NOVEL FITTED MULTI-POINT FLUX APPROXIMATION METHODS FOR OPTIONS PRICING

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mathematics and Applied Mathematics

UNIVERSITY OF CAPE TOWN

by

ROCK STEPHANE KOFFI

Supervised by:

Assoc. Prof ANTOINE TAMBUE

and

Dr FRANCOIS EBOBISSE

February 2020

Declaration

I, ROCK STEPHANE KOFFI, hereby declare that the work on which this thesis is based is my original work (except where indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University of Cape Town to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

SIGNATURE: Signed by candidate Date: 05/02/2020

Inclusion of Publication

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication(s) in my PhD thesis, and where co-authorships are involved, my co-authors have agreed that I may include the publication(s):

1. A fitted multi-point flux approximation method for pricing two options.
   Rock Stephane Koffi and Antoine Tambue. Comput Econ (2019). https ://doi.org/10.1007/s10614-019-09906-x,
   ArXiv preprint arXiv :1905.05052,2019c.

2. A fitted l-multi-point flux approximation method for pricing options
   Rock Stephane Koffi and Antoine Tambue.
   ArXiv preprint arXiv :1912.1274v1,2019b.

3. Convergence of the two point flux approximation method for pricing options.
   Rock Stephane Koffi and Antoine Tambue.
   ArXiv preprint arXiv :1912.12737,2019a.

SIGNATURE: Signed by candidate Date: 05/02/2020

Student Name: ROCK STEPHANE KOFFI Student Number: KFFROC001

Dedication

To my late father KOFFI KOUAME INNOCENT

To my late aunt GOLE Ahou Catherine

To my loving and supportive mother GORE AHOU ODETTE

Abstract

It is well known that pricing options in finance generally leads to the resolution of the second order Black-Scholes Partial Differential Equation (PDE). Several studies have been conducted to solve this PDE for pricing different type of financial options. However the Black-Scholes PDE has an analytical solution only for pricing European options with constant coefficients. Therefore, the resolution of the Black-Scholes PDE strongly relies on numerical methods. The finite difference method and the finite volume method are amongst the most used numerical methods for its resolution. Besides, the Black-Scholes PDE is degenerated when stock price approaches zero. This degeneracy affects negatively the accuracy of the numerical method used for its resolution, and therefore special techniques are needed to tackle this drawback. In this Thesis, our goal is to build accurate numerical methods to solve the multi-dimensional degenerated Black-Scholes PDE. More precisely, we develop in two dimensional domain novel numerical methods called fitted Multi-Point Flux Approximation (MPFA) methods to solve the multi-dimensional Black-Scholes PDE for pricing American and European options. We investigate two types of MPFA methods, the O-method which is the classical MPFA method and the most intuitive method, and the L-method which is less intuitive, but seems to be more robust. Furthermore, we provide rigorous convergence proofs of a fully discretized schemes for the one dimensional case of the corresponding schemes, which will be well known on the name of finite volume method with Two Point Flux Approximation (TPFA) and the fitted TPFA. Numerical experiments are performed and proved that the fitted MPFA methods are more accurate than the classical finite volume method and the standard MPFA methods.

Acknowledgement

Firstly, I would like to express my deepest gratitude to my supervisor Associate Professor Antoine TAMBUE, for his guidance, teaching and availability throughout this thesis. Through him, I have learnt what it takes to be a scientist at an international level. I would like also to thank my co-supervisor Dr Francois Ebobisse for his advices, encouragements and important inputs during this Thesis. I will always be grateful to you, supervisors.

Secondly, my gratitude goes to Prof Neil Turok for his vision in creating the African Institute for Mathematical Sciences (AIMS). This institute has played a major role in my life by training me academically and giving me another vision of Africa. I would like to thank also Prof Barry Green and the staff at AIMS-SA for creating a conducive environment for my research. Moreover, I am grateful to the Robert Bosch Stiftung through the AIMS ARETE Chair programme (Grant No 11.5.8040.0033.0) for financial support.
I would like also to thank the University of Cape Town, especially the department of Mathematics and Applied Mathematics for the support and for offering me the international and refugee scholarship through the UCT postgraduate office.
I am also thankful to the Western Norway University of Applied Sciences for funding my supervisor trips from Norway to South Africa, allowing physical meeting with my supervisor.
I am very grateful to Prof Daouda Sangare, Prof Adama Coulibaly, Prof Gozo Yoro, Prof Kangni Kinvi and Prof Youssouf Diagana, my lecturers back in COTE D'IVOIRE for the training in Mathematics.

A special mention to my PhD companion, David Attipoe for mutual support through pain and joy. I take this opportunity to thank Buri Gershom, Trust Chibawara, Mhlasakululeka Mvubu, Kossi Etekpo, Dr Yae Gaba, Zoe Hamel, Dr Landry Guessan and my beloved Gloria Burengengwa.
I would like to express my sincere gratitude to Prof Phillip Mashele for his advice and opportunities, Dr Rejoyce-Gahvi Molefe for encouragements and Dr Dennis IKPE for fruitful discussions.

I am deeply endebted to my family for the support during this Thesis. I will not be able to thank enough my mother, GORE AHOU ODETTE for unconditional love and support. I would like to thank Mr Kouakou Kan Marc; my brothers and sisters, Alex, Ruth, Eunice, Delaure, Evrard, Martin, Benedicte, Jean-Luc, Jean-Marc, Christelle and the family KOFFI without forgetting my uncle Alphonse KOFFI and my grand father Lazare KOFFI.

Contents

Introduction

A financial market is a market whereby investors, companies and governments meet to trade financial securities such as bonds, stocks, precious metal. A good number of transactions on financial markets are about buying and selling options. Indeed, an option is a contract which gives the right to buy (call) or to sell (put) an underlying asset at an agreed price (strike) on (European options) or before (American options) a specified date (maturity). In their seminal paper, Black and Scholes [1973], Fisher Black and Myron Scholes stated the famous Black-Scholes model. Under some assumptions, the derivation of the Black-Scholes model leads to a second order parabolic Partial Differential Equation (PDE) with respect to time and the underlying stock price. An analytical solution has been found for the Black-Scholes PDE only for pricing European options with constant coefficients. Moreover, pricing multi-assets options is of great interest in the financial industry (see Persson and von Persson [2007]). Multi-asset options are options based on more than one underlying. There are several kinds of multi-assets options, few of them are exchange options, rainbow options, baskets options, best or worst options, quotient options, foreign exchange options, quanto options, spread options, dual-strike options and out-performance options. Pricing these options leads to the resolution of the following second order degenerated Black-Scholes Partial Differential Equations (PDE)(see Persson and von Persson [2007])

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{ij} S_i S_j \frac{\partial^2 U}{\partial S_i \partial S_j} + r\sum_{i=1}^{n} S_i \frac{\partial U}{\partial S_i} - rU, \tag{1}$$

where r is the risk free interest, U is the option value at time τ, $\tau = T - t$ with t and T respectively the instantaneous and maturity time, S_i represents the asset i price, σ_i represents the volatility of asset i, ρ_{ij} represents the correlation between the assets i and j, with $i, j = 1, ..., n$. The main difference between multi-assets options is their payoff functions which represents the initial condition of the corresponding backward PDE. The spatial domain of the PDE is infinite, but for its numerical resolution, a truncation is required [Duffy, 2013, Chapter 3]. It has been observed that when the stock price S approaches the region near to zero, the Black-Scholes PDE is degenerated [Duffy, 2013, Chapter 30.3]. Moreover, the initial condition of the PDE has a discontinuity in its first derivative when the stock price is equal to the strike K. This discontinuity has an adverse impact on the accuracy when the finite difference method is used [Wilmott, 2005, Chapter 26]. Therefore, for the spatial discretization of the PDE, it is suitable to use non-uniform grids with more points in the region around $S = 0$ and $S = K$ in order to handle the degeneracy and the discontinuity. To overcome the above challenges, many methods have been proposed in the literature. In Wang [2004], a fitted finite volume method for one dimensional Black-Scholes PDE is proposed and rigorous convergence proof is provided. In Huang et al. [2006], a fitted finite volume discretization method for the two-dimensional Black-Scholes PDE is proposed and its rigorous convergence proof is analysed in Huang et al. [2009]. Although these two fitted finite volume methods are stable, they are only order 1 with respect to asset price variables.

In this Thesis, we present a novel discretization method for the Black Scholes PDE based on a special kind of finite volume method, the so-called Multi-Point Flux Approximation (MPFA) method. This method was introduced by Aavastmark (see Aavatsmark [2002]) and has been used in fluid dynamics for flow and transport equations (see Sandve et al. [2012] and references therein). Actually, the MPFA was designed to give a correct discretization of the flow equation for general grids including

fractures (see Aavatsmark [2002], Sandve et al. [2012]). The MPFA method is essentially based on the approximation of a linear function gradient over a triangle, the calculation and the continuity of flux through edges of this triangle. The convergence of MPFA method is usually second order in space domain on rough grids (see Aavatsmark [2007], Stephansen [2012]). Note that the MPFA is only used to approximate the second order operator. To build our new fitted MPFA method, we couple the standard MPFA with the upwind methods (first and second order) to approximate two dimensional options pricing operator. Besides, a fitted finite volume Huang et al. [2006] is used to handle the degeneracy of the PDE in the region where the stocks price approach zero (degeneracy region). In the region, where the PDE is not degenerated, we apply the MPFA method. The novel numerical technique from this combination is called fitted MPFA method and will obviously improve the accuracy of the current fitted finite volume in the literature, since more approximations involved are second order in space. Naturally, this method is applicable to other types of multi-asset options and also to financial models in incomplete markets such as jump-diffusion model (see Merton [1976]), Heston model (see Heston [1993]), and Bates Model (see Bates [1996]) on non-uniform grids.

This Thesis is subdivised in four Chapters whereby concepts and ideas are developed in details. We start, in Chapter 1, by presenting some basic mathematical tools in probability and option pricing theory. Moreover, we derive the multi-dimensional Black Scholes PDE and we set up the mathematical framework to prove the existence and uniqueness of the solution for the continuous option pricing problem.

The main contributions of this Thesis will be found in Chapters 2, 3 and 4. Thereby, in Chapter 2, we present the Two Point Flux Approximation (TPFA) method and the fitted Two Point Flux Approximation for solving the one dimensional degenerated Black Scholes PDE. Indeed, the fitted Two Point Flux Approximation method is the combination of the fitted finite volume introduced by Wang [2004] and the TPFA method. We provide a rigorous proof of convergence for the TPFA method and the fitted TPFA method. Numerical experiments are performed to confirm theoretical results.

Furthermore, in Chapter 3, we introduce a novel numerical method called the fitted Multi-Point Flux Approximation (MPFA) method to solve the two-dimensional degenerated Black-Scholes PDE for pricing European options. Like the fitted TPFA method, the fitted MPFA is also a combination of the fitted finite volume method and the MPFA method. Our focus will be on a type of MPFA method called O-MPFA method by giving details about the geometrical contruction of the method. Numerical experiments are provided and show that the fitted O-MPFA methods are more accurate than the O-MPFA methods which are also more accurate than the standard fitted finite volume the method.

Finally, we introduce another kind of Multi-Point Flux Approximation which is the L-MPFA method in Chapter 4. The L-MPFA methods and the fitted L-MPFA methods are applied to solve the Black-Scholes PDE for pricing American options. We provide numerical experiments that show, on one hand, that L-MPFA methods are more accurate than O-MPFA methods and the standard fitted finite volume method; On the other hand, we also show that the fitted L-MPFA methods are more accurate than L-MPFA methods.

Chapter 1

Options pricing problem: Basic notion

In this Chapter, we first recall some notions in probability theory and stochastic calculus in order to set a convenient mathematical framework for our study. Afterwards, we present the famous Black-Scholes model with its corresponding assumptions which will lead to the derivation of the multi-dimensional Black-Scholes Partial Differential Equation. Existence and uniqueness of the solution of the continuous option pricing problem are proven with help of Sobolev spaces.

1.1 Preliminaries

Let us introduce here some notions and definitions in probability theory and stochastic calculus which will be useful in our work.

1.1.1 Notion in Probability

Definition 1 [$\sigma - field$, *measurable space*] *Oksendal [1992]*
A class $\mathcal{F}$ of subsets of a given set Ω is a σ-field (or σ-algebra) of subsets of Ω if:

i) $\Omega \in \mathcal{F}$,

ii) $If \quad A \in \mathcal{F} \quad then \quad A^c \in \mathcal{F}$, *where* $A^c = \Omega \setminus A$ *is the complement of* A *in* Ω,

iii) $If \quad \{A_n, n \in \mathbb{N}\} \subset \mathcal{F} \quad then \quad \cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

The couple $(\Omega, \mathcal{F})$ is called measurable space. Moreover, the σ-field, $\sigma(\mathcal{F})$, generated by the class $\mathcal{F}$ is the smallest σ-field containing the class $\mathcal{F}$. Indeed

$$\sigma(\mathcal{F}) = \cap\{\mathfrak{F} : \mathfrak{F} \quad \sigma - field, \mathcal{F} \subset \mathfrak{F}\}. \tag{1.1}$$

The σ-field generated by the class of intervals in $\mathbb{R}$ is called the Borel σ-field denoted by $\mathcal{B}$.
Furthermore, an increasing sequence $\mathbb{F} = (\mathcal{F}_n)_{n \geqslant 0}$ of sub-σ-field of the σ-field $\mathcal{F}$ is a filtration.

Definition 2 [*Probability measure*] *Janssen and Manca [2007]*
A probability measure $\mathbf{P}$ on a measurable space $(\Omega, \mathcal{F})$ is
a function $\mathbf{P} : \mathcal{F} \longrightarrow [0; 1]$ such that:

- $\mathbf{P}(\emptyset) = 0, \qquad \mathbf{P}(\Omega) = 1.$
- *If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint (i.e $A_i \cap A_j = \emptyset \quad if \quad i \neq j$) then*

$$\mathbf{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbf{P}(A_i). \tag{1.2}$$

- *The triple $(\Omega, \mathcal{F}, \mathbf{P})$ is called a probability space.*

Definition 3 [*Random variable*] *Janssen and Manca [2007]*

Let $(\Omega, \mathcal{F}, \mathbf{P})$ *be a probability space. A random variable* X *is an application* $X : \Omega \longrightarrow \mathbb{R}^n$ *such that*

$$\forall B \in \mathcal{B}, X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}, \tag{1.3}$$

where $\mathcal{B}$ *is the borel* σ*-field.*

Definition 4 [*Independence of random variables*] *Kopp [2011]*
Let $(\Omega, \mathcal{F}, \mathbf{P})$ *be a probability space,* X_1 *and* X_2 *two random variables*
The random variables X_1 *and* X_2 *are independent if for every choice of borel sets* B_1, B_2*, we have*

$$\mathbf{P}(X_1 \in B_1, X_2 \in B_2) = \mathbf{P}(X_1 \in B_1) \times \mathbf{P}(X_2 \in B_2). \tag{1.4}$$

Definition 5 [*Expectation, variance, covariance of continuous random variable*] *Duffy [2013]*
Let X *be a random variable with density function* $f(x)$*. The expectation of* X *is defined by*

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx, \tag{1.5}$$

provided that the integral

$$\int_{-\infty}^{+\infty} \mid x \mid f(x) dx \tag{1.6}$$

is finite.
The variance of the random variable X *is defined as:*

$$Var(X) = \mathbb{E}\left[\Big(X - \mathbb{E}\big(X\big)\Big)^2\right] = \int_{-\infty}^{\infty} \Big(x - \mathbb{E}\big(X\big)\Big)^2 f(x) dx. \tag{1.7}$$

The variance is always non negative; in particular, for a deterministic variable it is zero. The standard deviation is defined as the square root of the variance by

$$\sigma = \sqrt{Var(X)}. \tag{1.8}$$

The covariance between two random variables X *and* Y *is defined as:*

$$\begin{aligned} Cov(X,Y) &= \mathbb{E}\Big[\Big(X - \mathbb{E}\big(X\big)\Big)\Big(Y - \mathbb{E}\big(Y\big)\Big)\Big] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Big(x - \mathbb{E}\big(X\big)\Big)\Big(y - \mathbb{E}\big(Y\big)\Big) g(x,y) dx dy, \end{aligned} \tag{1.9}$$

where $g(x, y)$ *is the so-called joint density function of the variables* X *and* Y*. In general, the variance is a special case of covariance, in particular* $Var(X) = Cov(X, X)$*. Another way to express the covariance is given by*

$$Cov(X,Y) = \mathbb{E}\big(XY\big) - \mathbb{E}\big(X\big)\mathbb{E}\big(Y\big). \tag{1.10}$$

In general, covariance can be negative, zero or positive. We define the correlation coefficient ρ *between* X *and* Y *by*

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}. \tag{1.11}$$

This factor can be negative, zero or positive. If ρ *is zero, we say that* X *and* Y *are uncorrelated, while if it is positive or negative they are said to be positively or negatively correlated, respectively.*

Let us provide some basic properties of expectation.

Properties 1.1.1.1 *Janssen and Manca [2007]*
Let X and Y be random variables with finite expected values, a and b two constant and $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ a continuous real valued function. The following properties are satisfied:

- $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
- *If X and Y are independent then we have*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y). \tag{1.12}$$

- $$\mathbb{E}\big(\Phi(X)\big) = \int_{-\infty}^{+\infty} \Phi(x) f_X(x) dx. \tag{1.13}$$

Definition 6 [*Normal distribution*] *Tankov [2003]*
The normal distribution is one of most used distribution as it approximates many natural phenomena. It tells us the probability that an observation in some context will fall between two real numbers.

- *Probability density function*

  Let X be a random variable distributed normally with μ as mean and σ^2 as variance. One denotes as following

$$X \sim \mathcal{N}(\mu, \sigma^2) \tag{1.14}$$

  and its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big[-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2} \Big]. \tag{1.15}$$

- *Log-normal distribution*

  The continuous random variable X is log-normally distributed with mean μ and standard deviation σ if the random variable $Y = \ln(X)$ is normally distributed with the mean μ and standard deviation σ. The probability density function of the random variable X is given by

$$f(x) = \mathbf{P}(X = x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\Big[-\frac{(\ln x - \mu)^2}{2\sigma^2} \Big] \qquad \textit{with} \quad x > 0. \tag{1.16}$$

1.1.2 Notion in stochastic calculus

Definition 7 [*Stochastic process*]
Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $(S, \sum)$ a measurable space. A stochastic process with values on the measure space S is a family of random variables

$$\{X_t, t \in T\}, \tag{1.17}$$

where for all t, $X_t : \Omega \longrightarrow E$ is $\mathcal{F}$-measurable i.e

$$\forall B \in S, X_t^{-1}(B) = \{\omega : X_t(\omega) \in B\} \in \mathcal{F}. \tag{1.18}$$

The set T is called parameter of the stochastic process. For every $\omega \in \Omega$, the mapping

$$t \longmapsto X_t(\omega), \tag{1.19}$$

defined on the set parameter T is called the trajectory or the sample path of the process. More details can be found in Oksendal [1992]

Definition 8 [*Adapted stochastic process*] *Kopp [2011]*
The process $\{X_t, t \in T\}$ is adapted to a given filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$ if X_t is $\mathcal{F}_t$-measurable for each $t \in T$.

Definition 9 [*Brownian motion*] *(Shreve [2004])*
Let $(\Omega, \mathcal{F}, \mathbf{P})$ *be a probability space. For each* $\omega \in \Omega$ *suppose there is a continuous function* $W(t)$ *of* $t \geq 0$*, that satifies* $W(0) = 0$ *and that depends on* ω*. Then* $W(t),\ t \geq 0$ *is a Brownian motion if for all* $0 = t_0 < t_1 < \ldots < t_m$ *the increments*

$$W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1}), \tag{1.20}$$

are independent and each of these increments is normally distributed with

$$\mathbb{E}\big[W(t_{i+1}) - W(t_i)\big] = 0, \tag{1.21}$$

$$Var\big[W(t_{i+1}) - W(t_i)\big] = t_{i+1} - t_i. \tag{1.22}$$

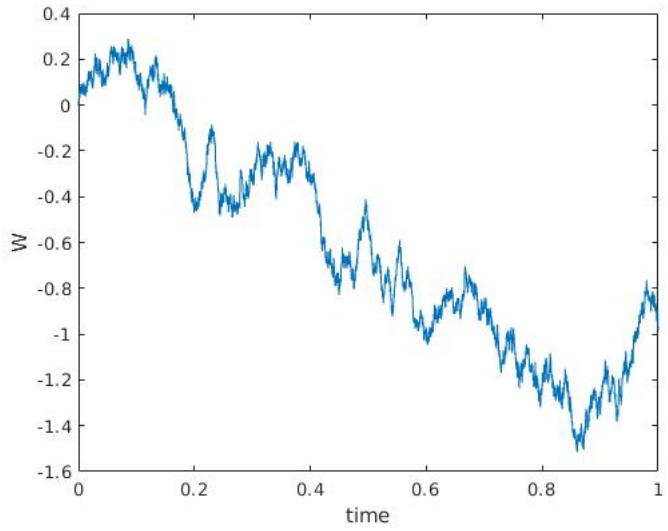


Figure 1.1: Trajectory of a Brownian motion

Definition 10 [*Itô process*] *(Shreve [2004])*
Let $W_t,\ t \geq 0$*, be a Brownian motion, and let* $\mathcal{F}_t,\ t \geq 0$ *be an associated filtration. An Itô process is a stochastic process of the form*

$$X_t = X_0 + \int_0^t u_s ds + \int_0^t v_s dW_s, \tag{1.23}$$

or in the differential form

$$dX_t = u_t dt + v_t dW_t, \tag{1.24}$$

where X_0 *is nonrandom,* u_s *and* v_s *are adapted stochastic processes.*

Theorem 1 [$Multi-dimensional\ Itô\ formula$] *(see Shreve [2004])*
Let

$$dX(t) = udt + vdW(t), \tag{1.25}$$

with

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \quad v = \begin{pmatrix} v_{11} & \ldots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \ldots & v_{nm} \end{pmatrix}, \quad dW(t) = \begin{pmatrix} dW_1(t) \\ \vdots \\ dW_n(t) \end{pmatrix}, \tag{1.26}$$

be a n-dimensional Itô process. Let $g(t,x) = (g_1(t,x), \ldots, g_p(t,x))$ *be a* C^2 *map from* $[0,\infty) \times \mathbb{R}^n$ *into* $\mathbb{R}^p$*. Then the process is again an Itô process whose component number* k, Y_k *is given by*

$$dY_k = \frac{\partial g_k}{\partial t}(t,X) + \sum_i \frac{\partial g_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j, \tag{1.27}$$

where $dW_i dW_j = \delta_{ij}dt, dW_i dt = dt dW_j = 0$.

1.1.3 Useful concepts and theorems

Degeneracy of Partial Differential Equations

Let us consider the following second order partial differential equation:

$$-\sum_{i,j=1}^{N} \frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial u}{\partial x_i}\right) + \sum_{j=1}^{N} b_j \frac{\partial u}{\partial x_j} + cu = f \qquad in\ \Omega \subset \mathbb{R}^N, \tag{1.28}$$

with boundary conditions on $\partial\Omega$ or on part of $\partial\Omega$, and the functions $a_{ij}(x) \in C^1(\overline{\Omega}), 1 \leq i,j \leq N$. The coefficients a_{ij} of PDE in (1.28) satisfy the **ellipticity condition** (see Brezis [2010]) if the following inequality holds

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \alpha|\xi|^2 \qquad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^N \ \ with\ \alpha > 0. \tag{1.29}$$

The PDE is said to be **degenerate** if the coefficients a_{ij} do not satisfy the ellipticity condition (1.29) but only

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq 0 \qquad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^N. \tag{1.30}$$

Let us recall recall some very important theorems in functional analysis for this work. We start with the following definition.

Definition 11 ***Weak derivatives*** *(Evans [1997])*
Suppose $u, v \in L^1_{loc}(\Omega)$ *and* α *is a multi-index. We say that* v *is the* α^{th}*-weak derivative of* u*, written*

$$\partial^\alpha u = v, \tag{1.31}$$

provided

$$\int_\Omega u\ \partial^\alpha \phi\ dx = (-1)^{|\alpha|}\int_\Omega v\phi\ dx, \tag{1.32}$$

for all test functions $\phi \in C_c^\infty(\Omega)$*, with* $L^1_{loc}(\Omega)$ *the space of functions locally integrable on* Ω*.*

Definition 12 ***Sobolev space-seminorm*** *[P. Knabner [2002], Definition 3.2, page 94]*
Suppose $\Omega \subset \mathbb{R}^d$ *is a (bounded) domain. The Sobolev space* $H^k(\Omega)$ *is defined by*

$$H^k(\Omega) := \left\{ v : \Omega \longrightarrow \mathbb{R} | v \in L^2(\Omega),\ \exists\ \partial^\alpha v \in L^2(\Omega), \forall \alpha\ with\ |\alpha| \leq k \right\}. \tag{1.33}$$

The seminorm $|\cdot|_l$ *for* $0 \leq l \leq k$ *in* $H^k(\Omega)$ *is defined by*

$$|v|_l = \left(\sum_{|\alpha|=l} ||\partial^\alpha v||_0^2 \right)^{1/2}, \tag{1.34}$$

such that

$$||v||_k = \left(\sum_{l=0}^{k} |v|_l^2 \right)^{1/2}. \tag{1.35}$$

Definition 13 ***Triangulation*** *[P. Knabner [2002], Definition 3.19, page 114]*
A triangulation $\mathcal{T}_h$ of a set $\Omega \subset \mathbb{R}^d$ consists of a finite number of subsets K of Ω with the following properties:

(T1) Every $K \in \mathcal{T}_h$ is closed.

(T2) For $K \in \mathcal{T}_h$ its nonempty interior $int(K)$ is a Lipschitz domain.

(T3) $\overline{\Omega} = \cup_{K \in \mathcal{T}_h} K$.

(T4) For different K_1 and K_2 of $\mathcal{T}_h$, the intersection of $int(K_1)$ and $int(K_2)$ is empty.

Moreover, A family of triangulation $(\mathcal{T}_h)_h$ is called regular if there exists some $\sigma > 0$ such that for all $h > 0$ and all $K \in \mathcal{T}_h$,

$$\varrho_K \geq \sigma h_K, \tag{1.36}$$

where

$$h_K := diam(K), \quad \varrho_K := \sup \left\{ diam(S) | S \text{ is a ball in } \mathbb{R}^d \text{ and } S \subset K \right\}, \qquad K \in \mathcal{T}_h, \tag{1.37}$$

with diam denoting the diameter.

Theorem 2 ***Projection theorem*** *[P. Knabner [2002], Theorem 3.29, page 138]*
Consider a family of Lagrange finite element discretization in $\mathbb{R}^d$ for $d \leq 3$ on a regular family of triangulation $(\mathcal{T}_h)_h$. For the respective local ansatz spaces P suppose $\mathcal{P}_k \subset P$ for some $k \in \mathbb{N}$. Then there exists some constant $C > 0$ such that for all $v \in H^{k+1}(\Omega)$ and $0 \leq m \leq k+1$

$$\left(\sum_{K \in \mathcal{T}_h} |v - I_K(v)|_{m,K}^2 \right)^{1/2} \leq C \cdot h^{k+1-m} |v|_{k+1}, \tag{1.38}$$

where I_K is a local interpolation operator over K.

Theorem 3 ***Sobolev embedding theorem*** *[Brezis [2010], Theorem 8.8, page 212/213]*

There exists a constant C_1 (depending only on $|I| \leq \infty$) such that

$$||u||_{L^\infty(I)} \leq C_1 ||u||_{W^{1,p}(I)} \quad \forall\ u \in W^{1,p}(I), \quad \forall\ 1 \leq p \leq \infty. \tag{1.39}$$

In other words, $W^{1,p}(I) \subset L^\infty(I)$ with continuous injection for all $1 \leq p \leq \infty$.

Further, if I is ***bounded*** *then*

(a) the ***injection*** *$W^{1,p}(I) \subset C(\bar{I})$ is* ***compact*** *for all $1 < p \leq \infty$,*

(b) the ***injection*** *$W^{1,p}(I) \subset L^q(I)$ is* ***compact*** *for all $1 \leq q < \infty$.*

where I is an open interval of $\mathbb{R}$ and $|I|$ is the measure of the interval I.

In the next paragrah, we are going to present a theorem for the existence and uniqueness of the solution in the case of time dependent problems (see Haslinger et al. [2013]).

Time dependent problems

Let V and H be real seperable Hilbert spaces. By $||\cdot||, \big((\cdot,\cdot)\big), V^*$ and $\langle\cdot,\cdot\rangle$, we denote the norm, the scalar product in V, the dual space to V and the duality pairing between V and V^*, respectively. The norm and the scalar product in H are denoted by $|\cdot|$ and $\big(\cdot,\cdot\big)$.
Let us denote by $L(0,T;X)$, $p\in[1,\infty)$, the set of all measurable functions $u:(0,T)\longrightarrow X$ for which

$$\int_0^T ||u(t)||_X^2 dt<\infty \qquad (in\ the\ Lebesgue\ sense). \tag{1.40}$$

Let us denote also by $W^{1,p}(0,T;V,H)$, the space defined as follows:

$$W^{1,p}(0,T;V,H)=\{u\in L^p(0,T;X)|\ u'\in L^q(0,T;X)\}, \tag{1.41}$$

where $1/p+1/q=1$.

Let $a(\cdot,\cdot;t):V\times V\longmapsto\mathbb{R}$ be an uniformly bounded and V-elliptic bilinear form with respect to $t\in[0,T]$ i.e

$$\exists M>0,\ |a(u,v;t)|\leq M||u||\cdot||v|| \qquad \forall u,v\in V,\ \forall t\in[0,T], \tag{1.42}$$

$$\exists\alpha>0,\ \beta\geq 0: a(v,v;t)\geq\alpha||v||^2-\beta|v|^2 \qquad \forall v\in V,\ \forall t\in[0,T]. \tag{1.43}$$

In addition, we assume that the function $t\longmapsto a(u,v;t)$ is measurable in (0,T) for every $u,v\in V$ and the function $f\in L^2(0,T;V^*)$, the initial value $u_0\in H$ being given. Let us consider the following problem

$$\begin{cases} Find\ u\in W^{1,2}(0,T;V,H)\ satisfying \\ \\ \langle u'(t),v\rangle+a\big(u(t),v;t\big)=\langle f(t),v\rangle \qquad \forall v\in V \quad and\ \ a.a.\ t\in(0,T), \\ \\ u(0)=u_0\ \ given. \end{cases} \tag{1.44}$$

Let us state now the existence and uniqueness theorem for the problem (1.44).

Theorem 4 *[Haslinger et al. [2013], Theorem 1.33, page 39]*
Let the bilinear form $a(\cdot,\cdot;t):V\times V\longmapsto\mathbb{R}$ satisfies (1.42) *and* (1.43)*. Then there exists a unique solution u of* (1.44) *for any $f\in L^2(0,T;V^*)$ and $u_0\in H$.*

1.2 Basic concepts in options theory

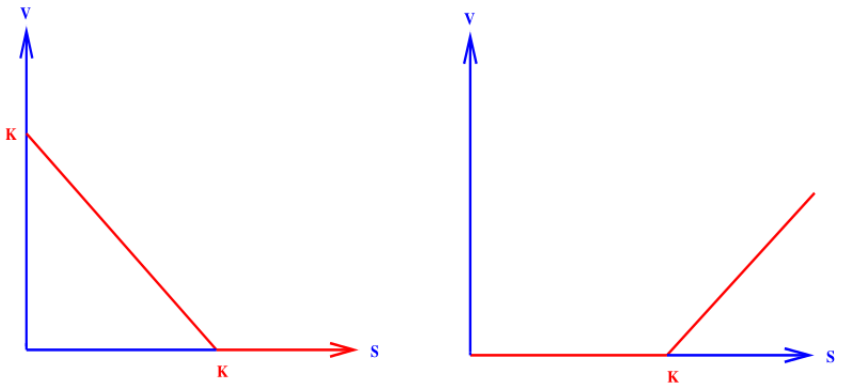
1.2.1 Option

An option is a contract between two parties which gives to the holder the right but not the obligation to buy or to sell an asset at an agreed price K called strike at a specified time T called maturity. A call is an option which gives the right to buy whereas a put gives the right to sell. Thereby, the payoff $\mathcal{P}$ of a call option is given by

$$\mathcal{P}=\max(S-K,0), \tag{1.45}$$

where S is the price of the underlying asset and the payoff for a put option, is given by

$$P=\max(K-S,0). \tag{1.46}$$



(a) Payoff of a put a otion (b) Payoff of a call option

Figure 1.2: Payoff function

Besides, there exist several types of options but the most common are European options and American options. An European option is an option which can be exercised only at maturity whereas an American option can be exercised anytime before the maturity.

In 1973, Robert Merton and Fisher Black stated a mathematical model which was a breakthrough revolutionizing the theory of corporate liability pricing. Nevertheless, their model was given under some conditions which will be discussed in the next paragraph.

1.2.2 Black-Scholes assumptions and model

Before giving the famous Black-Scholes model, Fisher Black and Robert Merton did some assumptions which are:

1. Frictionless Market: this means there no transaction costs of differential taxes.
2. No dividend payment.
3. The risk-free interest is avalaible and constant over time.
4. No restriction regarding value of the transaction and price developpement of the asset.
5. Short trading is not prohibited.
6. Stocks are randomly divisible.
7. All information are available to all the market participants.
8. No arbitrage possibilities.
9. The stock price S_t follows a geometric brownian motion given by:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{1.47}$$

where μ is the drift and σ is the volatility.

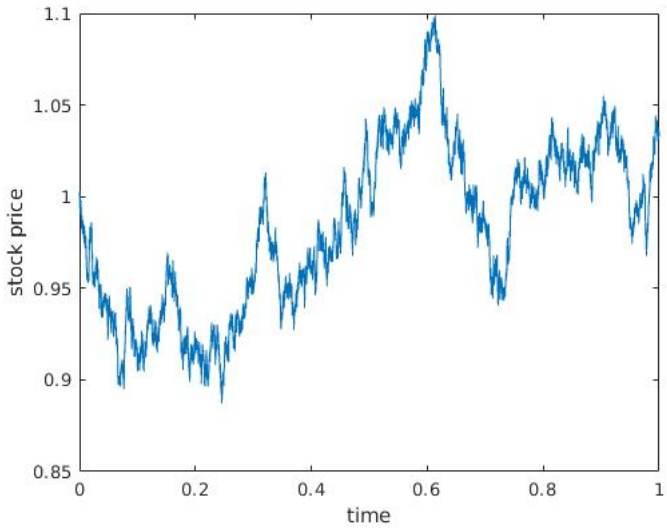


Figure 1.3: Sample of stock price

1.3 Multi assets option Black Scholes Partial Differential Equation

1.3.1 Multi-asset options

A multi-asset option is an option with more than one underlying correlated assets. There exist different types of multi-asset options which differ from each other mainly by their payoff function. Thereby, few of them are (Duffy [2013]):

- Exchange option which is one that gives the holder the right to exchange one asset for another. The payoff $\mathcal{P}$ is then given by

$$\mathcal{P} = \max\big(S_1(T) - S_2(T), 0\big), \tag{1.48}$$

  where S_i is the price of the asset $i,\ i = 1, 2$, and T is the maturity.

- Rainbow option is one that is linked to the performances of two or more underlying assets. It can be speculated on the maximum or minimum performance of all the underlying assets at one time. For two assets option, the payoff $\mathcal{P}$ on the maximum is given by

$$\mathcal{P} = \max\left(\Theta\Big(\max\big(S_1(T), S_2(T)\big) - \Theta K\Big), 0\right), \tag{1.49}$$

  where $\Theta = 1$ for a call, $\Theta = -1$ for a put, and K the strike price. Similarly, the payoff $\mathcal{P}$ for a two-colour rainbow option on the minimum of two assets is given by

$$\mathcal{P} = \max\left(\Theta\Big(\min\big(S_1(T), S_2(T)\big) - \Theta K\Big), 0\right). \tag{1.50}$$

- Basket options can be used by portfolio managers to hedge the risks of their portfolios. A basket option is defined as:

$$S(\tau) = \sum_{j=1}^{n} \alpha_j S_j(\tau), \tag{1.51}$$

where α_j is the the total investment in asset j (as a percentage), $S_j(\tau)$ is the price of the asset j and the sum of weight $\alpha_j,\ j = 1, 2, \ldots, n$ is equal to 1. The payoff $\mathcal{P}$ of a basket option is then given by:

$$\mathcal{P} = \max\left(\Theta\big[S(T) - K\big], 0\right). \tag{1.52}$$

where K is the strike and Θ is defined as in equation (1.49).

1.3.2 Derivation of the multi-dimensional Black-Scholes Partial Differential Equation

For a multi-asset option with n ($n \geq 2$) underlying assets, the Black-Scholes model is given by:

$$\begin{cases} dx_i(t) = rx_i(t)dt + \sigma_i x_i(t)dW_i(t) & i = 1, 2, \ldots, n, \\ \\ dW_i(t)dW_j(t) = \rho_{ij}dt \quad i \neq j & i, j = 1, 2, \ldots, n, \end{cases} \tag{1.53}$$

where x_i, σ_i, W_i represent respectively the stock price, the volatility and the Brownian motion associated to asset i; t is the current time, r is the risk-free interest and ρ_{ij} is the coefficient correlation between asset i and j . For simplicity, we denote $x_i(t)$ by $x_i,\ i = 1, 2, \ldots, N$.
Let us consider the Black-Scholes model of a multi-asset option with n underlying assets (1.53) and $V(x_1, x_2, \ldots, x_n, t)$ the value of that option; We set up a portfolio made of one option and δ_i assets $i,\ i = 1, 2, \ldots, n$. By applying the multi-dimensional Itô formula (1.27) to V, we have

$$\begin{aligned} dV(x_1, x_2, \ldots, x_n, t) = &\left[\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \mu_i x_i \frac{\partial V}{\partial x_i} + \sum_{i=1}^{n} \frac{1}{2}\sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i\neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j}\right] dt \\ &+ \sum_{i=1}^{n} \sigma_i x_i \frac{\partial V}{\partial x_i} dW_i. \end{aligned}$$

The dynamic of the portfolio P_t is given by

$$P_t = V(x_1, x_2, \ldots, x_n, t) - \sum_{i=1}^{n} \delta_i x_i.$$

The change in value of the portfolio P_t is given by

$$\begin{aligned} dP_t &= dV(x_1, x_2, \ldots, x_n, t) - \sum_{i=1}^{n} \delta_i dx_i \\ &= \left[\left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \mu_i x_i \frac{\partial V}{\partial x_i} + \sum_{i=1}^{n} \frac{1}{2}\sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i\neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j}\right) dt + \sum_{i=1}^{n} \sigma_i x_i \frac{\partial V}{\partial x_i} dW_i\right] \\ &\quad - \sum_{i=1}^{n} \delta_i \Big(\mu_i x_i(t)dt + \sigma_i x_i dW_i(t)\Big) \\ &= \left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \mu_i x_i \frac{\partial V}{\partial x_i} + \sum_{i=1}^{n} \frac{1}{2}\sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i\neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} - \sum_{i=1}^{n} \delta_i \mu_i x_i(t)\right) dt \\ &\quad + \sum_{i=1}^{n} \left(\sigma_i x_i \frac{\partial V}{\partial x_i} - \delta_i x_i \sigma_i\right) dW_i. \end{aligned} \tag{1.54}$$

Using the standard hedging (Arbitrage) argument, it gives

$$\sigma_i x_i \frac{\partial V}{\partial x_i} - \delta_i x_i \sigma_i = 0 \qquad i = 1, 2, \cdots, n, \tag{1.55}$$

which leads to

$$\delta_i = \frac{\partial V}{\partial x_i} \qquad i = 1, 2, \cdots, n. \tag{1.56}$$

Substituting δ_i in (1.54) leads to

$$\begin{aligned} dP_t &= \left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \mu_i x_i \frac{\partial V}{\partial x_i} + \sum_{i=1}^{n} \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} - \sum_{i=1}^{n} \mu_i x_i(t) \frac{\partial V}{\partial x_i} \right) dt \\ dP_t &= \left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} \right) dt. \end{aligned}$$

The portfolio is expected to grow at risk-free interest rate r, then

$$\mathbb{E}(dP_t) = r P_t dt.$$

Thereby

$$\mathbb{E}\left[\left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} \right) dt \right] = r P_t dt,$$

and then

$$\left(\frac{\partial V}{\partial t} + \sum_{i=1}^{n} \frac{1}{2} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} \right) dt = r \left(V - \sum_{i=1}^{n} x_i \frac{\partial V}{\partial x_i} \right) dt.$$

Therefore, we obtain

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i=1}^{n} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} + r \sum_{i=1}^{n} x_i \frac{\partial V}{\partial x_i} - rV = 0. \tag{1.57}$$

By setting $\tau = T - t$, we finally get multi-dimensional Black-Scholes Partial Differential Equation as follows:

$$\frac{\partial V}{\partial \tau} - \frac{1}{2} \sum_{i=1}^{n} \sigma_i^2 x_i^2 \frac{\partial^2 V}{\partial x_i^2} - \frac{1}{2} \sum_{i,j=1_{i \neq j}}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j} - r \sum_{i=1}^{n} x_i \frac{\partial V}{\partial x_i} + rV = 0. \tag{1.58}$$

As we can notice, when the stock price x_i goes to zero, the coefficients of the second order derivatives go to zero; then the ellipticity condition in (1.29) is not satistify. Therefore, the multi-dimensional Black-Scholes PDE is **degenerate**.

In our work, we call **degeneracy region**, the region denoted by $\mathcal{D}_{\mathcal{R}}$ and defined as:

$$\mathcal{D}_{\mathcal{R}} = \bigcup_{i=1}^{n} \left([0, x_i^1] \times \prod_{j=1_{j \neq i}}^{n} [0, x_j^{max}] \right), \tag{1.59}$$

where x_i represents the price of the asset i and $0, x_i^1, x_i^2, \ldots, x_i^{\max}$ are grid points in the increasing order, on the axis representing the asset price x_i.

1.4 The continuous problem: Black Scholes PDE for pricing multi-assets options

1.4.1 Weighted Sobolev spaces

Let us start by introducing the notations and functional spaces that we will be used in this work. For an open set $\Omega \subset R^n, n \in \mathbb{N}$, the space of square integrable functions is denoted $L^2(\Omega)$. We denote also by $C(\Omega)$ (respectively $C(\overline{\Omega})$) the set of continuous functions over Ω (respectively on $\overline{\Omega}$). For any Hilbert space $G(\Omega)$ of classes of functions defined on Ω, we let $L^2\big(0,T;G(\Omega)\big)$ denote the space defined by

$$L^2\Big(0,T;G(\Omega)\Big) = \Bigg\{ v(\cdot,t) \in G(\Omega) \ a.e \ in \ (0,T); ||v(\cdot,t)||_G \in L^2\big((0,T)\big) \Bigg\}, \tag{1.60}$$

where $||\cdot||_G$ denotes the natural norm on $G(\Omega)$. The norm on this space is denoted by $||\cdot||_{L^2\left(0,T;G(\Omega)\right)}$ i.e

$$||v||_{L^2\left(0,T;G(\Omega)\right)} = \left(\int_0^T ||v(\cdot,t)||_G^2 dt \right)^{1/2}. \tag{1.61}$$

The Black-Scholes operator being degenerated, we introduce a weighted L^2-norm $||\cdot||_\omega$ defined by

$$||v||_\omega = \left(\int_\Omega \sum_{i=1}^n x_i^2 v_i^2 dx \right)^{1/2}. \tag{1.62}$$

The space of all weighted square integrable functions is defined as

$$L^2_\omega(\Omega) = \Bigg\{ v : ||v||_\omega < \infty \Bigg\}, \tag{1.63}$$

and the corresponding weighted inner product on $L^2_\omega(\Omega)$ is given by

$$(u,v)_\omega = \int_\Omega \sum_{i=1}^n x_i^2 u_i v_i dx. \tag{1.64}$$

Thereby, we define the weighted Sobolev spaces as follows:

$$H^1_\omega(\Omega) = \Bigg\{ v \in L^2(\Omega) : \exists g \in L^2_\omega(\Omega) \quad \text{such that} \quad \int_\Omega v\varphi' = -\int_\Omega g\varphi \qquad \forall \varphi \in C^1_c(\Omega) \Bigg\}, \tag{1.65}$$

where $g = v'$ is the weak derivative; and also

$$H^1_{0,\omega}(\Omega) = \Bigg\{ v : v \in H^1_\omega(\Omega) \ \text{ and } \ v\big|_{\partial\Omega} = 0 \Bigg\}. \tag{1.66}$$

Using the inner products on $L^2(\Omega)$ and $L^2_\omega(\Omega)$, we define the inner product on $H^1_\omega(\Omega)$ by

$$\big(\cdot,\cdot\big)_H = \big(\cdot,\cdot\big) + \big(\cdot,\cdot\big)_\omega, \tag{1.67}$$

with the corresponding norm denoted by $||\cdot||_{1,\omega}$ and defined as follows:

$$\big|\big|v\big|\big|_{1,\omega} = \Bigg[\big|\big|v\big|\big|_0^2 + \big|\big|\nabla v\big|\big|_\omega^2 \Bigg]^{1/2} = \Bigg[(v,v) + (x\nabla v, x\nabla v) \Bigg]^{1/2}, \tag{1.68}$$

with

$$x\nabla v = \sum_{i=1}^n x_i \frac{\partial v}{\partial x_i}. \tag{1.69}$$

1.4.2 Existence and uniqueness results for the continuous problem solution

Let us recall the Black-Scholes PDE for pricing multi-asset option:

$$\mathcal{L}V := \frac{\partial V}{\partial t} - \frac{1}{2}\sum_{i,j=1}^{n} \sigma_i\sigma_j\rho_{ij}x_ix_j\frac{\partial^2 V}{\partial x_i\partial x_j} - r\sum_{i=1}^{n} x_i\frac{\partial V}{\partial x_i} + rV = 0, \tag{1.70}$$

with $(x_1, x_2, \ldots, x_n, \tau) \in \Omega\times(0,T],\ \Omega = (0, x_{\max})^n$, where x_i is the price of the asset i, r is the risk free interest, V is the option value, T is the maturity and t is the time to maturity. For $i, j = 1, 2, \ldots, n$, x_i represents the asset i price, σ_i represents the volatility of asset i, ρ_{ij} represents the correlation between the assets i and j. The corresponding initial and boundary conditions are:

$$\begin{cases} V(x_1, x_2, \ldots, x_n, 0) & = h_1(x_1, x_2, \ldots, x_n), \\ V(x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n, t) & = h_2^i(t), \\ V(x_1, x_2, \ldots, x_{i-1}, x_{\max}, x_{i+1}, \ldots, x_n, t) & = h_3^i(t), \end{cases} \tag{1.71}$$

where $h_1, h_2^i, h_3^i,\ \ i = 1, 2, \ldots, n$ are functions depending on the type of option. Our study is conducted under the following assumptions.

Assumption 1 *We assume that, for $i = 1, \ldots, n$, the coefficients r and σ_i satisfy*

$$\underline{r} \le r(t) \le \bar{r} \qquad\qquad \underline{\sigma} \le \sigma_i(t) \le \bar{\sigma}, \tag{1.72}$$

and we define by

$$\beta := \frac{1}{2}\sup_{t\in[0,T]}\sum_{i=1}^{n}\left(\sigma_i^2(t) + \frac{1}{2}\sum_{j=1_{i\neq j}}^{n}\rho_{ij}\sigma_i(t)\sigma_j(t)\right). \tag{1.73}$$

Without a loss of generality, we assume that the functions $h_1, h_2^i, h_3^i,\ i = 1, \ldots, n$, defined in (1.71) are equal to zero. Indeed, we can transform the nonhomogeneous boundary conditions (1.71) into homogeneous one. This can be done by substracting $\mathcal{L}V_0$ from both sides of the PDE (1.70), whereby V_0 is a known function satisfying the boundary conditions in (1.71) . A nonzero function g will appear on the right-hand side of equation (1.70) leading to the following new PDE

$$\mathcal{L}U := \frac{\partial U}{\partial t} - \frac{1}{2}\sum_{i,j=1}^{n} \sigma_i\sigma_j\rho_{ij}x_ix_j\frac{\partial^2 U}{\partial x_i\partial x_j} - r\sum_{i=1}^{n} x_i\frac{\partial U}{\partial x_i} + rU = g, \tag{1.74}$$

where $U = e^{\beta t}(V - V_0)$. In order to apply the finite volume method, it is convenient to transform (1.74) to its divergence form as follows:

$$\frac{\partial U}{\partial \tau} - \nabla\cdot(M\nabla U) - \nabla\cdot\big(fU\big) + \lambda U = g, \tag{1.75}$$

where $M = (m_{ij})_{1\le i,j\le n}$ is a $n\times n$ matrix such that

$$\begin{aligned} m_{ii} &= \frac{1}{2}\sigma_i^2x_i^2 \qquad i = 1, 2, \ldots, n, \\ m_{ij} &= \frac{1}{2}\rho_{ij}\sigma_i\sigma_jx_ix_j \qquad i \neq j \ \ i, j = 1, 2, \ldots, n, \end{aligned} \tag{1.76}$$

$f = (f_i)_{1\le i\le n}$ is a column vector of size $n\times 1$ such that

$$f_i = \left(r - \sigma_i^2 - \frac{1}{2} \sum_{j=1_{i \neq j}}^{n} \rho_{ij} \sigma_i \sigma_j \right) x_i, \tag{1.77}$$

and λ is a scalar such that

$$\lambda = \big(n+1\big) r + \beta - \sum_{i=1}^{n} \left(\sigma_i^2 + \frac{1}{2} \sum_{j=1_{i \neq j}}^{n} \rho_{ij} \sigma_i \sigma_j \right). \tag{1.78}$$

Assumption 2 *For $i, j = 1, 2, \ldots, n$, we assume that*

$$\sum_{j=1_{j \neq i}}^{n} \rho_{ij} < 1, \qquad \textit{and} \qquad \rho_{ij} \geq 0 \qquad i, j = 1, 2, \ldots, n. \tag{1.79}$$

Thereby using the weighted Sobolev space defined in (1.66), the corresponding variational problem can be formulated as follows:

Problem 1 *Find $u(t) \in H^1_{0,\omega}\big(\Omega\big)$ such that for all $v \in H^1_{0,\omega}\big(\Omega\big)$*

$$\begin{aligned} \Big(\frac{du}{dt}, v\Big) + A(u, v; t) &= (g, v) \qquad\qquad a.e \ in \ \big(0, T\big), \\ u_0 &= max(x - K, 0), \end{aligned} \tag{1.80}$$

where

$$A(u, v; t) = \Big(M \nabla u, \nabla v\Big) + \Big(fu, \nabla v\Big) + \lambda(u, v), \tag{1.81}$$

is a bilinear form and M, f and λ are defined in (1.76),(1.77) *and* (1.78) *and K is the strike price.*

Theorem 5 *There exists a unique solution to* ***Problem 1***

Proof

1. V-ellipticity of the bilinear form

   Using integration by part, for $v \in H^1_{0,\omega}(\Omega)$ we have:

$$\begin{aligned} \int_\Omega f v \cdot \nabla v d\Omega &= \int_{\partial\Omega} v^2 f \cdot n ds - \int_\Omega v \nabla \cdot (fv) d\Omega \\ &= -\int_\Omega f v \cdot \nabla v d\Omega - \int_\Omega v^2 \cdot \nabla f d\Omega, \end{aligned} \tag{1.82}$$

   where n is the outward vector normal to $\partial\Omega$. Since $v = 0$ on $\partial\Omega$, hence

$$\int_\Omega f v \cdot \nabla v d\Omega = -\frac{1}{2} \int_\Omega (\nabla \cdot f) \cdot v^2 d\Omega. \tag{1.83}$$

   As a result of that, we have

$$
\begin{aligned}
A(v,v;t) &= \Big(M\nabla v,\nabla v\Big)+\Big(fv,\nabla\cdot v\Big)+\Big(\lambda v,v\Big)\\
&= \Big(M\nabla v,\nabla v\Big)-\frac{1}{2}\Big((\nabla\cdot f)\cdot v,v\Big)+\Big(\lambda v,v\Big)\\
&= \Big(M\nabla v,\nabla v\Big)+\frac{1}{2}\Big(\big(2\lambda-\nabla\cdot f\big)\cdot v,v\Big)\\
&= \int_\Omega\sum_{i=1}^n\sum_{i=1}^n m_{ij}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}d\Omega+\frac{1}{2}\int_\Omega\sum_{i=1}^n\left(2\lambda-\sum_{i=1}^n\frac{\partial f_i}{\partial x_i}\right)v_i^2d\Omega\\
&= \int_\Omega\sum_{i=1}^n m_{ii}\Big(\frac{\partial v}{\partial x_i}\Big)^2d\Omega+\int_\Omega\sum_{i=1}^n\sum_{j=1_{i\neq j}}^n m_{ij}\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}d\Omega\\
&\quad+\frac{1}{2}\int_\Omega\sum_{i=1}^n\Big(2\lambda-\sum_{i=1}^n\frac{\partial f_i}{\partial x_i}\Big)v_i^2d\Omega\\
&= \frac{1}{2}\int_\Omega\sum_{i=1}^n\sigma_i^2x_i^2\Big(\frac{\partial v}{\partial x_i}\Big)^2d\Omega+\frac{1}{2}\int_\Omega\sum_{i=1}^n\sum_{j=1}^n\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}d\Omega\\
&\quad+\frac{1}{2}\int_\Omega\sum_{i=1}^n\left[2\left((n+1)r+\beta-\sum_{i=1}^n\Big(\sigma_i^2+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i\sigma_j\Big)\right)\right.\\
&\quad\left.-\sum_{i=1}^n\left(r-\sigma_i^2-\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i\sigma_j\right)\right]v_i^2d\Omega.
\end{aligned}
$$

$$
\begin{aligned}
A(v,v;t) &= \frac{1}{2}\int_\Omega\sum_{i=1}^n\sigma_i^2x_i^2\Big(\frac{\partial v}{\partial x_i}\Big)^2d\Omega+\frac{1}{2}\int_\Omega\sum_{i=1}^n\sum_{i=1}^n\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}d\Omega\\
&\quad+\frac{1}{2}\int_\Omega\left[\sum_{i=1}^n(n+2)rv_i^2+2\beta-\sum_{i=1}^n\Big(\sigma_i^2+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i\sigma_j\Big)\right]d\Omega.
\end{aligned}
\tag{1.84}
$$

Using the definition of β in (1.73), we have

$$
\begin{aligned}
2\beta-\sum_{i=1}^n\Big(\sigma_i^2+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i\sigma_j\Big) &= \sup_{t\in[0,T]}\sum_{i=1}^n\left(\sigma_i^2(t)+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i(t)\sigma_j(t)\right)\\
&\quad-\sum_{i=1}^n\Big(\sigma_i^2(t)+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i(t)\sigma_j(t)\Big)\\
2\beta-\sum_{i=1}^n\Big(\sigma_i^2+\frac{1}{2}\sum_{j=1_{i\neq j}}^n\rho_{ij}\sigma_i\sigma_j\Big) &\geq 0.
\end{aligned}
\tag{1.85}
$$

Besides, let us set

$$
P=\frac{1}{2}\int_\Omega\sum_{i=1}^n\sigma_i^2x_i^2\Big(\frac{\partial v}{\partial x_i}\Big)^2d\Omega+\frac{1}{2}\int_\Omega\sum_{i=1}^n\sum_{j=1}^n\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}d\Omega. \tag{1.86}
$$

Thereby, we have

$$\begin{aligned} P &= \frac{1}{2}\int_\Omega \sum_{i=1}^n \sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 d\Omega + \frac{1}{2}\int_\Omega \sum_{i=1}^n \sum_{j=1}^n \rho_{ij}\sigma_i\sigma_j x_i x_j \frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j} d\Omega \\ &= \frac{1}{2}\int_\Omega \left(\sum_{i=1}^n \sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 + \sum_{i=1}^n \sum_{j=1_{j\neq i}}^n \rho_{ij}\sigma_i\sigma_j x_i x_j \frac{\partial v}{\partial x_i}\frac{\partial v}{\partial x_j}\right) d\Omega \end{aligned}$$

$$\begin{aligned} P &= \frac{1}{2}\int_\Omega \left[\sum_{i=1}^n \sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 + \frac{1}{2}\sum_{i=1}^n \sum_{j=1_{j\neq i}}^n \rho_{ij}\left(\left(\sigma_i x_i \frac{\partial v}{\partial x_i} + \sigma_j x_j \frac{\partial v}{\partial x_j}\right)^2 - \sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 \right.\right. \\ &\quad \left.\left. -\sigma_j^2 x_j^2 \Big(\frac{\partial v}{\partial x_j}\Big)^2 \right)\right] \\ P &= \frac{1}{2}\int_\Omega \left[\sum_{i=1}^n \left(1 - \sum_{j=1_{j\neq i}}^n \rho_{ij}\right)\sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 + \sum_{i=1}^n \sum_{j=1_{j\neq i}}^n \rho_{ij}\left(\sigma_i x_i \frac{\partial v}{\partial x_i} + \sigma_j x_j \frac{\partial v}{\partial x_j}\right)^2 \right] \end{aligned} \tag{1.87}$$

and using (1.79), we have for $i, j = 1, 2, \ldots, n, \quad \rho_{ij} \geq 0$ then

$$\sum_{i=1}^n \sum_{j=1_{j\neq i}}^n \rho_{ij}\left(\sigma_i x_i \frac{\partial v}{\partial x_i} + \sigma_j x_j \frac{\partial v}{\partial x_j}\right)^2 \geq 0. \tag{1.88}$$

Furthermore, using (1.84) (1.87), (1.88) and (1.85), we get

$$\begin{aligned} A(v, v; t) &\geq \frac{1}{2}\int_\Omega \sum_{i=1}^n \left(1 - \sum_{j=1_{j\neq i}}^n \rho_{ij}\right)\sigma_i^2 x_i^2 \Big(\frac{\partial v}{\partial x_i}\Big)^2 d\Omega + \frac{1}{2}\int_\Omega (n+2) r v_i^2 d\Omega \\ &\geq \frac{(2-n)\sigma_i^2}{2}\Big(x\nabla v, x\nabla v\Big) + \frac{1}{2}\Big((n+2)rv, v\Big). \end{aligned} \tag{1.89}$$

Therefore

$$A(v, v, t) \geq C \cdot \big|\big|v\big|\big|_{1,\omega}^2, \tag{1.90}$$

with

$$C = \frac{1}{2}\min\left(|2-n|\underline{\sigma}^2, (n+2)\underline{r}\right), \qquad C > 0. \tag{1.91}$$

The bilinear form $A(\cdot, \cdot, t)$ is then coercive thus $H^1_\omega(\Omega)$-elliptic (taking $\beta = 0$ in (1.43)).

2. uniform bound of the bilinear form A

$$\big|A(u, v; t)\big| \leq \big|\big(M\nabla u, \nabla v\big)\big| + \big|\big(fu, \nabla v\big)\big| + \big|\big(\lambda u, v\big)\big|. \tag{1.92}$$

In one hand,

$$
\begin{aligned}
\big|(M\nabla u,\nabla v)\big| &= \left|\int_\Omega \frac{1}{2}\sum_{i=1}^{n}\sigma_i^2x_i^2\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}d\Omega+\frac{1}{2}\int_\Omega\sum_{i=1}^{n}\left(\sum_{j=1_{i\neq j}}\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\right)d\Omega\right| \\
&\leq \frac{1}{2}\left|\int_\Omega\sum_{i=1}^{n}\sigma_i^2x_i^2\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}d\Omega\right|+\frac{1}{2}\left|\int_\Omega\sum_{i=1}^{n}\left(\sum_{j=1_{i\neq j}}\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\right)d\Omega\right|.
\end{aligned}
\tag{1.93}
$$

Furthermore,

$$
\begin{aligned}
\frac{1}{2}\left|\int_\Omega\sum_{i=1}^{n}\sigma_i^2x_i^2\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}d\Omega\right| &\leq \frac{1}{2}\sum_{i=1}^{n}\left|\int_\Omega\sigma_i^2x_i^2\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}d\Omega\right| \\
&\leq \frac{1}{2}\sum_{i=1}^{n}\sigma_i^2\left|\int_\Omega\left(x_i\frac{\partial u}{\partial x_i}\times x_i\frac{\partial v}{\partial x_i}\right)d\Omega\right| \\
&= \frac{1}{2}\sum_{i=1}^{n}\sigma_i^2\left|\left(x_i\frac{\partial u}{\partial x_i},x_i\frac{\partial v}{\partial x_i}\right)\right| \\
&\leq \frac{1}{2}\sum_{i=1}^{n}\left[\left|\sigma_i^2\left(x_i\frac{\partial u}{\partial x_i},x_i\frac{\partial u}{\partial x_i}\right)\right|^{1/2}\cdot\left|\sigma_i^2\left(x_i\frac{\partial v}{\partial x_i},x_i\frac{\partial v}{\partial x_i}\right)\right|^{1/2}\right] \\
&= \frac{1}{2}\sum_{i=1}^{n}\left|\int_\Omega\Big(\sigma_ix_i\frac{\partial u}{\partial x_i}\Big)^2d\Omega\right|^{1/2}\cdot\left|\int_\Omega\Big(\sigma_ix_i\frac{\partial v}{\partial x_i}\Big)^2d\Omega\right|^{1/2} \\
&\leq \frac{1}{2}\left(\sum_{i=1}^{n}\int_\Omega\Big(\sigma_ix_i\frac{\partial u}{\partial x_i}\Big)^2d\Omega\right)^{1/2}\cdot\left(\sum_{i=1}^{n}\int_\Omega\Big(\sigma_ix_i\frac{\partial v}{\partial x_i}\Big)^2\right)^{1/2} \\
&\leq \frac{1}{2}\left(\int_\Omega\sum_{i=1}^{n}\sigma_i^2x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2d\Omega\right)^{1/2}\cdot\left(\int_\Omega\sum_{i=1}^{n}\sigma_i^2x_i^2\Big(\frac{\partial v}{\partial x_i}\Big)^2\right)^{1/2} \\
\frac{1}{2}\left|\int_\Omega\sum_{i=1}^{n}\sigma_i^2x_i^2\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}d\Omega\right| &\leq \frac{1}{2}\beta\big|\big|\nabla u\big|\big|_{0,\omega}\cdot\big|\big|\nabla v\big|\big|_{0,\omega}.
\end{aligned}
\tag{1.94}
$$

Moreover, let B be the expression given by

$$
B=\frac{1}{2}\left|\int_\Omega\sum_{i=1}^{n}\left(\sum_{j=1_{i\neq j}}^{n}\rho_{ij}\sigma_i\sigma_jx_ix_j\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\right)d\Omega\right|. \tag{1.95}
$$

it follows that:

$$\begin{aligned}
B &\leq \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n}\left|\int_{\Omega}\left(\rho_{ij}\sigma_i\sigma_j x_i x_j \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\right)d\Omega\right| \\
&\leq \frac{1}{2}\beta\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n}\left|\int_{\Omega}\left[x_i\frac{\partial u}{\partial x_i}\times x_j\frac{\partial v}{\partial x_j}\right]d\Omega\right| \\
&= \frac{1}{2}\beta\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n}\left|\left(x_i\frac{\partial u}{\partial x_i}, x_j\frac{\partial v}{\partial x_j}\right)\right| \\
&\leq \frac{\beta}{2}\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n}\left|\left(x_i\frac{\partial u}{\partial x_i}, x_i\frac{\partial u}{\partial x_i}\right)\right|^{\frac{1}{2}}\left|\left(x_j\frac{\partial u}{\partial x_j}, x_j\frac{\partial v}{\partial x_j}\right)\right|^{\frac{1}{2}} \\
&= \frac{\beta}{2}\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n}\left[\int_{\Omega}x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right]^{\frac{1}{2}}\left[\int_{\Omega}x_j^2\Big(\frac{\partial v}{\partial x_j}\Big)^2 d\Omega\right]^{\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}
B &= \frac{\beta}{2}\sum_{i=1}^{n}\left[\int_{\Omega}x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right]^{\frac{1}{2}}\sum_{j=1_{i\neq j}}^{n}\left[\int_{\Omega}x_j^2\Big(\frac{\partial v}{\partial x_j}\Big)^2 d\Omega\right]^{\frac{1}{2}} \\
&= \frac{\beta}{2}\sum_{i=1}^{n}\left[\int_{\Omega}x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right]^{\frac{1}{2}}\sum_{j=1_{i\neq j}}^{n}\left[\int_{\Omega}x_j^2\Big(\frac{\partial v}{\partial x_j}\Big)^2 d\Omega\right]^{\frac{1}{2}} \\
&= \frac{\beta}{2}\left[\int_{\Omega}\sum_{i=1}^{n}x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right]^{\frac{1}{2}}\left[\int_{\Omega}\sum_{j=1_{i\neq j}}^{n}x_j^2\Big(\frac{\partial v}{\partial x_j}\Big)^2 d\Omega\right]^{\frac{1}{2}} \\
&\leq \frac{\beta}{2}\left[\int_{\Omega}\sum_{i=1}^{n}x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right]^{\frac{1}{2}}\left[\int_{\Omega}\sum_{j=1}^{n}x_j^2\Big(\frac{\partial v}{\partial x_j}\Big)^2 d\Omega\right]^{\frac{1}{2}} \\
B &\leq \frac{\beta}{2}\big|\big|\nabla u\big|\big|_{0,\omega}\cdot\big|\big|\nabla v\big|\big|_{0,\omega}. \qquad (1.96)
\end{aligned}$$

Using (1.94) and (1.96), gives

$$\big|\big(M\nabla u,\nabla v\big)\big| \leq \beta\cdot\big|\big|\nabla u\big|\big|_{0,\omega}\cdot\big|\big|\nabla v\big|\big|_{0,\omega}. \qquad (1.97)$$

In another hand,

$$
\begin{aligned}
\big|(fu,\nabla v)\big| &= \left|\int_\Omega fu\cdot\nabla v d\Omega\right| \\
&= \left|\int_{\partial\Omega} uvf\cdot n\ dS-\int_\Omega v\nabla fu\ d\Omega\right| \\
&= \left|-\int_\Omega vf\nabla u\cdot d\Omega-\int_\Omega uv\nabla f\ d\Omega\right| \\
&\leq \left|\int_\Omega vf\nabla u\cdot d\Omega\right|+\left|uv\nabla f\ d\Omega\right| \\
&\leq \left|\int_\Omega vf\nabla u\cdot d\Omega\right|+M\big|\big|u\big|\big|_0\big|\big|v\big|\big|_0. \qquad (1.98)
\end{aligned}
$$

Furthermore, using Cauchy Schwartz inequality, leads to

$$
\begin{aligned}
\left|\int_\Omega vf\nabla u\cdot d\Omega\right| &= \left|\int_\Omega v\left[\sum_{i=1}^n\left(r-\sigma_i^2-\frac{1}{2}\sum_{j=1_{i\neq j}}\rho_{ij}\sigma_i\sigma_j\right)x_i\frac{\partial u}{\partial x_i}\right]d\Omega\right| \\
&\leq M_1\left|\int_\Omega v\sum_{i=1}^n x_i\frac{\partial u}{\partial x_i}d\Omega\right| \\
&\leq M_2\left(\int_\Omega v^2 d\Omega\right)^{1/2}\cdot\left(\int_\Omega\sum_{i=1}^n x_i^2\Big(\frac{\partial u}{\partial x_i}\Big)^2 d\Omega\right)^{1/2}. \qquad (1.99)
\end{aligned}
$$

Then we get

$$
\left|\int_\Omega vf\nabla u\cdot d\Omega\right| = M\big|\big|v\big|\big|_0\cdot\big|\big|\nabla u\big|\big|_\omega. \qquad (1.100)
$$

Besides, we have also

$$
\begin{aligned}
\big|(\lambda u,v)\big| &= \int_\Omega\lambda uv d\Omega \\
&\leq M_3\big|\big|u\big|\big|_0\cdot\big|\big|v\big|\big|_0. \qquad (1.101)
\end{aligned}
$$

Combining (1.97), (1.100) and (1.101) in (1.92), we get the following estimate

$$
\begin{aligned}
\big|A(u,v;t)\big| &\leq \big|(M\nabla u,\nabla v)\big|+\big|(fu,\nabla v)\big|+\big|(\lambda u,v)\big| \\
&\leq \beta\cdot\big|\big|\nabla u\big|\big|_\omega\cdot\big|\big|\nabla v\big|\big|_\omega+M\big|\big|v\big|\big|_0\cdot\big|\big|\nabla u\big|\big|_\omega+M_3\big|\big|u\big|\big|_0\cdot\big|\big|v\big|\big|_0 \\
\big|A(u,v;t)\big| &\leq M\cdot\big|\big|u\big|\big|_{1,\omega}\cdot\big|\big|v\big|\big|_{1,\omega}. \qquad (1.102)
\end{aligned}
$$

Besides, we can easily show that the function $v\to(f,v)$ is continuous in $H^1_\omega(\Omega)$. Then, the bilinear form is uniformly bounded and $H^1_\omega(\Omega)$-elliptic. Using Theorem 4, then there exists a unique solution to **Problem 1**.

Conclusion

In this Chapter, we have presented some basic concepts in measure theory, stochastic calculus and option theory. Thereby, using Itô formula and arbitrage arguments, we were able to derive the multi-dimensional degenerated Black-Scholes PDE. We ended the Chapter by proving that the multi-dimensional continuous pricing problem has a unique solution.

Chapter 2

Two-Point Flux Approximation methods and Fitted Two Point Flux Approximation methods for pricing European options

In this Chapter, we present the Two Point Flux Approximation (TPFA) method and introduce a novel numerical scheme called the fitted Two Point Flux Approximation by combining the Two Point Flux Approximation and the classical finite volume method for solving the one dimensional degenerated Black-Scholes Partial Differential Equations. Convergence proofs for the TPFA and fitted TPFA are provided. This Chapter is part of the preprint that can be found in Koffi and Tambue [2019a].

2.1 Introduction

It is well known that the Black-Scholes Partial Differential Equation is degenerate when the stock price approaches zero. This degeneracy may affect the accuracy of the numerical method used if sophisticated technique is not used. Thereby in Wang [2004], S. Wang proposed a fitted finite volume method with the corresponding convergence proof in space to tackle the degeneracy of the Black-Scholes PDE. Moreover, under less restrictive and more realistic assumptions, a convergence proof in space of the fitted finite volume method for pricing American options is proposed in Wang et al. [2006]. Furthermore, in Angermann and Wang [2007], a rigorous convergence proof of a fully discretized scheme using the fitted finite volume method and $\theta-$Euler method for pricing both American and European options is provided. Note that convergence proofs of the TPFA method are provided in Eymard et al. [2000], Tambue [2016] for non degenerated parabolic PDE. Their proofs are based on the fact that the diffusion coefficient can not reach zero, therefore such proofs can not be extended to the degenerated Black Scholes PDE where the diffusion coefficient is zero at $s = 0$ (stock price is zero). To the best of our knowledge, the convergence of classical TPFA method for degenerated PDE has been lacking in the literature due to the complexity of that degeneracy.
In this Chapter, we fill the gap by providing a rigorous convergence proof of a fully discretized scheme using the classical TPFA method for degenerated Black Scholes PDE in one dimension. Furthermore, we also derive the fitted TPFA method for the degenerated Black Scholes PDE by combining the classical TPFA method and the fitted finite volume method presented in Angermann and Wang [2007], Wang et al. [2006] and provide rigorous convergence proof of a fully discretized scheme where the time discretization is performed using the classical Euler methods. Note that the fitted finite volume method, presented in Angermann and Wang [2007], Wang et al. [2006], in this combination is meant to handle the degeneracy of the PDE when the stock price approaches zero.
This Chapter is organized as follows. In Section 2.2, notations and mathematical setting of the continuous problem are provided. In Section 2.3, the spatial discretization using the standard finite volume method with Two Point Flux Approximation are provided along with the corresponding novel

fitted scheme. The coercivity proofs of the corresponding discrete bilinear forms are also provided to ensure the existence and uniqueness of the discrete solution after TPFA spatial discretization and fitted TPFA spatial discretization. The full discretization of the Black-Scholes degenerated PDE and the convergence results are performed in Section 2.4. Note that the temporal discretization is performed using the standard $\theta-$ Euler method. Finally, numerical experiments are given in Section 2.5 to support theoretical results.

2.2 Mathematical setting for the one dimensional Black-Scholes PDE

As we introduced in paragraph 1.4.1, we define the weighted Sobolev spaces and the corresponding norms for solving the one dimensional option pricing problem. For an open set $\Omega \subset \mathbb{R}$, the space of square integrable functions is denoted $L^2(\Omega)$. We denote also by $C(\Omega)$ (respectively $C(\overline{\Omega})$) the set of continuous functions over Ω (respectively on $\overline{\Omega}$). For any Hilbert space $G(\Omega)$ of classes of functions defined on Ω, we let $L^2\big(0,T;G(\Omega)\big)$ denote the space defined by

$$L^2\Big(0,T;G(\Omega)\Big) = \left\{ v/v(\cdot,t) \in G(\Omega) \ a.e \ in \ (0,T); ||v(\cdot,t)||_G \in L^2\big(0,T\big) \right\}, \tag{2.1}$$

where $||\cdot||_G$ denotes the natural norm on $G(\Omega)$. The norm on this space is denoted by $||\cdot||_{L^2\left(0,T;G(\Omega)\right)}$ and is defined by

$$||v||_{L^2\left(0,T;G(\Omega)\right)} = \left(\int_0^T ||v(\cdot,t)||_G^2 dt \right)^{1/2}. \tag{2.2}$$

The Black-Scholes operator being degenerated, we introduce a weighted L^2-norm $||\cdot||_\omega$ defines by

$$||v||_\omega = \left(\int_\Omega x^2 v^2 dx \right)^{1/2}. \tag{2.3}$$

The space of all weighted square integrable functions is defined as

$$L^2_\omega(\Omega) = \left\{ v : ||v||_\omega < \infty \right\}, \tag{2.4}$$

and the corresponding weighted inner product on $L^2_\omega(\Omega)$ by

$$(u,v)_\omega = \int_\Omega x^2 uv dx. \tag{2.5}$$

Thereby, we define the weighted Sobolev spaces as follows:

$$H^1_\omega(\Omega) = \left\{ v \in L_2(\Omega) : \exists g \in L^2_\omega(\Omega) \quad such \ \ that \quad \int_\Omega v\varphi' = -\int_\Omega g\varphi \qquad \forall \varphi \in C_c(\Omega) \right\}. \tag{2.6}$$

Note that in (2.6), $g = v'$ is the weak derivative. We also denote by

$$H^1_{0,\omega}(\Omega) = \left\{ v : v \in H^1_\omega(\Omega) \ \ and \ \ v\big|_{\partial\Omega} = 0 \right\}. \tag{2.7}$$

Using the inner products on $L^2(\Omega)$ and $L^2_\omega(\Omega)$, we define the norm $||\cdot||_{1,\omega}$ on $H^1_\omega(\Omega)$ by

$$\big|\big|v\big|\big|_{1,\omega} = \left[\big|\big|v\big|\big|^2_{L^2(\Omega)} + \big|\big|v'\big|\big|^2_\omega \right]^{1/2} = \left[(v,v) + (xv', xv') \right]^{1/2}. \tag{2.8}$$

Without loss the generality [1], let us consider the Black-Scholes equation PDE

$$LV := \frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2(t)x^2\frac{\partial^2 V}{\partial x^2} - r(t)x\frac{\partial V}{\partial x} + r(t)V = 0, \tag{2.9}$$

in $(x,t) \in \Omega = (0, x_{\max}) \times (0,T]$, where V is the option value, x the stock price, σ the volatility, r the risk-free interest, T is the maturity time and t is the time to maturity. The corresponding initial and boundary conditions are:

$$\left\{ \begin{array}{lcl} V(x,0) & = & g_1(x) \\ V(0,t) & = & g_2(t) \\ V(x_{\max},t) & = & g_3(x), \end{array} \right. \tag{2.10}$$

where g_1, g_2 and g_3 are functions depending on type of options we are pricing. Our study is conducted under the following assumption.

Assumption 3 *We assume that the coefficient r and σ are sufficiently smooth and satisfy*

$$\underline{r} \leq r(t) \leq \bar{r} \qquad \underline{\sigma} \leq \sigma(t) \leq \bar{\sigma}, \tag{2.11}$$

and we denote as follows:

$$\beta := \sup_{t\in[0,T]} \sigma^2(t). \tag{2.12}$$

Multiplying by $e^{\beta t}$ and adding $f(x,t) = -e^{\beta t}LV_0(x)$ to both sides of (2.9), we can therefore transform the boundary conditions in (2.10) to homogeneous Dirichlet boundary conditions by using the following linear transformation

$$V_0(x,t) = g_2(t) + \frac{x}{x_{\max}}\Big(g_3(t) - g_2(t)\Big). \tag{2.13}$$

Furthermore, we introduce a new variable $u = e^{\beta t}\big(V - V_0\big)$ and we get this new PDE in its following divergence form:

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left[a(t)x^2\frac{\partial u}{\partial x} + b(t)xu\right] + c(t)u = f(x,t), \tag{2.14}$$

where

$$a(t) = \frac{1}{2}\sigma^2(t), \quad b(t) = r(t) - \sigma^2(t), \quad c(t) = 2r(t) + \beta - \sigma^2(t), \tag{2.15}$$

with the following initial and homogeneous boundary conditions

$$\left\{ \begin{array}{ll} u(0,t) = 0 = u(x_{\max},t) & t \in [0,T), \\ u(x,0) = g_1(x) - V_0(x) & x \in \Omega. \end{array} \right. \tag{2.16}$$

It can be proved that solving (2.14)-(2.16) is equivalent to solve the following problem:

Problem 2 *Find the function $u \in H^1_{0,\omega}(\Omega)$ such that*

$$\begin{array}{rcll} (u'(t), v) + A(u,v;t) & = & \Big(f, v\Big) & \forall\ v \in H^1_{0,\omega}(\Omega), \\ u_0 & = & \max(K - x, 0), & \end{array} \tag{2.17}$$

with

$$A(u,v;t) := \Big(ax^2u' + bxu, v'\Big) + \Big(cu, v\Big). \tag{2.18}$$

Theorem 6 *Under Asumption 3, **Problem 1** has a unique solution.*

Proof of theorem 5 The proof is a particular case of Theorem 5 proof for $n = 1$, in one dimension.

[1] The American options are solved by just adding a nonlinear term called Penalty

2.3 The finite volume formulation

2.3.1 Finite volume grid and discrete representation of the exact solution

Let Ω be subdivided into sub-intervals as follows:

$$\Omega_i = [x_i, x_{i+1}] \quad i = 0, 1, \ldots, N, \tag{2.19}$$

with $0 = x_0 < x_1 < x_2 < \ldots < x_N < x_{N+1} = x_{max}$ and $h_i = x_{i+1} - x_i$. We also defined the following mid-points of the intervals Ω_i by

$$x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2} \quad \text{and} \quad x_{i+\frac{1}{2}} = \frac{x_i + x_{i+1}}{2} \qquad \text{for} \quad i = 0, 1, \ldots, N+1,$$

with $x_{-\frac{1}{2}} = x_0$ and $x_{N+\frac{3}{2}} = x_{\max}$. These mid-points help us to define another partition K_i of Ω, called dual partition, defined by

$$K_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \qquad l_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, \qquad i = 0, 1.., N.$$

Assumption 4 *[Local quasi-uniformity of the spatial mesh]*
There exists a constant $c > 0$ such that

$$\frac{l_{i+1}}{c} \leq l_i \leq c l_{i+1} \qquad i = 0, 1, \ldots, N. \tag{2.20}$$

Since the dual partition K_i is linked to the partition Ω_i, Assumption 4 implies that

$$\frac{h_{i+1}}{c} \leq h_i \leq c h_{i+1} \quad i = 0, 1, \ldots, N. \tag{2.21}$$

We can now apply the finite volume method by integrating (2.14) over each interval K_i,
for $i = 1, 2, \ldots, N$, as follows:

$$\int_{K_i} \frac{\partial u}{\partial \tau} dx - \int_{K_i} \frac{\partial}{\partial x}\left[ax^2 \frac{\partial u}{\partial x} + bxu\right] dx + \int_{K_i} cu dx = \int_{K_i} f(x,t) dx. \tag{2.22}$$

Multiplying (2.22) by an arbitrary real number v_i for each $i = 1, 2, \ldots, N$, and summing them up, we get

$$\sum_{i=1}^{N} \int_{K_i} \frac{\partial u}{\partial \tau} v_i dx - \sum_{i=1}^{N} \int_{K_i} \frac{\partial}{\partial x}\left[ax^2 \frac{\partial u}{\partial x} + bxu\right] v_i dx + \sum_{i=1}^{N} \int_{K_i} cuv_i dx = \sum_{i=1}^{N} \int_{K_i} f(x,t) v_i dx. \tag{2.23}$$

Besides, for a function $v \in C(\overline{\Omega})$ we define the mass lumping operator L_h as follows:

$$\begin{aligned} L_h : C(\overline{\Omega}) &\longrightarrow L^\infty(\Omega) \\ v &\mapsto L_h v\big|_{K_i} := v(x_i), \qquad i = 1, 2, \ldots, N. \end{aligned}$$

Moreover, if the function v satisfies homogeneous Dirichlet boundary conditions,
we have $L_h v\big|_{K_0} = L_h\big|_{K_{N+1}} = 0$. Then using the operator L_h leads to

$$\sum_{i=1}^{N} \int_{K_i} u v_i dx = \sum_{i=1}^{N} \int_{K_i} u L_h v dx = \sum_{i=0}^{N+1} \int_{K_i} u L_h v dx = \int_{\Omega} u L_h v dx = \Big(u, L_h v\Big) \tag{2.24}$$

Using similar transformation as in (2.24), we can re-write (2.23) as follows:

$$\big(-\dot{u}(t), L_h v\big) + \hat{a}_h\big(u(t), v; t\big) = \big(f(t), L_h v\big), \tag{2.25}$$

where

$$\hat{a}_h\big(\omega, v; t\big) := \sum_{i=1}^{N} \Big(F(\omega(x_{i+\frac{1}{2}})) - F\big(\omega(x_{i-\frac{1}{2}})\big)\Big) L_h v(x_i) + \big(c(t)\omega, L_h \omega\big), \tag{2.26}$$

with F denoting the continuous flux defined by

$$F\big(\omega(x_{i+\frac{1}{2}})\big) := -a(t) x_{i+\frac{1}{2}}^2 \frac{\partial \omega}{\partial x}(x_{i+\frac{1}{2}}) - b(t) x_{i+\frac{1}{2}} \omega\big(x_{i+\frac{1}{2}}\big) \qquad i = 0, 1, \ldots, N. \tag{2.27}$$

Note that (2.25) is a representation of the exact solution on the dual partition K_i and will play a key role in our error analysis.

2.3.2 The Two Point Flux Approximation (TPFA) method

We are now ready to approximate u in the dual partition K_i. Indeed we denote by $u(x_i,t) \approx u_i(t)$. To approximate some integral terms of (2.22), we use the mid-quadrature rule as follows:

$$\int_{K_i} \frac{\partial u}{\partial \tau} dx \approx l_i \frac{du_i}{d\tau}, \quad \int_{K_i} cu dx \approx l_i c u_i, \quad \int_{K_i} f(x,t) dx \approx l_i f_i \qquad f_i = f(x_i,t). \tag{2.28}$$

Since the Black-Scholes equation (2.9) is one dimensional in space, instead of the Multi-Point Flux Approximation (MPFA) method we use the Two-Point Flux Approximation (TPFA) method to approximate the following second term of (2.22)

$$\int_{K_i} \frac{\partial}{\partial x}\left[ax^2 \frac{\partial u}{\partial x}\right] dx. \tag{2.29}$$

The Two-Point Flux Approximation (TPFA) method is used to approximate the term (2.29) as follows:

$$\int_{K_i} \frac{\partial}{\partial x}\left[a(t)x^2 \frac{\partial u}{\partial x}\right] dx \quad = \quad \left[a(t)x^2 \frac{\partial u}{\partial x}\right]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} = a(t)x^2 \frac{\partial u}{\partial x}\bigg|_{x_{i+\frac{1}{2}}} - a(t)x^2 \frac{\partial u}{\partial x}\bigg|_{x_{i-\frac{1}{2}}}.$$

Let us set

$$H(x) = k(x,t)\frac{\partial u}{\partial x} \qquad with \qquad k(x,t) = \frac{1}{2}\sigma^2(t)x^2. \tag{2.30}$$

Over an interval K_i, $k(x,t)$ in (2.30) will be replaced by its average value defined as follows:

$$k_i = \frac{1}{l_i}\int_{K_i} \frac{1}{2}\sigma^2(t)x^2 dx = \frac{1}{6}\sigma^2(t)\frac{x_{i+\frac{1}{2}}^3 - x_{i-\frac{1}{2}}^3}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}}, \tag{2.31}$$

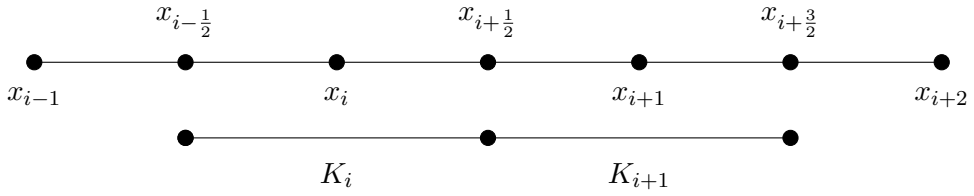


Figure 2.1: Interval

Thereby from Figure (2.1), $H_{i+\frac{1}{2}}$ can be evaluated at each side of $x_{i+\frac{1}{2}}$ as follows:

$$H_{i+\frac{1}{2}} = 2k_i \frac{u_{i+\frac{1}{2}} - u_i}{l_i}, \quad H_{i+\frac{1}{2}} = 2k_{i+1}\frac{u_{i+1} - u_{i+\frac{1}{2}}}{l_{i+1}}. \tag{2.32}$$

Using the continuity of the flux at the interface $x_{i+\frac{1}{2}}$, we can equate the two terms in(2.32). This leads to

$$u_{i+\frac{1}{2}} = \frac{\frac{k_i}{l_i}u_i + \frac{k_{i+1}}{l_{i+1}}u_{i+1}}{\frac{k_i}{l_i} + \frac{k_{i+1}}{l_{i+1}}}. \tag{2.33}$$

By setting $T_i = \frac{k_i}{l_i}$, we can rewrite $H_{i+\frac{1}{2}}$ in (2.32) as follows:

$$H_{i+\frac{1}{2}} = \tau_{i+\frac{1}{2}}(u_{i+1} - u_i), \quad \tau_{i+\frac{1}{2}} = \frac{2T_iT_{i+1}}{T_i + T_{i+1}}. \tag{2.34}$$

Besides, to approximate the third integral term of (2.22), we use the upwind method and it gives

$$\int_{K_i} \frac{\partial}{\partial x}\Big[b(t)xu\Big]dx \;=\; \Big[b(t)xu\Big]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \approx b(t)\Big(x_{i+\frac{1}{2}}u_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}u_{i-\frac{1}{2}}\Big) \tag{2.35}$$

with

$$u_{i+\frac{1}{2}} = \begin{cases} u_i & if \quad b > 0, \\ u_{i+1} & if \quad b < 0. \end{cases} \tag{2.36}$$

Therefore, the discrete formulation of (2.22) is given by

$$l_i\frac{du_i}{d\tau} + F_h(u_{i+\frac{1}{2}}) - F_h(u_{i-\frac{1}{2}}) - l_i c u_i = l_i f_i \qquad i = 1, 2, ..., N, \tag{2.37}$$

where

$$\begin{aligned} F_h(u_{i+\frac{1}{2}}) &= -\tau_{i+\frac{1}{2}}(u_{i+1} - u_i) - x_{i+\frac{1}{2}}\Big(b^+u_i + b^-u_{i+1}\Big), \\ F_h(u_{i-\frac{1}{2}}) &= -\tau_{i-\frac{1}{2}}(u_i - u_{i-1}) - x_{i-\frac{1}{2}}\Big(b^+u_{i-1} + b^-u_i\Big), \end{aligned} \tag{2.38}$$

with $b^+ = \max(b, 0)$ and $b^- = \min(b, 0)$. Moreover, in order to analyse the above scheme, it is convenient to rewrite it in a discrete variational form. Multiplying equation (2.37) by arbitrary real numbers v_i and summing the result over all the intervals K_i of Ω, we get:

$$\sum_{i=1}^{N} l_i\frac{du_i}{d\tau}v_i + \sum_{i=1}^{N}\Big(F_h(u_{i+\frac{1}{2}}) - F_h(u_{i-\frac{1}{2}})\Big)v_i + c\sum_{i=1}^{N} l_i u_i v_i = \sum_{i=1}^{N} l_i f_i v_i. \tag{2.39}$$

Let us denote by $V_h \subset H^1_\omega(\Omega)$ the space of continuous functions that are piecewise continuous over the grid (K_i) of Ω. Thereby, the TPFA method (2.37) is equivalent to

$$a_h(u_h, v_h) = a_h^1(u_h, v_h) + a_h^2(u_h, v_h) + c\sum_{i=1}^{N} l_i u_i v_i \qquad u_h, v_h \in V_h, \tag{2.40}$$

with

$$a_h^1(u_h, v_h) = \sum_{i=1}^{N}\Big[- \tau_{i+\frac{1}{2}}(u_{i+1} - u_i) + \tau_{i-\frac{1}{2}}(u_i - u_{i-1})\Big]v_i, \tag{2.41}$$

and

$$a_h^2(u_h, v_h) = \sum_{i=1}^{N}\Big[- x_{i+\frac{1}{2}}\Big(b^+u_{i+1} + b^-u_i\Big) + x_{i-\frac{1}{2}}\Big(b^+u_{i-1} + b^-u_i\Big)\Big]v_i. \tag{2.42}$$

Let us notice that we can rewrite the bilinear form in (2.40) as

$$a_h(u_h, v_h) = \sum_{i=1}^{N}\Big(F_h(u_{i+\frac{1}{2}}) - F_h(u_{i-\frac{1}{2}})\Big)v_i + c\sum_{i=1}^{N} l_i u_i v_i, \tag{2.43}$$

where F_h is the discrete flux in given in (2.38).

2.3.3 The fitted Two Point Flux Approximation method

Since the PDE (2.14) is degenerated when x approaches zero, the second term of (2.22), at the point $x = x_{1/2}$, will be approximated using the fitted finite volume method introduced Wang [2004]. The fitted finite volume method will consist to solve a two-point value problem over the interval K_1. Thereby, from Wang [2004], we have

$$ax^2\frac{\partial u}{\partial x} + bxu\Big|_{x_{1/2}} \;\approx\; \frac{1}{4}x_1[(a+b)u_1 - (a-b)u_0]. \tag{2.44}$$

On the the rest on the study domain ($K_i,\ i = 2, .., N$), we apply the TPFA method coupled to the upwind method introduced in Section 2.3.2. Therefore, the discrete approximation of (2.22) by the fitted TPFA method is given by

$$l_i \frac{du_i}{d\tau} + G_h(u_{i+\frac{1}{2}}) - G_h(u_{i-\frac{1}{2}}) - l_i c u_i = l_i f_i \qquad i = 1, 2, ..., N, \tag{2.45}$$

where

$$\begin{aligned} G_h(u_{i+\frac{1}{2}}) &= -\tau_{i+\frac{1}{2}}(u_{i+1} - u_i) - x_{i+\frac{1}{2}}\Big(b^+ u_i + b^- u_{i+1}\Big) \\ G_h(u_{i-\frac{1}{2}}) &= -\tau_{i-\frac{1}{2}}(u_i - u_{i-1}) - x_{i-\frac{1}{2}}\Big(b^+ u_{i-1} + b^- u_i\Big), \qquad i \neq 1 \\ G_h(u_{1/2}) &= -\frac{1}{4} x_1 (a+b) u_1, \end{aligned} \tag{2.46}$$

with $b^+ = \max(b, 0)$ and $b^- = \min(b, 0)$. Thereby, the discrete spatial formulation is given by

$$\sum_{i=1}^{N} l_i \frac{du_i}{d\tau} v_i + \sum_{i=1}^{N} \Big(G_h(u_{i+\frac{1}{2}}) - G_h(u_{i-\frac{1}{2}})\Big) v_i + c \sum_{i=1}^{N} l_i u_i v_i = \sum_{i=1}^{N} l_i f_i v_i. \tag{2.47}$$

We define the corresponding bilinear form b_h by

$$b_h(u_h, v_h) = b_h^1(u_h, v_h) + b_h^2(u_h, v_h) + c \sum_{i=1}^{N} l_i u_i v_i \qquad u_h, v_h \in V_h, \tag{2.48}$$

with

$$b_h^1(u_h, v_h) = \sum_{i=1}^{N} \Big[- \tau_{i+\frac{1}{2}}(u_{i+1} - u_i)\Big] + \sum_{i=2}^{N} \Big[\tau_{i-\frac{1}{2}}(u_i - u_{i-1})\Big] v_i, \tag{2.49}$$

and

$$b_h^2(u_h, v_h) = \frac{1}{4} x_1 (a+b) u_1^2 + \sum_{i=1}^{N} \Big[- x_{i+\frac{1}{2}}\Big(b^+ u_i + b^- u_{i+1}\Big)\Big] v_i + \sum_{i=2}^{N} \Big[x_{i-\frac{1}{2}}\Big(b^+ u_{i-1} + b^- u_i\Big)\Big] v_i. \tag{2.50}$$

The bilinear form b_h in (2.48) can be rewritten as:

$$b_h(u_h, v_h) = \sum_{i=1}^{N} \Big(G_h(u_{i+\frac{1}{2}}) - G_h(u_{i-\frac{1}{2}})\Big) v_i + c \sum_{i=1}^{N} l_i u_i v_i, \tag{2.51}$$

where G_h is the discrete flux given by the fitted TPFA in (2.46).

2.3.4 Coercivity and Flux consistency for TPFA and fitted TPFA

Let us denote by $\big(\cdot, \cdot\big)_h$ the scalar product on $C(\overline{\Omega}) \supset V_h$ given by

$$\big(u, v\big)_h = (L_h u, L_h v) = \sum_{i=1}^{N} l_i u_i v_i \qquad u, v \in C(\overline{\Omega}), \tag{2.52}$$

and its corresponding norm $||\cdot||_{0,h}$ given by

$$||v||_{0,h}^2 = \sum_{i=1}^{N} l_i v_i^2. \tag{2.53}$$

We define the discrete $H_0^1(\Omega)$ norm by

$$||u_h||_{0,\omega}^2 = \sum_{i=1}^{N} \tau_{i+\frac{1}{2}} |u_{i+1} - u_i|^2, \tag{2.54}$$

and weighted discrete $H^1_\omega-$ norm is then

$$||u_h||^2_{\omega,d} = ||u_h||^2_{0,\omega} + ||u_h||^2_{0,h}. \tag{2.55}$$

Indeed it is easy to show that $||.||_{0,\omega}$ is a semi-norm in V_h since $\tau_{i+\frac{1}{2}} > 0$.

Theorem 7 *[Coercivity of bilinear forms]*
Under the regularity of the mesh (see Assumption 4) and Assumption 3, there exists a constant $\alpha > 0$ independent of h such that,

$$a_h(u_h, u_h) \geq \alpha ||u_h||_{\omega,d} \qquad \forall u_h \in V_h. \tag{2.56}$$

where a_h is the bilinear form given by (2.40) *for the TPFA method. Simlilarly, when the fitted TPFA method* (2.45) *is used for the space discretization, there exists a constant $\gamma > 0$ independent of h such that*

$$b_h(u_h, u_h) \geq \gamma ||u_h||_{\omega,d} \qquad \forall u_h \in V_h. \tag{2.57}$$

where the bilinear form b_h is given by (2.48).

Remark 1 *Note that using the coercivity properties in* (2.56) *and* (2.57)*, with the fact that the linear mapping $v \to (f, v)_h$ is continuous in V_h, the existence and uniqueness of the discrete solution u_h is ensure for both the TPFA and fitted TPFA methods in* (2.39) *and* (2.47)*. The proof is done exactly as for the continuous case (see Theorem 4 or [Haslinger et al., 2013, Theorem 1.33]).*

Proof of Theorem 6 Here we distinguish two cases which are:

$\underline{1^{st}\textit{case}}$: The standard TPFA method is used for the spatial discretization.

In this case, the discrete flux is given by (2.46) and the corresponding bilinear form (2.40) is

$$a_h(u_h, v_h) = a^1_h(u_h, v_h) + a^2_h(u_h, v_h) + c\sum_{i=1}^{N} l_i u_i v_i, \qquad u_h, v_h \in V_h,$$

with

$$a^1_h(u_h, v_h) = \sum_{i=1}^{N} \Big[-\tau_{i+\frac{1}{2}}(u_{i+1} - u_i) + \tau_{i-\frac{1}{2}}(u_i - u_{i-1}) \Big] v_i,$$

and

$$a^2_h(u_h, v_h) = \sum_{i=1}^{N} \Big[-x_{i+\frac{1}{2}}\Big(b^+ u_{i+1} + b^- u_i\Big) + x_{i-\frac{1}{2}}\Big(b^+ u_{i-1} + b^- u_i\Big) \Big] v_i. \tag{2.58}$$

Thereby,

$$\begin{aligned}
a^1_h(u_h, u_h) &= \sum_{i=1}^{N} \Big[-\tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big) + \tau_{i-\frac{1}{2}}\big(u_i - u_{i-1}\big) \Big] u_i \\
&= \sum_{i=1}^{N} -\tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)u_i + \sum_{k=0}^{N-1} \tau_{k+\frac{1}{2}}\big(u_{k+1} - u_k\big)u_{k+1} \\
&= \tau_{1/2}\big(u_1 - u_0\big)u_1 + \sum_{i=1}^{N-1} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)\big(u_{i+1} - u_i\big) - \tau_{N+\frac{1}{2}}\big(u_{N+1} - u_N\big)u_N \\
&= \tau_{1/2}\big(u_1 - u_0\big)\big(u_1 - u_0\big) + \sum_{i=1}^{N-1} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)^2 + \tau_{N+\frac{1}{2}}\big(u_{N+1} - u_N\big)\big(u_{N+1} - u_N\big) \\
&= \tau_{1/2}\big(u_1 - u_0\big)^2 + \sum_{i=1}^{N} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)^2 \\
a^1_h(u_h, u_h) &\geq ||u_h||^2_{0,\omega},
\end{aligned} \tag{2.59}$$

and also

$$\begin{aligned}
a_h^2(u_h,u_h) &= \sum_{i=1}^{N}\Big[-x_{i+\frac{1}{2}}\Big(b^+u_{i+1}+b^-u_i\Big)+x_{i-\frac{1}{2}}\Big(b^+u_{i-1}+b^-u_i\Big)\Big]u_i\\
&= \sum_{i=1}^{N}-x_{i+\frac{1}{2}}\Big(b^+u_{i+1}+b^-u_i\Big)u_i+\sum_{i=1}^{N}x_{i-\frac{1}{2}}\Big(b^+u_{i-1}+b^-u_i\Big)u_i\\
&= \sum_{i=1}^{N}-x_{i+\frac{1}{2}}\Big(b^+u_{i+1}+b^-u_i\Big)u_i+\sum_{i=0}^{N-1}x_{i+\frac{1}{2}}\Big(b^+u_i+b^-u_{i+1}\Big)u_{i+1}\\
&= x_{1/2}\Big(b^+u_0+b^-u_1\Big)u_1+\sum_{i=1}^{N-1}x_{i+\frac{1}{2}}\Big[b^-\big(u_{i+1}^2-u_i^2\big)\Big]-x_{N+\frac{1}{2}}\Big(b^+u_{N+1}+b^-u_N\Big)u_i\\
&= -b^-x_{1/2}\big(u_0^2-u_1^2\big)-b^-\sum_{i=1}^{N-1}x_{i+\frac{1}{2}}\Big(u_i^2-u_{i+1}^2\Big)-b^-x_{N+\frac{1}{2}}\Big(u_N^2-u_{N+1}^2\Big)\\
&= -b^-\sum_{i=0}^{N}x_{i+\frac{1}{2}}\Big(u_i^2-u_{i+1}^2\Big).
\end{aligned}$$

Besides,

$$\begin{aligned}
a_h^2(u_h,u_h) &= -b^-\Bigg[\sum_{i=0}^{N}\Big(x_{i-\frac{1}{2}}+l_i\Big)u_i^2-\sum_{i=0}^{N}x_{i+\frac{1}{2}}u_{i+1}^2\Bigg]\\
&= -b^-\Bigg[\sum_{i=-1}^{N-1}\Big(x_{i+\frac{1}{2}}+l_{i+1}\Big)u_{i+1}^2-\sum_{i=0}^{N}x_{i+\frac{1}{2}}u_{i+1}^2\Bigg]\\
&= -b^-\Bigg[x_{-\frac{1}{2}}u_0^2+\sum_{i=-1}^{N-1}l_{i+1}u_{i+1}^2-x_{N+\frac{1}{2}}u_{N+1}^2\Bigg]\\
&= -b^-\sum_{i=0}^{N}l_iu_i^2\\
a_h^2(u_h,u_h) &\geq 0. \qquad (2.60)
\end{aligned}$$

Then, using (2.59) and (2.60) yields to

$$a_h(u_h,u_h)\geq \big|\big|u_h\big|\big|_{0,\omega}+c\sum_{i=1}^{N}l_iu_i^2. \qquad (2.61)$$

Moreover, we have

$$c=2r+\beta-\sigma^2>0, \qquad (2.62)$$

therefore, this yields to

$$a_h(u_h,u_h)\geq \alpha\Big(\big|\big|u_h\big|\big|_{0,\omega}+\big|\big|u_h\big|\big|_{0,h}\Big), \qquad (2.63)$$

with $\alpha=\min(1,c)$. Hence we get

$$a_h(u_h,u_h)\geq \alpha\big|\big|u_h\big|\big|_{\omega,d}. \qquad (2.64)$$

$\underline{2^{nd}\text{ case}}$ The fitted TPFA method is used for the spatial discretization.

Here, G_h defined in (2.45), gives the discrete flux. Thereby, the corresponding bilinear form (2.48) is given by

$$b_h(u_h,v_h)=b_h^1(u_h,v_h)+b_h^2(u_h,v_h)+c\sum_{i=1}^{N}l_iu_iv_i \qquad u_h,v_h\in V_h. \qquad (2.65)$$

Rearranging the terms of b_h^1 gives

$$\begin{aligned}
b_h^1(u_h, v_h) &= \sum_{i=1}^{N} \Big[-\tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)\Big] u_i + \sum_{i=2}^{N} \tau_{i-\frac{1}{2}} \Big[\big(u_i - u_{i-1}\big)\Big] u_i \\
&= -\tau_{\frac{3}{2}}\big(u_2 - u_1\big)u_1 + \sum_{i=2}^{N} \Big[-\tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big) + \tau_{i-\frac{1}{2}}\big(u_i - u_{i-1}\big)\Big] u_i \\
&= -\tau_{3/2}\big(u_2 - u_1\big)u_1 + \sum_{i=2}^{N} -\tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)u_i + \sum_{k=1}^{N-1} \tau_{k+\frac{1}{2}}\big(u_{k+1} - u_k\big)u_{k+1} \\
&= -\tau_{3/2}\big(u_2 - u_1\big)u_1 + \tau_{3/2}\big(u_2 - u_1\big)u_2 + \sum_{i=2}^{N-1} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)^2 + \tau_{N+\frac{1}{2}}\big(u_{N+1} - u_N\big)^2 \\
&= \tau_{3/2}\big(u_2 - u_1\big)^2 + \sum_{i=2}^{N} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)^2 \\
&= \sum_{i=1}^{N} \tau_{i+\frac{1}{2}}\big(u_{i+1} - u_i\big)^2 \\
b_h^1(u_h, u_h) &= ||u_h||_{0,\omega}^2. \qquad (2.66)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
b_h^2(u_h, u_h) &= \frac{1}{4}x_1(a+b)u_1^2 + \sum_{i=1}^{N} \Big[-x_{i+\frac{1}{2}}\Big(b^+u_i + b^-u_{i+1}\Big)\Big] u_i + \sum_{i=2}^{N} \Big[x_{i-\frac{1}{2}}\Big(b^+u_{i-1} + b^-u_i\Big)\Big] u_i \\
&= \frac{1}{4}x_1(a+b)u_1^2 + \sum_{i=1}^{N} \Big[-x_{i+\frac{1}{2}}\Big(b^+u_i + b^-u_{i+1}\Big)\Big] u_i + \sum_{i=1}^{N-1} \Big[x_{i+\frac{1}{2}}\Big(b^+u_i + b^-u_{i+1}\Big)\Big] u_{i+1} \\
&= \frac{1}{4}x_1(a+b)u_1^2 + \sum_{i=1}^{N-1} x_{i+\frac{1}{2}}b^-\big(u_{i+1}^2 - u_i^2\big) - x_{N+\frac{1}{2}}b^-u_N^2 \\
b_h^2(u_h, u_h) &= \frac{1}{4}x_1(a+b)u_1^2 + \sum_{i=1}^{N} x_{i+\frac{1}{2}}b^-\big(u_{i+1}^2 - u_i^2\big). \qquad (2.67)
\end{aligned}$$

1. $\underline{1^{st}case}$: $b > 0$.

   if $b > 0$ then $b^- = 0$. This gives

$$b_h^2(u_h, u_h) = \frac{1}{4}x_1(a+b)u_1^2 > 0 \qquad (2.68)$$

2. $\underline{2^{nd}case}$: $b \leq 0$

$$\begin{aligned}
b_h^2(u_h, u_h) &= \frac{1}{4}x_1(a+b)u_1^2 + \sum_{i=1}^{N} x_{i+\frac{1}{2}}b^-\big(u_{i+1}^2 - u_i^2\big) \\
&= \frac{1}{4}x_1(a+b)u_1^2 - b^-\left(\sum_{i=1}^{N}\Big(x_{i-\frac{1}{2}} + l_i\Big)u_i^2 - \sum_{i=1}^{N} x_{i+\frac{1}{2}}u_{i+\frac{1}{2}}^2\right) \\
&= \frac{1}{4}x_1(a+b)u_1^2 - b^-\left(\sum_{i=0}^{N-1}\Big(x_{i+\frac{1}{2}} + l_{i+1}\Big)u_{i+1}^2 - \sum_{i=1}^{N} x_{i+\frac{1}{2}}u_{i+\frac{1}{2}}^2\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4}x_1(a+b)u_1^2 - b^-\big(x_{1/2}+l_1\big)u_1^2 - \sum_{i=1}^{N} l_i u_i^2 - x_{N+\frac{1}{2}}u_{N+1}^2 \\
&= \frac{1}{4}x_1(a+b)u_1^2 - x_{3/2}b^- u_1^2 - b^- \sum_{i=1}^{N} l_i u_i^2 \\
&= \frac{1}{4}x_1 a u_1^2 + \big(\frac{1}{4}x_1 b - x_{3/2}b^-\big)u_1^2 - b^- \sum_{i=1}^{N} l_i u_i^2 \\
b_h^2(u_h,u_h) &\geq 0. \qquad (2.69)
\end{aligned}
$$

Thus for any b, we have

$$b_h^2(u_h,u_h) \geq 0, \qquad (2.70)$$

thereby, using (2.66) and (2.70) in (2.48), we have:

$$b_h(u_h,u_h) \geq \big|\big|u_h\big|\big|_{0,\omega} + c\sum_{i=1}^{N} l_i u_i^2. \qquad (2.71)$$

Besides, we have

$$c = 2r + \beta - \sigma^2 > 0. \qquad (2.72)$$

Therefore, this yields to

$$b_h(u_h,u_h) \geq \gamma\Big(\big|\big|u_h\big|\big|_{0,\omega} + \big|\big|u_h\big|\big|_{0,h}\Big), \qquad (2.73)$$

with $\gamma = \min(1,c)$. Hence we get

$$b_h(u_h,u_h) \geq \gamma\big|\big|u_h\big|\big|_{\omega,d}. \qquad (2.74)$$

Proposition 1 *Flux consistency*
Let I_h be the interpolation operator defines as follows

$$
\begin{aligned}
I_h : C(\overline{\Omega}) &\longrightarrow V_h \\
v &\mapsto I_h v(x) := \sum_{i=1}^{N} v(x_i)\phi_{x_i}(x), \qquad x \in \Omega
\end{aligned}
$$

where $\{\phi_{x_i}\}_{i=1}^N$, with $\phi_{x_i}(x_j) = \delta_{ij}$, is the nodal basis to $\{x_i\}_{i=1}^N$, $x_i \in K_i$. Let F be the total (continuous) flux function defined for $\omega \in \mathcal{C}(\bar{\Omega})$ by

$$F\big(\omega(x_{i+\frac{1}{2}})\big) := -k(x_{i+\frac{1}{2}})\frac{\partial \omega}{\partial x}\big(x_{i+\frac{1}{2}}\big) - b x_{i+\frac{1}{2}}\omega\big(x_{i+\frac{1}{2}}\big) \qquad x_{i+\frac{1}{2}} \in \Omega_i. \qquad (2.75)$$

When the TPFA method is applied for the spatial discretization i.e the discrete flux is given by F_h defined in (2.38), then for $\omega \in H^2_{0,\omega}(\Omega)$, there is a positive constant C_1 such that

$$\Big|F\big(w(x_{i+\frac{1}{2}})\big) - F_h\big(I_h w(x_{i+\frac{1}{2}})\big)\Big| \leq C_1 \int_{x_i}^{x_{i+1}} \Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)dx \qquad i = 0,1,...,N. \quad (2.76)$$

Similarly, when the fitted TPFA is applied for the spatial discretization i.e the discrete flux is given by G_h defined in (2.46), then for $\omega \in H^2_{0,\omega}(\Omega)$, there is a positive constant C_2 such that

$$\Big|F\big(w(x_{i+\frac{1}{2}})\big) - G_h\big(I_h w(x_{i+\frac{1}{2}})\big)\Big| \leq C_2 \int_{x_i}^{x_{i+1}} \Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)dx. \qquad i = 0,1,...,N. \quad (2.77)$$

Before proving the proposition (1), let us state the following lemma:

Lemma 1 *For $i = 1,2,\ldots,N$, there exist two constants C_3 and C_4 independent of h such that the transmissibility coefficient $\tau_{i+\frac{1}{2}}$ defined in (2.34) and its inverse are bounded as follows:*

$$\Big|\tau_{i+\frac{1}{2}}\Big| \leq C_3, \qquad \frac{1}{\tau_{i+\frac{1}{2}}} \leq C_4 h_i \qquad i = 0,1,\ldots,N. \qquad (2.78)$$

Proof of Lemma 1

On one hand, we have

$$\begin{aligned}
\left|\tau_{i+\frac{1}{2}}\right| &= \frac{k_i k_{i+1}}{l_{i+1}k_i + l_i k_{i+1}} \\
&= \frac{\sigma^2}{6}\frac{\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right)\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)}{l^2_{i+1}\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right) + l^2_i\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)} \\
&\leq \frac{\sigma^2}{6}\frac{\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right)}{l^2_i} \\
&\leq \frac{\sigma^2}{6}\frac{\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right)}{\left(x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}\right)^2} \\
&\leq \frac{\sigma^2}{6}\frac{x^2_{i+\frac{1}{2}} + x_{i+\frac{1}{2}}x_{i-\frac{1}{2}} + x^2_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} \\
&\leq \frac{\sigma^2}{6}\frac{x_{i+\frac{1}{2}}\left(1 + \frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}} + \left(\frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}\right)^2\right)}{1 - \frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}} \\
\left|\tau_{i+\frac{1}{2}}\right| &\leq \frac{\sigma^2 x_{\max}}{6}\frac{1}{1-X_i}\left(1 + X_i + X_i^2\right),
\end{aligned}$$

where

$$X_i = \frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}} \qquad \text{with} \quad 0 < X_i < 1. \tag{2.79}$$

Thereby, using the Taylor expansion, we have

$$\frac{1}{1-X_i} = 1 + X_i + \mathcal{O}(X_i^2). \tag{2.80}$$

Then there exists a constant M_1 such that

$$\frac{1}{1-X_i} \leq 1 + X_i + \mathbf{M_1}X_i^2 \leq 2 + \mathbf{M_1}. \tag{2.81}$$

Since $0 \leq X_i \leq 1$, we have also

$$0 \leq 1 + X_i + X_i^2 \leq 3, \tag{2.82}$$

thus using X_i in (2.81) and (2.82), we get

$$\left|\tau_{i+\frac{1}{2}}\right| \leq C_3 \tag{2.83}$$

with $C_3 = \frac{\beta}{2}\big(2 + \mathbf{M}_1\big)x_{\max}$.

On the other hand, we have

$$\begin{aligned}
\frac{1}{\tau_{i+\frac{1}{2}}} &= \frac{l_{i+1}k_i + l_i k_{i+1}}{k_i k_{i+1}} \\
&= \frac{l_{i+1} \times \frac{\sigma^2}{6l_i}\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right) + l_i \times \frac{\sigma^2}{6l_{i+1}}\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)}{\frac{\sigma^2}{6l_i}\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right) \cdot \frac{\sigma^2}{6l_{i+1}}\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)} \\
&= \frac{36 l_i l_{i+1}\left(l_{i+1} \times \frac{\sigma^2}{6l_i}\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right) + l_i \times \frac{\sigma^2}{6h_{i+1}}\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)\right)}{\sigma^4\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right)\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)} \\
&= \frac{6\sigma^2 l^2_{i+1}\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right) + 6\sigma^2 l^2_i\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)}{\sigma^4\left(x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}\right)\left(x^3_{i+\frac{3}{2}} - x^3_{i+\frac{1}{2}}\right)} \\
&= \frac{6l^2_{i+1}}{\sigma^2 x^3_{i+\frac{3}{2}}\left(1 - \left(\frac{x_{i+\frac{1}{2}}}{x_{i+\frac{3}{2}}}\right)^3\right)} + \frac{6l^2_i}{\sigma^2 x^3_{i+\frac{1}{2}}\left(1 - \left(\frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}\right)^3\right)} \\
&= \frac{6}{\sigma^2 x^2_{i+\frac{3}{2}}\left(1 - \left(\frac{x_{i+\frac{1}{2}}}{x_{i+\frac{3}{2}}}\right)^3\right)} \times \frac{l^2_{i+1}}{x_{i+\frac{3}{2}}} + \frac{6}{\sigma^2 x^2_{i+\frac{1}{2}}\left(1 - \left(\frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}\right)^3\right)} \times \frac{l^2_i}{x_{i+\frac{1}{2}}} \\
&\leq \frac{6}{\sigma^2 x^2_{\max}} \cdot \frac{1}{\frac{x^2_{i+\frac{1}{2}}}{x^2_{\max}}}\left[\frac{1}{1 - \left(\frac{x_{i+\frac{1}{2}}}{x_{i+\frac{3}{2}}}\right)^3} \times \frac{l^2_{i+1}}{x_{i+\frac{3}{2}}} + \frac{1}{1 - \left(\frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}\right)^3} \times \frac{l^2_i}{x_{i+\frac{1}{2}}}\right]. \qquad (2.84)
\end{aligned}$$

Moreover, we have

$$l_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \quad \textit{then} \quad l_i \leq x_{i+\frac{1}{2}} \qquad \text{thus} \qquad \frac{1}{x_{i+\frac{1}{2}}} \leq \frac{1}{l_i}.$$

$$\text{Similarly} \quad \frac{1}{x_{i+\frac{3}{2}}} \leq \frac{1}{l_{i+1}}.$$

Thereby, we get

$$\frac{l^2_i}{x_{i+\frac{1}{2}}} \leq l_i \qquad \text{and} \qquad \frac{l^2_{i+1}}{x_{i+\frac{3}{2}}} \leq l_{i+1}. \qquad (2.85)$$

Besides, we set

$$W_i = \frac{x_{i+\frac{1}{2}}}{x_{\max}}, \qquad Y_i = \frac{x_{i+\frac{1}{2}}}{x_{i+\frac{3}{2}}}, \qquad Z_i = \frac{x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}}}. \qquad (2.86)$$

Coming back to (2.84) and using Assumption 4, equations (2.85) and (2.20) lead to

$$\begin{aligned}
\frac{1}{\tau_{i+\frac{1}{2}}} &\leq \frac{6}{\sigma^2 x^2_{\max}} \cdot \frac{1}{W^2_i} \cdot \left[\frac{1}{1 - Y^3_i} l_{i+1} + \frac{1}{1 - Z^3_i} l_i\right], \\
\frac{1}{\tau_{i+\frac{1}{2}}} &\leq \frac{6}{\sigma^2 x^2_{\max}} \cdot \frac{1}{W^2_i} \cdot \left[\frac{c}{1 - Y^3_i} + \frac{1}{1 - Z^3_i}\right] l_i. \qquad (2.87)
\end{aligned}$$

Let us notice also that for $i = 0, 1, \ldots, N$, we have

$$0 < W_i < 1, \qquad 0 < Y_i < 1, \qquad 0 < Z_i < 1. \qquad (2.88)$$

Then

$$0 < W_i^2 < 1, \qquad 0 < Y_i^3 < 1, \qquad 0 < Z_i^3 < 1. \tag{2.89}$$

Using the Taylor expansion, gives

$$\begin{aligned} \frac{1}{W_i^2} &= 1 + (1 - W_i^2) + \mathcal{O}(W_i^4), \\ \frac{1}{1 - Y_i^3} &= 1 + Y_i^3 + \mathcal{O}(Y_i^6), \\ \frac{1}{1 - Z_i^3} &= 1 + Z_i^3 + \mathcal{O}(Z_i^6). \end{aligned}$$

This leads to

$$\begin{aligned} \frac{1}{W_i^2} &\leq 1 + (1 - W_i^2) + \mathbf{M_2} W_i^2 \leq 2 + \mathbf{M_2}, \\ \frac{1}{1 - Y_i^3} &\leq 1 + Y_i^3 + \mathbf{M_3} Y_i^6 \leq 2 + \mathbf{M_3}, \\ \frac{1}{1 - X_i} &\leq 1 + Z_i + \mathbf{M_4} X_i^2 \leq 2 + \mathbf{M_4}, \end{aligned} \tag{2.90}$$

where $\mathbf{M_2}, \mathbf{M_3}, \mathbf{M_4}$ are positive constants. Using (2.87),and (2.90), we get

$$\frac{1}{\tau_{i+\frac{1}{2}}} \leq \frac{6}{\sigma^2 x_{\max}^2} \cdot \Big(2 + \mathbf{M_2}\Big)\Big(2c + 2 + \mathbf{M_3} + \mathbf{M_4}\Big) l_i.$$

Therefore we have

$$\frac{1}{\tau_{i+\frac{1}{2}}} \leq C_4 h_i, \tag{2.91}$$

with $l_i \leq \frac{1}{2}\big(1 + c\big) h_i$ for $i = 0, \ldots, N$, and

$$C_4 = \frac{3(1+c)}{\underline{\sigma}^2 x_{\max}^2} \cdot \Big(2 + \mathbf{M_2}\Big)\Big(2c + 2 + \mathbf{M_3} + \mathbf{M_4}\Big).$$

Let us prove now, Proposition 1.

Proof of Proposition 1

Here we have two cases which are the following:

- $\underline{1^{st} case}$: The TPFA method is applied for the spatial discretization.

  Thereby, for $i = 0, 1, \ldots, N$ we have:

$$\begin{aligned} \Big|F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\Big| &= \Big| - \tau_{i+\frac{1}{2}}\Big(\omega(x_{i+1}) - \omega(x_i)\Big) - x_{i+\frac{1}{2}}\Big(b^+\omega(x_i) + b^-\omega(x_{i+1})\Big) \\ &\quad + k(x_{i+\frac{1}{2}})\omega'(x_{i+\frac{1}{2}}) + b x_{i+\frac{1}{2}}\omega\big(x_{i+\frac{1}{2}}\big)\Big| \\ \Big|F_h\big(\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\Big| &\leq \Big|k(x_{i+\frac{1}{2}})\Big| \cdot \Big|\omega'(x_{i+\frac{1}{2}}) - \frac{\omega(x_{i+1}) - \omega(x_i)}{h_i}\Big| \\ &\quad + x_{i+\frac{1}{2}}\Big|b\omega\big(x_{i+\frac{1}{2}}\big) - \Big(b^+\omega(x_i) + b^-\omega(x_{i+1})\Big)\Big| \\ &\quad + \Big|\frac{k(x_{i+\frac{1}{2}})}{h_i} - \tau_{i+\frac{1}{2}}\Big| \cdot \Big|\omega(x_{i+1}) - \omega(x_i)\Big|. \end{aligned}$$

Let us estimate P_1 defined as follows:

$$P_1 := k\big(x_{i+\frac{1}{2}}\big)\left(\frac{\omega(x_{i+1})-\omega(x_i)}{h_i} - \omega'\big(x_{i+\frac{1}{2}}\big)\right). \tag{2.92}$$

Indeed, using the Sobolev embedding, Theorem 3, as we are in dimension 1, $H^2(\Omega) \hookrightarrow C^1(\Omega)$,. Thereby, the Taylor expansion with integral remainder gives

$$\omega(x_{i+1}) = \omega\big(x_{i+\frac{1}{2}}\big) + \frac{h_i}{2}\omega'\big(x_{i+\frac{1}{2}}\big) + \int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \big(x_{i+1}-x\big)\omega''(x)dx. \tag{2.93}$$

Similarly,

$$\omega(x_i) = \omega\big(x_{i+\frac{1}{2}}\big) - \frac{h_i}{2}\omega'\big(x_{i+\frac{1}{2}}\big) + \int_{x_{i+\frac{1}{2}}}^{x_i} \big(x_i-x\big)\omega''(x)dx. \tag{2.94}$$

Using (2.93) and (2.94), we get

$$\frac{\omega(x_{i+1})-\omega(x_i)}{h_i} - \omega'(x_{i+\frac{1}{2}}) = \frac{1}{h_i}\int_{x_i}^{x_{i+\frac{1}{2}}} \big(x_i-x\big)\omega''(x)dx + \frac{1}{h_i}\int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \big(x_{i+1}-x\big)\omega''(x)dx. \tag{2.95}$$

Since

$$\begin{aligned}
\left|\int_{x_{i+\frac{1}{2}}}^{x_{i+1}} \big(x_{i+1}-x\big)\omega''(x)dx\right| &\leq \frac{h_i}{2}\int_{x_{i+\frac{1}{2}}}^{x_{i+1}} |\omega''|dx,\\
\left|\int_{x_i}^{x_{i+\frac{1}{2}}} \big(x_i-x\big)\omega''(x)dx\right| &\leq \frac{h_i}{2}\int_{x_i}^{x_{i+\frac{1}{2}}} |\omega''|dx.
\end{aligned}$$

Then

$$\left|\frac{\omega(x_{i+1})-\omega(x_i)}{h_i} - \omega'\big(x_{i+\frac{1}{2}}\big)\right| \leq \frac{1}{2}\int_{x_i}^{x_{i+1}} |\omega''|dx.$$

Besides, we have

$$k(x_{i+\frac{1}{2}}) = \frac{1}{2}\sigma^2 x_{i+\frac{1}{2}}^2.$$

Thus

$$|P_1| \leq \frac{\sigma^2 x_{i+\frac{1}{2}}^2}{4}\int_{x_i}^{x_{i+1}} |\omega''|dx \leq \frac{\sigma^2}{4}\int_{x_i}^{x_{i+1}} \Big(\frac{x_{i+\frac{1}{2}}}{x}\Big)^2 |x^2\omega''|dx. \tag{2.96}$$

For $x \in \Omega_i = [x_i; x_{i+1}]$, we have

$$x_i \leq x \leq x_{i+1} \Rightarrow \frac{1}{x_{i+1}} \leq \frac{1}{x} \leq \frac{1}{x_i} \Rightarrow \frac{x_{i+\frac{1}{2}}}{x_{i+1}} \leq \frac{x_{i+\frac{1}{2}}}{x} \leq \frac{x_{i+\frac{1}{2}}}{x_i}. \tag{2.97}$$

We have also, using (2.20)

$$\frac{x_{i+\frac{1}{2}}}{x_i} = \frac{x_i + x_{i+1}}{2x_i} = \frac{1}{2}\left(2 + \frac{h_i}{x_{i-1}+h_{i-1}}\right).$$

then

$$\frac{x_{i+\frac{1}{2}}}{x_i} \leq \frac{1}{2}\left(2 + \frac{h_i}{h_{i-1}}\right) \leq \frac{1}{2}\Big(2+c\Big) \quad \frac{x_{i+\frac{1}{2}}}{x_i} \leq 1 + \frac{c}{2}. \tag{2.98}$$

Coming back to (2.96), we have

$$|P_1| \leq \frac{\sigma^2}{4}\left(1+\frac{c}{2}\right)^2 \int_{x_i}^{x_{i+1}} |x^2\omega''|dx. \tag{2.99}$$

Using the continuous flux F defined in (2.75), yields

$$F'\Big((\omega(x)\Big) = -ax^2\omega''(x) - (2a+b)x\omega'(x) - b\omega(x), \tag{2.100}$$

then we have

$$a|x^2\omega''| \leq |F'\big(\omega(x)\big)| + |(2a+b)x| \cdot |\omega'(x)| + |b| \cdot |\ \omega(x)|. \tag{2.101}$$

Thus, we obtain

$$|P_1| \leq C_{14} \int_{x_i}^{x_{i+1}} \left(|F'\big(\omega(x)\big)| + |\omega'(x)| + |\omega(x)|\right)dx, \tag{2.102}$$

with

$$C_{14} = \frac{\sigma^2}{4}\left(1+\frac{c}{2}\right)^2 \max\Big(\big(\beta+\bar{r}\big)x_{\max}, \bar{r}+\beta, 1\Big).$$

Let us estimate P_2 defined by

$$P_2 := x_{i+\frac{1}{2}}\left(b\omega\big(x_{i+\frac{1}{2}}\big) - \Big(b^+\omega(x_i) + b^-\omega(x_{i+1})\Big)\right). \tag{2.103}$$

1. $\underline{1^{st}case}$: $b > 0$

$$P_2 = bx_{i+\frac{1}{2}}\Big(\omega\big(x_{i+\frac{1}{2}}\big) - \omega(x_i)\Big). \tag{2.104}$$

As in (2.93), by applying the Taylor theorem with integral remainder we have also

$$\omega(x_{i+\frac{1}{2}}) = \omega(x_i) + \int_{x_i}^{x_{i+\frac{1}{2}}} \omega'(x)dx, \tag{2.105}$$

then

$$|P_2| \leq |b|x_{\max} \int_{x_i}^{x_{i+1}} |\omega'(x)|dx, \tag{2.106}$$

2. $\underline{1^{st}case}$: $b \leq 0$:

$$P_2 = bx_{i+\frac{1}{2}}\Big(\omega\big(x_{i+\frac{1}{2}}\big) - \omega(x_{i+1})\Big). \tag{2.107}$$

As in (2.93), by applying the Taylor theorem with integral remainder we have also

$$\omega(x_{i+1}) = \omega(x_{i+\frac{1}{2}}) + \int_{x_i}^{x_{i+\frac{1}{2}}} \omega'(x)dx, \tag{2.108}$$

then

$$|P_2| \leq |b|x_{\max} \int_{x_i}^{x_{i+1}} |\omega'(x)|dx, \tag{2.109}$$

From (2.106) and (2.109), we have finally

$$|P_2| \leq \Big(\bar{r}+\beta\Big)x_{\max} \int_{x_i}^{x_{i+1}} |\omega'(x)|dx. \tag{2.110}$$

Similarly, we have

$$\omega(x_{i+1}) = \omega(x_i) + \int_{x_i}^{x_{i+1}} \omega'(x)dx. \tag{2.111}$$

then

$$|\omega(x_{i+1}) - \omega(x_i)| \leq \int_{x_i}^{x_{i+1}} |\omega'(x)|dx. \tag{2.112}$$

Let us bound

$$\Big|\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\Big|. \tag{2.113}$$

Indeed, we have

$$\Big|\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| \;\leq\; \Big|\tau_{i+\frac{1}{2}}\Big| + \Big|\frac{k(x_{i+\frac{1}{2}})}{h_i}\Big|. \tag{2.114}$$

Moreover,

$$\begin{aligned}
\Big|\frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| &= \frac{\sigma^2 x_{i+\frac{1}{2}}^2}{2h_i} = \frac{\sigma^2}{4}\frac{(x_i+x_{i+1})^2}{x_{i+1}-x_i} = \frac{\sigma^2}{4}\frac{x_{i+1}^2\Big(1+2\frac{x_i}{x_{i+1}}+\big(\frac{x_i}{x_{i+1}}\big)^2\Big)}{x_{i+1}\Big(1-\frac{x_i}{x_{i+1}}\Big)} \qquad (2.115)\\
&= \frac{\sigma^2}{4}x_{i+1}\times\frac{1+2Z_i+Z_i^2}{1-Z_i}\\
\Big|\frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| &\leq \frac{\sigma^2}{4}x_{\max}\times\Big(1+2Z_i+Z_i^2\Big)\times\frac{1}{1-Z_i}. \qquad (2.116)
\end{aligned}$$

with Z_i defined as follows

$$Z_i = \frac{x_i}{x_{i+1}}, \qquad\qquad 0 < Z_i < 1. \tag{2.117}$$

Using a similar argument as in (2.81) and (2.82), there exists a positive constant $\mathbf{M_4}$ such that

$$\Big|\frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| \leq \sigma^2 x_{\max}\times\big(2+\mathbf{M_4}\big). \tag{2.118}$$

On the other hand, using (2.83) we have,

$$\Big|\tau_{i+\frac{1}{2}}\Big| \leq \frac{\sigma^2}{2}\big(2+\mathbf{M}_1\big)x_{\max}. \tag{2.119}$$

Coming back to (2.114), and by using (2.118) and (2.83) we have

$$\begin{aligned}
\Big|\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| &\leq \Big|\tau_{i+\frac{1}{2}}\Big| + \Big|\frac{k(x_{i+\frac{1}{2}})}{h_i}\Big|\\
&\leq \frac{\sigma^2}{2}\big(2+\mathbf{M}_1\big)x_{\max} + \sigma^2 x_{\max}\Big(2+\mathbf{M_4}\Big)\\
\Big|\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\Big| &\leq \sigma^2 x_{\max}\Big(\mathbf{M_1}+\mathbf{M_4}+\frac{3}{2}\Big). \qquad (2.120)
\end{aligned}$$

Then the estimate of P_3 defined as follows

$$P_3 := \left(\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\right)\Big(\omega_{i+1}-\omega_i\Big), \tag{2.121}$$

and using (2.112) and (2.120). We therefore have

$$\begin{aligned}
|P_3| &\leq \left|\tau_{i+\frac{1}{2}} - \frac{k(x_{i+\frac{1}{2}})}{h_i}\right|\Big|\omega_{i+1}-\omega_i\Big|\\
&\leq \sigma^2 x_{\max}\Big(\mathbf{M_1}+\mathbf{M_4}+\frac{3}{2}\Big)\int_{x_i}^{x_{i+1}}|\omega'(x)|dx\\
|P_3| &\leq \sigma^2 x_{\max}\Big(\mathbf{M_1}+\mathbf{M_4}+\frac{3}{2}\Big)\int_{x_i}^{x_{i+1}}|\omega'(x)|dx. \qquad (2.122)
\end{aligned}$$

Using (2.96),(2.110) and (2.122) lead to

$$\begin{aligned}\Big|F_h\big(w(x_{i+\frac{1}{2}})\big)-F\big(w(x_{i+\frac{1}{2}})\big)\Big| &\leq |P_1|+|P_2|+|P_3|\\ &\leq C_{14}\int_{x_i}^{x_{i+1}}\Big(|F'(\omega)|+|\omega'|+|\omega|\Big)dx\\ &\quad +|r-\sigma^2|x_{\max}\int_{x_i}^{x_{i+1}}|\omega'(x)|dx\\ &\quad +\sigma^2 x_{\max}\Big(\mathbf{M_1}+\mathbf{M_4}+\frac{3}{2}\Big)\int_{x_i}^{x_{i+1}}|\omega'(x)|dx\\ \Big|F_h\big(w(x_{i+\frac{1}{2}})\big)-F\big(w(x_{i+\frac{1}{2}})\big)\Big| &\leq C_1\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)dx, \qquad (2.123)\end{aligned}$$

with

$$C_1 = C_{14}+\big(\bar{r}+\beta\big)x_{\max}+\beta x_{\max}\Big(\mathbf{M_1}+\mathbf{M_4}+\frac{3}{2}\Big),$$

- $\underline{2^{nd}\textit{case}:}$ The fitted TPFA method is applied for space discretization at $x_{1/2}$ we have:

$$\begin{aligned}\Big|G_h(\omega\big(x_{1/2}\big)-F(\omega\big(x_{1/2}\big)\Big| &= \Big|-\frac{1}{4}x_1(a+b)\omega(x_1)+k(x_{\frac{1}{2}})\omega'\big(x_{\frac{1}{2}}\big)+bx_{\frac{1}{2}}\omega\big(x_{\frac{1}{2}}\big)\Big|\\ &= \Big|-\frac{1}{4}x_1(a+b)\Big(\omega(x_1)-\omega(x_0)\Big)+k(x_{\frac{1}{2}})\omega'\big(x_{\frac{1}{2}}\big)+bx_{\frac{1}{2}}\omega\big(x_{\frac{1}{2}}\big)\Big|\\ &= \Big|k(x_{\frac{1}{2}})\Big(\omega'\big(x_{\frac{1}{2}}\big)-\frac{\omega(x_1)-\omega(x_0)}{h_0}\Big)+bx_{\frac{1}{2}}\Big(\omega(x_{1/2})-\omega(x_0)\Big)\\ &\quad +\Big(k(x_{1/2})-\frac{1}{4}x_1(a+b)\Big)\Big(\omega(x_1)-\omega(x_0)\Big)\Big|\\ \Big|G_h(\omega\big(x_{1/2}\big)-F(\omega\big(x_{1/2}\big)\Big| &\leq \big|k(x_{\frac{1}{2}})\big|\cdot\Big|\omega'\big(x_{\frac{1}{2}}\big)-\frac{\omega(x_1)-\omega(x_0)}{h_0}\Big|+|b|x_{\frac{1}{2}}\Big|\omega(x_{1/2})-\omega(x_0)\Big|\\ &\quad \Big|k(x_{1/2})-\frac{1}{4}x_1(a+b)\Big|\Big|\omega(x_1)-\omega(x_0)\Big|.\end{aligned}$$

It follows that:

$$\begin{aligned}\Big|k(x_{1/2})-\frac{1}{4}x_1(a+b)\Big| &\leq \Big|k(x_{1/2})\Big|+\Big|\frac{1}{4}x_1(a+b)\Big|\\ &\leq \Big|k(x_{1/2})\Big|+\frac{1}{4}x_{\max}\big(\bar{r}+\beta\big).\end{aligned}$$

Thereby, using (2.118), 2.92 and 2.102, (2.104) and 2.106, we get

$$\Big|G_h(\omega\big(x_{1/2}\big)-F(\omega\big(x_{1/2}\big)\Big|\leq C_2\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)dx$$

where

$$C_2 = C_{14}+\Bigg[\big(\bar{r}+\beta\big)+\Big(\beta\Big(\mathbf{M_1}+\frac{5}{2}\Big)+\frac{1}{4}\bar{r}\Big)\Bigg]x_{\max}.$$

Besides, since for $i=1,\ldots,N$,

$$G_h\big(w(x_{i+\frac{1}{2}})\big) = F_h(w(x_{i+\frac{1}{2}})\big), \qquad (2.124)$$

similarly to 2.123, we get

$$\Big|F\big(w(x_{i+\frac{1}{2}})\big)-G_h\big(I_h w(x_{i+\frac{1}{2}})\big)\Big| \leq C_2\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)dx,$$

with

$$C_2 = C_{14} + \big(\bar{r} + \beta\big)x_{\max} + \beta x_{\max}\Big(\mathbf{M_1} + \mathbf{M_4} + \frac{3}{2}\Big).$$

Finally, when the fitted TPFA method is applied for the space discretization, there exists a positive constant satisfying (2.77).

2.4 Full discretization and errors estimates

Let $0 := t_0 < t_1 < \ldots < t_{M-1} < t_M := T$ be a subdivision of the time interval $[0, T]$ with the step sizes $\Delta t_m = t_{m+1} - t_m$, $m \in \{0, \ldots, M-1\}$ and $\Delta t = \max_{1 \le m \le M-1} \Delta t_m$. The full discretization of (2.14) using the combination of the TPFA method with the parameter $\theta \in [0, 1]$ can be formulated as follows:
Find a sequence $u_h^1, \ldots, u_h^M \in V_h$ such that for $m \in \{0, \ldots, M-1\}$

$$\begin{cases} \left(\dfrac{u_h^{m+1} - u_h^m}{\Delta t_m}, v_h\right)_h + a_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big) &= \left(\theta f^{m+1} + (1-\theta)f^m, v_h\right)_h, \\ u^0 &= u_{oh}. \end{cases} \tag{2.125}$$

where $t_{m+\theta} = \theta t_{m+1} + (1-\theta)t_m$ and the bilinear form a_h is given by (2.43). Similarly, when the fitted TPFA method is applied for the the spatial discretization, the full discretization is formulated as follows:
Find a sequence $u_h^1, \ldots, u_h^M \in V_h$ such that for $m \in \{0, \ldots, M-1\}$

$$\begin{cases} \left(\dfrac{u_h^{m+1} - u_h^m}{\Delta t_m}, v_h\right)_h + b_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big), &= \left(\theta f^{m+1} + (1-\theta)f^m, v_h\right)_h, \\ u^0 &= u_{oh}. \end{cases} \tag{2.126}$$

where the bilinear form b_h is given by (2.51).

2.4.1 Errors estimates

Theorem 8 *Let us consider the unique solution u of (2.17) and ζ_h^m the numerical solution of the fully discretized scheme using the TPFA method (2.37)($\zeta_h^m = u_h^m$ for the TPFA method) or the fitted TPFA method (2.45) ($\zeta_h^m = z_h^m$ for fitted TPFA method). Let $\theta \in [1/2; 1]$, if $u \in H^1\Big(0, T; H^1(\Omega)\Big) \cap H^2\Big(0, T; L^2(\Omega)\Big)$ and $F(u) \in C(0, T, H^1(\Omega))$, then there exists a positive constant C, independent of h, Δt, M, and N such that*

$$\big|\big|u(t_m) - \zeta_h^m\big|\big|_{0,h} \le C(h + \Delta t). \tag{2.127}$$

Proof of Theorem 7

Indeed, the proofs follow the same lines as that in [Angermann and Wang, 2007, Theorem 7]. We summarise the keys steps. Here we have two cases.

<u>1^{st} case</u> When the TPFA method is applied for the space discretization.
Here, we take $\zeta_h^m = u_h^m$. Let us notice that

$$\big|\big|u(t_m) - u_h^m\big|\big|_{0,h} \le \big|\big|u(t_m) - I_h u(t_m)\big|\big|_{0,h} + \big|\big|I_h u(t_m) - u_h^m\big|\big|_{0,h}, \tag{2.128}$$

where I_h is the interpolation operator introduced in (2.75).

In order to bound the first term on the right hand side of (2.128), let us recall the following result. Since $u(t) \in H^2(\Omega)$ then there exists a constant C_{31} depending on u (see Theorem 4 or Theorem 3.25, page 138 in P. Knabner [2002]) such that

$$\big|\big|I_h u(t) - u(t)\big|\big|_{0,h} \leq C_{31} \cdot h^2 \cdot |u(t)|_2, \tag{2.129}$$

where $|\cdot|_2$ is the semi-norm of $H^2(\Omega)$. Furthermore, for $u \in C(\big(0,T\big), H^2(\Omega))$, there exists a positive constant $C_{32} = C_{31}(u,T) \cdot x_{\max}$ such that

$$\big|\big|I_h u(t_m) - u(t_m)\big|\big|_{0,h} \leq C_{32} \cdot h. \tag{2.130}$$

We now estimate $W^m := I_h u(t_m) - u_h^m$ in the discrete L^2-norm. We define the expression A by

$$A = \left(\frac{W^{m+1} - W^m}{\Delta t_m}, v_h\right)_h + a_h\Big(\theta W^{m+1} + (1-\theta)W^m, v_h; t_{m+\theta}\Big), \tag{2.131}$$

where a_h is the bilinear form given by (2.40) when the TPFA method is applied. By some arithmetic manipulations, we have

$$\begin{aligned}
A &= \sum_{i=1}^{N} l_i \frac{W_i^{m+1} - W_i^m}{\Delta t_m} v_i + a_h\Big(\theta W^{m+1} + (1-\theta)W^m, v_h; t_{m+\theta}\Big) \\
&= \sum_{i=1}^{N} l_i \frac{I_h u_i(t_{m+1}) - I_h u_i(t_m)}{\Delta t_m} v_i + a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h; t_{m+\theta}\Big) \\
&\quad - \sum_{i=1}^{N} l_i \frac{{u_h}_i^{m+1} - {u_h}_i^m}{\Delta t_m} v_i - a_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big) \\
&= \sum_{i=1}^{N} l_i \frac{I_h u_i(t_{m+1}) - I_h u_i(t_m)}{\Delta t_m} v_i + a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h; t_{m+\theta}\Big) \\
&\quad - \sum_{i=1}^{N} l_i \frac{{u_h}_i^{m+1} - {u_h}_i^m}{\Delta t_m} v_i - a_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big) \\
&\quad - \Big(\theta \dot{u}(t_{m+1}) + \big(1-\theta\big)\dot{u}(t_m), L_h v_h\Big) + \Big(\theta \dot{u}(t_{m+1}) + \big(1-\theta\big)\dot{u}(t_m), L_h v_h\Big) \\
&\quad + \theta \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+1}\big) + \big(1-\theta\big)\hat{a}_h\big(u(t_m), v_h; t_m\big) - \theta \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+1}\big) \\
&\quad - \big(1-\theta\big)\hat{a}_h\big(u(t_m), v_h; t_m\big).
\end{aligned} \tag{2.132}$$

We also have

$$\begin{aligned}
A &= \left[\sum_{i=1}^{N} l_i \frac{I_h u_i(t_{m+1}) - I_h u_i(t_m)}{\Delta t_m} v_i - \Big(\theta \dot{u}(t_{m+1}) + \big(1-\theta\big)\dot{u}(t_m), L_h v_h\Big)\right] \\
&\quad + \Bigg[a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h; t_{m+\theta}\Big) - \theta \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+1}\big) \\
&\quad - \big(1-\theta\big)\hat{a}_h\big(u(t_m), v_h; t_m\big)\Bigg] + \Bigg[\Big(\theta \dot{u}(t_{m+1}) + \big(1-\theta\big)\dot{u}(t_m), L_h v_h\Big) + \theta \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+1}\big) \\
&\quad + \big(1-\theta\big)\hat{a}_h\big(u(t_m), v_h; t_m\big)\Bigg] - \sum_{i=1}^{N} l_i \frac{{u_h}_i^{m+1} - {u_h}_i^m}{\Delta t_m} v_i - a_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big).
\end{aligned} \tag{2.133}$$

Remember that (see (2.39))

$$\sum_{i=1}^{N} l_i \frac{{u_h}_i^{m+1} - {u_h}_i^{m}}{\Delta t_m} v_i + a_h\Big(\theta u_h^{m+1} + (1-\theta)u_h^m, v_h; t_{m+\theta}\Big) = \Big(\theta f^{m+1} + (1-\theta) f^m, v_h\Big)_h, \quad (2.134)$$

and also

$$\begin{aligned}\Big(\theta \dot{u}(t_{m+1}) + \big(1-\theta\big)\dot{u}(t_m), L_h v_h\Big) + \theta \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+1}\big) \quad + \\ \big(1-\theta\big)\hat{a}_h\big(u(t_m), v_h; t_m\big) \quad = \quad \Big(\theta f^{m+1} + (1-\theta) f^m, L_h v\Big).\end{aligned} \quad (2.135)$$

Thereby, we get

$$A = Y_1^m + Y_2^m + Y_m^3, \quad (2.136)$$

where

$$Y_1^m = \sum_{i=1}^{N} l_i \frac{I_h u_i(t_{m+1}) - I_h u_i(t_m)}{\Delta t_m} v_i - \Big(\theta \dot{u}(t_{m+1}) + (1-\theta)\dot{u}(t_m), v_h\Big), \quad (2.137)$$

$$\begin{aligned}Y_2^m = a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h; t_{m+\theta}\Big) - \theta \hat{a}_h\Big(u(t_{m+1}), v_h; t_{m+1}\Big) \\ -(1-\theta)\hat{a}_h\Big(u(t_m), v_h; t_m\Big)\end{aligned} \quad (2.138)$$

and

$$Y_3^m = \Big(\theta f^{m+1} + (1-\theta) f^m, L_h v\Big) - \Big(\theta f^{m+1} + (1-\theta) f^m, v_h\Big)_h. \quad (2.139)$$

The estimation of Y_1^m is done exactly as in [Angermann and Wang, 2007, (54)] and we get

$$Y_1^m = \Big(\omega^m, L_h v_h\Big) \quad (2.140)$$

$$\omega^m := \frac{L_h u(t_{m+1}) - L_h u(t_m)}{\Delta t_m} - \theta \dot{u}(t_{m+1}) - (1-\theta)\dot{u}(t_m) \quad (2.141)$$

$$|Y_1^m| \le ||\omega^m||_{L^2(\Omega)} ||v_h||_{0,h} \quad (2.142)$$

with

$$||\omega^m||_{L^2(\Omega)} \le \mathcal{Q}_1^m\big(\Delta t_m, h\big) := \frac{1}{\Delta t_m} \int_{t_m}^{t_{m+1}} \big\|(L_h - I) \circ \dot{u}(s)\big\|_{L^2(\Omega)} ds + \int_{t_m}^{t_{m+1}} ||\ddot{u}||_{L^2(\Omega)} ds.$$

Let us estimate of Y_2^m.

$$Y_2^m = a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h\Big) - \theta \hat{a}_h\Big(u(t_{m+1}), v_h\Big) - (1-\theta)\hat{a}_h\Big(u(t_m), v_h\Big). \quad (2.143)$$

By adding and extracting the term $\hat{a}_h\Big(\theta u(t_{m+1}) + (1-\theta)u(t_m), v_h, t_{m+\theta}\Big)$ we get

$$\begin{aligned}Y_2^m \quad = \quad & a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h, t_{m+\theta}\Big) - \hat{a}_h\Big(\theta u(t_{m+1}) + (1-\theta)u(t_m), v_h, t_{m+\theta}\Big) \\ & + \hat{a}_h\Big(\theta u(t_{m+1}) + (1-\theta)u(t_m), v_h, t_{m+\theta}\Big) - \theta \hat{a}_h\Big(u(t_{m+1}), v_h, t_{m+1}\Big) \\ & -(1-\theta)\hat{a}_h\Big(u(t_m), v_h, t_m\Big) \\ =: \quad & Y_{21}^m + Y_{22}^m,\end{aligned} \quad (2.144)$$

where

$$Y_{21}^m = a_h\Big(\theta I_h u(t_{m+1}) + (1-\theta) I_h u(t_m), v_h; t_{m+\theta}\Big) - \hat{a}_h\Big(\theta u(t_{m+1}) + (1-\theta)u(t_m), v_h; t_{m+\theta}\Big), \quad (2.145)$$

and

$$Y_{22}^m = \hat{a}_h\Big(\theta u(t_{m+1}) + (1-\theta)u(t_m), v_h; t_{m+\theta}\Big) - \theta\hat{a}_h\Big(u(t_{m+1}), v_h; t_{m+1}\Big) - (1-\theta)\hat{a}_h\Big(u(t_m), v_h; t_m\Big), \quad (2.146)$$

with $t_{m+\theta} = \theta t_{m+1} + (1-\theta)t_m \qquad \theta \in [1/2, 1]$.

Note that

$$\begin{aligned} Y_{21}^m &= \theta\Big(a_h\big(I_h u(t_{m+1}), v_h; t_{m+\theta}\big) - \hat{a}_h\big(u(t_{m+1}), v_h; t_{m+\theta}\big)\Big) \\ &\quad +(1-\theta)\Big(a_h\big(I_h u(t_m), v_h; t_{m+\theta}\big) - \hat{a}_h\big(u(t_m), v_h; t_{m+\theta}\big)\Big). \end{aligned}$$

Let us consider the term

$$\delta_{21}(\omega, v_h, s) := a_h(I_h\omega, v_h; s) - \hat{a}_h(\omega, v_h; s). \quad (2.147)$$

Thereby, using (2.43) we have:

$$\begin{aligned} \delta_{21}(\omega, v_h, s) &= a_h(I_h\omega, v_h; s) - \hat{a}_h(\omega, v_h; s) \\ &= \sum_{i=1}^{N}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F_h\big(I_h\omega(x_{i-\frac{1}{2}})\big)\right)\cdot v_i + \Big(c(s)I_h\omega, v_h\Big)_h \\ &\quad -\sum_{i=1}^{N}\left(F\big(\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i-\frac{1}{2}})\big)\right)\cdot v_i - \Big(c(s)\omega, L_h v_h\Big) \\ &= \sum_{i=1}^{N}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)v_i - \sum_{i=1}^{N}\Bigg(F_h\big(I_h\omega(x_{i-\frac{1}{2}})\big) \\ &\quad -F\big(\omega(x_{i-\frac{1}{2}})\big)\Bigg)\cdot v_i + c(s)\left(\Big(I_h\omega, v_h\Big)_h - \Big(\omega, L_h v_h\Big)\right) \\ &= \sum_{i=1}^{N}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)\cdot v_i - \sum_{i=0}^{N-1}\Bigg(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) \\ &\quad -F\big(\omega(x_{i+\frac{1}{2}})\big)\Bigg)\cdot v_{i+1} + c(s)\left(\Big(L_h I_h\omega, L_h v_h\Big) - \Big(\omega, L_h v_h\Big)\right). \end{aligned} \quad (2.148)$$

Furthermore, rearranging the summation in (2.148) gives

$$\begin{aligned} \delta_{21}(\omega, v_h, s) &= -\left(F_h\big(I_h\omega(x_{\frac{1}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\right)v_1 + \sum_{i=1}^{N-1}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right) \\ &\quad \times\big(v_i - v_{i+1}\big) + \Big(F_h\big(I_h\omega(x_{N+\frac{1}{2}})\big) - F\big(\omega(x_{N+\frac{1}{2}})\big)\Big)\cdot v_N \\ &\quad +c(s)\left(\Big(L_h\omega, L_h v_h\Big) - \Big(\omega, L_h v_h\Big)\right) \\ &= \left(F_h\big(I_h\omega(x_{\frac{1}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\right)\cdot\Big(v_0 - v_1\Big) \\ &\quad + \sum_{i=1}^{N-1}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)\cdot\big(v_i - v_{i+1}\big) \\ &\quad +\Big(F_h\big(I_h\omega(x_{N+\frac{1}{2}})\big) - F\big(\omega(x_{N+\frac{1}{2}})\big)\Big)\cdot\Big(v_N - v_{N+1}\Big) + c(s)\Big(L_h\omega - w, L_h v_h\Big) \end{aligned}$$

$$\begin{aligned}\delta_{21}(\omega, v_h, s) &= \sum_{i=0}^{N}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)\cdot\big(v_i - v_{i+1}\big) + c(s)\Big(\big(L_h - I\big)\omega, L_h v_h\Big),\\ &= \delta_{211} + \delta_{212},\end{aligned} \tag{2.149}$$

where I is the identity operator and δ_{211} defined as follows:

$$\delta_{211} = \sum_{i=0}^{N}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)\cdot\big(v_i - v_{i+1}\big) \tag{2.150}$$

Moreover, we also have

$$\begin{aligned}|F_h\big(I_h\omega(x_{\frac{3}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\big(v_0 - v_1\big)| &= |\frac{1}{\tau_{1/2}}\left(F_h\big(I_h\omega(x_{\frac{3}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\right)\tau_{1/2}\big(v_0 - v_1\big)|\\ &\leq \frac{1}{\tau_{1/2}}\left(F_h\big(I_h\omega(x_{\frac{3}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\right)\tau_{1/2}|v_1|\\ &\leq C_4 h_0\int_{x_0}^{x_1}\Big(|F'(\omega)| + |\omega'| + |\omega|\Big)dx \times C_3|v_1|\\ &\leq C_{15}h_0\left[\int_{x_0}^{x_1}\Big(|F'(\omega)| + |\omega'| + |\omega|\Big)^2 dx\right]^{1/2}\sqrt{h_0 v_1^2}\\ |F_h\big(I_h\omega(x_{\frac{3}{2}})\big) - F\big(\omega(x_{\frac{1}{2}})\big)\big(v_0 - v_1\big)| &\leq C_{15}h_0\left[\int_{x_0}^{x_1}\Big(|F'(\omega)| + |\omega'| + |\omega|\Big)^2 dx\right]^{1/2}\sqrt{h_0 v_1^2},\end{aligned}$$

where the positive constant C_2 and C_3 are given in (2.78) and $C_{15} = \max\Big(C_3, C_4\Big)$. Let us denote by Q_i the following quantity

$$Q_i = \left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i+\frac{1}{2}})\big)\right)\Big(v_i - v_{i+1}\Big). \tag{2.151}$$

Besides, we have

$$\sum_{i=1}^{N} Q_i = \sum_{i=1}^{N}\frac{1}{\sqrt{\tau_{i+\frac{1}{2}}}}\left(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big) - F\big(\omega(x_{i-\frac{1}{2}})\big)\right)\times\sqrt{\tau_{i+\frac{1}{2}}}\Big(v_i - v_{i+1}\Big) \tag{2.152}$$

Thereby, using (2.76) and (2.78) in (2.152) leads to

$$\begin{aligned}\sum_{i=1}^{N}|Q_i| &\leq \sum_{i=1}^{N}\left(\sqrt{C_2 h_i}\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)dx \times \sqrt{\tau_{i+\frac{1}{2}}}|\Big(v_i - v_{i+1}\Big)|\right)\\ &\leq \sqrt{C_4}\sum_{i=1}^{N}\left(h_i\left[\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)^2 dx\right]^{1/2}\times\sqrt{\tau_{i+\frac{1}{2}}}|\Big(v_i - v_{i+1}\Big)|\right)\\ &\leq \sqrt{C_4}\sum_{i=1}^{N}\left(\left[\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)^2 dx\right]^{1/2}\times\sqrt{\tau_{i+\frac{1}{2}}}|\Big(v_i - v_{i+1}\Big)|\right)h.\end{aligned}$$

Furthermore,

$$\begin{aligned}\sum_{i=1}^{N}|Q_i| &\leq \sqrt{C_4}\left(\sum_{i=1}^{N}\int_{x_i}^{x_{i+1}}\Big(\big|F'(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)^2 dx\right)^{1/2}\times\left(\sum_{i=1}^{N}\tau_{i+\frac{1}{2}}\Big(v_i - v_{i+1}\Big)^2\right)^{1/2}h\\ \sum_{i=1}^{N}|Q_i| &\leq h\sqrt{C_4}\left[\int_{x_1}^{x_{N+1}}\Big(\big|F(\omega)\big| + \big|\omega'\big| + \big|\omega\big|\Big)^2 dx\right]^{1/2}||v_h||_{0,\omega}.\end{aligned} \tag{2.153}$$

Coming back to δ_{11} defined in (2.150), and using the equations (2.151), (2.153), and also the fact that $h_0 \leq cl_1$ from Assumption (4), we get:

$$\begin{aligned}
\delta_{211} &\leq C_{16}\Bigg(h_0\Bigg[\int_{x_0}^{x_1}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)^2dx\Bigg]^{1/2}\sqrt{h_0v_1^2} \qquad (2.154)\\
&\quad+\Bigg[\int_{x_1}^{x_{N+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)^2dx\Bigg]^{1/2}h\big|\big|v_h\big|\big|_{0,\omega}\Bigg)\\
&\leq C_{17}h\Bigg[\int_{x_0}^{x_{N+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)^2dx\Bigg]^{1/2}\Big(\sqrt{l_1v_1^2}+\big|\big|v_h\big|\big|_{0,\omega}\Big)\\
&\leq C_{18}h\Bigg[\int_{x_0}^{x_{N+1}}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)^2dx\Bigg]^{1/2}\Big(l_1v_1^2+\big|\big|v_h\big|\big|_{0,\omega}^2\Big)^{1/2}\\
&\leq C_{18}\Bigg[\int_{\Omega}\Big(\big|F'(\omega)\big|+\big|\omega'\big|+\big|\omega\big|\Big)^2dx\Bigg]^{1/2}\cdot h\cdot\big|\big|v_h\big|\big|_{\omega,d}\\
&\leq C_{211}\Big(\big|F(\omega)\big|_1+\big|\big|\omega\big|\big|_1\Big)\cdot h\cdot\big|\big|v_h\big|\big|_{\omega,d}. \qquad (2.155)
\end{aligned}$$

Note that $\|.\|_1$ and $|.|_1$ are respectively the $H^1(\Omega)$ norm and semi-norm.

For the second term of (2.149), δ_{212} is estimated as in Angermann and Wang [2007] as follows

$$\Big|c(s)\Big(\big(L_h-I\big)\omega,L_hv_h\Big)\Big| \leq C_{212}\cdot h\cdot\big|\big|\omega\big|\big|_{1,\omega}\cdot\big|\big|v_h\big|\big|_{0,h}.$$

Thus

$$\begin{aligned}
|\delta_{21}(\omega,v_h,s)| &\leq \Bigg|\sum_{i=0}^{N}\Big(F_h\big(I_h\omega(x_{i+\frac{1}{2}})\big)-F\big(\omega(x_{i+\frac{1}{2}})\big)\Big)\big(v_i-v_{i+1}\big)\Bigg|\\
&\quad+\Big|c(s)\Big(\big(L_h-I\big)\omega,L_hv_h\Big)\Big|\\
&\leq C_{211}\Big(\big|F(\omega)\big|_1+\big|\big|\omega\big|\big|_1\Big)\cdot h\cdot\big|\big|v_h\big|\big|_{\omega,d}+C_{212}\cdot h\cdot\big|\big|\omega\big|\big|_{1,\omega}\cdot\big|\big|v_h\big|\big|_{0,h} \qquad (2.156)\\
&\leq C_{211}\Big(\big|F(\omega)\big|_1+\big|\big|\omega\big|\big|_{1,\omega}\Big)\cdot h\cdot||v_h||_{\omega,d}+C_{212}\cdot h\cdot\big|\big|\omega\big|\big|_{1,\omega}\cdot\big|\big|v_h\big|\big|_{\omega,d}\\
|\delta_{21}(\omega,v_h,s)| &\leq C_{21}\Big(\big|F(\omega)\big|_1+\big|\big|\omega\big|\big|_{1,\omega}\Big)\cdot h\cdot||v_h||_{\omega,h}. \qquad (2.157)
\end{aligned}$$

Using (2.147), we have

$$\begin{aligned}
|Y_{21}^m| &\leq \theta\Big|\delta_{21}\Big(u(t_{m+1}),v_h,t_{m+\theta}\Big)\Big|+(1-\theta)\Big|\delta_{21}\Big(u(t_m),v_h,t_{m+\theta}\Big)\Big|\\
&\leq C_{21}\Big(\big|F(\omega)\big|_1+\big|\big|\omega\big|\big|_{1,\omega}\Big)\cdot h\cdot||v_h||_{\omega,h}. \qquad (2.158)
\end{aligned}$$

We estimate Y_{22}^m as is done in Angermann and Wang [2007] and we get

$$|Y_{22}^m|\leq C_{22}\Delta t_m||v_h||_{0,h}. \qquad (2.159)$$

Estimate of Y_3^m is done as in [Angermann and Wang, 2007, Y_4^m] and we have

$$\big|Y_3^m\big| \leq C_3 h \big|\big|v_h\big|\big|_{0,h}. \tag{2.160}$$

Coming back to the equation above

$$R = \sum_{i=1}^{N} l_i \frac{W_i^{m+1} - W_i^m}{\Delta t_m} v_i + a_h\Big(\theta W^{m+1} + (1-\theta)W^m, v_h; t_{m+\theta}\Big) = Y_1^m + Y_2^m + Y_3^m.$$

Using (2.142),(2.158),(2.159) and (2.160) we get

$$\begin{aligned} R &\leq \left(\frac{1}{\Delta t_m}\int_{t_m}^{t_{m+1}} ||\big(L_h - I\big)\dot{u}(s)||_{L^2(\Omega)} ds + \int_{t_m}^{t_{m+1}} ||\ddot{u}||_{L^2(\Omega)} ds\right) ||v_h||_{0,h} \\ &\quad + C_{21}\cdot h \cdot ||v_h||_{1,\omega} + C_{22}(\Delta t_m)||v_h||_{0,h} + C_3 h\big|\big|v_h\big|\big|_{0,h}. \end{aligned} \tag{2.161}$$

Replacing v_h by $W_h^\theta = \theta W^{m+1} + (1-\theta)W^m$ in (2.161),and, as in Angermann and Wang [2007], using the coercivity property of a_h, we obtain

$$\frac{1}{2\Delta t_m}\left[||W^{m+1}||_{0,h}^2 - ||W^m||_{0,h}^2\right] + \alpha||W_h^\theta||_{\omega,d}^2 \leq \mathcal{Q}^m(\Delta t_m, h)||W_h^\theta||_{\omega,d}, \tag{2.162}$$

with

$$\mathcal{Q}^m(\Delta t_m, h) \leq \mathcal{Q}_1^m(\Delta t_m, h) + C'\cdot\Big(h + \Delta t_m\Big),$$

and

$$\mathcal{Q}_1^m(\Delta t_m, h) = \frac{1}{\Delta t_m}\int_{t_m}^{t_{m+1}} ||\big(L_h - I\big)\dot{u}(s)||_{L^2(\Omega)} ds + \int_{t_m}^{t_{m+1}} ||\ddot{u}||_{L^2(\Omega)} ds.$$

Following Angermann and Wang [2007], it gives

$$\begin{aligned} \sum_{k=0}^{m-1} \Delta t_k\Big[\mathcal{Q}_1^k(\Delta t_k, h)\Big] &\leq 2\sum_{k=0}^{m-1}\left[\int_{t_{k+1}}^{t_k} ||(L_h - I)\dot{u}(s)||_{L^2(\Omega)}^2 ds + (\Delta t_k)^2\int_{t_{k+1}}^{t_k} ||\ddot{u}(s)||_{L^2(\Omega)}^2 ds\right] \\ &\leq 2\int_0^T ||(L_h - I)\dot{u}(s)||_{L^2(\Omega)}^2 ds + (\Delta t)^2\int_0^T ||\ddot{u}(s)||_{L^2(\Omega)}^2 ds, \end{aligned} \tag{2.163}$$

where $\Delta t := \max_{m=0,\ldots,M-1}|\Delta t_m|$. As in Angermann and Wang [2007], we have

$$||(L_h - I)\dot{u}(s)||_{L^2(\Omega)} \leq h|\dot{u}(s)|_1.$$

Then we get

$$\sum_{k=0}^{m-1} \Delta t_k\Big[\mathcal{Q}_1^k\big(\Delta t_k, h\big)\Big] \leq 2\left[h^2||\dot{u}||_{L^2\left(0,T;H^1(\Omega)\right)}^2 + (\Delta t)^2||\ddot{u}||_{L^2\left(0,T;L^2(\Omega)\right)}^2\right]. \tag{2.164}$$

Following [Angermann and Wang, 2007, (66)], (2.164) yields

$$||W^m||_{0,h}^2 \leq ||W^0||_{0,h}^2 + C'\Big(h^2 + (\Delta t)^2\Big). \tag{2.165}$$

By taking $u^0 = I_h u_0$, we have $||W^0||_{0,h} = 0$ and it leads to

$$||W^m||_{0,h} \leq C(h + \Delta t), \tag{2.166}$$

which is actually

$$||I_h u(t_m) - u_h^m||_{0,h} \leq C(h + \Delta t). \tag{2.167}$$

Therefore, using (2.130) and (2.167), we get

$$\begin{aligned} \big|\big|u(t_m) - u_h^m\big|\big|_{0,h} &\leq \big|\big|u(t_m) - I_h u(t_m)\big|\big|_{0,h} + \big|\big|I_h u(t_m) - u_h^m\big|\big|_{0,h} \\ &\leq C_{32}\cdot h + C'(h + \Delta t) \\ \big|\big|u(t_m) - u_h^m\big|\big|_{0,h} &\leq C(h + \Delta t). \end{aligned} \tag{2.168}$$

<u>2^{nd}*case:*</u> The proof for fitted TPFA method is done exactly in the same way.

2.5 Numerical experiments

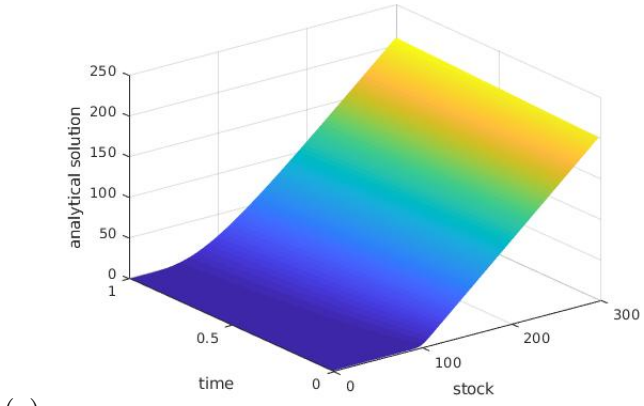
In this Section, we perform numerical experiments for an European call option pricing problem. The error are computed with respect to the following analytical solution of the Black-Scholes PDE (see Haug [2007]):

$$C(x,t) = xN(d_1) - Ke^{-rt}N(d_2), \tag{2.169}$$

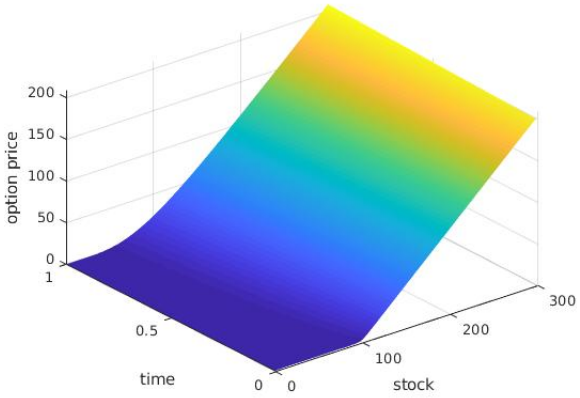
where

$$d_1 = \frac{\ln(\frac{x}{K}) + \left(r + \frac{\sigma^2}{2}\right)}{t}, \qquad d_2 = d_1 - \sigma\sqrt{t}. \tag{2.170}$$

with t the time to maturity and $N(\cdot)$ the standard cumulative normal distribution function. The computational domain is $\Omega = [0, x_{\max}] \times (0, T]$ with $x_{\max} = 300$ and the maturity time $T = 1$. These numerical experiments are performed using the risk free interest rate $r = 0.1$, the volatilty $\sigma = 0.5$ and the strike price $K = 100$.



(a)



(b)

Figure 2.2: Analytical solution in (a) and numerical solution from fitted TPFA in (b) at maturity time $T = 1$

Here we have performed space errors by fixing the time step at $dt = 1/100$ and vary the space step h.

Table 2.1: Table of space errors. The time step is fixed to be $dt = 1/100$.

Num meth \ Nb Grids pts	100	150	200	250	300	350	400	450	500
TPFA	0.0104	0.0069	0.0052	0.0042	0.0035	0.003	0.0026	0.0023	0.0021
Fitted TPFA	0.0103	0.0069	0.0052	0.0041	0.0034	0.0029	0.0026	0.0023	0.0021

For the time error, we fix the space step at $h = 0.25$, and vary the time step dt.

Table 2.2: Table of time errors. The space step is fixed to be $h = 0.25$.

Num meth \ Nb Time subdivisions	100	150	200	250	300	350	400
TPFA	$8.98.10^{-4}$	$8.83.10^{-4}$	$8.75.10^{-4}$	$8.71.10^{-4}$	$8.69.10^{-4}$	$8.66.10^{-4}$	$8.656.10^{-4}$
Fitted TPFA	$8.98.10^{-4}$	$8.83.10^{-4}$	$8.75.10^{-4}$	$8.71.10^{-4}$	$8.69.10^{-4}$	$8.66.10^{-4}$	$8.656.10^{-4}$

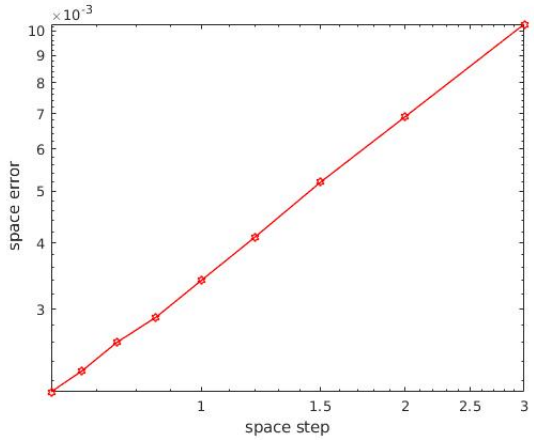


Figure 2.3: Space step versus L^2 Errors in loglog scale. This graph shows the convergence in space of the fitted TPFA. The order of convergence in space is $\mathcal{O}(h)$, this is in agreement with the theoretical result in Theorem 8. The time step is fixed to be $dt = 1/100$.

Conclusion

In this Chapter, we have presented two spatial numerical methods for spatial discretization of the Black-Scholes PDE for pricing options. The first scheme is the classical finite volume method with Two-Point Flux Approximation (TPFA) and the second scheme is a novel scheme called fitted Two-Point Flux Approximation (TPFA). The novel fitted Two-Point Flux Approximation (TPFA) combines the classical fitted finite volume method and the standard TPFA method. The classical fitted finite volume method is used to tackle the degeneracy of the Black-Scholes PDE near zero. The convergence analyses are performed along with numerical experiments to confirm the theoretical results.

Chapter 3

A Multi-Point Flux Approximation and fitted Multi-Point Flux Approximation method for two dimensional pricing options: The O-method

In this Chapter, we develop novel numerical methods based on the Multi-Point Flux Approximation (MPFA) method to solve partial differential equation (PDE) arising from pricing two-assets option. We should notice that the TPFA method introduced in the previous Chapter can only be extended to a two dimensions problem on K [1] orthogonal grids (see Aavatsmark [2007]), which are difficult to construct. The MPFA methods appear then as a solution to this shortcoming because they can be applied on general grids. Here, the O-MPFA method is used as our first method and it is coupled with a fitted finite volume to handle the degeneracy of the PDE and the corresponding scheme is called fitted MPFA method. The convection part is discretized using the upwinding methods (first and second order) that we have derived on non uniform grids. The time discretization is performed with the $\theta-$ Euler methods. Numerical simulations show that our new schemes are more accurate than the current fitted finite volume method proposed in the literature. This Chapter is published in Koffi and Tambue [2019c]

3.1 The finite volume formulation

An option with two underlying assets modeled by the Black Scholes equation is formulated as follows

$$\begin{cases} dx(t) &= \mu_1 x dt + \sigma_1 x dW_1, \\ dy(t) &= \mu_2 y dt + \sigma_2 y dW_2, \\ dW_1(t) dW_2(t) &= \rho dt, \end{cases} \tag{3.1}$$

where μ_i, σ_i, W_i are respectively the drift, the volatility and the Wiener process governing the stocks x, y and ρ is the correlation coefficient between the two Wiener processes. By applying the Ito's formula and using the standard arbitrage argument (see Section 1.3.2, Chapter 1), it is well known that the value of the option U follows the following two-dimensional Black-Scholes Partial differential equation on the domain $D = [0, +\infty) \times [0, +\infty) \times [0, T]$

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma_1^2 x^2 \frac{\partial^2 U}{\partial x^2} + \rho \sigma_1 \sigma_2 xy \frac{\partial^2 U}{\partial x \partial y} + \frac{1}{2}\sigma_2^2 y^2 \frac{\partial^2 U}{\partial y^2} + rx\frac{\partial U}{\partial x} + ry\frac{\partial U}{\partial y} - rU, \tag{3.2}$$

[1] K being the diffusion tensor, in our study it will be denoted by M.

where $\tau = T - t$, T is the maturity time, t the current time and r is the risk-free interest. For European rainbow option price on maximum of two risky assets, the following initial and boundary conditions are used

$$\begin{cases} U(x,y,0) = \max\left(\max(x,y) - K, 0\right), \\ U(0,y,\tau) = 0, \\ U(x,0,\tau) = 0, \end{cases} \tag{3.3}$$

with K the strike price. However, to compare our numerical solution with the existing fitted finite volume method, the exact solution will be used at the boundary. In order to apply the finite volume method, it is convenient to re-write the Partial Differential Equation (3.2) in the following divergence form

$$\frac{\partial U}{\partial \tau} = \nabla \cdot (\mathbf{M} \nabla U) + \nabla (fU) + \lambda U, \tag{3.4}$$

where

$$\mathbf{M} = \frac{1}{2} \begin{pmatrix} \sigma_1^2 x^2 & \rho\sigma_1\sigma_2 xy \\ \rho\sigma_1\sigma_2 xy & \sigma_2^2 y^2 \end{pmatrix}, f = \begin{pmatrix} (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)x \\ (r - \sigma_2^2 - \frac{1}{2}\rho\sigma_1\sigma_2)y \end{pmatrix},$$

$$\lambda = -3r + \sigma_1^2 + \sigma_2^2 + \rho\sigma_1\sigma_2.$$

Note that $\mathbf{M}$ does not satisfying the standard ellipticity condition (see [Tambue, 2016, (3)]), so the PDE (3.4) is degenerated. We will assume Dirichlet boundary condition in the entire domain.
Let us consider the domain of study $\Omega = I_x \times I_y \times [0,T]$ where $I_x = [0, x_{\max}]$ and $I_y = [0, y_{\max}]$. In the sequel of this work, the Black-Scholes partial differential equation (3.2) is considered over the truncated domain Ω.
At $x = x_{\max}$ and $y = y_{\max}$, the linear boundary condition will be applied (see Huang et al. [2006]). The intervals I_x and I_y will be subdivided into N part in the following way (see Huang et al. [2006, 2009]) without loss the generality as irregular grids such as triangular grids can be used.

$$I_{x_i} = [x_{i-1}; x_i],\ I_{y_j} = [y_{j-1}; y_j] \quad i,j = 1,2,...,N+1, \tag{3.5}$$

with $h_i = x_{i+1} - x_i, \quad l_j = y_{j+1} - y_j$.

Let us set the mid-points $x_{i-\frac{1}{2}}$ and $y_{j-\frac{1}{2}}$ as follows

$$x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2},\ y_{j-\frac{1}{2}} = \frac{y_{j-1} + y_j}{2}, \qquad i,j = 1,2,..,N, \tag{3.6}$$

with $k_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, \quad \gamma_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}, \quad$ and

$$x_{-\frac{1}{2}} = x_0 = 0, \qquad x_{N+\frac{3}{2}} = x_{N+1} = x_{\max},\ y_{-\frac{1}{2}} = y_0 = 0\,,\ y_{N+\frac{3}{2}} = y_{N+1} = y_{\max}.$$

For $i,j = 1,2,\ldots,N$, we denote by $\mathcal{C}_{ij} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ a control volume associated to our subdivision.

Note that the control volume $\mathcal{C}_{ij}$ is the area surrounding the grid point (x_i, y_j). Our goal is to approximate the option function U at (x_i, y_j) [2] by a function denoted $\mathcal{U}$. The matrix $\mathbf{M}$ in (3.4) will be replaced by its average value within each control volume as follows:

[2] center of the control volume $\mathcal{C}_{i,j}$

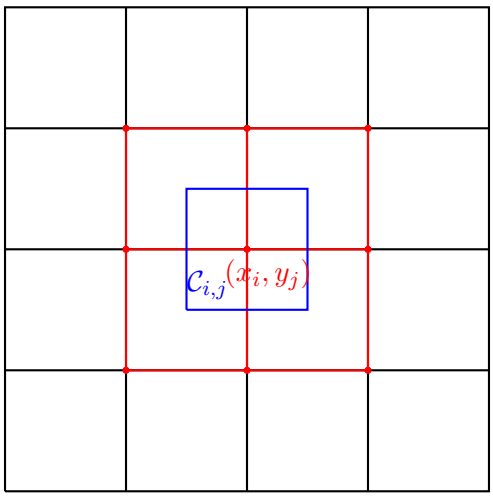


Figure 3.1: The control volume $\mathcal{C}_{i,j}$

$$\mathbf{M}^{ij} = \frac{1}{\text{meas}(\mathcal{C}_{i,j})} \int_{\mathcal{C}_{i,j}} \mathbf{M} dxdy, \quad i,j = 1, ..., N. \tag{3.7}$$

where $\text{meas}(\mathcal{C}_{ij})$ is the measure of $\mathcal{C}_{ij}$.

Thereby, we have

$$M^{ij} = \begin{bmatrix} \frac{\sigma_1^2}{6} \frac{x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} & \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) \\ \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) & \frac{\sigma_2^2}{6} \frac{y^3_{j+\frac{1}{2}} - y^3_{j-\frac{1}{2}}}{y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}} \end{bmatrix}.$$

Now let us consider the divergence form given in (3.4). Following the finite volume method's principle, we integrate the partial differential equation (3.4) over each control volume $\mathcal{C}_{ij}$ and we have

$$\int_{\mathcal{C}_{ij}} \frac{\partial U}{\partial \tau} d\mathcal{C} = \int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M} \nabla U) d\mathcal{C} + \int_{\mathcal{C}_{ij}} \nabla (fU) d\mathcal{C} + \int_{\mathcal{C}_{ij}} \lambda U d\mathcal{C}. \tag{3.8}$$

The next Section will be dedicated to spatial discretization of equation (3.8). For the term in the left hand side of (3.8) and for the last term in its right hand side, we use the mid-point quadrature rule for their approximations. More precisely

$$\int_{\mathcal{C}_{ij}} \frac{\partial U}{\partial \tau} d\mathcal{C} \approx \text{meas}(\mathcal{C}_{ij}) \frac{d\mathcal{U}}{d\tau}(x_i, y_j, \tau), \tag{3.9}$$

$$\int_{\mathcal{C}_{ij}} \lambda U d\mathcal{C} \approx meas(\mathcal{C}_{ij}) \lambda \mathcal{U}(x_i, y_j, \tau). \tag{3.10}$$

The diffusion term

$$\int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M} \nabla \mathcal{U}) d\mathcal{C}, \tag{3.11}$$

of (3.8) will be approximated using the **Multi-point flux approximation** (MPFA) method or our novel **fitted Multi-point flux approximation**. More details will be given in the next section. Besides, the convection term

$$\int_{\mathcal{C}_{ij}} \nabla (f\mathcal{U}) d\mathcal{C}, \tag{3.12}$$

of (3.8) will be approximated using the upwind methods (first or second order). Note that the standard two -point flux approximation in Tambue [2016] can only be consistent in the approximation of (3.11) if and only if the grid is $\mathbf{M}-$orthogonal.

3.2 The Multi-Point Flux Approximation (MPFA): O-method

Let us start by applying the divergence theorem to the diffusion term (3.11) as follows

$$\mathcal{F}^{ij} = \int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M}^{ij} \nabla \mathcal{U}) = \int_{\partial \mathcal{C}_{ij}} (\mathbf{M}^{ij} \nabla \mathcal{U}) \cdot \vec{n} d\mathcal{C}, \quad i, j = 1, 2, \ldots, N, \tag{3.13}$$

where $\vec{n}$ is the outward vector from the control volume.

Now, we can apply the so-called **Multi-Point Flux Aprroximation(MPFA)** to approximate the integral defined in (3.13).
Nevertheless, let us give a geometrical reminder which will be useful for the application of our method.

3.2.1 Geometrical reminder

Let us consider a triangle $x_1x_2x_3$ (see Figure 3.2).

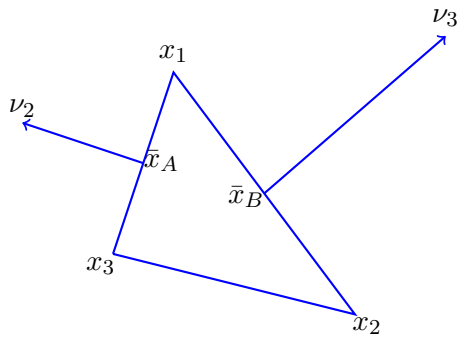


Figure 3.2: Triangle

Any linear function may be described by

$$\mathcal{U}(x) = \sum_{i=1}^{3} \mathcal{U}_i \phi_i(x).$$

Here, $\mathcal{U}_i$ is the value of $\mathcal{U}(x)$ at the vertex i, and $\phi_i(x)$ is the linear basis function defined by

$$\phi_i(x_j) = \delta_{ij}. \qquad (i)$$

Since we are in a 2-dimensional case, the linear basis function is under the following form:

$$\phi_i(x, y) = \alpha + \lambda x + \beta y = [1 \quad x \quad y] \begin{bmatrix} \alpha \\ \lambda \\ \beta \end{bmatrix} = p(x, y)\Gamma. \qquad (ii)$$

So for i=1,2,3 we have the following system of equations:

$$\begin{cases} \phi_i^1 = \phi_i(x_1, y_1) = \alpha + \lambda x_1 + \beta y_1, \\ \phi_i^2 = \phi_i(x_2, y_2) = \alpha + \lambda x_2 + \beta y_2, \\ \phi_i^3 = \phi_i(x_3, y_3) = \alpha + \lambda x_3 + \beta y_3. \end{cases}$$

This system of equations can be re-written as follows:

$$\varphi_i = M.\Gamma, \qquad (iii)$$

where

$$\varphi_i = \begin{bmatrix} \phi_i^1 \\ \phi_i^2 \\ \phi_i^3 \end{bmatrix}, M = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \quad and \quad \Gamma = \begin{bmatrix} \alpha \\ \lambda \\ \beta \end{bmatrix}.$$

From (iii), we have $\Gamma = M^{-1}\varphi_i$ and using the expression of Γ in (ii) we get

$$\phi_i(x,y) = p(x,y)M^{-1}\varphi_i = N(x,y)\varphi_i, \qquad (iv)$$

where $N(x,y) = p(x,y)M^{-1}$. Besides, we have

$$M^{-1} = \frac{1}{det(M)} \begin{bmatrix} x_2y_3 - x_3y_2 & -(x_1y_3 - x_3y_1) & x_1y_2 - x_2y_1 \\ -(y_3 - y_2) & y_3 - y_1 & -(y_2 - y_1) \\ x_3 - x_2 & -(x_3 - x_1) & x_2 - x_1 \end{bmatrix}.$$

Thereby, the expressions of the components of vector N are given by

$$N_1(x,y) = \frac{1}{det(M)} \left[(x_2y_3 - x_3y_2) - x(y_3 - y_2) + y(x_3 - x_2) \right],$$

$$N_2(x,y) = \frac{1}{det(M)} \left[-(x_1y_3 - x_3y_1) + x(y_3 - y_1) - y(x_2 - x_1) \right],$$

$$N_3(x,y) = \frac{1}{det(M)} \left[x_1y_2 - x_2y_1 + x(y_2 - y_1) + y(x_2 - x_1) \right],$$

with $det(M) = 2F$ and F is the area of the triangle $x_1x_2x_3$. From (iv), we have :

$$\phi_i(x,y) = \left(\phi_i^1 N_1(x,y) + \phi_i^2 N_2(x,y) + \phi_i^3 N_3(x,y) \right).$$

then the gradient of ϕ_i is given by

$$\begin{aligned} \nabla\phi_i &= \begin{bmatrix} \frac{\partial\phi_i(x,y)}{\partial x} \\ \\ \frac{\partial\phi_i(x,y)}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \phi_i^1 \frac{\partial N_1(x,y)}{\partial x} + \phi_i^2 \frac{\partial N_2(x,y)}{\partial x} + \phi_i^3 \frac{\partial N_3(x,y)}{\partial x} \\ \\ \phi_i^1 \frac{\partial N_1(x,y)}{\partial y} + \phi_i^2 \frac{\partial N_2(x,y)}{\partial y} + \phi_i^3 \frac{\partial N_3(x,y)}{\partial y} \end{bmatrix} \\ \nabla\phi_i &= \frac{1}{2F} \begin{bmatrix} -\phi_i^1(y_3 - y_2) + \phi_i^2(y_3 - y_1) - \phi_i^3(y_2 - y_1) \\ \\ \phi_i^1(x_3 - x_2) - \phi_i^2(x_3 - x_1) + \phi_i^3(x_2 - x_1) \end{bmatrix}. \end{aligned}$$

Since $\phi_i^j = \delta_{ij}$, therefore we have:

$$\nabla\phi_1 = \frac{1}{2F} \begin{bmatrix} -(y_3 - y_2) \\ \\ (x_3 - x_2) \end{bmatrix}, \ \nabla\phi_2 = \frac{1}{2F} \begin{bmatrix} (y_3 - y_1) \\ \\ -(x_3 - x_1) \end{bmatrix} \text{ and } \nabla\phi_3 = \frac{1}{2F} \begin{bmatrix} -(y_2 - y_1) \\ (x_2 - x_1) \end{bmatrix}.$$

We may notice that the vector $\omega_1 = \begin{bmatrix} -(y_3 - y_2) \\ \\ (x_3 - x_2) \end{bmatrix}$ is orthogonal to vector $\overrightarrow{X_2X_3}$, the same length with this vector and inner to the triangle $X_1X_2X_3$. It is similar for $\omega_2 = \begin{bmatrix} (y_3 - y_1) \\ \\ -(x_3 - x_1) \end{bmatrix}$ with vector $\overrightarrow{X_1X_3}$ and vector $\omega_3 = \begin{bmatrix} -(y_2 - y_1) \\ \\ (x_2 - x_1) \end{bmatrix}$ with vector $\overrightarrow{X_1X_2}$. Hence, the gradient is given by

$$\nabla\phi_i = -\frac{1}{2F}v_i, \tag{3.14}$$

where F is the area of the triangle , v_i is the outer normal vector of the edge located opposite of vertex i, the length of v_i equals the length of the edge to which it is normal.
For these normal vectors the following relations holds:

$$\sum_{i=1}^{3} v_i = 0. \tag{3.15}$$

Thus, the gradient expression of the potential in the triangle may be written in the form

$$\begin{aligned}
\nabla\mathcal{U} &= \nabla\left(\sum_{i=1}^{3}\mathcal{U}_i\phi_i(x)\right) \\
&= \sum_{i=1}^{3}\mathcal{U}_i\nabla\phi_i(x) \\
&= \sum_{i=1}^{3}\mathcal{U}_i(-\frac{1}{2F}v_i) \\
&= -\frac{1}{2F}\sum_{i=1}^{3}\mathcal{U}_i v_i \\
&= -\frac{1}{2F}(-\mathcal{U}_1(v_2+v_3)+\mathcal{U}_2 v_2+\mathcal{U}_3 v_3) \\
\nabla\mathcal{U} &= -\frac{1}{2F}\left[(\mathcal{U}_2-\mathcal{U}_1)v_2+(\mathcal{U}_3-\mathcal{U}_1)v_3\right].
\end{aligned} \tag{3.16}$$

In the following Section, the gradient of the function $\mathcal{U}$ over a triangle will be useful in the calculation of flux.

3.2.2 Flux through half edge inside an interaction volume

We first consider the control volume $\mathcal{C}_{ij}$ and its center is $x_k = (x_i, y_j)$ (see Figure 3.3). Using the local indices, the mid-points on the edges are denoted $\bar{x}_1$ and $\bar{x}_2$. We denote also by Γ_1 *and* Γ_2 the inner normal vector to edge located opposite of respectively vertex $\bar{x}_1$ and $\bar{x}_2$, with same length with the corresponding edge. The area of triangle $x_k\bar{x}_1\bar{x}_2$ is denoted F. Using the gradient of U expression (3.16) over the triangle $x_k\bar{x}_1\bar{x}_2$ and the fact that the normal vector of edges Γ_1 *and* Γ_2 are inner unlike normal vectors in Figure 4.2, we have

$$\begin{aligned}
\nabla\mathcal{U} &= -\frac{1}{2F}\left[(\bar{\mathcal{U}}_1-\mathcal{U}_k)(-\Gamma_1)+(\bar{\mathcal{U}}_2-\mathcal{U}_k)(-\Gamma_2)\right] \\
\nabla\mathcal{U} &= \frac{1}{2F}\left[\Gamma_1(\bar{\mathcal{U}}_1-\mathcal{U}_k)+\Gamma_2(\bar{\mathcal{U}}_2-U_k)\right].
\end{aligned} \tag{3.17}$$

We will let these normal vector pointing in the direction of the increasing cell indices and the horizontal ones will be denoted Γ_1 and the vertical ones Γ_2. This gives

$$\Gamma_1 = \frac{1}{2}\begin{bmatrix} h_i \\ \\ 0 \end{bmatrix}, \qquad \Gamma_2 = \frac{1}{2}\begin{bmatrix} 0 \\ \\ l_j \end{bmatrix}.$$

Each of the edges p can be associated with a global direction, defined through the unit normal n_p. It is convenient to let n_p point in the direction of increasing global cell indices. The flux through half edge p as seen from the control volume $\mathcal{C}_{ij}$ is denoted f_p^{ij}, and may now be determined from the gradient of the potential in the control volume.

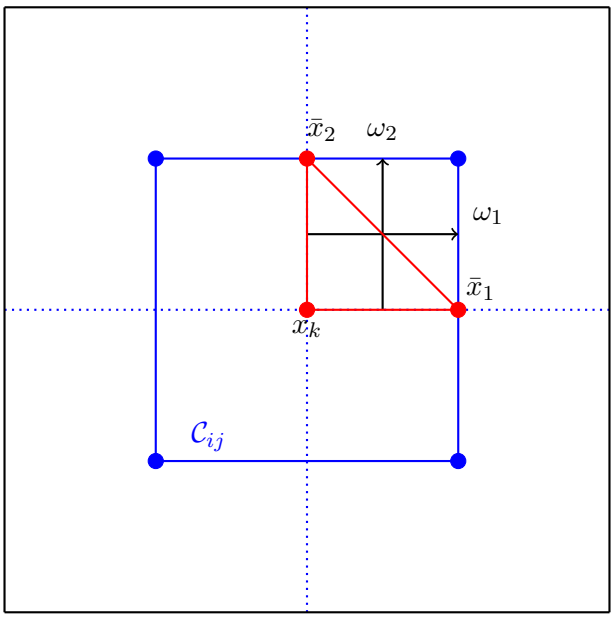


Figure 3.3: Triangle inside a control volume

The flux through an half edge p seen from the control volume $\mathcal{C}_{ij}$ is given by

$$f_p^{ij} = \Gamma_p n_p^T M^{ij} \nabla \mathcal{U}. \tag{3.18}$$

Here, we denote by f_h^{ij} (respectively f_v^{ij})the flux through an half horizontal edge (respectively through an half horizontal edge) seen from the control volume $\mathcal{C}_{ij}$ $\quad for\ i, j = 1, 2, \ldots, N$ and we have

$$\begin{bmatrix} f_h^{ij} \\ \\ f_v^{ij} \end{bmatrix} = \begin{bmatrix} \Gamma_1 n_1^T \\ \\ \Gamma_2 n_2^T \end{bmatrix} M^{ij} \nabla \mathcal{U}. \tag{3.19}$$

In our study, the interaction volume are rectangle, then:

$$\Gamma_p n_p = \frac{a_p}{2},$$

$$F = \frac{\mathcal{A}}{8},$$

$$\mathcal{A} = h_i l_j,$$

where $a_p,\ p = 1, 2$, is a vector parallel to Γ_p and twice the length of Γ_p and pointing toward increasing indices cell, F is the area of the considered triangle, and $\mathcal{A}$ is the area of the interaction volume.

It follows that

$$
\begin{aligned}
\begin{bmatrix} f_h^{ij} \\ \\ f_v^{ij} \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} a_1 \\ \\ a_2 \end{bmatrix} \mathbf{M}^{ij} \nabla \mathcal{U} \\
&= \frac{1}{2} \times \frac{1}{2\mathcal{A}} \begin{bmatrix} a_1 \\ \\ a_2 \end{bmatrix} \mathbf{M}^{ij} \times \frac{1}{2} \left[a_1(\bar{\mathcal{U}}_1 - \mathcal{U}_{ij}) + a_2(\bar{\mathcal{U}}_2 - \mathcal{U}_{ij}) \right] \\
&= \frac{1}{8\mathcal{A}} \begin{bmatrix} a_1 \\ \\ a_2 \end{bmatrix} \mathbf{M}^{ij} \left[a_1(\bar{\mathcal{U}}_1 - \mathcal{U}_{ij}) + a_2(\bar{\mathcal{U}}_2 - \mathcal{U}_{ij}) \right] \\
&= \frac{1}{8\mathcal{A}} \begin{bmatrix} a_1 \\ \\ a_2 \end{bmatrix} \mathbf{M}^{ij} \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} \bar{\mathcal{U}}_1 - \mathcal{U}_{ij} \\ \\ \bar{\mathcal{U}}_2 - \mathcal{U}_{ij} \end{bmatrix} \\
\begin{bmatrix} f_h^{ij} \\ \\ f_v^{ij} \end{bmatrix} &= \frac{1}{\mathcal{A}} \begin{bmatrix} a_1 \\ \\ a_2 \end{bmatrix} \mathbf{M}^{ij} \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} \bar{\mathcal{U}}_1 - \mathcal{U}_{ij} \\ \\ \bar{\mathcal{U}}_2 - \mathcal{U}_{ij} \end{bmatrix}, \qquad (3.20)
\end{aligned}
$$

where

$$\bar{\mathcal{U}}_p = \mathcal{U}(\bar{x}_p),$$

with $\bar{x}_p$ midpoint of the edge p (p=1,2,...,4) inside the interaction volume $\mathcal{R}_{ij}$. We get

$$\begin{bmatrix} f_h^{ij} \\ \\ f_v^{ij} \end{bmatrix} = G^{ij} \begin{bmatrix} \bar{\mathcal{U}}_1 - \mathcal{U}_{ij} \\ \\ \bar{\mathcal{U}}_2 - \mathcal{U}_{ij} \end{bmatrix}, \qquad (3.21)$$

where

$$G^{ij} = \begin{bmatrix} \frac{\sigma_1^2}{6} \frac{x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}}{y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}} & \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) \\ \\ \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) & \frac{\sigma_2^2}{6} \frac{y^3_{j+\frac{1}{2}} - y^3_{j-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} \end{bmatrix}. \qquad (3.22)$$

Besides, an interaction volume $\mathcal{R}_{ij}$(see red line in Figure 3.4) is a cell grid defined as follows

$$for \;\; i,j = 1,2,\ldots,N+1, \quad \mathcal{R}_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j].$$

We may also notice that the interaction volume $\mathcal{R}_{ij}$ is covering an area in the intersection of the controls volume $\mathcal{C}_{ij}$, $\mathcal{C}_{i+1,j}$, $\mathcal{C}_{i,j+1}$ and $\mathcal{C}_{i+1,j+1}$ (see Figure 3.4).

Remark 2 *Our goal here is to compute the flux through the half edges* $1, 2, 3$ *and* 4 *(see green lines in Figure 3.4) inside the interaction of volume* $\mathcal{R}_{ij}$ *by considering the triangles* $x_1\bar{x}_1\bar{x}_3$, $x_2\bar{x}_1\bar{x}_4$, $x_3\bar{x}_2\bar{x}_3$ *and* $x_3\bar{x}_2\bar{x}_4$ *where the vertices are* $x_1 = (x_i, y_j)$, $x_2 = (x_{i+1}, y_j)$
$x_3 = (x_i, y_{j+2})$ *and* $x_4 = (x_{i+1}, y_{j+1})$.

Thereby, the fluxes through the half edges 1 and 3 in the control volume $\mathcal{C}_{ij}$ are calculated as follows:

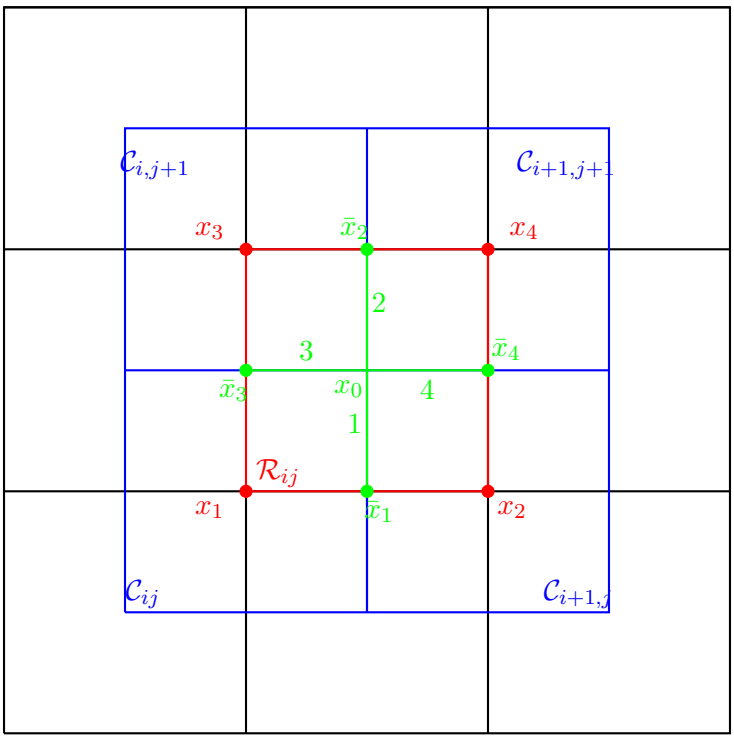


Figure 3.4: Interaction volume

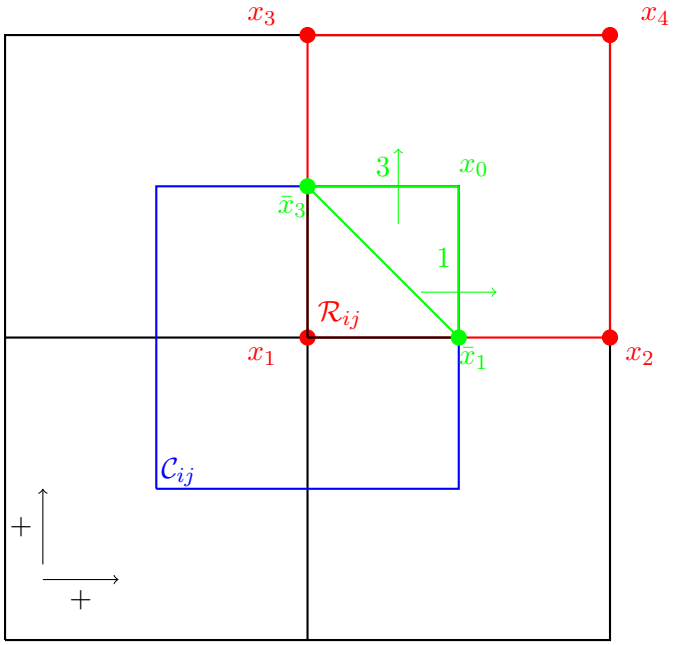


Figure 3.5: Triangle in an interaction volume

Let us consider the triangle $x_1\bar{x}_1\bar{x}_3$. Using the expression of the gradient $\nabla\mathcal{U}$ in (3.17), the gradient of $\mathcal{U}$ over the triangle $x_1\bar{x}_1\bar{x}_3$ is given by

$$\nabla\mathcal{U} = \frac{1}{4F}\left[a_1(\bar{\mathcal{U}}_1 - \mathcal{U}_{ij}) + a_2(\bar{\mathcal{U}}_3 - \mathcal{U}_{ij})\right]. \tag{3.23}$$

Since the triangles $x_1\bar{x}_1\bar{x}_3$ and $x_0\bar{x}_1\bar{x}_3$ (see Figure 3.5) are all contained in the control volume $\mathcal{C}_{ij}$ and $\mathcal{U}$ is constant over the control volume then the expression of the gradient $\nabla\mathcal{U}$ over the triangle $x_1\bar{x}_1\bar{x}_3$ is the same over the triangle $x_0\bar{x}_1\bar{x}_3$.
Thereby, using (3.21) and the fact that the fluxes through the edges 1 and 3 are in positive direction (increasing cell indices direction) then:

$$\begin{bmatrix} f_1^{ij} \\ \\ f_3^{ij} \end{bmatrix} = G^{ij} \begin{bmatrix} \bar{\mathcal{U}}_1 - \mathcal{U}_{ij} \\ \\ \bar{\mathcal{U}}_3 - \mathcal{U}_{ij} \end{bmatrix}. \tag{3.24}$$

Moreover, to calculate the fluxes through the half edges 1 and 4 of the control volume $\mathcal{C}_{i+1,j}$,

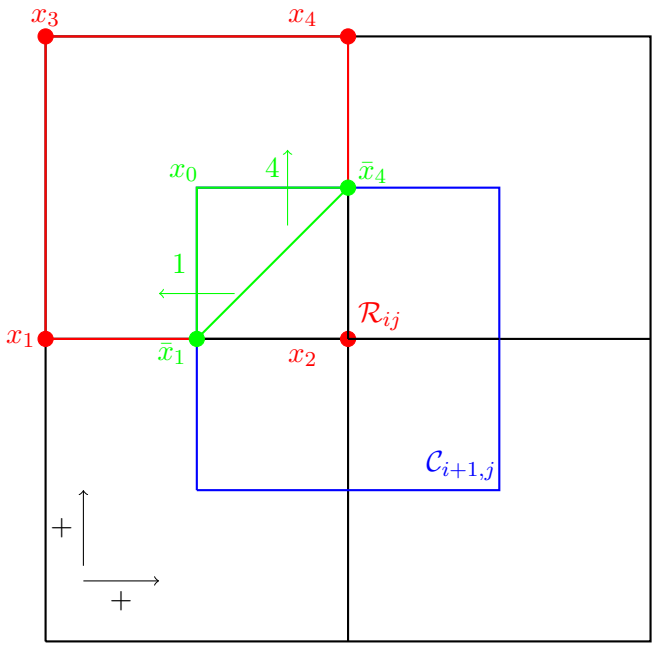


Figure 3.6: Triangle in an interaction volume

We consider the triangle $x_2\bar{x}_1\bar{x}_4$. Using the expression (3.17) of the gradient we have :

$$\nabla\mathcal{U} = \frac{1}{4F}\left[a_1(\bar{\mathcal{U}}_1 - \mathcal{U}_{i+1,j}) + a_2(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j})\right].$$

Since the triangles $x_2\bar{x}_1\bar{x}_4$ and $x_0\bar{x}_1\bar{x}_4$ are all contained in the control volume $\mathcal{C}_{i+1,j}$ and the fact that $\mathcal{U}$ is constant over the control volume then the expression of the gradient $\nabla\mathcal{U}$ over the triangle $x_2\bar{x}_1\bar{x}_4$ is the same over the triangle $x_0\bar{x}_1\bar{x}_4$ (see green lines in Figure 3.6).
Thereby, using (3.21) and the fact that the flux through the edges 1 in the negative direction and 4 is in positive direction (increasing cell indices direction) then:

$$\begin{bmatrix} f_1^{i+1,j} \\ \\ f_4^{i+1,j} \end{bmatrix} = G^{i+1,j} \begin{bmatrix} -(\bar{\mathcal{U}}_1 - \mathcal{U}_{i+1,j}) \\ \\ \bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j} \end{bmatrix} = G^{i+1,j} \begin{bmatrix} \mathcal{U}_{i+1,j} - \bar{\mathcal{U}}_1 \\ \\ \bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j} \end{bmatrix}. \tag{3.25}$$

Similarly, to compute the fluxes through the half edges 2 and 4 of the control volume $\mathcal{C}_{i+1,j+1}$, we will consider the interation volume $\mathcal{R}_{ij}$.

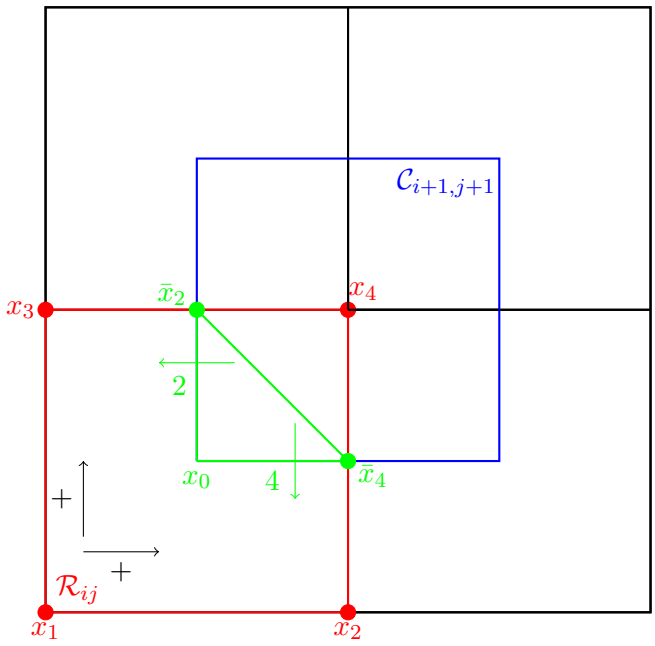


Figure 3.7: Triangle in an interaction volume

We consider the triangle $x_4\bar{x}_2\bar{x}_4$ (see Figure 3.7). Using the expression (3.17) of the gradient, gives

$$\nabla\mathcal{U} = \frac{1}{4F}\left[a_1(\bar{\mathcal{U}}_1 - \mathcal{U}_{i+1,j+1}) + a_2(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j+1})\right]. \tag{3.26}$$

Since the triangles $x_4\bar{x}_2\bar{x}_4$ and $x_4\bar{x}_2\bar{x}_4$ are all contained in the control volume $\mathcal{C}_{i+1,j+1}$ and $\mathcal{U}$ is constant over the control volume then the expression of the gradient $\nabla\mathcal{U}$ over the triangle $x_4\bar{x}_2\bar{x}_4$ is the same over the $x_0\bar{x}_2\bar{x}_4$.
Thereby, using (3.21) and the fact that the fluxes through the edges 2 and 4 in the negative direction (opposite of the increasing cell indices direction) then:

$$\begin{bmatrix} f_2^{i+1,j+1} \\ \\ f_4^{i+1,j+1} \end{bmatrix} = G^{i+1,j+1} \begin{bmatrix} -\left(\bar{\mathcal{U}}_2 - \mathcal{U}_{i+1,j+1}\right) \\ \\ -\left(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j+1}\right) \end{bmatrix} = G^{i+1,j+1} \begin{bmatrix} \mathcal{U}_{i+1,j+1} - \bar{\mathcal{U}}_2 \\ \\ \mathcal{U}_{i+1,j+1} - \bar{\mathcal{U}}_4 \end{bmatrix}. \tag{3.27}$$

Furthermore, to compute the fluxes through the half edges 2 and 3 in the control volume $\mathcal{C}_{i,j+1}$, we consider the interaction volume $\mathcal{R}_{ij}$.

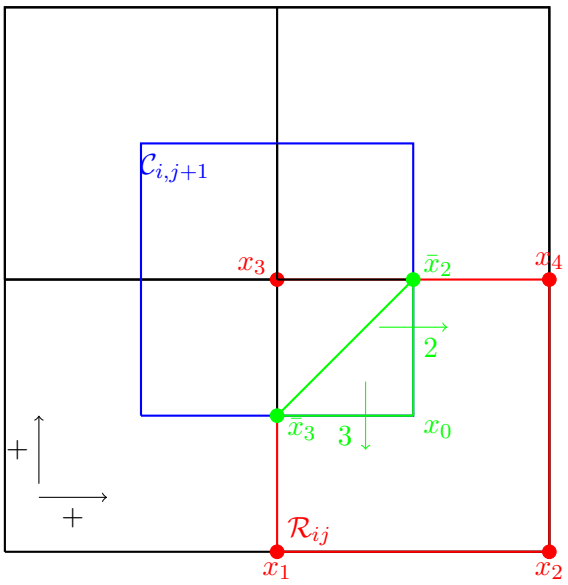


Figure 3.8: Triangle in an interaction volume

We consider the triangle $x_3\bar{x}_3\bar{x}_2$ (see Figure 3.8). Using the expression of the gradient (3.17), it follows that

$$\nabla\mathcal{U} = \frac{1}{4F}\left[a_1(\bar{\mathcal{U}}_2 - \mathcal{U}_{i,j+1}) + a_2(\bar{\mathcal{U}}_3 - \mathcal{U}_{i,j+1})\right]. \tag{3.28}$$

Since the triangles $x_3\bar{x}_3\bar{x}_2$ and $x_0\bar{x}_3\bar{x}_2$ are all contained in the control volume $\mathcal{C}_{i,j+1}$ and $\mathcal{U}$ is constant over the control volume then the expression of the gradient $\nabla\mathcal{U}$ (3.28) over the triangle $x_3\bar{x}_3\bar{x}_2$ is the same over the triangle $x_0\bar{x}_3\bar{x}_2$.
Thereby, using (3.21) and the fact that the flux through the edge 2 is in the positive direction and the flux through the half edge 3 is in the negative direction (opposite of the increasing cell indices direction) lead to

$$\begin{bmatrix} f_2^{i,j+1} \\ f_3^{i,j+1} \end{bmatrix} = G^{i,j+1}\begin{bmatrix} \bar{\mathcal{U}}_2 - \mathcal{U}_{i,j+1} \\ -\left(\bar{\mathcal{U}}_3 - \mathcal{U}_{i,j+1}\right) \end{bmatrix} = G^{i,j+1}\begin{bmatrix} \bar{\mathcal{U}}_2 - \mathcal{U}_{i,j+1} \\ \mathcal{U}_{i,j+1} - \bar{\mathcal{U}}_3 \end{bmatrix}. \tag{3.29}$$

Besides, the flux through an edge is continuous. Then we have

$$\begin{aligned} f_1 &= f_1^{ij} &&= f_1^{i,j+1}, \\ f_2 &= f_2^{i+1,j+1} &&= f_2^{i,j+1}, \\ f_3 &= f_3^{i,j+1} &&= f_3^{ij}, \\ f_4 &= f_4^{i+1,j} &&= f_4^{i+1,j+1}. \end{aligned} \tag{3.30}$$

Thereby, using (3.24),(3.25),(3.27) and (3.29), these equations become

$$
\begin{aligned}
f_1 &= g_{11}^{ij}(\bar{\mathcal{U}}_1 - \mathcal{U}_{ij}) + g_{12}^{ij}(\bar{\mathcal{U}}_3 - \mathcal{U}_{ij}), = -g_{11}^{i+1,j}(\bar{\mathcal{U}}_1 - \mathcal{U}_{i+1,j}) + g_{12}^{i+1,j}(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j}),\\
f_2 &= -g_{11}^{i+1,j+1}(\bar{\mathcal{U}}_2 - \mathcal{U}_{i+1,j+1}) - g_{12}^{i+1,j+1}(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j+1}) = g_{11}^{i,j+1}(\bar{\mathcal{U}}_2 - \mathcal{U}_{i,j+1}) - g_{12}^{i,j+1}(\bar{\mathcal{U}}_3 - \mathcal{U}_{i,j+1}),\\
f_3 &= g_{21}^{i,j+1}(\bar{\mathcal{U}}_2 - \mathcal{U}_{i,j+1}) - g_{22}^{i,j+1}(\bar{\mathcal{U}}_3 - \mathcal{U}_{i,j+1}) = g_{21}^{ij}(\bar{\mathcal{U}}_1 - \mathcal{U}_{ij}) + g_{22}^{ij}(\bar{\mathcal{U}}_3 - \mathcal{U}_{ij}),\\
f_4 &= -g_{21}^{i+1,j}(\bar{\mathcal{U}}_1 - \mathcal{U}_{i+1,j}) + g_{22}^{i+1,j}(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j}) = -g_{21}^{i+1,j+1}(\bar{\mathcal{U}}_2 - \mathcal{U}_{i+1,j+1}) - g_{22}^{i+1,j+1}(\bar{\mathcal{U}}_4 - \mathcal{U}_{i+1,j+1}).
\end{aligned}
\tag{3.31}
$$

Setting

$$
f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}, \qquad \mathcal{U} = \begin{bmatrix} \mathcal{U}_{ij} \\ \mathcal{U}_{i+1,j} \\ \mathcal{U}_{i,j+1} \\ \mathcal{U}_{i+1,j+1} \end{bmatrix}, \qquad \mathcal{V} = \begin{bmatrix} \bar{\mathcal{U}}_1 \\ \bar{\mathcal{U}}_2 \\ \bar{\mathcal{U}}_3 \\ \bar{\mathcal{U}}_4 \end{bmatrix}.
$$

We can rewrite (3.31) under the following form, using the left-hand side of each equality sign

$$
f = C\mathcal{V} + F\mathcal{U}, \tag{3.32}
$$

where

$$
C^{ij} = \begin{bmatrix} g_{11}^{ij} & 0 & g_{12}^{ij} & 0 \\ 0 & -g_{11}^{i+1,j+1} & 0 & -g_{12}^{i+1,j+1} \\ 0 & g_{21}^{i,j+1} & -g_{22}^{i,j+1} & 0 \\ -g_{21}^{i+1,j} & 0 & 0 & g_{22}^{i+1,j} \end{bmatrix},
$$

$$
F^{ij} = \begin{bmatrix} -g_{11}^{ij} - g_{12}^{ij} & 0 & 0 & 0 \\ 0 & 0 & 0 & g_{11}^{i+1,j+1} + g_{12}^{i+1,j+1} \\ 0 & 0 & -g_{21}^{i,j+1} + g_{22}^{i,j+1} & 0 \\ 0 & g_{21}^{i+1,j} - g_{22}^{i+1,j} & 0 & 0 \end{bmatrix}.
$$

Using the right-hand side of each equality sign of equation (3.31), one can write :

$$
A\mathcal{V} = B\mathcal{U}, \tag{3.33}
$$

with

$$
A^{ij} = \begin{bmatrix} g_{11}^{ij} + g_{11}^{i+1,j} & 0 & g_{12}^{ij} & -g_{12}^{i+1,j} \\ 0 & -g_{11}^{i+1,j+1} - g_{11}^{i,j+1} & g_{12}^{i,j+1} & -g_{12}^{i+1,j+1} \\ -g_{21}^{ij} & g_{21}^{i,j+1} & -g_{22}^{ij} - g_{22}^{i,j+1} & 0 \\ -g_{21}^{i+1,j} & g_{21}^{i+1,j+1} & 0 & g_{22}^{i+1,j} + g_{22}^{i+1,j+1} \end{bmatrix},
$$

$$B^{ij} = \begin{bmatrix} g_{11}^{ij} + g_{12}^{ij} & g_{11}^{i+1,j} - g_{12}^{i+1,j} & 0 & 0 \\ 0 & 0 & -g_{11}^{i,j+1} + g_{12}^{i,j+1} & -g_{11}^{i+1,j+1} - g_{12}^{i+1,j+1} \\ -g_{21}^{ij} - g_{22}^{ij} & 0 & g_{21}^{i,j+1} - g_{22}^{i,j+1} & 0 \\ 0 & -g_{21}^{i+1,j} + g_{22}^{i+1,j} & 0 & g_{21}^{i+1,j+1} + g_{22}^{i+1,j+1} \end{bmatrix}.$$

Thereby, $\mathcal{V}$ may be eliminated by solving (3.33) with respect to $\mathcal{V}$. By substituting $\mathcal{V} = A^{-1}BU$ in (3.32), the flux through the half edges 1,2,3 and 4 inside the interaction volume $\mathcal{R}_{ij}$ is given by

$$f = T^{ij}\mathcal{U}, \qquad i,j = 1,2,\ldots,N+1, \tag{3.34}$$

with

$$T^{ij} = C^{ij}[A^{ij}]^{-1}B^{ij} + F^{ij}, \tag{3.35}$$

where T^{ij} is called the transmissibility matrix of the interaction volume $\mathcal{R}_{ij}$.

Remark 3 *We should notice that at this point, to calculate the flux through an half edge of a control volume, we need to know the interaction volume to which it belongs and its position (position 1,2,3 or 4) in this interaction volume (see Figure 3.4).*

3.2.3 Flux through edges of a control volume

Let us recall that the Multi-Point Flux Approximation method is used to approximate the integral (3.11) defined as follows:

$$\begin{aligned} \int_{\mathcal{C}_{ij}} \nabla.(\mathbf{M}\nabla\mathcal{U})d\mathcal{C} &= \int_{\partial\mathcal{C}_{ij}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n}\,d\mathcal{C} \\ &= \int_{\mathcal{E}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{E}}}d\mathcal{C} + \int_{\mathcal{N}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{N}}}d\mathcal{C} \\ &\quad + \int_{\mathcal{W}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{W}}}d\mathcal{C} + \int_{S} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{S}}d\mathcal{C}. \end{aligned} \tag{3.36}$$

where $\mathcal{E}$,$\mathcal{N}$,$\mathcal{W}$ and S denote respectively the eastern, northern, western and southern of edge a control volume. In the rest of the document, we will denote by:

$$\begin{aligned} \mathcal{F}^{ij} = \int_{\mathcal{C}_{ij}} \nabla.(\mathbf{M}\nabla\mathcal{U})d\mathcal{C}_{ij} &= \int_{\partial\mathcal{C}_{ij}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n}\,d\mathcal{C}, \\ {}_{\mathcal{E}}f^{ij} &= \int_{\mathcal{E}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{E}}}d\mathcal{C}, \qquad {}_{\mathcal{N}}f^{ij} = \int_{\mathcal{N}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{N}}}d\mathcal{C}, \\ {}_{\mathcal{W}}f^{ij} &= \int_{\mathcal{W}} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{\mathcal{W}}}d\mathcal{C}, \qquad {}_{S}f^{ij} = \int_{S} (\mathbf{M}\nabla\mathcal{U})\overrightarrow{n_{S}}d\mathcal{C}. \end{aligned}$$

Let us consider, the following control volume $\mathcal{C}_{ij}$

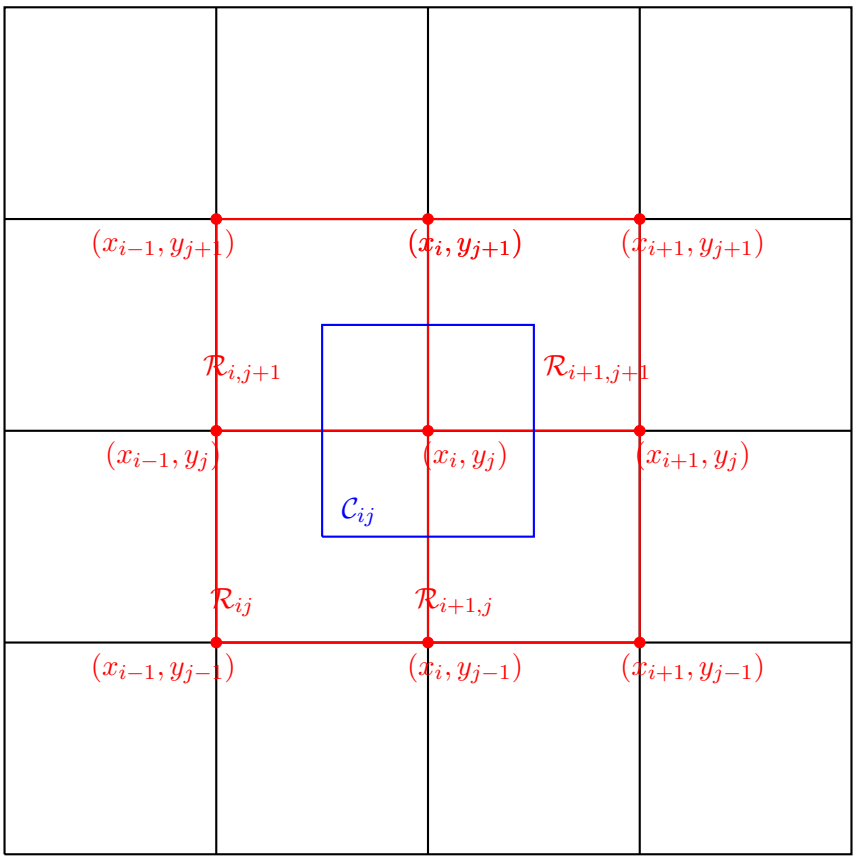


Figure 3.9: Control volume

We may notice that for calculating the flux through all the the edges of a control volume, we need to consider 4 interaction volumes (see Figure 3.9).

Flux $_{\mathcal{E}}f^{ij}$ through the eastern of the control volume $\mathcal{C}_{ij}$

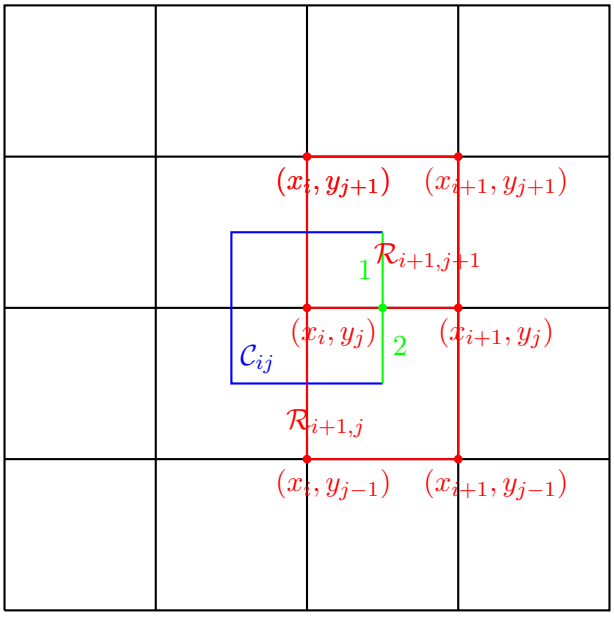


Figure 3.10: Eastern edge of a control volume

The lower eastern half edge is contained in the interaction volume $\mathcal{R}_{i+1,j}$ and it is in position 2 in the interaction of volume (see Figure 3.10). Thereby, by using (3.34) it follows that

$$\varepsilon f_d^{ij} = T_{21}^{i+1,j}\mathcal{U}_{i,j-1} + T_{22}^{i+1,j}\mathcal{U}_{i+1,j-1} + T_{23}^{i+1,j}\mathcal{U}_{ij} + T_{24}^{i+1,j}\mathcal{U}_{i+1,j}.$$

Similarly, the upper eastern half edge is contained in the interaction volume $\mathcal{R}_{i+1,j+1}$ and it is in position 1 in the interaction volume. So by using (3.34), it follows that

$$\varepsilon f_u^{ij} = T_{11}^{i+1,j+1}\mathcal{U}_{ij} + T_{12}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{13}^{i+1,j+1}\mathcal{U}_{i,j+1} + T_{14}^{i+1,j+1}\mathcal{U}_{i+1,j+1}.$$

Finally the flux through the eastern edge of the control volume $\mathcal{C}_{ij}$ will be the addition of εf_d^{ij} and εf_u^{ij}. Thereby we get

$$\begin{aligned}
\varepsilon f^{ij} &= \varepsilon f_d^{ij} + \varepsilon f_u^{ij} \\
&= T_{21}^{i+1,j}\mathcal{U}_{i,j-1} + T_{22}^{i+1,j}\mathcal{U}_{i+1,j-1} + T_{23}^{i+1,j}\mathcal{U}_{ij} + T_{24}^{i+1,j}\mathcal{U}_{i+1,j} + T_{11}^{i+1,j+1}\mathcal{U}_{ij} \\
&\quad + T_{12}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{13}^{i,j}\mathcal{U}_{i,j+1} + T_{14}^{i,j}\mathcal{U}_{i+1,j+1} \\
\varepsilon f^{ij} &= (T_{11}^{i+1,j+1} + T_{23}^{i+1,j})\mathcal{U}_{ij} + (T_{12}^{i+1,j+1} + T_{24}^{i+1,j})\mathcal{U}_{i+1,j} + T_{14}^{i+1,j+1}\mathcal{U}_{i+1,j+1} \\
&\quad + T_{13}^{i+1,j+1}\mathcal{U}_{i,j+1} + T_{21}^{i+1,j}\mathcal{U}_{i,j-1} + T_{22}^{i+1,j}\mathcal{U}_{i+1,j-1}.
\end{aligned} \tag{3.37}$$

Flux $_{\mathcal{N}}f^{ij}$ through the northern edge of the control volume $\mathcal{C}_{ij}$

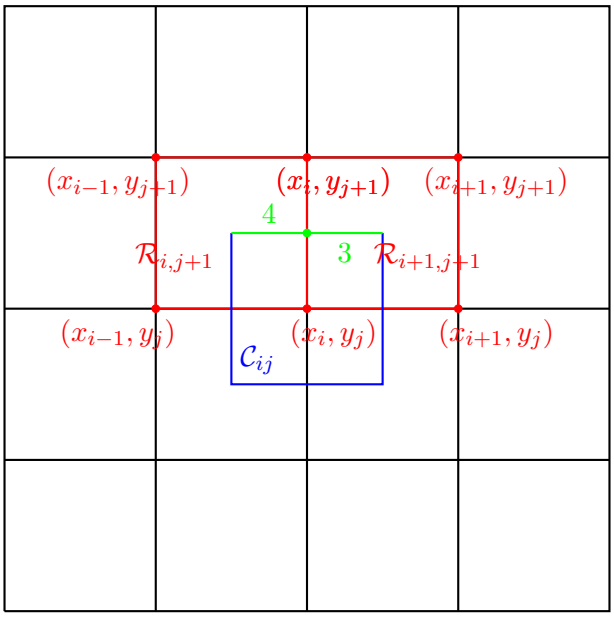


Figure 3.11: Northern edge of a control volume

Besides, the left northern half edge is contained in the interaction volume $\mathcal{R}_{i,j+1}$ and it is position 4 in the interaction volume (See Figure 3.11). then by using (3.34) we have

$$_{\mathcal{N}}f_{lf}^{ij} = T_{41}^{i,j+1}\mathcal{U}_{i-1,j} + T_{42}^{i,j+1}\mathcal{U}_{ij} + T_{43}^{i,j+1}\mathcal{U}_{i-1,j+1} + T_{44}^{i,j+1}\mathcal{U}_{i,j+1}.$$

The right northern half edge of the control volume $\mathcal{C}_{ij}$ is contained in the interaction volume $\mathcal{R}_{i+1,j+1}$ and it is in position 3 in the interaction volume. So by using (3.34), it follows that

$$_{\mathcal{N}}f_r^{ij} = T_{31}^{i+1,j+1}\mathcal{U}_{ij} + T_{32}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{33}^{i+1,j+1}\mathcal{U}_{i,j+1} + T_{34}^{i+1,j+1}\mathcal{U}_{i+1,j+1},$$

then the flux ${}_{\mathcal{N}}f^{ij}$ through the northern edge of the control volume $\mathcal{C}_{ij}$, is given by

$$\begin{aligned}
{}_{\mathcal{N}}f^{ij} &= {}_{\mathcal{N}}f_{lf}^{ij} + {}_{\mathcal{N}}f_{r}^{ij} \\
&= T_{41}^{i,j+1}\mathcal{U}_{i-1,j} + T_{42}^{i,j+1}\mathcal{U}_{ij} + T_{43}^{i,j+1}\mathcal{U}_{i-1,j+1} + T_{44}^{i,j+1}\mathcal{U}_{i,j+1} + T_{31}^{i+1,j+1}\mathcal{U}_{ij} \\
&\quad + T_{32}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{33}^{i+1,j+1}\mathcal{U}_{i,j+1} + T_{34}^{i+1,j+1}\mathcal{U}_{i+1,j+1} \\
{}_{\mathcal{N}}f^{ij} &= (T_{31}^{i+1,j+1} + T_{42}^{i,j+1})\mathcal{U}_{ij} + T_{32}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{34}^{i+1,j+1}\mathcal{U}_{i+1,j+1} \\
&\quad + (T_{33}^{i+1,j+1} + T_{44}^{i,j+1})\mathcal{U}_{i,j+1} + T_{43}^{i,j+1}\mathcal{U}_{i-1,j+1} + T_{41}^{i,j+1}\mathcal{U}_{i-1,j}.
\end{aligned} \tag{3.38}$$

Flux ${}_{\mathcal{W}}f^{ij}$ through the northern edge of the control volume $\mathcal{C}_{ij}$

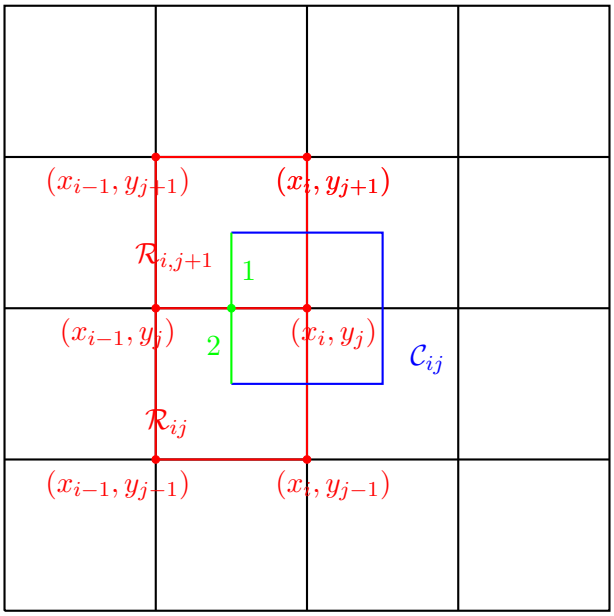


Figure 3.12: Western edge of a control volume

The lower western half edge is contained in the interaction volume $\mathcal{R}_{ij}$ and it is in position 2, then by using (3.34) it follows that

$${}_{\mathcal{W}}f_{d}^{ij} = T_{21}^{ij}\mathcal{U}_{i-1,j-1} + T_{22}^{ij}\mathcal{U}_{i,j-1} + T_{23}^{i,j}\mathcal{U}_{i-1,j} + T_{24}^{ij}\mathcal{U}_{ij}.$$

The upper western half edge is contained in the interaction volume $\mathcal{R}_{i,j+1}$ and it is in position 1, so by using (3.34), it follows that

$${}_{\mathcal{W}}f_{u}^{ij} = T_{11}^{i,j+1}\mathcal{U}_{i-1,j} + T_{12}^{i,j+1}\mathcal{U}_{ij} + T_{13}^{i,j+1}\mathcal{U}_{i-1,j+1} + T_{14}^{i,j+1}\mathcal{U}_{i,j+1}.$$

Thus the flux ${}_{\mathcal{W}}f^{ij}$ through the west edge of the control volume $\mathcal{C}_{ij}$ is given by

$$
\begin{aligned}
{}_{W}f^{ij} &= {}_{W}f_d^{i+1,j+1} + {}_{W}f_u^{i+1,j+1} \\
&= T_{21}^{ij}\mathcal{U}_{i-1,j-1} + T_{22}^{ij}\mathcal{U}_{i,j-1} + T_{23}^{ij}\mathcal{U}_{i-1,j} + T_{24}^{ij}\mathcal{U}_{ij} + T_{11}^{i,j+1}\mathcal{U}_{i-1,j} \\
&\quad + T_{12}^{i,j+1}\mathcal{U}_{ij} + T_{13}^{i,j+1}\mathcal{U}_{i-1,j+1} + T_{14}^{i,j+1}\mathcal{U}_{i,j+1} \\
{}_{W}f^{ij} &= (T_{12}^{i,j+1} + T_{24}^{ij})\mathcal{U}_{ij} + T_{14}^{i,j+1}\mathcal{U}_{i,j+1} + T_{13}^{i,j+1}\mathcal{U}_{i-1,j+1} + (T_{11}^{i,j+1} + T_{23}^{ij})\mathcal{U}_{i-1,j} \\
&\quad + T_{21}^{ij}\mathcal{U}_{i-1,j-1} + T_{22}^{ij}\mathcal{U}_{i,j-1}.
\end{aligned} \tag{3.39}
$$

Flux ${}_{S}f^{ij}$ through the northern edge of the control volume $\mathcal{C}_{ij}$

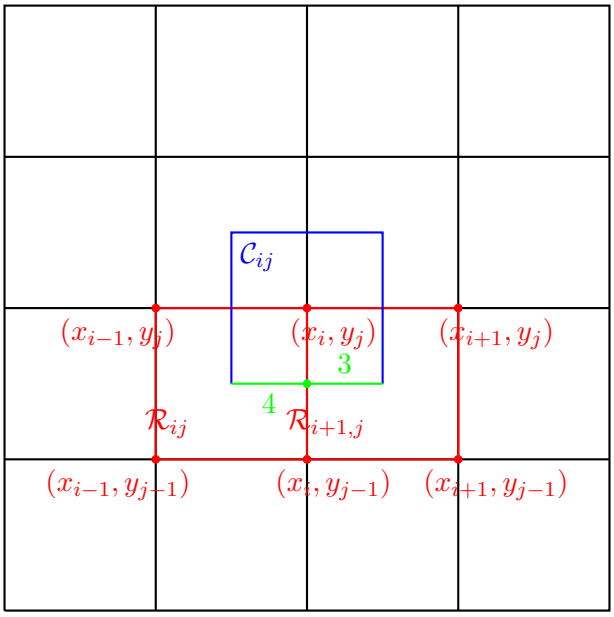


Figure 3.13: Southern edge of a control volume

The left southern edge is contained in the interaction volume $\mathcal{R}_{ij}$ and it is in position 4 in the interaction volume (see Figure 3.13), then by using (3.34) it follows that

$$
{}_{S}f_{lf}^{ij} = T_{41}^{ij}\mathcal{U}_{i-1,j-1} + T_{42}^{ij}\mathcal{U}_{i,j-1} + T_{43}^{ij}\mathcal{U}_{i-1,j} + T_{44}^{ij}\mathcal{U}_{ij}.
$$

The right southern half edge is contained in the interaction volume $\mathcal{R}_{i+1,j}$ and it is in position 3 in the interaction of volume(see Figure 3.13), then by using (3.34) it follows that

$$
{}_{S}f_{r}^{ij} = T_{31}^{i+1,j}\mathcal{U}_{i,j-1} + T_{32}^{i+1,j}\mathcal{U}_{i+1,j-1} + T_{33}^{i+1,j}\mathcal{U}_{ij} + T_{34}^{i+1,j}\mathcal{U}_{i+1,j}.
$$

Thus the flux ${}_{S}f^{ij}$ through the south edge of our control volume $\mathcal{C}_{ij}$ is given by

$$
\begin{aligned}
{}_S f^{ij} &= {}_S f_{lf}^{ij} + {}_S f_r^{ij} \\
&= T_{41}^{ij}\mathcal{U}_{i-1,j-1} + T_{42}^{ij}\mathcal{U}_{i,j-1} + T_{43}^{ij}\mathcal{U}_{i-1,j} + T_{44}^{ij}\mathcal{U}_{ij} + T_{31}^{i+1,j}\mathcal{U}_{i,j-1} \\
&\quad + T_{32}^{i+1,j}\mathcal{U}_{i+1,j-1} + T_{33}^{i+1,j}\mathcal{U}_{ij} + T_{34}^{i+1,j}\mathcal{U}_{i+1,j} \\
{}_S f^{ij} &= (T_{33}^{i+1,j} + T_{44}^{ij})\mathcal{U}_{ij} + T_{34}^{i+1,j}\mathcal{U}_{i+1,j} + T_{43}^{ij}\mathcal{U}_{i-1,j} + T_{41}^{ij}\mathcal{U}_{i-1,j-1} \\
&\quad + (T_{31}^{i+1,j} + T_{42}^{ij})\mathcal{U}_{i,j-1} + T_{32}^{i+1,j}\mathcal{U}_{i+1,j-1}.
\end{aligned}
\tag{3.40}
$$

Remark 4 *Let us remember that all the fluxes in the previous paragraph have been calculated in the direction of the increasing cell indices. Thereby, to take in account the outward normal vector direction, the fluxes through the western ${}_{\mathcal{W}}f$ and southern ${}_S f$ edge of a control volume will be counted negatively.*

Therefore, for $i, j = 1, 2, \ldots, N$, the outflux $\mathcal{F}^{ij}$ through the edges of the control volume $\mathcal{C}_{ij}$ is given by:

$$
\mathcal{F}^{ij} = {}_{\mathcal{E}}f^{ij} + {}_{\mathcal{N}}f^{ij} - {}_{\mathcal{W}}f^{ij} - {}_S f^{ij}.
$$

Thereby, we have

$$
\begin{aligned}
\mathcal{F}_{ij} &= {}_{\mathcal{E}}f^{ij} + {}_{\mathcal{N}}f^{ij} - {}_{\mathcal{W}}f^{ij} - {}_S f^{ij} \\
&= \Big((T_{11}^{i+1,j+1} + T_{23}^{i+1,j})\mathcal{U}_{ij} + (T_{12}^{i+1,j+1} + T_{24}^{i+1,j})\mathcal{U}_{i+1,j} + T_{14}^{i+1,j+1}\mathcal{U}_{i+1,j+1} \\
&\quad + T_{13}^{i+1,j+1}\mathcal{U}_{i,j+1} + T_{21}^{i+1,j}\mathcal{U}_{i,j-1} + T_{22}^{i+1,j}\mathcal{U}_{i+1,j-1} \Big) + \Big((T_{31}^{i+1,j+1} + T_{42}^{i,j+1})\mathcal{U}_{ij} \\
&\quad + T_{32}^{i+1,j+1}\mathcal{U}_{i+1,j} + T_{34}^{i+1,j+1}\mathcal{U}_{i+1,j+1} + (T_{33}^{i+1,j+1} + T_{44}^{i,j+1})\mathcal{U}_{i,j+1} + T_{43}^{i,j+1}\mathcal{U}_{i-1,j+1} \\
&\quad + T_{41}^{i,j+1}\mathcal{U}_{i-1,j} \Big) - \Big((T_{12}^{i,j+1} + T_{24}^{ij})\mathcal{U}_{ij} + T_{14}^{i,j+1}\mathcal{U}_{i,j+1} + T_{13}^{i,j+1}\mathcal{U}_{i-1,j+1} \\
&\quad + (T_{11}^{i,j+1} + T_{23}^{ij})\mathcal{U}_{i-1,j} + T_{21}^{ij}\mathcal{U}_{i-1,j-1} + T_{22}^{ij}\mathcal{U}_{i,j-1} \Big) - \Big((T_{33}^{i+1,j} + T_{44}^{ij})\mathcal{U}_{ij} + T_{34}^{i+1,j}\mathcal{U}_{i+1,j} \\
&\quad + T_{43}^{ij}\mathcal{U}_{i-1,j} + T_{41}^{ij}\mathcal{U}_{i-1,j-1} + (T_{31}^{i+1,j} + T_{42}^{ij})\mathcal{U}_{i,j-1} + T_{32}^{i+1,j}\mathcal{U}_{i+1,j-1} \Big).
\end{aligned}
$$

Finally, for $i, j = 1, ..., N$, the outflux $\mathcal{F}^{ij}$, through the edges of the control volume $\mathcal{C}_{ij}$, is given by

$$
\begin{aligned}
\mathcal{F}^{ij} &= a_{ij}\mathcal{U}_{ij} + b_{ij}\mathcal{U}_{i+1,j} + c_{ij}\mathcal{U}_{i+1,j+1} + d_{ij}\mathcal{U}_{i,j+1} + e_{ij}\mathcal{U}_{i-1,j+1} + \alpha_{ij}\mathcal{U}_{i-1,j} + \beta_{ij}\mathcal{U}_{i-1,j-1} \\
&\quad + \gamma_{ij}\mathcal{U}_{i,j-1} + \lambda_{ij}\mathcal{U}_{i+1,j-1},
\end{aligned}
\tag{3.41}
$$

where

$$a_{ij} = T_{11}^{i+1,j+1} + T_{23}^{i+1,j} + T_{31}^{i+1,j+1} + T_{42}^{i,j+1} - T_{12}^{i,j+1} - T_{24}^{ij} - T_{33}^{i+1,j} - T_{44}^{ij},$$

$$b_{ij} = T_{12}^{i+1,j+1} + T_{24}^{i+1,j} + T_{32}^{i+1,j+1} - T_{34}^{i+1,j},$$

$$c_{ij} = T_{14}^{i+1,j+1} + T_{34}^{i+1,j+1},\ d_{ij} = T_{13}^{i+1,j+1} + T_{33}^{i+1,j+1} + T_{44}^{i,j+1} - T_{14}^{i,j+1},\ e_{ij} = T_{43}^{i,j+1} - T_{13}^{i,j+1},$$

$$\alpha_{ij} = T_{41}^{i,j+1} - T_{11}^{i,j+1} - T_{23}^{ij} - T_{43}^{ij}, \quad \beta_{ij} = -T_{21}^{ij} - T_{41}^{ij},$$

$$\gamma_{ij} = T_{21}^{i+1,j} - T_{22}^{ij} - T_{31}^{i+1,j} - T_{42}^{ij},$$

$$\lambda_{ij} = T_{22}^{i+1,j} - T_{32}^{i+1,j}.$$

Let us notice that for the control volumes near to the boundary of the our domain, some terms from the boundary conditions will be involved in (3.41) .
Hence (3.11) becomes

$$\mathcal{F} = A_{mp}\mathcal{U} + F_{mp}, \tag{3.42}$$

where A_{mp} is a $N^2 \times N^2$ matrix and

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_{11} \\ \mathcal{F}_{12} \\ \vdots \\ \mathcal{F}_{1N} \\ \mathcal{F}_{21} \\ \mathcal{F}_{22} \\ \vdots \\ \vdots \\ \mathcal{F}_{NN} \end{bmatrix}, \quad \mathcal{U} = \begin{bmatrix} \mathcal{U}_{11} \\ \mathcal{U}_{12} \\ \vdots \\ \mathcal{U}_{1N} \\ \mathcal{U}_{21} \\ \mathcal{U}_{22} \\ \vdots \\ \vdots \\ \mathcal{U}_{NN} \end{bmatrix}, \quad A_{mp} = \begin{bmatrix} W_1 & X_1 & 0_N & \dots & \dots & \dots & \dots & 0_N \\ Y_2 & W_2 & X_2 & \ddots & & & & \vdots \\ 0_N & Y_3 & W_3 & X_3 & \ddots & & & \vdots \\ \vdots & \ddots & Y_4 & W_4 & X_4 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & & & & \ddots & Y_{N-1} & W_{N-1} & X_{N-1} \\ 0_N & \dots & \dots & \dots & \dots & 0_N & Y_N & W_N \end{bmatrix},$$

with 0_N is $N \times N$ null matrix , W_i, Y_i, X_i are tridiagonal matrices, and F_{mp} is a N^2 vector coming from the boundary. The structure of the diffusion matrix A_{mp} can be viewed in Figure 3.14

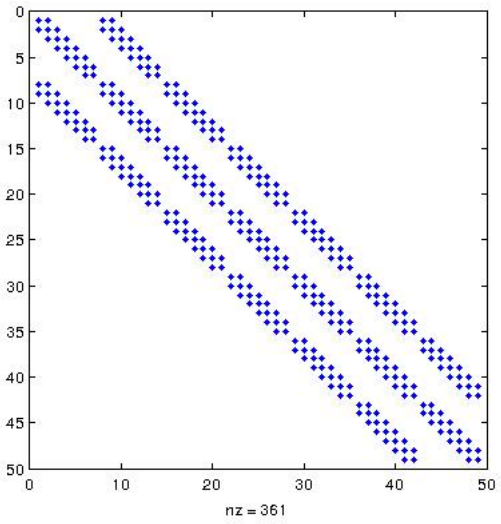


Figure 3.14: Structure of diffusion matrix coming from standard MPFA

In the next Section, the convection term will be discretized using the upwind 1^{st} order and 2^{nd} order methods.

3.3 Upwinds methods

The integral of convection term (3.12)

$$\int_{\mathcal{C}_{ij}} \nabla(f\mathcal{U})d\mathcal{C},$$

where

$$f = \begin{pmatrix} (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)x \\ \\ (r - \sigma_2^2 - \frac{1}{2}\rho\sigma_1\sigma_2)y \end{pmatrix},$$

will be approximated using the upwind methods (1^{st} and 2^{nd} order). We start by applying the divergence theorem, and we have for $i, j = 1, 2, ..., N$:

$$I^{ij} = \int_{\mathcal{C}_{ij}} \nabla(f\mathcal{U}) = \int_{\partial\mathcal{C}_{ij}} (f \cdot \mathcal{U}) \cdot \vec{n} d\mathcal{C}, \tag{3.43}$$

with $\vec{n}$ an outward unit normal vector. Since our control volume $\mathcal{C}_{ij}$ has four edges, then we have

$$\begin{aligned} I^{ij} &= I^{ij}_{\mathcal{E}} + I^{ij}_{\mathcal{N}} - I^{ij}_{\mathcal{W}} - I^{ij}_{S} \\ &= \int_{\mathcal{E}_{ij}} (f\mathcal{U}).\vec{n}_{\mathcal{E}} d\mathcal{C} + \int_{\mathcal{N}_{ij}} (f\mathcal{U}).\vec{n}_{\mathcal{N}} d\mathcal{C} - \int_{\mathcal{W}_{ij}} (f\mathcal{U}).\vec{n}_{\mathcal{W}} d\mathcal{C} - \int_{S_{ij}} (f\mathcal{U}).\vec{n}_{S} d\mathcal{C}, \end{aligned}$$

where $\mathcal{E}_{ij}, \mathcal{N}_{ij}, \mathcal{W}_{ij}$ and S_{ij} are respectively the eastern, northern, western and southern edges of the control volume $\mathcal{C}_{ij}$ and $\vec{n}_{\mathcal{E}}, \vec{n}_{\mathcal{N}}, \vec{n}_{\mathcal{W}}, \vec{n}_{S}$ are respectively the normal outward vector to the eastern, northern, western and southern edges of the control volume $\mathcal{C}_{ij}$.

3.3.1 First order upwind method

The **first order upwind method** discussed in [LeVeque, 2004, chapter 4.8] will be applied to evaluate the second term of (3.8).

I^{ij} is calculated by summing up the flux through the edges of the control volume $\mathcal{C}_{ij}$.
The flux through an edge using the first order upwind will depend on the sign of $f \cdot \vec{n}$ on this edge. If the sign of $f \cdot \vec{n}$ is positive, $\mathcal{U}_{ij}$ will be used to approximate $(f \cdot \vec{n}\mathcal{U})$ otherwise we will use the value of $\mathcal{U}$ in other side of the edge. This procedure is detailed in following paragraphs.

Flux through the eastern edge of control volume $\mathcal{C}_{ij}, i, j = 1, \ldots, N$

.

The normal outward vector to the eastern edge of the control volume $\mathcal{C}_{ij}$ is

$$\vec{n}_{\mathcal{E}} = \begin{bmatrix} 1 \\ \\ 0 \end{bmatrix}. \tag{3.44}$$

Thereby, we have

$$f.\vec{n}_{\mathcal{E}} = f_x = (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2)x,$$

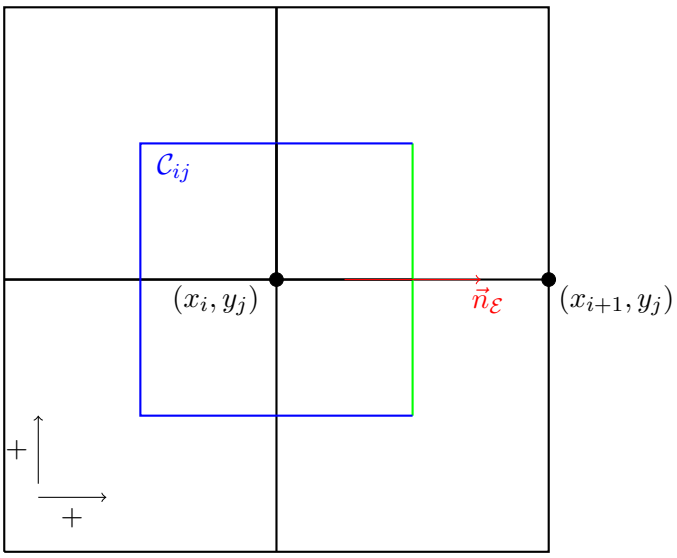


Figure 3.15: Eastern edge of a control volume

then, according the upwind method we have :

$$f_{\mathcal{E}} = (f.\vec{n}_{\mathcal{E}})\mathcal{U} = \begin{cases} f_x \mathcal{U}_{ij} & if \quad f_x \geq 0, \\ \\ f_x \mathcal{U}_{i+1,j} & if \quad f_x < 0. \end{cases} \tag{3.45}$$

Let us set

$$\mathcal{U}_{\mathcal{E}} = \begin{cases} \mathcal{U}_{ij} & if \quad f_x \geq 0, \\ \\ \mathcal{U}_{i+1,j} & if \quad f_x < 0. \end{cases} \tag{3.46}$$

Thus the integral $I_{\mathcal{E}}^{ij}$ over the east edge of the control volume $\mathcal{C}_{ij}$ is given by

$$\begin{aligned} I_{\mathcal{E}}^{ij} &= \int_{\mathcal{E}_{ij}} f_{\mathcal{E}} dy \\ &= \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}} dy \\ &= (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} dy \\ I_{\mathcal{E}}^{ij} &= \gamma_j (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}}. \end{aligned}$$

Hence

$$I_{\mathcal{E}}^{ij} = \gamma_j (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2) x_i \left[\max(f_x, 0)\mathcal{U}_{ij} + \min(f_x, 0)\mathcal{U}_{i+1,j} \right],$$

$$I_{\mathcal{E}}^{ij} = \gamma_j \left[\max(f_x^i, 0)\mathcal{U}_{ij} + \min(f_x^i, 0)\mathcal{U}_{i+1,j} \right], \tag{3.47}$$

with

$$f_x^i = (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2) x_{i+\frac{1}{2}}. \tag{3.48}$$

Flux through the western edge of the control volume $\mathcal{C}_{ij}$, $for \quad i,j = 1,...,N$.

The normal outward vector to the western edge of the control volume $\mathcal{C}_{ij}$ is

$$\vec{n}_{\mathcal{W}} = \begin{bmatrix} -1 \\ \\ 0 \end{bmatrix} = -\vec{n}_{\mathcal{E}}. \tag{3.49}$$

Using the similar approach as in the case of the eastern edge of the control volume $\mathcal{C}_{ij}$, we get

$$I^{ij}_{\mathcal{W}} = \gamma_j \left[max(f^{i-1}_x, 0)\mathcal{U}_{i-1,j} + min(f^{i-1}_x, 0)\mathcal{U}_{ij} \right], \tag{3.50}$$

with

$$f^{i-1}_x = (r - \sigma_1 - \frac{1}{2}\rho\sigma_1\sigma_2)x_{i-\frac{1}{2}}.$$

Flux through the north edge of the control volume $\mathcal{C}_{ij}$, $for \quad i,j = 1,...,N$.

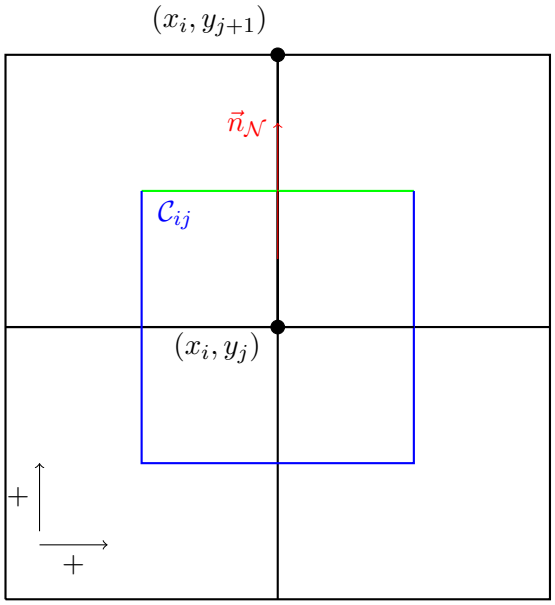


Figure 3.16: Northern edge of a control volume

The normal outward vector to the east edge of the control volume $\mathcal{C}_{ij}$ is given by

$$\vec{n}_{\mathcal{N}} = \begin{bmatrix} 0 \\ \\ 1 \end{bmatrix}. \tag{3.51}$$

Thereby, we have

$$f.\vec{n}_{\mathcal{N}} = f_y = (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2)y.$$

Then, according the upwind method we have :

$$f_{\mathcal{N}} = (f.\vec{n}_{\mathcal{N}})\mathcal{U} = \begin{cases} f_y\mathcal{U}_{ij} & if \quad f_y \geq 0, \\ \\ f_y\mathcal{U}_{i,j+1} & if \quad f_y < 0. \end{cases} \tag{3.52}$$

let us set

$$\mathcal{U}_{\mathcal{N}} = \begin{cases} \mathcal{U}_{ij} & if \quad f_y \geq 0, \\ \\ \mathcal{U}_{i,j+1} & if \quad f_y < 0. \end{cases} \tag{3.53}$$

Thus the integral $I_{\mathcal{N}}^{ij}$ over the east edge of the control volume $\mathcal{C}_{ij}$ is given by

$$\begin{aligned} I_{\mathcal{N}}^{ij} &= \int_{\mathcal{N}_{ij}} f_{\mathcal{N}} dx \\ &= \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}} dx \\ &= (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \\ I_{\mathcal{N}}^{ij} &= k_i (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}}. \end{aligned}$$

Hence we have

$$I_{\mathcal{N}}^{ij} = k_i (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j+\frac{1}{2}} \left[\max(f_\nu, 0)\mathcal{U}_{ij} + \min(f_\nu, 0)\mathcal{U}_{i,j+1} \right],$$

$$I_{\mathcal{N}}^{ij} = k_i \left[\max(f_y^j, 0)\mathcal{U}_{ij} + \min(f_y^j, 0)\mathcal{U}_{i,j+1} \right], \tag{3.54}$$

with

$$f_y^j = (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j+\frac{1}{2}}. \tag{3.55}$$

Flux through the southern edge of the control volume $\mathcal{C}_{ij}, \quad for \quad i, j = 1, ..., N$

. The normal outward vector to the southern edge of the control volume $\mathcal{C}_{ij}$ is

$$\vec{n}_S = \begin{bmatrix} 0 \\ \\ -1 \end{bmatrix} = -\vec{n}_{\mathcal{N}}. \tag{3.56}$$

Using the similar approach as in the case of the northern edge of the control volume $\mathcal{C}_{ij}$, we get

$$I_S^{ij} = k_i \left[\max(f_y^{j-1}, 0)\mathcal{U}_{i,j-1} + \min(f_y^{j-1}, 0)\mathcal{U}_{ij} \right], \tag{3.57}$$

with

$$f_y^{j-1} = (r - \sigma_2 - \frac{1}{2}\rho\sigma_1\sigma_2) y_{j-\frac{1}{2}}.$$

Therefore, the outflux through the edges of the control volume $\mathcal{C}_{ij} \quad \forall i, j = 1, \ldots, N$ is

$$
\begin{aligned}
I^{ij} &= I^{ij}_{\mathcal{E}} + I^{ij}_{\mathcal{N}} - I^{ij}_{\mathcal{W}} - I^{ij}_{S} \\
&= \left[\gamma_j\left(\max(f^i_x,0)\mathcal{U}_{ij} + \min(f^i_x,0)\mathcal{U}_{i+1,j}\right)\right] + \left[k_i\left(\max(f^j_y,0)\mathcal{U}_{ij} + \min(f^j_y,0)\mathcal{U}_{i,j+1}\right)\right] \\
&\quad -\left[\gamma_j\left(\max(f^{i-1}_x,0)\mathcal{U}_{i-1,j} + \min(f^{i-1}_x,0)\mathcal{U}_{ij}\right)\right] \\
&\quad -\left[k_i\left(\max(f^{j-1}_\nu,0)\mathcal{U}_{i,j-1} + \min(f^{j-1}_\nu,0)\mathcal{U}_{ij}\right)\right] \\
&= -\gamma_j\max(f^{i-1}_x,0)\mathcal{U}_{i-1,j} - k_i\max(f^{j-1}_y,0)\mathcal{U}_{i,j-1} \\
&\quad +\left[\gamma_j\left(\max(f^i_x,0) - \min(f^{i-1}_x,0)\right) + k_i\left(\max(f^j_y,0) - \min(f^{j-1}_y,0)\right)\right]\mathcal{U}_{ij} \\
&\quad +k_i\min(f^j_y,0)\mathcal{U}_{i,j+1} + \gamma_j\min(f^i_x,0)\mathcal{U}_{i+1,j}
\end{aligned}
$$

$$
I^{ij} = \epsilon_{ij}\mathcal{U}_{i-1,j} + \mu_{ij}\mathcal{U}_{i,j-1} + \Omega_{ij}\mathcal{U}_{ij} + \phi_{ij}\mathcal{U}_{i,j+1} + \Psi_{ij}\mathcal{U}_{i+1,j}, \tag{3.58}
$$

where

$$
\epsilon_{ij} = -\gamma_j\max(f^{i-1}_x,0), \qquad \mu_{ij} = -k_i\max(f^{j-1}_y,0),
$$

$$
\Omega_{ij} = \gamma_j\left(\max(f^i_x,0) - \min(f^{i-1}_x,0)\right), +k_i\left(\max(f^j_y,0) - \min(f^{j-1}_y,0)\right),
$$

$$
\phi_{ij} = k_i\min(f^j_y,0), \qquad \Psi_{ij} = \gamma_j\min(f^i_x,0).
$$

Equation (3.58) leads to a system of equations which can be written as follows:

$$
I = A_{up}\mathcal{U} + F_{up}, \tag{3.59}
$$

where A_{up} is a $N \times N$ matrix

$$
I = \begin{bmatrix} I^{11} \\ I^{12} \\ \vdots \\ I^{1N} \\ I^{21} \\ I^{22} \\ \vdots \\ \vdots \\ I^{NN} \end{bmatrix}, \qquad \mathcal{U} = \begin{bmatrix} \mathcal{U}_{11} \\ \mathcal{U}_{12} \\ \vdots \\ \mathcal{U}_{1N} \\ \mathcal{U}_{21} \\ \mathcal{U}_{22} \\ \vdots \\ \vdots \\ \mathcal{U}_{NN} \end{bmatrix}, \; A_{up} = \begin{bmatrix} H_1 & P_1 & 0_N & \dots & \dots & \dots & 0_N \\ Q_2 & H_2 & P_2 & \ddots & & & \vdots \\ 0_N & Q_3 & H_3 & P_3 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & Q_{N-2} & H_{N-2} & P_{N-2} & 0_N \\ \vdots & & & \ddots & Q_{N-1} & H_{N-1} & P_{N-1} \\ 0_N & \dots & \dots & \dots & 0_N & Q_N & H_N \end{bmatrix},
$$

with 0_N is $N \times N$ zeros matrix, H_i is a tridiagonal matrix, P_i, Q_i are diagonal matrices and F_{up} is a vector coming from the boundary conditions. Therefore, combining the $\mathbf{O-MPFA}$ method (3.41) and the first order upwind (3.59), we get

$$\frac{d\mathcal{U}}{d\tau} = A\mathcal{U} + F, \tag{3.60}$$

with

$$A = L^{-1}\Big(A_{mp} + A_{up} + A_L\Big) \qquad F = L^{-1}\Big(F_{mp} + F_{up}\Big),$$

where A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (3.10). The diagonal elements of A_L are $A_{ii} = h_i l_i \lambda$ with λ given in (3.4). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = h_i l_i$ for $i = 1, \ldots, N^2$.

3.3.2 Second order upwind method

For computing the different integral above, we are going to use the upwind method to approximate the term $(f\mathcal{U}).\vec{n}$, afterwards we will integrate this approximation over the corresponding face on the boundary of the control volume $\mathcal{C}_{ij}$.

Integral over the eastern face of the control volume $\mathcal{C}_{ij}$ for $i, j = 3, ..., N-2$

The normal outward vector to the eastern face of the control volume $\mathcal{C}_{ij}$ is given by

$$\vec{n}_{\mathcal{E}} = \begin{bmatrix} 1 \\ \\ 0 \end{bmatrix}. \tag{3.61}$$

Thereby, we have

$$f \cdot \vec{n_{\mathcal{E}}} = f_x = (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y)x.$$

Then, according the upwind method we have:

$$f_{\mathcal{E}} = (f.\vec{n}_{\mathcal{E}})\mathcal{U} = \begin{cases} f_x \frac{3\mathcal{U}_{ij} - \mathcal{U}_{i-1,j}}{2} & if \quad f_x \geq 0, \\ \\ f_x \frac{3\mathcal{U}_{i+1,j} - \mathcal{U}_{i+2,j}}{2} & if \quad f_x < 0. \end{cases} \tag{3.62}$$

Let us set

$$\mathcal{U}_{\mathcal{E}} = \begin{cases} \frac{3\mathcal{U}_{ij} - \mathcal{U}_{i-1,j}}{2} & if \quad f_x \geq 0, \\ \\ \frac{3\mathcal{U}_{i+1,j} - \mathcal{U}_{i+2,j}}{2} & if \quad f_x < 0. \end{cases} \tag{3.63}$$

Thus the integral $J_{\mathcal{E}}^{ij}$ over the eastern face of the control volume $\mathcal{C}_{ij}$ is given by

$$
\begin{aligned}
J_{\mathcal{E}}^{ij} &= \int_{\mathcal{E}_{ij}} (f\mathcal{U}) \cdot \vec{n_{\mathcal{E}}} dy \\
&= \int_{\mathcal{E}_{ij}} f_{\mathcal{E}} dy \\
&= \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}} dy \\
&= (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} dy \\
J_{\mathcal{E}}^{ij} &= \gamma_j (r - \sigma_x^2 - \frac{1}{2}\rho_{xy}\sigma_x\sigma_y) x_{i+\frac{1}{2}} \mathcal{U}_{\mathcal{E}}.
\end{aligned}
$$

Hence, we get

$$
J_{\mathcal{E}}^{ij} = \gamma_j \left[\frac{3}{2} \max(f_x^{i+1}, 0) \mathcal{U}_{ij} - \frac{1}{2} \max(f_x^{i+1}, 0) \mathcal{U}_{i-1,j} + \frac{3}{2} \min(f_x^{i+1}, 0) \mathcal{U}_{i+1,j} - \frac{1}{2} \min(f_x^{i+1}, 0) \mathcal{U}_{i+2,j} \right], \tag{3.64}
$$

with

$$
f_x^i = (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y) x_{i+\frac{1}{2}}.
$$

Integral over the western edge of the control volume $\mathcal{C}_{ij}$
for $i, j = 2, ..., N-1$

The normal outward vector to the western edge of the control volume $\mathcal{C}_{ij}$ is

$$
\vec{n}_{\mathcal{W}} = \begin{bmatrix} -1 \\ \\ 0 \end{bmatrix} = -\vec{n}_{\mathcal{E}}. \tag{3.65}
$$

Using a similar approach to the case of the eastern edge of the control volume $\mathcal{C}_{ij}$, we get:

$$
J_{\mathcal{W}}^{ij} = \gamma_j \left[\frac{3}{2} \max(f_x^i, 0) \mathcal{U}_{i-1,j} - \frac{1}{2} \max(f_x^i, 0) \mathcal{U}_{i-2,j} + \frac{3}{2} \min(f_x^i, 0) \mathcal{U}_{ij} - \frac{1}{2} \min(f_x^i, 0) \mathcal{U}_{i+1,j} \right], \tag{3.66}
$$

with

$$
f_x^{i-1} = (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y) x_{i-\frac{1}{2}}.
$$

Integral over the northern edge of the control volume $\mathcal{C}_{ij}$
for $i, j = 2, ..., N-1$

The normal outward vector to the northern edge of the control volume $\mathcal{C}_{ij}$ is

$$
\vec{n}_{\mathcal{N}} = \begin{bmatrix} .0 \\ \\ 1 \end{bmatrix}. \tag{3.67}
$$

Thereby, we have

$$
f \cdot \vec{n_{\mathcal{N}}} = f_y = (r - \sigma_y^2 - \frac{1}{2}\rho\sigma_x\sigma_y) y.
$$

Then, according the upwind method we have:

$$f_{\mathcal{N}} = (f.\vec{n}_{\mathcal{N}})\mathcal{U} = \begin{cases} f_y \mathcal{U}_{ij} & if \quad f_y \geq 0, \\ \\ f_y \mathcal{U}_{i,j+1} & if \quad f_y < 0. \end{cases} \tag{3.68}$$

Let us set

$$\mathcal{U}_{\mathcal{N}} = \begin{cases} \frac{3\mathcal{U}_{ij} - \mathcal{U}_{i,j-1}}{2} & if \quad f_y \geq 0, \\ \\ \frac{3\mathcal{U}_{i,j+1} - \mathcal{U}_{i,j+2}}{2} & if \quad f_y < 0. \end{cases} \tag{3.69}$$

Thus the integral $J_{\mathcal{N}}^{ij}$ over the northern edge of the control volume $\mathcal{C}_{ij}$ is

$$\begin{aligned} J_{\mathcal{N}}^{ij} &= \int_{\mathcal{N}_{ij}} (f\mathcal{U}) \cdot \vec{n_{\mathcal{N}}} dx \\ &= \int_{\mathcal{N}_{ij}} f_{\mathcal{N}} dx \\ &= \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (r - \sigma_y^2 - \frac{1}{2}\rho\sigma_x\sigma_y) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}} dx \\ &= (r - \sigma_y^2 - \frac{1}{2}\rho\sigma_x\sigma_y) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \\ J_{\mathcal{N}}^{ij} &= k_i (r - \sigma_y^2 - \frac{1}{2}\rho_{xy}\sigma_x\sigma_y) y_{j+\frac{1}{2}} \mathcal{U}_{\mathcal{N}}. \end{aligned}$$

Hence

$$J_{\mathcal{N}}^{ij} = k_i \left[\frac{3}{2} \max(f_y^{j+1}, 0)\mathcal{U}_{ij} - \frac{1}{2} \max(f_y^{j+1}, 0)\mathcal{U}_{i,j-1,} + \frac{3}{2} \min(f_y^{j+1}, 0)\mathcal{U}_{i,j+1} - \frac{1}{2} \min(f_y^{j+1}, 0)\mathcal{U}_{i,j+2} \right], \tag{3.70}$$

with

$$f_y^j = (r - \sigma_y^2 - \frac{1}{2}\rho\sigma_x\sigma_y) y_{j+\frac{1}{2}}.$$

Integral over the southern edge of the control volume $\mathcal{C}_{ij}$
$for \quad i, j = 2, ..., N-1$

The normal outward vector to the southern edge of the control volume $\mathcal{C}_{ij}$ is given by

$$\vec{n}_S = \begin{bmatrix} 0 \\ \\ -1 \end{bmatrix} = -\vec{n}_{\mathcal{N}}. \tag{3.71}$$

Using a similar approach as in the case of the northern edge of the control volume $\mathcal{C}_{ij}$, we get:

$$J_S^{ij} = k_i \left[\frac{3}{2} \max(f_y^j, 0)\mathcal{U}_{i,j-1} - \frac{1}{2} \max(f_y^j, 0)\mathcal{U}_{i,j-2} + \frac{3}{2} \min(f_y^j, 0)\mathcal{U}_{ij} - \frac{1}{2} \min(f_y^j, 0)\mathcal{U}_{i,j+1} \right], \tag{3.72}$$

with

$$f_y^{j-1} = (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y) y_{j-\frac{1}{2}}.$$

Therefore, the outflux through the edges of the control volume $\mathcal{C}_{ij}$, for $i,j = 1,\ldots,N$ is given by

$$\begin{aligned}
J^{ij} &= I^{ij}_{\mathcal{E}} + I^{ij}_{\mathcal{N}} - I^{ij}_{\mathcal{W}} - I^{ij}_{S} \\
&= \left[\gamma_j\left(\frac{3}{2}\max(f_x^{i+1},0)\mathcal{U}_{ij} - \frac{1}{2}\max(f_x^{i+1},0)\mathcal{U}_{i-1,j} + \frac{3}{2}\min(f_x^{i+1},0)\mathcal{U}_{i+1,j} - \frac{1}{2}\min(f_x^{i+1},0)\mathcal{U}_{i+2,j}\right)\right] \\
&\quad + \left[h\left(\frac{3}{2}\max(f_y^{j+1},0)\mathcal{U}_{ij} - \frac{1}{2}\max(f_y^{j+1},0)\mathcal{U}_{i,j-1,} + \frac{3}{2}\min(f_y^{j+1},0)\mathcal{U}_{i,j+1} - \frac{1}{2}\min(f_y^{j+1},0)\mathcal{U}_{i,j+2}\right)\right] \\
&\quad - \left[\gamma_j\left(\frac{3}{2}\max(f_x^{i},0)\mathcal{U}_{i-1,j} - \frac{1}{2}\max(f_x^{i},0)\mathcal{U}_{i-2,j} + \frac{3}{2}\min(f_x^{i},0)\mathcal{U}_{ij} - \frac{1}{2}\min(f_x^{i},0)\mathcal{U}_{i+1,j}\right)\right] \\
&\quad - \left[k_i\left(\frac{3}{2}\max(f_y^{j},0)\mathcal{U}_{i,j-1} - \frac{1}{2}\max(f_y^{j},0)\mathcal{U}_{i,j-2} + \frac{3}{2}\min(f_y^{j},0)\mathcal{U}_{ij} - \frac{1}{2}\min(f_y^{j},0)\mathcal{U}_{i,j+1}\right)\right]
\end{aligned}$$

$$\begin{aligned}
J^{ij} &= \epsilon_{ij}\mathcal{U}_{i-2,j} + \eta_{ij}\mathcal{U}i-1,j + \kappa_{ij}\mathcal{U}_{i,j-2} + \mu_{ij}\mathcal{U}_{i,j-1} + \Omega_{ij}\mathcal{U}_{ij} + \phi_{ij}\mathcal{U}_{i,j+1} + \Psi_{ij}\mathcal{U}_{i,j+2} + \Delta_{ij}\mathcal{U}_{i+1,j} \\
&\quad + \Pi_{ij}\mathcal{U}_{i+2,j}, \qquad (3.73)
\end{aligned}$$

where

$$\epsilon_{ij} = \frac{1}{2}\gamma_j\max(f_x^i,0), \qquad \eta_{ij} = -\frac{1}{2}\gamma_j\max(f_x^{i+1},0) - \frac{3}{2}\gamma_j\max(f_x^i,0),$$

$$\kappa_{ij} = +\frac{1}{2}k_i\max(f_y^j,0), \qquad \mu_{ij} = -\frac{1}{2}k_i\max(f_y^{j+1},0) - \frac{3}{2}k_i\max(f_y^j,0),$$

$$\Omega_{ij} = \frac{3}{2}\gamma_j\max(f_x^{i+1},0) + \frac{3}{2}k_i\max(f_y^{j+1},0) - \frac{3}{2}\gamma_j\min(f_x^i,0) - \frac{3}{2}k_i\min(f_y^j,0),$$

$$\phi_{ij} = \frac{3}{2}k_i\min(f_y^{j+1},0) + \frac{1}{2}k_i\min(f_y^j,0), \qquad \Psi_{ij} = -\frac{1}{2}k_i\min(f_y^{j+1},0),$$

$$\Delta_{ij} = \frac{3}{2}\gamma_j\min(f_x^{i+1},0) + \frac{1}{2}\gamma_j\min(f_x^i,0), \qquad \Pi_{ij} = -\frac{1}{2}\gamma_j\min(f_x^{i+1},0).$$

For the control volumes near the boundary of the study domain, the first order upwind method is used for the approximation of the flux through edges directly connected to the boundary. Equation (3.73) leads to a system of equations which can be written as:

$$J = A_{2up}\mathcal{U} + F_{2up}, \qquad (3.74)$$

where

$$J = \begin{bmatrix} J^{11} \\ J^{12} \\ \vdots \\ J^{1N} \\ J^{21} \\ J^{22} \\ \vdots \\ J^{NN} \end{bmatrix}, \qquad \mathcal{U} = \begin{bmatrix} \mathcal{U}_{11} \\ \mathcal{U}_{12} \\ \vdots \\ \mathcal{U}_{1N} \\ \mathcal{U}_{21} \\ \mathcal{U}_{22} \\ \vdots \\ \mathcal{U}_{NN} \end{bmatrix}, \qquad A_{2up} = \begin{bmatrix} K_1 & R_1 & G_1 & 0_N & \dots & \dots & 0_N \\ S_2 & K_2 & R_2 & G_2 & \ddots & & \vdots \\ H_3 & S_3 & K_3 & R_3 & G_3 & \ddots & \vdots \\ 0_N & \ddots & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & \ddots & H_{N-2} & S_{N-2} & K_{N-2} & R_{N-2} & G_{N-2} \\ \vdots & & \ddots & H_{N-1} & S_{N-1} & K_{N-1} & R_{N-1} \\ 0_N & \dots & \dots & 0_N & H_N & S_N & K_N \end{bmatrix},$$

where K_i is penta-diagonal matrice and R_i, G_i, S_i, k_i are diagonal matrices. F_{2up} is a vector coming from the boundary conditions.

Therefore, combining the $\mathbf{O-MPFA}$ method (3.41) and the second order upwind method (3.73), we have

$$\frac{d\mathcal{U}}{d\tau} = A\mathcal{U} + F, \tag{3.75}$$

with

$$A = L^{-1}\Big(A_{mp} + A_{2up} + A_L\Big), \qquad F = L^{-1}\Big(F_{mp} + F_{2up}\Big),$$

where A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (3.10). The diagonal elements of A_L are $A_{ii} = k_i l_i \lambda$, with λ given in (3.4). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = k_i l_i$ for $i = 1, \dots, N^2$.

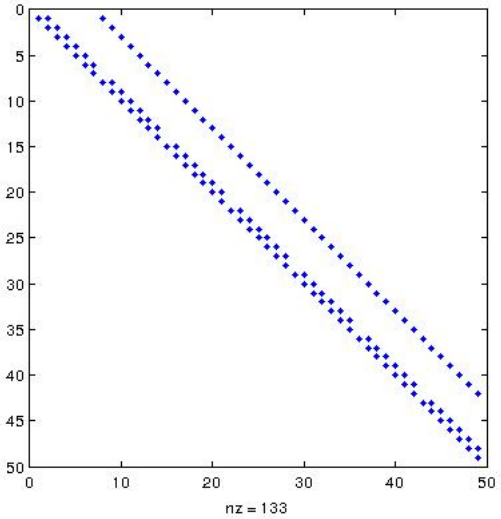


Figure 3.17: Structure of the advection matrix using the $2^{nd} order$ upwind

As we said previously, when the stock price approaches zero, the ellipcity conidtion (1.29) is not satisfied. The PDE is then degenerate. To overcome the degeneracy, we apply the fiited finite volume method in the degeneracy region $\mathcal{D}_{\mathcal{R}}$ defined as:

$$\mathcal{D}_{\mathcal{R}} = \Big([0, x_1] \times [0, y_{\max}]\Big) \bigcup \Big([0, x_{\max}] \times [0, y_1]\Big). \tag{3.76}$$

3.4 Fitted Multi-Point Flux Approximation

The fitted Multi-Point Flux Approximation is a combination of the fitted finite volume method (see Huang et al. [2006, 2009]) and the Multi-Point Flux Approximation method. The fitted finite volume helps to deal with the degeneracy of the PDE (3.2). We approximate simultaneously the diffusion term and the convection term in the degeneracy region by solving a two-points boundary problem. In the region where the PDE is not degenerated, we apply the standard Multi-point flux approximation to the diffusion term as described in the previous Section.
Let us set

$$k(\mathcal{U}) = \nabla \cdot (\mathbf{M}\nabla\mathcal{U} + f\mathcal{U}), \tag{3.77}$$

where $\mathbf{M}$ and f are defined in (3.4). Thereby, we have the following decomposition over a control volume $\mathcal{C}_{ij}$, for $i, j = 1, ..., N$

$$\begin{aligned}\int_{\mathcal{C}_{ij}} \nabla k(\mathcal{U}) d\mathcal{C} &= \int_{\mathcal{C}_{ij}} \nabla \cdot (M\nabla\mathcal{U} + f\mathcal{U}) d\mathcal{C} \\ &= \int_{\partial\mathcal{C}_{ij}} (M\nabla\mathcal{U} + f\mathcal{U}) \cdot \vec{n} d\partial\mathcal{C}\end{aligned}$$

$$\begin{aligned}\int_{\mathcal{C}_{ij}} \nabla k(\mathcal{U}) d\mathcal{C} &= \int_{(x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U} \right) dy \\ &\quad - \int_{(x_{i-\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U} \right) dy \\ &\quad + \int_{(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{21}\frac{\partial\mathcal{U}}{\partial x} + m_{22}\frac{\partial\mathcal{U}}{\partial y} + q\mathcal{U} \right) dx \\ &\quad - \int_{(x_{i-\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}})} \left(m_{21}\frac{\partial\mathcal{U}}{\partial x} + m_{22}\frac{\partial\mathcal{U}}{\partial y} + q\mathcal{U} \right) dx.\end{aligned} \tag{3.78}$$

with $\vec{n}$ is the outward unit normal vector, $m_{11}, m_{12}, m_{21}, m_{22}$ the coefficients of the matrix $\mathbf{M}$ and p, q coefficients of vector f defined in (3.4).
In their work, Huang et al. [2006, 2009] showed how the fitted finite volume method is used to approximate each of the integral in (3.78).

3.4.1 Fitted Finite volume method in the degeneracy region

Following Huang et al. [2006], the fitted finite volume method is used to approximate the flux through the edges which are effectively in the degeneracy region especially the western edge of the control volume $\mathcal{C}_{1,j}$ for $j = 1, \ldots, N$ and the southern edge of the control volume $\mathcal{C}_{i,1}$ for $i = 1, \ldots, N$.

Flux through the western edge of a control volume $\mathcal{C}_{1,j}$ $\quad for j = 1, \ldots, N$.

By applying the mid-quadrature rule, we have :

$$\int_{(x_{\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U} \right) dy \approx \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U} \right)_{(x_{\frac{1}{2}}, y_j)} \cdot l_j.$$

Besides, we have

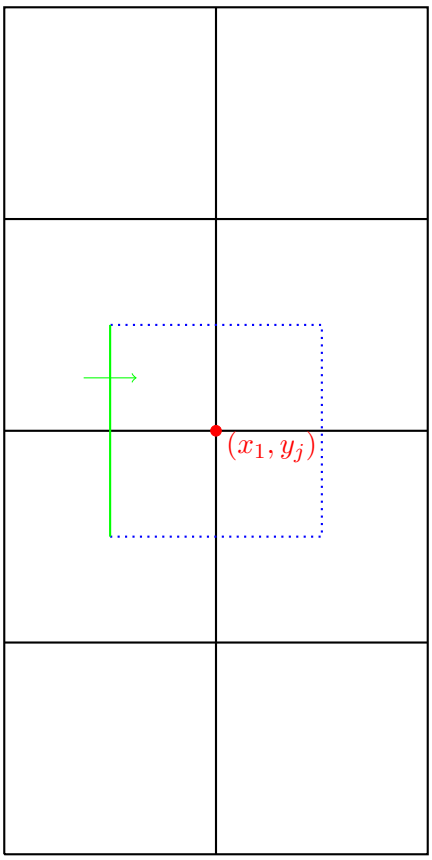


Figure 3.18: Eastern edge in the degeneracy region

$$
\begin{aligned}
m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} &= \frac{1}{2}\sigma_1^2 x^2 \frac{\partial \mathcal{U}}{\partial x} + \frac{1}{2}\rho\sigma_2\sigma_2 xy\frac{\partial \mathcal{U}}{\partial y} + (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)x\mathcal{U} \\
&= x\left(\frac{1}{2}\sigma^2 x\frac{\partial \mathcal{U}}{\partial x} + \frac{1}{2}\rho\sigma_2\sigma_2 y\frac{\partial \mathcal{U}}{\partial y} + (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)\mathcal{U}\right) \\
m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} &= x\left(ax\frac{\partial \mathcal{U}}{\partial x} + d\frac{\partial \mathcal{U}}{\partial y} + b\mathcal{U}\right),
\end{aligned}
$$

with $a = \frac{1}{2}\sigma_1^2$, $b = r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2$ and $d = \frac{1}{2}\rho\sigma_1\sigma_2 y$.

We want to approximate

$$
g(\mathcal{U}) = ax\frac{\partial \mathcal{U}}{\partial x} + b\mathcal{U}, \tag{3.79}
$$

by a constant over $I_{x_1} = (0, x_1)$ satisfying the following two-point boundary value problem

$$
\begin{cases}
g'(v) = \left(ax\frac{\partial v}{\partial x} + bv\right)' = K_1, \\
\\
v(0, y_j) = \mathcal{U}_{0,j} \qquad v(x_1, y_j) = \mathcal{U}_{1,j}.
\end{cases} \tag{3.80}
$$

By solving this problem, we obtain

$$\mathcal{U} = \mathcal{U}_{0,j} + (\mathcal{U}_{1,j} - \mathcal{U}_{0,j})\frac{x}{x_1}.$$

Thereby

$$\begin{aligned}
\left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U}\right)_{(x_{\frac{1}{2}}, y_j)} \cdot l_j &= x_{\frac{1}{2}} l_j \left(a(\mathcal{U}_{1,j} - \mathcal{U}_{0,j})\frac{x_{\frac{1}{2}}}{x_1} + b\left(\mathcal{U}_{0,j} + (\mathcal{U}_{1,j} - \mathcal{U}_{0,j})\frac{x_{\frac{1}{2}}}{x_1}\right) \right. \\
&\quad \left. + d\frac{\mathcal{U}_{1,j+1} - \mathcal{U}_{1,j}}{a} l_j \right) \\
&= x_{\frac{1}{2}} l_j \left(\frac{x_{\frac{1}{2}}}{x_1}\Big[(a+b)\mathcal{U}_{1,j} - (a-b)\mathcal{U}_{0,j})\Big] \right) + d_j \frac{\mathcal{U}_{1,j+1} - \mathcal{U}_{1,j}}{l_1} \Bigg) \\
&= \frac{x_0 + x_1}{2} l_1 \left(\frac{x_0 + x_1}{2}\frac{1}{x_1}\Big[(a+b)\mathcal{U}_{1,j} - (a-b)\mathcal{U}_{0,j})\Big] \right) \\
&\quad + d_j \frac{\mathcal{U}_{1,j+1} - \mathcal{U}_{1,j}}{l_j} \Bigg) \\
&= \frac{1}{2} x_1 l_j \left(\frac{1}{2}\Big[(a+b)\mathcal{U}_{1,j} - (a-b)\mathcal{U}_{0,j})\Big] + d_j \frac{\mathcal{U}_{1,j+1} - \mathcal{U}_{1,j}}{l_j} \right) \\
&= \frac{1}{2} x_1 l_j \left(\Big[\frac{1}{2}(a+b) - \frac{d}{l_j}\Big]\mathcal{U}_{1,j} + \frac{d_j}{l_j}\mathcal{U}_{1,j+1} - \frac{1}{2}(a-b)\mathcal{U}_{0,j} \right) \\
\left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U}\right)_{(x_{\frac{1}{2}}, y_j)} \cdot l_j &= \frac{1}{2} x_1 \Big[\frac{1}{2} l_j (a+b) - d_j\Big]\mathcal{U}_{1,j} + \frac{1}{2} d_j x_1 \mathcal{U}_{1,j+1} - \frac{1}{4} l_j x_1 (a-b)\mathcal{U}_{0,j}.
\end{aligned} \tag{3.81}$$

Flux through the southern edge of a control volume $\mathcal{C}_{i,1}$, $\quad i = 1, \ldots, N$

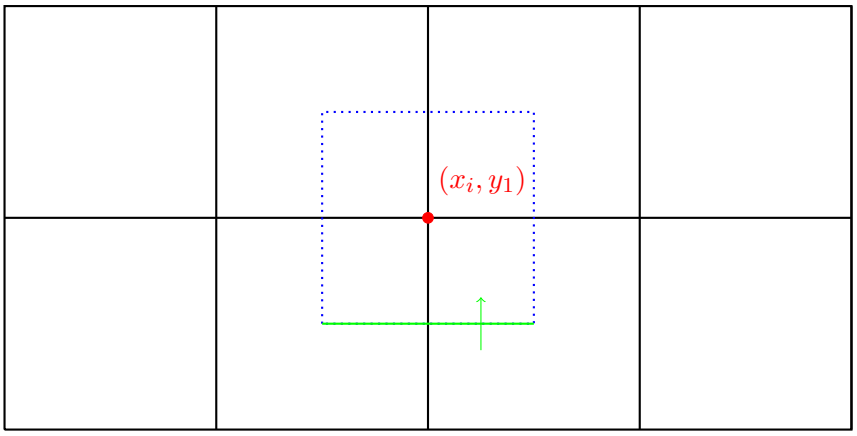


Figure 3.19: Southern edge in a degeneracy region

By applying the mid-quadrature rule, we get:

$$\int_{(x_{i-\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{\frac{1}{2}})} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U}\right) dx \approx \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U}\right)_{(x_i, y_{\frac{1}{2}})} \cdot h_i.$$

Besides, we have

$$\begin{aligned} m_{21}\frac{\partial \mathcal{U}}{\partial x}+m_{22}\frac{\partial \mathcal{U}}{\partial y}+q\mathcal{U} &= \frac{1}{2}\rho\sigma_1\sigma_2 xy\frac{\partial \mathcal{U}}{\partial x}+\frac{1}{2}\sigma_2^2 y^2\frac{\partial \mathcal{U}}{\partial y}+(r-\sigma_2^2-\frac{1}{2}\rho\sigma_1\sigma_2)y\mathcal{U} \\ &= y\left(\frac{1}{2}\sigma_2^2 y\frac{\partial \mathcal{U}}{\partial y}+\frac{1}{2}\rho\sigma_1\sigma_2 x\frac{\partial \mathcal{U}}{\partial x}+(r-\sigma_2^2-\frac{1}{2}\rho\sigma_1\sigma_2)\mathcal{U}\right) \\ m_{21}\frac{\partial \mathcal{U}}{\partial x}+m_{22}\frac{\partial \mathcal{U}}{\partial y}+q\mathcal{U} &= y\left(ey\frac{\partial \mathcal{U}}{\partial y}+h\frac{\partial \mathcal{U}}{\partial x}+k\mathcal{U}\right), \end{aligned}$$

with $e=\frac{1}{2}\sigma_2^2$, $k=r-\sigma_2^2-\frac{1}{2}\rho\sigma_1\sigma_2$ and $h'=\frac{1}{2}\rho\sigma_1\sigma_2 x$.
We want to approximate

$$f(\mathcal{U})=ey\frac{\partial \mathcal{U}}{\partial y}+k\mathcal{U},$$

by a constant over $I_{y_i}=(y_i,y_{i+1})$ satisfying the following two-point boundary value problem

$$\begin{cases} f'(v) \quad = \left(ey\frac{\partial v}{\partial y}+kv\right)' = K_1, \\ \\ v(x_i,0)=\mathcal{U}_{i,0} \qquad\qquad v(x_i,y_1)=\mathcal{U}_{i,1}. \end{cases} \tag{3.82}$$

By solving this problem, we obtain

$$\mathcal{U}=\mathcal{U}_{i,0}+(\mathcal{U}_{i,1}-\mathcal{U}_{i,0})\frac{y}{y_1}.$$

Thereby

$$\begin{aligned} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x}+m_{22}\frac{\partial \mathcal{U}}{\partial y}+q\mathcal{U}\right)_{(x_i,y_{\frac{1}{2}})}\cdot h_i &= y_{\frac{1}{2}}h_i\left(e(\mathcal{U}_{i,1}-\mathcal{U}_{i,0})\frac{y_{\frac{1}{2}}}{y_1}+k\left(\mathcal{U}_{i,0}+(\mathcal{U}_{i,1}-\mathcal{U}_{i,0})\frac{y_{\frac{1}{2}}}{y_1}\right)\right. \\ &\quad \left.+h_i'\frac{\mathcal{U}_{i+1,1}-\mathcal{U}_{i,1}}{h_i}\right) \\ &= y_{\frac{1}{2}}h_i\left(\frac{y_{\frac{1}{2}}}{y_1}\Big[(e+k)\mathcal{U}_{i,1}-(e-k)\mathcal{U}_{i,0}\Big]\right)+h_i'\frac{\mathcal{U}_{i+1,1}-\mathcal{U}_{i,1}}{h_i}\Bigg) \\ &= \frac{y_0+y_1}{2}h_i\left(\frac{y_0+y_1}{2}\frac{1}{y_1}\Big[(e+k)\mathcal{U}_{i,1}-(e-k)\mathcal{U}_{i,0}\Big]\right) \\ &\quad +h_i'\frac{\mathcal{U}_{i+1,1}-\mathcal{U}_{i,1}}{h_i}\Bigg). \end{aligned}$$

Then, we get

$$\left(m_{21}\frac{\partial \mathcal{U}}{\partial x}+m_{22}\frac{\partial \mathcal{U}}{\partial y}+q\mathcal{U}\right)_{(x_i,y_{\frac{1}{2}})}\cdot h_i=\frac{1}{2}y_1\Big[\frac{1}{2}h_i(e+k)-h_i'\Big]\mathcal{U}_{i,1}+\frac{1}{2}h_i'y_1\mathcal{U}_{i+1,1}-\frac{1}{4}y_1h_i(e-k)\mathcal{U}_{i,0}. \tag{3.83}$$

Furthermore, the flux through edges which are not in degeneracy region will be approximated using the MPFA method for the diffusion term and the upwind method for the advection term. Thereby, using (3.37) and (3.47), the flux through the eastern edge of a control volume $\mathcal{C}_{1,j},\ j=1,\ldots,N$, is given by

$$\begin{aligned}\int_{(x_{\frac{3}{2}},y_{j-\frac{1}{2}})}^{(x_{\frac{3}{2}},y_{j+\frac{1}{2}})}\left(m_{11}\frac{\partial\mathcal{U}}{\partial x}+m_{12}\frac{\partial\mathcal{U}}{\partial y}+p\mathcal{U}\right)dy \;\approx\; & \Big(T_{11}^{2,j+1}+T_{23}^{2,j}+\gamma_j\max(f_x^2,0)\Big)\mathcal{U}_{1,j}+\Big(T_{12}^{2,j+1}+T_{24}^{2,j}\\ & +\gamma_j\min(f_x^2,0)\Big)\mathcal{U}_{2,j}+T_{14}^{2,j+1}\mathcal{U}_{2,j+1}+T_{13}^{2,j+1}\mathcal{U}_{1,j+1}\\ & +T_{21}^{2,j}\mathcal{U}_{1,j-1}+T_{22}^{2,j}\mathcal{U}_{2,j-1}. \qquad (3.84)\end{aligned}$$

Using (3.38) and (3.54), the flux through the northern edge of a control volume $\mathcal{C}_{1,j},\ i=1,\ldots,N$ is given by

$$\begin{aligned}\int_{(x_{\frac{1}{2}},y_{j+\frac{1}{2}})}^{(x_{\frac{3}{2}},y_{j+\frac{1}{2}})}\left(m_{11}\frac{\partial\mathcal{U}}{\partial x}+m_{12}\frac{\partial\mathcal{U}}{\partial y}+p\mathcal{U}\right)dy \;\approx\; & \Big(T_{31}^{2,j+1}+T_{42}^{1,j+1}+k_1\max(f_y^{j+1},0)\Big)\mathcal{U}_{1,j}+T_{32}^{2,j+1}\mathcal{U}_{2,j}\\ & +T_{34}^{2,j+1}\mathcal{U}_{2,j+1}+\Big(T_{33}^{2,j+1}+T_{44}^{1,j+1}+k_1\min(f_y^{j+1},0)\Big)\mathcal{U}_{1,j+1}\\ & +T_{43}^{1,j+1}\mathcal{U}_{0,j+1}+T_{41}^{1,j+1}\mathcal{U}_{0,j}. \qquad (3.85)\end{aligned}$$

The flux through the southern edge of the control volume $\mathcal{C}_{1,j}\ , j=1,\ldots,N$, using (3.40) and (3.57), is given by:

$$\begin{aligned}\int_{(x_{\frac{1}{2}},y_{j-\frac{1}{2}})}^{(x_{\frac{3}{2}},y_{j-\frac{1}{2}})}\left(m_{21}\frac{\partial\mathcal{U}}{\partial x}+m_{22}\frac{\partial\mathcal{U}}{\partial y}+p\mathcal{U}\right)dy \;\approx\; & \Big(T_{33}^{2,j}+T_{44}^{1,j}+k_1\min(f_y^{j},0)\Big)\mathcal{U}_{1,j}+T_{34}^{2,j}\mathcal{U}_{2,j}\\ & +\Big(T_{31}^{2,j}+T_{42}^{1,j}+k_1\max(f_y^{j},0)\Big)\mathcal{U}_{1,j-1}+T_{32}^{2,j}\mathcal{U}_{2,j-1}\\ & +T_{43}^{1,j}\mathcal{U}_{0,j}+T_{41}^{1,j}\mathcal{U}_{0,j-1}. \qquad (3.86)\end{aligned}$$

Similarly, using (3.38) and (3.54), the flux through the northern edge of a control volume $\mathcal{C}_{i,1}$, $i=1,\ldots,N$ is given by

$$\begin{aligned}\int_{(x_{i-\frac{1}{2}},y_{\frac{3}{2}})}^{(x_{i+\frac{1}{2}},y_{\frac{3}{2}})}\left(m_{21}\frac{\partial\mathcal{U}}{\partial x}+m_{22}\frac{\partial\mathcal{U}}{\partial y}+q\mathcal{U}\right)dx \;\approx\; & \Big(T_{31}^{i+1,2}+T_{42}^{i,2}+k_i\max(f_y^{2},0)\Big)\mathcal{U}_{i,1}+T_{32}^{i+1,2}\mathcal{U}_{i+1,1}\\ & +T_{34}^{i+1,2}\mathcal{U}_{i+1,2}+\Big(T_{33}^{i+1,2}+T_{44}^{i,2}+k_i\min(f_y^{2},0)\Big)\mathcal{U}_{i,2}\\ & +T_{43}^{i,2}\mathcal{U}_{i-1,2}+T_{41}^{i,2}\mathcal{U}_{i-1,1}. \qquad (3.87)\end{aligned}$$

In the same way, using (3.37) and (3.47), the flux through the eastern edge of a control volume $\mathcal{C}_{i,1},\ i=1,\ldots,N$, is given by

$$\int_{(x_{i+\frac{1}{2}},y_{\frac{1}{2}})}^{(x_{i+\frac{1}{2}},y_{\frac{3}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dx \approx \Big(T_{11}^{i+1,2} + T_{23}^{i+1,1} + \gamma_1 \max(f_x^{i+1}, 0)\Big)\mathcal{U}_{i,1}$$
$$+\Big(T_{12}^{i+1,2} + T_{24}^{i+1,1} + \gamma_1 \min(f_x^{i+1}, 0)\Big)\mathcal{U}_{i+1,1}$$
$$+T_{14}^{i+1,2}\mathcal{U}_{i+1,2} + T_{13}^{i+1,2}\mathcal{U}_{i,2} + T_{21}^{i+1,1}\mathcal{U}_{i,0} + T_{22}^{i+1,1}\mathcal{U}_{i+1,0}. \tag{3.88}$$

Finally, the flux through the western edge of the control volume $\mathcal{C}_{i,1}$, $i = 1, \ldots, N$, using (3.39) and (3.50), is given by

$$\int_{(x_{i-\frac{1}{2}},y_{\frac{1}{2}})}^{(x_{i-\frac{1}{2}},y_{\frac{3}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dx \approx \Big(T_{12}^{i,2} + T_{24}^{i,1} + \gamma_1 \min(f_x^{i}, 0)\Big)\mathcal{U}_{i,1} + T_{14}^{i,2}\mathcal{U}_{i,2} + T_{13}^{i,2}\mathcal{U}_{i-1,2}$$
$$+\Big(T_{11}^{i,2} + T_{23}^{i,1} + \gamma_1 \max(f_x^{i}, 0)\Big)\mathcal{U}_{i-1,1} + T_{21}^{i,1}\mathcal{U}_{i-1,0}$$
$$+T_{22}^{i,1}\mathcal{U}_{i,0}. \tag{3.89}$$

Let us recall that we aim to approximate the integral of $k(\mathcal{U})$ (see (3.77)) by following the decomposition of this integral given in (3.78).

3.4.2 Flux through edges of control volume in the degeneracy region

The control volumes $\mathcal{C}_{11}, \mathcal{C}_{1,j}$ $j = 2, \ldots, N$, and $\mathcal{C}_{i,1}$, $i = 2, \ldots, N$, have at least one edge which is fully in the degeneracy region. The flux through the edges of these control volumes will be approximated either by the fitted finite volume method or the combination of the O-MPFA method and the upwind methods (1^{st} and 2^{nd} order).

Fitted MPFA-1^{st} order upwind

we call fitted MPFA- 1^{st} order upwind the combination of the fitted finite volume method and the MPFA method coupled to the 1^{st} order upwind method.

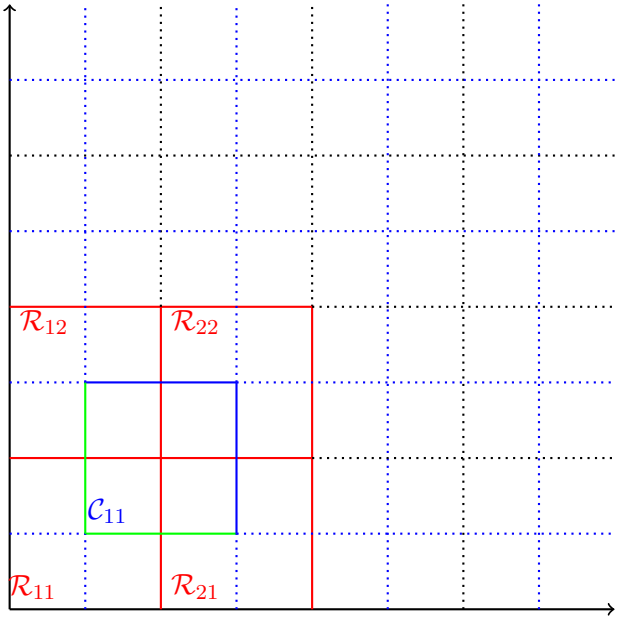


Figure 3.20: Control volume $\mathcal{C}_{11}$

As we can see on Figure 3.20, the western and the southern edge are in the degeneracy region. Thereby, using (3.84), (3.85), (3.81) and (3.83), the outflux through the edges of control volume $\mathcal{C}_{11}$ is given by

$$\begin{aligned}
\int_{\mathcal{C}_{11}} \nabla k(\mathcal{U}) &= \int_{\mathcal{C}_{11}} \nabla \cdot (M\nabla\mathcal{U} + f\mathcal{U}) \\
&= \int_{\mathcal{C}_{11}} (M\nabla\mathcal{U} + f\mathcal{U}) \cdot \vec{n} \\
&= \int_{(x_{\frac{3}{2}}, y_{\frac{1}{2}})}^{(x_{\frac{3}{2}}, y_{\frac{3}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dy + \int_{(x_{\frac{1}{2}}, y_{\frac{3}{2}})}^{(x_{\frac{3}{2}}, y_{\frac{3}{2}})} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U} \right) dx \\
&\quad - \int_{(x_{\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{\frac{1}{2}}, y_{\frac{3}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dy - \int_{(x_{\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{\frac{3}{2}}, y_{\frac{1}{2}})} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U} \right) dx
\end{aligned}$$

$$\int_{\mathcal{C}_{11}} \nabla k(\mathcal{U}) \approx a_{11}^1\mathcal{U}_{11} + b_{11}^1\mathcal{U}_{21} + c_{11}^1\mathcal{U}_{22} + d_{11}^1\mathcal{U}_{12} + \omega_{11}^1\mathcal{U}_{02} + \phi_{11}^1\mathcal{U}_{01} + r_{11}^1\mathcal{U}_{10} + s_{11}^1\mathcal{U}_{20}, \tag{3.90}$$

with

$$\begin{aligned}
a_{11}^1 &= T_{11}^{22} + T_{23}^{21} + T_{31}^{22} + T_{42}^{12} + \gamma_1 \max(f_x^2, 0) + k_1 \max(f_y^2, 0) - \frac{1}{2}x_1\Big[\frac{1}{2}l_1(a+b) - d_1\Big], \\
&\quad -\frac{1}{2}y_1\Big[\frac{1}{2}h_1(e+k) - h_1'\Big], \\
b_{11}^1 &= T_{12}^{22} + T_{24}^{21} + T_{32}^{22} + \gamma_1 \min(f_x^2, 0) - \frac{1}{2}h_1'y_1, \qquad c_{11}^1 = T_{14}^{22} + T_{34}^{22}, \\
d_{11}^1 &= T_{13}^{22} + T_{33}^{22} + T_{44}^{12} + k_1 \min(f_y^2, 0) - \frac{1}{2}d_1x_1, \qquad \omega_{11}^1 = T_{43}^{12}, \\
\phi_{11}^1 &= T_{41}^{12} + \frac{1}{4}\gamma_1x_1(a-b), \qquad r_{11}^1 = T_{21}^{21} + \frac{1}{4}k_1y_1(e-k), \qquad s_{11}^1 = T_{22}^{21}.
\end{aligned}$$

For the control volume $\mathcal{C}_{1,j},\ j = 2, \ldots, N$, only the western edge is in the degeneracy region; thereby

using (3.81), (3.86), (3.84) and (3.85), the outflux through the edges of the control volume $\mathcal{C}_{1,j}$, $j = 2, \ldots, N$, is given by

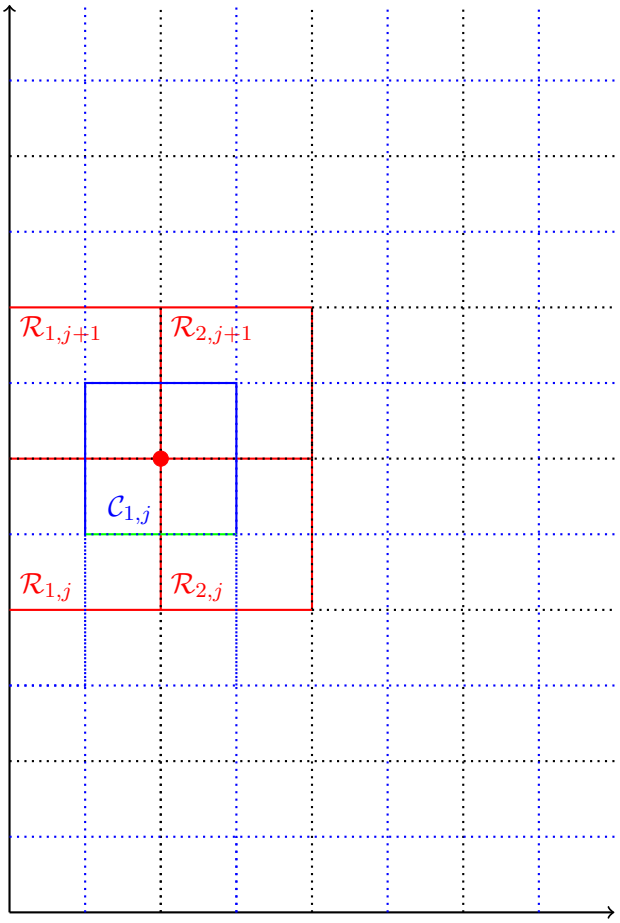


Figure 3.21: Control volume $C_{1,j}$ $j = 2, \ldots, N$

$$
\begin{aligned}
\int_{\mathcal{C}_{1,j}} \nabla k(\mathcal{U}) &= \int_{\mathcal{C}_{1,j}} \nabla \cdot (M\nabla\mathcal{U} + f\mathcal{U}) \\
&= \int_{\mathcal{C}_{1,j}} (M\nabla\mathcal{U} + f\mathcal{U}) \cdot \vec{n} \\
&= \int_{(x_{\frac{3}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{3}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dy + \int_{(x_{\frac{1}{2}}, y_{j+\frac{1}{2}})}^{(x_{\frac{3}{2}}, y_{j+\frac{1}{2}})} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U} \right) dx \\
&\quad - \int_{(x_{\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial \mathcal{U}}{\partial x} + m_{12}\frac{\partial \mathcal{U}}{\partial y} + p\mathcal{U} \right) dy - \int_{(x_{\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{3}{2}}, y_{j-\frac{1}{2}})} \left(m_{21}\frac{\partial \mathcal{U}}{\partial x} + m_{22}\frac{\partial \mathcal{U}}{\partial y} + q\mathcal{U} \right) dx.
\end{aligned}
$$

Thereby,

$$
\begin{aligned}
\int_{\mathcal{C}_{1,j}} \nabla k(\mathcal{U}) \approx\ & a^1_{1,j}\mathcal{U}_{1,j} + bb_{1,j}\mathcal{U}_{2,j} + c^1_{1,j}\mathcal{U}_{2,j+1} + d^1_{1,j}\mathcal{U}_{1,j+1} + \gamma^1_{1,j}\mathcal{U}_{1,j-1} + \lambda^1_{1,j}\mathcal{U}_{2,j-1} \\
& \omega^1_{1,j}\mathcal{U}_{0,j+1} + \phi^1_{1,j}\mathcal{U}_{0,j} + \Upsilon^1_{1,j}\mathcal{U}_{0,j-1}, \qquad (3.91)
\end{aligned}
$$

where

$$a^1_{1,j} = T^{2,j+1}_{11} + T^{2,j}_{23} + T^{2,j+1}_{31} + T^{1,j+1}_{42} - T^{2,j}_{33} - T^{1,j}_{44} - \frac{1}{2}x_1\Big(\frac{1}{2}l_j(a+b) - d_j\Big)$$

$$+l_j \max(f^2_x, 0) + h_1 \max(f^{j+1}_y, 0) - h_1 \min(f^j_y, 0),$$

$$b^1_{1,j} = T^{2,j+1}_{12} + T^{2,j}_{24} + T^{2,j+1}_{32} - T^{2,j}_{34} + l_j \min(f^2_x, 0), \qquad c^1_{1,j} = T^{2,j+1}_{14} + T^{2,j+1}_{34},$$

$$d^1_{1,j} = T^{2,j+1}_{13} + T^{2,j+1}_{33} + T^{1,j+1}_{44} + h_1 \min(f^{j+1}_y, 0) - \frac{1}{2}d_j x_1 +$$

$$\gamma^1_{1,j} = T^{2,j}_{21} - T^{2,j}_{31} - T^{1,j}_{42} - h_1 \max(f^j_y, 0), \qquad \lambda^1_{1,j} = T^{2,j}_{22} - T^{2,j}_{32},$$

$$\omega^1_{1,j} = T^{1,j+1}_{43}, \qquad \phi^1_{1,j} = T^{1,j+1}_{41} - T^{1,j}_{43} + \frac{1}{4}l_j x_1(a-b), \qquad \Upsilon^1_{1,j} = -T^{1,j}_{41}.$$

For the control volume $\mathcal{C}_{i,1},\ i = 2, \ldots, N$, only the southern edge is in degeneracy area; thereby, using (3.83),(3.88),(3.87) and (3.89), the outflux the edges of control volume $\mathcal{C}_{i,1},\ i = 2, \ldots, N$, is given by:

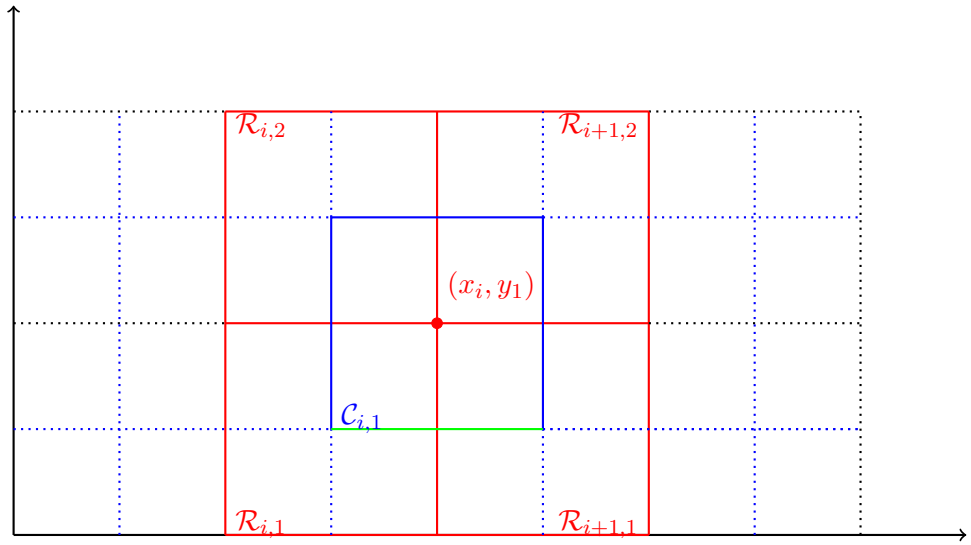


Figure 3.22: Control volume $\mathcal{C}_{i,1}\ i = 2, \ldots, N$

$$\begin{aligned}
\int_{\mathcal{C}_{i,1}} \nabla k(\mathcal{U}) &= \int_{\mathcal{C}_{i,1}} \nabla \cdot (M\nabla\mathcal{U} + f\mathcal{U}) \\
&= \int_{\mathcal{C}_{i,1}} (M\nabla\mathcal{U} + f\mathcal{U}) \cdot \vec{n} \\
&= \int_{(x_{i+\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{\frac{3}{2}})} \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U}\right)dy + \int_{(x_{i-\frac{1}{2}}, y_{\frac{3}{2}})}^{(x_{i+\frac{1}{2}}, y_{\frac{3}{2}})} \left(m_{21}\frac{\partial\mathcal{U}}{\partial x} + m_{22}\frac{\partial\mathcal{U}}{\partial y} + q\mathcal{U}\right)dx \\
&\quad - \int_{(x_{i-\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{i-\frac{1}{2}}, y_{\frac{3}{2}})} \left(m_{11}\frac{\partial\mathcal{U}}{\partial x} + m_{12}\frac{\partial\mathcal{U}}{\partial y} + p\mathcal{U}\right)dy \\
&\quad - \int_{(x_{i-\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{\frac{1}{2}})} \left(m_{21}\frac{\partial\mathcal{U}}{\partial x} + m_{22}\frac{\partial\mathcal{U}}{\partial y} + q\mathcal{U}\right)dx.
\end{aligned}$$

$$\int_{\mathcal{C}_{i,1}} \nabla k(\mathcal{U}) \approx a^1_{i,1}\mathcal{U}_{i,1} + b^1_{i,1}\mathcal{U}_{i+1,1} + c^1_{i,1}\mathcal{U}_{i+1,2} + d^1_{i,1}\mathcal{U}_{i,2} + e^1_{i,1}\mathcal{U}_{i-1,2} + \alpha^1_{i,1}\mathcal{U}_{i-1,1} + t^1_{i,1}\mathcal{U}_{i-1,0}$$
$$+r^1_{i,1}\mathcal{U}_{i,0} + s^1_{i,1}\mathcal{U}_{i+1,0}, \tag{3.92}$$

with

$$a^1_{i,1} = T^{i+1,2}_{11} + T^{i+1,1}_{23} + T^{i+1,2}_{31} + T^{i,2}_{42} - T^{i,2}_{12} - T^{i,1}_{24} - \frac{1}{2}y_1\Big[\frac{1}{2}h_i(e+k) - h'_i\Big]$$
$$+l_1\max(f^{i+1}_x, 0) + h_i\max(f^2_y, 0)\Big) - l_1\min(f^i_x, 0),$$

$$b^1_{i,1} = T^{i+1,2}_{12} + T^{i+1,1}_{24} + T^{i+1,2}_{32} + l_i\min(f^{i+1}_x, 0) - \frac{1}{2}h'_i y_1, \qquad c^1_{i,1} = T^{i+1,2}_{14} + T^{i+1,2}_{34},$$

$$d^1_{i,1} = T^{i+1,2}_{13} + T^{i+1,2}_{33} + T^{i,2}_{44} - T^{i,2}_{14} + h_i\min(f^2_y, 0), \qquad e^1_{i,1} = T^{i,2}_{43} - T^{i,2}_{13},$$

$$\alpha^1_{i,1} = T^{i,2}_{41} - T^{i,2}_{11} - T^{i,1}_{23} - l_1\max(f^i_x, 0), \qquad t^1_{i,1} = -T^{i,1}_{21},$$

$$r^1_{i,1} = T^{i+1,1}_{21} - T^{i,1}_{22} + \frac{1}{4}y_1 h_i(e-k), \qquad s^1_{i,1} = T^{i+1,1}_{22}.$$

As we already mentioned, for the control volumes which are not in the degeneracy region, we use the multi-Point flux approximation to approximate the diffusion term and the upwind methods (first and second order) to approximate the convection term. So by combining as before, we obtain the following ODE

$$\frac{d\mathcal{U}}{d\tau} = A\mathcal{U} + F, \tag{3.93}$$

where

$$\mathcal{U} = \begin{bmatrix} \mathcal{U}_{11} \\ \mathcal{U}_{12} \\ \vdots \\ \mathcal{U}_{1N} \\ \mathcal{U}_{21} \\ \mathcal{U}_{22} \\ \vdots \\ \mathcal{U}_{2N} \\ \vdots \\ \vdots \\ \mathcal{U}_{N,1} \\ \mathcal{U}_{N,2} \\ \vdots \\ \mathcal{U}_{NN} \end{bmatrix}, \qquad A = L^{-1}\Big(Z + A_L\Big),$$

with F the vector of boundary conditions, A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (3.10). The elements of A_L are $h_i l_j \lambda$ with λ given in (3.4). The matrix L is also a

diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $h_i l_j$ for $i, j = 1, \ldots, N$ and

$$Z = \begin{bmatrix} D_1 & K_1 & 0_N & \ldots & \ldots & \ldots & \ldots & 0_N \\ L_2 & D_2 & K_2 & \ddots & & & & \vdots \\ 0_N & L_3 & D_3 & K_3 & \ddots & & & \vdots \\ \vdots & \ddots & L_4 & D_4 & K_4 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & & & & \ddots & L_{N-1} & D_{N-1} & K_{N-1} \\ 0_N & \ldots & \ldots & \ldots & \ldots & 0_N & L_N & D_N \end{bmatrix}.$$

The fitted matrix Z uses the first order upwind method. The matrices D_i, K_i, L_i are tri-diagonal matrices defined as follows. For $i = 1, N$

$$k = 1, \ldots, N, \quad (D_i)_{kk} = a^1_{1,k}, \qquad k = 1, \ldots, N-1 \quad (D_i)_{k,k+1} = d^1_{1,k},$$

$$k = 2, \ldots, N, \quad (D_i)_{k,k-1} = \gamma^1_{1,k},$$

$$k = 1, \ldots, N, \quad (K_1)_{kk} = b^1_{1,k}, \qquad k = 1, \ldots, N-1 \quad (K_1)_{k,k+1} = c^1_{1,k},$$

$$k = 2, \ldots, N \quad (K_1)_{k,k-1} = \lambda^1_{1,k},$$

$$(L_N)_{11} = \alpha^1_{N,1}, \quad (L_N)_{12} = e^1_{N,1},$$

$$k = 2, \ldots, N \quad (L_N)_{kk} = \alpha_{N,k} + \epsilon_{N,k}, \qquad k = 1, \ldots, N-1 \quad (L_N)_{k,k+1} = e_{N,k},$$

$$k = 2, \ldots, N \quad (L_N)_{k,k-1} = \beta_{N,k},$$

For $i = 2, \ldots, N-1$,

$$(D_i)_{11} = a^1_{i,1} \; ; \; (D_i)_{12} = d^1_{i,1}; \qquad (K_i)_{11} = b^1_{i,1} \; ; \; (K_i)_{12} = c^1_{i,1}, \qquad (L_i)_{11} = \alpha_{i,1}, \quad (L_i)_{12} = e^1_{i,1},$$

$$k = 2, \ldots, N \qquad (D_i)_{kk} = a_{i,k} + \Omega_{i,k}, \qquad (K_i)_{kk} = b_{i,k} + \psi_{i,k}, \qquad (L_i)_{kk} = \alpha_{i,k} + \epsilon_{i,k},$$

$$k = 2, \ldots, N-1, \qquad (D_i)_{k,k+1} = d_{i,k} + \phi_{i,k}, \qquad (K_i)_{k,k+1} = c_{i,k}, \qquad (L_i)_{k,k+1} = e_{i,k},$$

$$k = 2, \ldots, N, \qquad (D_i)_{k,k-1} = \gamma_{i,k} + \mu_{i,k}, \qquad (K_i)_{k,k-1} = \lambda_{i,k}, \qquad (L_i)_{k,k-1} = \beta_{i,k},$$

where all the elements $a^1_{i,j}, b^1_{i,j}, c^1_{i,j}, d^1_{i,j}, e^1_{i,j}, \gamma^1_{i,j}, \lambda^1_{i,j}$ are defined in (3.90),(3.91),(3.92) and the others elements are defined in (3.41) and (3.58).

Fitted MPFA-2^{nd} order upwind

As discussed in the previous paragraph, the fitted MPFA-2^{nd} order upwind is the combination of the fitted finite volume method and the MPFA method coupled to the 2^{nd} order upwind method described in Section 3.3.2.

Thereby, combining the fitted finite volume method, the MPFA and the second order upwind method gives

$$\frac{d\mathcal{U}}{d\tau} = A\mathcal{U} + F, \tag{3.94}$$

where

$$\mathcal{U} = \begin{bmatrix} \mathcal{U}_{11} \\ \mathcal{U}_{12} \\ \vdots \\ \mathcal{U}_{1N} \\ \mathcal{U}_{21} \\ \mathcal{U}_{22} \\ \vdots \\ \mathcal{U}_{2N} \\ \vdots \\ \vdots \\ \mathcal{U}_{N,1} \\ \mathcal{U}_{N,2} \\ \vdots \\ \mathcal{U}_{NN} \end{bmatrix}, \qquad A = L^{-1}\Big(Y + A_L\Big),$$

with G the vector of boundary conditions, A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (3.10). The elements of A_L are $h_i l_j \lambda$ with λ given in (3.4). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose elements are $h_i l_j$ for $i, j = 1, \ldots, N$ and

$$Y = \begin{bmatrix} H_1 & P_1 & 0_N & 0 & \ldots & \ldots & \ldots & & 0_N & 0_N \\ Q_2 & H_2 & P_2 & R_2 & 0_N & & & & & 0_N \\ W_3 & Q_3 & H_3 & P_3 & R_3 & 0_N & & & & \vdots \\ 0_N & W_4 & Q_4 & H_4 & P_4 & R_4 & \ddots & & & \\ 0_N & 0_N & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0_N \\ & & & & \ddots & W_{N-2} & Q_{N-2} & H_{N-2} & P_{N-2} & R_{N-2} \\ \vdots & & & & & \ddots & W_{N-1} & Q_{N-1} & H_{N-1} & P_{i,N-1} \\ 0_N & \ldots & \ldots & \ldots & & \ldots & 0_N & 0_N & Q_N & H_N \end{bmatrix}.$$

The elements of matrix Y are matrices. Indeed 0_N is a zeros matrix of size $N \times N$. The matrices H_i, P_i, Q are tri-diagonal matrices and W_i, R_i are diagonal matrices defined as follows:

$(H_1)_{11} = a^1_{11}$ $(H_1)_{12} = d^1_{11}$, $\quad (P_1)_{11} = b^1_{11}$, $(P_1)_{12} = c^1_{11}$;

$k = 2, \ldots, N$ $(H_1)_{kk} = a^1_{1,k}$, $\quad k = 2, \ldots, N-1$ $(H_1)_{k,k+1} = d^1_{1,k}$, $\quad k = 2, \ldots, N$ $(H_1)_{k,k-1} = \gamma^1_{1,k}$;

$k = 2, \ldots, N$, $(P_1)_{kk} = b^1_{1,k}$, $\quad k = 2, \ldots, N-1$, $(P_1)_{k,k+1} = c^1_{1,k}$, $\quad k = 2, \ldots, N$ $(P_1)_{k,k-1} = \lambda^1_{1,k}$;

For $i = 2, \ldots, N-1$,

$$(H_i)_{11} = a^1_{i,1}, \quad (H_i)_{12} = d^1_{i,1}, \qquad (P_i)_{11} = b^1_{i,1} + \Delta_{i,1}, \quad (P_i)_{12} = c^1_{i,1}, \qquad (Q_i)_{11} = \alpha_{i,1} + \eta_{i,1}, \quad (Q_i)_{12} = e^1_{i,1};$$

$$k = 2, \ldots, N, \quad (H_i)_{kk} = a_{i,k} + \Omega_{i,k}, \qquad (P_i)_{kk} = b_{i,k} + \Delta_{i,k}, \qquad (Q_i)_{kk} = \alpha_{i,k} + \eta_{i,k};$$

$$k = 2, \ldots, N-1, \qquad (H_i)_{k,k+1} = d_{i,k} + \phi_{i,k}, \qquad (P_i)_{k,k+1} = c_{i,k}, \qquad (Q_i)_{k,k+1} = e_{i,k};$$

$$k = 2, \ldots, N, \qquad (H_i)_{k,k-1} = \lambda_{i,k} + \mu_{i,k}, \qquad (P_i)_{k,k-1} = \lambda_{i,k}, \qquad (Q_i)_{k,k-1} = \beta_{i,k};$$

$$k = 2, \ldots, N-2, \qquad (H_i)_{k,k+2} = \Psi_{i,k}, \qquad k = 3, \ldots, N \quad (H_i)_{k,k-2} = \kappa_{i,k};$$

and

$$(R_i)_{kk} = \Pi_{ik}, \quad i = 2, \ldots, N-2, \ k = 2, \ldots, N-1,$$
$$(W_i)_{kk} = \epsilon_{ik}, \quad i = 3, \ldots, N-1, \ \ = 2, \ldots, N-1,$$

where all the elements $a^1_{i,j}, b^1_{i,j}, c^1_{i,j}, d^1_{i,j}, e^1_{i,j}, \gamma^1_{i,j}, \lambda^1_{i,j}$ are defined (3.90),(3.91),(3.92), and the others elements are defined in (3.41) and (3.73).

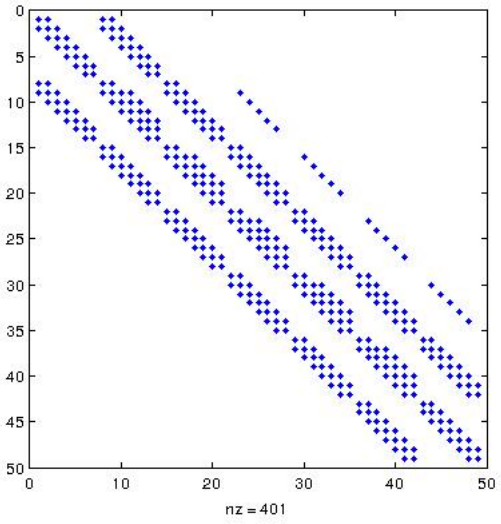


Figure 3.23: Structure of the fitted MPFA matrix using the 2nd order upwind

3.5 Time discretization

Let us consider the ODE stemming from the spatial dicretization and given by (3.60),(3.75),(3.93) and (3.94)

$$\frac{d\mathcal{U}}{d\tau} = A\mathcal{U} + F.$$

Using the θ-method for the time discretization, we have

$$\frac{\mathcal{U}^{n+1} - \mathcal{U}^n}{\Delta\tau} = \theta\Big(A\mathcal{U}^{n+1} + F^{n+1}\Big) + (1-\theta)\Big(A\mathcal{U}^n + F^n\Big). \tag{3.95}$$

Hence

$$\mathcal{U}^{n+1} = \Big(I - \theta\Delta\tau A\Big)^{-1}\left[\Big(I + (1-\theta)\Delta\tau A\Big)\mathcal{U}^n + \theta\Delta\tau F^{n+1} + (1-\theta)\Delta\tau F^n\right], \tag{3.96}$$

with

$$\mathcal{U}^n = \big[\mathcal{U}_{11}(\tau_n)\ \ \mathcal{U}_{12}(\tau_n)\ \ \ldots\ \ \mathcal{U}_{1N}(\tau_n)\ \ \mathcal{U}_{21}(\tau_n)\ \ \ldots\ \ \mathcal{U}_{2N}(\tau_n)\ \ \ldots\ \ \mathcal{U}_{N,1}(\tau_n)\ \ \ldots\ldots\ \mathcal{U}_{NN}(\tau_n)\big]^T,$$

$$F^n = F(\tau_n), \quad \tau_n = n\Delta\tau.$$

3.6 Numerical experiments

In this Section, we perform some numerical simulations and compare different numerical schemes developed in this work. More precisely, we compare the novel fitted MPFA method combined to the upwind methods, first order (fitted MPFA-1st upw) and second order (fitted MPFA-2nd upw), with the fitted finite volume method by Huang et al.[2006] and the standard MPFA method combined to the upwinds methods, first order (MPFA-1st upw) and second order (MPFA-2nd upw). The analytical solution of the PDE (3.2) is well known and given as

$$\begin{aligned} C(x,y,K,T) &= xe^{-rT}M(y_1,d;\rho_1) + ye^{-rT}M(y_2,-d+\sigma\sqrt{T};\rho_2) \\ &\quad -Ke^{-rT}\times\Big(1 - M(-y_1+\sigma_1\sqrt{T},-y_2+\sigma_2\sqrt{T};\rho)\Big), \end{aligned} \tag{3.97}$$

where

$$\begin{aligned} d &= \frac{\ln(x/y) + (b_1 - b_2 + \sigma_1^2/2)T}{\sigma\sqrt{T}}, \\ y_1 &= \frac{\ln(x/K) + (b_1 + \sigma_1^2/2)T}{\sigma_1\sqrt{T}}, \qquad y_2 = \frac{\ln(y/K) + (b_1 + \sigma_2^2/2)T}{\sigma_2\sqrt{T}}, \\ \sigma &= \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, \qquad \rho_1 = \frac{\sigma_1 - \rho\sigma_2}{\sigma}, \qquad \rho_2 = \frac{\sigma_2 - \rho\sigma_1}{\sigma}, \end{aligned}$$

with

$$M(a,b;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{a}\int_{-\infty}^{b} exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right)dxdy.$$

and illustrated in Figure (3.24) below

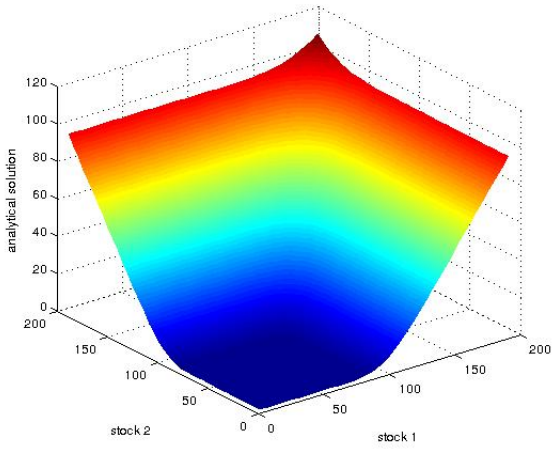
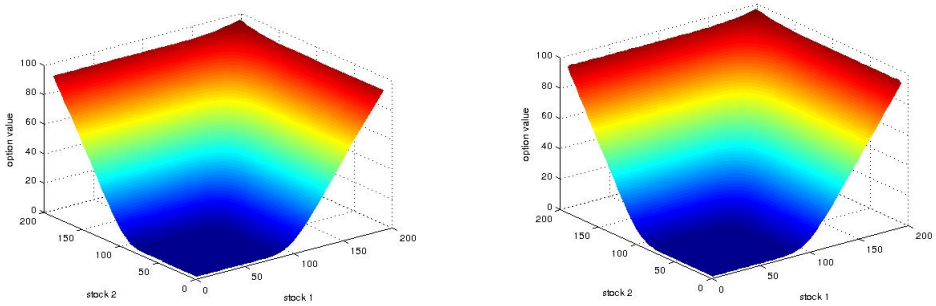


Figure 3.24: Analytical solution for the price at final T. The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/12,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0.03$ and $\Delta\tau = 1/100$.

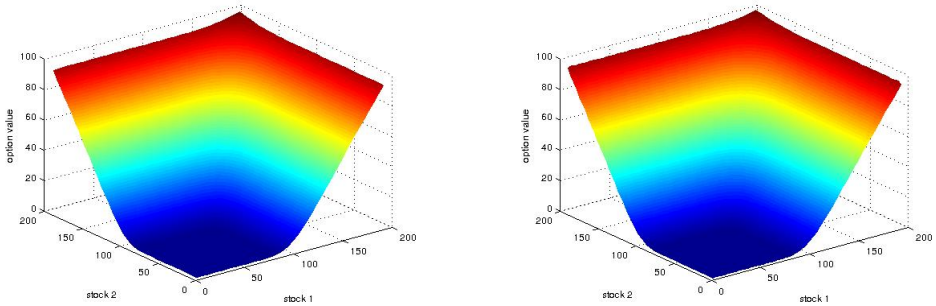
Note that in all our nuemrical schemes, the Dirichlet boundary conditions are used with the value equal to the analytical solution. The graphs of options price with differents methods are given in Figure (3.25) and Figure (3.26) below



(a) MPFA-upwind 1st order

(b) MPFA-upwind 2nd order

Figure 3.25: Option price for MPFA-upwind methods at final time T. The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/12,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0.03$ and $\Delta\tau = 1/100$.



(a) fitted MPFA-upwind 1st order

(b) fitted MPFA-upwind 2nd order

Figure 3.26: Option price for fitted MPFA-upwind methods at final time T. The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/12,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0.03$ and $\Delta = 1/100$.

In this paragraph, we consider the four numerical methods illustrated in the previous sections and the fitted finite volume Huang. We evaluate the error of these numeircal method with respect to the analytical solution (3.97). The L^2- norm is used to compute the error as follows:

$$err = \frac{\sqrt{\sum_{i,j=1}^{N} meas(\mathcal{C}_{ij})\big(\mathcal{U}_{ij} - U_{ij}^{ana}\big)^2}}{\sqrt{\sum_{i=1}^{n} meas(\mathcal{C}_{ij})\big(U_{ij}^{ana}\big)^2}}, \tag{3.98}$$

where $\mathcal{U}$ is the numerical solution, U^{ana} the analytical solution and $meas(\mathcal{C}_{ij})$ is the measure of the control volume $\mathcal{C}_{ij}$. This gives the following tables:

Table 3.1: Table of errors.The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/6,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0.1$ and $\Delta\tau = 1/100$.

Num method / Nb of grid pts	Fitted fin vol	MPFA-1^{st} upw	MPFA-2^{nd} upw	fitted MPFA-1^{st} upw	fitted MPFA -2^{nd} upw
50×50	0.0134	0.0060	0.0059	0.0060	0.0060
70×70	0.0133	0.0044	0.0044	0.0044	0.0044
85×85	0.0132	0.0037	0.0037	0.0037	0.0037
100×100	0.0132	0.0032	0.0032	0.0032	0.0032
150×150	0.0131	0.0024	0.0023	0.0023	0.0023

Table 3.2: Table of errors.The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/6,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0.08$ and $\Delta\tau = 1/100$.

Num method / Nb of grid pts	Fitted fin vol	MPFA-1^{st} upw	MPFA-2^{nd} upw	fitted MPFA-1^{st} upw	fitted MPFA -2^{nd} upw
50×50	0.0134	0.0060	0.0059	0.0060	0.0060
100×100	0.0104	0.0056	0.0055	0.0056	0.0055
150×150	0.0131	0.0056	0.0055	0.0056	0.0055

Table 3.3: Table of errors.The computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0, T]$ with $T = 1/6,\ K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficent is $\rho = 0.5$, the risk free interest $r = 0$ and $\Delta\tau = 1/100$.

Num method / Nb of grid pts	Fitted fin vol	MPFA-1^{st} upw	MPFA-2^{nd} upw	fitted MPFA-1^{st} upw	fitted MPFA -2^{nd} upw
100×100	0.0152	0.0239	0.0235	0.0240	0.0229
150×150	0.0151	0.0231	0.0228	0.0232	0.0229

Table 3.4: Table of errors.The computational domain of the problem is $\Omega = [0; 4] \times [0; 4] \times [0, T]$ with $T = 2,\ K = 1$, the volatilities $\sigma_1 = \sigma_2 = 1$. The correlation coefficent is $\rho = 0.3$, the risk free interest $r = 0.5$ and $\Delta\tau = 1/100$.

Num method / Nb of grid pts	Fitted fin vol	MPFA-1^{st} upw	MPFA-2^{nd} upw	fitted MPFA-1^{st} upw	fitted MPFA -2^{nd} upw
50×50	0.01208	0.0631	0.0669	0.0623	0.0659
150×150	0.01203	0.0572	0.0648	0.0559	0.0629

Table 3.5: Table of errors.The computational domain of the problem is $\Omega = [0;4] \times [0;4] \times [0,T]$ with $T = 2,\ K = 1$, the volatilities $\sigma_1 = \sigma_2 = 1$. The correlation coefficent is $\rho = 0.3$, the risk free interest $r = 0.5$ and $\Delta\tau = 1/10$.

Num method / Nb of grid pts	Fitted fin vol	MPFA-1^{st} upw	MPFA-2^{nd} upw	fitted MPFA-1^{st} upw	fitted MPFA -2^{nd} upw
50×50	0.01196	0.0562	0.0643	0.0555	0.0624
100×100	0.01201	0.0626	0.0664	0.0618	0.0654

As we can observe in Table 3.1-3.5, the errors from our fitted MPFA and MPFA methods are smaller compared to those of fitted finite volume in Huang et al. [2006]. We can also note that when r become smaller, the gaps between the errors of the fitted finite volume in Huang et al. [2006] and our fitted MPFA and MPFA methods reduce.

Conclusion

In this Chapter, we have presented the Multi-Point Flux Approximation (MPFA) to approximate the diffusion term of Black-Scholes Partial Differential Equation in its divergence form. The MPFA method coupled with the upwind methods (first and second order) have been used to solve numerically the Black-Scholes PDE.
To handle the degeneracy of the Black Scholes PDE, we have proposed a novel method based on a combination of the MPFA method and the fitted finite volume by Huang et al. [2006]. Besides, it is important to mention that the one dimensional version of the MPFA method is the TPFA method introduced in the previous Chapter. Thereby, the convergence proof for the MPFA will use similar arguments as in the case of the TPFA methods.
Moreover, we performed some numerical simulations which show that our fitted MPFA method coupled with the first or second order upwinding methods are more accurate than the fitted finite volume method Huang et al. [2006].

Chapter 4

A L-Multi-Point Flux Approximation method and a fitted L-Multi-point Flux Approximation method for pricing two dimensional options

In this Chapter, we introduce a special kind of finite volume method called Multi-Point Flux Approximation method (MPFA) to price European and American options in two dimensional domain. We focus on the L-MPFA method for space discretization of the diffusion term of Black-Scholes operator. The degeneracy of the Black Scholes operator is tackled using the standard fitted finite volume method. This combination of standard fitted finite volume method and L-MPFA method coupled to upwind methods gives us a novel scheme called the fitted L-MPFA method. Numerical experiments show the accuracy of the novel fitted L-MPFA method comparing to the O-MPFA methods presented in the previous chapter and well known schemes for pricing options. This Chapter is part of the preprint that can be found in Koffi and Tambue [2019b]

4.1 Introduction

In finance, there exist two main types of options which are European and American options. European options are options that can be exercised only at expiry date while American options can be exercised anytime before the expiry date. This flexibility of exercising American options leads to solve an optimal stopping time problem in the Black-Scholes framework which incorporate the early exercise. Many studies focused on the pricing problem of American options were conducted and the linear complementary problem approach was quite popular for pricing American options (see Kovalov et al. [2007], Topper [2005], Wang et al. [2006], Zhang et al. [2009]). This approach brings us to solve linear complementary problem stated as follows (see Topper [2005]):

$$\begin{cases} \mathcal{L}U & \geq & 0, \\ U - U^{\star} & \geq & 0, \\ \mathcal{L}U \cdot \left(U - U^{\star}\right) & = & 0. \end{cases} \tag{4.1}$$

where $\mathcal{L}$ is the following Black-Scholes operator

$$\mathcal{L}U = \frac{\partial U}{\partial t} - \frac{1}{2}\sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2 U}{\partial x_i \partial x_j} - r\sum_{i=1}^{n} x_i \frac{\partial U}{\partial x_i} + rU, \tag{4.2}$$

with r is the risk-free interest, t is the time to maturity T, U is the option value at time t, $U^{\star}$ is the payoff. For $i, j = 1, \ldots, n$, x_i represents the asset i price, σ_i represents the volatility of asset i, ρ_{ij}

represents the correlation between the assets i and j. Furthermore, Wang et al. [2006] proposed a power penalty method to solve the linear complementary problem for pricing American options. The power penalty problem is formulated as follows:

$$\mathcal{L}U + \eta\big[U^\star - U\big]_+^{1/k} = 0, \tag{4.3}$$

where η is penalty parameter and k is the power of the method. Let us notice that, when we take the penalty parameter $\eta = 0$ in (4.3), we get the Black-Scholes Equation for pricing European options, with the operator $\mathcal{L}$ defined in (4.2). However, the power penalty problem (4.3) can not be solved analytically, therefore numerical methods are required for its resolution. Nevertheless, the Black-Scholes operator (4.2) is degenerated when the stock price approaches zero. This degeneracy can affect the accuracy of the numerical method used for the resolution. To tackle this problem, several methods have been proposed. The fitted finite volume method, proposed by S.Wang in Wang [2004] whereby a rigorous proof of convergence is provided, appears to be more attractive. Moreover, the fitted finite volume method has been used for the resolution of the two dimensional second order Black Scholes PDE followed by the convergence proof in Huang et al. [2006]. In spite of the fact that the fitted finite volume methods perform well for the resolution of the Black-Scholes PDE, they are only of order 1 with respect to asset price variable. Besides, the fitted O-Multi-Point Flux Approximation (O-MPFA) method has been proposed in Chapter 3 to overcome the degeneracy problem of the Black-Scholes PDE. It has been shown that the O-MPFA is more accurate than the classical fitted finite volume method by Wang [2004]. However, the O-MPFA is computationally heavy, 9 points stencil method, and for more general grids, the convergence rate of the O-MPFA method may reduce (see Aavatsmark [2007]). In this Chapter, we focus on the L-MPFA method which is based on the approximation of a linear function gradient defined over a given triangle and the continuity of flux through the edges of this triangle.
Indeed, the L-MPFA method is a 7 points stencil method while the O-MPFA is a 9 points stencil method. This shows that the O-MPFA method can be computationally more expensive than the L-MPFA method. Moreover, for more general grids, the order reduction in convergence rate is larger for the O-MPFA than the L-MPFA (see Aavatsmark [2002]). Thereby, to approximate the solution of the second order Black-Scholes operator, we couple the L-MPFA method with the upwind methods (first and second order). Besides, the degeneracy of the Black-Scholes operator (4.2) is handled by the fitted finite volume (see in Wang [2004]) when the stock price is approaching zero. The L-MPFA method coupled with the upwind methods (1^{st} and 2^{nd} order) is used to approximate the solution of (4.3) when the Black-Scholes operator is not degenerated. We call fitted L-MPFA method that combination of the L-MPFA method and the fitted finite volume method. Numerical simulations show that the new fitted L-MPFA method is more accurate than the fitted O-MPFA method developed in the previous Chapter and the standard fitted finite volume method developed in Huang et al. [2006].
The Chapter is structured as follows. In Section 2, we present the power penalty problem with the corresponding initial and boundary conditions. The spatial discretization of the linear operator is developed in Section 3. Details on the L-MPFA method of the diffusion term discretization are provided. The convection term is discretized using the upwind methods (1^{st} and 2^{nd} method). At the end of Section 3, the novel fitted MPFA method is provided. The $\theta-$ Euler method is used for the time discretization method in Section 4. Numerical experiments are presented for the different numerical methods are presented in Section 5. The conclusions of our study are drawn in the last Section.

4.2 Formulation of the problem

We recall from Chapter 3 that an option with two underlying assets following the Black-Scholes model is given by

$$\begin{cases} dx(t) &= \mu_1 x dt + \sigma_1 x dW_1, \\ dy(t) &= \mu_2 y dt + \sigma_2 y dW_2, \\ dW_1(t)dW_2(t) &= \rho dt. \end{cases} \tag{4.4}$$

where μ_i, σ_i, W_i are respectively the drift, the volatility and the Wiener process governing the stocks x, y and ρ is the correlation coefficient between the two Wiener processes. As we have already discussed in Chapter 3, the value of the option U follows the two-dimensional Black-Scholes operator on the domain $D = \Omega \times (0,T),\ \Omega = [0, x_{\max}] \times [0, y_{\max}]$, given by

$$\mathcal{L}U := \frac{\partial U}{\partial t} - \frac{1}{2}\sigma_1^2 x^2 \frac{\partial^2 U}{\partial x^2} - \rho\sigma_1\sigma_2 xy \frac{\partial^2 U}{\partial x \partial y} - \frac{1}{2}\sigma_2^2 y^2 \frac{\partial^2 U}{\partial y^2} - rx\frac{\partial U}{\partial x} - ry\frac{\partial U}{\partial y} + rU, \tag{4.5}$$

where the initial and boundary conditions for an American put are given by

$$\begin{cases} U(x,y,0) = U^*(x,y) = \max\left(K - \alpha_1 x - \alpha_2 y, 0\right), \\ U(0,y,t,) = U(x,0,t,) = K, \\ \lim_{x,y \longrightarrow x_{\max}, y_{\max}} U(x,y,t) = 0. \end{cases} \tag{4.6}$$

K is the strike price, U^* is the payoff for basket options, and $\alpha_i,\ i = 1,2$, are weights such that $\alpha_1 + \alpha_2 = 1$. However, without a loss of generality, we can transform the non homogeneous boundary conditions (4.6) into homogeneous Dirichlet boundary conditions by substracting $\mathcal{L}U_0$ from both side of (4.5), with U_0 a function satisfying the boundary condition (4.6), and introducing a new variable V given by

$$V = e^{\beta t}(U - U_0) \ \text{ with } \ \beta = \frac{1}{2}\sup_{t\in[0,T]}\Big(\sigma_1^2(t) + \sigma_2^2(t) + \rho_{12}\sigma_1(t)\sigma_2(t)\Big). \tag{4.7}$$

Besides, in order to apply the finite volume method, it is convenient to re-write the Black-Scholes operator (4.5) in the following divergence form

$$\mathcal{L}V = \frac{\partial V}{\partial t} - \nabla\cdot(\mathbf{M}\nabla V) - \nabla(fV) + \lambda V, \tag{4.8}$$

where

$$\mathbf{M} = \frac{1}{2}\begin{pmatrix} \sigma_1^2 x^2 & \rho\sigma_1\sigma_2 xy \\ \rho\sigma_1\sigma_2 xy & \sigma_2^2 y^2 \end{pmatrix}, \quad f = \begin{pmatrix} (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)x \\ (r - \sigma_2^2 - \frac{1}{2}\rho\sigma_1\sigma_2)y \end{pmatrix},$$

$$\lambda = 3r + \beta - \sigma_1^2 - \sigma_2^2 - \rho\sigma_1\sigma_2.$$

4.2.1 Linear complementary approach

The linear complementary problem for pricing an American option with two underlying assets is given by

$$\begin{cases} \mathcal{L}v &\leq g, \\ v - v^\star &\leq 0, \\ \big(\mathcal{L}v - g\big)\cdot\big(v - v^\star\big) &= 0, \end{cases} \tag{4.9}$$

where $\mathcal{L}$ is the operator given by (4.8) , $g = e^{\beta t}\mathcal{L}U_0$, $v = -V$ with V defined in (4.7) and $v^\star = -e^{\beta t}\big(U^\star - U_0\big)$. Using the weighted Sobolev space defined in (1.66), let $\mathcal{G}$ be the set defined by

$$\mathcal{G} = \Big\{u \in H^1_{0,\omega}(\Omega) : u \leq v^\star\Big\}. \tag{4.10}$$

We should notice that $\mathcal{G}$ is a closed and convex subset of $H^1_{0,\omega}(\Omega)$. The variational formulation corresponding to the linear complementary problem (4.9) is then given by (see Wang et al. [2006])

Problem 3 *Find $v(t) \in \mathcal{G}$, such that, for all $u \in \mathcal{G}$,*

$$\Big(\frac{\partial v(t)}{\partial t}, u - v(t)\Big) + B\big(v(t), u - v(t); t\big) \geq \big(g(t), u - v(t)\big) \qquad a.e \in (0,T), \tag{4.11}$$

where

$$B\big(v, u; t\big) = \Big(M\nabla v, \nabla u\Big) + \Big(fv, \nabla u\Big) + \lambda(v,u), \qquad u, v \in H^1_{0,\omega}(\Omega), \tag{4.12}$$

is a bilinear form and M, f, λ are defined in (4.8).

Theorem 9 *There exists a unique solution in $\mathcal{G}$ to problem 3.*

Proof of Theorem 8 The proof is similar to the proof of Theorem 2.3 in Wang et al. [2006].

4.2.2 Power penalty approach

Pricing an American option with 2 underlying assets can also lead to solve the following power penalty problem with homogeneous boundary condition given by

$$\mathcal{L}V + \eta\big[V^\star - V\big]_+^{1/k} = g, \tag{4.13}$$

where the Black-Scholes operator $\mathcal{L}$ is defined in (4.8), V is defined in (4.7), η the penalty parameter and $g = \mathcal{L}U_0$. When the penalty parameter $\eta = 0$ in (4.13), we get the Black-Scholes Partial Differential Equation for pricing European options. We formulate the variational problem corresponding to the power penalty problem (4.13) as follows:

Problem 4 *Find $v(t) \in H^1_{0,\omega}(\Omega)$ such that for all $u \in H^1_{0,\omega}(\Omega)$,*

$$\Big(\frac{\partial v(t)}{\partial t}, u\Big) + B(v(t), u; t) + \eta\big(\big[v^\star - v\big]_+^{1/k}, v\big) = (g, u) \qquad a.e \in (0,T), \tag{4.14}$$

where

$$B\big(v, u; t\big) = \Big(M\nabla v, \nabla u\Big) + \Big(fv, \nabla u\Big) + \lambda(v,u) \quad u, v \in H^1_{0,\omega}(\Omega) \tag{4.15}$$

is a bilinear form and M, f, λ are defined in (4.8).

Theorem 10 *There exsits an unique solution v_η in $H^1_{0,\omega}(0,\omega)$ to Problem 4.*

Proof of Theorem 9 The proof is similar to the proof of Theorem 3.1 in Wang et al. [2006].

4.2.3 Convergence

Here, we recall an important result about the convergence of the solution of problem 3 to the solution of Problem 4. We first start by the following lemma.

Lemma 2 *Let v_η be the solution of Problem 3. If $v_\eta \in L^p(D)$, then there exists a positive constant C, independent of v_η and η, such that*

$$\big|\big|[v_\eta - v^\star]^+\big|\big|_{L^p(D)} \leq \frac{C}{\eta^k},$$

$$\big|\big|[v_\eta - v^\star]^+\big|\big|_{L^\infty\left(0,T;L^2(\Omega)\right)} + \big|\big|[v_\eta - v^\star]^+\big|\big|_{L^2\left(0,T;H^1_{0,\omega}(\Omega)\right)} \leq \frac{C}{\eta^{k/2}}. \tag{4.16}$$

where k is the penalty power, $p = 1 + 1/k$, and $[\cdot]^+ = \max(\cdot, 0)$.

Proof of Lemma 2 The proof is similar to the proof of Lemma 3.1 in Wang et al. [2006].

Theorem 11 *Let assume that the solution v to Problem 3 is such that $\frac{\partial u}{\partial t} \in L^{k+1}(D)$ and assumptions in lemma 2 are satisfied. Let assume v_η be the solution of Problem 4. Then, there exists a constant $C > 0$, independant of v, v_η, η such that*

$$||v - v_\eta||_{L^\infty\left(0,T;L^2(\Omega)\right)} + ||v - v_\eta||_{L^2\left(0,T;H^1_{0,\omega}(\Omega)\right)} \leq \frac{C}{\eta^{k/2}}. \tag{4.17}$$

where k is the penalty power.

Proof of Theorem 10 The proof is similar to the proof of theorem 4.1 in Wang et al. [2006].

Our goal in this Chapter is to approximate the value v_η. For simplicity v_η will be denoted by v.

4.2.4 Finite volume method

Let us consider the domain $D = \Omega \times (0, T)$ such that $\Omega = I_x \times I_y$ where $I_x = [0, x_{\max}]$ and $I_y = [0, y_{\max}]$.

At $x = x_{\max}$ and $y = y_{\max}$, the linear boundary condition will be applied Huang et al. [2006]. The intervals I_x and I_y will be subdivided into N without loss the generality, in the following way

$$I_{x_i} = [x_{i-1}; x_i], \; I_{y_j} = [y_{j-1}; y_j] \quad i, j = 1, ..., N+1. \tag{4.18}$$

Let us set the mid-points $x_{i-\frac{1}{2}}$ and $y_{j-\frac{1}{2}}$ as follows

$$x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2} \;\; y_{j-\frac{1}{2}} = \frac{y_{j-1} + y_j}{2} \qquad i, j = 1, ..., N, \tag{4.19}$$

with $h_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, \quad l_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$ and

$$x_{-\frac{1}{2}} = x_0 = 0, \qquad x_{N+\frac{3}{2}} = x_{N+1} = x_{\max}, \; y_{-\frac{1}{2}} = y_0 = 0\,, \;\; y_{N+\frac{3}{2}} = y_{N+1} = y_{\max}.$$

For $i, j = 1, \ldots, N$, we denote by $\mathcal{C}_{ij} = [x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}; y_{j+\frac{1}{2}}]$ a control volume associated our subdivision.

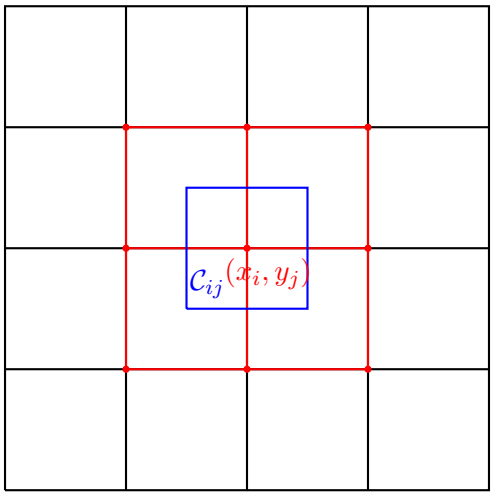


Figure 4.1: Control volume

Note that for $i, j = 1, \ldots, N$, the control volume $\mathcal{C}_{ij}$ is the area surrounding the grid point (x_i, y_j).

Our goal is to approximate the option function v at (x_i, y_j) [1] by a function denoted $\mathcal{V}$.

As in the previous Chapter, the matrix $\mathbf{M}$ in (4.8) will be replaced by its average value in each control volume.

$$\mathbf{M}^{ij} = \frac{1}{\text{meas}(\mathcal{C}_{ij})} \int_{\mathcal{C}_{ij}} \mathbf{M} dxdy, \;\; i,j = 1, ..., N, \tag{4.20}$$

where $\text{meas}(\mathcal{C}_{ij})$ is the measure of $\mathcal{C}_{ij}$. Thereby,

$$M^{ij} = \begin{bmatrix} \frac{\sigma_1^2}{6} \frac{x^3_{i+\frac{1}{2}} - x^3_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} & \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) \\ \frac{\rho\sigma_1\sigma_2}{8}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})(y_{j+\frac{1}{2}} + y_{j-\frac{1}{2}}) & \frac{\sigma_2^2}{6} \frac{y^3_{j+\frac{1}{2}} - y^3_{j-\frac{1}{2}}}{y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}} \end{bmatrix},$$

Now, let us consider the divergence form of equation (4.13). According to the finite volume method, we integrate the PDE (4.13) over each control volume $\mathcal{C}_{ij}$ as follows, for $i, j = 1, ..., N$,

$$\int_{\mathcal{C}_{ij}} \frac{\partial \mathcal{V}}{\partial t} d\mathcal{C} - \int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M} \nabla \mathcal{V}) d\mathcal{C} - \int_{\mathcal{C}_{ij}} \nabla (f\mathcal{V}) d\mathcal{C} - \int_{\mathcal{C}_{ij}} \lambda \mathcal{V} d\mathcal{C} + \int_{\mathcal{C}_{ij}} \eta \big[V^\star - \mathcal{V}\big]_+^{1/k} d\mathcal{C} = 0. \tag{4.21}$$

The next Section will be dedicated to spatial discretization of equation (4.21). For the term in the left hand side of the equality sign and for the last one in the right hand side of (4.21), we use the mid-point quadrature rule for their approximation as follows:

$$\int_{\mathcal{C}_{ij}} \frac{\partial \mathcal{V}}{dt} d\mathcal{C} \approx meas(\mathcal{C}_{ij}) \frac{\partial \mathcal{V}}{\partial t}(x_i, y_j, t) \approx h_i l_j \frac{\partial \mathcal{V}_{ij}}{\partial t}, \tag{4.22}$$

$$\int_{\mathcal{C}_{ij}} \lambda \mathcal{V} d\mathcal{C} \approx meas(\mathcal{C}_{ij}) \lambda \mathcal{V}(x_i, y_j, t) \approx h_i l_j \lambda \mathcal{V}_{ij}, \tag{4.23}$$

$$\int_{\mathcal{C}_{ij}} \eta \big[V^\star - \mathcal{V}\big]_+^{1/k} d\mathcal{C} \approx meas(\mathcal{C}_{ij}) \eta \big[V^\star - \mathcal{V}\big]_+^{1/k} \approx h_i l_j \eta \big[V_i^\star - \mathcal{V}_{ij}\big]_+^{1/k}. \tag{4.24}$$

The convection term

$$\int_{\mathcal{C}_{ij}} \nabla (f\mathcal{V}) d\mathcal{C}, \tag{4.25}$$

of (4.21) will be approximated using the upwind methods (first or second order). The diffusion term

$$\int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M} \nabla \mathcal{V}) d\mathcal{C}, \tag{4.26}$$

of (4.21) will be approximated using the **Multi-Point Flux Approximation** (MPFA) **L**-method or the **fitted multi-point flux approximation L-** method. More details about these methods will be given in the next Section.

4.3 Space discretization

The spatial discretization of (4.8) consists in approximating all terms in (4.21) over the control volumes of the study domain.

[1] center of the control volume $\mathcal{C}_{ij}$

4.3.1 Discretization of the diffusion term

Let us start by applying the divergence theorem to the diffusion term (4.26) as follows, for $i, j = 1, ..., N$:

$$\mathcal{F}^{ij} = \int_{\mathcal{C}_{ij}} \nabla \cdot (\mathbf{M}^{ij} \nabla \mathcal{V}) = \int_{\partial \mathcal{C}_{ij}} (\mathbf{M}^{ij} \nabla \mathcal{V}) \cdot \vec{n} d\mathcal{C}, \tag{4.27}$$

where $\vec{n}$ is the outward vector from the control volume.

Now, we can apply the so-called L-**Multi-Point Flux Aprroximation(MPFA)** method to approximate the integral defined in (4.27).

L -Multi-Point Flux Approximation (L-MPFA) method

The L-MPFA method takes its name from the fact that the curve connecting the three control volume centres considered for the application of the method, constitutes a stylized "L". Here, we follow the description of the L-method given in Aavatsmark [2002].

Let us consider the triangle $x_1x_2x_3$ (see Figure 4.2), and a linear function g defined over this triangle. we define

$$\mathbf{X}\nabla g = \begin{bmatrix} g(x_2) - g(x_1) \\ \\ g(x_3) - g(x_1) \end{bmatrix}, \tag{4.28}$$

where

$$\mathbf{X} = \begin{bmatrix} \big(x_2 - x_1\big)^T \\ \\ \big(x_3 - x_1\big)^T \end{bmatrix}. \tag{4.29}$$

Thereby, the gradient of the linear function g may be expressed as follows:

$$\nabla g = \frac{1}{T}\bigg(\nu_2\Big(g(x_2) - g(x_1)\Big) + \nu_3\Big(g(x_3) - g(x_1)\Big)\bigg), \tag{4.30}$$

where ν_2, ν_3 are respectively the normal vector to $x_2 - x_1$ and $x_3 - x_1$ defined by

$$\nu_2 = \mathbf{R}\big(x_2 - x_1\big), \qquad \nu_3 = -\mathbf{R}\big(x_3 - x_1\big),$$

with

$$\mathbf{R} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and

$$T = \nu_2^T \mathbf{R} \nu_3.$$

Let's notice that the matrix $\mathbf{R}$ is a rotation of angle $-\frac{\pi}{2}$. Thereby the vector ν_2 and ν_3 have same length with the vectors $x_2 - x_1$ and $x_3 - x_1$.

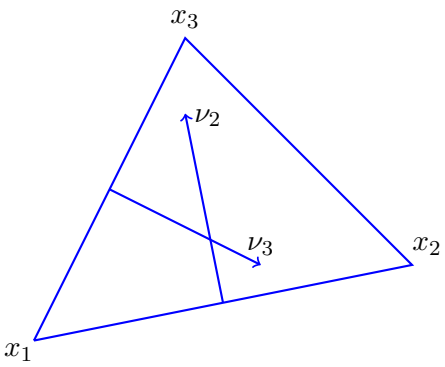


Figure 4.2: Triangle $x_1x_2x_3$

Let us called **interaction volume** $\mathcal{R}_{ij}$ a cell grid defined as follows

$$for \quad i,j = 1,\dots,N+1, \qquad \mathcal{R}_{ij} = [x_{i-1};x_i] \times [y_{j-1};y_j]. \tag{4.31}$$

We denote respectively by $x_1(x_{i-1},y_{j-1}), x_2(x_i,y_{j-1}), x_3(x_i,y_j)$ and $x_4(x_{i-1},y_j)$ the centre of the control volume $\mathcal{C}_{ij}, \mathcal{C}_{i+1,j}, \mathcal{C}_{i,j+1}$ and $\mathcal{C}_{i+1,j+1}$. We denote also by $\bar{x}_1, \bar{x}_2, \bar{x}_3$ *and* $\bar{x}_4$ the midpoints of the segment x_1x_2, x_3x_4, x_1x_3 *and* x_2x_4. We may notice that an interaction volume $\mathcal{R}_{ij}$ is covering an area in the intersection of the control volumes $\mathcal{C}_{ij}, \mathcal{C}_{i+1,j}, \mathcal{C}_{i,j+1}$ and $\mathcal{C}_{i+1,j+1}$. An interaction volume can be divided into 2 triangles such that the half edges $1,2$ are in the triangle $T_1 = x_1x_2x_3$ and the half edges $3,4$ are in the triangle $T_2 = x_1x_3x_4$ (see Figure 4.3).
Here, we follow the procedure in Aavatsmark [2007].

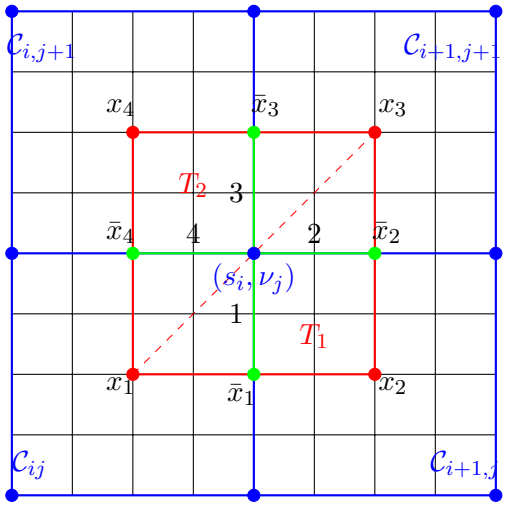


Figure 4.3: Interaction volume

In an interaction volume, we aim to compute the flux through the half edges $1,2,3$ and 4 (see Figure 4.3). Thereby, using (4.27), the flux f_p^{ij} through the half edge p seen from the centre of the control volume $\mathcal{C}_{ij}$ is expressed as follows:

$$f_p^{ij} = n_p^T \mathbf{M}^{ij} \nabla \mathcal{V}_{ij}, \tag{4.32}$$

where n_p is the vector normal to the half edge p with the same length.

Let us consider the triangle $T_1 = x_1x_2x_3$ from the interaction volume $\mathcal{R}_{ij}$

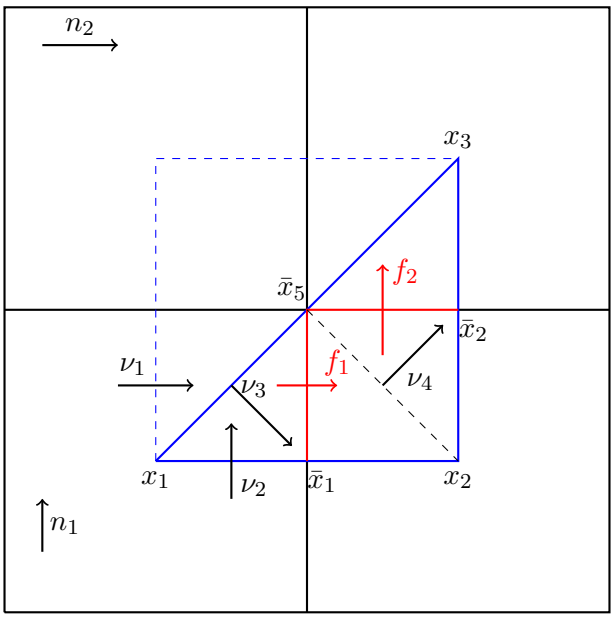


Figure 4.4: Triangle T_1

In the triangle $T_{12} = \bar{x}_1 x_2 \bar{x}_2$, using the flux expression (4.32) and following the gradient expression (4.30) lead to

$$\begin{aligned} f_1^{i,j-1} &= \omega_{12}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \omega_{11}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big), \\ f_2^{i,j-1} &= \omega_{22}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \omega_{21}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big), \end{aligned} \tag{4.33}$$

where

$$\omega_{11}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \times n_1^T \mathbf{M}^{i,j-1} \nu_1, \qquad \omega_{12}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \times n_1^T \mathbf{M}^{i,j-1} \nu_2,$$

$$\omega_{21}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \times n_2^T \mathbf{M}^{i,j-1} \nu_1, \qquad \omega_{22}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \times n_2^T \mathbf{M}^{i,j-1} \nu_2,$$

with

$$T_1^{i,j-1} = -\nu_2^T \mathbf{R} \nu_1.$$

Moreover, using the property of the matrix $\mathbf{R}$, we have

$$\nu_4 = \mathbf{R}(\bar{x}_5 - x_2). \tag{4.34}$$

Using also the equation (4.28) and the expression of gradient (4.30) in the triangle T_{12} lead to

$$\bar{\mathcal{V}}_5 = \mathcal{V}_{i,j-1} + \chi_{42}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \chi_{41}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big), \tag{4.35}$$

where

$$\chi_{41}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \nu_4^T \mathbf{R} \nu_1, \qquad \chi_{42}^{i,j-1} = \frac{1}{T_1^{i,j-1}} \nu_4^T \mathbf{R} \nu_2.$$

In the triangle $T_{11} = x_1\bar{x}_1\bar{x}_5$, we have

$$f_1^{i-1,j-1} = \omega_{13}^{i-1,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i-1,j-1}\Big) + \omega_{12}^{i-1,j-1}\Big(\bar{\mathcal{V}}_5 - \mathcal{V}_{i-1,j-1}\Big). \tag{4.36}$$

Replacing $\bar{\mathcal{V}}_5$ by its expression (4.35) in (4.36), gives

$$\begin{aligned} f_1^{i-1,j-1} &= \omega_{13}^{i-1,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i-1,j-1}\Big) + \omega_{12}^{i-1,j-1}\Bigg(\mathcal{V}_{i,j-1} - \mathcal{V}_{i-1,j-1} + \chi_{42}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) \\ &\quad -\chi_{41}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big)\Bigg), \end{aligned} \tag{4.37}$$

where

$$\omega_{13}^{i-1,j-1} = \frac{1}{T_1^{i-1,j-1}} \times n_1^T \mathbf{M}^{i-1,j-1}\nu_3, \qquad \omega_{12}^{i-1,j-1} = \frac{1}{T_1^{i-1,j-1}} \times n_1^T \mathbf{M}^{i-1,j-1}\nu_2,$$

with

$$T_1^{i-1,j-1} = \nu_3^T \mathbf{R}\nu_2.$$

Similarly, in the triangle $T_{13} = \bar{x}_5\bar{x}_2x_3$

$$f_2^{ij} = -\omega_{21}^{ij}\Big(\bar{\mathcal{V}}_5 - \mathcal{V}_{ij}\Big) + \omega_{23}^{ij}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{ij}\Big), \tag{4.38}$$

Replacing $\bar{\mathcal{V}}_5$ by its expression (4.35) in (4.38), gives

$$f_2^{ij} = -\omega_{21}^{ij}\Bigg(\mathcal{V}_{i,j-1} - \mathcal{V}_{ij} + \chi_{42}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \chi_{41}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big)\Bigg) + \omega_{23}^{ij}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{ij}\Big), \tag{4.39}$$

where

$$\omega_{21}^{ij} = \frac{1}{T_1^{ij}} \times n_2^T \mathbf{M}^{ij}\nu_1, \qquad \omega_{23}^{ij} = \frac{1}{T_1^{ij}} \times n_2^T \mathbf{M}^{ij}\nu_3,$$

with

$$T_1^{ij} = -\nu_3^T \mathbf{R}\nu_3.$$

Since the flux is continuous through edges, then using (4.33), (4.37) and (4.39) we have

$$\begin{cases} f_1 = \omega_{12}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \omega_{11}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big) \\ \quad = \omega_{13}^{i-1,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i-1,j-1}\Big) + \omega_{12}^{i-1,j-1}\Bigg(\mathcal{V}_{i,j-1} - \mathcal{V}_{i-1,j-1} + \chi_{42}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) \\ \qquad -\chi_{41}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big)\Bigg), \\ \\ f_2 = \omega_{22}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \omega_{21}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big) \\ \quad = -\omega_{21}^{ij}\Bigg(\mathcal{V}_{i,j-1} - \mathcal{V}_{ij} + \chi_{42}^{i,j-1}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{i,j-1}\Big) - \chi_{41}^{i,j-1}\Big(\bar{\mathcal{V}}_1 - \mathcal{V}_{i,j-1}\Big)\Bigg) + \omega_{23}^{ij}\Big(\bar{\mathcal{V}}_2 - \mathcal{V}_{ij}\Big). \end{cases} \tag{4.40}$$

by setting

$$g = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \qquad W = \begin{bmatrix} \mathcal{V}_{i-1,j-1} \\ \mathcal{V}_{i,j-1} \\ \mathcal{V}_{ij} \end{bmatrix}, \qquad \mathcal{V} = \begin{bmatrix} \bar{\mathcal{V}}_1 \\ \bar{\mathcal{V}}_2 \end{bmatrix}. \tag{4.41}$$

The system of equations (4.40) can be written as

$$g = C^{ij}\mathcal{V} + D^{ij}W, \tag{4.42}$$

where

$$C^{ij} = \begin{bmatrix} -\omega_{11}^{i,j-1} & \omega_{12}^{i,j-1} \\ -\omega_{21}^{i,j-1} & \omega_{22}^{i,j-1} \end{bmatrix}, \qquad D^{ij} = \begin{bmatrix} 0 & \omega_{11}^{i,j-1} - \omega_{12}^{i,j-1} & 0 \\ 0 & \omega_{21}^{i,j-1} - \omega_{22}^{i,j-1} & 0 \end{bmatrix}.$$

Using the expressions at both sides of the second equalities of system equations (4.40), it follows that

$$A^{ij}\mathcal{V} = B^{ij}W, \tag{4.43}$$

where

$$A^{ij} = \begin{bmatrix} \omega_{11}^{i,j-1} + \omega_{13}^{i-1,j-1} - \omega_{12}^{i-1,j-1}\chi_{41}^{i,j-1} & -\omega_{12}^{i,j-1} + \omega_{12}^{i-1,j-1}\chi_{42}^{i,j-1} \\ \omega_{21}^{i,j-1} + \omega_{21}^{ij}\chi_{41}^{i,j-1} & -\omega_{22}^{i,j-1} + \omega_{23}^{ij} - \omega_{21}^{ij}\chi_{42}^{i,j-1} \end{bmatrix},$$

$$B^{ij} = \begin{bmatrix} \omega_{13}^{i-1,j-1} + \omega_{12}^{i-1,j-1} & -\omega_{12}^{i,j-1} + \omega_{11}^{i,j-1} - \omega_{12}^{i-1,j-1}\big(1 + \chi_{41}^{i,j-1} - \chi_{42}^{i,j-1}\big) & 0 \\ 0 & -\omega_{22}^{i,j-1} + \omega_{21}^{i,j-1} + \omega_{21}^{ij}\big(1 + \chi_{41}^{i,j-1} - \chi_{42}^{i,j-1}\big) & -\omega_{21}^{ij} + \omega_{23}^{ij} \end{bmatrix}.$$

Thereby, by solving (4.43) with respect to $\mathcal{V}$ and replacing in (4.42) we get

$$g = R^{ij}\mathcal{V}, \tag{4.44}$$

where

$$R^{ij} = C^{ij}[A^{ij}]^{-1}B^{ij} + D^{ij}.$$

Now, considering the triangle T_2 (see Figure 4.3) and applying the above procedure used in the triangle T_1, we are able to the compute fluxes through the half edges 3 and 4 as follows:

$$h = S^{ij}V, \tag{4.45}$$

where

$$h = \begin{bmatrix} f_3 \\ f_4 \end{bmatrix}, \qquad V = \begin{bmatrix} \mathcal{V}_{i,j-1} \\ \mathcal{V}_{ij} \\ \mathcal{V}_{ij} \end{bmatrix}.$$

For simplicity, in an interaction volume $\mathcal{R}_{ij}$, the flux through the half edges $1, 2, 3$ and 4 are given by

$$f = T^{ij}\mathcal{V}, \tag{4.46}$$

where

$$f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}, \qquad \mathcal{V} = \begin{bmatrix} \mathcal{V}_{i-1,j-1} \\ \mathcal{V}_{i,j-1} \\ \mathcal{V}_{ij} \\ \mathcal{V}_{i-1,j} \end{bmatrix}, \tag{4.47}$$

and T^{ij} is 4×4 matrix coming from R^{ij} and S^{ij} defined in (4.44),(4.45). T^{ij} is called the transmissibility matrix of the interaction volume $\mathcal{R}_{ij}$.
Let us notice that the flux through a full edge will be the addition of the fluxes through its 2 half edges.

Let us recall that, from (4.27), our aim is to compute the flux through the edges of the control volume $\mathcal{C}_{ij}$. In order to cover a control volume, we need four interaction volumes.

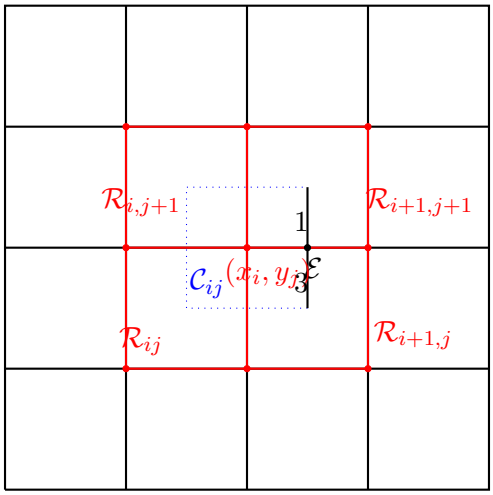


Figure 4.5: Control volume

Let us denote, for the volume control $\mathcal{C}_{ij}$, by εf_l^{ij} the flux through lower eastern half edge, by εf_u^{ij} the flux through the upper eastern half edge. The flux εf^{ij} through the eastern edge of the control volume $\mathcal{C}_{ij}$ is calculated as follows:
The lower eastern half edge is in position 3 in the triangle T_2 of the interaction volume $\mathcal{R}_{i+1,j}$ (see Figure (4.5)). So by using (4.46), it follows that

$$\varepsilon f_l^{ij} = T_{31}^{i+1,j}\mathcal{V}_{i,j-1} + T_{33}^{i+1,j}\mathcal{V}_{i+1,j} + T_{34}^{i+1,j}\mathcal{V}_{ij}.$$

Similarly, the upper half eastern edge is in position 1 in the triangle T_1 of the interaction volume $\mathcal{R}_{i+1,j+1}$ and it is in position 1 in the interaction volume. Thereby using (4.46), it follows that

$$\varepsilon f_u^{ij} = T_{11}^{i+1,j+1}\mathcal{V}_{ij} + T_{12}^{i+1,j+1}\mathcal{V}_{i+1,j} + T_{13}^{i+1,j+1}\mathcal{V}_{i+1,j+1}.$$

Finally, the flux through the eastern edge of the control volume $\mathcal{C}_{ij}$ will be the addition of εf_l^{ij} and εf_u^{ij} and it will be given as

$$\begin{aligned}\varepsilon f^{ij} &= \varepsilon f_d^{ij} + \varepsilon f_u^{ij} \\ \varepsilon f^{ij} &= T_{31}^{i+1,j}\mathcal{V}_{i,j-1} + \left(T_{33}^{i+1,j} + T_{12}^{i+1,j+1}\right)\mathcal{V}_{i+1,j} + \left(T_{11}^{i+1,j+1} + T_{34}^{i+1,j}\right)\mathcal{V}_{ij} + T_{13}^{i+1,j+1}\mathcal{V}_{i+1,j+1}.\end{aligned} \tag{4.48}$$

The same method is applied to calculate the flux through the northern, western and southern edges of the control volume $\mathcal{C}_{ij}$. The flux through the edges of the control volume $\mathcal{C}_{ij}$ is obtained by summing up the flux through the 4 edges. This gives :

$$\mathcal{F}_{ij} = a_{ij}\mathcal{V}_{ij} + b_{ij}\mathcal{V}_{i+1,j} + c_{ij}\mathcal{V}_{i+1,j+1} + d_{ij}\mathcal{V}_{i,j+1} + e_{ij}\mathcal{V}_{i-1,j} + \alpha_{ij}\mathcal{V}_{i-1,j-1} + \beta_{ij}\mathcal{V}_{i,j-1}, \tag{4.49}$$

where

$$a_{ij} = T_{11}^{i+1,j+1} + T_{34}^{i+1,j} + T_{41}^{i+1,j+1} + T_{22}^{i,j+1} - T_{12}^{i,j+1} - T_{33}^{ij} - T_{23}^{ij} - T_{44}^{i+1,j,}$$

$$b_{ij} = T_{33}^{i+1,j} + T_{12}^{i+1,j+1} - T_{43}^{i+1,j}, \qquad c_{ij} = T_{13}^{i+1,j+1} + T_{43}^{i+1,j+1} \qquad d_{ij} = T_{23}^{i,j+1} + T_{44}^{i+1,j+1} - T_{13}^{i,j+1},$$

$$e_{ij} = T_{21}^{i,j+1} - T_{11}^{i,j+1} - T_{34}^{ij}, \qquad \alpha_{ij} = -T_{31}^{ij} - T_{21}^{ij}, \qquad \beta_{ij} = T_{31}^{i+1,j} - T_{22}^{ij} - T_{41}^{i+1,j}.$$

This leads to a system of equations which can be written as follows:

$$\mathcal{F} = A_{mp}\mathcal{V} + F_{mp}, \tag{4.50}$$

where

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_{11} \\ \mathcal{F}_{12} \\ \vdots \\ \mathcal{F}_{1N} \\ \mathcal{F}_{21} \\ \mathcal{F}_{22} \\ \mathcal{F}_{NN} \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \mathcal{V}_{11} \\ \mathcal{V}_{12} \\ \vdots \\ \mathcal{V}_{1N} \\ \mathcal{V}_{21} \\ \mathcal{V}_{22} \\ \mathcal{V}_{NN} \end{bmatrix}, \quad A_{mp} = \begin{bmatrix} W_1 & X_1 & 0_N & \dots & \dots & \dots & \dots & 0_N \\ Y_2 & W_2 & X_2 & \ddots & & & & \vdots \\ 0_N & Y_3 & W_3 & X_3 & \ddots & & & \vdots \\ \vdots & \ddots & Y_4 & W_4 & X_4 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & & & & \ddots & Y_{N-1} & W_{N-1} & X_{N-1} \\ 0_N & \dots & \dots & \dots & \dots & 0_N & Y_N & W_N \end{bmatrix}.$$

with 0_N is $N \times N$ zeros matrix, W_i, X_i, Z_i are tridiagonal matrix and F_{mp} is a N^2 vector coming from the boundary conditions.
The diffusion matrix A_{mp} is under the following form:

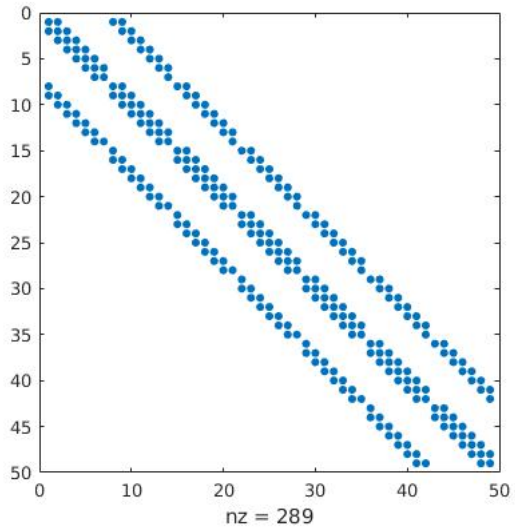


Figure 4.6: A structure of the diffusion matrix using L-MPFA method.

As we can see on Figure 4.6, the L-MPFA method is a 7 points stencil method, unlikely to the O-MPFA method (see Koffi and Tambue [2019c]) which is a 9 points stencil method.

4.3.2 Discretisation of the convection term

The integral of convection term

$$\int_{\mathcal{C}_{ij}} \nabla(f\mathcal{V})d\mathcal{C},$$

where

$$f = \begin{pmatrix} (r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2)x \\ \\ (r - \sigma_2^2 - \frac{1}{2}\rho\sigma_1\sigma_2)y \end{pmatrix},$$

will be approximated using the upwind methods (1^{st} and 2^{nd} order). We start by applying the divergence theorem as follows

$$I^{ij} = \int_{\mathcal{C}_{ij}} \nabla(f\mathcal{V})d\mathcal{C} = \int_{\partial\mathcal{C}_{ij}} (f \cdot \mathcal{V}) \cdot \vec{n}d\mathcal{C}, \qquad i,j = 1,\ldots,N, \tag{4.51}$$

with $\vec{n}$ an outward unit normal vector.

First order upwind method

The **first order upwind method** discussed in [LeVeque, 2004, chapter 4.8] and described in details in Section 3.3.1 of Chapter 3 will be applied to evaluate integral of the advection term given in (4.51).

I^{ij} is calculated by summing up the flux through the edges of the control volume $\mathcal{C}_{ij}$.
The flux through an edge using the first order upwind will depend on the sign of $f \cdot \vec{n}$ on this edge. If the sign of $f \cdot \vec{n}$ is positive, $\mathcal{V}_{ij}$ will be used to approximate $(f \cdot \vec{n}\mathcal{V})$ otherwise we will use the value of $\mathcal{V}$ in other side of the edge.
In doing so, we get

$$I^{ij} = \epsilon_{ij}\mathcal{V}_{i-1,j} + \mu_{ij}\mathcal{V}_{i,j-1} + \Omega_{ij}\mathcal{V}_{ij} + \phi_{ij}\mathcal{V}_{i,j+1} + \Psi_{ij}\mathcal{V}_{i+1,j}, \qquad i,j = 1,2,\ldots,N, \tag{4.52}$$

where

$$\epsilon_{ij} = -l_j f_x^{i-1} \max(f_x^{i-1}, 0); \qquad \mu_{ij} = -h_i f_y^{j-1} \max(f_y^{j-1}, 0),$$

$$\Omega_{ij} = l_j \Big(f_x^i \max(f_x^i, 0) - f_x^{i-1} \min(f_x^{i-1}, 0) \Big) + h_i \Big(f_y^j \max(f_y^j, 0) - f_y^{j-1} \min(f_y^{j-1}, 0) \Big),$$

$$\phi_{ij} = h_i f_y^j \min(f_y^j, 0), \qquad \Psi_{ij} = l_j f_x^i \min(f_x^i, 0).$$

Equation (4.52) will lead to a system of equations which will be written as follows

$$I = A_{up}\mathcal{V} + F_{up}, \tag{4.53}$$

where A_{up} is a $N^2 \times N^2$ matrix, I is a column vector of size N^2 defined by

$$I = \begin{bmatrix} I^{11} \\ I^{12} \\ \vdots \\ I^{1N} \\ I^{21} \\ I^{22} \\ \vdots \\ \vdots \\ I^{NN} \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \mathcal{V}_{11} \\ \mathcal{V}_{12} \\ \vdots \\ \mathcal{V}_{1N} \\ \mathcal{V}_{21} \\ \mathcal{V}_{22} \\ \vdots \\ \vdots \\ \mathcal{V}_{NN} \end{bmatrix}, \quad A_{up} = \begin{bmatrix} H_1 & P_1 & 0_N & \dots & \dots & \dots & 0_N \\ Q_2 & H_2 & P_2 & \ddots & & & \vdots \\ 0_N & Q_3 & H_3 & P_3 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & Q_{N-2} & H_{N-2} & P_{N-2} & 0_N \\ \vdots & & & \ddots & Q_{N-1} & H_{N-1} & P_{N-1} \\ 0_N & \dots & \dots & \dots & 0_N & Q_N & H_N \end{bmatrix},$$

with 0_N is $N \times N$ zeros matrix, H_i is a tridiagonal matrix, P_i, Q_i are diagonal matrices and F_{up} is a vector coming from the boundary conditions. Therefore, combining the L-MPFA method (4.50) and the first order upwind (4.53), we get

$$\frac{d\mathcal{V}}{dt} = A\mathcal{V} + G(\mathcal{V}) + F, \tag{4.54}$$

with

$$A = L^{-1}\Big(A_{mp} + A_{up} + A_L \Big), \quad G(\mathcal{V}) = \eta \Big[\max\Big(\mathcal{V}^* - \mathcal{V}, 0 \Big) \Big]^{1/k}, \quad F = L^{-1}\Big(F_{mp} + F_{up} \Big).$$

where A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (4.23). The diagonal elements of A_L are $A_{ii} = h_i l_i \lambda$ with λ given in (4.8). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = h_i l_i$ for $i = 1, 2, \dots, N^2$.

Second order upwind method

A second order upwind method approximation is used to calculated the flux defined in (4.51). Following the description of the method given in Section 3.3.2 of Chapter 3, the flux ${}_{\mathcal{E}}J^{ij}$ through the eastern edge of the control volume $\mathcal{C}_{ij}$ is given by

$${}_{\mathcal{E}}J^{ij} = \int_{\mathcal{E}_{ij}} (f \cdot \mathcal{V}) \cdot n_{\mathcal{E}}, \tag{4.55}$$

where

$$\vec{n}_{\mathcal{E}} = \begin{bmatrix} 1 \\ \\ 0 \end{bmatrix},$$

is the outward unit normal vector to the eastern edge $\mathcal{E}_{ij}$ of the control volume $\mathcal{C}_{ij}$. We set $f_x = f \cdot n_{\mathcal{E}}$ and we have

$$\mathcal{V} \approx \begin{cases} \frac{3\mathcal{V}_{ij} - \mathcal{V}_{i-1,j}}{2} & if \quad f_x \geq 0, \\ \\ \frac{3\mathcal{V}_{i+1,j} - \mathcal{V}_{i+2,j}}{2} & if \quad f_x < 0. \end{cases} \tag{4.56}$$

Then we get

$$_{\mathcal{E}}J^{ij} = l_j \left[\frac{3}{2} \max(f_x^{i+1}, 0)\mathcal{V}_{ij} - \frac{1}{2} \max(f_x^{i+1}, 0)\mathcal{V}_{i-1,j} + \frac{3}{2} \min(f_x^{i+1}, 0)\mathcal{V}_{i+1,j} - \frac{1}{2} \min(f_x^{i+1}, 0)\mathcal{V}_{i+2,j} \right], \tag{4.57}$$

with

$$f_x^{i+1} = (r - \sigma_x^2 - \frac{1}{2}\rho\sigma_x\sigma_y)x_{i+\frac{1}{2}}.$$

We use the same argument to calculate the flux $_{\mathcal{N}}J^{ij}, _{\mathcal{W}}J^{ij}, _{S}J^{ij}$ through the northern, western and southern edges of the control volume $\mathcal{C}_{ij}$ and after sum them up. We get then

$$\begin{aligned} J^{ij} &= \epsilon_{ij}\mathcal{V}_{i-2,j} + \eta_{ij}\mathcal{V}i - 1, j + \kappa_{ij}\mathcal{V}_{i,j-2} + \mu_{ij}\mathcal{V}_{i,j-1} + \Omega_{ij}\mathcal{V}_{ij} + \phi_{ij}\mathcal{V}_{i,j+1} + \Psi_{ij}\mathcal{V}_{i,j+2} + \Delta_{ij}\mathcal{V}_{i+1,j} \\ &\quad + \Pi_{ij}\mathcal{V}_{i+2,j}, \end{aligned} \tag{4.58}$$

where

$$\epsilon_{ij} = \frac{1}{2}l_j \max(f_x^i, 0), \qquad \eta_{ij} = -\frac{1}{2}l_j \max(f_x^{i+1}, 0) - \frac{3}{2}l_j \max(f_x^i, 0),$$

$$\kappa_{ij} = \frac{1}{2}h_i \max(f_y^j, 0), \qquad \mu_{ij} = -\frac{1}{2}h_i \max(f_y^{j+1}, 0) - \frac{3}{2}h_i \max(f_y^j, 0),$$

$$\Omega_{ij} = \frac{3}{2}l_j \max(f_x^{i+1}, 0) + \frac{3}{2}h_i \max(f_y^{j+1}, 0) - \frac{3}{2}l_j \min(f_x^i, 0) - \frac{3}{2}h_i \min(f_y^j, 0),$$

$$\phi_{ij} = \frac{3}{2}h_i \min(f_y^{j+1}, 0) + \frac{1}{2}h_i \min(f_y^j, 0), \qquad \Psi_{ij} = -\frac{1}{2}h_i \min(f_y^{j+1}, 0),$$

$$\Delta_{ij} = \frac{3}{2}l_j \min(f_x^{i+1}, 0) + \frac{1}{2}l_j \min(f_x^i, 0) \qquad \Pi_{ij} = -\frac{1}{2}l_j \min(f_x^{i+1}, 0).$$

For the control volumes near the boundary of the study domain, the first order upwind method is used for the approximation of the flux through edges directly connected to the boundary. Equation (4.58) leads to a system of equations which can be written as:

$$J = A_{2up}\mathcal{V} + F_{2up}, \tag{4.59}$$

where

$$J = \begin{bmatrix} J^{11} \\ J^{12} \\ \vdots \\ J^{1N} \\ J^{21} \\ J^{22} \\ \vdots \\ J^{NN} \end{bmatrix}, \qquad \mathcal{V} = \begin{bmatrix} \mathcal{V}_{11} \\ \mathcal{V}_{12} \\ \vdots \\ \mathcal{V}_{1N} \\ \mathcal{V}_{21} \\ \mathcal{V}_{22} \\ \vdots \\ \mathcal{V}_{NN} \end{bmatrix}, \quad A_{2up} = \begin{bmatrix} K_1 & R_1 & G_1 & 0_N & \dots & \dots & 0_N \\ S_2 & K_2 & R_2 & G_2 & \ddots & & \vdots \\ H_3 & S_3 & K_3 & R_3 & G_3 & \ddots & \vdots \\ 0_N & \ddots & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & \ddots & H_{N-2} & S_{N-2} & K_{N-2} & R_{N-2} & G_{N-2} \\ \vdots & & \ddots & H_{N-1} & S_{N-1} & K_{N-1} & R_{N-1} \\ 0_N & \dots & \dots & 0_N & H_N & S_N & K_N \end{bmatrix},$$

where for K_i is penta-diagonal matrice and R_i, G_i, S_i, H_i are diagonal matrices. F_{2up} is a vector coming from the boundary conditions.

Therefore, combining the L-MPFA method (4.50) and the second order upwind method (4.59) leads to

$$\frac{d\mathcal{V}}{dt} = A\mathcal{V} + G(\mathcal{V}) + F, \tag{4.60}$$

with

$$A = L^{-1}\bigg(A_{mp} + A_{2up} + A_L\bigg), \quad G(\mathcal{V}) = \eta\Big[\max\Big(\mathcal{V}^* - \mathcal{V}, 0\Big)\Big]^{1/k}, \quad F = L^{-1}\bigg(F_{mp} + F_{2up}\bigg),$$

where A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (4.23). The diagonal elements of A_L are $A_{ii} = h_i l_i \lambda$ with λ given in (4.8). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = h_i l_i$ for $i = 1, 2, \dots, N^2$.

Besides, the ellipticity condition for the PDE (4.2) is not satisfied when the stocks price ($x \to 0$ and/or $y \to 0$) is near to zero. This may cause some oscillations of the numerical solution when the PDE is degenerate.
Nevertheless, Wang [2004] suggested a fitted finite volume method to deal with the degeneracy of the PDE. Thereby, the fitted finite volume method will be applied in the degeneracy region ($x \to 0$ and/or $y \to 0$) in the next Section.

Fitted finite volume

The fitted finite volume method is used to approximated the flux through edges which are (fully) in the degeneracy region i.e the western edge of the control volume $\mathcal{C}_{1,j} \quad j = 1, .., N$ and the southern edge of the control volume $\mathcal{C}_{i,1} \quad i = 1, \dots, N$.

For the western edge of the control volume $\mathcal{C}_{1,j} \qquad j = 1, \dots, N$, using the mid-quadrature rule gives

$$\int_{(x_{\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial \mathcal{V}}{\partial x} + m_{12}\frac{\partial \mathcal{V}}{\partial y} + p\mathcal{V}\right) dy \approx \left(m_{11}\frac{\partial \mathcal{V}}{\partial x} + m_{12}\frac{\partial \mathcal{V}}{\partial y} + p\mathcal{V}\right)_{(x_{\frac{1}{2}}, y_j)} \cdot l_j. \tag{4.61}$$

Besides,

$$m_{11}\frac{\partial \mathcal{V}}{\partial x} + m_{12}\frac{\partial \mathcal{V}}{\partial y} + p\mathcal{V} \;=\; x\bigg(ax\frac{\partial \mathcal{V}}{\partial x} + d\frac{\partial \mathcal{V}}{\partial y} + b\mathcal{V}\bigg), \tag{4.62}$$

with $a = \frac{1}{2}\sigma_1^2$, $b = r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2$ and $d = \frac{1}{2}\rho\sigma_1\sigma_2 y$.

We want to approximate

$$g(\mathcal{V}) = ax\frac{\partial \mathcal{V}}{\partial x} + b\mathcal{V},$$

by a constant over $I_{x_1} = (0, x_1)$ satisfying the following two-point boundary value problem

$$\begin{cases} g'(v) \quad = \left(ax\frac{\partial v}{\partial x} + bv\right)' = K_1, \\ \\ v(0, y_j) = \mathcal{V}_{0,j} \qquad v(x_1, y_j) = \mathcal{V}_{1,j}. \end{cases} \tag{4.63}$$

By solving this problem, we get

$$\mathcal{V} = \mathcal{V}_{0,j} + (\mathcal{V}_{1,j} - \mathcal{V}_{0,j})\frac{x}{x_1}. \tag{4.64}$$

Thereby, using (4.61),(4.62), (4.63), (4.64) and the forward difference to approximate the first partial derivative $\frac{\partial \mathcal{V}}{\partial y}$ lead to

$$\int_{(x_{\frac{1}{2}}, y_{j-\frac{1}{2}})}^{(x_{\frac{1}{2}}, y_{j+\frac{1}{2}})} \left(m_{11}\frac{\partial \mathcal{V}}{\partial x} + m_{12}\frac{\partial \mathcal{V}}{\partial y} + p\mathcal{V}\right)dy \approx \frac{1}{2}x_1\Big[\frac{1}{2}l_j(a+b) - d_j\Big]\mathcal{V}_{i,1} + \frac{1}{2}d_j x_1 \mathcal{V}_{1,j+1} - \frac{1}{4}x_1 l_j(a-b)\mathcal{V}_{0,j}, \tag{4.65}$$

where

$$a = \frac{1}{2}\sigma_1^2, \qquad b = r - \sigma_1^2 - \frac{1}{2}\rho\sigma_1\sigma_2, \qquad d_j = \frac{1}{2}\rho\sigma_1\sigma_2 y_j, \qquad l_j = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}.$$

Similarly, for the flux through the southern edge of the control volume $\mathcal{C}_{i,1} \qquad i = 1, ..., N,$

$$\int_{(x_{i-\frac{1}{2}}, y_{\frac{1}{2}})}^{(x_{i+\frac{1}{2}}, y_{\frac{1}{2}})} \left(m_{21}\frac{\partial \mathcal{V}}{\partial x} + m_{22}\frac{\partial \mathcal{V}}{\partial y} + q\mathcal{V}\right)dx \approx \frac{1}{2}y_1\Big[\frac{1}{2}h_i(e+k) - h_i'\Big]\mathcal{V}_{i,1} + \frac{1}{2}h_i' y_1 \mathcal{V}_{i+1,1} - \frac{1}{4}y_1 h_i(e-k)\mathcal{V}_{i,0}, \tag{4.66}$$

where

$$e = \frac{1}{2}\sigma_2^2, \qquad k = r - \sigma_2^2 - \frac{1}{2}\rho\sigma_1\sigma_2, \qquad h_i' = \frac{1}{2}\rho\sigma_1\sigma_2 x_i, \qquad h_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}.$$

The fitted L-Multi-Point Flux Approximation method (with the 1^{st} order upwind method)

1. Fitted L-MPFA method (with 1^{st} order upwind method)

   Here the fitted finite volume method is combined with the first order upwind method . Thereby we have:

   For the control volume $\mathcal{C}_{11}$, the western and southern edges are (fully) in the degeneracy region. The integrals over the western and southern edges of the control volume $\mathcal{C}_{11}$ are then approximated using the fitted finite volume (4.65) and (4.66). The integrals over the eastern and northern edges of the control $\mathcal{C}_{11}$, which are not in the degeneracy region, are approximated using the L-MPFA method coupled to the upwind method (1^{st} and 2^{nd} order).

$$\int_{\mathcal{C}_{11}} \nabla k(\mathcal{V}) d\mathcal{C}_{11} \quad \approx \quad aa_{11}\mathcal{V}_{11} + bb_{11}\mathcal{V}_{21} + cc_{11}\mathcal{V}_{22} + dd_{11}\mathcal{V}_{12} + ee_{11}\mathcal{V}_{01} + \beta\beta_{11}\mathcal{V}_{10},$$

where

$$aa_{11} = T_{11}^{22} + T_{34}^{21} + T_{41}^{22} + T_{22}^{12} + l_1 \max(f_x^2, 0) + h_1 \max(f_y^2, 0) - \frac{1}{2}x_1\Big(\frac{1}{2}l_1(a+b) - d_1\Big)$$
$$-\frac{1}{2}y_1\Big(\frac{1}{2}h_1(e+k) - h_1'\Big),$$

$$bb_{11} = T_{33}^{21} + T_{12}^{22} + l_1 \min(f_x^2, 0) - \frac{1}{2}h_1' y_1, \qquad cc_{11} = T_{13}^{22} + T_{43}^{22},$$

$$dd_{11} = T_{23}^{12} + T_{44}^{22} + h_1 \min(f_y^2, 0) - \frac{1}{2}d_1 x_1, \qquad ee_{11} = T_{21}^{12} + \frac{1}{4}l_1 x_1 (a-b),$$

$$\beta\beta_{11} = T_{31}^{21} + \frac{1}{4}y_1 h_1 (e-k).$$

Similarly, for the control volume $\mathcal{C}_{1,j} \quad j = 1, \ldots, N$, only the southern edge is (fully) in the degeneracy region. Then the integral over this edge, is approximated using the fitted finite volume method (4.66). The integrals over the eastern, northern and western edges are approximated using the L-MPFA method coupled to the upwind methods(1^{st} and 2^{nd} order)

$$\begin{aligned}\int_{C_{1,j}} \nabla \mathbf{k}(\mathcal{V}) d\mathcal{C}_{1j} \quad &= \quad aa_{1,j}\mathcal{V}_{1,j} + bb_{1,j}\mathcal{V}_{2,j} + cc_{1,j}\mathcal{V}_{2,j+1} + dd_{1,j}\mathcal{V}_{1,j+1} + \beta\beta_{1,j}\mathcal{V}_{1,j-1} \\ &\quad + \alpha\alpha_{1,j}\mathcal{V}_{0,j-1} + ee_{1,j}\mathcal{V}_{0,j},\end{aligned} \tag{4.67}$$

where

$$aa_{1,j} = T_{11}^{2,j+1} + T_{34}^{2,j} + T_{41}^{2,j+1} + T_{22}^{1,j+1} - T_{23}^{1,j} - T_{44}^{2,j} + l_j \max(f_x^2, 0) + h_1 \max(f_y^{j+1}, 0)$$
$$-h_1 \min(f_y^j, 0) - \frac{1}{2}x_1\Big(\frac{1}{2}l_j(a+b) - d_j\Big),$$

$$bb_{1,j} = T_{33}^{2,j} + T_{12}^{2,j+1} - T_{43}^{2,j} + l_j \min(f_x^2, 0), \qquad cc_{1,j} = T_{13}^{2,j+1} + T_{43}^{2,j+1},$$

$$dd_{1,j} = T_{23}^{1,j+1} + T_{44}^{2,j+1} + h_1 \min(f_y^{j+1}, 0) - \frac{1}{2}d_j x_1, \qquad \beta\beta_{1,j} = T_{31}^{2,j} - T_{22}^{1,j} - T_{41}^{2,j} - h_1 \max(f_y^j, 0),$$

$$\alpha\alpha_{1,j} = -T_{21}^{1,j}, \qquad ee_{1,j} = T_{21}^{1,j+1} + \frac{1}{4}l_j x_1 (a-b).$$

Using the same argument as above, for the control volume $\mathcal{C}_{i,1} \quad i = 2, .., N$, the integral over the southern edge is approximated using the fitted finite volume (4.66). The integrals over the eastern, northern and western edges are approximated using the L-MPFA method combined with the upwind methods (1^{st} and 2^{nd} order)

$$\begin{aligned}\int_{\mathcal{C}_{i,1}} \nabla \mathbf{k}(\mathcal{V}) d\mathcal{C}_{i1} &= aa_{i,1}\mathcal{V}_{i,1} + bb_{i,1}\mathcal{V}_{i+1,1} + cc_{i,1}\mathcal{V}_{i+1,2} + dd_{i,1}\mathcal{V}_{i,2} + ee_{i,1}\mathcal{V}_{i-1,1} + \alpha\alpha_{i,1}\mathcal{V}_{i-1,0} \\ &\quad + \beta\beta_{i,1}\mathcal{V}_{i,0}.\end{aligned} \tag{4.68}$$

where

$$aa_{i,1} = T_{11}^{i+1,2} + T_{34}^{i+1,1} + T_{41}^{i+1,2} + T_{22}^{i,2} - T_{12}^{i,2} - T_{33}^{i,1} + l_1 \max(f_x^{i+1}, 0) + h_i \max(f_y^2, 0),$$

$$-l_1 \min(f_x^i, 0) - \frac{1}{2} y_1 \Big(\frac{1}{2} h_i (e + k) - h_i' \Big),$$

$$bb_{i,1} = T_{33}^{i+1,1} + T_{12}^{i+1,2} + l_1 \min(f_x^{i+1}, 0) - \frac{1}{2} h_i' y_1, \qquad cc_{i,1} = T_{13}^{i+1,2} + T_{43}^{i+1,2},$$

$$dd_{i,1} = T_{23}^{i,2} + T_{44}^{i+1,2} - T_{13}^{i,2} + h_i \min(f_y^2, 0), \qquad ee_{i,1} = T_{21}^{i,2} - T_{11}^{i,2} - T_{34}^{i,1} - l_1 \max(f_x^i, 0).$$

$$\alpha\alpha_{i,1} = -T_{31}^{i,1}, \qquad \beta\beta_{i,1} = T_{31}^{i+1,1} + \frac{1}{4} y_1 h_i (e - k).$$

Besides, for the control volume $\mathcal{C}_{ij}, \quad i, j = 2, ..., N$, the L-MPFA method is used to approximate the diffusion term and the upwind to approximate the advection term. This leads to the following semi-discrete equation

$$\frac{d\mathcal{V}}{dt} = A\mathcal{V} + G(\mathcal{V}) + F, \tag{4.69}$$

where

$$\mathcal{V} = \begin{bmatrix} \mathcal{V}_{11} \\ \mathcal{V}_{12} \\ \vdots \\ \mathcal{V}_{1N} \\ \mathcal{V}_{21} \\ \mathcal{V}_{22} \\ \vdots \\ \mathcal{V}_{NN} \end{bmatrix}, \quad A = L^{-1} \Big(Z + A_L \Big), \quad G(\mathcal{V}) = \eta \Big[\max \Big(\mathcal{V}^* - \mathcal{V}, 0 \Big) \Big]^{1/k},$$

with F the vector of boundary conditions. A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (4.23). The diagonal elements of A_L are $A_{ii} = h_i l_i \lambda$ for $i = 1, ..., N^2$ with λ given in (4.8). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = h_i l_i$ for $i = 1, \ldots, N^2$ and

$$Z = \begin{bmatrix} D_1 & K_1 & 0_N & \dots & \dots & \dots & \dots & 0_N \\ L_2 & D_2 & K_2 & \ddots & & & & \vdots \\ 0_N & L_3 & D_3 & K_3 & \ddots & & & \vdots \\ \vdots & \ddots & L_4 & D_4 & K_4 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & 0_N \\ \vdots & & & & \ddots & L_{N-1} & D_{N-1} & K_{N-1} \\ 0_N & \dots & \dots & \dots & \dots & 0_N & L_N & D_N \end{bmatrix}.$$

The elements of matrix Z are matrices. 0_N is a zeros matrix of size $N \times N$. The matrices D_i, K_i, L_i are tri-diagonal matrices defined as follows:

$$for \;\; i = 1 \; or \; i = N,$$

$$k = 1, \dots, N \;\; (D_i)_{kk} = aa_{1,k}, \qquad k = 1, \dots, N-1 \qquad (D_i)_{k,k+1} = dd_{1,k},$$

$$k = 2, \dots, N \;\; (D_i)_{k,k-1} = \beta\beta_{1,k},$$

$$for \; i = 1 \qquad k = 1, \dots, N \;\; (K_1)_{kk} = bb_{1,k} \qquad k = 1, \dots, N-1 \;\; (K_1)_{k,k+1} = cc_{1,k},$$

$$for \; i = N, \;\; (L_N)_{11} = ee_{N,1}; \;\; k = 2, \dots, N \; (L_N)_{kk} = e_{N,k} + \eta_{N,k}, \quad k = 2, \dots, N \quad (L_N)_{k,k+1} = \alpha_{N,k},$$

$$for \;\; i = 2, \dots, N-1$$

$$(D_i)_{11} = aa_{i,1} \; ; \; (D_i)_{12} = dd_{i,1}, \qquad (K_i)_{11} = bb_{i,1} \; , \; (K_i)_{12} = cc_{i,1}, \qquad (L_i)_{11} = ee_{i,1},$$

$$k = 2, \dots, N \qquad (D_i)_{kk} = a_{i,k} + \Omega_{i,k}, \qquad (K_i)_{kk} = b_{i,k} + \Delta_{i,k}, \qquad (L_i)_{kk} = e_{i,k} + \eta_{i,k},$$

$$k = 2, \dots, N-1 \qquad (D_i)_{k,k+1} = d_{i,k} + \phi_{i,k}, \qquad (K_i)_{k,k+1} = c_{i,k},$$

$$k = 2, \dots, N \qquad (D_i)_{k,k-1} = \beta_{i,k} + \mu_{i,k}, \qquad (L_i)_{k,k-1} = \alpha_{i,k}.$$

where the elements $aa_{ij}, bb_{ij}, cc_{ij}, dd_{ij}, ee_{ij}, \beta\beta_{ij}$ are defined in (4.67),(4.67),(4.68), and the elements $a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij}, \Omega_{ij} \; \Delta_{ij}, \beta_{ij}, \phi_{ij}, \alpha_{ij}, \mu_{ij}, \eta_{ij}$ are defined in (4.49) and (4.52).

2. Fitted Multi-Point Flux Approximation (2^{nd} order upwind)

Similarly, the fitted L-MPFA method deriving from the combination of the L-MPFA method and the 2^{nd} order upwind method leads to the following equation :

$$\frac{d\mathcal{V}}{dt} = A\mathcal{V} + G(\mathcal{V}) + F, \tag{4.70}$$

where

$$\mathcal{V} = \begin{bmatrix} \mathcal{V}_{11} \\ \mathcal{V}_{12} \\ \vdots \\ \mathcal{V}_{1N} \\ \mathcal{V}_{21} \\ \mathcal{V}_{22} \\ \vdots \\ \mathcal{V}_{NN} \end{bmatrix}, \quad A = L^{-1}\Big(Y + A_L\Big), \quad G(\mathcal{V}) = \eta\Big[\max\Big(\mathcal{V}^* - \mathcal{V}, 0\Big)\Big]^{1/k}.$$

with F the vector of boundary conditions. A_L is a diagonal matrix of size $N^2 \times N^2$ coming from the discretisation of (4.23). The diagonal elements of A_L are $A_{ii} = h_i l_i \lambda$ with λ given in (4.8). The matrix L is also a diagonal matrix of size $N^2 \times N^2$ whose diagonal elements are $L_{ii} = h_i l_i$ for $i = 1, \ldots, N^2$,

and

$$Y = \begin{bmatrix} H_1 & P_1 & R_1 & 0_N & \ldots & \ldots & \ldots & & 0_N & 0_N \\ Q_2 & H_2 & P_2 & R_2 & 0_N & & & & & 0_N \\ W_3 & Q_3 & H_3 & P_3 & R_3 & 0_N & & & & \vdots \\ 0_N & W_4 & Q_4 & H_4 & P_4 & R_4 & \ddots & & & \\ 0_N & 0_N & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0_N \\ & & & & \ddots & W_{N-2} & Q_{N-2} & H_{N-2} & P_{N-2} & R_{N-2} \\ \vdots & & & & & \ddots & W_{N-1} & Q_{N-1} & H_{N-1} & P_{i,N-1} \\ 0_N & \ldots & \ldots & \ldots & & \ldots & 0_N & W_N & Q_N & H_N \end{bmatrix}.$$

The elements of matrix Y are matrices. 0_N is a zeros matrix of size N. The matrices H_i are a penta-diagonal matrices and the matrices P_i, R_i, Q_i, W_i are diagonal matrices.

Furthermore, The θ- Euler method will be applied on the semi-discrete equations (4.54),(4.60), (4.69) and(4.70) for the spatial discretisation.

4.4 Time discretization

Let us consider the ODE stemming from the spatial discretization and given by (4.54),(4.60),(4.69) and(4.70)

$$\frac{d\mathcal{V}}{dt} = A\mathcal{V} + G(\mathcal{V}) + F,$$

By using the θ-Euler method for the time discretization, we have

$$for\ m = 0, \ldots, M-1,$$

$$\frac{\mathcal{V}^{m+1} - \mathcal{V}^m}{\Delta t} = \theta\Big(A\mathcal{V}^{m+1} + G(\mathcal{V}^{m+1}) + F(t_{m+1}\Big) + (1-\theta)\Big(A\mathcal{V}^m + G(\mathcal{V}^m) + F(t_m\Big), \tag{4.71}$$

At every time iteration, the nonlinear system where $\mathcal{V}^{m+1}$ is the solution, is solved using the Newton method. Note that

$$\mathcal{V}^m = \left[\mathcal{V}_{11}(t_m) \;\; \mathcal{V}_{12}(t_m) \;\; \ldots \mathcal{V}_{1N}(t_m) \;\; \mathcal{V}_{21}(t_m) \;\; \ldots \mathcal{V}_{2N}(t_m) \;\; \ldots \mathcal{V}_{N,1}(t_m) \;\; \ldots\ldots \;\; \mathcal{V}_{NN}(t_m)\right]^T,$$

$$t_m = m\Delta t.$$

where the time step is $\Delta t = \frac{T}{M}$, T being the maturity time.

4.5 Numerical experiments

In this Section, we perform some numerical simulations for the L-MPFA method combined to the upwind methods (first and second order) and for the fitted L-MPFA method combined to the upwind methods (first and second order).

4.5.1 Errors for European call options

The computational domain of the problem is $\Omega = [0;300] \times [0;300] \times [0;T]$ with T=1/12. The numerical experiments are performed with the strike price $E = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$, the correlation coefficient $\rho = 0.3$ and the risk free interest $r = 0.08$

Here, by taking $\beta = 0$ in (4.13), the L-MPFA method illustrated in the previous sections will be compared to the fitted finite volume method, Wang [2004], and the fitted O-MPFA methods for pricing multi-asset options for pricing options introduced in Koffi and Tambue [2019c]. The relative error will be computed with respect to the analytical solution of the Black-Scholes PDE defined in Haug [2007] as follows

$$\begin{aligned} C(x,y,K,T) &= xe^{-rT}M(z_1,d;\rho_1) + ye^{-rT}M(z_2,-d+\sigma\sqrt{T};\rho_2) \\ &\quad -Ke^{-rT} \times \left(1 - M(-z_1+\sigma_1\sqrt{T}, -z_2+\sigma_2\sqrt{T};\rho)\right), \end{aligned} \tag{4.72}$$

where

$$d = \frac{\ln(x/y) + (b_1 - b_2 + \sigma_1^2/2)T}{\sigma\sqrt{T}},$$

$$z_1 = \frac{\ln(x/K) + (b_1 + \sigma_1^2/2)T}{\sigma_1\sqrt{T}}, \qquad z_2 = \frac{\ln(y/K) + (b_1 + \sigma_2^2/2)T}{\sigma_2\sqrt{T}},$$

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, \qquad \rho_1 = \frac{\sigma_1 - \rho\sigma_2}{\sigma}, \qquad \rho_2 = \frac{\sigma_2 - \rho\sigma_1}{\sigma}.$$

and

$$M(a,b;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{a}\int_{-\infty}^{b} exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dxdy.$$

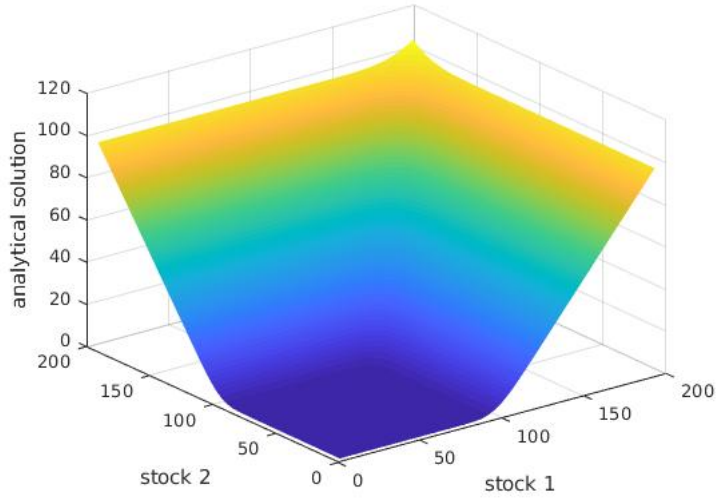


Figure 4.7: Analytical solution

The solution using the L-MPFA coupled to the 2^{nd} order upwind method is illustrated as below

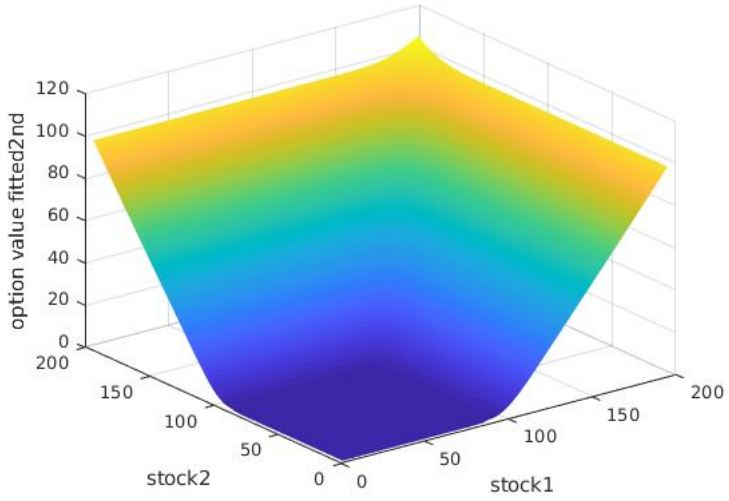


Figure 4.8: L-MPFA -upwind 2^{nd} order

The L^2-norm is used to compute the error is

$$err = \frac{\sqrt{\sum_{i,j=1}^{N} meas(\mathcal{C}_{ij})\left(\mathcal{V}_{ij} - V_{ij}^{ana}\right)^2}}{\sqrt{\sum_{i=1}^{n} meas(\mathcal{C}_{ij})\left(V_{ij}^{ana}\right)^2}}, \tag{4.73}$$

where $\mathcal{V}$ is the numerical solution, V^{ana} the analytical solution and $meas(\mathcal{C}_{ij})$ is the measure of the control volume $\mathcal{C}_{ij}$. This gives the following tables:

Table 4.1: Table of errors

	Fitted fin vol	O-MPFA-1^{st} upw	O-MPFA-2^{nd} upw	fit O-MPFA-1^{st} upw	fit O-MPFA -2^{nd} upw
50×50	0.0317	0.0224	0.0225	0.0212	0.0212
70×70	0.0329	0.0248	0.0248	0.0238	0.0238
85×85	0.0327	0.0260	0.0260	0.0251	0.0251

As we can observe in Table 4.1 and Table 4.2, the new fitted L-MPFA method is more accurate

Table 4.2: Table of errors

	L-MPFA-1stupw	L-MPFA-2nd upw	fit L-MPFA-1st upw	fit L-MPFA -2nd upw
50×50	0.0048	0.0049	0.0048	0.0047
70×70	0.0041	0.0041	0.0041	0.0041
85×85	0.0040	0.0040	0.0040	0.0040

than the fitted O-MPFA method developed in Koffi and Tambue [2019c] and the standard fitted finite volume method developed in Huang et al. [2006].

4.5.2 Errors for American put options

Since there is no analytical solution for the power penalty problem (4.13) for pricing American put options, and the numerical solution given by the fitted L-MPFA coupled to 2^{nd} order upwind method is more accurate when pricing European options (see Table 4.1 and Table 4.2), we have chosen for reference solution or "exact solution" the numerical solution given by the fitted L-MPFA coupled to 2^{nd} order upwind method with $dt = T/256$. The relative error of all the numerical methods used in this study will be performed with respect to this reference solution.

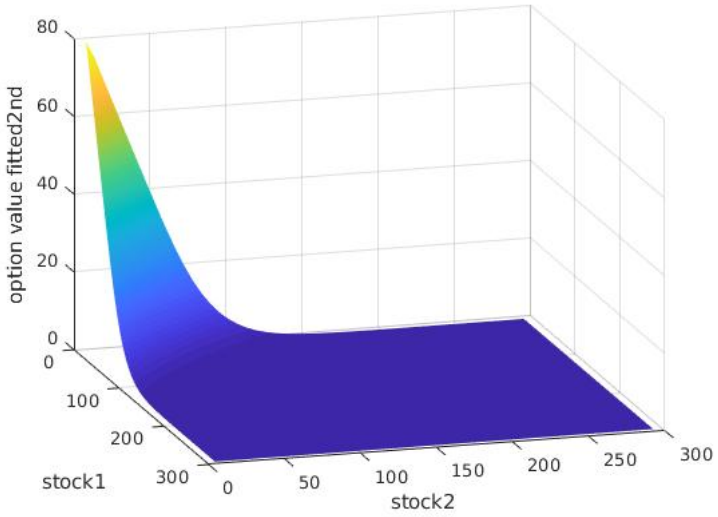


Figure 4.9: Reference solution

For the numerical simulations below, the computational domain of the problem is $\Omega = [0; 300] \times [0; 300] \times [0; T]$ with $T = 1/6, K = 100$, the volatilities $\sigma_1 = \sigma_2 = 0.3$. The correlation coefficient is $\rho = 0.3$, the risk free interest r = 0.08. The penalty parameter $\beta = 256$ and the power penalty $k = 1/2$.

Table 4.3: Table of errors for $\Delta t = T/64$

	Fitted Fin vol	L-MPFA-1stupw	L-MPFA-2nd upw	fit L-MPFA-1st upw	fit L-MPFA -2nd upw
50×50	0.0616	0.0610	0.0583	0.0611	0.0584
60×60	0.0277	0.0278	0.0276	0.0278	0.0277
70×70	0.0184	0.0183	0.0182	0.0182	0.0180
80×80	0.0104	0.0100	0.0098	0.0097	0.0095

Table 4.4: Table of errors for $\Delta t = T/128$

	Fitted Fin vol	L-MPFA-1^{st}upw	L-MPFA-2^{nd} upw	fit L-MPFA-1^{st} upw	fit L-MPFA -2^{nd} upw
50×50	0.0599	0.0520	0.0476	0.0522	0.0459
60×60	0.0227	0.0265	0.0249	0.0241	0.0220
70×70	0.0136	0.0148	0.0146	0.0146	0.0144
80×80	0.0087	0.0068	0.0065	0.0062	0.0059

Again we can observe in Table 4.3 and Table 4.4, the novel fitted L-MPFA coupled to the 2^{nd} order upwind method is more accurate than the standard fitted finite volume by Huang et al. [2006].

Conclusion

In this Chapter, the L-MPFA methods have been introduced to approximate the diffusion term of the Black-Scholes PDE. The upwind methods (1^{st} and 2^{nd} order) are used for space discretization of the convection term appearing in the two dimensional Black-Scholes PDE. We have provided novel schemes called the fitted L-MPFA method to handle the degeneracy of the Black-Scholes PDE by combining the standard fitted finite volume and the L-MPFA method coupled to the upwind methods. Numerical experiments are performed and comparison between the L-MPFA methods, the O-MPFA methods developed in Chapter 3 and the fitted finite methods by Huang et al. [2006] are performed. The results have shown that the fitted L-MPFA method coupled to the 2^{nd} order upwind method is more accurate than the fitted finite volume by Huang et al. [2006] and the O-MPFA methods developed in chapter 3 for pricing Europeans and American options.

Conclusion

In this Thesis, our goal was to develop numerical methods to solve accurately the degenerated Black-Scholes PDE for option pricing. The first step was, in Chapter 1, to prove the existence and uniqueness of the solution of the continuous degenerated multi-dimensional option pricing problem using weighted Sobolev spaces. Afterwards, in Chapter 2, we presented the TPFA method and the fitted TPFA methods. Despite the fact that the degeneracy of the Black-Scholes PDE does not allow a lower bound to the transmissibility coefficient of the TPFA method and fitted TPFA method, we were able to prove the flux consistency and the error estimate using appropriate Taylor expansion around zero. This led to the convergence proof of the TPFA and the fitted TPFA method for the degenerated Black-Scholes PDE.
Moreover, in Chapter 3, we introduced the MPFA method, the main numerical scheme of this work which helped to discretise the diffusion term of Black-Scholes PDE in its divergence form. The advection term was discretised using the upwind methods (first order and second order). To handle the degeneracy of the Black-Scholes PDE, we developed a novel numerical scheme called fitted MPFA method, which is a combination of the fitted finite volume method and the MPFA method. The fitted finite volume method was applied in the region where the stock prices (see Wang [2004]) are close to zero (degeneracy region) and over the rest of the study domain, the MPFA method coupled to upwinds methods were applied. The MPFA method used here was the O-MPFA methods. Numerical experiments have shown that the fitted O-MPFA methods coupled the upwind methods are more accurate than the fitted finite volume method.
Finally, in Chapter 4, we presented the L-MPFA method, which is another kind of MPFA method. The L-MPFA method is less intuitive and less computionally expensive than the O- MPFA method used in Chapter 3. Using a similar approach as the one in Chapter 3, we have developed the fitted L-MPFA methods to solve the degenerated Black-Scholes PDE for pricing american options. Numerical results have shown that the fitted L-MPFA methods coupled to upwind methods are more accurate than standard fitted finite volume method and O-MPFA methods introduced in Chapter 3.
We should note that the TPFA method is the one dimensional version of the MPFA methods. Thereby, the convergence proof for the O-MPFA methods and L-MPFA methods will be an extension of the convergence proof of the TPFA methods provided in Chapter 2. More precisely, we need to establish relation between the transmissibility coefficients and define appropriate weighted norm and bilinear forms corresponding to option pricing problem.

Our future work will be to develop new fitted schemes with high order of accuracy along with their rigorous convergence proofs. More precisely, we will develop the fitted MPFA schemes for pricing problems in dimension 3 and will also provide rigorous convergence proofs. Another advantage of our novel fitted MPFA methods is that it can easily be adapted to more structured commercial or open source software as the standard MPFA (see Lie et al. [2012]), thereby we will also build a MPFA and fitted MPFA toolbox for computational finance.

Bibliography

Ivar Aavatsmark. An introduction to multipoint flux approximation for quadrilateral grids. *Computational Geosciences*, 6(3-4):405–432, 2002.

Ivar Aavatsmark. Multipoint flux approximation methods for quadrilateral grids. In *9th International forum on reservoir simulation, Abu Dhabi*, pages 9–13, 2007.

Lutz Angermann and Song Wang. Convergence of a fitted finite volume method for the penalized black–scholes equation governing european and american option pricing. *Numerische Mathematik*, 106(1):1–40, 2007.

David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.

Daniel J Duffy. *Finite Difference methods in financial engineering: a Partial Differential Equation approach*. John Wiley & Sons, 2013.

Lawrence C. Evans. *Partial Differential Equations*, volume 19. American Mathematical Society, 1997.

Robert Eymard, Thierry Gallouët, and Raphaèle Herbin. Finite volume methods. *Handbook of numerical analysis*, 7:713–1018, 2000.

Jaroslav Haslinger, Markku Miettinen, and Panagiotis D Panagiotopoulos. *Finite element method for Hemivariational inequalities: theory, methods and applications*, volume 35. Springer Science & Business Media, 2013.

Espen Gaarder Haug. *The complete guide to option pricing formulas*, volume 2. McGraw-Hill New York, 2007.

Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

C-S Huang, C-H Hung, and Song Wang. A fitted finite volume method for the valuation of options on assets with stochastic volatilities. *Computing*, 77(3):297–320, 2006.

C-S Huang, C-H Hung, and Song Wang. On a convergence of a fitted finite-volume method for the valuation of options on assets with stochastic volatilities. *IMA journal of numerical analysis*, 30(4): 1101–1120, 2009.

Jacques Janssen and Raimondo Manca. *Semi-Markov risk models for finance, insurance and reliability*. 2007.

Rock S Koffi and Antoine Tambue. Convergence of the two point flux approximation and a novel fitted two-point flux approximation method for pricing options. *arXiv preprint arXiv:1912.12737*, 2019a.

Rock S Koffi and Antoine Tambue. A fitted l-multi-point flux approximation method for pricing options. *arXiv preprint arXiv:1912.1274v1*, 2019b.

Rock Stephane Koffi and Antoine Tambue. A fitted multi-point flux approximation method for pricing two options. *Computational Economics*, 2019c. doi: 10.1007/s10614-019-09906-x. URL `https://doi.org/10.1007/s10614-019-09906-x`.

Peter E. Kopp. *From measures to ITO integrals*. AIMS library series, 2011.

Pavlo Kovalov, Vadim Linetsky, and Michael Marcozzi. Pricing multi-asset american options: A finite element method-of-lines with smooth penalty. *Journal of Scientific Computing*, 33(3):209–237, 2007.

Randall J LeVeque. Finite volume methods for hyperbolic problems. *Cambridge Texts in Applied Mathematics*, 39(1):88–89, 2004.

Knut-Andreas Lie, Stein Krogstad, Ingeborg Skjelvåle Ligaarden, Jostein Roald Natvig, Halvor Møll Nilsen, and Bård Skaflestad. Open-source matlab implementation of consistent discretisation on complex grids. *Computational Geosciences*, 16(2):297–322, 2012.

Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.

Bernt Oksendal. *Stochastic differential equations*. Verlage Publication, 1992.

L. Angermann P. Knabner. *Numerical Methods for Elliptic and Parabolic Partial Differential Equations*, volume 44. Springer, 2002.

Jonas Persson and Lina von Persson. Pricing european multi-asset options using a space-time adaptive fd-method. *Computing and Visualisation in Science*, 10(4):173–183, 2007.

Tor Harald Sandve, Inga Berre, and Jan M Nordbotten. An efficient multi-point flux approximation method for discrete fracture–matrix simulations. *Journal of Computational Physics*, 231(9):3784–3800, 2012.

Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer Science & Business Media, 2004.

Annette F Stephansen. Convergence of the multipoint flux approximation l-method on general grids. *SIAM Journal on Numerical Analysis*, 50(6):3163–3187, 2012.

Antoine Tambue. An exponential integrator for finite volume discretization of a reaction–advection–diffusion equation. *Computers & Mathematics with Applications*, 71(9):1875–1897, 2016.

Peter Tankov. *Financial modelling with jump processes*. Chapman and Hall/CRC, 2003.

Jürgen Topper. *Financial engineering with finite elements*, volume 319. John Wiley & Sons, 2005.

S Wang, XQ Yang, and KL Teo. Power penalty method for a linear complementarity problem arising from american option valuation. *Journal of Optimization Theory and Applications*, 129(2):227–254, 2006.

Song Wang. A novel fitted finite volume method for the black–scholes equation governing option pricing. *IMA Journal of Numerical Analysis*, 24(4):699–720, 2004.

Paul Wilmott. *The Best of Wilmott 1: Incorporating the Quantative Finance Review*. John Wiley & Sons, 2005.

K Zhang, S Wang, X Yang, and Kok Lay Teo. A power penalty approach to numerical solutions of two-asset american options. *Numerical Mathematics: Theory, Methods and Applications*, 2:202–223, 2009.