ABSTRACT

Title of Thesis: HIGH RESOLUTION TIME-LIMITER SCHEMES FOR CONSERVATION LAWS

Jingcheng Lu
Master of Science, 2022

Thesis Directed by: Professor James D.Baeder
Department of Aerospace Engineering

This study investigates the high resolution time-limiter schemes for conservation laws. These schemes are proposed (K.Duraisamy, J.D.Baeder, J-G Liu, 2003) to enhance the stability of high order implicit time marching when the time step is beyond the original TVD limit. The improved stability is realized by taking local convex combination of a higher order oscillatory method (accurate mode) with a first order unconditionally TVD method (stable mode). The application of time-limiters, which detects the local smoothness, enables the self-adjusting switch between different modes. One of the main aspects of this work is employing time-limiters to improve the stability of the strongly S-stable DIRK3 scheme, which is shown to be non-SSP and thus may generate strong oscillations in non-smooth problems. The new Limited-DIRK3 scheme (L-DIRK3) is proposed. For convenience of applications to systems of equations, we also propose a new and convenient construction of time-limiter, which allows an arbitrary choice of reference quantity with minimal computation cost. Another innovation of our work is the extension of time-limiter schemes to multi-dimensional problems and convection-

diffusion problems. The numerical results for one- and two-dimensional problems confirm that the L-DIRK3 scheme generate high resolution and less oscillatory solutions under large time step. Particularly, the L-DIRK3 scheme shows a clear improvement against the original DIRK3 in convection-diffusion problems when a large CFL number is taken.

HIGH RESOLUTION TIME-LIMITER SCHEMES FOR CONSERVATION LAWS

by

Jingcheng Lu

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2022

Advisory Committee:
Professor James D.Baeder, Chair/Advisor
Professor Christoph Brehm
Professor Pino Martin

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The present work studies the high resolution numerical schemes for conservation laws,

$$\frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}f(u(x,t)) = 0, \quad u(x,0) = u_0(x), \quad x \in \Omega, \quad t \in \mathbb{R}_+. \tag{1.1}$$

To numerically solve (1.1), a common approach is the semidiscrete method-of-lines. In this framework, proper spatial discretization is first performed over the suitable domain, resulting in a set of ordinary differential equations (ODEs)

$$\frac{d}{dt}u_j(t) + \frac{\hat{f}_{j+\frac{1}{2}}(t) - \hat{f}_{j-\frac{1}{2}}(t)}{\Delta x} = 0, \quad j = 1, 2, \cdots, N. \tag{1.2}$$

Here N is the number of mesh points, $u_j(t)$ is the approximate solution to the point value $u(x_j,t)$/the cell average $\bar{u}(x_j,t) := \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x,t)dx$ in the framework of finite difference/finite volume methods, $\hat{f}_{j+\frac{1}{2}}(t)$ is the numerical flux at cell interface $x_{j+\frac{1}{2}}$ at time t. The resulting system of ODEs is then integrated with standard time integration schemes, e.g. Runge-Kutta methods, linear multistep methods. In many application problems, the ODE systems obtained from spatial discretization may be stiff, implicit time integration would be preferred so that we can apply large time steps.

Harten et al [1] have shown that the solutions of monotone schemes for conservation laws are only first order accurate in space and time. To achieve higher order of accuracy, less restrictive monotonicity conditions are usually considered, e.g. monotonicity preserving, total variation diminishing (TVD), and essentially non-oscillatory. Based on these frameworks, many high order non-oscillatory spatial discretization schemes have been proposed, for example, the TVD schemes [2], and the essentially non-oscillatory (ENO) type schemes [3]. However, ensuring non-oscillatory solutions of these high order schemes may impose severe restriction on the time step. Gottlieb, Shu, and Tadmor [4] have shown that even implicit time integration schemes become conditionally TVD when the order of accuracy in time is higher than one. Strong oscillations may occur if the time step is beyond the TVD limit. To enable the use of larger time step when using high order implicit methods, Duraisamy, Baeder and Liu [5] proposed the concepts of *time-limiter* schemes.

The main idea of limited time integration is to locally drop the order of accuracy in time near the discontinuities and extrema. When the time-limiter detects the occurrence of large solution gradient, the scheme is locally switched to a first order unconditionally TVD method. In this way, the monotonicity condition is locally satisfied where the time evolution of the solution is not smooth, while a high order of accuracy is still maintained in most of the solution domain. Duraisamy et al [5] proposed the Limited Trapezoidal (LTrap) and the Limited DIRK2 (L-DIRK2) schemes. These time-limiter schemes effectively reduce the oscillations of the Trapezoidal and the DIRK2 schemes near the discontinuities and extrema under large time steps, while the second order accuracy in time is still maintained.

In some modern applications such as turbulence simulations, very high order spatial

discretizations may be applied, such as the 5th order WENO scheme. Third order or even higher order implicit time integration would be preferred. The 3-stage, diagonally implicit Runge-Kutta (DIRK3) scheme [6] is a possible choice, but it can be highly oscillatory under large CFL numbers. The main work of this thesis is to apply the idea of limited time integration to improve the performance of DIRK3 scheme. The resulting scheme – the Limited-DIRK3 (L-DIRK3) is constructed. A new and simple construction of time-limiter is proposed in convenience of application to systems of equations. To examine the robustness and accuracy of L-DIRK3, the scheme is tested against a series of scalar/systems of equations. The extension of the scheme to multi-dimensions is validated through examples in two-space dimensions.

Chapter 2: Background: Strong Stability Preserving (SSP) Time Integration

In this chapter, we review the framework of Strong Stability Preserving (SSP) time integration schemes for solving time-dependent partial differential equations (PDEs). The concept of SSP is the cornerstone of nonlinear stability analysis of time integration schemes in a method-of-lines framework.

SSP time discretization methods were first studied in [7] and [8] (where they were called *TVD* time discretizations) and were further developed in [9], [4] and [10]. The idea is that assuming the first order forward Euler time discretization of the method-of-lines ODE is strongly stable under an *arbitrary* norm when the time step, Δt, is properly restricted, then a higher order SSP time discretization will maintain the strong stability for the same norm, perhaps under a modified time step restriction. Hence, the concept of strong stability preserving is a generalization of TVD stability.

In section 2.1, we illustrate the motivations of developing SSP time integration methods and introduce related concepts. In section 2.2 and section 2.3, critical results regarding strong stability of Runge-Kutta methods and multistep methods will be stated.

2.1 Motivation of SSP Methods

Consider the semidiscrete problem (1.2) for conservation laws (1.1)

$$\frac{d}{dt}u_j(t) = -\frac{\hat{f}_{j+\frac{1}{2}}(t) - \hat{f}_{j-\frac{1}{2}}(t)}{\Delta x} = L_j(\mathbf{u}(t)), \quad j = 1, 2, \cdots, N,$$

where the flux derivative, $\frac{\partial}{\partial x}f(u)$, is approximated by a TVD spatial discretization, e.g. [2], [11], [12]. This can be equivalently represented as a system of ODEs

$$\frac{d\mathbf{u}(t)}{dt} = \mathbf{L}(\mathbf{u}(t)), \tag{2.1}$$

where $\mathbf{u}(t) = [u_1(t), \cdots, u_N(t)]^\top$, $\mathbf{L}(\mathbf{u}(t)) = [L_1(\mathbf{u}(t)), \cdots, L_N(\mathbf{u}(t))]^\top$. Assume that the spatial discretization has the property that when it is combined with the first order forward Euler time discretization,

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t\mathbf{L}(\mathbf{u}^n), \tag{2.2}$$

for a sufficiently small time step restricted by CFL condition

$$\Delta t \leq \Delta t_{FE}, \tag{2.3}$$

the scheme is TVD, i.e. the total variation (TV) of one-dimensional solution

$$u^n = \sum_j u_j^n 1_{\{x_{j-\frac{1}{2}} \leq x \leq x_{j+\frac{1}{2}}\}}$$

does not increase in time

$$||\mathbf{u}^{n+1}||_{TV} \leq ||\mathbf{u}^n||_{TV}, \quad ||\mathbf{u}^n||_{TV} := \sum_j |u^n_{j+1} - u^n_j|. \tag{2.4}$$

The objective of developing the high order SSP Runge-Kutta or multistep time discretizations is to achieve higher order accuracy in time while preserving the strong stability (2.4) of the forward Euler, perhaps under a modified CFL condition

$$\Delta t \leq k \Delta t_{FE}. \tag{2.5}$$

We call the constant k the *SSP coefficient/TVD limit* of an SSP time integration scheme.

In [9], [4], numerical results were given to show that when using high order time discretization that is linearly stable but not strongly stable, oscillations may occur even if the spatial discretization is TVD when combined with the forward Euler time discretization. To illustrate the importance of strong stability, consider one-dimensional Burger's equation

$$u_t + (\frac{1}{2}u^2)_x = 0 \tag{2.6}$$

with Riemann initial data

$$u(x,0) = \begin{cases} 1 & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}. \tag{2.7}$$

The spatial discretization is obtained by second order MUSCL reconstruction [13], which is TVD when combined with forward Euler under suitable CFL condition. We consider

two different 2-stage Runge-Kutta methods

$$\mathbf{u}^{(1)} = \mathbf{u}^n - 20\Delta t\mathbf{L}(\mathbf{u}^n), \quad \mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{40}(41\mathbf{L}(\mathbf{u}^n) - \mathbf{L}(\mathbf{u}^{(1)})), \tag{2.8}$$

and

$$\mathbf{u}^{(1)} = \mathbf{u}^n + \Delta t\mathbf{L}(\mathbf{u}^n), \quad \mathbf{u}^{n+1} = \frac{1}{2}\mathbf{u}^n + \frac{1}{2}(\mathbf{u}^{(1)} + \Delta t\mathbf{L}(\mathbf{u}^{(1)})). \tag{2.9}$$

Both of these methods are of second order accuracy in time. It is worth observing that (2.9) can be represented as a convex combination of forward Euler steps (and therefore SSP [4]), while this is not the case for (2.8).

Figure 2.1 shows the results by the two different time-stepping methods at $t = 0.3$, with $N = 200$ mesh points under $CFL = 0.5$. It is seen that the solution by the non-SSP time integration method (2.8) has an overshoot near the shock. This numerical experiment shows that it is advantageous to use SSP time integration methods, as they better ensure the provable stability and do not increase the computational cost.

In fact, as pointed out in [4] the key feature of SSP time discretizations is the convexity, i.e. they can be represented as convex combinations of forward Euler steps. With this property, the SSP methods preserve (do not increase) not only the TV norm, but *any* convex function. Once the forward Euler time discretization is shown to be strongly stable under an arbitrary norm

$$||\mathbf{u}^n + \Delta t\mathbf{L}(\mathbf{u}^n)|| \leq ||\mathbf{u}^n||,$$

the high order SSP methods can be used to ensure the stability in the same norm.

Hence, the class of SSP methods have a wide range of applicability. For example, the weighted L^2 SSP higher order discretizations of spectral schemes were discussed in [14]; the L^∞ SSP higher order discretizations were discussed for discontinuous Galerkin and central schemes in [15] and [12]. Due to the preservation for arbitrary convex function, SSP time discretizations can be used to preserve the entropy stability, see e.g. [16] and [17].

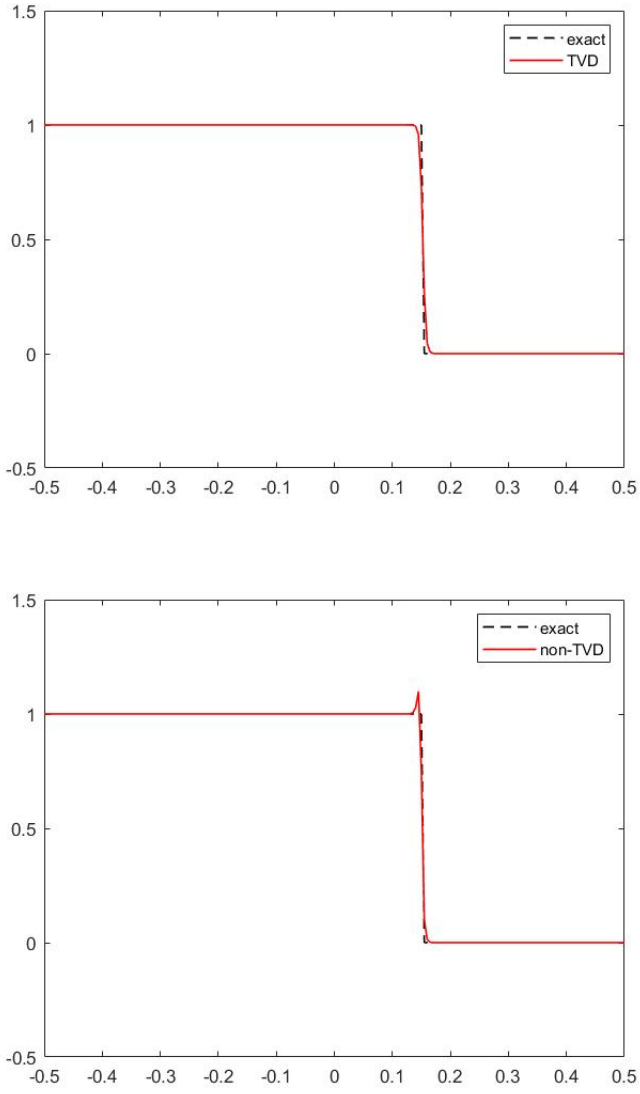


Figure 2.1: SSP method (2.9) vs non-SSP method (2.8), second order MUSCL in space, $N = 200$, $t = 0.3$, $CFL = 0.5$.

2.2 Runge-Kutta Methods

Our task is to approximate the solution $\mathbf{u} : [0, T] \mapsto \mathbb{R}^N$ of the initial value problem

$$\frac{d\mathbf{u}(t)}{dt} = \mathbf{L}(\mathbf{u}(t), t), \quad \mathbf{u}(0) = \mathbf{u}_0,$$

where $\mathbf{L} : \mathbb{R}^N \times [0, T] \mapsto \mathbb{R}^N$ is a sufficiently regular function. The idea of Runge-Kutta methods is to proceed from t_n to $t_{n+1} = t_n + \Delta t$ by approximating the integral in the exact solution

$$\mathbf{u}(t_{n+1}) = \mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} \mathbf{L}(\mathbf{u}(t), t)dt.$$

We choose quadrature points $c_1, \cdots, c_s$ and their weights $b_1, \cdots b_s$, the quadrature formula reads

$$\mathbf{u}(t_{n+1}) = \mathbf{u}(t_n) + \Delta t \sum_{i=1}^{s} b_u \mathbf{L}(\mathbf{u}(t_n + c_i\Delta t), t_n + c_i\Delta t) + error. \tag{2.10}$$

Suppose we have approximation $\mathbf{u}^n$ to $\mathbf{u}(t_n)$, and we approximate the values of $\mathbf{u}(t_n + c_i\Delta t)$ by $\mathbf{u}^{(i)}$, which are computed by numerical quadratures

$$\mathbf{u}^{(i)} = \mathbf{u}^n + \Delta t \sum_{j=1}^{s} a_{ij} \mathbf{L}(\mathbf{u}^{(j)}, t_n + c_j\Delta t), \quad i = 1, \cdots, s, \tag{2.11}$$

then we obtain the representation of a general $s-$stage Runge-Kutta (RK) method

$$\begin{aligned}\mathbf{u}^{(i)} &= \mathbf{u}^n + \Delta t \sum_{j=1}^{s} a_{ij}\mathbf{L}(\mathbf{u}^{(j)}, t_n + c_j\Delta t), \quad i = 1, \cdots, s, \\ \mathbf{u}^{n+1} &= \mathbf{u}^n + \Delta t \sum_{i=1}^{s} b_i\mathbf{L}(\mathbf{u}^{(i)}, t_n + c_i\Delta t),\end{aligned} \tag{2.12}$$

where $\mathbf{u}^{(1)}, \cdots, \mathbf{u}^{(s)}$ are the s intermediate stage values. The method (2.12) can be expressed in the form of the Butcher array

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} . \tag{2.13}$$

The Runge-Kutta method is explicit if the matrix (a_{ij}) is strictly lower triangular, and semi-implicit if (a_{ij}) is lower triangular. Otherwise, the RK method is fully implicit. A popular class of semi-implicit methods is the class of *diagonally-implicit* Runge-Kutta methods (DIRK), of which the diagonal entries a_{ii} are consistent.

In the following discussions, we assume for simplicity that the problem is autonomous, $\mathbf{L}(\mathbf{u}(t))$, otherwise we can consider

$$\frac{d}{dt}\begin{bmatrix} \mathbf{u}(t) \\ t \end{bmatrix} = \begin{bmatrix} \mathbf{L}(\mathbf{u}(t)) \\ 1 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{u}(0) \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_0 \\ 0 \end{bmatrix}.$$

For convenience of analysis, [7] represented a general $s-$stage RK method (2.12) in the

form

$$\begin{aligned}
\mathbf{u}^{(0)} &= \mathbf{u}^n, \\
\mathbf{u}^{(i)} &= \sum_{j=0}^{s+1} \alpha_{ij}\mathbf{u}^{(j)} + \Delta t \beta_{ij}\mathbf{L}(\mathbf{u}^{(j)}), \quad i = 1, \cdots, s+1, \\
\mathbf{u}^{n+1} &= \mathbf{u}^{(s+1)},
\end{aligned} \tag{2.14}$$

which is known as the *Shu-Osher* form of the Runge-Kutta method. In the following, we discuss explicit and implicit Runge-Kutta methods separately.

2.2.1 Explicit SSP Runge-Kutta methods

An explicit Runge-Kutta method admits the Shu-Osher form

$$\begin{aligned}
\mathbf{u}^{(0)} &= \mathbf{u}^n, \\
\mathbf{u}^{(i)} &= \sum_{j=0}^{i-1} \alpha_{ij}\mathbf{u}^{(j)} + \Delta t \beta_{ij}\mathbf{L}(\mathbf{u}^{(j)}), \quad i = 1, \cdots, s+1, \\
\mathbf{u}^{n+1} &= \mathbf{u}^{(s+1)}.
\end{aligned} \tag{2.15}$$

The central result is as follows.

Theorem 2.1 ([8]) *If the foward Euler method* (2.2) *is strongly stable under the CFL restriction* $\Delta t \leq \Delta t_{FE}$*, then the* explicit *Runge-Kutta method* (2.15) *with semipositive* α_{ij}*,* β_{ij} *is SSP,* $||\mathbf{u}^{n+1}|| \leq ||\mathbf{u}^n||$*, under the following CFL restriction*

$$\Delta t \leq k \Delta t_{FE}, \quad k = \min_{i,j} \frac{\alpha_{ij}}{\beta_{ij}}, \quad i = 1 \cdots s+1, \;\; j = 0 \cdots i-1.$$

The proof of Theorem 2.1 is given in Appendix A. The theorem indicates that if a RK method admits semipositive α_{ij} and β_{ij}, i.e. the scheme can be represented as a convex combiniation of forward Euler steps, then the strong stability follows from the stability of forward Euler. The assumption that $\beta_{ij} \geq 0$ may be relaxed under some circumstances [8], but the positivity of α_{ij} is necessary. The optimal construction of explicit SSP Runge-Kutta methods up to fourth order can be found in [9].

2.2.2 Implicit SSP Runge-Kutta methods

It is of interest to seek for coefficients that maximize the SSP coefficient. Gottlieb et al [4] have shown that the SSP coefficient of an $m-$stage, $mth-$order SSP Runge-Kutta method cannot exceed one. To loosen the time step restriction, one may resort to implicit methods.

We consider semi-implicit Runge-Kutta methods of the form

$$\begin{aligned}
\mathbf{u}^{(0)} &= \mathbf{u}^n, \\
\mathbf{u}^{(i)} &= \sum_{j=0}^{i-1} \alpha_{ij}\mathbf{u}^{(j)} + \Delta t \sum_{j=0}^{i} \beta_{ij}\mathbf{L}(\mathbf{u}^{(j)}), \quad i = 1, \cdots, s+1, \\
\mathbf{u}^{n+1} &= \mathbf{u}^{(s+1)}.
\end{aligned} \tag{2.16}$$

The simplest example is the first order backward Euler scheme

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathbf{L}(\mathbf{u}^{n+1}), \tag{2.17}$$

which is *unconditionally* SSP. To see this, we take $\gamma > 0$ and consider [18]

$$(1+\gamma)\mathbf{u}^{n+1} = \mathbf{u}^n + \gamma\mathbf{u}^{n+1} + \Delta t\mathbf{L}(\mathbf{u}^{n+1}),$$

and rewrite it as

$$\mathbf{u}^{n+1} = \frac{1}{1+\gamma}[\mathbf{u}^n + \gamma(\mathbf{u}^{n+1} + \frac{1}{\gamma}\Delta t\mathbf{L}(\mathbf{u}^{n+1}))].$$

Provided $\Delta t \leq \gamma\Delta t_{FE}$, we have $||\mathbf{u}^{n+1} + \frac{1}{\gamma}\Delta t\mathbf{L}(\mathbf{u}^{n+1})|| \leq ||\mathbf{u}^{n+1}||$ if the forward Euler is strongly stable for $\Delta t \leq \Delta t_{FE}$. Then we have

$$||\mathbf{u}^{n+1}|| \leq \frac{1}{1+\gamma}(||\mathbf{u}^n|| + \gamma||\mathbf{u}^{n+1}||),$$

the strong stability $||\mathbf{u}^{n+1}|| \leq ||\mathbf{u}^n||$ follows. Since γ can be arbitrary, the strong stability of backward Euler holds for arbitrarily large time steps, Δt.

Unfortunately, this result does not hold beyond methods of first order. Gottlieb, Shu and Tadmor [4] have shown the nonexistence of unconditionally SSP implicit RK methods of order higher than one. In general, following similar approach in [8] we have the following theorem.

Theorem 2.2 *If the foward Euler method* (2.2) *is strongly stable under the CFL restriction* $\Delta t \leq \Delta t_{FE}$*, then the* implicit *Runge-Kutta method* (2.16) *with semipositive* α_{ij}*,* β_{ij} *is SSP*

under the CFL restriction

$$\Delta t \le k\Delta t_{FE}, \quad k = \min_{i,j} \frac{\alpha_{ij}}{\beta_{ij}}, \quad i = 1 \cdots s+1, \;\; j = 0 \cdots i-1.$$

The proof of Theorem 2.2 is given in Appendix B.

2.3 Multistep Methods

Consider $s-$step multistep methods of the form

$$\mathbf{u}^{n+1} = \sum_{i=0}^{s} (\alpha_i \mathbf{u}^{n+1-i} + \Delta t \beta_i \mathbf{L}(\mathbf{u}^{n+1-i})) = \sum_{i=0}^{s} \alpha_i (\mathbf{u}^{n+1-i} + \Delta t \frac{\beta_i}{\alpha_i} \mathbf{L}(\mathbf{u}^{n+1-i})), \quad (2.18)$$

where $\mathbf{u}^{n+1-i}$ refers to the solution at past time t_{n+1-i}. The scheme is explicit $\alpha_0 = \beta_0 = 0$; otherwise, it is implicit. Through Taylor expansion, the conditions for order p are recovered

$$\sum_{i=0}^{s} i^j \alpha_i = j \sum_{i=0}^{s} i^{j-1} \beta_j, \quad j = 1, \cdots, p.$$

The commonly used families of multistep methods include Adams–Bashforth methods, Adams–Moulton methods, and Backward Differentiation Formulas (BDF).

It is clear that if α_i, β_i are semipositive, under the CFL condition

$$\Delta t \le \min_i \frac{\alpha_i}{\beta_i} \Delta t_{FE},$$

the scheme is SSP in the sense that

$$||\mathbf{u}^{n+1}|| \leq \max_{i=1\cdots s} ||\mathbf{u}^{n+1-i}||, \quad ||\cdot|| \text{ is an arbitrary norm.}$$

Unfortunately, the existence of unconditionally SSP multistep methods of order higher than one is also ruled out [4]. Lenferink [19, 20] analyzed the efficiency of SSP multistep methods, the main results are as follows.

Theorem 2.3 ([19]) *The SSP coefficient of an* $s-$*step* explicit *multistep method of order* $p > 1$ *satisfies*

$$k \leq \frac{s-p}{s-1}.$$

The theorem indicates that $s > p$ and that the SSP coefficient of an explicit multistep scheme cannot exceed one. The situation is not improved significantly for implicit methods.

Theorem 2.4 ([20]) *For* $p > 1$*, any* implicit *SSP multistep method has an SSP coefficient less than 2.*

Constructions of high order implicit methods can be found in [20, 18]. Due to the large number of past steps and the restrictive SSP coefficient, the use of very high order SSP multistep methods is limited.

Chapter 3: Concepts of Time-Limiter Schemes

In this chapter, we introduce the concepts of time-limiter schemes proposed by Duraisamy, Baeder and Liu [5]. The objective of applying time-limiters is to allow the use of large time steps beyond the monotonicity limits for high order implicit time integration methods without introducing severe oscillations. The basic idea is to satisfy monotonicity conditions locally by dropping the order of accuracy in the regions where the solution behaves unsmoothly. In chapter 4, such idea will be applied to construct the L-DIRK3 scheme, which is the main work of this thesis. To help illustrate the application of time-limiters, the constructions of the L-Trap and the L-DIRK2 schemes will be breifly reviewed in section 3.2. In section 3.3, a few simple numerical examples are presented to show the effectiveness of these limited time integration schemes.

3.1 Motivation of Limited Time Integration

For nonlinear conservation laws, the solutions may develop discontinuities even if the initial data is smooth. Numerical schemes must be properly constructed to prevent numerical oscillations in the nonsmooth regions. The early stage framework of non-oscillatory numerical methods for conservation laws is the class of monotone schemes,

i.e. the schemes of the form

$$u_j^{n+1} = G(u_{j-p}^n, u_{j-p+1}^n, \cdots, u_{j+q}^n) \tag{3.1}$$

that satisfy

$$\frac{\partial G}{\partial u_i}(u_{j-p}, u_{j-p+1}, \cdots, u_{j+q}) \geq 0, \quad \forall j-p \leq i \leq j+q. \tag{3.2}$$

Harten et al [1] have shown that the order of accuracy of monotone schemes cannot exceed first order. Based on less restrictive monotonicity conditions, such as the commonly used TVD condition (2.4), many high order non-oscillatory schemes have been proposed in the last few decades and successfully applied in hyperbolic problems, for example, the UNO schemes [21], the MP schemes [22], and the ENO/WENO type schemes [3]. However, ensuring non-oscillatory solutions of these high order schemes may impose severe time step restriction when used with explicit time integration. One way to resolve this is to use implicit time integration. But as discussed in chapter 2, the higher order implicit time integration methods can at most be conditionally SSP. Thus, the objective in using implicit time integration methods is lost since the allowable time steps are not much larger than those for explicit methods. Table 3.1 presents the SSP coefficients of some implicit time integration methods for the system of ODEs (2.1). The SSP coefficients can be computed according to Theorem 2.1–2.2. We observe that the high order implicit time integration methods may have restrictive SSP coefficients, or even worse, may not be SSP, even if they are $A-$stable.

To allow the use of time steps beyond the monotonicity limits of these high order implicit time integration schemes, Duraisamy, Baeder and Liu [5] proposed the concept of limited time integration. By applying time-limiters, the attempt is made at locally satisfying the TVD condition in the regions of high solution gradients, while maintaining high order of accuracy in time in the smooth regions.

Method	State update formula	SSP coefficient (k)
Imp. Euler	$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathbf{L}(\mathbf{u}^{n+1})$	∞
Imp. Trap.	$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2}(\mathbf{L}(\mathbf{u}^n) + \mathbf{L}(\mathbf{u}^{n+1}))$	2
Imp. BDF2	$\mathbf{u}^{n+1} = \frac{4\mathbf{u}^n - \mathbf{u}^{n-1} + 2\Delta t \mathbf{L}(\mathbf{u}^{n+1})}{3}$	not SSP
DIRK2	$\mathbf{u}^{(1)} = \mathbf{u}^n + \gamma \Delta t \mathbf{L}(\mathbf{u}^{(1)})$	
($\gamma = \frac{2-\sqrt{2}}{2}$)	$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t[(1-\gamma)\mathbf{L}(\mathbf{u}^{(1)}) + \mathbf{L}(\mathbf{u}^{n+1})]$	$\frac{1}{1-2\gamma} \approx 2.4142$

Table 3.1: SSP coefficients of some implicit time integration methods.

3.2 Review of Second Order Time-Limiter Schemes

We review the Limited-Trapezoidal (L-Trap) and the Limited-2 Stage Diagonally Implicit Runge-Kutta (L-DIRK2) schemes as examples to illustrate the idea of time-limiter schemes.

3.2.1 The L-Trap Scheme

Consider the family of θ methods for conservation laws (1.1)

$$u_j^{n+1} = u_j^n - \tau[(1-\theta)(\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}}) + \theta(\hat{f}^{n+1}_{j+\frac{1}{2}} - \hat{f}^{n+1}_{j-\frac{1}{2}})], \quad \tau = \frac{\Delta t}{\Delta x}. \tag{3.3}$$

The scheme can be written in the conservative form

$$u_j^{n+1} = u_j^n - \tau(\bar{f}_{j+\frac{1}{2}} - \bar{f}_{j-\frac{1}{2}}),$$

where

$$\bar{f}_{j+\frac{1}{2}} = (1-\theta)\hat{f}^n_{j+\frac{1}{2}} + \theta\hat{f}^{n+1}_{j+\frac{1}{2}}.$$

If we fix a constant $\theta \in [0,1]$ over the whole domain, the scheme is SSP with the SSP coefficient given by $k = \frac{1}{1-\theta}$. In particular, we recover the forward Euler method for $\theta = 0$; the Trapezoidal method for $\theta = 0.5$; and the backward Euler method. The schemes are as shown in Table 3.2.

θ	Method	Accuracy	SSP coefficient
1.0	Imp. Euler	1st order	∞
2.0	Imp. Trap	2nd order	2.0
0.0	Exp. Euler	1st order	1.0

Table 3.2: Accuracy and SSP coefficients of θ methods.

To maintain the stability of implicit Euler method near discontinuities and extrema as well as the second order accuracy in time of the Trapezoidal method in the smooth regions, the L-Trap scheme applies the *locally* defined values of θ, the numerical flux becomes

$$\bar{f}_{j+\frac{1}{2}} = (1-\theta_{j+\frac{1}{2}})\hat{f}^n_{j+\frac{1}{2}} + \theta_{j+\frac{1}{2}}\hat{f}^{n+1}_{j+\frac{1}{2}}, \quad \theta_{j+\frac{1}{2}} = \frac{\theta_j + \theta_{j+1}}{2}.$$

Here, θ_j is the *time-limiter* at point x_j, which is a real number between 0.5 and 1.0.

The idea is to define a θ_j such that in the smooth regions, we would apply $\theta_j \approx 0.5$

and thus attain second order accuracy in time locally. In the regions of large solution gradients, we would apply $\theta_j \approx 1.0$, thus locally dropping the order of accuracy to first order but improving the stability under large time steps.

3.2.2 The L-DIRK2 Scheme

The L-DIRK2 scheme is given by

$$u_j^{(1)} = u_j^n - \tau\gamma(\hat{f}^{(1)}_{j+\frac{1}{2}} - \hat{f}^{(1)}_{j-\frac{1}{2}}),$$
$$u_j^{n+1} = u_j^n - \tau[(a_{21,j+\frac{1}{2}}\hat{f}^{(1)}_{j+\frac{1}{2}} - a_{21,j-\frac{1}{2}}\hat{f}^{(1)}_{j-\frac{1}{2}}) + (a_{22,j+\frac{1}{2}}\hat{f}^{n+1}_{j+\frac{1}{2}} - a_{22,j-\frac{1}{2}}\hat{f}^{n+1}_{j-\frac{1}{2}})],$$

where

$$a_{21,j+\frac{1}{2}} = \gamma + \theta_{j+\frac{1}{2}}(1-2\gamma),$$
$$a_{22,j+\frac{1}{2}} = (1-\gamma) + \theta_{j+\frac{1}{2}}(2\gamma-1),$$
$$\theta_{j+\frac{1}{2}} = \frac{\theta_j + \theta_{j+1}}{2}, \;\; \theta_j \in [0,1], \;\; \gamma = \frac{2-\sqrt{2}}{2}.$$

For constant $\theta_j = \theta$, the scheme written in the Shu-Osher form

$$\mathbf{u}^{(1)} = \mathbf{u}^n + \gamma\Delta t\mathbf{L}(\mathbf{u}^{(1)}),$$
$$\mathbf{u}^{n+1} = \mathbf{u}^{(1)} + \theta(1-2\gamma)\Delta t\mathbf{L}(\mathbf{u}^{(1)}) + [(1-\gamma) + \theta(2\gamma-1)]\Delta t\mathbf{L}(\mathbf{u}^{n+1}).$$

The scheme is SSP for $\theta \in [0,1]$, and the SSP coefficient is given by $k = \frac{1}{\theta(1-2\gamma)}$. The scheme reduces to the second order DIRK2 when $\theta = 1.0$, with $k = 2.4142$, and to a first order 2-stage unconditionally SSP method if $\theta = 0.0$. Hence, we would like to use $\theta_j \approx 1.0$ in the smooth regions, and to use $\theta_j \approx 0.0$ in the regions where the solution behaves unsmoothly.

3.2.3 Construction of Time-Limiters

There are various possible ways to construct the time-limiter, θ_j. In the paper [5], the time-limiter θ_j is constructed as follows: for scalar problems we define the indicator function

$$r_j = minmod(\frac{2s^{n+\frac{1}{2}}}{L_j(\mathbf{u}^n)+\epsilon}, \frac{2s^{n+\frac{1}{2}}}{L_j(\mathbf{u}^{n+1})+\epsilon}, 1), \tag{3.4}$$

where ϵ is a small positive number (say $\epsilon = 10^{-10}$) introduced to enhance computational stability, and

$$s_j^{n+\frac{1}{2}} = \frac{u_j^{n+1}-u_j^n}{\Delta t}, \quad L_j(\mathbf{u}^n) = -\frac{\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}}}{\Delta x}.$$

Then the time-limiter is defined as $\theta_j = 1.0 - 0.5r_j$ for the L-Trap scheme and $\theta_j = r_j$ for the L-DIRK2 scheme. Here, the indicator r_j detects the temporal monotonicity of the solution. This construction is inspired by Huynh's lemma [23]: *Given the data* $f(q_1)$, $f(q_2)$ *and the derivatives* $f'(q_1)$ *and* $f'(q_2)$ *at points* q_1 *and* q_2, *the quadratic interpolant* $f(q)$ *is monotone in* $[q_1, q_2]$ *if* $f'(q_1), f(q_2) \in [0, 2s]$, *where* $s = \frac{f(q_2)-f(q_1)}{q_2-q_1}$. Thus, if $u_j(t)$ is monotonic in $t \in [t^n, t^{n+1}]$ we would have $r_j \approx 1.0$, which recovers the higher order mode. Otherwise, the change in monotonicity would result in $r_j \approx 0.0$, the scheme is then switched to the first order unconditionally SSP mode.

In scalar problems, all the quantities involved in the construction (3.4) are computed as part of the solution update. The extra computation cost of constructing limiter is minimal. For systems of equations, one may take an arbitrary variable to construct the time-limiter. In particular, it is preferred to use the quantity which is expected to encounter the most significant jumps. The optimal reference variable, however, may not be taken

from conservative variables, such as pressure and velocity for Euler equations. In this case, the evaluation of corresponding time derivative function, L_j, may require expensive extra work in complicated problems.

To allow more flexible choice of reference variable with minimal influence on efficiency, we propose a convenient alternative construction. For the construction of L-DIRK2, we define

$$\begin{aligned} &\theta_j = minmod(r_j, 1), \text{ where} \\ &r_j = \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{u_{j+1}^{n} - u_{j-1}^{n}}, \quad k = 1, 2, \quad j = 1, 2, \cdots N. \end{aligned} \tag{3.5}$$

For the L-Trap scheme we reset the limiter with $1.0 - 0.5\theta_j$. The indicator $r_j \approx \dfrac{\partial_x u_j^{n+1}}{\partial_x u_j^n}$ is used to detect the change in the solution monotonicity at point x_j. In the smooth regions where $r_j \approx 1.0$, we would have $\theta_j \approx 1.0$ and (roughly) recover the higher order method. If the solution changes the monotonicity at point x_j when proceeds from t^n to t^{n+1}, which usually implies the occurrence of discontinuities or extrema, r_j would be closed to zero or become negative. Then we have $\theta_j \approx 0.0$ and recover the first order method. It is noticed that the new construction does not require the evaluations of time derivative function and therefore saves the computation cost.

3.3 Numerical Examples

We illustrate the performance of the L-Trap and the L-DIRK2 through a few scalar examples. In the following computations, the spatial reconstructions are obtained from the 5th order WENO-JS finite volume methods [3]. The time-limiters are constructed with (3.5). The number of equally spaced mesh points in the domain is represented by N.

The first example is the linear advection equation

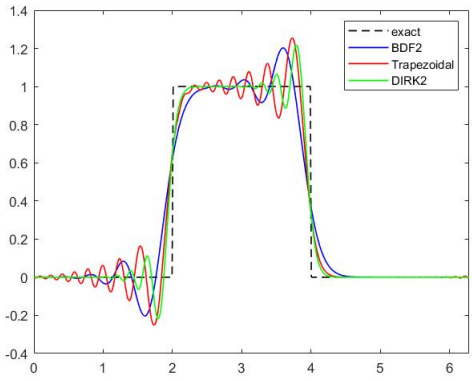
$$u_t + u_x = 0, \quad u_0(x) = \begin{cases} 1, & 2 < x < 4 \\ 0, & \text{else} \end{cases}$$

with periodic boundary conditions over the domain $[0, 2\pi]$. Figure 3.3 presents the after one period of revolution under the mesh size $N = 400$ at $CFL = 2$. It is seen that the Trapezoidal, BDF2 and DIRK2 schemes generate severe oscillations near the edges of the square wave, while the L-Trap and the L-DIRK2 schemes present no visible oscillation and are much less dissipative than the first order implicit Euler method.

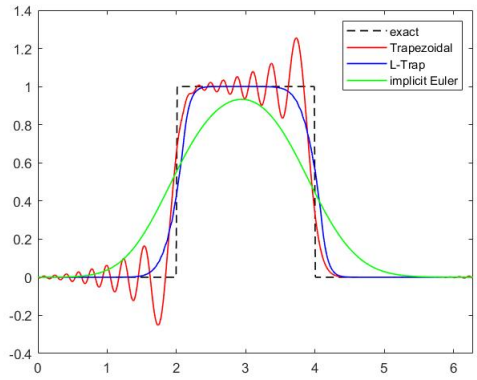
The second example is the one-dimensional Burger's equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) = 0$$

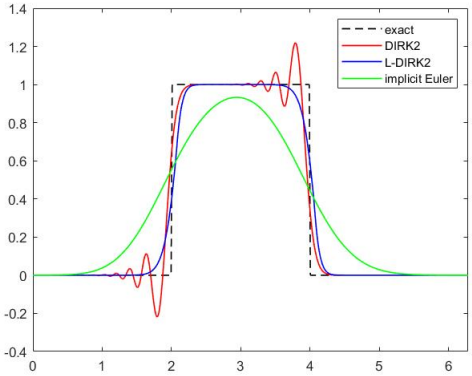
with periodic boundary conditions over the domain $[0, 2\pi]$, of the initial data comprises of an expansion wave and a compression wave. Figure 3.3 presents the solutions at $t = 2$ under the mesh size $N = 100$ at $CFL = 3$. It is seen that the solutions by the second order linear time integration methods present strong oscillations near the shock. The time-



(a) Linear Time Integration Methods



(b) L-Trap

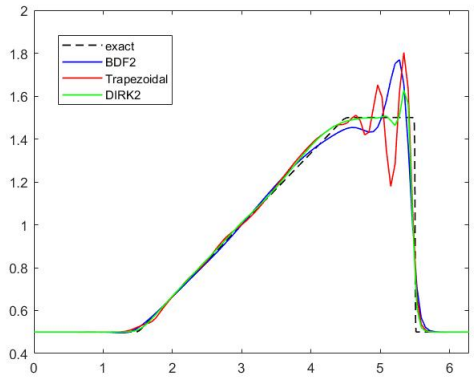


(c) L-DIRK2

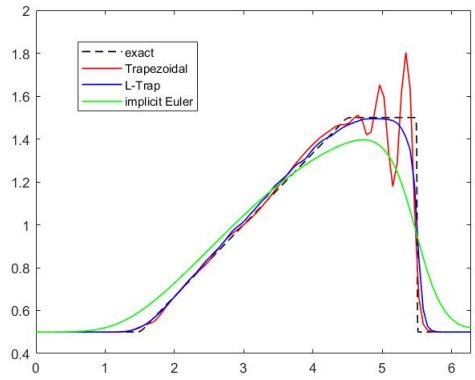
Figure 3.1: Linear advection, square wave, 2nd order time integration, 5th order WENO in space, N = 400, CFL = 2, one period of revolution.

limiter schemes are much less oscillatory than the unlimited second order methods, and their resolution is much better than the first order method.

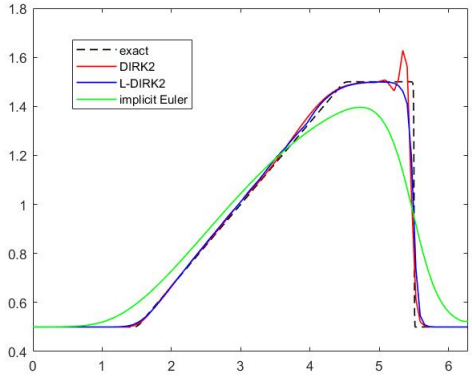
These examples show the significantly improved stability of the L-Trap and the L-DIRK2 schemes. More test cases in one dimension can be found [5]. In chapter 5, numerical studies for multidimensional problems will be presented.



(a) Linear Time Integration Methods



(b) L-Trap



(c) L-DIRK2

Figure 3.2: 1D Burger's equation, expansion and compression wave, 2nd order time integration, 5th order WENO in space, N = 100, CFL = 3, t = 2.

3.4 Unconditionally SSP variant of TR-BDF2

Besides the L-DIRK2 and the L-Trap schemes proposed by Duraisamy et al, we also mention the more recent effort by Bonaventura and Della Rocca [24] to improve the stability of the TR-BDF2 (Trapezoidal Rule-Backward Difference Formula 2) scheme. This method was first proposed by Bank et al [25] and then further reformulated and analyzed in [26]. The TR-BDF2 method consists of a stage of trapezoidal rule followed

by a stage of BDF2,

$$\begin{aligned}\mathbf{u}^{n+\gamma} - \frac{\gamma}{2}\Delta t\mathbf{L}(\mathbf{u}^{n+\gamma}) &= \mathbf{u}^n + \frac{\gamma}{2}\Delta t\mathbf{L}(\mathbf{u}^n),\\ \mathbf{u}^{n+1} - \frac{1-\gamma}{2-\gamma}\Delta t\mathbf{L}(\mathbf{u}^{n+1}) &= \frac{1}{\gamma(2-\gamma)}\mathbf{u}^{n+\gamma} - \frac{(1-\gamma)^2}{\gamma(2-\gamma)}\mathbf{u}^n,\end{aligned} \tag{3.6}$$

where the parameter $\gamma \in (0,1)$ determines the stability and monotonocity properties. The rationale of TR-BDF2 is to combine the high accuracy of the trapezoidal rule with the stability and damping of BDF2. Particularly, when $\gamma = 2-\sqrt{2}$ is taken both stages have the same Jacobian [25]. Following the lines of [26] the method can be written as a 3-stage second order DIRK method, as given by the Butcher array

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \frac{\gamma}{2} & \frac{\gamma}{2} & 0 \\ 1 & \frac{1}{2(2-\gamma)} & \frac{1}{2(2-\gamma)} & \frac{1-\gamma}{2-\gamma} \\ \hline & \frac{1}{2(2-\gamma)} & \frac{1}{2(2-\gamma)} & \frac{1-\gamma}{2-\gamma} \end{array}. \tag{3.7}$$

The TR-BDF2 scheme was proved to be L-stable [25] and S-stable [26]. However, in spite of the well guaranteed linear stability, the solution may still generate strong oscillations when the problem is not smooth. To obtain unconditional strong stability preservation, a hybrid method was constructed in [24] by introducing a weight parameter

$\alpha \in [0,1]$ in both stages,

$$\begin{aligned}
&\mathbf{u}^{n+\gamma} - \gamma\Delta t(1-\frac{\alpha}{2})\Delta t\mathbf{L}(\mathbf{u}^{n+\gamma}) = \mathbf{u}^n + \frac{\alpha}{2}\Delta t\mathbf{L}(\mathbf{u}^n),\\
&\mathbf{u}^{n+1} - \frac{1-\gamma}{\alpha(1-\gamma)+1}\Delta t\mathbf{L}(\mathbf{u}^{n+1})\\
&\qquad = \frac{\alpha(\frac{1}{\gamma}-1)+1}{\alpha(1-\gamma)+1}\mathbf{u}^{n+\gamma} - \frac{\alpha}{\alpha(1-\gamma)+1}\frac{(1-\gamma)^2}{\gamma}\mathbf{u}^n.
\end{aligned} \tag{3.8}$$

When $\alpha = 1$ we recover the original TR-BDF2 (3.6). When $\alpha = 0$, the scheme is equivalent to the succession of two implicit Euler steps of length $\gamma\Delta t$ and $(1-\gamma)\Delta t$, which is unconditionally SSP.

The hybrid method (3.8) can be written as a DIRK scheme associated with the Butcher array

$$\begin{array}{c|ccc}
0 & 0 & 0 & 0\\
\gamma & \alpha\frac{\gamma}{2} & \gamma(1-\frac{\alpha}{2}) & 0\\
1 & \frac{\alpha}{2}\frac{\alpha(1-\gamma)+\gamma}{\alpha(1-\gamma)+1} & (1-\frac{\alpha}{2})\frac{\alpha(1-\gamma)+\gamma}{\alpha(1-\gamma)+1} & \frac{1-\gamma}{\alpha(1-\gamma)+1}\\
\hline
 & \frac{\alpha}{2}\frac{\alpha(1-\gamma)+\gamma}{\alpha(1-\gamma)+1} & (1-\frac{\alpha}{2})\frac{\alpha(1-\gamma)+\gamma}{\alpha(1-\gamma)+1} & \frac{1-\gamma}{\alpha(1-\gamma)+1}
\end{array}.$$

For $\alpha = 0$, we obtain the first order unconditionally monotone implicit Euler method,

$$\begin{array}{c|ccc}
0 & 0 & 0 & 0\\
\gamma & 0 & \gamma & 0\\
1 & 0 & \gamma & 1-\gamma\\
\hline
 & 0 & \gamma & 1-\gamma
\end{array},$$

while for $\alpha = 1$ the TR-BDF2 tableau (3.7) is recovered. It is shown in [24] that the

hybrid TR-BDF2 method maintains L-stability for any value of $\alpha \in [0, 1]$.

To ensure the unconditionally SSP property, the strategy used by [24] is to define a global sensor function which evaluates the smoothness of the solution,

$$\sigma = s_g(\mathbf{u}^{n+1}) = \begin{cases} 1 & \text{if } ||\mathbf{u}^{n+1}|| > M \\ 0 & \text{otherwise} \end{cases}.$$

Here $||\cdot||$ is an arbitrary norm depending on our interest. At each time step, a tentative solution, $\mathbf{u}^*$, is computed with the original TR-BDF2 ($\alpha = 1$). If the sensor $\sigma = s_g(\mathbf{u}^*) = 0$, the tentative solution will be accepted as the new solution, $\mathbf{u}^{n+1} = \mathbf{u}^*$. If $\sigma = 1$, then the solution by TR-BDF2 violates the stability constraint, $||\mathbf{u}^{n+1}|| \leq M$. In this case, we set $\alpha = 0$ and repeat the current integration with implicit Euler method. Here, the weight parameter α enables the switch between a higher order and a first order methods, it plays a similar role with the time-limiters in the aforementioned L-Trap and L-DIRK2 schemes. In particular, if one employs the total variation semi-norm, $||\cdot||_{TV}$, and set the bound $M = ||\mathbf{u}^{n+1}||$, the method would be unconditionally TVD. For systems of equations, the construction can be extended in a componentwise manner.

The improved stability of the hybrid TR-BDF2 was verified through a series numerical examples [24]. However, the phenomenon of order reduction is reported at large time step due to the use of implicit Euler. Moreover, when solving non-smooth problems, the original TR-BDF2 may be too oscillatory such that the algorithm is restarted with implicit Euler at almost every time step, which leads to the risk of doubling the computation cost. To achieve a better compromise of the accuracy, stability and efficiency, one may borrow

the idea of local time limiting from the time-limiter schemes. We will leave it for future study.

Chapter 4: The Limited-DIRK3 Scheme

When a high order space discretization is applied in unsteady-state computations, applying a third order order or even higher order time integration method would better preserve the high resolution. In this chapter, the idea of limited time integration is applied to construct the L-DIRK3 scheme. The extension of the scheme to multi-dimensions is also presented.

4.1 Review of the DIRK3 Scheme

Before starting the construction of L-DIRK3, we briefly review the properties of the DIRK3 scheme and explain the motivation of applying time-limiters to DIRK3. The DIRK3 time integration scheme can be represented by the Butcher array

$$\begin{array}{c|ccc} \alpha & \alpha & 0 & 0 \\ \tau_2 & \tau_2 - \alpha & \alpha & 0 \\ 1 & b_1 & b_2 & \alpha \\ \hline & b_1 & b_2 & \alpha \end{array}, \tag{4.1}$$

where $\alpha \approx 0.435866521508459$ is the root of $x^3 - 3x^2 + \frac{3}{2}x - \frac{1}{6} = 0$ lying in $(\frac{1}{6}, \frac{1}{2})$,

$$\begin{aligned} \tau_2 &= \frac{1+\alpha}{2}, \\ b_1 &= -\frac{6\alpha^2 - 16\alpha + 1}{4}, \\ b_3 &= \frac{6\alpha^2 - 20\alpha + 5}{4}. \end{aligned} \tag{4.2}$$

It was shown in [6] that (4.1) is the unique *strongly S-stable* DIRK formula of order three in three stages. The S-stability of RK methods is defined as follows.

S-stability ([27]). A RK method is *S-stable* if for any bounded $g : [0, T] \mapsto \mathbb{R}$ having a bounded derivative, and any positive constant λ_0, there is a positive constant h_0 such that the numerical solution $\{y_n\}$, computed with time step h, to the equation

$$y' = g'(t) + \lambda(y - g(t))$$

satisfies

$$\left|\frac{y_{n+1} - g(t_{n+1})}{y_n - g(t_n)}\right| < 1$$

provided $y_n \neq g(t_n)$ for all $0 < h < h_0$ all complex λ with $Re(-\lambda) \geq \lambda_0$.

A RK method is *strongly S-stable* if

$$\frac{y_{n+1} - g(t_{n+1})}{y_n - g(t_n)} \to 0$$

as $Re(-\lambda) \to \infty$ for all $h > 0$ such that $[t_n, t_{n+1}] \subset [0, T]$.

The concept of $S-$stability was introduced in [27] to analyze the stability of a method when applied to problems of stiff equations. We notice that S-stability is stronger than A-stability, which can be recovered by taking $g \equiv 0$. However, being S-stable does *not* ensure non-oscillatory solutions. In fact, the oscillatory nature of DIRK3 can be expected. We look at the first two stages of DIRK3

stage 1:

$$\mathbf{u}^{(1)} = \mathbf{u}^n + \alpha \Delta t \mathbf{L}(\mathbf{u}^{(1)}).$$

stage 2:

$$\begin{aligned} \mathbf{u}^{(2)} &= \mathbf{u}^n + (\tau_2 - \alpha)\Delta t \mathbf{L}(\mathbf{u}^{(1)}) + \alpha \Delta t \mathbf{L}(\mathbf{u}^{(2)}) \\ &= \mathbf{u}^{(1)} + (\tau_2 - 2\alpha)\Delta t \mathbf{L}(\mathbf{u}^{(1)}) + \alpha \Delta t \mathbf{L}(\mathbf{u}^{(2)}). \end{aligned}$$

The first stage is an unconditionally TVD implicit Euler step, but the second stage includes a backward time stepping of $\mathbf{u}^{(1)}$ (notice that $\tau_2 - 2\alpha < 0$), which not TVD and may lead to numerical oscillations. Such instability motivates the application of time-limiters to enhance the stability of the scheme.

4.2 Construction of the L-DIRK3 Scheme

Now we construct the L-DIRK3 scheme for conservation laws (1.1). Denote $\tau := \dfrac{\Delta t}{\Delta x}$, the scheme is given by

$$\begin{aligned}
u_j^{(1)} &= u_j^n - \tau\alpha(\hat{f}^{(1)}_{j+\frac{1}{2}} - \hat{f}^{(1)}_{j-\frac{1}{2}}) \\
u_j^{(2)} &= u_j^n - \tau[(a_{21,j+\frac{1}{2}}\hat{f}^{(1)}_{j+\frac{1}{2}} - a_{21,j-\frac{1}{2}}\hat{f}^{(1)}_{j-\frac{1}{2}}) + (a_{22,j+\frac{1}{2}}\hat{f}^{(2)}_{j+\frac{1}{2}} - a_{22,j-\frac{1}{2}}\hat{f}^{(2)}_{j-\frac{1}{2}})] \\
u_j^{n+1} &= u_j^n - \tau[(a_{31,j+\frac{1}{2}}\hat{f}^{(1)}_{j+\frac{1}{2}} - a_{31,j-\frac{1}{2}}\hat{f}^{(1)}_{j-\frac{1}{2}}) + (a_{32,j+\frac{1}{2}}\hat{f}^{(2)}_{j+\frac{1}{2}} - a_{32,j-\frac{1}{2}}\hat{f}^{(2)}_{j-\frac{1}{2}}) \\
&\quad + (a_{33,j+\frac{1}{2}}\hat{f}^{n+1}_{j+\frac{1}{2}} - a_{33,j-\frac{1}{2}}\hat{f}^{n+1}_{j-\frac{1}{2}})]
\end{aligned} \qquad (4.3)$$

where

$$\begin{aligned}
a_{21,j+\frac{1}{2}} &= \alpha + \theta^{(1)}_{j+\frac{1}{2}}(\tau_2 - 2\alpha) \\
a_{22,j+\frac{1}{2}} &= \frac{1-\alpha}{2} + \theta^{(1)}_{j+\frac{1}{2}}(\frac{3\alpha-1}{2}) \\
a_{31,j+\frac{1}{2}} &= \alpha + \theta^{(2)}_{j+\frac{1}{2}}(b_1 - \alpha) \qquad , \quad \theta^{(k)}_{j+\frac{1}{2}} = \frac{\theta^{(k)}_j + \theta^{(k)}_{j+1}}{2}, \quad \theta^{(k)}_j \in [0,1]. \\
a_{32,j+\frac{1}{2}} &= \frac{1-\alpha}{2} + \theta^{(2)}_{j+\frac{1}{2}}(b_2 - \frac{1-\alpha}{2}) \\
a_{33,j+\frac{1}{2}} &= \frac{1-\alpha}{2} + \theta^{(2)}_{j+\frac{1}{2}}(\frac{3\alpha-1}{2})
\end{aligned}$$

The values of α, τ_2, b_1, b_2 are as defined in (4.2). $\theta^{(k)}_j$ is the time-limiter at point x_j at stage $k+1$. It is clear that the scheme (4.3) is conservative and consistent.

To illustrate the idea of construction, we first keep $\theta^{(k)}_j \equiv \theta$ as constant at all points

in the domain and at all stages. The resulting scheme can be written in the Butcher array

$$\begin{array}{c|ccc} \alpha & \alpha & 0 & 0 \\ \tau_2 & \alpha+\theta(\tau_2-2\alpha) & \frac{1-\alpha}{2}+\theta(\frac{3\alpha-1}{2}) & 0 \\ 1 & \alpha+\theta(b_1-\alpha) & \frac{1-\alpha}{2}+\theta(b_2-\frac{1-\alpha}{2}) & \frac{1-\alpha}{2}+\theta(\frac{3\alpha-1}{2}) \\ \hline & \alpha+\theta(b_1-\alpha) & \frac{1-\alpha}{2}+\theta(b_2-\frac{1-\alpha}{2}) & \frac{1-\alpha}{2}+\theta(\frac{3\alpha-1}{2}) \end{array} . \tag{4.4}$$

When $\theta = 1.0$ we recover the DIRK3 scheme (4.1), when $\theta = 0.0$ we obtain the unconditionally SSP successive implicit Euler steps,

$$\begin{array}{c|ccc} \alpha & \alpha & 0 & 0 \\ \tau_2 & \alpha & \frac{1-\alpha}{2} & 0 \\ 1 & \alpha & \frac{1-\alpha}{2} & \frac{1-\alpha}{2} \\ \hline & \alpha & \frac{1-\alpha}{2} & \frac{1-\alpha}{2} \end{array} , \tag{4.5}$$

or equivalently

$$\begin{aligned} \mathbf{u}^{(1)} &= \mathbf{u}^n + \alpha\Delta t L(\mathbf{u}^{(1)}), \\ \mathbf{u}^{(2)} &= \mathbf{u}^{(1)} + (\tau_2-\alpha)\Delta t L(\mathbf{u}^{(2)}) \cdot \\ \mathbf{u}^{n+1} &= \mathbf{u}^{(2)} + (1-\tau_2)\Delta t L(\mathbf{u}^{n+1}) \end{aligned}$$

We hope that the first order method (4.5) is applied locally near discontinuities and extrema, whereas in the smooth regions we would like to apply the third order accurate DIRK3 method. Hence, the local $\theta_j^{(k)}$ should be applied.

The constructions of time-limiters for second order methods, (3.4) and (3.5), can be

naturally extended to construct $\theta_j^{(k)}$. Here we follow the idea of limiter (3.5) and define

$$\begin{aligned} &\theta_j^{(k)} = minmod(r_j^{(k)}, 1), \text{ where} \\ &r_j^{(k)} = \frac{u_{j+1}^{(k+1)} - u_{j-1}^{(k+1)}}{u_{j+1}^{(k)} - u_{j-1}^{(k)}}, \quad k = 1, 2, \quad j = 1, 2, \cdots N. \end{aligned} \tag{4.6}$$

We use the indicator function $r_j^{(k)} \approx \dfrac{\partial_x u_j^{(k+1)}}{\partial_x u_j^{(k)}}$ to detect the change in the solution monotonicity at point x_j at stage k. If the solution is monotone when proceeds from stage k to stage $k+1$, we would expect $r_j^{(k)} \approx 1.0$, which recovers the DIRK3. Otherwise, we would have $\theta_j^{(k)} \approx 0.0$, the solution is updated with the implicit Euler stepping (4.5).

For systems of conservation laws, the concepts of monotonicity cannot be rigorously defined, the way of defining $r_j^{(k)}$ can be arbitrary. For example, for Euler equations we can use a density-based limiter

$$r_j^{(k)} = \frac{\rho_{j+1}^{(k+1)} - \rho_{j-1}^{(k+1)}}{\rho_{j+1}^{(k)} - \rho_{j-1}^{(k)}},$$

or a pressure-based limiter

$$r_j^{(k)} = \frac{p_{j+1}^{(k+1)} - p_{j-1}^{(k+1)}}{p_{j+1}^{(k)} - p_{j-1}^{(k)}}.$$

4.3 Multidimensional Extension

In this section, we discuss the formulation of the L-DIRK3 scheme in multi-dimensions. Without the loss of generality, we consider the two-dimensional conservation laws

$$\frac{\partial}{\partial t}u(x,y,t)+\frac{\partial}{\partial x}f(u(x,y,t))+\frac{\partial}{\partial y}g(u(x,y,t))=0. \tag{4.7}$$

The semidiscrete *finite difference* schemes for (4.7) are of the form

$$\frac{d}{dt}u_{i,j}(t)+\frac{\hat{f}_{i+\frac{1}{2},j}(t)-\hat{f}_{i-\frac{1}{2},j}(t)}{\Delta x}+\frac{\hat{g}_{i,j+\frac{1}{2}}(t)-\hat{g}_{i,j-\frac{1}{2}}(t)}{\Delta y}=0, \tag{4.8}$$

where $u_{i,j}(t)$ is the approximate solution to $u(x_i,y_j,t)$, $\hat{f}_{i+\frac{1}{2},j}(t)$ and $\hat{g}_{i,j+\frac{1}{2}}(t)$ are the $x-$ and $y-$numerical fluxes such that

$$\begin{aligned}\frac{\hat{f}_{i+\frac{1}{2},j}(t)-\hat{f}_{i-\frac{1}{2},j}(t)}{\Delta x}&=\frac{\partial f}{\partial x}(u(x_i,y_j,t))+O(\Delta x^m),\\ \frac{\hat{g}_{i,j+\frac{1}{2}}(t)-\hat{g}_{i,j-\frac{1}{2}}(t)}{\Delta y}&=\frac{\partial g}{\partial y}(u(x_i,y_j,t))+O(\Delta y^m),\end{aligned}$$

m is the desired order of accuracy in space.

The construction of L-DIRK3 in one dimension can be extended in a dimension-by-dimension manner with slight modifications on the time limiter. To be specific, denote

$\tau_x = \frac{\Delta t}{\Delta x}$ and $\tau_y = \frac{\Delta t}{\Delta y}$, the L-DIRK3 *finite difference* scheme reads

$$
\begin{aligned}
u_{i,j}^{(1)} &= u_{i,j}^n - \tau_x\alpha(\hat{f}_{i+\frac{1}{2},j}^{(1)} - \hat{f}_{i-\frac{1}{2},j}^{(1)}) - \tau_y\alpha(\hat{g}_{i,j+\frac{1}{2}}^{(1)} - \hat{g}_{i,j-\frac{1}{2}}^{(1)}) \\
u_{i,j}^{(2)} &= u_{i,j}^n - \tau_x(a_{21,i+\frac{1}{2},j}^x \hat{f}_{i+\frac{1}{2},j}^{(1)} - a_{21,i-\frac{1}{2},j}^x \hat{f}_{i-\frac{1}{2},j}^{(1)}) - \tau_y(a_{21,i,j+\frac{1}{2}}^y \hat{g}_{i,j+\frac{1}{2}}^{(1)} - a_{21,i,j-\frac{1}{2}}^y \hat{g}_{i,j-\frac{1}{2}}^{(1)}) \\
&\quad - \tau_x(a_{22,i+\frac{1}{2},j}^x \hat{f}_{i+\frac{1}{2},j}^{(2)} - a_{22,i-\frac{1}{2},j}^x \hat{f}_{i-\frac{1}{2},j}^{(2)}) - \tau_y(a_{22,i,j+\frac{1}{2}}^y \hat{g}_{i,j+\frac{1}{2}}^{(2)} - a_{22,i,j-\frac{1}{2}}^y \hat{g}_{i,j-\frac{1}{2}}^{(2)}), \\
u_{i,j}^{n+1} &= u_{i,j}^n - \tau_x(a_{31,i+\frac{1}{2},j}^x \hat{f}_{i+\frac{1}{2},j}^{(1)} - a_{31,i-\frac{1}{2},j}^x \hat{f}_{i-\frac{1}{2},j}^{(1)}) - \tau_y(a_{31,i,j+\frac{1}{2}}^y \hat{g}_{i,j+\frac{1}{2}}^{(1)} - a_{31,i,j-\frac{1}{2}}^y \hat{g}_{i,j-\frac{1}{2}}^{(1)}) \\
&\quad - \tau_x(a_{32,i+\frac{1}{2},j}^x \hat{f}_{i+\frac{1}{2},j}^{(2)} - a_{32,i-\frac{1}{2},j}^x \hat{f}_{i-\frac{1}{2},j}^{(2)}) - \tau_y(a_{32,i,j+\frac{1}{2}}^y \hat{g}_{i,j+\frac{1}{2}}^{(2)} - a_{32,i,j-\frac{1}{2}}^y \hat{g}_{i,j-\frac{1}{2}}^{(2)}) \\
&\quad - \tau_x(a_{33,i+\frac{1}{2},j}^x \hat{f}_{i+\frac{1}{2},j}^{n+1} - a_{33,i-\frac{1}{2},j}^x \hat{f}_{i-\frac{1}{2},j}^{n+1}) - \tau_y(a_{33,i,j+\frac{1}{2}}^y \hat{g}_{i,j+\frac{1}{2}}^{n+1} - a_{33,i,j-\frac{1}{2}}^y \hat{g}_{i,j-\frac{1}{2}}^{n+1}),
\end{aligned}
$$

where

$$
\begin{aligned}
a_{21,i+\frac{1}{2},j}^x &= \alpha + \theta_{i+\frac{1}{2},j}^{(1)}(\tau_2 - 2\alpha), & a_{21,i,j+\frac{1}{2}}^y &= \alpha + \theta_{i,j+\frac{1}{2}}^{(1)}(\tau_2 - 2\alpha), \\
a_{22,i+\frac{1}{2},j}^x &= \frac{1-\alpha}{2} + \theta_{i+\frac{1}{2},j}^{(1)}(\frac{3\alpha-1}{2}), & a_{22,i,j+\frac{1}{2}}^y &= \frac{1-\alpha}{2} + \theta_{i,j+\frac{1}{2}}^{(1)}(\frac{3\alpha-1}{2}), \\
a_{31,i+\frac{1}{2},j}^x &= \alpha + \theta_{i+\frac{1}{2},j}^{(2),x}(b_1 - \alpha), & a_{31,i,j+\frac{1}{2}}^y &= \alpha + \theta_{i,j+\frac{1}{2}}^{(2),y}(b_1 - \alpha), \\
a_{32,i+\frac{1}{2},j}^x &= \frac{1-\alpha}{2} + \theta_{i+\frac{1}{2},j}^{(2)}(b_2 - \frac{1-\alpha}{2}), & a_{32,i,j+\frac{1}{2}}^y &= \frac{1-\alpha}{2} + \theta_{i,j+\frac{1}{2}}^{(2)}(b_2 - \frac{1-\alpha}{2}), \\
a_{33,i+\frac{1}{2},j}^x &= \frac{1-\alpha}{2} + \theta_{i+\frac{1}{2},j}^{(2)}(\frac{3\alpha-1}{2}), & a_{33,i,j+\frac{1}{2}}^y &= \frac{1-\alpha}{2} + \theta_{i,j+\frac{1}{2}}^{(2)}(\frac{3\alpha-1}{2}),
\end{aligned}
$$

$$
\theta_{i+\frac{1}{2}.j}^{(k)} = \frac{\theta_{i,j}^{(k)} + \theta_{i+1,j}^{(k)}}{2}, \quad \theta_{i.j+\frac{1}{2}}^{(k)} = \frac{\theta_{i,j}^{(k)} + \theta_{i,j+1}^{(k)}}{2}.
$$

The time limiter $\theta_{i,j}^{(k)}$ is defined as

$$
\theta_{i,j}^{(k)} = minmod(r_{i,j}^{(k),1}, r_{i,j}^{(k),2}, r_{i,j}^{(k),3}, r_{i,j}^{(k),4}, 1),
$$

where $r_{i,j}^{(k),1}$, $r_{i,j}^{(k),2}$, $r_{i,j}^{(k),3}$, $r_{i,j}^{(k),4}$ are the monotonicity indicators at point (x_i, y_j) from

horizontal, vertical and diagonal directions

$$r_{i,j}^{(k),1} = \frac{u_{i+1,j}^{(k+1)} - u_{i-1,j}^{(k+1)}}{u_{i+1,j}^{(k)} - u_{i-1,j}^{(k)}}, \quad r_{i,j}^{(k),2} = \frac{u_{i,j+1}^{(k+1)} - u_{i,j-1}^{(k+1)}}{u_{i,j+1}^{(k)} - u_{i,j-1}^{(k)}},$$
$$r_{i,j}^{(k),3} = \frac{u_{i+1,j+1}^{(k+1)} - u_{i-1,j-1}^{(k+1)}}{u_{i+1,j+1}^{(k)} - u_{i-1,j-1}^{(k)}}, \quad r_{i,j}^{(k),4} = \frac{u_{i-1,j+1}^{(k+1)} - u_{i+1,j-1}^{(k+1)}}{u_{i-1,j+1}^{(k)} - u_{i+1,j-1}^{(k)}}.$$

The above two-dimensional constructions include the information from diagonal directions, this allows for the detection of potential discontinuities that are not aligned with the axis. The extension to higher dimensions can be obtained with similar idea.

The L-DIRK3 scheme can also be applied in the framework of finite volume methods, but the computation will be rather complicated due to the complexity of the quadratures for surface integrals, $\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(u(x_{i+\frac{1}{2}}, y))dy$ and $\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} g(u(x, y_{j+\frac{1}{2}}))dx$, when the order of accuracy in space is higher than two.

Chapter 5: Numerical Results

We present numerical results for several test cases. For all the computations of DIRK3 and L-DIRK3, we apply the 5th order WENO-JS finite difference scheme [3] for space reconstruction. The number of equally spaced mesh points in the domain is represented by N. The implicit systems are solved with Newton-type sub-iterations [28]. At each implicit stage, the iterations stop if the l^2-norm of the residual is reduced by four orders of magnitudes, or the number of iterations reaches the maximum of 30. The limiters $\theta_j^{(k)}$ are updated explicitly in the iterations.

5.1 Linear Advection Equation

The first test case is the linear advection equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$$

with periodic boundary conditions and a smooth initial condition $u_0(x) = \sin^4(\frac{x}{2})$ over the computational domain $[0, 2\pi]$. This test case is chosen to demonstrate the order of accuracy in time and the resolution at smooth extrema. Tables 5.1–5.2 present the $L^\infty-$, L^1- and L^2-errors of different methods under $N = 25, 50, 100, 200, 400$ at $CFL = 0.5$

after one period of revolution. The results indicate the second-order convergence rates of L-DIRK3. The uniform third order accuracy in time is not recovered since the scheme reduces to first order at the extrema. Figure 5.1 shows the solutions by L-DIRK3 and DIRK3 under $N = 80$. We observe that L-DIRK3, although suffers from order reduction, resolves the extrema as sharply as the unlimited DIRK3 under a moderate mesh size.

N	L^∞ err	L^1 err	L^2err	L^∞ rate	L^1 rate	L^2 rate
50	$1.58e^{-2}$	$2.19e^{-3}$	$3.67e^{-3}$	–	–	–
100	$4.72e^{-3}$	$6.91e^{-4}$	$1.05e^{-3}$	1.74	1.66	1.81
200	$1.24e^{-3}$	$1.87e^{-4}$	$2.78e^{-4}$	1.93	1.88	1.91
400	$3.85e^{-4}$	$4.60e^{-5}$	$6.89e^{-5}$	1.68	2.03	2.01
800	$1.09e^{-4}$	$1.02e^{-5}$	$1.66e^{-5}$	1.82	2.17	2.06

Table 5.1: Error norms for linear advection, L-DIRK3, CFL = 0.5, 5th order WENO in space, periodic bc, 1 period of revolution.

N	L^∞ err	L^1 err	L^2err	L^∞ rate	L^1 rate	L^2 rate
50	$6.07e^{-4}$	$1.96e^{-4}$	$2.37e^{-4}$	–	–	–
100	$1.71e^{-5}$	$9.01e^{-6}$	$1.01e^{-5}$	5.15	4.44	3.55
200	$1.71e^{-6}$	$8.16e^{-7}$	$9.53e^{-7}$	3.32	3.46	3.40
400	$2.00e^{-7}$	$1.02e^{-7}$	$1.16e^{-7}$	3.09	3.00	3.04
800	$2.47e^{-8}$	$1.27e^{-8}$	$1.44e^{-8}$	3.02	3.00	3.01

Table 5.2: Error norms for linear advection, DIRK3, CFL = 0.5, 5th order WENO in space, periodic bc, 1 period of revolution.

The second test case comprises of multiple waves over the domain $[0, 2\pi]$, including a $\sin^4(\pi x)$ distribution, a step function and a hat function. The mesh size $N = 400$ is applied. Figure 5.2 shows the solutions by DIRK3 and L-DIRK3 at $CFL = 2$ after one period of revolution. We also include the second order solution for comparison,

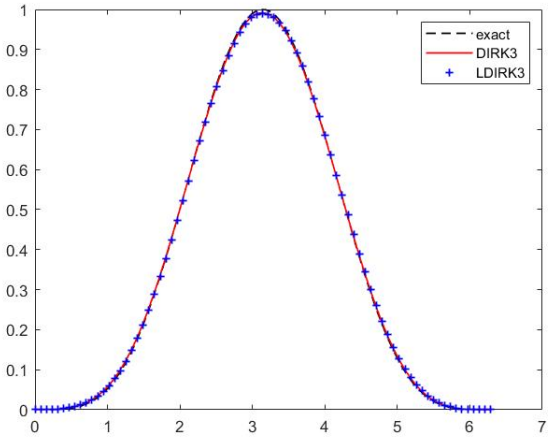


Figure 5.1: Linear advection, $\sin^4(\frac{x}{2})$ wave, DIRK3 vs L-DIRK3, $CFL = 0.5$, 5th order WENO in space, $N = 80$, periodic bc, 1 period of revolution.

which is computed with piecewise linear MUSCL extrapolation [13] (with minmod slope limiter) in space and explicit SSPRK3 [4] time integration at $CFL = 0.9$. It is seen that the solution by DIRK3 presents obvious overshoots and undershoots near the edges of square wave, whereas the L-DIRK3 generates non-oscillatory solution. In spite of slight clipping at the tops of sine wave and triangular wave, the solution by L-DIRK3 resolves the extremum obviously better than the second order solution.

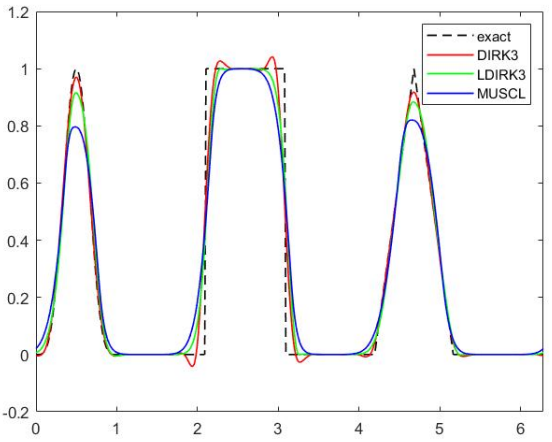


Figure 5.2: Linear advection, multiple waves, DIRK3 vs L-DIRK3, $CFL = 2$, 5th order WENO in space, $N = 400$, periodic bc, 1 period of revolution.

5.2 1D Burger's Equation

We present the results for the one-dimensional Burger's equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) = 0$$

with periodic boundary conditions over the domain $[0, 2\pi]$. The initial data comprises of an expansion wave and a compression wave. Figure 5.3 shows the solutions at $t = 2$ with mesh size $N = 100$ at $CFL = 3$. The solution comprises of an expansion wave and a shock wave. It is seen that the solution by DIRK3 has obvious overshoot near the shock under relatively large time step. The solution by L-DIRK3 is much less oscillatory and resolves the expansion wave and the shock well.

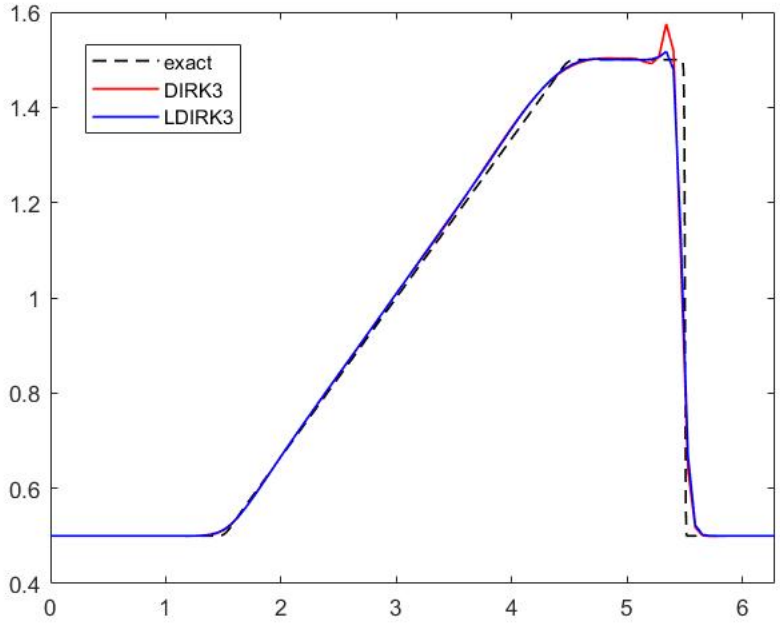


Figure 5.3: 1D Burger's equation, expansion and comrpession wave, DIRK3 vs L-DIRK3, $CFL = 3$, 5th order WENO in space, $N = 100$, $t = 2$.

5.3 1D Euler Equations

The one-dimensional Euler equations of gas dynamics is given by

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0,$$

where $\mathbf{U}$, the vector of conserved variables, and $\mathbf{F}$, the flux vector, are given by

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ e \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ (e+p)u \end{bmatrix},$$

ρ, u, p are density, velocity and pressure, e is the total energy per unit volume which is given by

$$e = \frac{p}{\gamma - 1} + \frac{\rho u^2}{2}, \quad \gamma = 1.4.$$

We test the performance of L-DIRK3 in the Sod's problem, Lax's problem and Shu-Osher problems. These test cases are Riemann problems in a constant area tube. The left and the right states are indexed by L and R.

The interface numerical fluxes are constructed with the characteristic-wise Lax Friedrich flux vector splitting

$$\begin{aligned} \mathbf{F}^{\pm}(\mathbf{U}) &= \frac{1}{2}(\mathbf{F}(\mathbf{U}) \pm \mathbf{R}\Lambda\mathbf{R}^{-1}\mathbf{U}), \quad \Lambda = diag(\max_{\mathbf{U}}\{|\lambda_i(\mathbf{U})|\}_i), \\ \hat{\mathbf{F}}_{j+\frac{1}{2}} &= \hat{\mathbf{F}}^{+}_{j+\frac{1}{2}} + \hat{\mathbf{F}}^{-}_{j+\frac{1}{2}}. \end{aligned} \tag{5.1}$$

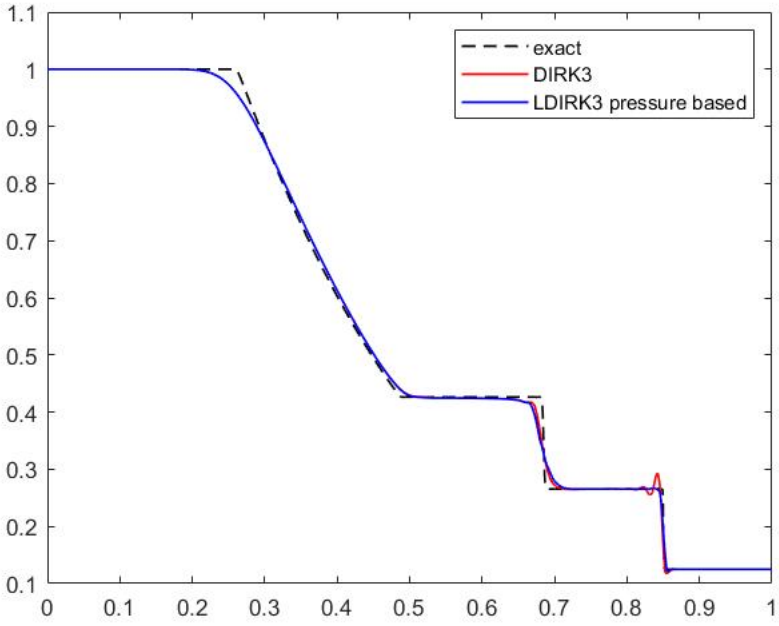
Here $\{\lambda_i\}$ are the eigenvalues of flux Jacobian $\dfrac{\partial \mathbf{F}}{\partial \mathbf{U}}$, $\mathbf{R}$ is the matrix of right eigenvectors, $\hat{\mathbf{F}}_{j+\frac{1}{2}}$ is the numerical flux at cell interface $x_{j+\frac{1}{2}}$, $\hat{\mathbf{F}}^{+}_{j+\frac{1}{2}}$, $\hat{\mathbf{F}}^{-}_{j+\frac{1}{2}}$ are the reconstructions of $\mathbf{F}^{+}$, $\mathbf{F}^{-}$ at $x_{j+\frac{1}{2}}$ from the left and the right.

Sod's problem

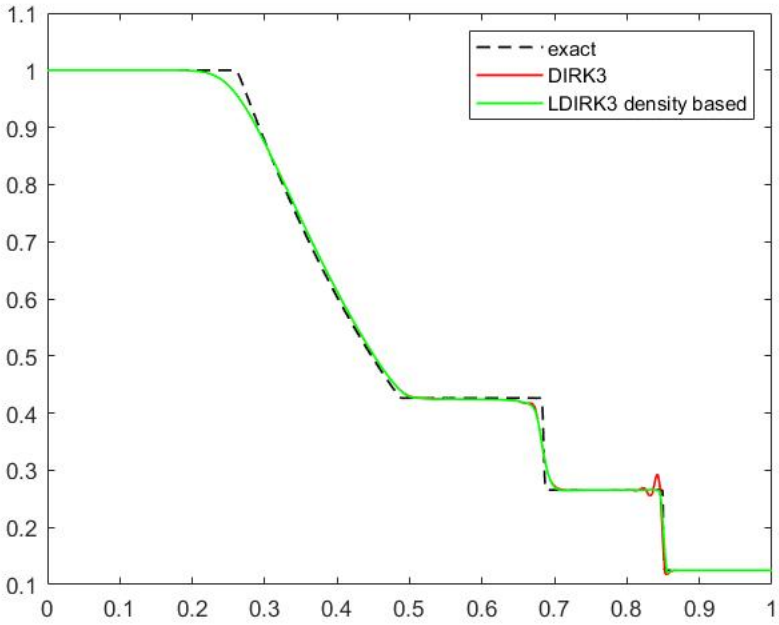
The computational domain of Sod's problem is [0, 1] and the interface is at $x = 0.5$. The initial data is given by

$$\begin{cases} p_L = 1.0, \quad p_R = 0.1 \\ \rho_L = 1.0, \quad \rho_R = 0.125 \\ u_L = 0.0, \quad u_R = 0.0 \end{cases} .$$

The numerical solution is computed with $N = 400$ under $CFL = 4$. We present the evolution of density at $t = 0.2$ in Figure 5.4. We see that under high CFL number the solution of DIRK3 has obvious oscillation at the shock. By comparison, the solutions by L-DIRK3, with density-based and pressure-based limiting, have no visible oscillation at the shock, while the resolution at the expansion wave and discontinuities remain comparable with the unlimited DIRK3.



(a) Pressure-Based



(b) Density-Based

Figure 5.4: Sod's problem (Density), $CFL = 4$, 5th order WENO in space, $N = 400$, $t = 0.2$.

Lax's problem

The initial data of Lax's problem is given by

$$\begin{cases} p_L = 3.528, \quad p_R = 0.571 \\ \rho_L = 0.445, \quad \rho_R = 0.5 \\ u_L = 0.698, \quad u_R = 0.0 \end{cases} .$$

over the domain $[0, 1]$ with interface at $x = 0.5$. We compute the solutions at $t = 0.14$ with mesh size $N = 300$ at $CFL = 3.5$. Figure 5.5 shows the density evolution. It is seen that the unlimited DIRK3 generates obvious oscillation at the edge of the square wave. The solutions by L-DIRK3, both pressure-based and density- based, present no visible oscillation and resolve the expansion wave and the shocks well.

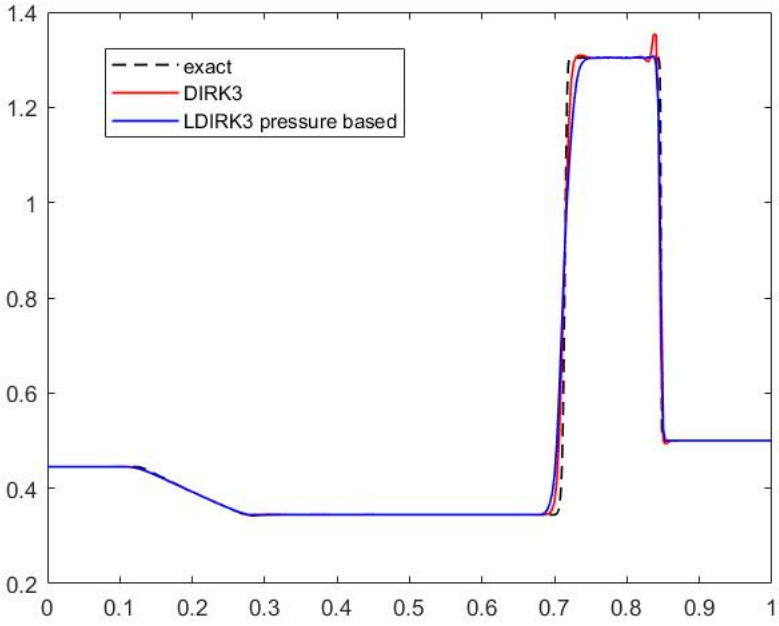
Shu-Osher problem

In Shu-Osher problem, the solution comprises of a discontinuity as well as a smooth harmonic density wave. We will examine the resolution of L-DIRK3 at extremum and its stability at the shock. The computational domain is [-5, 5] and the interface is at $x = -4$. The initial data is given by

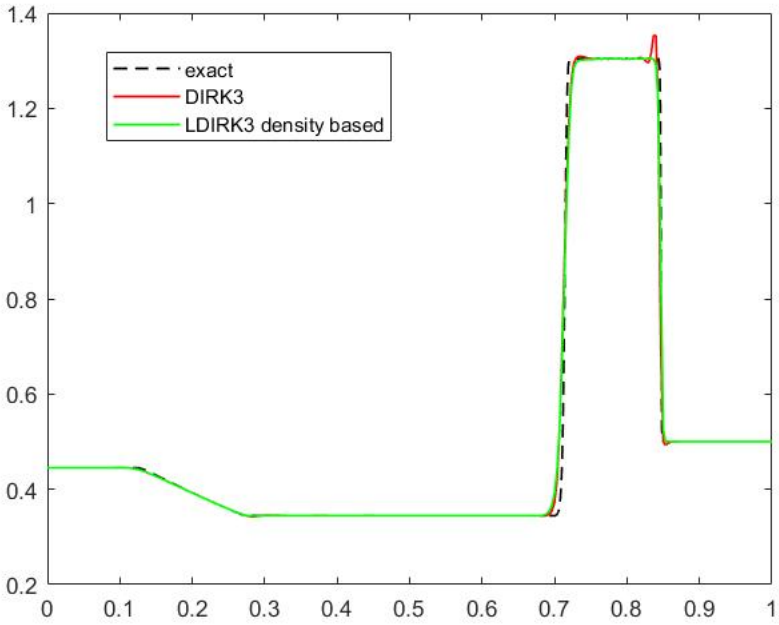
$$\begin{cases} p_L = 10.33333, \quad p_R = 1.0 \\ \rho_L = 3.857143, \quad \rho_R = 1 + 0.2\sin(5x) \\ u_L = 2.6293690, \quad u_R = 0.0 \end{cases} .$$

We compute the solution of density at $t = 1.8$. The mesh size $N = 400$ is applied. At first, we test with $CFL = 2$. The result is shown in Figure 5.6. We see that under relatively large CFL number, the solutions by L-DIRK3 have a little clipping at extremum but not too severe. The resolutions of L-DIRK3 and unlimited DIRK3 are overall comparable.

We also present the solutions under $CFL = 0.9$ in Figure 5.7. It is seen that when the smaller time step is applied, the solutions given by DIRK3 and L-DIRK3 become almost indistinguishable. For comparison, we include the second order solution under the same mesh size and the same CFL condition in the plot. For the second order solution, we apply the piecewise linear MUSCL extrapolation (with minmod slope limiter) for space discretization, the interface numerical fluxes are computed with the flux vector splitting (5.1), and we apply the explicit SSPRK3 for time integration. We see that the solutions by L-DIRK3, both density- and pressure-based, capture the extremum much more sharply than the typical second order solution. The high resolution of 5th order WENO scheme is well preserved despite of the time limiting.

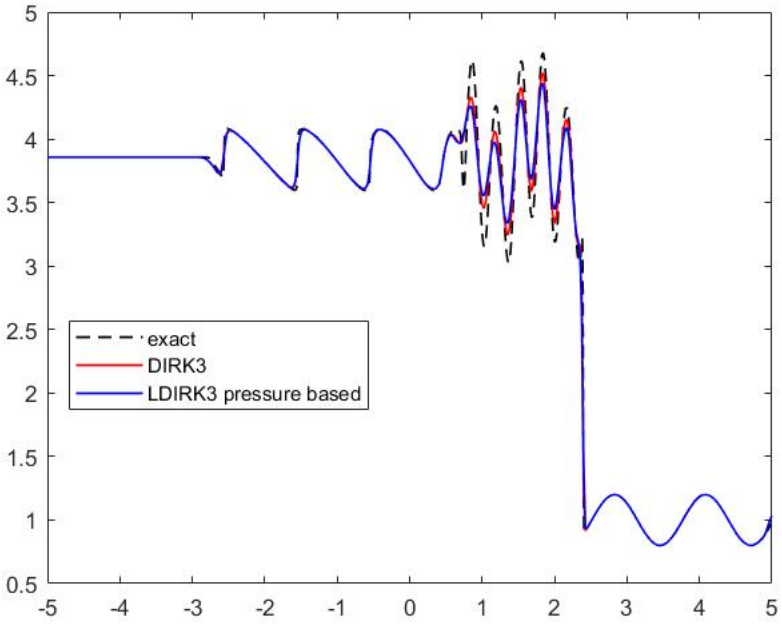


(a) Pressure-Based

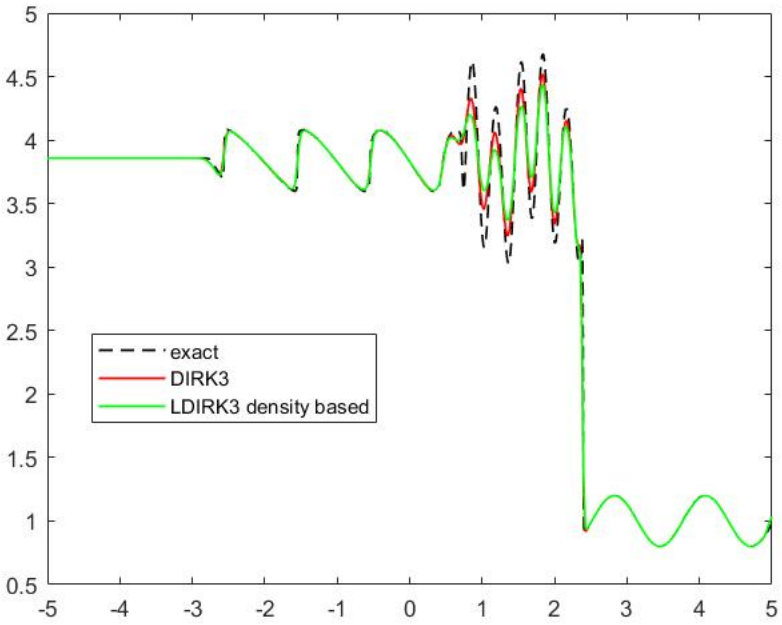


(b) Density-Based

Figure 5.5: Lax's problem (Density), $CFL = 3.5$, 5th order WENO in space, $N = 300$, $t = 0.14$.

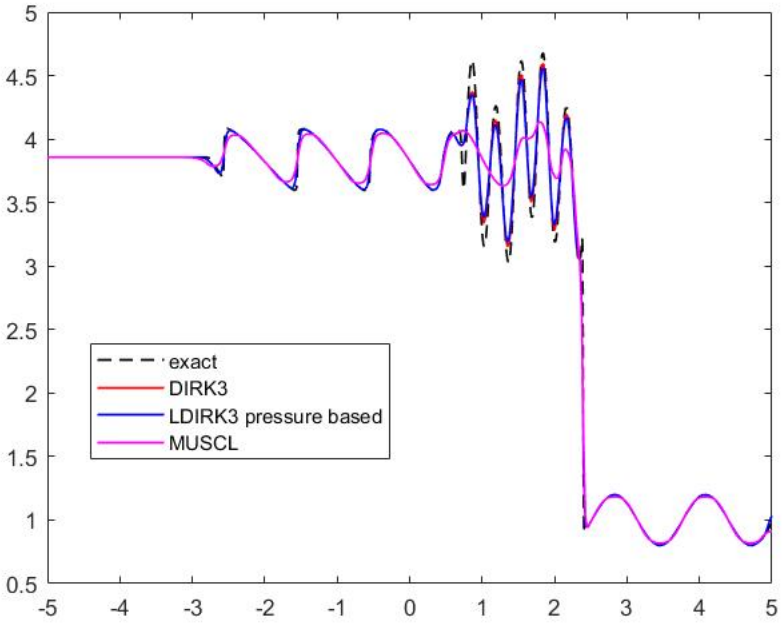


(a) Pressure-Based

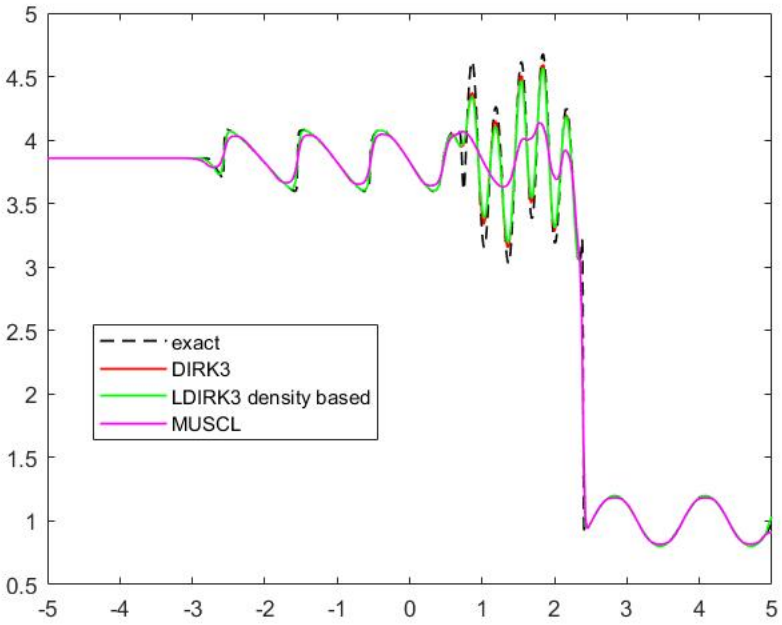


(b) Density-Based

Figure 5.6: Shu-Osher (Density), $CFL = 2$, 5th order WENO in space, $N = 400$, $t = 1.8$.



(a) Pressure-Based



(b) Density-Based

Figure 5.7: Shu-Osher (Density), $CFL = 0.9$, 5th order WENO in space, $N = 400$, $t = 1.8$.

5.4 1D Convection-Diffusion Equations

In actual applications, a scenario where implicit methods have clear advantage over explicit methods is solving convection-diffusion problems, where the time step step scales at $\Delta t \sim \Delta x^2$, in contrast to $\Delta t \sim \Delta x$ for convective problems. The diffusion term is usually discretized implicitly to get rid of the severe time step restriction. We first consider the viscous Burger's equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) = \epsilon \frac{\partial}{\partial x}(\nu(u)\frac{\partial u}{\partial x}),$$

which admits an equivalent conservative form

$$\frac{\partial u}{\partial t} + \frac{\partial \tilde{f}}{\partial x} = 0, \quad \tilde{f} = \frac{u^2}{2} - \epsilon\nu(u)\frac{\partial u}{\partial x}.$$

The associated diffusion coefficient is given by a discontinuous function

$$\nu(u) = \begin{cases} 1, & |u| \geq 0.5, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

The equation has hyperbolic nature when $u \in [-0.5, 0.5]$ and parabolic elsewhere.

In the following computations, the initial data are set to 2 and -2 on the intervals $[-0.9, -0.1]$ and $[0.1, 0.9]$ respectively over the domain $[-1.5, 1.5]$. the time step is

chosen according to the CFL condition

$$\Delta t(\frac{a}{\Delta x} + \max_u \frac{\epsilon\nu(u)}{\Delta x^2}) = CFL, \quad a = \max_u |f'(u)|.$$

The convection term is discretized with fifth order WENO method and the diffusion term is discretized with the usual second order central differencing. We take $\epsilon - 0.1$ and compute the solution at $t = 0.2$ under mesh size $N = 500$ a large CFL number of 10. Figure 5.4 compares the results by DIRK3 and L-DIRK3. The reference solution is computed with explicit SSPRK3 at $CFL = 0.6$ under the same space mesh size. Strong spurious oscillations are observed in the solution by DIRK3, whereas the L-DIRK3 scheme generates non-oscillatory solution with a much larger time step than the explicit method.

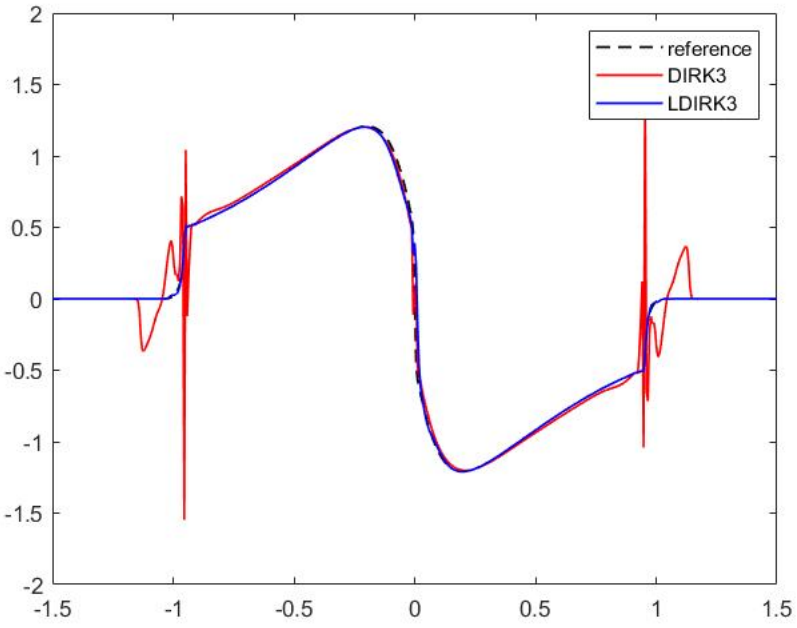


Figure 5.8: 1D viscous Burger's equation. DIRK3 vs L-DIRK3, $CFL = 10$, 5th order WENO for convection, 2nd order central differencing for diffusion, $N = 500$, $t = 0.5$.

Next, we consider the viscous Buckley-Leverett equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = \epsilon \frac{\partial}{\partial x}(\nu(u)\frac{\partial u}{\partial x}), \quad f(u) = \frac{u^2}{u^2 + (1-u)^2}.$$

This equation is a prototype model for oil reservoir (two-phase flow) simulation. The initial data are set to 0.9 for $|x + \frac{1}{\sqrt{2}}| < 0.4$ and -0.9 for $|x - \frac{1}{\sqrt{2}}| < 0.4$. The diffusion coefficient $\nu(u)$ is taken to 1 for $|u| \geq 0.2$ and zero otherwise. We compute the solution with respect to $\epsilon = 0.1$ under $CFL = 10$. Again, strong oscillations are observed in the solution by the unlimited DIRK3 scheme, whereas the result by L-DIRK3 presents a satisfying stability as well as a high resolution.

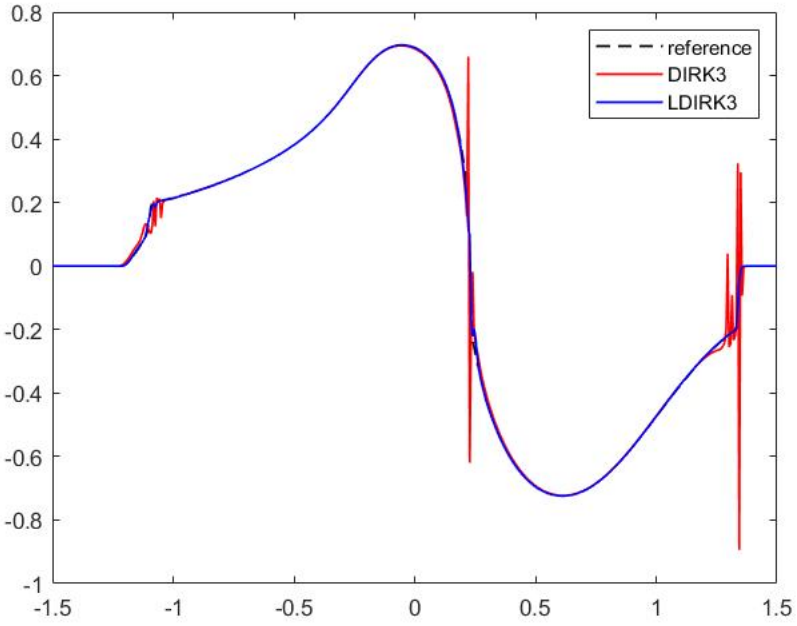


Figure 5.9: 1D viscous Buckley-Leverett equation. DIRK3 vs L-DIRK3, $CFL = 10$, 5th order WENO for convection, 2nd order central differencing for diffusion, $N = 500$, $t = 0.5$.

5.5 Two-Dimensional Problems

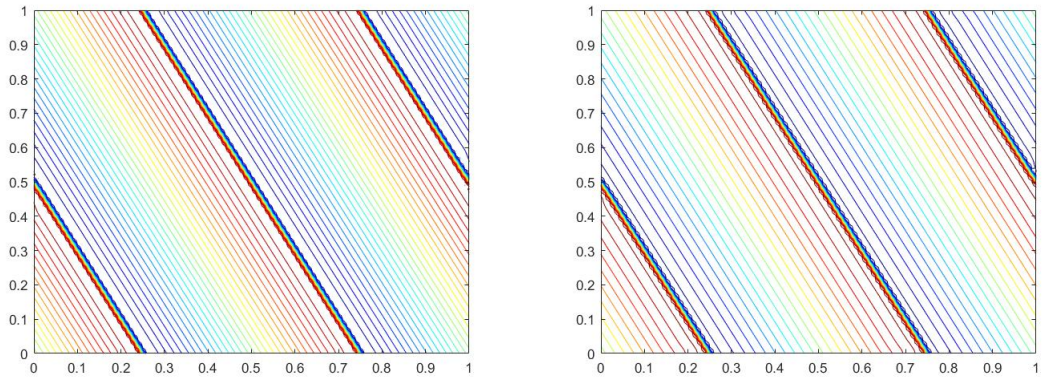
In section 4, we have discussed the construction of L-DIRK3 for 1D problems. For high-dimensional problems, the scheme can be extended dimension-by-dimension. We test the L-DIRK3 scheme against two-dimensional scalar and system of equations and examine the validity of high-dimensional extension.

2D Burger's equation

We solve the two-dimensional inviscid Burger's equation

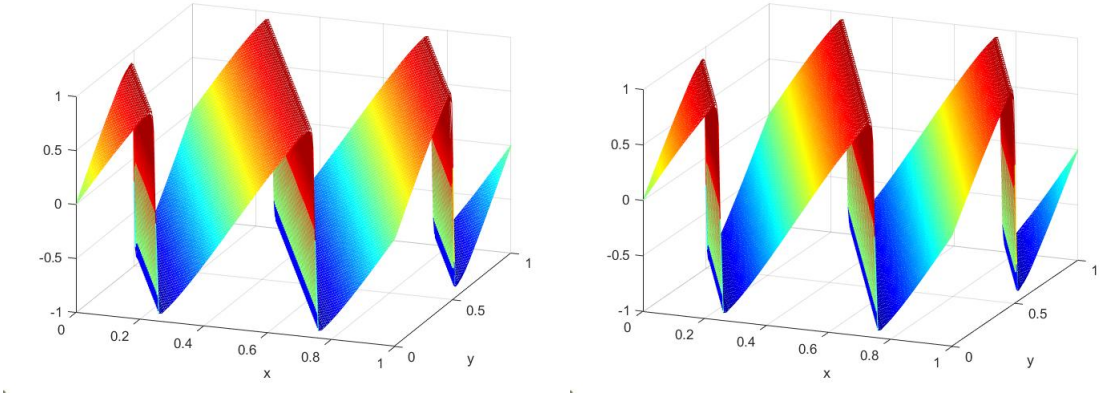
$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) + \frac{\partial}{\partial y}(\frac{u^2}{2}) = 0$$

with period boundary conditions, subject to initial data $u_0(x, y) = \sin(4\pi(x + \frac{y}{2}))$ over the domain $[0, 1] \times [0, 1]$. Due to the nonlinearity, the smooth initial condition steepens and forms stationary shocks, which eventually dissipate. We present the shock formation at $t = 0.1$ under mesh size $N_x = N_y = 100$ at $CFL = 2$. Figures 5.10–5.11 display the solution contours and surface plots. Figure 5.12 shows the solutions along the diagonal $x = y$. It is seen that the solutions by L-DIRK3 and DIRK3 are indistinguishable . Applying time-limiters does not smear the discontinuities in high dimension.



(a) DIRK3 (b) L-DIRK3

Figure 5.10: 2D Burger's equation surface plots, $CFL = 2$, 5th order WENO in space, $N_x = N_y = 100$, $t = 0.1$.



(a) DIRK3 (b) L-DIRK3

Figure 5.11: 2D Burger's equation, surface plots, $CFL = 2$, 5th order WENO in space, $N_x = N_y = 100$, $t = 0.1$.

2D Buckley-Leverett equation

Consider the two-dimensional inviscid Buckley-Leverett equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) + \frac{\partial}{\partial y} g(u) = 0,$$

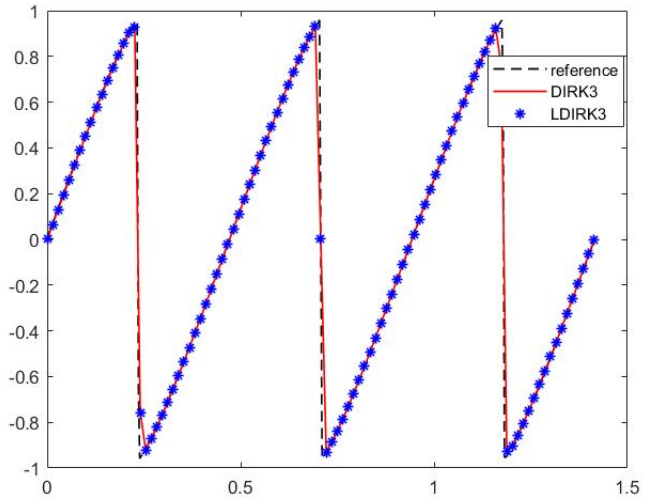


Figure 5.12: 2D Burger's equation, solutions along $x = y$, $CFL = 2$, 5th order WENO in space, $N_x = N_y = 100$, $t = 0.1$.

with the flux functions

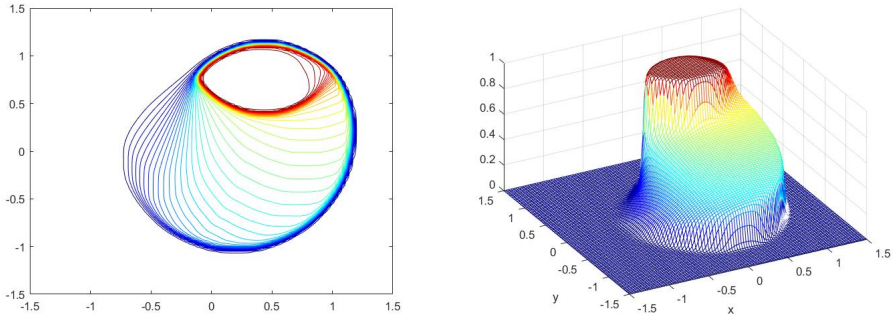
$$f(u) = \frac{u^2}{(1-u)^2 + u^2},$$

$$g(u) = f(u)(1 - 5(1-u)^2),$$

and the initial data

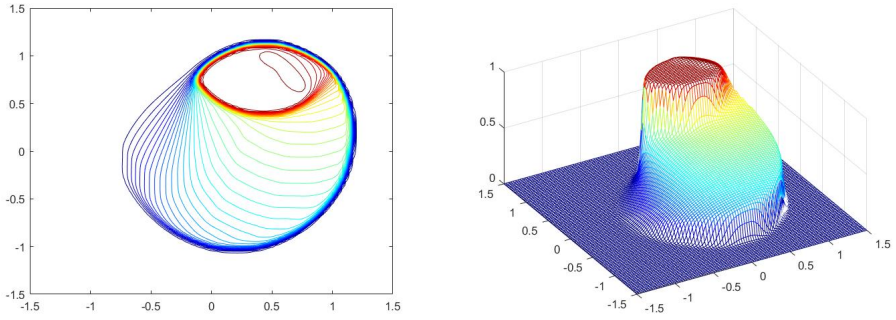
$$u(x, y, 0) = \begin{cases} 0.9 & x^2 + y^2 < 0.5 \\ 0 & \text{else} \end{cases}.$$

The reference solution, which is computed with SSPRK3-WENO5 scheme at $CFL = 0.9$, is shown in Figure 5.13. Figures 5.14–5.15 show the solutions by DIRK3 and L-DIRK3 at $t = 0.4$, under mesh size $N_x = N_y = 100$ at $CFL = 3$. Figure 5.16 compares the solutions by different methods along the diagonal $x = y$. It is seen that L-DIRK3, with the help of time limiting, has a smaller overshoot than DIRK3 near the shock.



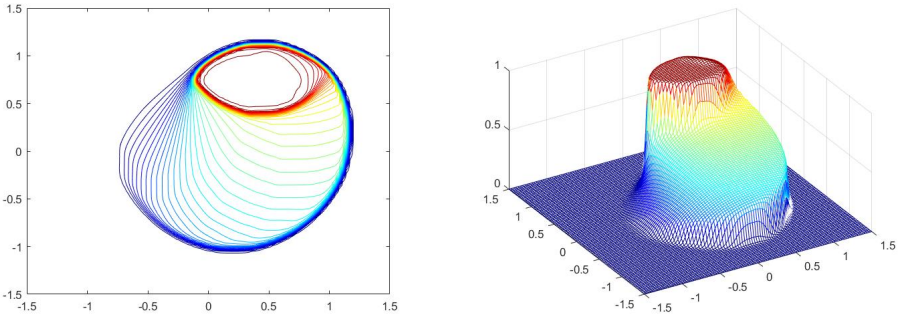
(a) Solution contours (b) Surface plot

Figure 5.13: 2D Buckley-Leverett equation, reference solution, $N_x = N_y = 100$, $t = 0.4$.



(a) Solution contours (b) Surface plot

Figure 5.14: 2D Buckley-Leverett equation, DIRK3, $CFL = 3$, 5th order WENO in space, $N_x = N_y = 100$, $t = 0.4$.



(a) Solution contours (b) Surface plot

Figure 5.15: 2D Buckley-Leverett equation, L-DIRK3, $CFL = 3$, 5th order WENO in space, $N_x = N_y = 100$, $t = 0.4$.

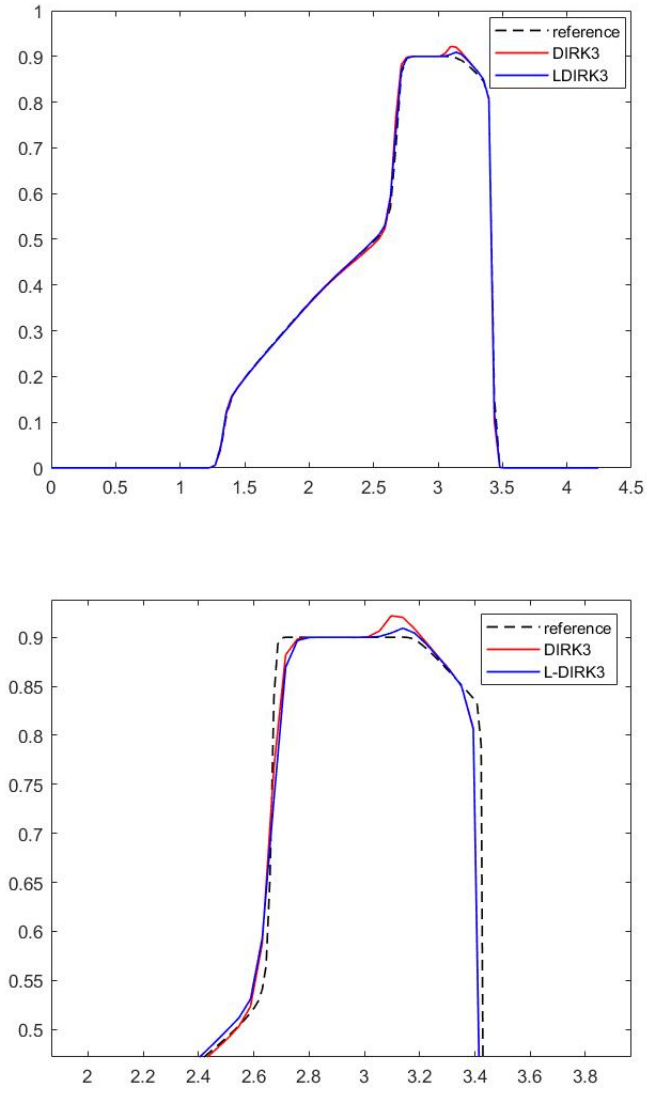


Figure 5.16: 2D Buckley-Leverett equation, solutions along $x = y$, $N_x = N_y = 100$, $t = 0.4$.

2D Euler equations

We apply the L-DIRK3 scheme to solve the two-dimensional Euler's equations for an ideal gas

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial y} = 0,$$

the conservative variables $\mathbf{U}$, the $x-$flux $\mathbf{F}$ and the $y-$flux $\mathbf{G}$ are given by

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ e \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (e+p)u \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (e+p)v \end{bmatrix},$$

where ρ, p are density and pressure, u, v are velocities in $x-$ and $y-$direction, the total energy e is given by

$$e = \frac{p}{\gamma - 1} + \frac{\rho(u^2 + v^2)}{2}$$

with $\gamma = 1.4$.

We first consider a 2D Riemann problem [29, 30, 31]. The computation domain is taken as $[0, 1] \times [0, 1]$, the initial data takes the form

$$\begin{cases} p = p_1, \rho = \rho_1, u = u_1, v = v_1, & \text{if } x > 0.5 \text{ and } y > 0.5 \\ p = p_2, \rho = \rho_2, u = u_2, v = v_2, & \text{if } x < 0.5 \text{ and } y > 0.5 \\ p = p_3, \rho = \rho_3, u = u_3, v = v_3, & \text{if } x < 0.5 \text{ and } y < 0.5 \\ p = p_4, \rho = \rho_4, u = u_4, v = v_4, & \text{if } x > 0.5 \text{ and } y < 0.5 \end{cases}.$$

We consider the following configuration

$p_2 = 0.3$	$\rho_2 = 0.5323$	$p_1 = 1.5$	$\rho_1 = 1.5$
$u_2 = 1.206$	$v_2 = 0$	$u_1 = 0$	$v_1 = 0$
$p_3 = 0.029$	$\rho_3 = 0.5323$	$p_4 = 0.3$	$\rho_4 = 0.5323$
$u_3 = 1.206$	$v_3 = 0$	$u_4 = 0$	$v_4 = 1.206$

.

In this test case, the shock waves are formed at the contact surfaces between each two subregions. At the intersection of the shocks, an interaction region with complex structures will be developed. We compute the density at $t = 0.3$ under mesh size 150×150. The reference solution given by the SSPRK3-WENO5 scheme is shown in Figures 5.17. Figures 5.18–5.19 show the solutions of the DIRK3 and the L-DIRK3 (density-based limiter) schemes under $CFL = 3$. It is seen that the L-DIRK3 scheme resolves the structures in the shock interaction region sharply under large time step in spite of the time limiting. Although the uniform convergence rate of L-DIRK3 does not achieve third order, its advantage, when combined with high order WENO method, over the typical second order scheme (Figure 5.20) is clear in terms of characterizing complex flows in multi-dimensions.

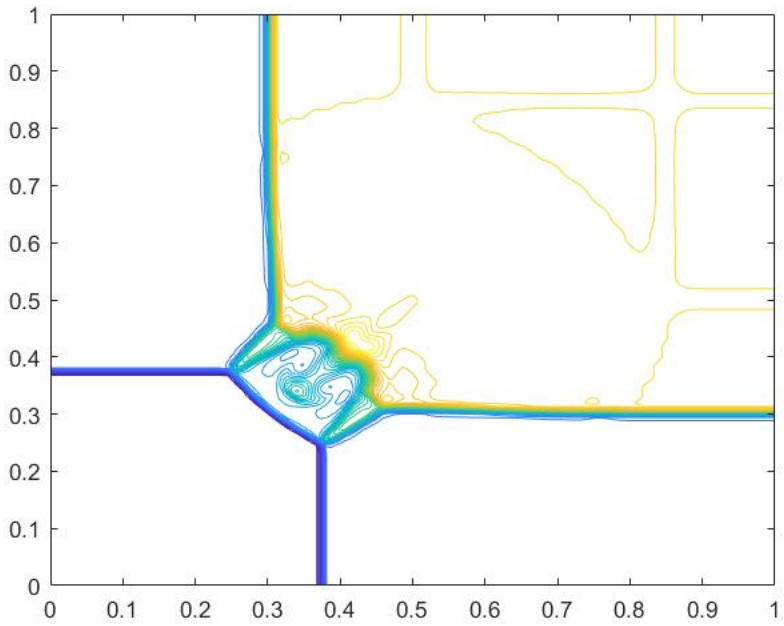


Figure 5.17: 2D Riemann problem for Euler equations, SSPRK3 - WENO5, $N_x = N_y = 150$, $CFL = 0.6$, $t = 0.3$.

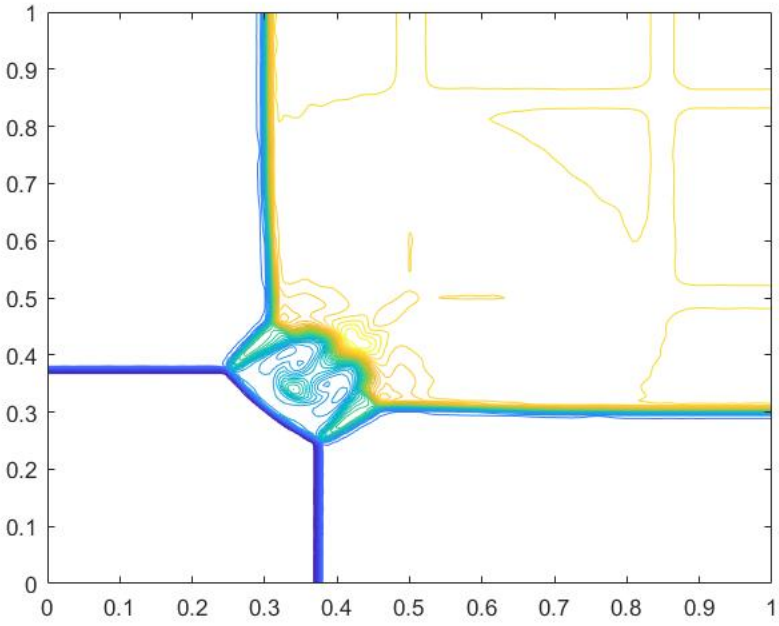


Figure 5.18: 2D Riemann problem for Euler equations, L-DIRK3 - WENO5, $N_x = N_y = 150$, $CFL = 3$, $t = 0.3$.

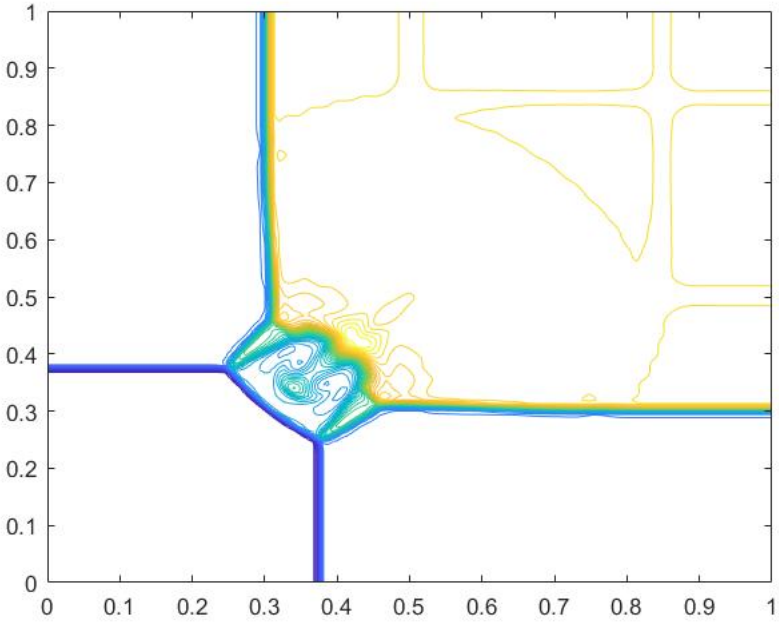


Figure 5.19: 2D Riemann problem for Euler equations, DIRK3 - WENO5, $N_x = N_y = 150$, $CFL = 3$, $t = 0.3$.

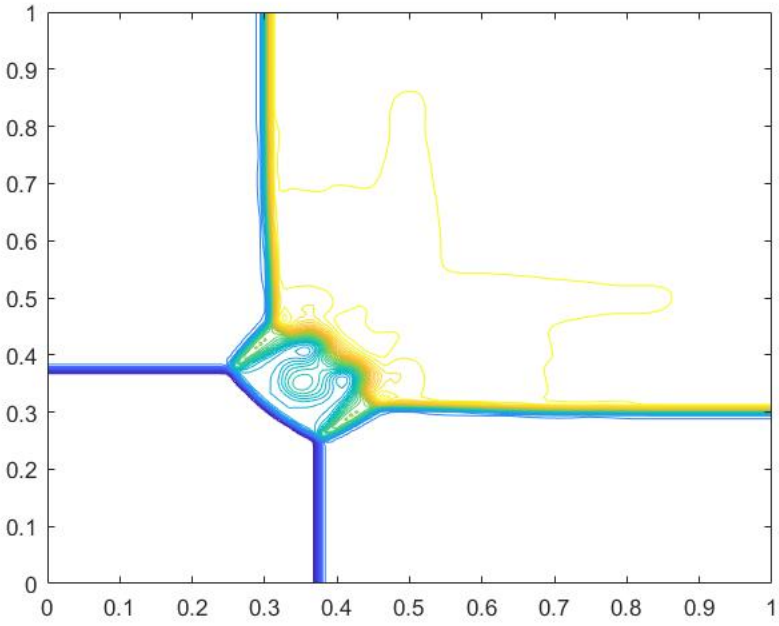


Figure 5.20: 2D Riemann problem for Euler equations, SSPRK3 - MUSCL, $N_x = N_y = 150$, $CFL = 0.6$, $t = 0.3$.

In the second test case, we consider the isentropic vortex advection problem. The mean flow is $\rho = 1$. $p = 1$, $u = 1$, $v = 1$. The temperature is given by $T = \dfrac{p}{\rho}$, and the entropy is defined as $S = \frac{p}{\rho^\gamma}$. We add the perturbations

$$(\delta u, \delta v) = \frac{\epsilon}{2\pi} e^{0.5(1-r^2)}(-\bar{y}, \bar{x}),$$
$$\delta T = -\frac{(\gamma - 1)\epsilon^2}{8\gamma\pi^2} e^{1-r^2}, \quad \delta S = 0,$$

where $(\bar{x}, \bar{y}) = (x - 5, y - 5)$, $r = \bar{x}^2 + \bar{y}^2$. The vortex strength is taken as $\epsilon = 5$. The exact solution of this problem is simply given by

$$u(x, y, t) = u(x - t, y - t, 0).$$

The computational domain is set to be $[0, 10] \times [0, 10]$ coupled with periodic boundary conditions in both directions. The vortex is advected along the diagonal direction. Figures 5.21–5.22 compares the solutions of density by the DIRK3 and the L-DIRK3 schemes along the diagonal $x = y$ after one period of revolution under mesh size 150×150 and $CFL = 4$. It is seen that the solution by DIRK3 presents obvious oscillation near the edge of the vortex whereas the result by L-DIRK3 is much less oscillatory. At the center of vortex, the L-DIRK3 is more dissipative since the order of accuracy in time locally drops to first order. Nevertheless, the clipping seems to be acceptable considering the large CFL number applied. Indeed, Table 5.3 compares the errors of density of the limited and unlimited schemes. It is seen that applying the time limiter increases the L^∞ error due to the stronger dissipation at smooth extrema. However, the L^1 and L^2 errors are reduced because of the removal of oscillations.

	L^∞ err	L^1 err	L^2 err
DIRK3	$2.39e-2$	$5.36e-3$	$6.72e-3$
L-DIRK3	$3.37e-2$	$2.72e-3$	$3.95e-3$

Table 5.3: Vortex advection, density error norms. DIRK3 vs L-DIRK3. $CFL = 4$. 5th order WENO in space, $N_x = N_y = 150$, periodic bc, 1 period of evolution.

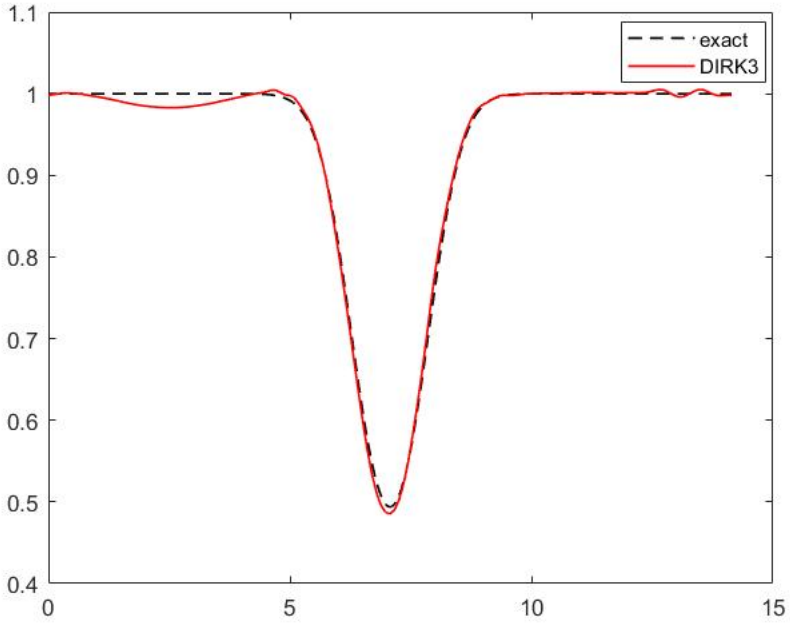


Figure 5.21: Vortex advection problem for 2D Euler equations, density along $x = y$ by DIRK3, $N_x = N_y = 150$, $CFL = 4$, $t = 10$.

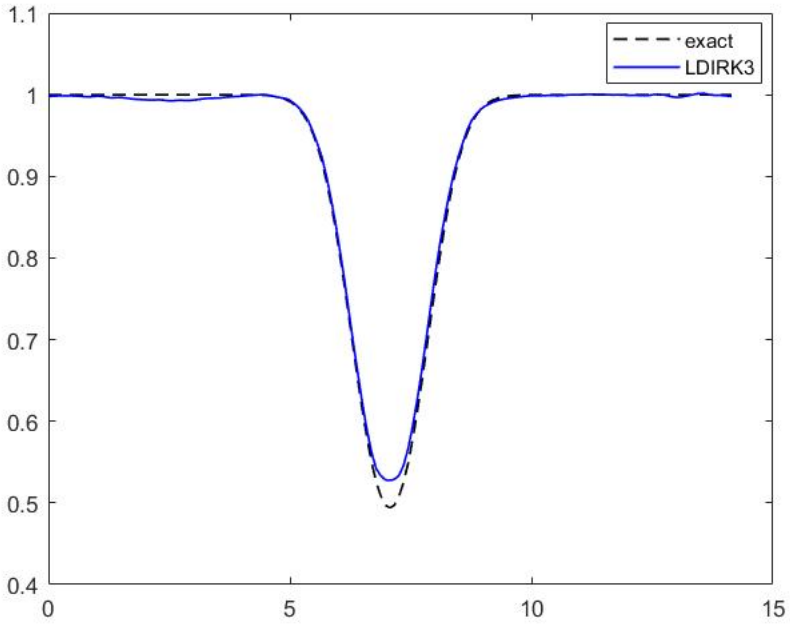


Figure 5.22: Vortex advection problem for 2D Euler equations, density along $x = y$ by L-DIRK3, $N_x = N_y = 150$, $CFL = 4$, $t = 10$.

2D Viscous Burger's equation

We consider the two-dimensional viscous Burger's equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) + \frac{\partial}{\partial y}(\frac{u^2}{2}) = \epsilon(\frac{\partial}{\partial x}(\nu(u)\frac{\partial u}{\partial x} + \frac{\partial}{\partial y}(\nu(u)\frac{\partial u}{\partial y}),$$

where the diffusion coefficient $\nu(u)$ is given by (5.2). The initial data is equal to 2 and -2 inside two circles of radius 0.4 centered at $(-0.5, -0.5)$ and $(0.5, 0.5)$. We discretize the convection and the diffusion terms with fifth order WENO and second order central differencing, respectively. The reference solution given by SSPRK3 at $CFL = 0.6$ is presented in Figure 5.23, where $\epsilon = 0.1$ is taken. The solutions along $y = 0$ are displayed in Figure 5.24, where $CFL = 10$ is taken for the computations of implicit methods. The results verify the improved stability of L-DIRK3.

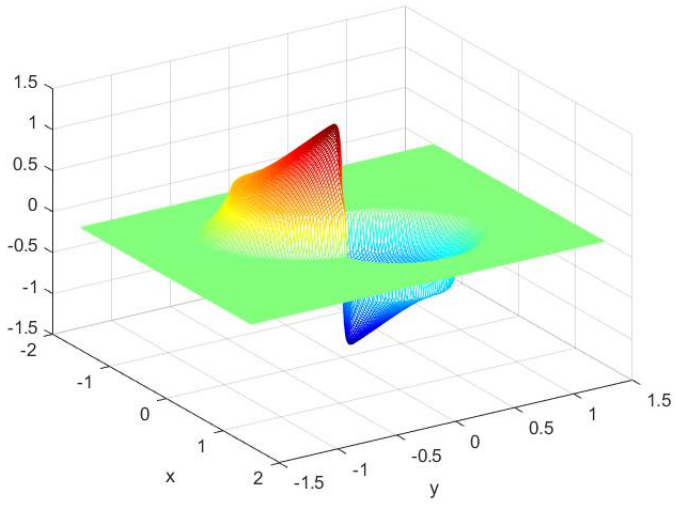


Figure 5.23: 2D viscous Burger's, reference solution, $N_x = N_y = 150$, $t = 0.5$.

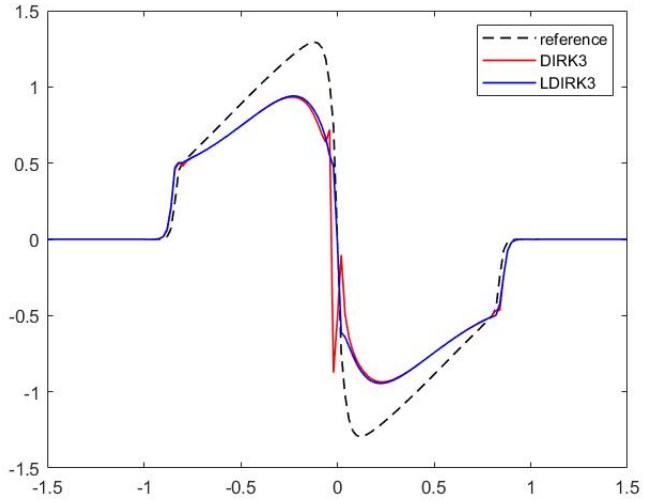


Figure 5.24: 2D viscous Burger's, solutions along $y = 0$, $N_x = N_y = 150$, $t = 0.5$.

From these numerical results, it is seen that L-DIRK3 is advantageous in the sense that it generates non-oscillatory solutions without damaging the resolution of high order spatial discretization schemes under large CFL numbers. The dimension-by-dimension extension of the scheme to multidimensional problems is validated. Further improvements may be done in several aspects. At first, the solution can still have small amplitude oscillations near discontinuities, which may be removed by improving the design of limiter $\theta_j^{(k)}$ and the way of defining $\theta_{j+\frac{1}{2}}^{(k)}$. Indeed, one may construct $\theta_{j+\frac{1}{2}}^{(k)}$ with different ways of averaging $\theta_j^{(k)}$ and $\theta_{j+1}^{(k)}$, such as $\theta_{j+\frac{1}{2}}^{(k)} = \min(\theta_j^{(k)}, \theta_{j+1}^{(k)})$. At smooth extrema, the solution has stronger dissipation since the order of accuracy in time is locally dropped to first order. This may be overcome by introducing additional shock detectors so that the time-limiter is switched off in the smooth regions.

Chapter 6: Conclusions

We conclude our study with the following.

6.1 Summary

In this thesis, we have reviewed the concepts and constructions of time-limiter schemes. The goal of these schemes is to loosen the time step restriction for high order implicit time marching, so that the increased cost by implicit iterations can be sufficiently compensated. The construction approach of time-limiter schemes is to take local convex combination of a higher order but oscillatory method with a first order unconditionally SSP method. The key ingredient is to employ local limiters to evaluate the smoothness and decide the switch between different methods. In this way, the hybrid scheme recovers the higher order time integration method in smooth regions, and locally satisfies the monotonicity condition near discontinuities by dropping the order of accuracy in time to first order.

In the earlier discussion by Duraisamy, Baeder and Liu, the time-limiters have been successfully applied to improve the stability of the second order accurate, conditionally SSP trapezoidal scheme and DIRK2 scheme under large time steps. To better preserve the high resolution of high order space discretizations such as fifth order WENO, we

extend the idea of limited time integration to the third order accurate, strongly S-stable but non-SSP DIRK3 method. The Limited-DIRK3 scheme is subsequently proposed. The previous construction of time-limiters by Duraisamy et al involves evaluations of time derivative functions, which may introduce much extra work if the reference quantity is not chosen from conservative variables. To allow greater flexibility when solving systems of equations, we proposed a simpler construction that only requires evaluations of variables. With the new limiter, one is allowed to choose an arbitrary interested reference variable based on the specific problem with only minimal cost.

In the numerical experiments, the L-DIRK3 scheme has been successfully applied in a series of one- and two-dimensional benchmark problems. The results have verified the improved stability of L-DIRK3 in non-smooth problems, where the original DIRK3 scheme can generate spurious oscillations when a large time step is taken. In spite of the reduced uniform convergence rate, the L-DIRK3 scheme well preserves the high resolution of high order WENO. We have also extended the discussion to convection-diffusion problems. In actual applications, convection-diffusion problems are the cases where a large CFL number is desired in order to produce a reasonable time step, and hence implicit methods would be preferred over explicit methods. Our computations for scalar convection-diffusion equations have shown the clear advantage of L-DIRK3 under large CFL number.

6.2 Future work

We point out some aspects of future work:

Firstly, the non-oscillatory property of time-limiter schemes is not rigorously proved. The existing SSP theory only applies to time integration methods with constant coefficients. New mathematical tools would be required to analyze the stability of time-limiter schemes which employ local and variable coefficients.

Secondly, with the current constructions, some portions of the solution by L-DIRK3 still exhibit some oscillations. Indeed, the original DIRK3 scheme is not SSP and it is challenging to completely remove the oscillations near discontinuities. It remains to explore better constructions of time-limiters in order to further remove the oscillations.

Thirdly, the application of time-limiters introduces strong dissipation at smooth extrema, leading to the order reduction of L-DIRK3. A possible approach to reduce the error is to introduce additional shock detecting mechanisms that distinguish disticontinuities and smooth extrema. This may allow us to switch off time-limiters in the smooth regions and only activate them when necessary.

Besides, the framework of time-limiter schemes may serve as a general approach to improve the stability of an arbitrary DIRK method, such as the TR-BDF2 scheme mentioned in Chapter 3. It would also be an interesting future work to examine the effect of time-limiters as applied to other DIRK methods.

With further improvements, the time-limiter schemes would have great potential to be successfully applied in complicated application problems such DNS.

Appendix A: Proof of Theorem 2.1

Proof. Consider any stage of (2.15), if we assume that $\alpha_{ij} \geq 0$ and $\beta_{ij} \geq 0$,

$$\begin{aligned}
||\mathbf{u}^{(i)}|| &= ||\sum_{j=0}^{i-1} \alpha_{ij}\mathbf{u}^{(j)} + \Delta t\beta_{ij}\mathbf{L}(\mathbf{u}^{(j)})|| = ||\sum_{j=0}^{i-1} \alpha_{ij}(\mathbf{u}^{(j)} + \Delta t\frac{\beta_{ij}}{\alpha_{ij}}\mathbf{L}(\mathbf{u}^{(j)}))|| \\
&\leq \sum_{j=0}^{i-1} \alpha_{ij}||\mathbf{u}^{(j)} + \Delta t\frac{\beta_{ij}}{\alpha_{ij}}\mathbf{L}(\mathbf{u}^{(j)})|| \\
&\leq \sum_{j=0}^{i-1} \alpha_{ij}||\mathbf{u}^{(j)}||,
\end{aligned}$$

provided $\Delta t\frac{\beta_{ij}}{\alpha_{ij}} \leq \Delta t_{FE}$. Applying this recursively from $i = 0$ to $s + 1$, we obtain

$$||\mathbf{u}^{(i)}|| \leq \sum_{j=0}^{i-1} \alpha_{ij}||\mathbf{u}^{n}|| = ||\mathbf{u}^{n}||, \quad i = 0, \cdots, s+1,$$

where, by consistency, $\sum_{j=0}^{i-1} \alpha_{ij} = 1$. Therefore, we have

$$||\mathbf{u}^{n+1}|| \leq ||\mathbf{u}^{n}||.$$

□

Appendix B: Proof of Theorem 2.2

Proof. Consider any stage of (2.16) and rearrange

$$||\mathbf{u}^{(i)} - \Delta t\beta_{ii}\mathbf{L}(\mathbf{u}^{(i)})|| = ||\sum_{j=0}^{i-1}\alpha_{ij}(\mathbf{u}^{(j)} + \Delta t\frac{\beta_{ij}}{\alpha_{ij}}\mathbf{L}(\mathbf{u}^{(j)}))||. \tag{B.1}$$

For the right hand side of (B.1), since $\alpha_{ij} \geq 0$ and $\beta_{ij} \geq 0$,

$$\begin{aligned}||\sum_{j=0}^{i-1}\alpha_{ij}(\mathbf{u}^{(j)} + \Delta t\frac{\beta_{ij}}{\alpha_{ij}}\mathbf{L}(\mathbf{u}^{(j)}))|| &\leq \sum_{j=0}^{i-1}\alpha_{ij}||\mathbf{u}^{(j)} + \Delta t\frac{\beta_{ij}}{\alpha_{ij}}\mathbf{L}(\mathbf{u}^{(j)})|| \\ &\leq \sum_{j=0}^{i-1}\alpha_{ij}||\mathbf{u}^{(j)}||,\end{aligned}$$

provided $\Delta t\frac{\beta_{ij}}{\alpha_{ij}} \leq \Delta t_{FE}$.

For the left hand side of (B.1), consider $\hat{\mathbf{u}} := \mathbf{u}^{(i)} - \Delta t\beta_{ii}\mathbf{L}(\mathbf{u}^{(i)})$. Since $\beta_{ii} \geq 0$,

$$||\mathbf{u}^{(i)}|| = ||\hat{\mathbf{u}} + (\beta_{ii}\Delta t)\mathbf{L}(\mathbf{u}^{(i)})|| \leq ||\hat{\mathbf{u}}|| = ||\mathbf{u}^{(i)} - \Delta t\beta_{ii}\mathbf{L}(\mathbf{u}^{(i)})||,$$

by the unconditional stability of backward Euler.

Hence, we have

$$||\mathbf{u}^{(i)}|| \leq \sum_{j=0}^{i-1}\alpha_{ij}||\mathbf{u}^{(j)}||.$$

Applying this recursively from $i = 0$ to $s + 1$ and notice that $\sum_{j=0}^{i-1} \alpha_{ij} = 1$, we recover

$$||\mathbf{u}^{(i)}|| \leq ||\mathbf{u}^n||, \quad i = 0, \cdots, s + 1.$$

The strong stability follows.

□

Appendix C: Verification of 5th Order WENO Finite Difference Method

To reinforce the discussions for temporal errors in chapter 5, we examine the convergence rate of our implementation of the 5th order WENO finite difference schemes in one- and two-dimensional smooth problems. The central idea of the finite difference formulation is to implement the $m-$th order WENO reconstruction on the *point values* of fluxes $\{f_j\}$, in which case the reconstructed values at interfaces, $\{f^{\pm}_{j+\frac{1}{2}}\}$, satisfy

$$\frac{f^{\pm}_{j+\frac{1}{2}} - f^{\pm}_{j-\frac{1}{2}}}{\Delta x} = \frac{\partial f}{\partial x}(x_j) + O(\Delta x^m).$$

Compared with the finite volume formulation, the advantage of the finite difference formulation is that when the order of accuracy is higher than two, the latter allows for dimension-by-dimension extension in multidimensional problems and hence requires much less computational cost than the finite volume counterpart of the same order, which involves expensive evaluations of flux surface integrals. For systems of equations, we apply the high order WENO schemes on characteristic variables in order to enhance the robustness for demanding problems. We refer more details to [3].

In the following test cases, we examine the convergence rate of the SSPRK3-WENO5 scheme. In theory, the error of the numerical solution scales at $O(\Delta x^5) +$

$O(\Delta t^3)$. To avoid the effect of temporal error, we reduce the CFL number by factor $\frac{1}{2^{\frac{2}{3}}}$ everytime the mesh size is halved, in which case we mimic a 5th order temporal convergence.

Table C.1: 1D linear advection, convergence rates of 5th order WENO finite difference method, periodic bc, 1 period of revolution.

N	CFL	L^1 err	L^2 err	L^∞ error	L^1 rate	L^2 rate	L^∞ rate
25	0.9	$4.36e^{-3}$	$5.07e^{-3}$	$9.85e^{-3}$	–	–	–
50	0.567	$2.35e^{-4}$	$2.64e^{-4}$	$6.15e^{-4}$	4.21	4.26	4.00
100	0.357	$7.05e^{-6}$	$8.05e^{-6}$	$1.91e^{-5}$	5.06	5.04	5.01
200	0.225	$1.95e^{-7}$	$2.20e^{-7}$	$4.03e^{-7}$	5.18	5.19	5.57
400	0.142	$5.29e^{-9}$	$6.16e^{-9}$	$1.23e^{-8}$	5.20	5.16	5.03

At first, we test the scheme against the one-dimensional linear wave equation $u_t + u_x = 0$ with period boundary conditions and the initial data $u_0 = \sin^4(\frac{x}{2})$ on the domain $[0, 2\pi]$. We compute the solution at $t = 2\pi$ (one period of revolution). The exact solution is as shown in Figure 5.1. Table C.1 shows the errors and convergence rates versus different mesh sizes and CFL numbers. The results indicate 5th order convergence rates in different norms when the mesh size and the CFL number are sufficiently reduced.

For two-dimensional problems, uniform mesh size $N_x = N_y = N$ is applied. We first consider the two-dimensional linear wave equation $u_t + u_x + u_y = 0$ with periodic boundary conditions on the domain $[-1, 1] \times [-1, 1]$. The initial data is given by $u_0(x, y) = \sin(\pi(x + y))$. The exact solution at time t is just given by $u(x, y, t) = u_0(x - t, y - t)$. The solutions are computed at $t = 1.0$. Figure C.1 shows the exact solution. The errors and convergence rates in different norms are given in Table C.2. The

5th order convergence rates are observed in this test case.

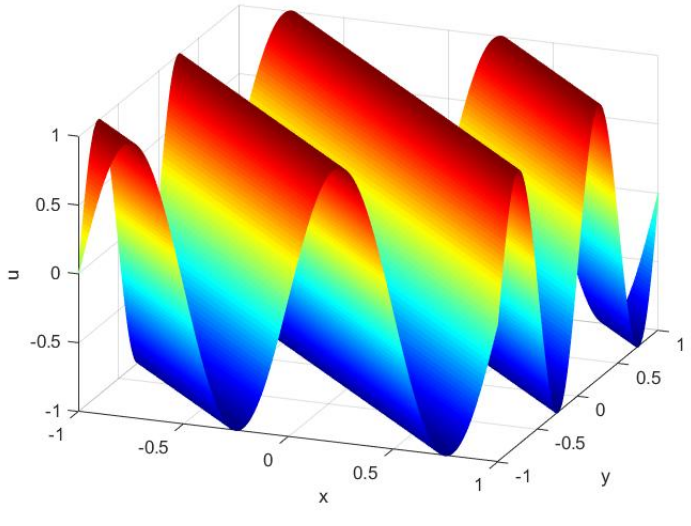


Figure C.1: 2D linear advection, $u_0(x, y) = \sin(\pi(x + y))$, exact solution at $t = 1.0$.

Table C.2: 2D linear advection, convergence rates of 5th order WENO finite difference method, periodic bc, 1 period of revolution.

N	CFL	L^1 err	L^2 err	L^∞ error	L^1 rate	L^2 rate	L^∞ rate
20	0.9	$3.46e^{-2}$	$3.78e^{-2}$	$5.73e^{-2}$	–	–	–
40	0.567	$1.13e^{-3}$	$1.25e^{-3}$	$1.76e^{-3}$	4.94	4.92	4.93
80	0.357	$3.57e^{-5}$	$3.97e^{-5}$	$5.61e^{-5}$	5.06	5.04	5.01
160	0.225	$1.95e^{-7}$	$2.20e^{-7}$	$4.03e^{-7}$	5.18	5.19	5.57
320	0.142	$5.29e^{-9}$	$6.16e^{-9}$	$1.23e^{-8}$	5.20	5.16	5.03

Next, we apply the scheme to solve a smooth travelling density perturbation for 2D Euler equations. The solution comprises of the periodic density perturbation

$$\rho(x, y, t) = 1 + 0.2 \sin(\pi(x + y - t(u + v))),$$

with constant $x-$ and $y-$velocities $u = 1.0$, $v = -0.5$, and a pressure of $p = 1.0$. Figure

C.2 presents the exact solution for density at $t = 1.0$. Table C.3 displays the errors and convergence rates during the reduction in the mesh size and CFL number. Again, a 5th order spatial convergence rate is confirmed, at least when the mesh is sufficiently refined.

These results show that our implementation of the 5th order WENO finite difference scheme achieves the expected order of accuracy. Therefore, the errors from time integration are dominating in the computations in chapter 5.

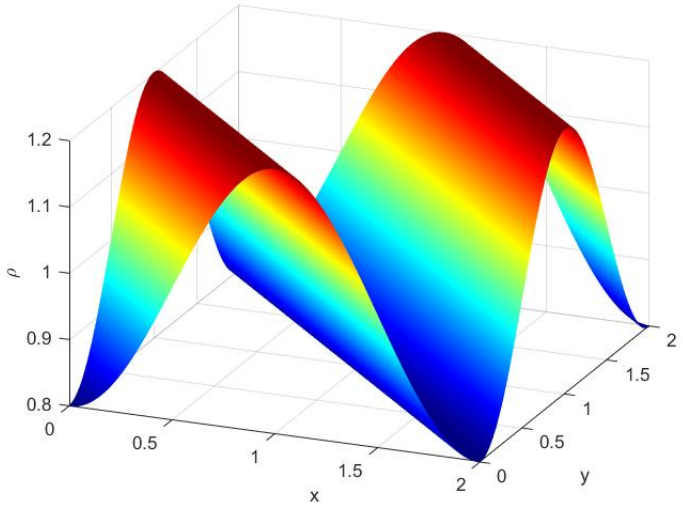


Figure C.2: Travelling 2D density perturbation for Euler equations, density at $t = 1.0$.

Table C.3: Traveling density perturbation for 2D Euler, convergence rates of 5th order WENO finite difference method in density, periodic bc, $t = 1.0$.

N	CFL	L^1 err	L^2 err	L^∞ error	L^1 rate	L^2 rate	L^∞ rate
10	0.9	$6.57e^{-3}$	$6.83e^{-3}$	$9.40e^{-3}$	–	–	–
20	0.567	$3.25e^{-4}$	$3.69e^{-4}$	$5.25e^{-4}$	4.34	4.21	4.16
40	0.357	$1.03e^{-5}$	$1.18e^{-5}$	$1.92e^{-5}$	4.98	4.97	4.77
80	0.225	$3.18e^{-7}$	$3.60e^{-7}$	$5.90e^{-7}$	5.02	5.03	5.02
160	0.142	$5.29e^{-9}$	$6.16e^{-9}$	$1.23e^{-8}$	5.20	5.16	5.03

Bibliography

[1] Amiram Harten, James M Hyman, Peter D Lax, and Barbara Keyfitz. On finite-difference approximations and entropy conditions for shocks. *Communications on pure and applied mathematics*, 29(3):297–322, 1976.

[2] Ami Harten. High resolution schemes for hyperbolic conservation laws. *Journal of computational physics*, 135(2):260–278, 1997.

[3] Chi-Wang Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In *Advanced numerical approximation of nonlinear hyperbolic equations*, pages 325–432. Springer, 1998.

[4] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43(1):89–112, 2001.

[5] Karthikeyan Duraisamy, James D Baeder, and Jian-Guo Liu. Concepts and application of time-limiters to high resolution schemes. *Journal of scientific computing*, 19(1):139–162, 2003.

[6] Roger Alexander. Diagonally implicit runge–kutta methods for stiff ode's. *SIAM Journal on Numerical Analysis*, 14(6):1006–1021, 1977.

[7] Chi-Wang Shu. Total-variation-diminishing time discretizations. *SIAM Journal on Scientific and Statistical Computing*, 9(6):1073–1084, 1988.

[8] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii. In *Upwind and High-Resolution Schemes*, pages 328–374. Springer, 1989.

[9] Sigal Gottlieb and Chi-Wang Shu. Total variation diminishing runge-kutta schemes. *Mathematics of computation*, 67(221):73–85, 1998.

[10] Raymond J Spiteri and Steven 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.

[11] Peter K Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM journal on numerical analysis*, 21(5):995–1011, 1984.

[12] Alexander Kurganov and Eitan Tadmor. New high-resolution central schemes for nonlinear conservation laws and convection–diffusion equations. *Journal of Computational Physics*, 160(1):241–282, 2000.

[13] Bram Van Leer. On the relation between the upwind-differencing schemes of godunov, engquist–osher and roe. *SIAM Journal on Scientific and statistical Computing*, 5(1):1–20, 1984.

[14] David Gottlieb and Eitan Tadmor. The cfl condition for spectral approximations to hyperbolic initial-boundary value problems. *Mathematics of Computation*, 56(194):565–588, 1991.

[15] Bernardo Cockburn and Chi-Wang 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.

[16] Ulrik S Fjordholm, Siddhartha Mishra, and Eitan Tadmor. Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM Journal on Numerical Analysis*, 50(2):544–573, 2012.

[17] Ulrik S Fjordholm, Siddhartha Mishra, and Eitan Tadmor. Entropy stable eno scheme. In *Hyperbolic Problems: Theory, Numerics and Applications (In 2 Volumes)*, pages 12–27. World Scientific, 2012.

[18] Sigal Gottlieb, David I Ketcheson, and Chi-Wang Shu. *Strong stability preserving Runge-Kutta and multistep time discretizations*. World Scientific, 2011.

[19] Hermanus WJ Lenferink. Contractivity preserving explicit linear multistep methods. *Numerische Mathematik*, 55(2):213–223, 1989.

[20] Hermanus WJ Lenferink. Contractivity-preserving implicit linear multistep methods. *Mathematics of computation*, 56(193):177–199, 1991.

[21] Ami Harten and Stanley Osher. Uniformly high-order accurate nonoscillatory schemes. i. In *Upwind and High-Resolution Schemes*, pages 187–217. Springer, 1997.

[22] A Suresh and HT Huynh. Accurate monotonicity-preserving schemes with runge–kutta time stepping. *Journal of Computational Physics*, 136(1):83–99, 1997.

[23] Hung T Huynh. Accurate monotone cubic interpolation. *SIAM Journal on Numerical Analysis*, 30(1):57–100, 1993.

[24] Luca Bonaventura and A Della Rocca. Unconditionally strong stability preserving extensions of the tr-bdf2 method. *Journal of Scientific Computing*, 70(2):859–895, 2017.

[25] Randolph E Bank, William M Coughran, Wolfgang Fichtner, Eric H Grosse, Donald J Rose, and R Kent Smith. Transient simulation of silicon devices and circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 4(4):436–451, 1985.

[26] ME Hosea and LF Shampine. Analysis and implementation of tr-bdf2. *Applied Numerical Mathematics*, 20(1-2):21–37, 1996.

[27] A Prothero and A Robinson. On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations. *Mathematics of Computation*, 28(125):145–162, 1974.

[28] Thomas Pulliam. Time accuracy and the use of implicit methods. In *11th Computational Fluid Dynamics Conference*, page 3360, 1993.

[29] Tung Chang, Gui-Qiang Chen, and Shuli Yang. On the 2-d riemann problem for the compressible euler equations i. interaction of shocks and rarefaction waves. *Discrete & Continuous Dynamical Systems*, 1(4):555, 1995.

[30] Peter D Lax and Xu-Dong Liu. Solution of two-dimensional riemann problems of gas dynamics by positive schemes. *SIAM Journal on Scientific Computing*, 19(2):319–340, 1998.

[31] Alexander Kurganov and Eitan Tadmor. Solution of two-dimensional riemann problems for gas dynamics without riemann problem solvers. *Numerical Methods for Partial Differential Equations: An International Journal*, 18(5):584–608, 2002.