Numerical Solutions of Two-factor Hamilton-Jacobi-Bellman Equations in Finance

by

Kai Ma

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

In this thesis, we focus on solving multidimensional HJB equations which are derived from optimal stochastic control problems in the financial market. We develop a fully implicit, unconditionally monotone finite difference numerical scheme. Consequently, there are no time step restrictions due to stability considerations, and the fully implicit method has essentially the same complexity per step as the explicit method. The main difficulty in designing a discretization scheme is development of a monotonicity preserving approximation of cross derivative terms in the PDE. We primarily use a wide stencil based on a local coordination rotation. The analysis rigorously show that our numerical scheme is ℓ_∞ stable, consistent in the viscosity sense, and monotone. Therefore, our numerical scheme guarantees convergence to the viscosity solution.

Firstly, our numerical schemes are applied to pricing two factor options under an uncertain volatility model. For this application, a hybrid scheme which uses the fixed stencil as much as possible is developed to take advantage of its accuracy and computational efficiency. Secondly, using our numerical method, we study the problem of optimal asset allocation where the risky asset follows stochastic volatility. Finally, we utilize our numerical scheme to carry out an optimal static hedge, in the case of an uncertain local volatility model.

Acknowledgements

First, I would like to thank Professor Peter Forsyth who supervised this thesis. At each struggling point in my research, I benefited tremendously from discussing with Peter. During these years, I am always amazed by his high-quality and high-efficiency work, encouraged by his enthusiasm on research, and inspired by his sharp thought. I am very grateful to him for this opportunity and for his time, guidance, patience.

I thank my committee members, Justin Wan, George Labahn and Ken Vetzal, for taking the time to review the thesis and give me valuable comments.

I would like to thank my external examiner, Matt Davison, for agreeing to serve on my dissertation committee.

I also benefit a lot from academic discussions with fellow schoolmates. Particular mention must be made of Stephen, Cui Zhenyu and Parsiad. Friendships from SciCom members, especially Zhang Haofan, Nian Ke, Eddie, Ken, Ryan, Amir, Ad, Xi Jiong and Han Dong, have given me many memorable moments in my five years in Waterloo.

My parents, Ma Jianping and Liu Shuzhen, have always supported and been behind me. They give me the freedom to pursue my goals. Last but not least, I thank my fiancee, Wang Hongning, for her love and encouragement.

Table of Contents

List of Tables **viii**

List of Figures **x**

1 Introduction **1**

1.1 Overview 1

1.1.1 Uncertain Volatility Model 2

1.1.2 Pre-commitment mean variance asset allocation 2

1.1.3 Uncertain Local Volatility 3

1.2 Contributions 4

1.3 Outline 6

2 An Unconditionally Monotone Numerical Scheme for the Two Factor Uncertain Volatility Model **8**

2.1 Formulation 8

2.2 Restriction of control set Z 9

2.3 Discretization 11

2.3.1 The fixed point stencil 12

2.3.2 Local coordinate rotation: the wide stencil 13

2.3.3 Boundary conditions 15

2.3.4 Avoid using points below the lower boundaries 16

2.3.5 Factoring the diffusion tensor 17

2.3.6 Maximal use of a fixed point stencil 18

2.3.7 Discretization form 20

2.3.8 The matrix form of the discrete equations 22

2.4 Numerical results 24
2.5 Summary 28

3 Convergence to the Viscosity Solution 32
3.1 Viscosity solution for the localized problem 32
3.2 Consistency 34
3.3 Stability 37
3.4 Monotonicity 38
3.5 Convergence 39
3.6 Summary 39

4 Solution of the Nonlinear Discrete Algebraic Equations 40
4.0.1 Convergence of the policy iteration 42
4.1 The optimal value for $\hat{Q}_\ell^k$ 44
4.2 Complexity: comparison of implicit and explicit methods 46
4.3 Summary 47

5 Numerical Solution for Continuous Time Mean Variance Asset Allocation under Stochastic Volatility 48
5.1 Mathematical formulation 48
5.1.1 Efficient frontiers and embedding methods 49
5.1.2 The value function problem 51
5.1.3 The expected wealth problem 52
5.1.4 Allowable portfolios 52
5.2 Numerical discretization of the HJB equation 53
5.2.1 Localization 53
5.2.2 Discretization 56
5.2.3 Matrix form of the discrete equations 59
5.2.4 Convergence to the viscosity solution 61
5.2.5 Policy iteration 62
5.3 Implementation details 63
5.3.1 Complexity 63

5.3.2 The efficient frontier . . . 63
5.3.3 Outside the computational domain . . . 64
5.3.4 An improved linear interpolation scheme . . . 65
5.4 Numerical experiments . . . 66
5.4.1 Effects of the improved interpolation scheme for the PDE method . . 67
5.4.2 Convergence analysis . . . 68
5.4.3 Sensitivity of Efficient Frontiers . . . 69
5.4.4 Comparison between constant volatility and stochastic volatility cases 72
5.5 Summary . . . 77

6 Option Pricing with Local Volatility Uncertainty 78
6.1 The mathematical formulation . . . 78
6.2 The local volatility bounds . . . 80
6.3 Lagrangian uncertain volatility model . . . 82
6.3.1 Solution of the optimization problem (6.17) . . . 84
6.4 Numerical results . . . 84
6.4.1 Local volatility bound construction . . . 84
6.4.2 Basket option pricing with uncertain local volatility . . . 85
6.4.3 Static hedging . . . 89
6.5 Summary . . . 92

7 Conclusions and Future Work 93
7.1 Future work . . . 94

APPENDICES 94

A Discrete equation coefficients in the fixed stencil 95

B The discretized equation for the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{w^*}$ 97

C Proof of the local consistency of $L^Q_{w^*}$ 99

D The discrete linear operator D^p_h 102

E Datasets 103

References 104

List of Tables

2.1 The domain definitions. 20

2.2 Model parameters for the max of two asset call option. 26

2.3 Convergence results for an at-the-money European call option with the payoff (2.50) and parameters as given in Table 2.2. 26

2.4 Convergence results for an at-the-money European call option with the payoff (2.50) and parameters as given in Table 2.2. 27

2.5 The test case of a European call option on the maximum of two assets. . . . 27

2.6 Convergence results for a worst-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). 28

2.7 Convergence results for a worst-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). 29

2.8 Convergence results for a best-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). 29

2.9 Convergence results for a best-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). 30

2.10 The test case for a worst-case (short) butterfly option on maximal of two assets. 30

2.11 Option values for various parameter choices with a butterfly payoff. 31

5.1 The domain definitions. 60

5.2 Parameter values in the Heston model 66

5.3 Input parameters for the mean variance portfolio allocation problem. 67

5.4 Grid and timestep refinement levels used during numerical tests. 67

5.5 The convergence table for the PDE method. Small standard deviation case with $\gamma = 540$. 69

5.6 The convergence table for the Hybrid method. Small standard deviation case with $\gamma = 540$. 70

5.7 The convergence table for the PDE method. Large standard deviation case with $\gamma = 1350$. 71

5.8 The convergence table for the Hybrid method. Large standard deviation case with $\gamma = 1350$. 71

5.9 Given a γ, the optimal portfolio allocation strategy is computed and stored assuming a control process, which is either GBM or stochastic volatility. . . . 76

6.1 The model Parameters. 89

6.2 The maximal and minimal prices of the basket call options with respect to strike price K and maturity T. 89

6.3 The prices of the European basket call options using the *best-fit* local volatility surfaces with respect to strike price K and maturity T. 90

6.4 Spread values for uncertain volatility pricing for a basket option without static hedging, and optimal static hedging with one asset call options 91

6.5 The optimal hedging positions. 91

E.1 For the S&P 500 dataset on October, 1995 103

E.2 Market implied volatilities for FTSE-100 index used in [49, 42]. 103

List of Figures

2.1 The seven-point stencil for $\rho \geq 0$ and $\rho < 0$. The seven points used in the stencil depend on the sign of ρ. 12

2.2 The wide stencil method based on local coordinate rotation. 15

2.3 The domain descriptions. 21

5.1 Close-up of efficient frontier for small standard deviations. (a) No Special Interpolation. (b) Special interpolation. 68

5.2 Convergence of frontiers in the PDE method and the Hybrid method. 70

5.3 Sensitivity analysis of the efficient frontiers with respect to different leverage constraints $p_{\max}$. 72

5.4 Sensitivity analysis of the efficient frontiers with respect to different risk premium factor ξ values. 73

5.5 Sensitivity analysis of the efficient frontiers with respect to different mean reversion level θ values. 73

5.6 Sensitivity analysis of the efficient frontiers with respect to different volatility of the variance σ values. 74

5.7 Sensitivity analysis of the efficient frontiers with respect to different correlation ρ values. 74

5.8 Sensitivity analysis of the efficient frontiers with respect to different mean variance speed κ. 75

5.9 Efficient Frontier Comparison between constant volatility and stochastic volatility cases. 76

6.1 For S&P500 dataset in Table E.1: using MCMC Metropolis Sampling, 249 surfaces from the posterior distribution were tracked. 86

6.2 The upper and lower volatility bounds for the S&P500 dataset in Table E.1. 86

6.3 The *best-fit* local volatility surface for the S&P500 dataset in Table E.1. . . . 87

6.4 The upper and lower volatility bounds for the FTSE-100 dataset in Table E.2. 87

6.5 The *best-fit* local volatility surface for the S&P500 dataset in Table E.1. 88

6.6 The terminal payoff of the unhedged basket option and the residual payoff with optimal static hedge (for the minimal price). 92

C.1 The region Ω_b. 99

Chapter 1

Introduction

1.1 Overview

A key sufficient requirement for ensuring convergence to the viscosity solution of multidimensional Hamilton-Jacobi-Bellman (HJB) equations is that the discretization be monotone [6, 7]. In this thesis, we are particularly interested in optimal stochastic control problems in the financial market where the control appears in the diffusion tensor. In this case, construction of a monotone scheme is a non-trivial matter. In general, no fixed stencil finite difference scheme can produce a monotone scheme for arbitrary multidimensional diffusion tensors [33]. To ensure monotonicity for problems with non-constant diffusion tensors, first order wide stencil methods have been suggested. That is, the stencil increases in size (relative to the node spacing) as the grid is refined. Previous work has focused on explicit wide stencil schemes [14, 30]. In this thesis, we focus on fully implicit methods (hence avoiding timestep restrictions due to stability considerations).

As to the discretization method, we will primarily use a wide stencil based on a local coordinate rotation. An alternative approach is based on factoring the diffusion tensor. This idea of factoring the diffusion tensor, has a long history in stochastic control, see for example [56, 17, 52]. For a recent overview of these methods, we refer the reader to [30]. Another variant of the wide stencil method is discussed in [14, 13]. However, as noted in [30], the computational complexity of computing the coefficients of the wide stencil technique in [14] is quite large, which leads to problems if the coefficients need to be recomputed at every node and every policy iteration (as would be required in our implicit approach).

As a concrete application of our method, we firstly develop the numerical scheme for pricing two-factor contingent claims under the well known uncertain volatility model. However, the reader should have no difficulty applying the techniques in this paper to other optimal stochastic control problems formulated as HJB equations. Secondly, we modify this scheme to study the pre-commitment mean variance asset allocation under stochastic volatility. Lastly, we provide a practical approach to value options with volatility uncertainty when calibrating local volatility values to market quotes.

1.1.1 Uncertain Volatility Model

The uncertain volatility model was first independently developed by Lyons [55] and Avellaneda *et al.* [5]. In this case, the volatility of the risky asset is assumed to lie within a range of values. As such, prices obtained under a no-arbitrage analysis are no longer unique. All that can be computed are the best-case and worst-case prices, for a specified long or short position. By assuming the worst case, an investor can hedge his/her position and obtain a non-negative balance in the hedging portfolio, regardless of the actual volatility movement, provided that volatility remains within the specified range.

Several studies have already considered the uncertain volatility for one factor problems. A European call option with transaction costs and uncertain volatility is considered in [32]. Barrier options under uncertain volatility were studied in [3] and [16], as well as American options and a portfolio of uncertain volatility options. American options were also studied in [66]. Using market bid-ask spreads, an uncertain volatility calibration method was suggested in [20]. A fully implicit PDE scheme is developed for discretely observed barrier options in [40]. These studies are all based on numerical solution of the HJB equations.

In the one-dimensional (single factor) case, it has been shown in [61] that seemingly reasonable discretizations of the uncertain volatility PDE may not converge to the viscosity solution, which is the financially relevant solution. Consequently, it is important to ensure that the numerical scheme is l_∞ stable, consistent in the viscosity sense, and monotone. These properties guarantee convergence to the viscosity solutions [7].

Two factor uncertain volatility models were discussed in [62], however, the scheme was not guaranteed to be monotone. The main difficulty in constructing compact multi-dimensional monotone schemes is the presence of the mixed derivative term, which appears in any case where there is a non-zero correlation between the two underlying assets. In certain cases, monotone schemes can be constructed for very restrictive grid spacing conditions and for certain classes of diffusion tensors [58], but this approach is not very general. In this thesis, we develop an unconditionally monotone numerical scheme for the two factor uncertain volatility model.

1.1.2 Pre-commitment mean variance asset allocation

Consider the following prototypical asset allocation problem: an investor can choose to invest in a risk free bond, or a risky asset, and can dynamically allocate wealth between the two assets, to achieve a pre-determined criteria for the portfolio over a long time horizon. In the continuous time mean variance approach, risk is quantified by variance, so that investors aim to maximize the expected return of their portfolios, given a risk level. Alternatively, they aim to minimize the risk level, given an expected return. As a result, mean variance strategies are appealing due to their intuitive nature, since the results can be easily interpreted in terms of the trade-off between the risk and the expected return.

In the case where the asset follows a Geometric Brownian Motion (GBM), there is considerable literature on the topic [53, 10, 76, 73]. The multi-period optimal strategy adopted in these papers is of pre-commitment type, which is not time-consistent as noted in [11, 8]. A comparison between time-consistent and pre-commitment strategies is given in [74], for continuous time mean variance optimization. We note that since a time consistent strategy can be constructed from a pre-commitment policy by adding a constraint [74], the time consistent strategy is sub-optimal compared to the pre-commitment policy, i.e., it is costly to enforce time consistency. In addition, it has been shown in [71] that pre-commitment strategies can also be viewed as a target-based optimization which involves minimizing a quadratic loss function. It is suggested in [71] that this is intuitive, adaptable to investor preferences, and is also mean variance efficient.

Most previous literature on pre-commitment mean variance optimal asset allocation has been based on analytic techniques [53, 76, 10, 75, 57]. These papers have primarily employed martingale methods [10, 75, 57] or tractable auxiliary problems [53, 76]. However, in general, if realistic constraints on portfolio selection are imposed, e.g., no trading if insolvent and a maximum leverage constraint, then a fully numerical approach is required. As shown in [72], in the case where the risky asset follows a GBM, realistic portfolio constraints have a significant effect on the efficient frontier.

Another modeling deficiency in previous work on pre-commitment mean variance optimal asset allocation is the common assumption that the risky asset follows a GBM. However, there is strong empirical evidence that asset return volatility is serially correlated, shocks to volatility are negatively correlated with asset returns, and the conditional variance of asset returns is not constant over time. As a result, it is highly desirable to describe the risky asset with a stochastic volatility model. In this case, the standard formulation of mean variance optimal asset allocation problems gives rise to a two-dimensional non-linear HJB PDE. In this thesis, we develop a numerical method for the pre-commitment mean variance portfolio selection problem when the underlying risky asset follows a stochastic volatility model.

1.1.3 Uncertain Local Volatility

Assume that the underlying asset satisfies the following stochastic differential equation under the risk-neutral measure

$$\frac{dS}{S} = r dt + \sigma(S,t) dW, \tag{1.1}$$

where W is a standard Brownian motion, and $\sigma(S,t)$ is a deterministic function that varies with both asset price S and time t. From Dupire's formula [35], the local volatility function can be explicitly and uniquely determined from European call prices and their sensitivities with respect to strike and maturity. However, this requires European call options for a continuum of strikes and maturities. Market European option prices are typically limited to a relatively few different strikes and maturities. Hence, this requirement is not met in practice. This makes the problem of determining the local volatility function a well-known ill-posed problem. Furthermore, the prices are only observed to within bid-ask spreads so there exists a noise component in the market data.

Many computational procedures, e.g., [22, 25], are developed to address the difficulties of calibrating a local volatility surface. All these techniques are based on similar ideas: loss functions, penalty functions and minimisation techniques. A loss function measures the difference between the market prices and the prices generated by the calibrated model. The penalty function imposes a regularised constraint to deal with the ill-posed calibration problem. Finally, the calibration problem is formulated as a minimisation problem where the objective function is involved in both the loss and penalty function. A *best-fit* local volatility model is then determined by the calibration procedure.

Choosing a pre-specified calibration method restricts the solution to a more well-behaved class, but the resulting solution does not contain information on its uncertainty. A calibration method depends on the discretization of the local volatility function; the choice of the loss function (absolute error or mean squared error); the choices of the weights for the calibrated instrument in the loss function; and the regularising assumptions. All these components play crucial roles in the calibration procedure. There is no reason to believe that this *best-fit* model is the best approximation of the underlying process. There are infinitely many models which can reproduce the market prices.

Realizing the shortcoming of the *best-fit* approach, another approach is to no longer focus on finding a *best-fit* solution, but finding an entire distribution of solutions. Thus, instead of calibrating the local volatility function which best replicates market prices, all models are tracked which replicate market prices to within an acceptable tolerance level of error. Using an evolutionary algorithm, Hamida and Cont [44] construct a family of local volatility models compatible with a given set of European call bid and ask prices. Gupta [41] recasts the calibration problem into a Bayesian framework. This approach makes it possible to assess the stability of the estimators by observing how the Bayesian posterior distribution varies with small changes in the prior and calibration prices.

In this thesis, we adapt the Bayesian framework [41] to determine the volatility bound for the local volatility function. In contrast to constant lower and upper volatility bounds in [4, 55], the range of the volatilities in our framework is bounded by state-dependent functions. We will refer to this model as uncertain local volatility to differentiate this model from the classic uncertain volatility model. However, the spread between minimal and maximal prices produced by this approach is sometimes so large as to make price bounds somewhat unrealistic. To address this issue, static hedging can be used to decrease the spread between the minimal and maximal prices.

1.2 Contributions

The major contributions of this thesis with regard to numerical solution for the two factor uncertain volatility model are as follows.

- A fully implicit, consistent, unconditionally monotone numerical scheme is first developed for a two factor uncertain volatility model. This method is based on a local coordinate rotation. The discretized algebraic equations are solved using policy iteration. Our discretization method results in a local objective function which can be a discontinuous function of the control. Hence some care must be taken when applying policy iteration [48]. Since we use implicit timestepping, there are no time step restrictions due to stability considerations, an advantage over the method in [30].

- Each policy iteration requires solution of an unstructured sparse M-matrix at each iterate. Since the stencil potentially changes at each policy iteration, this means that the data structure of the sparse matrix must be recomputed at each policy iteration. In this thesis, we use a preconditioned **Bi-CGSTAB** iterative method for solving the sparse matrix [64]. However, the cost of constructing the data structure and solving the matrix is in fact negligible in comparison to the cost of solving the local optimization problem at each grid node. Assuming that the number of policy iterations is bounded as the mesh size tends to zero (which is in fact observed experimentally) the fully implicit method has essentially the same complexity per step as the explicit method in [30].

- To make the numerical scheme consistent, the stencil length is increased to the order of $\sqrt{h}$, where h is the mesh discretization parameter. When solving PDEs on bounded domains, the naive numerical scheme may require points outside the computational domain. We propose a simple method to avoid this problem, which retains consistency.

- A monotone scheme is constructed by factoring the diffusion tensor in [30]. We compare this approach to the method based on a local coordinate system rotation. Although both of these wide stencils are first order, our numerical experiments indicate that the use of the locally rotated coordinate system seems to perform better than constructing a local coordinate system based on factoring the diffusion tensor.

- We prove that these methods converge to the viscosity solution of the uncertain volatility HJB equations.

- We also derive a hybrid numerical scheme that combines use of a fixed stencil [19, 58] and a wide stencil. The fixed stencil is a second-order approximation (for smooth test functions), but this discretization cannot ensure monotonicity at every node in general. We propose an algorithm which uses the fixed stencil as much as possible to take advantage of its accuracy and computational efficiency, while still keeping the numerical scheme monotone. This can be viewed as the multi-dimensional generalization of the standard "central differencing as much as possible" scheme in one dimension [72]. Our tests show that this hybrid technique is generally more smoothly convergent and more accurate than a pure wide stencil scheme. Note that use of an explicit scheme coupled with the hybrid discretization would not result in a practical method, due to the small timesteps required for stability.

The major contributions of this thesis with regard to numerical solution for continuous time mean variance asset allocation under stochastic volatility are as follows.

- A fully implicit, consistent, unconditionally monotone numerical scheme is developed for the HJB equation, which arises in the embedding formulation [76, 53] of the pre-commitment mean variance problem under our model set-up.
- Accurate efficient frontiers are constructed by using a semi-Lagrangian timestepping method to handle the drift term, and an improved method of linear interpolation at the foot of the characteristic in the semi-Lagrangian discretization. In particular, the improved interpolation method uses the exact solution value at a single point, dramatically increasing the accuracy of the numerical results. Any type of constraint can be applied to the investment policy.
- We prove that the scheme developed in this thesis converges to the viscosity solution of the nonlinear HJB value equation.
- In order to trace out the efficient frontier solution of our problem we use two techniques: the PDE method and the Hybrid (PDE - Monte Carlo) method [68]. We also demonstrate that the Hybrid method is superior to the PDE method.
- We carry out several numerical experiments, and illustrate the convergence of the numerical scheme, as well as the effect of modeling parameters on efficient frontiers.

The major contributions of this thesis with regard to pricing options with uncertain local volatility are as follows.

- We provide a practical framework to valuing options with volatility uncertainty when calibrating the local volatility function to the market quotes.
- By using the approach of reformulating the local volatility calibration problem into a Bayesian framework, we construct a reasonable pair of the lower and upper bounds for the local volatility values.
- We demonstrate how the use of an optimal static hedge can reduce model calibration risk.

1.3 Outline

Chapters 2 to 4 develop a numerical scheme for the two-factor uncertain volatility model. Chapter 2 describes the numerical scheme, and shows numerical results. In chapter 3 we prove that our numerical scheme guarantees convergence to the viscosity solution. Chapter

4 discusses the policy iteration algorithm used to solve the discretized algebraic equations using implicit timestepping.

Chapter 5 studies numerical solution of the HJB formulation for continuous time mean variance asset allocation under stochastic volatility. Chapter 6 discusses the uncertain local volatility model. Chapter 7 concludes.

Chapter 2

An Unconditionally Monotone Numerical Scheme for the Two Factor Uncertain Volatility Model

In this chapter, a fully implicit, consistent, unconditionally monotone numerical scheme is first developed for a two factor uncertain volatility model. This chapter is organized as follows: Section 2.1 gives a formulation of an associated HJB equation for the two factor option pricing problem under uncertain volatility model. We discuss the optimization problem involved in the HJB equation in Section 2.2. In Section 2.3, we detail the discretization of the HJB equation. Numerical results are presented and discussed in Section 2.4. This chapter focuses on a presentation of the scheme, and numerical tests. Proofs of convergence to the viscosity solution are postponed to Chapter 3.

2.1 Formulation

Let $\mathcal{U}(S_1, S_2, \tau)$ be the value of a European option contract written on asset prices S_1 and S_2, which both follow the stochastic processes under the risk neutral measure

$$\begin{aligned} \mathrm{d}S_1 &= (r - q_1)S_1\mathrm{d}t + \sigma_1 S_1 \mathrm{d}W_1, \\ \mathrm{d}S_2 &= (r - q_2)S_2\mathrm{d}t + \sigma_2 S_2 \mathrm{d}W_2, \end{aligned} \tag{2.1}$$

where r is the risk-free interest rate, $q_i,\ i = 1, 2$ are the dividend yields for S_i. $\sigma_i,\ i = 1, 2$ are volatilities, and $W_i,\ i = 1, 2$ are Wiener processes with $\mathrm{d}W_1\mathrm{d}W_2 = \rho\ \mathrm{d}t$. Following standard arguments, the underlying PDE for the no-arbitrage price of the option contract $\mathcal{U}(S_1, S_2, \tau)$ is given by

$$\begin{aligned} \mathcal{U}_\tau &= \mathcal{L}\mathcal{U}, \\ \mathcal{U}(S_1, S_2, 0) &= \mathcal{W}(S_1, S_2), \end{aligned} \tag{2.2}$$

which is defined over $(S_1, S_2, \tau) \in [0, +\infty] \times [0, +\infty] \times [0, T]$. $\mathcal{W}(S_1, S_2)$ is the terminal payoff of the option contract. The linear differential operator $\mathcal{L}$ is defined as

$$
\begin{aligned}
&\mathcal{L}\mathcal{U} = \mathbf{V} \cdot \nabla\mathcal{U} + (\mathbf{D}\nabla) \cdot \nabla\mathcal{U} - r\mathcal{U}, \\
&\mathbf{D} \in \mathbb{R}^2 \times \mathbb{R}^2; \quad \mathbf{V} \in \mathbb{R}^2; \\
&\nabla = \begin{pmatrix} \frac{\partial}{\partial S_1} \\ \frac{\partial}{\partial S_2} \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} (r - q_1)S_1 \\ (r - q_2)S_2 \end{pmatrix}, \quad \mathbf{D} = \frac{1}{2} \begin{pmatrix} \sigma_1^2 S_1^2 & \rho\sigma_1\sigma_2 S_1 S_2 \\ \rho\sigma_1\sigma_2 S_1 S_2 & \sigma_2^2 S_2^2 \end{pmatrix},
\end{aligned}
\tag{2.3}
$$

where ∇ is the gradient operator, $\mathbf{V}$ is the drift tensor, and $\mathbf{D}$ is the diffusion tensor. Note that the notation $(\mathbf{D}\nabla) \cdot \nabla\mathcal{U}$ is to be interpreted as

$$
(\mathbf{D}\nabla) \cdot \nabla\mathcal{U} = \frac{\sigma_1^2 S_1^2}{2}\, \mathcal{U}_{S_1 S_1} + \rho\sigma_1\sigma_1 S_1 S_2\, \mathcal{U}_{S_1 S_2} + \frac{\sigma_2^2 S_2^2}{2}\, \mathcal{U}_{S_2 S_2}\ . \tag{2.4}
$$

Now we generalize PDE (2.2) to the uncertain volatility model that was first developed in [5] and [55]. That is, σ_i is an uncertain volatility in the processes (2.1), but lies within a range, e.g., $\sigma_1 \in [\sigma_{1,\min}, \sigma_{1,\max}]$ and $\sigma_2 \in [\sigma_{2,\min}, \sigma_{2,\max}]$. In addition, uncertain correlation between the two underlying assets is permitted, e.g., $\rho_{\min} \leq \rho \leq \rho_{\max}$. When the volatilities σ_1, σ_2, and the correlation ρ are uncertain, the diffusion tensor in the PDE (2.2) for the price of an option contract is no longer unique. Nevertheless, the maximal and minimal solution values of an option contract are given by the following Hamilton-Jacobi-Bellman (HJB) PDEs

$$
\mathcal{U}_\tau = \mathbf{V} \cdot \nabla\mathcal{U} + \sup_{Q \in Z} \left((\mathbf{D}\nabla) \cdot \nabla\mathcal{U} \right) - r\mathcal{U}, \tag{2.5a}
$$

$$
\mathcal{U}_\tau = \mathbf{V} \cdot \nabla\mathcal{U} + \inf_{Q \in Z} \left((\mathbf{D}\nabla) \cdot \nabla\mathcal{U} \right) - r\mathcal{U}, \tag{2.5b}
$$

where the control $Q = (\sigma_1,\ \sigma_2,\ \rho)$, and the admissible set of the controls is given by

$$
\begin{aligned}
Z = [\sigma_{1,\min}, \sigma_{1,\max}] \times [\sigma_{2,\min}, \sigma_{2,\max}] \times [\rho_{\min}, \rho_{\max}], \\
\sigma_{1,\min} \geq 0, \quad \sigma_{2,\min} \geq 0, \quad -1 \leq \rho_{\min} \leq 1, \quad -1 \leq \rho_{\max} \leq 1.
\end{aligned}
\tag{2.6}
$$

We can consider equation (2.5a) to be the worst case cost of hedging a short position in the contract, while equation (2.5b) is the worst case cost of hedging a long position. Without loss of generality, we only consider the sup problem in the following discussion. All the results of the thesis hold in the inf case as well.

2.2 Restriction of control set Z

Before we introduce our discretization method, we take a short digression here to discuss the maximization of the right hand side of equation (2.5a). We consider (for the time being) that all the derivatives which appear on the right hand side of equation (2.5a) are constructed from known, smooth functions. Since consistency in the viscosity sense is defined in terms of smooth test functions [7], this will be relevant to our discretization approach.

To maximize the solution value in equation (2.5a), it suffices to maximize the diffusion terms. Let $\Gamma_{kl} \equiv \frac{\partial^2 \mathcal{U}}{\partial S_k \partial S_l}$, $k, l = 1, 2$. Assume for the moment that Γ_{kl} is constant, independent of the control. In this notation, the diffusion terms in (2.5a) become

$$\sup_{Q \in Z} \left((\mathbf{D}\nabla) \cdot \nabla \mathcal{U} \right) = \max_{(\sigma_1, \sigma_2, \rho) \in Z} \left(\frac{\sigma_1^2 S_1^2}{2} \Gamma_{11} + \rho \sigma_1 \sigma_2 S_1 S_2 \Gamma_{12} + \frac{\sigma_2^2 S_2^2}{2} \Gamma_{22} \right). \tag{2.7}$$

Since Z (2.6) is a compact set, the supremum is simply the maximum value.

It is easy to see that the optimal correlation value is a bang-bang control. That is, the optimal $\rho \in \{\rho_{\min},\ \rho_{\max}\}$, depends only on the sign of the cross derivative Γ_{12}.

$$\rho\left(\Gamma_{12}\right) = \begin{cases} \rho_{\max} & \Gamma_{12} \geq 0 \\ \rho_{\min} & \Gamma_{12} < 0. \end{cases} \tag{2.8}$$

With ρ given from (2.8), a quadratic-form optimization with linear constraints needs to be solved. The problem is formulated as

$$\max_{\sigma} \sigma^T \mathbf{M} \sigma \equiv \max_{\sigma_1, \sigma_2} \begin{pmatrix} \sigma_1 & \sigma_2 \end{pmatrix} \begin{pmatrix} \frac{S_1^2}{2} \Gamma_{11} & \rho(\Gamma_{12}) \frac{S_1 S_2}{2} \Gamma_{12} \\ \rho(\Gamma_{12}) \frac{S_1 S_2}{2} \Gamma_{12} & \frac{S_2^2}{2} \Gamma_{22} \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}, \tag{2.9}$$

subject to

$$\sigma_{1,\min} \leq \sigma_1 \leq \sigma_{1,\max}, \quad \sigma_{2,\min} \leq \sigma_2 \leq \sigma_{2,\max}. \tag{2.10}$$

Proposition 2.1. *Suppose that Γ_{ik} exist $\forall i, k$. The optimal value of the objective function in (2.7) can be determined by examining values only on the boundary of Z, denoted by ∂Z.*

$$\sup_{Q \in Z} \left((\mathbf{D}\nabla) \cdot \nabla \mathcal{U} \right) = \sup_{Q \in \partial Z} \left((\mathbf{D}\nabla) \cdot \nabla \mathcal{U} \right). \tag{2.11}$$

Proof. From equation (2.8), the choice of the optimal correlation ρ is either $\rho_{\max}$ or $\rho_{\min}$, depending on the sign of the cross derivative term. Thus, the optimal correlation is always either end of its range $[\rho_{\min}, \rho_{\max}]$.

The quadratic form in equation (2.9) is $\sigma^T \mathbf{M} \sigma$. A critical point is such that $\mathbf{M}\sigma = 0$. When $\mathbf{M}$ is a non-singular matrix, the critical point is $(0, 0)$, which is either outside Z or on the boundary of Z. When $\mathbf{M}$ is singular, the critical points are on the line

$$\left\{ (\sigma_1, \sigma_2) \middle| \frac{S_1^2}{2} \Gamma_{11} \sigma_1 + \frac{\rho(\Gamma_{12}) S_1 S_2}{2} \Gamma_{12} \sigma_2 = 0 \right\}.$$

If this line intersects Z, then the optimal value is attained on ∂Z. If this line does not intersect Z, then the optimal value is also on ∂Z. Hence, in all cases, the optimal value can be attained by examining the objective function on ∂Z. □

Remark 2.1. *Proposition 2.1 will prove useful when we design a numerical scheme. In the case when the discretization stencil depends on the control, no analytical expression is available for the optimal value. We can then discretize the control set and search over the boundary ∂Z, instead of the entire three dimensional set Z. Consistency in the viscosity sense is defined in terms of smooth test functions, hence our assumption that $\Gamma_{i,k}$ exist is not restrictive and we can then use Proposition 2.1 to prove that this is a consistent discretization (in the viscosity sense).*

2.3 Discretization

In the thesis, we develop an unconditionally monotone finite difference numerical scheme for the two factor uncertain volatility model. However, a standard finite difference scheme cannot ensure monotonicity due to the cross derivative term. For example, the fixed point stencil method in [58] requires a restrictive grid spacing, which cannot always be satisfied, to preserve monotonicity. In our problem, the diffusion tensor is non-constant and non-diagonally dominant. Monotonicity is formally defined in [7], and we will discuss this in more detail in Chapter 3. For now, it suffices to note that we can develop an unconditionally monotone scheme by choosing a discretization which eliminates the cross derivative term in equation (2.5a).

We will focus mainly on a wide stencil method based on a local coordinate rotation, but we include some comparisons with the factoring technique in [30]. Furthermore, we propose a hybrid algorithm which combines use of a fixed point stencil [19, 58] with a wide stencil. This algorithm uses the fixed point stencil as much as possible to take advantage of its accuracy and computational efficiency, but still keeping the numerical scheme monotone.

We discretize equation (2.5a) over a finite grid $N = N_1 \times N_2$ in the spatial space (S_1, S_2). Define a set of nodes $\{(S_1)_1, (S_1)_2, \ldots, (S_1)_{N_1}\}$ in the S_1 direction and $\{(S_2)_1, (S_2)_2, \ldots, (S_2)_{N_2}\}$ in the S_2 direction. Denote the n^{th} time step by $\tau^n = n\Delta\tau,\ n = 0, \ldots, N_\tau$, with $N_\tau = \frac{T}{\Delta\tau}$. Let $\mathcal{U}^n_{i,j}$ be the approximate solution of the equation (2.5a) at $((S_1)_i, (S_2)_j, \tau^n)$.

It will be convenient to define

$$\begin{aligned}
\Delta(S_1)_{\max} &= \max_i \left((S_1)_{i+1} - (S_1)_i\right), \quad \Delta(S_1)_{\min} = \min_i \left((S_1)_{i+1} - (S_1)_i\right), \\
\Delta(S_2)_{\max} &= \max_i \left((S_2)_{i+1} - (S_2)_i\right), \quad \Delta(S_2)_{\min} = \min_i \left((S_2)_{i+1} - (S_2)_i\right).
\end{aligned} \tag{2.12}$$

We assume that there is a mesh discretization parameter h such that

$$\Delta(S_1)_{\max} = C_1 h,\ \Delta(S_2)_{\max} = C_2 h,\ \Delta(S_1)_{\min} = C_1^{'} h,\ \Delta(S_2)_{\min} = C_2^{'} h,\ \Delta\tau = C_3 h, \tag{2.13}$$

where $C_1, C_2, C_1^{'}, C_2^{'}, C_3$ are constants independent of h.

2.3.1 The fixed point stencil

We use a seven-point stencil [19, 58] to discretize the cross-partial derivative $\frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_2}$. Denote

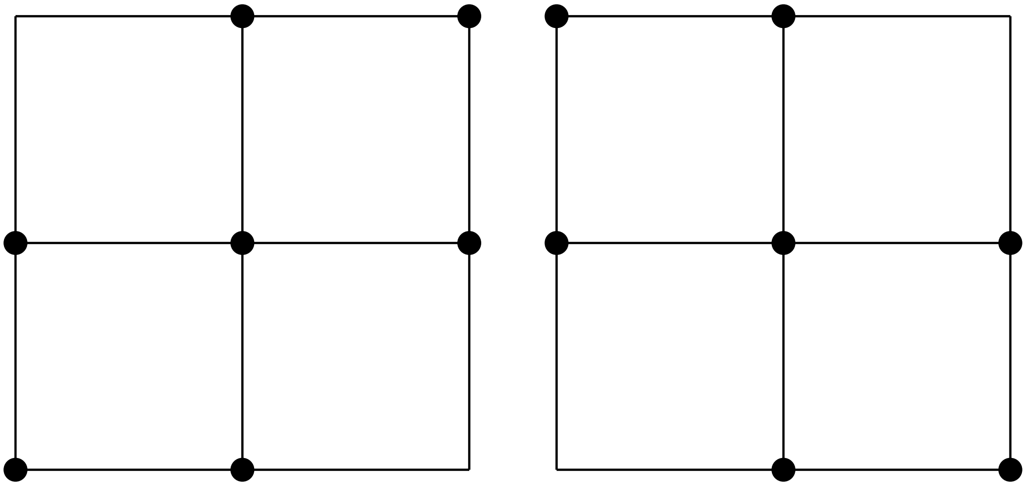
$$\begin{aligned} \Delta^+(S_1)_i &= (S_1)_{i+1} - (S_2)_i, \quad \Delta^-(S_1)_i = (S_1)_i - (S_1)_{i-1}, \\ \Delta^+(S_2)_j &= (S_2)_{j+1} - (S_2)_j, \quad \Delta^-(S_2)_j = (S_2)_j - (S_2)_{j-1}. \end{aligned} \tag{2.14}$$

We approximate the cross-partial derivative at $((S_1)_i, (S_2)_j, \tau^n)$ using one of the following stencils, as illustrated in Figure 2.1, depending on the sign of ρ. For $\rho \geq 0$, we use

$$\frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_2} \approx \frac{2\mathcal{U}^n_{i,j} + \mathcal{U}^n_{i+1,j+1} + \mathcal{U}^n_{i-1,j-1}}{\Delta^+(S_1)_i \Delta^+(S_2)_j + \Delta^-(S_1)_i \Delta^-(S_2)_j} - \frac{\mathcal{U}^n_{i+1,j} + \mathcal{U}^n_{i-1,j} + \mathcal{U}^n_{i,j+1} + \mathcal{U}^n_{i,j-1}}{\Delta^+(S_1)_i \Delta^+(S_2)_j + \Delta^-(S_1)_i \Delta^-(S_2)_j}. \tag{2.15}$$

For $\rho < 0$, we use

$$\frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_2} \approx -\frac{2\mathcal{U}^n_{i,j} + \mathcal{U}^n_{i+1,j-1} + \mathcal{U}^n_{i-1,j+1}}{\Delta^+(S_1)_i \Delta^-(S_2)_j + \Delta^-(S_1)_i \Delta^+(S_2)_j} + \frac{\mathcal{U}^n_{i+1,j} + \mathcal{U}^n_{i-1,j} + \mathcal{U}^n_{i,j+1} + \mathcal{U}^n_{i,j-1}}{\Delta^+(S_1)_i \Delta^-(S_2)_j + \Delta^-(S_1)_i \Delta^+(S_2)_j}. \tag{2.16}$$



(a) $\rho \geq 0$

(b) $\rho < 0$

Figure 2.1: The seven-point stencil for $\rho \geq 0$ and $\rho < 0$. The seven points used in the stencil depend on the sign of ρ.

Standard three point differences are used for the $\frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_1}$ and $\frac{\partial^2 \mathcal{U}}{\partial S_2 \partial S_2}$ terms. First order partial derivatives in (2.5a) are approximated with second order central differencing as much as possible. Algorithm A.1 in Appendix A shows how to select central, forward and backward differencing to minimize the appearance of negative coefficients in the discretization [72]. The linear differential operator $\mathcal{L}$ in (2.5a) is discretized to form the discrete linear operator L^Q_f.

$$\begin{aligned} L^Q_f \mathcal{U}^n_{i,j} = {} & (\alpha^{S_1}_{i,j} - \gamma_{i,j})\mathcal{U}^n_{i-1,j} + (\beta^{S_1}_{i,j} - \gamma_{i,j})\mathcal{U}^n_{i+1,j} + (\alpha^{S_2}_{i,j} - \gamma_{i,j})\mathcal{U}^n_{i,j-1} + (\beta^{S_2}_{i,j} - \gamma_{i,j})\mathcal{U}^n_{i,j+1} \\ & + 1_{\rho \geq 0}(\gamma_{i,j}\mathcal{U}^n_{i+1,j+1} + \gamma_{i,j}\mathcal{U}^n_{i-1,j-1}) + 1_{\rho < 0}(\gamma_{i,j}\mathcal{U}^n_{i+1,j-1} + \gamma_{i,j}\mathcal{U}^n_{i-1,j+1}) \\ & - (\alpha^{S_1}_{i,j} + \beta^{S_1}_{i,j} + \alpha^{S_2}_{i,j} + \beta^{S_2}_{i,j} - 2\gamma_{i,j} + r)\mathcal{U}_{i,j}, \end{aligned} \tag{2.17}$$

where $\alpha_{i,j}^{S_1}, \beta_{i,j}^{S_1}, \alpha_{i,j}^{S_2}, \beta_{i,j}^{S_2}$, and $\gamma_{i,j}$ are defined in Appendix A. The notation L_f^Q indicates that the equation coefficients are functions of the control Q.

The positive coefficient condition [39] is

$$\begin{aligned}
&\alpha_{i,j}^{S_1} - \gamma_{i,j} \geq 0, \quad \beta_{i,j}^{S_1} - \gamma_{i,j} \geq 0, \quad \alpha_{i,j}^{S_2} - \gamma_{i,j} \geq 0, \quad \beta_{i,j}^{S_2} - \gamma_{i,j} \geq 0, \\
&\gamma_{i,j} \geq 0, \quad \alpha_{i,j}^{S_1} + \beta_{i,j}^{S_1} + \alpha_{i,j}^{S_2} + \beta_{i,j}^{S_2} - 2\gamma_{i,j} + r \geq 0, \quad 1 \leq i < N_1, \ 1 \leq j < N_2.
\end{aligned} \tag{2.18}$$

We remind the reader that the positive coefficient condition is sufficient in order to ensure monotonicity [39].

Due to the presence of the $\gamma_{i,j}$ term in (2.17), the discretization does not ensure that the positive coefficient condition (2.18) is satisfied even if our choice of the seven-point operator ensures that $\gamma_{i,j} \geq 0$. However, our algorithm makes the positive coefficient condition hold on as many grid nodes as possible with a fixed stencil. Only when the cross derivative term disappears in the HJB equation (2.5a) can we guarantee that the positive coefficient condition always holds for a fixed point stencil.

Remark 2.2. *It is possible to carry out a logarithmic transformation on equation (2.5a). In the new coordinate system* $(\log S_1, \log S_2)$, *the diffusion tensor becomes constant for a fixed control. If we discretize the PDE on the space* $(\log S_1, \log S_2)$, *a positive coefficient discretization can be constructed for a very restrictive grid spacing condition [19], but this approach is not very general, and the diffusion tensor is not constant if local volatility surfaces are used, which is common in practice. Consequently, we prefer to use the more meaningful discretization in* (S_1, S_2) *coordinates.*

2.3.2 Local coordinate rotation: the wide stencil

We now consider the wide stencil discretization method. Suppose we discretize equation (2.5a) at grid node (i, j) for a fixed control. Consider a virtual rotation of the local coordinate system clockwise by

$$\theta_{i,j} = \frac{1}{2} \tan^{-1} \left(\frac{2\rho\sigma_1\sigma_2(S_1)_i(S_2)_j}{(\sigma_1(S_1)_i)^2 - (\sigma_2(S_2)_j)^2} \right). \tag{2.19}$$

That is, (y_1, y_2) in the transformed coordinate system is obtained by using the following matrix multiplication

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \begin{pmatrix} \cos\theta_{i,j} & -\sin\theta_{i,j} \\ \sin\theta_{i,j} & \cos\theta_{i,j} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{2.20}$$

We denote the rotation matrix in (2.20) as $\mathbf{R}_{i,j}$. This rotation operation will result in a zero correlation in the diffusion tensor of the rotated system. That is, the cross derivative term will be eliminated. Under this grid rotation, the second order terms in equation (2.2) are, in the transformed coordinate system (y_1, y_2),

$$a_{i,j} \frac{\partial^2 \mathcal{V}}{\partial y_1^2} + b_{i,j} \frac{\partial^2 \mathcal{V}}{\partial y_2^2}, \tag{2.21}$$

where $\mathcal{V}$ is the value function $\mathcal{V}(y_1, y_2, \tau)$ in the transformed coordinate system, and

$$
\begin{aligned}
a_{i,j} &= \left(\frac{(\sigma_1 \cos(\theta_{i,j})(S_1)_i)^2}{2} + \rho\sigma_1\sigma_2(S_1)_i(S_2)_j \sin(\theta_{i,j})\cos(\theta_{i,j}) + \frac{(\sigma_2 \sin(\theta_{i,j})(S_2)_j)^2}{2} \right), \\
b_{i,j} &= \left(\frac{(\sigma_1 \sin(\theta_{i,j})(S_1)_i)^2}{2} - \rho\sigma_1\sigma_2(S_1)_i(S_2)_j \sin(\theta_{i,j})\cos(\theta_{i,j}) + \frac{(\sigma_2 \cos(\theta_{i,j})(S_2)_j)^2}{2} \right).
\end{aligned}
\tag{2.22}
$$

The diffusion tensor in (2.21) is diagonally dominant with no off-diagonal terms, and consequently a standard finite difference discretization for the second partial derivatives is a positive coefficient scheme. The rotation angle $\theta_{i,j}$ depends on the grid node and the control, therefore it is impossible to rotate the global coordinate system by a constant angle and build a grid over the space (y_1, y_2). The local coordinate system rotation is only used to construct a virtual grid which overlays the original mesh. We have to approximate the values of $\mathcal{U}$ on our virtual local grid using an interpolant $\mathcal{J}_h\mathcal{U}$ on the original mesh. To keep the numerical scheme monotone, linear interpolation is the most accurate interpolation we can use. Thus, $\mathcal{J}_h$ is a linear interpolation operator. Moreover, to keep the numerical scheme consistent, we need to use the points on our virtual grid whose Euclidean distances are $O(\sqrt{h})$ from the central node, where h is the mesh discretization parameter (2.13). This results in a wide stencil method since the relative stencil length increases as the grid is refined ($\frac{\sqrt{h}}{h} \to \infty$ as $h \to 0$). The wide stencil method is illustrated in Figure 2.2. With a slight abuse of notation, we define the following

$$
\mathcal{U}^n(\mathbf{S}) \equiv \mathcal{U}(S_1, S_2, \tau^n), \quad \mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}, \quad \mathcal{V}^n(\mathbf{y}) \equiv \mathcal{V}(y_1, y_2, \tau^n), \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{2.23}
$$

Then, the second order terms in equation (2.5a) at $((S_1)_i, (S_2)_j, \tau^n)$ are approximated as

$$
\begin{aligned}
& a_{i,j} \frac{\mathcal{V}^n\left(\mathbf{y}_{i,j} + \sqrt{h}\mathbf{e}_1\right) + \mathcal{V}^n\left(\mathbf{y}_{i,j} - \sqrt{h}\mathbf{e}_1\right) - 2\mathcal{V}^n\left(\mathbf{y}_{i,j}\right)}{h} \\
& \quad + b_{i,j} \frac{\mathcal{V}^n\left(\mathbf{y}_{i,j} + \sqrt{h}\mathbf{e}_2\right) + \mathcal{V}^n\left(\mathbf{y}_{i,j} - \sqrt{h}\mathbf{e}_2\right) - 2\mathcal{V}^n\left(\mathbf{y}_{i,j}\right)}{h} \\
& \approx a_{i,j} \frac{\mathcal{J}_h\mathcal{U}^n(\mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_1) + \mathcal{J}_h\mathcal{U}^n(\mathbf{S}_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_1) - 2\mathcal{U}^n(\mathbf{S}_{i,j})}{h} \\
& \quad + b_{i,j} \frac{\mathcal{J}_h\mathcal{U}^n(\mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_2) + \mathcal{J}_h\mathcal{U}^n(\mathbf{S}_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_2) - 2\mathcal{U}^n(\mathbf{S}_{i,j})}{h},
\end{aligned}
\tag{2.24}
$$

where $\mathbf{S}_{i,j} = ((S_1)_i, (S_2)_j)$, $\mathbf{y}_{i,j} = \mathbf{R}_{i,j}^T \mathbf{S}_{i,j}$, $(\mathbf{R}_{i,j})_k$ is k-th column of the rotation matrix $\mathbf{R}_{i,j}$ (2.20), and

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

To satisfy the positive coefficient condition, we then use an upstream finite differencing to discretize the first order derivatives.

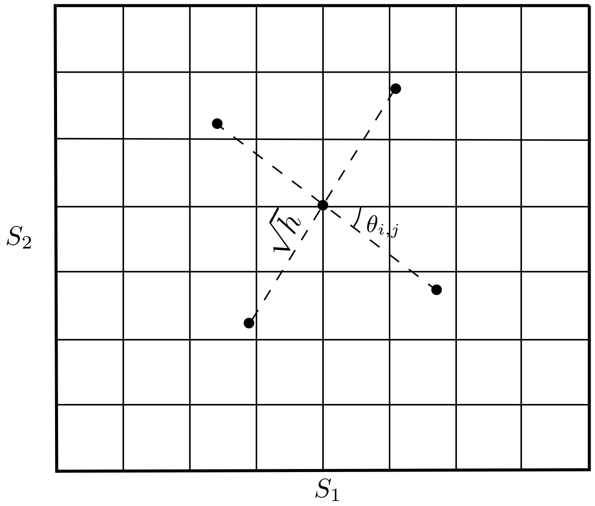


Figure 2.2: The wide stencil method based on local coordinate rotation.

2.3.3 Boundary conditions

We shall assume that the discretization is posed on a bounded domain for computational purposes. The discretization is applied to the localized finite region $(S_1, S_2) \in [0, S_{1,\max}] \times [0, S_{2,\max}]$.

No boundary condition is needed on the lower boundaries $S_1 = 0$ or $S_2 = 0$. The equation (2.2) reduces to

$$\frac{\partial \mathcal{U}}{\partial \tau} = \begin{cases} (r - q_2) S_2 \dfrac{\partial \mathcal{U}}{\partial S_2} + \dfrac{S_2^2 \sigma_2^2}{2} \dfrac{\partial^2 \mathcal{U}}{\partial S_2^2} - r\mathcal{U}, & \text{for } (S_1, S_2) \in \{0\} \times (0, S_{2,\max}), \\ (r - q_1) S_1 \dfrac{\partial \mathcal{U}}{\partial S_1} + \dfrac{S_1^2 \sigma_1^2}{2} \dfrac{\partial^2 \mathcal{U}}{\partial S_1^2} - r\mathcal{U}, & \text{for } (S_1, S_2) \in (0, S_{1,\max}) \times \{0\}, \\ -r\mathcal{U}, & \text{at } (S_1, S_2) = (0, 0). \end{cases} \tag{2.25}$$

The cross derivative term vanishes on the lower boundaries. Thus, we can use a standard finite difference stencil to construct a monotone scheme on the lower boundaries.

In order to preserve monotonicity of the discretization, a Dirichlet boundary condition is imposed on the upper boundaries $S_1 = S_{1,\max}$ or $S_2 = S_{2,\max}$. As pointed out in [6], we can expect any errors incurred by imposing approximate boundary conditions at $S_1 = S_{1,\max}$ or $S_2 = S_{2,\max}$ to be small in areas of interest if $S_{1,\max}$ or $S_{2,\max}$ is sufficiently large. As $S_1 \to \infty$ or $S_2 \to \infty$, we normally use financial reasoning to determine the asymptotic form of the solution. The upper boundary may be approximated by a time-dependent value

$$\mathcal{U}(S_1, S_2, \tau) \approx \bar{\mathcal{U}}(S_1, S_2, \tau) = C_0(\tau) + C_1(\tau) S_1 + C_2(\tau) S_2. \tag{2.26}$$

Equation (2.26) is substituted into equation (2.5a), and expressions for $C_1(\tau)$, $C_2(\tau)$ are easily obtained, since the diffusion terms vanish.

2.3.4 Avoid using points below the lower boundaries

To make the numerical scheme consistent in a wide stencil method, the stencil length needs to be increased to use the points beyond the nearest neighbors of the original grid. As shown in Section 2.3.2, we use the four points $\mathbf{S}_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k,\ k = 1, 2$ in (2.24), when we approximate the second order terms (2.21). Therefore, when solving the PDE on a bounded region, this numerical discretization (2.24) may require points outside the computational domain.

When a candidate point we use is outside the computational region at the upper boundaries, we directly use the asymptotic solution as specified in (2.26) at the point. However, we have to take special care when we may use a point below the lower boundaries $S_1 = 0$ or $S_2 = 0$. The possibility of using points below the lower boundaries only occurs when the node (i, j) falls in the region

$$[h, \sqrt{h}] \times (0, S_{2,\max}] \cup (0, S_{1,\max}] \times [h, \sqrt{h}]. \tag{2.27}$$

We propose a simple method to avoid this problem, which retains consistency. That is, when one of the four candidate points $\mathbf{S}_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k,\ k = 1, 2$ is below the lower boundaries, we then shrink its corresponding distance to h, instead of $\sqrt{h}$. This treatment ensures that all data required is within the computational domain. The details of the method are given in Algorithm 2.1. We will formally prove that this simple method retains consistency in Chapter 3.

Algorithm 2.1 Avoid using the points below the lower boundaries when approximating the $\frac{\partial^2 \mathcal{V}}{\partial y_k^2},\ k = 1, 2$

```
Let $\mathbf{S}_{k,left} = \mathbf{S}_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_k$ and $h_{k,left} = \sqrt{h}$
if $\mathbf{S}_{k,left}$ below the lower boundaries then
    $h_{left} = h$
end if
Let $\mathbf{S}_{k,right} = \mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_k$ and $h_{k,right} = \sqrt{h}$
if $\mathbf{S}_{k,right}$ below the lower boundaries then
    $h_{right} = h$
end if
The second derivative term $\frac{\partial^2 \mathcal{V}}{\partial y_k^2}$ at $\mathbf{y}_{i,j} = \mathbf{R}_{i,j}^T \mathbf{S}_{i,j}$ are approximated as
    $$\frac{\frac{\mathcal{J}_h \mathcal{U}\left(\mathbf{S}_{i,j} - h_{k,left}(\mathbf{R}_{i,j})_k\right) - \mathcal{U}(\mathbf{S}_{i,j})}{h_{k,left}} + \frac{\mathcal{J}_h \mathcal{U}\left(\mathbf{S}_{i,j} + h_{k,right}(\mathbf{R}_{i,j})_k\right) - \mathcal{U}(\mathbf{S}_{\mathbf{i},\mathbf{j}})}{h_{k,right}}}{\frac{h_{k,left} + h_{k,right}}{2}}. \tag{2.28}$$
```

2.3.5 Factoring the diffusion tensor

In [30], the wide stencil method based on factoring the diffusion tensor is surveyed. For the convenience of the reader, we briefly summarize this method here. For more details we refer readers to [30]. Let the diffusion tensor in (2.2) be

$$\mathbf{D} = \frac{1}{2}\mathbf{C}^T\mathbf{C}.$$

Then, the second order terms in (2.5a) are approximated as

$$((\mathbf{D}\nabla)\cdot\nabla\mathcal{U}) \approx \frac{1}{2}\Bigg(\frac{\mathcal{J}_h\mathcal{U}(\mathbf{S}+\sqrt{h}\mathbf{C}_1)+\mathcal{J}_h\mathcal{U}(\mathbf{S}-\sqrt{h}\mathbf{C}_1)-2\mathcal{U}(\mathbf{S})}{h} + \frac{\mathcal{J}_h\mathcal{U}(\mathbf{S}+\sqrt{h}\mathbf{C}_2)+\mathcal{J}_h\mathcal{U}(\mathbf{S}-\sqrt{h}\mathbf{C}_2)-2\mathcal{U}(\mathbf{S})}{h}\Bigg) + O(h), \tag{2.29}$$

where $\mathbf{C}_k$ is *k-th* column of $\mathbf{C}$. From the stochastic processes of the two asset prices (2.1), it is natural to choose

$$\mathbf{C} = \begin{pmatrix} \sigma_1 S_1 & 0 \\ \sigma_2\rho S_2 & \sigma_2\sqrt{1-\rho^2}S_2 \end{pmatrix}.$$

That is, $\mathbf{C}$ is the lower triangular matrix associated with the Cholesky decomposition of the diffusion tensor.

This consistent approximation is also a first order approximation and compatible with a monotone numerical scheme. Although the defining ideas, between this method and the local coordinate rotation introduced in Section 2.3.2, are different, we can relate them by re-interpreting the approximation (2.29). Firstly, we virtually transform the coordinate system as follows:

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \mathbf{C}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{2.30}$$

This transformation will result in a zero correlation in the diffusion tensor of the transformed system. After applying this local virtual coordinate transformation, we then construct a local discretization in a manner similar to the method used for the rotation method in Section 2.3.2. The transformation (2.30) is both a stretching and rotation of the coordinate system, not an orthogonal rotation (2.20) as in Section 2.3.2. Thus, in (2.29), we shall use points whose Euclidean distance from (S_1, S_2) are $|\sqrt{h}\mathbf{C}_k|,\ k = 1, 2$, which is state dependent on S_1 and S_2. For example, the points we use may be *far* away from the central node (i, j), especially when the grid state $(S_1)_i$ or $(S_2)_j$ is large. However, as noted in [14] and [52], it is highly desirable to limit the use of points that are *far* away from the central node. When we use the method of locally rotating coordinate system, the candidate points are always $\sqrt{h}|(\mathbf{R}_{i,j})_k| = \sqrt{h}$ away from the central node. In our numerical experiments, we will compare the performance of these two methods.

2.3.6 Maximal use of a fixed point stencil

We will derive a hybrid scheme which combines use of the fixed point stencil (Section 2.3.1) with the wide stencil based on a local coordinate rotation (Section 2.3.2). The fixed point stencil is a second-order approximation of the diffusion terms, but this discretization cannot ensure a positive coefficient method at every node in general. The computational cost is also highly increased when we use a wide stencil. This is due to the fact that we have an analytical solution for the local optimization problem for the fixed point stencil case. On the other hand, when using a wide stencil, we need to discretize the control set and then perform a linear search to find the optimal value for the control. Furthermore, the fixed point stencil uses seven neighbor nodes for a reference node, but the wide stencil method potentially uses as many as twenty nodes. Thus, when using the implicit timestepping method with solving sparse matrices, the fixed point stencil is more computational efficient than the wide stencil method. We propose an algorithm which uses the fixed point stencil as much as possible to take advantage of its accuracy and computational efficiency, while still satisfying the positive coefficient condition. Note that our algorithm is also applicable if we factor the diffusion tensor, as in [30].

Lemma 2.1. *The positive coefficient condition (2.18) for a fixed point stencil is satisfied for an arbitrary $Q = (\sigma_1, \sigma_2, \rho)$, if the following constraints hold*

(1) We must select equation (2.15) if $\rho \geq 0$ and equation (2.16) if $\rho < 0$ to approximate the cross derivative term.

(2) The following sufficient conditions are satisfied,

for $\rho \geq 0$

$$\frac{(S_2)_j \max(\Delta^+(S_1)_i, \Delta^-(S_1)_i)}{(S_1)_i} \frac{\Delta^+(S_1)_i + \Delta^-(S_1)_i}{\Delta^+(S_1)_i\Delta^+(S_2)_j + \Delta^-(S_1)_i\Delta^-(S_2)_j} \leq \frac{\sigma_1}{\sigma_2 \rho}, \tag{2.31a}$$

$$\frac{(S_2)_j}{(S_1)_i \max(\Delta^+(S_2)_j, \Delta^-(S_2)_j)} \frac{\Delta^+(S_1)_i\Delta^+(S_2)_j + \Delta^-(S_1)_i\Delta^-(S_2)_j}{\Delta^+(S_2)_j + \Delta^-(S_2)_j} \geq \frac{\sigma_1 \rho}{\sigma_2}, \tag{2.31b}$$

for $\rho < 0$

$$\frac{(S_2)_j \max\left(\Delta^+(S_1)_i, \Delta^-(S_1)_i\right)}{(S_1)_i} \frac{\Delta^+(S_1)_i + \Delta^-(S_1)_i}{\Delta^+(S_1)_i\Delta^-(S_2)_j + \Delta^-(S_1)_i\Delta^+(S_2)_j} \leq \frac{\sigma_1}{\sigma_2 |\rho|}, \tag{2.32a}$$

$$\frac{(S_2)_j}{(S_1)_i \max(\Delta^+(S_2)_j, \Delta^-(S_2)_j)} \frac{\Delta^+(S_1)_i\Delta^-(S_2)_j + \Delta^-(S_1)_i\Delta^+(S_2)_j}{\Delta^+(S_2)_j + \Delta^-(S_2)_j} \geq \frac{\sigma_1 |\rho|}{\sigma_2}. \tag{2.32b}$$

Proof. We select equation (2.15) if $\rho \geq 0$ and equation (2.16) if $\rho < 0$ to approximate the cross derivative term, this choice then ensures $\gamma_{i,j} \geq 0$. The condition (2) is sufficient to ensure that the following inequalities

$$\alpha_{i,j}^{S_1} - \gamma_{i,j} \geq 0, \ \beta_{i,j}^{S_1} - \gamma_{i,j} \geq 0, \ \alpha_{i,j}^{S_2} - \gamma_{i,j} \geq 0, \ \alpha_{i,j}^{S_2} - \gamma_{i,j} \geq 0,$$

hold. For more details see [58, Chapter 9.4]. □

Note that we use the seven point stencil method to approximate the cross-derivative term. An alternative discretization approach which is compatible with a monotone scheme is a weighted average of the approximations as described in [65]. A similar grid spacing condition is derived to ensure the positive coefficient condition in [65].

Theorem 2.1. *Assume that*

(1) We must select equation (2.15) if $\rho \geq 0$ and equation (2.16) if $\rho < 0$ to approximate the cross derivative term.

(2) The grid spacings satisfy the following conditions in terms of extreme values of the control $Q = (\sigma_1, \sigma_2, \rho)$.

$$\begin{cases} (2.31a) \text{ for } (\sigma_{1,\min}, \sigma_{2,\max}, \rho_{\max}) \text{ and } (2.31b) \text{ for } (\sigma_{1,\max}, \sigma_{2,\min}, \rho_{\max}), & \text{if } \rho_{\min} \geq 0, \\ (2.32a) \text{ for } (\sigma_{1,\min}, \sigma_{2,\max}, \rho_{\min}) \text{ and } (2.32b) \text{ for } (\sigma_{1,\max}, \sigma_{2,\min}, \rho_{\min}), & \text{if } \rho_{\max} \leq 0, \\ (2.31a) \text{ for } (\sigma_{1,\min}, \sigma_{2,\max}, \rho_{\max}),\ (2.31b) \text{ for } (\sigma_{1,\max}, \sigma_{2,\min}, \rho_{\max}), & \\ (2.32a) \text{ for } (\sigma_{1,\min}, \sigma_{2,\max}, \rho_{\min}), \text{ and } (2.32b) \text{ for } (\sigma_{1,\max}, \sigma_{2,\min}, \rho_{\min}), & \text{if } \rho_{\min} \leq 0 \leq \rho_{\max}. \end{cases} \tag{2.33}$$

With these conditions and Algorithm A.1, the positive coefficient condition (2.18) is satisfied for $\forall Q \in Z$. We denote the domain where the conditions (2.33) are satisfied by Ω_f.

Proof. For the case $\rho_{\min} \geq 0$, if the constraint (2.31) holds for all $Q \in Z$, we have

$$\begin{aligned} &\frac{(S_2)_j \max(\Delta^+(S_1)_i, \Delta^-(S_1)_i)}{(S_1)_i} \frac{\Delta^+(S_1)_i + \Delta^-(S_1)_i}{\Delta^+(S_1)_i\Delta^+(S_2)_j + \Delta^-(S_1)_i\Delta^-(S_2)_j} \\ &\leq \inf_{Q \in Z} \frac{\sigma_1}{\sigma_2 \rho} = \frac{\sigma_{1,\min}}{\sigma_{2,\max}\rho_{\max}}, \\ &\frac{(S_2)_j}{(S_1)_i \max(\Delta^+(S_2)_j, \Delta^-(S_2)_j)} \frac{\Delta^+(S_1)_i\Delta^+(S_2)_j + \Delta^-(S_1)_i\Delta^-(S_2)_j}{\Delta^+(S_2)_j + \Delta^-(S_2)_j} \\ &\geq \sup_{Q \in Z} \frac{\sigma_1 \rho}{\sigma_2} = \frac{\sigma_{1,\max}\rho_{\max}}{\sigma_{2,\min}}. \end{aligned} \tag{2.34}$$

The proof is similar for the other two cases. □

In Algorithm A.1, we select upstream differencing (forward or backward differencing) and central differencing for the first order derivative terms. When the conditions in Theorem 2.1 are satisfied, upstream differencing ensures that the positive coefficient condition holds. However, central differencing is used as much as possible to minimize discretization error. Consequently, given a control Q, if central differencing satisfies the positive coefficient condition, central differencing will be preferred.

Remark 2.3. *Grid spacing conditions in Theorem 2.1 depend on the space state* (S_1, S_2)*, thus the structure of a grid is not always such that these conditions are met everywhere. We shall not enforce these conditions, but indeed check whether they are satisfied at a given grid node.*

Our algorithm is summarized as follows. The domains are defined in Table 2.1. The fixed point stencil introduced in Section 2.3.1 is used in the domain Ω_f. For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w$, we need to use a wide stencil based on a local coordinate rotation to discretize the second derivative terms $(\mathbf{D}\nabla)\cdot\nabla\mathcal{U}$ in the HJB equation (2.5a). When using the wide stencil discretization, we use upstream finite differencing for the first order derivatives. We avoid using points below the lower boundaries for $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{w^*}$. We use the asymptotic solution (2.26) of the HJB equation at a point outside the computational region at the upper boundaries. From the discretization (2.24), we can see that the measure of Ω_{out} converges to zero as $h \to 0$ (2.13). Lastly, fully implicit time-stepping is used to ensure the unconditional monotonicity of our numerical scheme.

Ω	$[0, S_{1,\max}] \times [0, S_{2,\max}] \times [0, T]$
Ω_{τ_0}	$[0, S_{1,\max}] \times [0, S_{2,\max}] \times \{0\}$
Ω_{up}	$\{S_{1,\max}\} \times (0, S_{2,\max}] \times (0, T] \cup (0, S_{1,\max}] \times \{S_{2,\max}\} \times (0, T]$
Ω_{in}	$\Omega/\Omega_{\tau_0}/\Omega_{up}$
Ω_f	The region in Ω_{in} where conditions (2.33) in Theorem 2.1 hold.
Ω_b	$[h, \sqrt{h}] \times (0, S_{2,\max}] \times (0, T] \cup (0, S_{1,\max}] \times [h, \sqrt{h}] \times (0, T]$.
Ω_{w^*}	The region in Ω_b that does not satisfy the condition (2.33).
Ω_w	$\Omega_{in}/\Omega_f/\Omega_{w^*}$
Ω_{out}	$(S_{1,\max}, S_{1,\max} + \sqrt{h}] \times [0, S_{2,\max} + \sqrt{h}] \times (0, T] \cup [0, S_{1,\max}] \times (S_{2,\max}, S_{2,\max} + \sqrt{h}] \times (0, T]$

Table 2.1: The domain definitions.

2.3.7 Discretization form

We will give details of the discretization for the HJB equation (2.5a) in Ω_{in} in this section. For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_f$ where the fixed point stencil is used, the HJB equation (2.5a) has the following discretized form

$$\frac{\mathcal{U}_{i,j}^{n+1} - \mathcal{U}_{i,j}^{n}}{\Delta\tau} = \sup_{Q \in \partial Z} \left(L_f^Q \mathcal{U}_{i,j}^{n+1} \right), \tag{2.35}$$

where the discretized linear operator L_f^Q is defined in (2.17).

Remark 2.4. *(Restricting the control to the boundary) In the discrete equations* $L_f^Q \mathcal{U}_{i,j}^{n+1}$*, the numerical approximations of first order derivatives depend on the stencil, backward, forward*

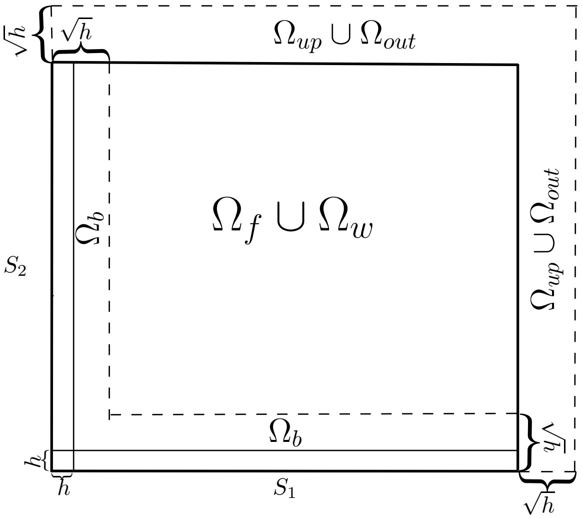


Figure 2.3: The domain descriptions.

or central differencing, which depend on the control. Thus, the discrete first order derivatives are also involved in the optimization of the discrete equations. In addition, the numerical approximation of the cross derivative term in (2.17) is dependent on the sign of the correlation ρ. In Proposition 2.1, the objective function contains just the diffusion terms, and we assume that $\Gamma_{kl},\ k,l = 1,2$ are constant and independent of the control. Therefore, the optimal value of the discrete equations is not necessarily attained at the boundary ∂Z. However, Proposition 2.1 holds for a smooth test function. Consequently, restricting the control to the boundary of the control set is a consistent approximation in the viscosity sense. Note that we also have an analytic expression for the optimal control for the discrete equations $L_f^Q \mathcal{U}_{i,j}^{n+1}$ when restricting $Q \in \partial Z$.

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w$ where the wide stencil is used, the discretized form of the linear differential operator $\mathcal{L}$ (2.3) is denoted by L_w^Q.

$$\begin{aligned} L_w^Q \mathcal{U}_{i,j}^{n+1} &= \frac{a_{i,j}}{h} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_1\right) + \frac{a_{i,j}}{h} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_1\right) \\ &\quad + \frac{b_{i,j}}{h} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_2\right) + \frac{b_{i,j}}{h} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_2\right) \\ &\quad + 1_{(r-q_1)\geq 0} \frac{(r-q_1)(S_1)_i}{\Delta^+(S_1)_i} \mathcal{U}_{i+1,j}^{n+1} - 1_{(r-q_1)<0} \frac{(r-q_1)(S_1)_i}{\Delta^-(S_1)_i} \mathcal{U}_{i-1,j}^{n+1} + 1_{(r-q_2)\geq 0} \frac{(r-q_2)(S_2)_j}{\Delta^+(S_2)_j} \mathcal{U}_{i,j+1}^{n+1} \\ &\quad - 1_{(r-q_2)<0} \frac{(r-q_2)(S_2)_j}{\Delta^-(S_2)_j} \mathcal{U}_{i,j-1}^{n+1} - \Bigg(1_{(r-q_1)\geq 0} \frac{(r-q_1)(S_1)_i}{\Delta^+(S_1)_i} - 1_{(r-q_1)<0} \frac{(r-q_1)(S_1)_i}{\Delta^-(S_1)_i} \\ &\quad + 1_{(r-q_2)\geq 0} \frac{(r-q_2)(S_2)_j}{\Delta^+(S_2)_j} - 1_{(r-q_2)<0} \frac{(r-q_2)(S_2)_j}{\Delta^-(S_2)_j} + \frac{2a_{i,j}}{h} + \frac{2b_{i,j}}{h} + r \Bigg) \mathcal{U}_{i,j}^{n+1}, \end{aligned} \tag{2.36}$$

where $a_{i,j}$ and $b_{i,j}$ are given in (2.22), and the presence of $\mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k\right)$, $k = 1,2$

is due to the discretization of the second derivative terms (2.24). As defined in (2.23), $\mathcal{U}^n(\mathbf{S}) \equiv \mathcal{U}(S_1, S_2, \tau^n)$, $\mathbf{S} = (S_1, S_2)$ and $\mathbf{S}_{i,j} = ((S_1)_i, (S_2)_j)$.

Remark 2.5. *The points $\mathbf{S}_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k$, $k = 1, 2$ used in (2.36) are control Q dependent. Therefore, the discretization in this case will depend on the control. We indicate this fact in the notation of the discrete linear operator L_w^Q.*

Since the numerical approximations of the diffusion terms depend on the control in the discrete equations $L_w^Q \mathcal{U}_{i,j}^{n+1}$, there is no simple analytic expression which can be used to maximize the discrete equations (2.36). We also do not have any known convexity properties of (2.36). For a bounded set of the controls, we must find the global maximum of (2.36) to ensure that our policy iteration algorithm converges. Hence, we discretize the control set Z (2.6), and maximize by linear search.

As explained in Remark 2.4, we will maximize the discrete equations $L_w^Q \mathcal{U}_{i,j}^{n+1}$ restricting the control to ∂Z. This significantly reduces the computational cost. We denote ∂Z_h as the discrete approximation of ∂Z

$$\partial Z_h = \{(\sigma_1)_1, \ldots, (\sigma_1)_{l_{\max}}\} \times \{(\sigma_2)_1, \ldots, (\sigma_2)_{k_{\max}}\} \times \{\rho_{\min}, \rho_{\max}\}, \tag{2.37}$$

where $(\sigma_1)_1 = \sigma_{1,\min}$, $(\sigma_1)_{l_{\max}} = \sigma_{1,\max}$, $(\sigma_2)_1 = \sigma_{2,\min}$, and $(\sigma_2)_{k_{\max}} = \sigma_{2,\max}$. Let

$$\max_i \left((\sigma_1)_i - (\sigma_1)_{i-1}\right) = C_4 h \text{ and } \max_i \left((\sigma_2)_i - (\sigma_2)_{i-1}\right) = C_5 h, \tag{2.38}$$

where h (2.13) is the mesh discretization parameter.

Finally, using fully implicit timestepping, the HJB equation (2.5a) has the following discretized form for this case

$$\frac{\mathcal{U}^{n+1} - \mathcal{U}^n}{\Delta\tau} = \sup_{Q \in \partial Z_h} \left(L_w^Q \mathcal{U}_{i,j}^{n+1}\right). \tag{2.39}$$

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{w^*}$, we need to adapt the discretized linear operator L_w^Q to avoid using points below the lower boundaries as described in Algorithm 2.1. The details of the discretized equation for this case are given in Appendix B.

2.3.8 The matrix form of the discrete equations

It is convenient to use a matrix form to represent the discretized equations for computational purposes. In this section we define a number of matrices and vectors to represent the discretized PDEs in (2.35), (2.39) and (B.2). Let $\mathcal{U}_{i,j}^n$ be the approximate solution of the equation (2.5a) at $((S_1)_i, (S_2)_j, \tau^n)$, $1 \le i \le N_1$, $1 \le j \le N_2$ and $0 \le \tau^n \le N_\tau$, and form the solution vector

$$\mathbf{U}^n = \left(\mathcal{U}_{1,1}^n, \mathcal{U}_{2,1}^n, \ldots, \mathcal{U}_{N_1,1}^n, \ldots, \mathcal{U}_{1,N_2}^n, \ldots, \mathcal{U}_{N_1,N_2}^n\right). \tag{2.40}$$

It will sometimes be convenient to use a single index when referring to an entry of the solution vector

$$\mathcal{U}_\ell^n = \mathcal{U}_{i,j}^n, \quad \ell = i + (j-1)N_1.$$

Let $N = N_1 \times N_2$, and we define the $N \times N$ matrix $\mathbf{L}^{n+1}(\mathcal{Q})$, where

$$\mathcal{Q} = \{Q_1, \ldots, Q_N\} \tag{2.41}$$

is an indexed set of N controls, and each Q_ℓ is in the set of admissible controls. $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{Q})$ is the entry on the ℓ-th row and k-th column, where $\ell = i+(j-1)N_1,\ i = 1, \ldots, N_1,\ j = 1, \ldots, N_2$.

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{up}$ where the Dirichlet boundary condition (2.26) is imposed, and we then have

$$\mathbf{L}_{\ell,k}^{n+1}(\mathcal{Q}) = 0, \quad k = 1, \ldots, N, \tag{2.42}$$

and define the vector $\mathbf{F}^{n+1}$ with entries

$$\mathbf{F}_\ell^{n+1} = \begin{cases} \mathcal{U}\left((S_1)_i, (S_2)_j, \tau^{n+1}\right), & \left((S_1)_i, (S_2)_j, \tau^{n+1}\right) \in \Omega_{up}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.43}$$

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_f$, the entries $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{Q})$ are constructed from the discrete linear operator L_f^Q (2.17). That is,

$$[\mathbf{L}^{n+1}(\mathcal{Q})\mathbf{U}^{n+1}]_\ell = L_f^Q \mathcal{U}_{i,j}^{n+1}. \tag{2.44}$$

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w$, we need to use the values at the following four off-grid points $\mathbf{S}_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k,\ k = 1, 2$ in the discrete linear operator L_w^Q (2.36). Let these four points be denoted as $P_{i,j}^m,\ m = 1, 2, 3, 4$, respectively. Note that these points may be outside the bounded domain Ω_{in}. When $P_{i,j}^m \in \Omega_{in}$, using linear interpolation, values at these four points are approximated as follows

$$\mathcal{J}_h \mathcal{U}^{n+1}\left(P_{i,j}^m\right) = \begin{cases} \displaystyle\sum_{\substack{d=0,1 \\ e=0,1}} \omega_{i,j}^{p_m+d, q_m+e} \mathcal{U}_{p_m+d, q_m+e}^{n+1}, & P_{i,j}^m \in \Omega_{in} \\ 0, & \text{Otherwise} \end{cases}. \tag{2.45}$$

For linear interpolation, we have that $\omega_{i,j}^{p_m+d, q_m+e} \geq 0$ and $\sum_{\substack{d=0,1 \\ e=0,1}} \omega_{i,j}^{p_m+d, q_m+e} = 1$. By inserting (2.45) in (2.36), the entries $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{Q})$ on the ℓ-th row are then specified. When a point $P_{i,j}^m$ is outside the domain Ω_{in} and inside the domain Ω_{out}, we then use its asymptotic solution at the point without extrapolating its value. We need to define the vector $\mathbf{B}^{n+1}(\mathcal{Q})$

to facilitate the construction of the matrix form in this situation when we use a point in the domain Ω_{out}.

$$\mathbf{B}_\ell^{n+1}(\mathcal{Q}) = \begin{cases} 1_{P^1_{i,j}\in\Omega_{out}} \frac{a_{i,j}}{h}\mathcal{U}(P^1_{i,j}) + 1_{P^2_{i,j}\in\Omega_{out}} \frac{a_{i,j}}{h}\mathcal{U}(P^2_{i,j}) & \\ \qquad + 1_{P^3_{i,j}\in\Omega_{out}} \frac{b_{i,j}}{h}\mathcal{U}(P^3_{i,j}) + 1_{P^4_{i,j}\in\Omega_{out}} \frac{b_{i,j}}{h}\mathcal{U}(P^4_{i,j}), & ((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w \cup \Omega_{w^*} \\ 0, & \text{otherwise} \end{cases}, \tag{2.46}$$

where $\mathcal{U}(P^m_{i,j})$ is the asymptotic solution (2.26) at the point. As a result, for the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w$, we have

$$[\mathbf{L}^{n+1}(\mathcal{Q})\mathbf{U}^{n+1}]_\ell + \mathbf{B}_\ell^{n+1}(\mathcal{Q})_\ell = L_w^Q \mathcal{U}_{i,j}^{n+1}. \tag{2.47}$$

Lastly, for $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{w^*}$, using the corresponding discrete linear operator $L_{w^*}^Q$ (B.1), the entries $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{Q})$ are constructed similarly to the previous case where $((S_1)_i, (S_2)_j, \tau^{n+1})$ $\in \Omega_w$.

Finally, the matrix form of the discretized equations is

$$\begin{aligned} &\left[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\hat{\mathcal{Q}})\right]\mathbf{U}^{n+1} = \mathbf{U}^n + \mathbf{F}^{n+1} - \mathbf{F}^n + \Delta\tau \mathbf{B}^{n+1}(\hat{\mathcal{Q}}), \\ &\qquad \hat{Q}_\ell \in \arg\sup_{Q\in\hat{Z}} \left[\mathbf{L}^{n+1}(\mathcal{Q})\mathbf{U}^{n+1} + \mathbf{B}^{n+1}(\mathcal{Q})\right]_\ell, \\ &\qquad\qquad \ell = i + (j-1)N_1, \quad i = 1,\dots,N_1 - 1, \quad j = 1,\dots,N_2 - 1, \end{aligned} \tag{2.48}$$

where we define $\hat{Z}$ as

$$\hat{Z} = \begin{cases} \partial Z, & \left((S_1)_i, (S_2)_j, \tau^{n+1}\right) \in \Omega_f, \\ \partial Z_h, & \left((S_1)_i, (S_2)_j, \tau^{n+1}\right) \in \Omega_w \cup \Omega_{w^*}, \end{cases} \tag{2.49}$$

and the notation $\arg\sup[\cdot]$ refers to the points which maximize the upper semi-continuous envelope of the argument. Chapter 3 proves that this numerical scheme converges to the viscosity solution, and Chapter 4 presents the policy iteration algorithm used to solve these nonlinear algebraic equations.

2.4 Numerical results

Our first test case is for a European call option on the maximum of two assets with a payoff

$$\max(\max(S_1, S_2) - K, 0), \tag{2.50}$$

All model parameters are given in Table 2.2. We consider the worst-case option value for a short position, i.e., the sup case in equation (2.5a). In this case, since the payoff is convex, and convexity is preserved [50], the worst case price can be analytically obtained for the value with the fixed parameters $\sigma_1 = \sigma_{1,\max}$, $\sigma_2 = \sigma_{2,\max}$, $\rho = \rho_{2,\min}$. The closed-form solution [67] with these volatility and correlation values is $\mathcal{U}(S_1 = 40, S_2 = 40, K = 40, t = 0) = 6.8477$. Thus, it is the solution to the HJB equation (2.5a).

The numerical solutions were computed on a sequence of uniformly refined grids, starting with 91×91 grid nodes. The initial discretization parameter h (2.13) is 0.4, and the initial timestep size is 0.01. At each grid refinement, the timestep is halved. We initiate the control set ∂Z_h with using twelve nodes along the boundary of volatility values and two nodes on two sides of the correlation range. At each refinement, only the number of volatility nodes is doubled. The relative convergence tolerance for nonlinear policy iteration is 10^{-6} (see Algorithm 4.1).

We use $(S_1)_{\max} = (S_2)_{\max} = 400$ (i.e. about ten times the asset values of interest). We carried out some tests using $(S_1)_{\max} = (S_2)_{\max} = 2000$. The solutions at $(S_1, S_2) = (40, 40)$ were the same to six digits.

Convergence results using a pure wide stencil method based on a local coordinate system and the hybrid scheme which uses the fixed point stencil as much as possible are given in Table 2.3. These correspond to the choice of sup in equations (2.5). Both the numerical results seem to be convergent to the benchmark. However, the hybrid scheme results are more accurate than those results obtained by the pure wide stencil method. We also carried out numerical experiments for the wide stencil based on factoring the diffusion tensor as shown in Table 2.4. The numerical results in Table 2.4 have larger errors than those in Table 2.3. Especially at the first two refinements, the pure wide stencil based on the factoring diffusion tensor performs poorly. Furthermore, the hybrid scheme significantly improves the accuracy of this pure wide stencil method.

Table 2.5 gives the average number of the policy iterations per time step in both the pure wide and the hybrid scheme method, which is about three. This result verifies our assumption that the number of the policy iterations is bounded as $h \to 0$, and hence the fully implicit method has the same complexity per step as an explicit method (for the pure wide stencil methods). Table 2.5 gives the ratio of the grid nodes where the fixed point stencil are used to the total number of nodes in the hybrid scheme. The ratio shows that the fixed stencil method cannot ensure monotonicity in general.

Note that the analytical result for the worst-case value is not immediately obvious, since even though Γ_{11} and Γ_{22} (2.7) are both non-negative, Γ_{12} is non-positive for a European call option on the maximal of two asset prices. Hence, maximizing or minimizing (2.7) is not necessarily trivial, although in this case it turns out that the same volatility ($\sigma_1 = 0.5$, $\sigma_2 = 0.5$) and correlation values ($\rho = 0.3$) should be chosen for the worst-case value in theory. Further, the numerical scheme did not always set the optimal controls to the same values as for the analytical values at all grid nodes for each time step. That is, the optimal controls for the discrete equations (2.48) are not the same as values obtained in (2.7). For example, the numerical approximations of the diffusion terms sometimes had different signs than would be expected from the theoretical values. Nevertheless, by optimizing the discrete equations, the numerical solution converges to the correct solution.

Parameter	Value
Type	Call
Time to expiry (T)	0.25
r	0.05
$\sigma_{1,\text{min}}$	0.3
$\sigma_{2,\text{max}}$	0.5
$\sigma_{2,\text{min}}$	0.3
$\sigma_{2,\text{max}}$	0.5
ρ_{min}	0.3
ρ_{max}	0.5

Table 2.2: Model parameters for the max of two asset call option.

		Hybrid Scheme (with rotation)			Pure Wide Stencil (rotation)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91 × 91	6.9182			7.4556		
50	181 × 181	6.8638	0.0544		7.1452	0.310	
100	361 × 361	6.8542	0.00962	5.62	6.9892	0.156	1.98
200	721 × 721	6.8506	0.00361	2.66	6.9208	0.0684	2.28

Table 2.3: Convergence results for an at-the-money European call option with the payoff (2.50) and parameters as given in Table 2.2. $S_1 = 40,\ S_2 = 40,\ K = 40$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on a local coordinate rotation, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference. Analytic solution in this case is 6.8477. Worst case short.

Our next test uses the same parameters as in Table 2.2. The payoff has been changed to a butterfly on the maximum of two assets. In particular, the payoff is

$$\begin{aligned} S_{\text{max}} &= \max(S_1, S_2), \\ \mathcal{W}(S_1, S_2) &= \max(S_{\text{max}} - K_1, 0) + \max(S_{\text{max}} - K_2, 0) - 2\max(S_{\text{max}} - (K_1 + K_2)/2, 0). \end{aligned} \tag{2.51}$$

This test is more challenging, since the payoff of the butterfly option is no longer convex, and thus the signs of the second order derivative terms change over the solution domain. Convergence results for the worst-case and best-case (short position) values are given in Tables 2.6 to 2.9. The numerical results in Table 2.6 and Table 2.8 are given by the wide stencil based on a local coordinate rotation. As shown in the tables, the convergence ratio

		Hybrid Scheme (with factoring)			Pure Wide Stencil (factoring)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91 × 91	6.9639			5.9476		
50	181 × 181	6.9302	0.0437		6.4910	0.543	
100	361 × 361	6.8966	0.0336	1.30	6.7168	0.226	2.40
200	721 × 721	6.8746	0.0221	1.52	6.7942	0.0774	2.92

Table 2.4: Convergence results for an at-the-money European call option with the payoff (2.50) and parameters as given in Table 2.2. $S_1 = 40,\ S_2 = 40,\ K = 40$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on factoring the diffusion tensor, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference. Analytic solution in this case is 6.8477. Worst case short.

	Average Iterations		
Time steps	Hybrid Scheme	Pure Wide	Fraction Fixed
25	3.3	3.1	0.38
50	3.3	2.9	0.42
100	3.0	2.5	0.44
200	2.8	2.4	0.45

Table 2.5: The test case of a European call option on the maximum of two assets. Average Iterations is the average number of the policy iterations per time step. Pure Wide stands for the wide stencil method based on a local coordinate rotation, while Hybrid Scheme stands for the hybrid scheme using the fixed point stencil as much as possible. Fraction Fixed gives the ratio of the grid nodes where the fixed point stencil are used to the total number of nodes in the hybrid scheme.

of the pure wide stencil method does not seem to be smooth. The best-case results seem to oscillate at the last two refinements. However, when we combine the wide stencil with use of a fixed point stencil as much as possible, the solution converges with an almost linear rate. Compared to the results in Table 2.7 and 2.9, which are given by the wide stencil method based on factoring the diffusion tensor, the performance of the wide stencil based on a local rotation seems to be superior. Both in the worst case and the best case scenarios, the errors of the pure wide stencil based on the factoring diffusion tensor are very large, especially at the first two refinements. Again, the hybrid scheme significantly improves the performance of the factoring method.

The average number of the policy iterations per time step are shown in Table 2.10 for the butterfly test case. The trends are the same as in Table 2.5, although both pure wide and hybrid stencil method tend to require more iterations on average. This is a direct result of this problem being truly nonlinear.

For comparison, Table 2.11 gives prices of the butterfly options on maximal of two assets using fixed volatility and correlation values. We see that the uncertain worst-case and best-case values form an upper and lower bound for the fixed parameter prices.

		Hybrid Scheme (with rotation)			Pure Wide Stencil (rotation)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91×91	2.7160			2.6371		
50	181×181	2.6946	0.0214		2.6397	0.00261	
100	361×361	2.6880	0.00655	3.27	2.6650	0.0252	0.10
200	721×721	2.6862	0.00184	3.60	2.6744	0.00940	2.67

Table 2.6: Convergence results for a worst-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). $S_1 = 40,\ S_2 = 40,\ K_1 = 34,\ K_2 = 46$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on a local coordinate rotation, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference.

2.5 Summary

The main results of this chapter are summarized as follows.

- We developed a fully implicit, unconditionally monotone finite difference numerical scheme for the two dimensional uncertain volatility HJB equation (2.5a).
- We derived a hybrid scheme which uses a fixed point stencil as much as possible and a wide stencil method as a complement to ensure monotonicity. Our numerical experiments showed that our hybrid scheme performs better than a pure wide stencil.

		Hybrid Scheme (with factoring)			Pure Wide Stencil (factoring)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91 × 91	2.8518			3.1129		
50	181 × 181	2.7733	0.0885		2.6121	0.501	
100	361 × 361	2.7282	0.0452	1.96	2.6083	0.00372	135
200	721 × 721	2.7085	0.0196	2.31	2.6196	−0.0113	−0.32

Table 2.7: Convergence results for a worst-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). $S_1 = 40,\ S_2 = 40,\ K_1 = 34,\ K_2 = 46$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on factoring the diffusion tensor, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference.

		Hybrid Scheme (with rotation)			Pure Wide Stencil (rotation)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91 × 91	0.9751			0.9787		
50	181 × 181	0.9420	0.0331		0.9213	0.0574	
100	361 × 361	0.9227	0.0193	1.72	0.9129	0.00842	1.69
200	721 × 721	0.9183	0.00435	4.44	0.9148	−0.00943	−0.89

Table 2.8: Convergence results for a best-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). $S_1 = 40,\ S_2 = 40,\ K_1 = 34,\ K_2 = 46$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on a local coordinate rotation, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference.

		Hybrid Scheme (with factoring)			Pure Wide Stencil (factoring)		
Time steps	Nodes	Value	Diff	Ratio	Value	Diff	Ratio
25	91 × 91	0.6448			2.3915		
50	181 × 181	0.7621	0.117		1.5937	0.796	
100	361 × 361	0.8621	0.0999	1.17	1.1287	0.465	1.71
200	721 × 721	0.8913	0.0293	3.41	1.0273	0.101	4.60

Table 2.9: Convergence results for a best-case (short) butterfly option with parameters as given in Table 2.2 and payoff specified by equation (2.51). $S_1 = 40,\ S_2 = 40,\ K_1 = 34,\ K_2 = 46$. Pure Wide Stencil shows the numerical solutions given by a wide stencil method based on factoring the diffusion tensor, and Hybrid Scheme shows results obtained using the fixed point stencil as much as possible. Diff is the value of the change in the solution as the grid refined. Ratio is the successive difference.

	Average Iterations		
Time steps	Hybrid Scheme	Pure Wide	Fraction Fixed
25	4.0	3.7	0.38
50	3.8	3.7	0.42
100	3.6	3.6	0.44
200	3.3	3.3	0.45

Table 2.10: The test case for a worst-case (short) butterfly option on maximal of two assets. Average Iterations is the average number of the policy iterations per time step. Pure Wide stands for the wide stencil based on a local coordinate rotation, while Hybrid Scheme stands for the hybrid scheme using the fixed point stencil as much as possible. Fraction Fixed gives the ratio of the grid nodes where the fixed point stencil are used to the total number of nodes in the hybrid scheme.

Test	Value
Uncertain worst-case	2.6862
$\sigma_1 = 0.3,\ \sigma_2 = 0.3,\ \rho = 0.3$	2.1910
$\sigma_1 = 0.3,\ \sigma_2 = 0.3,\ \rho = 0.5$	2.1891
$\sigma_1 = 0.4,\ \sigma_2 = 0.4,\ \rho = 0.4$	1.7404
$\sigma_1 = 0.5,\ \sigma_2 = 0.5,\ \rho = 0.3$	1.4480
$\sigma_2 = 0.5,\ \sigma_2 = 0.5,\ \rho = 0.5$	1.4364
Uncertain best-case	0.9183

Table 2.11: Option values for various parameter choices with a butterfly payoff. $S_1 = 40,\ S_2 = 40,\ K_1 = 34,\ K_2 = 46,\ T = 0.25$. The worst-case and best-case (short position) are obtained by the hybrid scheme using the fixed point stencil as much as possible and the wide stencil based on a local coordinate rotation.

- It appears that the rotation based scheme is superior to the method based on factoring the diffusion tensor.
- The number of policy iterations per timestep appears to be roughly constant as $h \to 0$ (h is the mesh discretization parameter).

Chapter 3

Convergence to the Viscosity Solution

In general, we cannot expect solutions to the HJB equation (2.5a) to be smooth. Hence, we seek the viscosity solution of the equation (2.5a). From [6], we find that a sufficient condition which guarantees convergence to the viscosity solution is that the numerical scheme is ℓ_∞ stable, consistent in the viscosity sense, and monotone. In this chapter, we will verify each of the properties in turn for our numerical scheme developed in Chapter 2.

3.1 Viscosity solution for the localized problem

To make the statement of the problem more precise in the context of viscosity solutions, we now write the localized problem in a compact form, which includes the terminal and boundary equations in a single equation. Let us define

$$\mathbf{x} = (S_1, S_2, \tau), \quad D\mathcal{U}(\mathbf{x}) = (\frac{\partial \mathcal{U}}{\partial S_1}, \frac{\partial \mathcal{U}}{\partial S_2}), \quad D^2\mathcal{U}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 \mathcal{U}}{\partial S_1^2} & \frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_2} \\ \frac{\partial^2 \mathcal{U}}{\partial S_1 \partial S_2} & \frac{\partial^2 \mathcal{U}}{\partial S_2^2} \end{pmatrix}.$$

The HJB equation for the value function (2.5a) on the localized domain $\Omega \cup \Omega_{out}$ is given by

$$F\mathcal{U} \equiv F\left(\mathbf{x}, \mathcal{U}(\mathbf{x}), D\mathcal{U}(\mathbf{x}), D^2\mathcal{U}(\mathbf{x})\right) = 0, \tag{3.1}$$

where the operator $F\mathcal{U}$ is defined by

$$F\mathcal{U} = \begin{cases} F_{in}\mathcal{U} \equiv F_{in}\left(\mathbf{x}, \mathcal{U}(\mathbf{x}), D\mathcal{U}(\mathbf{x}), D^2\mathcal{U}(\mathbf{x})\right), & \mathbf{x} \in \Omega_{in} = \Omega_f \cup \Omega_w \cup \Omega_{w^*}, \\ F_{\tau_0}\mathcal{U} \equiv F_{\tau_0}\left(\mathbf{x}, \mathcal{U}(\mathbf{x})\right), & \mathbf{x} \in \Omega_{\tau_0}, \\ F_{\max}\mathcal{U} \equiv F_{\max}\left(\mathbf{x}, \mathcal{U}(\mathbf{x})\right), & \mathbf{x} \in \Omega_{up} \cup \Omega_{out}. \end{cases} \tag{3.2}$$

Here,

$$\begin{aligned} F_{in}\mathcal{U} &= \mathcal{U}_\tau - \max_{Q \in Z}\left(\mathcal{L}\mathcal{U}\right), \ (2.5a) \\ F_0\mathcal{U} &= \mathcal{U} - \mathcal{W}(S_1, S_2), \\ F_{\max}\mathcal{U} &= \mathcal{U} - \bar{\mathcal{U}}(S_1, S_2, \tau), \end{aligned} \tag{3.3}$$

where $\bar{\mathcal{U}}(S_1, S_2, \tau)$ is the asymptotic solution defined in (2.26).

Before defining the viscosity solution of equation (3.1), we first recall the definitions of upper and lower semi-continuous envelopes. Given a function $f : \tilde{\Omega} \to \mathbb{R}$, $\tilde{\Omega} \subseteq \mathbb{R}^n$, the upper semi-continuous envelope of f, denoted by f^*, is defined as

$$f^*(\tilde{x}) = \lim_{\tilde{r} \to 0^+} \left[\sup \left\{ f(y) \;\middle|\; y \in B(\tilde{x}, \tilde{r}) \cap \tilde{\Omega} \right\} \right], \tag{3.4}$$

where $B(\tilde{x}, \tilde{r}) = \{y \in \mathbb{R}^n \mid |\tilde{x} - y| < \tilde{r}\}$. We also have the obvious definition for a lower semi-continuous envelope $f_*(\tilde{x})$.

We also define

$$\limsup_{y \to \tilde{x}} f(y) = \lim_{\tilde{r} \to 0^+} \left[\sup \left\{ f(y) \;\middle|\; y \in B(\tilde{x}, \tilde{r}) \cap \tilde{\Omega} - \tilde{x} \right\} \right], \tag{3.5}$$

with the corresponding definition of lim inf.

Definition 3.1. *(Viscosity solution of equation 3.1) A locally bounded function $\mathcal{U} : \Omega \cup \Omega_{out} \to \mathbb{R}$ is a viscosity sub-solution (resp. super-solution) of equation (3.1) if, for all test functions $\phi(\mathbf{x}) \in C^\infty(\Omega \cup \Omega_{out})$, and all $\mathbf{x}$, such that $\mathcal{U} - \phi$ has a strict global maximum (resp. minimum) with $\phi(\mathbf{x}) = \mathcal{U}^*(\mathbf{x})$ (resp. $\mathcal{U}_*(\mathbf{x})$), we have*

$$\begin{gathered} F_* \left(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x})\right) \leq 0, \\ \left(\textit{resp. } F^* \left(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x})\right) \geq 0 \right), \end{gathered} \tag{3.6}$$

where $F_*(\cdot)$ is the lower semi-continuous envelope of F (resp. the upper semi-continuous envelope F^*). $\mathcal{U}$ is a viscosity solution if it is both a viscosity sub-solution and a viscosity super-solution.

Proposition 3.1. *(Strong comparison) Suppose the payoff function $\mathcal{W}(S_1, S_2)$ at expiry time T is continuous with quadratic growth, then the value function satisfies a strong comparison result, hence there exists an unique continuous viscosity solution of the problem (2.5a) [59, 43].*

Proof. See [59]. □

Corollary 3.1. *Note that we restrict ourselves to a finite domain $\Omega \cup \Omega_{out}$ for the HJB equation $F\mathcal{U} = 0$ defined in (3.1), hence the value function (3.1) satisfies a strong comparison result.*

3.2 Consistency

For the purpose of proving convergence to the viscosity solution, it is more convenient to rewrite equations (2.35), (2.39) and (B.2) in an equivalent form. Let $\mathcal{G}(\cdot)$ be the discrete approximation to F_{in} for $\mathbf{x} \in \Omega_{in}$, and $\mathbf{x}_{i,j}^{n+1} = ((S_1)_i, (S_2)_j, \tau^{n+1})$. For $\mathbf{x}_{i,j}^{n+1} \in \Omega_f$, from (2.35), we have

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \{\mathcal{U}_{a,b}^{n+1}\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \{\mathcal{U}_{k,l}^{n}\}\right) = \frac{\mathcal{U}_{i,j}^{n+1} - \mathcal{U}_{i,j}^{n}}{\Delta\tau} - \sup_{Q\in\partial Z}\left(L_f^Q \mathcal{U}_{i,j}^{n+1}\right) = 0. \tag{3.7}$$

For $\mathbf{x}_{i,j}^{n+1} \in \Omega_w$, from (2.39), we have

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \{\mathcal{U}_{a,b}^{n+1}\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \{\mathcal{U}_{k,l}^{n}\}\right) = \frac{\mathcal{U}_{i,j}^{n+1} - \mathcal{U}_{i,j}^{n}}{\Delta\tau} - \sup_{Q\in\partial Z_h}\left(L_w^Q \mathcal{U}_{i,j}^{n+1}\right) = 0. \tag{3.8}$$

For $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*}$, from (B.2), we have

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \{\mathcal{U}_{a,b}^{n+1}\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \{\mathcal{U}_{k,l}^{n}\}\right) = \frac{\mathcal{U}_{i,j}^{n+1} - \mathcal{U}_{i,j}^{n}}{\Delta\tau} - \sup_{Q\in\partial Z_h}\left(L_{w^*}^Q \mathcal{U}_{i,j}^{n+1}\right) = 0. \tag{3.9}$$

Finally, we have

$$\mathcal{G}(\cdot) = 0 = \begin{cases} \mathcal{U}((S_1)_i, (S_2)_j, 0) - \mathcal{W}((S_1)_i, (S_2)_j), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{\tau_0}, \\ \mathcal{U}((S_1)_i, (S_2)_j, \tau^{n+1}) - \bar{\mathcal{U}}((S_1)_i, (S_2)_j, \tau^{n+1}), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{up} \cup \Omega_{out}. \end{cases} \tag{3.10}$$

The domains $\Omega_f, \ldots, \Omega_{out}$ are defined in Table 2.1.

Definition 3.2. *(Consistency) For any C^∞ function $\phi(S_1, S_2, \tau)$ in $\Omega \cup \Omega_{out}$, with $\phi_{i,j}^{n+1} = \phi(\mathbf{x}_{i,j}^{n+1}) = \phi((S_1)_i, (S_2)_j, \tau^{n+1})$, the numerical scheme $\mathcal{G}(\cdot)$ is consistent in the viscosity sense, if, $\forall \hat{\mathbf{x}} = (\hat{S}_1, \hat{S}_2, \hat{\tau})$ with $\mathbf{x}_{i,j}^{n+1} = ((S_1)_i, (S_2)_j, \tau^{n+1})$, the following holds*

$$\limsup_{\substack{h\to 0 \\ \psi\to 0 \\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ or\ b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \leq F^*\left(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}})\right), \tag{3.11}$$

and

$$\liminf_{\substack{h\to 0 \\ \psi\to 0 \\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ or\ b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \geq F_*\left(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}})\right). \tag{3.12}$$

Lemma 3.1. *(Local consistency). Suppose the mesh discretization parameter h is defined in (2.13) and the control discretization satisfies equation (2.38), then for any C^∞ function $\phi(S_1, S_2, \tau)$ in $\Omega \cup \Omega_{out}$, with $\phi_{i,j}^{n+1} = \phi\left((S_1)_i, (S_2)_j, \tau^{n+1}\right) = \phi(\mathbf{x}_{i,j}^{n+1})$, and for h, ψ sufficiently small, ψ a constant, we have that*

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ or\ b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \\
&= \begin{cases}
F_{in}\phi_{i,j}^{n+1} + O(h) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_f, \\
F_{in}\phi_{i,j}^{n+1} + O(h) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_w, \\
F_{in}\phi_{i,j}^{n+1} + O(\sqrt{h}) + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*}, \\
F_{\tau_0}\phi_{i,j}^{n+1} + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{\tau_0}, \\
F_{\max}\phi_{i,j}^{n+1} + O(\psi), & \mathbf{x}_{i,j}^{n+1} \in \Omega_{up} \cup \Omega_{out}.
\end{cases}
\end{aligned} \tag{3.13}$$

Proof. To be precise, define the following

$$\begin{aligned}
\mathcal{L}\phi_{i,j}^{n+1} &\equiv \mathcal{L}\phi((S_1)_i, (S_2)_j, \tau^{n+1}), \\
(\phi_\tau)_{i,j}^{n+1} &\equiv \phi_\tau((S_1)_i, (S_2)_j, \tau^{n+1}).
\end{aligned} \tag{3.14}$$

For the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_f$, $L_f^Q \phi_{i,j}^{n+1}$ (2.17) is a locally consistent discretization of the linear operator $\mathcal{L}$ (2.3), that is,

$$L_f^Q \phi_{i,j}^{n+1} = \mathcal{L}\phi_{i,j}^{n+1} + O(h), \tag{3.15}$$

which is easily proved by Taylor series, and note that

$$\begin{aligned}
L_f^Q \left(\phi_{i,j}^{n+1} + \psi\right) &= L_f^Q \phi_{i,j}^{n+1} - r\psi, \\
\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n}}{\Delta\tau} &= (\phi_\tau)_{i,j}^{n+1} + O(h).
\end{aligned} \tag{3.16}$$

Since ϕ is a smooth test function, and $\frac{\partial^2 \phi}{\partial S_k \partial S_l}$, $k, l = 1, 2$ are independent of the control, then, by Proposition 2.1, we have

$$\sup_{Q\in\partial Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) = \sup_{Q\in Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right), \tag{3.17}$$

and from equation (3.7) and (3.17), we then have the result

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \\
&= \frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n}}{\Delta\tau} - \sup_{Q\in\partial Z} \left(L_f^Q \phi_{i,j}^{n+1}\right) + O(\psi) \\
&= (\phi_\tau)_{i,j}^{n+1} - \sup_{Q\in\partial Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(\psi) + O(h) \\
&= (\phi_\tau)_{i,j}^{n+1} - \sup_{Q\in Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(\psi) + O(h) \\
&= F_{in}\phi_{i,j}^{n+1} + O(\psi) + O(h), \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_f
\end{aligned} \tag{3.18}$$

For the case where $\mathbf{x}_{i,j}^{n+1} \in \Omega_w$, $L_w^Q \phi_{i,j}^{n+1}$ (2.36) is also locally consistent,

$$L_w^Q \phi_{i,j}^{n+1} = \mathcal{L}\phi_{i,j}^{n+1} + O(h), \tag{3.19}$$

and note that

$$\begin{aligned} L_w^Q \left(\phi_{i,j}^{n+1} + \psi\right) &= L_w^Q \phi_{i,j}^{n+1} - r\psi, \\ \frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n}}{\Delta\tau} &= (\phi_\tau)_{i,j}^{n+1} + O(h). \end{aligned} \tag{3.20}$$

From equation (3.8), we then have

$$\begin{aligned} &\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \\ &= \frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n}}{\Delta\tau} - \sup_{Q\in\partial Z_h} \left(L_w^Q \phi_{i,j}^{n+1}\right) + O(\psi) \\ &= (\phi_\tau)_{i,j}^{n+1} - \sup_{Q\in\partial Z_h} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(\psi) + O(h). \end{aligned} \tag{3.21}$$

We discretize the set ∂Z and maximize the discrete equations by linear search. If the discretization step for the control is also $O(h)$, then this is a consistent approximation [72], since the equation coefficients are Lipschitz continuous functions of the controls. That is, using equation (3.17),

$$\sup_{Q\in\partial Z_h} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) = \sup_{Q\in\partial Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(h) = \sup_{Q\in Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(h). \tag{3.22}$$

Using equation (3.22) in equation (3.21), we then have the final result

$$\begin{aligned} &\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \\ &= (\phi_\tau)_{i,j}^{n+1} - \sup_{Q\in Z} \left(\mathcal{L}\phi_{i,j}^{n+1}\right) + O(\psi) + O(h), \\ &= F_{in}\phi_{i,j}^{n+1} + O(\psi) + O(h), \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_w. \end{aligned} \tag{3.23}$$

For the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*}$, the proof is similar to the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_w$, but the consistency of the discrete linear operator $L_{w^*}^Q$ is perhaps not obvious. A possible inconsistency may arise when we shrink the stencil length from $O(\sqrt{h})$ to $O(h)$ to avoid using points below the lower boundaries. However, consistency still holds for $L_{w^*}^Q$ (see the proof in Appendix C)

$$L_{w^*}^Q \phi_{i,j}^{n+1} = \mathcal{L}\phi_{i,j}^{n+1} + O(\sqrt{h}).$$

Following the same steps as the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_w$, we finally have

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) = F_{in}\phi_{i,j}^{n+1} + O(\psi) + O(\sqrt{h}), \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*}. \tag{3.24}$$

The remaining results in (3.13) can be proven using similar arguments. □

Lemma 3.2. *(Consistency) Provided that all conditions in Lemma 3.1 are satisfied, then scheme (3.7-3.10) is consistent according to Definition (3.2).*

Proof. This follows in straightforward fashion from Lemma 3.1, using the same steps as in, for example, [47]. We sketch the proof of the inequality (3.11) here. From the definition of lim sup, there exists sequences i_k, j_k, ψ_k and h_k such that

$$\text{as } m \to \infty,\ \ \mathbf{x}_{i_m,j_m}^{n_m+1} \to \hat{\mathbf{x}}, \psi_m \to 0,\ \ h_m \to 0,$$

and

$$\begin{aligned}
&\limsup_{m\to\infty} \mathcal{G}\left(h_m, \mathbf{x}_{i_m,j_m}^{n_m+1},, \left\{\phi_{i_m,j_m}^{n_m+1} + \psi_m\right\}_{\substack{a_m\neq i_m \\ \text{or } b_m\neq j_m}}, \left\{\phi_{k_m,l_m}^{n_m} + \psi_m\right\}\right) \\
&= \limsup_{\substack{h\to 0 \\ \psi\to 0 \\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right).
\end{aligned} \tag{3.25}$$

From Lemma 3.1, we have for m sufficiently large, there exist positive constants K_1, K_2 independent of m such that

$$\begin{aligned}
&\left|\limsup_{m\to\infty} \mathcal{G}\left(h_m, \mathbf{x}_{i_m,j_m}^{n_m+1},, \left\{\phi_{i_m,j_m}^{n_m+1} + \psi_m\right\}_{\substack{a_m\neq i_m \\ \text{or } b_m\neq j_m}}, \left\{\phi_{k_m,l_m}^{n_m} + \psi_m\right\}\right) - F\phi_{i_m,j_m}^{n_m+1}\right| \\
&\leq K_1 h_m + K_2 \psi_m.
\end{aligned} \tag{3.26}$$

From equations (3.25) and (3.26), we obatin

$$\begin{aligned}
&\limsup_{\substack{h\to 0 \\ \psi\to 0 \\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1} + \psi, \left\{\phi_{a,b}^{n+1} + \psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,l}^{n} + \psi\right\}\right) \\
&\leq \limsup_{m\to\infty} F\phi_{i_m,j_m}^{n_m+1} + \limsup_{m\to\infty} (K_1 h_m + K_2 \psi_m) \\
&\leq F^*\left(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}})\right).
\end{aligned} \tag{3.27}$$

Similary, we can prove the inequality (3.12). □

3.3 Stability

Definition 3.3. *(M-matrix) If a matrix $\mathbf{A}$ has elements $a_{ii} > 0$ and $a_{ij} < 0$ for $i \neq j$ and every row sum is non-negative with at least one row sum positive in each connected part of $\mathbf{A}$, then $\mathbf{A}$ is an M-matrix [70].*

Remark 3.1. *We remind the reader that a sufficient condition for a matrix $\mathbf{A}$ to be an M-matrix is that $\mathbf{A}$ has positive diagonals, non-positive offdiagonals, and is diagonally dominant [70].*

Lemma 3.3. *Providing the following conditions hold*

- *We only use the discrete linear operator L_f^Q (2.17) in the domain Ω_f,*
- *A linear interpolation operator $\mathcal{J}_h$ is used in (2.36) and (B.1).*

Then $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})]$ (2.48) is an M-matrix.

Proof. From the formation of matrix $\mathbf{L}$ in (2.42), (2.44) and (2.47), it is easily seen that $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})]$ has positive diagonals, non-positive offdiagonals, and the ℓ-th row sums for the matrix are

$$\sum_k \left[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})\right]_{\ell,k} = \begin{cases} 1 + r\Delta\tau & i = 1, \dots, N_1 - 1, \ , j = 1, \dots, N_2 - 1 \\ 1 & i = N_1 \text{ or } j = N_2 \end{cases}, \tag{3.28}$$

where $\ell = i + (j-1)N_1$. Thus, the matrix $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})]$ is diagonally dominant. □

Lemma 3.4. *(Stability) If the conditions for Lemma 3.3 are satisfied, the discretization (2.48), equivalently (3.7-3.10), is unconditionally l_∞ stable, as mesh discretization parameter (2.13) $h \to 0$, satisfying*

$$\|\mathbf{U}^n\|_\infty \le \max\left(\|\mathbf{U}^0\|_\infty,\ C_6\right), \tag{3.29}$$

where $C_6 = \max_n \|\mathbf{F}^n\|_\infty$, where $\mathbf{F}^n$ are determined by the asymptotic boundary condition (2.26) on the bounded region.

Proof. By Lemma 3.3, from the properties of M-matrices and equation (3.28), we have that

$$\|[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})]^{-1}\|_\infty \le \max_\ell \frac{1}{rowsum([\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})]_\ell)} \le 1, \tag{3.30}$$

and using a straightforward maximum analysis as in [31], the result follows. □

3.4 Monotonicity

Definition 3.4. *(Monotonicity) The discrete scheme is monotone if for all $\mathcal{Y}_{i,j}^n \ge \mathcal{X}_{i,j}^n$, $\forall i, j, n$*

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \{\mathcal{Y}_{a,b}^{n+1}\}_{\substack{a\neq i \\ or\ b\neq j}}, \{\mathcal{Y}_{k,l}^n\}\right) \le \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \{\mathcal{X}_{a,b}^{n+1}\}_{\substack{a\neq i \\ or\ b\neq j}}, \{\mathcal{X}_{k,l}^n\}\right). \tag{3.31}$$

Lemma 3.5. *(Monotonicity) If the scheme (3.7-3.10) satisfies the conditions required for Lemma 3.4, then the discretization is monotone, according to Definition 3.4.*

Proof. We write out

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \left\{\mathcal{Y}_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\mathcal{Y}_{k,l}^{n}\right\}\right)$$

and

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \left\{\mathcal{X}_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\mathcal{X}_{k,l}^{n}\right\}\right)$$

in one of the component forms (3.7-3.10). Since our discretization is a positive coefficient scheme $\forall Q \in \hat{Z}$, using the same steps as in [39], it is then easy to show that

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \left\{\mathcal{Y}_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\mathcal{Y}_{k,l}^{n}\right\}\right) - \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \mathcal{U}_{i,j}^{n+1}, \left\{\mathcal{X}_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\mathcal{X}_{k,l}^{n}\right\}\right) \leq 0. \tag{3.32}$$

□

3.5 Convergence

Theorem 3.1. *(Convergence) Assume that discretization (3.7-3.10) satisfies all the conditions required by Lemmas 3.2, 3.4 and 3.5, and that Proposition 3.1 holds, then numerical scheme (3.7-3.10) converges to the unique continuous viscosity solution of the problem* (3.1).

Proof. Since the scheme is monotone, consistent and ℓ_∞-stable, this follows from the results in [7]. □

3.6 Summary

The main result in this chapter is the following:

- We seek the viscosity solution of the equation (2.5a). Given a monotone scheme as proved in this chapter, it is straightforward to show that our scheme is ℓ_∞ stable [31]. We also prove that our numerical scheme is consistent in the viscosity sense. Consequently, we can prove that our numerical scheme guarantees convergence to the viscosity solution.

Chapter 4

Solution of the Nonlinear Discrete Algebraic Equations

Although we have established that discretization (2.48) is consistent, ℓ_∞ stable and monotone, fully implicit timestepping requires solution of highly nonlinear algebraic equations at each timestep. For the applications addressed in [39] an efficient method for solving the associated algebraic systems made use of a policy iteration scheme. However, our discretization method is control dependent, and consequently the local objective function may be a discontinuous function of the control [72, 48]. Hence some care must be taken when applying policy iteration. Recall that at every timestep τ^n, the nonlinear algebraic linear equations (2.48) can be represented as in the form

$$\sup_{\mathcal{Q} \in \hat{Z}} \left\{ -\mathbf{A}(\mathcal{Q})\mathbf{U}^{n+1} + \mathbf{C}(\mathcal{Q}) \right\} = 0, \tag{4.1}$$

where

$$\mathbf{A}(\mathcal{Q}) \equiv [\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{Q})], \tag{4.2}$$

and

$$\mathbf{C}(\mathcal{Q}) \equiv \mathbf{U}^n + \mathbf{F}^{n+1} - \mathbf{F}^n + \Delta\tau \mathbf{B}^{n+1}(\mathcal{Q}). \tag{4.3}$$

$\mathcal{Q} \in \hat{Z}$ (see the definition of $\hat{Z}$ in (2.49)) denotes that each $Q_\ell \in \hat{Z},\ \ell = 1, \ldots, N$. Equation (4.1) is to be understood in the row-wise sense, i.e. $\sup_{\mathcal{Q} \in \hat{Z}}[\ \cdot\ ]_\ell = 0\ ;\ \ell = 1, \cdots, N$.

Before proceeding with a discussion of Policy Iteration, for solution of equation (4.1), we list here a set of properties of $\mathbf{A}(\mathcal{Q})$, $\mathbf{C}(\mathcal{Q})$, $\hat{Z}$, which will prove useful in later sections.

Properties 4.1. *(Properties of $\mathbf{A}(\mathcal{Q})$, $\mathbf{C}(\mathcal{Q})$, $\hat{Z}$)*

(i) The set of controls $\hat{Z}$ (2.49) is compact.

(ii) The matrices and vectors have the property that $\mathbf{A}_{\ell,k}(\mathcal{Q})$ and $\mathbf{C}_\ell(\mathcal{Q})$ depend only on Q_ℓ. That is, $\mathbf{A}_{\ell,k}(\mathcal{Q}) = \mathbf{A}_{\ell,k}(Q_\ell)$ and $\mathbf{C}_\ell(\mathcal{Q}) = \mathbf{C}_\ell(Q_\ell)$.

(iii) $\mathbf{A}(\mathcal{Q})$ *is a diagonally dominant M-matrix* $\forall \mathcal{Q}$, *and* $\sum_k \mathbf{A}_{\ell,k}(\mathcal{Q}) \geq C_r > 0$, *where* C_r *is independent of* $\mathcal{Q}$ *and row* ℓ.

(iv) $\|\mathbf{A}(\mathcal{Q})\|_\infty$, $\|\mathbf{C}(\mathcal{Q})\|_\infty$, *and* $\|\mathbf{A}(\mathcal{Q})^{-1}\|_\infty$ *are bounded uniformly w.r.t.* $\mathcal{Q}$.

Lemma 4.1 (Verification of Properties 4.1)**.** *The discretization (2.48) satisfies Properties 4.1.*

Proof. Property (i) holds from the definition of $Z, \hat{Z}$, see equation (2.6) and equation (2.49). From the definitions of $\mathbf{A}$ and $\mathbf{C}$, in equations (4.2-4.3), (ii) follows from the fact that the control at discrete node ℓ depends only on the discretized equation at node ℓ. (iii) holds from Lemma 3.3, with $C_r = 1$ (equation (3.28)). From (i) and the definitions of $\mathbf{A}$ and $\mathbf{C}$, we have that $\|\mathbf{A}(\mathcal{Q})\|$ and $\|\mathbf{C}(\mathcal{Q})\|$ are bounded independent of $\mathcal{Q}$. From equation (3.30), it follows that $\|\mathbf{A}(\mathcal{Q})^{-1}\|$ is bounded independent of $\mathcal{Q}$ as well, hence (iv) is satisfied. □

Fix a vector $\mathbf{W}$. From Properties 4.1, there exists a sequence $\mathcal{Q}^k$, such that

$$\lim_{k\to\infty} \Big(-\mathbf{A}(\mathcal{Q}^k)\mathbf{W} + \mathbf{C}(\mathcal{Q}^k) \Big) = \sup_{\mathcal{Q}\in\hat{Z}} \{ -\mathbf{A}(\mathcal{Q})\mathbf{W} + \mathbf{C}(\mathcal{Q}) \} \ . \tag{4.4}$$

Since $\mathbf{A}(\mathcal{Q})$, $\mathbf{C}(\mathcal{Q})$ are bounded, then there is a convergent subsequence $\{\mathcal{Q}^{k_j}\}$ such that $\mathbf{A}(\mathcal{Q}^{k_j}) \to \widehat{\mathbf{A}}(\mathbf{W})$ and $\mathbf{C}(\mathcal{Q}^{k_j}) \to \widehat{\mathbf{C}}(\mathbf{W})$, for some $\widehat{\mathbf{A}}(\mathbf{W})$, $\widehat{\mathbf{C}}(\mathbf{W})$, satisfying

$$-\widehat{\mathbf{A}}(\mathbf{W})\mathbf{W} + \widehat{\mathbf{C}}(\mathbf{W}) = \sup_{\mathcal{Q}\in\hat{Z}} \{ -\mathbf{A}(\mathcal{Q})\mathbf{W} + \mathbf{C}(\mathcal{Q}) \} \ . \tag{4.5}$$

We also have the following result

Proposition 4.1. *If Properties 4.1 hold, with* $\widehat{\mathbf{A}}(\mathbf{W})$ *and* $\widehat{\mathbf{C}}(\mathbf{W})$ *defined in equation (4.5), then* $\widehat{\mathbf{A}}(\mathbf{W})$ *is an M-matrix, and* $\|\widehat{\mathbf{C}}(\mathbf{W})\|_\infty$ *and* $\|\widehat{\mathbf{A}}(\mathbf{W})^{-1}\|_\infty$ *are bounded uniformly w.r.t.* $\mathbf{W}$.

Proof. From Properties 4.1, every matrix in the sequence $\mathbf{A}(\mathcal{Q}^{k_j})$ has non-positive off-diagonals, and has $\sum_k \mathbf{A}_{\ell,k}(\mathcal{Q}^{k_j}) \geq C_r > 0$, independent of $\mathcal{Q}^{k_j}$, hence the limit of the sequence $\widehat{\mathbf{A}}(\mathbf{W})$ has these properties as well, and thus $\widehat{\mathbf{A}}(\mathbf{W})$ is an M-matrix with $\sum_k \widehat{\mathbf{A}}_{\ell,k}(\mathbf{W}) \geq C_r > 0$. Since $\|\widehat{\mathbf{A}}(\mathbf{W})^{-1}\|_\infty \leq 1/C_r$, then $\|\widehat{\mathbf{A}}(\mathbf{W})^{-1}\|_\infty$ is bounded independent of $\mathbf{W}$. Similarly, since $\widehat{\mathbf{C}}(\mathbf{W})$ is the limit of a sequence of $\mathbf{C}(\mathcal{Q}^{k_j})$, which are bounded independent of $\mathcal{Q}^{k_j}$, then $\widehat{\mathbf{C}}(\mathbf{W})$ is bounded independent of $\mathbf{W}$. □

Policy iteration is a well known iterative method for solution of problems of type (4.1) [46]. The policy iteration approach for solution of equation (4.1) is given in Algorithm 4.1.

The term *scale* in Algorithm 4.1 is used to ensure that unrealistic levels of accuracy are not required when the value is very small (typically *scale* for an option priced in dollars is

Algorithm 4.1 Policy Iteration

```
1: Let $\mathbf{W}^0$ = Initial solution vector $\mathbf{U}^n$; given $scale > 0,\ tolerance > 0$
2: for $k = 0, 1, 2, \ldots$ until converge do
3:    $-\widehat{\mathbf{A}}(\mathbf{W}^k)\mathbf{W}^k + \widehat{\mathbf{C}}(\mathbf{W}^k) = \sup_{\mathcal{Q}\in\hat{Z}} \left\{-\mathbf{A}(\mathcal{Q})\mathbf{W}^k + \mathbf{C}(\mathcal{Q})\right\}$
4:    Solve the linear system $\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\mathbf{W}^{k+1} = \widehat{\mathbf{C}}\left(\mathbf{W}^k\right)$
5:    if $\max_\ell \frac{|\mathbf{W}^{k+1} - \mathbf{W}^k|}{\max\left[scale, |(\mathbf{W}^{k+1}|\right]} < tolerance$ then
6:       break from the iteration
7:    end if
8: end for
9: $\mathbf{U}^{n+1} = \mathbf{W}^{k+1}$
```

unity). There are several possibilities for solving the linear system in the policy iteration method. In this paper, we use a preconditioned **Bi-CGSTAB** iterative method for solving the sparse matrix [64]. We use a level one *ILU* preconditioner. Note that in general, the stencil changes at each policy iteration, hence we must recompute the symbolic *ILU* at each policy iteration.

4.0.1 Convergence of the policy iteration

If $\mathbf{A}(\mathcal{Q}), \mathbf{C}(\mathcal{Q})$ are continuous functions of the control $\mathcal{Q}$, then convergence of the policy iteration is well known, see for example [52]. In fact, for the continuous case, superlinear convergence can be established [12]. However, we remind the reader that use of *central difference as much as possible* methods result in $\mathbf{A}(\mathcal{Q}), \mathbf{C}(\mathcal{Q})$ being possibly discontinuous functions of the control. Hence, in order to ensure convergence of of Algorithm 4.1 in the general case, we follow along the lines in [48].

Theorem 4.1. *(Convergence of policy iteration) If Properties 4.1 are satisfied, then Algorithm 4.1 converges to the unique solution of equation (4.1), for any initial iterate* $\mathbf{U}^n$*.*

Proof. For the convenience of the reader, we give a brief sketch of the proof of convergence of Policy Iteration here. See details in [48].

Note that step 4 in Algorithm 4.1 is

$$\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\mathbf{W}^{k+1} \;=\; \widehat{\mathbf{C}}\left(\mathbf{W}^k\right) \tag{4.6}$$

From Proposition 4.1, $\|\widehat{\mathbf{A}}\left(\mathbf{W}\right)^{-1}\|_\infty$, and $\|\widehat{\mathbf{C}}\left(\mathbf{W}\right)\|_\infty$ are bounded independent of $\mathbf{W}$. Then, from equation (4.6), we have that $\mathbf{W}^k$ is bounded $\forall k$.

Subtract $\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\mathbf{W}^k$ from both sides of equation (4.6) to give

$$\begin{aligned}
\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\left(\mathbf{W}^{k+1}-\mathbf{W}^k\right) &= -\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\mathbf{W}^k+\widehat{\mathbf{C}}\left(\mathbf{W}^k\right)\\
&= \sup_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{W}^k+\mathbf{C}(\mathcal{Q})\right\}\\
&\geq -\widehat{\mathbf{A}}\left(\mathbf{W}^{k-1}\right)\mathbf{W}^k+\widehat{\mathbf{C}}\left(\mathbf{W}^{k-1}\right)\\
&= 0 \qquad (4.7)
\end{aligned}$$

where the last line follows from writing equation (4.6) for $k-1$.

Since $\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)$ is an M-matrix, from equation (4.7), it follows that $\mathbf{W}^{k+1}-\mathbf{W}^k\geq 0$. Since $\mathbf{W}^{k+1}$ are nondecreasing and bounded, then the iteration converges to a vector $\mathbf{W}^{\infty}$. Since $\widehat{\mathbf{A}}$ is bounded, we have

$$\begin{aligned}
\lim_{k\to\infty}\widehat{\mathbf{A}}\left(\mathbf{W}^k\right)\left(\mathbf{W}^{k+1}-\mathbf{W}^k\right) &= 0\\
&= \lim_{k\to\infty}\sup_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{W}^k+\mathbf{C}(\mathcal{Q})\right\}\\
&= \sup_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{W}^{\infty}+\mathbf{C}(\mathcal{Q})\right\}\ , \qquad (4.8)
\end{aligned}$$

since $\sup(\cdot)$ is uniformly continuous w.r.t. $\mathbf{W}^k$. Hence $\mathbf{W}^{\infty}$ is a solution to equation (4.8). Suppose we have two solutions to (4.8), $\mathbf{X}$ and $\mathbf{Y}$, then

$$0 = \sup_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{Y}+\mathbf{C}(\mathcal{Q})\right\}-\sup_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{X}+\mathbf{C}(\mathcal{Q})\right\}\leq\sup_{\mathcal{Q}\in\hat{Z}}\left\{\mathbf{A}(\mathcal{Q})(\mathbf{X}-\mathbf{Y})\right\}(4.9)$$

Since $\mathbf{A}(\mathcal{Q})$ is bounded, $\exists$ a sequence $\mathcal{Q}^j$ such that $\mathbf{A}(\mathcal{Q}^j)\to\ \bar{\mathbf{A}}$, and

$$\lim_{j\to\infty}\mathbf{A}(\mathcal{Q}^j)(\mathbf{X}-\mathbf{Y})\to\sup_{\mathcal{Q}\in\hat{Z}}\left\{\mathbf{A}(\mathcal{Q})(\mathbf{X}-\mathbf{Y})\right\}=\bar{\mathbf{A}}(\mathbf{X}-\mathbf{Y})\geq 0 \qquad (4.10)$$

Using the same steps as in the proof of Proposition 4.1, $\bar{\mathbf{A}}$ is an M-matrix, hence $\mathbf{X}\geq\mathbf{Y}$. Interchanging $\mathbf{X}$ and $\mathbf{Y}$ gives $\mathbf{Y}\geq\mathbf{X}$, hence $\mathbf{X}=\mathbf{Y}$.

□

Remark 4.1. *For nodes where $\mathbf{A}(\mathcal{Q}),\mathbf{C}(\mathcal{Q})$ are continuous functions of $\mathcal{Q}$, or where the control set $\hat{Z}$ is finite (i.e. the control set is discretized) then trivially*

$$\begin{aligned}
\widehat{\mathbf{A}}(\mathbf{W})=\mathbf{A}(\hat{\mathcal{Q}})\quad ;\quad &\widehat{\mathbf{C}}(\mathbf{W})=\mathbf{C}(\hat{\mathcal{Q}})\\
&\hat{\mathcal{Q}}\in\arg\max_{\mathcal{Q}\in\hat{Z}}\left\{-\mathbf{A}(\mathcal{Q})\mathbf{W}+\mathbf{C}(\mathcal{Q})\right\}\ . \qquad (4.11)
\end{aligned}$$

More generally, since $\hat{Z}$ is compact, we can define the optimal control as

$$\hat{\mathcal{Q}}\in\arg\max_{\mathcal{Q}\in\hat{Z}}\left\{\left(-\mathbf{A}(\mathcal{Q})\mathbf{W}+\mathbf{C}(\mathcal{Q})\right)^*\right\}\ . \qquad (4.12)$$

where $(\cdot)^*$ refers to the upper semi-continuous envelope of the argument (as a function of $\mathcal{Q}$ for fixed $\mathbf{W}$). Note that in our case, we have only a finite number of possible discontinuities in $\mathbf{A}(\mathcal{Q}),\mathbf{C}(\mathcal{Q})$.

4.1 The optimal value for $\hat{Q}_\ell^k$

We give here some details of the method used to determine the optimal control. Recall that the optimal control can be defined in general as in Remark 4.1

$$\hat{\mathcal{Q}} \in \underset{\mathcal{Q}\in\hat{Z}}{\arg\max}\left\{\left(-\mathbf{A}(\mathcal{Q})\mathbf{W}+\mathbf{C}(\mathcal{Q})\right)^*\right\} \ , \tag{4.13}$$

given a policy iterate $\mathbf{W}$.

In our case, we have only simple discontinuities in $\mathbf{A}(\mathcal{Q}), \mathbf{C}(\mathcal{Q})$ which occur when the discretization changes from central to forward/backward or vice versa. Consequently, we can determine $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{C}}$ by first determining the optimal point $\hat{\mathcal{Q}}$, and, if this corresponds to a point of discontinuity, we take the appropriate limiting value of $\mathbf{A}(\mathcal{Q}), \mathbf{C}(\mathcal{Q})$.

For $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_w \cup \Omega_{w^*}$, we have to discretize the set ∂Z (2.6), and determine the optimal value for $\hat{Q}_\ell$ by using linear search over the discrete set ∂Z_h (2.37).

For $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_f$, we firstly determine the optimal $\hat{\rho}_\ell$. The discretized cross derivative term $\left(\Gamma_{12}^h(\rho)\right)_\ell$ (either (2.15) or (2.16)) depends on the sign of the correlation ρ. The choice of the optimal $\hat{\rho}_\ell$ is as follows:

$$\hat{\rho}_\ell = \begin{cases} \rho_{\max}, & \rho_{\max}\left(\Gamma_{12}^h(\rho_{\max})\right)_\ell \geq \rho_{\min}\left(\Gamma_{12}^h(\rho_{\min})\right)_\ell , \\ \rho_{\min}, & \rho_{\max}\left(\Gamma_{12}^h(\rho_{\max})\right)_\ell < \rho_{\min}\left(\Gamma_{12}^h(\rho_{\min})\right)_\ell . \end{cases} \tag{4.14}$$

Given an arbitrary pair of the volatility values (σ_1, σ_2), this choice maximizes the objective function.

Then, suppose that we only preselect a forward or backward difference depending on the sign of drift term terms (2.1) in order to discretize first order derivative terms. Then, the form of the discretized linear operator L_f^Q (2.17) is independent of the volatilities, and $\mathbf{A}(Q_\ell)$ is a continuous function of the volatilities. In addition, $\mathbf{C}_\ell(Q_\ell)$ (4.3) is constant with respect to Q_ℓ in this case. Therefore, we can determine the optimal volatilities $((\hat{\sigma}_1)_\ell, (\hat{\sigma}_2)_\ell)$ in a straightforward fashion. By inserting the optimal $\hat{\rho}_\ell$ and the discrete diffusion terms $\left(\Gamma_{kl}^h\right)_\ell, \ k,l = 1,2$ into (4.13), a quadratic-form optimization with linear constraints needs to be solved. The form is equivalent to inserting $\hat{\rho}_\ell$ and $\left(\Gamma_{kl}^h\right)_\ell$ into (2.9). Restricting the control set to ∂Z, then the linear constraint is

$$\begin{aligned}(\sigma_1, \sigma_2) \in \Sigma \equiv & \left\{\sigma_{1,\min} \times [\sigma_{2,\min}, \sigma_{2,\max}]\right\} \cup \left\{\sigma_{1,\max} \times [\sigma_{2,\min}, \sigma_{2,\max}]\right\} \\ & \cup \left\{\sigma_{2,\min} \times (\sigma_{1,\min}, \sigma_{1,\max})\right\} \cup \left\{\sigma_{2,\max} \times (\sigma_{1,\min}, \sigma_{1,\max})\right\} .\end{aligned} \tag{4.15}$$

We then can obtain an analytical solution to a quadratic optimization problem.

However, if central weighting for the first order derivative terms is used as much as possible in L_f^Q in order to discretize the first order derivative terms, the form of the discretization at $((S_1)_i, (S_2)_j, \tau^{n+1})$ is dependent on the volatilities, thus $\mathbf{A}_{\ell,k}(Q_\ell)$ (4.2) will not, in general, be a continuous of function of the volatilities. However, as shown in the last section, the proof of the convergence of the policy iterative algorithm does not require continuity of the local objective function. As in [72], we use Algorithm 4.2 to determine the optimal volatility values. Considering node $((S_1)_i, (S_2)_j, \tau^{n+1})$, with the current solution estimate $\mathbf{W}$ in Algorithm 4.1, the optimal $\hat{\rho}_\ell$ is determined as in (4.14). Suppose the subsets of (σ_1, σ_2), which give a positive coefficient discretization, for central, forward and backward differencing respectively, are $\Sigma_\ell^{forward}$, $\Sigma_\ell^{backward}$ and $\Sigma_\ell^{central}$. Without loss of generality, suppose the sign of the drift terms are positive in (2.1), thus we only need to select between forward and central differencing. Since central differencing is the most accurate, it should be used as much as possible. That is, $\Sigma_\ell^{forward} = \Sigma - \Sigma_\ell^{central}$.

Algorithm 4.2 Determining the Optimal Control $\hat{Q}_\ell$ and the Differencing Method

```
Determine the optimal $\hat{\rho}_\ell = \begin{cases} \rho_{\max}, & \rho_{\max}\left(\Gamma_{12}^h(\rho_{\max})\right)_\ell \geq \rho_{\min}\left(\Gamma_{12}^h(\rho_{\min})\right)_\ell \\ \rho_{\min}, & \rho_{\max}\left(\Gamma_{12}^h(\rho_{\max})\right)_\ell < \rho_{\min}\left(\Gamma_{12}^h(\rho_{\min})\right)_\ell \end{cases}$
Compute the positive coefficient sets $\Sigma_\ell^{central}$ and $\Sigma_\ell^{forward}$ for $(\sigma_1, \sigma_2)$.
differencing = central, $((\hat{\sigma}_1)_\ell, (\hat{\sigma}_2)_\ell) = (0, 0)$, $F_{\max} = -\infty$
for d = central, forward do
    Solve $(\sigma_1^d, \sigma_2^d) \in \arg\max_{(\sigma_1,\sigma_2)\in\bar{\Sigma}_\ell^d} \left[-\mathbf{A}(\sigma_1, \sigma_2, \hat{\rho}_\ell)\mathbf{W} + \mathbf{C}(\sigma_1, \sigma_2, \hat{\rho}_\ell)\right]_\ell^*$
    if $\left[-\mathbf{A}(\sigma_1^d, \sigma_2^d, \hat{\rho}_\ell)\mathbf{W} + \mathbf{C}(\sigma_1^d, \sigma_2^d, \hat{\rho}_\ell)\right]_\ell^* > F_{\max}$ then
        differencing = d, $((\hat{\sigma}_1)_\ell, (\hat{\sigma}_2)_\ell) = (\sigma_1^d, \sigma_2^d)$,
    end if
    $\hat{Q}_\ell = ((\hat{\sigma}_1)_\ell, (\hat{\sigma}_2)_\ell, \hat{\rho}_\ell)$
end for
```

In Algorithm 4.2, we compute the positive coefficients set $\Sigma_\ell^{central}$ and $\Sigma_\ell^{forward}$. For a given differencing method, the range of possible values of the volatilities is divided into segments where the objective function is smooth. That is, central differencing or forward differencing can be used on disjoint intervals of Σ (4.15). On each of the subintervals, we need to maximize a quadratic problem with a linear constraint. Thus, standard methods are then used to determine the maximum within each interval, and an analytic expression for the local objective function is available. Note that in Algorithm 4.2, we compute the maximum on the closure of the sets $\Sigma_\ell^{central}$, $\Sigma_\ell^{forward}$, which we denote by $\bar{\Sigma}_\ell^{central}$, $\bar{\Sigma}_\ell^{forward}$, which ensures that the maximum of the upper semi-continuous envelope is attained.

Remark 4.2. *For each spatial node (i, j), we can pre-compute the range of Σ (4.15), where central, forward and backward differencing give rise to a positive coefficient method, and use the precomputed ranges $\Sigma_\ell^{central}$, $\Sigma_\ell^{forward}$ and $\Sigma_\ell^{backward}$ at each step in the policy iteration.*

4.2 Complexity: comparison of implicit and explicit methods

Each time step requires the solution of a local optimization problem at each grid node. We consider the worst case where the wide stencil is used and the control is discretized. We have shown that the numerical scheme only needs to perform a linear search along the boundary of the control set, instead of the entire three dimensional space Z. This finding decreases the complexity of evaluating the objective function from $O(\frac{1}{h^3})$ to $O(\frac{1}{h})$ for each node. Thus, with total a $O(\frac{1}{h^2})$ nodes, this gives a complexity $O(\frac{1}{h^3})$ for solving the local optimization problems at each time step. When using a fully implicit timestepping method, we also need to use policy iterations to advance time. The time complexity of solving the sparse M-matrix in each policy iteration is $O((\frac{1}{h^2})^{5/4})$ [64]. Assuming that the number of policy iterations is bounded, as the mesh size tends to zero, which is in fact observed in our experiments, the complexity of the time advance is thus dominated by the solutions of the local optimization problems. Finally, the total complexity is $O(\frac{1}{h^4})$ with the number of time steps $O(\frac{1}{h})$.

In the existing literature [30, 14], the wide stencil method and an explicit timestepping technique is typically used to solve HJB equations. The complexity of our numerical scheme in the worst case is the same as for an explicit method, using a wide stencil method, since the spatial derivatives are computed on a mesh spacing of size $\sqrt{h}$ [30]. However, the complexity estimate also holds for the hybrid scheme, whereby a mixture of fixed and wide stencils are used, since fully implicit timestepping does not have any stability restrictions. On the contrary, if a fixed stencil is used at even a single node, the number of time steps for an explicit method becomes $O(\frac{1}{h^2})$ instead of $O(\frac{1}{h})$ (for a pure wide stencil scheme). Note that for nodes where a fixed stencil is used, the analytical solution of the local optimization problem has $O(1)$ complexity.

The worst case for the implicit method compared to an explicit method (e.g. see [30]) results in both methods having the same complexity per timestep. The implicit methods will undoubtedly have a larger constant in the order relation compared to an explicit method. Hence the overall efficiency will be purely dependent on the total number of timesteps. Since the number of timesteps for an implicit method is completely decoupled from the mesh size parameter h, we can certainly envision cases (e.g. barrier options) where a small spatial mesh parameter is required for accuracy. In this case, an explicit method would require that timesteps be directly tied to this mesh size, which may be very small, while the implicit method may use only the timestep required to minimize time truncation error. Of course, these effects will be highly problem dependent. Finally, we note that an implicit method, which is unconditionally stable, may be preferred in a production environment with inexperienced users.

4.3 Summary

The main results in this chapter are the following:

- We used fully implicit timestepping to build an unconditionally monotone numerical scheme. Implicit timestepping then requires solution of highly nonlinear algebraic equations at each time step, which are solved using the policy iteration algorithm. Our numerical discretization depends on the control, and thus results in a local discontinuous function of the control. However, we can prove that policy iteration is still guaranteed to converge.

- In our numerical scheme, the cost of constructing the data structure and solving the matrix at each timestep is dominated by the cost of solving the local optimization problems at each grid node. Therefore, the total complexity is the same as for an explicit method at each timestep using a wide stencil discretization, but there are no time step restrictions due to stability considerations. Unconditional stability also permits efficient use of the hybrid scheme (fixed stencil as much as possible).

Chapter 5

Numerical Solution for Continuous Time Mean Variance Asset Allocation under Stochastic Volatility

In this chapter, we study the numerical solution of the HJB formulation for continuous time mean variance asset allocation under stochastic volatility. This chapter is organized as follows: Section 5.1 describes the underlying processes and the embedding framework, and gives a formulation of an associated HJB equation and a linear PDE. Section 5.2 presents the discretization of the HJB equation. Section 5.3 highlights some important implementation details of the numerical method. Numerical results are presented and discussed in Section 5.4.

5.1 Mathematical formulation

Suppose there are two assets in the market: one is a risk free bond and the other is a risky equity index. The dynamics of the risk free bond B follows

$$dB(t) = rB(t)dt, \tag{5.1}$$

and an equity index S follows Heston's model [45] under the real probability measure

$$\frac{dS(t)}{S(t)} = (r + \xi V(t))dt + \sqrt{V(t)}dZ_1, \tag{5.2}$$

where the variance of the index, $V(t)$, follows a mean-reverting square-root process [24]:

$$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}dZ_2, \tag{5.3}$$

with dZ_1, dZ_2 being increments of Wiener processes. The instantaneous correlation between Z_1 and Z_2 is $dZ_1 dZ_2 = \rho dt$. The market price of volatility risk is $\xi V(t)$, which generates a risk premium proportional to $V(t)$. This assumption for the risk premium is based on Breeden's consumption-based model [15], and originates from [45]. Therefore, under this setup, the market is incomplete as trading in the risky asset and the bond cannot perfectly hedge the changes in the stochastic investment opportunity set.

An investor in this market is endowed at time zero with an initial wealth of w_0, and she can continuously and dynamically alter the proportion of wealth invested in each asset. In addition, let $W(t) = S(t) + B(t)$ denote the wealth at time t, let p denote the proportion of this wealth invested in the risky asset $S(t)$, consequently $(1-p)$ then denotes the fraction of wealth invested in the risk free bond $B(t)$. The allocation strategy is a function of the current state, i.e., $p(\cdot) : (W(t), V(t), t) \to p = p(W(t), V(t), t)$. Note that in using the shorthand notations $p(\cdot)$ for the mapping, p for the value $p = p(W(t), V(t), t)$, and the dependence on the current state is implicit. From (5.1) and (5.2), we see that the investor's wealth process follows:

$$dW(t) = (r + p\xi V(t))\, W(t)dt + p\sqrt{V} W(t) dZ_1. \tag{5.4}$$

5.1.1 Efficient frontiers and embedding methods

We assume here that the investor is guided by a pre-commitment mean variance objective based on the final wealth $W(T)$. The pre-commitment mean variance problem and its variations have been intensively studied in the literature [53, 76, 10, 75, 57]. To best of our knowledge, there is no explicit closed-form solution for the pre-commitment mean variance problem when the risky asset follows a stochastic volatility process along with leverage constraints.

To simplify notation, we define $x = (w, v) = (W(t), V(t))$ for the state space. Let $E^{x,t}_{p(\cdot)}[W(T)]$ and $Var^{x,t}_{p(\cdot)}[W(T)]$ denote the expectation and variance of the terminal wealth conditional on the state (x, t) and the control $p(\cdot)$. Given a risk level $Var^{x,t}_{p(\cdot)}[W(T)]$, an investor desires her expected terminal wealth $E^{x,t}_{p(\cdot)}[W(T)]$ to be as large as possible. Equivalently, given an expected terminal wealth $E^{x,t}_{p(\cdot)}[W(T)]$, she wishes the risk $Var^{x,t}_{p(\cdot)}[W(T)]$ to be as small as possible. That is, she desires to find controls $p(\cdot)$ which generate Pareto optimal points. For notational simplicity, let $E^{x,t}_{p(\cdot)}[W(T)] = \mathcal{E}$ and $Var^{x,t}_{p(\cdot)}[W(T)] = \mathcal{V}$. The problem is rigorously formulated as follows.

Define the achievable mean variance objective set as

$$\mathcal{Y} = \{(\mathcal{V}, \mathcal{E}) : \ p \in \mathcal{Z}\}, \tag{5.5}$$

where $\mathcal{Z}$ is the set of admissible strategies, and denote the closure of $\mathcal{Y}$ by $\bar{\mathcal{Y}}$.

Definition 5.1. *A point $(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}$ is Pareto mean variance optimal if there exists no admissible strategy $\bar{p} \in \mathcal{Z}$ such that*

$$\begin{aligned} Var^{x,t}_{\bar{p}}\{W(T)\} &\le \mathcal{V}, \\ E^{x,t}_{\bar{p}}\{W(T)\} &\ge \mathcal{E}, \end{aligned} \tag{5.6}$$

where at least one of the inequalities in equation is strict. We denote by $\mathcal{P}$ the set of Pareto mean variance optimal points. Note that $\mathcal{P} \subseteq \bar{\mathcal{Y}}$.

Although the above definition is intuitive, determining the points in $\mathcal{P}$ requires solution of a multi-objective optimization problem, involving two conflicting criteria. A standard scalarization method can be used to combine the two criteria into an optimization problem with a single objective. In particular, for each point $(\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}}$, and for an arbitrary scaler $\lambda > 0$, we define the set of points $\mathcal{Y}_{P(\lambda)}$ to be

$$\mathcal{Y}_{P(\lambda)} = \left\{ (\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}} : \inf_{(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}} (\lambda \mathcal{V}_* - \mathcal{E}_*) \right\}, \tag{5.7}$$

from which a point on the efficient frontier can be derived. The set of points on the efficient frontier are then defined as

$$\mathcal{Y}_P = \bigcup_{\lambda > 0} \mathcal{Y}_{P(\lambda)}. \tag{5.8}$$

Note that there is a difference between the set of all Pareto mean variance optimal points $\mathcal{P}$ (see Definition 5.1) and the efficient frontier $\mathcal{Y}_P$ (5.8) [69]. In general,

$$\mathcal{P} \subseteq \mathcal{Y}_P,$$

but the converse may not hold if the achievable mean variance objective set $\mathcal{Y}$ (5.5) is not convex. In this chapter, we restrict our attention to constructing $\mathcal{Y}_P$ (5.8).

Due to the presence of the variance term $Var_{p(\cdot)}^{x,t}[W(T)]$ in (5.7), a dynamic programming principle cannot be directly applied to solve this problem. To overcome this difficulty, we make use of the main result in [53, 76, 69] which essentially involves the embedding technique. This result is summarized in the following Theorem.

Assumption 5.1. *We assume that $\mathcal{Y}$ is a non-empty subset of $\{(\mathcal{V}, \mathcal{E}) \in \mathbb{R}^2 : \mathcal{V} > 0)\}$ and that there exists a positive scalarization parameter $\lambda_E > 0$ such that $\mathcal{Y}_{P(\lambda_E)} \neq \emptyset$.*

Theorem 5.1. *The embedded mean variance objective set $\mathcal{Y}_Q$ is defined by*

$$\mathcal{Y}_Q = \bigcup_{-\infty < \gamma < \infty} \mathcal{Y}_{Q(\gamma)}, \tag{5.9}$$

where

$$\mathcal{Y}_{Q(\gamma)} = \left\{ (\mathcal{V}_*, \mathcal{E}_*) \in \bar{\mathcal{Y}} : \mathcal{V}_* + \mathcal{E}_*^2 - \gamma \mathcal{E}_* = \inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}} (\mathcal{V} + \mathcal{E}^2 - \gamma \mathcal{E}) \right\}. \tag{5.10}$$

If Assumption 5.1 holds and $\lambda > \lambda_E$, then $\mathcal{Y}_{P(\lambda)} \neq \emptyset$. Assume $(\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{P(\lambda)}$. Then if

$$\lambda \mathcal{V}_0 - \mathcal{E}_0 = \inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}} (\lambda \mathcal{V} - \mathcal{E}), \tag{5.11}$$

then

$$\mathcal{V}_0 + \mathcal{E}_0^2 - \gamma \mathcal{E}_0 = \inf_{(\mathcal{V}, \mathcal{E}) \in \mathcal{Y}} (\mathcal{V} + \mathcal{E}^2 - \gamma \mathcal{E}), \quad \text{i.e. } (\mathcal{V}_0, \mathcal{E}_0) \in \mathcal{Y}_{Q(\gamma)}, \tag{5.12}$$

where $\gamma = \frac{1}{\lambda} + 2\mathcal{E}_0$. Consequently, $\mathcal{Y}_P \subseteq \mathcal{Y}_Q$.

Proof. See details in [53, 76, 29]. □

Theorem 5.1 states that the mean and variance $(\mathcal{V}, \mathcal{E})$ of $W(T)$ are embedded in a scalarization optimization problem with the objective function being $\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}$. Noting that

$$\begin{aligned} \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} &= E^{x,t}_{p(\cdot)}[W^2(T)] - (E^{x,t}_{p(\cdot)}[W(T)])^2 + (E^{x,t}_{p(\cdot)}[W(T)])^2 - \gamma E^{x,t}_{p(\cdot)}[W(T)] \\ &= E^{x,t}_{p(\cdot)}[W^2(T) - \gamma W(T)] \\ &= E^{x,t}_{p(\cdot)}[(W(T) - \frac{\gamma}{2})^2] + \frac{\gamma^2}{4}, \end{aligned} \tag{5.13}$$

and that we can ignore the constant $\frac{\gamma^2}{4}$ term for the purposes of minimization, we then define the value function

$$\mathcal{U}(x,t) = \inf_{p(\cdot)\in\mathcal{Z}} E^{x,t}_{p(\cdot)}[(W(T) - \frac{\gamma}{2})^2]. \tag{5.14}$$

Theorem 5.1 implies that there exists a γ, such that, for a given positive λ, a control p^* which minimizes (5.7) also minimizes equation (5.14). Dynamic programming can then be directly applied to equation (5.14) to determine the optimal control $p^*(\cdot)$.

The procedure for determining the points on the efficient frontier is as follows. For a given value of γ, the optimal strategy p^* is determined by solving for the value function problem (5.14). Once this optimal policy $p^*(\cdot)$ is known, it is then straightforward to determine a point $(Var^{x,t}_{p^*(\cdot)}[W(T)], E^{x,t}_{p^*(\cdot)}[W(T)])$ on the frontier. Varying γ traces out a curve in the $(\mathcal{V}, \mathcal{E})$ plane (see details in Section 5.3.2). Consequently, the numerical challenge is to solve for the value function (5.14). More precisely, the above procedure for constructing the efficient frontier generates points that are in the set $\mathcal{Y}_Q$. As pointed out in [69], the set $\mathcal{Y}_Q$ may contain spurious points, i.e., points which are not in $\mathcal{Y}_P$. For example, when the original problem is nonconvex, spurious points can be generated. An algorithm for removing spurious points is discussed in [69]. The set of points in $\mathcal{Y}_Q$ with the spurious points removed generates all points in $\mathcal{Y}_P$. Reference [29] also discusses the convergence of finitely sampled γ to the efficient frontier.

Remark 5.1 (Range of γ)**.** *As noted in [29], a solution to problem (5.10) generally exists $\forall\gamma \in (-\infty, +\infty)$. However, we know from the above discussion, that some of these solutions may be spurious. In some cases, we can use financial reasoning to reduce the range of γ so that obvious spurious points are eliminated. We discuss this further in Section 5.3.2.*

5.1.2 The value function problem

Following standard arguments, the value function $\mathcal{U}(w, v, \tau)$, $\tau = T - t$ (5.14) is the viscosity solution of the HJB equation

$$\mathcal{U}_\tau = \inf_{p\in\mathcal{Z}} \left\{ (r + p\xi v)w\mathcal{U}_w + \kappa(\theta - v)\mathcal{U}_v + \frac{1}{2}(p\sqrt{v}w)^2\mathcal{U}_{ww} + p\rho\sigma\sqrt{v}w\mathcal{U}_{wv} + \frac{1}{2}\sigma^2 v\mathcal{U}_{vv} \right\}, \tag{5.15}$$

on the domain $(w, v, \tau) \in [0, +\infty] \times [0, +\infty] \times [0, T]$, and with the terminal condition

$$\mathcal{U}(w, v, 0) = \left(w - \frac{\gamma}{2}\right)^2. \tag{5.16}$$

Remark 5.2. *In one of our numerical tests, we allow p to become unbounded, which may occur when $w \to 0$ [73]. However, although $p \to \infty$ as $w \to 0$, we must have $(pw) \to 0$ as $w \to 0$, i.e., the amount invested in the risky asset converges to zero as $w \to 0$. This is required in order to ensure that the no-bankruptcy boundary condition is satisfied [73]. As a result, we can then formally eliminate the problem with unbounded control by using $q = pw$ as the control, and assume q remains bounded. See details in [73].*

5.1.3 The expected wealth problem

The PDE formulation

Given the solution for the value function (5.14), with the optimal control $p^*(\cdot)$. We then need to determine the expected value $E^{x,t}_{p^*(\cdot)}[W(T)]$, denoted as

$$\mathcal{E}(w, v, t) = E^{x,t}_{p^*(\cdot)}[W(T)], \tag{5.17}$$

Then, $\mathcal{E}(w, v, \tau),\ \tau = T - t$ is given from the solution to the following linear PDE

$$\mathcal{E}_\tau = (r + p^*\xi v)w\mathcal{E}_w + \kappa(\theta - v)\mathcal{E}_v + \frac{1}{2}(p^*\sqrt{v}w)^2\mathcal{E}_{ww} + p^*\rho\sigma\sqrt{v}w\mathcal{E}_{wv} + \frac{1}{2}\sigma^2 v\mathcal{E}_{vv} \tag{5.18}$$

with the initial condition $\mathcal{E}(w, v, 0) = w$, where p^* is obtained from the solution of the HJB equation (5.15).

The Hybrid (PDE - Monte Carlo) method

Alternatively, given the stored control $p^*(\cdot)$ determined from the solution of equation (5.15), we can directly estimate $(Var^{x,t}_{p^*(\cdot)}[W(T)], E^{x,t}_{p^*(\cdot)}[W(T)])$ by using a Monte Carlo method, based on solving the SDEs (5.3-5.4). The details of the SDE discretization are given in Section 5.3.2. This hybrid(PDE - Monte Carlo) method was originally proposed in [68].

5.1.4 Allowable portfolios

In order to obtain analytical solutions, many previous papers typically make assumptions which allow for the possibility of unbounded borrowing and bankruptcy. Moreover, these models assume a bankrupt investor can still keep on trading. The ability to continue trading even though the value of an investor's wealth is negative is highly unrealistic. In this chapter, we enforce the condition that the wealth value remains in the solvency regions by applying certain boundary conditions to the HJB equation [72]. Thus, bankruptcy is prohibited, i.e.,

$$w \in [0, +\infty).$$

We will also assume that there is a leverage constraint, i.e., the investor must select an asset allocation satisfying

$$p = \frac{\text{The risky asset value}}{\text{The total wealth}} = \frac{pW(t)}{W(t)} < p_{\max},$$

which can be interpreted as the maximum leverage condition, and $p_{\max}$ is a known positive constant with typical value in $[1.0, 2.0]$. Thus, the control set

$$p \in \mathcal{Z} = [0, p_{\max}].$$

Note that when the risk premium ξ (5.2) is positive, it is not optimal to short the risky asset, since we have only a single risky asset in our portfolio. In some circumstances, it may be optimal to short the risky asset. This will be discussed in Section 5.2.1.

5.2 Numerical discretization of the HJB equation

5.2.1 Localization

We will assume that the discretization is posed on a bounded domain for computational purposes. The discretization is applied to the localized finite region $(w, v) \in [0, w_{\max}] \times [0, v_{\max}]$. Asymptotic boundary conditions will be imposed at $w = w_{\max}$ and $v = v_{\max}$ which are compatible with a monotone numerical scheme.

The localization of v

The proper boundary on $v = 0$ needs to be specified to be compatible with the corresponding SDE (5.3), which has a unique solution [37]. If $2\kappa\theta \geq \sigma^2$, the so-called Feller condition holds, and $v = 0$ is unattainable. If the Feller condition is violated, $2\kappa\theta < \sigma^2$, then $v = 0$ is an attainable boundary but is strongly reflecting [37]. The appropriate boundary condition can be obtained by setting $v = 0$ into equation (5.15). That is,

$$\mathcal{U}_\tau = rw\mathcal{U}_w + \kappa\theta\mathcal{U}_v, \tag{5.19}$$

and the equation degenerates to a linear PDE. On the lower boundary $v = 0$, the variance and the risk premium vanishes, according to (5.4), so that the wealth return is always the risk free rate r. The control value p vanishes in the degenerate equation (5.19), and we can simply define $p^*(w, v = 0, t) \equiv 0$ which we need in the estimation of $(Var_{p^*(\cdot)}^{x,t}[W(T)], E_{p^*(\cdot)}^{x,t}[W(T)])$

using the Monte Carlo simulation. In this case, since the equity asset has zero volatility with drift rate r, the distinction between the equity and risk free asset is meaningless.

The validity of this boundary condition is intuitively justified by the fact that the solution to the SDE for v is unique, such that the behavior of v at the boundary $v = 0$ is determined by the SDE itself, and hence the boundary condition is determined by setting $v = 0$ in equation (5.15). A formal proof that this boundary condition is correct is given in [36]. If the boundary at $v = 0$ is attainable, then this boundary behaviour serves as a boundary condition and guarantees uniqueness in the appropriate function spaces. On the other hand, if the boundary is non-attainable, then the boundary behaviour is not needed to guarantee uniqueness, but is nevertheless very useful in a numerical scheme.

On the upper boundary $v = v_{\max}$, $\mathcal{U}_v$ is set to zero. Thus, the boundary condition on $v_{\max}$ is set to

$$\mathcal{U}_\tau = \inf_{p \in \mathcal{Z}} \left\{ (r + p\xi v) w \mathcal{U}_w + \frac{1}{2}(p\sqrt{v}w)^2 \mathcal{U}_{ww} \right\}. \tag{5.20}$$

The optimal control p^* at $v = v_{\max}$ is determined by solving the equation (5.20). This boundary condition can be justified by noting that as $v \to \infty$, then the diffusion term in the w direction in equation (5.15) becomes large. As well, the initial condition (5.16) is independent of v. As a result, we expect that

$$\mathcal{U} \approx C'w + C'', \quad v \to \infty,$$

where C' and C'' are constants, and hence $\mathcal{U}_v \approx 0$ at $v = v_{\max}$.

The localization for w

We prohibit the possibility of bankruptcy ($W(t) < 0$) by requiring that $\lim_{w \to 0}(pw) = 0$ [73], so, on $w = 0$, the equation (5.15) reduces to

$$\mathcal{U}_\tau = \kappa(\theta - v)\mathcal{U}_v + \sigma^2 v \mathcal{U}_{vv}. \tag{5.21}$$

When $w \to +\infty$, we assume that the asymptotic form of the exact solution is

$$\mathcal{U}(w \to +\infty, v, \tau) = \bar{\mathcal{U}}(w) = H_2(\tau)w^2 + H_1(\tau)w + H_0(\tau), \tag{5.22}$$

and make the assumption that $p^*(w_{\max}, v, 0)$ at $w = w_{\max}$ is set to zero. That is, once the investor's wealth is very large, she prefers the risk free asset. This can be justified from the arguments in [26, 27].

Alternative localization for w

$\mathcal{U}(w, v, \tau)$ is the viscosity solution of the HJB equation (5.15). Recall that the initial condition for problem (5.14) is

$$\mathcal{U}(w, v, 0) = \left(W(T) - \frac{\gamma}{2} \right)^2.$$

For a fixed gamma, we define the discounted optimal embedded terminal wealth at time t, denoted by $W_{\text{opt}}(t)$, as

$$W_{\text{opt}}(t) = \frac{\gamma}{2} e^{-r(T-t)}. \tag{5.23}$$

It is easy to verify that $W_{\text{opt}}(t)$ is a globally minimum state of the value function $\mathcal{U}(w, v, t)$. Consider the state $(W_{\text{opt}}(t), v), t \in [0, T]$, and the optimal strategy $p^*(\cdot)$ such that $p^*(w, v, \mathcal{T}) \equiv 0,\ \mathcal{T} > t$. Under $p^*(\cdot)$, the wealth is all invested in the risk free bond without further re-balancing from time t. As a result, the wealth will accumulate to $W(T) = \frac{\gamma}{2}$ with certainty, i.e., the optimal embedded terminal wealth $\frac{\gamma}{2}$ is achievable. By definition (5.14), we have,

$$\mathcal{U}(W_{\text{opt}}(t), v, t) = \inf_{p(\cdot) \in \mathcal{Z}} \left\{ E^{x,t}_{p(\cdot)}[(W(T) - \frac{\gamma}{2})^2] \right\} = E^{x,t}_{p^*(\cdot)}[(W(T) - \frac{\gamma}{2})^2] = 0. \tag{5.24}$$

Since the value function is the expectation of a non-negative quantity, it can never be less than zero. Then, the exact solution for the value function problem at the special point $W_{\text{opt}}(t)$ must be zero. This result holds for both the discrete and continuous re-balancing case. For the formal proof, we refer the reader to [27].

Consequently, the point $w = \frac{\gamma}{2} e^{-r\tau}$ is a Dirichlet boundary $\mathcal{U}(\frac{\gamma}{2} e^{-r\tau}, v, \tau) = 0$, and information for $w > \frac{\gamma}{2} e^{-r\tau}$ is not needed. In principle, we can restrict the domain to $0 \le w \le \frac{\gamma}{2} e^{-r\tau}$. However, it is computationally convenient to restrict to the size of the computational domain to be $0 \le w \le \frac{\gamma}{2}$, which avoids issues with a moving boundary, at a very small additional cost. Note that the optimal control will ensure that $\mathcal{U}(\frac{\gamma}{2} e^{-r\tau}, v, \tau) = 0$ without any need to enforce this boundary condition. This will occur since we assume continuous rebalancing. This effect that $W(t) \le W_{\text{opt}}(t)$ is also discussed in [71]. It is interesting to note that, in the case of discrete rebalancing that it is optimal to withdraw cash from the portfolio if it is ever observed that $W(t) > W_{\text{opt}}(t)$. On the other hand, [9] show that if the market is complete, then it is never optimal to withdraw cash from the portfolio. This is discussed in [26, 27].

In the case of an incomplete market, such as discrete rebalancing or jump diffusions, if we do not allow withdrawing cash from the portfolio, then the investor has an incentive to lose money if $W(t) > W_{\text{opt}}(t)$, as pointed out in [26]. In this rather perverse situation, it may be optimal to short the risky asset, so that the admissible set in this would be $\mathcal{Z} = [p_{\min}, p_{\max}]$ with $p_{\min} < 0$.

We have verified, experimentally, that restricting the computational domain to $w \in [0, \gamma/2]$ gives the same results as the domain $w \in [0, w_{\max}]$, $w_{\max} \gg \frac{\gamma}{2}$, with asymptotic boundary condition (5.22).

Remark 5.3 (Significance of $W(t) \le W_{\text{opt}}(t)$)**.** *If we assume that initially $W(0) < W_{opt}(0)$ (otherwise the problem is trivial if we allow cash withdrawals), then the optimal control will ensure that $W(t) \le W_{opt}(t), \forall t$. Hence continuous time mean variance optimization is* time consistent in efficiency *[26]. Another interpretation is that continuous time mean variance optimization is equivalent to minimizing the quadratic loss with respect to the wealth target $W_{opt}(T)$[71].*

Remark 5.4 (Significance of $W(T) \leq \gamma/2$)**.** *From Remark 5.3 we have trivially that $W(T) \leq \gamma/2$, hence from equation (5.14), the investor is never penalized for large gains, i.e. the quadratic utility function (5.14) is always well behaved. Consequently, continuous time mean variance optimization is fundamentally different from the single period counterpart.*

5.2.2 Discretization

In the following section, we discretize equation (5.15) over a finite grid $N = N_1 \times N_2$ in the space (w, v). Define a set of nodes $\{w_1, w_2, \ldots, w_{N_1}\}$ in the w direction and $\{v_1, v_2, \ldots, v_{N_2}\}$ in the v direction. Denote the n^{th} time step by $\tau^n = n\Delta\tau,\ n = 0, \ldots, N_\tau$, with $N_\tau = \frac{T}{\Delta\tau}$. Let $\mathcal{U}_{i,j}^n$ be the approximate solution of the equation (5.15) at (w_i, v_j, τ^n).

It will be convenient to define

$$\begin{aligned} \Delta w_{\max} &= \max_i \left(w_{i+1} - w_i\right), \quad \Delta w_{\min} = \min_i \left(w_{i+1} - w_i\right), \\ \Delta v_{\max} &= \max_i \left(v_{i+1} - v_i\right), \quad \Delta v_{\min} = \min_i \left(v_{i+1} - v_i\right). \end{aligned} \tag{5.25}$$

We assume that there is a mesh discretization parameter h such that

$$\Delta w_{\max} = C_1 h,\ \Delta w_{\min} = C_2 h,\ \Delta v_{\max} = C_1^{'} h,\ \Delta v_{\min} = C_2^{'} h,\ \Delta\tau = C_3 h, \tag{5.26}$$

where $C_1, C_2, C_1^{'}, C_2^{'}, C_3$ are constants independent of h.

In the following sections, we will give the details of the discretization for a reference node $(w_i, v_j),\ 1 < i < N_1,\ 1 < j < N_2$.

The wide stencil

Following similar arguments in Chapter 2, we need to use a wide stencil method to construct a monotone and consistent discretization for equation (5.15). We will use the wide stencil method developed in Chapter 2 to discretize the second derivative terms. For more details, we refer the reader to Section 2.3.2. Suppose we discretize equation (5.15) at grid node (i, j) for a fixed control. For a fixed p, consider a virtual rotation of the local coordinate system clockwise by the angle $\eta_{i,j}$

$$\eta_{i,j} = \frac{1}{2} \tan^{-1} \left(\frac{2\rho p \sigma w_i v_j}{(p\sqrt{v_j} w_i)^2 - (\sigma\sqrt{v_j})^2} \right). \tag{5.27}$$

That is, (x_1, x_2) in the transformed coordinate system is obtained by using the following matrix multiplication

$$\begin{pmatrix} w \\ v \end{pmatrix} = \begin{pmatrix} \cos\eta_{i,j} & -\sin\eta_{i,j} \\ \sin\eta_{i,j} & \cos\eta_{i,j} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{5.28}$$

where $\mathbf{R}_{i,j}$ is the rotation matrix. Under this grid rotation, the second order terms in equation (5.18) are, in the transformed coordinate system (x_1, x_2),

$$a_{i,j}\frac{\partial^2 \mathcal{W}}{\partial x_1^2} + b_{i,j}\frac{\partial^2 \mathcal{W}}{\partial x_2^2}, \tag{5.29}$$

where $\mathcal{W}$ is the value function $\mathcal{W}(x_1, x_2, \tau)$ in the transformed coordinate system, and

$$\begin{aligned} a_{i,j} &= \left(\frac{1}{2}(p\sqrt{v_j}w_i)^2\cos(\eta_{i,j})^2 + \rho p\sigma w_i v_j \sin(\eta_{i,j})\cos(\eta_{i,j}) + \frac{1}{2}(\sigma\sqrt{v_j})^2\sin(\eta_{i,j})^2\right), \\ b_{i,j} &= \left(\frac{1}{2}(p\sqrt{v_j}w_i)^2\sin(\eta_{i,j})^2 - \rho p\sigma w_i v_j \sin(\eta_{i,j})\cos(\eta_{i,j}) + \frac{1}{2}(\sigma\sqrt{v_j})^2\cos(\eta_{i,j})^2\right). \end{aligned} \tag{5.30}$$

Let us rewrite the HJB equation (5.15) as

$$\sup_{p\in\mathcal{Z}} \{\mathcal{U}_\tau - (r + p\xi v)w\mathcal{U}_w - \mathcal{L}^p\mathcal{U}\} = 0, \tag{5.31}$$

where the linear operator $\mathcal{L}^p$ is defined as

$$\mathcal{L}^p\mathcal{U} = \kappa(\theta - v)\mathcal{U}_v + \frac{1}{2}(p\sqrt{v}w)^2\mathcal{U}_{ww} + p\rho\sigma\sqrt{v}w\mathcal{U}_{wv} + \frac{1}{2}\sigma^2 v\mathcal{U}_{vv}. \tag{5.32}$$

The drift term $\kappa(\theta - v)\mathcal{U}_v$ in equation (5.32) is discretized by a standard backward or forward finite differencing discretization, depending on the sign of $\kappa(\theta - v)$. Overall, the discretized form of the linear operator $\mathcal{L}^p$ is then denoted by L_h^p

$$\begin{aligned} L_h^p\mathcal{U}_{i,j}^{n+1} &= 1_{\kappa(\theta-v_j)\geq 0}\frac{\kappa(\theta - v_j)}{h}\mathcal{U}_{i,j+1}^{n+1} - 1_{\kappa(\theta-v_j)<0}\frac{\kappa(\theta - v_j)}{h}\mathcal{U}_{i,j-1}^{n+1} \\ &+ \frac{a_{i,j}}{h}\mathcal{J}_h\mathcal{U}^{n+1}\left(x_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_1\right) + \frac{a_{i,j}}{h}\mathcal{J}_h\mathcal{U}^{n+1}\left(x_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_1\right) \\ &+ \frac{b_{i,j}}{h}\mathcal{J}_h\mathcal{U}^{n+1}\left(x_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_2\right) + \frac{b_{i,j}}{h}\mathcal{J}_h\mathcal{U}^{n+1}\left(x_{i,j} - \sqrt{h}(\mathbf{R}_{i,j})_2\right) \\ &- \left(1_{\kappa(\theta-v_j)\geq 0}\frac{\kappa(\theta - v_j)}{h} - 1_{\kappa(\theta-v_j)<0}\frac{\kappa(\theta - v_j)}{h} + \frac{2a_{i,j}}{h} + \frac{2b_{i,j}}{h}\right)\mathcal{U}_{i,j}^{n+1}, \end{aligned} \tag{5.33}$$

where h is the discretization parameter, and the superscript p in L_h^p indicates that the discretization depends on the control p. $\mathcal{J}_h$ is a linear interpolation operator. $x_{i,j} = \begin{pmatrix} w_i \\ v_j \end{pmatrix}$, $a_{i,j}$ and $b_{i,j}$ are given in (5.30), and $(\mathbf{R}_{i,j})_k$ is k-th column of the rotation matrix.

Remark 5.5. *In Chapter 2, we proposed a hybrid method to minimize the use of a wide stencil discretization. Suppose that we use a fixed stencil method to discretize equation (5.15), and construct the uniform grid with grid sizes Δw and Δv. The standard finite difference* *method is used to approximate the second order derivative terms, and the seven point stencil method (see Section 2.3.1) is used to approximate the cross-derivative term. Then, the grid spacing condition under which a fixed stencil is monotone is, for equation (5.15),*

$$\frac{|\rho|\sigma w_i}{p\sqrt{v_j}} \leq \frac{\Delta v}{\Delta w} \leq \frac{\sigma\sqrt{v_i}}{|\rho| p w_i}.$$

However, for $\forall p \in [0, p_{\max}]$*, such a grid spacing does not exist to ensure that the fixed stencil method is monotone. Therefore, we only use the wide stencil method for this problem. It also reminds us of the necessity of a wide stencil method when discretizing a multidimensional HJB equation.*

Semi-Lagrangian timestepping scheme

When $p \to 0$, equation (5.15) degenerates, with no diffusion in the w direction. As a result, we will discretize the drift term $(r + p\xi v)w\mathcal{U}_w$ in equation (5.15) by a semi-Lagrangian timestepping scheme in this section. Initially introduced by [34, 60] for atmospheric and weather numerical prediction problems, semi-Lagrangian schemes can effectively reduce the numerical problems arising from convection dominated equations.

Firstly, we define the Lagrangian derivative $\frac{D\mathcal{U}}{D\tau}(p)$ by

$$\frac{D\mathcal{U}}{D\tau}(p) = \mathcal{U}_\tau - (r + p\xi v)w\mathcal{U}_w, \tag{5.34}$$

which is the rate of change of $\mathcal{U}$ along the characteristic $w = w(\tau)$ defined by the risky asset fraction p through

$$\frac{dw}{d\tau} = -(r + p\xi v)w. \tag{5.35}$$

We can then rewrite equation (5.31) as

$$\sup_{p \in \mathcal{Z}} \left\{ \frac{D\mathcal{U}}{D\tau} - \mathcal{L}^p\mathcal{U} \right\} = 0. \tag{5.36}$$

Solving equation (5.35) backwards in time from τ^{n+1} and τ^n, for a fixed w_i, gives the point at the foot of the characteristic

$$(w_{i^*}, v_j) = (w_i e^{(r+p\xi v_j)\Delta\tau^n}, v_j), \tag{5.37}$$

which in general is not on the PDE grid. We use the notation $\mathcal{U}^n_{i^*,j}$ to denote an approximation of the value $\mathcal{U}(w_{i^*}, v_j, \tau^n)$, which is obtained by linear interpolation to preserve monotonicity. The Lagrangian derivative at a reference node (i, j) is then approximated by

$$\frac{D\mathcal{U}}{D\tau}(p) \approx \frac{\mathcal{U}^n_{i,j} - \mathcal{U}^n_{i^*,j}(p)}{\Delta\tau^n}, \tag{5.38}$$

where $\mathcal{U}^n_{i^*,j}(p)$ denotes that w_{i^*} depends on the control p through equation (5.37). For the details of the semi-Lagrangian timestepping scheme, we refer the reader to [18].

Finally, by using the implicit timestepping method, combining the expressions (5.33) and (5.38), the HJB equation (5.36) at a reference point (w_i, v_j, τ^{n+1}) is then discretized as

$$\sup_{p \in \mathcal{Z}_h} \left\{ \left(\frac{\mathcal{U}^{n+1}_{i,j} - \mathcal{U}^n_{i^*,j}(p)}{\Delta \tau^n} \right) - L^p_h \mathcal{U}^n_{i,j} \right\} = 0, \tag{5.39}$$

where $\mathcal{Z}_h$ is the discrete control set. Since there is no simple analytic expression which can be used to minimize the discrete equations (5.39), we need to discretize the admissible control set $\mathcal{Z}$ and perform linear search. This guarantees that we find the global maximum of equation (5.39), since the objective function has no known convexity properties. If the discretization step for the controls is also $O(h)$, where h is the discretization parameter, then this is a consistent approximation [72].

5.2.3 Matrix form of the discrete equations

Our discretization is summarized as follows. The domains are defined in Table 5.1. For the case $(w_i, v_j) \in \Omega_{in}$, we need to use a wide stencil based on a local coordinate rotation to discretize the second derivative terms, and use the semi-Lagrangian timestepping scheme to handle the drift term $(r + p\xi v)w\mathcal{U}_w$. The HJB equation is discretized as (5.39), and the optimal p^* in this case is determined by solving (5.39). For the case $\Omega_{v_{\max}}$, the HJB equation degenerates to (5.20). In this case, the drift term is also handled by the semi-Lagrangian timestepping scheme. With a vanishing cross-derivative term, the degenerate linear operator $\mathcal{L}^p$ can be discretized by a standard finite difference method. The corresponding discretized form D^p_h is given in Appendix D. The value for case $\Omega_{w_{\max}}$ is obtained by the asymptotic solution (5.22), and the optimal p^* is set to zero. At the lower boundaries $\Omega_{w_{\min}}$ and $\Omega_{v_{\min}}$, the HJB equation degenerates to a linear equation. The wide stencil and the semi-Lagrangian timestepping scheme may require the value of the solution at a point outside the computational domain, denoted as Ω_{out}. Details on how to handle this case are given in Section 5.3.3. From the discretization (5.39), we can see that the measure of Ω_{out} convergences to zero as $h \to 0$. Lastly, fully implicit time-stepping is used to ensure unconditional monotonicity of our numerical scheme. We use policy iteration to solve nonlinear algebraic equations at each timestep. We refer the reader to [48, 39] and Chapter 4 for details of the policy iteration algorithm.

It is convenient to use a matrix form to represent the discretized equations for computational purposes. Let $\mathcal{U}^n_{i,j}$ be the approximate solution of the equation (5.15) at (w_i, v_j, τ^n), $1 \le i \le N_1$, $1 \le j \le N_2$ and $0 \le \tau^n \le N_\tau$, and form the solution vector

$$\mathbf{U}^n = \left(\mathcal{U}^n_{1,1}, \mathcal{U}^n_{2,1}, \ldots, \mathcal{U}^n_{N_1,1}, \ldots, \mathcal{U}^n_{1,N_2}, \ldots, \mathcal{U}^n_{N_1,N_2} \right). \tag{5.40}$$

Notation	The domain
Ω	$[0, w_{\max}] \times [0, v_{\max}]$
Ω_{in}	$(0, w_{\max}) \times (0, v_{\max})$
$\Omega_{w_{\max}}$	The upper boundary $w = w_{\max}$
$\Omega_{v_{\max}}$	The upper boundary $v = v_{\max}$
$\Omega_{w_{\min}}$	The lower boundary $w = 0$
$\Omega_{v_{\min}}$	The lower boundary $v = 0$
Ω_{out}	$(w_{\max}, +\infty) \times (0, +\infty) \cup (0, +\infty) \times (v_{\max}, +\infty)$

Table 5.1: The domain definitions.

It will sometimes be convenient to use a single index when referring to an entry of the solution vector

$$\mathcal{U}_\ell^n = \mathcal{U}_{i,j}^n, \quad \ell = i + (j-1)N_1.$$

Let $N = N_1 \times N_2$, and we define the $N \times N$ matrix $\mathbf{L}^{n+1}(\mathcal{P})$, where

$$\mathcal{P} = \{p_1, \ldots, p_N\} \tag{5.41}$$

is an indexed set of N controls, and each p_ℓ is in the set of admissible controls. $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{P})$ is the entry on the ℓ-th row and k-th column of the discretized matrix $\mathbf{L}^{n+1}(\mathcal{P})$. We also define a vector of boundary conditions $\mathbf{F}^{n+1}(\mathcal{P})$.

For the case $(w_i, v_j) \in \Omega_{w_{\max}}$ where the Dirichlet boundary condition (5.22) is imposed, we then have

$$\mathbf{F}_\ell^{n+1}(\mathcal{P}) = \bar{\mathcal{U}}(w_{\max}), \tag{5.42}$$

and

$$\mathbf{L}_{\ell,k}^{n+1}(\mathcal{P}) = 0, \quad k = 1, \ldots, N. \tag{5.43}$$

For the case $(w_i, v_j) \in \Omega_{v_{\min}} \cup \Omega_{w_{\min}} \cup \Omega_{v_{\max}}$, the differential operator degenerates, and the entries $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{P})$ are constructed from the discrete linear operator D_h^p (see the Appendix D, equation (D.1)). That is,

$$[\mathbf{L}^{n+1}(\mathcal{P})\mathbf{U}^{n+1}]_\ell = D_h^p \mathcal{U}_{i,j}^{n+1}. \tag{5.44}$$

For the case $(w_i, v_j) \in \Omega_{in}$, we need to use the values at the following four off-grid points $x_{i,j} \pm \sqrt{h}(\mathbf{R}_{i,j})_k$, $k = 1, 2$ in (5.33), and we denote those values by $\Psi_{i,j}^m$, $m = 1, 2, 3, 4$, respectively. When $\Psi_{i,j}^m \in \Omega$, using linear interpolation, values at these four points are approximated as follows

$$\mathcal{J}_h \mathcal{U}^{n+1}(\Psi_{i,j}^m) = \begin{cases} \sum_{\substack{d=0,1 \\ e=0,1}} \omega_{i,j}^{f_m+d, g_m+e} \mathcal{U}_{f_m+d, g_m+e}^{n+1}, & \Psi_{i,j}^m \in \Omega \\ 0, & \text{otherwise} \end{cases}. \tag{5.45}$$

For linear interpolation, we have that $\omega_{i,j}^{f_m+d,g_m+e} \geq 0$ and $\sum_{\substack{d=0,1\\e=0,1}} \omega_{i,j}^{f_m+d,g_m+e} = 1$. Then, inserting (5.45) in (5.33), the entries $\mathbf{L}_{\ell,k}^{n+1}(\mathcal{P})$ on ℓ-th row are specified. When we use $\Psi_{i,j}^m \in \Omega_{out}$, we directly use its asymptotic solution $\bar{\mathcal{U}}(\Psi_{i,j}^m)$ (5.22). Thus, we need to define the vector $\mathbf{G}^{n+1}(\mathcal{P})$ to facilitate the construction of the matrix form in this situation when we use a point in the domain Ω_{out}.

$$\mathbf{G}_\ell^{n+1}(\mathcal{P}) = \begin{cases} 1_{\Psi_{i,j}^1 \in \Omega_{out}} \frac{a_{i,j}}{h} \bar{\mathcal{U}}(\Psi_{i,j}^1) + 1_{\Psi_{i,j}^2 \in \Omega_{out}} \frac{a_{i,j}}{h} \bar{\mathcal{U}}(\Psi_{i,j}^2) & \\ \qquad + 1_{\Psi_{i,j}^3 \in \Omega_{out}} \frac{b_{i,j}}{h} \bar{\mathcal{U}}(\Psi_{i,j}^3) + 1_{\Psi_{i,j}^4 \in \Omega_{out}} \frac{b_{i,j}}{h} \bar{\mathcal{U}}(\Psi_{i,j}^4), & (w_i, v_j) \in \Omega_{in}, \\ 0, & \text{otherwise} \end{cases} \tag{5.46}$$

where $a_{i,j}$ and $b_{i,j}$ are defined in equation (5.30). As a result, for the case $(w_i, v_j) \in \Omega_{in}$,

$$[\mathbf{L}^{n+1}(\mathcal{P})\mathbf{U}^{n+1}]_\ell + \mathbf{G}_\ell^{n+1}(\mathcal{P}) = L_h^p \mathcal{U}_{i,j}^{n+1}, \tag{5.47}$$

where L_h^p is defined in equation (5.33).

Let $\Phi^{n+1}(\mathcal{P})$ be a linear Lagrange interpolation operator such that

$$[\Phi^{n+1}(\mathcal{P})\mathbf{U}]_l = \begin{cases} \mathcal{J}_h \mathcal{U}_{i^*,j}^n, & (w_{i^*}, v_j) \in \Omega \\ \bar{\mathcal{U}}(w_{i^*}) \text{ (5.22)}, & (w_{i^*}, v_j) \in \Omega_{out} \end{cases}, \tag{5.48}$$

where $\mathcal{U}_{i^*,j}^n$ is defined in (5.37).

The final matrix form of the discretized equations is then

$$\begin{aligned} &\left[\mathbf{I} - \Delta\tau^n \mathbf{L}^{n+1}(\hat{\mathcal{P}})\right] \mathbf{U}^{n+1} = \Phi^{n+1}(\mathcal{P})\mathbf{U}^n + \Delta\tau^n \mathbf{G}^{n+1}(\mathcal{P}) + \mathbf{F}^{n+1} - \mathbf{F}^n, \\ &\qquad\qquad \hat{p}_\ell \in \arg\min_{p \in \mathcal{Z}_h} \left[\Phi^{n+1}(\mathcal{P})\mathbf{U}^n + \Delta\tau^n \left(\mathbf{L}^{n+1}(\mathcal{P})\mathbf{U}^{n+1} + \mathbf{G}^{n+1}(\mathcal{P})\right)\right]_\ell, \\ &\qquad\qquad\qquad \ell = i + (j-1)N_1, \quad i = 2, \ldots, N_1 - 1, \quad j = 2, \ldots, N_2, \end{aligned} \tag{5.49}$$

where $\mathcal{Z}_h$ is the discretized control set $\mathcal{Z}$.

Remark 5.6. *Note that* $[\mathbf{I} - \Delta\tau^n \mathbf{L}^{n+1}(\mathcal{P})]_{\ell,k}$, $[\Phi^{n+1}(\mathcal{P})]_\ell$ *and* $[\mathbf{G}^{n+1}(\mathcal{P})]_\ell$ *depend only on* p_ℓ.

5.2.4 Convergence to the viscosity solution

Assumption 5.2. *If the control p is bounded, Equation (5.15) satisfies the strong comparison property, hence a unique continuous viscosity solution to equation (5.15) exists [30].*

Provided that the original HJB equation satisfies Assumption 5.2, we can show that the numerical scheme (5.49) is ℓ_∞ stable, consistent and monotone, and then the scheme converges to the unique and continuous viscosity solution [7]. We give a brief overview of the proof as follows.

- **Stability:** From the formation of matrix $\mathbf{L}$ in (5.43), (5.44) and (5.47), it is easily seen that $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{P})]$ (5.49) has positive diagonals, non-positive offdiagonals, and the ℓ-th row sum for the matrix is

$$\sum_k \left[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{P})\right]_{\ell,k} > 0, \quad i = 1, \dots, N_1, \ j = 1, \dots, N_2, \tag{5.50}$$

where $\ell = i + (j-1)N_1$, hence the matrix $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{P})]$ is diagonally dominant, and thus it is an M-matrix (see Definition 3.3 in Chapter 3). We can then easily show that the numerical scheme is l_∞ stable by a straightforward maximum analysis as in [31].

- **Monotonicity:** To guarantee monotonicity, we use a wide stencil to discretize the second derivative terms in the discrete linear operator L_h^p (5.33) (see proof in Section 3.4). Note that using linear interpolation to compute $\mathcal{U}_{i^*,j}^n$ (5.38) in the semi-Lagrangian timestepping scheme also ensures monotonicity.

- **Consistency:** A simple Taylor series verifies consistency. As noted in Section 5.3.3, we may shrink the wide stencil length to avoid using points below the lower boundaries. We can use the same proof in Section 3.2 to show this treatment retains local consistency. Since we have either simple Dirichlet boundary conditions, or the PDE at the boundary is the limit from the interior, then we need only use the classical definition of consistency here. The only case where the point $\mathcal{U}_{i^*,j}^n$ (5.38) in the semi-Lagrangian timestepping scheme is outside computational domain is through the upper boundary $w = w_{\max}$, where the asymptotic solution (5.22) is used. Thus, unlike the semi-Lagrangian timestepping scheme in [18], we do not need the more general definition of consistency [7] to handle the boundary data.

5.2.5 Policy iteration

Our numerical scheme requires the solution of highly nonlinear algebraic equations (5.49) at each timestep. We use the policy iteration algorithm [39] to solve the associated algebraic system [39, 48]. Regarding the convergence of the policy iteration, since the matrix $[\mathbf{I} - \Delta\tau \mathbf{L}^{n+1}(\mathcal{P})]$ (5.49) is an M-matrix and the control set $\mathcal{Z}_h$ is a finite set, it is easy to show that policy iteration is guaranteed to converge [39].

Remark 5.7. *In this application of policy iteration, the control set is discrete, thus it is straightforward to determine the optimal control, and guarantee the convergence of the algorithm.*

5.3 Implementation details

5.3.1 Complexity

Examination of the algorithm for solving discrete equations (5.49) reveals that each timestep requires the following steps.

- In order to solve the local optimization problems at each node, we perform a linear search to find the minimum for $p \in \mathcal{Z}_h$. Thus, with total $O(1/h^2)$ nodes, this gives a complexity $O(1/h^3)$ for solving the local optimization problems at each time step.
- As noted in Section 4.0.1, we use a preconditioned **Bi-CGSTAB** iterative method for solving the sparse matrix at each policy iteration. The time complexity of solving the sparse M-matrix is $O((1/h^2)^{\frac{5}{4}})$ [64].

Assuming that the number of policy iterations is bounded, as the mesh size tends to zero, which is in fact observed in our experiments, the complexity of the time advance is thus dominated by the solution of the local optimization problems. Finally, the total complexity is $O(1/h^4)$.

5.3.2 The efficient frontier

In order to trace out the efficient frontier solution of problem (5.7), we proceed in the following way. Pick an arbitrary value of γ and solve problem (5.14), which determines the optimal control $p^*(\cdot)$. There are then two methods to determine the quantities of interest $(Var_{p^*}^{x_0,0}[W(T)], E_{p^*}^{x_0,0}[W(T)])$, namely the PDE method and the Hybrid (PDE - Monte Carlo) method. We will compare the performance of these methods in the numerical experiments.

The PDE Method

For a fixed γ, given $\mathcal{U}(w_0, v_0, 0)$ and $\mathcal{E}(w_0, v_0, 0)$ obtained by solving the corresponding equations (5.15) and (5.18) at the initial time with $W_0 = w_0$ and $V_0 = v_0$, we can then compute the corresponding pair $(Var_{p^*(\cdot)}^{x_0,0}[W(T)], E_{p^*(\cdot)}^{x_0,0}[W(T)])$, where $x_0 = (w_0, v_0)$. That is,

$$\begin{aligned} E_{p^*(\cdot)}^{x_0,0}[W(T)] &= \mathcal{E}(w_0, x_0, 0), \\ Var_{p^*(\cdot)}^{x_0,0}[W(T)] &= \mathcal{U}(w_0, v_0, 0) - \gamma \mathcal{E}(w_0, x_0, 0) - \frac{\gamma^2}{4} - \mathcal{E}(w_0, v_0, 0)^2, \end{aligned} \tag{5.51}$$

which gives us a single candidate point $\mathcal{Y}_{Q(\gamma)}$. Repeating this for many values of γ gives us a set of candidate points.

We are effectively using the parameter γ to trace out the efficient frontier. From Theorem 5.1, we have that $\gamma = \frac{1}{\lambda} + 2\mathcal{E}_0$. If $\lambda \to \infty$, the investor is infinitely risk averse, and invests only the risk free bond, hence in this case, the smallest possible value of γ is

$$\gamma_{\min} = 2w_0 \exp(rT). \tag{5.52}$$

In practice, the interesting part of the efficient frontier is in the range $\gamma \in [\gamma_{\min}, 10\gamma_{\min}]$. Finally, the efficient frontier is constructed from the upper left convex hull of $\mathcal{Y}_Q$ [69] to remove spurious points. In our case, however, it turns out that all the points are on the efficient frontier, and there are no spurious points, if $\gamma \geq \gamma_{\min}$.

The Hybrid (PDE - Monte Carlo) discretization

In the hybrid method, given the stored optimal control $p^*(\cdot)$ from solving the HJB PDE (5.15), $(Var^{x_0,0}_{p^*(\cdot)}[W(T)], Var^{x_0,0}_{p^*(\cdot)}[W(T)])$ are then estimated by Monte Carlo simulations. We use the Euler scheme to generate the Monte Carlo simulation paths of the wealth (5.4), and an implicit Milstein scheme to generate the Monte Carlo simulation paths of the variance process (5.3). Starting with $W_0 = w_0$ and $V_0 = v_0$, the Euler scheme for the wealth process (5.4) is

$$W_{t+\Delta t} = W_t \exp\left(\left(r + p^*\xi V_t - 0.5(p^*\sqrt{V_t})^2\right)\Delta t + p^*\sqrt{V_t \Delta t}\phi_1\right), \tag{5.53}$$

and the implicit Milstein scheme of the variance process (5.3) [51] is

$$V_{t+\Delta t} = \frac{V_t + \kappa\theta\Delta t + \sigma\sqrt{V_t\Delta t}\phi_2 + \sigma^2\Delta t(\phi_2^2 - 1)/4}{1 + \kappa\Delta t}, \tag{5.54}$$

where ϕ_1 and ϕ_2 are standard normal variables with correlation ρ. Note that this discretization scheme will result in strictly positive paths for the variance process if $4\kappa\theta > \sigma^2$ [51]. For the cases where this bound does not hold, it will be necessary to modify (5.54) to prevent problems with the computation of $\sqrt{V_t}$. For instance, whenever V_t drops below zero, we could use the Euler discretization

$$V_{t+\Delta t} = V_t + \kappa(\theta - V_t^+)\Delta t + \sigma\sqrt{V_t^+}\sqrt{\Delta t}\phi_2, \tag{5.55}$$

where $V_t^+ = \max(0, V_t)$. [54] reviews a number of similar remedies to get around the problem when V_t becomes negative and concludes that the simple fix (5.55) works best.

5.3.3 Outside the computational domain

As a wide stencil method is used in our numerical scheme, the stencil length needs to be increased to use the points beyond the nearest neighbors of the original grid. Therefore, when solving the PDE in a bounded region, the numerical discretization may require points outside the computational domain. When a candidate point we use is outside the computational region at the upper boundaries, we can directly use its asymptotic solution (5.22). For a point outside the upper boundary $w = w_{\max}$, the asymptotic solution is specified by the equation (5.22). For a point outside the upper boundary $v = v_{\max}$, by the implication of the boundary condition $\mathcal{U}_v = 0$ on $v = v_{\max}$, we have,

$$\mathcal{U}(w, v, \tau) = \mathcal{U}(w, v_{\max}, \tau), \quad v > v_{\max}. \tag{5.56}$$

However, we have to take special care when we may use a point below the lower boundaries $w = 0$ or $v = 0$, because the equation (5.15) is defined over $[0, \infty] \times [0, \infty]$. We can use Algorithm 2.1 to avoid this problem by shrinking the stencil to remain inside the domain (see details in Section 2.3.4). This simple treatment ensures that all data required is within the domain of the HJB equation.

In addition, due to the semi-Lagrangian timestepping (Section 5.2.2) , we may need to evaluate the value of an off-grid point $(w_{i^*} = w_i e^{(r - p\xi v_j)\Delta\tau^n}, v_j)$ (5.37). This point may be outside computational domain through the upper boundary $w = w_{\max}$ (the only possibility). When this situation occurs, the asymptotic solution (5.22) is used.

5.3.4 An improved linear interpolation scheme

When solving the value function problem (5.15) or the expected value problem (5.18) on a computational grid, it is required to evaluate $\mathcal{U}(\cdot)$ and $\mathcal{E}(\cdot)$, respectively, at points other than a node of the computational grid. This is especially important when using semi-Lagrangian timestepping. Hence, interpolation must be used. As mentioned earlier, to preserve the monotonicity of the numerical schemes, linear interpolation for an off-grid node is used in our implementation. [28] introduces a special linear interpolation scheme applied along the w-direction to significantly improve the accuracy of the interpolation in a 2-D impulse control problem. We modify this algorithm in our problem set-up.

We then take advantage of the results in Section 5.2.1 to improve the accuracy of the linear interpolation. Assume that we want to proceed from timestep τ^n to τ^{n+1}, and that we want to compute $\mathcal{U}(\bar{w}, v_j, \tau^n)$ where $\bar{w}$ is neither a grid point in the w-direction nor the special value $W_{\text{opt}}(T - \tau^n)$, where W_{opt} is defined in equation (5.23). Furthermore, assume that $w_k < \bar{w} < w_{k+1}$ for some grid points w_k and w_{k+1}. For presentation purposes, let $w_{special} = W_{\text{opt}}(T - \tau^n)$ and $\mathcal{U}_{special} = 0$. An improved linear interpolation scheme along the w-direction for computing $\mathcal{U}(\bar{w}, v_j, \tau^n)$ is shown in Algorithm 5.1. Note that the interpolation along v-direction is a plain linear interpolation, thus we only illustrate the interpolation algorithm in w-direction.

Following the same line of reasoning used for the function value problem, we have that

$$\mathcal{E}(v, W_{\text{opt}}(t), t) = \frac{\gamma}{2}.$$

By using this result, a similar method as Algorithm 5.1 can be used to improve the accuracy of linear interpolation when computing the expected value $\mathcal{E}(\bar{w}, v_j, \tau^n)$.

Algorithm 5.1 Improved linear interpolation scheme along the w-direction for the function value problem

```
if $w_{special} < w_k$ or $w_{special} > w_{k+1}$ then
    set $w_{left} = w_k$, $\mathcal{U}_{left} = \mathcal{U}^n_{k,j}$, $w_{right} = w_{k+1}$, and $\mathcal{U}^n_{right} = \mathcal{U}^n_{k+1,j}$
else
    if $w_{special} < \bar{w}$ then
        set $w_{left} = w_{special}$, $\mathcal{U}_{left} = \mathcal{U}_{special}$, $w_{right} = w_{k+1}$, and $\mathcal{U}^n_{right} = \mathcal{U}^n_{k+1,j}$
    else
        set $w_{left} = w_k$, $\mathcal{U}_{left} = \mathcal{U}^n_{k,j}$, $w_{right} = w_{special}$, and $\mathcal{U}^n_{right} = \mathcal{U}_{special}$
    end if
end if
Apply linear interpolation to $(w_{left}, \mathcal{U}_{left})$ and $(w_{right}, \mathcal{U}_{right})$ to compute $\mathcal{U}(\bar{w}, v_j, \tau^n)$
```

Remark 5.8. *For the discretization of the expected value problem (5.18), we still use the semi-Lagrangian timestepping to handle the drift term $(r + p^*\xi v)w\mathcal{E}_w$. Since it may be necessary to evaluate $\mathcal{E}^n_{i^*,j}$ at points other than a node of the computational grid, we need to use linear interpolation.*

5.4 Numerical experiments

In this section, we present numerical results of solution of equation (5.15) applied to the continuous time mean variance portfolio allocation problem. In our problem, the risky asset (5.2) follows the Heston model. The parameter values of the Heston model used in our numerical experiments are taken from [1] based on empirical calibration from S&P 500 index and VIX index dataset during 1990 to 2004 (under the real probability measure). Table 5.2 lists the Heston model parameters, and Table 5.3 lists the parameters of the mean variance portfolio allocation problem.

κ	θ	σ	ρ	ξ
5.07	0.0457	0.48	-0.767	1.605

Table 5.2: Parameter values in the Heston model

For all the experiments, unless otherwise noted, the details of the grid, the control set, and timestep refinement levels used are given in Table 5.4.

Investment Horizon T	10
The risk free rate r	0.03
Leverage constraint $p_{\max}$	2
Initial wealth w_0	100
Initial variance v_0	0.0457

Table 5.3: Input parameters for the mean variance portfolio allocation problem.

Refinement	Timesteps	W Nodes	V Nodes	$\mathcal{Z}_h$ Nodes
0	160	112	57	8
1	320	223	113	15
2	640	445	225	31
3	1280	889	449	63

Table 5.4: On each refinement, a new grid point is placed halfway between all old grid points and the number of timesteps is doubled. A constant timestep size is used. $w_{\max} = 6 \times 10^6$ and $v_{\max} = 3.0$. The number of finitely sampled γ points is 50. Note that increasing $w_{\max}$ by an order of magnitude and doubling $v_{\max}$ results in no change to the points on the efficient frontier to five digits. Increasing the number of γ points did not result in any appreciable change to efficient frontier (no spurious points in this case).

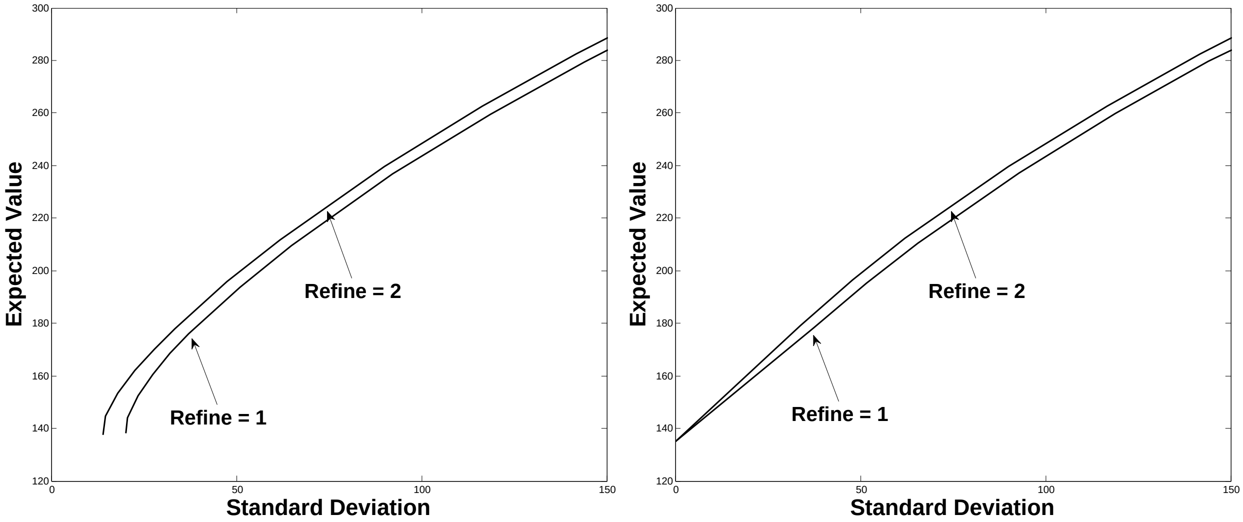
5.4.1 Effects of the improved interpolation scheme for the PDE method

In this subsection, we discuss the effects on numerical results of the linear interpolation scheme described in Section 5.3.4. We plot expected values against standard deviation, since both variables have the same units. Figure 5.1a illustrates the numerical efficient frontiers obtained using standard linear interpolation. It is clear that the results are very inaccurate for small standard deviations. It appears that the numerical methods were not able to construct the known point on the exact efficient frontier

$$(Var^{x,t}_{p^*(\cdot)}[W(T)], E^{x,t}_{p^*(\cdot)}[W(T)]) = (0, w_0 e^{rT}) \approx (0, 134.9859).$$

This trivial case corresponds to the case where $\gamma = \gamma_{\min}$ (5.52), and the investor invests only in the risk free bond and not in the risky asset. However, as shown in Figure 5.1a, in this special case, the standard deviation obtained by the numerical scheme using standard linear interpolation is far from the exact solution.

Figure 5.1b shows the numerical efficient frontiers obtained with the improved linear interpolation scheme, where Algorithm 5.1 is utilized. It is obvious that the numerical efficient frontiers obtained with the improved linear interpolation scheme are more reasonable, especially for the small standard deviation region. In particular, the special point where the variance is zero is now approximated accurately. This result illustrates the importance of using the optimal embedded terminal wealth $W_{\text{opt}}(t)$ and its function value for linear interpolation in constructing accurate numerical efficient frontiers. In all our numerical experiments in the following, the improved linear interpolation scheme is used.



(a) No Special Interpolation

(b) Special Interpolation

Figure 5.1: Close-up of efficient frontier for small standard deviations. (a) No Special Interpolation. (b) Special interpolation.

5.4.2 Convergence analysis

In this section, we illustrate the convergence of our numerical scheme, and compare the performance of two methods, namely the PDE method (Section 5.3.2) and the Hybrid method (5.3.2), for constructing the mean variance frontier under our model set-up.

Figure 5.2 shows that the mean standard deviation efficient frontiers computed by both the PDE method and the Hybrid method converge to the same frontier as the computational grid is refined. Our numerical results demonstrate that the Hybrid frontiers in general converge faster to the limit results than the pure PDE solutions. This same phenomenon was observed in [68]. As shown in Figure 5.2, the frontiers obtained by the Hybrid method are almost identical for refinement level 1 and 2. Note that for both methods, the optimal control is always computed by solving the HJB PDEs.

The same timesteps are used in both PDE method and Monte Carlo simulations, for each refinement level. For example, the frontiers labeled with "Refine = 1" for both methods in Figure 5.2 use the time steps as specified as Refinement level 1 in Table 5.4. To achieve small sampling error in Monte Carlo simulations, 10^6 simulations are performed for the numerical experiments. The standard error in Figure 5.2 can then be estimated. For example, consider a point on the frontier with the large standard deviation value which is about 350. For the expected value of $W(T)$, the sample error is approximately $350/\sqrt{10^6} \approx 0.35$, which could be negligible in Figure 5.2.

We will verify our conclusion by examining several specific points on these efficient frontiers in Figure 5.2. Table 5.5 and Table 5.6 show computed means and standard deviations for different refinement levels when $\gamma = 540$. The numerical results indicate first order convergence is achieved for both the PDE method and the Hybrid method. In this case, our numerical results demonstrate that the Hybrid frontiers converge faster to the limit results than the PDE solutions. Table 5.7 and Table 5.8 show computed means and standard deviations for different refinement levels when $\gamma = 1350$. The numerical results indicate first order convergence is achieved for the PDE method. In this case, our numerical results also demonstrate that the Hybrid frontiers converge faster to the limit results than the PDE solutions. However, the convergence ratio for the Hybrid method is erratic. As we noted before, in this case, the sample error for the estimate of the mean value is about $0.2 \simeq 200/\sqrt{10^6}$, which makes the convergence ratio estimates in Table 5.8 unreliable. The sample error may cause the phenomenon of the erratic convergence ratio in the Hybrid method results. To decrease the sample error to, for example, 0.02, the number of simulation paths would have to increase to 100×10^6, which is unaffordable in terms of the computational cost. Note that in the case $\gamma = 540$, with the small standard deviation, the sample error for the mean is about $0.05 \simeq 50/\sqrt{10^6}$.

Remark 5.9 (Efficiency of the Hybrid method.)**.** *We remind the reader that for both the Hybrid and PDE methods, the same (computed) control is used. The more rapid convergence of the Hybrid method is simply due to a more accurate estimate of the expected quantities (with a known control). This result is somewhat counter-intuitive, since it suggests that a low accuracy control can be used to generate high accuracy expected values. We also observe this from the fact that a fairly coarse discretization of the admissible set $\mathcal{Z}_h$ generates fairly accurate solutions.*

Refine	Mean	Change	Ratio	Standard Deviation	Change	Ratio
0	207.1434			71.3924		
1	210.4694	3.3260		65.5090	−5.88336	
2	212.1957	1.7263	1.92	62.0862	−3.42288	1.72
3	213.1481	0.95238	1.81	60.4738	−1.61237	2.12

Table 5.5: The convergence table for the PDE method. Small standard deviation case with $\gamma = 540$.

5.4.3 Sensitivity of Efficient Frontiers

In this subsection, we show some numerical sensitivity analysis for the major market parameters, namely the leverage constraints $p_{\max}$, the risk premium parameter ξ, the mean

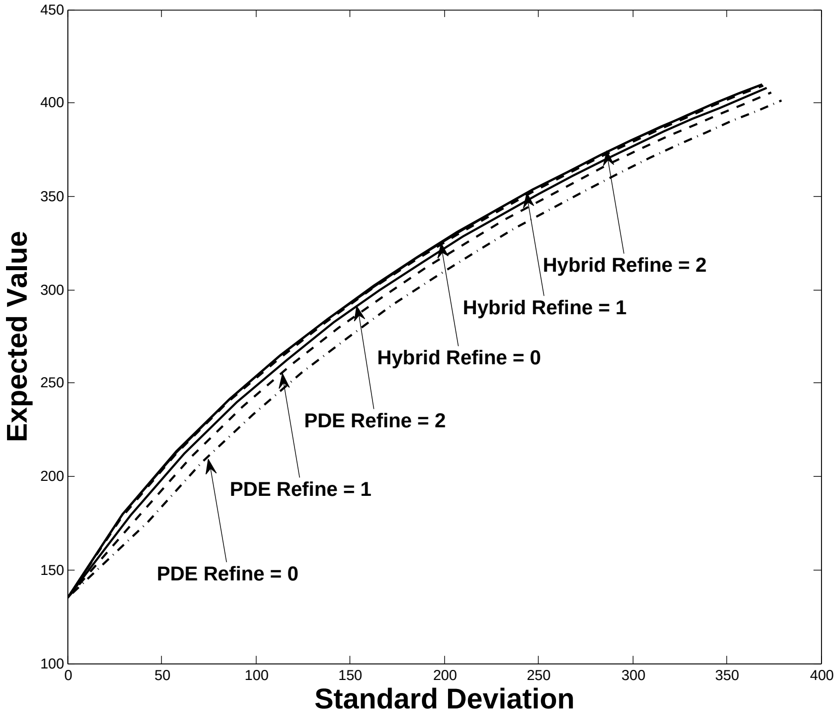


Figure 5.2: Convergence of frontiers in the PDE method and the Hybrid method. The frontiers labeled with “PDE” are obtained from the PDE method (Section 5.3.2). The frontiers labeled with “Hybrid” (Section 5.3.2) are obtained from a Monte Carlo simulation which uses the optimal controls determined by solving the HJB equation (5.15).

Refine	Mean	Change	Ratio	Standard Deviation	Change	Ratio
0	212.2993			56.6128		
1	213.2077	0.908		57.7652	1.152	
2	213.7573	0.550	1.65	58.2987	0.534	2.16
3	213.9903	0.233	2.36	58.5253	0.227	2.35

Table 5.6: The convergence table for the Hybrid method. Small standard deviation case with $\gamma = 540$.

Refine	Mean	Change	Ratio	Standard Deviation	Change	Ratio
0	320.5139			217.0009		
1	325.5443	5.030		212.1886	-4.812	
2	328.2670	2.723	1.85	209.8434	-2.345	2.05
3	329.8172	1.550	1.76	208.9045	-0.939	2.50

Table 5.7: The convergence table for the PDE method. Large standard deviation case with $\gamma = 1350$.

Refine	Mean	Change	Ratio	Standard Deviation	Change	Ratio
0	329.4411			206.0875		
1	330.5172	1.076		206.8351	0.748	
2	330.7066	0.189	5.68	207.1958	0.361	2.07
3	331.2820	0.575	0.33	207.3707	0.175	2.06

Table 5.8: The convergence table for the Hybrid method. Large standard deviation case with $\gamma = 1350$.

reversion level for the variance θ, the volatility of the variance σ, the correlation ρ between the risky asset and the variance, and the mean reversion speed κ. In our numerical tests, the corresponding frontiers are generated as the market parameter of interest changes, and the values of the remaining parameters are fixed and are listed in Table 5.2 and Table 5.3. We use the Hybrid method with the discretization level 2.

As observed in Figure 5.3, with $p_{\max} = \{1, 1.5, 2, +\infty\}$, larger values of the leverage constraints $p_{\max}$ result in much more dominant efficient frontiers. From Figure 5.4, with $\xi = \{0.5, 1.605, 2.5\}$, we can see that larger values of ξ result in much more dominant efficient frontiers. The maximal standard deviation point ($\gamma = +\infty$) on the efficient frontier with $\xi = 0.5$ is only about 191, which is much smaller than those with larger ξ values. From Figure 5.5, $\theta = \{0.01, 0.0457, 0.36\}$, we can see that larger values of the mean reversion level θ for the variance, result in much more dominant efficient frontiers. The maximal standard deviation point ($\gamma = +\infty$) on the efficient frontier with $\theta = 0.01$ is only about 108, which is much smaller than those with larger θ values. From Figure 5.6, $\sigma = \{0.2, 0.48, 0.7\}$, we can see that larger values of the volatility of the variance σ result in a slightly more dominant efficient frontiers in general. In particular, these efficient frontiers in large standard deviation region with different σ values values are almost identical.

On the other hand, from Figure 5.7, with $\rho = \{-0.767, -0.3, 0\}$, we can see that an increase in the correlation ρ produces frontiers with a slightly smaller expected value for a given standard deviation. These efficient frontiers in the large standard deviation region with different ρ values are almost identical. The effect of the κ values on the efficient frontier is more complex. From Figure 5.8, $\kappa = \{1, 5.07, 20\}$, in the small standard deviation region, an increase in κ produces frontiers with a smaller expected value for a given standard deviation. However, when the standard deviation increases to about 230, the larger values of κ gradually result in more significant dominant efficient frontiers.

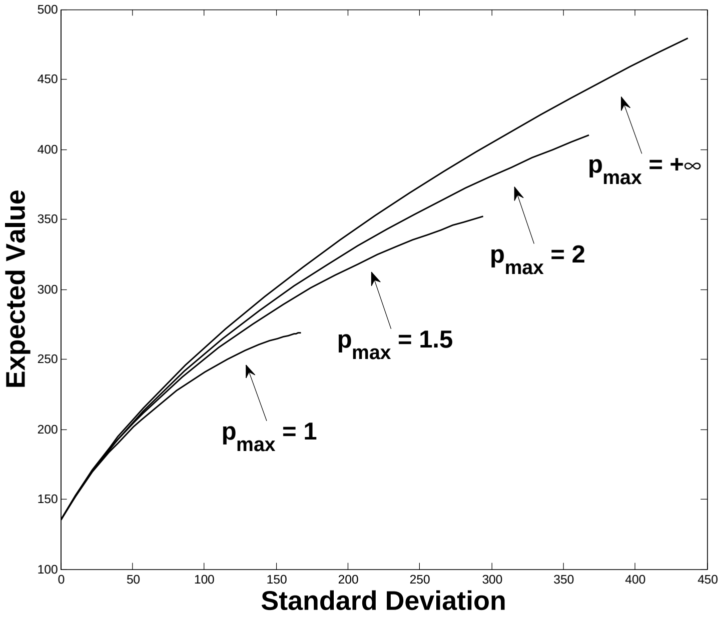


Figure 5.3: Sensitivity analysis of the efficient frontiers with respect to different leverage constraints $p_{\max}$. The Heston parameters and the remaining model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

5.4.4 Comparison between constant volatility and stochastic volatility cases

In this chapter, the risky asset follows the stochastic volatility model (5.2-5.3). In this Section, we will compare the constant volatility and stochastic volatility cases in terms of mean variance efficiency for the continuous time pre-commitment mean variance problem. With a constant volatility, the risky asset is the governed by the following geometric Brownian Motion (GBM) process:

$$\frac{dS}{S} = (r + \mu)dt + \sigma_S dZ_s. \tag{5.57}$$

To compare with the stochastic volatility case in Table 5.2, the constant volatility σ_S is set to $\sqrt{\theta} \approx 0.2138$, and the risky return over the risk free rate μ is set to $\xi\sigma_S^2 = 0.0733485$, which has the same mean premium of the volatility risk as the stochastic volatility model (5.2). This then corresponds to the case where the variance $V(t)$ in (5.2) is fixed to the mean reversion level θ. The remaining mean variance problem parameters are the same as listed in Table 5.3.

Figure 5.9 illustrates the fact that the efficient frontiers produced by using the stochastic volatility sightly dominates the curve produced by the constant volatility model. With the

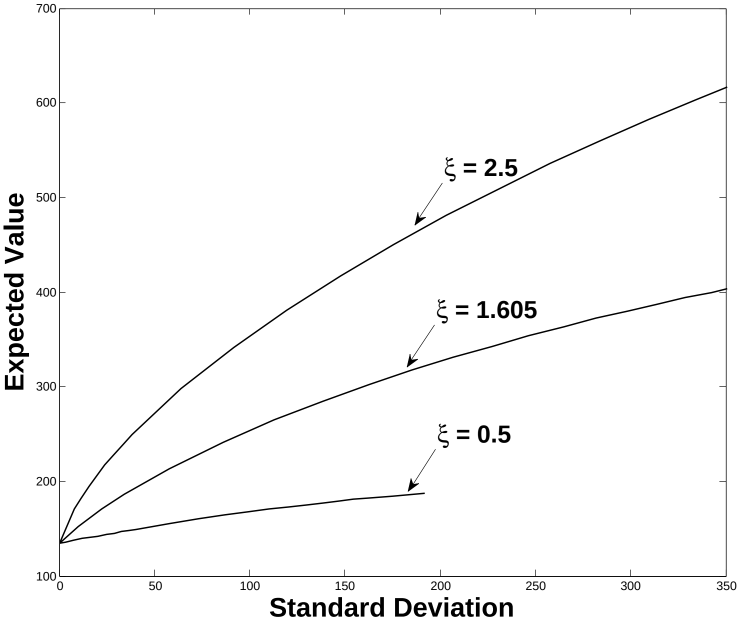


Figure 5.4: Sensitivity analysis of the efficient frontiers with respect to different risk premium factor ξ values. The remaining Heston parameters and the model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

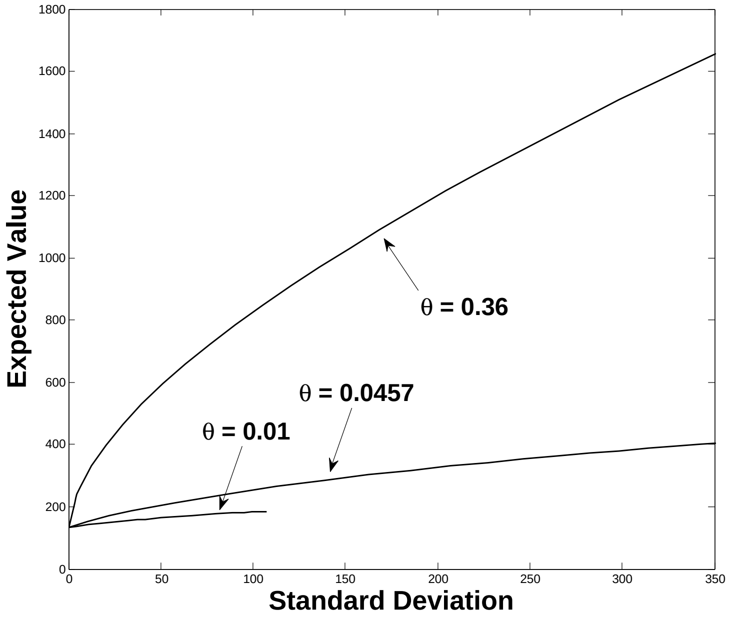


Figure 5.5: Sensitivity analysis of the efficient frontiers with respect to different mean reversion level θ values. The remaining Heston parameters and the model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

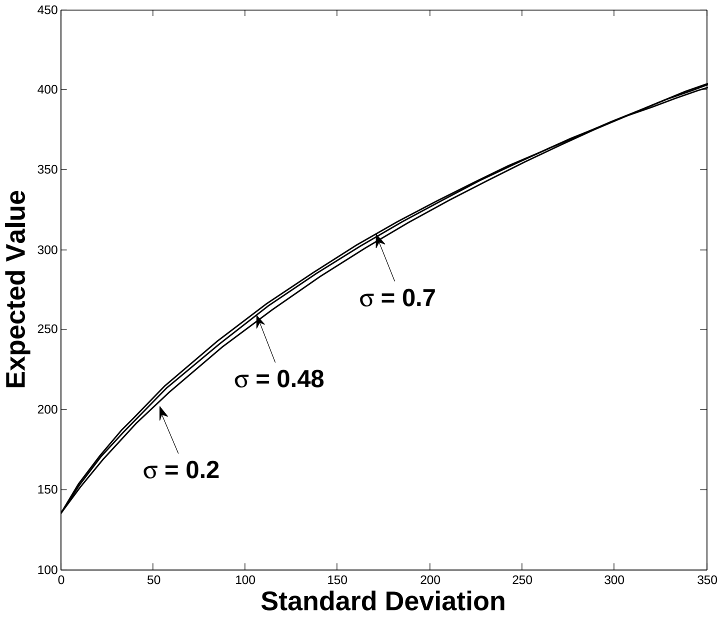


Figure 5.6: Sensitivity analysis of the efficient frontiers with respect to different volatility of the variance σ values. The remaining Heston parameters and the model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

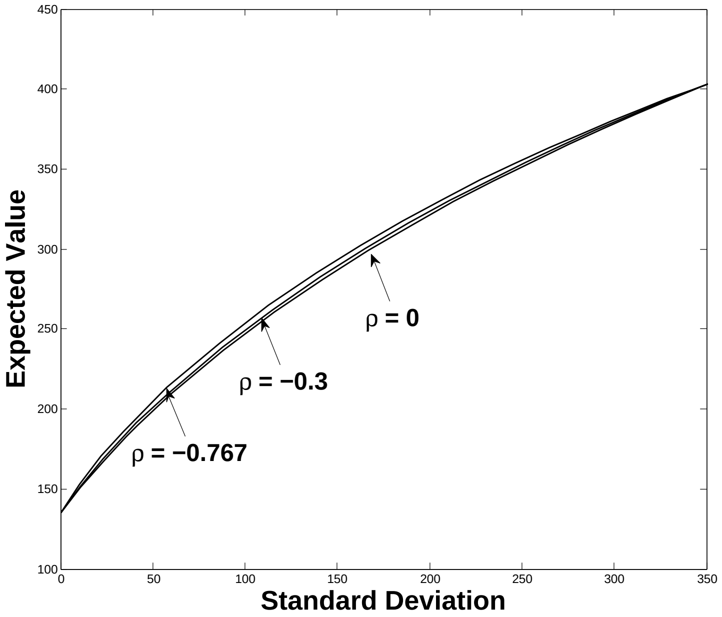


Figure 5.7: Sensitivity analysis of the efficient frontiers with respect to different correlation ρ values. The remaining Heston parameters and the model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

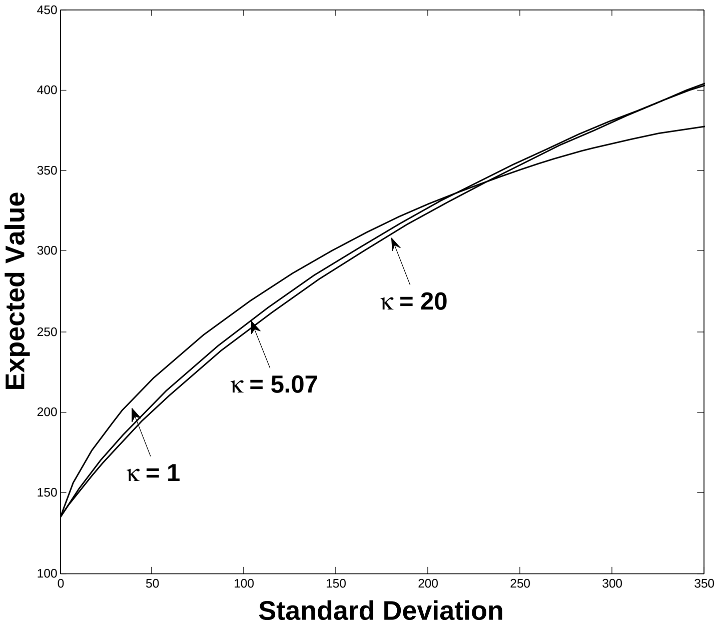


Figure 5.8: Sensitivity analysis of the efficient frontiers with respect to different mean variance speed κ. The Heston parameters and the remaining model parameters are given in Table 5.2 and Table 5.3. The Hybrid method with discretization level 2 is used.

Heston model's parameters in Table 5.2, we may conclude that the efficient frontier produced by the constant volatility is a good approximation of the frontier generated by the stochastic volatility model. From Figure 5.9, however, we see that if the mean reversion speed κ is set to a small value, e.g. one, in the stochastic volatility case, the efficient frontiers computed using a constant volatility model will be considerably different from those computed using the stochastic volatility model. The quantity $1/\kappa$ is measured in years and is related to the time over which a volatility shock dissipates. Specially, the half-life of a volatility shock is $\frac{\ln 2}{\kappa}$.

Finally, using the portfolio allocation strategy that is precomputed and stored from the constant volatility case, we then carry out a Monte Carlo simulation where the risky asset follows the stochastic volatility model. We then compare the results using this approximate control, with the optimal control computed using the full stochastic volatility model. From Table 5.9, we can see that the mean variance pairs computed using the optimal strategy are very close to the strategy computed using the GBM approximation. Based on several tests, a good heuristic guideline is that if $\kappa T > 40$, then the GBM control is a good approximation to the true (optimal control).

		$\gamma = 540$		$\gamma = 1350$	
Control Process	Price Process	Mean	Stnd Dev	Mean	Stnd Dev
GBM	GBM	209.50	59.68	330.09	213.01
GBM	Stoch. Vol.	212.68	58.42	329.13	207.23
Stoch. Vol.	Stoch. Vol	213.99	58.53	331.28	207.37

Table 5.9: Given a γ, the optimal portfolio allocation strategy is computed and stored assuming a control process, which is either GBM or stochastic volatility. The mean variance pairs are then estimated by Monte Carlo Simulation, using the stored controls, assuming that the actual price process follows either GBM or stochastic volatility. For the stochastic volatility case, the parameters are given in Table 5.2. For the GBM case, the variance is fixed to the mean value of the stochastic volatility case.

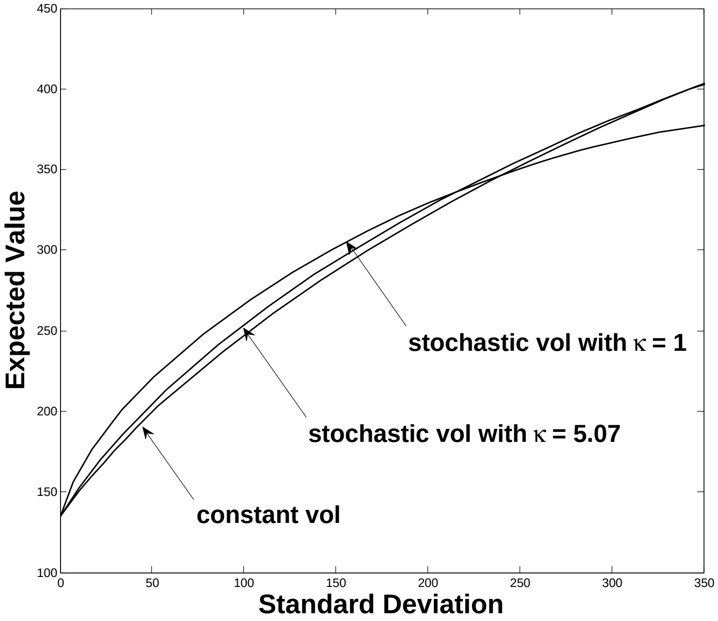


Figure 5.9: Efficient Frontier Comparison between constant volatility and stochastic volatility cases. For the stochastic volatility cases, $\kappa = 1,\ 5.07$, and the remaining stochastic volatility parameters are given in Table 5.2. The GBM parameters are given in Section 5.4.4.

5.5 Summary

The main results of this chapter are summarized as follows.

- We use the wide stencil method developed in Chapter 2 to overcome the main difficulty in designing a monotone approximation. We show that our numerical scheme is monotone, consistent, and ℓ_∞-stable. Hence, the numerical solution is guaranteed to converge to the unique viscosity solution of the corresponding HJB PDE, assuming that the HJB PDE satisfies a strong comparison property.
- Using semi-Lagrangian timestepping to handle the drift term and an improved method of linear interpolation, allows us to compute accurate efficient frontiers. Standard linear interpolation gives poor results for small values of standard deviation. When tracing out the efficient frontier solution of our problem, we demonstrate that the Hybrid (PDE - Monte Carlo) method [68] converges more rapidly than the pure PDE method. Similar results are observed in [68].
- We find that if the mean reversion time $\frac{1}{\kappa}$ is small compared to the investment horizon T, then a constant volatility GBM approximation to the stochastic volatility process gives a very good approximation to the optimal strategy. A reasonable rule of thumb is that if $\kappa T > 40$, then the GBM control is a good approximation to the optimal control.

Chapter 6

Option Pricing with Local Volatility Uncertainty

In this chapter, we value options with local volatility uncertainty. We formulate the uncertain local volatility model in Section 6.1. Section 6.2 reformulates the calibration problem into a Bayesian framework. Static hedging is used to decrease the spread between the minimal and maximal prices, and is introduced in Section 6.3. Finally, we carry out numerical experiments on pricing two-dimensional basket options under the uncertain local volatility model in Section 6.4.

6.1 The mathematical formulation

Consider a financial market with d reference assets, whose prices are denoted as $\mathbf{S} = (S_1, \dots, S_d)$. Suppose there is a family of local volatility models (pricing measures) $\mathbb{Q} \in \mathcal{Q}$. Under a risk neutral measure $\mathbb{Q}$, the dynamics of each asset price S_i is as follows

$$\frac{dS_i}{S_i} = (r - q_i)dt + \sigma_i dW_i, \quad i = 1, \dots, d, \tag{6.1}$$

where r is the risk-free interest rate, $q_i,\ i = 1, \dots, d$ are the dividend yields for S_i, σ_i are volatilities, and W_i are Wiener processes with $dW_i dW_j = \rho_{ij} dt$. Specifically, we assume that σ_i lies with a range, *e.g.*,

$$\sigma_i \in [\underline{\sigma}_i, \overline{\sigma}_i], \tag{6.2}$$

where σ_i is non-anticipative, and bounded in an interval. We emphasis that the bounds can be deterministic functions of price S_i and time t. That is, $\underline{\sigma}_i \equiv \underline{\sigma}_i(S_i, t)$ and $\overline{\sigma}_i \equiv \bar{\sigma}_i(S_i, t)$. In addition, uncertain correlation between the two underlying assets S_i and S_j is permitted, *e.g.*,

$$\rho_{ij} \in [\underline{\rho}_{ij}, \overline{\rho}_{ij}], \quad , i \neq j, \quad 1 \leq i, j \leq d. \tag{6.3}$$

Similarly, the correlation bounds could be deterministic functions of the underlying assets and time t, e.g., $\underline{\rho}_{ij} \equiv \underline{\rho}_{ij}(S_i, S_j, t)$. Let $\theta = (\sigma_1, \ldots, \sigma_d, \{\rho_{i,j}\}_{1\le i,j\le d})$, where σ_i represents the volatility function associated with the asset price S_i, and $\{\rho_{i,j}\}_{1\le i,j\le d}$ is the correlation matrix. Let $\mathcal{Q}^\theta$ denote the set of possible parameters defined in (6.2) (for volatilities) and (6.3) (for correlations). We call this model an uncertain local volatility model.

As in Chapter 2, when the volatilities σ_i and correlations ρ_{ij} are uncertain, the price of the option contract is no longer unique. Let $\mathcal{U}(\mathbf{S}, t)$ denote the value of a derivative at time t written on $\mathbf{S}_t$ with maturity T and final payoff $X(\mathbf{S})$, then at any time $0 \le t \le T$, we must have

$$\underline{\mathcal{U}}(\mathbf{S}, t) = \inf_{\theta\in\mathcal{Q}^\theta} E_t^\theta\left[e^{-rT}X(\mathbf{S})\right] \le \mathcal{U}(\mathbf{S}, t) \le \sup_{\theta\in\mathcal{Q}^\theta} E_t^\theta\left[e^{-rT}X(\mathbf{S})\right] = \bar{\mathcal{U}}(\mathbf{S}, t). \tag{6.4}$$

That is, the contingent claim value is bounded by $\underline{\mathcal{U}}(\mathbf{S}, t)$ and $\bar{\mathcal{U}}(\mathbf{S}, t)$. Then, we solve the option pricing problems in terms of backward time $\tau = T - t$, and denote $\bar{\mathcal{U}}(\mathbf{S}, \tau) = \bar{\mathcal{U}}\,(\mathbf{S}, T - t)$. By the principle of the dynamic programming, the maximal value $\bar{\mathcal{U}}(\mathbf{S}, \tau)$ is governed by the following HJB PDE,

$$\begin{aligned} \frac{\partial\bar{\mathcal{U}}}{\partial\tau} = \sum_{i=1}^d rS_i\frac{\partial\bar{\mathcal{U}}}{\partial S_i} + \sup_{\theta\in\mathcal{Q}^\theta}\left(\sum_{i,j=1}^d \frac{1}{2}\rho_{i,j}\sigma_i(S_i,\tau)\sigma_j(S_j,\tau)\frac{\partial^2\bar{\mathcal{U}}}{\partial S_i\partial S_j}\right) - r\bar{\mathcal{U}},\\ \overline{\mathcal{U}}(\mathbf{S}, 0) = X(\mathbf{S}), \end{aligned} \tag{6.5}$$

and similarly, the minimal value $\underline{\mathcal{U}}(\mathbf{S}, \tau)$ is governed by

$$\begin{aligned} \frac{\partial\underline{\mathcal{U}}}{\partial\tau} = \sum_{i=1}^d rS_i\frac{\partial\underline{\mathcal{U}}}{\partial S_i} + \inf_{\theta\in\mathcal{Q}^\theta}\left(\sum_{i,j=1}^d \frac{1}{2}\rho_{i,j}\sigma_i(S_i,\tau)\sigma_j(S_j,\tau)\frac{\partial^2\underline{\mathcal{U}}}{\partial S_i\partial S_j}\right) - r\underline{\mathcal{U}},\\ \underline{\mathcal{U}}(\mathbf{S}, 0) = X(\mathbf{S}). \end{aligned} \tag{6.6}$$

In general, (6.5) and (6.6) cannot be solved analytically, and we have to resort to the numerical scheme developed in Chapter 2.

Note that the maximal (minimal) prices could be connected to the superhedging strategy problem for a short (long) derivative position. Consider a one factor derivative with the payoff $X(S)$ written on the stock S as an example. Without any assumption on the volatility except that it lies in the $[\sigma_{\min}, \sigma_{max}]$ range, the maximal and minimal values are obtained by solving (6.5) and (6.6), respectively. Suppose we short the derivative with the maximal price $\bar{\mathcal{U}}(S, 0)$ and dynamically hedge the derivative with $\frac{\partial\bar{\mathcal{U}}(S,t)}{\partial S}$ shares of stock. This strategy is guaranteed to produce a non-negative value of the hedging portfolio whatever the realised path, provided that volatility lies in the specified range. This approach is termed a superhedging strategy approach to derivative pricing. In the worst-case scenario, the hedging portfolio will have value zero. Thus, the maximal price $\bar{\mathcal{U}}(S, 0)$ is the worst-case price for a short position. Note the price $\bar{\mathcal{U}}(S, 0)$ is cheapest superhedging price when a seller who shorts a delta-hedged derivative desires a non negative value. Similarly, the worst-case price for a long position is the minimal price $\underline{\mathcal{U}}(S, 0)$. From this perspective, the minimal and the maximal value can be interpreted as bid and ask prices for a long or short position.

6.2 The local volatility bounds

For ease of exposition, in this section, we restrict attention to a single asset local volatility. Suppose in the financial market, there are options written on S whose prices are observed. Denote the option payoffs as (C_i), $i \in I$ and their observed market prices by $(V_i^*)_{i\in I}$, which are used as model calibration inputs. In practice, instead of a unique price $(V_i^*)_{i\in I}$, we only have quotes for a bid price V_i^{bid} and ask price V_i^{ask}. We here use an interesting idea [42] which finds an entire distribution of model parameters (local volatility values on discrete nodes), instead of calibrating a unique *best-fit* local volatility model. This approach reformulates the calibration problem into a Bayesian framework to attain a posterior distribution for model parameters. The approach is outlined in the following section. We refer the reader to [42] for details. Using the posterior distribution of the model parameters, we can construct confidence intervals of the parameters, and then reasonable bounds of the local volatility function are obtained.

Firstly, the key characteristics expected of the local volatility surface that can be recast into a Bayesian prior are identified. There are three properties we could expect of $\sigma(S,t)$:

- Positive: $\sigma(S,t) > 0$ for all values of S and t.
- Smoothness: there should be no sharp spikes or troughs in the surface; this ensures pricing and hedging is stable.
- Asymptotics: For small values of t especially, $\sigma(S,t)$ should be close to today's at-the-money implied Black-Scholes volatility σ_{atm}.

The following approach is used to reformulate prior beliefs for σ (short for $\sigma(S,t)$) into a norm cost functional $\|\cdot\|$ so that parameters which better satisfy the prior beliefs have smaller norm. The natural Gaussian prior is then

$$p(\sigma) \propto \exp\left\{-\frac{1}{2}\beta_p \|\log(\sigma) - \log(\sigma_{atm})\|_\kappa^2\right\}, \tag{6.7}$$

where $\|\cdot\|_\kappa$ is a Sobolev norm given by

$$\|u(x,y)\|_\kappa = (1-\kappa)\|u(x,y)\|_2^2 + \kappa\||\nabla u(x,y)|\|_2^2, \tag{6.8}$$

where the gradient operator $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$. Working in the logarithmic space guarantees σ is positive. The first part of the norm is to ensure greater prior density is attached to σ that are closer to the at-the-money volatility, and the second part ensures that the volatility is smooth. β_p quantifies how strong our prior assumptions are: a higher value of β_p indicating greater confidence in our assumptions.

To make the method practically feasible, the local volatility surface $\sigma(S,t)$ is firstly restricted to a finite-dimensional space, and $\sigma(S,t)$ is represented by a grid of nodes whose positions are given by: $s_1 < \cdots < s_j < \ldots s_J$ in the spatial direction and $0 = t_1 < \cdots < t_L$ in the temporal direction. The discrete representation of $\sigma(S,t)$ is defined by the parameter vector

$$\theta = (\log \sigma_1, \log \sigma_2, \ldots, \log \sigma_M), \tag{6.9}$$

where $M = J \times L$, and as spline interpolant $\Theta(\cdot,\cdot)$ of θ. Then, $\sigma(S,t) = \exp(\Theta(S,t))$.

Let $V_i^* = \frac{1}{2}(V_i^{bid} + V_i^{ask})$ be the market observed price at time t of a European call and $V_i(\theta)$ be the corresponding model price for the same derivative when the model parameter is θ. Define δ_i as the basis point bid-ask spread $V_i^{ask} - V_i^{bid}$ for the i-th option. Initially, the basis point error for the i-th option is modeled as

$$\frac{10^4}{S_0}\left(V_i(\theta) - V_i^*\right) \sim N(0, \delta_i^2), \tag{6.10}$$

where S_0 is the spot stock price, and $N(\cdot,\cdot)$ is a normal density. Hence the tracking error is measured as basis points relative to the stock price. The $\frac{10^4}{S_0}$ in (6.10) is due to the fact that a basis point refers to a relative spread of $10^{-4}S_0$. This is interpreted as the observed price V_i^* is the model value $V_i(\theta)$ with a normally distributed noise. For computational purposes, the above definition will be slightly modified. The basis point square error function for a calibration problem is then defined as

$$G(\theta) = \sum_{i=1}^{I} w_i \left(\frac{10^4\left(V_i(\theta) - V_i^*\right)}{S_0}\right)^2, \tag{6.11}$$

where w_i are weights summing to one. The positive Bayesian posterior weight is only attached to parameters θ which on average reproduce prices to within the average basis point bid-ask spread. That is,

$$G(\theta) \leq \delta^2, \tag{6.12}$$

where $\delta^2 = \sum_{i \in I} w_i \delta_i^2$ is the pre-specified (determined by the bid-ask spread) average basis point square error tolerance. Then, the likelihood function $P(V|\theta)$, the probability of observing the data V given θ, is as follows

$$p(V|\theta) \propto 1_{G(\theta) \leq \delta^2} \exp\left\{-\frac{1}{2\delta^2} G(\theta)\right\}. \tag{6.13}$$

As a result, those parameters θ which reproduce prices closest to the market observed prices V_i^* give the greatest likelihood values.

Combining the prior (6.7) and likelihood functions (6.13), using the Bayes rule, we get the explicit form for the posterior function

$$p(\theta|V) \propto 1_{G(\theta) \leq \delta^2} \exp\left\{-\frac{1}{2\delta^2}\left(G(\theta) + \beta\|\theta - \theta_0\|_\kappa^2\right)\right\}, \tag{6.14}$$

where $\beta = \beta_p \delta^2$. Thus, any parameter θ with positive posterior density gives model prices for the calibration options within the bid-ask spread. This is how the Bayesian approach reformulates the traditional local volatility calibration methods into a unified and rigorous framework. Note that maximising the posterior (6.14) is equivalent to minimising the expression

$$-\frac{1}{2\delta^2}\left(G(\theta)+\beta\|\theta-\theta_0\|_\kappa^2\right), \tag{6.15}$$

which is precisely the form of functional [49, 25] use as an objective function to find their optimal calibration parameter.

6.3 Lagrangian uncertain volatility model

Recall that the minimal and maximal prices correspond to the bid-ask spread if option buyers and sellers price options using worst-case scenarios, from their own points of view. However, the spread is sometimes so large as to make an uncertain volatility method difficult to apply in practice, even if we use the method presented in 6.2 to construct reasonable volatility bounds. To address this issue, the Lagrangian uncertain volatility model was introduced in [5]. In this model, optimal static hedges are computed for the given problem using predetermined options with known market prices. The residual payoff is then priced using an uncertain volatility model.

Firstly, we briefly review the Lagrangian uncertain model. Consider an agent who must deliver the payoff of an option (a short position pricing problem), and wishes to statically hedge his exposure using a set of liquid options. Let $X(\mathbf{S})$ be the payoff of the target option. Assume that N options will be used to form a static hedge, and the market prices V_i^* exist for all N of these options. For simplicity, we will ignore bid-ask spreads in V_i^*, and assume that all options under consideration expire at the same time. If the static hedge is formed by holding λ_i of each underlying option, then the residual payoff is given by

$$W(\mathbf{S},T)=X(\mathbf{S})-\sum_{i=1}^{N}\lambda_i C_i(\mathbf{S}), \tag{6.16}$$

where $C_i(\mathbf{S})$ is the payoff of the ith option.

The maximal value of the static-hedged derivative is the cost of options $\sum_{i=1}^{N}\lambda_i V_i^*$ and the maximal value (i.e. cost of hedging for a short position) for the residual payoff (6.16), where $\lambda=(\lambda_1,\ldots,\lambda_N)$. Then, the optimal static hedge for maximal pricing is determined from the solution of

$$\inf_{\lambda}\sup_{\theta\in\mathcal{Q}^\theta}\left(E^\theta\left[e^{-rT}W(\mathbf{S},T)+\sum_{i=1}^{N}\lambda_i V_i^*\right]\right). \tag{6.17}$$

We can interpret equation (6.17) as: find the static hedge which gives the cheapest cost of hedging which ensures that the writer will have a non-negative balance in the hedging portfolio. Let

$$\bar{\mathcal{V}}(\mathbf{S},0;\lambda) = \sup_{\theta\in\mathcal{Q}^\theta}\left(E^\theta\left[e^{-rT}W(\mathbf{S},T)+\sum_{i=1}^{N}\lambda_i V_i^*\right]\right), \tag{6.18}$$

this problem (6.17) can be then written as

$$\inf_\lambda \bar{\mathcal{V}}(\mathbf{S},0;\lambda). \tag{6.19}$$

Due to the form of this optimization, the problem (6.17) is known as the Lagrangian uncertain volatility model. Suppose that

$$\hat{\lambda} = \arg\inf_\lambda \bar{\mathcal{V}}(\mathbf{S},0;\lambda),$$

and the value $\bar{\mathcal{V}}(\mathbf{S},0;\hat{\lambda})$ is then the final price for the target derivative. When the $\hat{\lambda}_i$ are all equal to zero, $\bar{\mathcal{V}}(\mathbf{S},0;\hat{\lambda})$ degenerates to the maximal price $\bar{\mathcal{U}}(\mathbf{S},0)$ under a regular uncertain volatility model. In general, however, the minimization in (6.17) is attained for $\hat{\lambda}_i \neq 0$. The residual will not be a convex function, and the full nonlinear uncertain volatility model must be used for pricing. Note that even if the payoff $X(\mathbf{S})$ is convex, the addition of different options for static hedging makes the residual payoff non-convex. Thus, the optimal volatilities lie inside their ranges, and the hedging portfolio is cheaper for a short position and more expensive for a long position than at the extreme volatility values $\bar{\sigma}_i(S_i,t)$ or $\underline{\sigma}_i(S_i,t)$. Accordingly, the total cost $\bar{\mathcal{V}}(\mathbf{S},0;\hat{\lambda})$ of the hedging portfolio obtained from (6.17) will be less than $\bar{\mathcal{U}}(\mathbf{S},0)$. In fact, in terms of the financial pricing framework, $\bar{\mathcal{V}}(\mathbf{S},0;\hat{\lambda})$ is the maximal arbitrage-free value for the derivative, conditional on the market prices V_i^* of the hedging instruments. More precisely, the problem (6.17) is equivalent to the following problem

$$\begin{aligned}&\sup_{\theta\in\mathcal{Q}^\theta} E^\theta\left[e^{-rT}X(\mathbf{S})\right],\\ &\text{subject to}\quad E^\theta\left[e^{-rT}C_i\right] = V_i^*, \quad V_i^* = 1,\ldots,N.\end{aligned} \tag{6.20}$$

Note the validation of this statement requires the condition that there exists at least one group of arbitrage-free parameters of $E^\theta\left[e^{-rT}C_i\right] = V_i^*, \quad i = 1,\ldots,N$ [5].

We will illustrate the superhedging strategy for a short position combined with static hedging. Suppose we are selling a one factor derivative with the exotic payoff $X(S)$ written on the stock S. Assume that we use a single European call option, whose market price is V^*, as the only instrument in the static hedge. By solving the corresponding problem (6.17), we determine the optimal number of contracts $\hat{\lambda}$ and the final value $\bar{\mathcal{V}}(S,0;\hat{\lambda})$ for the target derivative.

Initially, we will buy $\hat{\lambda}$ contracts of the European call option, and sell the derivative for the price $\bar{\mathcal{V}}(S,0;\hat{\lambda})$. We then perform dynamic hedging for the portfolio with $\frac{\partial\bar{\mathcal{V}}(S,t;\hat{\lambda})}{\partial S}$ shares of stock. This strategy is guaranteed to yield a non-negative value of the hedging portfolio, regardless of the realised volatility path, provided that volatility lies in the specified range. In the worst-case scenario, the hedging portfolio will have value zero at expiry. Thus, the price $\bar{\mathcal{V}}(S,0;\hat{\lambda})$ is the worst-case price for a short position with the optimal static hedge.

6.3.1 Solution of the optimization problem (6.17)

The algorithm for solving the optimization problem (6.17) is described here. Since the objective function $\bar{\mathcal{V}}(\mathbf{S}, 0; \lambda)$ (6.18) is a supremum (with respect to θ) of linear functions in λ_i, the problem is convex with respect to λ_i. Futhermore, the objective function $\bar{\mathcal{V}}(\mathbf{S}, 0; \lambda)$ is a differentiable function of λ_i as shown in [5]. Thus, we can use a standard optimization method, e.g., BFGS method for minimizing the problem [38]. The BFGS method requires the computation of the objective function as well as the first order derivatives, but the computation of a Hessian matrix is not required. In each optimization iteration, given a vector of λ, the objective function needs to be evaluated. The value $\bar{\mathcal{V}}(\mathbf{S}, 0; \lambda)$ is obtained by solving the corresponding HJB equation. The first order derivative with respect to λ_k is computed by the numerical approximation

$$\frac{\bar{\mathcal{V}}(\mathbf{S}, 0; \lambda + \epsilon e_k) - \bar{\mathcal{V}}(\mathbf{S}, 0; \lambda)}{\epsilon}, \tag{6.21}$$

where e_k is the unit basis vector. $\bar{\mathcal{V}}(\mathbf{S}, 0; \lambda + \epsilon e_k)$ is obtained by solving the corresponding HJB equation. To solve the optimization problems, we need to numerically solve the corresponding HJB equations $N+1$ times (one for the objective function value and N times for computing the first order derivatives with respect to λ_k, $k = 1, \ldots, N$) in each optimization iteration.

Similarly, the optimal static hedge for the minimal pricing is obtained by solving

$$\sup_{\lambda} \inf_{\theta \in \mathcal{Q}^\theta} \left(E^\theta \left[e^{-rT} W(\mathbf{S}, T) + \sum_{i=1}^{N} \lambda_i V_i^* \right] \right), \tag{6.22}$$

which finds the most expensive minimal value of the static-hedged derivative. In general, the final value obtained from (6.22) will be greater than the minimal price $\underline{\mathcal{U}}(\mathbf{S}, 0)$ computed using an uncertain volatility model.

6.4 Numerical results

We will carry out numerical experiments on pricing two-dimensional basket options as a concrete example under an uncertain local volatility model in this section. Firstly, we construct the reasonable volatility bounds using market data. We then investigate basket option prices with volatility uncertainty. Lastly, we use the static hedging method to decrease the spread between the maximal and minimal prices.

6.4.1 Local volatility bound construction

For pricing two-factor basket options, we here calibrate two market datasets. Firstly, we take real S&P implied volatility data with 10 strikes and 7 maturities used in [22] to determine the market prices of $I = 70$ corresponding European call options. The market implied volatilities in October, 1995 are given in Table E.1. Let the pre-specified basis point square error tolerance δ in (6.12) be 10 basis points, where δ is chosen based on the averaged bid-ask spreads. Given the price data E.1, we obtain the posterior distribution $P(\theta|V)$ (6.14). It is not feasible to analytically find $P(\theta|V)$. The best we can do is try to draw samples from this distribution using the Markov Chain Monte-Carlo (MCMC) Metropolis algorithm. As shown in Figure 6.1, 249 groups of the model parameters (6.9) with 50000 chains generated from the posterior distribution were tracked. For comparison, the *best-fit* local volatility surface is calibrated by minimizing equation (6.15). The trust-region method [21] is used to solve this non-convex optimization problem. The *best-fit* volatility surface is shown in Figure 6.3. The calibration error G (6.11) between market mid-price quotes and *best-fit* local volatility model prices is about 2.3 basis points.

Using this posterior distribution of the model parameters (6.9) in Figure 6.1, we can construct confidence intervals of the parameters. This facilitates construction of a reasonable pair of the lower and upper bounds for the local volatility values. Figure 6.2 shows the 95% pointwise confidence intervals of the 249 accepted samples. As shown in Figure 6.2, close to the spot price S_0, the bounds are very tight, and the further away from at-the-money, the larger these bounds become. Especially in the wings and for short times, the spread of volatilities are significant. These pointwise confidence intervals consequently serve as the lower and upper volatility bounds of the local volatility surface.

Similarly, we then calibrate European call options written on the FTSE-100 given in Table E.2 [49, 42]. Let the pre-specified basis point square error tolerance δ in (6.12) be 5 basis points, where δ is chosen based on the averaged bid-ask spreads. Given the price data in E.2, we obtain the posterior distribution $P(\theta|V)$ (6.14). Then, we draw samples from this distribution using the Markov Chain Monte-Carlo (MCMC) Metropolis algorithm. 172 groups of the parameters with 50000 chains generated from the posterior distribution were tracked. For comparison, the *best-fit* local volatility surface is also calibrated by minimizing equation (6.15). The calibration error G (6.11) between market quotes and *best-fit* local volatility model prices is about 0.9 basis points. We then construct the volatility bounds using the 95% pointwise confidence intervals of the tracked samples. Then, the upper and lower volatility bounds are illustrated in Figure 6.4. The *best-fit* volatility surface in this case is shown in Figure 6.5.

6.4.2 Basket option pricing with uncertain local volatility

Firstly, we carry out the numerical experiments for the two dimensional basket European call options. The payoff of the basket call option is as follows

$$\max(\omega_1 S_1 + \omega_2 S_2 - K, 0), \tag{6.23}$$

where K is the strike price, and ω_i is the weight for each individual asset price. Both of the individual assets are assumed to follow the uncertain local volatility model. The volatility

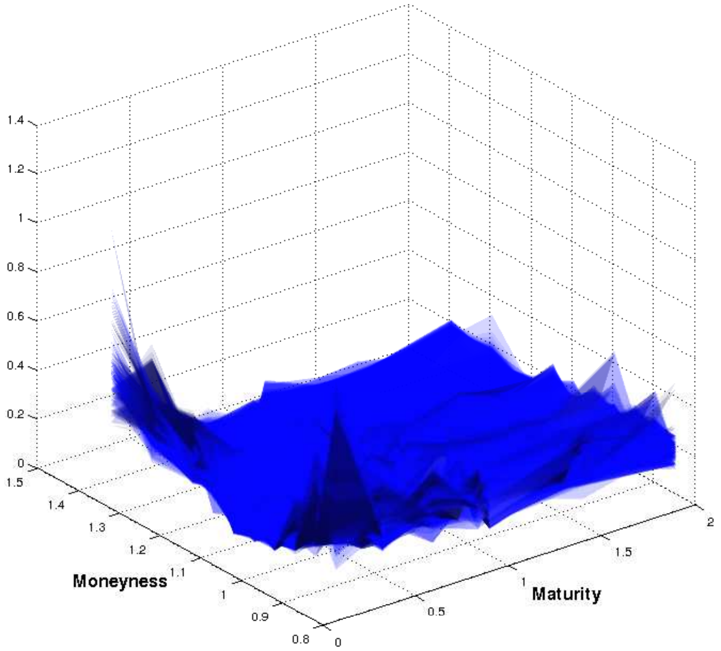


Figure 6.1: For S&P500 dataset in Table E.1: using MCMC Metropolis Sampling, 249 surfaces from the posterior distribution were tracked.

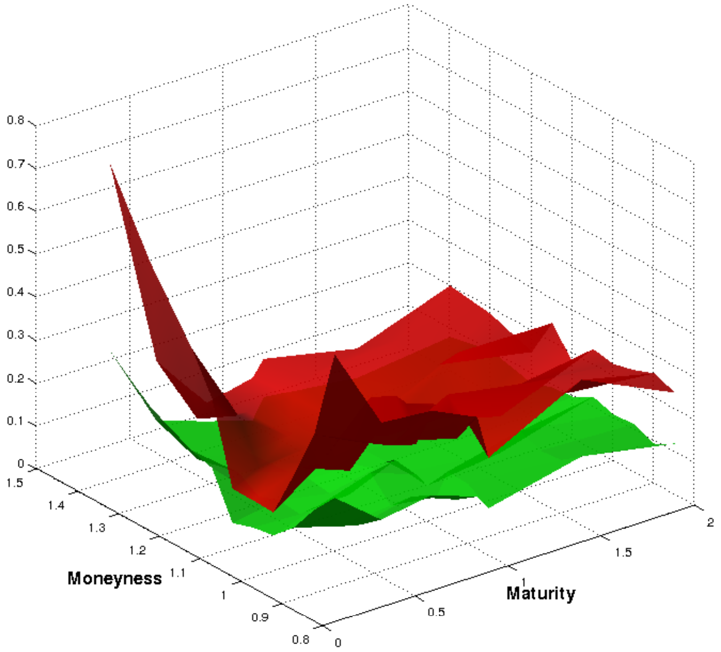


Figure 6.2: Using the 249 accepted local volatility surfaces in Figure 6.1, 95% confidence intervals are found pointwise. These confidence intervals are then used for the upper and lower volatility bounds.

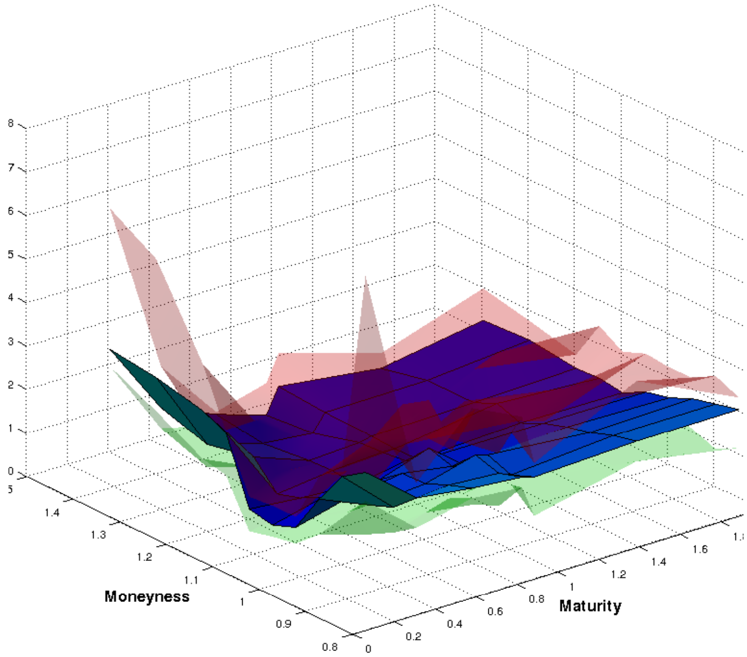


Figure 6.3: For the S&P500 dataset in Table E.1: the *best-fit* local volatility surface is calibrated by minimizing the error function (6.15). The middle(blue) surface is the *best-fit* local volatility surface. The top(red) and bottom(green) surfaces are upper and lower volatility bounds in Figure 6.2.

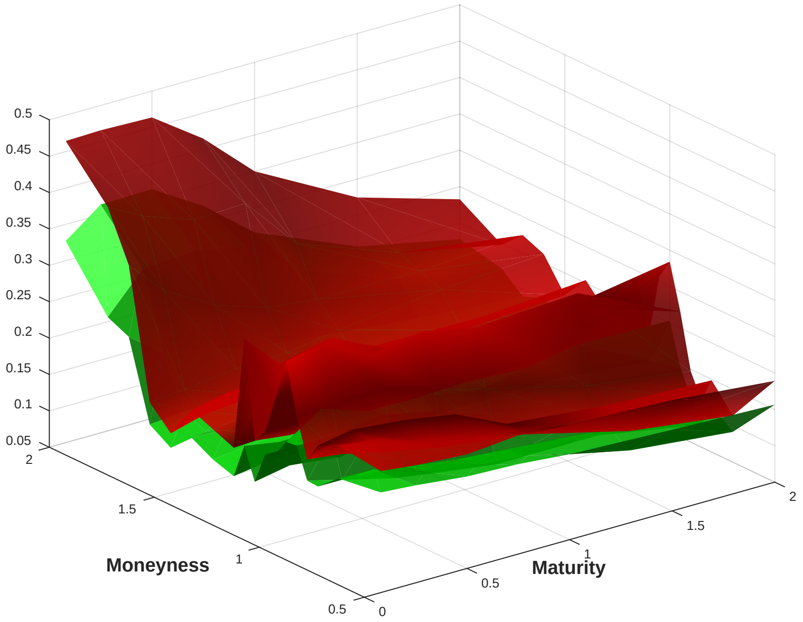


Figure 6.4: For S&P500 dataset in Table E.1: Using the 172 accepted local volatility surfaces, 95% confidence intervals are found pointwise. The top(red) and bottom(green) surfaces are upper and lower volatility bounds.

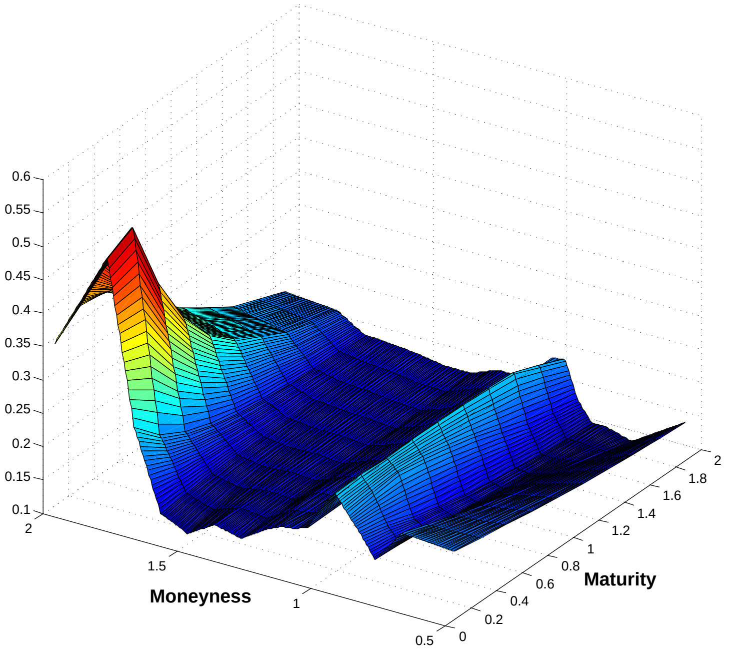


Figure 6.5: For the FTSE-100 dataset in Table E.2: the *best-fit* local volatility surface is obtained by minimizing the error function (6.15).

bounds constructed from the two market datasets in the last section will be used in the numerical experiments. The volatility bounds for the first individual asset S_1 are given in Figure 6.2, and the volatility bounds for the second individual asset S_2 are given in Figure 6.4. In general, the correlation between two assets is allowed to be uncertain, and lies in a range $[\rho_{\min}, \rho_{\max}]$. However, for these experiments, we force the correlation to be constant to investigate only the effect of the volatility uncertainty in the basket option pricing. The model parameters are given in Table 6.1. The maximal and minimal prices for the basket call options are given in Table 6.2. The maximal and minimal solutions are obtained by solving the numerical HJB equations (6.5) and (6.6).

For comparison, we then use the *best-fit* local volatility surface for S_1 shown in Figure 6.3 and the *best-fit* local volatility surface for S_2 shown in Figure 6.5. Using the *best-fit* volatility surface in the option pricing is the standard approach in the financial industry. The basket call option prices produced with these fixed local volatility surfaces are given in Table 6.3. The numbers in the parentheses in Table 6.2 are the relative distance between the maximal (minimal) prices and the values obtained by the *best-fit* volatility surfaces. Especially for out-of-the-money basket options, these relative distances are quite significant, as well as the spread between maximal and minimal prices. Nevertheless, a large spread signals the degree of risk in pricing and hedging these contracts due to the uncertainty in the calibrated local volatility surfaces [23].

Suppose that the volatility of the asset price S_1 and S_2 is assumed to lie within the upper and lower bound shown in Figure 6.2, regardless of the actual volatility movement. For a short position, the seller can dynamically hedge his position and obtain a non-negative balance in the hedging portfolio if the maximal values prices are offered. Thus, the prices by the uncertain local volatility model are obviously more robust. However, the *best-fit* prices are exposed to model risk when the model used in the derivative pricing does not coincide with the real asset price stochastic process.

Basket option type	European Call
Initial price S_1	100
Initial price S_2	100
Basket weight ω_1 and ω_2	0.5; 0.5
Interest rate r	0.05
ρ	0.5

Table 6.1: The model Parameters.

		$K = 80$	$K = 90$	$K = 100$	$K = 110$	$K = 120$
$T = 0.25$	Min	20.9945(0.05%)	11.1892(1.33%)	3.0158(13.78%)	0.1221(59.58%)	0.0003(96.74%)
	Max	21.0895(0.41%)	11.6563(2.79%)	4.1032(17.30%)	0.6099(101.89%)	0.0625(548.16%)
$T = 0.50$	Min	21.9836(0.23%)	12.4491(2.55%)	4.6792(12.40%)	0.6919(42.33%)	0.0275(78.78%)
	Max	22.2956(1.19%)	13.4474(5.26%)	6.3248(18.41%)	2.0063(67.24%)	0.4203(224.56%)
$T = 1.0$	Min	23.9463(0.63%)	14.8614(3.59%)	7.3632(10.99%)	2.4568(27.41%)	0.4480(52.44%)
	Max	24.6770(2.40%)	16.4822(6.92%)	9.6352(16.47%)	4.7246(39.60%)	1.8407(95.40%)
$T = 1.5$	Min	25.8683(0.90%)	17.0905(3.79%)	9.6741(3.90%)	4.3885(20.96%)	1.4450(37.65%)
	Max	26.8149(2.73%)	19.9978(12.58%)	12.3595(22.78%)	7.2699(30.94%)	3.8052(64.18%)
$T = 2.0$	Min	27.6968(1.16%)	19.1609(3.91%)	11.7920(9.31%)	6.2827(17.66%)	2.7431(29.93%)
	Max	28.8363(2.90%)	21.2918(6.77%)	14.4016(10.76%)	9.6349(26.28%)	5.8165(48.58%)

Table 6.2: The maximal and minimal prices of the basket call options with respect to strike price K and maturity T. The model parameters are given in Table 6.1. The volatility bounds for S_1 are illustrated in Figure 6.2, and the volatility bounds for S_2 are given in Figure 6.4. The number in the parentheses is the relative distance between the maximal (minimal) prices and the values obtained by the *best-fit* volatility surfaces.

6.4.3 Static hedging

In Table 6.2, the spread between the maximal and minimal prices sometimes seem too large to be used as the bid-ask offers. To decrease the spread, we can use a more conservative tolerable level of basis point error δ (6.12) in the posterior distribution function (6.14). Furthermore, when constructing the lower and upper volatility bounds, instead of using

	$K = 80$	$K = 90$	$K = 100$	$K = 110$	$K = 120$
$T = 0.25$	21.0044	11.3403	3.4979	0.3021	0.009636
$T = 0.5$	22.0343	12.7749	5.3414	1.1997	0.1295
$T = 1.0$	24.0990	15.4152	8.2727	3.3844	0.9420
$T = 1.5$	26.1024	17.7632	10.0666	5.5520	2.3177
$T = 2.0$	28.0227	19.9409	13.0029	7.6299	3.9148

Table 6.3: The prices of the European basket call options using the *best-fit* local volatility surfaces with respect to strike price K and maturity T. The *best-fit* local volatility surface for S_1 is given in Figure 6.3 and the *best-fit* local volatility surface for S_2 is given in Figure 6.5. The model parameters are given in Table 6.1.

95% pointwise confidence intervals, a smaller confidence intervals level, e.g. 68%, could be applied. All these approaches could make the volatility bounds more tight. Consequently, the spread between the maximal and minimal prices will be decreased. Here we shall use the optimal static hedging outlined in Section 6.3 to reduce this spread.

Firstly, we will show an interesting example. Consider three contracts (A, B, C) with terminal cash flows at the maturity

$$\begin{aligned} \text{Cashflow}_A &= \max(\omega_1 S_1 + \omega_2 S_2 - 90, 0), \\ \text{Cashflow}_B &= -\max(\omega_1 S_1 + \omega_2 S_2 - 100, 0), \\ \text{Cashflow}_C &= -\max(\omega_1 S_1 + \omega_2 S_2 - 110, 0), \end{aligned} \tag{6.24}$$

and all having the same maturity 0.5 year. The other parameters are given in Table 6.1. Consider another contract D, with the terminal cash flow

$$\text{Cashflow}_D = \text{Cashflow}_A + \text{Cashflow}_B + \text{Cashflow}_C.$$

If we price these contracts using an uncertain local volatility model, then we obtain the contract values

$$V_A = 13.4474,\ V_B = -4.6792,\ V_C = -0.6919,\ V_D = 7.3480. \tag{6.25}$$

The volatility bounds used for S_1 and S_2 are given in Figures 6.2 and 6.4, respectively. Note that,

$$V_D < V_A + V_B + V_C,$$

since

$$\begin{aligned} V_D &= \sup_{\theta \in \mathcal{Q}^\theta} E^\theta \left(e^{-rT} (\text{payoff}_A + \text{payoff}_B + \text{payoff}_C) \right) \\ &\le \sup_{\theta \in \mathcal{Q}^\theta} E^\theta (e^{-rT} \text{payoff}_A) + \sup_{\theta \in \mathcal{Q}^\theta} E^\theta (e^{-rT} \text{payoff}_B) + \sup_{\theta \in \mathcal{Q}^\theta} E^\theta (e^{-rT} \text{payoff}_C) \\ &= V_A + V_B + V_C. \end{aligned} \tag{6.26}$$

Method	Maximal	Minimal	Spread
No static hedge	6.3248	4.6792	1.6456
Optimal static hedge	5.6042	4.9296	0.6746

Table 6.4: Spread values for uncertain volatility pricing for a basket option without static hedging, and optimal static hedging with one asset call options. The strike price of the basket option is 100, and the maturity is 0.5. The model parameters are given in Table 6.1.

Problem	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Maximal pricing	0.46	0.49
Minimal pricing	0.28	0.37

Table 6.5: The optimal hedging positions.

Thus, under an uncertain local volatility model, the additive property of the contract prices is not preserved. Note that the value of derivatives is still homogeneous of degree one in the quantities: the value of n contracts is n times greater than one unit of the same contract.

We then illustrate the method of static hedging to decrease the bid-ask spread by a numerical experiment. Suppose that we price a European basket call option with strike $K = 100$ and maturity $T = 0.5$. The other model parameters are listed in Table 6.1. The static hedge portfolio will be composed of two European call options C_1 and C_2, each written on one single asset S_1 and S_2, with strikes of 100 and 100, respectively. The market value of C_1 for the one-factor call option written on S_1 is taken to 4.9725. This is simply the Black-Scholes price with the implied volatility provided in Table E.1. The market value of C_2 for the one-factor call option written on S_2 is 6.9407. This is also the Black-Scholes price with the implied volatility provided in Table E.2. The residual payoff (6.16) is then as follows

$$\text{Residual payoff} = \max(\omega_1 S_1 + \omega_1 S_2 - 100, 0) - \lambda_1 \max(S_1 - 100, 0) - \lambda_2 \max(S_2 - 100, 0), \tag{6.27}$$

where λ_i is the number of units of contract C_i. The task is then to find the optimal static hedging position λ_1 and λ_2 via solving the optimization problem (6.17) or (6.22), for maximal pricing or minimal pricing. After solving (6.17), we find that the maximal price is 5.6042, and the optimal static hedge position is to buy 0.46 units of C_1 and 0.49 units of C_2. After solving (6.22), the minimal price is 4.9296, and the optimal static hedge position is to buy 0.28 units of C_1 and 0.37 units of C_2. These spread values are summarized in Table 6.4. Hence, by statically hedging the basket option, the spread of has been greatly reduced from 1.6456 (6.3248 − 4.6792) to 0.6746 (5.6042 − 4.9296). This difference gives a more realistic bid-ask spread when pricing for buying or selling an option contract. Using this approach, we have reduced the uncertainty due to model calibration errors significantly.

The payoff of the original call option and the residual of the hedged option (for the minimal price) are shown in Figure 6.6. The magnitude of the payoff has been significantly reduced with static hedging. Note that although the payoff of the basket option is convex, the addition of two European call options for static hedging make the residual payoff non-convex.

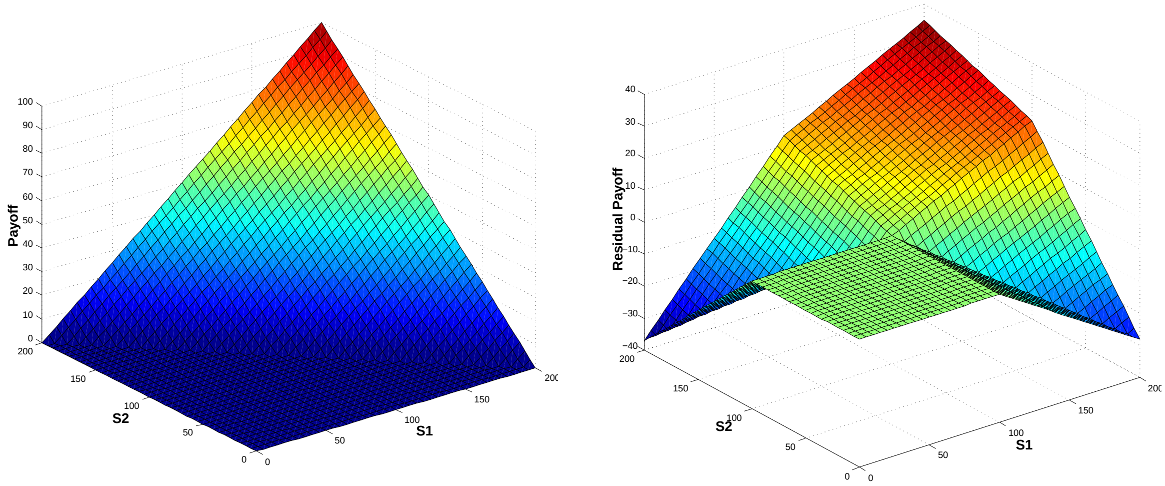


Figure 6.6: The terminal payoff of the unhedged basket option and the residual payoff with optimal static hedge (for the minimal price).

6.5 Summary

The main results of this chapter are as follows.

- By reformulating the local volatility calibration problem into a Bayesian framework, we construct a reasonable pair of the lower and upper bounds for the local volatility values.
- If the maximal and minimal values of a contract computed with the bounds for the local volatility functions give rise to a large spread, this indicates that using the best fit local volatility surfaces results in a large model uncertainty. By incorporating a static hedge, this model uncertainty can be reduced substantially.

Chapter 7

Conclusions and Future Work

We are particularly interested in solving multidimensional HJB equations which are derived from optimal stochastic control problems in the financial market. We developed a fully implicit, unconditionally monotone finite difference numerical scheme. Consequently, there are no time step restrictions due to stability considerations, and the fully implicit method has essentially the same complexity per step as the explicit method. The discretized algebraic equations are solved using policy iteration. In some cases our discretization method results in a local objective function which is a discontinuous function of the control. Therefore, we took particular care in guaranteeing the convergence of the algorithm when applying policy iteration.

The main difficulty in designing a discretization scheme is development of a monotonicity preserving approximation of the cross derivative term in the PDE. We primarily used a wide stencil based on a local coordination rotation. We also derived a hybrid numerical scheme which combines use of a fixed stencil to minimize the use of the wide stencil if the grid spacings could satisfy certain conditions.

The analysis rigorously showed that our numerical scheme is ℓ_∞ stable, consistent in the viscosity sense, and monotone. Therefore, our numerical scheme guarantees convergence to the viscosity solution.

Firstly, our numerical schemes have been applied to pricing two factor options under an uncertain volatility model. For this application, a hybrid scheme which uses the fixed stencil as much as possible was developed to take advantage of its accuracy and computational efficiency. Secondly, using our numerical method, we studied the problem of optimal asset allocation where the risky asset follows stochastic volatility. By applying a semi-Lagrangian timestepping method to handle the drift term and an improved linear interpolation, we could construct accurate efficient frontiers. An interesting fact is found that if the mean reversion time $\frac{1}{\kappa}$ is small compared to the investment horizon T, then a constant volatility GBM approximation to the stochastic volatility process gives a good approximation to the optimal strategy. Finally, we have utilized our numerical scheme to carry out an optimal static hedge, in the case of an uncertain local volatility model.

7.1 Future work

We suggest the following directions for future work.

- In the thesis, to keep the numerical scheme monotone, linear interpolation is the most accurate interpolation we can use. [63, 30] shed light on using higher order interpolation, and it would be interesting to use a higher order interpolation to improve the accuracy of a numerical scheme.
- Optimal trade execution is an important application of the HJB equations in the financial market. It would be interesting to study the simultaneous liquidation of multiple assets that have correlated price movements using the numerical scheme developed in the thesis.
- We could apply our numerical methods to three factor problems. For example, an optimal asset allocation problem which has three assets, i.e., risky stock index, a short term bond and a long term bond. This optimal allocation strategy could be investigated by solving a corresponding three factor HJB equation.

Appendix A

Discrete equation coefficients in the fixed stencil

The coefficients in the linear operator (2.17) are given in the following. We use three point operators for the first and second derivatives. Central Differencing in S_1 and S_2 direction:

$$\begin{aligned}
\alpha_{i,j}^{S_1,central} &= \left[\frac{(\sigma_1(S_1)_i)^2}{((S_1)_i-(S_1)_{i-1})((S_1)_{i+1}-(S_1)_{i-1})} - \frac{(r-q_1)(S_1)_i}{(S_1)_{i+1}-(S_1)_{i-1}}\right],\\
\beta_{i,j}^{S_1,central} &= \left[\frac{(\sigma_1(S_1)_i)^2}{((S_1)_{i+1}-(S_1)_i)((S_1)_{i+1}-(S_1)_{i-1})} + \frac{(r-q_1)(S_1)_i}{(S_1)_{i+1}-(S_1)_{i-1}}\right],\\
\alpha_{i,j}^{S_2,central} &= \left[\frac{(\sigma_2(S_2)_j)^2}{((S_2)_j-(S_2)_{j-1})((S_2)_{j+1}-(S_2)_{j-1})} - \frac{(r-q_2)(S_2)_j}{(S_2)_{j+1}-(S_2)_{j-1}}\right],\\
\beta_{i,j}^{S_2,central} &= \left[\frac{(\sigma_2(S_2)_j)^2}{((S_2)_{j+1}-(S_2)_j)((S_2)_{j+1}-(S_2)_{j-1})} + \frac{(r-q_2)(S_2)_j}{(S_2)_{j+1}-(S_2)_{j-1}}\right].
\end{aligned} \tag{A.1}$$

Forward/Backward Differencing in S_1 and S_2 direction (upstream):

$$\begin{aligned}
\alpha_{i,j}^{S_1,ups} &= \left[\frac{(\sigma_1(S_1)_i)^2}{((S_1)_i-(S_1)_{i-1})((S_1)_{i+1}-(S_1)_{i-1})} + \max\left(0, -\frac{(r-q_1)(S_1)_i}{(S_1)_i-(S_1)_{i-1}}\right)\right],\\
\beta_{i,j}^{S_1,ups} &= \left[\frac{(\sigma_1(S_1)_i)^2}{((S_1)_{i+1}-(S_1)_i)((S_1)_{i+1}-(S_1)_{i-1})} + \max\left(0, \frac{(r-q_1)(S_1)_i}{(S_1)_{i+1}-(S_1)_i}\right)\right],\\
\alpha_{i,j}^{S_2,ups} &= \left[\frac{(\sigma_2(S_2)_j)^2}{((S_2)_j-(S_2)_{j-1})((S_2)_{j+1}-(S_2)_{j-1})} + \max\left(0, -\frac{(r-q_2)(S_2)_j}{(S_2)_j-(S_2)_{j-1}}\right)\right],\\
\beta_{i,j}^{S_2,ups} &= \left[\frac{(\sigma_2(S_2)_j)^2}{((S_2)_{j+1}-(S_2)_j)((S_2)_{j+1}-(S_2)_{j-1})} + \max\left(0, \frac{(r-q_2)(S_2)_j}{(S_2)_{j+1}-(S_2)_j}\right)\right].
\end{aligned} \tag{A.2}$$

$$\gamma_{i,j} = \begin{cases} \frac{\rho(S_1)_i(S_2)_j\sigma_1\sigma_2}{((S_1)_{i+1}-(S_1)_i)((S_2)_{j+1}-(S_2)_j)+((S_1)_i-(S_1)_{i-1})((S_2)_j-(S_2)_{j-1})}, & \text{if } \rho >= 0,\\ -\frac{\rho(S_1)_i(S_2)_j\sigma_1\sigma_2}{((S_1)_{i+1}-(S_1)_i)((S_2)_j-(S_2)_{j-1})+((S_1)_i-(S_1)_{i-1})((S_2)_{j+1}-(S_2)_j)}, & \text{if } \rho < 0. \end{cases} \tag{A.3}$$

The algorithm to select upstream and central differencing is given as follows

Algorithm A.1 Differencing method in the $S_k,\ k = 1, 2$ direction

```
if $\alpha_{i,j}^{S_k,central} - \gamma_{i,j} \geq 0$ then
    $\alpha_{i,j}^{S_k} \leftarrow \alpha_{i,j}^{S_k,central}$
    $\beta_{i,j}^{S_k} \leftarrow \beta_{i,j}^{S_k,central}$
else
    $\alpha_{i,j}^{S_k} \leftarrow \alpha_{i,j}^{S_k,ups}$
    $\beta_{i,j}^{S_k} \leftarrow \beta_{i,j}^{S_k,ups}$
end if
```

Appendix B

The discretized equation for the case $\big((S_1)_i, (S_2)_j, \tau^{n+1}\big) \in \Omega_{w^*}$

For the case $((S_1)_i, (S_2)_j, \tau^{n+1}) \in \Omega_{w^*}$, using Algorithm 2.1 to avoid using points below the lower boundaries, the discrete linear operator L_w^Q (2.36) needs to be modified to the form $L_{w^*}^Q$.

$$
\begin{aligned}
L_{w^*}^Q \mathcal{U}_{i,j}^{n+1} = {} & \frac{a_{i,j}}{h_{1,left}(h_{1,left}+h_{1,right})} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} - h_{1,left}(\mathbf{R}_{i,j})_1\right) \\
& + \frac{a_{i,j}}{h_{1,right}(h_{1,left}+h_{1,right})} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} + h_{1,right}(\mathbf{R}_{i,j})_1\right) \\
& + \frac{b_{i,j}}{h_{2,left}(h_{2,left}+h_{2,right})} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} - h_{1,left}(\mathbf{R}_{i,j})_2\right) \\
& + \frac{b_{i,j}}{h_{2,right}(h_{2,left}+h_{2,right})} \mathcal{J}_h \mathcal{U}^{n+1}\left(\mathbf{S}_{i,j} + h_{2,right}(\mathbf{R}_{i,j})_2\right) \\
& + 1_{(r-q_1)\geq 0} \frac{(r-q_1)(S_1)_i}{\Delta^+(S_1)_i} \mathcal{U}_{i+1,j}^{n+1} - 1_{(r-q_1)<0} \frac{(r-q_1)(S_1)_i}{\Delta^-(S_1)_i} \mathcal{U}_{i-1,j}^{n+1} \\
& + 1_{(r-q_2)\geq 0} \frac{(r-q_2)(S_2)_j}{\Delta^+(S_2)_j} \mathcal{U}_{i,j+1}^{n+1} - 1_{(r-q_2)<0} \frac{(r-q_2)(S_2)_j}{\Delta^-(S_2)_j} \mathcal{U}_{i,j-1}^{n+1} \\
& - \Bigg(1_{(r-q_1)\geq 0} \frac{(r-q_1)(S_1)_i}{\Delta^+(S_1)_i} - 1_{(r-q_1)<0} \frac{(r-q_1)(S_1)_i}{\Delta^-(S_1)_i} + 1_{(r-q_2)\geq 0} \frac{(r-q_2)(S_2)_j}{\Delta^+(S_2)_j} \\
& + 1_{(r-q_2)<0} \frac{(r-q_2)(S_2)_j}{\Delta^-(S_2)_j} + \frac{a_{i,j}}{h_{1,left}(h_{1,left}+h_{1,right})} + \frac{a_{i,j}}{h_{1,right}(h_{1,left}+h_{1,right})} \\
& + \frac{b_{i,j}}{h_{2,left}(h_{2,left}+h_{2,right})} + \frac{b_{i,j}}{h_{2,right}(h_{2,left}+h_{2,right})} + r \Bigg) \mathcal{U}_{i,j}^{n+1},
\end{aligned}
\tag{B.1}
$$

where $h_{k,left},\ h_{k,right},\ k = 1, 2$ are determined by Algorithm 2.1. Then, using fully implicit timestepping, the HJB equation (2.5a) has the following discretized equation for this case

$$\frac{\mathcal{U}^{n+1} - \mathcal{U}^{n}}{\Delta\tau} = \sup_{Q \in \partial Z_h} \left(L^{Q}_{w^*} \mathcal{U}^{n+1}_{i,j} \right). \tag{B.2}$$

Appendix C

Proof of the local consistency of $L_{w^*}^Q$

Proof. We use the discrete linear operator $L_{w^*}^Q$ (B.1) in the region $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*}$. Ω_{w^*} is the region in Ω_b where the conditions (2.33) are not satisfied and then the wide stencil is used. As defined in Table 2.1, Ω_b is

$$\Omega_b \equiv [h, \sqrt{h}] \times (0, S_{2,\max}] \times (0, T] \cup (0, S_{1,\max}] \times [h, \sqrt{h}] \times (0, T], \tag{C.1}$$

where h (2.13) is a mesh discretization parameter.

We divide this region Ω_b into two parts. The first part Ω_{b_1} is defined as

$$\Omega_{b_1} \equiv [h, \sqrt{h}] \times [h, \sqrt{h}] \times (0, T], \tag{C.2}$$

and $\Omega_{b_2} = \Omega_b / \Omega_{b_1}$.

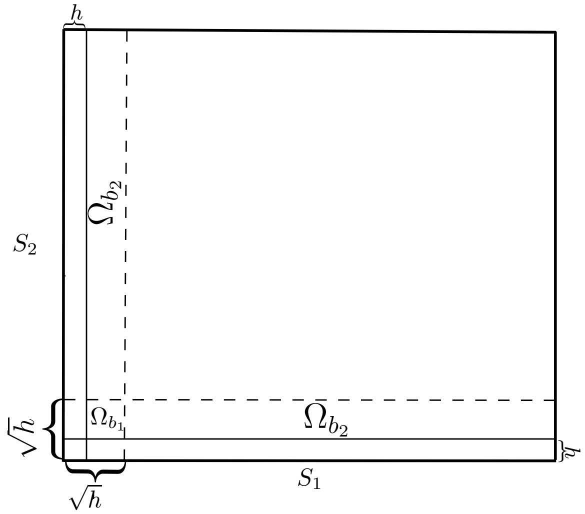


Figure C.1: The region Ω_b.

Algorithm 2.1 guides us as to how to shrink the stencil length to avoid using points below the lower boundaries when approximating the second order terms $\frac{\partial^2 \mathcal{V}}{\partial y_k^2}$, $k = 1, 2$ (2.21). If $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*} \cap \Omega_{b_2}$, we only need to change either the value of $h_{k,left}$ or $h_{k,right}$ from $\sqrt{h}$ to h, but not both. Only if $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*} \cap \Omega_{b_1}$, we may shrink $h_{k,left}$ and $h_{k,right}$ to h simultaneously.

For the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*} \cap \Omega_{b_2}$, without loss of generality, let $h_{k,left} = h$ and $h_{k,right} = \sqrt{h}$. Suppose ϕ is a smooth test function and we use linear interpolation operator $\mathcal{J}_h$, then we have

$$\begin{aligned}
&\frac{\frac{\mathcal{J}_h \phi^{n+1}(\mathbf{S}_{i,j} - h(\mathbf{R}_{i,j})_k) - \phi^{n+1}(\mathbf{S}_{i,j})}{h} + \frac{\mathcal{J}_h \phi^{n+1}\left(\mathbf{S}_{i,j} + \sqrt{h}(\mathbf{R}_{i,j})_k\right) - \phi^{n+1}(\mathbf{S}_{i,j})}{\sqrt{h}}}{\frac{h+\sqrt{h}}{2}} \\
&= \frac{\frac{\phi^{n+1}(\mathbf{y}_{i,j} - h\mathbf{e}_k) - \phi^{n+1}(\mathbf{y}_{i,j}) + O(h^2)}{h} + \frac{\phi^{n+1}\left(\mathbf{y}_{i,j} + \sqrt{h}\mathbf{e}_k\right) - \phi^{n+1}(\mathbf{y}_{i,j}) + O(h^2)}{\sqrt{h}}}{\frac{h+\sqrt{h}}{2}} \\
&= \frac{\frac{\phi^{n+1}(\mathbf{y}_{i,j} - h\mathbf{e}_k) - \phi^{n+1}(\mathbf{y}_{i,j})}{h} + \frac{\phi^{n+1}\left(\mathbf{y}_{i,j} + \sqrt{h}\mathbf{e}_k\right) - \phi^{n+1}(\mathbf{y}_{i,j})}{\sqrt{h}}}{\frac{h+\sqrt{h}}{2}} + O(\sqrt{h}) \\
&= \frac{\partial^2 \phi}{\partial y_k^2} + O(\sqrt{h}) + O(\sqrt{h}), \quad k = 1, 2
\end{aligned} \tag{C.3}$$

which follows from Taylor series expansion and that the error of linear interpolation for a smooth function ϕ is $O(h^2)$. Thus, our discretization to the second order terms at $\mathbf{x}_{i,j}^{n+1}$ is consistent.

For the case $\mathbf{x}_{i,j}^{n+1} \in \Omega_{w^*} \cap \Omega_{b_1}$, when we shrink $h_{k,left}$ and $h_{k,right}$ to h simultaneously, following the same steps in the previous case, we have

$$\frac{\frac{\mathcal{J}_h \phi^{n+1}(\mathbf{S}_{i,j} - h(\mathbf{R}_{i,j})_k) - \phi^{n+1}(\mathbf{S}_{i,j})}{h} + \frac{\mathcal{J}_h \phi^{n+1}(\mathbf{S}_{i,j} + h(\mathbf{R}_{i,j})_k) - \phi^{n+1}(\mathbf{S}_{i,j})}{h}}{\frac{h+h}{2}} = \frac{\partial^2 \phi}{\partial y_k^2} + O(1). \tag{C.4}$$

In this case, the approximation of $\frac{\partial^2 \phi}{\partial y_k^2}$ is locally inconsistent. However, by observing the fact that the value of $a_{i,j}$ and $b_{i,j}$ in the region Ω_{b_1} is $O(h)$ (see equation (2.22)), our discretization at $\mathbf{x}_{i,j}^{n+1}$ is still locally consistent. That is,

$$\begin{aligned}
&a_{i,j} \underbrace{\left(\left.\frac{\partial^2 \phi}{\partial y_1^2}\right|_{\mathbf{y}_{i,j}} + O(1) \right)}_{\text{approximation of } \frac{\partial^2 \phi}{\partial y_1^2}} + \; b_{i,j} \underbrace{\left(\left.\frac{\partial^2 \phi}{\partial y_2^2}\right|_{\mathbf{y}_{i,j}} + O(1) \right)}_{\text{approximation of } \frac{\partial^2 \phi}{\partial y_2^2}} \\
&= \left(a_{i,j} \frac{\partial^2 \phi}{\partial y_1^2} + b_{i,j} \frac{\partial^2 \phi}{\partial y_2^2} \right) + O(h) = ((\mathbf{D}\nabla) \cdot \nabla \phi) \big|_{\mathbf{x}_{i,j}^{n+1}} + O(h).
\end{aligned} \tag{C.5}$$

In $L_{w^*}^{Q}$, we use the standard forward or backward finite differencing, depending on the sign of drift $r - q_k$, $k = 1, 2$ to discretize the first order derivatives in (2.5a). The approximations of the first order derivatives are clearly locally consistent to $O(h)$. Finally, we have, in the worst case,

$$L_{w^*}^Q \phi_{i,j}^{n+1} = \mathcal{L}\phi_{i,j}^{n+1} + O(\sqrt{h}). \tag{C.6}$$

□

Appendix D

The discrete linear operator D_h^p

With vanishing cross-derivative term, the degenerate linear operator $\mathcal{L}^p$ (2.3) can be discretized by a standard finite difference method. The degenerate linear operators $\mathcal{L}^p$ in (5.19), (5.20), and (5.21) are approximated as the discrete form

$$D_h^p \mathcal{U}_{i,j}^n = \alpha_{i,j}^w \mathcal{U}_{i-1,j}^n + \beta_{i,j}^w \mathcal{U}_{i+1,j}^n + \alpha_{i,j}^v \mathcal{U}_{i,j-1}^n + \beta_{i,j}^v \mathcal{U}_{i,j+1}^n - (\alpha_{i,j}^w + \beta_{i,j}^w + \alpha_{i,j}^v + \beta_{i,j}^v)\mathcal{U}_{i,j}, \quad \text{(D.1)}$$

where $\alpha_{i,j}^w, \beta_{i,j}^w, \alpha_{i,j}^v$ and $\beta_{i,j}^v$ are defined as follows

$$\begin{aligned}
\alpha_{i,j}^w &= \frac{(\sqrt{v}pw_i)^2}{(w_i - w_{i-1})(w_{i+1} - w_{i-1})}, \\
\beta_{i,j}^w &= \frac{(\sqrt{v}pw_i)^2}{(w_{i+1} - w_i)(w_{i+1} - w_{i-1})}, \\
\alpha_{i,j}^v &= \left[\frac{(\sigma\sqrt{v_j})^2}{v_j - v_{j-1})(v_{j+1} - v_{j-1})} + \max\left(0, -\frac{\kappa(\theta - v_j)}{v_j - v_{j-1}}\right)\right], \\
\beta_{i,j}^v &= \left[\frac{(\sigma\sqrt{v_j})^2}{(v_{j+1} - v_j)(v_{j+1} - v_{j-1})} + \max\left(0, \frac{\kappa(\theta - v_j)}{v_{j+1} - v_j}\right)\right].
\end{aligned} \quad \text{(D.2)}$$

The coefficients $\alpha_{i,j}^w, \beta_{i,j}^w, \alpha_{i,j}^v$ and $\beta_{i,j}^v$ are all non-negative, and is compatible with a monotone scheme. On the upper boundary $v = v_{\max}$, the coefficients α_{i,N_2}^v and $\beta_{i,N_2}^v = 0$ degenerate to zero, and on the lower boundary $w = 0$, $\alpha_{1,j}^w$ and $\beta_{1,j}^w$ are set to zero. On the lower boundary $v = 0$, $\alpha_{i,1}^w = 0$, $\beta_{i,1}^w = 0$, $\alpha_{i,1}^v = 0$, and $\beta_{i,1}^w = \frac{\kappa\theta}{v_{j+1} - v_j}$, $j = 1$.

Appendix E

Datasets

Maturity	Strike (K), % of Spot									
	85	90	95	100	105	110	115	120	130	140
0.175	0.190	0.168	0.133	0.113	0.102	0.0970	0.120	0.142	0.169	0.200
0.425	0.177	0.155	0.138	0.125	0.109	0.1030	0.100	0.114	0.130	0.150
0.695	0.172	0.157	0.144	0.133	0.118	0.1040	0.100	0.101	0.108	0.124
0.940	0.171	0.159	0.149	0.137	0.127	0.1130	0.106	0.103	0.100	0.110
1.000	0.171	0.159	0.150	0.138	0.128	0.1150	0.107	0.103	0.099	0.108
1.500	0.169	0.160	0.151	0.142	0.133	0.1240	0.119	0.113	0.107	0.102
2.000	0.169	0.161	0.153	0.145	0.137	0.1300	0.126	0.119	0.115	0.111

Table E.1: For the S&P 500 dataset on October, 1995: market implied volatilities. The spot price of the underlying at time 0 is $S_0 = 590$, the interest rate is $r = 0.059$ and dividend rate $d = 0.026$. This dataset is often used as the benchmark dataset in the local volatility calibrations papers [22, 2].

Maturity	Strike (K), % of Spot												
	70	80	85	90	95	100	105	110	115	120	130	140	150
0.08	0.19	0.18	0.22	0.22	0.22	0.21	0.19	0.18	0.18	0.18	0.16	0.16	0.15
0.25	0.19	0.18	0.21	0.21	0.22	0.21	0.19	0.18	0.18	0.17	0.16	0.16	0.15
0.50	0.18	0.18	0.20	0.21	0.21	0.20	0.19	0.18	0.18	0.17	0.16	0.16	0.15
0.75	0.18	0.18	0.19	0.20	0.20	0.20	0.19	0.18	0.18	0.17	0.16	0.16	0.15
1.0	0.17	0.17	0.18	0.19	0.20	0.19	0.19	0.18	0.18	0.17	0.16	0.16	0.16
1.50	0.16	0.17	0.17	0.18	0.19	0.19	0.18	0.18	0.18	0.17	0.16	0.16	0.16
2.0	0.15	0.16	0.17	0.17	0.18	0.18	0.18	0.17	0.17	0.17	0.16	0.16	0.16

Table E.2: Market implied volatilities for FTSE-100 index used in [49, 42]. The spot price of the underlying at time 0 is $S_0 = 5000$, the interest rate is $r = 0.05$ and dividend rate $d = 0.03$. Strikes (K) are in percentage of initial spot and maturities (T) are measured in years.

References

[1] Y. Aıt-Sahalia and R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452, 2007.

[2] L. Andersen and R. Brotherton-Ratcliffe. The equity option volatility smile: an implicit finite-difference approach. *Journal of Computational Finance*, 1(2):5–37, 1998.

[3] M. Avellaneda and R. Buff. Combinatorial implications of nonlinear uncertain volatility models: the case of barrier options. *Applied Mathematical Finance*, 6(1):1–18, 1999.

[4] M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.

[5] M. Avellaneda and A. Paras. Managing the volatility risk of portfolios of derivative securities: the Lagrangian uncertain volatility model. *Applied Mathematical Finance*, 3(1):21 – 52, 1996.

[6] G. Barles, Ch. Daher, and M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Mathematical Models and Methods in Applied Sciences*, 5(1):125–143, 1995.

[7] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, 4(3):271–283, 1991.

[8] S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *Review of Financial Studies*, 23(8):2970–3016, 2010.

[9] N. Bauerle and S. Grether. Complete markets do not allow free cash flow streams. *Mathematical Methods of Operations Research*, 81:137–145, 2015.

[10] T.R. Bielecki, H. Jin, S.R. Pliska, and X.Y. Zhou. Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Mathematical Finance*, 15(2):213–244, 2005.

[11] T. Bjork and A. Murgoci. A general theory of Markovian time inconsistent stochastic control problems. *Available at SSRN 1694759*, 2010.

[12] O. Bokanowski, S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM Journal on Numerical Analysis*, 47(4):3001–3026, 2009.

[13] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. A fast algorithm for the two dimensional HJB equation of stochastic control. *M2AN Math. Model. Numer. Anal.*, 38(4):723–735, 2004.

[14] 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.

[15] D.T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, 1979.

[16] R. Buff. *Uncertain Volatility Models: Theory and Application*, volume 1. Springer-Verlag, Berlin, 2002.

[17] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO, Modelisation Mathematique Et Analyse Numerique*, 29(1):97–122, 1995.

[18] Z.L. Chen and P.A. Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30(1):339–368, 2007.

[19] 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.

[20] T.F. Coleman, C. He, and Y. Li. Calibrating volatility function bounds for an uncertain volatility model. *Journal of Computational Finance*, 13(4):63, 2010.

[21] T.F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[22] T.F. Coleman, Y. Li, and A. Verma. Reconstructing the unknown local volatility function. *Journal of Computational Finance*, 2:77–102, 1999.

[23] R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547, 2006.

[24] J.C. Cox, J. Ingersoll, E. Jonathan, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, 53(2):385–407, 1985.

[25] S. Crépey. Calibration of the local volatility in a generalized Black–Scholes model using Tikhonov regularization. *SIAM Journal on Mathematical Analysis*, 34(5):1183–1206, 2003.

[26] X. Cui, D. Li, S. Wang, and S. Zhu. Better than dynamic mean-variance: Time inconsistency and free cash flow stream. *Mathematical Finance*, 22(2):346–378, 2012.

[27] D.M. Dang and P.A. Forsyth. Better than pre-commitment mean-variance portfolio allocation strategies: a semi-self-financing Hamilton-Jacobi-Bellman equation approach. *Working paper, Cheriton School of Computer Science, University of Waterloo*, 2014.

[28] D.M. Dang and P.A. Forsyth. Continuous time mean-variance optimal portfolio allocation under jump diffusion: An numerical impulse control approach. *Numerical Methods for Partial Differential Equations*, 30(2):664–698, 2014.

[29] D.M. Dang, P.A. Forsyth, and Y. Li. Convergence of the embedded mean-variance optimal points with discrete sampling. To appear in Numerische Mathematik, 2015.

[30] K. Debrabant and E. Jakobsen. Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. *Mathematics of Computation*, 82(283):1433–1462, 2013.

[31] Y. d'Halluin, P.A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97(2):321–352, 2004.

[32] N.G. Dokuchaev and A.V. Savkin. The pricing of options in a financial market model with transaction costs and uncertain volatility. *Journal of Multinational Financial Management*, 8(2):353–364, 1998.

[33] 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.

[34] J.J. Douglas and T.F. Russell. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM Journal on Numerical Analysis*, 19(5):871–885, 1982.

[35] B. Dupire. *Pricing and hedging with smiles.* Mathematics of derivative securities. Dempster and Pliska eds., Cambridge University Press, 1997.

[36] E. Ekström and J. Tysk. The Black–Scholes equation in stochastic volatility models. *Journal of Mathematical Analysis and Applications*, 368(2):498–507, 2010.

[37] W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54(1):173–182, 1951.

[38] R. Fletcher. *Practical Methods of Optimization.* John Wiley & Sons, 2013.

[39] P.A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *Journal of Computational Finance*, 11(2):1–44, 2007.

[40] P.A. Forsyth and K.R. Vetzal. Implicit solution of uncertain volatility/transaction cost option pricing models with discretely observed barriers. *Applied Numerical Mathematics*, 36(4):427–445, 2001.

[41] A. Gupta. *A Bayesian approach to financial model calibration, uncertainty measures and optimal hedging*. PhD thesis, University of Oxford, 2010.

[42] A. Gupta and C. Reisinger. Robust calibration of financial models using Bayesian estimators. *Journal of Computational Finance*, 17(4):3–36, 2014.

[43] J. Guyon and P. Henry-Labordere. The uncertain volatility model: a Monte Carlo approach. *Journal of Computational Finance*, 14(3):37–71, 2011.

[44] S.B. Hamida and R. Cont. Recovering volatility from option prices by evolutionary optimization. *Journal of Computational Finance*, 8(4):43–76, 2005.

[45] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[46] R.A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, Boston, 1960.

[47] 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(1):320–351, 2012.

[48] Y. Huang, P.A. Forsyth, and G. Labahn. Combined fixed point and policy iteration for Hamilton–Jacobi–Bellman equations in finance. *SIAM Journal on Numerical Analysis*, 50(4):1861–1882, 2012.

[49] N. Jackson, E. Suli, and S. Howison. Computation of deterministic volatility surfaces. *Journal of Computational Finance*, 2(2):5–32, 1999.

[50] S. Janson and J. Tysk. Preservation of convexity of solutions to parabolic equations. *Journal of Differential Equations*, 206(1):182–226, 2004.

[51] C. Kahl and P. Jäckel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[52] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, 2001.

[53] D. Li and W.-L. Ng. Optimal dynamic portfolio selection: multiperiod mean-variance formulation. *Mathematical Finance*, 10(3):387–406, 2000.

[54] R. Lord, R. Koekkoek, and D.V. Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[55] T.J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance*, 2(2):117–133, 1995.

[56] J.L. Menaldi. Some estimates for finite difference approximations. *SIAM Journal on Control and Optimization*, 27(3):579–607, 1989.

[57] P. Nguyen and R. Portait. Dynamic asset allocation with mean variance preferences and a solvency constraint. *Journal of Economic Dynamics and Control*, 26(1):11–32, 2002.

[58] B.K. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions.* Springer-Verlag, Berlin, 2005.

[59] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2(1–2):506–549, 2005.

[60] O. Pironneau. On the transport-diffusion algorithm and its applications to the Navier-Stokes equations. *Numerische Mathematik*, 38(3):309–332, 1982.

[61] 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*, 23(2):241–267, 2003.

[62] D.M. Pooley, P.A. Forsyth, and K.R. Vetzal. Two factor option pricing with uncertain volatility. In *Computational Science and Its Applications ICCSA 2003*, volume 2669 of *Lecture Notes in Computer Science*, pages 158–167. Springer-Verlag, Berlin, 2003.

[63] C. Reisinger and P.A. Forsyth. Piecewise constant policy approximations to Hamilton-Jacobi-Bellman equations. *arXiv preprint arXiv:1503.05864*, 2015.

[64] Y. Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2004.

[65] S. Salmi, J. Toivanen, and L. von Sydow. An IMEX-scheme for pricing options under stochastic volatility models with jumps. *SIAM Journal on Scientific Computing*, 36(5):B817–B834, 2014.

[66] A.T. Smith. American options under uncertain volatility. *Applied Mathematical Finance*, 9(2):123–141, 2002.

[67] R.M. Stulz. Options on the minimum or the maximum of two risky assets: analysis and applications. *Journal of Financial Economics*, 10(2):161–185, 1982.

[68] S.T. Tse, P.A. Forsyth, J.S. Kennedy, and H. Windcliff. Comparison between the mean-variance optimal and the mean-quadratic-variation optimal trading strategies. *Applied Mathematical Finance*, 20(5):415–449, 2013.

[69] S.T. Tse, P.A. Forsyth, and Y. Li. Preservation of scalarization optimal points in the embedding technique for continuous time mean variance optimization. *SIAM Journal on Control and Optimization*, 52(3):1527–1546, 2014.

[70] R.S. Varga. *Matrix Iterative Analysis*. Springer-Verlag, Berlin, 2009.

[71] E. Vigna. On efficiency of mean–variance based portfolio selection in defined contribution pension schemes. *Quantitative Finance*, 14(2):237–258, 2014.

[72] 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.

[73] J. Wang and P.A. Forsyth. Numerical solution of the Hamilton Jacobi Bellman formulation for continuous time mean variance asset allocation. *Journal of Economic Dynamics and Control*, 34(2):207–230, 2010.

[74] J. Wang and P.A. Forsyth. Comparison of mean variance like strategies for optimal asset allocation problems. *International Journal of Theoretical and Applied Finance*, 15(2), 2012.

[75] Y. Zhao and W.T. Ziemba. Mean-variance versus expected utility in dynamic investment analysis. *Working paper, University of British Columbia*, 2000.

[76] X.Y. Zhou and D. Li. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics & Optimization*, 42(1):19–33, 2000.