_{5,3}\Delta t f(u^{(3)}) + \beta_{5,3}\frac{\Delta t}{\varepsilon}(s(u^{(3)}) + \mu u^{(3)})\right) \\ &+ \frac{p^{(5)}(c_4 z)}{p^{(5)}(c_5 z)}\left(\alpha_{5,4}u^{(4)} + \beta_{5,4}\Delta t f(u^{(4)}) + \beta_{5,4}\frac{\Delta t}{\varepsilon}(s(u^{(4)}) + \mu u^{(4)})\right). \end{aligned} \tag{2.27e}$$

The bound-preserving property of (2.27) is presented in the following and the proof is omitted.

Proposition 2.11. *For the five-stage fourth-order modified exponential RK scheme* (2.27), *if* $0 \le u^n \le M$, *and* $0 < \Delta t \le \Delta t_E$, $\mu \ge \sup_{0\le u\le 1.30453M}(-s'(u))$, *then*

$$0 \le u^{n+1} \le M,$$

The conclusion is similar for the negative value of u^n.

Remark 2.12. In constructing the fourth-order modified exponential RK method, we could also replace exponential functions in (2.25) by polynomials

$$p^{(5)}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{120}, \tag{2.28}$$

which is exactly the Taylor expansion series up to degree 5 for e^z at $z = 0$. In this case, the conclusion in Proposition 2.11 will be replaced by

$$0 \le u^{n+1} \le \frac{1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \gamma z^5}{1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{120}} M \le M.$$

The bound-preserving property also holds. However, this scheme behaves not well when $\varepsilon \ll \Delta t$ with inconsistent initial values. The convergence order is not uniform as the mesh refines.

Next, we construct exponential RK scheme based on the ten-stage fourth-order RK method in [21]:

$$u^{(1)} = e^{-\frac{\mu}{\varepsilon}c_1\Delta t}\left(u^n + \frac{1}{6}\Delta t(f(u^n) + \frac{1}{\varepsilon}(s(u^n) + \mu u^n))\right), \tag{2.29a}$$

$$u^{(i+1)} = e^{-\frac{\mu}{\varepsilon}(c_{i+1}-c_i)\Delta t}\left(u^{(i)} + \frac{1}{6}\Delta t(f(u^{(i)}) + \frac{1}{\varepsilon}(s(u^{(i)}) + \mu u^{(i)}))\right), \quad i = 1, 2, 3, \tag{2.29b}$$

$$u^{(5)} = \frac{3}{5}e^{-\frac{\mu}{\varepsilon}c_5\Delta t}u^n + \frac{2}{5}e^{-\frac{\mu}{\varepsilon}(c_5-c_4)\Delta t}\left(u^{(4)} + \frac{1}{6}\Delta t(f(u^{(4)}) + \frac{1}{\varepsilon}(s(u^{(4)}) + \mu u^{(4)}))\right), \tag{2.29c}$$

$$u^{(i+1)} = e^{-\frac{\mu}{\varepsilon}(c_{i+1}-c_i)\Delta t}\left(u^{(i)} + \frac{1}{6}\Delta t(f(u^{(i)}) + \frac{1}{\varepsilon}(s(u^{(i)}) + \mu u^{(i)}))\right), \quad i = 5, 6, 7, 8, \tag{2.29d}$$

$$\begin{aligned} u^{n+1} = {} & \frac{1}{25}e^{-\frac{\mu}{\varepsilon}c_{10}\Delta t}u^n + \frac{9}{25}e^{-\frac{\mu}{\varepsilon}(c_{10}-c_4)\Delta t}\left(u^{(4)} + \frac{1}{6}\Delta t(f(u^{(4)}) + \frac{1}{\varepsilon}(s(u^{(4)}) + \mu u^{(4)}))\right) \\ & + \frac{3}{5}e^{-\frac{\mu}{\varepsilon}(c_{10}-c_9)\Delta t}\left(u^{(9)} + \frac{1}{6}\Delta t(f(u^{(9)}) + \frac{1}{\varepsilon}(s(u^{(9)}) + \mu u^{(9)}))\right), \end{aligned} \tag{2.29e}$$

with

$$c_1 = \frac{1}{6}, \quad c_2 = c_5 = \frac{1}{3}, \quad c_3 = c_6 = \frac{1}{2}, \quad c_4 = c_7 = \frac{2}{3}, \quad c_8 = \frac{5}{6}, \quad c_9 = c_{10} = 1.$$

The corresponding modified scheme reads as:

$$u^{(1)} = \frac{1}{p^{(10)}(c_1 z)} \left(u^n + \frac{1}{6}\Delta t(f(u^n) + \frac{1}{\varepsilon}(s(u^n) + \mu u^n)) \right), \tag{2.30a}$$

$$u^{(i+1)} = \frac{p^{(10)}(c_i z)}{p^{(10)}(c_{i+1} z)} \left(u^{(i)} + \frac{1}{6}\Delta t(f(u^{(i)}) + \frac{1}{\varepsilon}(s(u^{(i)}) + \mu u^{(i)})) \right), \quad i = 1, 2, 3, \tag{2.30b}$$

$$u^{(5)} = \frac{3}{5}\frac{1}{p^{(10)}(c_5 z)} u^n + \frac{2}{5}\frac{p^{(10)}(c_4 z)}{p^{(10)}(c_5 z)} \left(u^{(4)} + \frac{1}{6}\Delta t(f(u^{(4)}) + \frac{1}{\varepsilon}(s(u^{(4)}) + \mu u^{(4)})) \right), \tag{2.30c}$$

$$u^{(i+1)} = \frac{p^{(10)}(c_i z)}{p^{(10)}(c_{i+1} z)} \left(u^{(i)} + \frac{1}{6}\Delta t(f(u^{(i)}) + \frac{1}{\varepsilon}(s(u^{(i)}) + \mu u^{(i)})) \right), \quad i = 5, 6, 7, 8, \tag{2.30d}$$

$$\begin{aligned} u^{n+1} = {} & \frac{1}{25}\frac{1}{p^{(10)}(c_{10} z)} u^n + \frac{9}{25}\frac{p^{(10)}(c_4 z)}{p^{(10)}(c_{10} z)} \left(u^{(4)} + \frac{1}{6}\Delta t(f(u^{(4)}) + \frac{1}{\varepsilon}(s(u^{(4)}) + \mu u^{(4)})) \right) \\ & + \frac{3}{5}\frac{p^{(10)}(c_9 z)}{p^{(10)}(c_{10} z)} \left(u^{(9)} + \frac{1}{6}\Delta t(f(u^{(9)}) + \frac{1}{\varepsilon}(s(u^{(9)}) + \mu u^{(9)})) \right), \end{aligned} \tag{2.30e}$$

with $z = \frac{\mu}{\varepsilon}\Delta t$ and $p^{(10)}(z)$ the stability function:

$$p^{(10)}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{17z^5}{2160} + \frac{7z^6}{6480} + \frac{z^7}{9720} + \frac{z^8}{155520} + \frac{z^9}{4199040} + \frac{z^{10}}{251942400}.$$

The bound-preserving properties of the ten-stage fourth-order exponential RK scheme (2.29) and the modified exponential RK scheme (2.30) are presented in the following and the proof is omitted as well.

Proposition 2.13. *For the ten-stage fourth-order exponential RK scheme* (2.29)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*,* $\mu \ge \sup_{0 \le u \le 1.976M}(-s'(u))$*, then*

$$0 \le u^{n+1} \le M.$$

For the ten-stage fourth-order modified exponential RK scheme (2.30)*, if* $0 \le u^n \le M$*, and* $0 < \Delta t \le \Delta t_E$*,* $\mu \ge \sup_{0 \le u \le 2.0584M}(-s'(u))$*, then*

$$0 \le u^{n+1} \le M.$$

The conclusion is similar for the negative value of u^n*.*

Before concluding this part, we give several remarks on the relations between our work and the existing work in the literature.

Remark 2.14. We remark that the modified RK methods developed above are only AP in the weak sense and thus the convergence orders in the unresolved region are nontrivial to analyze theoretically. We will test the accuracy in the numerical examples. They all have around first-order of accuracy with $\varepsilon \ll \Delta t$.

Remark 2.15. In [25], Li and Pareschi developed a class of exponential RK methods with positivity-preserving properties for the Boltzmann equations. We remark that our methods are different from theirs. As pointed out in Remark 4 in [25], for preserving the positivity of the distribution function, they require that $c_j \leq c_i$ for $j < i$ where c_j denotes the usual time nodes in RK methods. However, the classical third-order SSP method [31] does not satisfy this restriction.

Remark 2.16. Very recently, in [12], based on explicit SSP RK methods with non-decreasing abscissas, Gottlieb et al. constructed a class of explicit SSP integrating factor RK methods for problems with a linear component that is stiff and a nonlinear component that is not. If assuming that the abscissas are non-decreasing and only considering hyperbolic equations with linear source, our schemes before modification are exactly the same with theirs.

2.3 Multi-step method

In this section, we use the SSP multi-step methods (see Table 5.1 in [14]) to discretize (2.3) and obtain the exponential multi-step methods. The second-order one is given as

$$u^{n+1} = \frac{3}{4}e^{-\frac{\mu}{\varepsilon}\Delta t}\left(u^n + 2\Delta t f(u^n) + \frac{2\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right) + \frac{1}{4}e^{-\frac{3\mu}{\varepsilon}\Delta t}u^{n-2}, \tag{2.31}$$

the third-order one:

$$u^{n+1} = \frac{16}{27}e^{-\frac{\mu}{\varepsilon}\Delta t}\left((u^n + 3\Delta t f(u^n)) + \frac{3\Delta t}{\varepsilon}(s(u^n) + \mu u^n)\right)$$

$$+\frac{11}{27}e^{-\frac{4\mu}{\varepsilon}\Delta t}\left((u^{n-3}+\frac{12}{11}\Delta t f(u^{n-3}))+\frac{12\Delta t}{11\varepsilon}(s(u^{n-3})+\mu u^{n-3})\right), \tag{2.32}$$

and the fourth-order one:

$$u^{n+1}=\sum_{i=1}^{5}e^{-i\frac{\mu}{\varepsilon}\Delta t}\left(\alpha_i u^{n+1-i}+\Delta t\beta_i f(u^{n+1-i})+\beta_i\frac{\Delta t}{\varepsilon}(s(u^{n+1-i})+\mu u^{n+1-i})\right) \tag{2.33}$$

with

$$\alpha_1=\frac{1557}{32000},\quad \alpha_2=\frac{1}{32000},\quad \alpha_3=\frac{1}{120},\quad \alpha_4=\frac{2063}{48000},\quad \alpha_5=\frac{9}{10}$$

and

$$\beta_1=\frac{5323561}{2304000},\quad \beta_2=\frac{2659}{2304000},\quad \beta_3=\frac{904987}{2304000},\quad \beta_4=\frac{1567579}{768000},\quad \beta_5=0.$$

Proposition 2.17. *1. For the second-order exponential multi-step scheme* (2.31)*, if* $0\le u^{n-2},u^n\le M$*, and* $0<\Delta t\le\Delta t_E$*,* $\mu\ge\sup_{0\le u\le M}(-s'(u))$*, then*

$$0\le u^{n+1}\le\left(\frac{3}{4}e^{-\frac{\mu}{\varepsilon}\Delta t}(1+\frac{2\mu}{\varepsilon}\Delta t)+\frac{1}{4}e^{-\frac{3\mu}{\varepsilon}\Delta t}\right)M\le M, \tag{2.34}$$

2. For the third-order exponential multi-step scheme (2.32)*, if* $0\le u^{n-3},u^n\le M$*, and* $0<\Delta t\le\Delta t_E$*,* $\mu\ge\sup_{0\le u\le M}(-s'(u))$*, then*

$$0\le u^{n+1}\le\left(\frac{16}{27}e^{-z}(1+3z)+\frac{11}{27}e^{-4z}(1+\frac{12}{11}z)\right)M\le M, \tag{2.35}$$

with $z=\frac{\mu}{\varepsilon}\Delta t>0$*.*

3. For the fourth-order exponential multi-step scheme (2.33)*, if* $0\le u^{n-i}\le M$*, for* $i=0,1,\cdots,4$ *and* $0<\Delta t\le\Delta t_E$*,* $\mu\ge\sup_{0\le u\le M}(-s'(u))$*, then*

$$0\le u^{n+1}\le M, \tag{2.36}$$

with $z=\frac{\mu}{\varepsilon}\Delta t>0$*.*

The conclusion is similar for the negative value of u^n*.*

One important issue for the multi-step methods is the initialization. One usually uses the RK methods to start the multi-step methods [11]. However, the explicit RK methods

are not suitable for the stiff problems, and in the class of implicit methods only the implicit Euler backward methods has the desired bound-preserving property.

For the second-order exponential multi-step method (2.31), the first-order implicit-explicit (IMEX) method is enough to keep the accuracy and also has bound-preserving property:

$$u^{n+1} = u^n + \Delta t f(u^n) + \frac{\Delta t}{\varepsilon} s(u^{n+1}). \tag{2.37}$$

For the higher order exponential multi-step methods (2.32) and (2.33), if $\varepsilon \gg \Delta t$ or $\varepsilon \ll \Delta t$ with well-prepared initial value, we can use our (modified) exponential RK method designed in the last subsection. If $\varepsilon \ll \Delta t$ with an initial layer, our modified RK methods do not work well. The reason is that they are only weak AP and thus the numerical solutions in the first several time steps are not accurate. Hence, the multi-step methods shown above seem to be difficult to use for problems with initial layers. We will therefore mainly focus on the exponential RK methods in the following sections.

3 Bound-preserving DG methods

In this section, we couple the modified exponential RK methods with the semi-discrete DG scheme and discuss the bound-preserving property. We take the one dimensional case as an example and some remarks on two dimensional case and the finite volume ENO/WENO schemes will be given later.

We first discretize (1.1) in space following [8]. Denote the computational domain by $I \subset \mathbb{R}$. For each partition of the interval I, $\{x_{j+\frac{1}{2}}\}_{j=0}^N$, we set $I_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ for $j = 1, \cdots, N$. For simplicity, we assume that the partition is uniform with mesh size h. Define the finite element space:

$$V_h^k = \{v \in L^1(I) : v|_{I_j} \in \mathbb{P}^k(I_j), j = 1, \cdots, N\}, \tag{3.1}$$

where $\mathbb{P}^k(I_j)$ denotes the space of polynomials in I_j of degree at most $k \geq 0$. The semi-discrete DG scheme for (1.1) reads as

$$\int_{I_j} (u_h)_t v_h - \int_{I_j} f(u_h)(v_h)_x + \hat{f}_{j+\frac{1}{2}}(v_h)^-_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}(v_h)^+_{j-\frac{1}{2}} = \frac{1}{\varepsilon} \int_{I_j} s(u_h) v_h \tag{3.2}$$

for all j and all $v_h \in V_h^k$. Here $\hat{f}_{j+\frac{1}{2}} = \hat{f}(u^-_{j+\frac{1}{2}}, u^+_{j+\frac{1}{2}})$ is the numerical flux, which is chosen to be a Lipschitz continuous monotone flux, i.e., $\hat{f}(\cdot,\cdot)$ is nondecreasing in its first argument and nonincreasing in its second argument. For instance, the global Lax-Friedrichs flux is defined by

$$\hat{f}(u,v) = \frac{1}{2}(f(u) + f(v) - a(v-u)), \quad a = \max|f'(u)|. \tag{3.3}$$

We approximate the integral of the source term in (3.2) by quadrature rules:

$$\int_{I_j} g(x)\mathrm{d}x \approx \Delta x \sum_{\beta=1}^{L} \omega_\beta g(x_j^\beta). \tag{3.4}$$

Here $\omega_\beta \geq 0$ are integration weights and x_j^β are integration points. We denote the set of these integration points by

$$S_j^{SI} = \{x_j^\beta, \quad \beta = 1, \cdots, L\}. \tag{3.5}$$

Note that, to keep the accuracy, the quadrature rules are required to be exact for polynomials up to degree $2k$ [7, 17]. To present our scheme more clearly, we keep using the weak formulation and do not introduce basis functions. We introduce the notation [41]

$$H_j(u,v) := \int_{I_j} f(u)(v)_x - \hat{f}_{j+\frac{1}{2}}(v)^-_{j+\frac{1}{2}} + \hat{f}_{j-\frac{1}{2}}(v)^+_{j-\frac{1}{2}}, \tag{3.6}$$

and use $(\cdot,\cdot)_{I_j}$ to denote the inner product on $L^2(I_j)$ and $\langle\cdot,\cdot\rangle_{I_j}$ to denote the quadrature product in I_j which is defined as

$$\langle u, v\rangle_{I_j} = \Delta x \sum_{\beta=1}^{L} \omega_\beta u(x_j^\beta) v(x_j^\beta). \tag{3.7}$$

Then the semi-discrete DG scheme with source term approximated by quadrature rules can be rewritten as

$$((u_h)_t, v_h)_{I_j} = H_j(u_h, v_h) + \frac{1}{\varepsilon}\langle s(u_h), v_h\rangle_{I_j}, \tag{3.8}$$

which is reformulated as

$$\begin{aligned}((u_h)_t, v_h)_{I_j} =& H_j(u_h, v_h) + \frac{1}{\varepsilon}\langle s(u_h) + \mu u_h, v_h\rangle_{I_j} - \frac{\mu}{\varepsilon}\langle u_h, v_h\rangle_{I_j}, \\ =& H_j(u_h, v_h) + \frac{1}{\varepsilon}\langle s(u_h) + \mu u_h, v_h\rangle_{I_j} - \frac{\mu}{\varepsilon}(u_h, v_h)_{I_j},\end{aligned}$$

and is equivalent to

$$((e^{\frac{\mu}{\varepsilon}t}u_h)_t, v_h)_{I_j} = e^{\frac{\mu}{\varepsilon}t}\left(H_j(u_h, v_h) + \frac{1}{\varepsilon}\left\langle s(u_h) + \mu u_h, v_h\right\rangle_{I_j}\right). \tag{3.9}$$

We apply the modified exponential RK methods developed in section 2 to (3.9). For simplicity, we take the second-order method (2.13) as an example and deduce:

$$(u_h^{(1)}, v_h)_{I_j} = \frac{1}{1+\frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}\left((u_h^n, v_h)_{I_j} + \Delta t H_j(u_h^n, v_h) + \frac{\Delta t}{\varepsilon}\left\langle s(u_h^n) + \mu u_h^n, v_h\right\rangle_{I_j}\right), \tag{3.10a}$$

$$\begin{aligned}(u_h^{n+1}, v_h)_{I_j} =& \frac{1}{2(1+\frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2)}(u_h^n, v_h)_{I_j} \\ &+ \frac{1}{2}\left((u_h^{(1)}, v_h)_{I_j} + \Delta t H_j(u_h^{(1)}, v_h) + \frac{\Delta t}{\varepsilon}\left\langle s(u_h^{(1)}) + \mu u_h^{(1)}, v_h\right\rangle_{I_j}\right).\end{aligned} \tag{3.10b}$$

In the following, we discuss the bound-preserving property of the scheme (3.10). Setting the test function $v_h \equiv 1$ results in

$$\bar{u}_j^{(1)} = \frac{1}{1+\frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}\left((\bar{u}_j^n - \lambda(\hat{f}_{j+\frac{1}{2}}^n - \hat{f}_{j-\frac{1}{2}}^n)) + \frac{\Delta t}{\varepsilon}\sum_{\beta=1}^{L}\omega_\beta(s(u^n(x_j^\beta)) + \mu u^n(x_j^\beta))\right), \tag{3.11a}$$

$$\bar{u}_j^{n+1} = \frac{1}{2(1+\frac{\mu}{\varepsilon}\Delta t + \frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2)}\bar{u}_j^n + \frac{1}{2}\left(\bar{u}_j^{(1)} - \lambda(\hat{f}_{j+\frac{1}{2}}^{(1)} - \hat{f}_{j-\frac{1}{2}}^{(1)})\right) + \frac{\Delta t}{2\varepsilon}\sum_{\beta=1}^{L}\omega_\beta(s(u^{(1)}(x_j^\beta)) + \mu u^{(1)}(x_j^\beta)). \tag{3.11b}$$

First we briefly review the techniques introduced in [43] for controlling the bound of the convection part $(\bar{u}_j - \lambda(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}))$. Here we only list the main results and refer readers to [43] and [45] for details. Choose N to be the smallest integer satisfying $2N-3 \geq k$ and consider the N-point Legendre Gauss-Lobatto quadrature rule on the interval $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. We denote these quadrature points on I_j as

$$S_j^{GL} = \{x_{j-\frac{1}{2}} = \hat{x}_j^1, \hat{x}_j^2, \cdots, \hat{x}_j^{N-1}, \hat{x}_j^N = x_{j+\frac{1}{2}}\}. \tag{3.12}$$

Define $\hat{v}_\alpha = u_h(\hat{x}_j^\alpha)$ for $\alpha = 1, \cdots, N$, and let $\hat{\omega}_\alpha$ be the quadrature weights for the interval $[-\frac{1}{2}, \frac{1}{2}]$ such that $\sum_{\alpha=1}^N \hat{\omega}_\alpha = 1$.

Assume that $u_{j-\frac{1}{2}}^-$, $u_{j+\frac{1}{2}}^+$ and $\hat{v}_\alpha$ $(\alpha = 1, \cdots, N)$ are all in the range $[m, M]$. By splitting the cell average $\bar{u}_j$ into a convex combination of values at quadrature points and then writing

the convection term $(\bar{u}_j - \lambda(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}))$ into a convex combination of formally monotone schemes, under the condition $\lambda a \leq \hat{\omega}_1$ with a being the constant in the global Lax-Friedrichs flux defined in (3.3), we have the bound for the convective part:

$$m \leq \bar{u}_j - \lambda(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}) \leq M. \tag{3.13}$$

To ensure that $u^-_{j-\frac{1}{2}}$, $u^+_{j+\frac{1}{2}}$ and $\hat{v}_\alpha$ $(\alpha = 1, \cdots, N)$ are all in the range $[m, M]$, a scaling limiter is applied to keep the accuracy and conservativity [27]: The polynomial $p_j(x)$ (i.e. the polynomial on the interval I_j in the DG scheme) is modified to $\tilde{p}_j(x)$ which is calculated by

$$\tilde{p}_j(x) = \theta(p_j(x) - \bar{u}^n_j) + \bar{u}^n_j, \quad \theta = \min\{\left|\frac{M - \bar{u}^n_j}{M_j - \bar{u}^n_j}\right|, \left|\frac{m - \bar{u}^n_j}{m_j - \bar{u}^n_j}\right|, 1\}, \tag{3.14}$$

with

$$M_j = \max_{x \in S^{GL}_j} p_j(x), \quad m_j = \min_{x \in S^{GL}_j} p_j(x). \tag{3.15}$$

Now we analyze the bound-preserving property of our scheme (3.10). Assume that $m \leq \bar{u}^n_j \leq M$ for any j with $m \leq 0$ and $M \geq 0$. In the first stage (3.10a), by applying the limiter (3.14) with m_j and M_j replaced by

$$M_j = \max_{x \in S^{GL}_j \cup S^{SI}_j} p_j(x), \quad m_j = \min_{x \in S^{GL}_j \cup S^{SI}_j} p_j(x), \tag{3.16}$$

w have the bound for the convection term

$$m \leq \bar{u}^n_j - \lambda(\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}}) \leq M, \tag{3.17}$$

and the values at all quadrature points

$$m \leq u^n_h(x^\beta_j) \leq M. \tag{3.18}$$

Take $\mu = \sup_{m \leq u \leq M}(-s'(u))$. Then $s(u) + \mu u$ as a function of u is non-decreasing for $m \leq u \leq M$ and thus we have

$$\mu m \leq s(m) + \mu m \leq s(u^n_h(x^\beta_j)) + \mu u^n_h(x^\beta_j) \leq s(M) + \mu M \leq \mu M. \tag{3.19}$$

Remembering that $\sum_{\beta=1}^{L}\omega_\beta = 1$ and $\omega_\beta \geq 0$, by (3.11a) we have

$$\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}m \leq \bar{u}_j^{(1)} \leq \frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}M. \tag{3.20}$$

Now for the second stage (3.10b), we can get the bound of $\bar{u}_j^{n+1}$ in the same procedure, except that the constant m and M in the limiter should be replaced by $\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}m$ and $\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}M$, respectively.

Proposition 3.1. *Consider the modified second-order RK method coupled with the semi-discrete DG scheme* (3.10)*. Assume that the quadrature weights of the integration for the source term are non-negative and the CFL condition* $\lambda a \leq \hat{w}_1$*. In the first stage* (3.10a)*, if* $u_{j-\frac{1}{2}}^-$*,* $u_{j+\frac{1}{2}}^+$*,* $u^n(\hat{x}_j^\alpha)$ $(\alpha = 1, \cdots, N)$ *and* $u^n(x_j^\beta)$ $(\beta = 1, \cdots, L)$ *are all in the range* $[m, M]$ *with* $m \leq 0$ *and* $M \geq 0$*, then*

$$\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}m \leq \bar{u}_j^{(1)} \leq \frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}M. \tag{3.21}$$

In the second stage (3.10b)*, if* $u_{j-\frac{1}{2}}^-$*,* $u_{j+\frac{1}{2}}^+$*,* $u^{(1)}(\hat{x}_j^\alpha)$ $(\alpha = 1, \cdots, N)$ *and* $u^n(x_j^\beta)$ $(\beta = 1, \cdots, L)$ *are all in the range* $[\frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}m, \frac{1+\frac{\mu}{\varepsilon}\Delta t}{1+\frac{\mu}{\varepsilon}\Delta t+\frac{1}{2}(\frac{\mu}{\varepsilon}\Delta t)^2}M]$*, then*

$$m \leq \bar{u}_j^{n+1} \leq M. \tag{3.22}$$

The bound of these point values is guaranteed by applying the scaling limiter (3.14)*.*

Remark 3.2. Note that in the original paper on the maximum-principle-preserving schemes for scalar conservation laws, the constant m and M, i.e., the lower and upper bound of the solutions, are naturally taken to be the minimum and maximum of the given initial values [43]. However, in our case, the magnitudes of the solutions will decay with time due to the dissipation of the source term. Therefore, the constants m and M should vary with time and be taken as the minimum and maximum values of numerical solutions evaluated as polynomials. If the constants m and M are fixed to be the minimum and maximum values of the initial data, then the parameter μ will be fixed. In fact, with fixed μ, the scheme does not work well in the stiff case with inconsistent initial values. We remark that evaluating extrema of polynomials will increase the computational cost, especially in the 2D case.

Remark 3.3. Other types of spatial discretization such as the finite volume ENO/WENO scheme could be coupled with the modified exponential RK methods. Moreover, our time integrators could be applied to problems in the 2D case. The bound-preserving property could be derived in the same approach. The derivations are omitted here for saving space. We refer readers to the review paper [45] for details.

4 Numerical results for the ODE solver

In this part, we test the accuracy of our (modified) exponential RK and multi-step methods on the initial value problem of the scalar ODE (2.1). In the following part, for the (modified) exponential fourth-order RK method, we refer to the five-stage fourth-order one.

4.1 Non-stiff case

We first test the accuracy of our methods for the non-stiff case. In the ODE (2.1), we take $f(u) = -u^2$, $s(u) = -u^3$, $\varepsilon = 1$, the initial value $u_0 = 1$ and the final time $t = 1$.

The errors at $t = 1$ for the (modified) exponential RK methods are listed in Table 4.1. In the table, we denote errors for the exponential RK methods and the modified exponential RK methods by "error (exp)" and "error (modify)", respectively. In the computation, the parameter $\mu = \mu(u^n)$ is set to be the lower bound in the propositions. For example, $\mu = \sup_{0\le u\le 1.13652u^n}(-s'(u)) = 3(1.13652u^n)^2$ for the third-order modified exponential RK method. The table shows clear convergence orders except that the third-order modified RK method has some "superconvergence". Moreover, the errors for the modified RK methods are smaller than those for the exponential RK methods.

We also discuss the impact of the values of μ on the numerical behaviour presented in Table 4.2. Here we use the second-order (modified) exponential RK method. Note that in the propositions, for preserving the bound of the numerical solution, we only require the lower bound of the parameter μ. From the table, we can observe that, if μ is much larger than the lower bound in the proposition, there will be some order degeneration phenomenon.

Table 4.1: Error table for the (modified) exponential Runge-Kutta method for solving the scalar ODE in the non-stiff case.

	N	error (exp)	order	error (modify)	order
first-order	20	7.37e-03	-	1.09e-04	-
	40	3.63e-03	1.024	1.18e-04	-0.108
	80	1.80e-03	1.012	7.40e-05	0.667
	160	8.96e-04	1.006	4.08e-05	0.861
	320	4.47e-04	1.003	2.13e-05	0.936
second-order	20	8.13e-05	-	2.86e-04	-
	40	2.20e-05	1.887	6.98e-05	2.038
	80	5.66e-06	1.956	1.72e-05	2.021
	160	1.44e-06	1.981	4.27e-06	2.011
	320	3.61e-07	1.991	1.06e-06	2.005
third-order	20	1.23e-05	-	1.60e-06	-
	40	1.46e-06	3.073	9.85e-08	4.020
	80	1.77e-07	3.038	6.29e-09	3.969
	160	2.19e-08	3.020	4.21e-10	3.899
	320	2.71e-09	3.010	3.03e-11	3.797
fourth-order	20	2.90e-07	-	2.60e-07	-
	40	1.59e-08	4.191	1.57e-08	4.045
	80	9.23e-10	4.104	9.57e-10	4.038
	160	5.55e-11	4.055	5.88e-11	4.024
	320	3.33e-12	4.059	3.57e-12	4.041

Moreover, the modified RK method is better than the exponential RK method in the aspects of magnitudes of errors and convergence orders.

Table 4.2: Error table for different values of μ with the second order modified exponential RK method for solving the scalar ODE in the non-stiff case.

	N	error (exp)	order	error (modify)	order
$\mu = 3(u^n)^2$	20	8.13e-05	-	2.86e-04	-
	40	2.20e-05	1.887	6.98e-05	2.038
	80	5.66e-06	1.956	1.72e-05	2.021
	160	1.44e-06	1.981	4.27e-06	2.011
	320	3.61e-07	1.991	1.06e-06	2.005
$\mu = 3(3u^n)^2$	20	3.73e-02	-	3.51e-02	-
	40	1.40e-02	1.416	1.01e-02	1.798
	80	4.29e-03	1.703	2.78e-03	1.858
	160	1.17e-03	1.880	7.40e-04	1.912
	320	3.01e-04	1.954	1.92e-04	1.950
$\mu = 3(5u^n)^2$	20	1.85e-01	-	1.87e-01	-
	40	8.71e-02	1.084	6.27e-02	1.580
	80	4.88e-02	0.836	1.88e-02	1.736
	160	2.18e-02	1.163	5.41e-03	1.799
	320	7.45e-03	1.547	1.48e-03	1.867
$\mu = 3(10u^n)^2$	20	3.87e-01	-	4.84e-01	-
	40	3.67e-01	0.075	3.65e-01	0.408
	80	1.95e-01	0.915	1.87e-01	0.967
	160	1.32e-01	0.561	6.28e-02	1.571
	320	9.16e-02	0.528	1.90e-02	1.721

Next, we move to the exponential multi-step methods. In the non-stiff case, the exponential RK methods are good enough to start the multi-step methods. For example, we use the second-order exponential RK method to initialize the third-order multi-step method. The results are presented in Table 4.3 which shows clear convergence orders.

4.2 Stiff case

We now focus on the numerical behavior of the ODE solver on the stiff case. Set $f(u) = -u^2$, $s(u) = -u^5$ and $\varepsilon = 1 \times 10^{-4}$. Since the linear source term results in exponential decay in time, we take $s(u)$ to be a polynomial of high degree in order to guarantee that u is not that small for $t = O(1)$.

First we discuss the consistent initial value, i.e., $s(u_0) = O(\varepsilon)$. In the numerical examples, we take $u_0 = 0.1$. The errors for the RK methods and the multi-step methods at $t = 1$ are listed in Table 4.4 and Table 4.5, respectively. Here we use the exponential RK methods to start the multi-step methods. They all have full order of accuracy.

For the inconsistent initial value, we take $u_0 = 1$. For the RK methods, the errors are listed in Table 4.6. As expected, the exponential RK methods only output numerical solutions of zero value, which does not converge to the numerical solution with $\varepsilon \ll \Delta t$. The modified RK method all have similar convergence orders of around first-order and similar magnitudes of errors. We remark that the modified RK methods are only AP in the weak sense and thus the convergence orders in the unresolved region are nontrivial to analyze theoretically.

Next, we move to the exponential multi-step methods in the case of $\varepsilon \ll \Delta t$ with the inconsistent initial value. In order to verify the accuracy of our methods without the effects of the initial layer, we use exact solutions in the first several time steps to do initialization. As shown in Table 4.7, they all behave with a first-order convergence order. We also try to use

Table 4.3: Error table for the exponential multi-step methods for solving the scalar ODE in the non-stiff case.

	N	error	order
second-order	20	7.20e-04	-
	40	1.88e-04	1.934
	80	4.83e-05	1.965
	160	1.22e-05	1.982
	320	3.07e-06	1.991
third-order	20	1.44e-04	-
	40	1.97e-05	2.867
	80	2.58e-06	2.928
	160	3.32e-07	2.960
	320	4.21e-08	2.979
fourth-order	20	8.87e-05	-
	40	3.81e-06	4.541
	80	1.62e-07	4.554
	160	8.89e-09	4.190
	320	5.58e-10	3.994

Table 4.4: Error table for the (modified) exponential RK methods for the scalar ODE in the stiff case with consistent initial value.

	N	error (exp)	order	error (modify)	order
first-order	20	2.03e-03	-	6.78e-04	-
	40	1.06e-03	0.945	3.27e-04	1.053
	80	5.41e-04	0.965	1.60e-04	1.028
	160	2.74e-04	0.981	7.94e-05	1.014
	320	1.38e-04	0.990	3.95e-05	1.007
second-order	20	2.87e-05	-	1.61e-04	-
	40	9.14e-06	1.649	4.02e-05	2.005
	80	2.47e-06	1.886	1.00e-05	2.005
	160	6.38e-07	1.954	2.50e-06	2.003
	320	1.62e-07	1.979	6.24e-07	2.001
third-order	20	1.28e-05	-	1.93e-05	-
	40	1.14e-06	3.486	2.93e-06	2.719
	80	1.15e-07	3.311	3.97e-07	2.884
	160	1.27e-08	3.178	5.14e-08	2.949
	320	1.49e-09	3.095	6.53e-09	2.976
fourth-order	20	8.10e-06	-	3.24e-06	-
	40	4.72e-07	4.102	2.21e-07	3.874
	80	2.83e-08	4.061	1.41e-08	3.969
	160	1.73e-09	4.033	8.86e-10	3.993
	320	1.07e-10	4.017	5.55e-11	3.998

Table 4.5: Error table for the multi-step methods for the scalar ODE in the stiff case with consistent initial value (initialization with exponential RK methods).

	N	error	order
second-order	20	5.20e-04	-
	40	1.44e-04	1.853
	80	3.77e-05	1.933
	160	9.63e-06	1.970
	320	2.43e-06	1.986
third-order	20	6.51e-06	-
	40	7.36e-07	3.145
	80	1.02e-07	2.850
	160	1.43e-08	2.834
	320	1.93e-09	2.892
fourth-order	20	1.88e-06	-
	40	7.98e-08	4.555
	80	1.25e-08	2.672
	160	9.68e-10	3.693
	320	6.22e-11	3.962

Table 4.6: Error table for the (modified) exponential RK methods for the scalar ODE in the stiff case with inconsistent initial value.

	N	error (exponent)	order	error (polynomial)	order
first-order	20	6.79e-02	-	1.38e-02	-
	40	6.79e-02	0.000	5.44e-03	1.342
	80	6.79e-02	0.000	2.42e-03	1.168
	160	6.79e-02	0.000	1.12e-03	1.105
	320	6.79e-02	0.000	5.32e-04	1.079
second-order	20	6.79e-02	-	1.29e-02	-
	40	6.79e-02	0.000	4.83e-03	1.419
	80	6.79e-02	0.000	2.04e-03	1.241
	160	6.79e-02	0.000	8.97e-04	1.189
	320	6.79e-02	0.000	3.96e-04	1.178
third-order	20	6.79e-02	-	2.39e-02	-
	40	6.79e-02	0.000	7.13e-03	1.745
	80	6.79e-02	0.000	2.79e-03	1.351
	160	6.79e-02	0.000	1.16e-03	1.265
	320	6.79e-02	0.000	4.88e-04	1.253
fourth-order	20	6.79e-02	-	1.76e-01	-
	40	6.79e-02	0.000	1.69e-02	3.384
	80	6.79e-02	0.000	5.60e-03	1.591
	160	6.79e-02	0.000	2.22e-03	1.331
	320	6.79e-02	0.000	9.21e-04	1.273

the first-order IMEX method to start the second-order multi-step method. The first-order convergence is observed in Table 4.8.

Table 4.7: Error table for the multi-step methods for the scalar ODE in the stiff case with inconsistent initial value (initialization with the exact solution).

	N	error	order
second-order	20	3.00e-04	-
	40	1.84e-04	0.710
	80	1.04e-04	0.824
	160	5.59e-05	0.892
	320	2.92e-05	0.934
third-order	20	1.58e-04	-
	40	8.16e-05	0.958
	80	4.16e-05	0.971
	160	2.10e-05	0.985
	320	1.05e-05	0.995
fourth-order	20	1.88e-05	-
	40	1.19e-06	3.984
	80	6.27e-06	-2.398
	160	4.30e-06	0.544
	320	2.21e-06	0.958

Table 4.8: Error table for the second-order multi-step method for the scalar ODE in the stiff case with inconsistent initial value (initialization with first-order IMEX).

N	error	order
20	6.74e-04	-
40	2.85e-04	1.240
80	1.25e-04	1.186
160	5.65e-05	1.150
320	2.58e-05	1.133

We summarize the numerical results shown above:

i) In the non-stiff case, i.e., $\varepsilon \gg \Delta t$, the (modified) exponential RK and the multi-step methods all have full convergence orders. We also observe that the values of μ should not be set much larger than the lower bound, otherwise there will be order degeneration phenomenon.

ii) In the stiff case with well-prepared initial value, the (modified) exponential RK and

the multi-step methods all have full convergence orders. If the inconsistent initial value is given, the modified RK methods and the multi-step methods initialized with exact solutions all have first-order convergence.

Finally, we remark that the motivation for replacing the exponential functions by polynomials is to treat the stiff problems with inconsistent initial values without resolving the initial layer. However, for the multi-step method, the initialization problem is not easy to attack. Therefore, it seems that the exponential multi-step methods is not easily applicable for the stiff problems, and it is not very meaningful to develop modified multi-step methods following similar approaches for the RK methods. In the numerical examples for the PDEs below, we will not discuss the multi-step methods and will focus on the modified exponential RK methods only.

5 Numerical results for the DG scheme

In this section, we apply the modified exponential RK method to the semi-discrete DG scheme and discuss the accuracy and bound-preserving property of the scheme. In the numerical examples, unless otherwise stated, the CFL number is taken to be 1/3, 1/6 and 1/10 [43] for $k = 1, 2, 3$ where k denotes the polynomials of degree in the finite element space. The numerical flux is taken to be the global Lax-Friedrichs flux. For the numerical tests without explicit analytic solutions, the "exact" solutions are computed via the spectral method with an extremely refined mesh.

5.1 Non-stiff case

Example 5.1 (1D advection equation with source term)**.** We solve the 1D advection equation with source term:

$$\begin{aligned} u_t + u_x =& \frac{1}{\varepsilon} s(u), \quad 0 \le x \le 2\pi, \quad t > 0, \\ u(x,0) =& u_0(x), \end{aligned}$$

with periodic boundary condition.

In the computation, the initial condition is set to be

$$u(x,0) = u_0(x) = \frac{1}{2}(1+\sin(x)) \tag{5.1}$$

with the final time $t = 0.5$. The source term is $s(u) = -u$. The exact solution is

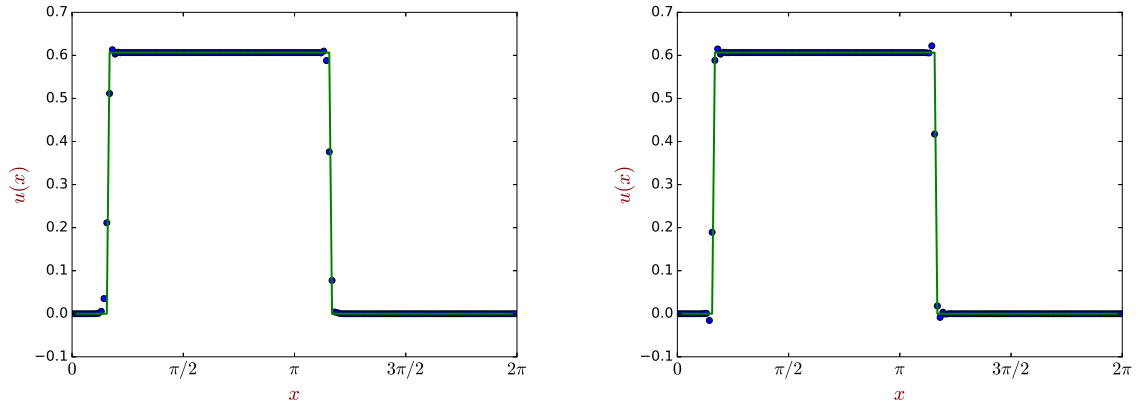
$$u(x,t) = e^{-\frac{1}{\varepsilon}t} u_0(x-t)$$

.

The errors with and without limiters are listed in Table 5.1. We observe designed order of accuracy without limiter and order degeneration with limiter, especially for high-order methods. The order degeneration phenomenon is also observed in [43]. The reason is that the high-order RK methods depend more heavily on the error cancellations in different stages, but the bound-preserving limiter after each stage will destroy this cancellation.

Table 5.1: Example 5.1: 1D advection equation with linear source term. Modified exponential RK method without and with limiter (left: no limiter; right: limiter). $\varepsilon = 1$.

	N	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
P^1	20	5.97e-03	-	4.55e-03	-	7.00e-03	-	8.35e-03	-
	40	1.49e-03	2.000	1.19e-03	1.930	1.76e-03	1.993	2.12e-03	1.981
	80	3.92e-04	1.927	2.88e-04	2.051	4.52e-04	1.958	6.37e-04	1.732
	160	9.66e-05	2.023	7.39e-05	1.962	1.08e-04	2.061	1.53e-04	2.056
	320	2.39e-05	2.015	1.91e-05	1.952	2.60e-05	2.063	4.03e-05	1.925
	640	6.07e-06	1.977	4.67e-06	2.033	6.47e-06	2.005	1.01e-05	1.991
P^2	20	1.63e-04	-	1.58e-04	-	2.74e-04	-	3.89e-04	-
	40	2.04e-05	2.999	1.97e-05	3.007	4.48e-05	2.616	1.03e-04	1.916
	80	2.55e-06	3.002	2.46e-06	2.997	6.21e-06	2.849	2.84e-05	1.858
	160	3.19e-07	3.000	3.08e-07	3.000	8.59e-07	2.856	1.02e-05	1.474
	320	3.99e-08	3.000	3.85e-08	3.000	1.21e-07	2.824	2.35e-06	2.121
	640	4.98e-09	3.000	4.81e-09	3.000	1.60e-08	2.921	4.87e-07	2.274
P^3	20	3.37e-06	-	3.82e-06	-	8.46e-06	-	1.64e-05	-
	40	2.06e-07	4.033	2.03e-07	4.234	1.07e-06	2.988	4.84e-06	1.757
	80	1.27e-08	4.024	1.39e-08	3.867	1.40e-07	2.934	1.26e-06	1.937
	160	7.89e-10	4.006	8.57e-10	4.021	1.92e-08	2.861	3.16e-07	2.000
	320	4.93e-11	4.000	5.37e-11	3.998	2.61e-09	2.882	8.05e-08	1.972
	640	3.27e-12	3.915	3.35e-12	4.000	3.59e-10	2.862	2.19e-08	1.881



(a) with limiter (b) without limiter

Figure 5.1: Example 5.1: 1D advection equation with piecewise constant initial value. Third-order modified exponential RK method with P^2-DG. The grid number $N = 160$. Solid line: exact solution; circle: numerical solution (cell averages). Left: with limiter; right: without limiter.

We also test another piecewise constant initial value for the same linear advection equation with linear source term:

$$u_0(x) = \begin{cases} 1, & 0 \le x \le \pi, \\ 0, & \pi < x \le 2\pi. \end{cases}$$

In the computation, the third-order modified exponential RK method is applied to the DG scheme. The final time $t = 0.5$ and the grid number $N = 160$. As shown in Figure 5.1, the numerical solution stays non-negative with the bound-preserving limiter. As a comparison, we also show the result of the same DG scheme without limiter. The undershoot near the discontinuity is then apparent.

Example 5.2 (1D Burgers' equation with source term)**.** We solve the 1D Burgers' equation with periodic boundary condition.

$$\begin{aligned} u_t + (\frac{u^2}{2})_x =& \frac{1}{\varepsilon} s(u), \quad 0 \le x \le 2\pi, \quad t > 0, \\ u(x,0) =& u_0(x), \end{aligned}$$

Take the initial value

$$u_0(x) = \frac{1}{2}(\sin(x) + 1) \tag{5.2}$$

and the source term $s(u) = -u$. The final time $t = 0.2$. The relaxation parameter $\varepsilon = 1$.

Again, we can clearly observe the designed order of accuracy without limiter in Table 5.2. In this case, the numerical results with limiter are almost the same with those without limiter.

Table 5.2: Example 5.2: 1D Burgers' equation with linear source term. Modified exponential RK method without and with limiter (left: no limiter; right: limiter). $\varepsilon = 1$.

	N	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
P^1	20	6.12e-03	-	5.28e-03	-	6.67e-03	-	6.98e-03	-
	40	1.70e-03	1.848	1.53e-03	1.783	1.76e-03	1.920	1.72e-03	2.018
	80	4.66e-04	1.865	4.30e-04	1.834	4.74e-04	1.895	4.69e-04	1.877
	160	1.23e-04	1.922	1.13e-04	1.923	1.24e-04	1.934	1.16e-04	2.021
	320	3.17e-05	1.957	2.88e-05	1.976	3.18e-05	1.962	2.88e-05	2.001
	640	8.07e-06	1.974	7.23e-06	1.997	8.09e-06	1.976	7.23e-06	1.997
P^2	20	2.31e-04	-	2.52e-04	-	2.39e-04	-	2.52e-04	-
	40	2.86e-05	3.015	3.73e-05	2.755	2.90e-05	3.043	3.73e-05	2.755
	80	3.57e-06	3.006	4.59e-06	3.021	3.59e-06	3.016	4.59e-06	3.021
	160	4.45e-07	3.003	5.90e-07	2.959	4.46e-07	3.007	5.90e-07	2.959
	320	5.55e-08	3.002	7.47e-08	2.982	5.56e-08	3.003	7.47e-08	2.982
	640	6.94e-09	3.001	9.40e-09	2.991	6.95e-09	3.002	9.40e-09	2.991
P^3	20	4.62e-06	-	6.40e-06	-	4.62e-06	-	6.40e-06	-
	40	2.71e-07	4.091	4.41e-07	3.858	2.71e-07	4.091	4.41e-07	3.858
	80	1.80e-08	3.910	2.81e-08	3.973	1.80e-08	3.910	2.81e-08	3.973
	160	1.14e-09	3.983	1.84e-09	3.930	1.14e-09	3.983	1.84e-09	3.930
	320	7.17e-11	3.992	1.19e-10	3.959	7.17e-11	3.992	1.19e-10	3.959
	640	4.63e-12	3.952	7.38e-12	4.005	4.63e-12	3.951	7.38e-12	4.005

Example 5.3 (2D advection equation with source term)**.** We solve the 2D advection equation with periodic boundary condition.

$$\begin{aligned} u_t + u_x + u_y =& \frac{1}{\varepsilon} s(u), \quad 0 \le x \le 2\pi, \quad t > 0, \\ u(x, y, 0) =& u_0(x, y), \end{aligned}$$

Take the initial value

$$u_0(x) = \exp(\sin(x + y)) - \exp(-1) \tag{5.3}$$

with the source term $s(u) = -u$. The exact solution is

$$u(x, y, t) = (\exp(\sin(x + y - 2t)) - \exp(-1)) \exp\left(-\frac{t}{\varepsilon}\right). \tag{5.4}$$

In the computation, the final time $t = 1.2$ and the parameter $\varepsilon = 1$. For the sake of simplicity, the uniform grid is used with $\Delta x = \Delta y$.

The errors are presented in Table 5.3. Once more, we observe designed order of accuracy without limiter and order degeneration with limiter, especially for high-order methods.

Table 5.3: Example 5.3: 2D advection equation with linear source term. Modified exponential RK method without and with limiter (left: no limiter; right: limiter). $\varepsilon = 1$.

	N	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
P^1	20	4.26e-03	-	3.29e-02	-	4.50e-03	-	3.26e-02	-
	40	8.71e-04	2.289	9.81e-03	1.745	9.07e-04	2.311	9.67e-03	1.754
	80	1.92e-04	2.181	2.71e-03	1.856	1.98e-04	2.199	2.68e-03	1.852
	160	4.55e-05	2.080	7.10e-04	1.934	4.65e-05	2.088	7.05e-04	1.926
	320	1.11e-05	2.032	1.81e-04	1.969	1.13e-05	2.041	1.81e-04	1.964
P^2	20	2.57e-04	-	5.58e-03	-	2.88e-04	-	5.75e-03	-
	40	2.97e-05	3.113	7.46e-04	2.903	3.34e-05	3.107	7.46e-04	2.946
	80	3.65e-06	3.024	9.37e-05	2.993	4.33e-06	2.950	9.37e-05	2.993
	160	4.54e-07	3.007	1.17e-05	2.996	5.75e-07	2.912	1.17e-05	2.996
	320	5.67e-08	3.002	1.47e-06	2.999	7.83e-08	2.875	2.05e-06	2.517
P^3	20	1.97e-05	-	8.33e-04	-	2.81e-05	-	8.33e-04	-
	40	1.21e-06	4.026	5.70e-05	3.868	1.69e-06	4.056	5.70e-05	3.868
	80	7.56e-08	4.003	3.60e-06	3.987	1.38e-07	3.613	3.60e-06	3.987
	160	4.72e-09	4.002	2.27e-07	3.989	1.41e-08	3.293	5.58e-07	2.688
	320	2.95e-10	4.001	1.42e-08	3.999	1.66e-09	3.080	1.45e-07	1.941

Example 5.4 (2D Burgers' equation with source term)**.** We solve the 2D Burgers' equation with periodic boundary condition.

$$\begin{aligned} u_t + (\frac{u^2}{2})_x + (\frac{u^2}{2})_y =& \frac{1}{\varepsilon} s(u), \quad 0 \le x \le 2\pi, \quad t > 0, \\ u(x, y, 0) =& u_0(x, y), \end{aligned}$$

Take the initial value

$$u_0(x) = \frac{1}{2}(\sin(x + y) + 1) \tag{5.5}$$

with the linear source $s(u) = -u$. The final time $t = 1.2$. The parameter is $\varepsilon = 1$.

Similar to the results in 1D Burgers' equation, the designed orders of accuracy with and without limiters are observed.

5.2 Stiff case

In this section, we investigate the numerical behaviour of our methods for solving the stiff problems.

As in the numerical examples for the ODE, since the linear source term results in exponential decay in time, with small ε, the solution decays to zero for $t = O(1)$. Hence, we take the source to be polynomial with high degree in this part. For saving space, we only present the numerical results with limiters.

We start with the 1D linear advection equation. In Example 5.1, take the source term to be $s(u) = -u^7$ and $\varepsilon = 1 \times 10^{-4}$. From Table 5.5, we observe that the converge orders are all around one, which is the same with the results for the ODE. Moreover, the errors with high order methods are slightly larger than those with low order methods.

Presented in Figure 5.2 are the numerical and exact solutions with stiff source for the stiff advection equation with the initial value (5.1). We observe that, with the mesh number $N = 160$, i.e., $\Delta x, \Delta t \ll \varepsilon$, the profiles of the solutions are captured well.

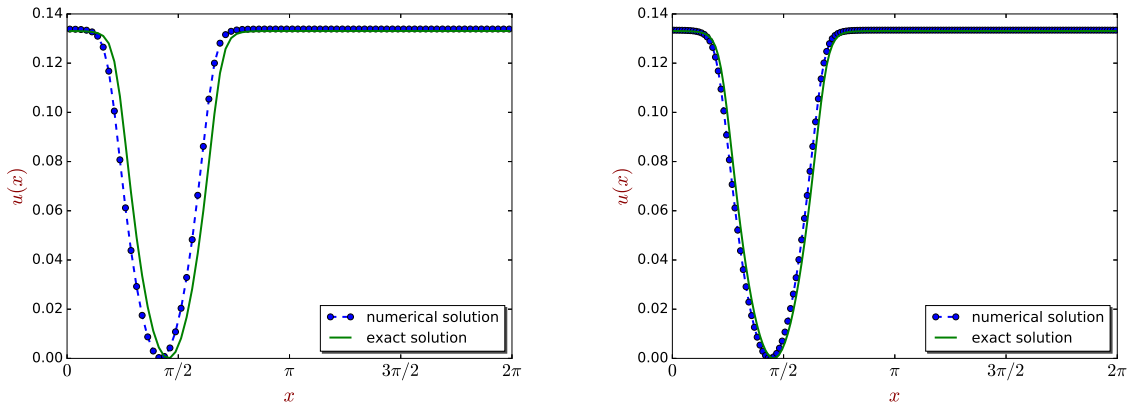
Next, we move to 1D Burgers' equation with stiff source term. As shown in Table 5.6, they

Table 5.4: Example 5.4: 2D Burgers' equation with linear source term. Modified exponential RK method without and with limiter (left: no limiter; right: limiter). $\varepsilon = 1$.

	N	L^1 error	order	L^∞ error	order	L^1 error	order	L^∞ error	order
P^1	20	1.66e-03	-	3.07e-02	-	1.72e-03	-	3.74e-02	-
	40	4.31e-04	1.942	1.15e-02	1.411	4.37e-04	1.972	1.16e-02	1.693
	80	1.08e-04	2.001	3.82e-03	1.595	1.08e-04	2.013	3.82e-03	1.596
	160	2.70e-05	1.996	1.07e-03	1.833	2.71e-05	2.002	1.07e-03	1.833
	320	6.78e-06	1.993	2.83e-04	1.921	6.79e-06	1.996	2.83e-04	1.921
P^2	20	3.39e-04	-	1.39e-02	-	3.43e-04	-	1.39e-02	-
	40	4.97e-05	2.769	4.47e-03	1.641	5.00e-05	2.781	4.47e-03	1.641
	80	6.70e-06	2.892	1.10e-03	2.028	6.70e-06	2.898	1.10e-03	2.028
	160	8.60e-07	2.960	1.56e-04	2.810	8.60e-07	2.962	1.56e-04	2.810
	320	1.09e-07	2.984	2.07e-05	2.920	1.09e-07	2.984	2.07e-05	2.920
P^3	20	8.48e-05	-	9.99e-03	-	9.73e-05	-	9.99e-03	-
	40	9.42e-06	3.169	1.53e-03	2.711	9.42e-06	3.368	1.53e-03	2.712
	80	6.06e-07	3.959	1.54e-04	3.310	6.06e-07	3.959	1.54e-04	3.310
	160	4.06e-08	3.902	1.17e-05	3.713	4.05e-08	3.902	1.17e-05	3.713
	320	2.59e-09	3.970	7.87e-07	3.897	2.59e-09	3.969	7.87e-07	3.897

Table 5.5: Example 5.1: 1D advection equation with stiff source term. Modified exponential RK method with limiter. $\varepsilon = 1 \times 10^{-4}$. The source term $s(u) = -u^7$.

	N	L^1 error	order	L^∞ error	order
P^1	20	7.88e-02	-	6.07e-02	-
	40	3.58e-02	1.137	3.18e-02	0.936
	80	1.63e-02	1.141	1.46e-02	1.126
	160	7.28e-03	1.159	6.35e-03	1.196
	320	3.21e-03	1.180	2.75e-03	1.210
P^2	20	1.53e-01	-	1.08e-01	-
	40	7.00e-02	1.126	5.78e-02	0.906
	80	3.11e-02	1.170	2.68e-02	1.110
	160	1.36e-02	1.196	1.18e-02	1.181
	320	5.81e-03	1.225	5.06e-03	1.222
P^3	20	3.11e-01	-	1.46e-01	-
	40	1.64e-01	0.927	1.15e-01	0.354
	80	7.49e-02	1.126	6.12e-02	0.905
	160	3.29e-02	1.185	2.82e-02	1.120
	320	1.42e-02	1.218	1.22e-02	1.203



(a) $N = 80$ (b) $N = 160$

Figure 5.2: Example 5.1: 1D linear advection equation with stiff source term. The the initial value (5.1) with limiter. Third-order modified exponential RK method with P^2-DG. $N = 80$ and $N = 160$, $t = 3$, $\Delta x = 2\pi/N$, $\Delta t = \Delta x/10$. Solid line: exact solution; circle: numerical solution (cell averages).

all have first-order of accuracy with $\Delta x, \Delta t \ll \varepsilon$. Presented in Figure 5.3 are the numerical and exact solutions with stiff source. With the grid number $N = 80$, the maximum value of the numerical solutions does not coincide with that of the exact solutions. With the grid number $N = 160$, the profiles of the solutions are captured well. This example indicates that the grid number should not be too small when using our methods.

Table 5.6: Example 5.2: 1D Burgers' equation with stiff source term. Modified exponential RK method with limiter, source term $s(u) = -u^7$.

	N	L^1 error	order	L^∞ error	order
P^1	20	1.91e-01	-	4.97e-02	-
	40	6.12e-02	1.644	1.44e-02	1.785
	80	2.49e-02	1.295	5.75e-03	1.326
	160	1.08e-02	1.212	2.50e-03	1.204
	320	4.73e-03	1.189	1.10e-03	1.177
P^2	20	7.96e-01	-	2.05e-01	-
	40	1.21e-01	2.719	2.76e-02	2.893
	80	4.15e-02	1.544	9.38e-03	1.559
	160	1.64e-02	1.339	3.72e-03	1.334
	320	6.65e-03	1.302	1.52e-03	1.289
P^3	20	1.55e+00	-	4.70e-01	-
	40	9.82e-01	0.659	2.63e-01	0.837
	80	1.42e-01	2.793	3.26e-02	3.013
	160	4.59e-02	1.626	1.04e-02	1.652
	320	1.77e-02	1.374	4.00e-03	1.378

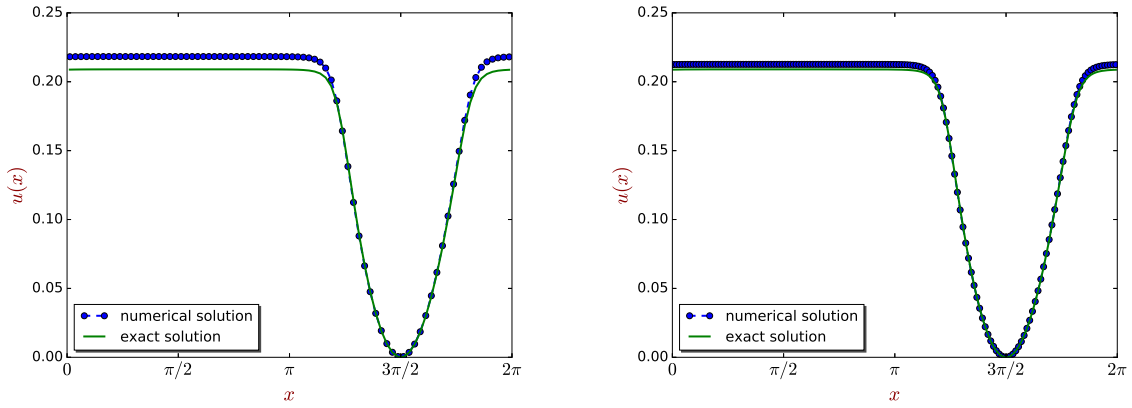
We also investigate the 2D linear equation with stiff source. The initial value is

$$u_0(x, y) = \frac{1}{2}(\sin(x + 3y) + 1) \tag{5.6}$$

and the parameter $\varepsilon = 1 \times 10^{-4}$.

Again, the first-order convergence orders are observed in Table 5.7. Illustrated in Figure 5.4 are the numerical and exact solutions. We observe good resolution of our scheme for this 2D example.

The results of our methods for the 2D Burgers' equation with stiff source are presented in Table 5.8, which shows around first-order accuracy. Here, the initial value is taken to be (5.6) and the parameter $\varepsilon = 1 \times 10^{-4}$.

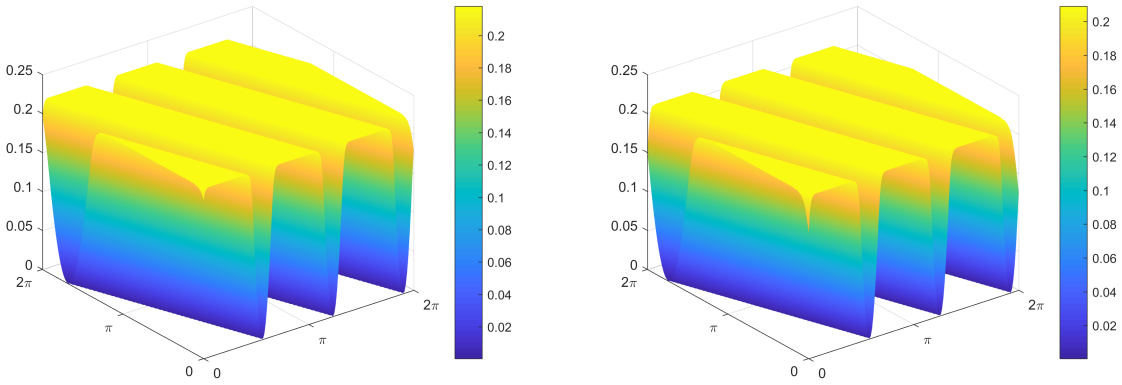


(a) $N = 80$ (b) $N = 160$

Figure 5.3: Example 5.2: 2D linear advection equation with stiff source term. The initial value (5.1) with limiter. Third-order modified exponential RK method with P^2-DG. $N = 80$ and $N = 160$, $t = 0.2$, $\Delta x = 2\pi/N$, $\Delta t = \Delta x/20/\max u$. Solid line: exact solution; circle: numerical solution (cell averages).

Table 5.7: Example 5.3: 2D linear advection equation with stiff source term. Modified exponential RK method with limiter, source term $s(u) = -u^7$.

	N	L^1 error	order	L^∞ error	order
P^1	20	2.27e-01	-	3.91e-01	-
	40	8.83e-02	1.362	2.17e-01	0.850
	80	2.46e-02	1.847	8.24e-02	1.399
	160	9.83e-03	1.322	3.35e-02	1.298
	320	4.28e-03	1.199	1.39e-02	1.270
P^2	20	2.72e-01	-	4.14e-01	-
	40	1.73e-01	0.652	2.86e-01	0.531
	80	4.51e-02	1.937	1.29e-01	1.152
	160	1.80e-02	1.329	5.76e-02	1.161
	320	7.45e-03	1.270	2.48e-02	1.216
P^3	20	3.30e-01	-	5.91e-01	-
	40	2.93e-01	0.172	4.72e-01	0.326
	80	2.02e-01	0.532	2.97e-01	0.667
	160	4.98e-02	2.024	1.34e-01	1.145
	320	1.91e-02	1.383	5.97e-02	1.170



(a) numerical solution (b) exact solution

Figure 5.4: Example 5.3: 2D linear advection equation with stiff source term. The initial value (5.6). Third-order modified exponential RK method with P^2-DG. $N = 160$, $t = 0.2$, $\Delta x = \Delta y = 2\pi/N$, $\Delta t = \Delta x/10$. Left: numerical solution; right: exact solution.

Table 5.8: Example 5.4: 2D Burgers' equation with stiff source term. The initial value: $u_0(x, y) = \frac{1}{2}(\sin(x + 3y) + 1)$. Modified exponential RK method with limiter, source term $s(u) = -u^7$.

	N	L^1 error	order	L^∞ error	order
P^1	20	1.88e-01	-	3.81e-01	-
	40	5.39e-02	1.801	1.03e-01	1.881
	80	1.17e-02	2.198	2.67e-02	1.953
	160	4.44e-03	1.402	8.14e-03	1.714
	320	1.92e-03	1.212	3.08e-03	1.403
P^2	20	2.28e-01	-	4.14e-01	-
	40	1.30e-01	0.805	2.05e-01	1.010
	80	2.12e-02	2.618	2.92e-02	2.812
	160	7.46e-03	1.509	1.12e-02	1.388
	320	2.98e-03	1.325	4.87e-03	1.199
P^3	20	3.00e-01	-	5.91e-01	-
	40	2.51e-01	0.255	4.72e-01	0.326
	80	1.60e-01	0.652	2.63e-01	0.840
	160	2.47e-02	2.695	3.27e-02	3.010
	320	8.22e-03	1.587	1.17e-02	1.484

6 Concluding remarks

In this work, we develop high-order modified exponential RK methods by replacing the exponential functions by polynomials without destroying the accuracy. The bound-preserving property and the weak asymptotic-preserving property are shown for these methods. By applying these time discretization methods to the semi-discrete DG schemes, we successfully obtain the bound-preserving DG schemes. Various numerical tests validate the theoretical analysis of our scheme.

We also mention several drawbacks of this work. Although our schemes are high-order accurate in resolved region and with consistent initial values in unresolved region, they degenerate to around first-order of accuracy with unresolved initial layers numerically. It is nontrivial to analyze this order degeneration phenomenon theoretically. New and powerful ideas need to be introduced to construct higher-order and uniformly accurate time integrators with bound-preserving property. We also hope to extend the idea in this work to specific hyperbolic systems with stiff sources, e.g. reactive Euler equations, for preserving the positivity of the density and pressure.

Moreover, we only focus on the time discretization for scalar hyperbolic equations with stiff source terms, with the standard DG spatial discretization. It is well-known that the spatial discretization of the source term should also mimic that of the convection term to avoid non-physical numerical wave propagations. In [15], a class of monotone finite-difference schemes and a split-step scheme are analyzed theoretically and numerically. We hope to extend the analysis in [15] to our newly developed time discretizations with different spatial discretizations. These issues constitute our ongoing work.

References

[1] W. Bao and S. Jin. The random projection method for hyperbolic conservation laws with stiff reaction terms. *Journal of Computational Physics*, 163(1):216–248, 2000.

[2] A. C. Berkenbosch, E. F. Kaasschieter, and J. H. M. T. Boonkkamp. *The numerical wave speed for one-dimensional scalar hyperbolic conservation laws with source terms*. Eindhoven University of Technology, Department of Mathematics and Computing Science, 1994.

[3] A. Chalabi. On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms. *Mathematics of Computation*, 66(218):527–545, 1997.

[4] A. Chertock, S. Cui, A. Kurganov, and T. Wu. Steady state and sign preserving semi-implicit Runge–Kutta methods for ODEs with stiff damping term. *SIAM Journal on Numerical Analysis*, 53(4):2008–2029, 2015.

[5] A. Chertock, S. Cui, A. Kurganov, and T. Wu. Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms. *International Journal for Numerical Methods in Fluids*, 78(6):355–383, 2015.

[6] A. J. Christlieb, Y. Liu, Q. Tang, and Z. Xu. High order parametrized maximum-principle-preserving and positivity-preserving WENO schemes on unstructured meshes. *Journal of Computational Physics*, 281:334–351, 2015.

[7] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Mathematics of Computation*, 54(190):545–581, 1990.

[8] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Mathematics of Computation*, 52(186):411–435, 1989.

[9] R. Donat, I. Higueras, and A. Martínez-Gavara. On stability issues for IMEX schemes applied to 1D scalar hyperbolic equations with stiff reaction terms. *Mathematics of Computation*, 80(276):2097–2126, 2011.

[10] M. Dumbser, C. Enaux, and E. F. Toro. Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws. *Journal of Computational Physics*, 227(8):3971–4001, 2008.

[11] C. W. Gear. Runge-Kutta starters for multistep methods. *ACM Transactions on Mathematical Software (TOMS)*, 6(3):263–279, 1980.

[12] S. Gottlieb, Z. J. Grant, and L. Isherwood. Strong stability preserving integrating factor Runge-Kutta methods. *arXiv preprint arXiv:1708.02595*, 2017.

[13] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of Computation*, 67(221):73–85, 1998.

[14] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43(1):89–112, 2001.

[15] D. Griffiths, A. Stuart, and H. Yee. Numerical wave propagation in an advection equation with a nonlinear source term. *SIAM journal on Numerical Analysis*, 29(5):1244–1260, 1992.

[16] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numerica*, 19:209–286, 2010.

[17] J. Huang and C.-W. Shu. Error estimates to smooth solutions of semi-discrete discontinuous Galerkin methods with quadrature rules for scalar conservation laws. *Numerical Methods for Partial Differential Equations*, 33(2):467–488, 2017.

[18] J. Huang and C.-W. Shu. A second-order asymptotic-preserving and positivity-preserving discontinuous Galerkin scheme for the Kerr–Debye model. *Mathematical Models and Methods in Applied Sciences*, 27(03):549–579, 2017.

[19] S. Jin. Runge-Kutta methods for hyperbolic conservation laws with stiff relaxation terms. *Journal of Computational Physics*, 122(1):51–67, 1995.

[20] S. Jin. Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review. *Lecture Notes for Summer School on "Methods and Models of Kinetic Theory" (M&MKT), Porto Ercole (Grosseto, Italy)*, pages 177–216, 2010.

[21] D. I. Ketcheson. Highly efficient strong stability-preserving Runge–Kutta methods with low-storage implementations. *SIAM Journal on Scientific Computing*, 30(4):2113–2136, 2008.

[22] J. F. B. M. Kraaijevanger. Contractivity of Runge–Kutta methods. *BIT Numerical Mathematics*, 31(3):482–528, 1991.

[23] S. N. Kružkov. First order quasilinear equations in several independent variables. *Mathematics of the USSR-Sbornik*, 10(2):217, 1970.

[24] R. J. LeVeque and H. C. Yee. A study of numerical methods for hyperbolic conservation laws with stiff source terms. *Journal of Computational Physics*, 86(1):187–210, 1990.

[25] Q. Li and L. Pareschi. Exponential Runge–Kutta for the inhomogeneous Boltzmann equations with high order of accuracy. *Journal of Computational Physics*, 259:402–420, 2014.

[26] C. Liang and Z. Xu. Parametrized maximum principle preserving flux limiters for high order schemes solving multi-dimensional scalar hyperbolic conservation laws. *Journal of Scientific Computing*, 58(1):41–60, 2014.

[27] X.-D. Liu and S. Osher. Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes i. *SIAM Journal on Numerical Analysis*, 33(2):760–779, 1996.

[28] T. Qin, C.-W. Shu, and Y. Yang. Bound-preserving discontinuous Galerkin methods for relativistic hydrodynamics. *Journal of Computational Physics*, 315:323–347, 2016.

[29] S. Ruuth. Global optimization of explicit strong-stability-preserving Runge-Kutta methods. *Mathematics of Computation*, 75(253):183–207, 2006.

[30] H. J. Schroll and R. Winther. Finite-difference schemes for scalar conservation laws with source terms. *IMA Journal of Numerical Analysis*, 16(2):201–215, 1996.

[31] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, 1988.

[32] R. J. Spiteri and S. J. Ruuth. A new class of optimal high-order strong-stability-preserving time discretization methods. *SIAM Journal on Numerical Analysis*, 40(2):469–491, 2002.

[33] T. Tang and Z.-H. Teng. Error bounds for fractional step methods for conservation laws with source terms. *SIAM Journal on Numerical Analysis*, 32(1):110–127, 1995.

[34] W. Wang, C.-W. Shu, H. Yee, and B. Sjögreen. High order finite difference methods with subcell resolution for advection equations with stiff source terms. *Journal of Computational Physics*, 231(1):190–214, 2012.

[35] K. Wu and H. Tang. High-order accurate physical-constraints-preserving finite difference WENO schemes for special relativistic hydrodynamics. *Journal of Computational Physics*, 298:539–564, 2015.

[36] Y. Xing and X. Zhang. Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes. *Journal of Scientific Computing*, 57(1):19–41, 2013.

[37] Y. Xing, X. Zhang, and C.-W. Shu. Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations. *Advances in Water Resources*, 33(12):1476–1493, 2010.

[38] T. Xiong, J.-M. Qiu, and Z. Xu. A parametrized maximum principle preserving flux limiter for finite difference RK-WENO schemes with applications in incompressible flows. *Journal of Computational Physics*, 252:310–331, 2013.

[39] Z. Xu. Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem. *Mathematics of Computation*, 83(289):2213–2238, 2014.

[40] Z. Xu and X. Zhang. Bound-Preserving High-Order Schemes. *Handbook of Numerical Analysis*, 18:81–102, 2017.

[41] Q. Zhang and C.-W. Shu. Stability analysis and a priori error estimates of the third order explicit Runge-Kutta discontinuous Galerkin method for scalar conservation laws. *SIAM Journal on Numerical Analysis*, 48(3):1038–1063, 2010.

[42] X. Zhang. On positivity-preserving high order discontinuous Galerkin schemes for compressible navier–stokes equations. *Journal of Computational Physics*, 328:301–343, 2017.

[43] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.

[44] X. Zhang and C.-W. Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *Journal of Computational Physics*, 229(23):8918–8934, 2010.

[45] X. Zhang and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proceedings of the Royal Society of London Series A*, 467:2752–2776, 2011.

[46] X. Zhang and C.-W. Shu. Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. *Journal of Computational Physics*, 230(4):1238–1248, 2011.

[47] Y. Zhang, X. Zhang, and C.-W. Shu. Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection–diffusion equations on triangular meshes. *Journal of Computational Physics*, 234:295–316, 2013.

[48] X. Zhao, Y. Yang, and C. E. Seyler. A positivity-preserving semi-implicit discontinuous Galerkin scheme for solving extended magnetohydrodynamics equations. *Journal of Computational Physics*, 278:400–415, 2014.