UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Numerical solution of advection-diffusion and convective Cahn-Hilliard equations

by

Hagos Hailu Gidey

Submitted in partial fulfillment of the requirements for the degree

Philosophiae Doctor

in the Department of Mathematics and Applied Mathematics

in the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

November 14, 2016

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Declaration

I, the undersigned, declare that the thesis which I hereby submit for the degree Philosophiae Doctor at the University of Pretoria, is my own, independent work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature:

Name: Hagos Hailu Gidey

Date: November 14, 2016

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

This thesis is dedicated to my late daughter Betarik.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Acknowledgements

It is my great pleasure to express my heartfelt gratitude appreciation to my supervisor, Dr Rao Appadu and my co-supervisor, Prof. Djoko Jules for their expertise, encouragement, continuous care, immeasurable support and intellectual guidance. Without their help, this thesis would not have been possible to take this form.

My acknowledgement and great appreciation also go to all the staff members of the Department of Mathematics and Applied Mathematics, University of Pretoria. Special mention goes to Prof. Roumen Anguelov for his positive support and willingness to assist with requests for assistance to attend conferences and Mrs Yvvone McDermot for her kind support during my application process to the University of Pretoria.

I also acknowledge all my friends and colleagues in the Department. Their support and encouragement were vital for the completion of this work. Special thanks goes to the following friends: Mr Tesfalem, Ms Claire, Mr Abey, Mr Bahru, Mr Basit, Mr Gediyon, Mr Jejeniwa, Dr Nega, Mr Tadele, Mr David, Mr Sesuai, Mr Koffi, Mr Usman, Ms Tigist, Dr Belayneh, Dr Yibeltal, Dr Zakaria, Dr Joseph and all other friends whom you all deserve to be appreciated.

I thank all the staff at AIMS who encouraged me to do my PhD. Thanks to Prof. Barry Green and Prof Jeff Sanders from AIMS-South Africa for their insightful guidance, indispensable advice and encouragement. I would also like to address my sincere thanks to Prof. Habtu Zegeye from International Institute of Science and Technology, Botswana and Dr Taddesse Zegeye from Bahir Dar University, Ethiopia who encouraged me to undergo

further studies.

My studies would not have been possible without the bursary provided by the University of Pretoria, AIMS-South Africa for the duration of my studies. My acknowledgement and appreciation go to DST-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS) for funding for the academic year 2016. I am grateful to Aksum University for granting me study leave.

I am grateful to my father Melake-Hiwot Hailu, to my brothers and sisters and to all my relatives who had assisted me in many ways for their kindness and unconditional love, encouragement and support throughout my entire life journey. It is because of their support and care that I have arrived at this stage.

Most special acknowledgement and profound thanks go to my wife's parents Mr Yemane Degnau and Alem Gobezay and her sister Mulu Yemane for their love and support, particularly in taking care of my lovely son.

My warm appreciation, love and respect goes to my lovely wife, Hewan Yemane, and my lovely son, Daniel Hagos, who have been sources of my strength and courage against all the hardships throughout my study. I would also like to thank everyone who made this possible. I couldn't have done it without you.

Finally, and most importantly, special thanks to the Ethiopian community in South Africa and other friends and staff of the Department of Mathematics and Applied Mathematics for your prayers, sympathy and kind support during our time of loss. Your very kind involvement and contribution in all aspects gave us a sense of strength and humour and it is godly gift to know that you are with us as we grieve our daughter's death.

Title	Numerical solution of advection-diffusion and convective Cahn-Hilliard equations
Name	Hagos Hailu Gidey
Supervisor	Dr A. R. Appadu
Co-supervisor	Prof. J. K. Djoko
Department	Mathematics and Applied Mathematics
Degree	Philosophiae Doctor

Abstract

In this PhD thesis, we construct numerical methods to solve problems described by advection-diffusion and convective Cahn-Hilliard equations. The advection-diffusion equation models a variety of physical phenomena in fluid dynamics, heat transfer and mass transfer or alternatively describing a stochastically-changing system. The convective Cahn-Hilliard equation is an equation of mathematical physics which describes several physical phenomena such as spinodal decomposition of phase separating systems in the presence of an external field and phase transition in binary liquid mixtures (Golovin et al., 2001; Podolny et al., 2005).

In chapter 1, we define some concepts that are required to study some properties of numerical methods. In chapter 2, three numerical methods have been used to solve two problems described by 1D advection-diffusion equation with specified initial and boundary conditions. The methods used are the third order upwind scheme (Dehghan, 2005), fourth order scheme (Dehghan, 2005) and Non-Standard Finite Difference scheme (NSFD) (Mickens, 1994). Two test problems are considered. The first test problem has steep boundary layers near the region $x = 1$ and this is challenging problem as many schemes are plagued by non-physical oscillation near steep boundaries. Many methods suffer from computational noise when modelling the second test problem especially when the coefficient of diffusivity is very small for instance 0.01. We compute some errors, namely L_2 and L_∞ errors, dissipation and dispersion errors, total variation and the total mean square error for both problems and

compare the computational time when the codes are run on a matlab platform. We then use the optimization technique devised by Appadu (2013) to find the optimal value of the time step at a given value of the spatial step which minimizes the dispersion error and this is validated by some numerical experiments.

In chapter 3, a new finite difference scheme is presented to discretize a 3D advection-diffusion equation following the work of Dehghan (2005, 2007). We then use this scheme and two existing schemes namely Crank-Nicolson and implicit Chapeau function to solve a 3D advection-diffusion equation with given initial and boundary conditions. We compare the performance of the methods by computing L_2- error, L_∞-error, dispersion error, dissipation error, total mean square error and some performance indices such as mass distribution ratio, mass conservation ratio, total mass and R^2 which is a measure of total variation in particle distribution. We also compute the rate of convergence to validate the order of accuracy of the numerical methods. We then use optimization techniques to improve the results from the numerical methods.

In chapter 4, we present and analyze four linearized one-level and multilevel (Bousquet et al., 2014) finite volume methods for the 2D convective Cahn-Hilliard equation with specified initial condition and periodic boundary conditions. These methods are constructed in such a way that some properties of the continuous model are preserved. The nonlinear terms are approximated by a linear expression based on Mickens' rule (Mickens, 1994) of nonlocal approximations of nonlinear terms. We prove the existence and uniqueness, convergence and stability of the solution for the numerical schemes formulated. Numerical experiments for a test problem have been carried out to test the new numerical methods. We compute L_2-error, rate of convergence and computational (CPU) time for some temporal and spatial step sizes at a given time. For the 1D convective Cahn-Hilliard equation, we present numerical simulations and compute convergence rates as the analysis is the same with the analysis of the 2D convective Cahn-Hilliard equation.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Nomenclature

Δt	time step size
Δx	spatial step size on x- direction
h	spatial step size
β	advection velocity
β_x	advection velocity in the direction of x
β_y	advection velocity in the direction of y
β_z	advection velocity in the direction of z
α	diffusivity constant
α_x	diffusion coefficient in the direction of x
α_y	diffusion coefficient in the direction of y
α_z	diffusion coefficient in the direction of z
$\mathbf{D}$	diffusivity tensor matrix
c	Courant number
c_x	Courant number in the direction of x
c_y	Courant number in the direction of y
c_z	Courant number in the direction of z
Re	Reynolds number
$\forall$	for all
γ	driving force
ε	dimensionless interfacial width

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Contents

CONTENTS

CONTENTS

CONTENTS

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

List of Figures

LIST OF FIGURES

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

List of Tables

LIST OF TABLES

LIST OF TABLES

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

List of Publications

Journals published/accepted

A. R. Appadu, J. K. Djoko and H. H. Gidey, A computational study of three numerical methods for some advection-diffusion problems, Applied Mathematics and Computation 272 (2016), 629-647.

A.R. Appadu and J.K.Djoko, H.H. Gidey and J. M. S. Lubuma, Analysis of Multilevel Finite Volume Approximation of 2D Convective Cahn-Hilliard Equation, published online in Japan Journal of Industrial and Applied Mathematics, doi:10.1007/s13160-017-0239-y.

Proceedings published/accepted

A. R. Appadu, J. K. Djoko and H. H. Gidey, A computational study of some numerical schemes for a test case with steep boundary layers, Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2014 (ICNAAM-2014), Vol. 1648, No. 1, AIP Publishing, 2015.

A. R. Appadu, J. K. Djoko and H. H. Gidey, Comparative study of some numerical methods to solve a 3-D advection-diffusion equation. To appear in Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2016 (ICNAAM-2016), AIP Publishing, (Rhodes, Greece).

Papers submitted for publication

A.R. Appadu and J.K.Djoko and H.H. Gidey, Performance of some finite difference methods for a 3D advection-diffusion equation, submitted to the Journal Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas.

A.R. Appadu and J.K.Djoko and H.H. Gidey, Analysis of Multilevel Finite Volume Approximation of Convective Cahn-Hilliard Equation, submitted to Afrika Mathematica.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Chapter 1

Introduction

In this thesis, some finite difference and finite volume methods have been used to discretize the advection-diffusion and convective Cahn-Hilliard equations, respectively.

Partial differential equations are used to describe a wide variety of phenomena such as physical, chemical and biological phenomena and are also used for economics, meteorological and financial forecasting and other fields. The analytical solutions of partial differential equations are not generally easy to find. Hence it is necessary to apply numerical methods to solve approximate solutions of these partial differential equations in order to investigate or forecast the predictions of the mathematical models.

Numerical methods are techniques by which mathematical problems are formulated so that they can be solved with arithmetic operations, which describe a particular real problem. There are many types of numerical methods used to solve ordinary/partial differential equations. Finite difference, finite volume, finite element, collocation, B-spline and cubic-splines are some of the most commonly used numerical methods to solve partial differential equations.

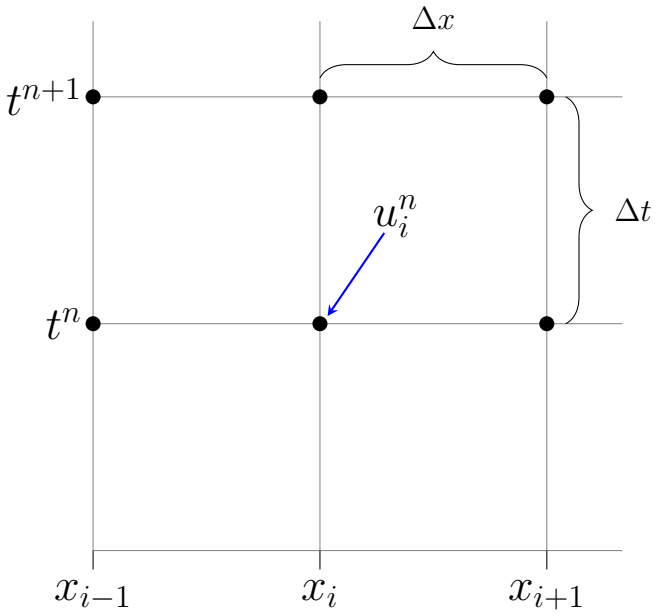


Figure 1.1: One dimensional finite difference discretization

1.1 Preliminaries and some important concepts

In this section, we recall some definitions and important concepts which are helpful for our study.

1.1.1 Finite difference methods

Finite difference methods for solving differential equations are numerical methods that proceed by transforming the differential equation into a large algebraic system of equations called difference equations. The domain is first partitioned into space and time and approximations of the solution are computed at the spacial and temporal grid points (called point wise approximation) as shown in Fig. 1.1. For 1D partial differential equation on the interval $(a, b), a < b$, the crossing points $(x_i = i\Delta x, t_n = n\Delta t), i = 0, 1, \ldots, N, n = 0, 1, \ldots,$ are called the grid points where $\Delta x = \dfrac{b-a}{N}$ and Δt is time step size. The finite difference solution, denoted by u_i^n, is the approximation of the solution at the grid point (x_i, t_n).

1.1.2 Finite volume methods

We start with the one dimensional conservation law of the form (for more details see Leveque (2004))

$$u_t + g(u)_x = 0, \tag{1.1.1}$$

where u is a conserved quantity and g is some flux function. Rather than point wise approximation at grid points, like finite difference methods, a finite volume method is based on subdividing the spatial domain into intervals called grid cells/finite volumes and approximating the integral of u over each of these grid cells. The approximate value of the average value over the i^{th} grid cell at time t_n is given by

$$u_i^n = \frac{1}{\Delta x}\int_{k_i} u(x,t_n)\,dx, \tag{1.1.2}$$

where $k_i = [x_{i-1/2}, x_{i+1/2}]$ is the i^{th} grid cell, $\Delta x = x_{i+1/2} - x_{i-1/2}$, see Fig. 1.2.

If $u(x,t)$ is a smooth function, then the integral in (1.1.2) agrees with the value of u at the mid point of the interval to $\mathcal{O}(\Delta x^2)$.

The integral form of the conservation law (1.1.2) gives

$$\begin{aligned}\frac{d}{dt}\int_{k_i} u(x,t)\,dx &= -\int_{k_i}\frac{d}{dx}g(u(x,t))\,dx \\ &= g(u(x_{i-1/2},t)) - g(u(x_{i+1/2},t)).\end{aligned} \tag{1.1.3}$$

Integrating (1.1.3) in time from t_n to t_{n+1} yields

$$\int_{k_i} u(x,t_{n+1})\,dx = \int_{k_i} u(x,t_n)\,dx + \int_{t_n}^{t_{n+1}} g(u(x_{i-1/2},t))\,dt - \int_{t_n}^{t_{n+1}} g(u(x_{i+1/2},t))\,dt. \tag{1.1.4}$$

Using the numerical fluxes, we have

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left[F_{i+1/2}^n - F_{i-1/2}^n\right], \tag{1.1.5}$$

or

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{F_{i+1/2}^n - F_{i-1/2}^n}{\Delta x} = 0, \tag{1.1.6}$$

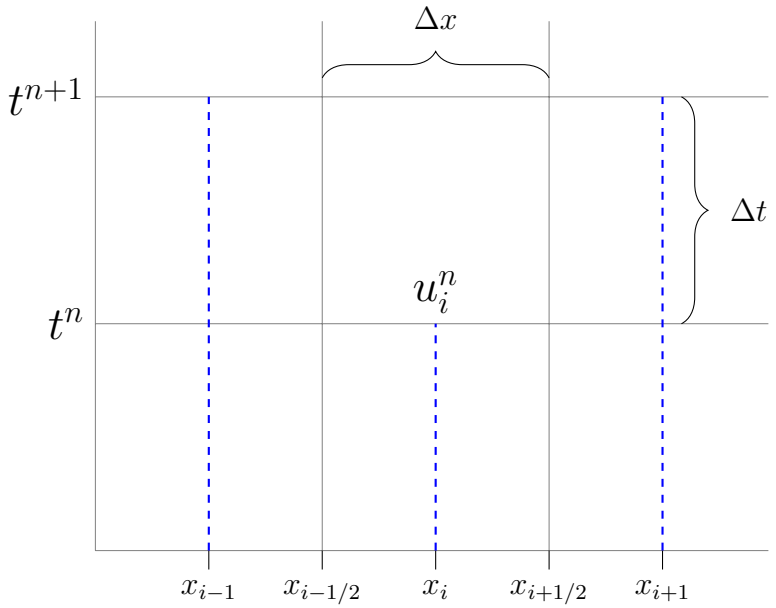


Figure 1.2: One dimensional finite volume discretization

where

$$F^n_{i-1/2} = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} g(u(x_{i-1/2}, t))\, dt.$$

Eq (1.1.6) is the discrete analogue of the conservation law (1.1.1), called a finite volume discretization of (1.1.1).

1.1.3 Some properties and definition of terms

Definition 1.1.1. *Let V is a vector space and F is a field. A function ψ,*

$$V \times V \to F,$$

is a bilinear map if it satisfies the following properties: for any $u, v, w \in V$

$$\psi(av + bv, w) = a\,\psi(v, w) + b\,\psi(u, w) \text{ and}$$
$$\psi(v, au + bw) = a\,\psi(v, w) + b\,\psi(v, w).$$

Definition 1.1.2. *A finite difference equation (FDE) is consistent with a partial differential equation (PDE) if the difference between the FDE and the PDE (i.e., the truncation error)*

vanishes as the size of the grid spacing goes to zero independently. The order of a FDE is the rate at which the global error decreases as the grid size approaches zero (Hoffman, 2001).

To obtain truncation error, we have to use Taylor's expansion of discrete terms.

Definition 1.1.3. *A finite difference/volume method is called stable in the norm $\|\cdot\|$ if there exist constants c and c_0 such that*

$$\|u^n\| \leq c_0 \exp(cn\Delta t)\|u^0\|, \tag{1.1.7}$$

where $c_0 > 0$ and c is independent of the spacial and temporal step sizes.

Definition 1.1.4. *A finite difference/volume method is convergent in the norm $\|\cdot\|$ if the discrete solution (i.e., the numerical values) approaches the exact solution of the partial differential equation in the norm $\|\cdot\|$ as the size of the grid spacing go to zero (Hoffman, 2001).*

Theorem 1.1.1 (Lax Equivalence Theorem (Morton and Mayers, 2005))**.** *For a consistent linear finite difference approximation to a well-posed linear partial differential equation, the stability of the scheme is necessary and sufficient for convergence.*

For the linear problems, i.e., advection-diffusion equations, it is required to show the finite difference methods are consistent and stable, and hence by Theorem 1.1.1 are convergent.

Poincaré inequality (Canuto et al., 2007): For $u \in H^1(\Omega)$, the Poincaré inequality states that there exists a constant C (depending upon Ω) such that

$$\|u\|^2 \leq C(\Omega)\left[|u|_1^2 + \int_\Omega u\,dx\right], \tag{1.1.8}$$

where $\|\cdot\|$ and $|\cdot|_1$ are the L^2 and H^1 norms, respectively.

Restricting ourselves to the case of functions vanishing on the boundary $\partial\Omega$ of the domain of definition Ω and zero mean in case of periodic boundaries, i.e., $\int_\Omega u\,dx = 0$, one has

$$\|u\|^2 \leq C(\Omega)|u|_1^2. \tag{1.1.9}$$

1.1 Preliminaries and some important concepts

The smallest constant C for which (1.1.9) and (1.1.8) hold is termed as the Poincaré constant of the domain Ω.

The same result holds if the domain Ω is simply connected and u only vanishes on a portion of $\partial\Omega$ of positive measure.

Given a linear operator $T : X \to U$, we say that T satisfies the existence property if there exists at least one solution $x \in X$ to $T(x) = u$ for every $u \in U$. The existence property is equivalent to the surjectivity of T and therefore, to the condition $rank(T) = dim(U)$.

We say that T satisfies the uniqueness property if there exists at most one solution $x \in X$ for each $u \in U$. The uniqueness property is equivalent to injectivity and therefore, to the condition $ker(T) = \{0\}$.

Lemma 1.1.1. *(Gockenbach, 2010) Let X and U be $d-$dimensional vector spaces over a field F, and let $T : X \to U$ be linear. Then T is surjective if and only if it is injective.*

The following inequalities are also important for the later discussion.

- For $x \in [0, \frac{1}{2}]$,

$$\left(\frac{1}{2}\right)^{2x} \leq 1 - x. \tag{1.1.10}$$

- For $x \in \mathbb{R}$,

$$1 + x \leq \exp(x). \tag{1.1.11}$$

- Young's inequality: For any $a, b \in \mathbb{R}$ and any $\delta > 0$, we have

$$ab \leq \frac{\delta}{2}a^2 + \frac{1}{2\delta}b^2. \tag{1.1.12}$$

- Cauchy-Schwarz's inequality: For $N \in \mathbb{N}$

$$\sum_{i=1}^{N} a_i b_i \leq \left(\sum_{i=1}^{N} a_i^2\right)^{1/2} \left(\sum_{i=1}^{N} b_i^2\right)^{1/2}. \tag{1.1.13}$$

1.2 Advection-diffusion equation

In nature, transport occurs in fluids through the combination of advection and diffusion. The advection-diffusion equation is one of the most challenging equations in science as it represents a superposition of two different transport processes: advection and diffusion (Szymkiewicz, 2010).

In practical applications, the advection-diffusion equation has been used to predict the movement of a pollutant in a body water (Szymkiewicz, 2010), describe heat transfer in a draining film (Isenberg and Gutfinger, 1972), water transport in soils (Parlarge, 1980), mass transfer (Guvanasen and Volker, 1983), flow in porous media (Kumar, 1983), charge transport in semi-conductor devices (Ehrhardt and Mickens, 2013), the spread of pollutants in rivers and streams (Chatwin and Allen, 1985), the transport of pollutants in the atmosphere (Zlatev et al., 1984), contaminant dispersion in shallow lakes (Salmon et al., 1980), oil reservoir flow (Dehghan, 2005) and thermal pollution in river systems (Chaudhry et al., 1983).

The advection-diffusion equation is given by

$$\frac{\partial u}{\partial t} + \vec{v} \cdot \nabla u = \nabla \cdot (\boldsymbol{D} \cdot \nabla u),\ t > 0,$$

where $\vec{v}$ is advecting velocity vector and D is the (constant) diffusivity tensor and u is the transported dependent variable.

Various numerical methods have been introduced to model accurately the interaction between advection and diffusion processes. The available numerical methods are sophisticated in order to avoid two undesirable features: oscillatory behaviour (dispersion) and numerical dissipation (damping). Modelling accurately the interaction between advection and diffusion processes in the numerical approximation of the partial differential equations is a challenging problem (Morton, 1996).

1.3 Von Neumann stability analysis

There are several techniques that can be used to obtain the stability of numerical methods. CFL condition (R. Courant and Lewy, 1967), matrix method (Siemieniuch and Gladwell, 1978) and von Neumann stability analysis are some of these methods (Richtmer and Morton, 1967). CFL conditions are only neccessary conditions for stability and matrix method is necessary but not always sufficient (Sousa, 2003). The von Neumann method is the most well-known method to determine the necessary and sufficient conditions for the stability of numerical methods (Sousa, 2003). Von Neumann's method is applicable only for linear finite difference equations with constant coefficients. To apply for more general problems, the governing finite difference equation must be linearised locally and any variable coefficients must be frozen at some constant value (Durran, 2010).

The von Neumann stability analysis is based on the decomposition of the discretized solution at some particular time step into a finite Fourier series as

$$u_i^n = \sum_{j=-N}^{N} \xi_j^n \exp(I\omega_j\, i),$$

where $I = \sqrt{-1}$; ξ_j is an amplification factor and θ_j is a wave number of the j-th Fourier component. The variable $\omega_j = \theta_j\, \Delta x$ is called the phase angle and covers the domain $(-\pi, \pi)$. The region around $\omega_j = 0$ correspond to the low frequencies and while around the region $\omega_j = \pi$ corresponds to the highest frequencies (Sousa, 2003). Since time evolution of a single Fourier mode, $\xi^n \exp(I\omega\, i)$, is determined by the same numerical scheme as the complete solution u_i^n, the stability of the numerical method is obtained by inserting $\xi_j^n \exp(I\omega_j\, i)$ into a numerical scheme (Durran, 2010). For simplicity, hereafter, the variables, ω, θ and ξ, are written without the subscript j representing the j-th Fourier component.

Definition 1.3.1. *An amplification factor, ξ, is said to satisfy the von Neumann condition if there is a constant K such that (Richtmer and Morton, 1967; Sousa, 2003; Durran, 2010)*

$$|\xi(\omega)| \leq 1 + K\Delta t, \forall\, \omega \in [-\pi, \pi]. \tag{1.3.1}$$

where K is independent of Δt and Δx.

In practice, the inequality (1.3.1) is replaced by the following strong condition

$$|\xi(\omega)| \leq 1, \forall\, \omega \in [-\pi, \pi], \tag{1.3.2}$$

which will ensure that (1.1.7) is satisfied.

Lemma 1.3.1. *A linear finite difference method is stable in the discrete L^2-norm if and only if the von Neumann condition is satisfied (Sod, 1988).*

In this thesis, we use this von Neumann stability analysis to find the stability conditions for the numerical methods used to solve the advection-diffusion equations.

1.4 Convective Cahn-Hilliard equation

Let Ω denote an open bounded set of $\mathbb{R}^d$, $d = 1, 2$ or 3. The convective Cahn-Hilliard equation is a fourth order nonlinear partial differential equation given by

$$u_t - \gamma u(\boldsymbol{\beta} \cdot \nabla u) = \Delta(f(u) - \varepsilon^2 \Delta u),\ t > 0, \tag{1.4.1}$$

where $u = u(x, t), x \in \Omega$, is a scalar unknown function, γ is the driving force, $\boldsymbol{\beta}$ is $d-$dimensional vector, ε is a dimensionless interfacial width and

$$f(u) = u^3 - u.$$

This equation is a successful model for the description of several physical phenomena: spinodal decomposition of phase separating systems in the presence of an external field (e.g. gravitational, magnetic and electronic) (Leung, 1990; Emmott and Bray, 1996; Yeung et al., 1992), formation of facets and corners in crystal growth (Golovin et al., 1998, 1999). Phase transition observed in alloys, glasses, polymer solution and binary liquid mixtures, faceting of growing thermodynamically unstable surfaces (see (Golovin et al., 2001; Podolny et al., 2005), and the references cited therein).

The convective Cahn-Hilliard equation is on one side related to the Cahn-Hilliard equation and on the other side it is related to the Kuramato-Sivashinsky equation (Eden and Kalantarov, 2007a; Golovin et al., 2001; Watson et al., 2003; Watson, 2003). When $\gamma \simeq 1$, the

morphologies of the solutions do not coarsen as time progresses but rather display periodic patterns (Watson, 2003).

In the absence of the driving field, i. e. $\gamma = 0$, the Eq. (1.4.1) reduces to the well known Cahn-Hilliard equation

$$u_t + \varepsilon^2 \Delta^2 u = \Delta f(u), \tag{1.4.2}$$

which is a model to describe the evolution of a concentration field for a binary mixture (Cahn, 1965) and phase separation of binary liquids or binary alloys (Cahn and Hilliard, 1958; Bray, 1994). This reduced model has been studied by several authors (Elliott and Songmu, 1986; Khiari et al., 2007; Song, 2015). In Song (2015), higher order schemes preserving the properties such as energy and large time behavior are constructed. The Cahn-Hilliard equation, (1.4.2), admits a Lyapunov (free energy) functional which guarantees that generically all solutions converge to an equilibrium.

For large γ, that is $\gamma \to \infty$, the transformation $u \to \dfrac{\bar{u}}{\gamma}$ reduces Eq. (1.4.1) to the well known Kuramato-Sivashinky equation, which describes the chaotic structure of dynamical systems. By ignoring the bar on u the Kuramato-Sivashinky is given by

$$u_t - u(\boldsymbol{\beta} \cdot \nabla u) = -\Delta(u + \varepsilon^2 \Delta u),\ x \in \Omega, t > 0.$$

The one-dimensional case of Eq. (1.4.1) has been studied by several researchers, theoretically and numerically. Analytical solutions have been obtained for a single interface in the presence of the driving force, i.e $\gamma \neq 0$, in an infinite system (Leung, 1990). The effect of this driving force on the coarsening dynamics of the one-dimensional Cahn-Hilliard equation at $T = 0$ has been studied by Emmott and Bray (1996) when $\varepsilon = 1$. They observed that the driving force γ has an asymmetric effect on the solution of a single stationary domain wall. They also noted that the behavior of the kink-anti kink pair (bubble) depends on γ^{-1} and the separation of the interfaces. Later, Golovin et al. (2001) demonstrated numerically that the one-dimensional convective Cahn-Hilliard equation exhibit a transition from coarsening to chaotic behaviour as γ increases. The presence of the driving force elucidates a fundamental asymmetry between kinks and anti-kinks which is not present in the Cahn-Hilliard theory (Watson et al., 2003). In Podolny et al. (2005), the dynamics of

domain walls (kinks) governed by the convective Cahn-Hilliard equation is studied by means of asymptotic and numerical methods. The bifurcations of stationary solutions for different values of γ with $\varepsilon = 1$ has been studied by Zaks et al. (2006). Eden and Kalantarov (2007a) proved the existence of compact attractor for the one dimensional convective Cahn-Hilliard equation s with periodic boundary condition. Zhao and Liu (2013) proved the existence of optimal solutions for the one dimensional convective Cahn-Hilliard Equation. Aderogba et al. (2014) solved the one dimensional convective Cahn-Hilliard equation numerically using fractional step-splitting methods for $\gamma = 0.1$ and $\varepsilon = 1$. The authors observe that the solution coarsens as t progresses and they tested numerically the transition of convective Cahn-Hilliard equation from coarsening to an order less pattern as γ increases, which is the behavior of Kuramoto-Sivashinsky equation.

For the convective Cahn-Hilliard equation in higher space dimensions, the transition to roughening and the structure of the steady states are not well understood (Emmott and Bray, 1996; Golovin et al., 2001; Eden and Kalantarov, 2007b). The existence of optimal solutions for the convective Cahn-Hilliard Equations in two dimensions has been proved by Zhao and Liu (2014). Eden and Kalantarov (2007b) considered a 3D convective Cahn-Hilliard equation with periodic boundary conditions and proved the existence of absorbing balls. In Zhao (2015), semidiscrete and completely discrete spectral approximations are constructed to study the long time behaviour of the 2D convective Cahn- Hilliard equation. Polata et al. (2009) show that a semi-group generated by 3D convective Cahn-Hilliard equation has a global attractor.

1.5 Thesis overview

The thesis is divided into three main chapters. Each chapter is self-contained and represents a scientific contribution (published, accepted or submitted). Chapters 2 and 3 are devoted to a study of some the finite difference methods for the discretization of the linear advection-diffusion equations. In chapter two, three numerical methods have been used to solve two 1D advection-diffusion problems. We compute some numerical errors and then implement

an optimization technique to find the optimal temporal step size when the spatial step size is chosen as $\Delta x = 0.1$. **The results in this chapter are the subject to the papers (Appadu et al., 2015, 2016).**

In chapter 3, three finite difference methods have been used to solve a 3D advection-diffusion equation with given initial and boundary conditions. A new finite difference scheme is presented to solve 3D advection-diffusion equation. We compare the performance of the methods by computing L_2-error, L_∞-error, dispersion error, dissipation error and some performance indices such as mass distribution ratio, mass conservation ratio, total mass and R^2. We then use an optimization technique to find an optimal time step size that minimizes dispersion errors. **This is the first time the fourth order finite difference method and an optimization technique applied to a 3D advection-diffusion equation. A part of this result is accepted for publication Appadu et al. (b) and a full version of this chapter is submitted for publication (Appadu et al., c).**

Chapter 4 deals with one-level and multilevel finite volume approximations of 1D and 2D convective Cahn-Hilliard equations. We construct four finite volume methods for solving 1D and 2D convective Cahn-Hilliard equations. We study the existence/uniqueness of solutions and stability and convergence of the finite volume methods for the 2D convective Cahn-Hilliard equation. We compute the computational time for some chosen temporal and spatial step sizes at some values of T. Numerical experiments are carried out to solve the 1D problem. We compute convergence rates for the finite volume methods constructed. **This is the first time the nonlocal approximation of the convective term preserving the properties of the continuous model applied to the convective Cahn-Hilliard equation and the multilevel approximation implemented on the convective Cahn-Hilliard equation. The results in this chapter are submitted for publication (Appadu et al., a,d).**

The last chapter summarizes the dissertation and some possible avenues for future research.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Chapter 2

A computational study of three numerical methods for some advection-diffusion problems

A version of this work has been published in (Appadu et al., 2016) *and a shorter version of this work published in* (Appadu et al., 2015).

2.1 Introduction

We consider the 1D advection-diffusion equation

$$u_t + u_x = \alpha\, u_{xx}, \quad 0 < x < 1, 0 < t \leq T, \tag{2.1.1}$$

with initial condition

$$u(x,0) = f(x),\ 0 \leq x \leq 1,$$

and boundary conditions

$$u(0,t) = g_0(t),\ 0 < t \leq T$$

$$u(1,t) = g_1(t),\ 0 < t \leq T,$$

where α is diffusivity constant.

In this chapter, three numerical methods, namely: third order upwind scheme, fourth order scheme and non-standard finite difference scheme (NSFD), have been used to solve (2.1.1) with given initial and boundary conditions. Also, an optimization technique has been implemented to find an optimal temporal step size that minimizes the dispersion error for the third order finite difference method for a fixed spatial step size.

This chapter is outlined as follows: In section 2.2, we study the dissipative and dispersive characteristics of some numerical methods for the 1D advection-diffusion equation. Section 2.3 describes the two test cases considered in this chapter. In section 2.4, we show how to quantify errors from numerical results into dispersion and dissipation errors using the technique by Takacs (1985). In sections 2.5-2.7, we describe how the three finite difference methods are constructed and obtain their stability conditions using the approach of Hindmarsh et al. (1984). Numerical results are presented in section 2.8. In section 2.9, we obtain the optimal temporal step size using a technique devised by Appadu (2013) to minimize the dispersion error at a given Reynolds number and spatial step size and this is validated by some numerical experiments for test case 1. In section 2.10, we highlight the salient features of the chapter.

2.2 Numerical dissipation and dispersion

Finite difference schemes used to solve partial differential equations will often lose energy as time t progresses, this property is called numerical dissipation (Trefethen, 1996). In the case of dispersive schemes, oscillations are generated in regions of discontinuity. We let the elementary solution of Eq. (2.1.1) be (Anderson et al., 1984)

$$u(x,t) = \exp(\gamma t)\exp(I\theta\, x), \tag{2.2.1}$$

where θ is the wave number and γ is dispersion relation. On plugging Eq. (2.2.1) into Eq. (2.1.1), we get

$$\gamma = -\theta I - \alpha\theta^2 \tag{2.2.2}$$

Hence

$$u(x,t) = \exp\left[(-\theta I - \alpha\theta^2)t\right] \exp(I\theta x). \tag{2.2.3}$$

From Eq (2.2.3), we deduce that the partial differential equation given by (2.1.1) represents a wave with exponentially decaying amplitude travelling at a constant speed. The exact amplification factor is obtained as (Anderson et al., 1984)

$$\xi_{exact} = \frac{u(x,t+\Delta t)}{u(x,t)} = \exp((-I\theta - \alpha\theta^2)\Delta t). \tag{2.2.4}$$

The modulus of the exact amplification factor is then given by

$$|\xi_{exact}| = \exp(-\alpha\theta^2\Delta t). \tag{2.2.5}$$

The numerical amplification factor, ξ_{num} is obtained using von Neumann stability analysis. The relative phase error (RPE) is obtained as (Anderson et al., 1984)

$$RPE = \frac{\arg(\xi_{num})}{\arg(\xi_{exact})} = -\frac{arg(\xi_{num})}{\theta\,\Delta t} = -\frac{arg(\xi_{num})}{c\,\omega}, \tag{2.2.6}$$

where c is courant number and $\omega = \theta\Delta x$ is phase angle, with Δx being the spatial step size. Eq. (2.2.6) can also be written as

$$RPE = -\frac{1}{c\,\omega}\arctan\left(\frac{\Im(\xi_{num})}{\Re(\xi_{num})}\right),$$

2.3 Test cases

In this section, we consider two test problems to compare the performance of the three numerical methods, namely third order upwind, fourth order and NSFD.

2.3.1 Test case 1

We solve a problem from Feng and Tian (2006) which is described by the partial differential equation

$$u_t + u_x = \frac{1}{Re} u_{xx}, \quad 0 < x < 1, \ t > 0, \tag{2.3.1}$$

with initial condition

$$u(x, 0) = 0, \ 0 < x < 1 \tag{2.3.2}$$

and boundary conditions

$$u(0, t) = 0, \ u(1, t) = 1, \ t > 0. \tag{2.3.3}$$

where Re is Reynolds number.

The analytical solution to this equation is obtained using the method of separation of variables. As $t \to \infty$, there exists a solution v of (2.3.1) independent of t and $v(0) = 0$ and $v(1) = 1$, then u can be written as

$$u(x, t) = v(x) + w(x, t),$$

where w is a solution of Eq. (2.3.1) with

$$w(0, t) = w(1, t) = 0, w(x, 0) = -v(x).$$

Substituting v into Eq. (2.3.1) and solving for v, one obtains

$$v(x) = \frac{\exp(\mathrm{Re}\, x) - 1}{\exp(\mathrm{Re}) - 1}.$$

Now we solve the following problem

$$\begin{cases} w_t + w_x = \dfrac{1}{\mathrm{Re}} w_{xx}, \\ w(0, t) = w(1, t) = 0, \\ w(x, 0) = \dfrac{\exp(\mathrm{Re}\, x) - 1}{1 - \exp(\mathrm{Re})}. \end{cases} \tag{2.3.4}$$

We first find A with

$$w(x, t) = A(x, t) p(x, t), \tag{2.3.5}$$

such that p satisfies the following partial differential equation (heat equation)

$$p_t = \frac{1}{\mathrm{Re}}\, p_{xx}. \tag{2.3.6}$$

Substituting (2.3.5) into Eq. (2.3.4) and solving for p_t, we get

$$p_t = \frac{1}{\mathrm{Re}} p_{xx} + \frac{-A + \frac{2}{\mathrm{Re}} A_x}{A} p_x + \frac{-A_t - A_x + \frac{1}{\mathrm{Re}} A_{xx}}{A} p. \tag{2.3.7}$$

To obtain Eq. (2.3.6), the coefficients of p and p_x in Eq. (2.3.7) must be zero. We now solve the following partial differential equations

$$-A + \frac{2}{\mathrm{Re}} A_x = 0 \tag{2.3.8}$$

$$-A_t - A_x + \frac{1}{\mathrm{Re}} A_{xx} = 0. \tag{2.3.9}$$

Solving Eq. (2.3.8), we obtain

$$A(x,t) = c(t) \exp\left(\frac{\mathrm{Re}}{2}\, x\right). \tag{2.3.10}$$

Substituting (2.3.10) into Eq. (2.3.9) and solving for $c(t)$ yields

$$c(t) = A_0 \exp\left(-\frac{\mathrm{Re}}{4}\, t\right),$$

where A_0 is constant. Hence we have

$$w(x,t) = A_0 \exp\left(-\frac{\mathrm{Re}}{4}\, t + \frac{\mathrm{Re}}{2}\, x\right) p(x,t). \tag{2.3.11}$$

By taking $A_0 = 1$, we solve Eq. (2.3.6) with initial condition

$$p(x,0) = \frac{\exp(\frac{\mathrm{Re}\, x}{2}) - \exp(\frac{-\mathrm{Re}\, x}{2})}{1 - \exp(\mathrm{Re})},$$

and boundary conditions

$$p(0,t) = p(1,t) = 0,$$

using the method of separation of variables. We write p as a product of two functions X and T such that

$$p(x,t) = X(x)\, T(t), \tag{2.3.12}$$

2.3 Test cases

with $X(0) = X(1) = 0$. Substituting 2.3.12 into Eq. (2.3.6), we get

$$\frac{\mathrm{Re}T'}{T} = \frac{X''}{X} = \lambda, \tag{2.3.13}$$

where λ is the coefficient of separation and it is given by

$$\lambda = -\omega^2, \omega \in \mathbb{R}^+.$$

We now solve the following two ordinary differential equations

$$\mathrm{Re}T' = -\omega^2\, T, \tag{2.3.14}$$

$$X'' = -\omega^2 X. \tag{2.3.15}$$

The general solutions for Eqs. (2.3.14) and (2.3.15) are respectively

$$T(t) = T_0\, \exp\left(\frac{-\omega^2\, t}{\mathrm{Re}}\right) \tag{2.3.16}$$

and

$$X(x) = a\, \cos(\omega\, x) + b\, \sin(\omega\, x),$$

for constants T_0, a and b. Using the boundary conditions we obtain $a = 0$ and $\omega_n = n\pi, n = 1, 2, \ldots$. Thus we have

$$p(x,t) = \sum_{n=1}^{\infty} B_n\, \exp\left(\frac{-n^2\, \pi^2\, t}{\mathrm{Re}}\right) \sin(n\, \pi\, x), \tag{2.3.17}$$

where

$$B_n = 2 \int_0^1 p(x,0)\, \sin(n\, \pi\, x) dx.$$

Using integration by parts, we obtain

$$B_n = \frac{n\, \pi (-1)^n}{n^2\, \pi^2 + \frac{\mathrm{Re}^2}{4}} \exp(-\frac{\mathrm{Re}}{2}).$$

Hence, we have

$$w(x,t) = \exp\left(-\frac{\mathrm{Re}}{4}\, t + \frac{\mathrm{Re}}{2}\, x\right) \sum_{n=1}^{\infty} \left[\frac{n\, \pi (-1)^n}{n^2\, \pi^2 + \frac{\mathrm{Re}^2}{4}} \exp\left(-\frac{\mathrm{Re}}{2}\right)\right] \exp\left(\frac{-n^2\, \pi^2\, t}{\mathrm{Re}}\right) \sin(n\, \pi\, x), \tag{2.3.18}$$

Therefore, the solution u of Eqs. (2.3.1)-(2.3.3) is given by

$$u(x,t) = \left[\frac{\exp(\mathrm{Re}\ \times x) - 1}{\exp(\mathrm{Re}) - 1}\right] + \sum_{m=1}^{\infty}\left\{\frac{(-1)^m m\pi}{(m\pi)^2 + \frac{\mathrm{Re}^2}{4}}\ \exp\left(\frac{\mathrm{Re} \times (x-1)}{2}\right)\right.$$
$$\left.\times \sin(m\pi x)\ \exp\left[-t\left(\frac{(m\pi)^2}{\mathrm{Re}} + \frac{\mathrm{Re}}{4}\right)\right]\right\}.$$

In Tian and Yu (2011), they have used two values of Re, namely 100 and 10,000 and two values of step-size (for coarse and fine grids), namely 0.1 and 0.025. The temporal step size was chosen as 0.01. This test case is quite challenging as for instance at Re = 100, numerical solutions from Crank-Nicolson method show non-physical oscillations while the scheme constructed by Ding and Zhang (2009) is not accurate on coarse mesh. Moreover, for Re = 10000, the scheme by Ding and Zhang gives inaccurate solutions due to dispersion.

In this work, we consider three values of Re, say 10, 100, and 10000, and two values of Δx, say 0.1 and 0.025. The temporal step size is chosen as 0.01. We compute the L_2, L_∞, total variation, dissipation and dispersion errors when the three schemes are used to solve problem 1 at time, $T = 1$. Table 2.1 gives the regions of stability of the three schemes at some values of Re and Δx. The errors are shown in Tables 2.2 to 2.4 and the numerical and exact plots are shown in Figs. 2.1a to 2.1e.

2.3.2 Test case 2

We consider the advection-diffusion equation

$$u_t + u_x = \alpha\, u_{xx}, \quad 0 < x < 1,\ t > 0 \tag{2.3.19}$$

with the boundary conditions

$$u(0,t) = u(1,t) = 0,$$

and the initial condition

$$u(x,0) = 3\sin(4\,\pi\,x).$$

The exact solution of the problem is given by (Chawla et al., 2000)

$$u(x,t) = \exp\left[\frac{1}{2\alpha}\left(x - \frac{t}{2}\right)\right] \sum_{j=1}^{\infty} \zeta_j \exp\left(-\alpha\, j^2\, \pi^2\, t\right)\, \sin(j\pi x),$$

where

$$\zeta_j = \frac{3}{2\alpha}\left[1 + (-1)^{j+1} \exp\left(-\frac{1}{2\alpha}\right)\right]\left[\frac{1}{(\frac{1}{2\alpha})^2 + (j-4)^2\,\pi^2} - \frac{1}{(\frac{1}{2\alpha})^2 + (j+4)^2\,\pi^2}\right].$$

The exact solution is obtained using the method of separation of variables as discussed in subsection 2.3.1.

In Chawla et al. (2000), they have used a one-parameter family of unconditionally stable third order time-integration scheme with temporal and spatial step sizes being $\Delta t = 0.25$ and $\Delta x = 0.05$, respectively and compared their results with Crank-Nicolson which is highly oscillatory. In this work, we consider four combinations of values of α and Δx; namely $\alpha = 0.01, \Delta x = 0.1$; $\alpha = 0.1, \Delta x = 0.05$; $\alpha = 1, \Delta x = 0.05$ and $\alpha = 1, \Delta x = 0.1$ and display the numerical results at time, $T = 1$. For each choice α and Δx, we consider four different values of the temporal step size for which the methods are stable and the regions of stability are depicted in Table 2.6.

2.4 Quantification of errors from numerical results

In this section, we describe how Takacs (1985) quantifies errors from numerical results into dispersion and dissipation errors.

The Total Mean Square Error (TMSE) is calculated as

$$TMSE = \frac{1}{N}\sum_{i=1}^{N}(u_i - v_i)^2,$$

where u_i represents the analytical solution and v_i, the numerical (discrete) solution at a given grid point i and N is the number of discrete points.

The Total Mean Square Error can be expressed as

$$\frac{1}{N}\sum_{i=1}^{N}(u_i - v_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(u_i - \bar{u})^2 + \frac{1}{N}\sum_{i=1}^{N}(v_i - \bar{v})^2 + \frac{2}{N}\sum_{i=1}^{N} u_i\bar{u} + \frac{2}{N}\sum_{i=1}^{N} v_i\bar{v}$$

$$-\frac{1}{N}\sum_{i=1}^{N}(\bar{u})^2-\frac{1}{N}\sum_{i=1}^{N}(\bar{v})^2-\frac{2}{N}\sum_{i=1}^{N}u_iv_i. \tag{2.4.1}$$

The right hand side of Eq. (2.4.1) can be rewritten as

$$\sigma^2(u)+\sigma^2(v)+2(\bar{u})^2+2(\bar{v})^2-(\bar{u})^2-(\bar{v})^2-\frac{2}{N}\sum_{i=1}^{N}u_iv_i,$$

where $\sigma^2(u)$ and $\sigma^2(v)$ denote the variance of u and v, respectively, $\bar{u}$ and $\bar{v}$ denote the mean values of u and v, respectively. Then we have

$$TMSE=\sigma^2(u)+\sigma^2(v)+(\bar{u}-\bar{v})^2-2\,Cov(u,v),$$

where $Cov(u,v)=\frac{1}{N}\sum_{i=1}^{N}u_iv_i-\bar{u}\bar{v}$. The Total Mean Square Error can also expressed as

$$(\sigma(u)-\sigma(v))^2+(\bar{u}-\bar{v})^2+2(1-\rho)\sigma(u)\sigma(v), \tag{2.4.2}$$

where $\rho=Cov(u,v)/(\sigma(u)\sigma(v))$ is the coefficient of correlation. The expression $2(1-\rho)\sigma(u)\sigma(v)$ measures the dispersion error and $(\sigma(u)-\sigma(v))^2+(\bar{u}-\bar{v})^2$ measures the dissipation error.

We also obtain values of the L_2 and L_∞ errors which are obtained by the following formulae:

$$L_2\text{-error}=\sqrt{\Delta x\sum_{i=1}^{N}(u_i-v_i)^2}, \tag{2.4.3}$$

$$L_\infty\text{-error}=\max|u_i-v_i|. \tag{2.4.4}$$

Numerical methods must have monotone and Total Variation Diminishing properties. A numerical scheme is said to be monotone if it produces a monotonic distribution after advection, given a distribution that is monotonic before advection. Monotonic methods neither create new extrema in the solution nor amplify existing extrema. Monotonic schemes are classified broadly into two classes: Flux-corrected transport (FCT) and Flux Limiter Method (FLM) (Boris and Book, 1973; Zalesak, 1979; Durran, 2010).

In FCT, the advective fluxes are essentially a weighted average of a lower-order monotonic scheme and a higher-order non-monotonic scheme. In FLM, the advective fluxes of a high-order scheme are modified so that the total variation of the solution does not increase with

time. This property is called Total Variation Diminishing (TVD). The Total Variation (TV) of a function u is defined as

$$TV = \sum_{i=1}^{N-1} |u_{i+1} - u_i|.$$

A TVD scheme ensures that $TV(u^{n+1}) \leq TV(u^n)$ which signifies that the overall amount of oscillations is bounded (Trac and Pen, 2003). All monotone schemes are TVD. All TVD schemes are monotonically preserving methods (for proof see (Ganzha and Vorozhtsov, 1996)).

2.5 Third order upwind explicit scheme

Dehghan (2005) developed high order efficient finite difference schemes (Eleiwi and Laleg-Kirati, 2014) based on the modified equivalent partial differential equations (MEPDE) as a means of estimating the order of accuracy of the methods. Third order upwind and fourth order schemes are some of these methods and have been used to solve a 1D problem described by a constant coefficient advection-diffusion equation with smooth initial conditions and quite good results with high accuracy have been obtained. These methods are explicit and can therefore be used to maximum advantage on a parallel computer. In this section and sections 2.6 and 2.7, we describe the three methods and obtain the order of accuracy and the region of stability. We now describe how the third order upwind scheme is constructed.

$$\left.\frac{\partial u}{\partial t}\right|_i^n \simeq \frac{u_i^{n+1} - u_i^n}{\Delta t}, \tag{2.5.1}$$

$$\left.\frac{\partial u}{\partial x}\right|_i^n \simeq \left(\frac{2c^2 + 3c + 12s - 2}{12}\right)\left(\frac{u_i^n - u_{i-2}^n}{2\Delta x}\right) + \left(\frac{2c^2 - 3c + 12s - 2}{12}\right)\left(\frac{u_{i+2}^n - u_i^n}{2\Delta x}\right) + \left(\frac{4 - c^2 - 6s}{3}\right)\left(\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}\right), \tag{2.5.2}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_i^n \simeq \left(\frac{6s - 12sc + 2c - 2c^3 + 3c^2}{6s}\right)\left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}\right) + \left(\frac{12sc - 2c + 2c^3 - 3c^2}{6s}\right)\left(\frac{u_{i+2}^n - 2u_i^n + u_{i-2}^n}{4(\Delta x)^2}\right). \tag{2.5.3}$$

2.5 Third order upwind explicit scheme

where $c = \dfrac{\Delta t}{\Delta x}$ and $s = \dfrac{\alpha\,\Delta t}{\Delta x^2}$. On substitution of Eqs. (2.5.1)-(2.5.3) into Eq. (2.1.1), we have the following:

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+\left(\frac{2c^2+3c+12s-2}{12}\right)\left(\frac{u_i^n-u_{i-2}^n}{2\Delta x}\right)+\left(\frac{2c^2-3c+12s-2}{12}\right)\left(\frac{u_{i+2}^n-u_i^n}{2\Delta x}\right)$$
$$+\left(\frac{4-c^2-6s}{3}\right)\left(\frac{u_{i+1}^n-u_{i-1}^n}{2\Delta x}\right)=\alpha\left[\left(\frac{6s-12sc+2c-2c^3+3c^2}{6s}\right)\left(\frac{u_{i+1}^n-2u_i^n+u_{i-1}^n}{(\Delta x)^2}\right)\right.$$
$$\left.+\left(\frac{12sc-2c+2c^3-3c^2}{6s}\right)\left(\frac{u_{i+2}^n-2u_i^n+u_{i-2}^n}{4(\Delta x)^2}\right)\right].$$

On simplification we get

$$u_i^{n+1}=\frac{1}{24}[4c^3+24sc-4c]u_{i-2}^n+\frac{1}{6}(6c-3c^3-18sc+6s+3c^2)u_{i-1}^n+$$
$$\left[1-\frac{c^2}{4}-\frac{1}{12}(24s-36sc+6c-6c^3+9c^2)\right]u_i^n+\frac{1}{6}(-2c-c^3-6sc+6s+3c^2)u_{i+1}^n.$$

Therefore, the third order upwind scheme is given by

$$u_i^{n+1}=A_1\,u_{i-2}^n+A_2\,u_{i-1}^n+A_3\,u_i^n+A_4\,u_{i+1}^n, \tag{2.5.4}$$

where

$$A_1=\frac{1}{6}c(c^2+6s-1), \qquad A_2=\frac{1}{2}(2c-c^3-6sc+2s+c^2),$$
$$A_3=\frac{1}{2}(2-2c^2-4s+6sc-c+c^3), \qquad A_4=\frac{1}{6}(1-c)(6s+c^2-2c).$$

The modified equation of the scheme is given by

$$u_t+u_x-\alpha\,u_{xx}=\frac{\Delta x^3}{24\,c}\left(12sc^2-2s-2c^3+2c+12s^2+6c^4-c^2-2cs\right)u_{xxxx}+\cdots, \tag{2.5.5}$$

and the leading error terms are dissipative in nature. The scheme is consistent and is third order accurate in space. The amplification factor of the scheme is given by

$$\xi=A_1\,\exp(-2I\omega)+A_2\,\exp(-I\omega)+A_3+A_4\,\exp(I\omega). \tag{2.5.6}$$

We use the Fourier analysis and the approach of Hindmarsh et al. (1984) to obtain the stability region. When $\omega=\pi$, on simplification of Eq. (2.5.6), we obtain

$$\xi=1-4s-\frac{4}{3}c+8sc-2c^2+\frac{4}{3}c^3. \tag{2.5.7}$$

Then we have

$$2c + 3c^2 - 2c^3 - 3 \leq s(12c - 6) \leq 2c + 3c^2 - 2c^3. \tag{2.5.8}$$

When $\omega \to 0$, using Taylor's expansion and on neglecting higher order terms, we have

$$|\xi|^2 \simeq 1 - 2s(\omega^2), \tag{2.5.9}$$

and therefore, we must have

$$s \geq 0. \tag{2.5.10}$$

Thus, the scheme is stable when both inequalities (2.5.8) and (2.5.10) are satisfied.

2.6 Fourth order explicit scheme

For this scheme the following approximations are used (Dehghan, 2005);

$$\left.\frac{\partial u}{\partial t}\right|_i^n \simeq \frac{u_i^{n+1} - u_i^n}{\Delta t}, \tag{2.6.1}$$

$$\left.\frac{\partial u}{\partial x}\right|_i^n \simeq \left(\frac{12s + 2c^2 - 3c - 2}{12}\right)\left(\frac{u_{i+2}^n - u_i^n}{2\Delta x}\right) + \left(\frac{12s + 2c^2 + 3c - 2}{12}\right)\left(\frac{u_i^n - u_{i-2}^n}{2\Delta x}\right) - \left(\frac{c^2 + 6s - 4}{3}\right)\left(\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}\right), \tag{2.6.2}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_i^n \simeq \left(\frac{-c^4 + 4c^2 - 12s^2 - 12sc^2 + 8s}{6s}\right)\left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}\right) + \left(\frac{c^4 - 4c^2 + 12s^2 + 12sc^2 - 2s}{6s}\right)\left(\frac{u_{i+2}^n - 2u_i^n + u_{i-2}^n}{4(\Delta x)^2}\right). \tag{2.6.3}$$

On substitution of Eqs. (2.6.1-2.6.3) into Eq. (2.1.1), we have

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \left(\frac{12s + 2c^2 - 3c - 2}{12}\right)\left(\frac{u_{i+2}^n - u_i^n}{2\Delta x}\right) + \left(\frac{12s + 2c^2 + 3c - 2}{12}\right)\left(\frac{u_i^n - u_{i-2}^n}{2\Delta x}\right) - \left(\frac{c^2 + 6s - 4}{3}\right)\left(\frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}\right) = \alpha\left[\left(\frac{-c^4 + 4c^2 - 12s^2 - 12sc^2 + 8s}{6s}\right)\left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}\right) + \left(\frac{c^4 - 4c^2 + 12s^2 + 12sc^2 - 2s}{6s}\right)\left(\frac{u_{i+2}^n - 2u_i^n + u_{i-2}^n}{4(\Delta x)^2}\right)\right],$$

which yields

$$u_i^{n+1} = \left\{1 + c\left[\left(\frac{12s + 2c^2 - 3c - 2}{24}\right) - \left(\frac{12s + 2c^2 + 3c - 2}{24}\right)\right]\right.$$

$$
\begin{aligned}
&- 2s\left[\left(\frac{-c^4+4c^2-12s^2-12sc^2+8s}{6s}\right)+\left(\frac{c^4-4c^2+12s^2+12sc^2-2s}{24s}\right)\right]\Bigg\}u_i^n\\
&+\left\{c\left(\frac{12s+2c^2+3c-2}{24}\right)+s\left(\frac{c^4-4c^2+12s^2+12sc^2-2s}{24s}\right)\right\}u_{i-2}^n\\
&+\left\{-c\left(\frac{c^2+6s-4}{6}\right)+s\left(\frac{-c^4+4c^2-12s^2-12sc^2+8s}{6s}\right)\right\}u_{i-1}^n\\
&+\left\{c\left(\frac{c^2+6s-4}{6}\right)+s\left(\frac{-c^4+4c^2-12s^2-12sc^2+8s}{6s}\right)\right\}u_{i+1}^n\\
&+\left\{-c\left(\frac{12s+2c^2-3c-2}{24}\right)+s\left(\frac{c^4-4c^2+12s^2+12sc^2-2s}{24s}\right)\right\}u_{i+2}^n.
\end{aligned}
\tag{2.6.4}
$$

and after some algebraic manipulation gives

$$
\begin{aligned}
u_i^{n+1} = &\frac{1}{24}[12s(s+c^2)+2s(6c-1)+c(c^3+2c^2-c-2)]u_{i-2}^n\\
&-\frac{1}{6}[12s(s+c^2)+2s(3c-4)+c(c^3+c^2-4c-4)]u_{i-1}^n\\
&+\frac{1}{4}[12s(s+c^2)-10s+((c)^2)^2-5c^2-4]u_i^n\\
&-\frac{1}{6}[12s(s+c^2)-2s(3c+4)+c(c^3-c^2-4c+4)]u_{i+1}^n\\
&+\frac{1}{24}[12s(s+c^2)-2s(6c+1)+c(c^3-2c^2-c+2)]u_{i+2}^n.
\end{aligned}
$$

On rearranging, we get the following finite difference scheme

$$
u_i^{n+1} = A\,u_{i-2}^n + B\,u_{i-1}^n + C\,u_i^n + D\,u_{i+1}^n + E\,u_{i+2}^n, \tag{2.6.5}
$$

where

$$
\begin{aligned}
A &= \frac{1}{24}\left(12s(s+c^2)+2s(6c-1)+c(c-1)(c+1)(c+2)\right),\\
B &= -\frac{1}{6}\left(12s(s+c^2)+2s(3c-4)+c(c-2)(c+1)(c+2)\right),\\
C &= \frac{1}{4}\left(12s(s+c^2)-10s+(c-1)(c-2)(c+1)(c+2)\right),\\
D &= -\frac{1}{6}\left(12s(s+c^2)-2s(3c+4)+c(c-2)(c-1)(c+2)\right),\\
E &= \frac{1}{24}\left(12s(s+c^2)-2s(6c+1)+c(c-1)(c+1)(c-2)\right).
\end{aligned}
$$

The modified equation of the scheme is given by

$$
u_t + u_x - \alpha\,u_{xx} = \frac{\Delta x^4}{120}\left(60s^2+20sc^2+c^4-5c^2+4-30s\right)\,u_{xxxxx}+\ldots. \tag{2.6.6}
$$

2.6 Fourth order explicit scheme

The scheme is essentially dispersive as the leading error terms are dispersive in nature due to the presence of the odd-order derivative term, u_{xxxxx}. The scheme is consistent and is fourth order accurate in space.

To obtain the stability conditions for the fourth order scheme we use the approach of Hindmarsh et al. (1984). The amplification factor of the scheme is given by

$$\xi = A\,\exp(-2I\omega) + B\,\exp(-I\omega) + C + D\,\exp(I\omega) + E\exp(2I\omega). \tag{2.6.7}$$

For $\omega = \pi$, we have

$$\xi = 1 - \frac{8\,c^2}{3} + \frac{2\,c^4}{3} + 8\,s^2 + 8\,s\,c^2 - \frac{16\,s}{3}. \tag{2.6.8}$$

For stability we need to have $|\xi| \leq 1$ and from (2.6.8), we obtain

$$-1 \leq 1 - \frac{8\,c^2}{3} + \frac{2\,c^4}{3} + 8\,s^2 + 8\,s\,c^2 - \frac{16\,s}{3} \leq 1,$$

which equivalently written as

$$\frac{c^4}{6} - \frac{5}{36} \leq \left(s - \left(\frac{1}{3} - \frac{c^2}{2}\right)\right)^2 \leq \frac{c^4}{6} + \frac{1}{9}. \tag{2.6.9}$$

Solving for s from the right hand side inequality of (2.6.9), we get

$$\frac{1}{3} - \frac{1}{2}c^2 - \frac{1}{6}\sqrt{4 + 6c^4} \leq s \leq \frac{1}{3} - \frac{1}{2}c^2 + \frac{1}{6}\sqrt{4 + 6c^4}, \tag{2.6.10}$$

and from the left hand side inequality, we have either

$$s \leq \frac{1}{3} - \frac{1}{2}c^2 - \frac{1}{6}\sqrt{6c^4 - 5} \tag{2.6.11}$$

or

$$s \geq \frac{1}{3} - \frac{1}{2}c^2 + \frac{1}{6}\sqrt{6c^4 - 5} \tag{2.6.12}$$

For $\frac{c^4}{6} \leq \frac{5}{36}$, it is only required to show the inequality (2.6.10) is satisfied. For the other case $\frac{c^4}{6} \geq \frac{5}{36}$, s and c should satisfy either of the inequalities

$$\frac{1}{3} - \frac{1}{2}c^2 - \frac{1}{6}\sqrt{4 + 6c^4} \leq s \leq \frac{1}{3} - \frac{1}{2}c^2 - \frac{1}{6}\sqrt{6c^4 - 5} \tag{2.6.13}$$

or

$$\frac{1}{3} - \frac{1}{2}c^2 + \frac{1}{6}\sqrt{6c^4 - 5} \leq s \leq \frac{1}{3} - \frac{1}{2}c^2 + \frac{1}{6}\sqrt{4 + 6c^4}. \tag{2.6.14}$$

When $\omega \to 0$, we use the following approximations

$$\cos(\omega) \simeq 1 - \frac{\omega^2}{2}; \quad \sin(\omega) \simeq \omega; \quad \sin(2\omega) \simeq 2\omega; \quad \cos(2\omega) \simeq 1 - 2\omega^2. \tag{2.6.15}$$

By neglecting the terms with higher degrees, we obtain

$$\xi^2 \simeq 1 - 2s\omega^2.$$

The inequality $|\xi| \leq 1$ holds only if

$$s \geq 0. \tag{2.6.16}$$

For the case $c^4 \leq \frac{5}{6}$, using (2.6.10), and (2.6.16), the scheme is stable if

$$0 \leq s \leq \frac{1}{3} - \frac{1}{2}c^2 + \frac{1}{6}\sqrt{4 + 6c^4}. \tag{2.6.17}$$

For the other case, i. e. $c^4 \geq \frac{5}{6}$, the inequalities (2.6.14) and (2.6.16) are satisfied only for $c \in [1, \sqrt{3}]$. Hence, for $c \in [1, \sqrt{3}]$ the scheme is stable if the conditions (2.6.14) and (2.6.16) are satisfied.

2.7 Non-standard finite difference scheme

In this section, we describe how the non-standard finite difference scheme (NSFD) has been constructed by Mickens (2000) for the 1D convection-diffusion equation.

The equation $u_t + u_x = \alpha\, u_{xx}$ has three sub-equations (Mickens, 1991, 1994) which are given by

$$u_t + u_x = 0, \tag{2.7.1}$$

$$u_x = \alpha\ u_{xx}, \tag{2.7.2}$$

$$u_t = \alpha\ u_{xx}. \tag{2.7.3}$$

2.7 Non-standard finite difference scheme

Eqs. (2.7.1) and (2.7.2) have known exact finite difference scheme which are

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+\frac{u_i^n-u_{i-1}^n}{\Delta x}=0,$$

and

$$\frac{u_i-u_{i-1}}{\Delta x}=\alpha\Big(\frac{u_{i+1}-2\,u_i+u_{i-1}}{\alpha\,\Delta x(\exp(\Delta x/\alpha)-1)}\Big),$$

respectively.

The NSFD is given by (Mickens, 1991, 2000)

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}+\frac{u_i^n-u_{i-1}^n}{\Delta x}=\alpha\Big(\frac{u_{i+1}^n-2u_i^n+u_{i-1}^n}{\alpha\,\Delta x(\exp(\Delta x/\alpha)-1)}\Big), \tag{2.7.4}$$

which on re-arranging gives

$$u_i^{n+1}=\beta_1\;u_{i+1}^n+(1-c-2\beta_1)\;u_i^n+(c+\beta_1)\;u_{i-1}^n, \tag{2.7.5}$$

where $c=\dfrac{\Delta t}{\Delta x}$ and $\beta_1=\dfrac{c}{\exp(\Delta x/\alpha)-1}$.

The modified equation of NSFD scheme is given by

$$u_t+u_x=\frac{\Delta x}{2}\left(-c+\frac{2\beta_1}{c}+1\right)u_{xx}+\frac{\Delta x^2}{6}(6s+c^2-1)u_{xxx}+\ldots, \tag{2.7.6}$$

and the leading error terms are dissipative in nature.

The square of the modulus of the amplification factor is given by

$$|\xi|^2=\Big((1-c-2\beta_1)+(c+2\;\beta_1)\;\cos(w)\Big)^2+(c\,\sin(w))^2. \tag{2.7.7}$$

For stability, $0<|\xi|\leq 1$ and this implies that $0<|\xi|^2\leq 1$. We now obtain the region of stability using the approach used in Hindmarsh et al. (1984).

We consider the case when $w=\pi$. The amplification factor of the NSFD is given by

$$\xi=1-2\;c-4\;\beta_1. \tag{2.7.8}$$

and therefore,

$$0\leq c+2\;\beta_1\leq 1. \tag{2.7.9}$$

Since c and β_1 are positive, we have $c + 2\beta_1 \geq 0$. Hence, we have the inequality

$$c + 2\beta_1 \leq 1. \tag{2.7.10}$$

We next consider the case when $w \to 0$. When $w \to 0$, $\cos(w) \approx 1 - \frac{1}{2}w^2$ and $\sin(w) \approx w$ and we get

$$c + 2\ \beta_1 - c^2 \geq 0. \tag{2.7.11}$$

From (2.7.10) and (2.7.11), we obtain $c^2 \leq 1$, which gives

$$\Delta t \leq \Delta x. \tag{2.7.12}$$

Thus the scheme is stable if it satisfies inequalities (2.7.10) and (2.7.12).

The stability can also be obtained from positivity property of nonstandard schemes. Here, we obtain

$$0 \leq c + 2\beta_1 \leq 1,$$

which is (2.7.9) and hence (2.7.10). This is one of the strengths of the NSFD methods.

Table 2.1 displays the regions of stability of the three schemes at three values of Re:10, 100, 10000 and also for two values of Δx namely 0.1 and 0.025. The three schemes are stable at Re=100 and Re=10000, when $\Delta x = 0.1$ and 0.025, with $\Delta t = 0.01$. However, the fourth order and NSFD schemes are not stable at Re=10, $\Delta x = 0.025$ and $\Delta t = 0.01$.

2.8 Numerical results

In this section, the numerical results obtained from the test problems, Test Case 1 and Test Case 2, at $T = 1$ are presented. For each test problem, the numerical profiles obtained using the three methods; third order upwind, fourth order and NSFD are plotted and some types of errors are tabulated.

2.8 Numerical results

Scheme	Re	Δx	Range of Δt for stability
Fourth Order	10	0.1	[0,0.0541]
	100	0.1	[0,0.1578]
		0.025	[0,0.0237]
	10000	0.1	[0,0.1001]∪[0.1725, 0.1995]
		0.025	$[0, 0.0251] \cup [0.0426, 0.0495]$
Third order	10	0.1	[0,0.0760]
	100	0.1	[0,0.1675]
		0.025	[0,0.0275]
	10000	0.1	[0,0.1996]
		0.025	[0,0.0496]
NSFD	10	0.1	[0,0.0333]
	100	0.1	[0,0.0833]
		0.025	[0,0.0138]
	10000	0.1	[0,0.0998]
		0.025	[0,0.0248]

Table 2.1: Stability regions when Re$= 10$, Re$= 100$ and Re$= 10000$ for Test Case 1.

Δx	Re	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
0.1	10	0.0012	0.0300	3.3114×10^{-7}	9.0556×10^{-7}	1.2370×10^{-6}
0.1	100	0.0635	0.1969	3.3448×10^{-4}	0.0033	0.0037
0.025	100	0.0011	0.0067	3.9951×10^{-8}	1.0893×10^{-6}	1.1293×10^{-6}
0.1	10000	0.1013	0.3052	9.0851×10^{-4}	0.0084	0.0093
0.025	10000	0.0252	0.1572	1.7086×10^{-5}	6.0090×10^{-4}	6.1798×10^{-4}

Table 2.2: Errors obtained from third order when $\Delta t = 0.01$, Re$= 10$, Re$= 100$ and Re$= 10,000$ for Test Case 1 at $T = 1$.

Δx	Re	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
0.1	10	0.0012	0.0030	3.5526×10^{-7}	8.8273×10^{-7}	1.2380×10^{-6}
0.1	100	0.1659	0.3973	0.0027	0.0224	0.0250
0.025	100	0.0029	0.0182	1.4444×10^{-7}	8.2459×10^{-6}	8.3904×10^{-6}
0.1	10000	0.3860	0.6031	0.0448	0.0906	0.1354
0.025	10000	0.0591	0.3102	1.4821×10^{-4}	0.0033	0.0034

Table 2.3: Errors obtained from fourth order when $\Delta t = 0.01$, Re$= 10$, Re$= 100$ and Re$= 10,000$ for Test Case 1 at $T = 1$.

Test Case 1

We tabulated the L_2, L_∞, dissipation, dispersion, total mean square errors and total variation at temporal step size 0.01 and Reynolds number: 10, 100 and 10000 for two values of Δx namely 0.1 and 0.025 for the three different methods in Table (2.2) to (2.5). For each method, it is seen that the errors are larger at higher Reynolds number for same values of Δx and Δt. The NSFD is by far the best scheme in terms of L_2, L_∞, dissipation and dispersion errors followed by the third order and fourth order schemes. The profiles from the three methods and exact profile are shown in Figs. (2.1a) to (2.1e). At Re$= 100$ and Re$= 10000$, with $\Delta x = 0.1$ (coarse grid), the results obtained from the third order and the fourth order are very oscillatory. However, the profile obtained using NSFD is very close to the exact profile. With $\Delta x = 0.025$, the profiles are less oscillatory as compared to $\Delta x = 0.1$ as expected. Again the NSFD is the most efficient shock-capturing scheme.

From Table 2.5, one can observe that at Re$= 100$ and Re$= 10000$, the total variation obtained from the third order and the fourth order are larger when $\Delta x = 0.1$ as compared to the case $\Delta x = 0.025$. In all cases considered the total variation for the NSFD scheme is much less than for the other methods.

Fig. 2.2 shows that the NSFD is TVD and the other two methods are not TVD. This is because the total variation for the third order upwind and fourth order methods are increasing near the initial time.

2.8 Numerical results

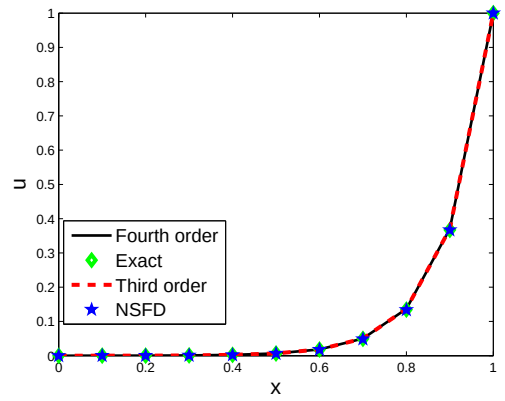
Δx	Re	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
0.1	10	3.6413×10^{-4}	6.7422×10^{-4}	5.8429×10^{-8}	6.2106×10^{-8}	1.2054×10^{-7}
0.1	100	3.8220×10^{-10}	1.2086×10^{-9}	1.3293×10^{-20}	3.6701×10^{-17}	1.3280×10^{-19}
0.025	100	1.4133×10^{-12}	8.8194×10^{-12}	7.0285×10^{-26}	1.0591×10^{-17}	1.9488×10^{-24}
0.1	10000	0	0	0	0	0
0.025	10000	4.2204×10^{-110}	2.6692×10^{-109}	0	5.2836×10^{-18}	5.2836×10^{-18}

Table 2.4: Errors obtained from NSFD when $\Delta t = 0.01$, Re$= 10$, Re$= 100$ and Re$= 10,000$ for Test Case 1 at $T = 1$.

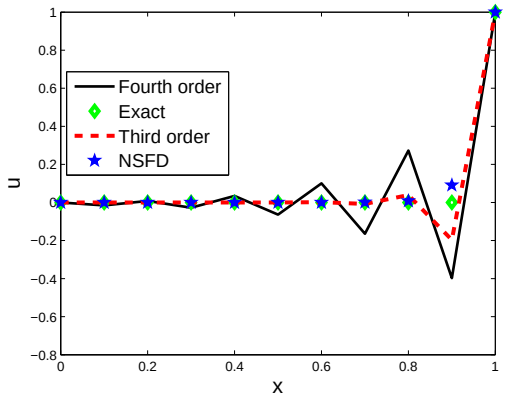
Numerical methods	Re	Δx	Total Variation	CPU time to solve and compute all errors	CPU time to solve u
Fourth Order	10	0.1	1	4.246	0.016
	100	0.1	3.1664	4.157	0.085
		0.025	1	8.718	0.099
	10000	0.1	7.6374	3.703	0.071
		0.025	2.3420	4.186	0.081
Third Order	10	0.1	1	4.363	0.072
	100	0.1	1.4902	4.106	0.084
		0.025	1	8.521	0.098
	10000	0.1	1.8783	3.545	0.074
		0.025	1.3730	3.961	0.095
NSFD	10	0.1	1	4.127	0.084
	100	0.1	1	4.262	0.073
		0.025	1	8.754	0.078
	10000	0.1	1	3.662	0.030
		0.025	1	4.295	0.075

Table 2.5: Total Variation and CPU time for different values of Δx when $\Delta t = 0.01$ for Test case 1 at $T = 1$.

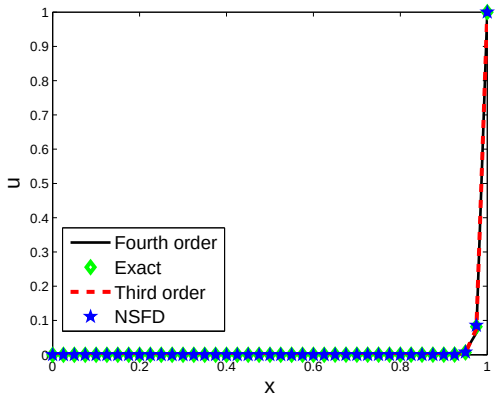
2.8 Numerical results



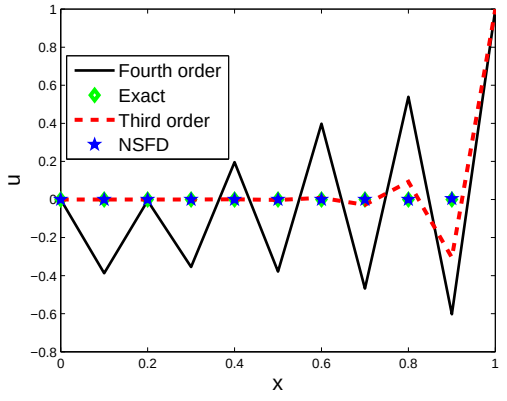
(a) $\Delta x = 0.1$ and Re=10



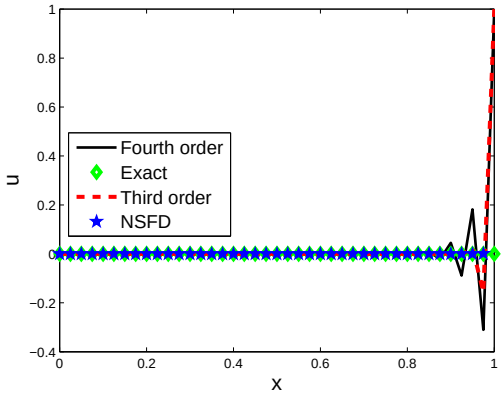
(b) $\Delta x = 0.1$ and Re=100



(c) $\Delta x = 0.025$ and Re=100

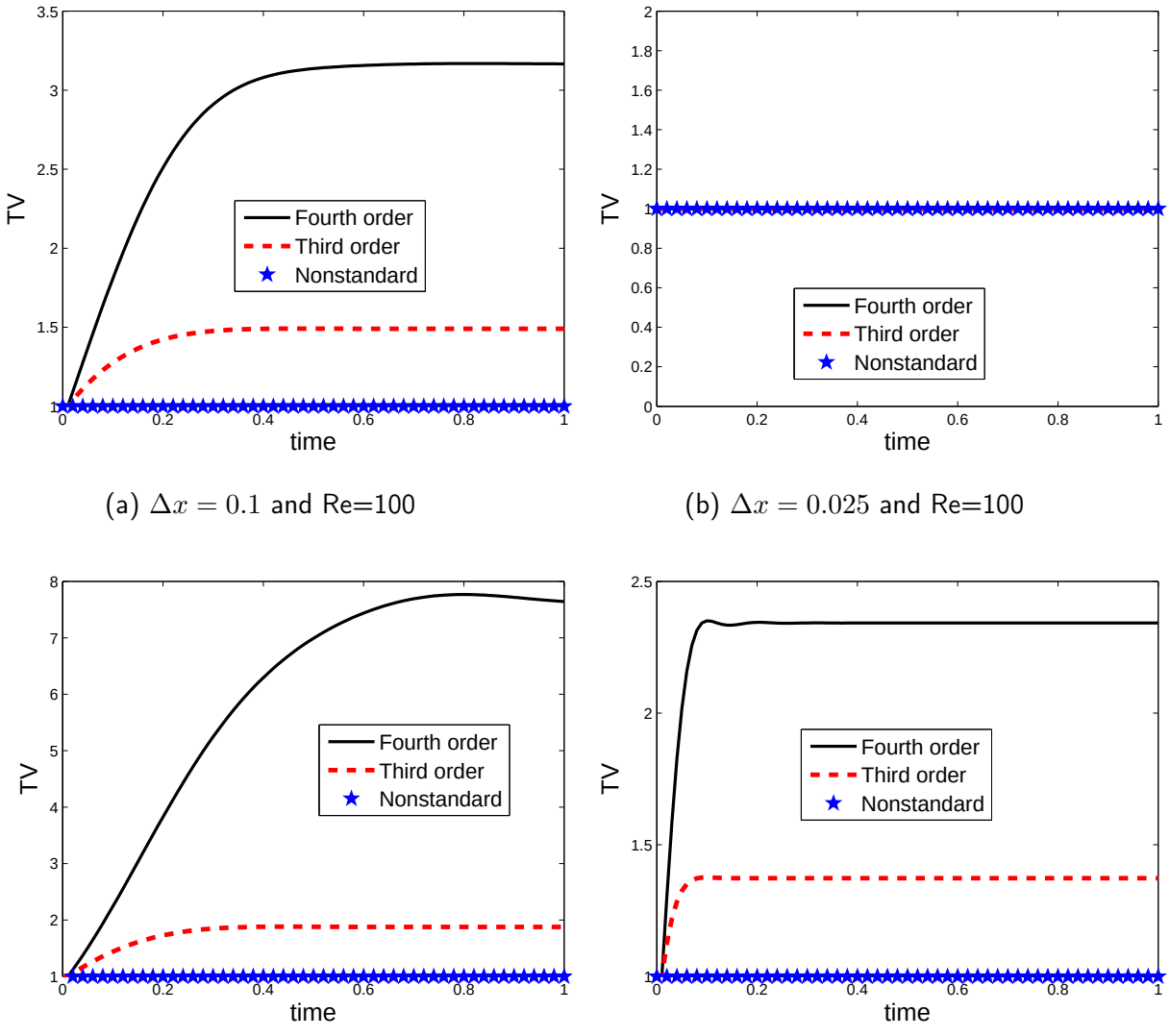


(d) $\Delta x = 0.1$ and Re=10000



(e) $\Delta x = 0.025$ and Re=10000

Figure 2.1: Comparison of the numerical schemes when Re=10, Re=100 and Re=10000 for Test case 1 at $T = 1$.



(a) $\Delta x = 0.1$ and Re=100

(b) $\Delta x = 0.025$ and Re=100

(c) $\Delta x = 0.1$ and Re=10000

(d) $\Delta x = 0.025$ and Re=10000

Figure 2.2: Total variation of the numerical schemes when Re=100 and Re=10000, $\Delta t = 0.01$ for Test case 1 at $T = 1$.

Test case 2

Test case 2 has been described in section 2.3.2 where we consider the following values of α and Δx:

(a) $\alpha = 0.01$, $\Delta x = 0.1$
(b) $\alpha = 0.1$, $\Delta x = 0.05$
(c) $\alpha = 1$, $\Delta x = 0.05$
(d) $\alpha = 1$, $\Delta x = 0.1$

Fig. 2.3 compares the profile when $\alpha = 0.01$, $\Delta x = 0.1$ at four different values of Δt. The fourth order scheme is very oscillatory and the third order is quite oscillatory. Table 2.7 compares the errors. It is seen that the NSFD performs the best in regard to L_2, L_∞ and Total Mean Square errors. The dissipation error is greater than dispersion error for the third and fourth order schemes.

For the case $\alpha = 0.1$ and $\Delta x = 0.05$, the exact profile is quite smooth as the coefficient of diffusivity is larger. The results from NSFD are very close to the exact profile. There is some dispersion and an overshoot in the peak from the third order and fourth order methods as shown Fig. (2.4). From Table 2.8, we can deduce the NSFD is the most effective scheme.

When $\alpha = 1$, the errors obtained using the three methods are very small as shown in Tables 2.9 and 2.10 and again the NSFD performs the best.

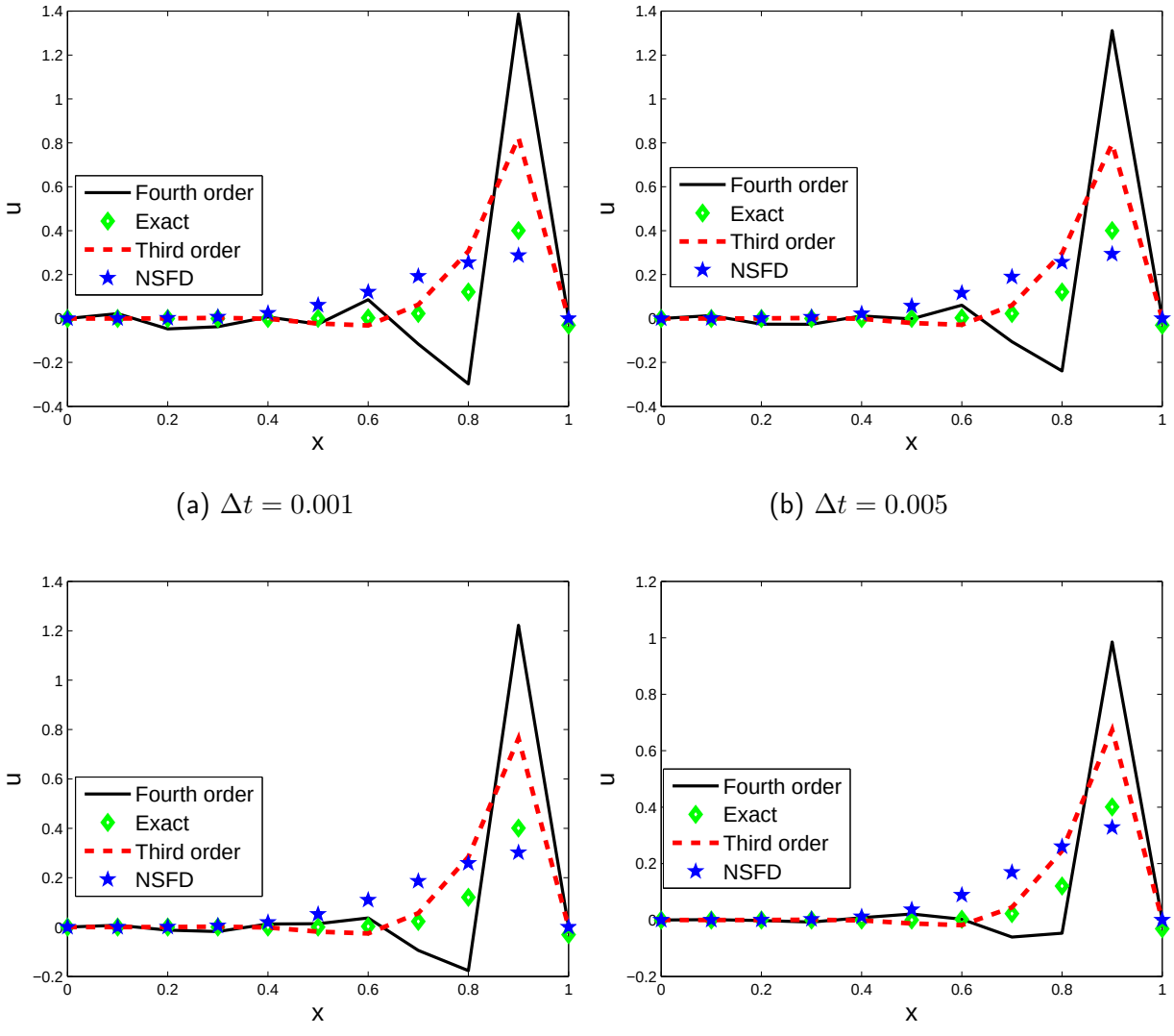
2.8 Numerical results

Scheme	α	Δx	Range of Δt for stability
Fourth Order	0.01	0.1	[0,0.1578]
	0.1	0.05	[0,0.0155]
	1	0.05	[0,0.0016]
	1	0.1	[0,0.0066]
Third order	0.01	0.1	[0,0.1675]
	0.1	0.05	[0,0.0314]
	1	0.05	[0,0.0012]
	1	0.1	[0,0.0053]
NSFD	0.01	0.1	[0,0.0999]
	0.1	0.05	[0,0.0122]
	1	0.05	[0,0.0012]
	1	0.1	[0,0.0049]

Table 2.6: Stability regions for three different values of α namely $0.01, 0.1$ and 1 for Test Case 2.

Scheme	Δt	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
Third Order	0.001	0.1475	0.4226	0.0194	3.9638×10^{-4}	0.0198
	0.005	0.1384	0.3958	0.0170	3.7389×10^{-4}	0.0174
	0.01	0.1274	0.3636	0.0144	3.4468×10^{-4}	0.0147
	0.025	0.0960	0.2733	0.0081	2.5683×10^{-4}	0.0084
Fourth Order	0.001	0.3436	0.9866	0.0939	0.0134	0.1073
	0.005	0.3133	0.9108	0.0780	0.0112	0.0892
	0.01	0.2792	0.8215	0.0619	0.0090	0.0709
	0.025	0.1944	0.5845	0.0297	0.0046	0.0344
NSFD	0.001	0.0888	0.1696	0.0017	0.0055	0.0072
	0.005	0.0868	0.1670	0.0016	0.0052	0.0069
	0.01	0.0842	0.1630	0.0015	0.0049	0.0064
	0.025	0.0748	0.1462	0.0012	0.0039	0.0051

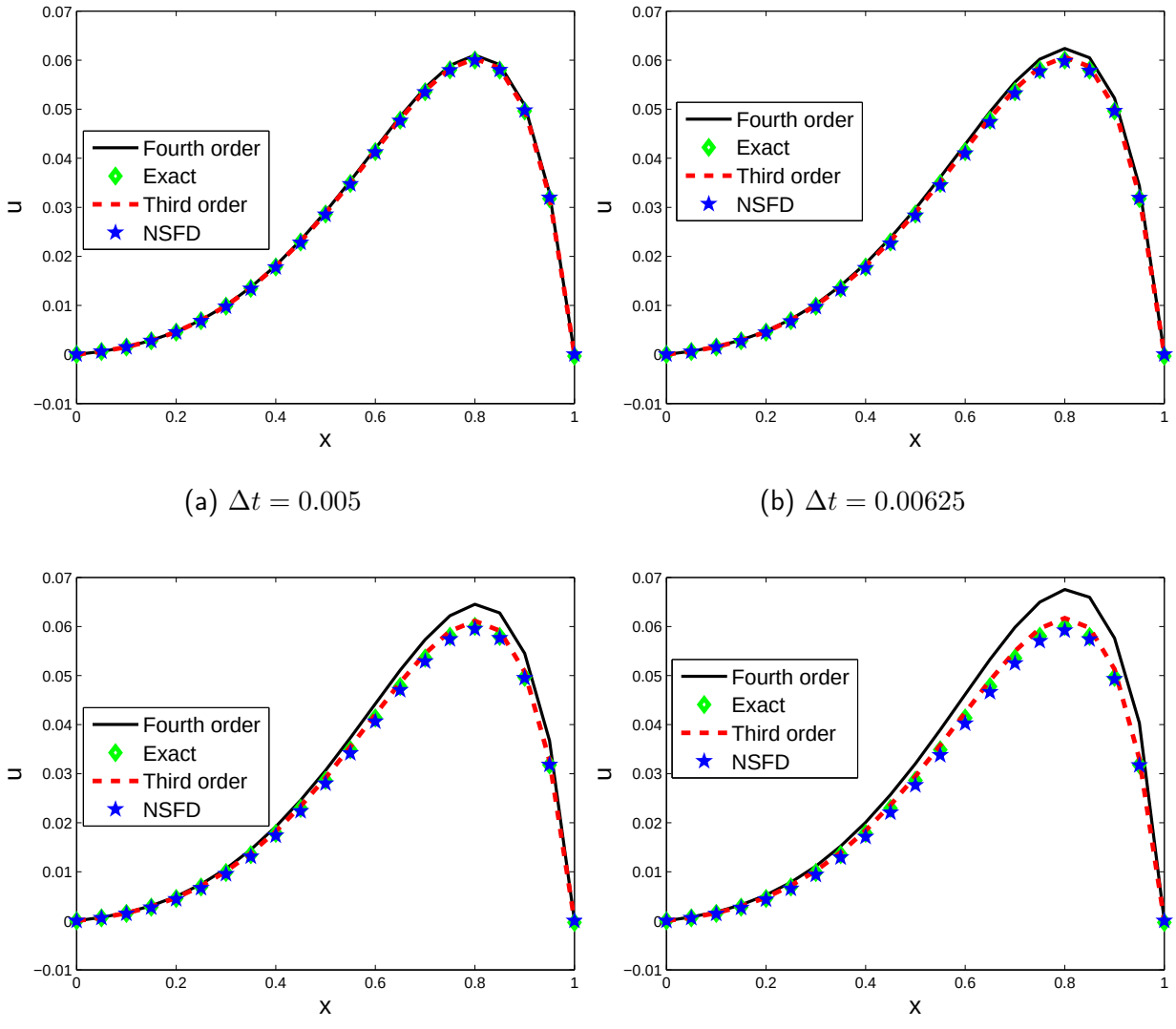
Table 2.7: Comparison of the numerical schemes when $\alpha = 0.01$ and $\Delta x = 0.1$ for Test Case 2 at $T = 1$.



(a) $\Delta t = 0.001$ (b) $\Delta t = 0.005$

(c) $\Delta t = 0.01$ (d) $\Delta t = 0.025$

Figure 2.3: Comparison of the numerical schemes when $\alpha = 0.01$ and $\Delta x = 0.1$ for Test Case 2 at $T = 1$.



(a) $\Delta t = 0.005$ (b) $\Delta t = 0.00625$

(c) $\Delta t = 0.008$ (d) $\Delta t = 0.01$

Figure 2.4: Comparison of the numerical schemes when $\alpha = 0.1$ and $\Delta x = 0.05$ for Test Case 2 at $T = 1$.

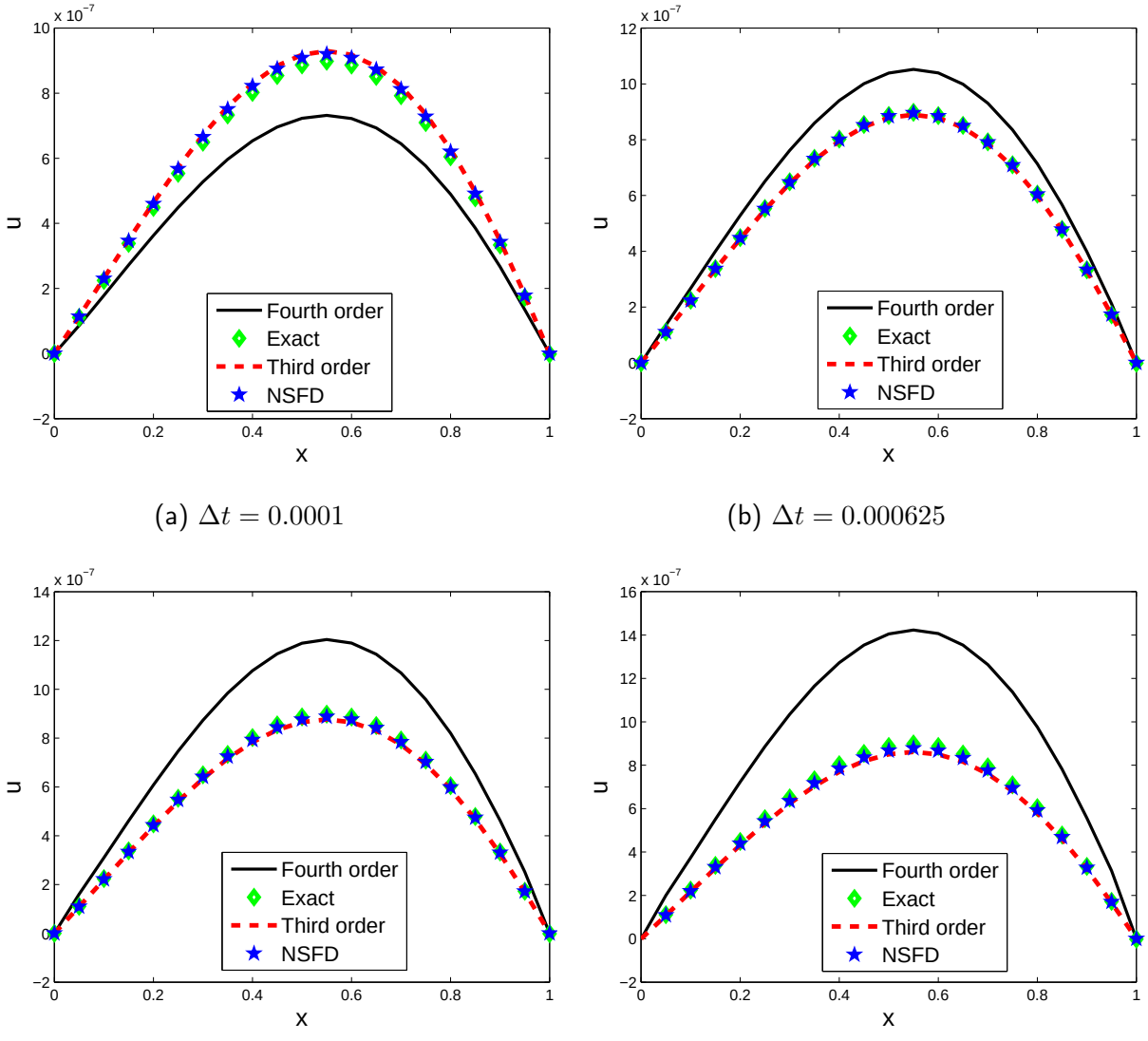
2.8 Numerical results

Scheme	Δt	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
Third Order	0.005	2.1355×10^{-4}	3.7906×10^{-4}	3.6010×10^{-8}	7.4241×10^{-9}	4.3434×10^{-8}
	0.00625	3.9782×10^{-4}	6.8387×10^{-4}	1.4248×10^{-7}	8.2386×10^{-9}	1.5072×10^{-7}
	0.008	6.7694×10^{-4}	0.0012	4.2648×10^{-7}	9.9494×10^{-9}	4.3643×10^{-7}
	0.01	0.0010	0.0018	9.7346×10^{-7}	1.2862×10^{-8}	9.8632×10^{-7}
Fourth Order	0.005	6.4476×10^{-4}	0.0012	3.6733×10^{-7}	2.8590×10^{-8}	3.9592×10^{-7}
	0.00625	0.0015	0.0026	1.9726×10^{-6}	1.0343×10^{-7}	2.0761×10^{-6}
	0.008	0.0028	0.0050	7.2691×10^{-6}	3.49477×10^{-7}	7.6186×10^{-6}
	0.01	0.0047	0.0086	2.0288×10^{-5}	1.0097×10^{-6}	2.1297×10^{-5}
NSFD	0.005	1.1822×10^{-4}	3.3094×10^{-4}	1.4911×10^{-9}	1.1819×10^{-8}	1.3310×10^{-8}
	0.00625	2.4401×10^{-4}	3.8193×10^{-4}	3.2902×10^{-8}	2.3803×10^{-8}	5.6706×10^{-8}
	0.008	4.5680×10^{-4}	7.2895×10^{-4}	1.4871×10^{-7}	5.0027×10^{-8}	1.9873×10^{-7}
	0.01	7.1049×10^{-4}	0.0011	3.8799×10^{-7}	9.2768×10^{-8}	4.8075×10^{-7}

Table 2.8: Comparison of the numerical schemes when $\alpha = 0.1$ and $\Delta x = 0.05$ for Test Case 2 at $T = 1$.

Scheme	Δt	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
Third Order	0.0001	2.2371×10^{-8}	3.1736×10^{-8}	4.7587×10^{-16}	7.7443×10^{-19}	4.7665×10^{-16}
	0.000625	5.7507×10^{-9}	8.4119×10^{-9}	3.1399×10^{-17}	9.7311×10^{-20}	3.1496×10^{-17}
	0.0008	1.5026×10^{-8}	2.1688×10^{-8}	2.1497×10^{-16}	6.7603×10^{-20}	2.1504×10^{-16}
	0.001	2.5597×10^{-8}	3.6796×10^{-8}	6.2384×10^{-16}	1.4553×10^{-19}	6.2399×10^{-16}
Fourth Order	0.0001	1.1829×10^{-7}	1.6605×10^{-7}	1.3322×10^{-14}	3.8972×10^{-18}	1.3326×10^{-14}
	0.000625	1.1035×10^{-7}	1.5428×10^{-7}	1.1593×10^{-14}	4.7517×10^{-18}	1.1597×10^{-14}
	0.0008	2.1916×10^{-7}	3.0609×10^{-7}	4.5725×10^{-14}	1.8239×10^{-17}	4.5743×10^{-14}
	0.001	3.7680×10^{-7}	5.2541×10^{-7}	1.3516×10^{-13}	5.6404×10^{-17}	1.3522×10^{-13}
NSFD	0.0001	1.5801×10^{-8}	2.2254×10^{-8}	2.3752×10^{-16}	2.6164×10^{-19}	2.3778×10^{-16}
	0.000625	1.5277×10^{-9}	2.2411×10^{-9}	1.9681×10^{-18}	2.5469×10^{-19}	2.2228×10^{-18}
	0.0008	7.0858×10^{-9}	1.0205×10^{-8}	4.7565×10^{-17}	2.5241×10^{-19}	4.7818×10^{-17}
	0.001	1.3492×10^{-8}	1.9322×10^{-8}	1.7311×10^{-16}	2.4982×10^{-19}	1.7336×10^{-16}

Table 2.9: Comparison of the numerical schemes when $\alpha = 1$ and $\Delta x = 0.05$ for Test Case 2 at $T = 1$.



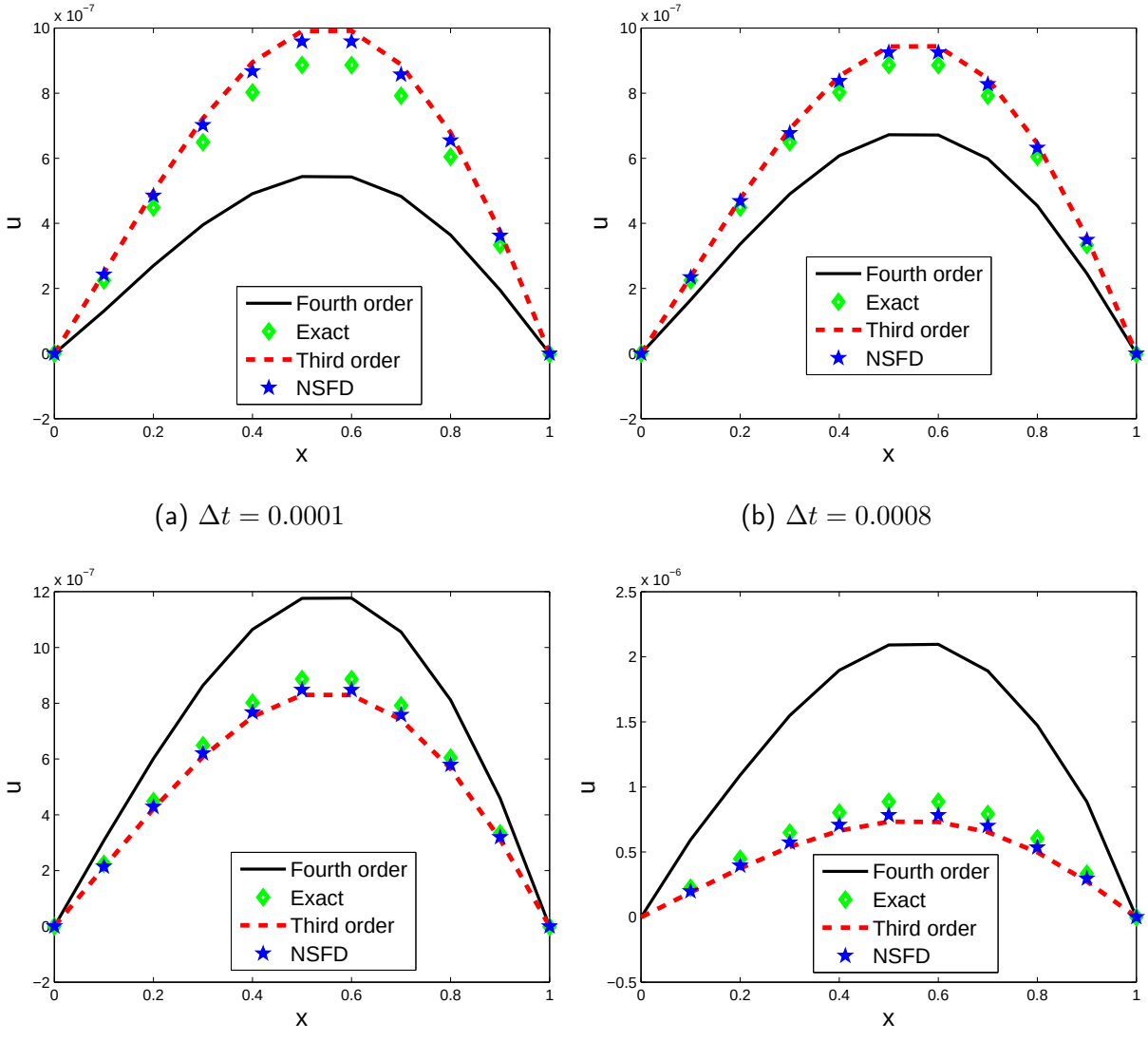
(a) $\Delta t = 0.0001$

(b) $\Delta t = 0.000625$

(c) $\Delta t = 0.0008$

(d) $\Delta t = 0.001$

Figure 2.5: Comparison of the numerical schemes when $\alpha = 1$ and $\Delta x = 0.05$ for Test Case 2 at $T = 1$.



(a) $\Delta t = 0.0001$ (b) $\Delta t = 0.0008$

(c) $\Delta t = 0.0025$ (d) $\Delta t = 0.004$

Figure 2.6: Comparison of the numerical schemes when $\alpha = 1$ and $\Delta x = 0.1$ for Test Case 2 at $T = 1$.

Scheme	Δt	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE
Third Order	0.0001	1.0558×10^{-7}	5.0247×10^{-15}	5.3754×10^{-18}	5.0301×10^{-15}	5.0301×10^{-15}
	0.0008	4.0576×10^{-8}	5.7541×10^{-8}	1.4946×10^{-15}	2.1423×10^{-18}	1.4968×10^{-15}
	0.0025	4.0199×10^{-8}	5.7023×10^{-8}	1.4687×10^{-15}	3.8842×10^{-19}	1.4690×10^{-15}
	0.004	1.0981×10^{-7}	1.5554×10^{-7}	1.0958×10^{-14}	4.3849×10^{-18}	1.0962×10^{-14}
Fourth Order	0.0001	2.4713×10^{-7}	3.4362×10^{-7}	5.5504×10^{-14}	1.7528×10^{-17}	5.5522×10^{-14}
	0.0008	1.5477×10^{-7}	2.1495×10^{-7}	2.1770×10^{-14}	7.3444×10^{-18}	2.1777×10^{-14}
	0.0025	2.1010×10^{-7}	2.9041×10^{-7}	4.0109×10^{-14}	1.9375×10^{-17}	4.0128×10^{-14}
	0.004	8.7923×10^{-7}	1.2089×10^{-7}	7.0240×10^{-13}	3.7892×10^{-16}	7.0278×10^{-13}
NSFD	0.0001	5.1603×10^{-8}	7.2322×10^{-8}	2.4205×10^{-15}	2.8981×10^{-19}	2.4208×10^{-15}
	0.0008	2.7793×10^{-8}	3.8923×10^{-8}	7.0193×10^{-16}	2.7971×10^{-19}	7.0221×10^{-16}
	0.0025	2.7513×10^{-8}	3.8847×10^{-8}	6.8789×10^{-16}	2.5627×10^{-19}	6.8815×10^{-16}
	0.004	7.3389×10^{-8}	1.0321×10^{-7}	4.8961×10^{-15}	2.3682×10^{-19}	4.8963×10^{-15}

Table 2.10: Comparison of the numerical schemes when $\alpha = 1$ and $\Delta x = 0.1$ for Test Case 2 at $T = 1$.

2.8 Numerical results

Numerical methods	α	Δx	Δt	Total Variation	CPU time to solve and compute all errors	CPU time to solve u
Fourth Order	0.01	0.1	0.001	3.7467	3.938	0.099
			0.005	3.7467	3.938	0.030
			0.01	3.7467	3.938	0.026
			0.025	3.7467	3.938	0.020
	0.1	0.05	0.005	0.1221	4.922	0.031
			0.00625	0.1248	4.923	0.027
			0.008	0.1291	4.947	0.026
			0.01	0.1351	4.954	0.025
	1	0.05	0.0001	1.4639×10^{-6}	6.238	1.382
			0.000625	2.1045×10^{-6}	5.022	0.122
			0.0008	2.4082×10^{-6}	4.979	0.096
			0.001	2.8468×10^{-6}	4.970	0.084
		0.1	0.0001	1.0875×10^{-6}	4.576	0.683
			0.0008	1.3444×10^{-6}	3.896	0.070
			0.0025	2.3537×10^{-6}	3.910	0.029
			0.004	4.1907×10^{-6}	3.838	0.021
Third Order	0.01	0.1	0.001	1.7159	3.757	0.062
			0.005	1.6560	3.749	0.030
			0.01	1.5843	3.751	0.020
			0.025	1.3859	3.775	0.018
	0.1	0.05	0.005	0.1206	4.892	0.037
			0.00625	0.1213	4.866	0.039
			0.008	0.1223	4.874	0.034
			0.01	0.1234	4.846	0.032
	1	0.05	0.0001	1.8595×10^{-6}	6.336	1.356
			0.000625	1.7792×10^{-6}	4.915	0.122
			0.0008	1.7526×10^{-6}	4.915	0.101
			0.001	1.7224×10^{-6}	4.913	0.084
		0.1	0.0001	1.9841×10^{-6}	4.310	0.653
			0.0008	1.8880×10^{-6}	3.769	0.048
			0.0025	1.6603×10^{-6}	3.771	0.021
			0.004	1.4648×10^{-6}	3.824	0.013
NSFD	0.01	0.1	0.001	0.5743	3.738	0.055
			0.005	0.5869	3.747	0.018
			0.01	0.6034	3.739	0.017
			0.025	0.6563	3.740	0.017
	0.1	0.05	0.005	0.1199	4.864	0.040
			0.00625	0.1196	4.868	0.033
			0.008	0.1190	4.847	0.032
			0.01	0.1184	4.859	0.015
	1	0.05	0.0001	1.8405×10^{-6}	6.058	1.305
			0.000625	1.7916×10^{-6}	4.954	0.116
			0.0008	1.7756×10^{-6}	4.936	0.097
			0.001	1.7573×10^{-6}	4.910	0.077
		0.1	0.0001	1.9179×10^{-6}	4.365	0.638
			0.0008	1.8511×10^{-6}	3.814	0.061
			0.0025	1.6960×10^{-6}	3.768	0.036
			0.004	1.5673×10^{-6}	3.762	0.016

Table 2.11: Total Variation and CPU time for different values of Δx and Δt for Test Case 2 at $T = 1$.

2.9 Optimization

From the numerical results obtained for Test Case 1, it is observed that all errors obtained from NSFD are very small. The results obtained from third order upwind shows that the dispersive errors are much greater than the dissipation errors for the case $\Delta x = 0.1$. In the case of fourth order scheme, both dispersion and dissipation errors are very large and are almost the same in magnitude.

Our aim in this section is to find an optimal value of Δt that minimizes the dispersion error when $\Delta x = 0.1$ and Re=100, using test case 1. Tam and Webb (1993), Bogey and Bailly (2004) and Berland et al. (2007) have implemented techniques which enable coefficients to be determined in high order numerical methods for computational aeroacoustics applications.

The Dispersion-Relation-Preserving (DRP) scheme was constructed (Tam and Webb, 1993) such that the dispersion relation of the finite difference scheme is formally the same as that of the original partial differential equation. The integrated error is calculated as

$$E = \int_{-\eta}^{\eta} |\theta^* \Delta x - \theta \Delta x|^2 \, d(\theta \, \Delta x),$$

where the quantities $\theta^* \Delta x$ and $\theta \, \Delta x$ represent the numerical and exact wave numbers respectively. The dispersion and dissipation error are obtained as $|\Re(\theta^* \Delta x) - \theta \, \Delta x|$ and $|\Im(\theta^* \Delta x)|$, respectively. Tam and Shen (1993) set η as 1.1 and optimize the coefficient in the numerical scheme such that the integrated error is minimized.

In Appadu (2012), these techniques were modified into equivalent forms so that the optimal CFL is computed for some known numerical schemes where the parameters were CFL and the phase angle. The author defined the following integrated errors integrated error from Tam and Webb (1993) (IETAM) and integrated error from Bogey and Bailly (2004) (IEBOGEY)

$$\text{IETAM} = \int_0^{\omega_1} |1 - RPE|^2 \, d\omega,$$
$$\text{IEBOGEY} = \int_0^{\omega_1} |1 - RPE| \, d\omega,$$

2.9 Optimization

with $\omega_1 = 1.1$ where $|1 - RPE|$ is the dispersion error.

In Appadu (2013), optimization techniques based on minimization of the dispersion error have been used to obtain the optimal Δt at a fixed value of Δx for the Lax-Wendroff and NSFD discretising a 1-D advection-diffusion equation and these techniques have been validated.

In this work, we use the same approach as in Appadu (2013). The amplification factor of the third order scheme approximating Eq. (2.1.1) with $\Delta x = 0.1$ is

$$\xi_{\text{Third}} = \Re(\xi) + I\Im(\xi),$$

and the relative phase error is given by

$$RPE_{\text{Third}} = \frac{-1}{10\,\Delta t\,\omega}\,\arctan(\frac{\Im(\xi)}{\Re(\xi)}),$$

where $\Re(\xi)$ and $\Im(\xi)$ are given by

$$\begin{aligned}\Re(\xi) &= 1 - 5.33\,\Delta t - 80\Delta t^2 + 60\Delta t^2\,\cos(\omega) + 20\Delta t^2\cos^2(\omega) + 333.33\,\Delta t^3\,\cos^2(\omega)\\ &\quad + 8.67\,\Delta t\,\cos(\omega) - 666.67\,\Delta t^3\,\cos(\omega) + 333.33\Delta t^3,\\ \Im(\xi) &= 3.33\,\Delta t\,\sin(\omega)\,\cos(\omega) - 13.33\,\Delta t\,\sin(\omega) - 333.33\,\Delta t^3\,\sin(\omega)\,\cos(\omega) + 20\Delta t^2\,\sin(\omega)\\ &\quad + 333.33\,\Delta t^3\,\sin(\omega) - 20\Delta t^2\,\sin(\omega)\,\cos(\omega).\end{aligned}$$

3D plots of the exact RPE versus $\Delta t \in [0, 0.1675]$ vs ω are shown in Figs. 2.7a and 2.7b and we have phase wrapping phenomenon. We thus obtain an approximation for the RPE till the terms $\mathcal{O}(\omega^4)$ using Taylor's series for $\omega \in [0, 1.1]$. The 3D plot of the approximated RPE vs Δt vs phase angle is shown in Fig. 2.7c.

The approximated RPE is obtained as

$$RPE_{\text{Approx}} = 1 - (0.0333 + 5\Delta t^2 - \Delta t + 50\Delta t^3 - 333.3333)\omega^4.$$

The integrated error is obtained as

$$\int_0^{1.1} (RPE_{\text{Approx}} - 1)^2 d\,\omega, \tag{2.9.1}$$

and is a function of Δt. A plot of the integrated error vs Δt is shown in Fig. 2.8a. The integrated error decreases as Δt increases from a value close to 0 to 0.05 and then it oscillates with local minimum near $\Delta t = 0.1$ and $\Delta t = 0.1414$ and then increases again as as shown in Fig. 2.8b. The scheme is stable if $\Delta t \in [0, 0.1675]$ and using NLPSolve, the optimal value of Δt is 0.1.

Table 2.12 compares the TV, L_2, L_∞, dissipation and dispersion errors, total mean square error and CPU time using the Third order upwind scheme at some different values of Δt with $\Delta x = 0.1$, Re$= 100$ and in Fig. 2.9, we obtain the plots of these error vs Δt.

It is seen that the errors initially decreases and reach a minimum when $\Delta t = 0.1$ and then increases again. We conclude that the variation of the integrated error in Fig. 2.8a mimic the actual variation of the L_2-error, L_∞-error, total variation, dissipation error, dispersion error and total mean square errors in Fig. 2.9. We conclude that the time step, $\Delta t = 0.1$ is indeed the optimal time step size which allows the method to perform at its best. Moreover, it can be observed from Table 2.12, there is no big change in CPU time when different values of Δt are used.

We also plot the exact RPE vs $\omega \in [0, \pi]$ at five different values of Δt namely $0.001, 0.08, 0.1, 0.125$ and $\frac{1}{6}$ as shown in Fig. 2.7d and it is seen that the scheme has best dispersion properties when $\Delta t = 0.1$.

2.10 Conclusion

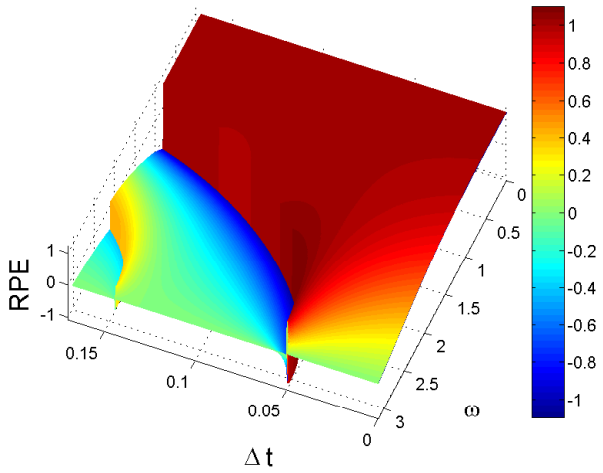
In this chapter, three numerical methods have been used to solve two test problems. The NSFD is much better than third order upwind and fourth order for all the cases considered. An optimization technique has been implemented for the third order scheme when Re=100 and Δx is fixed as 0.1 to find an optimal temporal step that minimizes the dispersion error. This optimal value is validated using numerical experiments. The computational time using NSFD is in general less as compared to the times using the fourth order and third order upwind schemes.

2.10 Conclusion

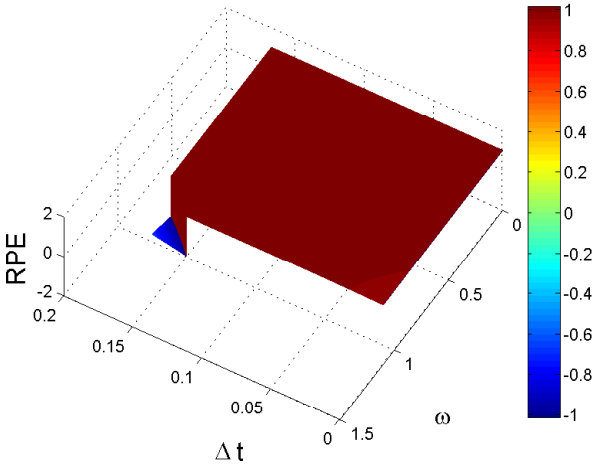
Δt	TV	L_2-error	L_∞-error	Diss. Error	Disp. Error	TMSE	CPU time
0.0001	1.6291	0.0779	0.2393	5.1174×10^{-4}	0.0050	0.0055	4.807
0.00025	1.6267	0.0777	0.2386	5.0844×10^{-4}	0.0050	0.0055	4.333
0.0005	1.6228	0.0773	0.2375	5.0299×10^{-4}	0.0049	0.0054	4.198
0.0008	1.6181	0.0768	0.2361	4.9656×10^{-4}	0.0049	0.0054	4.199
0.001	1.6150	0.0765	0.2352	4.9227×10^{-4}	0.0048	0.0053	4.199
0.0025	1.5922	0.0742	0.2285	4.6152×10^{-4}	0.0045	0.0050	4.128
0.005	1.5560	0.0705	0.2176	4.1457×10^{-4}	0.0041	0.0045	4.140
0.008	1.5156	0.0662	0.2050	3.6451×10^{-4}	0.0036	0.0040	4.178
0.01	1.4903	0.0635	0.1969	3.3448×10^{-4}	0.0033	0.0037	4.124
0.025	1.3333	0.0457	0.1429	1.7264×10^{-4}	0.0017	0.0019	4.139
0.05	1.1612	0.0237	0.0746	4.7694×10^{-5}	4.6144×10^{-4}	5.0914×10^{-4}	4.106
0.0625	1.1026	0.0155	0.0488	2.0753×10^{-5}	1.9668×10^{-4}	2.1743×10^{-4}	4.128
0.1	**1**	$\mathbf{1.4357 \times 10^{-5}}$	$\mathbf{4.5400 \times 10^{-5}}$	$\mathbf{1.8739 \times 10^{-11}}$	$\mathbf{1.6864 \times 10^{-10}}$	$\mathbf{1.8738 \times 10^{-10}}$	4.126
0.125	1	0.0025	0.0079	5.6869×10^{-7}	5.0676×10^{-6}	5.6363×10^{-6}	4.165
1/7	1.0060	9.6247×10^{-4}	0.0030	8.3942×10^{-8}	7.5819×10^{-7}	8.4213×10^{-7}	4.125
1/6	1.1319	0.0196	0.0619	3.3062×10^{-5}	3.1675×10^{-4}	3.4981×10^{-4}	4.148

Table 2.12: Errors obtained from third order when $\Delta x = 0.1$, Re$= 100$ for different values of Δt for Test case 1 at $T = 1$.

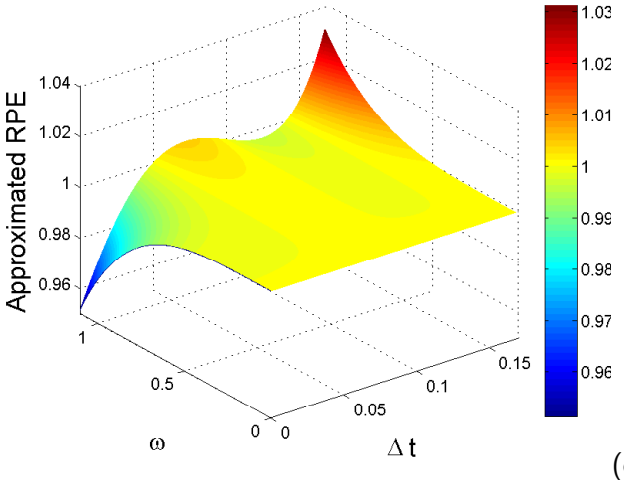
2.10 Conclusion



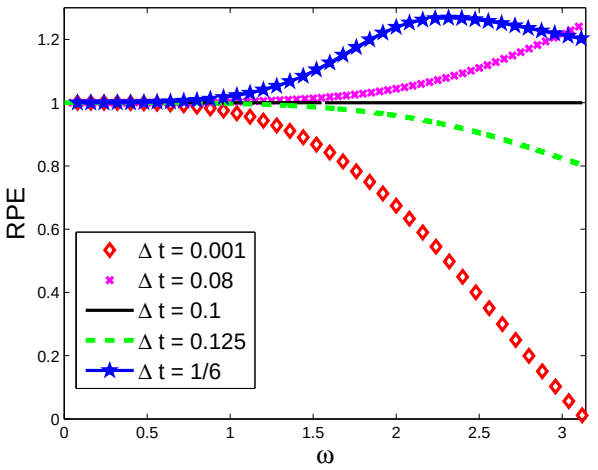
(a) Exact RPE versus Δt vs phase angle for $0 \leq \omega \leq \pi$



(b) Exact RPE versus Δt vs phase angle for $0 \leq \omega \leq 1.1$



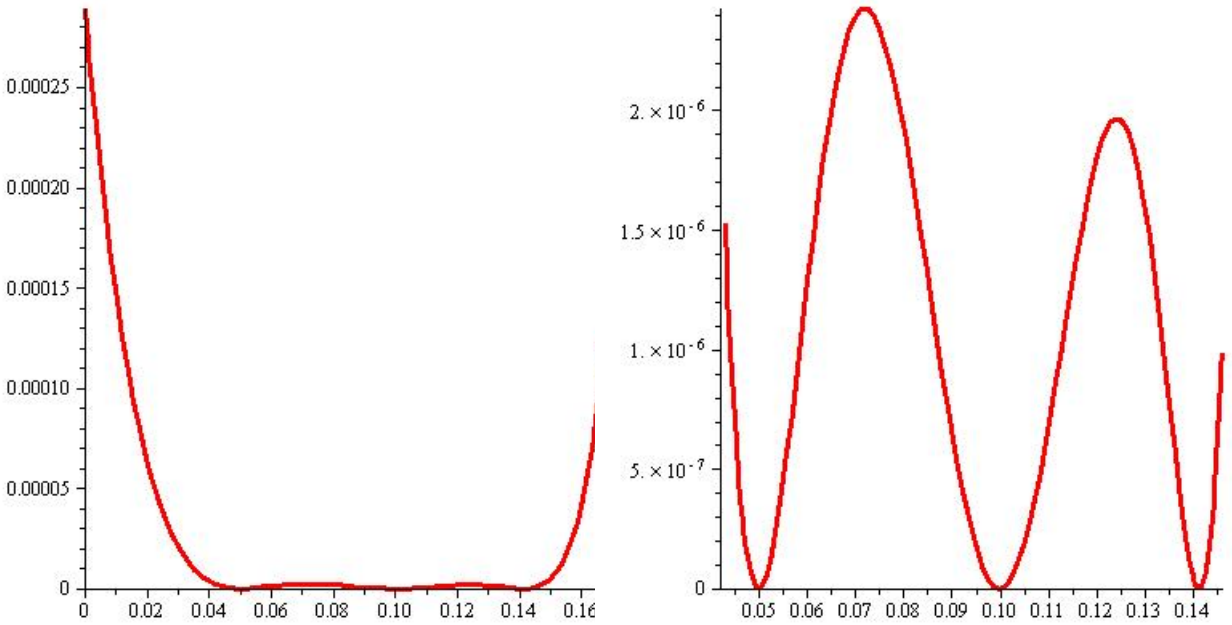
(c) Approximated RPE versus Δt vs phase angle for $0 \leq \omega \leq 1.1$



(d) Plots of Exact RPE vs $\omega \in [0, \pi]$ when $\Delta t = 0.001, 0.08, 0.1, 0.125$ and $\frac{1}{6}$ for the Third Order Upwind scheme.

Figure 2.7: Plots of RPE vs Δt vs ω for the third order upwind scheme for problem 1 with Re$= 100$ and $\Delta x = 0.1$.

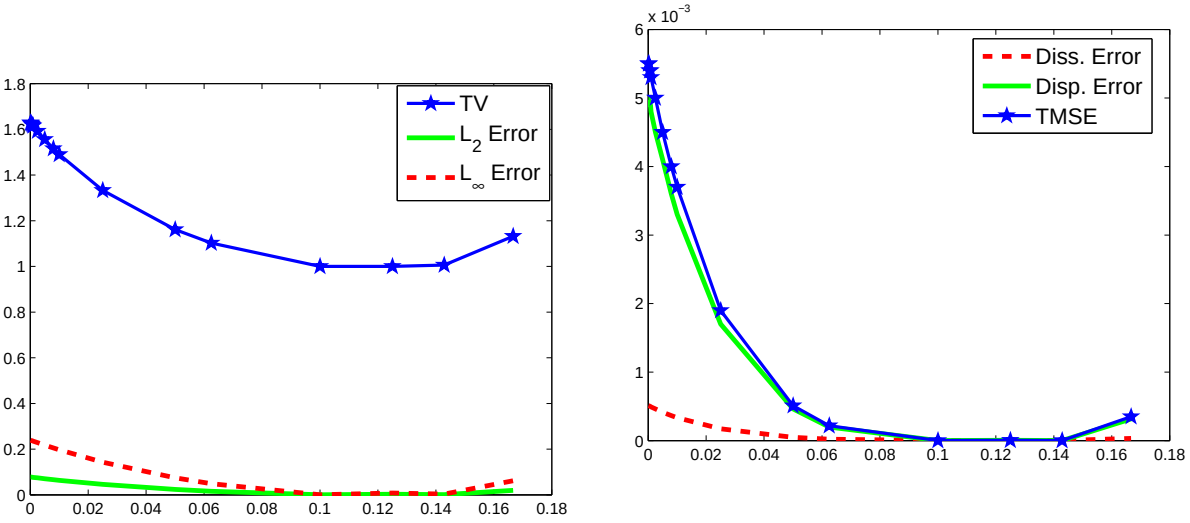
2.10 Conclusion



(a) Integrated error versus Δt

(b) Zoomed plot of Fig. 2.8a.

Figure 2.8: Integrated error versus Δt for third order upwind scheme when Re=100 and $h = 0.1$.



(a) L_2, L_∞ errors and TV versus Δt

(b) Dissipation, dispersion errors and TMSE versus Δt

Figure 2.9: Errors versus Δt for third order upwind scheme when Re=100 and $\Delta x = 0.1$ at $T = 1$.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Chapter 3

Performance of some finite difference methods for a 3D advection-diffusion equation

In this chapter, following the works of Dehghan (2005, 2007), a new finite difference method has been used to solve a 3D advection-diffusion equation and it is compared with two finite difference methods. An optimization technique has been implemented to find an optimal time step that minimizes dispersion errors and that performs better with the performance indices. This chapter is an extension of the work accepted for publication in Appadu et al. (b).

3.1 Introduction

The 3D advection-diffusion equation is given by:

$$\frac{\partial u}{\partial t} + \beta_x \frac{\partial u}{\partial x} + \beta_y \frac{\partial u}{\partial y} + \beta_z \frac{\partial u}{\partial z} = \alpha_x \frac{\partial^2 u}{\partial x^2} + \alpha_y \frac{\partial^2 u}{\partial y^2} + \alpha_z \frac{\partial^2 u}{\partial z^2}, \tag{3.1.1}$$

with initial condition

$$u(x, y, z, 0) = f(x, y, z), \quad 0 \leq x, y, z \leq 1, \tag{3.1.2}$$

and boundary conditions

$$\begin{aligned}
u(0,y,z,t) &= g_0(y,z,t), 0 \leq y,z \leq 1, 0 < t \leq T,\\
u(1,y,z,t) &= g_1(y,z,t), 0 \leq y,z \leq 1, 0 < t \leq T,\\
u(x,0,z,t) &= h_0(x,z,t), 0 \leq x,z \leq 1, 0 < t \leq T,\\
u(x,1,z,t) &= h_1(x,z,t), 0 \leq x,z \leq 1, 0 < t \leq T,\\
u(x,y,0,t) &= k_0(x,y,t), 0 \leq x,y \leq 1, 0 < t \leq T,\\
u(x,y,1,t) &= k_1(x,y,t), 0 \leq x,y \leq 1, 0 < t \leq T,
\end{aligned}$$

where $f, g_0, g_1, h_0, h_1, k_0$ and k_1 are known functions while u is unknown transported scalar variable; $\beta_x, \beta_y, \beta_z$ are constant velocity components of advection in the direction of x, y and z, respectively and α_x, α_y and α_z are constant diffusivities in the x, y and z directions, respectively.

Solving the analytical solution u of Eq. (3.1.1) is difficult and hence numerical method is advocated if one wants to get insight of the structure of solution (Yang et al., 1998).

In section 3.2, we discuss the discretization of the continuous region and describe the numerical experiment we consider. Section 3.3 discuses dispersion and dissipation of numerical methods. In section 3.4, we show how errors from numerical results are quantified into dispersion and dissipation errors and also define some performance indices. In Section 3.5, we describe how the numerical methods are constructed and obtain stability of the explicit scheme. Numerical results are presented in section 3.6 and in section 3.7, we obtain an optimal temporal step size that minimizes the dispersion error and which also performs better for the other performance indices and this is validated by numerical experiment.

3.2 Numerical experiment

The domain we consider is $0 \leq x,y,z \leq 1$. The grid points (x_i, y_j, z_k, t_n) are defined by

$$x_i = (i-1)\Delta x, i = 1, 2, \ldots, N_x,$$

3.2 Numerical experiment

$$y_j = (j-1)\Delta y, j = 1, 2, \ldots, N_y,$$
$$z_k = (k-1)\Delta z, k = 1, 2, \ldots, N_z,$$
$$t_n = (n-1)\Delta t, n = 1, 2, \ldots, N,$$

where $(N_x - 1)\Delta x = (N_y - 1)\Delta y = (N_z - 1)\Delta z = 1$ and $(N-1)\Delta t = T$.

Yang et al. (1998) studied the accuracy of three numerical methods, namely explicit central difference (ECD), implicit Crank-Nicolson (ICN) and implicit Chapeau function (ICF) methods for solving a 3D advection-diffusion equation as applied to spore and insect dispersal. The three numerical methods were used to simulate only the diffusion process. For the full advection-diffusion equation they only use the ICF for some advective and diffusive coefficients. The major criterion to test the accuracy was R^2 which represents the proportion of the total variation in particle distribution in all grid cells that is accounted for by the distribution through numerical solutions. High R^2 values were obtained for the ECD method with $\frac{\Delta t \alpha_z}{\Delta z^2} \leq 0.5$ and for the implicit methods with $\frac{\Delta t \alpha_z}{\Delta z^2} \leq 5$.

In this work, we compare the performance of three finite difference methods, namely: fourth order finite difference method (Dehghan, 2005), implicit Crank-Nicolson and implicit Chapeau function when used to solve the partial differential equation in (3.1.1) with initial conditions

$$u(x, y, z, 0) = \exp\left(-\frac{(x-0.5)^2}{\alpha_x} - \frac{(y-0.5)^2}{\alpha_y} - \frac{(z-0.5)^2}{\alpha_z}\right). \tag{3.2.1}$$

The analytical solution to this problem is (Dehghan, 2004)

$$u(x, y, z, t) = \frac{1}{(4t+1)^{3/2}} \exp\left[-\frac{(x-\beta_x t-0.5)^2}{\alpha_x\,(4t+1)} - \frac{(y-\beta_y t-0.5)^2}{\alpha_y\,(4t+1)} - \frac{(z-\beta_z t-0.5)^2}{\alpha_z\,(4t+1)}\right]. \tag{3.2.2}$$

The boundary conditions are obtained by substituting directly into Eq. (3.2.2).

Dehghan (2004) solved eq. (3.1.1) subject to the initial condition given by eq. (3.2.1) and using explicit methods: forward-time centred-space technique, forward-time backward-space centred-space scheme, Lax–Wendroff type technique and implicit methods: Backward

time centered-space fully implicit, the (7,7) fully implicit and implicit Crank-Nicolson type methods. Dehghan (2004) found that the explicit methods are very simple to implement and are economical to use as compared to the implicit methods. From the implicit methods described, it was found that the implicit Crank-Nicolson method is the most efficient method.

3.3 Numerical dispersion and dissipation

The term with the lowest-order even derivative in the truncation error produces amplitude error (loss of amplitude) in the approximate solution and this is responsible for *numerical dissipation*. The leading odd derivative in the truncation error produces small-scale waves, as different fourier components propagate at different phase speeds and this causes *numerical dispersion* (Durran, 1999).

The modulus of the amplification factor (AFM) is a measure of the stability of a scheme and it is also used to measure the dissipative characteristics of the scheme. The partial differential equation (3.1.1) is dissipative by nature due to the terms u_{xx} and u_{yy} in Eq. (3.1.1).

The relative phase error (RPE) is a measure of the dispersive characteristics of a scheme. The relative phase error of a scheme discretizing a 1D advection-diffusion equation, Eq. (2.1.1) from previous chapter, is obtained as

$$\text{RPE} = -\frac{1}{c\omega}\arctan\left(\frac{\Im(\xi_{num})}{\Re(\xi_{num})}\right),$$

where c is the Courant number, ω is phase angle, ξ_{num} is the amplification factor for a numerical scheme in approximating the 1D advection-diffusion equation and $\Re(\xi_{num})$ and $\Im(\xi_{num})$ are the real and imaginary parts of ξ_{num}, respectively.

We extend the work on relative phase error in Appadu et al. (2016) for the case of 3D advection-diffusion equation. We substitute u by $\exp(\gamma\, t)\, \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right)$ in equation (3.1.1), where $I = \sqrt{-1}$; t is time; θ_x, θ_y and θ_z are wave numbers in the

direction of x, y and z, respectively, and γ is the dispersion relation. Then we get

$$\frac{\partial u}{\partial t} = \gamma \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.1}$$
$$\frac{\partial u}{\partial x} = I\theta_x \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.2}$$
$$\frac{\partial u}{\partial y} = I\theta_y \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.3}$$
$$\frac{\partial u}{\partial z} = I\theta_z \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.4}$$
$$\frac{\partial^2 u}{\partial x^2} = -\theta_x^2 \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.5}$$
$$\frac{\partial^2 u}{\partial y^2} = -\theta_y^2 \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right), \tag{3.3.6}$$
$$\frac{\partial^2 u}{\partial z^2} = -\theta_z^2 \exp(\gamma\, t) \exp\left(I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right) \tag{3.3.7}$$

By substituting Equations (3.3.1)-(3.3.7) into Eqs. (3.1.1), we obtain

$$\gamma = -(\alpha_x\theta_x^2 + \alpha_y\theta_y^2 + \alpha_z\theta_z^2) - I(\beta_x\theta_x + \beta_y\theta_y + \beta_z\theta_z).$$

Hence we have

$$\begin{aligned} u(x,y,z,t) = &\exp\left[(-(\alpha_x\theta_x^2 I + \alpha_y\theta_y^2 I + \alpha_z\theta_z^2) - I(\beta_x\theta_x + \beta_y\theta_y + \beta_z\theta_z))\, t\right] \\ &\times \exp\left[I(\theta_x\, x + \theta_y\, y + \theta_z\, z)\right]. \end{aligned}$$

The exact amplification factor is obtained as (Anderson et al., 1984)

$$\begin{aligned} \xi_{exact} &= \frac{u(x,y,z,t+\Delta t)}{u(x,y,z,t)} \\ &= \exp\left[\left[-I(\beta_x\theta_x + \beta_y\theta_y + \beta_z\theta_z) - (\alpha_x\theta_x^2 + \alpha_y\theta_y^2 + \alpha_z\theta_z^2)\right]\Delta t\right]. \end{aligned}$$

The relative phase error of a numerical method is obtained using the relation (Anderson et al., 1984)

$$\text{RPE} = \frac{arg(\xi_{num})}{arg(\xi_{exact})},$$

where the argument of the exact amplification factor is given by

$$arg(\xi_{exact}) = -(\beta_x\theta_x + \beta_y\theta_y + \beta_z\theta_z)\,\Delta t,$$

and the amplification factor of the scheme denoted by ξ_{num}. Hence

$$\text{RPE} = \frac{arg(\xi_{num})}{-(\beta_x\theta_x + \beta_y\theta_y + \beta_z\theta_z)\,\Delta t}.$$

The amplification factor of the scheme can be written as $\xi_{num} = \xi_1 + I\xi_2$, where ξ_1 and ξ_2 are the real and imaginary parts of ξ_{num}. Hence, after some algebraic manipulations, we obtain

$$\text{RPE} = -\frac{1}{c_x\omega_x + c_y\omega_y + c_z\omega_z}\arctan\left(\frac{\xi_2}{\xi_1}\right),$$

where $\omega_x = \Delta x\,\theta_x$, $\omega_y = \Delta y\,\theta_y$ and $\omega_z = \Delta z\,\theta_z$ are phase angles in the direction of x, y and z, respectively; $c_x = \frac{\beta_x\Delta t}{\Delta x}$, $c_y = \frac{\beta_y\Delta t}{\Delta y}$ and $c_z = \frac{\beta_z\Delta t}{\Delta z}$ are the Courant numbers in the direction of x, y and z, respectively.

3.4 Performance indices and quantification of errors

To analyse and compare the performance of the three methods, we compute the L_2 and L_∞, dissipation, dispersion and total mean square errors and some performance indices such as total mass, mass conservation ratio (MCR), mass distribution ratio (MDR) and R^2 which is a measure of goodness fit (Kvalseth, 1985). We next describe how these performance indices are calculated. Let $u_{i,j,k}$ be the exact solution and $v_{i,j,k}$ the numerical (discrete) solution at a given grid point (i, j, k) and $\bar{v}$ and $\bar{u}$ their mean, respectively, then R^2, total mass, MCR and MDR are calculated as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(v_{i,j,k} - u_{i,j,k})^2}{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(v_{i,j,k} - \bar{v})^2},$$

$$\text{Total mass} = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} u^n_{i,j,k},$$

3.4 Performance indices and quantification of errors

$$\mathrm{MCR} = \frac{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} u^n_{i,j,k}}{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} v^n_{i,j,k}},$$

$$\mathrm{MDR} = \frac{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} (u^n_{i,j,k})^2}{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} (v^n_{i,j,k})^2}.$$

We note that MCR and MDR are calculated differently as compared to (Nguyen and Dabdub, 2001; Appadu et al., 2008), where the authors work with a 2D scalar advection equation.

Following the work of Appadu and Gidey (2013), we also extend the quantification of errors from numerical results into dispersion and dissipation errors discussed in section 2.4 for a 3D advection-diffusion problem. The Total Mean Square Error (TMSE) for the 3D case is calculated as

$$TMSE = \frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(u_{i,j,k} - v_{i,j,k})^2. \tag{3.4.1}$$

The total mean square error can be expressed as

$$\begin{aligned}\frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(u_{i,j,k} - v_{i,j,k})^2 &= \frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(u_{i,j,k} - \bar{u})^2 \\ &+ \frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(v_{i,j,k} - \bar{v})^2 + \frac{2}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} u_{i,j,k}\bar{u} \\ &+ \frac{2}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} v_{i,j,k}\bar{v} - \frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(\bar{u})^2 \\ &- \frac{1}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(\bar{v})^2 - \frac{2}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} u_{i,j,k}v_{i,j,k}.\end{aligned} \tag{3.4.2}$$

The right hand side of Eq. (3.4.2) can be rewritten as

$$\sigma^2(u) + \sigma^2(v) + 2(\bar{u})^2 + 2(\bar{v})^2 - (\bar{u})^2 - (\bar{v})^2 - \frac{2}{N_x\,N_y\,N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} u_{i,j,k}v_{i,j,k},$$

where $\sigma^2(u)$ and $\sigma^2(v)$ denote the variance of u and v, respectively, $\bar{u}$ and $\bar{v}$ denote the mean values of u and v, respectively. Then we have

$$TMSE = \sigma^2(u) + \sigma^2(v) + (\bar{u} - \bar{v})^2 - 2\,Cov(u, v),$$

where $Cov(u, v) = \dfrac{1}{N_x\,N_y\,N_z}\displaystyle\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}\Big(u_{i,j,k}v_{i,j,k} - \bar{u}\bar{v}\Big)$. The Total Mean Square Error can also expressed as

$$(\sigma(u) - \sigma(v))^2 + (\bar{u} - \bar{v})^2 + 2(1 - \rho)\sigma(u)\sigma(v), \tag{3.4.3}$$

where $\rho = Cov(u, v)/(\sigma(u)\sigma(v))$ is the coefficient of correlation. The expression $2(1 - \rho)\sigma(u)\sigma(v)$ measures the dispersion error and $(\sigma(u) - \sigma(v))^2 + (\bar{u} - \bar{v})^2$ measures the dissipation error.

We also obtain L_2 and L_∞ errors which are obtained using the following formulae:

$$L_2\text{-error} = \sqrt{\Delta x\Delta y\Delta z \sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z}(u_{i,j,k} - v_{i,j,k})^2}, \tag{3.4.4}$$

$$L_\infty\text{-error} = \max_{1\le i\le N_x,\, 1\le j\le N_y,\, 1\le k\le N_z} |u_{i,j,k} - v_{i,j,k}|. \tag{3.4.5}$$

The rate of convergence of a numerical method is obtained using $\log\left(\frac{e_1}{e_2}\right) \Big/ \log\left(\frac{h_1}{h_2}\right)$, where e_1 and e_2 are L_2-errors with h_1 and h_2 spatial step sizes, respectively.

3.5 Numerical methods and stability analysis

3.5.1 Stability analysis

The von Neumann stability analysis in three dimensions is the generalization of the one-dimensional case. We use the ansatz

$$u^n_{i,j,k} = \xi^n \exp\Big[I(\omega_x\, i + \omega_y\, j + \omega_z\, k)\Big],$$

where $I = \sqrt{-1}$; $\xi = \xi(\omega_x, \omega_y, \omega_z)$ is an amplification factor; $\omega_x = \theta_x\Delta x$, $\omega_y = \theta_y\Delta y$ and $\omega_z = \theta_z\Delta z$ are phase angles in the direction of x, y and z, respectively; θ_x, θ_y and θ_z are wave numbers in the direction of x, y and z, respectively.

The region around $(0, 0, 0)$ corresponds to the low frequencies and while around the region (π, π, π) corresponds to the highest frequencies (Sousa, 2003). An amplification factor, ξ, is said to satisfy the von Neumann condition if there is a constant K such that

$$|\xi(\omega_x, \omega_y, \omega_z)| \leq 1 + K\Delta t, \forall\, \omega_x, \omega_y, \omega_z \in [-\pi, \pi]. \tag{3.5.1}$$

In practice, we use the strong condition as in the one-dimensional case

$$|\xi(\omega_x, \omega_y, \omega_z)| \leq 1, \forall\, \omega_x, \omega_y, \omega_z \in [-\pi, \pi]. \tag{3.5.2}$$

According to Hindmarsh et al. (1984), the von Neumann stability condition holds if the inequality holds for $\omega_x = \omega_y = \omega_z = \pi$ and $\omega_x, \omega_y, \omega_z \to 0$.

Here, we use the approach given by Hindmarsh et al. (1984) to find the stability regions for the explicit numerical method considered to solve the 3D advection-diffusion equation.

3.5.2 Numerical methods

In this sub section, we see how the three finite difference methods are constructed to solve a 3D advection-diffusion equation. For simplicity, we use the following notations:

$$c_x = \frac{\beta_x \Delta t}{\Delta x}, \quad c_y = \frac{\beta_y \Delta t}{\Delta y}, \quad c_z = \frac{\beta_z \Delta t}{\Delta z},$$
$$s_x = \frac{\alpha_x \Delta t}{\Delta x^2}, \quad s_y = \frac{\alpha_y \Delta t}{\Delta y^2}, \quad s_z = \frac{\alpha_z \Delta t}{\Delta z^2}.$$

Implicit Crank-Nicolson method (ICN)

Since the 3D implicit methods are costly, we use fractional splitting methods, by first splitting the 3D problem into three 1D equations as done by Yang et al. (1998).

$$u_t = -\beta_x\, u_x + \alpha_x\, u_{xx} \tag{3.5.3a}$$

$$u_t = -\beta_y\, u_y + \alpha_y\, u_{yy} \tag{3.5.3b}$$

$$u_t = -\beta_z\, u_z + \alpha_z\, u_{zz}. \tag{3.5.3c}$$

3.5 Numerical methods and stability analysis

Each of Eqs. (3.5.3a)-(3.5.3c) are split into two equations. At the first fractional time step, Eq. (3.5.3a) is approximated by first separating into two equations

$$\frac{1}{2}u_t = -\beta_x\, u_x \tag{3.5.4}$$

$$\frac{1}{2}u_t = \alpha_x\, u_{xx}. \tag{3.5.5}$$

from which (3.5.4) is discretized as

$$\frac{1}{2}\frac{u_{i,j,k}^{n+1/3} - u_{i,j,k}^{n}}{\Delta t} = -\beta_x \left(\theta \frac{u_{i+1,j,k}^{n+1/3} - u_{i-1,j,k}^{n+1/3}}{2\Delta x} + (1-\theta)\frac{u_{i+1,j,k}^{n} - u_{i-1,j,k}^{n}}{2\Delta x} \right). \tag{3.5.6}$$

and (3.5.5) as

$$\frac{1}{2}\left(\frac{u_{i,j,k}^{n+1/3} - u_{i,j,k}^{n}}{\Delta t} \right) = \alpha_x \left(\theta \frac{u_{i+1,j,k}^{n+1/3} - 2u_{i,j,k}^{n+1/3} + u_{i-1,j,k}^{n+1/3}}{\Delta x^2} + (1-\theta)\frac{u_{i+1,j,k}^{n} - 2u_{i,j,k}^{n} + u_{i-1,j,k}^{n}}{\Delta x^2} \right). \tag{3.5.7}$$

Combining Eqs. (3.5.6) and (3.5.7), we get

$$A_x\, u_{i-1,j,k}^{n+1/3} + B_x\, u_{i,j,k}^{n+1/3} + C_x\, u_{i+1,j,k}^{n+1/3} = D_{i,j,k}^{n}, \tag{3.5.8}$$

where

$$A_x = -\theta\, c_x - 2\theta\, s_x, \quad B_x = 2 + 4\theta\, s_x, \quad C_x = \theta\, c_x - 2\theta\, s_x,$$

$$D_{i,j,k}^{n} = (1-\theta)[(c_x + 2s_x)u_{i-1,j,k}^{n} + (-c_x + 2s_x)u_{i+1,j,k}^{n}] + [2 - 4(1-\theta)s_x]u_{i,j,k}^{n}.$$

In a similar way, we obtain

$$A_y\, u_{i,j-1,k}^{n+2/3} + B_y\, u_{i,j,k}^{n+2/3} + C_y\, u_{i,j+1,k}^{n+2/3} = D_{i,j,k}^{n+1/3}, \tag{3.5.9}$$

and

$$A_z\, u_{i,j,k-1}^{n+1} + B_z\, u_{i,j,k}^{n+1} + C_z\, u_{i,j,k+1}^{n+1} = D_{i,j,k}^{n+2/3}, \tag{3.5.10}$$

from the second and third fractional time steps, respectively, where

$$A_y = -\theta\, c_y - 2\theta\, s_y, \quad B_y = 2 + 4\theta\, s_y, \quad C_y = \theta\, c_x - 2\theta\, s_x$$

$$D_{i,j,k}^{n+1/3} = (1-\theta)\left[(c_y + 2s_y)u_{i,j-1,k}^{n+1/3} + (-c_y + 2s_y)u_{i,j+1,k}^{n+1/3}\right] + [2 - 4(1-\theta)s_y]u_{i,j,k}^{n+1/3}$$

3.5 Numerical methods and stability analysis

$$A_z = -\theta\, c_z - 2\theta\, s_z, \quad B_z = 2 + 4\theta\, s_z, \quad C_z = \theta\, c_z - 2\theta\, s_z$$
$$D_{i,j,k}^{n+2/3} = (1-\theta)\left[(c_z + 2s_z)u_{i,j,k-1}^{n+2/3} + (-c_x + 2s_z)u_{i,j,k+1}^{n+2/3}\right] + [2 - 4(1-\theta)s_z]\, u_{i,j,k}^{n+2/3}.$$

The modified equations of (3.5.8), (3.5.9), and (3.5.10), are

$$\begin{aligned} u_t + \beta_x u_x - \alpha_x u_{xx} = &- \frac{\Delta t}{2} u_{tt} - \frac{\Delta t^2}{6} u_{ttt} - \frac{\Delta t^3}{24} u_{tttt} - \frac{\Delta t^4}{120} u_{ttttt} - \frac{\beta_x \Delta x^2}{6} u_{xxx} + \Delta t\, \theta\, \beta_x u_{xt} \\ &- \frac{\beta_x \Delta x^4}{120} u_{xxxxx} + \frac{\alpha_x \Delta x^2}{12} u_{xxxx} + \frac{\alpha_x \Delta x^4}{360} u_{xxxxxx} - \frac{\Delta t^2\, \theta\, \beta_x}{2} u_{xtt} \\ &- \frac{\Delta t^3\, \theta\, \beta_x}{6} u_{xttt} - \frac{\Delta t\, \theta\, \beta_x \Delta x^2}{6} u_{xxxt} - \frac{\Delta t^2\, \theta\, \beta_x \Delta x^2}{12} u_{xxxtt} + \Delta t\, \theta\, \alpha_x u_{xxt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_x}{2} u_{xxtt} + \frac{\Delta t^3\, \theta\, \alpha_x}{6} u_{xxttt} + \frac{\Delta t\, \theta\, \alpha_x \Delta x^2}{12} u_{xxxxt} - \frac{\Delta t^4\, \theta\, \alpha_x}{12\Delta x^2} u_{tttt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_x \Delta x^2}{24} u_{xxxxtt} - \frac{\Delta t^5\, \theta\, \alpha_x}{60\Delta x^2} u_{ttttt} + \cdots, \end{aligned} \tag{3.5.11}$$

$$\begin{aligned} u_t + \beta_y u_y - \alpha_y u_{yy} = &- \frac{\Delta t}{2} u_{tt} - \frac{\Delta t^2}{6} u_{ttt} - \frac{\Delta t^3}{24} u_{tttt} - \frac{\Delta t^4}{120} u_{ttttt} - \frac{\beta_y \Delta y^2}{6} u_{yyy} - \Delta t\, \theta\, \beta_y u_{yt} \\ &- \frac{\beta_y \Delta y^4}{120} u_{yyyyy} + \frac{\alpha_y \Delta y^2}{12} u_{yyyy} + \frac{\alpha_y \Delta y^4}{360} u_{yyyyyy} - \frac{\Delta t^2\, \theta\, \beta_y}{2} u_{ytt} \\ &- \frac{\Delta t^3\, \theta\, \beta_y}{6} u_{yttt} - \frac{\Delta t\, \theta\, \beta_y \Delta y^2}{6} u_{yyyt} - \frac{\Delta t^2\, \theta\, \beta_y \Delta y^2}{12} u_{yyytt} + \Delta t\, \theta\, \alpha_y u_{yyt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_y}{2} u_{yytt} + \frac{\Delta t^3\, \theta\, \alpha_y}{6} u_{yyttt} + \frac{\Delta t\, \theta\, \alpha_y \Delta y^2}{12} u_{yyyyt} - \frac{\Delta t^4\, \theta\, \alpha_y}{12\Delta y^2} u_{tttt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_y \Delta y^2}{24} u_{yyyytt} - \frac{\Delta t^5\, \theta\, \alpha_y}{60\Delta y^2} u_{ttttt} + \cdots, \end{aligned} \tag{3.5.12}$$

and

$$\begin{aligned} u_t + \beta_z u_z - \alpha_z u_{zz} = &- \frac{\Delta t}{2} u_{tt} - \frac{\Delta t^2}{6} u_{ttt} - \frac{\Delta t^3}{24} u_{tttt} - \frac{\Delta t^4}{120} u_{ttttt} - \frac{\beta_z \Delta z^2}{6} u_{zzz} - \Delta t\, \theta\, \beta_z u_{zt} \\ &- \frac{\beta_z \Delta z^4}{120} u_{zzzzz} - \frac{\alpha_z \Delta z^2}{12} u_{zzzz} + \frac{\alpha_z \Delta z^4}{360} u_{zzzzzz} - \frac{\Delta t^2\, \theta\, \beta_z}{2} u_{ztt} \\ &- \frac{\Delta t^3\, \theta\, \beta_z}{6} u_{zttt} - \frac{\Delta t\, \theta\, \beta_z \Delta z^2}{6} u_{zzzt} - \frac{\Delta t^2\, \theta\, \beta_z \Delta z^2}{12} u_{zzztt} + \Delta t\, \theta\, \alpha_z u_{zzt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_z}{2} u_{zztt} + \frac{\Delta t^3\, \theta\, \alpha_z}{6} u_{zzttt} + \frac{\Delta t\, \theta\, \alpha_z \Delta z^2}{12} u_{zzzzt} - \frac{\Delta t^4\, \theta\, \alpha_z}{12\Delta z^2} u_{tttt} \\ &+ \frac{\Delta t^2\, \theta\, \alpha_z \Delta z^2}{24} u_{zzzztt} - \frac{\Delta t^5\, \theta\, \alpha_z}{60\Delta z^2} u_{ttttt} + \cdots, \end{aligned} \tag{3.5.13}$$

respectively.

When the time and mixed derivatives in Eq. (3.5.11) are converted to space derivatives, we obtain

$$
\begin{aligned}
u_t + \beta_x\, u_x + \left[-\alpha_x + \Delta t\, \beta_x^2 \left(\frac{1}{2} - \theta\right)\right] u_{xx} = {} & \left[\Delta t\, \beta_x\, \alpha_x\, (1-2\theta) - \frac{\beta_x\, \Delta x^2}{6} + \frac{\Delta t^2\, \beta_x^3}{2}\left(\frac{1}{3} - \theta\right)\right] u_{xxx} \\
& + \left[\frac{\Delta t^3\, \beta_x^4}{6}\left(\theta - \frac{1}{4}\right) + \frac{\Delta t\, \theta\, \beta_x^2\, \Delta x^2}{6} + \Delta t\, \alpha_x^2 \left(\frac{1}{2} - \theta\right) + \frac{\Delta t^2\ \beta_x^2\, \alpha_x}{2}\,(3\theta - 1)\right. \\
& \left. + \frac{\alpha_x\, \Delta x^2}{12} - \frac{\Delta t^4\, \theta\, \alpha_x\, \beta_x^4}{12\Delta x^2}\right] u_{xxxx} + \left[\frac{\Delta t^4\, \beta_x^5}{120} + \frac{\Delta t^3\, \beta_x^3\, \alpha_x}{3}\left(\frac{1}{2} - 2\theta\right) + \frac{\Delta t^2\, \beta_x\, \alpha_x^2}{2}\,(3\theta - 1)\right. \\
& \left. + \frac{\Delta t^4\, \theta\, \alpha_x^2\, \beta_x^3}{3\Delta x^2} + \frac{\Delta t^5\, \theta\, \alpha_x\, \beta_x^5}{60\Delta x^2} - \frac{\Delta t^2\, \theta\, \beta_x^3\, \Delta x^2}{12} - \frac{\beta_x\, \Delta x^4}{120} - \frac{\Delta t\, \theta\, \beta_x\, \Delta x^2\, \alpha_x}{4}\right] u_{xxxxx} \\
& + \left[\frac{\Delta t^2\, \alpha_x^3}{2}\left(\theta - \frac{1}{3}\right) + \frac{\alpha_x\, \Delta x^4}{360} - \frac{\Delta t^5\, \theta\, \alpha_x^2\, \beta_x^4}{12\Delta x^2} + \frac{5\, \Delta t^2\, \theta\, \beta_x^2\, \Delta x^2\, \alpha_x}{24} - \frac{\Delta t^4\, \alpha_x\, \beta_x^4}{24}\right. \\
& \left. + \Delta t^3\, \beta_x^2\, \alpha_x^2 \left(\theta - \frac{1}{4}\right) + \frac{\Delta t\, \theta\, \alpha_x^2\, \Delta x^2}{12} - \frac{\Delta t^4\, \theta\, \alpha_x^3\, \beta_x^2}{2\Delta x^2}\right] u_{xxxxxx} + \cdots .
\end{aligned}
\tag{3.5.14}
$$

The modified equations in the directions of y and z can be obtained in the same way. Hence the method is a second order accurate in space and first order accurate in time for $\theta \neq \frac{1}{2}$. When $\theta = \frac{1}{2}$, we obtain the Crank-Nicolson method and from (3.5.14), its modified equation is

$$
\begin{aligned}
u_t + \beta_x\, u_x - \alpha_x\, u_{xx} = {} & -\left[\frac{\beta_x \Delta x^2}{6} + \frac{\Delta t^2 \beta_x^3}{12}\right] u_{xxx} + \left[\frac{\alpha_x\, \Delta x^2}{12} + \frac{\Delta t^3\, \beta_x^4}{24} + \frac{\Delta t\, \Delta x^2\, \beta_x^2}{12}\right. \\
& \left. + \frac{\Delta t^2\, \beta_x^2\, \alpha_x}{4} - \frac{\Delta t^4\, \alpha_x\, \beta_x^4}{24\Delta x^2}\right] u_{xxxx} - \left[\frac{\Delta t^2\, \beta_x\, \alpha_x^2}{4} + \frac{\Delta t^3\, \beta_x^3\, \alpha_x}{6} + \frac{\Delta t^2\, \Delta x^2\, \beta_x^3}{24}\right. \\
& \left. + \frac{\Delta t\, \Delta x^2\, \beta_x\, \alpha_x}{8} + \frac{\beta_x\, \Delta x^4}{120} - \frac{\Delta t^4\, \alpha_x^2\, \beta_x^3}{6\Delta x^2} - \frac{\Delta t^4\, \beta_x^5}{120} - \frac{\Delta t^5\, \alpha_x\, \beta_x^5}{120\Delta x^2}\right] u_{xxxxx} \\
& + \left[-\frac{\Delta t^5\, \alpha_x^2\, \beta_x^4}{24\Delta x^2} - \frac{\Delta t^4\, \alpha_x^3\, \beta_x^2}{4\Delta x^2} + \frac{\Delta t^2\, \alpha_x^3}{12} + \frac{5\, \Delta t^2\, \Delta x^2\, \alpha_x\, \beta_x^2}{48} + \frac{\Delta t\, \Delta x^2\, \alpha_x^2}{24}\right. \\
& \left. + \frac{\alpha_x\, \Delta x^4}{360} + \frac{\Delta t^3\, \beta_x^2\, \alpha_x^2}{4} - \frac{\Delta t^4\, \alpha_x\, \beta_x^4}{24}\right] u_{xxxxxx} + \cdots .
\end{aligned}
$$

The Crank-Nicolson scheme is second order accurate in space and time. Amplification factor is important to study dispersion and dissipation properties of numerical methods as well as to obtain stability of explicit methods. The amplification factor of the numerical

method is given by

$$\xi = \frac{(1-\theta)\big[(c_z + 2s_z)\,\exp(-I\omega_z) + (-c_z + 2s_z)\,\exp(I\omega_z)\big] + 2 - 4(1-\theta)s_z}{A_z \exp(-I\omega_z) + B_z + C_z \exp(I\omega_z)}\,\xi_2,$$

where

$$\xi_2 = \frac{(1-\theta)\big[(c_y + 2s_y)\,\exp(-I\omega_y) + (-c_y + 2s_y)\,\exp(I\omega_y)\big] + 2 - 4(1-\theta)s_y}{A_y \exp(-I\omega_y) + B_y + C_y \exp(I\omega_y)}\,\xi_1,$$

and

$$\xi_1 = \frac{(1-\theta)\big[(c_x + 2s_x)\,\exp(-I\omega_x) + (-c_x + 2s_x)\,\exp(I\omega_x)\big] + 2 - 4(1-\theta)s_x}{A_x \exp(-I\omega_x) + B_x + C_x \exp(I\omega_x)}.$$

For any Δx, Δy, $\Delta y > 0$, the method is unconditionally stable.

Implicit chapeau function method (ICF)

In this sub-section, we refer to Yang et al. (1998). We adopt same type of splitting as for the Implicit Crank-Nicolson method. We first integrate (3.5.3a) in the x-direction. Equations (3.5.4) and (3.5.5) are discretized as follows:

$$\frac{1}{12\Delta t}\left[(u^{n+1/3}_{i-1,j,k} - u^{n}_{i-1,j,k}) + 4(u^{n+1/3}_{i,j,k} - u^{n}_{i,j,k}) + (u^{n+1/3}_{i+1,j,k} - u^{n}_{i+1,j,k})\right] = -\beta_x \left(\theta \frac{u^{n+1/3}_{i+1,j,k} - u^{n+1/3}_{i-1,j,k}}{2\Delta x} + (1-\theta)\frac{u^{n}_{i+1,j,k} - u^{n}_{i-1,j,k}}{2\Delta x}\right), \tag{3.5.15}$$

$$\frac{1}{12\Delta t}\left[(u^{n+1/3}_{i-1,j,k} - u^{n}_{i-1,j,k}) + 4(u^{n+1/3}_{i,j,k} - u^{n}_{i,j,k}) + (u^{n+1/3}_{i+1,j,k} - u^{n}_{i+1,j,k})\right] = \alpha_x \left[\theta\left(\frac{u^{n+1/3}_{i+1,j,k} - 2u^{n+1/3}_{i,j,k} + u^{n+1/3}_{i-1,j,k}}{\Delta x^2}\right) + (1-\theta)\left(\frac{u^{n}_{i+1,j,k} - 2u^{n}_{i,j,k} + u^{n}_{i-1,j,k}}{\Delta x^2}\right)\right]. \tag{3.5.16}$$

Combining Eqs. (3.5.15) and (3.5.16), we get

$$A_x\, u^{n+1/3}_{i-1,j,k} + B_x\, u^{n+1/3}_{i,j,k} + C_x\, u^{n+1/3}_{i+1,j,k} = D^{n}_{i,j,k}, \tag{3.5.17}$$

where

$$A_x = 2 - 6\theta\, c_x - 12\theta\, s_x, \quad B_x = 8 + 24\theta\, s_x, \quad C_x = 2 + 6\theta\, c_x - 12\theta\, s_x$$

$$D^n_{i,j,k} = [2+(1-\theta)(6c_x+12s_x)]u^n_{i-1,j,k} + (8-24(1-\theta)s_x)u^n_{i,j,k} + [2+(1-\theta)(-6c_x+12s_x)]u^n_{i+1,j,k}$$

Using same procedures as those described by Eqs. (3.5.15)-(3.5.17), we obtain

$$A_y\, u^{n+2/3}_{i,j-1,k} + B_y\, u^{n+2/3}_{i,j,k} + C_y\, u^{n+2/3}_{i,j+1,k} = D^{n+1/3}_{i,j,k}, \tag{3.5.18}$$

and

$$A_z\, u^{n+1}_{i,j,k-1} + B_z\, u^{n+1}_{i,j,k} + C_z\, u^{n+1}_{i,j,k+1} = D^{n+2/3}_{i,j,k}, \tag{3.5.19}$$

for the integrations in the y-direction and z-direction, respectively, where

$$\begin{aligned}
&A_y = 2-6\theta\, c_y - 12\theta\, s_y, \quad B_y = 8+24\theta\, s_y, \quad C_y = 2+6\theta\, c_y - 12\theta\, s_y,\\
&D^{n+1/3}_{i,j,k} = [2+(1-\theta)(6c_y+12s_y)]u^{n+1/3}_{i,j-1,k} + (8-24(1-\theta)s_y)u^{n+1/3}_{i,j,k}\\
&\qquad + [2+(1-\theta)(-6c_y+12s_y)]u^{n+1/3}_{i,j+1,k},\\
&A_z = 2-6\theta\, c_z - 12\theta\, s_z, \quad B_z = 8+24\theta\, s_z, \quad C_z = 2+6\theta\, c_z - 12\theta\, s_z,\\
&D^{n+2/3}_{i,j,k} = [2+(1-\theta)(6c_z+12s_z)]u^{n+2/3}_{i,j,k-1} + (8-24(1-\theta)s_z)u^{n+2/3}_{i,j,k}\\
&\qquad + [2+(1-\theta)(-6c_z+12s_z)]u^{n+2/3}_{i,j,k+1}.
\end{aligned}$$

Making use of Taylor's expansion, the modified equations of (3.5.17), (3.5.18) and (3.5.19) are

$$\begin{aligned}
u_t + \beta_x\, u_x - \alpha_x\, u_{xx} = &-\frac{\Delta x^4}{72}\, u_{xxxxt} - \frac{\Delta x^2}{6}\, u_{xxt} - \frac{\Delta t^2\,\Delta x^2}{36}\, u_{xxttt} - \frac{\Delta t \Delta x^2}{12}\, u_{xxtt} - \Delta t\,\theta\,\beta_x\, u_{xt}\\
&- \frac{\Delta t\,\Delta x^4}{144}\, u_{xxxxtt} - \frac{\Delta t^3\,\theta\,\beta_x\,\Delta x^2}{36}\, u_{xxxttt} - \frac{\Delta t^4\,\theta\,\alpha_x}{12\Delta x^2}\, u_{tttt} - \frac{\Delta t^5\,\theta\,\alpha_x}{60\Delta x^2}\, u_{ttttt}\\
&- \frac{\Delta t^2\,\theta\,\beta_x\,\Delta x^2}{12}\, u_{xxxtt} - \frac{\Delta t\,\theta\,\beta_x\,\Delta x^2}{6}\, u_{xxxt} + \frac{\Delta t^2\,\theta\,\alpha_x\,\Delta x^2}{24}\, u_{xxxxtt}\\
&+ \frac{\Delta t\,\theta\,\alpha_x\,\Delta x^2}{12}\, u_{xxxxt} + \frac{\Delta t^2\,\theta\,\alpha_x}{2}\, uxxtt + \Delta t\theta\,\alpha_x\, u_{xxt} - \frac{\beta_x\,\Delta x^2}{6}\, u_{xxx}\\
&+ \frac{\alpha_x\,\Delta x^4}{360}\, u_{xxxxxx} - \frac{\beta_x\,\Delta x^4}{120}\, u_{xxxxx} + \frac{\alpha_x\,\Delta x^2}{12}\, u_{xxxx} - \frac{\Delta t^3\,\theta\,\beta_x}{6}\, u_{xttt}\\
&- \frac{\Delta t^2\,\theta\,\beta_x}{2}\, u_{xtt} + \frac{\Delta t^3\,\theta\,\alpha_x}{6}\, u_{xxttt} - \frac{\Delta t}{2}\, u_{tt} - \frac{\Delta t^2}{6}\, u_{ttt} - \frac{\Delta t^3}{36}\, u_{tttt}\\
&- \frac{\Delta t^4}{180}\, u_{ttttt} + \cdots,
\end{aligned} \tag{3.5.20}$$

$$
\begin{aligned}
u_t + \beta_y\, u_y - \alpha_y\, u_{yy} = &-\frac{\Delta y^4}{72}\, u_{yyyyt} - \frac{\Delta y^2}{6}\, u_{yyt} - \frac{\Delta t^2\, \Delta y^2}{36}\, u_{yyttt} - \frac{\Delta t \Delta y^2}{12}\, u_{yytt} - \Delta t\, \theta\, \beta_y\, u_{yt} \\
&- \frac{\Delta t\, \Delta y^4}{144}\, u_{yyyytt} - \frac{\Delta t^3\, \theta\, \beta_y\, \Delta y^2}{36}\, u_{yyyttt} - \frac{\Delta t^4\, \theta\, \alpha_y}{12 \Delta y^2}\, u_{tttt} - \frac{\Delta t^5\, \theta\, \alpha_y}{60 \Delta y^2}\, u_{ttttt} \\
&- \frac{\Delta t^2\, \theta\, \beta_y\, \Delta y^2}{12}\, u_{yyytt} - \frac{\Delta t\, \theta\, \beta_y\, \Delta y^2}{6}\, u_{yyyt} + \frac{\Delta t^2\, \theta\, \alpha_y\, \Delta y^2}{24}\, u_{yyyytt} \\
&+ \frac{\Delta t\, \theta\, \alpha_y\, \Delta y^2}{12}\, u_{yyyyt} + \frac{\Delta t^2\, \theta\, \alpha_y}{2}\, uyytt + \Delta t \theta\, \alpha_y\, u_{yyt} - \frac{\beta_y\, \Delta y^2}{6}\, u_{yyy} \\
&+ \frac{\alpha_y\, \Delta y^4}{360}\, u_{yyyyyy} - \frac{\beta_y\, \Delta y^4}{120}\, u_{yyyyy} + \frac{\alpha_y\, \Delta y^2}{12}\, u_{yyyy} - \frac{\Delta t^3\, \theta\, \beta_y}{6}\, u_{yttt} \\
&- \frac{\Delta t^2\, \theta\, \beta_y}{2}\, u_{ytt} + \frac{\Delta t^3\, \theta\, \alpha_y}{6}\, u_{yyttt} - \frac{\Delta t}{2}\, u_{tt} - \frac{\Delta t^2}{6}\, u_{ttt} - \frac{\Delta t^3}{36}\, u_{tttt} \\
&- \frac{\Delta t^4}{180}\, u_{ttttt} + \cdots, \qquad (3.5.21)
\end{aligned}
$$

and

$$
\begin{aligned}
u_t + \beta_z\, u_z - \alpha_z\, u_{zz} = &-\frac{\Delta z^4}{72}\, u_{zzzzt} - \frac{\Delta z^2}{6}\, u_{zzt} - \frac{\Delta t^2\, \Delta z^2}{36}\, u_{zzttt} - \frac{\Delta t \Delta z^2}{12}\, u_{zztt} - \Delta t\, \theta\, \beta_z\, u_{zt} \\
&- \frac{\Delta t\, \Delta z^4}{144}\, u_{zzzztt} - \frac{\Delta t^3\, \theta\, \beta_z\, \Delta z^2}{36}\, u_{zzzttt} - \frac{\Delta t^4\, \theta\, \alpha_z}{12 \Delta z^2}\, u_{tttt} - \frac{\Delta t^5\, \theta\, \alpha_z}{60 \Delta z^2}\, u_{ttttt} \\
&- \frac{\Delta t^2\, \theta\, \beta_z\, \Delta z^2}{12}\, u_{zzztt} - \frac{\Delta t\, \theta\, \beta_z\, \Delta z^2}{6}\, u_{zzzt} + \frac{\Delta t^2\, \theta\, \alpha_z\, \Delta z^2}{24}\, u_{zzzztt} \\
&+ \frac{\Delta t\, \theta\, \alpha_z\, \Delta z^2}{12}\, u_{zzzzt} + \frac{\Delta t^2\, \theta\, \alpha_z}{2}\, uzztt + \Delta t \theta\, \alpha_z\, u_{zzt} - \frac{\beta_z\, \Delta z^2}{6}\, u_{zzz} \\
&+ \frac{\alpha_z\, \Delta z^4}{360}\, u_{zzzzzz} - \frac{\beta_z\, \Delta z^4}{120}\, u_{zzzzz} + \frac{\alpha_z\, \Delta z^2}{12}\, u_{zzzz} - \frac{\Delta t^3\, \theta\, \beta_z}{6}\, u_{zttt} \\
&- \frac{\Delta t^2\, \theta\, \beta_z}{2}\, u_{ztt} + \frac{\Delta t^3\, \theta\, \alpha_z}{6}\, u_{zzttt} - \frac{\Delta t}{2}\, u_{tt} - \frac{\Delta t^2}{6}\, u_{ttt} - \frac{\Delta t^3}{36}\, u_{tttt} \\
&- \frac{\Delta t^4}{180}\, u_{ttttt} + \cdots, \qquad (3.5.22)
\end{aligned}
$$

respectively. When the time derivatives in (3.5.20) are converted to space derivatives using (3.5.17), we obtain

$$
\begin{aligned}
u_t + \beta_x\, u_x + \left[-\alpha_x + \Delta t \beta_x^2 \left(\frac{1}{2} - \theta\right)\right] u_{xx} = &\left[\Delta t\, \alpha_x\, \beta_x (1 - 2\theta) + \frac{\Delta t^2\, \beta_x^3}{2} \left(\frac{1}{3} - \theta\right)\right] u_{xxx} \\
&+ \left[\frac{\Delta t^2\, \beta_x^2\, \alpha_x}{2}(3\theta - 1) - \frac{\alpha_x\, \Delta x^2}{12} + \frac{\Delta t\, \theta\, \beta_x^2\, \Delta x^2}{6} + \Delta t \alpha_x^2 \left(\theta - \frac{1}{2}\right) - \frac{\Delta t^4 \theta\, \alpha_x\, \beta_x^4}{12 \Delta x^2}\right. \\
&\left. + \frac{\Delta t^3\, \theta\, \beta_x^4}{6} \left(\theta - \frac{1}{6}\right)\right] u_{xxxx} + \left[\frac{\beta_x\, \Delta x^4}{180} + \frac{\Delta t^2\, \beta_x\, \alpha_x^2}{2}\, (1 - 3\theta) + \frac{\Delta t^3\, \beta_x^3\, \alpha_x}{3} \left(\frac{1}{3} - 2\theta\right)\right.
\end{aligned}
$$

$$+\frac{\Delta t^2\,\beta_x^3\,\Delta x^2}{12}\left(\frac{1}{3}-\theta\right)+\frac{\Delta t\Delta x^2\,\beta_x\,\alpha_x}{2}\left(\frac{1}{3}-\frac{\theta}{2}\right)+\frac{\Delta t^5\,\theta\,\alpha_x\,\beta_x^5}{60\Delta x^2}+\frac{\Delta t^4\,\theta\,\alpha_x^2\,\beta_x^3}{3\Delta x^2}$$
$$+\frac{\Delta t^4\,\beta_x^5}{180}\Bigg]u_{xxxxx}+\Bigg[\Delta t^3\,\beta_x^2\,\alpha_x^2\left(\theta-\frac{1}{6}\right)+\frac{\Delta t\,\alpha_x^2\,\Delta x^2}{12}(\theta-1)-\frac{\Delta t^4\,\theta\,\alpha_x^3\,\beta_x^2}{2\Delta x^2}$$
$$-\frac{\alpha_x\,\Delta x^4}{90}-\frac{\Delta t^2\,\Delta x^2\,\alpha_x\,\beta_x^2}{24}(2-5\theta)+\frac{\Delta t^3\,\theta\,\beta_x^4\,\Delta x^2}{36}-\frac{\Delta t^5\,\theta\,\alpha_x^2\,\beta_x^4}{12\Delta x^2}-\frac{\Delta t\Delta x^2\,\beta_x^2}{12}$$
$$+\frac{\Delta t^2\,\alpha_x^3}{2}\left(\theta-\frac{1}{3}\right)-\frac{\Delta t^4\,\alpha_x\,\beta_x^4}{36}-\frac{\Delta t\,\Delta x^4\,\beta_x^2}{144}\Bigg]u_{xxxxxx}+\cdots=0. \qquad (3.5.23)$$

The modified equations in the directions of y and z can be obtained similarly. Hence the method is second order accurate in space and is first order accurate in time when $\theta \neq 1/2$. When $\theta=\frac{1}{2}$, we get the ICF method and from (3.5.23), its modified equation is given by

$$u_t+\beta_x\,u_x-\alpha_x\,u_{xx}=-\frac{\Delta t^2\,\beta_x^3}{12}\,u_{xxx}+\Bigg[\frac{\Delta t^2\,\beta_x^2\,\alpha_x}{4}-\frac{\alpha_x\,\Delta x^2}{12}+\frac{\Delta t\,\beta_x^2\,\Delta x^2}{12}-\frac{\Delta t^4\,\alpha_x\,\beta_x^4}{24\Delta x^2}$$
$$+\frac{\Delta t^3\,\theta\,\beta_x^4}{18}\Bigg]u_{xxxx}+\Bigg[\frac{\beta_x\,\Delta x^4}{180}-\frac{\Delta t^2\,\beta_x\,\alpha_x^2}{4}-\frac{2\Delta t^3\,\beta_x^3\,\alpha_x}{9}-\frac{\Delta t^2\,\beta_x^3\,\Delta x^2}{72}$$
$$+\frac{\Delta t\Delta x^2\,\beta_x\,\alpha_x}{24}+\frac{\Delta t^5\,\alpha_x\,\beta_x^5}{120\Delta x^2}+\frac{\Delta t^4\,\alpha_x^2\,\beta_x^3}{6\Delta x^2}+\frac{\Delta t^4\,\beta_x^5}{180}\Bigg]u_{xxxxx}+\Bigg[\frac{\Delta t^3\,\beta_x^2\,\alpha_x^2}{3}$$
$$-\frac{\alpha_x\,\Delta x^4}{90}+\frac{\Delta t^2\,\Delta x^2\,\alpha_x\,\beta_x^2}{48}+\frac{\Delta t^3\,\beta_x^4\,\Delta x^2}{72}-\frac{\Delta t^5\,\alpha_x^2\,\beta_x^4}{24\Delta x^2}-\frac{\Delta t^4\,\alpha_x^3\,\beta_x^2}{4\Delta x^2}+\frac{\Delta t^2\,\alpha_x^3}{12}$$
$$-\frac{\Delta t\Delta x^2\,\beta_x^2}{12}-\frac{\Delta t^4\,\alpha_x\,\beta_x^4}{36}-\frac{\Delta t\,\Delta x^4\,\beta_x^2}{144}\Bigg]u_{xxxxxx}+\cdots. \qquad (3.5.24)$$

The ICF scheme is second order accurate in space and time.

The amplification factor of the numerical method is given by

$$\xi=\frac{A_{1,z}\,\exp(-I\omega_z)\,\xi_2+A_{2,z}\,\xi_2+A_{3,z}\,\exp(I\omega_z)\,\xi_2}{A_z\,\exp(-I\omega_z)+B_z+C_z\,\exp(I\omega_z)},$$

where

$$\xi_2=\frac{A_{1,y}\,\exp(-I\omega_y)\,\xi_1+A_{2,y}\,\xi_1+A_{3,y}\,\exp(I\omega_y)\,\xi_1}{A_y\,\exp(-I\omega_y)+B_y+C_y\,\exp(I\omega_y)},$$

$$\xi_1=\frac{A_{1,x}\,\exp(-I\omega_z)+A_{2,x}+A_{3,x}\,\exp(I\omega_z)}{A_x\,\exp(-I\omega_x)+B_x+C_x\,\exp(I\omega_x)}.$$

and

$$A_{1,z}=2+(1-\theta)(6c_z+12s_z),\quad A_{2,z}=8-24(1-\theta)s_z,\quad A_{3,z}=2+(1-\theta)(-6c_z+12s_z),$$

$$A_{1,y} = 2 + (1-\theta)(6c_y + 12s_y), \quad A_{2,y} = 8 - 24(1-\theta)s_y, \quad A_{3,y} = 2 + (1-\theta)(-6c_y + 12s_y),$$
$$A_{1,x} = 2 + (1-\theta)(6c_x + 12s_x), \quad A_{2,x} = 8 - 24(1-\theta)s_x, \quad A_{3,x} = 2 + (1-\theta)(-6c_x + 12s_x).$$

The scheme is implicit and hence unconditionally stable.

Fourth order finite difference method (FOM)

Dehghan (2005) constructed a fourth order finite difference scheme for 1D advection-diffusion equation and then extended the scheme to discretize 2D advection-diffusion equation (Dehghan, 2007) using time-splitting procedures. In this work, we extend their work to obtain a scheme to discretize 3D advection-diffusion equation using fractional splitting techniques. The following approximations are used to solve (3.5.3a) at the first fractional step:

$$\left.\frac{\partial u}{\partial t}\right|_{i,j,k}^{n} \simeq \frac{u_{i,j,k}^{n+1/3} - u_{i,j,k}^{n}}{\Delta t}, \tag{3.5.25}$$

$$\left.\frac{\partial u}{\partial x}\right|_{i,j,k}^{n+1/3} \simeq \left(\frac{12s_x + 2c_x^2 - 3c_x - 2}{12}\right)\left(\frac{u_{i+2,j,k}^{n} - u_{i,j,k}^{n}}{2\Delta x}\right) + \left(\frac{12s_x + 2c_x^2 + 3c_x - 2}{12}\right)\left(\frac{u_{i,j,k}^{n} - u_{i-2,j,k}^{n}}{2\Delta x}\right)$$
$$- \left(\frac{c_x^2 + 6s_x - 4}{3}\right)\left(\frac{u_{i+1,j,k}^{n} - u_{i-1,j,k}^{n}}{2\Delta x}\right), \tag{3.5.26}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{i,j,k}^{n+1/3} \simeq \left(\frac{-c_x^4 + 4c_x^2 - 12s_x^2 - 12s_x c_x^2 + 8s_x}{6s_x}\right)\left(\frac{u_{i+1,j,k}^{n} - 2u_{i,j,k}^{n} + u_{i-1,j,k}^{n}}{(\Delta x)^2}\right)$$
$$+ \left(\frac{c_x^4 - 4c_x^2 + 12s_x^2 + 12s_x c_x^2 - 2s_x}{6s_x}\right)\left(\frac{u_{i+2,j,k}^{n} - 2u_{i,j,k}^{n} + u_{i-2,j,k}^{n}}{4(\Delta x)^2}\right). \tag{3.5.27}$$

On substitution of Eqs. (3.5.25)-(3.5.27) into Eq. (3.5.3a) and after some algebraic manipulation, we obtain

$$u_{i,j,k}^{n+1/3} = A_x\, u_{i-2,j,k}^{n} + B_x\, u_{i-1,j,k}^{n} + C_x\, u_{i,j,k}^{n} + D_x\, u_{i+1,j,k}^{n} + E_x\, u_{i+2,j,k}^{n}, \tag{3.5.28}$$

where

$$A_x = \frac{1}{24}\left(12s_x(s_x + c_x^2) + 2s_x(6c_x - 1) + c_x(c_x - 1)(c_x + 1)(c_x + 2)\right),$$
$$B_x = -\frac{1}{6}\left(12s_x(s_x + c_x^2) + 2s_x(3c_x - 4) + c_x(c_x - 2)(c_x + 1)(c_x + 2)\right),$$
$$C_x = \frac{1}{4}\left(12s_x(s_x + c_x^2) - 10s_x + (c_x - 1)(c_x - 2)(c_x + 1)(c_x + 2)\right),$$

$$D_x = -\frac{1}{6}\left(12s_x(s_x+c_x^2) - 2s_x(3c_x+4) + c_x(c_x-2)(c_x-1)(c_x+2)\right),$$
$$E_x = \frac{1}{24}\left(12s_x(s_x+c_x^2) - 2s_x(6c_x+1) + c_x(c_x-1)(c_x+1)(c_x-2)\right).$$

We use same procedure for the $y-$ and $z-$ directions at the second and third fractional time steps and obtain

$$u_{i,j,k}^{n+2/3} = A_y\, u_{i,j-2,k}^{n+1/3} + B_y\, u_{i,j-1,k}^{n+1/3} + C_y\, u_{i,j,k}^{n+1/3} + D_y\, u_{i,j+1,k}^{n+1/3} + E_y\, u_{i,j,k+2}^{n+1/3}, \tag{3.5.29}$$

and

$$u_{i,j,k}^{n+1} = A_z\, u_{i,j,k-2}^{n+2/3} + B_z\, u_{i,j,k-1}^{n+2/3} + C_z\, u_{i,j,k}^{n+2/3} + D_z\, u_{i,j,k+1}^{n+2/3} + E_z\, u_{i,j,k+2}^{n+2/3}, \tag{3.5.30}$$

where

$$A_y = \frac{1}{24}\left(12s_y(s_y+c_y^2) + 2s_y(6c_y-1) + c_y(c_y-1)(c_y+1)(c_y+2)\right),$$
$$B_y = -\frac{1}{6}\left(12s_y(s_y+c_y^2) + 2s_y(3c_y-4) + c_y(c_y-2)(c_y+1)(c_y+2)\right),$$
$$C_y = \frac{1}{4}\left(12s_y(s_y+c_y^2) - 10s_y + (c_y-1)(c_y-2)(c_y+1)(c_y+2)\right),$$
$$D_y = -\frac{1}{6}\left(12s_y(s_y+c_y^2) - 2s_y(3c_y+4) + c_y(c_y-2)(c_y-1)(c_y+2)\right),$$
$$E_y = \frac{1}{24}\left(12s_y(s_y+c_y^2) - 2s_y(6c_y+1) + c_y(c_y-1)(c_y+1)(c_y-2)\right),$$

and

$$A_z = \frac{1}{24}\left(12s_z(s_z+c_z^2) + 2s_z(6c_z-1) + c_z(c_z-1)(c_z+1)(c_z+2)\right),$$
$$B_z = -\frac{1}{6}\left(12s_z(s_z+c_z^2) + 2s_z(3c_z-4) + c_z(c_z-2)(c_z+1)(c_z+2)\right),$$
$$C_z = \frac{1}{4}\left(12s_z(s_z+c_z^2) - 10s_z + (c_z-1)(c_z-2)(c_z+1)(c_z+2)\right),$$
$$D_z = -\frac{1}{6}\left(12s_z(s_z+c_z^2) - 2s_z(3c_z+4) + c_z(c_z-2)(c_z-1)(c_z+2)\right),$$
$$E_z = \frac{1}{24}\left(12s_z(s_z+c_z^2) - 2s_z(6c_z+1) + c_z(c_z-1)(c_z+1)(c_z-2)\right).$$

The modified equations of Eqs. (3.5.28), (3.5.29) and (3.5.30) are (Appadu et al., 2016)

$$u_t + \beta_x\, u_x - \alpha_x\, u_{xx} = \frac{\Delta x^4}{120}\left(60s_x^2 + 20s_xc_x^2 + c_x^4 - 5c_x^2 + 4 - 30s_x\right)\, u_{xxxxx} + \cdots, \tag{3.5.31}$$

β	α	Stability region for Δt
0.8	0.01	[0,0.073865]
1	1	[0,0.001665]
2	0.01	[0, 0.039467]
0.01	0.01	[0, 0.166528]
0.001	0.01	[0, 0.166665]

Table 3.1: Stability region of FOM when for some values ofβ and α when $\Delta x = \Delta y = \Delta z = 0.05$.

$$u_t + \beta_y\, u_y - \alpha_y\, u_{yy} = \frac{\Delta y^4}{120}\left(60 s_y^2 + 20 s_y c_y^2 + c_y^4 - 5 c_y^2 + 4 - 30 s_y\right)\, u_{yyyyy} + \cdots, \tag{3.5.32}$$

and

$$u_t + \beta_z\, u_z - \alpha_z\, u_{zz} = \frac{\Delta z^4}{120}\left(60 s_z^2 + 20 s_z c_z^2 + c_z^4 - 5 c_z^2 + 4 - 30 s_z\right)\, u_{zzzzz} + \cdots, \tag{3.5.33}$$

respectively, and hence this scheme is fourth order accurate in space and third accurate in time.

Using von Neumann stability analysis, the amplification factor of FOM is given by

$$\xi = (A_z \exp(-2I\omega_z) + B_z \exp(-I\omega_z) + C_z + D_z \exp(I\omega_z) + E_z \exp(2I\omega_z))\xi_2,$$

where

$$\xi_2 = (A_y \exp(-2I\omega_y) + B_y \exp(-I\omega_y) + C_y + D_y \exp(I\omega_y) + E_y \exp(2I\omega_y))\xi_1,$$

and

$$\xi_1 = A_x \exp(-2I\omega_x) + B_x \exp(-I\omega_x) + C_x + D_x \exp(I\omega_x) + E_x \exp(2I\omega_x).$$

Using the approach of Hindmarsh et al. (1984), the region of stability at $h = 0.05$ for some values of advection velocities and diffusivities are given in Table 3.1.

3.6 Numerical results

The numerical results obtained from the three numerical methods at time $T = 0.05$ and 0.2 are shown in Tables 3.2 to 3.8. Figs. 3.1 to 3.3 present the numerical results when $z = 0.5$. For all computations, we consider the spatial step sizes $\Delta x = \Delta y = \Delta z = h = 0.05$ and the temporal step size, $\Delta t = 0.001$, where the advection and diffusion coefficients are the same in all directions, i. e. , $\beta_x = \beta_y = \beta_z = \beta$ and $\alpha_x = \alpha_y = \alpha_z = \alpha$, with β and α to be chosen. Some combinations of β and α are considered namely:

(a) $\beta = 0.8, \alpha = 0.01$.
(b) $\beta = 2, \alpha = 0.01$.
(c) $\beta = 1, \alpha = 1$.
(d) $\beta = 0.01, \alpha = 0.01$.
(e) $\beta = 0.001, \alpha = 0.01$.

Table 3.2 shows the errors at $h = 0.05, \Delta t = 0.001$ and $T = 0.05$ when $\beta = 0.8$ and $\alpha = 0.01$. Based on L_2, L_∞ and dispersion errors, ICF is slightly better than FOM. ICN is the worst performing scheme for this problem. Based on the positive definite character, ICF is best scheme out of the three methods, though it is not completely positive definite. MDR values from ICN and FOM are respectively 1.0129 and 1.0013 while this value from ICF is 0.9870. MDR is affected by mass spreading into areas of the grid that should remain free of mass, by damping of the peak concentration, by negative ripples or by diffusive events in the solution (Kenneth, 1983).

Table 3.3 shows errors when $\beta = 0.8$ and $\alpha = 0.01$ at $h = 0.05$, $\Delta t = 0.001$ and $T = 0.2$. Based on L_2, L_∞ and dispersion errors, MDR values, the most efficient scheme is ICF followed by FOM.

Table 3.4 shows errors when $\beta = 2$ and $\alpha = 0.01$ at $T = 0.05$. ICF performs better followed by FOM, except the MDR for which FOM better. As we progress in time, for instance $T = 0.2$, as shown in Table 3.5 the FOM performs well and the ICF is least efficient.

3.6 Numerical results

Method	L_2- error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5466	1	1	1	1.6640×10^{-32}	0.7419	0	0	0
ICN	0.0032	0.0799	44.5465	0.9932	1	1.0129	-3.9948×10^{-5}	0.7145	5.4840×10^{-8}	8.5946×10^{-6}	8.6494×10^{-6}
ICF	5.8236×10^{-4}	0.0210	44.5466	0.9998	1	0.9870	-1.2857×10^{-10}	0.7209	5.5783×10^{-8}	2.3718×10^{-7}	2.9297×10^{-7}
FOM	8.3415×10^{-4}	0.0226	44.5466	0.9995	1	1.0013	-1.2422×10^{-4}	0.7330	5.8748×10^{-10}	6.0047×10^{-7}	6.0106×10^{-7}

Table 3.2: Errors obtained when $\alpha = 0.01, \beta = 0.8$, $h = 0.05$ and $\Delta t = 0.001$, $T = 0.05$.

Method	L_2- error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5400	1	1	1	1.2227×10^{-32}	0.4072	0	0	0
ICN	0.0052	0.0916	44.3273	0.9658	0.9952	1.0238	-0.0027	0.4059	1.0541×10^{-7}	2.3197×10^{-5}	2.3302×10^{-5}
ICF	7.7228×10^{-4}	0.0163	44.8431	0.9992	1.0068	0.9778	-5.0893×10^{-10}	0.3910	9.5717×10^{-8}	4.2047×10^{-7}	5.1518×10^{-7}
FOM	9.6606×10^{-4}	0.0181	44.5894	0.9988	1.0011	1.0016	1.2227×10^{-32}	0.4110	4.7813×10^{-10}	8.0571×10^{-7}	8.0618×10^{-7}

Table 3.3: Errors obtained when $\alpha = 0.01, \beta = 0.8$, $h = 0.05$ and $\Delta t = 0.001$, $T = 0.2$.

Fig. 3.3c and Table 3.6 show that the ICN method is the worst method for the case $\alpha = 1$ followed by the ICF. Both methods are highly dissipative. For the cases α and β small, all the results obtained from all the three methods are close to the exact solution as shown in Figs. 3.3d and 3.3e and Tables 3.7 and 3.8, where least errors are obtained from FOM followed by ICF.

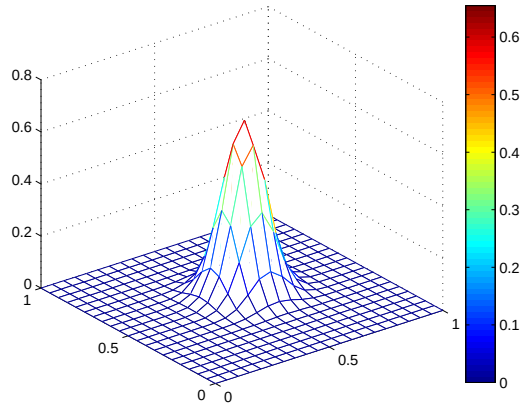
Based on dissipation errors obtained from all the cases considered, the FOM is the best scheme out of the three schemes.

We then obtain the rate of convergence of the three methods. Using Tables 3.9 and 3.10, we observe that the FOM is fourth order in space, ICN and ICF are second order in space.

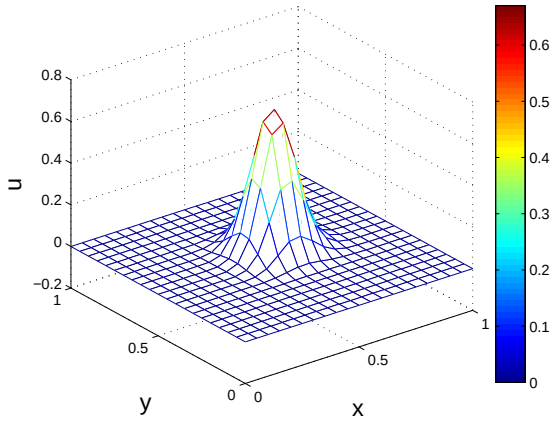
Method	L_2- error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5466	1	1	1	6.2334×10^{-40}	0.7607	0	0	0
ICN	0.0077	0.1712	44.5405	0.9600	0.9999	1.0130	-0.0063	0.6872	5.5188×10^{-8}	5.0758×10^{-5}	5.0813×10^{-5}
ICF	7.7692×10^{-4}	0.0217	44.5467	0.9996	1	0.9871	-1.0839×10^{-5}	0.7390	5.5330×10^{-8}	4.6609×10^{-7}	5.2142×10^{-7}
FOM	0.0020	0.0487	44.5488	0.9972	1	1.0001	-5.8074×10^{-4}	0.7436	5.1787×10^{-12}	3.5561×10^{-6}	3.5561×10^{-6}

Table 3.4: Errors obtained when $\alpha = 0.01, \beta = 2, h = 0.05$ and $\Delta t = 0.001$ at $T = 0.05$.

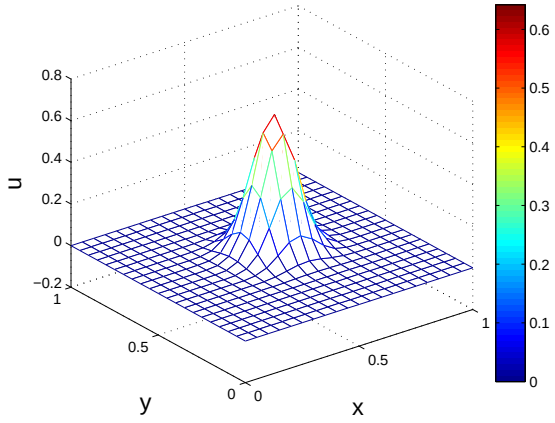
3.6 Numerical results



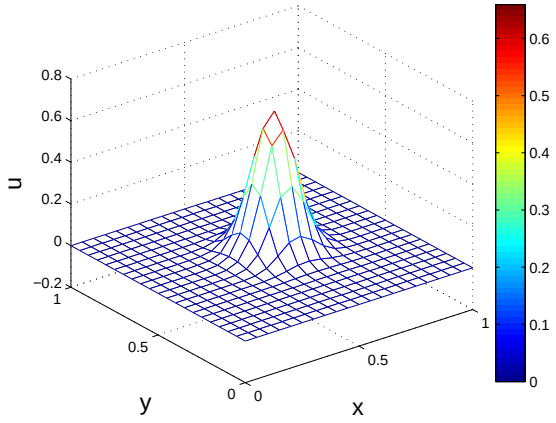
(a) Exact



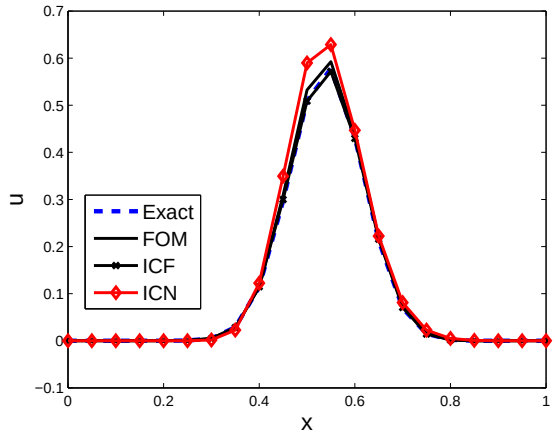
(b) ICN



(c) ICF



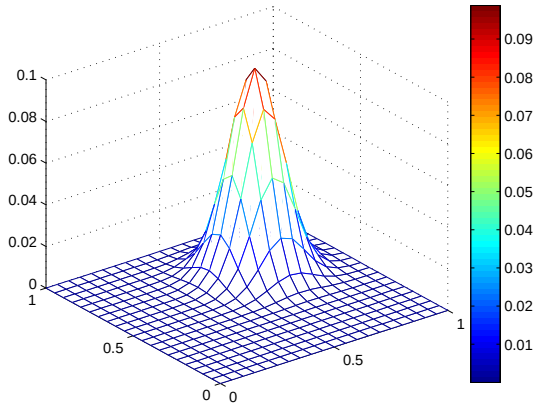
(d) FOM



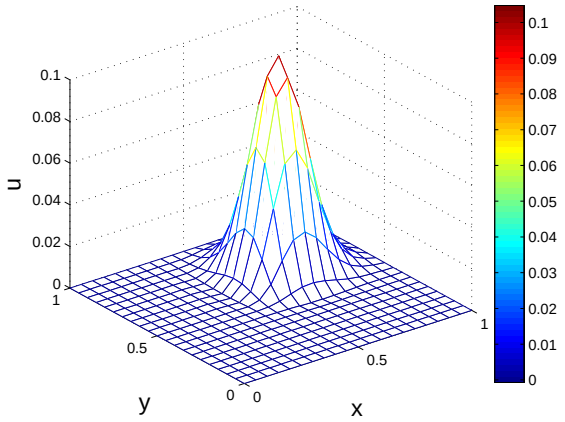
(e) u vs x when $y = 0.5$

Figure 3.1: Numerical results from FOM, ICF and ICN when $\alpha = 0.01, \beta = 0.8$, $h = 0.05$, $\Delta t = 0.001$ and $z = 0.5$ at $T = 0.05$.

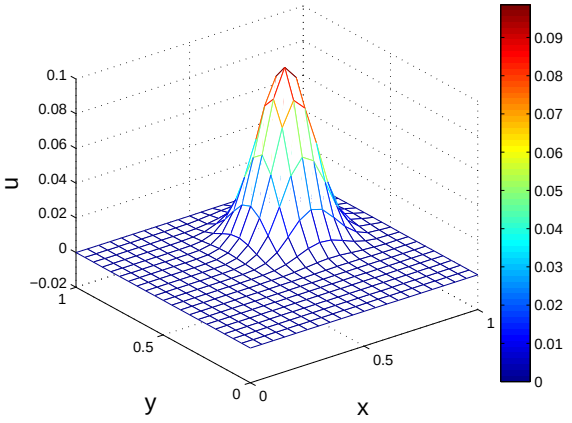
3.6 Numerical results



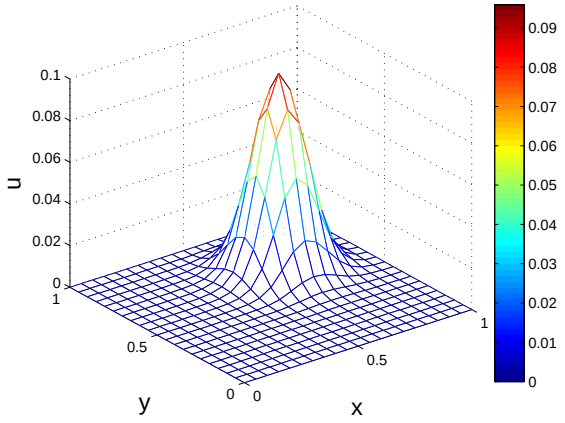
(a) Exact



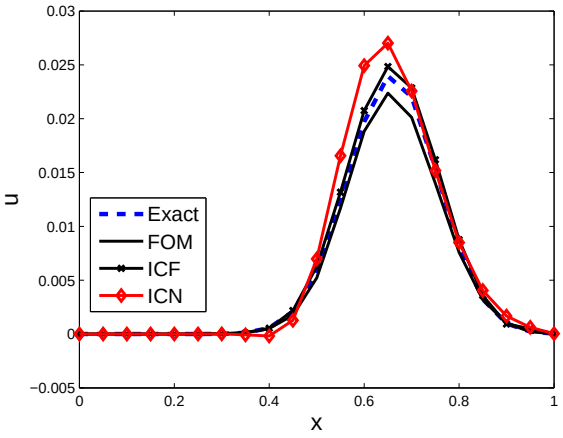
(b) ICN



(c) ICF



(d) FOM



(e) u vs x when $y = 0.5$

Figure 3.2: Numerical profiles from FOM, ICF and ICN when $\alpha = 0.01, \beta = 0.8, h = 0.05, \Delta t = 0.001$ and $z = 0.5$ at $T = 0.2$.

3.6 Numerical results

Method	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	33.4324	1	1	1	9.7126×10^{-60}	0.4141	0	0	0
ICN	0.0105	0.1864	36.2141	0.8488	1.0832	1.2867	-0.0375	0.3809	1.1695×10^{-5}	8.4049×10^{-5}	9.5743×10^{-5}
ICF	0.1760	6.8426	97.5905	-41.2453	2.9190	46.7061	-2.6598	7.1155	0.0220	0.0047	0.0268
FOM	0.0020	0.0422	33.4840	0.9946	1.0015	1.0229	-7.7893×10^{-4}	0.3891	8.5215×10^{-8}	3.3124×10^{-6}	3.3976×10^{-6}

Table 3.5: Errors obtained when $\alpha = 0.01, \beta = 2, h = 0.05$ and $k = 0.001$ at $T = 0.2$.

Method	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	3.1089×10^3	1	1	1	0.1830	0.4141	0	0	0
ICN	0.3201	0.4106	753.0659	-42.3521	0.2422	0.2235	-0.0163	0.3939	0.0733	0.0152	0.0885
ICF	0.0880	0.1585	2.4926×10^3	-2.2723	0.8017	0.6468	0.1665	0.3939	0.0044	0.0022	0.0067
FOM	0.0014	0.0020	3.1194×10^3	0.9991	1.0034	1.0068	0.1830	0.4156	1.3105×10^{-6}	4.5297×10^{-7}	1.7635×10^{-6}

Table 3.6: Errors obtained when $\alpha = 1, \beta = 1, h = 0.05$ and $\Delta t = 0.001$ at $T = 0.2$.

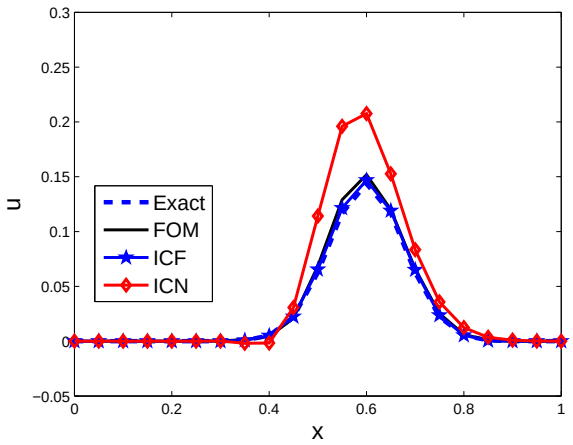
Method	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5466	1	1	1	2.3792×10^{-19}	0.4138	0	0	0
ICN	7.2622×10^{-4}	0.0204	44.5462	0.9993	1.0000	1.0239	2.3792×10^{-19}	0.4342	1.0230×10^{-7}	3.5328×10^{-8}	4.5558×10^{-7}
ICF	6.8262×10^{-4}	0.0179	44.5472	0.9994	1.0000	0.9776	-9.4898×10^{-11}	0.3959	9.2623×10^{-8}	3.0990×10^{-7}	4.0252×10^{-7}
FOM	1.0628×10^{-4}	0.0031	44.5468	1.0000	1.0000	1.0019	-4.9154×10^{-7}	0.4169	6.8583×10^{-10}	9.0715×10^{-9}	9.7573×10^{-9}

Table 3.7: Errors obtained when $\alpha = 0.01, \beta = 0.01, h = 0.05$ and $\Delta t = 0.001$ at $T = 0.2$.

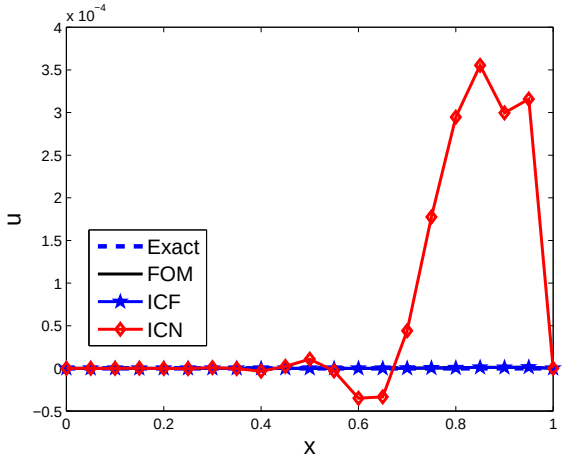
Method	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5466	1	1	1	3.2137×10^{-19}	0.4141	0	0	0
ICN	7.2327×10^{-4}	0.0204	44.5462	0.9993	1.0000	1.0239	3.2137×10^{-19}	0.4344	1.0230×10^{-7}	3.4958×10^{-7}	4.5189×10^{-7}
ICF	6.8262×10^{-4}	0.0179	44.5472	0.9994	1.0000	0.9776	-7.9385×10^{-11}	0.3962	9.2623×10^{-8}	3.0990×10^{-7}	4.0252×10^{-7}
FOM	1.0559×10^{-4}	0.0031	44.5468	1.0000	1.0000	1.0019	-1.7572×10^{-7}	0.4172	6.8586×10^{-10}	8.9457×10^{-9}	9.6315×10^{-9}

Table 3.8: Errors obtained when $\alpha = 0.01, \beta = 0.001, h = 0.05$ and $\Delta t = 0.001$ at $T = 0.2$.

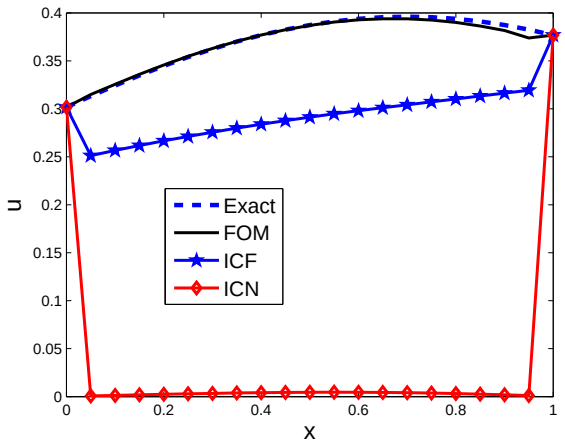
3.6 Numerical results



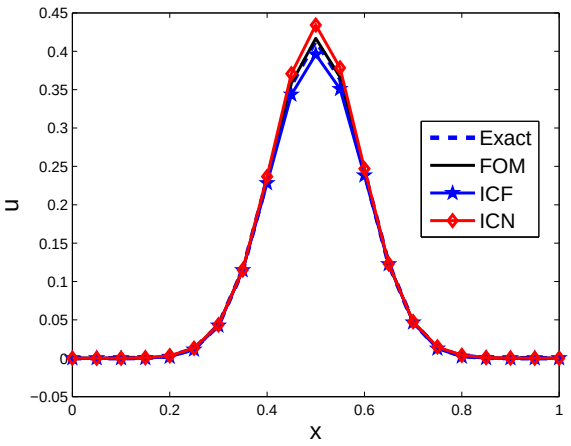
(a) u vs x when $\alpha = 0.01, \beta = 2$ at $T = 0.05$



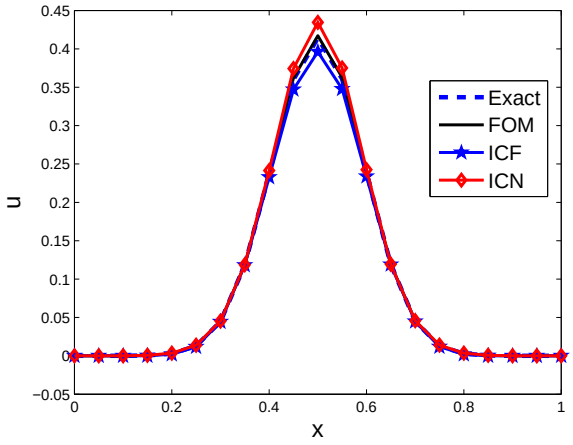
(b) u vs x when $\alpha = 0.01, \beta = 2$ at $T = 0.2$



(c) u vs x when $\alpha = 1, \beta = 1$ at $T = 0.2$



(d) u vs x when $\alpha = 0.01, \beta = 0.01$ at $T = 0.2$



(e) u vs x when $\alpha = 0.01, \beta = 0.001$ at $T = 0.2$

Figure 3.3: Numerical results when $h = 0.05$, $\Delta t = 0.001$ and $y = z = 0.5$.

Method	h	L_2-error	L_2-Rate
FOM	0.1	0.0013	
	0.05	2.6006×10^{-4}	2.3216
	0.025	2.0023×10^{-5}	3.6991
	0.0125	1.2781×10^{-6}	3.9696
ICN	0.1	0.0023	
	0.05	8.5445×10^{-4}	1.4286
	0.025	2.2617×10^{-4}	1.9176
	0.0125	5.573×10^{-5}	1.9792
ICF	0.1	4.0746×10^{-4}	
	0.05	1.8120×10^{-4}	1.1691
	0.025	4.0561×10^{-5}	2.1594
	0.0125	9.9999×10^{-6}	2.0201

Table 3.9: Convergence rate of FOM, ICN and ICF when $\beta = 0.8$ and $\alpha = 0.01$ with $\Delta t = 0.0001$ at $T = 0.01$.

Method	h	L_2-error	L_2-Rate	CPU time
FOM	0.05	2.4575×10^{-4}		0.368
	0.025	1.8858×10^{-5}	3.7039	4.222
	0.0125	1.2436×10^{-6}	3.9226	96.346
ICN	0.05	8.5474×10^{-4}		1.515
	0.025	2.2620×10^{-4}	1.9179	11.773
	0.0125	5.7365×10^{-5}	1.9794	339.296
ICF	0.05	1.8060×10^{-4}		0.657
	0.025	4.0517×10^{-5}	2.1562	11.430
	0.0125	9.9982×10^{-6}	2.0188	327.676

Table 3.10: Convergence rate of FOM, ICN and ICF when $\beta = 0.8$ and $\alpha = 0.01$ with $\Delta t = h^2$ at $T = 0.01$.

3.7 Optimal step size

In this section, we implement an optimization technique for the fourth order method to find an optimal time step size when the spatial step size is fixed as $h = 0.05$. Since the partial differential equation we consider is dissipative by nature and we observe from the numerical experiments carried out that the dissipative errors are very small as compared to the dispersion errors, we choose to minimize the dispersion error.

In this work, we extend the work in Appadu (2013), for which an optimization technique has been implemented to find an optimal temporal step size when solving 1D advection-diffusion equation. The relative phase error of a numerical method is given by

$$\text{RPE} = \frac{-1}{\left(c_x\omega_x + c_y\,\omega_y + c_z\,\omega_z\right)}\,\arctan\left(\frac{\Im(\xi_{num})}{\Re(\xi_{num})}\right),$$

where $\Re(\xi_{num})$ and $\Im(\xi_{num})$ are real and imaginary parts of the amplification factor of the scheme, respectively.

In Appadu (2013), the integrated error which minimizes dispersion error for a scheme discretizing the 1D advection-diffusion equation

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0.01\frac{\partial^2 u}{\partial x^2}, \tag{3.7.1}$$

is obtained as

$$\int_0^{1.1} (\text{RPE} - 1)^2 d\omega. \tag{3.7.2}$$

The spatial step size was chosen as 0.02 and the range of the temporal step size is determined. Then the integrated error is minimized and the optimal value is determined using NLPSolve function in maple.

Appadu and Gidey (2013) extended the work on 1D advection-diffusion (Appadu, 2013) to 2D advection-diffusion equation in order to find the optimal temporal step size of two finite difference methods: LOD Lax-Wendroff and LOD (1,5) when discretizing the 2D advection-diffusion equation

$$\frac{\partial u}{\partial t} + 0.8\frac{\partial u}{\partial x} + 0.8\frac{\partial u}{\partial y} = 0.01\frac{\partial^2 u}{\partial x^2} + 0.01\frac{\partial^2 u}{\partial y^2}, \tag{3.7.3}$$

and they define the integrated error as

$$\int_0^{1.1}\int_0^{1.1} (\text{RPE}-1)^2\, d\omega_x\, d\omega_y,$$

where the spacial step size was fixed as 0.025, and hence the RPE is a function of Δt, ω_x and ω_y. As there was phase wrapping, they made use of Taylor's expansion to obtain an approximation for the RPE upto the terms $\omega_x^i\,\omega_y^j$ with $i+j=5$.

Here, we extend the work on the 2D advection-diffusion equation by Appadu and Gidey (2013) to obtain an expression for the integrated error of the FOM when discretizing the partial differential equation

$$\frac{\partial u}{\partial t}+0.8\frac{\partial u}{\partial x}+0.8\frac{\partial u}{\partial y}+0.8\frac{\partial u}{\partial z}=0.01\frac{\partial^2 u}{\partial x^2}+0.01\frac{\partial^2 u}{\partial y^2}0.01\frac{\partial^2 u}{\partial z^2}, \tag{3.7.4}$$

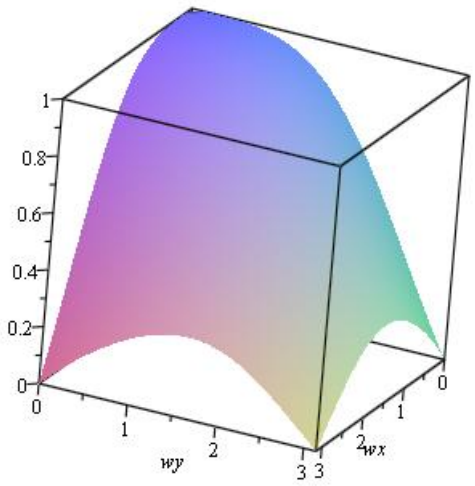
3D plots of the exact RPE versus ω_x versus ω_y for some values of Δt when $\omega_z=0$ are shown in Figs. 3.4a and 3.4b. Since phase wrapping phenomenon occurs for some value of Δt, as shown in Fig. 3.4b, we approximate the RPE till the terms $\mathcal{O}(\omega_x^i\,\omega_y^j\,\omega_z^k)$, with $i+j+k=5$, using Taylor's series for $\omega_x,\omega_y,\omega_z\in[0,1]$. The 3D plot of the approximated RPE versus ω_x versus ω_y, when $\omega_z=0$, is shown in Fig. 3.4d.

The integrated error is obtained as

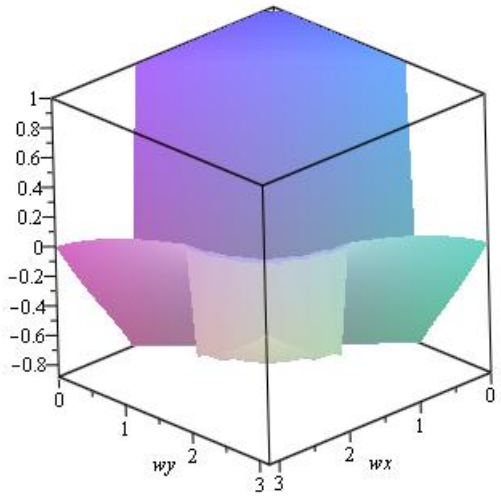
$$\int_0^1\int_0^1\int_0^1 (\text{RPE}_{\text{Approx}}-1)^2\, d\,\omega_x\, d\,\omega_y\, d\,\omega_z, \tag{3.7.5}$$

where RPE_{Approx} is the approximated relative phase error. A plot of the integrated error versus Δt is shown in Fig. 3.5. The integrated error obtained from FOM decreases as Δt increases from a value close to 0 to the value close 0.045 and then it increases as Δt increases, as shown in Figs. 3.5a and 3.5b. From Fig. 3.5a, it is observed that the integrated error is almost constant from $\Delta t=0.04$ to $\Delta t=0.05$ and it drastically increases after $\Delta t=0.05$.

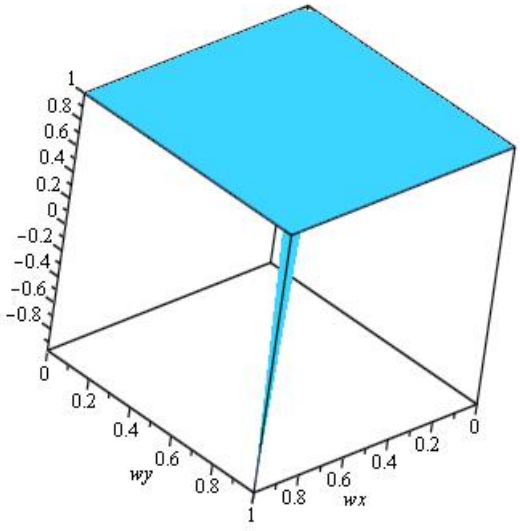
3.7 Optimal step size



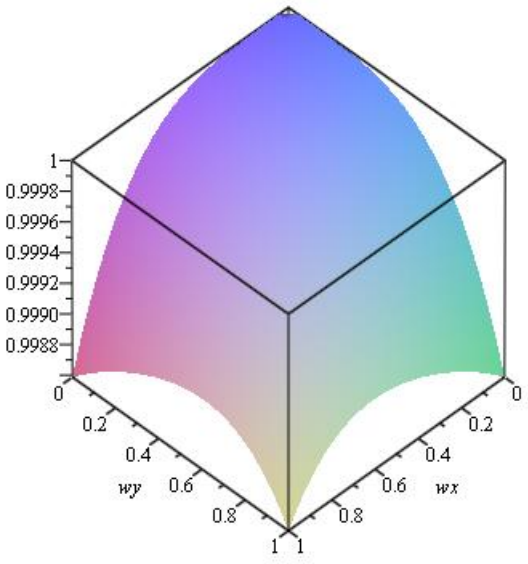
(a) Exact RPE vs ω_x vs ω_y when $\Delta t = 0.001$



(b) Exact RPE vs ω_x vs ω_y when$\Delta t = 0.05$



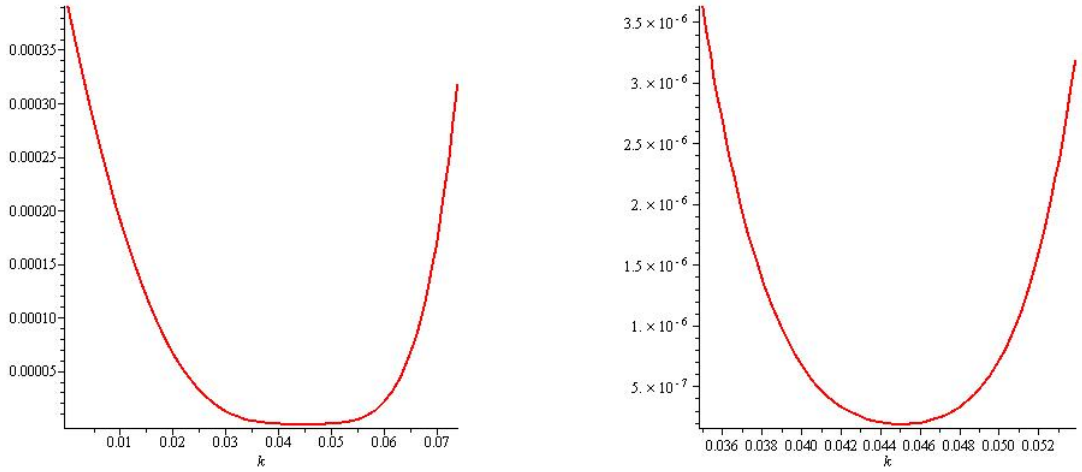
(c) RPE vs ω_x vs ω_y when $\Delta t = 0.05$ and $\omega_x, \omega_y \in [0, 1]$.



(d) Approximated RPE vs ω_x vs ω_y when $\Delta t = 0.05$ and $\omega_x, \omega_y \in [0, 1]$.

Figure 3.4: RPE versus ω_x versus ω_y when $\omega_z = 0$.

3.7 Optimal step size



(a) FOM for $\Delta t \in [0, 0.073865]$ (b) FOM ZOOMED

Figure 3.5: Integrated error versus Δt obtained from FOM, ICF and ICN.

Δt	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	44.5400	1	1	1	1.2227×10^{-32}	0.4072			
0.001	9.6606×10^{-4}	0.0181	44.5894	0.9988	1.0011	1.0016	1.2227×10^{-32}	0.4110	4.7813×10^{-10}	8.0571×10^{-7}	8.0618×10^{-7}
0.002	9.3842×10^{-4}	0.0175	44.5883	0.9989	1.0011	1.0013	1.2227×10^{-32}	0.4106	3.0461×10^{-10}	7.6040×10^{-7}	7.6071×10^{-7}
0.005	8.5490×10^{-4}	0.0162	44.5835	0.9991	1.0010	1.0004	1.2227×10^{-32}	0.4096	4.0174×10^{-11}	6.3130×10^{-7}	6.3134×10^{-7}
0.01	7.1512×10^{-4}	0.0140	44.5752	0.9994	1.0008	0.9991	1.2227×10^{-32}	0.4080	1.6340×10^{-10}	4.4160×10^{-7}	4.4176×10^{-7}
0.02	4.4447×10^{-4}	0.0093	44.5589	0.9997	1.0004	0.9979	-1.5821×10^{-10}	0.4060	8.3576×10^{-10}	1.6982×10^{-7}	1.7065×10^{-7}
0.04	6.0765×10^{-5}	0.0012	44.5407	1.0000	1.0000	0.9991	1.2227×10^{-32}	0.4060	1.4946×10^{-10}	3.0402×10^{-9}	3.1897×10^{-9}
0.05	4.3115×10^{-5}	9.0712×10^{-4}	44.5417	1.0000	1.0000	0.9996	1.2227×10^{-32}	0.4068	3.2830×10^{-11}	1.5730×10^{-9}	1.6058×10^{-9}
1/15	5.8212×10^{-4}	0.0118	44.5549	0.9996	1.0003	0.9935	-2.1744×10^{-8}	0.3998	7.7725×10^{-9}	2.8495×10^{-7}	2.9273×10^{-7}

Table 3.11: Numerical results from FOM when $\alpha = 0.01, \beta = 0.8, h = 0.05$ for some values of Δt at $T = 0.2$.

Δt	L_2-error	L_∞-error	Total mass	R^2	MCR	MDR	min u	max u	Diss. Error	Disp. Error	TMSE
Exact	0	0	0.0029	1	1	1	8.2051×10^{-46}	4.0398×10^{-4}			
0.001	5.9388×10^{-7}	5.0059×10^{-5}	0.0028	0.9929	0.9701	0.9896	-2.8931×10^{-9}	4.0398×10^{-4}	1.2319×10^{-15}	3.0344×10^{-13}	3.0467×10^{-13}
0.005	4.6819×10^{-7}	3.9332×10^{-5}	0.0029	0.9956	0.9740	0.9898	-5.3070×10^{-11}	4.0398×10^{-4}	1.1597×10^{-15}	1.8820×10^{-13}	1.8936×10^{-13}
0.02	1.8116×10^{-7}	1.4280×10^{-5}	0.0029	0.9993	0.9872	0.9944	8.2051×10^{-46}	4.0398×10^{-4}	3.4962×10^{-16}	2.8001×10^{-14}	2.8350×10^{-14}
0.04	2.1884×10^{-8}	1.6096×10^{-6}	0.0029	1	0.9982	0.9992	8.2051×10^{-46}	4.0398×10^{-4}	6.7095×10^{-18}	4.0700×10^{-16}	4.1371×10^{-16}
1/23	1.2772×10^{-8}	9.4243×10^{-7}	0.0029	1	0.9990	0.9995	8.2051×10^{-46}	4.0398×10^{-4}	2.2922×10^{-18}	1.3863×10^{-16}	1.4092×10^{-16}
1/22	1.0244×10^{-8}	7.5750×10^{-7}	0.0029	1	0.9992	0.9996	8.2051×10^{-46}	4.0398×10^{-4}	1.4639×10^{-18}	8.9188×10^{-17}	9.0652×10^{-17}
1/21	**9.8802** $\times 10^{-9}$	**7.2977** $\times 10^{-7}$	0.0029	1	**0.9992**	**0.9996**	8.2051×10^{-46}	4.0398×10^{-4}	**1.3541** $\times 10^{-18}$	**8.2972** $\times 10^{-17}$	**8.4326** $\times 10^{-17}$
0.05	1.2715×10^{-8}	9.3473×10^{-7}	0.0029	1	0.9990	0.9995	8.2051×10^{-46}	4.0398×10^{-4}	2.2534×10^{-18}	1.3741×10^{-16}	1.3966×10^{-16}
1/19	2.0497×10^{-8}	1.4999×10^{-6}	0.0029	1	0.9984	0.9993	8.2051×10^{-46}	4.0398×10^{-4}	5.8631×10^{-18}	3.5706×10^{-16}	3.6292×10^{-16}
0.0625	1.2431×10^{-7}	9.3980×10^{-6}	0.0029	0.9998	0.9905	0.9959	8.2051×10^{-46}	4.0398×10^{-4}	1.8503×10^{-16}	1.3163×10^{-14}	1.3348×10^{-14}
1/14	8.3678×10^{-7}	6.5463×10^{-5}	0.0028	0.9859	0.9592	0.9914	-7.9880×10^{-6}	4.0398×10^{-4}	9.3913×10^{-16}	6.0392×10^{-13}	6.0486×10^{-13}

Table 3.12: Numerical results from FOM when $\alpha = 0.01, \beta = 0.8, h = 0.05$ for some values of Δt at $T = 1$.

3.8 Conclusion

In this paper, three numerical methods have been used to solve a 3D advection-diffusion problem with spatial step size, $h = 0.05$ and temporal step size, $\Delta t = 0.001$ at some values of T. We compute L_2 error, L_∞ error, dissipation error, dispersion error, total mean square error, MCR, MDR, Minimum and Maximum values of u using the three methods. We observe that as we progress in time, the maximum value of u decreases as expected as the partial differential equation has dissipative terms. Based on the results obtained, we conclude that in general FOM is quite an efficient method to solve the problem for for some selected advective velocities and diffusivities. We also extend the optimization technique presented by Appadu and Gidey (2013) to a 3D problem to find an optimal temporal step size to minimize dispersion error when spatial step size is chosen as 0.05, for the case $\alpha = 0.01, \beta = 0.8$. This optimization works well and the optimal time step size is obtained from Figs. 3.5a and 3.5b. The optimal time step is validated using numerical experiments. Indeed, all the various errors are less at the optimal value of time step size as compared to other time step sizes. It is also shown that the FOM is fourth order accurate in space and the ICN and ICF are second order accurate in space and time.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Chapter 4

Analysis of multilevel finite volume approximation of convective Cahn-Hilliard equation

In this chapter, we solve 1D and 2D convective Cahn-Hilliard equations with specified initial condition and periodic boundary conditions using one-level and multilevel finite volume methods. The methods constructed for the 2D convective Cahn-Hilliard are analysed thoroughly. The existence/uniqueness, stability and convergence of the finite volume methods are proved. From the results obtained from the 2D problem, one can easily prove the existence/uniqueness, stability and convergence of the finite volume methods for the 1D convective Cahn-Hilliard equation. Hence, we only implement the corresponding finite volume methods for the 1D convective Cahn-Hilliard equation. Our results has been submitted for publication (Appadu et al., a,d).

4.1 Introduction

The general setting of this work is the 2D convective Cahn-Hilliard equation:

$$u_t - \gamma u(\boldsymbol{\beta} \cdot \nabla u) + \varepsilon^2 \Delta^2 u = \Delta f(u),\ (x, y) \in \mathcal{M}, t > 0, \tag{4.1.1}$$

with initial condition

$$u(x, y, 0) = u^0(x, y), \quad (x, y) \in \overline{\mathcal{M}}, \tag{4.1.2}$$

and periodic boundary conditions

$$\frac{\partial^j u}{\partial x^j}(-L_1, y, t) = \frac{\partial^j u}{\partial x^j}(L_1, y, t), \quad y \in (-L_2, L_2) \text{ and } 0 \leq t \leq T, \tag{4.1.3}$$

$$\frac{\partial^j u}{\partial y^j}(x, -L_2, t) = \frac{\partial^j u}{\partial y^j}(x, L_2, t), \quad x \in (-L_1, L_1) \text{ and } 0 \leq t \leq T, \tag{4.1.4}$$

where

$$f(u) = u^3 - u,$$

γ is the driving force, $j = 0, 1, 2, 3$, $\mathcal{M} = (-L_1, L_1) \times (-L_2, L_2)$, $\overline{\mathcal{M}}$ is the closure of $\mathcal{M}$, L_1 and L_2 are positive constants, $u^0 \in L^2(\mathcal{M})$, ε is a dimensionless interfacial width and $\boldsymbol{\beta}$ is a vector in 2D.

It is worth noting that (4.1.1) together with (4.1.3) and (4.1.4), for $j = 0$, leads to

$$\iint_{\mathcal{M}} u(x, y, t)\, dx\, dy = \iint_{\mathcal{M}} u^0(x, y)\, dx\, dy, \quad \forall t.$$

Hence for the analysis of (4.1.1)-(4.1.4), it is important to assume that (Temam, 2012)

$$\iint_{\mathcal{M}} u^0(x, y)\, dx\, dy = 0.$$

Our objective is to propose numerical techniques based on the work in Bousquet et al. (2013a, 2014) to compute the solution of (4.1.1)-(4.1.4) with $\gamma = 1$ and $\boldsymbol{\beta} = \langle 1, 1 \rangle$.

In many important phenomena (turbulence, excursion, etc) the solutions involves multi-scale analysis. Hence a reliable simulation requires a large number of degrees of freedom, which increases the calculation costs. Multilevel simulation in which the principle rely on

the separation of scales are therefore one possibility to describe these phenomena. Early contribution include (Chen and Temam, 1993; Folas et al., 1988; Marion and Temam, 1989; Temam, 1990), but there is now a vast literature in this research direction. It is worth mentioning that though the approach may differ from researchers to researchers, the common feature remains the same: the separation of scales.

The multilevel method we discus in this chapter was formulated by Bousquet et al. (2014), in which a hierarchical multilevel finite volume discretization is implemented for the 2D shallow water linearized around a constant flow. The contribution in Bousquet et al. (2014) is a followup of ideas started in Adamy et al. (2010) and Bousquet et al. (2013b). Our motivation in this work are as follows: formulate, analyse and implement the multilevel approach advocate in Bousquet et al. (2014) for the 2D nonlinear partial differential equation with high order derivative. One of the challenge as mentioned earlier is to discretize the nonlinear term $u(\boldsymbol{\beta} \cdot \nabla u)$ in a linear way while maintaining basic properties, and as a consequence saving computational time.

We construct two schemes associated with (4.1.1)-(4.1.4) based on the work of Bousquet et al. (2014). The schemes we construct are easy to implement and are respectively called:

(a) linear implicit multilevel finite volume approximation, and

(b) explicit multilevel finite volume approximation.

For the sake of comparison, we also formulate two one-level methods associated to the multilevel methods. One of the difficulties is to design an appropriate linear expression for the nonlinear term. We achieve that thanks to the nonlocal approximation of nonlinear quantity introduced by Mickens (1994) and Anguelov and Lubuma (2001). In particular, following Djoko (2008), we approximate the nonlinear term $u\,(\boldsymbol{\beta} \cdot \nabla u)$ in a linear way such that the property

$$\iint_{\mathcal{M}} u\,(\boldsymbol{\beta} \cdot \nabla u) u \, dx \, dy = 0 \tag{4.1.5}$$

is constructed at the discrete level.

After the construction of new schemes, we show the existence and uniqueness of the solution. At this step, we should bear in mind that since we are dealing with linear equations in finite dimension, existence of solutions is equivalent to uniqueness, thus, we provide conditions under which there is one solution. The third contribution of this work is the stability of the new schemes. We show that the implicit multilevel method is conditionally stable with a region of stability smaller than one obtained from the one-level implicit method on the fine mesh. The fourth contribution of our study is the convergence analysis of the implicit methods. Indeed, we show that the implicit methods are first order accurate in time and second order accurate in space. Our last contribution is numerical result that supports our theoretical findings. We compute L_2-error and rate of convergence for the proposed numerical methods. We also demonstrate that in all numerical tests, the multilevel methods are faster than the one-level methods on the fine mesh.

The remaining part of this chapter is organized as follows: in the next section, we recall some preliminaries and introduce some standard notations. We also discuss, in section 4.2, some properties of difference operators and the discrete analogue of L_2 space. In sections 4.3 and 4.4, we construct one-level and multilevel finite volume methods and proved those methods are conditionally stable and conditionally convergent. In section 4.5, we present some numerical results comparing computations done by one-level methods and computations done by the multilevel methods. In section 4.6, we solve 1D convective Cahn-Hilliard equation using one-level and multilevel finite volume methods. Lastly, conclusions are given in section 4.7.

4.2 Some preliminaries and space discretizations

In this section, we recall some preliminaries which are helpful to our discussion and we present the space discretization in a 2D rectangular region. To develop finite volume approximations that satisfy the discrete analogue of Eq. (4.1.5), we first introduce some standard notations and results. We partition $\mathcal{M}$ into $N_1 \times N_2$ control volumes $(k_{i,j})_{1\leq i\leq N_1, 1\leq j\leq N_2}$ of uniform area $\Delta x\,\Delta y$, where Δx and Δy are the spatial step sizes in the x- and y- di-

rections, respectively. It is assumed that the partition of the domain is conform, meaning that for two elements A and B one has, $A \cap B$ is either a face, a vertex or empty set. For $0 \leq i \leq N_1$ and $0 \leq j \leq N_2$,

$$x_{i+1/2} = i\Delta x - L_1,\ y_{j+1/2} = j\Delta y - L_2,$$

so that

$$k_{i,j} = (x_{i-1/2}, x_{i+1/2}) \times (y_{j-1/2}, y_{j+1/2}) \text{ for } 1 \leq i \leq N_1, 1 \leq j \leq N_2.$$

(x_i, y_j) is the centre of the (i,j) control volume, which is given by the formula

$$(x_i, y_j) = \left((i-1)\Delta x + \frac{\Delta x}{2} - L_1,\ (j-1)\Delta y + \frac{\Delta y}{2} - L_2 \right), \quad 1 \leq i \leq N_1, 1 \leq j \leq N_2.$$

In the rest of this work, we take $h = (\Delta x, \Delta y)$. The approximate solution to the control volume average of the true solution at $t_n = n\Delta t$ is denoted by $u^n_{i,j}$, i.e.

$$u^n_{i,j} \approx \frac{1}{\Delta x \Delta y} \iint\limits_{k_{i,j}} u(x, y, t_n)\, dx\, dy, 1 \leq i \leq N_1, 1 \leq j \leq N_2,$$

where Δt is the temporal step size such that $\Delta t M = T$, which is obtained recursively by starting with the initial average value, $u^0_{i,j}$, given by

$$u^0_{i,j} = \frac{1}{\Delta x \Delta y} \iint\limits_{k_{i,j}} u^0(x, y)\, dx\, dy, 1 \leq i \leq N_1,\ 1 \leq j \leq N_2.$$

Define the space $\mathcal{H}_h$ as

$$\mathcal{H}_h = \left\{ \mathbf{u} = \left(u_{i,j}\right)_{i,j \in \mathbb{Z}}, u_{i,j} \in \mathbb{R} |\ u_{i+N_1,j} = u_{i,j} = u_{i,j+N_2},\ \text{and}\ \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} u_{i,j} = 0 \right\},$$

equipped with the inner product and discrete L^2 norm

$$(\mathbf{u}, \mathbf{v})_h = \Delta x \Delta y \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} u_{i,j} v_{i,j} \quad \text{and} \quad \|\mathbf{u}\|_h = \left(\Delta x \Delta y \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} u^2_{i,j} \right)^{1/2},$$

respectively.

For $\mathbf{u} \in \mathcal{H}_h$, we introduce the following difference operators:

$$\nabla^-_{1,h} u_{i,j} = \frac{1}{\Delta x} \left(u_{i,j} - u_{i-1,j} \right),\ \nabla^+_{1,h} u_{i,j} = \frac{1}{\Delta x} \left(u_{i+1,j} - u_{i,j} \right), \tag{4.2.1}$$

$$\nabla^-_{2,h}u_{i,j} = \frac{1}{\Delta y}\left(u_{i,j} - u_{i,j-1}\right), \ \nabla^+_{2,h}u_{i,j} = \frac{1}{\Delta y}\left(u_{i,j+1} - u_{i,j}\right), \tag{4.2.2}$$

$$\Delta_{1,h}u_{i,j} = \frac{1}{\Delta x^2}\left(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}\right), \tag{4.2.3}$$

$$\Delta_{2,h}u_{i,j} = \frac{1}{\Delta y^2}\left(u_{i,j+1} - 2u_{i,j} + u_{i,j-1}\right), \tag{4.2.4}$$

$$\Delta^2_{1,h}u_{i,j} = \frac{1}{\Delta x^2}\left(\Delta_{1,h}u_{i+1,j} - 2\Delta_{1,h}u_{i,j} + \Delta_{1,h}u_{i-1,j}\right), \tag{4.2.5}$$

$$\Delta^2_{2,h}u_{i,j} = \frac{1}{\Delta y^2}\left(\Delta_{2,h}u_{i,j+1} - 2\Delta_{2,h}u_{i,j} + \Delta_{2,h}u_{i,j-1}\right). \tag{4.2.6}$$

From (4.2.1)-(4.2.6), we have

$$\boldsymbol{\beta}\cdot\nabla^\pm_h = \nabla^\pm_{1,h} + \nabla^\pm_{2,h}, \ \Delta_h = \Delta_{1,h} + \Delta_{2,h}, \ \Delta^2_h = \Delta^2_{1,h} + \Delta_{1,h}\Delta_{2,h} + \Delta_{2,h}\Delta_{1,h} + \Delta^2_{2,h}. \tag{4.2.7}$$

The discrete analogue of the derivative of product of functions is given as follows: for $\mathbf{u}, \mathbf{v} \in \mathcal{H}_h$,

$$(\boldsymbol{\beta}\cdot\nabla^+_h)(u_{i,j}v_{i,j}) = (\nabla^+_{1,h}u_{i,j})v_{i+1,j} + u_{i,j}(\nabla^+_{1,h}v_{i,j}) + (\nabla^+_{2,h}u_{i,j})v_{i,j+1} + u_{i,j}(\nabla^+_{2,h}v_{i,j}), \tag{4.2.8}$$

$$(\boldsymbol{\beta}\cdot\nabla^-_h)(u_{i,j}v_{i,j}) = (\nabla^-_{1,h}u_{i,j})v_{i-1,j} + u_{i,j}(\nabla^-_{1,h}v_{i,j}) + (\nabla^-_{2,h}u_{i,j})v_{i,j-1} + u_{i,j}(\nabla^-_{2,h}v_{i,j}).. \tag{4.2.9}$$

From the definition of $\mathcal{H}_h$ and the discrete product rules, (4.2.8) and (4.2.9), one obtains:

Lemma 4.2.1. *Let $\boldsymbol{u}, \boldsymbol{w} \in \mathcal{H}_h$. Then for any vector $\boldsymbol{\beta} = \langle\beta_1, \beta_2\rangle$ with $\beta_1, \beta_2 \in \mathbb{R}$*

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}(\boldsymbol{\beta}\cdot\nabla^+_h)u_{i,j} = -\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}(\boldsymbol{\beta}\cdot\nabla^-_h)w_{i,j}.$$

Proof. To prove this, we use the definition of $\mathcal{H}_h$.

$$\begin{aligned}
\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}(\boldsymbol{\beta}\cdot\nabla^+_h)u_{i,j} &= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}(\beta_1\,\nabla^+_{1,h}u_{i,j} + \beta_2\,\nabla^+_{2,h}u_{i,j}) \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}\left[\beta_1\left(\frac{u_{i+1,j}-u_{i,j}}{\Delta x}\right) + \beta_2\left(\frac{u_{i,j+1}-u_{i,j}}{\Delta y}\right)\right] \\
&= \frac{\beta_1}{\Delta x}\left(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i+1,j} - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\beta_2}{\Delta y}\left(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j+1}-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)\\
&=\frac{\beta_1}{\Delta x}\left(\sum_{i=1}^{N_1-1}\sum_{j=1}^{N_2} w_{i,j}u_{i+1,j}-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)\\
&\quad+\frac{\beta_2}{\Delta y}\left(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2-1} w_{i,j}u_{i,j+1}-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)\\
&\quad+\frac{\beta_1}{\Delta x}\sum_{j=1}^{N_2} w_{N_1,j}u_{N_1+1,j}+\frac{\beta_2}{\Delta y}\sum_{i=1}^{N_1} w_{i,N_2}u_{i,N_2+1}\\
&=\frac{\beta_1}{\Delta x}\left(\sum_{i=2}^{N_1}\sum_{j=1}^{N_2} w_{i-1,j}u_{i,j}-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)\\
&\quad+\frac{\beta_2}{\Delta y}\left(\sum_{i=1}^{N_1}\sum_{j=2}^{N_2} w_{i,j-1}u_{i,j}-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} w_{i,j}u_{i,j}\right)\\
&\quad+\frac{\beta_1}{\Delta x}\sum_{j=1}^{N_2} w_{N_1,j}u_{1,j}+\frac{\beta_2}{\Delta y}\sum_{i=1}^{N_1} w_{i,N_2}u_{i,1}\\
&=-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}\left[\beta_1\left(\frac{w_{i,j}-w_{i-1,j}}{\Delta x}\right)+\beta_2\left(\frac{w_{i,j}-u_{i,j-1}}{\Delta y}\right)\right]\\
&=-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)w_{i,j}.
\end{aligned}
$$

□

We define the following discrete semi-norms and norms for $\mathbf{u}=(u_{i,j})$, $1\le i\le N_1$, $1\le j\le N_2$.

$$
|\mathbf{u}|_{1,h}=\left(\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(\nabla_{1,h}^- u_{i,j})^2+(\nabla_{2,h}^- u_{i,j})^2\right]\right)^{\frac{1}{2}}, \tag{4.2.10}
$$

$$
\|\mathbf{u}\|_{1,h}^2=|\mathbf{u}|_{1,h}^2+\|\mathbf{u}\|_h^2,
$$

$$
|\mathbf{u}|_{2,h}=\left(\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\Delta_h u_{i,j})^2\right)^{\frac{1}{2}}, \tag{4.2.11}
$$

$$
\|\mathbf{u}\|_{\infty,h}=\max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}}|u_{i,j}|.
$$

In Eq. (4.2.10), $\nabla_{1,h}^-$ and $\nabla_{2,h}^-$ can be replaced by $\nabla_{1,h}^+$ and $\nabla_{2,h}^+$, respectively. Using Eqs.

(4.2.10) and (4.2.11), we have

$$|\mathbf{u}|_{1,h}^2 \leq 4\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\ \|\mathbf{u}\|_h^2, \tag{4.2.12}$$

and the following are obtained by direct computations

$$\|\mathbf{u}\|_h^2 \leq 4L_1L_2\|\mathbf{u}\|_{\infty,h}^2$$

and

$$\|\mathbf{u}\|_{\infty,h}^2 \leq \frac{1}{\Delta x\,\Delta y}\|\mathbf{u}\|_h^2. \tag{4.2.13}$$

Moreover, it is important to note that if $\mathbf{u}$ belongs to $\mathcal{H}_h$, then the discrete Poincaré's inequality holds; this is to say that there is $\eta > 0$, independent of Δx and Δy such that

$$\eta\,\|\mathbf{u}\|_h \leq |\mathbf{u}|_{1,h}. \tag{4.2.14}$$

Remark 4.2.1. *With* (4.2.14)*, we conclude that the semi-norm* $|\cdot|_{1,h}$ *is a norm on* $\mathcal{H}_h$ *equivalent to* $\|\cdot\|_{1,h}$.

For any $\mathbf{u},\mathbf{v}\in\mathcal{H}_h$, one easily obatains

$$2(\mathbf{u}-\mathbf{v},\mathbf{u})_h = \|\mathbf{u}\|_h^2 - \|\mathbf{v}\|_h^2 + \|\mathbf{u}-\mathbf{v}\|_h^2, \tag{4.2.15}$$

$$2(\mathbf{u}-\mathbf{v},\mathbf{v})_h = \|\mathbf{u}\|_h^2 - \|\mathbf{v}\|_h^2 - \|\mathbf{u}-\mathbf{v}\|\|_h^2. \tag{4.2.16}$$

The following result belong to calculus of multi-variable. It is an intermediate step in showing that the solution u of Eqs. (4.1.1)-(4.1.4), satisfies the property (4.1.5).

Lemma 4.2.2 (Green's Theorem).

$$\iint_{\mathcal{M}}\left(\frac{\partial M}{\partial x}-\frac{\partial L}{\partial y}\right)dx\,dy = \oint_C (Ldx+Mdy)dxdy.$$

Lemma 4.2.3. *The solution* u *of Eqs.* (4.1.1)- (4.1.4) *satisfy the relation:*

$$\iint_{\mathcal{M}} u^2(\boldsymbol{\beta}\cdot\nabla u)dA = 0. \tag{4.2.17}$$

Proof. The integrand can be expanded as follows:

$$u^2(\boldsymbol{\beta}\cdot\nabla)u = u^2u_x + u^2u_y = \frac{1}{3}\left(\frac{\partial}{\partial x}(u^3) - \frac{\partial}{\partial y}(-u^3)\right).$$

Using Green's Theorem, we obtain

$$\iint_{\mathcal{M}} u^2(\boldsymbol{\beta}\cdot\nabla)u dA = \frac{1}{3}\oint_C -u^3dx + u^3dy, \tag{4.2.18}$$

where $C = C_1 \cup C_2 \cup C_3 \cup C_4$, the contour line shown by Fig. 4.1. The right hand side of Eq. (4.2.18), we have

$$\begin{aligned}
\oint_C -u^3dx + u^3dy &= \int_{C_1} -u(x,-L_2)^3dx + \int_{C_2} u(L_1,y)^3dy \\
&\quad - \int_{C_3} u(x,L_2)^3dx + \int_{C_4} u(-L_1,y)^3dy \\
&= -\int_{-L_1}^{L_1} u(x,-L_2)^3dx + \int_{-L_2}^{L_2} u(L_1,y)^3dy \\
&\quad - \int_{L_1}^{-L_1} u(x,L_2)^3dx + \int_{L_2}^{-L_2} u(-L_1,y)^3dy \\
&= \int_{-L_1}^{L_1} \left(u(x,L_2)^3 - u(x,-L_2)^3\right)dx + \int_{-L_2}^{L_2} \left(u(L_1,y)^3 - u(-L_1,y)^3\right)dy \\
&= 0.
\end{aligned}$$

□

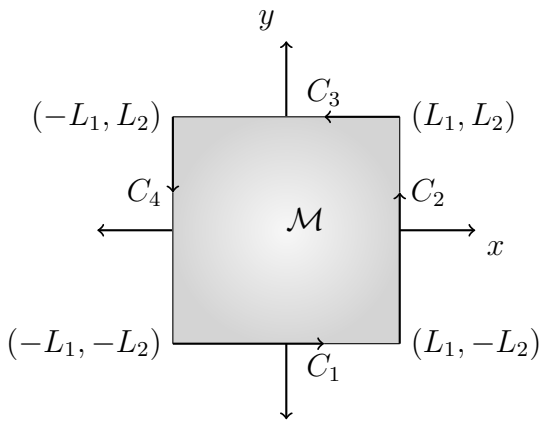


Figure 4.1: Representation of the contour diagram.

Hence, when discretizing (4.1.1), it is desirable to approximate $u\,(\boldsymbol{\beta}\cdot\nabla)\,u$ in such a way that the discrete counter part of (4.2.17) hold.

In order to approximate the nonlinear term, we introduce the bilinear map: $C_h : \mathcal{H}_h \times \mathcal{H}_h \to \mathbb{R}^{N_1\times N_2}$ in the form

$$\begin{aligned} C_h(\mathbf{u},\mathbf{v})_{i,j} = \alpha_1[u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)v_{i,j} + v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)u_{i,j} + v_{i+1,j}\nabla_{1,h}^+u_{i,j} + v_{i,j+1}\nabla_{2,h}^+u_{i,j}] \\ + \alpha_2[u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)v_{i,j} + v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)u_{i,j} + v_{i-1,j}\nabla_{1,h}^-u_{i,j} + v_{i,j-1}\nabla_{2,h}^-u_{i,j}], \end{aligned} \tag{4.2.19}$$

where α_1 and α_2 are constants. We use this bilinear map to approximate the nonlinear term $u(\boldsymbol{\beta}\cdot\nabla)u$ at t_{n+1} and t_n for the implicit and explicit methods, respectively.

Using Eqs. (4.2.8), (4.2.9) and Lemma 4.2.1, we prove the following.

Lemma 4.2.4. *For $\boldsymbol{u},\boldsymbol{v}\in\mathcal{H}_h$*

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(C_h(\boldsymbol{u},\boldsymbol{v}))_{i,j}u_{i,j} = 0. \tag{4.2.20}$$

Proof. For all $\mathbf{u},\mathbf{v}\in\mathcal{H}_h$, we have

$$\begin{aligned} \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}((\boldsymbol{\beta}\cdot\nabla_h^+)v_{i,j})u_{i,j} &= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)(v_{i,j}u_{i,j}) - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{1,h}^+u_{i,j})v_{i+1,j} \\ &\quad - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{2,h}^+u_{i,j})v_{i,j+1} \qquad \text{using (4.2.8)} \\ &= -\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)(u_{i,j}) - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{1,h}^+u_{i,j})v_{i+1,j} \\ &\quad - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{2,h}^+u_{i,j})v_{i,j+1} \qquad \text{using Lemma 4.2.1.} \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}((\boldsymbol{\beta}\cdot\nabla_h^-)v_{i,j})u_{i,j} &= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)(v_{i,j}u_{i,j}) - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{1,h}^-u_{i,j})v_{i-1,j} \\ &\quad - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}u_{i,j}(\nabla_{2,h}^-u_{i,j})v_{i,j-1} \qquad \text{using (4.2.9)} \end{aligned}$$

$$= -\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)(u_{i,j}) - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}(\nabla_{1,h}^- u_{i,j})v_{i-1,j}$$
$$-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}(\nabla_{2,h}^- u_{i,j})v_{i,j-1} \qquad \text{using Lemma 4.2.1.}$$

Thus we have

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}\left[u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)v_{i,j} + v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)u_{i,j} + v_{i+1,j}\nabla_{1,h}^+ u_{i,j} + v_{i,j+1}\nabla_{2,h}^+ u_{i,j}\right] = 0$$

and

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{i,j}\left[u_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^-)v_{i,j} + v_{i,j}(\boldsymbol{\beta}\cdot\nabla_h^+)u_{i,j} + v_{i-1,j}\nabla_{1,h}^- u_{i,j} + v_{i,j-1}\nabla_{2,h}^- u_{i,j}\right] = 0.$$

Therefore, the proof is complete. □

Remark 4.2.2. *For any d-dimensional space problem with $d \geq 3$, we can easily extend Eq.* (4.2.19) *such that an analogous of* (4.2.20) *holds. That is*

$$C_h(\boldsymbol{u},\boldsymbol{v})_{i_1,i_2,\ldots,i_d} = \alpha_1\Big[u_{i_1,i_2,\ldots,i_d}(\boldsymbol{\beta}\cdot\nabla_h^+)v_{i_1,i_2,\ldots,i_d} + v_{i_1,i_2,\ldots,i_d}(\boldsymbol{\beta}\cdot\nabla_h^-)u_{i_1,i_2,\ldots,i_d}$$
$$+\sum_{s=1}^{d} v_{s+}\nabla_{s,h}^+ u_{i_1,i_2,\ldots,i_d}\Big] + \alpha_2\Big[u_{i_1,i_2,\ldots,i_d}(\boldsymbol{\beta}\cdot\nabla_h^-)v_{i_1,i_2,\ldots,i_d}$$
$$+ v_{i_1,i_2,\ldots,i_d}(\boldsymbol{\beta}\cdot\nabla_h^+)u_{i_1,i_2,\ldots,i_d} + \sum_{s=1}^{d} v_{s-}\nabla_{s,h}^- u_{i_1,i_2,\ldots,i_d}\Big],$$

where $v_{s\pm} = v_{i_1,\ldots,i_s\pm1,\ldots,i_d}$, for $s = 1,2,\ldots,d$ and i_s is the position of the vector at the s^{th} coordinate.

Lemma 4.2.5. *For $\boldsymbol{u},\boldsymbol{w} \in \mathcal{H}_h$*

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}(\Delta_{1,h}u_{i,j})w_{i,j} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}(u_{i,j})\Delta_{1,h}(w_{i,j}).$$

Proof. For all $\mathbf{u},\mathbf{w} \in \mathcal{H}_h$, we have

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}(\Delta_{1,h}u_{i,j})w_{i,j} = \sum_{j=1}^{N_2}\Bigg[\sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i+2,j} - 2u_{i+1,j} + u_{i,j})w_{i,j}$$
$$-2\sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i+1,j} - 2u_{i,j} + u_{i-1,j})w_{i,j}$$
$$+\sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i,j} - 2u_{i-1,j} + u_{i-2,j})w_{i,j}\Bigg]. \qquad (4.2.21)$$

From the periodic boundary conditions, $\sum_{i=1}^{N_1} u_{i,j} = \sum_{i=1}^{N_1} u_{i-1,j} = \sum_{i=1}^{N_1} u_{i+1,j}$ for each $j = 1, \ldots, N_2$ and hence (4.2.21) yields

$$\begin{aligned}
\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}(\Delta_{1,h}u_{i,j})w_{i,j} &= \sum_{j=1}^{N_2}\Bigg[\sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i+1,j} - 2u_{i,j} + u_{i-1,j})w_{i-1,j} \\
&\qquad - 2\sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i+1,j} - 2u_{i,j} + u_{i-1,j})w_{i,j} \\
&\qquad + \sum_{i=1}^{N_1}\frac{1}{\Delta x^4}(u_{i+1,j} - 2u_{i,j} + u_{i-1,j})w_{i+1,j}\Bigg] \\
&= \sum_{j=1}^{N_2}\Bigg[\sum_{i=1}^{N_1}\frac{1}{\Delta x^2}(\Delta_{1,h}u_{i,j})w_{i-1,j} - 2\sum_{i=1}^{N_1}\frac{1}{\Delta x^2}(\Delta_{1,h}u_{i,j})w_{i,j} \\
&\qquad + \sum_{i=1}^{N_1}\frac{1}{\Delta x^2}(\Delta_{1,h}u_{i,j})w_{i+1,j}\Bigg] \\
&= \sum_{j=1}^{N_2}\sum_{i=1}^{N_1}(\Delta_{1,h}u_{i,j})\left[\frac{1}{\Delta x^2}(w_{i-1,j} - 2w_{i,j} + w_{i+1,j})\right] \\
&= \sum_{j=1}^{N_2}\sum_{i=1}^{N_1}(\Delta_{1,h}u_{i,j})(\Delta_{1,h}w_{i,j}).
\end{aligned}$$

□

Lemma 4.2.5 also holds when one (or two) of the operator(s) $\Delta_{1,h}$ is (are) replaced by $\Delta_{2,h}$.

Lemma 4.2.6. *For* $\boldsymbol{u}, \boldsymbol{w} \in \mathcal{H}_h$

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_h^2(u_{i,j})w_{i,j} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_h(u_{i,j})\Delta_h(w_{i,j}).$$

Proof. For any $\mathbf{u}, \mathbf{w} \in \mathcal{H}_h$ using (4.2.7), we have

$$\begin{aligned}
\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_h^2(u_{i,j})\,w_{i,j} &= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(\Delta_{1,h}^2 + \Delta_{1,h}\Delta_{2,h} + \Delta_{2,h}\Delta_{1,h} + \Delta_{2,h}^2)(u_{i,j})\right]w_{i,j} \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_{1,h}(\Delta_{1,h}u_{i,j})w_{i,j} + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_{1,h}(\Delta_{2,h}u_{i,j})w_{i,j} \\
&\quad + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_{2,h}(\Delta_{1,h}u_{i,j})w_{i,j} + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Delta_{2,h}(\Delta_{2,h}u_{i,j})w_{i,j}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}u_{i,j}\Delta_{1,h}w_{i,j} + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{2,h}u_{i,j}\Delta_{1,h}w_{i,j} \\
&\quad + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}u_{i,j}\Delta_{2,h}w_{i,j} \\
&\quad + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{2,h}u_{i,j}\Delta_{2,h}w_{i,j} \qquad \text{using Lemma 4.2.5} \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{1,h}u_{i,j}(\Delta_{1,h}w_{i,j} + \Delta_{2,h}w_{i,j}) \\
&\quad + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_{2,h}u_{i,j}(\Delta_{1,h}w_{i,j} + \Delta_{2,h}w_{i,j}) \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} (\Delta_{1,h}u_{i,j} + \Delta_{2,h}u_{i,j})(\Delta_{1,h}w_{i,j} + \Delta_{2,h}w_{i,j}) \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Delta_h u_{i,j}\Delta_h w_{i,j}.
\end{aligned}
$$

□

The following lemma will be used later.

Lemma 4.2.7. *For* $\boldsymbol{u} \in \{\boldsymbol{u} = (u)_{i,j}, u_{i,j} \in \mathbb{R} |\ u_{i+N_1,j} = u_{i,j} = u_{i,j+N_2},\ i,j \in \mathbb{Z}\}$*, the following inequality holds true*

$$|\boldsymbol{u}|_{1,h}^2 \leq |\boldsymbol{u}|_{2,h}\|\boldsymbol{u}\|_h.$$

Proof. Using (4.2.7), for $\mathbf{u} \in \{\mathbf{u} = (u)_{i,j}, u_{i,j} \in \mathbb{R} |\ u_{i+N_1,j} = u_{i,j} = u_{i,j+N_2},\ i,j \in \mathbb{Z}\}$, we have

$$
\begin{aligned}
(\Delta_h\mathbf{u}, \mathbf{u})_h &= (\Delta_{1,h}\mathbf{u} + \Delta_{2,h}\mathbf{u}, \mathbf{u})_h \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{\Delta x^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\Delta y^2}\right] u_{i,j} \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left(\frac{u_{i+1,j} - u_{i,j}}{\Delta x^2}\right) u_{i,j} - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left(\frac{u_{i,j} - u_{i-1,j}}{\Delta x^2}\right) u_{i,j} \\
&\quad + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left(\frac{u_{i,j+1} - u_{i,j}}{\Delta y^2}\right) u_{i,j} - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left(\frac{u_{i,j} - u_{i,j-1}}{\Delta y^2}\right) u_{i,j}. \qquad (4.2.22)
\end{aligned}
$$

Using periodic boundary conditions, we have

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i+1,j}-u_{i,j}}{\Delta x^2}\right)u_{i,j}=\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i,j}-u_{i-1,j}}{\Delta x^2}\right)u_{i-1,j} \tag{4.2.23}$$

and

$$\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i,j+1}-u_{i,j}}{\Delta y^2}\right)u_{i,j}=\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i,j}-u_{i,j-1}}{\Delta y^2}\right)u_{i,j-1}, \tag{4.2.24}$$

Combining Eqs (4.2.22), (4.2.23) and (4.2.24), we obtain

$$(\Delta_h\mathbf{u},\mathbf{u})_h=-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i,j}-u_{i-1,j}}{\Delta x}\right)^2-\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\frac{u_{i,j}-u_{i,j-1}}{\Delta y}\right)^2,$$

which is

$$(\Delta_h\mathbf{u},\mathbf{u})_h=-|\mathbf{u}|_{1,h}^2.$$

Using Cauchy-Schwarz's inequality, we have

$$\begin{aligned}|\mathbf{u}|_{1,h}^2&=-(\Delta_h\mathbf{u},\mathbf{u})_h\\&\leq\|\Delta_h\mathbf{u}\|_h\|\mathbf{u}\|_h=|\mathbf{u}|_{2,h}\|\mathbf{u}\|_h.\end{aligned}$$

□

For $\mathbf{u}\in\mathcal{H}_h$, Lemma 4.2.7 and Young's inequality implies the existence of η, positive constant independent of both Δy, and Δx such that

$$\eta|\mathbf{u}|_{1,h}\leq|\mathbf{u}|_{2,h}. \tag{4.2.25}$$

4.3 One-level finite volume methods

In this section, we present two traditional one-level finite volume methods: namely implicit finite volume method and explicit finite volume method. The existence, uniqueness and convergence of solution for the implicit method are proved and stability analysis is examined for both schemes. For both methods thirteen point stencils are used to approximate (4.1.1)-(4.1.4), as shown in Fig. 4.2. The introduction of these classical schemes is important at least for three reasons:

(a) comparison with multilevel methods;

(b) these schemes that are categorized as classical present significant challenges for their analysis as we will see;

(c) the analysis of these schemes will shed lights in the analysis of multilevel methods.

4.3.1 Implicit one-level finite volume method

The nonlinear term $u\,(\boldsymbol{\beta}\cdot\nabla u)$ at t_{n+1} is approximated linearly using the bilinear map defined in section 4.2, Eq. (4.2.19), and is given by

$$\left[u\,(\boldsymbol{\beta}\cdot\nabla)\,u\right]|_{i,j}^{n+1} \approx \left(C_h(\mathbf{u}^{n+1},\tilde{\mathbf{u}}^n)\right)_{i,j}, \tag{4.3.1}$$

where $\tilde{\mathbf{u}}^n$ is the approximation of $\mathbf{u}^{n+1}$, given by

$$\tilde{\mathbf{u}}^n = a_1\mathbf{u}^n + a_2\mathbf{u}^{n-1} + a_3\mathbf{u}^{n-2} + \cdots + a_{m_0}\mathbf{u}^{n-m_0+1}, \tag{4.3.2}$$

where $m_0 \in \{1,2,\ldots,n\}$ and $a_1, a_2, \ldots, a_{m_0}$ are coefficients that determine the approximation with $3(\alpha_1+\alpha_2)(a_1+a_2+\cdots+a_{m_0}) = 1$, ensuring consistency of the approximation. For $m < m_0 - 1$, the term $\tilde{\mathbf{u}}^m$ is given by the relation

$$\tilde{\mathbf{u}}^m = \mathbf{u}^m. \tag{4.3.3}$$

We approximate the nonlinear term on the right hand side of (4.1.1) at t_{n+1} by a linear second order accurate in space as follows:

$$\Delta f(u)|_{i,j}^{n+1} \approx \nabla_{1,h}^+(\varphi_{i-1/2,j}^n \nabla_{1,h}^- u_{i,j}^{n+1}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^n \nabla_{2,h}^- u_{i,j}^{n+1}). \tag{4.3.4}$$

where

$$\varphi_{i-1/2,j}^n = \frac{f'(u_{i,j}^n) + f'(u_{i-1,j}^n)}{2} \text{ and } \varphi_{i,j-1/2}^n = \frac{f'(u_{i,j}^n) + f'(u_{i,j-1}^n)}{2}.$$

Lemma 4.3.1. *For* $\boldsymbol{u}^n, \boldsymbol{u}^{n+1} \in \mathcal{H}_h$

$$\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\nabla_{1,h}^+(\varphi_{i-1/2,j}^n \nabla_{1,h}^- u_{i,j}^{n+1}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^n \nabla_{2,h}^- u_{i,j}^{n+1})\right] u_{i,j}^{n+1} \le |\boldsymbol{u}^{n+1}|_{1,h}^2.$$

Proof. For $\mathbf{u}^n, \mathbf{u}^{n+1} \in \mathcal{H}_h$, applying Lemma 4.2.1, we obtain

$$\begin{aligned}
&\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \nabla_{1,h}^{+}(\varphi_{i-1/2,j}^{n}\,\nabla_{1,h}^{-}u_{i,j}^{n+1})u_{i,j}^{n+1} + \nabla_{2,h}^{+}(\varphi_{i,j-1/2}^{n}\,\nabla_{2,h}^{-}u_{i,j}^{n+1})u_{i,j}^{n+1} \\
&\quad = -\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Big[\varphi_{i-1/2,j}^{n}\,(\nabla_{1,h}^{-}u_{i,j}^{n+1})^2 + \varphi_{i,j-1/2}^{n}\,(\nabla_{2,h}^{-}u_{i,j}^{n+1})^2\Big] \\
&\quad = -\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Big[\left(\frac{3}{2}((u_{i,j}^{n})^2 + (u_{i-1,j}^{n})^2) - 1\right)(\nabla_{1,h}^{-}u_{i,j}^{n+1})^2 \\
&\qquad + \left(\frac{3}{2}((u_{i,j}^{n})^2 + (u_{i,j-1}^{n})^2) - 1\right)(\nabla_{2,h}^{-}u_{i,j}^{n+1})^2\Big] \\
&\quad = -\frac{3}{2}\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Bigg[\left((u_{i,j}^{n})^2 + (u_{i-1,j}^{n})^2\right)(\nabla_{1,h}^{-}u_{i,j}^{n+1})^2 \\
&\qquad + \left((u_{i,j}^{n})^2 + (u_{i,j-1}^{n})^2\right)(\nabla_{2,h}^{-}u_{i,j}^{n+1})^2\Bigg] \\
&\qquad + \Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(\nabla_{1,h}^{-}u_{i,j}^{n+1})^2 + (\nabla_{2,h}^{-}u_{i,j}^{n+1})^2\right] \\
&\quad \le \Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(\nabla_{1,h}^{-}u_{i,j}^{n+1})^2 + (\nabla_{2,h}^{-}u_{i,j}^{n+1})^2\right] \\
&\quad = |\mathbf{u}^{n+1}|_{1,h}^2.
\end{aligned}$$

□

The fourth order derivative is discretized using the central difference method and combining together with Eqs. (4.3.1) and (4.3.4), and the implicit one-level finite volume discretization of Eqs. (4.1.1)-(4.1.4) is given as follows:

$$\begin{aligned}
&\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\Delta t} - C_h(\mathbf{u}^{n+1}, \tilde{\mathbf{u}}^{n})_{i,j} + \varepsilon^2\Delta_h^2 u_{i,j}^{n+1} = \nabla_{1,h}^{+}(\varphi_{i-1/2,j}^{n}\nabla_{1,h}^{-}u_{i,j}^{n+1}) \\
&\qquad\qquad + \nabla_{2,h}^{+}(\varphi_{i,j-1/2}^{n}\nabla_{2,h}^{-}u_{i,j}^{n+1}), \qquad (4.3.5a)\\
&u_{i,j}^{n} = u_{i+N_1,j}^{n} = u_{i,j+N_2}^{n}, \qquad (4.3.5b)\\
&u_{i,j}^{0} = \frac{1}{\Delta x\Delta y}\iint_{k_{i,j}} u^0(x)\,dx\,dy. \qquad (4.3.5c)
\end{aligned}$$

Remark 4.3.1. *It is worth noting thanks to* (4.2.19)*, that* (4.3.5a)-(4.3.5c) *is a linear system of equations, while* (4.1.1)-(4.1.4) *is nonlinear.*

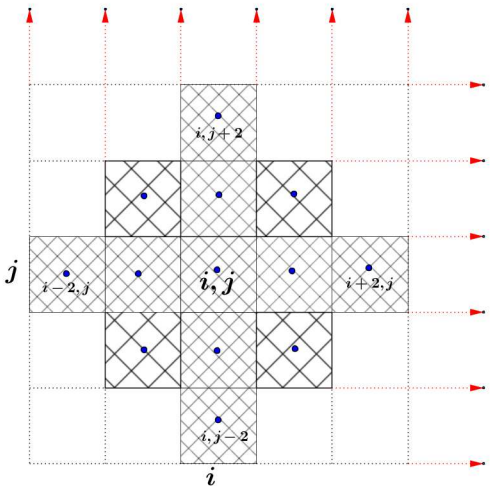


Figure 4.2: Finite volume discretization in 2D

Before discussing some qualitative properties of the solution of (4.3.5a)-(4.3.5c), we first address its feasibility.

Theorem 4.3.1. *If $\Delta t < 4\varepsilon^2$, then the approximate solution $\mathbf{u}^n$ of* (4.3.5a)-(4.3.5c) *is unique.*

Noting that equations (4.3.5a)-(4.3.5c) is a linear system in finite dimensional space, its existence is equivalent to uniqueness of solution (Gockenbach, 2010). Thus, we only show that the approximations $\mathbf{u}^1, \mathbf{u}^2, \ldots, \mathbf{u}^M$ satisfying (4.3.5a)-(4.3.5c) are unique.

Proof. For $n = 0, 1, 2, \ldots, M$, let $\mathbf{v}^n$ and $\mathbf{u}^n$ be two sequences of solutions of (4.3.5a)-(4.3.5c) with $\mathbf{v}_0 = \mathbf{u}_0$. Let $\mathbf{z}^n = \mathbf{u}^n - \mathbf{v}^n$ and clearly $\mathbf{z}^0 = 0$. We shall prove by induction that $\mathbf{z}^n = 0$ for all $n = 1, 2, \ldots, M$. We observe that $\mathbf{z}^{n+1}$ is a solution of

$$\begin{aligned}\frac{1}{\Delta t}(z_{i,j}^{n+1} - z_{i,j}^n) - (C_h(\mathbf{u}^{n+1}, \tilde{\mathbf{u}}^n))_{i,j} + (C_h(\mathbf{v}^{n+1}, \tilde{\mathbf{v}}^n))_{i,j} + \varepsilon^2 \Delta_h^2 z_{i,j}^{n+1} \\ = \nabla_{1,h}^+(\varphi_{i-1/2,j}^n \nabla_{1,h}^- u_{i,j}^{n+1}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^n \nabla_{2,h}^- u_{i,j}^{n+1}) \\ - \nabla_{1,h}^+(\psi_{i-1/2,j}^n \nabla_{1,h}^- v_{i,j}^{n+1}) + \nabla_{2,h}^+(\psi_{i,j-1/2}^n \nabla_{2,h}^- v_{i,j}^{n+1}),\end{aligned} \tag{4.3.6}$$

for $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$ and

$$\psi_{i-1/2,j}^n = \frac{f'(v_{i,j}^n) + f'(v_{i-1,j}^n)}{2} \text{ and } \psi_{i,j-1/2}^n = \frac{f'(v_{i,j}^n) + f'(v_{i,j-1}^n)}{2}.$$

By induction, we assume that $\mathbf{z}^n = 0$ and we want to show that $\mathbf{z}^{n+1} = 0$. It follows then that

$$z_{i,j}^{n+1} - \Delta t \left(C_h(\mathbf{z}^{n+1}, \tilde{\mathbf{u}}^n)\right)_{i,j} + \Delta t\, \varepsilon^2 \Delta_h^2 z_{i,j}^{n+1} = \Delta t \left[\nabla_{1,h}^+(\varphi_{i-1/2,j}^n \nabla_{1,h}^- z_{i,j}^{n+1}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^n \nabla_{2,h}^- z_{i,j}^{n+1})\right]. \tag{4.3.7}$$

Multiplying Eq. (4.3.7) by $\Delta t \Delta x \Delta y\, z_{i,j}^{n+1}$ and summing the resulting equalities for $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$, with help of (4.2.20), we obtain

$$\begin{aligned}\|\mathbf{z}^{n+1}\|_h^2 + \Delta t \varepsilon^2 |\mathbf{z}^{n+1}|_{2,h}^2 &= \Delta t \Delta x \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\nabla_{1,h}^+(\varphi_{i-1/2,j}^n \nabla_{1,h}^- z_{i,j}^{n+1}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^n \nabla_{2,h}^- z_{i,j}^{n+1})\right] z_{i,j}^{n+1} \\ &= -\Delta t \Delta x \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\varphi_{i-1/2,j}^n (\nabla_{1,h}^- z_{i,j}^{n+1})^2 + \varphi_{i,j-1/2}^n (\nabla_{2,h}^- z_{i,j}^{n+1})^2\right],\end{aligned}$$

which after the application of Lemma 4.3.1 gives

$$\|\mathbf{z}^{n+1}\|_h^2 + \Delta t \varepsilon^2 |\mathbf{z}^{n+1}|_{2,h}^2 \le \Delta t |\mathbf{z}^{n+1}|_{1,h}^2. \tag{4.3.8}$$

Applying Lemma 4.2.7 and Young's inequality, (4.3.8) implies that

$$(1 - \frac{\Delta t}{4\varepsilon^2})\|\mathbf{z}^{n+1}\|_h^2 \le 0. \tag{4.3.9}$$

For $\dfrac{\Delta t}{4\varepsilon^2} < 1$, we get

$$\|\mathbf{z}^{n+1}\|_h^2 \le 0.$$

Therefore, $\mathbf{z}^{n+1} = 0$. This completes the proof of uniqueness, hence the existence of solution. □

With the existence of solution being conditioned, it is quite clear that all possible results will be under at least the same condition, that is $\Delta t < 4\varepsilon^2$.

About the stability of the method (4.3.5a)-(4.3.5c), we have the following theorem.

Theorem 4.3.2. *The finite volume method defined by* (4.3.5a)-(4.3.5c)*, is conditionally stable in* $L^\infty(0, T; \mathcal{H}_h)$*, that is, for* $\Delta t \le \varepsilon^2$ *and* $1 \le n \le M$*,*

$$\|\boldsymbol{u}^n\|_h^2 \le 2^{\dfrac{2T}{\varepsilon^2}} \|\boldsymbol{u}^0\|_h^2.$$

Proof. By multiplying Eq. (4.3.5a) with $2\Delta x\Delta y\Delta t\ u_{i,j}^{n+1}$ and summing from $i=1$ to N_1 and $j=1$ to N_2, we obtain

$$\begin{aligned}\|\mathbf{u}^{n+1}\|_h^2-\|\mathbf{u}^n\|_h^2+\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2+2\Delta t\,\varepsilon^2(\Delta_h^2\mathbf{u}^{n+1},\mathbf{u}^{n+1})_h\\ =2\Delta x\Delta y\Delta t\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}\left[\nabla_{1,h}^+(\varphi_{i-1/2,j}^n\nabla_{1,h}^-u_{i,j}^{n+1})+\nabla_{2,h}^+(\varphi_{i,j-1/2}^n\nabla_{2,h}^-u_{i,j}^{n+1})\right]u_{i,j}^{n+1}.\end{aligned} \tag{4.3.10}$$

Using Lemmas 4.2.6 and 4.3.1 together with (4.3.10), we have

$$\|\mathbf{u}^{n+1}\|_h^2-\|\mathbf{u}^n\|_h^2+\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2+2\Delta t\varepsilon^2|\mathbf{u}^{n+1}|_{2,h}^2\le 2\Delta t|\mathbf{u}^{n+1}|_{1,h}^2. \tag{4.3.11}$$

Due to Lemma 4.2.7, Young's inequality and dropping the term $\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2$, (4.3.11) gives

$$\|\mathbf{u}^{n+1}\|_h^2-\|\mathbf{u}^n\|_h^2\le\frac{\Delta t}{2\varepsilon^2}\|\mathbf{u}^{n+1}\|_h^2,$$

which is re-written as follows

$$\left[1-\frac{\Delta t}{2\varepsilon^2}\right]\|\mathbf{u}^{n+1}\|_h^2\le\|\mathbf{u}^n\|_h^2. \tag{4.3.12}$$

Based on (1.1.10), for $\dfrac{\Delta t}{2\varepsilon^2}\le\dfrac{1}{2}$, (4.3.12) gives

$$\|\mathbf{u}^{n+1}\|_h^2\le 2^{\frac{2\Delta t}{\varepsilon^2}}\|\mathbf{u}^n\|_h^2.$$

By induction over n, we obtain

$$\|\mathbf{u}^n\|_h^2\le 2^{\frac{2n\,\Delta t}{\varepsilon^2}}\|\mathbf{u}^0\|_h^2\le 2^{\frac{2T}{\varepsilon^2}}\|\mathbf{u}^0\|_h^2.$$

Therefore, the proof is complete. □

Remark 4.3.2. *Starting with* (4.3.11) *and Lemma 4.2.7, one has*

$$\|\boldsymbol{u}^{n+1}\|_h^2-\|\boldsymbol{u}^n\|_h^2+\|\boldsymbol{u}^{n+1}-\boldsymbol{u}^n\|_h^2+2\Delta t\varepsilon^2\eta^2|\boldsymbol{u}^{n+1}|_{1,h}^2\le 2\Delta t|\boldsymbol{u}^{n+1}|_{1,h}^2. \tag{4.3.13}$$

Hence, we obtain

$$\|\boldsymbol{u}^{n+1}\|_h^2+2\Delta t(\varepsilon^2\eta^2-1)|\boldsymbol{u}^{n+1}|_{1,h}^2\le\|\boldsymbol{u}^n\|_h^2, \tag{4.3.14}$$

from which one deduces that if ε is big enough, then

$$\|\boldsymbol{u}^{n+1}\|_h^2 \leq \|\boldsymbol{u}^n\|_h^2 \leq \|\boldsymbol{u}^{n-1}\|_h^2 \leq \cdots \leq \|\boldsymbol{u}^0\|_h^2.$$

This alternative stability result requires both that $\Delta t < 4\varepsilon^2$ and the viscosity constant ε big enough.

Theorem 4.3.3. *Suppose that the solution $u(x,t)$ of Eqs.* (4.1.1)-(4.1.4) *is sufficiently smooth.*
Assume that $\Delta t < \min(4\varepsilon^2, c)$, with c given by (4.3.45)*, independent of Δx and Δy.*
Assume also that Δt, Δx and Δy satisfy the relation (4.3.47)*.*
Then, the solution of the finite volume discretization (4.3.5a)-(4.3.5c) *converges to the solution of the problems* (4.1.1)-(4.1.4) *in the discrete L^2 norm with rate of convergence $\mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2)$.*

Proof. For $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$, let

$$v_{i,j}^n = \frac{1}{\Delta x \Delta y} \iint_{k_{i,j}} u(x, y, t_n)\, dx\, dy,$$

be the cell average of the exact solution u of (4.1.1)-(4.1.4) at time t_n, for $0 \leq n \leq M$, on the cell $k_{i,j}$. Since u is smooth enough by assumption, (hence at least continuous on $[-L_1, L_1] \times [-L_2, L_2]$), we let

$$s = \max_{-L_1 \leq x \leq L_1, -L_2 \leq y \leq L_2, 0 \leq t \leq T} |u(x, y, t)|. \tag{4.3.15}$$

Also, the smoothness of u gives

$$v_{i,j}^n = u(x_i, y_j, t_n) + \mathcal{O}(\Delta x^2 + \Delta y^2), \quad 1 \leq i \leq N_1, 1 \leq j \leq N_2 \text{ and } 0 \leq n \leq M.$$

Making use of Taylor's expansion (see Appendix A), we obtain

$$\frac{v_{i,j}^{n+1} - v_{i,j}^n}{\Delta t} = u_t|_{i,j}^n + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2), \tag{4.3.16}$$

$$\Delta_h^2 v_{i,j}^{n+1} = \Delta^2 u|_{i,j}^n + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2), \tag{4.3.17}$$

$$\nabla_{1,h}^+(\psi_{i-\frac{1}{2},j}^n \nabla_{1,h}^- v_{i,j}^{n+1}) + \nabla_{2,h}^+(\psi_{i,j-\frac{1}{2}}^n \nabla_{2,h}^- v_{i,j}^{n+1}) = \Delta f(u)|_{i,j}^n + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2), \tag{4.3.18}$$

and

$$\begin{aligned}(C_h(\boldsymbol{v}^{n+1},\tilde{\boldsymbol{v}}^n))_{i,j} &= 3(\alpha_1+\alpha_2)(a_1+a_2+\cdots+a_{m_0})\,(u\,(\boldsymbol{\beta}\cdot\nabla)u)|_{i,j}^n \\ &+\frac{1}{2}(\alpha_1-\alpha_2)(a_1+a_2+\cdots+a_{m_0})(\Delta x\,(u\,u_{xx})|_{i,j}^n+\Delta y\,(u\,u_{yy})|_{i,j}^n) \\ &+(\alpha_1-\alpha_2)(a_1+a_2+\cdots+a_{m_0})(\Delta x\,(u_x^2)|_{i,j}^n+\Delta y\,(u_y^2)|_{i,j}^n) \\ &+\mathcal{O}(\Delta t+\Delta x^2+\Delta y^2). \end{aligned} \tag{4.3.19}$$

One observes that the numerical scheme is first order accurate when $\alpha_1 \neq \alpha_2$ and it is second order accurate in space if $\alpha_1 = \alpha_2$, from which (4.3.19) gives

$$C_h(\boldsymbol{v}^{n+1},\tilde{\boldsymbol{v}}^n)_{i,j} = (u(\boldsymbol{\beta}\cdot\nabla)u)|_{i,j}^n+\mathcal{O}(\Delta t+\Delta x^2+\Delta y^2). \tag{4.3.20}$$

In this study, we consider the case $\alpha_1 = \alpha_2$ to obtain second order method. Combining Eqs. (4.3.16)-(4.3.18) and (4.3.20), we obtain

$$\begin{cases} \dfrac{v_{i,j}^{n+1}-v_{i,j}^n}{\Delta t}-(C_h(\boldsymbol{v}^{n+1},\tilde{\boldsymbol{v}}^n))_{i,j}+\varepsilon^2\Delta_h^2 v_{i,j}^{n+1}=\nabla_{1,h}^+(\psi_{i-\frac{1}{2},j}^n\nabla_{1,h}^- v_{i,j}^{n+1})+\nabla_{2,h}^+(\psi_{i,j-\frac{1}{2}}^n\nabla_{2,h}^- v_{i,j}^{n+1})+r_{i,j}^n, \\ v_{i,j}^0=\dfrac{1}{\Delta x\Delta y}\iint_{k_{i,j}}u^0(x,y)\,dx\,dy, \end{cases} \tag{4.3.21}$$

where $r_{i,j}^n$ is the truncation error of the finite volume discretization (4.3.5a) for $0 \le n \le M-1$, $1 \le i \le N_1$ and $1 \le j \le N_2$. There exists a positive constant c_1 such that

$$\max_{i,j,n}|r_{i,j}^n| \le c_1(\Delta t+\Delta x^2+\Delta y^2),\ 0\le n\le M-1,\ 1\le i\le N_1,\ 1\le j\le N_2. \tag{4.3.22}$$

Let $\mathbf{e}^n = \boldsymbol{v}^n - \mathbf{u}^n, 0 \le n \le M$, where $u_{i,j}^n$ is the solution of Eqs. (4.3.5a)-(4.3.5c). Clearly $\mathbf{e}^n \in \mathcal{H}_h$ and $\mathbf{e}^0 = 0$.

Substituting $u_{i,j}^n = v_{i,j}^n - e_{i,j}^n$ into Eq. (4.3.5a), we obtain

$$\begin{aligned} &\frac{v_{i,j}^{n+1}-v_{i,j}^n}{\Delta t}-(C_h(\boldsymbol{v}^{n+1},\tilde{\boldsymbol{v}}^n))_{i,j}+\varepsilon^2\Delta_h^2 v_{i,j}^{n+1}-\left[\nabla_{1,h}^+(\psi_{i-\frac{1}{2},j}^n\nabla_{1,h}^- v_{i,j}^{n+1})+\nabla_{2,h}^+(\psi_{i,j-\frac{1}{2}}^n\nabla_{2,h}^- v_{i,j}^{n+1})\right] \\ &=\frac{e_{i,j}^{n+1}-e_{i,j}^n}{\Delta t}-(C_h(\mathbf{e}^{n+1},\tilde{\boldsymbol{v}}^n-\tilde{\mathbf{e}}^n))_{i,j}+\varepsilon^2\Delta_h^2 e_{i,j}^{n+1}-\left[\nabla_{1,h}^+(\varphi_{i-\frac{1}{2},j}^n\nabla_{1,h}^- e_{i,j}^{n+1})+\nabla_{2,h}^+(\varphi_{i,j-\frac{1}{2}}^n\nabla_{2,h}^- e_{i,j}^{n+1})\right] \\ &\quad-\nabla_{1,h}^+\left[\left(3(v_{i,j}^n\,e_{i,j}^n+v_{i-1,j}^n\,e_{i-1,j}^n)-\frac{3}{2}[(e_{i,j}^n)^2+(e_{i-1,j}^n)^2]\right)\nabla_{1,h}^- v_{i,j}^{n+1}\right] \\ &\quad-\nabla_{2,h}^+\left[\left(3(v_{i,j}^n\,e_{i,j}^n+v_{i,j-1}^n\,e_{i,j-1}^n)-\frac{3}{2}[(e_{i,j}^n)^2+(e_{i,j-1}^n)^2]\right)\nabla_{2,h}^- v_{i,j}^{n+1}\right] \end{aligned}$$

$$- (C_h(\boldsymbol{v}^{n+1}, \tilde{\mathbf{e}}^n))_{i,j}, \tag{4.3.23}$$

where

$$\psi^n_{i-1/2,j} = \frac{f'(v^n_{i,j}) + f'(v^n_{i-1,j})}{2} \text{ and } \psi^n_{i,j-1/2} = \frac{f'(v^n_{i,j}) + f'(v^n_{i,j-1})}{2}.$$

Using Eq. (4.3.21), Eq. (4.3.23) becomes

$$\begin{aligned}
\frac{e^{n+1}_{i,j} - e^n_{i,j}}{\Delta t} &- (C_h(\mathbf{e}^{n+1}, \tilde{\boldsymbol{v}}^n - \tilde{\mathbf{e}}^n))_{i,j} + \varepsilon^2 \Delta_h^2 e^{n+1}_{i,j} = \nabla^+_{1,h}(\varphi^n_{i-\frac{1}{2},j} \nabla^-_{1,h} e^{n+1}_{i,j}) + \nabla^+_{2,h}(\varphi^n_{i,j-\frac{1}{2}} \nabla^-_{2,h} e^{n+1}_{i,j}) \\
&+ \nabla^+_{1,h}\left[\left(3(v^n_{i,j}\, e^n_{i,j} + v^n_{i-1,j}\, e^n_{i-1,j}) - \frac{3}{2}[(e^n_{i,j})^2 + (e^n_{i-1,j})^2]\right) \nabla^-_{1,h} v^{n+1}_{i,j}\right] \\
&+ \nabla^+_{2,h}\left[\left(3(v^n_{i,j}\, e^n_{i,j} + v^n_{i,j-1}\, e^n_{i,j-1}) - \frac{3}{2}[(e^n_{i,j})^2 + (e^n_{i,j-1})^2]\right) \nabla^-_{2,h} v^{n+1}_{i,j}\right] \\
&+ (C_h(\boldsymbol{v}^{n+1}, \tilde{\mathbf{e}}^n))_{i,j} + r^n_{i,j}.
\end{aligned} \tag{4.3.24}$$

Multiplying Eq. (4.3.24) by $2\Delta t\, \Delta x\, \Delta y\, e^{n+1}_{i,j}$ and summing the corresponding equalities for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$, we obtain

$$\begin{aligned}
&2(\mathbf{e}^{n+1} - \mathbf{e}^n, \mathbf{e}^{n+1})_h + 2\Delta t \varepsilon^2 |\mathbf{e}^{n+1}|^2_{2,h} \\
&= 2\Delta t\, \Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\nabla^+_{1,h}(\varphi^n_{i-\frac{1}{2},j} \nabla^-_{1,h} e^{n+1}_{i,j}) + \nabla^+_{2,h}(\varphi^n_{i,j-\frac{1}{2}} \nabla^-_{2,h} e^{n+1}_{i,j})\right] e^{n+1}_{i,j} \\
&+ 2\Delta t\, \Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \nabla^+_{1,h}\left[\left(3(v^n_{i,j}\, e^n_{i,j} + v^n_{i-1,j}\, e^n_{i-1,j}) - \frac{3}{2}[(e^n_{i,j})^2 + (e^n_{i-1,j})^2]\right) \nabla^-_{1,h} v^{n+1}_{i,j}\right] e^{n+1}_{i,j} \\
&+ 2\Delta t\, \Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \nabla^+_{2,h}\left[\left(3(v^n_{i,j}\, e^n_{i,j} + v^n_{i,j-1}\, e^n_{i,j-1}) - \frac{3}{2}[(e^n_{i,j})^2 + (e^n_{i,j-1})^2]\right) \nabla^-_{2,h} v^{n+1}_{i,j}\right] e^{n+1}_{i,j} \\
&+ 2\Delta t (C_h(\boldsymbol{v}^{n+1}, \tilde{\mathbf{e}}^n), \mathbf{e}^{n+1})_h + 2\Delta t (\mathbf{r}^n, \mathbf{e}^{n+1})_h.
\end{aligned} \tag{4.3.25}$$

Using (4.2.15) and Lemmas 4.2.1 and 4.3.1, Eq. (4.3.25) yields

$$\begin{aligned}
\|\mathbf{e}^{n+1}\|^2_h - \|\mathbf{e}^n\|^2_h + \|\mathbf{e}^{n+1} - \mathbf{e}^n\|^2_h + 2\Delta t \varepsilon^2 |\mathbf{e}^{n+1}|^2_{2,h} &\le 2\Delta t |\mathbf{e}^{n+1}|^2_{1,h} \\
&- 6\Delta t\, \Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \Big[\left(v^n_{i,j}\, e^n_{i,j} + v^n_{i-1,j}\, e^n_{i-1,j}\right) \nabla^-_{1,h} v^{n+1}_{i,j} \nabla^-_{1,h} e^{n+1}_{i,j} \\
&+ \left(v^n_{i,j}\, e^n_{i,j} + v^n_{i,j-1}\, e^n_{i,j-1}\right) \nabla^-_{2,h} v^{n+1}_{i,j} \nabla^-_{2,h} e^{n+1}_{i,j}\Big]
\end{aligned}$$

$$+ 3\Delta t\,\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Big[[(e_{i,j}^n)^2 + (e_{i-1,j}^n)^2]\nabla_{1,h}^- v_{i,j}^{n+1}\nabla_{1,h}^- e_{i,j}^{n+1}$$
$$+ [(e_{i,j}^n)^2 + (e_{i,j-1}^n)^2]\nabla_{2,h}^- v_{i,j}^{n+1}\nabla_{2,h}^- e_{i,j}^{n+1}\Big]$$
$$+ 2\Delta t(C_h(\boldsymbol{v}^{n+1},\tilde{\mathbf{e}}^n),\mathbf{e}^{n+1})_h + 2\Delta t(\mathbf{r}^n,\mathbf{e}^{n+1})_h. \tag{4.3.26}$$

Since u is smooth then there exist constants c_2 and c_3 such that

$$\max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}} |\nabla_{1,h}^- v_{i,j}^n| \le c_2 \text{ and } \max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}} |\nabla_{2,h}^- v_{i,j}^n| \le c_3 \quad \text{for } n = 0,1,\cdots,M.$$

We estimate each of the terms on the right hand side of (4.3.26) as follows:

$$-2\Delta t\,\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} 3(v_{i,j}^n\, e_{i,j}^n + v_{i-1,j}^n\, e_{i-1,j}^n)\nabla_{1,h}^- v_{i,j}^{n+1}\nabla_{1,h}^- e_{i,j}^{n+1}$$
$$\le 6sc_2\,\Delta t\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}|e_{i,j}^n||\nabla_{1,h}^- e_{i,j}^{n+1}| + |e_{i-1,j}^n||\nabla_{1,h}^- e_{i,j}^{n+1}|\right)$$
$$\le 6s\,c_2\,\Delta t\left[2\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(e_{i,j}^n)^2\right)^{1/2}\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^- e_{i,j}^{n+1})^2\right)^{1/2}\right]$$
$$\le 12s\,c_2\,\Delta t\,\|\mathbf{e}^n\|_h\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^- e_{i,j}^{n+1})^2\right)^{1/2} \tag{4.3.27}$$

In a similar way, one obtains

$$-2\,\Delta t\,\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} 3(v_{i,j}^n\, e_{i,j}^n + v_{i,j-1}^n\, e_{i,j-1}^n)\nabla_{2,h}^- v_{i,j}^{n+1}\nabla_{2,h}^- e_{i,j}^{n+1}$$
$$\le 12s\,c_3\,\Delta t\,\|\mathbf{e}^n\|_h\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{2,h}^- e_{i,j}^{n+1})^2\right)^{1/2}. \tag{4.3.28}$$

Combining (4.3.27) and (4.3.28), we get

$$-2\Delta t\,\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Big[3(v_{i,j}^n\, e_{i,j}^n + v_{i-1,j}^n\, e_{i-1,j}^n)\nabla_{1,h}^- v_{i,j}^{n+1}\nabla_{1,h}^- e_{i,j}^{n+1}$$
$$+ 3(v_{i,j}^n\, e_{i,j}^n + v_{i,j-1}^n\, e_{i,j-1}^n)\nabla_{2,h}^- v_{i,j}^{n+1}\nabla_{2,h}^- e_{i,j}^{n+1}\Big]$$
$$\le 12sc_4\Delta t\|\mathbf{e}^n\|_h\left[\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^- e_{i,j}^{n+1})^2\right)^{1/2}\right.$$

$$
\begin{aligned}
&+\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{2,h}^{-}e_{i,j}^{n+1})^2\right)^{1/2}\Bigg]\\
&\leq 12\sqrt{2}\,s\,c_4\Delta t\|\mathbf{e}^n\|_h|\mathbf{e}^{n+1}|_{1,h}\\
&\leq 6\;\Delta t\left(\frac{2s^2\,c_4^2}{\delta_1}\|\mathbf{e}^n\|_h^2+\delta_1\,|\mathbf{e}^{n+1}|_{1,h}^2\right), \qquad (4.3.29)
\end{aligned}
$$

where $c_4=\max\{c_2,c_3\}$.

$$
\begin{aligned}
&3\Delta t\,\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(e_{i,j}^n)^2+(e_{i-1,j}^n)^2\right]\nabla_{1,h}^{-}v_{i,j}^{n+1}\nabla_{1,h}^{-}e_{i,j}^{n+1}\\
&\leq 3c_2\,\Delta t\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(e_{i,j}^n)^2+(e_{i-1,j}^n)^2\right]|\nabla_{1,h}^{-}e_{i,j}^{n+1}|\right)\\
&\leq 3c_2\,\Delta t\,\|\mathbf{e}^n\|_\infty\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[|e_{i,j}^n|+|e_{i-1,j}^n|\right]|\nabla_{1,h}^{-}e_{i,j}^{n+1}|\right)\\
&\leq 6c_2\,\Delta t\,\|\mathbf{e}^n\|_{\infty,h}\left[\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(e_{i,j}^n)^2\right)^{1/2}\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^{-}e_{i,j}^{n+1})^2\right)^{1/2}\right]\\
&=6c_2\,\Delta t\,\|\mathbf{e}^n\|_{\infty,h}\|\mathbf{e}^n\|_h\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^{-}e_{i,j}^{n+1})^2\right)^{1/2}. \qquad (4.3.30)
\end{aligned}
$$

Similarly, we get

$$
\begin{aligned}
&3\Delta t\,\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left[(e_{i,j}^n)^2+(e_{i,j-1}^n)^2\right]\nabla_{2,h}^{-}v_{i,j}^{n+1}\nabla_{2,h}^{-}e_{i,j}^{n+1}\\
&\leq 6c_3\,\Delta t\,\|\mathbf{e}^n\|_{\infty,h}\|\mathbf{e}^n\|_h\left(\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{2,h}^{-}e_{i,j}^{n+1})^2\right)^{1/2}. \qquad (4.3.31)
\end{aligned}
$$

Adding (4.3.30) and (4.3.31), one obtains

$$
\begin{aligned}
&3\Delta t\,\Delta x\,\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\Bigg[\left((e_{i,j}^n)^2+(e_{i-1,j}^n)^2\right)\nabla_{1,h}^{-}v_{i,j}^{n+1}\nabla_{1,h}^{-}e_{i,j}^{n+1}\\
&\qquad+\left((e_{i,j}^n)^2+(e_{i,j-1}^n)^2\right)\nabla_{2,h}^{-}v_{i,j}^{n+1}\nabla_{2,h}^{-}e_{i,j}^{n+1}\Bigg]\\
&\leq 6\sqrt{2}\,c_4\,\Delta t\,\|\mathbf{e}^n\|_{\infty,h}\,\|\mathbf{e}^n\|_h\,|\mathbf{e}^{n+1}|_{1,h}\\
&\leq 3\Delta t\left(\frac{2\,c_4^2}{\delta_2}\,\|\mathbf{e}^n\|_h^2+\delta_2\|\mathbf{e}^n\|_{\infty,h}^2\,|\mathbf{e}^{n+1}|_{1,h}^2\right). \qquad (4.3.32)
\end{aligned}
$$

From the definition of the bilinear map C_h, we have

$$
\begin{aligned}
&(C_h(\boldsymbol{\upsilon}^{n+1}, \tilde{\mathbf{e}}^n), \mathbf{e}^{n+1})_h \\
&= \Delta x\, \Delta y\, \alpha \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\upsilon_{i,j}^{n+1}(\boldsymbol{\beta}\cdot\nabla_h^+)\tilde{e}_{i,j}^n + \tilde{e}_{i,j}^n(\boldsymbol{\beta}\cdot\nabla_h^-)\upsilon_{i,j}^{n+1} + \tilde{e}_{i+1,j}^n\nabla_{1,h}^+\upsilon_{i,j}^{n+1} + \tilde{e}_{i,j+1}^n\nabla_{2,h}^+\upsilon_{i,j}^{n+1}\right] e_{i,j}^{n+1} \\
&+ \Delta x\, \Delta y\, \alpha \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\upsilon_{i,j}^{n+1}(\boldsymbol{\beta}\cdot\nabla_h^-)\tilde{e}_{i,j}^n + \tilde{e}_{i,j}^n(\boldsymbol{\beta}\cdot\nabla_h^+)\upsilon_{i,j}^{n+1} + \tilde{e}_{i-1,j}^n\nabla_{1,h}^-\upsilon_{i,j}^{n+1} + \tilde{e}_{i,j-1}^n\nabla_{2,h}^-\upsilon_{i,j}^{n+1}\right] e_{i,j}^{n+1}.
\end{aligned}
\tag{4.3.33}
$$

We now estimate each of the terms on the right hand side of (4.3.33).

$$
\begin{aligned}
&\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \upsilon_{i,j}^{n+1}(\boldsymbol{\beta}\cdot\nabla_h^-\tilde{e}_{i,j}^n)\, e_{i,j}^{n+1} = -\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \tilde{e}_{i,j}^n(\boldsymbol{\beta}\cdot\nabla_h^+)(\upsilon_{i,j}^{n+1}\, e_{i,j}^{n+1}) \\
&= -\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \tilde{e}_{i,j}^n \left[(\nabla_{1,h}^+\upsilon_{i,j}^{n+1})\, e_{i+1,j}^{n+1} + \upsilon_{i,j}^{n+1}\,\nabla_{1,h}^+ e_{i,j}^{n+1} + (\nabla_{2,h}^+\upsilon_{i,j}^{n+1})\, e_{i,j+1}^{n+1} + \upsilon_{i,j}^{n+1}\,\nabla_{2,h}^+ e_{i,j}^{n+1}\right] \\
&\leq (c_2+c_3)\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h \\
&\quad + s\|\tilde{\mathbf{e}}^n\|_h \left[\left(\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{1,h}^+ e_{i,j}^{n+1})^2\right)^{1/2} + \left(\Delta x\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\nabla_{2,h}^+ e_{i,j}^{n+1})^2\right)^{1/2}\right] \\
&\leq (c_2+c_3)\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h + \sqrt{2}s\,\|\tilde{\mathbf{e}}^n\|_h\,|\mathbf{e}^{n+1}|_{1,h} \\
&= c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h + \sqrt{2}s\,\|\tilde{\mathbf{e}}^n\|_h\,|\mathbf{e}^{n+1}|_{1,h}
\end{aligned}
\tag{4.3.34}
$$

where $c_5 = c_2 + c_3$.

$$
\begin{aligned}
\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \tilde{e}_{i,j}^n(\boldsymbol{\beta}\cdot\nabla_h^-\upsilon_{i,j}^{n+1})\, e_{i,j}^{n+1} &\leq \max_{\substack{1\leq i\leq N_1\\ 1\leq j\leq N_2}} |(\boldsymbol{\beta}\cdot\nabla_h^-)\upsilon_{i,j}^{n+1}| \left[\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |\tilde{e}_{i,j}^n|\,|e_{i,j}^{n+1}|\right] \\
&\leq \left[\max_{\substack{1\leq i\leq N_1\\ 1\leq j\leq N_2}} |\nabla_{1,h}^-\upsilon_{i,j}^{n+1}| + \max_{\substack{1\leq i\leq N_1\\ 1\leq j\leq N_2}} |\nabla_{2,h}^-\upsilon_{i,j}^{n+1}|\right] \left[\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |\tilde{e}_{i,j}^n|\,|e_{i,j}^{n+1}|\right] \\
&\leq c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h.
\end{aligned}
\tag{4.3.35}
$$

$$
\begin{aligned}
&\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\tilde{e}_{i+1,j}^n\nabla_{1,h}^+\upsilon_{i,j}^{n+1} + \tilde{e}_{i,j+1}^n\nabla_{2,h}^+\upsilon_{i,j}^{n+1}\right] e_{i,j}^{n+1} \\
&\qquad \leq \max_{\substack{1\leq i\leq N_1\\ 1\leq j\leq N_2}} |\nabla_{1,h}^+\upsilon_{i,j}^{n+1}| \left[\Delta x\, \Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |\tilde{e}_{i+1,j}^n|\,|e_{i,j}^{n+1}|\right]
\end{aligned}
$$

$$+ \max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}} |\nabla^+_{2,h} v^{n+1}_{i,j}| \left[\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} |\tilde{e}^n_{i,j+1}|\,|e^{n+1}_{i,j}|\right]$$
$$\le c_2\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h + c_3\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h = c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h. \tag{4.3.36}$$

In a similar fashion, we obtain

$$\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} v^{n+1}_{i,j}(\boldsymbol{\beta}\cdot\nabla^-_h \tilde{e}^n_{i,j})\, e^{n+1}_{i,j} \le \|\tilde{\mathbf{e}}^n\|_h \left[c_5\|\mathbf{e}^{n+1}\|_h + \sqrt{2}s\,|\mathbf{e}^{n+1}|_{1,h}\right]. \tag{4.3.37}$$

$$\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \tilde{e}^n_{i,j}(\boldsymbol{\beta}\cdot\nabla^+_h v^{n+1}_{i,j})\, e^{n+1}_{i,j} \le c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h. \tag{4.3.38}$$

$$\Delta x\,\Delta y \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \left[\tilde{e}^n_{i-1,j}\nabla^-_{1,h} v^{n+1}_{i,j} + \tilde{e}^n_{i,j-1}\nabla^-_{2,h} v^{n+1}_{i,j}\right] e^{n+1}_{i,j} \le c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h. \tag{4.3.39}$$

Combining the inequalities from (4.3.34)-(4.3.39), we have

$$\begin{aligned} 2\Delta t(C_h(\boldsymbol{v}^{n+1},\tilde{\mathbf{e}}^n),\mathbf{e}^{n+1})_h &\le 4\Delta t|\alpha|\left[s\sqrt{2}\|\tilde{\mathbf{e}}^n\|_h|\mathbf{e}^{n+1}|_{1,h} + 3c_5\|\tilde{\mathbf{e}}^n\|_h\|\mathbf{e}^{n+1}\|_h\right] \\ &\le \Delta t\left[\frac{8s^2\alpha^2}{\delta_4}\|\tilde{\mathbf{e}}^n\|_h^2 + \delta_4|\mathbf{e}^{n+1}|^2_{1,h}\right] \\ &\quad + \Delta t\left[\frac{36c_5^2\alpha^2}{\delta_5}\|\tilde{\mathbf{e}}^n\|_h^2 + \delta_5\|\mathbf{e}^{n+1}\|_h^2\right]. \end{aligned} \tag{4.3.40}$$

Finally, the last term of (4.3.26) is estimated as given bellow.

$$\begin{aligned} (\mathbf{r}^n,\mathbf{e}^{n+1})_h &\le \Delta x\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} r^n_{i,j}e^{n+1}_{i,j} \\ &\le N_1N_2\Delta x\Delta y \max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}}|r^n_{i,j}|\left(\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(e^{n+1}_{i,j})^2\right)^{1/2} \\ &\le \frac{16L_1^2L_2^2}{2\delta_3}\left(\max_{\substack{1\le i\le N_1\\ 1\le j\le N_2}}|r^n_{i,j}|\right)^2 + \frac{\delta_3}{2}\|\mathbf{e}^{n+1}\|_h^2 \\ &\le \frac{16L_1^2L_2^2c_1^2}{2\delta_3}\left(\Delta t+\Delta x^2+\Delta y^2\right)^2 + \frac{\delta_3}{2}\|\mathbf{e}^{n+1}\|_h^2. \end{aligned} \tag{4.3.41}$$

Thus, from (4.3.29), (4.3.32), (4.3.40) and (4.3.41), we get

$$\|\mathbf{e}^{n+1}\|_h^2 - \|\mathbf{e}^n\|_h^2 + 2\Delta t\varepsilon^2|\mathbf{e}^{n+1}|^2_{2,h} \le 2\Delta t|\mathbf{e}^{n+1}|^2_{1,h} + 6\,\Delta t\left(\frac{2s^2\,c_4^2}{\delta_1}\|\mathbf{e}^n\|_h^2 + \delta_1\,|\mathbf{e}^{n+1}|^2_{1,h}\right)$$
$$+ 3\Delta t\left(\frac{2\,c_4^2}{\delta_2}\,\|\mathbf{e}^n\|_h^2 + \delta_2\|\mathbf{e}^n\|^2_{\infty,h}\,|\mathbf{e}^{n+1}|^2_{1,h}\right)$$

$$+\Delta t\left[\frac{8s^2\alpha^2}{\delta_4}\|\tilde{\mathbf{e}}^n\|_h^2+\delta_4|\mathbf{e}^{n+1}|_{1,h}^2\right]+\Delta t\left[\frac{36c_5^2\alpha^2}{\delta_5}\|\tilde{\mathbf{e}}^n\|_h^2+\delta_5\|\mathbf{e}^{n+1}\|_h^2\right]$$
$$+\frac{16\Delta t\,L_1^2L_2^2c_1^2}{\delta_3}(\Delta t+\Delta x^2+\Delta y^2)^2+\Delta t\,\delta_3\|\mathbf{e}^{n+1}\|_h^2,$$

which on rearrangement gives

$$[1-\Delta t(\delta_3+\delta_5)]\|\mathbf{e}^{n+1}\|_h^2+2\Delta t\varepsilon^2|\mathbf{e}^{n+1}|_{2,h}^2\leq\Delta t[2+6\delta_1+\delta_4]|\mathbf{e}^{n+1}|_{1,h}^2$$
$$+\left[1+6\;\Delta t\left(\frac{2s^2\,c_4^2}{\delta_1}+\frac{c_4^2}{\delta_2}\right)\right]\|\mathbf{e}^n\|_h^2$$
$$+3\Delta t\,\delta_2\|\mathbf{e}^n\|_{\infty,h}^2\;|\mathbf{e}^{n+1}|_{1,h}^2+4\Delta t\,\alpha^2\left[\frac{2s^2}{\delta_4}+\frac{9c_5^2}{\delta_5}\right]\|\tilde{\mathbf{e}}^n\|_h^2$$
$$+\frac{16\Delta t\,L_1^2L_2^2c_1^2}{\delta_3}(\Delta t+\Delta x^2+\Delta y^2)^2. \tag{4.3.42}$$

Using (4.2.13), Lemma 4.2.7 and Young's inequality, (4.3.42) becomes

$$[1-\Delta t(\delta_3+\delta_5)]\|\mathbf{e}^{n+1}\|_h^2+2\Delta t\varepsilon^2|\mathbf{e}^{n+1}|_{2,h}^2\leq\frac{\Delta t(2+6\delta_1+\delta_4)^2}{4\varepsilon^2}\|\mathbf{e}^{n+1}\|_h^2+\Delta t\varepsilon^2|\mathbf{e}^{n+1}|_{2,h}^2$$
$$+\left[1+6\;\Delta t\left(\frac{2s^2\,c_4^2}{\delta_1}+\frac{c_4^2}{\delta_2}\right)\right]\|\mathbf{e}^n\|_h^2$$
$$+\frac{3\Delta t\,\delta_2}{\Delta x\Delta y}\|\mathbf{e}^n\|_h^2\;|\mathbf{e}^{n+1}|_{1,h}^2+4\Delta t\,\alpha^2\left[\frac{2s^2}{\delta_4}+\frac{9c_5^2}{\delta_5}\right]\|\tilde{\mathbf{e}}^n\|_h^2$$
$$+\frac{16\Delta t\,L_1^2L_2^2c_1^2}{\delta_3}(\Delta t+\Delta x^2+\Delta y^2)^2. \tag{4.3.43}$$

Using (4.2.25), (4.3.43) gives

$$[1-\Delta t\,c]\|\mathbf{e}^{n+1}\|_h^2+\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2\leq(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8\|\tilde{\mathbf{e}}^n\|_h^2$$
$$+\frac{12\Delta t\,\delta_2}{\Delta x\Delta y}\|\mathbf{e}^n\|_h^2\;|\mathbf{e}^{n+1}|_{1,h}^2+\Delta tc_7(\Delta t+\Delta x^2+\Delta y^2)^2, \tag{4.3.44}$$

where

$$c_7=\frac{16\,L_1^2L_2^2c_1^2}{\delta_3},\quad c=\delta_3+\delta_5+\frac{(2+6\delta_1+\delta_4)^2}{4\varepsilon^2}$$
$$c_9=6\;\left(\frac{2s^2\,c_4^2}{\delta_1}+\frac{c_4^2}{\delta_2}\right),\quad c_8=4\alpha^2\left[\frac{2s^2}{\delta_4}+\frac{9c_5^2}{\delta_5}\right].$$

For

$$\Delta t\leq\frac{1}{2c}\equiv c \tag{4.3.45}$$

then it follows from (4.3.44) that

$$\|\mathbf{e}^{n+1}\|_h^2+\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2\leq\;4^{\Delta tc}\Big[(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8\|\tilde{\mathbf{e}}^n\|_h^2+\frac{3\Delta t\,\delta_2}{\Delta x\Delta y}\|\mathbf{e}^n\|_h^2\;|\mathbf{e}^{n+1}|_{1,h}^2$$
$$+\Delta tc_7(\Delta t+\Delta x^2+\Delta y^2)^2\Big]. \tag{4.3.46}$$

Now for

$$\frac{3\delta_2\, c_7}{\Delta x\Delta y}\left[(\Delta t+\Delta x^2+\Delta y^2)^2\right] \leq \frac{c_{10}}{2}\varepsilon^2\eta^2 4^{-Tc}\exp\Big(-Tc_{10}\Big), \tag{4.3.47}$$

where $c_{10} = c_9 + m_0\, A\, c_8$ and $A = \sum_{i=1}^{m_0} |a_i|^2$, we prove by inductive method that

$$\|\mathbf{e}^{n+1}\|_h^2 + \frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2 \leq 4^{\Delta t\, c}\Big[(1+\Delta t c_9)\,\|\mathbf{e}^n\|_h^2 + \Delta t c_8\|\tilde{\mathbf{e}}^n\|_h^2 + \Delta t c_7(\Delta t+\Delta x^2+\Delta y^2)^2\Big]. \tag{4.3.48}$$

For $n=0$ from (4.3.46), one obtains

$$\|\mathbf{e}^1\|_h^2 + \Delta t\varepsilon^2\eta^2|\mathbf{e}^1|_{1,h}^2 \leq 4^{\Delta t\, c}\left[(1+\Delta t(c_9+c_8))\,\|\mathbf{e}^0\|_h^2 + \frac{3\Delta t\,\delta_2}{\Delta x\Delta y}\|\mathbf{e}^0\|_h^2\ |\mathbf{e}^1|_{1,h}^2\right] + \Delta t c_7(4^{\Delta t\, c})\left[(\Delta t+\Delta x^2+\Delta y^2)^2\right],$$

and hence,

$$\|\mathbf{e}^1\|_h^2 + \frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^1|_{1,h}^2 \leq 4^{\Delta t\, c}\left[(1+\Delta t(c_9+c_8))\,\|\mathbf{e}^0\|_h^2 + \Delta t c_7(\Delta t+\Delta x^2+\Delta y^2)^2\right],$$

which is (4.3.48) for $n=0$. Now suppose that (4.3.48) is true up to the order $n-1$. Thus, for $s=0,1,\ldots,n-1$,

$$\|\mathbf{e}^{s+1}\|_h^2 + \frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{s+1}|_{1,h}^2 \leq 4^{\Delta t\, c}\Big[(1+\Delta t c_9)\,\|\mathbf{e}^s\|_h^2 + \Delta t c_8\|\tilde{\mathbf{e}}^s\|_h^2 + \Delta t c_7(\Delta t+\Delta x^2+\Delta y^2)^2\Big]. \tag{4.3.49}$$

It is now remaining to treat the term $\|\tilde{\mathbf{e}}^s\|_h^2$. For $s < m_0 - 1$, it is clear from (4.3.3) that $\|\tilde{\mathbf{e}}^s\|_h = \|\mathbf{e}^s\|_h$. Thus, (4.3.49) becomes

$$\|\mathbf{e}^{s+1}\|_h^2 + \frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{s+1}|_{1,h}^2 \leq 4^{\Delta t\, c}\left[(1+\Delta t(c_9+c_8))\,\|\mathbf{e}^s\|_h^2 + \Delta t c_7(\Delta t+\Delta x^2+\Delta y^2)^2\right]. \tag{4.3.50}$$

It follows from (4.3.50) that

$$\|\mathbf{e}^{s+1}\|_h^2 \leq 4^{\Delta t\, c}\left[(1+\Delta t(c_9+c_8))\,\|\mathbf{e}^s\|_h^2 + \Delta t c_7(\Delta t+\Delta x^2+\Delta y^2)^2\right].$$

After $s+1$ iterations, we get

$$\|\mathbf{e}^{s+1}\|_h^2 \leq 4^{\Delta t(s+1)\, c}\Big[(1+\Delta t(c_9+c_8))^{s+1}\,\|\mathbf{e}^0\|_h^2$$

$$+ \Delta t c_7(\Delta t + \Delta x^2 + \Delta y^2)^2 \sum_{j=0}^{s}(1 + \Delta t(c_9 + c_8))^j\Big]$$
$$= 4^{\Delta t(s+1)\,c}\Big[\Delta t c_7(\Delta t + \Delta x^2 + \Delta y^2)^2 \sum_{j=0}^{s}(1 + \Delta t(c_9 + c_8))^j\Big]. \qquad (4.3.51)$$

For the case $s \geq m_0 - 1$ and $m_0 > 1$, it follows from (4.3.2) that

$$\|\tilde{\mathbf{e}}^s\|_h \leq \sum_{i=1}^{m_0} |a_i| \|\tilde{\mathbf{e}}^{s-i+1}\|_h,$$

which by Cauchy-Schwartz's inequality gives

$$\|\tilde{\mathbf{e}}^s\|_h^2 \leq A\Big[\|\mathbf{e}^s\|_h^2 + \|\mathbf{e}^{s-1}\|_h^2 + \cdots + \|\mathbf{e}^{s-m_0+1}\|_h^2\Big].$$

Hence, (4.3.49) gives

$$\|\mathbf{e}^{s+1}\|_h^2 \leq 4^{\Delta t\, c}\left[(1 + \Delta t c_9)\,\|\mathbf{e}^s\|_h^2 + \Delta t c_8 A\Big(\|\mathbf{e}^s\|_h^2 + \|\mathbf{e}^{s-1}\|_h^2 + \cdots + \|\mathbf{e}^{s-m_0+1}\|_h^2\Big)\right]$$
$$+ 4^{\Delta t\, c}\left[\Delta t c_7(\Delta t + \Delta x^2 + \Delta y^2)^2\right].$$

It follows easily that

$$\max\Big\{\|\mathbf{e}^{s+1}\|_h^2, \|\mathbf{e}^s\|_h^2, \ldots, \|\mathbf{e}^{s-m_0+2}\|_h^2\Big\} \leq 4^{\Delta t c}\Bigg[(1 + \Delta t\, c_{10})\, \max\Big\{\|\mathbf{e}^s\|_h^2, \|\mathbf{e}^{s-1}\|_h^2,$$
$$\ldots, \|\mathbf{e}^{s-m_0+1}\|_h^2\Big\} + \Delta t\, c_7\,(\Delta t + \Delta x^2 + \Delta y^2)^2\Bigg].$$

which after $s - m_0 + 2$ iterations gives

$$\max\Big\{\|\mathbf{e}^{s+1}\|_h^2, \|\mathbf{e}^s\|_h^2, \ldots, \|\mathbf{e}^{s-m_0+2}\|_h^2\Big\}$$
$$\leq 4^{(s-m_0+2)\Delta t c}\Bigg[(1 + \Delta t\, c_{10})^{s-m_0+2}\, \max\Big\{\|\mathbf{e}^{m_0-1}\|_h^2, \|\mathbf{e}^{m_0-2}\|_h^2, \ldots, \|\mathbf{e}^0\|_h^2\Big\}$$
$$+ \Delta t\, c_7(\Delta t + \Delta x^2 + \Delta y^2)^2 \sum_{j=0}^{s-m_0+1} (1 + \Delta t c_{10})^j\Bigg]. \qquad (4.3.52)$$

Combining (4.3.51) and (4.3.52), we get

$$\|\mathbf{e}^{s+1}\|_h^2 \leq 4^{(s+1)\Delta t c}\Big[\Delta t\, c_7(\Delta t + \Delta x^2 + \Delta y^2)^2 \sum_{j=0}^{s}(1 + \Delta t c_{10})^j\Big]$$
$$= 4^{(s+1)\Delta t c}\Big[\Delta t\, c_7(\Delta t + \Delta x^2 + \Delta y^2)^2 \left(\frac{(1 + \Delta t c_{10})^{s+1} - 1}{\Delta t c_{10}}\right)\Big]$$

$$\leq 4^{(s+1)\Delta tc}\exp((s+1)\Delta t\, c_{10})\Big[\frac{c_7}{c_{10}}(\Delta t+\Delta x^2+\Delta y^2)^2\Big] \tag{4.3.53}$$

and

$$\|\tilde{\mathbf{e}}^{s+1}\|_h^2 \leq m_0\, A\, 4^{(s+1)\Delta tc}\exp((s+1)\Delta t\, c_{10})\Big[\frac{c_7}{c_{10}}(\Delta t+\Delta x^2+\Delta y^2)^2\Big], \tag{4.3.54}$$

for $s=0,1,\ldots,n-1$.

Going back to (4.3.46), we have

$$\begin{aligned}\|\mathbf{e}^{n+1}\|_h^2+\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2 &\leq 4^{\Delta tc}\Big[(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8\|\tilde{\mathbf{e}}^n\|_h^2+\frac{3\Delta t\,\delta_2}{\Delta x\Delta y}\,\|\mathbf{e}^n\|_h^2\;|\mathbf{e}^{n+1}|_{1,h}^2\Big]\\ &\quad+\Delta tc_7\,4^{\Delta tc}(\Delta t+\Delta x^2+\Delta y^2)^2\\ &\leq 4^{\Delta tc}\Big[(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8|\tilde{\mathbf{e}}^n|_h^2+\Delta tc_7(\Delta t+\Delta x^2+\Delta y^2)^2\Big]\\ &\quad+\frac{3\Delta t\,\delta_2}{\Delta x\Delta y}\,4^{n\Delta tc}\exp(n\Delta t\, c_{10})\Big[\frac{c_7}{c_{10}}(\Delta t+\Delta x^2+\Delta y^2)^2\Big]\,|\mathbf{e}^{n+1}|_{1,h}^2,\end{aligned}$$

which by (4.3.47) gives

$$\begin{aligned}\|\mathbf{e}^{n+1}\|_h^2+\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2 &\leq 4^{\Delta tc}\Big[(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8\|\tilde{\mathbf{e}}^n\|_h^2+\Delta tc_7(\Delta t+\Delta x^2+\Delta y^2)^2\Big]\\ &\quad+\frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}\|\mathbf{e}^{n+1}\|_h^2+\frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+1}|_{1,h}^2 &\leq 4^{\Delta tc}\Big[(1+\Delta tc_9)\,\|\mathbf{e}^n\|_h^2+\Delta tc_8\|\tilde{\mathbf{e}}^n\|_h^2\\ &\quad+\Delta tc_7\,(\Delta t+\Delta x^2+\Delta y^2)^2\Big].\end{aligned} \tag{4.3.55}$$

Using (4.3.53) and (4.3.54), (4.3.55) gives

$$\begin{aligned}\|\mathbf{e}^n\|_h^2 &\leq 4^{n\Delta tc}\exp\left(n\,\Delta tc_{10}\right)\,\Big[\frac{c_7}{c_{10}}(\Delta t+\Delta x^2+\Delta y^2)^2\Big]\\ &\leq 4^{Tc}\exp(T\,c_{10})\,\Big[\frac{c_7}{c_{10}}(\Delta t+\Delta x^2+\Delta y^2)^2\Big],\end{aligned} \tag{4.3.56}$$

for $n=1,2,\ldots,M$. Thus, it follows from (4.3.56) that

$$\|\mathbf{e}^n\|_h \leq C(\Delta t+\Delta x^2+\Delta y^2),$$

for constant C independent of Δt, Δx and Δy. This completes the proof. □

Remark 4.3.3. *The assumption $\Delta t < 4\varepsilon^2$ in Theorem 4.3.3 is to ensure the existence of solution for* (4.3.5a)-(4.3.5c)*.*
It is worth noting that the convergence result Theorem 4.3.3 is conducted for $\alpha_1 = \alpha_2$. The case when $\alpha_1 \neq \alpha_2$ is treated in the same way but the rate of convergence will change.

4.3.2 Explicit one-level finite volume method

In this section, we approximate the solution of Eqs. (4.1.1)-(4.1.4) using an explicit finite volume method:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\Delta t} - (C_h(\mathbf{u}^n, \mathbf{u}^n))_{i,j} + \varepsilon^2 \Delta_h^2 u_{i,j}^n = \nabla_{1,h}^{+}(\varphi_{i-1/2,j}^n \nabla_{1,h}^{-} u_{i,j}^n) + \nabla_{2,h}^{+}(\varphi_{i,j-1/2}^n \nabla_{2,h}^{-} u_{i,j}^n), \tag{4.3.57a}$$

$$u_{i,j}^n = u_{i+N_1,j}^n = u_{i,j+N_2}^n, \tag{4.3.57b}$$

$$u_{i,j}^0 = \frac{1}{\Delta x \Delta y} \iint\limits_{k_{i,j}} u^0(x,y)dxdy, \tag{4.3.57c}$$

for $1 \leq n \leq M-1$ and $1 \leq i \leq N_1, 1 \leq j \leq N_2$. (4.3.57) is explicit, hence the solution $\mathbf{u}^n$ is computed at each time step. One important feature of this scheme is stated in the following.

Theorem 4.3.4. *We assume that the following are satisfied for some $\delta, 0 < \delta < 1$:*

$$\Delta t \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right)^2 \leq \frac{1-\delta}{64\varepsilon^2}, \tag{4.3.58}$$

$$16\Delta t \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right) \leq \varepsilon^2 \delta \eta^2 (1-\delta), \tag{4.3.59}$$

$$\frac{72\Delta t}{\Delta x \Delta y}\left((|\alpha_1| + |\alpha_2|)^2 + \frac{4}{\Delta x \Delta y}\left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right)\|\boldsymbol{u}^0\|_h^2\right)\|\boldsymbol{u}^0\|_h^2 \leq \varepsilon^2\delta^2\eta^2 \exp\left(\frac{-2T}{\varepsilon^2}\right), \tag{4.3.60}$$

where α_1 and α_2 are the constants in Eq. (4.2.19)*. Then the finite volume method defined by* (4.3.57) *is $L^\infty(0,T;\mathcal{H}_h)$ stable in the following sense:*

$$\|\boldsymbol{u}^n\|^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^{n-1}\|^2 \leq \cdots \leq \exp\left(\frac{n\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2 \leq \exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2, n = 1, 2 \ldots, M,$$

$$\frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\sum_{n=0}^{M-1}\exp\left(\frac{(M-n)\Delta t}{\varepsilon^2}\right)|\boldsymbol{u}^n|_{1,h}^2 \le \exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2.$$

Proof. To prove this assertion we use the approach of Temam (1979). Multiplying Eq. (4.3.57a) by $2\Delta t\Delta x\Delta y\, u_{i,j}^n$ and summing the equalities for $i = 1,\ldots,N_1$ and $j = 1,\ldots,N_2$, together with (4.2.16), (1.1.12) and Lemma 4.2.7, we arrive at

$$\|\mathbf{u}^{n+1}\|_h^2 - \|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2 + \Delta t\varepsilon^2|\mathbf{u}^n|_{2,h}^2 \le \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2. \tag{4.3.61}$$

Now we have to estimate the term $\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2$. By multiplying Eq. (4.3.57a) by $2\Delta t\Delta x\Delta y\,(u_{i,j}^{n+1}-u_{i,j}^n)$ and adding the corresponding equalities for $i = 1,\ldots,N_1$ and $j = 1,\ldots,N_2$, we obtain

$$\begin{aligned}
2\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2 &= 2\Delta t(C_h(\mathbf{u}^n,\mathbf{u}^n),\mathbf{u}^{n+1}-\mathbf{u}^n)_h - 2\Delta t\varepsilon^2(\Delta_h\mathbf{u}^n,\Delta_h(\mathbf{u}^{n+1}-\mathbf{u}^n))\\
&\quad - 2\Delta t\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}\left[\varphi_{i-1/2,j}^n\nabla_{1,h}^- u_{i,j}^{n+1}\right]\nabla_{1,h}^-(u_{i,j}^{n+1}-u_{i,j}^n)\\
&\quad - 2\Delta t\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}\left[\varphi_{i,j-1/2}^n\nabla_{2,h}^- u_{i,j}^{n+1}\right]\nabla_{2,h}^-(u_{i,j}^{n+1}-u_{i,j}^n)
\end{aligned} \tag{4.3.62}$$

Using (1.1.12) and (1.1.13), we majorize all the terms on the right hand side of (4.3.62) as follows:

$$2\Delta t(C_h(\mathbf{u}^n,\mathbf{u}^n),\mathbf{u}^{n+1}-\mathbf{u}^n)_h \le \frac{36\Delta t^2}{\Delta x\Delta y}(|\alpha_1|+|\alpha_2|)^2\|\mathbf{u}^n\|_h^2|\mathbf{u}^n|_{1,h}^2 + \frac{1}{4}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2;$$

$$-2\Delta t\varepsilon^2(\Delta_h\mathbf{u}^n,\Delta_h(\mathbf{u}^{n+1}-\mathbf{u}^n))_h \le 64\varepsilon^4\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)^2|\mathbf{u}^n|_{2,h}^2 + \frac{1}{4}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2;$$

$$\begin{aligned}
-2\Delta t\Delta x\Delta y\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}\Big[&\left(\varphi_{i-1/2,j}^n\nabla_{1,h}^- u_{i,j}^{n+1}\right)\nabla_{1,h}^-(u_{i,j}^{n+1}-u_{i,j}^n) + \left(\varphi_{i,j-1/2}^n\nabla_{2,h}^- u_{i,j}^{n+1}\right)\nabla_{2,h}^-(u_{i,j}^{n+1}-u_{i,j}^n)\Big]\\
&\le \frac{144\Delta t^2}{\Delta x^2\Delta y^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^n\|_h^4|\mathbf{u}^n|_{1,h}^2 + \frac{1}{4}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2\\
&\quad + 16\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)|\mathbf{u}^n|_{1,h}^2 + \frac{1}{4}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2.
\end{aligned}$$

Hence, (4.3.62) becomes

$$\begin{aligned}
\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2 \le{}& \frac{36\Delta t^2}{\Delta x\Delta y}(|\alpha_1|+|\alpha_2|)^2\|\mathbf{u}^n\|_h^2|\mathbf{u}^n|_{1,h}^2 + 64\varepsilon^4\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)^2|\mathbf{u}^n|_{2,h}^2\\
&+ \frac{144\Delta t^2}{\Delta x^2\,\Delta y^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^n\|_h^4|\mathbf{u}^n|_{1,h}^2 + 16\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)|\mathbf{u}^n|_{1,h}^2,
\end{aligned}$$

which by (4.3.58) gives

$$\begin{aligned}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|_h^2 \leq & \frac{36\Delta t^2}{\Delta x\Delta y}(|\alpha_1|+|\alpha_2|)^2\|\mathbf{u}^n\|_h^2|\mathbf{u}^n|_{1,h}^2+\Delta t\varepsilon^2(1-\delta)|\mathbf{u}^n|_{2,h}^2\\ &+\frac{144\Delta t^2}{\Delta x^2\,\Delta y^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^n\|_h^4|\mathbf{u}^n|_{1,h}^2\\ &+16\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)|\mathbf{u}^n|_{1,h}^2. \end{aligned} \tag{4.3.63}$$

On substitution of (4.3.63) back to (4.3.61), we arrive at

$$\begin{aligned}\|\mathbf{u}^{n+1}\|_h^2-\exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2+\Delta t\varepsilon^2\delta|\mathbf{u}^n|_{2,h}^2 \leq & 16\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)|\mathbf{u}^n|_{1,h}^2\\ &+\frac{36\Delta t^2}{\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2+\frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^n\|_h^2\right]\|\mathbf{u}^n\|_h^2|\mathbf{u}^n|_{1,h}^2. \end{aligned} \tag{4.3.64}$$

Using (4.2.25) and (4.3.59), (4.3.64) gives

$$\begin{aligned}\|\mathbf{u}^{n+1}\|_h^2-\exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2+\Delta t\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \leq & \frac{36\Delta t^2}{\Delta x\Delta y}\Bigg[(|\alpha_1|+|\alpha_2|)^2\\ &+\frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^n\|_h^2\Bigg]\|\mathbf{u}^n\|_h^2|\mathbf{u}^n|_{1,h}^2. \end{aligned} \tag{4.3.65}$$

We then need to show by induction on n, that

$$\|\mathbf{u}^{n+1}\|_h^2+\frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2. \tag{4.3.66}$$

For $n=0$, from (4.3.65), we obtain

$$\begin{aligned}\|\mathbf{u}^1\|_h^2+\Delta t\varepsilon^2\delta^2\eta^2|\mathbf{u}^0|_{1,h}^2 \leq & \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2\\ &+\frac{36\Delta t^2}{\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2+\frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^0\|_h^2\right]\|\mathbf{u}^0\|_h^2|\mathbf{u}^0|_{1,h}^2, \end{aligned}$$

which with (4.3.60) leads to

$$\|\mathbf{u}^1\|_h^2+\frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^0|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2,$$

which is (4.3.66) for $n=0$. Assuming now that (4.3.66) is true up to the order $n-1$, this is to say that for $s=0,2,\ldots,n-1$, we have

$$\|\mathbf{u}^s\|_h^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^{s-1}\|_h^2 \text{ and } \|\mathbf{u}^s\|_h^2 \leq \exp\left(\frac{s\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2. \tag{4.3.67}$$

Using (4.3.67) in (4.3.65), one obtains

$$\|\mathbf{u}^{n+1}\|_h^2 + \Delta t\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2 + \frac{36\Delta t}{\Delta x\Delta y}\exp\left(\frac{2n\Delta t}{\varepsilon^2}\right)\left[(|\alpha_1|+|\alpha_2|)^2 + \frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^0\|_h^2\right]\|\mathbf{u}^0\|_h^2|\mathbf{u}^n|_{1,h}^2,$$

which by (4.3.60) gives

$$\|\mathbf{u}^{n+1}\|_h^2 + \Delta t\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2,$$

re-written also as follows

$$\|\mathbf{u}^{n+1}\|_h^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2.$$

We then have

$$\begin{aligned}
\|\mathbf{u}^{n+1}\|_h^2 &\leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^n\|_h^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \\
&\leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\left[\exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^{n-1}\|_h^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{n-1}|_{1,h}^2\right] - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \\
&\leq \exp\left(\frac{2\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^{n-1}\|_h^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\exp\left(\frac{\Delta t}{\varepsilon^2}\right)|\mathbf{u}^{n-1}|_{1,h}^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \\
&\leq \exp\left(\frac{3\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^{n-2}\|_h^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\exp\left(\frac{2\Delta t}{\varepsilon^2}\right)|\mathbf{u}^{n-2}|_{1,h}^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\exp\left(\frac{\Delta t}{\varepsilon^2}\right)|\mathbf{u}^{n-1}|_{1,h}^2 \\
&\quad - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^n|_{1,h}^2 \\
&\vdots \\
&\leq \exp\left(\frac{(n+1)\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2 - \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\sum_{s=0}^{n}\exp\left(\frac{(n-s)\Delta t}{\varepsilon^2}\right)|\mathbf{u}^s|_{1,h}^2.
\end{aligned}$$

Hence we get

$$\|\mathbf{u}^{n+1}\|_h^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\sum_{s=0}^{n}\exp\left(\frac{(n-s)\Delta t}{\varepsilon^2}\right)|\mathbf{u}^s|_{1,h}^2 \leq \exp\left(\frac{(n+1)\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2.$$

Therefore, the proof is complete. □

4.4 Multilevel finite volume methods

Multilevel methods were introduced to improve calculation speed in the simulation of complex physical phenomena while maintaining a good level of accuracy, see (Bousquet et al.,

2013a, 2014; He and Liu, 2005; Faure et al., 2005; Bousquet and Temam, 14-17 June 2010; Adamy et al., 2010). This section is an application of the work presented in Bousquet et al. (2014), in which the shallow water equations is analyzed. Here, we are concerned with the two dimensional convective Cahn-Hilliard equation (4.1.1)-(4.1.4). We formulate in the spirit of Bousquet et al. (2014) two methods approximating (4.1.1)-(4.1.4), namely: implicit multilevel finite volume method and explicit multilevel finite volume method. These new methods are next studied thoroughly and comparison by stability and CPU time with the associated one-level methods discussed in section 4.3 are established. To make this text self-contained for the reader, we recall below the multilevel finite volume approximation as described in Bousquet et al. (2014).

Here N_1 and N_2 are assumed to be divisible by 3. Let N_1^0, N_2^0 and M_0 be integers such that $3N_1^0 = N_1, 3N_2^0 = N_2$ and $\Delta t M_0 = T$. We discretize $\mathcal{M}$ into fine meshes and coarse meshes. The fine mesh consists of $3N_1^0 \times 3N_2^0$ regular cells $(k_{i,j})_{1\le i\le 3N_1^0, 1\le j\le 3N_2^0}$ of uniform area $\Delta x \Delta y$.

The coarse mesh consists of $N_1^0 N_2^0$ control volumes $(K_{l,m})_{1\le l\le N_1^0, 1\le m\le N_2^0}$ of uniform area $9\Delta x\Delta y$, where

$$K_{l,m} = (x_{3l-2-1/2}, x_{3m+1/2}) \times (y_{3m-2-1/2}, y_{3m+1/2}).$$

We denote the approximate solutions on the fine grid by $u_{i,j}$, $1 \le i \le 3N_1^0, 1 \le j \le 3N_2^0$. The approximations (averages) on the coarse mesh are defined by

$$U_{l,m} = \frac{1}{9}\sum_{\alpha,\beta=0}^{2} u_{3l-\alpha,3m-\beta}, 1 \le l \le N_1^0, 1 \le m \le N_2^0,$$

and the incremental unknowns by

$$Z_{3l-\alpha,3m-\beta} = u_{3l-\alpha,3m-\beta} - U_{l,m}, \quad \alpha,\beta = 0,1,2, \tag{4.4.1}$$

which satisfy $\sum_{\alpha,\beta=0}^{2} Z_{3l-\alpha,3m-\beta} = 0$.

Let $p > 1$ and $q > 1$ be two fixed integers. We discretize (4.1.1) on the fine mesh by using time step $\Delta t/p$ and on the coarse mesh by using time step Δt. We assume that n is a

multiple of $q+1$ and $(u^n_{i,j})_{1\le i\le 3N_1^0,\,1\le j\le 3N_2^0}$ are known, where $u^n_{i,j}$ is an approximation of the average value of u over $k_{i,j}$ at the grid $t=n\Delta t$, for $i=1,\dots,3N_1^0$, $j=1,\dots,3N_2^0$. For $r=0,1,\dots,p$ and $s=1,2,\dots,q+1$, we let $u^{n+r/p}_{i,j}$ be the approximate solution of the mean values over $k_{i,j}$ at time $t_{n+t/p}=n\Delta t+r\Delta t/p$ for $i=1,\dots,3N_1^0, j=1,\dots,3N_2^0$ and $U^{n+s}_{l,m}$ be the approximate solution of the averages on the coarse mesh $K_{l,m}$ at time $t_{n+s}=(n+s)\Delta t$ for $l=1,\dots,N_1^0$ and $m=1,\dots,N_2^0$.

4.4.1 Implicit multilevel finite volume method

For $0\le r\le p-1$ and $1\le s\le q$, the following multilevel scheme is used to discretize Eqs. (4.1.1)-(4.1.4).

$$\frac{p}{\Delta t}(u^{n+(r+1)/p}_{i,j}-u^{n+r/p}_{i,j})-(C_h(\mathbf{u}^{n+(r+1)/p},\tilde{\mathbf{u}}^{n+r/p}))_{i,j}+\varepsilon^2\Delta_h^2u^{n+(r+1)/p}_{i,j}$$
$$=\nabla^+_{1,h}(\varphi^{n+r/p}_{i-1/2,j}\nabla^-_{1,h}u^{n+(r+1)/p}_{i,j})+\nabla^+_{2,h}(\varphi^{n+r/p}_{i,j-1/2}\nabla^-_{2,h}u^{n+(r+1)/p}_{i,j}),\quad(4.4.2a)$$

$$\frac{U^{n+s+1}_{l,m}-U^{n+s}_{l,m}}{\Delta t}-(C_{3h}(\mathbf{U}^{n+s+1},\tilde{\mathbf{U}}^{n+s}))_{l,m}+\varepsilon^2\Delta_{3h}^2U^{n+s+1}_{l,m}$$
$$=\nabla^+_{1,3h}(\Phi^{n+s}_{l-1/2,m}\nabla^-_{1,3h}U^{n+s+1}_{l,m})+\nabla^+_{2,3h}(\Phi^{n+s}_{l,m-1/2}\nabla^-_{2,3h}U^{n+s+1}_{l,m}),\quad(4.4.2b)$$

$$u^{n+(r+1)/p}_{i,j}=u^{n+(r+1)/p}_{i+3N_1^0,j}=u^{n+(r+1)/p}_{i,j+3N_2^0},\quad(4.4.2c)$$

$$U^{n+s+1}_{l,m}=U^{n+s+1}_{l+N_1^0,m}=U^{n+s+1}_{l,m+N_2^0},\quad(4.4.2d)$$

$$u^0_{i,j}=\frac{1}{\Delta x\Delta y},\iint_{k_{i,j}}u^0(x,y)dxdy,\quad(4.4.2e)$$

where $1\le i\le 3N_1^0, 1\le j\le 3N_2^0, 1\le l\le N_1^0$, $1\le m\le N_2^0$ and

$$\varphi^{n+r/p}_{i-1/2,j}=\frac{f'(u^{n+r/p}_{i,j})+f'(u^{n+r/p}_{i-1,j})}{2},\quad \varphi^{n+r/p}_{i,j-1/2}=\frac{f'(u^{n}_{i,j})+f'(u^{n+r/p}_{i,j-1})}{2},$$
$$\Phi^{n+s}_{l-1/2,m}=\frac{f'(U^{n+s}_{l,m})+f'(U^{n+s}_{l-1,m})}{2},\quad \Phi^{n+s}_{l,m-1/2}=\frac{f'(U^{n}_{l,m})+f'(U^{n+s}_{l,m-1})}{2}.$$

The multilevel discretization consists of alternating p steps on (4.4.2a) with smaller time step $\Delta t/p$, from t_n to t_{n+1} followed by q steps on (4.4.2b) with time step Δt, the incrementals being frozen at t_{n+1} from t_{n+1} to t_{n+q+1}. Then, using Eq. (4.4.1), we can go back to the fine mesh for p steps from t_{n+q+1} to t_{n+q+2}.

Since (4.4.2) is a succession of linear equation, we prove only uniqueness of solution and hence existence.

Theorem 4.4.1. *For $\Delta t < 4\varepsilon^2$, then the approximate solution $\boldsymbol{u}^n$ of* (4.4.2a)-(4.4.2e) *is unique.*

Proof. Suppose $\mathbf{v}_0 = \mathbf{u}_0$ and let $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^{M_0} \in \mathcal{H}_h$, $\Delta t M_0 = T$, such that $\mathbf{v}^n$ satisfies (4.4.2a) and (4.4.2b) with the relation (4.4.1). Let $\mathbf{z}^n = \mathbf{u}^n - \mathbf{v}^n$ and clearly $\mathbf{z}^0 = 0$. We use mathematical induction on n. Since $\mathbf{z}^0 = 0$, the theorem holds for $n = 0$. We assume that $\mathbf{z}^t = 0$ for $t = 1, 2, \ldots, n$. If n is multiple of $q+1$, we need to show the theorem holds by induction on r using Eq. (4.4.2a). We observe that $\mathbf{z}^{n+(r+1)/p}$ is a solution of

$$\begin{aligned}
&\frac{p}{\Delta t}(z_{i,j}^{n+(r+1)/p} - z_{i,j}^{n+r/p}) - (C_h(\mathbf{u}^{n+(r+1)/p}, \tilde{\mathbf{u}}^{n+r/p}))_{i,j} + (C_h(\mathbf{v}^{n+(r+1)/p}, \tilde{\mathbf{v}}^{n+r/p}))_{i,j} + \varepsilon^2 \Delta_h^2 z_{i,j}^{n+(r+1)/p} \\
&\quad = \nabla_{1,h}^+(\varphi_{i-1/2,j}^{n+r/p} \nabla_{1,h}^- u_{i,j}^{n+(r+1)/p}) + \nabla_{2,h}^+(\varphi_{i,j-1/2}^{n+r/p} \nabla_{2,h}^- u_{i,j}^{n+(r+1)/p}) \\
&\qquad - \nabla_{1,h}^+(\psi_{i-1/2,j}^{n+r/p} \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p}) - \nabla_{2,h}^+(\psi_{i,j-1/2}^{n+r/p} \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}),
\end{aligned}$$

for $i = 1, \ldots, 3N_1^0$, $j = 1, \ldots, 3N_2^0$ and

$$\psi_{i-1/2,j}^{n+r/p} = \frac{f'(v_{i,j}^{n+r/p}) + f'(v_{i-1,j}^{n+r/p})}{2} \text{ and } \psi_{i,j-1/2}^{n+r/p} = \frac{f'(v_{i,j}^{n+r/p}) + f'(v_{i,j-1}^{n+r/p})}{2}.$$

By assuming that $\mathbf{z}^{n+r/p} = 0$, we have

$$\begin{aligned}
\frac{z_{i,j}^{n+(r+1)/p}}{\Delta t} - (C_h(\mathbf{z}^{n+(r+1)/p}, \tilde{\mathbf{u}}^{n+r/p}))_{i,j} + \varepsilon^2 \Delta_h^2 z_{i,j}^{n+(r+1)/p} &= \nabla_{1,h}^+(\varphi_{i-1/2,j}^{n+r/p} \nabla_{1,h}^- z_{i,j}^{n+(r+1)/p}) \\
&\quad + \nabla_{2,h}^+(\varphi_{i,j-1/2}^{n+r/p} \nabla_{2,h}^- z_{i,j}^{n+(r+1)/p}). \qquad (4.4.3)
\end{aligned}$$

Multiplying (4.4.3) by $\dfrac{\Delta t\, \Delta x\, \Delta y}{p}\, z_{i,j}^{n+(r+1)/p}$ and summing from $i = 1$ to $i = 3N_1^0$ and $j = 1$ to $j = 3N_2^0$, we obtain

$$\begin{aligned}
\|\mathbf{z}^{n+(r+1)/p}\|_h^2 + \frac{\Delta t \varepsilon^2}{p} |\mathbf{z}^{n+(r+1)/p}|_{2,h}^2 = \frac{\Delta t\, \Delta x\, \Delta y}{p} \sum_{i=1}^{3N_1^0} \sum_{j=1}^{3N_2^0} \Bigg[&\nabla_{1,h}^+(\varphi_{i-1/2,j}^{n+r/p} \nabla_{1,h}^- z_{i,j}^{n+(r+1)/p}) \\
&+ \nabla_{2,h}^+(\varphi_{i,j-1/2}^{n+r/p} \nabla_{2,h}^- z_{i,j}^{n+(r+1)/p}) \Bigg].
\end{aligned}$$

Hence using Lemmas 4.3.1 and 4.2.7, we get

$$(1 - \frac{\Delta t}{4p\varepsilon^2}) \|\mathbf{z}^{n+(r+1)/p}\|_h^2 \leq 0,$$

which for $\dfrac{\Delta t}{4p\,\varepsilon^2} < 1$, and after p iterations gives

$$\|\mathbf{z}^{n+1}\|_h^2 \leq 0,$$

and implies

$$\|\mathbf{z}^{n+1}\|_h = 0.$$

For the case n is not multiple of $q+1$, we use Eq. (4.4.2b) to prove the theorem. By first calculating $\mathbf{U}^n$ and $\mathbf{V}^n$ using Eq. (4.4.1) from $\mathbf{u}^n$ and $\mathbf{v}^n$, respectively, we obtain

$$\begin{aligned}\frac{1}{\Delta t}(\mathcal{Z}_{l,m}^{n+1} - \mathcal{Z}_{l,m}^{n}) - \left[(C_{3h}(\mathbf{U}^{n+1}, \tilde{\mathbf{U}}^n))_{l,m} - (C_{3h}(\mathbf{V}^{n+1}, \tilde{\mathbf{V}}^n))_{l,m}\right] + \varepsilon^2 \Delta_{3h}^2 \mathcal{Z}_{l,m}^{n+1} \\ = \nabla_{3h}^{+}\left(\Phi_{l-\frac{1}{2},m}^{n} \nabla_{3h}^{-} U_{l,m}^{n+1}\right) - \nabla_{3h}^{+}\left(\Psi_{l-\frac{1}{2},m}^{n} \nabla_{3h}^{-} V_{l,m}^{n+1}\right), \qquad (4.4.4)\end{aligned}$$

where $\boldsymbol{\mathcal{Z}}^n = \mathbf{U}^n - \mathbf{V}^n$ for $l = 1, \ldots N_1^0, m = 1, \ldots, N_2^0$. Using the assumption $\tilde{\mathbf{U}}^n = \tilde{\mathbf{V}}^n$, (4.4.4) becomes

$$\frac{\mathcal{Z}_{l,m}^{n+1}}{\Delta t} - (C_{3h}(\boldsymbol{\mathcal{Z}}^{n+1}), \tilde{\mathbf{U}}^n))_{l,m} + \varepsilon^2 \Delta_{3h}^2 \mathcal{Z}_l^{n+1} = \nabla_{3h}^{+}\left(\Phi_{l-\frac{1}{2},m}^{n} \nabla_{3h}^{-} \mathcal{Z}_{l,m}^{n+1}\right).$$

Multiplying Eq. (4.4.4) by $6\Delta t\,\Delta x \Delta y\, \mathcal{Z}_{l,m}^{n+1}$ and adding the corresponding equalities for $l = 1, \ldots, N_0$, we obtain

$$\|\boldsymbol{\mathcal{Z}}^{n+1}\|_{3h}^2 + \Delta t \varepsilon^2 |\boldsymbol{\mathcal{Z}}^{n+1}|_{2,3h}^2 = 6\Delta t\,\Delta x \Delta y \sum_{l=1}^{N_1^0}\sum_{m=1}^{N_2^0} \nabla_{3h}^{+}\left(\Phi_{l-\frac{1}{2},m}^{n} \nabla_{3h}^{-} \mathcal{Z}_{l,m}^{n+1}\right) \mathcal{Z}_{l,m}^{n+1}.$$

Hence using Lemma 4.3.1, one obtains

$$\|\boldsymbol{\mathcal{Z}}^{n+1}\|_{3h}^2 + \Delta t \varepsilon^2 |\boldsymbol{\mathcal{Z}}^{n+1}|_{2,3h}^2 \leq \Delta t |\boldsymbol{\mathcal{Z}}^{n+1}|_{1,3h}^2. \qquad (4.4.5)$$

Applying Lemma 4.2.7 and Young's inequality, (4.4.5) leads to

$$(1 - \frac{\Delta t}{4\varepsilon^2})\|\boldsymbol{\mathcal{Z}}^{n+1}\|_{3h}^2 \leq 0,$$

which for $\dfrac{\Delta t}{4\varepsilon^2} < 1$, gives $\boldsymbol{\mathcal{Z}}^{n+1} = 0$. Therefore, $\mathbf{z}^{n+1} = 0$. This completes the proof of the theorem. □

Theorem 4.4.2. *The multilevel method defined by the equations* (4.4.2a)-(4.4.2e) *is conditionally stable in* $L^\infty(0, T; \mathcal{H}_h)$*, that is, if* $\Delta t \leq \varepsilon^2$ *and* $1 \leq n \leq M$*, then*

$$\|\boldsymbol{u}^n\|_h^2 \leq 2^{\dfrac{2T}{\varepsilon^2}} \|\boldsymbol{u}^0\|_h^2.$$

Proof. By multiplying (4.4.2a) by $\dfrac{2\Delta t\,\Delta x\,\Delta y}{p}\,u_{i,j}^{n+(r+1)/p}$ and adding the corresponding equalities for $i=1,\ldots,3N_1^0$ and $j=1,\ldots,3N_2^0$, we obtain

$$2\Delta x\Delta y\sum_{i=1}^{3N_1^0}\sum_{j=1}^{3N_2^0}(u_{i,j}^{n+(r+1)/p}-u_{i,j}^{n+r/p})u_{i,j}^{n+(r+1)/p}+\frac{2\Delta x\Delta y\,\Delta t}{p}\,\varepsilon^2\sum_{i=1}^{3N_1^0}\sum_{j=1}^{3N_2^0}\Delta_h^2u_{i,j}^{n+(r+1)/p}u_{i,j}^{n+(r+1)/p}$$
$$=\frac{2\Delta x\Delta y\,\Delta t}{p}\sum_{i=1}^{3N_1^0}\sum_{j=1}^{3N_2^0}\left[\nabla_{1,h}^+(\varphi_{i-1/2,j}^{n+r/p}\nabla_{1,h}^-u_{i,j}^{n+(r+1)/p})+\nabla_{2,h}^+(\varphi_{i,j-1/2}^{n+r/p}\nabla_{2,h}^-u_{i,j}^{n+(r+1)/p})\right]u_{i,j}^{n+(r+1)/p}. \tag{4.4.6}$$

Using Lemma 4.3.1, Eq. (4.4.6) becomes

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2-\|\mathbf{u}^{n+r/p}\|_h^2+\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2+2\frac{\Delta t}{p}\,\varepsilon^2|\mathbf{u}^{n+(r+1)/p}|_{2,h}^2\le\frac{2\Delta t}{p}\,|\mathbf{u}^{n+(r+1)/p}|_{1,h}^2 \tag{4.4.7}$$

And then using Young's inequality and Lemma 4.2.7, (4.4.7) yields

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2-\|\mathbf{u}^{n+r/p}\|_h^2+\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2\le\frac{\Delta t}{2p\varepsilon^2}\|\mathbf{u}^{n+(r+1)/p}\|_h^2.$$

Thus we have

$$\left[1-\frac{\Delta t}{2p\varepsilon^2}\right]\,\|\mathbf{u}^{n+(r+1)/p}\|_h^2\le\|\mathbf{u}^{n+r/p}\|_h^2.$$

Based on (1.1.10), for $\dfrac{\Delta t}{2p\,\varepsilon^2}\le\dfrac{1}{2}$, we have

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2\le 2^{\frac{2\Delta t}{p\,\varepsilon^2}}\|\mathbf{u}^{n+r/p}\|_h^2.$$

After p iterations, we obtain

$$\|\mathbf{u}^{n+1}\|_h^2\le 2^{\frac{2\Delta t}{\varepsilon^2}}\|\mathbf{u}^{n}\|_h^2. \tag{4.4.8}$$

We now perform q iterations on the coarse grid, Eq. (4.4.2b), using time step Δt and the relations (4.4.1). At time $t_{n+s}=(n+s)\Delta t, 2\le s\le q+1$, the incremental unknowns $Z_{i,j}$ are frozen at time $(n+1)\Delta t$. Multiplying Eq. (4.4.2b) by $18\Delta t\Delta x\Delta y\,U_{l,m}^{n+s+1}$ and adding the equalities for $l=1,\ldots,N_1^0$ and $m=1,\ldots,N_2^0$, we get

$$(\mathbf{U}^{n+s+1}-\mathbf{U}^{n+s},\mathbf{U}^{n+s+1})_{3h}+2\Delta t\varepsilon^2(\Delta_{3h}^2\mathbf{U}^{n+s+1},\mathbf{U}^{n+s+1})_{3h}$$

$$= 18\Delta x\Delta y\Delta t \sum_{l=1}^{N_1^0}\sum_{m=1}^{N_2^0}\left[\nabla^+_{1,3h}(\Phi^{n+s}_{l-1/2,m}\nabla^-_{1,3h}U^{n+s+1}_{l,m}) + \nabla^+_{2,3h}(\Phi^{n+s}_{l,m-1/2}\nabla^-_{2,3h}U^{n+s+1}_{l,m})\right]U^{n+s+1}_{l,m}. \tag{4.4.9}$$

Using Lemma 4.3.1 together with (4.4.9), we get

$$\|\mathbf{U}^{n+s+1}\|^2_{3h} - \|\mathbf{U}^{n+s}\|^2_{3h} + \|\mathbf{U}^{n+s+1} - \mathbf{U}^{n+s}\|^2_{3h} + 2\Delta t\,\varepsilon^2|\mathbf{U}^{n+s+1}|^2_{2,3h} \le 2\Delta t|\mathbf{U}^{n+s+1}|^2_{1,3h}. \tag{4.4.10}$$

Using Lemma 4.2.7 and Young's inequality, (4.4.10) becomes

$$\|\mathbf{U}^{n+s+1}\|^2_{3h} - \|\mathbf{U}^{n+s}\|^2_{3h} + \|\mathbf{U}^{n+s+1} - \mathbf{U}^{n+s}\|^2_{3h} \le \frac{\Delta t}{2\varepsilon^2}\|\mathbf{U}^{n+s+1}\|^2_{3h}.$$

Thus we have

$$\left[1 - \frac{\Delta t}{2\varepsilon^2}\right]\|\mathbf{U}^{n+s+1}\|^2_{3h} \le \|\mathbf{U}^{n+s}\|^2_{3h}.$$

Using (1.1.10), for $\dfrac{\Delta t}{2\varepsilon^2} \le \dfrac{1}{2}$, we obtain

$$\|\mathbf{U}^{n+s+1}\|^2_{3h} \le 2^{\frac{2\Delta t}{\varepsilon^2}}\|\mathbf{U}^{n+s}\|^2_{3h}. \tag{4.4.11}$$

From the definition of the increments $Z^{n+1}_{3l-\alpha,3m-\beta}$, we have

$$u^{n+s}_{3l-\alpha,3m-\beta} = U^{n+s}_{l,m} + Z^{n+1}_{3l-\alpha,3m-\beta}, \quad 1 \le l \le N_1^0, 1 \le m \le N_2^0, \quad \alpha,\beta = 0,1,2.$$

Taking the sum over α and β, we get

$$\sum_{\alpha,\beta=0}^{2}|u^{n+s}_{3l-\alpha,3m-\beta}|^2 = \sum_{\alpha,\beta=0}^{2}|U^{n+s}_{l,m} + Z^{n+1}_{3l-\alpha,3m-\beta}|^2 = 9|U^{n+s}_{l,m}|^2 + \sum_{\alpha,\beta=0}^{2}|Z^{n+1}_{3l-\alpha,3m-\beta}|^2.$$

For $s = 1,\ldots,q+1$, the following relation holds

$$\|\mathbf{u}^{n+s}\|^2_h = \|\mathbf{U}^{n+s}\|^2_{3h} + \|\mathbf{Z}^{n+1}\|^2_h \tag{4.4.12}$$

By adding $\|\mathbf{Z}^{n+1}\|^2_h$ to both sides of inequality (4.4.11) and using (4.4.12), we get

$$\|\mathbf{u}^{n+s+1}\|^2_h \le 2^{\frac{2\Delta t}{\varepsilon^2}}\|\mathbf{u}^{n+s}\|^2_h.$$

After q iterations, we have

$$\|\mathbf{u}^{n+q+1}\|^2_h \le 2^{\frac{2\Delta t\ q}{\varepsilon^2}}\|\mathbf{u}^{n+1}\|^2_h,$$

which by (4.4.8) gives

$$\|\mathbf{u}^{n+q+1}\|_h^2 \leq 2^{\frac{2\Delta t\ (q+1)}{\varepsilon^2}} \|\mathbf{u}^n\|_h^2.$$

By induction over n, we obtain

$$\|\mathbf{u}^n\|_h^2 \leq 2^{\frac{2n\Delta t}{\varepsilon^2}} \|\mathbf{u}^0\|_h^2 \leq 2^{\frac{2T}{\varepsilon^2}} \|\mathbf{u}^0\|_h^2.$$

This completes the proof. □

Theorem 4.4.3. *Suppose that the solution $u(x,t)$ of Eqs.* (4.1.1)-(4.1.4) *is sufficiently smooth.*

Assume that $\Delta t < 4\varepsilon^2$, (4.4.17)*, and* (4.4.20) *are satisfied.*

Assume also that $\Delta t, \Delta x$ and Δy satisfy (4.4.18).

Then, the solution of the finite volume discretization (4.4.2a)-(4.4.2e) *converges to the solution of Eq.* (4.1.1) *in the discrete L^2-norm with rate of convergence $\mathcal{O}(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)$.*

Proof. Let n is a multiple of $q+1$. Let

$$v_{i,j}^{n+r/p} = \iint_{k_{i,j}} u(x,y,t_{n+r/p})dxdy,$$

be the cell average of u at time $t_{n+r/p}$ on the cell $k_{i,j}$ for $i = 1,\ldots,3N_1^0$, $j = 1,\ldots,3N_2^0$, $r = 1,\ldots,p$. Denote

$$s = \max_{-L_1\leq x\leq L_1, -L_2\leq y\leq L_2, 0\leq t\leq T} |u(x,y,t)|.$$

Making use of Taylor expansion, we obtain

$$\begin{aligned}&\frac{v_{i,j}^{n+(r+1)/p} - v_{i,j}^{n+r/p}}{\Delta t/p} - (C_h(\boldsymbol{v}^{n+(r+1)/p}, \tilde{\boldsymbol{v}}^{n+r/p}))_{i,j} + \varepsilon^2 \Delta_h^2 v_{i,j}^{n+(r+1)/p} \\ &\quad = \nabla_{1,h}^+ \left[\psi_{i-\frac{1}{2},j}^{n+r/p} \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p}\right] + \nabla_{2,h}^+ \left[\psi_{i,j-\frac{1}{2}}^{n+r/p} \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}\right] + \tau_{i,j}^{n+r/p}, \quad (4.4.13)\end{aligned}$$

where $\tau_{i,j}^{n+r/p} \in \mathcal{H}_h$ is the truncation error of the finite volume discretization (4.4.2a) for $i = 1,\ldots,3N_1^0$, $j = 1,\ldots,3N_2^0$. There exists a positive constant c_1 such that

$$\max_{i,j,n,r} |\tau_{i,j}^{n+r/p}| \leq c_1 \left(\frac{\Delta t}{p} + \Delta x^2 + \Delta y^2\right),\ i = 1,\ldots,3N_1^0,\ j = 1,\ldots,3N_2^0,\ r = 1,\ldots,p.$$

Let $\mathbf{e}^{n+r/p} = \boldsymbol{v}^{n+r/p} - \mathbf{u}^{n+r/p}$, where $u_{i,j}^{n+r/p}$ is the solution of (4.4.2a). Substituting $u_{i,j}^{n+r/p} = v_{i,j}^{n+r/p} - e_{i,j}^{n+r/p}$ in Eq. (4.4.2a), we obtain

$$\frac{v_{i,j}^{n+(r+1)/p} - v_{i,j}^{n+r/p}}{\Delta t/p} - (C_h(\boldsymbol{v}^{n+(r+1)/p}, \tilde{\boldsymbol{v}}^{n+r/p}))_{i,j} + \varepsilon^2 \Delta_h^2 v_{i,j}^{n+(r+1)/p} - \nabla_{1,h}^+(\psi_{i-\frac{1}{2},j}^{n+r/p} \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p})$$
$$- \nabla_{2,h}^+(\psi_{i-\frac{1}{2},j-\frac{1}{2}}^{n+r/p} \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}) = \frac{e_{i,j}^{n+(n+1)/p} - e_{i,j}^{n+r/p}}{\Delta t/p} - (C_h(\mathbf{e}^{n+1}, \tilde{\boldsymbol{v}}^{n+r/p} - \tilde{\mathbf{e}}^{n+r/p}))_{i,j}$$
$$+ \varepsilon^2 \Delta_h^2 e_{i,j}^{n+(r+1)/p} - \left[\nabla_{1,h}^+\left(\varphi_{i-\frac{1}{2},j}^{n+r/p} \nabla_{1,h}^- e_{i,j}^{n+(r+1)/p}\right) + \nabla_{2,h}^+\left(\varphi_{i,j-\frac{1}{2}}^{n+r/p} \nabla_{2,h}^- e_{i,j}^{n+(r+1)/p}\right)\right]$$
$$+ \left[\nabla_{1,h}^+\left((\varphi_{i-\frac{1}{2},j}^{n+r/p} - \psi_{i-\frac{1}{2},j}^{n+r/p}) \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p}\right) + \nabla_{2,h}^+\left((\varphi_{i,j-\frac{1}{2}}^{n+r/p} - \psi_{i,j-\frac{1}{2}}^{n+r/p}) \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}\right)\right]$$
$$- (C_h(\boldsymbol{v}^{n+(r+1)/p}, \tilde{\mathbf{e}}^{n+r/p}))_{i,j}. \tag{4.4.14}$$

Using (4.4.13), Eq. (4.4.14) becomes

$$\frac{e_{i,j}^{n+(r+1)/p} - e_{i,j}^{n+r/p}}{\Delta t/p} - (C_h(\mathbf{e}^{n+(r+1)/p}, \tilde{\boldsymbol{v}}^{n+r/p} - \tilde{\mathbf{e}}^{n+r/p}))_{i,j} + \varepsilon^2 \Delta_h^2 e_{i,j}^{n+(r+1)/p}$$
$$= \nabla_{1,h}^+\left(\varphi_{i-\frac{1}{2},j}^{n+r/p} \nabla_{1,h}^- e_{i,j}^{n+(r+1)/p}\right) + \nabla_{2,h}^+\left(\varphi_{i,j-\frac{1}{2}}^{n+r/p} \nabla_{2,h}^- e_{i,j}^{n+(r+1)/p}\right) + (C_h(\boldsymbol{v}^{n+(r+1)/p}, \tilde{\mathbf{e}}^{n+r/p}))_{i,j}$$
$$+ \nabla_{1,h}^+\left[\left(3(v_{i,j}^{n+r/p} e_{i,j}^{n+r/p} + v_{i-1,j}^{n+r/p} e_{i-1,j}^{n+r/p}) - \frac{3}{2}[(e_{i,j}^{n+r/p})^2 + (e_{i-1,j}^{n+r/p})^2]\right) \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p}\right]$$
$$+ \nabla_{2,h}^+\left[\left(3(v_{i,j}^{n+r/p} e_{i,j}^{n+r/p} + v_{i,j-1}^{n+r/p} e_{i,j-1}^{n+r/p}) - \frac{3}{2}[(e_{i,j}^{n+r/p})^2 + (e_{i,j-1}^{n+r/p})^2]\right) \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}\right]$$
$$+ \tau_{i,j}^{n+r/p}. \tag{4.4.15}$$

Multiplying Eq. (4.4.15) by $\frac{2\Delta t \Delta x \Delta y}{p} e_{i,j}^{n+(r+1)/p}$ and adding for $i = 1, \ldots, 3N_1^0$ and $j = 1, \ldots, 3N_2^0$, we obtain

$$2(\mathbf{e}^{n+(r+1)/p} - \mathbf{e}^n, \mathbf{e}^{n+(r+1)/p})_h + \frac{2\Delta t}{p} \varepsilon^2 \|\mathbf{e}^{n+(r+1)/p}\|_{2,h}^2 = \frac{2\Delta t}{p}(C_h(\boldsymbol{v}^{n+(r+1)/p}, \tilde{\mathbf{e}}^{n+r/p}), \mathbf{e}^{n+(r+1)/p})_h$$
$$+ \frac{2\Delta t \Delta x \Delta y}{p} \sum_{i=1}^{3N_1^0} \sum_{j=1}^{3N_2^0} \Bigg\{ \left[\nabla_{1,h}^+\left(\varphi_{i-\frac{1}{2},j}^{n+r/p} \nabla_{1,h}^- e_{i,j}^{n+(r+1)/p}\right) + \nabla_{2,h}^+\left(\varphi_{i,j-\frac{1}{2}}^{n+r/p} \nabla_{2,h}^- e_{i,j}^{n+(r+1)/p}\right)\right] e_{i,j}^{n+(r+1)/p}$$
$$+ \nabla_{1,h}^+\left[\left(3(v_{i,j}^{n+r/p} e_{i,j}^{n+r/p} + v_{i-1,j}^{n+r/p} e_{i-1,j}^{n+r/p}) - \frac{3}{2}[(e_{i,j}^{n+r/p})^2 + (e_{i-1,j}^{n+r/p})^2]\right) \nabla_{1,h}^- v_{i,j}^{n+(r+1)/p}\right] e_{i,j}^{n+(r+1)/p}$$
$$+ \nabla_{2,h}^+\left[\left(3(v_{i,j}^{n+r/p} e_{i,j}^{n+r/p} + v_{i,j-1}^{n+r/p} e_{i,j-1}^{n+r/p}) - \frac{3}{2}[(e_{i,j}^{n+r/p})^2 + (e_{i,j-1}^{n+r/p})^2]\right) \nabla_{2,h}^- v_{i,j}^{n+(r+1)/p}\right] e_{i,j}^{n+(r+1)/p}$$
$$+ \frac{2\Delta t}{p}(\boldsymbol{\tau}^{n+r/p}, \mathbf{e}^{n+(r+1)/p})_h \Bigg\}. \tag{4.4.16}$$

Using (4.2.15) and Lemmas 4.3.1 and 4.2.1, Eq. (4.4.16) becomes

$$\|\mathbf{e}^{n+(r+1)/p}\|_h^2 - \|\mathbf{e}^{n+r/p}\|_h^2 + \|\mathbf{e}^{n+(r+1)/p} - \mathbf{e}^{n+r/p}\|_h^2 + \frac{2\Delta t}{p} \varepsilon^2 \|\mathbf{e}^{n+(r+1)/p}\|_{2,h}^2$$

$$
\begin{aligned}
&\leq \frac{2\Delta t}{p}\|\mathbf{e}_{i,j}^{n+(r+1)/p}\|_{1,h}^2 + \frac{2\Delta t}{p}(C_h(\boldsymbol{\upsilon}^{n+(r+1)/p}, \tilde{\mathbf{e}}^{n+r/p}), \mathbf{e}^{n+(r+1)/p})_h \\
&- \frac{2\Delta t\Delta x\Delta y}{p}\sum_{i=1}^{3N_1^0}\sum_{j=1}^{3N_2^0}\Bigg[3\Big((\upsilon_{i,j}^{n+r/p}\, e_{i,j}^{n+r/p} + \upsilon_{i-1,j}^{n+r/p}\, e_{i-1,j}^{n+r/p})\Big)\nabla_{1,h}^{-}\upsilon_{i,j}^{n+(r+1)/p}\nabla_{1,h}^{-}e_{i,j}^{n+(r+1)/p} \\
&\qquad - \frac{3}{2}\Big((e_{i,j}^{n+r/p})^2 + (e_{i-1,j}^{n+r/p})^2\Big)\nabla_{1,h}^{-}\upsilon_{i,j}^{n+(r+1)/p}\nabla_{1,h}^{-}e_{i,j}^{n+(r+1)/p} \\
&\qquad + 3\Big((\upsilon_{i,j}^{n+r/p}\, e_{i,j}^{n+r/p} + \upsilon_{i,j-1}^{n+r/p}\, e_{i,j-1}^{n+r/p})\Big)\nabla_{2,h}^{-}\upsilon_{i,j}^{n+(r+1)/p}\nabla_{2,h}^{-}e_{i,j}^{n+(r+1)/p} \\
&\qquad - \frac{3}{2}\Big((e_{i,j}^{n+r/p})^2 + (e_{i,j-1}^{n+r/p})^2\Big)\nabla_{2,h}^{-}\upsilon_{i,j}^{n+(r+1)/p}\nabla_{2,h}^{-}e_{i,j}^{n+(r+1)/p}\Bigg] \\
&+ \frac{2\Delta t}{p}(\boldsymbol{\tau}^{n+r/p}, \mathbf{e}^{n+(r+1)/p})_h.
\end{aligned}
$$

Using the approach implemented on the proof of Theorem 4.3.3, we deduce in the fine mesh taking

$$\Delta t c \leq \frac{p}{2}, \tag{4.4.17}$$

then

$$
\begin{aligned}
\|\mathbf{e}^{n+(r+1)/p}\|_h^2 + \frac{1}{p}\,\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+(r+1)/p}|_{1,h}^2 \leq 4^{\Delta t c}\Bigg[\left(1+\frac{\Delta t c_9}{p}\right)\|\mathbf{e}^{n+r/p}\|_h^2 + \frac{\Delta t\, c_8}{p}\|\tilde{\mathbf{e}}^{n+r/p}\|_h^2 \\
+ \frac{\Delta t}{p}\,c_7(\frac{\Delta t}{p}+\Delta x^2+\Delta y^2)^2\Bigg] + \frac{3\Delta t\,\delta_2\,4^{\Delta t c}}{p\Delta x\Delta y}\|\mathbf{e}^{n+r/p}\|_h^2\;|\mathbf{e}^{n+(r+1)/p}|_{1,h}^2.
\end{aligned}
$$

For

$$\frac{3\delta_2\, c_7}{\Delta x\Delta y}\left(\frac{\Delta t}{\min\{p,9\}}+\Delta x^2+\Delta y^2\right)^2 \leq \frac{1}{2T}\varepsilon^2\eta^2 4^{-Tc}\exp\Big(-T\,c_{10}\Big), \tag{4.4.18}$$

as shown in Theorem 4.3.3, we obtain

$$
\begin{aligned}
&\|\mathbf{e}^{n+(r+1)/p}\|_h^2 + \frac{1}{2p}\Delta t\varepsilon^2\eta^2|\mathbf{e}^{n+(r+1)/p}|_{1,h}^2 \\
&\qquad \leq 4^{\frac{\Delta t\, c}{p}}\left[\left(1+\frac{\Delta t c_9}{p}\right)\|\mathbf{e}^{n+r/p}\|_h^2 + \frac{\Delta t c_8}{p}\|\tilde{\mathbf{e}}^{n+r/p}\|_h^2 + \frac{\Delta t}{p}c_7\left(\frac{\Delta t}{p}+\Delta x^2+\Delta y^2\right)^2\right],
\end{aligned}
$$

which after p iterations gives

$$\|\mathbf{e}^{n+1}\|_h^2 \leq 4^{\Delta t c}\exp\left(\Delta t\, c_{10}\right)\left[\|\mathbf{e}^{n}\|_h^2 + \Delta t\, c_7\left(\frac{\Delta t}{p}+\Delta x^2+\Delta y^2\right)^2\right]. \tag{4.4.19}$$

In a similar way for $s = 1, \ldots, q$, and being on the coarse mesh and for

$$\Delta t c \leq \frac{1}{2}, \tag{4.4.20}$$

we get

$$\|\mathbf{E}^{n+s+1}\|_{3h}^2+\frac{1}{2}\Delta t\varepsilon^2\eta^2|\mathbf{E}^{n+s+1}|_{1,3h}^2\leq 4^{\Delta t\, c}\Big[(1+\Delta t c_9)\,\|\mathbf{E}^{n+s}\|_{3h}^2+\Delta t c_8\|\tilde{\mathbf{E}}^{n+s}\|_{3h}^2 + \Delta t c_7\left(\Delta t+(3\Delta x)^2+(3\Delta y)^2\right)^2\Big],$$

where $\mathbf{E}^{n+s}=\boldsymbol{\Upsilon}^{n+s}-\mathbf{U}^{n+s}$ and $\boldsymbol{\Upsilon}^{n+s}$ and $\mathbf{U}^{n+s}$ are exact cell average and numerical solutions on the coarse mesh, respectively. For $n+s<m_0-1$, we have

$$\|\mathbf{E}^{n+s+1}\|_{3h}^2\leq 4^{\Delta tc}\exp\left(\Delta t\,(c_9+c_8)\right)\Big[\|\mathbf{E}^{n+s}\|_{3h}^2+\Delta t\, c_7\left(\Delta t+(3\Delta x)^2+(3\Delta y)^2\right)^2\Big], \tag{4.4.21}$$

As we said at the beginning of this section, the numerical increments $Z_{i,j}$'s are fixed between steps $n+1$ and $n+q+1$ and therefore for $1\leq s\leq q$, $1\leq l\leq N_1^0$ and $1\leq m\leq N_2^0$,

$$Z_{3l-\alpha,3m-\beta}^{n+s+1}=Z_{3l-\alpha,3m-\beta}^{n+1}=u_{3l-\alpha,3m-\beta}^{n+1}-U_{l,m}^{n+1},\quad \alpha,\beta=0,1,2.$$

Using Eq. (4.4.1), we have

$$\begin{aligned}
e_{3l-\alpha,3m-\beta}^{n+s+1}&=\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}-u_{3l-\alpha,3m-\beta}^{n+s+1}\\
&=\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}-(U_{l,m}^{n+s+1}+Z_{3l-\alpha,3m-\beta}^{n+1})\\
&=\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}+\Upsilon_{l,m}^{n+s+1}-\Upsilon_{l,m}^{n+s+1}-(U_{l,m}^{n+s+1}+Z_{3l-\alpha,3m-\beta}^{n+1})\\
&=(\Upsilon_{l,m}^{n+s+1}-U_{l,m}^{n+1})+(\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}-\Upsilon_{l,m}^{n+s+1})-Z_{3l-\alpha,3m-\beta}^{n+1}\\
&=E_{l,m}^{n+s+1}+\zeta_{3l-\alpha,3m-\beta}^{n+s+1},
\end{aligned}$$

where

$$\zeta_{3l-\alpha,3m-\beta}^{n+s+1}=(\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}-\Upsilon_{l,m}^{n+s+1})-Z_{3l-\alpha,3m-\beta}^{n+1},$$

is the difference of the numerical increment from exact increment.

It is clear from the definition of increments that

$$\sum_{\alpha,\beta=0}^{2}(\upsilon_{3l-\alpha,3m-\beta}^{n+s+1}-\Upsilon_{l,m}^{n+s+1})=\sum_{\alpha,\beta=0}^{2}Z_{3l-\alpha,3m-\beta}^{n+1}=0,$$

and hence $\sum_{\alpha,\beta=0}^{2}\zeta_{3l-\alpha,3m-\beta}^{n+s+1}=0$. As a result for $s=1,\ldots,q$,

$$\sum_{\alpha,\beta=0}^{2}\left(e_{3l-\alpha,3m-\beta}^{n+s+1}\right)^2=9\left(E_{l,m}^{n+s+1}\right)^2+\sum_{\alpha,\beta=0}^{2}\left(\zeta_{3l-\alpha,3m-\beta}^{n+s+1}\right)^2. \tag{4.4.22}$$

Multiplying (4.4.22) by $\Delta x \Delta y$ and taking the sum for $l = 1, \ldots, N_1^0$ and $m = 1, \ldots, N_2^0$, we obtain

$$\|\mathbf{e}^{n+s+1}\|_h^2 = \|\mathbf{E}^{n+s+1}\|_{3h}^2 + \|\boldsymbol{\zeta}^{n+s+1}\|_h^2. \tag{4.4.23}$$

We now estimate the term $\boldsymbol{\zeta}^{n+s+1}$. From the definition of increments for $\mathbf{u} \in \mathcal{H}_h$, we have

$$U_{l,m} = u_{3l-\alpha,3m-\beta} + \mathcal{O}(\Delta x + \Delta y),$$

from which

$$\Upsilon_{l,m}^{n+s+1} - \upsilon_{3l-\alpha,3m-\beta}^{n+s+1} = \Upsilon_{l,m}^{n+s} - \upsilon_{3l-\alpha,3m-\beta}^{n+s} + \Delta t\, \mathcal{O}(\Delta x + \Delta y).$$

Hence

$$\boldsymbol{\zeta}^{n+s+1} = \boldsymbol{\zeta}^{n+s} + \Delta t\, \mathcal{O}(\Delta x + \Delta y). \tag{4.4.24}$$

(4.4.24) gives

$$\|\boldsymbol{\zeta}^{n+s+1}\|_h^2 = \|\boldsymbol{\zeta}^{n+s}\|_h^2 + \Delta t\, \mathcal{O}(\Delta t\, (\Delta x + \Delta y)^2),$$

which with the application of Young's inequality (1.1.12) gives

$$\|\boldsymbol{\zeta}^{n+s+1}\|_h^2 \leq \|\boldsymbol{\zeta}^{n+s}\|_h^2 + \Delta t\, c_{11}\, (\Delta t + \Delta x^2 + \Delta y^2)^2, \tag{4.4.25}$$

where c_{11} is a constant independent of Δt, Δx and Δy.

Combining (4.4.21), (4.4.23) and (4.4.25), we obtain

$$\|\mathbf{e}^{n+s+1}\|_h^2 \leq 4^{\Delta t c} \exp\left(\Delta t\, (c_9 + c_8)\right) \Big[\|\mathbf{e}^{n+s}\|_h^2 + \Delta t\, c_{12}\, (\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big],$$

which after s iterations gives

$$\|\mathbf{e}^{n+s+1}\|_h^2 \leq 4^{\Delta t\, s\, c} \exp\left(\Delta t\, s\ (c_9 + c_8)\right) \Big[\|\mathbf{e}^{n+1}\|_h^2 + \Delta t\, s\, c_{12}\, (\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big], \tag{4.4.26}$$

where $c_{12} = \max\{c_7, 4^{-\Delta t c} \exp\left(-\Delta t\, (c_9 + c_8)\right)\ c_{11}\}$.

Together with Eq. (4.4.19), (4.4.26) becomes

$$\|\mathbf{e}^{n+s+1}\|_h^2 \leq 4^{\Delta t\, (s+1)\, c} \exp\left(\Delta t\, s\ (c_9 + c_8)\right) \Big[\exp\left(\Delta t\, c_{10}\right) \Big[\|\mathbf{e}^{n}\|_h^2$$

$$+ \Delta t c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big] + \Delta t\, s\, c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big].$$

Since $m_0 A \geq 1$, it follows from this inequality that

$$\|\mathbf{e}^{n+s+1}\|_h^2 \leq 4^{\Delta t\,(s+1)\,c} \exp\left[\Delta t\, c_{10}\,(s+1)\right]\Big[\|\mathbf{e}^n\|_h^2 + \Delta t\,(s+1)\, c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big].$$

Thus, after n iterations, we get

$$\begin{aligned}\|\mathbf{e}^{n+s+1}\|_h^2 &\leq 4^{\Delta t\,(n+s+1)\,c} \exp\left[\Delta t\, c_{10}\,(n+s+1)\right]\Big[\|\mathbf{e}^0\|_h^2 \\ &\quad + \Delta t\,(n+s+1)\, c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big] \\ &= 4^{\Delta t\,(n+s+1)\,c} \exp\left[\Delta t\, c_{10}\,(n+s+1)\right]\Big[\Delta t\,(n+s+1)\, c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big].\end{aligned} \tag{4.4.27}$$

For the case $n + s \geq m_0 - 1$, we have

$$\begin{aligned}\|\mathbf{E}^{n+s+1}\|_{3h}^2 &\leq 4^{\Delta t c}\,(1 + \Delta t\, c_9)\,\|\mathbf{E}^{n+s}\|_{3h}^2 + \Delta t\, A\, c_8\, 4^{\Delta t c}\Big[\|\mathbf{E}^{n+s}\|_{3h}^2\|\mathbf{E}^{n+s-1}\|_{3h}^2 + \\ &\quad + \cdots + \|\mathbf{E}^{n+s-m_0+1}\|_{3h}^2\Big] + \Delta t\, c_7\, 4^{\Delta t c}(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2,\end{aligned}$$

which implies

$$\begin{aligned}&\max\Big\{\|\mathbf{E}^{n+s+1}\|_{3h}^2, \ldots, \|\mathbf{E}^{n+s-m_0+2}\|_{3h}^2\Big\} \\ &\qquad \leq 4^{\Delta t c} \exp\left(\Delta t\, c_{10}\right)\ \max\Big\{\|\mathbf{E}^{n+s}\|_{3h}^2, \|\mathbf{E}^{n+s-1}\|_{3h}^2, \ldots, \|\mathbf{E}^{n+s-m_0+1}\|_{3h}^2\Big\} \\ &\qquad\quad + \Delta t\, c_{12}\, 4^{\Delta t c}(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2.\end{aligned} \tag{4.4.28}$$

Using (4.4.23), (4.4.25) and (4.4.28), we obtain

$$\begin{aligned}&\max\Big\{\|\mathbf{e}^{n+s+1}\|_h^2, \ldots, \|\mathbf{e}^{n+s-m_0+2}\|_h^2\Big\} \\ &\qquad \leq 4^{\Delta t\, c} \exp\left(\Delta t\, c_{10}\right)\ \max\Big\{\|\mathbf{e}^{n+s}\|_h^2, \|\mathbf{e}^{n+s-1}\|_h^2, \ldots, \|\mathbf{e}^{n+s-m_0+1}\|_h^2\Big\} \\ &\qquad\quad + \Delta t\, c_{12}\, 4^{\Delta t c}(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2 \\ &\qquad \leq 4^{\Delta t(n+s-m_0+1)\,c} \exp\left[\Delta t\, c_{10}\,(n+s-m_0+1)\right]\Big[\ \max\Big\{\|\mathbf{e}^{m_0-1}\|_h^2, \\ &\qquad\quad \|\mathbf{e}^{m_0-2}\|_h^2, \ldots, \|\mathbf{e}^0\|_h^2\Big\} + \Delta t\,(n+s-m_0+1)\, c_{12}\,(\Delta t + (3\Delta x)^2 + (3\Delta y)^2)^2\Big].\end{aligned} \tag{4.4.29}$$

Using (4.4.27), (4.4.29) and induction on n, one obtains

$$\begin{aligned}\|\mathbf{e}^n\|_h^2 &\leq 4^{\Delta t(n)c}\exp\left[n\,\Delta t\,c_{10}\right]\left[\|\mathbf{e}^0\|_h^2 + n\,\Delta t c_{12}\,(\Delta t+\Delta x^2+\Delta y^2)^2\right]\\ &\leq 4^{Tc}\exp\left[T\,c_{10}\right]\left[Tc_{12}\,(\Delta t+(3\Delta x)^2+(3\Delta y)^2)^2\right],\end{aligned}$$

for $n=1,\ldots,M$. Therefore, we have

$$\|\mathbf{e}^n\|_h \leq C(\Delta t+(3\Delta x)^2+(3\Delta y)^2),$$

where C is a constant independent of $\Delta t, \Delta x$ and Δy. This completes the proof. □

4.4.2 Explicit multilevel finite volume method

For $0\leq r\leq p-1$ and $1\leq s\leq q$, we discretize Eq. (4.1.1) using explicit multilevel finite volume method.

$$\begin{aligned}&\frac{p}{\Delta t}(u_{i,j}^{n+(r+1)/p}-u_{i,j}^{n+r/p})-(C_h(\mathbf{u}^{n+r/p},\mathbf{u}^{n+r/p}))_{i,j}+\varepsilon^2\Delta_h^2 u_{i,j}^{n+r/p}\\ &\qquad=\nabla_{1,h}^+(\varphi_{i-1/2,j}^{n+r/p}\nabla_{1,h}^- u_{i,j}^{n+r+/p})+\nabla_{2,h}^+(\varphi_{i,j-1/2}^{n+r/p}\nabla_{2,h}^- u_{i,j}^{n+r/p}),\end{aligned}\tag{4.4.30a}$$

$$\begin{aligned}&\frac{U_{l,m}^{n+s+1}-U_{l,m}^{n+s}}{\Delta t}-(C_{3h}(\mathbf{U}^{n+s},\mathbf{U}^{n+s}))_{l,m}+\varepsilon^2\Delta_{3h}^2 U_{l,m}^{n+s}\\ &\qquad=\nabla_{1,3h}^+(\Phi_{l-1/2,m}^{n+s}\nabla_{1,3h}^- U_{l,m}^{n+s})+\nabla_{2,3h}^+(\Phi_{l,m-1/2}^{n+s}\nabla_{2,3h}^- U_{l,m}^{n+s}).\end{aligned}\tag{4.4.30b}$$

$$u_{i,j}^{n+r/p}=u_{i+3N_1^0,j}^{n+r/p}=u_{i,j+3N_2^0}^{n+r/p},\tag{4.4.30c}$$

$$U_{l,m}^{n+s}=U_{l+N_1^0,m}^{n+s}=U_{l,m+N_2^0}^{n+s},\tag{4.4.30d}$$

$$u_{i,j}^0=\frac{1}{\Delta x\Delta y}\iint\limits_{k_{i,j}}u^0(x,y)\,dx\,dy,\tag{4.4.30e}$$

where $1\leq i\leq 3N_1^0$, $1\leq l\leq N_1^0$, $1\leq j\leq 3N_2^0$ and $1\leq m\leq N_2^0$.

Theorem 4.4.4. *We assume that the following satisfied for some δ, $0<\delta<1$:*

$$32\Delta t\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)^2\leq\frac{1-\delta}{2\varepsilon^2}\min\{p,81\},\tag{4.4.31}$$

$$16\Delta t\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\leq\varepsilon^2\delta\eta^2(1-\delta)\min\{p,9\},\tag{4.4.32}$$

$$\frac{72\Delta t}{p\Delta x\Delta y}\left((|\alpha_1|+|\alpha_2|)^2+\frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\boldsymbol{u}^0\|_h^2\right)\|\boldsymbol{u}^0\|_h^2\leq\varepsilon^2\delta^2\eta^2\exp\left(\frac{-2T}{\varepsilon^2}\right), \tag{4.4.33}$$

$$\frac{8\Delta t}{\Delta x\Delta y}\left((|\alpha_1|+|\alpha_2|)^2+\frac{4}{81\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\boldsymbol{u}^0\|_h^2\right)\|\boldsymbol{u}^0\|_h^2\leq\varepsilon^2\delta^2\eta^2\exp\left(\frac{-2T}{\varepsilon^2}\right). \tag{4.4.34}$$

Then the multilevel method defined by the equations (4.4.30a)-(4.4.30e) *is* $L^\infty(0,T;\mathcal{H}_h)$ *stale in the following sense:*

$$\|\boldsymbol{u}^n\|^2\leq\exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^{n-1}\|^2\leq\cdots\leq\exp\left(\frac{n\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2\leq\exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2,\quad n=1,2\ldots,M_0,$$

$$\|\boldsymbol{u}^{z(q+1)+r/p}\|^2\leq\exp\left(\frac{r\Delta t}{p\varepsilon^2}\right)\|\boldsymbol{u}^{z(q+1)}\|^2,\quad r=1,2,\ldots,p,$$

where z *is non-negative integer such that* $z\,(q+1)<M_0$.

Proof. To prove this theorem we use the approach discussed in Theorem 4.3.4. We assume n is a multiple of $q+1$. Multiplying Eq. (4.4.30a) by $\frac{2\Delta t\,\Delta x\,\Delta y}{p}\,u_{i,j}^{n+r/p}$ and taking the sum for $i=1,\ldots,3N_1^0$ and $j=1,\ldots,3N_2^0$ together with (4.2.16) and Lemma 4.2.7, we obtain

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2-\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2+\frac{\Delta t}{p}\varepsilon^2|\mathbf{u}^{n+r/p}|_{2,h}^2\leq\exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2. \tag{4.4.35}$$

To estimate the term $\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2$, we multiply Eq. (4.4.30a) by $\frac{2\Delta t\Delta x\Delta y}{p}\,(u_{i,j}^{n+(r+1)/p}-u_{i,j}^{n+r/p})$ and summing from $i=1$ to $i=3N_1^0$ and from $j=1$ to $j=3N_2^0$,, we find

$$\begin{aligned}\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2\leq&\frac{36\Delta t^2}{p^2\,\Delta x\Delta y}(|\alpha_1|+|\alpha_2|)^2\|\mathbf{u}^{n+r/p}\|_h^2|\mathbf{u}^{n+r/p}|_{1,h}^2\\&+\frac{64\Delta t^2}{p^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)^2\varepsilon^4|\mathbf{u}^{n+r/p}|_{2,h}^2\\&+\frac{144\Delta t^2}{p^2\,\Delta x^2\Delta y^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^{n+r/p}\|_h^4|\mathbf{u}^n|_{1,h}^2\\&+\frac{16\Delta t^2}{p^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)|\mathbf{u}^{n+r/p}|_{1,h}^2.\end{aligned}$$

Using (4.4.31), we obtain

$$\|\mathbf{u}^{n+(r+1)/p}-\mathbf{u}^{n+r/p}\|_h^2\leq\frac{36\Delta t^2}{p^2\,\Delta x\Delta y}(|\alpha_1|+|\alpha_2|)^2\|\mathbf{u}^n\|_h^2|\mathbf{u}^{n+r/p}|_{1,h}^2+\frac{\Delta t}{p}\varepsilon^2(1-\delta)|\mathbf{u}^{n+r/p}|_{2,h}^2$$

$$+ 144\frac{\Delta t^2}{p^2\,\Delta x^2\Delta y^2}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^{n+r/p}\|_h^4|\mathbf{u}^{n+r/p}|_{1,h}^2 + \frac{16\Delta t^2}{p^2h^2}|\mathbf{u}^{n+r/p}|_{1,h}^2. \tag{4.4.36}$$

On substitution of (4.4.36) into (4.4.35), we get

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2 + \frac{\Delta t}{p}\varepsilon^2\delta|\mathbf{u}^{n+r/p}|_{2,h}^2 \le \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2 + \frac{36\Delta t^2}{p^2\,\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2 + \frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2\right]\|\mathbf{u}^{n+r/p}\|_h^2|\mathbf{u}^{n+r/p}|_{1,h}^2$$

Using 4.2.25 and (4.4.32), we obtain

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2 - \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2 + \frac{\Delta t}{p}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{n+r/p}|_{1,h}^2 - \frac{36\Delta t^2}{p^2\,\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2 + \frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2\right]\|\mathbf{u}^{n+r/p}\|_h^2|\mathbf{u}^{n+r/p}|_{1,h}^2 \le 0. \tag{4.4.37}$$

In a similar fashion, from (4.4.30b) together with the assumptions (4.4.31), (4.4.32) and (4.4.34), we obtain

$$\|\mathbf{U}^{n+m+1}\|_{3h}^2 - \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^{n+m}\|_{3h}^2 + \Delta t\varepsilon^2\delta^2\eta^2|\mathbf{U}^{n+m}|_{1,3h}^2 \le \frac{4\Delta t^2}{\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2 + \frac{4}{81\,\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{U}^{n+m}\|_{3h}^2\right]\|\mathbf{U}^{n+m}\|_{3h}^2|\mathbf{U}^{n+m}|_{1,3h}^2. \tag{4.4.38}$$

Now we need to prove the following by induction on n

$$\|\mathbf{u}^{n+(r+1)/p}\|_h^2 + \frac{\Delta t}{2p}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{n+r/p}|_{1,h}^2 \le \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{n+r/p}\|_h^2, \quad \text{for } r = 0,1,\ldots,p-1, \tag{4.4.39}$$

$$\|\mathbf{U}^{n+s+1}\|_{3h}^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{U}^{n+s}|_{1,3h}^2 \le \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^{n+s}\|_{3h}^2, \quad \text{for } s = 1,2,\ldots,q. \tag{4.4.40}$$

We first show (4.4.39) and (4.4.40) hold by induction on r and s when $n = 0$. We first show

$$\|\mathbf{u}^1\|_h^2 + \frac{\Delta t}{2p}\varepsilon^2\delta^2\eta^2\sum_{r=0}^{p-1}\exp\left(\frac{(p-1-r)\Delta t}{p\varepsilon^2}\right)|\mathbf{u}^{r/p}|_{1,h}^2 \le \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2, \tag{4.4.41}$$

holds.

For $n = 0$, the relation (4.4.37) becomes

$$\|\mathbf{u}^{(r+1)/p}\|_h^2 + \frac{\Delta t}{p}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{r/p}|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{r/p}\|_h^2 + 36\frac{\Delta t^2}{p^2\,\Delta x\Delta y}\left[(|\alpha_1|+|\alpha_2|)^2 + \frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^{r/p}\|_h^2\right]\|\mathbf{u}^{r/p}\|_h^2|\mathbf{u}^{r/p}|_{1,h}^2. \tag{4.4.42}$$

For $r = 0$ using (4.4.33), we get

$$\|\mathbf{u}^{1/p}\|_h^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{u}^0|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2.$$

Let us assume that (4.4.41) holds up to $r-1$. From the assumption for $s = 1, 2, \ldots, r-1$, we have

$$\|\mathbf{u}^{s/p}\|_h^2 \leq \|\mathbf{u}^{(s-1)/p}\|_h^2$$

and

$$\|\mathbf{u}^{s/p}\|_h^2 \leq \exp\left(\frac{s\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2$$

The relation (4.4.42) becomes

$$\begin{aligned}\|\mathbf{u}^{(r+1)/p}\|_h^2 + \frac{\Delta t}{p}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{r/p}|_{1,h}^2 &\leq \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{r/p}\|_h^2 + 36\frac{\Delta t^2}{p^2\,\Delta x\Delta y}\exp\left(\frac{2r\Delta t}{p\varepsilon^2}\right)\Big[(|\alpha_1|+|\alpha_2|)^2 \\ &\quad + \frac{4}{\Delta x\Delta y}\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)\|\mathbf{u}^0\|_h^2\Big]\|\mathbf{u}^0\|_h^2|\mathbf{u}^{r/p}|_{1,h}^2 \\ &\leq \exp\left(\frac{\Delta t}{p\varepsilon^2}\right)\|\mathbf{u}^{r/p}\|_h^2 + \frac{\Delta t}{2p}\varepsilon^2\delta^2\eta^2|\mathbf{u}^{r/p}|_{1,h}^2, \end{aligned} \tag{4.4.43}$$

which shows us that (4.4.39) is true for $n = 0$. From (4.4.43), we have

$$\|\mathbf{u}^1\|_h^2 + \frac{\Delta t}{2p}\varepsilon^2\delta^2\eta^2\sum_{r=0}^{p-1}\exp\left(\frac{(p-1-r)\Delta t}{p\varepsilon^2}\right)|\mathbf{u}^{r/p}|_{1,h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2,$$

which implies

$$\|\mathbf{u}^1\|_h^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^0\|_h^2. \tag{4.4.44}$$

We then show (4.4.40) by using induction on s for $n = 0$. From the definition of $\mathbf{U}$, we have

$$\|\mathbf{U}^n\|_{3h}^2 \leq \|\mathbf{u}^n\|_h^2. \tag{4.4.45}$$

For $s = 1$, from (4.4.31), we have

$$\|\mathbf{U}^2\|_{3h}^2 - \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^1\|_{3h}^2 + \Delta t\varepsilon^2\delta^2\eta^2|\mathbf{U}^1|_{1,3h}^2 - \frac{4\Delta t^2}{\Delta x\Delta y}\left[(|\alpha_1| + |\alpha_2|)^2 + \frac{4}{81\,\Delta x\Delta y}\left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right)\|\mathbf{U}^1\|_{3h}^2\right]\|\mathbf{U}^1\|_{3h}^2|\mathbf{U}^1|_{1,3h}^2 \leq 0.$$

Then using (4.4.44) and (4.4.45), we have

$$\|\mathbf{U}^2\|_{3h}^2 - \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^1\|_{3h}^2 + \Delta t\varepsilon^2\delta^2\eta^2|\mathbf{U}^1|_{1,3h}^2 - \frac{4\Delta t^2}{\Delta x\Delta y}\exp\left(\frac{2\Delta t}{\varepsilon^2}\right)\left[(|\alpha_1| + |\alpha_2|)^2 + \frac{4}{81\Delta x\Delta y}\left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right)\|\mathbf{u}^0\|_h^2\right]\|\mathbf{u}^0\|_h^2|\mathbf{U}^1|_{1,3h}^2 \leq 0,$$

and using (4.4.34), we arrive at

$$\|\mathbf{U}^2\|_{3h}^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{U}^1|_{1,3h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^1\|_{3h}^2.$$

We now assume that (4.4.40) holds true up to the order $q - 1$, that is

$$\|\mathbf{U}^q\|_{3h}^2 + \frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2|\mathbf{U}^{q-1}|_{1,3h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^{q-1}\|_{3h}^2.$$

and we observe that

$$\|\mathbf{U}^{s+1}\|_{3h}^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{U}^s\|_{3h}^2, \text{ for } s = 1, \ldots, q-1. \tag{4.4.46}$$

From (4.4.38) and (4.4.46) together with (4.4.34) we obtain the result. Thus using (4.4.1) and (4.4.12), we find

$$\|\mathbf{u}^{s+1}\|_h^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\mathbf{u}^s\|_h^2, \text{ for } s = 0, \ldots, q.$$

Now suppose that Eqs. (4.4.39) and (4.4.40) hold up to the order n. Using the same approach as in the case $n = 0$, it can be easily proved by induction on r and s. Hence, (4.4.39) and (4.4.40) hold for any $n = z(q + 1)$, where z is a positive integer. Therefore, the proof is complete. □

Remark 4.4.1. *By the subscript $3h$, we mean the discrete operators, discrete norms and semi-norms are applied on the coarse mesh discretization.*

Remark 4.4.2. *To compare the stability regions of the explicit finite volume methods, we use time steps $\frac{\Delta t}{p}$ on the fine mesh and Δt on the coarse mesh as discussed at the beginning of this section.*

- *When $p \leq 9$, the multilevel method has the same region of stability as the one-level method on the fine mesh but smaller region of stability than the one-level method on the coarse mesh.*
- *When $p \geq 81$, the multilevel method has the same region of stability as the one-level method on the coarse mesh but smaller region of stability than the one-level method on the fine mesh.*
- *When $9 < p < 81$, the multilevel method is less restrictive than the one-level method on the fine mesh and more restrictive than the one-level method on the coarse mesh.*

4.5 Numerical simulations

In this section, some numerical simulations of the 2D convective Cahn-Hilliard equation, (4.1.1), with specified initial condition and periodic boundary conditions at some values of T are presented. All the results are computed in a matlab platform using Windows 10 Intel CORE i3, 6G RAM PC and the parameters are chosen as: $\alpha_1 = \alpha_2 = \frac{1}{6}, p = 5$ and $q = 8$.

For the one-level finite volume methods, we use the following temporal and spatial step sizes

- One-level method on the fine mesh (Fine): time step $\Delta t/p$ and spatial step sizes $\Delta x = \Delta y$.
- One-level method on the coarse mesh (Coarse): time step size Δt and spatial step sizes $3\Delta x = 3\Delta y$.

4.5 Numerical simulations

For the implicit one-level method, we assume that

$$\tilde{\mathbf{u}}^n = \frac{1}{2}\left(\mathbf{u}^n + \mathbf{u}^{n-1}\right), \text{ for } n = 1, 2, \ldots, M-1.$$

Similarly for the implicit multilevel method, for a non-negative integer m and $n = m(q+1)$, we assume the following:

$$\begin{aligned}
\tilde{\mathbf{u}}^{m(q+1)+r/p} &= \frac{1}{2}\left(\mathbf{u}^{m(q+1)+r/p} + \mathbf{u}^{m(q+1)+(r-1)/p}\right), \text{ for } r = 1, \ldots, p-1, \\
\tilde{\mathbf{u}}^{m(q+1)} &= \mathbf{u}^{m(q+1)}, \\
\tilde{\mathbf{U}}^{m(q+1)+s} &= \frac{1}{2}\left(\mathbf{U}^{m(q+1)+s} + \mathbf{U}^{m(q+1)+s-1}\right), \text{ for } s = 1, \ldots, q,
\end{aligned}$$

and for both implicit methods $\tilde{\mathbf{u}}^0 = \mathbf{u}^0$.

To test the numerical methods, we consider the exact solution

$$u(x, y, t) = \sin\left(\frac{2\pi\, x}{L}\right) \sin\left(\frac{2\pi\, y}{L}\right) \cos(2\pi\, t),$$

where $L_1 = L_2 = L = 3$, from which the source term is obtained on substitution of Eq. (4.1.1). As shown by Figs. 4.3-4.4, it is observed that the numerical results obtained using the multilevel finite volume methods are close to the results obtained from one-level methods on the fine mesh as compared to the one-level method on the coarse mesh. There is no need to plot u versus y because of the similarities with u versus x.

Tables 4.1-4.2 show that we can save more time using the multilevel method as compared to the one-level methods on the fine mesh. From the numerical simulations, it is observed that all methods are second order accurate in space and the solutions obtained from the multilevel methods are intermediate between the ones obtained from one-level methods on the fine mesh and on the coarse mesh.

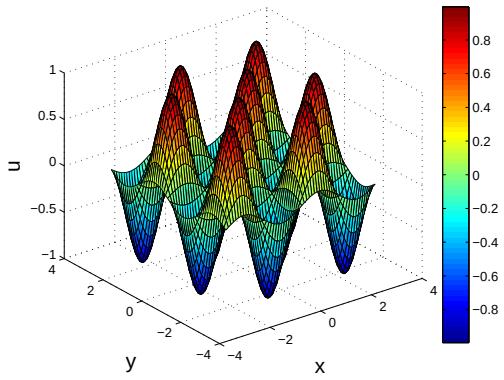
Method	$\Delta x(=\Delta y)$	Δt	L_2-error	CPU time	L_2 Rate
Fine	0.2	0.01	0.0518	1.032	
	0.1	0.0025	0.0131	15.134	1.9594
	0.05	0.000625	0.0033	1232.888	2
Coarse	0.2	0.01	0.3947	0.150	
	0.1	0.0025	0.1128	0.272	1.5661
	0.05	0.000625	0.0291	3.822	1.8806
	0.025	0.00015625	0.0073	159.934	1.9607
Multilevel	0.2	0.01	0.0518	1.341	
	0.1	0.0025	0.0279	4.643	0.8927
	0.05	0.000625	0.0098	271.597	1.5094
	0.025	0.00015625	0.0027	15701.626	1.8598

Table 4.1: Convergence rate, CPU time and L_2-error for some values of spatial step sizes and Δt for the implicit methods at $T = 0.01$.

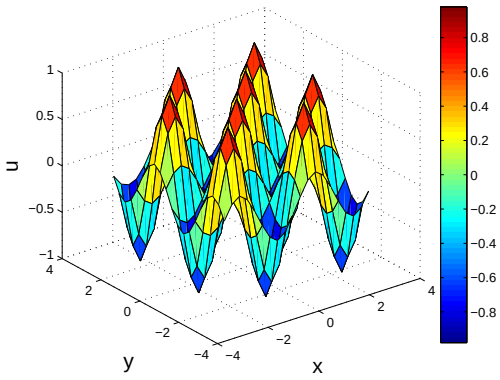
Method	$\Delta x(=\Delta y)$	Δt	L_2-error	CPU time	L_2 Rate
Fine	0.2	0.0002	0.0441	1.765	
	0.1	0.0000125	0.0112	73.639	1.9773
	0.05	0.00000078125	0.0031	10267.945	1.8532
Coarse	0.2	0.0002	0.3749	0.571	
	0.1	0.0000125	0.0983	2.079	1.9312
	0.05	0.00000078125	0.0249	117.542	1.9810
Multilevel	0.2	0.0002	0.0449	0.469	
	0.1	0.0000125	0.0143	10.725	1.6507
	0.05	0.00000078125	0.0043	723.252	1.7336

Table 4.2: Convergence rate, CPU time and L_2-error for some values of spatial step sizes and Δt for the explicit methods at $T = 0.001$.

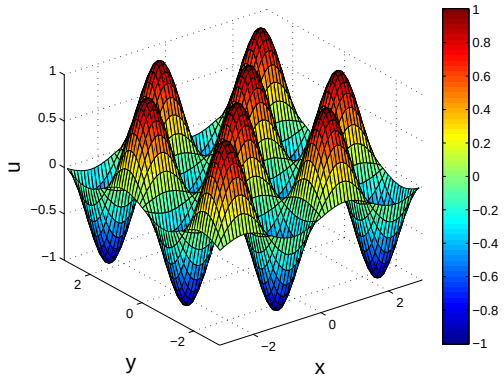
4.5 Numerical simulations



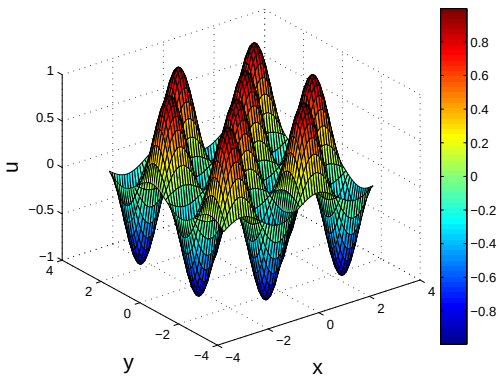
(a) One-level on the fine mesh when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0025$.



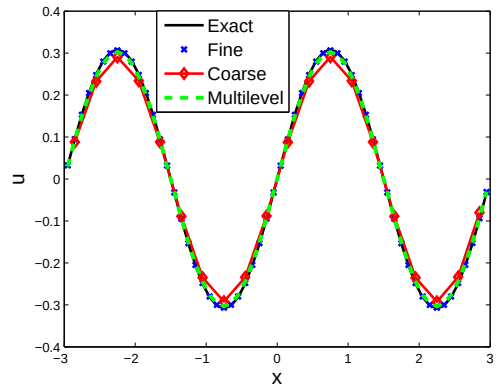
(b) One-level on the coarse mesh when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0025$.



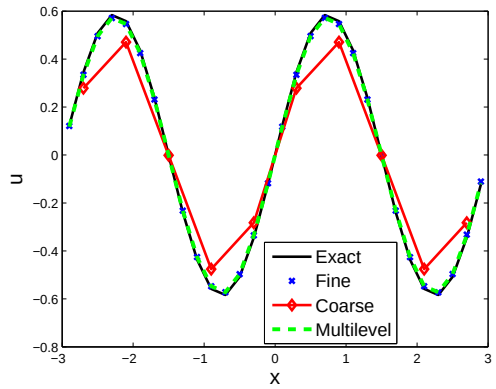
(c) Multilevel method when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0025$.



(d) Exact when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0025$.

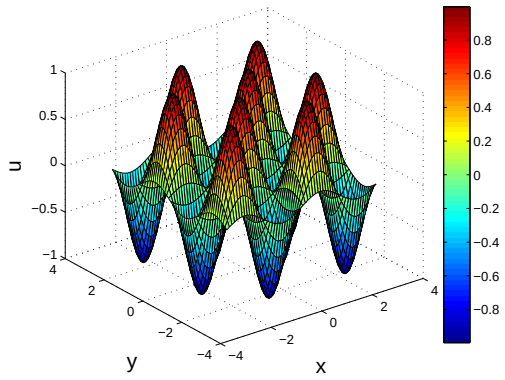


(e) u versus x when $\Delta x = \Delta y = 0.1, \Delta t = 0.0025$ where $y = 0.15$ is centre of the cells.

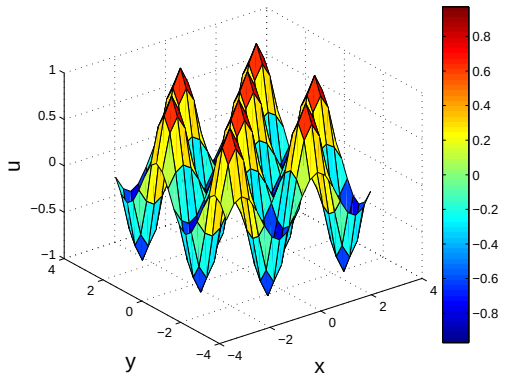


(f) u versus x when $\Delta x = \Delta y = 0.2, \Delta t = 0.01$ where $y = 0.3$ is centre of the cells.

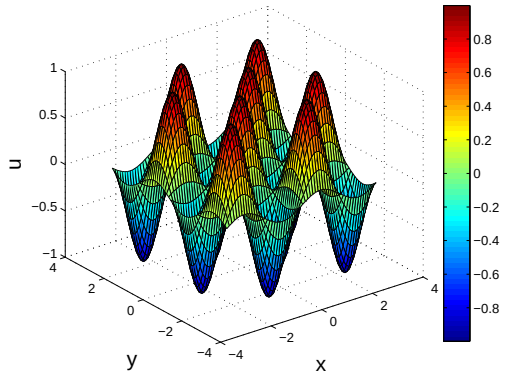
Figure 4.3: Numerical results from implicit methods for some values of Δt and spatial step sizes at $T = 0.01$.



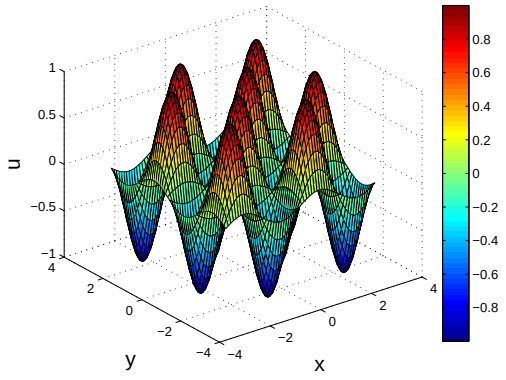
(a) Explicit one-level on the fine mesh when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0000125$.



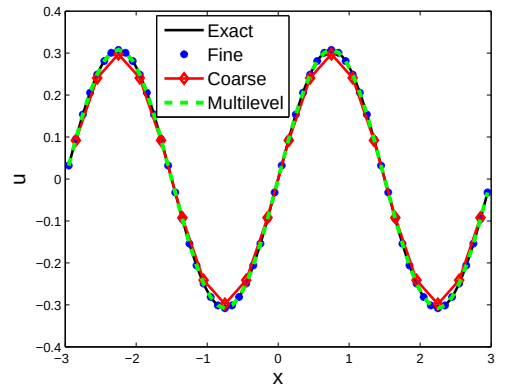
(b) Explicit one-level on the coarse mesh when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0000125$.



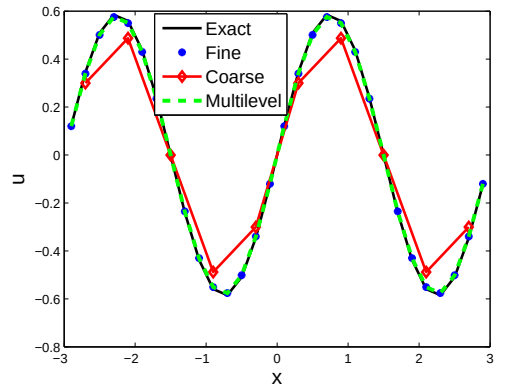
(c) Explicit multilevel method when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0000125$.



(d) Exact when $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0000125$.



(e) u versus x when $\Delta x = \Delta y = 0.1, \Delta t = 0.0000125$ where $y = 0.15$ is centre of the cells.



(f) u versus x when $\Delta x = \Delta y = 0.2, \Delta t = 0.0002$ where $y = 0.3$ is centre of the cells.

Figure 4.4: Numerical results from explicit methods for some values of spatial step sizes and Δt at $T = 0.001$.

4.6 Finite volume methods for 1D convective Cahn-Hilliard equation

In this section, we solve the 1D convective Cahn-Hilliard equation

$$u_t - \gamma u u_x + \varepsilon^2 u_{xxxx} = f(u)_{xx},\ x \in \mathcal{M}, t > 0, \tag{4.6.1}$$

with initial condition

$$u(x,0) = u^0(x), \quad x \in \overline{\mathcal{M}}. \tag{4.6.2}$$

and periodic boundary conditions

$$\frac{\partial^j}{\partial x^j} u(-L,t) = \frac{\partial^j}{\partial x^j} u(L,t), \quad j = 0,1,2,3 \text{ and } 0 < t \leq T. \tag{4.6.3}$$

We partition $\mathcal{M}$ into N cells (intervals) $(k_i)_{1 \leq i \leq N}$ of uniform length Δx with $N\Delta x = 2L$. For $0 \leq i \leq N$

$$x_{i+1/2} = i\Delta x - L, \text{ so that } k_i = (x_{i-1/2}, x_{i+1/2}).$$

We also consider the center of each cell:

$$x_i = \frac{x_{i-1/2} + x_{i+1/2}}{2} = (i-1)\Delta x + \frac{\Delta x}{2} - L, 1 \leq i \leq N.$$

We denote u_i^n as the approximate solution to the cell average of the true solution at $t_n = n\Delta t$ with $M\Delta t = T$, $0 \leq n \leq M$, i.e.

$$u_i^n \approx \frac{1}{\Delta x} \int_{k_i} u(x,t_n) dx, 1 \leq i \leq N,$$

which is obtained recursively by starting with the sequence u_i^0 given by

$$u_i^0 = \frac{1}{\Delta x} \int_{k_i} u^0(x) dx, 1 \leq i \leq N.$$

Define the space $\mathcal{H}_h$ as

$$\mathcal{H}_h = \{\mathbf{u} = (u_i), u_i \in \mathbb{R} | u_{i+N} = u_i, i \in \mathbb{Z} \text{ and } \sum_{i=1}^{N} u_i = 0\}.$$

We equip $\mathcal{H}_h$ with the inner product

$$(\mathbf{u},\mathbf{v})_h = \Delta x \sum_{i=1}^{N} u_i v_i,$$

and discrete L^2 norm

$$\|\mathbf{u}\|_h = \left(\Delta x \sum_{i=1}^{N} u_i^2\right)^{1/2}.$$

For $\mathbf{u},\mathbf{v} \in \mathcal{H}_h$, we introduce the following difference operators:

$$\begin{aligned}
&\nabla_h^- u_i = \frac{1}{\Delta x}(u_i - u_{i-1}),\ \ \nabla_h^+ u_i = \frac{1}{\Delta x}(u_{i+1} - u_i),\\
&\Delta_h u_i = \nabla_h^+(\nabla_h^- u_i) = \frac{1}{\Delta x^2}(u_{i+1} - 2u_i + u_{i-1}),\\
&\Delta_h^2 u_i = \Delta_h(\Delta_h u_i) = \frac{1}{\Delta x^2}(\Delta_h u_{i+1} - 2\Delta_h u_i + \Delta_h u_{i-1}),
\end{aligned}$$

We define the following semi-norm on $\mathcal{H}_h$

$$|\mathbf{u}|_{1,h} = \left(\Delta x \sum_{i=1}^{N} (\nabla_h^- u_i)^2\right)^{\frac{1}{2}}, \tag{4.6.4}$$

In Eq. (4.6.4), ∇_h^- can be replaced by ∇_h^+.

4.6.1 One-level implicit finite volume method

The one-level implicit finite volume discretization of Eqs. (4.6.1)-(4.6.3) is given as follows:

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} - (C_h(\mathbf{u}^{n+1}, \tilde{\mathbf{u}}^n)_i + \varepsilon^2 \Delta_h^2 u_i^{n+1} = \nabla_h^+ \left[\varphi_{i-\frac{1}{2}}^n \nabla_h^- u_i^{n+1}\right], \tag{4.6.5a}$$

$$u_i^n = u_{i+N}^n, \tag{4.6.5b}$$

$$u_i^0 = \frac{1}{\Delta x}\int_{k_i} u^0(x)dx, \tag{4.6.5c}$$

where

$$\varphi_{i-\frac{1}{2}}^n = \frac{f'(u_i^n) + f'(u_{i-1}^n)}{2}.$$

Remark 4.6.1. *Theorems 4.3.1 and 4.3.2 apply for the method* (4.6.5)*. Hence we only state the convergence theorem, its proof is similar to the proof of the corresponding 2D problem.*

Theorem 4.6.1. *Suppose that the solution* $u(x,t)$ *of Eqs.* (4.6.1)-(4.6.3) *is sufficiently smooth. Then, there exist positive constants* c, c_1, c_2 *and* c_3*, independent of* Δt *and* Δx *with*

$$\Delta t < \min(4\varepsilon^2, c)$$

and

$$\frac{(\Delta t + \Delta x^2)^2}{\Delta x} \leq c_2 \varepsilon^2 4^{-Tc_1} \exp\Big(-Tc_3\Big),$$

such that the solution of the finite volume discretization (4.6.5) *converges to the solution of Eq.* (4.6.1) *in the discrete* L^2 *norm with rate of convergence* $\mathcal{O}(\Delta t + \Delta x^2)$*.*

4.6.2 One-level explicit finite volume method

Here, we approximate the solution of Eq. (4.6.1) using an explicit finite volume method:

$$\begin{cases} \dfrac{u_i^{n+1} - u_i^n}{\Delta t} - (C_h(\mathbf{u}^n, \mathbf{u}^n))_i + \varepsilon^2 \Delta_h^2 u_i^n = \nabla_h^+ \left[\varphi_{i-\frac{1}{2}}^n \nabla_h^- u_i^n \right]. \\ u_i^n = u_{i+N}^n \\ u_i^0 = \frac{1}{\Delta x} \int_{k_i} u^0(x) dx, \end{cases} \tag{4.6.6}$$

for $1 \leq n \leq M-1$ and $1 \leq i \leq N$.

The stability theorem is stated as follows, for proof refer to Theorem 4.6.1.

Theorem 4.6.2. *We assume that the following are satisfied for some* $\delta, 0 < \delta < 1$*:*

$$\frac{\Delta t}{\Delta x^4} \leq \frac{1-\delta}{64\varepsilon^2} \tag{4.6.7}$$

$$\frac{16\Delta t}{\Delta x^2} \leq \varepsilon^2 \delta \eta^2 (1-\delta) \tag{4.6.8}$$

$$\frac{72\Delta t}{\Delta x} \left((|\alpha_1| + |\alpha_2|)^2 + \frac{4}{\Delta x^3} \|\boldsymbol{u}^0\|_h^2 \right) \|\boldsymbol{u}^0\|_h^2 \leq \varepsilon^2 \delta^2 \eta^2 \exp\left(\frac{-2T}{\varepsilon^2}\right). \tag{4.6.9}$$

Then the finite volume method defined by (4.6.6) *is stable is* $L^\infty(0,T;\mathcal{H}_h)$ *stable in the following sense:*

$$\|\boldsymbol{u}^n\|_h^2 \leq \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^{n-1}\|_h^2 \leq \cdots \leq \exp\left(\frac{n\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|_h^2 \leq \exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|_h^2, n = 1, 2\ldots, M,$$

$$\frac{\Delta t}{2}\varepsilon^2\delta^2\eta^2\sum_{n=0}^{M-1}\exp\left(\frac{(M-n)\Delta t}{\varepsilon^2}\right)|\boldsymbol{u}^n|_{1,h}^2 \leq \exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|_h^2$$

4.6.3 Multilevel finite volume methods

Let N_0 and M_0 be integers such that $3N_0\Delta x = 2L$ and $\Delta t M_0 = T$. We discretize $\mathcal{M}$ into fine meshes and coarser meshes. The fine mesh consisting of $3N_0$ cells $(k_i)_{1\leq i\leq 3N_0}$ of uniform length Δx and centres x_i, where

$$x_i = (i-1)\Delta x + \frac{\Delta x}{2} - L, 1 \leq i \leq 3N_0.$$

The coarse meshes consisting of N_0 cells $(K_l)_{1\leq l\leq N_0}$ of uniform length $3h$, where

$$K_l = (x_{3l-2-1/2}, x_{3l+1/2}).$$

The approximations of U on the coarse mesh is given by

$$U_l = \frac{1}{3}[u_{3l-2} + u_{3l-1} + u_{3l}], 1 \leq l \leq N_0,$$

and the incremental unknowns are defined by

$$\begin{cases} Z_{3l-2} &= u_{3l-2} - U_l, \\ Z_{3l-1} &= u_{3l-1} - U_l, \\ Z_{3l} &= u_{3l} - U_l. \end{cases}$$

Implicit multilevel finite volume method

For $r = 0, \ldots, p-1$ and $m = 1, 2, \ldots, q$, the following multilevel scheme is used to discretize Eqs. (4.6.1)-(4.6.3).

$$\frac{p}{\Delta t}(u_i^{n+(r+1)/p} - u_i^{n+r/p}) - (C_h(\mathbf{u}^{n+(r+1)/p}, \tilde{\mathbf{u}}^{n+r/p}))_i + \varepsilon^2\Delta_h^2 u_i^{n+(r+1)/p} = \nabla_h^+\left[\varphi_{i-\frac{1}{2}}^{n+r/p}\nabla_h^- u_i^{n+(r+1)/p}\right], \tag{4.6.10a}$$

$$\frac{U_l^{n+m+1}-U_l^{n+m}}{\Delta t}-(C_{3h}(\mathbf{U}^{n+m+1},\tilde{\mathbf{U}}^{n+m}))_l+\varepsilon^2\Delta_{3h}^2U_l^{n+m+1}=\nabla_{3h}^+\left[\Phi_{l-\frac{1}{2}}^{n+m}\nabla_{3h}^-U_l^{n+m+1}\right]. \tag{4.6.10b}$$

$$u_i^{n+(r+1)/p}=u_{i+3N_0}^{n+(r+1)/p} \tag{4.6.10c}$$

$$U_l^{n+m+1}=U_{l+N_0}^{n+m+1} \tag{4.6.10d}$$

$$u_i^0=\frac{1}{\Delta x}\int_{k_i}u^0(x)dx, \tag{4.6.10e}$$

where $1\le i\le 3N_0$ and $1\le l\le N_0$.

Remark 4.6.2. *The uniqueness, stability and convergence analysis can be done similarly as that of the 2D convective Cahn-Hilliard.*

Explicit multilevel finite volume method

For $0\le r\le p-1$ and $1\le m\le q$, we discretize Eq. (4.6.1) using explicit multilevel finite volume method.

$$\frac{p}{\Delta t}(u_i^{n+(r+1)/p}-u_i^{n+r/p})-(C_h(\mathbf{u}^{n+r/p},\mathbf{u}^{n+r/p}))_i+\varepsilon^2\Delta_h^2u_i^{n+r/p}=\nabla_h^+\left[\varphi_{i-\frac{1}{2}}^{n+r/p}\nabla_h^-u_i^{n+r/p}\right], \tag{4.6.11a}$$

$$\frac{U_l^{n+m+1}-U_l^{n+m}}{\Delta t}-(C_{3h}(\mathbf{U}^{n+m},\mathbf{U}^{n+m}))_l+\varepsilon^2\Delta_{3h}^2U_l^{n+m}=\nabla_{3h}^+\left[\Phi_{i-\frac{1}{2}}^{n+r/p}\nabla_{3h}^-U_l^{n+m}\right]. \tag{4.6.11b}$$

$$u_i^{n+r/p}=u_{i+3N_0}^{n+r/p}, \tag{4.6.11c}$$

$$U_l^{n+s}=U_{l+N_0}^{n+s}, \tag{4.6.11d}$$

$$u_i^0=\frac{1}{\Delta x}\int_{k_i}u^0(x)dx, \tag{4.6.11e}$$

where $1\le i\le 3N_0$, $1\le l\le N_0$.

Theorem 4.6.3. *We assume that the following satisfied for some $\delta, 0<\delta<1$:*

$$\begin{aligned}
&\frac{32\Delta t}{\Delta x^4}\le\frac{1-\delta}{2\varepsilon^2}\min\{p,81\}\\
&\frac{16\Delta t}{\Delta x^2}\le\varepsilon^2\delta\eta^2(1-\delta)\min\{p,9\}\\
&\frac{72\Delta t}{p\Delta x}\left((|\alpha_1|+|\alpha_2|)^2+\frac{4}{\Delta x^3}\|\boldsymbol{u}^0\|_h^2\right)\|\boldsymbol{u}^0\|_h^2\le\varepsilon^2\delta^2\eta^2\exp\left(\frac{-2T}{\varepsilon^2}\right),
\end{aligned}$$

$$\frac{24\Delta t}{\Delta x}\left((|\alpha_1|+|\alpha_2|)^2+\frac{4}{27\Delta x^3}\,\|\boldsymbol{u}^0\|_h^2\right)\|\boldsymbol{u}^0\|_h^2 \le \varepsilon^2\delta^2\eta^2\exp\left(\frac{-2T}{\varepsilon^2}\right).$$

Then the multilevel method defined by the equations (4.6.11a) - (4.6.11e) *is conditionally stable in* $L^\infty(0,T;\mathcal{H}_h)$ *in the following sense:*

$$\|\boldsymbol{u}^n\|^2 \le \exp\left(\frac{\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^{n-1}\|^2 \le \cdots \le \exp\left(\frac{n\Delta t}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2 \le \exp\left(\frac{T}{\varepsilon^2}\right)\|\boldsymbol{u}^0\|^2, \quad n=1,2\ldots,M_0,$$

$$\|\boldsymbol{u}^{s(q+1)+r/p}\|^2 \le \exp\left(\frac{r\Delta t}{p\varepsilon^2}\right)\|\boldsymbol{u}^{s(q+1)}\|^2, \quad r=1,2,\ldots,p.$$

4.6.4 Numerical simulations

Here, some numerical simulations of the 1D convective Cahn-Hilliard equation, (4.6.1), with specified initial condition and periodic boundary conditions at some values of T are presented. The parameters are chosen as: $L=3, \varepsilon=0.3, p=5, \alpha_1=\alpha_2=\frac{1}{6}$ and some values of q.

Here, we consider the exact solution

$$u(x,t)=\sin(\frac{2\pi x}{L})\,\cos(2\pi t), \tag{4.6.12}$$

for which the source term is obtained on substitution of u into Eq. (4.6.1).

The convergence rate is calculated based on the relation:

$$\text{Rate}=\log(e_1/e_2)\Big/\log(2),$$

where e_1 and e_2 are L_2-errors when the spatial step sizes are Δx and $\Delta x/2$, respectively. Tables 4.3 and 4.4 show the L_2-errors and the corresponding convergence rates due to each of the finite volume methods for some temporal and spatial step sizes at $T=0.1$. It is shown that all the methods are second order accurate in space. The numerical simulations of the exact solution and the corresponding numerical solutions are also shown by Figs 4.6 and 4.7.

The L_2-error for the multilevel methods is calculated by the formula:

$$L_2\text{-error}=\sqrt{\Delta x\sum_{i=1}^{3N_0}(u_i^{M_0}-v_i^{M_0})^2},$$

Method	Δx	Δt	L_2 error	Rate	CPU time
Fine	0.2	0.01	0.0432		
	0.1	0.0025	0.0111	1.9606	0.406
	0.05	0.000625	0.0028	1.9889	3.115
	0.025	0.00015625	6.9982×10^{-4}	1.9971	24.898
Coarse	0.2	0.01	0.4077		
	0.1	0.0025	0.0872	2.2250	0.117
	0.05	0.000625	0.0228	1.9370	0.282
	0.025	0.00015625	0.0058	1.9743	1.631
Multilevel	0.2	0.01	0.3521		
	0.1	0.0025	0.0623	2.4984	0.261
	0.05	0.000625	0.0151	2.0467	0.978
	0.025	0.00015625	0.0038	1.9984	18.546

Table 4.3: L_2-error and convergence rate of implicit methods when $q = 9$ for some values of Δt and Δx at $T = 0.1$

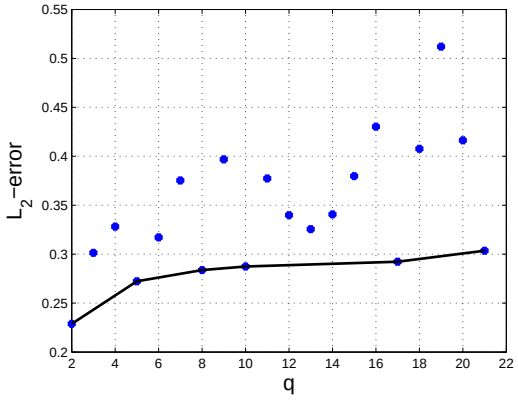
where $u_i^{M_0}$ and $v_i^{M_0}$ are numerical and exact solutions, respectively. Fig. 4.5 shows the 2D plot of q versus L_2-error obtained using the multilevel method for some values of Δx and Δt. Let m be an integer and $0 \leq s < 1$ such that

$$M_0 - 1 = m(q+1) + s(q+1).$$

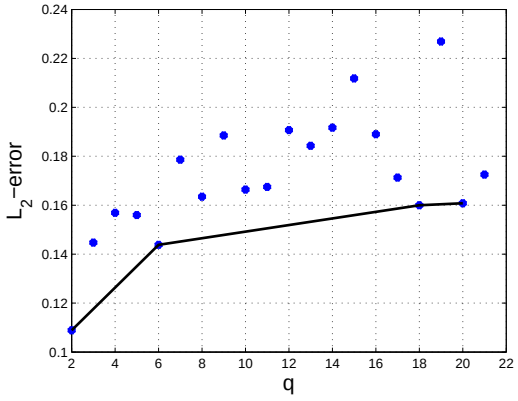
In Fig 4.5, the points, (q, L_2-error(q)), in which $s = 0$ (as in Fig 4.5b) or s is small as compared to the values of s for some neighbouring q's (as in Fig 4.5a), are connected by a curve. It is shown from these figures that the curve is increasing and all the L_2-errors obtained using the multilevel methods lie above these curves. Thus we note that the accuracy of the multilevel method relies on suitable choice of q.

Method	Δx	Δt	L_2 error	Rate	CPU time
Fine	0.2	0.002	0.0348		
	0.1	0.00025	0.0091	1.9364	2.871
	0.05	0.00003125	0.0023	1.9774	45.084
Coarse	0.2	0.002	0.3946		
	0.1	0.00025	0.0778	2.3418	0.331
	0.05	0.00003125	0.0204	1.9327	3.171
Multilevel	0.2	0.002	0.1485		
	0.1	0.00025	0.0287	2.3730	0.794
	0.05	0.00003125	0.0072	1.9935	9.516

Table 4.4: L_2-error and convergence rate of explicit methods for some values of Δt and Δx at $T = 0.1$

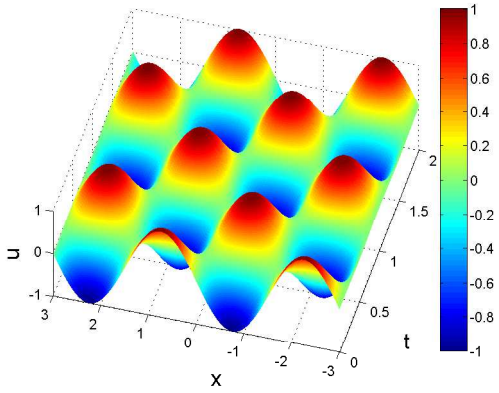


(a) Implicit multilevel when $\Delta x = 0.2, \Delta t = 0.01$ at $T = 2$.

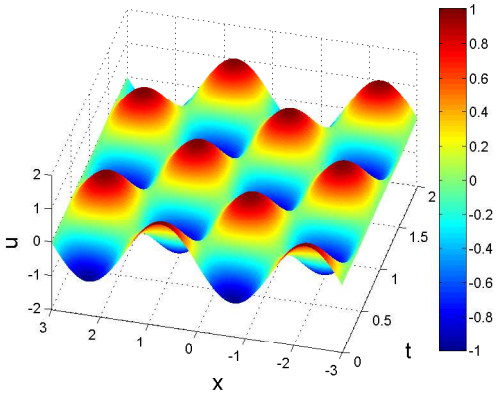


(b) Explicit multilevel when $\Delta x = 0.2, \Delta t = 0.005$ at $T = 2$.

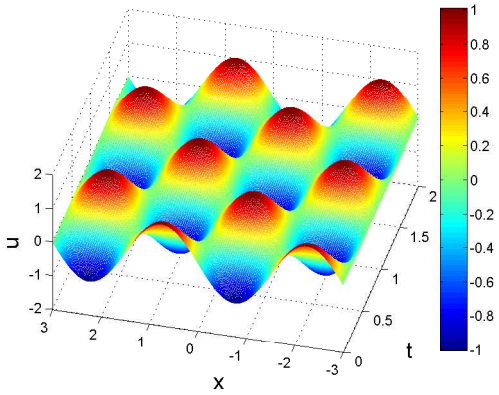
Figure 4.5: q versus L_2-error for multilevel methods



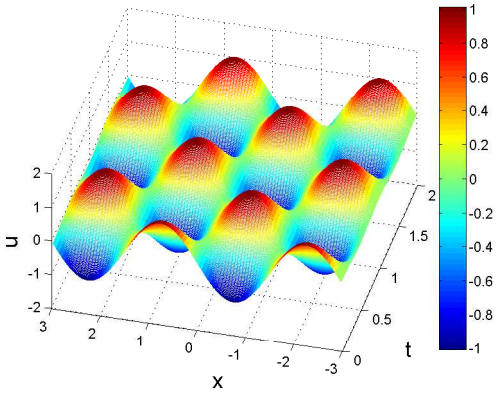
(a) Exact



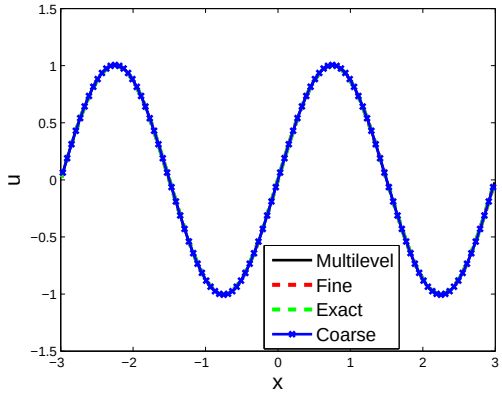
(b) One-level on the fine mesh



(c) Multilevel method

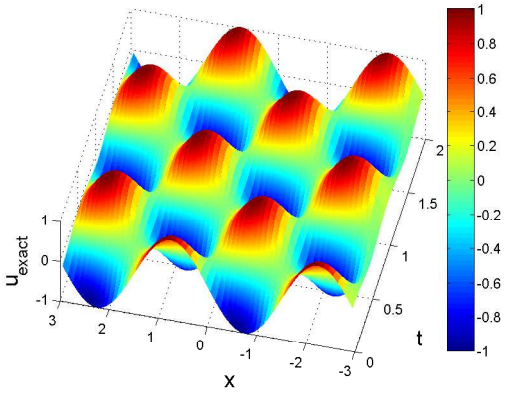


(d) One-level on the coarse mesh

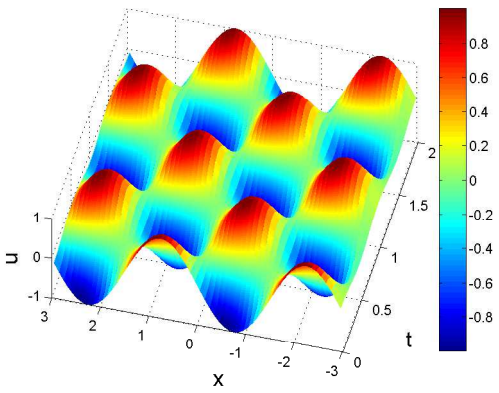


(e) u versus x at $T = 2$

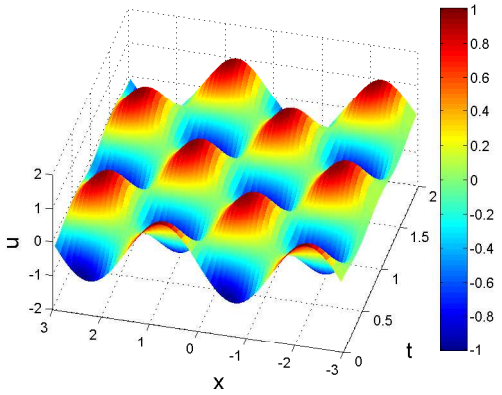
Figure 4.6: Numerical results from implicit one level and multilevel methods when when $q = 9$, $\Delta t = 0.005$ and $\Delta x = 0.02$.



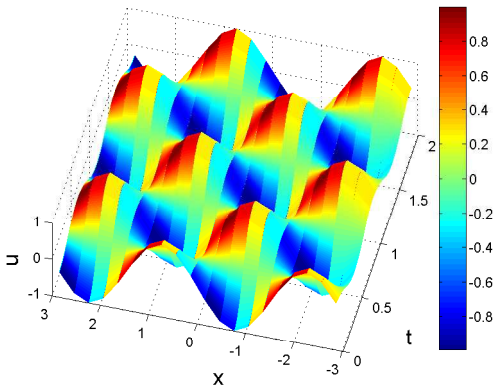
(a) Exact



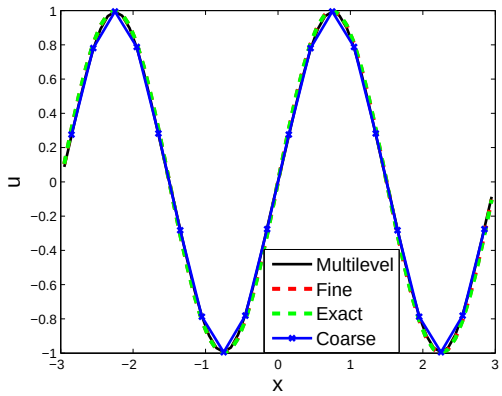
(b) One-level on the fine mesh



(c) Multilevel method



(d) One-level on the coarse mesh



(e) u versus x at $T = 2$

Figure 4.7: Numerical results from explicit one level and multilevel methods when when $q = 9$, $\Delta t = 0.0005$ and $\Delta x = 0.1$.

4.7 Conclusion

We have extended the work of Bousquet et al. (2013a, 2014) in two directions: first, we have considered a nonlinear equation in which the nonlinear term has been linearized following Mickens' rules. Secondly, we have shown that the method can be adapted to higher order partial differential equations. Four numerical methods have been presented and analyzed for the 2D convective Cahn-Hilliard equation. The implicit methods discussed here are linear and easy to implement. Existence, uniqueness of solutions for the schemes formulated are discussed and detailed convergence analysis of implicit schemes is furnished. We also study the stability of these schemes which allow us to make a classification based on region of stability. From the analysis done on the 2D problem, one can easily prove the analysis of the corresponding 1D problem. We compare the multilevel methods with the one-level methods by means of stability, convergence and CPU time. It is shown that the multilevel methods are faster than the one-level methods on the fine mesh. As the numerical experiments reveal, comparing these schemes only with the stability is misleading, hence the CPU time is good indicator for a classification. From the convergence analysis, it is proven that all the methods are second order accurate in space and it is validated by numerical experiments.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Chapter 5

Conclusion and future work

5.1 Conclusion

In this thesis, finite difference and finite volume methods have been used to solve advection-diffusion and convective Cahn-Hilliard equations, respectively. Some properties and applications of these equations are discussed in chapter 1.

In chapter 2, three finite difference methods have been used to solve two test problems described by 1D advection-diffusion equation. The NSFD is much better than third order and fourth order for all the cases considered. An optimization technique has been implemented for the third order scheme when Re=100 and Δx is fixed as 0.1 to find an optimal temporal step that minimizes the dispersion error. From Fig. 2.8a and 2.9, we observe that the errors obtained from the third order upwind scheme initially decreases and reach a minimum when $k = 0.1$ and then increases again. We conclude that the variation of the integrated error in Fig. 2.8a mimic the actual variation of the errors in Fig. 2.9. It is validated from Table 2.12 that the time step, $\Delta t = 0.1$ is indeed the optimal time step size which allows the third order scheme to perform at its best.

In chapter 3, three numerical methods have been used to solve a 3D advection-diffusion problem, with spatial step size, $\Delta x = \Delta y = \Delta z = 0.05$ and temporal step size, $\Delta t = 0.001$

at some values of T. We compute L_2 error, L_∞ error, dispersion error, dissipation error, total mean square error and some performance indices. We observe that as we progress in time, the maximum value of u decreases as expected as the partial differential equation has dissipative terms. Based on the results obtained, we conclude that in general FOM is quite an efficient method to solve the problem. We also extend the optimization technique presented by Appadu and Gidey (2013) for a 3D problem to find an optimal temporal step size to minimize dispersion error when spatial step size is chosen as 0.05, for the case $\alpha_x = \alpha_y = \alpha_z = 0.01, \beta_x = \beta_y = \beta_z = 0.8$. This optimization works well and the optimal time step size is obtained from Figs. 3.5a and 3.5b. The optimal time step is validated using numerical experiments. Indeed, all the various errors are least at the optimal value of time step size as compared to other time step sizes. It is shown that the FOM is fourth order accurate in space and the ICN and ICF are second order accurate in space and time using modified equation and this is tested numerically also by doing convergence tests.

In chapter 4, we have extended the work of Bousquet et al. (2013a, 2014) to a nonlinear fourth order equation in which the nonlinear term has been linearized following Mickens' rules. Four numerical methods have been presented and analyzed. The implicit methods discussed here are linear and easy to implement. Existence, uniqueness of solutions for the schemes formulated are discussed and detailed convergence analysis of implicit schemes is provided. It is shown that the multilevel methods are faster than the one-level methods on the fine mesh. We also study the stability of these schemes which allow us to make a classification based on region of stability. But as the numerical experiments reveal, comparing these schemes only with the stability is misleading, hence the CPU time is good indicator for a classification. It is proven that all the methods are second order accurate in space and this is validated by numerical experiments.

5.2 Future work

Our next plan is to extend this work in the following directions:

5.2 Future work

1. to construct efficient finite difference methods to solve:

   (a) advection-diffusion equations with non-constant advection velocities and diffusivities in Eq. (3.1.1), i. e.

   - $\beta_x = f(x)$,
   - $\beta_x = f(x, y, z), \beta_y = g(x, y, z), \beta_z = h(x, y, z)$,
   - $\beta_x = f(x)$ and $\alpha_x = g(x)$,
   - all coefficients non-constant.

   (b) advection-diffusion with von Neumann boundary conditions.

   (c) system of advection-equations with constant coefficients.

   (d) advection-diffusion equation with nonlinear convective term for instance Burgers' equation.

   (e) system of Burgers' equations.

   (f) advection-diffusion equation in heterogeneous media of the form

   $$\frac{\partial S(x,t)}{\partial t} + v(x, S(x,t)) \cdot \nabla S(x,t) + \nabla \cdot D(x, S(x,t)) \nabla S(x,t) = 0,$$

   where $S(x,t)$ is saturation of an immiscible fluid (Bolster et al., 2009).

2. Concerning the work on the convective Cahn-Hilliard equation, we would like to adapt the multilevel approach discussed;

   (a) to solve the singularly perturbed convective Cahn-Hilliard equation, that is when $\varepsilon \to 0$. It is well documented that naive schemes produce bad results when $\varepsilon \to 0$, so how does the multilevel method can take care of the singularity?

   (b) It is known that the multilevel schemes are cost effective for big system of equations. Hence, the natural follow up of this work is to investigate fluid flow phenomena (Navier-Stokes and its variants).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

References

K. Adamy, A. Bousquet, S. Faure, J. Lamine, and R. Temam. A multilevel method for finite volume discretization of the two-dimensional nonlinear shallow-water equations. *Ocean Modelling*, 33:235–256, 2010.

A. A. Aderogba, M. Chapwanya, and J. K. Djoko. On fractional step-splitting scheme for the Cahn-Hilliard equation. *Engineering Computations*, 31:1151–1168, 2014.

D. A. Anderson, J. C. Tannehill, and R. H. Pletcher. Computational fluid mechanics and heat transfer. 1984.

R. Anguelov and J. M. S. Lubuma. Contributions to the mathematics of the nonstandard finite difference method and applications. *Numerical Methods for Partial Differential Equations*, 17(5):518–543, 2001. ISSN 1098-2426.

A. R. Appadu. The Techniques of MIEELDLD in Computational Aeroacoustics. *Journal of Applied Mathematics*, 2012:Article ID 783101, 30 pages, 2012.

A. R. Appadu. Numerical solution of the 1-D advection-diffusion equation using standard and nonstandard finite difference schemes. *Journal of Applied Mathematics*, vol. 2013: Article Id 734374, 2013. 14 pages.

A. R. Appadu and H. H. Gidey. Time-Splitting Procedures for the Numerical Solution of the 2D Advection-Diffusion Equation. *Mathematical Problems in Engineering*, 2013, 2013. Article ID 634657, 20 pages.

A. R. Appadu, J. K. Djoko, and H. H. Gidey. Analysis of multilevel finite volume ap-

proximation of convective Cahn-Hilliard equation. *Submitted to Afrika Mathematica*, a.

A. R. Appadu, J. K. Djoko, and H. H. Gidey. Comparative study of some numerical methods to solve a 3D advection-diffusion equation. AIP Publishing, b. To appear in Proceedings of the International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2016).

A. R. Appadu, J. K. Djoko, and H. H. Gidey. Performance of some finite difference methods for a 3D advection-diffusion equation. *Submitted to the Journal Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, c.

A. R. Appadu, J. K. Djoko, H. H. Gidey, and J. M. S. Lubuma. Analysis of multilevel finite volume approximation of 2D convective Cahn-Hilliard equation. *Accepted for publication in Japan Journal of Industrial and Applied Mathematics*, d.

A. R. Appadu, M. Z. Dauhoo, and S. D. D. V. Rughooputh. Control of numerical effects of dispersion and dissipation in numerical schemes for efficient shock-capturing through an optimal Courant number. *Computers & Fluids*, 37(6):767–783, 2008.

A. R. Appadu, J. K. Djoko, and H. H. Gidey. A computational study of some numerical schemes for a test case with steep boundary layers. In *Proceedings of the International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2014)*, volume 1648. AIP Publishing, 2015.

A. R. Appadu, J. K. Djoko, and H. H. Gidey. A computational study of three numerical methods for some advection-diffusion problems. *Applied Mathematics and Computation*, 272:629–647, 2016.

J. Berland, C. Bogey, O. Marsden, and C. Bailly. High-order, low dispersive and low dissipative explicit schemes for multiple-scale and boundary problems. *Journal of Computational Physics*, 224:637–662, 2007.

C. Bogey and C. Bailly. A Family of Low Dispersive and Low Dissipative Explicit Schemes

for Computing the Aerodynamic Noise. *Journal of Computational Physics*, 194:194–214, 2004.

D. Bolster, M. Dentz, and J. Carrera. Effective two-phase flow in heterogeneous media under temporal pressure fluctuations. *Water resources research*, 45(5), 2009.

J. P. Boris and D. L. Book. Flux-corrected transport I: SHASTA, a fluid transport algorithm that works. *Journal of Computational Physics*, 11:38–69, 1973.

A. Bousquet and R. Temam. A Finite Volume Multilevel Approximation of the Shallow-Water Equations. In J. C. F. Pereira and A. Sequeira, editors, *European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010*, volume V, 14-17 June 2010.

A. Bousquet, M. Marion, M. Pectu, and R. Temam. Multilevel Finite Volume Methods and Boundary Value Conditions for Geophysical flows. *Computers and Fluids*, 74:66–90, 2013a.

A. Bousquet, M. Marion, and R. Temam. Finite volume multilevel approximation of the shallow water equations. *Chinese Annals of Mathematics, Series B*, 34:1–28, 2013b.

A. Bousquet, M. Marion, and R. Temam. Finite volume multilevel approximation of the shallow water equations with a time explicit scheme. *International Journal of Numerical Analysis and Modelling*, 11(4):762–786, 2014.

A. J. Bray. Theory of Phase-Ordering Kinetics. *Advances in Physics*, 43(3):357–459, 1994.

J. W. Cahn and J. E. Hilliard. Free Energy of a Nonunifrom System: I. Interfacial Free Energy. *The Journal of Chemical Physics*, 28:258–267, 1958.

W. Cahn. Phase separation by spinodal decomposition in isotropic systems. *Journal of Chemical Physics*, 42:93–99, 1965.

C. Canuto, A. Quarteroni, M. Y. Hussaini, and T. A. Zang. *Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics*. Springer Berlin Heidelberg New York, 2007.

REFERENCES

P. C. Chatwin and C. M. Allen. Mathematical models of dispersion in rivers and estuaries. *Annual Review of Fluid Mechanics*, 17:119–149, 1985.

M. H. Chaudhry, D. E. Cass, and J. E. Edinger. Modelling of unsteady-flow water temperatures. *Journal of Hydraulics Engineering*, 109(5):657–669, 1983.

M. M. Chawla, M. A. Al-Zanaidi, and M. G. Al-Aslab. Extended One-Step Time-Integration Schemes for Convection-Diffusion Equations. *Computers and Mathematics with Applications*, 39:71–84, 2000.

M. Chen and R. Temam. Incremental unknowns in finite differences: condition number of the matrix. *SIAM journal on Matrix analysis and applications*, 14(2):432–455, 1993.

M. Dehghan. Numerical solution of the three-dimensional advection-diffusion equation. *Applied Mathematics and Computation*, 150(1):5–19, 2004.

M. Dehghan. On the numerical solution of the one-dimensional convection-difussion equation. *Mathematical Problems in Engineering*, 1:61–74, 2005.

M. Dehghan. Time-splitting procedures for the solution of the two-dimensional transport equation. *Kybernetes*, 36(5/6):791–805, 2007.

H. Ding and Y. Zhang. A new difference scheme with high accuracy and absolute stability for solving convection- diffusion equations. *Journal of Computational and Applied Mathematics*, 230:600–606, 2009.

J. K. Djoko. On the Long-Time Stability of a Backward Euler Scheme for Burgers' Equation with Polynomial Force. *Numerical Methods for Partial Differential Equations*, 24:1371–1387, 2008.

D. R. Durran. *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*. Number 32 in Applied Mathematics. Springer Science and Business Media, New York, 1999.

D. R. Durran. *Numerical Methods for Fluid Dynamics: With applications to Geophysics*. Number 32 in Applied Mathematics. Springer, New York, 2010.

REFERENCES

A. Eden and V. K. Kalantarov. The convective Cahn-Hilliard equation. *Applied Mathematics Letters*, 20:455–461, 2007a.

A. Eden and V. K. Kalantarov. 3d convective Cahn-Hilliard equation. *Communications on Pure and Applied Analysis*, 6(4):1075–1086, 2007b.

M. Ehrhardt and R. E. Mickens. A nonstandard finite difference scheme for convection diffusion equations having constant coefficients. *Applied Mathematics and Computation*, 219:6591–6604, 2013.

F. Eleiwi and T. M. Laleg-Kirati. Dynamic modeling and optimization in membrane distillation system. In *IFAC Proceedings*, volume 47, 2014.

C. M. Elliott and Z. Songmu. On the Cahn-Hilliard equation. *Archive for Rational Mechanics and Analysis*, 96:339–357, 1986.

C. L. Emmott and A. J. Bray. Coarsening dynamics of a one-dimensional driven Cahn-Hilliard system. *Physical Review E*, 54(5):4568–4575, 1996.

S. Faure, J. Laminie, and R. Temam. Finite Volume Discretization and Multilevel Methods in Flow Problems. *Scientific Computing*, 25(1/2):231–261, 2005.

X. F. Feng and Z. F. Tian. Alternating group explicit method with exponential-type for the diffusion-convection equation. *International Journal of Computer Mathematics*, 83: 765–775, 2006.

C. Folas, O. Manley, and R. Temam. On the interaction of small and large eddles in two-dimensional turbulant flows. *Mathematical modelinc and numerical Analysis*, 22:99–114, 1988.

V. G. Ganzha and E. V. Vorozhtsov. *Numerical solutions for partial differential equations: problem solving using Mathematica*, volume 7. CRC Press, 1996.

M. S. Gockenbach. *Finite-Dimensional Linear Algebra*. CRC Press, Taylor and Francis Group, 2010.

REFERENCES

A. A. Golovin, S. H. Davis, and A. A. Nepomnyashcy. A convective Cahn-Hilliard model for the formation of facets and corners in crsytal growth. *Physica D*, 122:202–230, 1998.

A. A. Golovin, S. H. Davis, and A. A. Nepomnyashcy. A convective Cahn-Hilliard model for the formation of facets and corners in crsytal growth. *Physica D*, 198/199:1245–1250, 1999.

A. A. Golovin, A. A. Nepomnyashchy, S. H. Davis, and M. A. Zaks. Convective Cahn-Hilliard models: From coarsening to roughening. *Physical Review Letters*, 86(8):1550–1553, 2001.

V. Guvanasen and R. E. Volker. Numerical Solutions for Solute Transport in Unconfined Aquifers. *International Journal for Numerical methods in Fluids*, 3:103–123, 1983.

Y. He and K. Liu. A Multilevel Finite Element Method in Space-Time for the Navier-Stokes Problem. *Wiley InterScience*, pages 1052–1078, 2005.

A. C. Hindmarsh, P. M. Gresho, and D. F. Griffiths. The stability of explicit euler time-integration for certain finite difference approximations of the multi-dimensional advection–diffusion equation. *International Journal for Numerical Methods in Fluids*, 4: 853–897, 1984.

J. D. Hoffman. *Numerical Methods for Engineers and Scientists*. Marcel Dekker, Inc., 2001.

J. Isenberg and C. Gutfinger. Heat transfer to a draining film. *International Journal of Heat Transfer*, 16:505–513, 1972.

S. Kenneth. An evaluation of several numerical advection schemes. *Atmospheric Environment*, 17(10):1897–1907, 1983.

N. Khiari, T. Achouri, M. L. Ben Mohamed, and K. Omrani. Finite difference approximations for the Cahn-Hilliard equations. *Numerical methods for partial differential equations*, 23: 437–455, 2007.

N. Kumar. Unsteady flow against dispersion in finite porous media. *Journal of Hydrology*, 63:343–358, 1983.

REFERENCES

T. O. Kvalseth. Cautionary Note About R^2. *The American Statistician*, 39(4 (1)):279–285, 1985.

K. Leung. Theory of Morphological Instability in Driven Systems. *Statistical Physics*, 61 (1/2):345–364, 1990.

R. J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2004.

M. Marion and R. Temam. Nonlinear Galerkin methods. *SIAM Journal of Numerical Analysis*, 26(5):1139–1157, 1989.

R. E. Mickens. Analysis of a new finite-difference scheme for the linear advection-diffusion equation. *Journal of Sound and Vibration*, 146:342–344, 1991.

R. E. Mickens. *Nonstandard Finite Difference Models of Differential Equations*. World Scientific, Singapore,, 1994.

R. E. Mickens. *Applications of Nonstandard Finite Difference Schemes*. World Scientific Publishing Co. Pte. Ltd, 2000.

K. W. Morton. *Numerical solution of convection-diffusion equation*. London: Chapman and Hall, 1996.

K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations*. Cambridge University Press, 2005.

K. Nguyen and D. Dabdub. Two-level time-marching scheme using splines for solving the advection equation. *Atmospheric Environment*, 35(9):1627–1637, 2001.

J. Y. Parlarge. Water transport in soils. *Annual Review of Fluid Mechanics*, 2:77–102, 1980.

A. Podolny, M. A. Zaks, B. Y. Rubinstein, A. A. Golovin, and A. A. Nepomnyashchy. Dynamics of domin walls governed by the convective Cahn-Hilliard equation. *Physica D*, 201:291–305, 2005.

REFERENCES

M. Polata, A. O. Celebi, and N. Caliskanb. Global attractors for the 3D viscous Cahn–Hillard equations in an unbounded domain. *Applicable Analysis*, 88(8):1157–1171, 2009.

K. O. Friedrichs R. Courant and H. Lewy. On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11(2), 1967. 215-234.

R. D. Richtmer and K. W. Morton. *Difference Methods for Initial-Value Problems*. Wiley Interscience, New York, 1967.

J. R. Salmon, J. A. Liggett, and R. H. Gallager. Dispersion analysis in homogeneous lakes. *International Journal of Numerical methods in Engineering*, 15:1627–1642, 1980.

J. Siemieniuch and I. Gladwell. Analysis of explicit difference methods for the diffusion-convection equation. *International Journal for Numerical Methods in Engineering*, 12, 1978. 899-916.

G. A. Sod. *Numerical Methods in Fluid dynamics: Initial and Initial Boundary-Value Problems*. Cambridge University Press, 1988.

H. Song. Energy stable and large time-stepping methods for the Cahn-Hilliard equation. *International Journal of Computer Mathematics*, 92(10):2091–2108, 2015.

E. Sousa. The controversial stability analysis. *Applied Mathematics and Computation*, 145: 777–794, 2003.

R. Szymkiewicz. *Numerical Modelling in Open Channel Hydraulics*. Water Science and Technology Library, 2010.

L. L. Takacs. A Two-Step Scheme for the Advection Equation with Minimized Dissipation and Dispersion Errors. *Monthly Weather Review*, 113:1050–1065, 1985.

C. K. W. Tam and H. Shen. Direct Computation of Nonlinear Acoustic pulses using High-order Finite Differences schemes. *AIAA Paper 93-4325*, 1993.

C. K. W. Tam and J. C. Webb. Dispersion-Relation-Preserving Finite Difference Schemes for Computational Acoustics. *International Journal of Computational Physics*, 107:262–281, 1993.

REFERENCES

R. Temam. *Navier-Stokes Equations: Theory and Numerical Analysis*, volume 2. North-Holland, 1979.

R. Temam. Inertial Manifolds and Multigrid methods. *SIAM Journal of Mathematical Analysis*, 21(1):154–178, 1990.

R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68. Springer Science & Business Media, 2012.

Z. F. Tian and P. X. Yu. A high-order exponential scheme for solving 1d convection-diffusion equations. *Journal of Applied Mathematics*, 235:2477–2491, 2011.

H. Trac and U. Pen. A Primer on Eulerian Computational Fluid Dynamics for Astrophysics. *Astronomical Society of the Pacific*, 115:303–321, 2003.

L. N. Trefethen. *Finite Difference and Spectral Methods for Ordinary and Partial Differential Equations*. New York 14853, USA, 1996.

S. J. Watson. Crystal Growth, Coarsening and the Convective Cahn-Hilliard Equation. *International Series of Numerical Mathematics*, 147:329–341, 2003.

S. J. Watson, F. Otto, B. Y. Rubinstein, and S. H. Davis. Coarsening dynamics of the convective Cahn-Hilliard equation. *Physica D*, 178:127–148, 2003.

Y. Yang, L. T. Wilson, M. E. Makela, and M. A. Marchetti. Accuracy of numerical methods for solving the advection-diffusion equation as applied to spore and insect dispersal. *Ecological Modeling*, 109:1–24, 1998.

C. Yeung, T. Rogers, A. Hernandez-Machado, and D. Jasnow. Phase separation dynamics in driven diffusive systems. *Statistical Physics*, 66:1245–1250, 1992.

M. A. Zaks, A. Podolny, A. A. Nepomnyashchy, and A. A. Golovin. Periodic stationary patterns governed by a convective Cahn-Hilliard equation. *SIAM Journal of Applied Mathematics*, 66(2):700–720, 2006.

S. T. Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of Computational Physics*, 31:335–362, 1979.

REFERENCES

X. Zhao. Spectral approximations of attractors for convective Cahn-Hilliard equation in two dimensions. *Bull. Korean Mathematical Society*, 52(5):1445–1465, 2015.

X. Zhao and C. Liu. Optimal control of the convective Cahn-Hilliard equation. *Applicable Analysis*, 92(5):1028–1045, 2013.

X. Zhao and C. Liu. Optimal control of the convective Cahn-Hilliard equation in 2d case. *Applied Mathematcics and Optimization*, 70:61–82, 2014.

Z. Zlatev, R. Berkowicz, and L. P. Prahm. Implementation of a variable step-size variable formula in the time-integration part of a code for treatment of longrange transport of air pollutants. *J. Comput. Phys.*, 55:278–301, 1984.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Appendix A

Taylor's expansion of some nonlinear terms

In this section, we prove the Taylor's expansion given by Eqs. (4.3.17), (4.3.18) and (4.3.19). Note for simplicity that we omit $|_{i,j}^{n}$ on the expanded terms (right hand sides of each equations).

Proof of (4.3.17).

To find the Taylor's expansion of the approximation of the fourth order derivative, we use the relation

$$\Delta_h^2 v_{i,j}^{n+1} = \Delta_{1,h}^2 v_{i,j}^{n+1} + \Delta_{2,h}\Delta_{1,h} v_{i,j}^{n+1} + \Delta_{1,h}\Delta_{2,h} v_{i,j}^{n+1} + \Delta_{2,h}^2 v_{i,j}^{n+1}. \tag{A.1}$$

It is clear that

$$\Delta_{1,h}^2 v_{i,j}^{n+1} = u_{xxxx} + \mathcal{O}(\Delta t + \Delta x^2).$$
$$\Delta_{2,h}^2 v_{i,j}^{n+1} = u_{yyyy} + \mathcal{O}(\Delta t + \Delta y^2).$$

We only find the Taylor's expansion of the second term of the right hand side of Eq. (A.1) and hence the expansion of the third term can be obtained accordingly. Using central difference approximation, we have

$$\Delta_{2,h}\Delta_{1,h} v_{i,j}^{n+1} = \frac{\Delta_{1,h} v_{i,j+1}^{n+1} - 2\Delta_{1,h} v_{i,j}^{n+1} + \Delta_{1,h} v_{i,j-1}^{n+1}}{\Delta y^2}, \tag{A.2}$$

where

$$\Delta_{1,h}v_{i,j+1}^{n+1} = \frac{v_{i-1,j+1}^{n+1} - 2v_{i,j+1}^{n+1} + v_{i+1,j+1}^{n+1}}{\Delta x^2},$$
$$\Delta_{1,h}v_{i,j}^{n+1} = \frac{v_{i-1,j}^{n+1} - 2v_{i,j}^{n+1} + v_{i+1,j}^{n+1}}{\Delta x^2},$$
$$\Delta_{1,h}v_{i,j-1}^{n+1} = \frac{v_{i-1,j-1}^{n+1} - 2v_{i,j-1}^{n+1} + v_{i+1,j-1}^{n+1}}{\Delta x^2}.$$

The Taylor's expansion of the terms $v_{i\pm1,j\pm1}^{n+1}$, $v_{i\pm1,j}^{n+1}$ and $u_{i,j\pm1}^{n+1}$ are given by the following relations.

$$\begin{aligned}
v_{i\pm1,j\pm1}^{n+1} =& u \pm \Delta x \sum_{i=0}^{\infty} \frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}} + \Delta x^2 \sum_{i=1}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}} \pm \Delta y \sum_{i=0}^{\infty} \frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial y^{2i+1}} \\
&+ \Delta y^2 \sum_{i=1}^{\infty} \frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial y^{2i}} + \Delta t \sum_{i=1}^{\infty} \frac{(\Delta t)^{i-1}}{i!}\frac{\partial^{i}u}{\partial t^{i}} \\
&\pm \Delta x \left[\pm \Delta y \sum_{i=0}^{\infty} \frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u_x}{\partial y^{2i+1}} + \Delta y^2 \sum_{i=1}^{\infty} \frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u_x}{\partial y^{2i}} \right. \\
&\left. + \Delta t \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!}\frac{\partial^{j}}{\partial t^{j}} \left(u_x \pm \Delta y \sum_{i=0}^{\infty} \frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u_x}{\partial y^{2i+1}} + \Delta y^2 \sum_{i=1}^{\infty} \frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u_x}{\partial y^{2i}} \right) \right] \\
&+ \frac{\Delta x^2}{2} \left[\pm \Delta y\, u_{xxy} + \frac{\Delta y^2}{2} u_{xxyy} \pm \frac{\Delta y^3}{6} u_{xxyyy} + \Delta t \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!}\frac{\partial^{j}}{\partial t^{j}} \left(u_{xx} \pm \Delta y\, u_{xxy} \right) \right] \\
&\pm \frac{\Delta x^3}{6} \left[\pm \Delta y\, u_{xxxy} + \frac{\Delta y^2}{2} u_{xxxyy} \pm \frac{\Delta y^3}{6} u_{xxxyyy} + \Delta t \sum_{i=1}^{\infty} \frac{(\Delta t)^{i}}{i!}\frac{\partial^{i}}{\partial t^{i-1}} \left(u_{xxx} \pm \Delta y\, u_{xxxy} \right) \right] \\
&\pm \Delta y \frac{\partial}{\partial y} \left(\pm \Delta x \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}} + \Delta x^2 \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}} \right) \\
&+ \Delta t \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!}\frac{\partial^{j}}{\partial t^{j}} \left(\pm \Delta x \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}} + \Delta x^2 \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}} \right) \\
&+ \Delta t \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!}\frac{\partial^{j}}{\partial t^{j}} \left(\pm \Delta y \sum_{i=0}^{\infty} \frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial y^{2i+1}} + \Delta y^2 \sum_{i=1}^{\infty} \frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial y^{2i}} \right) \\
&+ \Delta t \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!}\frac{\partial^{j}}{\partial t^{j}} \left(\pm \Delta y \frac{\partial}{\partial y} \left(\pm \Delta x \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}} + \Delta x^2 \sum_{i=2}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}} \right) \right) \\
&+ \Delta x^2 \Delta y^2\, \mathcal{O}(\Delta x^2 + \Delta y^2). \qquad \text{(A.3)}
\end{aligned}$$

$$v_{i\pm1,j}^{n+1} = u \pm \Delta x \sum_{i=0}^{\infty} \frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}} + \Delta x^2 \sum_{i=1}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}} + \Delta t \sum_{i=1}^{\infty} \frac{(\Delta t)^{i-1}}{i!}\frac{\partial^{i}u}{\partial t^{i}}$$

$$+\Delta t\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\pm\Delta x\sum_{i=0}^{\infty}\frac{(\Delta x)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial x^{2i+1}}+\Delta x^2\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right). \tag{A.4}$$

$$\begin{aligned}v_{i,j\pm1}^{n+1}=&u\pm\Delta y\sum_{i=0}^{\infty}\frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial y^{2i+1}}+\Delta y^2\sum_{i=1}^{\infty}\frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial y^{2i}}+\Delta t\sum_{i=1}^{\infty}\frac{(\Delta t)^{i-1}}{i!}\frac{\partial^i u}{\partial t^i}\\&+\Delta t\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\pm\Delta y\sum_{i=0}^{\infty}\frac{(\Delta y)^{2i}}{(2i+1)!}\frac{\partial^{2i+1}u}{\partial y^{2i+1}}+\Delta y^2\sum_{i=1}^{\infty}\frac{(\Delta y)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial y^{2i}}\right).\end{aligned} \tag{A.5}$$

$$v_{i,j}^{n+1}=u+\Delta t\sum_{i=1}^{\infty}\frac{(\Delta t)^{i-1}}{i!}\frac{\partial^i u}{\partial t^i}. \tag{A.6}$$

Using Eqs. (A.3)-(A.6), we obtain

$$\begin{aligned}v_{i+1,j+1}^{n+1}-2v_{i,j+1}^{n+1}+v_{i-1,j+1}^{n+1}=&\,2\Delta x^2\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}+\frac{\Delta y^2\,\Delta x^2}{2}u_{xxyy}\\&+\Delta x^2\,\Delta t\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(u_{xx}+\Delta y\,u_{xxy}\right)+\Delta x^2\frac{\Delta y^3}{6}u_{xxyyy}\\&+\Delta t\,\Delta x^2\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\sum_{i=2}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right)+2\Delta y\,\Delta x^2\frac{\partial}{\partial y}\left(\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right)\\&+2\Delta t\,\Delta y\,\Delta x^2\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\frac{\partial}{\partial y}\left(\sum_{i=2}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right)\right)\\&+\Delta x^2\Delta y^2\,\mathcal{O}(\Delta t+\Delta x^2+\Delta y^2).\end{aligned}$$

$$v_{i+1,j}^{n+1}-2v_{i,j}^{n+1}+v_{i-1,j}^{n+1}=2\Delta x^2\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}+\Delta t\,\Delta x^2\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right).$$

$$\begin{aligned}v_{i+1,j-1}^{n+1}-2v_{i,j-1}^{n+1}+v_{i-1,j-1}^{n+1}=&\,2\Delta x^2\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}+\frac{\Delta y^2\,\Delta x^2}{2}u_{xxyy}-\Delta x^2\frac{\Delta y^3}{6}u_{xxyyy}\\&+\Delta t\,\Delta x^2\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(u_{xx}-\Delta y\,u_{xxy}\right)-2\Delta y\,\Delta x^2\frac{\partial}{\partial y}\left(\sum_{i=1}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right)\\&+\Delta t\,\Delta x^2\sum_{j=1}^{\infty}\frac{(\Delta t)^{j-1}}{j!}\frac{\partial^j}{\partial t^j}\left(\sum_{i=2}^{\infty}\frac{(\Delta x)^{2i-2}}{(2i)!}\frac{\partial^{2i}u}{\partial x^{2i}}\right)\end{aligned}$$

$$- 2\Delta t\,\Delta y\,\Delta x^2 \sum_{j=1}^{\infty} \frac{(\Delta t)^{j-1}}{j!} \frac{\partial^j}{\partial t^j} \left(\frac{\partial}{\partial y} \left(\sum_{i=2}^{\infty} \frac{(\Delta x)^{2i-2}}{(2i)!} \frac{\partial^{2i} u}{\partial x^{2i}} \right) \right)$$
$$+ \Delta x^2 \Delta y^2\, \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2).$$

Thus

$$\Delta_{2,h}\Delta_{1,h} v_{i,j}^{n+1} = u_{xxyy} + \mathcal{O}(\Delta t + \Delta x^2 + \Delta x \Delta y + \Delta y^2). \tag{A.7}$$

In a similar way, we obtain

$$\Delta_{1,h}\Delta_{2,h} v_{i,j}^{n+1} = u_{yyxx} + \mathcal{O}(\Delta t + \Delta x^2 + \Delta x \Delta y + \Delta y^2). \tag{A.8}$$

Therefore,

$$\Delta_h^2 v_{i,j}^{n+1} = \Delta^2 u|_{i,j}^n + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2).$$

Proof of (4.3.18).

$$\nabla_{1,h}^{+}(\psi_{i-\frac{1}{2},j}^{n} \nabla_{1,h}^{-} v_{i,j}^{n+1}) = \frac{3}{2\Delta x}\left((v_{i+1,j}^{n})^2 - (v_{i-1,j}^{n})^2\right) \frac{\left(v_{i+1,j}^{n+1} - v_{i,j}^{n+1}\right)}{\Delta x}$$
$$+ \frac{3}{2}\left((v_{i,j}^{n})^2 + (v_{i-1,j}^{n})^2\right) \Delta_{1,h} v_{i,j}^{n+1} - \Delta_{1,h} v_{i,j}^{n+1}. \tag{A.9}$$

Clearly $\Delta_{1,h} v_{i,j}^{n+1} = u_{xx} + \mathcal{O}(\Delta t + \Delta x^2)$. The remaining terms of the right hand side of Eq. (A.9) are given as follows:

$$(v_{i+1,j}^{n})^2 - (v_{i-1,j}^{n})^2 = (u + \Delta x u_x + \frac{\Delta x^2}{2} u_{xx})^2 - (u - \Delta x u_x + \frac{\Delta x^2}{2} u_{xx})^2 + \mathcal{O}(\Delta x^3)$$
$$= 4\Delta x u u_x + \mathcal{O}(\Delta x^3); \tag{A.10}$$
$$(v_{i,j}^{n})^2 + (v_{i-1,j}^{n})^2 = 2u^2 - 2\Delta x u u_x + \mathcal{O}(\Delta x^3); \tag{A.11}$$
$$v_{i+1,j}^{n+1} - v_{i,j}^{n+1} = \Delta x u_x + \frac{\Delta x^2}{2} u_{xx} + \mathcal{O}(\Delta x\,\Delta t + \Delta x^3). \tag{A.12}$$

The products are given as follows:

$$\frac{3}{2\Delta x^2}\left((v_{i+1,j}^{n})^2 - (v_{i-1,j}^{n})^2\right)\left(v_{i+1,j}^{n+1} - v_{i,j}^{n+1}\right) = \frac{6uu_x}{\Delta x}\left(\Delta x u_x + \frac{\Delta x^2}{2} u_{xx}\right) + \mathcal{O}(\Delta t + \Delta x^2)$$
$$= 6uu_x^2 + 3\Delta x u u_x\, u_{xx} + \mathcal{O}(\Delta t + \Delta x^2). \tag{A.13}$$
$$\frac{3}{2}\left((v_{i,j}^{n})^2 + (v_{i-1,j}^{n})^2\right)\Delta_{1,h} v_{i,j}^{n+1} = 3u^2 u_{xx} - 3u u_x u_{xx} + \mathcal{O}(\Delta t + \Delta x^2). \tag{A.14}$$

Thus we obtain

$$\nabla^+_{1,h}(\psi^n_{i-\frac{1}{2},j}\nabla^-_{1,h}v^{n+1}_{i,j}) = 6uu_x^2 + 3u^2u_{xx} - u_{xx} + +\mathcal{O}(\Delta t + \Delta x^2). \tag{A.15}$$

In a similar way, we obtain

$$\nabla^+_{2,h}(\psi^n_{i,j-\frac{1}{2}}\nabla^-_{2,h}v^{n+1}_{i,j}) = 6uu_y^2 + 3u^2u_{yy} - u_{yy} + +\mathcal{O}(\Delta t + \Delta y^2).$$

Combining Eqs. (A.15) and (A.16), we obtain

$$\nabla^+_{1,h}(\psi^n_{i-\frac{1}{2},j}\nabla^-_{1,h}v^{n+1}_{i,j}) + \nabla^+_{2,h}(\psi^n_{i,j-\frac{1}{2}}\nabla^-_{2,h}v^{n+1}_{i,j}) = \Delta f(u)|^n_{i,j} + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2). \tag{A.16}$$

Proof of (4.3.19). We recall that

$$\begin{aligned} C_h(\boldsymbol{v}^{n+1}, \tilde{\boldsymbol{v}}^n)_{i,j} =& \frac{\alpha_1}{\Delta x}\left[v^{n+1}_{i+1,j}\tilde{v}^n_{i+1,j} - v^{n+1}_{i-1,j}\tilde{v}^n_{i,j}\right]\right] + \frac{\alpha_2}{\Delta x}\left[v^{n+1}_{i+1,j}\tilde{v}^n_{i,j} - v^{n+1}_{i-1,j}\tilde{v}^n_{i-1,j}\right] \\ &+ \frac{\alpha_1}{\Delta y}\left[v^{n+1}_{i,j+1}\tilde{v}^n_{i,j+1} - v^{n+1}_{i,j-1}\tilde{v}^n_{i,j} + \frac{\alpha_2}{\Delta y}\left[v^{n+1}_{i,j+1}\tilde{v}^n_{i,j} - v^{n+1}_{i,j-1}\tilde{v}^n_{i,j-1}\right]\right]. \end{aligned} \tag{A.17}$$

The Taylor's expansion of the terms on the right hand side of Eq. (A.17) are given as follows

$$\begin{aligned} \tilde{v}^n_{i,j} &= c_0 u + \sum_{i=1}^{\infty} \frac{(-\Delta t)^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i}; \\ \tilde{v}^n_{i+1,j} &= c_0\left[u + \Delta x u_x + \frac{\Delta x^2}{2}u_{xx}\right] + \sum_{i=1}^{\infty} \frac{(-\Delta t)^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i} + \mathcal{O}(\Delta x\Delta t + \Delta x^3); \\ \tilde{v}^n_{i-1,j} &= c_0\left[u - \Delta x u_x + \frac{\Delta x^2}{2}u_{xx}\right] + \sum_{i=1}^{\infty} \frac{(-\Delta t)^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i} + \mathcal{O}(\Delta x\Delta t + \Delta x^3); \\ v^{n+1}_{i+1,j} &= u + \Delta x u_x + \frac{\Delta x^2}{2}u_{xx} + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i} + \mathcal{O}(\Delta x\Delta t + \Delta x^3); \\ v^{n+1}_{i-1,j} &= u - \Delta x u_x + \frac{\Delta x^2}{2}u_{xx} + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i} + \mathcal{O}(\Delta x\Delta t + \Delta x^3); \\ v^{n+1}_{i,j} &= u + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!}\frac{\partial^i u}{\partial t^i}. \end{aligned}$$

where

$$c_0 = a_1 + a_2 + \cdots + a_{m_0};$$

$$c_i = \sum_{j=2}^{m_0} (j-1)^i a_j, i = 1, 2, \ldots.$$

$$\begin{aligned} v_{i+1,j}^{n+1}\tilde{v}_{i+1,j}^{n} =& c_0 \left[u + \Delta x u_x + \frac{\Delta x^2}{2} u_{xx} + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i}\right] u + \sum_{i=1}^{\infty} \frac{(-\Delta t)^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i} \left(u + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i}\right) \\ &+ \Delta x\, c_0 \big[u + \Delta x u_x\big]\, u_x + \frac{c_0 \Delta x^2}{2} u\, u_{xx} + \mathcal{O}(\Delta x \Delta t + \Delta x^3). \end{aligned}$$

$$\begin{aligned} v_{i-1,j}^{n+1}\tilde{v}_{i,j}^{n} =& c_0 \left[u - \Delta x u_x + \frac{\Delta x^2}{2} u_{xx} + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i}\right] u + \sum_{i=1}^{\infty} \frac{(-\Delta t)^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i} \left(u + \sum_{i=1}^{\infty} \frac{\Delta t^i\, c_i}{i!} \frac{\partial^i u}{\partial t^i}\right) \\ &+ \mathcal{O}(\Delta x \Delta t + \Delta x^3). \end{aligned}$$

Thus we have

$$\frac{v_{i+1,j}^{n+1}\tilde{v}_{i+1,j}^{n} - v_{i-1,j}^{n+1}\tilde{v}_{i,j}^{n}}{\Delta x} = c_0 \left[3uu_x + \frac{\Delta x}{2} u\, u_{xx} + \Delta x\, u_x^2\right] + \mathcal{O}(\Delta t + \Delta x^2).$$

Similarly, we obtain

$$\frac{v_{i+1,j}^{n+1}\tilde{v}_{i,j}^{n} - v_{i-1,j}^{n+1}\tilde{v}_{i-1,j}^{n}}{\Delta x} = c_0 \left[3uu_x - \frac{\Delta x}{2} u\, u_{xx} - \Delta x\, u_x^2\right] + \mathcal{O}(\Delta t + \Delta x^2).$$

Hence

$$\begin{aligned} \frac{\alpha_1}{\Delta x}\left[v_{i+1,j}^{n+1}\tilde{v}_{i+1,j}^{n} - v_{i-1,j}^{n+1}\tilde{v}_{i,j}^{n}\right] + \frac{\alpha_2}{\Delta x}\left[v_{i+1,j}^{n+1}\tilde{v}_{i,j}^{n} - v_{i-1,j}^{n+1}\tilde{v}_{i-1,j}^{n}\right] &= 3c_0(\alpha_1+\alpha_2)\, uu_x \\ &+ \frac{\Delta x\, c_0}{2}(\alpha_1 - \alpha_2)\, u\, u_{xx} + \Delta x c_0(\alpha_1 - \alpha_2)\, u_x^2 + \mathcal{O}(\Delta t + \Delta x^2). \end{aligned} \tag{A.18}$$

In a similar way, one obtains

$$\begin{aligned} \frac{\alpha_1}{\Delta y}\left[v_{i,j+1}^{n+1}\tilde{v}_{i,j+1}^{n} - v_{i,j-1}^{n+1}\tilde{v}_{i,j}^{n}\right] + \frac{\alpha_2}{\Delta y}\left[v_{i,j+1}^{n+1}\tilde{v}_{i,j}^{n} - v_{i,j-1}^{n+1}\tilde{v}_{i,j-1}^{n}\right] &= 3c_0\,(\alpha_1+\alpha_2)\, uu_y \\ &+ \frac{\Delta y\, c_0}{2}(\alpha_1 - \alpha_2)\, u\, u_{yy} + \Delta y\, c_0(\alpha_1 - \alpha_2)\, u_y^2 + \mathcal{O}(\Delta t + \Delta y^2). \end{aligned} \tag{A.19}$$

Combining Eqs. (A.18) and (A.19), we obtain

$$\begin{aligned} (C_h(\boldsymbol{v}^{n+1}, \tilde{\boldsymbol{v}}^n))_{i,j} =& 3c_0(\alpha_1+\alpha_2)\,(u\,(\boldsymbol{\beta}\cdot\nabla)u)|_{i,j}^n + \frac{c_0}{2}(\alpha_1-\alpha_2)\,(\Delta x\,(u\,u_{xx})|_{i,j}^n + \Delta y\,(u\,u_{yy})|_{i,j}^n) \\ &+ c_0(\alpha_1-\alpha_2)\,(\Delta x\,(u_x^2)|_{i,j}^n + \Delta y\,(u_y^2)|_{i,j}^n) + \mathcal{O}(\Delta t + \Delta x^2 + \Delta y^2). \end{aligned}$$