IMEX and SEMI-IMPLICIT RUNGE-KUTTA SCHEMES for CFD SIMULATIONS

ARMAN ROKHZADI

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
PhD

Civil Engineering Department
Faculty of Engineering
University of Ottawa

Acknowledgements

I would like to express my sincere and deep gratitude to my supervisor, Dr. Abdolmajid Mohammadian, for his constant support, guidance, and inspiration. It is a great honor for me to know and work with Dr. Mohammadian, since he is not only a great scientist, but also a very kind, reliable, and respectful person.

The invaluable suggestions and comments received from the committee members, Drs. Ioan Nistor, Ousmane Seidou, Luc Mongeau and Amir Hakami, are highly appreciated.

This research was supported financially by Environment Canada and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

I extend my warmest appreciation and love to my wife, Sanaz, and my family, specifically my mother, Tahmineh, for her life-long dedication and support.

This thesis is dedicated to the loving memory of my grandfather, Raof, who I never stop feeling his presence.

Abstract

Numerical Weather Prediction (NWP) and climate models parametrize the effects of boundary-layer turbulence as a diffusive process, dependent on a diffusion coefficient, which appears as nonlinear terms in the governing equations. In the advection dominated zone of the boundary layer and in the free atmosphere, the air flow supports different wave motions, with the fastest being the sound waves. Time integrations of these terms, in both zones, need to be implicit otherwise they impractically restrict the stable time step sizes. At the same time, implicit schemes may lose accuracy compared to explicit schemes in the same level, which is due to dispersion error associated with these schemes. Furthermore, the implicit schemes need iterative approaches like the Newton-Raphson method. Therefore, the combination of implicit and explicit methods, called IMEX or semi-implicit, has extensively attracted attention. In the combined method, the linear part of the equation as well as the fast wave terms are treated by the implicit part and the rest is calculated by the explicit scheme. Meanwhile, minimizing the dissipation and dispersion errors can enhance the performance of time integration schemes, since the stability and accuracy will be restricted by these inevitable errors.

Hence, the target of this thesis is to increase the stability range, while obtaining accurate solutions by using IMEX and semi-implicit time integration methods. Therefore, a comprehensive effort has been made toward minimizing the numerical errors to develop new Runge-Kutta schemes, in IMEX and semi-implicit forms, to temporally integrate the governing equations in the atmospheric field so that the stability is extended and accuracy is improved, compared to the previous schemes.

At the first step, the A-stability and the Strong Stability Preserving (SSP) optimized properties were compared as two essential properties of the time integration schemes. It was shown that both properties attempt to minimize the dissipation and dispersion errors, but in two different aspects. The SSP optimized property focuses on minimizing the errors to increase the accuracy limits, while the A-stability property tries to extend the range of stability. It was shown that the combination of both properties is essential in the field of interest. Moreover, the A-stability property was found as an essential property to accelerate the steady state solutions.

Afterward, the dissipation and dispersion errors, generated by three-stage second order IMEX Runge-Kutta scheme were minimized, while the proposed scheme, so called IMEX-SSP2(2,3,2) enjoys the A-stability and SSP properties. A practical governing equation set in the atmospheric field, so called compressible Boussinesq equations set, was calculated using the new IMEX scheme and the results were compared to one well-known IMEX scheme in the literature, i.e. ARK2(2,3,2), which is an abbreviation of Additive Runge-Kutta. Note that, the ARK2(2,3,2) was compared to various types of IMEX Runge-Kutta schemes and it was found as the more efficient scheme in the atmospheric fields (Weller et al., 2013). It was shown that the IMEX-SSP2(2,3,2) could improve the accuracy and extend the range of stable time step sizes as well. Through the van der Pol test case, it was shown that the ARK2(2,3,2) with L-stability property may decline to the first order in the calculation of stiff limit, while IMEX-SSP2(2,3,2), with A-stability property, is able to retain the assigned second order of accuracy. Therefore, it was concluded that the L-stability property, due to restrictive conditions associated with, may weaken the time integration's performance, compared to the A-stability property. The ability of the IMEX-SSP2(2,3,2) was proved in solving different case, which is the inviscid Burger equation in spherical coordinate system by using a realistic initial condition dataset.

In the next step, it was attempted to maximize the non-negativity property associated with the numerical stability function of three-stage third order Diagonally Implicit Runge-Kutta (DIRK) schemes. It was shown that the non-negativity has direct relation with non-oscillatory behaviors. Two new DIRK schemes with A- and L-stability properties, respectively, were developed and compared to the SSP(3,3), which obtains the SSP optimized property in the same class of DIRK schemes. The SSP optimized property was found to be more beneficial for the inviscid (advection dominated) flows, since in the von Neumann stability analysis, the SSP optimized property provides more nonnegative region for the imaginary component of the stability function. However, in most practical cases, i.e. the viscous (advection diffusion) flows, the nonnegative property is needed for both real and imaginary components of the stability function. Therefore, the SSP optimized property, individually, is not helpful, unless mixed with the A-stability property. Meanwhile, the A- and L-stability properties were compared as well. The intention is to find how these properties influence the DIRK schemes' performances. The A-stability property was found as preserving the SSP property more than the L-stability property. Moreover, the proposed A-stable scheme tolerates larger Courant Friedrichs Lewy (CFL) number, while preserving the accuracy and non-oscillatory computations. This fact was proved in calculating different test cases, including compressible Euler and nonlinear viscous Burger equations.

Finally, the time integration of the boundary layer flows was investigated as well. The nonlinearity associated with the diffusion coefficient makes the implicit scheme impractical, while the explicit scheme inefficiently limits the stable time step sizes. By using the DIRK scheme, a new semi-implicit approach was proposed, in which the diffusion coefficient at each internal stage is calculated by a weight-averaged combination of the solutions at current internal stage and previous time step, in which the time integration can benefit from both explicit and implicit advantages. As shown, the accuracy was improved, which is due to engaging the explicit solutions and the stability was extended due to taking advantages of implicit scheme. It was found that the nominated semi-implicit method results in less dissipation error, more accurate solutions and less CPU time usage, compared to the implicit schemes, and it enjoys larger range of stable time steps than other semi-implicit approaches in the literature.

List of Figures

List of Tables

Glossary and Units

b (N) , buoyancy

c_p ($Jkg^{-1}K^{-1}$), specific heat capacity

C (Kgm^{-3}), concentration

f_c (s^{-1}), Coriolis frequency

F (Nm^{-2}), flux term

g (ms^{-2}), gravity constant

K_0 ($Wm^{-1}K^{-1}$), diffusion coefficient

L_p (Jkg^{-1}), latent heat

N (s^{-1}), Brunt–Väisälä frequency

P_0 (Nm^{-2}), reference pressure

P (Nm^{-2}), pressure

q (gg^{-1}), specific humidity

Q_j^* (J), j^{th} component of net radiation

R ($Jkg^{-1}K^{-1}$), gas constant

S_C (N), body source

t (s), time

T (K), temperature

u_i (ms^{-1}), the ith component of velocity vector

x_j (m), the jth component of spatial vector

z (m), height from the surface

ρ (Kgm^{-3}), density

δ_{i3}, Kronecker

τ_{ij} (Nm^{-2}), the ijth component of shear stress

ε_{ijk}, alternating tensor

Θ (K), potential temperature

ν_θ (m^2s^{-1}), thermal diffusivity

ν_C (m^2s^{-1}), molecular diffusivity

$\bar{\ }$, Mean value component

$'$, Perturbation component

Abbreviations

ABL: Atmospheric Boundary Layer

CFD: Computational Finite Difference

CFL: Courant Friedrichs Lewy

CPU: Central Processing Unit

DAE: Differential Algebraic Equations

DIRK: Diagonally Implicit Runge-Kutta

ERK: Explicit Runge-Kutta

EDIRK: Explicit Diagonally Implicit Runge-Kutta

FE: Forward Euler

FI-RK: Fully Implicit Runge-Kutta

GFD: Geophysical Fluid Dynamic

GPU: Graphical Processing Unit

GS: Guass-Seidel

HEVI: Horizontally Explicit Vertically Implicit

IMEX: Implicit-Explicit

IVP: Initial Value Problem

LHS: Left Hand Side

ME BDF: Modified Extended Backward Differentiation Formulae

MOL: Method of Line

NWP: Numerical Weather Prediction

ODE: Ordinary Differential Equation

PC: Personal Computer

PDE: Partial Differential Equation

RAA: Range of Acceptable Amplification

RAP: Range of Acceptable Phase

RKC: Runge-Kutta-Chebyshev

RMS: Root Mean Square

RHS: Right Hand Side

SE-RK : Semi Explicit Runge-Kutta

SI-PC : Semi Implicit Predictor Corrector

SI-RK : Semi Implicit Runge-Kutta

SOR: Successive Over Relaxation

SPP: Single Perturbation Problem

SSP: Strong Stability preserving

TVD: Total Variation Diminishing

Ufpreb: u forward pressure backward

Ufpref: u forward pressure forward

WENO: Weighted Essential Non Oscillatory

Contents

Chapter 1. Introduction

This chapter provides some introductory descriptions of the thesis, including the background of the subject (Section 1.1). Then, the motivation that encourages the author to develop this effort will be explained in Section 1.2. Section 1.3 refers to the objectives and the scope of the work. The novelty of the thesis will be addressed in Section 1.4. Finally, Section 1.5 outlines the structure of the thesis.

1.1.Background

The air flow plays an important role in the land surface climate and the exchange of heat, greenhouse gases and pollutant emission to the free atmosphere. Weather forecasting, i.e. the predictions of air flow behaviors, is based on computer models, used to simulate this unsteady and complicated flow through mass, momentum and energy conservation laws, in which many atmospheric factors should be taken into account as well. Massive computational power is needed to solve the equations that govern the atmosphere flow. The inaccuracy of forecasting is due to the chaotic nature of the atmosphere and the approximations implemented in the models as well as the numerical errors. Errors are involved in measuring the initial conditions, incomplete understanding of atmospheric processes, and in the numerical calculations. In addition to extensive efforts to further clarify how the physical processes involved in the practice could be best fit in the mathematical models, the extensive researches have been done to eliminate or minimize the numerical errors generated by the numerical methods, employed to discretize the derivative terms, including spatial and temporal ones. Otherwise they may lead the calculation to less accurate solutions or even the computations may blow up. Although, the numerical calculations is the great concern in all numerical computation fields, the NWP seems to be more distinctive since the numerical errors are going to be accumulated and spoil the quality of the predictions, especially, as the difference between current time and the time at which the forecast is made increases.

The Runge-Kutta schemes are recognized as single-step schemes since they just involve two time step levels. Therefore, they are not concerned with the computational error, which is the error associated with the multi-step methods due to required virtual initial conditions. At the meantime, they are called multi-stage methods because they use some internal stages between two time levels for the calculations. These temporal schemes have the ability to cover the well-known time integrations, such as Forward Euler (FE), Crank-Nicolson and Backward Euler methods. Meanwhile, they are able to offer some free parameters, more than enough, so that these parameters can be employed to feed the time integrations with other specific properties. In this thesis, it is aimed to study the Runge-Kutta scheme and examine its abilities in improving the stability limits while providing accurate solutions.

1.2. Motivation

Most explicit schemes have the ability to eliminate the dispersion errors (Higueras et al., 2014). Therefore they are likely to preserve the accuracy more than implicit types in the same class of order of accuracy (Higueras et al., 2014). On the other hand, the implicit schemes are more interesting since they can tolerate larger time step sizes. Within the thin boundary layer, near the earth surface, the diffusion terms, which are responsible for transporting the air qualities, are more dominant since the vertical gradients change rapidly. These terms mostly are simulated by nonlinear diffusion coefficients. This nonlinearity makes the explicit schemes too restrictive due to so small tolerable time step sizes. Meanwhile, the implicit schemes seem to be impractical due to expensive iteratively calculations. So many researches have been done through using predictor-corrector methods, e.g. Clancy and Pudykiewicz (2013) and Diamantakis et al. (2006), in which the semi implicit approaches were introduced so that the calculations can take advantages of accurate solutions associated with the explicit scheme as well as larger time step sizes offered by the implicit methods.

In the free stream zone, the role of acoustic and gravity terms become more significant. Since the air flows is nearly incompressible, the acoustic terms have no direct meteorological significance, but, they impose stability restrictions due to the fast frequencies associated with themselves. On the other hand, the gravity terms convey physically accurate solutions, which need them to be calculated more precisely. Extensive studies have been conducted, aiming at increasing the time range of the forecast, while preserving the reasonable accuracy tolerances for the computed data (Durran and Blossey, 2011; Giraldo et al., 2013 and Weller et al., 2013). The combination of the implicit and explicit schemes has been proposed in order to circumvent the stability restrictions of the fast wave terms, which are going to be treated by the implicit schemes, while resolving the slow wave terms, using the explicit schemes, in order to provide accurate solutions.

The Runge-Kutta schemes have been employed in many studies to calculate the atmospheric governing equations in both boundary layer and free stream zones (Nazari et al., 2014; Giraldo et al., 2013 and Weller et al., 2013), either the explicit and implicit schemes individually or as the IMEX combinations. It has been proposed to use a combination of implicit and explicit (IMEX) Runge-Kutta schemes such that the explicit part is responsible for physically accurate terms, which are mainly recognized as slow wave terms, and implicit schemes treat other terms, which restrict the time step sizes, e.g. fast-wave terms (Durran, 2010).

The inevitable numerical errors, corresponding to any numerical methods, could be classified as dissipation and dispersion errors, which may result in unstable calculations and inaccurate solutions. These errors would be augmented when the fast wave and nonlinear diffusive terms become dominating, compared to the remaining terms involved in the calculations, especially, when the fine gird sizes should be used to capture the small length-scale processes, the case which is known as stiff limits. Minimizing the mentioned errors could enhance the performance of the time integration schemes, applied to the atmospheric governing equations and eventually the stability and accuracy could be improved.

The ability of Runge-Kutta scheme in providing free parameters motivates the author to attempt to minimize the dissipation and dispersion errors. In this regard, the combination of the IMEX Runge-Kutta scheme applied to the free stream zone and the semi-implicit Runge-Kutta scheme applied to the thin boundary layer have been analyzed in the linearized Fourier space, such that the stability analysis is possible. Therefore, using the free parameters can help optimize the numerical errors and eventually offer improving the stability and accuracy.

1.3. Research scope and objectives

As described, the inaccuracy of the calculations is due to the chaotic nature of the atmosphere and the approximations implemented in the models as well as the numerical errors. Errors are involved in measuring the initial conditions, incomplete understanding of atmospheric processes, and in the numerical calculations. This thesis only focuses on the errors generated by the numerical discretization schemes. In this regard, the main focus is on the time integration methods, since in most cases of interest, the central scheme is applied to the spatial derivative terms. Therefore, the inaccuracy associated with the approximation of the turbulent shear stress terms within the boundary layer is excluded. The quality of the weather forecast depends on modelling the atmospheric disturbances that evolve on much slower timescales, e.g. Gravity and Rossby waves. Investigation of the possible inaccuracy corresponded to these disturbances is also found as out of scope of this study.

The objective of this thesis is to introduce new Runge-Kutta schemes, which are expected to help extend the range of stable time step sizes and improve the accuracy of the solutions, focusing on the NWP applications. These objectives include:

1- Minimizing the dissipation and dispersion errors through using a well-known integral function which combines the dissipation and dispersion errors. This integral was reform to fit the IMEX Runge-Kutta schemes applied to two frequencies, i.e. slow and fast waves, system.

2- It was shown that the A-stability and SSP optimized properties both attempt to minimize the dissipation and dispersion errors, but, for different application. Therefore it seems more effective to combine these properties in applying to the transient flows.

3- Different stability properties were also investigated to observe their influences on the performance of the time integration schemes. These properties include A- and L-stability properties as well as the SSP optimized and stiffly accurate properties.

4- The relation between nonnegative property, associated with the numerical stability function, and the non-oscillatory behaviors was shown. This non-negativity was maximized for the implicit Runge-Kutta schemes and in this regard the A- and L-stability properties were compared to the SSP optimized property as well.

5- A new semi-implicit approach was proposed to apply to the nonlinear governing equation, corresponding to the transient boundary layer flows. It was shown that this approach can help to further control the error propagations and eventually the stability and accuracy improvements were occurred compared to the previous methods.

The main concern is to use the combination of the implicit and explicit schemes to let the nominated schemes enjoy the advantages of both explicit and implicit schemes, either in the form of IMEX Runge-Kutta scheme, applied to the free stream zone, or in the form of semi-implicit Runge-Kutta scheme, applied to the boundary layer.

In the first step, the governing equations, whether those of the boundary layer or those of free stream zone, should be linearized in order to let the linear stability analysis be done. Eventually, it is attempted to optimize the associated errors in order to enhance the stability and quality of the solutions. Furthermore, this thesis investigates the influences of different stability properties corresponding to the Runge-Kutta schemes, i.e. A- and L-stability and SSP optimized properties, as well as the stiffly accurate property in order to find how they assist to achieve the target of the thesis, which is the stability and accuracy improvement. Finally, these new schemes are applied to the various test cases in order to observe and prove their abilities in promoting the quality of the computations. These cases include the governing equations, which may be served in the practice, e.g. compressible Boussinesq, or some other equations, in which the role of either advection, diffusion or source terms are intensified, to examine the performance of the time integrations in calculating these terms, e.g. vertically diffusion turbulent boundary layer equation, Euler compressible, viscous Burger equations or the linear advection with nonlinear source term.

1.4. Novelty of the research and advancement in the state of the art

The main concern in CFD modelling is more stable calculations and qualified solutions with enough accuracy. This thesis includes investigation of the Runge-Kutta schemes in atmospheric fields. The focus is mainly on increasing the range of stable time-step sizes as well as improving the accuracy. Due to involving some internal stages, the Runge-Kutta schemes are recognized as multi-stage methods. In most cases, using these internal stages leaves some free parameters that can be employed to optimize some desired properties of the time integration schemes. The optimization of dissipation and dispersion errors is considered as the main objective of this thesis, since they are known as the main source of instability and inaccuracy. This target is achieved for various arrangements of temporal scheme implementations. For the free stream zone, where the acoustic and gravity terms are more significant, the optimization is done using the IMEX Runge-Kutta scheme. For the boundary layer, in which the diffusion terms impose stability restriction, the semi-implicit Runge-Kutta scheme is employed. To capture the diffusion properties within the thin boundary layer, more effectively, a new approach is introduced, by which the implicit scheme can enjoy both explicit and implicit advantages. Therefore, the new Runge-Kutta schemes are developed, by which, as shown, the stability will be extended and the accuracy is improved for both boundary layer and free stream zones, compared to the other Runge-Kutta schemes introduced in the literature.

In this regard, different stability properties, associated with the Runge-Kutta schemes, were investigated as well to see how they may influence the performance of the Runge-Kutta scheme and how they help achieve the targets. The A-stability property imposes the stability function to be less than one in the infinite limits, i.e. $|R(-\infty)| \leq 1$, while the L-stability property impose the stability function to tend to zero in the infinite limit. It was shown that the L-stability property, compared to the A-stability property, may weaken the scheme's performance, in the sense of order reduction in the calculation of stiff limits. The SSP optimized property which is identified by a new parameter, called radius of absolute monotonicity, imposes the Runge-Kutta scheme to be a convex combination of the FE method, such that the monotonic behavior of the FE method can be held by the solution. It was shown that the A-stability and SSP optimized properties, both, aim at minimizing the dissipation and dispersion errors, but, for different applications. The SSP optimized property tries to extend the accuracy limits, while A-stability property tries to upper bound the stability in the infinite limits.

The non-negativity of the numerical stability function is shown that has direct relation with the non-oscillatory behavior. The optimization of this property associated with the stability function of the implicit Runge-Kutta schemes was another aspect involved in this study. The scheme with SSP optimized property was found as beneficial for the inviscid flows, or equivalently called advection dominated flows, since this property can help the imaginary component of the stability function be nonnegative for larger range of the variable. However, in practices, which are likely the viscous flows, or equivalently called advection diffusion flows, the SSP optimized property, individually, is not helpful, unless mixed with the A-stability property. It was shown that the combination of the SSP and A-stability properties can lead to larger range of non-negativity for the stability function, including both real and imaginary components, which help the time integration provide less oscillation in practical cases.

Time integration of the boundary layer flow is another concern, which is restricted by diffusion terms, due to rapid changes of the gradient terms. The diffusion terms, specifically, turbulent shear stress terms are approximated by a nonlinear model due to variable diffusion coefficient. This nonlinearity makes the explicit schemes too expensive and the implicit scheme is impractical, due to expensive iterative approach. By using the Diagonally Implicit Runge-Kutta (DIRK) scheme a new approach was proposed, which benefits from both explicit and implicit properties of the time integration. Therefore, as shown, it is expected that the new approach improves the accuracy of the calculation due to involving the explicit property as well as extends the range of stability, which is provided by the implicit property. Moreover, by comparison with other approaches, proposed in the literature, e.g. predictor-corrector method, it was shown that the new approach is more efficient, in terms of stability, accuracy and the calculation cost as well.

1.5. Organization of the thesis

This thesis is organized as follows. Chapter 2 will present the literature review, in which the theory of the governing equations will be explained (Section 2.1). The theory describes equations which govern the mass, momentum, heat and scalar transportations. This section also explains the turbulent decomposition method applied to both boundary layer and free atmosphere zones. Since this thesis has focused on the Runge-Kutta time integration schemes, the developing of these temporal schemes will be reviewed in Section 2.2. Moreover, the structure of Runge-Kutta schemes and their implementations to the time-dependent equations will be discussed. Afterward, the previous studies, in which the Runge-Kutta schemes have been applied to the atmospheric field will be explained (Section 2.3). In this section, the IMEX Runge-Kutta schemes employed in the parallel computations for the atmospheric field will be reviewed as well. Section 2.4 proves that the optimization process, employed in this thesis, could be applicable for many governing equations and any coordinate systems. The outline and contributions of this thesis will be found in Section 2.5, in which the achievements and improvements compared to previous works will be discussed. Section 2.6 suggested some possible future works. Concluding remarks and summary will be presented in Section 2.7.

The outcome of this thesis was constructed based on four publishable papers and one unpublished technical note, which will be provided in Chapters 3 to 7. Provided in Chapter 3 is an article, in which the relative stability function was discussed in the sense of optimization of dissipation and dispersion errors, and it was published in the peer reviewed Journal of Communication in Computational Physics. Chapter 4 is another article, in which the new IMEX scheme, so called IMEX-SSP2(2,3,2), was developed and it was published in Journal of Advances in Modeling Earth Systems. In the continuation of Chapter 4, in order to further examine the performance of the IMEX-SSP2(2,3,2), Chapter 5 contains an unpublished technical note, in which the performance of the IMEX-SSP2(2,3,2) was examined in different application. It is attempted to solve the inviscid Burger equation using the North America continent as the computational domain and a relevant realistic initial condition. Chapter 6 presents an article about the effect of the non-negativity property of the DIRK scheme, which is currently under review. Finally, Chapter 7 includes the article, in which the new semi-implicit DIRK scheme was proposed for the vertical diffusion of the boundary layer flows. This article is submitted to a peer reviewed journal and is still under review.

Chapter 2. Background and literature review

2.1. Introduction

In this chapter, the governing equations will be described in two parts; boundary layer processes and free atmosphere zone respectively (Section 2.2). Since this thesis focuses on the application of Runge-Kutta methods, the development of these schemes will be reviewed in the previous studies (Section 2.3). In this section the general formula and essential properties of these time integrators will be provided as well. Afterward, the origin of the linear stability analysis will be discussed in Section 2.4. The previous numerical investigations of the atmospheric field, relevant with this thesis, will be reviewed (Section 2.5). Research gaps and contributions of the present study will be explained in Section 2.6. Some suggested future works and the summary will end up this chapter in Sections 2.7 and 2.8 respectively.

2.2. Theoretical background (Governing equations)

The air flow governing equations are derived based on the conservation laws. Five equations form the foundation based on equation state and conservation of mass, momentum, moisture and heat. Additional equations for scalar quantities such as pollutant concentration may be added. In order to find how turbulent property contributes to the governing equations the common decomposition of the variables into their mean and fluctuating values will be presented for each equation as well.

2.2.1. Equation of state

The ideal gas can adequately describe the state of air

$$\mathrm{P} = \rho \mathrm{RT} \tag{1}$$

in which P is the pressure, ρ is the density of moist air, T is the absolute temperature and R is the gas constant, which is $\mathrm{R} = 287\ \mathrm{J.K^{-1}kg^{-1}}$ for the dry air.

Applying the Reynolds decomposition, splitting the variables into mean and fluctuation parts and then taking average, results in the following equation of state

$$\overline{P} = \overline{\rho} R \overline{T} + R \overline{\rho' T'} \tag{2}$$

It can be shown that the second term is negligible compared to the first one (Stull, 1988).

2.2.2. Conservation of mass

This equation, commonly called continuity equation, is written as

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_i}{\partial x_i} = 0 \tag{3}$$

or equivalently

$$\frac{D\rho}{Dt} + \rho \frac{\partial u_i}{\partial x_i} = 0 \tag{4}$$

It can be shown that, in the most cases of interest, $\frac{1}{\rho}\frac{D\rho}{Dt} \ll \frac{\partial u_i}{\partial x_i}$, which is the incompressibility approximation or Boussinesq approximation (Stull, 1988). Therefore, the continuity equation becomes

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{5}$$

The appropriate continuity equation associated with turbulent flow could be found as

$$\frac{\partial \overline{u}_i}{\partial x_i} = 0 \tag{6}$$

2.2.3. Conservation of momentum

The general form of the momentum equation included all terms would be as:

$$\frac{\partial u_i}{\partial t}+u_j\frac{\partial u_i}{\partial x_j}=-\delta_{i3}g-\varepsilon_{ijk}f_c u_j-\frac{1}{\rho}\frac{\partial P}{\partial x_i}+\frac{1}{\rho}\frac{\partial \tau_{ij}}{\partial x_j} \tag{7}$$

The first term in the Right Hand Side (RHS) represents the gravity acting in the vertical direction. The second term deals with Coriolis effect due to the Earth's rotation, which in most cases is neglected. The third term is gradient pressure and the last one shows the shear stresses due to viscosity.

The Reynolds averaging let the turbulent shear stresses appear in the mean momentum conservation equations:

$$\frac{\partial \bar{u}_i}{\partial t}+\bar{u}_j\frac{\partial \bar{u}_i}{\partial x_j}+\frac{\partial \overline{u_j' u_i'}}{\partial x_j}=-\delta_{i3}g-\varepsilon_{ij3}f_c\bar{u}_j-\frac{1}{\bar{\rho}}\frac{\partial \bar{P}}{\partial x_i}+\frac{1}{\bar{\rho}}\frac{\partial \overline{\tau_{ij}}}{\partial x_j} \tag{8}$$

The last term in the Left Hand Side (LHS) is known as turbulent shear stresses, which needs to be modelled using turbulence closures.

2.2.4. Conservation of moisture

The conservation of water substance can be written as:

$$\frac{\partial q}{\partial t}+u_j\frac{\partial q}{\partial x_j}=\nu_q\frac{\partial^2 q}{\partial x_j^2} \tag{9}$$

where q is the specific humidity. The convenient relation for turbulent flow can be found as:

$$\frac{\partial \bar{q}}{\partial t}+\bar{u}_j\frac{\partial \bar{q}}{\partial x_j}+\frac{\partial \overline{u_j' q'}}{\partial x_j}=\nu_q\frac{\partial^2 \bar{q}}{\partial x_j^2} \tag{10}$$

2.2.5. Conservation of heat

The first law of thermodynamics describes the conservation of enthalpy, which includes contributions from both sensible and latent heat transport. In other words, the water vapor in air not only transports sensible heat associated with its temperature, but also it has the potential to release or absorb additional latent heat during any phase changes might be occurred. To simplify the equations describing enthalpy conservation often the phase change information, E, is utilized in the moisture conservation equations (Stull, 1988).

$$\frac{\partial \theta}{\partial t} + u_j \frac{\partial \theta}{\partial x_j} = \nu_\theta \frac{\partial^2 \theta}{\partial x_j^2} - \frac{1}{\rho c_p} \frac{\partial Q_j^*}{\partial x_j} - \frac{L_p E}{\rho c_p} \tag{11}$$

where ν_θ is the thermal diffusivity and L_p is the latent heat associated with the phase change of E. Q_j^* is the component of net radiation in the j^{th} direction. The parameter θ is known as potential temperature and it is defined as:

$$\theta = T \left(\frac{P}{P_0}\right)^{-\frac{R}{c_p}} \tag{12}$$

where, P_0 is the constant reference pressure. Potential temperature is the more important quantity than the actual temperature, since it is not affected by the physical lifting or sinking associated with flow over obstacles or large-scale atmospheric turbulence. A parcel of air moving over a small mountain will expand and cool as it ascends the slope, then compress and warm as it descends on the other side. However, the potential temperature will not vary in the absence of heating, cooling, evaporation, or condensation. Since parcels with the same potential temperature can be exchanged without work or heating, lines of constant potential temperature are natural flow pathways. Similar to the above equations, the turbulent flow can be approximated as:

$$\frac{\partial \bar{\theta}}{\partial t} + \bar{u}_j \frac{\partial \bar{\theta}}{\partial x_j} + \frac{\partial \overline{u_j' \theta'}}{\partial x_j} = \nu_\theta \frac{\partial^2 \bar{\theta}}{\partial x_j^2} - \frac{1}{\bar{\rho} c_p} \frac{\partial \overline{Q_j^*}}{\partial x_j} - \frac{L_p E}{\bar{\rho} c_p} \tag{13}$$

2.2.6. Conservation of scalar quantity

Let C be the concentration of a scalar (mass per unit volume) such as a tracer in the atmosphere. The conservation of tracer mass requires that:

$$\frac{\partial C}{\partial t} + u_j \frac{\partial C}{\partial x_j} = \nu_C \frac{\partial^2 C}{\partial x^2} + S_C \tag{14}$$

where ν_C is the molecular diffusion coefficient and S_C is the body source term. The turbulence equation is:

$$\frac{\partial \bar{C}}{\partial t} + \bar{u}_j \frac{\partial \bar{C}}{\partial x_j} + \frac{\partial \overline{u_j' c'}}{\partial x_j} = \nu_c \frac{\partial^2 \bar{C}}{\partial x_j^2} + S_C \tag{15}$$

Boundary layer

Due to high Reynolds number, the Atmospheric Boundary Layer (ABL) is generally turbulent with intensive turbulent fluxes. In NWP and climate models, these fluxes are parametrized using low-order closure schemes. Such schemes approximate the effects of turbulence as a diffusive process with a diffusion coefficient. This is due to the thinness of the layers typically required to account for the large flux gradients near the Earth's surface. The diffusion coefficients are often nonlinear. Hence, within the boundary layer and close to the interfaces the diffusion terms become dominated. Therefore, the turbulence fluxes represented by F will be modeled as (Diamantakis et al., 2005):

$$F = K \frac{\partial X}{\partial z} \tag{16}$$

where z denotes the height from the surface of the Earth, X can be a wind component, temperature, moisture or any tracer and F is the flux of X due to vertical diffusion within the boundary layer. In this first-order closure scheme, K is called the diffusion coefficient, which often depends on X, i.e. $K \equiv K(X)$ implying that the PDE is nonlinear. F is modeled as follows, which is simple enough to analyze but captures the essential nonlinearity of the problem (Kalnay and Kanamitsu, 1988).

$$F = (K_0 X^P)X \tag{17}$$

with K_0, positive and constant, is a damping factor and plays the same role as the diffusion coefficient K(X). P is a non-negative parameter which determines the degree of nonlinearity of the system, with $P = 0$ corresponding to the linear case.

Free atmosphere

In the rest of boundary layer, far away from the earth surface, and in the free atmosphere, the advection transport is more significant so the equations are assumed as inviscid flow. The Euler equations for the inviscid isentropic motion of a perfect gas can be expressed as:

$$\frac{D\mathbf{v}}{Dt} + \frac{1}{\rho}\nabla p = -g\mathbf{k} \tag{18}$$

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho\mathbf{v}) = 0 \tag{19}$$

$$\frac{D\theta}{Dt} = 0 \tag{20}$$

where Coriolis forces have been neglected. Note that in many geophysical fluids the gravitational acceleration and the vertical pressure gradient are nearly balanced, i.e. hydrostatically balanced. Therefore, commonly, the variables are decomposed to steady and finite-amplitude perturbations about those reference profiles:

$$\rho(x, y, z, t) = \bar{\rho}(z) + \acute{\rho}(x, y, z, t) \tag{21}$$

$$p(x, y, z, t) = \bar{p}(z) + \acute{p}(x, y, z, t) \tag{22}$$

$$\theta(x, y, z, t) = \bar{\theta}(z) + \acute{\theta}(x, y, z, t) \tag{23}$$

in which $\frac{d\bar{p}}{dz} = -\bar{\rho}g$ is the hydrostatically balanced state.

The essential properties of the governing equations can be more simply examined through the two-dimensional assumption. The velocity components are decomposed as:

$$u(x, z, t) = U + \acute{u}(x, z, t) \tag{24}$$

$$w(x, z, t) = \acute{w}(x, z, t) \tag{25}$$

The reference basic state is assumed as hydrostatically balanced isothermal state, in which the vertical velocity is zero. Some simplifications can be achieved by removing the influence of the decreasing the mean density with height through the transformations:

$$b = -g\frac{\rho - \bar{\rho}(z)}{\rho_0} \tag{26}$$

$$P = \frac{p - \bar{p}(z)}{\rho_0} \tag{27}$$

$$N^2 = -\frac{g}{\rho_0}\frac{d\bar{\rho}(z)}{dz} \tag{28}$$

in which N is the Brunt–Väisälä frequency.

The following governing equations can be derived after some extra manipulations such as defining the Exner function pressure. The reader is referred to Durran (2010) for more details.

$$\frac{D\mathbf{v}}{Dt} + \nabla P = b\mathbf{k} \tag{29}$$

$$\frac{\mathrm{Db}}{\mathrm{Dt}} + \mathrm{N}^2\mathrm{w} = 0 \tag{30}$$

$$\frac{\mathrm{DP}}{\mathrm{Dt}} + \mathrm{c}_\mathrm{s}^2\nabla.\mathbf{v} = 0 \tag{31}$$

The Boussinseq approximation includes additional simplification rather than incompressibility approximation. This approximation that is common in geophysical fluid dynamics neglect the effects of density variations on the mass balance in the continuity and on inertia in the momentum equation, but includes the effect of density variations on buoyancy forces. These equations can also be regarded as the compressible Boussinesq equations. Similar to the standard Boussinesq approximation, the compressible Boussinesq system neglects the influence of density variations on inertia, while retaining the influence of density variations on buoyancy and assumes that buoyancy is conserved following a fluid parcel. In contrast to the standard Boussinesq system, the compressible Boussinesq system retains the influence of density fluctuations on pressure and thereby allows the formation of the prognostic pressure equation.

2.3. Runge-Kutta scheme (Literature and theory)

The governing Partial Differential Equations (PDEs) are spatially discretized to obtain a system of time dependent Ordinary Differential Equations (ODEs), so called Method of Lines (MOL.)

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{dt}} = \mathbf{f}(\mathrm{t}, \mathbf{u}), \quad \mathbf{u}(\mathrm{a}) = \boldsymbol{u_0}, \qquad \mathrm{t} \in [\mathrm{a}, \mathrm{b}] \tag{32}$$

Runge-Kutta time integration schemes calculate the source terms, $\mathrm{f}(\mathrm{t}, \mathbf{u})$, using correlations of solutions on specific internal stages, as well as the next time step, and then proceed one time step forward, while using the solutions in the last time step as the initial value. This could be considered as a memory saving benefit of multi-stages schemes such as Runge-Kutta schemes compared to multistep models. Traditionally, the internal stages and correlation weight coefficients are represented by vector **c** and matrix **A** and vector **b** respectively. These coefficients are illustrated in Butcher Tableau as shown in Table 2-1. The parameter s indicates the number of internal stages.

Table 2-1: Butcher Tableau for s-stage Runge-Kutta method

$\begin{array}{c\|c} c & A \\ \hline & b^T \end{array}$	$\begin{array}{c\|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ \vdots & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array}$

Applying the Runge-Kutta scheme to (32) results:

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \Delta t \sum_{j=1}^{s} b_i \mathbf{f}(\mathbf{Y}_{\boldsymbol{i}} + c_i \Delta t)$$

$$\mathbf{Y}_{\boldsymbol{i}} = \mathbf{u}_n + \Delta t \sum_{j=1}^{s} a_{ij} \mathbf{f}(\mathbf{Y}_{\boldsymbol{j}} + c_j \Delta t) \tag{33}$$

The Explicit Runge-Kutta (ERK) method is given by $a_{ij} = 0$ for $i \leq j$, so that Y_i is obtained explicitly in terms of the solutions in the preceding stages, i.e. $Y_j \;\; j = 1, \dots, i-1$.

The conditions to achieve specific order of accuracy, so called order conditions, were provided by Butcher (1987), Lambert (1991) and Hairer et al. (1993) within the concept of rooted trees. Through some simplifying assumptions and based on rooted trees, Butcher (1963), Butcher (1964), Dekker and Verwer (1984), Hairer et al. (1993) and Hairer and Wanner (1996) developed formulae for order conditions. Here, order conditions just up to order four are provided.

for first order of accuracy,

$$\sum_{i=1}^{s} b_i = 1 \tag{35}$$

for second order of accuracy,

$$\sum_{i=1}^{s} b_i c_i = \frac{1}{2} \tag{36}$$

for third order of accuracy,

$$\sum_{i=1}^{s} b_i c_i{}^2 = \frac{1}{3}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i a_{ij} c_j = \frac{1}{6} \tag{37}$$

and for fourth order of accuracy.

$$\sum_{i=1}^{s} b_i c_i{}^3 = \frac{1}{4}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i c_i a_{ij} c_j = \frac{1}{8}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i a_{ij} c_j{}^2 = \frac{1}{12} \tag{38}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s}\sum_{l=1}^{s} b_i a_{ij} a_{jl} c_l = \frac{1}{24}$$

Stage order:

Many authors have referred to the stage order as a true order of accuracy when a method is applied to arbitrarily stiff problems (Alt, 1971; Alt and Ceschino, 1972; Kurdi, 1974; Alt, 1978; Hairer and Wanner, 1999; Cameron et al., 2002; Gómez et al., 2002 and Ketcheson et al., 2009). Therefore, low stage order may result in order reduction while applying to stiff ODEs. The associated conditions, with the maximum stage order equal to $\tilde{p}$, are as follows (Dekker and Verwer, 1984).

$$\sum_{j=1}^{s} b_j c_j^{k-1} = \frac{1}{k} \qquad (1 \le k \le \tilde{p})$$

$$\sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k \qquad (1 \le k \le \tilde{p}) \tag{39}$$

As it is clear, the first condition already exists in the formal order conditions. Kraaijevanger (1991) proved two practical results, which could easily be comprehended from the above formulae:

1. A Runge-Kutta scheme with nonnegative coefficients, $A \geq 0$, must have stage order $\tilde{p} \leq 2$. Moreover, if $\tilde{p} = 2$ then the first row of A must be zero.

2. An ERK scheme cannot obtain stage order more than one, i.e. $\tilde{p} = 1$. This is true whether or not the coefficents are nonnegative.

Consequently, it would be concluded that for DIRK schemes applied to practical cases such as gas dynamic applications, the stage order cannot exceed two for implicit schemes and the explicit schemes are first-stage order. Consequently, second-stage order as an essential property for implicit scheme provides explicit approach for the first stage. These schemes, so called Explicit DIRK (EDIRK) schemes, have become popular.

Stiffly accurate property:

Stiffly accurate property has been introduced as a favorable property for implicit scheme when dealing with stiff problems (Alexander, 1977; Hairer and Wanner, 1996; Kennedy and Carpenter, 2001; Pareschi and Russo, 2005; Pareschi and Russo, 2004; Boscarino, 2007 and Boscarino and Russo, 2009; Giraldo, 2013 and Weller et al., 2013). This property implies that the solution at the next step is the same as the solution in the last internal stage. Therefore, not only it provides the implicit approach for the last internal stage but also it offers time saving because the calculation in the final step in unnecessary. This property could be achieved by the following condition:

$$b_i = a_{si}\ ,\quad i = 1, \dots, s \tag{40}$$

Most implicit schemes are susceptible to order reduction in stiff limits (Bijl et al., 2002 and Burrage and Chan, 1993). Also, Boscarino (2007) showed that most IMEX schemes, in stiff limits, reduce to lower order of accuracy (Order Reduction) and he provided some relations for true order of accuracy in solving Differential Algebraic Equations (DAEs).

SSP property:

The Strong Stability Preserving (SSP) schemes enjoy the monotonic behavior of Forward Euler (FE) method, as they are convex combination of this method. The explicit types of SSP schemes were first developed, so called Total Variation Diminishing (TVD), by Shu and Osher (1988). Kraaijevanger (1991) introduced the radius of absolute monotonicity as the largest ratio of time step size of SSP schemes to FE method's one. He also showed that nonnegative coefficients of schemes, i.e. **A** and **b**, guarantee SSP properties. These methods were further developed by Gottlieb and Shu (1998). They showed that oscillations associated with discontinuity and shock problems happen even with high order linearly stable schemes without SSP properties. Ketcheson (2004) obtained the maximum radius of absolute monotonicity for implicit SSP Runge–Kutta methods up to two stages. Afterward, SSP optimized implicit Runge-Kutta schemes up to order six with eleven stages were developed by Ketcheson et al. (2009). As they indicated, all SSP Runge-Kutta schemes with order of accuracy more than one preserve the monotonic behavior for limited radius of absolute monotonicity. Since the optimized SSP schemes with order of accuracy more than two do not acquire high level of stability, extensive efforts have been done to develop new schemes combining the SSP property and high level of linear stability properties (Pareschi and Russo, 2005; Constantinescu and Sandu, 2007; Boscarino, 2007; Higueras et al., 2014). Higueras and Roldán (2006) extended the associated conditions for SSP property to IMEX scheme including the region of absolute monotonicity.

IMEX methods:

The time-dependent governing equations, such as compressible Navier-Stokes Equations (NSEs), involve different type terms. The stiffness may happen either by diffusion, heat flux or source terms in the boundary layers (Zhong, 1995). The fine-grid spacing sizes, employed for discretization, lead to stiff systems of ODEs. As a result, if the governing equations are treated by explicit schemes, the computations will become inefficient because the stable time-step sizes are much less than the accuracy considerations. On the other side, the implicit methods take more computer time and memory. Hence, attentions have been devoted to the IMEX methods, in order to circumvent the stability restriction on the explicit models (Ascher et al., 1997; Calvo et al. 2001; Sayfy and Al-Refai, 2002; Bujanda and Jorge, 2002; Kennedy and Carpenter, 2003; Yoh and Zhong, 2004; Zuijlen and Bijl, 2004; Frutos, 2006; Higueras, 2006; Higueras and Roldán, 2006; Liu and Zou, 2006; Pantano, 2007; Bujanda and Jorge, 2007; Boscarino and Russo, 2009; Giraldo et al., 2013 and Higueras et al., 2014).

Spatial discretization of the governing equations yields an ODE as:

$$\frac{\partial \mathbf{u}(\mathrm{t})}{\partial \mathrm{t}} = \mathbf{S}(\mathbf{u}, \mathrm{t}) + \mathbf{F}(\mathbf{u}, \mathrm{t}) \tag{41}$$

In which **S** corresponds to nonstiff and slow-wave terms and **F** represents the terms including fast-wave term and stiffness parameters. The general form of the s-stage IMEX DIRK scheme applied to (41) is:

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \Delta t \sum_{i=1}^{s} \tilde{b}_i \mathbf{S}(\mathbf{Y}_i + \tilde{c}_i \Delta t) + \Delta t \sum_{i=1}^{s} b_i \mathbf{F}(\mathbf{Y}_i + c_i \Delta t) \tag{42}$$

$$\mathbf{Y}_i = \mathbf{u}_n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} \mathbf{S}(\mathbf{Y}_j + \tilde{c}_j \Delta t) + \Delta t \sum_{j=1}^{i} a_{ij} \mathbf{F}(\mathbf{Y}_j + c_j \Delta t)$$

The coefficients of the IMEX DIRK scheme can be represented in the double Butcher tableau as follow:

Table 2-2: Double Butcher Tableau for s-stage IMEX DIRK scheme

$\begin{array}{c\|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}^T \end{array}$	$\begin{array}{c\|cccc} 0 & 0 & 0 & 0 & 0 \\ \tilde{c}_1 & \tilde{a}_{j1} & 0 & 0 & 0 \\ : & : & : & 0 & 0 \\ \tilde{c}_s & \tilde{a}_{s1} & \tilde{a}_{s2} & \tilde{a}_{sl} & 0 \\ \hline & \tilde{b}_1 & \tilde{b}_2 & \ldots & \tilde{b}_s \end{array}$	$\begin{array}{c\|c} c & A \\ \hline & b^T \end{array}$	$\begin{array}{c\|cccc} c_1 & a_{11} & 0 & 0 & 0 \\ : & a_{21} & a_{22} & 0 & 0 \\ : & : & : & \ddots & 0 \\ c_s & a_{s1} & a_{s2} & a_{sl} & a_{ss} \\ \hline & b_1 & b_2 & \ldots & b_s \end{array}$

As the explicit and implicit parts of the IMEX schemes have interrelations in achieving the specific order of accuracy, there are some coupling order conditions that need to be satisfied in addition to the classical order conditions (Eqs. 35-38). Here, the coupling order conditions required for IMEX Runge-Kutta schemes up to order of accuracy three is provided.

Coupling order conditions:

second order of accuracy,

$$\sum_{i=1}^{s} b_i \tilde{c}_i = \frac{1}{2}$$

$$\sum_{i=1}^{s} \tilde{b}_i c_i = \frac{1}{2} \tag{43}$$

third order of accuracy,

$$\sum_{i=1}^{s} \tilde{b}_i {c_i}^2 = \frac{1}{3}$$

$$\sum_{i=1}^{s} b_i {\tilde{c}_i}^2 = \frac{1}{3} \tag{44}$$

$$\sum_{i=1}^{s} \tilde{b}_i \tilde{c}_i c_i = \frac{1}{3} \tag{45}$$

$$\sum_{i=1}^{s} b_i \tilde{c}_i c_i = \frac{1}{3}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} \tilde{b}_i \tilde{a}_{ij} c_j = \frac{1}{6} \tag{46}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} \tilde{b}_i a_{ij} \tilde{c}_j = \frac{1}{6} \tag{47}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} \tilde{b}_i a_{ij} c_j = \frac{1}{6} \tag{48}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i \tilde{a}_{ij} c_j = \frac{1}{6} \tag{49}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i a_{ij} \tilde{c}_j = \frac{1}{6} \tag{50}$$

$$\sum_{i=1}^{s}\sum_{j=1}^{s} b_i \tilde{a}_{ij} \tilde{c}_j = \frac{1}{6} \tag{51}$$

IMEX methods preserve the linear invariants of ODE if weight coefficients b and $\tilde{b}$ are the same (Iserles and Zanna, 2000). Moreover, Constantinescu and Sandu (2007) proved that this condition is required for conservation laws as well.

2.4. Linear stability analysis theory

In the computational modelling, first, the governing equation set may be discretized in terms of spatial derivative terms, which leads to an Initial Value Problem (IVP) with the general form as,

$$\frac{\partial u}{\partial t} = f(u), \qquad u(0) = f_0 \tag{52}$$

in which, u represents any tracer and $f(u)$ denotes the spatially discretized terms. The analytical and numerical solutions represented by u_e and u_n respectively are expected to satisfy the original governing equation as

$$\frac{\partial u_n}{\partial t} - f(u_n) \approx 0 \tag{53}$$

$$\frac{\partial u_e}{\partial t} - f(u_e) = 0 \tag{54}$$

A new equation for the numerical error propagation, defined as $\varepsilon = u_e - u_n$, could be obtained by subtracting (53) and (54) as,

$$\frac{\partial \varepsilon}{\partial t} - \big(f(u_e) - f(u_n)\big) = 0 \tag{55}$$

The second term in (55) could be simplified by using the Taylor series expansion as,

$$f(u_n) = f(u_e) + \frac{(u_n - u_e)}{1!} \frac{\partial f(u)}{\partial u}\bigg|_{u_e} + \cdots \tag{56}$$

in which, for the linear stability analysis purpose, just the first two terms of the Taylor series approximation is considered. Using the definition of the numerical error leads to a new equation as,

$$\frac{\partial \varepsilon}{\partial t} - \frac{\partial f(u)}{\partial u}\bigg|_{u_e} \varepsilon = 0 \tag{57}$$

The linearized form of this equation allows studying the performance of any time integration scheme in the sense of the linear stability analysis,

$$\frac{\partial \varepsilon}{\partial t} - \lambda \varepsilon = 0 \qquad \varepsilon(0) = 0 \tag{58}$$

One can find the main reason that why the stability of any new time integration is mostly analyzed by applying to a linear IVP as (58), since it is the linearized representative of the equation that governs the numerical error propagation. Obviously, there is no significance for specific governing equations or specific coordinate systems in the above analysis. Therefore, it ensures the validation of the optimization procedure, which is done in this thesis, for most of governing equations in any coordinate system. It can be noticed that the general initial condition of the original IVP in (52) does not influence the error propagation, as long as the problem is well posed, particularly, the associated function of the initial condition, i.e. f_0, is continuously differentiable.

One may notice that the numerical error parameter, ε, contains both errors from spatial and temporal discretization methods, but, the focus of this thesis is only accuracy improvement corresponding to the temporal methods. In order to remove the error contributions from applied spatial schemes, throughout this thesis, whenever the error distribution is illustrated, the solution is subtracted from the reference solution. The reference solution is obtained by using the same time integration with identical spatial discretization method and the same spatial grid resolution, but, with finer time step sizes. However, it does not remove the spatial discretization error completely, since there is interaction between spatial and temporal errors.

2.5. Atmospheric field (Literature review)

In the boundary layer flows, the vertical transports of heat, moisture and momentum by turbulent eddies are usually modeled by using nonlinear diffusion terms, in which the diffusion coefficient depends on the variables (Girard and Delage, 1990). The main concern in the Atmospheric Boundary Layer (ABL) prediction models is the nonlinearity involved in the diffusion terms. The diffusion coefficients can be relatively large in comparison with the time step and vertical grid space. Moreover, they often exceed the numerical stability limits of the explicit schemes and it may lead to very small time steps for the explicit schemes. Therefore, time integrations in NWP and climate models need to be at least partially implicit otherwise short time steps have to be used which seems to be impractical (Diamantakis et al., 2005). Typical partially implicit schemes, used to calculate the boundary layer, treat only the dependent variables implicitly. The diffusion coefficients, which are functions of the dependent variables, are calculated explicitly.

The result is that unconditional stability can only be guaranteed for linear or weakly nonlinear problems (Kalnay and Kanamitsu, 1988 and Girard and Delage, 1990). Furthermore, Kalnay and Kanamitsu (1988) showed that a simple first-order scheme such as Backward Euler is unconditionally stable if the whole discrete equation is treated implicitly, which is impractical. Girard and Delage (1990) and later Bénard et al. (2000) proposed a more practical alternative. However, both techniques are more difficult to implement in the boundary layer using both local and non-local schemes for calculating the diffusion coefficient profiles. Typically, in such parametrizations, a local scheme is used for stable boundary layers in which the diffusion coefficients depend on the Richardson number. A non-local scheme is used for unstable boundary layers in which separate methods for calculating the diffusion coefficient profiles are employed, depending on whether the turbulence is driven by the surface or by the cloud top. Semi implicit predictor-corrector methods have been utilized by many researchers including Clancy and Pudykiewicz (2013) who applied this approach to the nonlinear system of the shallow water equations.

Several articles exist in the literature regarding explicit and implicit Runge-Kutta methods for solving dependent coefficient governing equations. Verwer (1996) reviewed the various types of the explicit Runge-Kutta schemes, including the method of Lebedev and Chebyshev, applied to the nonlinear problem focusing on the diffusion reaction flows. He found the explicit Runge-Kutta schemes as special due to their ability in constructing the maximal stability interval with minimal internal stages. Ixaru (2012) studied the three-stage explicit Runge-Kutta methods in the form of the exponential fitting approach. He found a scheme, in which the coefficients are no longer constant, as in the standard version, but depend on the equation. He showed that this property can help increase the order of accuracy to four, although, in general the order is three. Moreover, as he reported, the stability properties are enhanced with A-stability property, a fact which is quite unusual for explicit methods. Nazari et al. (2013) developed a semi-implicit Modified Extended Backward Differentiation Formulae (ME BDF) scheme for the solution of the nonlinear diffusion equations, with a special focus on the atmospheric boundary layer diffusion. They showed that this scheme obtains A- and B-stability properties, which make it appropriate for a wide range of nonlinearity. They examined the performance of this scheme by applying it to a commonly used nonlinear ordinary differential equation, proposed and Kalnay and Kanamitsu (1988) for atmospheric boundary layers. They obtained accurate solutions with extended range of stability. As they indicated as well, this scheme is non-sensitive to the time step size, while being easy and efficient to implement. Nazari et al. (2014) optimized an implicit three-stage third order Runge-Kutta scheme for the boundary layer vertical nonlinear diffusion for the atmospheric applications. They showed that the developed scheme improves the stability for larger range of time steps, which makes it suitable to resolve different diffusive systems with various turbulence closure modellings, including the $\mathrm{E}-\mathrm{L}$ turbulence closure scheme, which is commonly used in the case of interest. This scheme maintains A-stability, which makes it appropriate for the solution of the stiff problems.

In the rest of airflow region, far away from the earth surface as well as the boundary layer, the advection transports become more significant than diffusion ones. The advection of a passive tracer is governed by a simple "wave equation" (Durran, 2010). Therefore, most of the investigated problems in the atmospheric fields include fluid systems with low viscosity and weak dissipation. The equations governing these flows are often nonlinear, but their solutions almost never develop energetic shocks or discontinuities (Durran, 2010). Therefore, the numerical methods are not primarily concerned with the treatment of shocks. The atmosphere supports several different types of waves, propagating at different speeds. KAR (2006) proposed a semi-implicit approach using a three stage Runge-Kutta scheme, so called siRK3, to solve the two dimensional nonlinear shallow-water equations in a conservative flux form on uniform grid network with gird spacing $\Delta x = \Delta y = 100$km. The linear gravity wave terms were treated implicitly using a second-order trapezoidal scheme and the nonlinear terms including horizontal advection and other terms were temporally integrated explicitly using a third-order Runge–Kutta scheme. They compared the performance of their new scheme to the third-order explicit Runge–Kutta, so called expRK3 and showed that although the siRK3 scheme treats the gravity waves with second-order accuracy in time, it provides a stable, reasonably accurate and numerically efficient solution with relatively large time steps, compared to the third-order accuracy of the expRK3 scheme. However, the siRK3 scheme costs about 30% more in terms of CPU time, since the siRK3 scheme requires a two-dimensional Helmholtz-type elliptic equation to be solved at each iterative step of the scheme.

The maximum stable time step will be limited by the speed of the fastest-moving wave, which is sound wave, while it plays no direct role in transportation of the physical variables and hence does not need to be accurately simulated. Whitaker and Kar (2013) found that siRK3 proposed by Kar (2006) is unstable for purely oscillatory problems. They considered the IMEX Runge-Kutta time integration schemes for the fast-wave–slow-wave problems for which the fast wave has low amplitude and need not be accurately simulated. By proposing two alternative schemes, they showed that stability is improved if the averaging of the implicit part of the scheme is decentered, although the second-order accuracy is sacrificed. One of the alternatives utilizes a 3-cycle Lorenz scheme for the slow-wave terms and a trapezoidal scheme for the fast-wave terms. It is worth mentioning that the Lorenz N-cycle schemes can be classified as a special Runge-Kutta scheme, since they can be represented in the Butcher tableau. The Lorenz 3-cycle has an accuracy between the leapfrog and explicit fourth order Runge-Kutta scheme. The other is a combination of two proposed schemes in the past, which is stable for purely oscillatory problems for large range of frequencies when the slow-wave frequency is less critical. They demonstrated their achievements by solving a global spectral shallow-water model.

The implicit time differences require the solution of a nonlinear system at each time step. However, they are generally less efficient than implicit–explicit (IMEX) methods (also known as semi-implicit methods) in which only those terms responsible for linear sound wave propagation are evaluated using implicit schemes and the rest are treated explicitly (Durran and Blossey, 2012 and Ullrich and Jablonowski, 2012). The presence of different waves in the atmospheric field has been discussed in the literature (e.g., Ghosh & Constantinescu, 2016; Nazari & Nair, 2017). Due to these various fast and slow waves in the non-hydrostatic models, extensive attention has been paid to time-split schemes (e.g., Nazari & Nair, 2017; Ullrich & Jablonowski, 2012) and more specifically to IMEX Runge-Kutta schemes (Durran & Blossey, 2012; Giraldo et al., 2013; Weller et al., 2013). It has been proposed to solve fast waves implicitly, as they have less significant physical effects (except for the inertial gravity waves) (Durran & Blossey, 2012; Giraldo et al., 2013; Weller et al., 2013). Therefore, their restriction on stability is circumvented and consequently, the maximum stable time step increases.

Recently two different approaches have been distinguished in applying IMEX schemes to atmospheric applications called semi-implicit and Horizontally Explicit Vertically Implicit (HEVI) forms. The semi-implicit models treat acoustic waves implicitly and the advection is usually treated by semi-Lagrangian or explicit Eulerian methods. In the semi-implicit form, gravity waves are treated either explicitly or implicitly (Durran and Blossey, 2012 and Giraldo et al., 2013). In the HEVI models, the horizontal waves are evaluated explicitly while the terms responsible for vertical propagation are treated implicitly. Therefore, the maximum stable time step decreases in comparison with the semi-implicit methods. However, explicit time-stepping is simpler and as there is no horizontal communication per time step, then the HEVI forms are likely to be cheaper in parallel computing. Durran and Blossey (2012) analyzed IMEX multistep methods and applied their scheme to compressible Boussinesq equations using semi-implicit form treating buoyancy terms whether explicitly or implicitly. Their proposed schemes are based on Adams and backwards methods. Weller et al. (2013) compared many different types of IMEX Runge-Kutta methods applied to non-hydrostatic compressible Boussinesq equations with both semi-implicit and HEVI forms. They introduced new arrangement for HEVI form in order to increase the maximum stable time step. In the linear stability analysis, they found that, unfortunately, many of the employed schemes are unconditionally unstable in the explicit limit. Other schemes have regions of instability at very small time-steps. Weller et al. (2013) simulated incompressible flows by using a coarser grid and applied larger time steps. They indicated that non-stiffly accurate schemes perform better in simulating buoyancy while stiffly accurate schemes simulate incompressible flow more accurately and eventually it would be more appropriate for global atmospheric modeling. Lock et al. (2014) considered von Neumann stability analysis for two linear systems as well as for acoustic wave equations. They investigated the behavior of amplification factor and phase angle produced by different IMEX schemes within specific range of CFL numbers compared to analytical ones. They found that all examined schemes exhibit damping and retarded phase for large values of vertical wave numbers. Furthermore, they indicated that group velocity is reversed for some IMEX schemes but within ranges close to instability region. Giraldo et al. (2013) worked on multistage and multistep methods. They developed an IMEX Runge-Kutta scheme called ARK2(2,3,2). They looked for an IMEX scheme that allowed unified representation of various non-hydrostatic flow regimes for three-dimensional Euler equations. It was indicated by Weller et al (2013), examining different test cases, that ARK2(2,3,2) has better performance than other schemes with the same level of accuracy. The implicit part of the IMEX scheme proposed by Giraldo et al. (2013) satisfies "second-stage-order" condition as well as it is being stiffly accurate and has the same weight coefficients as explicit part.

Gardner et al. (2017) investigated various types of the IMEX schemes to find the efficient simulation of the non-hydrostatic atmospheric dynamics and to overcome the explicit stability constraints on time step size arising from acoustic waves. The IMEX formulations considered included HEVI and semi-implicit approach, in which some horizontal terms is treated implicitly. The accuracy and efficiency were evaluated on a gravity wave and baroclinic wave test case. They indicated that using a HEVI approach, in which some vertical terms are treated explicitly do not have advantages in improving the quality of the solutions or calculation cost over the common HEVI formulation. While treating some horizontal terms implicitly increases the maximum stable step size, but, the gains do not overcome the additional cost of solving a globally coupled system. Overall, they found that the third-order ARS3(3,4,3) and ARK3(3,2,4) methods performed the best, followed by the second-order ARS2(2,3,2) and ARK2(2,3,2) methods.

Parallel computation potentials

The advantage of using IMEX or semi-implicit approaches is to separate the stiff terms in order to remove the stability restrictions associated with these terms and to solve the system of equations more efficiently. However, these approaches still take large computational effort, which seems to be compensated by parallel computing implementations to achieve acceptable calculation costs. Contrary to direct solvers, e.g. Gaussian elimination, LU, Cholesky, and QR decompositions, iterative methods, e.g. Guass-Seidel (GS), Jacobi and Successive Over Relaxation (SOR), approach the solution gradually. Therefore, they are likely to be more applicable in parallel computations.

The parallel calculation of the Runge-Kutta schemes on Personal Computer (PC) cluster has been studied in the literature (e.g. Bendtsen, 1997; Muir et al., 2003 and Mantas et al., 2005). The main difficulty in applying parallelization is to solve a linear system of equations, but, efficient speed-up was observed in using these methods. As indicated by Giraldo et al. (2013) as well, in practical cases like the NWP applications, in which the time step is restricted by the vertical resolution, using the IMEX scheme with the HEVI approach is efficient such that the horizontal direction is solved explicitly but the vertical direction is treated implicitly. This strategy results in the solution of a collection of small banded matrices that are stored on each processor and are decoupled from each other. This solution strategy requires using a two dimensional (2D) domain decomposition, where the vertical direction is on one processor, resulting in a parallel solution strategy. Abdi et al. (2017) discussed the acceleration of an IMEX scheme, including the ARK2(2,3,2), implementation for the non-hydrostatic atmospheric model on multi-core processors such as GPU architecture. In this effort, by using the HEVI approach, the fast acoustic and gravity waves were treated implicitly while the Rossby waves were calculated explicitly. Therefore, the maximum stable time step is governed by slow Rossby waves. Here it is worth mentioning that the only differences between the new IMEX scheme proposed in this thesis, i.e. IMEX-SSP2(2,3,2), and the ARK2(,2,3,2), proposed by Giraldo et al (2013), are i) IMEX-SSP2(2,3,2) enjoys different places for the internal stages served for explicit and implicit parts ii) the eigenvalues of the A matrix in the IMEX-SSP2(2,3,2) are different. However, these two differences impose no restrictions on the IMEX-SSP2(2,3,2) implementation in parallel computing since in the HEVI approach, the explicit and implicit components are decoupled from each other (Girarldo et al, 2013).

2.6. Research gaps and contributions of the present study

In this section, it is attempted to describe the gap remains from the literature, which motivates the author to tackle this study (section 2.6.1) and the contributions of this thesis will be explained in Section 2.6.2.

2.6.1. Research gaps

Extensive researches have been conducted in the past, in which different time integrations were studied in the field of the atmospheric applications. This includes multi-step methods, e.g. Durran and Blossey (2011), as well as the Runge-Kutta methods, known as the multi-stage schemes (Girarldo et al., 2013 and Weller et al., 2013). Predominantly, the semi-implicit and IMEX approaches give the opportunity to remove the stability restrictions, associated with the stiff terms so that the maximum CFL number is allowed to be increased. Using the Runge-Kutta schemes, different stability properties, e.g. I-, A-, B-, L-stability and SSP properties, have been utilized in different studies (Nazari et al., 2014; Girarldo et al., 2013; Weller et al, 2013, Lock et al., 2014). However, there were no significant studies on comparing these properties, especially, in terms of their applicability in the physical practice. Thus, they were not compared to observe their influential effects on improving the stability and accuracy. Furthermore, the non-negative property corresponding to the numerical stability was studied in a few researches, e.g. for parabolic equations by Higueras et al. (2014). However, this property has significant contributions in non-oscillatory behaviors, as indicated by Strikwerda (2004), apart from the type of the governing equations. Moreover, since the dissipation and dispersion errors are the main concerns in improving the quality of the computations, they deserve further elaborative research, in order to find how they can be controlled and diminished. Through reviewing the past research efforts, it was noticed that these errors, produced by the IMEX and semi-implicit Runge-Kutta schemes applied to the two frequencies system, were not studied yet. This was also true for the boundary layer flows, in which the main difficulty is how to treat the nonlinear diffusion coefficients by using the semi-implicit schemes. Therefore, it was targeted to study the optimization of dissipation and dispersion errors by using the semi-implicit and IMEX Runge-Kutta schemes for the field of interest. Moreover, different stability properties were examined by applying to different types of the PDEs in which the incorporated terms, i.e. advection, diffusion or source terms, were intensified significantly to observe the impact of these stability properties in the calculations.

2.6.2. Contributions of the present study

This thesis will be accomplished in four parts. In the first one, Chapter (3), the A-stability and SSP optimized properties will be further investigated to mathematically prove that both properties seek to optimize the dissipation and dispersion errors, but for different purposes. This effort will be done through utilizing one well-known integral, which combines the dissipation and dispersion errors. Furthermore, it will be shown that the combination of both A-stability and SSP optimized properties is essential for the sake of both stability and accuracy. Moreover, it will be shown that the relative stability concept, in which the numerical stability function is compared to analytical stability function, is not an appropriate approach to minimize the dissipation and dispersion errors. It will be shown that the relative stability concept just reveals the stability within the range of the order of accuracy. In this regard, it will be proved that there is no need to consider the viscous terms in the optimization process, since these terms have descending contributions to the stability function. Physically, it means that the inviscid (advection dominated) flows is the critical case for stability and accuracy investigation. This fact leads to the use of the absolute stability concept, in which the numerical stability function will be attempted to be upper-bounded. The result of this part is published in Journal of Communication in Computational Physics.

As discussed before, the dissipation and dispersion errors are the main obstacles toward stable calculations and accurate solutions. Besides, to the author's knowledge, the optimization of these errors, by using the IMEX Runge-Kutta schemes was never studied before. Therefore, the focus in the second part, Chapter (4), is on optimizing the dissipation and dispersion errors using the IMEX Runge-Kutta schemes. In this part, the performance of the optimized IMEX scheme will be examined through calculating the compressible Boussinesq equation. The compressible Boussinesq equation is popular in the atmospheric field, since it more realistically approximates the behavior of the air flow than others. The optimization of dissipation and dispersion errors will be done through using an integral, proved as a combination of these errors. This integral will be defined relevant to the discussion, such that it includes the effects of the ratio of fast- to slow-wave frequencies. Since in most cases of interest the central second order scheme discretizes the spatial terms, the higher order of accuracy for temporal scheme does not promise further advantages. Hence, the optimization will be done by using three-stage second order IMEX scheme, in which for the explicit part the dispersion error is eliminated. The implicit part is a DIRK scheme, that integrates the stiff terms in two internal stages and it deserves the second order of accuracy and second-stage order as well. The last property is necessary in order to prevent the order reduction through calculating the stiff limit cases. Most of the previous implicit Runge-Kutta schemes were developed based on preserving the L-stability property. In this part, it will be shown that the L-stability property unnecessarily weakens the scheme's performances, since the stability function does not need to obtain zero value in calculating the extremely large variables. Calculation of the test case, i.e. compressible Boussinesq equations, shows that the associated conditions with L-stability property result in poor behavior, in the sense of accuracy and stability concerns. In addition, the A- and L-stability properties will be further compared through calculation of a well-known ODE test case, i.e. van der Pol equations, in order to show that the L-stability property may decline to a lower order of accuracy in stiff limits. In other words the order reduction may happen by using those schemes with L-stability property rather than A-stability property. The other distinctive property of the proposed IMEX scheme, called IMEX-SSP2(2,3,2), is that the explicit and implicit parts calculate the source terms at different internal stages. It will be shown that this difference will not cause any phase errors, since the weight coefficients for both implicit and explicit schemes are nonnegative. This is due to preserving the SSP property for both explicit and implicit schemes. The new optimized scheme will be compared to a well-studied scheme in the literature, called ARK2(2,3,2), through solving the compressible Boussinesq equations on the actual length scale sizes. Thus, it will reveal the abilities of the new scheme in terms of stability and accuracy improvements. The A-stability property, compared to L-stability, could help the temporal scheme to tolerate larger stable time step sizes, while it retains the true order of accuracy. It will be shown that the proposed IMEX scheme can extend the range of stable time-step sizes and it improves the accuracy as well. This paper was published in Journal of Advances in Modeling Earth Systems.

Provided in Chapter 5 is an unpublished technical note, in which the performance of the IMEX-SSP2(2,3,2) is further examined in solving the inviscid Burger equation in the spherical coordinate system. The main purpose is to show that this scheme still can be helpful in actual problem. Therefore, the North America continent is selected as the computational domain and a realistic dataset is fed as the initial condition. As will be shown in Chapter 5, the new IMEX scheme can effectively improve the accuracy for this test case as well.

The non-negativity of stability function and its influential effects were discussed by Higueras et al. (2014), through calculating the parabolic governing equations. It was indicated that maximum nonnegativity may help temporal scheme in order to prevent the spurious oscillation, associated with dissipativity behavior of spatial schemes. The governing equations, in atmospheric fields, are mainly derived from compressible Euler equations, which is a well-known hyperbolic equation set. Meanwhile, the presence of various diffusion and source terms makes the governing equations more complicated. Therefore, in the third part, Chapter (6), the optimization of nonnegativity of stability function of the implicit Runge-Kutta scheme will be studied. The relation between non-negativity and non-oscillatory behavior will be explained. This part also includes further investigation of the A-stability property for the implicit Runge-Kutta schemes in the atmospheric field and especially in the boundary layer, in which the diffusion terms become significant (Stiff terms). Therefore, the ability of the DIRK schemes, preserving the A-stability property will be further compared to those obtaining the L-stability property in terms of non-negativity and non-oscillatory behaviors. Thus, the nonnegative region for the numerical stability function will be optimized by using three-stage third order DIRK schemes, which separately deserve two important properties, i.e. A-and L-stability properties, combined with SSP property. The output of this optimization is two different DIRK schemes with two different properties, i) A-stability ii) L-stability, which are combined with the SSP property. These schemes will be compared by calculating the simple prototype problems, which are the fundamental equations for deriving the atmospheric governing equations. As shown, the L-stability property shows poor behavior in terms of calculating different test cases, including compressible Euler and nonlinear viscous Burger equations. Moreover, it will be shown that the SSP optimized property is not sufficient for the field of interest, but the combination of A-stability and SSP property will be more helpful. This paper is currently under review.

In the fourth part, (Chapter 7), the boundary layer flow is studied in terms of the diffusion terms, which impose stability restrictions, due to the rapid changes of the gradient terms. The diffusion terms, are approximated by a nonlinear model due to variable diffusion coefficient. This nonlinearity makes the explicit schemes too expensive and the implicit scheme is impractical, due to expensive iterative approach. By using the DIRK scheme, a new approach was proposed, which benefits from both explicit and implicit properties of the time integration. It is shown that this semi-implicit approach improves the accuracy of the calculation due to involving the explicit property as well as extending the range of stability, which is provided by the implicit property.

2.7. Suggested future works

The explicit Runge-Kutta schemes obtain limited range of stability; therefore they impractically restrict the stability range for calculating less physically significant fast wave term and for stiff terms as well. At the same time these schemes are computationally efficient and they preserve the accuracy more than implicit scheme in the same class, because, in terms of the Runge-Kutta schemes, the stability function incorporates some definite terms, which let the ERK schemes eliminate the dispersion error. The new class of Runge-Kutta scheme, so called Runge-Kutta-Chebyshev (RKC) scheme, have been proposed recently, which enjoy larger stability range, compared to traditional Runge-Kutta methods. In general, the stability function of an s-stage explicit Runge-Kutta method is a polynomial of degree s, in which s represents the number of internal stages, i.e. $R(z) = \sum_{j=1}^{s} a_j z^j$. The idea of the RKC schemes is to use the Chebyshev basis instead of monomial basis. Hence, the approximation of larger variables becomes more accurate (Ketcheson and ahmadia, 2012). Therefore, the explicit RKC scheme can be utilized to compensate the stability restriction associated with traditional explicit scheme. This scheme can be very useful in serving to calculate in HEVI arrangement, since in this arrangement the stable time step is more limited than semi implicit approach, due to explicit calculation of the horizontal component of the acoustic waves. Moreover, the explicit RKC could be very helpful in combination with implicit Runge-Kutta scheme to improve the accuracy in semi-implicit approach as well. In this regard, the implicit Runge-Kutta scheme would be applied to fast acoustic wave terms and explicit RKC schemes served to calculate the advection and gravity terms. The explicit RKC can also help to calculate the nonlinear diffusion terms in the boundary layer.

As shown in this thesis, the combination of A-stability and SSP properties is helpful in both boundary layer and free stream zones. It was shown that in shock and discontinuity problems in viscous flows, the A-stability property leads to a non-oscillatory behavior for larger range of CFL numbers, while the optimized radius of absolute monotonicity, i.e. SSP optimized schemes, exhibits the oscillations. Therefore, it requires further investigation of A-stability property, more specifically, in the sense of entropy conservation. Tadmor (2004) proposed this approach to solve the hyperbolic equations. Entropy conservation brings an extra condition, related to the second thermodynamic law, imposing the numerical schemes to satisfy the positive entropy generation, and as shown in the literature, this property can help to preserve the non-oscillatory behavior, particularly, in shock and discontinuity problems.

Generally, the practical problems often reduced to IVPs, which is a system with very large ODEs. The numerical integration is computationally very demanding. Therefore, the system often does not allow using the explicit schemes, because these methods require very small time step sizes to ensure numerical stability. Instead, the problems are usually solved using implicit methods. However, these methods require solutions of an implicit system at each integration step. Recently, exponential integrators have been proposed as a potential alternative for efficiently solving large problems. The methods draw attention when the use of Krylov projection techniques allowed these matrix exponential terms to be evaluated efficiently (Loffeld and Tokman, 2013). Afterward, many exponential integrators for general stiff systems have been proposed. It is very interesting to employ the exponential integrator schemes for the case of interest. Therefore, all the efforts done in this thesis could be applied to the exponential integrators and it may help to further improve the stability and accuracy.

2.8. Summary and conclusion

This thesis investigates the performance of Runge-Kutta schemes for temporally integrating the atmospheric applications. The objective is to find the more efficient temporal schemes in order to calculate the governing equations reliably in the sense of large range of stable time steps and accurate solutions. The main obstacle in front has been recognized as the dissipation and dispersion errors, which limit the stability and accuracy. Therefore, the optimization of these inevitable errors has been studied in order to develop new schemes, which help extend the range of stability, while preserving the desired accuracy tolerances. In this regard, some well-studied stability properties are investigated, in order to observe how they affect the schemes' performances in terms of improving the stability and accuracy.

In Chapter 3, it is aimed at proving that the relative stability function, which is the ratio of absolute numerical stability function to the analytical one, is not informative for optimization of dissipation and dispersion errors. It is just useful to study the relation between stability and order of accuracy. In the literature, the optimization of dissipation and dispersion errors were studied by Hu et al. (1996) using an integral function as a combination of dissipation and dispersion errors. It was shown that the proposed integral by Hu et al. (1996) is sufficient for optimization purposes and it does not require inclusion of the viscous terms. Furthermore, it was shown that the proposed integral by Hu et al. (1996) incorporates the conditions associated with the A-stability and the SSP optimized properties. It was shown that the A-stability property tries to extend the limit of stability, while SSP optimized property tries to increase the limit of accuracy by decreasing the Truncation Errors (TEs) that its leading term is commonly called the error constant. Calculation of the nonlinear Burger viscous equation showed that the A-stability property is necessary for the DIRK schemes to tolerate large time step sizes and large diffusion terms; therefore, the A-stability property can accelerate achieving the steady state solutions.

Chapter 4 explains how an optimized stable and accurate IMEX Runge-Kutta scheme was developed for the atmospheric applications. The new IMEX scheme, so called IMEX-SSP2(2,3,2), was constructed with three-stage and second-order explicit and implicit components. The three-stage explicit part enjoys zero error constant, which may result in less phase error. Moreover, through von Neumann analysis, it was found that its stability function that intersects the imaginary axis at a point not much smaller than the maximum achievable value. In addition, this explicit scheme extends the stability region along the negative real axis farther than the explicit scheme which has the maximum value along the imaginary axis. The implicit part preserves the A-stability property and it satisfies stiffly accurate and second-stage order condition. These characteristics were found to be fundamental in stiff limits. Both explicit and implicit parts of the new IMEX scheme hold the SSP property with the region of absolute monotonicity much larger than a recent scheme in the literature, ARK2(2,3,2), which was shown that has the more appropriate performances in applying to the atmospheric applications (Weller et al., 2013). It was shown that the IMEX-SSP2(2,3,2) extends the range of stable time step sizes as well as it improves the accuracy, the fact that was proved by examining different test cases. The stability analysis was performed for acoustic wave equations as well, in which the new IMEX scheme did not improve the stability region significantly compared to ARK2(2,3,2). Two-frequency system was analyzed to compare the calculated phase distribution to the exact phase. The new IMEX scheme could improve the phase distribution, particularly for large values of wavenumber and for large time steps. This new IMEX scheme was employed to simulate the compressible Boussinesq equations in different arrangements, including the semi implicit and HEVI forms. It was demonstrated that IMEX-SSP2(2,3,2) could enhance the maximum stable time step for the semi-implicit in both forms of integration in comparison with ARK2(2,3,2). This improvement also includes the accuracy; the difference between buoyancy and divergence free, i.e., $\nabla . \mathrm{u} = 0$, fields calculated by both IMEX schemes and by reference solution indicated that the new scheme, compared to ARK2(2,3,2), decreased the level of numerical error. In order to show that the IMEX-SSP2(2,3,2) was not designed for the specific governing equation, the inviscid Burger equation was solved in the spherical coordinate system using a realistic data set of the North American continent as the initial condition. Chapter 5 presents the results for this test case, which reveals that the new scheme can still help improve the stability and accuracy for this problem as well.

Chapter 6 includes an article, in which the importance of the non-negativity property associated with the stability function is investigated. The three-stage third order DIRK scheme was undertaken, by which the non-negative region of the numerical stability function was optimized.

Eventually, two new DIRK schemes were developed such that both have the SSP property, but, the first one preserves the A-stability property and the other obtains L-stability property. The objective function for optimization is the maximum values of the roots of the numerical stability function, which produce optimum nonnegative region for the stability function. These schemes were compared to the three-stage third order SSP optimized scheme, called SSP(3,3). It was found that the SSP optimized property provides larger range of non-negativity for the imaginary component of the numerical stability function and this is the reason for the non-oscillatory behavior of these type schemes in calculation of inviscid (advection dominated) flows. It was found that the L-stability property can help enlarge the nonnegative region for the real component of the stability function, the property which makes the L-stability property more appropriate for the parabolic (diffusion dominated) equations. Finally, the A-stability property was found as a property which provides non-negativity for both real and imaginary components of the stability function. Therefore, the A-stability seems to be more beneficial in most practical applications, which are mainly viscous (advection diffusion) flows. These hypotheses were evaluated by examining the new schemes in solving various equations. Three different well-known test cases, i.e. the compressible Euler equations, nonlinear viscous Burger equation and linear advection with nonlinear source term were calculated. It was found that the A-stability property has the optimum performance, since this property provides large range of non-negativity for both real and imaginary components of the stability function. Finally, it is concluded that the essential property for the atmospheric field is the A-stability property combined with the SSP property.

Chapter 7 describes an effort, which has been done toward proposing a new approach for temporally integrating the boundary layer flows and to overcome the difficulties of nonlinearity associated with modelling of the diffusion process. In this new approach, which is a semi-implicit method, a DIRK scheme is employed, by which at each internal stage the nonlinear diffusion coefficient is calculated using a weighted average of the solutions in the previous time step and the current internal stage. This approach mixes the explicit and implicit properties so that the explicitness leads to improve the accuracy and implicitness helps to extend the range of stability. In the linear stability analysis, the optimum values of the weighted average coefficients were calculated subject to the minimum values of the numerical errors. One DIRK scheme, SSP(2,2), which has the SSP optimized property and preserve the second order of accuracy within two internal stages, was undertaken to temporally integrate the vertical diffusion boundary layer equation with different techniques, including implicit, so called FI-RK and common semi-implicit approach, called SE-RK as well as the new weighted-average method (weighted-SI-RK). Note that, since the tracers' properties are dominantly transferred by diffusion process within the thin boundary layer, the dispersion errors seem not to be so critical in this zone (Diamantakis et al., 2006). By switching to the Fourier space, the stability function just obtains the real variable, indicating the importance of dissipation error rather than dispersion error. The linear stability analysis showed that the SE-RK approach is less damping than other schemes; hence, it calculates the result more accurately, the fact that the numerical computation of the representative nonlinear equation, proposed by Kalnay and Kanamitsu (1988), proved. However, this test case showed that the SE-RK is less stable. The FI-RK scheme is impractical in real cases, since it is too expensive due to more CPU time usage. The proposed weighted-SI-RK, as shown, is able to improve the range of stable time step sizes, compared to the SE-RK, and to increase the accuracy compared to FI-RK and the proposed semi-implicit method by Diamantakis et al. (2006) (weighted-SI-PC).

Chapter 3

On dissipation and dispersion errors optimization, A-stability and SSP Properties[1]

Abstract:

In a recent paper (Du and Ekaterinaris, 2016) optimization of dissipation and dispersion errors was investigated. A Diagonally Implicit Runge-Kutta (DIRK) scheme was developed by using the relative stability concept, i.e. the ratio of absolute numerical stability function to analytical one. They indicated that their new scheme has many similarities to one of the optimized Strong Stability Preserving (SSP) schemes. They concluded that, for steady state simulations, time integration schemes should have high dissipation and low dispersion. In this note, dissipation and dispersion errors for DIRK schemes are studied further. It is shown that relative stability is not an appropriate criterion for numerical stability analyses. Moreover, within absolute stability analysis, it is shown that there are two important concerns, accuracy and stability limits. It is proved that both A-stability and SSP properties aim at minimizing the dissipation and dispersion errors. While A-stability property attempts to increase the stability limit for large time step sizes and by bounding the error propagations via minimizing the numerical dispersion relation, SSP optimized method aims at increasing the accuracy limits by minimizing the difference between analytical and numerical dispersion relations. Hence, it can be concluded that A-stability property is necessary for calculations under large time-step sizes and, more specifically, for calculation of high diffusion terms. Furthermore, it is shown that the oscillatory behavior, reported by Du and Ekaterinaris (2016), is due to Newton method and the tolerances they set and it is not related to the employed temporal schemes.

Key words: Diagonally Implicit Runge-Kutta methods, dissipation and dispersion, optimization, numerical stability, steady state.

---

[1]This chapter has been published as Arman Rokhzadi, Abdolmajid Mohammadian, Vol. 24, No. 1, pp. 268-285 July 2018, **Journal of Communication in Computational Physics**, 2018, doi: 10.4208/cicp.OA-2017-0015.

3.1. Introduction

Steady state solutions could be sought as the long-time mean solutions of the unsteady problems (Du and Ekaterinaris, 2016). However, dissipation and dispersion errors are the main obstacles to achieve long time-step sizes and accordingly to decrease the calculation time.

Among different temporal integrator schemes, Runge-Kutta methods have attracted attentions, as they are single-step methods and have free parameters which could lead to optimization of dissipation and dispersion errors. Much research in the literature has been made to control and bound dissipation and dispersion errors to increase the range of stability and accuracy. Consequently, numerous stability properties have been introduced including A-stability property. The reader is referred to Ordinary Differential Equations (ODEs) books for more details (e.g. Hairer and Wanner, 1996).

The Strong Stability Preserving (SSP) Runge-Kutta methods are well-known due to their non-oscillatory behavior in shock and discontinuity problems. These methods were designed as convex combinations, weighted average correlations with nonnegative weighted coefficients less than one, of Forward Euler (FE) method within limited radius of absolute monotonicity. The radius of absolute monotonicity is the ratio of the maximum stable time step tolerated by the Runge-Kutta scheme and the FE scheme. This class of methods was further developed by Gottlieb and Shu (1998). Then, Ketcheson (2009) developed optimal implicit SSP Runge–Kutta methods up to order six with eleven stages.

The objective of this paper is to further investigate the stability analysis within studying dissipation and dispersion errors, in order to discuss the conclusions of Du and Ekaterinaris (2016) and to discuss their proposed DIRK scheme, which was obtained using the relative stability function. Du and Ekaterinaris (2016) indicated that their developed scheme has many similarities with the three-stage fourth order SSP optimized DIRK scheme (Ketcheson et al., 2009). They also described their proposed scheme, so called DIRK-D, as a more accurate model for low wavenumber components than other schemes they employed. However, as will be discussed, in stability analyses of temporal schemes, the main attention is on time step sizes. The wavenumber is assumed as a fixed variable, which basically would be the highest one. It will be shown that the source of instability, imposed by grid mesh, is due to high wavenumbers.

Relative stability analysis, i.e. the ratio of absolute numerical stability function to analytical one, was introduced by Hairer and Wanner (1996) within the concept of “Order Star”. Du and Ekaterinaris (2016) used relative stability function to design the optimized three-stage fourth order DIRK scheme, DIRK-D, and to examine their performances. However, the Order Star is mainly useful in proving relation between stability and achievable order of accuracy and this idea is not useful for judging the stability. Indeed, absolute stability is the more practical one (Leveque, 2007).

In Section 3.2, the relative stability analysis is studied further in order to show that this concept is not useful for optimization of the dissipation and dispersion errors. Meanwhile, as indicated by Leveque (2007) and shown in the present paper, this concept just shows the order of accuracy and truncation error. Du and Ekaterinaris (2016) indicated that, for advection-diffusion system, the amplification factor needs to include contribution of physical and numerical diffusion. In contrast, it will be shown that this contribution is not required in minimizing the dissipation and dispersion errors (Section 3.3). Moreover this section proves that both A-stability and SSP properties cope with minimizing the dissipation and dispersion errors but for different applications. In Section 4, the non-linear viscous Burgers equation is examined with high order WENO-5 spatial scheme, which is very similar to WENO-5M used by Du and Ekaterinaris (2016). They used Newton iteration method, while, in this paper, the Gauss Seidel approach is used. It will be shown that the oscillatory behavior reported by Du and Ekaterinaris (2016) is not caused by temporal schemes and it is due to Newton method and the tolerances they set. Section 6 presents some concluding remarks.

3.2 Relative stability function

The advection-diffusion model is described by:

$$\frac{\partial \mathrm{u}}{\partial \mathrm{t}} + \mathrm{c}\frac{\partial \mathrm{u}}{\partial \mathrm{x}} = \nu \frac{\partial^2 \mathrm{u}}{\partial \mathrm{x}^2} \tag{3-1}$$

in which c and ν represents the advection speed and viscosity respectively.

physical coefficients Applying the Fourier transform and assuming the solutions as $\mathrm{u}(\mathrm{x},\mathrm{t}) = \hat{\mathrm{u}}(\mathrm{t})\mathrm{e}^{\mathrm{ikx}}$, this equation is simplified as:

$$\frac{\partial \hat{u}}{\partial t} = -(ick + k^2\nu)\hat{u}(t) \tag{3-2}$$

It could be found that the analytical solution is an exponential function. Consequently the analytical stability function, R_a, defined as the ratio of two successive time steps solutions, forms as:

$$\frac{\hat{u}^{n+1}}{\hat{u}^n} = R_a = e^z \tag{3-3}$$

in which,

$$z = -(ick + k^2\nu)\Delta t \tag{3-4}$$

As it is clear, $z \in \mathbb{C}^-$, where $\mathbb{C}^-$ includes all the complex variables with negative real part. Note that c represents the advection speed which is assumed to obtain negative value as well.
Applying the Fourier transform to time dependent function as $\hat{u}(t_n) = e^{-i\omega t_n}$, results in the analytical dispersion relation:

$$e^{-i\omega\Delta t} = e^z \tag{3-5}$$

in which ω is analytical frequency. Therefore, Eq. 3-5 is rearranged as:

$$e^{-i\omega\Delta t} = e^{-k^2\nu\Delta t}e^{-ick\Delta t} \tag{3-6}$$

As can be seen, the diffusion contribution to the equation is in the form of descending exponential function with maximum value equals one. This fact will be discussed in the optimization process in Section 3.3.2.

In order to analyze the performance of Runge-Kutta scheme, this time integration method would be applied to (3-2) using the following equations:

$$u_{n+1} = u_n + z\mathbf{b}^T\mathbf{Y} \tag{3-7}$$

$$\mathbf{Y} = u_n\mathbf{e} + z\mathbf{AY}$$

u_n represents the numerical solution at time step t_n, while **Y** represents the vector of solutions in internal stages. **e** is vector of ones of size $s \times 1$, and s is the number of internal stages. For DIRK scheme, matrix **A** and vector **b** are shown in Butcher tableau as Table 3-1.

Table 3-1: Butcher Tableau for s stage DIRK scheme

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b^T} \end{array} \qquad \begin{array}{c|cccc} c_1 & a_{11} & 0 & 0 & 0 \\ : & a_{21} & a_{22} & 0 & 0 \\ : & : & : & a_{33} & 0 \\ c_s & a_{s1} & a_{s2} & a_{sl} & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array}$$

Similar to the definition of the analytical stability function, the absolute numerical stability function, which is the ratio of the numerical solutions of two successive time steps, is in the form:

$$\frac{u^{n+1}}{u^n} = R_n = 1 + z\mathbf{b}^T(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{e} \tag{3-8}$$

in which **I** is $s \times s$ unit matrix.

It should be noted that the DIRK's coefficients which are specified by matrix **A** and vectors **b** and **c** need to satisfy the order conditions. The reader is referred to the literature for more details on order conditions, e.g. Nazari et al. (2015).

Du and Ekaterinaris (2016) described the relative stability function, called E, as the ratio of absolute numerical stability function to analytical one:

$$E = \frac{R_n}{R_a} = \frac{1 + z\mathbf{b}^T(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{e}}{e^z} \tag{3-9}$$

Taylor series expansion of the numerator reforms this equation as:

$$E = \frac{1 + z\mathbf{b}^{T}\mathbf{e} + z^{2}\mathbf{b}^{T}\mathbf{A}\mathbf{e} + z^{3}\mathbf{b}^{T}\mathbf{A}^{2}\mathbf{e} + \cdots + z^{p}\mathbf{b}^{T}\mathbf{A}^{p-1}\mathbf{e} + \sum_{n=p+1}^{\infty} z^{n}\mathbf{b}^{T}\mathbf{A}^{n-1}\mathbf{e}}{e^{z}} \tag{3-10}$$

This equation could be rearranged as:

$$E = \frac{1 + z\mathbf{b}^{T}\mathbf{e} + z^{2}\mathbf{b}^{T}\mathbf{A}\mathbf{e} + z^{3}\mathbf{b}^{T}\mathbf{A}^{2}\mathbf{e} + \cdots + z^{p}\mathbf{b}^{T}\mathbf{A}^{p-1}\mathbf{e}}{e^{z}} + \frac{\sum_{n=p+1}^{\infty} z^{n}\mathbf{b}^{T}\mathbf{A}^{n-1}\mathbf{e}}{e^{z}} \tag{3-11-a}$$

Obviously, in the first term of Right Hand Side (RHS), the polynomial's coefficients in numerator are part of the order conditions. Depending on the desired order of accuracy, p, imposed on DIRK scheme, this polynomial is an approximation of the Taylor series of exponential function up to order p. Hence, (3-11-a) could be rearranged as follows:

$$E = 1 - \frac{\sum_{n=p+1}^{\infty} \frac{z^{n}}{n!}}{e^{z}} + \frac{\sum_{n=p+1}^{\infty} z^{n}\mathbf{b}^{T}\mathbf{A}^{n-1}\mathbf{e}}{e^{z}} \tag{3-11-b}$$

Comparing the equations (3-11-a) & (3-11-b) implies that the Runge-Kutta methods try to revive the exponential function depending on the order of accuracy, p.

In stability analysis context, it is tried to design new schemes bearing large time step sizes, which is included in the variable z, (3-4). Note that decay rate of exponential function, denominator in RHS of Eq. A-11, is faster than any power law function. Hence, it is clear that the magnitude of relative stability function, i.e. $|E|$, diverges once the variable z, obtains extremely large negative values in the domain of interest, $z \in \mathbb{C}^{-}$. Moreover, to minimize the difference between the relative stability function and one, it is required to minimize the summation in numerator of the last term in RHS, (3-11), which is the Truncation Error (TE) term of the absolute numerical stability function, comparing (3-8) and numerator of (3-10). Therefore, it would be appropriate to consider the absolute numerical stability function in order to optimize the performance of Runge-Kutta schemes.

Leveque (2007) indicated that the relative stability function is just useful in studying the relation between stability and accuracy. In order to show this relation, Figure A-1 illustrates the relative stability region, i.e. $|E| \leq 1$, with blue color for some optimized SSP schemes. The coefficients of the examined SSP schemes are provided in tables 3-2 to 3-5 in Butcher tableau form. The first row compares the stable region of the optimized SSP Runge-Kutta schemes with second order of accuracy achieved in two and three stages, so called SSP(2,2) and SSP(3,2) respectively. It is clear that more internal stages, involved in the simulation, results in larger stable region. However, it is worth mentioning that these two schemes preserve the A-stability property, which means that their absolute numerical stability function remain less than one in the whole domain of interest, i.e. $z \in \mathbb{C}^-$. However, the relative stability function does not show this property.

To compare the relative stability region for third and fourth order of accuracy, SSP optimized Runge-Kutta schemes with three stages, so called SSP(3,3) and SSP(3,4) respectively, are illustrated in the second row of Figure 3-1. The blue fingers include the stability functions' roots (Hairer and Wanner, 1996). As can be seen, the higher order of accuracy results in more fingers and different region of stability. Further details of Order Star could be found in the literature (Hairer and Wanner, 1996; Wanner et al., 1978; Nørsett and Wanner, 1979; Nørsett and Trickett, 1984; Iserles and Nørsett, 1991 and Butcher, 2009).

Table 3-2: Butcher Tableau for SSP(2,2) scheme

$$\begin{array}{c|cc} 0.25 & 0.25 & 0 \\ 0.75 & 0.5 & 0.25 \\ \hline & 0.5 & 0.5 \end{array}$$

Table 3-3: Butcher Tableau for SSP(3,2) scheme

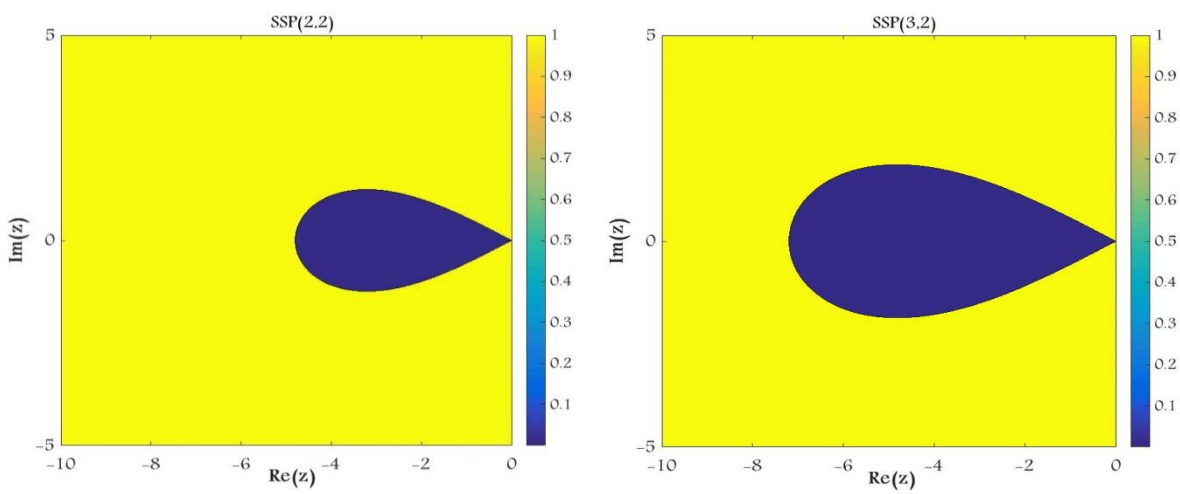
0.33333333333333	0.16666666666667	0	0
0.5	0.33333333333333	0.16666666666667	0
0.83333333333333	0.33333333333333	0.33333333333333	0.16666666666667
	0.33333333333333	0.33333333333333	0.33333333333333

Table 3-4: Butcher Tableau for SSP(3,3) scheme

0.146446609406726	0.146446609406726	0	0
0.5	0.353553390593275	0.146446609406726	0
0.853553390593274	0.353553390593273	0.353553390593273	0.146446609406726
	0.333333333333333	0.333333333333333	0.333333333333333

Table 3-5: Butcher Tableau for SSP(3,4) scheme

0.128886400515720	0.128886400515720	0	0
0.5	0.371113599484280	0.128886400515720	0
0.871113599484281	0.257772801031442	0.484454397937119	0.128886400515720
	0.302534578182651	0.394930843634698	0.302534578182651



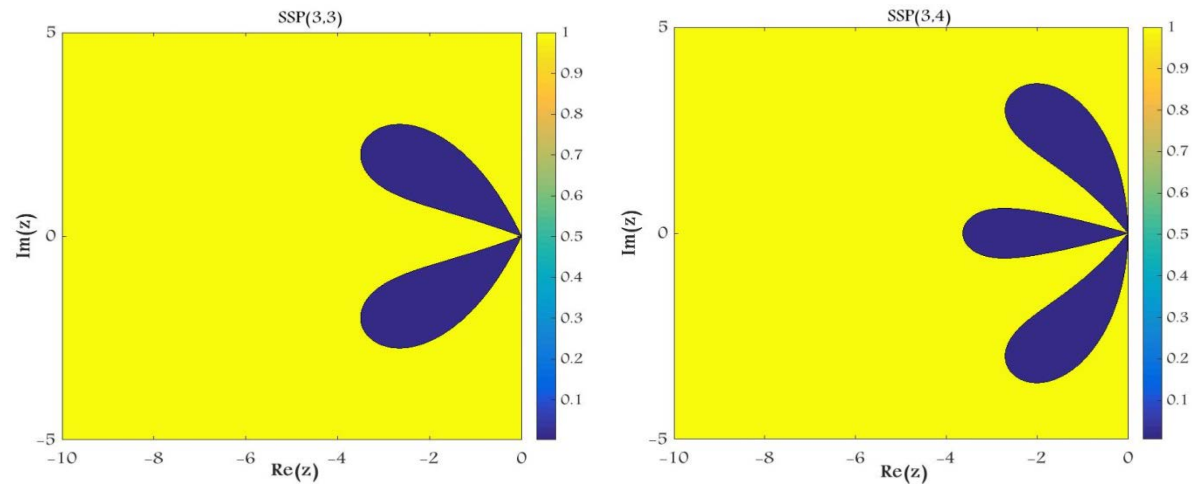


Figure 3-1: Relative stability analysis within domain of interest, $z \in C^-$, for SSP(2,2) and SSP(3,2) (first row) and SSP(3,3) and SSP(3,4) (second row), stable region, $|E| \leq 1$, shown in blue color.

Consequently, one may notice that the relative stability function is not an appropriate concept to provide the details of Runge-Kutta schemes. Moreover, as shown, it is just informative in studying the relation between order of accuracy and stability. This fact was also indicated by Leveque (2007).

Therefore, it is not more helpful than absolute stability function to minimize the dissipation and dispersion errors.

3.3. Dissipation and dispersion errors optimization

In this section, two issues, accuracy and stability limits, associated with the exact and numerical dispersion relations will be discussed. Afterward, the optimization of dissipation and dispersion errors and contribution of diffusion terms will be investigated. Finally, the A-stability and SSP properties will be described as the dissipation and dispersion optimization problems.

3.3.1. Accuracy and stability

Similar to the exact dispersion relation, (3-5), the numerical dispersion relation could be easily found by applying the Fourier transform to temporal term, i.e. $\hat{u}(t_n) = e^{-i\omega^* t_n}$ and substitution into (3-8):

$$e^{-i\omega^*\Delta t} = 1 + z\mathbf{b}^T(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{e} \tag{3-12}$$

in which $z = -(ick + k^2\nu)\Delta t$ and ω^* is called numerical frequency.

Furthermore, in stability analysis context, it is common to linearize the nonlinear governing equation by using perturbation technique. Hence, referring to (3-1), the variable is decomposed as $u = \bar{u} + \acute{u}$, in which $\bar{u}$ and $\acute{u}$ represent the mean and perturbed components. Therefore, the distribution of perturbed component, which represents the propagation of the errors as well, follows the original governing equation (with some approximation). Hence, one may expect that the absolute numerical stability function, (3-8), rules the error propagations as well. This means that Eq. 3-8, representing the ratio of errors in two successive time steps and the associated dispersion relation, (3-12), need to be less than one in order to have a stable scheme. Clearly, large values of the attributed variable, z, which includes time step and wavenumber, may result in large errors. Therefore, the main concern with instability, associated with spatial discretization, is the highest wavenumber corresponding to the finest grid size, i.e. $k_{max} = \frac{2\pi}{\lambda_{min}}$, in which $\lambda_{min} = 2\Delta x$.

Using (3-9) and (3-11-b) and the exact dispersion relation, (3-5), the following equation could be obtained:

$$e^{-i\omega^*\Delta t} = e^{-i\omega\Delta t} - \sum_{n=p+1}^{\infty} \frac{z^n}{n!} + \sum_{n=p+1}^{\infty} z^{\mathbf{n}}\mathbf{b}^T\mathbf{A}^{\mathbf{n-1}}\mathbf{e} \tag{3-13}$$

The last RHS summation in (3-13) presents the dissipation and dispersion error. It could be realized that, in order to have an accurate time integrator, the difference between numerical and exact dispersion relations, (3-13), needs to be minimized.

There are two important issues related to dispersion relations; the first one is related to accuracy and it is associated with (3-13). It means that minimizing the difference between exact and numerical dispersion relations in (3-13) increases the accuracy. The other issue is stability concern and corresponds to numerical dispersion relation in (3-12), which is necessary to be less than one. These issues were also indicated by Hu et al. (1996) within the concept of accuracy and stability limits.

In order to show these two issues, Figure 3-2 shows the magnitude of numerical dispersion relation, (3-12), in which the variable z just obtains the imaginary component. Referring to (3-5), the exact dispersion relation obtains the magnitude value equal to one. Hence, for accuracy reason, it is expected that time integrator schemes approximate the exact value, which is one, for large interval of the variable. In addition, due to the above discussion, the stable scheme needs to approximate the magnitude of numerical dispersion relation, (3-12), less than or equal one for large interval as well, which is due to error propagation.

Both schemes employed in Figure 3-2 are three-stage fourth order SDIRK schemes. However, the purple one is SSP optimized, so called SSP(3,4), which could be shown that it has less TE than the other one which is the unique three-stage fourth order A-stable scheme, developed by Crouzeix (1975).

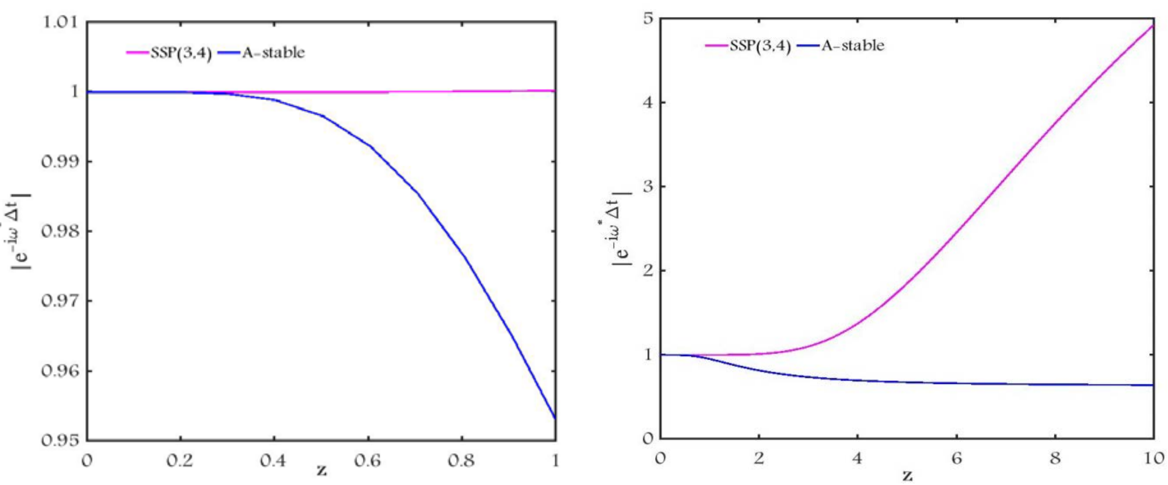


Figure 3-2: Comparison of absolute numerical stability functions' magnitude, SSP(3,4) (purple), three stage fourth order A-stable (blue), in terms of accuracy (left column) and stability (right column).

It is clear from left column that SSP(3,4) could remain more accurate due to less TE. However, as the absolute numerical stability function with A-stability property remains less or equal one for extremely large variable, it shows more stable behavior (right column). Therefore, it could be concluded that although SSP optimized schemes were developed from different concept and for different purposes, they try to minimize the dissipation and dispersion error to increase the accuracy limit, (3-13). Meanwhile, A-stability property tries to minimize these errors to increase the stability limit, (3-12).

3.3.2. Dissipation and dispersion optimization

In order to minimize the combination of dissipation and dispersion errors, Hu et al. (1996) introduced an integral function, which is a summation of the difference of exact and numerical phase relations in (3-13) for a specified interval. Note that the difference of exact and numerical dispersion relations is equivalent to the difference of exact and absolute numerical stability functions.

Du and Ekaterinaris (2016) optimized the dissipation and dispersion errors by maximizing the so called acceptable amplification (RAA) and acceptable phase shift (RAP). They indicated that the nominated errors need to include the contribution of physical and numerical diffusion terms. The purpose of the following discussion is to show that the objective function proposed by Hu et al. (1996) does not require including the diffusion terms.

Hu et al. (1996) tried to minimize the following integral as the objective function, while they proved it as a combination of dissipation and dispersion errors.

$$\mathrm{E_{rr}} = \min \int_0^{\Gamma} |\mathrm{R_n} - \mathrm{R_a}|^2 \mathrm{d}z \tag{3-14}$$

As they studied the hyperbolic compressible Euler system, the variable z in (3-3) and (3-8) just obtains imaginary component, i.e. $z = -\mathrm{i}ck\Delta t$, and consequently, a proposed objective function, (A-14), they employed was:

$$\mathrm{E_{rr}} = \min \int_0^{\Gamma} |\mathrm{R_n} - \mathrm{R_a}|^2 \mathrm{d}\sigma \tag{3-15}$$

in which $\mathrm{R_a} = e^{-\mathrm{i}\sigma}$ and $\sigma = -ck\Delta t$.

Hu et al. (1996) minimized the objective function in (3-15) in the interval $[0, \Gamma]$ while changing Γ manually. In the present paper, E_{rr} is defined as

$$E_{rr} = \min \int_0^{z_m} |R_n - R_a|^2 \, dz \tag{3-16}$$

a more general form of the error function in (3-15). It is the deviation of absolute numerical stability function from analytic stability function, or equivalently the difference of exact and numerical dispersion relations, in which variable z includes the real negative component, a representative of diffusion terms as well.

This integral would be calculated as the l_2-norm of the function $R_n - R_a$ in a discretized grid mesh. For any vector $\mathbf{X}$, the l_2-norm is bounded by l_∞-norm as follows (Boyd and Vandenberghe, 2004):

$$\|\mathbf{X}\|_\infty \leq \|\mathbf{X}\|_2 \leq \sqrt{m}\|\mathbf{X}\|_\infty \tag{3-17}$$

in which m is dimension of the vector $\mathbf{X}$. In this paper, it depends on the grid mesh size employed to calculate the integral (3-16). It should be mentioned that the l_∞-norm is the largest component of vector $\mathbf{X}$, which, hereafter, means the largest component of function $R_n - R_a$.

As the l_2-norm is lower bounded by l_∞-norm, it can be shown, for the present case, that maximizing the l_∞-norm results in the minimum value of l_2-norm. Consequently, the objective function in (3-16) could be replaced as follows:

$$E_{rr} = \max \|R_n - R_a\|_\infty \tag{3-18}$$

From triangular inequality, one obtains:

$$\|R_n\|_\infty - \|R_a\|_\infty \leq \|R_n - R_a\|_\infty \leq \|R_n\|_\infty + \|R_a\|_\infty \tag{3-19}$$

Referring to (3-3) and considering the fact that the desired domain is $z \in \mathbb{C}^-$, the analytical stability function is found as descending function with maximum value of one, i.e. $\|R_a\|_\infty = 1$.

Furthermore, one can find that "*max*" is a convex function (Boyd and Vandenberghe, 2004) and it preserves the inequality direction. Consequently, the following inequalities hold:

$$\max \|\mathrm{R_n}\|_\infty - 1 \leq \max \|\mathrm{R_n} - \mathrm{R_a}\|_\infty \leq \max \|\mathrm{R_n}\|_\infty + 1 \tag{3-20}$$

Hence, it is clear that maximizing the absolute numerical stability function results in maximum value for the objective function in (3-16), which is equivalent to minimizing the specified integral in (3-15). Therefore, in minimizing the summation of dissipation and dispersion errors, it is not required to include the contributions of diffusion terms and using the objective function proposed by Hu et al. (1996), Eq. 3-15, results in minimizing the dissipation and dispersion errors. In this regard, the constraint $\|\mathrm{R}_n\|_\infty \leq 1$ is necessary to impose on the optimization algorithm, as shown in Section 3.2.

In conclusion, in order to optimize the dissipation and dispersion errors using (3-15), the above discussions let us consider the analytical stability function without diffusion coefficient. Therefore, the objective function, (3-15), may be considered as:

$$\mathrm{E_{rr}^1} = \min \int_0^\Gamma |\mathrm{R_n} - 1|^2 \mathrm{d}\,\sigma \tag{3-21}$$

Moreover, by using (3-18), one may consider the objective function as:

$$\mathrm{E_{rr}^2} = \max \|\mathrm{R_n} - 1\|_\infty \tag{3-22}$$

Any above objective functions should be minimized subject to the constraint:

$$\|\mathrm{R}_n\|_\infty \leq 1 \tag{3-23}$$

Substituting the absolute numerical stability function, (3-8), into Eqs. 3-22 and 3-23, the new forms could be found as:

$$\mathrm{E_{rr}^2} = \max \|\mathrm{z}\mathbf{b}^\mathrm{T}(\mathbf{I} - \mathrm{z}\mathbf{A})^{-1}\mathbf{e}\|_\infty \tag{3-24}$$

subject to:

$$\|1 + z\mathbf{b}^{T}(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{e}\|_{\infty} \le 1 \tag{3-25}$$

Eq. 3-24 may be rearranged as:

$$E_{rr}^{2} = \max \|\mathbf{b}^{T}(z^{-1}\mathbf{I} - \mathbf{A})^{-1}\mathbf{e}\|_{\infty} \tag{3-26}$$

and the constraint may be reformed as:

$$-2 \le z\mathbf{b}^{T}(\mathbf{I} - z\mathbf{A})^{-1}\mathbf{e} \le 0 \tag{3-27}$$

As the desired domain in stability analysis is $z \in \mathbb{C}^{-}$, and z appears in the denominator in Eq. 3-26, it is clear that maximizing $\|z\|$ results in the maximum value for RHS of Eq. 3-26. Hence, the new objective function could be:

$$E_{rr}^{3} = \max \|z\|_{\infty} \tag{3-28}$$

subject to the constraint in (3-27).

3.3.3. A-stability property

According to Hairer and Wanner (1996), an implicit Runge-Kutta method is A-stable if and only if $R_n(z)$ is analytic for $Re(z) < 0$ and $|R_n(iy)| \le 1$ for all real y values.

Therefore it could be interpreted that A-stability condition implies bounding the magnitude of absolute numerical stability function in the whole left half plane, i.e. $z \in \mathbb{C}^{-}$ such that $\|R_n\|_{\infty} \le 1$.

One may find A-stability property as an optimization problem with objective function as (3-28) subject to the constraint in (3-27). Eq. 3-27 could be rearranged in an appropriate way for the discussion as:

$$-2 \le \mathbf{b}^{T}(z^{-1}\mathbf{I} - \mathbf{A})^{-1}\mathbf{e} \le 0 \tag{3-29}$$

The maximum value for $\|z\|_{\infty}$, which tends to infinity, lets the A-stability conditions become:

$$0 \le \mathbf{b}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{e} \le 2 \qquad (3\text{-}30)$$

Therefore, A-stability property aims at minimizing the dissipation and dispersion errors, which can be obtained by using the proposed integral in (3-15).

In addition, Eq. 3-4 shows that the real component of variable z includes the time step, Δt, the diffusion term, represented by ν, and wavenumber, k. It is clear that large values of time step and/or diffusion terms, in both physical and numerical forms, may move the numerical instability region farther along the negative real axis. This includes the maximum resolvable wavenumber, $k_{max} = \frac{2\pi}{\lambda_{min}}$, which represents the finest grid size as $\lambda_{min} = 2\Delta x$.

Consequently, A-stability property is necessary for calculation of large time steps and large diffusion terms. Although Du and Ekaterinaris (2016) correctly indicated that high dissipation is suitable for decreasing the calculation cost toward steady state solutions, their conclusion is inconsistent with their statement that A-stable schemes may not remain stable.

3.3.4. SSP property

The SSP DIRK schemes are convex combinations of Forward Euler (FE) method (Ketcheson et al., 2009). Hence, they preserve the monotonic behavior of FE method, but for limited range of time-step sizes, specified with absolute radius of monotonicity. This property prevents the non-oscillatory behavior in solving discontinuity and shock capturing. Ketcheson et al. (2009) calculated the optimized radius of monotonicity for SSP DIRK schemes up to order six with eleven stages. They described that the maximum radius of absolute monotonicity of the Runge-Kutta scheme is the largest $r \ge 0$ such that $(I + rA)^{-1}$ exists and

$$\begin{aligned} &\mathbf{K}(I + r\mathbf{A})^{-1} \ge 0 \\ &r\mathbf{K}(I + r\mathbf{A})^{-1}\mathbf{e}_s \le \mathbf{e}_{s+1} \end{aligned} \qquad (3\text{-}31)$$

in which $\mathbf{K} = \begin{pmatrix} \mathbf{A} \\ \mathbf{b}^{\mathrm{T}} \end{pmatrix}$ and $\mathbf{e}_s$ is $s \times 1$ vector of ones.

Clearly the objective function in SSP optimization problem is the same as (3-28) but with some additional constraints.

As the absolute numerical stability function is the ratio of solutions of two successive time steps, one may find that a requirement for non-oscillatory behavior is that the absolute numerical stability function needs to preserve the positive sign, which results in more restricted inequalities than (3-27) as:

$$-1 \leq \mathrm{z}\mathbf{b}^{\mathrm{T}}(\mathrm{I} - \mathrm{z}\mathbf{A})^{-1}\mathbf{e} \leq \mathbf{0} \tag{3-32}$$

Clearly, as the SSP optimization seeks positive value for radius of absolute monotonicity, r, simply replacing the variable z, in the interval $\mathrm{z} \in \mathbb{C}^{-}$, with $-\mathrm{r}$ gives the following condition:

$$0 \leq \mathrm{r}\mathbf{b}^{\mathrm{T}}(\mathrm{I} + \mathrm{r}\mathbf{A})^{-1}\mathbf{e} \leq \mathbf{1} \tag{3-33}$$

It is clear that equation set (3-31), i.e. constraints in SSP optimization problem, covers (3-33), the constraints in minimizing dissipation and dispersion errors problem. Consequently, SSP optimization aims at minimizing the dissipation and dispersion errors as well. Additionally, as it is clear from (3-31), SSP optimization imposes non negativity of all coefficients in DIRK scheme in order to guarantee the convex combination of FE (Ketcheson et al., 2009).

Therefore, It can be concluded that both A-stability and optimized SSP properties aim at minimizing the dissipation and dispersion errors. However, referring to the discussion in Section 3.3.1, A-stability tries to extend the range of stability while SSP optimization tries to extend the range of accuracy.

3.4. Non-linear viscous Burgers equation

Du and Ekaterinaris (2016), through calculation of non-linear viscous Burgers equation and using Newton method, compared their proposed scheme, DIRK-D, to the only A-stable three-stage fourth order SDIRK scheme (Crouzeix, 1975), so called DIRK-B, and reported some oscillatory behaviors. In this section, the non-linear viscous Burgers equation is examined to show the ability of A-stable schemes in obtaining the steady state solutions against DIRK-D which does not hold A-stability property.

The spatial terms are discretized with WENO-5 (Wang and Spiteri, 2007), which is not much different from WENO-5M employed by Du and Ekaterinaris (2016). The only difference is that they employed Newton iteration method to find the solutions, while here the Gauss Seidel approach is used.

The following equation represents the one-dimensional non-linear viscous Burgers equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2} \tag{3-34}$$

Neumann boundary conditions and initial condition are assumed as follows:

$$\begin{aligned} &u'(0,t) = u'(1,t) = 0 \\ &u(x,0) = R\cos(\pi x), \qquad\qquad 0 \le x \le 1 \end{aligned} \tag{3-35}$$

where R=5, and $\nu = 0.1$. The governing equation is calculated within domain $x \in [0,1]$, divided uniformly into 200 cells, and boundary conditions are set as symmetric. The analytical steady state solution for non-linear viscous equation was derived by Burns et al. (1998) as follows:

$$u(x) = 4.9799 \tanh\left[4.9799\frac{0.5 - x}{2\nu}\right] \tag{3-36}$$

In Figure 3-3, the l_2-norm of difference between numerical solutions in each time step and exact steady state solution, i.e. $\|u - u_{exact}\|_2$ is illustrated. Both schemes, DIRK-B and DIRK-D, have the same behavior with the same time-step size, $\Delta t = 3.2e - 4$. It is clear that the steady state solution is obtained by both schemes within 600 time-stepping. DIRK-B, thanks to A-stability property, tolerates larger time step, $\Delta t = 5.0e - 3$, but DIRK-D diverges under this time step. Therefore, DIRK-B reaches the steady state solution faster than DIRK-D, within less than 200 steps.

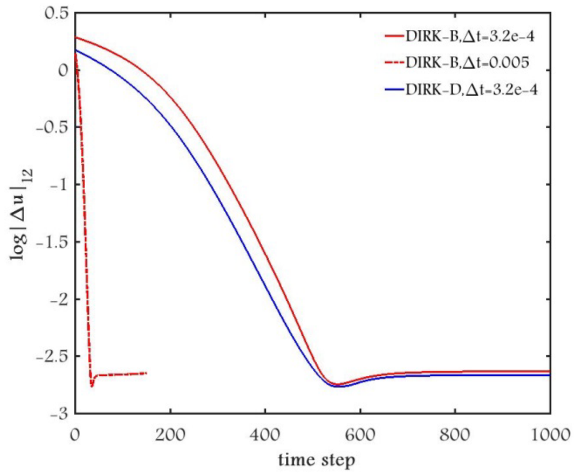


Figure 3-3: l_2-norm of difference between numerical solution and analytical steady state solution, $u - u_{exact}$, against number of time steps, DIRK-B (red) and DIRK-D (blue), $\Delta t = 3.2e-4$ (solid), $\Delta t = 5.0e-3$ (dash).

Following Du and Ekaterinaris (2016), the distribution of errors, defined as $\Delta u = u - u_{exact}$, at six successive time intervals, is illustrated in Figure 3-4 with time-step $\Delta t = 3.2e-4$s. Both DIRK-B and DIRK-D calculate identical error distributions. As mentioned, the Gauss Seidel approach was chosen for iteration. Comparing to the results presented by Du and Ekaterinaris (2016), which was calculated by Newton method, it is realized that Newton method predicts the results more accurately, as the level of errors is lower than those provided by Gauss Seidel.

However, it is clear that no oscillations appear for any of schemes using Gauss Seidel method. Consequently, one may find that the oscillation reported by Du and Ekaterinaris (2016) is related to Newton method and not to temporal integration methods. Moreover, Figure 5 shows the same distribution of errors generated by DIRK-B with larger time step, $\Delta t = 5.0e-3$s. This figure can demonstrate that oscillatory behavior, reported by Du and Ekaterinaris (2016), is due to Newton method and not due to the temporal integrators.

Figure 3-6 shows the CPU time distribution for both DIRK-B and DIRK-D schemes in solving the non-linear viscous Burger equation until t=0.76 s. DIRK-D, developed by Du and Ekaterinaris, enjoys the SSP property. This property let the time integrator prevent oscillatory behavior and therefore, it help to reduce the CPU time. As shown, although the stable time step is small for DIRK-D compared to DIRK-B, the CPU time is competitive.

3.5. Conclusion

In this note, the conclusions of Du and Ekaterinaris (2016) were further investigated. It was shown that the relative stability function, which is the ratio of absolute numerical stability function to analytical one, is not informative for optimization of dissipation and dispersion errors. It is just useful to study the relation between stability and order of accuracy. It was discussed that the high wavenumber is the main concern as a source of instability. In the literature, the optimization of dissipation and dispersion errors were studied by Hu et al. (1996) using an integral function as a combination of dissipation and dispersion errors. Du and Ekaterinaris (2016) included contribution of diffusion terms in their optimization procedure. It was shown that the proposed integral by Hu et al. (1996) is sufficient for optimization purposes and it does not require including the diffusion terms.

Moreover, referring to the analytical dispersion relation, (A-6), it is clear that the diffusion term appears as descending exponential function with maximum value equal one. This idea could be interpreted for numerical dispersion relation as well. Therefore, it is clear that the contribution of diffusion term does not require in the optimization.

Furthermore, it was shown that the proposed integral by Hu et al. (1996) covers the objectives of A-stability and SSP optimized properties. It was shown that A-stability property tries to extend the limit of stability by bounding the error propagation, while SSP optimized scheme tries to increase the limit of accuracy by decreasing the TE, which is commonly called error constant as well.

Calculation of non-linear viscous equation showed that A-stability property is necessary for stability under large time step sizes and large diffusion terms, and consequently, this property can accelerate achieving the steady state solutions. It was also shown that the oscillatory behavior reported by Du and Ekaterinaris (2016) is due to Newton method implementation and probably it was caused by the tolerances they set, as the Gauss Seidel method approaches the solution with no fluctuations.

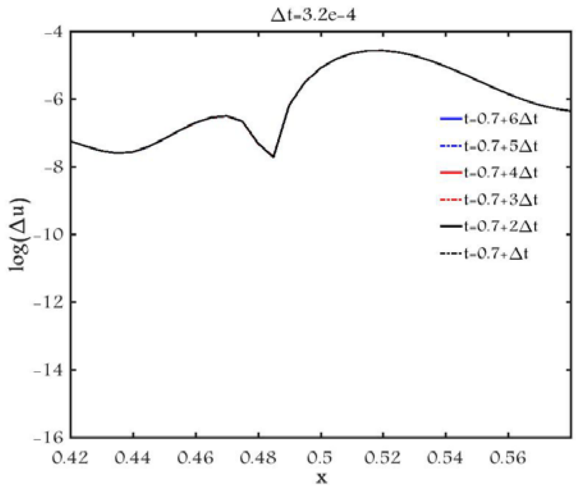


Figure 3-4: Distribution of errors for six successive time step, DIRK-B and DIRK-D (both identical), $\Delta t = 3.2e - 4s$.

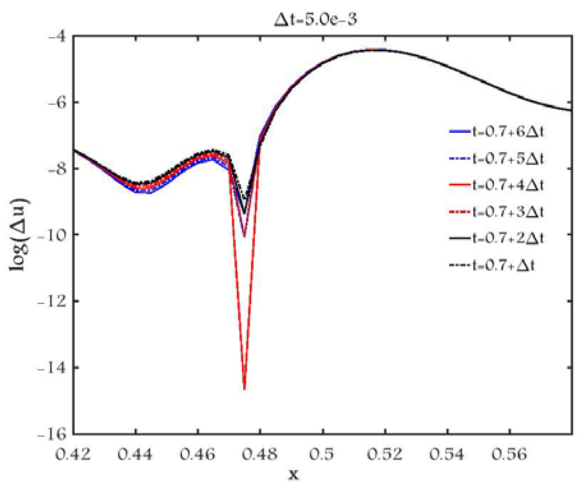


Figure 3-5: Distribution of errors for six successive time step, DIRK-B, $\Delta t = 5.0e - 3s$.

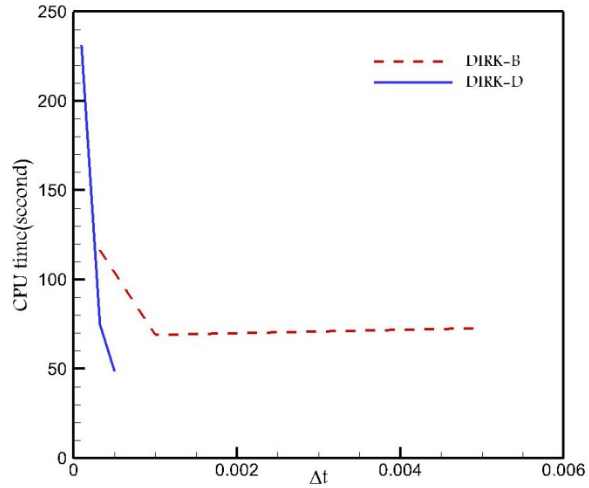


Figure 3-6: CPU time distribution in solving Burger equation till t=0.76 s, DIRK-B (red) and DIRK-D (blue).

Acknowledgment

The authors wish to acknowledge financial support from NSERC.

Chapter 4

An Optimized Stable and Accurate Second Order Runge-Kutta IMEX Scheme for Atmospheric Applications[2]

Abstract

The objective of this paper is to develop an optimized implicit-explicit (IMEX) Runge-Kutta scheme focusing on stability and accuracy for atmospheric applications. Following the common terminology, the proposed method is called IMEX-SSP2(2,3,2), as it has second order of accuracy and is composed of diagonally implicit two-stage and explicit three-stage parts. This scheme enjoys strong stability preserving (SSP) property for both parts. The proposed scheme is applied to the van der Pol nonlinear oscillation equation to examine its performance in stiff limits. This new scheme is applied to hydrostatic compressible Boussinesq equations in two different arrangements including (i) semi-implicit and (ii) horizontally explicit-vertically implicit (HEVI) forms. The new scheme preserves SSP property for larger region of absolute monotonicity compared to the other schemes in the same class. In addition, the numerical tests confirm that the IMEX-SSP2(2,3,2) improves the maximum stable time step as well as the level of accuracy compared to the other schemes in the same class. It is shown that the accuracy of this scheme is not reduced in stiff limits. It is demonstrated that A-stability property as well as satisfying "second-stage order" and stiffly accurate condition lead the proposed scheme to better performance than existing schemes for applications examined herein.

Keywords: IMEX, Runge–Kutta, Optimal Scheme, Atmosphere, Compressible Boussinesq, Hydrostatic.

4.1. Introduction

---

[2]This chapter has been published as Arman Rokhzadi, Abdolmajid Mohammadian, Martin Charron, Vol. 10, No. 1, pp. 18-42, January 2018**, Journal of Advances in Modeling Earth Systems Production**, 2018, doi: 10.1002/2017MS001065.

Due to the presence of different fast and slow waves in atmospheric field, extensive attention has been paid to IMEX schemes (Durran and Blossey, 2012; Ullrich and Jablonowski, 2012; Giraldo et al., 2013 and Weller et al., 2013). It has been proposed to resolve fast waves implicitly, as they have insignificant physical effects (Durran and Blossey, 2012; Ullrich and Jablonowski, 2012; Giraldo et al., 2013 and Weller et al., 2013). Therefore, their restriction on stability is circumvented and consequently, the maximum stable time step increases. Ascher et al. (1997) developed IMEX schemes up to four stages and third-order accuracy. They applied explicit schemes to the advection term and implicit schemes to the diffusion term. Stiffly accurate condition, which enables the schemes to remain accurate for stiff problems, was discussed. Pareschi and Russo (2004) developed new IMEX Runge-Kutta schemes with SSP property for explicit part and L-stability for implicit part, which is in Diagonally Implicit Runge-Kutta (DIRK) family. They defined the Asymptotic Preserving (AP) property which requires that an IMEX scheme with AP property, under some conditions, becomes a consistent discretization of the limit equilibrium equation characterized by explicit scheme. They also remarked that stiffly accurate condition could prevent the final layer errors and may guarantee the Asymptotic Accuracy (i.e. the scheme maintains its order of accuracy in the stiff limit). Those new schemes were tested on different applications such as shallow water equations, traffic flows and other cases. Boscarino (2007) performed an error analysis of IMEX schemes. The main attention was on A-stability property to impose less restriction than L-stability. He showed that, unfortunately, many popular schemes, more specifically those with order of accuracy more than two, suffer from order reduction when applied to stiff problems like van der Pol equation. In this regards, he provided relations for true order of accuracy in stiff limits and as he explained, satisfying stiffly accurate condition and higher "stage order condition", defined in Section 4.3.2, could enhance the accuracy. Higueras et al. (2014) investigated a family of optimized SSP IMEX schemes. They worked on some other properties for IMEX schemes rather than optimized radius of monotonicity; including L-stability, uniform convergence and inclusion of imaginary axis in the stability region. They presented the performance of their proposed schemes on advection-diffusion problems.

There are two different approaches in applying IMEX schemes to atmospheric applications called semi-implicit and HEVI forms. Semi-implicit models treat acoustic waves implicitly and advection is usually treated by semi-Lagrangian or explicit Eulerian methods. In semi-implicit form, gravity waves are treated explicitly. However, as Durran and Blossey (2012) and Giraldo et al. (2013) suggested, these terms could be calculated implicitly in order to increase the stability.

In HEVI models, the horizontal waves are evaluated explicitly while the terms responsible for vertical propagation are treated implicitly. In HEVI forms, the maximum stable time step decreases in comparison with semi-implicit methods. However, the HEVI forms are more efficient as the calculation cost per time step reduces.

Ullrich and Jablonowski (2012) proposed three Runge–Kutta IMEX schemes for the HEVI solution. They tested ARS(2,3,2) scheme proposed by Ascher et al. (1997) and concluded that they are less computationally expensive but nearly as accurate as Strang carryover scheme. As a novel approach, they used a Rosenbrock solution in order to treat all vertical terms implicitly rather than just the terms involved in wave propagation to circumvent the time-step restriction associated with vertical terms. Durran and Blossey (2012) analyzed IMEX multistep methods and applied their scheme to compressible Boussinesq equations using semi-implicit form of integration with buoyancy terms in both explicit and implicit parts. Their proposed schemes are based on Adams and backwards methods. However, multistep methods typically suffer from interaction of undamped computational modes with physical modes and often filtering is needed to avoid this problem. Weller et al. (2013) compared many different types of IMEX Runge-Kutta methods applied to hydrostatic compressible Boussinesq equations with both semi-implicit and HEVI forms. They introduced new arrangement for HEVI form in order to increase the maximum stable time step. It was, within the linear stability analysis, found out that, unfortunately, many of the employed schemes are unconditionally unstable in the explicit limit. Other schemes have regions of instability at very small time-steps. Weller et al. (2013) simulated incompressible flows by using a coarser grid and applied larger time steps. They indicated that non-stiffly accurate schemes perform better in simulating buoyancy while stiffly accurate schemes simulate incompressible flow more accurately and consequently would be better for global atmospheric modelling. Lock et al. (2014) considered von Neumman stability analysis for two linear systems as well as for acoustic wave equations. They investigated the behavior of amplification factor and phase angle produced by different IMEX schemes within specific range of CFL numbers compared to analytical ones. It was indicated that the analytical phase angle in the constant $k_x \Delta t$ decreases monotonically in terms of increasing k_z. This monotonic pattern demonstrates the constant group velocity. They found out that all examined schemes exhibit damping and retarded phase for large values of vertical wave numbers. Furthermore, they indicated that group velocity is reversed for some IMEX schemes but within ranges close to instability region. Some remarks were made on the analysis of acoustic wave equations as well. Giraldo et al. (2013) worked on multistage and on multistep methods. They developed an IMEX Runge-Kutta scheme called ARK2(2,3,2). They looked for an IMEX scheme that allowed unified representation of various non-hydrostatic flow regimes for three-dimensional Euler equations. It was indicated by Weller et al (2013), in their test cases, that ARK2(2,3,2) has better performance than other schemes with the same level of accuracy. The implicit part of the IMEX scheme proposed by Giraldo et al. (2013) satisfies "second-stage-order condition" as well as it is being stiffly accurate and has the same weight coefficients as explicit part. However, Giraldo et al. (2013) imposed the L-stability property and set the same internal calculation points for both explicit and implicit parts. Such conditions are typically unnecessary and restrict reliability of the scheme in achieving large time steps.

This paper considers maximization of stable time step for IMEX schemes while maintaining the acceptable level of accuracy and even improving it. Following Hu et al. (1996), Najafi-Yazdi and Mongeau (2013) and Nazari et al. (2015), we focus on minimizing an error function, which is the combination of dissipation and dispersion errors (Hu et al., 1996). In the IMEX scheme implementation, the fast frequency terms are treated implicitly and slow frequency ones are calculated explicitly. Hence, one can expect that the dissipation and dispersion errors would be augmented by large quantities of the ratio of fast to slow frequencies. Another issue considered in this paper is the order reduction as described in the following. It has been indicated in many papers such as Pareschi and Russo (2003), Carpenter and Kennedy (2003) and Boscarino (2007) that the order reduction phenomenon occurs for the most available IMEX schemes, especially when order of accuracy is more than two. Hence, we chose the IMEX Runge-Kutta scheme with three stages and second order of accuracy for both implicit and explicit parts. In addition, we considered the true order of accuracy in stiff limits as a constraint in the optimization process. This constraint is applied by solving the van der Pol equation which was recommended as a good test case for stiff problems (Boscarino, 2007 and Kværnø, 2004), in order to see the behavior of schemes in stiff limits (see Section 4.6.1).

This paper is structured as follows. In Section 4.2, governing compressible Boussinesq equations and two integration forms; semi-implicit and HEVI are presented. We introduce the IMEX Runge-Kutta formula, order conditions and desired properties in Section 4.3. Section 4.4 explains the optimization algorithm and the method to minimize the generated errors in calculation of large values of frequencies by IMEX schemes. Section 4.5 introduces the main characteristics of the proposed scheme and compares the linear stability analysis for different forms of integration to the well-studied IMEX scheme in the same class, ARK2(2,3,2). The results in Section 4.6 confirm the predictions presented in Section 4.5. The new IMEX scheme will be compared to ARK2(2,3,2) method through numerical simulations of the compressible Boussinesq equations for different semi-implicit and HEVI forms of integrations. Some concluding remarks complete the study in Section 4.7.

4.2. Governing Equations

Following Durran and Blossey (2012) and Weller et al. (2013), simulation of compressible Boussinesq governing equations is considered. In this section, different arrangements of temporal integration methods for compressible Boussinesq equations are reviewed.

In Geophysical Fluid Dynamics (GFD), wave-like motions of fluid associated with buoyancy perturbations (gravity waves), vorticity perturbations (Rossby waves) and acoustic waves (sound waves) support the transport of physical variables. In this study, we exclude the Coriolis terms and consequently the Rossby wave influences. Furthermore, the system is assumed to be two-dimensional. These two assumptions impose no restrictions on the analysis but simplify the calculations. The propagation of sound waves is described by hyperbolic PDEs. Gravity waves are also solutions to hyperbolic system but with some zero-order derivative terms of the unknown variables (Durran 2010, p. 12). These terms play no role in the classification of the governing equations as hyperbolic and simpler equations such as the Boussinesq equations closely approximate the gravity wave solutions (Durran 2010, p. 12). The compressible governing equations can thus be written as (Durran and Blossey, 2012),

$$\left(\frac{\partial}{\partial \mathrm{t}} + \mathbf{u}.\nabla\right)\mathbf{u} + \nabla \mathrm{P} = \mathrm{b}\hat{\mathbf{g}} \tag{4-1}$$

$$\left(\frac{\partial}{\partial \mathrm{t}}+\mathbf{u}.\nabla\right)\mathrm{b}+\mathrm{N}^2\mathrm{w}=0$$

$$\left(\frac{\partial}{\partial \mathrm{t}}+\mathbf{u}.\nabla\right)\mathrm{P}+\mathrm{c}_s^2\nabla.\mathbf{u}=0$$

where **u** is the velocity vector; w is the vertical component in the z direction, $\mathrm{p}=\frac{\mathrm{p}-\overline{\mathrm{p}}(\mathrm{z})}{\rho_0}$ is the deviation of pressure from a reference profile, $\overline{\mathrm{p}}(\mathrm{z})$, normalized by a reference density, ρ_0. The buoyancy term is $\mathrm{b}=\frac{\rho-\overline{\rho}(\mathrm{z})}{\rho_0}$ in which ρ(z) is a reference density profile; $\widehat{\mathrm{g}}$ is gravity vector and $\mathrm{N}^2=-\frac{\mathrm{g}}{\rho_0}\frac{\mathrm{d}\overline{\rho}}{\mathrm{dz}}$ where N is the buoyancy frequency (Brunt–Väisälä frequency).

4.2.1. Semi-implicit form

The sound waves with short amplitudes have little meteorological importance, but influence the stability limits. Thus maximum stable time step can be extended by applying numerical techniques to circumvent this restriction. Hence, the main idea of semi-implicit methods is to treat the terms responsible for sound waves implicitly. These terms are indicated by “fast” label in (4-2). Durran and Blossey (2012) and Giraldo et al. (2013) suggested that the buoyancy terms could also be treated implicitly in order to increase the stability range. Furthermore, Durran and Blossey (2012) emphasized that for hydrostatic approximation it is not practical to treat buoyancy explicitly. Therefore, the terms fundamental to buoyancy waves are labeled by “fast/slow”. The advection terms are specified by “slow” label and will be treated explicitly.

$$\begin{aligned}&\underbrace{\left(\frac{\partial}{\partial \mathrm{t}}+\mathrm{u}.\nabla\right)\mathrm{u}}_{\text{slow}}+\underbrace{\frac{\partial \mathrm{p}}{\partial \mathrm{x}}}_{\text{fast}}=0\\ &\underbrace{\left(\frac{\partial}{\partial \mathrm{t}}+\mathrm{u}.\nabla\right)\mathrm{w}}_{\text{slow}}+\underbrace{\frac{\partial \mathrm{p}}{\partial \mathrm{z}}}_{\text{fast}}=\underbrace{\mathrm{b}}_{\text{fast/slow}}\end{aligned}\qquad(4\text{-}2)$$

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right) b}_{\text{slow}} + \underbrace{N^2 w}_{\text{fast/slow}} = 0$$

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right) P}_{\text{slow}} + \underbrace{c_s^2\left(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}\right)}_{\text{fast}} = 0$$

4.2.2. HEVI form

In this form of arrangement, horizontal terms are integrated explicitly and vertical ones are calculated implicitly. The horizontal terms are described by "slow" label and the vertical terms are specified by "fast" label in (4-3). With the progress in parallel processing, this new form is getting more attractive due to less computational cost per time step, although the maximum stable time step is much lower than semi-implicit methods. Weller et al. (2013) called the classical HEVI form ufpref. The ufpref symbol is abbreviation to "u forward, pressure forward" since both u-momentum and pressure equations in (4-3) use forward (explicit) time differencing for horizontal gradients. They proposed a new HEVI form called ufpreb (u forward pressure backward). In ufpreb form, longitudinal derivative of divergence in pressure equation are treated implicitly in order to extend the stability region. This term is indicated by "fast/slow" label as:

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right) u}_{\text{slow}} + \underbrace{\frac{\partial p}{\partial x}}_{\text{slow}} = 0$$

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right) w}_{\text{slow}} + \underbrace{\frac{\partial p}{\partial z}}_{\text{fast}} = \underbrace{b}_{\text{fast}} \tag{4-3}$$

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right) b}_{\text{slow}} + \underbrace{N^2 w}_{\text{fast}} = 0$$

$$\underbrace{\left(\frac{\partial}{\partial t} + u.\nabla\right)}_{\text{slow}} P + c_s^2\left(\underbrace{\frac{\partial u}{\partial x}}_{\text{slow/fast}} + \underbrace{\frac{\partial w}{\partial z}}_{\text{fast}}\right) = 0$$

4.3. Runge–Kutta IMEX Methods and Desired Properties

A Runge-Kutta IMEX scheme is employed for any ordinary differential equations (ODE) of the form:

$$\frac{d\mathbf{y}}{dt} = s(\mathbf{y}, t) + f(\mathbf{y}, t) \tag{4-4}$$

where **y**, s and f include different flow variables. The s term presents long wavelengths and is treated by explicit part of the IMEX scheme while the f term deals with shorter waves and should be solved by implicit scheme. Two matrices $(A, \widetilde{A})$ and two vectors $(b, \tilde{b})$ define the implicit and explicit components of the IMEX scheme respectively, as described in the following. The IMEX scheme is defined by a double Butcher tableau as shown in Table (B-1). The explicit matrix, $\widetilde{A}$, is a strictly lower triangular matrix of $[\tilde{a}_{jk}]$ (k,j= 1, ..., ν) with $\tilde{a}_{jk} = 0$ for j ≥ k and ν is the number of internal stages. The implicit matrix A is lower triangular, with $a_{jk} = 0$ for $j > k$. This form of implicit matrix allows the nonlinear systems to approach the solution by forward substitutions of the previous stages, which is more time efficient. A Runge-Kutta IMEX solution is defined by ν internal stages and one final stage in order to advance from time t^n to t^{n+1} as:

$$Y^{(j)} = y^n + \Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} s(Y^{(k)}, t^n + \tilde{c}_k \Delta t) + \Delta t \sum_{k=1}^{j} a_{jk} f(Y^{(k)}, t^n + c_k \Delta t), \qquad j = 1, \ldots, \nu$$

(4-5)

$$y^{n+1} = y^n + \Delta t \sum_{j=1}^{\nu} \tilde{b}_j s(Y^{(j)}, t^n + \tilde{c}_j \Delta t) + \Delta t \sum_{j=1}^{\nu} b_j f(Y^{(j)}, t^n + c_j \Delta t),$$

IMEX methods preserve the linear invariants of ODE if weight coefficients b and $\tilde{b}$ are the same (Iserles and Zanna, 2000; Giraldo et al., 2013). Some IMEX schemes have been developed in which the vectors c and $\tilde{c}$ are set equal. However, it is not necessary to calculate implicit and explicit terms at the same intervals, which implies that vectors c and $\tilde{c}$ should not necessarily be equal.

Table 4-1: Double Butcher Tableau for v stage IMEX DIRK scheme

$$
\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}_t \end{array}
\qquad
\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \tilde{c}_1 & \tilde{a}_{j1} & 0 & 0 & 0 \\ : & : & : & 0 & 0 \\ \tilde{c}_\nu & \tilde{a}_{\nu 1} & \tilde{a}_{\nu 2} & \tilde{a}_{\nu l} & 0 \\ \hline & \tilde{b}_1 & \tilde{b}_2 & \dots & \tilde{b}_\nu \end{array}
\qquad
\begin{array}{c|c} c & A \\ \hline & b_t \end{array}
\qquad
\begin{array}{c|cccc} c_1 & a_{11} & 0 & 0 & 0 \\ : & a_{21} & a_{22} & 0 & 0 \\ : & : & : & a_{ll} & 0 \\ c_\nu & a_{\nu 1} & a_{\nu 2} & a_{\nu l} & a_{\nu\nu} \\ \hline & b_1 & b_2 & \dots & b_\nu \end{array}
$$

4.3.1. Order conditions

The following order conditions need to be satisfied for an IMEX Runge-Kutta scheme to be second order for both explicit and implicit parts and in IMEX combination.

For first order accuracy:

$$\sum_{i=1}^{\nu} \tilde{b}_i = 1, \qquad \sum_{i=1}^{\nu} b_i = 1 \tag{4-6-a}$$

for second order accuracy:

$$\sum_{i=1}^{\nu} \tilde{b}_i \tilde{c}_i = \frac{1}{2}, \quad \sum_{i=1}^{\nu} b_i c_i = \frac{1}{2} \tag{4-6-b}$$

and the coupling order conditions for second order of accuracy of the IMEX scheme,

$$\sum_{i=1}^{\nu} \tilde{b}_i c_i = \frac{1}{2}, \quad \sum_{i=1}^{\nu} b_i \tilde{c}_i = \frac{1}{2} \tag{4-6-c}$$

It can be observed that by setting $b_i = \tilde{b}_i$, no coupling between order conditions remains for maintaining second order of accuracy for combined IMEX scheme.

In addition, the following condition is imposed in this paper on the explicit scheme,

$$\sum_{i=1}^{\nu} \tilde{b}_i \tilde{a}_{ij} \tilde{c}_j = \frac{1}{6} \tag{4-6-d}$$

in order to eliminate the explicit error constant and satisfy one of the third order condition components as well.

4.3.2. Stage-order condition

Stage order is the order of accuracy of each internal stage, which is the maximum $\tilde{p}$ in the following additional condition:

$$\sum_{j=1}^{\nu} a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k, \quad 1 \leq k \leq \tilde{p} \tag{4-6-e}$$

The second-stage order condition was imposed on the proposed IMEX-SSP2(2,3,2) scheme. The reason is explained in the following. First, as described in Ketcheson et al. (2009), the stage order is the minimum order of accuracy when a scheme is applied to stiff problems. Second, the second-stage-order condition was imposed because the formal order of accuracy of the scheme is two. Therefore, the stage-order would be at most two. Hence, stage-order lower than two may cause slow convergence when computing a stiff ODE. The above additional condition (Eq. 4-6-e) implies that stage-order can not exceed one for explicit schemes (Ketcheson et al. 2009). In addition, for implicit DIRK schemes, this condition produces zero for the first row of the implicit matrix (A) components (Ketcheson et al. 2009). This is why the new explicit component of the IMEX scheme satisfies just the first-stage order condition and the implicit part obtaines zeros for the first row of A matrix.

4.3.3. Stiffly accurate condition

Another important condition that is described in the literature as necessary for stiff problems is stiffly accurate (Giraldo et al. 2013; Weller et al. 2013; Boscarino 2007 and Pareschi and Russo 2004). This condition, satisfied with $a_{\nu k} = b_k$ for $k = 1, \dots, \nu$, implies that the final stage is the same as the last sub stage. Clearly, this property enhances the performance of the scheme in terms of more memory requirements and time-efficiency as the final solution is already achieved by the last sub stage.

The L-stability property requires that implicit stability function, which is defined in the complex coordinates and in terms of complex variable (z), converges to zero as the real part of z goes to infinity. This condition clearly requires very strict constraints to be satisfied and it is just necessary for calculation of infinite quantities. It is known from wave number definition, $k_{max} = \frac{2\pi}{\lambda_{min}}$, that the shortest resolvable wave is related to the smallest grid spacing, $\lambda_{min} = 2\Delta z$. In most atmospheric applications, the smallest grid spacing which is used to capture very small boundary layer processes is usually Δz=10m (Lock et al., 2014). One can realize that this maximum wavenumber could generate large but finite quantities. Therefore, it can be concluded that the required L-stability condition imposes unnecessary strict constraints that limit the maximum stable time step. Hence, we seek only implicit A-stable schemes, i.e. the stability function remains less than unity in infinity limits (but not necessarily zero).

4.3.4. Performance in stiff problems and order reduction

Stiff problems are the other important concern in this study. There are diverse definitions for stiff problems in the literature. According to Dekker and Verwer (1984), the precise definition is; The solution varies slowly but there are some perturbations which are rapidly damped. This definition classifies the atmospheric applications as the stiff problems (Durran 2010, p. 72; Ullrich and Jablonowski, 2012). As Boscarino (2007) showed, many IMEX schemes suffer from order reduction in solving stiff problems. This means that the true order of accuracy of the IMEX scheme for stiff problems may become lower than the formal order. This is the case for stiff problems in the form:

$$y' = F(y) + \frac{1}{\varepsilon} G(y) \tag{4-7}$$

which is usually transformed into a partitioned system as:

$$
\begin{aligned}
y' &= f(y,z) \\
\varepsilon z' &= g(y,z)
\end{aligned}
\tag{4-8}
$$

Here, the y-component equation is calculated by explicit part while the z-component is treated by implicit part of the IMEX scheme. As Boscarino (2007) classified, schemes of the type-CK IMEX have the following form for implicit part,

$$
A = \begin{pmatrix} 0 & 0 \\ a & \hat{A} \end{pmatrix} \tag{4-9}
$$

in which **a** is a vector and $\hat{\mathbf{A}}$ is an invertible (ν-1) × (ν-1) matrix. The new IMEX method proposed in this paper is of the type-CK as it has zeros for the first row of the implicit A matrix. According to Theorem 3.2 in Boscarino (2007), for the type-CK schemes, the global errors are expected to satisfy the following relations for both non-stiff (y) and stiff (z) variables:

$$
\begin{aligned}
y_n - y(x_n) &= O\left(h^{\tilde{q}+2}, h^p\right) + O(\varepsilon h^2) \\
z_n - z(x_n) &= O\left(h^{\tilde{q}+1}, h^p\right) + O(\varepsilon h)
\end{aligned}
\tag{4-10}
$$

in which $\tilde{q}$ is the stage-order and p is the order of accuracy (both for the explicit part of the IMEX scheme). According to the previous section, stage order for explicit schemes can not exceed one ($\tilde{q} = 1$). Morover, the order of accuracy for the new scheme is two ($p = 2$). Thus, as we will show in the van der Pol test case, we expect that the IMEX-SSP2(2,3,2) does not experience order reduction because it preserves the second order of accuracy in global error prediction for both stiff and non-stiff variables based on (4-10). This will be confirmed by numerical experiments in Section 4.6.1.

4.3.5. SSP property

There are extensive studies on SSP property in the literature (e.g. Ketcheson et al., 2009 and Higueras et al., 2014). This property holds for schemes with non-negative coefficients, (A, b) and $(\tilde{A}, \tilde{b})$, within a limted range of time step called radius of monotonicity. According to

Higueras et al. (2014), A ν-stage Runge Kutta scheme has radius of monotonicity r if $(\mathrm{I}+\mathrm{r}\mathrm{A})^{-1}$ exists and

$$(\mathrm{I}+\mathrm{rK})^{-1}\mathrm{K} \geq 0$$
$$(\mathrm{I}+\mathrm{rK})^{-1}\mathrm{e} \geq 0 \tag{4-11}$$

in which

$$\mathrm{A} = \begin{pmatrix} \mathrm{A} & 0 \\ \mathrm{b}^{T} & 0 \end{pmatrix} \tag{4-12}$$

These inequalities should be considered component-wise in which e is a (ν-1)×1 vector with all components being equal to one. In the case of IMEX schemes (Higueras et al., 2014), the radius of monotonicity is $(-\mathrm{r_e}, -\mathrm{r_i})$ with $\mathrm{r_e} > 0, \mathrm{r_i} > 0$ if the matrix $\mathrm{I} + \mathrm{r_i}\mathrm{K} + \mathrm{r_e}\widetilde{\mathrm{K}}$ is nonsingular, i.e. $\left(\mathrm{I} + \mathrm{r_i}\mathrm{K} + \mathrm{r_e}\widetilde{\mathrm{K}}\right)^{-1}\mathrm{e} \geq 0$ and

$$(\mathrm{I} + \mathrm{r_i}\mathrm{K} + \mathrm{r_e}\widetilde{\mathrm{K}})^{-1}\mathrm{K} \geq 0$$
$$(\mathrm{I} + \mathrm{r_i}\mathrm{K} + \mathrm{r_e}\widetilde{\mathrm{K}})^{-1}\widetilde{\mathrm{K}} \geq 0 \tag{4-13}$$

4.4. Optimization algorithm

In this section, we describe the stability analysis for Runge–Kutta IMEX schemes for two-frequency systems (slow and fast components) with no diffusion or reaction terms. Then we will explain how to derive the new proposed scheme using an optimization procedure.

As the atmospheric applications are dominated by hyperbolic terms, the s and f terms in (4-4) are assumed like first order derivative convective terms. Hence, (4-4) becomes

$$\frac{\mathrm{dy}}{\mathrm{dt}} = -\mathrm{c_s}\frac{\mathrm{dy}}{\mathrm{dx}} - \mathrm{c_s}\frac{\mathrm{dy}}{\mathrm{dz}} \tag{4-14}$$

In von Neumann stability analysis, one assumes that the solution has the form $\mathrm{y} = \hat{\mathrm{y}}(\mathrm{t})\exp(\mathrm{i}\mathrm{k_x}\mathrm{x} + \mathrm{i}\mathrm{k_z}\mathrm{z})$. Substitution in (B-14) and replacing $\mathrm{c_s}\mathrm{k_x} = \mathrm{s}$ and $\mathrm{c_s}\mathrm{k_z} = \mathrm{f}$, yields

$$\frac{d\hat{y}}{dt} = -is\hat{y} - if\hat{y} \tag{4-15}$$

In this equation, f represents fast-wave terms with large quantities in comparison to slow-wave (s terms). We aim at finding an optimized IMEX scheme to deal reliably with the large range of the ratio $\frac{f}{s}$ such that the formal order of accuracy is not reduced. Therefore, we redefine the ODE problem (4-15) in terms of $\frac{f}{s}$ ratio.

$$\frac{d\hat{y}}{dt} = -i\hat{y} - i\frac{f}{s}\hat{y} \tag{4-16}$$

The analytical solution for this equation is

$$\hat{y}(t) = y_0 e^{-i(1+\frac{f}{s})t} \tag{4-17}$$

Following the von Neumann stability analysis, the amplification factor is defined between solutions in time t and $t + \Delta t$ like $\hat{y}(t + \Delta t) = R_a \hat{y}(t)$, where the analytical amplification factor is

$$R_a = e^{-i(1+\frac{f}{s})\Delta t} \tag{4-18}$$

Applying the general IMEX Runge-Kutta formula, (4-5), to (4-16), one obtains a linear relation similar to (4-5) as,

$$\begin{aligned} Y^{(j)} &= y^n - i\Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} Y^{(k)} - i\frac{f}{s}\Delta t \sum_{k=1}^{j} a_{jk} Y^{(k)}, \qquad j = 1, \dots, \nu \\ y^{n+1} &= y^n - i\Delta t \sum_{j=1}^{\nu} \tilde{b}_j Y^{(j)} - i\frac{f}{s}\Delta t \sum_{j=1}^{\nu} b_j Y^{(j)}, \end{aligned} \tag{4-19}$$

The numerical amplification factor is defined as the ratio of the numerical solution in all internal sub-stage solutions, $Y^{(j)}$, and final stage solution, y^{n+1}, to the numerical solution in time step n, y^n. Therefore, using (4-19) yields,

$$R^{(j)} = 1 - i\Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} R^{(k)} - i\frac{f}{s}\Delta t \sum_{k=1}^{j} a_{jk} R^{(k)}, \qquad j = 1, \dots, \nu$$

$$R = 1 - i\Delta t \sum_{j=1}^{\nu} \tilde{b}_j R^{(j)} - i\frac{f}{s}\Delta t \sum_{j=1}^{\nu} b_j R^{(j)}, \tag{4-20}$$

To obtain the optimized scheme, Hu et al. (1996), Najafi-Yazdi and Mongeau (2013) and Nazari et al. (2015) suggested minimizing an error function, which is the combination of dissipation and dispersion errors. For the IMEX Runge-Kutta method studied in this paper, the error function, E_{rr}, is defined as

$$E_{rr} = \int_0^{\Gamma} |R_n - R_a|^2 d(\frac{f}{s}) \tag{4-21}$$

where, E_{rr} is defined in terms of the variable $\frac{f}{s}$ in order to minimize the dissipation and dispersion errors for the large range of fast to slow frequency ratios. During the optimization procedure, we imposed to satisfy the order conditions in (4-6) with a very tight tolerances, $\mathrm{tol} = 10^{-12}$. Furthermore, the stiffly accurate condition for implicit scheme, which is satisfied with $a_{\nu k} = b_k$ for $k = 1, \dots, \nu$, was imposed on the system with $\mathrm{tol} = 10^{-9}$. The non-negativity of all coefficients, i.e. two matrices $(A, \widetilde{A})$ and two vectors $(b, \tilde{b})$, were also imposed to obtain the SSP property.

It is desirable to obtain the minimum value for the error function within the largest interval $[0, \Gamma]$. Therefore, Γ was varied subject to the condition that the IMEX scheme does not experience order reduction. As Pareschi and Russo (2003), Carpenter and Kennedy (2003) and Boscarino (2007) reported, most IMEX schemes suffer from order reduction. In this study, during the optimization process and calculating the error function in (4-21), it was realized that true order of accuracy is sensitive to the upper limit of the integral, Γ. Hence, Γ was varied such that the IMEX scheme does not experience the order reduction phenomenon. The van der Pol equation was tested for this purpose because it has been introduced as a good representative for stiff problems (Kværnø; 2004; Boscarino, 2007). To obtain A-stability property, the constraint $\left|\frac{R_n}{R_a}\right| \leq 1$ was applied to all grid points employed for calculation of the error function in (4-21).

4.5. A new optimized second order three-stage IMEX scheme

In this section, we present the coefficients for the new IMEX scheme in the form of double Butcher Tableau (Table 4-1) and the essential properties are discussed. Furthermore, we compare the linear stability region for semi-implicit and HEVI forms produced by the new scheme and ARK2(2,3,2) method which has been well studied in the literature. As described by Pareschi and Russo (2004), we use the denomination IMEX-SSP k(s,σ,p) in which k is the order of accuracy of explicit scheme, s and σ are number of stages for implicit and explicit schemes respectively, and p is the order of accuracy of the combined IMEX scheme. Moreover, as this new IMEX scheme holds the SSP property, it is introduced as IMEX-SSP2(2,3,2). The coefficients of the IMEX-SSP2(2,3,2) and ARK2(2,3,2) are presented in Table 4-2 and Table 4-3.

Table 4-2: Double Butcher Tableau for IMEX2(2,3,2) scheme

0	0	0	0
0.711664700366941	0.711664700366941	0	0
0.994611536833690	0.077338168947683	0.917273367886007	0
	0.398930808264688	0.345755244189623	0.255313947545689

0	0	0	0
0.707685730198550	0.353842865099275	0.353842865099275	0
1	0.398930808264689	0.345755244189622	0.255313947545689
	0.398930808264688	0.345755244189622	0.255313947545689

Table 4-3: Double Butcher Tableau for ARK2(2,3,2) scheme

0	0	0	0
$2\left(1-\frac{1}{\sqrt{2}}\right)$	$2\left(1-\frac{1}{\sqrt{2}}\right)$	0	0

$$
\begin{array}{c|ccc}
1 & 1-\frac{1}{6}\left(3+2\sqrt{2}\right) & \frac{1}{6}\left(3+2\sqrt{2}\right) & 0 \\
\hline
 & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 1-\frac{1}{\sqrt{2}}
\end{array}
$$

$$
\begin{array}{c|ccc}
0 & 0 & 0 & 0 \\
2\left(1-\frac{1}{\sqrt{2}}\right) & 1-\frac{1}{\sqrt{2}} & 1-\frac{1}{\sqrt{2}} & 0 \\
1 & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 1-\frac{1}{\sqrt{2}} \\
\hline
 & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 1-\frac{1}{\sqrt{2}}
\end{array}
$$

4.5.1. Characteristics of the proposed scheme IMEX-SSP2(2,3,2)

The following properties distinguish the proposed scheme from existing schemes in the same level:

1- The new scheme improves the maximum stable time step, as well as the accuracy, for all different forms of integration studied here, including HEVI and semi-implicit forms, compared to the most recent one in the literature, ARK2(2,3,2).
2- This new scheme has SSP property with radius of monotonicity region much larger than ARK2(2,3,2). The nonnegative coefficients for the IMEX scheme, (A, b) and $(\tilde{\mathrm{A}}, \tilde{\mathrm{b}})$, guarantee the SSP property (Higueras, 2014).
3- This new scheme shows the true order of accuracy in solving van der Pol test case as a representative for stiff problems.
4- As Norsett and Wolfbrandt (1977) proved, Singly Diagonally Implicit Runge-Kutta (SDIRK) schemes have the minimum error constant. The minimum achievable error constant for two-stage implicit schemes, as indicated by Burrage (1978), is $\mathrm{E}_{\min} \cong 0.041$. The implicit part of the new proposed IMEX scheme, with different a_{ii} components, produces the error constant $\mathrm{E} \cong 0.048$, which is not much larger than the minimum achievable value (see Appendix A for details). Instead, as we will show

later, in the numerical simulation of the compressible Boussinesq equations, it could extend the maximum stable time step as well as the accuracy.

Furthermore, the new proposed scheme holds the following properties as essential properties for reliable and efficient simulations:

5- The stability function of the implicit scheme has the upper bound less than unity in the infinity limit and contains the whole left half-plane. This property is necessary for stiff problems and is guaranteed by A-stability property (Hairer and Wanner, 1996).
6- Consistent with the results in the literature (Giraldo et al. 2013, Weller et al. 2013, Boscarino 2007, Pareschi and Russo 2004), we found stiffly accurate condition as a necessity for atmospheric applications. Furthermore, this property also leads to computational efficiency, because the final solution does not need to be calculated as it is the same as the last sub stage.
7- Many authors (e.g. Pareschi and Russo, 2003; Carpenter and Kennedy, 2003 and Boscarino, 2007) have stated that IMEX schemes with order of accuracy more than two suffer from order reduction. In addition, as Higueras et al. (2014) indicated, the accuracy is generally limited by spatial resolution. Therefore, third order accuracy for temporal integration schemes does not necessarily promise more benefits when the computational cost of extra stages is also taken into account. Hence, we chose the second order of accuracy for both explicit and implicit parts and for the IMEX combination as well.
8- The implicit scheme satisfies second-stage order condition, which guarantees the second order of accuracy for each internal stage. At the beginning of this investigation, it was tried to find a first-stage order IMEX scheme. However, it was realized that second-stage order IMEX method can remain stable for larger time steps, especially for semi-implicit form when buoyancy terms are treated explicitly.

The main purpose of using IMEX scheme is to circumvent the physically insignificant large fast frequency terms by applying the implicit part. Hence, a desired property for an implicit scheme is the ability to suppress those large terms. The authors believe that, for atmospheric applications, the L-stability property is not a necessary condition and may unjustifiably limit the scheme's performance. Therefore, an A-stable implicit scheme is sought. In Figure B-1, stability regions for both explicit and implicit parts of the new IMEX scheme, illustrated in white color, are compared to ARK2(2,3,2). The stability functions for proposed explicit and implicit schemes are as follows (Appendix A)

$$\begin{aligned} R_{\tilde{A}}(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} \\ R_A(z) &= \frac{-0.091z^2 + 0.391z + 1}{0.0903z^2 - 0.6092z + 1} \end{aligned} \tag{4-22}$$

For explicit part, the stability field $|R(z)| \leq 1$, should include the imaginary axis as much as possible, due to presence of the dominant hyperbolic terms treated explicitly in the atmospheric applications, i.e. advection terms, longitudinal pressure and velocity derivatives in some cases. As demonstrated by Higuares et al. (2014), the stability region of three-stage explicit schemes includes a part of the imaginary axis and the maximum achievable value is $\mathrm{Im}(z) = \pm 2$. This value for new IMEX-SSP2(2,3,2) and ARK2(2,3,2) schemes is $\mathrm{Im}(z) = \pm\sqrt{3}$ (See Appendix A). It is observed that the new scheme has a slightly lower value than the maximum achievable one. However, an explicit scheme, which achieves the maximum value along the imaginary axis, intersects the real negative axis in $\mathrm{Re}(z) = -2$, whereas the new proposed scheme extends the stability region, $|R(z)| \leq 1$, along the real negative axis to $\mathrm{Re}(z) \cong$ -2.513. This property, i.e. large stability region along negative real axis, is suited for parabolic terms such as diffusion (Hairer and Wanner, 1996). In addition, as indicated by Higuares et al. (2014), the stability function should be nonnegative for a large interval of the negative real axis. Indeed, this property is needed to stabilize the scheme for the dissipativity of the spatial discretization (Strikwerda, 2004). The stability function of the new explicit scheme remains nonnegative up to $\mathrm{Re}(z) \cong -1.596$, whereas the explicit scheme with the maximum value on the imaginary axis is nonnegative up to $\mathrm{Re}(z) \cong -1.3$. (Appendix).

For implicit scheme, A-stability property is a necessary and sufficient condition for stability along imaginary axis as well as in the whole negative real half plane (Hairer and Wanner, 1996).

In Figure 4-1, the stability regions for explicit schemes (in the first column) and implicit schemes (in the second column) are compared for new proposed scheme and ARK2(2,3,2). It can be seen that the explicit schemes are the same. However, the new proposed implicit scheme improves the stability region in the positive half plane, $\mathrm{Re}(z) \geq 0$ and consequently, as shown in Figure 2, this improvement extends the stability region of the combined IMEX scheme.

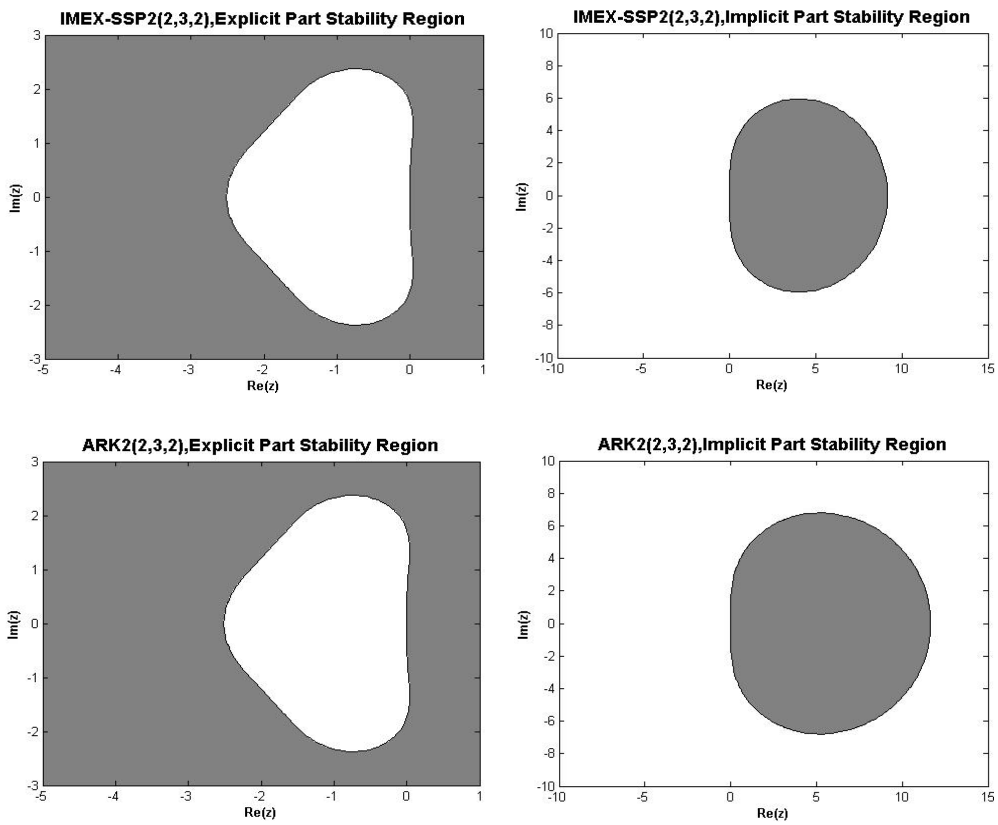


Figure 4-1: Stability region for the components of IMEX-SSP2(2,3,2) (first row) and ARK2(2,3,2) (second row) left: explicit part, right: implicit part, the stable region, i.e. $|R(z)| \leq 1$, is presented in white.

In Figure 4-2, we present the stability region of the proposed IMEX scheme and ARK2(2,3,2). The white region represents the stability region in which $|\mathrm{R(z)}| \leq 1$. The ARK2(2,3,2) has a narrower stability region. However, the two schemes have almost the same stability region along the imaginary axis.

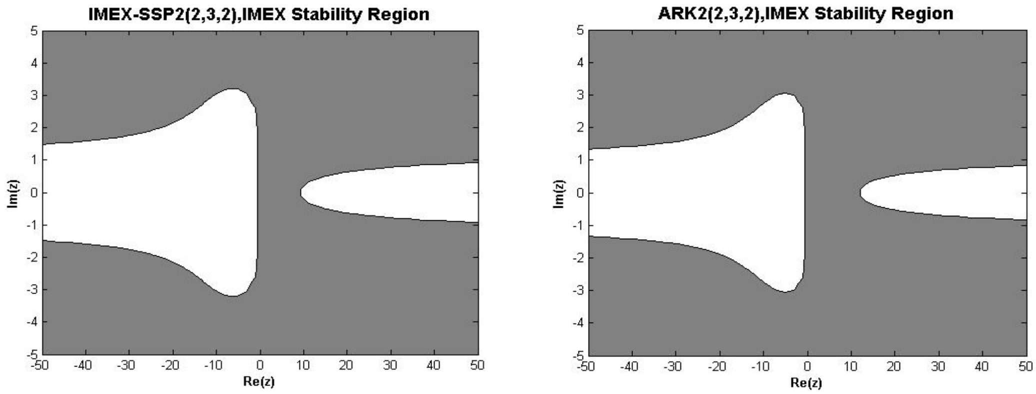


Figure 4-2: Stability region for IMEX-SSP2(2,3,2) (left) and ARK2(2,3,2) (right). The stable region, i.e. $|R(z)| \leq 1$, is presented in white.

In the case of radius of monotonicity, for IMEX-SSP2(2,3,2), we have calculated $\mathrm{R}_{\tilde{\mathrm{A}}} = 0.11841$ and $\mathrm{R}_{\mathrm{A}} = 2.3031$ for explicit and implicit parts, respectively. These data for ARK2(2,3,2) are $\mathrm{R}_{\tilde{\mathrm{A}}} = 0.0503$ and $\mathrm{R}_{\mathrm{A}} = 2.4142$ with significantly smaller radius for the explicit part. The regions of absolute monotonicity for both IMEX schemes are provided in Figure 3 which demonstrates a significantly larger region for IMEX-SSP2(2,3,2).

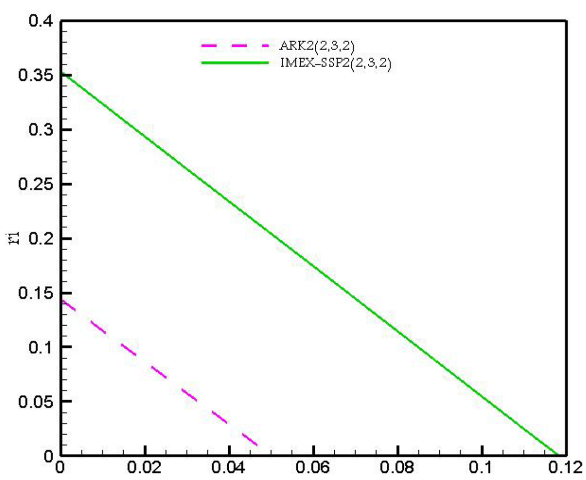


Figure 4-3: Region of absolute monotonicity IMEX-SSP2(2,3,2) solid line and ARK2(2,3,2) dashed line.

4.5.2. Linear stability analysis

In this section, we compare the new scheme and ARK2(2,3,2) via a linear stability analysis for two-frequency systems following Lock et al. (2014) and for acoustic wave equations in HEVI approach, following Weller et al. (2013). Then, linear stability analysis will be performed for compressible Boussinesq equations in semi-implicit approach as discussed in Durran and Blossey (2012).

4.5.2.1. Two-frequency system analysis

As it was described in Section 4.4, by applying the von Neumann stability analysis to (4-14), one can derive the following relation for analytical amplification factor,

$$R_a = \exp(-c_s k_x \Delta t - c_s k_z \Delta t) \tag{4-23}$$

Following Lock et al. (2014), it is interesting to see the performance of the IMEX scheme in terms of phase behavior and amplification factor for shortest resolvable waves with $k_{max} = \frac{2\pi}{\lambda_{min}}$, where λ_{min}= $2\Delta x$ in horizontal direction and $2\Delta z$ in vertical direction. Therefore, a non-dimensional time step which is defined as $\Delta t^* = \frac{c_s \pi \Delta t}{\Delta x}$ is equivalent to $\Delta t^* = c_s k_{x,max} \Delta t$. Substituting the new non-dimensional variables into (B-23) gives

$$R_a = \exp(-\Delta t^* - \frac{k_{z,max}}{k_{x,max}} \Delta t^*) \tag{4-24}$$

One can observe in (4-24) that the analytical magnitude of the amplification factor is one and the analytical phase is $\Theta = -\Delta t^* - \frac{k_{z,max}}{k_{x,max}}$. Figure 4-4 shows the contours of analytical phase for various Δt^* and $\frac{k_{z,max}}{k_{x,max}}$. As can be seen in this figure, the analytical phase monotonically decreases with the increase of Δt^* and $k_{z,max}$. This means that the analytical group velocity, which is defined as derivative of phase angle in terms of wavenumbers, i.e. $\frac{\partial \Theta}{\partial k}$, is negative.

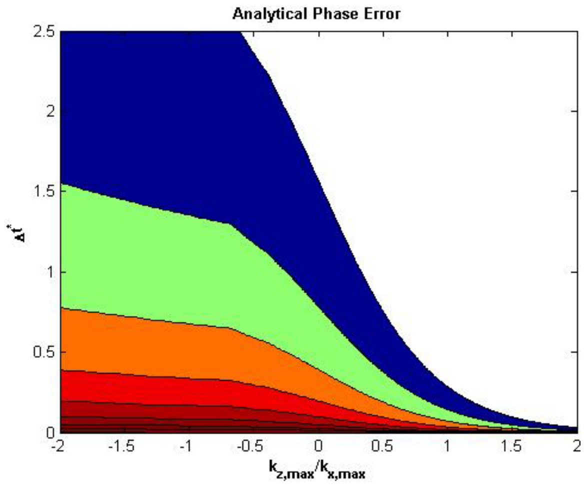


Figure 4-4: Exact phase for two-frequency systems, Eq. (4-24), plotted against Δt^* on vertical axis and logarithmic scale of $\frac{k_{z,max}}{k_{x,max}}$ on horizontal axis, in the specified ranges as Lock et al. (2014).

Numerical amplification factor and phase values are illustrated in Figure 4-5 for both IMEX schemes, i.e. new IMEX scheme and ARK2(2,3,2). The stability regions, where the amplification factor is less than one, are shown in the first row by white color. The new IMEX scheme maintains stability for larger time step in large values of vertical wavenumber. One can see this along vertical line at $\frac{k_{z,max}}{k_{x,max}} = 2$. In the second row of Figure B-5, the calculated phase angles for both schemes are compared. Within stability region far from instability boundary, both models preserve the direction of phase angle decreasing in both horizontal and vertical dimensions same as the analytical solution. However, the rate of decreasing is different from exact solution, particularly in the horizontal direction. This difference means the calculated group velocity is not the same as exact one. In stability region approaching the large values of vertical wavenumbers along horizontal axis, both schemes illustrate poor behavior as the phase angle increases for some range of wavenumbers. This means that group velocity is reversed. However, it is interesting that the new IMEX scheme improves the growth pattern of the phase angle at the large values of $\frac{k_{z,max}}{k_{x,max}}$ for a larger range of Δt^*. Therefore, one can conclude that the new scheme improves the calculated phase angle, specifically in large values of wavenumbers and for large time steps, which is the case in the atmospheric applications.

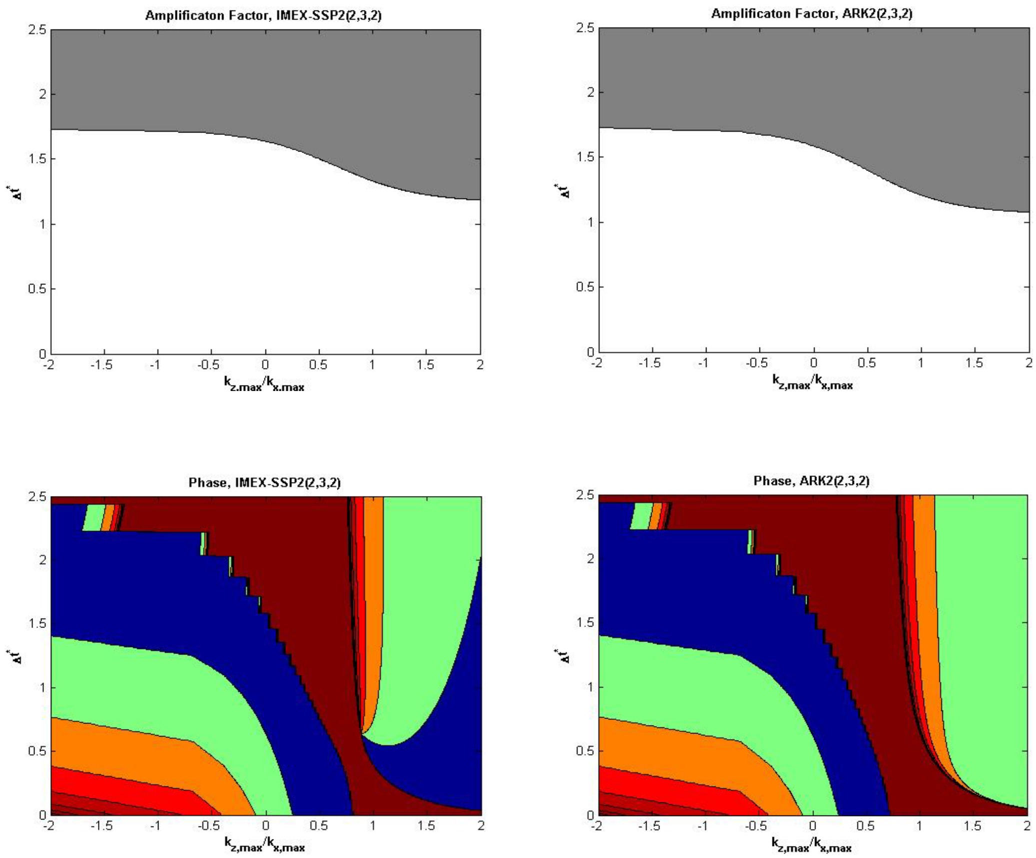


Figure 4-5: Two-frequency system, vertical axis represents $\Delta t^* \in [0,2.5]$, horizontal axis represents logarithmic scale ratio of maximum transverse wavenumber to maximum longitudinal one, $\frac{k_{z,max}}{k_{x,max}} \in [10^{-2}, 10^{2}]$ as Lock et al. (2014), first row; amplifiction factor, second row; Phase angle, IMEX-SSP2(2,3,2) (left), ARK2(2,3,2) (right).

4.5.2.2. 2D acoustic wave equations

Acoustic wave equations just support the sound wave propagation. These governing equations are derived from mass and momentum conservation laws with rearranging and neglecting small terms. The equations are written as

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

$$\frac{\partial w}{\partial t} + \frac{\partial p}{\partial z} = 0 \qquad (4\text{-}25)$$

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

As described in Weller et al. (2013), this system can be presented in vector form

$$y_t + Sy_x + Fy_x + Gy_z = 0 \qquad (4\text{-}26)$$

where $y = [u, w, p]^T$ and,

$$S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ c_s^2 & 0 & 0 \end{pmatrix}, G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & c_s^2 & 0 \end{pmatrix} \qquad (4\text{-}27)$$

In HEVI ufpref form, the S and F terms are treated explicitly and G term is calculated implicitly. Applying the general formulation of IMEX Runge-Kutta scheme (4-5), one obtains

$$\vec{Y}^{(j)} = \vec{y}^n - ik_x \Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} (S + F) \vec{Y}^{(k)} - ik_z \Delta t \sum_{k=1}^{j} a_{jk} G \vec{Y}^{(k)}, \qquad j = 1, \dots, \nu$$

$$(4\text{-}28)$$

$$\vec{y}^{n+1} = \vec{y}^n - ik_x \Delta t \sum_{j=1}^{\nu} \tilde{b}_j (S + F) \vec{Y}^{(j)} - ik_z \Delta t \sum_{j=1}^{\nu} b_j G \vec{Y}^{(j)},$$

Substituting the solutions of the form $\vec{y} = \vec{\hat{y}}(t) \exp(ik_x x + ik_z z)$ results,

$$\vec{Y}^{(j)} = \vec{\hat{y}}^n - ik_x \Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} (S + F) \vec{\hat{Y}}^{(k)} - ik_z \Delta t \sum_{k=1}^{j} a_{jk} G \vec{\hat{Y}}^{(k)}, \qquad j = 1, \dots, \nu \qquad (4\text{-}29)$$

$$\vec{y}^{n+1} = \vec{y}^{n} - ik_x\Delta t\sum_{j=1}^{\nu}\tilde{b}_j(S+F)\vec{\hat{Y}}^{(j)} - ik_z\Delta t\sum_{j=1}^{\nu}b_jG\vec{\hat{Y}}^{(j)},$$

The complex amplification matrix R is defined such that $\vec{y}^{n+1} = R\vec{y}^{n}$ in which R is a 3×3 matrix and similarly, the complex amplification matrices $R^{(j)}$ satisfy $\vec{Y}^{j} = R^{(j)}\vec{y}^{n}$. Substitution into (4-27) yields,

$$R^{(j)} = I - ik_x\Delta t\sum_{k=1}^{j-1}\tilde{a}_{jk}(S+F)R^{(k)} - ik_z\Delta t\sum_{k=1}^{j}a_{jk}GR^{(k)}, \qquad j = 1, \dots, \nu$$

$$R = I - ik_x\Delta t\sum_{j=1}^{\nu}\tilde{b}_j(S+F)R^{(j)} - ik_z\Delta t\sum_{j=1}^{\nu}b_jGR^{(j)}, \tag{4-30}$$

where I is the identity matrix. The amplification matrix for the HEVI ufpreb form is the same as (28) but the F term is moved into the implicit part. In Figure 6, the minimum and maximum eigenvalues of the amplification matrix, R, for acoustic wave equations are shown. For consistency of notation with Weller et al. (2013), the Courant numbers $c_sk_x\Delta t$ and $c_sk_z\Delta t$ associated with the horizontally and vertically propagating acoustic waves, are referred to as $s\Delta t$ and $f\Delta t$ respectively.

Figure B-6-a compares the performance of IMEX-SSP2(2,3,2) and ARK2(2,3,2) for HEVI ufpref form. The maximum eigenvalue fields show more stable region for the proposed scheme, especially, for large $\frac{f}{s}$ values. The contours for minimum eigenvalues, which indicate the greatest damping results of schemes, show that the IMEX-SSP2(2,3,2) scheme is less damping than ARK2(2,3,2). This is more clear in region with small $s\Delta t$ values and for different ranges of $f\Delta t$, which is the case in the atmospheric applications. It is seen from Figure B-6-b that the two IMEX schemes have almost the same performance in HEVI ufpreb form. Both schemes produce one as the order of magnitude for maximum eigenvalue and the minimum eigenvalue fields are almost the same.

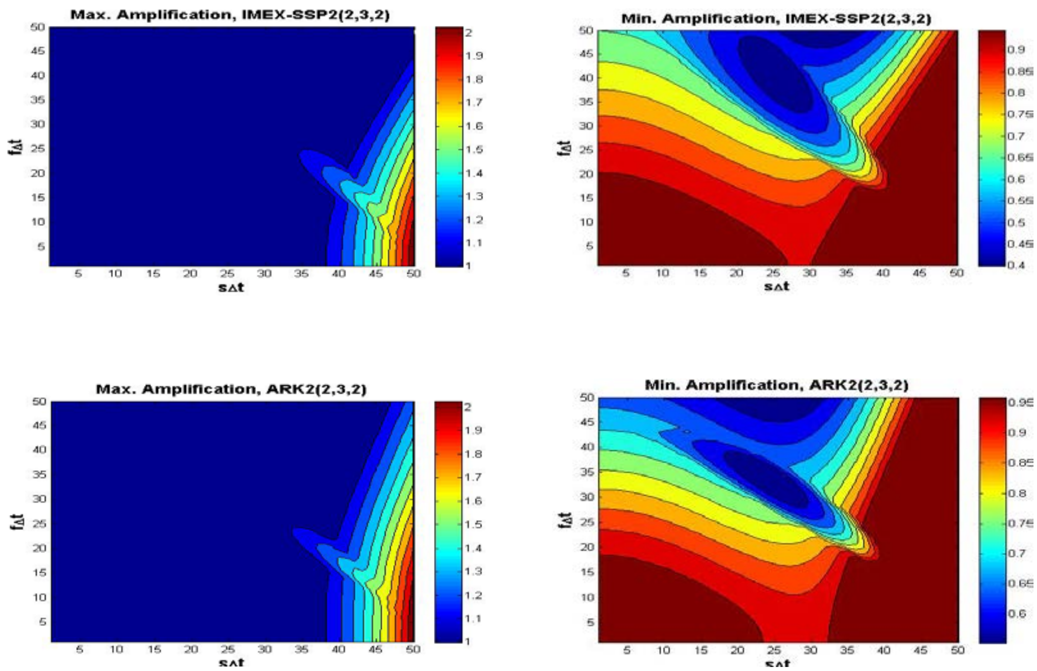


Figure 4-6-a: Maximum (left), minimum (right) eigenvalues of amplification factor, acoustic wave equations, HEVI ufpref form: first row IMEX-SSP2(2,3,2), second row ARK2(2,3,2).

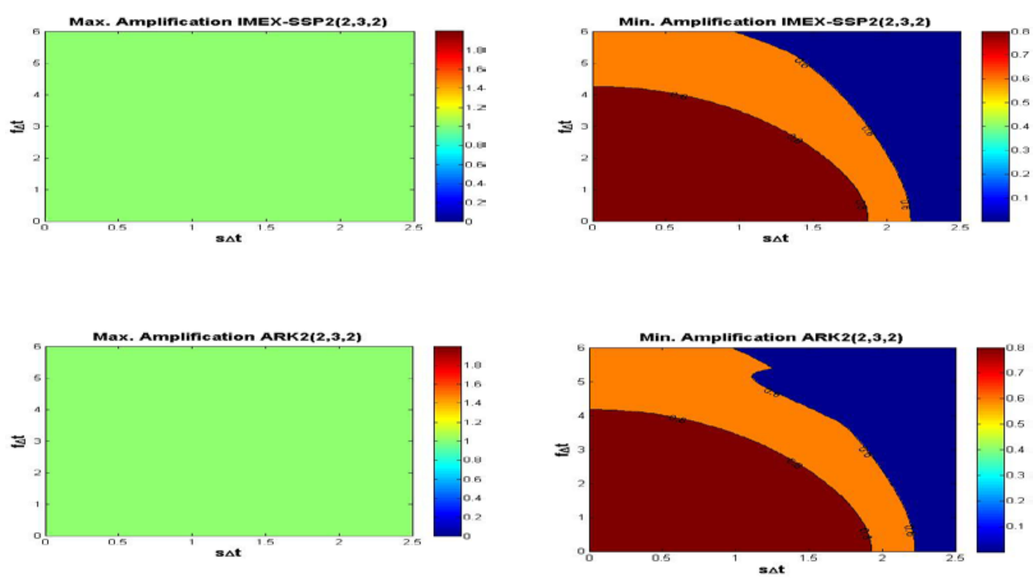


Figure 4-6-b: Maximum (left), minimum (right) eigenvalue of the amplification factor, acoustic wave equations, HEVI ufpreb form: first row IMEX-SSP2(2,3,2), second row ARK2(2,3,2).

4.5.2.3. Compressible Boussinesq equations

For linear stability analysis of compressible Boussinesq equations in semi-implicit form, we refer to the governing equations mentioned in Section 2 but in the linear form:

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)u + \frac{\partial p}{\partial x} = 0$$

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)w + \frac{\partial p}{\partial z} = b$$

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)b + N^2 w = 0 \tag{4-31}$$

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)p + c_s^2(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}) = 0$$

Here, we calculate the terms responsible for sound waves propagation, i.e. pressure gradients and divergence term, implicitly and advection terms explicitly. As discussed in Durran and Blossey (2012), the buoyancy terms could be integrated explicitly or implicitly.

4.5.2.3.1 Semi-implicit buoyancy explicit

The equation (B-29) is rearranged in vector form as

$$\vec{\tilde{y}}_t + \tilde{S}_1\vec{\tilde{y}}_x + \tilde{F}_1\vec{\tilde{y}}_x + \tilde{F}_2\vec{\tilde{y}}_z + \tilde{S}_2\vec{\tilde{y}} = 0 \tag{4-32}$$

where $\vec{\tilde{y}} = [u, w, b, p]^T$, $\tilde{S}_1$and $\tilde{S}_2$ will be treated explicitly while $\tilde{F}_1$and $\tilde{F}_2$ are integrated implicitly.

$$\tilde{S}_1 = \begin{pmatrix} U & 0 & 0 & 0 \\ 0 & U & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & U \end{pmatrix} \quad \tilde{S}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & N^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{F}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ c_s^2 & 0 & 0 & 0 \end{pmatrix} \quad \tilde{F}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & c_s^2 & 0 & 0 \end{pmatrix} \tag{4-33}$$

As described in Durran (2010, p. 14), to demonstrate that the preceding system is hyperbolic, one must perform a change of variables as $y = B\tilde{y}$ that renders $\tilde{S}_1$, $\tilde{F}_1$ and $\tilde{F}_2$ symmetric. The system (30) is then reformed as

$$\vec{y}_t + S_1\vec{y}_x + F_1\vec{y}_x + F_2\vec{y}_z + S_2\vec{y} = 0 \tag{4-34}$$

in which $\vec{y} = [u, w, \frac{b}{c_s}, \frac{p}{c_s}]^T$ and

$$S_1 = \begin{pmatrix} U & 0 & 0 & 0 \\ 0 & U & 0 & 0 \\ 0 & 0 & U & 0 \\ 0 & 0 & 0 & U \end{pmatrix} \quad S_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -c_s & 0 \\ 0 & c_s\frac{N^2}{c_s^2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$F_1 = \begin{pmatrix} 0 & 0 & 0 & c_s \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ c_s & 0 & 0 & 0 \end{pmatrix} \quad F_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_s \\ 0 & 0 & 0 & 0 \\ 0 & c_s & 0 & 0 \end{pmatrix} \tag{4-35}$$

We consider von Neumann linear stability analysis just for time discretization. Hence, we assume continuous spatial derivatives. The solutions admit the form $y = \hat{y}(t)\exp(ik_x x + ik_z z)$ which results in $\vec{y}_x = ik_x\vec{\hat{y}}$ and $\vec{y}_z = ik_z\vec{\hat{y}}$. Substitution in (4-32) and applying the same process as acoustic wave equations in Section 4.5.2.1, allow to define the amplification factors R and $R^{(j)}$ in terms of $c_s k_x \Delta t$, $c_s k_z \Delta t$ and $k_z \Delta t$. In this form of semi-implicit method, the terms including matrices S_1 and S_2 are treated explicitly and F_1 and F_2 are calculated implicitly.

$$R^{(j)} = I - ik_x\Delta t\sum_{k=1}^{j-1}\tilde{a}_{jk}(S_1 + S_2)R^{(k)} - ik_z\Delta t\sum_{k=1}^{j} a_{jk}(F_1 + F_2)R^{(k)}, \qquad j = 1, \dots, \nu$$
$$R = I - ik_x\Delta t\sum_{j=1}^{\nu}\tilde{b}_j(S_1 + S_2)R^{(j)} - ik_z\Delta t\sum_{j=1}^{\nu} b_j(F_1 + F_2)R^{(j)}, \tag{4-36}$$

Following Durran and Blossey (2012), we present the stability function, R, in terms of $\lambda_x = \frac{2\pi}{k_x}$ and Δt in the fixed $\lambda_z = \frac{2\pi}{2000}$ km. In Figure 4-7-a, the maximum eigenvalue of the stability matrix, R, for semi-implicit buoyancy explicit form are shown for both IMEX-SSP2(2,3,2) and ARK2(2,3,2) schemes. The white region shows stability region where the maximum eigenvalue is smaller than one and the gray field illustrates the maximum eigenvalue values larger than one. It can be seen that the new scheme, in a fixed longitudinal wavelength, is stable for larger time steps than ARK2(2,3,2).

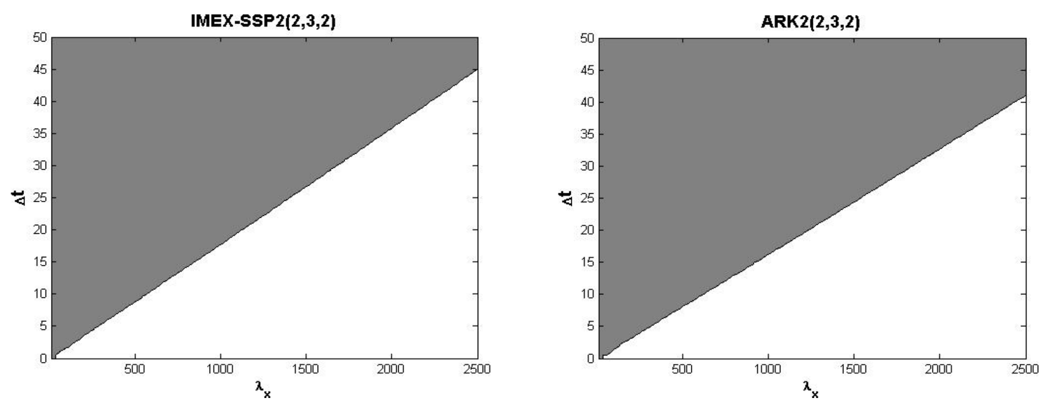


Figure 4-7-a: Maximum eigenvalue of amplification factor, compressible Boussinesq equations, semi-implicit buoyancy explicit form: left IMEX-SSP2(2,3,2), right ARK2(2,3,2). The stable region, i.e. $|R(z)| \leq 1$, is presented in white.

4.5.2.3.2. Semi-implicit buoyancy implicit

In this form of integration, the terms responsible for buoyancy wave propagation are calculated by the implicit scheme. Therefore, with reference to (4-34), we just move the S_2 matrix to implicit integration term,

$$R^{(j)} = I - ik_x\Delta t \sum_{k=1}^{j-1} \tilde{a}_{jk} S_1 R^{(k)} - ik_z\Delta t \sum_{k=1}^{j} a_{jk}(F_1 + F_2 + S_2)R^{(k)}, \qquad j = 1, \dots, \nu$$

$$R = I - ik_x\Delta t \sum_{j=1}^{\nu} \tilde{b}_j S_1 R^{(j)} - ik_z\Delta t \sum_{j=1}^{\nu} b_j(F_1 + F_2 + S_2)R^{(j)}, \tag{4-37}$$

Here again, the new proposed scheme performs better than ARK2(2,3,2). Surprisingly, comparing Figures 4-7-a and 4-7-b show no difference between buoyancy explicit and implicit forms. However, it will be shown in the test cases that treating buoyancy terms with implicit scheme can increase the maximum stable time step significantly.

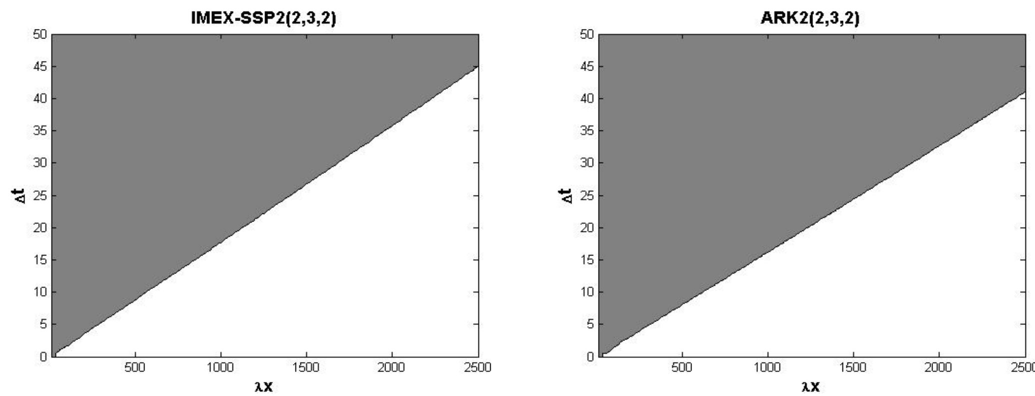


Figure 4-7-b: Maximum eigenvalue of amplification factor, compressible Boussinesq equations, semi-implicit buoyancy implicit form: left IMEX-SSP2(2,3,2), right ARK2(2,3,2). The stable region, i.e. $|R(z)| \leq 1$, is presented in white.

4.6. Numerical test cases

In this section, the performance of the IMEX-SSP2(2,3,2) scheme is examined in calculation of different problems. First, the ability of the proposed scheme is shown in solving stiff problems by applying it to the van der Pol equation. Afterwards, the hydrostatic test case of Durran and Blossey (2012) will be evaluated in solving the compressible Boussinesq equations with semi-implicit and HEVI forms.

4.6.1. van der Pol equation

As a model equation for stiff ODEs, we consider the Singular Perturbation Problem (SPP),

$$y' = f(y, z), \ \ \varepsilon z' = g(y, z), \ \ 0 < \varepsilon \ll 1 \qquad (4\text{-}38)$$

where f and g are smooth functions and initial values y(0) and z(0) are assumed to admit a smooth solution. As the second ODE in (4-38) conveys the stiffness parameter, ε, the implicit part of IMEX schemes treats this equation and explicit part calculates the first ODE. The more detail on Differential Algebraic Equations (DAE) is referred to literature, e.g. Boscarino (2007), Kværnø (2004) and Hairer et al. (1989). The van der Pol equation, with taken initial values from Kværnø (2004), is written as:

$$f(y,z) = z, \qquad g(y,z) = (1-y^2)z - y$$
$$y(0) = 2, \qquad z(0) = -\frac{2}{3} + \frac{10}{81}\varepsilon + \frac{292}{2187}\varepsilon^2 \tag{4-39}$$

in which stiffness parameter, ε, is set to 10^{-6}. Figure 4-8 shows the differences between calculated data, at time t_{end}=0.3 sec, by both IMEX schemes and the data generated by MATLAB's ode15s as the reference solution. Note that the tolerance for calculation of the reference solution was set as 10^{-15}. The l_2 norm of errors are shown against different time step sizes (h), ranged from 0.3 sec to 0.3×10^{-7} sec, in a log-log plot. The left row of Figure 4-8 shows that both IMEX schemes can capture the analytical solution of y-component equation with assigned second order of accuracy.

In the discretized formula for calculation of z-component, time step, h, and stiffness parameter, ε, appear in numerator and denominator respectively. Hence, the stiff limits will occur when $h \gg \varepsilon$. According to Kværnø (2004), for the given initial values, the stiff limit region, in which the order reduction may occur, was specified with $h \geq 0.1$. As seen in Figure 4-8 (right), the IMEX-SSP2(2,3,2) could preserve the second order of accuracy while ARK2(2,3,2) declines to the first order.

For $h \ll \varepsilon$, which is non stiff region, both schemes remain indifferent in terms of time step increasing. For the transient region which was specified by Kværnø (2004) as $5\times10^{-6} < h < 0.1$, ARK2(2,3,2) can remain second order while the new IMEX scheme reduces to the first order. With reference to Kværnø (2004) and Hairer et al. (1989), as the new scheme is A-stable with stability function's magnitude at infinity less than one but not equal to zero, i.e. $|R(-\infty)| = 0.21$, this new scheme calculates the error magnitude for z-component by order of accuracy one. However, ARK2(2,3,2), due to L-stability property and the fact that $|R(-\infty)| = 0$, shows second order of accuracy. As Kværnø (2004) indicated as well, low order is not the worst case in this interval, but propagation of the errors from the last solutions must remain well below the tolerances. It can be seen that the error for both IMEX schemes remains as low as $O(10^{-5})$.

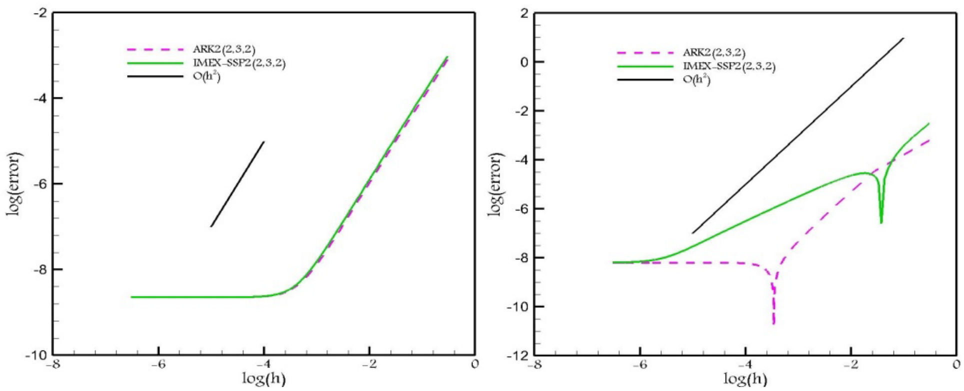


Figure 4-8: The logarithmic l_2 norm of errors for y (left) and z (right) in van der Pol equation, calculated at t_{end}=0.3, SSP-IMEX2(2,3,2) scheme (green solid line) and ARK2(2,3,2) scheme (purple dashed line), against different time step h.

4.6.2. Hydrostatic compressible Boussinesq test case

In this section, the results of solving the hydrostatic compressible Boussinesq eqautions is presented which was explained by Durran and Blossey (2012) as a nonlinear example for atmospheric application, using HEVI ufpref and ufpreb forms as well as semi-implicit buoyancy explicit and implicit. The calculation domain is considered two-dimensional and Rossby waves are neglected to simplify the calculations. Following Durran and Blossey (2012), the explicit streamfunction gradients are imposed in the momentum equations to enforce the flow. Hence the governing equation are:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} + \frac{\partial p}{\partial x} = -\frac{\partial \psi}{\partial z}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} + \frac{\partial p}{\partial z} = b + \frac{\partial \psi}{\partial x} \qquad (4\text{-}40)$$

$$\frac{\partial b}{\partial t} + u\frac{\partial b}{\partial x} + w\frac{\partial b}{\partial z} = -N^2 w$$

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

where the prognostic variables are $y=(u,w,b,p)^T$ and ψ is the streamfunction with the following profile,

$$\psi = \psi_0 \frac{\pi x}{L_x}\sin(\omega t)\exp[-\left(\frac{\pi x}{L_x}\right)^2 - \left(\frac{\pi z}{L_z}\right)^2](m^2 s^{-1}) \tag{4-41}$$

The parameters are defined as $N=0.02s^{-1}$, $c_s=350\ ms^{-1}$, $\omega=1.25\times10^{-4}\ s^{-1}$, $L_x=160$ km, $L_z=10$ km and $\psi_0=10\ m^2s^{-1}$. The calculation domain is defined as z ϵ [-5,5] km and x ϵ [-6000,6000] km. A uniform grid is employed with $\Delta x = 10$ km and $\Delta z = 250$ m. Calculations are performed on a staggered mesh where p, pressure, is located in the cell center. The velocities, u and w, and buoyancy, b, are located on the cell faces as $u_n = \mathbf{u}.\hat{S}_f$ and $b = \mathbf{b}\hat{g}.\hat{S}_f$ where $\hat{S}_f$ is the outward normal vector of the cell faces. The centeral discretization and avarageing are employed here as Durran and Blossey (2012),

$$\delta_{nx}f(x) = \frac{f(x + n\,\Delta x/2) - f(x - n\,\Delta x/2)}{n\Delta x}$$

$$\langle f(x)\rangle^{nx} = \frac{f(x + n\,\Delta x/2) + f(x - n\,\Delta x/2)}{2} \tag{4-42}$$

The governing equations are discretized in the form:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\delta_{2x}(u^2) + \langle\langle w\rangle^x \delta_z u\rangle^z + \delta_x p = -\frac{\partial \psi}{\partial z} - K[(\Delta x\delta_x)^2 + (\Delta z\delta_z)^2]u$$

$$\frac{\partial w}{\partial t} + \langle\langle u\rangle^z \delta_x w\rangle^x + \frac{1}{2}\delta_{2z}(w^2) + \delta_z p = b + \frac{\partial \psi}{\partial x} - K[(\Delta x\delta_x)^2 + (\Delta z\delta_z)^2]w \tag{4-43}$$

$$\frac{\partial b}{\partial t} + \langle\langle u\rangle^z \delta_x b\rangle^x + \langle\langle w\rangle^z \delta_z b\rangle^z = -N^2 w - K[(\Delta x\delta_x)^2 + (\Delta z\delta_z)^2]b$$

$$\frac{\partial p}{\partial t} + \langle u\delta_x p\rangle^x + \langle w\delta_z p\rangle^z + c_s^2(\delta_x u + \delta_z w) = 0$$

Here, turbulent mixing is simulated through the fourth-derivative hyperdiffusion terms with the value of $K = 1.17 \times 10 - 5\ s^{-1}$. This hyperdiffusion term prevents nonlinear instability (Durran and Blossey, 2012) and it will be treated explicitly in the calculation process.

The compressible Boussinesq equations are combined in a Helmholtz equation in one variable. The Helmholtz problem is obtained by temporally discretizing the buoyancy equation and substituting b into the momentum equations. Then the momentum equations are temporally discretized and substituted into the pressure equation. The formulation at each RK sub-stage is fully explained in Weller et al. (2013).

the comparison is made between extracted results and those obtained by the standard explicit Runge-Kutta 4^{th} order on the same grid but with a very small time step h=0.5, as a reference solution. The errors are monitored at $t=1.2\times10^5$s. The initial conditions are zero vertical velocity, zero pressure perturbations, zero buoyancy perturbations and horizontal velocity with the profile $u = 5 + z + 0.4(5 - z)(5 + z)\ ms - 1$. The x boundaries are assumed periodic and the z boundaries are free slip with b=0 and zero vertical pressure gradient.

In order to validate the numerical approach, Figure 4-9 shows the horizontal velocity contours generated by the fully explicit 4^{th} order Runge-Kutta scheme with Δt=0.5 s. In this figure, IMEX-SSP2(2,3,2) and ARK2(2,3,2) are compared in the semi-implicit buoyancy implicit case as well. ARK2(2,3,2) produces the velocity contours with maximum stable time step, i.e., Δt=100 s, while Δt=150 s is used for IMEX-SSP2(2,3,2). It should be noted that the new IMEX scheme shows some oscillatory behavior in the calculation of the horizontal contours for maximum stable time step, Δt=170 s. Comparing the reference solution to the same contours obtained by Weller et al. (2013), one observes that our calculated contours are the same but not identical, which is not surprising because the spatial discretization methods are not identical. Moreover, one may find that the new scheme capture the reference solution more accurately than ARK2(2,3,2).

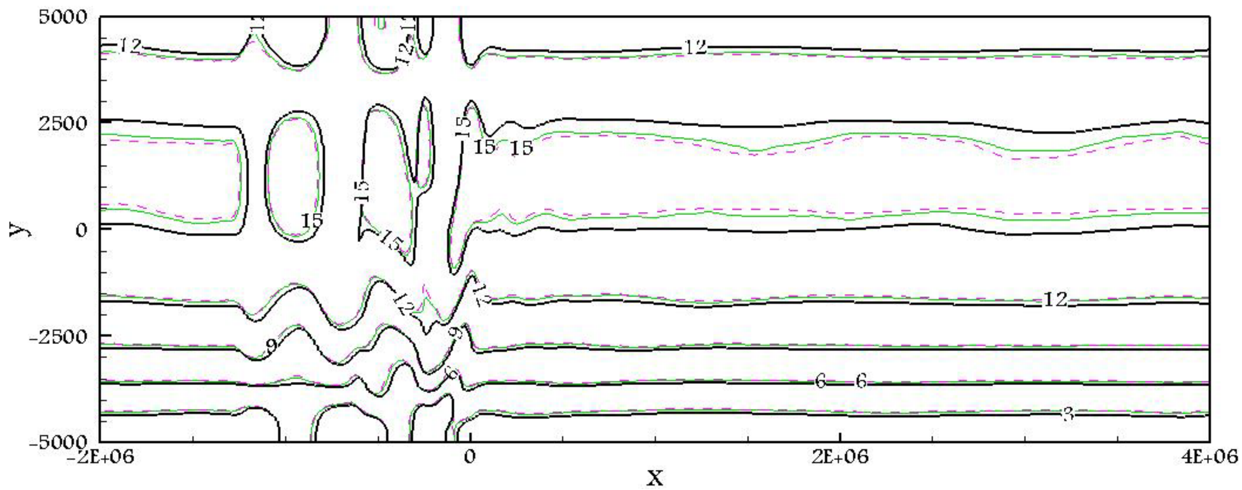


Figure 4-9: Horizontal velocity contours for hydrostatic compressible Boussinesq equations after 33.3 hr, fully explicit fourth-order Runge–Kutta with time step h=0.5 s (black), IMEX-SSP2(2,3,2) with time step h=150 s (green) and ARK2(2,3,2) with time step h=100 s (purple).

Figure 4-10 presents the RMS difference of buoyancy produced by both IMEX schemes and buoyancy of the fully explicit 4^{th} order Runge–Kutta scheme, b_r, in logarithmic scale. The errors, normalised by b_r, are calculated for different HEVI and semi-implicit forms of integrations. The schemes are tested at a variety of time-step sizes up to the maximum stable time-step. As Figure 4-8 illustrates, except for HEVI ufpref form in which ARK2(2,3,2) produces a little more accurate result, the proposed scheme performs better for different forms of integration. Moreover, the new scheme increases the maximum stable time step as indicated in Table 4-4.

Another good test case proposed by Weller et al. (2013) is the simulation of the incompressible flow. As mentioned by Weller et al. (2013), inaccurate simulation of incompressible flow, $\nabla.u=0$, may lead to spurious convergence on the predictions when other equations like moisture advection are included.

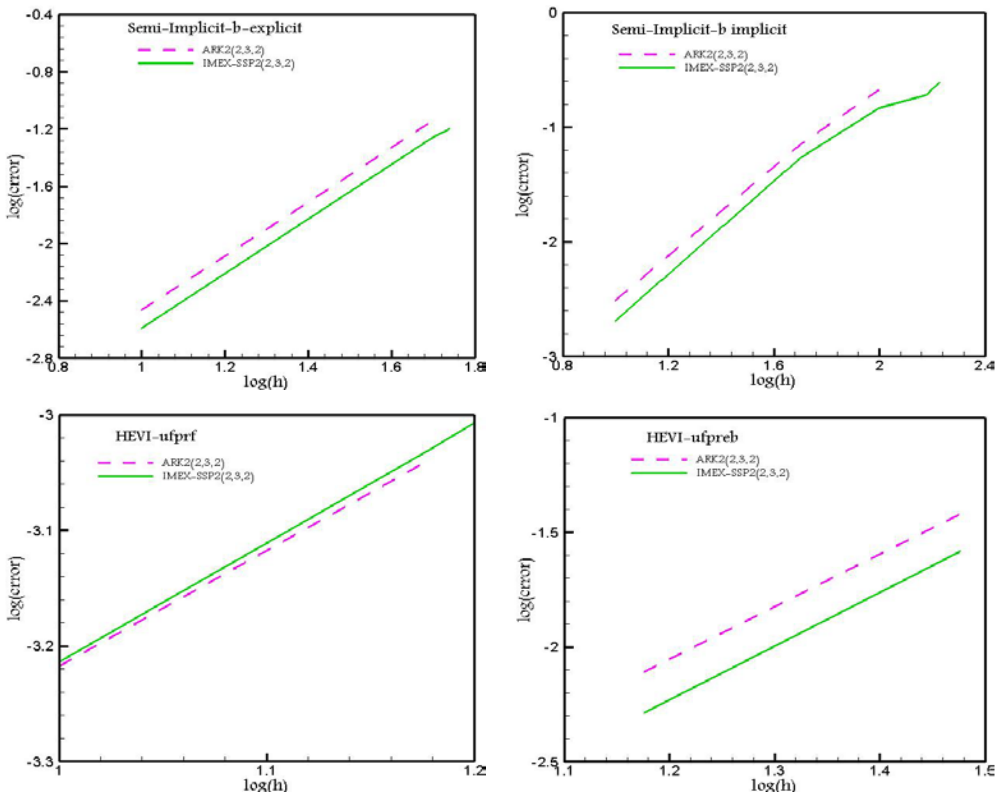


Figure 4-10: Normalised RMS buoyancy errors compared to standard Runge-Kutta fully explicit 4th order with Δt=0.5, hydrostatic compressible Boussinesq equations, semi-implicit form (first row), HEVI form (second row), IMEX-SSP2(2,3,2) (solid green line), ARK2(2,3,2) (dashed purple line).

Therefore, the hydrostatic compressible Boussinesq test case is solved on a coarser grid (100 km × 500 m instead of 10 km × 250 m) in order to avoid resolving the acoustic waves which generate divergent velocity and to allow much longer time-steps.

Table 4-4: Maximum stable Δt (sec), Different Forms of Integration, Compressible Boussinesq Equations, IMEX-SSP2(2,3,2) (first row), ARK2(2,3,2) (second row)

	semi-implicit-b-explicit	semi-implicit-b-implicit	HEVI-ufpref	HEVI-ufpreb
IMEX-SSP2(2,3,2)	55	170	20	30
ARK2(2,3,2)	50	100	15	30

The normalised RMS error for ∇.u=0 is shown in Figure 4-11. Here the reference solution is again obtained using a fully explicit 4th order standard Runge-Kutta scheme with time step h=1.0 sec. In this test case, the proposed scheme behaves more accurately than ARK2(2,3,2), except for semi-implicit form when buoyancy terms is calculated explicitly. However, as indicated by Durran and Blossey (2012), in hydrostatic approximation, it is not practical to treat buoyancy terms explicitly because in hydrostatic approximation gravity waves travel much faster than in nonhydrostatic case. In this test case the new scheme again demonstarates more reliability than ARK2(2,3,2). Table 4-5 shows the maximum stable time step size for both IMEX schemes for incompressible flow in different form of integrations. As shown, the proposed scheme leads to higher maximum time step sizes than ARK2(2,3,2).

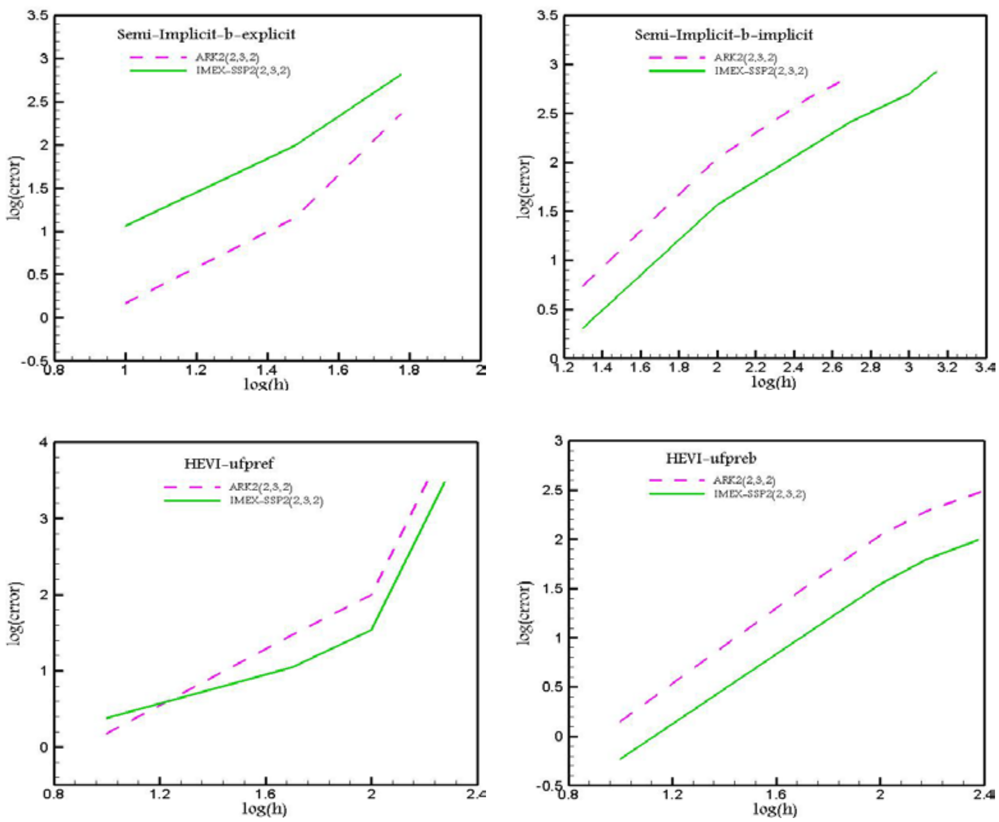


Figure 4-11: Normalised RMS errors in ∇.u in comparison with standard Runge-Kutta fully explicit 4th order with Δt=1.0, hydrostatic compressible Boussinesq equations, semi-implicit form (first row), HEVI form (second row), IMEX-SSP2(2,3,2) (solid green line), ARK2(2,3,2) (dashed purple line).

Table 4-5: Maximum stable Δt (s), different forms of integration, incompressible flow simulation, IMEX-SSP2(2,3,2) (first row), ARK2(2,3,2) (second row)

	semi-implicit-b-explicit	semi-implicit-b-implicit	HEVI-ufpref	HEVI-ufpreb
IMEX-SSP2(2,3,2)	60	1400	190	240
ARK2(2,3,2)	60	500	170	240

4.7. Conclusion

We have proposed an optimized stable and accurate Runge-Kutta IMEX scheme for atmospheric applications. The implicit part is applied to insignificant fast frequency waves in order to circumvent the stability restriction and the explicit part is used for the terms that should be calculated accurately. This new IMEX scheme was constructed by three-stage and second-order explicit and implicit components. The three-stage explicit part enjoys zero error constant and stability region intersection with imaginary axis which is not much smaller than maximum achievable value. This explicit scheme extends stability region along negative real axis further than the explicit scheme which has the maximum value along the imaginary axis. The implicit part preserves A-stability property as well as it satisfies stiffly accurate and second-stage order condition. These characteristics were found to be fundamental in stiff limits. Both explicit and implicit parts of the new IMEX scheme hold the SSP property with the region of absolute monotonicity much larger than the most recent scheme in the literature, ARK2(2,3,2). The von Neumann stability analysis was conducted following Durran and Blossey (2012). It was shown that the new proposed scheme extends the stability region for compressible Boussinesq equations in semi-implicit form when gravity waves are treated implicitly and explicitly. The stability analysis was performed for acoustic wave equations similar to Weller et al. (2013). In this test case, the IMEX-SSP2(2,3,2) improved the stability region for both HEVI ufpref and ufpreb forms. It has been shown that this new scheme obtains its true order of accuracy in solving stiff problems such as van der Pol equation. As it was suggested by Durran and Blossey (2012) and Weller et al. (2013), this new IMEX scheme was employed to simulate the compressible hydrostatic Boussinesq equations in different forms of integration including semi-implicit and HEVI forms. As it was demonstrated, the IMEX-SSP2(2,3,2) could enhance the maximum stable time step for different forms of integration in comparison to ARK2(2,3,2). This improvment also includes the accuracy; the difference between bouyancy field calculated by both IMEX schemes and by reference solution indicated that the new proposed scheme, compared to ARK2(2,3,2), decreased the level of bouyancy error. The incompressible flow was also simulated as it was recommended by Weller et al. (2013). The accurate results for this test case is important in order to correctly simulate other equations like moisture advcetion (Weller et al. 2013). Here, the new proposed scheme could improve the defined error in most cases.

Appendix A. Construction of Stability Functions

In this appendix, the details of stability function formulas and properties explained in section B.5.1 for both explicit and implicit components are provided. The new proposed IMEX scheme is based on explicit and implicit parts. The explicit scheme is three-stage second-order and implicit component is an A-stable two-stage scheme.

4.A.1 Explicit scheme

The general stability function formula was provided by Hairer and Wanner (1996),

$$R_{\tilde{A}}(z) = 1 + z\tilde{b}^T(I - z\tilde{A})^{-1}e \tag{4-44}$$

in which the term inside parentheses can be expanded by Taylor series,

$$(I - z\tilde{A})^{-1} = \sum_{n=0}^{\infty}(z\tilde{A})^n \tag{4-45}$$

The proposed explicit scheme has a strictly lower triangular $\tilde{A}$ matrix with 3×3 components,

$$\tilde{A} = \begin{pmatrix} 0 & 0 & 0 \\ \tilde{a}_{21} & 0 & 0 \\ \tilde{a}_{31} & \tilde{a}_{32} & 0 \end{pmatrix} \tag{4-46}$$

Consequently, the summation in (4-45) will be terminated after $n = 2$ because $\tilde{A}^3 = 0$. Thus, the stability function, (4-44), will be simplified as,

$$R_{\tilde{A}}(z) = 1 + z\tilde{b}^T e + z^2\tilde{b}^T\tilde{A}e + z^3\tilde{b}^T\tilde{A}^2 e \tag{4-47}$$

The coefficients in the right hand side are the order conditions explained in section 4.3.1. Therefore, the explicit stability function will be

$$R_{\widetilde{A}}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} \tag{4-48}$$

According to Burrage (1978), the error constant for schemes of the order of accuracy p is defined as,

$$E = \left|\tilde{b}^T\widetilde{A}^p e - \frac{1}{(p+1)!}\right| \tag{4-49}$$

The proposed explicit scheme has the order of accuracy two, $p = 2$. One can verify that the error constant for the new explicit stability function is zero.

It is desirable to find the intersection of the above explicit stability function, (4-48), with the imaginary axis. It can be obtained by seeking ω such that $|R_{\widetilde{A}}(i\omega)| \leq 1$. In our case, this leads to,

$$|R_{\widetilde{A}}(i\omega)| = \sqrt{(1 - \frac{\omega^2}{2})^2 + (\omega - \frac{\omega^3}{6})^2} \leq 1 \tag{4-50}$$

which admits its maximum solution at $\omega = \sqrt{3}$. The relation $|R_{\widetilde{A}}(-\omega)| \leq 1$ gives the intersection of stability function with negative real axis. The stability function along negative real axis has the form,

$$R_{\widetilde{A}}(-\omega) = 1 - \omega + \frac{\omega^2}{2} - \frac{\omega^3}{6} \tag{4-51}$$

Hence, the maximum value for the equation $|R_{\widetilde{A}}(-\omega)| \leq 1$ will be obtained at $\omega = -2.513$. The other important issue discussed in Section 5.1 is nonnegativity of the stability function on the negative real axis ($R_{\widetilde{A}}(-\omega) \geq 0$). By letting $R_{\widetilde{A}}(-\omega) = 0$, one can obtain the maximum value which is $\omega = -1.596$ for proposed explicit scheme.

4.A.2 Implicit scheme

Another definition for stability function which is more easily implemented for implicit schemes was discussed in Hairer and Wanner (1996),

$$R_A(z) = \frac{\det(I - zA + zeb^T)}{\det(I - zA)} \tag{4-52}$$

in which e represents the vector of ones. Using (4-52), the stability function for new implicit scheme can be obtained as,

$$R_A(z) = \frac{-0.091z^2 + 0.391z + 1}{0.0903z^2 - 0.6092z + 1} \tag{4-53}$$

According to definition of the error constant, (4-49), the proposed implicit scheme leads to $E \cong 0.048$, while the minimum value as described in Norsett and Wolfbrandt (1977) is $E \cong 0.041$.

Chapter 5

Performance of the IMEX-SSP2(2,3,2) in solving the nonlinear Burger equation using the spherical coordinate system with real initial condition[3]

Abstract

The intention of this unpublished technical note is to examine the performance of the IMEX-SSP2(2,3,2) in solving a different problem, chosen to be the nonlinear Burger equation in spherical coordinate system with a realistic initial condition. The computational domain was the North America continent, in which the flow dynamic initially started from January 1st , 2018, taken from International Research Institute (IRI) for Climate and Society. The latitude component of the momentum equation was solved using the implicit part of the IMEX and longitude direction was treated explicitly. As shown, the IMEX-SSP2(2,3,2) still effectively leads to improving the accuracy compared to the ARK2(2,3,2). However, the improvement is less significant compared to the compressible Boussinesq test case, provided in Chapter 4. The main reason is the that the purpose of using IMEX schemes is to overcome difficulties associated with the fast and slows waves, but the Burger equation is just an advection flow.

5.1. Introduction

In order to validate the performance of the IMEX-SSP2(2,3,2) so that the optimization of the dissipation and dispersion errors is depend neither on the coordinate system nor on the initial condition, the inviscid nonlinear Burger equation, in the spherical coordinate, was solved, using a realistic set initial condition. The initial condition, January 1st 2018, has been taken from International Research Institute (IRI) for Climate and Society. The flow field of the North America has been selected, where located in longitude direction, $20^{\mathrm{o}}\mathrm{W} \leq \theta \leq 170^{\mathrm{o}}\mathrm{W}$, and latitude direction, $10^{\mathrm{o}}\mathrm{N} \leq \varphi \leq 80^{\mathrm{o}}\mathrm{N}$ (figure 1) at the surface pressure level, i.e. $\mathrm{P} = 1000$ mb. Note that the boundary conditions were assumed periodic, in which the flow pattern was

[3] Arman Rokhzadi, Abdolmajid Mohammadian, unpublished technical note, 2018.

assumed to repeat within the inlet and outlet of the domain. The purpose of this test is just to examine the performance of the IMEX-SSP2(2,3,2) performances by applying to a practical case with a realistic initial condition. However, the inaccuracy associated with the boundary conditions as well as in the measurement of the initial condition may influence the results.

Figure 5.1: The North America continent (https://www.worldatlas.com/webimage/countrys/na.htm)

5.2. Governing equations

The two dimensional momentum equations in longitude-latitude directions are as follows,

$$\frac{\partial u_\theta}{\partial t} + \frac{u_\theta}{r\cos\varphi}\frac{\partial u_\theta}{\partial \theta} + \frac{u_\varphi}{r}\frac{\partial u_\theta}{\partial \varphi} + \frac{u_\theta u_\varphi \tan\varphi}{r} = 0$$

$$\frac{\partial u_\varphi}{\partial t} + \frac{u_\theta}{r\cos\varphi}\frac{\partial u_\varphi}{\partial \theta} + \frac{u_\varphi}{r}\frac{\partial u_\varphi}{\partial \varphi} - \frac{{u_\theta}^2 \tan\varphi}{r} = 0$$

in which the parameter r, the radius of the earth, is assumed as $r = 6371\text{km}$.

5.3. Numerical calculations

The central discretization and averaging are employed here as Durran and Blossey (2012),

$$\delta_{nx} f(x) = \frac{f(x + n\,\Delta x/2) - f(x - n\,\Delta x/2)}{n\Delta x}$$

$$\langle f(x) \rangle^{nx} = \frac{f(x + n\,\Delta x/2) + f(x - n\,\Delta x/2)}{2}$$

in which x represents the both longitude and latitude directions and f represents any velocity components. The governing equations are discretized in the form:

$$\frac{\partial u_\theta}{\partial t}+\frac{1}{2r\cos\varphi}\delta_{2\theta}(u_\theta{}^2)+\frac{1}{r}\langle u_\varphi\delta_\varphi u_\theta\rangle^\varphi+\frac{u_\theta u_\varphi \tan\varphi}{r}=0$$

$$\frac{\partial u_\varphi}{\partial t}+\frac{1}{r\cos\varphi}\langle u_\theta\delta_\theta u_\varphi\rangle^\varphi+\frac{1}{2r}\delta_{2\varphi}\left(u_\varphi{}^2\right)-\frac{u_\theta{}^2\tan\varphi}{r}=0$$

After spatial discretization, the IMEX Runge-Kutta scheme was applied to temporally integrate the equations. In this regards, those terms corresponding to latitude direction, φ, were calculated by implicit part of the IMEX scheme and the longitude terms were treated explicitly.

5.4. Results and discussions

Shown in figure 5.2 is the flow field transported by nonlinear Burger equation within one day. This flow field was calculated, employing the IMEX-SSP2(2,3,2) but with a very small time step size, i.e. $\Delta t = 3$sec. The solutions obtained by this time step for both IMEX schemes were considered as the reference solutions.

In order to ensure that both IMEX schemes work properly, figure 5.3 shows the same flow field, reproduced by the IMEX schemes but with a larger time step size, i.e. $\Delta t = 3\times10^4$sec. Clearly both IMEX schemes with larger time step size can calculate the same flow field as the reference solution.

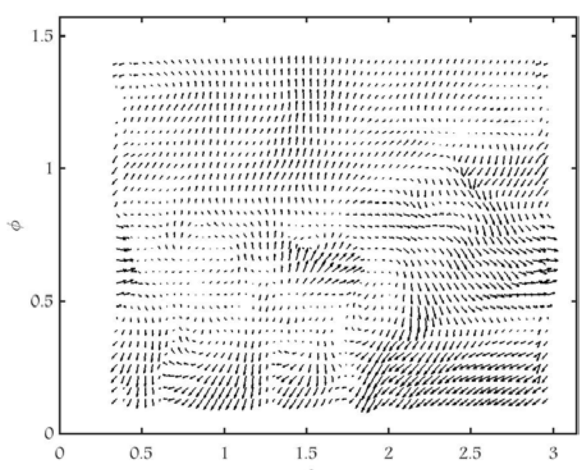


Figure 5.2: The flow field over the North America continent, calculated serving the IMEX-SSP2(2,3,2) with a very small time step size, $\Delta t = 3$sec, (reference solution).

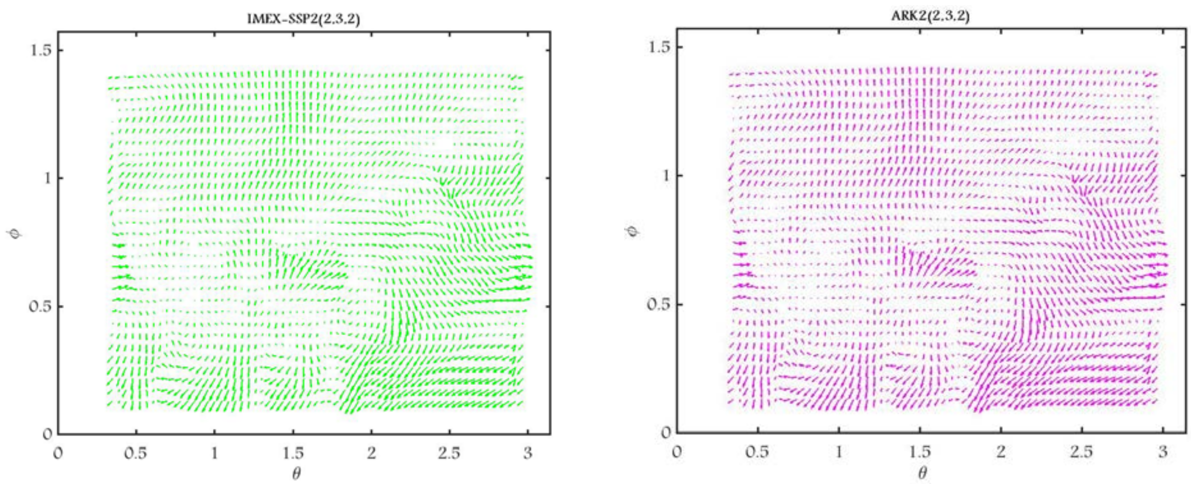


Figure 5.3: The flow field over the North America continent, using time step size $\Delta t = 3 \times 10^4$ sec, the IMEX-SSP2(2,3,2) (left), The ARK2(2,3,2) (right).

In order to observe how the new scheme, IMEX-SSP2(2,3,2), is more accurate than ARK2(2,3,2) in calculating the nonlinear viscous Burger equations, figure 5.4 presents the RMS difference of velocities in both longitudinal and latitudinal directions produced by both IMEX schemes and that of the reference solution in logarithmic scale. It is worth mentioning that the spatial schemes applied to both reference and the real cases are identical with the same grid resolution. Therefore, the numerical error, due to the spatial discretization, is eliminated through subtracting the reference and real solutions. The schemes are tested at a variety of time-step sizes. As figure 5.4 illustrates, the proposed scheme, i.e. IMEX-SSP2(2,3,2), improves the accuracy compared with ARK2(2,3,2), although, this improvement is not significant. The main reason is because only the advection terms were considered in the governing equations. As it is explained before, presence of acoustic waves with fast frequency and gravity waves with slow frequency is the main motivation in employing the IMEX schemes for the numerical calculations.

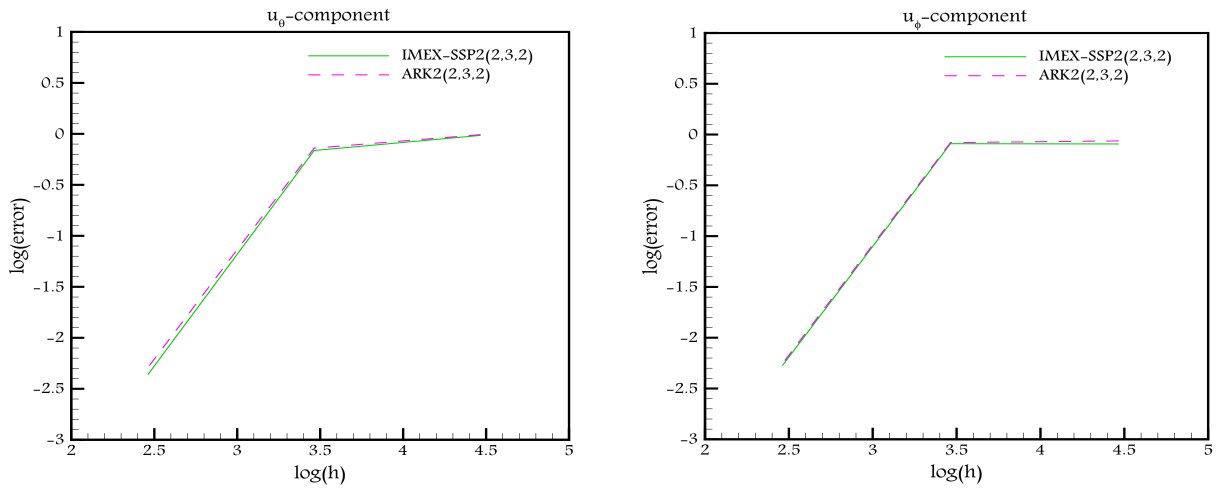


Figure 5.4: RMS errors, defined as the difference between results obtained by both IMEX schemes with large time step size, $\Delta t = 3 \times 10^4$sec, and small time step size, $\Delta t = 3$sec, IMEX-SSP2(2,3,2) (solid green line), ARK2(2,3,2) (dashed purple line), u_θ, longitudinal velocity (left) and u_φ, latitudinal velocity (right).

5.5. Conclusion

Regarding the theory of linear stability analysis, provided in section 2.4 and the above example, it can be concluded that the optimization of the dissipation and dispersion errors is not subject to either the specific coordinate system or initial condition. However, it is worth mentioning that some discrepancies are expected which are due to the linearization of the Jacobian matrix of the source terms (see Eq. 54).

Chapter 6

Optimal non-oscillatory temporal scheme for the advection diffusion reaction equations[4]

Abstract

This paper studies the oscillatory behaviour of the numerical methods in the advection diffusion reaction equations and investigates how time integration schemes can help control and prevent these oscillations. It was found that the nonnegative stability function ensures the non-oscillatory behaviour. Therefore, the linear stability analysis has been performed aiming at maximizing the nonnegative range, associated with the stability function. In this regard, the well-known stability properties, i.e. A- and L-stability properties and the maximum radius of absolute monotonicity, which is the characteristic of the optimized Strong Stability preserving (SSP) property have been investigated, in order to observe how they can affect the stability function in achieving larger nonnegative range. Various test cases, including the compressible Euler, nonlinear viscous Burger equation and linear advection with nonlinear source term have been examined.

It was found that the schemes with optimized SSP methods produce less oscillatory behaviours in the advection dominant flows, while the L-stability property is beneficial in diffusion dominant flows. However, the A-stability property was found to be more effective so that it can provide less oscillation in the practical viscous flows.

Keywords: Advection diffusion reaction; DIRK; Non-oscillatory; Non-negativity; A-stability; SSP

6.1. Introduction

The motivation for this article mainly is the need for reliable implicit schemes in the atmospheric applications, which is a perfect example of the advection diffusion reaction flows. Due to nearly incompressible air flow, the acoustic terms impose stability restriction, while they have less

[4] Arman Rokhzadi, Abdolmajid Mohammadian, under review, 2018.

contribution into the physical accuracy. These terms unnecessarily restrict stability range of the explicit schemes which are responsible to provide physically accurate solutions (Durran, 2010). Many spatial integration methods use flux limiters to avoid the generation of spurious overshoot and undershoot in solutions (Durran, 2010). The time integrators corresponding to such methods is often Forward-Euler (FE) differencing. The Strong Stability Preserving (SSP) Runge-Kutta schemes can further improve the accuracy in time while preserving the benefits of the flux limiters (Durran, 2010). However, SSP Runge-Kutta schemes preserve the monotonic behaviour under limited time-step sizes (Ketcheson et al., 2009).

In the linear stability analysis, using Fourier transform, the stability function, $\mathrm{R(z)}$, is derived based on a complex variable, z. In the literature, A-stability property has been described as a desirable property for temporal schemes (e.g. Hairer and Wanner, 1996). They defined this property such that the magnitude of the numerical stability function, remains less than one, i.e. $|\mathrm{R(z)}| \leq 1$, within the domain of interest, $\mathrm{z} \in \mathbb{C}^{-}$, whereas $\mathbb{C}^{-}$ includes all the complex numbers with negative real part. Crouzeix (1975) derived a formula for the only three-stage fourth order Singly Diagonally Implicit Runge-Kutta (SDIRK) scheme. This scheme does not preserve SSP property as some of the coefficients are negative. Nonnegative coefficients of temporal schemes are necessary for preserving SSP property (Ketcheson et al., 2009 and Higueras et al., 2014). Najafi and Mongeau (2013) developed a three-stage fourth order Runge-Kutta scheme, while minimizing the dissipation and dispersion errors. However, as Nazari et al. (2015) have shown, this scheme declines to lower order of accuracy, since the corresponded coefficients do not satisfy the first-stage order condition, $\mathrm{c_i} = \sum \mathrm{a_{ij}}$ (These parameters will be introduced in Section 2). The three-stage fourth order Runge-Kutta scheme developed by Nazari et al. (2015) does not hold SSP property due to existence of negative coefficients in their scheme. Moreover, its stability function is not analytic in the left half plan, as it is singular in one point. This could be found in the rational formula of the stability function, $\mathrm{R(z)}$, since it obtains one negative root in the denominator.

The atmospheric governing equations may be temporally integrated by Runge-Kutta schemes enjoying L-stability property (Giraldo et al., 2013). This property implies the A-stability and additionally, in the linear stability analysis context, the stability function decays to zero in the infinite limit of the variable, z, i.e. $|\mathrm{R}(-\infty)| = 0$, in order to dampen large errors and to remain stable in the calculation of large terms, provided by large time-step and fine mesh grid sizes. As indicated by Lock et al. (2014) a high resolution computational mesh is of order $\Delta x = 10\text{m}$, which is used to capture very small-scale boundary layer processes. Hence, in discretised numerical formula, very large values may occur due to appearance of the fine grid size values. However, in the physical sense, these values are still limited. Therefore, as shown in this paper, imposing highly restrictive condition on temporal schemes to satisfy L-stability property lead the schemes to a poor behaviour in the sense of stability.

In this investigation, the importance of nonnegative stability function is studied. Higueras et al. (2014) aimed at optimizing the nonnegative region of stability function for three-stage second order DIRK scheme, using parabolic equation. However, as will be shown later, this property is essential through entire domain of interest, i.e. $z \in C^{-}$.

This paper presents the behaviour of DIRK methods with order of accuracy three, obtained within three stages. Since accuracy of simulations is limited by spatial resolution, which commonly is the second order, more highly accurate temporal schemes may not provide further advantages.

First, a brief introduction of the general features of the three-stage third order DIRK schemes, including order conditions, will be provided (Section 6.2). It then moves to the linear stability analysis and general formula of the stability function of DIRK schemes in Section 6.3. The optimization procedure and relevant constraints are presented in Section 6.4. Section 6.5 presents the details of the new schemes and compares to the optimized SSP three-stage third order DIRK scheme, hereafter called SSP(3,3). Performances of the new schemes will be examined in Section 6.6 by solving various test cases so that they include the advection, diffusion and reaction terms. Concluding remarks will complete the study in Section 6.7.

6.2. Three-stage third order DIRK methods

In solving time-dependent Partial Differential Equations (PDEs), generally, the spatial terms are first discretized to obtain a semi-discrete Method of Lines (MOLs). This leads to a system of an Ordinary Differential Equations (ODE) in time, which can be integrated by an ODE solver.

A Runge-Kutta scheme integrates the ODE using selected internal stages between time steps t_n and t_{n+1}, presented by vector $\mathbf{c} = [c_1, c_2, \cdots, c_s]$. The subscript s is the number of internal stages; in this paper $s = 3$. This method calculates ODEs using a linear correlation of the solutions in the internal stages. The corresponding weight coefficients, specified by matrix **A** and vector **b,** and vector **c,** are commonly presented in the Butcher Tableau (Table 6-1).

Table 6-1: Butcher Tableau for s stage DIRK scheme

$\begin{array}{c\|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array}$	$\begin{array}{c\|cccc} c_1 & a_{11} & 0 & 0 & 0 \\ : & a_{21} & a_{22} & 0 & 0 \\ : & : & : & a_{ll} & 0 \\ c_s & a_{s1} & a_{s2} & a_{sl} & a_{ss} \\ \hline & b_1 & b_2 & \ldots & b_s \end{array}$

The advection diffusion reaction equations could be generalized as follows,

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{u}, t)}{\partial \mathbf{x}} = \frac{\partial^2 \mathbf{G}(\mathbf{u}, t)}{\partial^2 \mathbf{x}} + \mathbf{H}(\mathbf{u}, t), \tag{6-1}$$

Here, **u** includes vector of unknown variables and **F** represents vector of the fluxes, **G** is diffusion term and **H** represents the source terms. The general formula of s-stage DIRK schemes is:

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \sum_{i=1}^{s} b_i f(\mathbf{Y}_i, t_n + c_i \Delta t), \\ \mathbf{Y_i} &= \mathbf{u}_n + \Delta t \sum_{j=1}^{i} a_{ij} f(\mathbf{Y_j}, t_n + c_j \Delta t), \end{aligned} \tag{6-2}$$

Therefore, applying the Runge-Kutta integration scheme (6-2) to (6-1) results in:

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n + \Delta t \sum_{i=1}^{s} b_i \left(-\left(\frac{\partial \mathbf{F}(\mathbf{Y_i}, t_n + c_i \Delta t)}{\partial x} \right)_j + \left(\frac{\partial^2 \mathbf{G}(\mathbf{Y_i}, t_n + c_i \Delta t)}{\partial x^2} \right)_j + \mathbf{H}(\mathbf{Y_i}, t_n + c_i \Delta t) \right),$$
$$\mathbf{Y_i} = \mathbf{u}_j^n + \Delta t \sum_{k=1}^{i} a_{ik} \left(-\left(\frac{\partial \mathbf{F}(\mathbf{Y_k}, t_n + c_i \Delta t)}{\partial x} \right)_j + \left(\frac{\partial^2 \mathbf{G}(\mathbf{Y_k}, t_n + c_i \Delta t)}{\partial x^2} \right)_j + \mathbf{H}(\mathbf{Y_k}, t_n + c_k \Delta t) \right), \tag{6-3}$$

in which j represents the spatial discretization index.

To achieve the desired order of accuracy, the coefficients, matrix **A** and vectors **b** and **c**, must satisfy order conditions. The further details of the order conditions could be found in the literature, e.g. Rokhzadi et al. (2018).

6.3. Stability function

In linear stability analysis, it is required to assume that all functions in (6-1) are linearly dependent on the variable **u**. Without losing generality, but to simplify the calculation, only one variable is assumed varying in one spatial dimension. Hence (6-1) is simplified to:

$$\frac{\partial u}{\partial t} = -a \frac{\partial u}{\partial x} + \nu \frac{\partial^2 u}{\partial x^2} + \sigma u, \tag{6-4}$$

in which a, ν and σ are representing the physical coefficients.

After applying the spatial discretization schemes, the semi-discretized ODE will be achieved. In Fourier transform context, the solutions would be sought in the form $u = \hat{u}(t)\exp(ikx)$, in which k represents the wavenumbers, associated with spatial derivative discretization schemes. Substituting the Fourier mode into (6-4) results in:

$$\frac{\partial \hat{u}}{\partial t} = \hat{u}(t) g(k, a, \nu, \sigma), \tag{6-5}$$

This equation is briefly represented as:

$$\frac{\partial \hat{u}}{\partial t} = \lambda \hat{u}, \tag{6-6}$$

in which λ is a complex coefficient containing the parameters attributed to the physics of the fluid flows. Note that in solving the advection problem λ only contains imaginary numbers, which is due to the first order spatial derivative term. However, in the advection diffusion reaction problems λ obtains the real component as well, due to the diffusion and source terms.

Hereafter $\hat{u}$ is replaced with u. The analytical solution for (6-6) is:

$$u(t) = u_0 e^{\lambda t}, \tag{6-7}$$

The analytical stability function, defined as the ratio of the solutions at two successive time steps, i.e. t_{n+1} and t_n, is:

$$R_a = e^{\lambda \Delta t}, \tag{6-8}$$

where $\lambda\Delta t$ is simply called z. Therefore,

$$R_a(z) = e^z, \tag{6-9}$$

Applying the *s*-stage DIRK formula, (6-2), to (6-6), results in;

$$\begin{aligned} u_{n+1} &= u_n + z\sum_{i=1}^{s} b_i Y_i, \\ Y_i &= u_n + z\sum_{j=1}^{i} a_{ij} Y_j, \end{aligned} \tag{6-10}$$

An appropriate formula for the numerical stability function of the DIRK methods, $R_n(z)$, is given by

$$R_n(z) = \frac{\det(I - z\mathbf{A} + z\mathbf{e}\mathbf{b}^T)}{\det(I - z\mathbf{A})}, \tag{6-11}$$

which is obtained by using the Cramer's rule. The reader is referred to Hairer and Wanner (1996) for more details.

From the definition of the stability function, the solutions in two successive time steps must satisfy the following relationship:

$$\frac{u^{(n+1)}}{u^{(n)}} = R_n(z) \tag{6-12}$$

Since the variable, i.e. z, obtains both real and imaginary components, as discussed above, the stability function is a complex function, which can be decomposed as follows;

$$\frac{u^{(n+1)}}{u^{(n)}} = R_r + iR_I \tag{6-13}$$

in which R_r and R_I represent the real and imaginary components.

Clearly, the non-oscillatory behavior implies the same sign for the solutions, i.e. $u^{(n+1)}$ and $u^{(n)}$. The optimized SSP property was originally designed to compute the non-oscillatory solutions for the advection dominant problems, in which the variable z in (6-12) contains the imaginary component only. Therefore, as will be seen later (Section 5-3), in the complex coordinate of the Fourier space, the SSP optimized schemes provide larger non-negative region for the imaginary component of the stability function (R_I). However, in the advection, diffusion and reaction problems in which the variable z has both real and imaginary components, it is essential to further enlarge the non-negative region of both real and imaginary components of the stability function, i.e. R_r and R_I.

It is clear that the numerical stability function for implicit schemes (6-11) is a rational expression, in which both numerator and denominator are polynomials of degree s (three in this paper). Expanding this equation produces a rational polynomial of the order three:

$$\mathrm{R_n(z)} = \mathrm{A}\frac{(\mathrm{z}-\gamma_1)(\mathrm{z}-\gamma_2)(\mathrm{z}-\gamma_3)}{(\mathrm{z}-\frac{1}{a_{11}})(\mathrm{z}-\frac{1}{a_{22}})(\mathrm{z}-\frac{1}{a_{33}})}, \tag{6-14}$$

In this paper, the objective is to maximize the values of the roots of the stability function (14), i.e. γ_1, γ_2 and γ_3, which consequently produce a larger nonnegative stable region. As indicated by Strikwerda (2004) as well, the nonnegative property is directly related to the dissipativity of the spatial schemes and may prevent spurious oscillations.

6.4. Optimization algorithm

The main objective of this paper is to link the nonnegative property of the stability function to the non-oscillatory behaviour of the temporal schemes applied to the advection diffusion reaction problems. In this regard, three significant stability properties, A- and L-stability and SSP optimized properties are also investigated to observe how they may influence the nonnegative property.

In the advection diffusion reaction problems like the atmospheric flows, it is required to use very small grid sizes to capture small scales boundary layer processes (Lock et al., 2014). Using the formula of the Runge-Kutta schemes, (6-3), it could be found that these small grid sizes combined with the temporal scheme's coefficients, including time step resolution, produce very large coefficients. Hence, reliable temporal schemes with high level of stability are required to tolerate these large terms and to allow large stable time steps. This was also indicated by Giraldo et al. (2013) and Durran (2010).

Hairer and Wanner (1996) proved that A-stability property ensures the stability function being analytic in the negative real half plane. This condition implies that all roots of the denominator of the numerical stability function, (6-14), which are the inverse of the main diagonal components of the **A** matrix, must be in the positive real half-plane, i.e. C^{+}.

As indicated by Hairer and Wanner (1996), the other requirement for A-stability property is that the numerical stability function, (6-14), must remain less than or equal one along imaginary axis (I-stability). Furthermore, the A-stability property guarantees that the magnitude of numerical stability function remains less than or equal one, as the variable, i.e. z, obtains infinite values, i.e. $|\mathrm{R}(-\infty)| \leq 1$. This condition implies that magnitude of the A-factor in (6-14) must be at most one.

Therefore, this paper focuses on finding the maximum values for the roots of the stability function, (14), subject to following conditions;

6.4.1. All order conditions must be satisfied.

6.4.2. Positive values for Runge-Kutta scheme's components; this condition ensures to achieve SSP property, as indicated by Ketcheson et al. (2009) and Higueras et al. (2014),

6.4.3. The absolute value of A-factor in (6-14) must be less than or equal one.

The L-stability property requires that the stability function converges to zero while the variable, i.e. z, obtains infinite values, i.e. $|\mathrm{R}(-\infty)| = 0$. In order to show that this property imposes too restrictive conditions on the Runge-Kutta scheme, which seems to be unnecessary in practice, a three-stage third order DIRK scheme with the optimized nonnegative stability function combined with the L-stability property is developed as well.

Since the three-stage DIRK schemes are investigated, the numerator and denominator of stability function in (6-14) are polynomials with leading term of power three. The L-stability condition implies that the order of polynomial in the numerator of stability function must be less than the denominator's degree (Kennedy and Carpenter, 2016). Therefore, in order to achieve the L-stability property, in addition to the conditions 6.4.1 and 6.4.2, an extra constraint was imposed on the optimization algorithm:

- The coefficient of the leading term in the numerator of stability function (6-14) must be zero.

6.5. New optimized three-stage third order DIRK schemes

In this section, more details of the proposed schemes will be provided. First, the coefficients will be presented in the Butcher Tableau. The radius of absolute monotonicity, which measures to what time step sizes the schemes preserve the SSP property, will be compared with SSP(3,3) whose the radius of absolute monotonicity is the maximum for the three stage third order DIRK schemes. Then, the stability functions' formulae will be provided and the comparison will be made in terms of their achieved roots. Afterward, the linear stability analysis will be discussed in ODE and PDE approaches.

6.5.1 The coefficients of DIRK schemes

Two new DIRK schemes, called A-stable and L-stable, with A- and L-stability properties respectively, and with the optimized non-negative stability functions are compared to SSP(3,3).

Both new schemes preserve the SSP property. Tables 6-2 to 6-4 present the coefficients of SSP(3,3) scheme (Ferracina and Spijker, 2008) and the proposed schemes respectively.

6.5.2 Radius of absolute monotonicity

Radius of absolute monotonicity represents the maximum ratio of the time step for the chosen DIRK scheme to the maximum stable time step of the FE method such that monotonicity is preserved (Ketcheson et al., 2009). More details of the definition and associated conditions would be found in literature, e.g. Ketcheson et al. (2009).

Table 6-2: Butcher Tableau for three-stage third order SSP Optimized DIRK scheme (Ferracina and Spijker, 2007)

0.146446609406726	0.146446609406726	0	0
0.5	0.353553390593275	0.146446609406726	0
0.853553390593274	0.353553390593273	0.353553390593273	0.146446609406726
	0.333333333333333	0.333333333333333	0.333333333333333

Table 6-3: Butcher Tableau for three-stage third order A-stable DIRK scheme

0.159359567999120	0.159359567999120	0	0
0.744635996320139	0.597716998114124	0.146918998206015	0
0.744638924490261	0.341314697067372	0.249756248869499	0.153567978553390
	0.417984913235886	0.335857728607711	0.246157358156412

Table 6-4: Butcher Tableau for three-stage third order L-stable DIRK scheme

0.169752102061967	0.169752102061967	0	0
0.751288046819627	0.627187114014859	0.124100932804768	0
0.753964223884038	0.313028692059601	0.261339826878021	0.179595704946416
	0.433129323301426	0.345576123139623	0.221294553558950

Table 6-5 shows the calculated radius of absolute monotonicity for all schemes discussed in this paper. It is clear that the A-stable scheme has larger radius of absolute monotonicity than L-stable scheme, which means this scheme maintains SSP property for larger CFL numbers.

SSP(3,3) noticebly preserves the SSP property for larger radius than A-stable scheme. However, as discussed above and will be shown later, it is not the only property required in the advection diffusion reaction equations, while higher order of stability is a necessity for a temporal scheme in order to tolerate larger stable time steps.

Table 6-5: Radius of absolute monotonicity for SSP(3,3) and the new schemes

DIRK Schemes	**Radius of Absolute Monotonicity**
SSP(3,4)	4.8284271247461901
A-stable	2.2812434130240424
L-stable	2.1861028864641930

6.5.3 Stability functions' formulae

Using (6-11) and coefficients in Tables 6-2 to 6-4, the stability function for each scheme can be obtained:

SSP(3,3): (6-15)

$$R(z) = -2.6\frac{(z+9.313100922291955)(z+2.969308968665091-2.069904970162032i)(z+2.969308968665091+2.069904970162032i)}{(z-6.828503385561053)^3}$$

A-stable: (6-16)

$$R(z) = -\frac{(z+25.247316390883434)(z+2.757056602996978-1.847931639434187i)(z+2.757056602996978+1.847931639434187i)}{(z-6.806471676302732)(z-6.511774195506049)(z-6.275117412501512)}$$

L-stable: (6-17)

$$R(z) = \frac{(z+2.622465814681776-1.756014433335682i)(z+2.622465814681776+1.756014433335682i)}{(z-8.057957159541834)(z-5.890943251088293)(z-5.568061888219192)}$$

As explained in Section 4, this paper aims at optimizing the values of the stability function's roots that ensures this function obtains larger range of nonnegative property.

Following the discussion in Section 6.3 on (6-14), for A-stability property, absolute value of A-factor was imposed as a constraint to be less than or equal one. It is clear from stability function (6-15) that SSP(3,3) is not A-stable as A-factor is much larger than one (|A|=2.6).

Table 6-6: The admitted roots for numerical stability functions by the given DIRK schemes, Eqs. 6-15 to 6-17.

	$\mathbf{Z_1}$	$\mathbf{Z_2}$	$\mathbf{Z_3}$
SSP(3,3)	-9.313	-2.97 + 2.07i	-2.97 − 2.07i
A-stable	-25.25	−2.76 + 1.85i	−2.76 − 1.85i
L-stable		-2.62 + 1.76i	−2.62 − 1.76i

The roots of all discussed schemes are briefly provided in Table 6-6. Clearly, the real root of stability function for SSP(3,3), $z \approx -9.313$, is much smaller than proposed A-stable scheme, $z \approx -25.25$.

The stability functions admit two other conjugated roots, whose the imaginary component of the SSP(3,3) scheme is larger than the proposed schemes. As will be shown in the test cases, it is the reason that SSP scheme does not exhibit oscillations in the calculation of the advection dominant flows. However, the A-stable scheme enjoys larger values for both real and imaginary components compared to L-stable. Therefore, comparing the roots of all stability functions, it is expected that A-stable scheme performs more reliable than L-stable and SSP(3,3) schemes. This will be proved in the rest of the paper.

6.5.4 Linear stability analysis

In this section, the stability functions, formed by the schemes, will be compared in two forms. In the first one, the spatial derivative terms will be considered continuous. Hence it is called ODE stability analysis. In the second approach, the effects of spatial discretization method will be taken into account as well.

6.5.4.1 ODE approach

In this approach, the spatial derivative terms are assumed continuous in order to study the performance of the time integration schemes only. Figure 6-1 illustrates the contours of the magnitude of the stability function for the proposed schemes compared to the SSP(3,3). The stable region, in which $|R(z)| \leq 1$, is limited for SSP(3,3) while the A-stable and L-stable schemes provide stability for the variable in the whole left half plane, i.e. $z \in \mathbb{C}^{-}$, except for the limited area near the imaginary axis. However, as shown in this figure, the unstable region for these schemes is limited with the maximum value 1.3. Although, Hairer and Wanner (1996) indicated that A-stability covers the imaginary axis as well, one may find that physical diffusion and source terms, combined with numerical dissipations, which come from numerical discretization methods, move the variable z inward the domain along the negative real axis and consequently the required stable region is moved to somewhere far away from the imaginary axis. The fact that the magnitude of A-factor in (6-16) equals one guarantees the stability for A-stable scheme along the imaginary axis in the infinite limits.

In order to show the non-negativity property of the numerical stability function, the real and imaginary components of these functions are presented separately. Figure 6-2 illustrates the distributions of the real component of numerical stability functions, R_r.

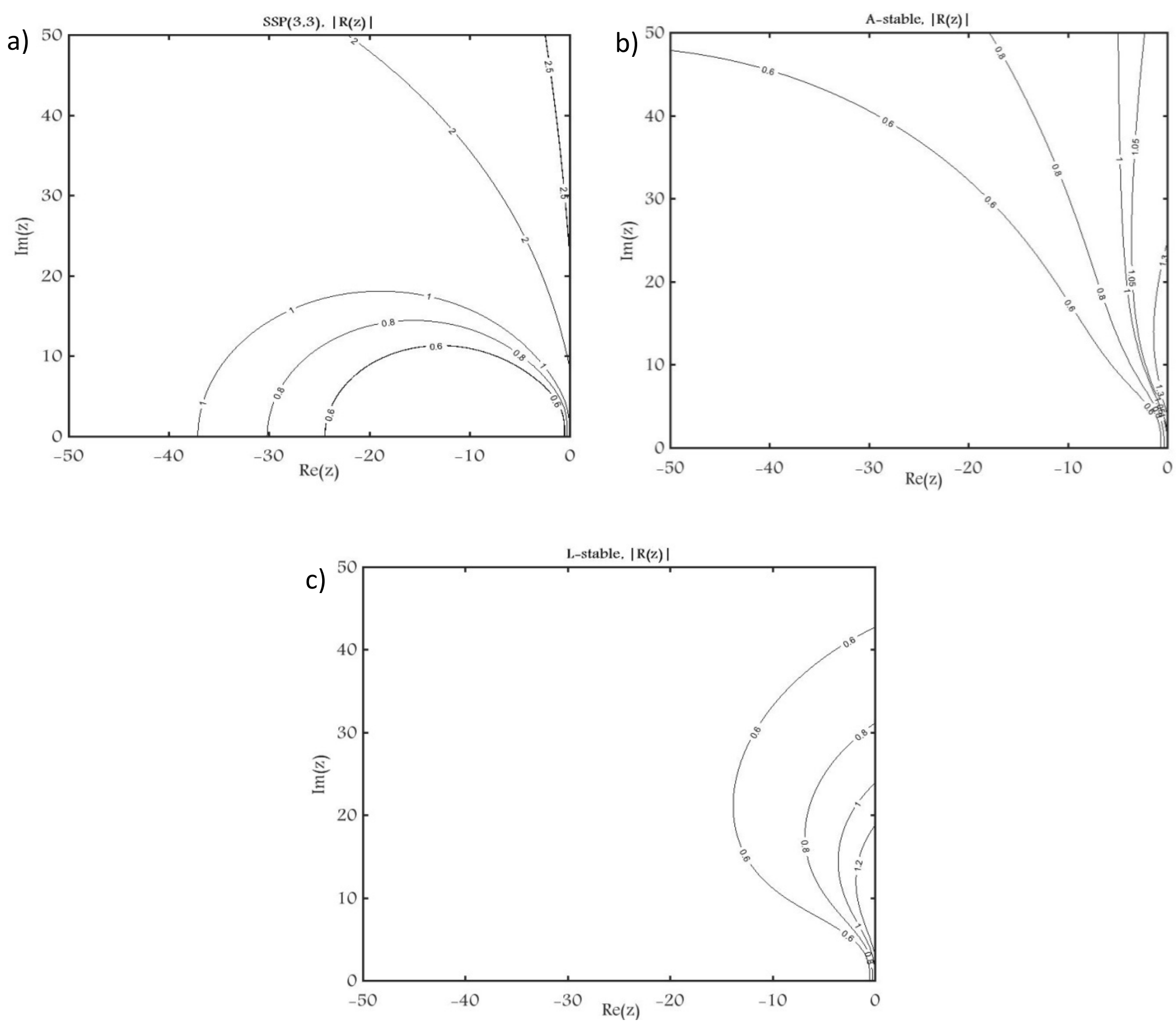
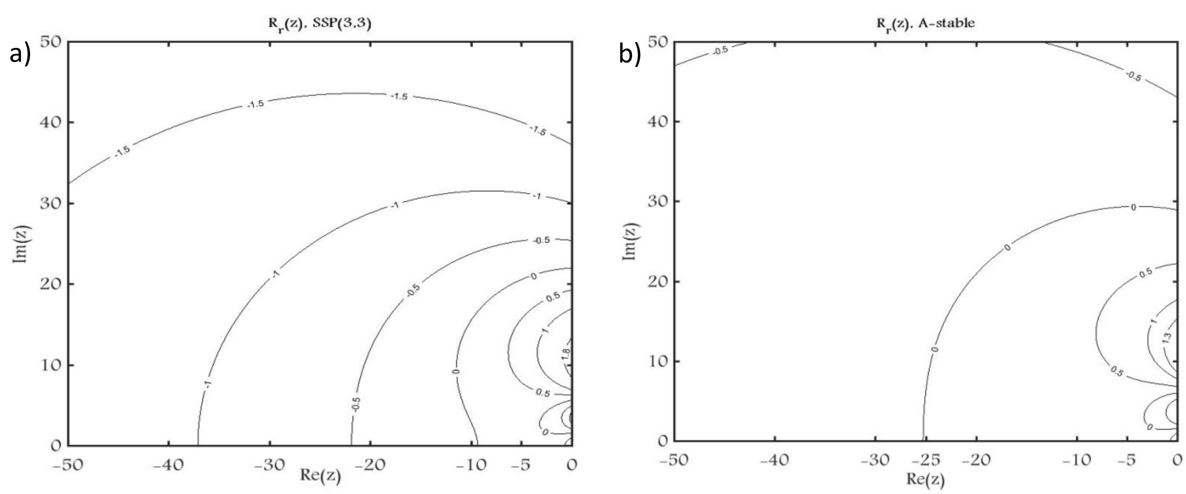


Figure 6-1: The distribution of the magnitude of the stability function, |R(z)|, for SSP(3,3) (a), A-stable (b), L-stable (c).

Except in a small region, formed as a bubble near the origin with $Re(z) < -5$, in which the real component of the stability function (R_r) of all three schemes are negative, the stability function of the L-stable scheme is non-negative through the entire domain. Considering the formula of the stability function, provided in (6-13), one may notice the reason that the L-stability property exhibits less oscillation in the diffusion dominant flows, which is the nonnegative property of the real component of the stability function for the larger range of the variable.

The non-negative zone, produced by the SSP(3,3) intersects the real axis in $Re(z) = -10$, while the A-stable scheme obtains non-negativity along real axis for $Re(z) \leq -25$. Along the imaginary axis, this property is preserved by SSP(3,3) for $Im(z) \leq -20$, and A-stable scheme is non-negative up to $Im(z) = -30$. Note that for larger values of the variable z both SSP(3,3) and A-stable schemes become negative, although the A-stable scheme is still stable, i.e. $R_r(z) \geq -1$.



c)

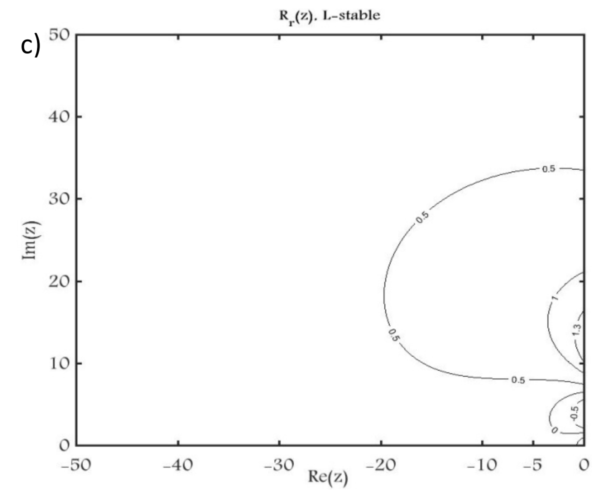


Figure 6-2: The contours of the real component of stability function, $\mathrm{R_r(z)}$, for SSP(3,3) (a), for A-stable (b) and L-stable (c).

In Figure 6-3, distributions of the imaginary component of the numerical stability function, $\mathrm{R_I}$, are shown. All three schemes produce non-negative region for the entire domain except near the origin. Clearly, the SSP(3,3) produces smaller negative zone, intersecting the real axis in $\mathrm{Re(z)} = -5$, than A-stable and L-stable schemes with $\mathrm{Re(z)} = -10$ and $\mathrm{Re(z)} = -20$ respectively. Referring to (6-13), one may find the reason for the less oscillatory behavior of the SSP optimized schemes in the advection dominated flows.

However, for the advection diffusion reaction flows, in which the Fourier variable, i.e. z, admits both real and imaginary components, the A-stability property is expected to be more efficient, since both the real and imaginary parts of its stability function preserve the nonnegative property for the moderate range of the variable.

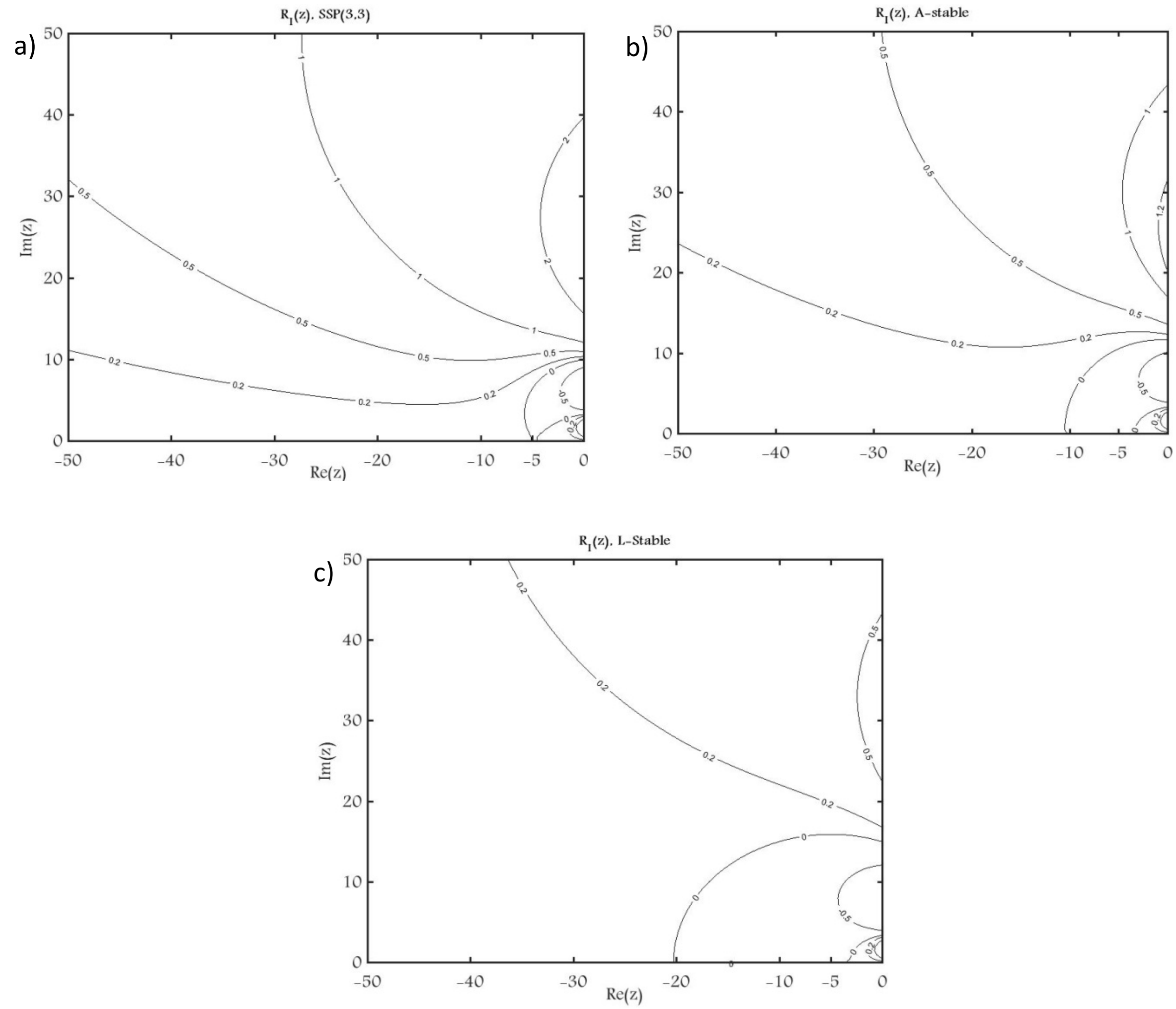


Figure 6-3: The contours of imaginary component of stability function (R_I), SSP(3,3) (a), A-stable (b) and L-stable (c).

6.5.4.2 PDE approach

In order to take into account the effect of spatial discretization schemes in the linear stability analysis and to observe the proposed schemes behaviors, the PDE linear stability analysis of advection-diffusion equation is performed, serving the central discretization scheme for the spatial derivatives.

Linear advection-diffusion equations can be written in the following form:

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2} \tag{6-18}$$

in which a represents the advection velocity and ν shows the diffusion coefficient.

The central scheme, applied to the first and second order derivatives, rearranges the PDE to ODE equation as follows:

$$\frac{\partial u_j}{\partial t} + a\frac{u_{j+1} - u_{j-1}}{2\Delta x} = \nu\frac{u_{j+1} - 2u_j + u_{j-1}}{\Delta x^2}, \tag{6-19}$$

Assuming the Fourier solution of the form $u_j(t) = \hat{u}(t)e^{ikx_j}$ and performing the von Neumann stability analysis lead to the following formula:

$$\frac{\partial \hat{u}}{\partial t} = -\left[i\frac{a}{\Delta x}\sin(2\omega) + 4\frac{\nu}{\Delta x^2}\sin^2(\omega)\right]\hat{u}, \tag{6-20}$$

in which $\omega = k\Delta x/2$. Comparing (6-20) to (6-6), it can be realized that the coefficient of the variable $\hat{u}$ in the Right Hand Side (RHS) is the same as λ. Consequently, the stability function is already the same as (6-11), in which $z = \lambda\Delta t$. Following Hundsdorfer and Ruuth (2007), the linear stability analysis is done in terms of Courant–Friedrichs–Lewy (CFL), hereafter called Courant, and Péclet numbers, defined as follow respectively:

$$\begin{aligned} \mu &= \frac{a\Delta t}{\Delta x}, \\ d &= \frac{a\Delta x}{\nu}, \end{aligned} \tag{6-21}$$

Figure 4 shows the distributions of the magnitude of the numerical stability functions in terms of Courant and Péclet numbers, i.e. $|R(\mu, d)|$. The propsoed A- and L-stable schemes show stability for larger Courant and smaller Péclet numbers, compared to SSP(3,3). This leads to a conclusion that high level of stability is required for simulations of viscous flows; for instance using fine mesh sizes to capture the rapidly changing diffusion properties, the case that often happens in practical applications. Note that these advantages include large stable time-step sizes, corresponding to the large Cournat number, as well.

The more efficient performance of the A-stable scheme compared to the L-stable will be clearer in Figure 4, as the A-stable scheme shows stability for the larger range of Courant and Péclet numbers.

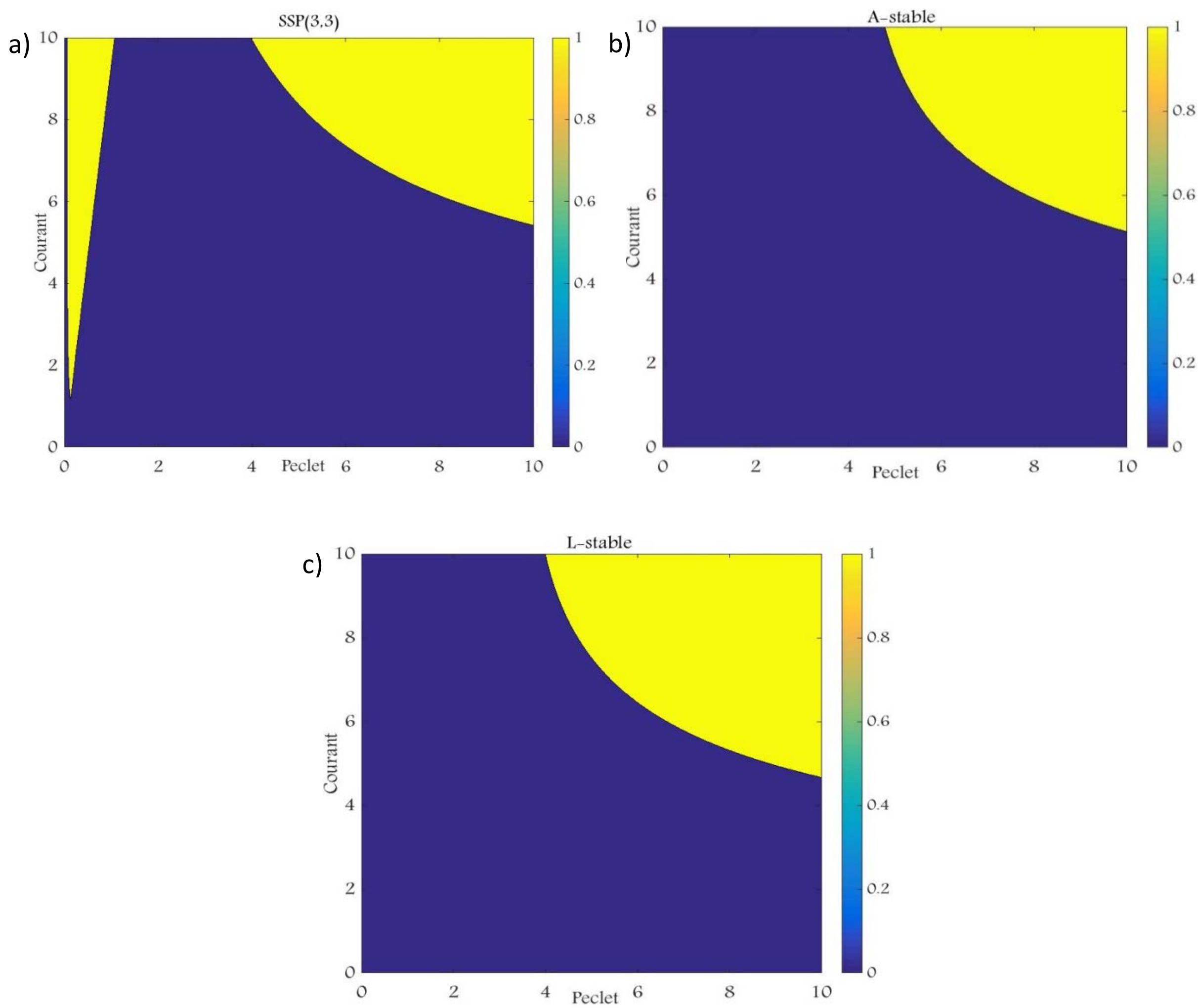


Figure 6-4: The distribution of the magnitude of the stability function, $|R_n|$, in terms of Courant and Péclet numbers, SSP(3,3) (a), A-stable (b) and L-stable (c), stable region, $|R(\mu, d)| \leq 1$, (blue), unstable region, $|R(\mu, d)| \geq 1$, (yellow).

6.6. Numerical test cases

In this section, performance of the new schemes is examined through calculating various problems such that they include the advection, diffusion and reaction terms. First, the ability of the proposed schemes is examined through solving the compressible Euler equations with Gaussian pulse initial condition. Afterwards, the nonlinear viscous Burger equation with shock occurring will be solved. Finally, the results of computing a linear advection equation with nonlinear source term will be discussed.

6.6.1 Compressible Euler equations

The compressible Euler equation set is the best representative of the advection dominated flows. However, in viscous flows, other terms including turbulent shear stress terms and forcing source terms are typically added to these equations as diffusion or source terms. Hence, the purpose of computation of the compressible Euler equations is to compare the performance of the new temporal schemes to the ones obtained by SSP(3,3). For simplicity, just one-dimensional form of compressible Euler equations is calculated.

The conservative form of these equations can be written as:

$$\frac{\partial \mathbf{Q}}{\partial \mathrm{t}} + \frac{\partial \mathbf{F}}{\partial \mathrm{x}} = 0, \tag{6-22}$$

in which **Q** is the vector of the variables and **F** is flux vector, shown as:

$$\mathbf{Q} = \begin{Bmatrix} \rho \\ \rho u \\ e \end{Bmatrix}, \qquad \mathbf{F} = \begin{Bmatrix} \rho u \\ \rho u^2 + P \\ u(e+P) \end{Bmatrix}, \tag{6-23}$$

Pressure variable, P, is derived from the equation of state:

$$P = (\gamma - 1)\left(e - \frac{\rho u^2}{2}\right), \tag{6-24}$$

6.6.1.1 Propagation of Initial Gaussian Pulse

This test simulates the propagation of a one-dimensional Gaussian pulse. The initial conditions are taken from Du and Ekaterinaris (2016):

$$\begin{cases} \rho_0 = 1.0 + \varepsilon e^{-x^2/b^2} \\ p_0 = 1.0 + \varepsilon e^{-x^2/b^2} \\ u_0 = -1.0 \end{cases} \tag{6-25}$$

where, $\varepsilon = 0.01$ and $b = 3.0$. The flow domain within the range of $x \in [-100,100]$ is divided uniformly into 500 cells and boundary conditions are assumed as periodic. For spatial discretisation of the flux terms, highly accurate WENO-5 scheme proposed by Wang and Spiteri (2007) is used. The initial pulse evolves into one entropy wave and two acoustic waves which propagate at the speeds u_0 and $u_0 \pm c$, respectively, where $c = 1.183$. The solutions were computed using CFL=2.1, while for the reference solution, optimal third order explicit Total Variation Diminishing (TVD) scheme (Gottlieb and Shu, 1998) with a very small time step, corresponding to CFL=0.01, was employed.

Figure 6-5 illustrates that the A-stable scheme exhibits less oscillations than the L-stable scheme, while the disturbances are disappeared by the SSP(3,3). The reason can be found within the discussions on non-negativity property. It can be observed that the governing equation, (6-22), is hyperbolic without any diffusion terms. Hence, according to the linear stability analysis, the variable z obtains larger imaginary values, compared to the real ones, as there is no physical diffusion terms in the governing equations. Therefore, referring to the distribution of the imaginary component of the stability function, Figure 6-3, it is expected that the SSP(3,3) shows less oscillations, due to the larger nonnegative staibity region. Due to the same reason, A-stable scheme performs better than L-stable scheme.

6.6.2 Nonlinear viscous Burgers equation

In order to observe the performance of the proposed temporal schemes in computation of viscous flows as well as to examine it in shock capturing, the one-dimensional nonlinear viscous Burgers equation is selected:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2}, \tag{6-26}$$

The periodical initial value is $u(x,0) = R\sin(x)$ where R=100 and $\nu = 1.0$. Due to the large initial amplitude, a discontinuity will form and is gradually damped by the large viscous term. The flow domain is $[0,2\pi]$ divided into 150 uniform cells, in which the spatial terms were discretized by WENO-5 scheme (Wang and Spiteri, 2007). For comparison, reference solution has been obtained by optimal explicit third order TVD scheme (Gottlieb and Shu, 1998) marching in time with CFL=0.05.

Figure 6 shows the velocity profile, u, and error distribution, $u - u_{ref}$, computed with CFL=6, at the time of shock forming, T=0.02s. Clearly, the A-stable scheme behaves more accurately with less oscillations, while the SSP(3,3) scheme shows more oscillatory behaviors after the L-stable scheme. One may find the reason as due to the larger range of non-negative region of the real component of the stability function provided by the A-stable scheme compared to the SSP optimized scheme and larger range of non-negativity region of the imaginary component of the stability function compared to the L-stable scheme. Hence, the importance of the non-negativity property combined with the A-stability property is revealed in calculation of the viscous flows, even though the SSP optimized scheme, i.e. SSP(3,3), produces the radius of absolute monotonicity as large as twice of the A-stable scheme.

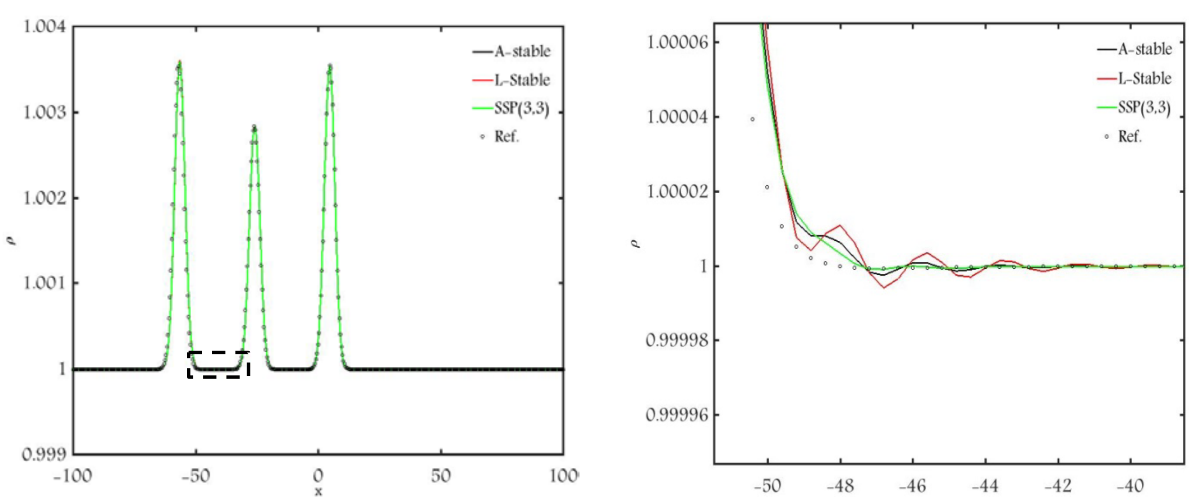


Figure 6-5: Distribution of density, with initial Gaussian Pulse, at T=26 with CFL=2.1, using WENO 5, reference solution is from TVD explicit third order with CFL=0.01.

6.6.3 Linear advection equation with a nonlinear source term

In order to examine the proposed schemes’ performance in calculating the nonlinear source terms, the linear advection equation with a nonlinear source term, recommended by Griffiths et al. (1992) and Nazari et al. (2015), is calculated and the results are shown in Figure 7. The governing equation, which would be considered as a model of nonequilibrium gas dynamics (Griffiths et al., 1992), is written as:

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = u - u^2, \tag{6-27}$$

The spatial derivative term is discretized by Central-Upwind scheme through 200 uniform cells in the domain, $x \in [0,10]$. The initial condition is $u(x,0) = 1$, whereas $x \in [2.3,3]$ and elsewhere $u(x,0) = 0$. The reference solution was obtained by optimal explicit TVD third order scheme (Gottlieb and Shu, 1998) marching on a very small time step (CFL=0.05), while the numerical results were calculated using CFL=2.6. The velocity, u, and error distribution, $u - u_{ref}$ are shown in Figure 7.

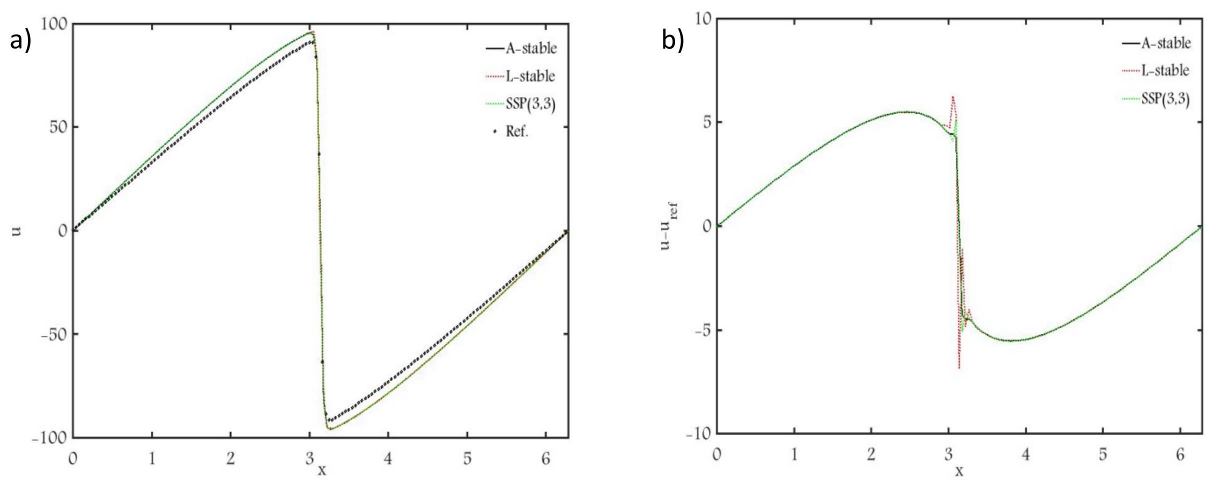


Figure 6-6: The velocity, u, (a) and error, $u - u_{ref}$, (b) distributions of the nonlinear Burgers equation at T=0.02, using WENO5 and CFL = 6, Reference solution TVD explicit third order with CFL=0.05.

It is clear that L-stable schemes show more oscillatory behaviors than other schemes. Clearly, SSP(3,3) scheme over-predicts the results, while A-stable scheme presents more reliable behavior.

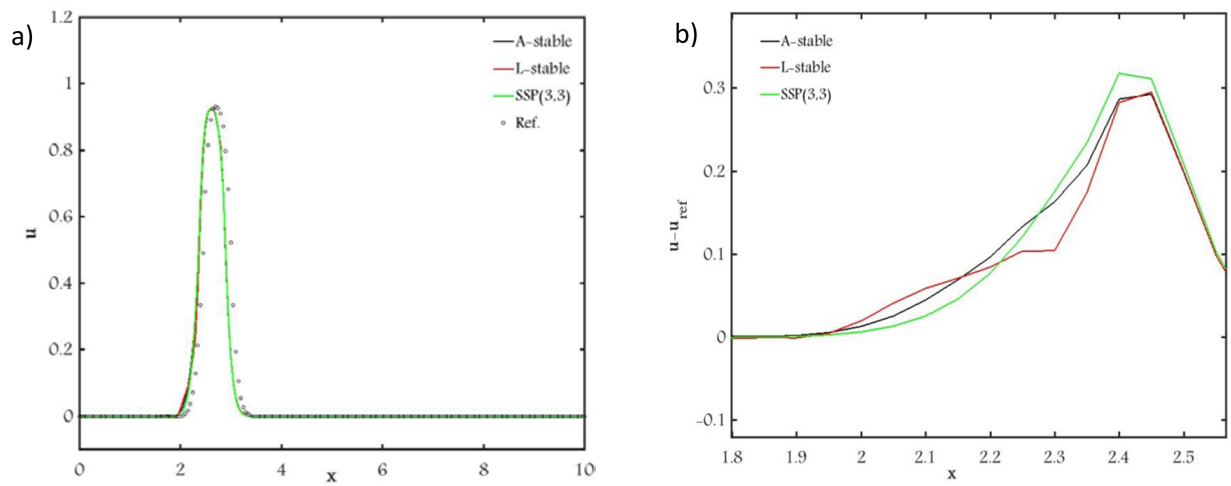


Figure 6-7: Distribution of velocity (a) and error (b), linear advection equation with nonlinear source term, T=1.4s and $\mathrm{CFL} = 2.6$, Reference solution obtained by TVD explicit third order with $\mathrm{CFL} = 0.05$.

6.7. Conclusion

The aim of this study was to reveal the significance of the nonnegative property, corresponding to the stability function, in providing non-oscillatory behaviour for the advection diffusion reaction problems. Moreover, it was aimed to observe how the well-known stability properties, i.e. A- and L-stability and SSP optimized properties, may influence achieving this non-negativity property. Therefore, two three-stage third order DIRK schemes with optimized nonnegative stability functions combined with A- and L-stability properties, called A- and L-stable schemes, were developed and compared to the SSP(3,3), which is in the same class of order of accuracy but with the maximum radius of absolute monotonicity.

It was found that the SSP(3,3) produces larger non-negative region for the imaginary component of the stability function, which is the main reason for non-oscillatory behavior in calculating the advection dominant flows. The L-stability property produces larger range of non-negativity but just for the real component of the stability function. However, the A-stability property was found as more efficient one, since the stability function moderately preserve the non-negativity property for both real and imaginary components. Furthermore, it was shown that the L-stability property and the associated restrictive conditions result in decreasing the radius of absolute monotonicity, more than the A-stability property.

These hypotheses were proved by computing the various types of test cases such that they include advection, diffusion and reaction terms. It was found that the SSP property with maximum radius of absolute monotonicity help to calculate the Euler compressible equations, a representative of the advection dominant flows, with insignificant oscillations. However, calculating the nonlinear viscous Burger and the linear advection with nonlinear source terms demonstrated that the maximum radius of absolute monotony is not the only essential property and the A-stability seems to be required as well.

To conclude, it was found that the A-stability property combined with SSP property more effectively can reduce the oscillations in computation of viscous flows, with advection, diffusion and reaction terms, so that the time integration scheme can tolerate larger CFL numbers with less oscillation.

Acknowledgment

The authors wish to acknowledge financial support from Natural Sciences and Engineering Research Council of Canada (NSERC).

Chapter 7

Efficient semi-implicit temporal scheme for boundary layer vertical diffusion[5]

Abstract:

Time integration of the boundary layer vertical diffusion equation has been investigated. The nonlinearity associated with the diffusion coefficient makes the fully-implicit approach impractical while the use of fully explicit scheme limits the stable time step sizes and consequently would be inefficient. By using diagonally implicit Runge-Kutta scheme, a new approach has been proposed, in which the diffusion coefficient at each internal stage is calculated by a weight-averaged combination of the solutions. Using the weight coefficient, α, offers more robust calculations, due to involving the implicit solutions and, as shown, it could improve the accuracy due to more engaging the explicit solutions. It has been found that the nominated semi-implicit method results in less dissipation error, more accurate solutions and eventually less CPU time usage, compared to the fully-implicit schemes, and it enjoys a larger range of stable time steps than other semi-implicit approaches.

Keywords: Vertical diffusion, Boundary layer, Semi-implicit, DIRK scheme, Stability, Accuracy.

7.1. Introduction

The turbulent diffusive mechanism is simulated by using the diffusion coefficients. To capture the flow properties, with rapidly changing flux gradients within the thin boundary layer, the time integrators need to be implicit; otherwise, the stable time step would be impractically too restricted in explicit schemes. However, the diffusion coefficients, which, in most cases, are dependent on the flow variables, result in nonlinear equations that make the fully-implicit scheme impractical. The use of semi-implicit schemes allows larger stable time steps, although some oscillatory behaviors and instability have been observed, e.g. Davies et al. (2005). In

---

[5] Arman Rokhzadi, Abdolmajid Mohammadian, under review, 2018.

common semi-implicit approaches, only the dependent variables are calculated implicitly and diffusion coefficients, function of the variables, are treated explicitly.

Kalnay and Kanamitsu (1988) investigated the vertical diffusion boundary layer using a nonlinear damping model, while they examined various two-time level schemes, including predictor-corrector, and different calculation techniques for the diffusion coefficients. They also introduced new approach, so called explicit exchange coefficient and implicit variable, which is followed by a time averaging. This new approach, as shown by Kalnay and Kanamitsu (1988), provides wider range of stability.

Girard and Delage (1990) studied the influence of varying diffusion coefficients on stability of the semi-implicit approach in calculating the vertical diffusion boundary layer. The diffusion coefficients, as a function of Richardson number, vary with static stability and wind shear. Later on, Bénard et al. (2000) proposed a modification to the Richardson number and showed some accuracy improvements compared to Girard and Delage (1990). Despite difficulties of implementing these techniques in the boundary layer calculation, they lead to more stability although at the expense of loss of accuracy.

Nazari et al. (2013) optimized an implicit Runge-Kutta scheme for the boundary layer vertical nonlinear diffusion for the atmospheric applications. By solving the turbulent boundary layer flow with $E-L$ turbulence closure scheme, they showed that this scheme can allow to extend the range of stable time steps and to improve the accuracy

A semi implicit predictor-corrector method has been utilized by many researchers including Clancy and Pudykiewicz (2013) who applied this approach to the nonlinear system of the shallow water equations. Diamantakis et al. (2006) studied general form of the semi-implicit predictor-corrector method, in which the diffusion coefficients would be calculated using over-weighted averaging technique, hereafter called weighted semi-implicit predictor corrector (weighted-SI-PC). Through asymptotic error analysis, they found the best choice of over-weighting coefficients, which provide better stability and more accurate solutions, compared to the approach employed by Kalnay and Kanamitsu (1988). Meanwhile, they indicated that unconditional stability can be achieved only for linear or weakly nonlinear problems.

In this paper, a Diagonally Implicit Runge-Kutta (DIRK) scheme has been employed to apply the weighted-averaging semi-implicit method to the vertical diffusion boundary layer equation. In this regard, the diffusion coefficients will be calculated by taking average over the solutions.

This paper is organized as follows. In Section 7.2, the nonlinear model introduced by Kalnay and Kanamitsu (1988) will be briefly described as well as the numerical method of Diamantakis et al. (2006). Afterward, the implementation of the DIRK scheme will be explained and further details of the proposed weighted-averaging approach will be provided. In Section 7.3, the stability function of different approaches, using the linearized method proposed by Kalnay and Kanamitsu (1988), will be presented. Section 7.4 explains how to optimize the weighted-averaging coefficients in applying with the nominated DIRK scheme in the calculation of nonlinear diffusion equation. In Section 7.5, the performance of different approaches will be examined in terms of computing the nonlinear equation and Section 7.6 will complete the paper with some concluding remarks.

7.2. Numerical calculation

The following partial differential equation (PDE) represents the governing equation in the vertically diffusive transportation, which occurs in practice and more specifically in the atmospheric fields

$$\frac{\partial X}{\partial t} = \frac{\partial F}{\partial z} + S, \qquad F = K\frac{\partial X}{\partial z} \tag{7-1}$$

in which K represents the diffusion coefficient, which is a function of the variable, $K = K(X)$, in which X could be any tracer such as velocity, moisture or temperature. The parameter z shows the vertical coordinate and the source terms, S, indicates a forcing term from the physical processes. The numerical computations may be performed through following discretization

$$\frac{X^{n+1} - X^n}{\Delta t} = \frac{\partial \bar{F}^{\xi}}{\partial z} + S, \tag{7-2}$$

where, n denotes the time-level index and Δt is the time step. The overbar on the flux term, $\bar{\ }^{\xi}$, shows the time-averaging value, obtained by the coefficient ξ,

$$\bar{F}^{\xi} = (1-\xi)F^{n} + \xi F^{n+1}, \tag{7-3}$$

Commonly, the implicit flux is calculated as

$$F^{n+1} = K(X^{n})\frac{\delta X^{n+1}}{\delta z}, \tag{7-4}$$

Kalnay and Kanamitsu (1988) demonstrated that the time-average coefficient less than one, i.e. $\xi \leq 1$, may not result in unconditional stability, hence, they suggested using an over-weighting such that $\xi > 1$, although, it may cause losing the accuracy (Diamantakis et al., 2006). Therefore, to achieve the unconditional stability, they suggested utilizing the fully-implicit method, where the diffusion coefficient in (7-4) is approximated as $K(X^{n+1})$ instead. However, due to nonlinearity, the calculation could be difficult to implement. Eventually, Diamantakis et al. (2006), to avoid difficulty of the fully-implicit scheme, but, to take its stability advantages, proposed a predictor-corrector scheme,

$$\begin{aligned}
\frac{X^{(1)}-X^{n}}{\Delta t} &= \frac{\partial}{\partial z}\left\{(1-\xi_1)\left(K^{n}\frac{\partial X^{n}}{\partial z}\right) + \xi_1\left(K^{n}\frac{\partial X^{(1)}}{\partial z}\right)\right\} + S, \\
\frac{X^{(2)}-X^{n}}{\Delta t} &= \frac{\partial}{\partial z}\left\{(1-\xi_2)\left(\bar{K}^{\alpha}\frac{\partial X^{n}}{\partial z}\right) + \xi_2\left(\bar{K}^{\alpha}\frac{\partial X^{(2)}}{\partial z}\right)\right\} + S, \\
X^{n+1} &= X^{(2)}
\end{aligned} \tag{7-5}$$

where,

$$\bar{K}^{\alpha} = (1-\alpha)K^{n} + \alpha K^{(1)}, \quad 0 \leq \alpha \leq 1 \tag{7-6}$$

They found the most efficient calculations with $\xi_1 = 2$, $\xi_2 = 1$ and $\alpha = \frac{1}{2}$.

In this paper, the DIRK scheme has been employed to calculate the vertical diffusion boundary layer equation, (7-1), using different approaches, including fully-implicit and semi-implicit techniques. Moreover, a new weighted-averaging semi-implicit approach will be proposed to improve the stability and accuracy.

Applying the ν-stage DIRK scheme to the general governing equation of the form $\frac{\partial X}{\partial t} = f(X)$ results in the following relations:

$$X^{n+1} = X^n + \Delta t \sum_{i=1}^{\nu} b_i f(Y_i), \tag{7-7}$$

$$Y_i = X^n + \Delta t \sum_{j=1}^{i} a_{ij} f(Y_j),$$

in which vector $\boldsymbol{b}$ and matrix $\boldsymbol{A}$ denote the weight coefficients of any Runge-Kutta schemes that must satisfy the so called order conditions. The reader is referred to the literature for further details of Runge-Kutta schemes and the corresponding order conditions, e.g. Dekker and Verwer (1984) and Kennedy and Carpenter (2016).

In the weighted-averaging semi-implicit method, as shown in (7-8), the implicit term, in which the diffusion coefficient is shown as $\overline{K}^{(i)}$, is separated, to demonstrate that the diffusion coefficient, which is a function of the variable X, is obtained by taking weighted-average over the solutions at the previous time step and the i$^{\text{th}}$ internal stage, i.e. $K(X^n)$ and $K(Y_i)$, as shown in (7-9). In this study, the proposed method has been named weighted semi-implicit Runge-Kutta (weighted-SI-RK). Since the main concern of this paper is the nonlinearity associated with the diffusion terms, the source term has been excluded from the discussion, while it is calculated implicitly.

$$Y_i = X^n + \Delta t \sum_{j=1}^{i-1} a_{ij} \left\{ \frac{\partial}{\partial z} K^{(j)} \frac{\partial Y_j}{\partial z} + S_j \right\} + \Delta t \; a_{ii} \left(\frac{\partial}{\partial z} \overline{K}^{(i)} \frac{\partial Y_i}{\partial z} + S_i \right)$$

$$X^{n+1} = X^n + \Delta t \sum_{i=1}^{\nu} b_i \left\{ \frac{\partial}{\partial z} K^{(i)} \frac{\partial Y_i}{\partial z} + S_i \right\} \tag{7-8}$$

$$\overline{K}^{(i)} = \alpha_i K(Y_i) + (1 - \alpha_i) K(X) \tag{7-9}$$

Note that the number of the weight coefficients, α_i, is the same as the number of internal stages, hence, $i \in [1, \dots, \nu]$. Clearly this method involves the fully-implicit approach, hereafter called FI-RK, which is obtained by setting $\alpha_i = 1$, as well as the conventional semi-implicit approach, where the exchange coefficients are calculated explicitly by using the solution at the previous time step, achieved by $\alpha_i = 0$. The latest approach, with $\alpha_i = 0$, is called semi-explicit Runge-Kutta (SE-RK) scheme.

7.3. Linear stability analysis

In this section the stability of the weighted-SI-RK is analyzed using the linearization method proposed by Kalnay and Kanamitsu (1988). They proposed the following damping nonlinear equation which can simply play the role of nonlinearity in the calculations.

$$\frac{\partial X}{\partial t} = -(K_0 X^P)X + S \tag{7-10}$$

Perturbation around the steady state solution, i.e. $X_0 = \left(\frac{S}{K_0}\right)^{1/(1+P)}$, leads to the linearized governing equation for the perturbed variable, δX, as

$$\frac{\partial(\delta X)}{\partial t} = -C(X_0)(\delta X + P\delta X) \tag{7-11}$$

where $X = X_0 + \delta X$ and $C(X) = K_0 X^P$. As indicated by Kalnay and Kanamitsu (1988), the second term in the right hand side (RHS) of (7-11) corresponds to the nonlinear diffusion coefficient in the vertical diffusion boundary layer equation. The analytical solution of the linearized governing equation,

$$\delta X = e^{-C(X_0)(1+P)t} \tag{7-12}$$

forms the analytical stability function, R_a, defined as the ratio of the solutions in two successive time steps,

$$R_a = \frac{\delta X(t+\Delta t)}{\delta X(t)} = e^{-C(X_0)(1+P)\Delta t} \tag{7-13}$$

Numerical solution of (7-11), obtained by ν-stage DIRK scheme, can be achieved by using the general formula in (7-7) as

$$\begin{aligned}
\delta Y_i &= \delta X^n + \Delta t \sum_{j=1}^{i-1} a_{ij} \left\{-C(\delta Y_j + P\delta Y_j)\right\} + \Delta t\, a_{ii}\{-C(\delta Y_i + P\delta Y_i)\} \\
\delta X^{n+1} &= \delta X^n + \Delta t \sum_{i=1}^{\nu} b_i \{-C(\delta Y_i + P\delta Y_i)\}
\end{aligned} \tag{7-14}$$

One may notice that (14) corresponds to the FI-RK approach, which is identical with the weighted-averaging parameter $\alpha = 1$ (see Equations 7-8 and 7-9). In FI-RK scheme the diffusion coefficient is calculated implicitly, while in weighted-SI-RK it is calculated by taking average over the solutions in the current internal stage and the previous time step as

$$\begin{aligned}
\delta Y_i &= \delta X^n + \Delta t \sum_{j=1}^{i-1} a_{ij} \left\{-C(\delta Y_j + P\delta Y_j)\right\} \\
&\quad + \Delta t\, a_{ii}\{-C(\delta Y_i + P(\alpha\delta Y_i + (1-\alpha)\delta X^n))\} \\
\delta X^{n+1} &= \delta X^n + \Delta t \sum_{i=1}^{\nu} b_i \{-C(\delta Y_i + P\delta Y_i)\}
\end{aligned} \tag{7-15}$$

It is clear that the SE-RK approach can be achieved by setting $\alpha = 0$ in (7-15). Some manipulations are required to render the solutions in vector form as

$$\begin{aligned}
\boldsymbol{\delta Y} &= \delta X^n \boldsymbol{e} - C\Delta t(\boldsymbol{A} + P\boldsymbol{A_1})\boldsymbol{\delta Y} - C\Delta t P((\boldsymbol{I} - \boldsymbol{\alpha})\boldsymbol{A}\delta X^n \boldsymbol{e} \\
\delta X^{n+1} &= \delta X^n - C\Delta t(P+1)\, \boldsymbol{b}^T \boldsymbol{\delta Y}
\end{aligned} \tag{7-16}$$

in which

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & a_{ll} & 0 \\ a_{v1} & a_{v2} & a_{vl} & a_{vv} \end{bmatrix}, \qquad \boldsymbol{A_1}=\begin{bmatrix} \alpha_1 a_{11} & 0 & 0 & 0 \\ a_{21} & \alpha_2 a_{22} & \dots & 0 \\ \vdots & \vdots & \alpha_l a_{ll} & 0 \\ a_{v1} & a_{v2} & a_{vl} & \alpha_v a_{vv} \end{bmatrix},$$

$$\boldsymbol{\alpha}=\begin{bmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \alpha_l & 0 \\ 0 & 0 & 0 & \alpha_v \end{bmatrix} \tag{7-17}$$

and $$\boldsymbol{\delta Y}=\begin{Bmatrix} \delta Y_1 \\ \vdots \\ \delta Y_v \end{Bmatrix}, \qquad \boldsymbol{e}=\begin{Bmatrix} 1 \\ \vdots \\ 1 \end{Bmatrix}, \qquad \boldsymbol{I}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Eventually, the numerical stability function for weighted-SI-RK approach, R, defined as the ratio of the solutions in two successive time steps, can be found as,

$$R = 1 - C\Delta t(P+1)\, \boldsymbol{b}^T\big(\boldsymbol{I} + C\Delta t(\boldsymbol{A} + P\boldsymbol{A_1})\big)^{\boldsymbol{-1}}(\boldsymbol{I} - C\Delta t P(\boldsymbol{I} - \boldsymbol{\alpha})\boldsymbol{A})\boldsymbol{e} \tag{7-18}$$

The stability function of the FI-RK and SE-RK schemes can be easily found by substituting $\boldsymbol{\alpha} = \boldsymbol{I}$ and $\boldsymbol{\alpha} = \boldsymbol{0}$ respectively in (7-16)-(7-18).

7.4. Optimization

This section explains how the averaging coefficients, α_i, can help minimize the dissipation errors, thus, improve the performance, including the stability and accuracy, of the DIRK scheme in terms of solving the nonlinear diffusion equation. In this regard, a well-known DIRK scheme is employed, so called SSP(2,2), which consists of two stages with second order of accuracy and preserves the strong stability property (SSP) (Ketcheson et al., 2009). The Butcher tableau of this DIRK scheme is shown in table 7-1. The Butcher tableau contains the $\boldsymbol{A}$ matrix whose i^{th} row shows the weight coefficients of the i^{th} internal stage solutions. Similarly, the vector $\boldsymbol{b}$ represents the weight coefficients for the solutions at the next time step, i.e. t^{n+1} and the vector $\boldsymbol{c}$ measures the distances of the internal stages from the previous time step, i.e. t^n. Further details may be found in the literature, e.g. Butcher (1963), Butcher (1987) and Butcher (2009).

Table 7-1: Butcher tableau of SSP(2,2) scheme

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b^T} \end{array} \qquad \begin{array}{c|cc} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Substituting the Fourier series of the time dependent solutions, i.e. $\delta X^n = e^{-i\omega t_n}$, in which the parameter ω is called frequency, into the analytical stability function, (7-13), let the analytical dispersion relation appear as

$$e^{-i\omega\Delta t} = e^{-C(X_0)(1+P)\Delta t} \tag{7-19}$$

Obviously, the magnitude of the analytical dispersion relation is one and the argument vanishes, since the RHS just holds the real values. Doing the same manipulations in the numerical stability function, (7-18), demonstrates the relationship between numerical dispersion and numerical stability function. Comparing the numerical stability function of the SE-RK, substituting $\boldsymbol{\alpha} = \mathbf{0}$ in (7-18), with the FI-RK's, substituting $\boldsymbol{\alpha} = \boldsymbol{I}$ in (7-18), reveals that in the SE-RK approach, the magnitude of the numerical stability function is closer to one (less dissipating), thus, the dispersion error is smaller. This fact can quickly be verified thanks to the nonnegative components of the matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}^T$(see table 7-1). Then, for the same reason, the weighted-SI-RK produces less dispersion error than FI-RK scheme. Eventually, as shown later, it can be expected that SE-RK can behave more accurately than weighted-SI-RK and then FI-RK.

To find the unconditional stability condition, letting the time step admit the infinite value, i.e. $\Delta t \to \infty$ allows (18) to be modified as,

$$R = 1 + C\Delta t P(P+1)\, \boldsymbol{b}^T(\boldsymbol{A} + P\boldsymbol{A_1})^{-1}(\boldsymbol{I} - \boldsymbol{\alpha})\boldsymbol{Ae} \tag{7-20}$$

Clearly, the FI-RK is unconditionally stable, realized by $\boldsymbol{\alpha} = \boldsymbol{I}$, compared to SE-RK with $\boldsymbol{\alpha} = \mathbf{0}$. However, in the weighted-SI-RK, the average coefficients, α_i, can be designed such that the last term in the RHS of (7-20) vanishes and eventually the unconditional stability would be achieved.

Obviously, the optimum averaging coefficients are different based on the chosen DIRK scheme and value of P, the power of the nonlinearity term. In this paper, the optimum averaging coefficients will be introduced using the SSP(2,2) and $P = 2$ and 4 (see table 7-2).

For the optimization purpose the sequential quadratic programming (SQP) algorithm of the optimization toolbox of MATLAB software was used. The minimum value of the last RHS term in (7-20) was selected as the objective function. Moreover, the following two extra conditions were found as essential:

7-1- The averaging coefficients, α_i, must satisfy the condition $0 \le \alpha_i \le 1$ ensuring that the solutions preserve the convexity, which means that the averaged solutions are still the combinations of the solutions. Meanwhile the non-negativity, i.e. $0 \le \alpha_i$, helps maintaining the SSP property, since the averaging coefficients multiply the DIRK scheme's coefficients, see (7-15).

7-2- It was found that the extra condition, i.e. $\sum \alpha_i = 1$, can further improve the performance of the weighted-SI-RK approach.

Table 7-2: The average coefficients for weighted-SI-RK approach, P $=2$ and 4

	$\boldsymbol{\alpha_1}$	$\boldsymbol{\alpha_2}$
$\boldsymbol{P = 2}$	0.411437827766148	0.588562172233852
$\boldsymbol{P = 4}$	0.447821961869480	0.552178038130520

7.5. Results and discussions

In this section, the performance of DIRK scheme using the weighted-SI-RK approach will be compared to the fully-implicit and common semi-implicit approaches through solving the nonlinear ordinary differential equation (ODE) nominated by Kalnay and Kanamitsu (1988).

Then the proposed approach will be examined by applying to the one dimensional vertical diffusion boundary layer turbulent flow with Coriolis effect, with the $E - l$ turbulence closure scheme.

7.5.1. Numerical simulation of the nonlinear problem (Kalnay and Kanamitsu, 1988)

Figure 7-1 shows the distribution of the magnitude of the numerical stability function, (7-18), using the SSP(2,2) and $P = 2$ and 4. Clearly, the SE-RK is less damping, since the stability function is closer to one for larger range of time step sizes, thus, this approach can be expected to result in more accurate solutions than weighted-SI-RK and FI-RK approaches respectively, as shown later.

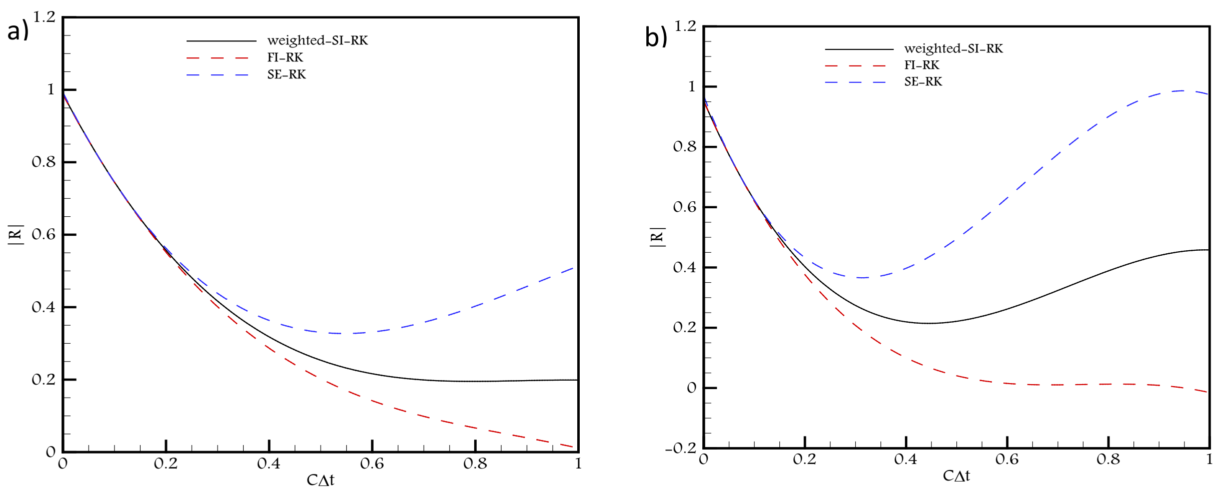


Figure 7-1: The distribution of magnitude of numerical stability function, $|R|$ (vertical axis) vs. time step, Δt, (horizontal axis); using SSP (2,2) scheme, weighted-SI-RK (black solid), FI-RK (red dashed) and SE-RK (blue dashed), a) $P = 2$ and b) $P = 4$.

Moreover, the equation (7-10) was solved by the SSP(2,2) scheme using three approaches and compared to the proposed weighted-SI-PC approach proposed by Diamantakis et al. (2006) (see Eq. 7-5). The forcing source term, S, was selected as

$$S = 1 + \sin(\frac{2\pi}{24}t) \qquad (7\text{-}21)$$

which is a typical form of the variation of a 24-hour diurnal cycle. Following the same work and the one by Kalnay and Kanamitsu (1988) K_0 and X_0 was set to 10 and 0.6 respectively. The governing equation was solved with different time steps up to the maximum stable size in the domain $t \in [0,48]$ hours. The reference solution was obtained by using the optimal explicit total variation diminishing (TVD) third order scheme (Gottlieb and Shu, 1998) with a very small time step ($\Delta t = 0.001s$).

The first row of figure 7-2 shows the distribution of variable, X, obtained by using $\Delta t = 0.25s$ for both cases, i.e. $P = 2$ and 4. Different approaches could capture the reference solution with a reasonable accuracy. However, in the second row the error distribution, defined as the difference between numerical solutions and reference solutions, illustrates that the weighted-SI-RK could significantly decrease the level of numerical error, after the SE-RK approach, but before FI-RK and weighted-SI-PC.

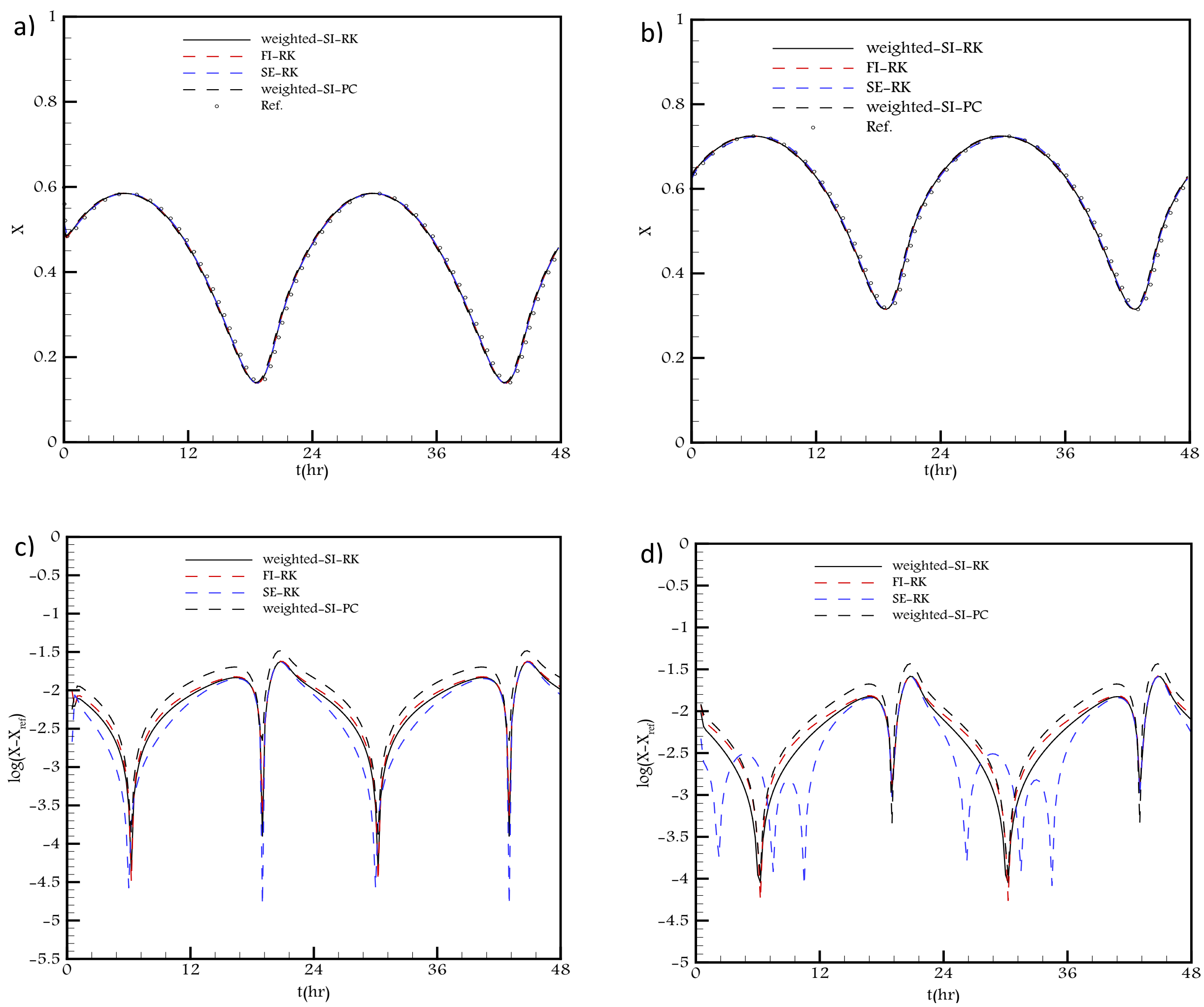


Figure 7-2: The distribution of the calculated variable, X, (first row) and the log scale of the error, i.e. the difference between numerical and reference solutions, $X - X_{ref}$, (second row) on the vertical axis vs. time, t, (horizontal axis); using SSP (2,2) scheme, weighted-SI-RK (black solid), FI-RK (red dashed) and SE-RK (blue dashed) and weighted-SI-PC (black dashed), $P = 2$ (a & c) and $P = 4$ (b & d).

Since the proposed DIRK schemes take advantage of implicitly calculating the exchange coefficients and therefore the Newton method is required for iterations, it is expected that the calculation cost increases. Figure 7-3 illustrates the CPU time usage, on vertical axis, of the different approaches against different stable time step sizes. Clearly, the proposed weighted-SI-PC scheme takes less calculation time than others, except for the case P = 2 with smaller time step sizes, in which the DIRK scheme converges faster than Predictor-Corrector approach. Among different approaches of DIRK scheme, the proposed weighted-SI-RK is competitive to the weighted-SI-PC approach, particularly with larger time step sizes.

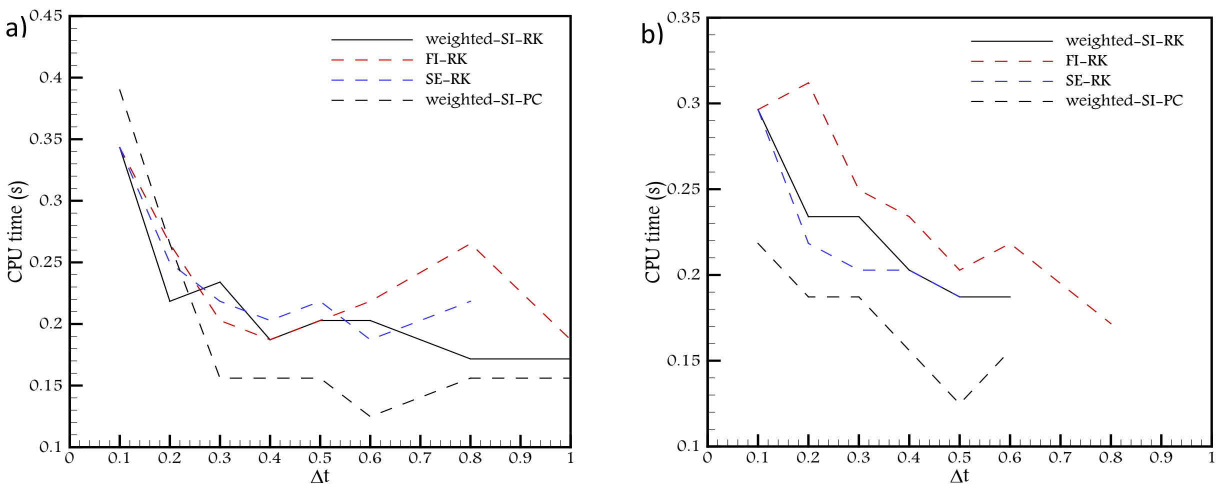


Figure 7-3: The CPU time of the calculation, against various stable time step, Δt, calculated by SSP(2,2), weighted-SI-RK (black solid), FI-RK (red dashed) and SE-RK (blue dashed) and weighted-SI-PC (black dashed), $P = 2$ (a) and $P = 4$ (b).

7.5.2. The E− l turbulence closure scheme with Coriolis effect

One dimensional, horizontally homogenous, dry boundary layer was also examined in order to observe how the weighted averaging technique can help improve the computation of the accurate solutions in the vertical diffusion boundary layer. Following André et al. (1978) the governing equation set, for the eastward and northward velocity components u and v and the potential temperature θ, is composed of,

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z}\left(K_m \frac{\partial u}{\partial z}\right) + f\left(v - v_g\right)$$

$$\frac{\partial v}{\partial t} = \frac{\partial}{\partial z}\left(K_m \frac{\partial v}{\partial z}\right) - f\left(u - u_g\right) \tag{7-22}$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left(K_h \frac{\partial \theta}{\partial z}\right)$$

in which the parameter f represents the Coriolis frequency and u_g and v_g represent the geostrophic wind components. The K_m and K_h are the momentum and heat eddy diffusivities, which are simulated by serving a prognostic equation for turbulent kinetic energy, shown by q^2 (23) and a diagnostic equation for turbulent mixing length scale l (7-24).

$$\frac{\partial q^2}{\partial t} = \frac{\partial}{\partial z}\left(K_e \frac{\partial q^2}{\partial z}\right) + 2K_e M^2 - 2K_h N^2 - \frac{2q^3}{16.6l} \tag{7-23}$$

where $K_e = 0.2lq$ is the eddy diffusivity of the turbulent kinetic energy equation and l is the turbulent mixing length, calculated as

$$\frac{1}{l} = \frac{1}{\kappa z} + \frac{1}{\lambda} \tag{7-24}$$

Note that λ is a limiting mixing length, which is set to 10m (Beare et al. 2006) and κ is the Von Kàrmàn constant ($\kappa = 0.4$). The parameters M and N represent the Prandtl and Brunt-Väisälä frequencies, respectively, and following Blackadar (1962) will be simulated as,

$$M^2 = \left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2$$

(7-25)

$$N^2 = \frac{g}{\theta_0}\frac{\partial \theta}{\partial z}$$

where g represents the gravity constant and θ_0 is the reference temperature ($\theta_0 = 273$K). The parameters K_m and K_h will be calculated as (Galperin et al., 1988)

$$K_m = lqS_m$$
$$K_h = lqS_h \tag{7-26}$$

where S_m and S_h are called dimensionless stability functions and parametrized as,

$$S_m = \frac{0.393 - 3.085G_h}{1 - 40.803G_h + 212.469{G_h}^2}$$
$$S_h = \frac{0.494}{1 - 34.676G_h} \tag{7-27}$$

where G_h is a function of the Brunt-Väisälä frequency,

$$G_h = -\frac{l^2N^2}{q^2} \tag{7-28}$$

The one dimensional vertically diffusive turbulent boundary layer flow equations, (7-22), were discretized using central spatial scheme, on a uniform grid network, with grid size $\Delta z = 20$ m, in the range of $z \in [0,400]$ m.

The potential temperature at the surface initially is assumed as 265 K and it remains fixed up to the height $z = 100$ m, then forward, it is going to increase with the rate 0.01 Km^{-1}. In terms of the boundary conditions, the potential temperature is assumed constant at top, i.e. $\theta_T = 268$ K and a cooling rate 0.25 Khr^{-1} is imposed at the surface. The initial wind speed is set to the geostrophic wind velocities with magnitudes as $u_g = 8$ ms^{-1} and $v_g = 0$. The no slip boundary condition was applied at the surface while at the top level the velocity components were set equal to the geostrophic values. As recommended by Mellor and Yamada (1974), the turbulent kinetic energy at the surface will be calculated as,

$$q_0^2 = B_1^{2/3}u_s^2 \tag{7-29}$$

in which $\mathrm{u_s}$ represents the friction velocity and will be calculated using

$$\mathrm{u_s} = -\sqrt{\mathrm{K_m}\left|\frac{\partial \mathrm{u}}{\partial \mathrm{z}}\right|} \tag{7-30}$$

The turbulent kinetic energy variable, q^2, initially is set to $0.4\,\mathrm{m^2s^{-2}}$ at the surface which is varying linearly up to the level $\mathrm{z} = 250\,\mathrm{m}$ and then it is fixed to a minimum value $5 \times 10^{-7}\,\mathrm{m^2s^{-2}}$. Note that due to stability concern the minimum value $1 \times 10^{-5}\,\mathrm{m^2s^{-1}}$, representing the molecular diffusivity, is added to the diffusion coefficients in all equations (7-22). Note that the Coriolis frequency is set to $f = 1.39 \times 10^{-4} s^{-1}$.

The governing equation is temporally integrated using SSP(2,2) and employing different approaches, including the weighted-SI-RK method, as described in Section 7.2. For all three approaches the vertical gradients will be calculated implicitly. For the weighted-SI-RK method, the non-linear coefficients, including the diffusion coefficients and the source terms, shown in (7-22) and (7-23), will be calculated, using the weighted average combination of the solutions in the current internal stage and the solution at the previous time step, i.e. Y_i and X^n. Note that the weighted averaging coefficients, corresponding to $P = 4$ provided in Table 7-2, were utilized in this computation. For the SE-RK, these non-linear coefficients will be calculated explicitly using the solutions at the previous time step and for the FI-RK method they will be calculated implicitly.

In order to compare the new approach with one introduced by Dimantakis et al. (2006), this scheme was also employed to calculate the equation set as well.

Shown in figure 7-4 is the distribution of the eastward velocity component and the error distribution, which is defined as the difference between solutions obtained by the nominated approaches and the reference solution calculated using an explicit Runge-Kutta scheme, proposed by Rokhzadi et al. (2018) employing very small time step resolution, i.e. $\Delta t = 1.0s$. The result was compared to the LES solution taken from Beare et al. (2005). Note that in all cases the grid resolution is $\Delta z = 10m$ and the time step size is $\Delta t = 25s$.

From the left graph, clearly, the height level, where the nocturnal jet, or super-geostrophic jet, occurs is over predicted, compared to the LES solution, which may be due to the nature of the turbulence modeling. However, the introduced approaches could reproduce the general features of the air flow including predicting the jet strength. The error distribution, illustrated in the right column, shows that the weighted-SI-PC method is less accurate than the other ones. Before the nocturnal jet occurs, the weighted-SI-RK improves the quality of the solution, compared to the FI-RK scheme, with less significant deviation from the SE-RK scheme, whilst, after the nocturnal jet the FI-RK scheme behaves more accurately.

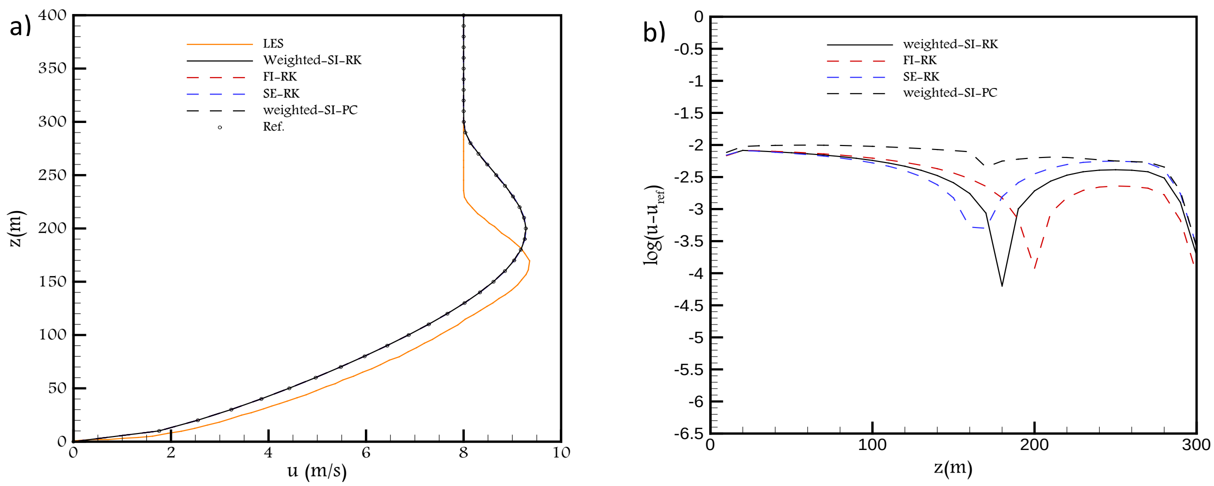


Figure 7-4: The distribution of the u component (a) and its error distribution, $u - u_{ref}$, (b) in the log scale, of the vertical diffusive turbulent boundary layer flow, after 8 hours, calculated by SSP(2,2) and serving the time step $\Delta t = 25\ s$.

Figure 7-5 illustrates the distribution of the potential temperature, θ, and the numerical error, $\theta - \theta_{ref}$, which is defined as the difference between solutions obtained by the nominated approaches and the reference solution calculated by an explicit Runge-Kutta scheme, proposed by Rokhzadi et al. (2018) employing very small time step resolution, i.e. $\Delta t = 1.0s$. As can be seen in the left graph, although there are some discrepancies between LES solutions and the ones obtained by the turbulence modelling, they still are able to produce the general features of the air flow. Referring to the right graph, it is clear that before the nocturnal jet all schemes can calculate the solution with the same quality, but, more accuracy is for the weighted-SI-PC and SE-RK schemes. However, afterward, the FI-RK shows less over prediction and the weighted-SI-RK scheme is more accurate than the weighted-SI-PC and SE-RK schemes.

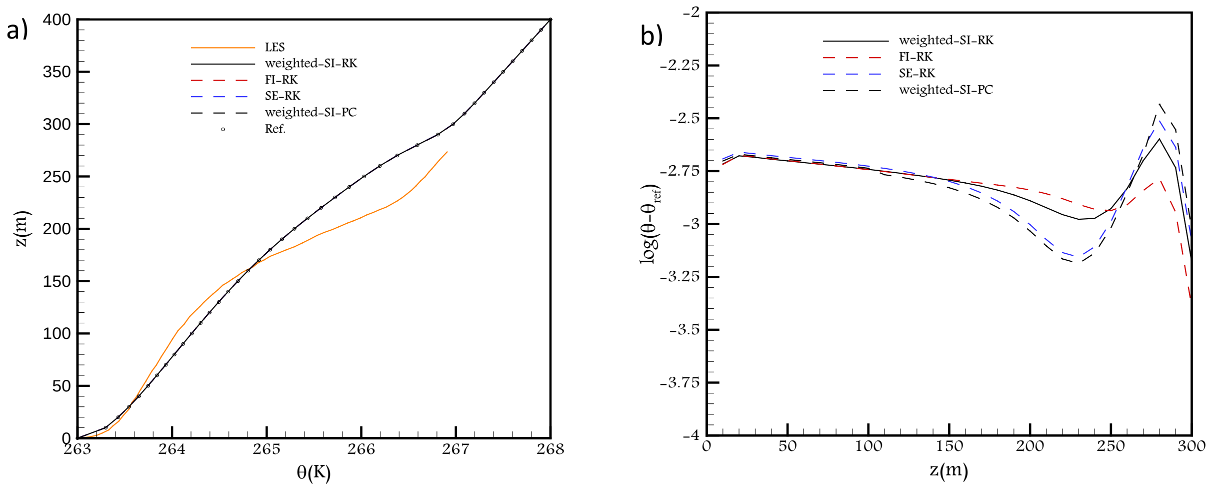


Figure 7-5: The potential temperature distribution of the vertical diffusive turbulent boundary layer flow, θ, (a) and its error distribution, $\theta - \theta_{ref}$, (b) in log scale, after 8 hours, calculated by SSP(2,2) and serving the time step $\Delta t = 25\ s$.

Provided in the figure 7-6 is the CPU time usage, on the vertical axis, by different approaches against different stable time step sizes, $10 \leq \Delta t \leq 25s$. Noticeably, the proposed weighted-SI-PC scheme takes less calculation time, while the other approaches of the Runge-Kutta scheme are almost the same. However, considering the more accuracy, provided by the weighted-SI-RK approach, one may find this scheme more beneficial than other schemes.

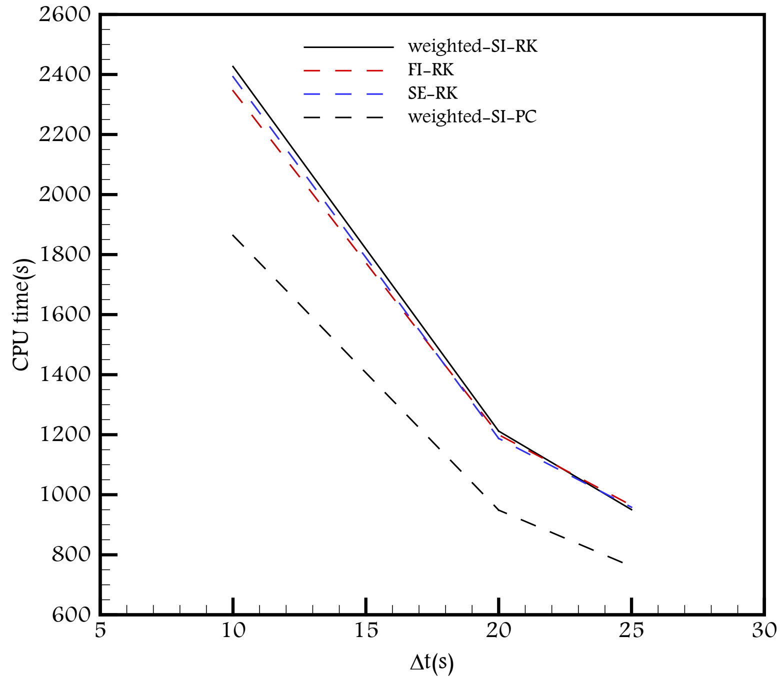


Figure 7-6: The CPU time of the calculation, against various stable time step, Δt, calculated by SSP(2,2), weighted-SI-RK (black solid), FI-RK (red dashed) and SE-RK (blue dashed).

7.6. Conclusion

Diagonally implicit Runge-Kutta scheme has been undertaken to temporally integrate the vertical diffusion boundary layer equation with different techniques, including fully implicit, so called FI-RK and common semi-implicit approach, called SE-RK schemes. Meanwhile, a new weighted-averaging method (weighted-SI-RK) was proposed. The linear stability analysis showed that the SE-RK approach is less damping than other schemes; hence, it calculates the result more accurately, the fact that the numerical computation of the representative nonlinear equation, proposed by Kalnay and Kanamitsu (1988), proved. However, this test showed that the SE-RK is less stable. The FI-RK scheme is impractical in real cases and it is too expensive due to the more CPU time usage. The proposed weighted-SI-RK showed the ability to improve the range of stable time step sizes, compared to the SE-RK, and to increase the accuracy compared to FI-RK and the proposed semi-implicit method by Diamantakis et al. (2006) (weighted-SI-PC). Therefore, the longer CPU time usage of weighted-SI-RK, compared to the weighted-SI-PC could be compensated by considering the relevant accuracy improvement. The benefits of using the proposed weighted averaging approach were also approved through solving the vertical diffusion turbulent boundary layer flow using the $E - l$ turbulence modelling.

Acknowledgment

The authors wish to acknowledge financial support from Natural Sciences and Engineering Research Council of Canada (NSERC).

References:

D. S. Abdi, F. X. Giraldo, E. M. Constantinescu, L. E. Carr, L. C. Wilcox and T. C. Warburton, Acceleration of the Implicit-Explicit Non-hydrostatic Unified Model of the Atmosphere (NUMA) on Manycore Processors, Int J High Perform Comput Appl, Accepted August 2017.

R. Alexander, Diagonally Implicit Runge-Kutta Methods for Stiff O.D.E.'S, Siam J. Numer. Anal., Vol. 14, No. 6, 1977.

R. Alt, Méthodes A-stables pour l'Intégration de Systémes Differentiels mal Conditionnés, Thesis (3rd cycle), Université de Paris 6 (1971). [In French]

R. Alt and F. Ceschino, Mise en æuvre de schémas A-stables du type de Runge-Kutta pour l'integration des systéms différentieles, C.R. Acad. Sci. Paris, Ser. A, 274(10) (1972) 846-849. [In French]

U. Ascher, S. Ruuth, R. Spiteri, Implicit–explicit Runge–Kutta methods for time-dependent partial differential equations, Appl. Numer. Math., 25 (2–3) (1997) 151–167.

P., Bénard, A., Marki, , P. N., Neytchev & M. T. Prtenjak, (2000). Stabilization of nonlinear vertical diffusion schemes in the context of NWP models. Monthly Weather Review, 128(6), 1937-1948.

H. Bijl, M.H. Carpenter, V.N. Vatsa and C.A. Kennedy, Implicit time integration schemes for the unsteady compressible Navier-Stokes equations: Laminar flow, J. Comp. Phys., 179(1) (2002) 313-329.

H. Bijl, S.J. Hulshoff and S. van Zuijlen, Accuracy and effciency of implicit time integration schemes for fluid-structure interaction simulations, In: Proc. of Fifth World Congress on Comp. Mech. (WCCM V), July 7-12, 2002, Vienna, Austria, Eds., H.A. Mang, F.G. Rammerstorfer and J. Eberhardsteiner, Vienna Univ. of Tech., Austria (2002) 11 pp. [ISBN 3-9501554-0-6, Paper-ID: 80695.

S. Boscarino, Error Analysis of IMEX Runge–Kutta Methods Derived from Differential-Algebraic Systems, SIAM J. Numer. Anal., 45 (4) (2007) 1600–1621.

S. Boscarino and G. Russo, On a class of uniformly accurate IMEX Runge-Kutta schemes and applications to hyperbolic systems with relaxation, SIAM J. Sci. Comp., 31(3) (2009) 1926-1945.

S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, ISBN 0 521 83378 7, 2004.

B. Bujanda and J.C. Jorge, Order conditions for linearly implicit fractional step Runge-Kutta methods, IMA J. Numer. Anal. 27(4) (2007) 781-797.

J. Burns, A. Balogh, D. S. Gilliam, V. I. Shubov, Numerical stationary solutions for a viscous Burgers's equation, Journal of Mathematical Systems, Estimation, and Control 8 (2) (1998) 1–16.

K. Burrage, A Special Family of Runge-Kutta Methods for Solving Stiff Differential Equations, BIT Numer. Math, 18 (1978) 22-41.

K. Burrage, R.P.K. Chan, On smoothing and order reduction effects for implicit Runge-Kutta formulae, Journal of Comput and Applied Math. 45 (1993) 17-27.

J.C. Butcher, Order and stability of generalized Padé approximations, Appl. Numer. Math., 59(3-4) (2009) 558-567.

J.C. Butcher, Coefficients for the study of Runge-Kutta integration processes, J. Australian Math. Soc., 3(2) (1963) 185-201.

J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods, John Wiley and Sons, Chichester (1987).

C., Clancy and J. A., Pudykiewicz, A class of semi-implicit predictor-corrector schemes for the time integration of atmospheric models, 250 (2013) 665-684.

M.P. Calvo, J. de Frutos and J. Novo, Linearly implicit Runge-Kutta methods for advection-reaction-diffusion equations, Appl. Numer. Math., 37(4) (2001) 535-549.

F. Cameron, M. Palmroth and R. Piché, Quasi stage order conditions for SDIRK methods, Appl. Numer. Math., 42(1-3) (2002) 61-75.

M. Carpenter and C. Kennedy, Additive Runge-Kutta Schemes for Convection-Diffusion-Reaction Equations, Appl. Numer. Math., 44 (2003) 139–181.

T., Davies, M. J. P., Cullen, A. J., Malcolm, M. H., Mawson, A., Staniforth, A. A. White, and N. Wood, 2005, A new dynamical core for the Met Office’s global and regional modelling of the atmosphere. Q. J. R. Meteorol. Soc., 131, 1759–1782.

C. Girard, Y. Delage, Stable schemes for nonlinear vertical diffusion in atmospheric circulation models, Month. Weath. Rev. 118 (1990) 737.

E.M. Constantinescu and A. Sandu, Multirate Timestepping Methods for Hyperbolic Conservation Laws, J Sci Comput 33 (2007), 239–278.

M. Crouzeix, Ph.D. thesis, Paris: Universite Paris VI; 1975. Sur l’approximation des equations differentielles operationelles lineaires par des methodes de Runge Kutta.

K. Dekker and J.G. Verwer, Stability of Runge-Kutta Methods for Stiff Non-linear Differential Equations, North-Holland, Amsterdam, Netherlands (1984).

M. Diamantakis, N. Wood and T. Davies, An improved implicit predictor–corrector scheme for boundary layer vertical diffusion, Q. J. R. Meteorol. Soc. (2006), 132, 959–978.

Y. Du and J. A. Ekaterinaris, Optimization of the Diagonally Implicit Runge-Kutta Methods for

Convection Diffusion Equations, under review, AIAA Journal (2016).

D. Durran, P. Blossey, Implicit–explicit multistep methods for fast-wave–slow-wave problems, Mon. Weather Rev. 140 (4) (2012) 1307–1325.

D. Durran, Numerical Methods for Fluid Dynamics with Applications to Geophysics, 2nd edition, Texts Appl. Math., vol. 32, Springer, ISBN 1441964118, 2010.

L. Ferracina, M. N. Spijker, Strong stability of singly-diagonally-implicit Runge–Kutta methods, J. Applied Numerical Mathematics 58 (2008) 1675–1686.

J. de Frutos, Implicit-explicit Runge-Kutta methods for financial derivatives pricing models, Euro. J. Oper. Res., 171(3) (2006) 991-1004.

B., Galperin, L.H., Kantha, S., Hassid and S., Rosati, A quasi-equilibrium turbulent energy model for geophysical flows, J.Atmos.Sci., 45 (1988), 55.

D. J. Gardner , J. E. Guerra , F. P. Hamon , D. R. Reynolds, P. A. Ullrich and C. S. Woodward, Implicit–explicit (IMEX) Runge-Kutta methods for non-hydrostatic atmospheric models, J. Geosci. Model Dev., 11 (2018), 1497–1515.

D., Ghosh and E. M., Constantinescu, (2016). Well-balanced, conservative finite difference algorithm for atmospheric flows. AIAA Journal, 54(4), 1370–1385.

F. Giraldo, J. Kelly and E. Constantinescu, SIAM Journal on Scientific Computing, 35(2013), B1162-B1194.

F. Giraldo, J. Kelly and E. Constantinescu, Implicit-Explicit Formulations of A Three-Dimensional Non-hydrostatic Unified Model of the Atmosphere (Numa), SIAM Journal on Scientific Computing, 35(2013), B1162-B1194.

C. Girard and Y. Delage, 1990 Stable schemes for nonlinear vertical diffusion in atmospheric circulation models. Mon. Weather Rev., 118, 737–745.

I. Gómez, I. Higueras and T. Rold_an, Starting algorithms for low stage order RKN methods, J. Comp. and Appl. Math., 140(1-2) (2002) 345-367.

S. Gottlieb and C.-W. Shu, Total variation diminishing Runge-Kutta schemes, Math. Comp., 67 (1998), 73–85.

D. F. Griffiths, A. M. Stuartt, and H. C. Yee, Numerical Wave Propagation in an Advection Equation with a Nonlinear Source Term, SIAM J. Numer. Anal., 29 (5) (1992), 1244-1260.

E. Hairer, S.P. Nørsett and G.Wanner, Solving Ordinary Differential Equations I, Nonstiff Problems, 2ed., Springer-Verlag, Berlin (1993).

E. Hairer and G. Wanner, Solving Ordinary Differential Equations II, Springer-Verlag, ISBN 978-3-642-05221-7, 1996.

E. Hairer and G.Wanner, Stiff differential equations solved by Radau methods, J. Comp. and Appl. Math., 111(1-2) (1999) 93-111.

I. Higueras, Strong stability for additive Runge-Kutta methods, SIAM J. Numer. Anal., 44(4) (2006) 1735-1758.

I. Higueras, N. Happenhofer, , O. Koch, , F. Kupka, Optimized strong stability preserving IMEX Runge–Kutta Methods, J. Comput. & Appl. Math., 272 (2014) 116–140.

I. Higueras and T. Roldán, Positivity-Preserving and Entropy-Decaying IMEX Methods, Monografías del Seminario Matemático García de Galdeano 33, 129–136 (2006).

F. Hu, M. Hussaini and J. Manthey, Low-Dissipation and Low-Dispersion Runge–Kutta Schemes for Computational Acoustics, J. Comput. Phys., 124 (1996) 177–191.

W. Hundsdorfer, S. J. Ruuth, IMEX extensions of linear multistep methods with general monotonicity and boundedness properties, J. Comput. Phys., 225 (2007) 2016–2042

A. Iserles, A. Zanna, Preserving Algebraic Invariants with Runge–Kutta Methods, J. Comput. & Applied Math., 125 (1–2) (2000) 69–81.

L. G. Ixaru, (2012). Runge–Kutta method with equation dependent coefficients. Computer Physics Communications, 183(1), 63-69.

E. Kalnay, and M. Kanamitsu, Time schemes for strongly nonlinear damping equations, J. Mon. Weather Rev., 116 (1988), 1945–1958.

S. K., Kar, A Semi-Implicit Runge–Kutta Time-Difference Scheme for the Two-Dimensional Shallow-Water Equations, J. Mon. Weather Rev., 134 (2006), 2916-2926.

O. A., Karakashian and W. Rust, On the parallel implementation of implicit Runge-Kutta methods. SIAM journal on scientific and statistical computing, 9(6) (1988), 1085-1090.

A. Kennedy, M.H. Carpenter, Diagonally Implicit Runge-Kutta Methods for Ordinary Differential Equations. A Review, NASA/TM-2016-219173.

D.I. Ketcheson, An algebraic characterization of strong stability preserving Runge–Kutta schemes, B.Sc. thesis, Brigham Young University, Provo, Utah, USA, 2004.

D. Ketcheson, C. Macdonald, S. Gottlieb, Optimal Implicit Strong Stability Preserving Runge–Kutta Methods, Appl. Numer. Math 59 (2009) 373–392.

J.F.B.M. Kraaijevanger, Contractivity of Runge-Kutta methods, BIT, 31(3) (1991) 482-528.

M.A. Kurdi, Stable High Order Methods for Time Discretization of Stiff Differential Equations, Ph.D. Thesis, Applied Mathematics, Univ. of California, Berkeley, Berkeley (1974).

A. Kværnø, Singly Diagonally Implicit Runge–Kutta Methods with an Explicit First Stage, BIT Numer. Math, 44 (2004) 489–502.

J.D. Lambert, Numerical Methods for Ordinary Differential Systems, The Initial Value Problem, John Wiley and Sons, Chichester, 1991.

R. J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations, SIAM, ISBN 978-0-89871-629-0, 2007.

H. Liu and J. Zou, Some new additive Runge-Kutta methods and their applications, J. Comp. and Appl. Math., 190(1-2) (2006) 74-98.

S. Lock, N. Wood and H. Weller, Numerical Analyses of Runge–Kutta Implicit–Explicit Schemes for Horizontally Explicit, Vertically Implicit Solutions of Atmospheric Models, Q. J. R. Meteorol., (140) (2014) 1654–1669.

J. M. Mantas, P. Gonzlez and J. A. Carrillo, Parallelization of implicit-explicit runge-kutta methods for cluster of PCs. In Euro-Par 2005 Parallel Processing, Springer Berlin Heidelberg (2005), 815-825.

L. Murray, GPU acceleration of Runge-Kutta integrators. Parallel and Distributed Systems, IEEE Transactions on, 23(1) (2012), 94-101.

A. Najafi-Yazdi, L. Mongeau, A Low-Dispersion and Low-Dissipation Implicit Runge-Kutta Scheme, J. Comput. Phys., 233 (2013) 315-323.

F. Nazari, A. Mohammadian, A. Zadra, M. Charron, A stable and accurate scheme for nonlinear diffusion equations: Application to atmospheric boundary layer, J. Comput. Phys., 236 (2013) 271-288.

F. Nazari, A. Mohammadian, M. Charron, and A. Zadra, Optimal high-order diagonally-implicit Runge–Kutta schemes for nonlinear diffusive systems on atmospheric boundary layer, J. Comput. Phys., 271 (2014) 118–130.

F. Nazari, A. Mohammadian, M. Charron, High-Order Low-Dissipation Low-Dispersion Diagonally Implicit Runge–Kutta Schemes, J. Comput. Phys., 286 (2015) 38–48.

F. Nazari and R. D., Nair, (2017). A Godunov type finite volume solver for nonhydrostatic Euler equations with a time-splitting approach. Journal of Advances in Modeling Earth Systems, 9, 465–481.

S.P. Nørsett and G. Wanner, The real-pole sandwich for rational approximations and oscillation equations, BIT, 19(1) (1979) 79-94.

S. Nørsett and A. Wolfbrandt, Attainable Order of Rational Approximations to the Exponential Function with Only Real Poles, BIT Numer. Math, 17 (1977) 200-208.

S.P. Nørsett and S.R. Trickett, Exponential fitting of restricted rational approximations to the exponential function, In: Rational Approximation and Interpolation, Lecture Notes in Mathematics, Vol. 1105, Eds: P. Russell Graves-Morris, E.B. Saff and R.S. Varga, Springer-Verlag, Berlin (1984) 466-476.

C. Pantano, An additive semi-implicit Runge-Kutta family of schemes for nonstiff systems, Appl. Numer. Math. 57(3) (2007) 297-303.

L. Pareschi and G. Russo, High Order Asymptotically Strong-Stability-Preserving Methods for Hyperbolic Systems with Stiff Relaxation in Hyperbolic Problems: Theory, Numerics, Applications, Springer-Verlag, Berlin, (2003) 241–251.

L. Pareschi and G. Russo, Implicit–Explicit Runge–Kutta Schemes and Applications to Hyperbolic Systems with Relaxation, J. Sci. Comput., 25 (1) (2005) 129–155.

A. Sayfy and M. Al-Refai, Additive Runge-Kutta methods with improved stability properties, Int. J. Diff. Eqns. and Appl., 4(4) (2002) 387-400.

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

J. Strikwerda, Finite Difference Schemes and Partial Differential Equations, 2nd edition, SIAM, Philadelphia, ISBN 0-89871-567-9, 2004.

R. B. Stull, An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, ISBN 9027727686, 1988.

P. Ullrich, C. Jablonowski, Operator-split Runge–Kutta–Rosenbrock methods for nonhydrostatic atmospheric models, Mon. Weather Rev., 140 (4) (2012), 1257–1284.

A. van Zuijlen and H. Bijl, Implicit and explicit higher order time integration schemes for fluid-structure interaction computations, In: Computational Science - ICCS 2004, Part IV, Lecture Notes in Computer Science, Vol. 3039, M. Bubak, G.D.v. Albada, P.M.A. Sloot and J. Dongarra, Eds., Springer-Verlag, Berlin (2004) 604-611.

J. G. Verwer, (1996). Explicit Runge-Kutta methods for parabolic partial differential equations. Applied Numerical Mathematics, 22(1), 359-379.

R. Wang and R. J. Spiteri, Linear Instability of the Fifth-Order WENO Method, SIAM J. Numer. Anal., 45(5) (2007), 1871–1901.

R. Wang and R. J. Spiteri, Linear Instability of the Fifth-Order WENO Method, SIAM J. Numer. Anal., 45(5) (2007), 1871–1901.

G. Wanner, E. Hairer and S. P. Nørsett, Order stars and stability theorems, BIT, 18(4) (1978a) 475-489.

H. Weller, S. Lock, N. Wood, Runge–Kutta IMEX schemes for the Horizontally Explicit/Vertically Implicit (HEVI) solution of wave equations, J. Comput. Phys., 252 (2013) 365–381.

J. S. Whitaker and S. K. Kar, Implicit–Explicit Runge–Kutta Methods for Fast–Slow Wave Problems, Mon. Weather Rev., 141 (2013), 3426-3434.

J. J. Yoh and X. Zhong, New hybrid Runge-Kutta methods for unsteady reactive ow simulation, AIAA J., 42(8) (2004) 1593-1600.

J. J. Yoh and X. Zhong, New hybrid Runge-Kutta methods for unsteady reactive flow simulation: Applications, AIAA J., 42(8) (2004) 1601-1611.

X. Zhong, Additive Semi-Implicit Runge–Kutta Methods for Computing High-Speed Nonequilibrium Reactive Flows, J. Comput Phys., 128, 19–31 (1996).