Monotone Numerical Methods for Nonlinear Systems and Second Order Partial Differential Equations

by

Swathi Amarala

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2015

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Multigrid methods are numerical solvers for partial differential equations (PDEs) that systematically exploit the relationship between approximate solutions on multiple grids to arrive at a solution whose accuracy is consistent with the finest grid but for considerably less work. These methods converge in a small number of constant iterations independent of the grid size and hence, are often dramatically more efficient than others. In this thesis, we develop multigrid methods for three different classes of PDEs. In addition, we also develop discretization schemes for two model problems.

First, we propose multigrid methods based on upwind interpolation and restriction techniques for computing the steady state solutions for systems of one and two-dimensional nonlinear hyperbolic conservation laws. We prove that the two-grid method is total variation diminishing and the multigrid methods are consistent and convergent for one-dimensional linear systems.

Second, we propose a fully implicit, positive coefficient discretization that converges to the viscosity solution for a two-dimensional system of Hamilton-Jacobi-Bellman (HJB) PDEs resulting from dynamic Bertrand duopoly. Furthermore, we develop fast multigrid methods for solving the systems of discrete nonlinear HJB PDEs. The new multigrid methods are general and can be applied to other systems of HJB and HJB-Isaacs (HJBI) PDEs resulting from American options under regime switching and American options with unequal lending/borrowing rates and stock borrowing fees under regime switching, respectively. We provide a theoretical analysis for the smoother, restriction and interpolation operators of the multigrid methods.

Finally, we develop a fully implicit, unconditionally monotone finite difference numerical scheme, that converges to the viscosity solution of the three-dimensional PDE to price European options under a two-factor stochastic volatility model. The presence of cross derivative terms in high dimensional PDEs makes the construction of monotone discretization schemes challenging. We develop a wide stencil discretization based on a local coordinate transformation to eliminate the cross derivative terms. But, wide stencil discretization is first order accurate and computationally expensive compared to the second order fixed stencil discretization. Therefore, we use a hybrid stencil in which fixed stencil is used as much as possible and a wide stencil when the fixed stencil discretization does not satisfy the positive coefficient condition. We also develop fast multigrid methods to solve the discrete linear system.

Acknowledgements

I would like to express my deep gratitude to my advisor and mentor Justin Wan for his support, patience and insightful comments at every stage of the thesis. Thanks for believing in me and making the journey of my PhD smooth. I would also like to thank my committee members Peter Forsyth and George Labahn for their invaluable comments and suggestions at various points in my research. I also value the intellectual discussions with my labmates in the scientific computing lab.

I am very thankful to have a great family that stood by me throughout this process. Especially my grandparents, Rajagopal Naidu and Jayamma, and parents, Damodar Naidu and Prameela, who raised me with love and supported me in all my pursuits. I would also like to thank my dearest uncle Ravi for instilling a love of science in me. My brother, Prasanth, and sister, Sruthi, for being there for me all the time. And most of all, for my loving husband, Sunil Reddy Bommu, for his endless support and belief in me.

Dedication

To my husband, Sunil Reddy Bommu.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Multigrid Methods

A wide variety of phenomena such as fluid dynamics, heat flow, option pricing are modeled using PDEs. These equations often do not have a closed form solution. A typical way to solve such equations is to discretize them and solve the resulting large matrix problems. While direct solvers are efficient for small systems, they are not scalable to large systems both in terms of cost and memory. Iterative methods such as Jacobi, Gauss-Siedel and Krylov subspace methods are often used for large systems. For Poisson equation, Jacobi and Gauss-Siedel iterations converge at the rate of $1 - O(h^2)$, where h is the grid size. It is clear that the convergence of these methods deteriorates as the grid size is refined. Krylov subspace methods have a convergence rate of $1 - O(h)$ and hence are faster than Jacobi and Gauss-Siedel methods. Nevertheless, they are also very slow for very fine grids. However, it is noticed that the error components corresponding to the high frequency modes are damped very rapidly. The slow convergence is due to the difficulty in damping the low frequency modes. But many of these modes are mapped naturally into high frequency modes on a coarser mesh. Multigrid techniques exploit this property by moving to a coarser mesh to eliminate the corresponding error components [18, 47]. As a result, multigrid methods have convergence rates that are independent of the grid size. Multigrid methods were first developed for elliptic PDEs. They were extended to a wide variety of problems later on [18, 47, 106].

The idea of multigrid methods is to accelerate the convergence of a relaxation scheme by removing the low frequency error efficiently. First, a relaxation method, such as Jacobi or Gauss-Siedel iterations, is applied to the fine grid problem. The resulting error is smooth

and hence can be accurately represented on a coarse grid. Since the coarse grid is much smaller than the fine grid, it is much less expensive to work on the coarse grid. Resolving the error on the coarse grid is effective for low frequency error reduction. The fine grid solution is then updated with the interpolated coarse grid error followed by post-relaxation iterations.

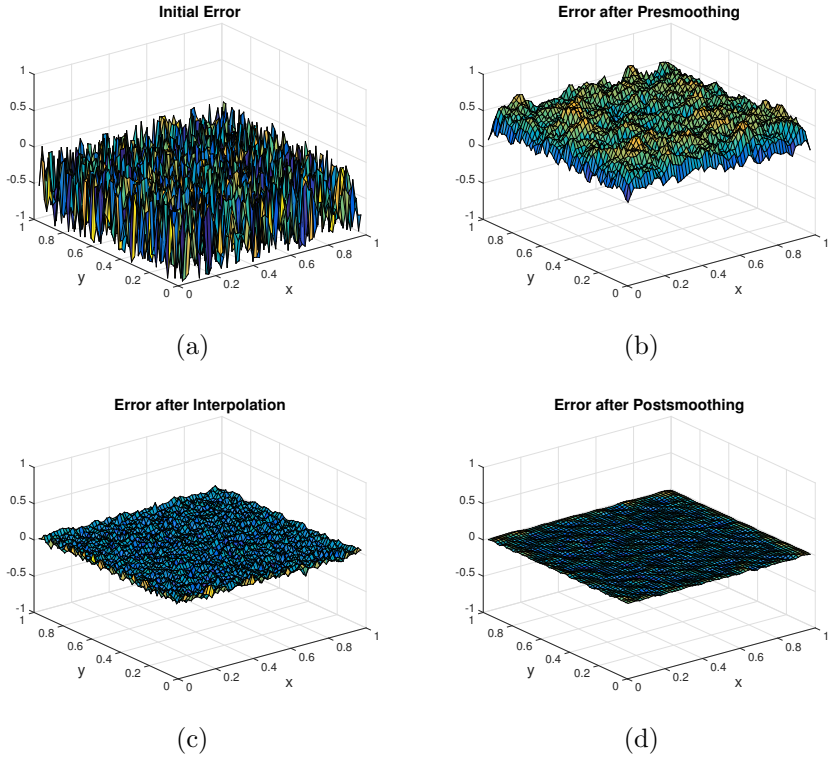


Figure 1.1: Illustration of error reduction of a two-grid method applied to a two-dimensional Poisson equation.(a) Initial error.(b) Error after pre-smoothing.(c) Error after coarse grid correction.(d) Error after post-smoothing.

Figure 1.1 illustrates the two-grid method applied to the two-dimensional Poisson equation. The initial error shown in Figure 1.1(a) is highly oscillatory. Two iterations of Gauss-Siedel smoothing efficiently smooths the error as shown in Figure 1.1(b). The coarse grid correction significantly reduces the error as shown in Figure 1.1(c). The error after one iteration of the two-grid method is very close to zero as shown in Figure 1.1(d)

Multigrid methods were first proposed for solving linear PDEs and later extended to

nonlinear problems as well. Full approximation scheme (FAS) is a multigrid method that directly handles the nonlinearity [18]. Consider a nonlinear problem $A^h(U^h) = B^h$ on the fine grid Ω_h. Given $U^{n,h}$, the two-grid FAS V-cycle to compute $U^{n+1,h}$ is given in Algorithm 1. Recursively applying the two-grid method gives the multigrid method.

Algorithm 1 Two-grid FAS V-cycle where Ω_h and Ω_H are the fine and coarse grids respectively.

(1) Pre-smoothing

Compute $\bar{U}^{n,h}$ by applying ν_1 smoothing iterations $\mathcal{S}(.)$ to $U^{n,h}$:

$\bar{U}^{n,h} = \mathcal{S}^{\nu_1}\left(U^{n,h}, A^h, B^h\right)$

(2) Coarse Grid Correction

Compute the residual: $\bar{R}^{n,h} = B^h - A^h(\bar{U}^{n,h})$

Restrict the residual using $\mathcal{R}_r$: $R^{n,H} = \mathcal{R}_r \bar{R}^{n,h}$

Restrict the solution using $\mathcal{R}_u$: $U^{n,H} = \mathcal{R}_u \bar{U}^{n,h}$

Compute the right hand side: $B^H = R^{n,H} + A^H(U^{n,H})$

Solve $A^H(\widetilde{U}^{n,H}) = B^H$ for $\widetilde{U}^{n,H}$

Compute correction: $E^{n,H} = \widetilde{U}^{n,H} - U^{n,H}$

Interpolate the correction using $\mathcal{P}$: $E^{n,h} = \mathcal{P} E^{n,H}$

Correct the approximation: $\hat{U}^{n,h} = \bar{U}^{n,h} + E^{n,h}$

(3) Post-smoothing

Compute $U^{n+1,h}$ by applying ν_2 smoothing iterations $\mathcal{S}(.)$ to $\hat{U}^{n,h}$:

$U^{n+1,h} = \mathcal{S}^{\nu_2}\left(\hat{U}^{n,h}, A^h, B^h\right)$

The core components of multigrid methods are the smoothing procedure, coarsening strategy, coarse grid operators, restriction and interpolation operators. In general relaxation-type iterative methods are efficient as smoothers. Full coarsening is the most commonly used coarsening technique. There are also semi-coarsening and other techniques [100]. A common choice for the coarse grid operator A^H is the direct discretization of the operator on the coarse grid. Another choice is the Galerkin coarse grid operator $A^H = \mathcal{R}_u A^h \mathcal{P}$, where $\mathcal{R}_u$ and $\mathcal{P}$ are the solution restriction and interpolation operators, respectively. The choice of the restriction and interpolation operators are closely related to the choice of the coarse grid. For standard full-coarsening, typical choice for the restriction and interpolation operators are the full-weighting and bilinear interpolation, respectively.

The multigrid components are very specific to each class of PDE. Although, it is well known how to choose suitable components for large classes of problems, it is still very difficult to define the right or reasonable ones for complicated new applications. For example, multigrid methods using fully-weighted restriction and bilinear interpolation are

efficient for elliptic PDEs. However, this approach does not work for hyperbolic problems [104]. One major difficulty is that the discretization matrices of hyperbolic equations are in general non-symmetric, and hence the smoothing property of relaxation methods, and the minimization property of Galerkin coarse grid correction, both of which are essentially based on symmetry and positive definiteness, may not hold for hyperbolic equations.

Another area in which standard multigrid methods are slow or may not work are for HJB and HJBI PDEs with jump in control. For efficient convergence of the multigrid method, it is important that the consistency of control between the fine and the coarse grid is preserved during restriction and interpolation. Also, the optimal control at the jump locations must be accurately captured during interpolation [48]. Standard restriction and interpolation techniques fail to address these issues.

It is also a challenge to design multigrid methods for high dimensional PDEs. A possible issue when solving higher dimensional PDEs is the inherent anisotropy making the use of pointwise relaxation ineffective [25, 100]. Standard coarsening and grid transfer operators may not work for anisotropic problems. Thus, significant theoretical and numerical research is required to design effective multigrid components for these different PDEs.

In this thesis, we develop fast multigrid methods for solving different types of PDEs, in particular, systems of hyperbolic PDEs, systems of HJB/HJBI PDEs and finally, a three-dimensional PDE.

1.2 Discretization

Discretization is the first step in numerically solving any given PDE. Multigrid methods are then used to solve the resulting discrete system. In [89], it was shown that seemingly reasonable discretizations of nonlinear PDEs may not converge to the viscosity solution, which is the financially relevant solution. Therefore, it is important to design discretization schemes that converge to the viscosity solution. A key sufficient requirement for ensuring convergence to a viscosity solution of a PDE is that the discretization scheme be monotone [11]. Monotone discretization schemes have not been developed for all the model problems we consider in this thesis. In particular, for the two-dimensional nonlinear system of HJB PDEs from dynamic Bertrand duopoly and the three-dimensional PDE from pricing European options under a two-factor stochastic volatility model. Therefore, we develop a fully implicit, positive coefficient discretization which results in an unconditionally monotone numerical scheme for these PDEs.

The presence of cross derivative terms in high dimensional PDEs makes the construction of positive coefficient discretization non-trivial. A skewed co-ordinate system can be used to transform the PDEs such that it results in a zero diffusion correlation. One can also rotate the grid by an appropriate angle such that it eliminates the cross derivative term [109]. The latter approach has the advantage that it preserves the orthogonality of the co-ordinate system. Another alternative is to enforce a spacing restriction on the original finite difference grid such that a positive coefficient condition results [26]. Enforcing grid spacing restriction is computationally inexpensive compared to the other approaches. However, it is in general difficult to enforce such a constraint when three cross derivative terms are present in a PDE. We address these issues and develop fully implicit, positive coefficient discretizations which results in an unconditionally monotone numerical scheme for the dynamic Bertrand oligopoly problem in Chapter 3 and the three-dimensional PDE in Chapter 4.

1.3 Model Problems

In this thesis, we choose three different types of PDEs for which standard multigrid methods do not work well and design the right components to achieve fast convergence of the multigrid method. The model problems considered arise from computational fluid dynamics, dynamic games and computational finance applications, which are described in the following sections.

1.3.1 Systems of Hyperbolic Conservation Laws

Hyperbolic PDEs arise in various applications where wave motion or advective transport is important. Such applications include, but are not limited to, gas dynamics, acoustics, electrodynamics, optics, geophysics and biomechanics. A special case of compressible gas dynamics, where the effects of viscosity and heat conduction are neglected, reduce to hyperbolic systems, in particular Euler equations. These equations result from aerodynamics, astrophysics and related applications where shock waves arise. Euler equations are nonlinear. This nonlinearity along with the shock formation in solution leads to computational challenges. The hyperbolic terms of the PDE, in general, pose the most stringent requirements on the discretization techniques [99].

Non-conservative schemes do not converge to the correct solution in the presence of shocks [56]. On the other hand, conservative numerical schemes, if convergent, do converge

to the weak solution of the conservation law [75]. Conservative methods in which the discretization is performed according to the direction of propagation of information are called upwind numerical methods. Upwind schemes incorporate the salient features of the physical phenomena modeled by the conservation laws. There are essentially two approaches for identifying upwind directions: the Godunov approach [44] and flux vector splitting (FVS) [96, 78] approach. FVS approach is simpler and hence more efficient than the Godunov approach. Therefore, we use the FVS scheme to discretize systems of hyperbolic conservation laws. We focus on developing efficient multigrid methods for systems of hyperbolic conservation laws.

1.3.2 Dynamic Bertrand Oligopoly

There are different models describing the operation of an oligopolistic market. Cournot oligopolies are markets in which the firms compete by using quantity as their strategic variable and price is determined by the market through an inverse demand function [27]. Bertrand oligopolies are competitive markets in which a small number of firms producing similar goods use price as their strategic variable under randomly fluctuating demands [15]. In reality, some markets can be better modeled as Cournot and some as Bertrand. In this thesis, we consider continuous time Bertrand models. The firms in this market sell differentiated but substitutable goods. Many products that are sold in markets fit this structure. For example, in the energy market, oil, coal and natural gas are commodities that can be substituted for one another. However, they have different prices per unit of energy produced. The price strategies of the firms in Bertrand oligopoly are characterized using the solution to a system of N_p-dimensional nonlinear HJB PDEs, where N_p is the number of firms.

Bertrand oligopolies under linear demand functions were analyzed by [77] using an asymptotic approximation in the limit of small competition. Numerical solutions were further used to analyze cases with high degree of substitutability. In particular, they analyze the effects of substitutability and relative firm size on prices, demands and profits. The main finding is that customers benefit the most when a market is composed of many firms of the same relative size producing highly substitutable goods.

The systems of nonlinear HJB PDEs, resulting from dynamic oligopoly problems, do not admit an explicit solution, except possibly in the monopoly case. Therefore, one needs numerical methods to compute the value functions and the equilibrium strategies. Naive numerical methods quickly become computationally infeasible as the number of players goes beyond three. Moreover, even in the two player case, these PDEs are highly coupled when

the competition is strong and are hard to handle. In this thesis, we develop an efficient discretization scheme and a fast numerical solver for the two-dimensional nonlinear HJB system, which has not been addressed in the literature so far.

1.3.3 Computational Finance

The Black-Scholes model has been widely used for option pricing. According to this model, the implied volatility of an option should be independent of its strike and expiration. The volatility surface of index options was indeed fairly flat prior to the stock market crash of October 1987 [30]. Since then, the volatility surface of index options has become skewed. This phenomenon, termed as the volatility smile, has spread to stock options, interest-rate options, currency options and almost every other volatility market. The Black-Scholes model cannot account for the smile. Recent research shows that models based on stochastic volatility, jump diffusion and regime switching processes better fit market data. There is no general agreement as to which model is the best, but each market has its own favorite or two. In this thesis, we focus on regime switching and stochastic volatility models.

Regime switching models are intuitively appealing and computationally inexpensive compared to the stochastic volatility and jump diffusion models. Regime switching models allow for different mean and volatility for the underlying variable in different regimes. This implies that the underlying variable is governed by different distributions in various regimes subject to the changing market conditions. Regime switching models have been widely used in different markets such as insurance, electricity, natural gas, optimal forestry management, trading strategies, valuation of stock loans, convertible bond pricing and interest rate dynamics [59]. A two state regime switching model with constant parameters can reproduce a volatility smile [107]. We consider American options under regime switching which results in systems of nonlinear HJB PDEs and American options with unequal lending/borrowing rates and stock borrowing fees under regime switching which result in systems of nonlinear HJBI PDEs. We use a fully implicit, positive coefficient, finite difference discretization, which ensures convergence to the viscosity solutions [37]. We then develop fast multigrid solvers for the discrete system.

A number of single-factor stochastic volatility models [60, 52, 94] have been developed. These models were motivated by the empirical study of stock price returns in which estimated volatility has random characteristics. The distribution of the returns from stochastic volatility models have thicker tails compared with the normal distribution, therefore modeling more extreme events which are characteristic of the real market. Single-factor stochastic volatility models can generate smiles and smirks. However, these models have

constant correlation between variance and stock returns over time. This limits the model's ability to capture the time varying nature of the smirk [24, 39]. It is also assumed that the variance drifts towards a long run constant mean. Historical values show that these parameters are not constant [14]. These shortcomings can be overcome by adding another factor to the single-factor stochastic volatility model [41].

In the past, stochastic volatility models with additional factors have been developed [35, 93]. Dupire's model [35] is hard to implement because local variances are not tradable. Stochastic implied volatility [93] is not practical as implied volatilities are not tradable. Variance swaps are tradable and are martingales under the risk neutral measure. In addition, forward variance swaps are natural hedges for exotic options. Therefore, it seems natural to impose dynamics on forward variance swaps [42]. A two-factor variance curve functional consistent with double mean reverting dynamics provides more flexibility to model the volatility term structure [20, 24]. We develop efficient numerical methods for the three-dimensional PDE modeling European option pricing under a two-factor stochastic volatility model.

1.4 Outline of Contributions

The main contributions of this thesis with regard to the multigrid methods for systems of hyperbolic PDEs are as follows.

- We develop total variation diminishing (TVD) methods for computing the steady state solutions for systems of hyperbolic conservation laws. An efficient multigrid method should avoid spurious numerical oscillations. This can be achieved by designing methods which preserve monotonicity and TVD properties through the use of upwind interpolation and restriction techniques. Such multigrid methods have been developed for scalar conservation laws [104]. However for hyperbolic systems, the upwinding directions are not apparent. We generalize the upwind techniques to systems by formulating the interpolation as solving a local Riemann problem. Upwind biased restriction is performed on the positive and negative components of the residual. This idea stems from the fact that the flux vector can be split into positive and negative components.

- We theoretically prove that the two-grid method is TVD for one-dimensional linear systems. We also prove that both the additive and multiplicative multigrid schemes are consistent and convergent for one-dimensional linear systems.

- For two-dimensional systems, we introduce a novel coarsening technique and extend the upwind interpolation and restriction techniques, which together capture the characteristics of the underlying system of hyperbolic equations.
- We demonstrate the effectiveness of our method by numerical examples including Euler equations.

The main contributions of this thesis with regard to the numerical methods for systems of nonlinear HJB/HJBI PDEs are as follows.

- We first construct a fully implicit, consistent, unconditionally l_∞ stable and monotone discretization scheme that converges to the viscosity solution for the HJB PDE resulting from dynamic Bertrand monopoly problem.
- The main challenge for the two-dimensional system of HJB PDEs, resulting from duopoly problem, is the discretization of the cross derivative term. We use a seven point stencil while imposing a constraint on the grid spacing to develop a positive coefficient discretization. We prove that the resulting discretization scheme converges to the viscosity solution.
- We develop a multigrid method based on the FAS to solve the systems of discrete HJB and HJBI PDEs. A weighted relaxation scheme is used as the smoother. We prove that the weighted relaxation scheme is globally convergent and show by a local Fourier analysis that the relaxation scheme with a damping parameter 0.67 is efficient in damping the high frequency error components. We develop two novel interpolation techniques which efficiently capture the optimal control in the presence of jumps. For HJB/HJBI problems, it is important that the consistency of control is preserved during restriction. We use a combination of penalty method and injection operator to satisfy this constraint.
- We show by a two grid Fourier analysis that the multigrid method gives efficient convergence. We also prove that the multigrid method is monotone, which ensures smooth convergence.
- We numerically test the efficiency of the multigrid method on HJB/HJBI systems resulting from dynamic Bertrand oligopoly and pricing American options under regime switching.

The main contributions of this thesis with regard to the numerical methods for three-dimensional PDE for pricing European options under a two-factor stochastic volatility model are as follows.

- We develop a fully implicit, unconditionally monotone finite difference numerical scheme for the three-dimensional partial differential equation (PDE) to price European options under a two-factor stochastic volatility model. The presence of cross derivative terms in high dimensional PDEs makes the construction of monotone discretization schemes challenging. We develop a wide stencil discretization based on a local coordinate transformation to eliminate the cross derivative terms.

- The wide stencil discretization is only first order accurate and computationally expensive compared to the second order fixed stencil discretization. Therefore, we use a hybrid stencil in which a fixed stencil is used as much as possible and a wide stencil when the fixed stencil discretization does not satisfy the positive coefficient condition [83].

- We perform a discretization analysis to show that the hybrid scheme is monotone, l_∞ stable and consistent in the viscosity sense and hence converges to the viscosity solution.

- We develop fast multigrid methods to solve the discrete linear system for uniform grids. The inherent anisotropy in high-dimensional PDEs makes the use of standard pointwise relaxation ineffective. We address this issue by using a block relaxation scheme, where each block consists of all the strongly coupled unknowns, for better smoothing of the error.

- We numerically test the convergence of the hybrid stencil discretization on non-uniform grids for double Heston and double CEV models. The discrete system is solved using a preconditioned Bi-CGSTAB iterative method [92]. We also numerically test the convergence of multigrid methods for uniform grids.

1.5 Outline

Chapter 2 presents multigrid methods for systems of hyperbolic conservation laws. Chapter 3 presents numerical methods for systems of nonlinear HJB/HJBI PDEs. Chapter 4 presents numerical methods for the three-dimensional PDE for pricing European options under a two-factor stochastic volatility model. Chapter 5 concludes.

Chapter 2

Multigrid Methods for Systems of Hyperbolic Conservation Laws

2.1 Introduction

While the standard theorems in the theory of multigrid methods generally assume ellipticity, it was first proposed by [63, 86] that it is possible to apply multigrid techniques to hyperbolic systems as well. The idea of their approach is to accelerate the wave propagation on multiple grids by using larger time steps on coarse grids without violating the Courant-Friedrichs-Lewy (CFL) condition. Thus the low frequency disturbances are rapidly expelled through the outer boundary while the high frequency modes are locally damped by the smoother. The smoother must be sufficiently dissipative to reduce the high frequency disturbances present in the initial approximation. Ni [86] performs smoothing by means of a control volume centered integration method with fluxes interpolated from corner values. In [23, 68, 69], Johnson applies the popular MacCormack scheme. Spekreijse [95] proposed a multigrid method for a second order accurate monotone upwind discretization of hyperbolic conservation laws. Dick [32] developed a multigrid method using flux difference splitting for discretization of the Euler equations.

For time-dependent hyperbolic problems, a standard explicit iterative scheme should suffice. The time step of an explicit scheme is limited by the CFL condition. If one is only interested in the steady state solution of the problem, multigrid methods using fully implicit schemes are unconditionally stable and have been used with success for the Euler equations [50, 51, 65, 84, 97]. In [64], Jameson has shown that implicit schemes based on approximate factorization can be used with multigrid techniques to achieve rapidly

convergent algorithm for computing the steady state solution of Euler equations. However, the implicit methods require the solution of the system at each time step. In [66], Jespersen gives an extensive review of the advantages and disadvantages of multigrid methods using implicit schemes.

The coarse grid corrections for multigrid methods using explicit time stepping schemes have an accelerating effect since they move the disturbances of the steady state over the distance of many mesh cells in one time step. Jespersen [67] mathematically supported this argument by showing that, under certain assumptions, such multigrid time stepping schemes on an L-grid are consistent and first order accurate, with effective time step, $\Delta t = \sum_k^L \Delta t_k$, where Δt_k is the time step taken on grid k. Further, the optimal speed of wave propagation is theoretically estimated to be $2^L - 1$ by [46, 82] using Fourier analysis.

Numerical oscillations can delay the propagation of the wave substantially [104]. Thus, it is desirable to design nonoscillatory multigrid schemes. Two multigrid time stepping schemes were proposed by [62] for computing the steady state solution of the one-dimensional linear wave equation. The idea is to capture the characteristics of the underlying PDE using upwind biased interpolation and residual restriction operators. The coarse grid update formula is modified to provide smooth results for the hyperbolic equations. It was proved that for two levels, these schemes preserve monotonicity. Wan et al. [104] extended the analysis to multilevel and proved that the multigrid methods satisfy the total variation diminishing property as well. They also extended these schemes to nonlinear scalar conservation laws, which are able to capture discontinuities arising from shocks and rarefactions without violating the entropy conditions.

A number of variants of the characteristic multigrid method have been developed. Koren et al. [72] used an upwind prolongation operator and restriction operator as an approximate adjoint of the upwind prolongation operator. The numerical results show that the improved multigrid method performs significantly better than a standard nonlinear multigrid method. Leclercq et al. [76] emphasized that the use of characteristic transfer operators allows for efficient resolution of shocks. This is not possible with the classical multigrid approach. Grasso et al. [45] proposed a multigrid technique based on directional grid transfer operators for hypersonic viscous flows.

In this thesis, we extend the multigrid schemes based on the characteristics approach [104] to systems of nonlinear hyperbolic conservation laws. However, a main challenge is that the system of equations is composed of a mixture of characteristics. In addition, the characteristic directions change from grid point to grid point. As a result, the upwind biased restriction and interpolation operators defined for scalar conservation laws cannot be directly extended to hyperbolic systems. We address these issues and develop novel upwind

biased interpolation and restriction techniques for systems of equations which result in efficient multigrid methods. We extend these upwind biased techniques to two-dimensional systems and introduce a novel coarsening technique. We also provide an analysis of the two-grid time stepping scheme for one-dimensional linear systems and prove that it satisfies the TVD property. This is a desirable property to design nonoscillatory multigrid schemes. Consistency and convergence proofs for multilevel schemes are also provided.

In Section 2.2, we present the FVS scheme, a finite difference scheme used for the discretization of systems of hyperbolic conservation laws. In Section 2.3, we describe the details of our multigrid method for one-dimensional hyperbolic conservation laws and in Section 2.4, we present an analysis for the TVD, consistency and convergence properties of the proposed multigrid scheme. In Section 2.5, we extend the multigrid scheme to two-dimensional hyperbolic systems. Finally, in Section 2.6, we demonstrate the effectiveness of our multigrid scheme by testing it on a number of hyperbolic systems including Euler equations.

2.2 Systems of Conservation Laws

Consider a multi-dimensional hyperbolic system of conservative laws,

$$U_t + \sum_{s=1}^{d} F_s(U)_{x_s} = 0, \qquad 0 < x_s < 1, \qquad t > 0, \tag{2.1}$$

where $U = (u_1, \cdots, u_m)^T \in \mathbf{R}^m$, d is the number of dimensions, $x = (x_1, \cdots, x_d) \in \mathbf{R}^d$, $F_s(U)$ is an m-dimensional vector of flux function and t denotes time. We are interested in finding the steady state solution of the system of equations given the initial condition,

$$U(x, 0) = U_0(x).$$

The boundary conditions are either of the Dirichlet or the Neumann type, which will be made more precise in the theoretical analysis in Section 2.4 and in the numerical results in Section 2.6.

There are various finite difference methods [49, 74, 81, 88, 98, 99] to discretize (2.1). We will use the FVS scheme, where discretization is performed according to the upwind direction. In Section 2.3, we will also use it as a smoother for the multigrid method. FVS [99] is a conservative numerical scheme of the form

$$U_J^{n+1} = U_J^n - \sum_{s=1}^{d} \lambda_s (F_{J+\frac{1}{2}e_s} - F_{J-\frac{1}{2}e_s}), \tag{2.2}$$

where $\lambda_s = \Delta t/\Delta x_s$ is the CFL number, J is the index of the grid point, Δt is the time step, Δx_s is the spatial step in the x_s direction and e_s is the unit vector in the x_s direction. FVS numerical flux is given by

$$F_{J+\frac{1}{2}e_s} = F_J^+(U_J^n) + F_{J+e_s}^-(U_{J+e_s}^n).$$

The positive and negative fluxes are given by

$$F_J^\pm(U_J^n) = A^\pm U_J^n,$$

where A^+ and A^- are the positive and negative components of the Jacobian matrix $A = \frac{\partial F}{\partial U}$. Let K be the matrix of right eigenvectors of A. Then,

$$A^\pm = K\Lambda^\pm K^{-1},$$

where Λ^+ and Λ^- are the diagonal matrices whose entries are the positive and negative eigenvalues of A, respectively.

For constant coefficient systems, the FVS scheme (2.2) can be rewritten as

$$U_J^{n+1} = U_J^n - \sum_{s=1}^{d}(\lambda_s A^+(U_J^n - U_{J-e_s}^n) - \lambda_s A^-(U_{J+e_s}^n - U_J^n)). \tag{2.3}$$

We note that we use the first order FVS scheme for illustration purposes only; other high resolution schemes, such as the positive schemes [74, 81], can also be used.

2.3 One-dimensional hyperbolic conservation laws

We design multigrid methods for hyperbolic systems based on the multigrid approach for the scalar conservation laws [104]. This multigrid method preserves monotonicity and is total variation diminishing for scalar conservation laws, which are two key properties to obtain efficient multigrid methods. We first briefly review the multigrid approach for scalar conservation laws and then extend the idea for systems of conservation laws.

Consider a scalar conservation law:

$$u_t + f(u)_x = 0, \qquad x \in (0,1),\ t > 0, \tag{2.4}$$

with appropriate boundary and initial conditions. The interval $[0,1]$ is discretized uniformly by taking the N points: $x_i^h = i \times h,\ i = 0, 1, \ldots, N-1$, where $h = 1/(N-1)$ denotes

the fine grid. Let u_i^h be the approximation of the exact solution $u(x_i^h)$. The boundary values u_0^h and u_{N-1}^h are either of the Dirichlet or the Neumann type. Let b_i^n be the fine grid right hand side (RHS) function which is initially zero. In general, the RHS function is nonzero for coarse grids. Given $\{x_i^h\}$, $i = 0, 1, \ldots, N-1$, the grid points with even indices are selected as coarse grid points; i.e., $\{x_i^H\}$, $i = 0, 2, \ldots, N-1$, where H denotes the coarse grid. In a two grid method, the fine grid solution is evolved by one time step using the FVS scheme (2.2). The evolved solution is denoted by $\bar{u}_i^h$. The solution $\bar{u}_i^h$ and the residual $\bar{r}_i^h$ are restricted to the coarse gird Ω^H using the restriction operators $\mathcal{R}_u$ and $\mathcal{R}_r$ respectively. $\mathcal{R}_u$ is an injection operator. $\mathcal{R}_r$ is an upwind restriction operator, the details of which are given in Section 2.3.2. The coarse grid RHS function is denoted by b_i^H. The solution on the coarse grid is evolved by one time step, where $\Delta t^H = 2\Delta t^h$, to obtain $\tilde{u}_i^H$. The updated solution u_i^{n+1} is computed by correcting the fine grid solution $\bar{u}_i^h$ with the interpolated coarse grid error $\mathcal{P}(\tilde{u}^H - \mathcal{R}_u u^n)_i$, where $\mathcal{P}$ is the interpolation operator. To summarize, given the solution u_i^n on the fine grid, the two-grid algorithm to compute u_i^{n+1} is given below.

Fine grid smoothing:
$$\bar{u}_i^h = u_i^n - \lambda(f^n_{i+\frac{1}{2}} - f^n_{i-\frac{1}{2}}) + \Delta t^h b_i^n \qquad i = 1, 2, \ldots, N-2.$$

Fine grid residual:
$$\bar{r}_i^h = \frac{1}{\Delta x^h}(\bar{f}^h_{i+\frac{1}{2}} - \bar{f}^h_{i-\frac{1}{2}}) - b_i^n \qquad i = 1, 2, \ldots, N-2.$$

Restriction of solution:
$$u_i^H = \mathcal{R}_u \bar{u}_i^h \qquad i = 2, 4, \ldots, N-3. \tag{2.5}$$

Coarse grid RHS:
$$b_i^H = \frac{1}{\Delta x^H}(f^H_{i+\frac{1}{2}} - f^H_{i-\frac{1}{2}}) - \mathcal{R}_r \bar{r}_i^h \qquad i = 2, 4, \ldots, N-3. \tag{2.6}$$

Coarse grid evolution:
$$\tilde{u}_i^H = u_i^H - \lambda(f^H_{i+\frac{1}{2}} - f^H_{i-\frac{1}{2}}) + \Delta t^H b_i^H \qquad i = 2, 4, \ldots, N-3. \tag{2.7}$$

Fine grid update:
$$\begin{cases} u_i^{n+1} = \tilde{u}_i^H \\ u_{i-1}^{n+1} = \bar{u}_{i-1}^h + \mathcal{P}(\tilde{u}^H - \mathcal{R}_u u^n)_{i-1} \end{cases} \qquad i = 2, 4, \ldots, N-1. \tag{2.8}$$

A multigrid algorithm is obtained by recursively applying the two-grid algorithm on the coarser grids. Since on each coarse grid, we use the updated solution, $\bar{u}^h$, from the

finer grid, we call this method the multiplicative scheme. Another approach is to restrict and propagate u^n on all coarse grids; i.e., (2.5) is replaced by

$$u_i^H = \mathcal{R}_u u_i^n, \ i = 2, 4, \ldots, N-3.$$

The resulting algorithm is called the additive scheme. For the additive scheme, b_i^H can be computed in two ways. We can either use (2.6) or set it to zero. Both the approaches have similar convergence rates. Setting $b_i^H = 0$ enables parallel computation of the algorithm.

The multigrid algorithm for hyperbolic equations is different from the FAS V-cycle in Algorithm 1. FAS is mainly used for elliptic or parabolic PDEs. For hyperbolic equations, the problem on the coarse grid is evolved by one time step as opposed to a complete solve in Algorithm 1. A modified fine grid update formula (2.8) is used for smooth convergence of the solution for hyperbolic case. There are no post-smoothing iterations, in general, for the hyperbolic case.

We will extend the above multigrid methods to systems of hyperbolic conservation laws in one dimension:

$$U_t + F(U)_x = 0, \qquad 0 < x < 1, \ t > 0, \tag{2.9}$$

where U is an m-dimensional vector. Given the appropriate initial and boundary conditions, we are interested in computing the steady state solution of the system (2.9). The basic principle of the multiplicative and additive multigrid methods is essentially the same as those for scalar conservation laws. We perform smoothing on the fine grid, restrict the solution to the coarser grids and accelerate the propagation on the coarser grids. However, the restriction and interpolation operators need to be redefined as there is a mixture of characteristics in a system.

2.3.1 Interpolation

The idea of upwind biased interpolation is to interpolate values based on the direction of propagation of information. Consider the linear wave equation: $u_t + u_x = 0$. In a two-grid method, given the coarse grid values $\tilde{u}_{i-2}^H$ and $\tilde{u}_i^H$, $i = 2, 4, \ldots, N-1$, we want to interpolate the solution at x_{i-1}^h. Since the wave is propagating from left to right, the interpolated value is given by $\tilde{u}_{i-1}^h = \tilde{u}_{i-2}^H$. For general scalar hyperbolic conservation laws, the direction of propagation of information can be determined using the sign of the characteristic speeds. If the sign is positive on both coarse grid points, the information is propagating from left to right and the fine grid solution is interpolated from the left. Conversely, if the sign

is negative, the interpolated value is from the right. For mixed signs (e.g. shocks and rarefactions), upwind interpolation can also be defined [104].

Schemes for applying upwind biased interpolation to systems of hyperbolic equations, however, is not apparent since the characteristics for some characteristic variables can be to the right while others are to the left. This is equivalent to the fact that some eigenvalues of the Jacobian of F can be positive and some negative. One possible solution is to apply upwind interpolation to each characteristic variable explicitly. This method works well when the right eigenvectors (i.e. K) are constant; e.g. $F(U) = AU$ where A is a constant matrix. In this case, let $W = K^{-1}U$. Then (2.9) becomes a decoupled system in the new characteristic variables W. One can easily apply the upwind interpolation to this system. In general, however, the right eigenvectors are not constant. Then it would not be a good idea to apply upwind interpolation to the characteristic variables ($W_i = K_i^{-1}U_i$) from one grid point to interpolate the values of the characteristic variables ($W_{i-1} = K_{i-1}^{-1}U_{i-1}$) at another grid point when K_i and K_{i-1} are different.

There is another more subtle issue of using characteristics explicitly for defining upwind interpolation. Recall that the coarse grid correction step updates the fine grid solution by interpolating the error from the coarse grid as given by (2.8). When (2.8) is directly applied to systems, we have the fine grid update, $\widetilde{U}_{i-1}^h$, as

$$\widetilde{U}_{i-1}^h = \overline{U}_{i-1}^h + \mathcal{P}(\widetilde{U}^H - \mathcal{R}_u U^n)_{i-1}, \qquad i = 2, 4, \ldots, N-1.$$

For nonlinear equations, upwind interpolation based on characteristics will depend on the solution being interpolated. Thus, if one tries to interpolate the coarse grid error to the fine grid, the characteristics of the error will be used to define the upwind direction, which may not necessarily capture the behavior of the solution. One may interpolate $\widetilde{U}^H$ by $\mathcal{P}_1$ based on the characteristics of $\widetilde{U}^H$ and separately interpolate $\mathcal{R}_u U^n$ by $\mathcal{P}_2$ based on the characteristics of $\mathcal{R}_u U^n$. In general $\mathcal{P}_1$ and $\mathcal{P}_2$ are not the same, and hence different interpolation will be performed at the same grid point for $\widetilde{U}^H$ and $\mathcal{R}_u U^n$. In our experience, this inconsistency leads to slow convergence in general. One may pick one of the two interpolation operators. In [104], the authors apply $\mathcal{P}_1$ to interpolate $\mathcal{R}_u U^n$. However, in general, it is not clear whether one should choose $\mathcal{P}_1$ or $\mathcal{P}_2$, and whether the upwind interpolation based on one solution would be good for the other.

To address this issue, we note that upwind interpolation can be formulated as solving a local Riemann problem [104]. Consider the linear wave equation and the following local

boundary value problem:

$$\begin{aligned} e_\tau + e_x &= 0, & x_{i-2}^h < x < x_i^h, \\ e(x,0) &= \begin{cases} e_{i-2}^H & x_{i-2}^h < x < x_{i-1}^h, \\ e_i^H & x_{i-1}^h < x < x_i^h, \end{cases} \end{aligned} \tag{2.10}$$

where τ is an artificial time, e_{i-2}^H and e_i^H are two given constant values as shown in Figure 2.1. The steady state solution of (2.10) at x_{i-1}^h gives the interpolated value. In this case, the value is e_{i-2}^H, which is the same as the value given by the upwind biased interpolation.

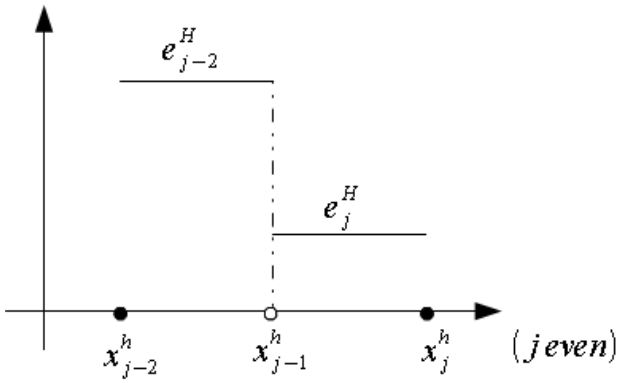


Figure 2.1: The initial data for a Riemann problem for defining an upwind interpolation. The data consists of two constant states separated by a discontinuity at x_{i-1}^h.

We extend the local Riemann formulation for upwind interpolation to systems of hyperbolic equations. In this case, we do not interpolate the coarse grid error onto the fine grid and then update the fine grid solution. Instead, we update the fine grid solution at the coarse grid points using the coarse grid solution. At the noncoarse grid points, we determine the values by upwind interpolation. More precisely, for $i = 2, 4, \ldots, N-1$, we solve the following local Riemann problem:

$$\begin{aligned} U_\tau + F(U)_x &= 0, & x_{i-2}^h < x < x_i^h, \\ U(x,0) &= \begin{cases} \widetilde{U}_{i-2}^h & x_{i-2}^h < x < x_{i-1}^h, \\ \widetilde{U}_i^h & x_{i-1}^h < x < x_i^h, \end{cases} \end{aligned} \tag{2.11}$$

where $\widetilde{U}_{i-2}^h$ and $\widetilde{U}_i^h$ are the fine grid values of U at x_{i-2}^h and x_i^h respectively after coarse grid correction. We compute the steady state solution by using the FVS scheme (2.2) for each noncoarse grid point. The upwind interpolation is essentially carried out by the upwind

discretization of the FVS scheme. Since the local Riemann problem only involves one grid point, the computation is very efficient.

We note that the interpolation can be performed by computing the exact solution to the local Riemann problem. For the linear wave equation example, the exact solution at x_{i-1}^h is $\tilde{u}_{i-2}^H$, which is the same as (2.12). In general, however, not all the hyperbolic equations have analytical solutions. In order for the method to be applicable to general hyperbolic systems, we describe the interpolation procedure as solving a local Riemann problem numerically to steady state.

The proposed interpolation approach resolves the two issues addressed above. The FVS scheme captures the different characteristics through the two flux functions. Thus, we do not need to determine the characteristic direction for each characteristic variable in order to apply the upwind interpolation. Second, the local Riemann problem is based on the fine grid values of U and hence we avoid the inconsistency inherent in choosing whether to apply the upwind interpolation to $\widetilde{U}^H$ or $\mathcal{R}_u U^n$.

We remark that this interpolation approach results in a different coarse grid correction. Consider the 1D linear wave equation: $u_t + u_x = 0$. In [104], the coarse grid update at a noncoarse grid point is:

$$\tilde{u}_{i-1}^h = \bar{u}_{i-1}^h + (\tilde{u}_{i-2}^H - u_{i-2}^n), \qquad i = 2, 4, \ldots, N-1.$$

The new approach performs upwind interpolation directly on $\tilde{u}^h$ from coarse grid points to noncoarse grid points. Thus,

$$\tilde{u}_{i-1}^h = \tilde{u}_{i-2}^H, \qquad i = 2, 4, \ldots, N-1. \tag{2.12}$$

We will show in Section 2.4 that the new interpolation approach will still result in a multigrid scheme which is total variation diminishing.

2.3.2 Residual Restriction

For scalar hyperbolic equations, an upwind biased residual restriction can be defined based on the characteristic direction [104]. Consider the scalar conservation law (2.4) again. The characteristic direction can be determined by the sign of the derivative of the flux, $f'(u)$. Given the fine grid residual, r_i^h, we want to compute the residual on the coarse grid, r_i^H. Suppose $f_i'(u_{i-1})$, $f_i'(u_i)$ and $f_i'(u_{i+1})$ are all positive, then the information is propagating from left to right and hence averaging for the residual is performed from the grid point towards its left as given by

$$r_i^H = 0.5[r_i^h + r_{i-1}^h], \qquad i = 2, 4, \ldots, N-1, \tag{2.13}$$

and $r_0^H = 0$. Similarly, when the wave is locally propagating to the left, averaging for residual is performed from the right. For details on special situations such as shocks and rarefactions, we refer the reader to [104].

The upwind residual restriction technique described above works well since the scalar unknown is either moving forward or backward at each grid point. However, a system of hyperbolic conservation laws is composed of a mixture of characteristics; i.e., it contains both forward and backward moving components. Thus, this approach cannot be directly extended to the system case. The main issue is that it is not apparent what the upwind side is at each grid point for systems and hence averaging for the residual based on the upwind direction becomes infeasible.

Nevertheless, the concept of propagation direction of information can still be exploited for systems of equations. The FVS intercell numerical flux is composed of two contributions:

$$F_{i+\frac{1}{2}} = F_i^+(U_i^h) + F_{i+1}^-(U_{i+1}^h), \qquad i = 1, 2, \ldots, N-2. \tag{2.14}$$

where one comes from the forward component of the flux, F_i^+ and the other comes from the backward component of the flux, F_{i+1}^-, as shown in Figure 2.2. Using the idea of the FVS

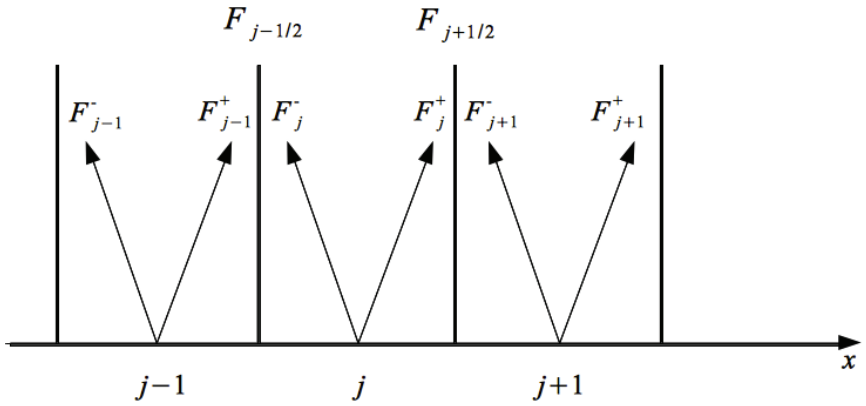


Figure 2.2: The forward and backward flux components of the intercell flux functions $F_{i-\frac{1}{2}}$ and $F_{i+\frac{1}{2}}$.

numerical flux splitting (2.14), we will split the fine grid residual, $R_i^h, i = 1, 2, \ldots, N-2$, similarly into two parts, based on the forward and backward flux components. More precisely, let $(R_i^h)^+$ be the residual from the forward component of fluxes and $(R_i^h)^-$ be

the residual from the backward component of fluxes. Then $R_i^h = (R^h)_i^+ + (R^h)_i^-$ where

$$\begin{aligned}(R^h)_i^+ &= \frac{1}{\Delta x^h}[F_i^+(U_i^h) - F_{i-1}^+(U_{i-1}^h)] - (B^h)_i^+, \\ (R^h)_i^- &= \frac{1}{\Delta x^h}[F_{i+1}^-(U_{i+1}^h) - F_i^-(U_i^h)] - (B^h)_i^- \\ & \qquad\qquad\qquad\qquad i = 1, 2, \ldots, N-2.\end{aligned}$$

We note that $R_0^h = R_{N-1}^h = 0$. $(B^h)_i^+$ and $(B^h)_i^-$ are the positive and negative RHS functions which will be made more precise later in this section.

In order to define the upwind restriction for the residual, we note that the Jacobian matrix, $A^+ = \frac{\partial F^+}{\partial U}$ is positive semidefinite and $A^- = \frac{\partial F^-}{\partial U}$ is negative semidefinite. The eigenvalues determine the characteristic direction of the system. A^+ has positive or zero eigenvalues and the information is propagating from left to right. Similarly, A^- has negative or zero eigenvalues and the information is propagating from right to left. We exploit these directionality properties of the FVS scheme to perform upwind biased residual restriction. For the positive component of the residual, averaging is performed from the grid point towards its left as given by

$$(R^H)_i^+ = 0.5[(R^h)_i^+ + (R^h)_{i-1}^+], \qquad i = 2, 4, \ldots, N-3.$$

For the negative component of the residual, averaging is performed from the grid point towards its right,

$$(R^H)_i^- = 0.5[(R^h)_i^- + (R^h)_{i+1}^-], \qquad i = 2, 4, \ldots, N-3.$$

The coarse grid residual, R_i^H, is then given by the sum of the positive and negative coarse grid residuals,

$$\begin{aligned}R_i^H &= (R^H)_i^+ + (R^H)_i^-, \qquad i = 2, 4, \ldots, N-3, \\ R_i^H &= 0, \qquad\qquad\qquad\quad i = 0, N-1.\end{aligned} \tag{2.15}$$

Following a similar idea, the coarse grid RHS, B_i^H, is then defined by

$$\begin{aligned}B_i^H &= (B^H)_i^+ + (B^H)_i^-, \qquad i = 2, 4, \ldots, N-3. \\ B_i^H &= 0, \qquad\qquad\qquad\quad i = 0, N-1.\end{aligned}$$

where,

$$\begin{aligned}(B^H)_i^+ &= \frac{1}{\Delta x}[F_i^+(U_i^H) - F_{i-2}^+(U_{i-2}^H)] - 0.5[(R^h)_i^+ + (R^h)_{i-1}^+], \\ (B^H)_i^- &= \frac{1}{\Delta x}[F_{i+2}^-(U_{i+2}^H) - F_i^-(U_i^H)] - 0.5[(R^h)_i^- + (R^h)_{i+1}^-], \\ & \qquad\qquad\qquad\qquad i = 2, 4, \ldots, N-3.\end{aligned}$$

This residual restriction approach extends the previous approach for scalar conservation laws. Consider the scalar equation (2.4) again. Suppose the wave is locally propagating to the right. Then the residual restriction based on characteristics is given by (2.13). In our new approach, $r_i^H = (r^H)_i^+ + (r^H)_i^-$, $i = 2, 4, \ldots, N-1$. However, since the sign of the characteristic speed is positive, the backward moving component of the flux is zero; i.e. $(r^H)^- = (r^h)^- = 0$. As a result, $r_i^H = (r^H)_i^+ = 0.5[(r^h)_i^+ + (r^h)_{i-1}^+] = 0.5[(r^h)_i + (r^h)_{i-1}]$, which is the same as (2.13). By a similar calculation, for other characteristic directions including cases of shocks and rarefactions, it can be shown that the new approach recovers the same formulas as given by the upwind restriction technique defined for the scalar case. This generalization has the advantage that it does not require an explicit knowledge of the characteristics as in [104], and hence can be easily applied to the system case.

2.3.3 Algorithm

The multiplicative multigrid time stepping scheme for one-dimensional systems of hyperbolic conservation laws is given in Algorithm 2. We denote the functions on the finest grid by (1), the second grid by (2), and so on. We denote the functions on the coarsest grid by (CG). FVS scheme (2.2) is used as the smoother. The interpolation operator $\mathcal{P}$ solves a local Riemann problem to the steady state; see (2.11).

On each coarse grid k, we restrict and propagate the most recently updated solution $\overline{U}^{(k-1)}$ from the finer grid in the multiplicative scheme as given in Algorithm 2. If we restrict and propagate $U^{(k)}$ on all the coarse grids as

$$U_i^{(k+1)} = U_i^{(k)} \qquad (i = 2^k, 2 \cdot 2^k, \ldots),$$

then we obtain the additive multigrid time stepping scheme.

In the next section, we prove that both the two-grid multiplicative and additive time stepping schemes satisfy the TVD, consistency and convergence properties for constant coefficient linear systems of hyperbolic conservation laws. In Appendix A.1, we prove that the two-grid schemes for the linear wave equation preserve monotonicity. These properties are important in ensuring that the multigrid methods are nonoscillatory and in turn result in smooth and fast convergence of the multigrid solution.

2.4 Linear Analysis

In this section, we analyze the TVD, consistency and convergence properties of multigrid time stepping schemes for solving an $m \times m$ hyperbolic system of linear conservation laws

Algorithm 2 Multiplicative multigrid scheme for one-dimensional systems of hyperbolic conservation laws.

$U^{(1)} = U^n, B^{(1)} = 0$
Define $\widetilde{U}^{(k)} = \mathrm{MG}(U^{(k)}, B^{(k)})$:
if $k = CG$ **then**
  $\widetilde{U}_i^{(k)} = U_i^{(k)} - \lambda(F_{i+\frac{1}{2}}^{(k)} - F_{i-\frac{1}{2}}^{(k)}) + \Delta t^{(k)} B_i^{(k)} \qquad (i = 2^{k-1}, 2\cdot 2^{k-1}, \ldots)$
else
  $\overline{U}_i^{(k)} = U_i^{(k)} - \lambda(F_{i+\frac{1}{2}}^{(k)} - F_{i-\frac{1}{2}}^{(k)}) + \Delta t^{(k)} B_i^{(k)} \qquad (i = 2^{k-1}, 2\cdot 2^{k-1}, \ldots)$
  $\overline{R}_i^{(k)} = \frac{1}{\Delta x^{(k)}}(\overline{F}_{i+\frac{1}{2}}^{(k)} - \overline{F}_{i-\frac{1}{2}}^{(k)}) - B_i^{(k)} \qquad (i = 2^{k-1}, 2\cdot 2^{k-1}, \ldots)$
  $U_i^{(k+1)} = \mathcal{R}_u \overline{U}_i^{(k)} \qquad (i = 2^k, 2\cdot 2^k, \ldots)$
  $B_i^{(k+1)} = \frac{1}{\Delta x^{(k+1)}}(F_{i+\frac{1}{2}}^{(k+1)} - F_{i-\frac{1}{2}}^{(k+1)}) - \mathcal{R}_r \overline{R}_i^{(k)} \qquad (i = 2^k, 2\cdot 2^k, \ldots)$
  $\widetilde{U}^{(k+1)} = \mathrm{MG}(U^{(k+1)}, B^{(k+1)})$
  $\widetilde{U}_i^{(k)} = \widetilde{U}_i^{(k+1)} \qquad (i = 2^k, 2\cdot 2^k, \ldots)$
  $\widetilde{U}_{i-2^{k-1}}^{(k)} = \mathcal{P}(\widetilde{U}_{i-2\cdot 2^{k-1}}^{(k+1)}, \overline{U}_{i-2^{k-1}}^{(k)}, \widetilde{U}_i^{(k+1)}) \qquad (i = 2^k, 2\cdot 2^k, \ldots)$
end if
$U^{n+1} = \mathrm{MG}(U^{(1)}, B^{(1)})$

given by,

$$U_t + AU_x = 0, \qquad 0 < x < 1,\ t > 0, \tag{2.16}$$

where A is an $m \times m$ constant matrix. Without loss of generality, we assume a zero boundary condition and a zero RHS. We use the FVS scheme (2.2) as the smoother for the multigrid algorithm. We first define the TVD and consistency preserving properties.

Definition 2.4.1. *(TVD) Given any set of data U^n, let U^{n+1} be computed by one multigrid cycle. Then the multigrid method is called total variation diminishing (TVD) [79], if*

$$TV(U^{n+1}) \le TV(U^n),$$

where the total variation of a function U is defined in terms of its vector norm as

$$TV(U) = \sum_i \|U_i - U_{i-1}\|. \tag{2.17}$$

Definition 2.4.2. *(Consistency) Let ϕ be any numerical scheme for solving a PDE. $U^{n+1} = \phi(U^n)$ is a consistent scheme if $U^* = \phi(U^*)$, where U^* is the exact solution.*

Before we present the analysis, we make a useful observation about the coarse grid RHS B_i^H.

Lemma 2.4.1. *The coarse grid RHS function is zero for constant coefficient linear systems.*

Proof. For the two-grid, multiplicative scheme given by Algorithm 1, we rewrite B_i^H, $i = 2, 4, \ldots, N-3$ for constant coefficient linear systems as

$$\begin{aligned} B_i^H &= \frac{1}{\Delta x^H}(A^+(U_i^H - U_{i-2}^H) + A^-(U_{i+2}^H - U_i^H)) - \frac{1}{2}((\bar{R}_i^h)^+ + (\bar{R}_{i-1}^h)^+) \\ &\quad - \frac{1}{2}((\bar{R}_{i-1}^h)^- + (\bar{R}_i^h)^-), \\ &= \frac{1}{2\Delta x^h}(A^+(\bar{U}_i^h - \bar{U}_{i-2}^h) + A^-(\bar{U}_{i+2}^h - \bar{U}_i^h)) - \frac{1}{2\Delta x^h}(A^+(\bar{U}_i^h - \bar{U}_{i-1}^h) \\ &\quad + A^+(\bar{U}_{i-1}^h - \bar{U}_{i-2}^h)) - \frac{1}{2\Delta x^h}(A^-(\bar{U}_{i+1}^h - \bar{U}_i^h) + A^-(\bar{U}_{i+2}^h - \bar{U}_{i+1}^h)), \\ &= 0. \end{aligned}$$

□

For constant coefficient linear systems, it is possible to prove stability by measuring the total variation in terms of wave strengths instead of using the standard vector norms [79]. Let K be the matrix of right eigenvectors of A and define $W = K^{-1}U$. The total variation for constant coefficient linear systems can then be computed in terms of its characteristic variables.

Lemma 2.4.2. *[79] The total variation of an m-dimensional vector U can be computed in terms of W as*

$$TV(U) = \sum_{l=1}^{m} TV(W_l),$$

where $TV(W_l)$ is the total variation of the l-th characteristic component.

We assume that A has m_1 positive eigenvalues and m_2 negative eigenvalues, where $m_1 + m_2 = m$. Without loss of generality, we assume that the vector of positive eigenvalues, α^+, of A are ordered before the vector of negative eigenvalues, α^-. Thus we write W as

$$W = \begin{bmatrix} p \\ q \end{bmatrix}, \tag{2.18}$$

where $p = [p_1, p_2, \ldots, p_{m_1}]$ is an m_1-dimensional vector with positive sign for the characteristic speeds and $q = [q_{m_1+1}, q_{m_1+2}, \ldots, q_m]$ is an m_2-dimensional vector with negative sign for the characteristic speeds.

In the following theorem, we analyze the TVD properties of the multigrid method for the system of equations (2.16).

Theorem 2.4.1. *Let U^n be the solution at time step n, then,*

$$TV(U^{n+1}) \leq TV(U^n)$$

for both the two-grid multiplicative and additive multigrid time stepping schemes.

Proof. From Lemma 2.4.2 and (2.18), we have

$$TV(U^{n+1}) = \sum_{l=1}^{m_1} TV(p_l^{n+1}) + \sum_{l=m_1+1}^{m} TV(q_l^{n+1}). \tag{2.19}$$

We now compute $TV(p_l^{n+1})$, $l = 1, 2, \ldots, m_1$ using Lemma A.2.2. Let $\mathcal{I} = \{2, 4, \ldots, N-1\}$. We assume zero boundary conditions and $(p_l)_i^n = 0,\ i \leq 0$. From the CFL condition, we have $0 \leq |\lambda\alpha_l^+| \leq 1$, and

$$\begin{aligned}
TV(p_l^{n+1}) &= \sum_{i\in\mathcal{I}} |(p_l)_i^{n+1} - (p_l)_{i-1}^{n+1}| + |(p_l)_{i-1}^{n+1} - (p_l)_{i-2}^{n+1}| \\
&\leq (1-\lambda\alpha_l^+)^2|((p_l)_i^n - (p_l)_{i-1}^n)| + (1-\lambda\alpha_l^+)^2|((p_l)_{i-1}^n - (p_l)_{i-2}^n)| \\
&\quad + \lambda\alpha_l^+(1-\lambda\alpha_l^+)|((p_l)_{i-1}^n - (p_l)_{i-2}^n)| + \lambda\alpha_l^+(1-\lambda\alpha_l^+)|((p_l)_{i-2}^n - (p_l)_{i-3}^n)| \\
&\quad + \lambda\alpha_l^+(1-\lambda\alpha_l^+)|((p_l)_{i-2}^n - (p_l)_{i-3}^n)| + \lambda\alpha_l^+(1-\lambda\alpha_l^+)|((p_l)_{i-3}^n - (p_l)_{i-4}^n)| \\
&\quad + (\lambda\alpha_l^+)^2|((p_l)_{i-3}^n - (p_l)_{i-4}^n)| + (\lambda\alpha_l^+)^2|((p_l)_{i-4}^n - (p_l)_{i-5}^n)| \\
&\leq ((1-\lambda\alpha_l^+)^2 + 2\lambda\alpha_l^+(1-\lambda\alpha_l^+) + (\lambda\alpha_l^+)^2)TV(p_l^n) \\
&= TV(p_l^n).
\end{aligned} \tag{2.20}$$

Similarly $TV(q_l^{n+1})$, where $l = m_1+1, m_1+2, \ldots, m$ is computed using Lemma A.2.2. We note that with zero boundary conditions, $(q_l)_i^n = 0,\ i \geq N-1$.

$$\begin{aligned}
TV(q_l^{n+1}) &= \sum_{i\in\mathcal{I}} |(q_l)_i^{n+1} - (q_l)_{i-1}^{n+1}| + |(q_l)_{i-1}^{n+1} - (q_l)_{i-2}^{n+1}| \\
&= \sum_{i\in\mathcal{I}} |(\lambda\alpha_l^-)^2((q_l)_{i+3}^n - (q_l)_{i+2}^n) + (\lambda\alpha_l^-)^2((q_l)_{i+2}^n - (q_l)_{i+1}^n) \\
&\quad - \lambda\alpha_l^-(1+\lambda\alpha_l^-)((q_l)_{i+2}^n - (q_l)_{i+1}^n) - \lambda\alpha_l^-(1+\lambda\alpha_l^-)((q_l)_{i+1}^n - (q_l)_i^n) \\
&\quad - \lambda\alpha_l^-(1+\lambda\alpha_l^-)((q_l)_{i+1}^n - (q_l)_i^n) - \lambda\alpha_l^-(1+\lambda\alpha_l^-)((q_l)_i^n - (q_l)_{i-1}^n) \\
&\quad + (1+\lambda\alpha_l^-)^2((q_l)_i^n - (q_l)_{i-1}^n) + (1+\lambda\alpha_l^-)^2((q_l)_{i-1}^n - (q_l)_{i-2}^n)|.
\end{aligned} \tag{2.21}$$

We have the CFL condition as $0 \leq |\lambda\alpha_l^-| \leq 1$. Since $\alpha_l^- < 0$, (2.21) can be rewritten as

$$\begin{aligned} TV(q_l^{n+1}) &\leq \sum_{i\in\mathcal{I}} (\lambda\alpha_l^-)^2|((q_l)_{i+3}^n - (q_l)_{i+2}^n)| + (\lambda\alpha_l^-)^2|((q_l)_{i+2}^n - (q_l)_{i+1}^n)| \\ &\quad - \lambda\alpha_l^-(1+\lambda\alpha_l^-)|((q_l)_{i+2}^n - (q_l)_{i+1}^n)| - \lambda\alpha_l^-(1+\lambda\alpha_l^-)|((q_l)_{i+1}^n - (q_l)_i^n)| \\ &\quad - \lambda\alpha_l^-(1+\lambda\alpha_l^-)|((q_l)_{i+1}^n - (q_l)_i^n)| - \lambda\alpha_l^-(1+\lambda\alpha_l^-)|((q_l)_i^n - (q_l)_{i-1}^n)| \\ &\quad + (1+\lambda\alpha_l^-)^2|((q_l)_i^n - (q_l)_{i-1}^n)| + (1+\lambda\alpha_l^-)^2|((q_l)_{i-1}^n - (q_l)_{i-2}^n)| \\ &\leq ((\lambda\alpha_l^-)^2 - 2\lambda\alpha_l^-(1+\lambda\alpha_l^-) + (1+\lambda\alpha_l^-)^2)TV(q_l^n) \\ &= TV(q_l^n). \end{aligned} \tag{2.22}$$

Hence by (2.19), (2.20) and (2.22), the two-grid multiplicative time stepping scheme is TVD for systems of constant coefficient linear hyperbolic conservation laws.

Similarly for a two-grid additive scheme, $TV(p_l^{n+1})$, where $l = 1, 2, \ldots, m_1$ is computed using Lemma A.2.3,

$$\begin{aligned} TV(p_l^{n+1}) &\leq \sum_{i\in\mathcal{I}} (1-\lambda\alpha_l^+)|((p_l)_i^n - (p_l)_{i-1}^n)| + (1-\lambda\alpha_l^+)|((p_l)_{i-1}^n - (p_l)_{i-2}^n)| \\ &\quad + \lambda\alpha_l^+|((p_l)_{i-2}^n - (p_l)_{i-3}^n)| + \lambda\alpha_l^+|((p_l)_{i-3}^n - (p_l)_{i-4}^n)| \\ &\leq (1-\lambda\alpha_l^+ + \lambda\alpha_l^+)TV(p_l^n) \\ &= TV(p_l^n). \end{aligned} \tag{2.23}$$

$TV(q_l^{n+1})$, $l = m_1+1, m_1+2, \ldots, m$ is obtained from Lemma A.2.3 as,

$$\begin{aligned} TV(q_l^{n+1}) &\leq \sum_{i\in\mathcal{I}} -\lambda\alpha_l^-|((q_l)_{i+2}^n - (q_l)_{i+1}^n)| - \lambda\alpha_l^-|((q_l)_{i+1}^n - (q_l)_i^n)| \\ &\quad + (1+\lambda\alpha_l^-)|((q_l)_i^n - (q_l)_{i-1}^n)| + (1+\lambda\alpha_l^-)|((q_l)_{i-1}^n - (q_l)_{i-2}^n)| \\ &\leq (-\lambda\alpha_l^- + 1 + \lambda\alpha_l^-)TV(q_l^n) \\ &= TV(q_l^n). \end{aligned} \tag{2.24}$$

Thus, by (2.19), (2.23) and (2.24) the additive scheme for constant coefficient linear systems also satisfies the TVD property. □

The analysis for consistency and convergence properties are presented in the following theorems.

Theorem 2.4.2. *Both the multiplicative and additive time stepping schemes for constant coefficient linear systems are consistent.*

Proof. The characteristics form of the linear system (2.16) is given by

$$W_t + \Lambda W_x = 0, \qquad 0 < x < 1,\ t > 0. \tag{2.25}$$

We rewrite (2.25) as two separate equations: one for vector p associated with positive eigenvalues α^+ and the other for vector q associated with negative eigenvalues α^- using (2.18).

$$p_t + \alpha^+ p_x = 0, \tag{2.26}$$

$$q_t + \alpha^- q_x = 0. \tag{2.27}$$

Suppose $p^n = p^*$ is the exact steady state solution of (2.26), which is given by $p_i^* = p_0^*,\ i = 0, 1, \ldots, N-1$. We then compute $p_i^{n+1},\ i = 1, 2, \ldots, N-1$ using a k-level additive scheme. By Lemma A.2.3,

$$p_i^{n+1} \quad = \quad (1 - \lambda\alpha^+ + \lambda\alpha^+)p_0^* = p_0^*.$$

Similarly, if $q^n = q^*$ is the exact steady state solution of (2.27), which is given by $q_i^* = q_{N-1}^*,\ i = 0, 1, \ldots, N-1$. $q_i^{n+1},\ i = 0, 1, \ldots, N-2$ is computed using Lemma A.2.3,

$$q_i^{n+1} \quad = \quad (1 + \lambda\alpha^- - \lambda\alpha^-)q_{N-1}^* = q_{N-1}^*.$$

By Definition 4.4.1, the k-level additive scheme is consistent for each scalar conservation law in the W-space and hence in the U-space as well. The consistency of the multilevel multiplicative scheme can be proved using similar reasoning, we omit the details. □

Theorem 2.4.3. *Both the additive and multiplicative time stepping schemes for constant coefficient linear systems are convergent.*

Proof. Consider the linear system (2.16). Without loss of generality, we assume that the RHS and boundary conditions are zero. For the two-grid multiplicative scheme, $(p_l)_i^{n+1},\ i = 2, 3, \ldots, N-1$ is given by Lemma A.2.2. In matrix notation, we have $p_l^{n+1} = M_{\text{Mult}}^l p_l^n$, where M_{Mult}^l is an $(N-2) \times (N-2)$ iteration matrix. For notational convenience, we define $c = \lambda\alpha_l^+$. The iteration matrix is then given by

$$M_{\text{Mult}}^l = \begin{bmatrix}
(1-c)^2 & & & & & & & & \\
(1-c)^2 & 0 & & & & & & & \\
c(1-c) & c(1-c) & (1-c)^2 & & & & & & \\
c(1-c) & c(1-c) & (1-c)^2 & 0 & & & & & \\
0 & c^2 & c(1-c) & c(1-c) & (1-c)^2 & & & & \\
 & c^2 & c(1-c) & c(1-c) & (1-c)^2 & 0 & & & \\
 & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & & c^2 & c(1-c) & c(1-c) & (1-c)^2 & 0 & \\
 & & & & 0 & c^2 & c(1-c) & c(1-c) & (1-c)^2
\end{bmatrix}$$

Note that M^l_{Mult} is a lower triangular matrix and hence its eigenvalues $(\mu^l_{\text{Mult}})_i$, $i = 1, 2, \ldots, N-2$, are given by the entries on the diagonal; i.e., they are either equal to either 0 or $(1 - \lambda\alpha^+_l)^2$. With $0 < |\lambda\alpha^+_l| \leq 1$, the eigenvalues of M^l_{Mult} satisfy $|(\mu^l_{\text{Mult}})_i| < 1$. Similar results are obtained for the scalar equations with negative characteristic speeds using Lemma A.2.2. Hence, the two-grid multiplicative scheme is convergent. We note that convergence can also be proved for the multilevel case, details of which are omitted.

For a k-level additive scheme, we have $p^{n+1}_l = M^l_{\text{Add}} p^n_l$, where M^l_{Add} is an $(N - 2^{k-1}) \times (N - 2^{k-1})$ iteration matrix constructed from Lemma A.2.3. M^l_{Add} is a lower triangular matrix with eigenvalues, $(\mu^l_{\text{Add}})_i$, $i = 1, 2, \ldots, N-2^{k-1}$, equal to either 0 or $(1-\lambda\alpha^+_l)$. The CFL condition $0 < |\lambda\alpha^+_l| \leq 1$ implies that $|(\mu^l_{\text{Add}})_i| < 1$. Therefore, the multilevel additive scheme is convergent for all the scalar characteristic equations with positive characteristic speeds. It can be proved in a similar way that the scheme is convergent for the scalar equations with negative characteristic speeds using Lemma A.2.3. The multilevel additive scheme is convergent in the W-space and hence convergent in the U-space. □

2.5 Two-Dimensional Systems of Hyperbolic Conservation Laws

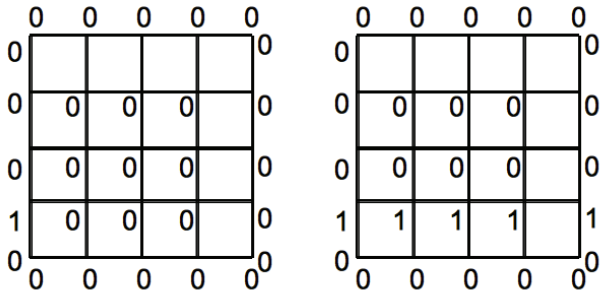
A two-dimensional system of hyperbolic conservation laws can be written as

$$U_t + F(U)_{x_1} + G(U)_{x_2} = 0, \quad 0 < x_1, x_2 < 1, \quad t > 0, \tag{2.28}$$

where $F(U)$ and $G(U)$ are m-dimensional flux vectors in the x_1 and x_2-dimensions respectively and U is an m-dimensional vector of conserved quantities. We will apply the idea of the one-dimensional multigrid time stepping scheme to two dimensions. In particular, we will describe the details of the construction of the coarse grids, the interpolation, and the restriction operators.

Standard multigrid methods for two-dimensional problems use either the typical fully coarsened grid or a semi-coarsened grid. There are a few inherent issues with these types of coarse grids for hyperbolic problems which may results in slow convergence of the multigrid method. Suppose we are solving the scalar wave equation $u_t + u_x = 0$ in two dimensions with initial conditions as given in Fig. 2.3(a) using a standard fully-coarsened grid. Note that the steady state solution, shown in Fig. 2.3(b), is nonzero only along a single noncoarse grid line and zero elsewhere. If injection is used for restriction of the solution, then this information on the noncoarse grid line will completely go unseen on the coarse grid. Hence, the coarse grid evolution of the solution would not help in accelerating the wave propagation. Full

weighting restriction would spread out the solution so that the solution on the coarse grid is not zero, but still it does not resolve the fundamental issue.



(a) (b)

Figure 2.3: Demonstration of the scenario in which multigrid method using standard fully coarsened grid would fail.(a) The initial and boundary conditions for $u_t + u_x = 0$.(b) Steady state solution is nonzero only along a single noncoarse grid line.

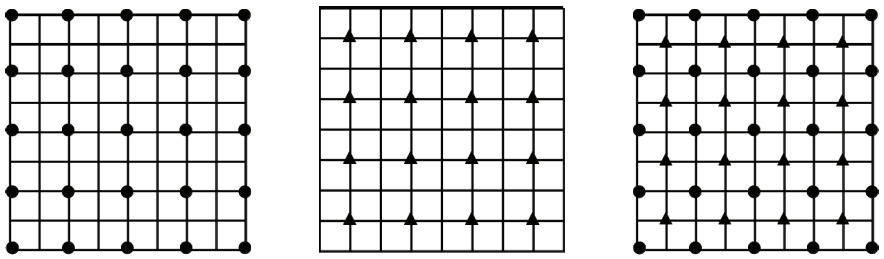
One possible solution to avoid the problem of missing information is to use a semi-coarsened grid [85, 104]. However with the use of explicit time marching schemes for hyperbolic equations, it is necessary to satisfy the CFL condition, which limits the size of the time step by the minimum of the spatial grid size. In semi-coarsening, the spatial grid size increases by a factor of two along one dimension while the spatial grid size in the other dimension remains unchanged from that of the fine grid. Hence, the time step size on the coarse grid cannot be larger than the time step size on the fine grid. As a result, using a semi-coarsened grid would not accelerate wave propagation as there is no increase in the size of the time step on the coarse grid.

To address the issue, we propose to use two fully coarsened grids as shown in Figs. 2.4(a) and 2.4(b). Given a fine grid $((x_1)_i^h, (x_2)_j^h)$, $i, j = 0, 1, \cdots, N-1$, the coarse grid shown in Fig. 2.4(a) is defined such that the grid points with even indices for i and j are selected as coarse grid points: $((x_1)_i^{1,H}, (x_2)_j^{1,H})$, $i, j = 0, 2, \ldots, N-1$. The other coarse grid as shown in Fig. 2.4(b) is defined such that the grid points with odd indices for i and j are selected as coarse grid points: $((x_1)_i^{2,H}, (x_2)_j^{2,H})$, $i, j = 1, 3, \ldots, N-2$. The superscripts h, $\{1, H\}$ and $\{2, H\}$ represent functions on the fine and the two coarse grids, respectively.

The grid size for the coarse grids $\Omega^{1,H}$ and $\Omega^{2,H}$ increases by a factor of two from that of the fine grid in both dimensions as shown in Figs. 2.4(a) and 2.4(b). This means that a time step of size $2\Delta t^h$ can be used for both these coarse grids without violating the CFL

condition, in contrast with the semi-coarsening approach. Furthermore, from the combined coarse grid, Ω^H, shown in Figure 2.4(c), we can see that coarse grid points are selected in such a way that coarse grid points exist on every grid line both in the x_1 and x_2 directions. This ensures that there is no loss of information due to disturbances along the noncoarse grid lines as in the case of full coarsening.

Further coarse grids are obtained by recursively applying the coarsening technique described above to each of the grids $\Omega^{1,H}$ and $\Omega^{2,H}$. Figure 2.5 shows the coarsening of the grid $\Omega^{1,H}$. The computational complexity of our approach is the same as that of a multigrid method using a semi-coarsened grid. For three-dimensional hyperbolic equations, our proposed coarsening technique would include four coarse grids such that coarse grid points exist on every plane and line, with the coarse grid size double that of the fine grid size in all dimensions. This avoids pathological cases such as those in Figure 2.3. The complexity of this coarsening is also same as that of semi-coarsening along one-dimension.

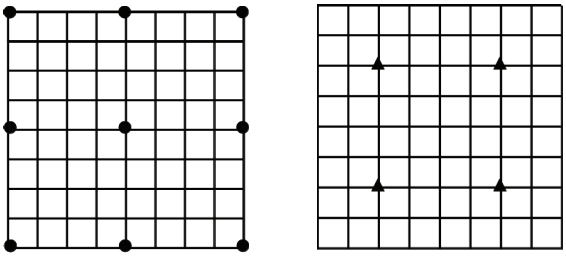


(a) (b) (c)

Figure 2.4: The coarse grids used in the two-dimensional multigrid time stepping scheme.(a) The dots represent the coarse grid points on $\Omega^{1,H}$.(b) The triangles represent the coarse grid points on $\Omega^{2,H}$.(c) Ω^H, the grid obtained when the two coarse grids $\Omega^{1,H}$ and $\Omega^{2,H}$ are put together.

2.5.1 Interpolation

In this section, we describe how a coarse grid solution is interpolated to the fine grid. At the coarse grid points, we just copy the values to the fine grid as shown in Fig. 2.4(c). For interpolation of the solution at the noncoarse grid points, $((x_1)_i^h, (x_2)_j^h)$, we will apply the same idea as in one dimension. More precisely, we solve a local boundary value problem at $((x_1)_i^h, (x_2)_j^h)$; see Figure 2.6. Notice that all the neighboring grid points of $((x_1)_i^h, (x_2)_j^h)$ are



(a) (b)

Figure 2.5: The grids obtained from coarsening the grid $\Omega^{1,H}$.

coarse grid points whose values are known from the coarse grid solution. The interpolated value of $U^h_{i,j}$ is then given by the steady state solution of (2.28).

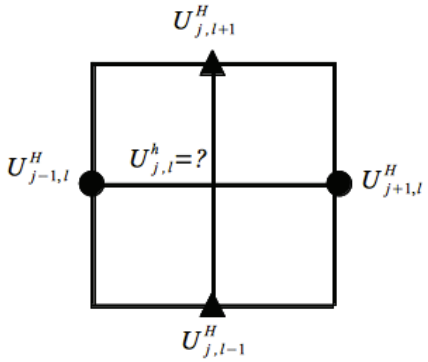


Figure 2.6: Construction of a local boundary value problem for interpolation of the solution at $((x_1)^h_i, (x_2)^h_j)$.

2.5.2 Residual Restriction

The idea for residual restriction in two dimensions is the same as that described for one-dimensional systems in Section 2.3.2. We will split the residual into the forward and backward components and then perform upwind restriction on the corresponding components accordingly. On the fine grid, for $i, j = 1, 2, \ldots, N-2$, the forward and backward

components of the residual are defined as

$$(R^h)^+_{i,j} = \frac{1}{\Delta x^h_1}[F^+_{i,j}(U^h_{i,j}) - F^+_{i-1,j}(U^h_{i-1,j})] + \frac{1}{\Delta x^h_2}[G^+_{i,j}(U^h_{i,j}) - G^+_{i,j-1}(U^h_{i,j-1})] - (B^h)^+_{i,j},$$
$$(R^h)^-_{i,j} = \frac{1}{\Delta x^h_1}[F^-_{i+1,j}(U^h_{i+1,j}) - F^-_{i,j}(U^h_{i,j})] + \frac{1}{\Delta x^h_2}[G^-_{i,j+1}(U^h_{i,j+1}) - G^-_{i,j}(U^h_{i,j})] - (B^h)^-_{i,j},$$

where B^h is the RHS on the fine grid, F^+, G^+ and F^-, G^- are the forward and backward flux components, Δx^h_1 and Δx^h_2 are the fine grid sizes in the x_1 and x_2-dimensions, respectively. We note that $(R^h)^+_{i,j} = (R^h)^-_{i,j} = 0,\ i,j = 0, N-1$.

We first consider the restriction of the residual on the coarse grid, $\Omega^{1,H}$. For the forward component of the residual, the information is propagating in the positive x_1 and x_2-dimensions. We compute the residual on the coarse grid by averaging along the upwind directions:

$$(R^{1,H})^+_{i,j} = 0.25[(R^h)^+_{i,j} + (R^h)^+_{i-1,j}] + 0.25[(R^h)^+_{i,j} + (R^h)^+_{i,j-1}],\ i,j = 2,4,\ldots,N-3.$$

Similarly, we compute the backward component of the residual on the coarse grid by averaging along the right and top sides:

$$(R^{1,H})^-_{i,j} = 0.25[(R^h)^-_{i,j} + (R^h)^-_{i+1,j}] + 0.25[(R^h)^-_{i,j} + (R^h)^-_{i,j+1}],\ i,j = 2,4,\ldots,N-3.$$

Combining the two components gives the coarse grid residual on $\Omega^{1,H}$:

$$R^{1,H}_{i,j} = (R^{1,H})^+_{i,j} + (R^{1,H})^-_{i,j}, \qquad i,j = 2,4,\ldots,N-3.$$

For the other coarse grid, $\Omega^{2,H}$, the restriction of the residual is computed similarly as

$$R^{2,H}_{i,j} = (R^{2,H})^+_{i,j} + (R^{2,H})^-_{i,j}, \qquad i,j = 1,3,\ldots,N-2,$$

where

$$\begin{aligned}(R^{2,H})^+_{i,j} &= 0.25[(R^h)^+_{i,j} + (R^h)^+_{i-1,j}] + 0.25[(R^h)^+_{i,j} + (R^h)^+_{i,j-1}],\\ (R^{2,H})^-_{i,j} &= 0.25[(R^h)^-_{i,j} + (R^h)^-_{i+1,j}] + 0.25[(R^h)^-_{i,j} + (R^h)^-_{i,j+1}],\\ &\qquad\qquad i,j = 1,3,\ldots,N-2.\end{aligned}$$

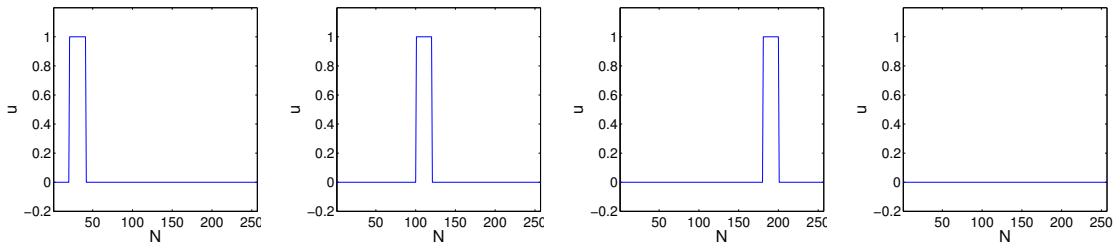
2.6 Numerical Results

In this section, we demonstrate the effectiveness of our multigrid time stepping scheme numerically by several examples. The stopping criteria is that the residual norm is smaller than 10^{-6}. First we provide the results for a simple linear system and then for other nonlinear systems including the Euler equations.

Example 2.6.1. We numerically verify that the multigrid time stepping schemes are nonoscillatory and have optimal speed of propagation in reaching the steady state solution. The linear system of hyperbolic conservation laws is given by

$$\begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} v \\ u \end{pmatrix}_x = 0, \qquad 0 < x < 1, \;\; t > 0.$$

The boundary condition is chosen as 0 and the initial condition is a square wave for both u and v. The steady state solution for this system is $(u, v) \equiv 0$ as given in Figure 2.7(d). We apply the three level multiplicative multigrid scheme to solve the system with a CFL number of 1. The plots for u at different multigrid steps are shown in Figure 2.7 for the multiplicative scheme. It can be seen that the square wave remains a square wave as it propagates. We omit the results for v as they are similar to the solution u. While it takes 236 iterations to propagate to the steady state on a fine grid with a grid size of $\Delta x^h = 1/256$, the three level algorithm converges in 29 iterations. Thus, the speed up of the three level method is 8.14. Our multigrid method achieves speed up close to 2^L, where L is the number of multigrid levels, while the theoretical optimal speed up of a standard multigrid method was predicted to be $2^L - 1$ by [46, 66].



(a) (b) (c) (d)

Figure 2.7: The numerical solution given by the three level multiplicative scheme for one-dimensional linear system, $\frac{\Delta t}{\Delta x} = 1$.(a) multigrid step = 0,(b) multigrid step = 10,(c) multigrid step = 20,(d) multigrid step = 29.

Example 2.6.2. We consider a system of nonlinear hyperbolic conservation laws described by

$$\begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} \frac{1}{2}u^2 + \frac{1}{2}v^2 \\ \frac{1}{2}v^2 + \frac{1}{2}uv \end{pmatrix}_x = 0, \qquad 0 < x < 1, \; t > 0,$$

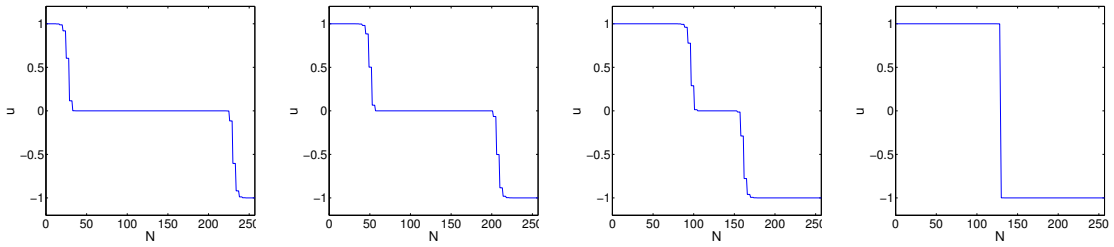
with the Neumann boundary condition and the initial condition

$$u(x,0) = v(x,0) = \begin{cases} 1 & \text{if } x = 0, \\ -1 & \text{if } x = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The steady state solution is obtained as

$$u = v = \begin{cases} 1 & \text{if } x < \frac{1}{2}, \\ -1 & \text{if } x > \frac{1}{2}. \end{cases}$$

Figure 2.8 shows the plots for u at different multigrid steps for a four level multiplicative scheme with a fine grid size of $\Delta x^h = 1/256$ and a CFL number of 0.1. The steady state solution is given by Figure 2.8(d). The solution v shows similar results which are not presented here. It is observed from Figure 2.8 that there are no oscillations in the solution and the shock is resolved as a sharp discontinuity without any wiggles. This ensures for the optimal speed of propagation of the wave resulting in convergence in a small number of multigrid cycles.



(a) (b) (c) (d)

Figure 2.8: The numerical solution given by the four level multiplicative scheme, $\frac{\Delta t}{\Delta x} = 0.1$.(a) multigrid step = 20,(b) multigrid step = 40,(c) multigrid step = 80,(d) multigrid step = 120.

The speed ups for different multigrid levels and different grid sizes for multiplicative scheme are given in Table 2.1. The speed up consistently increases with the increase in the number of coarse grids used.

Example 2.6.3. The one-dimensional Euler equations of gas dynamics constitute a hy-

Table 2.1: Speed up for different multigrid levels and different grid spacings for a one-dimensional nonlinear hyperbolic system

	1/256	1/512	1/1024	1/2048	1/4096
$L = 2$	3.91	3.96	3.98	3.99	3.99
$L = 3$	10.98	11.25	11.44	11.51	11.56
$L = 4$	26.14	27.82	28.75	29.17	29.46
$L = 5$	56.83	63.10	66.84	68.91	69.99

perbolic system of conservation laws given by

$$\begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}_t + \begin{pmatrix} \rho u \\ \rho u^2 + P \\ u(E+P) \end{pmatrix}_x = 0, \tag{2.29}$$

where ρ, ρu and E are conserved quantities of density, momentum and energy. u is the velocity and P is the pressure. The equation of state is given by

$$P = (\gamma - 1)(E - 0.5\rho u^2), \ \gamma = 1.4. \tag{2.30}$$

We compute the steady state solution for (2.29) using two types of initial value problems (Sod tube and Lax tube) and the Neumann boundary condition. The initial data for the Sod tube problem is given by

$$(\rho, \rho u, E)(x, 0) = \begin{cases} (1, 0, 2.5) & \text{if } x < \frac{1}{2}, \\ (0.125, 0, 0.25) & \text{if } x > \frac{1}{2}. \end{cases}$$

Figures 2.9, 2.10 and 2.11 show the density, velocity and pressure plots respectively at different intermediate multigrid steps computed by a four level multiplicative scheme with a CFL number of 0.1 for a grid size of 1/1024. Note that the multiplicative scheme introduces minor oscillations. This behavior is expected as the multiplicative scheme in general does not preserve monotonicity and is not TVD when the number of grids is greater than two [104]. However, these oscillations are very small and do not affect the speed of convergence of the multigrid algorithm as shown in Table 2.2.

Example 2.6.4. We apply the multigrid schemes to (2.29) for the Lax tube problem in which the initial conditions are given by

$$(\rho, \rho u, E)(x, 0) = \begin{cases} (0.445, 0.311, 8.928) & \text{if } x < \frac{1}{2}, \\ (0.5, 0, 1.4275) & \text{if } x > \frac{1}{2}. \end{cases}$$

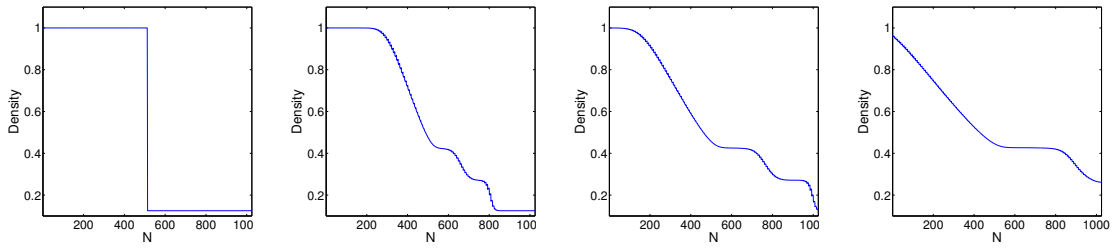


Figure 2.9: Density plots for Sod tube problem given by the four level multiplicative scheme, $\frac{\Delta t}{\Delta x} = 0.1$.(a) multigrid step = 0,(b) multigrid step = 60,(c) multigrid step = 100,(d) multigrid step = 150.

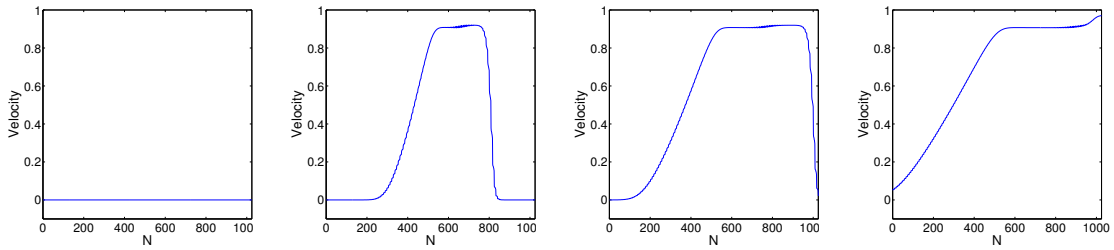
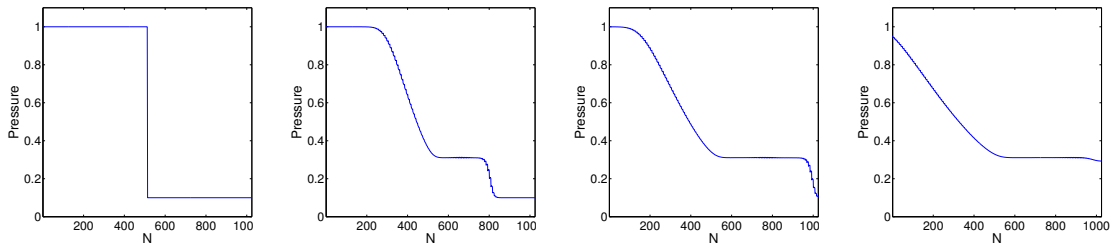


Figure 2.10: Velocity plots for Sod tube problem given by the four level multiplicative scheme, $\frac{\Delta t}{\Delta x} = 0.1$.(a) multigrid step = 0,(b) multigrid step = 60,(c) multigrid step = 100,(d) multigrid step = 150.

Table 2.2: Speed up for Sod tube problem

	1/64	1/128	1/256	1/512	1/1024
$L = 2$	3.26	3.35	3.41	3.44	3.49
$L = 3$	8.39	8.17	8.63	8.80	8.98
$L = 4$	19.40	18.00	18.83	19.82	20.33
$L = 5$	42.73	37.24	38.11	40.83	42.87



(a) (b) (c) (d)

Figure 2.11: Pressure plots for Sod tube problem given by the four level multiplicative scheme, $\frac{\Delta t}{\Delta x} = 0.1$.(a) multigrid step = 0,(b) multigrid step = 60,(c) multigrid step = 100,(d) multigrid step = 150.

Table 2.3: Speed up table for Lax tube problem.

	1/64	1/128	1/256	1/512	1/1024
$L = 2$	3.86	3.89	3.89	3.91	3.95
$L = 3$	9.77	9.99	10.00	9.93	10.23
$L = 4$	21.34	22.23	22.57	22.31	22.37
$L = 5$	41.59	45.75	47.79	47.57	47.37

The convergence results for Lax tube problem for different levels of the multigrid method are given in Table 2.3. The speed up results again show that taking larger time steps on coarser grids results in faster multigrid convergence.

Example 2.6.5. The two-dimensional Euler equations constitute a system of hyperbolic conservation laws described by

$$\begin{pmatrix} \rho \\ \rho u \\ \rho v \\ E \end{pmatrix}_t + \begin{pmatrix} \rho u \\ \rho u^2 + P \\ \rho uv \\ u(E+P) \end{pmatrix}_{x_1} + \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + P \\ v(E+P) \end{pmatrix}_{x_2} = 0,$$

and the two-dimensional equation of state is given by

$$P = (\gamma - 1)(E - 0.5\rho(u^2 + v^2)), \qquad \gamma = 1.4.$$

We conduct our experiments on Riemann problems for two-dimensional gas dynamics in which the initial data are constant in each quadrant [81]. The initial data are restricted so that the only types of behavior allowed at interfaces are a one-dimensional elementary wave, a one-dimensional shock, a one-dimensional rarefaction and a contact discontinuity. There are a total of nineteen genuinely different configurations for a polytropic gas [81]. We test our algorithm on six different initial configurations.

The rectangular domain for the computations is $\{(x_1, x_2)|0 < x_1, x_2 < 1\}$. The initial conditions are listed as (P_l, ρ_l, u_l, v_l), where l refers to the quadrant number. The initial data for the six configurations are given in Table 2.4. Here C denotes configuration and the numbers listed for each configuration refer to those used in [81].

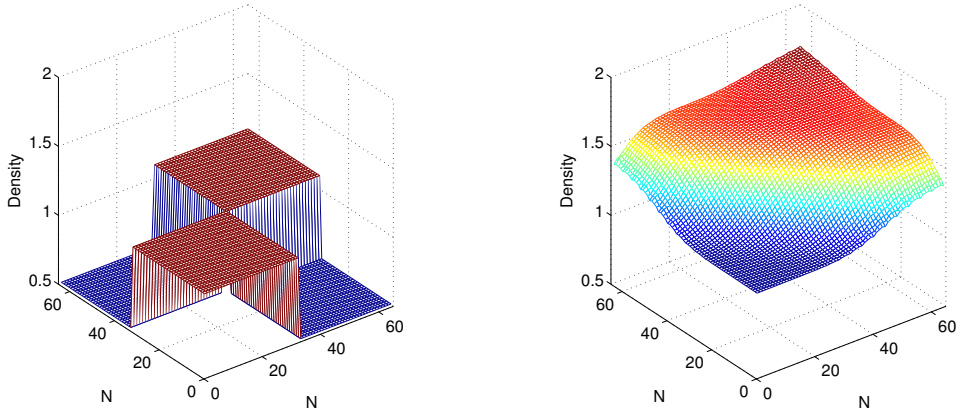
We apply the multiplicative multigrid scheme to compute the steady state solution of the Euler equations. Figures 2.12 and 2.13 demonstrates the propagation of the density and velocity respectively using a three level multiplicative scheme on a fine grid size of 65×65 for the initial configuration 4. It can be seen that the solution converges smoothly to the steady state without any major oscillations. The steady state solutions for density and velocity are shown in Figures 2.12(d) and 2.13(d), respectively. On a single grid, the FVS scheme (2.2) takes 4340 iterations to converge to the steady state. The three level multiplicative scheme takes 510 multigrid cycles giving a speed up of 8.51. In Table 2.5, we present the speed up for different levels of the multiplicative method for the six initial configurations for a fine grid size of 65×65. While the theoretical optimal speed up of an L-grid method was computed asymptotically as $2^L - 1$ [46, 66], in practice our multigrid scheme achieves speed up much higher than the theoretical optimal for certain initial configurations as shown in Table 2.5.

Table 2.4: Initial data for the six configurations for two-dimensional Euler equations.

Quadrant 2: (P_2, ρ_2, u_2, v_2)	Quadrant 1: (P_1, ρ_1, u_1, v_1)
C1: $(0.4, 0.5197, -0.7259, 0)$ C2: $(0.4, 0.5197, -0.7259, 0)$ C4: $(0.35, 0.5065, 0.8939, 0)$ C5: $(1, 2, -0.75, 0.5)$ C9: $(1, 2, 0, -0.3)$ C11: $(0.4, 0.5313, 0.8276, 0)$	C1: $(1, 1, 0, 0)$ C2: $(1, 1, 0, 0)$ C4: $(1.1, 1.1, 0, 0)$ C5: $(1, 1, -0.75, -0.5)$ C9: $(1, 1, 0, 0.3)$ C11: $(1, 1, 0.1, 0)$
Quadrant 3: (P_3, ρ_3, u_3, v_3)	Quadrant 4: (P_4, ρ_4, u_4, v_4)
$C1$: $(0.0439, 0.1072, -0.7259, -1.4045)$ $C2$: $(1, 1, -0.7259, -0.7259)$ $C4$: $(1.1, 1.1, 0.8939, 0.8939)$ $C5$: $(1, 1, 0.75, 0.5)$ $C9$: $(0.4, 1.039, 0, -0.8133)$ $C11$: $(0.4, 0.8, 0.1, 0)$	$C1$: $(0.15, 0.2579, 0, -1.4045)$ $C2$: $(0.4, 0.5197, 0, -0.7259)$ $C4$: $(0.35, 0.5065, 0, 0.8939)$ $C5$: $(1, 3, 0.75, -0.5)$ $C9$: $(0.4, 0.5197, 0, -0.4259)$ $C11$: $(0.4, 0.5313, 0.1, 0.7276)$

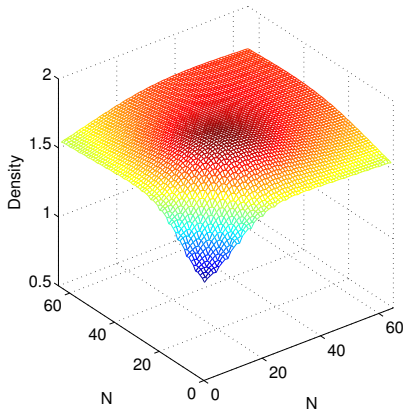
Table 2.5: Speed ups of multiplicative scheme for the 6 initial configurations for two-dimensional Euler equations for a fine grid size of 65×65.

	C1	C2	C4	C5	C9	C11
$L = 2$	3.22	4.19	3.71	4.35	4.76	3.87
$L = 3$	9.11	12.04	8.51	6.13	22.43	7.83
$L = 4$	18.69	25.78	12.88	22.44	40.37	25.58

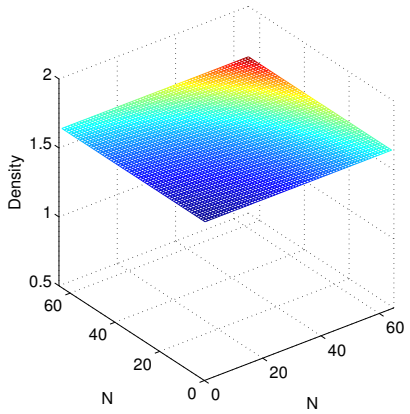


(a)

(b)



(c)



(d)

Figure 2.12: Density plots for two-dimensional Euler equations (configuration 4) given by the three level multiplicative scheme, $\lambda = 0.1$.(a) multigrid step = 0,(b) multigrid step = 50,(c) multigrid step = 100,(d) multigrid step = 510.

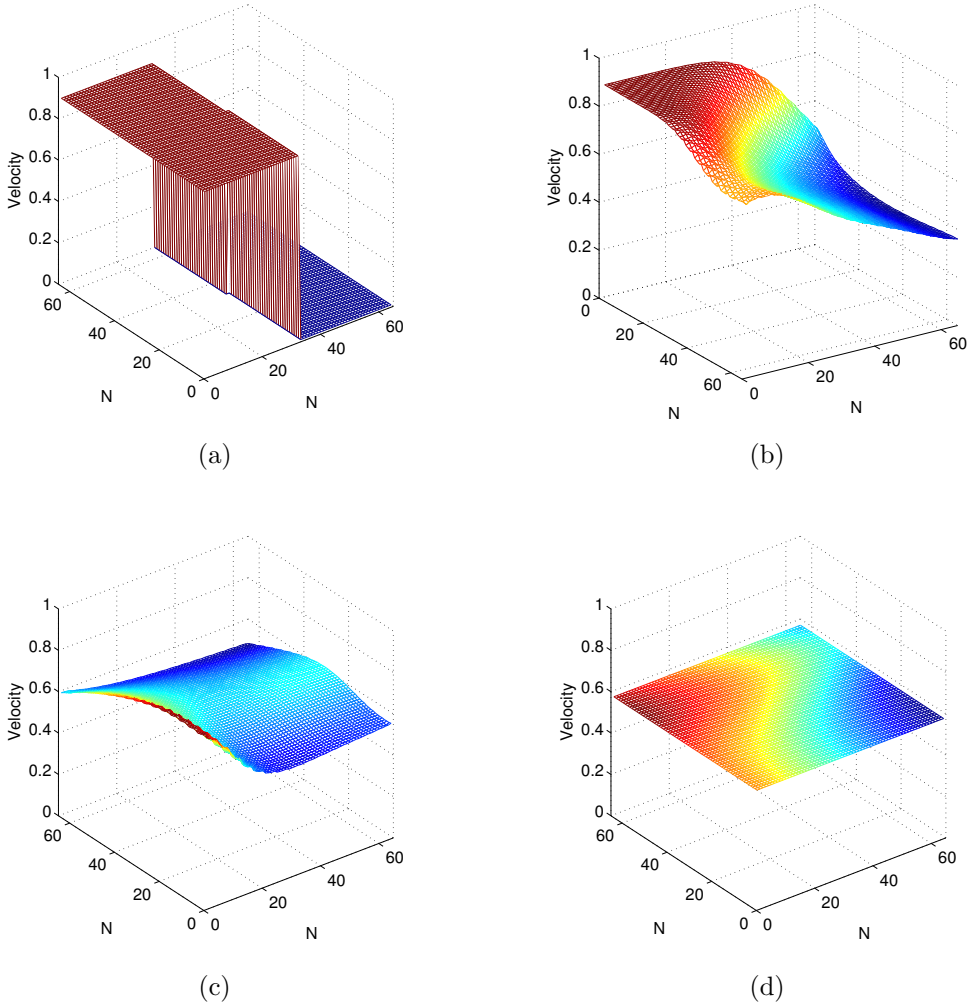


Figure 2.13: Velocity plots for two-dimensional Euler equations (configuration 4) given by the three level multiplicative scheme, $\lambda = 0.1$.(a) multigrid step = 0,(b) multigrid step = 50,(c) multigrid step = 100,(d) multigrid step = 510.

As a comparison, Kanarachos et al. [70] developed a multigrid method for the Euler equations using full weighting and bilinear interpolation techniques. With three grid levels, they obtain a reduction in the number of work units by a factor of 5 to 9 (compared to a single grid), while our upwind biased multigrid technique gives a factor of 4 to 13 for the different initial configurations.

2.7 Conclusion

As with the scalar conservation laws, we have shown that the upwind biased restriction and interpolation techniques result in efficient multigrid methods for systems of hyperbolic conservation laws as well. Although the systems are a mixture of characteristics, we can still exploit the direction of propagation of information for the restriction and interpolation operations. The interpolation is formulated as solving a local Riemann problem to steady state and we perform upwind restriction by splitting the residual into positive and negative components based on the idea of flux vector splitting. We theoretically proved that both the additive and multiplicative multigrid schemes satisfy the total variation diminishing property for one-dimensional linear systems. We also proved that these schemes are consistent and convergent for one-dimensional linear systems. We showed that the upwind techniques can be extended to higher dimensional systems and also introduced a novel coarsening technique. The numerical experiments for both one and two-dimensional hyperbolic systems have consistently shown that our multigrid methods result in smooth and fast convergence to the steady state solution. The upwind multigrid methods reduce the number of work units by a factor of 4 to 13 (compared to the single grid), while factors of 3 to 9 are reported in the literature.

Chapter 3

Numerical Methods for HJB and HJBI Systems

3.1 Introduction

In this chapter, we develop numerical methods for systems of nonlinear HJB PDEs resulting from dynamic Bertrand oligopoly, American options under regime switching and systems of nonlinear HJBI PDEs resulting from American options with unequal lending/borrowing rate under regime switching. A consistent, l_∞ stable and monotone scheme converges to a viscosity solution [11]. Among these properties monotonicity is, in general, hard to achieve. Discretization schemes that converge to the viscosity solution for HJB/HJBI equations associated with American options under regime switching applications have been studied by [37, 59]. Monotone discretization schemes for HJB equations from dynamic Bertrand oligopoly have not been developed before.

We first construct a fully implicit, positive coefficient discretization which results in an unconditionally monotone numerical scheme for the scalar HJB PDE resulting from dynamic Bertrand monopoly case. For duopoly problem, the Bertrand model results in a two-dimensional system of HJB PDEs, for which we use the extended definition of the viscosity solution from [61]. We still develop positive coefficient discretization as this is essential for a monotone scheme. The presence of cross derivative terms makes this task non-trivial. Since there is only one cross derivative term in the two-dimensional HJB system, we adopt the computationally efficient approach of enforcing a spacing restriction on the original finite difference grid such that a positive coefficient condition results [26].

We theoretically prove that the discretization schemes converge to the viscosity solution for both the monopoly and duopoly cases.

Among the solvers, policy iteration [12, 57, 80] is an efficient and convergent solver for the discrete HJB equations. However, convergence is not guaranteed for HJBI equations [16, 103]. An alternative is to use relaxation-type iterative methods [9], which are convergent for both HJB and HJBI equations. A major drawback for relaxation methods is its slow convergence. Therefore, we propose to develop multigrid methods based on FAS. The components of the multigrid method are very specific for the underlying PDE. To the best of our knowledge, multigrid methods have not been developed for systems of HJB and HJBI PDEs previously.

The control and the solution of the HJB/HJBI equations are highly nonlinearly coupled. Standard multigrid techniques do not work well for problems with jumps in control [48], which often happens in practice. In [1, 3, 2, 4], multigrid-Howard and full multigrid-Howard are used to solve portfolio selection problems modeled as HJB equations. They use policy iteration and standard multigrid to solve the linearized problem in each iteration. Multigrid methods for HJB equations proposed by [54] can be directly applied to the nonlinear problem. This approach is different from FAS, in particular, the way of constructing the coarse grid problem. We refer the reader to [48] for the details. The convergence of these methods is slow because jumps in control are ignored. Han et al. [48] propose a multigrid method using a damped relaxation scheme as the smoother and grid transfer techniques which address the issue of jumps in the control. This method is computationally inefficient as the control set size increases.

A number of multigrid methods have been proposed for linear complementarity problems (LCP) and variational inequalities [19, 55, 73, 87, 91]. Kornhuber [73] proposed a two-stage method consisting of a globally convergent descent method and a subsequent constrained Newton linearization for fast solution of variational inequalities. In [91], one-sided grid transfer operators are used at the free boundary to preserve complementarity conditions on all grid levels for LCPs arising from American options. In some special cases, the complementarity problems can be formulated as HJB equations. Multigrid methods for LCP with an application to American options were proposed by [87]. They use projected pointwise Gauss-Siedel smoother and standard grid transfer operators. The LCP formulation treats the unknown boundary explicitly in a post processing step. Therefore, there is no issue of jumps in control in this formulation. Moreover, not all HJB and HJBI equations can be formulated as complementarity problems. Therefore, techniques for one formulation are not generally applicable to the other. Furthermore, none of these methods are designed for systems of nonlinear HJB and HJBI PDEs.

3.2 Model Problems

In this section, we present the model problems and its discretization in detail starting with dynamic Bertrand oligopoly, followed by American options under regime switching and American options with unequal lending/borrowing rates with stock borrowing fees under regime switching.

3.2.1 Dynamic Bertrand Oligopoly

We consider a market with two firms which make their decisions dynamically through time. Each firm has a fixed lifetime capacity of production at time $t = 0$ denoted by $x_l(0), l = 1, 2$. At any later time t, the remaining capacity is given by $x_l(t)$. When $x_l(t) = 0$, the firm has exhausted its capacity and is out of business. The cost of production in the dynamic game is assumed to be zero. However it is noticed that shadow costs associated with the scarcity of goods as they run down are introduced in the system [77].

The price of the good for each firm in general depends on the capacity of all firms, i.e., $p_l = p_l(\mathbf{x}(t))$, where $\mathbf{x}(t) = (x_1(t), x_2(t))$ and is chosen by a Markovian dynamic strategy. Given these prices, each firm expects the market to demand at a rate $D_l(p_1, p_2)$, which is affine in prices [77]:

$$D_l(p_1, p_2) = a_1 - a_2\, p_l + a_3\, p_m, \qquad l, m = 1, 2,\ m \neq l,$$

where a_1, a_2, a_3 are positive parameters such that $a_2 > a_3$. The intercept parameter a_1 is a measure of general level of demand due to business cycles and recessions and a_3/a_2 is the measure of substitutability. The actual demand from the market $d_l(t)$, however, undergoes short term unpredictable fluctuations, which is modeled by

$$d_l(t) = D_l(p_1, p_2) - \sigma_l\, \vartheta_l(t), \qquad l = 1, 2,$$

where $\vartheta_l(t)$ are correlated Gaussian white noise sequences and σ_l is the volatility of the demand of firm l. The lifetime capacity of each firm depletes over time according to the market demand for its good. As a result, the dynamics of the lifetime capacity of the firms are given by $dx_l(t) = -d_l(t)dt$. Consequently, the stochastic differential equation for the lifetime capacity is given by

$$dx_l(t) = -D_l\left(p_1, p_2\right) dt + \sigma_l\, dW_l(t), \quad \text{if } x_l > 0,\ l = 1, 2,$$

where $W_l(t)$ are correlated Brownian motions. If $x_l(t) = 0$, then for all $s \geq t$, $x_l(s) = 0$.

Given the initial lifetime capacity $x_l(0) > 0$, the players seek to maximize their expected discounted lifetime profit, also known as the value function $U_l(x_1, x_2)$, in the Nash equilibrium sense. Each player l maximizes its value function by assuming that the other player is using its equilibrium pricing strategy p_m^*:

$$U_l(x_1,x_2) = \sup_{p_l \geq 0} \mathbb{E}\left\{ \int_0^\infty e^{-rt} p_l(\mathbf{x}(t)) D_l\left(p_l(\mathbf{x}(t)), p_m^*(\mathbf{x}(t))\right) \mathbb{1}_{\{x_l(t)>0\}} dt \right\}, \; l,m = 1,2, \; m \neq l, \quad (3.1)$$

where $r > 0$ is the discount rate. Using a dynamic programming argument for nonzero sum differential games [33], the equations for the value functions (3.1) can be reformulated as a system of coupled backward nonlinear HJB PDEs:

$$(U_l)_\tau = \sup_{p_l \geq 0} \left\{ \mathcal{L}^{p_l} U_l \right\}, \quad l = 1,2, \quad (3.2)$$

where the differential operators $\mathcal{L}^{p_l}$ are given by

$$\begin{aligned}
\mathcal{L}^{p_1} &= \frac{1}{2}\sigma_1^2 \frac{\partial^2 U_1}{\partial x_1^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 U_1}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 U_1}{\partial x_2^2} - (a_1 - a_2 p_1 + a_3 p_2^*) \frac{\partial U_1}{\partial x_1} \\
&+ \left[\frac{\gamma}{\eta}(a_1 - a_2 p_1 + a_3 p_2^*) - \frac{\kappa - p_2^*}{\eta} \right] \frac{\partial U_1}{\partial x_2} - r U_1 + p_1 (a_1 - a_2 p_1 + a_3 p_2^*). \\
\mathcal{L}^{p_2} &= \frac{1}{2}\sigma_1^2 \frac{\partial^2 U_2}{\partial x_1^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 U_2}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 U_2}{\partial x_2^2} + \left[\frac{\gamma}{\eta}(a_1 - a_2 p_2 + a_3 p_1^*) - \frac{\kappa - p_1^*}{\eta} \right] \frac{\partial U_2}{\partial x_1} \\
&- (a_1 - a_2 p_2 + a_3 p_1^*) \frac{\partial U_2}{\partial x_2} - r U_2 + p_2 (a_1 - a_2 p_2 + a_3 p_1^*),
\end{aligned}$$

where ρ is the correlation coefficient of the Brownian motions. The parameters κ, η and γ are positive and are defined as [77]

$$\gamma = \frac{a_3}{(a_2 - a_3)(a_2 + a_3)}, \quad \kappa = \gamma \, a_1 \left(\frac{a_2}{a_3} + 1 \right), \quad \eta = \gamma \, \frac{a_2}{a_3}.$$

The parameter γ gives a measure of degree of substitutability. We note that coupling in the system of HJB PDEs (3.2) is due to the equilibrium pricing strategies (p_1^*, p_2^*).

The domain of the PDE is $x_1 > 0, \; x_2 > 0, \; \tau > 0$. For computational purposes, the domain is truncated to $\Omega^D = (x_1, x_2, \tau) \in [0, x_{1,\max}] \times [0, x_{2,\max}] \times [0, T]$. When one firm runs out of capacity, the other has a monopoly. If $u(x, \tau)$ is the value function of a monopolist, then on $x_1 = 0, \; x_2 > 0$, we have $U_1(0, x_2, \tau) = 0, U_2(0, x_2, \tau) = u(x_2, \tau)$. Similarly when $x_1 > 0, \; x_2 = 0$, we have $U_1(x_1, 0, \tau) = u(x_1, \tau)$ and $U_2(x_1, 0, \tau) = 0$. We look for solutions

in which $\lim_{x_l \to \infty} \partial U_l / \partial x_l = 0,\ l = 1, 2$ and $\lim_{x_m \to \infty} \partial U_l / \partial x_m = 0,\ l, m = 1, 2,\ m \neq l$ [77]. As a result, we use Neumann conditions on $x_1 = x_{1,\max}$ and $x_2 = x_{2,\max}$. For a sufficiently large T, the capacities of all firms are exhausted and hence the terminal condition is given by $U_1(x_1, x_2, 0) = U_2(x_1, x_2, 0) = 0$ [21].

When $\gamma = 0$, both firms are monopolists in disjoint markets of their own goods and hence $U_l(x_1, x_2, \tau) = u(x_l, \tau)$. The value function $u(x, \tau)$ of the monopoly firm is given by the following HJB equation [33, 77]

$$u_\tau = \sup_{p \geq 0} \left\{ \mathcal{L}^{M,p} u \right\}, \tag{3.3}$$

where

$$\mathcal{L}^{M,p} u = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} - \frac{1}{\eta}(\kappa - p)\frac{\partial u}{\partial x} - ru + \frac{p}{\eta}(\kappa - p), \tag{3.4}$$

where x denotes the firm's remaining life time capacity and p is the price of the firm's good. The domain of the PDE is $x > 0,\ \tau > 0$. The boundary conditions are $u(0) = 0$ and $\lim_{x \to \infty} \partial u / \partial x = 0$. For computational purposes, the domain is truncated to $\Omega^M = (x, \tau) \in [0, x_{\max}] \times [0, T]$.

Ledvina et al. [77] mainly consider the analysis of the HJB PDEs (3.2) and (3.3). In this thesis, we develop a fully implicit, positive coefficient, finite difference discretization schemes for (3.2) and (3.3). We prove that our discretization schemes converge to the viscosity solution in Section 3.3.

3.2.1.1 Discretization

We first briefly provide the discretization details for the scalar monopoly problem (3.3). The spatial domain is discretized into a set of nodes $\{x_0, x_1, \ldots, x_{N-1}\}$ with a uniform grid spacing Δx. Let u_i^n be the approximate solution of (3.3) at (x_i, τ^n). We assume a mesh and control discretization parameter h such that

$$\Delta x = C_1 h, \quad \Delta \tau = C_2 h, \quad \Delta p = C_3 h, \tag{3.5}$$

where C_1, C_2 and C_3 are constants independent of h.

The differential term $\mathcal{L}^{M,p} u$ in (3.3) is discretized using a fully implicit finite difference discretization, which results in

$$\mathcal{L}_h^{M,p} u_i^n = \alpha_i(p) u_{i-1}^n + \beta_i(p) u_{i+1}^n - (\alpha_i(p) + \beta_i(p) + r) u_i^n + \frac{p}{\eta}(\kappa - p), \tag{3.6}$$

where $\alpha_i(p)$ and $\beta_i(p)$ are given in Algorithm 3 in Appendix B. A combination of central and upstream differencing is used for the drift term such that the coefficients $\alpha_i(p)$ and $\beta_i(p)$ are positive, while ensuring that central differencing is used as much as possible [105]. Using a fully implicit time stepping and (3.6), the discrete form of (3.3) is then given by

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = \sup_{p \geq 0} \left\{ \mathcal{L}_h^{M,p} u_i^{n+1} \right\}. \tag{3.7}$$

We next consider the two-dimensional HJB system (3.2). The spatial domain is discretized into a set of nodes $\{(x_1)_0, (x_1)_1, \ldots, (x_1)_{N_1-1}\} \times \{(x_2)_0, (x_2)_1, \ldots, (x_2)_{N_2-1}\}$ with a uniform grid spacing of size Δx_1 and Δx_2 in the x_1 and x_2 directions, respectively. Let $(U_l)_{i,j}^n$ be the approximate solution of (3.2) at $((x_1)_i, (x_2)_j, \tau^n)$ for $l = 1, 2$. We assume a mesh and control discretization parameter h such that

$$\Delta x_1 = C_4\, h, \quad \Delta x_2 = C_5\, h, \quad \Delta\tau = C_6\, h, \quad \Delta p_1 = \Delta p_2 = C_7\, h, \tag{3.8}$$

where C_4, C_5, C_6 and C_7 are constants independent of h.

A seven point stencil [26] is used to discretize the cross derivative term. For $\rho \geq 0$, the stencil in Figure 3.1(a) is used and the finite difference formula is given by

$$\frac{\partial^2 U_l}{\partial x_1 \partial x_2} \approx \frac{2\,(U_l)_{i,j}^n + (U_l)_{i+1,j+1}^n + (U_l)_{i-1,j-1}^n - (U_l)_{i+1,j}^n - (U_l)_{i-1,j}^n - (U_l)_{i,j+1}^n - (U_l)_{i,j-1}^n}{2\Delta x_1 \Delta x_2}. \tag{3.9}$$

For $\rho < 0$, the stencil in Figure 3.1(b) is used and the corresponding formula is given by

$$\frac{\partial^2 U_l}{\partial x_1 \partial x_2} \approx \frac{-2\,(U_l)_{i,j}^n - (U_l)_{i+1,j-1}^n - (U_l)_{i-1,j+1}^n + (U_l)_{i+1,j}^n + (U_l)_{i-1,j}^n + (U_l)_{i,j+1}^n + (U_l)_{i,j-1}^n}{2\Delta x_1 \Delta x_2}. \tag{3.10}$$

Standard three point central differencing is used for $\frac{\partial^2 U_l}{\partial x_1^2}$ and $\frac{\partial^2 U_l}{\partial x_2^2}$. The first order derivatives are discretized using central differencing as much as possible and forward or backward differencing when central differencing fails to satisfy positive coefficient discretization. The discrete form of the objective function in (3.2) is then given by

$$\begin{aligned}
\mathcal{L}_h^{p_l} U_{i,j}^{n+1} \;=\; & \left((\alpha_l)_{i,j}^{x_1} - \zeta_{i,j}\right)(U_l)_{i-1,j}^{n+1} + \left((\beta_l)_{i,j}^{x_1} - \zeta_{i,j}\right)(U_l)_{i+1,j}^{n+1} \\
& + \left((\alpha_l)_{i,j}^{x_2} - \zeta_{i,j}\right)(U_l)_{i,j-1}^{n+1} + \left((\beta_l)_{i,j}^{x_2} - \zeta_{i,j}\right)(U_l)_{i,j+1}^{n+1} \\
& + \mathbb{1}_{\rho \geq 0}\, \zeta_{i,j} \left((U_l)_{i+1,j+1}^{n+1} + (U_l)_{i-1,j-1}^{n+1}\right) + \mathbb{1}_{\rho < 0}\, \zeta_{i,j} \left((U_l)_{i+1,j-1}^{n+1} + (U_l)_{i-1,j+1}^{n+1}\right) \\
& - \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r\right)(U_l)_{i,j}^{n+1} \\
& + p_l\,(a_1 - a_2 p_l + a_3 (p_m^*)_{i,j}), \qquad\qquad l, m = 1, 2, \quad m \neq l,
\end{aligned} \tag{3.11}$$

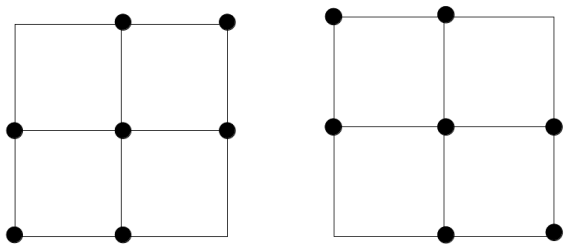


Figure 3.1: Seven point stencil for finite differencing. (a) $\rho \geq 0$. (b) $\rho < 0$.

where $(\alpha_l)_{i,j}^{x_1}, (\beta_l)_{i,j}^{x_1}, (\alpha_l)_{i,j}^{x_2}, (\beta_l)_{i,j}^{x_2}$ and $\zeta_{i,j}$ are given in Algorithm 4 in Appendix B. Using a fully implicit time stepping and (3.11), the discrete form of (3.2) is given by

$$\frac{(U_l)_{i,j}^{n+1} - (U_l)_{i,j}^{n}}{\Delta\tau} = \sup_{p_l \geq 0} \left\{ \mathcal{L}_h^{p_l} (U_l)_{i,j}^{n+1} \right\}, \quad l = 1, 2. \tag{3.12}$$

We note that the presence of the cross derivative term poses a challenge to construct a positive coefficient discretization. The following theorem illustrates how the constraint on grid spacing results in a positive coefficient discretization.

Theorem 3.2.1. *If $\rho \neq 0$ and with the use of a seven point stencil, if the grid spacing is chosen such that*

$$\left| 2\rho \frac{\sigma_2}{\sigma_1} \right| \leq \frac{\Delta x_2}{\Delta x_1} \leq \left| \frac{1}{2\rho} \frac{\sigma_2}{\sigma_1} \right|, \tag{3.13}$$

then the positive coefficient conditions are satisfied, i.e.,

$$(\alpha_l)_{i,j}^{x_1} - \zeta_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_1} - \zeta_{i,j} \geq 0, \tag{3.14}$$

$$(\alpha_l)_{i,j}^{x_2} - \zeta_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_2} - \zeta_{i,j} \geq 0, \tag{3.15}$$

$$\zeta_{i,j} \geq 0, \tag{3.16}$$

$$l = 1, 2, \quad 1 \leq i \leq N_1 - 1, \quad 1 \leq j \leq N_2 - 1.$$

Proof. The use of the seven point stencil ensures that $\zeta_{i,j} \geq 0$ irrespective of the sign of the correlation ρ. This is obvious from (B.9). From Algorithm 4 in Appendix B, conditions (3.14) and (3.15) can be rewritten as

$$(\alpha_l)_{i,j}^{x_1,ups} - \zeta_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_1,ups} - \zeta_{i,j} \geq 0, \tag{3.17}$$

$$(\alpha_l)_{i,j}^{x_2,ups} - \zeta_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_2,ups} - \zeta_{i,j} \geq 0. \tag{3.18}$$

From (B.1), (B.2), (B.5) and (B.6), it is clear that the conditions (3.17) are satisfied when

$$\frac{\sigma_1^2}{2\Delta x_1^2} - \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2} \geq 0. \tag{3.19}$$

Similarly for the x_2 dimension, it is clear from (B.3), (B.4), (B.7) and (B.8), that the conditions (3.18) hold when

$$\frac{\sigma_2^2}{2\Delta x_2^2} - \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2} \geq 0. \tag{3.20}$$

Conditions (3.19) and (3.20) then result in the restriction on the grid spacing as given in (3.13). □

3.2.2 American Options under Regime Switching

Consider a regime switching model with N_m regimes and a finite set of discrete volatilities $\sigma_j,\ j = 1, 2, \ldots, N_m$. A continuous Markov chain process controls the shifts between these regimes. The stochastic process for the underlying asset S under the real world measure is

$$dS = \mu_j^{\mathbb{P}}\ S\ dt + \sigma_j\ S\ dZ + \sum_{m=1}^{N_m} (\xi_{jm} - 1)\ S\ dX_{jm}, \quad j = 1, \ldots, N_m, \tag{3.21}$$

where dZ is the increment of a Wiener process and $\mu_j^{\mathbb{P}}$ is the drift in regime j. The superscript $\mathbb{P}$ denotes the objective probability measure. The term dX_{jm} is given by

$$dX_{jm} = \begin{cases} 1, & \text{with probability } \lambda_{jm}^{\mathbb{P}} dt + \delta_{jm}, \\ 0, & \text{with probability } 1 - \lambda_{jm}^{\mathbb{P}} dt - \delta_{jm}, \end{cases}$$

where the transition probability $\lambda_{jm}^{\mathbb{P}} \geq 0,\ j \neq m$ and $\lambda_{jj}^{\mathbb{P}} = -\sum_{m \neq j} \lambda_{jm}^{\mathbb{P}}$. The asset price jumps from S to $\xi_{jm} S$ when a transition from j to m occurs and $\xi_{jj} = 1$. The jump amplitudes ξ_{jm} are assumed to be deterministic functions of (S, t). In practice, the quantities ξ_{jm} and λ_{jm} are determined by calibration to market prices [6].

Let $U_j(S, \tau)$ be the no-arbitrage value of the contingent claim in regime j, where $\tau = T - t$ with T being the expiry time of the contingent claim and t the time variable. Consider a hedging portfolio P such that

$$P = -U_j + \Delta_s S + \sum_{m=1}^{N_m - 1} \Delta_m F_m,$$

where Δ_s is the number of units of underlying asset with price S and Δ_m is the number of units of additional hedging instruments with price F_m. It is possible to set up a perfect hedge under the assumption that the set of assets $\{S, F_1, \ldots, F_{N_m-1}\}$ forms a non-redundant set [71]. The existence of a perfect hedge allows us to define the risk neutral transition probabilities λ_{jm} and the quantities ρ_j and λ_j as [59]

$$\rho_j = \sum_{m \neq j} \lambda_{jm}(\xi_{jm} - 1), \quad \lambda_j = \sum_{m \neq j} \lambda_{jm}, \quad \lambda_{jj} = -\lambda_j. \tag{3.22}$$

Let $U = [U_1, U_2, \ldots, U_{N_m}]^T$. The differential operators $\mathcal{L}_j U_j$ and $\mathcal{J}_j U$ are defined as

$$\mathcal{L}_j U_j = \frac{\sigma_j^2 S^2}{2} \frac{\partial^2 U_j}{\partial S^2} + (r - \rho_j) S \frac{\partial U_j}{\partial S} - (r + \lambda_j) U_j, \tag{3.23}$$

$$\mathcal{J}_j U = \sum_{m \neq j} \frac{\lambda_{jm}}{\lambda_j} U_m(\xi_{jm} S, \tau), \tag{3.24}$$

where r is the risk free interest rate. The no-arbitrage price of the American option $U_j(S, \tau)$ is then given by a system of equations [71]

$$\min\left[U_{j,\tau} - \mathcal{L}_j U_j - \lambda_j \mathcal{J}_j U,\ U_j - U^*\right] = 0, \quad j = 1, \ldots, N_m, \tag{3.25}$$

where $U^*(S)$ is the payoff function. We consider the truncated domain $(S, \tau) \in [0, S_{\max}] \times [0, T]$ for computational purposes. At $S = 0$ boundary, we simply solve (3.25) with no additional boundary conditions as there are no incoming characteristics. At $S = S_{\max}$, we follow the standard approach and use a Dirichlet condition with $U(S_{\max}, \tau) = U^*(S_{\max})$ [59]. The initial condition is given by the payoff function at $\tau = 0$, which is denoted by

$$U(S, 0) = U^*(S).$$

The minimization problem (3.25) can be solved in different ways. A straight forward approach is to enforce the constraint explicitly. But, the resulting solution is inconsistent and the option delta is not continuous across the early exercise boundary [36]. Alternatively, (3.25) can be reformulated using different optimal control formulations [59] to overcome the issues. In this thesis, we adopt the penalty method. The penalized form of (3.25) [38] is given by a system of HJB PDEs

$$U_{j,\tau} = \mathcal{L}_j U_j + \lambda_j \mathcal{J}_j U + \max_{\varphi \in \{0,1\}} \left[\varphi \frac{(U^* - U_j)}{\epsilon}\right], \tag{3.26}$$

where ϵ is the penalty parameter and $\varphi \in \{0,1\}$ is the control parameter. For efficient convergence of the multigrid method, it is important that the consistency of control from the fine to the coarse grids is maintained during restriction [48]. This can be easily enforced when the penalty formulation (3.26) is used. We discuss this in detail in Section 3.4.3.

Equation (3.26) can be written in the general form as

$$U_{j,\tau} = \max_{\varphi \in \{0,1\}} \left\{ a_j(S,\tau,\varphi)\frac{\partial^2 U_j}{\partial S^2} + b_j(S,\tau,\varphi)\frac{\partial U_j}{\partial S} - c_j(S,\tau,\varphi)U_j + d_j(S,\tau,\varphi) + \lambda_j \mathcal{J}_j U \right\}, \quad (3.27)$$

where

$$\begin{aligned} a_j(S,\tau,\varphi) &= \tfrac{\sigma_j^2 S^2}{2},\ b_j(S,\tau,\varphi) = S(r-\rho_j), \\ c_j(S,\tau,\varphi) &= (r+\lambda_j+\tfrac{\varphi}{\epsilon}),\ d_j(S,\tau,\varphi) = \varphi\tfrac{U_j^*}{\epsilon}. \end{aligned} \quad (3.28)$$

Fully implicit finite differencing is used to discretize (3.27) as described in the following section.

3.2.2.1 Discretization

We discretize (3.27) using a fully implicit, positive coefficient, finite difference discretization, which ensures convergence to the viscosity solution [37]. The spatial domain is discretized into a set of nodes $\{S_0, S_1, \ldots, S_{N-1}\}$ and the n-th time step is denoted by $\tau^n = n\Delta\tau$, where $\Delta\tau$ is the time step size. Let $U_{i,j}^n$ be the discrete approximation to $U_j(S_i, \tau^n)$. The discretized differential terms in (3.27) are represented as

$$\begin{aligned} &\left(a_j(S,\tau,\varphi)\frac{\partial^2 U_j}{\partial S^2} + b_j(S,\tau,\varphi)\frac{\partial U_j}{\partial S} - c_j(S,\tau,\varphi)U_j \right)_i^{n+1} = \\ &\alpha_{i,j}(\varphi)\, U_{i-1,j}^{n+1} + \beta_{i,j}(\varphi)\, U_{i+1,j}^{n+1} - (\alpha_{i,j}(\varphi) + \beta_{i,j}(\varphi) + c_{i,j}(\varphi))\, U_{i,j}^{n+1}. \end{aligned} \quad (3.29)$$

A weighted average of central and upstream differencing [59] is used such that the positive coefficient condition ($\alpha_{i,j} \geq 0$, $\beta_{i,j} \geq 0$) is satisfied and central differencing is used as much as possible. The details are given in Algorithm 5 in Appendix B.

We append all the discrete vectors of the approximation U_j^n to form a long vector U^n of size NN_m,

$$U^n = \left[U_{0,1}^n, \ldots, U_{N-1,1}^n, \ldots, U_{0,N_m}^n, \ldots, U_{N-1,N_m}^n\right]^T.$$

Let $\mathcal{J}_j^h$ denote the discrete form of the operator $\mathcal{J}_j$, the discretization for the regime switching term $\mathcal{J}_j U$ is then given by

$$[\mathcal{J}_j^h U^{n+1}]_{i,j} = \sum_{m \neq j} \frac{\lambda_{jm}}{\lambda_j} U_m(\min(S_{\max}, \xi_{jm} S_i), \tau^{n+1}), \tag{3.30}$$

where $U_m(\min(S_{\max}, \xi_{jm} S_i), \tau^{n+1})$ is approximated by linear interpolation, which is given by

$$U_m(\min(S_{\max}, \xi_{jm} S_i), \tau^{n+1}) = w_m U_{i_m,m}^{n+1} + (1 - w_m) U_{i_m+1,m}^{n+1}, \quad w_m \in [0, 1]. \tag{3.31}$$

Note that we truncate any jumps that require data outside the computational domain in (3.30). The error due to this approximation is small when $S_{\max}$ is sufficiently large [71]. The HJB system (3.27) is then discretized using (3.29), (3.30) and a fully implicit time stepping as

$$\begin{aligned} \frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta\tau} &= \max_{\varphi \in \{0,1\}} \left\{ \alpha_{i,j}(\varphi) U_{i-1,j}^{n+1} + \beta_{i,j}(\varphi) U_{i+1,j}^{n+1} - (\alpha_{i,j}(\varphi) + \beta_{i,j}(\varphi) + c_{i,j}(\varphi)) U_{i,j}^{n+1} \right. \\ &\quad \left. + d_{i,j}(\varphi) + [\mathcal{J}_j^h U^{n+1}]_{i,j} \right\}, \qquad i < N-1;\ j = 1, \ldots, N_m, \\ U_{i,j}^{n+1} &= U_{i,j}^*, \qquad i = N-1;\ j = 1, \ldots, N_m. \end{aligned} \tag{3.32}$$

3.2.3 American Options with unequal lending/borrowing rates under Regime Switching

Consider the model where the cash borrowing rate, r_b, and the lending rate, r_l, are not necessarily equal (with $r_b \geq r_l$) along with the stock borrowing fees, r_f. Such models result in nonlinear HJB PDEs [13, 37]. We consider these models under a N_m-state regime switching process (3.21) combined with the American early exercise, which results in a system of HJBI PDEs. Let $U_j(S, \tau)$ be the no-arbitrage value of the contingent claim in regime j and $U = [U_1, U_2, \ldots, U_{N_m}]^T$. We define the following differential operators for a long position in the contingent claim:

$$\begin{aligned} \mathcal{L}_j^Q U_j &= \frac{\sigma_j^2 S^2}{2} \frac{\partial^2 U_j}{\partial S^2} + (q_3 q_1 + (1 - q_3)(r_l - r_f) - \rho_j) S \frac{\partial U_j}{\partial S} - (q_3 q_1 + q_2(1 - q_3) + \lambda_j) U_j \\ \mathcal{J}_j U &= \sum_{m \neq j} \frac{\lambda_{jm}}{\lambda_j} U_m(\xi_{jm} S, \tau), \end{aligned}$$

where ρ_j and λ_j are given by (3.22). When combined with American early exercise, we have,

$$\min\left[U_{j,\tau} - \inf_{Q\in\hat{Q}}\{\mathcal{L}_j^Q U_j\} - \lambda_j \mathcal{J}_j U,\ U_j - U^*\right] = 0, \tag{3.33}$$

with $Q = (q_1, q_2, q_3)$ and $\hat{Q} = (\{r_l, r_b\}, \{r_l, r_b\}, 0, 1)$. As mentioned in Section 3.2.2, directly solving the minimization problem (3.33) leads to an inconsistent solution. Therefore, we reformulate (3.33) to a penalty form, which results in the following system of HJBI PDEs,

$$U_{j,\tau} = \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\left\{\mathcal{L}_j^Q U_j + \lambda_j \mathcal{J}_j U + \varphi\frac{(U^* - U_j)}{\epsilon}\right\}. \tag{3.34}$$

We rewrite (3.34) in the general form as

$$\begin{aligned} U_{j,\tau} = \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\Big\{ & a_j(S,\tau,Q,\varphi)\frac{\partial^2 U_j}{\partial S^2} + b_j(S,\tau,Q,\varphi)\frac{\partial U_j}{\partial S} - c_j(S,\tau,Q,\varphi)U_j \\ & + \ d_j(S,\tau,Q,\varphi) + \lambda_j \mathcal{J}_j U\Big\}, \end{aligned} \tag{3.35}$$

where

$$\begin{aligned} a_j(S,\tau,Q,\varphi) = \frac{\sigma_j^2 S^2}{2},\ b_j(S,\tau,Q,\varphi) = S(q_3 q_1 + (1-q_3)(r_l - r_f) - \rho_j), \\ c_j(S,\tau,Q,\varphi) = (q_3 q_1 + q_2(1-q_3) + \lambda_j + \frac{\varphi}{\epsilon}),\ d_j(S,\tau,Q,\varphi) = \varphi\frac{U_j^*}{\epsilon}. \end{aligned}$$

3.2.3.1 Discretization

Following the discretization procedure in Section 3.2.2.1, i.e., a fully implicit positive coefficient discretization and linear interpolation (3.30) for $\mathcal{J}_j U$, we obtain the following for (3.35),

$$\begin{aligned} U_{i,j}^{n+1} &= U_{i,j}^n + \Delta\tau \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\left\{\alpha_{i,j}(Q,\varphi)U_{i-1,j}^{n+1} + \beta_{i,j}(Q,\varphi)U_{i+1,j}^{n+1}\right. \\ &\quad \left. - \ (\alpha_{i,j}(Q,\varphi) + \beta_{i,j}(Q,\varphi) + c_{i,j}(Q,\varphi))\,U_{i,j}^{n+1} + d_{i,j}(Q,\varphi) + \lambda_j[\mathcal{J}_j^h U^{n+1}]_{i,j}\right\}, \quad i < N-1, \\ U_{i,j}^{n+1} &= U_{i,j}^*, \qquad i = N-1. \end{aligned} \tag{3.36}$$

The coefficients $\alpha_{i,j}$ and $\beta_{i,j}$ are defined using a weighted average of central and upstream differencing as described in Algorithm 5 in Appendix B.

Remark 3.2.1. Equations (3.25) and (3.33) are special cases of the more general systems of variational inequalities (VIs) considered by [28], where it is shown that such VIs have unique viscosity solutions. The definition of a viscosity solution must be generalized for systems of PDEs [28]. The discretization schemes (3.32) and (3.36) for American options under regime switching can be shown to be unconditionally l_∞ stable, consistent and monotone and hence converge to the viscosity solution in a straightforward way by using methods in [37].

In the next section, we prove that the discretization schemes (3.7) and (3.12) converge to the viscosity solution.

3.3 Discretization Analysis

For nonlinear second-order PDEs, any monotone, consistent and l_∞ stable discretization scheme converges to the viscosity solution provided that the strong comparison property holds [11, 61]. In Section 3.3.1, we verify these properties for the scalar HJB equation (3.3) resulting from dynamic Bertrand monopoly. For the system of HJB PDEs (3.2), we note that there is no coupling of derivative terms among the individual PDEs and hence the extended definition of the viscosity solution from [61] can be applied here. In Section 3.3.2, we verify that our discretization (3.12) converges to the viscosity solution of the system (3.2).

3.3.1 Monopoly Problem

We provide some definitions and new notations for the purpose of compactness in further analysis. Let

$$\mathbf{x} = (x, \tau)\,, \qquad Du(\mathbf{x}) = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial \tau}\right), \qquad D^2u(\mathbf{x}) = \frac{\partial^2 u}{\partial x^2}.$$

Definition 3.3.1. *The domain Ω^M is partitioned into*

$$\begin{aligned}
\Omega^M_{in} &= (x,\tau) \in (0, x_{max}] \times (0, T], \\
\Omega^M_{x_0} &= (x,\tau) \in \{0\} \times (0, T], \\
\Omega^M_{\tau_0} &= (x,\tau) \in [0, x_{max}] \times \{0\}.
\end{aligned}$$

The HJB equation (3.3) is then rewritten in the compact form as

$$Fu \equiv F\left(\mathbf{x}, u(\mathbf{x}), Du(\mathbf{x}), D^2u(\mathbf{x})\right) = 0, \tag{3.37}$$

where Fu is defined as

$$Fu = \begin{cases} F_{in}u \equiv F_{in}\left(\mathbf{x}, u(\mathbf{x}), Du(\mathbf{x}), D^2u(\mathbf{x})\right) & = u_\tau - \sup_{p\geq 0}\left\{\mathcal{L}^{M,p}u\right\}, & \mathbf{x}\in\Omega^M_{in}, \\ F_{x_0}u \equiv F_{x_0}\left(\mathbf{x}, u(\mathbf{x})\right) & = u - u(x,\tau), & \mathbf{x}\in\Omega^M_{x_0}, \\ F_{\tau_0}u \equiv F_{\tau_0}\left(\mathbf{x}, u(\mathbf{x})\right) & = u - u_0(x), & \mathbf{x}\in\Omega^M_{\tau_0}, \end{cases}$$

where $u_0(x)$ is the initial condition.

We first introduce the notations of upper and lower semi-continuous envelopes, which will later be used in the definition of viscosity solution. For a real-valued function $f : C \mapsto \mathbf{R}$ defined on the closed set C, the upper semi-continuous envelope f^* and lower semi-continuous envelope f_* are defined as

$$\begin{aligned} f^*(y) &= \limsup_{\epsilon\to 0}\left\{f(z) \text{ s.t. } |y-z|<\epsilon, z\in C\right\}, \\ f_*(y) &= \liminf_{\epsilon\to 0}\left\{f(z) \text{ s.t. } |y-z|<\epsilon, z\in C\right\}. \end{aligned}$$

Definition 3.3.2. (Viscosity solution of (3.37)) A locally bounded function $u : \Omega^M \to \mathbf{R}$ is a viscosity sub-solution (resp. super-solution) of (3.37) if for all test functions $\phi(\mathbf{x}) \in C^\infty(\Omega^M)$ and all $\mathbf{x}$, such that $u-\phi$ has a strict global maximum (resp. minimum) with $\phi(\mathbf{x}) = u^*(\mathbf{x})$ (resp. $u_*(\mathbf{x})$), we have

$$\begin{gathered} F_*\left(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x})\right) \leq 0, \\ \left(\text{resp. } F^*\left(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x})\right) \geq 0\right), \end{gathered}$$

where $F_*(.)$ is the lower semi-continuous envelope of F (resp. the upper semi-continuous envelope F^*). The function u is a viscosity solution if it is both a viscosity sub-solution and a super-solution.

Assumption 3.3.1. *(Strong comparison property) If u is an upper semi-continuous sub-solution of (3.37) and if v is a lower semi-continuous super-solution of (3.37), then*

$$u \leq v.$$

The strong comparison property has been proven for first order equations for all kinds of classical equations and boundary conditions. It has also been proven for second order equations with Neumann boundary conditions and classical Dirichlet boundary conditions [7, 10, 22]. Only fully degenerate equations are not well understood. As such, it is clear that the strong comparison property holds for all PDEs in this thesis.

3.3.1.1 Consistency

In this section, we prove that the discretization scheme (3.7) is a consistent approximation to the PDE (3.3) in the viscosity solution sense. Let $\mathcal{G}(.)$ be the discrete approximation to F_{in} for $\mathbf{x} \in \Omega_{in}^M$ and $\mathbf{x}_i^{n+1} = (x_i, \tau^{n+1})$. Then for $\mathbf{x}_i^{n+1} \in \Omega_{in}^M$, we have the following from (3.7),

$$\mathcal{G}\left(h, \mathbf{x}_i^{n+1}, u_i^{n+1}, \left\{u_{i'}^{n+1}\right\}_{i' \neq i}, \{u_{i'}^n\}\right) = \frac{u_i^{n+1} - u_i^n}{\Delta\tau} - \sup_{p \geq 0}\left\{\mathcal{L}_h^{M,p} u_i^{n+1}\right\} = 0. \tag{3.38}$$

We also have

$$\mathcal{G}(.) = 0 = \begin{cases} u(x_i, 0) - u_0(x_i), & x_i^{n+1} \in \Omega_{\tau_0}^M, \\ u(x_i, \tau^{n+1}) - u(x_i, \tau^{n+1}), & x_i^{n+1} \in \Omega_{x_0}^M. \end{cases} \tag{3.39}$$

Definition 3.3.3. (Consistency) For any C^∞ function $\phi(x, \tau)$ in Ω^M with $\phi_i^{n+1} = \phi(\mathbf{x}_i^{n+1}) = \phi(x_i, \tau^{n+1})$, the discretization scheme $\mathcal{G}(.)$ is consistent in the viscosity sense if $\forall \hat{\mathbf{x}} = (\hat{x}, \hat{\tau})$ with $\mathbf{x}_i^{n+1} = (x_i, \tau^{n+1})$, the following holds

$$\limsup_{\substack{h \to 0 \\ \psi \to 0 \\ \mathbf{x}_i^{n+1} \to \hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_i^{n+1}, \phi_i^{n+1} + \psi, \left\{\phi_{i'}^{n+1} + \psi\right\}_{i' \neq i}, \left\{\phi_{i'}^n + \psi\right\}\right) \leq F^*\left(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}})\right),$$

and

$$\liminf_{\substack{h \to 0 \\ \psi \to 0 \\ \mathbf{x}_i^{n+1} \to \hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_i^{n+1}, \phi_i^{n+1} + \psi, \left\{\phi_{i'}^{n+1} + \psi\right\}_{i' \neq i}, \left\{\phi_{i'}^n + \psi\right\}\right) \geq F_*\left(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}})\right).$$

Lemma 3.3.1. *(Local consistency) Suppose the mesh and control discretization parameter h satisfies (3.5), then for any C^∞ function $\phi(x, \tau)$ in Ω^M, with $\phi_i^{n+1} = \phi(x_i, \tau^{n+1}) = \phi(\mathbf{x}_i^{n+1})$, and for h and a constant ψ sufficiently small, we have that*

$$\begin{aligned} &\mathcal{G}\left(h, \mathbf{x}_i^{n+1}, \phi_i^{n+1} + \psi, \left\{\phi_{i'}^{n+1} + \psi\right\}_{i' \neq i}, \left\{\phi_{i'}^n + \psi\right\}\right) \\ &= \begin{cases} F_{in}\phi_i^{n+1} + O(h) + O(\psi), & \mathbf{x}_i^{n+1} \in \Omega_{in}^M, \\ F_{x_0}\phi_i^{n+1} + O(\psi), & \mathbf{x}_i^{n+1} \in \Omega_{x_0}^M, \\ F_{\tau_0}\phi_i^{n+1} + O(\psi), & \mathbf{x}_i^{n+1} \in \Omega_{\tau_0}^M. \end{cases} \end{aligned} \tag{3.40}$$

Proof. Let

$$\begin{aligned}\mathcal{L}^{M,p}\phi_i^{n+1} &\equiv \mathcal{L}^{M,p}\phi\left(x_i,\tau^{n+1}\right),\\ (\phi_\tau)_i^{n+1} &= \phi_\tau\left(x_i,\tau^{n+1}\right).\end{aligned}$$

For $\mathbf{x}_i^{n+1}\in\Omega_{in}^M$, $\mathcal{L}_h^p\phi_i^{n+1}$ given by (3.6) is a locally consistent discretization of the linear operator $\mathcal{L}^{M,p}$, i.e., by the use of Taylor series, we have

$$\mathcal{L}_h^{M,p}\phi_i^{n+1}=\mathcal{L}^{M,p}\phi_i^{n+1}+O(h).$$

Note that

$$\begin{aligned}\mathcal{L}_h^{M,p}\left(\phi_i^{n+1}+\psi\right) &= \mathcal{L}_h^{M,p}\phi_i^{n+1}-r\psi,\\ \frac{\phi_i^{n+1}-\phi_i^n}{\Delta\tau} &= (\phi_\tau)_i^{n+1}+O(h).\end{aligned}$$

From (3.38), we then have

$$\begin{aligned}\mathcal{G}\left(h,\mathbf{x}_i^{n+1},\phi_i^{n+1}+\psi,\left\{\phi_j^{n+1}+\psi\right\}_{j\neq i},\left\{\phi_k^n+\psi\right\}\right) &= \frac{\phi_i^{n+1}-\phi_i^n}{\Delta\tau}-\sup_{p\geq 0}\left\{\mathcal{L}_h^{M,p}\phi_i^{n+1}\right\}+O(\psi)\\ &= (\phi_\tau)_i^{n+1}-\sup_{p\geq 0}\left\{\mathcal{L}^{M,p}\phi_i^{n+1}\right\}+O(h)+O(\psi)\\ &= F_{in}\phi_i^{n+1}+O(h)+O(\psi),\quad \mathbf{x}_i^{n+1}\in\Omega_{in}^M.\end{aligned}$$

The remaining results in (3.40) are proved using similar arguments. □

Lemma 3.3.2. *(Consistency) Provided that all the conditions in Lemma 3.3.1 are satisfied, then the scheme (3.38) - (3.39) is consistent according to Definition 3.3.3.*

Proof. This follows in a straightforward fashion from Lemma 3.3.1 and following the analysis in [58]. □

3.3.1.2 Stability

Definition 3.3.4. *(Stability) Discretization (3.7) is stable if*

$$||u^n||_\infty\leq C_8,$$

for $0\leq n\leq N$ and $h\to 0$, where C_8 is independent of h.

Lemma 3.3.3. *(Stability) Given the positive coefficient conditions* $(\alpha_i(p) \geq 0,\ \beta_i(p) \geq 0,\ i = 1, \ldots, N-1)$ *are satisfied, the discretization* (3.7) *is unconditionally* l_∞ *stable, as the mesh parameter* (3.5) $h \to 0$*, satisfying*

$$||u^n||_\infty \leq ||u^0||_\infty + C_9, \tag{3.41}$$

where $C_9 = T\frac{\kappa^2}{4\eta}$.

Proof. The discrete equations are

$$\begin{aligned} u_i^{n+1} &= u_i^n - \Delta\tau\left(\alpha_i + \beta_i + r\right)u_i^{n+1} + \Delta\tau\alpha_i u_{i-1}^{n+1} + \Delta\tau\beta_i u_{i+1}^{n+1} \\ &\quad + \Delta\tau\frac{p_i^{n+1}}{\eta}(\kappa - p_i^{n+1}). \end{aligned} \tag{3.42}$$

To avoid notational clutter, we suppressed the p dependence of the discrete coefficients in (3.42). It is implied that the coefficients are the limiting values at the optimal p. The following inequality is obtained from (3.42)

$$\begin{aligned} |u_i^{n+1}|\left(1 + \Delta\tau\left(\alpha_i + \beta_i + r\right)\right) &\leq ||u^n||_\infty + \Delta\tau\left(\alpha_i + \beta_i\right)||u^{n+1}||_\infty \\ &\quad + \Delta\tau\left|\frac{p_i^{n+1}}{\eta}(\kappa - p_i^{n+1})\right|. \end{aligned} \tag{3.43}$$

If $||u^{n+1}||_\infty = |u_j^{n+1}|$, then (3.43) gives

$$||u^{n+1}||_\infty\left(1 + r\Delta\tau\right) \leq ||u^n||_\infty + \Delta\tau\left|\frac{p_j^{n+1}}{\eta}(\kappa - p_j^{n+1})\right|.$$

Letting $\hat{d}_{max}^{n+1} = \max_j\left(\frac{p_j^n}{\eta}(\kappa - p_j^n)\right) = \frac{\kappa^2}{4\eta}$, we obtain

$$||u^{n+1}||_\infty \leq ||u^n||_\infty + \Delta\tau\hat{d}_{max}^{n+1},$$

which then results in (3.41). □

3.3.1.3 Monotonicity

Definition 3.3.5. (Monotonicity) The discrete scheme (3.38)-(3.39) is monotone if for all $y_i^n \geq z_i^n$, $\forall i$

$$\mathcal{G}\left(h, \mathbf{x}_i^{n+1}, u_i^{n+1}, \{y_{i'}\}_{i'\neq i}, \{y_{i'}^n\}\right) \leq \mathcal{G}\left(h, \mathbf{x}_i^{n+1}, u_i^{n+1}, \{z_{i'}\}_{i'\neq i}, \{z_{i'}^n\}\right).$$

Lemma 3.3.4. *(Monotonicity) If the scheme (3.38)-(3.39) satisfies the conditions required for Lemma 3.3.3, then the discretization is monotone, according to Definition 3.3.5.*

Proof. The discretization is a positive coefficient scheme $\forall p \geq 0$, hence monotonicity is proved using the same steps from [37]. □

3.3.1.4 Convergence

Theorem 3.3.1. *(Convergence) Assume that the discretization (3.38)-(3.39) satisfies all the conditions required by Lemma 3.3.2, 3.3.3 and 3.3.4 and Assumption 3.3.1, then the discretization converges to the unique viscosity solution of the problem (3.37).*

Proof. Since the scheme is monotone, consistent and l_∞ stable, the convergence follows from the results in [11]. □

3.3.2 Duopoly Problem

In this section, we analyze the convergence properties of the discretization (3.12) of the system of HJB PDEs (3.2). For compactness of analysis, let

$$\mathbf{x} = (x_1, x_2, \tau), \qquad DU_l(\mathbf{x}) = \left(\frac{\partial U_l}{\partial x_1}, \frac{\partial U_l}{\partial x_2}, \frac{\partial U_l}{\partial \tau}\right), \qquad D^2U_l(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 U_l}{\partial x_1^2} & \frac{\partial^2 U_l}{\partial x_1 \partial x_2} \\ \frac{\partial^2 U_l}{\partial x_1 \partial x_2} & \frac{\partial^2 U_l}{\partial x_2^2} \end{pmatrix}.$$

Definition 3.3.6. *The domain Ω^D is partitioned into*

$$\begin{aligned} \Omega^D_{in} &= (x_1, x_2, \tau) \in (0, (x_1)_{max}] \times (0, (x_2)_{max}] \times (0, T], \\ \Omega^D_{(x_1)_0} &= (x_1, x_2, \tau) \in \{0\} \times (0, (x_2)_{max}] \times (0, T], \\ \Omega^D_{(x_2)_0} &= (x_1, x_2, \tau) \in (0, (x_1)_{max}] \times \{0\} \times (0, T], \\ \Omega^D_{\tau_0} &= (x_1, x_2, \tau) \in [0, (x_1)_{max}] \times [0, (x_2)_{max}] \times \{0\}. \end{aligned}$$

The system of HJB equations (3.2) is then written in compact form as

$$F^l U_l = F^l\left(\mathbf{x}, U_l(\mathbf{x}), DU_l(\mathbf{x}), D^2U_l(\mathbf{x}), \{U_m(\mathbf{x})\}_{m \neq l}\right) = 0, \quad l, m = 1, 2, \tag{3.44}$$

where $F^l U_l$ is defined as

$$F^l U_l = \begin{cases} F^l_{in} U_l & \equiv F^l_{in}\left(\mathbf{x}, U_l(\mathbf{x}), DU_l(\mathbf{x}), D^2U_l(\mathbf{x}), \{U_m(\mathbf{x})\}_{m\neq l}\right) \\ & = (U_l)_\tau - \sup_{p_l\geq 0}\left\{\mathcal{L}^{p_l}U_l\right\}, \quad \mathbf{x}\in\Omega^D_{in}, \\ F^l_{(x_1)_0} U_l & \equiv F^l_{(x_1)_0}\left(\mathbf{x}, U_l(\mathbf{x}), DU_l(\mathbf{x}), D^2U_l(\mathbf{x}), \{U_m(\mathbf{x})\}_{m\neq l}\right) \\ & = \begin{cases} U_1 \\ U_2 - \sup_{p_2\geq 0}\left\{\mathcal{L}^{M,p_2}U_2\right\} \end{cases}, \quad \mathbf{x}\in\Omega^D_{(x_1)_0}, \\ F^l_{(x_2)_0} U_l & \equiv F^l_{(x_2)_0}\left(\mathbf{x}, U_l(\mathbf{x}), DU_l(\mathbf{x}), D^2U_l(\mathbf{x}), \{U_m(\mathbf{x})\}_{m\neq l}\right) \\ & = \begin{cases} U_1 - \sup_{p_1\geq 0}\left\{\mathcal{L}^{M,p_1}U_1\right\} \\ U_2 \end{cases}, \quad \mathbf{x}\in\Omega^D_{(x_2)_0}, \\ F^l_{\tau_0} U_l & \equiv F^l_{\tau_0}\left(\mathbf{x}, U_l(\mathbf{x})\right) \\ & = U_l - (U_l)_0(x_1,x_2), \quad \mathbf{x}\in\Omega^D_{\tau_0}, \end{cases}$$

where $(U_l)_0(x_1,x_2)$ is the initial condition.

Definition 3.3.7. *(Viscosity solution of the system of PDEs (3.44)) A $\mathbf{R}^2$-valued function $U=(U_1,U_2)$, where each $U_l:\Omega^D\to\mathbf{R}$ is locally bounded, is called a viscosity subsolution (respectively supersolution) of the system of PDEs (3.44) if and only if for all smooth test functions $\phi_l\in C^\infty(\Omega^D)$, and for all maximum (respectively minimum) points $\mathbf{x}$ of $U^*_l-\phi_l$ (respectively $U_{l*}-\phi_l$), one has*

$$F^l_*\left(\mathbf{x}, U^*_l(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{U^*_m(\mathbf{x})\}_{m\neq l}\right)\leq 0$$
$$\left(\text{respectively } F^{l*}\left(\mathbf{x}, U_{l*}(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{U_{m*}(\mathbf{x})\}_{m\neq l}\right)\geq 0\right).$$

A locally bounded function U is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

There is no coupling of derivative terms among the individual PDEs of the system (3.44), hence the test function for the l-th equation is scalar valued and replaces only the l-th component of the solution U, as in the above definition of the viscosity solution. It is in this sense that we extend the convergence result of [11] to system of PDEs that arise in dynamic Bertrand duopoly. Related work that also generalize the result of [11] to systems of PDEs are [61, 90, 101].

3.3.2.1 Consistency

In this section, we prove that the discretization scheme (3.12) is a consistent approximation to the system of PDEs (3.2) in the viscosity sense. Let $\mathcal{G}^l(.)$ be the discrete approximation

to F_{in}^l for $\mathbf{x} \in \Omega_{in}^D$ and $\mathbf{x}_{i,j}^{n+1} = ((x_1)_i, (x_2)_j, \tau^{n+1})$. Then for $\mathbf{x}_{i,j}^{n+1} \in \Omega_{in}^D$, we rewrite (3.12) as

$$\begin{aligned}
&\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \{(U_l)_{i',j'}^{n+1}\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \{(U_l)_{i',j'}^{n}\}, \{(U_m)_{i',j'}^{n+1}\}_{m\neq l}\right) \\
&\quad = \frac{(U_l)_{i,j}^{n+1} - (U_l)_{i,j}^{n}}{\Delta\tau} - \sup_{p_l \geq 0}\left\{\mathcal{L}_h^{p_l}(U_l)_{i,j}^{n+1}\right\} = 0, \quad l, m = 1, 2.
\end{aligned} \tag{3.45}$$

For $\mathbf{x}_{i,j}^{n+1} \in \Omega_{(x_1)_0}^D$, we have

$$\begin{aligned}
&\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \{(U_l)_{i',j'}^{n+1}\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \{(U_l)_{i',j'}^{n}\}, \{(U_m)_{i',j'}^{n+1}\}_{m\neq l}\right) \\
&\quad = \begin{cases} U_l^{n+1} = 0, & l = 1, \\ \frac{(U_l)_{i,j}^{n+1} - (U_l)_{i,j}^{n}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}_h^{M,p_l}(U_l)_{i,j}^{n+1}\right\} = 0, & l = 2. \end{cases}
\end{aligned} \tag{3.46}$$

Similarly, for $\mathbf{x}_{i,j}^{n+1} \in \Omega_{(x_2)_0}^D$, we have

$$\begin{aligned}
&\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \{(U_l)_{i',j'}^{n+1}\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \{(U_l)_{i',j'}^{n}\}, \{(U_m)_{i',j'}^{n+1}\}_{m\neq l}\right) \\
&\quad = \begin{cases} \frac{(U_l)_{i,j}^{n+1} - (U_l)_{i,j}^{n}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}_h^{M,p_l}(U_l)_{i,j}^{n+1}\right\} = 0, & l = 1, \\ U_l^{n+1} = 0, & l = 2 \end{cases}
\end{aligned} \tag{3.47}$$

Finally, for $\mathbf{x}_{i,j}^{n+1} \in \Omega_{\tau_0}^D$, we have

$$\begin{aligned}
&\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \{(U_l)_{i',j'}^{n+1}\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \{(U_l)_{i',j'}^{n}\}, \{(U_m)_{i',j'}^{n+1}\}_{m\neq l}\right) \\
&\quad = (U_l)((x_1)_i, (x_2)_j, 0) - (U_l)_0((x_1)_i, (x_2)_j) = 0, \quad l, m = 1, 2.
\end{aligned} \tag{3.48}$$

Definition 3.3.8. *Let $\{d_m\}_{m\neq l}$ be a set of real values d_m. We use the notation*

$$\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \left\{(U_l)_{i',j'}^{n+1}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \left\{(U_l)_{i',j'}^{n}\right\}, \{d_m\}_{m\neq l}\right)$$

to mean

$$\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \left\{(U_l)_{i',j'}^{n+1}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \left\{(U_l)_{i',j'}^{n}\right\}, \left\{(U_m)_{i',j'}^{n+1} = d_m\right\}_{m\neq l}\right),$$

which implies that for a fixed m, $(U_m)_{i',j'}^{n+1}$ all have the same value d_m.

Definition 3.3.9. *(Consistency) For any C^∞ function $\phi_l(x_1, x_2, \tau)$ in Ω^D with $(\phi_l)_{i,j}^{n+1} = \phi_l\left(\mathbf{x}_{i,j}^{n+1}\right) = \phi_l\left((x_1)_i, (x_2)_j, \tau^{n+1}\right)$, the discretization scheme $\mathcal{G}^l(.)$ is consistent in the viscosity sense if $\forall \hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{\tau})$ with $\mathbf{x}_{i,j}^{n+1} = \left((x_1)_i, (x_2)_j, \tau^{n+1}\right)$ and $l, m = 1, 2$ and for a small constant ψ, the following holds*

$$\limsup_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (\phi_l)_{i,j}^{n+1} + \psi, \left\{(\phi_l)_{i',j'}^{n+1} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}}, \left\{(\phi_l)_{i',j'}^{n} + \psi\right\}, \{d_m\}_{m\neq l}\right)$$
$$\leq F^{l*}\left(\hat{\mathbf{x}}, \phi_l(\hat{\mathbf{x}}), D\phi_l(\hat{\mathbf{x}}), D^2\phi_l(\hat{\mathbf{x}}), \{d_m\}_{m\neq l}\right),$$

and

$$\liminf_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (\phi_l)_{i,j}^{n+1} + \psi, \left\{(\phi_l)_{i',j'}^{n+1} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}}, \left\{(\phi_l)_{i',j'}^{n} + \psi\right\}, \{d_m\}_{m\neq l}\right)$$
$$\geq F_*^{l}\left(\hat{\mathbf{x}}, \phi_l(\hat{\mathbf{x}}), D\phi_l(\hat{\mathbf{x}}), D^2\phi_l(\hat{\mathbf{x}}), \{d_m\}_{m\neq l}\right).$$

Lemma 3.3.5. *(Local consistency) Suppose the mesh and control discretization parameter h satisfies (3.8), then for any C^∞ function $\phi_l(x_1, x_2, \tau)$ in Ω^D, with $(\phi_l)_{i,j}^{n+1} = \phi_l\left((x_1)_i, (x_2)_j, \tau^{n+1}\right) = \phi_l\left(\mathbf{x}_{i,j}^{n+1}\right)$, and for h and for a sufficiently small constant ψ, we have that*

$$\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (\phi_l)_{i,j}^{n+1} + \psi, \left\{(\phi_l)_{i',j'}^{n+1} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}}, \left\{(\phi_l)_{i',j'}^{n} + \psi\right\}, \{d_m\}_{m\neq l}\right)$$
$$= \begin{cases} F_{in}^l\left(\phi_l\right)_{i,j}^{n+1} + O(h) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{in}^D,\\ F_{(x_1)_0}^l\left(\phi_l\right)_{i,j}^{n+1} + O(h) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{(x_1)_0}^D,\\ F_{(x_1)_0}^l\left(\phi_l\right)_{i,j}^{n+1} + O(h) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{(x_2)_0}^D,\\ F_{\tau_0}^l\left(\phi_l\right)_{i,j}^{n+1} + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{\tau_0}^D. \end{cases} \tag{3.49}$$

Proof. Let

$$\begin{aligned} \mathcal{L}^{p_l}\left(\phi_l\right)_{i,j}^{n+1} &\equiv \mathcal{L}^{p_l}\phi_l\left((x_1)_i, (x_2)_j, \tau^{n+1}\right),\\ \left((\phi_l)_\tau\right)_{i,j}^{n+1} &= (\phi_l)_\tau\left((x_1)_i, (x_2)_j, \tau^{n+1}\right). \end{aligned}$$

For $\mathbf{x}_{i,j}^{n+1} \in \Omega_{in}^D$, $\mathcal{L}_h^{p_l}(\phi_l)_{i,j}^{n+1}$ given by (3.11) is a locally consistent discretization of the linear operator $\mathcal{L}^{p_l}$, i.e., by use of Taylor series, we get

$$\mathcal{L}_h^{p_l}(\phi_l)_{i,j}^{n+1} = \mathcal{L}^{p_l}(\phi_l)_{i,j}^{n+1} + O(h).$$

Further, we have

$$\begin{aligned}\mathcal{L}_h^{p_l}\left((\phi_l)_{i,j}^{n+1}+\psi\right) &= \mathcal{L}_h^{p_l}(\phi_l)_{i,j}^{n+1}-r\psi,\\ \frac{(\phi_l)_{i,j}^{n+1}-(\phi_l)_{i,j}^{n}}{\Delta\tau} &= ((\phi_l)_\tau)_{i,j}^{n+1}+O(h).\end{aligned}$$

From (3.45), we then have

$$\begin{aligned}&\mathcal{G}^l\left(h,\mathbf{x}_{i,j}^{n+1},(\phi_l)_{i,j}^{n+1}+\psi,\{(\phi_l)_{i',j'}^{n+1}+\psi\}_{\substack{i'\neq i\\ or\ j'\neq j}},\{(\phi_l)_{i',j'}^{n}+\psi\},\{d_m\}_{m\neq l}\right)\\ =\ &\frac{(\phi_l)_{i,j}^{n+1}-(\phi_l)_{i,j}^{n}}{\Delta\tau}-\sup_{p_l\geq 0}\left\{\mathcal{L}_h^{p_l}(\phi_l)_{i,j}^{n+1}\right\}+O(\psi)\\ =\ &((\phi_l)_\tau)_{i,j}^{n+1}-\sup_{p_l\geq 0}\left\{\mathcal{L}^{p_l}(\phi_l)_{i,j}^{n+1}\right\}+O(h)+O(\psi)\\ =\ &F_{in}^l(\phi_l)_{i,j}^{n+1}+O(h)+O(\psi),\quad \mathbf{x}_{i,j}^{n+1}\in\Omega_{in}^D.\end{aligned}$$

The remaining results in (3.49) are proved similarly using Taylor series and (3.46)-(3.48). □

Lemma 3.3.6. *(Consistency) Provided all the conditions in Lemma 3.3.5 are satisfied, the scheme (3.45)-(3.48) is consistent according to the Definition 3.3.9.*

Proof. The proof follows in a straightforward fashion from Lemma 3.3.5 and following the analysis in [58] □

3.3.2.2 Stability

Definition 3.3.10. *(Stability) Discretization (3.12) is stable if*

$$||U_l^n||_\infty \leq C_{10},$$

for $0\leq n\leq N$ and $h\to 0$, where C_{10} is independent of h.

Lemma 3.3.7. *(Stability) Given the positive coefficient conditions (3.14)-(3.16) are satisfied, the discretization (3.12) is unconditionally l_∞ stable, as the mesh parameter (3.8) $h\to 0$ satisfying*

$$||U_l^n||_\infty \leq ||U_l^0||_\infty + C_{11}, \tag{3.50}$$

where $C_{11}=T\frac{(a_1+a_3(p_m^*)_{max})^2}{4a_2}$ and $(p_m^*)_{max}=\max_{i,j}(p_m^*)_{i,j}$.

Proof. The discrete equations given by (3.12) are

$$\begin{aligned}(U_l)_{i,j}^{n+1} &= (U_l)_{i,j}^{n} - \Delta\tau\left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r\right)(U_l)_{i,j}^{n+1} \\ &+ \Delta\tau\left((\alpha_l)_{i,j}^{x_1} - \zeta_{i,j}\right)(U_l)_{i-1,j}^{n+1} + \Delta\tau\left((\beta_l)_{i,j}^{x_1} - \zeta_{i,j}\right)(U_l)_{i+1,j}^{n+1} \\ &+ \Delta\tau\left((\alpha_l)_{i,j}^{x_2} - \zeta_{i,j}\right)(U_l)_{i,j-1}^{n+1} + +\Delta\tau\left((\beta_l)_{i,j}^{x_2} - \zeta_{i,j}\right)(U_l)_{i,j+1}^{n+1} \\ &+ \mathbb{1}_{\rho\geq 0}\,\Delta\tau\zeta_{i,j}\left((U_l)_{i+1,j+1}^{n+1} + (U_l)_{i-1,j-1}^{n+1}\right) + \mathbb{1}_{\rho<0}\,\Delta\tau\zeta_{i,j}\left((U_l)_{i+1,j-1}^{n+1} + (U_l)_{i-1,j+1}^{n+1}\right) \\ &+ \Delta\tau(p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{i,j}\right). \end{aligned} \tag{3.51}$$

Similar to the stability analysis for monopoly case, we suppressed the p_l dependence of the discrete coefficients in (3.51) to avoid notational clutter. We emphasize that the coefficients are the limiting values at the optimal p_l. The following inequality is obtained from (3.51)

$$\begin{aligned}&\left|(U_l)_{i,j}^{n+1}\right|\left(1 + \Delta\tau\left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r\right)\right) \\ &\leq ||U_l^n||_\infty + \Delta\tau\left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r\right)||U_l^{n+1}||_\infty \\ &+ \Delta\tau\left|(p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{i,j}\right)\right|. \end{aligned} \tag{3.52}$$

Let $||U_l^{n+1}||_\infty = |(U_l)_{i',j'}^{n+1}|$, then (3.52) becomes

$$||(U_l)_{i,j}^{n+1}||_\infty\,(1 + r\Delta\tau) \leq ||U_l^n||_\infty + \Delta\tau\left|(p_l)_{i',j'}^{n+1}\left(a_1 - a_2(p_l)_{i',j'}^{n+1} + a_3(p_m^*)_{i',j'}\right)\right|.$$

Letting $(p_m^*)_{max} = \max_{i,j}(p_m^*)_{i,j}$ and $\hat{d}_{max}^{n+1} = \max_{i,j}\left((p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{max}\right)\right) = \frac{(a_1 + a_3(p_m^*)_{max})^2}{4a_2}$, we get

$$||U_l^{n+1}||_\infty \leq ||U_l^n||_\infty + \Delta\tau\hat{d}_{max}^{n+1},$$

which then results in (3.50). □

3.3.2.3 Monotonicity

The notion of monotonicity in [11] needs to be extended for systems of PDEs. Quasi-monotone property is an important assumption in the theory of viscosity solution for systems of PDEs [61]. We first show that the system (3.44) satisfies the quasi-monotone property.

Proposition 3.3.1. *Let $w_1, w_2 \in \mathbf{R}^2$ and $l \in \{1, 2\}$. We use the notation $w_1 \geq_l w_2$, which means that $w_1 \geq w_2$ component-wise and $(w_1)_l = (w_2)_l$. The system of PDEs (3.44) is called quasi-monotone [61], if for all $\mathbf{x} \in \Omega^D$ and $\phi_l \in C^\infty$, whenever $w_1 \geq_l w_2$, then*

$$\begin{aligned}&F^l\left(\mathbf{x}, (w_1)_l(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{(w_1)_m(\mathbf{x})\}_{m\neq l}\right) \\ \leq\ &F^l\left(\mathbf{x}, (w_2)_l(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{(w_2)_m(\mathbf{x})\}_{m\neq l}\right). \end{aligned}$$

Proof. This follows from a straightforward calculation by noting that the coupling among the individual PDEs in (3.44) is only due to the control and not the solution. □

We now prove that our discretization (3.45)-(3.48) is monotone. Note that the definition of monotonicity with respect to the last argument of $\mathcal{G}^l(.)$ in (3.53) below is a discrete version of the quasi-monotone property given in Proposition 3.3.1.

Definition 3.3.11. *(Monotonicity) The discretization scheme (3.45)-(3.48) is monotone if for any two $\mathbf{R}^2$-valued discrete functions W_h and U_h defined on Ω^D such that $W_h \geq U_h$ and $(W_l)^{n+1}_{i,j} = (U_l)^{n+1}_{i,j}$,*

$$\begin{aligned} & \mathcal{G}^l\left(h, \mathbf{x}^{n+1}_{i,j}, (W_l)^{n+1}_{i,j}, \left\{(W_l)^{n+1}_{i',j'}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \left\{(W_l)^{n}_{i',j'}\right\}, \left\{(W_m)^{n+1}_{i',j'}\right\}_{m\neq l}\right) \\ \leq\ & \mathcal{G}^l\left(h, \mathbf{x}^{n+1}_{i,j}, (U_l)^{n+1}_{i,j}, \left\{(U_l)^{n+1}_{i',j'}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}}, \left\{(U_l)^{n}_{i',j'}\right\}, \left\{(U_m)^{n+1}_{i',j'}\right\}_{m\neq l}\right). \end{aligned} \tag{3.53}$$

Lemma 3.3.8. *(Monotonicity) If the scheme (3.45)-(3.48) satisfies the conditions required for Lemma 3.3.7, then the discretization is monotone according to Definition 3.3.11.*

Proof. Our discretization (3.45)-(3.48) satisfies positive coefficient conditions $\forall p_l \geq 0$, therefore monotonicity is proved using the same steps from [37]. □

3.3.2.4 Convergence

Theorem 3.3.2. *(Convergence) Assuming that the discretization (3.45)-(3.48) satisfies all the conditions required by Lemma 3.3.6. 3.3.7 and 3.3.8 and that Assumption 3.3.1 holds for (3.44), then the numerical scheme converges to the unique viscosity solution of the system (3.44).*

Proof. Since the scheme is monotone, consistent and l_∞ stable, the convergence follows from the results in [11, 61]. □

3.4 Multigrid Method for HJB and HJBI Systems

We develop a multigrid method based on the FAS, given in Algorithm 1, for solving (3.32), (3.12) and (3.36). The problem is defined on the fine grid Ω_h as

$$A^h(.) = B^h. \tag{3.54}$$

For the HJB system (3.12) due to dynamic Bertrand oligopoly,

$$\begin{aligned} A^h\left(U_l^{n+1,h}\right) &\equiv U_l^{n+1,h} - \Delta\tau \sup_{p_l \geq 0}\left\{\mathcal{A}^{n+1}\left(p_l, p_m^*\right) U_l^{n+1,h} + \mathcal{B}^{n+1}\left(p_l, p_m^*\right)\right\}, \\ B^h &\equiv U_l^{n,h}, \qquad\qquad l, m = 1, 2; \quad m \neq l, \end{aligned} \tag{3.55}$$

where $\mathcal{A}^{n+1}\left(p_l, p_m^*\right) U_l^{n+1,h}$ is the matrix form of the differential operator $\mathcal{L}_h^{p_l}$ dependent on $U_l^{n+1,h}$ and $\mathcal{B}^{n+1}\left(p_l, p_m^*\right)$ is the vector form of $p_l\left(a_1 - a_2 p_l + a_3 (p_m^*)_{i,j}\right)$ in (3.12).

For HJB system (3.32) the operators are defined by

$$A^h\left(U^{n+1,h}\right) \equiv U^{n+1,h} - \Delta\tau \max_{\varphi \in \{0,1\}}\left\{\mathcal{A}^{n+1}(\varphi)\, U^{n+1,h} + \mathcal{B}^{n+1}(\varphi)\right\}, \quad B^h \equiv U^{n,h}, \tag{3.56}$$

where $\mathcal{A}^{n+1}(\varphi)\, U^{n+1,h}$ is the matrix form of the objective function dependent on $U^{n+1,h}$ and $\mathcal{B}^{n+1}(\varphi)$ is the vector form of $d_{i,j}(\varphi)$ in (3.32).

For HJBI system (3.36), we have

$$A^h\left(U^{n+1,h}\right) \equiv U^{n+1,h} - \Delta\tau \sup_{\varphi \in \{0,1\}} \inf_{Q \in \hat{Q}}\left\{\mathcal{A}^{n+1}(Q, \varphi)\, U^{n+1,h} + \mathcal{B}^{n+1}(Q, \varphi)\right\}, \quad B^h \equiv U^{n,h} \tag{3.57}$$

where $\mathcal{A}^{n+1}(Q, \varphi)\, U^{n+1,h}$ is the matrix form of the objective function dependent on $U^{n+1,h}$ and $\mathcal{B}^{n+1}(Q, \varphi)$ is the vector form of $d_{i,j}(\varphi)$ in (3.36).

For HJB and HJBI systems, in addition to the solution and residual, the control should also be carefully considered during restriction and interpolation. Standard FAS techniques using fully weighted restriction and linear interpolation, in general, work well when the control is continuous and bounded. However, when the control is discrete with large jumps, the convergence of the standard FAS deteriorates or it may not converge at all in certain situations [48]. For American options formulated in HJB/HJBI form, there is typically a large jump in control. For efficient convergence of the multigrid method, it is important that the consistency of control between the fine and the coarse grid is preserved during restriction and interpolation. Also, the optimal control at the jump locations must be accurately captured during interpolation. We address these issues and develop efficient multigrid methods for the HJB and HJBI systems. A weighted relaxation scheme, described in section 3.4.1, is used as the smoother. We propose novel interpolation techniques which are presented in section 3.4.2. The restriction operator is chosen such that it preserves consistency of the control, the details are presented in Section 3.4.3.

3.4.1 Weighted Relaxation Smoother

Relaxation type iterative methods are efficient in damping the high frequency error components [48, 100]. We use weighted relaxation scheme as the smoother. Consider the HJB system resulting from American options under regime switching. Rearranging (3.32), we have

$$\begin{aligned}\max_{\varphi\in\{0,1\}} \Big\{ & U_{i,j}^{n} + \Delta\tau\alpha_{i,j}^{n+1}(\varphi)U_{i-1,j}^{n+1} + \Delta\tau\beta_{i,j}^{n+1}(\varphi)U_{i+1,j}^{n+1} + \Delta\tau d_{i,j}^{n+1}(\varphi) + \Delta\tau[\mathcal{J}_j^h U^{n+1}]_{i,j} \\ & - \left(1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)\right) U_{i,j}^{n+1} \Big\} = 0. \end{aligned} \tag{3.58}$$

We note that the coefficient of $U_{i,j}^{n+1}$ in (3.58) is non-negative and hence the equation can be rewritten as

$$\max_{\varphi\in\{0,1\}} \left\{ -U_{i,j}^{n+1} + \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)U_{i-1,j}^{n+1} + \beta_{i,j}^{n+1}(\varphi)U_{i+1,j}^{n+1} + d_{i,j}^{n+1}(\varphi) + [\mathcal{J}_j^h U^{n+1}]_{i,j}\right) + U_{i,j}^{n}}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)} \right\} = 0. \tag{3.59}$$

Let $\bar{U}^k$ be the k^{th} estimate for U^{n+1}. A relaxation scheme can then be derived from (3.59) as

$$\bar{U}_{i,j}^{k+1} = \max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)\bar{U}_{i-1,j}^{k} + \beta_{i,j}^{n+1}(\varphi)\bar{U}_{i+1,j}^{k} + d_{i,j}^{n+1}(\varphi) + [\mathcal{J}_j^h \bar{U}^k]_{i,j}\right) + U_{i,j}^{n}}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)} \right\}. \tag{3.60}$$

This relaxation scheme is not efficient in reducing the high frequency components. Therefore, we introduce a damping factor ω to obtain a weighted relaxation scheme, which is given by

$$\begin{aligned}\bar{U}_{i,j}^{k+1} = & (1-\omega)\bar{U}_{i,j}^{k} \\ & + \omega \max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)\bar{U}_{i-1,j}^{k} + \beta_{i,j}^{n+1}(\varphi)\bar{U}_{i+1,j}^{k} + d_{i,j}^{n+1}(\varphi) + [\mathcal{J}_j^h \bar{U}^k]_{i,j}\right) + U_{i,j}^{n}}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)} \right\}. \end{aligned} \tag{3.61}$$

Following similar derivation, the weighted relaxation scheme for the HJB system (3.12) resulting from dynamic Bertrand duopoly is obtained as

$$
\begin{aligned}
(\bar{U}_l)_{i,j}^{k+1} \;=\; & (1-\omega)\,(\bar{U}_l)_{i,j}^{k} \\
& + \omega \sup_{p_l \geq 0} \left\{ \frac{\Delta\tau \left(\left((\alpha_l)_{i,j}^{x_1} - \zeta_{i,j} \right) (\bar{U}_l)_{i-1,j}^{k} + \left((\beta_l)_{i,j}^{x_1} - \zeta_{i,j} \right) (\bar{U}_l)_{i+1,j}^{k} \right)}{1 + \Delta\tau \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r \right)} \right. \\
& + \frac{\Delta\tau \left(\left((\alpha_l)_{i,j}^{x_2} - \zeta_{i,j} \right) (\bar{U}_l)_{i,j-1}^{k} + \left((\beta_l)_{i,j}^{x_2} - \zeta_{i,j} \right) (\bar{U}_l)_{i,j+1}^{k} \right)}{1 + \Delta\tau \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r \right)} \\
& + \frac{\Delta\tau \left(\mathbb{1}_{\rho \geq 0}\, \zeta_{i,j} \left((\bar{U}_l)_{i+1,j+1}^{k} + (\bar{U}_l)_{i-1,j-1}^{k} \right) + \mathbb{1}_{\rho < 0}\, \zeta_{i,j} \left((\bar{U}_l)_{i+1,j-1}^{k} + (\bar{U}_l)_{i-1,j+1}^{k} \right) \right)}{1 + \Delta\tau \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r \right)} \\
& \left. + \frac{\Delta\tau \left(p_l \left(a_1 - a_2 p_l + a_3 \left(p_m^* \right)_{i,j}^{k} \right) \right) + (U_l)_{i,j}^{n}}{1 + \Delta\tau \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\zeta_{i,j} + r \right)} \right\}, \quad l, m = 1, 2;\ m \neq l. \qquad (3.62)
\end{aligned}
$$

Similarly, for HJBI systems, the weighted relaxation scheme for (3.36) is given by

$$
\begin{aligned}
& \bar{U}_{i,j}^{k+1} = (1-\omega)\bar{U}_{i,j}^{k} \\
& + \omega \sup_{\varphi \in \{0,1\}} \inf_{Q \in \hat{Q}} \left\{ \frac{\Delta\tau \left(\alpha_{i,j}^{n+1}(Q,\varphi)\bar{U}_{i-1,j}^{k} + \beta_{i,j}^{n+1}(Q,\varphi)\bar{U}_{i+1,j}^{k} + d_{i,j}^{n+1}(Q,\varphi) + [\mathcal{J}_j^h \bar{U}^k]_{i,j} \right) + U_{i,j}^{n}}{1 + \Delta\tau \left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi) \right)} \right\} \qquad (3.63)
\end{aligned}
$$

LFA shows that $\omega = 0.67$ results in efficient reduction of the high frequency error components for both HJB and HJBI systems. We present the details of the analysis in Section 3.5.1. We now prove that the weighted relaxation scheme is globally convergent.

Theorem 3.4.1. *(HJBI Systems) Suppose the discretization (3.36) satisfies a positive coefficient condition [37], then the weighted relaxation scheme (3.63) is globally convergent for any initial condition given that* $0 < \omega < 2/(1+\varrho)$. *Furthermore, we have*

$$
||\bar{U}^{k+1} - \bar{U}^{k}||_\infty \leq (|1-\omega| + \omega\varrho)||\bar{U}^{k} - \bar{U}^{k-1}||_\infty
$$

where

$$
\varrho = \max_{i,j} \sup_{\phi \in \{0,1\}} \sup_{Q \in \hat{Q}} \left\{ \frac{\Delta\tau \left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + \lambda_j \right)}{1 + \Delta\tau \left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi) \right)} \right\}.
$$

Proof. Using (3.63), we have $|\bar{U}^{k+1}_{i,j} - \bar{U}^{k}_{i,j}|$ as

$$
\begin{aligned}
&|\bar{U}^{k+1}_{i,j} - \bar{U}^{k}_{i,j}| \\
&\leq |1-\omega||\bar{U}^{k}_{i,j} - \bar{U}^{k-1}_{i,j}| \\
&+ \omega \left| \sup_{\varphi\in\{0,1\}} \inf_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau \left(\alpha^{n+1}_{i,j}(Q,\varphi)\bar{U}^{k}_{i-1,j} + \beta^{n+1}_{i,j}(Q,\varphi)\bar{U}^{k}_{i+1,j} + d^{n+1}_{i,j}(Q,\varphi) + \lambda_j[\mathcal{J}^h_j \bar{U}^k]_{i,j}\right) + U^n_{i,j}}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + c^{n+1}_{i,j}(Q,\varphi)\right)} \right\} \right. \\
&\left. - \sup_{\varphi\in\{0,1\}} \inf_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau \left(\alpha^{n+1}_{i,j}(Q,\varphi)\bar{U}^{k-1}_{i-1,j} + \beta^{n+1}_{i,j}(Q,\varphi)\bar{U}^{k-1}_{i+1,j} + d^{n+1}_{i,j}(Q,\varphi) + \lambda_j[\mathcal{J}^h_j \bar{U}^{k-1}]_{i,j}\right) + U^n_{i,j}}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + c^{n+1}_{i,j}(Q,\varphi)\right)} \right\} \right|
\end{aligned}
$$

Using the properties of sup-inf operators [37] and replacing the regime switching term $[\mathcal{J}^h_j \bar{U}^k]_{i,j}$ by (3.30) and (3.31), we have

$$
\begin{aligned}
&|\bar{U}^{k+1}_{i,j} - \bar{U}^{k}_{i,j}| \\
&\leq |1-\omega||\bar{U}^{k}_{i,j} - \bar{U}^{k-1}_{i,j}| \\
&+ \omega \sup_{\varphi\in\{0,1\}} \sup_{Q\in\hat{Q}} \left\{ \left| \frac{\Delta\tau \left[\alpha^{n+1}_{i,j}(Q,\varphi)\left(\bar{U}^{k}_{i-1,j} - \bar{U}^{k-1}_{i-1,j}\right) + \beta^{n+1}_{i,j}(Q,\varphi)\left(\bar{U}^{k}_{i+1,j} - \bar{U}^{k-1}_{i+1,j}\right)\right]}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + c^{n+1}_{i,j}(Q,\varphi)\right)} \right. \right. \\
&\left. \left. + \frac{\lambda_j\Delta\tau \left(\sum_{m\neq j} \frac{\lambda_{jm}}{\lambda_j} \left[w_m \left(\bar{U}^{k}_{i_m,m} - \bar{U}^{k-1}_{i_m,m}\right) + (1-w_m)\left(\bar{U}^{k}_{i_m+1,m} - \bar{U}^{k-1}_{i_m+1,m}\right)\right]\right)}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + c^{n+1}_{i,j}(Q,\varphi)\right)} \right| \right\} \\
&\leq |1-\omega||\bar{U}^{k}_{i,j} - \bar{U}^{k-1}_{i,j}| \\
&+ \omega \sup_{\varphi\in\{0,1\}} \sup_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + \lambda_j\right)}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(Q,\varphi) + \beta^{n+1}_{i,j}(Q,\varphi) + c^{n+1}_{i,j}(Q,\varphi)\right)} \right\} |\bar{U}^{k}_{i,j} - \bar{U}^{k-1}_{i,j}|.
\end{aligned}
$$

Therefore,

$$
||\bar{U}^{k+1} - \bar{U}^{k}||_\infty \quad \leq \quad (|1-\omega| + \omega\varrho)\, ||\bar{U}^{k} - \bar{U}^{k-1}||_\infty.
$$

Note that $\varrho < 1$ as $\alpha^{n+1}_{i,j}(\varphi), \beta^{n+1}_{i,j}(\varphi), c^{n+1}_{i,j}(\varphi)$ and λ_j are all non-negative. Therefore, the weighted relaxation scheme converges if $0 < \omega < 2/(1+\varrho)$. □

By a similar argument, it can be shown that the relaxation schemes (3.62) and (3.61) are globally convergent. We omit the proof here.

3.4.2 Interpolation

We use the HJB system (3.56) resulting from American options under regime switching as an example to explain the interpolation and restriction operators, but they work well for the Bertrand duopoly (3.55) and the unequal lending/borrowing (3.57) cases as demonstrated in the numerical results.

Consider a two grid method. Given $\{S_i^h\}$, $i = 0, 1, \ldots, N-1$, the grid points with even indices are selected as coarse grid points, i.e., $\{S_i^H\}$, $i = 0, 2, \ldots, N-1$. Given the coarse grid solution $\widetilde{U}_{i,j}^{n,H}$, $i = 0, 2, \ldots, N-1$, we want to interpolate the solution on S_i^h, $i = 0, 1, \ldots, N-1$ for all $j = 1, 2, \ldots, N_m$. Let the interpolated solution be denoted by $\hat{U}^{n,h}$.

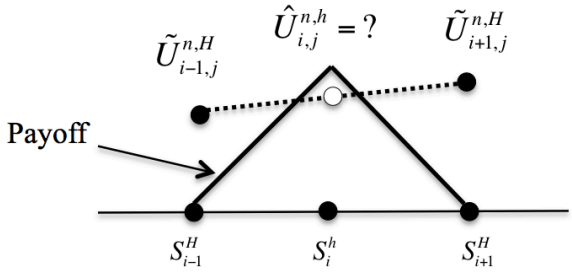


Figure 3.2: Illustration of a scenario where standard linear interpolation fails to capture the accurate optimal control.

For the example problem (3.56), the solution after interpolation must satisfy the American constraint $\hat{U}^{n,h} \geq U^*$. Consider Figure 3.2, where $\widetilde{U}_{i-1,j}^{n,H}$ and $\widetilde{U}_{i+1,j}^{n,H}$ denoted by black circles are solutions from the coarse grid at S_{i-1}^H and S_{i+1}^H, respectively. Suppose the payoff function is a hat function as shown in Figure 3.2. If linear interpolation is used for the noncoarse grid point S_i^h, then the solution $\hat{U}_{i,j}^{n,h}$, denoted by the white circle, lies below the payoff function. It is clear from this example that standard linear interpolation fails to capture the optimal control and hence the correct solution. In the following sections, we present two novel interpolation techniques which address this issue.

3.4.2.1 Direct Interpolation of the Solution

We propose a new interpolation technique which is derived from the relaxation scheme and hence we can ensure that the optimal control is accurately captured. We first copy the

coarse grid solution to the fine grid points which coincide with the coarse grid points, i.e.,

$$\hat{U}^{n,h}_{i,j} = \widetilde{U}^{n,H}_{i,j}, \qquad i = 0, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N_m. \tag{3.64}$$

For the noncoarse grid points, $i = 1, 3, \ldots, N-2$, $j = 1, 2, \ldots, N_m$ we interpolate $\hat{U}^{n,h}_{i,j}$ using one iteration of the relaxation scheme (3.60). We rewrite it here for convenience:

$$\hat{U}^{n,h}_{i,j} = \max_{\varphi \in \{0,1\}} \left\{ \frac{\Delta\tau \left(\alpha^{n+1}_{i,j}(\varphi) \hat{U}^{n,h}_{i-1,j} + \beta^{n+1}_{i,j}(\varphi) \hat{U}^{n,h}_{i+1,j} + d^{n+1}_{i,j}(\varphi) + [\mathcal{J}^h_j \hat{U}^{n,h}]_{i,j} \right) + U^{n,h}_{i,j}}{1 + \Delta\tau \left(\alpha^{n+1}_{i,j}(\varphi) + \beta^{n+1}_{i,j}(\varphi) + c^{n+1}_{i,j}(\varphi) \right)} \right\} \tag{3.65}$$

where $\hat{U}^{n,h}_{i-1,j}, \hat{U}^{n,h}_{i+1,j}$ are given by (3.64). Whenever there is a switch from regime j to regime m, the term $[\mathcal{J}^h_j \hat{U}^{n,h}]_{i,j}$ is approximated with the values $\hat{U}^{n,h}_{i_m,m}$ and $\hat{U}^{n,h}_{i_m+1,m}$ (See (3.31)). If i_m is odd and $m > j$, then $\hat{U}^{n,h}_{i_m,m}$ is an unknown. In such cases, we use standard linear interpolation to approximate $\hat{U}^{n,h}_{i_m,m}$, i.e.,

$$\hat{U}^{n,h}_{i_m,m} = \bar{U}^{n,h}_{i_m,m} + 0.5 \left(E^{n,H}_{i_m-1,m} + E^{n,H}_{i_m+1,m} \right),$$

where $\bar{U}^{n,h}_{i_m,m}$ is the solution after presmoothing and

$$E^{n,H}_{i_m-1,m} = \widetilde{U}^{n,H}_{i_m-1,m} - U^{n,H}_{i_m-1,m}.$$

A similar approximation is used for $\hat{U}^{n,h}_{i_m+1,m}$ when i_m is even and $m > j$. This approximation does not hamper the convergence of our multigrid method. Since we use the relaxation iteration (3.65), it is guaranteed that the optimal control is accurately captured. We theoretically prove that the constraint $\hat{U}^{n,h} \geq U^*$ is satisfied when (3.64) - (3.65) is used and the resulting multigrid method is monotone in Section 3.5.3. Similarly, for the HJB system (3.55) and HJBI system (3.57), we use their respective relaxation iterations for interpolation.

Note that this interpolation is different from the standard interpolation, where the fine grid solution is corrected with the interpolated coarse grid error. Instead, we directly interpolate the solution and the interpolation formula depends on the underlying PDE. We also develop another interpolation technique which is based on the traditional idea of interpolating the coarse grid error. This approach again depends on the underlying PDE as detailed in the following section.

3.4.2.2 Interpolation of the Error

Let the exact solution for time step $n+1$ be $\mathring{U}$ and the approximate solution after interpolation in the k-th iteration be $\hat{U}^k = \mathring{U} + \varepsilon^k$, where ε^k is the error after interpolation in the k-th approximation. Using the relaxation scheme (3.60), we obtain

$$\begin{aligned}\mathring{U}_{i,j}^{k+1} + \varepsilon_{i,j}^{k+1} &= \max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau \left[\alpha_{i,j}^{n+1}(\varphi)(\mathring{U}_{i-1,j}^{k} + \varepsilon_{i-1,j}^{k}) + \beta_{i,j}^{n+1}(\varphi)(\mathring{U}_{i+1,j}^{k} + \varepsilon_{i+1,j}^{k})\right]}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)} \right. \\ &\quad \left. + \frac{\Delta\tau\left[d_{i,j}^{n+1}(\varphi) + \lambda_j[\mathcal{J}_j^h(\mathring{U}^k + \varepsilon^k)]_{i,j}\right] + U_{i,j}^{n}}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi) + \beta_{i,j}^{n+1}(\varphi) + c_{i,j}^{n+1}(\varphi)\right)} \right\}. \end{aligned} \tag{3.66}$$

Let $\varphi_{i,j}^k = \mathring{\varphi}_{i,j}$ for all i,j, where $\mathring{\varphi}_{i,j}$ is the optimal control of the exact solution $\mathring{U}_{i,j}$. Then (3.66) reduces to

$$\varepsilon_{i,j}^{k+1} = \frac{\Delta\tau\left[\alpha_{i,j}^{n+1}(\mathring{\varphi}_{i,j})\varepsilon_{i-1,j}^{k} + \beta_{i,j}^{n+1}(\mathring{\varphi}_{i,j})\varepsilon_{i+1,j}^{k} + \lambda_j[\mathcal{J}_j^h\varepsilon^k]_{i,j}\right]}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\mathring{\varphi}_{i,j}) + \beta_{i,j}^{n+1}(\mathring{\varphi}_{i,j}) + c_{i,j}^{n+1}(\mathring{\varphi}_{i,j})\right)}. \tag{3.67}$$

Given the coarse grid error $E^{n,H} = \widetilde{U}^{n,H} - U^{n,H}$, we first copy the error to the fine grid points that coincide with the coarse grid points, i.e.,

$$E_{i,j}^{n,h} = E_{i,j}^{n,H}, \quad i = 0, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N_m.$$

For interpolation at noncoarse grid points, S_i^h, $i = 1, 3, \ldots, N-2$, $j = 1, 2, \ldots, N_m$, we use (3.67), i.e.,

$$E_{i,j}^{n,h} = \frac{\Delta\tau\left[\alpha_{i,j}^{n+1}(\mathring{\varphi}_{i,j})E_{i-1,j}^{n,h} + \beta_{i,j}^{n+1}(\mathring{\varphi}_{i,j})E_{i+1,j}^{n,h} + \lambda_j[\mathcal{J}_j^h E^{n,h}]_{i,j}\right]}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\mathring{\varphi}_{i,j}) + \beta_{i,j}^{n+1}(\mathring{\varphi}_{i,j}) + c_{i,j}^{n+1}(\mathring{\varphi}_{i,j})\right)}. \tag{3.68}$$

A major challenge in using (3.68) is that the exact solution is unknown and hence its optimal control $\mathring{\varphi}$ is also unknown. We address this issue by approximating $\mathring{\varphi}$ using the relaxation iteration, i.e., for every noncoarse grid point S_i^h, $i = 1, 3, \ldots, N-2$, we set $\mathring{\varphi}_{i,j}$ to be the optimal control which maximizes

$$\max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left[\alpha_{i,j}^{n+1}(\varphi)\widetilde{U}_{i-1,j}^{n,H} + \beta_{i,j}^{n+1}(\varphi)\widetilde{U}_{i+1,j}^{n,H} + d_{i,j}^{n+1}(\varphi) + \lambda_j[\mathcal{J}_j^h\widetilde{U}^{n,H}]_{i,j}\right] + U_{i,j}^{n,h}}{1 + \Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi_{i,j}) + \beta_{i,j}^{n+1}(\varphi_{i,j}) + c_{i,j}^{n+1}(\varphi_{i,j})\right)} \right\} \tag{3.69}$$

The regime switching terms $[\mathcal{J}_j^h \widetilde{U}^{n,H}]_{i,j}$ and $[\mathcal{J}_j^h E^{n,h}]_{i,j}$ are handled as described in Section 3.4.2.1. The fine grid solution is then updated using the standard procedure as

$$\hat{U}_{i,j}^{n,h} = \widetilde{U}_{i,j}^{n,h} + E_{i,j}^{n,h}, \quad i = 0, 1, \ldots, N-1, \quad j = 1, 2, \ldots, N_m, \tag{3.70}$$

Similar derivation applies for interpolation of the error for (3.12) and (3.36) as well. We skip the details here.

3.4.3 Restriction

In the FAS scheme, we restrict the solution and the residual from the fine to the coarse grid. For HJB and HJBI problems, with jumps in the control, the control also needs to be carefully considered while performing restriction. Given the solution and its optimal control on the fine grid, we have to perform restriction such that the optimal control on the coarse grid is consistent with that on the fine grid. Let the fine grid control be given by the first plot of Figure 3.3, with a jump in control between grid points 4 and 5. The coarse grid control may assume one of the three possible values shown in the last three plots of Figure 3.3 based on the choice of the restriction operator. The coarse grid controls in the third and fourth plots of Figure 3.3 are off by one grid point from that of the fine grid. The desired coarse grid control is the one which is consistent with the fine grid control as given in the second plot of Figure 3.3. This can be justified by the FAS coarse grid right hand side B^H, which from Algorithm 1 is given by

$$B^H = \mathcal{R}_r \bar{R}^{n,h} + A^H(\mathcal{R}_u \bar{U}^{n,h}). \tag{3.71}$$

Suppose $A^H(.)$ is a direct discretization on the coarse grid. Then $A^H(\mathcal{R}_u \bar{U}^{n,h})$ depends on the optimal control on the coarse grid, $\varphi^{*,H}$, which is computed from the restricted fine grid solution, $\mathcal{R}_u \bar{U}^{n,h}$. Whereas, the fine grid residual $\bar{R}^{n,h}$ depends on the optimal control on the fine grid, $\varphi^{*,h}$. When $\varphi^{*,h} \neq \varphi^{*,H}$ as shown in the third and fourth plots of Figure 3.3, then the two components of B^H are inconsistent with each other, leading to slow convergence of the multigrid method [48]. This inconsistency can occur due to improper choice of the restriction operator and is most visible near the jumps in control. Hence preserving consistency of control from the fine to the coarse grid is important.

There are two possible ways to achieve the consistency in control. One is to choose the restriction operator of the solution $\mathcal{R}_u$ such that the optimal control remains consistent. The other is to fix the controls on the coarse grid to be consistent with that on the fine grid and solve a local linear problem to update the coarse grid solution accordingly. The latter approach was used by [48]. We choose $\mathcal{R}_u$ to be the injection operator, which in a certain sense is constraint preserving as given in Theorem 3.4.2.

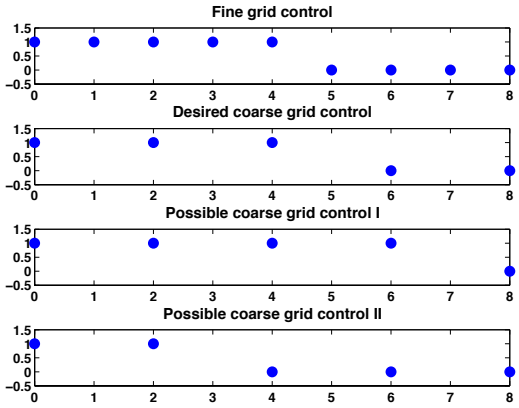


Figure 3.3: The different possibilities for restriction of control.

Theorem 3.4.2. *Let $\varphi^{*,h}$ and $\varphi^{*,H}$ be the optimal control on the fine and coarse grids respectively for problems discretized using penalty method. Let $\mathcal{R}_u$ be an injection operator to restrict the solution U^h, then $\varphi^{*,H} = \varphi^{*,h}$ for all the coarse grid points.*

Proof. The optimal control for the HJB and HJBI systems in the penalty form (3.26) and (3.34) is determined by

$$\max_{\varphi\in\{0,1\}} \left[\frac{\varphi}{\epsilon}(U_i^* - U_{i,j}^n)\right]. \tag{3.72}$$

We note that the optimal control at the grid point (S_i, τ^n) depends only on $U_{i,j}^n$ for any regime j. By choosing $\mathcal{R}_u$ to be an injection operator, we have

$$U_{i,j}^{n,H} = U_{i,j}^{n,h}, \qquad i = 0, 2, \ldots, N-1;\ j = 1, 2, \ldots, N_m. \tag{3.73}$$

Using (3.73), the solution at the coarse grid points is unchanged from that of the fine grid points. Therefore, the optimal control given by (3.72) on the coarse grid is consistent with that of the fine grid, i.e., $\varphi^{*,H} = \varphi^{*,h}$. □

The residual restriction operator $\mathcal{R}_r$ can either be a fully weighted restriction or an injection operator.

3.5 Theoretical Analysis

In this section, we perform LFA to analyse the smoothing property of the weighted relaxation scheme and a two grid Fourier analysis to analyse the convergence behavior of the multigrid method. We also prove that the multigrid method using direct interpolation of the solution is monotone.

3.5.1 Smoothing Analysis

We perform LFA for the weighted relaxation scheme to determine its efficiency as a smoother. LFA is a popular tool for the quantitative analysis of multigrid methods [100]. It is applied on linear discrete operators with constant coefficients, which are obtained by locally linearizing the nonlinear discrete operators with non-constant coefficients. LFA is based on grid functions of the form $\psi(\theta, x) = e^{\mathrm{i}\theta x/h}$, where $\mathrm{i} = \sqrt{-1}$.

We present a detailed smoothing analysis of the weighted relaxation scheme (3.61) for the HJB system resulting from American options under regime switching. We transform (3.26) into the log domain for simplicity of the analysis, which is a common practice in option pricing literature. Using $X = \log S$, we rewrite (3.26) in the log domain as

$$U_{j,\tau} = \max_{\varphi\in\{0,1\}} \left\{ \frac{\sigma_j^2}{2} U_{j,XX} + (r - \rho_j - \frac{\sigma_j^2}{2}) U_{j,X} - (r + \lambda_j + \frac{\varphi}{\epsilon}) U_j + \varphi \frac{U^*}{\epsilon} + \lambda_j \mathcal{J}_j^X U \right\},$$

where

$$\mathcal{J}_j^X U = \sum_{m\neq j} \frac{\lambda_{jm}}{\lambda_j} U_m(\xi_{jm} e^X, \tau).$$

The coefficients for (3.27) in the log domain are given by

$$a_j(\tau,\varphi) = \frac{\sigma_j^2}{2}, \quad b_j(\tau,\varphi) = r - \rho_j - \frac{\sigma_j^2}{2}, \quad c_j(\tau,\varphi) = r + \lambda_j + \frac{\varphi}{\epsilon}, \quad d_j(\tau,\varphi) = \frac{U^*}{\epsilon}. \tag{3.74}$$

We note that the coefficients on the log grid are independent of X, which is a desirable property for LFA.

Let $\mathring{U}$ be the exact solution for time step $n+1$ and let the approximate solution after the k-th smoothing iteration be $\bar{U}^k = \mathring{U} + \varepsilon^k$, where ε^k is the error after the k^{th} iteration. Since LFA is applied on linear operators, we will assume the optimal control for every grid point will not change from iteration to iteration. Let $\mathring{\varphi}_l$ be the optimal control corresponding

to the exact solution $\mathring{U}$. Let $\varphi_l^k = \mathring{\varphi}_l$ be the optimal control for $\bar{U}_l^k$ and let $\mathring{\alpha}_l = \alpha_l^{n+1}(\varphi_l^k)$, $\mathring{\beta}_l = \beta_l^{n+1}(\varphi_l^k)$, $\mathring{c}_l = c_l^{n+1}(\varphi_l^k)$ and $\mathring{\kappa}_l = \mathring{\alpha}_l + \mathring{\beta}_l + \mathring{c}_l$. Then using (3.60), we obtain the error equation as

$$\varepsilon^{k+1} = \mathcal{H}^k \cdot \varepsilon^k,$$

where $\mathcal{H}^k$ is an $NN_m \times NN_m$ matrix. For $l = (j-1)N + i,\ i = 0, 1, \ldots N-1,\ j = 1, 2, \ldots, N_m$ and $l_m = (m-1)N + i_m,\ m = 1, 2, \ldots, N_m,\ m \neq j$, the elements of $\mathcal{H}^k$ are given by

$$\begin{aligned} \mathcal{H}^k_{l,l-1} =& \frac{\Delta\tau\mathring{\alpha}_l}{1+\Delta\tau\mathring{\kappa}_l}, \quad \mathcal{H}^k_{l,l+1} = \frac{\Delta\tau\mathring{\beta}_l}{1+\Delta\tau\mathring{\kappa}_l}, \\ \mathcal{H}^k_{l,l_m} =& \frac{\Delta\tau w_m\lambda_{jm}}{1+\Delta\tau\mathring{\kappa}_l}, \quad \mathcal{H}^k_{l,l_m+1} = \frac{\Delta\tau(1-w_m)\lambda_{jm}}{1+\Delta\tau\mathring{\kappa}_l}, \\ \mathcal{H}^k_{l,l'} =& 0, \quad l' \neq l-1, l+1, l_m, l_m+1. \end{aligned} \tag{3.75}$$

Using (3.75), we obtain the symbol of the smoothing operator for the relaxation scheme (3.60) as

$$\widetilde{\mathcal{H}}_l^k(\theta) = \frac{\Delta\tau\left[\mathring{\alpha}_l e^{-\mathrm{i}\theta} + \mathring{\beta}_l e^{\mathrm{i}\theta} + \sum\limits_{m\neq j}\lambda_{jm}\left(w_m e^{\mathrm{i}(l_m-l)\theta} + (1-w_m)e^{\mathrm{i}(l_m-l+1)\theta}\right)\right]}{1+\Delta\tau\mathring{\kappa}_l}.$$

Since we use the weighted relaxation scheme (3.61) for smoothing, the symbol of its smoothing operator is given by

$$\widetilde{\mathcal{H}}_l^k(\theta,\omega) = (1-\omega) + \omega\frac{\Delta\tau\left[\mathring{\alpha}_l e^{-\mathrm{i}\theta} + \mathring{\beta}_l e^{\mathrm{i}\theta} + \sum\limits_{m\neq j}\lambda_{jm}\left(w_m e^{\mathrm{i}(l_m-l)\theta} + (1-w_m)e^{\mathrm{i}(l_m-l+1)\theta}\right)\right]}{1+\Delta\tau\mathring{\kappa}_l} \tag{3.76}$$

We are interested in the smoothing effect, i.e., the reduction of high frequency error components. Hence, the smoothing factor $\mu(\mathcal{H}_l^k)$ is defined as

$$\mu(\mathcal{H}_l^k) = \sup\left\{|\widetilde{\mathcal{H}}_l^k(\theta,\omega)| : \theta \in [-\pi,\pi) \setminus \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)\right\}.$$

Since generating a useful analytical expression for $\mu(\mathcal{H}_l^k)$ is very complicated, we consider specific frequency values and analyse the behavior of the smoother. In addition,

we demonstrate the efficiency of the smoother for the entire frequency range by plotting $|\widetilde{\mathcal{H}}_l^k(\theta,\omega)|$ for $\theta \in [-\pi,\ \pi)$ and different ω.

Consider a high frequency point $\theta = -\pi$. We note that $l_m - l$ is an integer. We then have $\cos(z\theta) = -1$ if z is odd and $\cos(z\theta) = 1$ if z is even and $\sin(z\theta) = 0$ for all z. Using these values in (3.76), we have

$$\widetilde{\mathcal{H}}_l^k(\theta,\omega) = \frac{(1-\omega)[1+\Delta\tau\mathring{\kappa}_l] + \omega\Delta\tau\left[-(\mathring{\alpha}_l + \mathring{\beta}_l) + \sum\limits_{m\neq j}\lambda_{jm}w_m\right]}{1+\Delta\tau\mathring{\kappa}_l}, \tag{3.77}$$

where

$$w_m^* = \begin{cases} 2w_m - 1 & \text{if } l_m - l \text{ is even,} \\ 1 - 2w_m & \text{if } l_m - l \text{ is odd,} \end{cases}$$

with $|w_m^*| \leq 1$ since $w_m \in [0,1]$. Adding and subtracting $\omega\Delta\tau\mathring{c}_l$ from the numerator of (3.77), we get

$$\begin{aligned}
\widetilde{\mathcal{H}}_l^k(\theta,\omega) &= \frac{(1-\omega)[1+\Delta\tau\mathring{\kappa}_l] + \omega\Delta\tau\left[-\left(\mathring{\alpha}_l + \mathring{\beta}_l + \mathring{c}_l\right) + \mathring{c}_l + \sum\limits_{m\neq j}\lambda_{jm}w_m^*\right]}{1+\Delta\tau\mathring{\kappa}_l} \\
&= \frac{(1-\omega)[1+\Delta\tau\mathring{\kappa}_l] + \omega\Delta\tau\left[-\mathring{\kappa}_l + \mathring{c}_l + \sum\limits_{m\neq j}\lambda_{jm}w_m^*\right]}{1+\Delta\tau\mathring{\kappa}_l} \\
&= \frac{(1-2\omega)[1+\Delta\tau\mathring{\kappa}_l] + \omega + \omega\Delta\tau\left[\mathring{c}_l + \sum\limits_{m\neq j}\lambda_{jm}w_m^*\right]}{1+\Delta\tau\mathring{\kappa}_l} \\
&= (1-2\omega) + \frac{\omega\left[1+\Delta\tau\left(\mathring{c}_l + \sum\limits_{m\neq j}\lambda_{jm}w_m^*\right)\right]}{1+\Delta\tau\mathring{\kappa}_l}
\end{aligned} \tag{3.78}$$

Using (B.10), (B.11) and (3.74) in (3.78), we have

$$\widetilde{\mathcal{H}}_l^k(\theta,\omega) = (1-2\omega) + \frac{\omega\left[1+\Delta\tau\left(r + \lambda_j + \frac{\varphi_l^k}{\epsilon} + \sum\limits_{m\neq j}\lambda_{jm}w_m^*\right)\right]}{1+\Delta\tau\left(\frac{2\sigma_j^2}{h^2} + r + \lambda_j + \frac{\varphi_l^k}{\epsilon}\right)}. \tag{3.79}$$

We consider a three regime model with the following parameters:

$$r = 0.02,\ \Delta\tau = 10^{-3},\ \epsilon = 10^{-4}\Delta\tau,$$
$$\sigma = \begin{bmatrix} 0.2 \\ 0.15 \\ 0.3 \end{bmatrix},\ \lambda = \begin{bmatrix} -3.2 & 0.2 & 3.0 \\ 1.0 & -1.08 & .08 \\ 3.0 & 0.2 & -3.2 \end{bmatrix},\ \xi = \begin{bmatrix} 1.0 & 0.9 & 1.1 \\ 1.2 & 1.0 & 1.3 \\ 0.95 & 0.8 & 1.0 \end{bmatrix}. \tag{3.80}$$

As $h \to 0$ and for the parameters given in (3.80), equation (3.79) reduces to $\widetilde{\mathcal{H}}_l^k(\theta,\omega) = 1 - 2\omega$. For convergence, we should have $|\widetilde{\mathcal{H}}_l^k(\theta,\omega)| < 1$, which is satisfied when $\omega \in (0,1)$. The smoothing factor is minimized when $\omega = 0.5$. Similarly, for low and medium frequency components, $|\widetilde{\mathcal{H}}_l^k(\theta,\omega)| < 1$ when $\omega \in (0,2)$. Optimal smoothing is obtained with the ω which satisfies

$$\min_{\omega} \sup_{\theta} \left\{ |\widetilde{\mathcal{H}}_l^k(\theta,\omega)| : \theta \in [-\pi,\pi) \backslash [-\frac{\pi}{2},\frac{\pi}{2}) \right\}.$$

Since analytically evaluating this expression is very complicated, we numerically determine the optimal ω by plotting $|\widetilde{\mathcal{H}}_l^k(\theta,\omega)|$ against θ for different $\omega \in (0,1]$ and for different grid sizes as $h \to 0$. Figure 3.4(a) shows the smoothing factor when $\varphi = 0$ and $h = 0.0125$ and for different $\omega \in (0,1]$. The relaxation scheme is convergent for the entire frequency range. Furthermore, $\omega = 0.67$ and 0.8 have small smoothing factors for the high frequency range. We now analyse the smoothing property for $\omega = 0.5, 0.67$ and 0.8 as $h \to 0$. We are interested in $\omega = 0.5$ as it minimizes the smoothing factor at $\theta = \pm\pi$ as $h \to 0$. Figures 3.4(b), 3.4(c) and 3.4(d) show the smoothing factors for $\omega = 0.5, 0.67$ and 0.8, respectively. Among the different choices, $\omega = 0.67$ has small smoothing factor, i.e., for all $\theta \in [-\pi,\pi)\backslash[-\frac{\pi}{2},\frac{\pi}{2})$ and different h, $|\widetilde{\mathcal{H}}_l^k(\theta, 0.67)| \leq 1/3$. The smoothing factor has the same upper bound with $\omega = 0.67$ for $\varphi = 1$ as well. Therefore, we use $\omega = 0.67$ as the damping parameter.

Similar results are obtained for the weighted relaxation schemes for the dynamic Bertrand oligopoly and the HJBI systems. Therefore, we use $\omega = 0.67$ for all model problems.

3.5.2 Two-grid Analysis

We apply LFA to the two grid operator to study the convergence properties of the multigrid method. An analysis of the two grid method in general provides sufficient insight into the behavior of multigrid methods. Let ε^k be the error after the k-th two grid iteration, then

$$\varepsilon^{k+1} = \mathcal{V}_h^H \varepsilon^k,$$

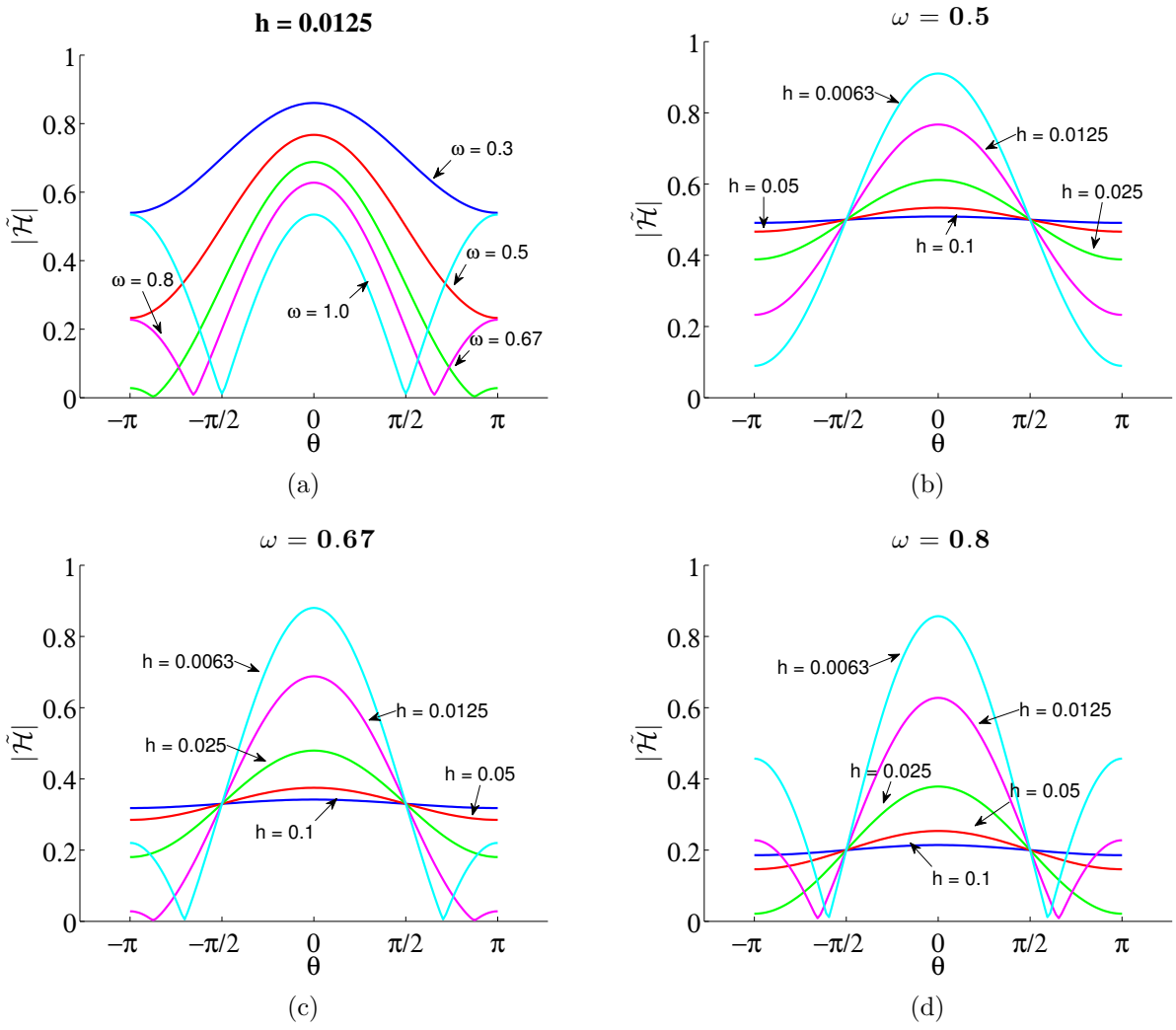


Figure 3.4: Smoothing factor for (3.61) with $\varphi = 0$. (a) $h = 0.0125$ and $\omega \in (0, 1]$, (a) $\omega = 0.5$, (c) $\omega = 0.67$, (d) $\omega = 0.8$.

where $\mathcal{V}_h^H$ is the two grid operator,

$$\mathcal{V}_h^H = \mathcal{H}^{h,\nu_1}\mathcal{K}_h^H\mathcal{H}^{h,\nu_2},$$

where $\mathcal{H}^h$ is the smoothing operator and ν_1 and ν_2 are the number of pre and post smoothing iterations, respectively. $\mathcal{K}_h^H$ is the coarse grid correction operator given by

$$\mathcal{K}_h^H = I^h - \mathcal{P}(L^H)^{-1}\mathcal{R}_u L^h,$$

where I^h is an identity matrix, L^h is the fine grid discrete operator, L^H is the coarse grid operator, $\mathcal{P}$ and $\mathcal{R}_u$ are the interpolation and restriction operators. The spectral radius of $\mathcal{V}_h^H$ gives an indication of the asymptotic rate of multigrid convergence. The convergence factors of $\mathcal{V}_h^H$ are computed by analyzing how the operators $\mathcal{H}^h$, L^h, $\mathcal{R}_u$, L^H and $\mathcal{P}$ act on the Fourier components $\psi(\theta, x) = e^{\mathrm{i}\theta x/h}$. Let $x^h = \{ih, i \in \mathbb{Z}\}$ be the infinite fine grid and $x^H = \{iH, i \in \mathbb{Z}\}$ be the corresponding coarse grid with $H = 2h$.

The Fourier space $\mathcal{F}^h = \text{span}\{e^{\mathrm{i}\theta x/h} : \theta \in (-\pi, \pi]\}$ contains any infinite grid function on x^h [100]. The current approximation $U^{k,h}$ and the error $\varepsilon^{k,h}$ can be represented as linear combinations of the basis functions $e^{\mathrm{i}\theta x/h} \in \mathcal{F}^h$. We note that $\mathcal{F}^h$ can be divided into two dimensional subspaces, also called the harmonics:

$$\mathcal{F}^{\theta,h} = \text{span}\left[\psi(\theta^\xi, x) : \gamma \in \{0, 1\}\right], x \in x^h,$$

where

$$\theta^0 \in (-\pi/2, \pi/2], \quad \theta^1 = \theta^0 - \text{sign}\left(\theta^0\right)\pi.$$

For an arbitrary $\theta \in (-\pi/2, \pi/2]\backslash\{\theta : \hat{L}^H(2\theta^0) = 0\}$, $\mathcal{F}^{\theta,h}$ is invariant under the coarse grid correction operator $\mathcal{K}_h^H$ and the smoother,

$$\mathcal{K}_h^H : \mathcal{F}^{\theta,h} \to \mathcal{F}^{\theta,h}, \quad \mathcal{H}^h : \mathcal{F}^{\theta,h} \to \mathcal{F}^{\theta,h}.$$

Hence $\mathcal{V}_h^H$ is orthogonally equivalent to a 2×2 block matrix given by

$$\hat{\mathcal{V}}_h^H(\theta, \omega) = \hat{\mathcal{H}}^{h,\nu_2}(\theta, \omega)\,\hat{\mathcal{K}}_h^H(\theta)\,\hat{\mathcal{H}}^{h,\nu_1}(\theta, \omega), \tag{3.81}$$

where

$$\hat{\mathcal{K}}_h^H(\theta) = I^h - \hat{\mathcal{P}}(\theta)\left(\hat{L}^H(2\theta)\right)^{-1}\hat{\mathcal{R}}_u(\theta)\,\hat{L}^h(\theta),$$

where $\hat{L}^h$, $\hat{\mathcal{R}}_u$, $\hat{L}^H$, $\hat{\mathcal{P}}$ and $\hat{\mathcal{H}}^h$ are matrices built with the Fourier symbols of their respective multigrid operators. We present these matrices for the case of American options under

regime switching (3.32).

Discrete fine grid operator $\hat{L}^h(\theta)$
$\hat{L}^h(\theta)$ is a 2×2 matrix given by

$$\hat{L}^h(\theta) = \begin{pmatrix} \tilde{L}^h(\theta^0) & \\ & \tilde{L}^h(\theta^1) \end{pmatrix},$$

where

$$\begin{aligned} \tilde{L}^h(\theta^\gamma) &= -\Delta\tau \mathring{\alpha}_l^h e^{-i\theta^\gamma} + \left(1 + \Delta\tau \mathring{\kappa}_l^h\right) - \Delta\tau \mathring{\beta}_l^h e^{i\theta^\gamma} \\ &\quad - \Delta\tau \sum_{m \neq j} \lambda_{jm} \left(w_m e^{\mathrm{i}(l_m - l)\theta^\gamma} + (1 - w_m) e^{\mathrm{i}(l_m - l + 1)\theta^\gamma}\right) \end{aligned} \tag{3.82}$$

Restriction operator $\hat{\mathcal{R}}_u(\theta)$
The restriction operator is denoted by the following 1×2 matrix

$$\hat{\mathcal{R}}_u(\theta) = \begin{bmatrix} \widetilde{\mathcal{R}}_u(\theta^0) & \widetilde{\mathcal{R}}_u(\theta^1) \end{bmatrix}. \tag{3.83}$$

For injection the Fourier symbols $\widetilde{\mathcal{R}}_u(\theta^\gamma)$ for all frequencies are 1.

Interpolation operator $\hat{\mathcal{P}}$
The interpolation matrix is in general given by

$$\hat{\mathcal{P}}(\theta) = \begin{pmatrix} \widetilde{\mathcal{P}}(\theta^0) \\ \widetilde{\mathcal{P}}(\theta^1) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \delta \\ 1 - \delta \end{pmatrix}$$

where for (3.32), we have

$$\delta = \frac{\Delta\tau \left[\mathring{\alpha}_l^h e^{-i\theta^0} + \mathring{\beta}_l^h e^{i\theta^0} + \sum_{m \neq j} \left(w_m e^{\mathrm{i}(l_m - l)\theta^0} + (1 - w_m) e^{\mathrm{i}(l_m - l + 1)\theta^0} \right) \right]}{1 + \Delta\tau \mathring{\kappa}_l^h}. \tag{3.84}$$

Smoothing operator $\hat{\mathcal{H}}^h$
The smoothing operator is a 2×2 matrix,

$$\hat{\mathcal{H}}^h(\theta, \omega) = \begin{pmatrix} \widetilde{\mathcal{H}}^h(\theta^0, \omega) & \\ & \widetilde{\mathcal{H}}^h(\theta^1, \omega) \end{pmatrix},$$

where $\widetilde{\mathcal{H}}^h$ is given by (3.76).

Coarse grid discrete operator $\hat{L}^H(2\theta)$
$\hat{L}^H(2\theta)$ is a 1×1 matrix whose symbol is given by

$$\begin{aligned}\tilde{L}^H(2\theta) &= -\Delta\tau\mathring{\alpha}_l^H e^{-i2\theta} + \left(1 + \Delta\tau\mathring{\kappa}_l^H\right) - \Delta\tau\mathring{\beta}_l^H e^{i2\theta} \\ &\quad - \Delta\tau \sum_{m\neq j} \lambda_{jm}\left(w_m e^{\mathrm{i}(l_m-l)2\theta} + (1-w_m)e^{\mathrm{i}(l_m-l+1)2\theta}\right). \end{aligned} \tag{3.85}$$

We then construct $\hat{\mathcal{V}}_h^H(\theta)$ using (3.76), (3.82), (3.83), (3.84) and (3.85). We can now determine the spectral radius of $\mathcal{V}_h^H$ by calculating the spectral radius of $\hat{calV}_h^H(\theta)$:

$$\rho\left(\mathcal{V}_h^H\right) = \max_{\theta\in(-\pi/2,\pi/2]} \rho\left(\hat{\mathcal{V}}_h^H(\theta)\right).$$

We recall that Fourier analysis is exact only for linear operators with constant (or frozen) coefficients [100]. Therefore, we fix the parameters as given in (3.80) and $\varphi = 0$. Figure 3.5(a) shows the plot of $\rho(\hat{\mathcal{V}}_h^H(\theta))$ against θ for different h for the HJB system (3.32). As $h \to 0$, $\rho\left(\mathcal{V}_h^H\right) \to 0.12$, which is a very satisfactory convergence rate. Figure 3.5(b) shows the convergence rates for HJBI systems. Similar results are observed for $\varphi = 1$, the details are omitted here.

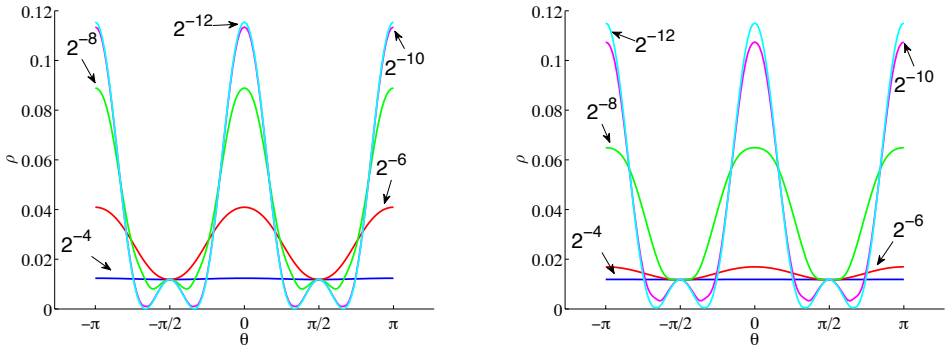


Figure 3.5: Convergence rate of the two-grid algorithm for different grid sizes and $\varphi = 0$. (a) HJB system and (b) HJBI system.

3.5.3 Monotonicity

Monotonicity properties often result in smooth and fast convergence of the multigrid solution [5]. In this section, we present detailed analysis of the monotonicity property for the HJB system resulting from American options under regime switching (3.25). We note that (3.25) can also be formally stated in a linear complementarity form:

$$\begin{aligned} U_{j,\tau} - \mathcal{L}_j U_j - \lambda_j \mathcal{J}_j U \geq 0, \\ U_j - U^* \geq 0, \\ \left(U_{j,\tau} - \mathcal{L}_j U_j - \lambda_j \mathcal{J}_j U\right)\left(U_j - U^*\right) = 0. \end{aligned} \tag{3.86}$$

We use the linear complementarity formulation to define the monotonicity property.

Definition 3.5.1. *[53] A multigrid method for the linear complimentary problem (3.86) is monotone if, as $\epsilon \to 0$, the interpolated fine grid solution $\hat{U}^{n,h}$ satisfies the constraint*

$$\hat{U}^{n,h} - U^{*,h} \geq 0. \tag{3.87}$$

Theorem 3.5.1. *The multigrid method using direct interpolation for the solution for the HJB system (3.32) is monotone as $\epsilon \to 0$.*

Proof. Given $U^{n,h}$, let $\hat{U}^{n,h}$ be the solution after interpolation, which is given by (3.64) and (3.65). The updated coarse grid solution $\widetilde{U}_{i,j}^{n,H}$ is obtained by solving the following coarse grid problem:

$$\begin{aligned} \widetilde{U}_{i,j}^{n,H} \quad - \quad \Delta\tau \Bigg[\alpha_{i,j}^{n,H} \widetilde{U}_{i-1,j}^{n,H} + \beta_{i,j}^{n,H} \widetilde{U}_{i+1,j}^{n,H} - \left(\alpha_{i,j}^{n,H} + \beta_{i,j}^{n,H} + c_{i,j}^{n,H}\right) + \lambda_j \left[\mathcal{J}_j^H \widetilde{U}^{n,H}\right]_{i,j} + \\ \max_{\varphi \in \{0,1\}} \left(\frac{\varphi_{i,j}^{n,H}}{\epsilon} \left(U_{i,j}^{*,H} - \widetilde{U}_{i,j}^{n,H} \right) \right) \Bigg] = B_{i,j}^{n,H}, \qquad i = 0, 2, \ldots, N-3, \end{aligned} \tag{3.88}$$

where $B_{i,j}^{n,H}$ is the coarse grid right hand side (See Algorithm 1) and $U_{i,j}^{*,H}$ is the payoff function on the coarse grid, which is given by

$$U_{i,j}^{*,H} = U_{i,j}^{*,h}, \qquad i = 0, 2, \ldots, N-3. \tag{3.89}$$

Suppose we use relaxation scheme to solve the coarse grid problem. Let $(\widetilde{U}^{n,H})^0 = U^{n,H}$,

then the relaxation iteration for (3.88) is given by

$$\left(\widetilde{U}_{i,j}^{n,H}\right)^{k+1} = \max_{\varphi\in\{0,1\}}\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,H}\left(\widetilde{U}_{i-1,j}^{n,H}\right)^{k}+\beta_{i,j}^{n,H}\left(\widetilde{U}_{i+1,j}^{n,H}\right)^{k}+d_{i,j}^{n,H}+\left[\mathcal{J}_j^H\left(\widetilde{U}^{n,H}\right)^{k}\right]_{i,j}\right)+B_{i,j}^{n,H}}{1+\Delta\tau\left(\alpha_{i,j}^{n,H}+\beta_{i,j}^{n,H}+c_{i,j}^{n,H}\right)}\right], \quad (3.90)$$

$$i=0,2,\ldots,N-3;\quad j=1,2,\ldots,N_m;\quad k=0,1,\ldots,\text{until convergence.}$$

Using $c_{i,j}^{n,H}$ and $d_{i,j}^{n,H}$ from (3.28) in (3.90), we obtain

$$\left(\widetilde{U}_{i,j}^{n,H}\right)^{k+1} = \max\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,H}\left(\widetilde{U}_{i-1,j}^{n,H}\right)^{k}+\beta_{i,j}^{n,H}\left(\widetilde{U}_{i+1,j}^{n,H}\right)^{k}+\left[\mathcal{J}_j^H\left(\widetilde{U}^{n,H}\right)^{k}\right]_{i,j}\right)+B_{i,j}^{n,H}}{1+\Delta\tau\left(\alpha_i^{n,H}+\beta_i^{n,H}+r+\lambda_j\right)},\right.$$

$$\left.\frac{\Delta\tau\left(\alpha_{i,j}^{n,H}\left(\widetilde{U}_{i-1,j}^{n,H}\right)^{k}+\beta_{i,j}^{n,H}\left(\widetilde{U}_{i+1,j}^{n,H}\right)^{k}+\frac{1}{\epsilon}U_{i,j}^{*,H}+\left[\mathcal{J}_j^H\left(\widetilde{U}^{n,H}\right)^{k}\right]_{i,j}\right)+B_{i,j}^{n,H}}{1+\Delta\tau\left(\alpha_{i,j}^{n,H}+\beta_{i,j}^{n,H}+r+\lambda_j+\frac{1}{\epsilon}\right)}\right].$$

As $\epsilon\to 0$, we have

$$\left(\widetilde{U}_{i,j}^{n,H}\right)^{k+1} = \max\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,H}\left(\widetilde{U}_{i-1,j}^{n,H}\right)^{k}+\beta_{i,j}^{n,H}\left(\widetilde{U}_{i+1,j}^{n,H}\right)^{k}+\left[\mathcal{J}_j^H\left(\widetilde{U}^{n,H}\right)^{k}\right]_{i,j}\right)+B_{i,j}^{n,H}}{1+\Delta\tau\left(\alpha_i^{n,H}+\beta_i^{n,H}+r+\lambda_j\right)},\ U_{i,j}^{*,H}\right]. \quad (3.91)$$

From (3.91), it is clear that

$$\left(\widetilde{U}_{i,j}^{n,H}\right)^{k+1} \geq U_{i,j}^{*,H}, \qquad i=0,2,\ldots,N-3 \quad (3.92)$$

for any k. From (3.64), (3.89) and (3.92), we have

$$\hat{U}_{i,j}^{n,h} \geq U_{i,j}^{*,h}, \qquad i=0,2,\ldots,N-3. \quad (3.93)$$

For the noncoarse grid points, the interpolated fine grid solution is given by (3.65), which as $\epsilon\to 0$ becomes:

$$\hat{U}_{i,j}^{n,h} = \max\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,h}\hat{U}_{i-1,j}^{n,h}+\beta_{i,j}^{n,h}\hat{U}_{i+1,j}^{n,h}+\left[\mathcal{J}_j^h\hat{U}^{n,h}\right]_{i,j}\right)+U_{i,j}^{n,h}}{1+\Delta\tau\left(\alpha_{i,j}^{n,h}+\beta_{i,j}^{n,h}+r+\lambda_j\right)}, U_{i,j}^{*,h}\right], \quad i=1,3,\ldots,N-2,$$

which results in

$$\hat{U}_{i,j}^{n,h} \geq U_{i,j}^{*,h}, \qquad i = 1, 3, \ldots, N-2. \tag{3.94}$$

From (3.93), (3.94) and Definition 3.5.3, the two grid method using direct interpolation for the solution is monotone. Using induction, we can prove that a L-grid method ($L \geq 2$) is monotone. □

3.6 Numerical Results

We test our multigrid method with two pre and post-smoothing steps on the model problems presented in Section 3.2. We present the results using direct interpolation of the solution (3.65). The convergence using interpolation of the error (3.70) is very similar and hence we omit the details here.

Example 3.6.1. *2D HJB System: Dynamic Bertrand Duopoly* (3.12).
We use the parameters $T = 0.25$, $r = 1$, $\sigma_1 = \sigma_2 = 0.6$, $\rho = 0.1, \gamma = 0.1$, $\kappa = 6$, $\eta = 1$, $\Delta\tau = 0.025$ and a convergence tolerance of 10^{-6}. The two dimensional grid is coarsened using the multiple coarsening strategy [5]. We use multiple grids such that the coarsest grid has 17×17 grid points. Since the convergence is similar in each time step, we only show the convergence results for the very first time step in Table 3.1. The relaxation scheme alone takes 452 iterations for the grid size of 1025×1025, whereas our multigrid method converges in only 9 iterations.

Table 3.1: Multigrid convergence for the HJB system (3.12).

$N_1 \times N_2$	65×65	129×129	257×257	513×513	1025×1025
$\gamma = 0.1$	3	3	4	5	9
$\gamma = 0.5$	3	3	4	6	9

Example 3.6.2. *HJB System: American Options under Regime Switching* (3.32).
We consider a three regime model for evaluating the multigrid method. The transition probabilities λ, jump amplitudes ξ and the volatilities are given in (3.95). The other parameters are given in Table 3.2. We consider American options with three different payoffs: put, straddle and butterfly. Numerical tests are performed on a uniform log grid. Multiple grids are used for different grid sizes such that the coarsest grid had only 9

grid points. The convergence results for the very first timestep for different grid sizes are given in Table 3.3. The multigrid method converges in a very small number of iterations irrespective of the grid size.

$$\lambda = \begin{bmatrix} -3.2 & 0.2 & 3.0 \\ 1.0 & -1.08 & .08 \\ 3.0 & 0.2 & -3.2 \end{bmatrix}; \ \xi = \begin{bmatrix} 1.0 & 0.9 & 1.1 \\ 1.2 & 1.0 & 1.3 \\ 0.95 & 0.8 & 1.0 \end{bmatrix}; \ \sigma = \begin{bmatrix} .2 \\ .15 \\ .3 \end{bmatrix}. \tag{3.95}$$

Table 3.2: Parameters used for American options.

Expiry Time, T	0.5
Exercise	American
Strike, K	100
Butterfly Parameters, K_1, K_2	$90, 110$
Risk free interest rate, r	.02
Time step, $\Delta\tau$	10^{-3}
Penalty Parameter, ϵ	$10^{-4}\Delta\tau$
Convergence Tolerance	10^{-6}

Table 3.3: Multigrid convergence for the HJB system (3.32).

N	65	129	257	513	1025	2049
Put	2	3	3	3	3	4
Straddle	3	3	3	3	3	4
Butterfly	3	3	3	3	3	5

Example 3.6.3. *HJBI System: American Option and Stock Borrowing Fees* (3.36).
We use the parameters given in Table 3.2 and (3.95) for the HJBI systems under a three regime model. The borrowing rate $r_b = 0.05$, lending rate $r_l = 0.03$ and the stock borrowing fee $r_f = 0.004$. The convergence results for different initial conditions are given in Table 3.4. Similar to the case of HJB systems, the multigrid method for HJBI system also converges in a very small number of iterations independent of the grid size.

Table 3.4: Multigrid convergence for the HJBI system (3.36).

N	65	129	257	513	1025	2049
Put	2	3	3	3	3	4
Straddle	3	3	3	3	3	4
Butterfly	3	3	3	3	4	5

3.7 Conclusion

We constructed fully implicit, consistent, unconditionally l_∞ stable and monotone discretization schemes that converges to the viscosity solution for the HJB PDE resulting from dynamic Bertrand monopoly problem and the two-dimensional systems of nonlinear HJB PDEs from duopoly problem. We developed multigrid methods for discrete systems of nonlinear HJB and HJBI PDEs associated with dynamic Bertrand duopoly and regime switching applications. Weighted relaxation scheme is used as the smoother. We show that the smoother is convergent for both HJB and HJBI systems, in contrast to the policy iteration which may not converge for HJBI problems. A smoothing analysis shows that the weighted relaxation scheme with $\omega = 0.67$ effectively damps the high frequency error components. We choose injection for restriction, which preserves the consistency of the control from the fine to the coarse grids. We introduce new interpolation techniques which efficiently capture the optimal control in the presence of jumps. We analyze the convergence behavior of the multigrid method through a two grid Fourier analysis, which gives a convergence factor as low as 0.12 as $h \to 0$. Numerical tests on practical examples show that the multigrid method converges in very small number of iterations irrespective of the grid size.

Chapter 4

Numerical Methods for Two-factor Stochastic Volatility Models

4.1 Introduction

The European option pricing under a two-factor stochastic volatility model is formulated as a three-dimensional partial differential equation. We develop an unconditionally monotone finite difference discretization scheme for the three-dimensional PDE. A sufficient condition that guarantees convergence to the viscosity solution is that the discretization scheme be l_∞ stable, monotone and consistent in the viscosity sense [11]. For higher dimensional PDEs with cross derivative terms, a standard finite difference scheme on a fixed stencil, in general, cannot ensure monotonicity due to the cross derivative terms [34]. However, a monotone scheme can be constructed on a fixed stencil by enforcing a constraint on the grid spacing for certain two-dimensional PDEs [26]. It is in general difficult to enforce such a constraint on the grid spacing in the presence of three cross derivative terms.

Explicit monotone wide stencil schemes were developed by [17, 29]. These schemes have the disadvantage of time step restrictions due to stability considerations. A fully implicit wide stencil method based on a local coordinate rotation to eliminate the cross derivative terms that results in a monotone discretization was proposed by [83]. The stencil size here is $O(\sqrt{h})$, hence the name wide stencil. We propose to perform a similarity transformation using the eigenvector matrix of the diffusion matrix, which results in the diagonalization of the diffusion matrix. Noting that similarity transformation is essentially a coordinate transformation, we now have a diffusion process in a new coordinate system with no cross

derivative terms. This approach is equivalent to the local coordinate rotation of [83]. In our approach, however, we do not have to explicitly determine the angle of rotation.

An issue with the wide stencil method is that for grid points near the boundaries, the discretization may include points outside the computational domain. This issue was addressed in [83] by shrinking the stencil of those points lying outside the domain to $O(h)$. When only one point lies outside the domain, this results in a discretization which is $O(\sqrt{h})$ consistent with the diffusion term. On the other hand, when both the sides of the three point stencil are shrunk to $O(h)$, this method is inconsistent . Fortunately in [83], the coefficients of the diffusion term are $O(h)$ in this region, thus preserving consistency. However, this is not the case for the three-dimensional PDE. We propose a novel way to address the issue of points in the discretization lying outside the computational domain.

We use a hybrid stencil which is a combination of the fixed stencil and wide stencil [26, 83]. For any smooth test function, the fixed stencil is second order accurate. But the fixed stencil cannot ensure monotonicity at every grid point, in general. On the other hand, wide stencil discretization guarantees monotonicity but is only first order accurate and is computationally expensive. We construct an algorithm which uses fixed stencil as much as possible to take advantage of its computational efficiency and accuracy, while still ensuring monotonicity of the numerical scheme. This is analogous to the "central differencing as much as possible" scheme in one dimension [105].

We develop multigrid methods to solve the linear system resulting from the discretization of the three-dimensional PDE. Multigrid methods converge in a small constant number of iterations independent of the grid size. Therefore, these methods are computationally efficient compared to the relaxation-type and Krylov subspace iterative methods. Efficient multigrid methods for higher dimensional PDEs have been developed in the literature [91, 100, 108]. Various block smoothers are tested for equations of Black-Scholes type in several dimensions in [91]. Bi-CGSTAB solver with multigrid preconditioner was developed for high-dimensional parabolic PDEs in [108]. However, multigrid methods are very specific to the underlying problem and these methods have not been investigated for two-factor stochastic volatility models in the past.

In Section 4.2, we present the two-factor stochastic volatility model and the formulation of the PDE to price European options. In Section 4.3.1, we present the non-monotone fixed stencil discretization for the three-dimensional PDE. In Section 4.3.2, the monotone wide stencil discretization and the novel way to handle the boundary regions are presented. The hybrid stencil is then presented in 4.3.3. The discretization analysis to prove that the numerical scheme using a hybrid stencil converges to the viscosity solution is presented in Section 4.4. Multigrid methods are presented in Section 4.5. Finally the numerical results

are presented in Section 4.6.

4.2 Formulation

Let $U(S,\nu,\xi,\tau)$ be the value of the European option under a two-factor stochastic volatility model. The asset price S, variance ν and the variance factor ξ follow the stochastic process (under the risk neutral measure),

$$\begin{aligned} dS &= rS\,dt + \sqrt{\nu}S\,dZ_1, \\ d\nu &= C_1\left(\xi-\nu\right)\,dt + C_2\nu^{\alpha}\,dZ_2, \\ d\xi &= C_3\left(C_4-\xi\right)\,dt + C_5\xi^{\beta}\,dZ_3, \end{aligned} \tag{4.1}$$

where r is the risk-free interest rate, $\tau = T-t$, where T is the expiry time, $C_{k,k=1\ldots5},\alpha,\beta$ are constants and dZ_i are increments of correlated Brownian motions. The correlations between the Brownian motions are denoted by ρ_{ij}:

$$d\langle Z_i, Z_j\rangle_t = \rho_{ij}\,dt, \quad i,j = 1,2,3.$$

The stochastic differential equations (SDEs) (4.1) can be interpreted as a model with short term variance ν, that reverts to a long term level C_4, at possibly a slower mean reversion speed $C_3 < C_1$. Gatheral [42] terms the choice $\alpha=\beta=0.5$ as the double Heston, $\alpha=\beta=1$ as the double lognormal and the general case as the double CEV. Gatheral [42] finds that the double lognormal model fits better than the double Heston and the double CEV with $\alpha=\beta=0.94$ fits with even better parameter stability to model SPX and VIX options.

Following standard arguments for SDEs of the kind (4.1), the pricing function for European options satisfies the Feynman-Kac partial differential equation (PDE) (See Sections 1.9.3 and 2.5.4, [39]) given by

$$\begin{aligned} \frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu S^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{12}C_2\nu^{\alpha}\sqrt{\nu}S\frac{\partial^2 U}{\partial S\partial\nu} + \rho_{13}C_5\xi^{\beta}\sqrt{\nu}S\frac{\partial^2 U}{\partial S\partial\xi} \\ &\quad + \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta}\frac{\partial^2 U}{\partial\nu\partial\xi} + rS\frac{\partial U}{\partial S} + C_1\left(\xi-\nu\right)\frac{\partial U}{\partial\nu} + C_3\left(C_4-\xi\right)\frac{\partial U}{\partial\xi} - rU. \end{aligned} \tag{4.2}$$

with the terminal condition $U(S,\nu,\xi,0) = U^*(S)$, where $U^*(S)$ is the option payoff. Equation (4.2) is defined over the domain $(S,\nu,\xi,\tau)\in[0,\infty]\times[0,\infty]\times[0,\infty]\times[0,T]$.

For computational simplicity, (4.2) is transformed to the log domain along the S-dimension. Using $x = \log S$, (4.2) is rewritten in the log domain $(x, \nu, \xi, \tau) \in [-\infty, \infty] \times [0, \infty] \times [0, \infty] \times [0, T]$ as

$$U_\tau = \mathcal{L}U, \tag{4.3}$$

where the linear differential operator $\mathcal{L}$ is given by

$$\begin{aligned}\mathcal{L}U \;=\; & \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{12}C_2\nu^{\alpha}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \nu} + \rho_{13}C_5\xi^{\beta}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \xi} \\ & + \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta}\frac{\partial^2 U}{\partial \nu \partial \xi} + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} + C_1\left(\xi - \nu\right)\frac{\partial U}{\partial \nu} + C_3\left(C_4 - \xi\right)\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.4}$$

As $x \to \infty$, one can use financial reasoning to determine the asymptotic form of the solution [26], which is denoted by

$$U(x, \nu, \xi, \tau) \approx d_o(\tau) + d_1(\tau)\, \mathrm{e}^x. \tag{4.5}$$

For computational purposes, we assume the discretization is posed on the truncated domain

$$\Omega = (x, \nu, \xi, \tau) \in [x_{\min}, x_{\max}] \times [0, \nu_{\max}] \times [0, \xi_{\max}] \times [0, T]\,. \tag{4.6}$$

The error due to this approximation is small when the upper limits are sufficiently large and the lower limit $x_{\min}$ is sufficiently small [26].

4.2.1 Boundary conditions

Figure 4.1 illustrates the different boundary regions of the three dimensional domain. There are six boundary planes, eight boundary lines and four points for which boundary conditions must be specified. The boundary planes include only the inner region of the plane. The $\Omega_{x_{\max}}$ plane, highlighted in blue in Figure 4.1, is the only exception as it includes the entire plane. The four corner points of $\Omega_{x_{\min}}$: Ω_{cp1}, Ω_{cp2}, Ω_{cp3} and Ω_{cp4} are denoted by red dots in Figure 4.1. The definitions of these boundary regions are given in the Table 4.1. We now present the boundary conditions for each of these regions.

For $\Omega_{x_{\max}}$, we follow the standard approach [26, 83] and use a Dirichlet boundary condition with $U\left(x_{\max}, \nu, \xi, \tau\right) = U^*\left(x_{\max}\right)$. In order to specify a boundary condition for $\Omega_{x_{\min}}$, we first consider the PDE (4.2) in S domain. On the lower boundary $S = 0$. Hence

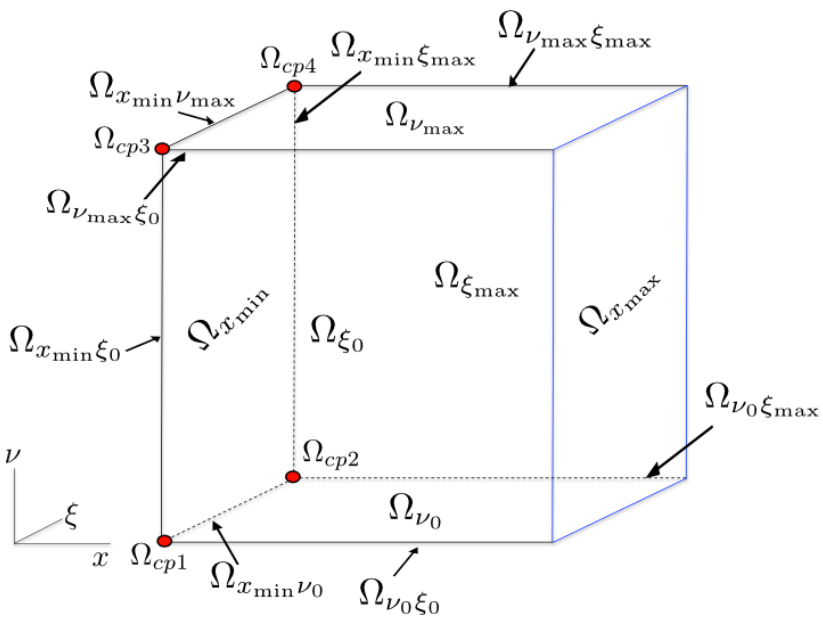


Figure 4.1: Boundaries Regions of the three dimensional domain.

$\Omega_{x_{\min}}$	$\{x_{\min}\} \times (0, \nu_{\max}) \times (0, \xi_{\max})$	$\Omega_{x_{\max}}$	$\{x_{\max}\} \times [0, \nu_{\max}] \times [0, \xi_{\max}]$
Ω_{ν_0}	$(x_{\min}, x_{\max}) \times \{0, \} \times (0, \xi_{\max})$	$\Omega_{\nu_{\max}}$	$(x_{\min}, x_{\max}) \times \{\nu_{\max}\} \times (0, \xi_{\max})$
Ω_{ξ_0}	$(x_{\min}, x_{\max}) \times (0, \nu_{\max}) \times \{0\}$	$\Omega_{\xi_{\max}}$	$(x_{\min}, x_{\max}) \times (0, \nu_{\max}) \times \{\xi_{\max}\}$
$\Omega_{\nu_0\xi_0}$	$(x_{\min}, x_{\max}) \times \{0, \} \times \{0\}$	$\Omega_{x_{\min}\nu_0}$	$\{x_{\min}\} \times \{0, \} \times (0, \xi_{\max})$
$\Omega_{x_{\min}\xi_0}$	$\{x_{\min}\} \times (0, \nu_{\max}) \times \{0\}$	$\Omega_{\nu_0\xi_{\max}}$	$(x_{\min}, x_{\max}) \times \{0, \} \times \{\xi_{\max}\}$
$\Omega_{x_{\min}\xi_{\max}}$	$\{x_{\min}\} \times (0, \nu_{\max}) \times \{\xi_{\max}\}$	$\Omega_{x_{\min}\nu_{\max}}$	$\{x_{\min}\} \times \{\nu_{\max}, \} \times (0, \xi_{\max})$
$\Omega_{\nu_{\max}\xi_{\max}}$	$(x_{\min}, x_{\max}) \times \{\nu_{\max}, \} \times \{\xi_{\max}\}$	$\Omega_{\nu_{\max}\xi_0}$	$(x_{\min}, x_{\max}) \times \{\nu_{\max}, \} \times \{0\}$
Ω_{cp1}	$\{x_{\min}\} \times \{0\} \times \{0\}$	Ω_{cp2}	$\{x_{\min}\} \times \{0\} \times \{\xi_{\max}\}$
Ω_{cp3}	$\{x_{\min}\} \times \{\nu_{\max}\} \times \{0\}$	Ω_{cp4}	$\{x_{\min}\} \times \{\nu_{\max}\} \times \{\xi_{\max}\}$

Table 4.1: Domain definitions for the boundary regions.

(4.2) reduces to

$$\begin{aligned}\frac{\partial U}{\partial \tau} &= \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta}\frac{\partial^2 U}{\partial \nu \partial \xi} \\ &\quad + C_1(\xi-\nu)\frac{\partial U}{\partial \nu} + C_3(C_4-\xi)\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.7}$$

Note that $S \to 0$ as $x_{\min} \to -\infty$. For practical purposes, we choose a sufficiently small $x_{\min}$ and update the solution on $\Omega_{x_{\min}}$ by solving the approximation (4.7).

In Ω_{ν_0}, $\nu = 0$, hence (4.3) reduces to

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + r\frac{\partial U}{\partial x} + C_1\xi\frac{\partial U}{\partial \nu} + C_3(C_4-\xi)\frac{\partial U}{\partial \xi} - rU. \tag{4.8}$$

The characteristics in the ν dimension are outgoing in (4.8) and hence we simply solve (4.8) to update the solution in the Ω_{ν_0}.

Similarly for the region Ω_{ξ_0} in Figure 4.1, using $\xi = 0$ in (4.3) results in

$$\begin{aligned}\frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \rho_{12}C_2\nu^{\alpha}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \nu} \\ &\quad + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} - C_1\nu\frac{\partial U}{\partial \nu} + C_3C_4\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.9}$$

Here again, the characteristics in the ξ direction are outgoing and hence (4.9) can be solved with no additional boundary conditions. Using similar arguments, the following regions do not require additional boundary conditions and the solution is updated by solving the corresponding PDE.

$$\Omega_{x_{\min}\nu_0}: \quad \frac{\partial U}{\partial \tau} = \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + C_1\xi\frac{\partial U}{\partial \nu} + C_3(C_4-\xi)\frac{\partial U}{\partial \xi} - rU. \tag{4.10}$$

$$\Omega_{x_{\min}\xi_0}: \quad \frac{\partial U}{\partial \tau} = \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} - C_1\nu\frac{\partial U}{\partial \nu} + C_3C_4\frac{\partial U}{\partial \xi} - rU. \tag{4.11}$$

$$\Omega_{\nu_0\xi_0}: \quad \frac{\partial U}{\partial \tau} = r\frac{\partial U}{\partial x} + C_3C_4\frac{\partial U}{\partial \xi} - rU. \tag{4.12}$$

$$\Omega_{cp1}: \quad \frac{\partial U}{\partial \tau} = C_3C_4\frac{\partial U}{\partial \xi} - rU. \tag{4.13}$$

As $\nu \to \infty$, we have $\partial U/\partial \nu = 0$ [26]. Therefore, by choosing a sufficiently large $\nu_{\max}$, we

use the condition $\partial U/\partial \nu = 0$ for the region $\Omega_{\nu_{\max}}$ in Figure 4.1. Hence, (4.3) reduces to

$$\begin{aligned}\frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{13}C_5\xi^{\beta}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \xi} \\ &\quad + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} + C_3\left(C_4 - \xi\right)\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.14}$$

The following regions also use the condition $\partial U/\partial \nu = 0$:

$$\Omega_{x_{\min}\nu_{\max}}: \quad \frac{\partial U}{\partial \tau} = \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + C_3\left(C_4 - \xi\right)\frac{\partial U}{\partial \xi} - rU. \tag{4.15}$$

$$\Omega_{\nu_{\max}\xi_0}: \quad \frac{\partial U}{\partial \tau} = \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} + C_3C_4\frac{\partial U}{\partial \xi} - rU. \tag{4.16}$$

$$\Omega_{cp3}: \quad \frac{\partial U}{\partial \tau} = C_3C_4\frac{\partial U}{\partial \xi} - rU. \tag{4.17}$$

As $\xi \to \infty$, from the mean reverting property in (4.1), $\nu \to \infty$ and hence if follows that $\partial U/\partial \nu = 0$ and $\partial U/\partial \xi = 0$. Therefore in the regions $\Omega_{\xi_{\max}}$ and $\Omega_{\nu_{\max}\xi_{\max}}$, (4.3) reduces to

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} - rU. \tag{4.18}$$

The following regions also use the additional conditions $\partial U/\partial \nu = 0$ and $\partial U/\partial \xi = 0$:

$$\Omega_{x_{\min}\xi_{\max}},\ \Omega_{cp2} \text{ and } \Omega_{cp4}: \quad \frac{\partial U}{\partial \tau} = -rU. \tag{4.19}$$

$$\Omega_{\nu_0\xi_{\max}}: \quad \frac{\partial U}{\partial \tau} = r\frac{\partial U}{\partial x} - rU. \tag{4.20}$$

We next present methods to discretize (4.4), (4.7)-(4.20) in the following section.

4.3 Discretization

Wide stencil methods based on a local coordinate rotation to eliminate the cross derivative terms have been developed for two-dimensional HJB PDE [83]. In the absence of cross derivative terms, a standard finite differencing will result in a monotone scheme. In this thesis, we extend the wide stencil method to three-dimensional PDE (4.3). We make use of a hybrid stencil which is a combination of a fixed stencil and wide stencil, in which a fixed stencil is used as much as possible for accuracy and computational efficiency.

The pricing PDE (4.3) is linear and therefore a non-monotone finite difference scheme is adequate. Nevertheless, we develop a monotone discretization scheme as it serves as a framework for three-dimensional nonlinear PDEs, such as pricing American options under two-factor stochastic volatility model.

The spatial domain (x, ν, ξ) is discretized into a set of nodes $\{x_0, x_1, \dots x_{n_x-1}\} \times \{\nu_0, \nu_1, \dots \nu_{n_\nu-1}\} \times \{\xi_0, \xi_1, \dots, \xi_{n_\xi-1}\}$ with a total number of grid points $N = n_x \times n_\nu \times n_\xi$. Let the n-th time step be denotes by $\tau^n = n\Delta\tau,\ n = 0, \dots, N_\tau$ with $N_\tau = T/\Delta\tau$. Let $U^n_{i,j,k}$ be the approximate solution of (4.3) at $(x_i, \nu_j, \xi_k, \tau^n)$. For convenience, we sometimes use an alternative notation U^n_l, $l = k \cdot n_x \cdot n_\nu + j \cdot n_x + i$, for the approximation $U^n_{i,j,k}$. We further define

$$\begin{aligned} \Delta^+ x_i &= x_{i+1} - x_i, \quad \Delta^- x_i = x_i - x_{i-1}, \\ \Delta^+ \nu_j &= \nu_{j+1} - \nu_j, \quad \Delta^- \nu_j = \nu_j - \nu_{j-1}, \\ \Delta^+ \xi_k &= \xi_{k+1} - \xi_k, \quad \Delta^- \xi_k = \xi_k - \xi_{k-1}, \end{aligned}$$

and

$$\begin{aligned} (\Delta x)_{\max} &= \max_i (x_{i+1} - x_i), \quad (\Delta x)_{\min} = \min_i (x_{i+1} - x_i), \\ (\Delta \nu)_{\max} &= \max_i (\nu_{i+1} - \nu_i), \quad (\Delta \nu)_{\min} = \min_i (\nu_{i+1} - \nu_i), \\ (\Delta \xi)_{\max} &= \max_i (\xi_{i+1} - \xi_i), \quad (\Delta \xi)_{\min} = \min_i (\xi_{i+1} - \xi_i). \end{aligned}$$

We assume a mesh discretization parameter h such that

$$\begin{aligned} (\Delta x)_{\max} = \mathcal{B}_1 h, \quad (\Delta \nu)_{\max} = \mathcal{B}_2 h, \quad (\Delta \xi)_{\max} = \mathcal{B}_3 h, \\ (\Delta x)_{\min} = \mathcal{B}'_1 h, \quad (\Delta \nu)_{\min} = \mathcal{B}'_2 h, \quad (\Delta \xi)_{\min} = \mathcal{B}'_3 h, \quad \Delta\tau = \mathcal{B}_4 h, \end{aligned} \tag{4.21}$$

where $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}'_1, \mathcal{B}'_2, \mathcal{B}'_3$ and $\mathcal{B}_4$ are constants independent of h. We present the fixed stencil discretization in the following section and the wide stencil method in Section 4.3.2.

4.3.1 The fixed point stencil

The cross derivative terms are discretized using a seven point stencil [26]. For example, consider the cross derivative term $\rho_{12} C_2 \nu^\alpha \sqrt{\nu} \frac{\partial^2 U}{\partial x \partial \nu}$. For $\rho_{12} > 0$, the stencil in Figure 3.1(a) is used with the finite difference formula given by

$$\frac{\partial^2 U}{\partial x \partial \nu} \approx \frac{2U^n_{i,j,k} + U^n_{i+1,j+1,k} + U^n_{i-1,j-1,k} - U^n_{i+1,j,k} - U^n_{i-1,j,k} - U^n_{i,j+1,k} - U^n_{i,j-1,k}}{\Delta^+ x_i \Delta^+ \nu_j + \Delta^- x_i \Delta^- \nu_j}.$$

For $\rho_{12} < 0$, the stencil in Figure 3.1(b) is used with the corresponding formula given by

$$\frac{\partial^2 U}{\partial x \partial \nu} \approx \frac{-2U^n_{i,j,k} - U^n_{i+1,j-1,k} - U^n_{i-1,j+1,k} + U^n_{i+1,j,k} + U^n_{i-1,j,k} + U^n_{i,j+1,k} + U^n_{i,j-1,k}}{\Delta^+ x_i \Delta^- \nu_j + \Delta^- x_i \Delta^+ \nu_j}.$$

Standard three point central differencing is used for the terms $\frac{\partial^2 U}{\partial x^2}$, $\frac{\partial^2 U}{\partial \nu^2}$ and $\frac{\partial^2 U}{\partial \xi^2}$. The first order derivatives are discretized with central differencing as much as possible and forward or backward differencing when central differencing does not satisfy the positive coefficient condition (4.24)-(4.28). The discrete form $\mathcal{L}_f$ of the differential operator $\mathcal{L}$ (4.4) is then given by

$$\begin{aligned}
\mathcal{L}_f U^{n+1}_{i,j,k} \;=\; & \left(\alpha^x_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{x\xi}_{i,j,k}\right) U^{n+1}_{i-1,j,k} + \left(\beta^x_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{x\xi}_{i,j,k}\right) U^{n+1}_{i+1,j,k} \\
& + \left(\alpha^\nu_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k}\right) U^{n+1}_{i,j-1,k} + \left(\beta^\nu_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k}\right) U^{n+1}_{i,j+1,k} \\
& + \left(\alpha^\xi_{i,j,k} - \zeta^{x\xi}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k}\right) U^{n+1}_{i,j,k-1} + \left(\beta^\xi_{i,j,k} - \zeta^{x\xi}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k}\right) U^{n+1}_{i,j,k+1} \\
& + \mathbb{1}_{\rho_{12}>0}\, \zeta^{x\nu}_{i,j,k} \left(U^{n+1}_{i+1,j+1,k} + U^{n+1}_{i-1,j-1,k}\right) + \mathbb{1}_{\rho_{12}<0}\, \zeta^{x\nu}_{i,j,k} \left(U^{n+1}_{i+1,j-1,k} + U^{n+1}_{i-1,j+1,k}\right) \\
& + \mathbb{1}_{\rho_{13}>0}\, \zeta^{x\xi}_{i,j,k} \left(U^{n+1}_{i+1,j,k+1} + U^{n+1}_{i-1,j,k-1}\right) + \mathbb{1}_{\rho_{13}<0}\, \zeta^{x\xi}_{i,j,k} \left(U^{n+1}_{i+1,j,k-1} + U^{n+1}_{i-1,j,k+1}\right) \\
& + \mathbb{1}_{\rho_{23}>0}\, \zeta^{\nu\xi}_{i,j,k} \left(U^{n+1}_{i,j+1,k+1} + U^{n+1}_{i,j-1,k-1}\right) + \mathbb{1}_{\rho_{23}<0}\, \zeta^{\nu\xi}_{i,j,k} \left(U^{n+1}_{i,j+1,k-1} + U^{n+1}_{i,j-1,k+1}\right) \\
& - \left(\alpha^x_{i,j,k} + \beta^x_{i,j,k} + \alpha^\nu_{i,j,k} + \beta^\nu_{i,j,k} + \alpha^\xi_{i,j,k} + \beta^\xi_{i,j,k} - 2\zeta^{x\nu}_{i,j,k} - 2\zeta^{x\xi}_{i,j,k} - 2\zeta^{\nu\xi}_{i,j,k} + r\right) U^{n+1}_{i,j,k}, \qquad (4.22)
\end{aligned}$$

where the coefficients $\alpha^x_{i,j,k}, \beta^x_{i,j,k}, \alpha^\nu_{i,j,k}, \beta^\nu_{i,j,k}, \alpha^\xi_{i,j,k}, \beta^\xi_{i,j,k}, \zeta^{x\nu}_{i,j,k}, \zeta^{x\xi}_{i,j,k}$ and $\zeta^{\nu\xi}_{i,j,k}$ are given in Algorithm 6 in Appendix C. Using a fully implicit time stepping and (4.22), the discrete form of (4.3) is given by

$$\frac{U^{n+1}_{i,j,k} - U^n_{i,j,k}}{\Delta\tau} = \mathcal{L}_f U^{n+1}_{i,j,k}. \qquad (4.23)$$

The PDEs (4.7) - (4.20) along the boundaries are discretized in the same way as the PDE in the inner region. For example, consider the $\Omega_{x_{\min}}$ region in Figure 4.1. The PDE in this region is given by (4.7), which is a two-dimensional PDE in ν and ξ dimensions and independent of derivatives in the x dimension. Hence, the discretization of this PDE results in (4.22) with $\alpha^x_{i,j,k}$, $\beta^x_{i,j,k}$, $\zeta^{x\nu}_{i,j,k}$ and $\zeta^{x\xi}_{i,j,k}$ equal to zero. The discretization of the remaining PDEs (4.8) - (4.20) will also result in (4.22) with appropriate coefficients set to zero. To be concise, we list, in Table 4.2, the non-zero coefficients of (4.22) when the PDEs along the boundaries (4.7) - (4.20) are discretized.

Region	Non-zero Coefficients
$\Omega_{x_{\min}}$	$\alpha^{\nu}_{i,j,k}$, $\beta^{\nu}_{i,j,k}$, $\alpha^{\xi}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$, $\zeta^{\nu\xi}_{i,j,k}$
Ω_{ν_0}	$\beta^{x}_{i,j,k}$, $\beta^{\nu}_{i,j,k}$, $\alpha^{\xi}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
Ω_{ξ_0}	$\alpha^{x}_{i,j,k}$, $\beta^{x}_{i,j,k}$, $\alpha^{\nu}_{i,j,k}$, $\beta^{\nu}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$, $\zeta^{x\nu}_{i,j,k}$
$\Omega_{x_{\min}\nu_0}$	$\beta^{\nu}_{i,j,k}$, $\alpha^{\xi}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
$\Omega_{x_{\min}\xi_0}$	$\alpha^{\nu}_{i,j,k}$, $\beta^{\nu}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
$\Omega_{\nu_0\xi_0}$	$\beta^{x}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
Ω_{cp1}	$\beta^{\xi}_{i,j,k}$
$\Omega_{\nu_{\max}}$	$\alpha^{x}_{i,j,k}$, $\beta^{x}_{i,j,k}$, $\alpha^{\xi}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$, $\zeta^{x\xi}_{i,j,k}$
$\Omega_{x_{\min}\nu_{\max}}$	$\alpha^{\xi}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
$\Omega_{\nu_{\max}\xi_0}$	$\alpha^{x}_{i,j,k}$, $\beta^{x}_{i,j,k}$, $\beta^{\xi}_{i,j,k}$
Ω_{cp3}	$\beta^{\xi}_{i,j,k}$
$\Omega_{\xi_{\max}}$, $\Omega_{\nu_{\max}\xi_{\max}}$	$\alpha^{x}_{i,j,k}$, $\beta^{x}_{i,j,k}$
$\Omega_{x_{\min}\xi_{\max}}$, Ω_{cp2}, Ω_{cp4}	-
$\Omega_{\nu_0\xi_{\max}}$	$\beta^{x}_{i,j,k}$

Table 4.2: Coefficients of the discretized PDEs along the boundaries.

The positive coefficient condition [37] is given by

$$\alpha^x_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{x\xi}_{i,j,k} \geq 0, \quad \beta^x_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{x\xi}_{i,j,k} \geq 0, \tag{4.24}$$

$$\alpha^\nu_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0, \quad \beta^\nu_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0, \tag{4.25}$$

$$\alpha^\xi_{i,j,k} - \zeta^{x\xi}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0, \quad \beta^\xi_{i,j,k} - \zeta^{x\xi}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0, \tag{4.26}$$

$$\alpha^x_{i,j,k} + \beta^x_{i,j,k} + \alpha^\nu_{i,j,k} + \beta^\nu_{i,j,k} + \alpha^\xi_{i,j,k} + \beta^\xi_{i,j,k} - 2\zeta^{x\nu}_{i,j,k} - 2\zeta^{x\xi}_{i,j,k} - 2\zeta^{\nu\xi}_{i,j,k} + r \geq 0, \tag{4.27}$$

$$\zeta^{x\nu}_{i,j,k} \geq 0, \quad \zeta^{x\xi}_{i,j,k} \geq 0, \quad \zeta^{\nu\xi}_{i,j,k} \geq 0. \tag{4.28}$$

It is essential for the discretization scheme to satisfy the positive coefficient condition (4.24)-(4.28) as it guarantees monotonicity. The seven point operator for the cross derivative terms ensures (4.28). But, the rest of the conditions (4.24)-(4.27) cannot be guaranteed due to the presence of the cross derivative terms. To address this issue, the idea is to use a wide stencil discretization to eliminate cross derivatives by a local coordinate transformation. As a result, the positive coefficient conditions can be easily guaranteed. The details of the wide stencil method are given in the following section.

4.3.2 The Wide Stencil

We perform a local coordinate transformation for the diffusion terms such that the cross derivative terms vanish in the new coordinate system (y_1, y_2, y_3). Consider the diffusion matrix of (4.4), which is given by

$$D = \frac{1}{2}\begin{pmatrix} \nu & \rho_{12}C_2\nu^\alpha\sqrt{\nu} & \rho_{13}C_5\xi^\beta\sqrt{\nu} \\ \rho_{12}C_2\nu^\alpha\sqrt{\nu} & C_2^2\nu^{2\alpha} & \rho_{23}C_2C_5\nu^\alpha\xi^\beta \\ \rho_{13}C_5\xi^\beta\sqrt{\nu} & \rho_{23}C_2C_5\nu^\alpha\xi^\beta & C_5^2\xi^{2\beta} \end{pmatrix}$$

Note that D is a symmetric positive semi-definite matrix. Let R be a matrix whose columns are comprised of the linearly independent eigenvectors of D. The diffusion matrix D can then be regarded as a diagonal matrix Λ that has been re-expressed in coordinates of the basis R. The elements of the diagonal matrix Λ are the eigenvalues of D. Thus, we have

$$\Lambda = R^T D R = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Using this similarity transformation, the second order terms in (4.4) are locally transformed to

$$\lambda_1\frac{\partial^2\mathcal{U}}{\partial y_1^2} + \lambda_2\frac{\partial^2\mathcal{U}}{\partial y_2^2} + \lambda_3\frac{\partial^2\mathcal{U}}{\partial y_3^2}, \tag{4.29}$$

where $\mathcal{U}$ is the value function $\mathcal{U}(y_1, y_2, y_3, \tau)$ in the local transformed coordinate system. Note that (4.29) is free of cross derivative terms and hence a standard three point central differencing will now yield a positive coefficient discretization.

The local transformed coordinate system (y_1, y_2, y_3) is a virtual grid which overlays the original grid. Therefore, the values of $\mathcal{U}$ on the virtual local grid are approximated by using an interpolant $\mathcal{J}_h U$ on the original grid. We choose a linear interpolation operator for $\mathcal{J}_h$ as it preserves monotonicity. In order to preserve consistency, the discretization is performed on the virtual grid with grid points that are at a distance of $O(\sqrt{h})$ from the central node [83], where h (4.21) is the mesh discretization parameter. This results in a wide stencil method as the length of the relative stencil increases as $h \to 0$, more precisely, $\sqrt{h}/h \to \infty$ as $h \to 0$. The wide stencil is illustrated in Figure 4.2.

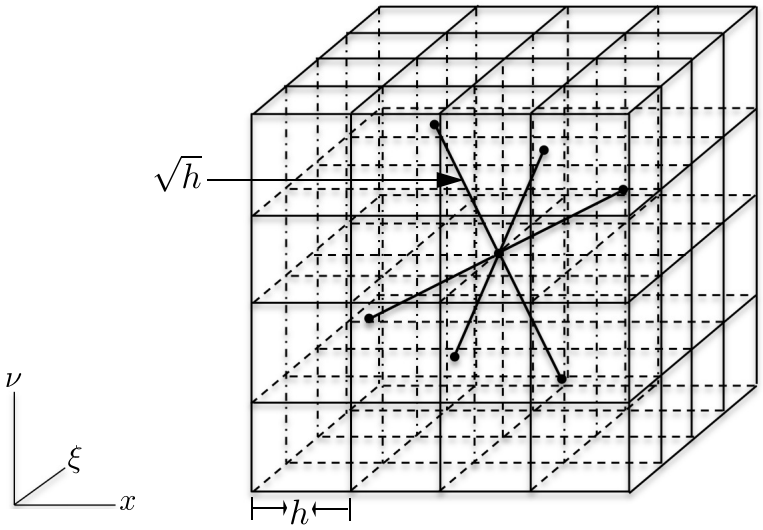


Figure 4.2: Illustration of wide stencil.

We now present the details of discretization for (4.29). For notational convenience, we define the following

$$U^n(X) \equiv U(x, \nu, \xi, \tau^n), \quad X = \begin{pmatrix} x \\ \nu \\ \xi \end{pmatrix}, \quad \mathcal{U}^n(Y) \equiv \mathcal{U}(y_1, y_2, y_3, \tau^n), \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

The second order derivatives in (4.4) are approximated at $(x_i, \nu_j, \xi_k, \tau^n)$ in terms of (4.29)

as

$$
\begin{aligned}
(\lambda_1)_{i,j,k} & \frac{\mathcal{U}^n\left(Y_{i,j,k}+\sqrt{h}\mathbf{e}_1\right)+\mathcal{U}^n\left(Y_{i,j,k}-\sqrt{h}\mathbf{e}_1\right)-2\,\mathcal{U}^n\left(Y_{i,j,k}\right)}{h} \\
& + (\lambda_2)_{i,j,k}\frac{\mathcal{U}^n\left(Y_{i,j,k}+\sqrt{h}\mathbf{e}_2\right)+\mathcal{U}^n\left(Y_{i,j,k}-\sqrt{h}\mathbf{e}_2\right)-2\,\mathcal{U}^n\left(Y_{i,j,k}\right)}{h} \\
& + (\lambda_3)_{i,j,k}\frac{\mathcal{U}^n\left(Y_{i,j,k}+\sqrt{h}\mathbf{e}_3\right)+\mathcal{U}^n\left(Y_{i,j,k}-\sqrt{h}\mathbf{e}_3\right)-2\,\mathcal{U}^n\left(Y_{i,j,k}\right)}{h} \\
\approx \quad & (\lambda_1)_{i,j,k}\frac{\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\left(R_{i,j,k}\right)_1\right)+\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\left(R_{i,j,k}\right)_1\right)-2\,U^n\left(X_{i,j,k}\right)}{h} \\
& + (\lambda_2)_{i,j,k}\frac{\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\left(R_{i,j,k}\right)_2\right)+\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\left(R_{i,j,k}\right)_2\right)-2\,U^n\left(X_{i,j,k}\right)}{h} \\
& + (\lambda_3)_{i,j,k}\frac{\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\left(R_{i,j,k}\right)_3\right)+\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\left(R_{i,j,k}\right)_3\right)-2\,U^n\left(X_{i,j,k}\right)}{h},
\end{aligned} \tag{4.30}
$$

where $X_{i,j,k}=(x_i,\nu_j,\xi_k)$, $X_{i,j,k}=R_{i,j,k}Y_{i,j,k}$ and $(R_{i,j,k})_{\hat{l}}$ is the $\hat{l}$-th column of the eigenvector matrix and

$$
\mathbf{e}_1=\begin{pmatrix}1\\0\\0\end{pmatrix},\quad \mathbf{e}_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad \mathbf{e}_3=\begin{pmatrix}0\\0\\1\end{pmatrix}.
$$

The wide stencil discretization (4.30) is $O(h)$ consistent, a detailed analysis of consistency is provided in Section 4.4.1. Using (4.30) and an upstream finite differencing for the first order derivatives to ensure positive coefficient condition, the wide stencil discrete

operator $\mathcal{L}_w$ of the differential operator $\mathcal{L}$ (4.4) is given by

$$\begin{aligned}
\mathcal{L}_w U^n_{i,j,k} \;=\; & \frac{(\lambda_1)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\,(R_{i,j,k})_1\right)+\frac{(\lambda_1)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\,(R_{i,j,k})_1\right)\\
& +\frac{(\lambda_2)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\,(R_{i,j,k})_2\right)+\frac{(\lambda_2)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\,(R_{i,j,k})_2\right)\\
& +\frac{(\lambda_3)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\,(R_{i,j,k})_3\right)+\frac{(\lambda_3)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\,(R_{i,j,k})_3\right)\\
& +\mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i}U^n_{i+1,j,k}-\mathbb{1}_{(r-\frac{\nu_j}{2})<0}\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i}U^n_{i-1,j,k}\\
& +\mathbb{1}_{(\xi_k-\nu_j)\geq 0}\frac{C_1\,(\xi_k-\nu_j)}{\Delta^+\nu_j}U^n_{i,j+1,k}-\mathbb{1}_{(\xi_k-\nu_j)<0}\frac{C_1\,(\xi_k-\nu_j)}{\Delta^-\nu_j}U^n_{i,j-1,k}\\
& +\mathbb{1}_{(C_4-\xi_k)\geq 0}\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}U^n_{i,j,k+1}-\mathbb{1}_{(C_4-\xi_k)<0}\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k}U^n_{i,j,k-1}\\
& -\Bigg(\mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i}-\mathbb{1}_{(r-\frac{\nu_j}{2})<0}\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i}+\mathbb{1}_{(\xi_k-\nu_j)\geq 0}\frac{C_1\,(\xi_k-\nu_j)}{\Delta^+\nu_j}\\
& -\mathbb{1}_{(\xi_k-\nu_j)<0}\frac{C_1\,(\xi_k-\nu_j)}{\Delta^-\nu_j}+\mathbb{1}_{(C_4-\xi_k)\geq 0}\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}-\mathbb{1}_{(C_4-\xi_k)<0}\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k}\\
& +\frac{2\,(\lambda_1)_{i,j,k}}{h}+\frac{2\,(\lambda_2)_{i,j,k}}{h}+\frac{2\,(\lambda_3)_{i,j,k}}{h}+r\Bigg)U^n_{i,j,k}. \qquad (4.31)
\end{aligned}$$

Using fully implicit timestepping, (4.3) has the following discretized form in the wide stencil,

$$\frac{U^{n+1}_{i,j,k}-U^n_{i,j,k}}{\Delta\tau}=\mathcal{L}_w U^{n+1}_{i,j,k}. \qquad (4.32)$$

When applied to the two-dimensional problem in [83], our approach is equivalent to the wide stencil method based on a local coordinate rotation. The main difference is that [83] uses a geometric approach to locally transform the coordinate system, whereas we algebraically transform the local coordinate system. In doing so, we do not have to explicitly determine the rotation angle and hence our approach can be easily extended to any number of higher dimensions.

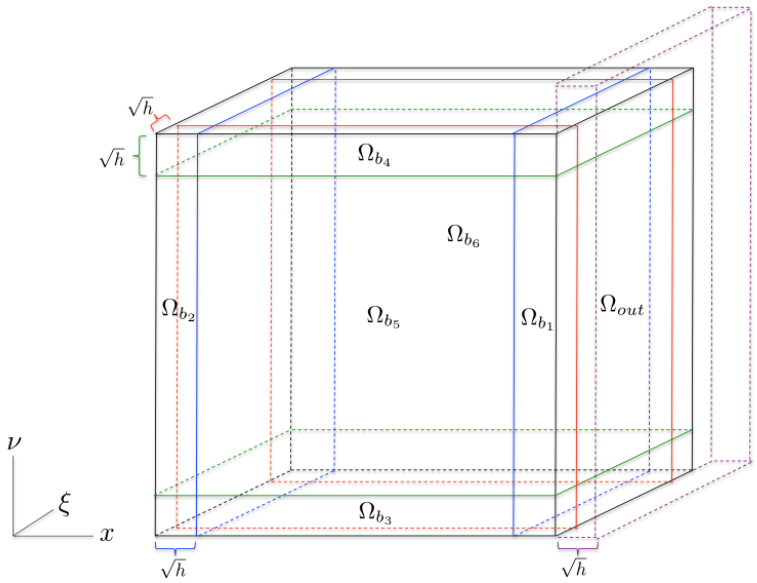


Figure 4.3: Boundary regions of wide stencil discretization

4.3.2.1 Boundary Regions for the Wide Stencil Discretization

One or more of the grid points $X_{i,j,k} \pm \sqrt{h}\,(R_{i,j,k})_{\hat{l}}$, $\hat{l} = 1, 2, 3$ in the wide stencil discretization (4.30) may lie outside the computational domain Ω. When this happens, the grid point (x_i, ν_j, ξ_k) lies in the one of the regions: $\Omega_{b_{\hat{l}}}$, $\hat{l} = 1, \ldots, 6$ shown in Figure 4.3. These regions along with Ω_{in} and Ω_{out} are defined in Table 4.3.

When $(x_i, \nu_j, \xi_k) \in \Omega_{b_1}$, points in the wide stencil discretization (4.31) lie in the region $\Omega_{in} \cup \Omega_{out}$. No special consideration is required if the points lie in Ω_{in}. However, when they lie in Ω_{out}, the asymptotic solution (4.5) is used at that point [83]. Therefore, the real issue is when $(x_i, \nu_j, \xi_k) \in \Omega_{b_{\hat{l}}}$, $\hat{l} = 2, \ldots, 6$.

For notational convenience, we denote the cross derivative terms in (4.4) as

$$\gamma_{x\nu} = \rho_{12} C_2 \nu^{\alpha} \sqrt{\nu} \frac{\partial^2 U}{\partial x \partial \nu}, \quad \gamma_{x\xi} = \rho_{13} C_5 \xi^{\beta} \sqrt{\nu} \frac{\partial^2 U}{\partial x \partial \xi}, \quad \gamma_{\nu\xi} = \rho_{23} C_2 C_5 \nu^{\alpha} \xi^{\beta} \frac{\partial^2 U}{\partial \nu \partial \xi}.$$

Recall that $\alpha, \beta \in [0.5, 1]$, ρ_{12}, ρ_{13} ρ_{23}, C_2, C_5 are constants [42] and ν, ξ are bounded. We now study the behavior of $\gamma_{x\nu}$, $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ in Ω_{b_5} in Figure 4.3. In this region, $\xi \in [0, \sqrt{h}]$ and hence $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ are $O(h^{\frac{\beta}{2}})$ for smooth test function ϕ. The term $\gamma_{x\nu}$ may or may not be small. To further analyze this, consider a cross-section of Ω_{b_5} as shown in Figure 4.4.

Ω_{τ_0}	$[x_{\min}, x_{\max}] \times [0, \nu_{\max}] \times [0, \xi_{\max}] \times \{0\}$
Ω_{in}	$\Omega / \Omega_{\tau_0} / \Omega_{x_{\max}}$
Ω_{out}	$\left(x_{\max}, x_{\max} + \sqrt{h}\right] \times \left[0, \nu_{\max} + \sqrt{h}\right] \times \left[0, \xi_{\max} + \sqrt{h}\right] \times (0, T]$
Ω_{b_1}	$\left[x_{\max} - \sqrt{h}, x_{\max}\right) \times (0, \nu_{\max}) \times (0, \xi_{\max}) \times (0, T]$
Ω_{b_2}	$\left(x_{\min}, x_{\min} + \sqrt{h}\right] \times (0, \nu_{\max}) \times (0, \xi_{\max}) \times (0, T]$
Ω_{b_3}	$(x_{\min}, x_{\max}) \times \left(0, \sqrt{h}\right] \times (0, \xi_{\max}) \times (0, T]$
Ω_{b_4}	$(x_{\min}, x_{\max}) \times \left[\nu_{\max} - \sqrt{h}, \nu_{\max}\right) \times (0, \xi_{\max}) \times (0, T]$
Ω_{b_5}	$(x_{\min}, x_{\max}) \times (0, \nu_{\max}) \times \left(0, \sqrt{h}\right] \times (0, T]$
Ω_{b_6}	$(x_{\min}, x_{\max}) \times (0, \nu_{\max}) \times \left[\xi_{\max} - \sqrt{h}, \xi_{\max}\right) \times (0, T]$

Table 4.3: Domain definitions for wide stencil discretization.

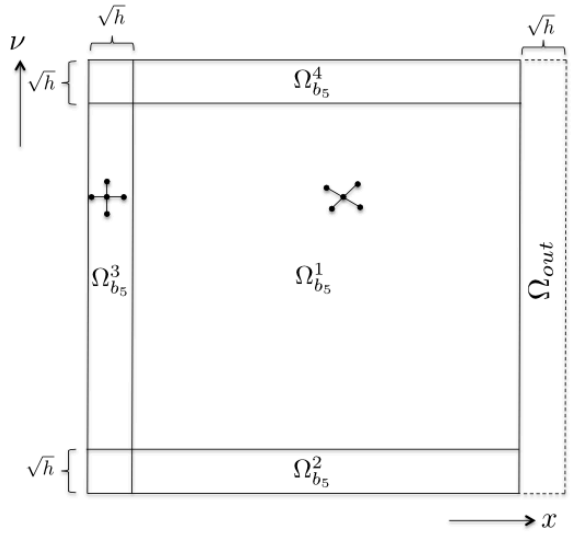


Figure 4.4: Cross sections of Ω_{b_5}. Illustration of locally transformed coordinate system in the inner region ($\Omega^1_{b_5}$) and in the boundary region ($\Omega^3_{b_5}$) as $h \to 0$.

In $\Omega^2_{b_5}$, $\nu \in (0, \sqrt{h}]$, therefore, $\gamma_{x\nu}$ is $O(h^{\frac{2\alpha+1}{4}})$. In $\Omega^3_{b_5}$ in Figure 4.4, $x \in (x_{\min}, x_{\min} + \sqrt{h}]$. Recall that at $x = x_{\min}$, we set $S = 0$ and hence in $\Omega^3_{b_5}$, $S \in (0, \sqrt{h}]$. Since $\frac{\partial^2 U}{\partial x \partial \nu} = S\frac{\partial^2 U}{\partial S \partial \nu}$, $\gamma_{x\nu}$ is $O(\sqrt{h})$ in $\Omega^3_{b_5}$. Next consider $\Omega^4_{b_5}$ in Figure 4.4, where $\nu \in [\nu_{\max} - \sqrt{h}, \nu_{\max})$. Recall that at $\nu = \nu_{\max}$, we use $\frac{\partial U}{\partial \nu} = 0$. Therefore, when $\nu \in [\nu_{\max} - \sqrt{h}, \nu_{\max})$, $\frac{\partial U}{\partial \nu}$ is $O(\sqrt{h})$ and it follows that $\gamma_{x\nu}$ is $O(\sqrt{h})$. As $h \to 0$, all the cross derivative terms $\gamma_{x\nu}$, $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ in $\Omega^2_{b_5}$, $\Omega^3_{b_5}$ and $\Omega^4_{b_5}$ also approach zero and hence a local coordinate transformation in these regions will not result in a discretization with grid points outside Ω. For example, consider $\gamma_{x\nu}$ in Ω_{b_5}. We note that a local coordinate transformation to eliminate $\gamma_{x\nu}$ is geometrically equivalent to a local rotation of grid in the (x, ν) dimension. When a grid point lies in $\Omega^1_{b_5}$, the wide stencil could possibly result in a locally rotated grid as shown in Figure 4.4. But, for grid points near the boundary regions $\Omega^{\hat{l}}_{b_5}$, $l = 2, 3, 4$ and as $h \to 0$, the angle of rotation to eliminate the cross term $\gamma_{x\nu}$ would be very small or negligent such that the locally rotated grid is almost perpendicular to the boundary. This is shown for a grid point in $\Omega^3_{b_5}$ in Figure 4.4. Hence, no points of the discretization lie outside the computational domain when grid points lie in $\Omega^{\hat{l}}_{b_5}$, $l = 2, 3, 4$ and as $h \to 0$.

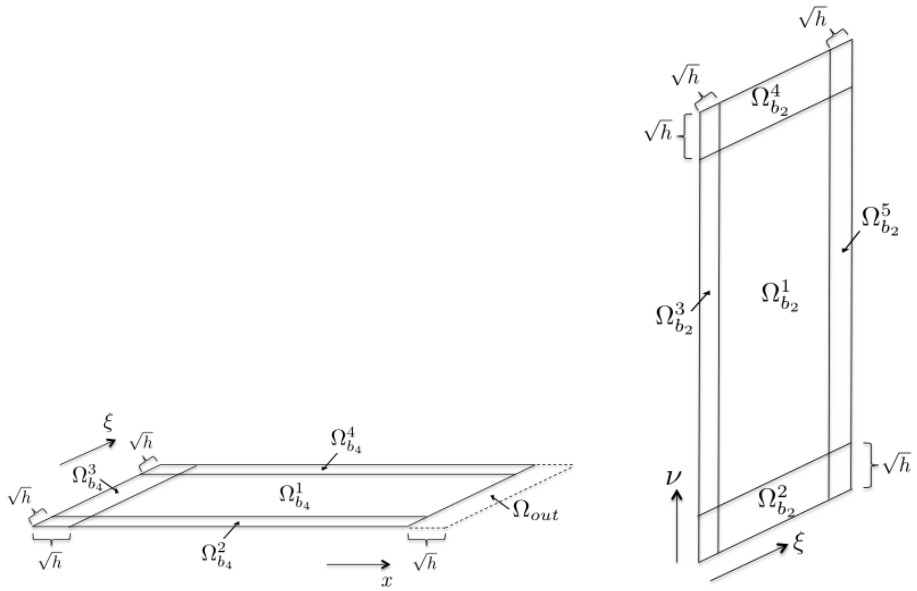


Figure 4.5: Cross sections of Ω_{b_4} and Ω_{b_2}.

Finally in $\Omega^1_{b_5}$ in Figure 4.4, $\gamma_{x\nu}$ may not be small in general but $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ are small.

A local rotation in the (x, ν) dimension will not result in points outside the computational domain as $\Omega^1_{b_5}$ is more than $\sqrt{h}$ distance away from the boundaries as shown in Figure 4.4. Hence, all the grid points in the discretization (4.31) lie within $\Omega_{in} \cup \Omega_{out}$. Hence, in theory, when $(x_i, \nu_j, \xi_k) \in \Omega_{b_5}$ and for sufficiently small h, the points of the wide stencil discretization will lie in $\Omega_{in} \cup \Omega_{out}$. Similar idea applies to the other regions $\Omega_{b_{\hat{l}}}$, $\hat{l} = 2, \ldots, 4$ and Ω_{b_6} in Figure 4.3 as well, the details are omitted here.

In practice, h is not sufficiently small, therefore the idea presented above may not hold. Hence, we have to address the issue of points lying outside the domain in the discretization (4.31). We propose to solve approximations to the PDE (4.3) in the boundary regions, $\Omega_{b_{\hat{l}}}$, $\hat{l} = 2, \ldots, 6$ such that its discretization does not result in grid points outside the domain. Once again, we consider Ω_{b_5} in Figure 4.4 to illustrate the idea. In $\Omega^1_{b_5}$, we completely ignore the cross derivative terms $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ in (4.3) and solve the following approximation

$$\begin{aligned} \frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{12}C_2\nu^{\alpha}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \nu} \\ &\quad + \left(r - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + C_1\left(\xi - \nu\right)\frac{\partial U}{\partial \nu} + C_3\left(C_4 - \xi\right)\frac{\partial U}{\partial \xi} - rU, \end{aligned} \tag{4.33}$$

which is an $O(h^{\frac{\beta}{2}})$ approximation to (4.3). The only cross derivative term in (4.33) is $\gamma_{x\nu}$ and the wide stencil discretization will have all its points in $\Omega_{in} \cup \Omega_{out}$. The wide stencil discretization of (4.33) will result in (4.31), where R is now the eigenvector matrix of the diffusion matrix of (4.33) given by

$$D_1 = \frac{1}{2}\begin{pmatrix} \nu & \rho_{12}C_2\nu^{\alpha}\sqrt{\nu} & 0 \\ \rho_{12}C_2\nu^{\alpha}\sqrt{\nu} & C_2^2\nu^{2\alpha} & 0 \\ 0 & 0 & C_5^2\xi^{2\beta} \end{pmatrix}$$

In regions $\Omega^2_{b_5}$, $\Omega^3_{b_5}$ and $\Omega^4_{b_5}$, the presence of one or more of the cross derivative terms may cause the discretization to include points outside $\Omega_{in} \cup \Omega_{out}$. Hence we ignore all the cross derivative terms $\gamma_{x\nu}$, $\gamma_{x\xi}$ and $\gamma_{\nu\xi}$ and solve the following, which is $O(h^{\frac{\beta}{2}})$ approximation to (4.3).

$$\begin{aligned} \frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \left(r - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} \\ &\quad + C_1\left(\xi - \nu\right)\frac{\partial U}{\partial \nu} + C_3\left(C_4 - \xi\right)\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.34}$$

With no cross derivative terms in (4.34), discretization is performed on a standard seven point fixed stencil. The discretization results in (4.22) with the coefficients $\zeta^{x\nu}_{i,j,k}$, $\zeta^{x\xi}_{i,j,k}$ and

$\zeta^{\nu\xi}_{i,j,k}$ equal to zero. Since all cross derivative terms are zero, this discretization naturally satisfies the positive coefficient condition (4.24)-(4.28) as per the definition of the discrete coefficients in Algorithm 6 in Appendix C.

Applying similar idea to other regions in $\Omega_{b_{\hat{l}}}$, $\hat{l} = 2, \ldots, 4$ and Ω_{b_6}, we solve approximations of (4.3) with appropriate cross derivative terms eliminated. We now present the approximations used in these regions along with the order of error of this approximation to (4.3). Equation (4.34) is used in Ω_{b_6} and Ω_{b_3}, which is $O(\sqrt{h})$ and $O(h^{\frac{1}{4}})$ approximations to (4.3), respectively.

The region Ω_{b_4} is divided into four sub-regions as shown in Figure 4.5(a). We use the following in $\Omega_1^{b_4}$,

$$\begin{aligned}\frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{13}C_5\xi^{\beta}\sqrt{\nu}\frac{\partial^2 U}{\partial x \partial \xi} \\ &\quad + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} + C_1(\xi - \nu)\frac{\partial U}{\partial \nu} + C_3(C_4 - \xi)\frac{\partial U}{\partial \xi} - rU, \end{aligned} \tag{4.35}$$

which is an $O(\sqrt{h})$ approximation of (4.3). A wide stencil discretization of (4.35) will result in (4.31), where R is now the eigenvector matrix of the diffusion matrix of (4.35). In $\Omega_2^{b_4}$, $\Omega_3^{b_4}$ and $\Omega_4^{b_4}$ we solve (4.34) which is an $O(h^{\frac{\beta}{2}})$ approximation of (4.3).

The region Ω_{b_2} is divided into five sub-regions as shown in Figure 4.5(b). The following $O(\sqrt{h})$ approximation of (4.3) is used in $\Omega_{b_2}^1$:

$$\begin{aligned}\frac{\partial U}{\partial \tau} &= \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta}\frac{\partial^2 U}{\partial \nu \partial \xi} \\ &\quad + (r - \frac{1}{2}\nu)\frac{\partial U}{\partial x} + C_1(\xi - \nu)\frac{\partial U}{\partial \nu} + C_3(C_4 - \xi)\frac{\partial U}{\partial \xi} - rU. \end{aligned} \tag{4.36}$$

A wide stencil discretization of (4.36) will result in (4.31), where R is now the eigenvector matrix of the diffusion matrix of (4.36). In $\Omega_{b_2}^{\hat{l}}$, $\hat{l} = 2, \ldots, 5$, we solve (4.34) which is an $O(h^{\gamma})$ approximation of (4.3), where $\gamma = \min(\frac{\alpha}{2}, \frac{\beta}{2})$.

4.3.2.2 Boundaries

The boundary conditions are given by (4.7)-(4.20). Among these, three PDEs (4.7), (4.9) and (4.14) involve cross derivative terms. We use the wide stencil method in Section 4.3.2 to discretize these PDEs. Note that the problem of points lying outside the domain still

exists here. So, following the idea proposed in Section 4.3.2.1, we solve approximations to these PDEs in the appropriate regions as described below.

First consider the region $\Omega_{x_{\min}}$ in Figure 4.1. This region is further divided into five regions as shown in Figure 4.5(b). In $\Omega^1_{x_{\min}}$, we solve (4.7) using a wide stencil discretization with all points inside the domain $\Omega_{x_{\min}}$. The wide stencil discretization formula is given by

$$\begin{aligned}
\mathcal{L}_w \;=\; & \frac{(\lambda_1^{x_{\min}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\,(R_{i,j,k})_1^{x_{\min}}\right)+\frac{(\lambda_1^{x_{\min}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\,(R_{i,j,k})_1^{x_{\min}}\right)\\
& +\frac{(\lambda_2^{x_{\min}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}-\sqrt{h}\,(R_{i,j,k})_2^{x_{\min}}\right)+\frac{(\lambda_2^{x_{\min}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k}+\sqrt{h}\,(R_{i,j,k})_2^{x_{\min}}\right)\\
& +\mathbb{1}_{(\xi_k-\nu_j)\geq 0}\,\frac{C_1\,(\xi_k-\nu_j)}{\Delta^+\nu_j}U^n_{i,j+1,k}-\mathbb{1}_{(\xi_k-\nu_j)<0}\,\frac{C_1\,(\xi_k-\nu_j)}{\Delta^-\nu_j}U^n_{i,j-1,k}\\
& +\mathbb{1}_{(C_4-\xi_k)\geq 0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}U^n_{i,j,k+1}-\mathbb{1}_{(C_4-\xi_k)<0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k}U^n_{i,j,k-1}\\
& -\left(\mathbb{1}_{(\xi_k-\nu_j)\geq 0}\,\frac{C_1\,(\xi_k-\nu_j)}{\Delta^+\nu_j}-\mathbb{1}_{(\xi_k-\nu_j)<0}\,\frac{C_1\,(\xi_k-\nu_j)}{\Delta^-\nu_j}+\mathbb{1}_{(C_4-\xi_k)\geq 0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}\right.\\
& \left.-\mathbb{1}_{(C_4-\xi_k)<0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k}+\frac{2\,(\lambda_1^{x_{\min}})_{i,j,k}}{h}+\frac{2\,(\lambda_2^{x_{\min}})_{i,j,k}}{h}+r\right)U^n_{i,j,k},
\end{aligned}\tag{4.37}$$

where $R^{x_{\min}}$ and $\lambda_1^{x_{\min}}$ and $\lambda_2^{x_{\min}}$ are the eigenvector matrix and eigenvalues of the diffusion matrix of (4.7) given by

$$D^{x_{\min}}=\frac{1}{2}\begin{pmatrix} C_2^2\nu^{2\alpha} & \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta}\\ \rho_{23}C_2C_5\nu^{\alpha}\xi^{\beta} & C_5^2\xi^{2\beta}\end{pmatrix}.$$

In regions $\Omega^{\hat{l}}_{x_{\min}}$, $\hat{l}=2,\ldots,5$, we ignore $\gamma_{\nu\xi}$ and solve

$$\frac{\partial U}{\partial \tau}=\frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2}+\frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2}+C_1\,(\xi-\nu)\,\frac{\partial U}{\partial \nu}+C_3\,(C_4-\xi)\,\frac{\partial U}{\partial \xi}-rU,\tag{4.38}$$

which is an $O(h^{\gamma})$ approximation to (4.7), where $\gamma=\min(\frac{\alpha}{2},\frac{\beta}{2})$. A standard seven point stencil is used to discretize (4.38), which results in (4.22) with $\alpha^x_{i,j,k}$, $\beta^x_{i,j,k}$, $\zeta^{x\nu}_{i,j,k}$, $\zeta^{x\xi}_{i,j,k}$ and $\zeta^{\nu\xi}_{i,j,k}$ set to zero.

Next consider Ω_{ξ_0} in Figure 4.1. This region is further divided into four regions as shown in Figure 4.4. In $\Omega^1_{\xi_0}$, we solve (4.9) and a local coordinate transformation to eliminate

$\gamma_{x\nu}$ will have all the points inside the domain $\Omega_{\xi_0} \cup \Omega_{out}$. The wide stencil discretization results in

$$\begin{aligned}
\mathcal{L}_w \;=\; & \frac{\left(\lambda_1^{\xi_0}\right)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} - \sqrt{h}\,(R_{i,j,k})_1^{\xi_0}\right) + \frac{\left(\lambda_1^{\xi_0}\right)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} + \sqrt{h}\,(R_{i,j,k})_1^{\xi_0}\right) \\
& + \frac{\left(\lambda_2^{\xi_0}\right)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} - \sqrt{h}\,(R_{i,j,k})_2^{\xi_0}\right) + \frac{\left(\lambda_2^{\xi_0}\right)_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} + \sqrt{h}\,(R_{i,j,k})_2^{\xi_0}\right) \\
& + \mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i}U^n_{i+1,j,k} - \mathbb{1}_{(r-\frac{\nu_j}{2})<0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i}U^n_{i-1,j,k} + \mathbb{1}_{(\xi_k-\nu_j)\geq 0}\,\frac{C_1\left(\xi_k-\nu_j\right)}{\Delta^+\nu_j}U^n_{i,j+1,k} \\
& - \mathbb{1}_{(\xi_k-\nu_j)<0}\,\frac{C_1\left(\xi_k-\nu_j\right)}{\Delta^-\nu_j}U^n_{i,j-1,k} + \frac{C_3\left(C_4-\xi_k\right)}{\Delta^+\xi_k}U^n_{i,j,k+1} - \Bigg(\mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i} \\
& - \mathbb{1}_{(r-\frac{\nu_j}{2})<0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i} + \mathbb{1}_{(\xi_k-\nu_j)\geq 0}\,\frac{C_1\left(\xi_k-\nu_j\right)}{\Delta^+\nu_j} - \mathbb{1}_{(\xi_k-\nu_j)<0}\,\frac{C_1\left(\xi_k-\nu_j\right)}{\Delta^-\nu_j} + \frac{C_3\left(C_4-\xi_k\right)}{\Delta^+\xi_k} \\
& + \frac{2\left(\lambda_1^{\xi_0}\right)_{i,j,k}}{h} + \frac{2\left(\lambda_2^{\xi_0}\right)_{i,j,k}}{h} + r\Bigg)U^n_{i,j,k}, \qquad (4.39)
\end{aligned}$$

where R^{ξ_0} is the eigenvector matrix and $\lambda_1^{\xi_0}$ and $\lambda_2^{\xi_0}$ are the eigenvalues of the diffusion matrix of (4.9) given by

$$D = \frac{1}{2}\begin{pmatrix} \nu & \rho_{12}C_2\nu^\alpha\sqrt{\nu} \\ \rho_{12}C_2\nu^\alpha\sqrt{\nu} & C_2^2\nu^{2\alpha}\end{pmatrix}.$$

In regions $\Omega^2_{\xi_0}$, $\Omega^3_{\xi_0}$ and $\Omega^4_{\xi_0}$, we solve

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_2^2\nu^{2\alpha}\frac{\partial^2 U}{\partial \nu^2} + (r-\frac{1}{2}\nu)\frac{\partial U}{\partial x} - C_1\nu\frac{\partial U}{\partial \nu} + C_3C_4\frac{\partial U}{\partial \xi} - rU, \qquad (4.40)$$

which is an $O(h^\gamma)$ approximation to (4.9), where $\gamma = \min(\frac{1}{2}, \frac{2\alpha+1}{4})$. A standard seven point stencil is used to discretize (4.40), which results in (4.22) with $\alpha^\xi_{i,j,k}$, $\zeta^{x\nu}_{i,j,k}$, $\zeta^{x\xi}_{i,j,k}$ and $\zeta^{\nu\xi}_{i,j,k}$ set to zero.

Finally consider $\Omega_{\nu_{\max}}$ in Figure 4.1. This region is further divided into four regions as shown in Figure 4.5(a). In $\Omega^1_{\nu_{\max}}$, a wide stencil discretization to solve (4.14) will have all

the points inside the domain $\Omega_{\nu_{\max}} \cup \Omega_{out}$. The discretization formula is given by

$$\begin{aligned}
\mathcal{L}_w \;=\; & \frac{(\lambda_1^{\nu_{\max}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} - \sqrt{h}\,(R_{i,j,k})_1^{\nu_{\max}}\right) + \frac{(\lambda_1^{\nu_{\max}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} + \sqrt{h}\,(R_{i,j,k})_1^{\nu_{\max}}\right) \\
& + \frac{(\lambda_2^{\nu_{\max}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} - \sqrt{h}\,(R_{i,j,k})_2^{\nu_{\max}}\right) + \frac{(\lambda_2^{\nu_{\max}})_{i,j,k}}{h}\mathcal{J}_h U^n\left(X_{i,j,k} + \sqrt{h}\,(R_{i,j,k})_2^{\nu_{\max}}\right) \\
& + \mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i}U^n_{i+1,j,k} - \mathbb{1}_{(r-\frac{\nu_j}{2})<0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i}U^n_{i-1,j,k} \\
& + \mathbb{1}_{(C_4-\xi_k)\geq 0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}U^n_{i,j,k+1} - \mathbb{1}_{(C_4-\xi_k)<0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k}U^n_{i,j,k-1} \\
& - \left(\mathbb{1}_{(r-\frac{\nu_j}{2})\geq 0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^+ x_i} - \mathbb{1}_{(r-\frac{\nu_j}{2})<0}\,\frac{\left(r-\frac{\nu_j}{2}\right)}{\Delta^- x_i} + \mathbb{1}_{(C_4-\xi_k)\geq 0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^+\xi_k}\right. \\
& \left. - \mathbb{1}_{(C_4-\xi_k)<0}\,\frac{C_3\,(C_4-\xi_k)}{\Delta^-\xi_k} + \frac{2\,(\lambda_1^{\nu_{\max}})_{i,j,k}}{h} + \frac{2\,(\lambda_2^{\nu_{\max}})_{i,j,k}}{h} + r\right)U^n_{i,j,k}, \qquad (4.41)
\end{aligned}$$

where $R^{\nu_{\max}}$ is the eigenvector matrix and $\lambda_1^{\nu_{\max}}$ and $\lambda_2^{\nu_{\max}}$ are the eigenvalues of the diffusion matrix of (4.14) given by

$$D = \frac{1}{2}\begin{pmatrix} \nu & \rho_{13}C_5\xi^\beta\sqrt{\nu} \\ \rho_{13}C_5\xi^\beta\sqrt{\nu} & C_5^2\xi^{2\beta}\end{pmatrix}.$$

In regions $\Omega^2_{\nu_{\max}}$, $\Omega^3_{\nu_{\max}}$ and $\Omega^4_{\nu_{\max}}$, we solve

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\nu\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}C_5^2\xi^{2\beta}\frac{\partial^2 U}{\partial \xi^2} + \left(r - \frac{1}{2}\nu\right)\frac{\partial U}{\partial x} + C_3\,(C_4-\xi)\,\frac{\partial U}{\partial \xi} - rU, \qquad (4.42)$$

which is an $O(h^{\frac{\beta}{2}})$ approximation to (4.14). A standard seven point stencil is used to discretize (4.42), which results in (4.22) with $\alpha^\nu_{i,j,k}$, $\beta^\nu_{i,j,k}$, $\zeta^{x\nu}_{i,j,k}$, $\zeta^{x\xi}_{i,j,k}$ and $\zeta^{\nu\xi}_{i,j,k}$ set to zero.

4.3.3 Hybrid Stencil

The fixed stencil (Section 4.3.1) uses second order approximation for the diffusion terms and second order as much as possible for the drift terms. A disadvantage with the fixed stencil is that the positive coefficient condition (4.24)-(4.28) is not guaranteed across the entire domain. The wide stencil (Section 4.3.2) is only first order accurate for the diffusion terms. But, the wide stencil results in a positive coefficient discretization. On the other hand, the wide stencil is computationally expensive compared to the fixed stencil. In order

to take advantage of both the methods, we use a hybrid stencil, which is a combination of the fixed and wide stencils [83]. For the hybrid stencil, we follow the approach of using different approximations to the PDE (4.3) in different regions of Ω_{in} as detailed in Section 4.3.2. As for the discretization method, for a given grid point (x_i, ν_j, ξ_k), we first check if the discretization on the fixed stencil (4.22) satisfies the positive coefficient condition (4.24)-(4.28). If it does, the standard fixed stencil is used. Otherwise, the wide stencil discretization (one of (4.31), (4.37), (4.39), (4.41)) is used. Therefore, in the hybrid stencil, Ω_{in} is composed of two regions: Ω_f, the region in which the positive coefficient condition (4.24)-(4.28) holds and Ω_w, the region in which (4.24)-(4.28) does not hold. Discretization on hybrid stencil uses fixed stencil discretization as much as possible to take advantage of its accuracy and computational efficiency, while still satisfying the positive coefficient condition (4.24)-(4.28).

4.3.3.1 The matrix form of discrete equations

In this section, the discrete PDEs (4.23) and (4.32) are written in matrix form for notational convenience for computational purposes. Let U^n be the solution vector defined by

$$U^n = \begin{bmatrix} U_0^n & U_1^n & \dots & U_N^n \end{bmatrix}.$$

Let $\mathbf{L}$ be an $N \times N$ discretization matrix and $\mathbf{L}_{l,l'}$ be its entries for row l and column l'. Recall that $l = k \cdot n_x \cdot n_\nu + j \cdot n_x + i$, $i = 0, \dots, n_x - 1$, $j = 0, \dots, n_\nu - 1$ and $k = 0, \dots, n_\xi - 1$. For $(x_i, \nu_j, \xi_k) \in \Omega_{x_{\max}}$, where the Dirichlet boundary condition is imposed, we have

$$\mathbf{L}_{l,l'} = 0, \quad l' = 0, \dots, N-1. \tag{4.43}$$

We also define the vector $\mathbf{F}$ as

$$\mathbf{F}_l = \begin{cases} U(x_i, \nu_j, \xi_k), & (x_i, \nu_j, \xi_k) \in \Omega_{x_{\max}} \\ 0, & \text{otherwise.} \end{cases} \tag{4.44}$$

For the case $(x_i, \nu_j, \xi_k) \in \Omega_f$, we use the discrete linear operator $\mathcal{L}_f$ (4.22) to populate the entries $\mathbf{L}_{l,l'}$,

$$\left[\mathbf{L}U^{n+1}\right]_l = \mathcal{L}_f U_{i,j,k}^{n+1}. \tag{4.45}$$

For the case, $(x_i, \nu_j, \xi_k) \in \Omega_w$, the wide stencil discretization $\mathcal{L}_w$ (4.32) results in six off-grid points $X_{i,j,k} \pm \sqrt{h}(R_{i,j,k})_{\hat{l}}$, $\hat{l} = 1, 2, 3$. Let these six points be denoted as $P_{i,j,k}^m$, $m = 1, \dots, 6$, respectively. Note that these points can lie either in Ω_{in} or Ω_{out}. First, we consider that

all the six points lie inside Ω_{in}. The values at these points are approximated using linear interpolation as

$$\mathcal{J}_h U^{n+1}\left(P^m_{i,j,k}\right) = \sum_{a,b,c=0,1} w^{p_m+a,q_m+b,r_m+c}_{i,j,k} U^{n+1}_{p_m+a,q_m+b,r_m+c}, \quad m = 1,\ldots,6, \tag{4.46}$$

where $w^{p_m+a,q_m+b,r_m+c}_{i,j,k} \geq 0$ and $\sum_{a,b,c=0,1} w^{p_m+a,q_m+b,r_m+c}_{i,j,k} = 1$. Using (4.46) in (4.31), the entries $\mathbf{L}_{l,l'}$ of the l-th row are given as

$$\left[\mathbf{L}U^{n+1}\right]_l = \mathcal{L}_w U^{n+1}_{i,j,k}. \tag{4.47}$$

Next, consider a case where a point $P^m_{i,j,k}$ is in Ω_{out} in which the asymptotic solution is used at that point. We define a vector $\mathbf{G}$ to accommodate points in Ω_{out},

$$\mathbf{G} = \begin{cases} \mathbb{1}_{P^1_{i,j,k}\in\Omega_{out}} \frac{(\lambda_1)_{i,j,k}}{h} U\left(P^1_{i,j,k}\right) + \mathbb{1}_{P^2_{i,j,k}\in\Omega_{out}} \frac{(\lambda_1)_{i,j,k}}{h} U\left(P^2_{i,j,k}\right) \\ \quad + \mathbb{1}_{P^3_{i,j,k}\in\Omega_{out}} \frac{(\lambda_2)_{i,j,k}}{h} U\left(P^3_{i,j,k}\right) + \mathbb{1}_{P^4_{i,j,k}\in\Omega_{out}} \frac{(\lambda_2)_{i,j,k}}{h} U\left(P^4_{i,j,k}\right) \\ \quad + \mathbb{1}_{P^5_{i,j,k}\in\Omega_{out}} \frac{(\lambda_3)_{i,j,k}}{h} U\left(P^5_{i,j,k}\right) + \mathbb{1}_{P^6_{i,j,k}\in\Omega_{out}} \frac{(\lambda_3)_{i,j,k}}{h} U\left(P^6_{i,j,k}\right), & (x_i,\nu_j,\xi_k)\in\Omega_w, \\ 0, & \text{otherwise}, \end{cases} \tag{4.48}$$

where $U(P^m_{i,j,k})$ is the asymptotic solution (4.5) at $P^m_{i,j,k}$. The matrix form of the discretized equations is then given by

$$\left[\mathbf{I} - \Delta\tau\mathbf{L}\right] U^{n+1} = U^n + \mathbf{F}^{n+1} - \mathbf{F}^n + \Delta\tau\mathbf{G}. \tag{4.49}$$

4.4 Discretization Analysis

In this section, we perform discretization analysis for the hybrid stencil discretization (4.49). A sufficient condition which guarantees convergence to a viscosity solution is that the numerical scheme be l_∞ stable, consistent in the viscosity sense and monotone [8]. We analyze these properties for the hybrid numerical scheme.

We rewrite (4.3) in a compact form for notational convenience. We first define the following

$$\mathbf{z} = (x,\nu,\xi,\tau), \quad DU(\mathbf{z}) = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial \nu}, \frac{\partial U}{\partial \xi}, \frac{\partial U}{\partial \tau}\right), \quad D^2U(\mathbf{z}) = \begin{pmatrix} \frac{\partial^2 U}{\partial x^2} & \frac{\partial^2 U}{\partial x\partial \nu} & \frac{\partial^2 U}{\partial x\partial \xi} \\ \frac{\partial^2 U}{\partial x\partial \nu} & \frac{\partial^2 U}{\partial \nu^2} & \frac{\partial^2 U}{\partial \nu\partial \xi} \\ \frac{\partial^2 U}{\partial x\partial \xi} & \frac{\partial^2 U}{\partial \nu\partial \xi} & \frac{\partial^2 U}{\partial \xi^2} \end{pmatrix}.$$

The equation (4.3) is rewritten on the localized domain $\Omega \cup \Omega_{out}$ as

$$FU \equiv F\left(\mathbf{z}, U\left(\mathbf{z}\right), DU\left(\mathbf{z}\right), D^2U\left(\mathbf{z}\right)\right) = 0, \tag{4.50}$$

where FU is defined as

$$FU = \begin{cases} F_{in}U, & \Omega_{in} = \Omega_f \cup \Omega_w, \\ F_{\tau_0}, & \Omega_{\tau_0}, \\ F_{max}, & \Omega_{x_{\max}} \cup \Omega_{out}, \end{cases}$$

where

$$\begin{aligned} F_{in}U \equiv F_{in}\left(\mathbf{z}, U\left(\mathbf{z}\right), DU\left(\mathbf{z}\right), D^2U\left(\mathbf{z}\right)\right) &= U_\tau - \mathcal{L}U, && \mathbf{z} \in \Omega_{in} = \Omega_f \cup \Omega_w, \\ F_{\tau_0}U \equiv F_{\tau_0}\left(\mathbf{z}, U\left(\mathbf{z}\right)\right) &= U - U^*\left(x\right), && \mathbf{z} \in \Omega_{\tau_0}, \\ F_{max}U \equiv F_{max}\left(\mathbf{z}, U\left(\mathbf{z}\right)\right) &= U - U\left(x, \nu, \xi, \tau\right), && \mathbf{z} \in \Omega_{x_{\max}} \cup \Omega_{out}. \end{aligned}$$

For the definition of viscosity solution of (4.50), we refer to Definition 3.3.2 with a locally bounded function $U : \Omega \to \mathbf{R}$ and the test functions $\phi(\mathbf{z}) \in C^\infty(\Omega)$.

4.4.1 Consistency

For convenience, we rewrite (4.23) and (4.32) in an equivalent form. Let $\mathcal{G}(.)$ be the discrete approximation of F_{in} for $\mathbf{z} \in \Omega_{in}$ and $\mathbf{z}_{i,j,k}^{n+1} = (x_i, \nu_j, \xi_k, \tau^{n+1})$. For the region $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_f$, we rewrite (4.23) as

$$\mathcal{G}\left(h, \mathbf{z}_{i,j,k}^{n+1}, U_{i,j,k}^{n+1}, \left\{U_{i',j',k'}^{n+1}\right\}_{\substack{i' \neq i \\ or\ j' \neq j, \\ or\ k' \neq k}}, \left\{U_{i',j',k'}^{n}\right\}\right) = \frac{U_{i,j,k}^{n+1} - U_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_f U_{i,j,k}^{n+1} = 0. \tag{4.51}$$

For $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_w$, we have from (4.32),

$$\mathcal{G}\left(h, \mathbf{z}_{i,j,k}^{n+1}, U_{i,j,k}^{n+1}, \left\{U_{i',j',k'}^{n+1}\right\}_{\substack{i' \neq i \\ or\ j' \neq j, \\ or\ k' \neq k}}, \left\{U_{i',j',k'}^{n}\right\}\right) = \frac{U_{i,j,k}^{n+1} - U_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_w U_{i,j,k}^{n+1} = 0. \tag{4.52}$$

Finally, we have

$$\mathcal{G}(.) = 0 = \begin{cases} U\left(x_i, \nu_j, \xi_k, 0\right) - U^*\left(x_i, \nu_j, \xi_k\right), & \mathbf{z}_{i,j,k}^{n+1} \in \Omega_{\tau_0}, \\ U\left(x_i, \nu_j, \xi_k, \tau^{n+1}\right) - U\left(x_i, \nu_j, \xi_k, \tau^{n+1}\right), & \mathbf{z}_{i,j,k}^{n+1} \in \Omega_{x_{\max}} \cup \Omega_{out}. \end{cases} \tag{4.53}$$

Definition 4.4.1. *(Consistency) For any C^∞ function $\phi(x,\nu,\xi,\tau)$ in $\Omega\cup\Omega_{out}$ with $\phi^{n+1}_{i,j,k} = \phi\left(\mathbf{z}^{n+1}_{i,j,k}\right) = \phi\left(x_i,\nu_j,\xi_k,\tau^{n+1}\right)$, the numerical scheme $\mathcal{G}(.)$ is consistent in the viscosity sense, if, $\forall\hat{\mathbf{z}} = \left(\hat{x},\hat{\nu},\hat{\xi},\hat{\tau}\right)$ with $\mathbf{z}^{n+1}_{i,j,k} = \left(x_i,\nu_j,\xi_k,\tau^{n+1}\right)$, the following holds*

$$\limsup_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{z}^{n+1}_{i,j,k}\to\hat{\mathbf{z}}}} \mathcal{G}\left(h,\mathbf{z}^{n+1}_{i,j,k},\phi^{n+1}_{i,j,k}+\psi,\left\{\phi^{n+1}_{i',j',k'}+\psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j,\\ or\ k'\neq k}},\left\{\phi^{n}_{i',j',k'}+\psi\right\}\right) \leq F^*\left(\hat{\mathbf{z}},\phi(\hat{\mathbf{z}}),D\phi(\hat{\mathbf{z}}),D^2\phi(\hat{\mathbf{z}})\right),$$

and

$$\liminf_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{z}^{n+1}_{i,j,k}\to\hat{\mathbf{z}}}} \mathcal{G}\left(h,\mathbf{z}^{n+1}_{i,j,k},\phi^{n+1}_{i,j,k}+\psi,\left\{\phi^{n+1}_{i',j',k'}+\psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j,\\ or\ k'\neq k}},\left\{\phi^{n}_{i',j',k'}+\psi\right\}\right) \geq F^*\left(\hat{\mathbf{z}},\phi(\hat{\mathbf{z}}),D\phi(\hat{\mathbf{z}}),D^2\phi(\hat{\mathbf{z}})\right),$$

Lemma 4.4.1. *(Local consistency). Suppose the mesh discretization parameter h defined in (4.21). Then for any C^∞ function $\phi(x,\nu,\xi,\tau)$ in $\Omega\cup\Omega_{out}$, with $\phi^{n+1}_{i,j,k} = \phi\left(x_i,\nu_j,\xi_k,\tau^{n+1}\right) = \phi\left(\mathbf{z}^{n+1}_{i,j,k}\right)$, and for h, ψ sufficiently small, ψ a constant, we have that*

$$\begin{aligned}
&\mathcal{G}\left(h,\mathbf{z}^{n+1}_{i,j,k},\phi^{n+1}_{i,j,k}+\psi,\left\{\phi^{n+1}_{i',j',k'}+\psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j\\ or\ k'\neq k}},\left\{\phi^{n}_{i',j',k'}+\psi\right\}\right)\\
&\quad= \begin{cases} F_{in}\phi^{n+1}_{i,j,k} + O(h^{\frac{1}{4}}) + O(\psi), & \mathbf{z}^{n+1}_{i,j,k}\in\Omega_f\\ F_{in}\phi^{n+1}_{i,j,k} + O(h^{\frac{\beta}{2}}) + O(\psi), & \mathbf{z}^{n+1}_{i,j,k}\in\Omega_w\\ F_{\tau_0}\phi^{n+1}_{i,j,k} + O(\psi), & \mathbf{z}^{n+1}_{i,j,k}\in\Omega_{\tau_0},\\ F_{max}\phi^{n+1}_{i,j,k} + O(\psi), & \mathbf{z}^{n+1}_{i,j,k}\in\Omega_{up}\cup\Omega_{out}. \end{cases}
\end{aligned} \tag{4.54}$$

Proof. Let

$$\begin{aligned}
\mathcal{L}\phi^{n+1}_{i,j,k} &\equiv \mathcal{L}\phi\left(x_i,\nu_j,\xi_k,\tau^{n+1}\right),\\
(\phi_\tau)^{n+1}_{i,j,k} &\equiv \phi_\tau\left(x_i,\nu_j,\xi_k,\tau^{n+1}\right).
\end{aligned}$$

For notational compactness, we define

$$\Omega_b = \Omega_{b_2}\cup\Omega_{b_3}\cup\Omega_{b_4}\cup\Omega_{b_5}\cup\Omega_{b_6}.$$

For the region, $\mathbf{z}_{i,j,k}^{n+1} \in (\Omega_{in} \setminus \Omega_b \setminus \Omega_{x_{\min}} \setminus \Omega_{\xi_0} \setminus \Omega_{\nu_{\max}}) \cap \Omega_f$, $\mathcal{L}_f\phi_{i,j,k}^{n+1}$ given by (4.22) is a locally consistent discretization of the linear operator $\mathcal{L}$ (4.4), i.e., we have from Taylor series

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h). \tag{4.55}$$

In the region, $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_2} \cap \Omega_f$, an approximation of $O(h^\gamma)$, $\gamma = \min(\frac{1}{2}, \frac{\alpha}{2}, \frac{\beta}{2})$ is used. Hence we have

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^\gamma) + O(h). \tag{4.56}$$

For the region, $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_3} \cap \Omega_f$, an $O(h^{\frac{1}{4}})$ approximation (4.34) to (4.4) is used. Therefore,

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{1}{4}}) + O(h). \tag{4.57}$$

For the region, $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_4} \cap \Omega_f$, an $O(h^{\frac{\beta}{2}})$ approximation ((4.34), (4.35)) to (4.4) is used, resulting in

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}) + O(h). \tag{4.58}$$

For the region, $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_5} \cap \Omega_f$, an $O(h^{\frac{\beta}{2}})$ approximation ((4.33), (4.34)) to (4.4) is used leading to

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}) + O(h). \tag{4.59}$$

For the region, $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_6} \cap \Omega_f$, an $O(\sqrt{h})$ approximation (4.34) to (4.4) is used and hence

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(\sqrt{h}) + O(h). \tag{4.60}$$

Similarly for region $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{x_{\min}} \cap \Omega_f$, we have from (4.7) and (4.38),

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^\gamma) + O(h), \tag{4.61}$$

where $\gamma = \min(\frac{\alpha}{2}, \frac{\beta}{2})$. For the region $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{\xi_0} \cap \Omega_f$, taylor series on (4.9) and (4.40) results in

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(\sqrt{h}) + O(h). \tag{4.62}$$

For the region $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{\nu_{\max}} \cap \Omega_f$, taylor series on (4.14) and (4.42) results in

$$\mathcal{L}_f\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}) + O(h). \tag{4.63}$$

From (4.55)-(4.63), we have for $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_f$,

$$\mathcal{L}_f^{in}\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{1}{4}}).$$

We also have

$$\mathcal{L}_f\left(\phi_{i,j,k}^{n+1} + \psi\right) \;=\; \mathcal{L}_f\phi_{i,j,k}^{n+1} - r\psi.$$

Recall that $\Delta\tau = \mathcal{B}_4 h$ from (4.21) and hence we have

$$\frac{\phi_{i,j,k}^{n+1} - \phi_{i,j,k}^{n}}{\Delta\tau} \;=\; (\phi_\tau)_{i,j,k}^{n+1} + O(h).$$

Therefore, from (4.51), we have

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{z}_{i,j,k}^{n+1}, \phi_{i,j,k}^{n+1} + \psi, \left\{\phi_{a,b,c}^{n+1} + \psi\right\}_{\substack{a\neq i \\ orb\neq j \\ orc\neq k}}, \left\{\phi_{e,f,g}^{n} + \psi\right\}\right) \\
=\;& \frac{\phi_{i,j,k}^{n+1} - \phi_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_f\left(\phi_{i,j,k}^{n+1} + \psi\right) \\
=\;& \frac{\phi_{i,j,k}^{n+1} - \phi_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_f\left(\phi_{i,j,k}^{n+1}\right) + O(\psi) \\
=\;& (\phi_\tau)_{i,j,k}^{n+1} - \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{1}{4}}) + O(\psi) \\
=\;& F_{in}\phi_{i,j,k}^{n+1} + O(h^{\frac{1}{4}}) + O(\psi), \quad \mathbf{z}_{i,j,k}^{n+1} \in \Omega_f.
\end{aligned}$$

For the regions where $\mathbf{z}_{i,j,k}^{n+1} \in (\Omega_{in}/\Omega_b) \cap \Omega_w$, $\mathcal{L}_w\phi_{i,j,k}^{n+1}$ given by (4.31) is also locally consistent, i.e.,

$$\mathcal{L}_w^{in}\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h). \tag{4.64}$$

In $\mathcal{L}_w$, the first order derivatives are discretized using either forward or backward differencing and hence they are consistent to $O(h)$. For the second order derivatives, we provide a short proof for consistency. Using a smooth test function ϕ and the linear interpolation

operator $\mathcal{J}_h$, we have

$$\begin{aligned}
& \frac{\frac{\mathcal{J}_h\phi^{n+1}\left(X_{i,j,k}-\sqrt{h}\left(R_{i,j,k}\right)_{\hat{l}}\right)-\phi^{n+1}\left(X_{i,j,k}\right)}{\sqrt{h}}+\frac{\mathcal{J}_h\phi^{n+1}\left(X_{i,j,k}+\sqrt{h}\left(R_{i,j,k}\right)_{\hat{l}}\right)-\phi^{n+1}\left(X_{i,j,k}\right)}{\sqrt{h}}}{\frac{\sqrt{h}+\sqrt{h}}{2}} \\
= & \frac{\frac{\phi^{n+1}\left(Y_{i,j,k}-\sqrt{h}\mathbf{e}_{\hat{l}}\right)-\phi^{n+1}\left(Y_{i,j,k}\right)+O(h^2)}{\sqrt{h}}+\frac{\phi^{n+1}\left(Y_{i,j,k}+\sqrt{h}\mathbf{e}_{\hat{l}}\right)-\phi^{n+1}\left(Y_{i,j,k}\right)+O(h^2)}{\sqrt{h}}}{\frac{\sqrt{h}+\sqrt{h}}{2}} \\
= & \frac{\frac{\phi^{n+1}\left(Y_{i,j,k}-\sqrt{h}\mathbf{e}_{\hat{l}}\right)-\phi^{n+1}\left(Y_{i,j,k}\right)}{\sqrt{h}}+\frac{\phi^{n+1}\left(Y_{i,j,k}+\sqrt{h}\mathbf{e}_{\hat{l}}\right)-\phi^{n+1}\left(Y_{i,j,k}\right)}{\sqrt{h}}}{\frac{\sqrt{h}+\sqrt{h}}{2}}+O(h) \\
= & \frac{\partial^2\phi}{\partial z_{\hat{l}}^2}+O(h), \quad \hat{l}=1,2,3,
\end{aligned}$$

which follows from Taylor series expansion and that the error of linear interpolation for a smooth function ϕ is $O(h^2)$.

For the regions where $\mathbf{z}_{i,j,k}^{n+1} \in (\Omega_{b_2} \cup \Omega_{b_4}) \cap \Omega_w$, $O(\sqrt{h})$ approximations ((4.35), (4.36)) to (4.4) are solved. Hence $\mathcal{L}_w\phi_{i,j,k}^{n+1}$ given by (4.31) is also locally consistent, i.e.,

$$\mathcal{L}_w\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h) + O(\sqrt{h}). \tag{4.65}$$

Similarly for region $\mathbf{z}_{i,j,k}^{n+1} \in \Omega_{b_5} \cap \Omega_w$, $O(h^{\frac{\beta}{2}})$ approximation (4.33) is solved and hence $\mathcal{L}_w\phi_{i,j,k}^{n+1}$ given by

$$\mathcal{L}_w\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h) + O(h^{\frac{\beta}{2}}). \tag{4.66}$$

From (4.64) - (4.66), we have

$$\mathcal{L}_w\phi_{i,j,k}^{n+1} = \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}). \tag{4.67}$$

We also have

$$\begin{aligned}
\mathcal{L}_w\left(\phi_{i,j,k}^{n+1}+\psi\right) &= \mathcal{L}_w\phi_{i,j,k}^{n+1} - r\psi, \\
\frac{\phi_{i,j,k}^{n+1}-\phi_{i,j,k}^{n}}{\Delta\tau} &= (\phi_\tau)_{i,j,k}^{n+1} + O(h).
\end{aligned}$$

Therefore, we have from (4.52)

$$
\begin{aligned}
& \mathcal{G}\left(h, \mathbf{z}_{i,j,k}^{n+1}, \phi_{i,j,k}^{n+1}+\psi, \left\{\phi_{a,b,c}^{n+1}+\psi\right\}_{\substack{a\neq i \\ or b\neq j \\ or c\neq k}}, \left\{\phi_{e,f,g}^{n}+\psi\right\}\right) \\
= & \frac{\phi_{i,j,k}^{n+1}-\phi_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_w\left(\phi_{i,j,k}^{n+1}+\psi\right) \\
= & \frac{\phi_{i,j,k}^{n+1}-\phi_{i,j,k}^{n}}{\Delta\tau} - \mathcal{L}_w\left(\phi_{i,j,k}^{n+1}\right) + O(\psi) \\
= & (\phi_\tau)_{i,j,k}^{n+1} - \mathcal{L}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}) + O(\psi) \\
= & F_{in}\phi_{i,j,k}^{n+1} + O(h^{\frac{\beta}{2}}) + O(\psi), \quad \mathbf{z}_{i,j,k}^{n+1} \in \Omega_w.
\end{aligned}
$$

The remaining cases in (4.54) can be proven in a similar way and we omit the details. □

Lemma 4.4.2. *(Consistency) Provided that all the conditions in Lemma 4.4.1 are satisfied, then the scheme ((4.51)-(4.53)) is consistent according to the Definition 4.4.1.*

Proof. The proof is straightforward and can be constructed from Lemma 4.4.1 and the corresponding steps in [58]. □

4.4.2 Stability

Definition 4.4.2. *(M-matrix) A real $n \times n$ matrix Q with $Q_{i,j} \leq 0$ for all $i \neq j$ is an M-matrix if Q is nonsingular and $Q^{-1} \geq 0$. (See Definition 3.22, Page 91 [102]).*

Corollary 4.4.1. *If a matrix Q has elements $Q_{ii} > 0$ and $Q_{ij} \leq 0$ for $i \neq j$ and every row sum is non-negative with at least one row sum positive in each connected part of Q, then $Q^{-1} \geq 0$. (See Corollary 3.20, Page 91 [102]).*

Remark 4.4.1. *From Definition 4.4.2 and Corollary 4.4.1, a sufficient condition for a matrix Q to be an M-matrix is that Q has positive diagonals, non-positive off-diagonals, and is diagonally dominant.*

Lemma 4.4.3. *Given that we use the discretization operators $\mathcal{L}_f$ in the domain Ω_f and a linear interpolation operator $\mathcal{J}_h$ is used in (4.31), (4.37), (4.39) and (4.41), we then have that the matrix $[\mathbf{I} - \Delta\tau\mathbf{L}]$ (4.49) is an M-matrix.*

Proof. From (4.43), (4.45) and (4.47) (include more), it is clear that $[\mathbf{I} - \Delta\tau\mathbf{L}]$ has positive diagonals, non-positive off-diagonals and the l-th row sum of the matrix is given by

$$\sum_k [\mathbf{I} - \Delta\tau\mathbf{L}]_{l,m} = \begin{cases} 1 + r\Delta\tau, & i = 0, \ldots, n_x - 2, \quad j = 0, \ldots n_\nu - 1, \quad k = 0, \ldots, n_\xi - 1 \\ 1, & i = n_x - 1, \quad j = 0, \ldots n_\nu - 1, \quad = 0, \ldots, n_\xi - 1, \end{cases}$$

where $l = k \cdot n_x \cdot n_\nu + j \cdot n_x + i$. Thus the matrix $[\mathbf{I} - \Delta\tau\mathbf{L}]$ is diagonally dominant. □

Lemma 4.4.4. *(Stability) Given the conditions for Lemma 4.4.3 are satisfied, the discretization (4.49), equivalently ((4.51)-(4.53)), is unconditionally l_∞ stable, as the mesh parameter (4.21) $h \to 0$, satisfying*

$$||U^n||_\infty \leq \max\left(||U^0||_\infty, \mathcal{B}_5\right),$$

where $\mathcal{B}_5 = \max_n ||F^n||_\infty$, where F^n is determined by the asymptotic boundary condition (4.5).

Proof. The result follows from Lemma 4.4.3 and a maximum analysis [31]. □

4.4.3 Monotonicity

Definition 4.4.3. (Monotonicity) A discrete scheme is monotone if for all $\mathcal{R}^n_{i,j,k} \geq \mathcal{S}^n_{i,j,k}$, $\forall i, j, k, n$

$$\mathcal{G}\left(h, \mathbf{z}^{n+1}_{i,j,k}, U^{n+1}_{i,j,k}, \{\mathcal{R}_{i',j',k'}\}_{\substack{i'\neq i \\ \text{or } j'\neq j \\ \text{or } k'\neq k}}, \left\{\mathcal{R}^n_{i',j',k'}\right\}\right) \leq \mathcal{G}\left(h, \mathbf{z}^{n+1}_{i,j,k}, U^{n+1}_{i,j,k}, \{\mathcal{S}_{i',j',k'}\}_{\substack{i'\neq i \\ \text{or } j'\neq j \\ \text{or } k'\neq k}}, \left\{\mathcal{S}^n_{i',j',k'}\right\}\right).$$

Lemma 4.4.5. *(Monotonicity) If the scheme ((4.51)-(4.53)) satisfies the conditions required for Lemma 4.4.4, then the discretization is monotone as per the Definition A.1.1.*

Proof. Since our discretization (4.49) is a positive coefficient scheme, monotonicity follows using similar proof given in [37]. □

4.4.4 Convergence

Theorem 4.4.1. *(Convergence) Given that the discretization* (4.51)-(4.53) *satisfies all the conditions in Lemma 4.4.2, 4.4.4 and 4.4.5 and that Assumption 3.3.1 holds for* (4.50)*, then the numerical scheme converges to the unique continuous viscosity solution of* (4.50)*.*

Proof. Since the scheme is monotone, consistent and l_∞ stable, the convergence follows from the results in [11]. □

4.5 Multigrid

In this section, we develop multigrid methods for the three-dimensional PDE (4.3). We first investigate multigrid methods for fixed stencil discretization of (4.3). We then apply the idea to multigrid methods for hybrid stencil discretization. FAS scheme directly handles the nonlinearity of the underlying PDE. When applied to linear problems, FAS scheme is essentially the standard multigrid method for linear PDEs. The linear system under a fixed stencil discretization is given by

$$A^h U^{n+1,h} = B^h, \tag{4.68}$$

where $A^h = [\mathbf{I} - \Delta\tau\mathbf{L}]$ with entries of $\mathbf{L}$ given by (4.45) for $l = 0, \ldots, N-1$. The right hand side B^h is given by

$$B^h = U^n + \mathbf{F}^{n+1} - \mathbf{F}^n,$$

where $\mathbf{F}$ is given by (4.44).

Efficient multigrid methods for higher dimensional PDEs have been developed in the literature [100]. However, these methods require implementations that are specific to the underlying problem and since they have not been investigated for the three-dimensional PDE (4.4) in the past, we propose to develop multigrid methods to efficiently solve (4.4).

A known issue for the application of multigrid methods to higher dimensional PDEs is the inherent anisotropy making the use of point-wise relaxation ineffective [25, 100]. There are two possible ways to fix this issue. First, it is observed that errors become smooth if strongly connected unknowns are updated collectively [100]. A standard coarse grid, where the number of coarse grid points $N_H = N/8$ for three-dimensional problem, can then be used for coarsening. The relaxation-type smoothers for anisotropic problems are essentially block iterations in which each block of unknowns correspond to a group of

strongly connected unknowns. Gauss-Siedel or Jacobi type line or plane relaxations are particularly efficient smoothers if the anisotropy is aligned with the grid along a line or a plane, respectively. For (4.4), a line relaxation leads to tridiagonal systems that are easily solved. On the other hand, in a plane relaxation, a discrete two-dimensional problem is solved for each plane, which are expensive to solve. A computationally efficient way to perform a plane relaxation is to use two-dimensional multigrid in a plane [40]. For anisotropic problems with varying coefficients, a standard approach is to use alternating plane relaxations [100] since the direction of coupling between unknowns is not clear. But, this approach is computationally expensive.

Second, we can keep the point-wise relaxation and use semicoarsening strategies instead of block relaxations and standard coarsening. Here again, semicoarsening should be performed along the direction in which the unknowns are strongly coupled. If semicoarsening is performed along one dimension, then $N_H = N/2$ and when it is performed along two dimensions, then $N_H = N/4$. In either case the complexity of the FAS V-cycle is $O(N)$.

We use block relaxation and standard coarsening in our multigrid method because this approach is relatively easy to implement and analyze. We use the standard trilinear interpolation and a fully weighted restriction for the grid transfer operators $\mathcal{P}$ and $(\mathcal{R}_u, \mathcal{R}_r)$, respectively. The smoother is presented in the following section.

4.5.1 Smoother

The three-dimensional PDE (4.3) has varying coefficients, which may lead to anisotropies in different directions. A robust smoother for such cases is alternating plane relaxation [100]. Each smoothing step consists of three plane relaxations ((x,ν), (ν,ξ) and (x,ξ) plane relaxation), which is computationally expensive. If 2D multigrid is used as the plane solver, alternating line relaxation within each plane solver is required to guarantee good smoothing properties for different choices of the coefficients in (4.3).

In order to avoid this overhead in the computational expense, we analyze (4.68) for both double Heston and double CEV models, with a fixed set of parameters, to find if there is any strong coupling in unknowns. For double CEV, the model parameters are: $\alpha = 0.94,\ \beta = 0.94,\ C_1 = 5.5,\ C_2 = 2.873,\ C_3 = 0.1,\ C_4 = 0.078,\ C_5 = 0.302,\ \rho_{12} = -0.992,\ \rho_{13} = -0.615,\ \rho_{23} = 0.59,\ r = 0.01$. For any node (x_i, ν_j, ξ_k), it is observed that the unknowns are relatively strongly coupled with only two other neighboring nodes: (x_i, ν_{j-1}, ξ_k) and (x_i, ν_{j+1}, ξ_k) for different h. In other words, the elements of the difference

operator are related by

$$|A^h_{l,l}| > \max\{|A^h_{l,l-n_x}|, |A^h_{l,l+n_x}|\} > \max\big\{|A^h_{l,l-n_xn_\nu-n_x}|, |A^h_{l,l-n_xn_\nu}|, |A^h_{l,l-n_xn_\nu+1}|, |A^h_{l,l-n_x+1}|, |A^h_{l,l-1}|, |A^h_{l,l+1}|, |A^h_{l,l+n_x-1}|, |A^h_{l,l+n_xn_\nu-1}|, |A^h_{l,l+n_xn_\nu}|, |A^h_{l,l+n_xn_\nu+n_x}|\big\}. \quad (4.69)$$

To illustrate (4.69), consider the difference operator at $(x_{27}, \nu_{34}, \xi_1)$ in a $64 \times 64 \times 64$ grid, which is given by

$$\left[\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.2954 & -0.0142 \\ 0 & -0.2831 & 0 \end{bmatrix} \begin{bmatrix} 0 & -114.15 & 0 \\ 5.525 & 234.48 & 5.577 \\ 0 & -118.73 & -5.846 \end{bmatrix} \begin{bmatrix} 0 & -0.283 & 0 \\ -0.0142 & 0.2854 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right]. \quad (4.70)$$

The elements of (4.70) clearly agree with (4.69). We observe similar behavior elsewhere in the grid for different grid sizes. Similar results are obtained for double Heston model with $\alpha = 0.5,\ \beta = 0.5,\ C_1 = 12,\ C_2 = 0.7,\ C_3 = 0.34,\ C_4 = 0.0421,\ C_5 = 0.14,\ \rho_{12} = -0.9,\ \rho_{13} = -0.7,\ \rho_{23} = 0.59,\ r = 0.01$. Fortunately, strong coupling is along a line aligned with the grid in the ν-dimension. Hence line smoothing in the ν direction will efficiently smooth the error.

We now derive a Gauss-Siedel type line relaxation scheme for (4.68). Rearranging (4.68) at the grid point (x_i, ν_j, ξ_k), we have

$$\begin{aligned} U^{n+1}_l = \frac{-1}{A^h_{l,l}} \big(& A^h_{l,l-1}U^{n+1}_{l-1} + A^h_{l,l+1}U^{n+1}_{l+1} + A^h_{l,l-n_x}U^{n+1}_{l-n_x} + A^h_{l,l+n_x}U^{n+1}_{l+n_x} + A^h_{l,l-n_xn_\nu}U^{n+1}_{l-n_xn_\nu} \\ & + A^h_{l,l+n_xn_\nu}U^{n+1}_{l+n_xn_\nu} + \mathbb{1}_{\rho_{12}\geq 0}\ \big(A^h_{l,l+n_x+1}U^{n+1}_{l+n_x+1} + A^h_{l,l-n_x-1}U^{n+1}_{l-n_x-1}\big) \\ & + \mathbb{1}_{\rho_{12}<0}\ \big(A^h_{l,l-n_x+1}U^{n+1}_{l-n_x+1} + A^h_{l,l+n_x-1}U^{n+1}_{l+n_x-1}\big) \\ & + \mathbb{1}_{\rho_{13}\geq 0}\ \big(A^h_{l,l+n_xn_\nu+1}U^{n+1}_{l+n_xn_\nu+1} + A^h_{l,l-n_xn_\nu-1}U^{n+1}_{l-n_xn_\nu-1}\big) \\ & + \mathbb{1}_{\rho_{13}<0}\ \big(A^h_{l,l-n_xn_\nu+1}U^{n+1}_{l-n_xn_\nu+1} + A^h_{l,l+n_xn_\nu-1}U^{n+1}_{l+n_xn_\nu-1}\big) \\ & + \mathbb{1}_{\rho_{23}\geq 0}\ \big(A^h_{l,l+n_xn_\nu+n_x}U^{n+1}_{l+n_xn_\nu+n_x} + A^h_{l,l-n_xn_\nu-n_x}U^{n+1}_{l-n_xn_\nu-n_x}\big) \\ & + \mathbb{1}_{\rho_{23}<0}\ \big(A^h_{l,l-n_xn_\nu+n_x}U^{n+1}_{l-n_xn_\nu+n_x} + A^h_{l,l+n_xn_\nu-n_x}U^{n+1}_{l+n_xn_\nu-n_x}\big) - B^h_l\big). \end{aligned} \quad (4.71)$$

Let $\widetilde{U}^m$ be the m-th estimate for U^{n+1}. A point-wise Gauss-Siedel relaxation scheme

can then be derived from (4.71) as

$$\begin{aligned}
\tilde{U}_l^{m+1} &= \frac{-1}{A_{l,l}^h}\Big(A_{l,l-1}^h\tilde{U}_{l-1}^{m+1} + A_{l,l+1}^h\tilde{U}_{l+1}^m + A_{l,l-n_x}^h\tilde{U}_{l-n_x}^m + A_{l,l+n_x}^h\tilde{U}_{l+n_x}^m + A_{l,l-n_xn_\nu}^h\tilde{U}_{l-n_xn_\nu}^{m+1} \\
&+ A_{l,l+n_xn_\nu}^h\tilde{U}_{l+n_xn_\nu}^m + \mathbb{1}_{\rho_{12}\geq 0}\left(A_{l,l+n_x+1}^h\tilde{U}_{l+n_x+1}^m + A_{l,l-n_x-1}^h\tilde{U}_{l-n_x-1}^{m+1}\right) \\
&+ \mathbb{1}_{\rho_{12}<0}\left(A_{l,l-n_x+1}^h\tilde{U}_{l-n_x+1}^m + A_{l,l+n_x-1}^h\tilde{U}_{l+n_x-1}^{m+1}\right) \\
&+ \mathbb{1}_{\rho_{13}\geq 0}\left(A_{l,l+n_xn_\nu+1}^h\tilde{U}_{l+n_xn_\nu+1}^m + A_{l,l-n_xn_\nu-1}^h\tilde{U}_{l-n_xn_\nu-1}^{m+1}\right) \\
&+ \mathbb{1}_{\rho_{13}<0}\left(A_{l,l-n_xn_\nu+1}^h\tilde{U}_{l-n_xn_\nu+1}^{m+1} + A_{l,l+n_xn_\nu-1}^h\tilde{U}_{l+n_xn_\nu-1}^m\right) \\
&+ \mathbb{1}_{\rho_{23}\geq 0}\left(A_{l,l+n_xn_\nu+n_x}^h\tilde{U}_{l+n_xn_\nu+n_x}^m + A_{l,l-n_xn_\nu-n_x}^h\tilde{U}_{l-n_xn_\nu-n_x}^{m+1}\right) \\
&+ \mathbb{1}_{\rho_{23}<0}\left(A_{l,l-n_xn_\nu+n_x}^h\tilde{U}_{l-n_xn_\nu+n_x}^{m+1} + A_{l,l+n_xn_\nu-n_x}^h\tilde{U}_{l+n_xn_\nu-n_x}^m\right) - B_l^h\Big). \qquad (4.72)
\end{aligned}$$

A ν-line Gauss-Siedel relaxation is obtained from (4.72) by collectively updating all the unknowns along the ν dimension for a fixed x_i and ξ_k. Using a stencil notation, ν-line Gauss-Siedel relaxation for (4.23) can be written as

$$\begin{aligned}
&\left[A_{l,l-n_x}^h \quad A_{l,l}^h \quad A_{l,l+n_x}^h\right]\tilde{U}^{m+1} \\
&= B_l^h - A_{l,l-1}^h\tilde{U}_{l-1}^{m+1} - A_{l,l+1}^h\tilde{U}_{l+1}^m - A_{l,l-n_xn_\nu}^h\tilde{U}_{l-n_xn_\nu}^{m+1} - A_{l,l+n_xn_\nu}^h\tilde{U}_{l+n_xn_\nu}^m \\
&- \mathbb{1}_{\rho_{12}>0}\left(A_{l,l+n_x+1}^h\tilde{U}_{l+n_x+1}^m + A_{l,l-n_x-1}^h\tilde{U}_{l-n_x-1}^{m+1}\right) \\
&- \mathbb{1}_{\rho_{12}<0}\left(A_{l,l-n_x+1}^h\tilde{U}_{l-n_x+1}^m + A_{l,l+n_x-1}^h\tilde{U}_{l+n_x-1}^{m+1}\right) \\
&- \mathbb{1}_{\rho_{13}>0}\left(A_{l,l+n_xn_\nu+1}^h\tilde{U}_{l+n_xn_\nu+1}^m + A_{l,l-n_xn_\nu-1}^h\tilde{U}_{l-n_xn_\nu-1}^{m+1}\right) \\
&- \mathbb{1}_{\rho_{13}<0}\left(A_{l,l-n_xn_\nu+1}^h\tilde{U}_{l-n_xn_\nu+1}^{m+1} + A_{l,l+n_xn_\nu-1}^h\tilde{U}_{l+n_xn_\nu-1}^m\right) \\
&- \mathbb{1}_{\rho_{23}>0}\left(A_{l,l+n_xn_\nu+n_x}^h\tilde{U}_{l+n_xn_\nu+n_x}^m + \mathbf{L}_{l,l-n_xn_\nu-n_x}\tilde{U}_{l-n_xn_\nu-n_x}^{m+1}\right) \\
&- \mathbb{1}_{\rho_{23}<0}\left(A_{l,l-n_xn_\nu+n_x}^h\tilde{U}_{l-n_xn_\nu+n_x}^{m+1} + A_{l,l+n_xn_\nu-n_x}^h\tilde{U}_{l+n_xn_\nu-n_x}^m\right). \qquad (4.73)
\end{aligned}$$

A ν-line Gauss-Siedel relaxation scheme for the linear system (4.49) resulting from hybrid stencil discretization is derived using the same procedure used for the fixed stencil system (4.68) above. We omit the details here.

The ν-line Gauss-Siedel relaxation (4.73) involves solving a tridiagonal system of equations. Figure 4.6 demonstrates the order in which unknowns are relaxed for ν-line Gauss-Siedel relaxation.

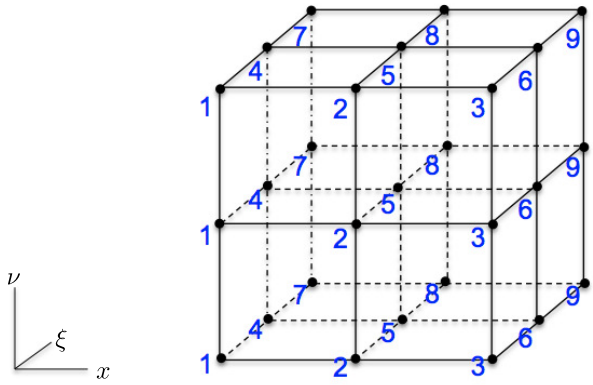


Figure 4.6: Order in which unknowns are collectively solved for ν-line Gauss-Siedel relaxation.

4.6 Numerical Results

We numerically price European call options under a two-factor stochastic volatility model. In particular, we consider the double Heston and double CEV models, whose parameters are calibrated to fit VIX options dated 03-April 2007 [41, 43] as given in Table 4.4. The payoff of the European call option is $\max(S - K, 0)$, where K is the strike price.

The convergence of the hybrid stencil discretization proposed in Section 4.3.3 is tested on a series of uniformly refined grids, starting with a $40 \times 31 \times 30$ non-uniform grid. The initial time step is 0.1. For each grid refinement, the time step and the grid size are halved. ILU preconditioned BiCGStab [92] with a convergence tolerance of 10^{-6} is used as the solver for the linear system. The convergence results for the double Heston and double CEV model at $U(S, \nu, \xi, t) = U(93.55, 0.01, 0.02, 0)$ are given in Tables 4.5 and 4.6, respectively. Due to the absence of analytical solution, we use Monte Carlo simulations as a benchmark to compare solutions. We observe that the convergence rate of the hybrid stencil discretization is between first and second order for both the model parameters. Tables 4.5 and 4.6 also show the number of iterations taken by BiCGStab in the first timestep. The convergence is similar in the rest of the timesteps and hence we only present the iterations from the first timestep.

We numerically test the performance of the multigrid method for the fixed stencil discretization (4.68). A series of uniformly refined grids, starting with a $33 \times 33 \times 33$ uniform grid are used. The initial time step is 0.1. The time step and the grid size are halved in each successive refinement. Two pre-smoothing and post-smoothing iterations

Table 4.4: Double Heston and double CEV model parameters for 03-April 2007 VIX options.

	Double Heston	Double CEV
α	0.5	0.94
β	0.5	0.94
C_1	12	5.5
C_2	0.7	2.689
C_3	0.34	0.1
C_4	0.0421	0.078
C_5	0.14	0.302
ρ_{12}	-0.9	-0.992
ρ_{13}	-0.7	-0.615
Common Parameters		
ρ_{23}	0.59	
r	0.01	
T	1	
K	93.55	
$x_{\min}$	-5	
$x_{\max}$	7	
$\nu_{\max}$	1.5	
$\xi_{\max}$	1	

Table 4.5: Convergence results of the hybrid stencil for double Heston model under an European call option with parameters given in Table 4.4. $S = 93.55, \nu = 0.01, \xi = 0.02, t = 0$. Difference is the change in value of the solution for successive grids. Ratio is the ratio of the successive difference values. Monte Carlo value: 5.77733, standard error: 0.0068.

Time steps	$n_x \times n_\nu \times n_\xi$	# iterations in first time step	Value	Difference	Ratio
10	$40 \times 31 \times 30$	7	3.48056		
20	$79 \times 61 \times 59$	12	4.83075	1.35019	
40	$157 \times 117 \times 121$	17	5.53985	0.70910	1.9
80	$313 \times 233 \times 241$	28	5.71968	0.17983	3.94

Table 4.6: Convergence results of the hybrid stencil for double CEV model under an European call option with parameters given in Table 4.4. $S = 93.55, \nu = 0.01, \xi = 0.02, t = 0$. Difference is the change in value of the solution for successive grids. Ratio is the ratio of the successive difference values. Monte Carlo value: 5.35409, standard error: 0.00643.

Time steps	$n_x \times n_\nu \times n_\xi$	# iterations in first time step	Value	Difference	Ratio
10	$40 \times 31 \times 30$	7	3.18273		
20	$79 \times 61 \times 59$	10	4.48730	1.30457	
40	$157 \times 117 \times 121$	14	5.24634	0.75903	1.72
80	$313 \times 233 \times 241$	25	5.42414	0.17781	4.27

are used in each multigrid cycle. Multiple coarse grids are used such that the coarsest level has only 9 grid points in each of the three dimensions. The convergence tolerance is 10^{-6}. The number of iterations for multigrid convergence in the first timestep for the double Heston and double CEV models are shown in Table 4.7. The multigrid method converges in small number of iterations for different grid sizes. For a $257 \times 257 \times 257$ grid discretized using a fixed stencil, an ILU preconditioned BiCGStab solver takes 24 and 46 iterations for double Heston and double CEV models, respectively. Whereas, our multigrid method converges in just 4 and 14 iterations as shown in Table 4.7.

Table 4.7: Multigrid convergence for fixed stencil discretization (4.68).

$n_x \times n_\nu \times n_\xi$	$33 \times 33 \times 33$	$65 \times 65 \times 65$	$129 \times 129 \times 129$	$257 \times 257 \times 257$
Double Heston	2	2	3	4
Double CEV	4	7	8	14

We also test the convergence of the multigrid method on the linear system (4.49) resulting from the hybrid stencil discretization. We use two ν-line pre-smoothing and post-smoothing iterations in each multigrid cycle, similar to the multigrid for the fixed stencil. The convergence results for the first timestep for double Heston model are presented in Table 4.8. Note that the percentage of grid points using wide stencil increases as the grid is refined. The convergence results are very similar to the multigrid for double Heston using a fixed stencil given in Table 4.7.

Table 4.8: Multigrid convergence for double Heston using hybrid stencil discretization (4.49).

$n_x \times n_\nu \times n_\xi$	# iterations	% fixed stencil	% wide stencil
$33 \times 33 \times 33$	2	91.65	8.35
$65 \times 65 \times 65$	3	77.63	22.37
$129 \times 129 \times 129$	3	60.93	39.07
$257 \times 257 \times 257$	5	44.63	55.37

4.7 Conclusion

A standard fixed stencil discretization for (4.3) will not satisfy the positive coefficient condition across the domain due to the presence of the cross derivative terms. The positive coefficient condition is essential for monotonicity. We propose a wide stencil discretization based on a local coordinate transformation to eliminate the cross derivative terms. Hence a positive coefficient discretization can be constructed on the wide stencil. The wide stencil discretization in the boundary regions may result in points lying outside the computational domain. It is observed that in these regions the cross derivative terms are small in magnitude, hence we propose to solve approximations to (4.3) by eliminating those cross derivative terms. To take advantage of the computational efficiency and accuracy of the fixed stencil, we use a hybrid stencil in which the fixed stencil is used as much as possible and a wide stencil when the fixed stencil does not satisfy the positive coefficient conditions. We prove that the hybrid numerical scheme converges to the viscosity solution. We perform numerical tests on double Heston and double CEV models to show the convergence of the hybrid stencil discretization. We also numerically test the multigrid solver both on the fixed and hybrid stencil discretization and show that it converges in a small number of iterations for different grid sizes.

Chapter 5

Conclusion and Future Work

5.1 Multigrid Methods

Multigrid methods are very efficient when the components are specifically designed for the underlying physical problem. This is clearly seen in the three different types of PDEs we consider. Hyperbolic PDEs model the wave motion or advective transport. Therefore, it is important to construct upwind based interpolation and restriction techniques. Hyperbolic systems have a mixture of characteristics and hence extending the upwind techniques from the scalar case is not straightforward. We propose to solve a local Riemann problem to perform upwind interpolation for hyperbolic systems. Upwind restriction is performed by splitting the residual into positive and negative components based on flux vector splitting. For two-dimensional hyperbolic systems, we propose a novel coarsening technique and extend the upwind grid transfer techniques. The upwind multigrid methods reduce the number of work units by a factor of 4 to 13 compared to the single grid, while factors of 3 to 9 are reported in the literature.

For HJB and HJBI systems with jump in control, multigrid methods should preserve the consistency of control between the fine and coarse grids. Otherwise, the multigrid method can be slow or may not converge in certain situations. We choose injection for restriction, which preserves the consistency of the control from the fine to the coarse grids. We introduce two new interpolation techniques which efficiently capture the optimal control in the presence of jumps. A two-grid Fourier analysis shows that our multigrid method has a convergence factor 0.12 as $h \to 0$. Numerical tests on dynamic games and computational finance applications show that the multigrid method converges in very small number of iterations irrespective of the grid size.

For the three-dimensional anisotropic PDE from European option pricing under a two-factor stochastic volatility model, we observed that the unknowns are strongly connected along the ν-dimension. Hence, we develop a multigrid method using ν-line relaxation, standard coarsening and grid transfer techniques. Line relaxation results in a tridiagonal linear system and can be easily solved. Numerical results show that the multigrid converges in as low as 5 iterations for a grid size of $257 \times 257 \times 257$ for double Heston model using hybrid stencil discretization.

5.2 Discretization

In general, the solutions for HJB equations are not smooth. Hence, we construct a fully implicit, consistent, unconditionally l_∞ stable and monotone discretization scheme that converges to the viscosity solution for the HJB PDE (3.3) resulting from dynamic Bertrand monopoly problem and for the two-dimensional system of nonlinear HJB PDEs (3.2) from dynamic Bertrand duopoly.

The three-dimensional PDE (4.3) from pricing European options under a two-factor stochastic volatility model is linear and a non-monotone discretization scheme is adequate. However, we still develop an unconditionally monotone scheme that converges to the viscosity solution of (4.3) to develop a framework for nonlinear equations under a two-factor stochastic volatility model, for example pricing American options. The cross derivative terms in (4.3) makes the construction of monotone schemes non-trivial. We propose a wide stencil discretization based on a local coordinate transformation to eliminate the cross derivative terms. Hence a positive coefficient discretization can be constructed on the wide stencil. A hybrid stencil is used to take advantage of the computational efficiency and accuracy of the fixed stencil while still preserving monotonicity. We perform a discretization analysis to prove that the hybrid stencil converges to the viscosity solution. We use a preconditioned Bi-CGSTAB to numerically test the convergence of the hybrid stencil.

5.3 Future Work

Some directions for future research are given below.

- The smoothing property of relaxation methods are essentially based on the symmetry of the discretization matrix. Wide stencil discretization results in a highly non-symmetrical discretization matrix. Further research is required to design the right or reasonable components of multigrid methods for hybrid discretization.
- It would be interesting to test the convergence of hybrid stencil discretization for non-linear PDEs under a two-factor stochastic volatility model such as pricing American options.
- Multigrid methods developed in this thesis are applied to uniformly discretized grids. The use of non-uniform grids is popular in computational finance applications. Existing multigrid methods are largely based on the symmetry of the discretization matrix. Finite difference methods using non-uniform grids result in asymmetric matrices and hence standard multigrid methods do not work well. This is an interesting area for further research.

APPENDICES

Appendix A

Theoretical Analysis for Hyperbolic PDE

A.1 Monotonicity Analysis for the Linear Wave Equation

We present the monotonicity preserving properties of the multigrid time stepping schemes for solving the linear wave equation

$$\begin{aligned} u_t + u_x &= 0, & 0 < x < 1, \\ u(0,t) &= 0, & u(x,0) = u_0(x). \end{aligned} \tag{A.1}$$

Definition A.1.1. *Suppose u^n is a non-increasing (non-decreasing) function. Let u^{n+1} be computed by one multigrid cycle. Then the multigrid method is said to preserve monotonicity if u^{n+1} is also non-increasing (non-decreasing).*

Theorem A.1.1. *Both the two level multiplicative and additive multigrid time stepping schemes preserve monotonicity for the linear wave equation.*

Proof. Given u^n, let u^{n+1} be the solution after one iteration of the two level multiplicative scheme. We assume a Dirichlet boundary condition as given in (A.1). For $i = 2, 4, \ldots, N-1$, u_i^{n+1} and u_{i-1}^{n+1} are computed using (2.8) and (2.12),

$$u_i^{n+1} = \tilde{u}_i^H \ = \bar{u}_i^h - \lambda(\bar{u}_i^h - \bar{u}_{i-2}^h), \tag{A.2}$$

$$u_{i-1}^{n+1} = \tilde{u}_{i-2}^H = \bar{u}_{i-2}^h - \lambda(\bar{u}_{i-2}^h - \bar{u}_{i-4}^h), \tag{A.3}$$

where $\bar{u}_i^h$ is the solution after fine grid upwind smoothing:

$$\bar{u}_i^h = u_i^n - \lambda(u_i^n - u_{i-1}^n). \tag{A.4}$$

Subtracting (A.3) from (A.2),

$$\begin{aligned} u_i^{n+1} - u_{i-1}^{n+1} &= \bar{u}_i^h - \lambda(\bar{u}_i^h - \bar{u}_{i-2}^h) - \bar{u}_{i-2}^h + \lambda(\bar{u}_{i-2}^h - \bar{u}_{i-4}^h), \\ &= (1-\lambda)\bar{u}_i^h + (2\lambda - 1)\bar{u}_{i-2}^h - \lambda\bar{u}_{i-4}^h. \end{aligned} \tag{A.5}$$

After substituting (A.4) into (A.5) and simplifying, we have

$$\begin{aligned} u_i^{n+1} - u_{i-1}^{n+1} &= (1-\lambda)^2(u_i^n - u_{i-1}^n) + (1-\lambda)(u_{i-1}^n - u_{i-2}^n) \\ &\quad + 2\lambda(1-\lambda)(u_{i-2}^n - u_{i-3}^n) + \lambda(u_{i-3}^n - u_{i-4}^n) \\ &\quad + \lambda^2(u_{i-4}^n - u_{i-5}^n). \end{aligned} \tag{A.6}$$

Similarly for $i-1$ being odd, we have

$$u_{i-1}^{n+1} - u_{i-2}^{n+1} = \bar{u}_{i-2}^H - \bar{u}_{i-2}^H = 0. \tag{A.7}$$

Thus, given that $0 \leq \lambda \leq 1$ (CFL condition), u^{n+1} is non-increasing (non-decreasing) if u^n is non-increasing (non-decreasing).

Following similar calculations for additive scheme, the solution when $i = 2, 4, \ldots, N-1$ is given by,

$$u_i^{n+1} = \tilde{u}_i^H \quad = u_i^n - \lambda(u_i^n - u_{i-2}^n), \tag{A.8}$$

$$u_{i-1}^{n+1} = \tilde{u}_{i-2}^H = u_{i-2}^n - \lambda(u_{i-2}^n - u_{i-4}^n). \tag{A.9}$$

After simplification,

$$\begin{aligned} u_i^{n+1} - u_{i-1}^{n+1} &= (1-\lambda)(u_i^n - u_{i-1}^n) + (1-\lambda)(u_{i-1}^n - u_{i-2}^n) \\ &\quad + \lambda(u_{i-2}^n - u_{i-3}^n) + \lambda(u_{i-3}^n - u_{i-4}^n), \end{aligned} \tag{A.10}$$

$$u_{i-1}^{n+1} - u_{i-2}^{n+1} = 0. \tag{A.11}$$

Thus, the additive scheme also preserves monotonicity for $0 \leq \lambda \leq 1$. □

A.2 Linear Systems

We derive expressions required for the analysis of the TVD, consistency and convergence properties in Section 2.4.

Let $\widetilde{K}_1$ be an $m \times m_1$ block matrix containing the eigenvectors corresponding to α^+ and $\widetilde{K}_2$ be an $m \times m_2$ block matrix containing the eigenvectors corresponding to α^-. Then we write K as

$$K = \begin{bmatrix} \widetilde{K}_1 & \widetilde{K}_2 \end{bmatrix}, \tag{A.12}$$

Similarly, we write the matrix K^{-1} in terms of another set of block matrices as

$$K^{-1} = \begin{bmatrix} \widehat{K}_1 \\ \widehat{K}_2 \end{bmatrix}, \tag{A.13}$$

where $\widehat{K}_1$ is an $m_1 \times m$ block matrix and $\widehat{K}_2$ is an $m_2 \times m$ block matrix.

Lemma A.2.1. *For any $m \times m$ invertible matrix K, the following identities hold.*

$$\begin{bmatrix} \widehat{K}_1\widetilde{K}_1 & \widehat{K}_1\widetilde{K}_2 \\ \widehat{K}_2\widetilde{K}_1 & \widehat{K}_2\widetilde{K}_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Lemma A.2.2. *Given U_i^n and a zero boundary condition, the analytical expressions for p^{n+1} and q^{n+1} given by the two level multiplicative scheme are*

$$\begin{aligned}
p_i^{n+1} &= (1-\lambda\alpha^+)^2 p_i^n + \lambda\alpha^+(1-\lambda\alpha^+)p_{i-1}^n + \lambda\alpha^+(1-\lambda\alpha^+)p_{i-2}^n \\
&\quad + (\lambda\alpha^+)^2 p_{i-3}^n, \qquad i = 2,4,\ldots,N-3, && \text{(A.14)}\\
p_{i-1}^{n+1} &= (1-\lambda\alpha^+)^2 p_{i-2}^n + \lambda\alpha^+(1-\lambda\alpha^+)p_{i-3}^n + \lambda\alpha^+(1-\lambda\alpha^+)p_{i-4}^n \\
&\quad + (\lambda\alpha^+)^2 p_{i-5}^n, \qquad i = 2,4,\ldots,N-1, && \text{(A.15)}\\
q_i^{n+1} &= (1+\lambda\alpha^-)^2 q_i^n - \lambda\alpha^-(1+\lambda\alpha^-)q_{i+1}^n - \lambda\alpha^-(1+\lambda\alpha^-)q_{i+2}^n \\
&\quad + (\lambda\alpha^-)^2 q_{i+3}^n, \qquad i = 2,4,\ldots,N-3, && \text{(A.16)}\\
q_{i-1}^{n+1} &= (1+\lambda\alpha^-)^2 q_i^n - \lambda\alpha^-(1+\lambda\alpha^-)q_{i+1}^n - \lambda\alpha^-(1+\lambda\alpha^-)q_{i+2}^n \\
&\quad + (\lambda\alpha^-)^2 q_{i+3}^n, \qquad i = 2,4,\ldots,N-1. && \text{(A.17)}
\end{aligned}$$

Proof. Given U_i^n and a zero boundary condition, we note that $U_i^n = 0$ for any $i \leq 0$ and $i \geq N-1$. Let U_i^{n+1}, $i = 1,2,\ldots,N-2$ be the solution obtained after one iteration of

a two level multiplicative scheme. For the grid points which coincide with the coarse grid points, i.e., $i = 2, 4, \ldots, N-3$, U_i^{n+1} is obtained from Algorithm 2 as

$$U_i^{n+1} = \widetilde{U}_i^H, \tag{A.18}$$

where $\widetilde{U}_i^H$ is the solution after coarse grid evolution. For x_{i-1}^{n+1}, $i = 2, 4, \ldots, N-1$, the interpolated solution is the steady state solution of the local Riemann problem using $\widetilde{U}_{i-2}^H$ and $\widetilde{U}_i^H$ as the boundary values as described in Section 2.3. Analytically, the steady state solution is not apparent from the U-space as the system consists of a mixture of characteristics. However, for constant coefficient systems, the steady state solution in the U-space is equivalent to the steady state solution in the W-space. For the characteristic equations associated with p, the sign of the characteristic speed is positive, i.e., the wave is moving to the right and hence the interpolated value at x_{i-1}^{n+1} is $\tilde{p}_{i-2}^H$. In contrast, the sign of the characteristic speed for equations associated with q is negative and hence $q_{i-1}^{n+1} = \tilde{q}_i^H$. Using the relations $W = K^{-1}U$ and (A.13), we have

$$W_{i-1}^{n+1} = \begin{bmatrix} \tilde{p}_{i-2}^H \\ \tilde{q}_i^H \end{bmatrix} = \begin{bmatrix} \widehat{K}_1 \widetilde{U}_{i-2}^H \\ \widehat{K}_2 \widetilde{U}_i^H \end{bmatrix}.$$

U_{i-1}^{n+1} is then computed as KW_{i-1}^{n+1},

$$\begin{aligned} U_{i-1}^{n+1} &= \begin{bmatrix} \widetilde{K}_1 & \widetilde{K}_2 \end{bmatrix} \begin{bmatrix} \widehat{K}_1 \widetilde{U}_{i-2}^H \\ \widehat{K}_2 \widetilde{U}_i^H \end{bmatrix}, \\ &= \widetilde{K}_1 \widehat{K}_1 \widetilde{U}_{i-2}^H + \widetilde{K}_2 \widehat{K}_2 \widetilde{U}_i^H. \end{aligned} \tag{A.19}$$

For the multiplicative scheme, (A.18) is rewritten using (2.3) as

$$\begin{aligned} U_i^{n+1} &= (I - \lambda A^+ + \lambda A^-) U_i^n + \lambda A^+ (I - \lambda A^+) U_{i-1}^n + \lambda A^+ (I - \lambda A^+) U_{i-2}^n \\ &\quad + (\lambda A^+)^2 U_{i-3}^n - \lambda A^- (I + \lambda A^-) U_{i+1}^n - \lambda A^- (I + \lambda A^-) U_{i+2}^n \\ &\quad + (\lambda A^-)^2 U_{i+3}^n. \end{aligned} \tag{A.20}$$

Multiplying both sides of (A.20) by $K^{-1}A$, we obtain a decoupled equation in characteristic variables,

$$\begin{aligned} \Lambda W_i^{n+1} &= (\Lambda - 2\lambda(\Lambda^+)^2 + \lambda^2(\Lambda^+)^3 + 2\lambda(\Lambda^-)^2 + \lambda^2(\Lambda^-)^3) W_i^n \\ &\quad + \lambda(\Lambda^+)^2 (I - \lambda\Lambda^+) W_{i-1}^n - \lambda(\Lambda^-)^2 (I + \lambda\Lambda^-) W_{i+1}^n \\ &\quad + \lambda(\Lambda^+)^2 (I - \lambda\Lambda^+) W_{i-2}^n + \lambda^2(\Lambda^+)^3 W_{i-3}^n - \lambda(\Lambda^-)^2 (I + \lambda\Lambda^-) W_{i+2}^n \\ &\quad + \lambda^2(\Lambda^-)^3 W_{i+3}^n. \end{aligned} \tag{A.21}$$

We rewrite (A.21) as two separate equations, one for the vector $p = [p_1, p_2, \ldots, p_{m_1}]$ associated with positive eigenvalues and the other for vector $q = [q_{m_1+1}, q_{m_1+2}, \ldots, q_m]$ associated with negative eigenvalues as given in (A.14) and (A.16) respectively.

Similarly, for the odd grid points, we multiply both sides of (A.19) by $K^{-1}A$ and simplify using Lemma A.2.1 and (2.3) as

$$\begin{aligned}\Lambda W_{i-1}^{n+1} &= \Lambda^+(I-\lambda\Lambda^+)^2 W_{i-2}^n + \lambda(\Lambda^+)^2(I-\lambda\Lambda^+)W_{i-3}^n + \lambda(\Lambda^+)^2(I-\lambda\Lambda^+)W_{i-4}^n \\ &\quad + \lambda^2(\Lambda^+)^3 W_{i-5}^n + \Lambda^-(I+\lambda\Lambda^-)^2 W_i^n - \lambda(\Lambda^-)^2(I+\lambda\Lambda^-)W_{i+1}^n \\ &\quad - \lambda(\Lambda^-)^2(I+\lambda\Lambda^-)W_{i+2}^n + \lambda^2(\Lambda^-)^3 W_{i+3}^n. \end{aligned} \tag{A.22}$$

Rewriting the decoupled set of m scalar equations (A.22) as two separate equations, we have (A.15) and (A.17) respectively.

□

Lemma A.2.3. *Given U_i^n and a zero boundary condition, the analytical expressions for p^{n+1} and q^{n+1} given by a k-level additve scheme are*

$$p_i^{n+1} = (1-\lambda\alpha^+)p_i^n + \lambda\alpha^+ p_{i-2^{k-1}}^n, \tag{A.23}$$

$$\left.\begin{matrix} p_{i-1}^{n+1} \\ p_{i-2}^{n+1} \\ \vdots \\ p_{i-2^{k-1}+1} \end{matrix}\right\} = (1-\lambda\alpha^+)p_{i-2^{k-1}}^n + \lambda\alpha^+ p_{i-2\cdot 2^{k-1}}^n, \tag{A.24}$$

$$q_i^{n+1} = (1+\lambda\alpha^-)q_i^n - \lambda\alpha^- q_{i+2^{k-1}}^n, \tag{A.25}$$

$$\left.\begin{matrix} q_{i-1}^{n+1} \\ q_{i-2}^{n+1} \\ \vdots \\ q_{i-2^{k-1}+1} \end{matrix}\right\} = (1+\lambda\alpha^-)q_i^n - \lambda\alpha^- q_{i+2\cdot 2^{k-1}}^n. \tag{A.26}$$

$$i = 2^{k-1}, 2\cdot 2^{k-1}, \ldots$$

Proof. For a k-level additive scheme and $i = 2^{k-1}, 2\cdot 2^{k-1}, \ldots$, the solution after the coarse grid evolution is as shown below

$$\widetilde{U}_i^{(k)} = (I - \lambda A^+ + \lambda A^-)U_i^n + \lambda A^+ U_{i-2^{k-1}}^n - \lambda A^+ U_{i+2^{k-1}}^n.$$

The interpolated solution on level $(k-1)$ is given by,

$$\begin{aligned}\widetilde{U}_i^{(k-1)} &= \widetilde{U}_i^{(k)}, \\ \widetilde{U}_{i-2^{k-2}}^{(k-1)} &= \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k)}.\end{aligned}$$

Similarly, for level $(k-2)$, we have,

$$\begin{aligned}
\widetilde{U}_i^{(k-2)} &= \widetilde{U}_i^{(k-1)} && = \widetilde{U}_i^{(k)},\\
\widetilde{U}_{i-2^{k-3}}^{(k-2)} &= \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-2}}^{(k-1)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k-1)} && = \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k)},\\
\widetilde{U}_{i-2^{k-2}}^{(k-2)} &= \widetilde{U}_{i-2^{k-2}}^{(k-1)} && = \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k)},\\
\widetilde{U}_{i-3\cdot 2^{k-3}}^{(k-2)} &= \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k-1)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_{i-2^{k-2}}^{(k-1)} && = \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k)}.
\end{aligned}$$

By induction, the solution on the fine grid, U_i^{n+1} is given by

$$U_i^{n+1} = \widetilde{U}_i^{(k)}, \tag{A.27}$$

$$\left.\begin{matrix} U_{i-1}^{n+1}\\ U_{i-2}^{n+1}\\ \vdots\\ U_{i-2^{k-1}+1}\end{matrix}\right\} = \widetilde{K}_1\widehat{K}_1\widetilde{U}_{i-2^{k-1}}^{(k)} + \widetilde{K}_2\widehat{K}_2\widetilde{U}_i^{(k)}. \tag{A.28}$$

Multiplying both sides of (A.27) and (A.28) by $K^{-1}A$ and simplifying reduces the system into a set of m-decoupled equations,

$$\begin{aligned}
\Lambda W_i^{n+1} &= (I - \lambda\Lambda^+ + \lambda\Lambda^-)W_i^n + \lambda\Lambda^+ W_{i-2^{k-1}}^n - \lambda\Lambda^- W_{i+2^{k-1}}^n,\\
\left.\begin{matrix} \Lambda W_{i-1}^{n+1}\\ \Lambda W_{i-2}^{n+1}\\ \vdots\\ \Lambda W_{i-2^{k-1}+1}\end{matrix}\right\} &= \Lambda^+(I-\Lambda^+)W_{i-2^{k-1}}^n + \lambda(\Lambda^+)^2 W_{i-2\cdot 2^{k-1}}^n + \Lambda^-(I+\lambda\Lambda^-)W_i^n\\
&\quad - \lambda(\Lambda^-)^2 W_{i+2\cdot 2^{k-1}}^n.
\end{aligned}$$

We rewrite these decoupled equations as two separate equations. One for all the characteristic equations with positive sign for the characteristic speeds: (A.23) and (A.24). Another set of equations for the characteristic equations with negative sign for the characteristic speeds: (A.25) and (A.26).

□

Appendix B

Discretization Coefficients for HJB and HJBI Systems

B.1 Dynamic Bertrand Oligopoly

B.1.1 Monopoly Problem: Discrete Equation Coefficients

The discrete coeffecients α_i and β_i for (3.6) are presented here. Standard central differencing for all derivatives of (3.4) results in

$$\begin{aligned} \alpha_i^{cent} &= \frac{\sigma^2}{2\Delta x^2} - \frac{\kappa - p}{2\eta \Delta x}, \\ \beta_i^{cent} &= \frac{\sigma^2}{2\Delta x^2} + \frac{\kappa - p}{2\eta \Delta x}. \end{aligned}$$

Using upstream differencing for first order terms and central for second order terms, we have

$$\begin{aligned} \alpha_i^{ups} &= \frac{\sigma^2}{2\Delta x^2} + \max\left\{0, -\frac{\kappa - p}{2\eta \Delta x}\right\}, \\ \beta_i^{ups} &= \frac{\sigma^2}{2\Delta x^2} + \max\left\{0, \frac{\kappa - p}{2\eta \Delta x}\right\}. \end{aligned}$$

A combination of central and upstream differencing is then chosen according to Algorithm 3 on a node by node basis.

Algorithm 3 Differencing method for monopoly problem

if $\alpha_i^{cent} \geq 0$ **and** $\beta_i^{cent} \geq 0$ **then**
  $\alpha_i \leftarrow \alpha_i^{cent}$
  $\beta_i \leftarrow \beta_i^{cent}$
else
  $\alpha_i \leftarrow \alpha_i^{ups}$
  $\beta_i \leftarrow \beta_i^{ups}$
end if

B.1.2 Duopoly Problem: Discrete Equation Coefficients

The coefficients of the discrete equations (3.12) are given here. Using standard central differencing for the first and second order derivatives, we have for player 1,

$$\begin{aligned}
(\alpha_1)_{i,j}^{x_1,cent} &= \frac{\sigma_1^2}{2\Delta x_1^2} + \frac{a_1 - a_2 p_1 + a_3 (p_2^*)_{i,j}}{2\Delta x_1} \\
(\beta_1)_{i,j}^{x_1,cent} &= \frac{\sigma_1^2}{2\Delta x_1^2} - \frac{a_1 - a_2 p_1 + a_3 (p_2^*)_{i,j}}{2\Delta x_1} \\
(\alpha_1)_{i,j}^{x_2,cent} &= \frac{\sigma_2^2}{2\Delta x_2^2} - \frac{1}{2\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3 (p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right) \\
(\beta_1)_{i,j}^{x_2,cent} &= \frac{\sigma_2^2}{2\Delta x_2^2} + \frac{1}{2\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3 (p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)
\end{aligned}$$

Similarly for player 2, we have

$$\begin{aligned}
(\alpha_2)_{i,j}^{x_1,cent} &= \frac{\sigma_1^2}{2\Delta x_1^2} - \frac{1}{2\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3 (p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right) \\
(\beta_2)_{i,j}^{x_1,cent} &= \frac{\sigma_1^2}{2\Delta x_1^2} + \frac{1}{2\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3 (p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right) \\
(\alpha_2)_{i,j}^{x_2,cent} &= \frac{\sigma_2^2}{2\Delta x_2^2} + \frac{a_1 - a_2 p_2 + a_3 (p_1^*)_{i,j}}{2\Delta x_2} \\
(\beta_2)_{i,j}^{x_2,cent} &= \frac{\sigma_2^2}{2\Delta x_2^2} - \frac{a_1 - a_2 p_2 + a_3 (p_1^*)_{i,j}}{2\Delta x_2}
\end{aligned}$$

Using upstream for first order derivatives and central for second order derivatives results in

$$(\alpha_1)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, \frac{(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j})}{\Delta x_1}\right\} \tag{B.1}$$

$$(\beta_1)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, -\frac{(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j})}{\Delta x_1}\right\} \tag{B.2}$$

$$(\alpha_1)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, -\frac{1}{\Delta x_2}\left(\frac{\gamma}{\eta}(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)\right\} \tag{B.3}$$

$$(\beta_1)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, \frac{1}{\Delta x_2}\left(\frac{\gamma}{\eta}(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)\right\} \tag{B.4}$$

$$(\alpha_2)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, -\frac{1}{\Delta x_1}\left(\frac{\gamma}{\eta}(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)\right\} \tag{B.5}$$

$$(\beta_2)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, \frac{1}{\Delta x_1}\left(\frac{\gamma}{\eta}(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)\right\} \tag{B.6}$$

$$(\alpha_2)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, \frac{(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j})}{\Delta x_2}\right\} \tag{B.7}$$

$$(\beta_2)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, -\frac{(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j})}{\Delta x_2}\right\}. \tag{B.8}$$

The coefficient due to the cross derivative term for both the players is given by

$$\zeta_{i,j} = \begin{cases} \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2}, & \text{if } \rho \geq 0 \\ \frac{-\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2}, & \text{if } \rho < 0 \end{cases} \tag{B.9}$$

A combination of central and upstream differencing is then chosen according to Algorithm 4.

Algorithm 4 Differencing in the x_k, $k = 1, 2$ and for each player $l = 1, 2$.

if $(\alpha_l)_{i,j}^{x_k,cent} - \zeta_{i,j} \geq 0$ **then**
  $(\alpha_l)_{i,j}^{x_k} \leftarrow (\alpha_l)_{i,j}^{x_k,cent}$
  $(\beta_l)_{i,j}^{x_k} \leftarrow (\beta_l)_{i,j}^{x_k,cent}$
else
  $(\alpha_l)_{i,j}^{x_k} \leftarrow (\alpha_l)_{i,j}^{x_k,ups}$
  $(\beta_l)_{i,j}^{x_k} \leftarrow (\beta_l)_{i,j}^{x_k,ups}$
end if

B.2 American Option under Regime Switching: Discrete Equation Coefficients

The coefficients of the discrete equations (3.29) and (3.36) are given here. Using standard three point stencil and central differencing for the first and second order derivatives, we have

$$\alpha_{i,j}^{cent} = \frac{2a_{i,j}\left(S,\tau,\varphi\right)}{\left(S_i - S_{i-1}\right)\left(S_{i+1} - S_{i-1}\right)} - \frac{b_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_{i-1}\right)}, \tag{B.10}$$

$$\beta_{i,j}^{cent} = \frac{2a_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_i\right)\left(S_{i+1} - S_{i-1}\right)} + \frac{b_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_{i-1}\right)}. \tag{B.11}$$

Using upstream (forward/backward) differencing for the first order derivative and central differencing for the second order derivative, we have

$$\alpha_{i,j}^{ups} = \frac{2a_{i,j}\left(S,\tau,\varphi\right)}{\left(S_i - S_{i-1}\right)\left(S_{i+1} - S_{i-1}\right)} + \max\left\{0, -\frac{b_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_{i-1}\right)}\right\}, \tag{B.12}$$

$$\beta_{i,j}^{ups} = \frac{2a_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_i\right)\left(S_{i+1} - S_{i-1}\right)} + \max\left\{0, \frac{b_{i,j}\left(S,\tau,\varphi\right)}{\left(S_{i+1} - S_{i-1}\right)}\right\}. \tag{B.13}$$

A weighted average of central and upstream differencing is used on a node by node basis such that a positive coefficient discretization is obtained as detailed in Algorithm 5.

Algorithm 5 Differencing method.

for $i = 1, 2, \ldots$ **do**
  $\omega = 1$
  if $\alpha_{i,j}^{cent} < 0$ **then**
    $\omega = \frac{\alpha_{i,j}^{ups}}{\alpha_{i,j}^{ups} - \alpha_{i,j}^{cent}}$
  else
    if $\beta_{i,j}^{cent} < 0$ **then**
      $\omega = \frac{\beta_{i,j}^{ups}}{\beta_{i,j}^{ups} - \beta_{i,j}^{cent}}$
    end if
  end if
  $\alpha_{i,j} = \omega \alpha_{i,j}^{cent} + (1 - \omega) \alpha_{i,j}^{ups}$
  $\beta_{i,j} = \omega \beta_{i,j}^{cent} + (1 - \omega) \beta_{i,j}^{ups}$
end for

Appendix C

Discretization Coefficients for Two-factor Stochastic Volatility Model

Using standard central differencing for the first and second order derivatives (not cross derivatives), we have

$$
\begin{aligned}
\alpha_{i,j,k}^{x,cent} &= \frac{\nu_j}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})} - \frac{r - \nu_j/2}{(x_{i+1} - x_{i-1})}, \\
\beta_{i,j,k}^{x,cent} &= \frac{\nu_j}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})} + \frac{r - \nu_j/2}{(x_{i+1} - x_{i-1})}, \\
\alpha_{i,j,k}^{\nu,cent} &= \frac{C_2^2 \nu_j^{2\alpha}}{(\nu_j - \nu_{j-1})(\nu_{j+1} - \nu_{j-1})} - \frac{C_1(\xi_k - \nu_j)}{(\nu_{j+1} - \nu_{j-1})}, \\
\beta_{i,j,k}^{\nu,cent} &= \frac{C_2^2 \nu_j^{2\alpha}}{(\nu_{j+1} - \nu_j)(\nu_{j+1} - \nu_{j-1})} + \frac{C_1(\xi_k - \nu_j)}{(\nu_{j+1} - \nu_{j-1})}, \\
\alpha_{i,j,k}^{\xi,cent} &= \frac{C_5^2 \xi_j^{2\beta}}{(\xi_k - \xi_{k-1})(\xi_{k+1} - \xi_{k-1})} - \frac{C_3(C_4 - \xi_k)}{(\xi_{k+1} - \xi_{k-1})}, \\
\beta_{i,j,k}^{\xi,cent} &= \frac{C_5^2 \xi_j^{2\beta}}{(\xi_{k+1} - \xi_k)(\xi_{k+1} - \xi_{k-1})} + \frac{C_3(C_4 - \xi_k)}{(\xi_{k+1} - \xi_{k-1})}.
\end{aligned}
$$

Using upstream for the the first order derivatives and central for second order derivatives results in

$$
\begin{aligned}
\alpha_{i,j,k}^{x,ups} &= \frac{\nu_j}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})} + \max\left(-\frac{(r - \nu_j/2)}{(x_i - x_{i-1})}, 0\right), \\
\beta_{i,j,k}^{x,ups} &= \frac{\nu_j}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})} + \max\left(\frac{(r - \nu_j/2)}{(x_{i+1} - x_i)}, 0\right), \\
\alpha_{i,j,k}^{\nu,ups} &= \frac{C_2^2 \nu_j^{2\alpha}}{(\nu_j - \nu_{j-1})(\nu_{j+1} - \nu_{j-1})} + \max\left(\frac{-C_1(\xi_k - \nu_j)}{(\nu_j - \nu_{j-1})}, 0\right), \\
\beta_{i,j,k}^{\nu,ups} &= \frac{C_2^2 \nu_j^{2\alpha}}{(\nu_{j+1} - \nu_j)(\nu_{j+1} - \nu_{j-1})} + \max\left(\frac{C_1(\xi_k - \nu_j)}{(\nu_{j+1} - \nu_j)}, 0\right), \\
\alpha_{i,j,k}^{\xi,ups} &= \frac{C_5^2 \xi_j^{2\beta}}{(\xi_k - \xi_{k-1})(\xi_{k+1} - \xi_{k-1})} + \max\left(\frac{-C_3(C_4 - \xi_k)}{(\xi_k - \xi_{k-1})}, 0\right), \\
\beta_{i,j,k}^{\xi,ups} &= \frac{C_5^2 \xi_j^{2\beta}}{(\xi_{k+1} - \xi_k)(\xi_{k+1} - \xi_{k-1})} + \max\left(\frac{C_3(C_4 - \xi_k)}{(\xi_{k+1} - \xi_k)}, 0\right).
\end{aligned}
$$

The coefficients due to the cross derivative terms are given by

$$
\zeta_{i,j,k}^{x\nu} = \begin{cases} \frac{\rho_{12} C_2 \nu_j^{\alpha} \sqrt{\nu_j}}{(x_{i+1} - x_i)(\nu_{j+1} - \nu_j) + (x_i - x_{i-1})(\nu_j - \nu_{j-1})}, & \text{if } \rho_{12} \geq 0, \\ \frac{-\rho_{12} C_2 \nu_j^{\alpha} \sqrt{\nu_j}}{(x_{i+1} - x_i)(\nu_j - \nu_{j-1}) + (x_i - x_{i-1})(\nu_{j+1} - \nu_j)}, & \text{if } \rho_{12} < 0. \end{cases}
$$

$$
\zeta_{i,j,k}^{x\xi} = \begin{cases} \frac{\rho_{13} C_5 \xi_k^{\beta} \sqrt{\nu_j}}{(x_{i+1} - x_i)(\xi_{k+1} - \xi_k) + (x_i - x_{i-1})(\xi_k - \xi_{k-1})}, & \text{if } \rho_{13} \geq 0, \\ \frac{-\rho_{13} C_2 \nu_j^{\alpha} \sqrt{\nu_j}}{(x_{i+1} - x_i)(\xi_k - \xi_{k-1}) + (x_i - x_{i-1})(\xi_{k+1} - \xi_k)}, & \text{if } \rho_{13} < 0. \end{cases}
$$

$$
\zeta_{i,j,k}^{\nu\xi} = \begin{cases} \frac{\rho_{23} C_2 C_5 \nu_j^{\alpha} \xi_k^{\beta}}{(\nu_{j+1} - \nu_j)(\xi_{k+1} - \xi_k) + (\nu_j - \nu_{j-1})(\xi_k - \xi_{k-1})}, & \text{if } \rho_{23} \geq 0, \\ \frac{-\rho_{23} C_2 C_5 \nu_j^{\alpha} \xi_k^{\beta}}{(\nu_{j+1} - \nu_j)(\xi_k - \xi_{k-1}) + (\nu_j - \nu_{j-1})(\xi_{k+1} - \xi_k)}, & \text{if } \rho_{23} < 0. \end{cases}
$$

The coefficients of the discrete operator (4.22) are then constructed using a combination of central and upstream differencing according to Algorithm 6.

Algorithm 6 Differencing in the x, ν and ξ dimensions

if $\alpha^{x}_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{x\xi}_{i,j,k} \geq 0$ **then**
  $\alpha^{x}_{i,j,k} \leftarrow \alpha^{x,cent}_{i,j,k}$
  $\beta^{x}_{i,j,k} \leftarrow \beta^{x,cent}_{i,j,k}$
else
  $\alpha^{x}_{i,j,k} \leftarrow \alpha^{x,ups}_{i,j,k}$
  $\beta^{x}_{i,j,k} \leftarrow \beta^{x,ups}_{i,j,k}$
end if
if $\alpha^{\nu}_{i,j,k} - \zeta^{x\nu}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0$ **then**
  $\alpha^{\nu}_{i,j,k} \leftarrow \alpha^{\nu,cent}_{i,j,k}$
  $\beta^{\nu}_{i,j,k} \leftarrow \beta^{\nu,cent}_{i,j,k}$
else
  $\alpha^{\nu}_{i,j,k} \leftarrow \alpha^{\nu,ups}_{i,j,k}$
  $\beta^{\nu}_{i,j,k} \leftarrow \beta^{\nu,ups}_{i,j,k}$
end if
if $\alpha^{\xi}_{i,j,k} - \zeta^{x\xi}_{i,j,k} - \zeta^{\nu\xi}_{i,j,k} \geq 0$ **then**
  $\alpha^{\xi}_{i,j,k} \leftarrow \alpha^{\xi,cent}_{i,j,k}$
  $\beta^{\xi}_{i,j,k} \leftarrow \beta^{\xi,cent}_{i,j,k}$
else
  $\alpha^{\xi}_{i,j,k} \leftarrow \alpha^{\xi,ups}_{i,j,k}$
  $\beta^{\xi}_{i,j,k} \leftarrow \beta^{\xi,ups}_{i,j,k}$
end if

References

[1] M. Akian, J. L. Menaldi, and A. Sulem. On the investment-consumption model with transaction costs. *Research Report RR-1926, INRIA*, 1993.

[2] M. Akian, J. L. Menaldi, and A. Sulem. Multi-asset portfolio selection problem with transaction costs. *Mathematics and Computers in Simulation*, 38:163–172, 1995.

[3] M. Akian, P. Sequier, and A. Sulem. Analysis of the multigrid method applied to first order systems. In *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 2193–2198, 1995.

[4] M. Akian, A. Sulem, and M. I. Taksar. Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logarithmic utility. *Mathematical Finance*, 11:153–188, 2001.

[5] S. Amarala and J. W. L. Wan. Multigrid methods for systems of hyperbolic conservation laws. *SIAM Journal on Multiscale Modeling and Simulation*, 11(2):586–614, 2013.

[6] E. Ayache. Equity-credit problem. *In R. Cont, editor, Encyclopedia of Quantitative Finance, Wiley, New York*, pages 571–575, 2010.

[7] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance. *Numerical Methods in Finance*, pages 1–21, 1997.

[8] G. Barles, C. Daher, and M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Mathematical models and methods in applied sciences*, 5(01):125–143, 1995.

[9] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM Journal on Numerical Analysis*, 43:540–558, 2005.

[10] G. Barles and E. Rouy. A strong comparison result for the Bellman equation arising in stochastic exit time control problems and applications. *Comm. Partial Differential Equations*, 23(1995-2033), 1998.

[11] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283, 1991.

[12] R. E. Bellman. Introduction to the mathematical theory of control processes: Nonlinear processes. *Mathematics in Science and Engineering. Academic Press*, 1971.

[13] Y. Bergman. Option pricing with differential interest rates. *Review of Financial Studies*, 8:475–500, 1995.

[14] L. Bergomi. Smile dynamics. Risk Pages 117-123, September 2004.

[15] J. Bertrand. Théorie mathématique de la richesse sociale. *J. Savants*, 67(499-508), 1883.

[16] O. Bokanowski, S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM Journal on Numerical Analysis*, 47:3001–3026, 2009.

[17] J. F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM Journal on Numerical Analysis*, 41(3):1008–1021, 2003.

[18] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, 31(138):333–390, 1977.

[19] A. Brandt and C. W. Cryer. Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems. *SIAM J. Sci. Statist. Comput*, 4:655–684, 1983.

[20] H. Buehler. Consistent variance curve models. *Finance and Stochastics*, 10:178–203, 2006.

[21] P. Chan and R. Sircar. Bertrand and Cournot mean field games. *To appear in Applied Mathematics and Optimization*, 2015.

[22] S. Chaumont. A strong comparison result for viscosity solutions to Hamilton-Jacobi-Bellman equations with dirichlet conditions on a non-smooth boundary. *Preprint, Institute Elie Cartan, Universitè Nancy I*, 2004.

[23] R. V. Chima and G. M. Johnson. Efficient solution of the Euler and Navier Stokes equations with a vectorized multi grid algorithm. *AIAA, (Danvers 1983)*, 1983.

[24] P. Christoffersen, S. L. Heston, and K. Jacobs. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 55(12):1914–1932, 2009.

[25] N. Clarke and K. Parrot. Multigrid for American option pricing with stochastic volatility. *Applied Mathematical Finance*, 6(177-195), 1999.

[26] S. S. Clift and P. A. Forsyth. Numerical solution of two asset jump diffusion models for option valuation. *Applied Numerical Mathematics*, 58(6):743–782, 2008.

[27] A. Cournot. Recherches sur les principes mathématique de la théorie des richesses. Hachette, Paris (1838). *English translation by N.T. Bacon, published in Economic Classics, Macmillan, 1897, and reprinted in 1960 by Augustus M. Kelley*, 1838.

[28] S. Crepey. About pricing equations in finance. In A.R. Carmona, editor, *Paris-Princeton Lectures on Mathematical Finance*, pages 63–203. Springer, Berlin, 2010.

[29] K. Debrabant and E. Jakobsen. Semi-lagrangian schemes for linear and fully non-linear diffusion equations. *Mathematics of Computation*, 82(283):1433–1462, 2013.

[30] E. Derman. Introduction to the smile; the principles of valuation. E4718 Lecture Series, Spring 2007.

[31] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for american options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004.

[32] E. Dick. Multigrid acceleration of a flux-difference splitting method for steady Euler equations. *Journal of Computational and Applied Mathematics*, 28:173–180, 1989.

[33] E. Dockner, S. Jorgensen, N. V. Long, and G. Sorger. *Differential games in economics and management science*. Cambridge University Press, 2000.

[34] H. Dong and N. V. Krylov. On the rate of convergence of finite difference approximations for Bellman equations with constant coefficients. *St Petersburg Mathematical Journal*, 17(2):295–314, 2006.

[35] B. Dupire. *A unified theory of volatility*. In Peter Carr, editor, Derivatives Pricing: The Classic Collection, Risk Books, pages 185-196, 2004.

[36] P. A. Forsyth. Numerical computation for financial modelling. *Course notes, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada.*, 2012.

[37] P. A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *Journal of Computational Finance*, 2007.

[38] P. A. Forsyth and K. R. Vetzal. Quadratic convergence of a penalty method for valuing American options. *SIAM Journal on Scientific Computing*, 23:2095–2122, 2002.

[39] J. P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Cambridge University Press, 2011.

[40] U. Gärtel. Parallel multigrid solver for 3d anisotropic elliptic problems. *Hypercube and Distributed Computers (eds F. André and J.P. Versus) Noth Holland, Amsterdam*, pages 37–47, 1989.

[41] J. Gatheral. Developments in volatility derivatives pricing. Presentation: Global Derivatives, Paris, 2007.

[42] J. Gatheral. Developments in volatility derivatives pricing. Presentation: New York Quantitative Finance Seminar, 2008.

[43] J. Gatheral, C. Bayer, and M. Karlsmark. Fast ninomiya-victoir calibration of the double-mean-reverting model. October 2013.

[44] S. K. Godunov. A finite difference method for the computation of discontinuous solutions of the equations of fluid dynamics. *Mat. Sb.*, 47:357–393, 1959.

[45] F. Grasso and M. Marini. Solutions of hypersonic viscous flows with total variation diminishing multigrid techniques. *Computers and Fluids*, 24:571–592, June 1995.

[46] B. Gustafsson and P. Lötstedt. Analysis of the multigrid method applied to first order systems. In *Mandel, J. et al. (ed.) Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods*, pages 181–233. SIAM Philadelphia, PA, 1989.

[47] W. Hackbush. *Multigrid Methods and applications*. Springer, Berlin Heidelberg New York, 1985.

[48] D. Han and J. W. L. Wan. Multigrid methods for second order Hamilton-Jacobi-Bellman and Hamilton-Jacobi-Bellman-Isaacs equations. *SIAM Journal on Scientific Computing*, 35(5):S323–S344, 2013.

[49] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 135:260–278, 1997.

[50] P. W. Hemker and S. P. Spekreijse. Multigrid solution of the steady Euler equations. D.Braess, W.Hackbush and U.Trottenberg (eds.). Advances in Multi-grid Methods, Notes on Numerical Fluid Dynamics, Viewing Publ. Comp., Braunschweig, 1985.

[51] P. W. Hemker and S. P. Spekreijse. Multiple grid and Osher's scheme for the efficient solution of the steady Euler equations. *Applied Numerical Mathematics*, 2:475–493, December 1986.

[52] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[53] M. Holtz and A. Kunoth. B-spline based monotone multigrid methods. *SIAM J. Numer. Anal.*, 45(3):1175–1199, 2007.

[54] R. H. W. Hoppe. Multi-grid methods for Hamilton-Jacobi-Bellman equations. *Numerische Mathematik*, 49:239–254, 1986.

[55] R. H. W. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer, Anal.*, 24:1046–1065, 1987.

[56] T. Y. Hou and P. LeFloch. Why non-conservative schemes converge to the wrong solutions: error analysis. *Mathematics of Computation*, 62(206):497–530, April 1994.

[57] R. A. Howard. Dynamic programming and Markov process. *MIT Press, Cambridge, MA, USA*, 1960.

[58] Y. Huang and P. A. Forsyth. Analysis of a penalty method for pricing a guaranteed minimum withdrawal benefit (GMWB). *IMA Journal of Numerical Analysis*, 32:320–351, 2012.

[59] Y. Huang, P. A. Forsyth, and G.Labahn. Methods for pricing American options under regime switching. *SIAM Journal on Scientific Computing*, 33:2144–2168, 2011.

[60] J. C. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

[61] H. Ishii and S. Koike. Viscosity solutions for monotone systems of second order elliptic pdes. *Comm. Partial Differential Equations*, 16:1095–1128, 1991.

[62] A. Jameson. Unpublished manuscript.

[63] A. Jameson. Solution of the Euler equations for two-dimensional transonic flow by a multigrid method. *Applied Mathematical Computation*, 13:327–356, 1983.

[64] A. Jameson. Multigrid solution of the Euler equations using implicit schemes. *AIAA*, 1985.

[65] D. C. Jespersen. Design and implementation of a multigrid code for the Euler equations. *Applied Mathematics and Computation*, 13:357–374, 1983.

[66] D. C. Jespersen. Recent developments in multigrid methods for the steady Euler equations. Lecture Notes, von Karman Institute, Rhode-St.Genese, Belgium, March 1984.

[67] D. C. Jespersen. A time-accurate multiple-grid algorithm. *Computational Fluid Dynamics Conference*, pages 58–66, 1985.

[68] G. M. Johnson. Multiple grid acceleration of Lax Wendroff algorithms. NASA Technical Memorandum, 1982.

[69] G. M. Johnson. Accelerated solution of the steady state Euler equations. W.G. Habashi (eds). Recent Advances in Numerical Methods in Fluids. Vol 4., Pineridge Press., 1983.

[70] A. Kanarachos and I. Vournas. Multigrid technique applied to the compressible Euler equations. *Acta Mechanica, Springer-Verlag*, 97:23–40, 1993.

[71] J. S. Kennedy. Hedging contingent claims in markets with jumps. *PhD Thesis, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada*, 2007.

[72] B. Koren and P. W. Hemker. Damped, direction-dependent multigrid for hypersonic flow computations. *Applied Numerical Mathematics*, 7:309–328, 1991.

[73] R. Kornhuber. Montone iterations for elliptic variational inequalities. *II. Numer. Math.*, 72, 1998.

[74] A. Kurganov and E. Tadmor. New high resolution central schemes for nonlinear conservation laws and convection-diffusion equations. *Journal of Computational Physics*, 160:241–282, 2000.

[75] P. D. Lax and B. Wendroff. Systems of conservation laws. *Communications on Pure and Applied Mathematics*, 13:217–237, 1960.

[76] M. P. Leclercq and B. Stoufflet. Characteristic multigrid method application to solve the Euler equations with unstructured and unnested grids. *Journal of Computational Physics*, 104:329–346, 1993.

[77] A. Ledvina and R. Sircar. Dynamic Bertrand oligopoly. *Applied Mathematics and Optimization*, 63(1):11–44, February 2011.

[78] B. V. Leer. Flux vector splitting for the euler equations. In *8th International Conference on Numerical Methods in Fluid Dynamics*, pages 507–512. Springer-Verlag, 1982.

[79] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2004.

[80] P. L. Lions and B. Mercier. Approximation numerique des equations hamilton-jacobi-bellman. *RAIRO - Analyse numerique*, 14:369–393, 1980.

[81] X. D. Liu and P. D. Lax. Positive schemes for solving multi-dimensional hyperbolic systems of conservation laws. *Journal of Computational Physics*, 187:428–440, 2003.

[82] P. Lötstedt and B. Gustafsson. Fourier analysis of multigrid methods for general systems of PDEs. *Mathematics of Computation*, 60(202):473–493, April 1993.

[83] K. Ma and P. A. Forsyth. An unconditionally monotone numerical scheme for the two factor uncertain volatility model. *Submitted to IMA Journal of Numerical Analysis*, 2014.

[84] W. A. Mulder. Multigrid relaxation for the Euler equations. *Journal of Computational Physics*, 60:235–252, 1985.

[85] W. A. Mulder. A new multigrid approach to convection problems. *Journal of Computational Physics*, 83:303–323, August 1989.

[86] Ni and R. H. A multiple-grid scheme for solving the Euler equations. *AIAA*, 20:1565–1571, 1982.

[87] C. W. Oosterlee. On multigrid for linear complimentarity problems with application to American-style options. *Electronic Transactions on Numerical Analysis*, 15:165–185, 2003.

[88] S. Osher and F. Solomon. Upwind difference schemes for hyperbolic systems of conservation laws. *Mathematics of Computation*, 38:339–374, 1982.

[89] D. M. Pooley, P. A. Forsyth, and K. R. Vetzal. Numerical convergence properties of option pricing PDEs with uncertain volatility. *IMA Journal of Numerical Analysis*, pages 241–267, 2003.

[90] C. Reisinger and P. A. Forsyth. Piecewise constant policy approximations to Hamilton-Jacobi-Bellman equations. *Techinical Report*, 2015.

[91] C. Reisinger and G. Wittum. On multigrid for anisotropic equations and variational inequalities. *Computing and Visualizaion in Science*, 7:189–197, 2004.

[92] Y. Saad. Iterative methods for sparse linear systems. *SIAM, Philadelphia*.

[93] P. J. Schonbucher. A market model for stochastic implied volatility. *The Royal Society*, 357:2071–2092, 1999.

[94] L. O. Scott. Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Finacial and Quantitative Analysis*, 22:419–438, 1987.

[95] S. Spekreijse. Multigrid solution of monotone second-order discretizations of hyperbolic conservation laws. *Mathematics of Computation*, 49(179):135–155, July 1987.

[96] J. L. Steger and R. F. Warming. Flux vector splitting of the inviscid gas dynmaic equations with applications to finite difference methods. *Journal of Computational Physics*, 40:263–293, 1981.

[97] J.L. Steger. A preliminary study of relaxation methods for the inviscid conservative gas dynamics equations using flux splitting. NASA Contractor Report 3415, 1981.

[98] P. K. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 21:995–1011, 1984.

[99] E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer-Verlag, 2009.

[100] U. Trottenberg, C. W. Oosterlee., and A. Schuller. *Multigrid*. Academic Press, 2001.

[101] S. T. Tse. Numerical methods for optimal trade execution. *PhD Thesis, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada*, 2012.

[102] R. S. Varga. *Matrix Iterative Analysis*. Springer-Verlag, Berlin, 2009.

[103] J. Wal. Discounted Markov games: Generalized policy iteration method. *Journal of Optimization Theory and Applications*, 25:125–138, 1978.

[104] J. W. L. Wan and A. Jameson. Monotonicity preserving multigrid time stepping schemes for conservation laws. *Computing and Visualizaion in Science*, 11:41–58, 2007.

[105] J. Wang and P. A. Forsyth. Maximal use of central differencing for Hamilton-Jacobi-Bellman PDEs in finance. *SIAM Journal on Numerical Analysis*, 46(3):1580–1601, 2008.

[106] P. Wesseling. *An introduction to multigrid methods*. John Wiley and Sons, 1992.

[107] D. Yao, Q. Zhang, and X. Y. Zhou. A regime switching model for European options. *In Stochastic Processes, Optimization and Control Theory: Applications in Financial Engineering, Queueing Networks and Manufacturing Systems. Springer International Series in Operations Research and Management Science*, 94:281–300, 2006.

[108] H. B. Zubair, C. C. W. Leentvaar, and C. W. Oosterlee. Efficient d-multigrid preconditioners for sparse-grid solution of high-dimensional partial differential equations. *International Journal of Computer Mathematics*, 84, 2007.

[109] R. Zvan, P. A. Forsyth, and K. R. Vetzal. A finite volume approach for contingent claims valuation. *IMA Journal of Numerical Analysis*, 21:703–731, 2001.